

Data-Driven insights on Olympic Sports Participation and Performance for data analytics

Team:

Chintada Vamsimohan

Chukka Saiswaroop

Danda Sampathvenakatagovindaaditya

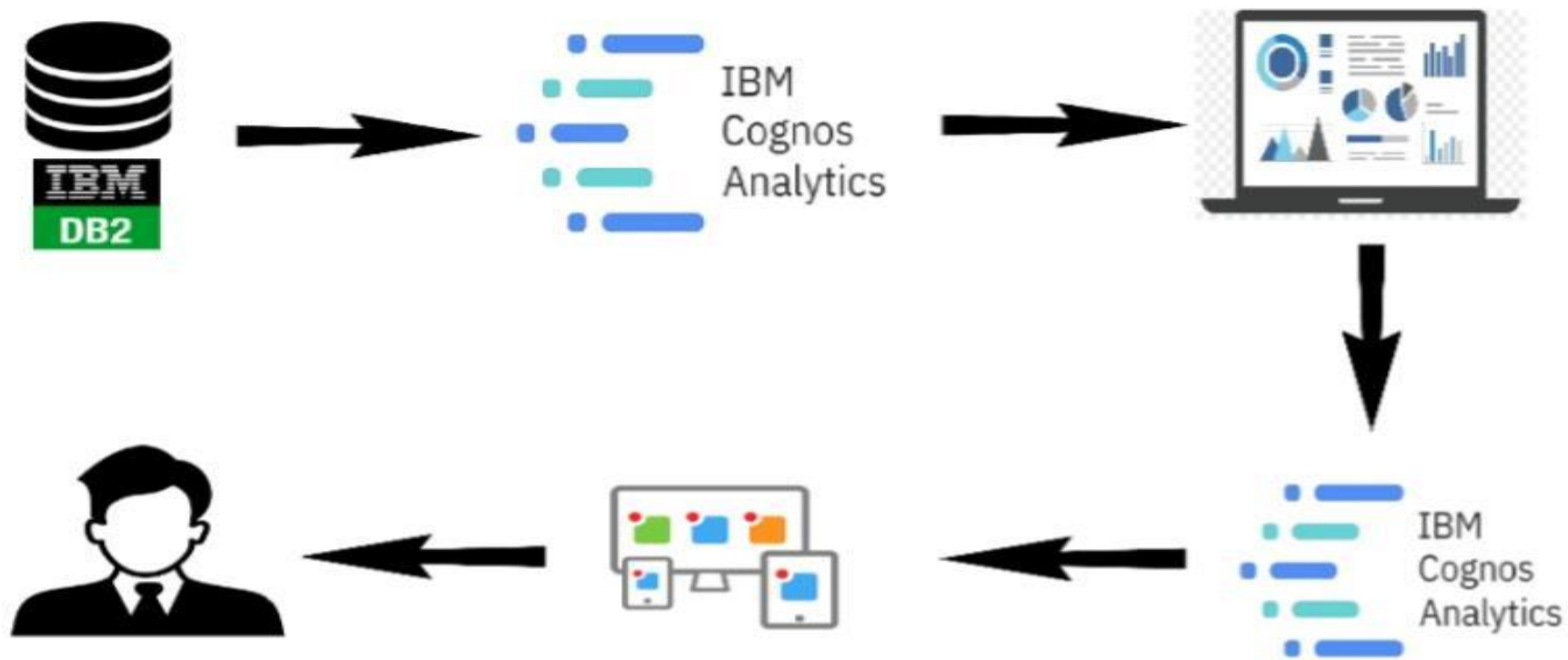
Devara Devakiran

Project Description:

The modern Olympic Games or Olympics are leading international sporting events featuring summer and winter sports competitions in which thousands of athletes from around the world participate in a variety of competitions. The Olympic Games are considered the world's foremost sports competition with more than 200 nations participating. The Olympic Games are held every four years, with the Summer and Winter Games alternating by occurring every four years but two years apart.

The evolution of the Olympic Movement during the 20th and 21st centuries has resulted in several changes to the Olympic Games. Some of these adjustments include the creation of the Winter Olympic Games for snow and ice sports, the Paralympic Games for athletes with a disability, the Youth Olympic Games for athletes aged 14 to 18, the five Continental games (Pan American, African, Asian, European, and Pacific), and the World Games for sports that are not contested in the Olympic Games. The Deaflympics and Special Olympics are also endorsed by the IOC. The IOC has had to adapt to a variety of economic, political, and technological advancements. As a result, the Olympics has shifted away from pure amateurism, as envisioned by Coubertin, to allowing participation of professional athletes. The growing importance of mass media created the issue of corporate sponsorship and commercialisation of the Games. World wars led to the cancellation of the 1916, 1940, and 1944 Games. Large boycotts during the Cold War limited participation in the 1980 and 1984 Games. The latter, however, attracted 140 National Olympic Committees, which was a record at the time.

Technical architecture:



What is data analytics?

Data analytics is the collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision making.

Data analytics is often confused with data analysis. While these are related terms, they aren't exactly the same. In fact, data analysis is a subcategory of data analytics that deals specifically with extracting meaning from data. Data analytics, as a whole, includes processes beyond analysis, including data science (using data to theorize and forecast) and data engineering (building data systems).

How is data analytics used?

Data is everywhere, and people use data every day, whether they realize it or not. Daily tasks such as measuring coffee beans to make your morning cup, checking the weather report before deciding what to wear, or tracking your steps throughout the day with a fitness tracker can all be forms of analyzing and using data.

Data analytics is important across many industries, as many business leaders use data to make informed decisions. A sneaker manufacturer might look at sales data to determine which designs to continue and which to retire, or a health care administrator may look at inventory data to determine the medical supplies they should order. At Coursera, we may look at enrollment data to determine what kind of courses to add to our offerings.

What is the importance of data analytics?

- Data analytics is the science of analyzing raw data to make conclusions about that information.
- Data analytics help a business optimize its performance, perform more efficiently, maximize profit, or make more strategically-guided decisions.
- The techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.
- Various approaches to data analytics include looking at what happened (descriptive analytics), why something happened (diagnostic analytics), what is going to happen (predictive analytics), or what should be done next (prescriptive analytics).
- Data analytics relies on a variety of software tools ranging from spreadsheets, data visualization, and reporting tools, data mining programs, or open-source languages for the greatest data manipulation.

Introduction to Business Intelligence:

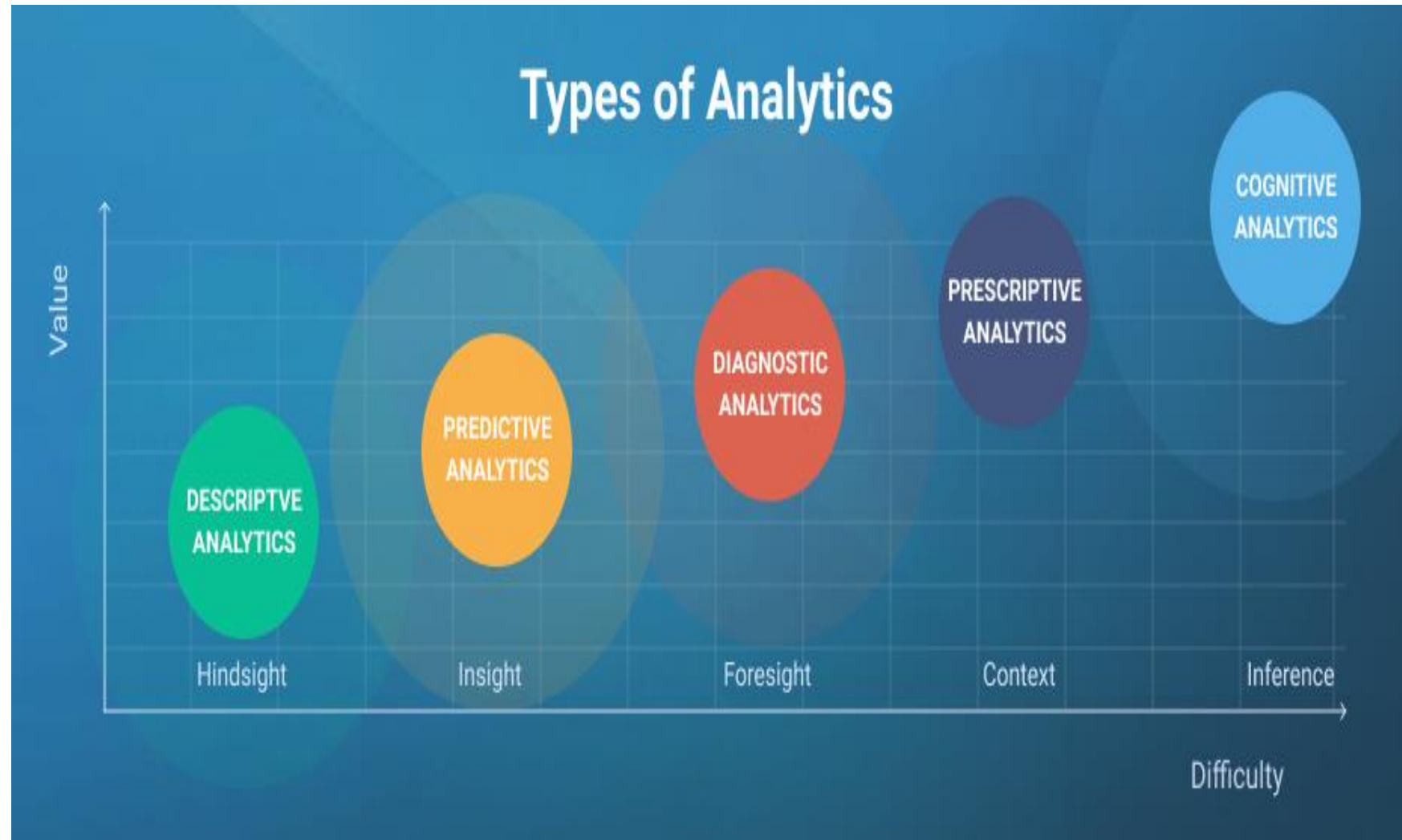
Business intelligence combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations make more data-driven decisions. In practice, you know you've got modern business intelligence when you have a comprehensive view of your organization's data and use that data to drive change, eliminate inefficiencies, and quickly adapt to market or supply changes. Modern BI solutions prioritize flexible self-service analysis, governed data on trusted platforms, empowered business users, and speed to insight.

THE MODERN ANALYTICS WORKFLOW



Types of Data Analytics

- Descriptive Analytics
- Diagnostic Analytics
- Predictive Analytics
- Prescriptive Analytics
- Cognitive Analytics



Working with IBM Cognos Analytics:

Cognos Analytics is a set of business intelligence tools available on cloud or on-premise. The primary focus is in the area of Descriptive Analytics, to help users see the information in your data through dashboards, professional reporting and self-service. IBM Cognos Analytics and Microsoft Power BI are two of the top business intelligence (BI) and data analytics software. Both of these applications are in heavy demand, as organizations seek to harness the vast repositories of data – with ever more created hourly. Whether from unstructured data, social media, relational databases, or overworked enterprise applications, there is a vast amount of data subjected to data analytics. Instead of a small team of business intelligence pros mining data, now there are teams from marketing, sales, and IT all utilizing analytics in search of the benefits of digital transformation data exploration.

Why is Data Analytics in Sports Useful?

The global sports market grew from \$354.96 billion in 2021 to \$496.52 billion in 2022, according to the business research company's research, making it one of the biggest markets in the world. As spending in the global sports market increases, sporting teams are much more committed to investing in sports data analytics for a competitive edge, hence the projection of the sports analytics market reaching upwards of \$4.5 billion by 2025. Sports data analytics can be a useful tool for making important strategic decisions. For example, when Thomas Tuchel, former Chelsea Football Club manager, was asked why he brought on goalkeeper Kepa Arrizabalaga late into extra time, he responded, "So we had some statistics, we were well prepared, that Kepa is the best in percentage in saving penalties." He added, "There is proof that Kepa is better at this discipline." Chelsea went on to win the penalty shootout, with Kepa saving two penalties in the process.

How Different Sports Use Analytics:

Although the fundamental purpose of sports data analytics is universal – to gain a competitive advantage through statistics and data analysis – different sports deploy varying methods to collect and analyze data efficiently for their sport.

- **Soccer-** data analytics forms an integral aspect of off-field decision-making in soccer. Soccer clubs worldwide have invested heavily in data science and related technology to help boost players' on-field performance and improve off-field decision-making.
- **Baseball-**As one of the first sports to embrace sports analytics, Baseball has been setting the benchmark for several years. Some of the brightest minds in the game, like Theo Epstein, have never even set foot into a major or minor league baseball game. Instead, these people depend on their advanced education and passion for crunching numbers to help make decisions on and off the field.

Why is data important in sports performance?

Data can be used to evaluate player or team performance over time, recognizing areas of progress and tracking the progress. This helps players and coaches identify where they need to put focus while preparing and improving their performance on the field.

Overall, sports analytics and data-driven decision-making are important because they allow teams and coaches to make better decisions, optimize performance, and achieve greater success both on and off the field. By utilizing data and analytics, teams can gain a competitive advantage and improve their chances of success.

How data analysis in sports is changing the game?

Using advanced analytics technologies, companies can improve human resources practices and customer relationship management by using astute data analysis in sports. Teams and associations can make key decisions about their core products and services to help improve the experience for customers and maximize revenue.

At the MIT Sloan Sports Analytics Conference, for example, Evan Wasch, senior vice president of basketball strategy and analytics for the NBA, described an intricate web of decision points that affect the quality of the product, ranging from scheduling and playoff structure to draft lotteries. Data can prompt small changes that make a big difference, he asserted.

What is the future of data analytics in sport?

Improved Player and Team Performance With the help of AI algorithms, coaches and players can analyze vast amounts of data about games, opponents, and their own performances. This data can be used to identify strengths and weaknesses, develop strategies, and make informed decisions in real-time. AI can also help with injury prevention by tracking player movements and analyzing biomechanical data to identify any movements that could lead to injury.

- **Data-Driven insights on Olympic Sports**
Participation and Performance for data analytics:

Introduction:

Olympics is considered as most important event worldwide, which provides common platform to players from various nations to show their talents. Olympics has been started at 1896, which is being conducted once in every four years. The goal of this paper is to analyze performance and participation of nations in Olympics from 1896 to 2012. In addition, the field of sports of particular country in particular year, in which they have contributed the maximum can be identified. The comparison of the performance of each sports with other can be done. The field of sports, that has to have more participation can be identified and necessary action can be taken by players and nations to enhance themselves in future contributions towards Olympics.

LITERATURE SURVEY :

Performance measure for a country in Olympics can be predicted using their win using maximum value scored by them in previous participation, the chance of winning gold in 2016 has been identified[1]. If a person wins a medal in an Olympics during a year, the chance of winning a medal in upcoming Olympics was predicted[1]. Having sports performance data, predicting one's future performance has been done [2]. Their performance can also be increased, if they are not performing good in certain areas and then placing them accordingly in the training program will provide considerable measure in their outcomes[2]. Machine learning techniques were used for heuristics prediction of Olympic medals of a country[4]. Estimation of Olympics success of a country can be done by efficiency analysis and importance of sports in society[5]. When analyzing the sports categories they are mainly being more presentive towards view point based content rather being an view point which is spatio temporal. The video content analysis has significance of providing more interior information than structured collected data [3]. In addition to these techniques, the exploratory data analysis uses visual methods to provide deep understanding and statistical summary about the data .

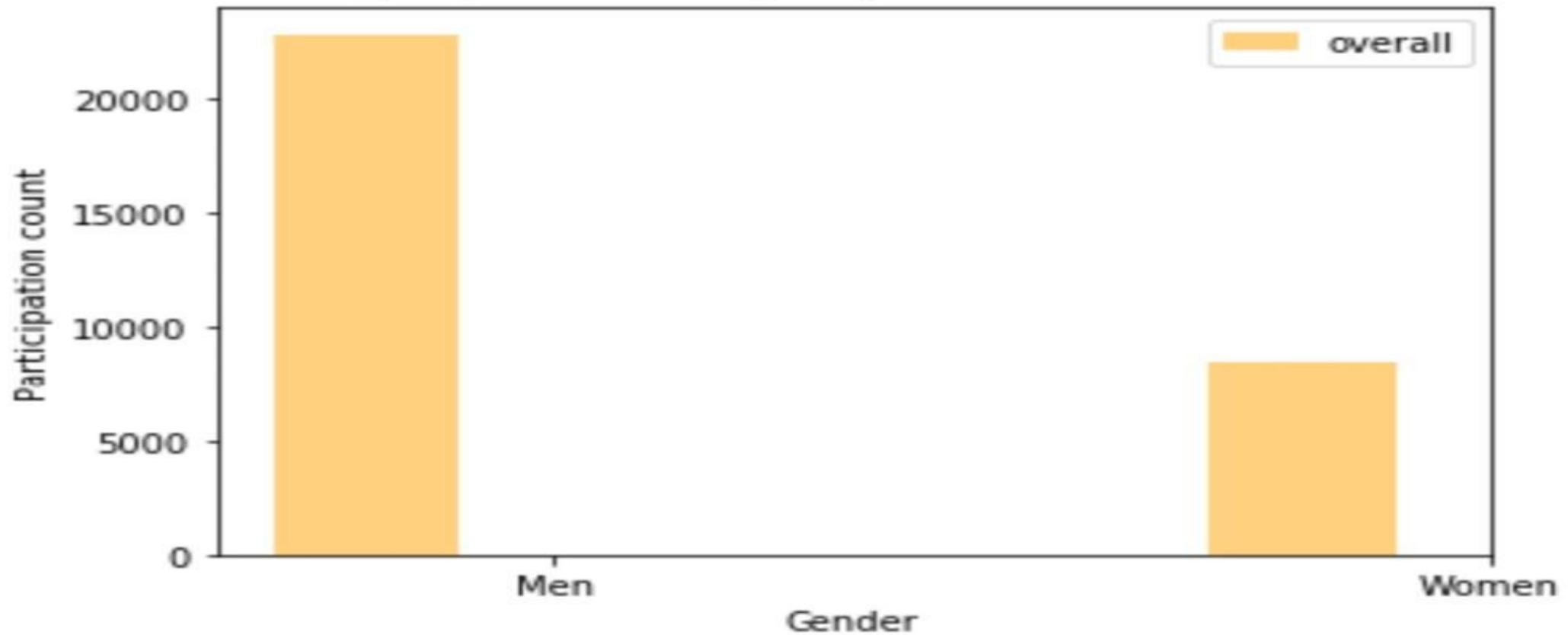
ANALYSIS AND VISUALIZATION:

The Summer Olympics dataset with data collections from 1896 to 2012 has been analysed. This dataset contains around 30,000 rows and 9 columns. The fields include Year, Sports, Discipline, Medal, Gender, Country, City, Event and Athlete.

Identifying Contribution of Men And Women Participants In Olympics (1896-2012):

The total number of men and women participants in Olympics from 1896-2012 is analysed and the ratio between men and women participants can be obtained. The analysis represents contribution of men is higher than women among all over the world. The figure shows gender wise contribution of players in Olympics.

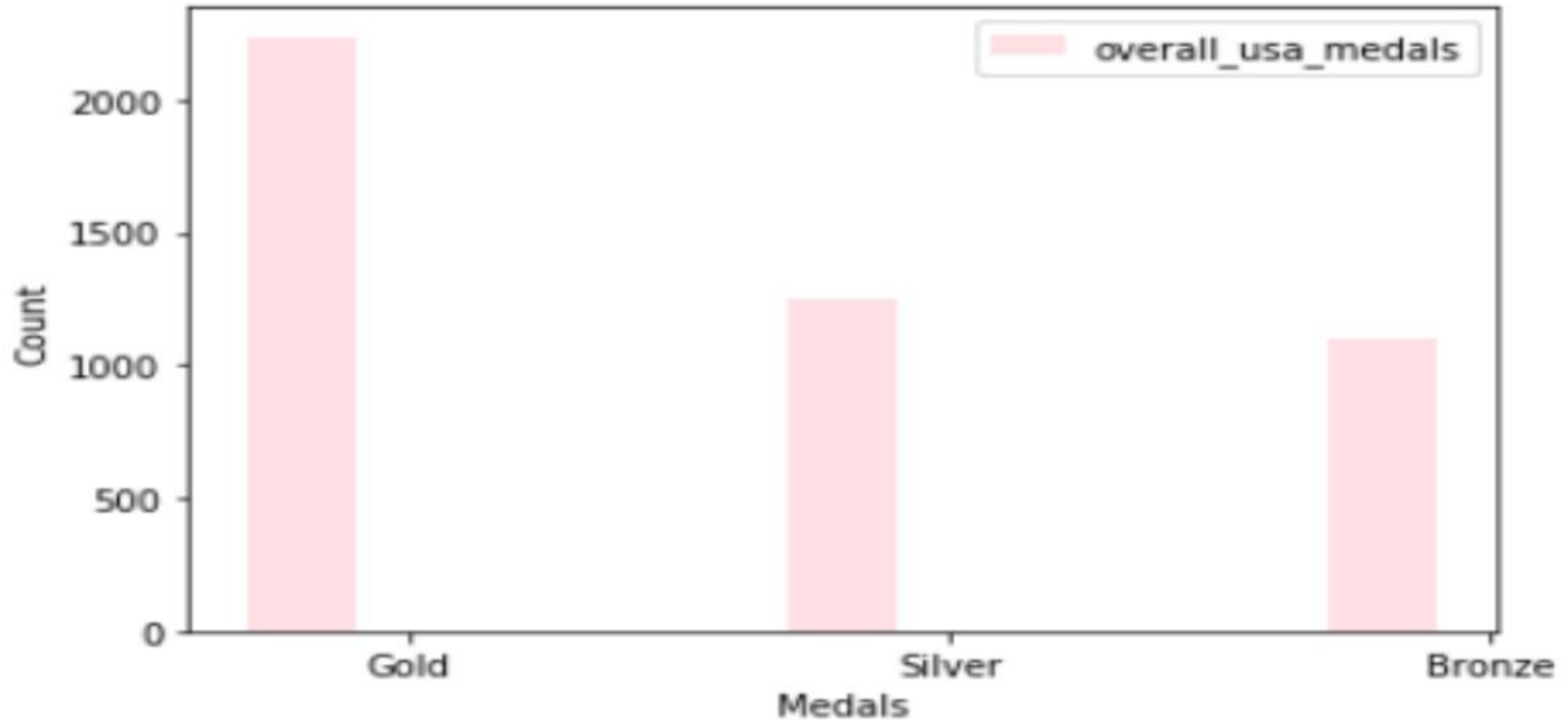
Gender Wise Contribution in Olympics:



Performance Analysis in Olympic Games using Exploratory Data Analysis Techniques :

In this analysis, total number of gold, silver and bronze medals won by the participants from all countries in Olympics from 1896 to 2012 can be identified. The count includes number of individuals who were contributed separately or as a team to receive medals for their nations. The following results are obtained in the analysis. (i) USA has won the highest number of gold medals when compared to other medals and almost equal percent of silver and bronze medals. (ii) Australia has received least number of gold medals when compared to other medals and won the highest number of bronze medals. Japan has least number of gold than other medals. France has less number of gold and high number of silver and bronze medals

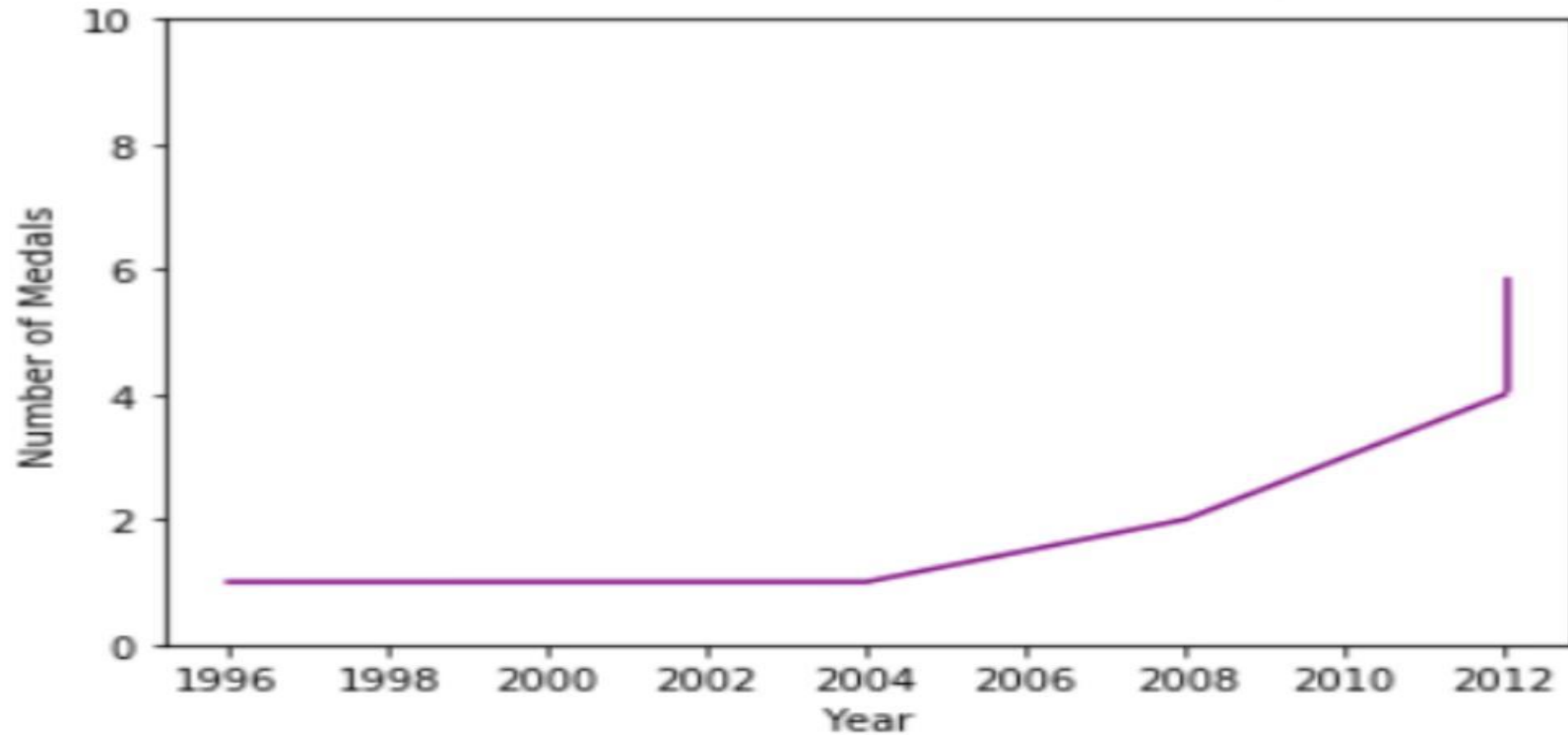
Total Participants of USA Contributed for Winning Medals in Olympics:



Identifying the performance of Particular Country in Olympics (1992-2012):

Excellence of a country in Olympics can be viewed by number of medals won by a country. This analysis identifies the performance a particular country in Olympics from 1992 to 2012. This can be processed by calculating the total medals won by particular country in particular year from 1992 to 2012. Data visualization can be carried out to represent the result of particular country. The results are (i) Performance of India was gradually increasing from 1992 with no medals, 1996 with 1 medal and finally in 2012 with 6 medals. (ii) Performance of USA was found like zig-zag graph from 1992 with 220 medals, 1996 with 260 medals, suddenly performance has decreased in 2000 with 240 medals, increased gradually from 2004, contributed best in 2008 with 350 medals. (iii) France's Performance was gradually increasing from 1996 to 2008 with medals within range of 40 and has performed well in 2012 with 80 medals. (iv) Performance of Australia was better during 1992 Olympics with 60 medals and there was a sudden increase in its performance with almost 200 medals over the period of 2000 and there has been gradual decrease in performance from 2004 to 2012. (v) Initially, performance of Japan was not so good, but over the period of 2000 and 2004 there was a drastic increase in it and gained 100 medals which was higher than the rest.

Performance of India in Olympics:



Comparing the performance between the countries in Olympics(1996-2012) :

The analysis compares the performance between the countries by medals won by the participants from selected countries in Olympics from 1996 to 2012. Countries such as USA, Hungary, France, Japan, Australia are selected for analysis. From this analysis, the following results have been inferred.

(i) In 1996 Olympics, among the five selected countries, USA is the leading country with contribution of 7.53%, Australia is the second country with 2.25%, Japan with 1.61%, Hungary with 0.75% and France is the least country with 0.69% contribution among them.

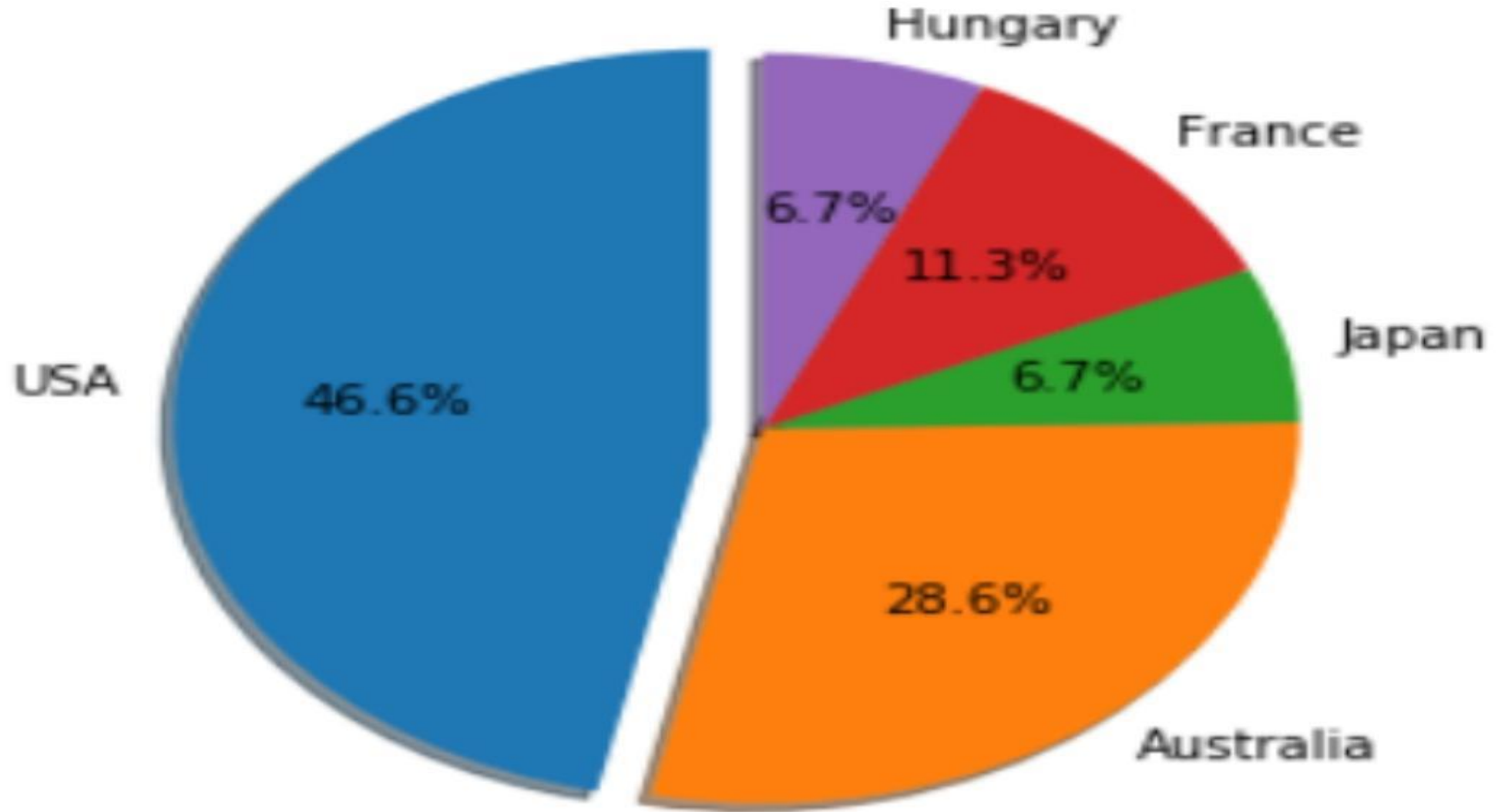
(ii) In 2000 Olympics, USA is the leading country with contribution of 6.55%, Australia is the second country with 4.019%, France with 1.58% and Hungary & Japan are the least country with 0.94% among them.

(iii) In 2004 Olympics, USA is the leading country with contribution of 8.05%, Australia is the second country with 3.3%, Japan with 2.1%, France with 1.6% and Hungary is the least country with 0.9% among them.

(iv) In 2008 Olympics, USA is the leading country with contribution of 8.52%, Australia is the second country with 3.72%, France with 1.71%, Japan with 1.42% and Hungary is the least country with 0.88% among them.

(v) In 2012 Olympics, USA is the leading country with contribution of 8.5%, Australia is the second country with 3.01%, Japan with 1.8%, France with 1.6% and Hungary is the least country with 0.9% among them.

Comparing the Performance of Selected Countries in Olympics:



Identifying the Best Performed Field of Sports for Particular Country in Olympics (2000-2012):

The analysis represents the performance from participants of particular country and their best performed field of sport in Olympics from 2000 to 2012. To identify field of sports of particular country in particular year and to analyse which field of sport has to have more participation. This provides information to enhance themselves in future contributions towards Olympics.

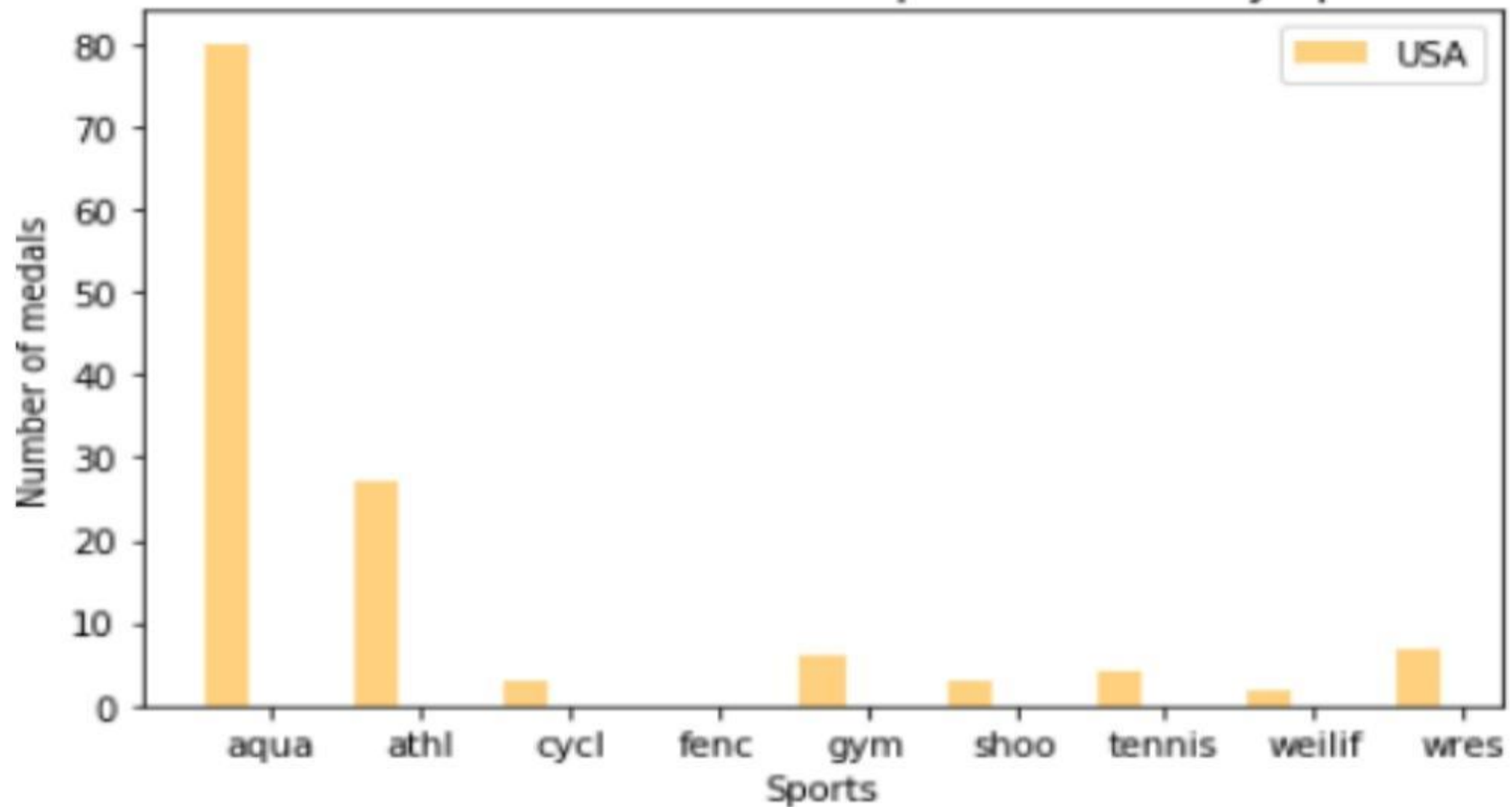
(i) In 2000, USA has performed best in the field of Aquatics and has performed least in the field of Weightlifting.

(ii) In 2000, Australia has performed best in the field of Aquatics and has performed least in the field of Gymnastics.

(iii) In 2000, France has performed best in the field of Fencing and has performed least in the field of Tennis.

(iv) In 2000, Australia has performed best in the field of Aquatics and has performed least in the field of Athletics.

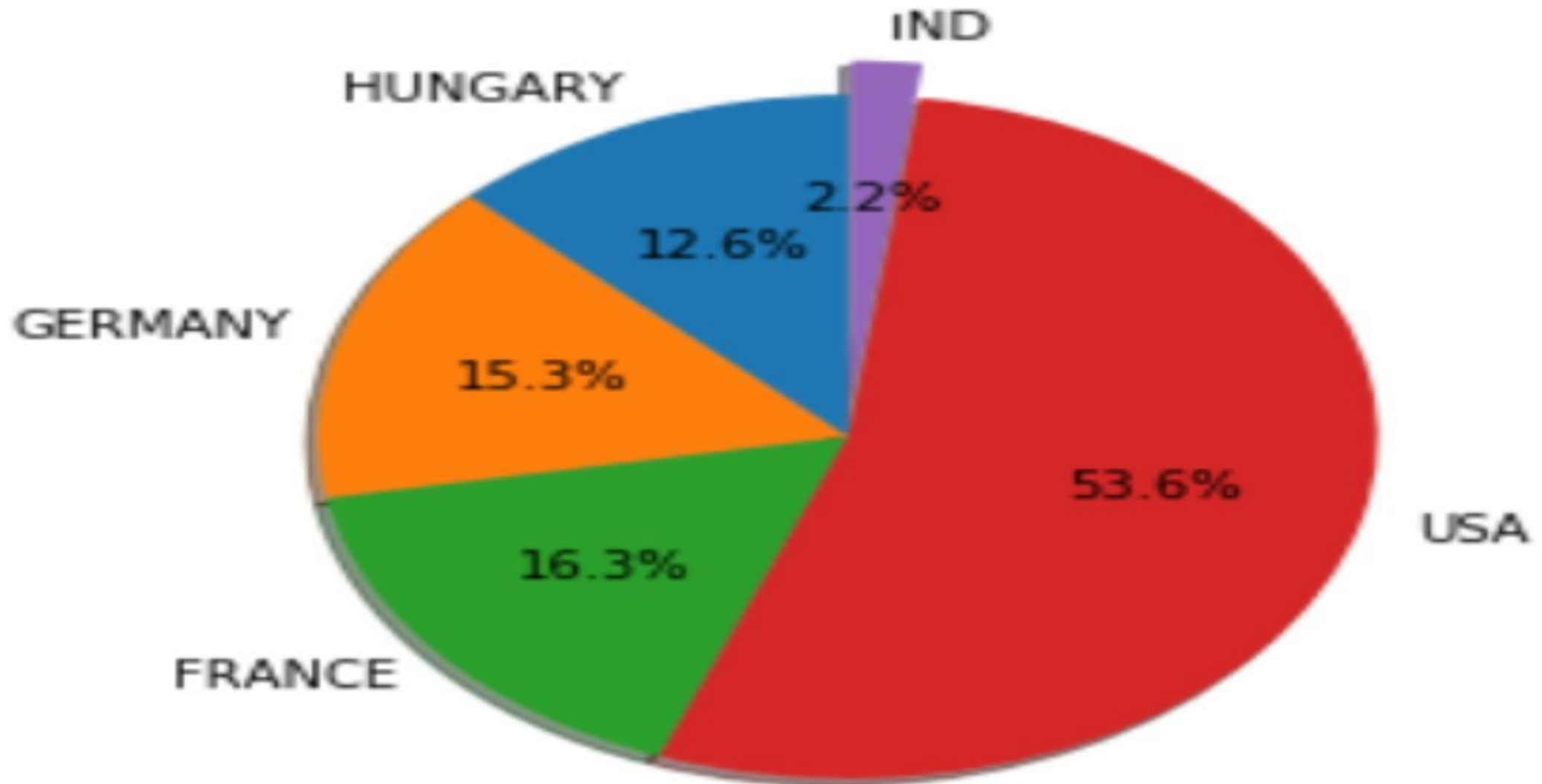
Performance of USA in the Year 2000:



Comparison of Overall Performance of Selected Countries with India in Olympics (1896-2012):

This analysis helped to find the performance of Selected countries and compare those with a particular country. Here India's performance was compared with USA, Hungary, France, Germany, France. This has been carried out by finding the percentage of medals won by each countries. The resulted percent has been plotted to analyze the performance. Among the selected countries, USA is the top most country with 53.6% which has contributed higher than other countries. India is one which has least percent of 2.2%. France is next to USA with 16.3%, Germany is the next to France with 15.3%. Hungary with 12.6%.

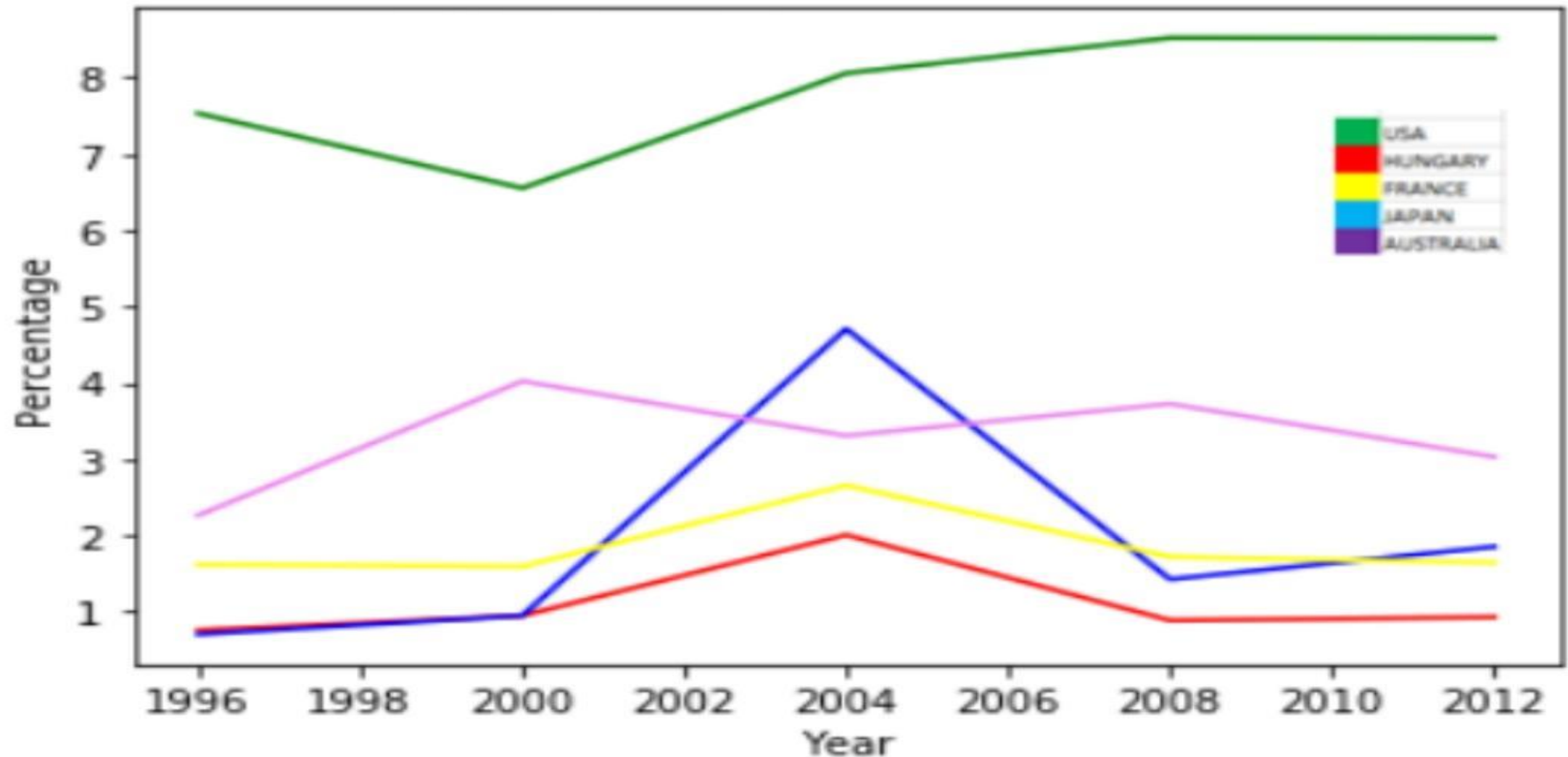
Comparing the Performance of India with Selected Countries:



Analyzing the Highest Performance of a Country in Given Period From 1996 to 2012:

From this analysis, USA has performed well in Olympics 2012 compared to other years. Australia has performed well in Olympics 2000. France has performed well in Olympics 2008 compared to others. Japan has performed well in the year 2004. Hungary's performance was good in the Olympics conducted in the year 2004.

Highest Performance of a Country in Olympics from 1996 to 2012:



Future Scope:

We all know that any Analysis is not perfect and it consists of some limitations which define the Future scope of the Research Work. This project work also contains some limitations which we are considering as the Future Scope of the Project. We can also describe the data in other formats like Geographical format where we can depict the countries on the World map. We can also apply various Machine Learning Algorithms to the data set after Analysis and can create a Predictive Model which can predict the statistics of the Future Olympic Games.

Advantages of Data Analytics:

- Data analytics helps an organization make better decisions.
- Increase the efficiency of the work.
- The analytics keeps you updated of your customer behavioural changes.
- Personalization of products and services.
- Improving quality of products and services.

Disadvantages of Data Analytics:

- Lack of alignment within teams
- Lack of commitment and patience
- Low quality of data
- Privacy concerns
- Complexity & Bias

Conclusion:

The exploratory data analysis on Olympic dataset provides statistical and visual representation of performance of nations, players in Olympics from the year 1896 to 2012. From the above analyses, it is useful to identify the country that needs more skills, the field of sports in which players are performing well and players who need practice to enhance themselves in upcoming Olympics. The contribution of women in Olympics have to be encouraged. Country which has least performance have to find the steps to improve their performance. Country which has performed best so far will also consider to take measures to increase their performance.