

Olympic Medals (1896-2014)

Overview – we will be conducting some analysis of medalling trends for the Summer and Winter Olympic Games

Project Description/Outline:

We will analyse a master dataset and split into – Summer Olympics & Winter Olympics – in order to answer the following questions relating to medal count.

Research Questions:

- 1) Who were the top ten performing nations for each Olympic Games and how does the top ten table change over time?
- 2) Who medals in each new sport and what was their rank in total medals table for the previous games
- 3) Taking three games prior and three games post the host Olympics – what is the impact on medalling increase/decrease % ? (correlation)

Dataset to be used:

Olympic Sports and Medals 1896-2014

<https://www.kaggle.com/the-guardian/olympic-games>

Bonus

- 1) Distance of Host country from Country of Origin - % increase/ decrease impact on medal count

Rough Breakdown of Tasks

- Understand the data as a group and identify limitations of the data Vs. intended questions
 - Clean data in pandas DataFrame
- 1) Identify top ten nations
 - 2) Identify the year and location of each games – host nations
 - 3) Identify when new values appears in the sports category
 - 4) Pull out individual dataframes for each host nation +3/-3 games
 - 5) Identify new sport medalists & get rank from previous games for comparison
 - 6) Decide upon most suitable visual for each question
 - 7) Create plot for each question
 - 8) Group discussion on findings
 - 9) Write written conclusion and analysis for each individual question
 - 10) Build slides for presentation

Tentative Schedule

Question 1 – Finish on Thursday, 5th

Question 2 – Saturday, 7th

Question 3 – Tuesday, 10th

Slides complete by Thursday 12th

