Winter Sports Search Interest

SkillBuilder

This SkillBuilder uses data from Google Trends, which reports the relative popularity of search terms used on the search engine over time. In this particular dataset, you have been provided search interest about nine winter sports from 2004 to early 2022, from searches made in the United States.

You might notice that the Google Trends data values range from 0 to 100. These numbers do not correspond with the raw search interest, but instead represent *relative* search interest. According to the FAQ, each search term's raw popularity score is based on the relative share of searches for that term compared to the given region's overall search volume. The raw popularity values are then scaled from 0 to 100 based on a comparison to the most popular month and search term being queried.

As a result, the month and sport with the highest search interest in the dataset will be given a score of 100. Every other data point will be scaled against that value: a 50 means that month and sport had about half the search interest as the highest point. This means that we can compare the relative interest of various sports based on their average values and their trends over time. And in this SkillBuilder, we'll explore those trends through the use of visualizations.

1. Create a line chart that captures the trend in interest levels in "curling", over the time span recorded in the data.
    - Don't just plot the data, don't forget to also format your chart nicely, as though you were to present it to someone else to demonstrate your findings. In particular, consider how you would format:
        - Title
        - Y-axis limits and tick marks (What is a meaningful range of values?)
        - X-axis tick marks (What is a meaningful gap between ticks?)
2. Summarize what the chart you created can tell us about search interest in curling. Are there any repeating patterns? Where does this pattern come from?
    - Don't think you need to answer this question just based on the data alone. Make use of an internet search engine to dig deeper, by using curling and a month as search terms.

3. Create a line chart that captures the trend in interest levels in “luge”, over the time span recorded in the data.

4. Compare this chart to the previous one (from question 1). How does the overall search interest in luge compare to curling? Are there any notable outliers? If there are, what made those months so notable?

5. Create a line chart that captures the trend in interest levels in “hockey” and “snowboarding”, over the time span recorded in the data.

    - This line chart will include two lines on it, one for each sport. Don’t forget to select both columns of data to create this chart. Make sure the legend labels the lines according to their sport.

    - As a further hint, make sure that the values aren’t stacked. You should see some overlap and crossing between the lines for the two sports.

6. Describe the general trend in search interest in hockey and snowboarding based on the chart. How do they differ from the previous charts, and how are they similar?

7. Create a bar chart that captures the overall interest in the remaining sports that have not yet been plotted: speed skating, Nordic combined, biathlon, alpine skiing, and bobsleigh. You should use a stacked bar chart, only on months in which the Olympics took place.

    - To easily select just the Olympic-year months to create the chart, you can use a filter on the months column. You can isolate only the Februaries for each year by putting “-02” in the filter search box!

    - The final chart should have one bar for each sport, divided into sub-bars, one for each Olympics Games year. Make sure you consider the orientation of the chart and the colors used for each bar when polishing up your plot.

8. Which of the five Olympic sports plotted has the highest overall search interest across the five Olympic games months included in the data?

9. In which year did that Olympic sport (from question 8) have the highest search interest?