**1 Identify & Gather Data**

**Customer Demographics Table:**

The Customer Demographics table provides background information on customers, including attributes such as age, gender, marital status, and income level. These demographic factors often correlate with customer retention or churn behaviour. By analysing this data, we can identify segments that are more likely to churn, enabling the bank to design targeted retention strategies and personalized services.

**Transaction History Table:**

This table provides a record of transactions made by each customer. Transaction data serves as an important indicator of customer engagement and financial activity. Patterns such as declining transaction frequency or reduced spending could signal dissatisfaction or disengagement, both of which are closely linked to churn. Aggregating this data allows us to derive key features like total spend, average transaction value, and changes in behaviour over time.

**Customer Service Table:**

Data from this table captures interactions between customers and the bank’s support team. This data is critical, as frequent or unresolved service issues often correlate with a higher likelihood of churn. By analysing interaction frequency, resolution status, and the types of interactions (e.g., complaints vs. inquiries), we can quantify customer satisfaction levels and identify at-risk individuals.

**Online Activity Table:**

This table reflects customers’ engagement with digital services, including metrics like last login date, login frequency, and overall service usage. A decline in digital activity is a common behavioural signal of disengagement. These features serve as strong predictors in churn modelling, helping to detect when a customer is becoming inactive or losing interest.

**Churn Status Table:**

The Churn Status table indicates whether a customer has churned (1) or remained active (0). This is the target variable for our supervised learning model. It allows us to train, validate, and evaluate predictive models that aim to forecast future churn based on the behavioural and demographic characteristics derived from the other tables.