Unit 1 | Assignment - KickStart My Chart

1. What are three conclusions we can make about Kickstarter campaigns given the provided data?
   1. Going by category, considering all countries in the dataset, “music”, “theater”, “film & video” possess the highest success rates of 77%, 60% and 58% respectively. On the other hand, “food”, “games” and “publishing” have the highest failure rates of 70%, 64% and 54% respectively.
   2. Based on the sample size, the highest number of projects pitched were for “plays” under the “theater” category which had a success rate of 65% in receiving funding.
   3. Projects pitched in May have the highest success rate (61%), while projects pitched in December have the highest failure rates (47%). Further, projects pitched in July and November have the highest cancelation rate (11%).
2. What are some of the limitations of this dataset?
   1. Sample size skewing the outcome – Out of all categories, “theater” alone constitutes about 34% of the sample size, thus having a larger spread across the state of the project. In contrast, certain projects have been pitched only once and have been successful, resulting in a 100 % success rate. Thus, finalizing to pitch on a category just based on this data may not guarantee funding.
   2. Critic/ viewer ratings – While this dataset delineates the success or failure of the project in securing funding, there is no account of critic or customer feedback because when a successful project is accompanied by positive critic/ viewer feedback, it is then an indication of prolonged viewership.
   3. Number of viewers in each country – If provided with the number of viewers for each project by country, it will help determine the popularity of each category in the country.
3. What are some other possible tables/graphs that we could create?
   1. Success rate based on year i.e. what was the most successful category for each year
   2. Success rate based on country i.e. what is the most successful category in each country
   3. Relation between staff picks and success rate? Do staff picks always result in success?
   4. A sentiment analysis using text mining to determine the theme outlined in the “blurb” column can be undertaken to correlate success/ failure rates with theme.