
Python for Data Analytics

Course numbers:	410 / 510
Course credits:	4
Prerequisites:	Business Analytics II and admitted business major.
Instructor:	Saeed Piri
Contact:	spiri@uoregon.edu
Class time and location:	Monday/Wednesday 2:00 pm -3:50 pm, Chiles 225
Office hours and location:	Monday/Wednesday 12:00 pm -1:50 pm, or by appointment Lillis 434
Virtual office hours:	Monday/Wednesday 12:00 pm -1:50 pm, or by appointment Zoom (meeting ID: 983 1325 6320) Meeting URL: https://uoregon.zoom.us/j/98313256320

1. Lundquist College of Business Undergraduate Learning Goals

- Our students will be creative and analytical problem solvers.
 - Apply relevant data analytic tools, quantitative methods, and qualitative methods to identify, analyze, and solve business problems
 - Be able to conduct research to identify theories and evidence applicable to organizational or business problems
 - Consider multiple and divergent perspectives to analyze and solve problems
- Our students will be effective communicators
 - Give an oral presentation that is logical, compelling, and clear to the audience
 - Create a written document that is logical, compelling, grammatically correct, and clear to the intended audience
 - Communicate the results of the data analysis in a manner that is clear to a non-technical audience
- Our students will be able to effectively work in a global business environment
 - Identify cross-country differences in cultural, economic, social, and legal factors, and evaluate their implications for business decisions
 - Demonstrate cultural sensitivity in interpersonal interactions
- Our students will be skilled in leading and working successfully in diverse teams
 - Interact effectively with peers in a diverse work environment
 - Be able to manage team dynamics and recognize team members' strengths to facilitate effective team processes
 - Know how to create a work climate that respects diversity in its members
- Our students will be able to recognize and evaluate the ethical, legal and sustainability implications of business decisions
 - Recognize the environmental and social impacts of business decisions and evaluate the implications of alternative solutions
 - Recognize ethical dilemmas and evaluate the implications of alternative solutions
 - Recognize and apply the legal and regulatory implications of business decisions
- Our students will demonstrate proficiency in fundamental and concentration-specific business knowledge

Demonstrates proficiency in each of the foundational business disciplines; accounting, finance, marketing, management, and operations and business analytics

2. Overview of the Course

The focus of this course is on **Python** and its application to data analytics. **Python** is the most popular programming language for data science; therefore, our objective is to teach programming fundamentals in this language. Major data analytics libraries in **Python** will be discussed in this course. These libraries are Numpy, Pandas, and Scikit-learn.

This course introduces the basics concepts of **data analytics**, and the audiences of this course are senior undergraduate, Master's, and MBA students. **Data analytics** is one of the most popular fields in business. Analyzing historical data and applying various statistical and machine learning techniques enables us to develop models that can be used in different classification and pattern recognition problems such as credit card fraud detection, disease diagnosis, and predicting the success rate of a marketing campaign, among numerous other applications.

In this course, the main concepts of various techniques such as k-nearest neighbors, linear regression, logistic regression, lasso, ridge, decision tree, random forest, and support vector machine are introduced. In this course, theoretical concepts of **data analytics** will be fortified by applying them to real datasets.

Students who successfully complete this course will:

- Be able to use **Python** programming language and its data analytics library
- Be able to analyze data and develop predictive analytics models using **Python**
- Be able to establish the thinking process of defining and developing a real-world data-driven analytics project
- Learn the fundamentals of different data analytics techniques and how to implement them
- Learn how to evaluate various models and select the best model

3. Textbook

The following textbooks are **recommended (not required)**:

- **Introduction to Machine Learning with Python**, by Sarah Guido, Andreas Müller
<http://shop.oreilly.com/product/0636920030515.do>

E-version is available on UO Library for **free**:

[E-book- Introduction to Machine Learning with Python](#)

- **The Elements of Statistical Learning, Data Mining, Inference, and Prediction**, 2nd Edition by Hastie, T., Tibshirani, R., Friedman, J.

Available online for **free** at:

<https://web.stanford.edu/~hastie/ElemStatLearn/>

These books are not required, and they are both available online for free. While I will provide my own lecture notes throughout the term, these books serve as great compliments of the lecture notes and provide a lot more details.

4. Tentative Topic schedule¹

	Topic and learning objectives		Topic and learning objectives	
	Monday		Wednesday	
Week 1 Sep 26 th			Introduction to Data Analytics Introduction to Python	<ul style="list-style-type: none"> Intro to Data Analytics Introduction to programming in Python Variables, their types Functions, new function
Week 2 Oct 3 th	Introduction to Python	<ul style="list-style-type: none"> Strings Lists Dictionaries 	Numpy Library Pandas Library Data Preprocessing	<ul style="list-style-type: none"> Using Numpy library Using Pandas library Series Creating a DataFrame
Week 3 Oct 10 th	Pandas Library Data Preprocessing	<ul style="list-style-type: none"> Reading data into a DataFrame DataFrame components Sorting DataFrames Subsetting DataFrames 	Pandas Library Data Preprocessing	<ul style="list-style-type: none"> Subsetting DataFrames Adding/Removing data to/from DataFrames Changing labels and values in DataFrames Descriptive analytics Querying Data
Week 4 Oct 17 th	Pandas Library Scikit-learn k-Nearest Neighbors	<ul style="list-style-type: none"> Querying Data Missing data Review of k-nearest neighbors (classification) 	Scikit-learn k-Nearest Neighbors Fair Evaluation	<ul style="list-style-type: none"> Review of k-nearest neighbors (Regression) Train test splits Modeling in Python using Scikit-learn
Week 5 Oct 24 th	Scikit-learn Fair Evaluation Linear Regression	<ul style="list-style-type: none"> Cross-Validation Generalization, Overfitting, and Underfitting Review of linear regression Modeling in Python 	Scikit-learn Lasso and Ridge	<ul style="list-style-type: none"> Review of ridge regression, lasso regression Modeling in Python using Scikit-learn
Week 6 Oct 31 st	Scikit-learn Categorical variables Multiclass	<ul style="list-style-type: none"> Handling categorical variables Multiclass problems Modeling in Python using Scikit-learn 	Project Proposal Presentations	
Week 7 Nov 7 th	Scikit-learn Logistics Regression	<ul style="list-style-type: none"> Review of logistics regression Modeling in Python using Scikit-learn 	Scikit-learn Linear Support Vector Machine	<ul style="list-style-type: none"> Linear support vector machine Modeling in Python using Scikit-learn
Week 8 Nov 14 th	MIDTERM EXAM		Scikit-learn Decision Tree	<ul style="list-style-type: none"> Classification trees Modeling in Python using Scikit-learn
Week 9 Nov 21 th	Scikit-learn Decision Tree (Continued)	<ul style="list-style-type: none"> Regression trees Modeling in Python using Scikit-learn 	Scikit-learn Random Forest	<ul style="list-style-type: none"> Random Forests for Classification and Regression Term Project Q&A
Week 10 Nov 28 th	Term Project Presentations		Term Project Presentations	

¹ Subject to change.

5. Assessment

Assessment	Relative Weight
In-class Activities	5%
Quizzes	24%
Midterm Exam	33%
Term Project	38%
Total	100%

Note: if you notice any issues with your grades (quiz, project, or midterm) and want to discuss it with me, you have only **one week** after the grades are posted on Canvas to contact me (or the TA). After a week, all grades are considered final, and no changes can be made.

- **In-class Activities - 5%**

- You participate in in-class activities. For most activities, you will be asked to upload your in-class work (**Jupyter Notebook**) to Canvas **at the end of the class** to receive points based on effort.
- If you miss a class session, you can still work on the code, and submit your Jupyter notebook **within 24 hours** of the class time. You may do this **only up to 3 times**. Any other late submissions won't be credited.

- **Quizzes – 24%**

- There will be **eight quizzes** throughout the term.
- You will take the quizzes **during our Class time** on Canvas.
- Questions will be short-answer, multiple-choice, or true/false, **both coding and conceptual**.
- Quizzes are open book. However, they **are timed**. Therefore, you must be prepared; otherwise, if you want to go back and forth to your notes, you won't be able to finish on time!

- **Midterm Exam - 33% (Week 8, Monday (November 14th) regular class time/room)**

- The midterm is an individual open book exam; and, **no communication of any kind with others is allowed**, either verbal, written, electronic, or internet.
- **The exam is timed**. Therefore, you must be prepared; otherwise, if you want to go back and forth to your notes, you won't be able to finish on time!
- Cell phones must be turned off.
- The Midterm exam will test your data analytics conceptual and Python knowledge: identifying the appropriateness of different techniques for different business scenarios; identifying the strengths and shortcomings of these techniques; interpreting results of analyses in python; coding in Python.
- Midterm questions are similar to quiz questions in terms of form and difficulty.

- **Term Project – 38% (Proposal – 10%, Final Presentation – 10%, Final Report – 18%)**

The term project will pose a relevant business question, gather and clean data, perform data analysis, and report your results in a detailed write-up.

- You are to work in a group of **THREE students**
- You need to specify a business problem and find a relevant dataset. The business context could be in any area, including but not limited to, healthcare, operations, marketing, finance, and social media. If you do not have any business problem in mind, we can meet and define a problem together.
- You need to **present your proposal (in about 5 minutes)** in class during **week 6**. In this proposal, you should clearly define the problem, your objectives, and how you will acquire the data.
- You will be asked to make **a 10-15 minute final presentation** of your project and results during the last week of class (**week 10**).
- The final project report will be due **December 5th, 8:00 am or earlier**. **Absolutely no late submission will be accepted for the final report in any circumstances**, so make sure to have your report ready before the due date.
- The final report will be a formal business report, including the executive summary, introduction, problem description, data exploration, analysis results, conclusions, and recommendations. It should be **5-7 pages** (with tables and graphs).

OBA 510 Students' requirement: The term project for Master's students (enrolled in OBA 510) will go beyond the analogous 410-level project in terms of data analysis expectations. Those higher expectations may also include, depending on the project, more demanding data preparation and cleaning processes and more in-depth and/or comprehensive analyses.

- **Peer-Assessment- Team Project Grading Adjustments:**
 - Should you not contribute to a team assignment, your team is instructed not to include your name on the assignment. If this occurs, you will receive a 0 for that assignment.
 - Furthermore, your overall grade on the team project components of the class (Proposal, Final Presentation & Final Report) is **subject to adjustments either up or down based on peer evaluations** from your teammates as reported in and calculated from the Peer Evaluations form. Adjustments will be made according to the following scale:

Peer Evaluation Grade Adjustments

Raw Peer Evaluation Score	Grade Adjustments *
0-50	-25%
51-55	-20%
56-60	-17.5%
61-65	-15%
66-70	-12.5%
71-75	-10%
76-80	-7.5%
81-85	-5%
86-90	- 2.5%
91-110	NO ADJUSTMENTS MADE
111-120	+ 2.5%
121-130	+ 5%
131-140	+ 7.5%
141 & up	+ 10%

**Grade adjustments will be calculated by applying the percentage adjustment to the total points earned by your group from all of the major team assignments (i.e., Proposal, Final Presentation & Final Report).*

6. Grading Scale

Weighted Grade	Letter Grade
93% and above	A
90% to 92.99%	A-
87% to 89.99%	B+
83% to 86.99%	B
80% to 82.99%	B-
77% to 79.99%	C+
73% to 76.99%	C
70% to 72.99%	C-
67% to 69.99%	D+
63% to 66.99%	D
60% to 62.99%	D-
Below 60%	F

7. Misc. Course Policies

We will adhere to the following policies, the motivations for which should be self-explanatory.

- A missed test will result in a score of zero for that test, so you should be sure to check whether you have a conflict with the announced test dates and times.
- Entering class late (either at the beginning or following a break) is disruptive and reflects negatively on professional interaction.

8. Academic Honesty

The University Student Conduct Code (available at <http://dos.uoregon.edu/conduct>) defines academic misconduct. Students are prohibited from committing or attempting to commit any act that constitutes academic misconduct. By way of example, students should not give or receive (or attempt to give or receive) unauthorized help on assignments or examinations without express permission from the instructor. Students should properly acknowledge and document all sources of information (e.g. quotations, paraphrases, ideas) and use only the sources and resources authorized by the instructor. If there is any question about whether an act constitutes academic misconduct, it is the students' obligation to clarify the question with the instructor before committing or attempting to commit the act. Additional information about a common form of academic misconduct, plagiarism, is available at: <http://researchguides.uoregon.edu/citing-plagiarism/whycite>

Business students are also expected to adhere to the principles and values expressed in the Lundquist Code of Professional Business Conduct. We as a community aspire to be open, respectful, and honest, possessing no tolerance for inappropriate behavior such as cheating, plagiarism, or disorderly conduct.

9. Prohibited Discrimination and Harassment Reporting

Any student who has experienced sexual assault, relationship violence, sex or gender-based bullying, stalking, and/or sexual harassment may seek resources and help at safe.uoregon.edu. To get help by phone, a student can also call either the UO's 24-hour hotline at 541-346-7244 [SAFE], or the non-confidential Title IX Coordinator at 541-346-8136. From the SAFE website, students may also connect to Callisto, a confidential, third-party reporting site that is not a part of the university.

Students experiencing any other form of prohibited discrimination or harassment can find information at respect.uoregon.edu or aaeo.uoregon.edu or contact the non-confidential AAEO office at 541-346-3123 or the Dean of Students Office at 541-346-3216 for help. As UO policy has different reporting requirements based on the nature of the reported harassment or discrimination, additional information about reporting requirements for discrimination or harassment unrelated to sexual assault, relationship violence, sex or gender based bullying, stalking, and/or sexual harassment is available at [Discrimination & Harassment](#).

I am a student-directed employee. Specific details about confidentiality of information and reporting obligations of employees can be found at titleix.uoregon.edu.

10. University of Oregon ADA Policy

The University of Oregon is committed to making available to all its students the opportunity for an excellent and rewarding education. The Americans with Disabilities Act of 1990 and Section 504 of the Rehabilitation Act of 1973 provide federal guidelines which help the University ensure that students with documented disabilities have equal access to this opportunity. If you have a documented disability and anticipate needing accommodations in this course, please make arrangements to meet with me soon. Please request that the Counselor for Students with Disabilities send a letter verifying your disability.