

SOLUTION APPROACH

1. Objective Clarity

- Analyse a **highly imbalanced credit card transaction dataset**.
- Extract **actionable insights** that can inform future fraud detection modeling.

2. Structured Workflow

2.1 Data Understanding

- Loaded the dataset and explored columns, types, and distributions.
- Verified **absence of missing values** to ensure clean EDA.

2.2 Exploratory Data Analysis (EDA)

- Checked **class imbalance** (fraud vs. non-fraud).
- Explored **distribution of transaction times** to check when fraud typically occurs.
- Visualized **transaction amounts for fraud and non-fraud cases**.
- Analyzed **hourly transaction patterns** by engineering an Hour feature from Time.

2.3 Feature Engineering

- Created:
 - Hour feature for temporal analysis.
 - Aggregations by Hour and Class to compute:
 - Total transaction amounts
 - Number of transactions
 - Min, max, median transaction amounts

2.4 Visualization

- Bar plots for class distribution.
- Density plots using Plotly to visualize time-based transaction distributions.

- Line plots for hourly trends of:
 - Total transaction amounts
 - Transaction counts
 - Min, max, median amounts

3. Insights Extraction

- Fraudulent transactions are **distributed uniformly across time**, unlike legitimate transactions.
- Fraudulent transactions happen even during **low-activity hours**, indicating stealth patterns.
- Fraudulent transactions are **not consistently high-value**, indicating models must also capture low-value fraud cases.
- Severe imbalance in the dataset necessitates **careful metric selection** for future modeling.