Olympic Games

Connor Hudziak

3/4/2022

Introduction

The Olympics data set is composed of information from every Olympic games from 1896-2016. The set includes information about all athletes such as age and height. It also includes geographical data that includes where the games were held and if it was summer or winter. I wanted to focus closely on the winter Olympic events and how the average ages of males and females related to each sport. I then took the highest and lowest average age by sport to look closer at each specific athlete through a histogram.

Data wrangling

In this section I have taken the Olympics data set and grouped it according to sport and sex. I then filtered for the Winter games as well as remove the Alpinism sport which I will explain further in the next section. Following that I calculated the average of all the winter sports under the Olympics MF variable.

After concluding that the highest average age for both male and female was in Curling, I grouped the OlympicsMF by sport and filtered for only Curling athletes. I then found the youngest average age was Figure skating and then did the same as the previous.

```
olympicsMF <- olympics %>%
  group_by(sport,sex) %>%
  filter(season == "Winter",
         sport != "Alpinism") %>%
  summarise(avg_age = mean(age, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'sport'. You can override using the `.groups` argument.
```

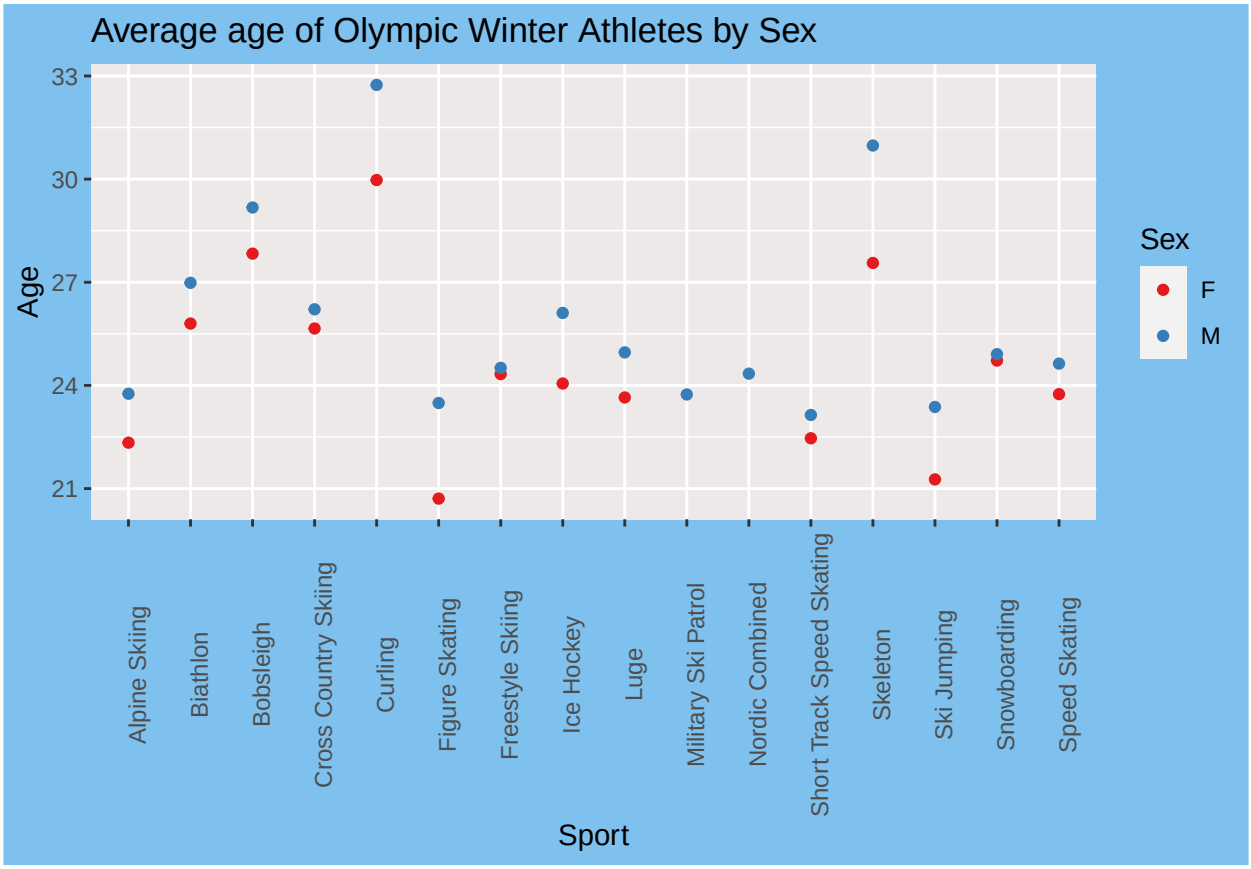
```
olympics_Curl <- olympics %>%
  group_by(sport) %>%
  filter(sport == "Curling")



olympics_FS <- olympics %>%
  group_by(sport) %>%
  filter(sport == "Figure Skating")
```

Average Age by Winter Olympic Sport

In this scatter plot I have visualized the average age of males and females according to their Olympic sport. I have also omitted one sport as it was an outlier and that was Alpinism. Alpinism was only in the Olympics

in 1924 summer and winter as well as summer of 1936. The sport consisted of climbing summits and every participant was a male except for one female. The average age was 40 and did not accurately portray the goal that I was achieving. In this visualization, most of the averages fall between 21-30 years old. It can also been seen that females are almost always younger than the male athletes in the same sport which could be attributed to age at which males and females mature at different ages. The highest average sport was Curling with an average of 29.97 for females and 32.74 for males. The youngest average age for a sport was figure skating with 20.71 for females and 23.49 for male athletes.

```
ggplot(olympicsMF, aes(x = sport,
                       y = avg_age,
                       color = sex)) +
  geom_point() +
  theme(axis.text.x = element_text(angle = 90),
        plot.background = element_rect(fill = "skyblue2"),
        panel.background = element_rect(fill = "snow2"),
        legend.background = element_rect(fill = "skyblue2")) +
  labs(title = "Average age of Olympic Winter Athletes by Sex",
       x = "Sport",
       y = "Age",
       col = "Sex",
       col.axis = "grey0",
       col.lab = "grey0") +
  scale_color_brewer(palette = "Set1")
```

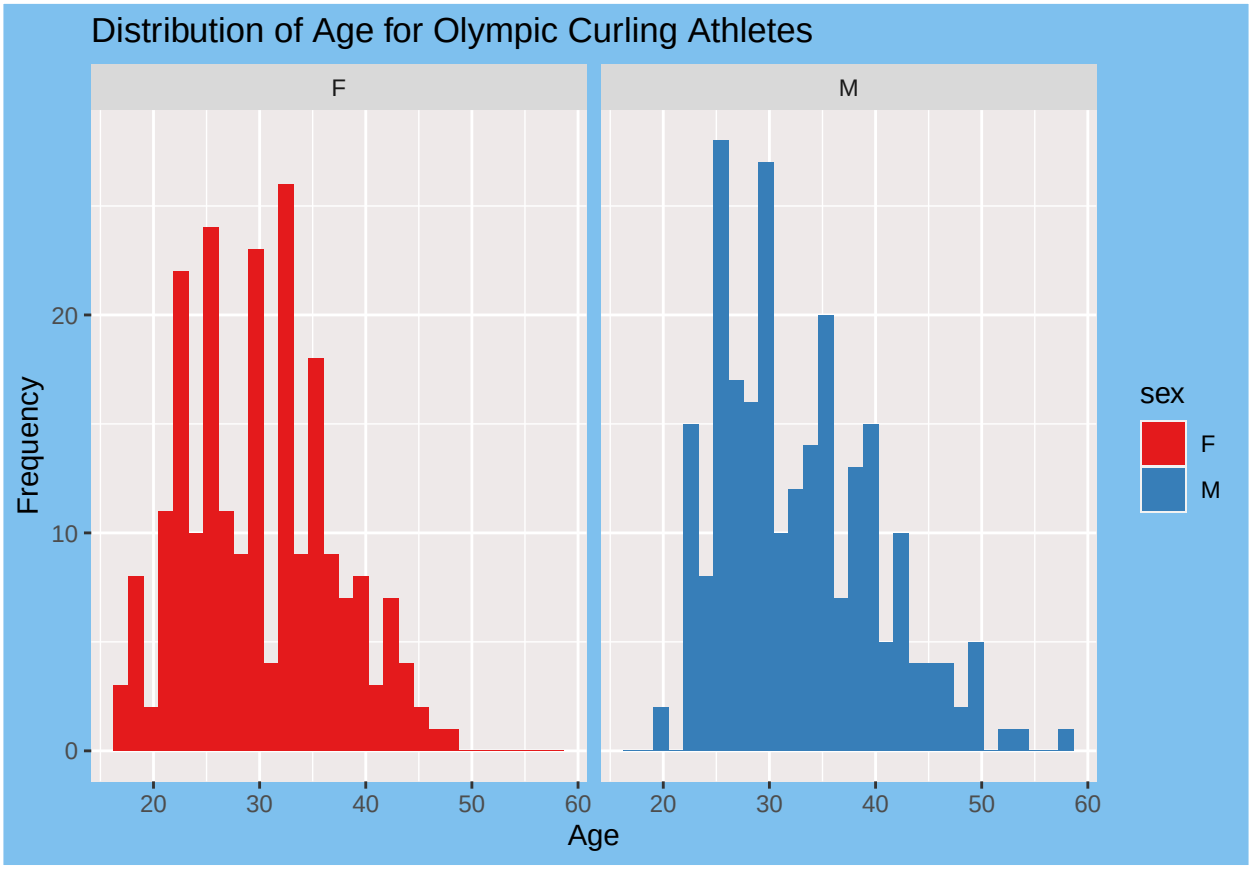


Olympic Curling Ages by Sex

Curling had the highest average athlete age compared to all the winter Olympic sports. The females had a lower average than the males with reason to the difference in maturity ages. However, the sport of Curling is historically an older crowd such as even older than 50 years old competing in the Olympics. Upon further research I had concluded that Curling is a sport that does not take vigorous physical ability such as hockey. I have read that Curling was a sport where you could play a match and then sit back and have a few beers with some friends after. It was more of a social event than a mainstream sport. This could indicate as to why the average age of the sport was significantly higher compared to the rest.

```
ggplot(olympics_Curl, aes(x = age,
                          fill = sex))+
  geom_histogram() +
  facet_grid(cols = vars(sex)) +
  labs(title = "Distribution of Age for Olympic Curling Athletes ",
       x = "Age",
       y = "Frequency",
       col = "Sex") +
  theme(plot.background = element_rect(fill = "skyblue2"),
        panel.background = element_rect(fill = "snow2"),
        legend.background = element_rect(fill = "skyblue2")) +
  scale_fill_brewer(palette = "Set1")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



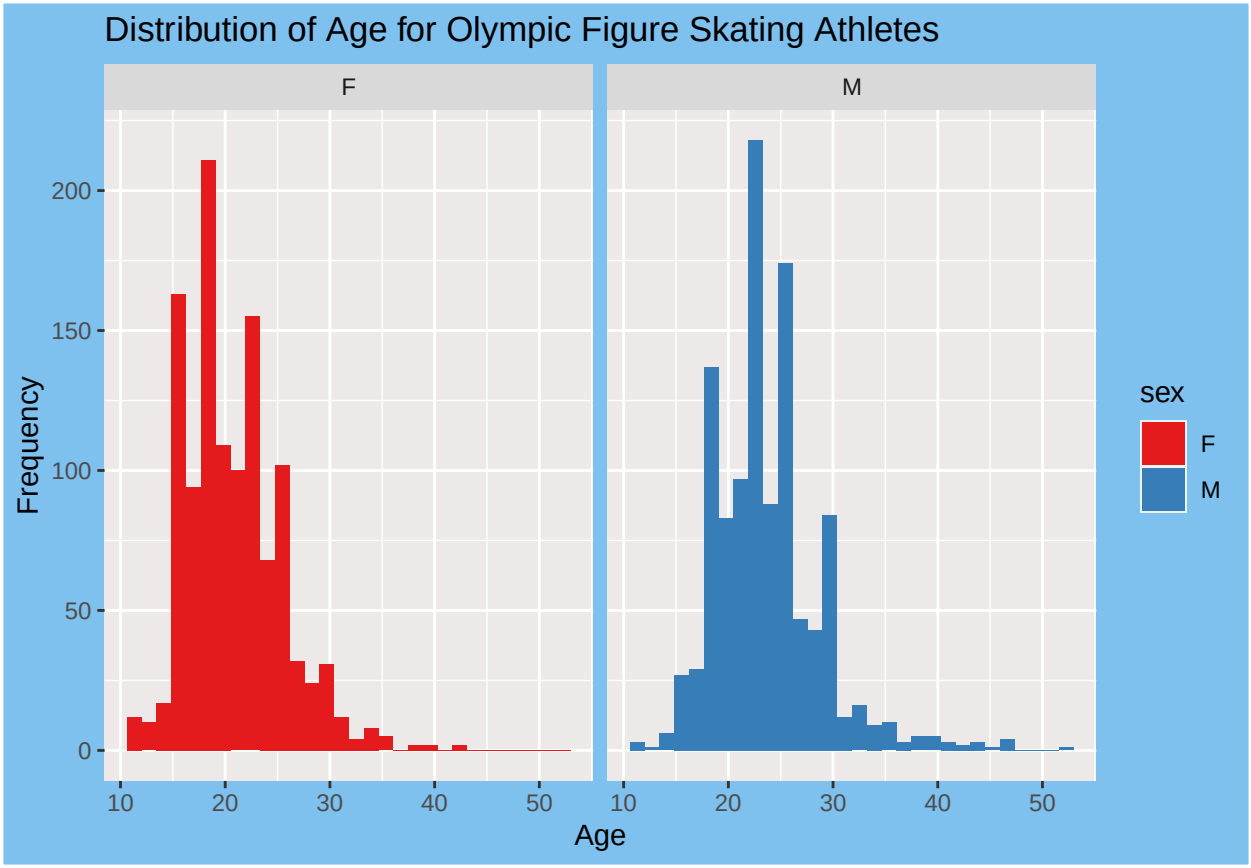
Olympic Figure Skating Ages by Sex

Figure Skating was the youngest average age of athletes in winter Olympic sports. The females still had a much lower average compared to the men as depicted below. After researching why the age is historically low, I concluded that it had to do with maturity and the fact that it was easier to perform certain tricks at a younger age. Some jumps such as a quad jump is much easier for a 15 year old compared to a 20 year old because they have not grow to the point where they need to require more balance for themselves. This concluded why so many athletes for females fell below the 20 year old label. Overall, the difference in average age between males and females in Figure Skating is clear due to the maturity level. The sport itself is requiring younger athletes to have a better chance at success.

```
ggplot(olympics_FS, aes(x = age,
                        fill = sex))+
  geom_histogram() +
  facet_grid(cols = vars(sex)) +
  labs(title = "Distribution of Age for Olympic Figure Skating Athletes ",
       x = "Age",
       y = "Frequency",
       col = "Sex") +
  theme(plot.background = element_rect(fill = "skyblue2"),
        panel.background = element_rect(fill = "snow2"),
        legend.background = element_rect(fill = "skyblue2")) +
  scale_fill_brewer(palette = "Set1")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 24 rows containing non-finite values (stat_bin).
```



Conclusion

In all, the Olympic data set provided information and I was able to discover and analyze a specific interest that interested me such as the average age of athletes in winter sports. Furthermore, I would like to analyze to see if over the years the ages have a trend of younger for curlers and figure skaters.

Sources:

1. https://www.nytimes.com/2018/02/18/sports/olympics/curlers-fit-buff.html
2. https://triblive.com/sports/the-age-old-question-in-figure-skating-how-old-is-too-old/#:~:text=But%20the%20reality
3. https://github.com/rfordatascience/tidytuesday/blob/master/data/2021/2021-07-27/readme.md