BUSINESS ANALYTICS WITH SPREADSHEET MODELLING

(MGT1058)

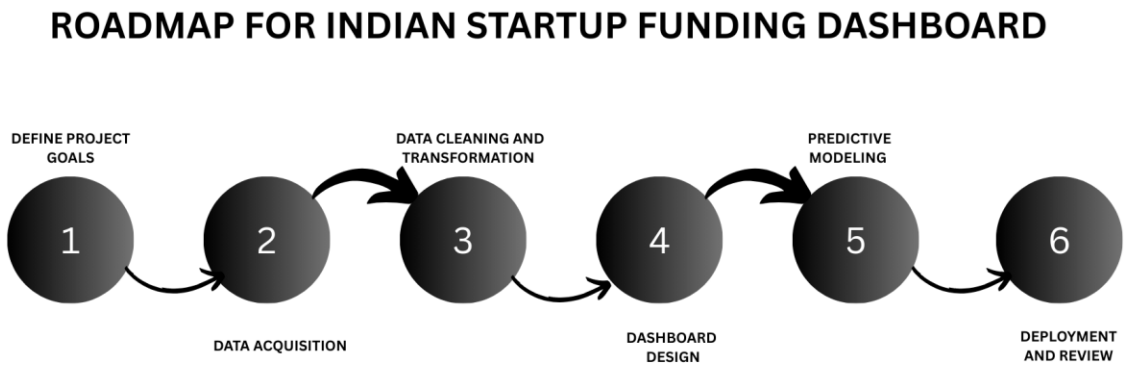
DA-2

TEAM MEMBERS:

R Abishek - 22MIA1003
M Vishal - 22MIA1014
A R Keerthana - 22MIA1017
Phoobesh S - 22MIA1072

INDIAN STARTUP FUNDING DASHBOARD

ROADMAP :



PHASE 1 - Define Project Goals

- **Objective**: Clearly define the goals of the project to provide valuable insights into the Indian startup funding ecosystem.

- The dataset's attributes, such as **Startup Name**, **City**, **Industry**, **Funding Amount**, and **Date**, help identify funding trends, top-performing startups, and leading industries. By analyzing these attributes, the project can uncover investment patterns, geographic hotspots, and industry preferences. Predictive insights into future funding trends can also be derived to support strategic decision-making.

PHASE 2 - Data Acquisition

- **Objective**: Obtain and import the dataset into the project environment.

- The first and foundational step of the project involves gathering a robust dataset. Data was primarily sourced from Kaggle and supplemented by other curated public platforms to ensure comprehensive coverage of startup funding events in India. Once collected, the raw data underwent extensive cleaning and preprocessing. This included handling missing values, standardizing funding amounts, and removing duplicate or irrelevant entries. The goal was to retain only meaningful fields that would support both visualization and predictive modeling, such as startup names, industries, investor names, funding amounts, cities, and investment types

PHASE 3 - Data Cleaning and Transformation

- **Objective**: Prepare the data for analysis by handling inconsistencies and ensuring uniformity.

- With a clean dataset in hand, exploratory analysis was conducted to understand underlying patterns and trends. This involved visualizing funding trends over the years, identifying the most active cities and industries, and determining the most common investment types. Key performance indicators (KPIs), such as total funding per year, average investment size, and the number of investors per round, were derived. EDA helped in identifying areas of focus for the dashboard and informed the direction of the machine learning model.

PHASE 4 - Dashboard Design

- **Objective**: Create a user-friendly interface to visualize and interact with insights.

- The core of the project is the development of a user-friendly and interactive dashboard using **Streamlit** for the web interface and **Plotly** for dynamic visualizations. The dashboard was designed to be intuitive, with multi-tab navigation that separates data exploration from predictive tools. Filters were added to allow users to narrow results by year, city, industry, and investment type. Visual elements like bar charts, line graphs, scatter plots, and pie charts were integrated to provide a rich and engaging user experience.

PHASE 5 - Predictive Modeling

- **Objective**: Enhance the dashboard with forecasting and advanced analytics.

- To enhance the dashboard's functionality, a predictive analytics module was integrated using a combination of machine learning and time series forecasting techniques. Two models were used: **ARIMA**, to capture time-based funding trends, and **Gradient Boosting Regressor**, to estimate funding based on input parameters like year, industry, city, and investment type. The final predicted funding amount is calculated by averaging the outputs of both models. This hybrid approach enables users to get more accurate and reliable funding predictions, offering valuable insights for early-stage planning and strategic decision-making.

PHASE 6 - Deployment and Review

- **Objective**: Deploy the dashboard and gather feedback for iterative improvements.

- Thorough testing was carried out to ensure the dashboard's reliability and accuracy. This included validating the functionality of filters, verifying the correctness of visualizations, and testing the machine learning predictions with known data. Once validated, the application was deployed for public access, ensuring it runs smoothly on different platforms and devices. Emphasis was placed on responsive design and low-latency data interaction.

DASHBOARD DESIGN :

I) VISUALIZATION DESIGN

II) PREDICTION DESIGN