Homework #1

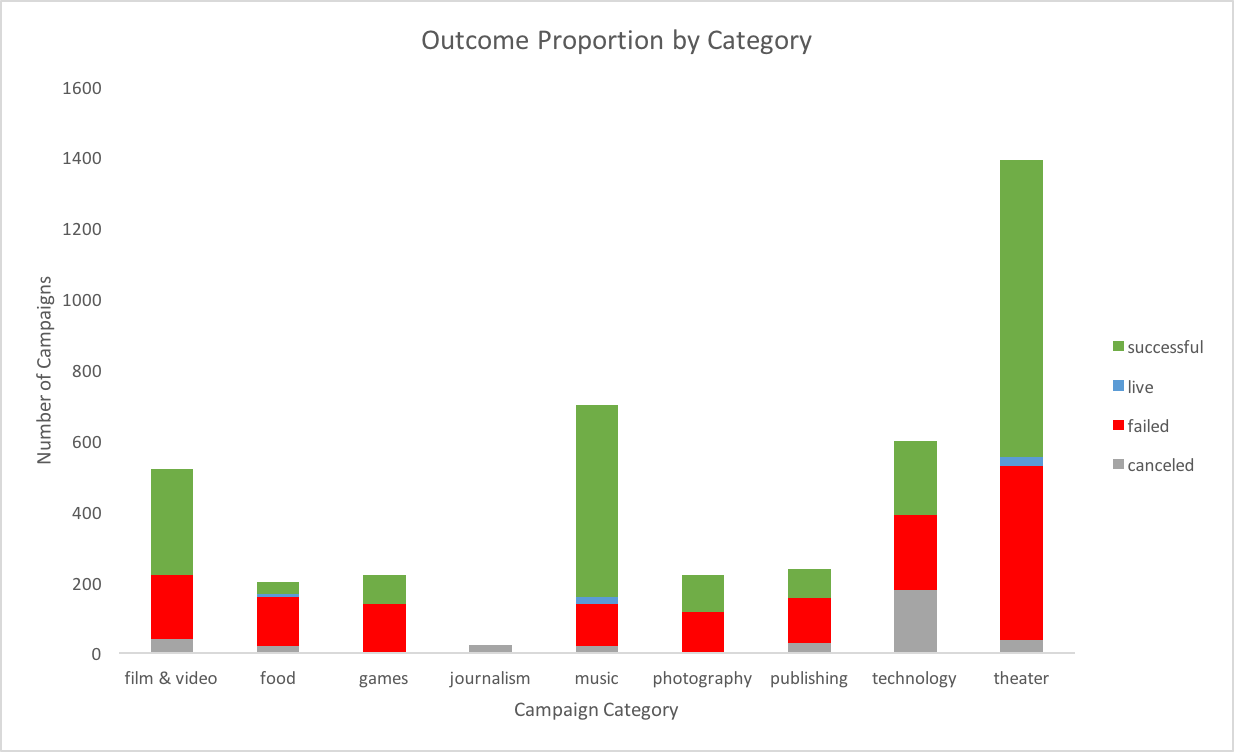
Max Parry

Kickstarter Analysis Report:

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?

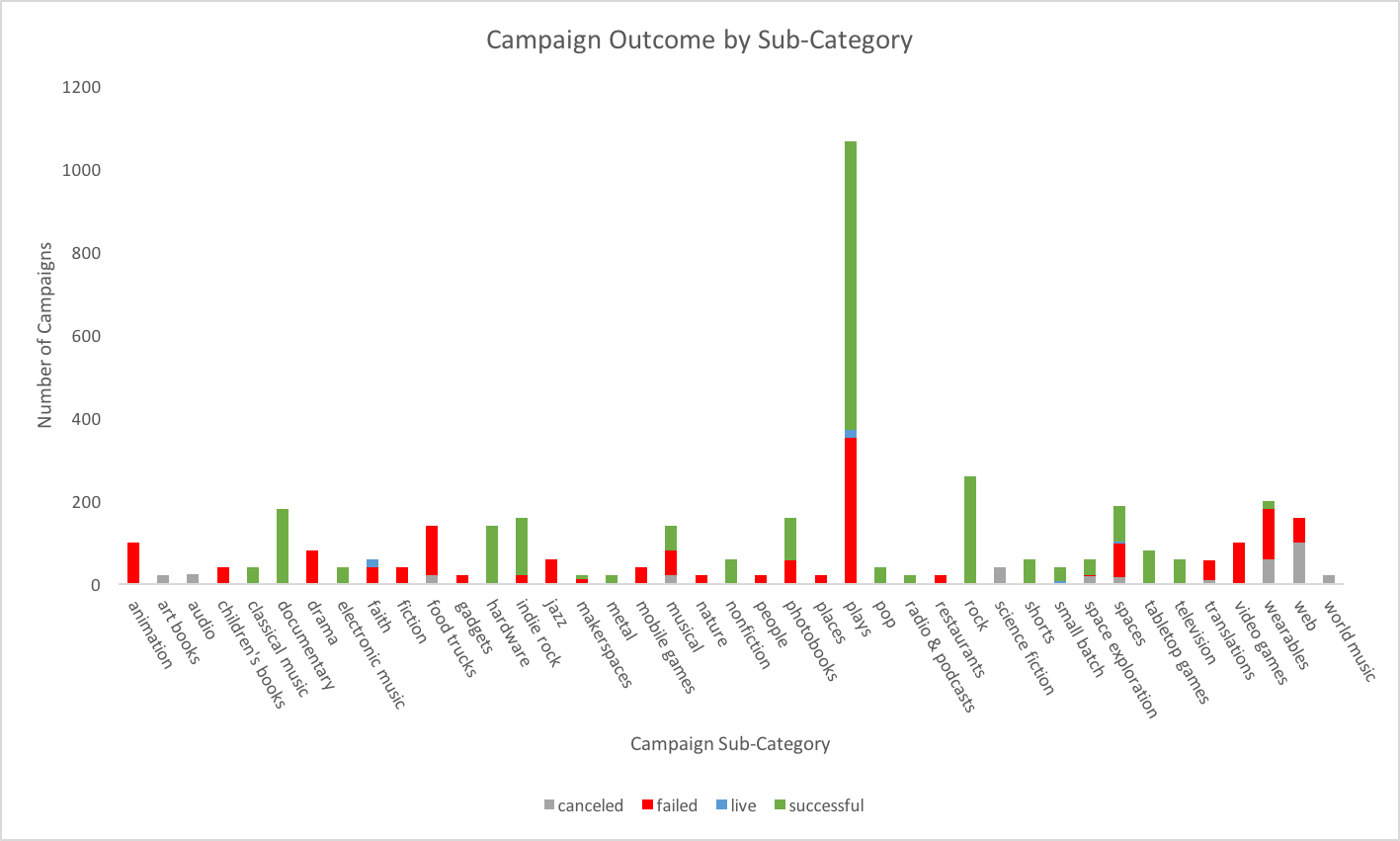
Results of the first visualization shows that some categories have a higher number of campaigns than others, and that success rate varies across categories. A quick look at the visualization below tells the following:

* Theater is the largest category for Kickstarter campaigns
* Journalism, food, games, and publishing have low rates of success
* Music has a high success rate



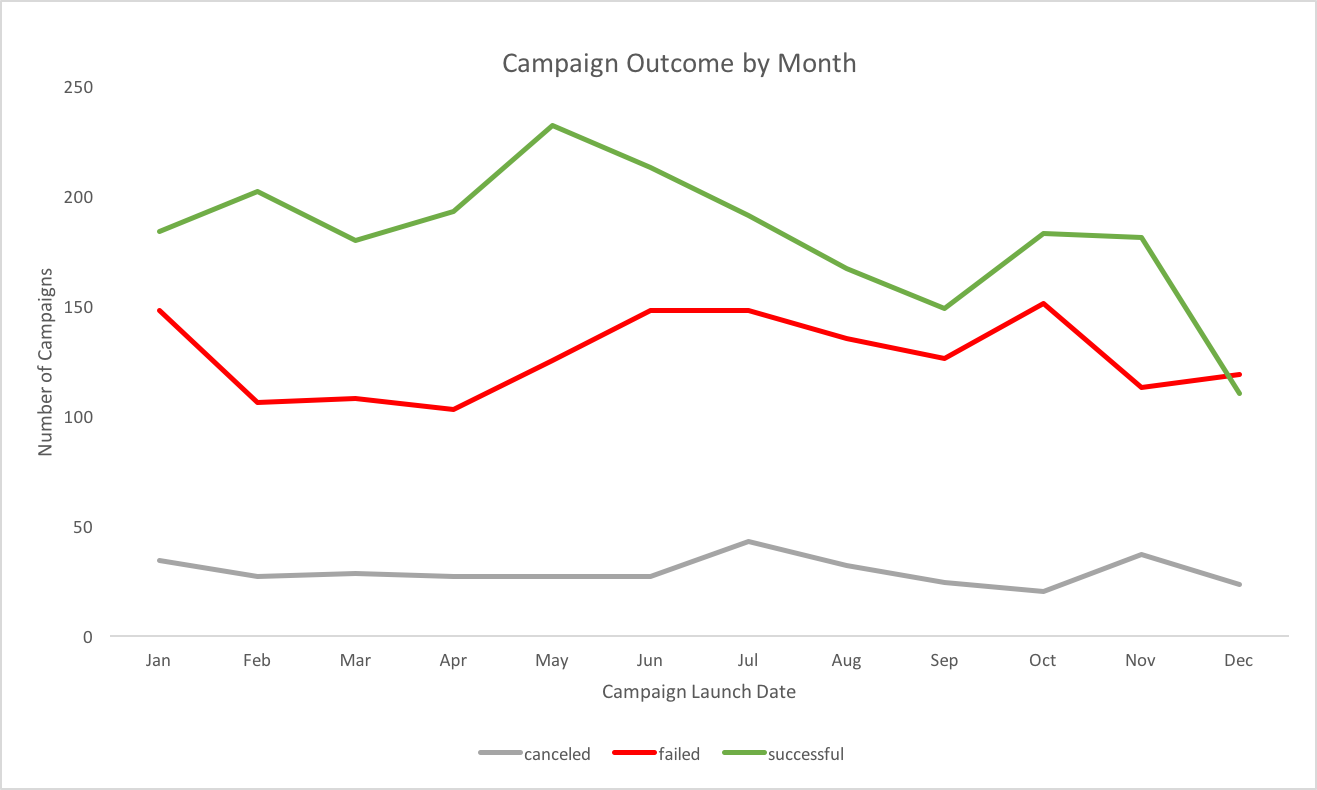
The next visualization shows proportion of campaign outcome at a more granular level. Review of the plot will show:

* Plays outnumber other campaigns by a wide margin
* Documentary, hardware, and rock stand out as categories that have both relatively high campaign numbers and very low fail rates
* Animation, food trucks, and video games rank among categories with a near 100% fail rate



The third visualization shows that success, fail, and cancellation rates vary across the year. Below, the plot shows:

* Higher success rates across all campaigns in May
* Lower failure rates in April
* Highest cancellation rates in July



1. What are some limitations of this dataset?

In the context of predicting successful trends, the dataset has the following limitations:

* Limited sample size
* Lack of information on individual pledge datetimes and amounts
* Lack of data past March 2017
* Lack of information on average number of visits to Kickstarter website during campaign periods

1. What are some other possible tables/graphs that we could create?

Many. We could create:

* A plot showing campaign success probability (successful ÷ (failures + cancellations)) for category, subcategory, and launch date
* A winner-take all series of visualizations that showed the best category, subcategory, and time of year to launch a campaign for the highest chance of success
* A series of small multiples for each of the included visualizations filtered by:
  + Year
  + Category
  + Subcategory
  + A series of graphs comparing the effect of spotlighting or no spotlighting
  + A graph showing success and failure rate by:
    - campaign length
    - Goal amount
* We could also alter the graphs to reflect the amount of goal funded, rather than just whether it was successful, failed, or cancelled. This would remove the distortion inherent in summarizing percent funded with the outcome.
* We could filter the above graphs by country to look at additional differences between countries