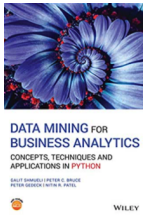
Module 1 Resources

Required Readings



Shmueli, G., Bruce, P. C., Gedeck, P., & Patel, N. R. (2020). *Data mining for business analytics: Concepts, techniques and applications in Python.* Wiley.

- Chapter 7
- Chapter 10
- Review Chapter 2
- Review Chapter 5
- **Python Code for All Chapter Examples** (https://sandiego.instructure.com/courses/13937/files/1936536?wrap=1) (https://sandiego.instructure.com/courses/13937/files/1936536/download?download_frd=1) (ZIP file)
- **Datasets** (https://sandiego.instructure.com/courses/13937/files/1936502?wrap=1) (https://sandiego.instructure.com/courses/13937/files/1936502/download?download_frd=1) (ZIP file)
- **Textbook Errata** (https://www.dataminingbook.com/content/errata-python-edition)
- Hint: Due to compatibility with third-party packages, it is recommended that you install and use older versions of Python, such as 3.6.X
- Read the **Application of k-Nearest Neighbor (kNN) Approach for Predicting Economic Events: Theoretical Background** (https://sandiego.instructure.com/courses/13937/files/1936529?wrap=1) article.

Recommended Readings

- Read the **Gentle Introduction to Scikit-Learn: A Python Machine Learning Library** (https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/) article.
- Read the **K-Nearest Neighbors for Machine Learning** (https://machinelearningmastery.com/k-nearest-neighbors-for-machine-learning/#:~:text=KNN%20for%20Classification&text=If%20you%20are%20using%20K,an%20odd%20number%20of%20classes.) article.
- Read the **Tune Hyperparameters for Classification Machine Learning Algorithms** (https://machinelearningmastery.com/hyperparameters-for-classification-machine-learning-algorithms/) article.
- Read the **How to Choose Optimal Value of K in KNN Algorithm** (https://sandiego.instructure.com/courses/13937/files/1936544?wrap=1) article.

Required Media

- Watch the following lectures by Dr. Shmueli at NTHU:
  - **BADM 1.1: Data Mining Applications** (https://www.youtube.com/watch?v=u9zIWkkCX4g&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=2) (11:58) Closed captioning on site.
  - **BADM 1.2: Data Mining in a Nutshell** (https://www.youtube.com/watch?v=ZXW1lDaZvzQ&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=3) (11:03) Closed captioning on site.
  - **BADM 1.3: The Holdout Set** (https://www.youtube.com/watch?v=v8SgRpn0EZE&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=4) (06:29) Closed captioning on site.

- **BADM 2.1: Data Visualization** (https://www.youtube.com/watch?v=hdGO9dvWhlU&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=5) (16:21) Closed captioning on site.
- **BADM 2.2: Data Preparation** (https://www.youtube.com/watch?v=_M_nF7aTQzQ&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=6) (12:24) Closed captioning on site.
- **BADM 7.1 K-Nearest Neighbors** (https://www.youtube.com/watch?v=X4lc1ui3jD0) (19:10) Closed captioning on site.
- **BADM 9.2 Logistic Regression for Classification** (https://www.youtube.com/watch?v=LaOzjAaHvwg&list=PLoK4olB1jeK0VeChKj1F0s6w6jl3C2RUa&index=28) (18:32) Closed captioning on site.

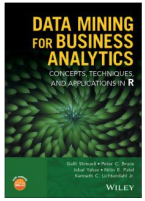
Recommended Media

- Watch the **KNN Algorithm Using Python | How KNN Algorithm Works | Data Science For Beginners | Simplilearn** (https://www.youtube.com/watch?v=4HKqjENq9OU) video. (27:42) Closed captioning on site.
- Watch the **KNN Algorithm using Python | How KNN Algorithm Works | Python Data Science Training | Edureka** (https://www.youtube.com/watch?v=6kZ-OPLNcgE) video. (19:51) Transcript available upon request.
- Watch the **KNN Algorithm - Python Scikit-Learn Tutorial** (https://www.youtube.com/watch?v=t-DJiq7RIPc) video. (09:50) Transcript available upon request.
- Watch the **Logistic Regression in Python | Logistic Regression Example | Machine Learning Algorithms | Edureka** (https://www.youtube.com/watch?v=VCJdg7YBbAQ&t=2394s) video. (53:40) Transcript available upon request.
- Watch the **StatQuest: Logistic Regression** (https://www.youtube.com/watch?v=yIYKR4sgzI8) video. (08:47) Closed captioning on site.
- Watch the **Logistic Regression in Python Step by Step in 10 Minutes** (https://www.youtube.com/watch?v=HYcXgN9HaTM) video. (10:06) Closed captioning on site.

Supplemental R Programming Resources

The authors and publisher of the *Data Mining for Business Analytics* textbook offer content using R programming. The list of recommended readings, resources, and media is provided for students who are interested in accumulating knowledge in R.

- Read the **R Markdown Cheat Sheet** (https://rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf) .
- Watch the **KNN Algorithm Using R | KNN Algorithm Example | Data Science Training | Edureka** (https://www.youtube.com/watch?v=XSoau_q0kz8&t=3s) video. (24:58) Transcript available upon request.
- Watch the **Logistic Regression in R | Machine Learning Algorithms | Data Science Training | Edureka** (https://www.youtube.com/watch?v=Z5WKQr4H4Xk) video. (01:09:11) Closed captioning on site.
- Watch the **Applying and Understanding K-Nearest Neighbors (KNN) in R** (https://www.youtube.com/watch?v=htnZp__02qw) video. (12:19) Closed captioning on site.
- Watch the **Logistic Regression Models in R** (https://www.youtube.com/watch?v=jAc2SNPTmLY) video. (06:22) Closed captioning on site.



Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl Jr., K. C. (2018). *Data mining for business analytics: Concepts, techniques, and applications in R.* Wiley.

- Chapter X
- **R Code for All Chapter Examples** (https://sandiego.instructure.com/courses/13937/files/1936539?wrap=1) (https://sandiego.instructure.com/courses/13937/files/1936539/download?download_frd=1) (ZIP file)
- **Datasets** (https://sandiego.instructure.com/courses/13937/files/1936541?wrap=1) (https://sandiego.instructure.com/courses/13937/files/1936541/download?download_frd=1) (ZIP file)
- **Textbook Errata** (https://www.dataminingbook.com/content/errata-r-edition)