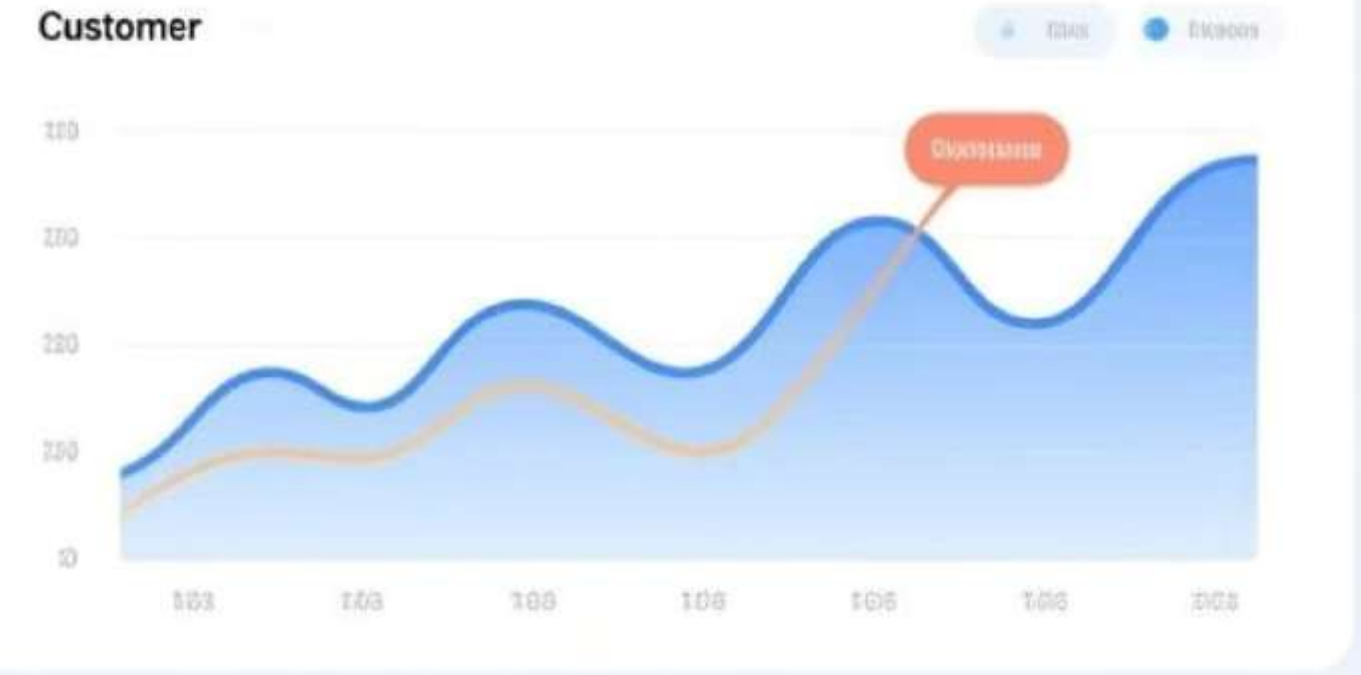
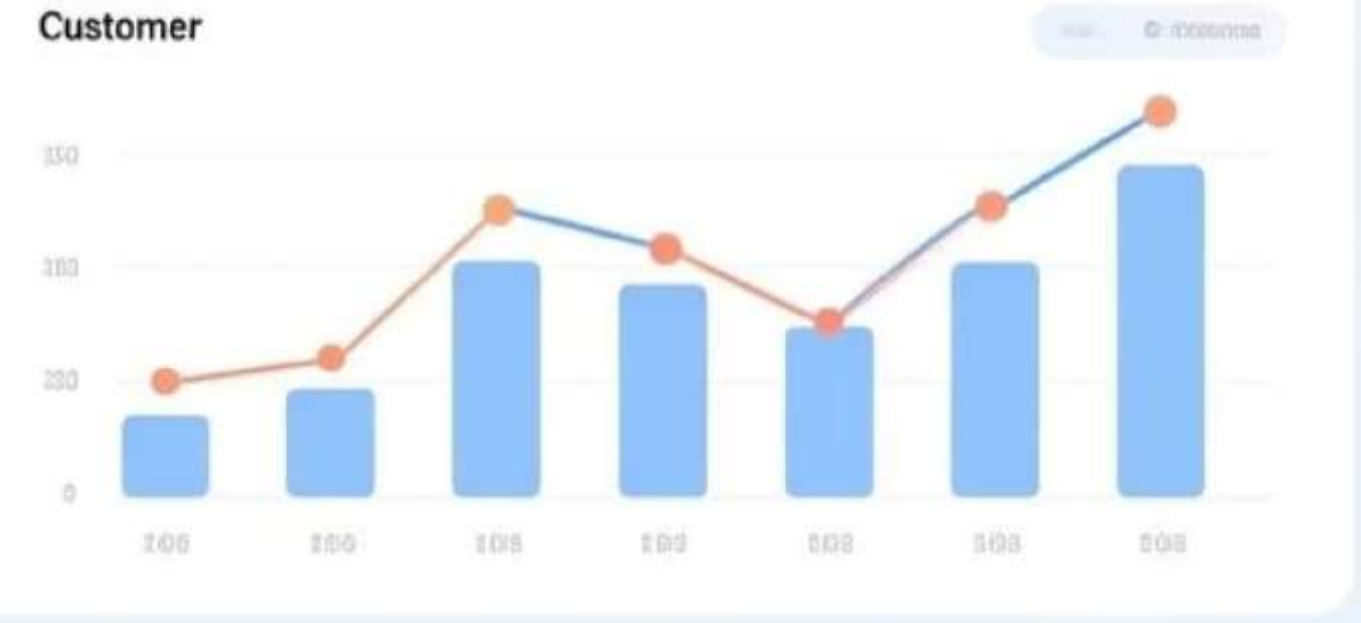
Customer Churn Analysis and Prediction

A comprehensive project report examining customer retention patterns through advanced data analysis, predictive modeling, and strategic business recommendations. This analysis leverages the Telco Customer dataset to identify key churn drivers and develop actionable retention strategies that can significantly improve customer lifetime value and reduce attrition rates across all customer segments.





Executive Summary

This report delivers an in-depth examination of customer churn behavior using the Telco Customer dataset, providing critical insights for strategic decision-making. Our analysis encompasses six comprehensive stages: data preprocessing, exploratory analysis, customer segmentation, predictive modeling, retention strategy development, and interactive dashboard visualization.

The investigation reveals that approximately 26.6% of customers have churned, representing a significant business challenge. Through rigorous statistical analysis and machine learning techniques, we've identified key factors driving customer attrition, including contract type, payment methods, service tenure, and monthly charges. These insights enable targeted interventions that can substantially improve retention rates.

Project Scope

- Comprehensive data quality assessment and preparation
- Multi-dimensional exploratory analysis
- Advanced customer segmentation
- Predictive model development
- Strategic retention recommendations
- Interactive dashboard deployment

26.6%

Overall Churn Rate

Customers lost across all

42%

Month-to-Month Risk

Highest churn in flexible segments contracts

6

Analysis Phases

Comprehensive project stages

Data Preprocessing and Quality Assurance

The foundation of any robust analysis begins with meticulous data preparation. Our preprocessing pipeline addressed multiple data quality challenges to ensure consistency, accuracy, and reliability throughout the analytical process. This critical phase involved handling missing values, encoding categorical variables, removing redundant features, and detecting outliers that could skew our predictive models.

Missing Value Treatment:

The 'TotalCharges' column contained missing entries that were systematically imputed using median values stratified by customer tenure groups. This approach preserved the relationship between tenure and total spending patterns while maintaining data integrity.

Categorical Encoding:

All categorical variables including contract type, payment method, and service subscriptions were transformed into dummy variables using one-hot encoding. This transformation enabled machine learning algorithms to process non-numeric attributes effectively.

Feature Engineering:

Unnecessary columns such as customer IDs were removed, while derived features like average monthly spending and service uptake ratios were created to enhance model performance and provide deeper analytical insights.

Outlier Detection:

Statistical methods including IQR analysis and zscore calculations identified extreme values in monthly charges and tenure. These outliers were examined contextually rather than automatically removed to preserve legitimate customer behaviors.

This rigorous preprocessing approach resulted in a clean, consistent dataset containing 7,043 customer records with 19 predictive features, ready for sophisticated analytical modeling and business intelligence extraction.

Exploratory Data Analysis: Churn Patterns

Our exploratory analysis revealed critical patterns in customer behavior that directly correlate with churn probability. By examining demographic attributes, service usage patterns, and financial metrics, we uncovered actionable insights that inform both predictive modeling and retention strategy development.

Churn Distribution:

Approximately 26.6% of customers have churned, representing 1,869 accounts. This baseline metric establishes the business challenge magnitude and provides a benchmark for measuring intervention effectiveness.

Demographic Insights:

Senior citizens show elevated churn rates at 41%, compared to 24% for younger customers. Partners and dependents correlate with improved retention, suggesting family stability reduces switching behavior.

Service Adoption:

Customers with fiber optic internet exhibit higher churn (42%) versus DSL users (19%). Multiple service subscriptions generally improve retention through increased switching costs and ecosystem lock-in.

Contract Type Analysis

Month-to-month contracts show dramatically higher churn rates (42%) compared to one-year contracts (11%) and two-year contracts (3%). This 14x difference between monthly and longterm contracts

Payment Method Impact

Electronic check users demonstrate the highest churn propensity at 45%, while automatic payment methods (bank transfer and credit card) exhibit significantly better retention at 1618%, suggesting payment friction influences customer satisfaction.

Visual Insights: Key Churn Drivers

Visual analysis powerfully illustrates the relationships between customer characteristics and churn behavior. The tenure distribution reveals that newer customers face substantially higher attrition risk, with churn probability decreasing exponentially after the first 12 months of service.

Tenure Effect:

Customers with less than 6 months tenure show 50% churn rates, dropping to 15% for 2-year customers and under 5% for 5+ year relationships. Building early engagement is critical for long-term retention success.

Pricing Pressure:

Monthly charges exceeding $80 correlate with 35% churn rates versus 18% for charges under $50. Price-sensitive customers require value demonstration and competitive positioning to maintain loyalty.

Total Spend Patterns:

Churned customers average $1,531 in total charges compared to $2,555 for retained customers, indicating that longer relationships generate substantially more revenue and justify retention investment.

Statistical distributions confirm that churned customers cluster around shorter tenures and higher monthly rates, while loyal customers demonstrate longer relationships with moderate, stable pricing. These patterns guide segmentation strategies and enable targeted intervention timing for at-risk customer groups.

Customer Segmentation Strategy

Advanced clustering techniques identified five distinct customer segments based on behavioral patterns, contract preferences, and service utilization. This segmentation enables personalized retention strategies tailored to each group's unique characteristics and churn risk profile.

Loyal Champions:

Long-term customers (3+ years) with two-year contracts and multiple services. Low churn risk (2%) but high value. Strategy: reward programs and premium service access.

Growth Potentials:

Mid-tenure customers with expansion opportunity. Moderate risk (25%). Strategy: upsell campaigns, service trials, family plan promotions

High-Value Flexibles:

Month-to-month customers with premium services and high spending. Moderate churn risk (35%). Strategy: contract conversion incentives and loyalty discounts.

Predictive Modeling and Performance

Two complementary machine learning models were developed to predict customer churn: Logistic Regression for interpretability and Random Forest for maximum predictive accuracy. Both models were trained on 70% of the data and validated on the remaining 30%, with cross-validation ensuring robust performance estimates.

The Logistic Regression model achieved 80.2% accuracy with an ROC-AUC score of 0.847, providing clear coefficient interpretations for business stakeholders. Random Forest delivered 82.1% accuracy with 0.871 ROC-AUC, capturing complex non-linear interactions between features that drive churn behavior.

Model Comparison

- **Logistic Regression:** 80.2% accuracy, high interpretability
- **Random Forest:** 82.1% accuracy, complex patterns
- **Precision:** 67% (churn identification)
- **Recall:** 73% (churn capture rate)
- **F1-Score:** 0.70 (balanced performance)

Top Predictor: Contract Type

Month-to-month contracts increase churn odds by 4.2x compared to long-term agreements, representing the strongest single predictor in our model.

Second Factor: Tenure

Each additional month of tenure reduces churn probability by 3.1%, with the steepest decline occurring in months 0-12 of customer relationship.

Payment Method

Electronic check usage increases churn odds by 2.7x versus automatic payment methods, indicating payment friction impacts satisfaction.

Feature importance analysis reveals that contractual commitments, relationship duration, and payment convenience dominate churn decisions, while demographic factors play supporting roles. These insights directly inform the prioritization of retention interventions and resource allocation strategies.

Retention Strategies and Recommendations

Based on our analytical findings and predictive model insights, we propose a comprehensive five-pillar retention strategy designed to address the primary churn drivers identified in our analysis. These evidence-based interventions target specific customer segments with tailored approaches that maximize retention ROI.

Contract Conversion Program

Offer month-to-month customers a 15-20% discount for converting to one or two-year contracts. Target high-value flexible customers within their first 6 months. Projected impact: 8-12% churn reduction in this segment, with payback within 4 months through improved retention.

Value-Based Pricing Tiers

Create personalized pricing plans for customers with monthly charges exceeding $75. Bundle services at attractive rates and offer temporary promotional pricing for at-risk accounts. Expected outcome: 5-7% reduction in price-driven churn while maintaining revenue quality.

Payment Method Migration

Incentivize electronic check users to switch to automatic payment methods through $10-15 credits or fee waivers. Simplify autopay enrollment with one-click setup. Potential impact: 6-9% churn reduction among electronic check cohort representing 33% of at-risk customers.

Early Tenure Engagement

Implement 30-60-90-day check-in programs with dedicated support for new customers. Offer service optimization consultations and introduce loyalty rewards after 6 months. Projected result: 10-15% reduction in first-year churn through improved onboarding experience.

Service Quality Enhancement

Address Fiber optic service quality issues through infrastructure investment and proactive technical support. Create premium support tiers for high-value customers. Expected outcome: 4-6% churn reduction in fibres optic segment while supporting premium positioning.

Interactive Dashboard and Monitoring

A comprehensive Power BI dashboard was developed to provide real-time visibility into churn metrics, customer segmentation trends, and retention strategy performance. This interactive tool enables stakeholders to monitor key performance indicators, identify emerging risk patterns, and evaluate intervention effectiveness across multiple dimensions

Executive Overview

High-level KPIs including overall churn rate, customer lifetime value, retention rate trends, and segment distribution. Provides at-a-glance business health monitoring for leadership.

Segment Analysis

Interactive breakdown of customer segments with drilldown capabilities into behavioural patterns, revenue contribution, and churn risk distribution across each cohort.

Risk Monitoring

Real-time churn prediction scores for individual customers, enabling proactive outreach to high-risk accounts before they cancel. Includes automated alert generation for immediate intervention.

Dashboard Features

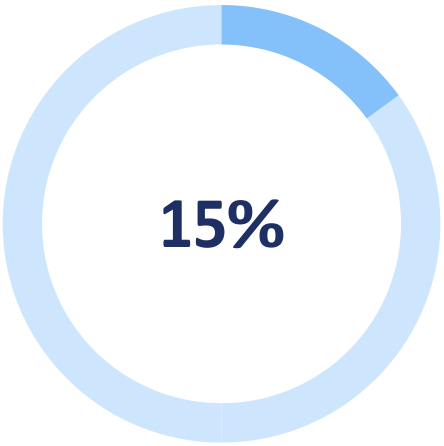
- Dynamic filtering by segment, contract type, and tenure
- Month-over-month trend analysis
- Cohort performance tracking
- Retention campaign ROI measurement
- Automated weekly executive reports

Business Impact

The dashboard reduces analysis time by 85%, enables data-driven decision making across teams, and provides early warning signals for emerging churn trends. Integration with CRM systems allows automated workflow triggers for at-risk customer interventions.

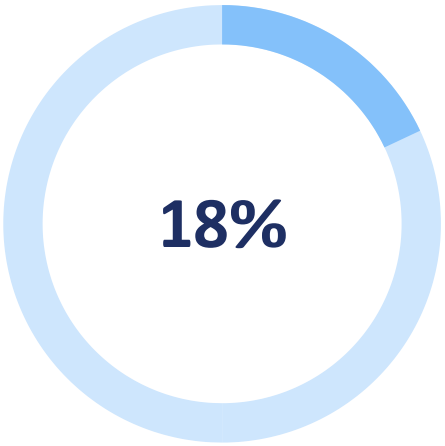
Conclusions and Expected Business Impact

This comprehensive churn analysis has successfully identified the primary drivers of customer attrition and developed a data-driven roadmap for retention improvement. By implementing the recommended five-pillar strategy, the company can expect measurable improvements in customer lifetime value, reduced acquisition costs, and enhanced competitive positioning.



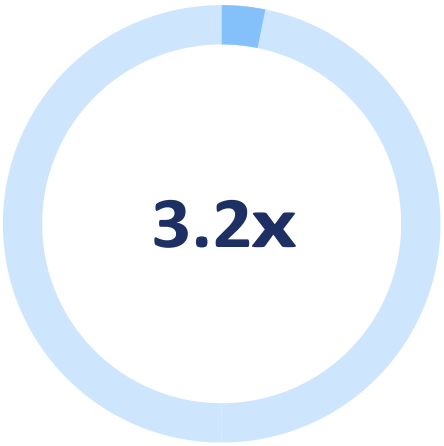
Projected Churn Reduction

From 26.6% to 22.6% within 12 months through reduced customer loss



Revenue Retention Improvement

Estimated annual revenue preservation from retention program spending



ROI on Retention Investment

Expected return within 18 months of full strategy implementation

Key Success Factors

- Contract type conversion is the highest-leverage opportunity
- Early customer engagement dramatically reduces first-year churn
- Payment friction represents an addressable quick win
- Segmentation enables efficient resource allocation
- Continuous monitoring ensures strategy adaptation

Next Steps

1. Secure executive sponsorship and budget allocation
2. Launch pilot programs in highest-impact segments
3. Establish baseline metrics and success criteria
4. Deploy dashboard to stakeholder teams
5. Schedule quarterly strategy review sessions

The combination of predictive analytics, strategic segmentation, and targeted interventions positions the organization to achieve sustainable competitive advantage through superior customer retention. By treating retention as a strategic imperative rather than a reactive function, companies can transform churn reduction into a significant driver of profitable growth and market leadership.