



Skrypt 290

Mirosława Szewczyk, Magdalena Ciesielska

Podstawy statystyczne badań marketingowych

Skrypt dla studentów



ISSN 1427-9932
ISBN 978-83-62736-15-7
Opole 2010

Politechnika Opolska

Komitet Redakcyjny

Andrzej Knapik, Jan Kubik,
Tadeusz Łagoda – przewodniczący
Mariusz Migąła, Iwona Mulicka,
Jan Sadecki, Małgorzata Wróblewska

Komitet Redakcyjny Wydawnictw Politechniki Opolskiej, ul. S. Mikołajczyka 5

Recenzent: dr hab. Krystyna Hanusik, prof. Uniwersytetu Opolskiego

Redaktor: Małgorzata Wróblewska

Skład: Oficyna Wydawnicza Politechniki Opolskiej.

Nakład: 115 + 25 + 10 egz. Ark. wyd. 13,8. Ark. druk. 13,8.

Druk i oprawa: Sekcja Poligrafii Politechniki Opolskiej.

Spis treści

WSTĘP	7
1. WPROWADZENIE DO BADAŃ MARKETINGOWYCH	9
1.1. Podejmowanie decyzji	9
1.2. Definicja badań marketingowych	10
1.3. Obszary badań marketingowych	11
1.4. System informacji marketingowej a badania marketingowe	13
2. CELE I ETAPY BADANIA MARKETINGOWEGO	17
2.1. Cele badań marketingowych	17
2.2. Etapy badania marketingowego	19
2.3. Przykłady z omówieniem	26
2.4. Zadania do samodzielnego opracowania	30
3. POMIAR W BADANIACH MARKETINGOWYCH	33
3.1. Skala pomiaru	33
3.2. Trafność i rzetelność pomiaru	39
3.3. Wybrane skale pomiaru postaw	40
3.4. Przykłady z omówieniem	46
3.5. Zadania do samodzielnego opracowania	54
4. KWESTIONARIUSZ – PODSTAWOWY INSTRUMENT POMIARU	57
4.1. Wprowadzenie	57
4.2. Podstawowe zasady konstrukcji kwestionariusza	57
4.3. Formułowanie pytań i odpowiedzi	59
4.4. Przykłady z omówieniem	62
4.5. Zadania do samodzielnego opracowania	64
5. GROMADZENIE DANYCH WTÓRNYCH	67
5.1. Typy danych wtórnych	67
5.2. Zalety i wady danych wtórnych	73
5.3. Przykłady z omówieniem	75
5.4. Zadania do samodzielnego opracowania	77
6. GROMADZENIE DANYCH PIERWOTNYCH	79
6.1. Wywiady	79
6.2. Badania ankietowe	81
6.3. Eksperyment	83
6.4. Obserwacja	86
6.5. Przykłady z omówieniem	88
6.6. Zadania do samodzielnego opracowania	90

7.	WYBRANE KRYTERIA SEGMENTACJI RYNKU	93
7.1.	Pojęcie segmentacji rynku	93
7.2.	Cechy demograficzne, socjologiczne, psychologiczne i stylu życia	94
7.3.	Zadania do samodzielnego opracowania	97
8.	DOBÓR PRÓBY DO BADAŃ	99
8.1.	Zdefiniowanie populacji	99
8.2.	Dobór losowy i dobór nielosowy	99
8.3.	Ustalenie wykazu jednostek badanej populacji	103
8.4.	Określenie liczebności próby	105
8.5.	Minimalna liczebność próby losowej	106
8.6.	Błędy w badaniach marketingowych	109
8.7.	Przykłady z omówieniem	116
8.8.	Zadania do samodzielnego opracowania	117
9.	WSTĘPNE OPRACOWANIE DANYCH	119
9.1.	Podstawowe pojęcia	119
9.2.	Kontrola pomiaru	119
9.3.	Redakcja danych	119
9.4.	Grupowanie	120
9.5.	Kodowanie	120
9.6.	Tabulacja	121
9.7.	Przykłady z omówieniem	121
10.	PREZENTACJA DANYCH	129
10.1.	Wprowadzenie	129
10.2.	Prezentacja danych na wykresie	130
10.3.	Podstawowe charakterystyki liczbowe	131
10.4.	Przykłady z omówieniem	135
10.5.	Zadania do samodzielnego opracowania	142
11.	ANALIZA I INTERPRETACJA DANYCH – ANALIZA KORELACJI I REGRESJI	145
11.1.	Analiza korelacji	145
11.2.	Współczynniki korelacji – dwie zmienne nominalne	146
11.3.	Współczynniki korelacji – dwie zmienne mierzone na skali porządkowej	146
11.4.	Współczynniki korelacji – dwie zmienne mierzone na skali przedziałowej	149
11.5.	Analiza regresji	151
11.6.	Przykłady z omówieniem	153
11.7.	Zadania do samodzielnego opracowania	164
12.	1. Dekompozycja szeregu czasowego	167
12.2.	Miary dynamiki	168

12.3. Przykłady z omówieniem	169
12.4. Zadania do samodzielnego opracowania	171
13. WNIOSKOWANIE STATYSTYCZNE	175
13.1. Wnioskowanie statystyczne - podstawowe pojęcia	175
13.2. Systematyka testów	178
13.3. Testy parametryczne dla wnioskowania o własnościach populacji jednowymiarowej - testy dla średniej	180
13.4. Test niezależności χ^2	183
13.5. Przykłady z omówieniem	184
13.6. Zadania do samodzielnego opracowania	188
14. SEGMENTACJA Z WYKORZYSTANIEM METODY TAKSONOMII WROCLAWSKIEJ	191
14.1. Klasyfikacja obiektów	191
14.2. Normalizacja	192
14.3. Metoda taksonomii wrocławskiej	193
14.4. Przykład z omówieniem	194
TABLICE STATYSTYCZNE	198
SŁOWNIK NAJWAŻNIEJSZYCH POJĘĆ	202
LITERATURA	212
SPIS TABEL	215
SPIS WYKRESÓW	217
SPIS RYSUNKÓW	218

WSTĘP

Badania marketingowe są wieloetapowym procesem, który polega na systematycznym gromadzeniu, analizie i prezentacji danych, w celu doskonalenia procesów decyzyjnych. W skrypcie uwagę poświęcono zarówno teoretycznym, jak i praktycznym zagadnieniom związanym z realizacją badań marketingowych, m.in.: przekładaniu problemu decyzyjnego na problem badawczy, doborowi odpowiednich metod, technik i narzędzi badawczych, doborowi próby, analizie i interpretacji danych (z zastosowaniem metod statystycznych), raportowaniu i prezentacji wyników badań. Firmy najczęściej decydują się na przeprowadzenie badań, gdy wprowadzają na rynek nowy produkt lub usługę, szukają nowych nabywców, gdy spada sprzedaż, albo gdy zmienia się dotychczasowy profil działalności firmy. Jednak nie każda firma dysponuje wystarczającym budżetem, aby zlecić takie badania profesjonalnej firmie badawczej. Warto zatem posiadać podstawową wiedzę w zakresie prowadzenia badań marketingowych. A jeśli własne siły okażą się zbyt skromne, to może okazać się przydatne umiejętności współpracy z agencją badawczą oraz nadzorowania jej działań.

Opracowanie obejmuje swym zakresem materiał pomocny w samodzielnym przeprowadzaniu badań marketingowych lub/i współpracy przy ich przeprowadzaniu z profesjonalną agencją badawczą. Praca zawiera omówienie podstawowych pojęć, które student powinien opanować w trakcie studiowania przedmiotu *badania marketingowe*, a także przykłady z omówieniem. Możliwość utrwalenia nabytej wiedzy dają zadania do samodzielnego rozwiązania, umieszczone na końcach rozdziałów.

Skrypt jest przeznaczony głównie dla studentów odbywających zajęcia z *badania marketingowych* na kierunku *zarządzanie*. Student, przystępujący do analizy materiału zawartego w skrypcie, powinien posiadać podstawowy zasób wiadomości z marketingu oraz statystyki opisowej.

Skrypt, dzięki przyjętej konstrukcji, na którą składają się omówienia pojęć, przykłady z rozwiązaniami oraz zadania do samodzielnego rozwiązania, powinien ułatwić studentom nabycie umiejętności zastosowania podstawowych metod statystycznych w badaniach marketingowych. Zawarte w pracy zadania i pytania kontrolne powinny być ze szczególną starannością przygotowane przez studenta i w razie potrzeby skonsultowane z osobą prowadzącą zajęcia. Jednocześnie student powinien sięgnąć do zalecanej literatury przedmiotu, aby uzupełnić wiedzę zawartą w opracowanych materiałach dydaktycznych. Pomoże to we właściwym przygotowaniu do zaliczenia przedmiotu.

Oddając do rąk studentów niniejsze opracowanie, autorki zwracają się z prośbą o przekazywanie wszelkich uwag i spostrzeżeń odnośnie zawartego materiału. Życzymy studentom, by dzięki tej pomocy dydaktycznej dobrze opanowali materiał i pomyślnie zdali egzamin.

Mirosława Szewczyk

Magdalena Ciesielska

1. WPROWADZENIE DO BADAŃ MARKETINGOWYCH

1.1. Podejmowanie decyzji

Dla przeciętnego człowieka badania marketingowe kojarzą się przede wszystkim z badaniami ankietowymi przeprowadzanymi na grupie konsumentów. Jednak w rzeczywistości badania marketingowe obejmują dużo szerszy zakres. Mogą one być przeprowadzane nie tylko na rynku konsumenckim, ale także na rynku przemysłowym oraz rynku business-to-business (B2B). W celu lepszego zrozumienia istoty badań marketingowych, należy rozważyć ich rolę w podejmowaniu decyzji.

Wszystkie podmioty, które działają na rynku, muszą szybko reagować na zmiany w ich otoczeniu zewnętrznym. Przedsiębiorstwa działają w specyficznym dla siebie otoczeniu prawnym, ekonomicznym, ekologicznym, technologicznym i konkurencyjnym. Szybka reakcja przedsiębiorstwa na zmiany pozwala menedżerom wypracować przewagę konkurencyjną, a w związku z tym osiągnąć wyższe zyski. Spóźnione działania marketingowe stanowią dla przedsiębiorstwa potencjalne źródło zagrożenia. Właśnie w celu zminimalizowania podejmowanego ryzyka oraz lepszego dostosowania działań do zajmowanej na rynku pozycji, menedżerowie podejmują szereg decyzji marketingowych. Decyzje te przede wszystkim dotyczą dwóch elementów otoczenia zewnętrznego organizacji: konkurentów i konsumentów.

Decyzje są podejmowane przez decydentów – menedżerów działających w danej organizacji. Poszczególne decyzje marketingowe są zależne od wielu czynników, ale przede wszystkim od rodzaju branży, w jakiej działa organizacja. Zadania marketingu w przedsiębiorstwie skupiają się na zaspokajaniu potrzeb konsumentów. Wykonując swoje funkcje, menedżerowie mają do dyspozycji cztery instrumenty marketingu-mix: produkt (product), cenę (price), promocję (promotion), dystrybucję (place), zwane także 4P. Podstawowym celem menedżerów jest ułożenie takiej kompozycji instrumentów marketingu-mix, która zapewni maksymalną sprzedaż oraz przyniesie przedsiębiorstwu największe zyski. Osiągnięcie takiej efektywności nie jest zadaniem prostym, gdyż podejmowanie decyzji wiąże się z ryzykiem. Im większa niepewność, co do oczekiwanych zdarzeń, tym większe ryzyko i niższy zwrot z poniesionych nakładów. Największym ryzykiem obarczone są decyzje, których skutków nie jesteśmy w stanie przewidzieć. Gdyby menedżerowie mogli określić wszystkie czynniki, które mogłyby wpłynąć na wielkość sprzedaży produktu bądź usługi, wówczas działałoby w warunkach pewności. Niestety takie warunki występują rzadko bądź wcale. Istnieje szereg czynników, których menedżer nie jest w stanie przewidzieć. Minimalizując ryzyko menedżerowie muszą określić nie tylko zachowanie indywidualnych klientów, ale także posiadać informację na

temat zmian zachodzących w otoczeniu bliższym i dalszym przedsiębiorstwa. Menedżer posiada wpływ na zastosowanie ceny, promocji, dystrybucji czy produktu na danym rynku docelowym. Natomiast poza jego kontrolą są zmiany otoczenia polityczno-prawnego, kulturowego, społecznego, ekonomicznego czy technologicznego. Wszystkie te czynniki mają znaczenie przy podejmowaniu decyzji marketingowych. Tak więc badania marketingowe mają pomóc menedżerom w podejmowaniu decyzji marketingowych, dostarczając im odpowiednich danych i informacji.

1.2. Definicja badań marketingowych

W literaturze istnieje wiele definicji badań marketingowych. Wśród dostępnych definicji można wyróżnić podział na dwie podstawowe grupy: skoncentrowanych na **funkcji** badań marketingowych oraz skoncentrowanych wokół **procesu** badawczego.

Do pierwszej grupy (definicji skoncentrowanych na **funkcji** badań marketingowych) zaliczamy m.in. definicję P. Green, D. Tull, D. Albaum – „Badania marketingowe to systematyczne i obiektywne poszukiwanie i analiza informacji właściwych do określenia i rozwiązywania problemów marketingowych”¹.

Jak zwraca uwagę E. Duliniec – „Badania marketingowe mogą spełniać następujące funkcje:

- deskryptywną, która wyraża się w charakteryzowaniu podmiotów i przedmiotów rynku oraz zdarzeń i procesów zachodzących na rynku,
- eksplikacyjną, która wyraża się w identyfikacji zależności mających miejsce na rynku,
- prognostyczną, która wyraża się w symulacji przyszłych zdarzeń i procesów rynkowych,
- innowacyjną, która wyraża się w tworzeniu nowych rozwiązań w zakresie produktów i działań rynkowych,
- kontrolną, która wyraża się w ocenie skuteczności różnorodnych działań marketingowych”².

Do drugiej grupy (definicji skoncentrowanych wokół **procesu** badawczego) zaliczamy następujące definicje:

- P. Kotler – „Badania marketingowe to systematyczne planowanie, zbieranie, analiza i przekazywanie danych i informacji istotnych dla sytuacji marketingowej, w jakiej znajduje się przedsiębiorstwo”³.

¹ P. Green, D. Tull, D. Albaum, Research for Marketing Decisions, Prentice Hall, New York 1985.

² M. Rószkiewicz, Metody ilościowe w badaniach marketingowych. Wydawnictwo Naukowe PWN, Warszawa 2002, s. 13

- S. Kaczmarczyk – „Badania marketingowe wiążą organizację z jej otoczeniem rynkowym. Organizacja ta projektuje badania, gromadzi, analizuje i interpretuje dane, aby pomóc kierownictwu w zrozumieniu otoczenia, identyfikowaniu problemów i sprzyjających sytuacji oraz w rozwoju i przeprowadzaniu działań na określonych rynkach”⁴.
- Z. Kędzior, K. Karcz – „Badania marketingowe to włączony do systemu informacji marketingowej obiektywny proces gromadzenia, przetwarzania oraz prezentacji informacji na potrzeby podejmowania decyzji marketingowych”⁵.

Oficjalną definicję badań marketingowych stanowi definicja uznana w 1987 r. przez Amerykańskie Stowarzyszenie Marketingu (*American Marketing Association*) – „Badania marketingowe są to działania wiążące daną organizację z jej rynkowym otoczeniem za pomocą informacji. Informacje te są wykorzystywane do zidentyfikowania i określenia problemów oraz szans marketingowych do wywołania, udoskonalenia i oceny działań marketingowych, do monitorowania efektów tych działań oraz do coraz lepszego poznania procesów marketingowych”⁶.

1.3. Obszary badań marketingowych

Badania marketingowe wykonywane są we wszystkich obszarach problemowych. Badania mogą dotyczyć uzyskania pomiaru pragnień i potrzeb konsumentów, wpływu zastosowanego narzędzia marketingu-mix na wyniki sprzedaży, studiów dotyczących reklamy i promocji, studiów politycznych, itd. Badania marketingowe mają za zadanie odpowiedzieć na pytania stawiane przez decydentów w konkretnych obszarach ich niewiedzy np.: Kto jest głównym konkurentem bezpośrednim? Jakie są główne cechy konkurencji? Jaki jest wizerunek przedsiębiorstwa wśród klientów? Czym się kieruje konsument w procesie zakupu produktu? Jaki wpływ na sprzedaż ma opakowanie produktu? Jaki wpływ ma zmiana ceny produktu na popyt? Jakie segmenty rynku możemy wyróżnić? Jaka jest efektywność zastosowanej promocji?

Badania marketingowe można klasyfikować z podziałem na obszary badawcze:

³ P. Kotler, *Marketing. Analiza, planowanie, wdrażanie i kontrola*. Gebethner i S-ka, Warszawa 1994, s. 139.

⁴ S. Kaczmarczyk. *Badania marketingowe. Metody i techniki*. Polskie Wydawnictwo Ekonomiczne Warszawa 2003, s. 15.

⁵ Z. Kędzior, K. Karcz. *Badania marketingowe w praktyce*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2007, s. 21.

⁶ American Marketing Association. <http://www.marketingpower.com/AboutAMA/Pages/DefinitionofMarketing.aspx>.

- Badanie produktu – celem badania jest określenie obecnej lub przyszłej pozycji produktu na rynku
 - identyfikacja potrzeby, którą produkt ma zaspokoić
 - dostarczanie informacji o nowych produktach
 - opracowanie koncepcji nowego produktu
 - tworzenie marki handlowej
 - projektowanie opakowania
 - analiza produktów oferowanych przez konkurencję
 - rynek próbny
 - analiza produktów oferowanych przez rynek
- Cena – celem badania jest określenie wrażliwości cenowej konsumentów oraz zróżnicowania cenowego
 - analiza struktury cen
 - analiza trendów cen
 - prognozowanie skutków zmiany cen
 - analiza kosztów
 - prognozowanie zysku
 - analiza polityki cenowej konkurencji
 - analiza popytu
- Promocja – badanie to obejmuje badanie reklamy, zarówno w formie pre-testu (badania potencjalnego odbioru i skutków reklamy) oraz post-testu (badania oceniające efektywność reklamy). Badania promocji nie koncentrują się jedynie na reklamie, dotyczą one także kosztów promocji czy charakterystykę kanałów przekazu.
 - badania mediów
 - badania motywacji
 - analiza kosztów promocji
 - analiza reklamy konkurencji
 - ocena skuteczności promocji
 - badanie wizerunku
 - określanie form sprzedaży
 - rozwój koncepcji reklamy
- Dystrybucja – badanie ma na celu charakterystykę kanału dystrybucji lub całej sieci dystrybucji
 - analiza lokalizacji magazynów
 - analiza zasięgu dystrybucji
 - analiza efektywności kanału dystrybucji
 - analiza eksportu
- Rynek – badanie koncentruje się na charakterystyce rynku, tzn. jego wielkości, dynamiki, podaży, popytu oraz udziału w nim, a także podziału rynku na poszczególne kategorie produktów i konsumentów.

- charakterystyka rynku
- segmentacja rynku
- trendy panujące na rynku
- analiza udziału w rynku
- główni konkurenci na rynku
- struktura dystrybucji
- udział marek w rynku
- Konsumenci – badanie potencjalnych grup nabywców produktu oraz charakterystyka procesu zakupu
 - preferencje konsumentów co do marki
 - wizerunek marki
 - identyfikacja podstawowych potrzeb
 - zachowania nabywcze
 - świadomość marki
 - motywy zakupu
- Otoczenie zewnętrzne przedsiębiorstwa – badanie pozycji przedsiębiorstwa na rynku, także jej pozycji wobec pozycji konkurentów. Charakterystyka otoczenia zewnętrznego przedsiębiorstwa, a także identyfikacja kluczowych czynników sukcesu występujących wewnątrz firmy.
 - analiza trendów gospodarczych
 - analiza stanu prawnego działalności przedsiębiorstwa
 - analiza uregulowań prawnych dotyczących ochrony środowiska
 - analiza wartości społecznych.

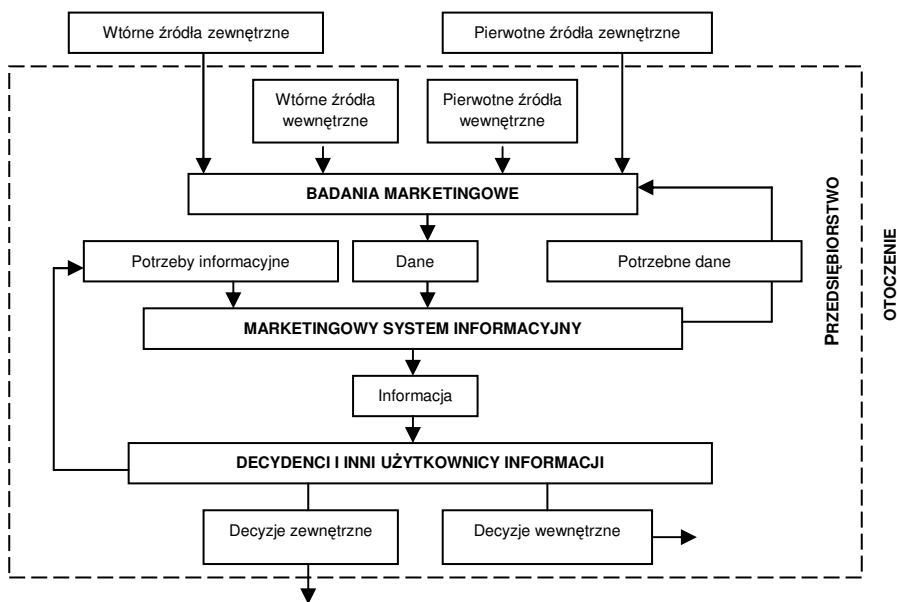
1.4. System informacji marketingowej a badania marketingowe

Każda firma powinna zorganizować przepływ informacji tak, aby umożliwić zarządzającym podejmowanie trafnych decyzji marketingowych. **System informacji marketingowej (SIM)** składa się z ludzi, sprzętu oraz technik gromadzenia, porządkowania, analizy i oceny, a następnie przekazywania na czas potrzebnej i dokładnej informacji do osób podejmujących decyzje dotyczące marketingu⁷. System informacji marketingowej powinien zapewnić zarządzającym możliwie pełne, komplementarne i aktualne

⁷ P. Kotler, Marketing, Gebethner&Ska, Warszawa 1994, s. 115, Koncepcja systemu informacji marketingowej pojawia się w wielu ujęciach. Por. Kaczmarczyk S., Badania marketingowe. Metody i techniki, PWE, Warszawa 1995; J. Jerschina, Całościowy system informacji marketingowej w przedsiębiorstwach i instytucjach finansowych [<http://www.statsoft.pl/czytelnia>]

informacje o rynku, umożliwiające podejmowanie trafnych decyzji marketingowych⁸.

Podstawową składową systemu (por. rys. 1) jest suma informacji pochodzących z badań marketingowych (zarówno ze źródeł pierwotnych, jak i wtórnych). Odpowiednio zebrane, opracowane i zinterpretowane dane o uwarunkowaniach rynkowych przedsiębiorstwa trafiają do poszczególnych kierowników (w zależności od struktury organizacyjnej przedsiębiorstwa i potrzeb informacyjnych). Efektywny system informacji marketingowej powinien dostarczać optymalnej ilości informacji, zgodnej z potrzebami. Użytkownicy SIM oczekują często pomocy w podejmowaniu decyzji oraz kontroli rezultatów decyzji. Zarówno użytkownicy, jak i dostawcy informacji, winni być świadomi podejmowanych decyzji oraz powinni być obeznani z potrzebami każdego etapu procesu decyzyjnego.



Rys. 1. Schemat marketingowego systemu informacyjnego

Źródło: Kaczmarczyk S., *Badania marketingowe. Metody i techniki*, PWE, Warszawa 1995, s. 29.

⁸ Więcej: J. Jerschina, Całościowy system informacji marketingowej w przedsiębiorstwach i instytucjach finansowych, <http://www.statsoft.pl/czytelnia> (11.03.2010).

Oto warunki, które warunkują powstanie efektywnego systemu informacji marketingowej:

- właściwy, odpowiadający strategii przedsiębiorstwa sposób zhierarchizowania i zintegrowania informacji wytwarzanych przez system
- plan badań marketingowych, zapewniający systematyczne zbieranie informacji marketingowych
- metody badań pozwalające na zdobycie porównywalnych (zwłaszcza w czasie) i komplementarnych informacji
- metody przetwarzania danych (przekładalność informacji o rynku na rekomendacje marketingowe)
- metody komunikowania zarządzającym uzyskanych informacji i rekomendacji⁹.

Dokumentacja przedsiębiorstwa dostarcza danych o tym, co już miało miejsce (danych o rezultatach działalności przedsiębiorstwa), natomiast system wywiadu marketingowego dostarcza kierownictwu danych o tym, co może się wydarzyć. System wywiadu marketingowego to zbiór procedur i źródeł wykorzystywanych przez kierownictwo, aby pozyskać codzienne informacje na temat rozwoju sytuacji w otoczeniu marketingowym¹⁰. Kierownicy obserwują otoczenie za pomocą obserwacji nieukierunkowanych, obserwacji ukierunkowanych, poszukiwań nieformalnych oraz poszukiwań formalnych. Poszukiwania formalne mają charakter działania przemyślanego, poprzedzonego zazwyczaj zatwierdzonym planem (lub/i metodologią) w celu uzyskania określonych informacji. System informacji marketingowej jest zawsze elementem przedsiębiorstwa. Natomiast badania marketingowe mogą być prowadzone przez zlecenie badań na zewnątrz – przez profesjonalne firmy. Badania marketingowe są traktowane jako element systemu informacji marketingowej (por. rys. 1). Badania marketingowe wiążą organizację z jej otoczeniem rynkowym. Organizacja

⁹ Więcej: J. Jerschina, Całościowy system informacji marketingowej w przedsiębiorstwach i instytucjach finansowych, <http://www.statsoft.pl/czytelnia> (11.03.2010).

¹⁰ P. Kotler, Marketing, Gebethner&Ska, Warszawa 1994, s. 119.

projektuje badania, gromadzi, analizuje i interpretuje dane, aby wspomóc kierownictwo w zrozumieniu otoczenia, identyfikowaniu szans i zagrożeń rozwoju¹¹.

¹¹ Dane wykorzystywane w działaniach marketingowych są gromadzone i przechowywane w bazach danych. Hurtownie danych coraz częściej stają się fundamentem i podstawowym środkiem realizacji nowoczesnego zarządzania kontaktami z klientami. Jednym z najciekawszych przypadków wdrożenia aktywnej hurtowni danych w celu polepszenia efektywności prowadzonych działań marketingowych jest przykład sieci sklepów Wal-Mart w USA. Wal-Mart, jako pierwsza sieć handlowa, już w latach 80. zaczął rejestrować elektronicznie zdarzenia w całej sieci i agregować je w hurtowni danych. Wnioski, jakie dzięki temu można było wyciągnąć z historii sprzedaży, umożliwiły globalną reorganizację logistyki i obniżkę kosztów, a co z tym związane również obniżkę cen oferowanych produktów. Hurtownia danych przechowuje stany magazynowe wszystkich spośród ponad 100 000 oferowanych produktów, aktualizując je na bieżąco po dokonaniu przez klienta zakupu każdego egzemplarza, a dokładniej z chwilą zeskanowania jego kodu kreskowego podczas sprzedaży w kasie. Użytkownikami hurtowni są dostawcy towarów, którzy w dowolnym momencie mają wgląd w aktualne stany magazynowe wszystkich dostarczanych przez siebie towarów. Mogą więc niemal natychmiast reagować na najmniejsze nawet zmiany popytu, zmniejszając lub zwiększając podaż poprzez dodatkowe dostawy do sklepów lub ograniczenie produkcji. Nie mają oczywiście wglądu w wyniki sprzedaży towarów dostarczanych przez konkurencję. Sieć Wal-Mart płaci dostawcom jedynie za towary, które zostały sprzedane, a nie tak jak w przypadku sieci klasycznych hipermarketów z góry za określoną dostawę. Dodatkowo znacząco zmniejszają się koszty magazynowania, ponieważ sami dostawcy w celu maksymalizacji własnych zysków troszczą się o proporcjonalność i regularność dostaw. Dzięki temu możliwe jest utrzymywanie niskich marż, a tym samym konkurencyjnych cen oferowanych towarów. Kolejnym, niebagatelnym zyskiem, który przyniosło wdrożenie hurtowni jest możliwość bardzo szybkiej reakcji na wszelkie zmiany popytu na rynku. Najlepszym egzaminem systemu hurtowni danych sieci Wal-Mart okazał się okres bezpośrednio po tragicznym zamachu terrorystycznym 11.09.2001 r. na World Trade Center. Na rynku w okamgnieniu wzrósł popyt na żywność o długim terminie przydatności do spożycia, maski gazowe, a także różne symbole narodowe (przede wszystkim amerykańskie flagi). Sieć Wal-Mart dzięki posiadaniu zintegrowanej aktywnej hurtowni danych prawie natychmiast zidentyfikowała nowy trend, a następnie dzięki szybkiej reakcji dostawców i użytkowników hurtowni jednocześnie oraz sprawnej logistyce zarobiła miliony dolarów na sprzedaży towarów, których w innych sieciach zabrakło. Historię opisanego wyżej wdrożenia można znaleźć w książce Paula Westermana: „Data Warehousing. Using the Wal-Mart Model”. Efekt prac w zakresie składowania i analizy dużych wolumenów danych zaowocował rozwojem technologii hurtowni danych w zastosowaniach komercyjnych. Pojęcie Business Intelligence definiuje się zwykle jako wspomaganie decyzji biznesowych przy wykorzystaniu narzędzi informatycznych, które umożliwiają w czasie rzeczywistym na analizę danych, które są krytyczne z punktu widzenia działania firmy. W przypadku organizacji handlowych, wynik już wieloletnich doświadczeń począwszy od podejścia Wal-Mart, przyczynił się do ustalenia całego zestawu analiz biznesowych. Por. R. Jesionek, IT w branży sprzedaży detalicznej, CIO - Magazyn Dyrektorów IT, 1-2/2007, <http://cio.cxo.pl/artykuly/54278/IT.w.branzy.sprzedazy.detalicznej.html>, <http://aneksy.pwn.pl/zarzadzanie/pdf/Waters17-Wal-Mart.pdf>.

2. CELE I ETAPY BADANIA MARKETINGOWEGO

2.1. Cele badań marketingowych

Krystyna Mazurek-Łopacińska wskazuje trzy typy celów badań marketingowych: pomoc w zrozumieniu, pomoc w decydowaniu i pomoc w kontroli. Wymienione cele, jak twierdzi autorka, mają charakter komplementarny i odnoszą się zarówno do marketingu strategicznego, jak i marketingu operacyjnego (por. rys. 2).



Rys. 2. Cele badań marketingowych

Źródło: *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 15.

Należy pamiętać, że badania marketingowe są działaniami, które mogą generować duże koszty. Badacz powinien sporządzić swego rodzaju bilans uzasadniający ekonomiczną zasadność podejmowanego badania. Mazurek-Łopacińska wymienia czynniki, które należy wziąć pod uwagę, rozważając zasadność podjęcia badań marketingowych: ilość czasu do dyspozycji przed podjęciem decyzji marketingowej, dostępność potrzebnych informacji, naturę decyzji, wartość dodatkowej informacji w stosunku do kosztów ich uzyskania¹².

¹² *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 17.

Badania marketingowe mają dostarczyć decydującym odpowiednich danych, koniecznych do rozwiązania problemu. Jednocześnie muszą zaspokoić potrzeby informacyjne decydentów w możliwie jak najkrótszym czasie. Jest to możliwe tylko wtedy, gdy problemy i cele badania zostaną jasno i precyzyjnie zdefiniowane, a badanie właściwie zaprojektowane.

W celu prawidłowej identyfikacji problemu menedżera konieczna jest analiza otoczenia zewnętrznego przedsiębiorstwa. Rolą badacza jest poprawna identyfikacja problemu menedżera oraz przedstawienie odpowiednich danych, które pomogą ten problem rozwiązać. Otoczenie zewnętrzne przedsiębiorstwa stwarza dla menedżerów wiele szans i zagrożeń. Owe szanse i zagrożenia są wynikiem zmian w otoczeniu demograficznym, ekonomicznym, konkurencyjnym, technologicznym, politycznym, prawnym. Zmiany wpływają na decyzje marketingowe menedżerów, a także kształtują nową rzeczywistość rynkową. Istotnym elementem dostosowania się do rynku jest szybkość reakcji na zmiany. Im szybsza reakcja na zmiany w otoczeniu, tym szybsze dostosowanie się do zmian w otoczeniu, co sprawia, że organizacja jest konkurencyjna, a może być także innowacyjna na rynku. Przejawem tych działań może być np. wprowadzenie na rynek nowego produktu lub usługi, czy też zastosowanie innowacyjnej technologii w produkcji¹³.

Zmiany dotyczą nie tylko otoczenia zewnętrznego, ale także otoczenia wewnętrznego przedsiębiorstwa. Poważne zmiany udziału w rynku, spadek sprzedaży, utrata zaufanych i stałych klientów, zmiany w strukturze dystrybucji, skłaniają przedsiębiorców do poszukiwania przyczyn takiego zachowania rynku. Wówczas badanie marketingowe będzie skoncentrowane na poszukiwaniu odpowiedzi w otoczeniu wewnętrznym organizacji.

Gilbert A. Churchill rozróżnia trzy podstawowe źródła szans i trudności w rozpatrywanych badaniach marketingowych: nieprzewidziane zmiany, zmiany planowane, szczęśliwy traf w formie nowych idei¹⁴. Nieplanowane zmiany nie zależą bezpośrednio od działalności przedsiębiorstwa. Zmiany planowane są zbieżne z celami strategicznymi przedsiębiorstwa. Firmy planując zwiększenie udziału w rynku, rozwinięcie sieci dystrybucji czy też zmianę wizerunku firmy, mają na celu wzmocnienie pozycji przedsiębiorstwa na rynku i stworzenie długotrwałej przewagi konkurencyjnej. Czysty traf definiowany jest przez Churchilla jako przypadkowe idee, na które przedsiębiorstwo zostało naprowadzone przez list lub spostrzeżenie konsumenta.

¹³ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 21.

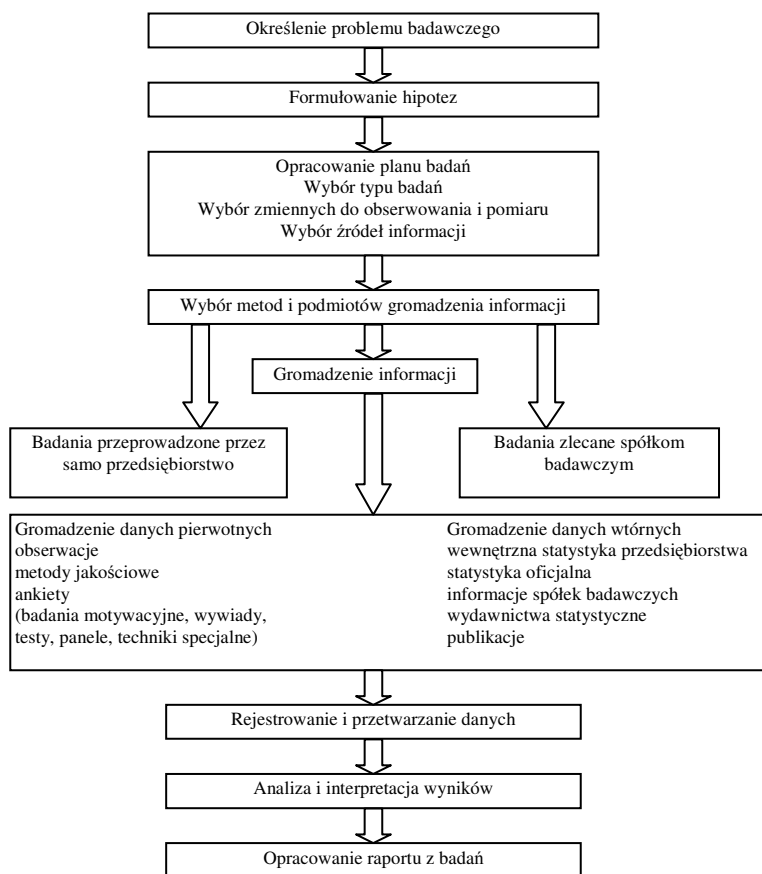
¹⁴ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 105.

2.2. Etapy badania marketingowego

Według Mazurek-Łopacińskiej „podstawowe decyzje, podejmowane na etapie opracowania planu badań, obejmują: wybór typu badań, wybór zmiennych do obserwowania i pomiaru, wybór źródeł informacji, wybór metod gromadzenia i rejestrowania informacji, wybór instrumentów pomiarowych oraz metod analizy informacji”¹⁵ (por. rys. 3).

Sformułowanie problemu badawczego to pierwszy z etapów badania marketingowego i jeden z najważniejszych. Na tym etapie zostaje podjęta decyzja o prowadzeniu, bądź nieprowadzeniu badania marketingowego. Głównym zadaniem badacza na tym etapie jest poprawne zrozumienie i odnalezienie prawdziwego problemu decyzyjnego. Problem decyzyjny zostaje sformułowany na podstawie szeregu rozmów i spotkań z decydentem. W firmie decydentem może być zarówno zarząd jak i menedżer danej specjalizacji. Następnie problem decyzyjny jest przekształcany w problem badawczy. Problem badawczy jest uszczegółowieniem problemu decyzyjnego – jego opisem w formie pozwalającej na analizę danych. Prawidłowe określenie problemu badawczego jest podstawą do opracowania w dalszej części badania hipotez. Pozwala także na lepszą komunikację pomiędzy menedżerem a badaczem. Kluczowe w badaniach marketingowych jest wzajemne zrozumienie problemów, które mogą wystąpić podczas prowadzonego badania. W tym etapie dokonywana jest także analiza kosztów i korzyści badania. Jeżeli koszty badania przewyższają jego ewentualne korzyści, wówczas badanie nie zostaje podjęte. W przeciwnym wypadku przystępuje się do kolejnego etapu.

¹⁵ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 26-33.



Rys. 3. Proces badania marketingowego według Mazurek-Łopacińskiej

Źródło: *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 24.

Początkowe działania mają na celu prawidłowe zdefiniowanie problemu decyzyjnego. W kolejnym etapie badania należy przekształcić problem decyzyjny w problem badawczy. Churchill definiuje problem badawczy, jako: „przeformułowanie problemu decyzyjnego, w kategorii badawcze”¹⁶. Problem decyzyjny dotyczy konkretnych działań strategicznych, a problem badawczy obejmuje postać szeregu pytań konkretyzujących zbadanie problemu decyzyjnego. Innymi słowy problem badawczy

¹⁶ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 114.

określa, jakiego rodzaju dane i informacje są niezbędne do rozwiązania problemu decyzyjnego. Na tym etapie zakładamy, że zarówno prowadzący badanie, jak i menedżer, mają dość precyzyjnie określony problem decyzyjny. W przeciwnym przypadku sformułowany problem badawczy rozwiąże źle sformułowany problem. W tym celu badacz powinien osiągnąć zrozumienie: decydenta i otoczenia, alternatywnych kierunków działania, celów decydenta, skutków alternatywnych działań¹⁷.

Formułowanie hipotez to etap, w którym następuje przekształcenie problemu badawczego w hipotezy. Mazurek-Łopacińska definiuje hipotezę, jako przypuszczenie, orzeczenie niezupełnie pewne o występowaniu danego zjawiska bądź o jego wielkości, częstotliwości, stosunku do innych zjawisk lub o zależności między zjawiskami¹⁸. Hipoteza jest więc pewnym twierdzeniem, które w ciągu badania zostanie potwierdzone bądź obalone. Hipotezy odgrywają dość istotną rolę w badaniu: pozwalają jasno wytyczyć kierunek badań oraz wskazują na zależności pomiędzy zmiennymi w badaniu.

Jeżeli zapadła decyzja, aby powierzyć przeprowadzenie badania wyspecjalizowanej firmie¹⁹, kolejnym krokiem będzie zebranie ofert od różnych agencji badania rynku²⁰. W celu umożliwienia obiektywnego

¹⁷ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 118.

¹⁸ *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 26.

¹⁹ Powody, dla których badanie należy przeprowadzić samodzielnie: stopień skomplikowania poszukiwanych informacji nie jest zbyt duży, spodziewamy się, iż we własnym zakresie szybciej zgromadzimy interesujące nas dane, badanie wymaga zachowania szczególnej tajemnicy, dane, które są nam potrzebne, są na tyle poufne, że nie można powierzyć ich gromadzenia firmie zewnętrznej (np. dane o transakcjach bankowych), mamy znaczne ograniczenia budżetowe, nie stać nas na wynajęcie profesjonalnej firmy. Powody, dla których badanie należy zlecić firmie zewnętrznej to: stopień skomplikowania poszukiwanych danych jest zbyt duży, aby przeprowadzić badanie samodzielnie; potrzebujemy np. zgromadzić wiele specjalistycznych ekspertyz; zleceniodawca badania nie ma wystarczających zasobów, żeby takie badanie przeprowadzić lub nie jest w stanie zapewnić odpowiedniej jakości informacji, badanie wymaga specjalistycznego podejścia badawczego, potrzebne jest zewnętrzne, obiektywne spojrzenie na problem, koszty są niższe niż w przypadku samodzielnego wykonania, firma badawcza może szybciej dostarczyć wyniki, zlecenie firmie zewnętrznej pozwoli na znacznie szybsze uzyskanie wyników, wyniki otrzymane od profesjonalnej firmy badawczej są wiarygodne i rzetelne. Por. <http://www.tns-global.pl/przewodnik> (11.03.2010).

²⁰ Czasami trudno jest jednak porównywać oferty, ponieważ zwykle istnieje więcej niż jedno rozwiązanie dla danego pytania/problemu badawczego. Poniższa lista może okazać się pomocna w podejmowaniu decyzji: Czy agencja lub badacze

porównania ofert i cen, pozycje w ofertach powinny być mniej więcej zbieżne. Jeżeli któraś z pozycji nie jest analogiczna do pozostałych, porównywanie cen będzie trudniejsze. W takim przypadku należy ocenić, która oferta najlepiej odpowiada na pytanie badawcze i zawiera najatrakcyjniejszą cenę²¹. Jeśli decydujemy się na zlecenie badania firmie specjalistycznej, to następnym krokiem powinno być przygotowanie profesjonalnego zapytania ofertowego, rozesłanie do kilku/kilkunastu firm, a następnie ocena otrzymanych ofert. Uwaga praktyczna: duże i znane firmy zazwyczaj żądają za przeprowadzenie badania nawet 2-3-krotnie więcej niż działające na rynku od kilku lat mniejsze firmy.

W końcu lat osiemdziesiątych zaczęły w Polsce powstawać agencje, organizacje, instytuty, firmy, centra oraz fundacje związane z badaniami marketingowymi, społecznymi oraz badaniami opinii publicznej. Dwie polskie organizacje zrzeszające firmy badawcze to OFBOR – Organizacja Firm Badania Opinii i Rynku (adres <http://www.ofbor.pl>) i PTBRiO – Polskie Towarzystwo Badaczy Rynku i Opinii (adres <http://www.ptbrio.pl>). W obu organizacjach zrzeszonych jest ponad 30 ośrodków badawczych, kilka ośrodków należy do obu organizacji jednocześnie. Niektóre z firm specjalizują się w danej dziedzinie, oferując na przykład jedynie badania w terenie, inne natomiast są wyspecjalizowane w pewnym typie prac, jak np. badania przemysłu chemicznego. Ośrodki badawcze zrzeszone w obu polskich organizacjach przestrzegają standardów badań wyznaczanych przez ESOMAR – The World Association of Research Professionals (<http://www.esomar.org>) oraz przez WAPOR - World Association for Public Opinion Research (<http://www.unl.edu/WAPOR>). Na stronach internetowych poszczególnych ośrodków badawczych są umieszczane podstawowe interesujące nas informacje o zakresie badań, rodzaju badań oraz o metodach i technikach badawczych. Do grona członków Organizacji Firm Badania Opinii i Rynku (OFBOR) zalicza się wiodące agencje badawcze w Polsce, co świadczy o roli i znaczeniu organizacji w branży badawczej.

Natura stojącego przed badaczem problemu warunkuje wybór określonej metodologii. Obrana metodologia decyduje zaś o zastosowaniu konkretnych technik gromadzenia danych, sposobów ich analizy i interpretacji. Dokonanie właściwego wyboru metodologii ma kluczowe

posiadają wiedzę, doświadczenie oraz zaplecze potrzebne do tego, żeby prowadzić badania? Czy badacz zrozumiał problem badawczy, czy wykazuje się znajomością problemu? Jak zostało zaprojektowane badanie? Czy informacje wynikające z badania przyczynią się do rozwiązania problemu? Czy badanie ma dobrą strukturę, czy realizacja badania została dobrze zaplanowana? Ile czasu potrzeba na całościową realizację badania i przygotowanie wyników? Czy koszty są proporcjonalne do wagi problemu? Czy oferta badawcza jest wyczerpująca?

²¹ Por. <http://www.tns-global.pl/przewodnik> (11.03.2010).

znaczenie dla pomyślnej realizacji każdego przedsięwzięcia badawczego. Stefan Nowak określa metodę jako określony, powtarzalny i wyuczony sposób postępowania, świadomie skierowanego na realizację pewnego celu²². Żadna z metod zbierania danych nie jest doskonała, każda ma swoje zalety i wady. Różne problemy badawcze wymagają różnych metod. David Silverman pisze tak: „Nie istnieją ani dobre, ani złe metody. Są jedynie metody właściwe dla twojego projektu badawczego (...)”²³.

Badacze nie są zgodni, co do podziału metod zbierania danych. Silverman wymienia cztery główne metody stosowane przez badaczy jakościowych: obserwacja, analiza tekstów i dokumentów, wywiady, nagrywanie i transkrybowanie. Wśród pięciu głównych metod ilościowych wymienia natomiast: sondaż społeczny, eksperyment, statystyki urzędowe, obserwację ustrukturyzowaną oraz analizę treści²⁴. Chava Frankfort-Nachmias i David Nachmias jako metody zbierania danych podają: obserwacje, wywiady, badania sondażowe (w tym: wywiady bezpośrednie, wywiady telefoniczne, badania kwestionariuszowe), wtórną analizę danych i badania jakościowe (w tym: badania terenowe, obserwacja uczestnicząca)²⁵. Philip Kotler wyróżnia cztery podstawowe metody zbierania danych: obserwacja, wywiad zogniskowany, ankieta, eksperyment. Natomiast jako narzędzia badawcze wymienia kwestionariusz i urządzenia mechaniczne²⁶. Małgorzata Rószkiewicz dzieli metody pierwotnych badań marketingowych na nieingerujące (w tym: obserwacje, wywiady, sondaże) oraz ingerujące (eksperyment)²⁷. Każda z metod ma swoją specyfikę, każda metoda ma zalety i wady. Dzięki łączeniu różnych metod w jednym badaniu można, przynajmniej częściowo, pokonać ograniczenia wynikające z zastosowania tylko jednej metody. Mazurek-Łopacińska²⁸ określa trzy podstawowe kryteria, jakimi mogą posługiwać się badacze wybierając określoną metodę

²² S. Nowak, *Metodologia badań społecznych*, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 19.

²³ D. Silverman, „Prowadzenie badań jakościowych”, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 153.

²⁴ Por. D. Silverman, „Prowadzenie badań jakościowych”, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 39, 54.

²⁵ Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 221.

²⁶ P. Kotler, *Marketing*, Gebethner & Ska, Warszawa 1994, s. 123.

²⁷ M. Rószkiewicz, *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 36.

²⁸ *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 30-31.

gromadzenia danych: typ badań, wizja rynku, dostępne zasoby finansowe i ludzkie.

Proces analizy i interpretacji danych można podzielić na następujące etapy: edycja, kodowanie, tabulacja, analiza i interpretacja. Edycja – jest to sprawdzenie kompletności i poprawności danych²⁹. Podczas edycji uzupełnia się brakujące dane, poprawia ich czytelność oraz sprawdza się je pod względem merytorycznym. Celem edycji jest przygotowanie danych do kodowania. Kodowanie – to proces polegający na przydzieleniu danym odpowiedniego symbolu. Symbole przydzielane są w celu lepszej identyfikacji oraz uporządkowania danych. Tabulacja jest to proces polegający na odzwierciedleniu danych w tabelach, które pomagają w charakterystyce i analizie danych. Analiza polega na sporządzeniu charakterystyki i opisu badanego zjawiska, populacji. Analiza prowadzona jest metodami statystycznymi. Analiza ma za zadanie takie przygotowanie danych, aby mogły one zostać zinterpretowane. Interpretacja danych to proces polegający na sformułowaniu wniosków na podstawie zanalizowanych danych. Dane w tym etapie mogą zostać poddane indukcji, analogii bądź dedukcji. Indukcja polega na formułowaniu wniosków „od szczegółu do ogółu”. Analogia jest to proces polegający na formułowaniu wniosków na podstawie wniosków wyciągniętych na podstawie innego podobnego (analogicznego) zjawiska. Dedukcja ma za zadanie taką interpretację danych, aby na jej podstawie możliwe było stworzenie odpowiedniego modelu lub schematu danego zjawiska.

Typowe badanie marketingowe kończy się opracowaniem raportu i jego prezentacją. Raport z badań często jest jedynym dokumentem, z którym zapoznaje się menedżer i na podstawie którego podejmuje decyzję marketingową. Dlatego poprawnie sporządzony raport powinien odzwierciedlać cały przebieg badania marketingowego oraz zawierać poprawne wnioski końcowe.

Wielokrotnie zdarza się, iż decydenci nie znają szczegółów, dość skomplikowanego procesu projektowania i realizacji badań, zaś decyzję podejmują na podstawie skróconego przedstawienia wniosków pochodzących z raportu. Tworząc prezentację dla kierownictwa (raport) należy mieć na uwadze, iż kierownictwo firmy w znacznej mierze podejmie decyzje o kolejnym uruchomieniu procesu badawczego dotyczącego określonego problemu decyzyjnego w sytuacji, gdy dotychczas przeprowadzone badania dostarczały przydatnych informacji, które ułatwiają podjęcie decyzji.

Raport z przeprowadzonych badań – opisuje przebieg rezultaty i wnioski z przeprowadzonych badań. Raport musi być przedstawiony w

²⁹ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 33.

sposób czytelny i zrozumiały. Raport pełni trzy zasadnicze funkcje: gromadzi w sposób uporządkowany informacje analizy i rezultaty przeprowadzonych badań, stanowi podstawę do oceny jakości badań, wpływa na podejmowanie decyzji przez zawarte w nim wnioski i konkluzje

Ogólna struktura raportu:

- spis treści
- wstęp (cel, tło badania)
- metodologia badania
- analiza i interpretacja
- podsumowanie i konkluzje
- załączniki.

Niektóre raporty koncentrują się na próbie wstępnego, ogólnego zrozumienia i przedstawienia jakiegoś zjawiska. Czasami cel raportu jest opisowy, czyli badacz tylko opisuje zestaw różnych cech danej populacji (np. wiek, płeć, wykształcenie). Najczęściej jednak w raporcie łączy się cel opisowy ze wskazywaniem zależności przyczynowo-skutkowych. Skutkuje to głębszą i bardziej szczegółową analizą danych.

Metodologia badania powinna zawierać: określenie populacji, czyli grupy ludzi, na temat której będą formułowane wnioski, określenie metody doboru próby – opis w jaki sposób wybrano grupę badawczą, określenie wielkości próby – ile osób zostało wytypowanych do wzięcia udziału w badaniu, a ile osób wzięło faktyczny udział (przyczyny ewentualnej rozbieżności), określenie terminu badania – opis kiedy badanie zostało przeprowadzone, odsetek braków odpowiedzi – czy w ankiecie znalazły się takie pytania, na które nie udzielono odpowiedzi lub udzielono ich niewiele; która z badanych grup nie chciała lub nie mogła, czy też nie była w stanie na nie odpowiedzieć; opis metod analizy danych

Analiza i interpretacja – najważniejsza i zarazem najtrudniejsza do napisania część raportu. Następuje tu prezentacja danych, ich przekształcenie oraz własna interpretacja i konstruowanie wniosków. Wszystkie te elementy powinny stanowić logiczną całość, każda hipoteza musi być podparta konkretnymi wynikami czy opisami.

Raport powinien się kończyć stwierdzeniem odkryć oraz sugestiami co do kierunku przyszłych badań. W podsumowaniu mogą się znaleźć również „mocne” i „słabe” strony przedmiotu badań, np. w przypadku oceny działalności badanego przedsiębiorstwa – które obszary działalności spełniają wymogi, a które są słabsze i wymagają poprawy³⁰.

Przy tworzeniu raportu z badań marketingowych należy pamiętać o następujących zasadach:

³⁰ Podstawy prowadzenia badań marketingowych. Etapy procesu badań marketingowych – teoria procesu badawczego, CRON, Warszawa 2006, s. 95-99.

- Należy przedstawić maksimum informacji (bez nadmiernej, zbędnej szczegółowości), pamiętając o celu raportu. W raporcie nie ma potrzeby uwzględniania wszystkich aspektów; wybieramy tylko najważniejsze.
- Prezentacje danych ilościowych w tabelach zwiększają przejrzystość raportu.
- Należy unikać dublowania tych samych danych przedstawionych w różnych formach (tabela, wykres i opis).
- Tabele, wykresy lub rysunki należy włączyć do tekstu w sąsiedztwie omawiającego je fragmentu.
- Wnioski należy formułować jasno, precyzyjnie (dobrze widziane jest wyróżnienie wniosków z zalewu informacji).
- Należy używać jasnego, zwięzłego, przejrzystego stylu. Unikać skomplikowanych konstrukcji językowych, zbyt rozbudowanych zdań oraz żargonu.
- Zbyt krótkie sprawozdanie z badań, może być równie mało przydatne, jak zbyt szczegółowe i rozwlekłe.
- Język raportu powinien być zrozumiały dla czytającego, np. specjalista w danej dziedzinie nie potrzebuje szczegółowego wyjaśnienia pojęć, ale laik – potrzebuje wyjaśnień.
- Używamy bezosobowej formy czasowników.

Dobry raport powinien być: zgodny z celami badania, na temat, ciekawy, uporządkowany, obiektywny, wyczerpujący, udokumentowany, kompletny, przejrzysty, zrozumiały, estetyczny, dać konkretne wskazówki, na czas³¹.

Duże firmy przeprowadzają dość często badania ogólnopolskie, z których część wyników ukazuje się bezpłatnie w postaci raportu cząstkowego. Cały raport można nabyć odpłatnie, zwykle w formacie MS Word lub PowerPoint. Całkowity koszt raportu to na ogół kilka tysięcy złotych. Zainteresowane osoby proszone są o kontakt e-mailowy lub telefoniczny w celu przekazania szczegółowych informacji. Na ogół raport cząstkowy pojawia się (bezpłatnie) na stronach internetowych firmy.

2.3. Przykłady z omówieniem

Przykład 1. Rozważmy oświadczenie menedżera ds. marketingu firmy X o spadku sprzedaży odkurzaczy na polskim rynku. Na tej podstawie formułujemy problem badawczy:

- Czy spadek sprzedaży odkurzaczy wynika z cyklu życia produktu?
- Czy spadek sprzedaży odkurzaczy wynika ze zmiany struktury popytu?

³¹ <http://www.inquiry.com.pl> (11.03.2010)

- Czy spadek sprzedaży odkurzaczy wynika ze złej kompozycji marketingu-mix na rynku polskim?
- Czy spadek sprzedaży odkurzaczy wynika ze zmian w otoczeniu konkurencyjnym przedsiębiorstwa?
- Czy spadek sprzedaży odkurzaczy wynika ze zmian w rynku docelowym?

Hipotezy, sformułowane na podstawie pytań badawczych, wyglądają następująco:

- Rynek odkurzaczy jest nasycony.
- Na rynku pojawił się nowy model odkurzacza.
- Firma powinna zmienić formę promocji odkurzaczy.
- Na rynku pojawił się silny globalny konkurent.
- Konsumenci zmienili swoje postawy względem odkurzaczy firmy X.

Przykład 2. Załóżmy, że badanie prowadzone jest dla przedsiębiorstwa „Iskra” produkującego świeczki. Ogólnym celem badania jest identyfikacja cech produktu, dla których jest on kupowany oraz określenie rzeczywistych reakcji respondentów na zapach świec i ich opakowanie. Chcemy się dowiedzieć, czy produkt podoba się klientom firmy „Iskra”, oraz czy konsumenci są skłonni do jego zakupu.

Identyfikujemy obszary badawcze:

- Identyfikacja cech świeczek, dla których są one kupowane.
- Identyfikacja zależności pomiędzy zapachem świeczki, a wartością produktu dla klienta.
- Identyfikacja zależności pomiędzy opakowaniem świeczki a jej wartością dla klienta.

Cele szczegółowe można sformułować następująco:

- Określenie, które z przedstawionych cech są najważniejsze dla klienta rozważającego zakup świeczki zapachowej: cena, opakowanie, rekomendacja przyjaciela lub znajomego, dostępność, zapach, kolor, wzór na świecy, zadymienie, które powoduje.
- Określenie, które z następujących zapachów klienci kupują najchętniej: malinowy, pomarańczowy, brzoskwiniowy, melona, miodowy.
- Określenie, jak konsumenci oceniają świeczki firmy „Iskra” pod względem kluczowych cech.
- Określenie fizjologicznej reakcji konsumentów na zapach świeczek i opakowania.

Docelowych respondentów³² w omawianym przykładzie możemy określić, jako istniejących i potencjalnych klientów firmy „Iskra” oraz konkurencji. Badanie w zależności od celów strategicznych przedsiębiorstwa możemy przeprowadzić na istniejących klientach według wyodrębnionego wzorca, bądź na klientach potencjalnych, charakteryzujących się tymi samymi kryteriami demograficznymi (np. kobiet w wieku 20-50 lat, mieszkańców miast i wsi, o dochodzie 1500-3000 zł miesięcznie).

Przykład 3. Przykład ukazuje proces transformacji problemu decyzyjnego w problem badawczy

Tabela 1.

Proces transformacji problemu decyzyjnego w problem badawczy

Problem decyzyjny	Problem badawczy
Zmiana kanału dystrybucji	Ocena obecnego kanału dystrybucji, ocena rozpatrywanego kanału dystrybucji
Powiększenie grupy docelowej	Ocena obecnego wizerunku firmy, analiza postaw wybranej grupy docelowej, analiza środków promocji
Zwiększenie obszaru geograficznego dystrybucji produktów	Analiza obecnego kanału dystrybucji, analiza kanału dystrybucji na nowym obszarze, analiza sprzedaży w poszerzonym obszarze geograficznym
Opracowanie nowego opakowania dla starego produktu	Ocena postaw konsumentów wobec obecnego opakowania, analiza wizerunku firmy, ocena propozycji nowego opakowania

Źródło: opracowanie własne.

³² R.J. Kaden formułuje pytania, na które należy odpowiedzieć, definiując rynek docelowy: „Jakie rynki docelowe stwarzają największe szanse jak najszybszego rozwoju firmy? Jakie rynki będą ważne dla długoterminowego rozwoju firmy? Jakie grupy będzie można najtaniej przekonać do zwiększenia zakupów w naszej firmie? Jacy respondenci mogliby dostarczyć informacji, które chętnie byśmy poznali, lecz ich pozyskanie do współpracy byłoby zbyt kosztowne? Czy znając postawy respondentów z jakiegoś jednego określonego rynku docelowego, moglibyśmy ekstrapolować je na inne rynki bez ich badania? Czy nasz budżet na badania pozwala nam zarzucić sieci bardzo szeroko oraz czy kiedy zgromadzimy dane równocześnie z wielu rynków docelowych, będziemy mogli później, gdy zapoznamy się już z danymi, ustalić, na których rynkach sprzedaż jest najłatwiejsza?”. Por. R.J. Kaden, Badanie marketingowe, Polskie Wydawnictwo Ekonomiczne, Warszawa 2008, s. 32.

Przykład 4. Przykład raportu cząstkowego z badania zrealizowanego przez Centrum Badań Marketingowych INDICATOR³³ nt. wykorzystania usług bankowych i planów na 2009 rok w małych i średnich firmach. W tym częściowym raporcie znajdują się m.in. informacje, kto przeprowadził badanie, na czyje zlecenie, jaką metodą, jaka była liczność próby, z kim należy się skontaktować w sprawie uzyskania szczegółowych informacji o badaniu.

„Z przyjemnością informujemy, że są już dostępne standardowe raporty z realizacji XI edycji badania syndykatowego „Zapotrzebowanie małych i średnich przedsiębiorstw na usługi poszczególnych banków”. Badanie wykonane zostało we współpracy z Warszawskim Instytutem Bankowości. Badanie zostało przeprowadzone w I kwartale 2009 roku metodą zestandaryzowanych indywidualnych wywiadów kwestionariuszowych. Realizowaliśmy łącznie blisko 1200 wywiadów, z przedstawicielami małych przedsiębiorstw (o obrotach do 3,6 mln PLN rocznie) oraz średnich przedsiębiorstw (o obrotach powyżej 3,6 mln do 30 mln PLN rocznie).(...)

Raporty zawierają następujące zagadnienia:

- Rynekowa rozpoznawalność banków wśród małych i średnich przedsiębiorstw
- Banki o podstawowym znaczeniu dla małych i średnich przedsiębiorstw,
- Lojalność względem banków obsługujących małe i średnie firmy
- Czynniki decydujące o wyborze banku
- Aktywność ofertowa banków w segmencie małych i średnich przedsiębiorstw
- Korzystanie małych i średnich przedsiębiorstw z produktów i usług bankowych
- Alternatywne kanały dystrybucji
- Opinie na temat kosztów korzystania z produktów i usług bankowych
- Poziom zadowolenia małych i średnich firm ze współpracy z bankami
- Kierunki lokowania przez małe i średnie przedsiębiorstwa wolnych środków finansowych
- Rynek kredytów bankowych dla małych i średnich przedsiębiorstw
- Uwarunkowania rozwoju sektora małych i średnich przedsiębiorstw w Polsce
- Poziom korzystania ze środków z funduszy europejskich

W sprawach związanych z tym badaniem prosimy o kontakt:

Wojciech Terlikowski

wterlikowski@indicator.pl

tel. (22) 595 1 595³⁴

³³ <http://www.indicator.pl> (11.03.2010)

³⁴ <http://www.indicator.pl> (11.03.2010).

2.4. Zadania do samodzielnego opracowania

Zadanie 1. W każdej z przedstawionych sytuacji zidentyfikuj źródło problemu, problem decyzyjny oraz możliwe problemy badawcze.

a) Firma „Pod Złotym Krążkiem” jest lokalnym zakładem jubilerskim usytuowanym w bocznej uliczce miejscowego rynku. Ostatnio, w wyniku otwarcia w centrum rynku salonu jednego z zagranicznych producentów jubilerskich, sprzedaż „Pod Złotym Krążkiem” spadła w ciągu dwóch miesięcy o 30%. Badanie przeprowadzone samodzielnie przez właściciela „Pod Złotym Krążkiem”, wśród klientów zakładu, ujawniło, że głównym powodem spadku sprzedaży jest atrakcyjny system rabatów zastosowany przez konkurencję.

b) Firma „IT Profesor” cieszyła się dużym powodzeniem na rynku krajowym w sprzedaży oprogramowania „Drawing”, przeznaczonego do rysowania skomplikowanych procesów technologicznych, biznesowych i aranżacyjnych. Przeprowadzona dwa lata temu promocja programu „Drawing” wśród średnich i dużych przedsiębiorstw, przyniosła oszałamiający skutek, w postaci zamówień oprogramowanie przez 80% przedsiębiorstw. Jednakże konieczność ciągłych nowelizacji programu „Drawing” oraz brak jego kompatybilności z systemami operacyjnymi, stanowi przyczynę niezadowolenia klientów. Spośród dotychczasowych klientów jedynie 30% deklaruje zakup nowego oprogramowania „Drawing Professional”.

c) Supermarket „Hipcio” zlokalizowany jest na dużym osiedlu mieszkaniowym na obrzeżach 100 tys. miasta. Kierownik sklepu przyjął nowego pracownika na okres 6 miesięcy w ramach umowy z Powiatowym Urzędem Pracy. W ciągu pierwszych 3 dni pracy, pracownik został skierowany do pracy przy kasach, polegającej na pakowaniu zakupionego przez klientów towaru do siatek. W ciągu następnego tygodnia liczba osób, robiących zakupy w supermarkecie „Hipcio” podwoiła się.

Zadanie 2. Odpowiedz na pytania:

- a) Jakie są zalety przeprowadzania badań marketingowych we własnym zakresie? Jakie są wady przeprowadzania badań marketingowych we własnym zakresie?
- b) Jakie są zalety przeprowadzania badań marketingowych przez wynajętą profesjonalną firmę?
- c) Jakie są wady przeprowadzania badań marketingowych przez wynajętą profesjonalną firmę?

- d) Wymień nazwy trzech dużych, znanych firm zajmujących się badaniami marketingowymi.
- e) Czy w województwie, w którym mieszkasz znajdują się firmy zajmujące się badaniami marketingowymi? Jeśli nie wiesz – sprawdź np. w „Panoramie Firm” lub Internecie.

3. POMIAR W BADANIACH MARKETINGOWYCH

3.1. Skala pomiaru

Po określeniu problemu badawczego oraz hipotez badawczych, badacz staje przed problemem pomiaru. Pomiaru, czyli przyporządkowania właściwościom empirycznym – symboli, zgodnie z określonymi zasadami.

Pomiar w badaniach marketingowych oznacza przypisanie określonych symboli cechom obiektów takich jak przedmioty, zjawiska, konsumenci, zgodnie z określonym zbiorem reguł³⁵. Pomiaru dokonujemy na cechach obiektu, a nie na obiekcie samym w sobie. Przedmiotem pomiaru osoby jest jej cecha: płeć, wykształcenie, klasa społeczna, waga, wzrost. Skale przypisujemy i odczytujemy za pomocą przyjętej i zrozumiałej skali pomiaru. Mierząc wzrost danej osoby, narzędziem pomiaru jest centymetr z określoną skalą wyrażoną w centymetrach. Jest to skala powszechnie znana i zrozumiała. Mierzenie postaw osoby badanej nie jest już zadaniem tak prostym, dlatego istotne jest zrozumienie zastosowanej skali oceny. Rozważmy właściwości metra, przyjętego do pomiaru wzrostu osoby badanej. Metr składa się z podziałki na centymetry. Badacz przyjmuje, że liczba 1 oznacza jeden centymetr, liczba 2 – dwa centymetry itd. Badacz jest w stanie powiedzieć, że 1 centymetr jest mniejszy od 2 centymetrów, oraz że 2 centymetry są mniejsze od 3 centymetrów. Prowadzący badanie może na podstawie przyjętej skali stwierdzić, że 3 centymetry są trzykrotnie większe niż 1 centymetr, a 4 centymetry są dwukrotnie większe niż 2 centymetry. Przyjęta skala pomiaru jest właściwa dla pomiaru przyjętej cechy, jaką jest wzrost. Jednakże gdybyśmy przypisali tę samą skalę pomiaru do zmierzenia postaw, okazałaby się ona niewłaściwa. Nie moglibyśmy stwierdzić jednoznacznie, że osoba badana numer 1, dwa razy bardziej gustuje w czekoladzie niż osoba badana numer 1. Skala ta nie mogłaby zostać zastosowana. Skala jest więc odwzorowaniem, czyli modelem rzeczywistych zjawisk i ich relacji. Skalowaniem nazywamy czynność odwzorowania mierzonej cechy za pomocą wybranej skali³⁶.

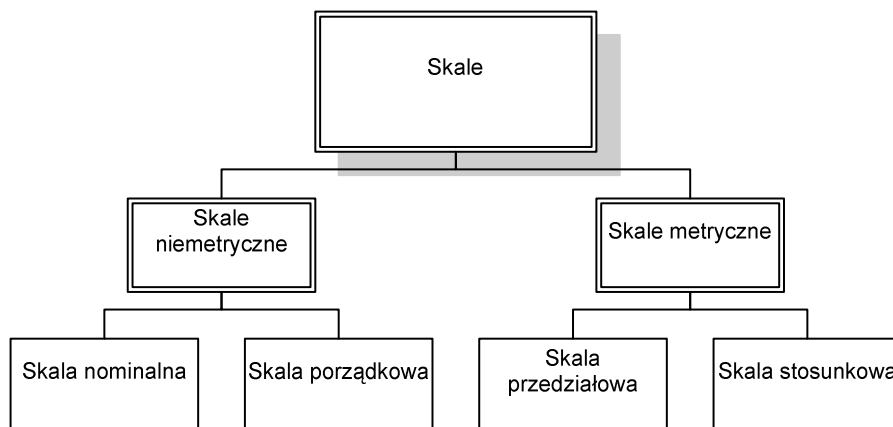
Earl Babbie łączy rozróżnianie danych ilościowych i jakościowych w badaniach z rozróżnianiem danych liczbowych i nieliczbowych³⁷. Należy jednak pamiętać, iż liczby nie zawsze mają znaczenie ilościowe (np. za pomocą liczb możemy określać miesiące, województwa, studentów danej

³⁵ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 74.

³⁶ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 85.

³⁷ Por. E. Babbie, *Badania społeczne w praktyce*, Wydawnictwo Naukowe PWN, Warszawa 2003, s. 48.

uczelnii, itp). W zależności od rodzaju danych, jakie skale generują, możemy rozróżnić skale metryczne (ilościowe) i niemetryczne (jakościowe). Rozróżniamy różne typy skali, według których mierzymy cechy obiektów: skalę nominalną, skalę porządkową, skalę przedziałową (interwałową), skalę stosunkową (ilorazową) (por. rys. 4).



Rys. 4. Skale pomiaru

Źródło: opracowanie własne.

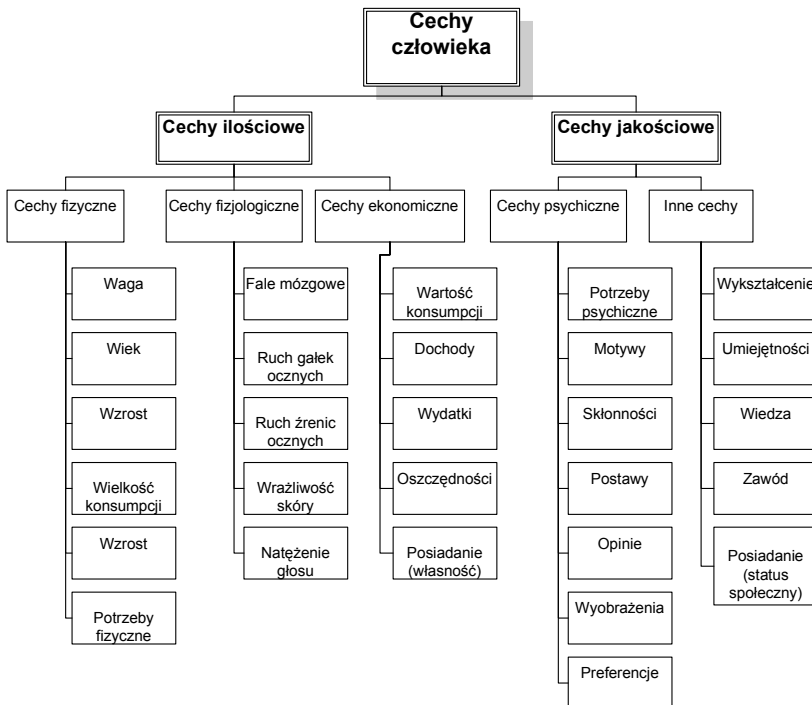
Cechy obiektów dzielimy na jakościowe i ilościowe. Pomiar cechy ilościowej jest łatwy do wyobrażenia: wzrost, wiek, waga. Natomiast cechy jakościowe są bardziej skomplikowane w swoim pomiarze. Na poniższym schemacie zilustrowane są przykładowe cechy człowieka, które mogą stanowić przedmiot pomiaru. Operacje matematyczne i statystyczne (por. tabela 4), które możemy przeprowadzić na otrzymanych w badaniu danych, zależą od przyjętej skali pomiaru (nominalna, porządkowa, przedziałowa, stosunkowa).

Tabela 2.

Skale pomiaru

Skala	Podstawowe porównania	Typowe przykłady	Miary
Nominalna	Tożsamość	Mężczyzna – kobieta Używający – nieużywający Klasyfikacja zawodów Numeracja zawodników	Dominanta
Porządkowa	Porządek	Preferencja marki Klasa społeczna Twardość materiałów Klasa, jakości drewna	Dominanta, mediana
Przedziałowa	Porównanie podziałów	Skala temperatury (°C) Postawa wobec marki	Dominanta, mediana, średnia arytmetyczna
Stosunkowa	Porównanie wielkości absolutnych	Sprzedane sztuki Liczba nabywców Prawdopodobieństwo zakupu Waga	Dominanta, mediana, średnia arytmetyczna

Źródło: G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 404.



Rys. 5. Cechy jakościowe i ilościowe człowieka

Źródło: S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 78.

Najniższym poziomem pomiaru jest poziom nominalny, klasyfikujący zbiór obiektów na wyczerpujące³⁸, wzajemnie rozłączne³⁹ kategorie. Właściwością skali nominalnej jest tożsamość. Numeracja użyta w skali nominalnej jedynie identyfikuje jednostkę przypisaną dla niej liczbą⁴⁰. Osoby badane są identyfikowane, jako posiadacze danej cechy lub nie. Jeśli chcemy zmierzyć cechę „płeć” wówczas badanym płci męskiej przypisujemy liczbę 0, natomiast badanym płci żeńskiej przypisujemy liczbę 1. Prowadzący badanie za pomocą zastosowanej skali nie jest w stanie określić nic innego oprócz liczby mężczyzn i liczby kobiet. Fakt, iż kobietom nadano numer 1, a mężczyznom numer 0, nie oznacza, iż mężczyźni są gorsi od kobiet, i nie oznacza to także, że jest ich więcej niż kobiet. W wyniku zastosowania skali nominalnej do płci możemy uzyskać jedynie wynik „kobieta”, „mężczyzna”. Jediną dopuszczalną operacją matematyczną jest sumowanie, policzenie liczby uzyskanych odpowiedzi należących do kategorii płci, za pomocą przyjętych symboli. Prowadzący badanie może określić, że w badanej próbie znajduje się 70% kobiet i 30% mężczyzn. Nie może on natomiast powiedzieć, że w danej grupie średnia osób wynosi 2,3.

Skale nominalne możemy podzielić na:

- skale nominalne alternatywne – z dwudzielną klasyfikacją
- skale nominalne niealternatywne – z wielodzielnią klasyfikacją⁴¹.

Skale nominalne niealternatywne posiadają co najmniej trzy klasy. Charakterystyczną cechą tej skali jest dowolna kolejność klas. Jedynym wymaganiami, jakie musi spełniać skala nominalna niealternatywna, jest jednoznaczność klasyfikacji. Klasyfikacja jednoznaczna powinna być zarazem wyczerpywalna i rozłączna. Rozłączność oznacza, że jeden element skali może znaleźć się w skali, tylko w jednej klasie. Spełnienie tego warunku jest konieczne w celu dokonania prawidłowego pomiaru.

Właściwością skali porządkowej jest porządek. Obiekty są szeregowane za pomocą skali porządkowej w relacji do posiadanej cechy. Przyporządkowanie liczb badanym cechom wskazuje nie tylko relacje równości lub różności, lecz również ich uporządkowanie ze względu na tę cechę. Przyporządkowanie to ma znaczenie względne. Oznacza to, że liczba większa występuje w większym nasileniu niż liczba mniejsza⁴². Można

³⁸ Wyczerpujące – zawierające wszystkie obiekty określonego rodzaju).

³⁹ Wzajemnie rozłączne – każdej kategorii przypisujemy inny symbol, a żaden z obiektów nie może być jednocześnie zaliczony do więcej niż jednej kategorii.

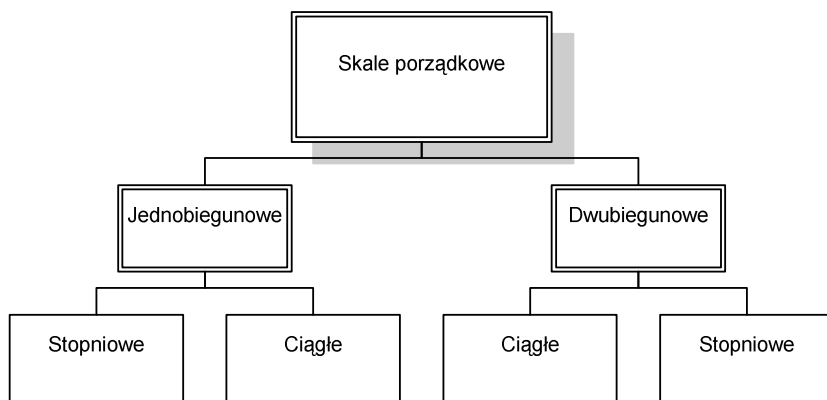
⁴⁰ Por. G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 404.

⁴¹ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 85-86.

⁴² A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 72.

powiedzieć, że kolejne liczby 1, 2, 3, 4 oznaczają nasilenie właściwości. Uszeregowanie liczby dzieci szkoły podstawowej: 1 – uczniowie klas pierwszych, 2 – uczniowie klas drugich, 3 – uczniowie klas trzecich, 4 – uczniowie klas czwartych świadczy także o ich zaawansowaniu i poziomie wiedzy, którą posiadają. Nie jesteśmy natomiast w stanie stwierdzić nic ponad to. Należy pamiętać, że zastosowanie skali porządkowej musi być przemyślane. Prowadzący badanie musi tak dobrać badaną cechę, aby ta sama w sobie charakteryzowała się porządkiem. Cechy porządkuje się według relacji: większy – mniejszy, silny – słaby, mało – dużo⁴³. Zastosowanie skali porządkowej nie da badaczowi informacji na temat wielkości różnic pomiędzy kolejnymi przedziałami. Wiemy, że 2 jest większe od 1, ale nie wiemy o ile. Nie możemy tego stwierdzić, ponieważ nie posiadamy jednostki miary.

Skale porządkowe są szeroko stosowane w badaniach marketingowych. Skale porządkowe mogą być klasyfikowane ze względu na kryterium ich konstrukcji: skale jednobiegunowe, skale dwubiegunowe, skale stopniowe, skale ciągłe⁴⁴.



Rys. 6. Podział skal porządkowych

Źródło: S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, PWE, Warszawa 1995, s. 89.

Skale jednobiegunowe charakteryzują się posiadaniem jednego bieguna (drugi biegun zostaje samoczynnie, z góry określony). Skale dwubiegunowe, polegają na tym, że drugi biegun skali nie został określony przez pierwszy. Drugi biegun skali stanowi przeważnie przeciwieństwo, lecz

⁴³ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 87.

⁴⁴ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 88-90.

nie stanowi negacji. Skale stopniowe zawierają w sobie podział na przedziały. Przedziały w skalach stopniowych nie są sobie równe, natomiast ich graficzna ilustracja przedstawia to w sposób przeciwny. Celem graficznego przekazu jest nieprzekazywanie respondentowi żadnych sugestii, co do wyboru. Skale ciągłe stanowią jak sama nazwa wskazuje, ciąg wartości od najniższej do najwyższej.

Skala przedziałowa należy do skal metrycznych. Skala ta posiada określone odległości między wyznaczonymi przedziałami. Można powiedzieć, że skala przedziałowa jest mocniejsza od skali porządkowej, gdyż dostarcza informacji na temat wielkości różnic pomiędzy przedziałami. Skala przedziałowa nie pozwala prowadzącemu badania na porównanie absolutnej wielkości liczb, gdyż punkt zerowy na skali przedziałowej ustalany jest arbitralnie. Oznacza to, że każda modyfikacja formy $y=a+bx$, gdzie b jest dodatnie, x jest liczbą wyjściową, a y jest liczbą przekształconą, zachowa właściwość skali⁴⁵. Relację tę można opisać za pomocą pomiarów temperatury. Badacz dokonując pomiaru temperatury w danym roku, stwierdza, że maksymalna temperatura powietrza w ciągu roku osiągnęła 70°F, a najniższa 35°F. Stwierdzenie, że najwyższa temperatura roku jest dwukrotnie większa niż najniższa temperatura roku, jest nieprawdziwe. Prowadzący badanie, dokonując przeliczenia temperatury na stopnie Celsjusza, dowiaduje się, że najwyższa temperatura roku wyniosła 21,1°C, a najniższa temperatura roku wyniosła 1,7°C. Jak widzimy relacja wyrażona w stopniach Celsjusza jest zupełnie inną relacją niż ta wyrażona w stopniach Farenheita. Kolejnym przykładem ilustrującym zastosowanie skali przedziałowej może być ocena produktów pod względem ostrości. Każdemu produktowi A, B, C, D możemy przypisać liczbę z przedziału [1, 7], gdzie: 1 – zdecydowany brak ostrości, 2 – nieznacznie ostry, 3 – lekko ostry, 4 – średnio ostry, 5 – ostry, 6 – bardzo ostry, 7 – wyjątkowo ostry. Przypisując produktom liczby A=7, B=3, C=2, D=3, E=4, badacz nie jest w stanie stwierdzić, że produkt E jest dwa razy bardziej ostry od produktu C, gdyż skala ta nie posiada miary, ani ustalonego punktu zerowego.

Skala stosunkowa wyróżnia się spośród wszystkich skal tym, że posiada punkt zerowy, co do położenia którego nie ma wątpliwości ani zastrzeżeń (zero absolutne). Skala stosunkowa charakteryzuje się tym, że posiada równe różnice pomiędzy badanymi cechami, którym przyporządkowane zostają równe liczby, a przyporządkowanie to ma charakter proporcjonalny do stopnia, w jakim cecha występuje w mierzonych obiektach⁴⁶. Punkt zerowy w skali stosunkowej oznacza fizyczny brak

⁴⁵ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 406.

⁴⁶ A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 73.

występowania danej cechy. Skala stosunkowa wykorzystuje miary centymetrowe, kilogramowe. Oczywistymi cechami, które możemy badać za pomocą skali stosunkowej są waga, wzrost. Skale stosunkowe pozwalają na proporcjonalne przekształcenie wartości. Proporcjonalne przekształcenie wartości można przedstawić, jako $y=bx$, gdzie x reprezentuje wartości pierwotne, a y wartości przekształcone, b jest pewną dodatnią stałą. Skala stosunkowa jest najdokładniejszą ze wszystkich dotychczas wymienionych skal⁴⁷.

3.2. Trafność i rzetelność pomiaru

Dokonany przez badacza pomiar wartości, ażeby był godny zaufania, powinien spełniać dwa podstawowe kryteria: trafności i rzetelności. Pomiar jest trafny, jeśli mierzy tę cechę, którą powinien mierzyć. Pomiar jest rzetelny, wtedy gdy powtórzony w tych samych warunkach daje te same wyniki. Prowadzący badanie może mówić o wiarygodności pomiaru, jeśli jest on pozbawiony błędów systematycznych i błędów przypadkowych. Istnieją cztery główne kombinacje błędów w pomiarze.

Tabela 3.

Macierz błędów systematycznych i przypadkowych

	Rzetelność duża	Rzetelność mała
Trafność duża	Pomiary trafne i rzetelne – są osiągane wtedy, gdy wynik dokonanego pomiaru jest bliski wynikowi oczekiwanemu. Wówczas mówimy o dużej trafności i rzetelności. Pomiar taki daje małe błędy systematyczne i przypadkowe.	Pomiary trafne, ale nierzetelne – wyniki uzyskane z takiego pomiaru nie stanowią realizacji celu badania. Pomiar taki obarczony jest małymi błędami systematycznymi oraz poważnymi błędami przypadkowymi.
Trafność mała	Pomiary rzetelne, ale nietrafne – wyniki uzyskane z takiego pomiaru, nie stanowią realizacji celu badania. Pomiar taki obarczony jest małymi błędami przypadkowymi, lecz poważnymi błędami systematycznymi	Pomiary nierzetelne i nietrafne – obarczone zarówno błędami systematycznymi jak i przypadkowymi, nie mogą dostarczać poprawnych wyników

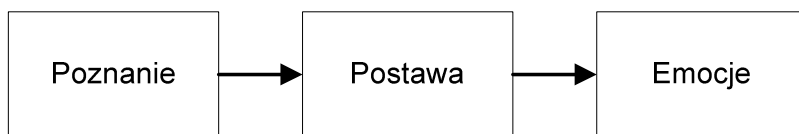
Źródło: opracowanie własne na podstawie S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 93.

⁴⁷ Por. G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 407, A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 73.

Trafność pomiaru jest miarą dokładności instrumentu pomiarowego. W badaniach marketingowych ocenia się: trafność prognostyczną, trafność diagnostyczną, trafność wewnętrzną i trafność teoretyczną⁴⁸. Trafność prognostyczna polega na wyliczeniu korelacji pomiędzy wynikami pomiaru a zachowaniem się mierzonej cechy w późniejszym okresie. Prowadzący badanie sprawdza, na ile dokonana prognoza zostanie potwierdzona w dalszym okresie czasu. Trafność diagnostyczna polega na obliczeniu współczynnika korelacji, zwanego współczynnikiem trafności diagnostycznej, pomiędzy wynikami dokonanego pomiaru a wynikami pomiaru tej samej cechy w tym samym czasie za pomocą innego instrumentu. Jedynym zastrzeżeniem jest, żeby drugi pomiar także był trafny. Trafność wewnętrzna jest osiągnięta, jeśli definicja operacyjna przedmiotu pomiaru jest poprawna. Trafność teoretyczna polega na udowodnieniu powiązania wyników pomiaru do serii pomiarów próbnych lub kontrolnych.

3.3. Wybrane skale pomiaru postaw

Postawę definiujemy jako indywidualną predyspozycję do oceniania przedmiotu bądź pewnego aspektu świata w korzystny lub niekorzystny sposób⁴⁹. Postawy wobec produktów bądź usług mogą powstawać poprzez skojarzenia emocjonalne. Nabywcy wykształcają postawy długookresowo, dlatego menedżerowie ds. marketingu przywiązują tak dużą wagę do postaw. Zmiana postawy wymaga intensywnej promocji, wpływu zarówno na przekonania jak i uczucia nabywcy⁵⁰.



Rys. 7. Wpływ procesu poznania na kształtowanie się postaw i emocji

Źródło: opracowanie własne

⁴⁸ Por. S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 95-96.

⁴⁹ A. Gerrit, W.F. van Raaij, *Zachowanie konsumenta. Podręcznik akademicki*, Wydawnictwo Naukowe PWN, Warszawa 2003, s. 212.

⁵⁰ K. Mazurek-Łopacińska, *Zachowania nabywców jako podstawa strategii marketingowej*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1997, s. 205.

Niezgodność pomiędzy postawą a zachowaniem nabywcy określa się mianem dysonansu poznawczego. Jest to negatywny stan emocjonalny, który powodowany jest wystąpieniem negatywnych emocji. Menedżerowie podejmują szereg działań marketingowych mających na celu zminimalizowanie dysonansu poznawczego przejawiających się poprzez:

- dostarczanie dodatkowych informacji o produkcie, możliwościach jego wykorzystania,
- reklamę skierowaną do aktualnych użytkowników, podkreślającą walory zakupionego produktu i umacniające odczucia pozakupowe,
- utrzymywanie bezpośredniego kontaktu z nabywcą już po dokonaniu przez niego zakupu produktu.

W literaturze wskazuje się na zasady, którymi powinien kierować się każdy badacz przy skalowaniu odpowiedzi w pomiarze postaw (skalowanie zrównoważone i niezrównoważone, graficzna forma skali, parzysta lub nieparzysta liczba kategorii skali, skalowanie wymuszające i nie wymuszające).

Skalowanie zrównoważone polega na określeniu klasy środkowej oraz wyznaczeniu identycznych przedziałów (klas) o identycznej intensywności po obu stronach klasy środkowej⁵¹. Innymi słowy skalowanie zrównoważone polega na wyznaczeniu jednakowej liczby odpowiedzi pozytywnych jak i negatywnych po obu stronach skali środkowej. Skalowanie niezrównoważone jest to takie skalowanie, w którym liczba klas nie jest identyczna, bądź intensywność klas nie jest identyczna po obu stronach klasy środkowej. Innymi słowy nie posiada tej samej liczby odpowiedzi pozytywnych i negatywnych.

Każda skala zawiera kategorie, które mogą być przedstawione w różny sposób. Precyzja w opisie kategorii skali jest istotna. Im dokładniejsza kategoria skali, tym pomiar jest bardziej rzetelny. Graficzną formę skali stosuje się przy pomiarze, w którym grupę respondentów stanowią dzieci lub dorośli, nie umiejący pisać ani czytać⁵². Wówczas kategorię skali przedstawia się w formie obrazków. Użycie graficznej formy skali polega na ułatwieniu respondentowi wskazania odpowiedniego punktu na linii biegnącej od jednej skrajności do drugiej.

Skalowanie wymuszające polega na wymuszeniu określenia postawy przez respondenta. Gdy respondent nie jest w stanie wyrazić swojej postawy, wówczas stosuje się skalowanie niewymuszające. Skalowanie niewymuszające pozwala na zajęcie stanowiska neutralnego przez respondenta.

⁵¹ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 115.

⁵² S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 117.

Liczba kategorii skali zależy od przyjętej przez prowadzącego badanie formy pomiaru, metody analizy danych oraz samej postawy respondenta. Stosując nieparzystą liczbę kategorii, z możliwością zaznaczenia postawy neutralnej, jesteśmy w stanie uzyskać więcej informacji od respondenta. Skale parzyste, bez możliwości zaznaczenia postawy neutralnej, zmuszają respondentów do określenia swojej postawy. Skale pięciopunktowe dają respondentowi możliwość wybrania postawy neutralnej⁵³. W praktyce zdania są podzielone między zwolennikami skali czteropunktowej, którzy uważają, że postawy nie są neutralne, a innymi, którzy twierdzą, że liczba kategorii skali nie stanowi większej różnicy⁵⁴. Decyzja jakiej skali użyć, zależy przede wszystkim od typu respondentów, z jakimi mamy do czynienia oraz cechą, która ma być badana. Respondenci posiadający doświadczenie w przedmiocie pytania, nie będą wskazywać postawy neutralnej. Istnieje duże prawdopodobieństwo, że wskażą oni postawę pozytywną lub negatywną względem danego zjawiska. Wówczas nie istnieje potrzeba stosowania skali nieparzystej⁵⁵.

Prowadzący badanie dokonując wyboru skali musi kierować się przede wszystkim rzetelnością pomiaru. Rzetelność pomiaru jest bardzo ważnym wskaźnikiem jakości pomiaru. Zapewnienie wiarygodności pomiaru zależy przede wszystkim od jego rzetelności.

⁵³ R.J. Kaden, *Badanie marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2008, s. 142.

⁵⁴ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 118.

⁵⁵ R.J. Kaden, *Badanie marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2008, s. 143.

Tabela 4.

Wpływ wybranych cech pomiaru na jego rzetelność

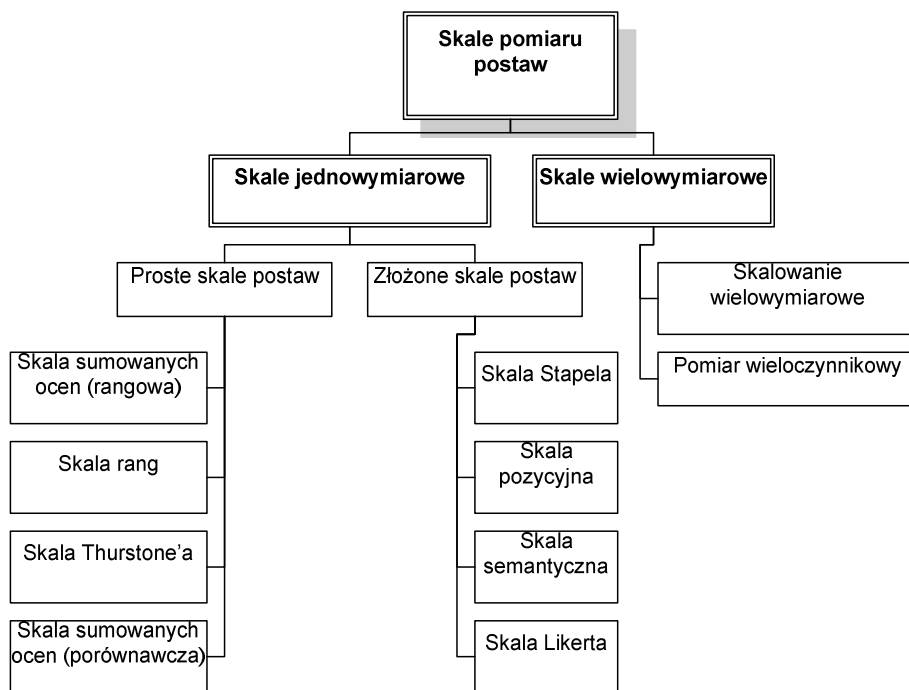
Cecha pomiaru	Wniosek
Liczba pozycji na ostatecznej skali	Hipoteza o istnieniu dodatniej zależności pomiędzy liczbą użytych w skali pozycji i rzetelnością pomiaru jest uzasadniona
Trudność pozycji	Hipoteza o istnieniu zależności pomiędzy trudnością punktowania pozycji i rzetelnością pomiaru nie jest uzasadniona.
Odwrócona punktacja	Hipoteza, iż skale z odwróconą punktacją będą miały niższą rzetelność niż skala bez niej, nie jest uzasadniona.
Typ skali	Nie dokonano <i>a priori</i> przewidywania o wyższości jednego typu skali, nie stwierdzono związku między typami skali i rzetelnością pomiaru.
Liczba punktów skali	Hipoteza o dodatniej zależności pomiędzy liczbą punktów skali ponad zakres normalny i rzetelnością pomiaru jest uzasadniona
Typ oznakowań	Nie przyjęto <i>a priori</i> założenia o wyższości oznakowań numerycznych i werbalnych nad jedynie werbalnymi lub <i>vice versa</i> nie stwierdzono też żadnego związku pomiędzy typem opakowania i rzetelnością pomiaru
Zakres opisu punktów skali	Hipoteza, iż skale, dla których wszystkie punkty są oznaczone, wykazują wyższą rzetelność niż skale, dla których oznaczono tylko punkty biegunowe, nie znajduje uzasadnienia
Niepewność lub niewiedza respondenta	Hipoteza, iż skale z punktami neutralnymi mają wyższą rzetelność niż skale o wymuszonym wyborze, nie znajduje uzasadnienia.

Źródło: G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 432.

Na rysunku 8 przedstawiona jest klasyfikacja skal do pomiaru postaw. Skala rang jest to skala, przy pomocy której respondent porządkuje cechy obiektu według podanego kryterium. Kwestionariusz zawiera odpowiedzi wyskalowane w skali nominalnej. Respondent dokonując oceny wybranych cech, przekształca skalę nominalną w skalę porządkową. Skale rang wykorzystywane są do pomiaru preferencji⁵⁶.

⁵⁶ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 120.

Skala sumowanych ocen – respondent jest proszony o to, aby wymienionym przez badacza cechom przydzielił zgodnie ze swoimi preferencjami punkty. Punkty powinny być rozdzielone tak, aby dawały podaną przez badacza sumę. Najczęściej stosuje się sumę 10 lub 100 punktów.



Rys. 8. Klasyfikacja skal do pomiaru postaw

Źródło: S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 120.

Skala pozycyjna składa się przede wszystkim ze skal nominalnych. Każda skala nominalna oceniana jest przez respondentów za pomocą jednej i tej samej skali porządkowej. Respondent ocenia każdą z wymienionych cech obiektu wyrażonych poprzez skale nominalne, zgodnie ze swoimi preferencjami za pomocą skali porządkowej⁵⁷.

⁵⁷ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 124.

Skala semantyczna (inaczej: skala zróżnicowania semantycznego lub skala dyferencjału semantycznego) jest wykorzystywana do pomiaru znaczeń, jakie respondent przypisuje obiektowi badania. Skala semantyczna skonstruowana jest jako szereg prostych skal dwubiegunowych. W literaturze wyróżnia się następujące etapy budowy skali semantycznej⁵⁸:

- Dobór obiektu badań – nazwy firm, marek, funkcji lub cech produktów, sloganów reklamowych,
- Określenie około 10-14 par opozycyjnych terminów – wskazane jest użycie przymiotników, opisujących dany obiekt. Przymiotniki powinny odpowiadać trzem podstawowym wymiarom:
 - wartościowaniu (np. dobry – zły, kwaśny – słodki, smaczny – niesmaczny)
 - potencji (np. silny – słaby, niski – wysoki, głęboki – płytki)
 - aktywności (np. szybki – wolny, głośny – cichy)
- Opracowanie wzorców odpowiedzi z wykorzystaniem szal szacunkowych.

Skala semantyczna składa się z kilku skal porządkowych, najczęściej siedmiopunktowych, dwubiegunowych. Środkowy przedział każdej skali jest neutralny. W skali semantycznej istotne jest także ułożenie sformułowań negatywnych i pozytywnych. Rozmieszczenie sformułowań negatywnych raz po lewej stronie, a raz po prawej stronie, zapobiega powstawaniu tzw. „efektu halo”, czyli prostemu zakreślaniu odpowiedzi przez respondenta⁵⁹. Skala zostaje następnie przedstawiona grupie respondentów, proszonych o zakreślenie właściwego okienka odzwierciedlającego jego odczucia. Każdy z respondentów zakreśla swoje preferencje dla tej samej skali semantycznej, celem porównania wyników pomiaru. Prezentacja wyników ma postać wykresu, nazywanego profilem polaryzacji⁶⁰. Następnie porównujemy ze sobą profile pochodzące od różnych respondentów. Przy budowie profilu dla ułatwienia wszystkie pozytywne przymiotniki badanej cechy umieszcza się po prawej stronie wykresu. Poszczególnym punktom na skali semantycznej możemy przypisać oznaczenia : -3, -2, -1, 0, 1, 2, 3 lub 1, 2, 3, 4, 5, 6, 7⁶¹. Następnie dokonujemy analizy zebranych wyników polegającej na obliczeniu dystansów semantycznych między obiektami.

⁵⁸ A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 83-84.

⁵⁹ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 125.

⁶⁰ A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 85.

⁶¹ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 421.

Skala Stapela jest zestawem skal jednobiegunowych, dziesięcio- lub jedenastopunktowych. Przeważnie bieguny oznaczane są -5, -4, -3, -2, -1, +1, +2, +3, +4, +5. Respondenci przyznają liczby dodatnie dla swoich pozytywnych postaw. Punkty ujemne są przyznawane dla postaw negatywnych. Punkt neutralny, jeśli w skali występuje, oznacza brak postawy respondenta. Do podstawowych zalet skali Stapela zaliczamy: łatwość opracowania oraz łatwość zastosowania. W przypadku skali Stapela prowadzący badanie nie jest zmuszony do dokładnego doboru przymiotników opisujących cechy obiektu badanego. Skala Stapela jest wykorzystywana do pomiaru opinii i wyobrażeń o nowych produktach czy miejscach sprzedaży⁶².

Skala Likerta jest najbardziej popularną skalą stosowaną w badaniach marketingowych. Zastosowanie skali Likerta polega na dokonaniu oceny respondenta według podanych kryteriów co do zawartych w skali twierdzeń. Twierdzenia powinny być dobrane pod kątem prowadzonego badania, ale także powinny umożliwiać respondentowi odpowiedź „tak”, „nie”, lub „nie wiem”. Stwierdzenia powinny być dobrane w ten sposób, aby ich charakter był stopniowalny⁶³. Każda pozycja skali Likerta jest sama w sobie skalą, przeważnie dwubiegunową, zazwyczaj pięciostopniową.

3.4. Przykłady z omówieniem

Przykład 1. Odpowiedzmy na pytanie: jak wybrać skalę pomiaru? Procedurę wyboru skali pomiaru możemy uporządkować w kilku podstawowych krokach.

Krok 1. Zidentyfikuj jednostkę doboru próby. W tym celu prowadzący badanie powinien odpowiedzieć na pytanie: „Co chcę badać?”. Badaniu może podlegać gospodarstwo domowe, zespół, miasto, kraj, województwo czy studenci Politechniki Opolskiej.

np. Prowadzący badanie chce określić liczbę studentów, którzy zaliczyli egzamin z badań marketingowych. Jednostką doboru próby są: studenci, którzy przystąpili do egzaminu z badań marketingowych.

Krok 2. Zdecyduj, co możesz zrobić z jednostką doboru próby? Określ jednostkę pomiaru. Prowadzący badanie ma do wyboru dwie opcje. Opcja 1: Jednostce pomiaru może przypisać kategorię, np.: płeć, narodowość, wykształcenie. Opcja 2: Jednostce pomiaru należy przypisać atrybuty i je zmierzyć, np.: poziom motywacji, poziom inteligencji.

⁶² S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 127.

⁶³ A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 89.

np. Prowadzący badanie określa jednostkę pomiaru: Jeden student, który przystąpił do egzaminu z badań marketingowych. Następnie odpowiada na pytanie: „Czy jednostce pomiaru można przypisać kategorię?”. Odpowiedź brzmi: Tak. Kategoria 1: studenci, którzy zdali egzamin z badań marketingowych. Kategoria 2: studenci, którzy nie zdali egzaminu z badań marketingowych.

Krok 3. Jeżeli wybrałeś opcję 1, wówczas odpowiedz na pytanie: „Czy kategorię można uporządkować (uszeregować)?”. Jeżeli można, wówczas stosujemy skalę porządkową. Jeżeli nie, stosujemy skalę nominalną.

np. Kategorię można uporządkować ze względu na osiągnięte wyniki w nauce. Skalą pomiaru jest skala porządkowa.

Jeżeli wybrałeś opcję 2, wówczas odpowiedz na pytanie: „Czy jednostka pomiaru posiada punkt zerowy?”. Należy pamiętać, iż jedynie cechy fizyczne i niektóre zachowania posiadają punkt zerowy, np. wzrost, waga, ilość sprzedanych sztuk, natomiast cechy psychologiczne takie jak: iloraz inteligencji, potrzeba samorealizacji nie posiadają punktu zerowego. Jeżeli jednostka pomiaru posiada punkt zerowy, wówczas stosujemy skalę stosunkową. Jeżeli jednostka pomiaru nie posiada punktu zerowego, wówczas stosujemy skalę przedziałową.

Przykład 2. Określ skalę pomiaru dla następujących przykładów:

a) Prowadzący wykład na Politechnice Opolskiej jest zainteresowany zbadaniem zależności, pomiędzy liczbą obecności na wykładzie a osiąganą oceną końcową z egzaminu. Na każdym wykładzie sprawdzana jest obecność według listy wykładowej. Dla każdego studenta z listy notuje się wynik egzaminu końcowego (2.0, 3.0, 3.5, 4.0, 4.5, 5.0). Według jakiej skali prowadzący wykład zmierzy liczbę obecności na wykładzie, a w jakiej wynik egzaminu końcowego?

Liczba obecności na wykładzie: stosujemy skalę stosunkową, gdyż istnieje punkt zerowy. Zero w tym przypadku oznacza fizyczny brak uczestnictwa studenta na wykładach.

„Proszę zakreślić liczbę obecności na wykładach studenta X:

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15.”

Wynik egzaminu końcowego zmierzony zostanie skalą porządkową. Kategorię można uporządkować: ocena 4.0 jest lepsza od oceny 3.0.

„Proszę zakreślić wynik egzaminu końcowego osiągniętego przez studenta X:

2.0, 3.0, 3.5, 4.0, 4.5, 5.0”

b) Kierownik zakładu produkcyjnego chce wiedzieć, jaki wpływ na wydajność pracowników ma temperatura w pomieszczeniu roboczym. Kierownik dokonuje pomiaru temperatury w trzech różnych pomiesz-

zeniach zakładu: pomieszczeniu A, B i C, w których temperatura wynosi odpowiednio 16°C, 18°C oraz 20°C. Wydajność pracowników mierzona jest średnią liczbą wyprodukowanych sztuk wyrobu na dzień. Jaką skalą należy zmierzyć: temperaturę pomieszczenia, a jaką wydajność pracowników?

Temperatura pomieszczenia zostanie zmierzona skalą przedziałową, gdyż dla temperatury nie istnieje punkt zerowy.

„Proszę podać temperaturę panującą w pomieszczeniu A: _____”

Wydajność pracowników zostanie zmierzona skalą stosunkową, ponieważ istnieje punkt zerowy. Zero oznacza, zero sztuk wyrobu wyprodukowanych przez pracowników zakładu w ciągu dnia.

„Proszę podać liczbę wyprodukowanych sztuk produktu przez pracowników w pomieszczeniu A: _____”

Przykład 3. Określ skalę pomiaru dla następujących przykładów:

a) „Płeć:

- 1 - Kobieta
- 2 - Mężczyzna”

W powyższym przykładzie numer „1” jest przypisany do „kobiety”, a numer „2” do „mężczyzny”. Możemy równie dobrze przypisać numer „1” do „mężczyzny”, a numer „2” do „kobiety”. Celem numeru jest jedynie nazwanie cechy lub nadanie jej tożsamości. Jest to przykład skali nominalnej. Zmiana numeracji nie ma wpływu na dane. Liczba mężczyzn i kobiet nadal pozostaje taka sama.

b) „Proszę uszeregować proszki do prania stosownie do stopnia preferencji dla każdego z nich, przypisując pozycję 1 proszkowi najbardziej lubianemu, a pozycję 4 najmniej lubianemu.”

- Vizir
- Persil
- Ariel
- Bonux

Odpowiedź respondenta A:

- Ariel 1
- Persil 2
- Vizir 3
- Bonux 4

Ze względu na uporządkowanie, numery nie są rozpatrywane jako arbitralne, tak jak w przypadku skali nominalnej. Dla respondenta A najbardziej lubianym proszkiem do prania jest proszek Ariel. Zamieńmy teraz numerację przypisaną dla proszków przez respondenta A:

- Ariel 2
- Persil 1
- Vizir 3
- Bonux 4

Zmieniając numery, zmieniamy tym samym znaczenie preferencji. W tym przypadku, najbardziej lubianym proszkiem do prania dla respondenta A jest Persil. Jest to przykład skali porządkowej. Zakłada się, że odległości między punktami skali są nierówne. Tak więc, odległość między rangą „1” a „2” niekoniecznie jest taka sama, jak odległość między rangą „3” a „4”. Jeżeli respondent A lubi proszek Ariel, ale ma równie silną skłonność do proszku Persil ocenionego jako „2”. Uważa, że proszek Vizir jest dobry, ale proszku Bonux w ogóle nie lubi. W takim przypadku odległość między „3” i „4” jest dużo większa, pomimo iż numerycznie odległość jest taka sama. Ta nierówność psychologiczna wygląda następująco:

Persil	Ariel	Vizir	Bonux
1	2	3	4

Przykład 4. Przykład graficznej formy skali⁶⁴.

Proszę ocenić poszczególne cechy w kategoriach znaczenia, które przypisuje Pani/Pan każdej z nich osobiście, umieszczając znak „X” na linii w miejscu, które najlepiej oddaje Pani/Pana odczucie.

Cecha	Bez znaczenia	Bardzo ważna
Uprzejma obsługa	_____	
Dogodne położenie	_____	
Dogodne godziny pracy	_____	
Niskooprocentowany kredyt	_____	

Przykład 5. Przykłady skalowania zrównoważonego i niezrównoważonego.

a) Skalowanie zrównoważone – Jak się Panu/Pani podoba nasze czasopismo?

- Bardzo mi się podoba.
- Raczej mi się podoba.
- Ani mi się podoba, ani mi się nie podoba.
- Raczej mi się nie podoba.
- Bardzo mi się nie podoba.

⁶⁴ A. Churchill, Badania marketingowe. Podstawy metodologiczne, PWN, Warszawa 2002, s. 427.

Jak widzimy, odpowiedzi są jednakowo ułożone po obu stronach skali środkowej „ani mi się podoba, ani mi się nie podoba”, która odzwierciedla brak opinii w danej sprawie. Respondent może wybrać odpowiedź całkowicie negatywną lub całkowicie pozytywną, a także raczej negatywną lub raczej pozytywną.

b) Skalowanie zrównoważone – Jaka część surowców potrzebnych do produkcji w Waszej firmie jest importowana?

- 0-20%
- 21-40%
- 41-60%
- 61-80%
- 81-100%

c) Skalowanie niezrównoważone – Jak się Panu/Pani podoba nasze czasopismo?

- Bardzo mi się podoba.
- Raczej mi się podoba.
- Troszkę mi się podoba.
- Troszkę mi się nie podoba.
- Bardzo mi się nie podoba.

Przedstawione odpowiedzi wykazują pewną tendencję bądź pozytywną, bądź negatywną. W niektórych przykładach respondent ma do wyboru trzy odpowiedzi pozytywne oraz dwie negatywne.

d) Skalowanie niezrównoważone – Ile razy w tym roku podróżowała Pan/Pani pociągiem TLK po kraju?

- Ani razu.
- 1 raz.
- 2-4 razy.
- 5-9 razy.
- 10-19 razy.
- 20 i więcej razy.

e) Skalowanie niezrównoważone – Proszę przypomnieć sobie ostatnią podróż pociągiem. Jak ocenia Pan/Pani przewoźnika kolejowego, z którego usług Pan/Pani korzystała?

- Znakomite.
- Bardzo pozytywnie.
- Pozytywnie.
- Obojętnie.
- Negatywnie.

Przykład 6. Poniżej przedstawiono przykład skalowania niewymuszającego. Jak ocenia Pani/Pan firmę XYZ w porównaniu z konkurentami na rynku polskim? Proszę ująć w kółka te liczby, które najlepiej wyrażają Pani/Pana opinię:

Lepiej niż konkurenci		Nie różnią się od konkurentów		Gorzej niż konkurenci		Nie wiem
6	5	4	3	2	1	0

Przykład 7. Przykłady skali parzystych i nieparzystych.

a) Czteropunktowa skala ważności:

- Niezwykle ważne.
- Dość ważne.
- Niezbyt ważne.
- Zupełnie nieważne.

b) Pięciopunktowa skala ważności:

- Niezwykle ważne.
- Dość ważne.
- Ani ważne, ani nieważne.
- Niezbyt ważne.
- Zupełnie nieważne.

c) Czteropunktowa skala postaw:

- Bardzo mi się podoba
- Raczej mi się podoba
- Raczej mi się nie podoba
- Bardzo mi się nie podoba

d) Pięciopunktowa skala postaw:

- Bardzo mi się podoba
- Raczej mi się podoba
- Ani mi się podoba, ani mi się nie podoba
- Raczej mi się nie podoba
- Bardzo mi się nie podoba

e) Czteropunktowa skala zadowolenia:

- Bardzo zadowolony
- Raczej zadowolony
- Raczej niezadowolony
- Bardzo niezadowolony

f) Pięciopunktowa skala zadowolenia:

- Bardzo zadowolony
- Raczej zadowolony
- Ani zadowolony, ani nie zadowolony
- Raczej niezadowolony
- Bardzo niezadowolony

g) Czteropunktowa skala skłonności do zakupu:

- Zdecydowanie bym kupił
- Raczej bym kupił
- Raczej bym nie kupił
- Zdecydowanie bym nie kupił

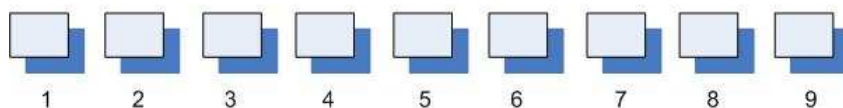
h) Pięciopunktowa skala skłonności do zakupu:

- Zdecydowanie bym kupił
- Raczej bym kupił
- Może bym kupił, może nie
- Raczej bym nie kupił
- Zdecydowanie bym nie kupił

Przykład 8. Przykład zastosowania skali rang⁶⁵

Oto lista cech samochodu, do których może Pan(i) przywiązać większą lub mniejszą wagę. Proszę wpisać kolejno do kratek skali od 1 do 9 odpowiadające poszczególnym cechom symbole literowe tak, aby cecha, która dla Pana(-i) ma największą wagę, znalazła się w pierwszej kratce, druga najważniejsza cecha- w drugiej itd.

A- producent	D- łatwość obsługi	G- szybkość
B- cena	E- zużycie paliwa	H- przyspieszenie
C- kolor	F- moc silnika	I- pojemność bagażnika



⁶⁵ S. Kaczmarczyk, Badania marketingowe. Metody i techniki, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 121.

Przykład 9. Przykład zastosowania skali sumowanych ocen

Proszę podzielić 10 punktów między każdą z trzech podanych par różnych proszków do prania zgodnie z Pani/Pana preferencjami.

- A. Vizir (.....) + Persil (.....) = 10
 B. Rex (.....) + Persil (.....) = 10
 C. Rex (.....) + Vizir (.....) = 10

Przykład 8. Przykład zastosowania skali dyferencjału semantycznego

Proszę dokonać oceny firmy XYZ.

Obsługa jest nieuprzejma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Obsługa jest uprzejma
Położenie jest dogodne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Położenie jest niedogodne
Godziny pracy są niedogodne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Godziny pracy są dogodne
Oprocentowanie kredytów jest wysokie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Oprocentowanie kredytów jest niskie

Obsługa jest nieuprzejma	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Obsługa jest uprzejma
Położenie jest niedogodne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Położenie jest dogodne
Godziny pracy są niedogodne	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Godziny pracy są dogodne
Oprocentowanie kredytów jest wysokie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Oprocentowanie kredytów jest niskie

Firma XYZ _____

Firma ABC — — — — —

Przykład 9. Przykład zastosowania skali Likerta⁶⁶

	Zdecydowanie nie	nie	Nie mam zdania	tak	Zdecydowanie tak
1. Każdy Polak ma moralny obowiązek kupować polskie produkty bez względu na ich cenę.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Jeżeli produkt polski i zagraniczny mają te same cechy, powinno się kupować towar polski.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Popieranie produktów polskich przyczynia się do wzrostu dobrobytu wszystkich obywateli.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Polskie produkty są równie dobre jak zagraniczne.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

⁶⁶ A. Sagan, Badania marketingowe. Podstawowe kierunki, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 89.

3.5. Zadania do samodzielnego opracowania

Zadanie 1. Określ typ skali (nominalna, porządkowa, przedziałowa, stosunkowa).

- a) Jakie są twoje trzy ulubione marki dżinsów? Uszereguj je od 1 do 3 stosownie do preferencji, oznaczając markę nr 1 najbardziej preferowaną.
- Levis
 - Orsay
 - New Yorker
 - Big Star
 - H&M
 - Bershka
- b) W jakim miesiącu się urodziłeś?
- | | |
|-----------------------------------|--------------------------------------|
| <input type="checkbox"/> styczeń | <input type="checkbox"/> lipiec |
| <input type="checkbox"/> luty | <input type="checkbox"/> sierpień |
| <input type="checkbox"/> marzec | <input type="checkbox"/> wrzesień |
| <input type="checkbox"/> kwiecień | <input type="checkbox"/> październik |
| <input type="checkbox"/> maj | <input type="checkbox"/> listopad |
| <input type="checkbox"/> czerwiec | <input type="checkbox"/> grudzień |
- c) Proszę rozdzielić 100 punktów pomiędzy następujące batoniki stosownie do stopnia preferencji.
- | | |
|-----------------------------------|-------|
| <input type="checkbox"/> MilkyWay | |
| <input type="checkbox"/> Snickers | |
| <input type="checkbox"/> Mars | |
| <input type="checkbox"/> Twix | |
| <input type="checkbox"/> Grzeński | |
| <input type="checkbox"/> 3Bit | |
| <input type="checkbox"/> Bounty | |
- d) Ile czasu spędzasz codziennie w drodze do pracy?
- mniej niż 5 min
 - 5-10 min
 - 11-15 min
 - 16-20 min
 - 21-30 min
 - 31 min i więcej

- e) W jakim stopniu jesteś zadowolony z magazynu „Nature”?
- bardzo zadowolony
 - zadowolony
 - ani zadowolony, ani nie zadowolony
 - niezadowolony
 - bardzo niezadowolony
- f) Jaki jest dochód twojego gospodarstwa domowego (w zł)?
- g) Czy słyszałeś o walentynkach?
- Tak
 - Nie
- h) Ile planujesz wydać na prezent walentynkowy?
- do 20 zł
 - 21-50 zł
 - 51-100 zł
 - 101-200 zł
 - powyżej 200 zł
 - trudno powiedzieć
- i) Wyraż swój stosunek do stwierdzenia „Agent marketingowy przejmuje funkcje działu marketingowego od producentów, chociaż główną jego funkcją jest sprzedaż produktu.
- całkowicie się zgadzam
 - zgadzam się
 - ani tak, ani nie
 - nie zgadzam się
 - całkowicie się nie zgadzam
- j) Ile paczek papierosów wypalasz w ciągu jednego dnia?
- mniej niż ½ paczki
 - ½ do 1 paczki
 - więcej niż 1 paczkę

Zadanie 2. Dokonaj charakterystyki poniższych skal (zrównoważona, niezrównoważona, parzysta, nieparzysta, graficzna, wymuszająca, niewymuszająca, sumowanych ocen, pozycyjna, dyferencjału semantycznego, Likerta).

- a) Jak ocenia Pan/Pani szybkość obsługi w naszym salonie optycznym?
- bardzo szybka
 - szybka
 - ani szybka, ani wolna
 - wolna
 - bardzo wolna

- b) Jak zareagowała Pani na ukazanie się nowego czasopisma „Woman” na rynku polskim?
- entuzjastycznie
 - bardzo pozytywnie
 - pozytywnie
 - obojętnie
 - negatywnie

- c) Proszę zaznaczyć swoje oceny najlepiej odzwierciedlające Pana/Pani opinię o wymienionych cechach banku X. Każda skala reprezentuje jedną cechę o siedmiostopniowej intensywności.

pewny	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	niepewny
sztywny	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	elastyczny
szybki	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	opieszaly
drogi	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	tani

- d) Proszę ocenić w jakim stopniu zgadza się Pan/Pani z poniższymi stwierdzeniami, poprzez zaznaczenie właściwego pola.

	Zdecydowanie nie	nie	tak	Zdecydowanie tak
1. Uważam, że dobrze wykonuję swoje obowiązki w pracy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Uważam, że mam dobre stosunki ze współpracownikami.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Jestem dumny z moich umiejętności radzenia sobie z trudnymi zadaniami.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Mogę powiedzieć, że zespół pracowników jest zadowolony, że z nimi pracuję.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- e) Proszę ocenić cechy czasopisma “Woman” w zależności od stopnia, w jakim spełnia te określenia. Jeżeli ocenia Pan/Pani pozytywnie, proszę zaznaczyć na skali od +1 do +3, jeżeli negatywnie od -1 do -3.

	-3	-2	-1	+1	+2	+3
Jakość edycji	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dobór tematyki	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Okładka	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Łatwość przyswajania treści	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

4. KWESTIONARIUSZ – PODSTAWOWY INSTRUMENT POMIARU

4.1. Wprowadzenie

Kwestionariusz jest podstawowym instrumentem pomiaru w badaniach marketingowych. Wykorzystywany jest w wywiadach oraz badaniach ankietowych. Stanisław Kaczmarczyk definiuje kwestionariusz w następujący sposób: „Kwestionariusz jest to zbiór pytań zapisanych w określony sposób na kartkach papieru lub w innej formie w celu wywołania żądanych odpowiedzi”⁶⁷. Kwestionariusz wykorzystywany w badaniach ankietowych, nazywamy kwestionariuszem ankietowym. Wykorzystanie kwestionariusza ankietowego polega na wręczeniu respondentowi, który sam odczytuje pytania i sam odnotowuje odpowiedzi w kwestionariuszu. Kwestionariusz wykorzystywany w wywiadach, nazywamy kwestionariuszem wywiadu. Kwestionariusz wywiadu tym różni się od kwestionariusza ankietowego, że pytania w nim zawarte, są odczytywane przez ankietera.

Podstawową funkcją kwestionariusza jest dokonanie pomiaru. Celem pomiaru jest zmierzenie lub zidentyfikowanie cech fizycznych, psychicznych, demograficznych, społecznych i ekonomicznych respondenta. W badaniach marketingowych kwestionariusz często jest wykorzystywany do badania postaw konsumentów względem określonych produktów lub marek.

4.2. Podstawowe zasady konstrukcji kwestionariusza

W literaturze⁶⁸ wymienia się podstawowe zasady, zgodnie z którymi prowadzący badanie powinien konstruować kwestionariusz. Przy projektowaniu kwestionariusza prowadzący badanie musi zastanowić się nad sensem zadawanych pytań. Pomocną czynnością jest opracowanie dokumentu zawierającego zdefiniowane problemy badawcze, do których prowadzący badanie może wracać na każdym etapie konstrukcji kwestionariusza. Podstawowym celem kwestionariusza jest dokonanie pomiaru. Prowadzący badanie musi podjąć cztery podstawowe decyzje

⁶⁷ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 97.

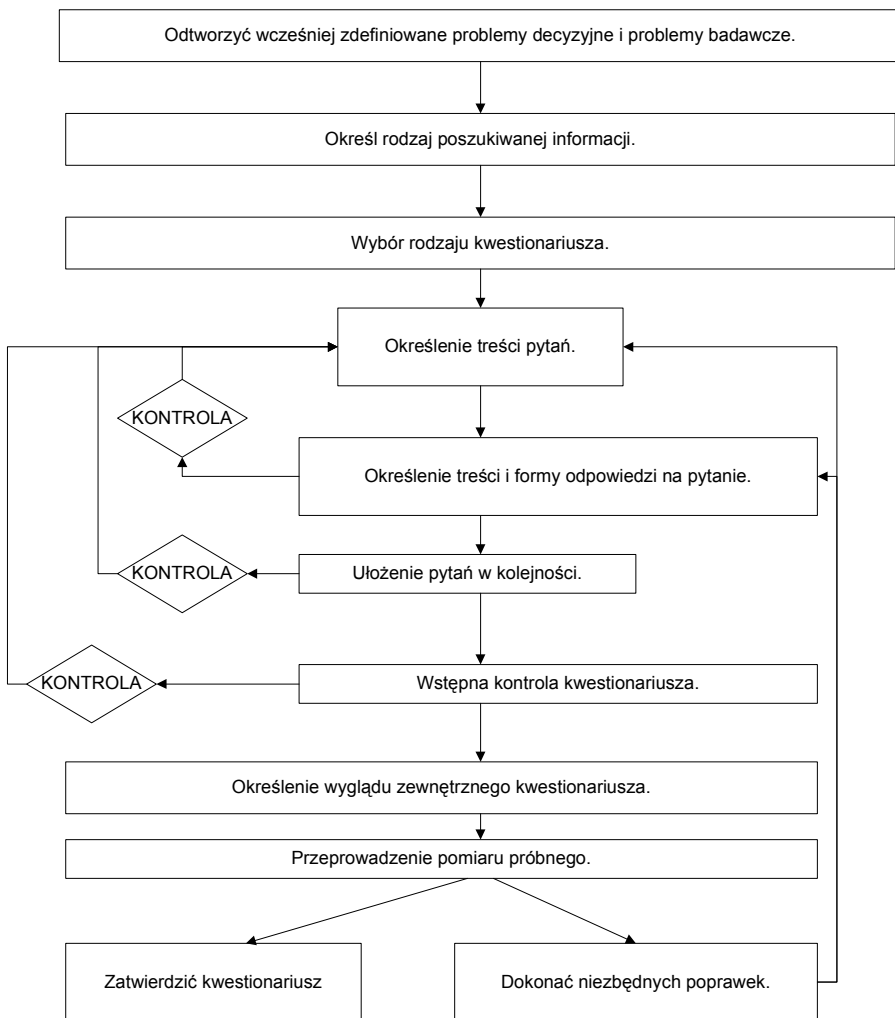
⁶⁸ Zobacz S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003; S. Kaczmarczyk, *Zastosowanie badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007; R.J. Kaden, *Badanie marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2008.

dotyczące: rodzaju poszukiwanych informacji, treści pytania, formy odpowiedzi oraz stylu i formy gramatycznej pytania⁶⁹.

Kwestionariusze ankietowe i wywiadu powinny na początku zawierać wprowadzenie. Wprowadzenie w kwestionariuszu upewnia respondenta, co do celu badania i konieczności jego przeprowadzenia. Kwestionariusz powinien zawierać także takie informacje jak: dane ankietera, datę przeprowadzenia wywiadu. Istotnym faktem, z punktu widzenia funkcji kwestionariusza, jest jego wygląd zewnętrzny. Kwestionariusz powinien być estetyczny, uporządkowany, wydrukowany na białym papierze. Pytania zawarte w kwestionariuszu powinny płynnie przechodzić z jednego do drugiego, tworzyć logiczną całość i w miarę możliwości być pogrupowane tematycznie.

Poprawnie sformułowane pytania wynikają bezpośrednio z przygotowanych problemów badawczych. Pytania powinny oddawać charakter poszukiwanej informacji. Dane, które zostaną pozyskane po zadaniu pytania, powinny mieć istotne znaczenie przy ich interpretacji oraz formułowaniu wniosków w końcowym etapie procesu badania marketingowego. Pytania w kwestionariuszu powinny być ułożone w odpowiedniej kolejności. Prowadzący badanie dostosowuje treść pytań i ich formę do rodzaju danych, które mają być pozyskane. Podczas układania pytań w kwestionariuszu badacz powinien zastanowić się nad możliwymi wariantami odpowiedzi (prowadzący badanie powinien mieć na uwadze system kodowania zebranych danych). W badaniach dąży się do jak najprostszego i czytelnego kodowania danych zebranych w trakcie pomiaru. Ostatnią wskazówką istotną przy niektórych formach przeprowadzania wywiadu/ankiety, jest opracowanie szczegółowych wskazówek dla osoby przeprowadzającej badanie, co do formy i sposobu przeprowadzania badania. Ostatnim etapem procesu budowy kwestionariusza jest przeprowadzenie pomiaru próbnego. Przykładowy proces formułowania kwestionariusza został przedstawiony na rysunku 10.

⁶⁹ A. Sagan, *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 41.



Rys. 9. Etapy budowy kwestionariusza

Źródło: opracowanie własne.

4.3. Formułowanie pytań i odpowiedzi

Pytania w kwestionariuszu spełniają dwie funkcje: gwarancyjną oraz optymalizacyjną. Funkcja gwarancyjna jest realizowana poprzez gwarancję takiego sformułowania pytania, które będzie dostarczać potrzebnych danych. Funkcja optymalizacyjna polega na takim sformułowaniu pytania, które zapewni optymalne warunki dla respondenta do udzielenia odpowiedzi. Przez optymalne warunki rozumiemy warunki, w których możliwa jest swobodna i wyczerpująca wypowiedź.

Pierwszą kwestią, jaką powinien rozważyć prowadzący badanie przy formułowaniu treści pytań, jest niezbędność pytania. Drugą kwestią, która powinna zostać rozwiązana przez badacza, jest liczba pytań, która wyczerpuje dany temat.

Badacz musi dokonać oceny, czy przeciętny respondent będzie dysponował informacją, której oczekujemy. W takich sytuacjach należy zastosować pytania filtrujące. Celem pytań filtrujących jest określenie zdolności respondenta do udzielenia informacji w danym temacie.

Respondenci nie zawsze są chętni do dzielenia się informacją z prowadzącym badanie. Fakt, iż respondent posiada informację, nie oznacza, iż będzie on chciał wziąć udział w badaniu. Gotowość do udzielenia informacji zależy od wysiłku, jaki jest związany z udzieleniem odpowiedzi, drażliwości pytania oraz jego sformułowania⁷⁰. Pytania drażliwe dotyczą danych personalnych osoby badanej, są to pytania wprawiające respondenta w zakłopotanie lub wywołujące w nim poczucie zagrożenia. Pytania drażliwe dotyczą sfery seksualnej, spożywania napojów alkoholowych, uprzedzeń rasowych, religii, spożywania niedozwolonych używek. W przypadku konieczności ich zadania, prowadzący badanie powinien odpowiednio je sformułować i umieścić na odpowiedniej pozycji w kwestionariuszu. Zasady postępowania z pytaniami drażliwymi są następujące: przejście od pytań ogólnych do pytań drażliwych nie może być nagłe i zaskakujące respondenta, pytania drażliwe powinny być umieszczane na końcu kwestionariusza, zastosowanie pytań stwarzających poczucie bezpieczeństwa u respondenta (zastosowanie długich pytań wprowadzających oraz wydłużenie skali odpowiedzi np. „Czy sądzi Pani/Pan, że większość osób nie płaci abonamentu telewizyjno-radiowego? Dlaczego?”), powołanie się na powszechność występowania danego zjawiska (respondent jest bardziej skłonny do podania szczerzej odpowiedzi, gdy twierdzenie drażliwe dotyczy większości społeczeństwa, a nie tylko jego samego np. „Najnowsze badania wskazują, że co trzeci Polak nie spłaca zadłużenia kredytowego”), sformułowanie odpowiedzi według kategorii klas (np. w przypadku pytania „Jaki jest Pana/Pani dochód miesięczny brutto” stosuje się przedziały klasowe, np. 0-2000 zł, 2001-4000 zł, 4001 i więcej zł).

Pytania w kwestionariuszu powinny zostać sformułowane w odniesieniu do typu respondentów, do których badanie jest skierowane. Dobór słownictwa w pytaniach zależy przede wszystkim od poziomu wykształcenia osoby badanej. Jeżeli w pytaniu muszą wystąpić specjalistyczne sformułowania, prowadzący badanie powinien je wyjaśnić respondentowi. Zdania powinny być zwięzłe, precyzyjne, jasne i jednoznaczne. Nie powinny

⁷⁰ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 361.

pozostawiać wątpliwości, co do ich interpretacji. Badacz musi się wystrzegać pytań sugestywnych.

Do ważniejszych reguł stosowanych przy formułowaniu pytań należą:

- Używaj prostych słów – jeżeli prowadzący badanie ma do wyboru użycie trudnego słowa, bądź prostego słowa, powinien wybrać słowa rozumiane przez wielu respondentów.
- Unikaj wieloznacznych sformułowań – słownictwo użyte w kwestionariuszu powinno charakteryzować się prostotą i jednoznacznością. Do słów wieloznacznych zaliczamy słowa: „wszystko” (np. „Czy to prawda, że wszystkie kobiety nie potrafią prowadzić samochodu?”), „około” (np. „około 50% badanych twierdzi, że ...”, słowo „około” oznacza „gdzieś w pobliżu” około 50% może oznaczać zarówno 41% jak i 59%), „zawsze”, „czasami”, „często”, „okazjonalnie”, „nigdy”.
- Unikaj pytań sugestywnych.
- Unikaj ukrytych założeń – w tym przypadku w zależności od posiadanej przez respondenta wiedzy w danym temacie, może on wiązać konsekwencje z udzieloną odpowiedzią. Na pytanie „Czy jest Pan/Pani za umocnieniem się złotego w stosunku do euro?” można uzyskać różną odpowiedź, w zależności od poziomu wykształcenia respondenta i podstawowej znajomości ekonomii.
- Unikanie szacunków i uogólnień – pytanie „Z iloma kontrahentami zawarł Pan/Pani kontrakty w zeszłym roku?” wymaga od respondenta dokonania skomplikowanych obliczeń. W takim przypadku prowadzący badanie może zadać pytanie dotyczące krótszego okresu czasu, np. typowego kwartału (przy założeniu *ceteris paribus*) i pomnożyć wynik przez cztery.
- Unikanie pytań „dubeltowych” – pytanie dubeltowe, to pytanie, które wymaga od respondenta dwóch odpowiedzi. „Jak Pan/Pani ocenia kolor i kształt tej kanapy?” Odpowiadając na to pytanie respondent musi odnieść swoją odpowiedź osobno do koloru i kształtu⁷¹.

Kolejność pytań w kwestionariuszu jest ważnym elementem ułatwiającym pomiar. Osoba badana powinna bez trudu przechodzić od jednego pytania do drugiego. Pytania następujące po sobie muszą charakteryzować się ciągiem logicznym. Podejmując decyzję o kolejności pytań w kwestionariuszu, prowadzący badanie powinien kierować się następującymi zasadami:

- Zasada pierwszeństwa pytań ogólnych nad pytaniami szczegółowymi. Zgodnie z tą zasadą prowadzący badanie stosuje na początku kwestionariusza pytania szerokie, a następnie zawęża ich zakres.

⁷¹ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 372-379.

- Zasada wywołania zaciekawienia. Na samym początku kwestionariusza umieszczamy pytania ciekawe, otwierające kwestionariusz. Największe znaczenie ma pierwsze pytanie. Jeżeli pierwsze pytanie kwestionariusza jest na tyle ciekawe, iż wzbudzi zainteresowanie respondenta, wówczas istnieje duże prawdopodobieństwo przeprowadzenia wywiadu w całości. W przeciwnym razie respondent może odmówić udziału w badaniu.
- Zasada staranności projektowania pytań rozgałęziających. Pytania rozgałęziające, są to pytania, kierujące respondenta do innego miejsca w kwestionariuszu. Pytania te powinny być tak zaprojektowane, aby respondent swobodnie poruszał się po kwestionariuszu.
- Zasada umieszczania pytań drażliwych na końcu kwestionariusza.

Odpowiedzi na zadane pytanie są ważnym etapem procesu projektowanie kwestionariusza. W momencie, w którym prowadzący badanie określił już treść pytania, musi dokonać wyboru, jaką formę ma przybrać odpowiedź na nie. Czy pytanie ma nie dawać wariantów odpowiedzi? Czy przedstawić respondentowi warianty odpowiedzi? Jaką zastosować skalę odpowiedzi?

Pytania otwarte polegają na tym, że respondentom zapewnia się swobodę wypowiedzi. Respondenci nie są ograniczeni wytypowanymi przez badacza postawami oraz mają możliwość wyrażenia postawy swoimi słowami.

Pytania zamknięte charakteryzują się tym, że odpowiedzi na nie są wyskalowane. Respondenci są proszeni o zaznaczenie wariantu, który jest najbliższy ich postawie. Skalowanie polega na użyciu w odpowiedzi jednej z czterech podstawowych skal: nominalnej, porządkowej, przedziałowej i stosunkowej. Pytania zamknięte są łatwiejsze do stosowania w porównaniu do pytań otwartych. Ich analiza nie stwarza większych problemów. W wyniku zastosowania skali klasyfikacja, kodowanie, redukcja i analiza danych są ułatwione. Istotną wadą pytań zamkniętych jest możliwość popełnienia błędów przy formułowaniu wariantów odpowiedzi. Liczba wariantów odpowiedzi może być zbyt duża lub zbyt mała dla respondenta. W wyniku tego, dane dostarczone przez respondenta nie będą rzetelne.

Pytania dychotomiczne są to pytania o dwóch wariantach odpowiedzi. Respondent ma do wyboru tylko jedną z nich.

4.4. Przykłady z omówieniem

Przykład 1. Oto przykłady pytań dychotomicznych:

- a) Czy w przyszłym roku zamierza Pani pojechać na wczasy?
 Tak Nie

b) Jakie jest Pana/Pani, stanowisko w sprawie podatków od posiadania kota?

- Jestem za Jestem przeciw

Pytania dychotomiczne możemy przekształcić w pytania o wielu wariantach odpowiedzi. I tak:

Ad. a) Czy w przyszłym roku zamierza Pani pojechać na wczasy?

- zdecydowanie tak,
 raczej tak,
 ani tak, ani nie,
 raczej nie,
 zdecydowanie nie.

Ad. b) Jakie jest Pana/Pani, stanowisko w sprawie podatków od posiadania kota?

- zdecydowanie popieram,
 raczej popieram,
 ani popieram, ani nie popieram,
 raczej nie popieram,
 zdecydowanie nie popieram.

Przykład 2. Dokonaj oceny poniższych pytań i odpowiedzi:

1 Czy jeździsz na rowerze?

- często,
 czasami,
 rzadko,
 nigdy.

2 Czy uważasz, że wobec kiboli należy zaostrzyć wymiar kary oraz poddać ich przymusowym pracom na rzecz społeczeństwa?

3 Ile książek przeczytałeś w zeszłym roku?

- 0
 1 -5
 6-10

4 Co myślisz o niesłusznych atakach opozycji na działania rządu?

- zdecydowanie je popieram,
 popieram,
 ani popieram, ani nie popieram,
 zdecydowanie ich nie popieram.

Ad 1. Problem w tym pytaniu to kategorie odpowiedzi. Dla różnych respondentów odpowiedź „często”, „czasami”, „rzadko” może mieć różne znaczenie. Należy użyć kategorii odpowiedzi ze zdefiniowaną częstością, np.: „raz w tygodniu”, „raz na miesiąc”. Odpowiedzi powinny być ułożone w ciągu logicznym: od najmniejszej do największej częstości.

Ad 2. Jest to pytanie „dubeltowe”. W tym pytaniu zawarte są dwa odrębne pytania: o zaostrzenie wymiaru kary oraz o przymusowe prace na rzecz społeczeństwa. Ponadto w pytaniu użyto sformułowania „kiboli”, które nasuwa negatywne skojarzenia z kibicami piłki nożnej, i jest sformułowaniem w języku potocznym.

Ad 3. Pytanie to zawiera ukrytą tendencyjność. Respondent może chcieć wyrzucić lepsze wrażenie zaznaczając odpowiedź z wyższą liczbą przeczytanych książek. Kategorie odpowiedzi są źle sformułowane. Nie ma kategorii dla osób, które przeczytały więcej niż 10 książek. Należy więc dodać kategorię „więcej, ile? ____ „

Ad 4. Pytanie zawiera sformułowanie o zabarwieniu emocjonalnym. używając sformułowania „niesłuszne ataki” zakłada się, że wszystkie działania rządu są słuszne. Zakłada się także, iż opozycja krytykując działania rządu, „atakuję” go. Pytanie zakłada także, pewną wiedzę respondenta w temacie krytyki opozycji na rząd. Założenie to jest niesłuszne, gdyż nie uwzględnia respondentów, którzy nie interesują się polityką i nie posiadają wiedzy w danym temacie. Kategorie odpowiedzi powinny zawierać skalę pięciostopniową: „Zdecydowanie popieram”, „Popieram”, „Ani popieram, ani nie popieram”, „Nie popieram”, „Zdecydowanie nie popieram”.

4.5. Zadania do samodzielnego opracowania

Zadanie 1. Oceń poprawność następujących pytań i wariantów odpowiedzi. Zaproponuj ewentualne korekty.

Który z następujących programów oglądasz regularnie?

- „Jak oni śpiewają?”
- „Taniec z gwiazdami.”
- „Idol”

a) Czy często kupuje Pan/Pani karmę dla kotów „Mruczek”?

- tak
- nie

- b) Czy zgadza się Pan/Pani ze stwierdzeniem, że rząd powinien dofinansowywać polskie stocznie?
- całkowicie się zgadzam
 - zgadzam się
 - ani się zgadzam, ani nie zgadzam
 - nie zgadzam się
 - całkowicie się nie zgadzam
- c) Jak często kupuje Pan/Pani płyn do mycia naczyń?
- raz w tygodniu
 - raz na dwa tygodnie
 - raz na trzy tygodnie
 - raz na dwa miesiące
- d) Proszę uszeregować następujące marki według preferencji:
- Vizir
 - Rex
 - Persil
 - Lovela
- e) Gdzie Pan/Pani zazwyczaj kupuje materiały biurowe?
- f) Jak Pan/Pani ocenia polskiego premiera i rząd?
- g) Czy sądzi Pan/Pani, że należy realizować politykę ograniczania świadczeń socjalnych?
- tak
 - nie
- h) Jak często w ciągu siedmiu dni tygodnia, spożywa Pan/Pani kolację?
- codziennie
 - 4-5 razy w tygodniu
 - 2-4 razy w tygodniu
 - raz w tygodniu
 - nigdy

5. GROMADZENIE DANYCH WTÓRNYCH

5.1. Typy danych wtórnych

Dane wtórne są to wyniki przeprowadzonych badań, które zostały już opublikowane. Badania prowadzone na danych wtórnych, określa się mianem *desk research*, czyli tzw. „badań zza biurka”. Dane wtórne możemy podzielić na wewnętrzne oraz zewnętrzne (por. tabela 5).

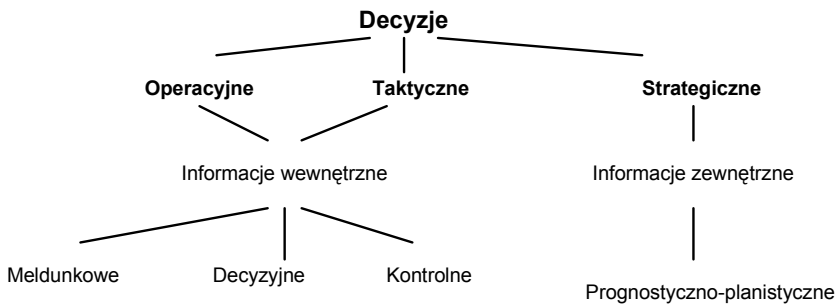
Tabela 5.

Podział danych wtórnych

Dane wtórne	wewnętrzne		faktura sprzedaży, raporty agentów handlowych, rozliczenia kosztów handlowych
	zewnętrzne	publikowane	periodyki, źródła statystyczne, dokumentacja finansowa, skorowidze
		komercyjne	dane geodemograficzne, dane z badań panelowych, dane z inwentaryzacji sklepowej, dane z rejestracji skaningowej

Źródło: G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 259.

Dane wtórne możemy podzielić ze względu na kryteria rodzajów decyzji, dla których są one przygotowywane oraz miejsca ich przetwarzania. Rysunek 5 przedstawia podział informacji ze względu na kryterium rodzajów decyzji. Informacje wewnętrzne są pozyskiwane w celu podjęcia decyzji operacyjnych i taktycznych. Natomiast pozyskiwanie informacji zewnętrznych jest uzasadnione w przypadku, gdy badanie podejmowane jest ze względu na konieczność podjęcia decyzji strategicznej.



Rys. 10. Podział informacji ze względu na kryterium rodzajów decyzji

Źródło: opracowanie własne na podstawie *(Badania marketingowe. Podstawowe metody i obszary zastosowań,* pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 65.

Podział informacji wewnętrznych ze względu na miejsce ich powstawania w organizacji, dotyczy przede wszystkim działu marketingu, sprzedaży, finansów i logistyki. Jednak w procesie badania marketingowego dwie jednostki organizacyjne przejmują na siebie ciężar dostarczenia informacji prowadzącemu badanie. Jest to dział marketingu oraz dział sprzedaży. Dział sprzedaży posiada w swojej dyspozycji dane dotyczące:

- wielkości sprzedaży, ze względu na: obszar geograficzny, rodzaj rynku, rodzaj produktu, agenta handlowego, klientów,
- liczby zamówień zrealizowanych, oczekujących na wysyłkę,
- branży klienta,
- kanałów dystrybucji.

Tabela 6.

Wybrane użyteczne źródła danych wtórnych wewnętrznych

Komórki przedsiębiorstw	Przykłady źródeł wtórnych
Zarząd	roczny raport z działalności przedsiębiorstwa, roczne sprawozdanie finansowe, sprawozdania z posiedzeń zarządu, zgromadzenia wspólników, zgromadzenia rady nadzorczej, strategie i plany działania przedsiębiorstw, biznesplany, programy rozwoju, plany inwestycyjne
Sprzedaż i marketing	zestawienia na temat wielkości i dynamiki sprzedaży poszczególnych produktów, grup asortymentowych, strategie i plany marketingowe, sprawozdania i protokoły z kontroli placówek handlowych własnych i współpracujących, punktów dealerskich, hurtowni, placówek serwisowych, sprawozdanie z negocjacji z klientem, zapytania ofertowe, korespondencja z nabywcami, dokumentacja kontaktów posprzedażowych, reklamacje, raporty z przeprowadzonych wcześniej badań i analiz rynku
Księgowość i finanse	bilans, rachunek wyników, księgi i kartoteki inwentarzowe majątku przedsiębiorstwa, oświadczenia i rozliczenia podatkowe, faktury sprzedaży, analiza finansowa przedsiębiorstwa, analiza kosztów produkcji i zysku.
Pozostałe komórki organizacyjne (zaopatrzenie, produkcja, transport, zaplecze badawczo-rozwojowe, dział personalny)	raporty o wielkości i strukturze produkcji, plany zaopatrzenia, produkcji, plany wprowadzania nowego produktu, innowacje, listy teleadresowe dostawców, sprawozdania z negocjacji z dostawcami, opracowania systemów kontroli jakości, stan zatrudnienia i charakterystyka pracowników

Źródło: A. Kaniewska-Sęba, G. Leszczyński, B. Pilarczyk, *Badania marketingowe na rynku business-to-business*, Oficyna Ekonomiczna, Kraków 2006, s. 60-61.

Prowadzący badanie na podstawie danych wtórnych wewnętrznych organizacji, dla której prowadzone jest badanie, jest w stanie:

- określić udział poszczególnych klientów w zysku,
- określić udział poszczególnych rynków w zysku,
- określić udział w danym rynku i zakupach poszczególnych klientów,
- dokonać pomiaru zdolności produkcyjnych,
- dokonać pomiaru kosztów dostawy produktów,
- dokonać analizy ofert i zrealizowanych zamówień,
- określić obecny i przyszły kształt popytu⁷².

Prowadzący badanie, analizując dane wtórne zewnętrzne, może poczuć się oszołomiony ich ilością i różnorodnością. Przebicie się przez ten tzw. „gąszcz informacji”, wymaga od badacza wiele cierpliwości, a przede wszystkim ustalonego planu działania. Churchill proponuje zestaw siedmiu kroków, mających pomóc w przystąpieniu do analizy danych wtórnych zewnętrznych⁷³. Krok pierwszy nakazuje ustalenie informacji, które należy zdobyć. Powinny one zostać określone najszerzej, obejmując zarówno publikacje na dany temat, aż po dane osobowe. Krok drugi polega na opracowaniu najważniejszych nazwisk i terminów, które będą wykorzystywane w dalszym etapie poszukiwań. Taka lista ułatwia poszukiwania i oszczędza czas prowadzącego badanie na dotarcie do tych faktów w późniejszym okresie. Krok trzeci wiąże się ściśle z poszukiwaniami referatów i sprawozdań. W tym celu należy się udać do biblioteki. Poszukiwania opublikowanych danych powinny być prowadzone w oparciu o wcześniej sporządzoną listę terminów z terminami pojęć zawartymi w słowniczku danej publikacji. Krok następny opiera się na analizie zebranych źródeł danych w oparciu o ich przydatność w stosunku do potrzeb. Prowadzący badanie w zależności od uzyskanych wyników może zagubić się w analizowanej literaturze, bądź wręcz przeciwnie, szybko uporać się z wytypowanymi źródłami. Nieocenioną pomocą są programy komputerowe do katalogowania i przeszukiwania danych. Krok szósty odsyła badacza do zewnętrznych źródeł specjalistycznych. Mogą to być prace ogólne, informatory, bazy danych, czasopisma i periodyki. Jeśli poszukiwania wśród publikacji specjalistycznych nie przyniosły zakładanego rezultatu, wówczas wskazane jest zasięgnięcie opinii u podmiotu, który ma jakąś wiedzę na ten temat. Może być to autorytet naukowy w danej dziedzinie, organizacja, osoba.

⁷² Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 68.

⁷³ G.A. Churchill, Badania marketingowe. Podstawy metodologiczne, PWN, Warszawa 2002, s. 263.

Liczba zewnętrznych źródeł wtórnych nieustannie wzrasta. Wybrane źródła informacji biznesowej na rynku polskim prezentuje tabela 7.

Tabela 7.

Wybrane źródła informacji biznesowej na rynku polskim

Źródła publikowane	
Książki	Stanowią nie tylko źródło nowej wiedzy, ale i przedstawiają sposób działania konkurentów (zwłaszcza, gdy autor był pracownikiem takiej firmy).
Książki adresowe	Źródło numer jeden na początku projektu – informuje, kto jest w branży, na kim skoncentrować działania.
Prasa ekonomiczna	Najczęściej czytana w gospodarce. Pozwala zdobyć niezbędne informacje, ale i zorientować się, jaki jest stan wiedzy konkurentów.
Prasa lokalna	Źródło szczegółowych informacji dla uważnego czytelnika (plany rozwoju na lokalnym rynku, konflikty z pracownikami i wiele innych)
Pisma fachowe i branżowe	Szczegółowe informacje fachowe, często napisane językiem zrozumiałym tylko dla specjalistów.
Źródła rządowe i centralne	
Urzędy	Urzędy takie jak Komisja Papierów Wartościowych i Giełd, czy Główny Urząd Statystyczny dysponują olbrzymią bazą danych. Jej część jest publikowana, jednak o ujawnienie reszty trzeba zabiegać. Polecamy wizytę w biurach patentowych – tam dobrze widać, nad czym pracuje konkurencja.
Ministerstwa	Nie dość, że dysponują cennymi informacjami, to podejmują decyzje o znaczeniu dla całej gospodarki (a zawsze można znaleźć rozgadanego urzędnika).
Sejm	Uchwalane ustawy i ich projekty mają wpływ na działalność gospodarczą (a niektórzy posłowie, jak urzędnicy, czasem powiedzą za dużo).
Oficjalne dokumenty informacyjne	Zazwyczaj mają pokazać Polskę w dobrym świetle, często jednak zawierają ciekawe informacje, strzępy planów na przyszłość.
Informacje od konkurentów	
Przemówienia, komunikaty prasowe	Są to informacje z pierwszej ręki, udostępniane często w przełomowych momentach dla firmy. Zazwyczaj mają dużą wartość dla wywiadowcy.
Strony internetowe	Zazwyczaj nie ma na nich nic ciekawego (w Polsce). Czasem jednak stanowi zaskakująco bogate źródło danych.

Dokumenty oficjalne	Takie jak sprawozdania roczne – ukazują priorytety, plany inwestycyjne, czasem wytłumaczenie popełnionych błędów. Z drugiej jednak strony mogą być teorią dorabianą do nieudolnych działań.
Dokumenty wewnętrzne	Np. gazetki, przesłania zarządu – wyjaśniają działania, plany na przyszłość, prezentują firmę. Nie zawsze są łatwe do zdobycia, ale warto zadać sobie trud.
Pracownicy	Najwspanialsze ze wszystkich źródeł, wymaga jednak umiejętności nawiązania i pokierowania rozmową. Jest to bardzo trudne, ale bywa bardzo opłacalne.
Wycieczki po zakładach	Organizowane dla chętnych gapiów (za opłatą), albo przy okazji prowadzonych rozmów. Dobre źródło dla spostrzegawczych.
Źródła operacyjne	
Konkurenci naszego klienta	Czasem warto zjednoczyć siły i wymieniać informacje z innymi firmami. Jest to duża pomoc przy próbie zrozumienia działań firmy na innych rynkach.
Klienci	Klienci znają dokładnie produkty konkurenta – ich wady i zalety, mogą otrzymywać informacje o rychłym wprowadzeniu nowego produktu.
Dostawcy	Nawet, jeśli nie dostarczają produktów konkurentowi, z pewnością wiedzą, co się dzieje na rynku.
Związki zawodowe	Znakomite źródło informacji zarówno przekrojowych (mają dane o tym, co dzieje się w wielu zakładach), jak i w poszczególnych firmach.
Dziennikarze	Zazwyczaj wiedzą więcej niż napisali w artykule i chętnie informacje te udostępniają, zwłaszcza, gdy pracują już nad innym materiałem.
Izby handlowe	Jednym z powodów ich istnienia jest wymiana informacji. Nie można więc do nich nie należeć.
Organizacje konsumenckie	Organizacje zbierające informacje po to, by dzielić się nimi z innymi, np. żeby ostrzegać przed niesolidnymi producentami. Uzyskanie informacji nie stanowi większego problemu, a ich jakość często jest bardzo wysoka.
Inne	

Źródło: M. Jaworski, *Wywiad gospodarczy na wewnętrzny użytek*, 2002, <http://ebib.oss.wroc.pl/2002/40/jaworski.php> (20.04.2010).

5.2. Zalety i wady danych wtórnych

Dane wtórne niosą ze sobą wiele korzyści. Przede wszystkim oferują oszczędności czasu i kosztów. Prowadzący badanie rozpoczynając zbieranie danych powinien przede wszystkim udać się do biblioteki, w której może znaleźć odpowiadającą mu pozycję oraz wykorzystać zawarte w niej dane. Całość tego procesu nie powinna zająć badaczowi więcej niż kilka godzin, gdy tymczasem przygotowanie całej procedury do zbierania pierwotnych danych, jest zajęciem czasochłonnym. Badacz musiałby opracować ramę doboru, ułożyć pytania w kwestionariuszu, przetestować kwestionariusz, zatrudnić ankietatorów, przeszkolić ankietatorów, a następnie zebrać dane mając na uwadze wszystkie rodzaje błędów badania. Zbieranie danych pierwotnych jak widzimy jest procesem długim, który może trwać nawet kilka miesięcy. Dlatego należy pamiętać, aby w żadnym wypadku nie omijać tego stadium badań marketingowych. Zbieranie danych należy rozpocząć właśnie od analizy danych wtórnych, i dopiero, gdy okażą się one niewystarczające, wówczas przygotować się do zbierania danych pierwotnych.

Dane wtórne mogą dostarczyć tyle informacji, że zbieranie danych pierwotnych nie jest już konieczne. Okazuje się, że zamawiający w celu redukcji kosztów badania, prowadzi jedynie studia nad danymi wtórnymi. Rozwiązanie niektórych problemów jest możliwe poprzez dokładną analizę literatury branżowej, naukowej, czy danych opublikowanych przez odpowiednie organy rządowe. Całkowite rozwiązanie problemu na podstawie analizy danych wtórnych nie jest częste, jednak pomagają one poprawnie sformułować problem badawczy, wskazują, które metody zbierania danych pierwotnych będą właściwe dla danego studium, dostarczają porównywalnych danych. Wady danych wtórnych są skoncentrowane wokół dwóch podstawowych problemów: dopasowania do problemu i dokładności⁷⁴.

Prowadzący badanie analizując dane wtórne szybko zorientuje się, że nie są one dokładnie dopasowane do problemu, którym się zajmuje. Wyróżnia się trzy podstawowe przyczyny problemu niedopasowania: problem jednostki miary, problem definicji klas, problem aktualności⁷⁵. Problem związany z jednostką miary związany jest z faktem, iż wiele danych zapisywanych jest w inny sposób, niż jest to potrzebne do studium.

Kolejną wadą danych wtórnych jest aktualność. Problem aktualności danych pojawia się często podczas prowadzenia badań marketingowych. Należy pamiętać, że opublikowane dane mogą pochodzić sprzed długiego okresu i nie zawsze można posłużyć się nimi w badaniu z uwagi na zmiany

⁷⁴ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 553.

⁷⁵ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 553.

zachodzące na rynku. W zależności od rodzaju podejmowanej decyzji marketingowej, dane wtórne mogą być z dużą, średnią bądź małą aktualnością. Przy podejmowaniu decyzji operacyjnych od źródła danych wtórnych oczekuje się bardzo dużej aktualności danych, natomiast w przypadku podejmowania decyzji strategicznych oczekiwania menedżerów nie są już takie duże i dopuszcza się średnią aktualność danych⁷⁶.

Wiarygodność danych wtórnych określona jest przez pięć czynników: dostępność, dokładność, aktualność, trafność, porównywalność⁷⁷. Dostępność danych wtórnych wyznaczana jest przez możliwości zespołu badawczego. Im lepsze przygotowanie zespołu pod względem wiedzy i doświadczenia, tym okres poszukiwania właściwych danych się skraca. Dokładność danych zależy od doboru metody gromadzenia danych. Badacz powinien także zwrócić uwagę na źródło oferujące dane wyniki. Trafność określana jest jako stopień użyteczności danych w procesie decyzyjnym z dostosowaniem do zakresu czasowego, podmiotowego, przestrzennego i przedmiotowego.

Problem ścisłości danych wtórnych zależy przede wszystkim od tego, czy dane wtórne zostały zebrane ze źródeł pierwotnych lub wtórnych. Źródło pierwotne jest to źródło, które wytworzyło dane. Źródło wtórne to źródło, które pozyskało dane ze źródła pierwotnego⁷⁸. Prowadzący badanie powinien ograniczyć poszukiwania jedynie do danych pierwotnych. Niezachowanie ostrożności podczas *desk research*, może doprowadzić do sytuacji, w której wynik badania będzie opierał się na danych wtórnych, co do których nie ma pewności, czy zostały zebrane przy zastosowaniu prawidłowej metody gromadzenia danych. Źródła pierwotne są przede wszystkim dokładne. Badacz jest w stanie ocenić sposób gromadzenia danych oraz wykryć popełnione w badaniu błędy. Kolejny aspekt korzystania z niesprawdzonych danych wtórnych, to cel publikacji. Badania publikowane i opracowywane na potrzeby kampanii reklamowych, czy zniesławienia konkurenta, nie mogą być traktowane wiążąco dla prowadzącego badanie. Wynika to z tego, iż rozwój informacji wtórnej

⁷⁶ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 69.

⁷⁷ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 81.

⁷⁸ G.A. Churchill, Badania marketingowe. Podstawy metodologiczne, PWN, Warszawa 2002, s. 255.

związany jest z dynamicznym wzrostem rynku, szczególnie na rynku dóbr usług konsumpcyjnych⁷⁹.

Bardzo istotnym ułatwieniem dostępu do źródeł wtórnych jest Internet. Do niewątpliwych zalet Internetu należą: oszczędność kosztów, prowadzący badanie jest w stanie dotrzeć do interesujących go informacji w krótkim czasie, dane są niekiedy publikowane w sieci zaraz po ich powstaniu, wielorakie kryteria wyszukiwania haseł, nieograniczona liczba egzemplarzy, większa dostępność do komentarzy, wykresów, opracowań. Do wad korzystania z danych internetowych zaliczamy m.in.: wiarygodność źródeł i odpłatność niektórych serwisów.

5.3. Przykłady z omówieniem

Przykład 1. Ze wszystkich danych wtórnych wewnętrznych istniejących w przedsiębiorstwie, faktura jest dokumentem dostarczającym najwięcej informacji:

1. Dane nabywcy – nazwa klienta, siedziba, forma prowadzenia działalności gospodarczej, dane osoby fizycznej.
2. Nazwa towaru/usługi – określenie zakupionego produktu, usługi.
3. Wartość towaru/usługi – cena netto, jaką nabywca zapłacił.
4. Miejscowość, data – data dokonania transakcji kupna-sprzedaży, miejsce dokonania transakcji.
5. Ilość zakupionego produktu/usługi.
6. Dane osoby upoważnionej do wystawienia faktury, często w tę rubrykę wpisuje się dane agenta handlowego.
7. Sposób zapłaty – określa formę przepływu wartości towaru: przelew, gotówka.
8. Dodatkowe informacje – mogą być także umieszczane na odwrocie faktury, np. opis ostatecznego przeznaczenia produktu/usługi; środki transportu użyte do wysyłki; określenie kosztów rodzajowych, do których można zakwalifikować ten dokument księgowy.

⁷⁹ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s. 82.

FAKTURA VAT NR								
Miejscowość, data: 4					ORYGINAŁ / KOPIA			
Data sprzedaży:					Nabywca: 1			
Sprzedawca:					Termin płatności:			
2					3			
5					7			
Sposób zapłaty: gotówka / przelew.								
L.p.	Nazwa towaru lub usług	Symbol PKWiU	J.m.	Ilość	Cena jedn. netto	Wartość towaru/usługi netto	Podatek VAT	Wartość brutto
Słownie:					RAZEM		0%	
					WTYM:		7%	
							22%	
							ZW.	
Imię i nazwisko oraz podpis osoby upoważnionej do odbioru faktury			6 Imię i nazwisko oraz podpis osoby upoważnionej do wystawienia faktury			DO ZAPŁATY		

Przykład 2. Na przedstawionym poniżej przykładzie widzimy, jakiego typu problemy możemy napotkać w trakcie przeprowadzania studium nad danymi wtórnymi. Ludność została pogrupowana według wieku i płci. Jeśli badacz zamierza zbadać osoby między 52 a 57 rokiem życia, wówczas występuje problem definicji klas, gdyż jak widzimy osoby w wieku 52-57 lat mieszczą się w dwóch klasach. Prezentowane dane nie będą przydatne z punktu widzenia prowadzonego studium.

LUDNOŚĆ WEDŁUG PŁCI, WIEKU I WOJEWÓDZTW W 2007 R.
 POPULATION BY SEX, AGE AND VOIVODSHIPS IN 2007
 Stan w dniu 30 VI
 As of June 30

WOJEWÓDZTWA VOIVODSHIPS WIEK AGE	Ogółem Total	Mężczyźni Males	Kobiety Females	Miasta <i>Urban areas</i>			Wieś <i>Rural areas</i>		
				razem total	mężczyźni males	kobiety females	razem total	mężczyźni males	kobiety females
POLSKA POLAND	38115967	18417074	19698893	23254482	11031728	12222754	14861485	7385346	7476139
0 lat <i>0 years</i>	377981	194516	183465	219263	112759	106504	158718	81757	76961
1—4	1427123	733903	693220	815851	419915	395936	611272	313988	297284
5—9	1888577	968674	919903	1026835	527020	499815	861742	441654	420088
10—14	2266490	1158694	1107796	1213654	620278	593376	1052836	538416	514420
15—19	2748281	1405084	1343197	1525076	779168	745908	1223205	625916	597289
20—24	3242517	1648625	1593892	1922026	973842	948184	1320491	674783	645708
25—29	3142989	1593343	1549646	1992656	995090	997566	1150333	598253	552080
30—34	2883250	1459721	1423529	1828092	914872	913220	1055158	544849	510309
35—39	2447095	1236687	1210408	1481463	739109	742354	965632	497578	468054
40—44	2407351	1209305	1198046	1438878	701496	737382	968473	507809	460664
45—49	2829230	1399485	1429745	1765220	836587	928633	1064010	562898	501112
50—54	3018291	1464697	1553594	1995041	927340	1067701	1023250	537357	485893
55—59	2648538	1252293	1396245	1780185	814706	965479	868353	437587	430766
60—64	1667319	758254	909065	1111826	496553	615273	555493	261701	293792
65—69	1478153	633936	844217	946583	397780	548803	531570	236156	295414
70—74	1382265	554446	827819	860195	340513	519682	522070	213933	308137
75—79	1147664	419838	727826	683847	249927	433920	463817	169911	293906
80—84	717779	220588	497191	418211	126775	291436	299568	93813	205755
85 lat i więcej <i>85 years and more</i>	395074	104985	290089	229580	57998	171582	165494	46987	118507

Źródło: Rocznik Statystyczny 2009, GUS, Warszawa 2009, s. 46.

5.4. Zadania do samodzielnego opracowania

Zadanie 1. Określ główne wtórne źródła informacji dla następujących sytuacji:

- Pan Nowak planuje otworzyć sklep spożywczy. Chce się dowiedzieć, jak wygląda ta branża w jego rodzinnym mieście.
- Pani Kowalska słyszała, że branża kosmetyczna zapewnia wysoką marżę zysku. Branża ta zawsze ją intrygowała. Postanowiła przeprowadzić badania, celem ustalenia czy odpowiada to prawdzie.
- Państwo Krajewscy planują otwarcie nowego domu weselnego w swoim mieście. Postanowili znaleźć informacje źródłowe, aby ocenić potencjał rynku.
- Kierownik działu marketingu producenta farb musi przygotować kompleksowy raport na temat tej branży.

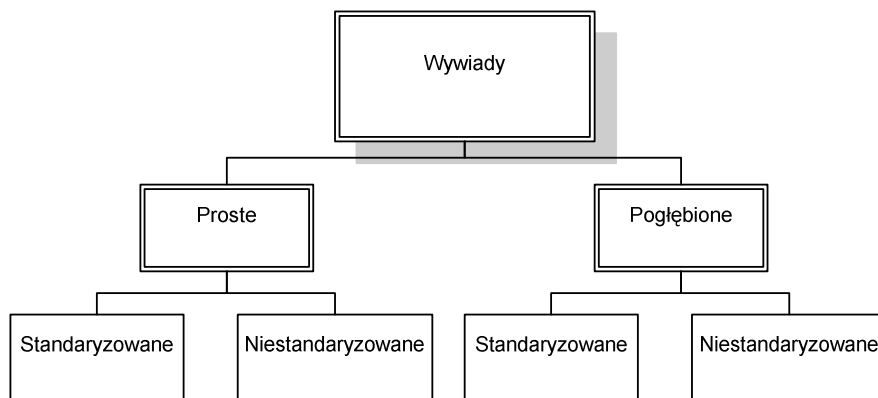
6. GROMADZENIE DANYCH PIERWOTNYCH

6.1. Wywiady

Wywiady w badaniach marketingowych służą bezpośredniej komunikacji ankietera z respondentem. W badaniach marketingowych najważniejsze są dwie cechy kwestionariuszy wywiadów, ze względu, na które dokonuje się podziału. Są to mianowicie standaryzacja i głębokość. Im bardziej standaryzowany wywiad, tym prowadzący badania pozostawia mniej pola do działania prowadzącemu wywiad i respondentowi. Badacz nie ma możliwości prowadzenia wywiadu według własnych kryteriów, a respondent nie może odpowiadać na pytania własnymi słowami. Wywiady, które są standaryzowane w literaturze mogą być określane mianem wywiadów ustrukturyzowanych lub wywiadów kwestionariuszowych.

Głębokość wywiadu oznacza stopień ukrycia jego celu przed respondentem. Wywiady, które nie pozostawiają respondentowi wątpliwości co do ich celu, nazywamy wywiadami prostymi lub wywiadami jawnymi. A contrario wywiady, w których cel badania jest ukryty przed respondentem, nazywamy wywiadami pogłębionymi lub wywiadami niejawnymi. Ukrycie celu badania przed respondentem pełni ważną funkcję w badaniach marketingowych. Respondent, który nie jest świadomy prawdziwego celu badania, dostarcza więcej informacji na temat motywów jego postępowania. Wywiady pogłębione są jedną z metod przeprowadzania badań motywacyjnych.

Z uwagi na wymienione cechy wywiady dzielimy na cztery podstawowe rodzaje (por. rysunek 11): wywiad standaryzowany prosty, wywiad niestandaryzowany prosty, wywiad standaryzowany głęboki, wywiad niestandaryzowany głęboki.



Rys. 11. Podstawowe rodzaje wywiadów

Źródło: opracowanie własne

Wywiad standaryzowany prosty jest wywiadem, który jest najczęściej stosowany w badaniach marketingowych. Istota tego rodzaju wywiadu polega na standaryzacji pytań i odpowiedzi w kwestionariuszu. Celem standaryzacji pytań i odpowiedzi jest uniemożliwienie respondentom udzielenia odpowiedzi własnymi słowami. Zestaw pytań i odpowiedzi zawartych w kwestionariuszu zostaje opracowany dla całej badanej grupy. Każdy respondent z danej grupy podlega pomiarowi za pomocą tego samego kwestionariusza.

Wywiad niestandaryzowany prosty charakteryzuje się tym, że cel badania jest znany respondentowi. Nie ma natomiast opracowanej struktury badania. Badanie jest prowadzone w sposób dynamiczny. Dyskusja pomiędzy prowadzącym badanie a respondentem niejako wymusza sformułowanie pytań w czasie trwania wywiadu. Pytanie sformułowane jest w sposób jasny i prosty. Nie pozostawia żadnych wątpliwości, co do celu badania, np. Co sądzi Pani/Pan na temat prowadzenia działalności gospodarczej w Polsce? W literaturze można się spotkać z nazwą wywiad niestandaryzowany pogłębiony lub głęboki⁸⁰. Wywiad niestandaryzowany prosty charakteryzuje się poważnymi problemami przy analizie uzyskanych wyników. Interpretacja odpowiedzi jest zadaniem wyszkolonego zespołu psychologów, co generuje dodatkowe koszty badania.

Wywiad standaryzowany głęboki posługuje się kwestionariuszem. Kwestionariusz ma na celu ujawnienie świadomych bądź nieświadomych postaw respondentów. Zastosowanie kwestionariusza do mierzenia postaw, ma związek z charakterem badanej postawy. Kwestionariusz wywiadu jest skonstruowany w oparciu o pewne założenie, co do reprezentowanej przez respondenta postawy. Założenie to zakłada, że wiedza, postrzeganie i pamięć respondenta są uzależnione od jego postawy. Kwestionariusz wywiadu zawiera więc pytania o wiedzę respondenta. Kwestionariusz zawiera pytania pośrednie z wyskalowanymi odpowiedziami. Im większa wiedza respondenta w danym temacie, tym większa siła i ukierunkowanie postawy. Wywiady głębokie standaryzowane można prowadzić zarówno z jednym respondentem, jak i grupami respondentów.

Wywiady głębokie niestandaryzowane są znane także jako badania motywacyjne. Prowadzący badanie powinien mieć świadomość, iż respondenci niechętnie ujawniają swoje postawy i opinie. Zastosowanie wywiadu niestandaryzowanego może przynieść wyniki inne od oczekiwanych. Niektóre pytania w kwestionariuszu nie ujawnią ukrytych postaw respondenta. Dlatego badania nad motywacją są dość ważną dziedziną badań marketingowych. Respondent może nie chcieć, bądź nie być w stanie, opisać motywów swojego zachowania. Zadaniem prowadzącego badanie jest

⁸⁰ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 310.

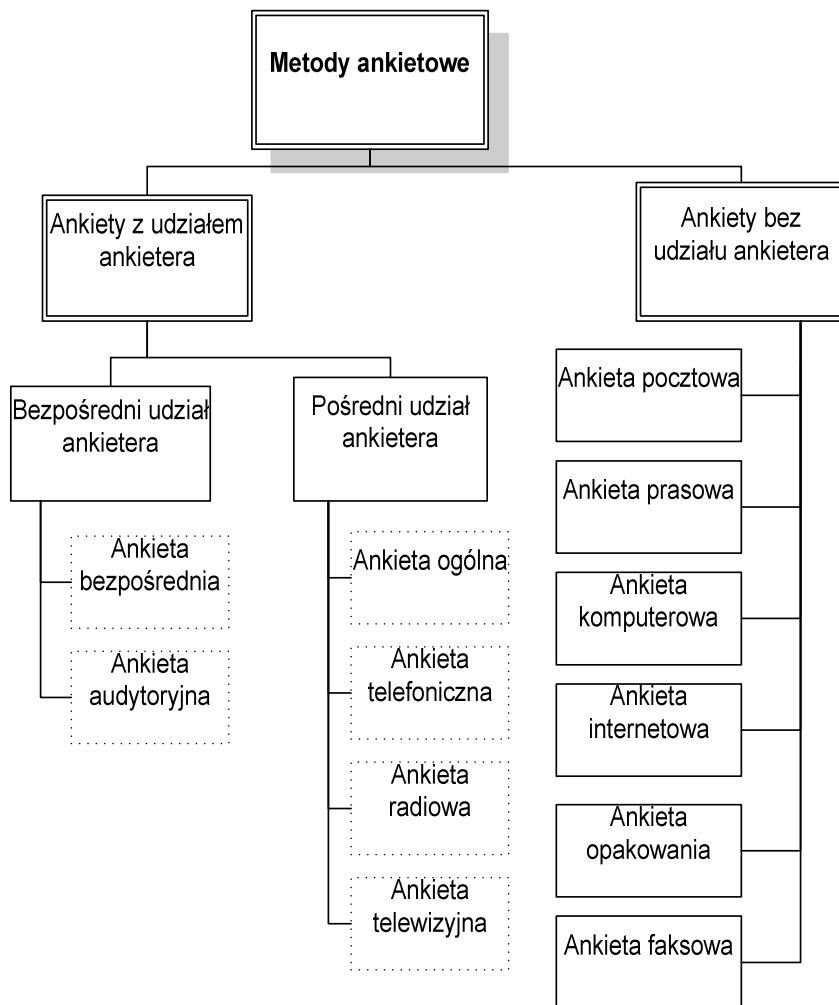
ominięciu niechęci respondenta do udzielenia informacji z wykorzystaniem specjalnie opracowanych technik. Techniki te nazywamy technikami projekcyjnymi⁸¹.

6.2. Badania ankietowe

Wywiady są formą pomiaru bezpośredniego, natomiast w przypadku badań ankietowych mamy do czynienia z formą pisemną. Ankieta jest więc metodą prowadzenia badania, w której prowadzący badanie komunikuje się z respondentem za pomocą kwestionariusza, który respondent wypełnia samodzielnie, a następnie zwraca prowadzącemu badanie. Jeżeli w badaniach ankietowych mamy do czynienia z ankieterem, jego rola ogranicza się do rozdawania i zbierania kwestionariuszy. Pośrednia forma kontaktu z respondentem może być oceniana dwojako. Z jednej strony pozytywnie, gdyż osoba ankietera nie wpływa na wynik badania, lecz z drugiej strony, respondent nie uzyskuje pomocy ze strony ankietera przy wypełnianiu kwestionariusza.

Kryteria podziału ankiet ze względu na kryterium formy kontaktu z respondentem przedstawione są na rysunku 16.

⁸¹ Respondent oceniając innych ludzi, ich postawy, motywy, oceny, przypisuje im nieświadomie swoje cechy. To przeniesienie cech na osoby lub rzeczy nazywamy projekcją. Inną definicję projekcji przedstawia Maison „podmiot postrzega otoczenie i odpowiada na nie zgodnie z własnymi zainteresowaniami, postawami, przyzwyczajeniami, trwałymi lub chwilowymi stanami emocjonalnymi, oczekiwaniami, pragnieniami”. Por. D. Maison, A. Noga-Bogomilski, *Badania marketingowe. Od teorii do praktyki*. Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2007, s. 47. Techniki projekcyjne zakładają użycie niejawnego bodźca. Bodziec ten powinien zostać opisany, rozwinięty przez respondenta, a tym samym odkryć prawdę o respondencie. Techniki projekcyjne stosuje się do zidentyfikowania wyobrażeń respondenta dotyczących korzyści produktu, marki oraz wyjaśnienia przyczyn dokonanego zakupu. Respondenci nie ujawniają swoich postaw w pytaniach wprost z następujących przyczyn: nie rozumieją, dlaczego robią, to, co robią („Zawsze to kupuję, tak z przyzwyczajenia”, „Kupuję to, bo wiem, że jest najlepsze”, „Po prostu najbardziej mi smakuje”), zdają sobie sprawę, co stanowi przyczynę ich zachowania, ale nie są w stanie tego wyrazić („Trudno jest mi mówić o swoich odczuciach”, „Nie jestem pewien, co czuję, doprawdy nie wiem”, „Tak trudno wyrazić to słowami, to się po prostu czuje i już”), nie chcą wyjawiać swoich motywów z poczucia wstydu, obawy czy zniechęcenia („Lepiej działa”, „Jest bezpieczniejszy”, „Ma dobrą jakość”).



Rys. 12. Klasyfikacja ankiet

Źródło: S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 179.

Ankieta pocztowa jest metodą często wykorzystywaną w badaniach marketingowych. Istota zastosowania tej techniki polega na rozesłaniu kwestionariusza do respondentów za pomocą poczty. Respondent udziela odpowiedzi na pytania zawarte w kwestionariuszu. Następnie odsyła za pośrednictwem poczty kwestionariusz pod wskazany adres. W porównaniu do innych metod zbierania danych jest ona stosunkowo tania. Za jej pomocą prowadzący badanie ma możliwość dotarcia do dużej grupy respondentów rozproszonych terytorialnie.

Pomiar z pomocą ankiety prasowej polega na zamieszczeniu w czasopiśmie kwestionariusza. Respondenci wypełniają załączone kwestionariusze, a następnie odsyłają pod wskazany adres. W praktyce okazuje się, że przeprowadzenie badania za pomocą ankiety prasowej, jest dużo bardziej efektywne, niż ankiety pocztowej. Okazuje się, że ankiety dołączone do czasopisma budzącego zaufanie wśród respondentów, są częściej zwracane prowadzącym badanie.

Ankieta audytoryjna polega na rozdaniu kwestionariusza osobom, np. podczas konferencji, sympozjów. Kwestionariusze rozdawane są respondentom w jednym miejscu i czasie. Wypełnione kwestionariusze zbierane są przez ankietera albo wrzucane do wcześniej przygotowanej urny. Ankieta audytoryjna, z uwagi na proces jej przeprowadzenia, jest metodą szybką i tanią. Pozwala ona w krótkim czasie dotrzeć do wielkiej ilości respondentów.

Ankieta ogólna polega na umieszczeniu kwestionariuszy w miejscach ogólnodostępnych. Osoby zainteresowane wzięciem udziału w badaniu mogą podejść i wypełnić kwestionariusz. Wypełnione formularze można złożyć w wyznaczonym do tego miejscu bądź przesać je na wskazany adres.

Ankieta bezpośrednia polega na dostarczaniu kwestionariusza bezpośrednio za pomocą ankietera. Osoby poddawane badaniu są proszone o samodzielnie wypełnienie otrzymanego kwestionariusza, a następnie jego zwrot. Zwrot kwestionariusza może polegać na wysłaniu go pocztą lub na odebraniu kwestionariusza przez ankietera.

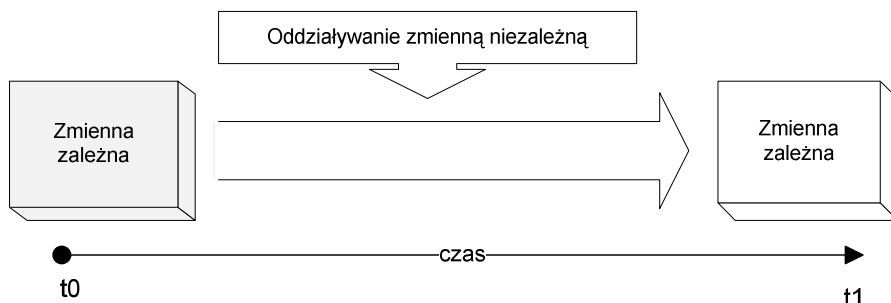
Ankiety internetowe stają się coraz bardziej popularną metodą prowadzenia badań marketingowych. Badania za pomocą ankiet internetowych dotyczą przede wszystkim rynku B2B⁸². Wykorzystanie ankiet internetowych w badaniu marketingowym posiada wiele zalet. Wykorzystanie Internetu, jako instrumentu pośredniego, wiąże się z oszczędnością czasu i kosztów. Badania prowadzone za pomocą Internetu dają możliwość przeprowadzenia badania na całym świecie. Ograniczeniem prowadzenia badań internetowych jest dostępność Internetu. Problemem, który pojawia się przy stosowaniu ankiet wysyłanych pocztą elektroniczną jest odrzucanie ich przez respondentów traktujących je jako spam. Kolejnym problemem, na który może natknąć się prowadzący badanie, jest niezgodność oprogramowania pomiędzy nim a respondentem.

6.3. Eksperyment

Eksperyment jest metodą polegającą na zbadaniu zależności pomiędzy jedną zmienną a drugą. Prowadzący badanie steruje nasileniem zmiennej pierwszej, nazywanej zmienną niezależną (np. liczba emisji reklamy), w celu

⁸² B2B – business-to-business.

wywołania zmiany w nasileniu drugiej zmiennej, nazywanej zmienną zależną (np. zamiar zakupu, dokonanie zakupu). Zmienna niezależna i zmienna zależna są podstawowymi składnikami każdego eksperymentu. Obserwacja zmiany w nasileniu zmiennej zależnej, pozwala wnioskować prowadzącemu badanie na temat istnienia i kształtu związku przyczynowo-skutkowego między nimi. Rysunek 13 ilustruje przebieg eksperymentu.



Rys. 13. Schemat przebiegu eksperymentu

Źródło: opracowanie własne.

Badanie w formie eksperymentu powinno zostać podjęte, jeśli spełnione są trzy warunki:

- istnieje prawdopodobieństwo, że zachodzi związek pomiędzy przyczyną a jej skutkiem⁸³,

⁸³ Pierwszą ważną zaletą eksperymentu jest samo stwierdzenie, czy istnieje związek przyczynowo-skutkowy pomiędzy zmienną niezależną a zmienną zależną. Dla menedżerów istotne jest potwierdzenie bądź zaprzeczenie istnienia związku pomiędzy działaniami marketingowymi a zachowaniami konsumentów. Przykładem obrazującym jak ważne jest potwierdzenie związku przyczynowo-skutkowego jest badanie kampanii reklamowej kawy instant Nescafe przeprowadzone w latach 50. Przeprowadzona kampania reklamowa kawy instant Nescafe nie przyniosła oczekiwanych przez menedżerów rezultatów. Wówczas przy wykorzystaniu technik projekcyjnych przeprowadzono eksperyment mający wykazać istnienie i siłę związku pomiędzy działaniami promocyjnymi a zachowaniami konsumentów. Grupę badanych osób podzielono na dwie grupy. Obu zaprezentowano listę zakupów. Jedna grupa otrzymała listę zakupów, na której była kawa instant Nescafe. Druga grupa otrzymała listę, zawierającą kawę sypaną Maxwell House. Przeprowadzone badanie pokazało, że nie istnieje związek pomiędzy intensywnością promocji a zachowaniem konsumenta. Niska sprzedaż kawy Nescafe spowodowana była złym wizerunkiem potencjalnego jej użytkownika wśród konsumentów. Osoba dokonująca zakupów kawy instant Nescafe określona została jako „zła gospodyni”,

- zależność ta możliwa jest do odnotowania,
- możliwe jest wyeliminowanie lub kontrolowanie czynników niezależnych od badacza, które także mogłyby wywołać badany skutek⁸⁴.

Prowadzenie badań w formie eksperymentu posiada dwie podstawowe zalety:

- możliwość badania związków przyczynowo-skutkowych oraz ich siły
- unikanie zadawania pytań wprost.

W badaniach prowadzonych na konsumentach często wykorzystywaną techniką badań jest zadawanie pytań wprost. Pytania wprost obarczone są jedną podstawową wadą – opierają się na subiektywnych deklaracjach respondentów. Prowadzący badanie przyjmuje, że uzyskane odpowiedzi są odzwierciedleniem rzeczywistych postaw i opinii. Jednak rzeczywistość może się różnić od tej, którą postrzega konsument. Eksperyment jest metodą, która jest w stanie określić rzeczywiste motywy zachowania konsumentów⁸⁵.

Każdy eksperyment prowadzony jest w określonych warunkach. Mogą to być warunki naturalne bądź sztuczne. Nasilenie sztuczności i naturalności eksperymentu może być różny. Eksperymenty możemy podzielić na: naturalne i sztuczne.

Eksperyment naturalny nie posiada tak dużego stopnia kontroli w porównaniu z eksperymentem laboratoryjnym. Zaletą eksperymentu naturalnego są naturalne warunki jego prowadzenia. Eksperymenty naturalne przeprowadza się w celu określenia preferencji konsumentów wobec nowych produktów, nowych opakowań, zmiany cen, zmiany narzędzia promocji. W ramach eksperymentu naturalnego stosuje się dwie jego odmiany: eksperyment standardowy oraz eksperyment kontrolowany. Eksperyment standardowy (test sprzedaży lub test rynkowy) polega na tym, iż wyznacza się próbę punktów sprzedaży z pewnego obszaru geograficznego (np. miasta). Produkt sprzedaje się przez funkcjonującą sieć dystrybucji, a następnie wprowadza się zmiany w oferowanej produkcie, jego cenie, sposobie promocji. Pomiaru dokonuje się za pomocą ankiety lub inwentaryzacji. Kluczowym czynnikiem powodzenia eksperymentu jest wyodrębnienie zasięgu geograficznego oddziaływania eksperymentu. W tym celu stosuje się następujące kryteria: reprezentatywności kryterium eksperymentalnego i kontrolnego dla całej badanej populacji, wielkość dobranych terytoriów musi być optymalna ze względu na koszty i liczbę danych, wyznaczone terytoria powinny odzwierciedlać sytuację danego

„leniwa”. Wyniki badania pomogły menedżerom skupić się na prawdziwej przyczynie zachowań konsumenta, a skutkiem tego zwiększyć sprzedaż.

⁸⁴ A. Kaniewska-Sęba, G. Leszczyński, B. Pilarczyk, *Badania marketingowe na rynku business-to-business*, Oficyna Ekonomiczna, Kraków 2006, s. 105.

⁸⁵ D. Maison, A. Noga-Bogomilski, *Badania marketingowe. Od teorii do praktyki*, Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2007, s. 136.

rynku pod względem demograficznym i konkurencyjnym, terytoria nie powinny być podatne na wpływy innych terytoriów, ani same nie powinny wpływać na te terytoria⁸⁶.

Większość eksperymentów rynkowych prowadzonych jest bez grupy kontrolnej z jednym pomiarem po wprowadzeniu zmiennej niezależnej. W warunkach dużej konkurencji na rynku przeprowadzenie eksperymentu bez oddziaływania konkurencji jest praktycznie niemożliwe. Drugim rodzajem eksperymentu naturalnego jest eksperyment kontrolowany. Polega on na zleceniu przeprowadzenia eksperymentu agencji badawczej, która posiada własną, przeznaczoną głównie do celów eksperymentalnych sieć magazynów i sieć dystrybucyjną. Powodzenie eksperymentu jest uwarunkowane prowadzeniem go w małych miastach, na małych segmentach rynku. Zalety eksperymentu kontrolowanego to także: krótki czas prowadzenia, niskie koszty, ograniczony wpływ konkurentów.

Eksperyment sztuczny umożliwia maksymalne kontrolowanie wpływu zmiennych niezależnych na wynik pomiaru. Eksperyment przeprowadza się w warunkach sztucznych specjalnie zaaranżowanych do tego celu. Wadą tej metody jest wpływ eksperymentatorów i stworzonych przez nich warunków na wynik badania.

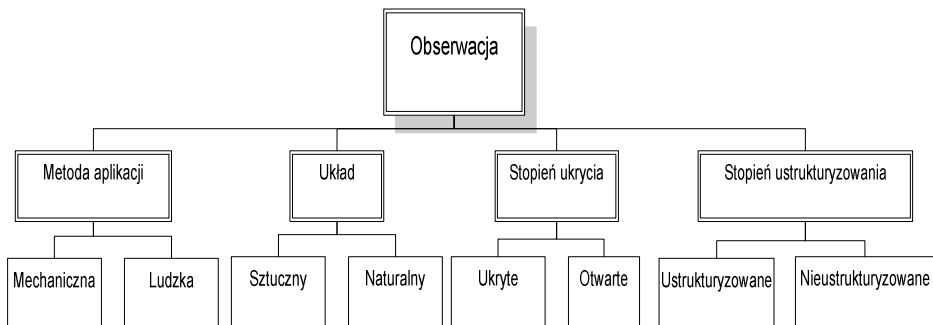
6.4. Obserwacja

Obserwacja jest metodą badań marketingowych wykorzystywaną rzadziej niż pozostałe metody. Obserwacja jest metodą prowadzenia badań jakościowych. Początki obserwacji datowane są na lata dwudzieste XX wieku. Obserwacje wykorzystywano przede wszystkim w antropologii kulturowej, etnografii oraz w naukach przyrodniczych. Rozwój metody obserwacji zawdzięczamy angielskim badaczom, którzy rozwinęli jako pierwsi szeroko zakrojone badania społeczności brytyjskiej nazywane masową obserwacją (mass observation)⁸⁷. Obecnie obserwacja wykorzystywana jest w badaniach marketingowych jako metoda wstępna, rozpoznawcza. Obserwacja polega na postrzeganiu celowym, zamierzonym i prowadzonym według ustalonego z góry planu. Jest ona zjawiskiem z życia codziennego. Obserwujemy innych ludzi i zjawiska w celu pozyskania informacji o otaczającym nas świecie. Obserwacje badawcze są planowane i rejestrowane tak, aby mogły się odnosić do określonego interesującego badacza zjawiska. Planowanie obserwacji nie świadczy o stopniu jej skomplikowania.

⁸⁶ S. Kaczmarczyk, *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003, s. 247.

⁸⁷ A. Nikodemska-Wołowik, *Jakościowe badania marketingowe*, Polskie Wydawnictwo Ekonomiczne, Warszawa 1999, s. 162.

Wyróżnia się trzy podstawowe wymiary obserwacji: obserwację kontrolowaną i niekontrolowaną, obserwację jawną i ukrytą oraz obserwację ustrukturyzowaną i nieustrukturyzowaną. Obserwacja standaryzowana to taka obserwacja, której problem został zdefiniowany wystarczająco dokładnie, aby pozwalał na jasne a priori określenie zachowań, które będą przedmiotem obserwacji i kategorii, które będą używane do rejestrowania i analizy sytuacji. Obserwacja nieustrukturyzowana jest stosowana w badaniach, w których nie ma precyzyjnie określonych problemów. Obserwatorzy mają dużo swobody co do przedmiotu obserwacji oraz sposobu jej rejestracji.



Rys. 14. Klasyfikacja obserwacji

Źródło: G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002 s. 304.

Obserwacja ukryta polega na tym, że obiekt pomiaru nie wie, że jest obserwowany. Obserwacja jawna zakłada, że obiekt pomiaru wie, że jest obserwowany. Przyczyną ukrycia obecności obserwatora jest skłonność ludzka do zachowywania się w sposób odmienny niż normalnie, przy świadomości, że jest się obserwowanym.

Obserwacja może być prowadzona w warunkach naturalnych lub sztucznych. Warunki naturalne mogą być całkowicie naturalne bądź możemy mieć do czynienia z pewną manipulacją eksperymentalną. Przykładowo możemy obserwować czynności, które dzieją się zwykle przy wyborze czekolady. Możemy także wprowadzić do sklepu materiały informacyjne i mierzyć ich skuteczność. Miarą skuteczności jest zakres wzbudzonego zainteresowania czekoladą, poszukiwanie jej i zastanawianie się nad zakupem określonej marki. Oba przykłady obserwacji prowadzone są w sklepie w warunkach naturalnych.

Obserwacja prowadzona w warunkach sztucznych (laboratoryjnych) pozwala na większą kontrolę wpływów zewnętrznych przez prowadzącego badanie. Wpływy zewnętrzne mogą znacząco oddziaływać na interpretację

obserwowanych zachowań nabywców. Wadą obserwacji prowadzonej w warunkach sztucznych jest fakt, iż stworzone warunki także mogą wpływać na zachowania nabywców. Powstają wówczas wątpliwości, co do wiarygodności wyników badania. W warunkach sztucznych prowadzący badanie nie oczekuje na samoistne wydarzenie się oczekiwanego zdarzenia, lecz może polecić uczestnikom zaangażowanie się w pożądaný typ zachowania. Z punktu widzenia procesu badania jest to bardzo duża oszczędność czasu i kosztów. Oznacza to, że badanie może zostać przeprowadzone w krótkim czasie (kilku dni lub tygodni). Warunki laboratoryjne pozwalają także na szersze wykorzystanie specjalistycznego sprzętu badawczego niż badanie w warunkach naturalnych.

W badaniach marketingowych wykorzystuje się częściej obserwację za pomocą notowania. Obserwacja prowadzona jest systematycznie przy udziale specjalnie wyszkolonych w tym celu osób. Obserwacja mechaniczna lub też elektroniczna wykorzystywana jest w badaniach marketingowych od dłuższego czasu. Szybki postęp nowych technologii ustawicznie rozszerza jej rolę i znaczenie w badaniach. Niektóre z urządzeń mechanicznych zostały zaprojektowane specjalnie w celu obserwacji reklamy. Dziedzinę badań marketingowych zajmującą się obserwacją reakcji fizjologicznych obiektów pomiaru za pomocą wyspecjalizowanych urządzeń pomiarowych, nazywamy pomiarem fizjologicznym. Pomiar fizjologiczny wykorzystuje takie narzędzia pomiaru jak: galwanometr, tachistoskop, kamera do obserwacji ruchu gałek ocznych, badanie fal mózgowych.

6.5. Przykłady z omówieniem

Przykład 1. Rozważmy następujący przykład – co Pani/Pan sądzi na temat zmniejszenia zakresu regulacji prawnej dotyczącej swobody działalności gospodarczej?

- Potrzebują zmniejszenia
- Potrzebują zwiększenia
- Ani zmniejszenia, ani zwiększenia
- Nie mam zdania

Osoby, które uważają, że zmniejszenie regulacji prawnej dotyczącej swobody działalności gospodarczej nie zmieni obecnego stanu rzeczy, gdyż konieczna jest także zmiana regulacji prawnej w zakresie odpowiedzialności materialnej urzędników państwowych, nie wyrażą swojej postawy. Odpowiedzi według ustalonych wariantów deformują ich zgodność ze stanem faktycznym. Możliwe jest, że same kategorie odpowiedzi, będą charakteryzowały się pewną tendencyjnością. Prowadzący badanie może popełnić błędy na etapie badań pilotażowych kwestionariusza, wobec czego

pytanie istotne z punktu widzenia odzwierciedlenia postaw respondenta zostanie pominięte.

Przykład 2. Budowa formularzy obserwacyjnych. Prowadzący badanie przy konstrukcji kwestionariusza musi zwrócić uwagę na to, co ma być obserwowane. Od badacza wymaga się podjęcia konkretnych decyzji na temat przedmiotu obserwacji, jednostek i kategorii, którymi należy się posłużyć do rejestracji zachowania. Warunkiem poprawnego przeprowadzenia obserwacji jest wcześniejsze określenie przedmiotu obserwacji, zanim przystąpi się do badania. Prowadzący badanie musi pamiętać, że każde wydarzenie można opisać na kilka różnych sposobów. Kiedy obserwujemy osobę kupującą czekoladę, jej zachowanie możemy opisać w wielu wariantach:

- „Osoba kupiła tabliczkę czekolady”
- „Kobieta kupiła tabliczkę czekolady”
- „Kobieta kupiła tabliczkę czekolady marki Milka”
- „Kobieta kupiła tabliczkę mlecznej czekolady z nadzieniem truskawkowym marki Milka”
- „Kobieta zapytała o czekoladę z nadzieniem truskawkowym Lindt i kiedy jej powiedziano, że takiej nie posiadają, kupiła czekoladę z nadzieniem truskawkowym Milka”

Aby obserwacja była skuteczna, należy z góry określić, które aspekty zachowania są istotne. Decyzja o tym, co należy obserwować, wymaga rozstrzygnięcia następujących kwestii:

- Kto ma być obserwowany? Każdy wchodzący do sklepu? Każdy kupujący? Każdy kupujący czekoladę?
- Jakie aspekty zakupu mają być rejestrowane? Nabyta marka? Marka, o którą prosi klient w pierwszej kolejności? A sam kupujący? Czy klient kupił czekoladę słodzoną czy gorzką? Czy należy zanotować jego płeć? Czy należy oszacować wiek tej osoby? Czy ma to jakieś znaczenie, że dana osoba jest sama czy w grupie?
- Kiedy należy dokonać obserwacji? W którym dniu tygodnia? O jakiej porze? Czy należy odnotować dzień czy godzinę? Czy obserwację należy zarejestrować dopiero po zakupie, czy trzeba rejestrować też klienta zwracającego się do sprzedawcy, nawet, jeśli nie dojdzie do transakcji?
- Gdzie należy prowadzić obserwację? W jakiego typu sklepie? W jaki sposób należy wybrać ten sklep? Jak należy go opisać na formularzu obserwacyjnym? Czy należy odnotować zakupy w automatach?⁸⁸

⁸⁸ Por. G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 394.

6.6. Zadania do samodzielnego opracowania

Zadanie 1. Odpowiedz na pytania:

- a) Co to jest eksperyment?
- b) Czym różni się eksperyment w warunkach naturalnych od tego w warunkach sztucznych?
- c) Co oznacza głębokość wywiadu, a co oznacza standaryzacja?
- d) Co to jest kwestionariusz standaryzowany prosty?
- e) Co to jest kwestionariusz niestandaryzowany głęboki?

Zadanie 2. Dokończ rozpoczęte zdania:

- a) Obserwacja ... polega na tym, że obiekt pomiaru nie wie, że jest obserwowana.
- b) Obserwacja pośrednia polega na obserwacji
- c) Obserwacja prowadzona w warunkach sztucznych (laboratoryjnych) pozwala na ... przez prowadzącego badanie.

Zadanie 3. Wpisz TAK lub NIE dla poniższych twierdzeń:

- a) Wywiady w badaniach marketingowych służą pośredniej komunikacji badacza z respondentem.
- b) Im bardziej standaryzowany wywiad, tym prowadzący badania pozostawia mniej pola do działania ankietarowi i respondentowi.
- c) Odpowiedzi w wywiadzie standaryzowanym prostym nie powinny być ograniczone do wybranych przez prowadzącego badanie wariantów.
- d) Wywiady, w których cel badania jest ukryty przed respondentem, nazywamy wywiadami prostymi.
- e) Celem standaryzacji pytań i odpowiedzi jest uniemożliwienie respondentom udzielenia odpowiedzi własnymi słowami.
- f) Kwestionariusz w wywiadzie standaryzowanym głębokim ma na celu ujawnienie świadomych bądź nieświadomych postaw respondentów.

Zadanie 4. Średniej wielkości producent mebli sosnowych wprowadzał na rynek nowy model systemu meblowego. Właściciel firmy musiał zdecydować na jeden z dwóch programów promocyjnych tego produktu: program A lub program B. Wybrał program promocyjny A, będąc przekonany, że zwiększy on sprzedaż w większym stopniu niż program promocyjny B. Spotkanie z menedżerem ds. marketingu ma rozstrzygnąć, w jaki sposób zostanie przeprowadzone badanie. Wyobraźcie sobie, że jesteście menedżerem ds. marketingu. Jakie badanie polecić właścicielowi i dlaczego?

Zadanie 5. Lokalna firma zajmująca się dostawą Internetu dla mieszkańców średniej wielkości miasta drogą bezprzewodową, prowadzona przez dwóch kolegów z czasów studiów, rozważa sposoby poprawy swoich usług. W szczególności firma chciała dowiedzieć się, czy klienci są zadowoleni z obecnych usług, ewentualnie wyjaśnić przyczyny tego niezadowolenia. Jakie badanie polecić właścicielowi i dlaczego?

7. WYBRANE KRYTERIA SEGMENTACJI RYNKU

7.1. Pojęcie segmentacji rynku

Nabywcy różnią się od siebie pod wieloma względami (np. siły nabywczej, lokalizacji geograficznej, preferencji zakupu, cech demograficznych). Segmentacja rynku oznacza „podział rynku na oddzielne grupy nabywców, którzy mogą wymagać osobnych produktów i/lub odrębnego marketingu mix”⁸⁹. Segmentacja umożliwia identyfikację i wybór docelowych rynków działania.

Segmenty rynku konsumentów próbuje się definiować według charakterystycznych cech konsumenta. Segmentacja geograficzna wymaga podziału rynku na jednostki geograficzne (np. państwa, regiony, gminy). Segmentacja demograficzna wymaga podziału na grupy na podstawie cech demograficznych (np. dochodu, zawodu, wykształcenia, wieku, płci, liczby członków rodziny, cyklu życia rodziny, religii, rasy, narodowości). Segmentacja według cech psychograficznych wymaga podziału rynku na grupy pod względem przynależności do klasy społecznej, stylu życia i/lub cech osobowości. Segmentacja behawioralna wymaga podziału na grupy na podstawie wiedzy, postawy, sposobu użytkowania i reakcji nabywców na produkt⁹⁰.

Kryteria charakteryzujące wzorce zakupów sprowadzają się do wyodrębnienia jednorodnych grup nabywców na podstawie motywów zakupu. Można dokonać podziału nabywców na podstawie kryterium: lojalności wobec firmy, częstotliwości zakupów, metody zbierania informacji (odróżnia się grupy zainteresowane głównie niskimi cenami od grup zwracających uwagę na wysoką jakość produktu i obsługi).

Segmenty rynku przedsiębiorstw próbuje się definiować według cech demograficznych, zmiennych operacyjnych, podejścia do aktu kupna, czynników sytuacyjnych oraz cech osobowych. Segmentacja demograficzna związana jest z gałęzią gospodarki, wielkością przedsiębiorstwa oraz określeniem zainteresowań obszarami geograficznymi. Wśród zmiennych operacyjnych wymienia się technologię, status użytkownika i wymagania klienta. W podejściu do aktu kupna uwzględnia się organizację funkcji zakupu, strukturę władzy, charakter istniejących stosunków, ogólną politykę oraz kryteria dokonywania zakupu. Wśród czynników sytuacyjnych wymienia się nagłą potrzebę, szczególne zastosowanie oraz wielkość zamówienia. Wśród cech osobowych uwzględnia się podobieństwo między nabywcą a sprzedawcą, postawę wobec ryzyka oraz lojalność⁹¹.

⁸⁹ P. Kotler, Marketing, Gebethner&Ska, Warszawa 1994, s. 243.

⁹⁰ Ibidem, s. 243-255.

⁹¹ Ibidem, s. 255-256.

7.2. Cechy demograficzne, socjologiczne, psychologiczne i stylu życia

Dla menedżerów marketingu szczególnie interesujące są: cechy demograficzne i socjologiczne badanej grupy, cechy psychologiczne i cechy stylu życia, postawy, świadomość, motywacja, zachowanie. Zebranie danych demograficznych i socjologicznych grupy badanej jest konieczne dla prawidłowego sporządzenia wniosków końcowych. Dane demograficzne grupy badanej to wiek, płeć, dochód, stan cywilny, miejsce zamieszkania, klasa społeczna. Identyfikacja kluczowych cech konsumentów prowadzi bezpośrednio do wyodrębnienia segmentów rynku.

Cechy charakteryzujące konsumenta:

- Społeczno – ekonomiczne
 - dochód,
 - wykształcenie,
 - zawód,
 - miejsce zamieszkania,
 - kryteria geograficzne.
- Demograficzne
 - płeć, wiek,
 - wielkość i struktura rodziny,
 - faza cyklu życia.
- Psychograficzne
 - styl życia, osobowość,
 - aktywność,
 - zainteresowania,
 - postawy, opinie.

Sytuacje w jakich użytkuje się produkt:

Wynikające z organizacji życia społecznego:

- codzienne sytuacje,
- tradycyjne święta,
- uroczystości i inne wydarzenia rodzinne
- życie towarzyskie.



Rys. 15. Kryteria segmentacji rynku

Źródło: K. Mazurek-Łopacińska, *Zachowania nabywców jako podstawa strategii marketingowej*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1997, s. 146.

Kolejnym typem danych, które interesują menedżerów są dane psychologiczne i styl życia konsumentów. Cechy osobowości mają duży

wpływ na zachowanie konsumenta, są stałe i niezależne od czynników zewnętrznych. Identyfikacja cech osobowości konsumentów pozwala na określenie i wyjaśnienie ich zachowania⁹².

Kaczmarczyk podaje, że „styl życia to specyficzny zespół codziennych zachowań określonej zbiorowości ludzkiej, który jest wyrazem położenia społecznego i umożliwia jej społeczną identyfikację. Styl życia to nie tylko zachowanie ludzkie, lecz także psychofizyczne mechanizmy będące podstawami owych zachowań, a zwłaszcza motywy, postawy, systemy wartości”⁹³. Styl życia jest zjawiskiem złożonym, na które składają się następujące elementy: ogólny profil konsumenta, elementy behawioralne, poziom i struktura konsumpcji, sposoby spędzania wolnego czasu, postawy obyczajowe, kontakty z mediami i referencje w tym zakresie, postrzeganie samego siebie, wartości życiowe i rodzinne, upodobania, pasje i zainteresowania, aspiracje krótko- i długookresowe, postawy wobec przemian w otoczeniu, stosunek do pracy, strategie przystosowawcze⁹⁴.

Przeprowadzone w Stanach Zjednoczonych badania nad stylem życia, wartościami i zachowaniem konsumenta (VALS) wyodrębniły następujące segmenty nabywców⁹⁵:

- Zmagający się (*Survivors*) – koncentrują się na zaspokajaniu potrzeb, nie są otwarci na zmiany, liczy się dla nich poczucie bezpieczeństwa i pewności.
- Nowatorzy (*Innovators*) – są to osoby o wysokim poczuciu własnej wartości, są otwarci na zmiany, aktywni, doświadczeni, dążący do osiągnięcia sukcesu, niezależni. Łączą w sobie wszystkie trzy motywacje.
- Ufni (*Believers*) – jest to grupa, dla której ważne są takie wartości jak rodzina, religia, duma narodowa. Są to konserwatyści o niskich dochodach.

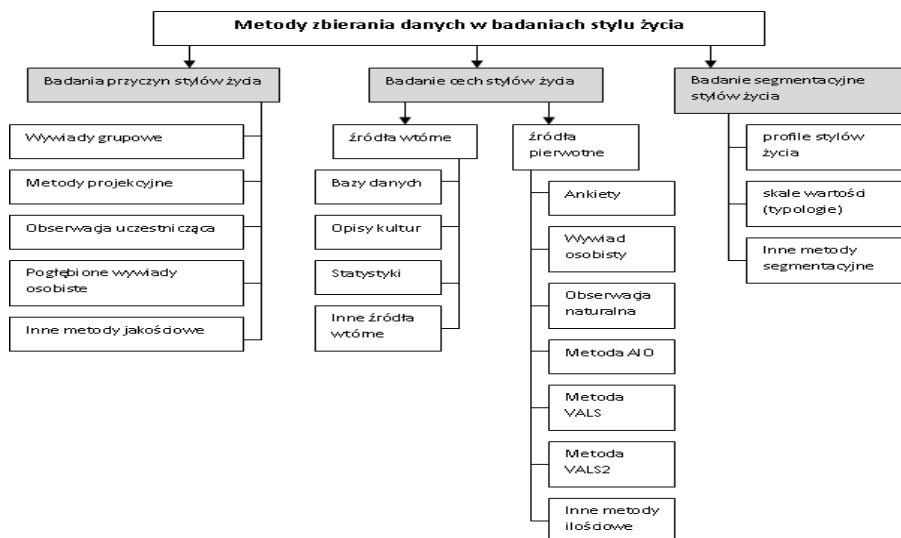
⁹² A. Gerrit, W.F. van Raaij, Zachowanie konsumenta. Podręcznik akademicki, Wydawnictwo Naukowe PWN, Warszawa 2003, s. 186.

⁹³ S. Kaczmarczyk, Zastosowanie badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007, s. 132.

⁹⁴ S. Kaczmarczyk, Zastosowanie badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007, s. 131-132.

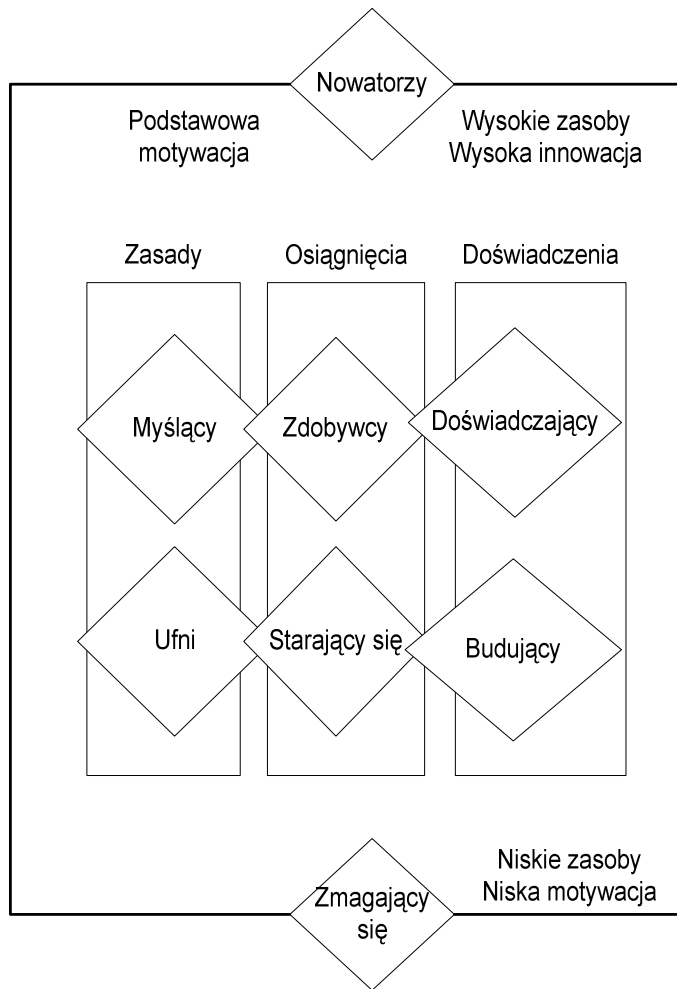
⁹⁵ Por. D. Maison, A. Noga-Bogomilski, Badania marketingowe. Od teorii do praktyki, Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2007, s. 183-184. Wartości i normy stanowią podstawę zachowania nabywcy. Wartości to np.: dokładność, życzliwość, szczęście, sprawiedliwość, bezpieczeństwo. Normy to przekonania na temat tego, co jest dozwolone, a co nie. Por. A. Gerrit, W.F. van Raaij, Zachowanie konsumenta. Podręcznik akademicki, Wydawnictwo Naukowe PWN, Warszawa 2003, s. 386.

- Myślący (*Thinkers*) – dojrzały, zadowoleni, cenią porządek, wiedzę, odpowiedzialność, jako konsumenci są konserwatywni, poszukują produktów dobrej jakości.
- Starający się (*Strivers*) – mniejsze zasoby, rozrywkowi, wrażliwi na modę, zależy im na opinii innych ludzi, dążą do bogactwa, jako konsumenci są aktywni, wybierają produkty markowe.
- Zdobywcy (*Achievers*) – większe dochody, zorientowani na rozwój kariery, mają wyznaczony cel, rodzina ma dla nich dużą wartość, cenią stabilność, przewidywalność, jako konsumenci są aktywni, wybierają sprawdzone produkty.
- Budujący (*Makers*) – niskie dochody, osiągnięcia życiowe wyznaczają poprzez posiadanie domu, dzieci, praktyczni, niepodążający za nowościami, jako konsumenci kupują produkty podstawowe.
- Doświadczający (*Experiencers*) – wyższe dochody, młodzi, entuzjastyczni, otwarci na nowości, podejmujący ryzyko, aktywni, jako konsumenci kupują przede wszystkim odzież.



Rys. 16. Metody zbierania danych w badaniach stylów życia

Źródło: S. Kaczmarczyk, *Zastosowanie badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007, s. 133.



Rys. 17. Segmenty nabywców

Źródło: SRI Consulting Business Experiences. [Online] <http://www.sric-bi.com/VALS/types.shtml> (20.04.2010).

7.3. Zadania do samodzielnego opracowania

Zadanie 1. Zaproponuj kryteria segmentacji:

- a) nabywców zmywarek
- b) nabywców usług z zakresu medycyny estetycznej
- c) nabywców margaryny
- d) nabywców gumy do żucia.

Zadanie 2. Przeprowadź procedurę segmentacji rynku nieruchomości w Opolu w następujących krokach:

- a) określenie pełnego rynku
- b) określenie kryteriów segmentacji
- c) identyfikacja segmentów
- d) charakterystyka odbiorców w segmentach
- e) ocena atrakcyjności i dostępności segmentów
- f) wybór segmentów docelowych.

Zadanie 3. Do podanych poniżej rynków proszę dobrać odpowiednie kryteria segmentacji i wyodrębnić na ich podstawie segmenty.

Rynek	Kryteria	Segmenty
Książek		
Cukru		
Usług turystycznych		
Szamponów		
Usług ubezpieczeniowych		
Leków		
Mebli		
Przelotów pasażerskich		
Kwiatów ciętych		
Samochodów		
Telefonów komórkowych		

Zadanie 4. Dla podanych poniżej produktów proszę przeanalizować wzorce konsumpcji (częstotliwość użycia, lojalność wobec marki, itp.), warunki zakupu (rodzaj sklepu, czas zakupu, wielkość jednorazowego zakupu, częstotliwość zakupu, itd.), oferowane korzyści:

- a) sól
- b) batony czekoladowe
- c) samochód
- d) serwis do kawy
- e) chleb.

8. DOBÓR PRÓBY DO BADAŃ

8.1. Zdefiniowanie populacji

Proces doboru próby składa się następujących etapów: zdefiniowanie badanej zbiorowości (populacji), ustalenie wykazu badanej populacji, wybór metody doboru próby, określenie liczebności próby, pobranie próby.

Populacja generalna (populacja) to zbiór dowolnych elementów, nieidentycznych z punktu widzenia badanej cechy. Może być skończonym lub nieskończonym zbiorem jednostek, które zamierzamy obserwować empirycznie i analizować. **Próba** to podzbiór populacji, podlegający bezpośrednio badaniu ze względu na ustaloną cechę, w celu wyciągnięcia wniosków o kształtowaniu się tej cechy w populacji⁹⁶.

Badać można wszystkie elementy populacji generalnej (badanie kompletne, całkowite, wyczerpujące) albo tylko ich część (badanie częściowe). Badanie kompletne dostarcza pełnej informacji o badanej cesze populacji generalnej. Często jednak jest niecelowe lub niewykonalne (np. podczas badań niszczących), bywa również kosztowne i czasochłonne. Wtedy badaniu poddaje się jedynie próbę.

Jedną z ważniejszych dla rzetelności badań marketingowych jest kwestia odpowiedniego doboru próby. Wykorzystywana w badaniach próba powinna być reprezentatywna, tzn. powinna opisywać strukturę populacji generalnej z przyjętą dokładnością. Reprezentatywność próby ma zatem decydujące znaczenie dla jakości wyciągniętych wniosków⁹⁷. Reprezentatywność próby zależy od sposobu doboru próby oraz od liczebności próby.

8.2. Dobór losowy i dobór nielosowy

Metody **doboru próby** dzielą się na dwie grupy:

- metody oparte na rachunku prawdopodobieństwa (**dobór losowy**)
- metody nie oparte na rachunku prawdopodobieństwa (**dobór nielosowy**).

W doborze losowym można określić prawdopodobieństwo, że dana jednostka lub zespół jednostek znajdzie się w próbie. Określenie operatu losowania to procedura, według której poszczególne jednostki losowania

⁹⁶ Por. Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 193; S. Mynarski, *Praktyczne metody analizy danych rynkowych i marketingowych*, Kantor Wydawniczy Zakamycze, Zakamycze 2000, s. 20-21; S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza FOGRA, Kraków 1995, s. 19-26.

⁹⁷ Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 197.

dostają się do próby. Podstawą operatu losowania jest wykaz wszystkich jednostek reprezentujących populację⁹⁸. Proces wyłonienia próby zwykle rozpoczynamy od zdefiniowania populacji, z której będzie dobierana próba i na którą zostaną uogólnione wnioski z badań.

Badania prowadzone są często w warunkach, w których niemożliwe jest zastosowanie doboru losowego. Przyczyn takiego stanu rzeczy może być wiele (np. niemożność stworzenia operatu losowania). Co więcej, istnieją sytuacje, w których dobór losowy nie wydaje się właściwy, choć jest możliwy do przeprowadzenia. Badacze są często szczególnie zainteresowani badaniem przypadków odbiegających od normy. Ostatecznie to od badacza zależy ocena zalet i wad różnych metod doboru próby w odniesieniu do własnych potrzeb, a co za tym idzie wybór najbardziej adekwatnej metody doboru próby do badań. Metodę doboru próby do badań wybieramy, kierując się m.in. zapewnieniem zgodności kryteriów z celami badania oraz racjonalnością wykorzystania dostępnych środków.

Do najbardziej rozpowszechnionych schematów **losowania** można zaliczyć: **losowanie indywidualne nieograniczone**, **losowanie indywidualne systematyczne**, **losowanie warstwowe**, **wielostopniowe**⁹⁹.

Losowanie **indywidualne nieograniczone** związane jest z ponumerowaniem wszystkich jednostek danej populacji (N) i dokonaniu wyboru losowego pewnej ich liczby (n), najczęściej wykorzystując losowanie za pomocą tablic liczb losowych. Tablice te zbudowane są tak, że liczby czytane w dowolnym porządku (pionowo, poziomo, wprzód, wstecz, całościowo, segmentowo), zachowują zawsze cechy przypadkowości, niezależnie od miejsca ich występowania. Losowanie za pomocą tablic liczb losowych polega na odczytaniu w ustalonej kolejności liczb w nich umieszczonych. Zaczynając zatem od dowolnego miejsca tablicy, wybieramy według ustalonej zasady tyle liczb, ile jednostek ma liczyć dana próba losowa. Pomija się przy tym wszystkie liczby większe od liczebności badanej zbiorowości. Zanotowane liczby z tablic liczb losowych wskazują numery jednostek, które zostały wylosowane do próby.

Losowanie **systematyczne** polega na wyborze z uporządkowanego zbioru odpowiedniej liczby jednostek w równych odstępach (interwałach). Na początku ustala się liczebność całej zbiorowości (N), a następnie liczebność próby (n). Na tej podstawie ustala się interwał losowania

⁹⁸ Por. Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 195; S. Mynarski, *Praktyczne metody analizy danych rynkowych i marketingowych*, Kantor Wydawniczy Zakamycze, Zakamycze 2000, s. 21; S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza FOGRA, Kraków 1995.

⁹⁹ Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 200-205.

$k=N/n$. Poczynając od losowo obranej jednostki pierwszego interwału dobiera się kolejno co „k” jednostek z każdego interwału po jednej jednostce, aż do osiągnięcia pożądanej wielkości próby.

Losowanie **warstwowe** zaleca się w sytuacji, gdy badana zbiorowość wykazuje silne zróżnicowanie pewnej interesującej nas cechy. Metoda ta wymaga podziału całej zbiorowości na warstwy, w obrębie których następuje oddzielny dobór losowy. Losowania dokonuje się zazwyczaj w stosunku wprost proporcjonalnym do liczebności danej warstwy.

Dobór **wielostopniowy** stosuje się, kiedy możliwe jest podzielenie, przy wykorzystaniu określonego kryterium, badanej zbiorowości na kolejne, coraz mniejsze grupy, np. podział kraju na województwa, podział województw na powiaty, podział powiatów na gminy itd. Na początku wybiera się jednostki losowania pierwszego stopnia, następnie jednostki te dzieli się na mniejsze zespoły (jednostki losowania drugiego stopnia) itd., aż ostatecznie dochodzi się do jednostek badania.

Dobór **nielosowy** (subiektywny) uzależniony jest od osoby prowadzącej badanie. W przypadku doboru nielosowego nie ma możliwości określenia prawdopodobieństwa znalezienia się jednostki w próbie. Dobór nielosowy polega na wyborze konkretnych jednostek o ustalonych z góry charakterystykach. W naukach społecznych najczęściej wykorzystuje się dobór **celowy** oraz dobór **kwotowy**. W doborze **celowym** osoba prowadząca badanie wybiera do próby jednostki populacji świadomie (kierując się swoją wiedzą i przekonaniem w wyborze jednostek „typowych”). Dobór wygodny, jednostek typowych oraz metodą kuli śnieżnej to także przykłady doboru nielosowego¹⁰⁰.

W doborze **kwotowym** wybiera się określone segmenty próby, w proporcji odpowiadającej strukturze zbiorowości generalnej. Im więcej wyróżnia się cech zbiorowości, tym uzyskuje się więcej segmentów i tym trudniej skompletować odpowiedni skład zbiorowości próbnej. Dlatego też dla skompletowania składu liczebnego najczęściej poprzestaje się na 2-3 cechach dających nie więcej niż 10 segmentów. Dobór jednostek do próby nie ma charakteru losowego, zwłaszcza w końcowej fazie kompletowania struktury zbioru¹⁰¹. Za punkt wyjścia w konstrukcji próby przyjmuje się założenie, że pod względem cech próba będzie stanowiła miniaturę populacji. Wybór tych cech jest zwykle arbitralny, co stwarza możliwość popełnienia błędów. Badacz musi mieć pewną wiedzę wstępną o populacji,

¹⁰⁰ Por. S. Mynarski, Praktyczne metody analizy danych rynkowych i marketingowych, Kantor Wydawniczy Zakamycze, Zakamycze 2000, s. 32-34; S. Mynarski, Badania rynkowe w warunkach konkurencji, Oficyna Wydawnicza FOGRA, Kraków 1995.

¹⁰¹ Por. Ch. Frankfort-Nachmias, D. Nachmias, Metody badawcze w naukach społecznych, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 198-199.

tzn. musi znać jej strukturę ze względu na wybrane do badania cechy (nie zawsze można tę wiedzę uzyskać, dane statystyczne mogą być przecież nieaktualne, nieadekwatne itd.). Jeśli konstruuje się próbę kwotową w oparciu o kilka cech, badacz musi znać ich rozkład w badanej populacji. Jeżeli zna te rozkłady, to w próbie te proporcje muszą zostać odwzorowane. Każdy ankieter otrzymuje **kwotę**, czyli przydział osób dla wywiadów, które ma przeprowadzić. Ankieter będzie szukał respondentów na zasadzie „chybił – trafił”. Nie występuje tutaj problem odmów – nie są one rejestrowane. I tak na przykład, jeżeli wiadomo, że populacja składa się z kobiet i mężczyzn, a udział kobiet wynosi 51,2%, to do próby kwotowej o liczebności $n=205$ należy wybrać 105 kobiet i 100 mężczyzn.

Dobór jednostek typowych – polega na doborze do próby takich jednostek, które zostały uznane za przeciętne, typowe. W badaniach marketingowych dość często stosuje się także **dobór wygodny**, który obrazują takie przypadki, jak np. ankieta uliczna. O dostępności respondentów decyduje przypadek. **Dobór metodą kuli śnieżnej** stosujemy w wypadku badań, w których populacja badana składa się z jednostek, do których jest bardzo trudno dotrzeć (niemożność ustalenia operatu losowania). Gdyby firma X zamierzała przeprowadzić badanie wśród dealerów narkotyków, to zapewne spróbuje dotrzeć do jednego z nich, a ten może pomoże dotrzeć do kolejnego, itd.

Tabela 8.

Reprezentatywność próby

Próba reprezentatywna	Próba niereprezentatywna
Prosty dobór losowy Dobór systematyczny Dobór warstwowy Dobór wielostopniowy	Dobór celowy Dobór metodą kuli śnieżnej Próba okolicznościowa Próba wygodna (np. w ankiecie ulicznej)
Dobór kwotowy (próba quasi-reprezentatywna)	

Źródło: opracowanie własne.

Żaden ze sposobów doboru próby nie jest doskonały, każdy ma swoje zalety i wady. W przypadku doboru losowego brak wstępnej wiedzy o populacji generalnej nie wpływa negatywnie na reprezentatywność próby, dobór losowy, pozwala skorygować wiedzę o populacji, umożliwia oszacowanie metodami probabilistycznymi stopnia reprezentatywności uzyskanej próby (i to w zakresie wszystkich możliwych cech, ich rozkładów oraz zależności między nimi). Niestety w przypadku doboru losowego mogą wystąpić trudności związane z efektywnym zdefiniowaniem populacji,

zgodnej z celem badania, a co z tym związane – listy obiektów tworzących populację generalną. Próba losowa musi być dostatecznie liczna (co ze względów technicznych lub finansowych może stanowić pewną trudność). Tak więc m.in. wygoda i czynniki ekonomiczne często skłaniają badaczy ku doborowi nielosowemu.

Dobór losowy łączy się zazwyczaj z dużymi kosztami. Nie zawsze niezbędne są próby losowe. Wprawdzie wnioskowanie statystyczne na podstawie prób nielosowych nie jest uzasadnione, niemniej jednak – zwykle sprawdza się w praktyce.

Jeśli wiemy dużo o tym, jakie czynniki wpływają na badane postawy, opinie bądź zachowania, często korzystamy z prób celowych – np. w testach produktów. Tak jest prościej i taniej.

Tabela 9.

Wady i zalety wybranych metod doboru próby

Rodzaj próby	Zalety	Wady
Prosta próba losowa	Możliwość wnioskowania o populacji	Wysokie koszty, zwykle dłuższy czas realizacji
Dobór systematyczny	Możliwość wnioskowania o populacji	Wysokie koszty, zwykle dłuższy czas realizacji
Dobór kwotowy	Oddaje strukturę populacji	Możliwość uogólnień, wiążące się jednak z ryzykiem pominięcia ważnego parametru
Dobór celowy	Tania, mało czasochłonna, Przy odpowiednio określonych kryteriach są pewne podstawy do uogólnień	Brak podstaw teoretycznych do uogólnień
Próba wygodna	Tania, mało czasochłonna	Brak podstaw do uogólnień

Źródło: opracowanie własne.

8.3. Ustalenie wykazu jednostek badanej populacji

Jednym z powszechnie występujących trudności w podjęciu tematyki doboru próby jest zebranie listy obiektów tworzących zbiorowość generalną. Ze względu na brak ogólnie dostępnych i bezpłatnych baz danych, szczególnie jeśli chodzi o nabywców indywidualnych, sytuacja taka jest

kłopotliwa dla badacza. Dodatkowo, gromadzenie i przetwarzanie danych oraz wykorzystanie informacji o nabywcach indywidualnych podlega dość restrykcyjnym obwarowaniom które narzuca Ustawa o ochronie danych osobowych. Pod pojęciem respondentów instytucjonalnych należy rozumieć podmioty będące firmami i jednostkami prowadzącymi działalność gospodarczą, instytucjami administracji rządowej i samorządowej oraz wszelkie inne jednostki, działające zgodnie z zasadą non-profit (do których należą m.in. fundacje i stowarzyszenia). Podmioty wymienione powyżej nie są objęte Ustawą o ochronie danych osobowych, ich wykazy są powszechnie dostępne w książkach telefonicznych, bazach adresowych oraz w Internecie.

Pozyskanie wykazu jednostek tworzących zbiorowość generalną może odbywać się z wykorzystaniem następujących źródeł:

- Klasyczne książki telefoniczne klasyczne (drukowane) – zawierają: nazwę podmiotu, adres, numer telefonu. Informacje te są przydatne dla zgromadzenia bazy danych podmiotów z danego regionu kraju. Niestety nie zawierają dostatecznie szczegółowych danych, które pozwalałyby klasyfikować podmioty.
- Bezpłatne bazy adresowe – np.: Polskie Książki Telefoniczne (<http://www.pkt.pl>), Panorama Firm (<http://www.pf.pl>), Ditel (<http://www.ditel.pl>). Bazy są bezpłatne, jeśli chodzi o korzystanie on-line. Można je również zamówić w postaci danych na CD. Przewagą tego typu baz danych jest możliwość klasyfikowania podmiotów ze względu na branżę oraz lokalizację.
- Płatne bazy adresowe – do najpopularniejszych należą: baza REGON (którą można zakupić np. w Głównym Urzędzie Statystycznym w Warszawie, czy też w Urzędzie Statystycznym w Opolu; <http://www.stat.gov.pl>) oraz bazy firm komercyjnych, które specjalizują się w gromadzeniu danych (np. HBI Polska, która gromadzi szczegółowe dane o średnich i dużych przedsiębiorstwach w Polsce; <http://www.hbi.pl>).

Zgromadzenie wykazu zbiorowości generalnej jest najtrudniejszym i najbardziej kosztownym elementem procesu doboru próby badawczej.

Dobry operat losowania musi spełniać warunki:

- warunek adekwatności – powinien odpowiadać całej populacji, którą chcemy badać;
- warunek kompletności – powinien zawierać wszystkie jednostki wchodzące w skład danej populacji;
- brak powtórzeń – każda jednostka musi w nim wystąpić tylko jeden raz;
- warunek dokładności – nie może zawierać jednostek nie istniejących, bądź też nie należących do populacji;
- warunek techniczny – łatwość korzystania z operatu.

8.4. Określenie liczebności próby

Wyznaczenie niezbędnej liczebności próby należy do podstawowych problemów badawczych. Na tym etapie badań należy odpowiedzieć na pytanie, jak liczna ma być próba, aby na podstawie badań możliwe było wyciągnięcie wniosków o badanej populacji lub/i populacji generalnej.

Można wskazać następujące sposoby określania liczebności próby:

- badacz wybiera próbę na podstawie własnych osądów
- liczebność próby jest określona poprzez minimalne liczby potrzebnych w tablicy kontyngencji (por. testowanie hipotez nieparametrycznych – test niezależności χ^2)
- liczebność próby zostaje ograniczona w związku z kosztami (ograniczenia budżetowe)
- liczebność próby zostaje ustalona na podstawie określonego z góry poziomu precyzji (konstruowanie przedziałów ufności).

Tabela 10.

Określenie liczebności próby nielosowej – wielkość próby stosowana w różnego typu badaniach marketingowych

Rodzaj badania	Typowa wielkość próby (liczba osób)	Minimalna wielkość próby (liczba osób)
Marketingowe badania rynku	1000 – 1500	500
Test rynkowy	300 – 500	200
Test produktu (lub produktów)	200 – 300	200
Test nazwy	200 – 300	100
Test opakowania	200 – 300	100
Test reklamy telewizyjnej	200 – 300	150
Test reklamy radiowej	200 – 300	150
Test reklamy prasowej	200 – 300	150

Źródło: *Podstawy prowadzenia badań marketingowych. Etapy procesu badań marketingowych – teoria procesu badawczego*, CRON, Warszawa 2006, s. 25.

Tabela 11.

Zwyczajowa wielkość próby stosowana w różnego typu badaniach marketingowych

Wymagany podział na podgrupy	Respondenci indywidualni lub gospodarstwa domowe		Firmy	
	Próba ogólnopolska	Próba regionalna, wybrany segment, itp.	Próba ogólnopolska	Próba regionalna, branżowa, itp.
Brak lub nieliczne podgrupy	800-1000	200-400	200-400	50-200
Przeciętny	1001-2000	401-800	401-800	201-600
Liczne podgrupy	>2000	>800	>800	>600

Źródło: <http://www.inquiry.com.pl> (11.03.2010).

8.5. Minimalna liczebność próby losowej

Wzrost liczebności próby losowej polepsza precyzję estymatora, ale jednocześnie zwiększa koszty badania. Minimalną liczebność próby wyznacza się zazwyczaj dla dwóch szacowanych parametrów: średniej oraz frakcji¹⁰².

Symbol α (tzw. poziom istotności), który pojawia się we wzorach, oznacza prawdopodobieństwo popełnienia błędu. Zazwyczaj przyjmuje się $\alpha=0,05$. Prawdopodobieństwo $1-\alpha$ to tzw. poziom ufności.

Minimalną liczebność próby losowej, gdy szacuje się średnią, a populacja jest nieskończona, wyznacza się w oparciu o wzór:

$$n = \frac{u_{\alpha}^2 \sigma^2}{d^2}$$

n – minimalna liczebność próby; u_{α} – wartość odczytana z tablic standaryzowanego rozkładu normalnego dla poziomu istotności α (dla poziomu ufności $1-\alpha$); σ – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd szacunku średniej

¹⁰² Zamiast określenia „frakcja” używane są często inne, takie jak np.: „proporcja”, „wskaźnik struktury”.

Jeśli odchylenie standardowe nie jest znane, to należy je oszacować na podstawie próby wstępnej o liczebności n_0 .

$$s^2 = \frac{1}{n_0 - 1} \sum_{i=1}^{n_0} (x_i - \bar{x})^2$$

Wzór na minimalną liczebność próby ma wówczas postać:

$$n = \frac{t_\alpha^2 s^2}{d^2}$$

n – minimalna liczebność próby; n_0 – liczebność próby wstępnej; t_α – wartość odczytana z tablic rozkładu t-Studenta dla poziomu istotności α (dla poziomu ufności $1-\alpha$) oraz n_0-1 stopni swobody; s – odchylenie standardowe badanej zmiennej w próbie; d – maksymalny dopuszczalny błąd szacunku średniej.

Minimalną liczebność próby losowej, gdy szacuje się frakcję, a populacja jest nieskończona, wyznacza się w oparciu o wzór:

$$n = \frac{u_\alpha^2 pq}{d^2}$$

n – minimalna liczebność próby; p – ocena frakcji populacji posiadającej wyróżnioną cechę; $q=1-p$; u_α – wartość odczytana z tablic standaryzowanego rozkładu normalnego dla poziomu istotności α (dla poziomu ufności $1-\alpha$); s – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd, różnica pomiędzy frakcją próby a frakcją populacji.

Jeśli iloczyn pq należy oszacować na podstawie próby wstępnej o liczebności n_0 , to przy wyznaczaniu minimalnej liczebności próby korzystamy ze wzoru:

$$n = \frac{t_\alpha^2 pq}{d^2}$$

n – minimalna liczebność próby; p – ocena frakcji (w próbie) posiadającej wyróżnioną cechę; $q=1-p$; t_α – wartość odczytana z tablic rozkładu t-Studenta dla poziomu istotności α (dla poziomu ufności $1-\alpha$) oraz n_0-1

stopni swobody; s – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd, różnica pomiędzy frakcją próby a frakcją populacji.

Minimalną licznosc próby losowej, gdy szacuje się średnią, a populacja jest skończona, wyznacza się w oparciu o wzór:

$$n = \frac{\sigma^2}{\frac{d^2}{u_\alpha^2} + \frac{\sigma^2}{N}}$$

N – liczebność populacji generalnej; n – minimalna liczebność próby; u_α – wartość odczytana z tablic standaryzowanego rozkładu normalnego dla poziomu istotności α (dla poziomu ufności $1-\alpha$); σ – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd szacunku średniej.

Jeśli odchylenie standardowe nie jest znane, to należy je oszacować na podstawie próby wstępnej o liczebności n_0 .

$$s^2 = \frac{1}{n_0 - 1} \sum_{i=1}^{n_0} (x_i - \bar{x})^2$$

Wzór na minimalną liczebność próby ma wówczas postać:

$$n = \frac{s^2}{\frac{d^2}{t_\alpha^2} + \frac{s^2}{N}}$$

N – liczebność populacji generalnej; n – minimalna liczebność próby; n_0 – liczebność próby wstępnej; t_α – wartość odczytana z tablic rozkładu t-Studenta dla poziomu istotności α (dla poziomu ufności $1-\alpha$) oraz n_0-1 stopni swobody; s – odchylenie standardowe badanej zmiennej w próbie; d – maksymalny dopuszczalny błąd szacunku średniej.

Minimalną licznosc próby losowej, gdy szacuje się frakcję, a populacja jest skończona, wyznacza się w oparciu o wzór:

$$n = \frac{pq}{\frac{d^2}{u_\alpha^2} + \frac{pq}{N}}$$

N – liczebność populacji generalnej; n – minimalna liczebność próby; p – ocena frakcji populacji posiadającej wyróżnioną cechę; q=1-p; u_α – wartość odczytana z tablic standaryzowanego rozkładu normalnego dla poziomu istotności α (dla poziomu ufności 1- α); s – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd, różnica pomiędzy frakcją próby a frakcją populacji.

Jeśli iloczyn pq należy oszacować na podstawie próby wstępnej o liczebności n_0 , to przy wyznaczaniu minimalnej liczebności próby korzystamy ze wzoru:

$$n = \frac{pq}{\frac{d^2}{t_\alpha^2} + \frac{pq}{N}}$$

N – liczebność populacji generalnej; n – minimalna liczebność próby; p – ocena frakcji (w próbie) posiadającej wyróżnioną cechę; q=1-p; t_α – wartość odczytana z tablic rozkładu t-Studenta dla poziomu istotności α (dla poziomu ufności 1- α) oraz n_0-1 stopni swobody; s – odchylenie standardowe badanej zmiennej w populacji (znane lub szacowane na podstawie wcześniejszych badań); d – maksymalny dopuszczalny błąd, różnica pomiędzy frakcją próby a frakcją populacji¹⁰³.

Uwaga praktyczna: im mniejszy zakładany błąd – tym większa próba, im wyższy poziom ufności – tym większa próba.

8.6. Błędy w badaniach marketingowych

Błędy w badaniu marketingowym są powszechne i występują w każdym badaniu. Dlatego istotne jest zrozumienie źródeł ich powstawania oraz ich wpływu na wynik badania. Rozróżniamy dwa podstawowe rodzaje błędów w badaniach marketingowych: błędy związane z doбором próby oraz błędy niezwiązane z doбором próby. Błędy doboru próby są „różnicą

¹⁰³ Zagadnienie minimalnej liczności próby zostało opracowane na podstawie: S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza FOGRA, Kraków 1995, s. 26-35.

między obserwowanymi wartościami zmiennej i długofalową średnią wartości obserwowanych wyników powtarzanych pomiarów”¹⁰⁴. Podstawową zaletą błędu doboru próby jest to, że może on zostać zmniejszony poprzez zwiększenie wielkości próby.

Błędy w badaniach marketingowych niezwiązane z doбором próby odzwierciedlają wszystkie czynniki, które występują podczas badania marketingowego. W trakcie badania ciężko jest określić wielkość i ukierunkowanie danego błędu. Oto możliwe źródła błędów niezwiązanych z doбором próby: mylnego wnioskowania logicznego, nieprawidłowych założeń projektowych, niepoprawnego zinterpretowania danych, błędnej tabulacji danych, błędnego kodowania, błędnego opracowania wniosków końcowych, błędów obrachunkowych, błędów statystycznych, niewłaściwych metod wywiadu, obserwacji lub pomiaru z wykorzystaniem kwestionariuszy, niewłaściwego dobrania ankietera.

Łączny błąd badania jest sumą błędów z próby oraz błędów z nią niezwiązanych. Łączny błąd badania jest więc odzwierciedleniem z jednej strony błędów związanych z wielkością próby, a z drugiej strony z błędami wynikającymi z błędnego przyjęcia procedur w trakcie zbierania danych i ich przetwarzania.



Rys. 18. Łączny błąd badania

Źródło: opracowanie własne.

Badania przeprowadzone w kierunku zmierzenia wielkości i skali występowania błędów podczas zbierania danych w badaniach marketingowych doprowadziły do sformułowania trzech wniosków:

1. Badania opierały się na faktycznej dostępności do nie wypaczonej informacji, która następnie służyła za sprawdzian wiarygodności.
2. Używały porównań podzielonych serii, w których respondent otrzymywał w różnej formie to samo pytanie, tak, by można było zmierzyć konsekwencje respondenta w udzielaniu odpowiedzi.

¹⁰⁴ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 569.

3. Mierzyły konsekwencje odpowiedzi tego samego respondenta w czasie¹⁰⁵.

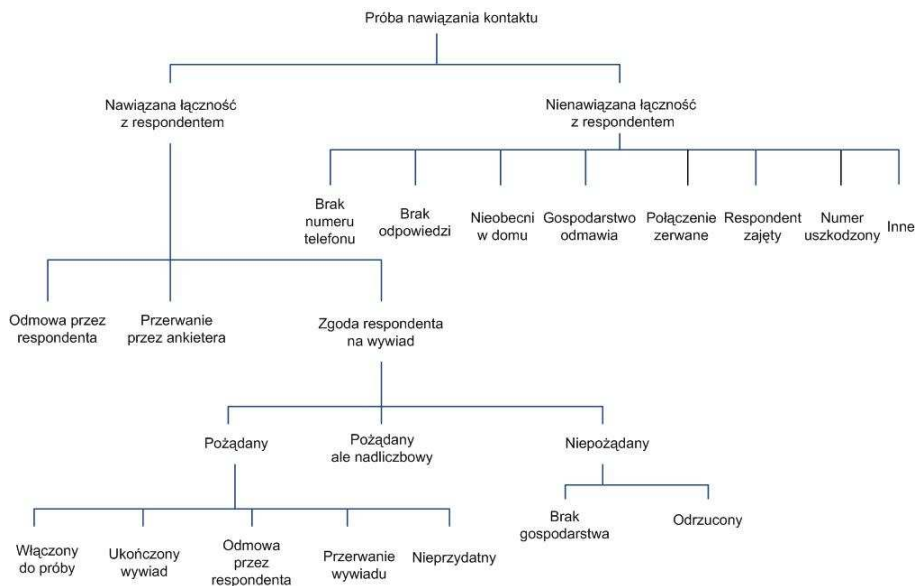
Błąd sformułowania problemu badawczego jest to błąd podejmowany podczas pierwszego etapu procesu badania marketingowego. Niepoprawnie sformułowany problem badawczy rzutuje negatywnie na cały proces badania. Prowadzący badanie jest w stanie popełnić błąd już na etapie formułowania problemu decyzyjnego. Przekształcenie problemu decyzyjnego na problem badawczy także może być obarczone błędem. W efekcie badanie jest przeprowadzone od samego początku źle i nie dostarcza żadnych przydatnych danych.

Nieobjęcie oznacza niewłączenie jakichś jednostek lub całych części zdefiniowanej populacji generalnej do faktycznych operacyjnych ram próby¹⁰⁶. W badaniach marketingowych prowadzący badanie wyznacza pewien obszar, w którym będzie wyznaczał elementy próby. Obszar ten może być wyznaczony zmiennymi, takimi jak np.: granice geograficzne, posiadanie telefonu czy samochodu. W zależności od rodzaju badania obszar ten może się różnić. W badaniach zdarza się, że obszar wyznaczony przez badacza jest nieprawidłowy. Zdarza się to w przypadku posiadania np. nieaktualnego planu miasta lub starej książki telefonicznej. Przy tak dużych niedoskonałościach wysoce prawdopodobne jest dwukrotnie zebranie danych od tych samych elementów lub pominięcie innych. Błąd nieobjęcia może wpływać na wyniki badania w dwojaki sposób: jeśli elementy próby nie są reprezentowane na liście elementów próby oraz jeśli elementy próby są duplikowane. W pierwszym przypadku skutkuje to zerowym prawdopodobieństwem uwzględnienia w badaniu. W drugim przypadku błąd ten prowadzi do błędu nadmiernego pokrycia, który może skutkować pewnym zbliżeniem do siebie wyników. Należy podkreślić, że zarówno wielkość, jak i skutek wystąpienia błędu nieobjęcia nie są proste do zauważenia w trakcie procesu badawczego. Prowadzący badanie nie wie, jak szeroki jest błąd nieobjęcia. Nie posiada także wiedzy na temat możliwości jego zredukowania. Zauważenie i określenie wielkości błędu nieobjęcia jest możliwe jedynie w przypadku istnienia źródła zewnętrznego. Konieczne jest, aby informacja pochodząca z zewnętrznego źródła, była informacją porównywalną.

Błąd braku odpowiedzi jest to błąd wynikający z niezyskania informacji od niektórych elementów populacji badanej.

¹⁰⁵G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 573.

¹⁰⁶G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 574.



Rys. 19. Możliwe przyczyny wystąpienia braku kontaktu z elementami badanymi

Źródło: G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 579.

Błąd braku odpowiedzi jest obliczany przez różne agencje badań marketingowych w odmienny sposób. W wyniku zastosowania różnych obliczeń wyznaczona stopa odpowiedzi waha się od 10 do 90 procent. W tym celu CASRO opracowało jednolitą definicję stopy odpowiedzi:

$$\text{stopa odpowiedzi} = \frac{\text{liczba wywiadów wypełnionych przez odpowiadające jednostki}}{\text{liczba uprawnionych odpowiadających jednostek w próbie}}$$

Warunkiem koniecznym do prawidłowego obliczenia stopy odpowiedzi według podanego wzoru jest prawidłowe określenie osób uprawnionych.

Istnieją dwie główne przyczyny występowania błędu braku odpowiedzi: nieobecność respondenta w domu oraz odmowa udziału w badaniu. Rozmiary, w jakich błąd wynikający z nieobecności respondenta w domu w momencie zgłoszenia się ankietera do badania występuje, jest uzależnione od pory dnia, w której ankieter zgłasza się do badania. Bardziej prawdopodobne jest, iż w godzinach rannych i wczesnopołudniowych, do dyspozycji ankietera będą kobiety zamężne, prawdopodobnie wychowujące dzieci. Kobiety niezamężne i pracujące będą natomiast dostępne w godzinach popołudniowych i wieczornych. Prowadzący badanie układając instrukcje dla ankietera powinien uwzględnić, czy badanie jest przepro-

wadzane w dni powszednie, czy w weekendy, czy jest ono prowadzone w okresie urlopowym, lub w tzw. „długie weekendy”.

Istnieje szereg możliwości zniwelowania zasięgu występowania tego błędu. Ankieter może: umówić się z respondentem telefonicznie, ponowić wizytę, zastosować technikę „wizyty zmodyfikowanej”. Wcześniejsze umówienie się z respondentem telefonicznie jest możliwe w stosunku do osób, które prowadzą dynamiczne życie zawodowe. Ułatwia to zarówno respondentowi, jak i ankieterowi, zaplanowanie dnia oraz wcześniejsze przygotowanie się do wywiadu. Ponowienie wizyty jest skuteczną techniką, przy założeniu, że ponowiona wizyta odbywa się w innych porach dniach niż pierwotna. Natomiast technika „wizyty zmodyfikowanej” polega na pozostawieniu kwestionariusza wywiadu do samodzielnego wypełnienia przez respondenta w przypadku braku jakiegokolwiek z nim kontaktu. Warunkiem koniecznym uzyskania odpowiedzi jest pozostawienie wraz z kwestionariuszem zaadresowanej koperty zwrotnej¹⁰⁷.

Istotnym czynnikiem w występowaniu błędu braku odpowiedzi jest sama postawa ankietera. Dlatego szczególnie ważną rolę spełnia szkolenie ankieterów. Umiejętności ankietera względem planowania i organizacji własnego czasu są nie do przecenienia. Właściwe nastawienie do respondenta oraz umiejętności łatwego nawiązywania kontaktu, pomagają ankieterom w zdobywaniu zaufania respondentów, a co za tym idzie, do lepszego odsetka udzielanych odpowiedzi. Przyjętym wskaźnikiem umożliwiającym porównanie pracy ankieterów jest stopa kontaktu (K)¹⁰⁸.

$$K = \frac{\text{Liczba uprawnionych jednostek próbnych, z którymi nawiązano kontakt}}{\text{Łączna liczba uprawnionych jednostek próbnych, do których się zwracano}}$$

Niska wartość wskaźnika stopy kontaktu, powinna zasygnalizować prowadzącemu badanie istnienie problemu. Nie należy jednak zakładać z wyprzedzeniem, że problem dotyczy bezpośrednio i wyłącznie osoby ankietera. Należy pamiętać, że podczas prowadzenia badania mogą wystąpić czynniki niezależne od niego, takie jak: trudny dostęp do obszaru wyznaczonego do badania, badanie prowadzone jest na terenie miasta, gdzie jest wyższy odsetek osób pracujących.

Błąd braku odpowiedzi z powodu odmowy wzięcia udziału w badaniu występuje w badaniach marketingowych bardzo często. Powodami, dla których udzielana jest odmowa wzięcia udziału w badaniu są; niedogodność, przedmiot badania, niechęć do stosowanych technik sprzedaży, osoba

¹⁰⁷ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 583.

¹⁰⁸ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 584.

ankietera, projekt. Okolicznościami, które mają wpływ na odmowę respondenta, mogą być: okoliczności, w jakich nawiązano kontakt, charakter tematu, osoba ankietera¹⁰⁹.

Można wyróżnić trzy podstawowe techniki zwiększenia stopy odpowiedzi: zwiększenie stopy pierwotnych reakcji, kontynuowanie kontaktów, oszacowanie odmów i skorygowanie wniosków. Zwiększenie stopy pierwotnych reakcji jest działaniem, które nasuwa się wręcz naturalnie. Ankieter może podjąć działania zachęcające respondenta do wzięcia udziału w studium. Jedną z możliwości jest dostarczenie respondentowi szerszej informacji na temat przedmiotu badania oraz znaczenia jego roli przy wyznaczaniu wyników badania. Inną możliwością jest zagwarantowanie respondentowi pełnej anonimowości przy prowadzeniu wywiadu. Istotną barierą może być zamawiający badania – sponsor badania. Wówczas wydaje się być słusznym ukrycie sponsora badania przed respondentem.

Metoda polegająca na kontynuowaniu kontaktów zależy bezpośrednio od charakteru braku reakcji. W większości przypadków, podczas prowadzenia bezpośredniego wywiadu w gospodarstwach domowych, respondenci, którzy pierwotnie nie wzięli udziału w badaniu z powodu np. złego stanu zdrowotnego, nieobecności w pracy lub dużego nakładu obowiązków domowych, w wyniku podjęcia drugiej próby kontaktu, zgadzają się na przeprowadzenie badania. W przypadku, kiedy prowadzimy badanie metodą ankiety pocztowej, sytuacja się komplikuje. Identyfikacja respondentów, którzy odmówili wzięcia udziału w badaniu, nie jest zadaniem prostym i tanim. Ponowne rozesłanie kwestionariuszy nie musi się spotkać z pozytywnym odzewem.

Trzecia strategia polega na modyfikacji wniosków z uwzględnieniem błędu braku odpowiedzi. W przypadku rozwiniętych agencji badań marketingowych, można pozyskać brakujące dane, opierając się na danych zebranych we wcześniejszych badaniach. Oczywiście prowadzący badanie muszą dołożyć staranności podczas analizy porównawczej danych, pod względem kryteriów doboru. Jednym z problemów, na które może natknąć się ankieter, jest brak odpowiedzi jedynie na niektóre punkty ankiety. Może to świadczyć o złej konstrukcji kwestionariusza. Forma, w jakiej zadano pytania, czy też kolejności zadawania pytań jest kwestią istotną przy układaniu kwestionariusza.

Błędy terenowe powstają na etapie uczestnictwa respondenta w badaniu. Respondent udziela odpowiedzi na pytania, lecz odpowiedź nie odzwierciedla jego faktycznych postaw wobec danego zjawiska. Błąd ten występuje także wtedy, gdy respondent odmawia udzielenia odpowiedzi na

¹⁰⁹ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 584.

jedno z pytań w kwestionariuszu. Błędy terenowe określane są często jako błędy odpowiedzi.



Rys. 20. Proces udzielania odpowiedzi przez respondenta

Źródło: opracowanie własne na podstawie G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 593.

Proces udzielenia odpowiedzi przez respondenta składa się z pięciu części (por. rysunek 29). Pierwszy etap składa się ze zrozumienia pytania przez respondenta. Zrozumienie pytania nie zależy tylko i wyłącznie od jego poprawnego sformułowania w kwestionariuszu oraz od postawy ankietera. Zrozumienie pytania zależy także od zdolności psychofizycznych respondenta. Drugi etap, poznawcze przetwarzanie, polega na dokonywanej przez respondenta ocenie odpowiedzi pod względem jej trafności oraz na przypominaniu sobie własnych postaw względem danego zagadnienia. Następnie wstępnie ocenioną informację poddaje się kolejnej ocenie z punktu widzenia ochrony swojego wizerunku, czy też zaspokojenia oczekiwań ankietera. Ostatni etap, czyli podanie odpowiedzi, wiąże się tak samo jak w przypadku etapu pierwszego z możliwościami psychofizycznymi danego respondenta, czyli możliwością i trafnością jego wysławiania się.

Błędy odpowiedzi występują na linii respondent – ankieter i dotyczą trzech głównych obszarów: czynników psychologicznych, czynników behawioralnych, czynników tła¹¹⁰. Czynniki psychologiczne są to postawy, oczekiwania ankietera, które wpływają na udzielane przez respondenta odpowiedzi. Czynniki behawioralne przejawiają się poprzez zachowanie ankietera. Można wyróżnić trzy rodzaje zachowania, które prowadzi do ujawnienia pewnych tendencji, skłonności ankietera do udzielanych odpowiedzi: błędy w zadawaniu pytań i sondowaniu pytaniami uzupełniającymi, błędy w rejestrowaniu odpowiedzi, błędy spowodowane oszustwem¹¹¹. Czynniki tła wpływają na udzielane odpowiedzi poprzez podobieństwa i różnice pomiędzy ankieterem a respondentem. W badaniu możliwe jest uzyskanie lepszej współpracy, jeśli zarówno respondent, jak i ankieter, są pod pewnymi względami do siebie podobni. Ankieter powinien

¹¹⁰ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 593.

¹¹¹ G.A. Churchill, *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002, s. 596.

być dobrany do respondenta pod względem płci, wieku, rasy, dochodu. Im więcej podobieństw, tym większe prawdopodobieństwo uzyskania wymaganych informacji.

Błędy pomiaru wynikają z faktu, że wartość obserwowana lub mierzona odbiega od rzeczywistej wartości elementów próby. Innymi słowy, że wartość x_i dla X_i nie jest wartością prawdziwą. Źródłem tych błędów jest błędne przekształcenie celu badawczego w odpowiadające mu pytania kwestionariuszowe. Błędy pomiaru wynikają ze złego nagrywania, kodowania i przetwarzania odpowiedzi. Tak więc dotyczą one pomiaru na poziomie poszczególnych elementów próby. Przykładem błędu pomiaru może być niedokładne opracowanie ramy. Wprowadzenie do niej mylnych pojęć lub definicji, na podstawie których zbudowany zostaje kwestionariusz. Wyniki badania za pomocą źle skonstruowanego kwestionariusza są niekompletne i niejednoznaczne dla ankierów. W wyniku tego, ankierzy mogą interpretować pytania zawarte w kwestionariuszu na dowolny sposób. Prowadzi to do nieścisłości w zebranych danych. Kolejnym źródłem błędów mogą być niedokładne instrukcje opracowane dla ankierów. Nieprecyzyjność instrukcji pozostawia ankierom dużą swobodę co do sposobu przeprowadzania wywiadu w terenie. Ankierzy sami w sobie mogą być źródłem błędów.

8.7. Przykłady z omówieniem

Zadanie 1. Jak liczną próbę należy wylosować, aby oszacować przeciętny roczny wydatek Polaków na wyjazdy wakacyjne (per capita) przy odchyleniu standardowym 500 zł z dopuszczalnym błędem 50 zł na poziomie ufności 95%.

$$n = \frac{1,96^2 \cdot 500^2}{50^2} = 385$$

Odpowiedź: należy wylosować próbę złożoną z 385 osób.

Przykład 2. Błąd nieobjęcia może wystąpić podczas prowadzenia osobistych wywiadów w centrach handlowych.

Podczas zbierania danych w centrach handlowych zakładamy, że ankier nie dysponuje listą elementów, które mają zostać poddane badaniu. Wywiady są przeprowadzane wśród osób aktualnie przebywających w danym centrum handlowym. Zakładając, że badanie jest prowadzone przez 5 dni w tygodniu w godzinach od 10.00 – 15.00, gromadzenie danych będzie uzależnione od co najmniej kilku czynników: częstości dokonywanych przez respondenta zakupów, posiadania pracy przez respondenta, miejsca zamieszkania. Większe prawdopodobieństwo na objęcie badaniem

posiadają osoby, które nie pracują (np. matki wychowujące dzieci), często robiące zakupy lub mieszkające w pobliżu supermarketu.

8.8. Zadania do samodzielnego opracowania

Zadanie 1. Firma „Źródło” dostarcza wodę źródłaną do zakładów pracy w województwie opolskim. Zamierza przeprowadzić badania wśród przedsiębiorstw z terenu województwa opolskiego. Co może być operatem losowania? Jak firma „Źródło” może pozyskać taki operat?

Zadanie 2. Odpowiedz na pytania:

- Czym charakteryzuje się dobór losowy?
- Czym charakteryzuje się dobór nielosowy?
- Jakie znasz przykłady doboru losowego?
- Jakie znasz przykłady doboru nielosowego?
- Jakie są zalety, a jakie wady doboru losowego.
- Jakie są zalety, a jakie wady doboru nielosowego.

Zadanie 3. Jak liczną próbę należy wylosować, aby oszacować przeciętny roczny wydatek gospodarstwa domowego na remont mieszkania przy odchyleniu standardowym 4 tys. zł z dopuszczalnym błędem 0,5 tys. zł na poziomie ufności 95%.

Zadanie 4. Jak liczną próbę należy wylosować, aby oszacować przeciętny roczny wydatek gospodarstwa domowego na remont mieszkania z dopuszczalnym błędem 0,5 tys. zł na poziomie ufności 95%.

Zadanie 5. Jak liczną próbę należy wylosować, aby oszacować stan wyposażenia gospodarstw domowych w kuchenki mikrofalowe z dopuszczalnym błędem 5% na poziomie ufności 95%. Rozważ różne możliwości (p – pozyskane ze źródeł wtórnych, p – oszacowane na podstawie próby wstępnej).

Zadanie 6. Jak liczną próbę należy wylosować, aby oszacować przeciętny roczny wydatek gospodarstwa domowego studentów PO (przyjmując $N=10\ 000$) na remont mieszkania przy odchyleniu standardowym 4 tys. zł z dopuszczalnym błędem 0,5 tys. zł na poziomie ufności 95%.

Zadanie 7. Jak liczną próbę należy wylosować, aby oszacować przeciętny roczny wydatek gospodarstwa domowego studentów PO (przyjmując $N=10\ 000$) na remont mieszkania z dopuszczalnym błędem 0,5 tys. zł na poziomie ufności 95%.

Zadanie 8. Jak liczną próbę należy wylosować, aby oszacować stan wyposażenia gospodarstw domowych studentów PO (przyjąć $N=10\ 000$) w kuchenki mikrofalowe z dopuszczalnym błędem 5% na poziomie ufności 95%.

Uwaga praktyczna (zadania 3-8): przy wyznaczaniu minimalnej liczności próby można porównać otrzymane wyniki z wynikami uzyskanymi z kalkulatora do obliczania wielkości próby. Można skorzystać z kalkulatora do obliczania wielkości próby zamieszczonego na stronie jednej z firm zajmujących się badaniami np. <http://www.pbssopot.com.pl/x.php?x=139/Dobor-proby.html>

Zadanie 9. Omów potencjalne błędy, które mogą się wiązać z wymienionymi ramami doboru próby. W których przypadkach może wystąpić błąd nieobjęcia?

- a) Książka telefoniczna.
- b) Mapa.
- c) Lista obecnych klientów.

Zadanie 10. Pan Arkadiusz Nowak jest właścicielem salonu optycznego w centrum średniej wielkości miasta. Od jakiegoś czasu rozważa wprowadzenie zmian w oferowanych przez niego usługach. Chciałby zaoferować klientom indywidualne karty magnetyczne, które ułatwią identyfikację klientów, zmianę wyglądu salonu na nowocześniejszy, a także umieszczenie w rogu salonu nowoczesnej poczekalni (sofa, telewizor plazmowy, przy dłuższym oczekiwaniu kawa na koszt salonu). Przed wprowadzeniem tych zmian postanowił rozprzecznić w salonie krótki kwestionariusz, adresowany do co trzeciego klienta wchodzącego do salonu. Pracownicy mieli przeprowadzać krótkie osobiste wywiady z klientami salonu. Pan Nowak dał wyraźne polecenie pracownikom, aby nie obrażali klientów. Zidentyfikuj główne źródła błędów i wyjaśnij ich przyczyny.

9. WSTĘPNE OPRACOWANIE DANYCH

9.1. Podstawowe pojęcia

Zebrany w badaniach kwestionariuszowych materiał empiryczny jest materiałem surowym, wymagającym przetworzenia. Przetwarzanie danych polega na uporządkowaniu i przedstawieniu danych uzyskanych w badaniu marketingowym w postaci opisowej, tabelarycznej lub graficznej. Proces przetwarzania danych składa się z kilku czynności, które muszą być wykonane, aby dane uzyskane w trakcie badania mogły zostać wykorzystane do podjęcia decyzji. Redukcja danych jest procesem przystosowywania surowych danych do ich analizy. W procesie redukcji danych surowych są one kontrolowane, redagowane, wprowadzane do komputera, kodowane, klasyfikowane, zliczane oraz poddawane tabulacji i agregacji. Wynikiem redukcji jest uporządkowanie i wstępna prezentacja danych w postaci opisowej, tabelarycznej i graficznej.

9.2. Kontrola pomiaru

Redukcja danych rozpoczyna się od kontroli pomiaru. Budowa instrumentu pomiarowego pozwala na zidentyfikowanie osoby prowadzącej pomiar, daty, czasu trwania, przedmiotu, miejsca oraz innych elementów tego procesu. Kontrola pomiaru polega m.in. na sprawdzeniu, od ilu jednostek objętych badaniem uzyskano odpowiedzi oraz czy odpowiedzi są kompletne i nie zawierają błędów logicznych. Czynność tę powinien wykonywać prowadzący badanie i w razie możliwości powinien spowodować, aby respondent uzupełnił braki i poprawił błędy.

Kontrola winna uwzględniać:

- kompletność danych (tzw. kontrola ilościowa, formalna)
- poprawność danych (tzw. kontrola jakościowa, merytoryczna).

Szczególną uwagę należy zwrócić na najczęściej występujące błędy systematyczne i przypadkowe.

9.3. Redakcja danych

Redakcja danych polega na sprawdzeniu otrzymanych danych z punktu widzenia ich czytelności i dokładności. Weryfikacja formalna skupia się na sprawdzeniu zebranych danych pod względem czytelności i dokładności. Zdarzają się pytania bez odpowiedzi: niektóre pytania, a nawet całe strony mogą być pominięte przez przeprowadzającego wywiad, czasami respondenci świadomie odmawiają udzielenia odpowiedzi na niektóre pytania bądź w ogóle nie poddają się pomiarowi, a czasami zapisy

odpowiedzi w kwestionariuszu są nieczytelne. Weryfikacja merytoryczna polega na logicznej analizie uzyskanych danych i sprawdzeniu prawdziwości respondentów¹¹².

9.4. Grupowanie

Zebrany i zweryfikowany materiał badawczy musi zostać poddany zabiegowi grupowania. Grupowanie może łączyć się z **klasyfikacją** (zliczanie jednostek o podobnych lub takich samych cechach) lub **typologizacją** (wyszczególnienie zjawisk charakterystycznych, najczęściej występujących). Typologizacja przydatna jest wówczas, gdy badacz spotyka się z niezmiernie niejednorodnym, niemożliwym do jednoznacznej klasyfikacji materiałem.

Klasyfikacja musi spełniać następujące warunki:

- rozłączności (nie może być wątpliwości, do której klasy należy zaliczyć daną jednostkę)
- wyczerpywalności (wszystkie jednostki populacji muszą zostać objęte klasyfikacją, żadna jednostka nie może pozostać poza klasami)
- jednorodności (wyszczególnione klasy zjawisk winny być wewnętrznie maksymalnie jednorodne).

Zabiegi klasyfikacyjne i typologizacyjne znajdują swój wyraz w czynności kodowania.

9.5. Kodowanie

Po zakończeniu procesu redakcji dane są zliczane w ramach poszczególnych klas według różnych kryteriów. Zliczanie danych polega na stwierdzeniu, ilu respondentów wybrało określony wariant cechy.

Kodowanie to proces grupowania i przyporządkowywania symboli różnym danym. Kodowanie odbywa się z zastosowaniem tzw. książki kodowej, która zawiera instrukcje dotyczące postępowania przy kodowaniu.

Książka kodowa zawiera:

- symbole zmiennych (cech);
- typ zmiennych;
- wartość zmiennych (skal);
- nazwy kategorii skali (wartości zmiennej);
- oznaczenia braku danych;
- pełne nazwy zmiennych;
- krótkie instrukcje kodowania.

¹¹² Podstawy prowadzenia badań marketingowych. Etapy procesu badań marketingowych – teoria procesu badawczego, CRON, Warszawa 2006, s. 57-60.

Inaczej koduje się odpowiedzi na pytania zamknięte, a inaczej na pytania otwarte. Kodowanie pytań zamkniętych jest stosunkowo proste i nie wymaga na ogół specjalnych przygotowań. Czasami postępuje się tak, że sposób kodowania pytań przygotowuje się podczas ich budowy. W tym celu każdej odpowiedzi przypisuje się w kwestionariuszu numer (np. tak – 1, nie – 2). Kodowanie staje się wówczas zwykłym powtórzeniem, tzn. przepisaniem numeru odpowiedzi do odpowiedniej rubryki przeznaczonej na symbole kodowe.

9.6. Tabulacja

Tabulacja danych to przedstawienie danych w postaci tabel. Stosowanie tablic (jednodzielnych, dwudzielnych oraz wielodzielnych) ułatwia analizę danych. Tabulacja to zliczanie przypadków należących do różnych kategorii. Tabulacja może przybrać postać tabulacji prostej lub złożonej:

Tabulacja prosta może zostać zastosowana do:

- określenia zakresu braku odpowiedzi na pytanie,
- umiejscowienia błędów;
- umiejscowienia odpowiedzi wyraźnie odbiegających od innych,
- określenia rozkładu empirycznego rozpatrywanej zmiennej,
- obliczenia zbiorowych statystyk.

Tabulacja złożona stanowi najważniejsze narzędzie studiowania związków pomiędzy zmiennymi. Zliczamy przypadki wykazujące dwie wspólne cechy. Tabele są budowane na podstawie odpowiedniej kombinacji cech korelowanych (zestawianych) ze sobą. Jeśli skorelowane są ze sobą dwie cechy, to mamy tabele dwudzielcze.

9.7. Przykłady z omówieniem

Przykład 1. Przeprowadzić procedurę kodowania, przy założeniu, że dysponujemy kwestionariuszem ankiety, w którym pojawiły się m.in. następujące pytania:

1. Czy lubi Pan/Pani słuchać Radia Park?

- | | | |
|--------------------------|----------------|------------------------------------|
| <input type="checkbox"/> | Tak | <i>Proszę przejść do pytania 2</i> |
| <input type="checkbox"/> | Raczej tak | <i>Proszę przejść do pytania 2</i> |
| <input type="checkbox"/> | Raczej nie | <i>Proszę przejść do pytania 3</i> |
| <input type="checkbox"/> | Nie | <i>Proszę przejść do pytania 3</i> |
| <input type="checkbox"/> | Nie wiem | <i>Proszę przejść do pytania 3</i> |
| <input type="checkbox"/> | Nie mam zdania | <i>Proszę przejść do pytania 3</i> |

2. Jeśli w pytaniu 1 padła odpowiedź „tak” lub „raczej tak”, to dlaczego?

.....

10. Co jest dla Pana/Pani ważne przy zakupie sprzętu AGD?

Proszę wybrać trzy odpowiedzi

- Cena
- Marka
- Wygląd
- Wskazania sprzedawcy
- Moje doświadczenie
- Długość okresu gwarancyjnego

11. Co jest dla Pana/Pani ważne przy zakupie samochodu?

Proszę wybrać trzy odpowiedzi spośród podanych i uszeregować je według ważności w ten sposób, że 1 oznacza „najważniejsze”, 2 – „mniej ważne”, 3 – „najmniej ważne”

- Cena
- Marka
- Wygląd
- Parametry samochodu

Tabela 12.

Fragment instrukcji kodowej do pytania zamkniętego

Nr pytania	Treść pytania	Kategorie odpowiedzi	Nr kolumny	Nazwa kolumny	Symbol kodu
1	Czy lubi Pan/Pani słuchać Radia Park?	Tak Raczej tak Raczej nie Nie Nie wiem Nie mam zdania Brak odpowiedzi	1	Pyt_1	1 2 3 4 5 6 0

Źródło: opracowanie własne.

Kolumna „Treść pytania” jest pomocna w instrukcji kodowej, choć nie zawsze występuje w instrukcjach. Czasami podawany jest tylko numer pytania (numer pytania w instrukcji pokrywać się musi z numerem pytania w kwestionariuszu). „Kategorie odpowiedzi” zawierają przepisane z kwestionariusza odpowiedzi do wyboru, zamieszczone pod każdym pytaniem

zamkniętym. Numer kolumny (w niektórych instrukcjach „Numer zmiennej” lub „Nazwa zmiennej”) to kolejny numer informacji, którą kodujemy.

W przypadku kodowania pytań zamkniętych alternatywnych i dysjunktywnych wystarczający jest system wielowartościowy wskazany w przykładzie 1. W przypadku pytań koniunktywnych sytuacja jest nieco bardziej skomplikowana. Jeśli badacz chce uchwycić wszystkie podkreślone przez respondenta odpowiedzi, to może stosować system wielowartościowy z zastosowaniem w bazie danych tylu kolumn, ile odpowiedzi mogli wskazać respondenci.

Tabela 13.

Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego (system wielowartościowy, Pyt_10a, Pyt_10b, Pyt_10c w bazie danych)

Nr pytania	Treść pytania	Kategorie odpowiedzi	Nr kolumny	Nazwa kolumny	Symbol kodu
10	Co jest dla Pana/Pani ważne przy zakupie sprzętu AGD? <i>Proszę wybrać trzy odpowiedzi</i>	Brak odpowiedzi Cena Marka Wygląd Wskazania sprzedawcy Moje doświadczenie Długość okresu gwarancyjnego	15 16 17	Pyt_10a Pyt_10b Pyt_10c	0 1 2 3 4 5 6
Uwaga: Jeżeli respondent zaznaczył tylko jedną odpowiedź, to należy wpisać jej symbol do kolumny Pyt_10a, natomiast w Pyt_10b i Pyt_10c należy wpisać 98. Jeżeli respondent zaznaczył dwie odpowiedzi, to należy wpisać ich symbole do kolumn Pyt_10a i Pyt_10b, natomiast w Pyt_10c należy wpisać 98.					

Źródło: opracowanie własne.

Kodowanie wtórne (przedstawione poniżej w tabeli), którego wynik wpisuje się w pozostawioną do tego celu specjalną kolumnę, daje możliwość skonstruowania typologii przedstawiającej charakter preferowanych przez respondenta odpowiedzi.

Tabela 14.

Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego
(system wielowartościowy, Pyt_10 razem w bazie danych)

Nr pytania	Treść pytania	Kategorie odpowiedzi	Nr kolumny	Nazwa kolumny	Symbol kodu
10	Typologia kombinacji odpowiedzi do pytania: „Co jest dla Pana/Pani ważne przy zakupie sprzętu AGD? <i>Proszę wybrać trzy odpowiedzi</i> ”	odpowiedzi: 1, 2, 3 odpowiedzi: 1, 2, 4 ...	18	Pyt_10razem	1 2 ...

Źródło: opracowanie własne.

Jeden ze sposobów kodowania pytań z rangowaniem uwzględnia, ile razy dana kategoria kafeterii została umieszczona na pierwszym miejscu, ile razy dana kategoria została umieszczona na drugim miejscu, itd. Odpowiedni fragment instrukcji kodowej ilustruje poniższy przykład.

Tabela 15.

Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego
(system wielowartościowy, Pyt_11a, Pyt_11b, Pyt_11c w bazie danych)

Nr pytania	Treść pytania	Kategorie odpowiedzi	Nr kolumny	Nazwa kolumny	Symbol kodu
11	Co jest dla Pana/Pani ważne przy zakupie samochodu? <i>Proszę wybrać trzy odpowiedzi i uszeregować je według ważności w ten sposób, że 1 oznacza „najważniejsze”, 2 – „mniej ważne” 3 – „najmniej ważne”</i>	1. Cena 2. Marka 3. Wygląd 4. Parametry samochodu	19 20 20	Pyt_11a Pyt_11b Pyt_11c	Podać nr odpowiedzi wskazanej na pierwszym miejscu Podać nr odpowiedzi wskazanej na drugim miejscu Podać nr odpowiedzi wskazanej na trzecim miejscu

Źródło: opracowanie własne.

Kodowanie odpowiedzi na pytania otwarte przeprowadzane jest w kilku etapach:

- sporządzenie listy wszystkich odpowiedzi na pytania otwarte;
- uporządkowanie i ostateczne sklasyfikowanie odpowiedzi i sporządzenie listy ogólnych kategorii odpowiedzi powstałych z jednorodnych grup odpowiedzi respondentów według kryterium logiczności, podobnie jak w przypadku skalowania odpowiedzi na pytania zamknięte;
- przyporządkowanie kodów poszczególnym pozycjom listy¹¹³.

Gdy wypowiedzi respondentów są jednowątkowe, a badaczowi zależy przede wszystkim na policzeniu, ile razy w wypowiedziach respondentów pojawił się jeden wątek, to wystarczające będzie zastosowanie systemu wielowartościowego (każdy nowy wątek otrzymuje kolejny numer symbolu kodowego), a na zakodowanie wystarcza jedna kolumna w bazie danych.

Tabela 16.

Fragment instrukcji kodowej do pytania otwartego systemem wielowartościowym z wykorzystaniem jednej kolumny w bazie danych

Nr pytania	Treść pytania	Kategorie odpowiedzi (rodzaj wątku)	Wskaźniki kategorii odpowiedzi	Nr kolumny	Nazwa kolumny	Symbol kodu
2	Jeśli tak lub raczej tak, to dlaczego?	- użyteczność	- są ogłoszenia „zatrudnię pracownika” - podają dużo informacji o Kędzierzynie - dzięki radiu Park znalazły się moje dokumenty	2	Pyt_2	1
		- brak alternatywy	- moje radio odbiera jedynie tę stację - nie odpowiada mi żadna inna stacja			2
		- nie potrafię uzasadnić	- nie potrafię uzasadnić			3
		- brak odpowiedzi	- brak odpowiedzi			0

Źródło: opracowanie własne.

¹¹³ Podstawy prowadzenia badań marketingowych. Etapy procesu badań marketingowych – teoria procesu badawczego, CRON, Warszawa 2006, s. 60-61.

Przykład 2. Zapisywanie kodowanych danych w kolumnach arkusza kalkulacyjnego:

- Jednej skali nominalnej – przypisujemy jedną kolumnę w przypadku skali dysjunkcyjnej, a tyle kolumn, ile jest pozycji skali – w przypadku skali koniunkcyjnej.
- Każdej skali porządkowej przypisuje się maksymalnie tyle kolumn, ile jest porządkowanych obiektów (cech produktu, marek, firm). Kolumny odpowiadają rangowanym obiektom, lub kolejności rangowej obiektów.
- Każdej skali przedziałowej odpowiada jedna kolumna (w której wpisuje się symbole cyfrowe odpowiadające poziomom ocen).
- Każdej skali ilorazowej przypisuje się jedną kolumnę Wpisuje się rzeczywiste, niekodowane wartości cechy.

Tabela 17.

Zakodowane i zapisane w arkuszu kalkulacyjnym odpowiedzi

Nr	Pyt_1	Pyt_2a	Pyt_2b	Pyt_2c	Pyt_3	Pyt_4
1	1	1	2	3	21	18
2	2	1	3	2	21	33
3	2	2	1	3	23	74
...

Źródło: opracowanie własne, dane umowne

Objaśnienia do tabeli:

Kolumna 1 (Pyt_1) – „Płeć” (skala nominalna, 1 – kobieta, 2 – mężczyzna)

Kolumna 2-4 (Pyt_2a, Pyt_2b, Pyt_2c) – Pytanie: „Uszereguj marki kawy X, Y, Z w kolejności preferencji” (skala porządkowa, rezerwujemy trzy kolumny, kolumna 2 - „najbardziej preferowana”, ..., kolumna 4 – „najmniej preferowana”)

Kolumna 5 (Pyt_3) – Pytanie „Jaka temperatura powietrza (podaj w °C)

w samochodzie odpowiada Panu/Pani najbardziej?” (skala przedziałowa)

Kolumna 6 (Pyt_4) – Pytanie „Ile masz lat?” (skala ilorazowa)

Przykład 3.

Tabela 18.

Przykład tabulacji prostej – miejsca sprzedaży produktów rolniczych z wykorzystaniem bezpośredniego kanału zbytu (n=200)

Wyszczególnienie	Liczba odpowiedzi	Udział (%)
Sprzedaż na targowisku rolnym	148	74
Sprzedaż na poboczu drogi	66	33
Prowadzenie sklepu lub straganu w gospodarstwie	44	22
Sprzedaż obwoźna	12	6
Inne	10	5

Źródło: opracowanie własne, dane umowne

Tabela 19.

Przykład tabulacji złożonej – dane surowe – opinie respondenta o produkcie (n=280)

Wiek respondenta	Opinie o produkcie		Suma
	pozytywne	negatywne	
Poniżej 18 lat	100	10	110
18-30 lat	50	10	60
31-40 lat	30	10	40
41-50 lat	30	20	50
51 lat i więcej	10	10	20
Suma	220	60	280

Źródło: opracowanie własne, dane umowne

Na podstawie tabeli z danymi surowymi, w zależności od potrzeby, można obliczyć odpowiednie proporcje obserwacji względem sum w każdym wierszu, w każdej kolumnie lub w stosunku do wszystkich obserwacji.

Tabela 20.

Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280)

Wiek respondenta	Opinie o produkcie	
	pozytywne	negatywne
Ponizej 18 lat	35,7%	3,6%
18-30 lat	17,9%	3,6%
31-40 lat	10,7%	3,6%
41-50 lat	10,7%	7,1%
51 lat i więcej	3,6%	3,6%

Źródło: opracowanie własne, dane umowne

Tabela 21.

Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280)

Wiek respondenta	Opinie o produkcie		Suma
	pozytywne	negatywne	
Ponizej 18 lat	90,9%	9,1%	100,0%
18-30 lat	83,3%	16,7%	100,0%
31-40 lat	75,0%	25,0%	100,0%
41-50 lat	60,0%	40,0%	100,0%
51 lat i więcej	50,0%	50,0%	100,0%

Źródło: opracowanie własne, dane umowne

Tabela 22.

Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280)

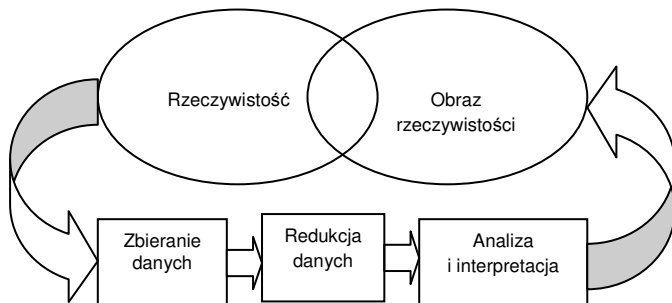
Wiek respondenta	Opinie o produkcie	
	pozytywne	negatywne
Ponizej 18 lat	45,5%	16,7%
18-30 lat	22,7%	16,7%
31-40 lat	13,6%	16,7%
41-50 lat	13,6%	33,3%
51 lat i więcej	4,5%	16,7%
Suma	100,0%	100,0%

Źródło: opracowanie własne, dane umowne

10. PREZENTACJA DANYCH

10.1. Wprowadzenie

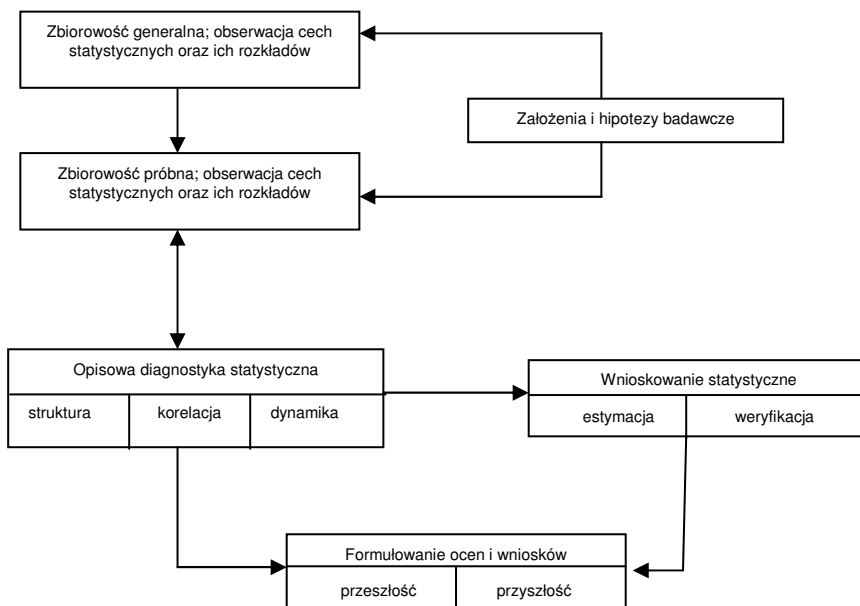
Proces zbierania i analizy danych nigdy nie daje idealnego obrazu rzeczywistości (rys. 21).



Rys. 21. Odniesienie wyników badań marketingowych do rzeczywistości

Źródło: opracowanie własne.

Zakres analizy danych obrazuje poniższy rysunek.



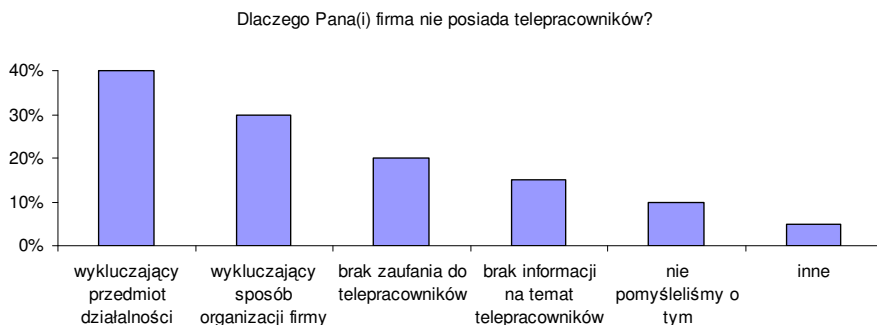
Rys. 22. Zakres analizy statystycznej

Źródło: opracowanie własne na podstawie A. Luszniwicz, T. Słaby, *Statystyka stosowana*, PWE, Warszawa 1997, s. 25.

10.2. Prezentacja danych na wykresie

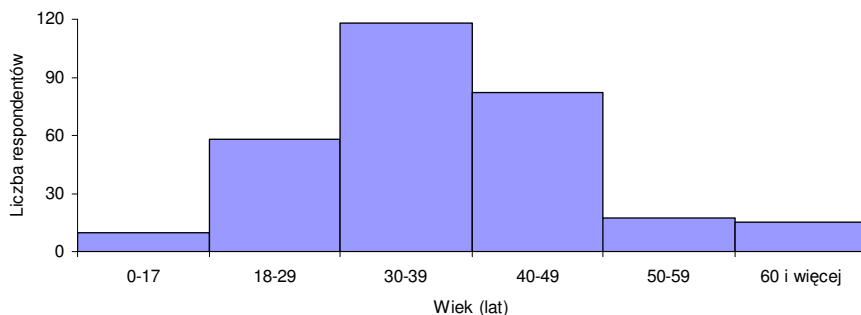
Jedną z bardziej efektywnych metod prezentacji danych jest przedstawienie danych na wykresie. Niektórzy praktycy badań marketingowych oraz wszelakich analiz stosują dla poparcia tej tezy stwierdzenie, iż dobry wykres wart jest więcej niż 100 słów. Najczęściej w opracowaniach pojawiają się histogramy, wykresy słupkowe, wykresy kołowe oraz wykresy rozrzutu.

Wykres 1. Przykład prezentacji danych na wykresie (wykres słupkowy)



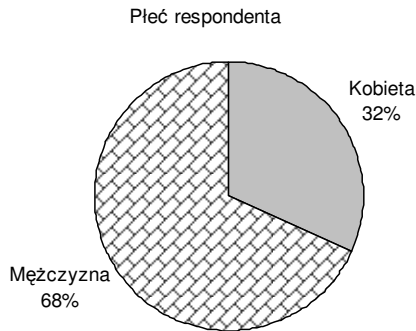
Źródło: dane umowne, opracowanie własne.

Wykres 2. Przykład prezentacji danych na wykresie (histogram)



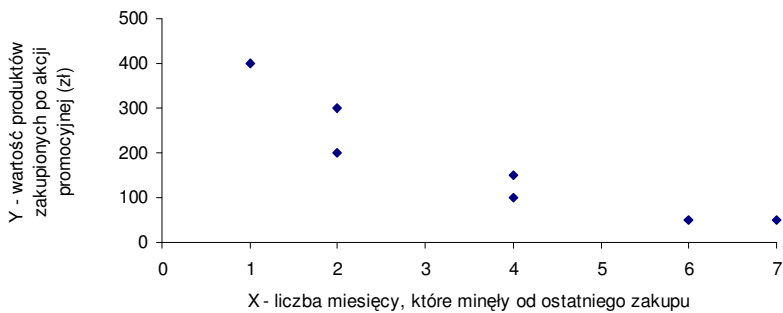
Źródło: dane umowne, opracowanie własne.

Wykres 3. Przykład prezentacji danych na wykresie (wykres kołowy)



Źródło: dane umowne, opracowanie własne.

Wykres 4. Przykład prezentacji danych na wykresie (wykres rozrzutu)



Źródło: dane umowne, opracowanie własne.

10.3. Podstawowe charakterystyki liczbowe

Do analizy danych wykorzystuje się m.in. różne charakterystyki liczbowe, które pozwalają na wykrycie prawidłowości oraz wyciągnięcie wniosków. Najczęściej wykorzystywany jest **wskaźnik struktury**, który obrazuje, jaką część całej zbiorowości objętej badaniem posiada dany wariant cechy. Inną często wykorzystywaną miarą jest **średnia arytmetyczna**, która obrazuje przeciętny poziom zjawiska. Kolejnymi miernikami statystycznymi wykorzystywanymi w badaniach marketingowych są **dominanta** i **kwartyle**. **Dominanta** pozwala na wskazanie tej wartości cechy, która występuje najczęściej w badanej zbiorowości. **Mediana (kwartył**

drugi) dzieli zbiorowość prezentowaną w postaci szeregu na dwie liczebnie równe części w ten sposób, że połowa jednostek ma wartości cechy mniejsze lub równe medianie. **Kwartył dolny** dzieli zbiorowość na dwie części w ten sposób, że 25% jednostek zbiorowości ma wartości cechy niższe bądź równe kwartyłowi pierwszemu, a 75% równe bądź wyższe od tego kwartyła. **Kwartył górny** dzieli zbiorowość na dwie części w ten sposób, że 75% jednostek zbiorowości ma wartości cechy niższe bądź równe kwartyłowi trzeciemu, a 25% równe bądź wyższe od tego kwartyła.

Zasadnicze znaczenie w analizie danych ma odpowiednie rozpoznanie poziomu pomiaru (skali w jakiej dokonywany jest pomiar). Celem poprawnego prowadzenia analizy danych konieczne jest posiadanie wiedzy odnośnie stosowanych w badaniach marketingowych skal pomiarowych. Stosowane wobec różnych skal pomiarowych charakterystyki liczbowe wykorzystywane w celu analizy danych przedstawia poniższe zestawienie (tabela 23).

Tabela 23.

Charakterystyki liczbowe stosowane w analizie danych w zależności od użytej skali pomiarowej

Skala charakterystyki	Nominalna	Porządkowa	Interwałowa (przedziałowa)	Ilorazowa (stosunkowa)
wskaźnik struktury dominanta (moda)	tak	tak	tak	tak
mediana kwartyle	nie	tak	tak	tak
średnia	nie	nie	tak	tak

Źródło: opracowanie własne.

Wzory dla szeregu szczegółowego:

✓ Średnia arytmetyczna

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

✓ Mediana

$$Me = \begin{cases} \frac{x_{n+1}}{2} & \text{gdy } n \text{ jest nieparzyste} \\ \frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{gdy } n \text{ jest parzyste} \end{cases}$$

- ✓ W przypadku szeregów szczegółowych kwartyle pierwszy i trzeci wyznacza się, korzystając ze wzoru na medianę. Zbiorowość dzieli się na dwie równe części (pierwszą – której jednostki przyjmują wartości nie większe od mediany, drugą – złożoną z pozostałych jednostek). Dla każdej z tych części można wyznaczyć medianę. Dla pierwszej części wartość jej mediany odpowiada kwartyłowi dolnemu, dla drugiej części – kwartyłowi górnemu.

- ✓ Rozstęp

$$R = x_{\max} - x_{\min}$$

- ✓ Wariancja

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

- ✓ Odchylenie standardowe

$$s = \sqrt{s^2}$$

- ✓ Typowy klasyczny obszar zmienności cechy

$$\bar{x} - s < x < \bar{x} + s$$

- ✓ Współczynnik zmienności

$$V = \frac{s}{\bar{x}}$$

- ✓ Odchylenie ćwiartkowe

$$Q = \frac{Q_3 - Q_1}{2}$$

- ✓ Typowy pozycyjny obszar zmienności cechy

$$Me - Q < x < Me + Q$$

Wzory dla szeregu rozdzielczego punktowego:

- ✓ Średnia arytmetyczna ważona

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i n_i$$

- ✓ Kwartył dolny, mediana, kwartył górny, odchylenie ćwiartkowe, moda – analogicznie jak w przypadku szeregu szczegółowego.

✓ Wariancja

$$s^2 = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x})^2 n_i$$

✓ Odchylenie standardowe, rozstęp, klasyczny i pozycyjny obszar zmienności – analogicznie jak w przypadku szeregu szczegółowego.

Wzory dla szeregu rozdzielczego przedziałowego:

✓ Średnia arytmetyczna ważona

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k x_i' n_i$$

✓ Kwartył dolny

$$Q_1 = x_{Q_1} + \frac{\frac{n}{4} - \sum_{i=1}^{k-1} n_i}{n_{Q_1}} h$$

gdzie

x_{Q_1} – dolna granica przedziału zawierającego kwartył dolny, h – rozpiętość przedziału klasowego,

n_{Q_1} – liczebność klasy zawierającej kwartył dolny,

$\sum_{i=1}^{k-1} n_i$ – suma liczebności klas od pierwszej do tej, która bezpośrednio poprzedza klasę zawierającą kwartył dolny

✓ Mediana

$$Me = x_{Me} + \frac{\frac{n}{2} - \sum_{i=1}^{k-1} n_i}{n_{Me}} h$$

gdzie

x_{Me} – dolna granica przedziału zawierającego medianę, h – rozpiętość przedziału klasowego,

n_{Me} – liczebność klasy zawierającej medianę,

$\sum_{i=1}^{k-1} n_i$ – suma liczebności klas od pierwszej do tej, która bezpośrednio poprzedza klasę zawierającą medianę.

✓ Kwartył górny

$$Q_3 = x_{Q_3} + \frac{\frac{3n}{4} - \sum_{i=1}^{k-1} n_i}{n_{Q_3}} h$$

gdzie

x_{Q3} – dolna granica przedziału zawierającego kwartył górny, h – rozpiętość przedziału klasowego,

n_{Q3} – liczebność klasy zawierającej kwartył górny,

$\sum_{i=1}^{k-1} n_i$ – suma liczebności klas od pierwszej do tej, która bezpośrednio poprzedza klasę zawierającą kwartył górny.

✓ Moda

$$Mo = x_{Mo} + \frac{n_{Mo} - n_{Mo-1}}{(n_{Mo} - n_{Mo-1}) + (n_{Mo} - n_{Mo+1})} h$$

gdzie

x_{Mo} – dolna granica przedziału zawierającego modę, h – rozpiętość przedziału klasowego,

n_{Mo} – liczebność klasy zawierającej modę, n_{Mo-1} – liczebność klasy poprzedzającej klasę zawierającą modę, n_{Mo+1} – liczebność klasy następującej po klasie zawierającej modę.

✓ Wariancja

$$s^2 = \frac{1}{n} \sum_{i=1}^k (x'_i - \bar{x})^2 n_i$$

✓ Odchylenie standardowe, rozstęp, klasyczny i pozycyjny obszar zmienności – analogicznie jak w przypadku szeregu szczegółowego.

10.4. Przykłady z omówieniem

Przykład 1. Przykład zastosowań wskaźników struktury – badania wydatków na reklamę przeprowadzane przez AGB Nielsen

Dane z okresu 8.12.2008 - 14.12.2008; pomiar przeprowadzono dla grupy celowej: wszystkie osoby w wieku 4+ (uniwers¹¹⁴: 35 886 000, średnia wielkość próby: 3906).

Przeciętny Polak w wieku 4+ oglądał telewizję średnio 254 minuty dziennie.

¹¹⁴ Uniwersum = populacja

Tabela 24.

TOP 20 Tygodniowe wydatki na reklamę w podziale na stacje

	Stacja	Wydatki	Udział
1	TVN	68 155 090,00 zł	26,06 %
2	Polsat	42 140 065,00 zł	16,11 %
3	TVP1	38 616 245,00 zł	14,77 %
4	TVP2	25 005 825,00 zł	9,56 %
5	Jetix	6 713 255,00 zł	2,57 %
6	PULS	6 116 400,00 zł	2,34 %
7	AXN	6 026 865,00 zł	2,30 %
8	TV4	5 375 400,00 zł	2,06 %
9	Cartoon Network/TCM	4 998 700,00 zł	1,91 %
10	MiniMini	4 617 055,00 zł	1,77 %
11	TVN7	3 947 550,00 zł	1,51 %
12	TVN24	3 412 029,50 zł	1,30 %
13	TVP3*	3 276 215,00 zł	1,25 %
14	MTV Polska	3 188 730,00 zł	1,22 %
15	Viva Polska	3 006 940,00 zł	1,15 %
16	Discovery	3 003 380,00 zł	1,15 %
17	Hallmark	2 437 595,00 zł	0,93 %
18	National Geographic	2 294 290,00 zł	0,88 %
19	Polsat2	2 209 245,00 zł	0,84 %
20	Tele5	1 787 390,00 zł	0,68 %
	RAZEM	261 532 001,75 zł	

Wydatki podano według oficjalnych cenników stacji TV.

* - wydatki obejmują jedynie wspólne pasmo stacji regionalnych

Źródło: <http://www.agbnelsen.pl/2008-12-14,1310.html#statinv>

Przykład 2. Na podstawie danych zawartych w tabeli 25:

- Zbudować szereg szczegółowy, wyznaczyć (korzystając ze wzorów): średnią, odchylenie standardowe, współczynnik zmienności, modę, medianę, kwartył dolny i kwartył górny; przedstawić interpretacje otrzymanych wyników; naszkicować i omówić wykres „pudełko z wąsami”, odszukać obserwacje nietypowe
- Zbudować szereg punktowy, sporządzić rozkłady liczości, liczości skumulowanych i względnych; wyznaczyć (korzystając ze wzorów): średnią, odchylenie standardowe, współczynnik zmienności, modę, medianę, kwartył dolny i kwartył górny. Przedstawić interpretacje otrzymanych wyników.
- Zbudować szereg rozdzielczy przedziałowy, wyznaczyć (korzystając ze wzorów): średnią, odchylenie standardowe, współczynnik zmienności, modę (o ile to możliwe), medianę, kwartył dolny i kwartył górny. Przedstawić interpretacje otrzymanych wyników. Sporządzić histogram liczości.

Tabela 25.

Miesięczne płace brutto (w zł) uzyskane w badaniu pracowników na stanowiskach robotniczych w przedsiębiorstwie X (n=35)

900	1200	1000	900	1000	1100	900
1300	900	1200	1300	1100	1100	1100
2000	1500	900	1300	1100	1100	900
1000	1800	1000	1300	900	1000	1200
1900	1700	1600	900	1100	1100	900

Źródło: dane umowne.

Rozwiązanie

- Szereg szczegółowy: 900, 900, 900, 900, 900, 900, 900, 900, 900, 900, 1000, 1000, 1000, 1000, 1000, 1100, 1100, 1100, 1100, 1100, 1100, 1100, 1100, 1100, 1200, 1200, 1200, 1300, 1300, 1300, 1300, 1500, 1600, 1700, 1800, 1900, 2000.

Wyniki obliczeń na podstawie wzorów:

$$\bar{x} = \frac{41\,200}{35} = 1177,14$$

$$s = \sqrt{\frac{3\,101\,714,29}{35}} = 297,69$$

$$v = 0,25$$

$$M_o = 900$$

$$Q_1 = 950$$

$$M_e = 1100$$

$$Q_3 = 1300$$

Interpretacje wyników:

Przeciętna płaca brutto na stanowisku robotniczym w przedsiębiorstwie X wynosi 1 177,14 zł. Płace brutto różnią się przeciętnie od średniej płacy o $\pm 297,69$ zł. Grupa nie jest jednorodna, płace są dość zróżnicowane, ponieważ współczynnik zmienności wynosi aż 25%. Najczęściej na stanowisku robotniczym płaca brutto wynosi 900 zł. 25% badanych otrzymuje płacę 950 zł lub niższą. Połowa badanych otrzymuje płacę 1100 zł lub niższą. 75% badanych otrzymuje płacę 1300 zł lub niższą.

Obliczenia pomocnicze do wykresu pudełko z wąsami

$$Q_1=950$$

$$M_e=1100$$

$$Q_3=1300$$

$$Q=175$$

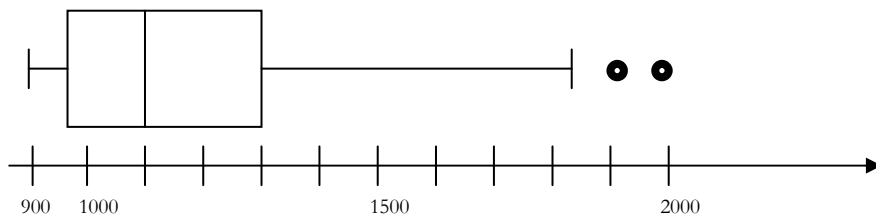
$$X_{\min}=900$$

$$Q_1-3Q=425$$

$$X_{\max}=2000$$

$$Q_3+3Q=1825$$

Wykres 5. Wykres pudełko z wąsami.



Źródło: opracowanie własne.

Interpretacja wykresu:

Widoczna jest asymetria prawostronna oraz dwie obserwacje odstające (płace 1 900 zł i 2 000 zł). Najniższa płaca to 900 zł. 25% badanych otrzymuje płacę 950 zł lub niższą. Połowa badanych otrzymuje płacę 1 100 zł lub niższą. 75% badanych otrzymuje płacę 1 300 zł lub niższą. Najwyższa płaca to 2 000 zł. Większość badanych otrzymuje płacę poniżej średniej (→asymetria prawostronna).

Ten przykład obrazuje przypadek, kiedy policzenie wyłącznie wartości średniej arytmetycznej może wprowadzać w błąd, ponieważ największe liczebności skupiają się wokół najniższych wartości cechy (900 zł). Dwie obserwacje odstające przyczyniają się do tego, że średnia wynosi 1 177 zł. Wyznaczenie charakterystyk takich jak moda i kwartale wydaje się w tym przypadku niezwykle ważne dla uzupełnienia opisu zbiorowości.

b) Szereg rozdzielczy punktowy

i	Płaca brutto (w zł)	Liczebności (n_i)	Liczebności skumulowane ($n_{sk\ i}$)	Liczebności względne (w_i)	Obliczenia do średniej	Obliczenia do odchylenia standardowego
1	900	9	9	0,26	8 100	691 273,47
2	1 000	5	14	0,14	5 000	156 897,96
3	1 100	8	22	0,23	8 800	47 608,16
4	1 200	3	25	0,09	3 600	1 567,35
5	1 300	4	29	0,11	5 200	60 375,51
6	1 500	1	30	0,03	1 500	104 236,73
7	1 600	1	31	0,03	1 600	178 808,16
8	1 700	1	32	0,03	1 700	273 379,59
9	1 800	1	33	0,03	1 800	387 951,02
10	1 900	1	34	0,03	1 900	522 522,45
11	2 000	1	35	0,03	2 000	677 093,88
suma		35		1,00	41 200	3 101 714,29

Oczywiście należy się spodziewać takich wyników (i ich interpretacji), jakie wyszły w przykładzie a.

c) Szereg rozdzielczy przedziałowy

Ustalenie liczby klas k

$$k \approx \sqrt{35}$$

Przyjmijmy $k=6$

Ustalenie rozpiętości przedziałów klasowych h

$$h \approx \frac{2000 - 900}{6}$$

Przyjmijmy $h=185$

Ustalenie granic poszczególnych klas

i	Płaca brutto (w zł)	Środek przedziału (x_i')	Liczności (n_i)	Liczności skumulowane ($n_{sk\ i}$)	Liczności względne (w_i)	Obliczenia do średniej	Obliczenia do odchylenia standardowego
1	900-1085	992,5	14	14	0,40	13 895	723 223,14
2	1085-1270	1 177,5	11	25	0,31	12 953	19 668,90
3	1270-1455	1 362,5	4	29	0,11	5 450	81 469,47
4	1455-1640	1 547,5	2	31	0,06	3 095	214 793,31
5	1640-1825	1 732,5	2	33	0,06	3 465	525 751,88
6	1825-2010	1 917,5	2	35	0,06	3 835	973 610,45
suma			35		1,00	42 693	2 538 517,00

Wyniki obliczeń na podstawie wzorów:

$$\bar{x} = \frac{42\ 693}{35} = 1\ 220$$

$$s = \sqrt{\frac{2\ 538\ 517}{35}} = 269$$

$$v = 0,22$$

Mo – brak

$$Q_1 = 1\ 016$$

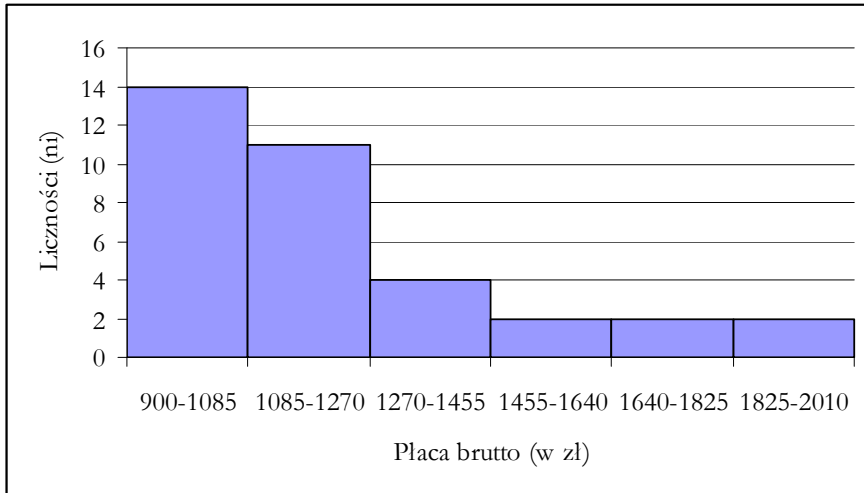
$$Me = 1\ 144$$

$$Q_3 = 1\ 328$$

Interpretacje wyników:

Przeciętna płaca brutto na stanowisku robotniczym w przedsiębiorstwie X wynosi 1 220 zł. Płace brutto różnią się przeciętnie od średniej płacy o ± 269 zł. Grupa nie jest jednorodna, płace są dość zróżnicowane, ponieważ współczynnik zmienności wynosi aż 22%. Nie można obliczyć mody, ponieważ najliczniejsza jest pierwsza klasa. 25% badanych otrzymuje płacę 1 016 zł lub niższą. Połowa badanych otrzymuje płacę 1 144 zł lub niższą. 75% badanych otrzymuje płacę 1328 zł lub niższą.

Wykres 6. Histogram licznosci



Źródło: opracowanie własne.

Przykład 3. Poniżej zamieszczono wyniki badań stopnia zadowolenia pasażerów z funkcjonowania komunikacji zbiorowej w Krakowie¹¹⁵. Respondenci udzielali odpowiedzi na pytanie o ogólny stopień zadowolenia z funkcjonującej komunikacji zbiorowej w zakresie oferty przewozowej MPK SA w Krakowie. Stopień zadowolenia wyrażali, mając do dyspozycji przedział 0-100%, gdzie 0% oznaczało zupełny brak zadowolenia, natomiast 100% – całkowite zadowolenie. Respondenci mieli do dyspozycji pełne „dziesiątki”, skalę potraktowano ja przedziałową.

¹¹⁵ A. Ciastoń, Statystyczne porównanie wyników badań stopnia zadowolenia pasażerów z funkcjonowania komunikacji zbiorowej w Krakowie; http://www.transport.wshe.lodz.pl/documents/statystyczne_porownanie.pdf

Tabela 26.

Wyznaczone charakterystyki liczbowe

	n	\bar{x}	Mo	Me	Q ₁	Q ₃	s	v
Badanie 1 (2001 r.)	449	55,94	62,56	57,15	44,62	68,32	16,52	0,295
Badanie 2 (2001 r.)	482	58,80	67,18	62,52	46,18	74,23	21,45	0,365
Badanie 3 (2002 r.)	540	60,69	56,50	61,67	46,99	76,45	20,66	0,340
Badanie 4 (2005 r.)	500	57,40	57,18	57,50	49,75	66,41	16,04	0,297

Źródło: A. Ciastoń, Statystyczne porównanie wyników badań stopnia zadowolenia pasażerów z funkcjonowania komunikacji zbiorowej w Krakowie;
http://www.transport.wshe.lodz.pl/documents/statystyczne_porownanie.pdf (14.12.2008)

gdzie:

n – liczność próby, \bar{x} – średnia z próby, Mo – moda (dominanta),
 Me – mediana, Q₁ – kwartył dolny, Q₃ – kwartył górny, s – odchylenie standardowe, v – współczynnik zmienności

10.5. Zadania do samodzielnego opracowania

Zadanie 1. Pan Kowalski jest cukiernikiem. Pan Kowalski wykonuje na zamówienie m.in. torty. Tabela przedstawia wagę zamówionych w ciągu ostatnich 6 miesięcy tortów.

Waga tortu (kg)	Liczba tortów (n _i)
0-1	60
1-2	80
2-3	54
3-4	6

a) Sporządź histogram licznosci.

b) Uzupełnij następujące zdania:

- Najczęściej zamawiany tort waży ... kg.
- Przeciętnie zamawiany tort waży ... kg.
- Przeciętne odchylenie wagi tortu od średniej wynosi ... kg.
- 25% zamawianych tortów waży nie więcej niż ... kg.
- Połowa zamawianych tortów waży nie więcej niż ... kg.
- 75% zamawianych tortów waży nie więcej niż ... kg.

Zadanie 2. Przeprowadzić w swojej grupie ćwiczeniowej badanie dotyczące miesięcznych wydatków na ksero. Wyznaczyć podstawowe charakterystyki liczbowe. Zinterpretować wyniki.

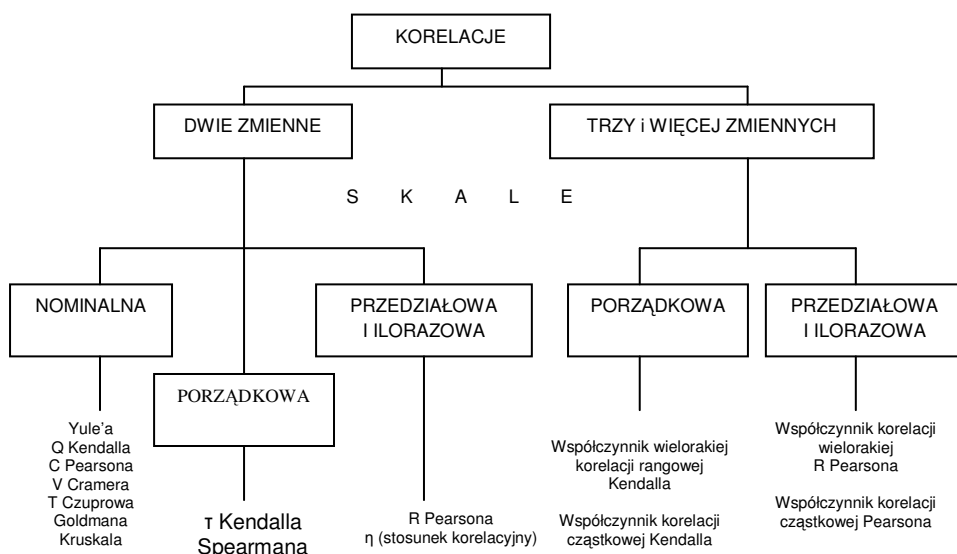
Zadanie 3. Przeprowadzić w swojej grupie ćwiczeniowej badanie dotyczące korzystania ze środków transportu w celu dojazdu na studia (skupić się na głównym środku transportu: MPK, PKS, PKP, inne). Wyznaczyć wskaźniki struktury, dominantę. Zinterpretować wyniki.

Zadanie 4. Przeprowadzić w swojej grupie ćwiczeniowej badanie dotyczące kwestii posiadania, częstości korzystania oraz wad i zalet korzystania z kuchenki mikrofalowej. Zapropionować sposób pomiaru. Dokonać analizy i interpretacji uzyskanych wyników.

11. ANALIZA I INTERPRETACJA DANYCH – ANALIZA KORELACJI I REGRESJI

11.1. Analiza korelacji

Badając zjawiska, czasami stwierdzić można, że konkretne zjawisko jest przyczynowo uwarunkowane innym (lub innymi). Stwierdzenie istnienia współzależności zjawisk odnosi się do sytuacji, w której zmiany natężenia jednego zjawiska wywołują zmiany w natężeniu drugiego zjawiska. Uwzględniając liczbę zmiennych oraz skalę ich pomiaru można dokonać podziału metod analizy korelacji jak na poniższym rysunku.



Rys. 23. Podział metod analizy korelacji

Źródło: S. Mynarski, *Metody ilościowe i jakościowe badań rynkowych i marketingowych*, <http://www.statsoft.pl/czytelnia> (12.03.2010 r.)

Zazwyczaj na przedmiocie „statystyka opisowa” studenci rozpoczynają analizę zależności od sporządzenia **korelacyjnego wykresu rozrzutu** (o ile jest to możliwe). W prostokątnym układzie współrzędnych na osi odciętych zaznaczamy wartości tej cechy, którą określamy jako **zmienną niezależną** (inaczej **zmienną objaśniającą**). Na osi rzędnych zaznaczamy wartości cechy określanej mianem **zmiennnej zależnej (zmienna objaśniana)**. **Korelacja dodatnia** występuje wtedy, gdy wzrostowi wartości jednej cechy odpowiada wzrost średnich wartości drugiej cechy. **Korelacja**

ujemna występuje wtedy, gdy wzrostowi wartości jednej cechy odpowiada spadek średnich wartości drugiej cechy.

Najliczniejszą grupę metod analizy korelacji stanowią miary korelacji dla dwóch zmiennych nominalnych. W grupie tej znajdują zastosowanie proste współczynniki: Yule'a, Q-Kendalla, T-Czuprowa, V-Cramera czy C-Pearsona. W przypadku mierzonych przynajmniej na skali porządkowej (rangowej) najczęściej wykorzystywaną miarą korelacji jest współczynnik korelacji rang Kendalla oraz współczynnik korelacji rang R Spearmana. Współczynnik ten może być z powodzeniem wykorzystywany do badania profilów zbudowanych na zasadzie dyferencjału semantycznego, tj. tam, gdzie można uporządkować natężenie cechy (od najmniejszego do maksymalnego). Gdy dwie zmienne mierzone są na skali przedziałowej lub ilorazowej najczęściej wykorzystywaną miarą korelacji jest współczynnik korelacji liniowej r-Pearsona oraz stosunek korelacyjny η . Jeśli z kolei trzy i więcej zmiennych jest mierzonych na skali przedziałowej lub ilorazowej, wówczas znajdują zastosowanie współczynniki korelacji wielorakiej R-Pearsona i współczynniki korelacji cząstkowej Pearsona.

11.2. Współczynniki korelacji – dwie zmienne nominalne

Zakładając, że współzależność badanych zmiennych jest statystycznie istotna, można określić siłę tej zależności za pomocą miar wyprowadzonych ze statystyki χ^2 (np. za pomocą współczynnika zbieżności T-Czuprowa). Zaletą współczynnika zbieżności T-Czuprowa jest to, że może on być stosowany zarówno do mierzenia korelacji zmiennych mierzalnych jak i niemierzalnych. Wadą współczynnika zbieżności T-Czuprowa jest to, że nie mierzy kierunku korelacji, bowiem bez względu na to, czy korelacja jest dodatnia czy ujemna, współczynnik ten jest liczbą z przedziału $\langle 0,1 \rangle$. Więcej informacji na temat współczynnika zbieżności T-Czuprowa znajduje się w podrozdziale 13.4 poświęconemu testowi niezależności.

11.3. Współczynniki korelacji – dwie zmienne mierzone na skali porządkowej

Założmy, że mamy n jednostek (obiektów) opisanych za pomocą dwóch cech. Jednostki można uporządkować ze względu na wartości każdej cechy oddzielnie. Numery określające miejsce jednostki w uporządkowaniu nazywa się rangami, a procedurę nadawania rang – rangowaniem. Dla uproszczenia przyjmijmy, że porządkujemy jednostki malejąco. W przypadku mierzonych przynajmniej na skali porządkowej (rangowej) najczęściej wykorzystywaną miarą korelacji jest **współczynnik korelacji rang Kendalla** oraz **współczynnik korelacji rang R Spearmana**.

Współczynnik korelacji rang Spearmana wyznacza się ze wzoru:

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

gdzie d_i – różnice między rangami odpowiadających sobie wartości cechy x_i i cechy y_i ($i=1, \dots, n$). W sytuacji, gdy wystąpią jednakowe wartości cechy dla różnych jednostek, przypisuje się tym jednostkom wartości równe średniej arytmetycznej z kolejnych rang.

Współczynnik korelacji rang Spearmana służy do opisu siły korelacji dwóch cech w przypadku gdy:

- cechy są mierzalne, a badana zbiorowość jest nieliczna,
- cechy mają charakter jakościowy i istnieje możliwość ich uporządkowania.

Współczynnik korelacji rang Spearmana przyjmuje wartości z przedziału $\langle -1, 1 \rangle$. Znak współczynnika korelacji informuje o kierunku korelacji („+” oznacza korelację dodatnią, a „-” korelację ujemną). Wartość bezwzględna współczynnika korelacji informuje o sile związku – im bliżej $|r_s|$ jest liczby 1, tym silniejsza jest badana zależność. Wartość współczynnika korelacji rang bliska zera ($r_{XY} \approx 0$) oznacza brak zależności.

Współczynnik korelacji rang τ Kendalla daje ocenę podobieństwa uporządkowań zbioru danych dla dwóch zmiennych mierzonych na skali porządkowej. Współczynnik Kendalla przyjmuje wartości z przedziału $\langle -1, 1 \rangle$. Wartość 1 oznacza pełną zgodność uporządkowań, wartość 0 brak zgodności, a wartość -1 pełną ich przeciwstawność. Współczynnik korelacji rang τ Kendalla zostanie wyznaczony według następującej procedury. Najpierw próbę należy uporządkować w taki sposób, aby rangi jednej z cech (np. X) ustawione były w kolejności naturalnej (1, 2, 3, ..., n), a rangi drugiej cechy (np. Y) odpowiadały rangom pierwszej cechy z pierwszego ułożenia. Następnie oblicza się różnice pomiędzy kolejnymi rangami.

Współczynnik korelacji rang τ Kendalla występuje w trzech wariantach:

- τ_a dla rang niepowiązanych
- τ_b dla rang powiązanych
- τ_c dla rang powiązanych w układzie tablicowym¹¹⁶.

¹¹⁶ S. Mynarski, Praktyczne metody analizy danych rynkowych i marketingowych, Kantor Zakamycze, Zakamycze 2000, s. 136.

Współczynnik korelacji rang τ Kendalla oblicza się według następującego wzoru:

$$\tau_a = \frac{P - Q}{\frac{n(n-1)}{2}}$$

gdzie n to liczność próby.

Współczynnik korelacji rang τ_b Kendalla oblicza się według następującego wzoru:

$$\tau_b = \frac{P - Q}{\sqrt{\frac{n(n-1)}{2} - T} \sqrt{\frac{n(n-1)}{2} - U}}$$

gdzie n to liczność próby.

Współczynnik korelacji rang τ_c Kendalla oblicza się według następującego wzoru:

$$\tau_c = \frac{P - Q}{\frac{n^2(m-1)}{2m}}$$

gdzie n to liczność próby, a m oznacza mniejszą z liczby wierszy i kolumn.

Chcąc obliczyć τ , należy zestawić obserwacje z próby we wszystkie możliwe pary, a następnie podzielić te pary na trzy możliwe kategorie:

- **pary zgodne** – porównywane zmienne w obrębie tych dwóch obserwacji zmieniają się w tę samą stronę, tzn. albo w pierwszej obserwacji obydwie są większe niż w drugiej, albo obydwie mniejsze. Liczba takich par w próbie będzie dalej oznaczana przez P .
- **pary niezgodne** – zmienne zmieniają się w przeciwną stronę, to znaczy jedna z nich jest większa dla tej obserwacji w parze, dla której druga jest mniejsza. Liczba takich par w próbie będzie oznaczana przez Q .
- **pary związane** – jedna (lub obie) ze zmiennych ma równe wartości w obydwu obserwacjach. Wyznaczamy T i U wg wzorów:

$$T = \frac{\sum (t_i^2 - t_i)}{2}, \quad U = \frac{\sum (u_i^2 - u_i)}{2}$$

gdzie t_i jest liczbą rang w każdej grupie rang wiązanych występujących w X , a u_i jest liczbą rang w każdej grupie rang wiązanych występujących w Y .

Współczynnik τ Kendalla stanowi różnicę między prawdopodobieństwem, że porównywane zmienne będą układały się w tym samym porządku dla dwóch obserwacji, a prawdopodobieństwem, że ułożą się w przeciwnym porządku. τ Kendalla przyjmuje wartości od -1 do 1 włącznie. Wartość +1 oznacza, że każda ze zmiennych rośnie przy wzroście drugiej. Wartość -1 oznacza, że każda maleje przy wzroście drugiej. Tym samym τ Kendalla jest miarą monotonicznej zależności zmiennych.

Współczynnik τ Kendalla, co do swych założeń, odpowiada współczynnikiowi Spearmana. Dotyczy to również ich statystycznej mocy. Jednakże wielkości obu współczynników zwykle nie pokrywają się, gdyż ich podstawy logiczne oraz formuły obliczeniowe bardzo się różnią. Co ważniejsze, współczynniki te posiadają różną interpretację: Współczynnik Spearmana można traktować podobnie jak współczynnik korelacji liniowej Pearsona, tzn. w kategoriach procentu wyjaśnianej zmienności, przy czym Spearmana jest wyliczany w oparciu o rangi. Z kolei współczynnik τ Kendalla opiera się na prawdopodobieństwie, tzn. różnicy między prawdopodobieństwem tego, że dwie zmienne układają się w tym samym porządku w obrębie obserwowanych danych, a prawdopodobieństwem, że ich uporządkowanie się różni.

11.4. Współczynniki korelacji – dwie zmienne mierzone na skali przedziałowej

Najszerzej wykorzystywanym współczynnikiem korelacji jest współczynnik korelacji liniowej Pearsona. Wartość współczynnika korelacji nie zależy od używanych jednostek miary; na przykład, korelacja pomiędzy wysokością i masą ciała będzie taka sama niezależnie od tego czy jako jednostki stosowano cale i funty, czy centymetry i kilogramy. Współczynnik korelacji Pearsona zakłada, że dane dwie zmienne zostały zmierzone przynajmniej na skalach przedziałowych.

Współczynnik korelacji liniowej Pearsona r_{XY} , przyjmujący wartości z przedziału $\langle -1, 1 \rangle$, jest miarą kierunku i siły związku liniowego między cechami X i Y. Znak współczynnika korelacji informuje o kierunku korelacji („+” oznacza korelację dodatnią, a „-” korelację ujemną). Wartość bezwzględna współczynnika korelacji informuje o sile związku – gdy $r_{XY} = -1$ lub $r_{XY} = 1$, wówczas między zmiennymi zachodzi zależność w postaci funkcji liniowej (punkty na wykresie układają się wzdłuż linii prostej). Wartość współczynnika korelacji liniowej bliska zera ($r_{XY} \approx 0$) oznacza brak zależności liniowej (nie wyklucza istnienia zależności nieliniowej).

Dla szeregów szczegółowych współczynnik korelacji liniowej wyznacza się ze wzoru:

$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

gdzie \bar{x} , \bar{y} – średnie.

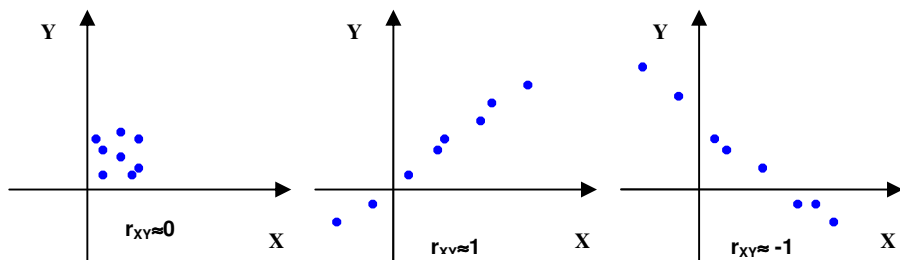
Jeżeli dane podane są w postaci tablicy korelacyjnej, to współczynnik korelacji liniowej wyznacza się ze wzoru:

$$r_{XY} = \frac{\sum_{i=1}^r \sum_{j=1}^s (\dot{x}_i - \bar{x})(\dot{y}_j - \bar{y})n_{ij}}{\sqrt{\sum_{i=1}^r (\dot{x}_i - \bar{x})^2 n_{i.} \sum_{j=1}^s (\dot{y}_j - \bar{y})^2 n_{.j}}}$$

gdzie \dot{x}_i , \dot{y}_j – środki przedziałów klasowych cech (odpowiednio X i Y), a $n_{i.}$, $n_{.j}$, n_{ij} – odpowiednimi licznosciami tablicy korelacyjnej.

W analizach statystycznych zwykle przyjmuje się, że jeżeli wartość bezwzględna współczynnika korelacji liniowej Pearsona wynosi:

- mniej niż 0,2, to praktycznie brak jest związku liniowego między badanymi cechami
- 0,2-0,4 – zależność jest słaba
- 0,4-0,7 – zależność jest umiarkowana
- 0,7-0,9 – zależność jest znacząca
- powyżej 0,9 – zależność jest bardzo silna.



Rys. 24. Zależność współczynnika korelacji liniowej Pearsona od zakresu zmienności cech

Źródło: opracowanie własne.

W badaniach niejednokrotnie zachodzi konieczność ustalenia zależności między dwiema cechami X i Y, z których obie (lub przynajmniej jedna) mają charakter jakościowy, np. :

- miejsce zamieszkania – preferencje kulinarne
- płeć– zainteresowanie programami TV
- wykształcenie – źródła informacji o nowym produkcie.

W takim przypadku materiał statystyczny umieszczony w zbudowanej do tego celu tablicy wielozdzielczej (kontyngencji) o określonej liczbie kolumn i wierszy, w której zamieszczamy liczebność poszczególnych wariantów cech.

11.5. Analiza regresji

W analizie zależności między cechami mierzalnymi możliwe jest nie tylko liczbowe określenie siły związku, ale również matematyczne odzwierciedlenie zależności za pomocą funkcji regresji (najczęściej rozważa się istnienie związku liniowego). Uwzględnia się przy tym podział na zmienne zależne i niezależne. Zmiennymi zależnymi są wszystkie te zmienne, które zależą od zmian innych zmiennych. Zmienne te nazywają się również zmiennymi objaśnianymi, ponieważ zależą od całego szeregu czynników wewnętrznych i zewnętrznych. W przeciwieństwie do nich, zmienne niezależne same są czynnikami wywołującymi zmiany. Dlatego nazywają się one również zmiennymi objaśniającymi.

Po ustaleniu, że pomiędzy rozważanymi cechami istnieje zależność korelacyjna, można oszacować parametry funkcji regresji. Oznacza to badanie wpływu jednej (regresja prosta) lub kilku zmiennych, objaśniających na zmienną objaśnianą (regresja wieloraka). Analiza regresji polega na estymacji parametrów równania teoretycznego, odzwierciedlającego badaną zależność. Najprostsze modele regresji zakładają występowanie zależności liniowych istniejących pomiędzy zmienną objaśnianą a zmiennymi objaśniającymi¹¹⁷.

$$\text{Rozważmy model} \quad Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k + \varepsilon$$

Najpowszechniej stosowaną w praktyce metodą estymacji parametrów jest klasyczna metoda najmniejszych kwadratów (KMNK)¹¹⁸. Idea KMNK:

¹¹⁷ W liniowym modelu regresji wielorakiej zmienne objaśniające powinny odznaczać się następującymi właściwościami: powinny być silnie skorelowane ze zmienną objaśnianą, powinny być słabo skorelowane między sobą, powinny wykazywać odpowiednio dużą zmienność.

¹¹⁸ Zapropionowana ponad 200 lat temu przez Legendre'a i Gaussa KMNK służy do estymacji parametrów strukturalnych modeli liniowych i sprowadzalnych do liniowych. Założenia KMNK są następujące: model jest niezmienniczy ze względu na obserwacje, zakłada się stabilność relacji występujących między badanymi

należy ustalić takie wartości ocen parametrów strukturalnych a_0, a_1, \dots, a_k , dla których suma kwadratów odchyłeń wartości empirycznych zmiennej objaśnianej (y_i) od wartości teoretycznych (\hat{y}_i) jest najmniejsza

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min$$

Model po oszacowaniu parametrów przyjmuje postać:

$$\hat{Y} = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k$$

gdzie \hat{Y} – wartość teoretyczna zmiennej objaśnianej,
 a_0, a_1, \dots, a_k – oceny parametrów strukturalnych.

Interpretacja oceny a_i parametru strukturalnego α_i ($i=1, 2, \dots, k$) występującego przy zmiennej X_i jest następująca: ocena parametru strukturalnego a_i wskazuje, o ile przeciętnie zmieniła się (wzrosła lub zmalała) wartość zmiennej objaśnianej Y , jeżeli przy niezmiennych wartościach innych zmiennych objaśniających (*ceteris paribus*) wartość zmiennej X_i zmieniła się (wzrosła lub zmalała) o jednostkę. Zasada *ceteris paribus* znajduje zastosowanie w interpretacji ocen parametrów strukturalnych modeli ekonometrycznych, ponieważ ocena parametru informuje o bezpośrednim efekcie jednostkowej zmiany zmiennej objaśniającej, pod warunkiem, że pozostałe zmienne objaśniające są constans.

W przypadku funkcji regresji zmiennej zależnej Y przy danych wartościach zmiennej niezależnej X

$$Y = \alpha_0 + \alpha_1 X + \varepsilon$$

oceny parametrów strukturalnych otrzymamy, korzystając ze wzorów:

$$a_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$

zmiennymi; model jest liniowy; zmienne objaśniające są nielosowe; zmienne objaśniające są liniowo niezależne. Żadna ze zmiennych nie może stanowić kombinacji liniowych pozostałych; liczba obserwacji n musi być większa od liczby szacowanych parametrów strukturalnych; składnik losowy ma rozkład normalny o wartości oczekiwanej zero i stałej skończonej wariancji; nie występuje autokorelacja składników losowych, czyli składniki losowe pochodzące z różnych okresów nie są wzajemnie zależne; składnik losowy nie jest skorelowany z żadną ze zmiennych objaśniających.

Interpretacja współczynnika regresji a_1 jest następująca: wskazuje, o ile przeciętnie zmieniała się (wzrosła lub zmalała) wartość zmiennej objaśnianej (zależnej – Y), jeżeli wartość zmiennej objaśniającej (niezależnej – X) zmieniła się (wzrosła lub zmalała) o jednostkę.

Do oceny dopasowania funkcji regresji do danych empirycznych najczęściej stosowanymi miarami są:

✓ **współczynnik zbieżności** wyznaczany ze wzoru

$$\varphi^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

✓ **współczynnik determinacji** wyznaczany ze wzoru

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

gdzie y_i oznaczają wartości empiryczne cechy Y, a \hat{y}_i – wartości teoretyczne wyznacza się na podstawie równania prostej regresji.

Współczynnik zbieżności określa, w jakim stopniu zmiany jednej cechy nie są wyjaśnione przez zmiany drugiej cechy. Współczynnik determinacji określa, w jakim stopniu zmiany jednej cechy są wyjaśnione przez zmiany drugiej cechy. Współczynnik zbieżności przyjmuje wartości z przedziału $\langle 0, 1 \rangle$, przy czym im mniejszą wartość przyjmie, tym lepsze jest dopasowanie funkcji regresji do danych empirycznych. Współczynnik determinacji przyjmuje wartości z przedziału $\langle 0, 1 \rangle$, przy czym im większą wartość przyjmie, tym lepsze jest dopasowanie funkcji regresji do danych empirycznych. Funkcja regresji może służyć do przewidywania (prognozowania) wartości jednej cechy przy ustalonym poziomie drugiej z nich.

11.6. Przykłady z omówieniem

Przykład 1. Małżeństwo poproszono o uszeregowanie rangowe dziesięciu filmów, które mogą wypożyczyć z wypożyczalni kaset wideo (1 – najlepszy film, ..., 10 – najgorszy film). Wyniki ocen filmów prezentuje tabela:

Wyszczególnienie	A	B	C	D	E	F	G	H	I	J
Mąż	1	2	3	4	5	6	7	8	9	10
Żona	8	7	10	9	1	2	6	5	4	3

- Wykorzystując współczynnik τ Kendalla, sprawdzić czy występuje zależność pomiędzy wyborami męża i żony,
- Wykorzystując współczynnik korelacji rang Spearmana, sprawdzić czy występuje zależność pomiędzy wyborami męża i żony.

Rozwiązanie

- Jest to przykład z rangami niepowiązanymi. W celu prześledzenia istniejących powiązań między rangami, zestawmy wszystkie porównania parami:

AB-									
AC+	BC+								
AD+	BD+	CD-							
AE-	BE-	CE-	DE-						
AF-	BF-	CF-	DF-	EF+					
AG-	BG-	CG-	DG-	EG+	FG+				
AH-	BH-	CH-	DH-	EH+	FH+	GH-			
AI-	BI-	CI-	DI-	EI+	FI+	GI-	HI-		
AJ-	BJ-	CJ-	DJ-	EJ+	FJ+	GJ-	HJ-	IJ-	

W przykładzie współczynnik korelacji rang τ Kendalla dla badanych cech wynosi:

$$\tau_a = \frac{13 - 32}{\frac{10(10-1)}{2}} = \frac{-19}{45} = -0,45$$

Współczynnik korelacji rang τ Kendalla przyjął wartość $-0,45$, co oznacza, iż istnieje umiarkowana korelacja ujemna między ulubionymi rodzajami filmów wybranymi przez męża i przez żonę.

- b) W celu prześledzenia istniejących powiązań między rangami, obliczymy poszczególne różnice między rangami (d_i), a następnie d_i^2 :

Wyszczególnienie	A	B	C	D	E	F	G	H	I	J
Mąż	1	2	3	4	5	6	7	8	9	10
Żona	8	7	10	9	1	2	6	5	4	3
d_i	-7	-5	-7	-5	4	4	1	3	5	7
d_i^2	49	25	49	25	16	16	1	9	25	49

W przykładzie współczynnik korelacji rang Spearmana wynosi

$$r_s = 1 - \frac{6 \cdot 264}{10(100 - 1)} = -0,60$$

co oznacza, iż istnieje umiarkowana korelacja ujemna między ulubionymi rodzajami filmów wybranymi przez męża i przez żonę.

Przykład 2. Małżeństwo poproszono o uszeregowanie rangowe dziesięciu filmów, które mogą wypożyczyć z wypożyczalni kaset wideo (1 – najlepszy film, ..., 10 – najgorszy film). Wyniki ocen filmów prezentuje tabela:

Wyszczególnienie	A	B	C	D	E	F	G	H	I	J
Mąż	1	2	4	4	4	6	7	8	9,5	9,5
Żona	8	7	10	9	2	2	6	5	4	2

- a) Wykorzystując współczynnik τ Kendalla, sprawdzić czy występuje zależność pomiędzy wyborami męża i żony,
 b) Wykorzystując współczynnik korelacji rang Spearmana, sprawdzić czy występuje zależność pomiędzy wyborami męża i żony.

Rozwiązanie

- a) Jest to przykład z rangami powiązanymi (C, D, E oraz I, J w ocenie męża, E, F, J w ocenie żony) W celu prześledzenia istniejących powiązań między rangami, zestawmy wszystkie porównania parami:

AB-									
AC+	BC+								
AD+	BD+	CD*							
AE-	BE-	CE*	DE*						
AF-	BF-	CF-	DF-	EF*					
AG-	BG-	CG-	DG-	EG+	FG+				
AH-	BH-	CH-	DH-	EH+	FH+	GH-			
AI-	BI-	CI-	DI-	EI+	FI+	GI-	HI-		
AJ-	BJ-	CJ-	DJ-	EJ*	FJ*	GJ-	HJ-	IJ*	

W przykładzie współczynnik korelacji rang τ Kendalla dla badanych cech wynosi:

$$\tau_b = \frac{10 - 28}{\sqrt{\frac{10(10-1)}{2} - 4} \sqrt{\frac{10(10-1)}{2} - 3}} = \frac{-18}{\sqrt{41}\sqrt{42}} = -0,44$$

gdzie T i U obliczono następująco: $T = \frac{1}{2}[(3^2-3)+(2^2-2)] = 4$, $U = \frac{1}{2}[(3^2-3)] = 3$

Współczynnik korelacji rang τ Kendalla przyjął wartość $-0,44$, co oznacza, iż istnieje umiarkowana korelacja ujemna między ulubionymi rodzajami filmów wybranymi przez męża i przez żonę.

b) W celu prześledzenia istniejących powiązań między rangami, obliczymy poszczególne różnice między rangami (d_i), a następnie d_i^2 :

Wyszczególnienie	A	B	C	D	E	F	G	H	I	J
Mąż	1	2	4	4	4	6	7	8	9,5	9,5
Żona	8	7	10	9	2	2	6	5	4	2
d_i	-7	-5	-6	-5	2	4	1	3	5,5	7,5
d_i^2	49	25	36	25	4	16	1	9	30,25	56,25

W przykładzie współczynnik korelacji rang Spearmana wynosi

$$r_s = 1 - \frac{6 \cdot 251,5}{10(100-1)} = -0,52$$

co oznacza, iż istnieje umiarkowana korelacja ujemna między ulubionymi rodzajami filmów wybranymi przez męża i przez żonę.

Przykład 3. Rozpatrzmy zainteresowanie udziałem w koncertach organizowanych przez Miejski Ośrodek Kultury a wiekiem respondentów. Dane prezentuje tabela:

Wyszczególnienie	Uczęszczanie na koncerty			Razem
	Rzadko	Często	Bardzo często	
Wiek przedprodukcyjny	32	127	251	410
Wiek produkcyjny	158	289	153	600
Wiek poprodukcyjny	51	11	8	70
Razem	241	427	412	1080

Wykorzystując współczynnik τ Kendalla, sprawdzić czy występuje zależność pomiędzy udziałem w koncertach organizowanych przez Miejski Ośrodek Kultury a wiekiem respondentów.

Rozwiązanie:

Każdy z elementów pierwszej kolumny, przy obliczaniu sumy rang, otrzyma znak plus w połączeniu z każdym elementem położonym w prawo i w dół, czyli:

$$32(289+11+153+8)=14\ 752$$

$$158(11+8)=3\ 002$$

Elementy drugiej kolumny, przy obliczaniu sumy rang, otrzymają znak plus w połączeniu z każdym elementem leżącym w prawo i w dół oraz znak minus w połączeniu z każdym elementem położonym w lewo i w dół, czyli:

$$127(153+8-158-51)=-6\ 096$$

$$289(8-51)=-12\ 427$$

Elementy trzeciej kolumny, przy obliczaniu sumy rang, zawierają tylko ujemne składniki sumy:

$$251(-289-11-158-51)=-127\ 759$$

$$153(-11-51)=-9\ 486$$

$$\tau_c = \frac{17\ 754 - 155\ 768}{(1\ 080)^2 (3-1)} = \frac{-138\ 014}{388\ 800} = -0,35$$

Występuje nieznaczna ujemna zależność pomiędzy udziałem w koncertach organizowanych przez Miejski Ośrodek Kultury a wiekiem respondentów (im starszy respondent, tym rzadziej uczestniczy w koncertach).

Przykład 4. Analiza korelacji i regresji

Niech: X – liczba miesięcy, które upłynęły od ostatniego zakupu produktów w firmie firmy Kogucik, Y – wartość produktów (w zł) zakupionych po akcji promocyjnej w firmie Kogucik. Sporządź wykres rozrzutu. Wyznacz współczynnik korelacji liniowej Pearsona (skomentuj wynik). Rozważ model liniowy, model wykładniczy, potęgowy i logarytmiczny (porównaj dopasowanie modeli) – skorzystaj z wykresu sporządzonego w programie Excel.

Tabela 27.

Dane do przykładu 4.

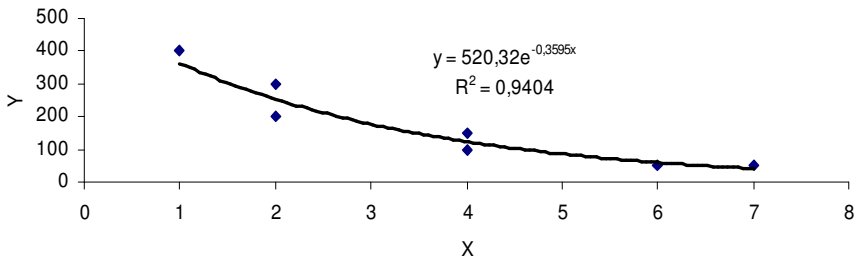
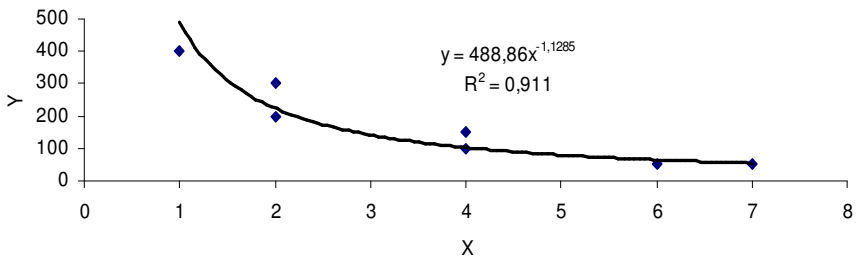
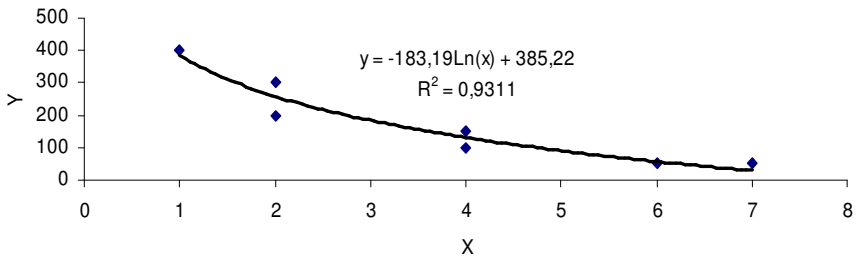
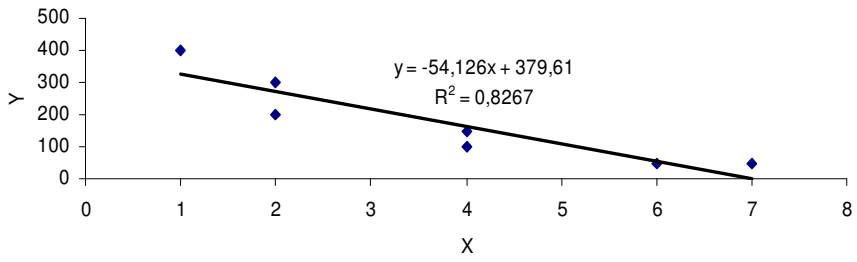
X	Y
1	400
2	300
2	200
4	100
4	150
6	50
7	50

Źródło: dane umowne.

Rozwiązanie

Jest to przykład korelacji ujemnej (współczynnik korelacji liniowej Pearsona wynosi $r = -0,91$). Najlepiej dopasowany jest model wykładniczy ($R^2 = 0,94$) (wykres 7).

Wykres 7. Wykres rozrzutu, porównanie modeli – przykład 4



Źródło: dane umowne, obliczenia własne.

Przykład 5. Analiza korelacji i regresji

Niech: X – ile razy klient dokonał zakupu produktów firmy Kogucik w ciągu roku, Y – wartość produktów (w zł) zakupionych po akcji promocyjnej firmy Kogucik. Sporządź wykres rozrzutu. Wyznacz współczynnik korelacji liniowej Pearsona (skomentuj wynik). Rozważ model liniowy, model wykładniczy, potęgowy i logarytmiczny (porównaj dopasowanie modeli) – skorzystaj z wykresu sporządzonego w programie Excel.

Tabela 28.

Dane do przykładu 5.

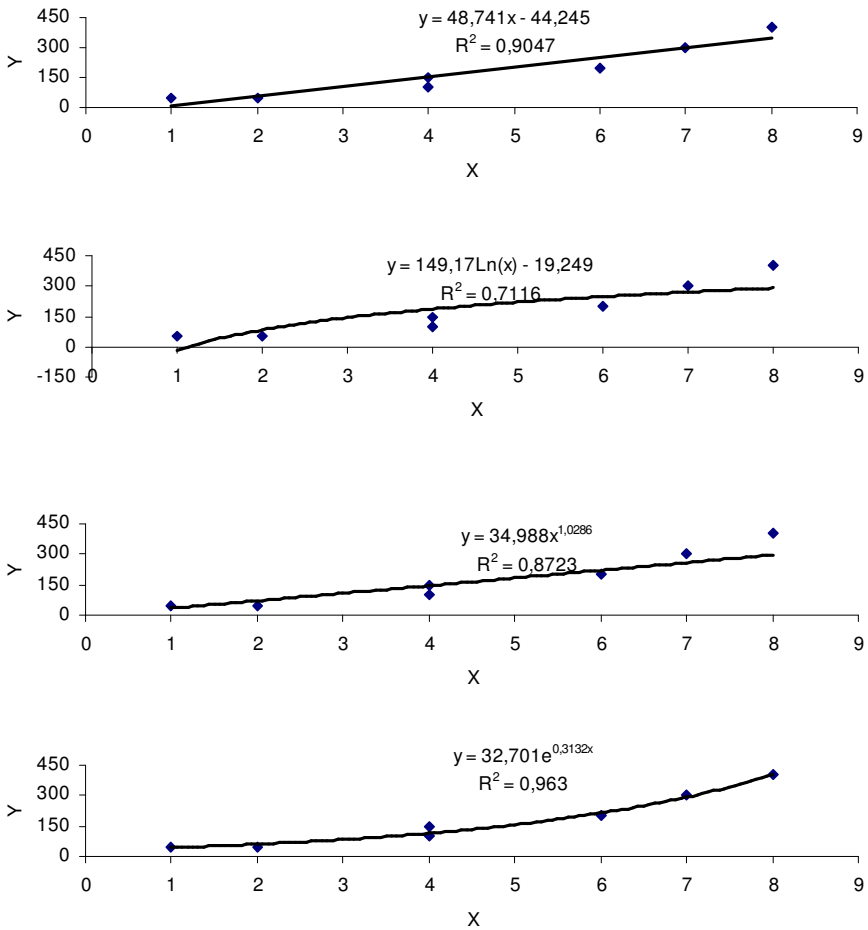
X	Y
8	400
7	300
6	200
4	100
4	150
2	50
1	50

Źródło: dane umowne.

Rozwiązanie

Jest to przykład korelacji dodatniej (współczynnik korelacji liniowej Pearsona wynosi $r = 0,95$). Najlepiej dopasowany jest model wykładniczy ($R^2 = 0,91$) (wykres 8).

Wykres 8. Wykres rozrzutu, porównanie modeli – przykład 5



Źródło: dane umowne, obliczenia własne.

Przykład 6. Zastosowanie funkcji REGLINP do szacowania parametrów ekonometrycznego modelu liniowego. Niech: X_1 – liczba miesięcy, które upłynęły od ostatniego zakupu produktów w firmie firmy Kogucik, X_2 – ile razy klient dokonał zakupu produktów firmy Kogucik w ciągu roku, Y – wartość produktów (w zł) zakupionych po akcji promocyjnej w firmie Kogucik.

X_1	X_2	Y
1	8	400
2	7	300
2	6	200
4	4	100
4	4	150
6	2	50
7	1	50

- c) Dla danych z przykładu 6 oszacuj parametry modelu $Y = \alpha_0 + \alpha_1 X_1 + \varepsilon$ i odpowiedz na pytanie, jakiej wartości zakupów należy się spodziewać, gdy od ostatniego zakupu minęły trzy miesiące?.
- d) Dla danych z przykładu 6 oszacuj parametry modelu $Y = \beta_0 + \beta_1 X_2 + \varepsilon$ i zinterpretuj uzyskaną ocenę b_1 .
- e) Dla danych z przykładu 6 oszacuj parametry modelu $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$ i zinterpretuj uzyskane oceny a_1 oraz a_2 . Porównaj otrzymaną ocenę a_1 ze współczynnikiem korelacji pomiędzy X_1 oraz Y ($r = -0,91$). Czy jest coś, co powinno nas zaniepokoić?

Rozwiązanie

Korzystając z funkcji REGLINP (program Excel) otrzymujemy:

- oceny parametrów strukturalnych (pierwszy wiersz tablicy)
- standardowe błędy ocen (drugi wiersz tablicy)
- współczynnik determinacji R^2 oraz odchylenie standardowe reszt¹¹⁹ S_e (trzeci wiersz tablicy)
- wartość statystyki Fishera-Snedecora F oraz liczbę stopni swobody df (czwarty wiersz tablicy)
- regresyjną sumę kwadratów ss_{reg} oraz resztową sumę kwadratów ss_{resid} (piąty wiersz tablicy).

Tablica ma zawsze 5 wierszy, natomiast liczba kolumn zależy od liczby szacowanych parametrów strukturalnych (zanim zastosujemy funkcję REGLINP należy zaznaczyć odpowiedni obszar, natomiast po wypełnieniu pól „znane_y”, „znane_x”, „stała”, „statystyka” należy użyć kombinacji klawiszy Ctrl, Shift oraz Enter).

¹¹⁹ Odchylenie standardowe reszt informuje, o ile średnio rzecz biorąc, zaobserwowane wartości zmiennej objaśnianej odchylają się od wartości wyznaczonych na podstawie oszacowanego modelu.

Tabela 29.

Wygląd tablicy REGLINP

a_k	...	a_1	a_0
$S(a_k)$...	$S(a_1)$	$S(a_0)$
R^2	S_e		
F	Df		
SS_{Reg}	SS_{resid}		

Źródło: opracowanie własne.

- a) Dla modelu $Y = \alpha_0 + \alpha_1 X_1 + \varepsilon$ otrzymujemy, korzystając z funkcji REGLINP:

$$\hat{Y} = 379,61 - 54,13X_1$$

Po upływie trzech miesięcy należy się spodziewać, iż wartość zakupionych produktów wyniesie

$$379,61 - 54,13 \cdot 3 = 217,23 \text{ zł}$$

- b) Dla modelu $Y = \beta_0 + \beta_1 X_2 + \varepsilon$ otrzymujemy, korzystając z funkcji REGLINP:

$$\hat{Y} = -44,245 + 48,741X_2$$

Jeżeli klient dokona dodatkowego zakupu, to należy się spodziewać wzrostu wartości o 48,74 zł zakupionych w firmie Kogucik po akcji promocyjnej produktów.

- c) Dla modelu $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$ otrzymujemy, korzystając z funkcji REGLINP:

$$\hat{Y} = -817,52 + 101,59X_1 + 135,35X_2$$

Jeżeli liczba miesięcy, które upłynęły od ostatniego zakupu produktów w firmie Kogucik (X_1) wzrośnie o 1, a ilość razy, kiedy klient dokonał zakupu produktów firmy Kogucik w ciągu roku (X_2), nie ulegnie zmianie, to można oczekiwać, że wartość produktów zakupionych po akcji promocyjnej w firmie Kogucik (Y) wzrośnie o 101,59 zł. Ocena a_1 oraz współczynnik korelacji powinny być tego samego znaku (a nie są).

Jeżeli ilość razy, kiedy klient dokonał zakupu produktów firmy Kogucik w ciągu roku (X_2) wzrośnie o 1, a liczba miesięcy, które upłynęły od ostatniego zakupu produktów w firmie Kogucik (X_1) nie ulegnie zmianie, to można oczekiwać, że wartość produktów zakupionych po akcji promocyjnej w firmie Kogucik (Y) wzrośnie o 135,35 zł.

11.7. Zadania do samodzielnego opracowania

Zadanie 1. Konsumentów poproszono o rozdzielenie 100 punktów proszkom marek A, B, C, D, E (im więcej punktów otrzyma dana marka, tym jest lepiej postrzegana). Następnie konsumenci obejrżeli spoty reklamowe wszystkich pięciu marek proszków i ponownie rozdzielili 100 punktów. Wyniki prezentuje tabela (dane surowe):

Wyszczególnienie	A	B	C	D	E
Przed spotem reklamowym	20	30	40	5	5
Po spocie reklamowym	20	25	30	10	15

- Wykorzystując współczynnik τ Kendalla sprawdzić, czy występuje zależność pomiędzy oceną marki proszku przed i po obejrzeniu spotów reklamowych.
- Wykorzystując współczynnik korelacji rang Spearmana sprawdzić, czy występuje zależność pomiędzy ocenami przed i po obejrzeniu spotów reklamowych.

Zadanie 2. Pan Nowak jest cukiernikiem. Pan Nowak wykonuje na zamówienie m.in. torty. Ich cena zależy m.in. od wagi. Tabela przedstawia wagę (X) i cenę (Y) zamówionych w ciągu ostatniego tygodnia tortów.

Waga tortu (kg)	Cena tortu (zł)
3,2	60
3,8	80
2,2	54
1,8	36
1,5	26
1,4	26
1,2	22
0,4	15
0,9	20
0,8	18
1,2	22
3,9	80

Uzupełnij następujące zdania:

- a) Zamawiany tort waży przeciętnie... kg.
- b) Pomiędzy wagą a ceną tortu istnieje ... (silna/słaba) korelacja ... (dodatnia/ujemna).
- c) Równanie prostej regresji Y względem X (X – zmienna objaśniająca, Y – zmienna objaśniana)

$$\hat{Y} = \dots X + \dots$$

- d) Wartość współczynnika determinacji wynosi $R^2 = \dots$
- e) ... % zmienności ceny tortu zostało wytłumaczone przez wagę tortu.
- f) Jeśli waga tortu wzrośnie o 1 kg, to należy się spodziewać ... (wzrostu/spadku) ceny o ... zł.

Zadanie 3. Badaniu poddano pięć wybranych punktów sprzedaży luksusowej bielizny firmy „Lux”. Zbadano powierzchnię ekspozycji w m^2 (X) oraz miesięczną wartość sprzedaży w tys. zł (Y). Wyniki badania prezentuje tabela:

Powierzchnia ekspozycji w m^2 (X)	44	64	16	32	24
Wartość sprzedaży w tys. zł (Y)	45	67	32	44	32

- a) Wyznacz współczynnik korelacji liniowej Pearsona. Zinterpretuj wynik.
- b) Wyznacz równanie prostej regresji Y względem X (X – zmienna objaśniająca, Y – zmienna objaśniana) oraz wartość współczynnika determinacji. Zinterpretuj wyniki.
- c) Jaka jest przewidywana wartość sprzedaży przy powierzchni ekspozycji $20 m^2$?

Zadanie 4. W pewnym przedsiębiorstwie przez 9 kolejnych miesięcy przeprowadzono badanie, związane z wprowadzaniem do produkcji nowego proszku do prania. Niech: Y – wielkość produkcji (ton), X_1 – koszty produkcji (tys. zł), X_2 – doświadczenie kierownika mierzone stażem pracy (w miesiącach). Wyniki badania prezentuje tabela:

Y	X ₁	X ₂
2	3	0
1	2	1
8	8	2
5	5	3
6	6	4
4	4	5
7	6	6
9	8	7
8	7	8

- c) Oszacuj parametry modelu $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$ korzystając z funkcji REGLINP.
- d) Zinterpretuj uzyskane oceny a_1 oraz a_2 .
- e) Oblicz i zinterpretuj wartość współczynnika determinacji.

Zadanie 5. W pewnym powiecie przeprowadzono badanie na grupie 10 losowo wybranych osób.. Niech: Y – roczne wydatki na zakup ubrania w gospodarstwie domowym (tys. zł), X₁ – liczba osób w rodzinie, X₂ – roczne dochody netto w gospodarstwie domowym (tys. zł). Wyniki badania prezentuje tabela:

Y	X ₁	X ₂
0,8	1	21
1,4	2	25
0,3	1	10
2,3	2	37
3,8	2	48
2,1	3	27
3,3	4	35
2,9	3	35
3,0	5	25
4,0	3	51

- a) Oszacuj parametry modelu $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$ korzystając z funkcji REGLINP.
- b) Zinterpretuj uzyskane oceny a_1 oraz a_2 .
- c) Oblicz i zinterpretuj wartość współczynnika determinacji.
- d) Jakich rocznych wydatków na zakup ubrania należy się spodziewać w pięcioosobowym gospodarstwie domowym o dochodach netto 35 tys. zł?

12. ANALIZA I INTERPRETACJA DANYCH – ANALIZA SZEREGÓW CZASOWYCH

12. 1. Dekompozycja szeregu czasowego

Analiza szeregów czasowych ma na celu :

- wykrywanie natury zjawiska reprezentowanego przez sekwencję obserwacji
- prognozowanie (przewidywanie przyszłych wartości szeregu czasowego)

Oba te cele wymagają zidentyfikowania i opisanie, w sposób mniej lub bardziej formalny, elementów szeregu czasowego. Tak jak w większości innych analiz, w analizie szeregów czasowych zakłada się, że w danych można wyodrębnić składnik systematyczny oraz losowy, który utrudnia identyfikację struktury zjawiska. Większość technik analizy szeregów czasowych wiąże się z pewnymi formami filtrowania szumu w celu uwidocznienia składnika systematycznego.

Proces wyodrębniania poszczególnych składowych szeregu czasowego nazywa się dekompozycją szeregu czasowego. Strukturę większości szeregów czasowych można opisać przy pomocy dwóch podstawowych klas składników: trendu i sezonowości. Trend (tendencja rozwojowa) jest długookresową skłonnością do jednokierunkowych zmian (wzrostu lub spadku) wartości badanej zmiennej. Jest rozpatrywana jako konsekwencja działania stałego zestawu czynników. Może być wyznaczana, gdy dysponuje się długim ciągiem obserwacji. Modelem tendencji rozwojowej nazywamy taki model, w którym zmienną objaśniającą jest zmienna czasowa lub jej funkcja. Do najczęściej spotykanych modeli tendencji rozwojowej należą: model liniowy, model wykładniczy, model potęgowy, model logarytmiczny. Stały poziom prognozowanej zmiennej występuje wówczas, gdy w szeregu czasowym nie ma tendencji rozwojowej, zaś wartości prognozowanej zmiennej oscylują wokół pewnego (stałego) poziomu.

Wahania sezonowe są wahaniami wartości obserwowanej zmiennej wokół tendencji rozwojowej lub stałego poziomu tej zmiennej. Mają skłonność do powtarzania się w określonym czasie, odzwierciedlają wpływ pór roku na działalność gospodarczą. Wahania cykliczne wyrażają się w postaci długookresowych, rytmicznych wahań wartości zmiennej wokół tendencji rozwojowej lub stałego (przeciętnego) poziomu tej zmiennej. W ekonomii są one na ogół związane z cyklem koniunkturalnym gospodarki.

szereg czasowy



Rys. 25. Dekompozycja szeregu czasowego

Źródło: opracowanie własne.

12.2. Miary dynamiki

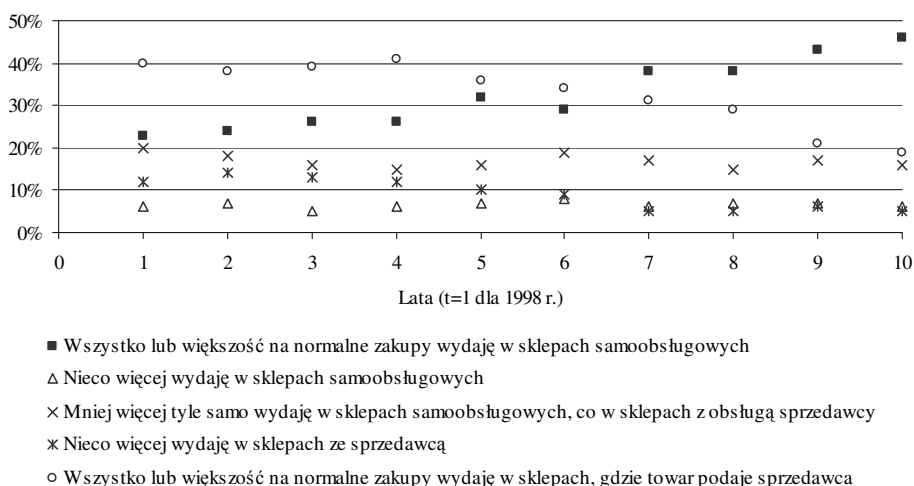
Do liczbowego określenia tempa i intensywności zmian zjawiska w czasie służą miary dynamiki, wśród których wyróżnia się: przyrosty i indeksy. **Miary dynamiki o podstawie stałej** (tzw. **jednopodstawowe**) służą do określenia zmian, jakie nastąpiły w poziomie zjawiska w kolejnych okresach (momentach) w porównaniu z okresem (momentem) przyjętym jako podstawowy (bazowy). Jako podstawę przyjmuje się poziom zjawiska w wyróżnionym okresie, przy czym najczęściej jest to poziom z pierwszego okresu. **Miary dynamiki o podstawie ruchomej** (tzw. **łańcuchowe**) służą do określenia zmian, jakie nastąpiły w poziomie zjawiska z okresu (momentu) na okres (moment). Jako podstawę odniesienia przyjmuje się poziom zjawiska w okresie poprzednim. **Przyrosty absolutne** – informują o ile wzrósł (zmalął) poziom zjawiska w okresie badanym w porównaniu z jego poziomem w okresie bazowym. Są to wielkości mianowane, wyrażone w takich samych jednostkach miary jak badane zjawisko. **Przyrosty względne** (wskaźnik tempa przyrostu) – stosunek przyrostu

absolutnego zjawiska do jego poziomu w okresie bazowym. Jeśli przyrosty względne przemnoży się przez 100, to otrzyma się procentowe przyrosty względne, zwane tempem zmian. **Tempo zmian** informuje, o ile procent poziom zjawiska w danym okresie jest wyższy (lub niższy) od poziomu w okresie przyjętym za podstawę porównań. Przez indeksy dynamiki rozumie się mierniki określające stosunek wielkości zjawiska w dwóch różnych okresach. Indeksy dynamiki mogą dotyczyć zjawisk jednorodnych opisanych pojedynczym szeregiem czasowym (**indywidualne indeksy dynamiki**) bądź zespołu zjawisk (**indeksy agregatowe**)¹²⁰.

12.3. Przykłady z omówieniem

Przykład 1. Rozważmy szereg czasowy, w którym obserwujemy systematyczne zmiany w czasie (trend) oraz towarzyszące im zmiany przypadkowe. Zmienność zaprezentowaną na wykresie „Alokacja wydatków przeznaczanych na normalne zakupy w ciągu tygodnia” można przybliżyć funkcją liniową (na wykresie 10 przykładowo wyznaczono trendy liniowe dla udziału osób, które „wszystko lub większość na normalne zakupy wydają w sklepach samoobsługowych” oraz tych, które „wszystko lub większość na normalne zakupy wydają w sklepach, gdzie towar podaje sprzedawca”).

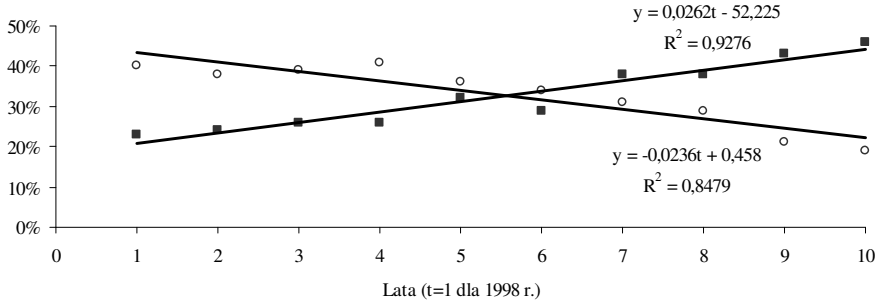
Wykres 9. Alokacja wydatków przeznaczanych na normalne zakupy w ciągu tygodnia



Źródło: Pentor (http://www.pentor.pl/52980.xml?doc_id=11280 z dnia 23.10.2008)

¹²⁰ S. Ostasiewicz, Z. Rusnak, U. Siedlecka, Statystyka. Elementy teorii i zadania, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1997, s. 300.

Wykres 10. Linie trendu



- Wszystko lub większość na normalne zakupy wydają w sklepach samoobsługowych
- Wszystko lub większość na normalne zakupy wydają w sklepach, gdzie towar podaje sprzedawca
- Liniowy (Wszystko lub większość na normalne zakupy wydają w sklepach samoobsługowych)
- Liniowy (Wszystko lub większość na normalne zakupy wydają w sklepach, gdzie towar podaje sprzedawca)

Źródło: Obliczenia własne na podstawie danych Pentor
http://www.pentor.pl/52980.xml?doc_id=11280 z dnia 23.10.2008)

Na podstawie danych dotyczących alokacji wydatków przeznaczanych na normalne zakupy w ciągu tygodnia można stwierdzić, iż:

- z roku na rok maleje (średnio o 2,36 punktu procentowego) liczba osób, które wydają wszystko lub większość na normalne zakupy w sklepach, gdzie towar podaje sprzedawca
- z roku na rok wzrasta (średnio o 2,62 punktu procentowego) liczba osób, które wydają wszystko lub większość na normalne zakupy w sklepach samoobsługowych

Przykład 2. Przykład wykorzystania miar dynamiki

Według danych ACNielsen w okresie czerwiec 2004 – maj 2005 sprzedano ogólnie w placówkach handlu detalicznego 20 379,5 ton herbat o łącznej wartości 827 463 tys. złotych i w porównaniu z okresem czerwiec 2003 - maj 2004 zanotowano spadek pod względem wielkości (o 1,3%) oraz wzrost wartości sprzedaży (o 1,9%). Rynek herbaty można podzielić ze względu na rodzaj oraz sposób pakowania herbaty. Zdecydowanie największy segment zarówno pod względem sprzedaży w ujęciu ilościowym (71% całego rynku herbaty), jak i wartościowym (58,1%) tworzą herbaty czarne. Warto jednak zauważyć, że udziały herbaty czarnej zmniejszyły się w okresie czerwiec 2004 – maj 2005 w porównaniu z analogicznym okresem roku poprzedniego o 1,6% pod względem wielkości i 2% pod względem wartości sprzedaży. Drugim liczącym się segmentem są herbaty owocowo-

ziołowe, stanowiące 21,2% wielkości i 31,6% wartości sprzedaży. Wymieniony segment odnotował wzrost udziałów w całym rynku herbat (o 0,6% w udziałach ilościowych i 0,9% w udziałach wartościowych), co może być spowodowane trendami konsumpcyjnymi lansującymi zdrowy tryb życia, poszukiwaniem nowych smaków lub znudzeniem klasycznymi produktami. Powyższą opinię potwierdzają pozostałe „młode segmenty” odgrywające mniejsze znaczenie, które zanotowały ilościowy oraz wartościowy wzrost w okresie czerwiec 2004 - maj 2005: herbaty czarne typu Earl Grey (5,9% wielkości i 7,2% wartości rynku), herbaty aromatyzowane (odpowiednio 1,6% i 2,8%) oraz herbaty w zestawach (0,2% i 0,3%).

Zauważalny wzrost udziałów w sprzedaży herbaty odnotowały nowoczesne kanały dystrybucji – w okresie czerwiec 2004 - maj 2005 w porównaniu z analogicznym okresem poprzedniego roku – nastąpił wzrost udziałów hipermarketów i supermarketów z 48,7% do 52,9%. Biorąc pod uwagę proporcję udziałów sprzedaży – znaczenie supermarketów wyraźnie wzrosło (o 3,5%) w porównaniu z hipermarketami (wzrost o 0,7%). Na podobnym poziomie kształtuje się sprzedaż w dużych sklepach tradycyjnych (wzrost o 0,1%). W innych tradycyjnych kanałach dystrybucji zanotowano spadek¹²¹.

12.4. Zadania do samodzielnego opracowania

Zadanie 1. Pan Nowak jest mężem pani Nowakowej od pięciu lat. Pan Nowak kupuje żonie prezent (najczęściej biżuterię) w każdą rocznicę ślubu. Tabela przedstawia wartość prezentu (Y) w omawianym okresie.

Rocznica ślubu	Wartość prezentu (zł)
1	600
2	620
3	540
4	360
5	400

Uzupełnij następujące zdania:

a) Równanie linii trendu i wartość współczynnika determinacji:

$$\hat{y} = \dots t + \dots$$

$$R^2 = \dots$$

¹²¹ Rynek kawy i herbaty w Polsce, raport ACNielsen

(<http://www.poradnikhandlowca.com.pl/bezcms/archiwum/online05/09/raport1.html> z dnia 03.12.2009)

- b) ... % zmienności wartości prezentu zostało wytłumaczone przez wpływ czasu.
- c) Z roku na rok wartość prezentu (rośnie/maleje) o zł.

Zadanie 2. Zbadaj dynamikę sprzedaży, dokonaj prognozy sprzedaży na dwa kolejne lata.

Tabela 30.

Dane do przykładu

Rok	2004	2005	2006	2007	2008	2009
Obserwowana wartość sprzedaży (mln zł)	17	24	26	25	29	35

Źródło: dane umowne.

Zadanie 3. Zinterpretuj dane zawarte w tabeli (dane miesięczne, I – styczeń, II – luty).

Tabela 31.

Produkcja ważniejszych wyrobów przemysłowych wg Polskiej Klasyfikacji Wyrobów i Usług (PKWiU 2008) / PRODPOL

Wyroby	Jednostka miary	Liczby bezwzględne		II	I	I-II
		2009=100	2010=100	2009=100	2010=100	2009=100
Węgiel: kamienny	tys. t	6 269	12 677	98,7	97,8	95,8
brunatny	tys. t	4 752	9 863	98,1	93,0	95,6
Rudy i koncentraty miedzi	tys. t	2 600	5 261	95,5	97,7	94,0
Rudy i koncentraty ołowiu	tys. t	5,2	10,2	63,1	104,8	60,7

Źródło: Produkcja ważniejszych wyrobów przemysłowych, GUS, <http://www.stat.gov.pl> (25.06.2010).

Zadanie 4. Odszukać w rocznikach statystycznych GUS lub w Internecie przykłady szeregów czasowych, w których podczas dekompozycji można wyróżnić:

- stały poziom zmiennej
- trend
- trend (lub stały poziom) i wahania sezonowe
- trend (lub stały poziom) i wahania cykliczne.

Zadanie 5. Sporządzić wykres prezentujący zamieszczone w tabeli dane. Dokonać dekompozycji szeregu (wyodrębnić poszczególne składowe danego szeregu poprzez ocenę wzrokową sporządzonego wykresu). Wyjaśnić pojęcie „cykl świński”.

Tabela 32.

Trzoda chlewna (w mln sztuk) w Polsce (stan w czerwcu)

Lata	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Trzoda chlewna	19,5	21,9	22,1	18,9	19,5	20,4	18,0	18,1	19,2	18,5	17,1	17,1	18,7	18,6	17,0

Źródło: Rocznik Statystyczny Rolnictwa, GUS, Warszawa 2001, s. XXXIX oraz Rocznik Statystyczny Rolnictwa i Obszarów Wiejskich, GUS, Warszawa 2005, s. 287.

Zadanie 6. Dane miesięczne dotyczące skupu mleka w Polsce (w mln litrów) w okresie od stycznia 2000 r. do grudnia 2004 r. zawiera tabela 33. Sporządzić wykres szeregu, dokonać dekompozycji. Ocenic charakter wahań.

Tabela 33.

Skup mleka w Polsce (w mln litrów)

Miesiąc	Skup mleka	Miesiąc	Skup mleka	Miesiąc	Skup mleka	Miesiąc	Skup mleka	Miesiąc	Skup mleka
01.2000	428,3	01.2001	477,7	01.2002	494,5	01.2003	525,1	01.2004	554,7
02.2000	414,2	02.2001	445,2	02.2002	468,6	02.2003	488,1	02.2004	536,6
03.2000	468,0	03.2001	517,1	03.2002	536,1	03.2003	572,8	03.2004	592,2
04.2000	482,2	04.2001	539,8	04.2002	576,1	04.2003	579,8	04.2004	607,9
05.2000	624,0	05.2001	669,7	05.2002	691,5	05.2003	680,4	05.2004	712,9
06.2000	631,0	06.2001	690,8	06.2002	675,0	06.2003	686,8	06.2004	717,3
07.2000	650,0	07.2001	675,5	07.2002	689,7	07.2003	695,5	07.2004	731,8
08.2000	646,2	08.2001	649,0	08.2002	665,4	08.2003	655,3	08.2004	703,5
09.2000	608,1	09.2001	607,9	09.2002	618,6	09.2003	619,2	09.2004	651,1
10.2000	576,2	10.2001	586,1	10.2002	577,7	10.2003	586,9	10.2004	632,5
11.2000	481,6	11.2001	490,1	11.2002	500,8	11.2003	517,9	11.2004	570,0
12.2000	476,9	12.2001	483,0	12.2002	509,3	12.2003	542,5	12.2004	590,0

Źródło: Rolnictwo w I półroczu 2005 r., GUS, Warszawa 2005, s. 106-107; Rolnictwo w 2003 r., GUS, Warszawa 2004, s. 109; Rolnictwo w 2001 r., GUS, Warszawa 2002, s. 109.

13. WNIOSKOWANIE STATYSTYCZNE

13.1. Wnioskowanie statystyczne – podstawowe pojęcia

Z wnioskowaniem statystycznym student zapoznaje się na przedmiocie *Statystyka matematyczna* (studia II stopnia). Tymczasem przedmiot *badania marketingowe* (na którym można wykorzystać elementy statystyki matematycznej) pojawia się już na studiach I stopnia. Poniżej zaprezentowano wybrane elementy statystyki matematycznej¹²². Liczne przykłady weryfikacji hipotez znajdują się w pracy prof. S. Mynarskiego: S. Mynarski, *Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2003. Przykłady zastosowania wnioskania statystycznego w badaniach stopnia zadowolenia pasażerów z funkcjonowania komunikacji zbiorowej w Krakowie można znaleźć np. w pracy A. Ciastoń, *Statystyczne porównanie wyników badań stopnia zadowolenia pasażerów z funkcjonowania komunikacji zbiorowej w Krakowie*¹²³.

Wnioskowanie statystyczne polega na uogólnianiu wyników uzyskanych dla próby na całą populację generalną z wykorzystaniem metod probabilistycznych. Reprezentatywność próby¹²⁴ ma zatem decydujące znaczenie dla jakości wyciągniętych wniosków. Metody wnioskania statystycznego obejmują estymację parametrów zbiorowości generalnej oraz weryfikację hipotez statystycznych.

Proces weryfikacji hipotezy przebiega według pewnego schematu postępowania, nazywanego testem statystycznym (por. rysunek). Weryfikując daną hipotezę ponosimy zawsze pewne ryzyko podjęcia błędnej decyzji (por. tabela 34).

¹²² Zainteresowanych czytelników zachęcamy do sięgnięcia na przykład do pozycji S. Ostasiewicz, Z. Rusnak, U. Siedlecka, *Statystyka. Elementy teorii i zadania*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1997; W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach, tom II*, Wydawnictwo Naukowe PWN, Warszawa 1995.

¹²³ http://www.transport.wshe.lodz.pl/documents/statystyczne_porownanie.pdf

¹²⁴ Stopień w jakim próba reprezentuje populację generalną w celu uogólnienia na nią prawidłowości zaobserwowanych w próbie.

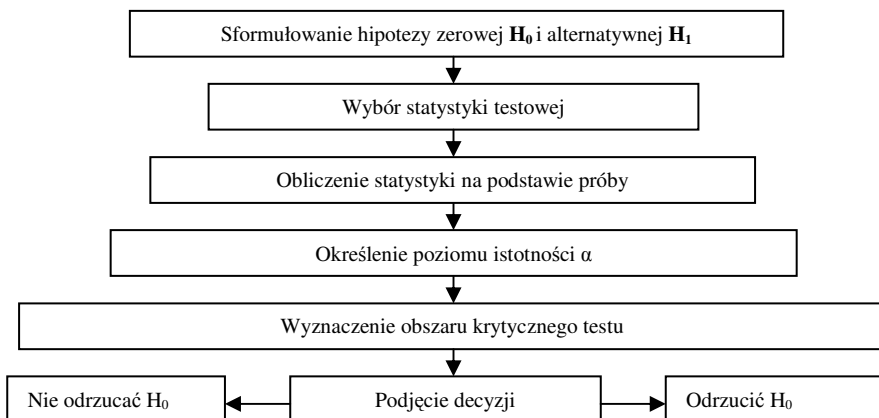
Tabela 34.

Możliwe decyzje w procesie testowania hipotez

	PRAWDA (badacz niestety jej nie zna)	
WNIOSEK badacza na podstawie badania marketingowego	Hipoteza H_0 jest prawdziwa	Hipoteza H_0 jest fałszywa
Brak podstaw, aby odrzucić weryfikowaną hipotezę H_0	TAK – decyzja poprawna	BŁĄD – błąd II rodzaju Prawdopodobieństwo $=\beta$
Odrzucić weryfikowaną hipotezę H_0	BŁĄD – błąd I rodzaju Prawdopodobieństwo $=\alpha$	TAK – decyzja poprawna Moc testu $=1 - \beta$

Źródło: opracowanie własne na podstawie W. Kryszicki, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach*, tom II, Wydawnictwo Naukowe PWN, Warszawa 1995, s. 79 oraz <http://inquiry.com.pl>.

Hipotezą statystyczną nazywamy każde przypuszczenie dotyczące nieznanego rozkładu badanej cechy populacji, o prawdziwości lub fałszywości, które wnioskuje się na podstawie pobranej próbki. **Hipotezą zerową H_0** nazywamy hipotezę sprawdzaną (testowaną, weryfikowaną). **Hipotezą alternatywną H_1** nazywamy hipotezę, którą jesteśmy skłonni przyjąć, gdy odrzucamy hipotezę H_0 .



Rys. 26. Przebieg procedury weryfikacyjnej

Źródło: opracowanie własne na podstawie A. Balicki, W. Makać, *Reguły wnioskowania statystycznego*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2000, s. 136.

Poziomem istotności α nazywamy prawdopodobieństwo popełnienia błędu I rodzaju. Przyjmujemy je arbitralnie (najczęściej $\alpha=0,01$ lub $\alpha=0,05$ lub $\alpha=0,1$). **Sprawdzianem hipotezy** nazywamy taką statystykę T (o znanym rozkładzie), której wartość t policzona na podstawie próby losowej, pozwala na podjęcie decyzji czy odrzucić hipotezę H_0 .

Zbiorem (obszarem) krytycznym nazywamy zbiór tych wartości sprawdzianu hipotezy, które przemawiają za odrzuceniem hipotezy H_0 . Rozkład sprawdzianu hipotezy określa, z jakich tablic należy odczytać wartość krytyczną wyznaczającą obszar krytyczny. Obszar krytyczny może być w zależności od hipotezy alternatywnej zbiorem jednostronnym (prawostronnym lub lewostronnym) bądź zbiorem dwustronnym.

Testy statystyczne są zróżnicowane ze względu na:

- rodzaj porównywanych charakterystyk
- skale pomiaru
- liczbę badanych prób
- zależność bądź niezależność w obrębie prób.

Hipotezy, które dotyczą wyłącznie wartości parametru określonej klasy rozkładów nazywamy **hipotezami parametrycznymi**. Każdą hipotezę, która nie jest parametryczna, nazywamy **nieparametryczną**.

Wykorzystanie parametrycznych testów statystycznych do opracowywania wyników badań naukowych jest ograniczone określonymi założeniami (zmienne mierzalne o rozkładzie normalnym, jednorodność zbioru itd.). Warunkiem użycia tych testów jest więc sprawdzenie założeń. Jeśli nie zostały one spełnione, wyciągnięte wnioski nie są w pełni poprawne lub tracą wiarygodność. Testy te stają się też bezużyteczne dla danych jakościowych i danych typu porządkowego. W tych wszystkich przypadkach stosuje się odpowiednie testy nieparametryczne¹²⁵.

W przypadku pojedynczej próby losowej testy statystyczne ograniczają się w zasadzie do porównywania zgodności charakterystyk z próby z odpowiednimi parametrami populacji (testy parametryczne) bądź do porównywania rozkładów empirycznych z próby z odpowiednimi rozkładami hipotetycznymi w populacji (testy nieparametryczne). W przypadku większej liczby prób, ze względu na charakter zależności między próbami, tworzy się znacznie większe zróżnicowanie w grupie testów parametrycznych i nieparametrycznych.

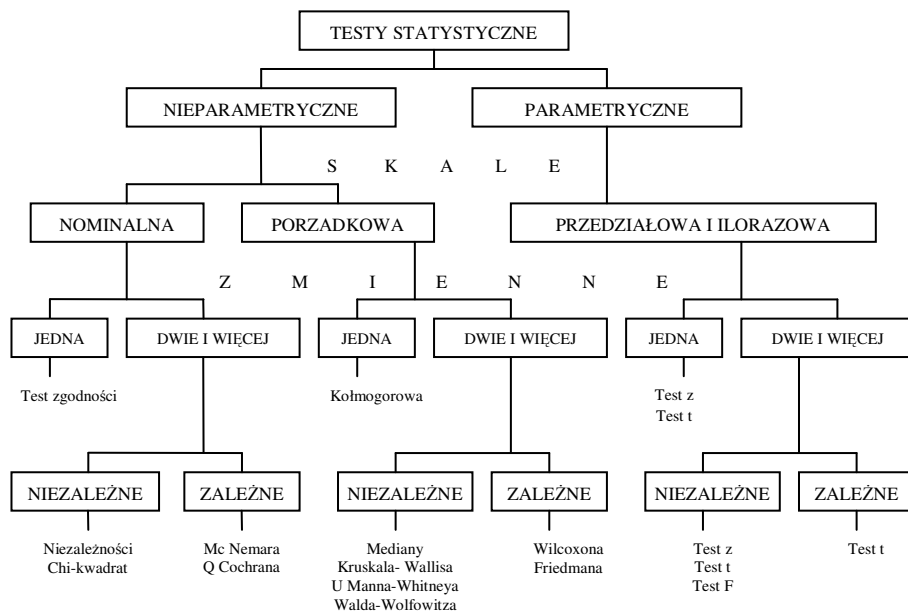
Próby niezależne to próby nie związane ze sobą żadnymi relacjami czasowymi bądź przestrzennymi. Najczęściej są to próby różnicowe, na których dokonuje się pomiarów jednokrotnych i niezależnych w czasie i

¹²⁵ Zob. S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza Fogra, Kraków 1995, s. 55-61; A. Stanisławski, *Przystępny kurs statystyki w oparciu o program STATISTICA PL na przykładach z medycyny*, Kraków 1998, s. 263-292.

w przestrzeni. W odróżnieniu od prób niezależnych, próby zależne są powiązane ze sobą pewnymi relacjami. Na próbach tych dokonuje się najczęściej pomiarów wielokrotnych (powtarzalnych w różnych odcinkach czasu).

13.2. Systematyka testów

Systematykę testów, uwzględniającą poszczególne kryteria (rodzaj porównywanych charakterystyk, skale pomiaru, liczbę badanych prób, zależność bądź niezależność w obrębie prób) przedstawia rysunek 27. Kolejne podrozdziały zawierają omówienie wybranych testów parametrycznych i nieparametrycznych.



Rys. 27. Podział testów statystycznych

Źródło: S. Mynarski, *Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2003, s. 21.

Testy statystyczne mogą być szeroko stosowane w analizach danych uzyskanych w badaniach ankietowych¹²⁶. W przypadku pojedynczej próby losowej w grę wchodzi testy zgodności rozkładu zmiennej losowej, a mianowicie testy nieparametryczne: test zgodności *chi-kwadrat* i *test Kołmogorowa* (Λ) oraz testy parametryczne: *test z* i *test t*. Testy te znajdują szerokie zastosowanie między innymi w badaniach jednorodności próby losowej pochodzącej np. od różnych ankietów, rozkładów charakterystyk z próby, np. dochodów i wydatków wśród respondentów, natężenia zmiennej w próbie, np. nurt nabywców w sklepie itp.

W przypadku dwóch prób losowych niezależnych mamy do czynienia z kilkoma testami nieparametrycznymi: testem serii dla cech jakościowych mierzonych na skali nominalnej (w badaniach losowości pojawiania się cech korzystnych i niekorzystnych), testem Walda-Wolfowitza dla cech ilościowych mierzonych na skali porządkowej (do badania uszeregowania rangowego obrotów sklepów położonych w centrum miasta i na peryferiach), testem U Manna-Whitneya dla rang łączonych (w badaniach rangowanych ocen, gdy niektórym z nich przypisuje się te same wagi) i testem Kołmogorowa-Smirnowa dla rozkładów kumulacyjnych (przy porównywaniu struktur preferencji), a także z dwoma testami parametrycznymi: testem *z* i testem *t* dla dwóch prób niezależnych. Wszystkie te testy pozwalają na stwierdzenie, czy dwie próby niezależne pochodzą z tej samej populacji.

W przypadku trzech i więcej prób losowych niezależnych znajdują zastosowanie następujące testy nieparametryczne: test Kruskala-Wallisa i test Mediany dla rang łączonych (np. w badaniach zakupów różnych marek produktów w różnych typach sklepów) oraz test niezależności Chi-kwadrat dla rozkładów częstości, a także testy parametryczne: test *z*, test *t* i test *F* (ilorazu wariancji).

W przypadku dwóch prób losowych zależnych znajdują zastosowanie następujące testy nieparametryczne: test znaków dla cech jakościowych mierzonych na skali nominalnej, test Mc Nemara dla rozkładu częstości w tablicy dwudzielczej czteropolowej oraz test Wilcoxon dla rangowanych znaków. Najczęstszymi przypadkami zastosowania wszystkich wymienionych testów są badania skuteczności środków promocji, polegające na porównywaniu obrotów przed i po zastosowaniu odpowiedniego środka reklamowego.

¹²⁶ Polecamy zwłaszcza pozycje prof. S. Mynarskiego, np. S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza Fogra, Kraków 1995; S. Mynarski, *Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2003.

W przypadku trzech i więcej prób losowych zależnych znajdują zastosowanie następujące testy nieparametryczne: test Q Cochra dla cech jakościowych oraz test Friedmana (z rangami) dla cech ilościowych. Obydwa testy pozwalają na badanie skuteczności działań promocyjnych nie tylko przed i po ich stosowaniu, ale również w trakcie trwania.

13.3. Testy parametryczne dla wnioskowania o własnościach populacji jednowymiarowej – testy dla średniej

W testach tych hipoteza H_0 jest hipotezą „o równości”. Hipoteza alternatywna, w zależności od pytania sformułowanego w zadaniu, może być prostym zaprzeczeniem hipotezy H_0 , może być hipotezą „o większości” lub „o mniejszości”.

$H_0: m=m_0$

$H_1: m \neq m_0$ lub $H_1: m < m_0$ lub $H_1: m > m_0$

a) Populacja generalna ma rozkład $N(m, \sigma)$; σ - znane

Obliczamy statystykę testową

$$t = \frac{\bar{X} - m_0}{\sigma} \sqrt{n}$$

Porównujemy ją z wartością krytyczną (lub wartościami krytycznymi) odczytaną z tablic standaryzowanego rozkładu normalnego, przy założonym poziomie istotności α .

Dla **$H_1: m \neq m_0$** wartość statystyki obliczona z próby $t \leq -u_\alpha$ lub $t \geq u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny dwustronny).

Dla **$H_1: m < m_0$** wartość statystyki obliczona z próby $t \leq -u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny lewostronny).

Dla **$H_1: m > m_0$** wartość statystyki obliczona z próby $t \geq u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny prawostronny).

b) Rozkład populacji – dowolny, wariancja nie jest znana, próba – duża

Obliczamy statystykę testową

$$t = \frac{\bar{X} - m_0}{S} \sqrt{n}$$

Porównujemy ją z wartością krytyczną (lub wartościami krytycznymi) odczytaną z tablic standaryzowanego rozkładu normalnego, przy założonym poziomie istotności α .

Dla $H_1: m \neq m_0$ wartość statystyki obliczona z próby $t \leq -u_\alpha$ lub $t \geq u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny dwustronny).

Dla $H_1: m < m_0$ wartość statystyki obliczona z próby $t \leq -u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny lewostronny).

Dla $H_1: m > m_0$ wartość statystyki obliczona z próby $t \geq u_\alpha$ powoduje odrzucenie H_0 (obszar krytyczny prawostronny).

c) Populacja generalna ma rozkład $N(m, \sigma)$; σ – nie jest znane, próba – mała

Obliczamy statystykę testową

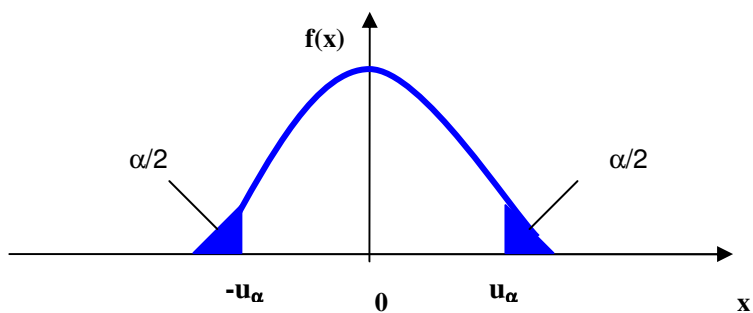
$$t = \frac{\bar{X} - m_0}{S} \sqrt{n-1} = \frac{\bar{X} - m_0}{S_1} \sqrt{n}$$

Porównujemy ją z wartością krytyczną (lub wartościami krytycznymi) odczytaną z tablic rozkładu t-Studenta dla poziomu istotności α oraz $n-1$ stopni swobody.

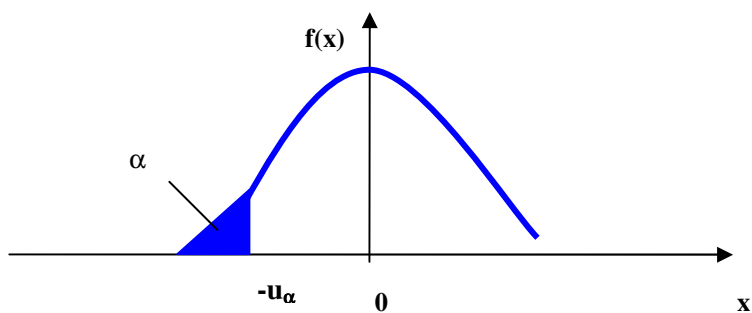
Dla $H_1: m \neq m_0$ wartość statystyki obliczona z próby $t \leq -t_{\alpha, n-1}$ lub $t \geq t_{\alpha, n-1}$ powoduje odrzucenie H_0 (obszar krytyczny dwustronny).

Dla $H_1: m < m_0$ wartość statystyki obliczona z próby $t \leq -t_{\alpha, n-1}$ powoduje odrzucenie H_0 (obszar krytyczny lewostronny).

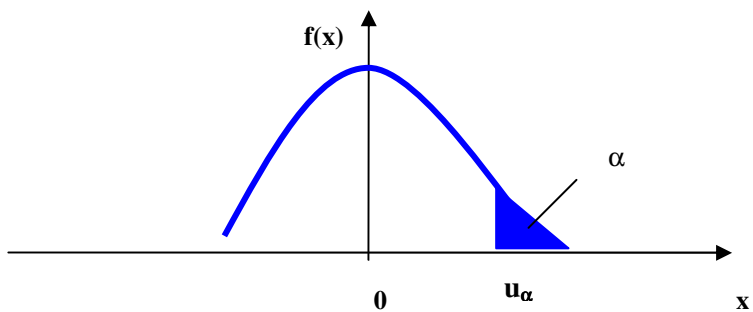
Dla $H_1: m > m_0$ wartość statystyki obliczona z próby $t \geq t_{\alpha, n-1}$ powoduje odrzucenie H_0 (obszar krytyczny prawostronny).



dwustronny obszar krytyczny



lewostronny obszar krytyczny



prawostronny obszar krytyczny

Rys. 28. Obszary krytyczne dla różnych postaci hipotezy alternatywnej (rozkład normalny standaryzowany)

Źródło: opracowanie własne.

13.4. Test niezależności χ^2

Test niezależności χ^2 pozwala na sprawdzenie, czy dwie badane cechy (niekoniecznie mierzalne) są niezależne. Wymogiem jest duża liczebność próby, której wyniki zostały podzielone na odpowiednie grupy wartości (kategorie) ze względu na obie cechy jednocześnie¹²⁷. Sporządza się zatem tzw. tablicę wielodzzielczą (lub inaczej tablicę niezależności / kontyngencji).

H_0 : X i Y są cechami niezależnymi

H_1 : X i Y są cechami zależnymi

Ocena skojarzenia cech opiera się na statystyce χ^2 , która pokazuje odchylenie zaobserwowanych liczebności dla wyodrębnionych klas obu cech od liczebności, których należałoby oczekiwać, gdyby cechy były niezależne. Statystykę χ^2 oblicza się dla tablicy o wymiarach $r \times s$, która powstaje w wyniku grupowania badanej zbiorowości według dwóch cech i składa się z r wierszy odpowiadających wariantom jednej cechy oraz s kolumn odpowiadających wariantom drugiej cechy (najmniejsze wymiary tablicy wynoszą 2×2 – jest to tzw. tablica czteropolowa). Ze względu na wymaganą liczebność w każdej klasie, zachodzi czasem konieczność połączenia za pomocą spójnika „lub” dwóch kategorii danej cechy w jedną.

Tabela 35.

Możliwe decyzje w procesie testowania hipotez – tablica wielodzzielcza

$X_i \backslash Y_j$	y_1	y_2	...	y_s	$n_{i.}$
x_1	n_{11}	n_{12}	...	n_{1s}	$n_{1.}$
x_2	n_{21}	n_{22}	...	n_{2s}	$n_{2.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_r	n_{r1}	n_{r2}	...	n_{rs}	$n_{r.}$
$n_{.j}$	$n_{.j}$	$n_{.j}$...	$n_{.j}$	n

Źródło: opracowanie własne.

¹²⁷ Należy jednak pamiętać, że klasy, na jakie dzieli się wyniki próby, powinny być odpowiednio liczne. Większość autorów określa licznosc klas na „co najmniej 5”, ale J. Greń podaje już „co najmniej 8”. Por. J. Greń, Modele i zadania statystyki matematycznej, PWN, Warszawa 1970, s. 104; A. Luszniwicz, T. Słaby, Statystyka z pakietem komputerowym STATISTICATM PL, Wydawnictwo C.H. BECK, Warszawa 2001, s. 210.

Prawdopodobieństwa brzegowe: $p_{i.}=n_{i.}/n$ oraz $p_{.j}=n_{.j}/n$

Prawdopodobieństwo $p_{ij}=p_{i.}p_{.j}$

Dla każdego pola tablicy wyznacza się liczebności teoretyczne

$$\hat{n}_{ij} = np_{ij} = \frac{n_{i.} \cdot n_{.j}}{n}$$

Po ustaleniu liczebności teoretycznych, zakładających niezależność cech, statystykę oblicza się według wzoru:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$$

Obszar krytyczny (prawostronny) w tym teście określa relacja $P\{\chi^2 \geq \chi^2_{\alpha}\} = \alpha$, gdzie χ^2_{α} jest wartością krytyczną odczytaną z tablicy rozkładu χ^2 dla danego poziomu istotności α oraz dla $(r-1)(s-1)$ stopni swobody.

Jeżeli $\chi^2 \geq \chi^2_{\alpha}$, to hipotezę H_0 o niezależności badanych cech należy odrzucić na rzecz hipotezy alternatywnej, natomiast gdy $\chi^2 < \chi^2_{\alpha}$, to nie ma podstaw do odrzucenia hipotezy H_0 .

Miarą zależności jest współczynnik zbieżności T-Czuprowa (bazujący na wartości statystyki χ^2) postaci:

$$T = \sqrt{\frac{\chi^2}{n\sqrt{(r-1)(s-1)}}$$

Gdy $r=s$ to $0 \leq T \leq 1$. Gdy $r \neq s$ to T może być znacznie mniejsze od 1.

13.5. Przykłady z omówieniem

Przykład 1. Przykłady hipotez

H_0 : Opakowanie kawy A jest równie atrakcyjne jak opakowanie B

H_1 : Opakowanie kawy A jest bardziej atrakcyjne niż opakowanie B

H_0 : Średnie miesięczne wydatki na zakup kawy wynoszą 20 zł

H_1 : Średnie miesięczne wydatki na zakup kawy przewyższają 20 zł

H_0 : Wygląd opakowania kawy nie zależy od ceny kawy

H_1 : Wygląd opakowania kawy zależy od ceny kawy

H_0 : $m_1 = m_2$ (program lojalnościowy nie wpływa na wielkość konsumpcji kawy)

H_1 : $m_1 < m_2$ (program lojalnościowy zwiększa wielkość konsumpcji kawy)

gdzie m_1 – średnie spożycie kawy przed udziałem w programie lojalnościowym, m_2 – średnie spożycie kawy po udziale w programie lojalnościowym

Przykład 2. Według danych Agencji Konsumentów Amerykanie wrzucają do Coca-coli przeciętnie 25 kostek lodu. Sformułowano przypuszczenie, że w 2010 roku norma ta zostanie przekroczone. W celu weryfikacji hipotezy przebadano 9 losowo wybranych obywateli USA. Dla próby uzyskano średnią 2,5 kostki. Wiadomo, że rozkład liczby wrzucanych kostek lodu jest rozkładem normalnym o odchyleniu standardowym $\sigma=1$. Sformułować hipotezę zerową i hipotezę badawczą Przeprowadzić weryfikację na poziomie istotności $\alpha=0,05$.

Rozwiązanie:

$$H_0: m=5$$

$$H_a: m>5$$

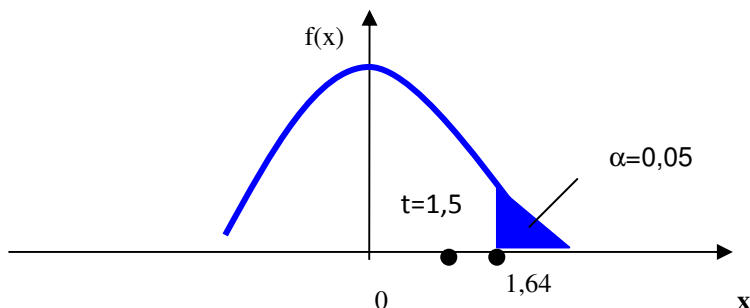
Populacja generalna ma rozkład $N(m, \sigma)$; σ - znane ($\sigma=1$); $m_0=2$; $n=9$; $\alpha=0,05$

$$\bar{x} = 2,5;$$

Obliczamy wartość statystyki testowej:

$$t = \frac{2,5 - 2}{1} \sqrt{9} = 1,5$$

Z tablic standaryzowanego rozkładu normalnego odczytujemy wartość krytyczną (przy założonym poziomie istotności $\alpha=0,05$) 1,64. Obszar krytyczny jest prawostronny (por. rys. 29)



Rys. 29. Obszar krytyczny

Źródło: opracowanie własne.

Wniosek: nie ma podstaw do odrzucenia hipotezy H_0 , a zatem nie można twierdzić, że w 2010 r. liczba kostek wrzucanych do Coca-coli zostanie przekroczona.

Przykład 3. Przeprowadzono badania zależności między ceną a jakością ręcznych wyciskarek do soków. Wyniki badań przedstawia tabela 36.

Tabela 36.

Dane do przykładu

Cena	Jakość	
	Dobra	Zła
60,00 zł i mniej	40	160
60,01-100,00 zł	50	50
100,01 zł i więcej	90	10

Źródło: opracowanie własne, dane umowne.

Na poziomie istotności $\alpha=0,05$ zweryfikować hipotezę, że cena wyciskarki zależy od jakości.

Rozwiązanie:

H_0 : cena wyciskarki nie zależy od jakości

H_a : cena wyciskarki zależy od jakości

Tabela 37.

Obliczenia pomocnicze – obliczenia liczebności teoretycznych

Cena	Jakość		n_i
	Dobra	Zła	
60,00 zł i mniej	40	160	200
60,01-100,00 zł	50	50	100
100,01 zł i więcej	90	10	100
n_j	180	220	400

Cena	Jakość	
	Dobra	Zła
60,00 zł i mniej	90	110
60,01-100,00 zł	45	55
100,01 zł i więcej	45	55

Źródło: opracowanie własne.

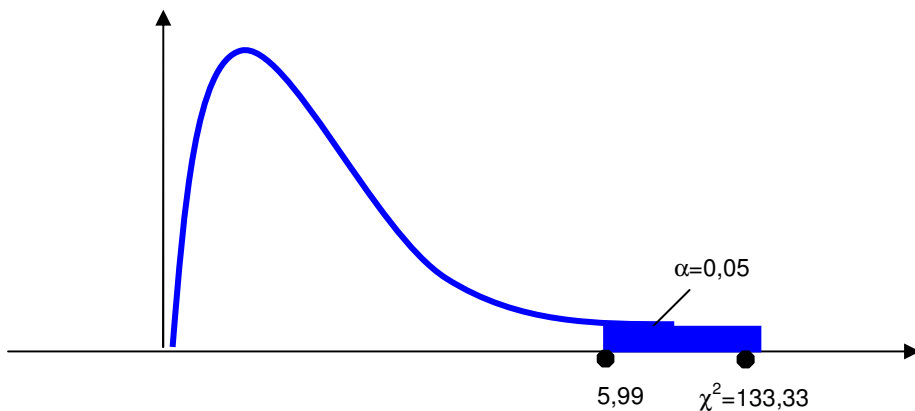
Tabela 38.
Obliczenia pomocnicze – obliczenie statystyki χ^2

Cena	Jakość	
	Dobra	Zła
60,00 zł i mniej	27,78	22,73
60,01-100,00 zł	0,56	0,45
100,01 zł i więcej	45,00	36,82

Źródło: opracowanie własne.

Wartość statystyki testowej: $\chi^2=133,33$.

Z tablic rozkładu chi-kwadrat, dla $(r-1) \cdot (s-1) = (3-1) \cdot (2-1) = 2$ stopni swobody oraz dla założonego poziomu istotności $\alpha=0,05$, odczytujemy wartość krytyczną $\chi^2_{0,05}=5,99$. Obszar krytyczny jest prawostronny (por. rysunek 30).



Rys. 30. Obszar krytyczny

Źródło: opracowanie własne.

Wniosek: Hipotezę H_0 odrzucamy na rzecz hipotezy H_a , a zatem cena wyciskarki zależy od jej jakości.

13.6. Zadania do samodzielnego opracowania

Zadanie 1. Firma produkująca napoje energetyczne martwi się o spadek udziału w rynku. Napoje tej firmy mają stosunkowo wysoką cenę. Wszystkie składniki napojów (poza wodą i cukrem) sprowadzane są z Chin. Napoje produkowane są w zakładzie w Austrii i stamtąd rozprowadzane na całą Europę. Wskaż przykłady hipotez, jakie mogą wystąpić w badaniu.

Zadanie 2. Przeprowadzono badanie wpływu informacji marketingowej (promocja supermarketu) na wybór miejsca zakupu. Wyniki badania prezentuje tabela.

Tabela 39.

Wyniki badania

Grupa	Preferowane miejsce zakupu				Razem
	Hipermarket	Supermarket	Sklep osiedlowy	Targowisko	
Bez informacji marketingowej	300	200	100	200	800
Z informacją marketingową	200	600	50	150	900
Razem	500	700	150	350	1700

Źródło: dane umowne.

Czy promocja ma wpływ na wybór preferowanego miejsca zakupu?
Przeprowadź wnioskowanie na poziomie ufności 95%.

Zadanie 3. Wydawca książek próbuje ustalić, czy popularność danej książki uwarunkowana jest specyfiką rynku regionalnego. Przeprowadź wnioskowanie na poziomie ufności 90%. Wyniki badania prezentuje tabela.

Tabela 40.

Wyniki badania

Książka \ Miasto	Opole	Wrocław	Katowice
Książka A	81	60	182
Książka B	78	93	95
Książka C	241	247	123

Źródło: dane umowne.

Zadanie 4. Pracownik działu kredytów pewnego banku uważa, iż na zakup samochodu klienci biorą przeciętnie kredyt w wysokości 20 tys. zł, natomiast pracownik salonu samochodowego uważa, że jest to mniej niż 20 tys. zł. W losowo wybranej próbie złożonej z dziewięciu osób okazało się, iż kredyty na zakup samochodu zaciągano na kwoty: 24 tys. zł, 22,9 tys. zł, 11,1 tys. zł, 13,4 tys. zł, 13,8 tys. zł, 9,3 tys. zł, 18,4 tys. zł, 8,2 tys. zł, 14,6 tys. zł. Przeprowadź wnioskowanie na poziomie ufności 95%.

14. SEGMENTACJA Z WYKORZYSTANIEM METODY TAKSONOMII WROCŁAWSKIEJ

14.1. Klasyfikacja obiektów

Ustalanie segmentów rynku (wyodrębnienie homogenicznych wewnątrznie i zróżnicowanych zewnętrznie grup) ma istotne znaczenie dla określenia skali i profilu działalności przedsiębiorstwa. Efektywna segmentacja, pozwalająca precyzyjnie zidentyfikować docelowe grupy konsumentów, jest warunkiem osiągnięcia sukcesu. Znajomość kryteriów segmentów rynku jest podstawowym elementem wiedzy umożliwiającym przeprowadzenie badań jej dotyczących. Efektywność segmentacji zależy od mierzalności, opłacalności, dostępności i homogeniczności segmentów.

Przez **klasyfikację** rozumie się potocznie systematyczny podział przedmiotów, zjawisk na klasy, działy, poddziały dokonywany według określonej zasady. Badane zjawiska bywają często złożonej natury. Przedmiotem klasyfikacji są zbiory obserwacji (obiektów) opisanych zwykle wieloma cechami. W praktyce, badając zjawiska złożone (opisywane za pomocą wielu pojedynczych cech diagnostycznych), zdarza się, że wartości zmiennych wyrażone są np. w różnych jednostkach miary lub wykazują znaczne zróżnicowanie. Jednym z głównych celów statystyki jest uproszczenie złożoności. Zebrane dane statystyczne często zatem podlegają przekształceniom, w celu ułatwienia analizy.

Wielowymiarowa analiza porównawcza (WAP) polega na przeprowadzeniu wielowymiarowych badań porównawczych, które pozwolą wykryć prawidłowości w zbiorowościach statystycznych, w których jednostki są opisane przez stosunkowo liczny zespół cech. Badania porównawcze znajdują liczne zastosowania na przykład w badaniach społeczno-gospodarczych, ponieważ umożliwiają między innymi:

- ✓ hierarchizację obiektów opisanych wieloma cechami z punktu widzenia pewnej charakterystyki, której nie można zmierzyć w sposób bezpośredni (np. przydatność poszczególnych wariantów lokalizacyjnych, standard życia, poziom rozwoju społeczno-gospodarczego)
- ✓ wyodrębnienie homogenicznych podzbiorów, zawierających obiekty podobne do siebie z punktu widzenia pewnej charakterystyki, a jednocześnie niepodobne do obiektów tworzących pozostałe podzbiory.

Klasyfikacja hierarchiczna polega na liniowym uporządkowaniu zbioru obiektów według określonej charakterystyki, która nie podlega bezpośredniemu pomiarowi, ale może być opisana za pomocą co najmniej dwóch cech. Klasyfikacja ta porządkuje obiekty od „najlepszego” do „najgorszego” (lub odwrotnie) na podstawie syntetycznego miernika, który jest agregatem informacji zawartych w poszczególnych cechach. **Klasyfikacja niehierarchiczna** polega na podziale niejednorodnego zbioru

obiektów na pewną liczbę klas, w których znajdują się obiekty podobne do siebie ze względu na wartości cech.

14.2. Normalizacja

Normowanie jest działaniem mającym na celu przysposobienie zmiennych diagnostycznych do roli kryteriów cząstkowych w procesie oceny zjawiska złożonego. Uwzględniając potrzebę pozbycia mian oraz ujednoczenia zakresów liczbowych zmiennych diagnostycznych, metody normujące służą transformacji bezwzględnych wartości na wartości względne. W tym sensie każda z metod normowania (oprócz metody rangowej) jest pewnym przekształceniem ilorazowym, które w końcowym wyniku daje zmienną diagnostyczną transformowaną. Tradycyjny podział procedur normowania obejmuje: metody standaryzacji, metody unitaryzacji, przekształcenia ilorazowe, metody rangowe.

Metody standaryzacyjne stanowią taką formę przekształcenia ilorazowego, w której wartości cechy normowanej (lub wartości te same cechy pomniejszonej o jej średnią) są odnoszone do wartości odchylenia standardowego cechy:

a) metoda przedstawiona w pracy J. Perkała¹²⁸

$$z_{ij} = \frac{x_{ij} - \bar{X}_j}{S(X_j)} \quad \text{gdy } X_j \in S$$

$$z_{ij} = \frac{\bar{X}_j - x_{ij}}{S(X_j)} \quad \text{gdy } X_j \in D$$

S – stymulanta, D – destymulanta

b) metoda przedstawiona przez M. Cieślak¹²⁹

$$z_{ij} = \frac{x_{ij}}{S(X_j)} \quad \text{gdy } X_j \in S$$

$$z_{ij} = -\frac{x_{ij}}{S(X_j)} \quad \text{gdy } X_j \in D$$

¹²⁸ J. Perkał, O wskaźnikach antropologicznych, [w:] Przegląd Antropologiczny 1953, t. 19.

¹²⁹ M. Cieślak, Taksonomiczna procedura programowania rozwoju gospodarczego i określania zapotrzebowania na kadry kwalifikowane, [w:] Przegląd Statystyczny 1974, z. 1.

c) inne metody np.

$$z_{ij} = \frac{x_{ij}}{S(X_j)} \quad \text{gdy } X_j \in S$$

$$z_{ij} = \frac{S(X_j)}{x_{ij}} \quad \text{gdy } X_j \in D$$

Metody unitaryzacyjne charakteryzują się przyjęciem stałego punktu odniesienia, który stanowi rozstęp zmiennej normowanej

$$R(X_j) = \max_i x_{ij} - \min_i x_{ij}$$

Takie podejście sprawia, że rozstęp cechy unormowanej Z_j jest stały i wynosi 1.

np.
$$z_{ij} = \frac{x_{ij}}{R(X_j)} \quad \max_i x_{ij} \neq \min_i x_{ij}$$

Metodę rangową stosuje się na ogół do porządkowania obiektów ze względu na zmienne jakościowe mierzone na skali porządkowej. Informacja o kolejności obiektów nie umożliwia jednakże określenia dokładnych różnic dzielących te objekty.

14.3. Metoda taksonomii wrocławskiej

Na początku XX wieku w naukach przyrodniczych pojawiły się numeryczne metody klasyfikacji określane mianem **metod taksonomicznych**¹³⁰. Inicjatorem badań taksonomicznych i stosowania metod numerycznych w antropologii był polski uczoney J. Czekanowski¹³¹. Duże znaczenie dla rozwoju metod taksonomicznych mają również prace H. Steinhausa i kierowanego przez niego zespołu matematyków wrocławskich. Metoda, znana pod nazwą **taksonomii wrocławskiej**, jest do dziś powszechnie stosowana nie tylko w Polsce, ale i na świecie.

Metoda taksonomii wrocławskiej należy do niehierarchicznych procedur taksonomicznych, w których wyniki grupowania przedstawia się w postaci dendrytu, zdefiniowanego jako graf spójny i otwarty. Celem klasyfikacji niehierarchicznej jest podział niejednorodnego zbioru obiektów na pewną liczbę klas, w której znajdują się objekty do siebie najbardziej

¹³⁰ Ich nazwa pochodzi od dwóch greckich słów: taxis – układ, porządek; nomos – prawo, zasada.

¹³¹ J. Czekanowski, analizując populację Pigmejów, dążył do wydzielenia podgrup osobników tak, by w obrębie podgrupy zachodziły jak najmniejsze różnice między poszczególnymi osobnikami. Opracowaną w 1909 r. metodę określa się mianem metody różnic przeciętnych.

podobne, a w różnych klasach obiekty do siebie najmniej podobne (ze względu na wartości cech). Do określania stopnia podobieństwa obiektów wykorzystuje się najczęściej miary odległości.

Schemat postępowania w metodzie taksonomii wrocławskiej jest następujący:

1. Wybór zbioru obiektów. Wybór cech opisujących problem badawczy.
2. Konstrukcja macierzy danych X .
3. Doprowadzenie wartości cech do porównywalności poprzez standaryzację (macierz znormalizowanych wartości).
4. Budowa macierzy odległości

$$D = \begin{bmatrix} 0 & d_{12} & \dots & d_{1n} \\ d_{21} & 0 & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}$$

gdzie

$$d_{kl} = \sqrt{\sum_{j=1}^p (z_{jk} - z_{jl})^2}$$

5. W każdym wierszu ($k=1, 2, \dots, n$) macierzy D szuka się elementu najmniejszego

$$d_i = \min_{k \neq i} \{d_{kl}\}$$

wskazując tym samym obiekty, które należy połączyć. Otrzymane w ten sposób połączenia tworzą na ogół graf niespójny, który nazywa się dendrytem pierwszego stopnia.

6. W dendrycie pierwszego stopnia znajduje się pary obiektów (należące do różnych podgrafów) o najmniejszej odległości. Łączy się te podgrafy, otrzymując w ten sposób dendryt drugiego stopnia.

Kontynuując to postępowanie można wyznaczyć dendryt kolejnych stopni, otrzymując w końcu pewien graf spójny łączący wszystkie obiekty.

14.4. Przykład z omówieniem

Założmy, że analizując cechy dwunastu nabywców proszku do prania wyróżniono: X_1 – średni miesięczny dochód na osobę w gospodarstwie domowym (w tys. zł) oraz X_2 – liczba osób w gospodarstwie domowym. W celu wyodrębnienia jednorodnych grup konsumentów można zastosować metodę taksonomii wrocławskiej. A profile wyznaczonych segmentów można opisać za pomocą średnich wyznaczonych dla klientów zgrupowanych w jeden segment (tabela 43). W trakcie obliczeń przechodzimy od

obliczenia średniej, odchylenia standardowego i współczynnika zmienności, poprzez standaryzację zmiennych (tabela 41), do macierzy odległości (tabela 42) i profili wyznaczonych segmentów (tabela 43).

Tabela 41.

Wyniki badania (dane umowne) wraz z cząstkowymi wynikami obliczeń

Lp.	Średni dochód na osobę (w tys. zł)	Liczba osób w gospodarstwie	Z ₁ (X ₁ po standaryzacji)	Z ₂ (X ₂ po standaryzacji)
1	5,0	1	0,26	-1,52
2	0,6	4	-0,52	0,27
3	7,0	2	0,62	-0,92
4	0,4	7	-0,56	2,07
5	1,0	3	-0,45	-0,33
6	1,6	4	-0,34	0,27
7	1,2	4	-0,41	0,27
8	0,7	4	-0,50	0,27
9	0,7	5	-0,50	0,87
10	20,0	1	2,94	-1,52
11	0,5	4	-0,54	0,27
12	2,0	3	-0,27	-0,33

$$\bar{x}_i = \quad 3,5 \quad 3,5$$

$$s_i = \quad 5,6 \quad 1,7$$

$$v_i = \quad 1,6 \quad 0,5$$

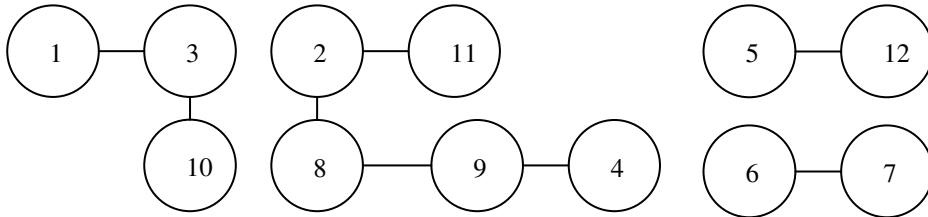
Źródło: dane umowne, obliczenia własne.

Tabela 42.

Macierz odległości

	1	2	3	4	5	6	7	8	9	10	11	12	d_i
1	0,000	1,960	0,697	3,683	1,394	1,895	1,919	1,952	2,514	2,680	1,967	1,311	0,697
2	1,960	0,000	1,655	1,795	0,603	0,179	0,107	0,018	0,599	3,903	0,018	0,649	0,018
3	0,697	1,655	0,000	3,216	1,228	1,537	1,583	1,643	2,119	2,398	1,668	1,075	0,697
4	3,683	1,795	3,216	0,000	2,396	1,808	1,801	1,796	1,198	5,015	1,795	2,410	1,198
5	1,394	0,603	1,228	2,396	0,000	0,608	0,599	0,601	1,198	3,599	0,605	0,179	0,179
6	1,895	0,179	1,537	1,808	0,608	0,000	0,071	0,161	0,620	3,746	0,197	0,603	0,071
7	1,919	0,107	1,583	1,801	0,599	0,071	0,000	0,089	0,605	3,808	0,125	0,615	0,071
8	1,952	0,018	1,643	1,796	0,601	0,161	0,089	0,000	0,598	3,887	0,036	0,642	0,018
9	2,514	0,599	2,119	1,198	1,198	0,620	0,605	0,598	0,000	4,197	0,599	1,219	0,598
10	2,680	3,903	2,398	5,015	3,599	3,746	3,808	3,887	4,197	0,000	3,919	3,431	2,398
11	1,967	0,018	1,668	1,795	0,605	0,197	0,125	0,036	0,599	3,919	0,000	0,656	0,018
12	1,311	0,649	1,075	2,410	0,179	0,603	0,615	0,642	1,219	3,431	0,656	0,000	0,179

Źródło: obliczenia własne.



Rys. 31. Dendryt I rzędu

Źródło: obliczenia własne.

Tabela 43.

Profile wyznaczonych segmentów

Segment	Średni dochód na osobę (w tys. zł)	Średnia liczba osób w gospodarstwie	Udział segmentu (%)
Segment 1 (1, 3, 10)	10,66	1,3	25,0
Segment 2 (2, 4, 8, 9, 11)	0,58	4,8	41,7
Segment 3	1,5	3	16,7
Segment 4	1,4	4	16,7

Źródło: obliczenia własne.

Tablice statystyczne

Tabela 44.

Wartości dystrybuanty rozkładu normalnego standaryzowanego $P(X < u_\alpha) = 1 - \alpha$

	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

38	0,013	0,025	0,063	0,127	0,255	0,388	1,051	1,304	1,686	2,024	2,429	2,712
39	0,013	0,025	0,063	0,126	0,255	0,388	1,050	1,304	1,685	2,023	2,426	2,708
40	0,013	0,025	0,063	0,126	0,255	0,388	1,050	1,303	1,684	2,021	2,423	2,704
50	0,013	0,025	0,063	0,126	0,255	0,388	1,047	1,299	1,676	2,009	2,403	2,678
60	0,013	0,025	0,063	0,126	0,254	0,387	1,045	1,296	1,671	2,000	2,390	2,660
70	0,013	0,025	0,063	0,126	0,254	0,387	1,044	1,294	1,667	1,994	2,381	2,648
80	0,013	0,025	0,063	0,126	0,254	0,387	1,043	1,292	1,664	1,990	2,374	2,639
90	0,013	0,025	0,063	0,126	0,254	0,387	1,042	1,291	1,662	1,987	2,368	2,632
100	0,013	0,025	0,063	0,126	0,254	0,386	1,042	1,290	1,660	1,984	2,364	2,626

Tabela 46.

Tablice rozkładu chi-kwadrat $P(X > \chi^2_{\alpha}) = \alpha$

$\alpha \backslash k$	0,99	0,98	0,95	0,9	0,8	0,2	0,1	0,05	0,02	0,01
1	0,000	0,001	0,004	0,016	0,064	1,642	2,706	3,841	5,412	6,635
2	0,020	0,040	0,103	0,211	0,446	3,219	4,605	5,991	7,824	9,210
3	0,115	0,185	0,352	0,584	1,005	4,642	6,251	7,815	9,837	11,345
4	0,297	0,429	0,711	1,064	1,649	5,989	7,779	9,488	11,668	13,277
5	0,554	0,752	1,145	1,610	2,343	7,289	9,236	11,070	13,388	15,086
6	0,872	1,134	1,635	2,204	3,070	8,558	10,645	12,592	15,033	16,812
7	1,239	1,564	2,167	2,833	3,822	9,803	12,017	14,067	16,622	18,475
8	1,646	2,032	2,733	3,490	4,594	11,030	13,362	15,507	18,168	20,090
9	2,088	2,532	3,325	4,168	5,380	12,242	14,684	16,919	19,679	21,666
10	2,558	3,059	3,940	4,865	6,179	13,442	15,987	18,307	21,161	23,209
11	3,053	3,609	4,575	5,578	6,989	14,631	17,275	19,675	22,618	24,725
12	3,571	4,178	5,226	6,304	7,807	15,812	18,549	21,026	24,054	26,217
13	4,107	4,765	5,892	7,042	8,634	16,985	19,812	22,362	25,472	27,688
14	4,660	5,368	6,571	7,790	9,467	18,151	21,064	23,685	26,873	29,141
15	5,229	5,985	7,261	8,547	10,307	19,311	22,307	24,996	28,259	30,578
16	5,812	6,614	7,962	9,312	11,152	20,465	23,542	26,296	29,633	32,000
17	6,408	7,255	8,672	10,085	12,002	21,615	24,769	27,587	30,995	33,409
18	7,015	7,906	9,390	10,865	12,857	22,760	25,989	28,869	32,346	34,805
19	7,633	8,567	10,117	11,651	13,716	23,900	27,204	30,144	33,687	36,191
20	8,260	9,237	10,851	12,443	14,578	25,038	28,412	31,410	35,020	37,566
21	8,897	9,915	11,591	13,240	15,445	26,171	29,615	32,671	36,343	38,932
22	9,542	10,600	12,338	14,041	16,314	27,301	30,813	33,924	37,659	40,289
23	10,196	11,293	13,091	14,848	17,187	28,429	32,007	35,172	38,968	41,638
24	10,856	11,992	13,848	15,659	18,062	29,553	33,196	36,415	40,270	42,980

25	11,524	12,697	14,611	16,473	18,940	30,675	34,382	37,652	41,566	44,314
26	12,198	13,409	15,379	17,292	19,820	31,795	35,563	38,885	42,856	45,642
27	12,879	14,125	16,151	18,114	20,703	32,912	36,741	40,113	44,140	46,963
28	13,565	14,847	16,928	18,939	21,588	34,027	37,916	41,337	45,419	48,278
29	14,256	15,574	17,708	19,768	22,475	35,139	39,087	42,557	46,693	49,588
30	14,953	16,306	18,493	20,599	23,364	36,250	40,256	43,773	47,962	50,892
31	15,655	17,042	19,281	21,434	24,255	37,359	41,422	44,985	49,226	52,191
32	16,362	17,783	20,072	22,271	25,148	38,466	42,585	46,194	50,487	53,486
33	17,074	18,527	20,867	23,110	26,042	39,572	43,745	47,400	51,743	54,776
34	17,789	19,275	21,664	23,952	26,938	40,676	44,903	48,602	52,995	56,061
35	18,509	20,027	22,465	24,797	27,836	41,778	46,059	49,802	54,244	57,342
36	19,233	20,783	23,269	25,643	28,735	42,879	47,212	50,998	55,489	58,619
37	19,960	21,542	24,075	26,492	29,635	43,978	48,363	52,192	56,730	59,893
38	20,691	22,304	24,884	27,343	30,537	45,076	49,513	53,384	57,969	61,162
39	21,426	23,069	25,695	28,196	31,441	46,173	50,660	54,572	59,204	62,428
40	22,164	23,838	26,509	29,051	32,345	47,269	51,805	55,758	60,436	63,691
50	29,707	31,664	34,764	37,689	41,449	58,164	63,167	67,505	72,613	76,154
60	37,485	39,699	43,188	46,459	50,641	68,972	74,397	79,082	84,580	88,379
70	45,442	47,893	51,739	55,329	59,898	79,715	85,527	90,531	96,388	100,425
80	53,540	56,213	60,391	64,278	69,207	90,405	96,578	101,879	108,069	112,329
90	61,754	64,635	69,126	73,291	78,558	101,054	107,565	113,145	119,648	124,116
100	70,065	73,142	77,929	82,358	87,945	111,667	118,498	124,342	131,142	135,807

Słownik najważniejszych pojęć

Analiza korelacji i regresji – dział statystyki zajmujący się badaniem związków i zależności pomiędzy rozkładami dwóch lub większej liczby cech w populacji generalnej¹³².

Ankieta – zestaw pytań kierowanych do respondenta, ułożonych w określonym porządku, sformułowanych zgodnie z kanonami wiedzy metodologicznej i przedstawiony osobie badanej w formie papierowego wydruku lub zaprezentowany mu w wersji komputerowej na ekranie monitora, w Internecie¹³³.

Ankieter – przeszkolony pracownik agencji badawczej lub firmy/instytucji prowadzącej badania, który zajmuje się pozyskiwaniem informacji od respondentów za pomocą opracowanych przez badaczy kwestionariuszy wywiadów¹³⁴.

Badania eksplanacyjne to inaczej badania opisowe i przyczynowo-skutkowe. Badania opisowe mają na celu odpowiedź na pytania: Co? Jak? Gdzie? Kto? Ich przeprowadzenie ma na celu scharakteryzowanie pewnych zjawisk oraz procesów. Badania przyczynowo-skutkowe mają na celu wyjaśnienie rozwoju i kształtowania się zjawisk przez określenie związków przyczynowo-skutkowych pomiędzy zmiennymi¹³⁵.

Badania eksploracyjne są to badania, „które mają na celu zidentyfikowanie natury problemu, formułowanie hipotez, które następnie będą wyrażone w formie pytań badawczych”¹³⁶. Badanie eksploracyjne polegają na rozpoznaniu i zidentyfikowaniu pewnych zjawisk.

Badania ilościowe – szeroka kategoria sondażowych i eksperymentalnych metod badawczych, które charakteryzują się zestawem pewnych podstawowych cech wspólnych:

- ukierunkowanie na ilościowy/liczbowy opis i wyjaśnianie rzeczywistości ustalanie jak często określone opinie, fakty, zjawiska występują w badanej zbiorowości;

¹³² J. Greń, Modele i zadania statystyki matematycznej, PWN, Warszawa 1970, s. 16.

¹³³ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010)

¹³⁴ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010)

¹³⁵ Badania przyczynowe koncentrują się na opisie relacji pomiędzy przyczyną danego zjawiska i jej skutkiem. Są prowadzone najczęściej w formie eksperymentu i pozwalają m.in. na dokonanie oceny zasadności wydatków na promocję czy też kształtowanie się popytu w wyniku zmiany ceny na dany produkt.

¹³⁶ Badania marketingowe. Podstawowe metody i obszary zastosowań, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1998, s.27.

- wykorzystywanie dużych prób respondentów dobranych w sposób losowy lub kwotowy, stanowiących reprezentację całej badanej zbiorowości/populacji;
- zastosowanie precyzyjnych metod statystycznych na etapie doboru próby i kalkulowania wyników badania;
- użycie wystandaryzowanych narzędzi badawczych kwestionariuszy, ankiet, schematów eksperymentalnych, które ujednolicają fazę kontaktu z respondentami, gwarantują, iż wszyscy badani poddani są działaniu bardzo zbliżonych bodźców;
- dążenie do generalizowania wyników pomiaru wobec zbiorowości liczniejszej niż tylko badana próba;
- ukierunkowanie na rozstrzygające konkluzje badawcze, jednoznaczne i konkretne odpowiedzi na pytania badawcze¹³⁷.

Badania jakościowe – ogólna kategoria rozmaitych metod i technik badawczych cechujących się następującymi wspólnymi właściwościami:

- skoncentrowanie na poznaniu, opisie, wyjaśnianiu różnorodności zjawisk, a nie na diagnozie częstości ich występowania;
- wykorzystanie w charakterze informatorów pojedynczych osób lub małych grup dobranych w sposób celowy podporządkowany wąskiemu zwykle tematowi badawczemu;
- użycie niewystandaryzowanych narzędzi badawczych ogólnego scenariusza wywiadu, elastycznych reguł dyskusji, indywidualnej obserwacji;
- podejmowanie pytań o charakterze eksploracyjnym, koncentracja na niewymiernych zagadnieniach badawczych na wyjaśnianiu motywacji, postaw i zachowań, ujawnianiu ukrytych przekonań i emocji, odkrywaniu repertuarów skojarzeń, odtwarzaniu stereotypów, wzorów reagowania, oceniania, wyrażania myśli, etc;
- zaangażowanie specjalistycznej grupy badaczy o nastawieniu psychologicznym bezpośrednich uczestników, obserwatorów, animatorów badania;
- brak standardowych metod analizy i interpretacji wyników bazowanie na doświadczeniu, kapitale wiedzy, intuicji badaczy;
- unikanie generalizacji rezultatów badania – formułowania wniosków ogólnych, dotyczących całej populacji¹³⁸.

Najpopularniejsze techniki badań jakościowych to zogniskowane wywiady grupowe (tzw. dyskusje fokusowe), swobodne i pogłębione wywiady indywidualne, obserwacje uczestniczące¹³⁹.

¹³⁷ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010)

¹³⁸ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

Badania marketingowe – systematyczne planowanie, zbieranie, analiza, przekazywanie danych i informacji istotnych dla sytuacji marketingowej w jakiej znajduje się przedsiębiorstwo¹⁴⁰.

Business-to-business – w terminologii badawczej – projekty realizowane w środowisku podmiotów gospodarczych badania, w których respondentami/jednostkami obserwacji są niejako firmy (bezpośrednio zaś osoby zarządzające / pracownicy wypowiadający się w kwestiach związanych ze swą aktywnością zawodową). Popularny skrótowiec terminu *business-to-business* to b2b¹⁴¹.

CAPI – (od ang. *Computer-Assisted Personal Interviewing* – wywiad osobisty ze wspomaganie komputerowym) technika badań ilościowych, w której bezpośredniemu kontaktowi ankietera z respondentem towarzyszy komputer (zwykle przenośny), zaopatrzone w specjalistyczne oprogramowanie wspomagające realizację wywiadu. Ankieter, posługujący się techniką CAPI, odczytuje wystandaryzowane pytania kwestionariuszowe wprost z ekranu komputerowego, a po uzyskaniu odpowiedzi rejestruje je w pamięci urządzenia. Oprogramowanie komputerowe zapewnia utrzymanie prawidłowego scenariusza wywiadu, weryfikuje poprawność wprowadzanych danych, wariantuje logiczny układ kwestionariusza zgodnie ze specyfiką kolejnych odpowiedzi respondenta. Komputer pozwala tu też na prezentację respondentowi bodźców multimedialnych, na głosową rejestrację jego odpowiedzi, jest też narzędziem kontroli jakości pracy ankietera¹⁴².

CASI – (od ang. *Computer-Assisted Self Interviewing* – komputerowa ankieta wypełniana samodzielnie przez respondenta) technika badań ilościowych zbliżona do systemu CAPI. Respondent, zaopatrzone w komputer ze specjalistycznym oprogramowaniem, samodzielnie odczytuje z ekranu treść kolejnych pytań kwestionariuszowych i własnoręcznie wprowadza do komputera swoje odpowiedzi. Podobnie jak w przypadku systemu CAPI, komputer nadzoruje, by wypełnianie ankiety było prawidłowe, wyświetla na ekranie niezbędne instrukcje i weryfikuje jakość oraz logikę wprowadzanych odpowiedzi. Najbardziej istotną odmiernością

¹³⁹ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

¹⁴⁰ P. Kotler, Marketing, Gebethner&Ska, Warszawa 1994, s. 120.

¹⁴¹ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

¹⁴² P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

CASI w stosunku do systemu CAPI jest nieobecność ankietera, którego rolę przejmuje całkowicie komputer¹⁴³.

CATI – (od ang. *Computer-Assisted Telephone Interviewing* – wywiad telefoniczny ze wspomaganie komputerowym) technika badań ilościowych, w której pozyskiwanie informacji od respondentów, odbywa się za pośrednictwem łączy telefonicznych i przy wykorzystaniu specjalistycznego oprogramowania komputerowego. Wywiady telefoniczne ze wspomaganie komputerowym (CATI) prowadzone są zwykle z profesjonalnie zaaranżowanego studia, gdzie odpowiednio przeszkoleni ankieterzy łączą się telefonicznie z wylosowanymi do badania rozmówcami i prowadzą z nimi wystandaryzowane wywiady kwestionariuszowe. Przebieg tych wywiadów podlega ścisłemu nadzorowi systemu komputerowego. Specjalne oprogramowanie, wspomagające pracę teleankieterów, reguluje kolejność zadawanych pytań, rejestruje odpowiedzi, weryfikuje logiczną poprawność gromadzonych danych¹⁴⁴. Do zalet wywiadu telefonicznego należą: niski koszt prowadzenia badania, szeroki dostęp do respondentów, możliwość kontroli pracy ankietera, wyższy odsetek reakcji w porównaniu do innych metod, krótki czas zbierania danych, prostota metody, pomoc programów komputerowych. Do podstawowych wad wywiadu telefonicznego zaliczamy: trudności z uzyskaniem dostępu do pełnej listy abonentów telefonicznych, brak możliwości prezentacji graficznej, trudności w nawiązaniu kontaktu przez telefon, niepewność co do cech elementu badanego.

CAWI – (od ang. *Computer-Assisted Web Interviewing*) – wywiad internetowy nadzorowany przez system komputerowy) popularyzująca się technika badań ilościowych, w której pytania kwestionariuszowe pobierane są ze strony internetowej organizatora badania i przekazywane za pośrednictwem sieci do dowolnego punktu, w którym znajduje się respondent wraz z komputerem podłączonym do Internetu. Osoba badana w systemie CAWI, samodzielnie lub w asyście ankietera, odczytuje z ekranu treść pytań i udziela odpowiedzi, które rejestrowane są na docelowym serwerze. Oprogramowanie komputerowe obsługujące ankietę internetową dba o zachowanie właściwej kolejności pytań przesyłanych respondentowi, weryfikuje poprawność logiczną wprowadzanych odpowiedzi. Technika CAWI ułatwia ankietowanie niektórych grup respondentów (zbiorowości trudno dostępnych, rozproszonych geograficznie), pozwala na jednoczesne

¹⁴³ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

¹⁴⁴ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010).

prorowadzenie dużej ilości niezależnych pomiarów¹⁴⁵. Do zalet CAWI można zaliczyć krótki czas realizacji, wysoką penetrację w najbardziej interesujących segmentach populacji (np. osób o wyższych zarobkach, lepiej wykształconych, młodszych) oraz możliwość wykorzystania atrakcyjnych materiałów badawczych (np. zdjęć, filmów, utworów muzycznych). Ogólnodostępne ankiety WWW są przeważnie umieszczane na portalach branżowych lub stronie WWW organizacji. Dużym ograniczeniem stosowania tej metody jest brak kontroli nad próbą oraz błąd braku reakcji. Ankiety WWW z ograniczonym dostępem polegają na dokonaniu wstępnej selekcji respondentów. Wybranej grupie respondentów wysyłane jest zaproszenie do udziału w badaniu wraz z linkiem do strony, na której ono się znajduje. Zaproszenia wysyłane są często w formie e-maili, co nie wyklucza wysłania ich pocztą czy też zaproszeń telefonicznych.

Ciągłe badania – badania prowadzone w sposób cykliczny, powtarzalny, analizujące niezmiennie ten sam zasadniczy zakres tematów, realizowane w kolejnych etapach (tzw. falach) przy użyciu tych samych, bądź bardzo zbliżonych narzędzi badawczych. Badania ciągłe służą ocenie dynamiki postaw, zachowań, zwyczajów badanej zbiorowości, analizie rynkowych trendów etc. Każda kolejna fala badania ciągłego realizowana jest z nową grupą respondentów reprezentujących całą badaną populację. Czynnikiem ten odróżnia badania ciągłe od projektów typu panelowego. Przykładem badań ciągłych może być np: monitoring popularności stacji telewizyjnych, badanie dynamiki zwyczajów żywieniowych Polaków, analiza trendów związanych z czytelnictwem książek etc.¹⁴⁶

Dane pierwotne – informacje pozyskane za pomocą procedury badawczej zaplanowanej specjalnie w tym celu. Dane, które do chwili badania nie istniały, nie były dostępne¹⁴⁷.

Dane wtórne – informacje pozyskane ze źródeł już istniejących, nie wymagające realizacji bezpośrednich badań, a jedynie scalenie i uporządkowania rozproszonego dotąd materiału informacyjnego¹⁴⁸.

Desk research (badania gabinetowe) – metoda badań społecznych lub marketingowych, która zakłada szczegółową analizę istniejących już i dostępnych danych. Realizacja projektu typu Desk research nie jest zatem związana z pozyskiwaniem nowych informacji, a jedynie scaleniem,

¹⁴⁵ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

¹⁴⁶ P. Dmitruk, „Słownik badawczy”, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

¹⁴⁷ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

¹⁴⁸ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

przetworzeniem i analizą danych rozproszonych dotychczas wśród rozmaitych źródeł¹⁴⁹.

Dobór losowy – jest taką metodą doboru próby, która zapewnia każdemu elementowi populacji jednakową szansę wejścia do próby. Spełnienie powyższego warunku jest niezwykle trudne, dlatego Kerlinger zaproponował definicję bardziej realistyczną: „Dobór losowy jest metodą pobierania porcji (próby) z populacji w ten sposób, że wszystkie możliwe próby o określonej wielkości n cechuje takie samo prawdopodobieństwo wyselekcjonowania”. Do najbardziej rozpowszechnionych schematów losowania można zaliczyć: losowanie indywidualne nieograniczone, losowanie indywidualne systematyczne, losowanie warstwowe, wielostopniowe¹⁵⁰.

Dobór nielosowy – uzależniony jest od osoby prowadzącej badanie. W przypadku doboru nielosowego nie ma możliwości określenia prawdopodobieństwa znalezienia się jednostki w próbie. Dobór nielosowy polega na wyborze konkretnych jednostek o ustalonych z góry charakterystykach. Do najbardziej rozpowszechnionych należą: dobór celowy oraz dobór kwotowy¹⁵¹.

Eksperyment – badanie, w którym rozmówca/respondent/uczestnik testu, nie jest całkowicie świadom jego ostatecznych celów, nie udziela też wprost odpowiedzi na ustanowione pytania badawcze. Ostateczny rezultat eksperymentu ustalany jest na podstawie szczegółowej analizy zachowań osób badanych, ich pośrednich deklaracji, porównań pomiędzy różnymi grupami eksperymentalnymi, odwołań do teorii. Badania eksperymentalne podejmowane są najczęściej w odniesieniu do tematów słabo uświadamianych przez samych respondentów, które obarczone są silną presją stereotypów i racjonalizacji, a ich bezpośrednia analiza natrafia na istotne bariery psychologiczne, językowe etc.

Indywidualne wywiady pogłębione (IDI) (z ang. *Individual In-Depth Interviews*) – jedna z bardziej popularnych metod badań jakościowych, polegająca na szczegółowej, wnikliwej rozmowie z informatorem/respondentem, której celem jest dotarcie do jakichś precyzyjnych informacji, poszerzenie wiedzy związanej z tematem badania. W trakcie wywiadu indywidualnego podejmowane są pytania badawcze o charakterze eksplo-

¹⁴⁹ P. Dmítruk, Słownik badawczy, http://www.pentor.pl/badania_slovník.xml (10.03.1010).

¹⁵⁰ Por. F. N. Kerlinger, *Foundations of Behavioral Research*, Holt, Rinehart & Winston, New York 1983; Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 200-205.

¹⁵¹ Por. S. Mynarski, *Praktyczne metody analizy danych rynkowych i marketingowych*, Kantor Wydawniczy Zakamycze, Zakamycze 2000, s. 32-34; S. Mynarski, *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza FOGRA, Kraków 1995.

racyjnym próby wyjaśniania/zrozumienia zjawisk, motywacji, postaw, zachowań. Wywiady pogłębione prowadzone są przez wyspecjalizowanych badaczy o nastawieniu psychologicznym, którzy umiejętnie wprowadzają kolejne tematy rozmowy, odpowiednio ukierunkowują wypowiedź respondenta, pomagają w ujawnianiu przekonań, opinii, postaw. Przebieg indywidualnego wywiadu pogłębionego jest zazwyczaj rejestrowany na kasetach audio lub wideo, a podstawą interpretacji wyników staje się pogłębiona analiza informacji uzyskanych w całej serii niezależnych wywiadów. Wywiady indywidualne prowadzone są zwykle w przypadkach trudnej dostępności określonego rodzaju respondentów, bądź w sytuacjach, gdy podejmowane zagadnienia mają charakter drażliwy, kontrowersyjny, intymny¹⁵².

Instrukcja kodowa (klucz kodowy) – zestaw reguł uwzględniających przede wszystkim zasady budowania i wyróżniania klas¹⁵³.

Korelacja – współzależność, czyli wzajemne oddziaływanie dwóch zjawisk lub cech tej samej zbiorowości¹⁵⁴.

Kwestionariusz – w rozumieniu ogólnym, każdy zestaw pytań skierowanych do respondenta, ułożonych w określonym porządku i sformułowanych zgodnie z kanonami wiedzy metodologicznej niezależnie od stopnia ich standaryzacji i oczekiwanej precyzji odpowiedzi. W wąskim i bardziej popularnym rozumieniu kwestionariusz to kompozycja pytań do respondentów wykorzystywana w badaniu ilościowym, układ precyzyjnie sformułowanych zapytań, zadawanych dużej grupie respondentów, zaopatrzonej zwykle w listę możliwych odpowiedzi (kafeterię), instrukcje dotyczące ich odczytywania, interpretacji¹⁵⁵.

Metoda – określony, powtarzalny i wyuczony sposób postępowania, świadomie skierowanego na realizację pewnego celu poprzez dobór środków odpowiednich do tego celu¹⁵⁶.

PAPI - (od ang. *Paper and Pencil Interview*) termin określający zbiorczo techniki badań ilościowych, bazujące na konwencjonalnej – papierowej wersji kwestionariuszy badawczych, w których pisemnie zaznaczane są odpowiedzi respondentów. Określenie upowszechnione wraz z rozwojem metod komputerowych wspomagających badania, stosowane dla odróżnienia

¹⁵² P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml

¹⁵³ Por. Gruszczyński L.A., *Kwestionariusze w socjologii. Budowa narzędzi do badań surveyowych*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2003, s. 137.

¹⁵⁴ Makać W., Urbanek-Krzysztofiak D., *Statystyka opisowa*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 1992, s. 115.

¹⁵⁵ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

¹⁵⁶ S. Nowak, *Metodologia badań społecznych*, Wydawnictwo Naukowe PWN, Warszawa 2008, s. 19.

tychże od sytuacji wykorzystania standardowego kwestionariusza papierowego¹⁵⁷.

Pomiar – przyporządkowanie właściwościom empirycznym – symboli, zgodnie z określonymi zasadami. Z pojęciem pomiaru związane są pojęcia trafności i rzetelności¹⁵⁸.

Populacja generalna (populacja) to zbiór dowolnych elementów, nieidentycznych z punktu widzenia badanej cechy. Może być skończonym lub nieskończonym zbiorem jednostek, które zamierzamy obserwować empirycznie i analizować¹⁵⁹.

Poziom istotności α – prawdopodobieństwo mierzące szansę popełnienia podczas weryfikacji hipotezy statystycznej błędu pierwszego rodzaju¹⁶⁰. Najczęściej bierze się pod uwagę następujące wartości poziomu istotności: 0,1; 0,05; 0,01 i 0,001¹⁶¹.

Poziom ufności (współczynnik ufności) – prawdopodobieństwo $1-\alpha$ oznaczające prawdopodobieństwo z jakim nieznaną wartość parametru Q objęta jest (pokryta) przez ten przedział¹⁶².

Próba (próbka) – podlegający badaniu skończony podzbiór populacji generalnej¹⁶³.

Regresja, funkcja regresji – podstawowe narzędzie do badania charakteru i kształtu związku pomiędzy rozkładami cech¹⁶⁴.

¹⁵⁷ P. Dmítruk, Słownik badawczy, http://www.pentor.pl/badania_sloownik.xml (10.03.1010).

¹⁵⁸ Por. Ch. Frankfort-Nachmias, D. Nachmias, Metody badawcze w naukach społecznych, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 169, 180-187.

¹⁵⁹ Por. Ch. Frankfort-Nachmias, D. Nachmias, Metody badawcze w naukach społecznych, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 193; S. Mynarski, Praktyczne metody analizy danych rynkowych i marketingowych, Kantor Wydawniczy Zakamycze, Zakamycze 2000, s. 20-21; S. Mynarski, Badania rynkowe w warunkach konkurencji, Oficyna Wydawnicza FOGRA, Kraków 1995, s. 19-26.

¹⁶⁰ M. Maliński, Statystyka matematyczna wspomagana komputerowo, Wydawnictwo Politechniki Śląskiej, Gliwice 2000, s. 95.

¹⁶¹ J. Greń, Modele i zadania statystyki matematycznej, PWN, Warszawa 1970, s. 15.

¹⁶² M. Maliński, Statystyka matematyczna wspomagana komputerowo, Wydawnictwo Politechniki Śląskiej, Gliwice 2000, s. 94.

¹⁶³ W. Krysicki, J. Bartos, K. Królikowska, M. Wasilewski, Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach, Wydawnictwo Naukowe PWN, Warszawa 1995, s. 5.

¹⁶⁴ J. Greń, Modele i zadania statystyki matematycznej, PWN, Warszawa 1970, s. 16.

Respondent – uczestnik badania społecznego lub marketingowego, który jest źródłem informacji badawczych, uczestniczy w wywiadzie, wypełnia ankietę, jest poddawany obserwacji lub eksperymentowi.

Rzetelność określa, czy i na ile wyniki otrzymane z wykorzystaniem instrumentu pomiarowego nie zmieniają się wraz z kolejnym pomiarem¹⁶⁵.

System informacji marketingowej (SIM) składa się z ludzi, sprzętu oraz technik gromadzenia, porządkowania, analizy i oceny, a następnie przekazywania na czas potrzebnej i dokładnej informacji do osób podejmujących decyzje dotyczące marketingu¹⁶⁶.

System wywiadu marketingowego to zbiór procedur i źródeł wykorzystywanych przez kierownictwo, aby pozyskać codzienne informacje na temat rozwoju sytuacji w otoczeniu marketingowym¹⁶⁷.

Szereg czasowy – ciąg obserwacji pewnego zjawiska w kolejnych jednostkach czasu¹⁶⁸.

Test statystyczny – „narzędzie statystyczne”, za pomocą którego dokonuje się weryfikacji hipotez statystycznych. Wynikiem zastosowania testu statystycznego jest decyzja o przyjęciu lub odrzuceniu sprawdzanej hipotezy statystycznej¹⁶⁹. Test statystyczny to reguła postępowania, która na podstawie wyników próby ma doprowadzić do decyzji przyjęcia lub odrzucenia postawionej hipotezy statystycznej¹⁷⁰.

Trafność określa, czy instrument pomiarowy mierzy to, co miał mierzyć¹⁷¹.

Wnioskowanie statystyczne polega na uogólnianiu wyników uzyskanych dla próby na całą populację generalną z wykorzystaniem metod probabilistycznych¹⁷².

Zogniskowane wywiady grupowe (FGI) – (z ang. *Focus Group Interviews*) jedna z najpopularniejszych metod badań jakościowych, polegająca na wspólnej dyskusji grupy respondentów/uczestników wywiadu, na zadany z góry temat/grupę tematów. W trakcie wywiadu grupowego podejmowane są pytania badawcze o charakterze eksploracyjnym próby wyjaśniania/

¹⁶⁵ Por. Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 169, 180-187.

¹⁶⁶ P. Kotler, *Marketing*, Gebethner&Ska, Warszawa 1994, s. 115.

¹⁶⁷ P. Kotler, *Marketing*, Gebethner&Ska, Warszawa 1994, s. 119.

¹⁶⁸ W. Starzyńska, *Statystyka praktyczna*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 146.

¹⁶⁹ M. Maliński, *Statystyka matematyczna wspomagana komputerowo*, Wydawnictwo Politechniki Śląskiej, Gliwice 2000, s. 95.

¹⁷⁰ J. Greń, *Modele i zadania statystyki matematycznej*, PWN, Warszawa 1970, s. 15.

¹⁷¹ Por. Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001, s. 169, 180-187.

¹⁷² A. Balicki, W. Makać, *Metody wnioskowania statystycznego*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2000, s. 7.

zrozumienia zjawisk, motywacji, postaw, zachowań bez intencji wyrażania badanej rzeczywistości w sposób liczbowy i czysto opisowy. Uczestnicy dyskusji stymulują się wzajemnie, inspirują, prowokują do wyrażania opinii i konstruowania pomysłów. Zogniskowane wywiady grupowe prowadzone są przez wyspecjalizowanych badaczy/moderatorów, którzy nadzorują przebieg dyskusji, ukierunkowują rozmowę zgodnie z zaplanowanym scenariuszem, pomagają w dochodzeniu do konkluzji. Wywiad realizowany jest zwykle w pomieszczeniach/salach umożliwiających zlecniodawcy badania bieżący podgląd uczestników dyskusji poprzez lustro weneckie lub za pomocą monitorów. Przebieg dyskusji grupowej jest też zazwyczaj rejestrowany na kasetach wideo, a podstawą interpretacji wyników staje się pogłębiona analiza przebiegu wywiadu dokonywana przez prowadzącego ją badacza¹⁷³.

¹⁷³ P. Dmitruk, Słownik badawczy, http://www.pentor.pl/badania_slownik.xml (10.03.1010). Wywiady grupowe jako metodę badawczą zaczęto wykorzystywać na początku lat czterdziestych XX wieku w Stanach Zjednoczonych. Pierwsze przeprowadzone badanie dotyczyło audytorium programów radiowych i zostało przeprowadzone przez dwóch socjologów F.P. Lazarsfelda oraz R.K. Mertona. Wywiady grupowe są narzędziem badania jakościowego. Ich celem jest: poznanie schematów myślenia respondentów, poznanie wyobrażeń i percepcji, poznanie zakresu upodobań i niechęci, wysłuchanie uwag i pomysłów, wysłuchanie argumentów za i przeciw, poznanie pozytywnych i negatywnych przykładów z doświadczeń respondentów, odkrycie pozytywnych i negatywnych skojarzeń. Por. R.J. Kaden, *Badanie marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 200, s. 93

LITERATURA

Opracowania zwarte

- [1] BABBIE E.: *Badania społeczne w praktyce*, Wydawnictwo Naukowe PWN, Warszawa 2003.
- [2] *Badania marketingowe. Podstawowe metody i obszary zastosowań*, pod red. K. Mazurek-Łopacińskiej, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 2002.
- [3] CHURCHILL G.A.: *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002.
- [4] DULINIEC E.: *Badania marketingowe w zarządzaniu przedsiębiorstwem*, Wydawnictwo Naukowe PWN, Warszawa 1999.
- [5] Frankfort-Nachmias Ch., Nachmias D., *Metody badawcze w naukach społecznych*, Wydawnictwo Zysk i S-ka, Poznań 2001.
- [6] GARBARSKI L., RUTKOWSKI I., WRZOSEK W.: *Marketing. Punkt zwrotny nowoczesnej firmy*, PWE, Warszawa 1996.
- [7] GRUSZCZYŃSKI L.A.: *Kwestionariusze w socjologii. Budowa narzędzi do badań surveyowych*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2003.
- [8] GERRIT A., FRED VAN RAAIJ W.: *Zachowanie konsumenta. Podręcznik akademicki*, Wydawnictwo Naukowe PWN, Warszawa 2003.
- [9] HAGUE P.: *Badanie marketingowe. Planowanie, metodologia i ocena wyników*, Wydawnictwo Helion, Gliwice 2006.
- [10] KACZMARCZYK S.: *Badania marketingowe. Metody i techniki*, PWE, Warszawa 1995.
- [11] KACZMARCZYK S.: *Badania marketingowe. Metody i techniki*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2003.
- [12] KACZMARCZYK S.: *Zastosowanie badań marketingowych. Zarządzanie marketingowe i otoczenie przedsiębiorstwa*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007.
- [13] KADEN R.J.: *Badanie marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2008.
- [14] KANIEWSKA-SĘBA A., LESZCZYŃSKI G., PILARCZYK B.: *Badania marketingowe na rynku business-to-business*, Oficyna Ekonomiczna, Kraków 2006.
- [15] KĘDZIOR Z., KARCZ K.: *Badania marketingowe w praktyce*, Polskie Wydawnictwo Ekonomiczne, Warszawa 2007.
- [16] KOTLER P.: *Marketing*, Gebethner&Ska, Warszawa 1994.

- [17] KRYSICKI W., BARTOS J., DYCZKA W., KRÓLIKOWSKA K., WASILEWSKI M.: *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach*, tom II, Wydawnictwo Naukowe PWN, Warszawa 1995.
- [18] LUSZNIEWICZ A., SŁABY T.: *Statystyka stosowana*, PWE, Warszawa 1997.
- [19] MAISON D., NOGA-BOGOMILSKI A.: *Badania marketingowe. Od teorii do praktyki*. Gdańskie Wydawnictwo Psychologiczne, Gdańsk 2007.
- [20] *Marketing interaktywny 2006*, Gemius SA, Warszawa 2006.
- [21] MAZUREK-ŁOPACIŃSKA K. (red.): *Badania marketingowe. Teoria i praktyka*, PWN, Warszawa 2005.
- [22] MAZUREK-ŁOPACIŃSKA K.: *Zachowania nabywców jako podstawa strategii marketingowej*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 1997.
- [23] MYNARSKI S.: *Analiza danych rynkowych i marketingowych z wykorzystaniem programu STATISTICA*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2003.
- [24] MYNARSKI S.: *Badania rynkowe w warunkach konkurencji*, Oficyna Wydawnicza Fogra, Kraków 1995.
- [25] MYNARSKI S.: *Praktyczne metody analizy danych rynkowych i marketingowych*, Kantor Wydawniczy Zakamycze, Zakamycze 2000.
- [26] NIKODEMSKA-WOŁOWIK A.: *Jakościowe badania marketingowe*. Polskie Wydawnictwo Ekonomiczne, Warszawa 1999.
- [27] OSTASIEWICZ S., RUSNAK Z., SIEDLECKA U.: *Statystyka. Elementy teorii i zadania*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego, Wrocław 1997.
- [28] *Podstawy prowadzenia badań marketingowych. Etapy procesu badań marketingowych – teoria procesu badawczego*, CRON, Warszawa 2006.
- [29] RÓSZKIEWICZ M.: *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002.
- [30] SAGAN A.: *Badania marketingowe. Podstawowe kierunki*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004.

Inne

- [1] CIEŚLAK M.: Taksonomiczna procedura programowania rozwoju gospodarczego i określania zapotrzebowania na kadry kwalifikowane, [w:] *Przegląd Statystyczny* 1974, z. 1.
- [2] PERKAL J.: O wskaźnikach antropologicznych, [w:] *Przegląd Antropologiczny* 1953, t. 19.

- [3] American Marketing Association. <http://www.marketingpower.com/AboutAMA/Pages/DefinitionofMarketing.aspx>.
- [4] JERSCHINA J.: *Całościowy system informacji marketingowej w przedsiębiorstwach i instytucjach finansowych*, <http://www.statsoft.pl/czytelnia>
- [5] JAWORSKI M.: Wywiad gospodarczy na wewnętrzny użytek. [Online] 2002. <http://ebib.oss.wroc.pl/2002/40/jaworski.php>.
- [6] <http://aneksy.pwn.pl/zarzadzanie/pdf/Waters17-Wal-Mart.pdf>
- [7] <http://cio.cxo.pl/artykuly/54278/IT.w.branzy.sprzedazy.detalicznej.html>
- [8] Business Intelligence w organizacjach handlowych, http://www.img.pl/Advice_Corner__BI_in_Retail__PL.IMG?ActiveID=2691
- [9] DMITRUK P.: *Słownik badawczy*, http://www.pentor.pl/badania_sloownik.xml
- [10] SRI Consulting Business Experiences. [Online] <http://www.sricbi.com/VALS/types.shtml>.
- [11] Główny Urząd Statystyczny <http://www.stat.gov.pl>
- [12] American Marketing Association. <http://www.marketingpower.com/AboutAMA/Pages/DefinitionofMarketing.aspx>.
- [13] HBI Polska <http://www.hbi.pl>
- [14] Główny Urząd Statystyczny <http://www.stat.gov.pl>
- [15] <http://www.pf.pl>
- [16] <http://www.ditel.pl>
- [17] <http://www.ofbor.pl>
- [18] <http://www.ptbrio.pl>
- [19] <http://www.4prm.com>
- [20] <http://www.ACNielsen.com>
- [21] *Zatrudnienie i wynagrodzenia w gospodarce narodowej w I kwartale 2009 r.*, Główny Urząd Statystyczny, Warszawa 2009.
- [22] Nielsen, <http://www.agbnielsen.pl>
- [23] <http://www.opiniaabr.pl>
- [24] ASM, <http://www.asm-poland.com.pl>
- [25] CEM, <http://www.cem.pl>
- [26] Indicato, <http://www.indicator.pl>
- [27] Gfk Polonia, <http://www.gfk.pl>
- [28] <http://www.memrb.com>
- [29] Pentor, <http://www.pentor.pl>
- [30] BSM, <http://www.bsm.com.pl>
- [31] TNS Global, <http://www.tns-global.pl>
- [32] Schober Group [Online]. - <http://www.schober.pl/site/index.php?id=1863>.

Spis tabel

Tabela 1.	Proces transformacji problemu decyzyjnego w problem badawczy.....	28
Tabela 2.	Skale pomiaru.....	35
Tabela 3.	Macierz błędów systematycznych i przypadkowych.	39
Tabela 4.	Wpływ wybranych cech pomiaru na jego rzetelność.....	43
Tabela 5.	Podział danych wtórnych.	67
Tabela 6.	Wybrane użyteczne źródła danych wtórnych wewnętrznych: ...	69
Tabela 7.	Wybrane źródła informacji biznesowej na rynku polskim.....	71
Tabela 8.	Reprezentatywność próby	102
Tabela 9.	Wady i zalety wybranych metod doboru próby	103
Tabela 10.	Określenie liczebności próby nielosowej – wielkość próby stosowana w różnego typu badaniach marketingowych.	105
Tabela 11.	Zwyczajowa wielkość próby stosowana w różnego typu badaniach marketingowych.....	106
Tabela 12.	Fragment instrukcji kodowej do pytania zamkniętego.....	122
Tabela 13.	Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego (system wielowartościowy, Pyt_10a, Pyt_10b, Pyt_10c w bazie danych)	123
Tabela 14.	Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego (system wielowartościowy, Pyt_10 razem w bazie danych).....	124
Tabela 15.	Fragment instrukcji kodowej do pytania zamkniętego koniunktywnego (system wielowartościowy, Pyt_11a, Pyt_11b, Pyt_11c w bazie danych)	124
Tabela 16.	Fragment instrukcji kodowej do pytania otwartego systemem wielowartościowym z wykorzystaniem jednej kolumny w bazie danych.....	125
Tabela 17.	Zakodowane i zapisane w arkuszu kalkulacyjnym odpowiedzi	126
Tabela 18.	Przykład tabulacji prostej – miejsca sprzedaży produktów rolniczych z wykorzystaniem bezpośredniego kanału zbytu (n=200).....	127
Tabela 19.	Przykład tabulacji złożonej – dane surowe – opinie respondenta o produkcie (n=280).....	127
Tabela 20.	Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280).....	128
Tabela 21.	Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280).....	128
Tabela 22.	Przykład tabulacji złożonej – rozkłady procentowe – opinie respondenta o produkcie (n=280).....	128

Tabela 23.	Charakterystyki liczbowe stosowane w analizie danych w zależności od użytej skali pomiarowej.....	132
Tabela 24.	TOP 20 Tygodniowe wydatki na reklamę w podziale na stacje.....	136
Tabela 25.	Miesięczne płace brutto (w zł) uzyskane w badaniu pracowników na stanowiskach robotniczych w przedsiębiorstwie X (n=35).....	137
Tabela 26.	Wyznaczone charakterystyki liczbowe	142
Tabela 27.	Dane do przykładu 4.	158
Tabela 28.	Dane do przykładu 5.	160
Tabela 29.	Wygląd tablicy REGLINP	163
Tabela 30.	Dane do przykładu	172
Tabela 31.	Produkcja ważniejszych wyrobów przemysłowych wg Polskiej Klasyfikacji Wyrobów i Usług (PKWiU 2008) / PRODPOL.....	172
Tabela 32.	Trzoda chlewna (w mln sztuk) w Polsce (stan w czerwcu).....	173
Tabela 33.	Skup mleka w Polsce (w mln litrów).	173
Tabela 34.	Możliwe decyzje w procesie testowania hipotez.....	176
Tabela 35.	Możliwe decyzje w procesie testowania hipotez – tablica wielodzielcza.....	183
Tabela 36.	Dane do przykładu	186
Tabela 37.	Obliczenia pomocnicze – obliczenia liczebności teoretycznych	186
Tabela 38.	Obliczenia pomocnicze – obliczenie statystyki χ^2	187
Tabela 39.	Wyniki badania.	188
Tabela 40.	Wyniki badania.	188
Tabela 41.	Wyniki badania (dane umowne) wraz z cząstkowymi wynikami obliczeń	195
Tabela 42.	Macierz odległości	196
Tabela 43.	Profile wyznaczonych segmentów	197
Tabela 44.	Wartości dystrybuanty rozkładu normalnego standaryzowanego $P(X < u_\alpha) = 1 - \alpha$	198
Tabela 45.	Tablice rozkładu t-Studenta $P(X > t_\alpha) = \alpha$	199
Tabela 46.	Tablice rozkładu chi-kwadrat $P(X > \chi_\alpha^2) = \alpha$	200

Spis wykresów

Wykres 1. Przykład prezentacji danych na wykresie (wykres słupkowy).	130
Wykres 2. Przykład prezentacji danych na wykresie (histogram).....	130
Wykres 3. Przykład prezentacji danych na wykresie (wykres kołowy)	131
Wykres 4. Przykład prezentacji danych na wykresie (wykres rozrzutu)...	131
Wykres 5. Wykres pudełko z wąsami.....	138
Wykres 6. Histogram licznosci.....	141
Wykres 7. Wykres rozrzutu, porównanie modeli – przykład 4.	159
Wykres 8. Wykres rozrzutu, porównanie modeli – przykład 5.	161
Wykres 9. Alokacja wydatków przeznaczanych na normalne zakupy w ciągu tygodnia	169
Wykres 10. Linie trendu	170

Spis rysunków

Rys. 1.	Schemat marketingowego systemu informacyjnego	14
Rys. 2.	Cele badań marketingowych	17
Rys. 3.	Proces badania marketingowego według Mazurek-Łopacińskiej	20
Rys. 4.	Skale pomiaru.....	34
Rys. 5.	Cechy jakościowe i ilościowe człowieka	35
Rys. 6.	Podział skal porządkowych	37
Rys. 7.	Wpływ procesu poznania na kształtowanie się postaw i emocji ..	40
Rys. 8.	Klasyfikacja skal do pomiaru postaw	44
Rys. 9.	Etapy budowy kwestionariusza.	59
Rys. 10.	Podział informacji ze względu na kryterium rodzajów decyzji....	68
Rys. 11.	Podstawowe rodzaje wywiadów.....	79
Rys. 12.	Klasyfikacja ankiet	82
Rys. 13.	Schemat przebiegu eksperymentu	84
Rys. 14.	Klasyfikacja obserwacji.....	87
Rys. 15.	Kryteria segmentacji rynku	94
Rys. 16.	Metody zbierania danych w badaniach stylów życia	96
Rys. 17.	Segmenty nabywców	97
Rys. 18.	Łączny błąd badania	110
Rys. 19.	Możliwe przyczyny wystąpienia braku kontaktu z elementami badanymi	112
Rys. 20.	Proces udzielania odpowiedzi przez respondenta.....	115
Rys. 21.	Odniesienie wyników badań marketingowych do rzeczywistości.....	129
Rys. 22.	Zakres analizy statystycznej	129
Rys. 23.	Podział metod analizy korelacji.....	145
Rys. 24.	Zależność współczynnika korelacji liniowej Pearsona od zakresu zmienności cech	150
Rys. 25.	Dekompozycja szeregu czasowego	168
Rys. 26.	Przebieg procedury weryfikacyjnej	176
Rys. 27.	Podział testów statystycznych	178
Rys. 28.	Obszary krytyczne dla różnych postaci hipotezy alternatywnej.	182
Rys. 29.	Obszar krytyczny	185
Rys. 30.	Obszar krytyczny	187
Rys. 31.	Dendryt I rzędu.....	196