**HUMBER INSTITUTE OF TECHNOLOGY**

**AND ADVANCED LEARNING**

**(HUMBER COLLEGE)**

**Logo, company name

Description automatically generated**

**GROUP PROJECT – TABLEAU**

**Team: 1**

|  |  |  |
| --- | --- | --- |
| **Last name** | **First name** | **Student number** |
| **Avila** | **Janna** | **N01512064** |
| **Ganem** | **Farid** | **N01510374** |
| **Garg** | **Deepanshi** | **N01510660** |
| **Gupta** | **Sidharth** | **N01467753** |
| **Rocha** | **Milena** | **N01510465** |
| **Saxena** | **Lakshya** | **N01510799** |
| **Sharma** | **Kapil** | **N01501686** |

**Submitted to: Professor Guneet Kaur Sawhney**

**Course: Communication & Visualization - BIA-5400-0LA**

**Submission Date: August 18th, 2023**

Table of Contents

[**1.**](#_heading=h.gjdgxs) **Abstract** 3

[**2.**](#_heading=h.30j0zll) **Business Problem Statement** 4

[**3.**](#_heading=h.1fob9te) **Data explanation** 5

[**4.**](#_heading=h.3znysh7) **Data Preparation** 6

[**5.**](#_heading=h.2et92p0) **Dashboards** 8

[**6.**](#_heading=h.tyjcwt) **Challenges** 15

[**7.**](#_heading=h.3dy6vkm) **Conclusion & Recommendations** 17

[**8.**](#_heading=h.1t3h5sf) **References** 20

# **Abstract:**

This report presents a comprehensive analysis of the "Telco Customer Churn" dataset within the telecommunications industry. The central objective of this analysis is to determine the main reason for customer churn, a critical challenge impacting revenue streams, market positioning, and overall sustainability. This study searches and shows insights into the drivers underlying customer attrition by exploring demographic indicators, service utilization patterns, billing behaviors, and other pertinent variables. The analysis was made in different stages such as data preprocessing, creation of diagnostic dashboards, and formulation of strategic solutions. The crucial role of data-informed decision-making, tailored retention strategies, and predictive modeling are underscored in addressing the multifaceted issue of churn. By finding patterns and offering strategic recommendations, this report aims to empower telecommunications enterprises to effectively navigate the challenges posed by customer churn, optimize their efforts in retaining customers, and flourish within the competitive business landscape.

# **Business Problem Statement**

We aim to study the 'Telco Customer Churn' dataset and determine which customers are leaving the telecommunications company and make different visualizations for it. We'll look at age, products used, how they use services, bills, and more to understand why they leave or move to other companies. We aim to dissect the 'Telco Customer Churn' dataset and unravel the pivotal factors contributing to customer attrition within the telecommunications sector. We explore demographics, service utilization patterns, billing dynamics, and pertinent attributes to decipher the underlying determinants propelling customer disengagement. Furthermore, we constructed informative dashboards in Tableau capable of showing churned customers by age, geographic location, churned base on contracts, services acquired based on churned, between other visualizations.

# **Data explanation**

The Telco Customer Churn Analysis dataset was obtained from Kaggle, available at <https://www.kaggle.com/datasets/ylchang/telco-customer-churn-1113>. Designed to explore the factors that influence customer churn within the telecommunications industry, the dataset comprises various customer attributes, details about service subscriptions, and indications of churn status, collectively offering a holistic view of customer interactions. The extracted dataset initially contained a total of 52 columns and 7043 observations, with certain columns lacking essential information. In alignment with our problem statement, we opted to narrow down our focus to select columns that would be conducive to our modeling and visualization objectives. Presented below are descriptions of several key columns commonly used in our analysis:

1. Customer ID: A unique identifier assigned to each customer.
2. Gender: Reflects the gender of the customer (Male or Female).
3. Senior Citizen: Indicates whether the customer is classified as a senior citizen (Yes or No).
4. Partner: Indicates whether the customer has a partner (Yes or No).
5. Dependents: Indicates whether the customer has dependents, such as children (Yes or No).
6. Tenure: Represents the duration, in months, that the customer has been subscribed to the service.
7. Phone Service: Indicates whether the customer has an active phone service (Yes or No).
8. Multiple Lines: Indicates whether the customer has multiple phone lines (Yes, No, or No phone service).
9. Internet Service: Specifies the type of internet service the customer has (DSL, Fiber optic, or No).
10. Online Security: Indicates whether the customer has subscribed to online security service (Yes, No, or No internet service).
11. Online Backup: Indicates whether the customer has subscribed to online backup service (Yes, No, or No internet service).
12. Device Protection: Indicates whether the customer has device protection service (Yes, No, or No internet service).
13. Tech Support: Indicates whether the customer has tech support service (Yes, No, or No internet service).
14. Streaming TV: Indicates whether the customer has streaming TV service (Yes, No, or No internet service).
15. Streaming Movies: Indicates whether the customer has streaming movies service (Yes, No, or No internet service).
16. Contract: Specifies the contract term chosen by the customer (Month-to-month, One year, or Two year).
17. Paperless Billing: Indicates whether the customer has opted for paperless billing (Yes or No).
18. Payment Method: Specifies the preferred method of payment (Electronic check, Mailed check, Bank transfer, or Credit card).
19. Monthly Charges: Represents the monthly charge for the subscribed services.
20. Total Charges: Represents the cumulative charges billed to the customer.
21. Churn: Indicates whether the customer has churned, i.e., canceled the service (Yes or No).

The Telco Customer Churn Analysis dataset is highly versatile and can be utilized for various analytical purposes, including:

* Predictive Modeling: Developing predictive models to anticipate customer churn based on their attributes and behaviors.
* Customer Segmentation: Grouping customers based on shared characteristics to tailor retention strategies.
* Feature Importance Analysis: Identifying the most influential factors contributing to churn.
* Trend Analysis: Investigating patterns and trends in customer behavior over time.
* Service Enhancement: Identifying opportunities for service improvements based on customer preferences.

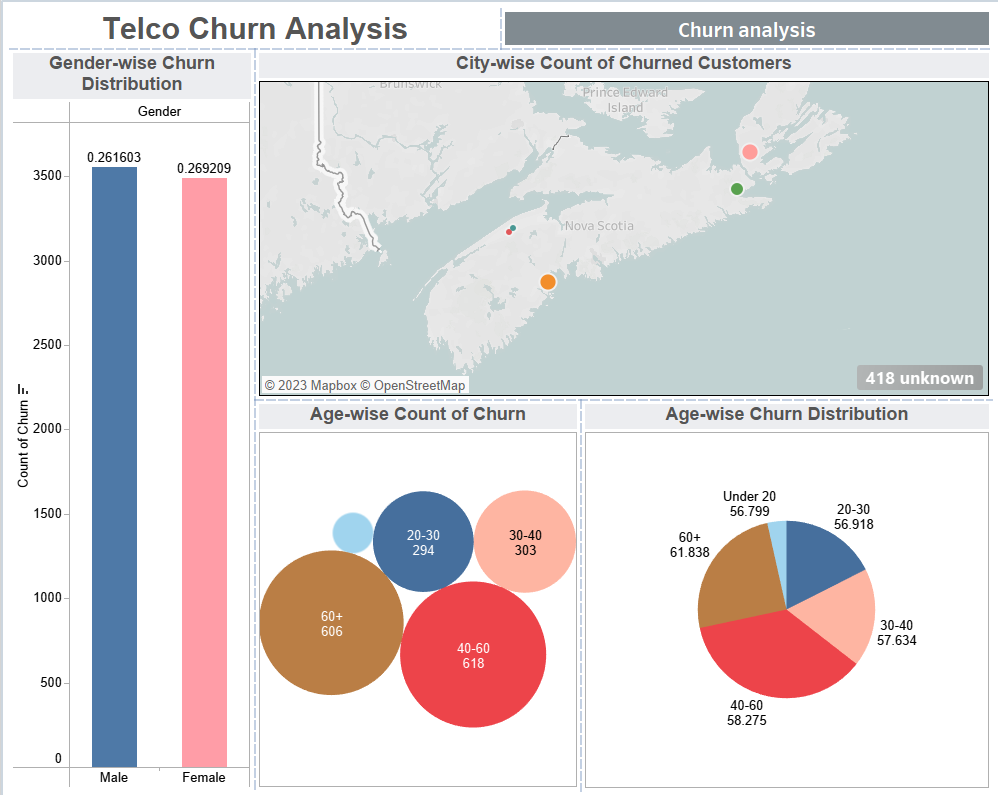
# **Data Preparation**

In Tableau, we used the data interpreter feature to assist in importing and preparing data for analysis. It is designed to help users clean and structure their data more efficiently before creating visualizations and dashboards. With the data interpreter, data cleaning was performed. The data interpreter in Tableau helps identify and fix common data issues, such as extra rows, columns, headers, and formatting problems. To sum up, we did not require further preparation and did not find any null and duplicated values for the dataset.

It also handles Headers and Footers when importing data for ana. Headers, footers, or summary rows might need to be excluded from the analysis. The data interpreter recognizes these elements and assists in removing or ignoring them to ensure accurate analysis. This Tableau tool also deals with data Structure Inference merging Similar Columns, handling Unstructured Data, identifying Data Types, recognizing hierarchical data: and suggesting Pivot and Split Operations.

Moreover, we create some calculations for further illustrations. One of the calculations we implemented was the count of Churned; count the amount of clients that churned, the Age category; depending on the customer age split them in categories, the Senior City category; and the Internet\_Phone column. In this last column, we wanted to check if a client had just one service or both of them.

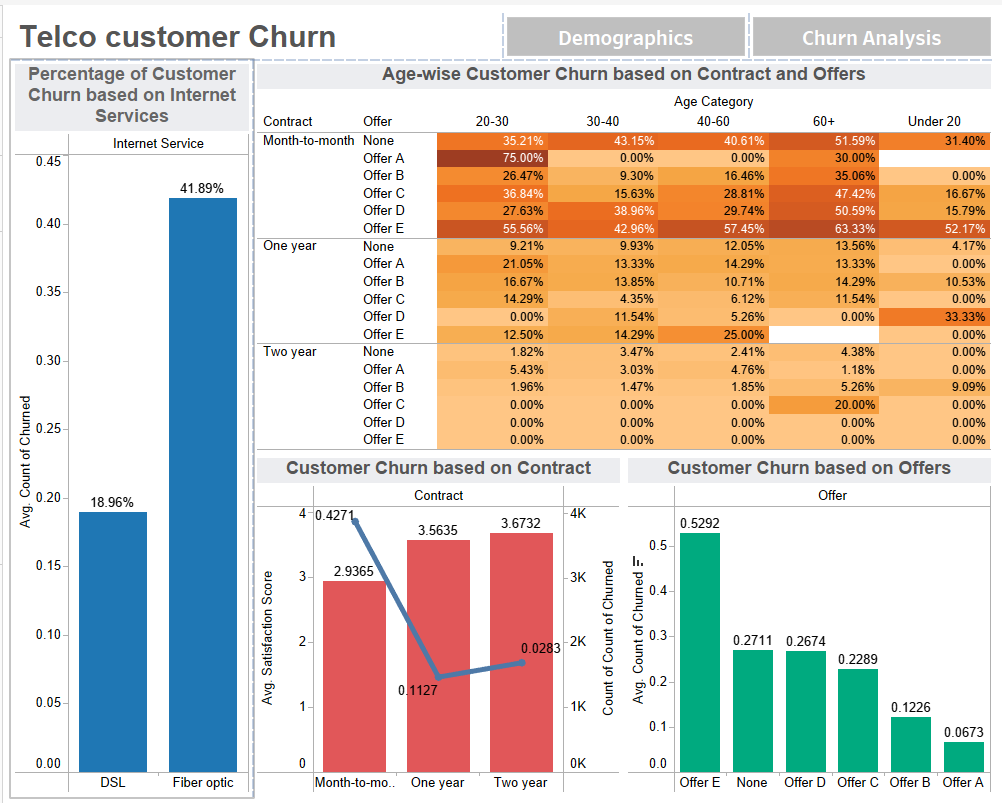
# **Dashboards**



This is the first set of visuals we created to understand the demographics of customers leaving:

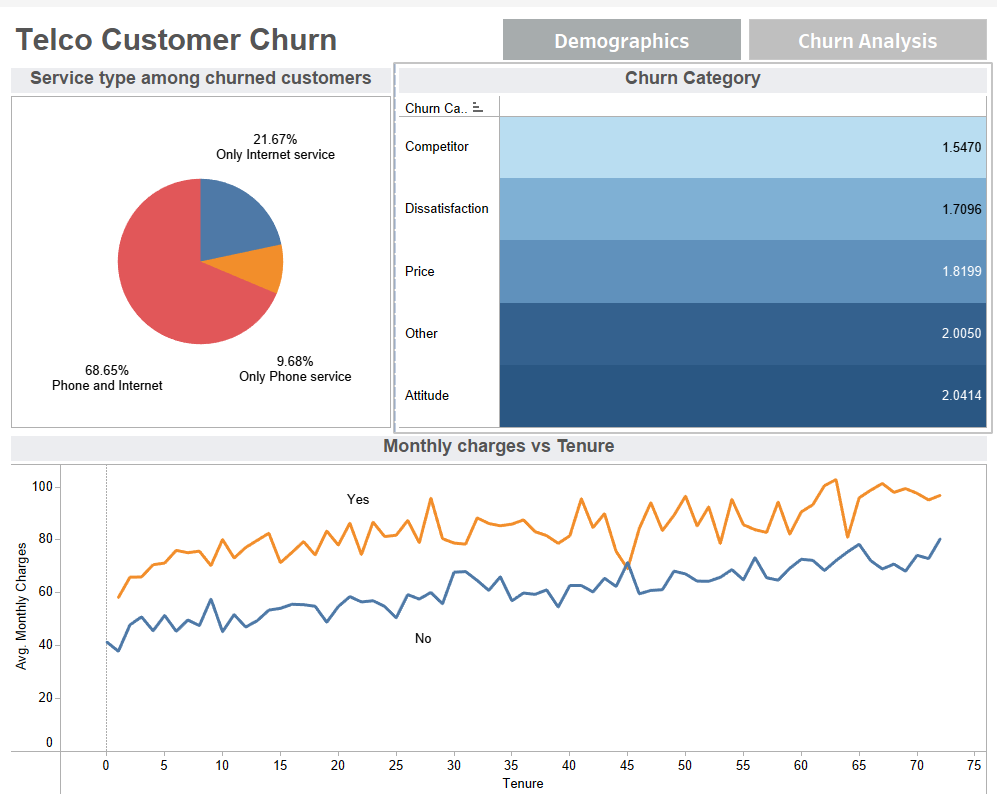
* Gender-wise Churn Distribution bar graph shows how many customers left based on gender. It looks like slightly more females (27%) left than males (24%).
* The City-wise Count of Churned Customers map displays the number of customers who live in different cities. Some cities have more leaving customers. For instance, Toronto had the most leaving customers (76), followed by Montreal (58) and Vancouver (54).
* There are also bubble and pie charts named "Age-wise Count of Churn" that tell us how many and what percentage of customers left based on their age group. The age group with the most leaving customers is 31-40 years old (112 customers, 28%), followed by 41-50 years old (98 customers, 24%), and 21-30 years old (86 customers, 21%).

These charts help us see why customers leave, like their gender, where they live, and how old they are. But remember, other things might also affect this, like how satisfied they are, the quality of the service, or the prices. We made another set of visuals focusing on these factors to look deeper into this.



For the second set of visuals we focused on service traits that might influence customer churn:

* The first chart, "Percentage of Customer Churn based on Internet Services," displays the percentage of customers who left the company based on their internet service type. It turns out that customers with fiber optic internet service had the highest churn rate (41%), followed by those with DSL internet service (19%). The lowest churn rate was for customers without internet service (8%).
* The second chart, "Customer Churn based on Contract and Offers," shows the percentage of customers who left depending on their contract type and whether they got any special offers. The chart clearly shows that customers with month-to-month contracts who didn't receive any offers had the highest churn rate (43%). Next are customers with month-to-month contracts who received one offer (28%). On the other hand, customers with two-year contracts who received two offers had the lowest churn rate (1%).
* The third chart, "Contract based on Customer Churn," reveals the percentage of customers who left based on their contract type. It's most common for customers with month-to-month contracts to churn (42%), followed by those with one-year agreements (11%). Customers with two-year contracts had the lowest churn rate (3%).
* Lastly, the fourth chart, "Customer Churn based on Age and Offers," breaks down the percentage of customers who left by age group and whether they got any offers. The highest churn rate is among customers aged 21-30 who received no offers (35%). Customers aged 31-40 without recommendations follow (29%). However, those aged 51-60 who received two offers had the lowest churn rate (4%).



The third dashboard showcases additional relevant data about customer churn drivers in the Telecommunications sector:

* The first visualization corresponds to a pie chart which indicates the type of service that churned customers had. We can observe that most of the customers that left the company had both internet and phone services (68%), in second place we find customers which only had internet service (22%) and lastly customers with phone service only (9%). This information can give us insights that probably we will need to focus our efforts on improving the internet service provided by the company.
* The second visualization contains a highlight table with information about the reasons given by customers that churned and its correlation with the satisfaction level. A lighter-coloured box indicates that customers leaving due to better proposals obtained from competitors usually have the lowest satisfaction score (1.54), followed by the dissatisfaction reason (1.70).
* The third visualization contains a line chart with average monthly charges and tenure. We can observe that customers that churned usually have higher monthly charges compared to those who remained in the company. Moreover, we can see a slight upward trend in the line, meaning that as a customer stays longer with the company, it pays higher monthly charges for their telecommunication services.

# **Challenges**

While making the analysis of the "Telco Customer Churn" dataset, several challenges emerged in the project. These challenges were:

1. Data Quality and Integrity: Making sure the dataset is accurate and reliabile is a primary challenge. Incomplete or erroneous data could potentially lead to skewed insights and inaccurate model predictions that will make wrong analysis.
2. Feature Selection: Determining which features to include in the analysis requires careful consideration due to the selected business problem. While some attributes may seem irrelevant at first, they might hold hidden patterns crucial for understanding customer churn.
3. Ethical Considerations: Extracting insights from customer data must adhere to ethical guidelines and data privacy regulations. Safeguarding customer information and ensuring responsible data usage is a significant challenge
4. Business Impact: Translating analysis findings into tangible business strategies and actions that enhance customer retention is a challenge that requires alignment with organizational goals.
5. Changing Dynamics: The telecommunications industry is dynamic, with customer preferences and market trends evolving rapidly. Analyzing historical data to predict future churn patterns must consider this ever-changing landscape.

By realizing and approaching these challenges we were able to approach the business problem in a correct way. We analyzed meaningful insights from the "Telco Customer Churn" dataset, which can make a better business to devise effective strategies for mitigating churn and ensuring long-term success in the company.

# **Conclusion & Recommendations**

In conclusion, the examination of the "Telco Customer Churn" dataset has provided valuable insights into the factors fostering customer attrition in the telecommunications sector that we have studied. Our approach of data preprocessing, exploratory analysis, and informative dashboards have deepened our comprehension of the relevant determinants influencing customer turnover. This comprehension is important for companies to maintain their current customer base.

Several demographic and service-oriented aspects wield considerable influence over customer churn. Gender, age bracket, contract type, and internet service can be considered as notorious predictors of churn within our analysis. Customers with month-to-month contracts and those utilizing fiber optic internet service exhibit greater churn rates. Moreover, the absence of special offers or discounts seems to have an impact on customer attrition.

Based on the insights obtained from our analysis, the following recommendations are applicable:

* Improve Internet services: Focus on improving internet service quality and performance, given its significant impact on churn, especially the issues related to fiber optic internet to enhance customer satisfaction and reduce churn in this segment.
* Contract Incentives: Encourage customers to opt for longer contracts by offering discounts, rewards, or other incentives to promote loyalty. Longer contracts are proven to carry lower churn rates.
* Targeted Promotions: Design personalized offers for specific age groups, particularly in the middle-aged segments (ages 40-60), to capture their interest and loyalty since they have demonstrated to be more willing to churn.
* Competitive Pricing: Ensure competitive pricing strategies to prevent customers from leaving due to better proposals from competitors.
* Regional Strategies: Address issues specific to high-churn cities like Toronto, Montreal, and Vancouver through targeted marketing campaigns and service improvements.
* Customer Satisfaction Enhancement: Continuously monitor customer satisfaction and address issues promptly to prevent dissatisfaction from leading to churn.

Although churn is a complex issue influenced by various factors, by regularly tracking and analyzing customer feedback, service quality, and market trends the company can help refine these recommendations over time and stay competitive in the business arena.

# **References**

1. Telco customer churn (11.1.3+). (n.d.). Kaggle. Retrieved August 16, 2023, from <https://www.kaggle.com/datasets/ylchang/telco-customer-churn-1113>