Milestone 3: Beyond Descriptive Stats

My client/dataset is SportsStats. It is a sports analysis firm partnering with local news and elite personal trainers to provide "interesting" insights to help their partners. Insights could be patterns/trends highlighting certain groups/events/countries, etc. for the purpose of developing a news story or discovering key health insights.

Go broader

1. What jumps out at you now?
    - I was able to find a positive correlation (linear regression) between the number of athletes per country and the number of medals won by that team: the more athletes you have in your team, the more medals you can win.
2. Use the descriptive stats to point you to features that you may now want to consider.
    - I deep dived into my data considering men swimmers and trying to find some correlation between Age / Height / Weight and Total medal won. Unfortunately, there is no correlation between these parameters.

What key terms did you discover in any text analysis, for whom? Any themes? If you are not analyzing text, summarize what other things you are considering in your analysis?

- I discovered that it is hard to determine any correlation between the number of medals won and athlete's physical characteristics. I can only consider average values (Age, Height and Weight) to give a global indication on the archetype of the winning athlete, but it is not something to rely on to predict future winner.

New metric

Create 1 or 2 new metrics to track relationships of data you discovered. Explain why you created them.

- My cross-category metric is the **total number of medals** won. It helps me verify that any correlation, relationship between categories is valid or not.