# CSC/SDS 235: Visual Analytics

*Fall 2024*

## Final Project: A Visual Analysis

*This is a group assignment, work in groups of 3-5. Groups of different sizes must be approved.*

***Goals:***

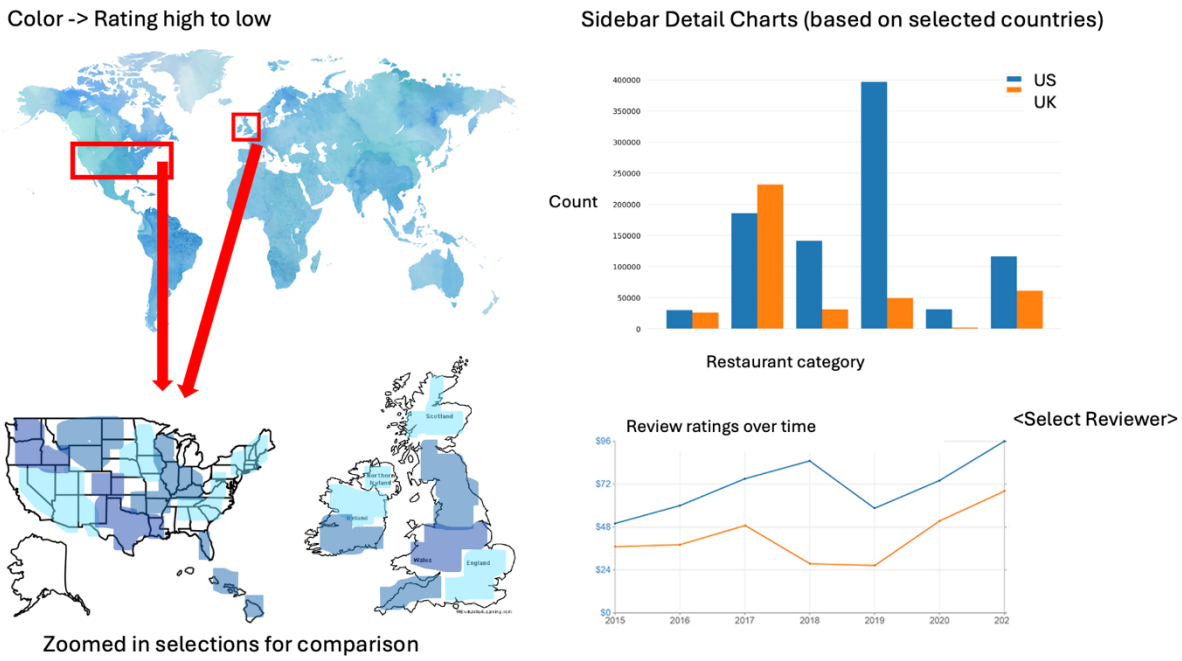
- Perform an analysis of data via a custom visual analytics system

### Instructions

For your final project you will work with classmates to design and create a visual analytic system that enables you to analyze a dataset of interest to you. You'll want to choose data that is interesting and complicated enough for you to use the visual analytics skills you've learned this semester for analysis. Below is an example:

- yelp open dataset: https://www.yelp.com/dataset
    - Example questions you might ask:
        - Cultural Trends: What makes a particular city different? For example, in which countries are Yelpers sticklers for service?
        - Location Mining: How much of a business' success is really just location, location, location? Do reviewers' behaviors change when they travel?
        - Seasonal Trends: What about seasonal effects: are there more reviews for sports bars on major game days?
        - Infer Categories: Do you see any non-intuitive correlations between business categories e.g., how many karaoke bars also offer Korean food, and vice versa?
        - Social Graph Mining: Can you figure out who the trend-setters are? For example, who found the best waffle joint before waffles were cool?

- Example Visual Analytics Systems for cultural trends question:




- Other potential datasets:
  - US Data: https://data.gov/
  - Kaggle: kaggle.com
  - Smith Data: https://www.smith.edu/your-campus/offices-services/institutional-research/data-about-smith
  - VAST Challenge Data: https://vast-challenge.github.io/2024/
  - SNAP: https://www.fns.usda.gov/data-research/data-visualization
  - Massachusetts Data: https://data.mass.gov/
  - NYC Open Data: https://opendata.cityofnewyork.us/
  - CDC Data: https://data.cdc.gov/browse?category=500+Cities+%26+Places
  - There are tons more, pick what's interesting to you and google around!

This project is a large portion of your final grade, and is broken up into milestones, described below. Be sure to submit each milestone and to put your best effort into all pieces. Your final project grade will combine all milestone grades. Milestones are due as listed on the course website. Be aware that some are due before class instead of 11:59pm. It is recommended that you read all instructions before beginning.

## Milestone 1: Proposal- 26 points

For your project proposal you will identify your project group, a plan for working together, a topic, a dataset, and an understanding of your data. Type up a document that answers the following questions.

1. Group
    a. With whom do you plan to work? Groups must be 3 – 5 members (you must speak with me if you would like to work with a different size group). [2pts]
    b. Talk about your schedules. Detail a plan for coordinating your work throughout the project. [2pts]
2. Topic
    a. What topic area will you work with? Why is this area important/interesting? [2pts]
    b. What questions does your analysis seek to answer? What is the broader impact of the answer to that question? [2pts]
3. Data
    a. Find the data that you will analyze. You may need to combine multiple datasets, so having more than one is okay. Provide a link to your data source(s). [2pts]
    b. If you plan to use more than one data source, how will you combine your datasets? What is/are the common identifiers across datasets on which you can link them? [2pts]
    c. Investigate the source of your data.
        - Who collected it? [2pts]
        - Who provided the funds for data collection? [2pts]
        - Why was the data collected? [2pts]
        - What possible data biases do the answers to these questions highlight? [2pts]
    d. Which variables in your data will you use for analysis? How does visualizing those variables relate to your answer to 2.b above? [2pts]
4. Roadblocks
    a. What roadblocks do you anticipate hitting as you complete this project? Name at least three. [2pts]
    b. Detail your plan for dealing with these roadblocks. How will you overcome them, or if you cannot, how will you pivot your project? [2pts]

## Milestone 2: Design & Prototype- 24 points

You will design your visual analytic system and produce a low-to-medium-fidelity prototype to bring to class for testing/feedback.

1. Task Analysis
    a. Revisit the analysis question guiding your work. Identify three tasks that your visual analytic system must support in order to answer that analysis question. For each task:
        i. Name the task [2pt per task]
        ii. Identify the high-level interaction needed to support the task [2pts per task]
        iii. Identify the low-level interaction needed to support the task [2pts per task]
2. Visualization Sketch
    a. Sketch out your visual analytic tool in detail. Your sketch must include:
        i. Clear (and appropriate) data – visual channel mappings [2pts]
        ii. Notation where appropriate (Ex. To explain interaction or coordinated views) [2pts]
3. Prototype and Feedback
    a. Create a low-fidelity paper prototype, or a medium-fidelity computer-based prototype of your visualization. [2pts]
    b. You will bring your prototype to class for testing and feedback as listed on the course schedule.

## Milestone 3: Redesign- 14 points

You will pair up with 3 other groups to test your prototype in class. The goal is to identify any parts of your design that are unintuitive, or confusing to an end user.

1. Testing
    a. In class, find another group to pair up with. Ask one person in that group to complete each of the tasks you identified for milestone 2. Do not provide any feedback, help, or cues while they complete the tasks, but do ask them to "think aloud" and take notes on what you observe. [2pts]
    b. Repeat the step above with another group. [2pts]
    c. Repeat the step above with another group. [2pts]
    d. Review the notes you have from your three tests. Synthesize common stumbling blocks/design issues you noticed across the three tests. [2pts]
2. Redesign
    a. Based on the results of your testing, re-work your visual analytic tool sketch. Include:
        i. Clear (and appropriate) data – visual channel mappings [2pts]
        ii. Notation where appropriate (Ex. To explain interaction or coordinated views) [2pts]
        iii. Notation of changes from the original design [2pts]

## Milestone 4: Final Product & Presentation- 20 points

You will demonstrate your final product to the class during one of our last two class periods. **You must upload your submission to Gradescope before class on the first day of presentations.**

1. Visual Analytic Tool
   a. Includes at least 3 *different* visual encodings. [3pts]
   b. Includes at least 1 interactive feature. [1pt]
   c. Includes at least 2 coordinated views. [1pt]
   d. All visual encodings are labeled appropriately. [2pts]
   e. All visual encodings are titled appropriately. [2pts]
   f. All data-visual mappings are appropriate. [1pt]
2. Presentation
   a. An introduction to your analysis, including the motivation behind it. [1pt]
   b. A brief overview of the context for your data. [1pt]
   c. A demonstration of your visual analytic tool. To demonstrate, show us the completion of the three tasks your tool supports. [2pts per task]
   d. A conclusion—answer the analysis question you posed to start. [1pt]
   e. Time for Q&A. In total, your presentation should be ~10 minutes. [1pt]

## *Submission*

Submit your milestones on Gradescope as a group (https://guides.gradescope.com/hc/en-us/articles/21863861823373-Adding-Group-Members-to-a-Submission).