This project performs an **Exploratory Data Analysis (EDA)** on a Telco customer churn dataset to identify patterns and factors contributing to customer loss.

Project Overview

The primary goal is to analyze customer behavior and demographic data to understand why customers leave the service (churn). The dataset contains information for **7,043 customers** with **21 different attributes**, including demographics, account information, and subscribed services.

Key Data Analysis Steps

- **Data Cleaning:**
  - Handled missing or blank values in the `TotalCharges` column by replacing them with "0" and converting the data type to float for numerical analysis.
  - Simplified the `SeniorCitizen` column by converting binary values (0/1) into "Yes/No" for better readability during visualization.
- **Data Profiling:**
  - Verified that there are no missing values across all 21 columns after initial cleaning.
  - Confirmed the uniqueness of records by checking for duplicate `customerID` entries, ensuring no redundant data influenced the results.
- **Statistical Analysis:**
  - Used descriptive statistics to determine that the average customer tenure is approximately **32 months**, with average monthly charges of **$64.76**.

Data Attributes Analyzed

The analysis considers several categories of data:

- **Demographics:** Gender, Senior Citizen status, Partner, and Dependents.
- **Services:** Phone service, multiple lines, internet service (DSL, Fiber optic), and security/streaming add-ons.
- **Account Information:** Contract type (Month-to-month, One year, Two year), payment method, and billing preferences.

Visualizations

The project utilizes `matplotlib` and `seaborn` to create visual distributions, specifically focusing on how different **Payment Methods** (e.g., Electronic check, Mailed check, Bank transfer) correlate with churn rates.

S

Advanced Data Cleaning and Preprocessing

- **TotalCharges Correction**: The initial data had blank values in the `TotalCharges` column for customers with zero tenure. These were replaced with "0" and the entire column was converted to a `float` data type to enable mathematical calculations.
- **Categorical Encoding**: For easier interpretation in charts, binary values (0 and 1) in the `SeniorCitizen` column were mapped to "yes" and "no".
- **Uniqueness Verification**: A specific check for duplicate `customerID` values confirmed that there are zero duplicate entries, ensuring each record represents a unique customer.

Deep-Dive Statistical Insights

- **Monthly Spending Patterns**: The project's descriptive statistics show a wide range of customer spending, with **MonthlyCharges** varying from a minimum of **$18.25** to a maximum of **$118.75**.
- **Tenure Extremes**: While the average tenure is 32 months, the dataset includes new customers (0 months) as well as long-term loyalists with up to **72 months** (6 years) of service.
- **Lifetime Value**: The `TotalCharges` column reveals that some customers have contributed as much as **$8,684.80** in total revenue to the company.

Service & Contractual Analysis

- **Technical Add-ons**: The analysis tracks several security and convenience services, including **Online Security**, **Online Backup**, **Device Protection**, and **Tech Support**, which are potential factors in reducing churn.
- **Entertainment Services**: The data distinguishes between customers who use **Streaming TV** and **Streaming Movies**, allowing for an analysis of how entertainment packages impact retention.
- **Billing Preferences**: The project examines the impact of **Paperless Billing** and various **Contract types** (Month-to-month vs. Multi-year) on a customer's likelihood to churn.

Visualization and Patterns

- **Churn by Payment Method**: The project specifically utilizes `seaborn` to create count plots that compare Churn (Yes/No) across different payment methods.
- **Impactful Visuals**: By adding bar labels directly onto the charts, the analysis provides immediate clarity on the exact number of customers leaving versus staying for each category (e.g., how many "Electronic check" users churned compared to "Credit card" users).