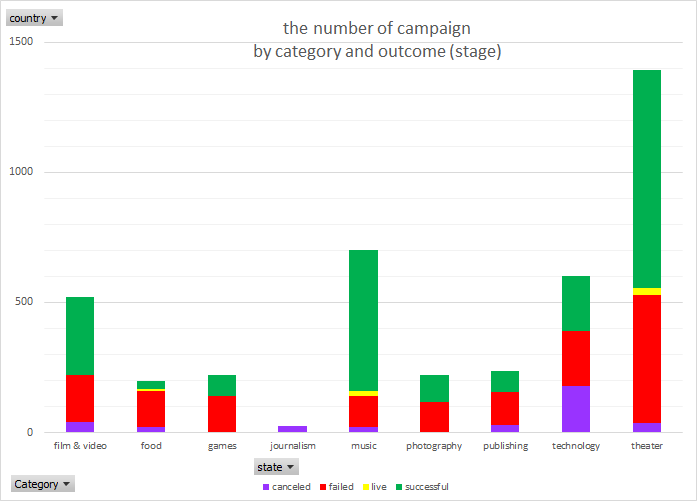
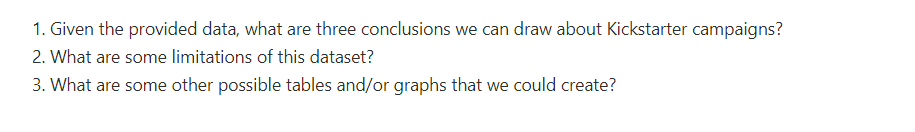


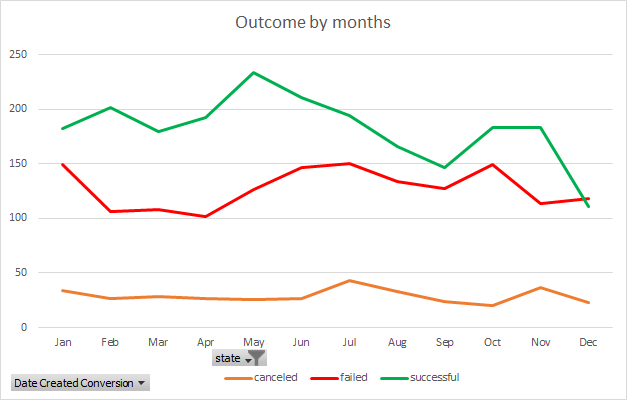
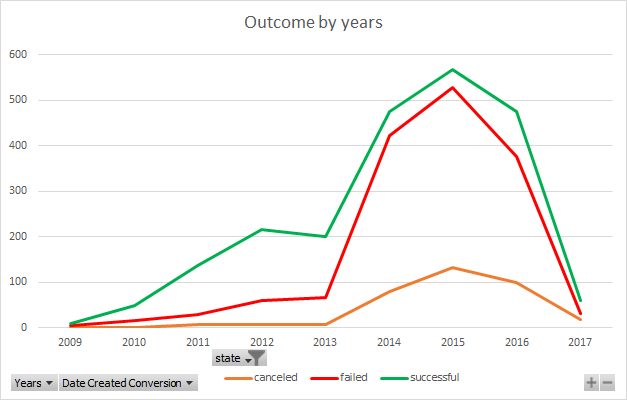
For this homework we are asked to modify tables, reorganize data to make them more readable with pivot tables and charts, and answer some questions.   
  
I will start with answering the questions, posting tables and charts as needed to answer questions. On the second session I will explain the various ways I get into the required tables and charts .

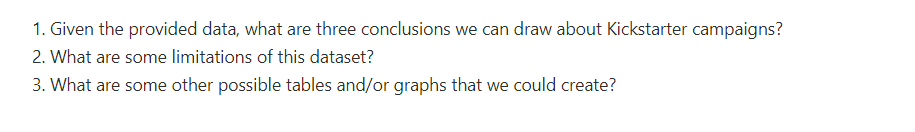
QUESTIONS :



1. The three conclusion that I get from this data are

* The type of projects that is being back-up in Kickstarter campaigns  
  Most of them are in the arts fields, with the non art fields; journalism, games, technology and food industry represent only a quarter of the project. Because of that, more projects in these fields failed to be funded or canceled completely.  
  Theater, music, and film&video seem to be the most popular projects and the most successful in getting funded as well. With their success level above 50%
* The number of backers and the size of their pledge matters in the success and failure of a campaign. Successful campaigns have a medium of 62 backers, while the failed one only generates a medium of 4 backers.
* The kickstarter programs became very popular in 2014, reaching top in 2015. The 2009 and 2017 data is distorted because it only covers three months of data.   
  The outcome of successful campaigns generally is greater in the beginning of the year.





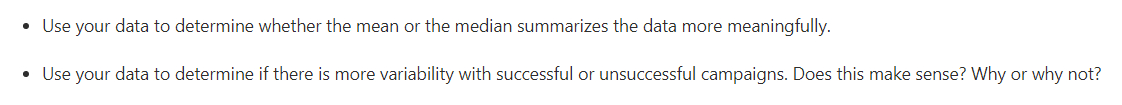
1. The data set does not show the quality of the product, experience, and promotion quality of these campaigns to evaluate why these campaigns fail.   
   It doesn’t also show the variety size of the contribution of the backers.
2. There are a lot of possible tables/graphs that can be created. For example :

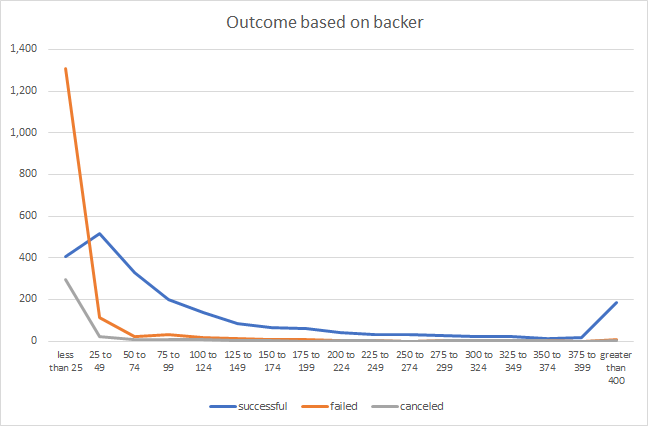
* The relationship between targeted campaign goal and the amount of the pledge.
* The size of targeted campaign goal and the amount of pledge based on field.
* The time that is needed for a campaign to fulfilled based on the outcome
* Weather staff pick resulted in a better outcome

**BONUS QUESTIONS**

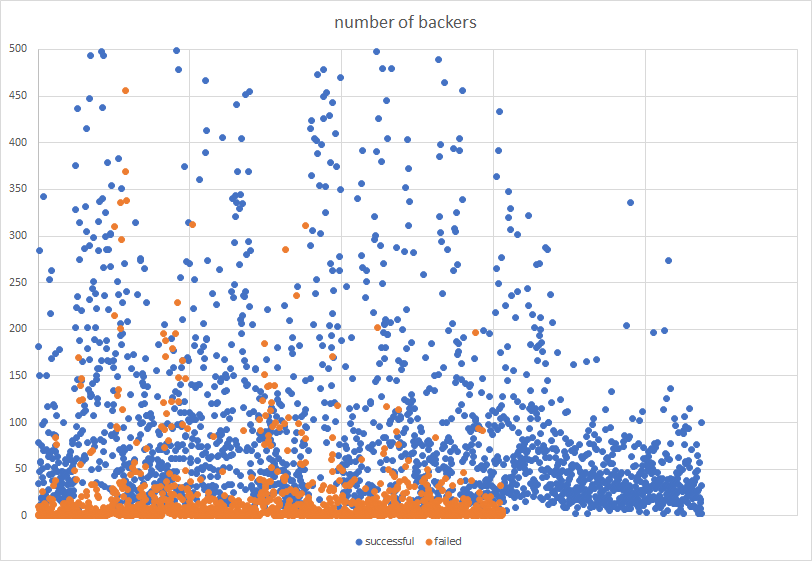


|  |  |  |
| --- | --- | --- |
|  | successful | failed |
| Mean (average) | 194.43 | 17.71 |
| Median | 62 | 4 |
| Min value | 1 | 0 |
| Max Value | 26,457 | 1,293 |
| Variance | 712,841 | 3,773 |
| Standard Dev | 844.30 | 61.43 |



From the data that I analyze, the median give much better representation of these data because 

* The max value are much higher than median
* There are high number of outliers
* The outliers are extreme and thus skew the result of mean calculation
* This data is not evenly distributed and skewed to the lower quarter.
* The variance, which has an extremely high number shows the high non-uniformity of these data. A high variance shows that many of the data differ from the trendline.



The data can be used to determine the success and failure of a campaign, especially after these outliers are put aside. As can be seen in this chart, successful campaigns have so much more backers than the failure ones.

EXPLANATIONS

1. As being requested, I created column “Percent Funded” and “Average Donation” and put conditional formatting on “state” and “percent funded” column added column and
2. I use the text function to create column split for “Category” and “Sub-Category”
3. Created column “Date Ended Conversion” and “Date Created Conversion” and converted the data with =(I2/86400)+DATE(1970,1,1) formula.
4. Created pivot table to determine the number of campaigns by category and outcome (stage) and its related graph. (worksheet “Pivot table 1”)
5. Copy worksheet “Pivot table 1” and rename it “Pivot table 2”, anjust the entry to accommodate sub categories
6. Created pivot table to determine outcome (stage) by month and created the related graph. (worksheet “Pivot table 3”)
7. Created table to calculate outcome based on funding goal (on “Table4 worksheet”). And populate it with the formula below

=COUNTIFS(data!$D:$D,"<"&$A4,data!$F:$F,"="&C$3)

1. I filtered successful and failed campaigns and copy special to a new worksheet (Worksheet Statistical Analysis). Then created a table of various statistical analysis, and put formulas to calculate their values.  
   Beside using pivot table - I used various ways to extract to the data, including parching and sorting it manually (Sheet 2), using index function\* (sheet 1), VBA (Table5)

\* The index formula that I use.   
 =IFERROR(INDEX($B$2:$B$4201, MATCH(0, COUNTIF(D$1:D1,IF($A$2:$A$4201=D$1,$B$2:$B$4201,D$1)), 0)),"")