

**Microsoft Excel** jest naprawdę wszechstronnym narzędziem. Umożliwia tworzenie raportów, inteligentnych modeli, a także prowadzenie złożonych analiz statystycznych. Profesjonalni badacze, studenci i biznesmeni, którzy zajmują się analizą danych i statystyką, właśnie Excela traktują jako ulubione narzędzie pracy. Osiągnięcie biegłości w posługiwaniu się nim wymaga odrobiny wysiłku, jednak zdobyta w ten sposób wiedza okazuje się bardzo przydatna!

Niniejsza książka jest bardzo **praktycznym przewodnikiem po analizie statystycznej i funkcjach statystycznych Excela**. Dzięki licznym przykładom nauczysz się dobierać właściwe narzędzia do rozwiązania konkretnego problemu. Dowiesz się, jak korzystać z korelacji, regresji i analizy wariancji oraz kowariancji. Zastosujesz Excela do testowania hipotez statystycznych z zastosowaniem rozkładów normalnych, dwumianowych, t-Studenta i F-Snedecora. Zapoznasz się ze znaczeniem najważniejszych pojęć statystycznych i unikniesz typowych błędów. W tym wydaniu książki uwzględniono nowe funkcje, które pojawiły się w Excelu 2016. Dowiesz się, do czego służą i jak je można wykorzystać.

Najważniejsze zagadnienia ujęte w książce:

- podstawowe pojęcia w statystyce, statystyce opisowej i wnioskowaniu statystycznym
- ważniejsze funkcje statystyczne Excela
- testy z, testy t i dodatek Analiza danych Excela
- histogramy i wykresy w Excelu
- analiza wariancji i kowariancji oraz metody regresji

## **Analiza statystyczna: z Excelem to łatwe!**

---

**Dr Conrad Carlberg** jest uznanym autorytetem w dziedzinie statystyki i analizy danych. Doskonale zna takie aplikacje, jak MS Excel, SAS i Oracle. Wielokrotnie otrzymywał nagrodę MVP. Od ponad 20 lat doradza firmom, które chcą podejmować decyzje biznesowe na podstawie analizy danych. Z radością pisze o tych technikach, ze szczególnym upodobaniem dzieląc się swoją ogromną wiedzą o MS Excelu, którego uważa za najpopularniejszy w świecie program do analiz numerycznych.

## **Spis treści**

### **Wstęp (13)**

- Stosowanie Excela do analizy statystycznej (13)
  - Czytelnicy i Excel (14)
  - Porządkowanie terminów (15)
  - Upraszczenie spraw (16)
  - Zły produkt? (17)
  - Odwracanie kota ogonem (19)
- Co zawiera książka? (20)

### **1. Zmienne i wartości (21)**

- Zmienne i wartości (21)

- Zapisywanie danych w postaci list (22)
- Korzystanie z list (23)
- Skale pomiarowe (25)
  - Skale nominalne (25)
  - Skale liczbowe (27)
  - Określanie wartości przedziałowych na podstawie wartości tekstowych (28)
- Graficzna prezentacja zmiennych liczbowych w Excelu (31)
  - Graficzna prezentacja dwóch zmiennych (31)
- Pojęcie rozkładów liczebności (33)
  - Stosowanie rozkładów liczebności (36)
  - Budowanie rozkładu liczebności na podstawie próby (40)
  - Tworzenie symulowanych rozkładów liczebności (49)

## **2. Jak się skupiają wartości (51)**

- Obliczanie średniej arytmetycznej (53)
  - Funkcje, argumenty i wyniki (54)
  - Formuły, wyniki i formaty (56)
  - Minimalizowanie rozproszenia (58)
- Obliczanie mediany (64)
  - Decyzja o użyciu mediany (65)
  - Stabilna czy raczej odporna? (66)
- Obliczanie wartości modalnej (67)
  - Otrzymywanie wartości modalnej kategorii za pomocą formuły (73)
- Od tendencji centralnej do rozrzutu (80)

## **3. Rozrzut - jak się rozpraszają wartości (81)**

- Mierzenie rozproszenia za pomocą rozstępu (82)
  - Rozstęp a liczebność próby (84)
  - Zmienność na bazie rozstępu (85)
- Koncepcja odchylenia standardowego (87)
  - Dopasowanie do standardu (88)
  - Myślenie w kategoriach odchyłeń standardowych (89)
- Obliczanie odchylenia standardowego i wariancji (91)
  - Podnoszenie odchyłeń do kwadratu (94)
  - Parametry populacji i przykładowe statystyki (95)
  - Dzielenie przez N-1 (96)
- Obciążoność estymatora a stopnie swobody (98)
- Funkcje Excela do mierzenia rozproszenia (100)
  - Funkcje odchylenia standardowego (100)
  - Funkcje wariancji (101)

## **4. Jak zmienne wspólnie się zmieniają - korelacja (103)**

- Pojęcie korelacji (103)
  - Wyznaczanie współczynnika korelacji (105)
  - Korzystanie z funkcji WSP.KORELACJI() (112)
  - Korzystanie z narzędzi analitycznych (115)
  - Korzystanie z narzędzia Korelacja (117)
  - Korelacja nie oznacza przyczynowości (120)

- Stosowanie korelacji (122)
  - Usuwanie efektów skali (123)
  - Korzystanie z funkcji Excela (125)
  - Prognozowanie wartości (127)
  - Szacowanie funkcji regresji (129)
- Stosowanie funkcji REGLINW() do regresji wielorakiej (131)
  - Łączenie predyktorów (131)
  - Najlepsza kombinacja liniowa (133)
  - Pojęcie współdzielonej zmienności (136)
  - Dodatek techniczny: algebra macierzowa i regresja wieloraka w Excelu (138)

## 5. Tworzenie wykresów (141)

- Właściwości wykresów w Excelu (142)
  - Osie wykresów (142)
  - Wartości daty a oś kategorii (144)
  - Inne wartości liczbowe a oś kategorii (146)
- Histogramy (148)
  - Używamy tabeli przestawnej do zliczania obiektów (148)
  - Używamy zaawansowanego filtra i funkcji CZĘSTOŚĆ() (151)
  - Histogram jako część dodatku Analiza danych (154)
  - Histogram na palecie wykresów (154)
  - Serie danych i ich adresy (156)
- Wykresy pudełkowe (157)
  - Obserwacje odstające (160)
  - Badamy asymetrię (160)
  - Porównujemy rozkłady (160)

## 6. Jak zmienne są wspólnie klasyfikowane - tabele kontyngencji (163)

- Jednowymiarowe tabele przestawne (163)
  - Przeprowadzanie testu statystycznego (167)
- Stawianie założeń (173)
  - Dobór losowy (173)
  - Niezależność elementów (175)
  - Wzór na prawdopodobieństwo w rozkładzie dwumianowym (175)
  - Korzystanie z funkcji ROZKŁ.DWUM.ODWR() (177)
- Dwuwymiarowe tabele przestawne (183)
  - Prawdopodobieństwa i zdarzenia niezależne (187)
  - Sprawdzanie niezależności klasyfikacji (189)
  - O regresji logistycznej (195)
- Efekt Yule'a i Simpsona (196)
- Podsumowanie funkcji  $\chi^2$  (199)
  - Korzystanie z funkcji ROZKŁ.CHI() (199)
  - Korzystanie z funkcji ROZKŁ.CHI.PS() i ROZKŁAD.CHI() (200)
  - Korzystanie z funkcji ROZKŁ.CHI.ODWR() (202)
  - Korzystanie z funkcji ROZKŁ.CHI.ODWR.PS() i ROZKŁAD.CHI.ODW() (202)
  - Korzystanie z funkcji CHI.TEST() i TEST.CHI() (203)
  - Stosowanie mieszanych i bezwzględnych odwołań do obliczenia oczekiwanych częstości (204)
  - Korzystanie z wyświetlania indeksu tabeli przestawnej (205)

## **7. Praca z rozkładem normalnym w Excelu (207)**

- Opis rozkładu normalnego (207)
  - Charakterystyki rozkładu normalnego (207)
  - Standaryzowany rozkład normalny (213)
- Funkcje Excela dla rozkładu normalnego (214)
  - Funkcja ROZKŁ.NORMALNY() (214)
  - Funkcja ROZKŁ.NORMALNY.ODWR() (216)
- Przedziały ufności i rozkład normalny (220)
  - Znaczenie przedziału ufności (220)
  - Konstruowanie przedziału ufności (221)
  - Funkcje arkusza Excela, które wyznaczają przedziały ufności (225)
  - Korzystanie z funkcji UFNOŚĆ.NORM() i UFNOŚĆ() (226)
  - Korzystanie z funkcji UFNOŚĆ.T() (229)
  - Zastosowanie dodatku Analiza danych do przedziałów ufności (230)
  - Przedziały ufności i testowanie hipotez (232)
- Centralne twierdzenie graniczne (233)
  - O pewnej osobliwości tabel przestawnych słów kilka (234)
  - Upraszczenie spraw (235)
  - Ulepszanie spraw (238)

## **8. Prawdopodobieństwo statystyki (239)**

- Kontekst wnioskowania statystycznego (240)
  - Zapewnienie trafności wewnętrznej (241)
  - Zagrożenia trafności wewnętrznej (243)
- Problemy z dokumentacją Excela (247)
- Test F z dwiema próbami dla wariancji (249)
  - Po co przeprowadzać ten test? (250)
- Replikowalność badań (262)
- Uwagi końcowe (265)

## **9. Testowanie różnic pomiędzy średnimi - podstawy (267)**

- Testowanie średnich - przesłanki (268)
  - Stosowanie testu z (270)
  - Stosowanie błędu standardowego średniej (272)
  - Tworzenie wykresów (277)
- Stosowanie testu t zamiast testu z (285)
  - Definiowanie reguły decyzyjnej (287)
  - Pojęcie mocy statystycznej (292)

## **10. Testowanie różnic pomiędzy średnimi - dalsze zagadnienia (299)**

- Stosowanie funkcji Excela ROZKŁ.T() i ROZKŁ.T.ODWR() do weryfikacji hipotez (300)
  - Hipotezy jednostronne a hipotezy dwustronne (300)
  - Dobieranie funkcji rozkładu t-Studenta w Excelu do stawianych hipotez (302)
  - Uzupełnienie obrazu za pomocą funkcji ROZKŁ.T() (310)
- Korzystanie z funkcji T.TEST() (311)
  - Stopnie swobody w funkcjach Excela (312)
  - Równe i nierówne liczebności grup (312)

- Składnia funkcji T.TEST() (315)
- Korzystanie z narzędzi do testów t w dodatku Analiza danych (329)
  - Wariacje grupowe w testach t (329)
  - Wizualizacja mocy statystycznej (335)
  - Kiedy unikać testów t (336)

## **11. Testowanie różnic pomiędzy średnimi - analiza wariancji (337)**

- Dlaczego nie testy t? (338)
- Koncepcja analizy wariancji (340)
  - Dzielenie wyników (340)
  - Porównywanie wariancji (343)
  - Test F (348)
- Stosowanie funkcji F arkusza Excela (352)
  - Korzystanie z funkcji ROZKŁ.F() i ROZKŁ.F.PS() (352)
  - Korzystanie z funkcji ROZKŁ.F.ODWR() i ROZKŁAD.F.ODW() (353)
  - Rozkład F (355)
- Nierówne liczebności grup (357)
- Procedury porównań wielokrotnych (358)
  - Procedura Scheffégo (360)
  - Planowane kontrasty ortogonalne (365)

## **12. Analiza wariancji - dalsze zagadnienia (369)**

- Czynnikiowa analiza wariancji (369)
  - Inne przesłanki dla zastosowania wielu czynników (371)
  - Korzystanie z narzędzia do dwuczynnikowej analizy wariancji (373)
- Znaczenie interakcji (376)
  - Istotność statystyczna interakcji (377)
  - Obliczanie efektu interakcji (379)
- Problem nierównych liczebności grup (384)
  - Powtarzane obserwacje - analiza dwuczynnikowa bez powtórzeń (387)
- Funkcje i narzędzia Excela - ograniczenia i rozwiązania (388)
  - Modele mieszane (390)
  - Moc testu F (390)

## **13. Planowanie eksperymentu a ANOVA (393)**

- Czynniki skrzyżowane i czynniki zagnieżdżone (393)
  - Prawidłowy opis eksperymentu (395)
  - Czynniki uciążliwe (397)
- Czynniki stałe i czynniki losowe (397)
  - Narzędzia ANOVA dostępne w dodatku Analiza danych (399)
  - Układ danych (402)
- Wyznaczamy wartości statystyki F (403)
  - Dostosowujemy dodatek Analiza danych do czynników losowych (403)
  - Idea testu F (404)
  - Model mieszany: wybór postaci mianownika (406)
  - Dostosowujemy dodatek Analiza danych do czynników zagnieżdżonych (408)
  - Układ danych dla schematu zagnieżdżonego (409)
  - Sumy kwadratów (410)

- Statystyka F dla czynnika zagnieżdżającego (411)
- Bloki zrandomizowane (412)
  - Interakcja między czynnikami a blokami (413)
  - Test nieaddytywności Tukeya (415)
  - Zwiększamy moc statystyczną (418)
  - Bloki: stałe czy losowe? (419)
- Schemat czynnikowy split-plot (420)
  - Tworzymy schemat split-plot (420)
  - Analiza schematu split-plot (422)

#### **14. Moc statystyczna (427)**

- Kontrola ryzyka (428)
  - Testy jednostronne i dwustronne (428)
  - Zmiana liczebności próby (429)
  - Wizualizacja mocy testu (429)
- Moc statystyczna testów t (433)
  - Test dwustronny (434)
  - Testy jednostronne (437)
  - Zwiększanie rozmiaru próby (438)
  - Test t dla grup zależnych (439)
- Parametr niecentralności w rozkładzie F (441)
  - Oszacowania wariancji (442)
  - Parametr niecentralności a funkcja gęstości prawdopodobieństwa (446)
- Obliczamy moc testu F (448)
  - Wyznaczamy wartość dystrybuanty rozkładu F (449)
  - Wykorzystanie mocy testu do optymalizacji liczebności próby (450)

#### **15. Analiza regresji wielorakiej i rekodowanie zmiennych nominalnych - podstawy (455)**

- Regresja wieloraka a analiza wariancji (456)
  - Stosowanie rekodowania zmiennych (458)
  - Rekodowanie zmiennych - ogólne zasady (459)
  - Inne sposoby kodowania (461)
- Regresja wieloraka a alokacja wariancji (461)
  - Gładkie przejście od analizy wariancji do regresji (464)
  - Znaczenie rekodowania zmiennych (467)
- Rekodowanie zmiennych w Excelu (469)
- Korzystanie z narzędzia Regresja w Excelu do analizy grup o nierównych liczebnościach (472)
- Rekodowanie zmiennych, regresja i schematy czynnikowe w Excelu (474)
  - Stosowanie kontroli statystycznej z korelacjami semicząstkowymi (476)
  - Stosowanie kwadratów współczynników korelacji semicząstkowej do otrzymania prawidłowej sumy kwadratów (478)
- Stosowanie funkcji REGLINW() zamiast kwadratów współczynników korelacji semicząstkowej (479)
  - Praca z resztami (482)
  - Stosowanie bezwzględnego i względnego adresowania Excela do wyznaczania kwadratów współczynników korelacji semicząstkowej (484)

#### **16. Analiza regresji wielorakiej i rekodowanie zmiennych nominalnych - dalsze zagadnienia (489)**

- Analiza niezrównoważonych schematów czynnikowych za pomocą regresji wielorakiej (490)
  - W schemacie zrównoważonym zmienne nie są skorelowane (491)
  - W schemacie niezrównoważonym zmienne są skorelowane (493)
  - Kolejność zmiennych w schemacie zrównoważonym nie jest istotna (494)
  - Kolejność zmiennych w schemacie niezrównoważonym jest istotna (497)
  - Wahające się udziały wariancji (499)
- Schematy eksperymentalne, badania obserwacyjne i korelacja (500)
- Kompletny zestaw wyników funkcji REGLINP() (504)
- Tajniki funkcji REGLINP() (511)
  - Jak działa REGLINP() (512)
  - Współczynniki regresji (514)
  - Sumy kwadratów dla regresji oraz reszt (518)
  - Statystyki diagnostyczne regresji (521)
  - Jak funkcja REGLINP() radzi sobie ze współliniowością (525)
  - Restrykcje zerowe na wyraz wolny (531)
  - Excel 2007 (532)
- Nierówne liczebności grup w prawdziwym eksperymencie (540)
- Nierówne liczebności grup w badaniach obserwacyjnych (543)

### **17. Analiza kowariancji - podstawy (547)**

- Cele analizy kowariancji (548)
  - Większa moc (548)
  - Redukcja obciążenia (549)
- Stosowanie analizy kowariancji w celu zwiększenia mocy statystycznej (549)
  - Analiza wariancji nie znajduje znaczącej różnicy średnich (550)
  - Dodawanie zmiennej towarzyszącej do analizy (552)
- Testowanie średniego współczynnika regresji (560)
- Usuwanie obciążenia - inny przypadek (563)

### **18. Analiza kowariancji - dalsze zagadnienia (569)**

- Korygowanie średnich za pomocą funkcji REGLINP() i rekodowania zmiennych (569)
- Rekodowanie zmiennych a skorygowane średnie grup (575)
- Wielokrotne porównania po analizie kowariancji (578)
  - Metoda Scheffégo (579)
  - Kontrasty planowane (584)
- Analiza kowariancji wielorakiej (586)
  - Decyzja o zastosowaniu wielu zmiennych towarzyszących (586)
  - Dwie zmienne towarzyszące - przykład (587)
- Kiedy nie stosować metody ANCOVA (589)
  - Grupy zdeterminowane (589)
  - Ekstrapolacja (591)

### **Skorowidz (593)**