

# CAPSTONE PROJECT OUTLINE DOCUMENT

**PROJECT TITLE:** SYNTHETIC FINANCIAL DATA ANALYSIS

**TEAM NAME:** GROWTH GURUS

**PROJECT MEMBERS:**

- AKANSHA SHETTY
- CHIMIRALA KOWSTUBHA
- KAPAROTU VENKATA SURYA THARANI

**PROJECT GOAL:**

**Objective:**

The project aims to analyze synthetic financial data, with a specific focus on transaction details such as amount, timestamp, and fraud indicators.

Exploratory Data Analysis (EDA): The primary methodology employed involves conducting an EDA on the dataset.

**This process entails:**

- **Characterization:** Understanding the fundamental characteristics of the financial data.
- **Pattern Identification:** Identifying patterns, trends, and distributions within the dataset.
- **Anomaly Detection:** Recognizing any irregularities or outliers that may indicate potential fraudulent activities.
- **Further Investigation:** Highlighting areas within the data that require deeper investigation or analysis.

**Insight Generation:** Through the EDA process, the project aims to generate actionable insights essential for:

- **Informed Decision-Making:** Providing stakeholders with valuable information for making strategic decisions.

- **Risk Management:** Facilitating the identification and mitigation of financial risks associated with fraudulent activities.
- **Business Optimization:** Identifying opportunities for process optimization or improvement within financial operations.

### **Methodology:**

The analysis will utilize statistical techniques, visualization methods, and exploratory tools to comprehensively explore and interpret the financial dataset.

### **Importance:**

This analysis is crucial for enhancing understanding, detecting patterns, and uncovering potential areas of concern within the synthetic financial data. It serves as a foundational step in ensuring data integrity and reliability for subsequent analyses and decision-making processes.

### **PROJECT SCOPE:**

The project involves a thorough examination of synthetic financial transaction data with an emphasis on **pattern recognition** and **fraud detection**.

We seek to gain insights into transaction characteristics through careful **Exploratory Data Analysis** (EDA), facilitating well-informed decision-making.

Accurate fraud detection will be made possible by the development of strong predictive models, which will be made possible by feature engineering.

Meticulous documentation will be used to record every stage of the project, giving future endeavors a transparent and replicable framework.

### **PROJECT DELIVERABLES:**

- Source code for the analysis
- Comprehensive EDA report
- Visualizations
- Recommendations and business perspectives for further exploration
- PPT on insights of the project

## ONUS

- **Insightful Reporting:** Combining all the analysis work and providing insights. Done by **Akansha, Kowstubha** and **Tharani** for their respective works.

**Name - Akansha Shetty**

### **Responsibilities -**

- **Enhanced Descriptive Analysis:** Tasked with succinctly summarizing and vividly portraying the intrinsic attributes of the dataset, enabling comprehensive understanding.
- **Advanced Multivariate Analysis:** Delving into numerous variables concurrently to unveil intricate patterns and interconnections within the dataset, facilitating nuanced insights.
- **Precision Anomaly Detection:** Proficiently spotting aberrant patterns or outliers within the dataset, crucial for pinpointing potential fraudulent activities with utmost accuracy.
- **Communication:** Communicate analysis methodologies and findings, and provide guidance on data-driven decision-making.

**Name - Chimirala Kowstubha**

### **Responsibilities -**

- **Bivariate Analysis:** Will examine the relationship between pairs of variables within the dataset to identify correlations or patterns that may indicate fraudulent behaviour.
- **Correlation Analysis:** Analysis on the correlations between the different variables.
- **Fraud Rate Analysis:** Analysis on the fraud rate of different variables and finding the variables which are responsible for the fraud.

- **Collaboration:** Combining all the analysis work and formatting.

**Name - Kaparotu Venkata Surya Tharani**

**Responsibilities -**

- **Data Preprocessing** - Data preprocessing involves cleaning and preparing the synthetic financial transaction dataset by handling missing values, encoding categorical variables, scaling features, and transforming data to ensure it is suitable for analysis and modelling which aims to enhance data quality, consistency, and compatibility with machine learning algorithms.
- **Univariate Analysis:** Analysis on individual variables within the dataset to understand their distributions and characteristics.
- **Temporal Analysis:** Involves examining trends and patterns in transaction data over time intervals facilitating the detection of suspicious activities and understanding how transaction behaviour evolves over time.
- **Customer Segmentation:** Analysis to identify distinct groups of customers based on transaction behaviours and spending patterns.