# **Introduction**

This report presents the results of an analysis of the Crowdfunding platforms Book using Excel pivot tables. This analysis aimed to provide a comprehensive view of the relationship between successful and failed projects. The data used in this analysis was generated by edX Boot Camps LLC, and is intended for educational purposes only.

Pivot tables are a powerful tool for organizing and summarizing large data sets, allowing us to explore and analyze data from multiple angles quickly and easily. In this report, I will use datasheet & pivot tables to analyze:

* **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?**

# **Methodology**

The methodology used in this analysis involved the following steps:

1. Data preparation: The raw data was cleaned and organized to ensure that it was ready for analysis. This included checking for missing or inconsistent values and ensuring that the data was in the appropriate format for use in datasheet & pivot tables.
2. Pivot table creation: Pivot tables were created to summarize and organize the data in a way that facilitated analysis. This involved choosing the appropriate rows, columns, and values for the pivot table, as well as any relevant filters or slicers.
3. Pivot table analysis: The pivot tables were then used to explore and analyze the data, looking for trends, patterns, and other insights. This involved manipulating the pivot tables to view the data from different perspectives and using pivot table features such as calculated fields and conditional formatting to enhance the analysis.
4. Data visualization: To help communicate the analysis results, data visualizations such as charts and graphs were created using the pivot table data. These visualizations were used to highlight key findings and make the results of the analysis more accessible and understandable.

# **Results**

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

* 1. Pivot per category shows that the top 3 (by number) successful projects belong to Theater, Film& Video & Music. Using these preliminary results, however, the technology and journalism categories have the highest percentage (66.6 % & 100%) of successful projects. These two categories have the highest chances of succeeding.



* 1. Pivot per Sub-Category demonstrates that the Plays have the highest number of successful projects (54%), and at the same time, it has a high percentage of failed projects (40%).
  2. Goal Analysis chart points out that Ranges 15000-19999, 20000-24999, and 30000-34999 have a 100% of success rate. The ranges >1000;5000-9999;10000 -14999; >50000 have the highest rate of failures. But this chart does not explain the reason behind those numbers but demonstrates the result.
  3. United States is the country with the highest number of successful projects.



**QUESTION 2: What are some limitations of this dataset?**

1. No error control or accidental data change will lead to a complete change of the data in the datasheet.
2. Challenge with Scalability. Excel is not the best program to be used with large data sets and will take a lot of time to process.
3. Additionally, the table only includes information on projects that were launched between January 2010 and Jan 2020, so it may not be representative of Kickstarter projects more generally. Additionally, the table does not include any information on the performance of the projects after they were funded or completed, so it is not possible to determine their long-term success or impact.

**QUESTION 3: What are some other possible tables and/or graphs that we could create, and what additional value would they provide??**

I would add additional information related to the after-funding period of each project to correlate their success or failure in the future to make a more robust decision-making process. Based on the current datasheet, we can judge based on the historical data, and no further information is available to prove further that whoever succeeded in crowdfunding was continue succeeding even after that.

**QUESTION 4: A brief and compelling justification of whether the mean or median better summarizes the data??**

Normally in the data analysis, I would use the mean (Average) value to see an average number of series of values. At the same time, the median demonstrates the value in the middle of the series. It is a measure of central tendency that is not affected by the presence of outliers, while the mean is dependent on the presence of the outliers. In other words, the median will refer to a single number in the middle of a chain values.