**Exploratory Data Analysis – Case study**

**Customer Churn Project Documentation**

Project is to explore the customer churn dataset available to gain some insights and share with stakeholders.

**Key Observations Explained (Columns in the dataset)**

1. **Customer ID**: Unique identifier for each customer.
2. **Gender**: Male or Female.
3. **SeniorCitizen**: Binary indicator (0 or 1) representing whether the customer is a senior citizen.
4. **Partner & Dependents**: Indicators for customers having a partner or dependents.
5. **Tenure**: The number of months a customer has been with the company.
6. **Services**: Multiple columns represent services (e.g., PhoneService, InternetService, OnlineSecurity).
7. **Contract**: Type of contract (e.g., Month-to-month, One year, Two year).
8. **Billing and Payment**: Columns like PaperlessBilling, PaymentMethod, MonthlyCharges, and TotalCharges.
9. **Churn**: Target variable indicating if the customer has churned (Yes or No).

* Data sourced from <https://www.kaggle.com/code/bhartiprasad17/customer-churn-prediction/notebook>

**Objectives**

Explore the dataset available and share some insights.

Conducted Exploratory analysis and shared Insights using Data Visualization.

**Exploratory Analysis**

* Previewed five rows to understand the columns in the dataset.
* Checked the total number of customers (sampled) in the dataset.
* Checking for unique values in the dataset.
  1. Column gender has unique values.
* Male
* Female
  1. Column internet service has unique values.
* DSL
* Fiber Optic
* No
  1. Column payment type has unique values.
* Electronic check
* Mailed check
* Bank Transfer (automatic)
* Credit card (automatic)
  1. Column Contract has unique values.
* Monthly
* One year
* Two years
* Checked for missing values.
* Checked for duplicates but found none.
* Checked number of customers who churned, and the total amount of revenue lost.

Total number of customers who churned = 1,869

Revenue lost = 2,862,926.90

* Checked which internet service customers churned.
* Checked which contract types these churned customers had.
* Checked whether these churned customers had tech support.
* Checked whether the churned customers had dependents.

**Insights**

* The dataset contains 7,043 entries with 21 columns.
* Average tenure: **32.4 months** (range: 1–72 months).
* Average monthly charges: **$64.80** (range: $18.25–$118.75).
* Average total charges: **$2,283.30** (range: $18.80–$8,684.80).
* **Churn Distribution**:

**73.4% of customers** did not churn.

**26.6% of customers** churned.

This indicates that churn is a minority class, which may require specific attention in predictive modeling.

* **Tenure and Churn**:

Average tenure of non-churning customers: **37.7 months**.

Average tenure of churning customers: **18 months**.

This means that customers with shorter tenures are more likely to churn.

* **Monthly Charges and Churn**:

Average monthly charges for non-churning customers: **$61.31**.

Average monthly charges for churning customers: **$74.44**.

This indicates that customers with higher monthly charges are more likely to churn.

* **Churn by Contract Type**:

**Month-to-month**: 42.7% churn rate (highest).

**One-year**: 11.3% churn rate.

**Two-year**: 2.8% churn rate (lowest).

This indicates that longer-term contracts significantly reduce churn