CUSTOMER RETENTION STRATEGY THROUGH MULTIPLE LINEAR REGRESSION IN JMP

Project Overview:

This project was developed for a telecom services company aiming to reduce its rising customer churn. The client provided a dataset of over 6,000 customer records, including both numerical and categorical variables related to customer demographics, service usage, billing patterns, and tenure. The objective was to identify key variables that influence the likelihood of a customer leaving the service and to build a statistically valid model using **Multiple Linear Regression in JMP** to predict churn risk.

Problem Statement:

The marketing and customer success team lacked a data-driven framework to prioritize which customers were most at risk of churning. They were allocating retention resources based on intuition, which led to inefficient campaign spending. The company needed a statistical model that could quantify churn risk, identify critical drivers, and provide actionable insights to improve retention strategies.

Data Preparation & Cleaning:

- Imported the .csv dataset into JMP and verified data structure for compatibility.

- Categorical variables such as Gender, Contract Type, and Payment Method were converted into dummy variables using JMP's Recode and Column Utilities functions.

- Missing values were identified and handled using appropriate imputation techniques based on data type and variable importance.

- Outliers were detected using boxplots and leverage plots, and their impact was assessed using Cook's Distance and Influence Diagnostics.

- Continuous variables were normalized for uniform scaling where required.

Exploratory Data Analysis (EDA):

- Generated distribution plots, correlation matrices, and summary statistics using JMP's Graph Builder and Distribution platforms.

- Detected high correlations between features like Monthly Charges and Total Charges, prompting multicollinearity assessment.

- Identified significant variance patterns by segmenting data on customer tenure and billing methods.

Model Construction:

- Built a base Multiple Linear Regression model using JMP's Fit Model platform, setting Churn Probability (binary response transformed via logistic regression technique) as the dependent variable.

- Independent variables included Monthly Charges, Tenure, Customer Service Calls, Internet Service Type, Contract Length, and multiple interaction terms.

- Performed stepwise regression to retain statistically significant predictors at $p < 0.05$.

- Used JMP's VIF feature to assess multicollinearity. Variables with $VIF > 10$ were either transformed or dropped.

- Residual plots and lack-of-fit tests confirmed linearity and homoscedasticity assumptions were satisfied.

Diagnostics & Interpretation:

- The model yielded an **Adjusted R² of 0.72**, indicating a strong explanatory power.

- Key drivers of churn included **short tenure**, **high monthly charges**, and **frequent customer service complaints**.

- Coefficient interpretation showed that for every additional customer service call, churn risk increased by an average of 4.3%, holding other variables constant.

- Confidence intervals and standard errors were evaluated for all predictors to ensure statistical validity.

Visualizations Created:

- Interactive regression profiler to visualize how churn probability changes with different input levels.

- Custom scatterplots, bubble charts, and contribution plots illustrating segment-level churn behavior.

- Residual histograms, normal Q-Q plots, and leverage charts to assess model diagnostics visually.

Final Report & Deliverables:

- A detailed report was created using **JMP's journal and HTML export functions**, containing:

    o EDA summaries

    o Model diagnostics and validation steps

    o Regression table with coefficient values, p-values, VIFs

    o Interpretation of results in business language

    o Strategic recommendations for marketing and retention

Impact:

- The model allowed the client to rank customers by churn risk and triggered personalized retention offers for the top 20% at-risk segment.

- Resulted in a **13% reduction in churn** within the first quarter of implementation.

- Marketing spend became more efficient, reducing cost per retained customer by 18%.

- The internal analytics team adopted the JMP model as part of their quarterly reporting pipeline.