Financial Risk Analysis with Python

Project overview:

You are hired as a Financial Data Analyst by Goldman Sachs, a major global banking institution. As digital transactions rise and customers diversify their account usage, the executive team wants to uncover patterns in customer financial behavior, account performance, and transaction risks to improve customer service and minimize financial risk.

The leadership is particularly interested in understanding:

- How do customers interact with different account types (e.g., savings, current, credit)?
- What are the trends in debit vs. credit transactions across segments?
- Can we identify accounts with unusually high risk or inconsistent financial activity?
- What transaction behaviors correlate with lower account balances or overdraft incidents?

Objective: Build a complete Customer Financial Behavior and Risk Analysis Report in Excel using Python for data handling. This report should clean and consolidate Barclays' customer account data, uncover transaction insights, and help decision-makers align customer strategies with usage trends and financial risk indicators

Dataset: **Goldman Sachs Financial Transactional Data**

Dataset Overview: This dataset provides a transactional summary of Goldman Sachs' customer accounts, including account numbers, transaction dates, transaction types (credit/debit), transaction amounts, account types, and available balances.

Column Name	Description
TransactionID	Unique identifier for each transaction
CustomerID	Unique identifier for each customer
AccountID	Unique identifier for each bank account
AccountType	Type of bank account (e.g., Savings, Loan, Credit)
TransactionType	Nature of transaction (e.g., Deposit, Withdrawal, Transfer, Payment)
Product	Financial product involved (e.g., Credit Card, Mutual Fund, Home Loan)
Firm	Name of the Firm where the transaction occurred
Region	Geographical region of the branch (e.g., East, West, North)
Manager	Branch manager responsible for the transaction or customer
TransactionDate	Date when the transaction occurred
TransactionAmount	Amount involved in the transaction
AccountBalance	Account balance after the transaction
RiskScore	Risk score associated with the customer or transaction
CreditRating	Credit rating score of the customer
TenureMonths	Duration (in months) the customer has held the account or relationship

Tasks to be performed:

Task 1: Data Cleaning and Formatting (10 Marks)

- Remove/treat any special characters or non-numeric entries from financial fields.
- Convert currency amounts into numerical format.
- Validate and format date columns.
- Ensure account types and transaction categories are standardized.

Task 2: Descriptive Transactional Analysis (20 Marks)

- Calculate monthly and yearly summaries of total credits, debits, and net transaction volume.
- Plot trends in total credits vs. debits over time.
- Identify top and bottom performing accounts based on net inflow.
- Identify and flag accounts as dormant or inactive if there is a gap of two months or more between consecutive transactions.

Task 3: Customer Profile Building (20 Marks)

- Group accounts by activity levels: High, Medium, Low based on transaction frequency on your analysis and rubrics. Do not forget to mention the rubric in the headings.
- Segment customers by average balance and transaction volume.
- Create profiles for:
    - High-net inflow accounts
    - High-frequency low-balance accounts
    - Accounts with negative or near-zero balances

Task 4: Financial Risk Identification (10 Marks)

- Track accounts with frequent large withdrawals or overdrafts.
- Calculate balance volatility using standard deviation or coefficient of variation.
- Use IQR or z-score methods to detect anomalies.
- Highlight customers with irregular or suspicious transaction behavior.

Task 5: Visualisation (10 Marks)

- Conduct extensive exploratory data analysis with attractive visualizations for your findings

Task 6: Hypothesis Testing (10 Marks)

- Test whether high-volume transaction accounts have statistically higher average balances than low-volume accounts.
- Conduct hypothesis testing based on segmentation.

Task 7: Video Presentation (20 Marks)

- Record a short video summarizing findings and insights.
- Highlight what drives customer transaction behavior and financial risk.
- Discuss data-backed recommendations for customer engagement or monitoring.

Note:
The explanation should be in your own words. Avoid scripted or AI-generated narration.
Use charts from your analysis as visual aids in the video.
If you set your own criteria (e.g., for flagging risk years or segmenting periods), clearly mention them in the workbook.
Maintain separate workbooks/sheets for each task.
Ensure plagiarism-free content. Original work will be appreciated and rewarded.

Submission Guidelines:

- Follow basic coding ethics (headlines, pointers for insights and comments wherever required)
- Place all code files, and summary PDF in one folder
- Add the video link in a text or Excel file
- Compress the folder into a .zip file
- Upload the zipped folder to your project dashboard