**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.

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   * *There has been more funding in campaigns related to the entertainment (film & video, games, music, theater) and technology industry, regardless of outcomes.*

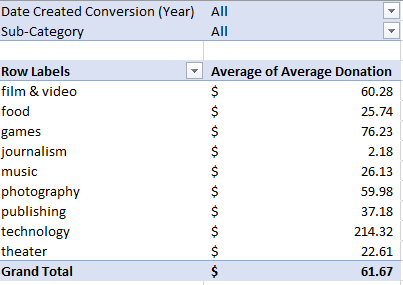
* + *Some categories such as food, games, journalism (100% failed), photograph, and publishing were not successful in most of the campaigns.*
  + *Campaigns launched between April and May had better outcome, for all years analysed.*
  + *Most campaigns had their outcome during the first 50 days. However, failed campaigns last longer than the successful ones, which had positive outcome in the first 40 days.*
  + *By confronting “subcategory” and “status” it is possible to identify which subcategories were successful or not, for example, “rock” and “indie rock” have stood out from other campaigns. While "jazz", "world music" and "faith" would be a great risk of failure. Therefore, the data provides enough information to decide which category to fund / launch a campaign.*

1. What are some limitations of this dataset?

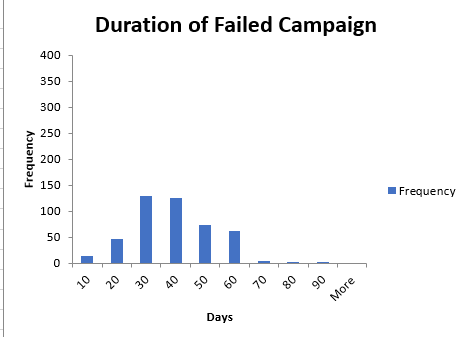
*There are limitations:*

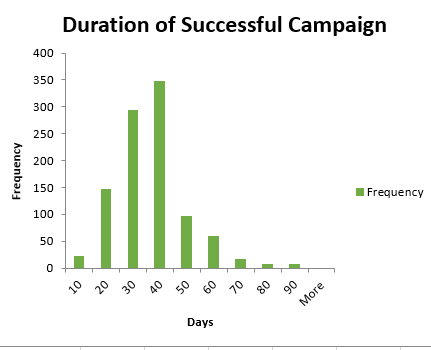
* + *The dataset goes up to March 2017, which means it is not fresh information.*
  + *There is no information why campaigns were canceled.*
  + *There is no information why some campaigns asked for only 1 dollar.*

1. What are some other possible tables and/or graphs that we could create?
   * *Average donation by Category, filtering by Year and Subcategory: leading to understand how much people can invest.*



* + *Duration of Successful and Failed Campaigns by Day*



**

Statistical Bonus Questions:

1. **Use your data to determine whether the mean or the median summarizes the data more meaningfully**.  
     
   *In this case, the median gives a more realistic picture of the data because by checking other metrics - quartiles, outliers and boundaries - it is clear that the mean has been affected by outliers.*
2. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?***The "successful campaigns" shows more variability and it makes sense because its dataset has more spread out values than "failed campaigns" dataset.*