Crowdfunding Written Report

In this report I am going to the addressing the following questions:

* + Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
  + What are some limitations of this dataset?
  + What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

The most common crowdfunding projects are from the theatre/play subcategory with 344 campaigns out of 1000 from the crowdfunding dataset which is approximately 34% of the data. The least common category is the music/world music which has 3 out of the 1000 dataset making it 0.3% of the data. (Source: Pivot Table1)

The most common parent category crowdfunding categories are plays with 344 projects and least common is world music with 3 projects. (Source: Pivot Table2)

There are 7 counties represented in the dataset which are, Australia (AU), Canada (CA), Switzerland (CH), Denmark (DK), Great Britain (GB), Itay (IT) and United States of America (US). The county with the most crowdfunding campaigns is US which has a total of 763 out of 1000 which is approximately 76% of the dataset. There could be several reasons why crowdfunding is popular in USA one of them being that the US is one of the leading countries in the entertainment industries. The USA has a strong entrepreneurial culture, and crowdfunding provides a platform for startups and small businesses raise funds directly from the public, bypassing traditional financial institutions. The dataset shows that Switzerland has the least number of campaigns of 23 which is approximately 2%. The rest of the countries have less than 50 campaigns which makes USA an outlier in this dataset. Another conclusion observed from the dataset is that theatre/play subcategory is top category for all included countries. (Source: Pivot Table1)

The dataset shows some impact caused by the period on the campaign. The most successful period for campaigns is the 2nd quarter of the year with the highest number of 147 campaigns. The least period of successful campaigns is the 3rd quarter. Even though the highest pick of successful campaigns is in July (3rd quarter), August has the most cancellations. The data set shows the January has the highest number of failed campaigns. These timelines suggest that the best period to launch a crowdfunding campaign in 2nd quarter of the year, but other factors should be considered, like country and subcategory of campaign which influence the outcome. (Source: Pivot Table3)

Some of the limiting factors of this dataset are that most of the projects are in the entertainment industry which is influenced and promoted by social lights which might give a biased successful rate. Hence the dataset is not representative of the entire population or target group, the conclusions drawn may not be applicable to the broader context. The data set time stamp is from the year 2010 to 2020 which means that changes in technology should be considered. Technology has advanced and more people now have access to the internet and can easily access online banking. Deciding on crowdfunding in 2023 might require a most recent dataset. Environment and economic shift since the COVID19 pandemic my have positively or negatively impacted the crowdfunding projects.

Other graphs to consider would be a pie charts as it an excellent way to illustrate how different parts like country, timeline and subcategory contribute to the whole. They make it easy to understand the distribution of a single category. Pie charts allow visual comparison immediately without overwhelming viewers with complex information. Pie charts also have their limits as in the datasets as we are comparing a big set of data.

Box plots are good at providing concise summary of the distribution dataset, including the median, quartiles and the presence of outliers. This allows a quick understanding of central tendency and the spread of data. Box plots can assist in decision making by revelling the skewness. It is important to note that box plots have limitations too. They may not show the fine details of data distribution.