

Sample Business Requirement:

You are tasked with analyzing credit card fraud transactions to gain insights into patterns, customer behavior, and merchant risks. The analysis should cover transactional details, customer and merchant demographics, fraud trends over time, and possible fraud indicators.

Key Objectives:

1. Identify patterns of fraudulent activity across different regions, time periods, and merchant categories.
2. Provide insights into high-risk merchants and customers.
3. Track the correlation between customer distance and fraud likelihood.
4. Highlight potential fraud indicators such as anomalous transaction volumes or inconsistencies in customer data.
5. Offer detailed exception reporting for further investigation by analysts.

Business Questions:

- How many transactions are fraudulent? What percentage of total transactions do they represent?
- Which cities and states are fraud hotspots, and how does population size correlate with fraud?
- What is the relationship between customer-merchant distance and the likelihood of fraud?
- Which merchants or categories are most associated with fraudulent transactions?
- Can patterns in customer behavior (e.g., multiple names per CC number) or transaction volumes highlight potential fraud?

Sample Reports / Dashboard Layout:

1. Overview Dashboard (Fact Table: Fact_Transactions, Dimensions: Dim_Customer, Dim_Merchant, Dim_Location)

- **1.1 Filters** (Time and Geography Dimensions):
 - Month
 - Date Range
- **1.2 Transaction Demographics:**
 - **Total Transaction Volume** (from Fact_Transactions): Sum of all transactions.
 - **Fraud Percentage** (from Fact_Transactions): Calculated as [Number of Fraudulent Transactions / Total Transactions].
 - **City Population vs Transaction Volume** (from Dim_Location and Fact_Transactions): Scatter plot comparing population size with transaction volume.
 - **City Population vs Fraud Volume**: Similar to above but focused on fraudulent transactions.
 - **Customer-Merchant Distance vs Fraud** (from Fact_Transactions): Analyze the correlation between distance and fraud occurrence.
 - **Last Data Refresh Date**: Display last data refresh timestamp.

- **1.3 Business Metrics:**

- **Fraud Over Time** (from Fact_Transactions): Line chart displaying fraud count across time (e.g., daily, monthly).
- **Fraud by State** (from Dim_Location and Fact_Transactions): Map visualization with fraud per state.
- **Fraud by Merchant** (from Dim_Merchant and Fact_Transactions): Treemap to highlight top merchants linked to fraud.
- **Fraud by Category** (from Dim_Merchant): Donut chart showing fraud volume by merchant category.

2. Detailed Analysis (Fact Table: Fact_Transactions, Dimensions: Dim_Customer, Dim_Location)

- **2.1 Transaction Patterns:**

- **Hourly Transaction Distribution:** Heat map of transactions by time of day.
- **Day of Week Analysis:** Column chart analyzing transactions by the day.
- **Transaction Amount Distribution:** Histogram of transaction amounts.

- **2.2 Geographical Insights:**

- **Fraud Hotspots:** Heat map of cities or regions with high fraud rates.
- **State-wise Fraud Rate:** Filled map displaying fraud intensity by state.
- **City Population vs Fraud Rate:** Scatter chart linking city population to fraud rates.

- **2.3 Customer Segmentation:**

- **Age Group Analysis** (from Dim_Customer): Column chart showing fraud by age groups.
- **Gender Distribution** (from Dim_Customer): Pie chart showing fraud breakdown by gender.
- **Top Fraud-Prone Customer Segments:** Table showing customer segments (e.g., high transaction volume, frequent fraud).

3. Merchant Analysis (Fact Table: Fact_Transactions, Dimension: Dim_Merchant)

- **3.1 Merchant Performance:**

- **Top Merchants by Transaction Volume:** Bar chart of top merchants.
- **Merchant Category Performance:** Stacked column chart showing merchant categories and fraud distribution.
- **Merchant Fraud Rate Comparison:** Scatter chart comparing merchants by fraud rates.

- **3.2 Merchant Risk Assessment:**

- **High-Risk Merchants:** Table showing merchants with the highest fraud count.
- **Merchant Fraud Trend:** Line chart showing fraud over time per merchant.
- **Merchant Category Risk Matrix:** Matrix visual showing fraud by merchant category.

4. Exception Reports (Fact Table: Fact_FraudTransactions, Dimensions: Dim_Customer, Dim_Merchant, Dim_Location)

- **4.1 Fraud Customer List:**

- Detailed customer info with fraudulent transactions.
- Transaction history per customer (timeline view).
- Associated merchants (network graph).
- **4.2 Fraud-Related Merchants:**
 - **Merchant Fraud Summary:** Table showing merchant fraud count and total transaction volume.
 - **Fraud Occurrence Trend:** Sparklines in table showing trends.
 - **Transaction vs Fraud Rate:** Scatter chart comparing transaction volume to fraud rate.
- **4.3 Potential Fraud Indicators:**
 - **Multiple Names per CC Number:** List of customers using the same card with different names.
 - **Anomaly Transaction Volume:** Conditional formatting to highlight unusual transaction volumes by location.
 - **Unusual Transaction Patterns:** Chart showing anomalies detected in transaction patterns.

Star Schema Model:

1. **Fact_Transactions:** Contains transaction-related data such as transaction_id, amount, trans_date, fraud_flag, customer_id, merchant_id, location_id, distance_km.
2. **Dim_Customer:** Includes customer details such as customer_id, name, age, gender, address.
3. **Dim_Merchant:** Includes merchant details such as merchant_id, merchant_name, category, city.
4. **Dim_Location:** Contains location information such as location_id, city, state, population.