IBM Telco Customer Churn Data Analysis Project Report

Project Overview

The **IBM Telco Customer Churn** project aimed to analyze customer churn behavior to uncover patterns and insights that could help the telecommunications company improve customer retention. Utilizing the IBM Telco Customer Churn dataset and MySQL Workbench, the analysis focused on understanding the factors influencing churn, such as payment methods, customer demographics, contract types, and customer feedback.

Objectives

The primary objective was to address three key queries using MySQL, derive actionable insights, and propose recommendations to reduce customer churn:

- **Query 1**: How can personalized offers be tailored based on age, gender, and contract type for the top 5 groups with the highest average monthly charges among churned customers to potentially improve customer retention rates?
- **Query 2**: What are the feedback or complaints from those churned customers?
- **Query 3**: How does the payment method influence churn behavior?

Query 1: Considering the top 5 groups with the highest average monthly charges among churned customers: how can personalized offers be tailored based on age, gender, and contract type to potentially improve customer retention rates?

CODE:

```
Select
    d.gender, s.contract, round(avg(s.monthly_charge),2) as avg_monthly_charge, round(avg(s.tenure_in_months),2) as avg_tenure,
    Case
        when d.age < 30 then "Under 30"
        when d.age between 30 and 50 then "30-50"
        when d.age between 51 and 70 then "50-70"
        else "above 70" end as age_group

from status as st
join services as s on st.customer_id = s.customer_id
join demographics as d on st.customer_id = d.customer_id
where st.churn_label = "yes"
group by age_group, d.gender, s.contract
order by avg_monthly_charge desc
limit 5;
```

In this analysis, the top 5 groups with the highest average monthly charges among churned customers are segmented based on three key criteria: **Gender, Contract, Age Group**

Logic and Flow of the Query:

- The query starts by joining the necessary tables (status, services, demographics) to pull in all relevant customer data.
- It filters the data to include only churned customers (WHERE st.churn_label = "Yes").
- The customers are then grouped by age group, gender, and contract type.
- Within each group, the query calculates the average monthly charge and tenure.
- Finally, it orders the results by the highest average monthly charges and limits the output to the top 5 groups.

How This Query Helps:

This query identifies the top 5 customer segments (based on age group, gender, and contract type) that have the highest average monthly charges among churned customers. These high-value groups can then be targeted with personalized retention strategies

1. **SELECT Clause:** The SELECT clause determines the data you want in the result set, including fields or expressions. The order in the SELECT clause dictates how the output columns are displayed but doesn't impact the query's logic. In our case, the SELECT clause is used to create age groups, pull gender and contract details, and calculate averages for monthly charges and tenure.
2. **JOIN Clause:** The JOIN clauses combine data from multiple tables based on a common key (customer_id). The sequence of JOIN operations (in our case, joining services and demographics to status) is not crucial for inner joins, as they produce the same result regardless of the order. However, for outer joins, the sequence can impact the result. Here, INNER JOIN ensures you're only working with customers that exist in all tables.
3. **GROUP BY Clause:** The GROUP BY clause groups the result set by the specified columns (age group, gender, and contract). The sequence here defines the hierarchy of grouping (i.e., the rows are first grouped by age group, then by gender, and finally by contract). Changing the order can affect how the data is grouped and summarized, but in our case, any order of these fields would yield useful results because they are all relevant for aggregation.
4. **WHERE Clause:** The WHERE clause filters the data before grouping, ensuring that only churned customers are considered (churn_label = 'Yes'). It is positioned before GROUP BY to limit the rows processed by the grouping and aggregation functions, focusing the analysis specifically on churned customers.

QUERY RESULT

gender	contract	avg_monthly_charge	avg_tenure	age_group
Female	One Year	104.36	50.12	above 70
Female	Two Year	100.26	54.88	30-50
Male	Two Year	97.34	66.1	30-50
Female	One Year	90.38	43.71	Under 30
Female	One Year	87.62	44.7	50-70

ANALYSIS: TOP 5 GROUPS:

1.High-Value Customers with Long Tenure:

- **Male, 30-50, Two-Year Contract, Avg Monthly Charge: 97.34, Avg Tenure: 66.1 months**:
    - This group has high average monthly charges and a long tenure. This suggests that they are high-value, loyal customers who were on a two-year contract but still churned.
    - **Retention Strategy**: Offering loyalty rewards, such as exclusive discounts, free device upgrades, or premium service add-ons (e.g., increased data or additional services like streaming) at renewal time could incentivize these customers to stay. Emphasizing stability and long-term benefits will likely resonate with this group.

2. Elderly Female Customers:

- **Female, Above 70, One-Year Contract, Avg Monthly Charge: 104.36, Avg Tenure: 50.12 months**:
    - This segment has the highest average monthly charge but is on a one-year contract, which might suggest they prefer shorter commitments despite having a relatively long tenure. Elderly customers might seek stability and reliability in services.
    - **Retention Strategy**: Personalized offers could focus on simplified billing plans or senior-friendly services, like enhanced customer support and discounts for long-term contracts. Bundling home services or family plans might also appeal to them, given that they're paying a premium for their current services.

3. Female, 30-50, Two-Year Contract:

- **Female, 30-50, Two-Year Contract, Avg Monthly Charge: 100.26, Avg Tenure: 54.88 months**:
    - Similar to the male group in the same age range, this group has relatively high average monthly charges and long tenure, indicating their value.
    - **Retention Strategy**: These customers could be motivated by family-oriented plans, additional data packages, or enhanced service quality (e.g., better network coverage). Offering flexibility in switching between contract types or providing family plans might help in retaining them.

4. Younger Female Customers:

- **Female, Under 30, One-Year Contract, Avg Monthly Charge: 90.38, Avg Tenure: 43.71 months**:
    - This group is younger, with a shorter tenure, and is on a one-year contract. They have lower monthly charges compared to other segments, which suggests they might be more cost-sensitive and are likely looking for flexibility.
    - **Retention Strategy**: Offering flexible, no-commitment plans with incentives like discounted rates for extending their contracts or referral bonuses could appeal to this

group. They may also value mobile data plans with streaming partnerships or loyalty points to earn rewards for staying longer.

5. Middle-Aged Female Customers (50-70):

- **Female, 50-70, One-Year Contract, Avg Monthly Charge: 87.62, Avg Tenure: 44.7 months**:
    - This group pays moderately high charges and has a relatively short tenure. They are in the middle of their working lives and might seek affordability and stability in their plans.
    - **Retention Strategy**: Offering discounts on bundled services like family plans, home internet, or additional data packages could be effective for this segment. Providing flexible renewal options or promotions for shifting to longer-term contracts could help reduce churn.

Query 2: What are the feedback or complaints from those churned customers?

CODE: Customer Count by churn_category and churn_reason

```
select churn_category, churn_reason, count(customer_id) as customer_count,
        rank() over(partition by churn_category order by count(customer_id)) as category_rank
from status
where churn_label = "Yes"
group by churn_category, churn_reason
order by churn_category, category_rank;
```

Explanation of the Query:

Query Purpose: The query aims to retrieve feedback or complaints from churned customers by counting how many customers provided each churn reason within each churn category. Additionally, it ranks these reasons based on their frequency.

1. **PARTITION BY churn_category:**
    - This part divides the result set into separate groups or "partitions" based on the values in the churn_category column. Each unique churn category will form its own partition.

- For example, if there are three churn categories (e.g., "Pricing", "Service Quality", "Customer Support"), the ranking will be calculated independently within each of these categories.

2. **ORDER BY COUNT(customer_id):**
    - Within each partition (each churn category), the rows are ordered based on the count of customers (using COUNT(customer_id)).
    - The rank is assigned starting from 1 for the churn reason with the highest count (most frequently reported), followed by the next highest count, and so on.
    - If two churn reasons have the same count, they will receive the same rank. For instance, if "Pricing" and "Service Quality" both have 10 customers reporting them, they might both get a rank of 1, and the next unique churn reason would get a rank of 3.

By ranking churn reasons within each category, you can quickly identify which reasons are most common among churned customers. This is crucial for understanding customer sentiment and addressing key issues that may be causing churn.

The ranking mechanism provides a systematic way to handle ties. If two or more churn reasons have the same number of customer reports, they are given the same rank. This is important in real-world scenarios where customer feedback can be nuanced and overlapping.

QUERY RESULT

churn_category	churn_reason	customer_count	category_rank
Attitude	Attitude of support person	220	1
Attitude	Attitude of service provider	94	2
Competitor	Competitor had better devices	313	1
Competitor	Competitor made better offer	311	2
Competitor	Competitor offered more data	117	3
Competitor	Competitor offered higher download speeds	100	4
Dissatisfaction	Product dissatisfaction	77	1
Dissatisfaction	Network reliability	72	2
Dissatisfaction	Service dissatisfaction	63	3
Dissatisfaction	Limited range of services	37	4
Dissatisfaction	Lack of self-service on Website	29	5
Dissatisfaction	Poor expertise of online support	13	6
Dissatisfaction	Poor expertise of phone support	12	7
Other	Don't know	130	1
Other	Moved	46	2
Other	Poor expertise of online support	18	3
Other	Deceased	6	4
Price	Price too high	78	1
Price	Long distance charges	64	2
Price	Extra data charges	39	3
Price	Lack of affordable download/upload speed	30	4

CODE: Running customer total by churn category

```sql
with t1 as (
select churn_category, churn_reason, count(customer_id) as customer_count
from status
where churn_label = "Yes"
group by churn_category, churn_reason
)

select churn_category, churn_reason, customer_count,
        sum(customer_count) over (partition by churn_category order by churn_reason) as running_total
from t1
order by churn_category, churn_reason;
```

Explanation of Using a CTE

1. **Separation of Logic**: The CTE (t1) **aggregates customer counts** based on churn_category and churn_reason before applying any window functions. This keeps aggregation and running total calculations distinct and clear.
2. **Avoiding SQL Conflicts**: Without the CTE, combining COUNT() for aggregation and SUM() for running totals in the same query would lead to errors, as all non-aggregated columns must be in the GROUP BY clause.
3. **Facilitating Window Functions**: The CTE provides a structured dataset for the window function to operate on, allowing for smooth calculations of running totals.

Logic of the Code

- The CTE calculates the count of churned customers, creating a temporary table.
- The main query selects from this CTE to compute the running total of those counts, ordered by churn_reason.
- This structure allows for clear analysis of churn behavior without SQL errors.

Query Result

churn_category	churn_reason	customer_count	running_total
Attitude	Attitude of service provider	94	94
Attitude	Attitude of support person	220	314
Competitor	Competitor had better devices	313	313
Competitor	Competitor made better offer	311	624
Competitor	Competitor offered higher download speeds	100	724
Competitor	Competitor offered more data	117	841
Dissatisfaction	Lack of self-service on Website	29	29
Dissatisfaction	Limited range of services	37	66
Dissatisfaction	Network reliability	72	138
Dissatisfaction	Poor expertise of online support	13	151
Dissatisfaction	Poor expertise of phone support	12	163
Dissatisfaction	Product dissatisfaction	77	240
Dissatisfaction	Service dissatisfaction	63	303
Other	Deceased	6	6
Other	Don't know	130	136
Other	Moved	46	182
Other	Poor expertise of online support	18	200
Price	Extra data charges	39	39
Price	Lack of affordable download/upload speed	30	69
Price	Long distance charges	64	133
Price	Price too high	78	211

Trends and Patterns:

1. **Dominance of Competitor-Related Reasons:**
    - The most cited churn reasons are related to competitors, with "Competitor had better devices" and "Competitor made better offer" being the top two reasons for churn. This indicates that a significant number of customers are leaving due to perceived advantages offered by competing providers, such as better technology and promotional offers.
    - The strong emphasis on competitor offerings suggests that the company may need to conduct thorough market research to understand what competitors are providing and how they can enhance their value proposition (e.g., device upgrades, data packages).
2. **Customer Service Training:**

    - The "Attitude of support person" and "Attitude of service provider" are the next most common complaints. These suggest that customer service quality plays a crucial role in customer retention. Customers appear to value not just the product but also the interactions they have with support staff.

- Given the prominence of service-related complaints, there is a potential opportunity to invest in training programs for customer service representatives. Improving the customer experience through better support could mitigate churn and enhance loyalty.

3. **Product and Service Dissatisfaction:**

   - The category of "Dissatisfaction" is notably broad, encompassing various issues like "Product dissatisfaction," "Network reliability," and "Limited range of services." This suggests that while the customer may appreciate certain aspects of the service, they may also feel that their expectations are not being met.
   - The recurring themes of network reliability and service dissatisfaction indicate that technical improvements may be necessary. This could involve investing in infrastructure or offering higher-quality service plans to retain customers.

Query 3: How does the payment method influence churn behavior?

CODE: Churned Customers across PMT Methods; Revenue across PMT Methods

```
-- Query 3: How does the payment method influence churn behavior?

with t1 as (
select customer_id from status
where churn_label = "yes"
)

select s.payment_method, count(s.customer_id) as customer_count, round(sum(s.total_revenue), 2) as TotalRevenue_PMT_Method
from services as s
join t1 as t1 on s.customer_id = t1.customer_id
group by payment_method
order by TotalRevenue_PMT_Method desc;
```

Query Result

payment_method	customer_count	TotalRevenue_PMT_Method
Bank Withdrawal	1329	2759053.71
Credit Card	398	835124.46
Mailed Check	142	90281.65

Analysis of Query Results:

1. Bank Withdrawal (1329 churned customers, $2,759,053.71 in revenue):

- **High Churn Volume**: Bank withdrawal has the highest number of churned customers, accounting for a significant portion of the churned customer base. This suggests that customers using this payment method might be more likely to churn compared to others.
- **High Total Revenue**: Despite the high churn, this payment method also contributes the most revenue to the company. This indicates that customers using bank withdrawal may represent a high-value segment.
- **Possible Reasons for Churn**:
    - **Automatic Withdrawals**: Customers on automatic bank withdrawals might be less engaged with the service since their payments are automated. They may be less attentive to price changes or service issues, potentially leading to churn when dissatisfaction builds up.
    - **Bank Fees**: Some customers may be deterred by fees or issues related to their bank, such as overdrafts, insufficient funds, or transaction failures. This could lead to frustration and churn.
    - **Retention Opportunity**: Since this group contributes the most revenue, it would be beneficial for the company to focus retention efforts here, possibly through loyalty programs or providing incentives like discounted services for continuous bank withdrawals.

2. Credit Card (398 churned customers, $835,124.46 in revenue):

- **Moderate Churn Volume**: Credit card users are the second-largest group of churned customers, though significantly smaller than bank withdrawal. However, their churn rate could still be notable relative to their overall base.
- **Moderate Revenue Contribution**: The total revenue from churned credit card users is considerable, though much lower than bank withdrawal. This could indicate that while these customers are less likely to churn, they might not be as high-value as bank withdrawal customers.
- **Possible Reasons for Churn**:
    - **Payment Flexibility**: Credit card users tend to have more flexibility in managing their payments. However, they may also be more aware of recurring charges, leading them to cancel services if they feel they aren't receiving value.

- **Credit Card Expirations/Declines**: Churn could be triggered by expired cards or declined payments, causing service interruptions.
- **Retention Opportunity**: Offering improved payment tracking and flexibility (e.g., allowing alternative payment methods when cards expire) might reduce churn in this segment.

3. Mailed Check (142 churned customers, $90,281.65 in revenue):

- **Low Churn Volume**: Mailed check users have the smallest churn volume. This indicates that relatively few customers rely on this method, making it a niche segment.
- **Low Revenue Contribution**: The revenue from mailed check users is very low compared to the other payment methods. This suggests that customers using mailed checks are either lower-value customers or are more traditional, low-usage clients.
- **Possible Reasons for Churn**:
  - **Inconvenience**: Mailed checks are less convenient, and customers may churn due to the manual nature of this payment method, especially in an era where automated and digital payments are the norm.
  - **Customer Demographics**: The smaller volume and revenue could point to an older or more traditional customer base, who may not adopt new technology as easily, contributing to churn.
  - **Retention Opportunity**: Encouraging mailed check users to switch to more modern, automated payment methods (like bank withdrawal or credit card) through incentives could reduce churn in this segment.

CODE: Churn % across platforms

```sql
-- CTE for total customers
with TotalCustomers as (
select customer_id, payment_method from services
),

-- CTE for churned customers
ChurnedCustomers as (
select customer_id from status
where churn_label = "yes"
)

select
    t1.payment_method, count(t1.customer_id) as total_customers, count(t2.customer_id) as churned_customers,
    round((count(t2.customer_id)/count(t1.customer_id) *100), 2) as churn_rate_percentage
from TotalCustomers as t1
left join ChurnedCustomers as t2 on t1.customer_id = t2.customer_id  -- left join to include all results from t1 and not just the common ones B/w t1 & t2
group by t1.payment_method
order by churn_rate_percentage desc;
```

How the Query Works:

1. The TotalCustomers CTE selects all customers with their payment methods from the services table.
2. The ChurnedCustomers CTE selects only those customers who have churned from the status table.
3. The main query performs a LEFT JOIN between TotalCustomers and ChurnedCustomers, ensuring that all customers are included and that churned customers can be identified.
4. The query calculates the total number of customers for each payment method (total_customers), the number of churned customers (churned_customers), and then computes the churn rate for each payment method.

Purpose of LEFT JOIN:

- A LEFT JOIN ensures that all customers from TotalCustomers are included in the result, even if they haven't churned. If a customer hasn't churned, their corresponding t2.customer_id will be NULL, which allows us to differentiate churned from non-churned customers.
- This join links the TotalCustomers dataset (t1) with the ChurnedCustomers dataset (t2) based on the customer_id. It ensures that we can count both churned and non-churned customers for each payment method.

- **ROUND((COUNT(t2.customer_id) / COUNT(t1.customer_id)) * 100, 2) AS churn_rate_percentage:** This calculates the **churn rate** by dividing the number of churned customers by the total number of customers for each payment method, then multiplying by 100 to get a percentage. The ROUND() function rounds the result to 2 decimal places.

Query Result

payment_method	total_customers	churned_customers	churn_rate_percentage
Bank Withdrawal	2722	1329	48.82
Mailed Check	324	142	43.83
Credit Card	1458	398	27.3

Analysis of Payment Method Churn Results

1. **Bank Withdrawal**:
    - **Total Customers**: 2,722
    - **Churned Customers**: 1,329
    - **Churn Rate**: 48.82%
    - **Insight**: Bank withdrawal accounts for the largest customer base but also exhibits the highest churn rate. Nearly half of these customers have churned, indicating significant dissatisfaction or engagement issues. This high churn rate suggests that customers might feel disconnected due to the automatic nature of withdrawals, leading to churn when they experience dissatisfaction or find better alternatives.

2. **Mailed Check**:
    - **Total Customers**: 324
    - **Churned Customers**: 142
    - **Churn Rate**: 43.83%
    - **Insight**: The churn rate for mailed checks is also relatively high, though the overall customer base is much smaller compared to bank withdrawals. The manual process of

mailing checks may lead to inconvenience, prompting customers to seek easier payment methods. However, the smaller customer base indicates that this payment method might already be on the decline as customers transition to digital options.

3. **Credit Card**:
    - **Total Customers**: 1,458
    - **Churned Customers**: 398
    - **Churn Rate**: 27.30%
    - **Insight**: Credit card users represent a moderate churn rate compared to the other methods, suggesting that these customers may be more engaged with their accounts and aware of their payments. This lower churn rate indicates that credit card users might be more likely to perceive value in the service, or they are better equipped to manage their subscriptions and respond to billing issues.