

1. Record Nr.	UNINA9910806102003321
Autore	Fortino Andres G.
Titolo	Text analytics for business decisions : a case study approach / / Andres G. Fortino
Pubbl/distr/stampa	Dulles : , : Mercury Learning & Information, , [2021] ©2021
ISBN	1-68392-664-1 1-68392-665-X
Descrizione fisica	1 online resource (332 pages)
Disciplina	650.02855369
Soggetti	Business - Decision making - Computer programs Text processing (Computer science)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Contents -- Preface -- On the Companion Files -- Acknowledgements -- Chapter 1 : Framing Analytical Questions -- Data is the New Oil -- The World of the Business Data Analyst -- How Does Data Analysis Relate to Decision Making? -- How Do We Frame Analytical Questions? -- What are the Characteristics of Well-framed Analytical Questions? -- Exercise 1.1 - Case Study Using Dataset K: Titanic Disaster -- What are Some Examples of Text-Based Analytical Questions? -- Additional Case Study Using Dataset J: Remote Learning Student Survey -- References -- Chapter 2 : Analytical Tool Sets -- Tool Sets for Text Analytics -- Excel -- Microsoft Word -- Adobe Acrobat -- SAS JMP -- R and RStudio -- Voyant -- Java -- Stanford Named Entity Recognizer (NER) -- Topic Modeling Tool -- References -- Chapter 3 : Text Data Sources and Formats -- Sources and Formats of Text Data -- Social Media Data -- Customer opinion data from commercial sites -- Email -- Documents -- Surveys -- Websites -- Chapter 4 : Preparing the Data File -- What is Data Shaping? -- The Flat File Format -- Shaping the Text Variable in a Table -- Bag-of-Words Representation -- Single Text Files -- Exercise 4.1 - Case Study Using Dataset L: Resumes -- Exercise 4.2 - Case Study Using Dataset D: Occupation Descriptions -- Additional Exercise 4.3 - Case Study Using Dataset I: NAICS Codes -- Aggregating Across Rows and Columns --

Exercise 4.4 - Case Study Using Dataset D: Occupation Descriptions -- Additional Advanced Exercise 4.5 - Case Study Using Dataset E: Large Data Files -- Additional Advanced Exercise 4.6 - Case Study Using Dataset F: The Federalist Papers -- References -- Chapter 5 : Word Frequency Analysis -- What is Word Frequency Analysis? -- How Does It Apply to Text Business Data Analysis? -- Exercise 5.1 - Case Study Using Dataset A: Training Survey.

Exercise 5.2 - Case Study Using Dataset D: Job Descriptions -- Exercise 5.3 - Case Study Using Dataset C: Product Reviews -- Additional Exercise 5.4 - Case Study Using Dataset B: Consumer Complaints -- Chapter 6 : Keyword Analysis -- Exercise 6.1 - Case Study Using Dataset D: Resume and Job Description -- Exercise 6.2 - Case Study Using Dataset G: University Curriculum -- Exercise 6.3 - Case Study Using Dataset C: Product Reviews -- Additional Exercise 6.4 - Case Study Using Dataset B: Customer Complaints -- Chapter 7 : Sentiment Analysis -- What is Sentiment Analysis? -- Exercise 7.1 - Case Study Using Dataset C: Product Reviews - Rubbermaid -- Exercise 7.2 - Case Study Using Dataset C: Product Reviews-Windex -- Exercise 7.3 - Case Study Using Dataset C: Product Reviews-Both Brands -- Chapter 8 : Visualizing Text Data -- What Is Data Visualization Used For? -- Exercise 8.1 - Case Study Using Dataset A: Training Survey -- Exercise 8.2 - Case Study Using Dataset B: Consumer Complaints -- Exercise 8.3 - Case Study Using Dataset C: Product Reviews -- Exercise 8.4 - Case Study Using Dataset E: Large Text Files -- References -- Chapter 9 : Coding Text Data -- What is a Code? -- What are the Common Approaches to Coding Text Data? -- What is Inductive Coding? -- Exercise 9.1 - Case Study Using Dataset A: Training -- Exercise 9.2 - Case Study Using Dataset J: Remote Learning -- Exercise 9.3 - Case Study Using Dataset E: Large Text Files -- Affinity Diagram Coding -- Exercise 9.4 - Case Study Using Dataset M: Onboarding Brainstorming -- References -- Chapter 10 : Named Entity Recognition -- Named Entity Recognition -- What is a Named Entity? -- Common Approaches to Extracting Named Entities -- Classifiers - The Core NER Process -- What Does This Mean for Business? -- Exercise 10.1 - Using the Stanford NER -- Exercise 10.2 - Example Cases.

Exercise 10.2 - Case Study Using Dataset H: Corporate Financial Reports -- Additional Exercise 10.3 - Case Study Using Dataset L: Corporate Financial Reports -- Exercise 10.4 - Case Study Using Dataset E: Large Text Files -- Additional Exercise 10.5 - Case Study Using Dataset E: Large Text Files -- References -- Chapter 11 : Topic Recognition in Documents -- Information Retrieval -- Document Characterization -- Topic Recognition -- Exercises -- Exercise 11.1 - Case Study Using Dataset G: University Curricula -- Exercise 11.2 - Case Study Using Dataset E: Large Text Files -- Exercise 11.3 - Case Study Using Dataset E: Large Text Files -- Exercise 11.4 - Case Study Using Dataset E: Large Text Files -- Exercise 11.5 - Case Study Using Dataset E: Large Text Files -- Additional Exercise 11.6 - Case Study Using Dataset P: Patents -- Additional Exercise 11.7 - Case Study Using Dataset F: Federalist Papers -- Additional Exercise 11.8 - Case Study Using Dataset E: Large Text Files -- Additional Exercise 11.9- Case Study Using Dataset N: Sonnets -- References -- Chapter 12 : Text Similarity Scoring -- What is Text Similarity Scoring? -- Text Similarity Scoring Exercises -- Exercise 12.1 - Case Study Using Dataset D: Occupation Description -- Analysis using R -- Exercise 12.2 - Case D: Resume and Job Description -- Reference -- Chapter 13 : Analysis of Large Datasets by Sampling -- Using Sampling to Work with Large Data Files -- Exercise 13.1 - Big Data Analysis -- Additional Case Study Using Dataset E: BankComplaints Big Data File -- Chapter 14 :

Installing R and RStudio -- Installing R -- Install R Software for a Mac System -- Installing RStudio -- Reference -- Chapter 15 : Installing the Entity Extraction Tool -- Downloading and Installing the Tool -- The NER Graphical User Interface -- Reference -- Chapter 16 : Installing the Topic Modeling Tool.

Installing and Using the Topic Modeling Tool -- Install the tool -- For Macs -- For Windows PCs -- UTF-8 caveat -- Setting up the workspace -- Workspace Directory -- Using the Tool -- Select metadata file -- Selecting the number of topics -- Analyzing the Output -- Multiple Passes for Optimization -- The Output Files -- Chapter 17 : Installing the Voyant Text Analysis Tool -- Install or Update Java -- Installation of Voyant Server -- The Voyant Server -- Downloading VoyantServer -- Running Voyant Server -- Controlling the Voyant Server -- Testing the Installation -- Reference -- INDEX.

Sommario/riassunto

With the rise in data science development, we now have many remarkable techniques and tools to extend data analysis from numeric and categorical data to textual data. Sifting through the open-ended responses from a survey, for example, was an arduous process when performed by hand. Using a case study approach, this book was written for business analysts who wish to increase their skills in extracting answers for text data in order to support business decision making. Most of the exercises use Excel, today's most common analysis tool, and R, a popular analytic computer environment. The techniques covered range from the most basic text analytics, such as key word analysis, to more sophisticated techniques, such as topic extraction and text similarity scoring. Companion files with numerous datasets are included for use with case studies and exercises. FEATURES: Organized by tool or technique, with the basic techniques presented first and the more sophisticated techniques presented later. Uses Excel and R for datasets in case studies and exercises. Features the CRISP-DM data mining standard with early chapters for conducting the preparatory steps in data mining. Companion files with numerous datasets and figures from the text. The companion files are available online by emailing the publisher with proof of purchase at info@merclearning.com.
