

Fall 2023

Instructor: Sergio Luna, Ph.D.

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

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

Contact Information: salunafong@utep.edu

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

Office Hours: Thursday 5:00pm - 8:00 p.m. MST or by appointment

Office session URL: [Link to Zoom session](#)

Teaching Assistant Hours: Mr. Ian Torres

Email: iatorres5@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 mostly on Python, the presented concepts and techniques 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 types of 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-process data while manipulating 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 developing 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 key 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: [Link to Blackboard shell](#)
- Virtual office hours will be held Wednesday from 3:00 p.m. - 4:30 p.m. MST or by appointment. Please email me at 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 First Name, such as "assignment1_LastName_FirstName.pdf". This makes it easier for me 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	Aug 30 – Sep 5 th	Orientation & Introduction and Course Overview From Data Visualization to Data Analytics	Required: <ul style="list-style-type: none"> Data visualization literacy: Definitions, conceptual frameworks, exercises. Principles of Effective Data Visualization Top considerations for effective visualizations 	- HW: Data Visualization in Your Life - Quiz
Week 02	Sep 4 th – Sep 10 th	Recap of Python	Required: <ul style="list-style-type: none"> Data Structures Data Structures in Python 	- HW Data Structures - Quiz
Week 03	Sep 6 th – Sept 12 th	Data Manipulation 1	Required: <ul style="list-style-type: none"> The Data Life Cycle 8 Steps in the data life cycle CRISP-DM for Data Science: Strengths, Weaknesses, and Potential Next Steps Optional: <ul style="list-style-type: none"> Pandas 10 minutes 	- HW: Data Manipulation 1 - Quiz
Week 04	Sep 13 th – 19 th	Data Manipulation 2	Required: <ul style="list-style-type: none"> Long vs Wide Data 	- HW: Data Manipulation 2 - Quiz
Week 05	Sep 20 th – 26 th	Developing Basic Charts	Required: <ul style="list-style-type: none"> Midterm Project Instructions Matplotlib examples Optional: <ul style="list-style-type: none"> Matplotlib package documentation 	- HW: Basic Graphs - Quiz
Week 06	Sept 27 th – Oct 3 rd	Enhancing Basic Graphs	Required: <ul style="list-style-type: none"> Midterm Project Instructions Seaborn examples Optional: <ul style="list-style-type: none"> Seaborn package documentation 	- HW: Enhancing Basic Graphs - Quiz
Week 07	Oct 4th – Oct 10th	Midterm Project - Storytelling	Required: <ul style="list-style-type: none"> Midterm Project Instructions 	Midterm project

Week 08	Oct 11 th – 17 th	Spatial Data and Time Series	Required: <ul style="list-style-type: none"> Midterm Project Instructions Seaborn examples 	- HW: Spatial Data - Quiz
Week 09	Oct 18th – 24th	Interactive Visualizations & Heatmaps	Required: <ul style="list-style-type: none"> These Heatmaps Show How Retailers Shop Optional: <ul style="list-style-type: none"> Seaborn and heatmaps 	- HW: Heatmaps and Treemaps - Quiz
Week 10	Oct 25 th – Oct 31 st	Interactive Visualizations: Introduction to Dashboards	Required: <ul style="list-style-type: none"> Five key properties of interactive data visualization 3 Ways Data Dashboards Can Mislead You Build a Better Dashboard for Your Agile Project Optional: <ul style="list-style-type: none"> Bokeh package capabilities 	- HW: Introduction to Dashboards
Week 11	Nov 1 st – Nov 7 th	Text Analysis Visualization	Required: <ul style="list-style-type: none"> Review Final Project Instructions Text Analytics 101: The C-Suite Guide to Mining The Voice of The Customer Natural Language Processing (NLP) Actionable Strategies in Real-Time Marketing Using Sentiment Analysis 	- HW: Text Analysis
Week 12	Nov 8 th – 14 th	Network Graphs	Required: <ul style="list-style-type: none"> An introduction to Graph, The Essential Data Analysis Tool 	- HW: Network Graphs - Quiz
Week 13	Nov 15 th – 21 st	Off-the-shelf visualization platforms	Required <ul style="list-style-type: none"> Final project Instructions 	Focus on Final Project No deliverable for this week
Week 14	Nov 22 nd – Nov 28 th	CRISP-DM Presentation	Required <ul style="list-style-type: none"> Poster template 	
Week 15	Nov 29th – Dec 5th	Final Project		Final Project due: 12/9

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) required 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, rather than simply viewing lecture content, you will be asked to read, listen to, and watch a variety of media. 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. In some sessions, you will work collaboratively with your classmates 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, data 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 it 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 conducted 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:
 - 1 research poster (data story, must be coherent and stay with only 1 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 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. Failure to do so will result in 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 followed:

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, 3.8.12 [Download 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, accommodations and support services to help them pursue their academic, graduation, and career goals. For more information concerning services for students with disabilities, please contact 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, it is vitally important for us to create an educational environment of inclusion and mutual respect. This includes the ability for all students to have their chosen gender pronoun(s) and chosen name affirmed. If the class roster does not align with your name and/or pronouns, please inform the instructor of the necessary changes.

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