Name: Mahyar Mohammadyani (Mika M-Yani)

Course: Bootcamp UTA-VIRT-DATA-PT-12-2022-U-LOLC-MTTH

Topic: Module 1 Challenge

* Given the provided data, what conclusions can we draw about crowdfunding campaigns?
  1. By analyzing the data from the Parent Category, Sub-Category & Date tabs, projects which were successful and unsuccessful shows that subcategories such as music, documentary, electronic music, hardware, metal, non-fiction, pop, radio and podcasts, rock, shorts, small batch, tabletop games, and television have a 100% success ratio. However, some projects have had no success at all or a 100% failure ratio; these include animation, children’s books, drama, faith, fiction, food trucks, gadgets, jazz, mobile games, nature, people, places, restaurants, translations, videogames, and web. Considering the relatively small sample size, we should note that these findings may not be statistically significant and require deeper statistical analysis.
  2. July is the peak month for the most successful campaigns.
  3. The theater parent category has the most pledges and has excellent potential for future campaigns.
* What are some limitations of this dataset?
  + The dates that each campaign was held were sometimes years different than other parameters that were being compared, which in the digital world can have a huge impact when looking at the growth happening in this industry over time.
  + Crowdfunding is becoming more and more prevalent, which can make it hard to estimate the future prospects for the same type of campaign.
  + The currency difference makes the data entirely irrelevant for comparison unless the exchange rates are being applied for each filter.
  + The source of the funding for each campaign is not clear.
* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
  + Graphs for year over year category
  + The success/failed ratio for the parent categories over the time
  + Excluding the subcategories that are not as popular as in the past.
  + Using Z-score to interpret how far is the data point from the mean for future reference.
  + Excluding the failed campaigns that have not had enough time to evolve, Some got failed in only two days, which makes our data set not being realistic.