DATA SCIENCE STRATEGY & ANALYTICS

CONSULTING FOR EU-PARK

A data-driven roadmap designed for the board of directors to optimize wait times, guest satisfaction, and operational revenue.

PRESENTED BY

Mehmet Fatih Özdemir
Kasun Gayashan Pinto Ranwalage
Maulik Dilipbhai Chopda

DATE

15 January 2026

DATA INTEGRITY & METHODOLOGY

We established a reliable operational baseline by filtering ~40% of 'noise' from the raw data.

Audit Findings

The raw dataset contained non-operational hours (22:00–08:00) where the system defaulted to a static '5-minute' wait, heavily distorting reality.

Corrective Action

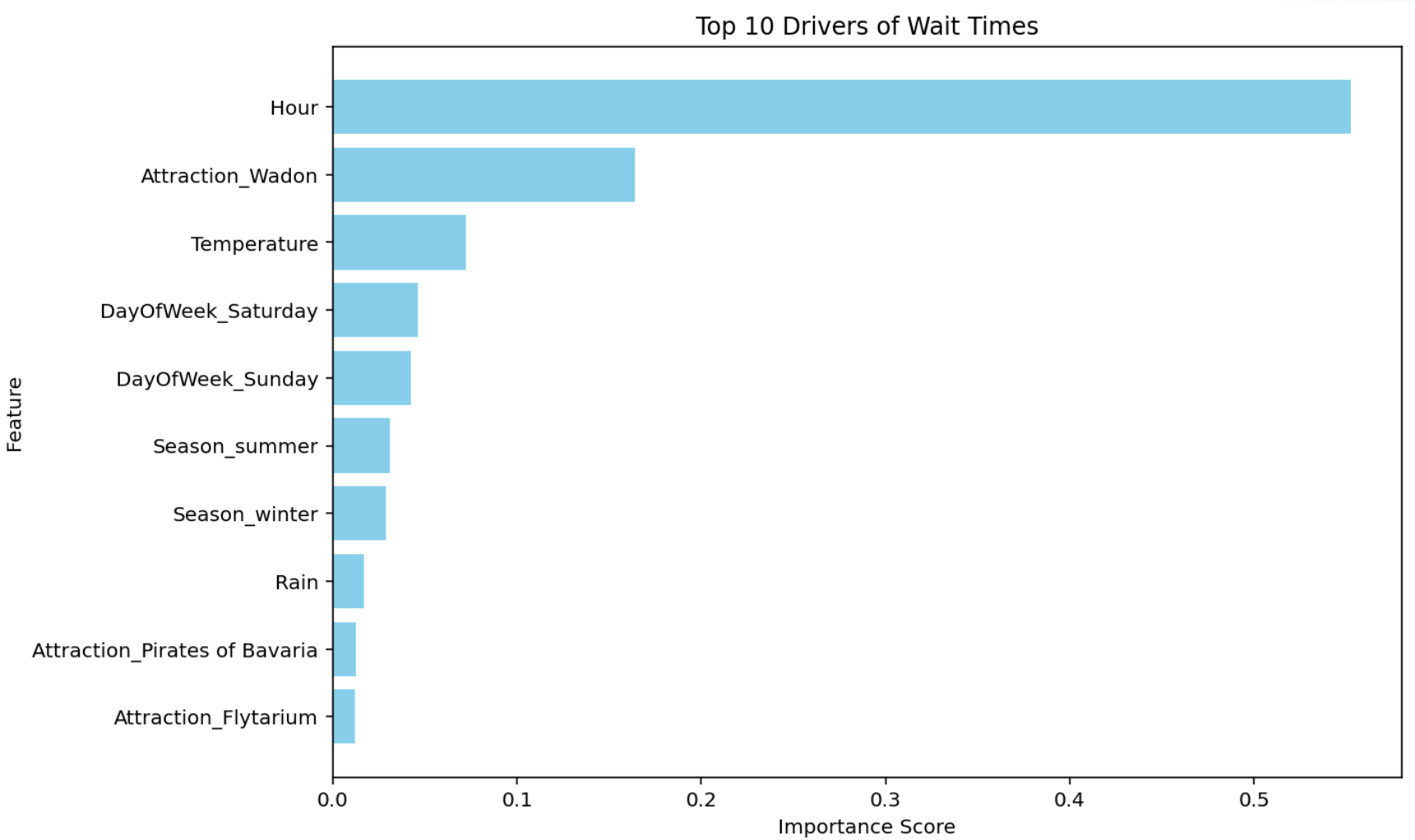
Applied strict time filter (08:00–21:00) and removed technical error codes (values < 0 or > 300 minutes) to isolate valid operational data.

Impact

Cleaning reduced the model's RMSE (error metric) from ~200 minutes to **9.8 minutes**, transforming it from unusable to precise.

STRATEGIC DRIVERS OF CONGESTION

Congestion is predictable: 'Time of Day' and specific attractions dictate 71% of wait time variance.



MODEL VALIDATION & RELIABILITY

The forecasting model achieves ~85% accuracy, meeting strict operational standards.

- **Precision:** The model demonstrates a Mean Absolute Error (MAE) of just 6.50 minutes.
- **Context:** On an average wait time of 42 minutes, this represents a 15.5% error margin, which is highly acceptable for guest expectation management.
- **Stability:** The Root Mean Squared Error (9.83 min) is significantly lower than the standard deviation, proving the model is robust against outliers.

Metric	Value	Assessment
Model Error (MAE)	**6.5 min**	High Precision
Average Wait Time	42.0 min	Baseline
Error Margin	15.5%	Within Tolerance
Model Accuracy	~85%	Operational Grade

STRATEGIC RECOMMENDATIONS

We recommend targeted interventions to maximize ROI on operational spend.

1

'Wadon' Virtual Queue

Since one ride drives 16% of congestion, implement a Virtual Queue specifically for 'Wadon' to disperse crowds.

2

Dynamic Staffing

Shift staff break times away from the model-identified peak window (11:00–15:00) to increase throughput capacity.

3

Guest Transparency

Display 'Forecasted Wait vs. Actual Wait' in the app to build trust using the model's 6.5-minute accuracy confidence interval.

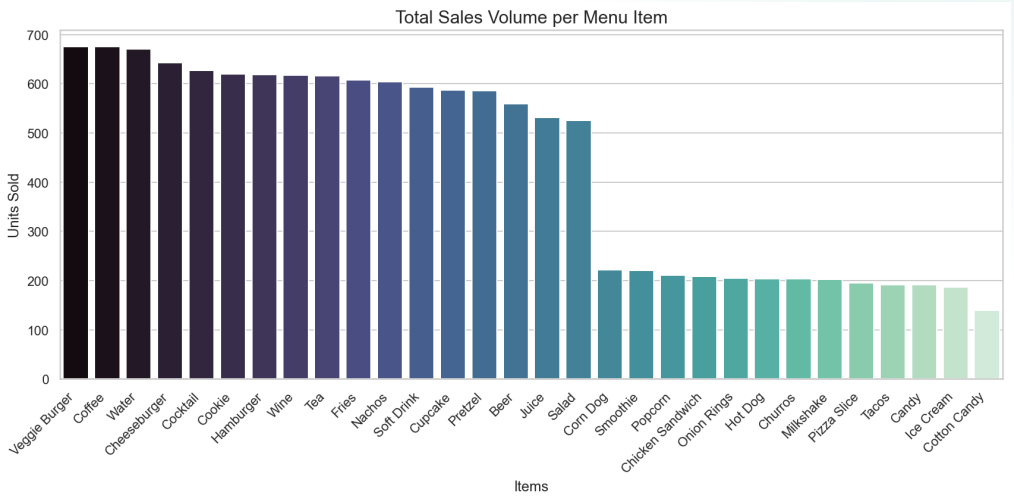
Best Seller & Basket Metrics

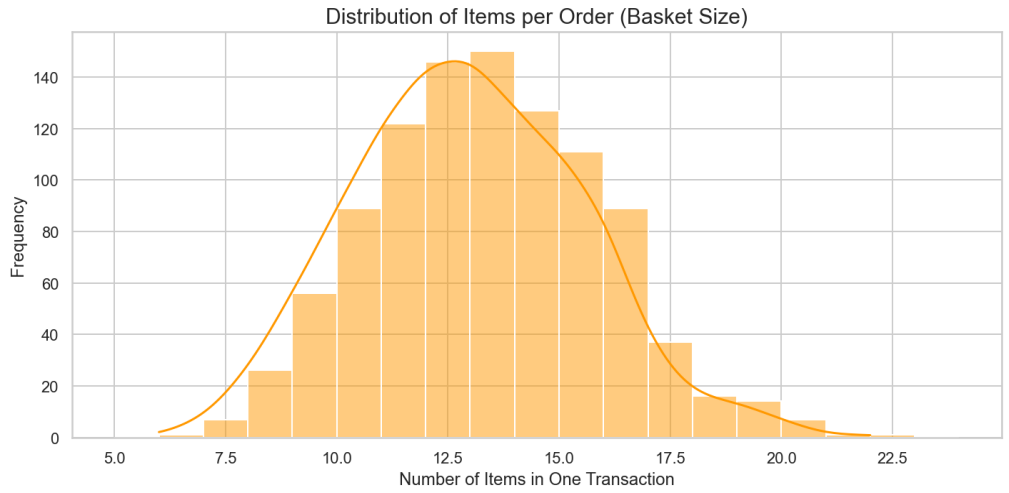
Key Sales Insights

· **Top Anchors:** Veggie Burger, Coffee, and Water represent the highest sales volume.

· **Basket Density:** The average transaction contains 13 items, indicating group dining is the primary behavior.

· **Opportunity:** Individual popularity doesn't capture the "Total Table Value" potential.





Visualizing Item Connectivity

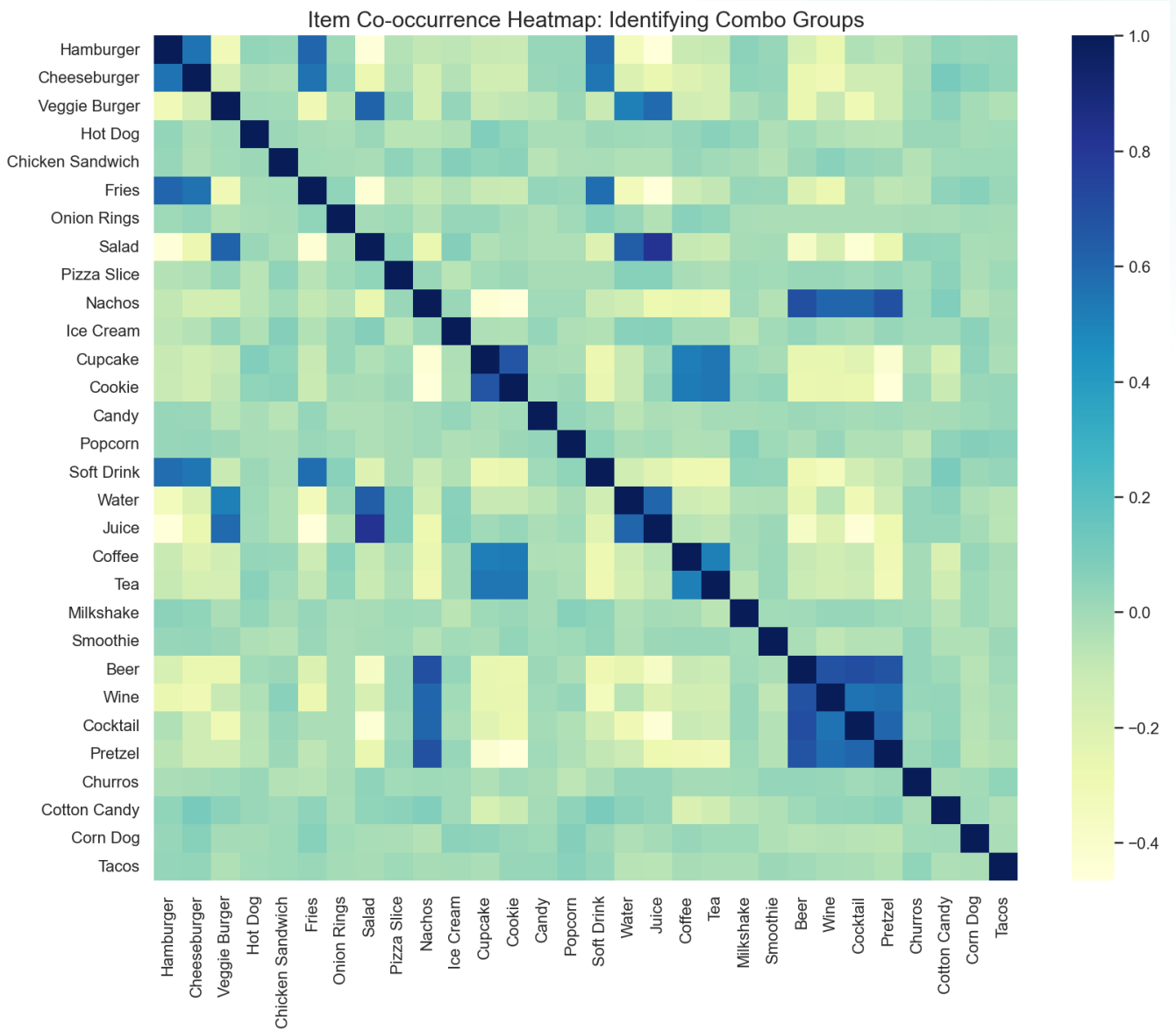
Correlation Heatmap Analysis

Our co-occurrence analysis identifies distinct clusters of items that are mathematically linked:

· **The Social Cluster**: Beer, Wine, Nachos, and Pretzels show the strongest link in the park.

· **The Healthy Cluster**: Salad and Juice have a near-perfect correlation (0.81).

· **The Cafe Cluster**: High connectivity between coffee, tea, cookies, and cupcakes.



Association Rule Validation

Antecedents (If)	Consequents (Then)	Confidence	Lift
Nachos, Pretzel, Wine	Beer, Cocktail	99%	3.77
Hamburger, Cheeseburger	Soft Drink, Fries	91%	3.56
Salad	Water, Juice, Veggie Burger	85%	3.37
Coffee, Tea	Cupcake, Cookie	88%	3.12

Note: A Lift of 4.0 indicates items are 4x more likely to be bought together than by chance.

Proposed Gold Pass Menu Bundles

Social Combo

Nachos, Pretzel, Beer, and Wine. Targets the highest-lift social drinking segment identified in the data.

Classic Diner Feast

Burger, Fries, and Soft Drink. The reliable high-volume anchor for group dining with 91% confidence.

Fresh Wellness Reward

Salad, Juice, and Veggie Burger. Capitalizes on the 0.81 correlation between healthy mains and sides.

Afternoon Sweet Retreat

Coffee, Cupcake, and Cookie. Optimized for the mid-afternoon snack window with 88% confidence.

Executive Summary

Objective

Distinguish drivers

behind ‘Gold’ and ‘Silver’ pass purchases

The Model

XGBClassifier

High-performance gradient boosting identifies non-linear behavior patterns.

Key Performance

86%

accuracy rate in predicting customer pass type

Technical Performance

Framework

XGBoost using Multi-class Log Loss (mlogloss)

Precision and Recall

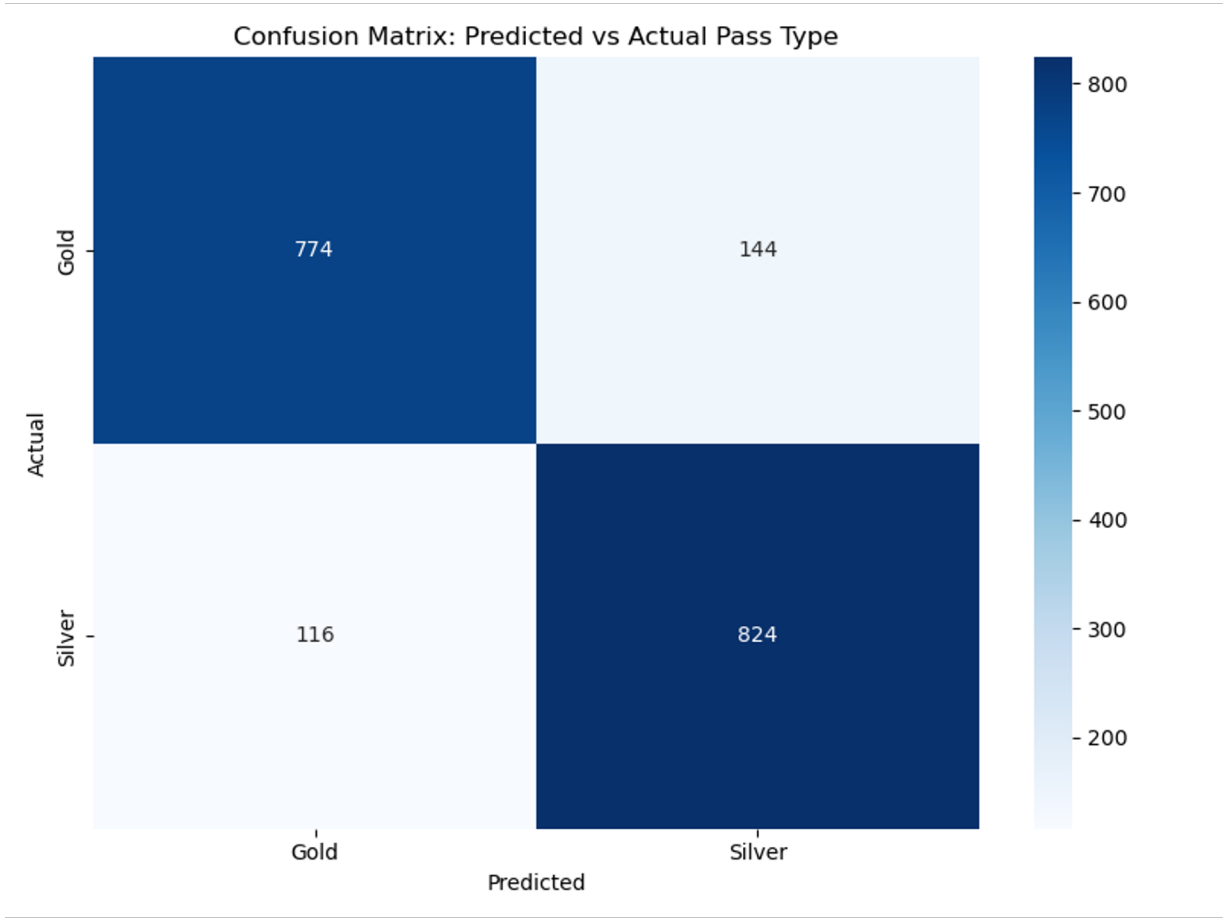
The model shows high confidence in distinguishing between ‘Gold’ and ‘Silver’ holders

	precision	recall
Gold	0.87	0.84
Silver	0.85	0.88

Data Integrity

Successfully handled missing values and excluded non-predictive features to prevent overfitting

Confusion Matrix



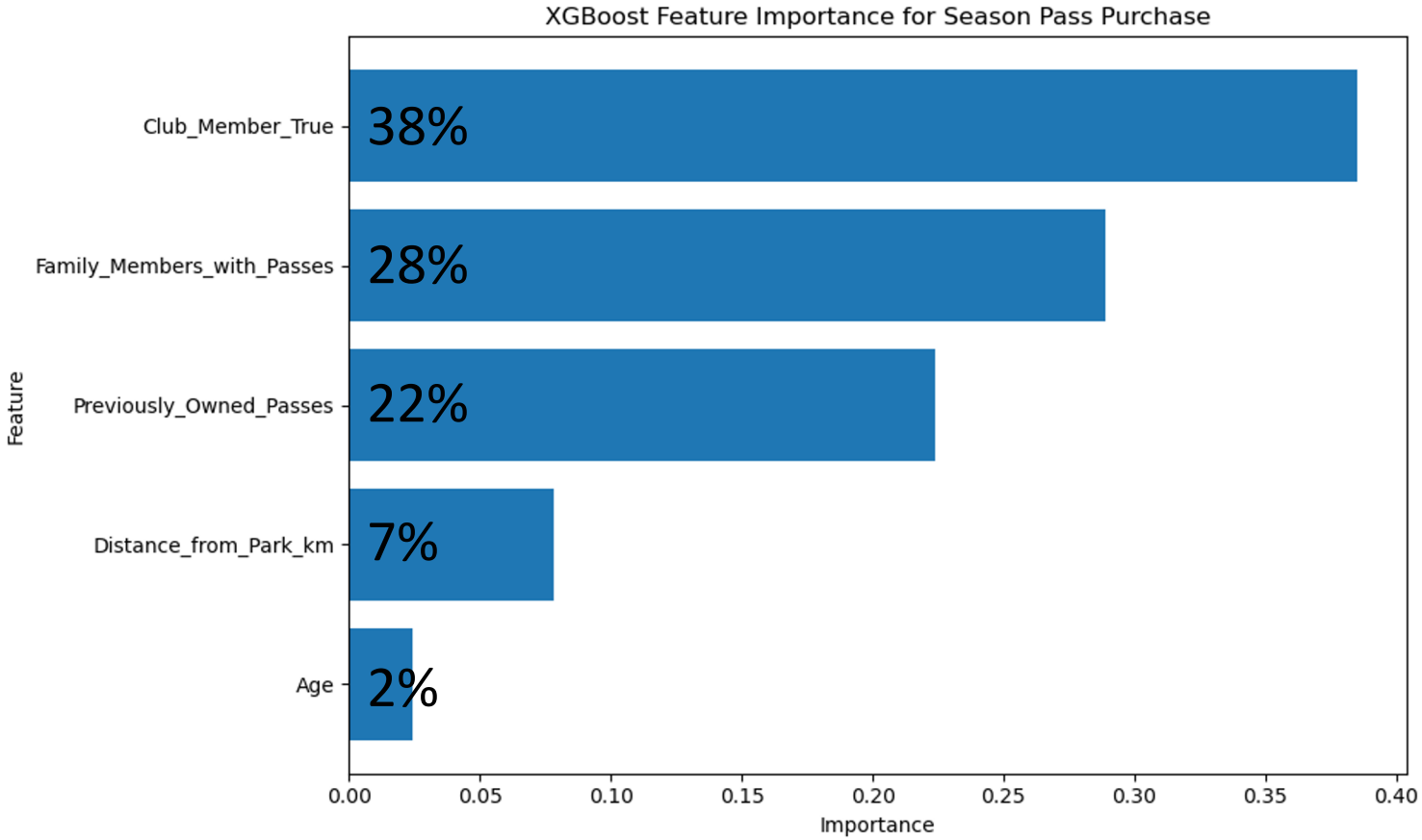
Key Performance

86%

accuracy rate in predicting customer pass type

Feature Importance

Ranked in descending order based on feature importance by XGBoost Feature Importance



THANK YOU

For your attention

NEXT STEPS

Q&A
Session

Image Sources

https://media.istockphoto.com/id/1170460641/photo/harbin-skyline.jpg?s=612x612&w=0&k=20&c=wiWmz44EhKJhZv24bkBQWQasgHglH3eYAePpyVknjX4=

Source: www.istockphoto.com