श्री रामस्वरूप
मेमोरियल यूनिवर्सिटी

SHRI RAMSWAROOP
MEMORIAL UNIVERSITY

Topic - Banking Data Analysis using (Pandas & Matplotlib)

Submitted By -
Syed Mohd Raza Rizvi (086)
Arman Rizvi (076)

Submitted to – Mrs. Deepika Ma’am

Subject – Data Analytics & Reporting
Group – 33rd

OBJECTIVE

To clean, organize, and visualize *banking customer data* using **Microsoft Excel** and **Python (Pandas + Matplotlib)** to make it structured, reliable, and easy to analyze.

- **Work done Using Pandas and Matplotlib**
  - **Data Cleaning with Pandas**
    Loaded raw Excel data using pd.read
    Removed duplicate and null entries.
    Filled missing values:
  - **Data Analysis**
    (Describing,info,Sorting,Grouping)
  - **Data Visualization with Matplotlib**
- **Work Done Using Microsoft Excel**
  - **Data Inspection & Formatting**.
  - **Conditional Formatting**
  - **Pivot Tables & Summaries**
  - **Trend Observation**

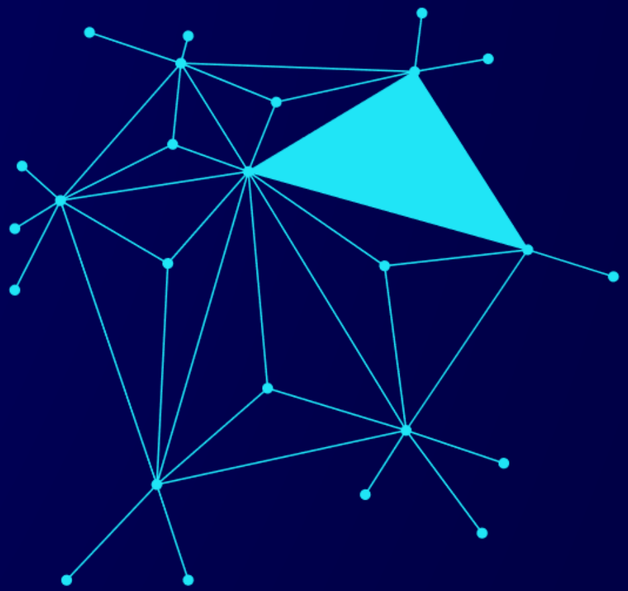
Tools and Library Used

Tool	Purpose
Pandas	Data handling, cleaning, grouping
Matplotlib	Visualizations (bar, histograms, etc.)
Excel	Pivot Tables, Conditional Formatting
Python (Jupyter/VS Code)	Code execution and reporting

Data overview

Dataset Name –
Banking_Data (2500 Records, 14 Columns)

Dataset Description
This dataset contains **bank customer details** such as demographics, account types, balances, and credit-related information.
It is used to perform **data cleaning, visualization, and financial trend analysis** using Excel and Python (Pandas + Matplotlib).



Initial Insights
Dataset is **balanced and clean**, minor missing values in four columns.
Balance and **Income** show **high variability** (large range).
Age range: 18 – 69 years → good diversity among customers.
Credit Score mostly between 300–900, ideal for financial risk analysis

Data overview

Dataset Structure

Column	Description	Data Type
Customer_ID	Unique customer identifier	Object
Customer_Name	Customer full name	Object
Age	Age in years	Numeric
Gender	Gender (Male/Female)	Categorical
City	Customer’s city	Categorical
Job	Occupation	Categorical
Marital_Status	Marital status	Categorical
Education	Education level	Categorical
Account_Type	Type of account	Categorical
Balance	Account balance	Numeric
Credit_Score	Customer’s credit score	Numeric
Loan	Loan category	Categorical
Response	Customer feedback / response	Categorical
Annual_Income	Annual income (₹)	Numeric

Key Statistics (From Pandas Analysis)

Metric	Observation
Total Rows (Records)	2500
Total Columns	14
Missing Values	25 missing each in Education, Balance, Credit_Score, and Loan
Duplicate Rows	0
Numeric Columns	Age, Balance, Credit_Score, Annual_Income
Categorical Columns	10 columns
Average Age	43.7 years
Average Balance	₹ 4.94 lakh
Average Credit Score	597
Average Annual Income	₹ 10.6 lakh

Excel Work

Objective 1: Clean, organize, and format raw banking data in Excel

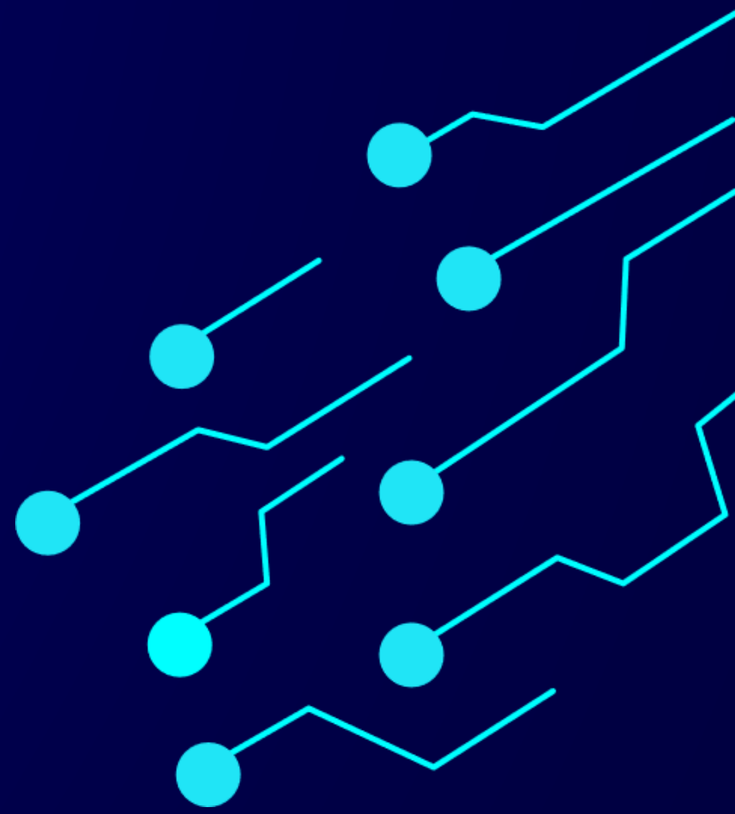
Step 1: Open and inspect the data
Check for:
- **Missing values** (NaN, blank cells)
- **Duplicates**
- **Inconsistent data** (e.g., Gender = “M” and “Male” both)

Step 2: Remove duplicates
1.Select all data.
2.Go to **Data** → **Remove Duplicates**.
3.Select all columns → **OK**.

Step 3: Handle missing data
1.Use **Filter** to find blank cells.
2.Fill them logically (e.g., unknown → “Not Provided”)
or leave blank if not known.

Objective 2: Apply Conditional Formatting for Key Metrics

Condition	Where to Apply	Formatting
Balance > 5,00,000	Balance column	Green fill
Credit_Score < 500	Credit_Score column	Red fill
Loan = “yes”	Loan column	Light Blue fill
Annual_Income top 10%	Annual_Income column	Gold fill



Steps:

1. Select column → Go to **Home** → **Conditional Formatting**.
2. Choose **Highlight Cell Rules** or **Top/Bottom Rules**.
3. Set your conditions and color styles

Objectives 3 - Use Pivot Tables and Charts

Step 1:
Insert Pivot Table
1. Select entire data table.
2.Go to **Insert → PivotTable**.
3.Place Pivot Table on a new sheet.

Step 2:
Create Key Pivot Tables
Examples:
1.Average Credit Score by City
1. Rows → City
2. Values → Average of Credit_Score

2.Total Balance by Account Type
1. Rows → Account_Type
2. Values → Sum of Balance

3.Loan Approval by Gender
1. Rows → Gender
2. Columns → Loan
3. Values → Count of Customer_ID

Step 3:
Add Charts
1.Click Pivot Table → Go to Insert → Recommended Charts.
2.Use:
1. **Column chart for city-wise comparison**
2. **Pie chart for loan approval rate**
3. **Line chart for age vs credit score**

Pandas Work

Data Overview (Using Pandas

1 Data Loading
• Dataset imported in Pandas from Excel file by pd.read(“Excel file”)
• Verified data by displaying top 5 rows by Rizvi.head()

2 Dataset Summary
• Checked total rows and columns (data size).
• Listed all column names.
• Viewed data types and non-null counts using info().
• Generated statistical summary (mean, min, max, etc.) using describe()

3 Missing Value Analysis

- Checked if any missing values exist.
- Counted missing values in each column.
- Filled missing categorical and numeric values appropriately.

4 Duplicate Handling

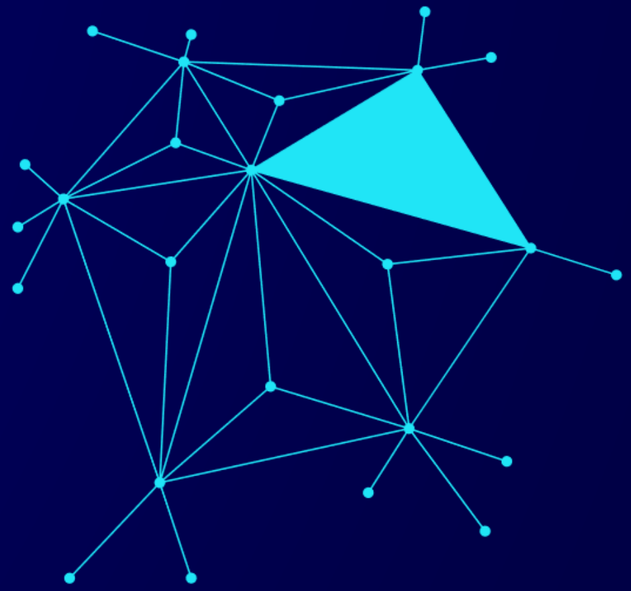
- Checked for duplicate records.
- Removed all duplicate rows to ensure data consistency.

5 Balance-Based Insights

- Identified customers having balance more than ₹5,00,000.
- Sorted records based on descending order of balance.
- Calculated average balance according to account type.

6 Outcome

- Dataset cleaned and organized.
- Missing and duplicate values handled.
- Data ready for further visualization and analysis



Thank You

Have a any Question ?