

C S 48A: DATA VISUALIZATION

Foothill College Course Outline of Record

| | |
|------------------------------------|---|
| Effective Term: | Summer 2024 |
| Units: | 4.5 |
| Hours: | 4 lecture, 2 laboratory per week (72 total per quarter) |
| Advisory: | MATH 10, MATH 17, PSYC 7 or SOC 7. |
| Degree & Credit Status: | Degree-Applicable Credit Course |
| Foothill GE: | Non-GE |
| Transferable: | CSU/UC |
| Grade Type: | Letter Grade (Request for Pass/No Pass) |
| Repeatability: | Not Repeatable |

Student Learning Outcomes

- Students will be able to communicate analysis insight through data visualization to an intended audience.
- The student will be able to Identify the key techniques and theory used in data visualization
- The student will be able to create and design static and interactive visualizations using a data visualization package

Description

Introduction to the effective processing and communication of data. Topics include identifying the key techniques and theory used in data visualization, creating and designing static and interactive visualizations using data, and communicating insight through data visualization to an

intended audience. Students will use a data visualization package, such as R, Tableau, or Matplotlib in Python.

Course Objectives

The student will be able to:

1. Identify the key techniques and theory used in data visualization.
2. Create and design static and interactive visualizations using a data visualization package.
3. Communicate analysis insight through data visualization to an intended audience.

Course Content

1. Identify the key techniques and theory used in data visualization
 - a. Introduction to data visualization
 - i. Describe data visualization
 - ii. Compare descriptions of exploratory vs. explanatory data visualizations
 - iii. Identify Gestalt Principles in visual perception
 - iv. Describe cognitive load in visual perception to identify cognitive load best techniques
 - v. Compare static and dynamic visualizations represented as use cases
 - vi. Describe effective design in visual perception to identify effective design best techniques
 - vii. Describe ethical design in visual perception to identify ethical design best techniques
 - viii. Identify most common graphical representations and their key characteristics
 - ix. Recognize the context of use for each graphical representation and what approaches are appropriate for each situation
 - x. Identify context or setting to properly story-tell with data visualization best practices
 - xi. Identify plot to properly story-tell using data visualization best practices
 - xii. Identify message/interpretation to properly story-tell using data visualization best practices
 - xiii. Identify purpose or implications to properly story-tell using data visualization best practices

- b. Introduction to a visualization package, such as R, Tableau, or Matplotlib
 - i. Describe the purpose and applications of the package
 - ii. Identify and navigate the package environment
 - iii. Import from an external data source
 - iv. Develop a basic data object
 - v. Apply data visualization customization, such as adding colors, labels, and formatting
 - vi. Develop a variety of data visualizations, including histograms, bar graphs, box plots, pie charts, scatterplots, etc.
- 2. Create and design static and interactive visualizations using a data visualization package
 - a. Working with data
 - i. Prepare data for use in visualization
 - ii. Merge multiple data sources
 - iii. Apply data cleaning techniques
 - i. Outliers
 - ii. Bad data
 - iii. Missing data
 - iv. Identify distinct data types in imported data
 - v. Identify and use different types of data sources:
 - i. CSV
 - ii. Excel
 - iii. SQL
 - iv. URL
 - v. Other
 - vi. Develop filtering techniques
 - vii. Develop table calculation techniques
 - b. Interactive design
 - i. Apply interactivity of plots with highlight, filter, and URL actions
 - ii. Identify dynamic data visualization best practices
 - iii. Develop dynamic data presentations
 - iv. Identify and develop mapped visualizations

- v. Apply custom geocoding to map visualization
 - vi. Develop custom maps
 - vii. Apply background images to map visualizations
- 3. Communicate analysis insight through data visualization to an intended audience
 - a. Visual analytics
 - i. Develop statistical visualizations used in exploratory data analysis
 - ii. Develop trends analytics techniques
 - iii. Develop distributions analytics techniques
 - iv. Develop forecasting analytics techniques
 - v. Apply mapping techniques
 - b. Sharing visualizations
 - i. Identify the value of data visualization in communicating to an intended audience
 - ii. Select most appropriate data visualization for intended outcome
 - iii. Apply design principles, such as eliminating clutter, focusing attention, and thinking like a designer to improve data visualizations
 - iv. Compare the pros and cons of sharing live data vs. extracting data

Lab Content

- 1. Navigating the visualization package
 - a. Installing and running
 - b. Creating a simple plot
- 2. Demonstrating visual design principles
 - a. Identify effective and ineffective use of design principles in sample products
 - b. Modify sample products to improve effective communication
- 3. Explanatory data analysis and visualization
 - a. Create plots from existing datasets
 - b. Choose from among various representations (scatterplot, histogram, boxplot, etc.) to communicate the data most effectively
 - c. Apply labels, colors, and formatting

- d. Apply statistical techniques to summarize the data
- 4. External data sources
 - a. Find and download data from a repository
 - b. Scrape data from webpages
 - c. Prepare external data for import
 - i. Identify and handle missing data
 - ii. Identity and handle bad data
 - iii. Identify and handle outliers
 - d. Import data into the visualization package
 - e. Merge data from multiple sources
- 5. Exploratory data analysis
 - a. Using complex or large datasets, identify correlations between data elements
 - b. Using complex or large datasets, identify trends in time-series data
 - c. Using complex or large datasets, produce and analyze predictive models
 - d. Choose from among various representations (time series, facets, heat map, tree map, etc.) to communicate the data most effectively
- 6. Geospatial visualization
 - a. Summarize geotagged data in the context of a map
 - b. Create a custom geospatial and data-rich object, such as a stadium or parking lot
- 7. Dynamic visualization
 - a. Create animated visualizations
 - b. Create interactive data applications or webpages

Special Facilities and/or Equipment

1. The college will provide access to a computer laboratory with an installed data visualization package.
2. The college will provide a website or course management system with an assignment posting component (through which all lab assignments are to be submitted) and a forum component (where students can discuss course material and receive help from the instructor). This applies to all sections, including on-campus (i.e., face-to-face) offerings.

3. When taught online, the college will provide a fully functional and maintained course management system through which the instructor and students can interact.
4. When taught online, students must have currently existing email accounts and ongoing access to computers with internet capabilities.

Method(s) of Evaluation

Methods of Evaluation may include but are not limited to the following:

Written discussion of concepts on class forums

Formative activities, such as quizzes, assignments, and graphic organizers

Final course project

Method(s) of Instruction

Methods of Instruction may include but are not limited to the following:

Lectures which include motivation for the use of data visualization concepts, example code and visualization objects, and analysis of these products

Online labs (for all sections, including those meeting face-to-face/on campus), consisting of:

1. A programming assignment webpage located on a college-hosted course management system or other department-approved internet environment. Here, the students will review the specification of each programming assignment and submit their completed lab work
2. A discussion webpage located on a college-hosted course management system or other department-approved internet environment. Here, students can request assistance from the instructor and interact publicly with other class members

Detailed review of programming assignments, which includes model solutions and specific comments on the student submissions

In-person or online discussion which engages students and instructor in an ongoing dialog pertaining to all aspects of designing, implementing, and analyzing programs

When course is taught fully online:

1. Instructor-authored lecture materials, handouts, syllabus, assignments, tests, and other relevant course material will be delivered through a college-hosted course management system or other department-approved internet environment
2. Additional instructional guidelines for this course are listed in the addendum of CS department online practices

Representative Text(s) and Other Materials

Schwabish, Jonathan. Better Data Visualization, a Guide for Scholars, Researchers and Wonks. 2021.

Types and/or Examples of Required Reading, Writing, and Outside of Class Assignments

1. Reading

- a. Textbook assigned reading averaging 30 pages per week
- b. Reading the supplied handouts and modules averaging 10 pages per week
- c. Reading online resources as directed by instructor though links pertinent to data visualization
- d. Reading library and reference material directed by instructor through course handouts

2. Writing

- a. Writing to communicate analysis of existing products
- b. Writing to document processes used in analyzing and visualizing data
- c. Writing to describe trends or correlations found in data

3. Other projects

- a. Plots based on datasets, along with explanatory text
- b. Other forms of graphical objects that communicate data

Discipline(s)

Computer Science