

**Spring 2024**

**Instructor: Sergio Luna, Ph.D.**

Course Web Address: [Link to Blackboard shell](#)

Course Schedule: Thursday (6:00 pm – 8:50 pm MST)

Contact Information: [salunafong@utep.edu](mailto:salunafong@utep.edu)

Live session: [Link to live session](#)

Office Hours: Thursday 3:00 pm - 4:30 p.m. MST or by appointment

Teaching Assistant Hours: Mr. Jesus Alejandro Gutierrez Araiza

Email: [jgutierrez42@miners.utep.edu](mailto:jgutierrez42@miners.utep.edu)

#### COURSE DESCRIPTION

This course provides a formal introduction to the process of data manipulation, analysis, and visualization to drive the strategic decision-making process. It leverages such methods with the corresponding problem-solving and story-telling skills to complement data visualization. Although the course is based mainly on Python, the concepts and techniques presented are general enough to extend to other statistical or data analysis tools. The course covers basic plotting techniques, including bar and scatter plots, time series, heatmaps, geographic maps, and network graphs, among other visualization charts. Additionally, students will get hands-on experience with basic data pre-processing techniques, including summaries, data cleaning and transformation, introduction to interactive visualizations, and fundamentals of machine learning.

No (proficient) programming experience is required.

#### STUDENT LEARNING OUTCOMES

After successful completion of this course, students will be able to:

- **Understand** the conventional approach to manipulating and pre-processing data while using large datasets. Transformed data will then be analyzed to generate insights and provide stakeholders with data-driven recommendations.
- **Implement** data manipulation techniques and create high-quality plots to understand the phenomena of interest.
  - The student will develop critical thinking skills by addressing 5W questions when creating visualizations. What and So-What will be further explored.
- **Generate** data-driven insights by loading, cleaning, transforming, and visualizing datasets using Python.
- **Formulate** a data analytics project manuscript and presentation describing clear project objectives, explaining the data analytics process implemented, and summarizing the study into critical insights and recommendations for stakeholders.

## COURSE FORMAT AND STRUCTURE

- Our weeks will run from Wednesday to Tuesday. I will post information (online activities, discussion starters, etc.) in Blackboard for the upcoming week by Tuesday evening so you can begin the new week when you log in on Wednesday.
    - Course Web Address:
  - Virtual office hours will be held Thursday from 3:00 p.m. - 4:30 p.m. MST or by appointment. Please email me at [salunafong@utep.edu](mailto:salunafong@utep.edu) to schedule a meeting.
  - Assignments are due by 11:59 p.m. MST on the due date listed in the course schedule. **Late submissions will be accepted up to 24 hours after the due date. However, 50% of the maximum will be deducted from the assignment score. No assignments will be accepted after 24 hours from the due date.**
  - An assignment file should be appended with the respective assignment number, your Last Name, and your First Name, such as "assignment1\_LastName\_FirstName.pdf". This makes it easier to manage assignment files when downloading to my computer.
- **Please submit the following files:**
    - Working **Python file, script.** ("assignment1\_LastName\_FirstName.py") - # Commented
    - A **PDF file** describing the process and steps taken while manipulating data
    - The **dataset (raw data)** used to run the Python-submitted file
    - **NOTE: Points will be deducted if any of the above items are missing from your submission because we will not be able to grade your submission completely**

## TENTATIVE COURSE SCHEDULE

The schedule below is subject to change. If, for any reason, I am required to make any amendments, I will inform you via Blackboard, email, and MS Teams.

Assignments are due by 11:59 p.m. MST on the day specified in the “Assignment” column seen below.

### Tentative Schedule

Module	Date	Topic(s)	Readings	Assignment
Week 01	Jan 16 <sup>th</sup> – 22 <sup>nd</sup> , 2024	Orientation & Introduction and Course Overview  From Data Visualization to Data Analytics	Required: <ul style="list-style-type: none"> <li>Data visualization literacy: Definitions, conceptual frameworks, exercises.</li> <li>Principles of Effective Data Visualization</li> <li>Top considerations for effective visualizations</li> </ul>	- HW: Data Visualization in Your Life - Quiz
Week 02	Jan 23 <sup>rd</sup> – 29 <sup>th</sup>	Recap of Python	Required: <ul style="list-style-type: none"> <li>Data Structures</li> <li>Data Structures in Python</li> </ul>	- HW Data Structures - Quiz
Week 03	Jan 30 <sup>th</sup> – Feb 5 <sup>th</sup>	Data Manipulation 1	Required: <ul style="list-style-type: none"> <li>The Data Life Cycle</li> <li>8 Steps in the data life cycle</li> <li>CRISP-DM for Data Science: Strengths, Weaknesses, and Potential Next Steps</li> </ul> Optional: <ul style="list-style-type: none"> <li>Pandas 10 minutes</li> </ul>	- HW: Data Manipulation 1 - Quiz
Week 04	Feb 6 <sup>th</sup> – 12 <sup>th</sup>	Data Manipulation 2	Required: <ul style="list-style-type: none"> <li>Long vs Wide Data</li> </ul>	- HW: Data Manipulation 2 - Quiz
Week 05	Feb 13 <sup>th</sup> – 19 <sup>th</sup>	Developing Basic Charts	Required: <ul style="list-style-type: none"> <li>Midterm Project Instructions</li> <li>Matplotlib examples</li> </ul> Optional: <ul style="list-style-type: none"> <li>Matplotlib package documentation</li> </ul>	- HW: Basic Graphs - Quiz
Week 06	Feb 20 <sup>th</sup> – 26 <sup>th</sup>	Enhancing Basic Graphs	Required: <ul style="list-style-type: none"> <li>Midterm Project Instructions</li> <li>Seaborn examples</li> </ul> Optional: <ul style="list-style-type: none"> <li>Seaborn package documentation</li> </ul>	- HW: Enhancing Basic Graphs - Quiz
<b>Week 07</b>	Feb 27 <sup>th</sup> – March 4 <sup>th</sup>	<b>Midterm Project - Storytelling</b>	<b>Required:</b> <ul style="list-style-type: none"> <li><b>Midterm Project Instructions</b></li> </ul>	<b>Midterm project</b>

Week 08	March 5 <sup>th</sup> – 11 <sup>th</sup>	Spatial Data and Time Series	Required: <ul style="list-style-type: none"> <li>• Midterm Project Instructions</li> <li>• Seaborn examples</li> </ul>	- HW: Spatial Data - Quiz
<b>Week 09</b>	<b>March 12<sup>th</sup> – 18<sup>th</sup></b>	Interactive Visualizations & Heatmaps	Required: <ul style="list-style-type: none"> <li>• These Heatmaps Show How Retailers Shop</li> </ul> Optional: <ul style="list-style-type: none"> <li>• Seaborn and heatmaps</li> </ul>	- HW: Heatmaps and Treemaps - Quiz
Week 10	March 19 <sup>th</sup> – 25 <sup>th</sup>	Interactive Visualizations: Introduction to Dashboards	Required: <ul style="list-style-type: none"> <li>• Five key properties of interactive data visualization</li> <li>• 3 Ways Data Dashboards Can Mislead You</li> <li>• Build a Better Dashboard for Your Agile Project</li> </ul> Optional: <ul style="list-style-type: none"> <li>• Bokeh package capabilities</li> </ul>	- HW: Introduction to Dashboards
Week 11	March 26 <sup>th</sup> – April 1 <sup>st</sup>	Text Analysis Visualization	Required: <ul style="list-style-type: none"> <li>• Review Final Project Instructions</li> <li>• Text Analytics 101: The C-Suite Guide to Mining The Voice of The Customer</li> <li>• Natural Language Processing (NLP)</li> <li>• Actionable Strategies in Real-Time Marketing Using Sentiment Analysis</li> </ul>	- HW: Text Analysis
Week 12	April 2 <sup>nd</sup> – 8 <sup>th</sup>	Network Graphs	Required: <ul style="list-style-type: none"> <li>• An introduction to Graph, The Essential Data Analysis Tool</li> </ul>	- HW: Network Graphs - Quiz
Week 13	April 9 <sup>th</sup> – 15 <sup>th</sup>	Off-the-shelve visualization platforms	Required <ul style="list-style-type: none"> <li>• Final project Instructions</li> </ul>	Focus on Final Project No deliverable for this week
Week 14	April 16 <sup>th</sup> – 22 <sup>nd</sup>	CRISP-DM Presentation	Required <ul style="list-style-type: none"> <li>• Poster template</li> </ul>	
<b>Week 15</b>	<b>April 28<sup>th</sup></b>	<b>Final Project</b>		<b>Final Project due: 4/28</b>

## COURSE MATERIALS

**No textbook is required.** Students will be provided with a compilation of materials (slide decks, reference readings, sample scripts, pre-recorded lectures, and tool tutorials) needed to complete the course successfully.

**OPTIONAL reading material includes** Yau, N. (2013). *Data points: visualization that means something*. John Wiley & Sons. Link to book: [Data Points: Visualization that means something 1st edition Course Requirements](#)

- **Attendance:** Attendance is not mandatory but recommended.

As you move through the topics in each module, you will be asked to read, listen to, and watch various media rather than simply viewing lecture content. You will also be regularly prompted to evaluate your knowledge as you build it actively.

The components in each module are designed to be completed sequentially in order. In addition to videos, readings, and interactives, we want to bring your attention to a few learning activities you'll encounter.

- **Live Sessions:** Live sessions are scheduled in each module. You will work with your classmates in some sessions to solve a problem. In others, you will work through hands-on virtual labs to help you complete your lab assignments.
- **Homeworks:** To help reinforce the material covered in the module, homework exercises will be assigned each week, which will involve formulating and solving a small but practically relevant homework problem.
- **Quizzes:** These graded quizzes evaluate your mastery of concepts shortly after they are introduced. You are expected to complete these checks (and your doing so will contribute to your overall participation grade).
- **Midterm Project:** The midterm project reviews foundational concepts in data visualization, manipulation, and basic graphs. Students conduct a study where data needs to be cleaned, transformed, summarized, and visualized in basic charts. In addition, students implement concepts of data lifecycle phases to create a storytelling report.
  - The midterm project consists of developing a report summarizing a dataset through charts. A description of the work is as follows:
    - 1 Project description
    - Provide context of the assigned dataset
    - Discuss the objectives, scope, and data source for the analysis to be conducted
    - Describe the steps undertaken to manipulate the dataset effectively
    - Summarize Results - Graphs supporting your statements
    - Insights - What are the visualizations telling you? Remember to implement the storytelling approach
- **Final Project:** The final project captures the knowledge and skills acquired during the semester. It evaluates students' approaches when developing and communicating data-driven insights through visualizations.
  - The final project consists of developing the following:
    - One research poster (data story must be coherent and stay with only one topic). All graphs must be related. A minimum of 4 charts must be developed.
    - 1 Recorded presentation of your poster – 10 -15min presentation
    - 1 PDF document with lifecycle phases
    - All CSV files were used as part of the analysis.
- There are some instances where subsets are extracted, or the provided data is joined with an external file you found interesting. Please attach ALL files used as part of your final project; I must be able to replicate your analysis (run your code) using the provided file. Please do so to avoid losing points for replicability.

## GRADING PROCEDURES

Grades will be based on the following weights:

Assignment	Percentage of Final Grade	Due Date
Homework assignments	35%	Week 1-6, Week 8-12
Mid-term project	25%	Week 7
Final project	30%	Week 15
Quizzes	10%	Week 1-6, Week 8-10, Week 12

The final grading rubric will be as follows:

A	90 - 100
B	80-89
C	70 - 79
D	60 – 69
F	< 60

## TECHNOLOGY REQUIREMENTS

Required Software:

- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint
- Python, 3.12.1 [Download the Python website](#)
- Python IDLE
  - PyCharm <https://www.jetbrains.com/pycharm/>

## LEARNING ACCOMMODATIONS

The Center for Accommodations and Support Services (CASS) aspires to provide students with disabilities with accommodations and support services to help them pursue their academic, graduation, and career goals. For more information concerning services for students with disabilities, please get in touch with the Center for Accommodations and Support Services at <https://www.utep.edu/student-affairs/cass/>

## INCLUSIVITY

### **Name and Pronoun Usage**

As this course includes group work and class discussion, we must create an educational environment of inclusion and mutual respect. This consists of the ability for all students to have their chosen gender pronoun(s) and chosen name affirmed. Please inform the instructor of the necessary changes if the class roster does not match your name and pronouns.

You are expected to treat your instructor and all other participants in the course with courtesy and respect. Disrespectful conduct and harassing statements will not be tolerated and may result in disciplinary actions.