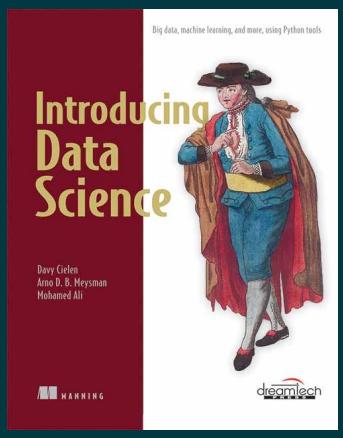
WILEY



Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools

By Davy Cielen, Arno D.B. Meysman, Mohamed Ali

Paperback 9789351199373 **₹849.00**

Pages: 320

Description

Introducing Data Science explains vital data science concepts and teaches you how to accomplish the fundamental tasks that occupy data scientists. You'll explore data visualization, graph databases, the use of NoSQL, and the data science process. You'll use the Python language and common Python libraries as you experience firsthand the challenges of dealing with data at scale. Discover how Python allows you to gain insights from data sets so big that they need to be stored on multiple machines, or from data moving so quickly that no single machine can handle it.

About the Author

Davy Cielen, Arno D.B. Meysman, Mohamed Ali

Davy Cielen

Table of Contents

1. Data science in a Big Data world

1.1. Benefits and uses of data science and Big Data

1.2. Facets of data

1.3. The data science process

1.4. The Big Data ecosystem and data science

1.5. An introductory working example of Hadoop

1.6. Summary

2. The data science process

2.1. Overview of the data science process

2.2. Step 1: defining research goals and creating a project charter

2.3. Step 2: retrieving data

2.4. Step 3: cleansing, integrating, and transforming data

2.5. Step 4: exploratory data analysis

2.6. Step 5: Build the models

2.7. Step 6: Presenting findings and building applications on top of them

2.8. Summary

3. Machine learning

3.1. What is machine learning and why should you care about it?

3.2. The modelling process

3.3. Types of machine learning

3.4. Semi-supervised learning

3.5. Summary

4. Handling large data on a single computer

4.1. The problems you face when handling large data

4.2. General techniques for handling large volumes of data

4.3. General programming tips for dealing with large datasets

4.4. Case study 1: predicting malicious URLs

4.5. Case study 2: building a recommender system inside a database

4.6. Summary

5. First steps in Big Data

5.1. Distributing data storage and processing with frameworks

5.2. Case study: assessing risk when loaning money

5.3. Summary

6. Join the NoSQL movement

6.1. Introduction to NoSQL

6.2. Case study: what disease is that?

6.3. Summary

7. The rise of graph databases

7.1. Introducing connected data and graph databases

7.2. Introducing Neo4j: a graph database

7.3. Connected data example: a recipe recommendation engine

7.4. Summary

8. Text mining and text analytics

8.1. Text mining in the real world

8.2. Text mining techniques

8.3. Case study: classifying Reddit posts

8.4. Summary

9. Data visualization to the end user

To purchase this product, please visit:
https://www.wileyindia.com/introducing-data-science-big-data-machine-learning-and-more-using-python-tools.html

Scan to buy