Maven Analytics Kickstarter Projects

In this data analytics project, I am analyzing a dataset from Maven Analytics(https://www.mavenanalytics.io/).

\*\* Kickstarter Dataset—This dataset contains projects from different domain from different countries & their current status.

|  |  |
| --- | --- |
| ID | Internal Kickstarter id |
| Name | Name of the project |
| Category | Project category |
| Subcategory | Project subcategory |
| Country | Country the project is from |
| Launched | Date the project was launched |
| Deadline | Deadline date for crowdfunding |
| Goal | Amount of money the creator needs to complete the project (USD) |
| Pledged | Amount of money pledged by the crowd (USD) |
| Backers | Number of backers |
| State | Current condition the project is in (as of 2018-01-02) |

Following the steps of data analytics from Google Data Analytics Professional Certificate to complete this project.

--Ask

--Prepare

--Process

--Analyze

--Share

1. Ask—

In this step we ask proper questions & try to understand the business problems and what are the Stakeholders expectation. Since this is my personal project, So I am trying to understand the data insights without any proper business question but on the basis of my Hypothesis which I have created.

Research Questions: --

What are the variables which affect the project’s success.

How can we make most of the project successful.

Hypothesis: -

Location & Goal have a huge effect on the success of the project.

1. Prepare—

In this step we prepare the data for analysis. Since this is my personal project so I have decided to use the dataset from Maven Analytics. Since this available data is provided by a credible organization, I am assuming that it is secure & have no licensing or privacy issues. Data is almost accurate, complete and reliable for our analysis but it does require some cleaning. I have downloaded the dataset from Maven Analytics in Zip format then extracted the data which contains two csv files containing.

--Data

--Metadata about the data

\*\*Data Exploration Steps—

-First, I looked at the table in the first csv file which contains the data & to understand the data I took help from the csv file containing the Metadata about the data.

|  |  |
| --- | --- |
| Field | Field |
| ID | Internal Kickstarter id |
| Name | Name of the project |
| Category | Project category |
| Subcategory | Project subcategory |
| Country | Country the project is from |
| Launched | Date the project was launched |
| Deadline | Deadline date for crowdfunding |
| Goal | Amount of money the creator needs to complete the project (USD) |
| Pledged | Amount of money pledged by the crowd (USD) |
| Backers | Number of backers |
| State | Current condition the project is in (as of 2018-01-02) |

1. Process—

In this step we Perform required cleaning and preparatory steps on data to get it ready for analysis.

Steps—

1. Firstly, I used Excel to take a first look at the data. Then I moved to SQL and python for further analysis and computations. Then I used tableau for data visualization.
2. I checked for data types of each column, then the data range & data format. I also checked for duplicate data, incomplete data, inconsistent data.
3. I checked for duplicates, removed extra spaces in text columns, removed blanks from each column also in python with the help of data visualization I cleaned the outliers & data accuracy with respect to the information it is representing.

\*\*\*\*Documentation of cleaning and manipulation process---

1. Checked each column for correct data type.
2. Checked each column for blank cell
3. Checked for duplicate rows
4. Checked the country names for accuracy & correctness
5. Checked text columns for correct & duplicate values
6. Checked numerical columns for significant values like in “Backers” column values should not be less than zero

4.Analyze—

I started analysis in Google BigQuery to get an idea about the data & query the required data to analyze it in Excel & create visualization for the report.

<https://console.cloud.google.com/bigquery?sq=1088166550175:77e899a5ac5a4888b4fba32ca8de1f31>

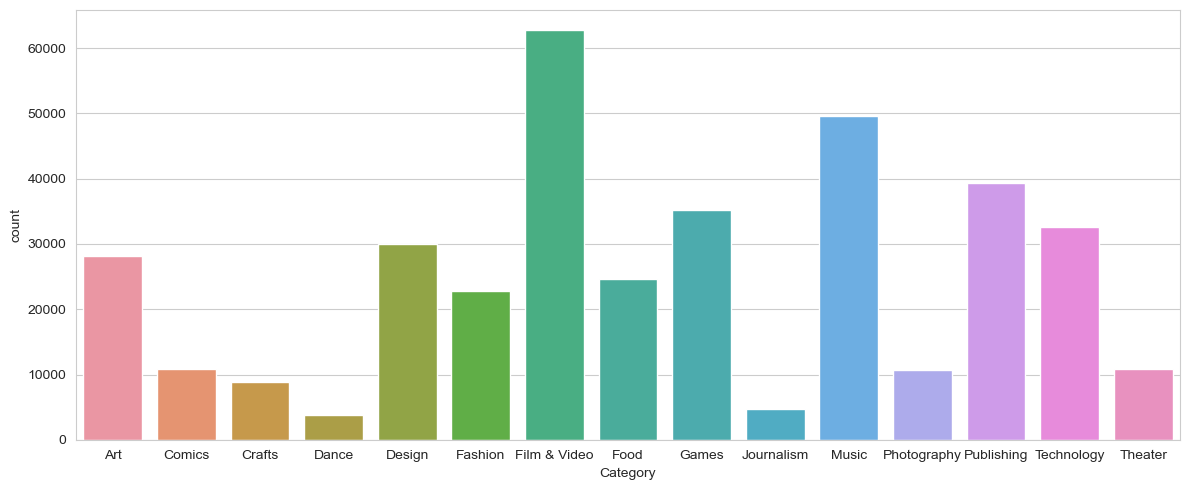
Then I moved to python to analyze the data as it is to analyze and visualize the data in python.

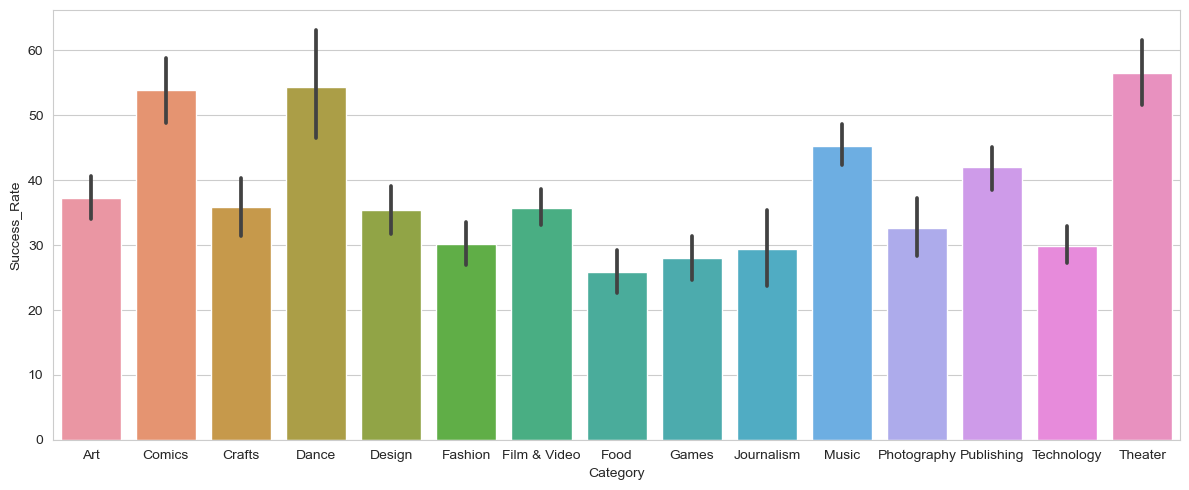
Also, I found that Tableau is a great help in data exploration, analysis & making interactive dashboard.

Analysis Finding & Results

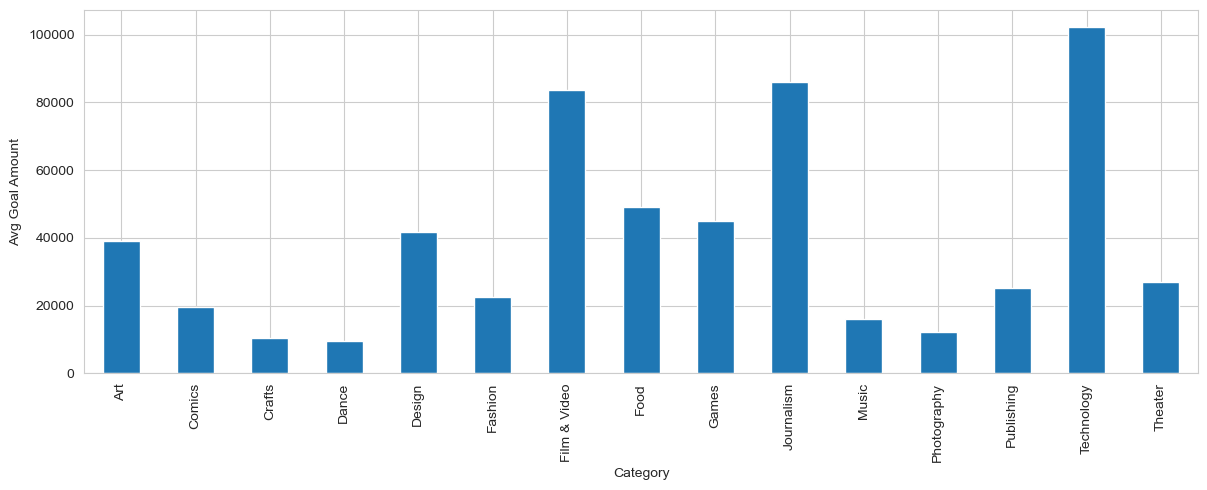
First, I analyzed the data considering the Categorical variables like Category, Subcategory & then I moved to analyze the data according to Country then according to the Goal & Pledged amount.

Overall Category Distribution & Their Success Rate



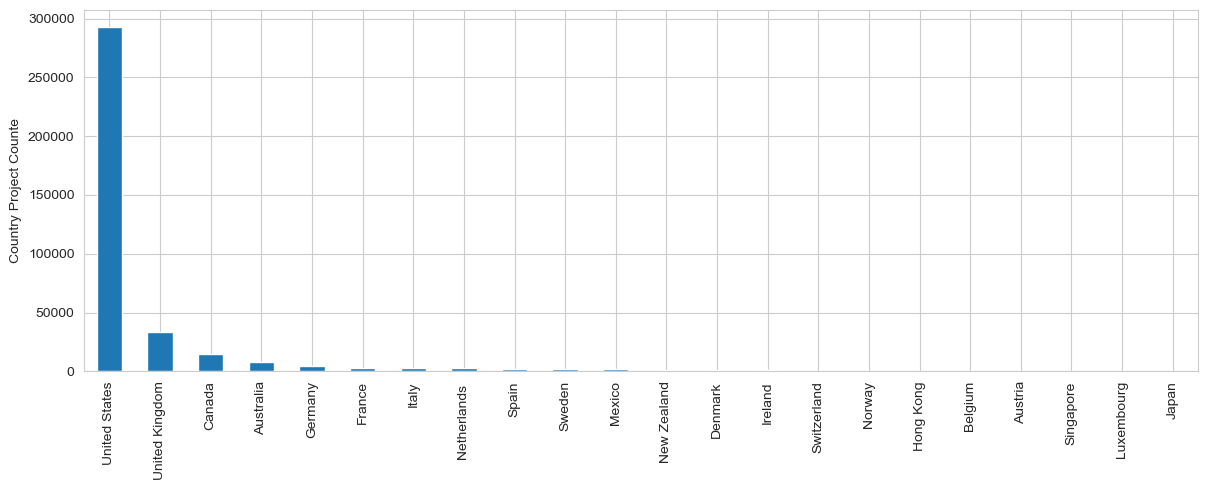


Here we can see that there is no direct relation between Project count Category wise & their Success Rate so we will analyze other parameter which is affecting the Success Rate of the Projects. Also, we can see that the Success Rate of projects are in the range of 20-60 %.

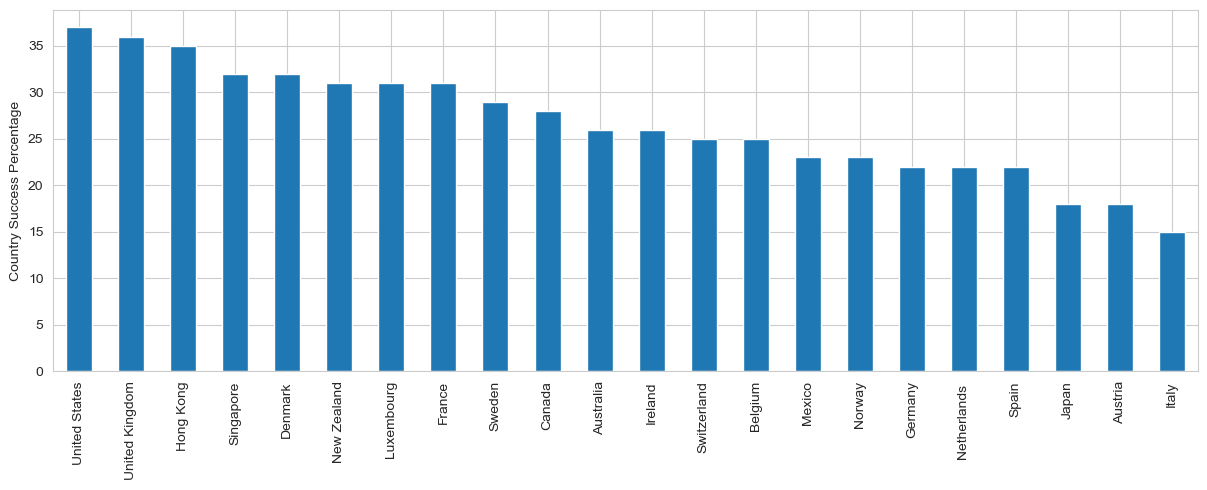


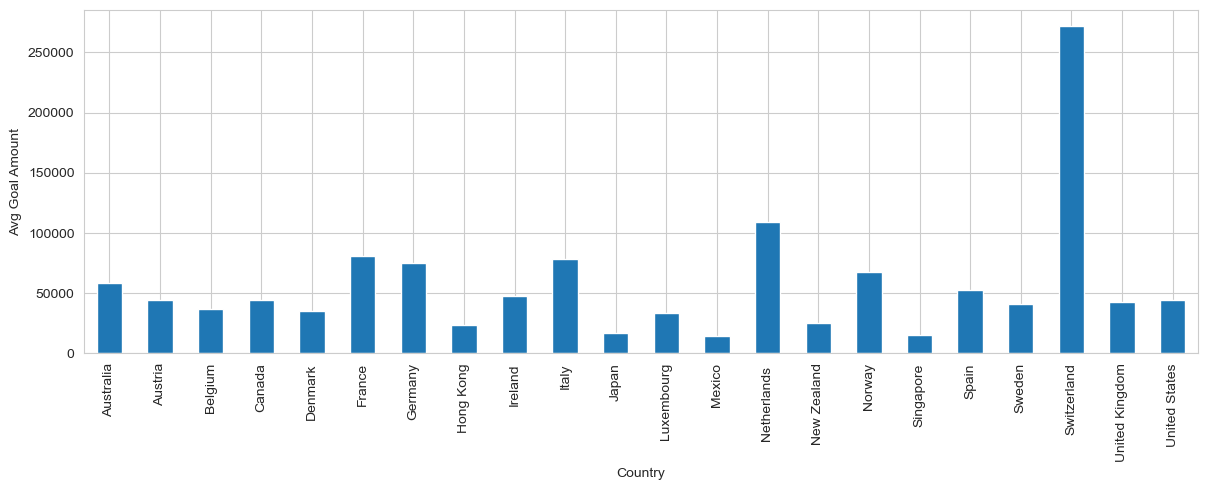
Here we can conclude that the Average Goal Amount does affects the success of the Projects. As here the Category with the comparatively Small Goal Amounts have fairly good Success Rate.

Overall Category Distribution & Their Success Rate Country Wise



Here we can see that the most of the projects are from USA, UK & Canada and some other countries from Europe & Asia have very less projects comparatively.





Here we can conclude that the Average Success Rate is does affected by the Location & Avg Goal as here country with the less Avg goal tends to have more success rate than the countries with comparatively larger Avg Goal.

Tableau Dashboard Link

(https://public.tableau.com/app/profile/sandeep.kumar6589/viz/MavenAnalyticskickstarterprojectsAnalysis/Dashboard2?publish=yes)