Module 1 Challenge | Conclusions

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

1. Success rates for crowdfunding campaigns tend to be higher during the months of May, June, and July. This would be summer in the Northern Hemisphere and Winter in the Southern Hemisphere.
2. Success rates tend to be higher when crowdfunding campaigns set goals between $1,000 and $35,000. It seems to trend that if the goals are lower or higher than this range, their success rates are noticeably reduced.
3. The theater category and more specifically the plays sub-category far exceed other crowdfunding campaign categories. The number of backers, total number of crowdfunding campaigns, and their overall success rates relative to the increased quantities tend to show that this is a comparatively lucrative category for fundraising.

***What are some limitations of this dataset?***

1. This dataset is overwhelmingly skewed toward crowdfunding campaigns in the United States. Other countries represent the vast minority of campaigns reported and therefore a complete international representation is not present here.
2. The date range of this dataset is limited to 10 years between 2010 and 2020. A larger date range would offer a more detailed report of trendlines and statistical data.

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

1. A pivot table establishing a correlation using the data launched and date ended and the amount of funds donations received would be beneficial to determine the average range of time needed for an active campaign to generate a certain amount of funds.
2. A pivot table showing the correlation between average donation and category and sub-category could be beneficial to see the types of donations each industry generate interest for.

***Statistical Analysis – does mean or median better*** summarize ***the data?***

With this data set, the median calculation better summarizes the data because there are quite a few outliers on the high and low side which drastically skews the mean calculation.

***Statistical Analysis – is there more variability in successful or unsuccessful campaigns? Does this make sense? Why or why not?***

There is more variability in successful crowdfunding campaigns. This is clearly noticed with the higher variance and standard deviation calculations from the datasets. This does make sense and can be expected because for most unsuccessful campaigns, the number of backers are comparatively low. This is possibly one of the contributing or correlating factors as to why they failed. There are outlier campaigns with a high number of backers that failed which brings the mean and median up, but there are more results on the lower end of the backer count scale compared to successful campaigns.

However, for the successful campaigns, the variance in the amounts stated as the goal amount is directly related to the variance in how many backers are required to meet that goal. Lower goals on average require fewer backers and higher goals on average require more backers. This variance is expected and seen in the calculations of variance and standard deviation done on the datasets.