An Analysis of Summer/Winter Olympic Games Throughout History (1986 - 2016)

James Wang, Kushal Dave, Patrick Lee & Serim Jang

Background

- The first Modern Olympic games were started in the year 1896 in Athens, Greece with athletes only from 12 countries. Since then, the summer Olympics Games were held for more than 25 times and the athletes from each country have increased in addition to the participating countries, adding to the friendly spirit of competition.

- This analysis seeks to find general patterns in relationships between the countries and the performances of the athletes, and more specifically the impacts that the physical aspects of the athletes bring to different performance levels, ages, and other factors.

- All visualizations were conducted in R using the Olympic dataset on Kaggle.

Countries That Won the Most Medals

1   982
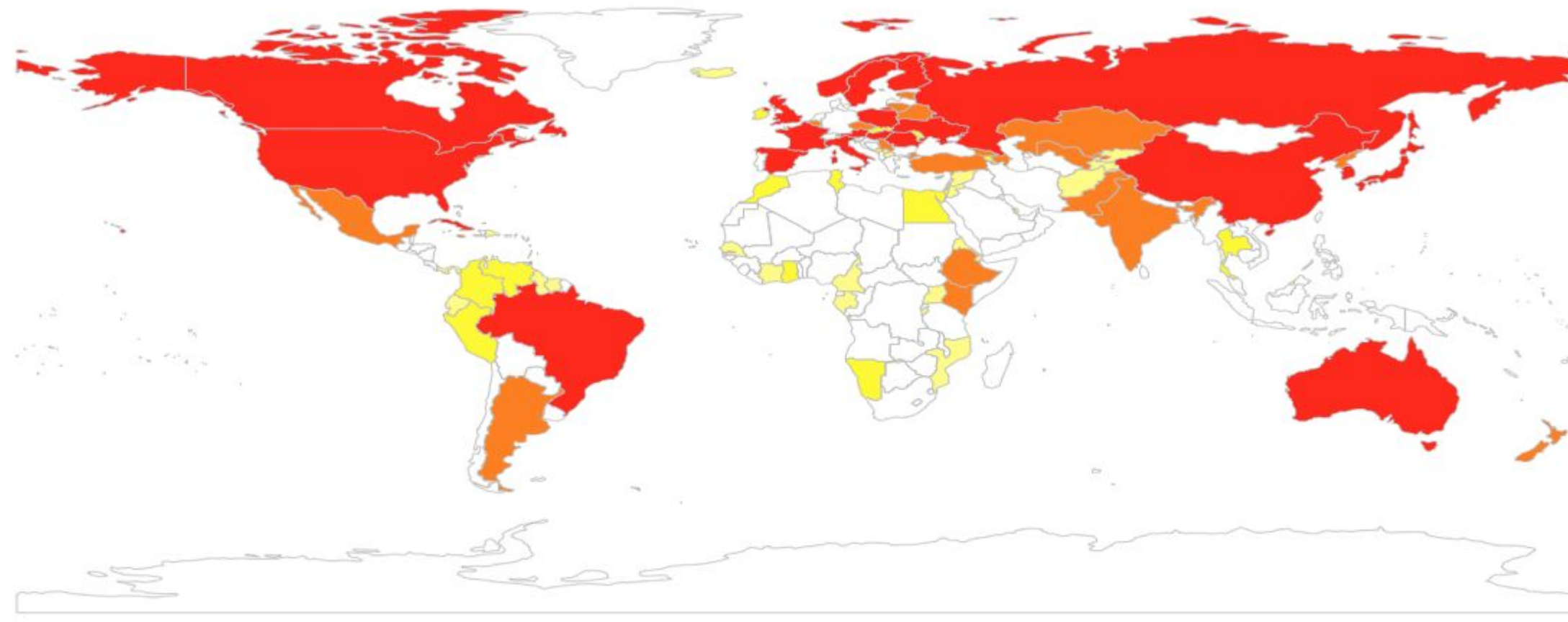
- In the graph, we observe that the countries that win a lot of medals at the Olympic games are the developed countries in North America, Asia and Europe.
- Less developed countries in Africa and South America are less likely to win medals.

Analysis of the **Physical Aspects** of Athletes and Their Effects

**Density of Height of Athletes' by Winning**

Won   FALSE   TRUE
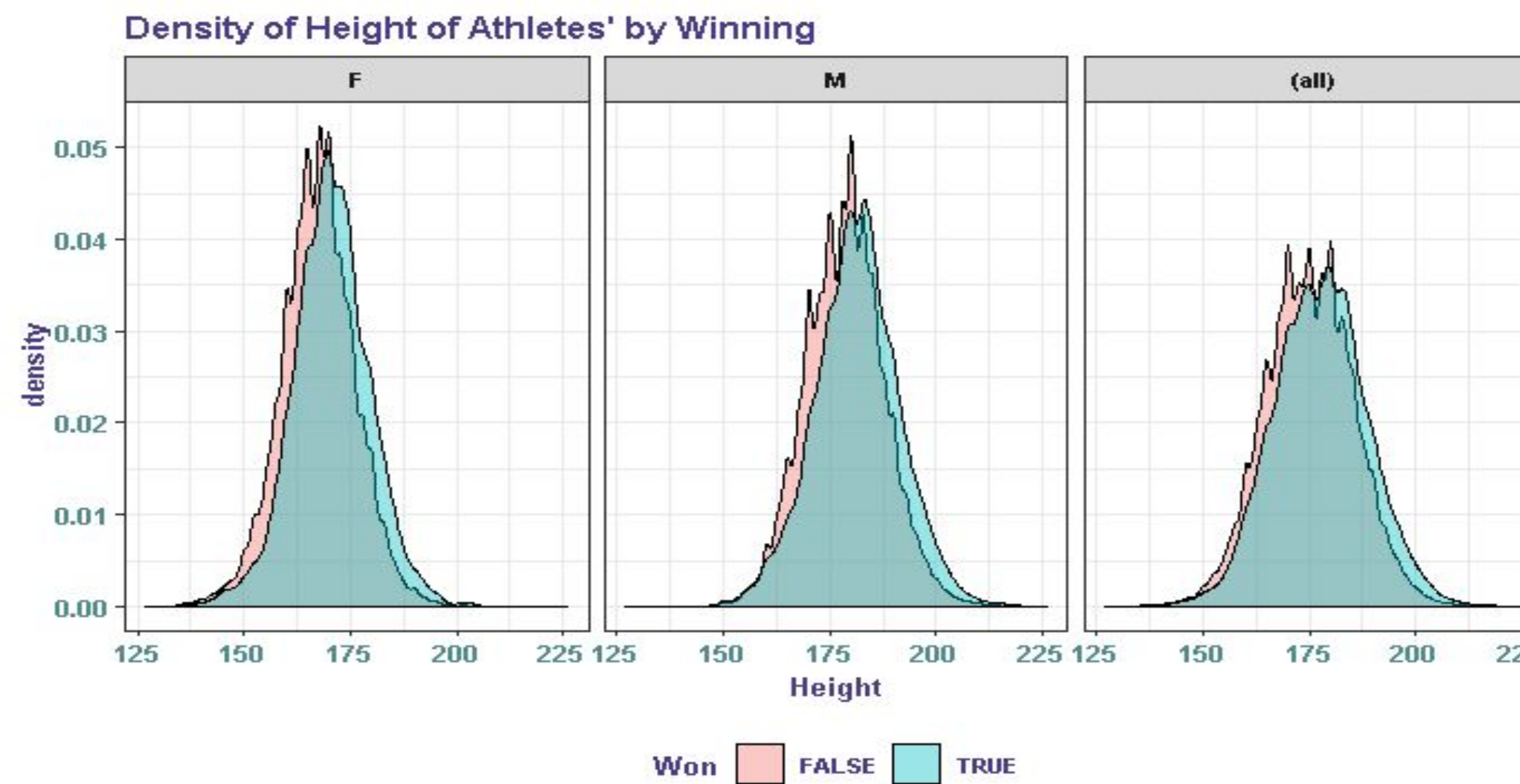
- Across genders, winners have a higher normal distribution in height than non-winners.
- The mean height difference between winners and non-winners is 3 cm.
- The distribution are both normal for females and males, although male height is centered around 180 cm and female around 172 cm.

**Cluster Dendrogram of Winter Olympic Sports by BMI of Athletes**
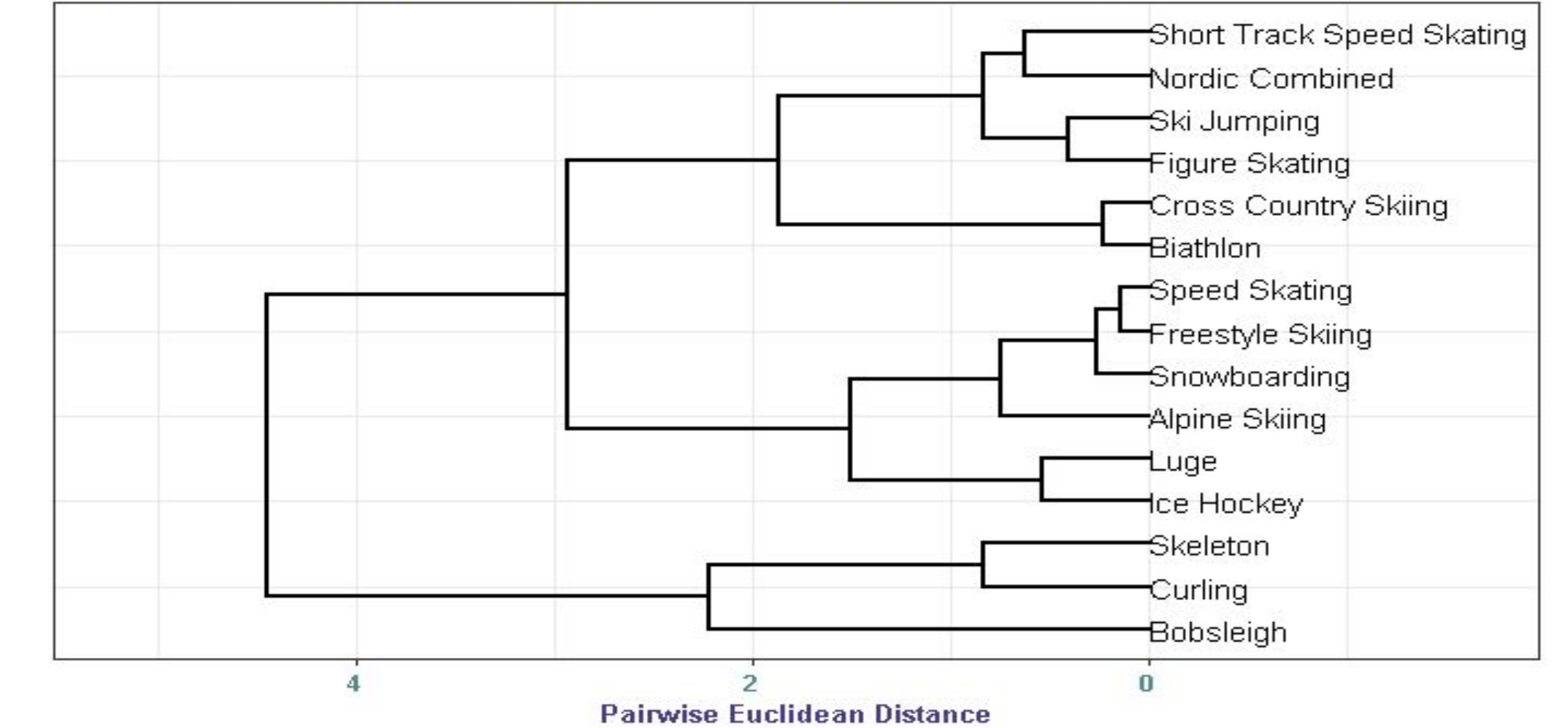
- Among winter sports, the average BMI of the athletes varies considerably.
- There are sports where the athletic builds of athletes are similar, such as freestyle skiing and speed skating.
- In the example of freestyle skiing and speed skating, the similar BMIs emphasize the importance of versatility.
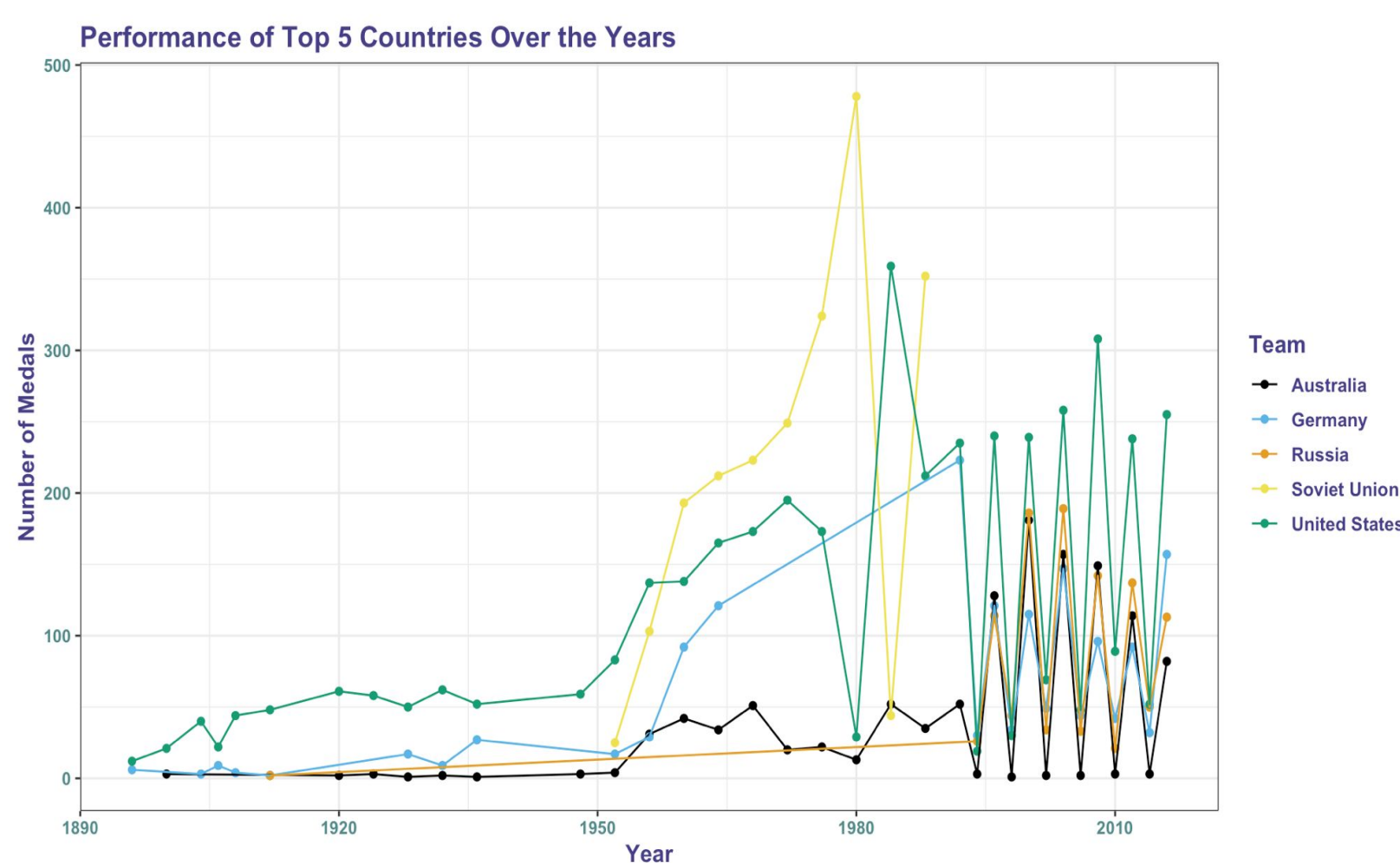
Age of Athletes by Age Over the Years
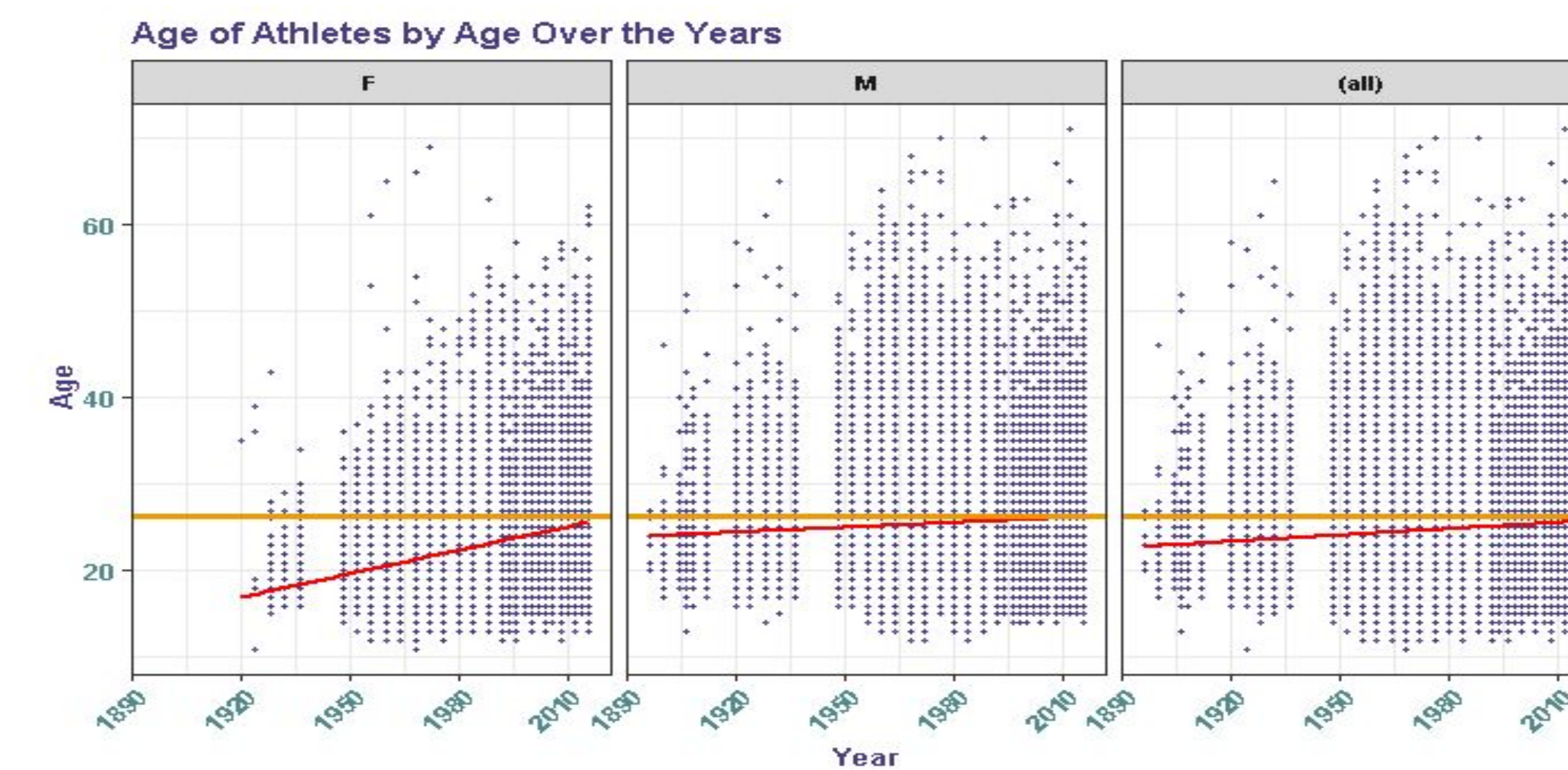
- The vertical orange lines indicate the average age of athletes in 2016.
- The overall age increased over the years since the inception of the Olympics in 1896. However, the increase is much more pronounced for females compared to males.
- We believe that the belief in the past that females after a certain age are expected to retire played a role, but the increase in age shows that this belief is not common in the present.

Changes in BMI by Sport for Female and Male

Sport: Baseball, Basketball, Figure Skating, Gymnastics, Volleyball
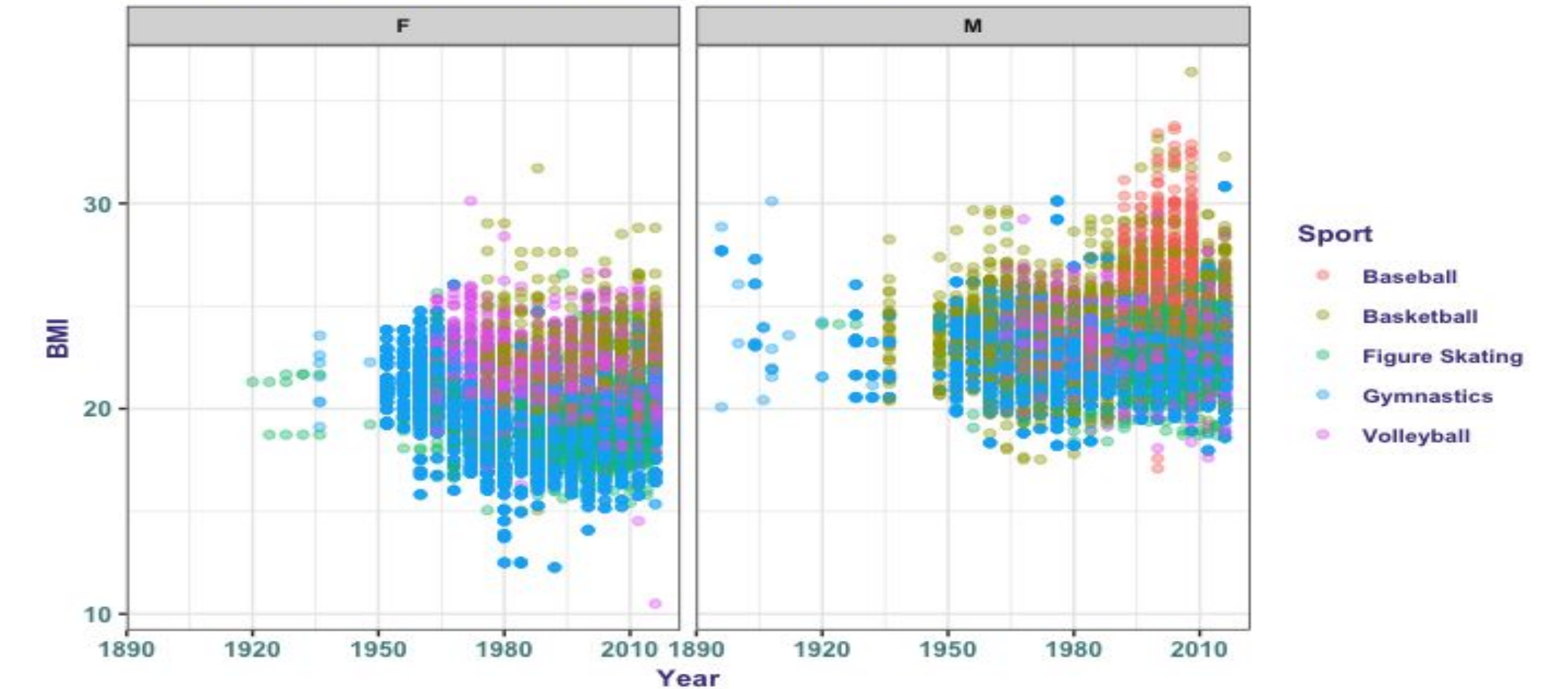
- Changes in BMI by year with information on which particular sports have a higher or lower BMI by gender is illustrated.
- We see a difference in general BMI for different sports; for example, gymnasts tend to have a lower BMI while basketball players and volleyball players tend to have a higher BMI. We also notice a significantly higher BMI for male baseball players.
- There were many more male sports players than female sports players before 1950. Females tended to be figure skaters and males tended to be gymnasts at this time.
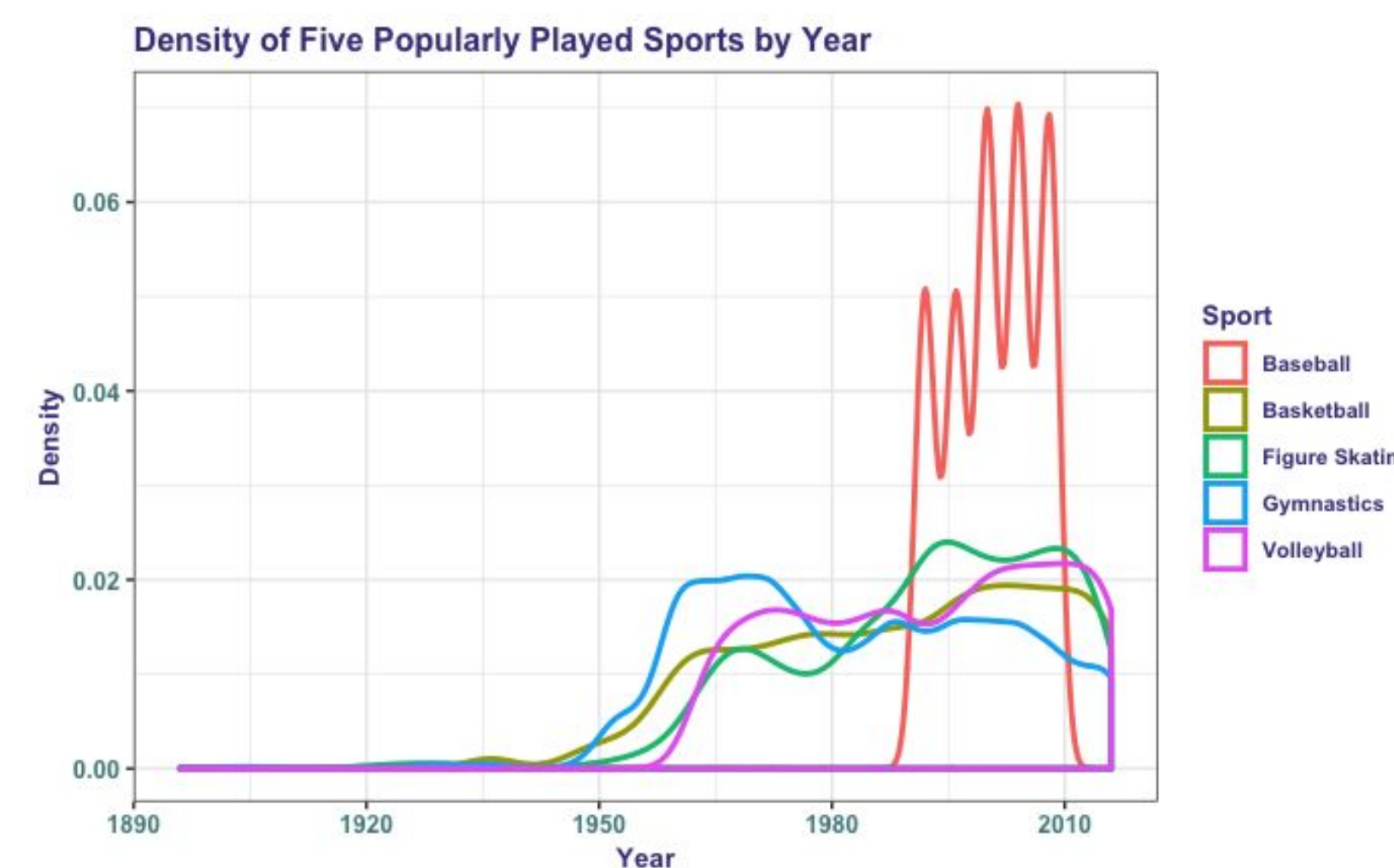
Exploratory Data Analysis

Performance of Top 5 Countries Over the Years

Team: Australia, Germany, Russia, Soviet Union, United States

- USA has been the consistent highest performer in the Olympic Games, except for a period between 1950-1980, where the Soviet Union won more medals than USA.
- The Soviet Union won about 450 medals in the 1980 games, which is the greatest number of medals won by a country, while Australia won the lowest number of medals.
- Performance of Germany skyrocketed between 1950-1980, then diminished in the following years.

Top 20 Athletes in the History of Olympic Games

Type of Medal: Bronze, Gold, Silver
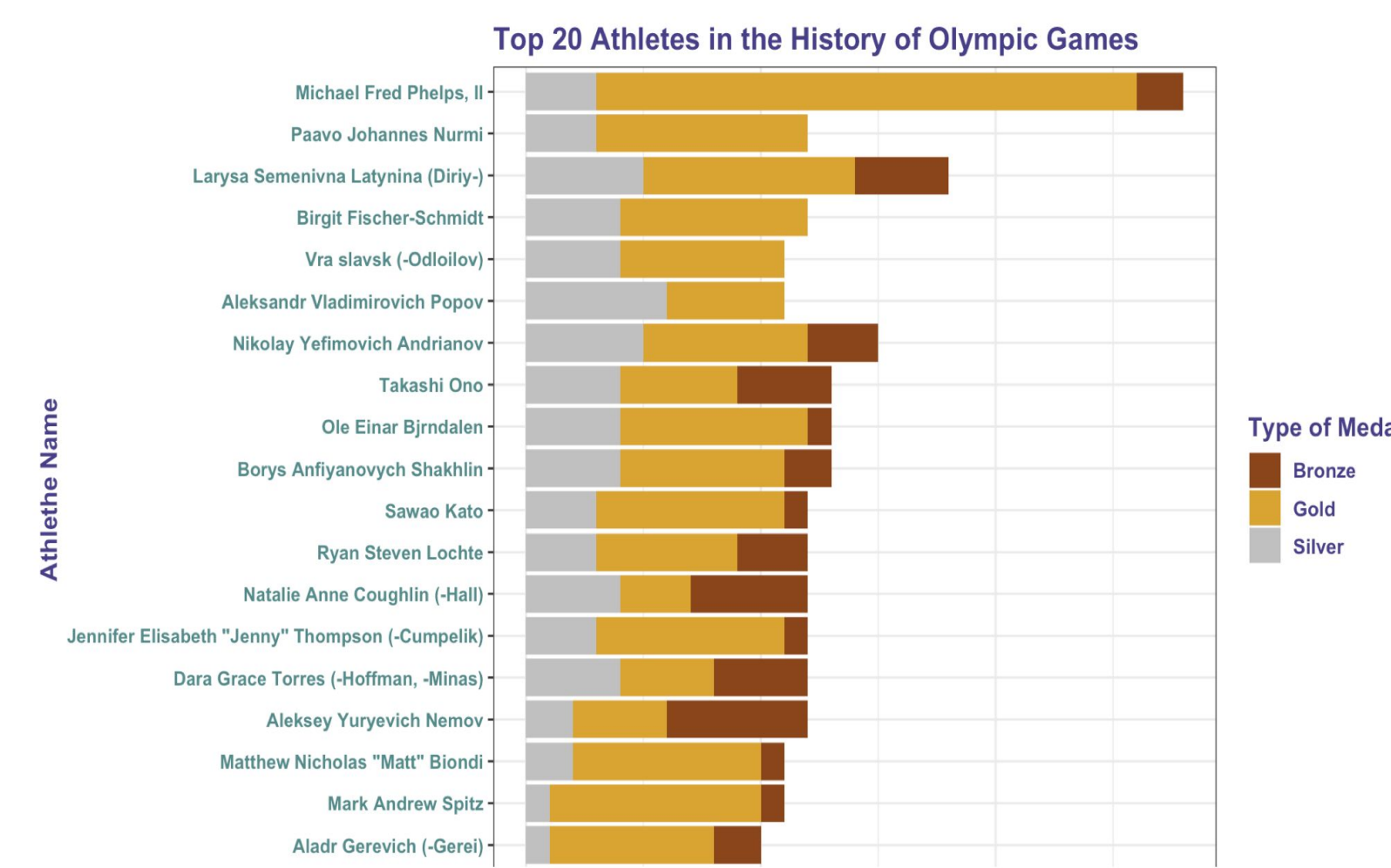
- Michael Phelps, Larisa Latynina and Nikolay Andrianov are the most successful athletes over the history of Olympic Games by the total number of medals won.
- Michael Phelps is much ahead of every other athlete with twenty two medals, as can be seen by the difference in number.
- Also, a vast majority of his medals are gold medals.

Density of Five Popularly Played Sports by Year

Sport: Baseball, Basketball, Figure Skating, Gymnastics, Volleyball

- Changes in the popularity of playing each of the five sports chosen is illustrated, and through the density curves we can observe the rise and fall of popularity by year.
- There is a rise in popularity in playing sports in general starting around 1950. Then we see variation, such as how gymnastics has the highest density in the beginning but starts to lose popularity, and how figure skating steadily gains players.
- Baseball is by far the most popular sport. The rapid up and downward movement of the curve indicate that this is a sport only played in the summer.

Discussion

- The EDA showed interesting findings. For example, Michael Phelps won almost twice the medal than the runner-ups. Change in popularity of sports is also observed. EDA also shows the strong performance of the Soviet Union leading to its breakup, and the United States taking lead after that.

- From the analysis of physical aspects of Olympic athletes, we can see that height is positively correlated with chance of winning across sports. While BMI has a huge range, it varies a lot by sport type. An interesting point to note is that more women, including older women, are competing in the Olympics in a more variety of sports.

Acknowledgements