



Practical Data Science for Information Professionals

David Stuart

List price £50.00

Product Details

Format: Paperback

ISBN: 9781783303441

Published: 24th Jul 2020

Publisher: Facet Publishing

Dimensions: 234mm x 156mmmm

Description

Practical Data Science for Information Professionals provides an accessible introduction to a potentially complex field, providing readers with an overview of data science and a framework for its application. It provides detailed examples and analysis on real data sets to explore the basics of the subject in three principle areas: clustering and social network analysis; predictions and forecasts; and text analysis and mining.

As well as highlighting a wealth of user-friendly data science tools, the book also includes some example code in two of the most popular programming languages (*R* and *Python*) to demonstrate the ease with which the information professional can move beyond the graphical user interface and achieve significant analysis with just a few lines of code.

After reading, readers will understand:

- the growing importance of data science
- the role of the information professional in data science
- some of the most important tools and methods that information professionals can use.

Bringing together the growing importance of data science and the increasing role of information professionals in the management and use of data, *Practical Data Science for Information Professionals* will provide a practical introduction to the topic specifically designed for the information community. It will appeal to librarians and information professionals all around the world, from large academic libraries to small research libraries. By focusing on the application of open source software, it aims to reduce barriers for readers to use the lessons learned within.

Contents

Contents

Figures

Tables**Boxes****Preface****1 What is data science?**

Data, information, knowledge, wisdom

Data everywhere

The data deserts

Data science

The potential of data science

From research data services to data science in libraries

Programming in libraries

Programming in this book

The structure of this book

2 Little data, big data

Big data

Data formats

Standalone files

Application programming interfaces

Unstructured data

Data sources

Data licences

3 The process of data science

Modelling the data science process

Frame the problem

Collect data

Transform and clean data

Analyse data

Visualise and communicate data

Frame a new problem

4 Tools for data analysis

Finding tools

Software for data science

Programming for data science

5 Clustering and social network analysis

Network graphs

Graph terminology

Network matrix

Visualisation

Network analysis

6 Predictions and forecasts

Predictions and forecasts beyond data science

Predictions in a world of (limited) data

Predicting and forecasting for information professionals

Statistical methodologies

7 Text analysis and mining

Text analysis and mining, and information professionals

Natural language processing

Keywords and n-grams

8 The future of data science and information professionals

Eight challenges to data science

Ten steps to data science librarianship

The final word: play

References

Appendix – Programming concepts for data science

Variables, data types and other classes

Import libraries

Functions and methods

Loops and conditionals

Final words of advice

Further reading

Index

Author

David Stuart is an independent information professional and an honorary research fellow at the University of Wolverhampton, and was previously a research fellow at King's College London and the University of Wolverhampton. He regularly publishes in peer-reviewed academic journals and professional journals on information science, metrics, and semantic web technologies. He is author of *Practical Ontologies for Information Professionals* (2016), *Web Metrics for Library and Information Professionals* (2014), and *Facilitating Access to the Web of Data* (2011).