**Jonathan Young**

**Module 1 Challenge Writeup**

**4/25/24**

The sample in this data set had an extremely high preference for plays, even beating out the second two highest event types by a long shot. Most of the success for plays, rock music and web were from the US (being the most popular campaigns in both the US and in total). The US has seemed to take over and skew the data due to its popular demand for such events.

Limitations on the data seem to be based off outliers and missing information about other countries that are not included. The fact that plays had an unusually high success rate might be from a certain show gaining popularity but is the biggest outlier in this data set. Missing information from campaigns in other non-listed countries may limit knowledge for general planning. While that is the broad case, this data set does make it easier to see the positive and negative results in a specific country (such as the US) and can show researchers what is and isn’t successful. For sheets 1 and 2, a pie chart for tell a similar data story as the bar chart but have a more comparative view. Each pie chart would have to be specific with one column and not utilize the whole graph. A box and whisker chart would be able to show a similar shape as the bar chart and give researchers a different view on the comparisons.

In “successful” the mean is higher than the median. It is likely that outliers that make the mean that high. This also skews the distribution to the right. The “failed” category is similar but less extreme. The median may better summarize the data because it is not affected by outliers like the mean. The “successful” category had more of a variety because of the higher number of campaigns. The comparison of the min and max is also wider than the “failed” category.