



From Insight to Action: A Data-Driven Retention Strategy to Predict Customer Churn

SyriaTel: Telecom Customer Churn Prediction

Presented by: Samwel Ongechi

Date: Sunday, July 20, 2025

Technical Mentor: George Kamundia



Introduction



The telecommunications industry is characterized by intense competition and rapid technological advancements. In such an environment, the cost of acquiring new customers continues to rise, making customer retention not just an operational goal but a strategic imperative for long-term profitability and market leadership. Our ability to keep existing customers directly impacts our revenue stability and growth potential.

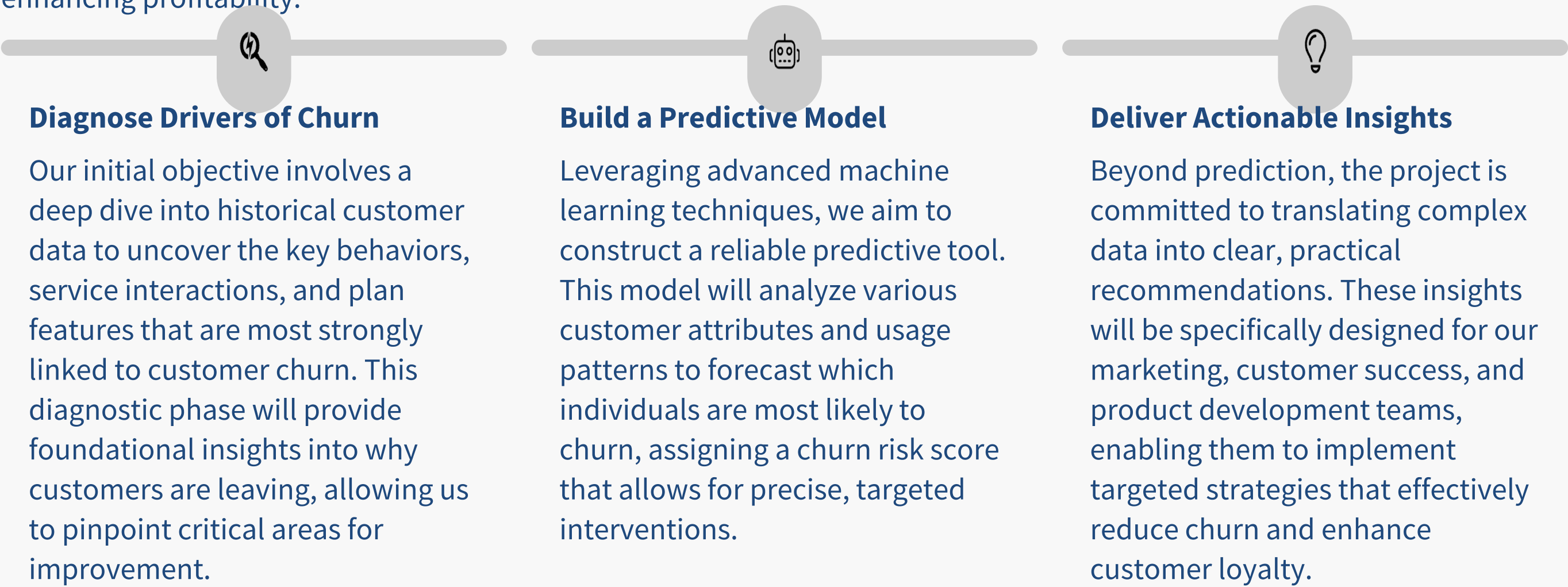
Acquiring a new customer is significantly more expensive than retaining an existing one, often costing **five to ten times more**. This disparity underscores the immense value of every loyal customer. By focusing on retention, SyriaTel can optimize its marketing spend, reduce operational overhead associated with customer onboarding, and cultivate a stable customer base that generates consistent revenue.

Understanding why and when customers leave (churn) is therefore a top priority. Churn erodes market share, diminishes brand value, and ultimately impacts the bottom line. By proactively addressing the root causes of churn, we can protect our existing revenue streams and strengthen our competitive position in the Syrian telecom market. This strategic focus on retention is essential for sustainable growth.



From Reactive to Predictive Retention

This project marks a pivotal shift in SyriaTel's customer retention strategy, moving from a reactive approach—where we respond to churn after it occurs—to a proactive, data-driven model that identifies at-risk customers *before* they decide to leave. Our goal is to empower SyriaTel with the foresight to intervene effectively, safeguarding our customer base and enhancing profitability.



By achieving these objectives, SyriaTel will be able to implement more effective retention campaigns, personalize customer interactions, and ultimately foster stronger, more lasting relationships with our subscribers.



Dataset Overview: The Foundation of Our Insights

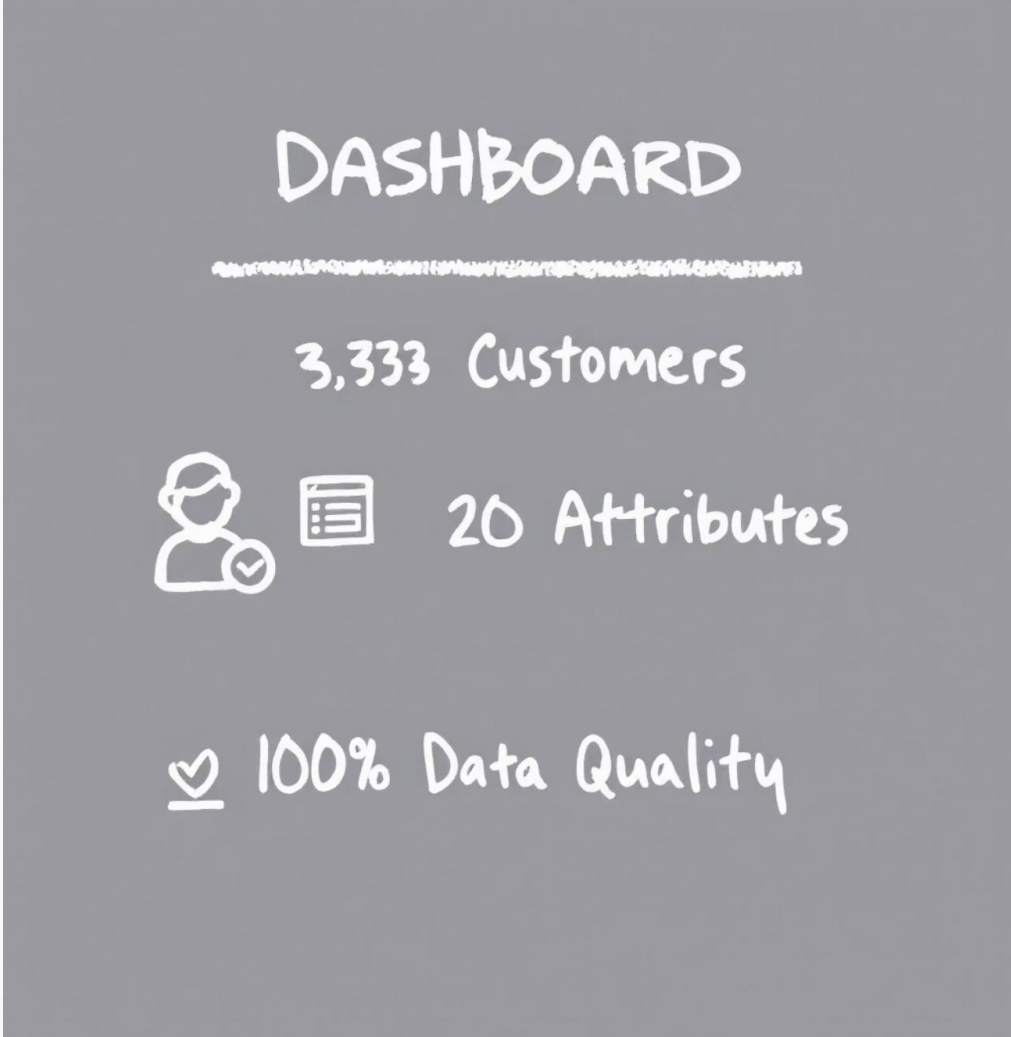
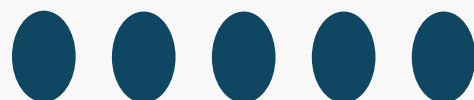
Our comprehensive analysis was built upon a meticulously curated dataset comprising detailed information for **3,333 SyriaTel customers**. This robust dataset served as the bedrock for our predictive modeling, offering a holistic view of customer interactions and service usage.

The dataset encompassed a wide array of attributes, providing rich context for our analysis. Key categories of information included:

- **Customer Account Information:** Details such as account length, contract type, and payment methods.
- **Plan Details:** Information on subscribed plans, including international plans, voicemail plans, and data bundles.
- **Usage Patterns:** Granular data on call minutes (day, evening, night, international), number of calls, and customer service call history.

Crucially, the dataset exhibited **100% data quality** with no missing values, which significantly streamlined the data preparation phase and ensured the reliability of our findings. This impeccable data integrity provided a solid and trustworthy foundation for all subsequent analytical and modeling efforts.

The core focus of our analysis was the **"churn" variable**. This binary indicator clearly identified whether a customer had discontinued their service with SyriaTel or remained a loyal subscriber. This target variable enabled us to build a model specifically designed to predict customer attrition.



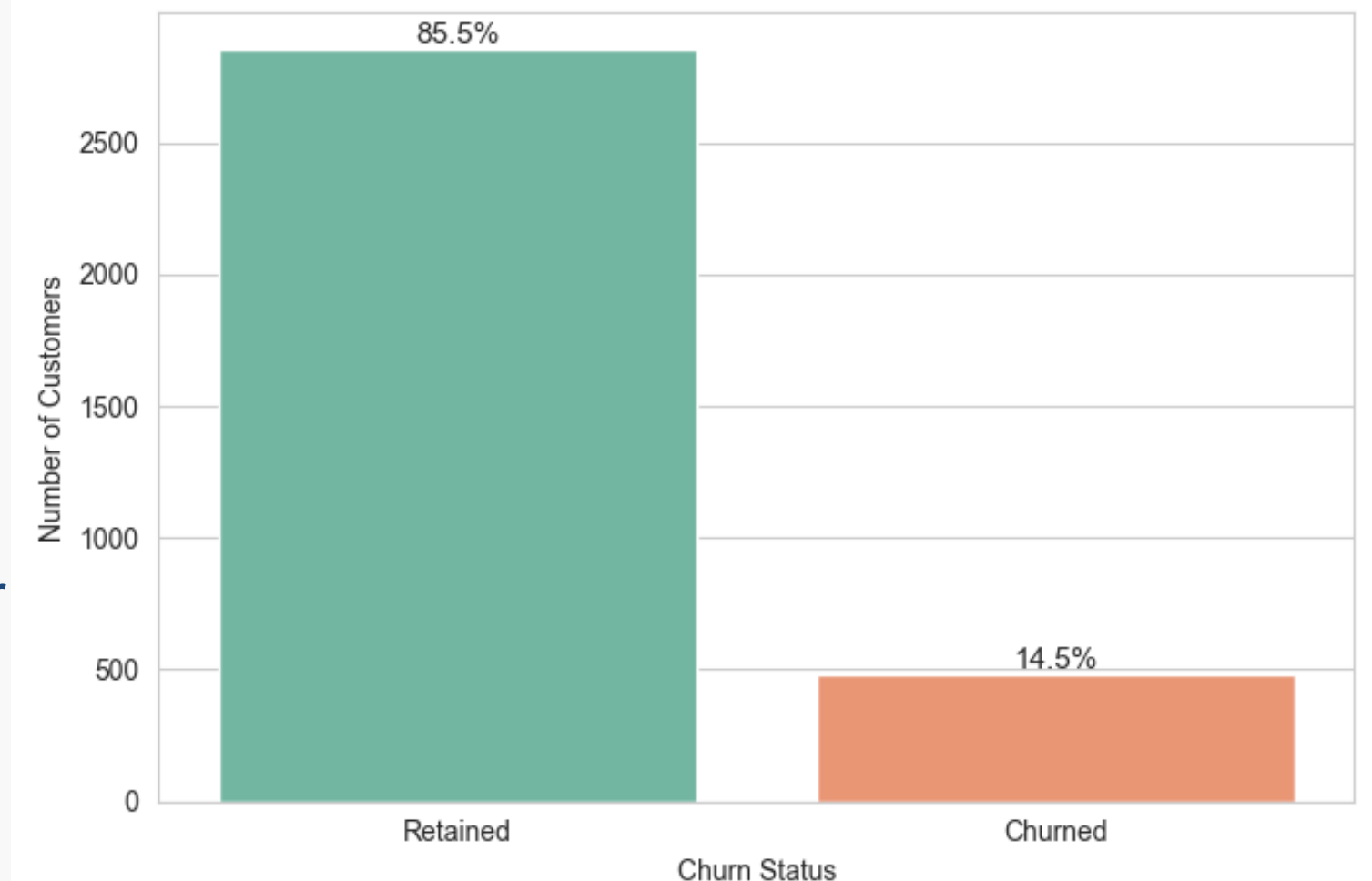
Churn is a Significant, but Minority, Event

Our initial exploratory data analysis revealed a crucial insight into SyriaTel's customer base: approximately **14.5% of customers have churned** from our services. While the vast majority—**85.5%**—of our customers demonstrate loyalty and continue their subscriptions, this churn rate, though seemingly a minority, represents a substantial loss in potential revenue and market share.

The implications of this finding are significant. Although churners constitute a smaller segment of our total customer base, their cumulative impact on our bottom line is considerable. Protecting this segment is paramount for sustainable business growth. Furthermore, the minority nature of the churn group presents a unique challenge for our predictive model: it must be exceptionally effective at identifying this "needle in the haystack" to ensure our retention efforts are precisely targeted and resource-efficient. The model's ability to accurately pinpoint these at-risk customers will be key to successful intervention strategies.



Overall Customer Churn Distribution



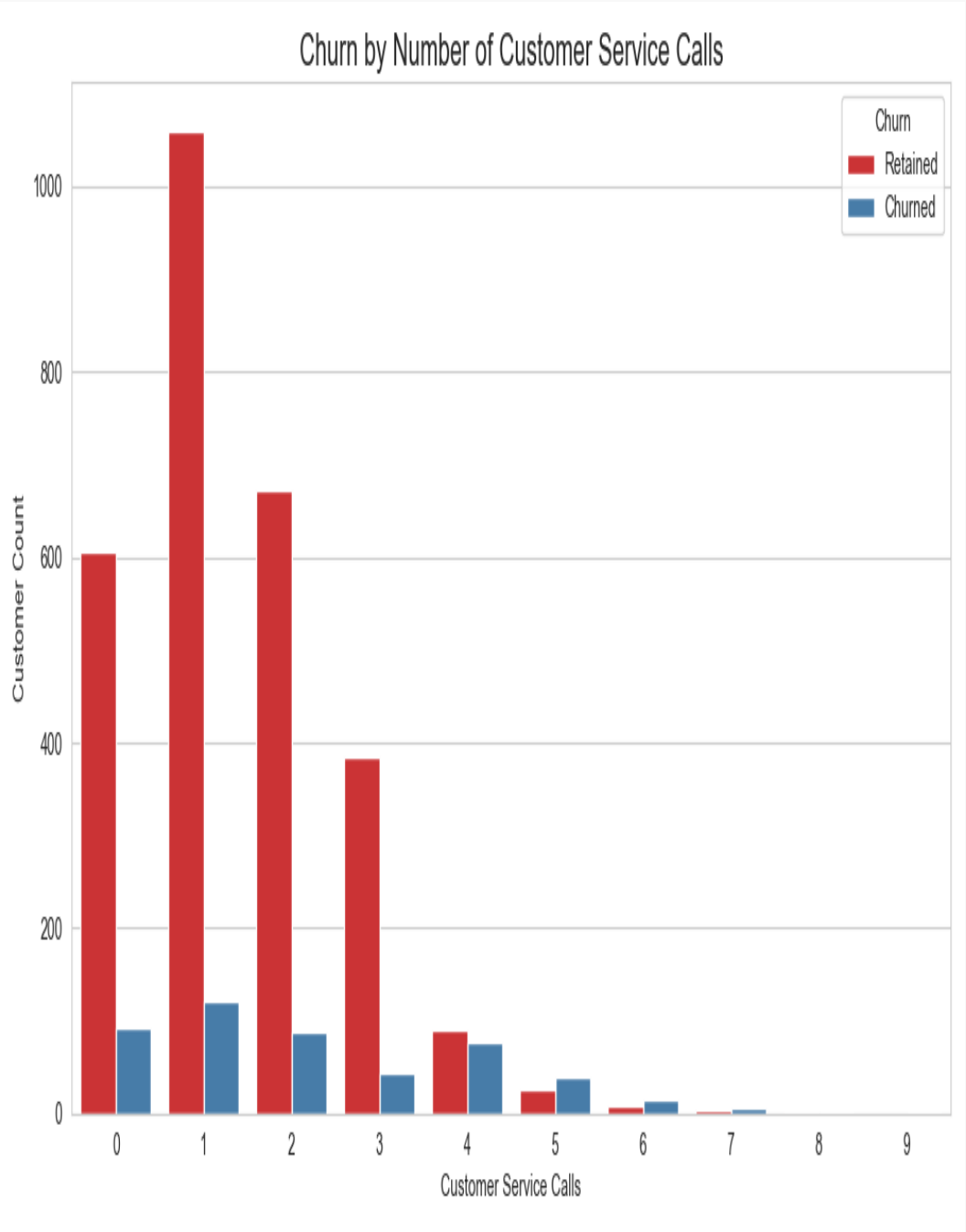
More Calls, Higher Churn Risk



One of the most compelling insights derived from our analysis is the strong correlation between the number of customer service calls and the likelihood of churn. Customers who frequently contact our support channels are inadvertently signaling their increasing dissatisfaction or unresolved issues, acting as a critical early warning system for potential attrition.

Our data distinctly shows that the churn rate **skyrockets for customers who make four or more service calls**. This threshold represents a critical point where frustration levels appear to escalate significantly, making customers highly susceptible to churning. The more interactions a customer has with support, especially without a satisfactory resolution, the higher their propensity to leave.

This finding provides a powerful **actionable insight**: the number of support calls can serve as a potent early warning sign. By proactively monitoring call volumes, SyriaTel can intervene with targeted support or personalized offers before a customer's cumulative frustration reaches the point of no return. This proactive engagement can significantly mitigate churn risks among this vulnerable segment of our customer base.



The International Plan Paradox



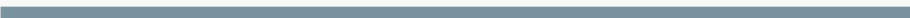
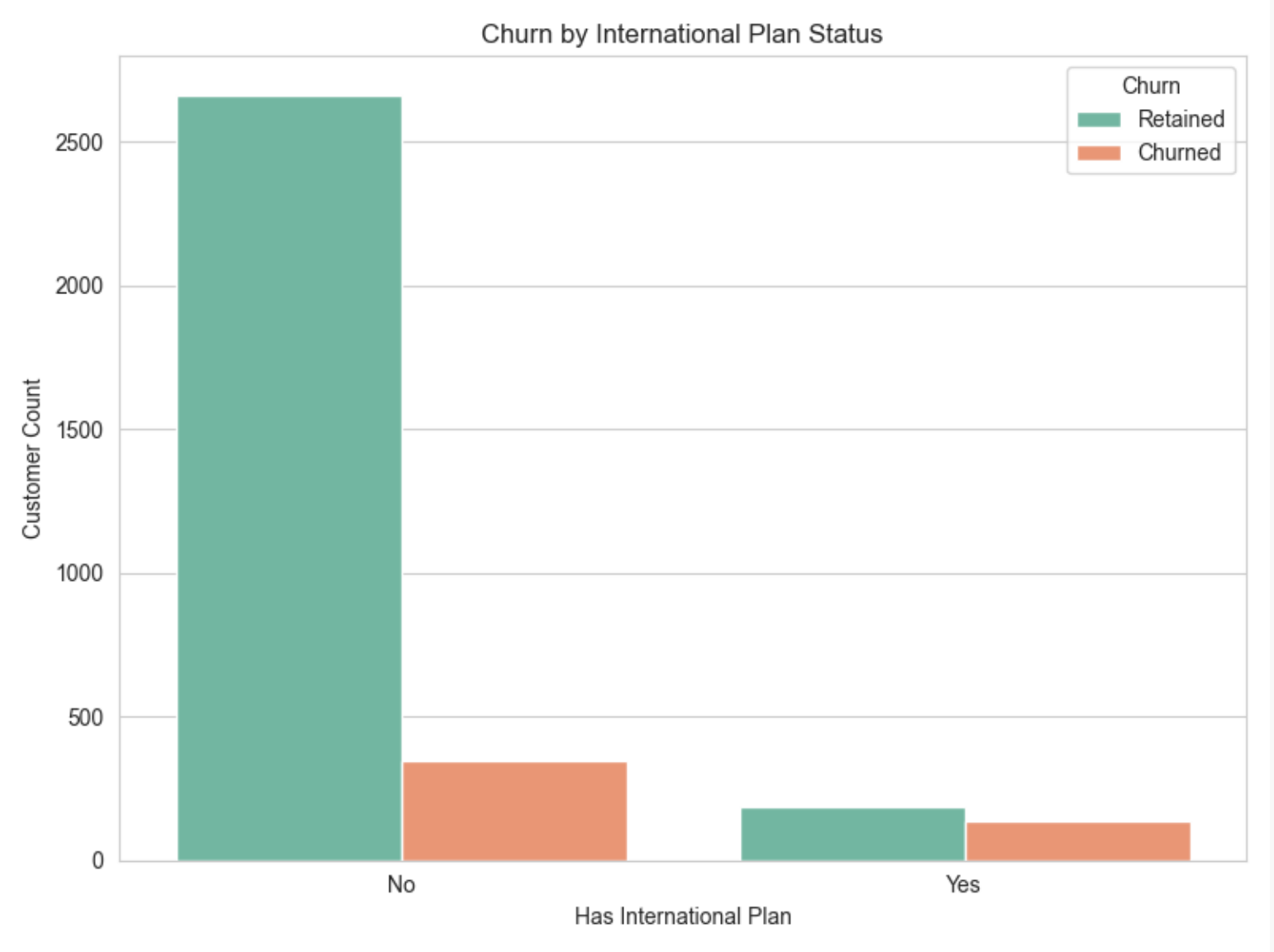
Customers with **international plans** are **more likely to churn** than those without — **contradicting expectations** that premium users are more loyal.

Possible Root Causes:

- Pricing Discrepancies:** Perceived as overpriced vs. competitors.
- Service Quality Issues:** Pain points in reliability, usability, or support.
- Lack of Value Proposition:** Customers unclear on the plan’s benefits.

Action Opportunity:

- Deep-dive into these drivers to:
- Refine international offerings
- Improve satisfaction & communication
- Convert at-risk users into loyal, profitable customers**

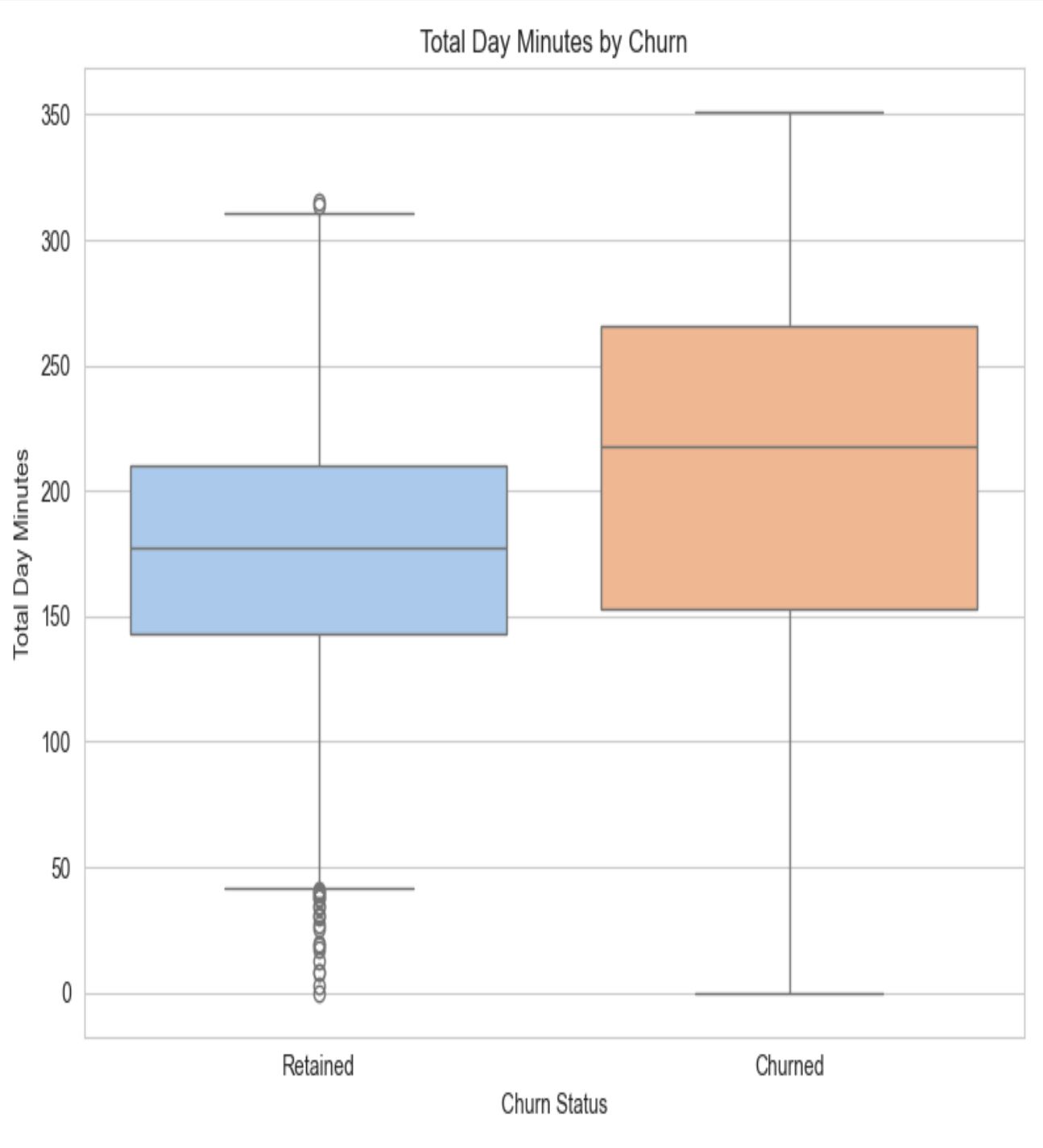


High Usage Can Signal Risk

Another counter-intuitive yet critical finding from our analysis is the correlation between high daytime call usage and an increased propensity to churn. Traditionally, high engagement might be interpreted as a sign of a satisfied and loyal customer. However, our data reveals that customers who churn tend to use **more daytime minutes** on average compared to those who remain with SyriaTel.

This finding suggests an important hypothesis: heavy users, particularly those with extensive daytime calling, may be experiencing "bill shock." They might feel that the value they receive does not align with the cost of their usage, leading them to actively seek more cost-effective solutions from competitors. While they are actively using our services, their satisfaction could be eroding due to perceived overpayment or a lack of sufficient value for their high consumption.

The implication is profound: **high engagement does not always equate to high loyalty.** Instead, it might signal an unmet need for better-value plans tailored to our most active users. This highlights an opportunity for SyriaTel to reassess its pricing strategies and plan structures for heavy daytime users. By introducing more predictable billing models, tiered plans, or loyalty incentives for high usage, we can transform a potential churn risk into a segment of highly satisfied, long-term advocates for SyriaTel. This proactive approach will be vital in retaining our most active and potentially valuable customers.



Building the Predictive Model & Top Factors Influencing Churn

How We Predict Churn

To transform our insights into actionable predictions, we leveraged the key customer behaviors and attributes identified in our extensive data analysis to train a sophisticated predictive model. The objective was not merely to predict churn, but to create a tool that is both highly accurate and exceptionally effective at identifying the specific, smaller group of customers most likely to churn.

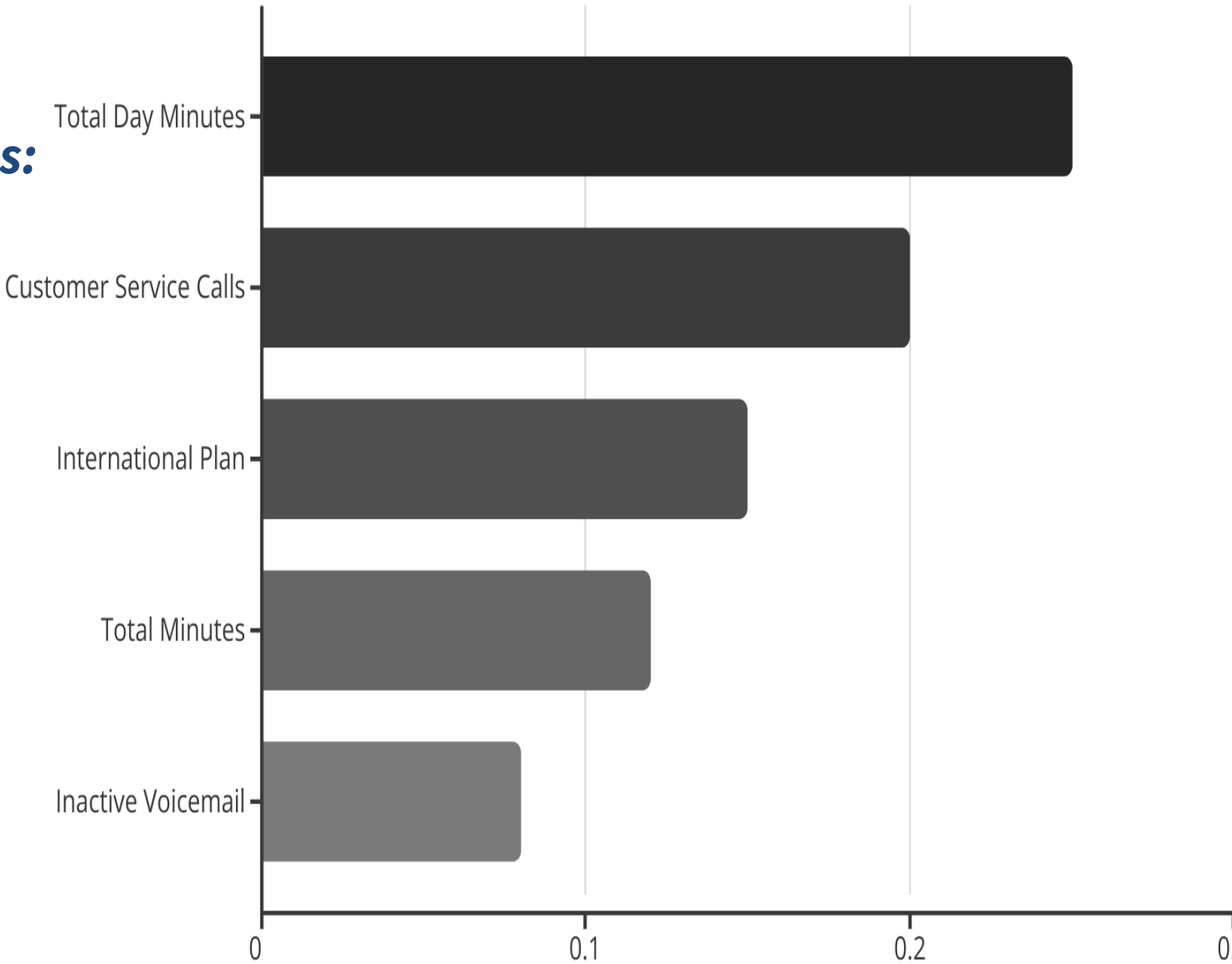
Our approach involved rigorously testing two distinct machine learning models:

- **Logistic Regression:** A reliable and interpretable baseline model that provides a solid foundation for binary classification tasks like churn prediction.
- **Random Forest:** A more powerful and advanced ensemble model, known for its ability to handle complex datasets, reduce overfitting, and deliver superior predictive performance.

The selection of these models was strategic, aiming to balance robust performance with the ability to interpret the drivers behind their predictions. The ultimate goal was to build a model that not only predicts churn with high accuracy but also excels at recalling the small, yet critical, group of customers who will churn, thereby maximizing the effectiveness of our retention efforts.

The Top Factors Influencing Churn

The predictive model's insights strongly align with the key factors uncovered in our initial data exploration. This convergence validates that our identified insights are indeed the most crucial drivers of churn, providing a robust foundation for our data-driven recommendations. The model consistently highlighted the following features as having the highest predictive importance:



Model Results: Random Forest is the Clear Winner

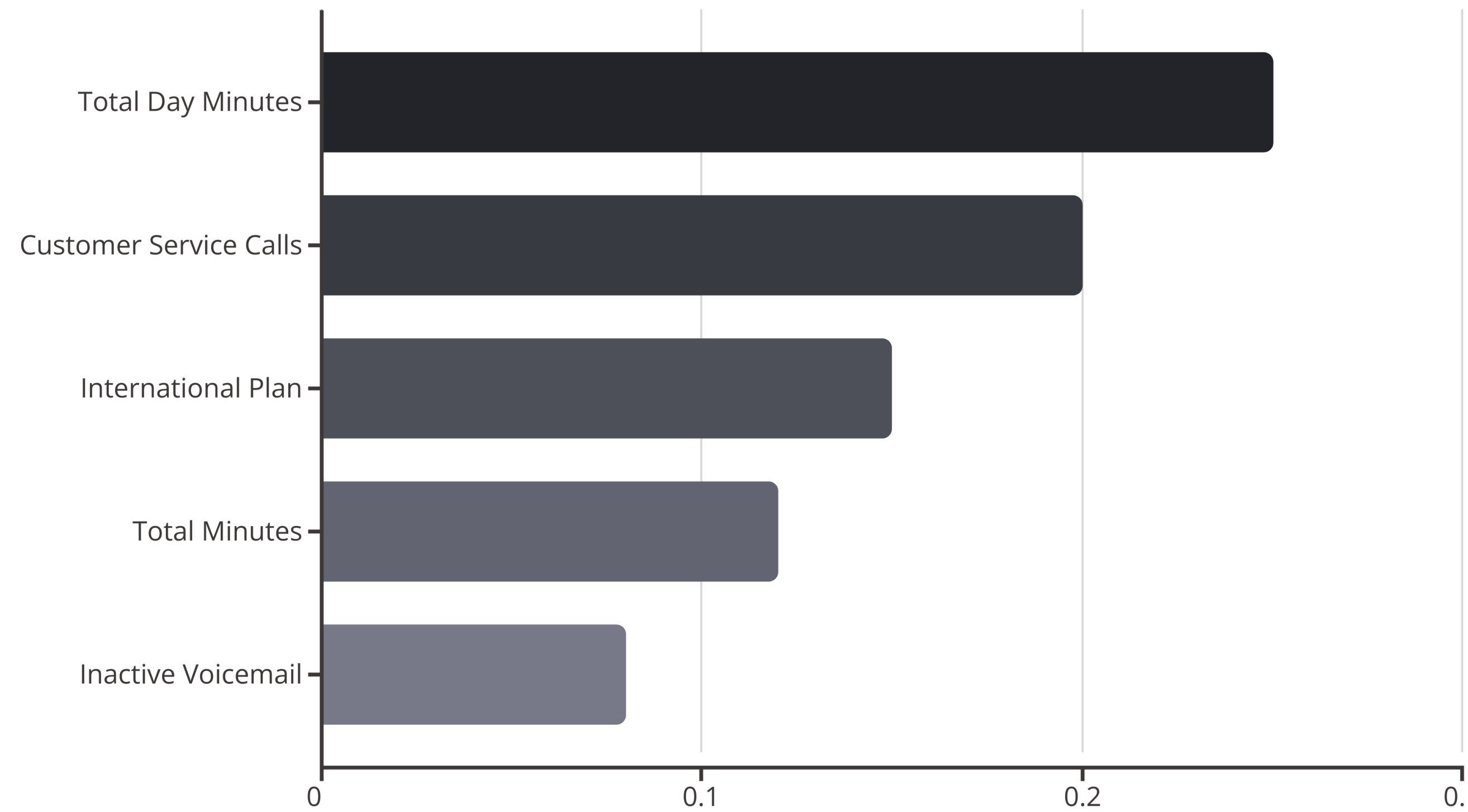
The Random Forest model stands out as an exceptionally accurate and reliable predictor of customer churn. Its robust performance enables targeted retention efforts.

Overall Accuracy	94%	The model is correct 94% of the time.
Churn Recall	62%	We successfully identify 62% of all customers who are going to churn.
Churn Precision	97%	When the model predicts a customer will churn, it is correct 97% of the time.
ROC AUC Score	0.90	An excellent score indicating the model is great at distinguishing between churners and non-churners.

The Random Forest model provides a powerful and reliable tool to focus our retention efforts with minimal wasted resources.

What Drives the Predictions?

The model's predictions are heavily influenced by the same key factors identified in our initial analysis, confirming alignment between insights and data.



Understanding these drivers allows us to pinpoint specific areas for intervention to reduce churn effectively.

Data-Driven Recommendations: Turning Insights into Action



Proactive Retention for High-Call Customers

Overhaul the International Plan

Create Value for High-Usage Users

Automatically flag customers in the CRM after their third service call, enabling retention specialists to intervene with tailored support or offers before churn.

Survey international plan users to understand pain points and benchmark against competitors. Redesign the plan for improved value, pricing, or usage bundles.

Introduce tiered or capped plans for heavy daytime users to prevent "bill shock." Offer predictable billing and loyalty benefits, transforming churn risks into loyal advocates.

Next Steps: Enhancing Our Predictive Power



Enrich Our Data

Integrate payment history, customer tenure, and service contract details for a more powerful model.



Advanced Techniques

Utilize techniques like SMOTE to improve churn prediction accuracy for the smaller churn group.



Deployment

Operationalize the model to provide real-time churn risk scores to customer success and marketing teams.

.....

Thank you

Q & A?

.....