CSC/SDS 109: Communicating with Data

Fall 2024

# HW 02: First Visualizations

This is an individual or pair assignment-- you pick!

**Goals:**

* **Work with real data**
* **Use Tableau to visualize data**

## Instructions

Step 1: Obtain Data

* Download a dataset from [awesome-public-datasets](https://github.com/smuthubabu/awesome-public-datasets), [Tableau Public](https://public.tableau.com/s/resources?qt-overview_resources=1#qt-overview_resources), [Kaggle](https://www.kaggle.com/datasets), or any other source.
* Load the dataset into Tableau and explore the dimensions. For example, if you download the [Auto dataset](https://jcrouser.github.io/CSC109/datasets/Auto.csv), which contains [information on various types of cars](https://rdrr.io/cran/ISLR/man/Auto.html), you'll see something like this:A screenshot of a graph

  Description automatically generated

Step 2: Deliverable

* Create THREE different visualizations that highlight something interesting in your data.
* Create a document that contains:
  + A link to your dataset.
  + An overview of your dataset :
    - Where does it come from?
    - Who collected the data and why?
    - Is there any important context to consider?
    - Do you anticipate any ethical issues or biases in the data?
  + Each visualization paired with a description of the interesting thing the visualization shows.

## Submission

Submit your deliverable(s) as a PDF on Gradescope. If you worked with a partner, submit as a group (<https://guides.gradescope.com/hc/en-us/articles/21863861823373-Adding-Group-Members-to-a-Submission>).

## Rubric

The following matches the rubric you will see on Gradescope.

|  |  |  |
| --- | --- | --- |
|  | Points | Criteria |
|  | 1 | Submission is well-formatted and easy to read. |
|  | 1 | Submission includes a link to the original dataset. |
|  | 4 | Submission includes an overview of the dataset (1pt per bullet point above) |
|  | 3 | Submission contains THREE different visualizations (1pt per visualization) |
|  | 3 | Submission contains a brief, readable, and accurate description of each visualization’s interesting thing (1pt per visualization) |
|  | 3 | Each visualization uses an appropriate data🡪visual mapping as discussed in lecture. Eg. If a bar chart is present, is it being used to compare related quantities? (1pt per visualization) |
| TOTAL | 15 |  |