

# BUS 536 Special Topics in Business Analytics

## Spring 2024

Time: Tuesday 6:30 p.m. – 9:30 p.m.

Location: Sierra Hall 1242

Office Hours: Tuesday 5:30 p.m. – 6:30 p.m.

Instructor: Dr. Minder Chen, Professor of MIS  
Email: [Minder.Chen@csuci.edu](mailto:Minder.Chen@csuci.edu)

Office location: Martin V. Smith Hall 1177  
Course website: Canvas

## Course Objectives

In-depth analysis of current topics in business analytics. Topics may vary each semester. This comprehensive course is designed for MBA students who aim to gain a deep understanding of Business Analytics (BA) while simultaneously developing practical skills in Python programming. The course is structured to provide a balanced mix of theoretical knowledge and hands-on experience in both business analytics and Python, catering to the evolving needs of the business world.

## Textbooks

1. Jeffrey D. Camm, etc., *Business Analytics*, 5<sup>th</sup> edition, Cengage Learning, 2024, <https://www.cengage.com/c/business-analytics-5e-camm/9780357902202/> (Don't use any other website. Get the 4-month rental will be sufficient.)
2. Wes McKinney, *Python for Data Analysis*, 3<sup>rd</sup> edition, September 2022, O'Reilly Media; open edition by the author at <https://wesmckinney.com/book/> and code at <https://wesmckinney.com/book/>

## References

3. \*Foster Provost and Tom Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*, 2013, O'Reilly Media, Inc., p. 20  
<https://learning.oreilly.com/library/view/data-science-for/9781449374273/> & Audiobook version at <https://learning.oreilly.com/videos/data-science-for/9781663728265/>
4. Allen Downey, *Think Python: How to Think Like a Computer Scientist*, 2nd Edition, 2016, Green Tea Press, free version by the author at <https://greenteapress.com/thinkpython2/thinkpython2.pdf>
5. Jake VanderPlas, *Python Data Science Handbook*, Nov. 2016, O'Reilly Media, Inc., free version by the author at <https://jakevdp.github.io/PythonDataScienceHandbook/> Particularly Chapter 5. Code on Github at <https://github.com/jakevdp/PythonDataScienceHandbook/>
6. Leo Porter and Daniel Zingaro, *Learn AI-Assisted Python Programming with GitHub Copilot and ChatGPT*, September 2023, Manning. <https://www.manning.com/books/learn-ai-assisted-python-programming>

7. <https://www.edx.org/learn/python/harvard-university-cs50-s-introduction-to-programming-with-python>
8. Ossama Embarak, etc., *Data Analysis and Visualization Using Python*, Apress, 2018. At [https://www.academia.edu/37886932/Data\\_Analysis\\_and\\_Visualization\\_Using\\_Python\\_Dr\\_Ossama\\_Embarak\\_pdf](https://www.academia.edu/37886932/Data_Analysis_and_Visualization_Using_Python_Dr_Ossama_Embarak_pdf)
9. *Machine Learning Systems with TinyML*, [https://harvard-edge.github.io/cs249r\\_book/](https://harvard-edge.github.io/cs249r_book/) (This is mainly for embedded systems but contains some general contents applicable to this course.)

## Python

10. Python.org, *The Python Tutorial*, at <https://docs.python.org/3/tutorial/index.html>
11. <https://www.w3schools.com/python/default.asp> (Python Tutorial) and [https://www.w3schools.com/python/python\\_ml\\_getting\\_started.asp](https://www.w3schools.com/python/python_ml_getting_started.asp) (Machine Learning)
12. Pandas Cheat Sheet at [https://pandas.pydata.org/Pandas\\_Cheat\\_Sheet.pdf](https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf)
13. Python Pandas Getting Started Tutorial at [https://pandas.pydata.org/docs/getting\\_started/intro\\_tutorials/](https://pandas.pydata.org/docs/getting_started/intro_tutorials/)
14. <https://scikit-learn.org/stable/tutorial/index.html>

## Schedule

<i>Date</i>	<i>Topics</i>	<i>Comments</i>
1/23	BA: Introduction to Business Analytics PY: Introduction to Python Programming & IDE	BA:Ch. 1 PY:Ch. 1, 2
1/30	BA: Descriptive Statistics PY: Python Built-in Data Types & Data Structures	BA:Ch. 2 PY:Ch. 3
2/6	BA: Data Visualization PY: <b>MatPlotLib</b> for Data Visualization	BA:Ch. 3 PY:Ch. 9
2/13	BA: Data Wrangling PY: <b>Pandas</b> : The Excel in Python	BA:Ch. 4 PY:Ch. 5, 6, 7, 8
2/20	BA: Probability PY: <b>Numpy</b> : Multidimensional array for scientific computing in Python	BA:Ch. 5 PY:Ch. 4
2/27	BA: Descriptive Data Mining: Dimension Reduction and Cluster Analysis PY: Principal Components Analysis(CPA), K-Means, & Hierarchical Clustering	BA:Ch. 6.1 & 6.2
3/5	BA: Descriptive Data Mining: Association Rules and Text Mining PY: MLxtend (Machine Learning Extensions) & NLTK (Natural Language Toolkit)	BA:Ch. 6.3 & 6.4

3/12	Orange Data Mining Tool <a href="https://orangedatamining.com/">https://orangedatamining.com/</a>	BA:Ch. 6 Appendix
3/19	<b>Spring Break (No class)</b>	
3/26	BA: Statistical Inference PY: Statsmodels	BA:Ch. 7
4/2	BA: Linear Regression PY: Scikit-Learn	BA:Ch. 8
4/9	BA: Predictive Data Mining: Regression Tasks PY: Scikit-Learn	BA:Ch. 10
4/16	BA: Predictive Data Mining: Classification Tasks PY: Scikit-Learn	BA:Ch. 11
4/23	BA: Time Series Analysis and Forecasting PY: Statsmodels	BA:Ch. 9 PY:Ch. 11
4/30	Deep Learning: TensorFlow.Keras Final Project Presentation	
5/7	Final Project Presentation	
5/14	Final Exam (Online) 7:00 p.m. to 9:00 p.m.	

## Grading Information

1. Class participation (including attendance & class discussion): 5%
2. Individual Assignments (Excel, Python, Orange, etc.): (50%)
3. Final Project: (15%)
4. Final Exam: 30% (Quiz and Written)

## Grading Scale

A	93 +	B	83 – 87	C	70 – 77
A -	90 – 92	B -	80 – 82	D	60 – 69
B +	88 – 89	C +	78 – 79	F	60<

## Readings

Additional reading materials will be assigned on Canvas course website.

## Attendance

Attendances are required except for excused absences and advanced notification of the instructor; attendance will be counted as part of your participation in the class.

## Equipment

You must have access to a PC at home or on campus. The PC should have adequate software products such as Microsoft Office and other software recommended by the instructor.

## CSUCI Disability Statement

Cal State Channel Islands is committed to equal educational opportunities for qualified students with disabilities in compliance with Section 504 of the Federal Rehabilitation Act of 1973 and the Americans with Disabilities Act (ADA) of 1990. The mission of Disability Accommodation Services is to assist students with disabilities to realize their academic and personal potential. Students with physical, learning, or other disabilities are encouraged to contact the Disability Accommodation Services office at (805) 437-8510 for personal assistance and accommodations. Please refer to CSUCI's [Disability Accommodations and Support Services](#) (DASS) website;).

## Academic honesty

Academic honest is expected by all the students in this class. Any academic dishonesty will not be tolerated. Academic dishonesty includes such things as cheating, inventing false information or citations, plagiarism and helping someone else commit an act of academic dishonesty. It usually involves an attempt by a student to show possession of a level of knowledge or skill that he/she does not possess. Students in this course should abide by Policy on Academic Dishonesty at [SP 13-06](#).

## Disclaimer

The syllabus is subject to change to reflect new materials, assignments, and background of students. Students should check the online version of the syllabus frequently.

## Orange

Download Orange from: <https://orangedatamining.com/download/>

Chapter 6: Case 2 & Problem 27-41

Chapter 11: Case 1 & Problem 40-56

[Getting Started with Orange - YouTube](#) Playlist

## Textbook Table of Contents

1. Jeffrey D. Camm, etc., *Business Analytics*, 5<sup>th</sup> edition, Cengage Learning, 2024, <https://www.cengage.com/c/business-analytics-5e-camm/9780357902202/> (Don't use any other website.)

### Table of Contents

1. Introduction.
2. Descriptive Statistics.
3. Data Visualization.
4. Data Wrangling.
5. Probability: An Introduction to Modeling Uncertainty.
6. Descriptive Data Mining.
7. Statistical Inference.
8. Linear Regression.
9. Time Series Analysis and Forecasting.
10. Predictive Data Mining: Regression.
11. Predictive Data Mining: Classification.
12. Spreadsheet Modeling.
13. Monte Carlo Simulation.
14. Linear Optimization Models.
15. Integer Linear Optimization Models.
16. Nonlinear Optimization Models.
17. Decision Analysis.

Appendix A: Basics of Excel.

Appendix B: Database Basics with Microsoft Access.

Appendix C: Solutions to Even-Numbered Questions (online).

Wes McKinney, *Python for Data Analysis*, 3<sup>rd</sup> edition, September 2022, O'Reilly Media; open edition by the author at <https://wesmckinney.com/book/>

## Preface

- 1 Preliminaries
- 2 Python Language Basics, IPython, and Jupyter Notebooks
- 3 Built-In Data Structures, Functions, and Files
- 4 NumPy Basics: Arrays and Vectorized Computation
- 5 Getting Started with pandas
- 6 Data Loading, Storage, and File Formats
- 7 Data Cleaning and Preparation
- 8 Data Wrangling: Join, Combine, and Reshape
- 9 Plotting and Visualization
- 10 Data Aggregation and Group Operations
- 11 Time Series
- 12 Introduction to Modeling Libraries in Python
- 13 Data Analysis Examples

## Appendices

- A Advanced NumPy
- B More on the IPython System

# Python Data Science Handbook

*Jake VanderPlas*

Code on Github at <https://github.com/jakevdp/PythonDataScienceHandbook/>

## Table of Contents

### [Preface](#)

## 1. IPython: Beyond Normal Python

- [Help and Documentation in IPython](#)
- [Keyboard Shortcuts in the IPython Shell](#)
- [IPython M agic Commands](#)
- [Input and Output History](#)
- [IPython and Shell Commands](#)
- [Errors and Debugging](#)
- [Profiling and Timing Code](#)
- [M ore IPython Resources](#)

## 2. Introduction to NumPy

- [Understanding Data Types in Python](#)
- [The Basics of NumPy Arrays](#)
- [Computation on NumPy Arrays: Universal Functions](#)
- [Aggregations: M in, M ax, and Everything In Between](#)
- [Computation on Arrays: Broadcasting](#)
- [Comparisons, M asks, and Boolean Logic](#)
- [Fancy Indexing](#)
- [Sorting Arrays](#)
- [Structured Data: NumPy's Structured Arrays](#)

## 3. Data M anipulation with Pandas

- [Introducing Pandas Objects](#)
- [Data Indexing and Selection](#)
- [Operating on Data in Pandas](#)
- [Handling M issing Data](#)
- [Hierarchical Indexing](#)
- [Combining Datasets: Concat and Append](#)
- [Combining Datasets: Merge and Join](#)
- [Aggregation and Grouping](#)
- [Pivot Tables](#)

- [Vectorized String Operations](#)
- [Working with Time Series](#)
- [High-Performance Pandas: eval\(\) and query\(\)](#)
- [Further Resources](#)

## [4. Visualization with Matplotlib](#)

- [Simple Line Plots](#)
- [Simple Scatter Plots](#)
- [Visualizing Errors](#)
- [Density and Contour Plots](#)
- [Histograms, Binnings, and Density](#)
- [Customizing Plot Legends](#)
- [Customizing Colorbars](#)
- [Multiple Subplots](#)
- [Text and Annotation](#)
- [Customizing Ticks](#)
- [Customizing Matplotlib: Configurations and Stylesheets](#)
- [Three-Dimensional Plotting in Matplotlib](#)
- [Geographic Data with Basemap](#)
- [Visualization with Seaborn](#)
- [Further Resources](#)

## [5. Machine Learning](#)

- [What Is Machine Learning?](#)
- [Introducing Scikit-Learn](#)
- [Hyperparameters and Model Validation](#)
- [Feature Engineering](#)
- [In Depth: Naive Bayes Classification](#)
- [In Depth: Linear Regression](#)
- [In-Depth: Support Vector Machines](#)
- [In-Depth: Decision Trees and Random Forests](#)
- [In Depth: Principal Component Analysis](#)



- [In-Depth: Manifold Learning](#)
- [In Depth: k-Means Clustering](#)
- [In Depth: Gaussian Mixture Models](#)
- [In-Depth: Kernel Density Estimation](#)
- [Application: A Face Detection Pipeline](#)
- [Further Machine Learning Resources](#)

## [Appendix: Figure Code](#)