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

1. The most successful parent category is theater.
2. When we analyze deeper into the parent category theater will find that “Plays” is the most successful sub-category.
3. The most successful campaign was around May – August.

* What are some limitations of this dataset?
  1. It’s not showing the budget and time that was spent on advertising and marketing to give a realistic data.
  2. In the OutcomeVsParentcategory table, it’s not showing the ration between the successful/failed to represent the relative amount of success.
  3. The currency is not standardized and that will not give a clear comparison between the data set.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  1. We could use a Goal vs. Pledged table with stacked column graph to show unrealistic pre-campaign crowd funding goals.
* Use your data to determine whether the mean or the median better summarizes the data.
  + For the successful campaign table, the mean value of 851 indicates that, on average, the number of successful backers is relatively high compared to the median value of 201. That means the successful backers may not be evenly spread, and there may be some outliers with high numbers of successful backers that are driving up the mean value. Inconclusion, the number of backers required for successful campaign is better represented by the median number because 50% of the successful backers are 201 or less.

On the other hand, the mean value of failed backers is 586, which is lower than the mean value of successful backers, indicating that the number of failed backers is relatively lower. Inconclusion, the median number of failed backers are better represented by the median number because 50% of the failed backer are 115 or less.

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

After we have calculated the variance, the standard deviation and the IQR we can analyze that there is more variability in the number of successful backers compared to the number of failed backers. This is because the variance and standard deviation for successful campaigns are higher than those for failed campaigns.

The variance of successful campaigns is 1,603,373, which is higher than the variance of failed campaigns, which is 921,574. This indicates that the successful campaign data has a wider range of values than the failed campaign data. the standard deviation of successful campaigns is 1,266, which is higher than the standard deviation of failed campaigns, which is 959. This indicates that the successful campaign data more outliers than the failed campaign data.

In addition to the variance and standard deviation, the IQR which is 1161 for successful backers is lower than the standard deviation which is 1266. That means the outliers are affecting the data negatively when we try to analyze it.

There are more factors that lead to the success of a campaign than numbers of backers like funding, strategies and time spent.