CSC/SDS 109: Communicating with Data

HW 02: First Tableau Visualizations

This is a pair assignment! Individual submissions must be pre-approved.

Goals:

- Work with real data
- Use Tableau to visualize data

Instructions

Choose Data

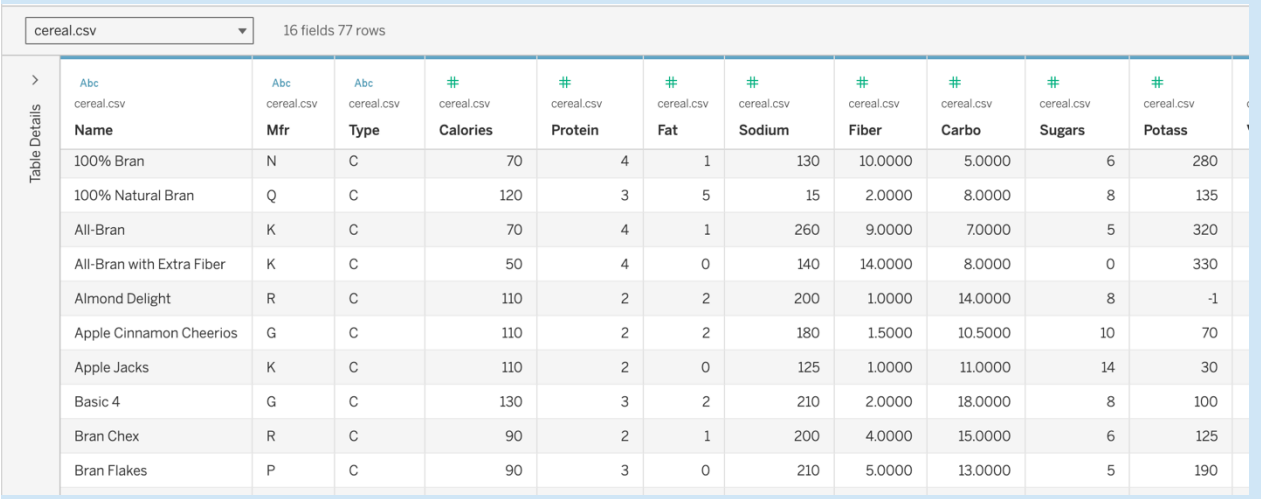
Choose one of the datasets under the "In-class" tab on the course website:
****Use the dataset on the course website – they've been cleaned for you**

- College.csv (source: https://www.kaggle.com/datasets/yashgpt/us-college-data)
- cereal.csv (source: https://perso.telecom-paristech.fr/eagan/class/igr204/datasets)
- palmerpenguins.csv (source: https://allisonhorst.github.io/palmerpenguins/articles/intro.html)
- bluebikes-tripdata_sm.xlsx (source: https://data.boston.gov/dataset/blue-bikes-system-data)

Deliverables

Research the dataset you chose. You may need to do additional research beyond clicking the source link above. Answer the questions in the attached worksheet about your dataset.

Load the dataset into Tableau and explore it. For example, if you download the cereal dataset, you'll see something like this:



cereal.csv — 16 fields 77 rows

Table Details

Name	Mfr	Type	Calories	Protein	Fat	Sodium	Fiber	Carbo	Sugars	Potass
100% Bran	N	C	70	4	1	130	10.0000	5.0000	6	280
100% Natural Bran	Q	C	120	3	5	15	2.0000	8.0000	8	135
All-Bran	K	C	70	4	1	260	9.0000	7.0000	5	320
All-Bran with Extra Fiber	K	C	50	4	0	140	14.0000	8.0000	0	330
Almond Delight	R	C	110	2	2	200	1.0000	14.0000	8	-1
Apple Cinnamon Cheerios	G	C	110	2	2	180	1.5000	10.5000	10	70
Apple Jacks	K	C	110	2	0	125	1.0000	11.0000	14	30
Basic 4	G	C	130	3	2	210	2.0000	18.0000	8	100
Bran Chex	R	C	90	2	1	200	4.0000	15.0000	6	125
Bran Flakes	P	C	90	3	0	210	5.0000	13.0000	5	190

Using Tableau, create three DIFFERENT visualizations that each show something interesting in your data. Each visualization must:

- Show unique variables of the dataset (i.e. not the same as your other two visualizations)
- Be a unique visual encoding (i.e. not the same as your other two visualizations)
- Include:
    - Descriptive title
    - Readable axis titles
    - Readable axis labels
    - A legend when necessary
    - Zero double encodings
    - Appropriate data-visual mappings

Save each of your visualizations as a .png (use Export Image then Save As png). For each visualization, add the PNG to the attached worksheet and answer the questions that follow.

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
General:	0.5	Raw data collector
	0.5	Raw data collection time
	0.5	Raw data observations
	0.5	Raw data variables
	1	Raw data biases
For each visualization:	1	PNG
	1	Unique dimensions
	1	Unique visual encoding
	1	Descriptive title
	1	Readable & descriptive axis titles
	0.5	Readable axis labels
	0.5	Legend if necessary
	1	Zero double encodings
	1	Appropriate data-visual mappings
	1	Description of interesting trend
TOTAL	30	

SDS/CSC 109 hw02 Worksheet

Fill out this worksheet with respect to the dataset you chose.

Data Investigation

Who collected the raw data?

When was the raw data collected?

What does one observation (row) in the raw data represent?

What variables (columns) are in the raw data?

What biases are present in the raw data?

Fill out this worksheet with your three visualizations.

Visualization 1

Add your png.

What is the data → visual mapping in the visualization?

Describe one interesting trend shown in the visualization.

Visualization 2

Add your png.

What is the data → visual mapping in the visualization?

Describe one interesting trend shown in the visualization.

Visualization 3

Add your png.

What is the data → visual mapping in the visualization?

Describe one interesting trend shown in the visualization.