*Raul Martinez - Ivan Ulloa*

Exercise 1 Report: *Gender Representation in the Olympics*

Motivation

     In recent years, availability, and access to information over the internet has allowed the public to discuss and share their thoughts over a wide range of controversial topics. The issue of gender inequality has been frequently showcased and empowered in the media where the underrepresentation gap of women in many aspects of society is discussed. Historically, women have faced an uphill battle across all fronts and their inclusion into sports and athletics has not been the exception.

Tasks

     A sample dataset of the Winter Olympics covering the years from 1924 through 2014 is to be explored and analyzed to answer the following questions: Has women inclusion in the Olympics increased or decreased? Has women participation in the different sports increased over time? Is there a noticeable change in the rate of increase or decrease in the ratio between genders?
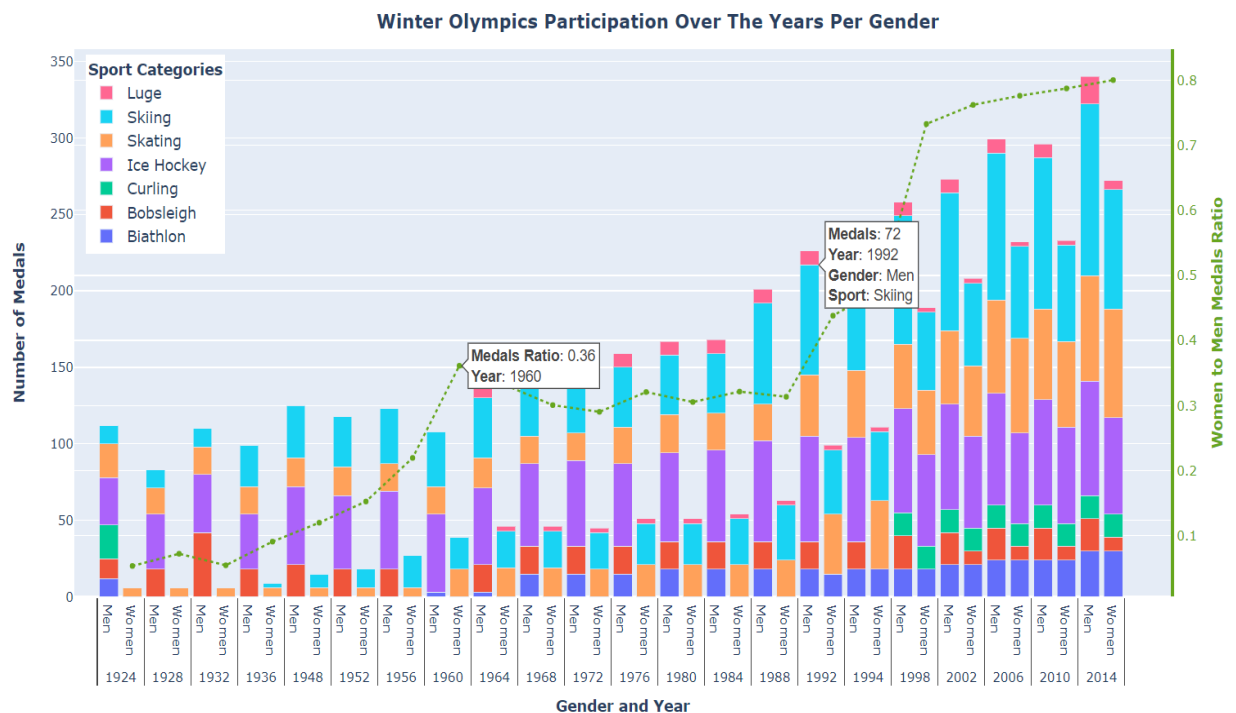
Expressiveness of design

     The visualization components are designed to show trends related to gender imbalance for the number of medals given during the winter Olympics, and how the differences have progressively become leveled over time. It also provides insight for available and emerging sports participation for each gender.

Raul Martinez - Ivan Ulloa

Data analysis and exploration revealed a linear and exponential pattern for the number of medals given to men and women, respectively. It also shows large differences per year for Women-to-Men medal ratio where Men always get more medals, and this is likely because of more participation, but as time goes by this ratio increases somewhat linearly as more medals are generated for Women. The largest increase occurred in the years of 1960 and 1998 from 22% to 36% and 48% to 73% respectively. The data also implies that available sports for participation has increased over time. From all seven known sports in the latest data (2014) women would only participate in 1-2 in the early years which included Skating and Skiing while men participated in 4-6 sports and this pattern persisted for almost 40 years, since 2002 both now participate in all sports. The least represented sports over time have been Luge and Curling. Luge is the last sport introduced in 1964 for both genders and Curling started in 1924 for males only and was suspended until 1998. All these observations are encoded in the data and the attributes described below demonstrate how the design works together to convey such information.

The dataset attributes are expressed with two major channels. First, stacked bars are used to show the total number of medals per gender as well as the fraction per sport with a side-by-side comparison per year, each bar is colored by the different sports at the time, and gender categories are ordered with Men on the left and Women on the right. Second, a trend line is superimposed to show the women to men ratio of medals given for each year to decode gender participation differences. And lastly, a multi-level x-axis is used to group related information together and years are sorted increasingly to show trends from past to present.

Effectiveness of the solution

To evaluate the trends and ideas previously mentioned among the number of medals given, sports categories, and gender participation over time, several different channels were accurately considered for each attribute within the visualization.

The increasing pattern for the number of medals over the years is organized using gender categories following a consistent order and years are sorted in an ascending format to facilitate gender comparisons over time. Each bar has at most seven stacked colors pertaining to each sport, the legend defines the color position for which the sports are represented inside each stacked bar. This attribute makes identification of existing and missing sports simpler, especially for years where not many sports are present, or they are included but are represented with a very thin stack. For example, if all sports exist it is expected that Biathlon is always the first stack at the bottom and Luge is the final stack at the Top. In addition, the thin white gap within each stack provides better discrimination between the colors. The dotted line chart provides a visual estimate for the upward trend of gender medals ratio over time where the thin segmented format allows the data behind to not be obstructed, additionally, circled markers provide 2-D spatial positions to each data-point and they give a good reference to the years in the x-axis.

Raul Martinez - Ivan Ulloa

The visual has a variety of well-contrasted light colors for the bars and a darker color for the line. First and secondary y-axis are differentiated with coloring and thickness; the secondary axis is matched with the trend line in color. Horizontal white lines in the background allow relative comparisons between bar heights, as shown in the visualization some of the lines match the bar's height for better reference. White background for the hover tool is also adjusted to contrast with the plot's light-blue general background when the interaction happens, and metadata is shown. The legend is placed inside the left empty area of the plot to provide more space to the secondary axis on the right side, and a white background is used to provide contrast with the graph.

And lastly, font size is adjusted accordingly to be readable and bold font is used for the visual title, axis, legend, and hover data labels to emphasize text.

Interaction

The graph provides three main features that facilitate readability of the visualization. The first feature is the ability to hover over a section of any data point including bar chart and trend line, this will automatically display a small container which includes metadata for that object. The second feature is contained in the legend section where the names of the sports are located, here the viewer may toggle the names of the sports to hide or display the data associated with them. The third feature allows the user to zoom in to a specific section of the chart by creating a box around it. These three features are included to reduce clutter or to obtain targeted information.

Conclusion

Overall, the visualization effectively provides multiple layers of information that the viewer may use to explore the relationships for gender participation, number of medals, and sport categories for the winter Olympics. Women inclusion has increased over the years and is quite evident from a high-level view of the visualization, but it has not been a simple task. Understanding in depth gender balance trends over time, women incorporation among the different sports, and the rate at which this is occurring is outside the boundaries of the data available and it requires more historical context, however, the data shows women today have more representation and inequality has been reduced as compared to twenty years ago. It is evident that much of the progress has been made in only the last couple of decades.