**Analysis Starter Book**

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 to uncover any hidden trends.

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?

Three conclusions based on data sheets and graphs:

* 1. Theatre (839), Music (540), Film & Video (300 ) are the top three most successful, but of these top three, theatre had half failed (493), compared to music (120) and film & video (180).
  2. In the Category of theatre, the sub-category of plays (694) was most successful, followed by music’s sub-category of rock (260 ), and film & video’s sub-category of documentary (180). But of the three sub-categories, only plays had failures of 353 and none in rock nor documentary.
  3. Campaigns peak in the months of April through May, where cancellations are minimal (at a plateau), but then failures of campaigns begin to increase from April through June.

1. What are some limitations of this dataset? The dataset for the Kickstarter service does not have information to answer some of the basic questions of the Kickstarter business which are,

* Who are the stakeholders involved in the project?
* Where does the problem occur?
* Once the where of the problem is identified, what needs to be solved?
* When does the problem occur? When does the project need to be completed?
* Why should this problem be solved? Why does the problem occur?
* Locations of the fundraisers under Kickstarter.

1. What are some other possible tables and/or graphs that we could create? Other possible tables and/or graphs that can be created are
   1. Identifying successful, failed, and canceled campaigns at peak periods /month.
   2. Identifying successful campaigns that remained consistently successful throughout the years, those that begin and/or continue to fail and/or were canceled.
   3. Identifying successful campaigns with the highest target goals which fell into less than 1000 and 1000 to 4999.

Statistical Analysis

1. Use your data to determine whether the mean or the median summarizes the data more meaningfully.

The mean summarizes the overall picture of the Kickstart data set to identify the average successful and failed campaigns. The median cannot be used to determine an exact midpoint since there are several missing information of data under sub-categories and outliers.

1. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

Variability measures the spread from the mean. There is more variability with the successful campaigns versus the unsuccessful campaigns. This means that the measure of spread is greater in the successful campaigns, and the measure of spread is less in the failed/unsuccessful campaigns. Hence, it will be easier to predict which campaigns will fail versus which will be successful. This may be the direction to consider when looking into which campaigns to cancel or dissolve.