# Project 1

In this project, you will be looking for a relationship between two measurement (quantitative) variables.

**Step 1:** Choose two measurement variables that interest you and may possibly be linearly related. Your assignment is to determine if the variables are linearly related (and how strongly they are) using the regression techniques discussed in class. Make sure you are interested in the topic. While height and shoe size will make a dandy regression model, it is VERY BORING for me to read and I will be grumpy when I grade your paper. It is the same thing with GPA- it is very depressing if that is all you care about. These project ideas will not be fun, and it will make this project torture for you to do.

**Step 2:** Collect **at least** 20 observations of those two measurement variables. You can do this by conducting surveys, researching data, or setting up an experiment and measuring data yourself. Try to avoid sampling bias.

**Step 3:** Look closely at your data. Are there any oddities? Are there any outliers?

**Step 4:** Fit a regression line to your data. Is it a good fit? Would you be willing to use it for prediction?

**Step 5:** Write up your findings in a paper that includes the following section. Shoot for a text length of about a page. *Any number you put in the written report must be interpreted in terms of your topic.* Good papers use a professional style and are interesting because they show evidence of critical thought.

## Write Up & Grading

**Section 1:** (15 total pts) Introduction:

What question are you trying to answer?

**Section 2:** (20 total pts) Data Collection:

(10 pt) How was your data collected?

(10 pt) What did you do to avoid sampling bias? What factors were beyond your control?

**Section 3:** Regression Analysis: (50 total pts)

(20 pts) Give a scatterplot with the regression line. Make sure the graph is neatly done, is clear, and includes what it needs to, and only what it needs. It should be titled, have axes labeled, and have the regression equation and $r^2$ value. (You should probably use Excel for this!) NOTE: Do you need the box that says "Series and Linear" that Excel automatically creates? Probably not, huh. (Maybe you should delete it.)

(15 pts) Interpret the slope, intercept and coefficient of determination.

(15 pts) Are there any outliers? If so, are they influential? How did you deal with them? If there are no outliers, discuss why not. Is it impossible for outliers to exist?

**Section 4:** Conclusion: (15 total pts)

Are your two variables linearly related? If so or if not, what does this mean? Be careful not to overstate your findings, just talk about what you measured and observed in your study or experiment. Sum it all up. ☺

*Note: This is an individual project, not a group project.*