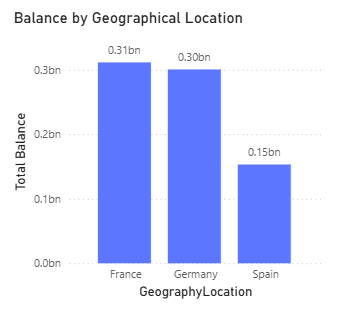
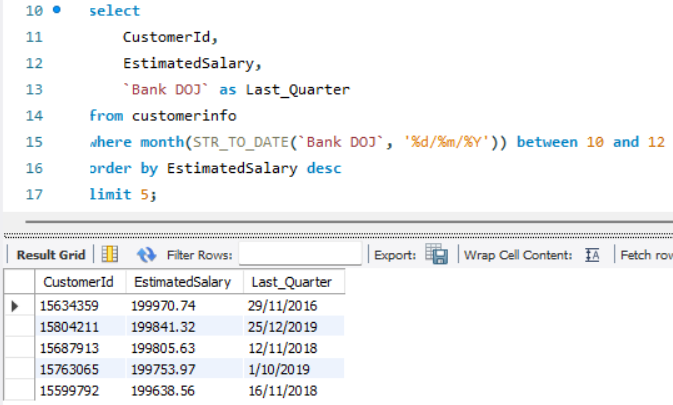
**Objective Questions:**

1. What is the distribution of account balances across different regions?

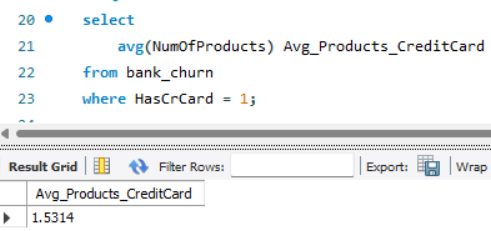
Ans: Balances across different regions are-

****

1. Identify the top 5 customers with the highest Estimated Salary in the last quarter of the year. (SQL)

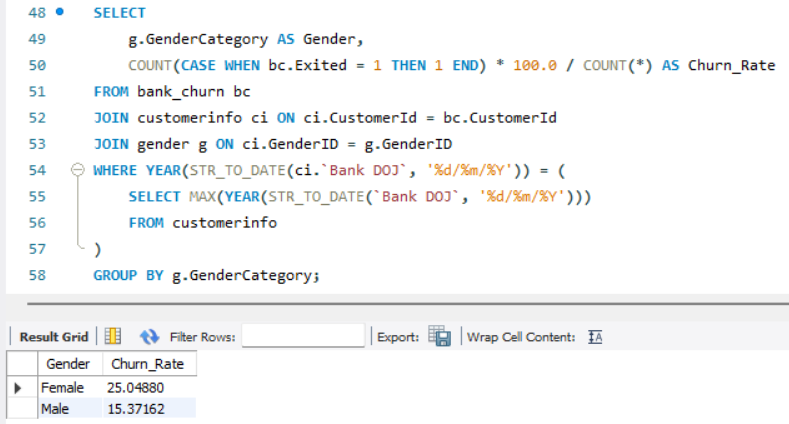
Ans: The top 5 customers with the highest estimated salary in the last quarter of the year are-  
  


1. Calculate the average number of products used by customers who have a credit card. (SQL)

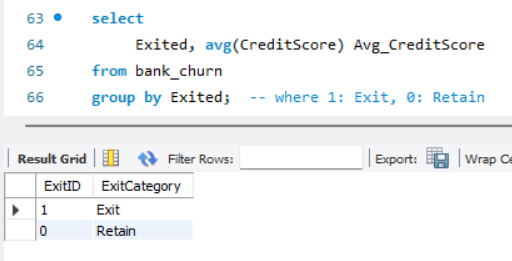
Ans:   
  
 

1. Determine the churn rate by gender for the most recent year in the dataset.

Ans:



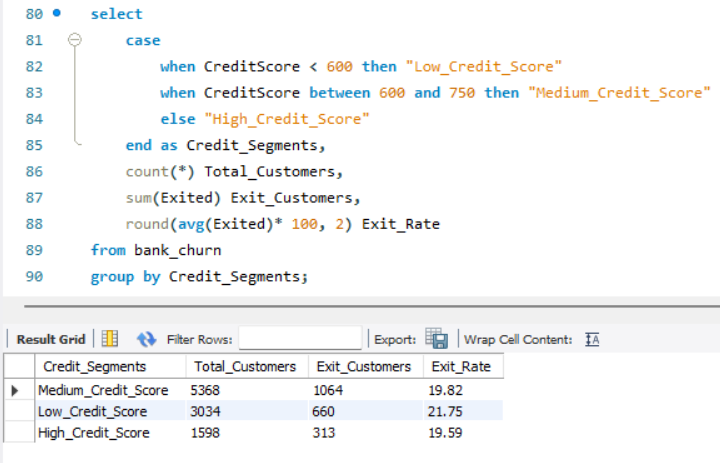
1. Compare the average credit score of customers who have exited and those who remain. (SQL)

Ans:   


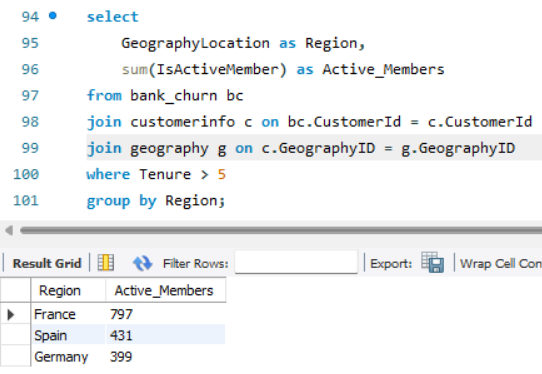
1. Which gender has a higher average estimated salary, and how does it relate to the number of active accounts? (SQL)

Ans:   
  
 

1. Segment the customers based on their credit score and identify the segment with the highest exit rate. (SQL)

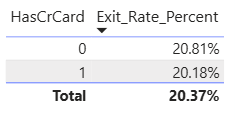
Ans:   
  


1. Find out which geographic region has the highest number of active customers with a tenure greater than 5 years. (SQL)

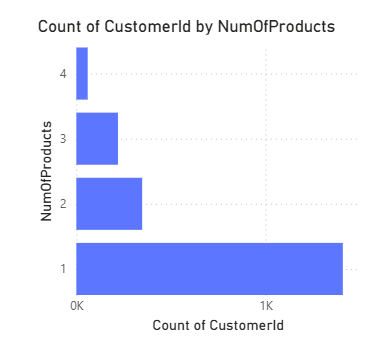
Ans:   
 

1. What is the impact of having a credit card on customer churn, based on the available data?

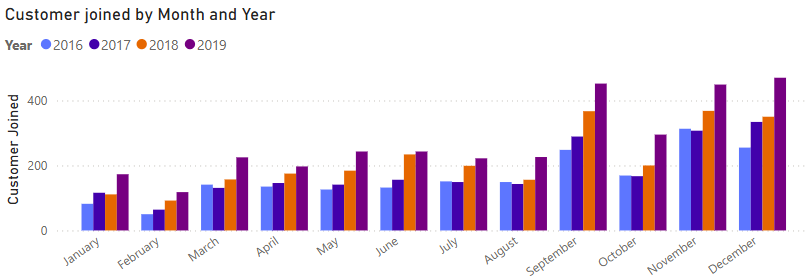
Ans: The Formula used to find Exit\_Rate\_Percent is:

**Exit\_Rate\_Percent = AVERAGE ('Bank\_Churn xlsx - Bank\_Churn'[Exited])**  
  
Change the format to percentage and set decimal to 2 places.  
   
 

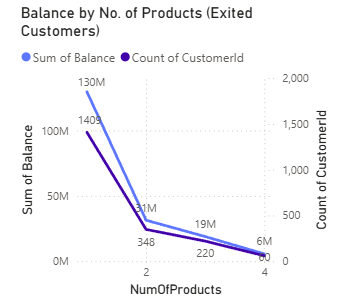
1. For customers who have exited, what is the most common number of products they have used?

Ans: According to the chart, **product 1** is the most common among existing customers.  
  
 

1. Examine the trend of customers joining over time and identify any seasonal patterns (yearly or monthly). Prepare the data through SQL and then visualize it.

Ans:   
  
  
  


1. Analyze the relationship between the number of products and the account balance for customers who have exited.

Ans:   
 

So, based on this visual, we can observe that-  
1. **Customers with only 1 product** account for the **highest churn** and hold the **highest total balance**, making them **financially valuable but weakly engaged**.

2. As the **number of products increases**, both churn count and total balance among exited customers **drop sharply**.

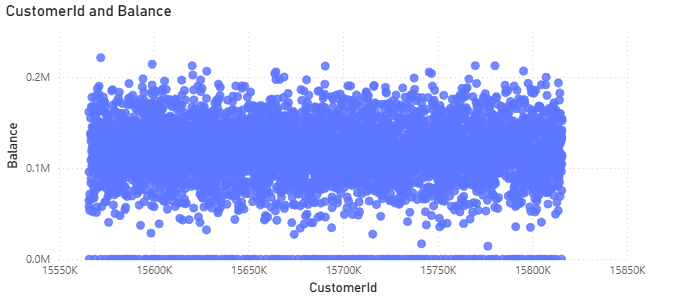
3. Customers with **4 products show zero churn**, highlighting a **strong link between product engagement and retention**.

**Recommendations**

1. **Cross-sell to 1-product customers** through personalized offers and bundled services.
2. **Target high-balance, low-engagement customers** with retention-focused outreach.
3. **Introduce a Product Engagement Score** to identify and act on at-risk profiles.
4. **Enhance communication and value proposition** of additional banking products.

**Conclusion:** Boosting product engagement is crucial to reduce churn and retain high-value customers.

1. Identify any potential outliers in terms of balance among customers who have remained with the bank.

Ans:   
 

**Observations:**

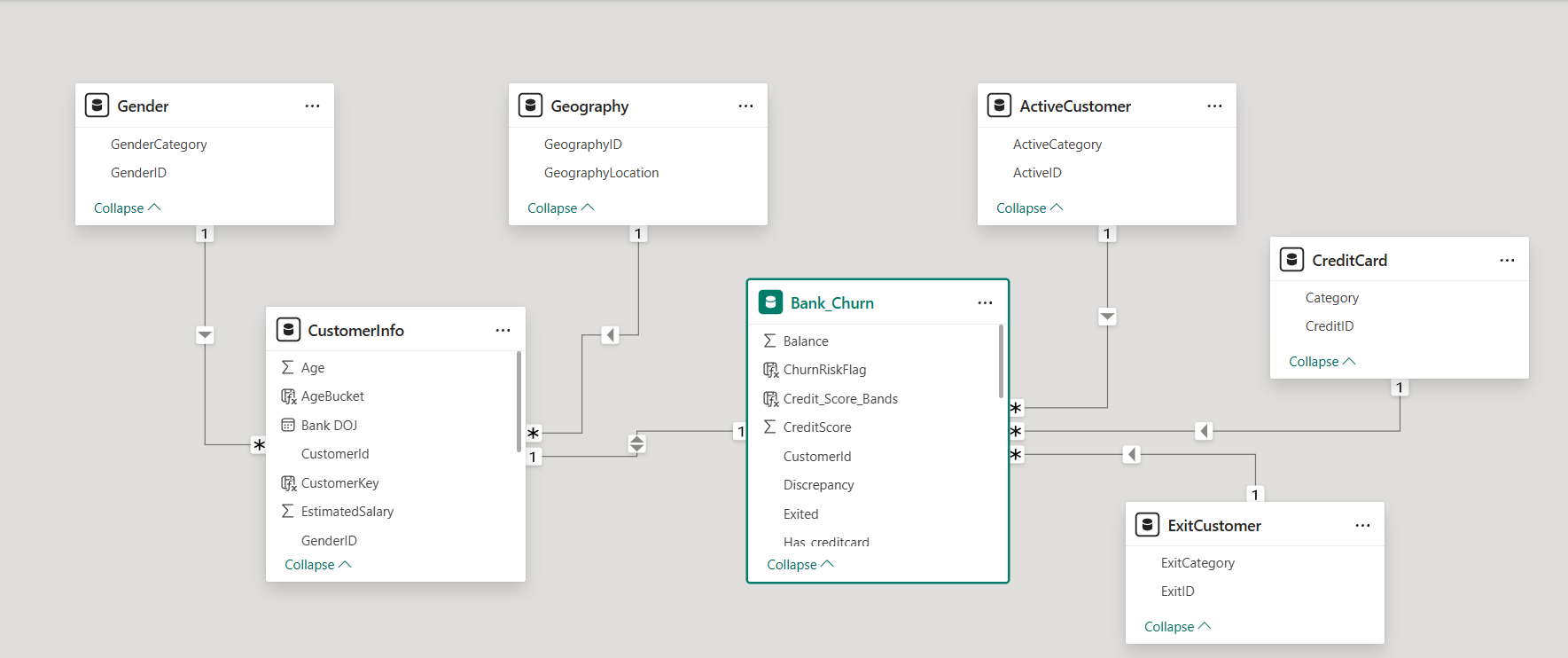
1. **Majority Cluster**
   * Most customers have balances between **₹60K–₹180K**, forming a **dense horizontal band**.
   * This indicates a fairly consistent financial profile among retained customers.
2. **Lower-End Outliers**
   * A noticeable number of customers have a **balance close to ₹0**, yet they haven’t churned.
   * These accounts might be inactive or used for specific low-volume purposes.
3. **Upper-End Outliers**
   * A few dots lie **above ₹200K**, marking **high-value accounts** that are retained.
   * These are likely **priority customers** and may need **dedicated relationship management**.

**Recommendations:**

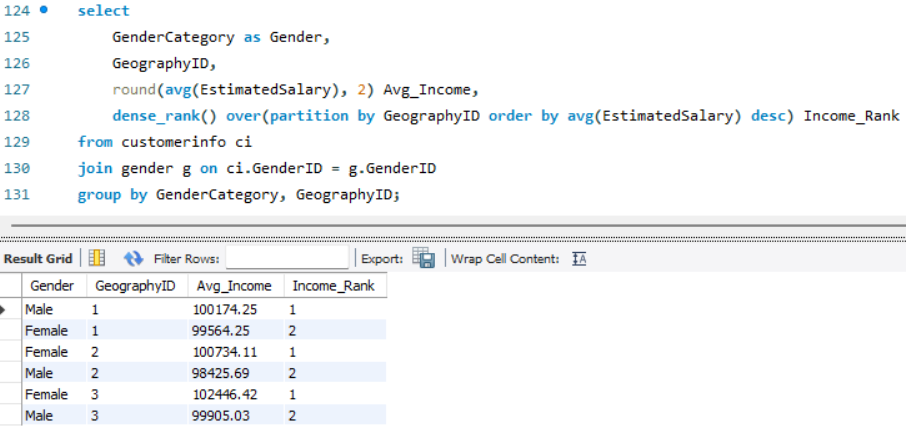
1. **Investigate Zero-Balance Retentions**
   * Understand why zero-balance accounts are still active — these could be candidates for **cross-selling or closure**.
2. **Reward High-Balance Loyalty**
   * Offer **exclusive benefits or upgrades** to customers with unusually high balances to **ensure retention and increase satisfaction**.
3. **Segmentation-Based Follow-Up**
   * Use this balance distribution to **segment customers** into:
     + High Net Worth (₹ 200 K+)
     + Mid-range Active (₹60K–₹180K)
     + At-risk Low Value (<₹10K)
4. How many different tables are given in the dataset? Out of these tables, which table consists only of categorical variables?

Ans: There are seven tables, of which five consist solely of categorical variables. They are-

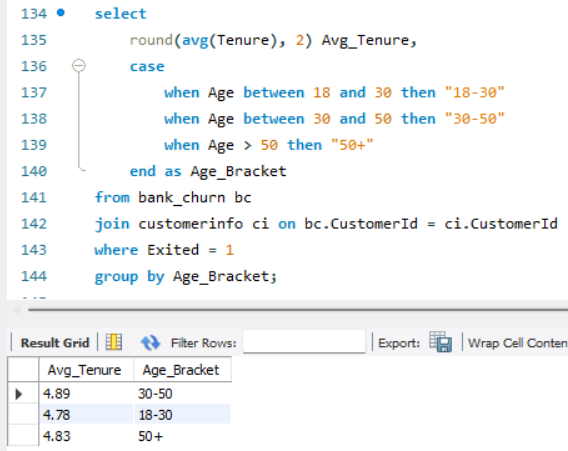
* Gender
* Geography
* ActiveCustomer
* CreditCard
* ExitCustomer



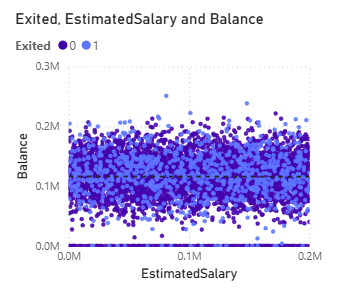
1. Using SQL, write a query to find out the gender-wise average income of males and females in each geography id. Also, rank the gender according to the average value. (SQL)

Ans:   
  


1. Using SQL, write a query to find out the average tenure of the people who have exited in each age bracket (18-30, 30-50, 50+).

Ans:   
 

1. Is there any direct correlation between salary and the balance of the customers? And is it different for people who have exited or not?

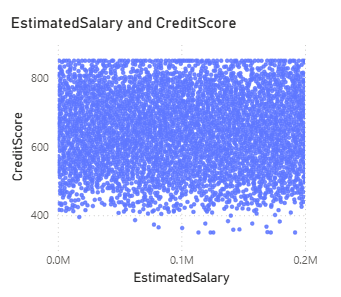
Ans:   
   
  
So, through this analysis, we can observe that-  
  
**1. No Strong Visible Correlation:**

* The dots are **evenly spread across the salary range**, regardless of balance.
* This suggests **no clear linear relationship** between EstimatedSalary and Balance.

**2. Churned vs. Not Churned:**

* Both exited (purple) and retained (blue) customers are **similarly distributed**.
* There's **no major separation or cluster** that would indicate salary or balance is a strong standalone driver of churn.

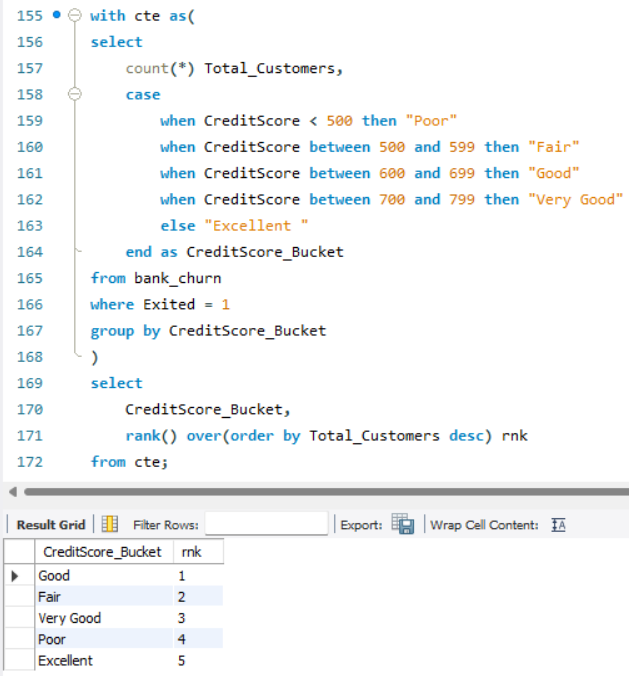
1. Is there any correlation between the salary and the Credit score of customers?

Ans:   
 

### **Correlation Between Estimated Salary and Credit Score-**

* **Observation:** Scatter plot shows credit scores are widely distributed across all salary levels.
* **Insight:** No visible pattern or trend between EstimatedSalary and CreditScore.
* **Conclusion:**  
  There is **no significant correlation** between a customer's salary and credit score.  
  ➤ This means income does **not directly influence** creditworthiness in your dataset.

1. Rank each bucket of credit score as per the number of customers who have churned the bank.

Ans:   
  


1. According to the age buckets find the number of customers who have a credit card. Also retrieve those buckets that have lesser than average number of credit cards per   
   bucket.

Ans:



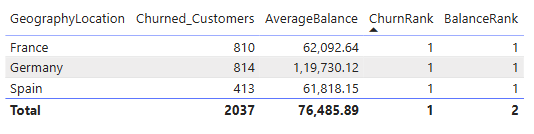
1. Rank the Locations as per the number of people who have churned the bank and the average balance of the customers.

Ans: We have calculated these in Power BI using the DAX formula as  
  
Step 1: Create Two Measures  
  
**Count of churned customers:**ChurnedCustomers = CALCULATE (COUNTROWS (bank\_churn), bank\_churn[Exited] = 1

**Average balance:**AvgBalance = AVERAGE (bank\_churn[Balance])  
  
Step 2: Create Rank Measures

**Rank by churn:**ChurnRank = RANKX (ALL (customerinfo[LocationID]), [ChurnedCustomers], , DESC)

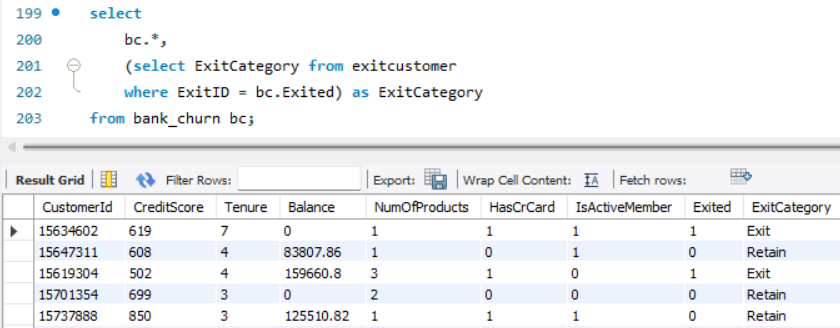
**Rank by balance:**BalanceRank = RANKX (ALL (customerinfo[LocationID]), [AvgBalance], , DESC)



1. As we can see that the “CustomerInfo” table has the CustomerID and Surname, now if we have to join it with a table where the primary key is also a combination of CustomerID and Surname, come up with a column where the format is “CustomerID\_Surname”.

Ans: To achieve this in Power BI, we’ll create a column using DAX-  
  
 CustomerKey = CustomerInfo[CustomerId] & "\_" & CustomerInfo[Surname]

1. Without using “Join”, can we get the “ExitCategory” from ExitCustomers table to the Bank\_Churn table? If yes do this using SQL.

Ans:   
  


1. Were there any missing values in the data, using which tool did you replace them and what are the ways to handle them?

Ans: No missing values were found in the data. If any had been present, we could have managed them using Power BI's Power Query features.

* The “Column Quality” and “Column Distribution” tools were utilised to identify any missing or null values.
* Various transformations, such as "Replace Values," "Remove Nulls," and "Fill Down," were applied depending on the context of each column.

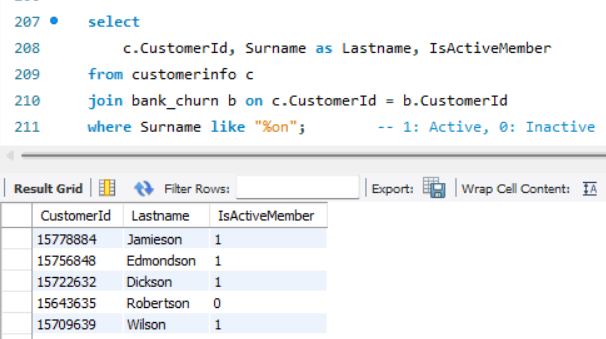
**Handling Missing Data:**

1. Numerical Columns (e.g., Balance, Credit Score):   
 - Missing values were replaced with the median to avoid skew caused by outliers.

2. Categorical Columns (e.g., Gender, Geography):  
 - Missing values were filled in with the most frequent category (mode) or labelled as "Unknown."

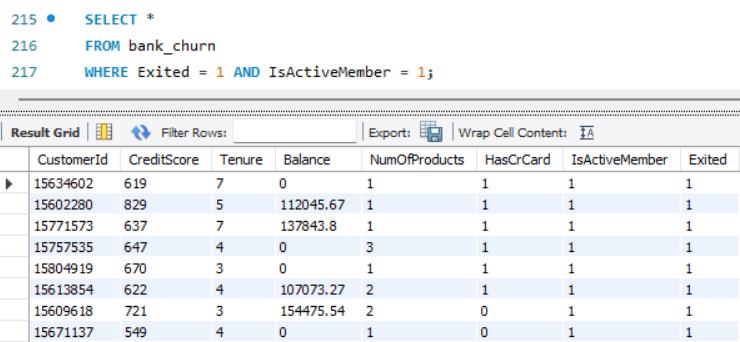
3. Date Columns (e.g., Bank Date of Joining):  
 - Rows with missing critical dates were excluded to ensure the accuracy of the analysis.

1. Write the query to get the customer IDs, their last name, and whether they are active or not for the customers whose surname ends with “on”.

Ans:   
  


1. Can you observe any data discrepancy in the Customer’s data? As a hint, it’s present in the IsActiveMember and Exited columns. One more point to consider is that the data in the Exited Column is absolutely correct and accurate.

Ans: So, according to the question Exited Column is accurate; therefore, if a customer has exited the bank (Exited = 1), then they should not be an active member (IsActiveMember should be 0).

Because once a customer leaves (churns), they are no longer active.  
  
  
  
Hence, by analysing the above query, we can see that there is a **potential data discrepancy** in the IsActiveMember column when compared to Exited. Since churned customers (Exited = 1) logically shouldn’t be active, those mismatched records reflect **bad data**.

**Subjective Question:**

1. Customer Behaviour Analysis: What patterns can be observed in the spending habits of long-term customers compared to new customers, and what might these patterns suggest about customer loyalty?

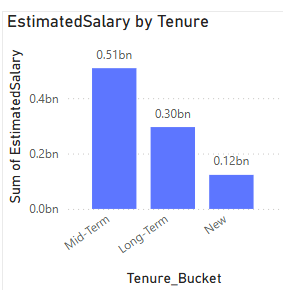
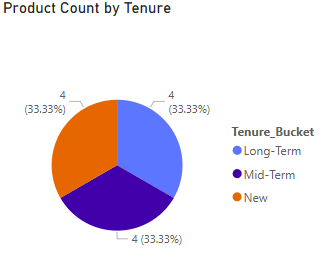
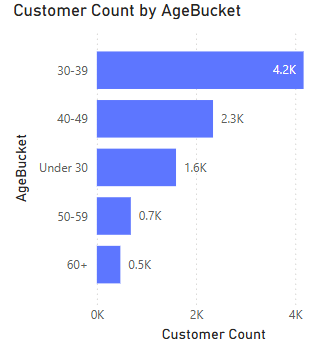
Ans: **Approach**

* **Data Segmentation**: Customers are categorised into three tenure buckets – New, Mid-Term, and Long-Term.

**Formula**-  
Tenure\_Bucket = SWITCH (TRUE (), Bank\_Churn [Tenure] <= 3, "New”, Bank\_Churn [Tenure] <= 5, "Mid-Term", "Long-Term")

* **Metrics Analysed**:
  + **Average Balance**
  + **Estimated Salary by Tenure**
  + **Product Count by Tenure**
  + **Customer Distribution by Age Group**

**Insights**

1. **Average Balance**:
   * Remains consistent across tenure categories:
     + New: **76K**, Mid-Term: **75K**, Long-Term: **76K**
   * Suggests that the average account balance does not drastically change with tenure.
2. **Estimated Salary by Tenure**:
   * Mid-Term customers have the highest total estimated salary (**0.51bn**).
   * Long-Term customers: **0.30bn**
   * New customers: **0.12bn**  
      
   * High-earning individuals may not transition into long-term customers.
   * Possible factors: career shifts, retirement, attrition due to dissatisfaction.
3. **Product Count by Tenure**:
   * Even distribution: All tenure buckets have **4 products**, each representing **33.33%**.
   * Product engagement is similar regardless of tenure, which could mean consistent cross-sell efforts.
4. **Customer Count by Age Bucket**:
   * The majority of customers fall in the **30–39** age group (**4.2K**) & **40–49** (**2.3K**).
   * Fewer older customers (60+ only **0.5K**) suggest potential churn or lower digital engagement in older demographics.  
       
       

**Conclusion**

* **Loyalty Reflected in Stability:**  
  Long-term customers maintain similar balances and product usage as new ones, suggesting loyalty is driven more by consistency than by expanding product adoption.
* **Potential Retention Gap Among High Earners:**  
  The decline in estimated salary for long-term users may indicate that high-income customers are less likely to stay engaged over time.
* **Mid-Term Segment Holds Strategic Value:**  
  Mid-term customers contribute the highest estimated salary. This group presents a strong opportunity for long-term retention efforts.
* **Consistent Cross-Sell, Limited Growth:**  
  Product usage remains consistent across tenure groups, showing effective cross-sell strategies but also suggesting limited upsell or expansion potential.
* **Recommended Actions:**  
  Prioritise engagement strategies for mid-term, high-income customers to improve retention. Explore personalised offers and loyalty initiatives to retain older and long-tenured segments.

1. Product Affinity Study: Which bank products or services are most commonly used together, and how might this influence cross-selling strategies?

Ans: **Approach** To identify product affinity patterns and inform cross-selling strategies, we analysed customer data using the following steps:

1. Segmented Customers by Number of Products Used (1 to 4).
2. Studied product combinations, especially those involving the Credit Card (CC).
3. Analysed churn behaviour across different product combinations.
4. Introduced tenure-based segmentation (New, Mid-Term, Long-Term) to understand when customers churn after product adoption.
5. Evaluated product usage with respect to balance, salary, and credit card holding behaviour.

**Insights  
  
**

#### 1. **Most Common Product Combinations**

* The most frequent product usage combinations are:
  + **Product 1 only**
  + **Product 1 + Credit Card**
  + **Product 2 + Credit Card**

#### 2. **Highest Churn Combo: Product 1 + Credit Card**

* **676 churned customers** used only **Product 1 + Credit Card**, making it the **most churn-prone combination**.
* Indicates **insufficient product depth** or value to retain customers.

#### 3. **2 Products = Sweet Spot for Retention**

* Customers using **2 products** churn **less than those with only 1,** representing a strong user base.
* These customers have **higher account activity**, with balanced Credit Card holding distribution.

#### 4. **Tenure Segmentation Adds Critical Timing Insight**

* While churn rates are **similar across all tenure groups (~14%)**, the **Mid-Term group** (2–5 years) has the **highest churn volume**.
* **New customers** have **lower balances and still churn**, signalling early-stage disengagement.

5. **High-Balance Customers Are Still Leaving Early**

* Many churners have significant balances (₹90K+), especially among Mid-Term users.
* This shows that **having money in the account doesn’t always prevent churn**—product relevance does.  
    
  **Conclusions & Cross-Selling Strategies**

1. **Avoid single-product engagements**, especially **Product 1 + CC**. Customers using only basic services are most at risk of leaving.
   * **Strategy:** Promote bundles that include **loan products, savings schemes, or investment products** early.
2. **Target Mid-Term customers (2–5 years)** with personalised cross-sell offers.
   * This group is **high-value** but also **most prone to churn**.
3. **Upsell from 1 to 2+ products quickly**, ideally within the **first year** of engagement.
   * Improves stickiness and reduces early churn seen in new customers.
4. Use **tenure-aware marketing**:
   * For **New users**, focus on **onboarding journeys** and early nudges for multi-product use.
   * For **Mid-Term**, use **loyalty offers, financial planning tools, or bundled benefits** to deepen engagement.
5. Geographic Market Trends: How do economic indicators in different geographic regions correlate with the number of active accounts and customer churn rates?

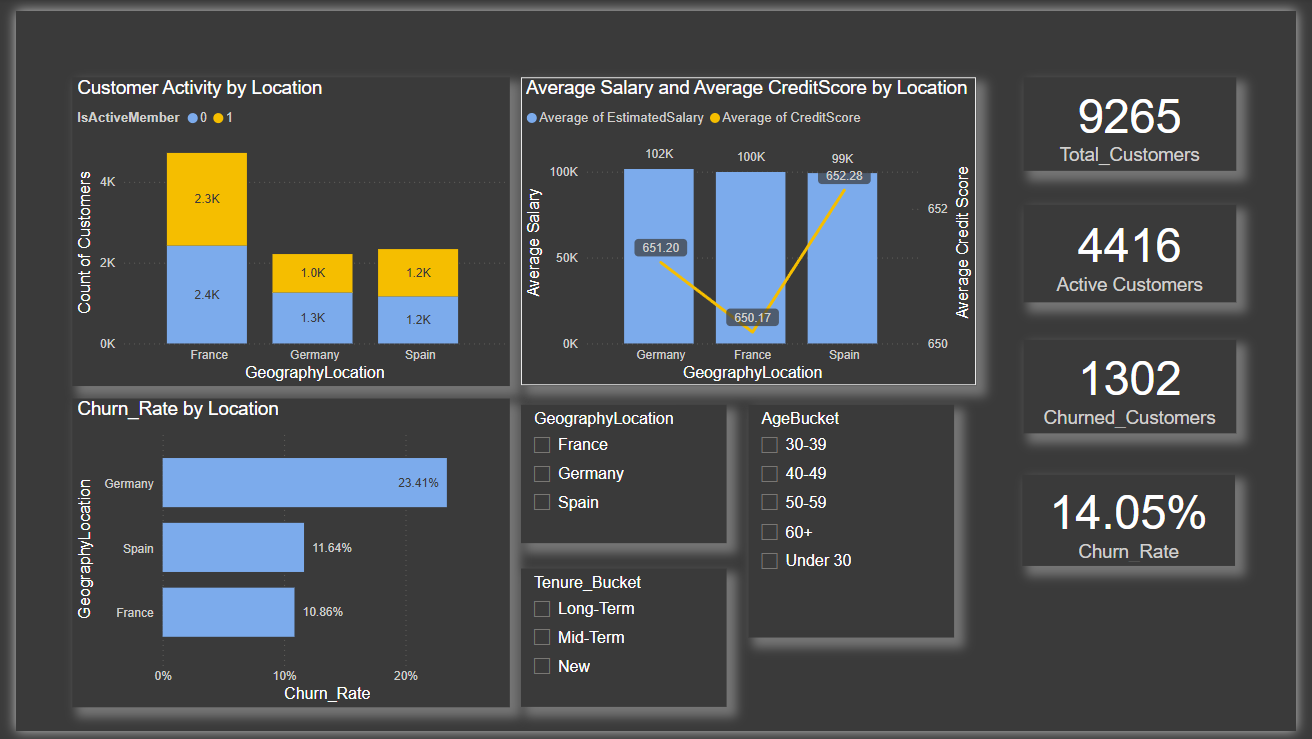
### Ans: **Approach**

I explored customer data across France, Germany, and Spain to understand how different regions are performing. The focus was on comparing active vs. churned customers and seeing how economic indicators like average salary and credit score relate to churn rates.

I built a Power BI dashboard showing customer activity, churn rates, and economic indicators by location. Slicers for age and tenure allow for interactive exploration of trends based on different segments.

**Insights**

* **France** has the biggest customer base, with over 4,700 customers, and a good portion of them are still active. It also has the lowest churn rate at around 10.8%.
* **Germany** stands out—but not in a good way. Even though the average salary there is the highest (around 102K), Germany has the **highest churn rate** at 23.4%. That’s a red flag and suggests people are leaving despite earning more.
* **Spain** does better than expected. It has the **highest average credit score** and a relatively low churn rate (~11.6%), even though the salary is the lowest of the three.
* So, it’s not just about how much customers earn. Credit score seems to have a stronger connection with churn than salary does.



**Conclusion**

Even though Germany has the highest-earning customers, it’s losing them faster than any other region. This tells us that salary alone doesn’t keep customers loyal—maybe they’re not happy with the service, or they don’t find enough value in the products.

On the other hand, France and Spain are doing a better job of keeping their customers, even with slightly lower credit scores or salaries.

To understand what’s going on in Germany, we should dig deeper—maybe look at product usage, complaints, or tenure trends. Something’s not clicking with customers there.

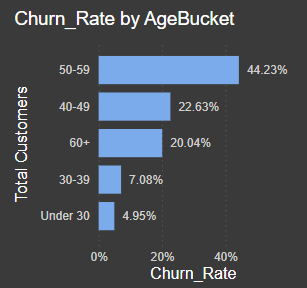
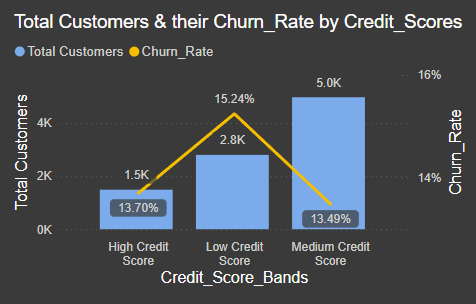
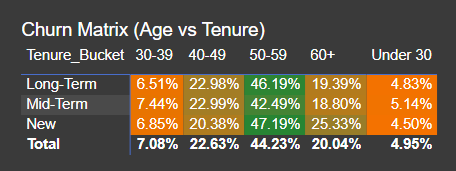
1. Risk Management Assessment: Based on customer profiles, which demographic segments appear to pose the highest financial risk to the bank, and why?

Ans: **Approach**We explored the **churn behaviour across multiple customer dimensions** — including **age group**, **tenure bucket**, and **credit score bands** — to assess which segments might pose a **higher financial risk** to the bank. We used:

* **Churn Rate by AgeBucket**
* **Churn Rate by Tenure\_Bucket**
* **Credit Score vs Salary scatter plot**
* **Churn Matrix (Age vs Tenure)**
* **Credit Score Band Analysis**

**Insights**

1. **Highest Risk Age Segment: 50–59 Age Group**
   * This group has a **churn rate of 44.23%**, the **highest across all age groups**.
   * It remains consistently high across all tenure categories (peaks at **47.19%** for New customers).
   * Possibly nearing retirement or shifting financial priorities.
2. **Credit Score Risk**
   * **Low Credit Score customers** have the **highest churn rate at 15.24%**.
   * Even though the medium and high score groups have more customers, they show **lower churn rates** (~13.5%).
   * Indicates that customers with **lower creditworthiness are less stable**.
3. **Tenure Influence**
   * Surprisingly, **New, Mid-Term, and Long-Term customers** have **very similar churn rates (~14%)** overall.
   * But, within the **50–59 age group**, all tenure segments still churn heavily, indicating **age is a stronger churn driver than tenure** here.
4. **Salary vs Credit Score Scatter (\*)**
   * Across locations, **salary does not significantly impact churn**, but **lower credit scores** appear more frequent in churn-heavy zones.
5. **Age + Tenure Risk Matrix**
   * The **50–59 + New** and **50–59 + Mid-Term** combinations show churn rates above **46%**, highlighting this intersection as especially risky.
   * The **40–49 Mid-Term** group is also elevated at ~23%.

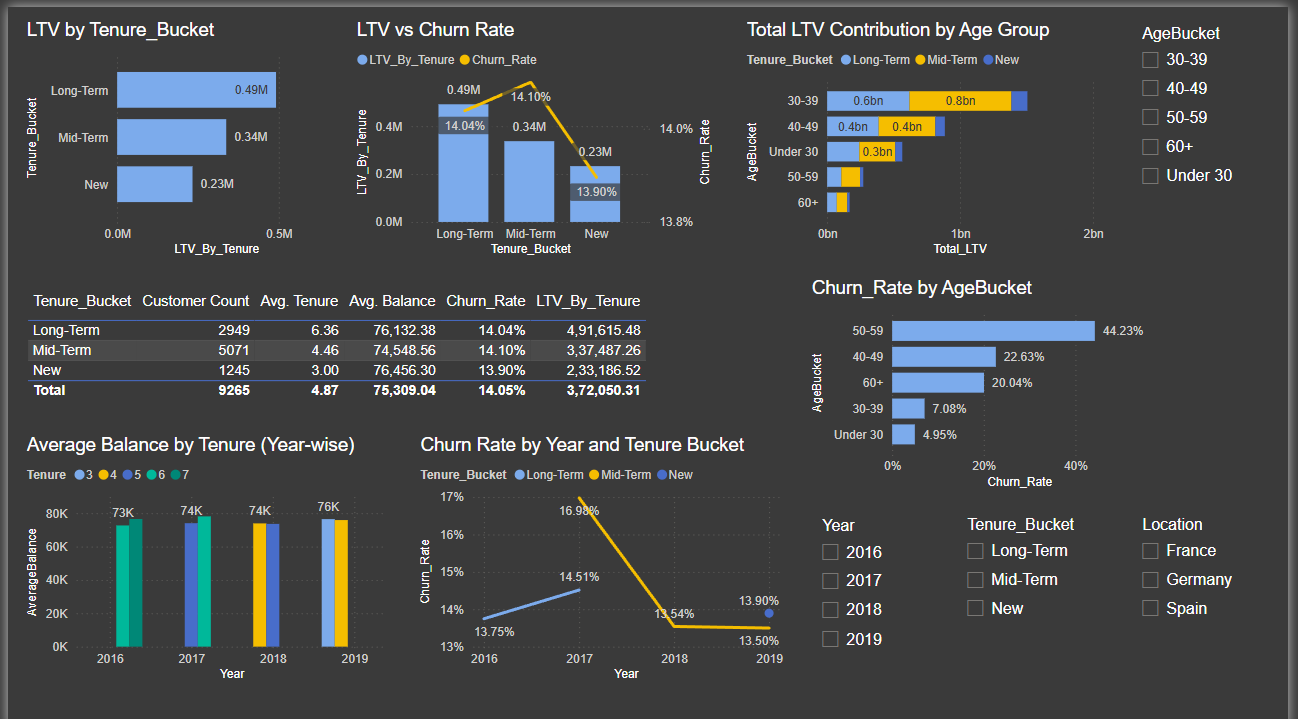
**Recommendations**

1. **Flag and monitor the 50–59 Age Group closely**
   * Particularly those with **low credit scores and short tenure**.
   * Consider proactive engagement, loyalty incentives, or financial planning offers.
2. **Credit Risk-Based Segmentation**
   * Introduce targeted retention efforts for **low-credit-score** customers, such as credit counselling or bonus interest schemes for improved scores.
3. **Refine Risk Models**
   * Incorporate **age group** and **credit score bands** as stronger weights in churn or risk prediction models.
4. **Onboarding & Early Engagement**
   * Despite tenure having a flatter churn impact, **new customers aged 40+** show sharp churn spikes, suggesting onboarding and early engagement could be key.
5. Customer Tenure Value Forecast: How would you use the available data to model and predict the lifetime (tenure) value in the bank of different customer segments?

Ans: **Approach**

To model and forecast Customer Lifetime Value (LTV) across different customer segments, we analyse key influencing factors such as **average tenure, average balance, churn rate, and age group distribution**. By segmenting customers into tenure buckets—**New, Mid-Term, and Long-Term**—and observing their behaviour over time, we simulate LTV forecasts using a rule-based approach derived from historical trends. We also adjust LTV by factoring in churn to understand the true value of retaining customers in each segment.

**Insights**



* **Long-term customers** have the **highest average LTV (₹4.91L)** but also face a churn rate of **14.04%**, almost equal to others.
* **Mid-Term customers** contribute the **largest share of total LTV (₹1.68B)**, despite a slightly higher churn rate (14.10%) and lower individual LTV (₹ 3.37 L).
* **New customers** show the **lowest churn (13.90%)** and are growing steadily in average balance, but their current contribution to overall LTV is minimal (₹0.23L per customer).
* The **age group 30–39 years** contributes the **highest total LTV**, indicating strong value potential from younger-mid professionals.
* **Churn Rate by Age** shows a sharp rise in the **50–59 age group (44.23%)**, suggesting a risk of value erosion in that segment.

**Recommendations**

* **Retain Mid-Term customers** by offering tailored engagement strategies—they have high growth potential and dominate the LTV contribution.
* **Focus on converting New customers to Mid-Term** through early engagement campaigns, as they show the lowest churn and can be nurtured into valuable long-term clients.
* **Address churn risk in the 50–59 age group** with targeted retention initiatives—this segment has high value, but the highest churn rate.
* Consider using **Churn-adjusted LTV** as a more reliable measure of customer value when forecasting future trends or prioritising retention budgets.

1. Marketing Campaign Effectiveness: How could you assess the impact of marketing campaigns on customer retention and acquisition within the dataset? What extra information would you need to solve this?

Ans: **Approach**

To assess the impact of marketing campaigns on **customer retention and acquisition**, we would compare **pre- and post-campaign performance metrics**, such as customer count, churn rate, average balance, and LTV, across different customer segments and periods. This requires linking campaign exposure to specific customers or cohorts, then evaluating how those exposed to campaigns performed relative to those who weren’t (control vs treatment).

**Insight**

Based on the current dataset:

* We can track **churn trends over time**, LTV by tenure, and **customer count changes**, but we **cannot isolate the impact of any marketing effort**, since there’s no campaign attribution data.
* Some positive trends, like increasing balances among New and Mid-Term customers, may hint at engagement success, but we can’t directly link this to marketing without more data.

**Recommendation**

To effectively evaluate marketing campaign performance, the following additional data is needed:

**Essential Additional Data:**

1. **Campaign Details:**
   * Campaign ID, type (email, SMS, ad), start/end dates, target audience.
2. **Customer Exposure:**
   * Which customers were exposed to which campaigns (Campaign-Customer join table).
3. **Customer Acquisition Source:**
   * Whether a new customer was acquired through a campaign or organically.
4. **Campaign Cost (optional but helpful):**
   * To perform ROI and cost-per-retention or cost-per-acquisition calculations.

**Suggested Metrics to Measure Impact:**

* **Retention Rate Change:** Before vs after campaign for exposed vs non-exposed groups.
* **Churn Rate Drop:** Especially in high-risk segments like age 50–59.
* **LTV Lift:** Did the LTV of campaign-engaged customers improve?
* **New Customers Added:** In campaign periods vs non-campaign periods.
* **Customer Behaviour Post-Campaign:** Change in average balance, product usage.

1. Customer Exit Reasons Exploration: Can you identify common characteristics or trends among customers who have exited that could explain their reasons for leaving?

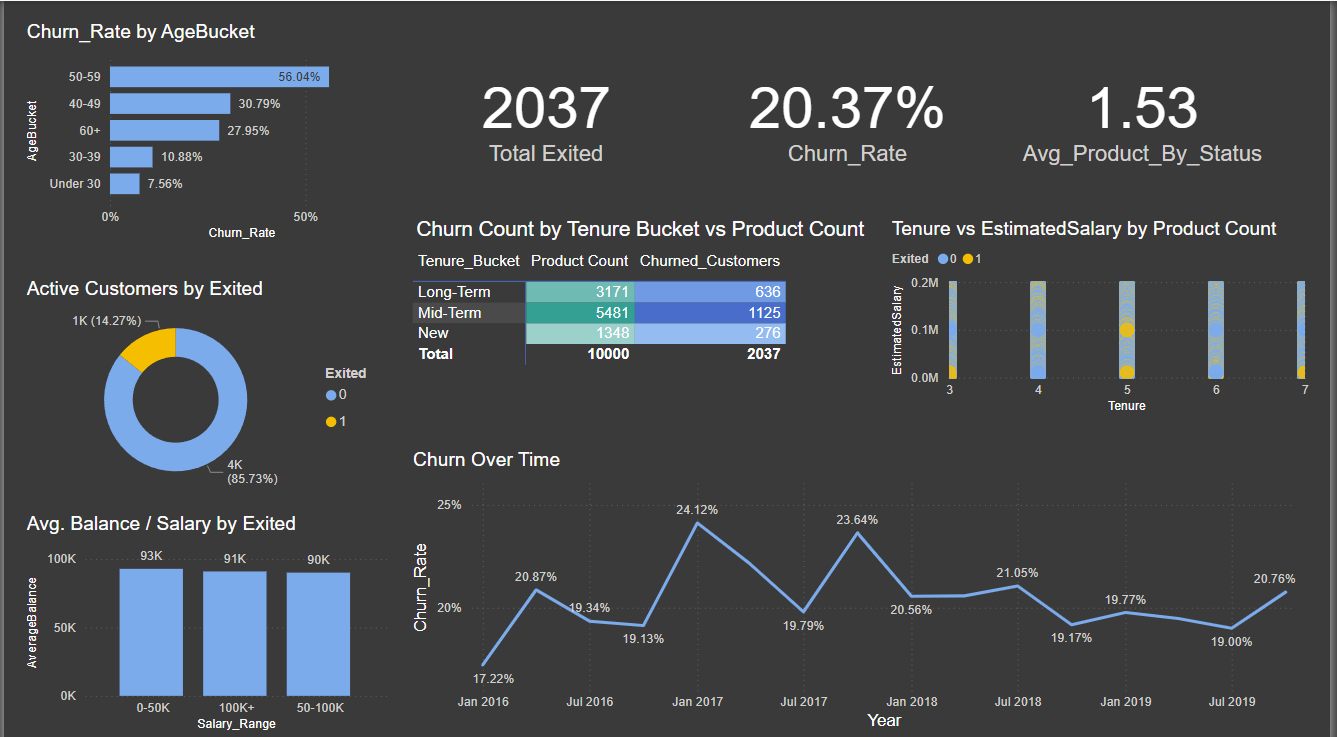
Ans: **Approach**

We analysed churned customers based on key demographic and behavioural attributes such as:

* **Tenure\_Bucket**
* **Product Count**
* **AgeBucket**
* **Estimated Salary**
* **Average Balance**
* **Exited status over time**

We evaluated the relationship between these fields to uncover common patterns among customers who exited the bank.

**Insights**

****

1. **High Churn in Mid-Term Customers (2–4 years):**
   * Mid-Term group had the **highest churn count (1,125)** and owned **5,481 products**, showing deep engagement before exiting.
2. **Age Group 50–59 Shows Peak Churn Rate (56.04%):**
   * This age segment is **disproportionately churning** compared to others, especially under-30s or 30–39s.
3. **Churned Customers Hold Significant Balances:**
   * Despite salary differences, churned customers across all salary buckets had **average balances close to ₹ 90 K+**, indicating profitable customer loss.
4. **High Product Count Does Not Ensure Retention:**
   * The churned group held **~1.5 products on average**, showing that multiple product adoption **does not equate to loyalty**.
5. **Time-Based Churn Spikes in Jan and Jul 2017:**
   * Notable churn rate spikes occurred during these quarters, suggesting **event-driven dissatisfaction** (e.g., policy changes or external market conditions).
6. **Active Customers are the Majority, but Exited Ones Are High-Value:**
   * While only ~14% exited, the **financial value (products + balances)** of those lost is significant.

**Recommendations**

1. **Retention Triggers for Mid-Term Customers:**
   * Deploy **customer health scoring** between years 2–3 to flag early signs of dissatisfaction.
   * Offer **renewal perks, loyalty bonuses, and personalised offers** during this period.
2. **Age-Specific Engagement Programs:**
   * Build **age-friendly onboarding**, digital literacy support, and assign **dedicated relationship managers** to older customers.
   * Include **offline support** and **traditional communication** channels for the 50–59 segment.
3. **Personalised Financial Planning for Low-Income High-Balance Users:**
   * These customers need **custom products, safety-net options**, or savings growth strategies.
   * Launch **financial health advisory campaigns** for them.
4. **Evaluate Product Relevance Post-Adoption:**
   * Track **post-sale satisfaction** and create **product-usage engagement journeys**.
   * Avoid aggressive cross-selling; instead, **match product recommendations** using behavioural insights.
5. **Investigate Policy or Service Events in 2017:**
   * Audit internal records and customer service tickets from Jan & Jul 2017.
   * Prepare **change management communication plans** to avoid sudden dissatisfaction from future policy rollouts.
6. **Build Churn Prediction & Early Intervention Framework:**
   * Use **historical churn data to build a predictive model**.
   * Enable **real-time alerts** for at-risk customers based on tenure, engagement drop, and support history.
7. Are 'Tenure', 'NumOfProducts', 'IsActiveMember', and 'EstimatedSalary' important for predicting if a customer will leave the bank?

Ans: **Variable Importance for Churn Prediction-**

**1. Tenure ✅ Important**

**Evidence:**

* Mid-Term customers (Tenure 2–4 years) account for **over 55% of churn** (1,125 out of 2,037 churned).
* Long-term customers churn less despite having significant product usage.
* New customers have lower churn numbers, suggesting **churn risk peaks in mid-tenure**.

**Conclusion:** Tenure influences churn patterns and is **strongly predictive** of customer exit behaviour.

**2. NumOfProducts ✅ Important, but Non-Linear**

**Evidence:**

* Churned customers had **moderate to high product counts**, with an average of around **1.53**.
* High product ownership **did not reduce churn**—in fact, it may indicate pressure selling or dissatisfaction post-sale.
* This shows a **non-linear** effect: beyond a point, more products **do not guarantee retention**.

**Conclusion:** NumOfProducts is important, but **its relationship with churn is complex** and needs to be used carefully in modelling.

**3. IsActiveMember ✅ Highly Important**

**Evidence (from previous visuals):**

* The pie chart of **Exited vs Active Customers** suggests that **inactive customers churn more**.
* Active membership is **positively correlated with retention**.
* This variable likely acts as a **behavioural signal**—inactive customers show early signs of detachment.

**Conclusion:** A **key churn predictor**. Strongly consider this in your model or scorecard.

**4. EstimatedSalary ❌ Less Important (by itself)**

**Evidence:**

* Churn happens across all salary buckets, with **no strong pattern**.
* Average balances and churn rates are relatively **flat across income levels**.
* The **50–100K**, **0–50K**, and **100 K+** buckets show similar average balances, and no bucket dominates churn% %.

**Conclusion:** Salary alone is **not a strong predictor**, though it may add value **when combined with other features** (e.g., balance-to-income ratio).

1. Utilise SQL queries to segment customers based on demographics and account details.

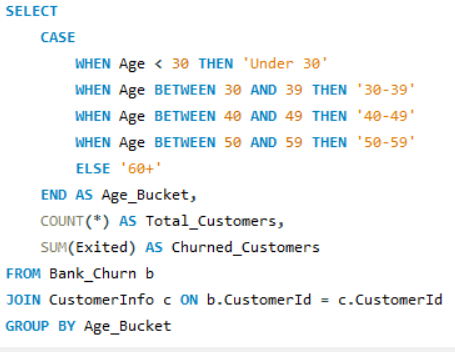
Ans: **Approach**

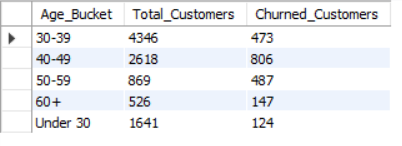
Here, I’ve performed segmentation of bank customers based on key demographic and account attributes to identify patterns in customer churn. The segmentation was done using SQL queries on the following dimensions:

1. Age Buckets
2. Gender
3. Tenure
4. Number of Products
5. Activity Status
6. Balance and Salary Range
7. Combined Segmentation (Gender, Age, Tenure, Product, Activity)

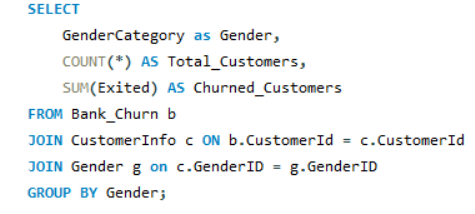
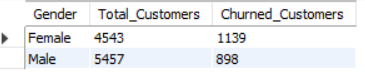
Each query calculates:

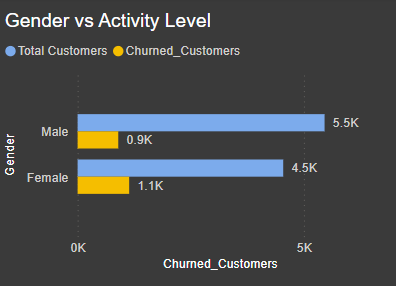
* Total number of customers in each segment
* Number of churned customers (SUM(Exited))

**1. Segment by Age Bucket**  
  


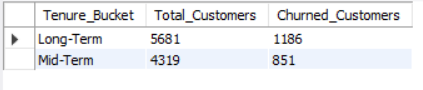

**2. Segment by Gender**

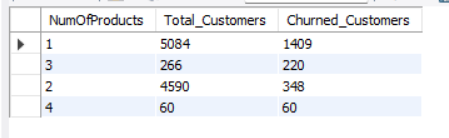
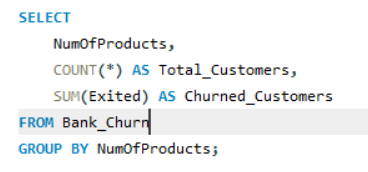
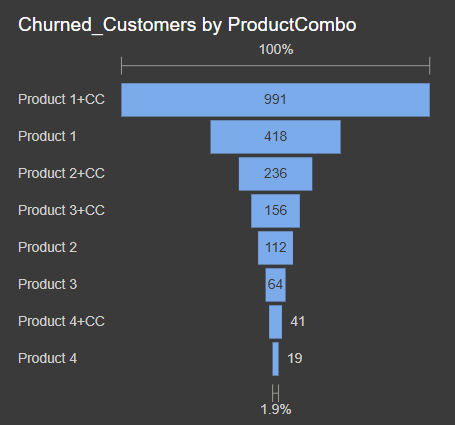


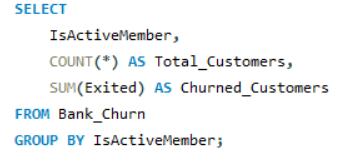
**3. Segment by Tenure Bucket**

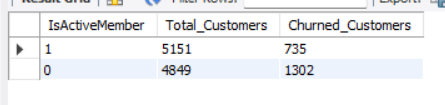


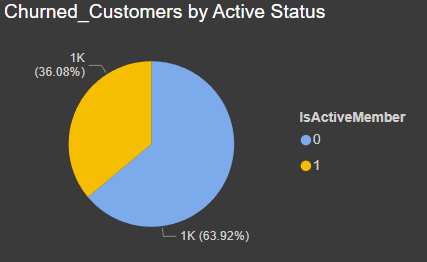


**4. Segment by Product Holdings**

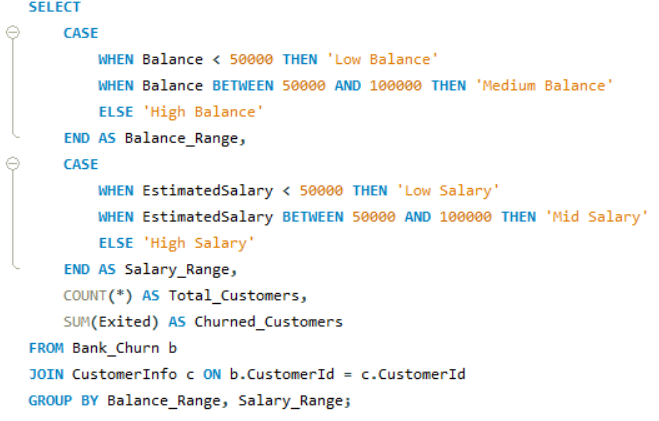
** **

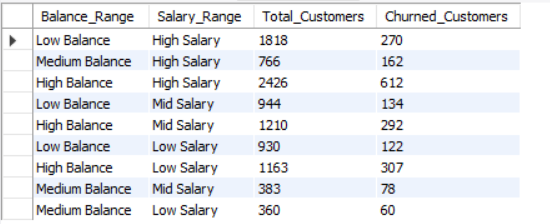
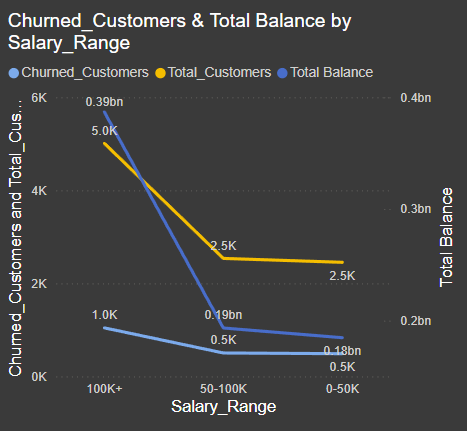
1. **Segment by Activity Status**

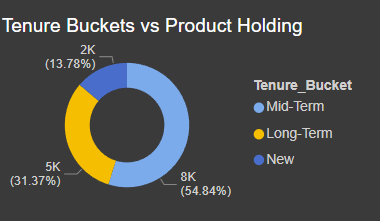
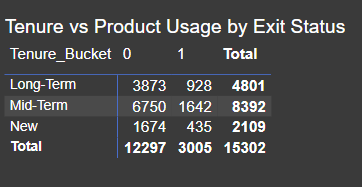
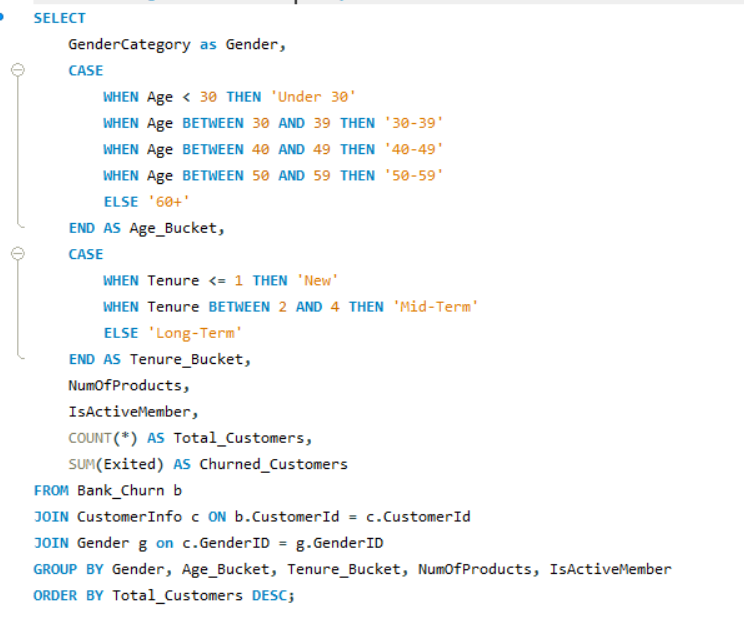
****

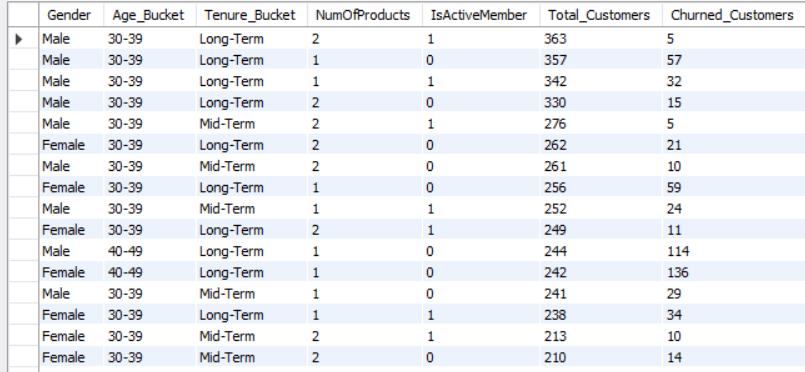
****

1. **Segment by Balance and Salary Range**

****

**  
  
**

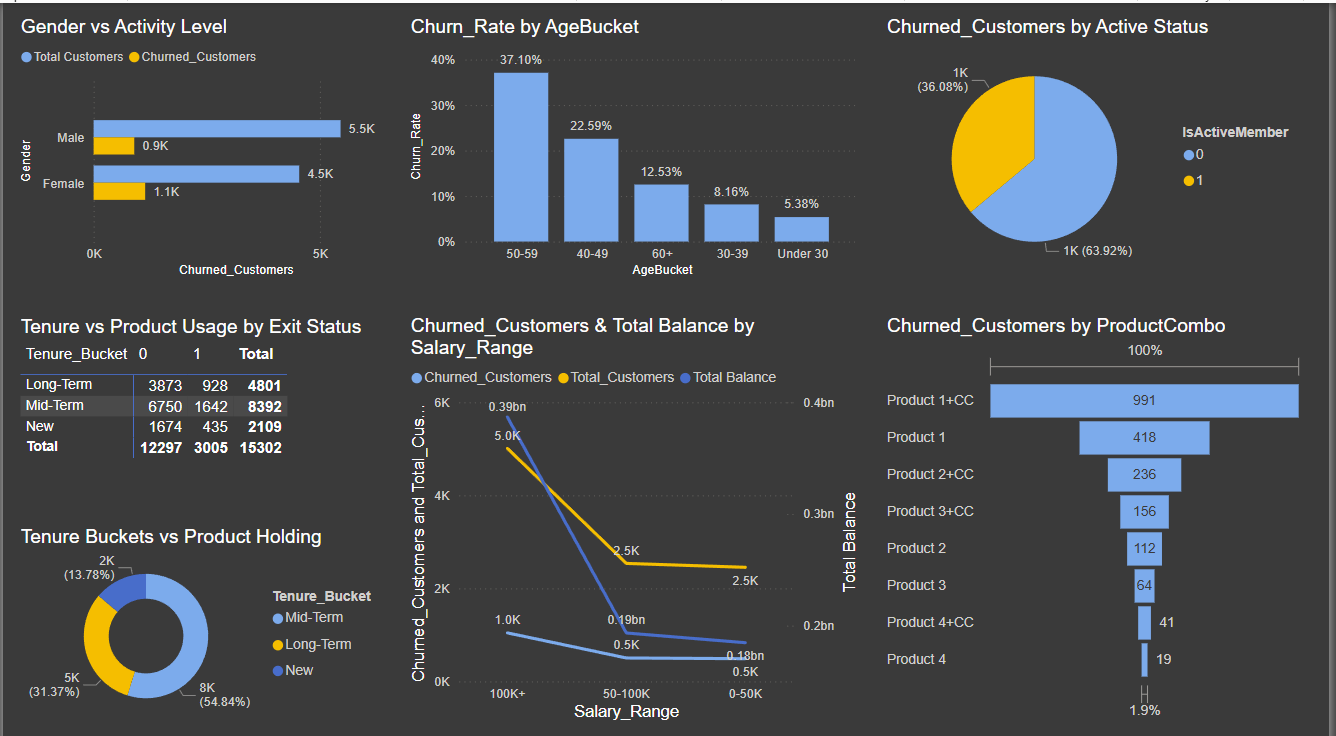
1. **Full Segmentation: Combine Key Dimensions  
     
    **

****

This **strategic and multidimensional** approach covers key facets of customer profiling like:

* **Demographic Segmentation**: Age buckets and Gender to understand who is churning.
* **Behavioural Segmentation**: Tenure, Number of Products, and Activity Status to analyse engagement and loyalty patterns.
* **Financial Segmentation**: Balance and Estimated Salary to capture the monetary value and financial strength of customers.
* **Combined Segmentation**: Blending all above dimensions to build granular customer profiles and identify high-risk cohorts.

This methodology supports both **descriptive churn insights** and forms the foundation for **predictive churn modelling**.

**Insights  
  
**

**1. Age-Driven Churn Trends**

* Churn rises significantly after age 40, with the 50–59 segment showing the highest churn rate at 37.1%.
* Customers under 30 have the lowest churn (5.38%), indicating stronger loyalty among younger segments.

**2. Gender-Based Churn**

* While male customers are more in number, female customers show higher churn volume (1.1K vs. 0.9K).
* This highlights a disproportionate churn risk among female customers.

**3. Tenure Impact on Loyalty**

* Mid-term customers (2–4 years) form the largest base.
* Long-term customers churn significantly less, suggesting increased loyalty with time.

**4. Product Holdings Influence**

* Highest churn is linked to customers with only 1 product, particularly those with a Product 1 + Credit Card combination.
* These combinations may indicate low engagement or perceived value.

**5. Balance vs. Salary Patterns**

* High-balance, high-salary customers, though financially strong, still experience churn, implying that financial value alone doesn't ensure retention.
* Churn doesn’t consistently decrease with higher salary, suggesting other factors (e.g., service, personalisation) play a role.

**6. Activity Status Matters**

* Active members contribute ~64% of total churn volume, likely due to a larger customer base.
* A churn rate comparison, not just count, is needed to draw clearer conclusions about engagement vs. risk.

**Recommendations**

**1. Age-Based Retention Plans**

* Focus retention efforts on customers aged 40–59 with:
  + Lifecycle-based communications
  + Exclusive benefits refreshers
  + Engagement campaigns tailored to mid-life financial goals

**2. Gender-Personalised Strategy**

* Address churn among female customers by:
  + Developing financial products and communication tailored to women's preferences
  + Exploring service gaps or unmet needs in female segments

**3. Tenure-Led Loyalty Programs**

* Strengthen customer relationships through:
  + Tenure-based rewards or loyalty tiers
  + Proactive outreach to mid-term customers nearing the high-churn phase

**4. Product Portfolio Optimisation**

* Reevaluate and enhance low-value product combinations like Product 1 + Credit Card by:
  + Offering value-based bundles
  + Introducing personalised product upgrades

**5. Move Beyond Salary Segmentation**

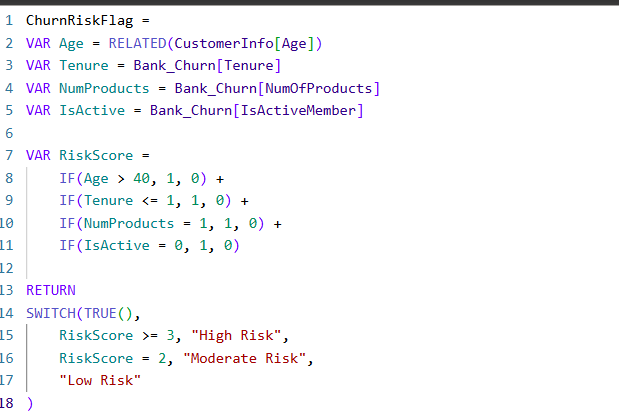
* Salary alone is not predictive of churn.
  + Use a layered segmentation approach by combining financials with product usage, digital engagement, and tenure.

1. How can we create a conditional formatting setup to visually highlight customers at risk of churn and to evaluate the impact of credit card rewards on customer retention?

Ans: **Approach**

1 **Conditional Formatting for Churn Risk Identification**

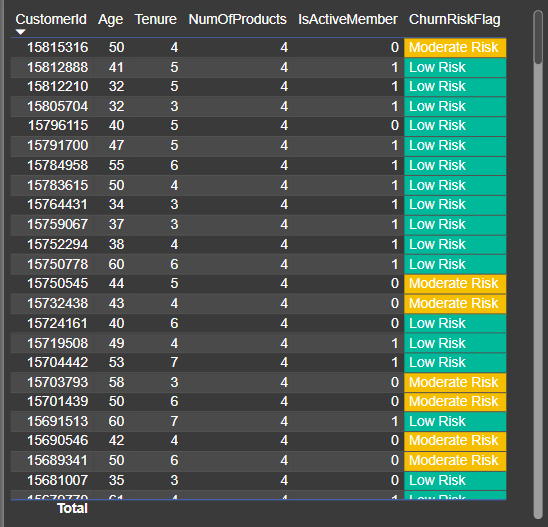
* Built a calculated column using a **Risk Scoring Model** based on key churn predictors:
  + Age > 50
  + Tenure ≤ 3 years
  + Only 1 product
  + Inactive member (IsActiveMember = 0)
* Each factor contributes to a risk score (0–4):
  + **High Risk**: Score ≥ 3
  + **Moderate Risk**: Score = 2
  + **Low Risk**: Score < 2
* Created a ChurnRiskFlag column and applied **traffic-light conditional formatting**:
  + 🔴 High Risk
  + 🟡 Moderate Risk
  + 🟢 Low Risk

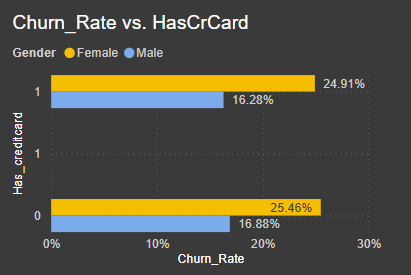


* This setup allows quick visibility of vulnerable customers across dashboards and tables.

2. **Evaluating Credit Card Reward Impact**

* Segmented customers by:
  + HasCrCard (Credit Card Ownership)
  + Gender (Male/Female)
* Compared churn rates between cardholders and non-holders.
* Visualized churn performance across these segments using:
  + Bar Charts: Gender × Credit Card Ownership
  + Cross-Tabs: Risk level overlaid with Credit Card participation
  + Tables: Enhanced with ChurnRiskFlag-based formatting



 **Insights**

**1. Risk-Based Customer Profiles Are Clear**

* Customers with:
  + **Low tenure**,
  + **Older age**,
  + **One product**, and
  + **Inactive status**  
    → are significantly more likely to churn.
* Risk flagging through conditional formatting makes these customers easy to isolate and act on.

**2. Credit Card Ownership Has Limited Retention Value**

* Churn rates:
  + **Females**:
    - With Card = 24.91%
    - Without Card = 25.46%
  + **Males**:
    - With Card = 16.28%
    - Without Card = 16.88%
* Indicates that while **credit card ownership slightly improves retention**, the difference is **marginal**.
* Suggests **reward programs currently lack engagement or perceived value**.

**3. Gender-Based Churn Patterns Persist**

* **Female customers** churn more than males, regardless of credit card ownership.
* Points to **possible disconnect in product appeal, communication, or value perception** for female users.

**Recommendations**

**1. Operationalize Risk Flagging**

* Refresh risk scores and apply conditional formatting **weekly**.
* Share reports with **Customer Success Teams** to prioritize high-risk outreach using traffic-light logic.

**2. Rebuild the Rewards Program**

* Current rewards have a **low retention impact**. Enhance by:
  + **Tiering**: Better benefits with deeper engagement
  + **Personalization**: Tailor based on customer segments (e.g., women, high-balance)
  + **Gamification**: Use progress bars, surprise bonuses, or milestone rewards

**3. Launch Targeted Outreach for At-Risk Female Customers**

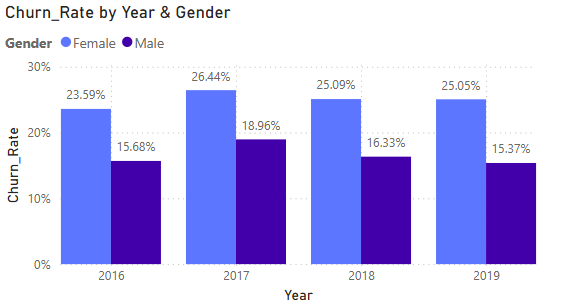
* Combine gender + risk scoring to **identify high-churn female segments**.
* Conduct qualitative research: surveys, interviews, and usability studies to understand pain points.
* Use preferred channels and messaging tone to **better align with expectations**.

While this analysis is built on rule-based segmentation, a logical next step is to implement a predictive churn model using machine learning (e.g., logistic regression or decision tree). This would help assign churn probabilities to each customer for even more proactive and precise retention strategies. Such a model can be explored collaboratively with the data science team in the future.

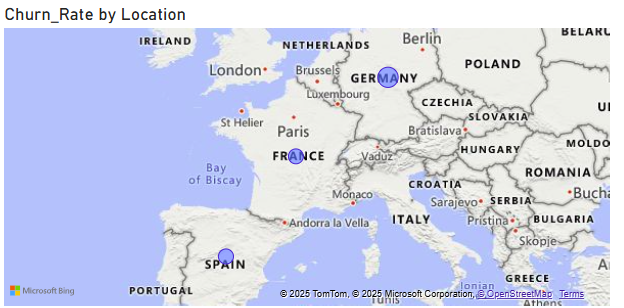
1. What is the current churn rate per year and overall, as well as in the bank? Can you suggest some insights to the bank about which kind of customers are more likely to churn and what different strategies can be used to decrease the churn rate?

Ans: **Approach:** To assess the yearly and overall churn trends, and identify risk groups, I analysed:

1. Churn Rate by Year & Gender
   * Visualized churn trends over time, split by gender, to observe patterns and disparities.



1. Churn Rate by Location
   * Used a map visual to highlight churn concentration by country, identifying regional risk pockets.

****

**Insights**

1. **Female Customers Consistently Churn More**
   * Across all years:
     + Female churn ranges from **23.59% to 26.44%**
     + Male churn stays lower: **15.37% to 18.96%**
   * Despite having fewer female customers, **their churn rate is significantly higher**, showing a **gender-based churn disparity**.
2. **Churn Rate Hasn’t Improved Over the Years**
   * Female churn has **remained above 25%** in most years.
   * Male churn has also plateaued, hovering around **16–18%**.
   * Indicates that previous retention efforts may not have been sufficiently targeted or impactful.
3. **Germany Shows the Highest Churn**
   * Among all locations, **Germany has the largest churn concentration**, followed by **France** and **Spain**.
   * May be due to region-specific dissatisfaction, service issues, or local competitors.

**Recommendations**

**1. Implement Gender-Specific Retention Strategies**

* Introduce **tailored financial products** and **communication strategies** for female customers.
* Launch campaigns that resonate with different life stages or financial goals (e.g., wealth planning, flexible savings).
* Conduct **surveys, interviews, or NPS analysis** to understand the key drivers of dissatisfaction among female segments.

**2. Develop Location-Focused Interventions**

* Investigate the **cause of high churn in Germany and France**—are customers leaving due to service gaps, competitive offerings, or cultural misalignment?
* Allocate **regional customer success teams** or advisors to handle churn-prone areas with more personalized support.
* Adjust marketing or rewards programs to **match local customer expectations**.

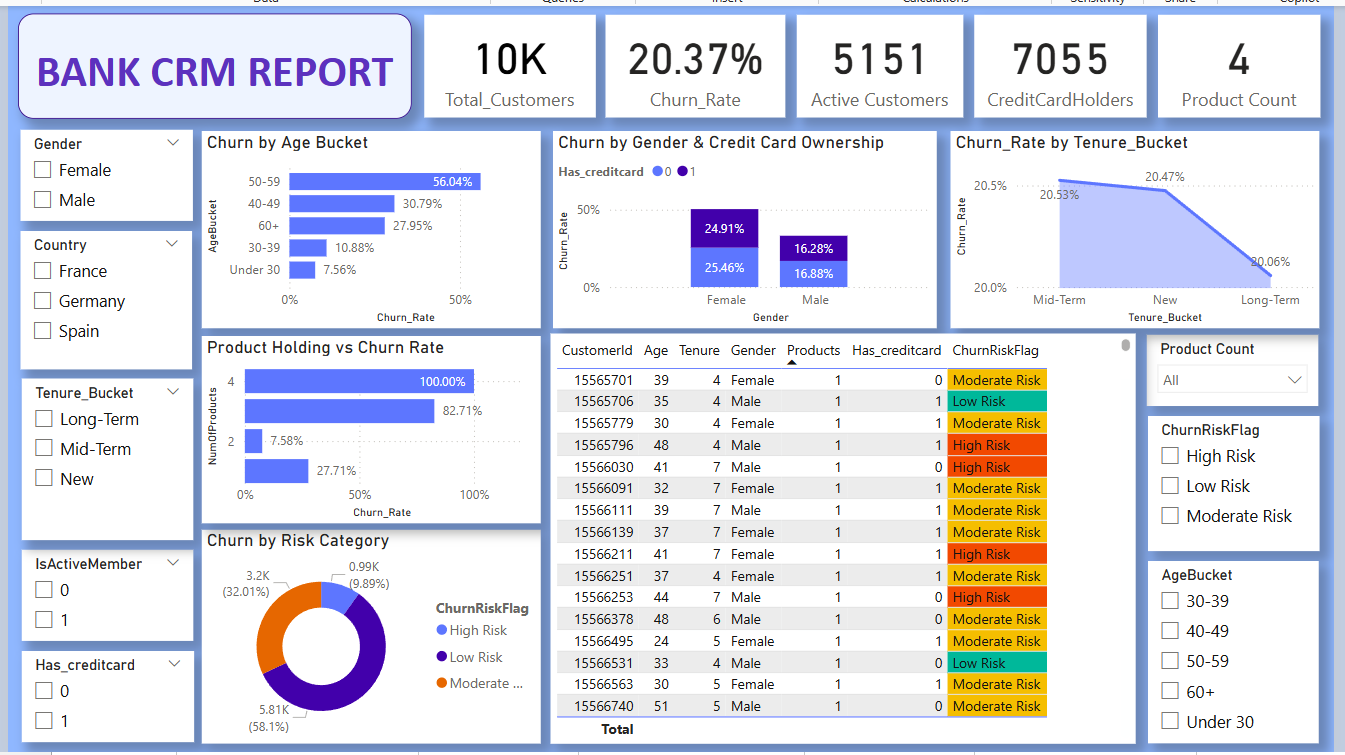
**3. Integrate Churn KPIs into Executive Dashboards**

* Use visuals like the **gender-wise churn bar chart** and **churn map** in ongoing reports.
* Add drill-through options by location and gender to enable **actionable segmentation**.
* Refresh visuals monthly to detect rising churn trends before they escalate.

4. **Re-evaluate Retention Programs**

* As churn is high even in financially strong countries, this suggests that **monetary value alone doesn’t ensure loyalty**.
* Focus on **engagement, personalization, and digital experience** rather than only financial incentives.
* For example, customers in Germany may benefit more from simplified UX and better mobile banking support than traditional rewards.

1. Create a dashboard incorporating all the KPIs and visualisation-related metrics. Use a slicer in order to assist in selection on the dashboard.

Ans: Here is the dashboard that has been created using all the necessary KPIs and visuals.  
  


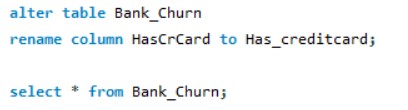
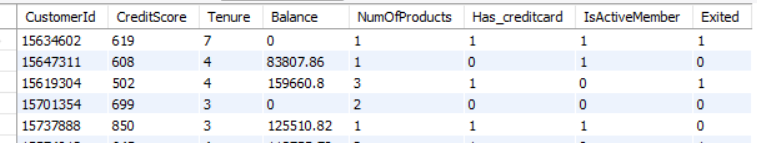
1. How would you approach this problem if the objective and subjective questions weren't given?

Ans: If the objective and subjective questions weren’t provided, I would start by clearly defining the **business goal** through exploratory data analysis (EDA). My steps would be:

1. **Understand the Domain and Dataset**
   * Since this is a bank dataset, the natural business objective could be around **customer retention**, **churn reduction**, or **segmentation**.
   * I’d start by looking at the data columns like Exited, Tenure, Balance, NumOfProducts, etc., to identify what kind of behaviour we’re trying to understand or improve.
2. **Initial Exploration**
   * Run summary stats and distributions using SQL or Power BI visuals to:
     + Understand churn percentages.
     + Identify outliers or patterns.
     + Check relationships between customer attributes and churn.
3. **Frame My Own Questions**
   * Based on the patterns, I’d frame objective questions like:
     + Which age group has the highest churn?
     + Does tenure affect loyalty?
     + Is inactivity a strong churn predictor?
   * And subjective ones like:
     + Why might older customers be leaving?
     + Is our reward system ineffective?
4. **Segment the Data Strategically**
   * Perform segmentation on key factors (age, gender, tenure, product usage).
   * Combine these to create risk profiles or customer personas.
5. **Design Visual & Analytical Framework**
   * Create a report/dashboard to monitor churn drivers.
   * Apply techniques like:
     + Conditional formatting to flag risks.
     + Cohort analysis or time-based comparisons.
6. **Insights & Recommendations**
   * Translate findings into **actionable strategies**, such as reward redesign, tenure-based campaigns, or targeted outreach.

So, even without predefined questions, I would let the **data speak**, guide the **business problem discovery**, and move from raw exploration to meaningful insights.

1. In the “Bank\_Churn” table, how can you modify the name of the “HasCrCard” column to “Has\_creditcard”?

Ans: Using SQL or Power BI Power Query, we can rename the column from “HasCrCard” to “Has\_creditcard”.  
  
SQL Query:  
  
  
  
  
Power Query:  
  
