BAX 431: Data Visualization

Instructor: Mak Ahmad, Adjunct Professor
Phone: (415) 935-3188
Email: shahmad@ucdavis.edu
Office Hours: By appt (over VC)

Course Description

Course topics include an overview of vocabulary, theory and principles as well as the algorithms and methods for structured and unstructured data visualization. The course will teach students how to present information in an understandable and efficient way for the purposes of data analysis and effective communication of results. The course also empowers students to critically analyze good and bad visualizations, and develop ways to improve on them.

Course Objectives

Data visualization is a particularly broad field of study, with research areas ranging from computer science to human-computer interaction to visual design. The goal of this course is to offer students a solid foundation in visualization theory and principles, and then help them apply this knowledge in weekly projects and workshops aimed at developing their skills in creating and evaluating data visualizations. Sessions will cover major areas of visualization including studies of visual perception; time-based, categorical, and geospatial analysis; graphs and networks; and visual storytelling, as well as the practical creation of visualizations using tools including Tableau, Salesforce Einstein Analytics, Python and Looker.

Learning Objectives

- Understand basic principles of data visualization
- Evaluate and think critically about data visualization examples
- Create effective visualizations for data analysis and presentations using a range of tools

Class Rules

Academic Honor Code: All students are expected to adhere to the University of California, Davis' Code of Conduct as noted here: http://sja.ucdavis.edu/files/cac.pdf. Please also note that if "academic misconduct is admitted or is determined by adjudication to have occurred," per Regulation 550 the student could potentially receive a grade of "F" not only for the assignment or project in question, but also for the *entire course*.

Use of Electronics in Class: You are allowed to use your laptop/tablet in class – however this use is conditional. You cannot use your laptops for any other activity other than those pertaining to the class. Hence, surfing, emailing, chatting, facebook visits and other related activities are not allowed. If such activity is observed in class, then laptop privileges will be revoked – *for the entire class*. Smartphone/phone use is not allowed in class – you are welcome to step out of class, with my permission, if you need to answer a call. Failure to comply with this policy will result in the creation of 'laptop zones' (e.g., only the first two rows will be allowed to use laptops).

Course Material

Optional Texts

- Now You See It, Stephen Few
- The Visual Display of Quantitative Information, 2nd Ed, Edward R Tufte
- Information Visualization, Perception for Design, Colin Ware
- Visualizing Data, Ben Fry
- The Functional Art, Alberto Cairo
- How Maps Work, Alan M. MacEachren

Grading

- 10% Class Participation
- 50% Assignments
- 25% Team Project
- 15% Final Exam

NOTE: Late assignments, exams, or projects, will not be accepted.

Assignments

See Canvas for details.

Team Project

See Canvas for details.

Lecture Outline

Lecture #1: Introduction to Data Visualization

Theory

- We will cover principles of data visualization, including how to model human perception in information visualization research, mapping from data types to visual marks and channels, and a taxonomy of visualizations, the process of visualization (Ben Fry, Visualizing Data, 2018), and laying out the fundamentals for the rest of the course.

Practice

- Download and install Tableau Desktop
- Download and install Anaconda (and be able to launch Jupyter notebooks)

Reading

- Tidy Data, Hadley Wickham
- Good enough to great: A quick guide for better data visualizations (see Canvas)
- ***Optional***: The Visual Display of Quantitative Information, Edward Tufte, Chapter 1: Graphical Excellence

Lecture #2: Elements of Data Visualization

Theory

- We will review the elements that make up data visualizations, including marks, scales, color, legends, labels, and interactivity. We will also introduce several archetypical charts often used in visual analytics, and how to choose between them based on the dataset in question (univariate, bivariate, trivariate, multi-dimensional data). We will also look at examples of advanced usage of these archetypes and real-word applications.

Reading

- Machine Learning for Hackers, John Myles White & Drew Conway: Chapter 2: Data Exploration
- ***Optional***: Now You See It, Stephen Few, Chapter 10: Distribution Analysis

Lecture #3: Charts + Python + Introduction to Tableau

Theory

- We will cover basic chart types, how chart types encode data and how they can be used. We will also discuss problems with various chart types, and we will provide an introduction to Tableau and Python.

Practice

- Python (using Seaborn library)
  - Horizontal bar plots
  - Annotated heatmaps
  - Scatterplot with varying point sizes and hues

Reading

- Tableau Public Tip Sheet
- Student Reference - Charts Overview
- ***Optional***: Now You See It, Stephen Few, Chapter 3: Thinking with our Eyes

Lecture #4: Visualizations with Tableau & Looker

Theory

- We will dive deeper into Tableau Desktop capabilities such as simple calculations, time between dates, data validations, etc. We will also cover how to analyze data with Looker and create dashboards.

Reading

- Register for Getting started with Looker and watch all the intro videos

Practice

- Download and play with the “the_tableau_interface_finished.twbx” file from Canvas
- Building Gapminder in Tableau (see Canvas if you want to follow along)

Lecture #5: *Visual Analysis for All!* By guest speaker: Jock Mackinlay - 10/31 10am-12pm **(combined class)**

Guest Speaker

- Jock Mackinlay PhD visualization expert, first Technical Fellow at Tableau

Theory

- This talk describes how evolution has equipped every person with the essential skills for seeing and understanding data. Individual and organizational data work is described with a Cycle of Visual Analysis. The talk then dives into data, view, and story (key components of the cycle) to explain why visual analysis is for everyone.
- You will learn:
    - What psychologists, statisticians and others have learned about using the human visual system to work effectively with data
    - Fundamental principles for designing effective dashboards for yourself and others
    - How to expand the use of visual analysis in your organization

Lecture #6: Visualizations for Business Intelligence

Theory

- We will cover the concept of a dashboard, a set of interconnected charts organized to support a specific task. We will look at the types of questions that business intelligence tools answer, the building blocks of dashboards, and the relationships between different components in a business intelligence application. We will review Analytical Visualization and different ways of interacting with data that are very useful, as well as best practices. In addition, we will cover crucial steps to consider while building dashboards.

Reading

- ***Optional***: Now You See It, Stephen Few, Chapter 4: Analytical Interaction and Navigation
- ***Optional***: Now You See It, Stephen Few, Chapter 5: Analytical Techniques and Practices

Lecture #7: *Interpretable ML Visualizations* by guest speaker: Nathan Mannheimer - 11/14 10am-12pm (combined class)

Guest Speaker

- Nathan Mannheimer, Senior Product Manager at Tableau

Theory

- Predictions are an increasingly important part of business analytics and decision-making processes. However, many of the decision makers don't have a deep understanding of the 'prediction machines' that create these guesses about the future including how they can make mistakes. This gap leads to an inability to ask critical questions about predictive modeling development and reasonably account for decisions incorporating predictions. Techniques in explainable AI allow both model developers and consumers to better understand what a model is learning and how it is making predictions, allowing for more interpretable inputs to decisions and better auditing of the expansion of ML into the business.

Reading

- ***Optional***: Now You See It, Stephen Few, Chapter 7: Time Series Analysis
- ***Optional***: Now You See It, Stephen Few, Chapter 9: Deviation Analysis

Lecture #8: Storytelling with Data

Theory

- We will cover the general principles of storytelling with data: exploration vs explanation, understanding the audience, mechanisms of attention and engagement. We'll present various narrative patterns for different purposes and explore successful examples.

Reading

- Narrative Visualization: Telling Stories with Data, Segel & Heer 2010
- NPR Guide to Hypothesis-Driven Design
- UI Checklist for Dashboard Design (download UI Checklist at the bottom of the page)

Lecture #9: *Salesforce Einstein Analytics* by guest speaker: Ziad Fayad - 12/05 10am-12pm (combined class)

Guest Speaker

- Ziad Fayad, Director of Analytics Cloud at Salesforce

Theory

- We will cover visualizations as they relate to business domains specifically what users in CRM systems typically use to monitor their business KPIs. We will also touch on how AI & Visualizations co-exist to provide the latest Augmented Analytics solutions to the business. Highlights include:
    - Visualizations in the daily business workflow
    - Examples Visualizations in the Sales CRM (desktop and mobile)
    - Recent Shifts in dashboards authoring tools
    - Augmented Analytics bridging hard core AI & dashboards personas

Practice

- Hands on exercise starting with identifying the user persona, building a dashboard, applying smart design techniques, and exploring augmented analytics common visualizations. You will receive an email a week before with a registration link to spin out Salesforce dev orgs to use (all cloud no need for installation)

Lecture #10: *Geospatial Visualization* by guest speaker: Nick Rabinowitz - 12/12 10am-12pm (combined class)

Guest Speaker

- Nick Rabinowitz, Senior Visualization Engineer at Uber, prev. UC Davis Adjunct Professor

Theory

- We will take a look at the theory and practice of visualizing geospatial data with maps. Starting by discussing the theory behind how we understand maps, we will review several common types of mapping visualizations

Practice

- Demo using Tableau and/or Kepler.gl

Reading

- How Maps Work: Representation, Visualization, and Design, Alan M. MacEachren
    - Chapter 1: How meaning is derived from maps
    - Chapter 2. How maps are imbued with meaning
- ***Optional***: How to Lie with Maps, Mark Monmonier, Chapter 2: Elements of the Map

Course Calendar

Date	Lecture Topic	Notes
10/03	Introduction to Data Visualization	
10/10	Elements of Data Visualization	
10/17	Charts + Python + Introduction to Tableau	Be ready to work with Python and Tableau
10/24	Visualizations with Tableau & Looker	
*10/31, 10AM	Visual Analysis for All!	Guest speaker: Jock Mackinlay
11/07	Visualizations for Business Intelligence	
*11/14, 10AM	Interpretable ML Visualizations	Guest speaker: Nathan Mannheimer
11/21	Storytelling with Data	
11/28	---	Thanksgiving Break
*12/05, 10AM	Salesforce Einstein Analytics	Guest speaker: Ziad Fayad
*12/12, 10AM	Geospatial Visualization	Guest speaker: Nick Rabinowitz

***NOTE**: These lectures will be combined sessions.*