Problem Set 1

Q1. What is your name and email address?

Q2. Sign up for GitHub (www.github.com) and go through the GitHub Bootcamp. What is your GitHub username?

The International Capybara Racing League is concerned that the scoring system they use is too easy to game, and that it reflects how *many* races a Capybara is entered into rather than how fast that Capybara actually runs. They have hired you to determine if there is a clear association between the number of races attended and a Capybara's score for the season.

Q3. Import 'Scoring.csv' into R, and plot races by score. Does there seem to be a relationship? Include your plot in your answer.

Q4. Now fit a simple regression predicting score by races. What is the coefficient you obtain, and what does it suggest?

Q5. Plot your model fit. Do you find it satisfactory? Justify your answer.

Q6. Fit a model with a squared term for races, and plot the model fit. Does this seem like a better fit for the data? How would you interpret your results?