



IE 5390
Special Topics Industrial Engineering
“Data Visualization for Decision Making”

College of Engineering

Spring 2020

Instructor: Dr. Sergio Luna

Course Web Address:

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

Contact Info:

Virtual Office Hours: Mondays 6:00 - 7:30 p.m. MST or by appointment

Virtual session URL: <https://stevens.zoom.us/j/952179803>

Prerequisite(s): Department Approval, [CR] Level--UG, GR, DR, [CR] Campus Location, CAPP (MECH Eng Elective)

Cross-listed with: None

COURSE DESCRIPTION

This course provides a formal introduction to data manipulation, analysis, and visualization techniques. It leverages such techniques with the corresponding problem-solving skills necessary to complement data visualization into specific strategic decision-making. Although the course is based mostly on R statistical programming language, the techniques presented are general enough that can be used with other statistical/data analysis tools. The course covers basic plotting techniques such as scatter plots, bar plots, line charts, heatmaps, geographic maps, among others. Additionally, students will get hands-on-experience with basic data pre-processing techniques including summaries and data cleaning, introduction to interactive visualization, and basic machine learning skills.

No (proficient) programming experience is required, basic working knowledge of R and RStudio is strongly recommended.

STUDENT LEARNING OUTCOMES

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

- Understand and implement an analytical framework for conducting data exploration and analysis. In addition, the student will be able to apply 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.
- Understand the conventional approach to manipulate and pre-process data before making any inferences.
 - The students will be able to load, manipulate, transform, and clean data using R and RStudio.
- Use R statistical language to generate data visualizations.
 - The student will be able to create different plots (Time series, Bubble graphs, Histograms, etc.) by creating R scripts.
- Develop project planning, public speaking and project management experience by creating a multi-chart static poster or interactive dashboard while disseminating the key insights to an audience.
 - Students will develop 2 major projects that will need be managed in terms of data extraction, research scope, analysis, insights generated, insights presentation and suggested outcomes. Text documents and oral presentation skills are being practiced and evaluated.

COURSE FORMAT AND STRUCTURE

This course is offered in-person; however the materials will be available in the class module at the university's respective platform. To access the course, Blackboard

Course Logistics

- Virtual office hours will be held Monday evenings from 8:00 - 9:30 p.m. EST or by appointment. Please email me at sluna@stevens.edu to schedule a meeting.
- Assignments are due by 6:00 p.m. MST on the due date listed in the course schedule. Late submissions will be accepted, however for every day (every 24 hours) the assignment is late after due date, 20% of the maximum will be deducted from the assignment score. No assignments will be accepted once they are five or more days late.
- An assignment file should be appended by your Last Name and First Name, such as "assignment1_LastName_FirstName.pdf". This may make it easier for me to manage assignment files you download to my computer.
- Please submit your working R file ("assignment1_LastName_FirstName.R") as well as a PDF file with your conclusions.

TENTATIVE COURSE SCHEDULE

The schedule below is subject to change. If for any reason I am required to make any amendments, I will be informing you via Canvas and email.

Assignments are due by 6:00 p.m. MST on the day specified in the "Assignment" column seen below.

Tentative Course Schedule

Module	Date	Topic(s)	Assignment
Week 01	Jan 21 st – 26 th , 2020	Orientation & Introduction and Course Overview	
Week 02	Jan 27 th – Feb 2 nd , 2020	From Data Visualization to Data Analytics	
Week 03	Feb 3 rd – 9 th , 2020	Introduction to R	
Week 04	Feb 10 th – 16 th , 2020	Data Manipulation	HW #1 due 2/18 Data Manipulation
Week 05	Feb 17 th – 23 rd , 2020	Basic Graphs	
Week 06	Feb 24 th – March 1 st , 2020	Graphs with Enhancement	HW #2 due 3/8 Basic graphs & enhancement
Week 07	March 2 nd – 8 th , 2020	Heat Map	
Week 08	March 9 th – 15 th , 2020	Spatial Data	HW #3 due 3/17 Heatmap & Spatial Data Release Mid-term exam
Week 09	March 16th – 22nd, 2020	Spring Break	Spring Break
Week 10	March 23 rd – 29 th , 2020	Categorical Data	
Week 11	March 30 th – April 5 th , 2020	Time Series	Mid-term due on 3/29 at 11:59pm
Week 12	April 6 th – 12 th , 2020	Multivariate Data	
Week 13	April 13 th – 19 th , 2020	Interactive Visualization	Define Final Project
Week 14	April 20 th – 26 th , 2020	Network Graphs	HW #4 due 4/14 Categorical – Time Series & Multivariate
Week 15	April 27 th – May 3 rd , 2020	Text Analysis	HW #5 due 4/28
Week 16		Final Project	Final Project due 5/6 Presentations Submit deliverables on 5/3 before 11:59pm.

COURSE MATERIALS

Textbook(s): No textbook is required for this course. Students will be presented with slides, and online tutorials. In case students are interested, the following books are recommended:

[R Graphics Cookbook Practical Recipes for Visualizing Data](#)
(Available also in an online version)

[Storytelling with Data: A Data Visualization Guide for Business Professionals](#)

COURSE REQUIREMENTS

- **Attendance:** Attendance is NOT mandatory, but recommended.
- **Assignments:** A total of 6 homework assignments will be assigned throughout the semester.

Assignments are due by 6:00 p.m. MST on the due date listed in the course schedule. Late submissions will be accepted, however for every day the assignment is late after due date, 20% of the maximum will be deducted from the assignment score. No assignments will be accepted once they are four or more days late.

An assignment file should be appended by your Last Name and First Name, such as "assignment1_LastName_FirstName.pdf". This may make it easier for me to manage assignment files you download to my computer

Please submit your working R file (filename.R) along with a PDF file with your conclusions.

- **Exams:** 1 Mid-term exam will be assigned.

Students will work individually with assigned project and deliver a "data story" report.

Grading rubric:

- Context: Clearly defined topic that answers a specific question.
- Data Cleanliness: The extent to which the data has been manipulated to remove potential bias injection to the resulting visualizations. Identifying incomplete, incorrect, inaccurate or irrelevant parts of the data.
- Importance: the extent to which the visualization addresses problems facilitates decision making.
- Relevancy: Visualization contains no color, symbolism, or text that is irrelevant to the question the visualization seeks to answer.
- Grammar of graphics: Applies appropriate graphic variable types for the data type and scale.
- Aesthetic Design: Meticulous care given to colors, shape, size, background, annotation and overall design.
- Key insights communication. Clearly communicate the purpose, objective, analysis and insights and recommendations to the audience.

- **Project(s):** 1 Final Project will be assigned.

Students will work individually on their chosen project and create a poster and "data story" report.

The final project aims to capture the knowledge and skills learned throughout the semester. It evaluates the students' approach when creating visualizations that support the decision-making process.

Grading rubric:

- *All of the criteria* evaluated at during the mid-term exam, plus:
- Evidence: Data comes from multiple highly reputable sources; data is supplemented with references to relevant scholarship; methodology behind how the data was collected is explained.
- Layout: design multi-chart static poster or interactive dashboard.
- Key insights communication. Clearly communicate the purpose, objective, analysis and insights and recommendations to the audience.

GRADING PROCEDURES

Grades will be based on:

Homework assignments	35%
Mid-term project	30%
Final Project	35%
Mini-Challenges	Extra Credit Professor Discretion