Points: 28 for part 2. Part 1 is the completion of the survey and it has 12 points. So the total points are 40 for homework 1.

Out: September 14, 2020 (Monday)

Due: 11:59PM, September 18, 2020 (Friday). No late homework will be accepted.

Handin method and requirement: name your file as **yourlastname-firstname-studentID-hw1.pdf**. For example, if your name is Amy Zhang, the file should be named as zhang-amy-5678910-hw1.pdf. Upload your pdf file to Canvas.

-----

**Homework overview:** Now we have collected the top 5 character strength data from our class, let us use what we learned in Exploratory Data Analysis to answer the following question:

Do female and male students tend to have different character strengths?

For example, you can generate the histogram of 24 character strengths for female and male students, respectively. This will allow us to observe whether these two histograms are conserved intra-gender while highly different between the two genders.

**Requirement and grading:** please use two visualization techniques (feel free to use any technique that is not included in class) to draw an observable conclusion.

- 1) Tell us which two visualization techniques you used and show their plots for our data
  - a. Correct choice and application of the visualization techniques: 14 pts (7 for each)
- 2) Do a brief analysis based on your figures: do you think female and male students tend to have different character strengths?
  - a. Clear and easy to understand analysis/explanation. Be brief and clear. 10pts
- 3) You can use excel or any other tools/software to help generate the figures. Please tell us which tool you used (4 pts)

## Data:

Canvas → files/homework, "character-report-1.csv" and "character-report-2.csv". Each complete sample contains 6 attributes: the top 5 character strengths and the gender. The column names in the csv file are self-explanatory. Combine both files into one file and remove lines with incomplete attributes (pre-processing).