Emma Braden

Excel HW 6/13/20

1. Given the data in the Kickstarter dataset, we can see that the most common category of campaign to launch is theater. Plays were most especially popular of those with 1066 different campaigns in the time frame of the dataset. The least common campaign category is journalism, making up only 24 of the campaigns in that time frame, although there is likely something about Kickstarter’s platform that is not suitable for journalism since all 24 were canceled. The most successful category (meaning with the highest percentage of fully funded campaigns of the total) was music, with over three-quarters of campaigns achieving their goal.
2. One limitation of the dataset is that the goals are in different currencies. It makes direct comparisons of projects cost and success challenging, although the majority of analysis can be done using the percentage of the goal and comparing rates is appropriate. Another limitation is the existence of outliers, particularly if one wanted to test for correlation and causation. There are many Kickstarter campaigns that floundered, reaching deadlines without raising any money, but their significant number may be obfuscated by a select few wildly successful examples—some of which earned more than ten times their goal. It is also limiting that there are several entries that were live at the time the data was collected, meaning their outcomes may not be accurately represented. It may be advisable in for some analysis to drop those data points.
3. An interesting graph to create from the Kickstarter data set would be to track the number of campaigns across time using a line graph, filtered by category. Seeing the trends of different projects at different points in time could reveal some about Kickstarter’s role from the point of view of contributors. Filtering the same graph by outcome would also reveal which categories were most appealing at which points in time. Filtered by year and subcategory, a scatterplot of each campaigns’ percent funded against the month would create up to 12 lines of points and could reveal a correlation between competition and the overall success of a project. A time with many of the same subcategory of campaigns may deplete the funding of all the projects or draw attention to a type of product or project.