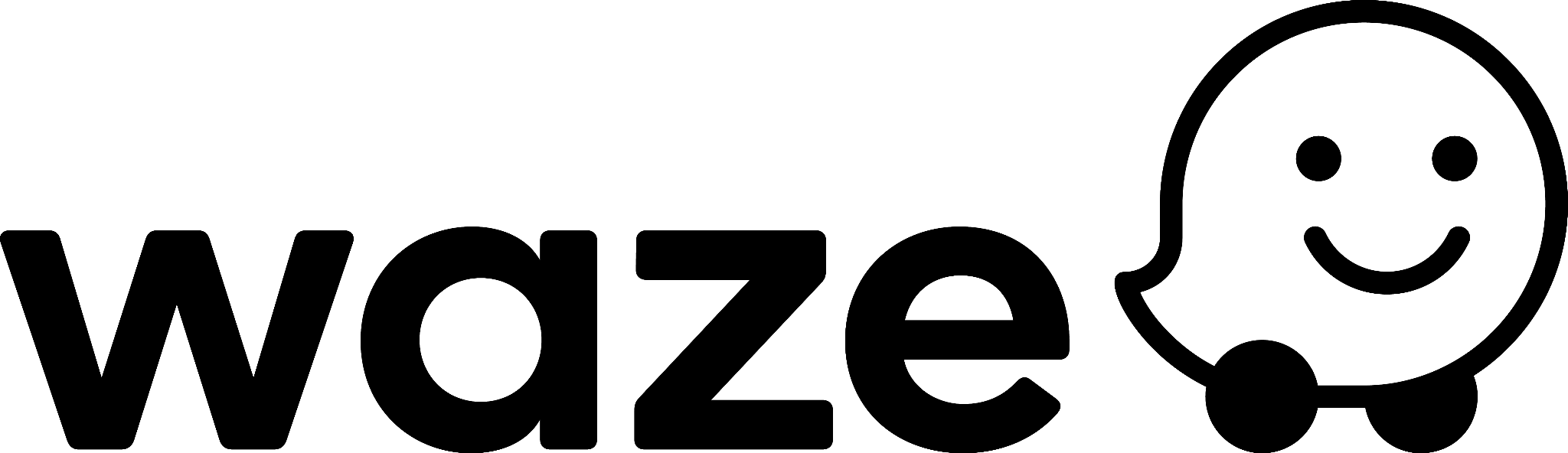
# **Executive Summary**

**Waze Navigation App User Churn Project**



**Waze User Churn Analysis**

**Executive Overview**

Our comprehensive analysis of Waze user behavior reveals critical insights into user retention patterns and churn drivers. With an overall churn rate of **17.74%**, we've identified key behavioral indicators that can guide strategic decision-making for user retention initiatives.

**Key Findings**

**User Retention Landscape**

* **82.27% of users remain active** on the platform, indicating strong overall user satisfaction
* **17.74% churn rate** provides opportunity for targeted retention strategies
* **iPhone users dominate** our user base at 64.5%, compared to 35.5% Android users

**Usage Patterns and Engagement**

**Session Behavior**

Our analysis reveals highly skewed usage patterns across the user base:

* **Median monthly sessions: 56** (50th percentile)
* **Significant power users exist** with over 700 sessions monthly
* **Right-skewed distribution** indicates most users have moderate engagement levels

**Driving Activity**

Similar patterns emerge in driving behavior:

* **Median monthly drives: 48**
* **Heavy users reach 400+ drives** monthly
* **Strong correlation** between session frequency and driving activity

**Critical User Tenure Insights**

**User Onboarding Distribution**

* **Nearly uniform tenure distribution** from new users to 9.5-year veterans
* **Median user tenure: ~5 years** (1,750 days)
* **Balanced representation** across all tenure segments

**Activity vs. Driving Engagement**

A fascinating behavioral pattern emerges:

* **Median monthly app opens: 16 days**
* **Users consistently engage** with the app beyond driving needs
* **App utilization exceeds driving frequency** at nearly every engagement level

**Risk Factors and Churn Drivers**

**High-Risk User Segments**

**Heavy Distance Drivers**

* **Users with high daily driving distances show elevated churn risk**
* **Potential one-time or tourist users** using Waze for extended trips
* **40% churn rate** among users with zero driving activity

**Low Engagement Users**

* **Inverse relationship** between driving frequency and churn probability
* **Users with 30 active days show 0% churn rate**
* **Engagement frequency is a strong predictor** of retention

**Unexpected Findings**

**Session Rate Paradox**

Our analysis uncovered a counterintuitive pattern:

* **Churned users demonstrate higher session rates** (0.2 sessions/day)
* **Retained users show lower session rates** (0.1 sessions/day)

This apparent contradiction is explained by:

* **Short-term intensive usage** before churning
* **Tenure bias** affecting long-term user metrics
* **Recent heavy usage** potentially indicating last-attempt engagement

**Data Quality and Methodology Notes**

**Data Integrity Assessment**

* **Overall data quality is robust** with consistent internal logic
* **Draving days cannot exceed Activity days** - validation confirmed
* **Some outliers identified** in distance metrics requiring further investigation

**Methodological Considerations**

* **Potential measurement discrepancy** between activity\_days (31 max) and driving\_days (30 max)
* **Possible different tracking periods** for monthly metrics
* **Recommendation for data team consultation** on variable definitions

**Strategic Recommendations**

**Immediate Actions**

1. **Implement early engagement monitoring** for users with high daily distances
2. **Develop retention campaigns** targeting low-frequency users
3. **Investigate and address** data collection inconsistencies

**Long-term Strategy**

1. **Leverage tenure-based segmentation** for personalized retention approaches
2. **Focus on consistent engagement** rather than intensive short-term usage
3. **Monitor session rate patterns** as early churn indicators

**Business Impact**

**Retention Opportunity**

With 17.74% churn rate across our user base, targeted interventions could significantly impact:

* **User lifetime value improvement**
* **Reduced acquisition costs** through better retention
* **Enhanced product-market fit** through behavior-driven improvements

**User Experience Optimization**

Understanding that users engage with Waze beyond driving creates opportunities for:

* **Feature development** supporting non-driving use cases
* **Engagement strategies** that don't require active driving
* **Community building** around location-based services

**Conclusion**

Our analysis reveals that Waze maintains strong user retention overall, but significant opportunities exist to reduce churn through targeted interventions. The key lies in identifying and supporting users who show early risk indicators while maintaining engagement strategies that recognize the diverse ways users interact with our platform.

The data suggests that **consistent, moderate engagement is more valuable than intensive short-term usage**, providing a clear framework for user success metrics and retention strategies moving forward.

*This analysis provides the foundation for data-driven decision making in user retention strategies. I recommend follow-up analysis focusing on the identified risk segments and validation of our data collection methodologies.*

**Waze User Churn Prediction Model**

**Overview**

Our data science team has developed and evaluated a machine learning model to predict user churn for the Waze platform. This executive summary presents our key findings, model performance metrics, and strategic recommendations for leadership consideration.

**Model Performance Summary**

The logistic regression model achieved an **83% overall accuracy** in predicting user churn behavior. However, a deeper analysis of the performance metrics reveals important nuances:

**Key Performance Metrics**

* **Overall Accuracy**: 83%
* **Precision for Churn Detection**: 57%
* **Recall for Churn Detection**: 8%
* **F1-Score for Churn Detection**: 0.15

**Model Predictions Breakdown**

* **True Negatives**: 3,481 users correctly predicted as retained
* **True Positives**: 64 users correctly predicted as churned
* **False Positives**: 697 users incorrectly predicted as retained (but actually churned)
* **False Negatives**: 48 users incorrectly predicted as churned (but actually retained)

**Critical Findings**

**1. Primary Predictive Factor**

**Activity Days** emerged as the most influential predictor of user churn, showing a strong negative correlation with churn behavior. This finding aligns with our exploratory data analysis, where we observed that users with higher activity levels demonstrate significantly lower churn rates.

**2. Unexpected Model Behavior**

Contrary to our initial exploratory analysis, **kilometers per driving day** proved to be a weaker predictor than anticipated. While our preliminary analysis showed this variable had the strongest positive correlation with churn, the multivariate model ranked it as the second-least important factor. This discrepancy highlights the complex interactions between variables in predictive modeling.

**3. Model Limitations**

The model demonstrates a significant weakness in **recall performance (8%)**, meaning it fails to identify 92% of users who actually churn. This represents a critical limitation for practical business applications.

**Strategic Implications**

**Business Impact Assessment**

While the model shows promise with 83% overall accuracy, the extremely low recall rate for churn detection presents substantial challenges for business implementation. The model's inability to effectively identify at-risk users limits its utility for proactive retention strategies.

**Variable Interactions**

The multivariate nature of our model reveals complex feature interactions that, while improving predictive accuracy, make the model less interpretable. This trade-off between predictive power and explainability is a common challenge in machine learning applications.

**Final Recommendations**

**Immediate Actions**

1. **Limited Deployment**: The current model should **not** be used for high-stakes business decisions or automated retention campaigns due to its poor recall performance.
2. **Exploratory Use**: The model can serve as a valuable tool for guiding further research and hypothesis generation about user behavior patterns.

**Future Development Priorities**

1. **Model Enhancement**: Invest in advanced modeling techniques to improve recall performance, potentially including:
   * Ensemble methods
   * Feature engineering
   * Alternative algorithms (Random Forest, Gradient Boosting)
   * Class imbalance handling techniques
2. **Data Collection**: Expand data collection efforts to capture additional user behavior indicators that may improve predictive accuracy.
3. **Threshold Optimization**: Investigate alternative classification thresholds to balance precision and recall based on business objectives.

**Conclusion**

Our initial churn prediction model provides valuable insights into user behavior patterns, particularly highlighting the critical importance of user activity levels. However, the model's current limitations prevent immediate deployment for business-critical applications. We recommend treating this as a foundational step in developing a more robust churn prediction system.

The next phase should focus on model improvement techniques and expanded data collection to enhance our ability to identify at-risk users effectively. With continued development, this approach has the potential to significantly impact user retention strategies and business outcomes.

*For technical details, methodology, and detailed performance metrics, please refer to the complete technical documentation provided separately.*