# TELCO CUSTOMER CHURN ANALYSIS

Prepared for: Telco Communications Management Team

Presented by: Nicolette Woolery

## Project Summary

Customer churn is a critical concern for Telco, directly impacting revenue, customer retention, and growth strategy. Churn occurs when customers discontinue their subscription or contract, often due to competitive alternatives or dissatisfaction with service. The goal of this analysis is to uncover patterns in churn behavior, identify high-risk customer groups, and develop actionable solutions that support customer retention and loyalty.

This project focuses on:

- Creating 3 high-risk customer profiles linked to churn
- Analysing service scrutiny: identifying high-value customers using CLTV and monthly charges.

Insights were drawn using Python, data mining, and clustering analysis techniques.

## Objectives

- Investigate churn patterns using training/test datasets.
- Identify 3 profiles with high churn risk.
- Analyze the relationship between Monthly Charges and Customer Lifetime Value (CLTV).
- Propose strategic, data-driven recommendations for reducing churn.

## Methodology Overview

- Tools: Python 3.10.9, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn
- Datasets: Train/Test from Telco (CSV, Apache 2.0 License)
- Techniques:
- Data Cleaning & Encoding
- Correlation Analysis
- K-Means Clustering
- Feature Selection & Visualization

## Assumptions

- Focus is on early churn (tenure < 12 months)
- Customers are sensitive to pricing, service quality, and customer support.
- "High-value" is based on CLTV, Monthly Charges, and tenure.

## Problem Solving Process

Step 1: Set the Goal

I began by understanding Telco's biggest pain point — customer churn. From there, I defined clear objectives to guide the analysis.

Step 2: Get the Data Ready

I loaded the customer data from CSV files into Python using Pandas. Then cleaned things up — fixing missing values and making sure everything was consistent.

Step 3: Prepare the Data

To get the data model-ready, I encoded all the categories (like Yes/No responses) and made sure all numerical values were in the right format.

Step 4: Explore the Patterns

I used correlation analysis to identify relationships between features — for example, to see how strongly churn is associated with attributes like contract type, monthly charges, or tenure. This helped us narrow down which variables were most relevant for predicting customer behaviour. K-Means clustering was applied to group customers into meaningful segments based on their behaviors and characteristics. Since we didn't know exactly what kinds of customer profiles existed beforehand, this unsupervised learning technique allowed the data to reveal its own structure.

## Problem Solving Process continued

Step 5: Pinpoint What Matters

I identified the most influential features that signaled whether a customer was likely to churn — contract type, payment method, and more.

Step 6: Visualize the Insights

I created clear, compelling visuals to tell the story of churn — from heatmaps to cluster plots, making the findings easy to digest.

Step 7: Recommend Action

With the insights in hand, I developed practical, data-backed recommendations to help Telco retain high-value customers and reduce churn.

## Profile Creation - Data Preprocessing

- Removed irrelevant columns: CustomerID, City, etc.
- Retained Zip Code for geographic churn trends
- Encoded categorical features (Yes/No -> 1/0)
- Cleaned missing and inconsistent data

## Feature Selection & Clustering

- Features selected: Contract, Internet Service, Payment Method, Paperless Billing, etc.
- K-Means Clustering created 3 customer profiles
- Churn probability calculated per profile

## Customer Churn Profiles

Profile 1:

- No Contract, Internet Service = DSL, No Phone Service
- Medium risk (Attrition ~7%)

Profile 2:

- Paperless Billing = Yes
- Higher risk (Attrition ~8%)

Profile 3:

- Electronic Check, No Internet Service, Paperless Billing
- Highest risk (Attrition ~25%)

## Service Scrutivy - High-Value Customers

Goal: Identify customers with high CLTV

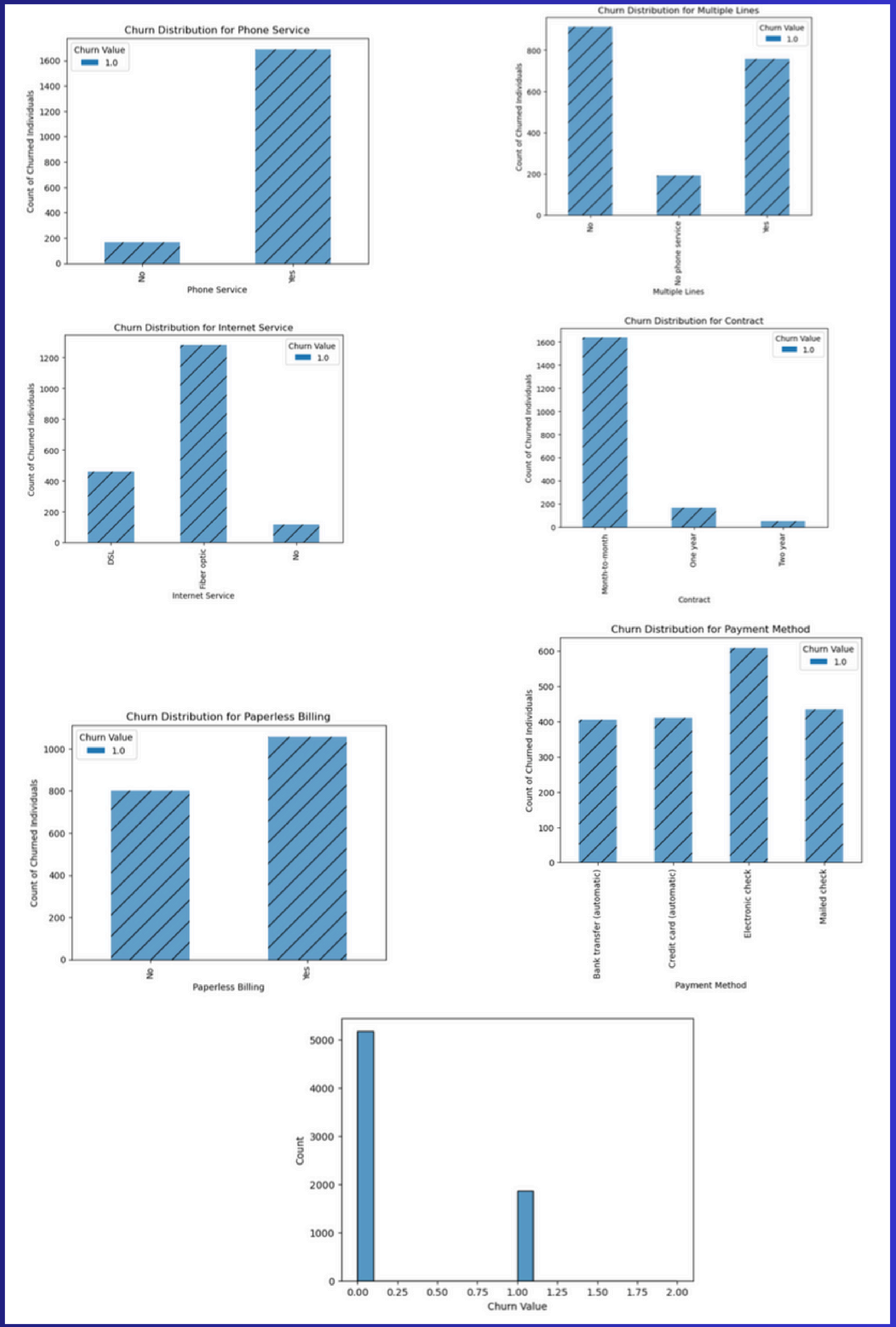
Analyzed correlation between Monthly Charges and CLTV (r = 0.0987)

Weak correlation -> high CLTV is not solely dependent on monthly charges

## Attritbutes of High-Value Targets

Attributes positively associated with CLTV:

- Monthly Charges
- Multiple Lines
- Paperless Billing
- Streaming Services
- Long Tenure



## Tools Evaluation

Python: Best suited for this task due to flexibility and community support

Pandas/NumPy: For data cleaning and transformation

Scikit-learn: Clustering & modeling

Seaborn/Matplotlib: Data visualization

## Recommendations

- Focus marketing on Profile 3 customers with targeted retention offers
- Introduce bundles or incentives for paperless billing users
- Offer security services to high CLTV customers with streaming services
- Consider follow-up with new customers during first 12 months

## Future Work

- Integrate sentiment analysis from customer reviews or support tickets
- Test alternative clustering methods (e.g. DBSCAN, hierarchical)
- Real-time churn prediction models with adaptive retraining
- Incorporate IoT or behavioral data for precision targeting

Thank you.