**Crowdfunding Data Analysis Report**

Large amounts of data are created every minute by both organisations and individual users. This data can only be made into a useful asset if it is analysed, understood, and used to enhance operations or benefit users. A significant number of web-based crowd-funding initiatives are one such source that contributes a lot of data each year.

The use of crowdfunding to introduce new items and create buzz is growing with fast pace, but not every initiative has been successful. So, with the help of given data we’ll be able to analyse different trends. The database consists of 1000 sample projects.

Based on campaign dataset, we can perform certain analysis and make judgement on how successful the campaign will be and what other factors we can possibly consider to make data more useful.

**Question & Answers**

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?

* World music and audio category has best success to failure ratio then any other category with 100% successful projects and 0% failure.
* Plays category have unfavorable success to failure ratio with 187 successful projects and 132 failed projects. Even though they have the most profitable initiatives, they also have the most unsuccessful ones.
* Throughout the span of 10 years from 2010 to 2020 the highest crowdfunding we got was in USD, so on the basis of that we can say U.S stands top among all other country in crowdfunding.

1. What are some limitations of this dataset?

- Consideration of data from a wide variety of years

- More variables should be considered when calculating success

rates.

1. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

* We can create pivot table using year in rows and outcome in values.
* With the help of this picot table we can create bar graph which can

provide summarized data of yearly based progress of outcomes.

1. Use your data to determine whether the mean or the median better summarizes the data.

- For normal number distributions, which contain few outliers, the mean is taken into account. I think Mean summarizes the data better

1. Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

* The mean for successful campaigns is 851.14, whereas standard deviation is around 1266.24, showing that there is a significant difference between the outcome data and the statistical average, making it less credible. Hence less accuracy is achieved since the values are not particularly close making data unreliable.