ANALYZING CUSTOMER CHURN PATTERNS IN A U.S. TELECOM FIRM USING SAS

1. Overview

Client:

A mid-sized U.S. telecom provider offering wireless, broadband, and VoIP services

Objective:

To identify key factors contributing to customer churn using SAS and provide data-driven recommendations to reduce churn and improve retention strategies.

2. Background

Churn directly affects telecom profitability. The client had a growing subscriber base but observed a sudden increase in monthly cancellations. While marketing and support teams had hypotheses, they lacked a statistically sound analysis to confirm churn predictors and segment at-risk customers.

3. Data Summary

Dataset:

Customer-level data of 10,000+ subscribers (12-month activity window)

Key Variables:

Variable	Type	Description
Churned (0/1)	Binary	Whether the customer left within the 12-month period
Tenure_Months	Continuous	Number of months since activation
Contract_Type	Categorical	Month-to-Month / One Year / Two Year
Monthly_Charges	Continuous	Monthly bill (USD)
Internet_Service_Type	Categorical	DSL / Fiber / None
Customer_Support_Calls	Integer	Number of customer support calls made
Payment_Method	Categorical	Credit Card / Bank Transfer / Mailed Check

Addon_Services_Count	Integer	Number of optional services subscribed

4. Methodology

Software Used:

SAS 9.4

SAS Workflow:

1. **Data Preparation:**
    - Imported .csv with PROC IMPORT
    - Categorical variables encoded using PROC FORMAT and dummy creation in DATA step
    - Checked and imputed missing values using PROC STDIZE
2. **Exploratory Data Analysis:**
    - PROC FREQ, PROC UNIVARIATE, and PROC MEANS
    - Churn rate by contract type and service package
    - Correlation matrix of numeric predictors
3. **Statistical Analysis:**
    - PROC LOGISTIC for churn prediction
    - Stepwise model selection with AIC
    - Odds ratio interpretation for business teams
    - Used PROC HPLOGISTIC to validate model on large subsets
4. **Data Visualization:**
    - Bar charts of churn vs. tenure
    - Heatmaps for service usage vs. churn
    - Lift charts to evaluate model performance

5. Key Results

Predictor	Odds Ratio	p-value	Interpretation
Month-to-Month Contract	3.7	<0.001	Highest churn risk (vs. 2-Year contract)
Customer_Support_Calls	1.4	<0.001	Each additional call increases churn odds
Fiber Internet	1.2	0.042	Fiber users more likely to churn than DSL users
Addon_Services_Count	0.85	<0.01	More addons → lower churn risk
Mailed_Check Payment	1.6	0.008	Traditional payment methods linked to higher churn

Model Accuracy:

- AUC = 0.87
- Sensitivity = 83%, Specificity = 78%
- Lift at decile 1 = 2.3× baseline churn rate

6. Visual Outputs (SAS):

- Lift chart showing model performance by decile
- Churn rate plotted by tenure buckets
- Mosaic plot of contract type vs. churn
- Coefficient plot with confidence intervals

7. Deliverables

- Clean, well-commented .sas code for all procedures
- Full analysis report (21 pages) with:
    - Business-friendly explanations of statistical findings
    - Data preparation pipeline and model diagnostics
    - Segmentation of high-risk customer profiles

- Executive presentation deck (5 slides):
    - Key churn drivers
    - Retention strategy recommendations
    - Projected retention improvement from model usage

8. Application & Outcome

- Client integrated churn prediction model into CRM
- Enabled proactive outreach to high-risk customers
- Reduced churn rate by **6.2%** in first two quarters post-implementation

9. Strategic Value Delivered

- Identified **contract structure and support interaction** as actionable churn levers
- Translated SAS output into **practical business interventions**
- Delivered a **repeatable analysis framework** for future churn studies