**Credit Card Fraud Detection and Visualization**

**Overview**

This project analyzes credit card transaction data to identify patterns and detect potential fraudulent activities. The project utilizes Python for data preprocessing and Power BI for visualizations to gain actionable insights.

**Project Objectives**

1. Identify high-risk (fraudulent) transactions using statistical methods.

2. Visualize transaction trends, fraud patterns, and geolocation-based insights.

3. Enable dynamic exploration of data through interactive Power BI dashboards.

**Features**

1. Data Cleaning:

- Removed currency symbols from the `Amount` column and converted it to numeric.

- Handled missing and inconsistent data.

2. Outlier Detection:

- Used Z-Score method to identify transactions with values beyond 3 standard deviations.

3. Feature Engineering:

- Extracted additional insights from the `Date` column:

- Year, Month, Day, Day of Week, Hour, and Day Part (e.g., Morning, Evening).

- Created a new column for `Fraud` categorization (e.g., Fraud vs. No Fraud).

4. Visualizations:

- Time-Series Analysis:

- Line chart showing transaction trends over time.

- Fraud Analysis:

- Bar chart comparing fraudulent vs. non-fraudulent transactions.

- Pie chart showing the proportion of frauds.

- Geolocation Insights:

- Heatmap visualizing transactions by country/region.

- High-Risk Analysis:

- Scatter plots to highlight high-risk transactions.

5. Interactive Dashboard:

- Slicers for dynamic filtering by:

- Bank

- Fraud/No Fraud

- Date Range

- Type of Transaction

**Data Description**

- Dataset Name\*\*: `CreditCardData.csv`

- Columns:

- `Transaction ID`: Unique ID for each transaction.

- `Date` & `Time`: Date and time of transaction.

- `Amount`: Transaction amount.

- `Fraud`: Indicates whether the transaction is fraudulent (`1` for fraud, `0` for no fraud).

- `Type of Transaction`: Transaction category (e.g., Online, POS, ATM).

- `Country of Transaction`: Geolocation of the transaction.

- And other relevant fields.

**Tools and Technologies**

Python (for data preprocessing and feature engineering)

- Libraries used:

- `pandas` for data manipulation.

- `matplotlib` and `seaborn` for basic visualizations.

- `scikit-learn` for encoding categorical variables and statistical analysis.

PowerBI

- Built an interactive dashboard with:

- Line charts, bar charts, pie charts, and maps.

- Slicers for filtering by fields like date, bank, fraud, and transaction type.