# **CSDA 5320**

# **FINAL PROJECT**

**Data set Source link:** [**https://www.kaggle.com/datasets/blastchar/telco-customer-churn**](https://www.kaggle.com/datasets/blastchar/telco-customer-churn)

**Initial Question/Hypothesis**  
The primary question guiding this analysis is:

**“What factors contribute most significantly to customer churn?”**

The hypothesis is that specific customer attributes, such as high monthly charges, longer tenure, and lower engagement with certain services, are associated with higher churn rates. Analyzing this will provide actionable insights for retention strategies.

**Project Summary**

This project involves analyzing customer churn data to identify key factors influencing churn behavior. The dataset contains 7,043 entries with 21 variables, including demographic details, service usage, and billing information. The objective is to derive actionable insights for improving customer retention and reducing churn rates.

**Data Analysis Summary**

1. **Dataset Overview**:
   * The dataset was cleaned to remove 11 rows with missing values in the TotalCharges column.
   * Key variables analyzed include Churn (target variable), tenure, MonthlyCharges, Contract, and InternetService.
2. **Key Visualizations**:
   * **Churn Distribution**: ~26% of customers churned, while ~74% retained.
   * **Monthly Charges**: Higher monthly charges are associated with a higher likelihood of churn.
   * **Tenure**: Customers with shorter tenure are more likely to churn.
   * **Contract Type**: Month-to-month contracts exhibit a higher churn rate compared to one- and two-year contracts.
   * **Payment Method**: Out of all payment methods people prefer electronic check.
   * **Considering Gender**: From this data, we can say that compared to males significant amount of females are using our services. Maybe we can attract new customers or stop the existing customers from churning by providing offers/discounts concentrating on females.

**Statistical/Hypothetical Question**

**Do contract type and tenure significantly predict customer churn?**

* Hypothesis: Customers on month-to-month contracts with shorter tenure are more likely to churn.

**Outcome of EDA**

1. **Contract Type**:
   * Customers on month-to-month contracts have the highest churn rates.
   * Long-term contracts (one- or two-year) correlate with lower churn.
2. **Tenure**:
   * Average tenure for customers who churned: ~18 months.
   * Average tenure for retained customers: ~38 months.
3. **Monthly Charges**:
   * Median monthly charges for churned customers: ~$79.
   * Median monthly charges for retained customers: ~$64.
4. **Internet Service**:
   * Fiber optic users exhibit higher churn compared to DSL or no internet service.

**Project Discoveries, Obstacles, and Lessons Learned by me:**

* **Discoveries**:
  + Flexible month-to-month contracts are a key driver of churn, likely due to fewer barriers to exit.
  + High monthly charges increase the likelihood of churn, suggesting potential pricing concerns.
  + Retention strategies should focus on early-stage customers with lower tenure.
* **Obstacles**:
  + Missing values in the TotalCharges column required careful handling.
  + Ambiguous entries in categorical variables (e.g., "No internet service") required thoughtful interpretation.
* **Lessons Learned**:
  + Data cleaning and visualization are crucial for identifying meaningful patterns.
  + Customer behavior is influenced by a combination of contract type, pricing, and service quality.

**This Was Missed During the Analysis:**

* **Potential Misses**:
  + A deeper dive into demographics (e.g., SeniorCitizen status, gender) might provide further insights.
  + Interaction effects between variables, such as the combined impact of tenure and InternetService, were not fully explored.
* **Variables That Could Help:**
  + Customer satisfaction scores or complaint data could provide a more nuanced view of churn behavior.
  + External factors, such as competitors' pricing and service quality, might impact churn.

**Incorrect Assumptions**

* Initially assumed that higher tenure guarantees customer retention, but some high-tenure customers also churn due to dissatisfaction or better alternatives.
* Assumed Total Charges would be clean, but it required significant preprocessing due to its incorrect data type and missing values.

**Challenges and Gaps in Understanding**

* Difficulty in quantifying the impact of qualitative variables like PaymentMethod and Paperless Billing.
* Understanding why some long-term customers churn despite long tenure and affordable pricing.

**Summary of Findings**

This analysis highlights the importance of:

* Offering incentives to month-to-month contract customers.
* Revising pricing strategies to address high monthly charges.
* Proactively engaging with new customers to increase their tenure.
* We can say that compared to males’ significant number of females are using our services. Maybe we can attract new customers or stop the existing customers from churning by providing offers/discounts concentrating on females.

The findings will help in creating targeted retention strategies and improving overall customer satisfaction.