

DATASCI 400: Main Textbooks

All of the books below are available for free to UW students via <https://learning.oreilly.com/home/>. You will need to use your UW NetID to access this page. Sign in using your UW NetID email and click on the *Using Single Sign-on (SSO)?* link above the *Password* field. You will be directed to a UW sign-in and then back to learning.oreilly.com. Both sites may take a few seconds to load.

| | |
|---------------------|--|
| Title | Python Data Science Handbook |
| Author | Jake VanderPlas |
| Publisher | O'Reilly Media, Inc. |
| ISBN | 9781491912058 |
| Abbreviation | PDSH |

| | |
|---------------------|--|
| Title | Data Science from Scratch, 2nd Edition |
| Author | Joel Grus |
| Publisher | O'Reilly Media, Inc. |
| ISBN | 9781492041139 |
| Abbreviation | DSS |

| | |
|--|--|
| Lesson 1: Intro to Data Science | DSS Chapter 2 - A Crash Course in Python |
| Lesson 2: Nature of Data | PDSH Chapter 2 - Introduction to NumPy DSS Chapter 24 - Databases and SQL DSS Chapter 4 - Linear Algebra |
| Milestone 1: Data Flow Diagram | |
| Lesson 3: Aberrant Data | DSS Chapter 10 - Cleaning and Munging DSS Chapter 5 - Statistics |
| Lesson 4: Numeric Data | PDSH Chapter 3 - Data Manipulation with Pandas PDSH Chapter 4 - Visualization with Matplotlib DSS Chapter 5 - Statistics |
| Lesson 5: Categorical Data | PDSH Chapter 5 - Categorical Features |
| Lesson 6: Data Structures | PDSH Chapter 3 - Working with Time Series |
| Milestone 2: Data Preparation | PDSH Chapter 5 - Imputation of Missing Data PDSH Chapter 5 - Categorical Features PDSH Chapter 3 - Data Manipulation with Pandas DSS Chapter 9 - Scraping the Web |
| Lesson 7: Unsupervised Learning | DSS Chapter 20 - Clustering |
| Lesson 8: Predictive Analytics - Intro to Supervised Learning | PDSH Chapter 5 - In Depth: Linear Regression |
| Lesson 9: Accuracy Measures | |
| Lesson 10: Simple Statistical Fallacies | |
| Milestone 3: Data Model | PDSH Chapter 5 - Introducing Scikit-Learn |

Additional Textbooks

All of the books below are available for free to UW students via <https://learning.oreilly.com/home/>

| | |
|------------------|---|
| Title | Principles of Data Science - Second Edition |
| Author | Sunil Kakade, Sinan Ozdemir |
| Publisher | Packt Publishing |
| ISBN | 9781789804546 |

| | |
|------------------|--|
| Title | Python Data Science Essentials - Third Edition |
| Author | Luca Massaron, Alberto Boschetti |
| Publisher | Packt Publishing |
| ISBN | 9781789537864 |

| | |
|------------------|---|
| Title | Mastering pandas - Second Edition |
| Author | Ashish Kumar |
| Publisher | Packt Publishing |
| ISBN | 9781789343236 |

| | |
|------------------|---|
| Title | Python Feature Engineering Cookbook |
| Author | Soledad Galli |
| Publisher | Packt Publishing |
| ISBN | 9781789806311 |

| | |
|------------------|--|
| Title | Pandas for Everyone: Python Data Analysis, First Edition |
| Author | Daniel Y. Chen |
| Publisher | Addison-Wesley Professional |
| ISBN | 9780134547046 |

| | |
|------------------|--|
| Title | Modern Python Cookbook |
| Author | Steven F. Lott |
| Publisher | Packt Publishing |
| ISBN | 9781786469250 |

| | |
|------------------|---|
| Title | Hands-On Unsupervised Learning Using Python |
| Author | Ankur A. Patel |
| Publisher | O'Reilly Media, Inc. |
| ISBN | 9781492035640 |

| | |
|------------------|---|
| Title | Mastering Machine Learning with scikit-learn - Second Edition |
| Author | Gavin Hackeling |
| Publisher | Packt Publishing |
| ISBN | 9781788299879 |

| | |
|------------------|---|
| Title | Essential Math for Data Science |
| Author | Hadrien Jean |
| Publisher | O'Reilly Media, Inc. |
| ISBN | 9781098115562 |
| Notes | To be published in September 2020 |

| | |
|---------------|---|
| Title | Visual Models for Software Requirements |
| Author | Joy Beatty and Anthony Chen |

| | |
|------------------|-----------------|
| Publisher | Microsoft Press |
| ISBN | 9780735667730 |