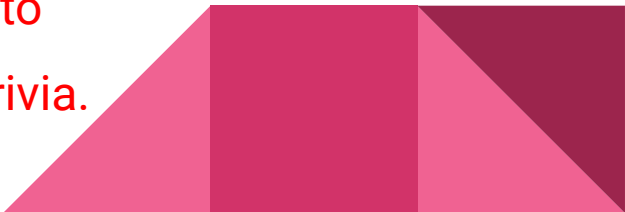# Presenting Your Findings

Milestone 4 Peer reviewed Assignment by M SURIYA RAM
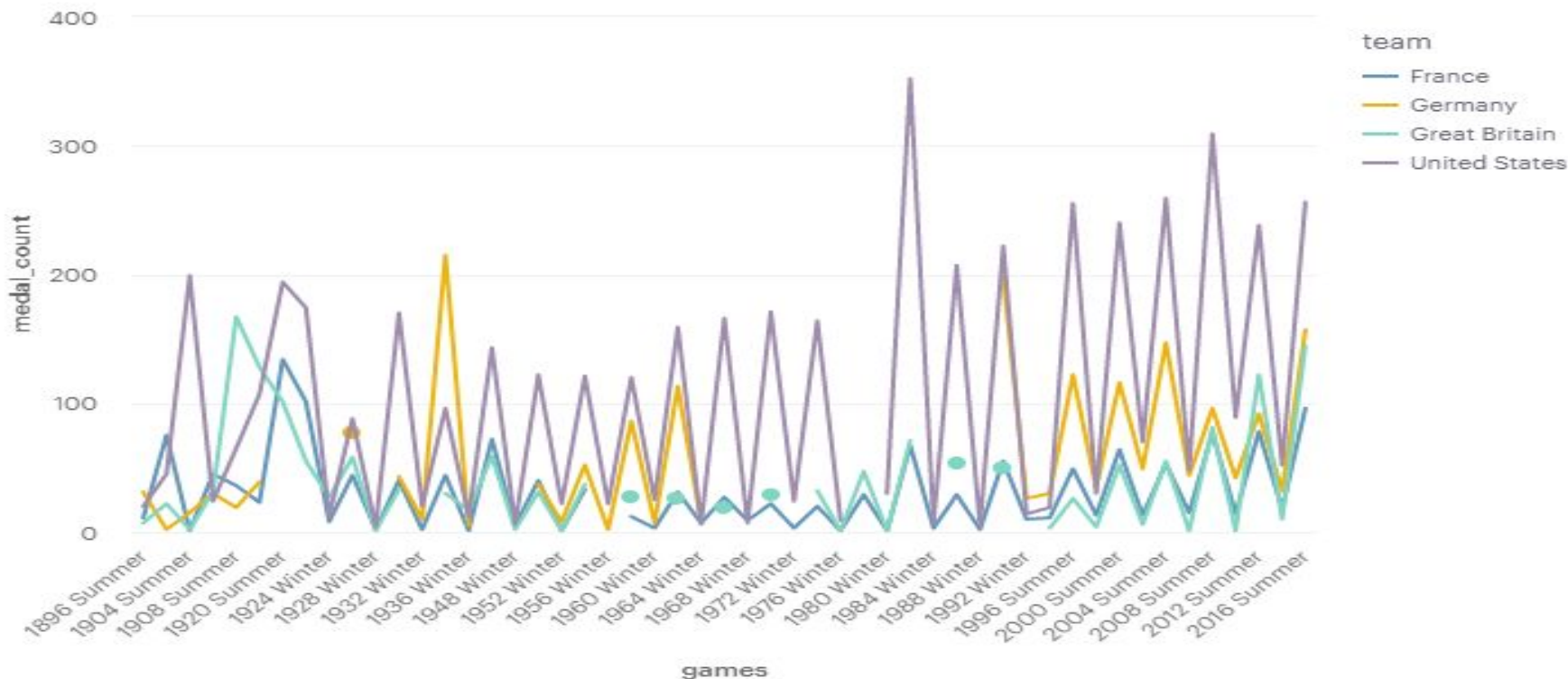
Based on work done from the SportsStack Olympics Dataset
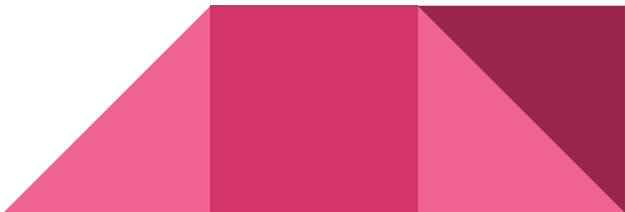
The Client and some history

- The dataset consists of 2 tables with details like the names of athletes, sports, medals won, countries, etc for the last 120 years of the Olympics.
- I am looking to study the top performing countries and find some reason for their success while looking at other crucial aspects that determine winning.
- My client is any sports magazine/media who is interested in writing/running a feature on the history of Olympics and a look back at time.
- My target audience is sports history buffs who are into documentaries, reading and are interested in some trivia.

Initial Analysis of the Dataset

Medals won over the years by the top 4 countries

Some findings and assumptions:

- One of the countries that showed up in the top 5 best performing countries in the Olympics was Soviet Union.

- This was not considered as part of the study because it was dissolved.

- The 4 countries considered for the entire study are: The United States, Germany, France and Great Britain.

Questions to answer- THE BIG PICTURE

- While looking at the graph obtained during the exploratory data analysis, there is an up-down periodic rise and slump in the number of medals won by the top 4 teams. What is the reason for this trend?
- Is there a relationship between sex and sport? For example, are men known to predominantly win in a specific sport and women in another sport? If yes, why the discrepancy? To be concise, is there any relation between the sport that is played and the more dominant sex in that game?
- What impact does the season have on the medal-winning count on the country?

Initial HYPOTHESES

- The United States has won so many medals only because they have sent the most number of participants.

- The majority of a country's medals are concentrated in a certain number of sports. This is common for all countries.

- The event, in general, does not matter compared to sports. For example, an athlete has an equal chance of winning a gold medal in any of the events that fall under athletics, and so on.

HYPOTHESIS 1- The US won most medals because they sent most participants

- For this, we will be looking at the percentage of people who won a medal out of the total people who participated in the Olympics.

	team	gold_percent	silver_percent	bronze_percent	win_percent
1	United States	13.862273097038269	8.472011983394623	6.908724457025528	29.24301028251648
2	France	3.7954621016979218	4.320987686514854	4.813146591186523	12.92959600687027
3	Great Britain	4.551034793257713	5.1034726202487946	5.015783756971359	14.670291543006897
4	Germany	7.280720770359039	6.723139435052872	7.2699978947639465	21.273857355117798

- The win percent for the US is the highest out of all top performing countries. This corroborates with our initial hypothesis, thus proving it right.

HYPOTHESIS 2- Majority of the medals are concentrated in a select few sports

	cumulative_percent	sport	percent_contribution
1	27.76049766718507	Athletics	27.76049766718507
2	55.39139450492483	Swimming	27.630896837739762
3	64.23017107309487	Basketball	8.838776568170035
4	72.8615863141524	Rowing	8.631415241057542
5	80.01555209953344	Ice Hockey	7.153965785381026
6	85.01814411612234	Shooting	5.002592016588906
7	89.60601347848626	Gymnastics	4.58786936236392
8	93.2348367029549	Diving	3.6288232244686367
9	96.65629860031103	Equestrianism	3.421461897356143
10	99.99999999999999	Water Polo	3.3437013996889577

	cumulative_percent	sport	percent_contribution
1	29.2177191328935	Fencing	29.2177191328935
2	43.44957587181904	Cycling	14.231856738925542
3	53.81715362865222	Athletics	10.367577756833176
4	62.20546654099906	Rowing	8.388312912346843
5	76.7200754005655	Equestrianism	7.257304429783224
6	76.7200754005655	Handball	7.257304429783224
7	83.41187558906692	Swimming	6.691800188501413
8	89.91517436380774	Shooting	6.503298774740811
9	95.38171536286524	Gymnastics	5.466540999057493
10	100.00000000000001	Judo	4.618284637134779

FRANCE

THE UNITED STATES

HYPOTHESIS 2- Majority of the medals are concentrated in a select few sports

	cumulative_percent	sport	percent_contribution
1	15.156017830609212	Rowing	15.156017830609212
2	28.677563150074292	Hockey	13.521545319465082
3	40.49034175334324	Canoeing	11.812778603268944
4	51.931649331352155	Athletics	11.441307578008916
5	63.22436849925706	Swimming	11.292719167904904
6	73.62555720653789	Equestrianism	10.401188707280832
7	81.35215453194651	Football	7.726597325408618
8	88.33580980683507	Cycling	6.983655274888559
9	94.27934621099554	Biathlon	5.943536404160475
10	100	Gymnastics	5.720653789004458

GERMANY

	cumulative_percent	sport	percent_contribution
1	25.61349693251534	Athletics	25.61349693251534
2	42.25460122699387	Rowing	16.641104294478527
3	55.21472392638037	Cycling	12.960122699386503
4	64.95398773006136	Swimming	9.739263803680982
5	74.00306748466258	Hockey	9.049079754601227
6	80.67484662576688	Shooting	6.671779141104295
7	87.0398773006135	Equestrianism	6.365030674846625
8	91.79447852760737	Sailing	4.754601226993866
9	96.08895705521473	Boxing	4.294478527607362
10	100.00000000000001	Gymnastics	3.9110429447852764

GREAT BRITAIN

HYPOTHESIS 2- Majority of the medals are concentrated in a select few sports

- The tables from the previous slides represent the top 10 sports of each country in terms of medals won.
- It also denotes the percentage contribution of each sport towards the total medal winning tally of the country.
- As we can see from the cumulative percentage values, even though there are 66 different sports played across the history of the Olympics, the top 10 sports constitute almost all the medals won by that country.
- Therefore, our second hypothesis is true. This is also verified by the fact that most number of participants from the country, fall under these top 10 sports.

HYPOTHESIS 3-Not much correlation between sports and events in terms of medal winners

- For this, we look at the participants in each sport and whether they won medals in multiple events.
- This hypothesis fails because events are categorised into sports based on factors like playing conditions, equipments required, etc.
- Therefore a sport like athletics may consist of events like 400m running, 100m running as well as shot-put, javelin throw etc.
- What this means is that even though they belong to the same sport, the skillset required to excel in these games is different.
- Therefore the more important relationship is between a player's physical attributes and event, more than sport and event itself.

Reason for up-down curve in medal history graph

- This trend is observed where the curve rises for medal tally in summer games while it drops for winter games.
- This is because the sports conducted depend on the weather.
- Games like ice hockey occur in winter while athletic events happen in summer.

	season	participation
1	Summer	42234
2	Winter	8331

- We can see that participation in summer games is 5x more than that of winter games.
- Combining with the fact that the percentage of winning for top countries is more or less the same, we can conclude that:

- The up-down trajectory is because of conducting different sports and number of participants.

Do both sexes excel in similar sports?

	sex	sport	medal_count
1	F	Swimming	625
2	F	Athletics	389
3	F	Rowing	178
4	F	Football	147
5	F	Basketball	131
6	F	Equestrianism	122
7	F	Gymnastics	120
8	F	Hockey	113
9	F	Ice Hockey	101
10	F	Fencing	78

	sex	sport	medal_count
1	M	Athletics	1280
2	M	Swimming	791
3	M	Rowing	665
4	M	Cycling	421
5	M	Fencing	402
6	M	Shooting	355
7	M	Equestrianism	310
8	M	Basketball	248
9	M	Gymnastics	243
10	M	Hockey	217

- As we can see, most of the top 10 sports in the list are similar.
- Thus we can say that both sexes do excel in similar sports.
- This is because both sexes are subjected to the same level of training programs in their respective countries.

TO SUMMARISE

HYPOTHESIS 1- The US won most medals because they sent most participants.

HYPOTHESIS 2- Majority of the medals are concentrated in a select few sports.

HYPOTHESIS 3- There is not much correlation between sports and events in terms of medal winning capabilities.

- There is an up-down trajectory in the medals won across history curve because of difference in seasonal games and participation.

- Irrespective of sex, most of the sports that contribute to top medal winning counts are similar.

THE STORY IS NOT OVER YET...

- There are a lot of other relationships left to explore.
- For example, looking at the top 10 sports of the countries, they have something to do with the popularity and cultural significance in the country.
- For example, basketball is a very popular sport in the US and it is also 1 of the top 10 sports in which the US has won bulk of its medals from.
- The same can also be said for cycling with respect to France.
- Better popularity, leads to better funding and allocation of resources in developing world class training programs that enable athletes to become champions.
- We can also do a historical analysis and the journey of a specific country, looking at their star athletes' performances, explore some interesting incidents and stats, etc.

RESOURCES USED

- Google Docs- To prepare project proposal document
- Google Slides- To prepare the presentation
- Mode Analytics- To run the SQL queries as well as to obtain graphs

THANK YOU !! :)