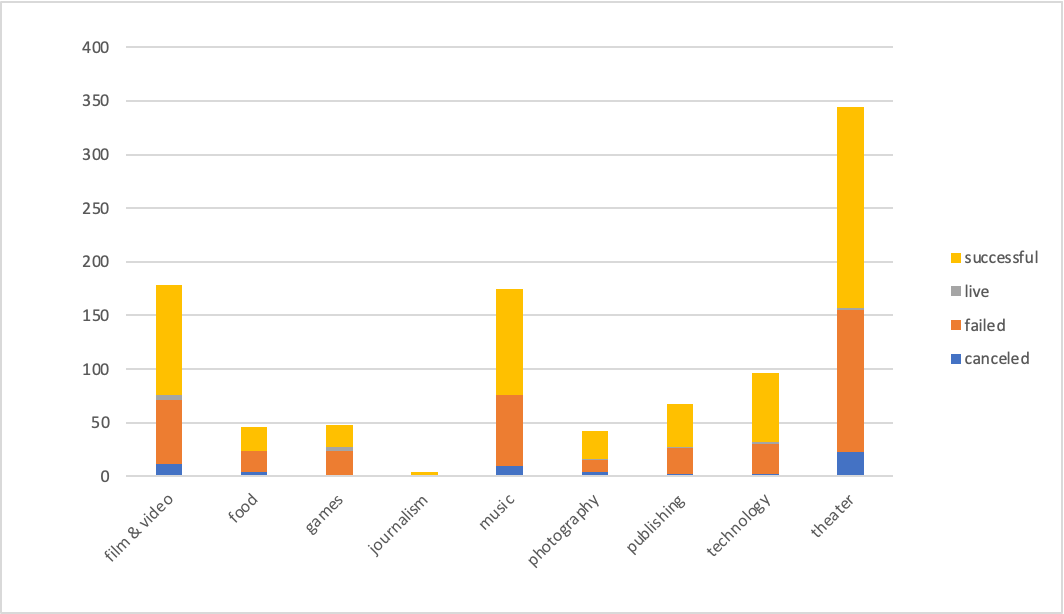
Excel-Challenge 1

**Conclusions:**

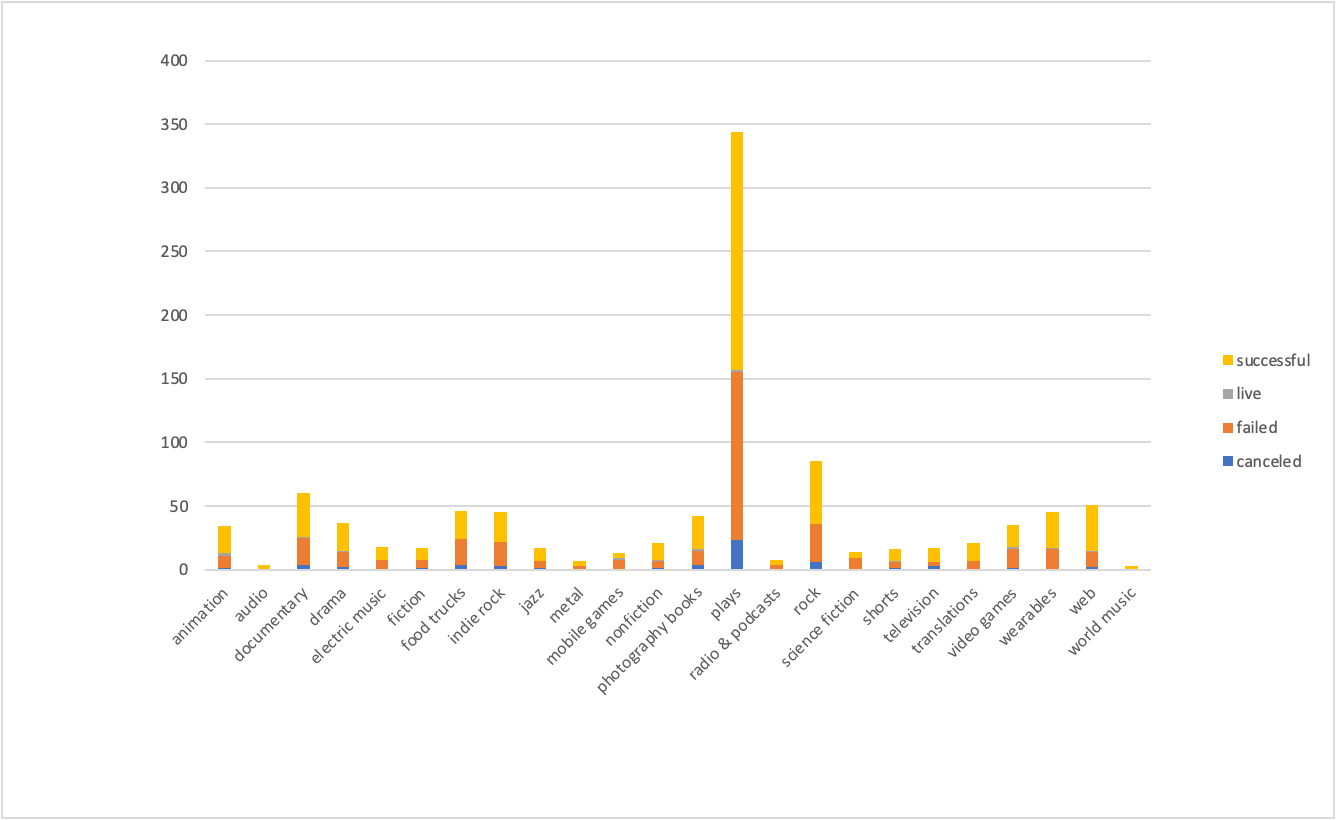
Conclusion #1: From the 1000 dataset of crowdfunding projects, there are 56.50% chances of getting successful, 36.40% chances of getting failed and 5.7% chances of getting canceled.

It purely shows that, if your project comes under the category of Technology(66.66%), Theater(54.36%), film & Video(57.30%) or Music(56.57%) then chances of getting your campaign successful is way higher the project comes under the category of game(43.75%) and food(47.82%).

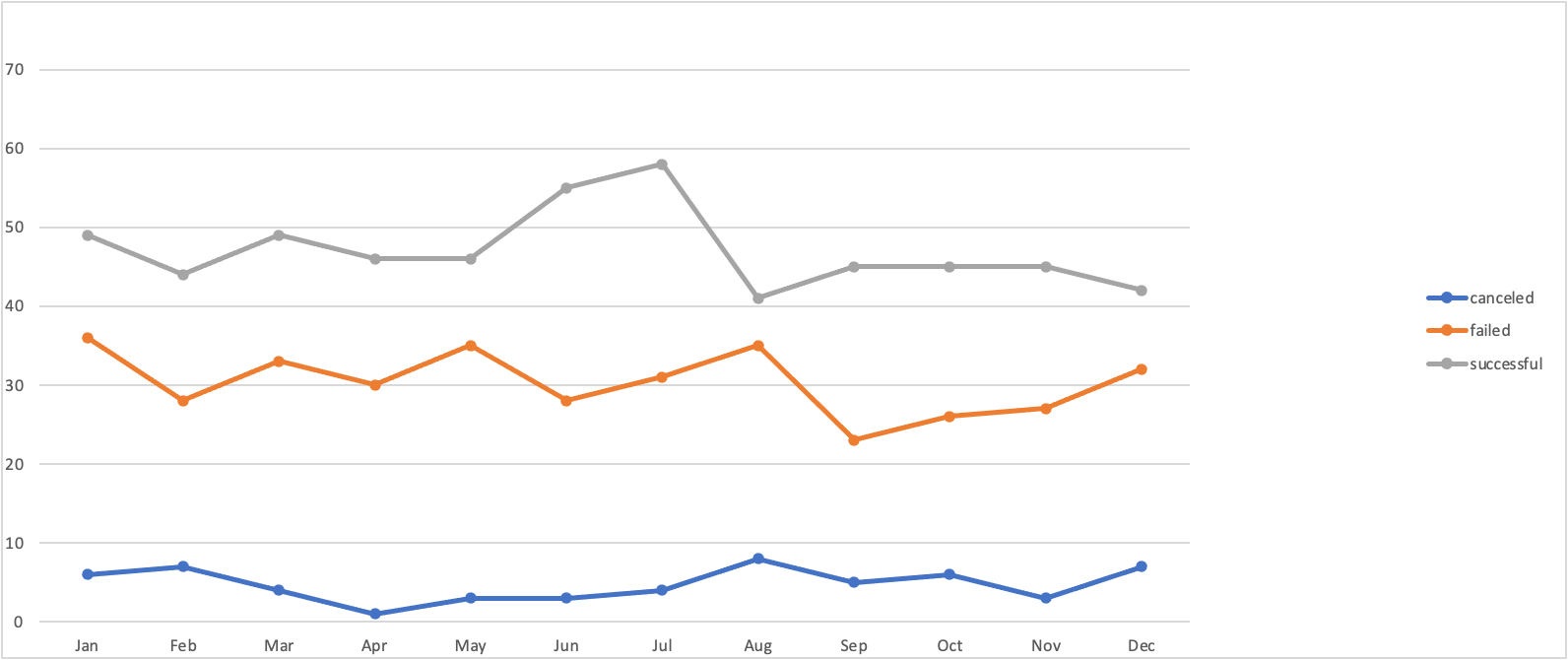
The data suggest that Journalism has 100% chances of getting success.



Conclusion #2: Looking at the following chart there are about 9 parent categories and 24 sub-category, out of which audio and world music have all the success rate. Sub-categories like plays,food trucks,rock,audio documentary, indie rock, Photography books and television have higher chances for getting it canceled than any other sub category in the data set which is provided for the analysis. That is something to keep in mind that if our campaign comes from those categories which have higher chances of getting it canceled then you can reevaluate your project to reach to the point that it atleast does not get canceled (but we need more information on that in our dataset).



Conclusion #3: Last, from the following chart I can conclude that projects in summer have a higher chance of getting successful than other months. The observation from the month of December shows that it has a higher number of canceled campaigns, failed campaigns and lesser number of successful campaigns.



**Limitations:**

* The data size of 1000 is not large enough to draw valid conclusions equally for all the categories and subcategories.
* The dataset needs more information like the reasons for which the campaigns failed, canceled or unsuccessful.

**Additional Values & Tables:**

* We could also use the goal amount to do the analysis of the successful campaign. The smaller the goal is higher the possibilities of getting it successful. Smaller the goal is, higher is the chance of getting backers. We could make the table of goals vs backers.
* We could also use the average donation to evaluate the data for each parent category.
* We need more information about the outcome of the datasets to evaluate upcoming projects to get it successful.

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

* The median summarizes the data better vs the mean.In this data set the Mean for both successful and unsuccessful campaigns is very far from the median and is higher than the median which tells us that the data is right skewed and has outliers.

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

* Describing Variability with Variance:

The variance of both successful and unsuccessful campaigns are very high, which tells us that both the categories have higher variability but The variance of successful campaigns is greater than the variance of unsuccessful campaigns which shows that the variability in successful campaigns is more than the variability in unsuccessful campaigns. There is another reason the variance of successful campaigns(Count:655) is higher because it has almost double the sample size than of unsuccessful campaigns(Count:365).

* Describing Variability with Standard Deviation:

The Standard Deviation of both the campaigns are also very high and the range of 1 Std.Dev. from the mean is also very big which tells us that both the projects have high variability and some outlier data. But, successful campaigns have higher standard deviation than unsuccessful campaigns which means that the successful projects have higher variability.

As stated above, the second reason for higher Standard Deviation of successful campaigns(Count:655) is the size of the sample set compared to unsuccessful campaigns(Count:365).

* Describing Variability with Mean and Median:

Mean and Median are away from each other means there is more variability in the data set.

1.In successful campaigns the difference between mean(851.51) and median(201) is 650.15.

2. In an unsuccessful campaign the difference between mean(585.62) and median(114.5) is 471.12.

Which concludes that the variability in successful campaigns is more than the variability in unsuccessful campaigns.