PROJECT REPORT

Introduction

The project accesses powerful visualization functions through **Exploratory Data Analysis (EDA)** and statistical methods like **Trend Analysis** to analyze the investment data from the first season of the reality show *Shark Tank India*. The aim was to convert the raw dataset into actionable business intelligence – dashboard through **Power BI**. The primary focus of this analysis is: identifying the **market success rates** and **implied valuations** across various business domains (Idea), and profiling the **investment patterns** of the individual investors (Sharks). It provides the outcome of the quantitative framework for understanding the Indian startup funding landscape as showcased on the program.

Abstract

The project analyzed the pitch entries of the *Shark Tank India* dataset to derive the key investment metrics. The findings revealed interesting revelations about the significant disparities in investment probability and valuation across sectors. The **Pitch Success Rate** for potentially sustainable domains like - *Solar Energy & Healthy Snacks* – was a near-100% deal rate. While the **Post Valuation Analysis** showed sectors like - *Digital Media Content* - commanding the highest post-money worth. Analysis of the **Sharks** (also known as Shark Profiling) led to interesting insights into their investment behavior such as – high volume investors (**Aman**) & high-ticket-sized-investors (**Namita**). The project demonstrates proficiency in data preparation, feature engineering (creating the *Deal_Valuation* metric), and statistical aggregation using the Python data stack.

Tools Used

Programming Language Used : *Python*

Libraries:-

- Pandas - Essential for loading the CSV, handling the **DataFrame** structure, data cleaning, and complex aggregation operations (e.g., **groupby**).
- NumPy - Used for efficient vectorized numerical operations, specifically for implementing conditional logic (**np.where**) to calculate metrics and prevent division-by-zero errors.
- Matplotlib & Seaborn - Used to create static visuals (histograms and bar charts) to display the distribution of successful investment amounts and compare Shark investment activity.

Visualization Tool Used : *Power BI*

- Suggested tool for dashboarding and creating dynamic, interactive reports and visuals like Line Charts, Tree Map & Stacked Bar Chart.

Steps Involved in Building the Project

The project followed a standard data analysis pipeline, with key steps focused on cleaning and engineering financial metrics:

1. Data Acquisition and Cleaning : The CSV file was loaded using – **pd.read_csv()**
2. Feature Engineering : Equity – changed the **messy strings** and converted it into **numeric** values for *efficient calculation*, Deal Status – a binary (**0** or **1**) column obtained to *identify the successful deal pitches*, Post Valuation Calculation – using the formula [**(Investment Account/Equity)*100**]
3. Investment Trend Analysis : The data was grouped by the 'Idea' using – **df.groupby()**, Aggregate Statistics were created - **Total_Pitches**, **Successful_Deals**, and **Total_Investment** – the Success Rate was calculated for all the pitches.
4. Shark Investment Profiling : The dataset was reshaped to separate the investment column using – **pd.melt()**, this data was aggregated to calculate the Total Deals and Average Investment made by each individual Shark.
5. Visualization and Reporting : Final insights were presented using formatted markdown tables and visualizations to communicate the trends and Power BI to create an interactive dashboard.

Conclusion

The project was made successful using Python's ecosystem (Pandas/NumPy/Matplotlib & Seborn) to conduct a deep analytical dive into the Shark Tank India dataset. The engineering of sophisticated financial metrics (like **Post-Money Valuation**), the analysis moved beyond simple counts to generate commercially relevant insights into investor behavior and market preference. The results of the analysis show in which direction the market and investment behavior is moving towards, giving entrepreneurs a peek into their chances of getting an investor who will use these pitches as their benchmark of the investment performance. The exercise demonstrated strong skills in end-to-end data analysis, methods to deal with complex data, and presenting results in visuals to give the accurate picture.