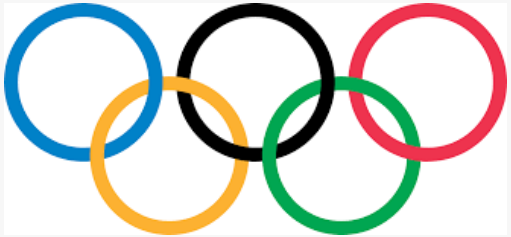


CODELYMPICS

- Mahir Dave, Parth Nanwani, Priyanka Chib,
Smriti Gangwani, Vishwa Thakkar, Zaid Dar

INDEX

01

PROJECT OVERVIEW

Broad overview of the project

02

DATA ACQUISITION & DATA CLEANING

Describe the data and data processing tasks

03

DATA ANALYSIS

Visualizations and modeling

04

CONCLUSION

Justifying the selections

01

PROJECT OVERVIEW

INTRODUCTION

- The Olympic Games, also known as the Games of the Olympiad, are a major international multi-sport event normally held once every four years (every two years before 1992)
- There are over 400 total events in 40 sports in the Olympic Tournament
- The project consists of processing and analyzing the dataset of the Olympics games from the year 1896 to 2016
- The data helped us to interpret which factors are the key to winning, while showing us trends and interesting details that were used to make inferences regarding the various fields of data.

MOST POPULAR SPORTS

ATHLETICS

Athletics is a versatile sport often played by both Male and Women consisting of sprints, hurdles, relays, races, marathons, etc. making it most played sports amongst women too.

SWIMMING/GYMNASTICS

Swimming is the most played sports for men and Gymnastics the most played for women

02

DATA ACQUISITION & DATA CLEANING

DATA ACQUISITION & DATA CLEANING

DATA ACQUISITION

1. Retrieved two datasets from www.sports-reference.com and acquired from kaggle.com

2. Data in the first set contains all the information categories like type of games, name of the player, physical characteristics of players, country representation, assigned team, games participated in, season, event and medal ranging from 1896 to 2016

3. The other dataset consists of the NOC and Region which is useful to determine what players play for which region

DATA CLEANING

1. Consistent data: Split the first set by column value and assigned the first column as the row value

2. Missing data: Utilized group sets and found mean of a biometric to replace NaN

3. For cases where NaN values could not be replaced, calculated mean of biometric value

4. For redundant data, dropping the notes column and created hierarchical index based on played ID and row Index

03

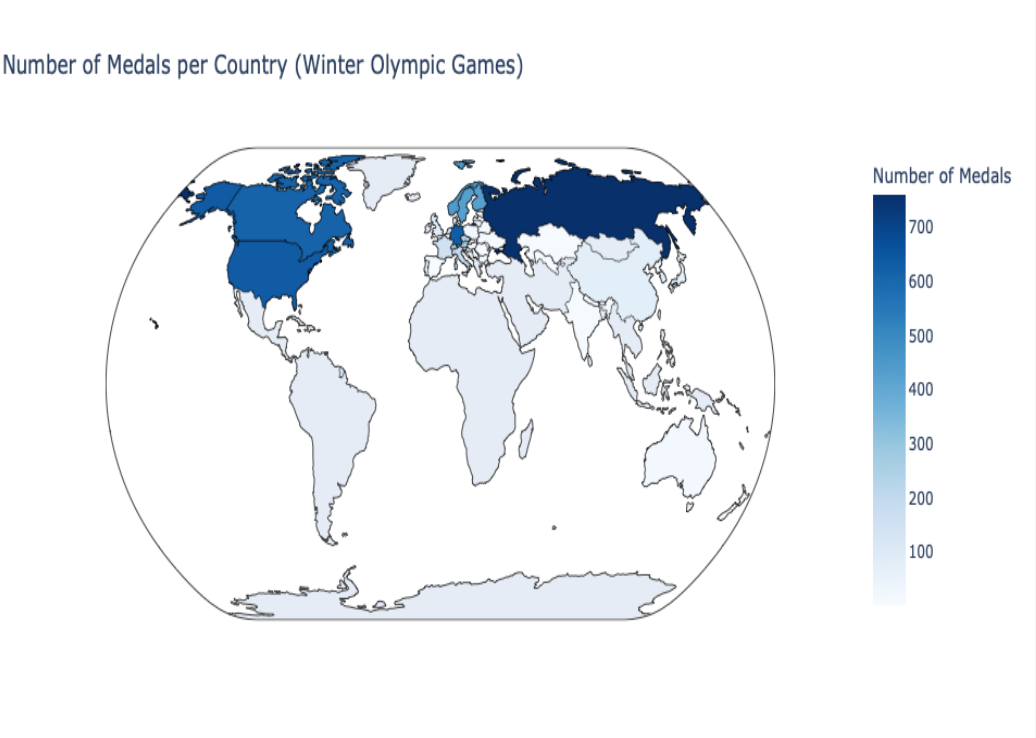
DATA ANALYSIS

IMPACT OF GEOGRAPHICAL LOCATION ON WINTER OLYMPICS

Countries with colder climates (Canada, Russia) perform much better than other countries

Athletes from countries with colder climates are able to train Winter sports at a much higher capacity than other countries (Skiing, Snowboarding)

We can expect countries with colder climates to perform well in future Winter Olympic games



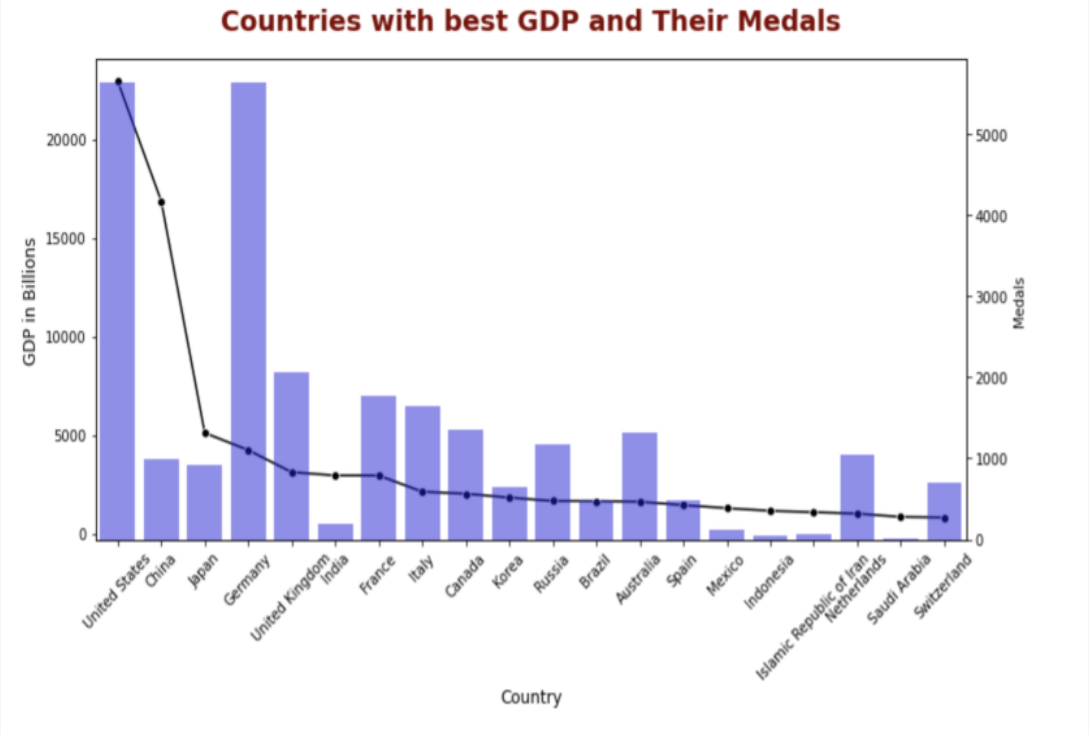
IMPACT OF GDP ON COUNTRIES PERFORMANCE

Countries with a higher GDP perform better in the Olympics

Richer countries can provide their athletes with the world class facilities for training

These countries can hire the best coaches and the best nutritionists along with providing better infrastructure which would assist for better training

These facilities do have an impact on the athletes performance as visible from the graph



IMPACT OF BMI ON WINNING FOR EACH SPORT FOR MALES

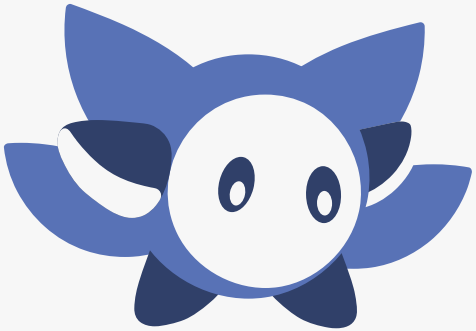
For sports that require more physical strength such as Weight lifting and Rugby, the players are found to be heavier in weight

The heavier a player, the more it outweighs the lighter weight player

We can further infer that for sports that require more flexibility and grace such as Gymnastics, the player needs to be lighter in weight

Sport	BMI of loser	BMI of winner
Tug-Of-War	28.841756	28.660159
Weightlifting	27.369212	28.095725
Bobsleigh	27.121967	27.248596
Judo	25.527043	26.057466
Rugby Sevens	25.312075	25.869383
Baseball	25.635023	25.742927
Wrestling	25.116385	25.341874
Ice Hockey	25.206615	25.244001
Shooting	24.683585	24.955228
Water Polo	24.752512	24.930969

Sport	BMI of loser	BMI of winner
Rhythmic Gymnastics	17.371665	16.885826
Synchronized Swimming	19.647038	19.769782
Ski Jumping	20.990095	20.177024
Triathlon	20.396695	20.319454
Figure Skating	20.820736	20.954077
Nordic Combined	21.520404	21.119267
Trampolining	21.339424	21.228316
Taekwondo	21.664081	21.365381
Diving	21.909241	21.467918
Gymnastics	21.507387	21.477622
Table Tennis	22.075338	21.739265



04

CONCLUSION

FACTORS AFFECTING WINNING IN OLYMPICS

WINTER SEASON

Countries with colder climates perform better in Winter Olympics

MORE PARTICIPANTS VS MORE WINNERS

Focus on quality of participants and NOT quantity of participants

THE GDP EFFECT

Countries with high GDP are better equipped to train and develop athletes

OTHER INTERESTING FINDINGS

POWER OF BMI

BMI parameter is partly correlated with the level of sports competence mostly in disciplines where the athletes need larger muscle mass

EFFECT OF POLITICS

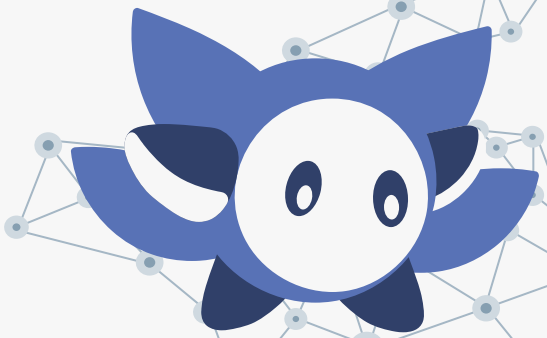
Olympic games have been canceled due to external circumstances (World War I, World War II) and have been boycotted by countries due to political reasons.

SUMMER SEASON

Climate did not have substantial effect on country performance in Summer Olympic Games

THANKS

Do you have
any questions?



REFERENCES

- https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results
- https://staff.math.su.se/hoehle/blog/2016/08/21/gapMedal.html
- https://knoema.com/jesoqmb/historical-gdp-per-capita-by-country-statistics-from-the-world-bank-1960-2018
- https://iopscience.iop.org/article/10.1088/1757-899X/1099/1/012058/pdf
- https://www.history.com/news/olympic-boycotts
- https://statisticstimes.com/economy/projected-world-gdp-ranking.php