Exercise 1 Report: *Visualizing Gender Disparity in Winter Olympics (1924-2014)*

Eleanor Hope-Bell & Ifit Mirza

Motivation: The Olympic Games[1] are leading international sporting events across two competitions, summer and winter, that alternative every two years. Thousands of athletes from around the world participate in a number of sports with a multitude of events in each sport. More than 200 nations participate[2] in the Olympic Games and new sports and events are being added continually.

The Olympic Games is an event during which the world unites to celebrate athleticism. The International Olympic Committee (IOC) states that its aims is to promote sport competition that is free of any discrimination, "in a spirit of friendship, solidarity and fair play."[3]

Even though the Olympics officially began in 1896, women were permitted to compete from 1900 onwards and initially could compete in only three sports namely, tennis, golf, and sailing. The first Winter Olympics was held in January 1924 and women were eligible to compete in one sport, namely skating and in that sport only two events: figure skating and mixed pairs. Women athletes could compete in only skating until 1948 when skiing was opened as a competitive Olympic sport to women.

Women are now eligible to compete in all Olympic sports categories and women represent almost half of the athletes competing.[4]

Tasks: Using the data for the Winter Olympics from 1924 through 2014, the aim of the analysis was to understand how the gender mix has changed over the 90 year time period for each Olympic winter sport. Is the gender split similar in each sport and in each country represented? Does the country the athlete represents influence the gender split? Do women have equal access to [winter] sports? Are there economic or political factors in any country beyond its mere location that may help identify the reasons behind the gender gap in the Winter Olympics?

Visualization: The visualization in the first chart for the Winter Olympic data is a stacked bar chart that displays the total medals awarded in each of the seven winter sports. Each bar is further broken out by gender. The bars are placed on a common scale and the total length is used to express the total number of medals awarded.

The light blue-grey box below that appears slightly to the right of the horizontal box for each sport expresses the values of total medals won by gender. Within each stacked bar, there are two distinct and bright colors that represent the total number of medals awarded to each gender provided in the dataset which further identifies the categorical attributes.

Within the chart, each sport is placed horizontally and sorted alphabetically. The legend is located outside the edges of the chart to coordinate the colors with the genders.

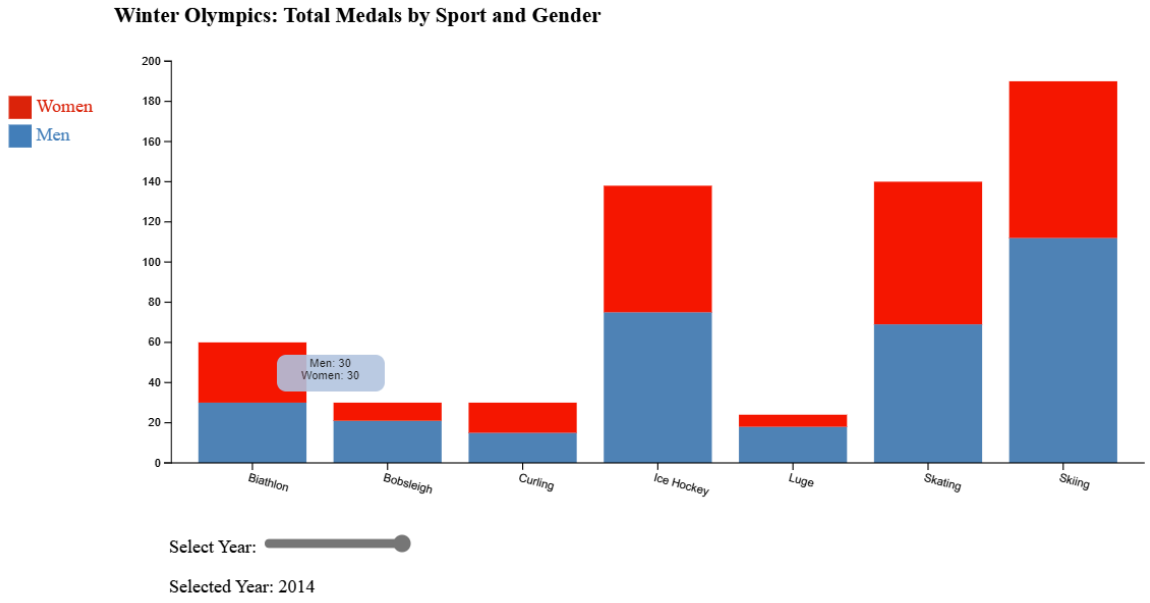
---

[1] Olympic Games - Wikipedia

[2] Fun fact: There are 195 universally recognized countries in the world (WorldAtlas) but the Olympics has 205 teams represented. How is that possible? - Quora

[3] The importance and influence of Olympic Games (Britannia)

[4] There are two sporting disciplines that are solely for women and men are not eligible: synchronized swimming and rhythmic gymnastics.Women at the Olympic Games



In HTML, a user is able to hover over the bar and have the total number of medals awarded to men and women in each sport is identified. In HTML a user can enable the slider below the chart to adjust fluidly through different years. The interactivity of the slider in the chart allows for quick access to view changes in gender split and value by winter sport throughout the complete years provided in the dataset between 1924 to 2014.

The genders are categorical and are represented with a qualitative scheme since it does not imply magnitude differences between the sports. Using a qualitative scheme of two classes, the two bright colors serves as a useful tool for distinguishing the color hues for each category.

As more women have become eligible and participate in Winter Olympic sports, the chart shows the general trends of the increased proportion of total medals that were awarded to women in each Winter Sport throughout time.

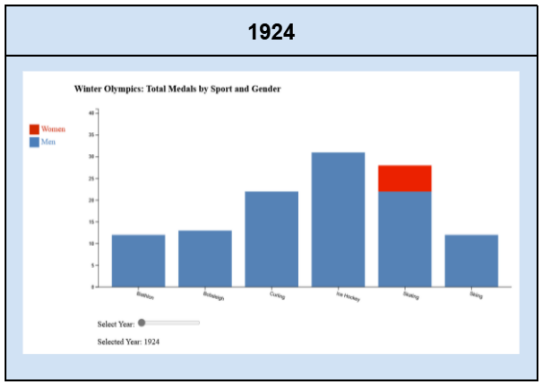
The Expressiveness principle is met because the visual encoding expresses effectively, is ordered and identifies the categorical attributes. It also correctly matched (both ordinal and quantitative)

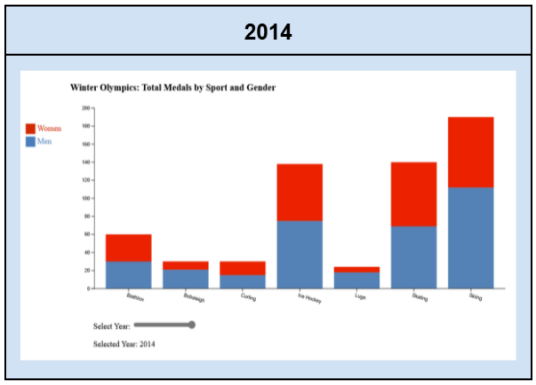
The Effectiveness principle is met because the visual matches the salience of the channel that are encoded with the most effective channel in order to be most noticeable.

Data Omission: No data from the original data set was augmented. However, there were a few attributes that may have explanatory contributions through visualizations. The data omitted for the visuals above includes:

1. **countries** that each athlete represented, the events in each sport, and;
2. **name** each athlete.

Furthermore, the dataset has two columns that represent the same meaning so representing both would be redundant. One column is 'Country' that names the country that the athlete represented as a full string name and a second column representing the country as a three digit string 'Code'.





The dataset was in one file called ‘olympics.csv. The tools that were used to create the visual chart were:

1. Text Editors:
    a. CodePen; and,
    b. Visual Studio Code
2. Hypertext Markup Language (HTML)
3. JavaScript (JS)
4. Cascading Style Sheets (CSS)
5. D3.js for visualization.
6. Jupyter Notebook Python (JPYNB)
7. Repository: Github.com

Improvements and Conclusions: There are many improvements that could be made to visualize the data in order to extract additional insights. The data, for example, is limited to the Winter Olympics. Adding the Summer Olympic data would enrich the ability to draw more meaningful conclusions on the gender disparity of Olympic sports by country.
In addition, creating an **interactive** world **map** to represent the gender split for all the medals by gender in each country and by year would may help more easily identify the access women have to winter sports and/or whether the political landscape in that country influences women’s participation and therefore ability to attain a medal and how it has changed over the years.

Taking the data even further down to the athletes themselves and understanding if there are outliers that need to be considered and removed from the analysis? This analysis could be done with a visual representation and filtering out known outlier athletes. Lastly, contextualizing the data with economic and political factors may also lead to important if not critical insights and conclusions.

Demo:
Use the following link to look at the visualization demo and the code.

https://codepen.io/Ifti007/pen/ZEXZzNN