This textbook provides future data analysts with the tools, methods, and skills needed to answer data-focused, real-life questions: to carry out data analysis; and to visualize and interpret results to support better decisions in business, economics, and public policy. Data wrangling and exploration, regression analysis, machine learning, and causal analysis are comprehensively covered, as well as when, why, and how the methods work, and how they relate to each other. As the most effective way to communicate data analysis, running case studies play a central role in this textbook. Each case starts with an industry-relevant question and answers it by using real-world data and applying the tools and methods covered in the textbook. Learning is then consolidated by 360 practice questions and 120 data exercises. Extensive online resources, including raw and cleaned data and codes for all analysis in Stata, R, and Python, can be found at www.gabors-data-analysis.com.

‘This exciting new text covers everything today’s aspiring data scientist needs to know, managing to be comprehensive as well as accessible. Like a good confidence interval, the Gabors have got you almost completely covered!’

Joshua Angrist, Massachusetts Institute of Technology

‘A beautiful integration of econometrics and data science that provides a direct path from data collection and exploratory analysis to conventional regression modeling, then on to prediction and causal modeling. Exactly what is needed to equip the next generation of students with the tools and insights from the two fields.’

David Card, University of California, Berkeley

Adopted at over 80 universities worldwide in economics and business, finance, public and health policy, business analytics and data science. www.gabors-data-analysis.com/courses-using