Homework 1 – Answers

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

(i) Using the first bar-chart. It is clear that Kickstarter campaigns in “film & video”, “music” and “theatre” categories were successful. Music was the most successful category with 77.1% success rate. All campaigns in “journalism” had to cancel and there was very little success in the food industry.

(ii) After analyzing the second chart it is clear even in the “music” category there were failures. All campaigns in “jazz” and “faith” sub-categories have failed. There is a high chance that the remaining campaigns in the “faith” category too will end up with failure. The most notable category is the “rock” sub-category. All 260 campaigns in this category were successful. Also, all campaigns in “classical music”, “electronic music”, “metal” and “pop” were successful. In the theater category “plays” sub-category was more successful compared to other two categories in raising funds. “Film & video” category, “documentary”, ”shorts”, and “television” sub-categories show 100% success rate. However, number of campaigns in these categories is much lower compared to successful campaigns under the “music” category. All campaigns in “animation” and “drama” sub-categories were failed. This trend is also the same in USA.

(iii) The 3rd chart shows the success-failure dependence with the month of the year. The is a high chance to be successful in the months of May, June and October while there is a high chance to fail in the months of January, September and December.,

2. What are some limitations of this dataset?

There are not sufficient data from the other countries. Most campaigns are from USA.

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

We could create a graph with percent-funded with each category and sub-category. Also, it will be useful to graph total pledged with how long campaigns lasted.

Part 2- Bonus Statistical Analysis

In. both cases median explains the data more accurately. In the successful campaigns, the max. no. of backers is 26,457 while the min. no. of backers is 1. This makes the mean very high. Also, in the second case median no. backers is 4 and it explains why these campaigns were failed. There is more variability in successful campaigns. This makes sense because all unsuccessful campaigns have very few backers, so the variability is lower. Successful campaigns have large number of backers in some cases increasing the variability.