The Summer Olympic Games are an international athletics event held every four years and hosted in different countries around the world. Rowing was added to the Olympics in 1896 and has been in every Summer Olympics since. Rowing races in the Olympic context are typically regatta style, meaning that there are multiple boats racing head-to-head against each other in multiple lanes. Since 1912, the standard distance for Olympic regattas has been 2000m. The boat that is first to cross the finish line is awarded a gold medal, the second a silver medal, and the third a bronze. Over the course of its time as an Olympic sport there have been 25 different event entries.

In this dataset (<rowing_medals.csv>), the medals are counted as one medal towards each boat as opposed to each athlete in the boat. In looking at the total medals and total points for each nation, it is interesting to see which nations dominate in Olympic rowing. Additionally, looking at the overall distribution of the medals for all countries provides insight on just how lopsided medaling can be in rowing at the Olympic level.

|  |  |
| --- | --- |
| **Variable** | **Description** |
| NOC | The nation competing.  |
| medals | The total number of medals for that country in that event. |
| points | $$points=(3\*Gold)+(2\*Silver)+(1\*Bronze)$$ |
| gold | The total number of gold medals for that country. |
| silver | The total number of silver medals for that country. |
| bronze | The total number of bronze medals for that country. |

1. Create a histogram of *points* for all countries. Describe the distribution of points.
2. Obtain the summary statistics for *points* and fill them in below.

**Minimum**: **Lower Quartile:**  **Median**: **Mean**:

**Upper Quartile**: **Maximum:**

1. Check to see if there are any other outliers for *points*.
2. Create a bar plot of *NOC* ranked by *medals* to showcase the distribution of *medals* per nation. Describe any noticeable patterns in the plot.
3. What could be a reason for some nations winning more than others?
4. There is a lot of debate about how to best weigh the *points* for the different types of medals. Read [this link](https://www.topendsports.com/events/summer/medal-tally/rankings-weighted.htm) (https://www.topendsports.com/events/summer/medal-tally/rankings-weighted.htm) about different medal point weighing and decide on a method you think would be best or make up your own weighting system. How would the new method alter the data?
5. Using the gold, silver, and bronze variables, create a new *points* variable with a different name based on your idea for medal scaling. Create a histogram of the points with this new variable and comment on the skew.