

1. Record Nr.	UNINA9911007182203321
Autore	Campeato Oswald
Titolo	Statistics Using Python
Pubbl/distr/stampa	Bloomfield : , : Mercury Learning & Information, , 2023 ©2023
ISBN	9781683928782 1683928784 9781683928799 1683928792
Edizione	[1st ed.]
Descrizione fisica	1 online resource (273 pages)
Soggetti	Python (Computer program language) - Statistical methods MATHEMATICS / Probability & Statistics / General
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Front Cover -- Half-Title Page -- LICENSE, DISCLAIMER OF LIABILITY, AND LIMITED WARRANTY -- Title Page -- Copyright Page -- Dedication -- Contents -- Preface -- CHAPTER 1: Working with Data -- What is Data Literacy? -- Exploratory Data Analysis (EDA) -- Dealing with Data: What Can Go Wrong? -- An Explanation of Data Types -- Working with Data Types -- What is Drift? -- Discrete Data Versus Continuous Data -- Binning Data Values -- Correlation -- Working with Synthetic Data -- Summary -- CHAPTER 2: Introduction to Probability -- What is Set Theory? -- Open, Closed, Compact, and Convex Sets (Optional) -- Concepts in Probability -- Set Theory and Probability -- Coin Tossing Probabilities -- Dice Tossing Probabilities -- Card Drawing Probabilities -- Container-Based Probabilities -- Children-Related Probabilities -- Summary -- CHAPTER 3: Introduction to Statistics -- Introduction to Statistics -- Basic Concepts in Statistics -- The Variance and Standard Deviation -- The Moments of a Function (Optional) -- Random Variables -- Multiple Random Variables -- Sampling Techniques for a Population -- What is Bias? -- Two Important Results in Probability -- Summary -- CHAPTER 4: Metrics in Statistics -- The Confusion Matrix -- The ROC Curve and AUC Curve -- The sklearn.

metrics Module (Optional) -- Statistical Metrics for Categorical Data -- Metrics for Continuous Data -- MAE, MSE, and RMSE -- Approximating Linear Data with `np.linspace()` -- Summary -- CHAPTER 5: Probability Distributions -- PDF, CDF, and PMF -- Two Types of Probability Distributions -- Discrete Probability Distributions -- Continuous Probability Distributions -- Advanced Probability Functions -- Non-Gaussian Distributions -- The Best-Fitting Distribution for Data -- Summary -- CHAPTER 6: Hypothesis Testing -- What is Hypothesis Testing? -- Components of Hypothesis Testing -- Test Statistics. Working with p-values -- Working with Alpha Values -- Point Estimation, Confidence Level, and Confidence Intervals -- What is A/B Testing? -- The Lifespan of an A/B Test -- Maximum Likelihood Estimation (MLE) -- Summary -- Appendix A: Introduction to Python -- Tools for Python -- Python Installation -- Setting the PATH Environment Variable (Windows Only) -- Launching Python on Your Machine -- Identifiers -- Lines, Indentation, and Multi-Line Statements -- Quotation Marks and Comments -- Saving Your Code in a Module -- Some Standard Modules -- The `help()` and `dir()` Functions -- Compile Time and Runtime Code Checking -- Simple Data Types -- Working with Numbers -- Working with Fractions -- Unicode and UTF-8 -- Working with Strings -- Slicing and Splicing Strings -- Search and Replace a String in Other Strings -- Remove Leading and Trailing Characters -- Printing Text without New Line Characters -- Text Alignment -- Working with Dates -- Exception Handling -- Handling User Input -- Python and Emojis (Optional) -- Command-Line Arguments -- Summary -- Appendix B: Introduction to Pandas -- What is Pandas? -- A Pandas Data Frame with a NumPy Example -- Describing a Pandas Data Frame -- Boolean Data Frames -- Data Frames and Random Numbers -- Reading CSV Files in Pandas -- The `loc()` and `iloc()` Methods -- Converting Categorical Data to Numeric Data -- Matching and Splitting Strings -- Converting Strings to Dates -- Working with Date Ranges -- Detecting Missing Dates -- Interpolating Missing Dates -- Other Operations with Dates -- Merging and Splitting Columns in Pandas -- Reading HTML Web Pages -- Saving a Pandas Data Frame as an HTML Web Page -- Summary -- Index.

Sommario/riassunto

This book is designed to offer a fast-paced yet thorough introduction to essential statistical concepts using Python code samples, and aims to assist data scientists in their daily endeavors. The ability to extract meaningful insights from data requires a deep understanding of statistics. The book ensures that each topic is introduced with clarity, followed by executable Python code samples that can be modified and applied according to individual needs. Topics include working with data and exploratory analysis, the basics of probability, descriptive and inferential statistics and their applications, metrics for data analysis, probability distributions, hypothesis testing, and more. Appendices on Python and Pandas have been included. From foundational Python concepts to the intricacies of statistics, this book serves as a comprehensive resource for both beginners and seasoned professionals.
