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

By filtering the "Outcome of Campaigns Based on Category" and "Outcome of Campaigns Based on Sub-Category" stacked-column pivot charts by country, we can see that a significant percentage of the 1000 crowdfunding projects being analyzed correspond to campaigns in the United States. This indicates that crowdfunding is significantly more popular and widely used in the United States, than the other countries being analyzed.

Another conclusion we can draw from the data, by looking at the "Outcome of Campaigns Based on Category" pivot chart, is that crowdfunding platforms are very commonly used to support "theater" projects. This conclusion is consistent with the data provided in the "Outcome of Campaigns Based on Sub-Category" pivot table, since by analyzing the respective chart, we can see that the count of campaigns associated with "plays", the only sub-category of the "theater" parent category, accounts for a large percentage of the total count. Specifically, 344 out of the 1000 projects.

Analyzing the "Outcome of Campaigns Based on Date Created" pivot chart, we can see that, regardless of the month, there is a larger count of successful campaigns, than other outcomes. We can also see that most successful campaigns that were conducted started in the middle of the year, around the months of June and July. Considering that most crowdfunding projects take place in the United States, we can infer that individuals feel more inclined to contribute to campaigns during the summer.

- What are some limitations of this dataset?

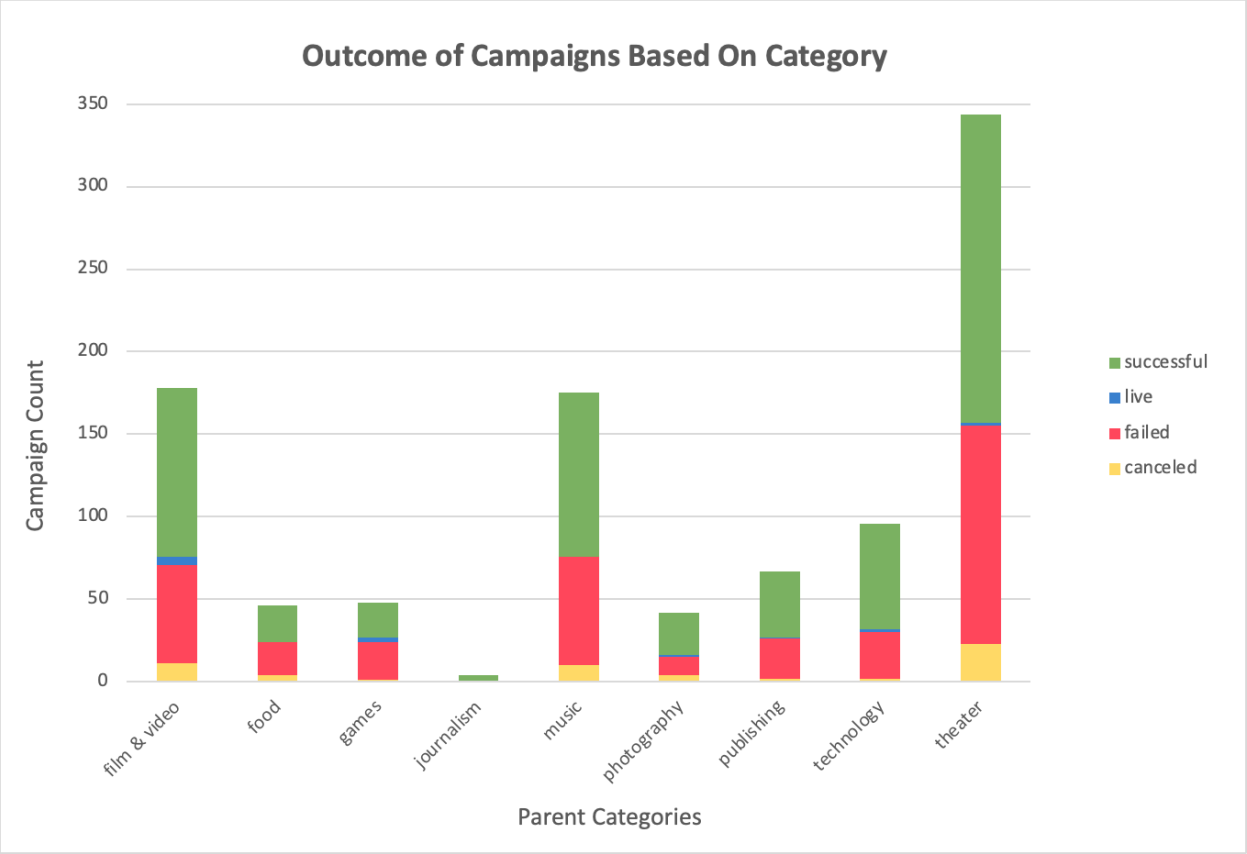
A possible limitation of this dataset is that it does not take into account all of the crowdfunding platforms that currently exist throughout the world. The outcome of a crowdfunding project is likely influenced by the platform used, as more popular platforms might reach a larger audience, and ultimately generate better results. This has not been taken into account when analyzing the outcome of specific categories and sub-categories. Additionally, assuming that the dataset corresponds to projects run by more popular platforms, the data might be skewed so that successful outcomes are more highly represented.

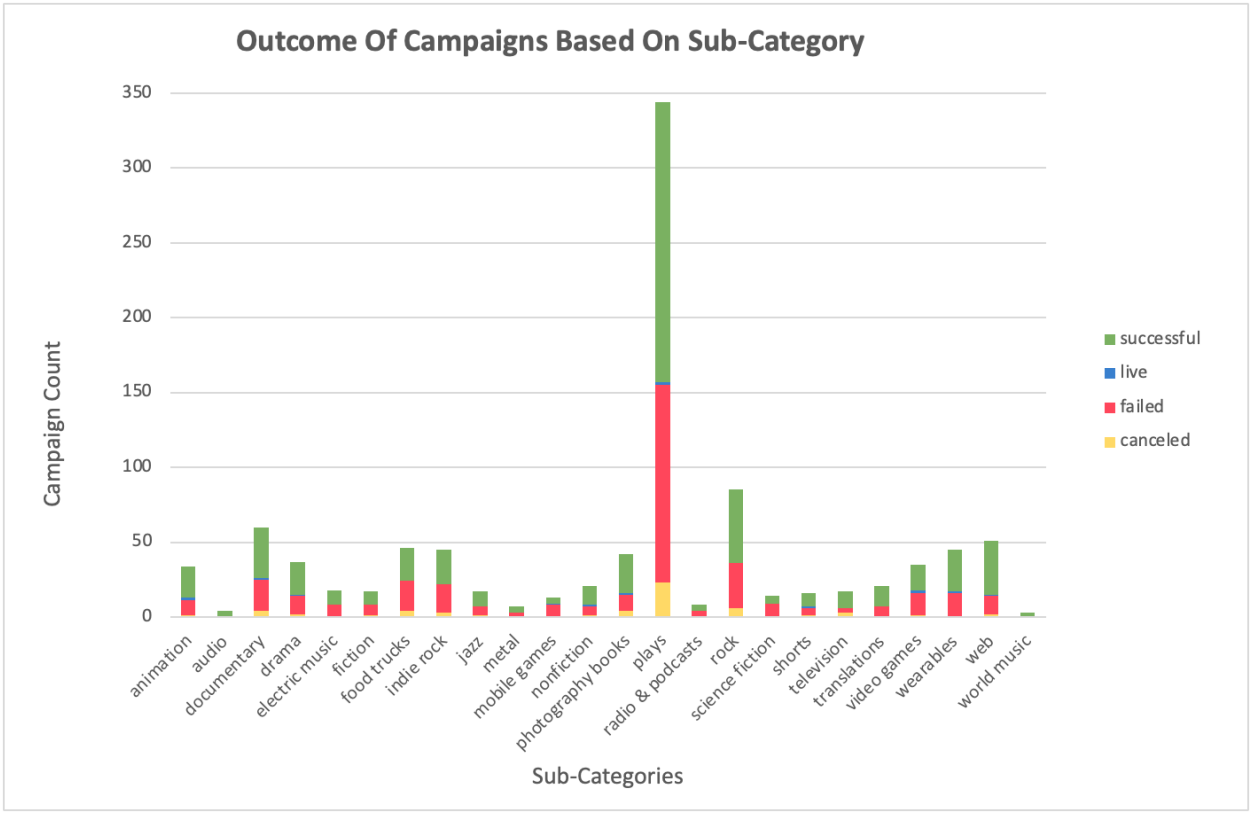
Although crowdfunding is organized through online platforms, and is therefore available by almost anyone in the world, the audience actually reached by certain campaigns is also likely to be influenced by how well it is publicized by the project's creator, which is something that can not be quantitatively measured and analyzed.

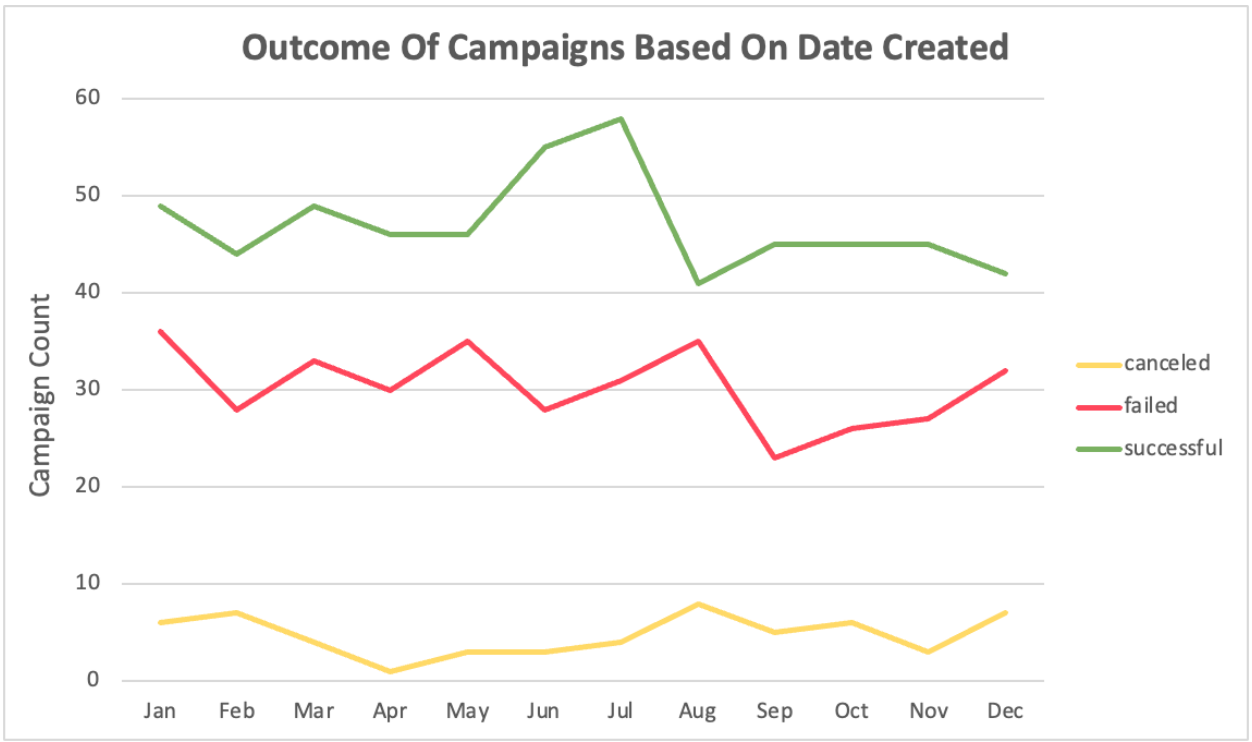
- What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

By creating an additional column in the dataset indicating how long each project ran for, we could analyze if the length of campaigns have a correlation with their outcome. This information could be taken into consideration when planning a project, and determining its duration to maximize the chance of success.

Campaign outcomes (success/failure) correspond to the percent funded, and ultimately to the established goal for each project. With the data provided, we can also analyze trends in the amount pledged, as well as the count of backers, and the average donation for each respective campaign category. This allows us to analyze how much support specific campaigns receive from the audience, based on factors that are not independent of whether or not the project achieved the proposed goal, which might not always be based on realistic expectations.




Outcome Of Campaigns Based On Sub-Category
Campaign Count
350
300
250
200
150
100
50
0
animation
audio
documentary
drama
electric music
fiction
food trucks
indie rock
jazz
metal
mobile games
nonfiction
photography books
plays
radio & podcasts
rock
science fiction
shorts
television
translations
video games
wearables
web
world music
Sub-Categories
successful
live
failed
canceled


Outcome Of Campaigns Based On Date Created
Campaign Count
60
50
40
30
20
10
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
canceled
failed
successful

- Does the mean or median best summarize the data?

Utilizing interquartile ranges to determine outliers within the dataset, we can see that several campaigns surpass the outlier upper boundary in terms of backers count, in both successful and failed campaigns. From this information, and from observing that the calculated mean is significantly larger than the median, we can deduce that the dataset is skewed right. Since the data does not follow a symmetric distribution, the median is a better central tendency measure to summarize the data.

- Is there more variability with successful or unsuccessful campaigns? Does this make sense? Why or why not?

According to the statistical analysis conducted, we can see that successful campaigns have a higher variance and standard deviation, and therefore, more variability. This makes sense since several projects which were successful were funded by a significantly larger number of backers, than the calculated mean, which is consistent with the fact that campaigns backed by a larger number of individuals are likely to raise more money, and ultimately be considered successful. Since variability measurements take into account the difference between independent values, and the calculated mean, outliers result in a larger variance within the successful category.

Statistical Analysis	Successful Campaigns	Unsuccessful Campaigns
mean	851.15	585.62
median	201	114.5
minimum	16	0
maximum	7295	6080
variance	1603373.73	921574.68
standard deviation	1266.24	959.99