#1: Tableau

Due Sep 2, 2019 by 11:59pm **Points** 10
Submitting a file upload **File Types** zip
Available Aug 27, 2019 at 12am - Sep 2, 2019 at 11:59pm 7 days

This assignment was locked Sep 2, 2019 at 11:59pm.

This homework allows you to grasp what visualizing data is like using a commercial visual analysis software.

The objective of this assignment is to practice using visualization techniques to explore, analyze, and explain a dataset (and also get some HTML coding practice, which you must know to successfully complete this class). For this, we will use Tableau (**https://www.tableau.com (https://www.tableau.com/)**), a popular software for "drag-n-drop" data visualization. It's a nice software for exploring and initially visualizing a dataset, and is great for prototyping visualizations before implementing them using libraries like D3 and ggplot.

For this assignment, you will download and install Tableau, learn how to use it, and then use it to explore a dataset.

- **Tableau for Students** **(https://www.tableau.com/academic/students)** provides free licenses to students. Use it to download and install Tableau.
- Next, learn how to use Tableau. The Tableau site provides a nice **learning section (https://www.tableau.com/learn)** .
- Then, find a dataset online that you'd like to study from **Kaggle** **(https://www.kaggle.com/)** , a website that hosts thousands of datasets. You may pick any dataset that you like, but try not to use one that is too simple. Your exploration will be too shallow!

Generate several questions about your chosen dataset. Then you will use Tableau to visualize the data and find answers. You will log this exploratory process as a "data exploration story." You should ask at least three questions about the dataset. You will document each step of the process (e.g., question → visualization → answer) and will turn in a single HTML webpage that you create. For each question on your web page, you should have the following:

- The question posed
- An image/png of your created visualization that you used to answer the question
- The answer to the question
- A brief description of how or what about that visualization helps you answer the question

You can pose your questions in an iterative manner. For example, suppose you first load your dataset into Tableau. You might start by posing a broad, high-level question. Then, using Tableau, create a chart that tries to answer this question. (Note: Before doing so, you might need to assess the fitness of your dataset. Maybe you need to clean it first! Or aggregate some data points together!) Upon creating this initial chart, you may find your initial assumption has changed---perhaps a new question now emerges. Maybe the data shows something different than expected? Or maybe now you can ask a more specific query about some subset of the data. Based on this new question, repeat the chart creation process.

You must create a total of at least 3 charts. Each chart should be a different visualization technique (i.e., if you first used a scatter plot, you cannot re-use it again). This is an organic exploration process, so we are intentionally leaving it open-ended as to what constitutes the "correct" analysis or the "correct" charts; instead, focus on trying to discover aspects, patterns, outliers, and/or trends of the data using visualization, so use your judgment and follow what interests you!

Screenshot each chart to save it as an image. With your set of questions and created charts, come up with what you feel is a good data story that explains your analysis and exploration of the dataset. Create one HTML page that tells this story (all 3 charts on the same page, scrolling is okay). Include a short introduction (with a link to your dataset) before your questions/charts and a short summary at the end. Include a title and your name at the top of the webpage. Remember that you are telling a narrative about your discovery process; you want to inform and explain your thought process, and hopefully tell something interesting about the dataset.

Beginning the Assignment

Since GitHub accounts aren't set up yet, homework #1 will be directly submitted here on Canvas. There is a template set of HTML code in the **Files** tab that you may use for your assignment. It contains the following:

- An **index.html**
- A linked CSS style sheet (the Bootstrap library)
- A folder **imgs/** for storing your Tableau screenshots.

You're may use a differently-styled site if you prefer, it's up to you! But make it look presentable and nice.

The completed assignment will be the web page containing your data story. Be sure to use relative links to use image files and CSS file so they render correctly.

Zip all the website files together and submit here on Canvas. **Be sure to double check your work, to make sure it was uploaded correctly! And remember, no late work is accepted.**

Grading Criteria

- **3 pts: Exploratory Process**
  - The questions asked are applicable to the chosen dataset
  - Analysis on the data makes sense (i.e., it's not trivial)
  - Follow-up questions drill deeper down into the data
- **4 pts: Created Visualizations**
  - At least 3 visualization techniques are used
  - The designs make sense (techniques used are appropriate for the data, color choices make sense, etc.)
- **3 pts: Web Page Design**
  - The data story is contained in a single page
  - The page contains the necessary charts, story text, title, intro, and summary
  - Page design and styling is nicely done