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, 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 lob-sided medaling can be in rowing at the Olympic level.

|  |  |
| --- | --- |
| **Variable** | **Description** |
| NOC | The nation competing.  |
| medals | The total number of medals for that country. |
| 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. The below graphic is a histogram of *points* for all countries. Describe the distribution of points.



1. The summary statistics for *points* are provided below.



* 1. This table shows the top 5 nations ranked by *points*. Using the summary statistics, determine if there are any other outliers for *points*.



1. Using your answer to Question 2 draw a boxplot of *points*.
2. The below bar plot shows NOC ranked by medals to showcase the distribution of *medals* per nation. Can you think of any possible reasons why some nations win more than others?



1. In the bar plot in Question 4, notice that the German Democratic Republic (GDR) or East Germany is listed as a country. Since the end of the Cold War the GDR is no longer a country. What would you do with countries in the dataset that are no longer existent today? How would it impact the data?
2. 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#google\_vignette) about different medal point weighing and decide on a method you think would be best. How would the new method alter the data?