Data Dictionary Report

This data dictionary explains each dataset variable in simple business language and highlights how each field reflects real-world retail banking activity. It also identifies how each variable can support trend analysis, opportunity detection, and AI-based forecasting.

1. Data Dictionary Table

Column Name	Description	Potential Business Relevance
customer_id	Unique identifier assigned to each customer	Enables tracking customer behavior over time and identifying loyal or high-value customers
account_no	Unique bank account number	Helps analyze account-level activity such as balances, deposits, and withdrawals
region	Geographic location or branch region of the customer	Identifies regional growth, decline, or opportunity areas
product_type	Type of banking product (Savings, Current, Loan, RD, FD, Credit Card, etc.)	Helps detect which products are gaining or losing popularity
transaction_type	Nature of transaction (deposit, withdrawal, transfer, payment)	Indicates customer spending and saving behavior
transaction_value	Amount involved in each transaction	Helps identify high-value transactions and revenue-generating customers
channel	Medium used for transaction (Branch, ATM, Mobile, Internet)	Reflects digital banking adoption and channel preference trends
timestamp	Date and time when transaction occurred	Enables time-series analysis and forecasting of banking activity
account_balance	Current balance in the customer's account	Indicates financial health and saving patterns
customer_age	Age of the customer	Supports demographic segmentation and targeted product offerings
income_bracket	Income group classification of the customer	Useful for identifying premium vs mass-market opportunities
loan_amount	Total loan amount sanctioned (if applicable)	Helps forecast credit demand and lending growth
interest_rate	Interest rate applied to the product or loan	Allows analysis of rate sensitivity and policy impact
monthly_installment	Regular loan repayment amount	Useful for detecting repayment stress and default risk
credit_score	Measure of customer's creditworthiness	Helps identify low-risk and high-risk customers
customer_tenure	Duration of relationship with the bank	Indicates customer loyalty and churn risk
month	Month extracted from timestamp	Used for monthly trend and seasonality analysis

age_group	Customer age grouped into ranges	Simplifies segment-wise behavior analysis

2. Business Interpretation Summary

- Customer-centric fields (customer_id, age, income, tenure) help understand who the customers are.
- Transaction-related fields (transaction_type, value, channel, timestamp) explain how customers interact with the bank.
- Product and credit fields (product_type, loan_amount, interest_rate) show what products drive revenue.
- Regional and time fields (region, month) support where and when trends emerge.

3. Relevance to AI Forecasting & Opportunity Detection

Use Case	**Supporting Fields**
Trend Forecasting	timestamp, transaction_value, product_type, region
Digital Adoption Analysis	channel, region, customer_age
Credit Demand Forecasting	loan_amount, interest_rate, income_bracket
Customer Segmentation	age_group, income_bracket, account_balance
Opportunity Identification	product_type, channel, transaction_type
Risk Detection	credit_score, monthly_installment, account_balance