**STA 404/504 Homework 1**
**Basic Plots with ggplot2**

Submission through Canvas - Due Monday Feb 11th by 5:00pm

Learning Objectives:

- Reading data into R
- Identifying basic R object characteristics
- Building basic visualizations
- Writing short descriptions of plot stories

Start-up R code:

```
library(ggplot2)
music <- read.csv("https://kmaurer.github.io/documents/data/shuffled_playlist.csv")
```

Assignment: Load the *shuffledplaylist.csv* dataset into R. This data includes information on songs from a shuffled Spotify playlist. Answer the following questions:

1. [2 pt] Answer each of the following:
    a. How many songs are in the data frame?
    b. What song has the highest “valence” score?
    c. What is the average energy for hip hop songs?
    d. What percentage of songs the data are by the artist Dessa?
2. [2 pt] Create a scatterplot of the relationship between the energy and tempo of songs. Describe in a few sentences what this display tells us about the relationship between the variables.
3. [2 pt] Create and describe a plot that tells us about the distribution of danceability in this playlist.
4. [2 pt] How do the tempo values differ by genre? Use a side-by-side style of plot to help tell a story of how they differ. Describe in a few sentences what this display tells us about the relationship between the variables. Why did you choose this type of display?
5. [2 pt] Create an additional plot of your choosing that tells an interesting story about the data. Describe this plot in a few sentences.

Submission Format: **R code file** showing your work **+ a pdf or word document** contain your written solutions. Note that plots need to be included in your write-up!

( Note: R markdown is also an acceptable alternative for those who know how to use it; if used please submit the .Rmd and pdf document. )

What will we look for in grading?
Correct values, axis labels and titles, labeled units, correct plots, proper grammar in write-up, insight into data.