Data Visualization Proposal

Project Title: Analyzing the Effects of Demographics and Behaviors on Social Media Use and Changes from 2018 to 2019

Team Members:

Jonathan Sloan

Samihah Islam

GitHub Repository: https://github.com/PudgeTheFish/DataVisProject

Background and Motivation:

We chose this project because over the past few years, attitudes towards big technology companies, especially in the media, have appeared to grow increasingly negative. Paradoxically, teenagers and adults are still using the social media sites owned by these companies in droves. We'd like to investigate the trends of social media and internet use from 2018-2019 and analyze the effects of demographics and other behaviors.

Objectives:

We'd like to discover how attitudes towards technology changed from 2018 to 2019 among different demographics and behaviors of Americans. We hope that controlling for these factors can help us understand where technology usage is shifting in the American populus. This could be the baseline to helping understand how media itself is shifting public perception.

For example, conservative media often criticizes large technology companies, but has that affected how convervative individuals use technology? What about for liberal individuals? Young people tend to be more appreciative of technology, does their usage reflect that?

What we hope we find is patterns amongst public usage of technology by different cohorts, which could be the first step in identifying why those differences exist.

Data:

We are using data from the Pew Research Center, which runs an annual survey of trends in technology. The survey has a sample size of 2,000 Americans, which research has shown to be large enough to get an accurate sampling of the American populus.

Links:

https://www.pewinternet.org/dataset/jan-3-10-2018-core-trends-survey/

https://www.pewinternet.org/dataset/core-trends-survey/

Data Processing

The survey includes some questions that were only asked to respondents based on earlier responses so we need to remove those data points so that our data we have answers from all respondents. We also have some extraneous data points such as household size and location.

We will need to be able to filter data by certain attributes such as race, sex, etc so that each can be displayed separatey. We will also have to sum the number of responses for each frequency of use level and convert them to percentages of the total responses so that the data can be comparable across different cohorts and social media apps. This shouldn't take too much time, but it will be tedious so it might take a few hours.

Must-Have Features

Based on the problem question, some essential features include:

- View overall changes in trends between 2018-2019 for all participants
- View changes in trends between 2018-2019 for individual demographic breakdowns
- Compare social media and internet usage between demographic groups and other defined cohorts
- Selectively hide and show data (filtering) on the graph, which is useful for comparing data within one or a few cohorts

Optional Features: List the features that you think would be nice to have, but not critical.

- Provide custom querying for comparing specific demographics to specific questions
- Add more years of data (cannot find source yet)

Project Schedule:

10/27-11/2 Project kickoff

11/3-11/9 Clean data & set up page (Samiyah clean, Jonathan set up page)

11/10-11/16 Complete data nests & set up scales (Jonathan nest, Samiyah scales)

11/17-11/23 Complete charts (divide evenly)

12/1-12/7 Add interactions (divide evenly)

Visualization Design

See Design.pdf for the design pages.