Colin Powell School
for Civic and Global Leadership

The City College of New York

Python for Business Analytics – ECO 32500-ONL (Fall, 2023)
Mon / Wed 2 – 3:15pm (asynchronous), Aug 25 – Dec 20, 2023

Instructor: John Droescher | jdroescher@ccny.cuny.edu | Office Hours: Mon / Wed 5-6 pm

Course Description:
This course teaches the use of SQL and basic Python programming to analyze business and financial datasets. This course will be economics and business based, but the ability to retrieve, clean, and analyze large datasets using SQL queries and Python algorithms is useful beyond data science at technology firms and is beneficial to a career in marketing, business operations, finance, and economics.

Learning Outcomes:
- Develop an understanding of the logic behind programming solutions
- Learn how to write SQL queries to organize and retrieve data.
- Use of fundamental Python programming, Pandas, NumPy, and MatPlotLib to analyze data
- Ability to create and examine various models

Required Books:
Python for Data Analysis: *Data Wrangling with Pandas, NumPy, and IPython*, Author: McKinney, Wes, Publisher: O'Reilly Media, Incorporated, Edition: 2, Year Published: 2018 (ISBN: 978-1-491-95766-0)

Learning SQL: *Generate, Manipulate, and Retrieve Data*, Author: Beaulieu, Alan, Publisher: O'Reilly Media, Incorporated, Edition: 3, Year Published: 2020 (ISBN: 978-1-492-05761-1)

Course Requirements and Grading:
This course has weekly assignments as well as 3 quizzes, an SQL project, and a final project. Attendance is not required but attendance and participation is highly encouraged (for your benefit). Most days will have a short homework assignment. Assignments and quizzes will be submitted through Blackboard while projects will be submitted through GitHub.

Grade Breakdown	
Daily assignments	20%
Quizzes (3x)	10% each
Final Project	50%

Cheating and Plagarism:
Copying other's work and passing it off as your own will not be tolerated. Cheating and/or plagiarism will lead to failing the course and possibly being expelled from City College. If you have any doubts about what plagiarism is, please ask me or visit the following website to familiarize yourself with the rules of academic honesty: https://www.cuny.edu/about/administration/offices/legal-affairs/policies-procedures/academic-integrity-policy

Technology:
This course, of course, requires extensive use of Python and SQL, along with Teams and Github. Python can be installed on your personal computer and accessed using the interface of your choice. For a local install, I recommend VSCode. I also have a server that is accessible remotely that hosts JupyterLab (and can be switched to Jupyter Classic). Additionally, we will be using Microsoft's version of SQL (MSSQL or SQL Server) and the Microsoft SQL Server Management software, VSCode, or Microsoft Azure Data Studio must be installed and usable.

For communication, we will use Teams. Zoom will be used for lectures, although watching posted lectures does not require installation of Zoom software. Submission of assignments will be performed through a Github repository.

Course Outline:

Day 0:

Welcome to the course, homework assignments, project overview, and introduction to the formal analytical process.

Week 1:

SQL Chapters 2 and 3 – Creating and Populating a Database. Query Primer: SELECT, FROM, WHERE, GROUP BY, HAVING, ORDER BY

SQL Chapters 4 and 5 – Filtering and Querying Multiple Tables. An introduction to conditional statements and joins

Introduce Factors and Project – Maven Toys, Texas Deathrow, Weekly Logic Assignments.

Logic Assignment – Airplane and PB&J (in class)

Week 2:

SQL Chapters 6 and 7 – Working with Sets and Data Types. Set theory and operators: union, intersect, except. Working with different types of data: strings, numbers, and time.

SQL Chapters 8 and 9 – Grouping, Aggregates, and Subqueries. Simple and complex grouping. Implicit and explicit aggregation. What is a Subquery and when to use it.

Logic Assignment – Do a load of laundry

Week 3:

SQL Chapters 10 and 11 – Joins and Conditional Logic revisited. Outer, Inner, and Cross joins. Use of case to implement conditional logic.

Python Fundamentals – Installing JupyterLab, Installing Python Libraries, Calling Libraries, and Review of Fundamental Python Structures (Lists, Tuples, Dictionaries, Sets).

Logic Assignment – Set the table for dinner

Week 4:

Python Chapter 3 – Python Data Types, Fundamentals of Python programming. Data structures and sequences: tuples, lists, dictionaries, and sets. List comprehensions. Basics of functions.

Logic Assignment – Make a phone call

Week 5:

Python Chapter 4 – NumPy Basics: Arrays and Linear Algebra. The array as a multidimensional object. Creating arrays, indexing and slicing, universal functions, array-oriented programming.

Writing Good Code – There is more to writing code than just whether it works.

Logic Assignment – Go to Google.com

Week 6:

Python Chapter 5 – Getting Started with Pandas. Introduction to Pandas data structures, essential functionality, and descriptive statistics.

Logic Assignment – Put on a coat

Week 7:

Python Chapter 6 – Data Loading, Storage, and File Formats. The basics of Python Input/Output operations. Web scraping. Interacting with Databases and Web APIs.

Week 8:

Python Chapter 7 – Data Cleaning and Preparation. Handling missing data, removing duplicates, detecting and filtering outliers, binning, and random sampling.

Logic Assignment – Load a dishwasher

Week 9:

Python Chapter 8 – Data Wrangling: Join, Combine, and Reshape. Hierarchical indexing, summary statistics by level, and indexing with dataframe columns. Combining and merging datasets. Reshaping and pivoting dataframes and tables.
Logic Assignment – Order a PB&J from your local bodega

Week 10:

Python Chapter 9 – Plotting and visualization. A brief MatPlotLib primer. Figures and subplots. Colors, markers, and line styles. Ticks, labels, and legends. Plotting with Pandas and Seaborn.
Logic Assignment – Mail a letter

Week 11:

Python Chapter 10 – Data Aggregation and Group Operations. Groupby mechanics, iterating over groups, data aggregation, split-apply-combine, quantile and bucket analysis.
Logic Assignment – Feed the pet

Week 12:

Python Chapter 11 – Time Series. Date and time data types and tools. Converting between string and datetime. Indexing, Selection, Slicing. Date ranges. Handling time zones. Periods and period based arithmetic. Resampling and frequency conversion.
Logic Assignment – Peel a banana

Week 13:

Efficient Frontier Discussion
Python Chapter 12 – Advanced Pandas. Categorical data. Advanced Groupby use. Techniques for method chaining.
Logic Assignment – Brush your teeth

Week 14:

Python Chapter 13 – Introduction to Modeling Libraries in Python. Estimating linear models and time series processes. Introduction to SciKit-Learn
Logic Assignment – Take out the trash

Week 15:

Python Chapter 14 – Real world data analysis examples. Maven Toys analysis.
Logic Assignment – Watch *Spider-Man: Into the Spider-Verse* on Netflix

Week 16:

Semester Wrap – Final Project Presentations

Oct 2: Quiz 1 Due: SQL basics
Nov 1: Quiz 2 Due: Python Fundamentals, Pandas, and Data Cleaning
Nov 29: Final Project Draft
Nov 29: Quiz 3 Due: More Advanced Python methods
Dec 14: Final Project Due