Business Document

1.Table (Bank_data):

Branch_ID	City	Region	Firm_Revenue	Expenses	Profit_Margin
1000	Solapur	East	422443	346471	56.31
1001	Mumbai	North	211155	216256	79.53
1002	Nashik	West	985006	71777	55.31
1003	Solapur	North	55113	253655	2.57

Purpose:

Stores branch-level geographic and financial data.

Row Meaning (Grain):

Each row represents:

- One branch (*Branch_ID*)
- Its location (*City, Region*)
- Its financial metrics (*Firm_Revenue, Expenses, Profit_Margin*)

Observations / Notes:

- There is no date information indicating when these financial figures were captured.
- There is no explanation of how *Profit_Margin* is calculated or whether it is derived from Firm_Revenue and Expenses.
- Firm_Revenue and Expenses do not appear to be directly related mathematically.

Business Assumption:

Each row represents:

- The latest snapshot of each branch's financial status
- Data is assumed to be updated periodically (e.g., monthly)
- Financial metrics are calculated using unknown business logic and should be treated as independent measures.

2.Table (Customer_data):

Customer_ID	Age	Customer_Type	City	Region	Bank_Name	Branch_ID
200000	56	Employee	Pune	North	HDFC Bank	1350
200001	69		Kolhapur	East	HDFC Bank	1648
200002	46	Business	Pune	West	HDFC Bank	1798
200003	32	Individual	Pune	West	HDFC Bank	1900

Purpose:
Stores customer demographic, geographic, and banking information.

Row Meaning:
Each row represents:
- One *Customer_ID*
- Demographic attributes (*Age*)
- Geographic attributes (*City, Region*)
- Banking attributes (*Bank_Name, Branch_ID, Customer_Type*)

Observations / Notes:
- *Customer_ID* appears to be a primary key, which implies:
  - One customer belongs to only one bank
  - One customer belongs to only one branch
  - One customer has only one *Customer_Type*
- This raises a concern:
  - What if a real customer has multiple accounts, across multiple branches?

Business Assumption:
Each row actually represents:
- One customer account, not a real-world customer
- *Customer_ID* is acting as an *Account_ID*
- Each account belongs to one bank and one branch

3.Table (Transaction_data):

Transaction_ ID	Customer_ ID	Account_ Type	Total_ Balance	Transaction_ Amount	Investment_ Amount	Investment_ Type	Transaction_ Date
300000	209689	Business	69339	4794	42580	Fixed Deposit	12/8/2024
300001	206124	Business	12825	3500	46605	Fixed Deposit	9/9/2022
300002	207501	Current	67753	2401	17027	Fixed Deposit	4/28/2022
300003	208675	Savings	67061	2952	3054	Fixed Deposit	5/22/2023

Purpose:

Stores transaction-level data along with account-related and investment information.

Row Meaning:

Each row represents:

- One transaction (*Transaction_ID*)
- Transaction details (*Transaction_Amount, Transaction_Date*)
- Account balance after the transaction (*Total_Balance*)
- Account and customer context (*Customer_ID, Account_Type*)
- Investment-related information (*Investment_Amount, Investment_Type*)

Observations:

- The only clearly event-driven attributes are:
  - *Transaction_Amount*
  - *Transaction_Date*
- Other attributes (*Account_Type, Investment_Type, Investment_Amount*) appear to be account-level attributes, not transaction-specific.
- *Total_Balance* changes over time but represents account state, not a transactional measure.

Business Assumption:

- *Account_Type, Investment_Type,* and *Investment_Amount* belong to the account, not the transaction.
- *Total_Balance* represents a snapshot value after each transaction.