(part 1) What are three conclusions we can make about Kickstarter campaigns given the provided data?

Note: I created more graphs then requested in the homework problem set to answer part 1.

1)

Parent category theatre, followed by music and technology had the most campaigns. Among the sub-category, plays, wearables, and documentary were the top 3 popular campaigns.

2)

However, when looking at successful % by sub-category, wearables which was in the top 3 campaigns only had 10% success rate. Documentary had 100% success rate as well as 11 other campaigns. Plays which had the highest number of campaign had 65% success rate. All other campaigns were not at all successful. Parent category also showed high variability with music, theatre, and film & video more than 50% success rate and rest under 50% success rate. Surprisingly, journalism with 24 total campaigns did not have any successful campaigns.

3)

Most campaigns were created in the years 2014-2016. In the earlier years of kickstarter, success rate was significantly higher than the later years of kickstarter.

(part 2) What are some of the limitations of this dataset?

The outcomes by goal is not a standardized currency so we cannot use this graph to gain insight on outcomes by goal. The data should be furthered cleaned to convert all currency to one type and analyze success vs failure rate by set goal.

(part 3) What are some other possible tables/graphs that we could create?

Please see answers to question part 1. We can graph success rate % by parent or sub-category, and breakout state by month and year. More analysis should be conducted to see if % of success is due to the “novelty” of kickstarter campaigns or if type of campaign and its success rate and if it also deviates by time.