1. **[Statement of Research Question (Due 8-29)](https://mcdaniel.blackboard.com/webapps/assignment/uploadAssignment?content_id=_1372071_1&course_id=_171342_1&group_id=&mode=view)**

A statement of your research question.

Roughly speaking, an empirical research paper is a research paper that includes data analysis. What topic are you interested in researching? I encourage you to propose a topic of personal interest. It is okay if modifications or changes need to be made later. What variables would be most relevant for your analysis? How would these variables be measured (i.e. continuous, discrete, qualitative)? What type of graphs or descriptive statistics would be most useful for describing your dataset?

It is acceptable for you to specify a research question that can be addressed with the datasets available in the ANA 500 course site. Those datasets are also located in this Blackboard site in the Datasets folder.

**My submission:**

**I am interested in researching the relative fitness capabilities of individuals; how do the factors of height, weight, age, gender, etc affect performance (if at all)? My current dataset is from 2019 and it contains data from various powerlifting meets from around the world.**

[**https://www.kaggle.com/open-powerlifting/powerlifting-database**](https://www.kaggle.com/open-powerlifting/powerlifting-database)

**One of the things I found interesting this summer (not reflected in my current dataset as it was not powerlifting) was that during the first event of the 2021 Crossfit Games, men and women appeared to have a similar performance -- their gender did not have a significant impact on how quickly they swam. In fact, it was noted by the commentators early on during the competition that there were more women in the lead than men.**

Instructor feedback:

This is an excellent research question! It has clear focus and purpose. In this research question, you have already pointed out the dependent and independent variables in this study. Your independent variables include categorical variables (e.g., gender) and continuous variables (e.g., weight). You could use a multiple regression model to answer your research question. Or you could run some basic analysis (e.g., correlation, simple linear regression) to explore their relationship among these factors.

Many results have shown that women will not run, jump, swim, or ride as fast as men. However, I am surprised to learn that men and women appeared to have similar performance during the first event of the 2021 CrossFit Games. It is a bit anti-intuitive but very interesting, and I encourage you to explore the gender gap in sports performance.