Module 1 Challenge

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

According to the results, we can see that the theatre field has the most crowdfunding compared to other fields. To be more specific, plays have the most crowdfunding in all sub-categories. The percentage of outcomes is stable for most of the year, and there are some waves for the trends of successful percentage and failed percentage between Jun to Sep. The percentage of successful campaigns is usually about 50%.

1. What are some limitations of this dataset?

There can have some data are missing and there are two unexplained variables, spotlight and staff pick.

1. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

We can create a pivot table to see the relationship between outcomes of crowdfunding and spotlights or staff pick.

* Use your data to determine whether the mean or the median better summarises the data.

Considering there are many outliers and the data is spread out (According to high variance and standard deviation), the median better summarizes the data. Because mean will be affected highly by those outliers.

* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

There is more variability with successful campaigns than unsuccessful campaigns because both variances and standard deviations of successful campaigns are larger than unsuccessful campaigns. It makes sense because successful campaigns have a larger amount of data and outliers, and the difference between minimum and maximum is greater than unsuccessful campaigns.