Yazan Khayata

Kickstarting With Excel

1: Overview of Project:

The main purpose of this analysis was to detect how different Kickstarter campaigns fared in relation to both their launch dates and their funding goals. In order for Louise to meet her fundraising goal, she must look at these different relationships and determine if there are any outstanding correlations that are note-worthy.

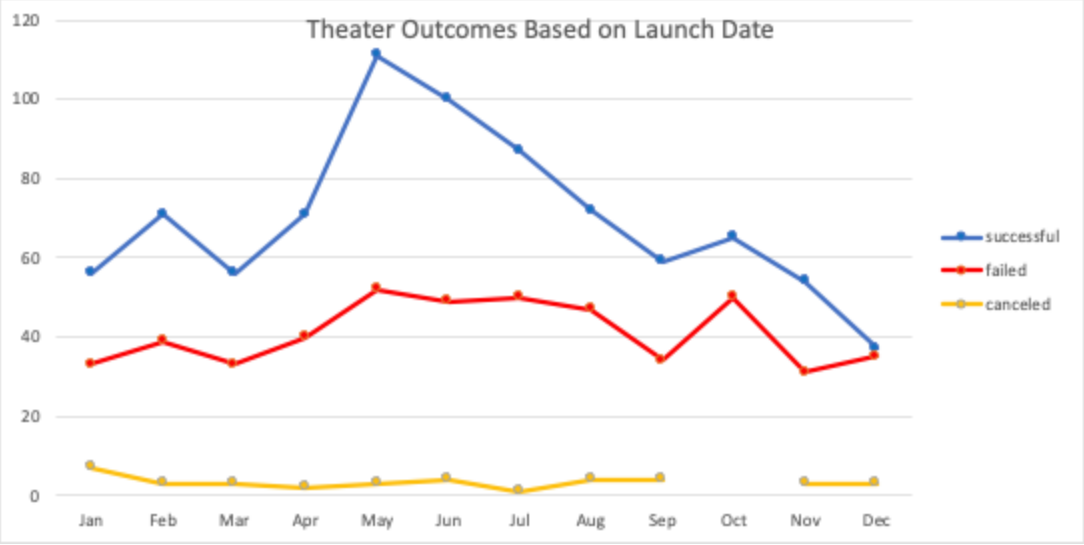
2: Analysis and Challenges:

A) Outcomes by Launch Date:

The first thing we wanted to do was visualize the campaign outcomes based on launch date. In order to perform this task, we needed to create a pivot table and filter the Parent Category field for “theater”. Next, we made sure the rows demonstrated the months of the year, while the columns/values showed the count of outcomes (successful, failed, or canceled). Please see screenshot of Pivot Table below:

Parent Category	theater				
Year	(All)				
Count of outcomes	Column Labels				
Row Labels	successful		failed	canceled	Grand Total
Jan		56	33	7	96
Feb		71	39	3	113
Mar		56	33	3	92
Apr		71	40	2	113
May		111	52	3	166
Jun		100	49	4	153
Jul		87	50	1	138
Aug		72	47	4	123
Sep		59	34	4	97
Oct		65	50		115
Nov		54	31	3	88
Dec		37	35	3	75
Grand Total		**839**	**493**	**37**	**1369**

Next, to better visualize the data we're dealing with, we went ahead and created a line chart to summarize the table in a rather simple visual. (See below)



B) Outcome Based on Goals:

Next, we wanted to visualize the relationship between outcomes and goal amount ($) to see if there are any noteworthy correlations that Louise may use in her future campaigning work. Our ultimate goal was to visualize the percentage of successful, failed, and canceled plays based on

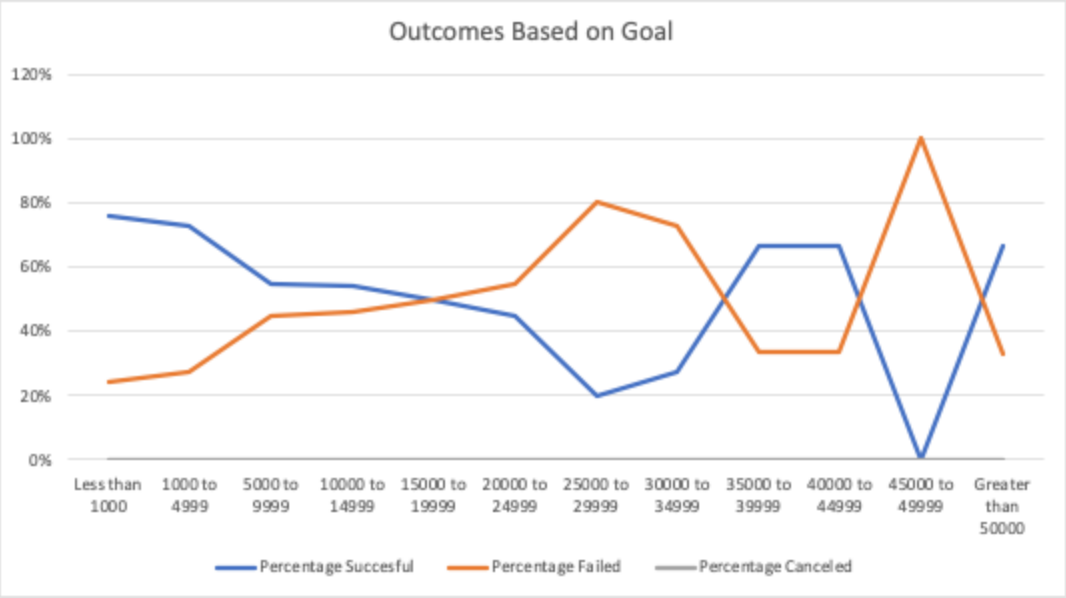
the goal amounts, and luckily we had the data to go ahead and accomplish that. We created our own table on a new sheet, that consisted of different goal ranges in dollar amounts to determine whether a correlation is valid. We used the “COUNTIFS()” function to populate the "Number Successful," "Number Failed," and "Number Canceled" columns by filtering on the Kickstarter "outcome" column, and filtered the “Subcategory” column, using “plays” as the criteria. Please see the table below:

Goal	Number Successful	Number Failed	Number Canceled	Total Projects	Percentage Succesful	Percentage Failed	Percentage Canceled
Less than 1000	141	45	0	186	76%	24%	0%
1000 to 4999	388	146	0	534	73%	27%	0%
5000 to 9999	93	76	0	169	55%	45%	0%
10000 to 14999	39	33	0	72	54%	46%	0%
15000 to 19999	12	12	0	24	50%	50%	0%
20000 to 24999	9	11	0	20	45%	55%	0%
25000 to 29999	1	4	0	5	20%	80%	0%
30000 to 34999	3	8	0	11	27%	73%	0%
35000 to 39999	4	2	0	6	67%	33%	0%
40000 to 44999	2	1	0	3	67%	33%	0%
45000 to 49999	0	1	0	1	0%	100%	0%
Greater than 50000	692	343	0	1035	67%	33%	0%

Example of the “COUNTIFS()” function pertaining to this specific table:

=COUNTIFS('Kickstarter (Filtered to Plays)'!$F:$F, "=successful", 'Kickstarter (Filtered to Plays)'!$D:$D, "<1000")

Finally, we inserted another line graph based on the data in the table above, to visualize the relationship between "Percentage Successful," "Percentage Failed," and "Percentage Canceled" relative to the goal dollar amount ranges. See below:



C) Challenges:

The main challenge that I came across putting this all together was integrating the new formulas that were introduced in the module. I have very little excel experience so I had to go over the module a couple of times before I fully understood everything that was being explained. Thankfully, I was able to integrate all these new formulas in the challenge.

3: Results:

A) What are two conclusions you can draw about the Outcomes based on Launch Date?

i: The Summer months (May, June, July) tend to have a higher success rate, while the success rate towards the end of the year tends to dramatically decrease and hits its lowest point in December. What does this tell us? Well, it shows the summer time may be a peak time for theater campaigns, and for the future, one who is willing to start a campaign should focus on these summer months.

ii: Another conclusion we can draw is that the cancellation rate for the theaters subcategory is very low, and this remains consistent with all months. This shows that the

months of the year does not really affect the cancellations, and in fact, proves most entrepreneurs go through with their campaigns until all else fails.

B) What can you conclude about the Outcomes based on Goals?

i: Looking at our "Outcomes Based on Goals" line chart, we can conclude that as the goal dollar amount increases for the "plays" subcategory, the successful percentage decreases. In other words, as the goal dollar amount increases, the failure percentage increases. The goal range from $30,000 to $45,000 shows the success rate going up, but this is simply because there were much fewer projects in these ranges, compared to the rest which can suggest this is possibly an outlier. But, when looking at goal amounts greater than $50,000, we see that the successful rate skyrockets back up again, while the failure rate plummets. This suggests that campaigns with very high goal amounts tend to succeed heavily and my assumption would be because these campaigns are well-run and well-managed.

C) What are some limitations of this dataset?

Some limitations may include the outliers (specifically in the 'plays' data) and how it played a significant role in the line graph. Another limitation is the data could be biased...maybe successful campaigns are more likely to be reported than failed campaigns?

D) What are some other possible tables/graphs we can create?

We can create a table/graph that visualizes the relationship between plays and successful outcomes, and determine how much money was actually pledged vs what the goal was.

Once we see the variance, we can determine if there's a correlation between high variance of goal vs pledged and successful outcomes.