1. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**
   1. An obvious conclusion is that the number of backers has an influence on success rates of a project. Generally, the more backers a project has the better chance of reaching the goal of money sought.
   2. The type of project makes a difference. Certain categories, especially those related to the arts, like theatre plays, music, and film seem to have a better rate of successful projects than others.
   3. Collecting the goal amount is the determining factor on whether the project is successful or not. If this amount is not realistic in relation to pledge efforts or other factors, then it can directly affect the number of successful and failed projects. Keeping that in mind, when goal amounts hit $20K, the percentage of successful projects drops below the percentage of failed projects and continues this trend as goal amounts increase.
2. **What are some limitations of this dataset?**

One of the limitations to be considered would be that of the sample size of the dataset. Are the 4,000 projects used in the dataset a good representation of the more than 300,000 projects launched on Kickstarter?

Another limitation of this dataset is that it does not include data on pledge budgets or demographic and financial data of individuals pledging and donating. As mentioned above, the goal amount is instrumental in determining the project’s success or failure. Just having the goal amount does not give us an idea if the goal was realistic. What kind of methodology was used to come up with that amount? Inconsistencies in stating these goal amounts could affect the successful and failure rates of this dataset.

Also, the fact that these projects are conducted all across the world adds a dynamic that the dataset cannot fully address. Economic conditions, cultural differences of the project’s country of origin, as well as people’s attitudes toward services like Kickstarter can influence several factors that can determine a project’s success.

**What are some other possible tables and/or graphs that we could create?**

One factor that seems to have an influence on the success of a campaign is obviously the number of backers of each project. I think a table showing details on the advertising/marketing budget of a project could show if there was a correlation between

(cont.)

budget expenditures and success rates of a project. And if possible, break down specifics of the budget (online, tv/radio, cold calling, etc…)

Also break down the pledge information. Obtain info on how the pledge was made and how much was pledged to see if that can help predict if a pledge will be fulfilled.

Demographic/financial data on the backers could possibly predict whether a pledge is likely to become an actual donation.

**Bonus Statistical Analysis**

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

In the case of the successful campaigns, there are significant outliers in the dataset that affect the mean and therefore does not make it an accurate depiction of the data. The median will represent the dataset better. The failed projects data does have outliers, but not to the extent of the successful projects. In this case, using the mean and median together can give a better summarization of the dataset.

**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 campaign has more variability due to some extreme outliers within the sampling. I think it does make sense because the successful campaigns generally have more backers and with a higher number of backers than the unsuccessful campaigns it is likely to have a higher occurrence of outliers which will cause the variability.