**CLIENT REQUIREMENTS**

**🔹 Fraud Analytics**

1. **Fraud Trends Over Time**
   * How does the number (and % share) of fraudulent transactions change by **day, week, or month**?
   * Do we see spikes in fraud attempts at particular times (holidays, weekends, year-end)?
2. **High-Risk Transaction Types**
   * Which **transaction types** (DEBIT, CASH\_OUT, TRANSFER, etc.) have the highest fraud rate?
3. **Geographical Fraud Patterns**
   * Are certain **branches/locations** (e.g., Kathmandu vs Syangja) more prone to fraud?
   * Can we map fraud hotspots?
4. **System Accuracy**
   * Compare is\_fraud vs flagged\_fraud:
     + True Positives, False Positives, False Negatives, True Negatives.
   * Calculate **precision, recall, accuracy, and F1-score** for the fraud flagging system.

**🔹 Customer Behavior**

1. **Top Customers by Transaction Volume & Value**
   * Who are the top 10 customers with the highest **transaction amounts** and **number of transactions**?
2. **Account Type Insights**
   * Compare behavior of **Savings vs Current accounts**:
     + Average balance before/after transactions.
     + Average transaction size.
3. **Transaction Timing Patterns**
   * Which **time of day** (Morning, Afternoon, Night) sees the most activity?
   * Do frauds cluster at certain times of day?

**🔹 Financial Analytics**

1. **Cash Flow Analysis**
   * Track **balances before vs after transactions** to understand customer liquidity trends.
2. **Destination Accounts**
   * Identify accounts that **receive large inflows** repeatedly (possible mule accounts).
3. **Unusual Login Patterns**

* Do customers with unusual\_login = 1 have higher chances of fraud?

**🔹 Predictive Insights**

1. **Fraud Risk Factors**

* Which features (amount, transaction type, branch, unusual login, time of day, etc.) are **most correlated with fraud**?

1. **Fraud Probability Model** *(optional, if ML allowed)*

* Build a simple model to predict the probability of fraud based on transaction features.

✅ Deliverables I (as your client) would expect:

* Cleaned dataset with handled NULLs.
* Summary tables & visualizations (fraud rates, top customers, branch comparison).
* Fraud detection system evaluation (precision/recall).
* Business recommendations (e.g., “CASH\_OUT in rural branches at night = higher fraud risk”).

Would you like me to **prioritize these into a roadmap** (step 1 → step 2 → …) so you can start with the essentials (data cleaning, fraud trend reports) and then move towards advanced analytics (fraud prediction)?