INDEX

Abstract

Each country tries to give their best performance during the event. Despite a lot of hard work, many countries or players are unable to perform well during the events and grab medals whereas there are many countries that perform very well in the event and secure many medals. An analysis needs to be done by each country to evaluate the previous statistics which will detect the mistakes which they have done previously and will also help them in future development. Visualization of the data over various factors will provide us with the statistical view of the various factors which lead to the evolution of the Sporting Games and Improvement in the performance of various Countries/Players over time. The primary objective of this Research paper is to analyze the large Sport dataset using Exploratory Data Analysis to evaluate the evolution of the Sporting Games over the years. An analysis can also be done by the host country to find out the mistakes in the arrangements of the event which will help them in overcoming these mistakes and host the event accurately. This analysis will provide detailed and accurate information regarding various factors which lead to the evolution of the Sporting Games and the improvement of Countries/Players over time in a visual format.

Key Words: Sports Data Analysis, Classification, Linear Regression, Data Collection, Data Exploration

1. INTRODUCTION

This project proposes a handy visualization analysis of sporting games between 1896 to 2016. The primary goal of the dashboard is to explain how the user may benefit from the developed system. This application utilizes concrete analysis examples and claim to provide efficient, effective, functional, and convenient model for users. Sports analytics is the use of historical data and advanced statistics to measure performance, make decisions and predictions regarding performance and outcomes, in order to gain an advantage over competitors. Performance prediction is the commonest task in sports analytics. Sport analysts process data regarding players and teams with an intended goal the prediction of match results, tournament winners or team and individual player efficiency. Forecasts may be related to short–term or long– term events. For that reason, diverse methods and algorithms have been deployed. Clubs use sophisticated devices and software (i.e. GPS tracking systems) to gather and analyze data generated by players during training sessions and matches. They process these data to use for short– term decision making and long–term organization development. Also, extensive analysis of all data available is a prerequisite for betting companies. Finally, fans are also very interested in advanced statistics and how they affect football. For all the above reasons, the use of sports analytics has increased during the last few years. Football was selected for our research because of the abundance of statistical categories and historical data, its fame, as well as the simplicity of its rules and of national championships formats. On the other hand, there are special difficulties, which make football long–term prediction challenging. The abundance of online data regarding football is an asset, but requires filtering and proper data for team and player performance prediction. Unfortunately, this is not always easy. Additionally, team and player performance can be affected by incidents not depicted in the data collected; a team is rated higher than it should be when their opponents underperform. A player might have a low rating performance when coming into action after a serious injury.

2. LITERATURE REVIEW

Kabita Paul, Elif Demir, Anjali Bapat (May 2019):

They proposed system which uses data visualization technique to create application which utilizes concrete analysis examples and claim to provide efficient, effective, functional, and conventional model for users. And explain how the user benefits from developed system experiencing visual representation. It cannot predict the winning probability. It cannot find constellation of winning medal with their factors.[1]

Sacha Schmidt, Limas, Wunderlich, Dominik Schreger (December 2020):

They proposed system which uses Random Forest Algorithm to forecast and provide accurate results for future predictions. It forecasts the number of Olympic medals for each nation. Various factors are taken into consideration to benchmark the performance of their teams and evaluate drivers to success. They are trying to improve performance model in the future. Novel approach for missing data is missing. [2]

Rahul Pradhan (January 2021):

They proposed a system which exploratory data analysis for analyzing and visualizing factors and correlation between factors. It aims to analyze vast history of Olympic Games and determine the evolution of Olympic Games. To determine factors and perform comparative study on these factors. They are trying to visualize information in other data format. Machine learning is to be explored.[3]

3. METHODOLOGY

Data Collection: -

Sports are played globally thus data collected from it is in interest of world map. Collecting data from different sources like official websites of Sports association, Wikipedia and different media outlets and converting the raw data into datasets for exploration. While dealing with Sports related data, we have to deal with sports related costs, deep game analytics of all games all the year.

The dataset contains 271116 rows and 15 columns. Each row corresponds to an individual athlete competing in an individual Olympic event (athlete-events). The columns are:

ID - Unique number for each athlete

Name - Athlete's name

Sex - M or F

Age - Integer

Height - In cm

Weight - In kilograms

Team - Team name

NOC - National Olympic Committee 3-letter code

Games - Year and season

Year - Integer

Season – summer/winter

City - Host city

Sport - Sport

Event - Event Medal **-** Gold, Silver, Bronze, or NA

Data Exploration: -

The collected data sets are structured and indexed using Microsoft Excel. After importing the collected data from data sets, two or more data sets are merged together depending on the user’s requirements. Redundant data may negatively affect the usefulness; therefore, transformation is carried out to remove it.

Data Visualization: -

User friendly dashboard is created for user interaction. Depending on user requirement, specific data is provided in a more convenient and user-friendly manner. Visualization provides rapid, simple and detailed information in several ways as per user requirement.

4. SYSTEM ARCHITECTURE

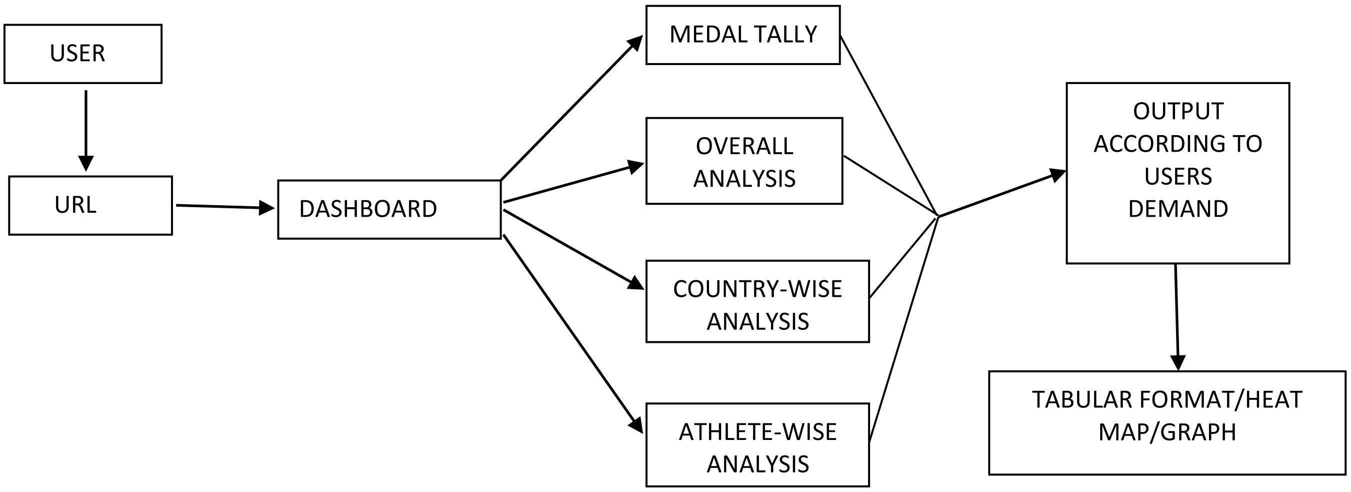


Fig 4.1: Architecture

The web application is hosted via a Streamlit interface and is accessible through a URL. The landing page presents a dashboard featuring four primary options: Medal Tally, Overall Analysis, Country-wise Analysis, and Athlete-wise Analysis. Users interact with the system by selecting one of these options. The application then generates outputs displayed in various formats including tables, heat maps, and graphs based on the user's input.

5. FRAMEWORK

The dataset is about each the number of medal and country in relationship to sporting events. These variables have two relationships are: Total GDP (Gross domestic product) = Population × Gross domestic product per capita.

Population Density = Population ÷ Country Area

We are using Linear Regression algorithm. Linear regression is a linear approach to modelling the relationship between a dependent variable and one or more independent variables. The case of one explanatory variable is called simple linear regression, so we can use this regression model to predict the Y when the X is known,
This mathematical equation can be generalized as

$Y = \beta0 +\beta1X +\varepsilon$

Where $\beta0$ is the intercept and $\beta1$ is the slope,

So, they are called regression co-efficient. ε is the error term, the part of Y the regression models are unable to explain.

On the other hand, the case of some explanatory variable is called Multiple Linear Regression, then the mathematical equation can be generalized as

$Y = \beta0 +\beta1X1 +\beta2X2 +\beta3X3 + \cdots+\beta PXp +\varepsilon$

6. Code

# 1. Import Libraries

```
import plotly.express as px
import streamlit as st
import pandas as pd
import preprocesser, helper
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.figure_factory as plf
import numpy as np
```

# 2. Read Data

```
df = pd.read_csv('athlete_events.csv')
df_region = pd.read_csv('noc_regions.csv')
```

# 3. **[Exploratory Data Analysis](#data_preparation)**

```
df =df[df['Season']=='Summer']
df =df.merge(df_region,on='NOC',how='left')
df['region'].unique().shape
df.isnull().sum()
df.duplicated().sum()
df.drop_duplicates(inplace=True)
df['Medal'].value_counts()
df =pd.concat([df,pd.get_dummies(df['Medal'])],axis=1)
df.head()
```

# 4. Medal Tally

```
def fetch_medal_tally(df,year,country):
    medal_tally =
df.drop_duplicates(subset=['Team','NOC','Year','City','Sport','Medal','Event',])
    flag =0
    if year=='Overall' and country == "Overall":
        temp_df = medal_tally
    elif year != 'Overall' and country == 'Overall':
        temp_df = medal_tally[medal_tally['Year'] == int(year)]
    elif year == 'Overall' and country != 'Overall':
        flag =1
        temp_df = medal_tally[medal_tally['region'] == country]
    elif year != 'Overall' and country != 'Overall':
        temp_df = medal_tally[(medal_tally['Year'] == int(year)) &
(medal_tally['region']==country)]
    if flag ==1:
        x=
temp_df.groupby('Year').sum()[['Gold',"Silver",'Bronze']].sort_values('Year',ascendin
g=True).reset_index()
    else:
        x=
temp_df.groupby('region').sum()[['Gold',"Silver",'Bronze']].sort_values('Gold',ascendi
ng=False).reset_index()
    x['total']=x['Gold']+x['Bronze']+x['Silver']
    return x
```

# 5. Filter Data

```
nations_over_time =
df.drop_duplicates(['Year','region'])['Year'].value_counts().reset_index()
nations_over_time.rename(columns={'Year':'year','count':'n_country'},inplace=True)
```

# 6. Athlete-wise Data

```
temp_df =df.dropna(subset=['Medal'])
temp_df = temp_df[temp_df['region'] == "USA"]
temp_df
=temp_df['Name'].value_counts().reset_index().head(15).merge(df,on='Name',how='le
ft')[['Name','Sport','count']].drop_duplicates('Name')
```

```
temp_df.rename(columns={'count':'Medal'},inplace=True)
temp_df.reset_index()
temp_df



def most_successful(df,Country):
    temp_df =df.dropna(subset=['Medal'])

    temp_df = temp_df[temp_df['region'] == Country]
    temp_df =temp_df['Name'].value_counts().reset_index().head(15).merge(df,left_on='index',right_on='Name',how='left')[['index','Name_x','Sport']].drop_duplicates('index')
    return temp_df
most_successful(df,'India')




temp_df =df.drop_duplicates(['Name','region'])
x1 =temp_df['Age'].dropna()
x2= temp_df[temp_df['Medal']=='Gold']['Age'].dropna()
x3= temp_df[temp_df['Medal']=='Silver']['Age'].dropna()
x4= temp_df[temp_df['Medal']=='Bronze']['Age'].dropna()
fig =ff.create_distplot([x1,x2,x3,x4],['Age Distribution','Gold Medalist','Silver Medalist','Bronze Medalist'],show_hist=False,show_rug =False)
fig.show()

temp_df =df.drop_duplicates(['Name','region'])
temp_df = temp_df[temp_df['Sport']=='Athletics']
temp_df['Medal'].fillna('No Medal',inplace=True)
plt.figure(figsize=(10,10))
sns.scatterplot(temp_df['Weight'],temp_df['Height'],hue=temp_df['Medal'],style=temp_df['Sex'],s=100)
```

```
plt.show()
```

# 7. Male-Female Comparison

```
temp_df =df.drop_duplicates(['Name','region'])
men = temp_df[temp_df['Sex']=='M'].groupby('Year').count()['Name'].reset_index()
women = temp_df[temp_df['Sex']=='F'].groupby('Year').count()['Name'].reset_index()
final =women.merge(men,on='Year',how='right')
final.fillna(0,inplace=True)
final.rename(columns={'Name_x':'Female','Name_y':'Male'},inplace=True)
```

# 8. Medal category wise Analysis

```
temp_df =df.drop_duplicates(['Name','region'])
x1 =temp_df['Age'].dropna()
x2= temp_df[temp_df['Medal']=='Gold']['Age'].dropna()
x3= temp_df[temp_df['Medal']=='Silver']['Age'].dropna()
x4= temp_df[temp_df['Medal']=='Bronze']['Age'].dropna()
fig =ff.create_distplot([x1,x2,x3,x4],['Age Distribution','Gold Medalist','Silver
Medalist','Bronze Medalist'],show_hist=False,show_rug =False)
fig.show()
```

# 9. Overall Analysis

```
if user_menu == 'Overall Analysis':
    st.title('Top Statistics')
    athletes = df.Name.unique().shape[0]
    country = df.region.unique().shape[0]
    events = df.Event.unique().shape[0]
    sports = df.Sport.unique().shape[0]
    cities = df.City.unique().shape[0]
    editions = df.Year.unique().shape[0] - 1

    col1, col2, col3, = st.columns(3)
    with col1:
        st.header('Editions')
        st.title(editions)
    with col2:
        st.header('Hosts')
        st.title(cities)
    with col3:
        st.header('Sports')
        st.title(sports)
    col1, col2, col3, = st.columns(3)
    with col1:
        st.header('Events')
        st.title(events)
    with col2:
        st.header('Nations')
        st.title(country)
    with col3:
        st.header('Athletes')
        st.title(athletes)
```

# 10. National over the year Analysis

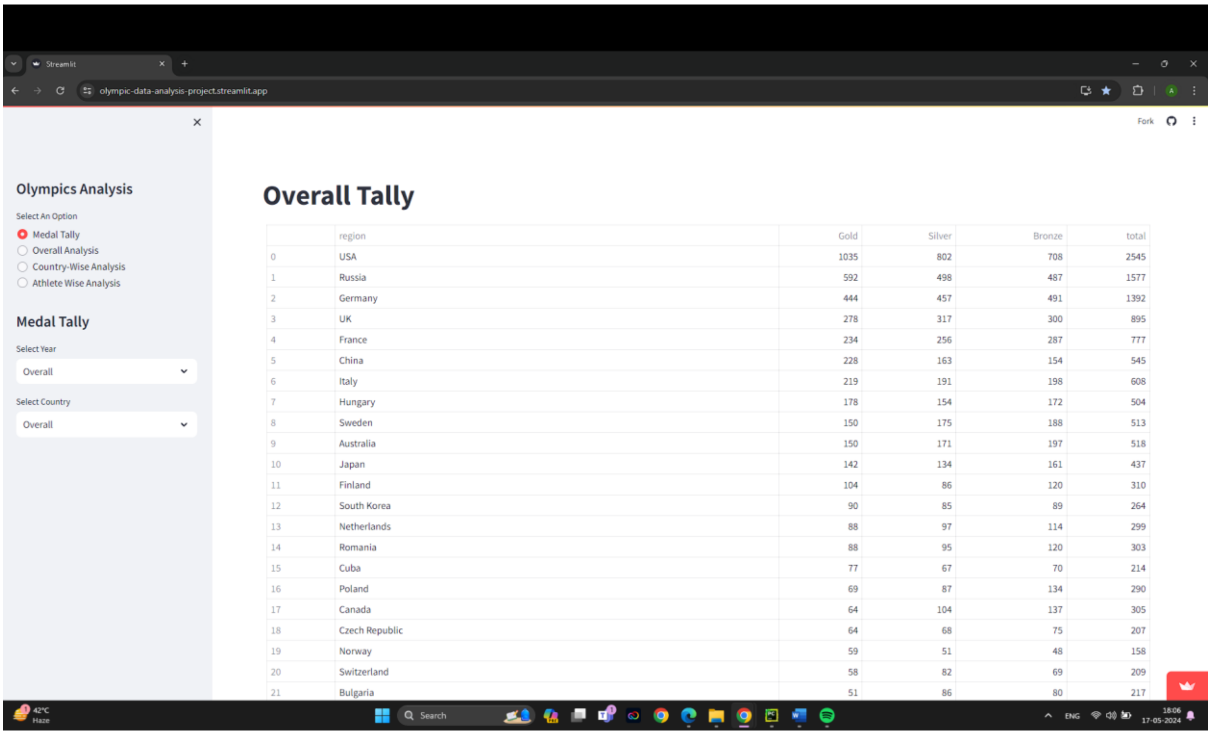
```
    st.title('Participating Nations Over The Year')
    nations_over_time_df = helper.data_over_time(df, 'region', 'countries')
    fig = px.line(nations_over_time_df, x='year', y='countries')
    st.plotly_chart(fig)
    st.title('Events Over The Year')
    nations_over_time_df = helper.data_over_time(df, 'Event', 'events')
    fig = px.line(nations_over_time_df, x='year', y='events')
    st.plotly_chart(fig)
    st.title('Athlete Over The Year')
    nations_over_time_df = helper.data_over_time(df, 'Name', 'athletes')
    fig = px.line(nations_over_time_df, x='year', y='athletes')
    st.plotly_chart(fig)
    st.title('No Of Events Overtime Of Sports')
    fig, ax = plt.subplots(figsize=(25, 20))
    x = df.drop_duplicates(['Year', 'Sport', 'Event'])
    ax = sns.heatmap(
        x.pivot_table(index='Sport', columns='Year', values='Event',
aggfunc='count').fillna(0).astype(int), annot=True)
    st.pyplot(fig)
    st.title('Most Successful Players')
    sport_list = df['Sport'].unique().tolist()
    sport_list.sort()
    sport_list.insert(0, 'Overall')
    selected_sport = st.selectbox('Select A Sport', sport_list)
    st.table(helper.most_successful(df, selected_sport))
```

# 11. National over the year Analysis

```
famous_sports = ['Basketball', 'Judo', 'Football', 'Tug-Of-War', 'Athletics',
                     'Swimming', 'Badminton', 'Sailing', 'Gymnastics',
                     'Art Competitions', 'Handball', 'Weightlifting', 'Wrestling',
                     'Water Polo', 'Hockey', 'Rowing', 'Fencing',
                     'Shooting', 'Boxing', 'Taekwondo', 'Cycling', 'Diving', 'Canoeing',
                     'Tennis', 'Golf', 'Softball', 'Archery',
                     'Volleyball', 'Synchronized Swimming', 'Table Tennis', 'Baseball',
                     'Rhythmic Gymnastics', 'Rugby Sevens',
                     'Beach Volleyball', 'Triathlon', 'Rugby', 'Polo', 'Ice Hockey']
    x = []
    name = []
    for sport in famous_sports:
        temp_df = player_df[player_df['Sport'] == sport]
        x.append(temp_df[temp_df['Medal'] == 'Gold']['Age'].dropna())
        name.append(sport)
    fig = ff.create_distplot(x, name, show_hist=False, show_rug=False)
```

```
fig.update_layout(autosize=False, width=880, height=600)
st.header('Distribution Of Age With Respect To Sport Who Won Gold Medal')
st.plotly_chart(fig)
```

7. RESULTS



Olympics Analysis

Select An Option

● Medal Tally
○ Overall Analysis
○ Country-Wise Analysis
○ Athlete Wise Analysis

Medal Tally

Select Year: Overall

Select Country: Overall

Overall Tally

	region	Gold	Silver	Bronze	total
0	USA	1035	802	708	2545
1	Russia	592	498	487	1577
2	Germany	444	457	491	1392
3	UK	278	317	300	895
4	France	234	256	287	777
5	China	228	163	154	545
6	Italy	219	191	198	608
7	Hungary	178	154	172	504
8	Sweden	150	175	188	513
9	Australia	150	171	197	518
10	Japan	142	134	161	437
11	Finland	104	86	120	310
12	South Korea	90	85	89	264
13	Netherlands	88	97	114	299
14	Romania	88	95	120	303
15	Cuba	77	67	70	214
16	Poland	69	87	134	290
17	Canada	64	104	137	305
18	Czech Republic	64	68	75	207
19	Norway	59	51	48	158
20	Switzerland	58	82	69	209
21	Bulgaria	51	86	80	217

Fig 7.1: Overall Tally

The overall medal table provides a comprehensive overview of the country's performance at every Olympic Games from its inception in 1896 to 2016. This historical data shows the total gold medals, silver and bronze medals won by all participating countries. This is a testament to the country's success and consistency in various Olympic Games over the years. This joint record demonstrates the passion, skill and determination of athletes, as well as the sports cooperation between countries on the world's biggest stage. This wealth of information provides a unique perspective on the evolution of international sports and the history of Olympic success.

The analysis compares countries' performance based on the number of medals won by athletes from selected countries at the Olympic Games from 1896 to 2016. The United States, Hungary, France, Japan, Australia and other countries are selected for review.

The following results were obtained from the analysis:

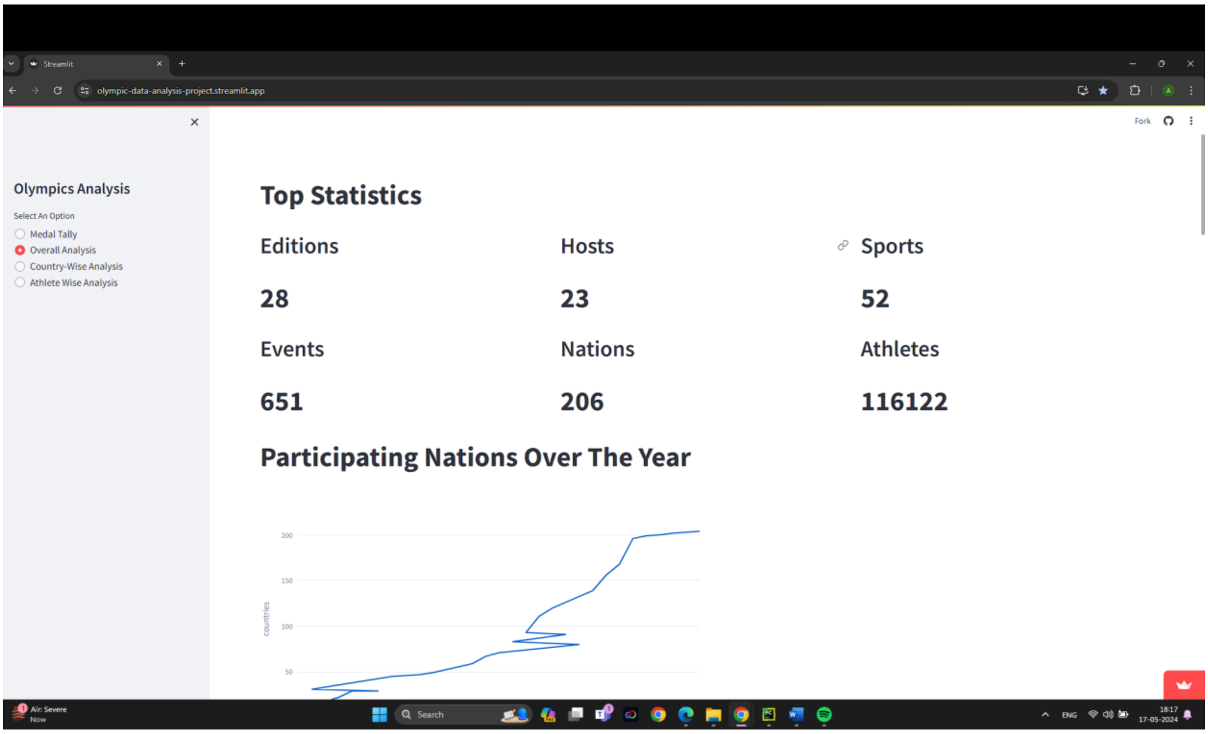
(i) Among the five countries participating in the 1996 Olympic Games, the United States ranked first with 7.53% participation, followed by Australia with 2.25% and Japan with 1.61%. Hungary and France rank first with 0.75%. The country with the lowest rate is 0.69 percent.

(ii) In the 2000 Olympic Games, the United States led with 6.55% participation, followed by Australia with 4.019%, France with 1.58% and Hungary and Japan with 0.94%.

(iii) In the 2004 Olympic Games, the United States came first with a contribution of 8.05%, followed by Australia with 3.3%, Japan with 2.1%, France with 1.6% and Hungary with at least 0.9%. has done.

(iv) In the 2008 Olympic Games, the United States led with participation of 8.52%, followed by Australia with 3.72%, France with 1.71%, Japan with 1.42% and the lowest participation with 0.88%. Hungary followed.

(v) In the 2016 Olympic Games, the USA came first with a contribution of 8.5%, Australia came second with a contribution of 3.01%, Japan 1.8%, France 1.6% and Hungary with the minimum contribution and maximum 0. 9%.



Olympics Analysis

Select An Option

○ Medal Tally

● Overall Analysis

○ Country-Wise Analysis

○ Athlete Wise Analysis

Top Statistics

Editions	Hosts	Sports
28	23	52
Events	Nations	Athletes
651	206	116122

Participating Nations Over The Year

Fig 7.2: Top Statistics

The statistics above give an overview of the differences between the Games. These statistics include the number of host countries, the total number of specific sports, the number of events held and the diversity of sports types participating in the event. A fascinating moment in the rich history of the Olympic Games in terms of growth, participation and global significance. The document summarizes the development of the Games and shows how they evolved into different sports, countries and sports.

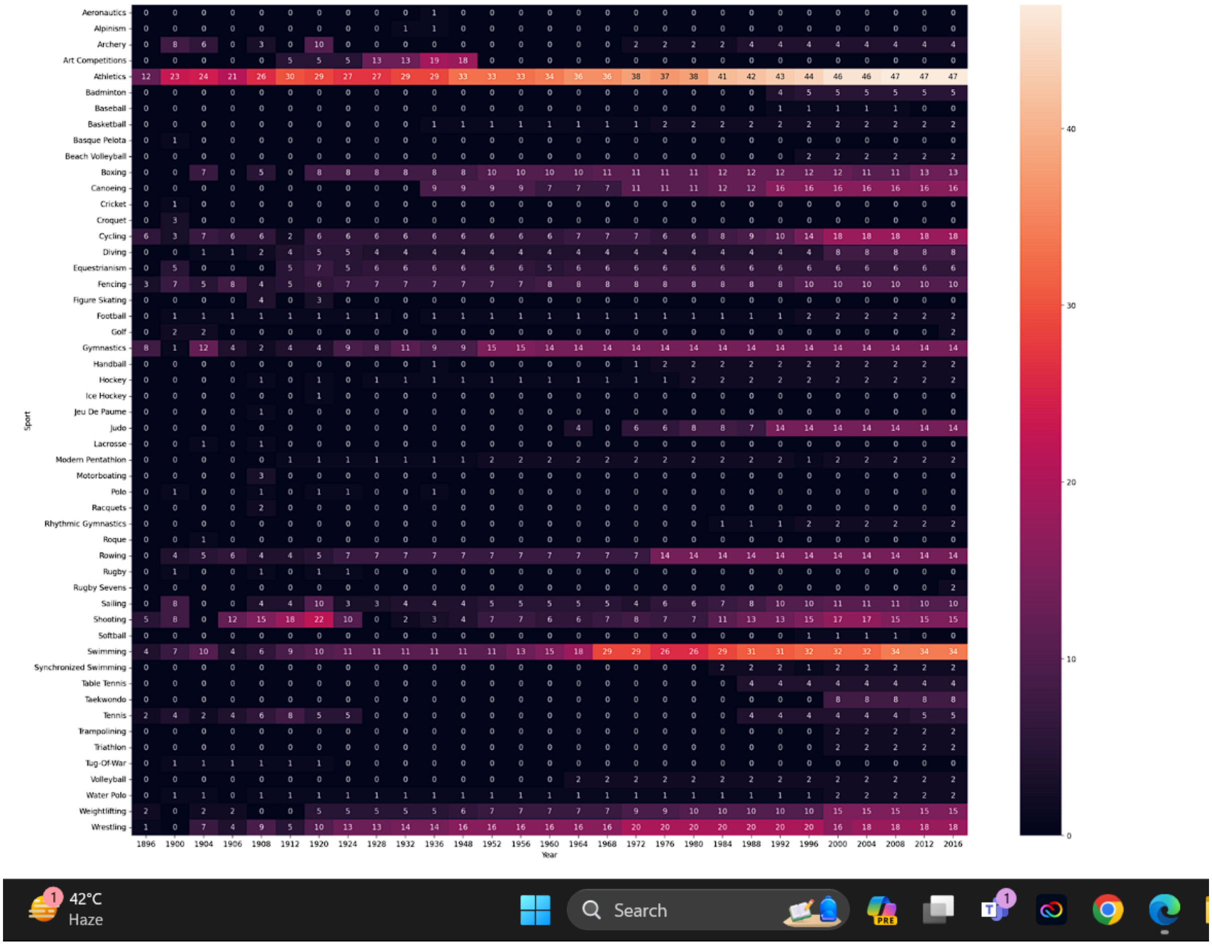


Fig 7.3: Heat map

The above Heat map shows the number of events that took place over time for every sport played at the Olympic Games. Suppose for wrestling, in 1896 only one event took place whereas in 1904 seven events took place. For athletics in 1896 only twelve events took place whereas in 2016 a total number of forty-seven events took place. This showed the rise in the popularity of that sport over the years.

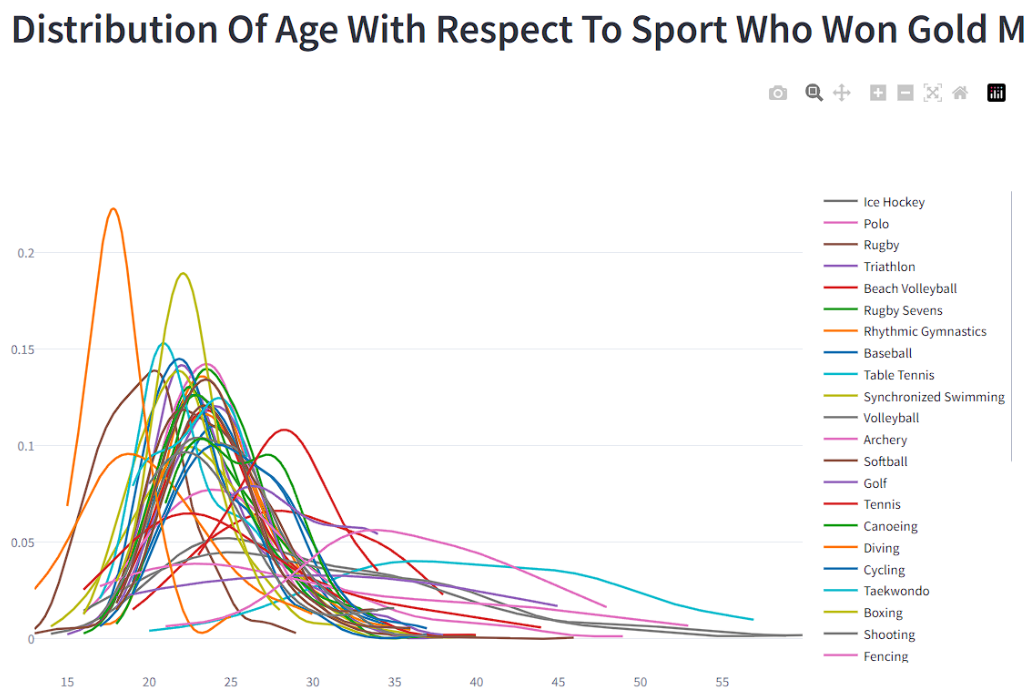


Fig 7.4: Age Distribution

The age group of the sport is a graph showing the number of athletes in different age groups participating in different sports. Here the coloured lines represent different types of movement. This information helps understand the interests of athletes of different ages and helps identify potential areas for recruiting or retaining athletes.

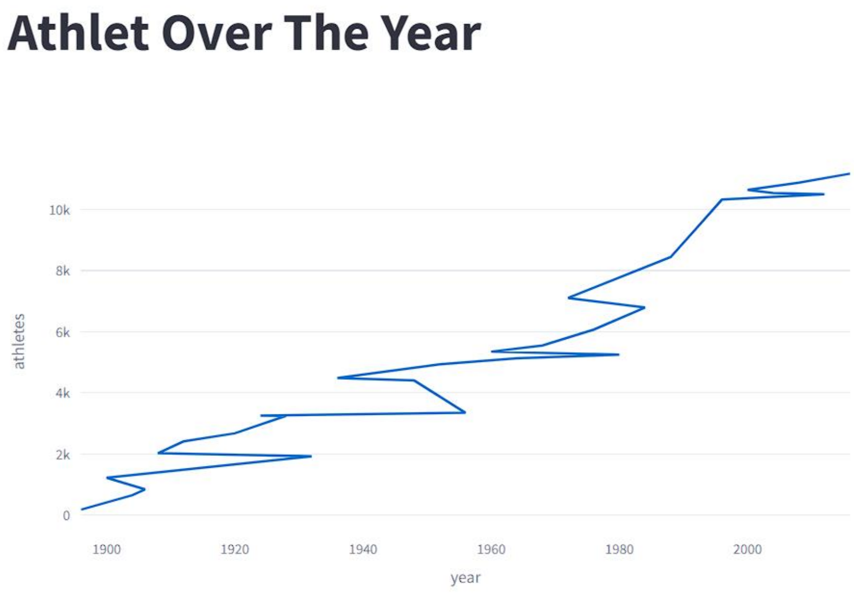


Fig 7.5: Athletes over the years

A country's performance in the Olympics can be measured by the number of medals it received. An analysis of how well countries performed at the Olympics from 1992 to 2016. This can be solved by counting the total number of medals won by a particular country in a year from 1992 to 2016. Data visualization can be done to represent the results of a particular country.

The results are as follows:

(i) India's performance improved from no awards in 1992, 1 award in 1996 and finally 6 awards in 2016.

(ii) US performance showed a zigzag pattern from 1992 (220 awards) to 1996. The number of awards dropped sharply in 2000, and the awards numbered 240. It gradually increased from 2004 onwards, with the best contribution being made in 2008 with 350 medals.

(iii) The performance of the French team increased with awards from 1996 to 2008. Around 40. It was successful in 2016 and won 80 awards.

(iv) Australia performed better in the 1992 Olympics, winning 60 medals. In 2000, Australia's performance skyrocketed, winning nearly 200 awards. From 2004 to 2016, its performance gradually declined.

(v) At first the Japanese team's results were poor, but between 2000 and 2004 the results improved greatly, winning 100 more awards than any other country.

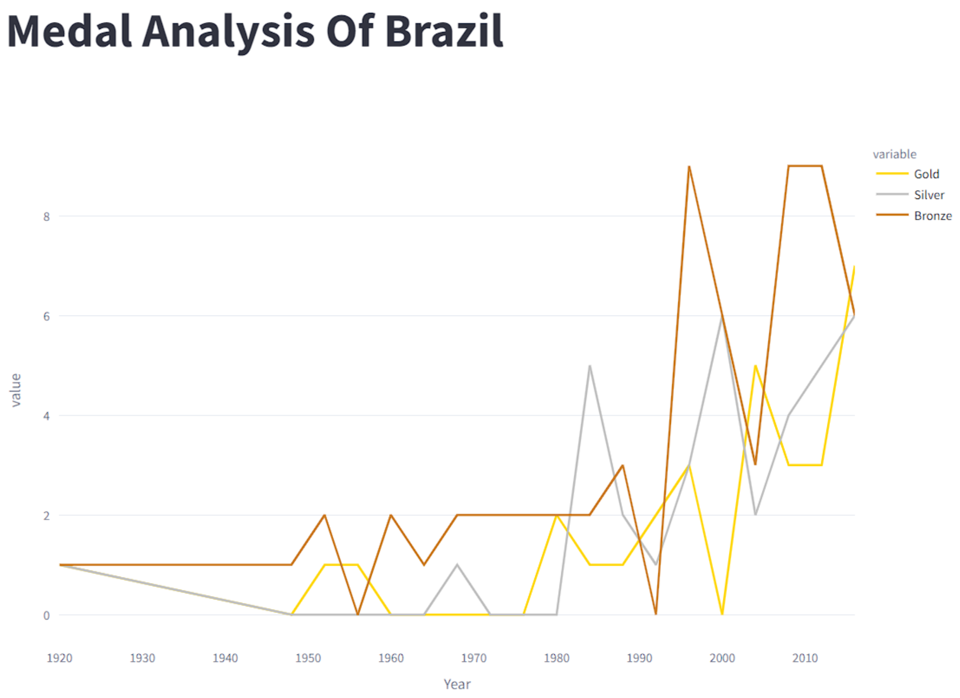


Fig 7.6: Country-wise Medal Tally

Due to the fluctuating medal counts across diverse editions, it was necessary to create visual representations for each country to accurately capture the variations in their respective performance over time. So, we conducted comprehensive country-based analysis, generating distinct line graphs depicting the varying numbers of medals won by each country over different editions.

In this analysis, it is possible to determine all the medals won by participants from all countries in the Olympic Games between 1896 and 2016, by the number of individuals who contributed individually or as a team or success of the country.

The analysis provided the following results:

(i) The USA has won the most gold medals with almost equal number of silver and bronze medals compared to other countries.

(ii) Compared to other tournaments, Australia won the fewest gold medals but the most medals. Japan won fewer gold medals than other countries. France has fewer gold medals and more silver and bronze medals.

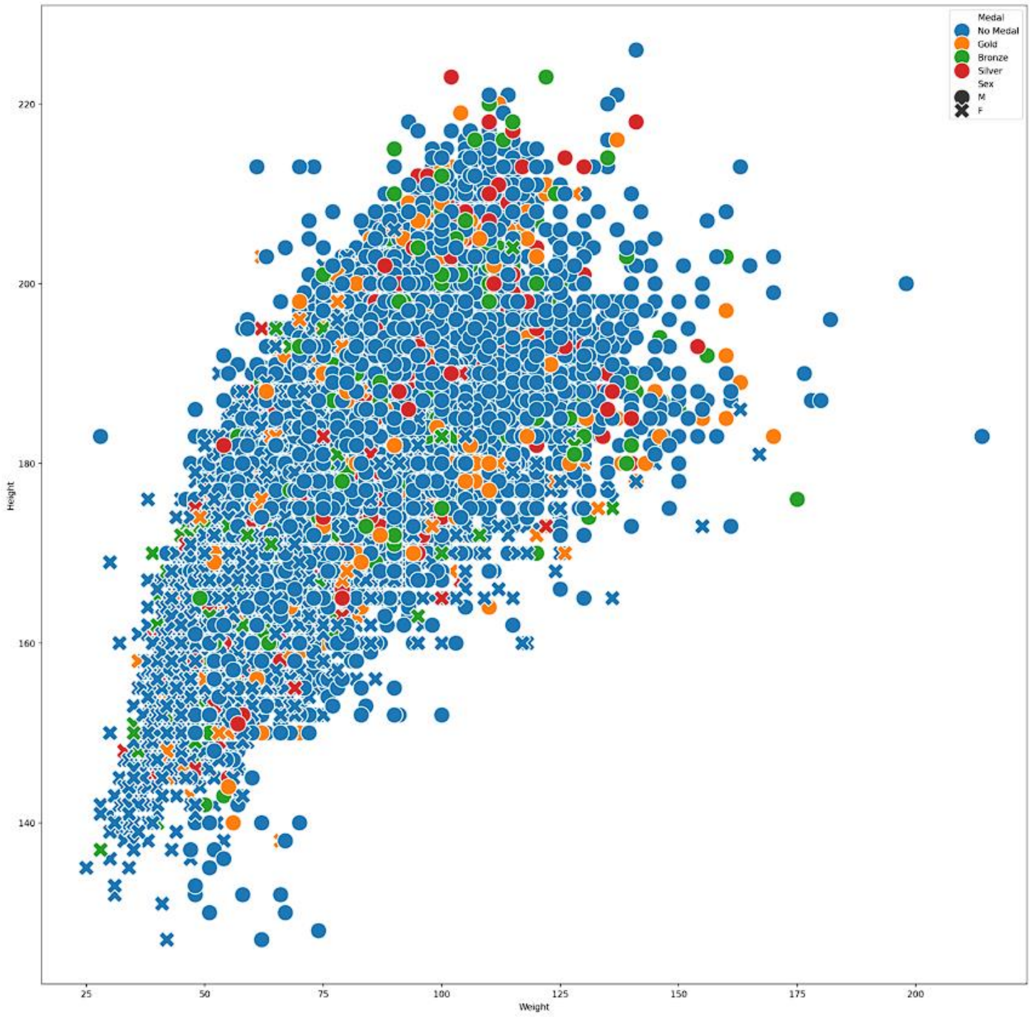
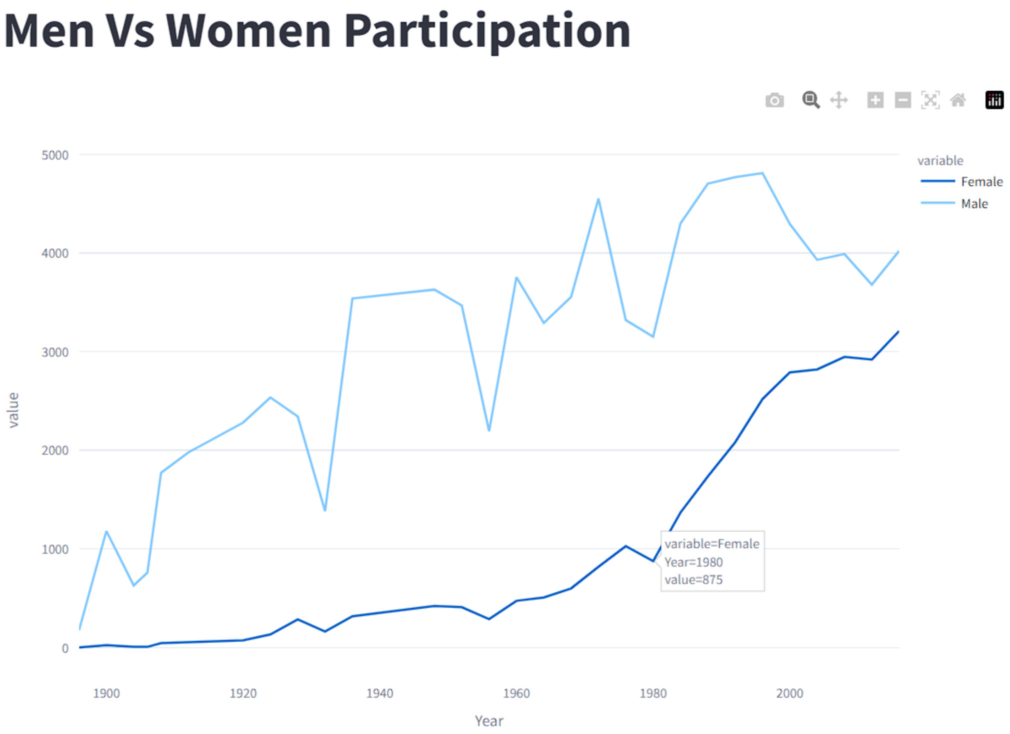


Fig 7.7: Height VS Weight

From the Height VS Weight analysis, we concluded that Most Females who have won the medal are between 160-180cm tall and their weight class wide ranging from 50-150kg. The Number of Gold winners has performed well irrespective of their weight but seen density high at range of 175-180cm height.



Men Vs Women Participation

Fig 7.8: Participation of Men VS Women over the years

The above figure shows the gender distribution of athletes in the Olympic Games All male and female participants in the Olympic Games from 1896 to 2016 were analysed to obtain the ratio of male and female participants. Analysis shows that men are more productive than women around the world.

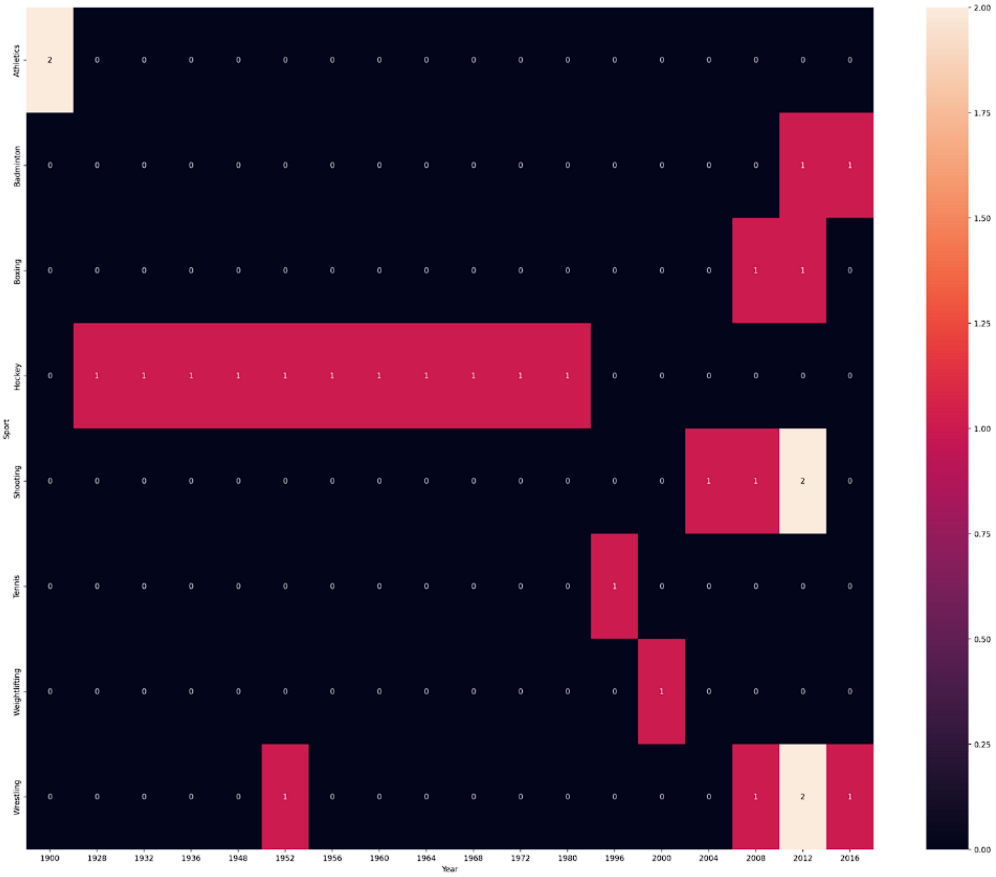


Fig 7.9: Country-wise Sports Analysis

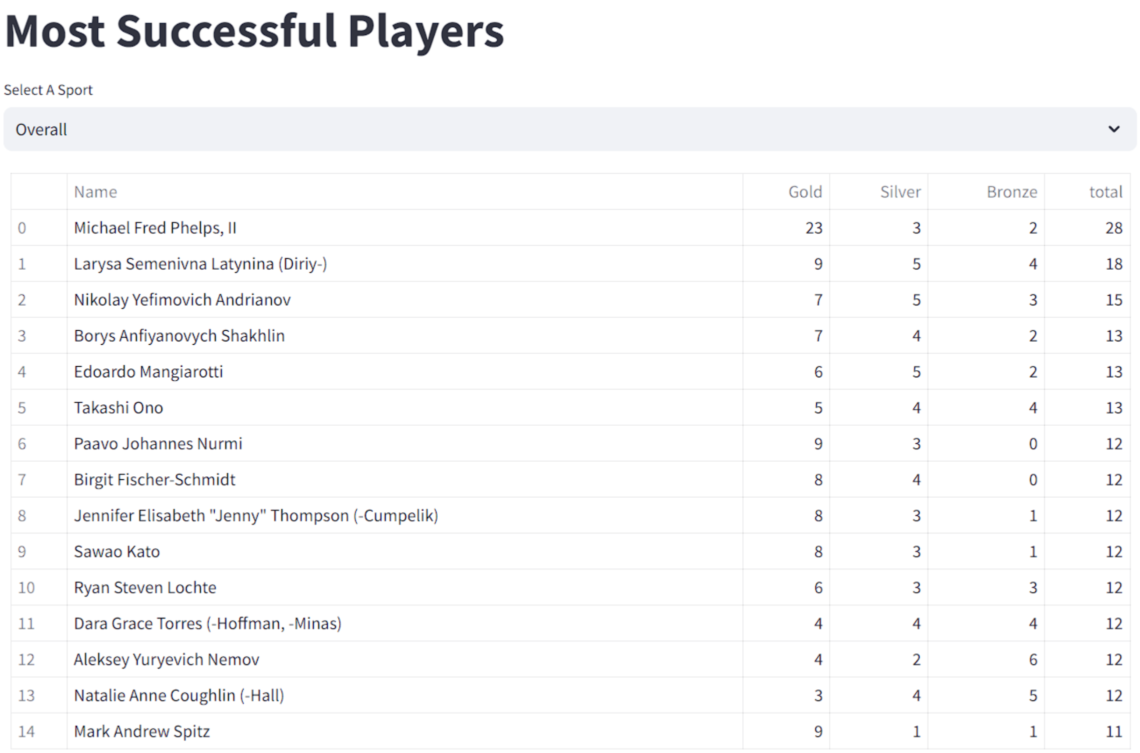
The analysis represents the performance of participating countries and top sports at the Olympics between 2000 and 2016. Analyze a country's sports field in a given year and det ermine which sports facilities should be more involved. This provides information to improve their future participation in the Olympics.

A. In 2000, the United States was the best country in swimming and the worst in football.

B. In 2000, Australia performed best in swimming and worst in gymnastics.

C. In 2000, France had the best performance in wrestling and the worst performance in tennis.

D. In 2000, Australia per for med best in the ocean and worst in sports.



Most Successful Players

Select A Sport

Overall

	Name	Gold	Silver	Bronze	total
0	Michael Fred Phelps, II	23	3	2	28
1	Larysa Semenivna Latynina (Diriy-)	9	5	4	18
2	Nikolay Yefimovich Andrianov	7	5	3	15
3	Borys Anfiyanovych Shakhlin	7	4	2	13
4	Edoardo Mangiarotti	6	5	2	13
5	Takashi Ono	5	4	4	13
6	Paavo Johannes Nurmi	9	3	0	12
7	Birgit Fischer-Schmidt	8	4	0	12
8	Jennifer Elisabeth "Jenny" Thompson (-Cumpelik)	8	3	1	12
9	Sawao Kato	8	3	1	12
10	Ryan Steven Lochte	6	3	3	12
11	Dara Grace Torres (-Hoffman, -Minas)	4	4	4	12
12	Aleksey Yuryevich Nemov	4	2	6	12
13	Natalie Anne Coughlin (-Hall)	3	4	5	12
14	Mark Andrew Spitz	9	1	1	11

Fig 7.10: Most Successful Athletes

This report shows the most successful athletes of the year.

We see that the number of countries and regions participating in the Olympics increases with each edition. There is also a very high increase in many sports participating in the Olympic Games.

In the table we present the best athletes in Olympic history, which shows us that Michael Fred Phelps of USA has the most medals in the history of Olympic swimming, his baton counts to a total of 28. Users can find the best athletes in a particular sport in the app itself. We saw that the chance of winning a medal is directly proportional to the number of athletes participating in the Olympics. For example, Team USA has won the most awards so far and has a large number of participants.

8. CONCLUSIONS AND FUTURE WORK

The main objective was to analyze the various factors which have contributed to the Evolution of the Olympic Games over the years. In the country-wise analysis, we saw that the country's performance changes over time and is affected by the sports it participates in. In athlete-wise analysis, we see that an athlete's individual performance is not always related to the overall success of the country. This study provides a comprehensive analysis of Olympic sports data, shedding light on the historical development of the Olympics and providing context for the performance of various countries over the years. Our findings have important implications for researchers, sports organizations, and policymakers seeking to understand the dynamics of the Olympic Games and participating countries. We have visualized our data in Graphical format and also performed Correlation Analysis on the data set to analyze the relationship between two continuous independent variables.

We have exclusively used Exploratory Data Analysis for Data Analysis. In the future, we may apply several Machine Learning Algorithms to the data set and develop a Predictive Model that can forecast the statistics of future Olympic and Asian Games
Our Project contains some imperfections and weakness. We plan to overcome some of these weaknesses in future ans see these limitations as future scope. These are:

● No Prediction is done - We have used the data related to Olympics and analysed it thoroughly but have not predicted anything. So we can feed this analysed data to Machine Learning Algorithms to Predictive something related to the same.

● We have only added information regarding Olympics in Sports Snooze, other major sporting events like Common wealth games, Asian games etc can be made a part of Sports Snooze.

● We can update live scores during the events itself, addition of text commentary features etc.

9. REFERENCES

[1] Kabita Paul, Elif Demir, Anjali Bapat: Olympic Data Analysis Project (May 2019)

[2] Sacha Schmidt, Limas, Wunderlich, Dominik Schreger (December 2020): Olympic Data Analysis Project

[3] Pradhan, Rahul & Agrawal, Kartik & Nag, Anubhav. (2021). Analysing Evolution of the Olympics by Exploratory Data Analysis using Python.

[4] http://www.sports-reference.com

Annexure I

Responsibility Chart

Roll No.	**Name**	**Responsibilities**
2201560049	**Sandeep Kumar**	**Fronted, Deployment and Research**
2201560033	**Rohit Yadav**	**Medal Tally and Analytics**
2201560045	**Sparsh Garg**	**Overall Analysis**
2201560044	**Jeetu Poswal**	**Athlete-wise Analysis**