Assignment 1

Investigation of Skating Events in the Winter Olympics

Chris Dong

MSDS 622: DATA VISUALIZATION

CONTEXT

This data pertains to the Winter Olympics from 1924 to 2006. It consists of 2311 rows and 11 columns. Each row records a medal awarded from the Winter Olympics. A description of a subset of the variables can be seen in the following table:

VARIABLE NAME	DESCRIPTION
Year	1924, 1928, … , 1992, 1994, … , 2006
Sport	Skiing, Skating, Biathlon, etc. (7 unique)
Discipline	Figure skating, Bobsleigh, Ice Hockey, etc. (15 unique)
Gender	Gender of medal recipient (Men, Women)
Medal	Type of medal received (Bronze, Silver, Gold)
Country	Country name of medal winner

MOTIVATION

There are interesting trends that occurred or changed over time with respect to the event skating. With a total of 5 visualizations; the topics to be explored pertain to event popularity, gender, and medal count.

1. **Event Popularity:** This graph shows that Skiing and Skating were not so popular when the Winter Olympics first started but quickly became the most popular sports today in terms of the number of events allocated to them. The events Biathlon, Bobsleigh, Curling, Ice Hockey, and Luge were grouped into one because they did not change significantly by themselves and to reduce clutter. A line plot is used to show the growth of the sports from 1924 to 2006; specific points are also added to make it easier to see.

2. **Type of Skating Event:** A grouped bar plot is chosen because it is effective in showing the number of men and women for a specific skating event. When the skating event only has one color, it means that the event is unique for that gender; for example, only women participate in the 3000m event.

3. **Discipline of Skating Event**: This next plot focuses on the disciplines of the Skating events and how they changed over time with respect to the number of Men and Women participating in them. It is represented with a stacked bar chart where each pane is a separate discipline within Skating.

4. **Medal Count:** A map visualization is created to show the number of Skating medals won by each country. The visualization is the world map where darker shades of green represent a higher medal count and lighter shades represent a lower medal count.

5. **Medal Count - Skating:** The last plot, specifically for the event Skating, examines the distribution of medals for each country. It is also displayed with a stacked bar plot because it makes the most sense in order to show the overall count as well as the count of the categories of Bronze, Silver, and Gold.

APPROACH

In the first plot, the number of lines is limited to three to focus the attention to Skating and Skiing. Many of the other sports are overlapping otherwise, which is why they were grouped into Other. Although the rotation of the x-axis labels is not optimal, it looks less cluttered this way. For the second plot, clutter is reduced by removing two events that only have 3 occurrences -- "Team pursuit" and "combined(4 events)". For the third plot, only the top row is annotated because there is less variation between the different years. The clutter in the world map is minimized by using annotations to highlight some of the country names (instead of having the text overlayed). In the last plot, countries are filtered if they won less than 8 medals. Overall, attention is drawn with the use of color and text. Each text (that pertains to a category) is color-coded with the corresponding hex value in the visualization. Annotations are used as a way to give clarifications as well as the most important message to be learned from the visualizations.

TAKE HOME MESSAGE

The take home message is that the Winter Olympics in 1924 was definitely not the same as it was in 2006. Some events (skiing, skating) became much more popular. The proportion of men and women events have equalized over time; for skating in particular, women did not participate in Speed skating events until 1960 onwards. The diversity of events has increased over time. Interestingly, when the number of medals won by each country from 1924 to 2006 for Skating Events is compared to the overall medal count in the recent PyeongChang 2018 Winter Olympics Games, it is pretty similar. The top 5 from 1924 to 2006 in *Skating* are: United States, Russia, Germany, Norway, Netherlands. In contrast, the top 5 from 2018 in *All Sports* are: Norway, Germany, Canada, United States, Netherlands. Note that Canada was number 6 in *Skating* from 1924 to 2006. Also, Russia was banned from the 2018 Olympics, which may have affected the results.