**Overview Of Project**

Crowdfunding platforms like Kickstarter and Indiegogo have surged in popularity since the late 2000s. They offer a unique opportunity for creators, from individuals to celebrities, to fund their projects. However, not every endeavor succeeds. Success hinges on factors like funding goals, campaign duration, backer engagement, compelling project descriptions, attractive reward tiers, consistent updates, social media presence, and creator reputation. By meticulously analyzing a database of 1,000 sample projects, we aim to uncover hidden trends, shedding light on what separates triumph from failure in crowdfunding. Our findings will provide actionable insights for future project creators, offering a roadmap to crowdfunding success.

**Aim or Purpose**

The purpose of this project is to analyze a database of 1,000 crowdfunding campaigns to identify hidden trends and key success factors on platforms like Kickstarter and Indiegogo. By examining funding goals campaign duration, backer engagement, and other variables, the goal is to provide valuable insights and recommendations for project creators to increase their chances of success in the competitive crowdfunding landscape. Ultimately, this analysis aims to demystify the "trick" to successful crowdfunding and assist organizations and individuals in achieving their funding goals.Top of Form

**Step by Step analysis details for identifying Market Trends**

**Campaign outcome vs funding**

* Chose the column name Outcome. From there select the campaign names: “Successful”, “Failed”, “Currently live” and “Cancelled”. Used different color to highlight them.
* Made a new column with the name of “Percent Funded”, using the formula to fill the money a campaign made relative to its initial funding goal.
* With conditional formatting color coded “Percent Funded” to identify the range of “Percent Funding”.
* X.png is attached in resources folder for reference.

**Outcome of analysis:**

As can be seen, the successful campaign had higher percentage of fundings, and failed and cancelled campaigns had lower percentage of funding.

**Campaign Outcome vs Category**

* Created a new column named “Average Donation” that used formula to find how much each project backer paid on average.
* Created two new columns, one called “Parent Category”, and another called “Sub-Category”, that use formulas to split the “Category and Sub-Category” column into the two new, separate columns.
* Created a Pivot Table based on the “Outcome column”. Filtered the “Country”, Created columns and value to get the “Outcome”. “Put percent” category on “Rows”.
* Made a new sheet with a pivot table that analyzes initial worksheet to the count how many campaigns were “successful”, “failed”, “canceled”, or are “currently live” per category. Also created visualization (column stacked chart) using the Pivot table.

**Outcome of analysis:**

* The distribution of successful, and failed campaigns is mostly similar for all categories. So, the success or failure of campaign did not depend on the category.

**Campaign Outcome vs Subcategory**

* Created a new stacked pivot chart that filtered by “Country” and “Parent Category” based. Rows with “Subcategory” and added “Outcome” in Value and columns.
* By using the formula have created two new columns in the Crowdfunding sheet. Selected “Launched at” column then create new column “Date Created Conversion” . tfor the other new column selected “Deadline” column then made “Date Ended Conversion”.
* Created a new sheet with a pivot table that has a column of outcome, rows of Date Created Conversion, values based on the count of outcome, and filters based on parent category and Years.
* Created a pivot-chart line graph that can visualizes this new table.

**Outcome of Analysis:**

* By Creating a new columns and pivot graph we can see the success and failure data and visualize years and goals. The distribution of successful and failed campaigns is similar for all subcategories. Hence, we can conclude the success vs failure did not depend on subcategory of the campaign.

**Analysis results**

Based on analysis of this data, we will try to answer the following questions.

1. Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?
2. We can draw following conclusions based on our analysis till now.
3. Success or failure of the campaign has not been dependent on category/subcategory of campaign. As we could see the data distribution was similar for all categories between success and failure of a campaigns.
4. We could identify the similarity of percent funded and outcome of the campaign. If a campaign is well funded, its most likely to be successful and failure otherwise.
5. The launched campaigns in the month of June and July had more success rate, while launched in August the success rate was low.
6. What are some limitations of this dataset?
7. There are multiple limitations in the dataset provided,
8. The launch date is provided as timestamp, which is not readable, and multiple calculations had to be performed to read the date, month and year form that.
9. Category and Subcategory was not declared in separate columns and was provided as a single text, which had to be separated.
10. Live outcome is not useful. Live is not outcome of any campaign, which should have been concluded to be used in the computation, but live records are just noise and not very useful to draw any conclusion.
11. Staff Picks and Spotlight data sets were provided which were not useful for deriving any conclusion. Similar for currency and country.
12. Backer data was very limited, where backers played very important role in making a campaign successful or failed.
13. What are some other possible tables and/or graphs that we could create, and what additional value would they provide?
14. We could do following additional analysis to provide better understanding,
15. We could create Country/Currency based graph to visualize the outcome count per country/currency. This would give the information if the campaigns were getting more successful in specific country/region.
16. We could create analysis based on Staff Picks and Spotlight columns, which would indicate if the staff picking the right campaigns and promoting them right or not.
17. Additionally, if we were provided data, we could have looked at the country of backers, or interest areas of backers to identify which backers were better supporting which campaign.

**Crowdfunding Goal Analysis**

* In this part of the challenge, we created a new sheet with 8 columns “Goal” “Number Successful” “Number Failed” “Number Canceled” “Total Projects” “Percentage Successful” “Percentage Failed” “Percentage Canceled”.
* Made 12 rows Starting from the less then 1000 till then 50000 to show all the data points on goals.
* Used CountIf() formula to count the “success”, “failure” and “cancelled” counts. Then added them to count Total, which was used to calculate the percentage of success, failed and cancelled campeigns.
* Created a line graph to show the goal numbers of projects which got “success”, “failure” and “canceled” etc.

**Outcome of Analysis:**

* In this analysis we observed that the campaigns within the goals of 15000 till 35000 had more chances of getting success. Some of the ranges found 100% success in this data set as well. This could be a good goal for future campaign owners to set the goals at that level for achieving success.
* The data also shows the goals between 5000 to 14999 the failure rates were higher.

**Statistical Analysis Backers Count:**

* Created an Excel sheet to evaluate the outcome of successful and unsuccessful campaign.
* Made 4 columns and two different colors to show the outcome, green for successful, and red for failed. The other 2 columns to store the count of backers for each successful and failed campaigns.
* Using formulas calculated “Mean”, “Median”, “Minimum Numbers”, “Maximum Numbers”, “Variance” and “Standard Deviation” for backer count on the dataset.
* By creating this summery we can more understand about the campaigns to evaluate how the number of backers impacted successful and unsuccessful campaigns.

**Outcome of Analysis:**

* **Mean or Median (which better summarizes)**

Median better represents the data set than mean. In both cases of successful and failed backers count, the median is smaller than mean, which represents that majority of numbers are smaller and there are some big numbers in the data set which are outliers, which are impacting the mean calculation.

Median represents the middle number of datasets, and the difference of median from mean shows most of the dataset numbers are smaller, hence median is a smaller number.

* **Which dataset has more variability? (Successful vs failure)**

In this analysis we also calculated variance and standard deviation for successful and failed events.

The variance for successful backer count is 1603373, which is bigger than variance of failed backer count 921574. This indicates that the variability of successful backer count is higher than failed once.

We can also reach this conclusion using the difference between Mean and Median of the successful vs failed events. The Difference between Mean and Median highlights that how disperse the dataset it, and how many outliers are in the dataset. The difference of Mean vs Median is almost 650 for successful dataset, where the same difference for failure is approx. 560, this also indicates the variability is higher for successful data set.