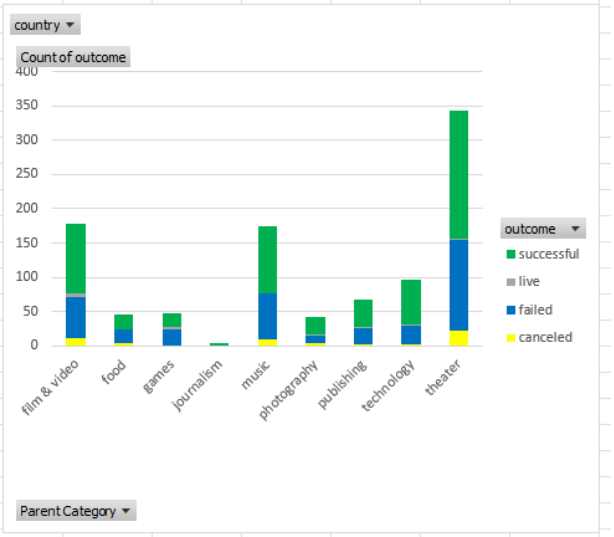
Report

1. Draw three conclusions from the data.

Answer:

Conclusion 1. The Kickstarter campaign has shown a higher frequency in the theater category compared to other groups. Not only theater but entertainment in general, is the best place to get more people's attention, in the data next to theater music, film & video is most prevalent followed by the music category. These may be due to theater is the best-suited place to get a lot of people at once and conduct a campaign.

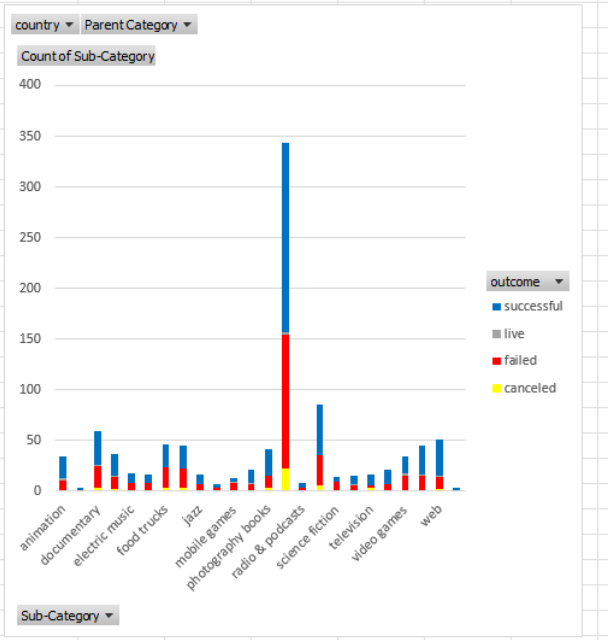
country	(All)				
Count of outcome	**Column Labels**				
Row Labels	**canceled**	**failed**	**live**	**successful**	**Grand Total**
film & video	11	60	5	102	178
food	4	20		22	46
games	1	23	3	21	48
journalism				4	4
music	10	66		99	175
photography	4	11	1	26	42
publishing	2	24	1	40	67
technology	2	28	2	64	96
theater	23	132	2	187	344
Grand Total	**57**	**364**	**14**	**565**	**1000**



Conclusion 2. In general, the Kickstarter campaign has greater success than failed, canceled, or live. However, the sample size is yet limited to generalize. When we come to successful campaign, "Plays" is the top. The percentage of successful play campaigns from the total is higher than the other groups. Rock music has next success rate.

country	(All)
Parent Category	(All)

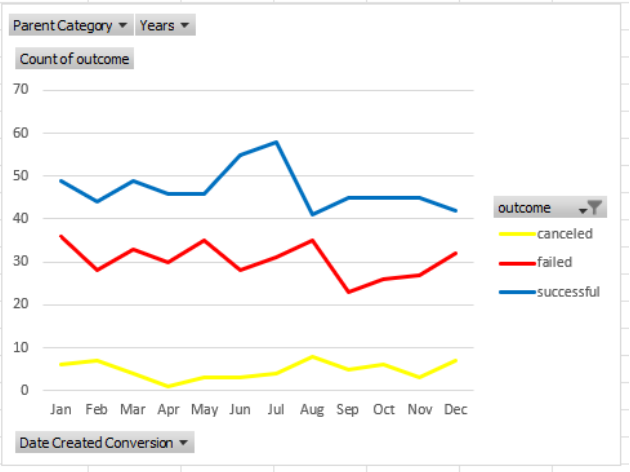
Count of Sub-Category	Column Labels				
Row Labels	canceled	failed	live	successful	Grand Total
animation	1	10	2	21	34
audio				4	4
documentary	4	21	1	34	60
drama	2	12	1	22	37
electric music		8		10	18
fiction	1	7		9	17
food trucks	4	20		22	46
indie rock	3	19		23	45
jazz	1	6		10	17
metal		3		4	7
mobile games		8	1	4	13
nonfiction	1	6	1	13	21
photography books	4	11	1	26	42
plays	23	132	2	187	344
radio & podcasts		4		4	8
rock	6	30		49	85
science fiction		9		5	14
shorts	1	5	1	9	16
television	3	3		11	17
translations		7		14	21
video games	1	15	2	17	35
wearables		16	1	28	45
web	2	12	1	36	51
world music				3	3
Grand Total	**57**	**364**	**14**	**565**	**1000**



Conclusion 3. The time serious analysis of Kickstarter campaign has shown successful result in July. This may be due to summer for people to go out for entertainment or theater. In the other category the beginning of the year is also the peak it may be due to a new year resolution started by people to be implemented.

Parent Category	(All)
Years	(All)

Count of outcome	Column Labels			
Row Labels	canceled	failed	successful	Grand Total
Jan	6	36	49	91
Feb	7	28	44	79
Mar	4	33	49	86
Apr	1	30	46	77
May	3	35	46	84
Jun	3	28	55	86
Jul	4	31	58	93
Aug	8	35	41	84
Sep	5	23	45	73
Oct	6	26	45	77
Nov	3	27	45	75
Dec	7	32	42	81
Grand Total	**57**	**364**	**565**	**986**



Conclusion 4. The mean and the outlier result on the successful vs failed backers count showed that there is more variability in a successful group than the unsuccessful group. However, for both groups the data is not normally distributed, it is positively skewed (left-skewed). The probability distribution graph also looks identical to the data graph.

The median, the quartile, and outlier tests also showed an upper bound outlier for both successful and failed groups. There is higher variability in a successful group than the failed group. Based on the given data the median summarizes the data more meaningfully.

2. State limitations of the dataset.

Answer:

- This dataset only includes a limited set of the population from many projects. To have a better insight, we need more samples.
- The data is not normally distributed we need more samples to generalize about the data set.

3. Suggestions for additional tables and/or graphs that we could create?

Answer:

- Make outlier testing
- Creating a graph that shows the campaign result quarterly, and yearly not only by month.
- Conducting some statistical analysis, creating a table and graph that shows the difference across outcomes, categories and time
- Creating a graph and table by making a comparison across different countries on the campaign being successful or failed.
- Creating a comparison table or graph which shows average donation across campaign outcome.