

Background

You are a **Data Analyst** at **Metropolitan National Bank**, a multi-branch retail and commercial banking institution. The bank's executive team is preparing for the **Q3 2024 Quarterly Business Review (QBR)**, and your manager has assigned you to extract critical insights from the bank's customer, account, transaction, and loan databases.

The Chief Risk Officer (CRO) and Chief Financial Officer (CFO) need answers to strategic questions that will help them:

- Understand customer profitability and segmentation
- Assess portfolio risk across different account types
- Identify transaction patterns and potential anomalies
- Evaluate lending performance and customer relationships

Your analysis will directly inform executive decision-making on:

- Branch performance evaluation
- Product pricing strategies
- Risk management policies
- Customer retention programs
- Resource allocation for Q4 2024

Your Mission

Complete the following **4 analytical tasks** within the next hour. Each task requires you to write SQL queries that answer specific business questions. Use the hints provided to guide your query construction.

Database Tables Available:

- `branches` - Branch locations and information
 - `customers` - Customer demographics and credit profiles
 - `accounts` - Customer accounts (Checking, Savings, CD, Money Market)
 - `transactions` - All account transactions
 - `loans` - Customer loans (Auto, Home, Business, Personal, Student)
-

TASK 1: Customer Segmentation & Profitability Analysis

Business Question:

The CFO wants to understand **which customer segments are most profitable** and which branches are serving high-value customers. This will inform Q4 marketing spend and branch resource allocation.

What You Need to Find:

Identify customer segments based on their **total account balances** and **credit scores**, then analyze which branches have the most high-value customers.

Required Analysis:

1. Create a customer segmentation that shows:
 - Customer name and customer number
 - Their home branch (branch name and city)
 - Total balance across ALL their accounts
 - Their credit score
 - A **segment classification** based on total balance:
 - "Premium" for customers with total balance \geq \$100,000
 - "Gold" for customers with total balance \geq \$50,000 but $<$ \$100,000
 - "Silver" for customers with total balance \geq \$10,000 but $<$ \$50,000
 - "Standard" for customers with total balance $<$ \$10,000
2. Filter to show only customers who have:
 - Credit score \geq 650 (creditworthy customers)
 - At least one active account
3. Order results by total balance (highest first)

Hints:

- You'll need to **JOIN** the customers, accounts, and branches tables
 - Use **SUM()** with **GROUP BY** to calculate total balance per customer
 - Use a **CASE statement** to create the segment classification
 - Use **HAVING** clause to filter aggregated results
 - Don't forget to handle customers with multiple accounts at potentially different branches (use the first account's branch as their "home branch")
-

TASK 2: Account Portfolio Risk Assessment

Business Question:

The CRO needs a **risk assessment of the bank's account portfolio** broken down by account type and branch. This will help identify which branches may need additional risk management oversight and which products are showing concerning patterns.

What You Need to Find:

Analyze account balances by type and identify accounts that may require attention due to low balances or other risk factors.

Required Analysis:

1. Create a comprehensive account portfolio view that shows:
 - Branch name and city
 - Account type
 - Number of accounts of that type
 - Total balance for that account type
 - Average balance for that account type
 - Minimum balance for that account type
 - Maximum balance for that account type
 - Count of "at-risk" accounts (accounts with balance < \$5,000 for Checking/Savings, or < \$50,000 for CD/Money Market)
2. Filter to show only:
 - Account types with **at least 3 accounts**
 - Branches that have **total balances ≥ \$100,000** for that account type
3. Order by branch name, then account type

Hints:

- JOIN accounts and branches tables
- Use **COUNT()**, **SUM()**, **AVG()**, **MIN()**, **MAX()** aggregate functions
- Use **CASE statement with SUM()** to count at-risk accounts (this is a common pattern: `SUM(CASE WHEN condition THEN 1 ELSE 0 END)`)
- Use **GROUP BY** with multiple columns (branch and account type)
- Use **HAVING** to filter groups that meet the criteria
- Remember: aggregate functions cannot be used in WHERE clause, use HAVING instead

TASK 3: Transaction Pattern Analysis

Business Question:

The Operations team has noticed some **unusual transaction patterns** in August 2024 and wants to identify high-velocity transaction accounts that may require additional monitoring for fraud prevention or customer service follow-up.

What You Need to Find:

Identify accounts with high transaction volumes in August 2024, analyze their transaction patterns, and extract time-based insights.

Required Analysis:

1. For all transactions in **August 2024**, create a report showing:
 - Account number
 - Customer name (first and last name combined)
 - Total number of transactions
 - Total deposit amount (sum of all Deposit transactions)
 - Total withdrawal amount (sum of all Withdrawal and ATM Withdrawal transactions - these should be positive numbers in your output)
 - Net transaction amount (deposits minus withdrawals)
 - Number of weekend transactions (Saturday or Sunday)
 - Number of late-night transactions (between 10 PM and 6 AM)
2. Filter to show only accounts with:
 - **At least 5 transactions** in August 2024
 - **At least one weekend OR late-night transaction**
3. Order by total number of transactions (highest first)

Hints:

- Use **EXTRACT()** or date functions to filter for August 2024
 - Use **EXTRACT(DOW FROM timestamp)** to identify weekends (0 = Sunday, 6 = Saturday)
 - Use **EXTRACT(HOUR FROM timestamp)** to identify late-night transactions
 - Use **CASE statements with SUM()** for conditional aggregation (separate deposits from withdrawals)
 - For withdrawals, you may need to use **ABS()** or multiply by -1 to show positive amounts
 - Use **CONCAT()** or **||** to combine first and last names
 - JOIN transactions, accounts, and customers tables
 - Use **HAVING** to filter for accounts with ≥ 5 transactions
-

TASK 4: Comprehensive Customer Relationship Value

Business Question:

The Relationship Management team wants to identify "**Total Relationship Value**" for customers who have both deposits (accounts) and loans with the bank. This metric helps identify the bank's most important customers and prioritize retention efforts.

What You Need to Find:

Calculate the total relationship value for customers, combining their deposit balances with their loan balances, and identify their primary relationship branch.

Required Analysis:

1. Create a comprehensive customer view that shows:
 - Customer number and full name
 - Primary branch name (from their accounts)
 - Total deposit balance (sum of all account balances)
 - Number of accounts they have
 - Total loan amount (sum of original loan amounts)
 - Total outstanding loan balance (sum of remaining balances)
 - Number of loans they have
 - **Total Relationship Value** = Total deposit balance + Total outstanding loan balance
 - Credit score
 - Annual income
 - **Customer tenure** in years (based on when their first account was opened, relative to August 1, 2024)
2. Include **ONLY customers who have BOTH accounts AND loans**
3. Order by Total Relationship Value (highest first)
4. Show top 20 customers only

Hints:

- This requires JOINing **customers, accounts, AND loans** tables
- Use **INNER JOIN** for accounts and loans to ensure customers have both
- Use **GROUP BY** customer information
- Use **MIN(date_opened)** to find the first account opening date
- Use **AGE()** or **EXTRACT(YEAR FROM AGE())** to calculate tenure
- You'll need to calculate multiple SUMs and COUNTs
- Use **COALESCE()** if you want to handle potential NULL values gracefully (though not needed with INNER JOINS)

Real-World Application

These exact types of analyses are performed daily in investment banking for:

- **Customer Segmentation:** Identifying high-value clients for private banking services
 - **Risk Management:** Monitoring account and loan portfolios for risk exposure
 - **Fraud Detection:** Identifying unusual transaction patterns
 - **Relationship Management:** Calculating customer lifetime value and prioritizing retention
 - **Regulatory Reporting:** Providing data for compliance and regulatory requirements
 - **Strategic Planning:** Informing executive decision-making with data-driven insights
-

Good luck! Remember: Start simple, test your queries incrementally, and use the hints to guide you.