Exploratory Data Analysis

For practice in exploratory data analysis and reporting, you will pick 1 or 2 datasets from the Tidy Tuesday sets https://github.com/rfordatascience/tidytuesday (or if you have some other data you want to explore ask me) to ask and answer 10 questions about it. Tidy Tuesday is a weekly event run by the R4DS Online Learning Community for people to practice their data wrangling skills and sharpen their exploratory work, so they're perfect to practice with.

There is a short example that you can base your analysis on, as well as a template .Rmd file to get you started. This will be due next **Thursday 7/22 by 9AM**. You will give a short 10 minute presentation to your fellow students and a few people from Coriell the following day **Friday 7/23 at 12PM**.

The Report

Here's an outline of what you need to include in your report:

A. Introduction
    a. Needs to include a short description of the dataset. One (or two max) paragraphs similar to how datasets have been described through the assignments.
    b. Also needs to include a bulleted list describing the columns, again like you've seen throughout the assignment.
B. Data Preparation
    a. You need to read the table in this section
    b. You should also do any general data wrangling to tidy up the table in this section and explain why you reformatted/reshaped the data the way you did. This doesn't mean that you can't or shouldn't wrangle the data later, but you should do anything that would generally apply for all your questions
C. Questions
    a. This section should contain the 10 questions and your answers
    b. Each question should have:
        i. A question
        ii. A plot illustrating the question
        iii. A statistical test to answer the question
        iv. A short explanation of how you interpret your results
    c. This is also laid out in the example
D. References
    a. Primarily just reference where your data came from. This should be the Tidy Tuesday repository for the dataset, and if the repository references anything else please cite that as well
    b. If you need any other information to understand the dataset and/or write the conclusion include that as well

---

The Presentation

This should include:

- Any background necessary to understand the dataset
- 3-5 of your plots and their conclusions