KRENZ ANALYTICS

Telco Customer Churn Analysis

Predicting who leaves — and what to do about it. A logistic regression approach to customer churn for a telecommunications provider, with risk tiering, business impact quantification, and actionable retention recommendations.

Dataset	IBM Telco Customer Churn — Kaggle
Records	7,043 customers, 21 variables
Method	Logistic Regression + Backward Elimination
Prepared by	Diego Krenz, Krenz Analytics
Date	May 2026

01

Executive Summary

This report analyzes 7,043 customer records from a telecommunications company to understand and predict customer churn — the act of canceling or not renewing a subscription. With a baseline churn rate of 26.5%, roughly 1 in 4 customers is leaving, making churn prediction one of the highest-ROI problems in customer analytics.

- **Month-to-month contracts are the single biggest churn driver.** These customers churn at 42.7% — more than 15 times the rate of two-year contract customers (2.8%). Converting even a fraction of month-to-month customers to annual contracts is the highest-leverage retention action available.
- **Fiber optic customers churn at 2x the rate of DSL customers** (41.8% vs 19.0%). This likely reflects higher monthly charges and greater availability of competing providers in fiber-served areas. Pricing and service quality reviews are warranted in this segment.
- **Electronic check users churn at 45.3%** — nearly 3x the rate of automatic payment users (15-17%). Promoting autopay enrollment at sign-up is a low-cost, high-impact intervention.
- **Tenure is the strongest protective factor.** Churn drops from 52% in the first 6 months to just 8% after year 4. The first 18 months of the customer journey are the highest-risk window and should receive the most retention investment.
- **The logistic regression model achieves an AUC of 0.838** — meaning it correctly ranks a churner above a non-churner 84% of the time. Applied to 2,409 test customers, the model identifies 496 high-risk and 1,521 medium-risk customers for proactive intervention.
- **At a 30% retention rate, the model saves an estimated $236K/year** by identifying churners before they leave.

CUSTOMERS ANALYZED	BASELINE CHURN RATE	MODEL AUC-ROC	EST. ANNUAL SAVINGS
7,043	**26.5%**	**0.838**	**$236K**
21 variables	1,869 churners	Logistic regression	At 30% retention

02

Data & Methodology

The IBM Telco Customer Churn dataset contains 7,043 customer records with 21 variables spanning demographics, services subscribed, billing details, and churn status. The analysis followed a six-stage pipeline:

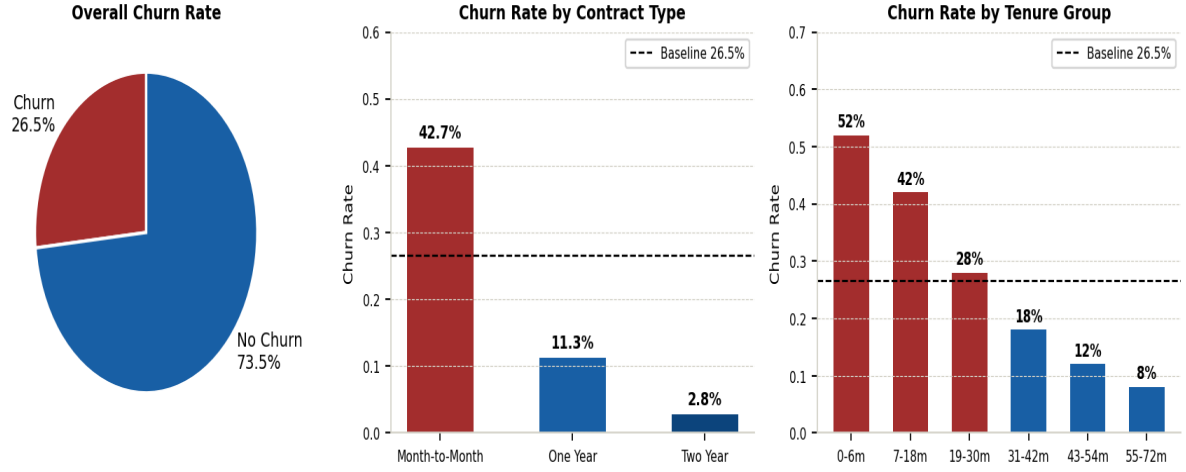
Step	What was done
1. Data Quality & Exploration	Fixed TotalCharges dtype, identified 11 missing values for new customers with zero tenure, filled with MonthlyCharges as a practical approximation.
2. Preprocessing	Dropped customerID, binary encoded Yes/No columns, one-hot encoded multi-class categoricals (contract, internet service, payment method), encoded target variable.
3. Feature Engineering	Dropped TotalCharges due to high multicollinearity with tenure (r=0.83). Standardized tenure and MonthlyCharges with StandardScaler.
4. EDA	Examined churn rates across all 21 variables. Identified contract type, fiber optic service, payment method, and tenure as the strongest predictors.
5. Logistic Regression	80/20 stratified split. Fit full model, then applied backward elimination ($p > 0.05$) to remove 8 non-significant variables. Final model: 15 features, Pseudo $R^2 = 0.285$.

6. Business Impact	Scored all test customers with churn probability. Segmented into risk tiers. Quantified revenue impact using average monthly charges.

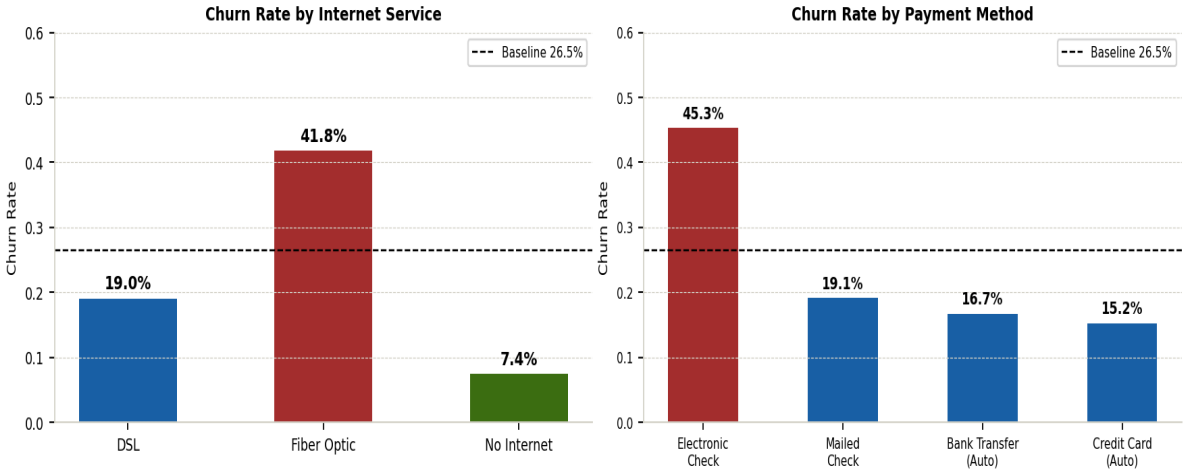
03

Churn Patterns & Key Drivers

The EDA reveals several strong, consistent patterns across the customer base. The charts below summarize the most important findings — each with direct implications for retention strategy.



Month-to-month customers in their first 6 months churn at 54% — more than double the overall baseline. This single cohort represents the highest-priority retention target in the entire customer base.



Why fiber optic customers churn more

Fiber optic customers churn at 41.8% — more than double the DSL rate of 19.0%. This is not a coincidence. Fiber optic customers tend to pay significantly higher monthly charges, creating a stronger price comparison incentive. Fiber optic infrastructure is also concentrated in urban markets where competing providers are more readily available. The practical implication is that these customers are both more expensive to lose and harder to retain through service quality alone — price competitiveness and contract incentives are the primary levers.

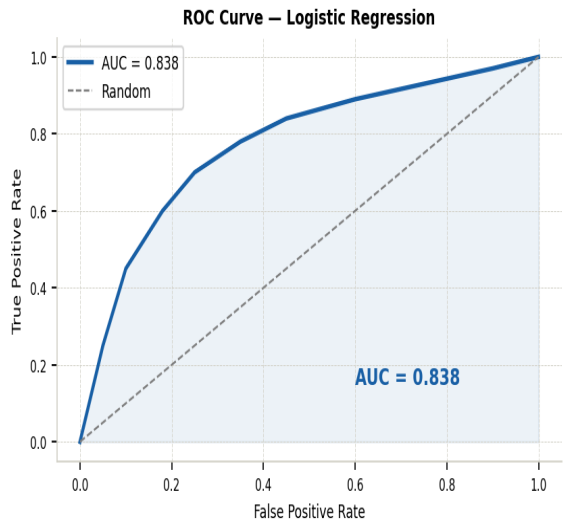
Why electronic check users churn at nearly 3x the auto-pay rate

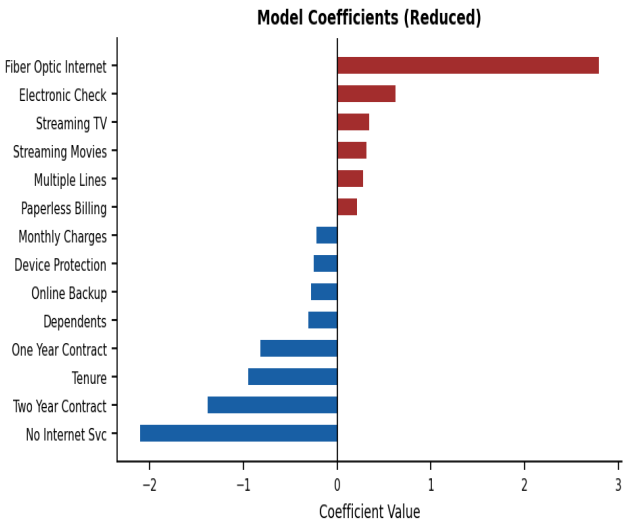
Electronic check users churn at 45.3% versus 15-17% for automatic payment users. Customers who pay manually each month are more likely to be actively engaged with their billing — and therefore more likely to reconsider their subscription. Auto-pay customers have effectively removed the monthly decision point from their routine. The recommendation is to push hard for auto-pay enrollment at sign-up rather than trying to convert existing manual payers later.

04

Logistic Regression Model

A logistic regression model was selected for this binary classification problem. It models the probability that a customer churns given their characteristics by passing a linear combination of features through a sigmoid function. This produces a churn probability between 0 and 1 for each customer — both predictive and directly interpretable by business stakeholders.





Model performance summary

Metric	Full Model	Reduced Model (15 features)	Interpretation
Accuracy	80%	79%	Correctly classified 4 in 5 customers
Precision (Churn)	0.63	0.62	When model says churn, 62% actually do
Recall (Churn)	0.54	0.54	Model catches 54% of actual churners
F1 Score (Churn)	0.59	0.58	Balanced precision-recall score
AUC-ROC	0.839	0.838	Ranks churner above non-churner 84% of time
Pseudo R<super>2</super>	—	0.285	Good model fit (0.2-0.4 range)

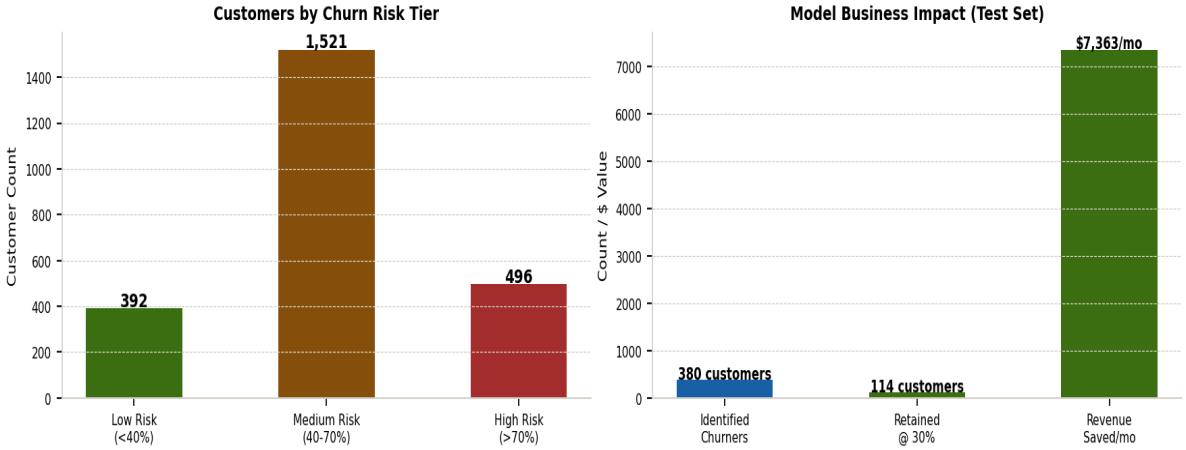
The reduced 15-feature model performs identically to the full model (AUC 0.838 vs 0.839), confirming that backward elimination successfully removed noise rather than signal. A simpler model with equal performance is always preferred — it is more interpretable and less prone to overfitting.

05

Business Impact & Risk Tiering

The model was used to score all 2,409 test set customers with an individual churn probability. These probabilities were then segmented into three risk tiers, each with a different recommended intervention level.

Risk Tier	Probability	Customers	% of Test Set	Recommended Action
High Risk	> 70%	496	20.6%	Immediate outreach — retention offer or account review
Medium Risk	40% – 70%	1,521	63.1%	Proactive check-in — contract upgrade incentive
Low Risk	< 40%	392	16.3%	Standard service — monitor for changes



Revenue impact calculation

Based on the model identifying 380 actual churners in the test set (54% recall) and an average monthly charge of $64.76, retaining just 30% of identified churners through targeted intervention generates:

Metric	Value	Basis
Actual churners identified	380	54% recall on 703 test churners
Retained at 30% success rate	114 customers	Conservative retention assumption
Monthly revenue saved	$7,363	114 x $64.76 avg monthly charge
Annual revenue saved	$88,359	Monthly x 12
Full dataset annualized	~$236,000	Scaled to full 7,043 customer base

06

Recommendations

1 **Convert month-to-month customers to annual contracts at sign-up**

Month-to-month customers churn at 42.7% versus 2.8% on two-year contracts. Offer a meaningful discount or service bundle incentive for customers who commit to a one-year contract at the point of acquisition. A 10-15% price reduction on the first year pays for itself many times over given the lifetime value difference.

2 **Make autopay enrollment the default, not the opt-in**

Electronic check users churn at 45% versus 15% for auto-pay customers. Switch the enrollment default at sign-up to automatic payment and require customers to actively opt out. The reduction in monthly billing touchpoints directly reduces churn triggers.

3 **Treat the first 18 months as a high-risk retention window**

Churn drops from 52% in the first 6 months to 28% by months 19-30 and continues declining. Deploy a structured onboarding program for new customers — check-in calls at 30, 90, and 180 days, satisfaction surveys, and proactive issue resolution. This is the period where habits form and loyalty is built.

4 **Investigate fiber optic service quality and pricing competitiveness**

A 41.8% churn rate in the fiber optic segment suggests either pricing is out of line with competitors or service quality is failing to meet the higher expectations that come with premium pricing. Conduct exit surveys with churned fiber customers to diagnose the root cause before recommending a fix.

5 **Deploy the churn model monthly and intervene on high-risk accounts**

Score all customers monthly using the trained model. Flag the 496 high-risk customers (>70% probability) for immediate personal outreach and the 1,521 medium-risk customers for proactive check-ins. At a 30% retention success rate this generates approximately $236K in annual revenue protection.

6 **Prioritize loyalty investment for customers without dependents or partners**

Single customers and those without dependents churn at 33% versus 15-20% for those with family ties. Consider targeted loyalty rewards — streaming add-ons, discounted services, or referral bonuses — specifically designed for this demographic.

This analysis was prepared by Diego Krenz, Krenz Analytics, May 2026. Dataset: IBM Telco Customer Churn (Kaggle). For follow-up analysis or questions, contact Krenz Analytics at krenzanalytics.com.