Thomas Stephens

Data Science Excel Homework Feb. 13, 2021

1. Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?
   1. The success line is highest… so more Kickstarter campaigns succeed than fail or cancel (though combined fail and cancel are larger than Success).
   2. The best time to start a Kickstarter is May or June. September is bad and December is the worst.
   3. Theater, and particularly Plays are the most popular Kickstarter campaigns
2. What are some of the limitations of this dataset?
   1. Overall a very nice dataset. It’s clearly labelled and was able to do all calculations easily. Could be larger. I thought it was strange that many sub-groups are 100% failed, cancelled, etc. More data would give a clearer picture of the distribution of subgroup state.
3. What are some other possible tables and/or graphs that we could create?
   1. We could go as far as graphing every event on a big timeline using colors for state and width for funding amount. We could create quarterly breakdowns or show how backer donation amount effects success.

Bonus:

I completed all bonus activities. Both the mean and median showed that successful campaigns did indeed have more backers. Median is the more useful metric in this situation due to the high standard of deviation. That being said, the standard of deviation is much lower for Failed campaigns than Successful. This is due to the consistently low average of failed backers compared with the viral success of some successful campaigns ranging from 0 - ~26000 rather than 0-1293.