Bank Customer Churn Prediction:

A Bank wants to take care of customer retention for its product: savings accounts. The bank wants us to identify customers likely to churn balances below the minimum balance. We have the customer's information such as age, gender, demographics along with their transactions with the bank. Our task as a data scientist would be to predict the propensity to churn for each customer.

Data Dictionary

There are multiple variables in the dataset which can be cleanly divided into 3 categories:

I. Demographic information about customers

- **customer_id** - Customer id
- **vintage** - Vintage of the customer with the bank in a number of days
- **age** - Age of customer
- **gender** - Gender of customer
- **dependents** - Number of dependents
- **occupation** - Occupation of the customer
- **city** - City of the customer (anonymized)

II. Customer Bank Relationship

- **customer_nw_category** - Net worth of customer (3: Low 2: Medium 1: High)
- **branch_code** - Branch Code for a customer account
- **days_since_last_transaction** - No of Days Since Last Credit in Last 1 year

III. Transactional Information

- **current_balance** - Balance as of today
- **previous_month_end_balance** - End of Month Balance of previous month
- **average_monthly_balance_prevQ** - Average monthly balances (AMB) in Previous Quarter
- **average_monthly_balance_prevQ2** - Average monthly balances (AMB) in previous to the previous quarter

- **current_month_credit** - Total Credit Amount current month
- **previous_month_credit** - Total Credit Amount previous month
- **current_month_debit** - Total Debit Amount current month
- **previous_month_debit** - Total Debit Amount previous month
- **current_month_balance** - Average Balance of current month
- **previous_month_balance** - Average Balance of previous month
- **churn** - Average balance of customer falls below minimum balance in the next quarter (1/0)

Dataset: The Dataset of this problem statement is available on the Kaggle Patform.