**Team Project One**

Potential data topics plus a few others just for shits n giggles:

1. Marijuana acceptance (by age group, state, political affiliation etc…?)
2. Student loans (amount, age, gender, etc…?)
3. Alcoholism statistics (age, socioeconomic status, DUIs, alcohol-related deaths by age/gender etc…?)
4. Google searches (Popularity, category, count by day/week/month/over time)
5. Apples (popularity, Granny Goose, Fuji, Honeycrisp, Gala, Delicious Red, sales by state, sales by type)

Things to consider:

1. Some topics may not have sufficient data sets. If you find yourself struggling to find data for a topic, have a topic b or topic c (plan b, plan c etc) and move on to it. It will tough enough to write the code for this, so finding a topic with sufficient data will help immensely.
2. Related to above. How many questions can you ask, and is there data for your questions.

Everything below is just filler to help us getting started

# Hypothesis

## Interested in showing whether there is a correlation between x and y. We believe that the data will show\_\_\_\_\_\_\_\_\_\_\_\_.

# Data Source:

* 1. <https://www.kaggle.com/explore-projects> (etc)
  2. Data source 2 (not sure….if necessary?)
  3. Data source 3 (not sure…if necessary?)

# Data elements to look at

* 1. Age range
  2. Gender
  3. Socioeconomic status
  4. Cost

1. Sample size
   1. How many? 100? 500? 1000?
2. Data to plot/graph (need 6-8 total, rough estimate at least 2 per question asked of data ):
   1. Age range and gender
      1. Type of chart: bar
   2. Age range and count
      1. Type of chart: Pie chart
   3. Gender and socioeconomic status
      1. Type of chart: Scatter
3. Data Summary
   1. Conclusions
   2. Implications
   3. Story-telling!
4. Finish Line!!!