Excel week homework report

# What are 3 conclusions we can draw about Kickstarter campaigns?

1. Based on the data within the excel book there are a few observations first Theatre is by far the most popular category on Kickstarter however it also has one of the largest failure rates as well it would appear that with Theatre that the majority of campaigns were plays.
2. All journalism projects were cancelled this suggests that if you would like a successful campaign do not attempt a journalism project. Also, it seems to be a sign of the times most people today are unwilling to pay for news and this translates into the data.
3. December seems to be to have more failed campaigns as a proportion of all months this suggests people are less likely to back a campaign around the Christmas period. This makes inherent sense as Kickstarter campaigns are typically not returning goods immediately and consumers may be focused on other channels for example retail to purchase goods during these times.

# What are some of the limitations of the data set?

The first limitation I would like to discuss is the sample size and sampling method it appears that for some categories the data may have been sampled down and in others this may not have been the case. To illustrate this check tab ‘Category Pivot’ in the tab on the grand total column of the pivot table you can see that there are round numbers on all categories except for theatre and publishing. This suggest that the sampling method may have been flawed because it’s very rare to have such round numbers EG ‘520-film and video , 200-food ’ and then 1,393 for theatre. This may be due to the total number of campaigns available there may not have been enough total campaigns in theatre available to make up the total the data gatherer wanted to get and so they selected all the available projects. This is a problem because it suggests they did not gather the data proportionally and the data set may be skewed as a result.

The data set is small compared to the total amount of projects and as previously discussed may not have been gathered appropriately and would likely not be representative of the larger body of data.

# What are some other possible tables and or graphs that we could create?

I would create a table to illustrate the average goal and pledged amounts by category this would be useful as it would tell us how much we should select as a goal within our category as to not be over or under-priced more analysis would be needed to determine complexity of the projects, but this would provide a general price guideline.

# Bonus statistical analysis

The median summarizes the data more effectively because there are a lot of high range outliers in the dataset. There is significantly more variability in successful projects vs failed this makes sense because there are less high range outliers as these projects were generally less popular and were much less likely to attract significant media attention which would cause the number of backers to increase.