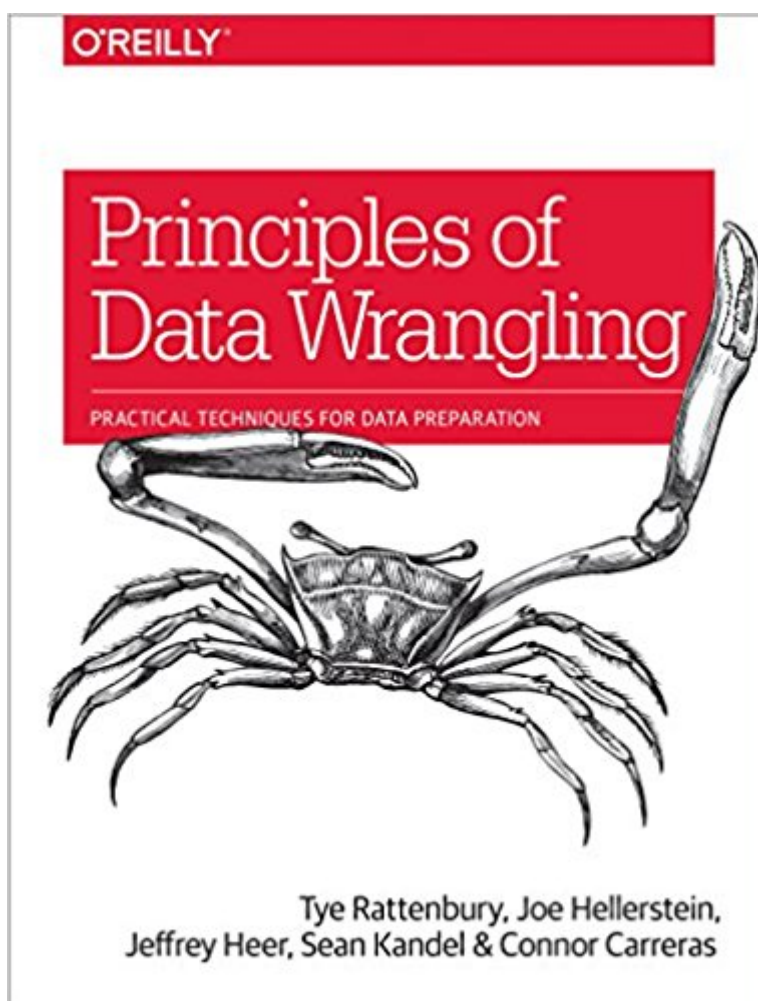


The book was found

Principles Of Data Wrangling: Practical Techniques For Data Preparation



Synopsis

A key task that any aspiring data-driven organization needs to learn is data wrangling, the process of converting raw data into something truly useful. This practical guide provides business analysts with an overview of various data wrangling techniques and tools, and puts the practice of data wrangling into context by asking, "What are you trying to do and why?" Wrangling data consumes roughly 50-80% of an analyst's time before any kind of analysis is possible. Written by key executives at Trifacta, this book walks you through the wrangling process by exploring several factors—time, granularity, scope, and structure—that you need to consider as you begin to work with data. You'll learn a shared language and a comprehensive understanding of data wrangling, with an emphasis on recent agile analytic processes used by many of today's data-driven organizations. Appreciate the importance and the satisfaction of wrangling data the right way. Understand what kind of data is available. Choose which data to use and at what level of detail. Meaningfully combine multiple sources of data. Decide how to distill the results to a size and shape that can drive downstream analysis.

Book Information

File Size: 2846 KB

Print Length: 96 pages

Page Numbers Source ISBN: 1491938927

Simultaneous Device Usage: Unlimited

Publisher: O'Reilly Media; 1 edition (June 29, 2017)

Publication Date: June 29, 2017

Sold by: Amazon Digital Services LLC

Language: English

ASIN: B073HMH8XG

Text-to-Speech: Enabled

X-Ray: Not Enabled

Word Wise: Not Enabled

Lending: Not Enabled

Enhanced Typesetting: Not Enabled

Best Sellers Rank: #540,443 Paid in Kindle Store (See Top 100 Paid in Kindle Store) #8

in Kindle Store > Kindle eBooks > Nonfiction > Science > Mathematics > Research #45

in Books > Science & Math > Mathematics > Research #154 in Books > Business &

Customer Reviews

View larger From the Forward The phrase data wrangling, born in the modern context of agile analytics, is meant to describe the lion's share of the time people spend working with data. There is a common misperception that data analysis is mostly a process of running statistical algorithms on high-performance data engines. In practice, this is just the final step of a longer and more complex process; 50 to 80 percent of an analyst's time is spent wrangling data to get it to the point at which this kind of analysis is possible. Not only does data wrangling consume most of an analyst's workday, it also represents much of the analyst's professional process: it captures activities like understanding what data is available; choosing what data to use and at what level of detail; understanding how to meaningfully combine multiple sources of data; and deciding how to distill the results to a size and shape that can drive downstream analysis. These activities represent the hard work that goes into both traditional data 'curation' and modern data analysis. And in the context of agile analytics, these activities also capture the creative and scientific intuition of the analyst, which can dictate different decisions for each use case and data source. We have been working on these issues with data-centric folks of various stripes—from the IT professionals who fuel data infrastructure in large organizations, to professional data analysts, to data-savvy 'enthusiasts' in roles from marketing to journalism to science and social causes. Much is changing across the board here. This book is our effort to wrangle the lessons we have learned in this context into a coherent overview, with a specific focus on the more recent and quickly growing agile analytic processes in data-driven organizations. Hopefully, some of these lessons will help to clarify the importance—and yes, the satisfaction—of data wrangling done well.

[Download to continue reading...](#)

Principles of Data Wrangling: Practical Techniques for Data Preparation Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython Data Analytics: What Every Business Must Know About Big Data And Data Science (Data Analytics for Business, Predictive Analysis, Big Data Book 1) Data Analytics: Applicable Data Analysis to Advance Any Business Using the Power of Data Driven Analytics (Big Data Analytics, Data Science, Business Intelligence Book 6) Big Data For Business: Your Comprehensive Guide to Understand Data Science, Data Analytics and Data Mining to Boost More Growth and Improve Business - Data Analytics Book, Series 2 Of Mess and Moxie:

Wrangling Delight Out of This Wild and Glorious Life Analytics: Data Science, Data Analysis and Predictive Analytics for Business (Algorithms, Business Intelligence, Statistical Analysis, Decision Analysis, Business Analytics, Data Mining, Big Data) Data Mining: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques, Third Edition (Morgan Kaufmann Series in Data Management Systems) Data Mining: Practical Machine Learning Tools and Techniques, Second Edition (Morgan Kaufmann Series in Data Management Systems) Certified Paralegal Review Manual: A Practical Guide to CP Exam Preparation (Test Preparation) Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data Analytics: Business Intelligence, Algorithms and Statistical Analysis (Predictive Analytics, Data Visualization, Data Analytics, Business Analytics, Decision Analysis, Big Data, Statistical Analysis) Data Analytics and Python Programming: 2 Bundle Manuscript: Beginners Guide to Learn Data Analytics, Predictive Analytics and Data Science with Python Programming Data Analytics For Beginners: Your Ultimate Guide To Learn and Master Data Analysis. Get Your Business Intelligence Right – Accelerate Growth and Close More Sales (Data Analytics Book Series) Discovering Knowledge in Data: An Introduction to Data Mining (Wiley Series on Methods and Applications in Data Mining) Principles & Techniques of Patient Care, 4e (Principles and Techniques of Patient Care) Basics of R/C Model Aircraft Design: Practical Techniques for Building Better Models: Practical Techniques for Building Better Models

[Contact Us](#)

[DMCA](#)

[Privacy](#)

[FAQ & Help](#)