**📄 Customer Churn Analysis and Prediction**

**Task 2: Exploratory Data Analysis (EDA)**

**Objective**

The aim of this task is to perform exploratory data analysis on the customer churn dataset of a telecommunications company. The analysis seeks to uncover key patterns and relationships that influence customer churn and to provide initial insights for predictive modeling.

**Dataset**

* **Source:** IBM Telco Customer Churn Dataset
* **Records:** ~7,000 customers
* **Key features:**
  + Demographics (gender, senior citizen status)
  + Service usage (tenure, contract type, payment method)
  + Churn label (Yes/No)

**EDA Activities**

✅ **Calculated overall churn rate**

* Identified proportion of customers who have churned vs those who stayed.

✅ **Explored customer distribution by key demographics**

* Gender
* Senior citizen status

✅ **Analyzed tenure distribution**

* Visualized how tenure varies between churned and retained customers.

✅ **Investigated contract type and payment method relationships with churn**

* Checked how contract types and payment methods impact churn rates.

✅ **Generated correlation heatmap**

* Examined numeric features (e.g. tenure, charges) for interrelationships.

**Key Findings**

* The overall churn rate was approximately **26-27%**.
* **Month-to-month contract customers** showed significantly higher churn compared to those with longer contracts.
* Customers paying via **electronic check** exhibited higher churn rates.
* Churn was relatively similar across gender groups.
* Customers with lower tenure were more likely to churn.

**Visualizations**

(Include screenshots of your generated plots here in the report)

* Churn distribution bar chart
* Tenure histogram (churn vs non-churn)
* Contract type vs churn stacked bar
* Payment method vs churn bar
* Correlation heatmap

**Tools Used**

* Python (pandas, matplotlib, seaborn)
* VS Code / Google Colab

**Next Steps**

* Perform feature engineering (e.g. binning tenure, encoding categorical variables)
* Build predictive models (e.g. logistic regression, decision tree, random forest)
* Evaluate model performance using accuracy, precision, recall, F1 score

**Conclusion**

The EDA provided valuable insights into customer churn drivers. The identified patterns will guide the development of robust predictive models aimed at improving customer retention.

⚡ **Tip:**