Graham Penrose

UCD Data Analytics Boot Camp

Week 1 Excel Homework

Due: 9/28/19

1. Three conclusions I can draw about Kickstarter Campaigns after creating the pivot charts and tables to analyze the relationship between the success of a campaign and category, subcategory, starting date, or funding goal are as follows:
   1. Primarily, the highest probability of successfully reaching a funding goal (achieving success as campaign state) is in the parent category “Music”, followed by “Theater” then “Film and Video”. The percent success (count of successful projects/ count of failed projects) for those categories are 450%, 170%, and 167% percent respectively.
   2. With regards to subcategories, within the above three most successful parent categories, the most successful subcategories were “Rock” for music, “Plays” for theater, and “Documentaries” for film and video. An interesting note about subcategories is that within the Film and Video parent category, all documentaries, shorts, and television campaigns were successful, but there wasn’t a single successful animation, science fiction, or drama campaign.
   3. When looking at successful campaigns over time, based on the month the campaign was started, it seems like the most successful month to start a campaign in is May. This is across all the years of the dataset, so a more general trend. The mean of the number of projects started per month is 343 with a standard deviation of 38 projects and May’s grand total for projects started is 386 (just barely more than one standard deviation from the mean). This simply means that the higher count of successful projects that start in May isn’t due to may having an usually large amount of projects starting in May regardless of their final state of success, failure, or cancelation.
   4. Finally, when examining campaign state across funding goals, there is a general trend of decreasing number of successful campaigns as the funding goal increases. Also, inversely there are generally increasing numbers of canceled and failed campaigns as the funding goal increases. This trend makes logical sense, but the numbers from the dataset confirm the hunch.
2. While this Kickstarter Campaign dataset has quite a few dimensions provided for analysis, there were a few limitations in it that reduce the ability to analyze underlying factors that may impact if a campaign is successful or not. One bit of data that would be nice to have in order to analyze successful campaigns would be more data on the backers for each campaign. If there was more data about average age, primary sex, backer location, etc., it may be useful information in understanding what type of person is more likely to support what type of campaign. Also, if there is a data point collected by Kickstarter that captures how many unique people look at a campaigns description vs how many backers they actually capture, it may be useful in assessing how well the blurb is written to describe the campaign and capture backers interest.
3. The graphs and tables we made do a pretty thorough job at covering the valuable insights that can be uncovered from analyzing the given data. A few more tables/ graphs that could have been created include:
   1. Campaign duration (deadline date – launched date) by state of success/failed/canceled
      1. To understand impact of campaign timeline on success
   2. Percent Funded by parent and subcategory
      1. To see if a certain category has a tendency to get funded above their goal
   3. Average Donation vs. State
      1. To see if successful campaigns are usually created through a lot of small donations or fewer large donations
   4. Finally, could look at Spotlight and Staff Pick vs State
      1. To examine if those two ways of emphasizing a campaign impact the outcome of the campaign