DATA 690 | Statistical Analysis & Visualization with Python (Spring 2023)

Instructor: **Sunela Thomas**
Email: **sunelat1@umbc.edu**

Course Description

This course aims to provide –
  i.   an introduction to Python programming
  ii. the fundamental statistical concepts and methods, which are frequently used in exploratory data analysis
  iii. how to carry out these analyses and visualize the findings in Python and Tableau

Prerequisites: Enrollment in the Data Science Program

Recommended References:

- Python for Data Analysis by Wes McKinney
- Python Programming: An Introduction to Computer Science by John Zelle
- Programming in Python 3: A Complete Introduction to the Python Language by M. Summerfield
- Allen B. Downey, Think Stats, 2nd edition, O'Reilly
- Statistics for Business & Economics by Anderson et al.
- Python package* user guides (i.e., NumPy User Guide, Pandas Reference Guide, etc.)

Python packages
  - Analysis: NumPy and SciPy
  - Handling Data: Pandas
  - Plotting: Matplotlib and Seaborn

Software

The course will be using Python 3 with the following libraries: NumPy, Pandas, Matplotlib, Seaborn, and Stats. It is the student's responsibility to have a working environment. If you'd like to have the environment installed locally, Anaconda is a Python distribution that has all required libraries. You can use Jupyter Notebook through Anaconda or use Google's Colab which is available from your UMBC account. **All code in class lecture will be done in a Jupyter Notebook.**

Course Content (tentative)

Week	Date	Description
1	1/30	Orientation, Introduction to Python and Notebooks, basic data types
2	2/6	Python Statements, Loops, and Functions
3	2/13	Python: IO, Files, and Numerical Calculations with NumPy
4	2/20	Python: Basic plotting with Matplotlib and data handling with Pandas
5	2/27	Python: Advanced Plotting and Time Series
6	3/6	Tableau – Part 1
7	3/13	Tableau – Part 2
8	3/20	SPRING BREAK
9	3/27	Tableau – Part 3 or Power BI Intro or Neo4j Intro
10	4/3	Introduction to Statistics & Descriptive Statistics-1
11	4/10	Descriptive Statistics-2
12	4/17	Introduction to Probability
13	4/24	Discrete and Continuous Probability Distributions
14	5/1	Sampling, Distribution, and Interval Estimation
15	5/8	Hypothesis Testing and Inference
16	5/15	Final Project Due

Course Requirements

Successful students will: (a) actively participate during class and engage substantially in the breakout activities; and (b) complete the individual assignments and project. Attendance in class is required and students are expected to turn on their camera during the class.

Course Format and Assignments

Students will respond to questions in class, participate in breakout sessions, complete homework/assignments and a final project. The assignments will give students an opportunity to gain practical insights with Python programming, explanatory statistical analysis, and

visualization/Tableau. The project will give students the opportunity to practice various topics studied in class.

Grading Criteria

Class/Break Out Participation	5%
"Live" Class Attendance	5%
Homework/Assignments	65%
Final Project	25%

Final Grade is computed as:

90% - 100%	A
85% - 89.99%	A-
80% - 84.99%	B+
75% - 79.99%	B
70% - 74.99%	B-
65% - 69.99%	C+
60% - 64.99%	C
50% - 59.99%	D
49.99% and below	F

Course Policies

UMBC provides a range of writing assistance, which can be found in the following:
- The Writing Center https://lrc.umbc.edu/tutor/writing-center/
- Research Guides & Tutorials https://lib.guides.umbc.edu/tutorial

Failure to follow guidelines for each assignment, including the required format, style, length, submission date, etc., may result in at least one-letter-grade reduction on the paper depending on the type and/or number of transgressions.

Late/Incomplete assignments will not be accepted unless an extension has been agreed to in advance. Emergency situations will be handled on a case-by-case basis with appropriate justification and/or documentation. Incomplete grades will not be entertained unless extenuating circumstances warrant, and your request is made before the last week of class. This course does

have any **extra** grade assignments or activities. Assignments are expected to be completed by the due date. Late assignments will receive a grade of "**zero**".

Academic Integrity

By enrolling in this course, each student assumes the responsibilities of an active participant in UMBC's scholarly community in which everyone's academic work and behavior are held to the highest standards of honesty. Cheating, fabrication, plagiarism, and helping others to commit these acts are all forms of academic dishonesty, and they are wrong. Academic misconduct could result in disciplinary action that may include, but is not limited to, failure, suspension, or dismissal.

The use of previously released materials, including Discussion posts, assignments, and papers will be considered a violation of academic integrity. Failure to adequately cite or quote direct language is also subject to a penalty.

Refer to the UMBC policy at:
https://catalog.umbc.edu/content.php?catoid=14&navoid=718#academic-integrity