INNOVATIVE ASSIGNMENT

ANSHUL PATEL(23BEI007)

BHAVISH GUPTA(23BEI016)

INTERIM SUBMISSION

COURSE-DATA ANALYSIS AND VISUALISATION

1) Identify the Problem for your Application Domain

Managing personal finances is often challenging for individuals due to multiple expense categories (rent, groceries, entertainment, etc.) and financial goals (savings, debt repayment, etc.). Many people struggle to track spending patterns, identify overspending areas, and compare desired savings with actual savings.

Problem: Lack of a structured, data-driven way to monitor income, expenses, and savings opportunities.

2) Find the Dataset with Raw Data

- The dataset you're using has **20,000 rows & 27 columns** containing:
    - **Income & Expenses** (Rent, Insurance, Groceries, etc.)
    - **Savings Goals** (Desired_Savings, Desired_Savings_Percentage)
    - **Potential Savings per Category**
    - **Demographics** (Age, Occupation, City_Tier, Dependents)
- This acts as the **raw financial dataset** for the project.

---

3) Set the Objectives for the Study

1. Track **total expenses vs income** for individuals.
2. Compare **desired savings vs actual disposable income**.
3. Identify **overspending areas** using expense breakdown.
4. Estimate **potential savings** if expenses are optimized.
5. Provide **visual insights** via Tableau dashboard for easy decision-making.

---

4) Scope of the Study & Stakeholders

Scope:

- The study covers personal finance management at an individual level.
- It identifies spending trends, savings gaps, and potential opportunities.
- Outputs are presented in an **interactive Tableau dashboard** for better accessibility.

Stakeholders:

- **Primary Users (Individuals)** → Want to track income, expenses, and savings.
- **Financial Advisors** → Use insights to suggest better financial planning.
- **Banks/Fintech Apps** → Can integrate insights for customer engagement.
- **Researchers/Analysts** → Analyze financial behavior patterns.

---

5) Know Your Data Characteristics & Make Plots

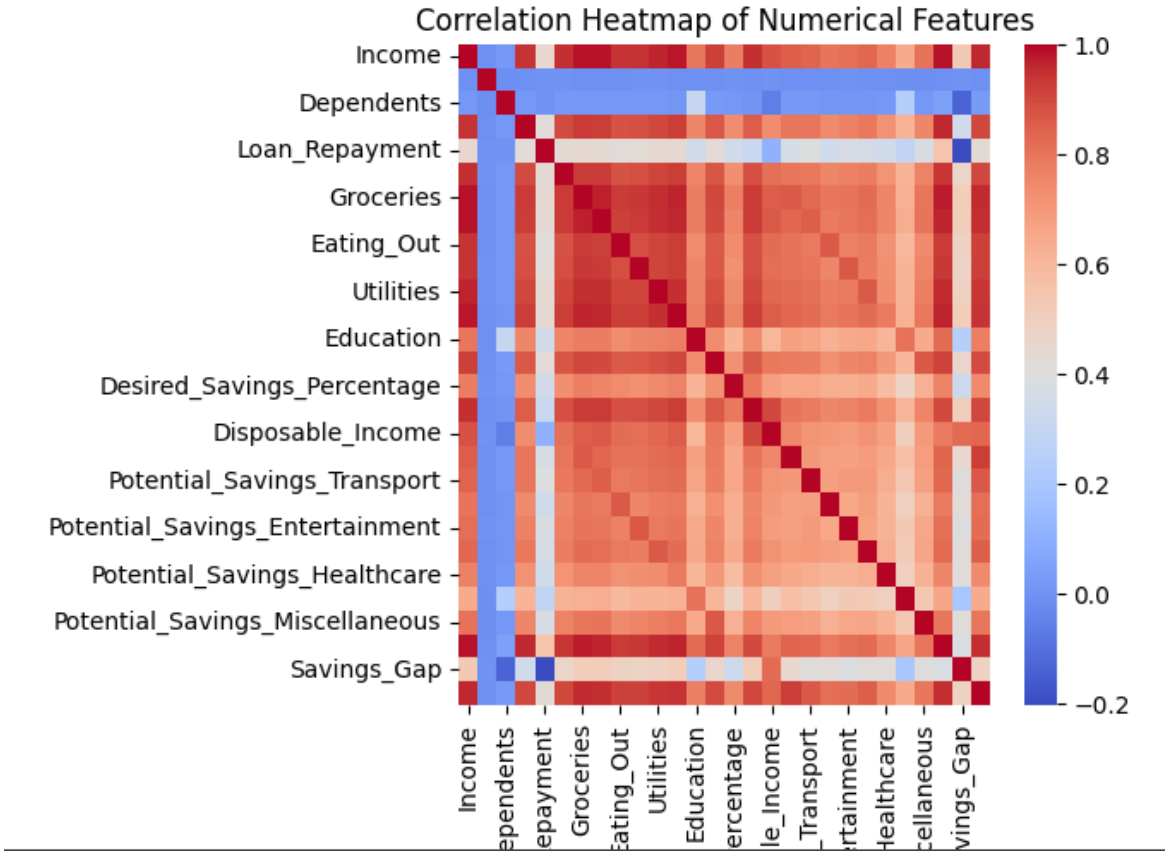
- **Data Characteristics:**
    - Numerical: Income, Expenses, Savings, Potential Savings.
    - Categorical: Occupation, City_Tier.
    - Distribution ranges from very small to extreme outliers (e.g., Rent, Income).
- **Plots to Explore Data:**
    - Histogram → Income, Disposable Income distribution.
    - Boxplots → To detect outliers in expenses and income.
    - Correlation Heatmap → Relationship between income, expenses, and savings.
    - Pie/Bar Chart → Expense distribution by category.

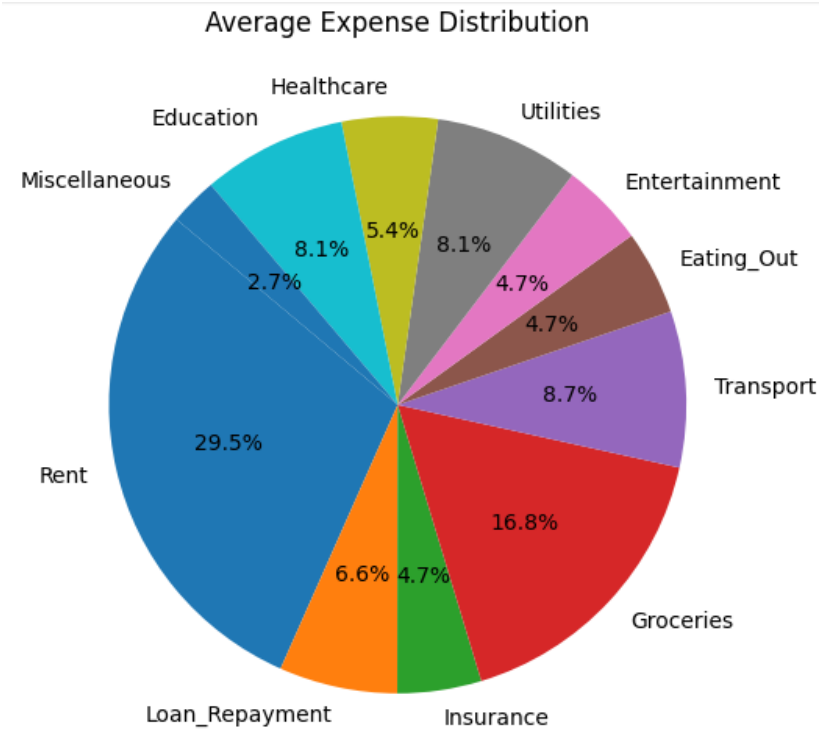
---

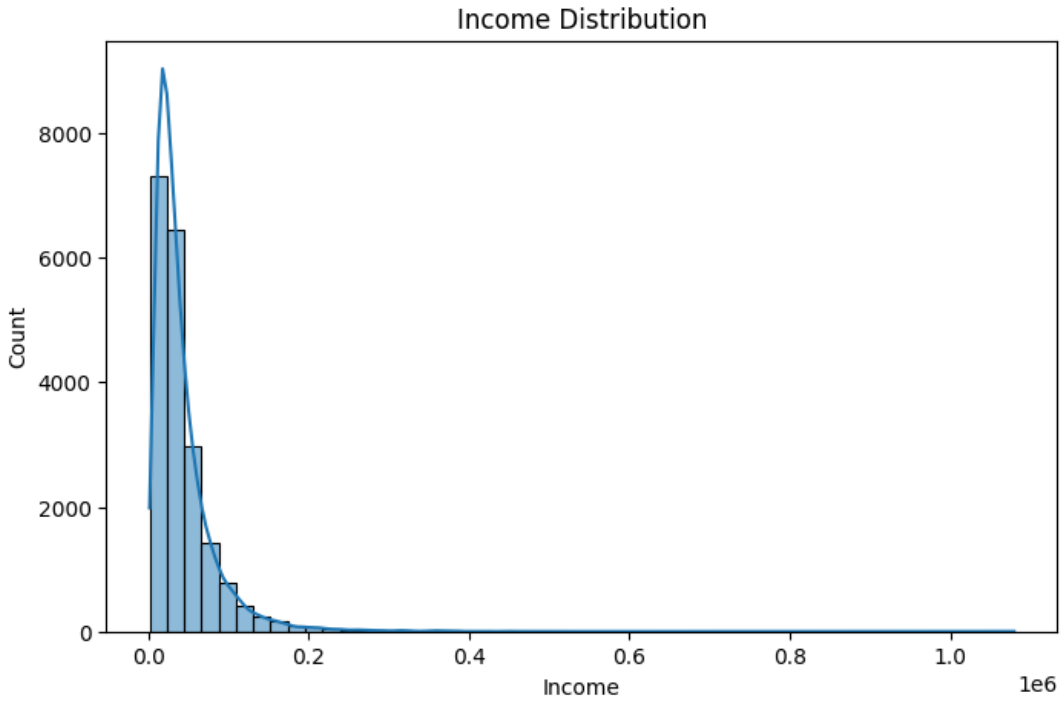
6) List the Questionnaire / Needs of Each Stakeholder

- **Individuals:**
    - Am I spending more than I earn?
    - Which category takes most of my money?
    - Am I meeting my savings goals?
    - Where can I save more?

- **Financial Advisors:**
    - Which expense categories need optimization for a client?
    - What percentage of clients are not meeting savings goals?
    - How much potential savings can be unlocked?
- **Banks/Fintech Apps:**
    - What are common spending patterns across city tiers?
    - Which customer segments have the highest disposable income?
    - How can personalized financial products be offered?
- **Researchers/Analysts:**
    - How do demographics (Age, Occupation, Dependents) affect savings?
    - What correlations exist between income, expenses, and savings gap?
    - What factors drive financial health?




Average Expense Distribution
Healthcare
Education
Utilities
Miscellaneous
Entertainment
Eating_Out
Transport
Rent
Groceries
Loan_Repayment
Insurance
8.1%
5.4%
8.1%
2.7%
4.7%
4.7%
8.7%
29.5%
16.8%
6.6%
4.7%


Income Distribution
Count
8000
6000
4000
2000
0
0.0
0.2
0.4
0.6
0.8
1.0
Income
1e6