

Data Anaysis

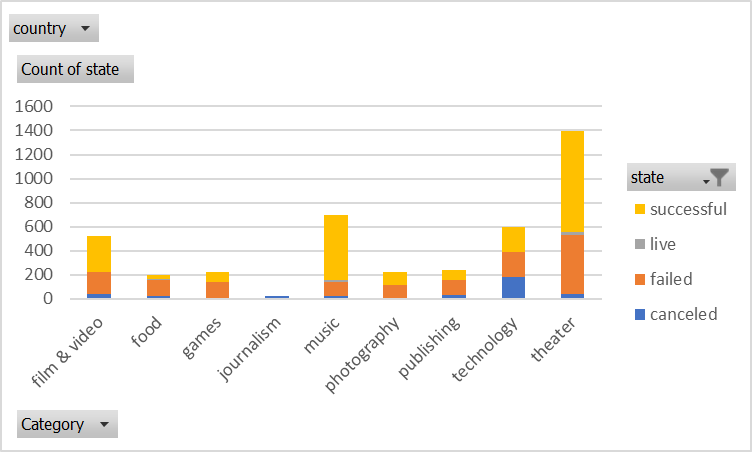
KickStarter Campaigns

Neethu John | 6/10/2020

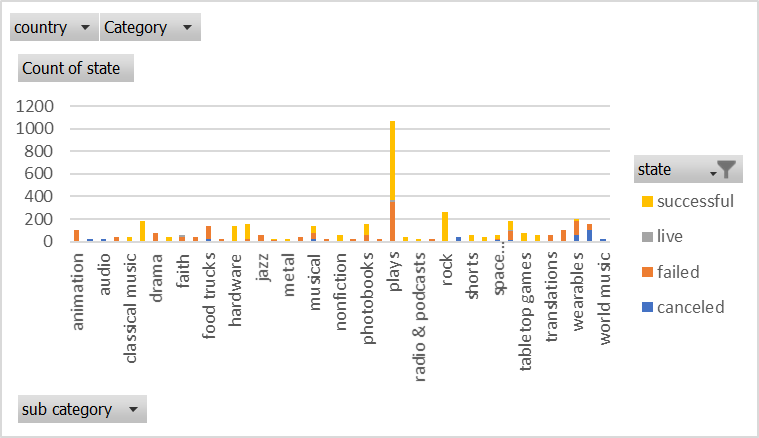
# Background

Over $2 billion has been raised using the massively successful crowdfunding service, Kickstarter, but not every project has found success. Of the more than 300,000 projects launched on Kickstarter, only a third have made it through the funding process with a positive outcome.

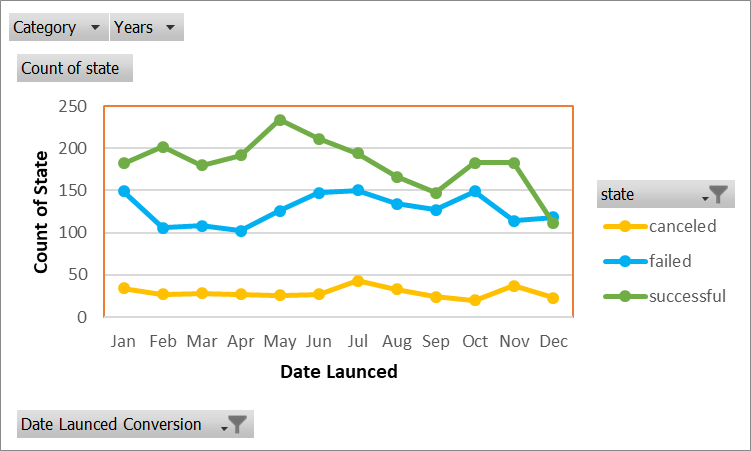
Getting funded on Kickstarter requires meeting or exceeding the project's initial goal, so many organizations spend months looking through past projects in an attempt to discover some trick for finding success. For this week's homework, you will organize and analyze a database of 4,000 past projects in order to uncover any hidden trends.



**Count of campaigns which were successful, failed, canceled, or are currently live per Sub- category.**

****

**Kickstart Campaigns success,failure and cancel rates**

****

# Conclusion

1. The Theater category had the highest number of kickstart campaigns and Music had the highest success ratio.
2. The Sub-Category, Plays, had substantially higher number of kickstart campaigns compared to the rest.
3. There was a higher number of Successful kickstarts throughout the year except in December where the number of successful campaigns were closer to the number of failed campaigns.

# Data Limitations

Analysis based on demographic characteristics like age, gender or income cannot be done.

# Other Tables/Graphs

Two graphs could be plotted, one showing the amount raised by country per Category and another one per Sub-Category, which will give an analysis of the contributions by country.

Also plotting the time period available between the launch and deadline dates vs the pledged amount could give an idea for getting better understanding of what would work for future projects to raise the required goal amount.

# Bonus Analysis: Successful and failed vs backers count.

In my opinion for the available data sets, median is more meaningful because it shows that more than half of the successful kickstarts were backed by less than 62 people each and were still successful. Also, the median of the failed kickstarts shows half of the those had less than 4 backers each which would have caused it to be failures.

The variance is a measure of how far the data are from each other and the mean value. In this case the variance of successful kickstarts are way higher than the failed ones, because the data points are too spread out with the minimum count at 1 and maximum count at 26457 for the successful ones compared to the minimum count at 0 and maximum count only at 1293 for failed ones .