



Impact of Surrounding Amenities on Physical Point of Sales Performance

June 11th, 2018

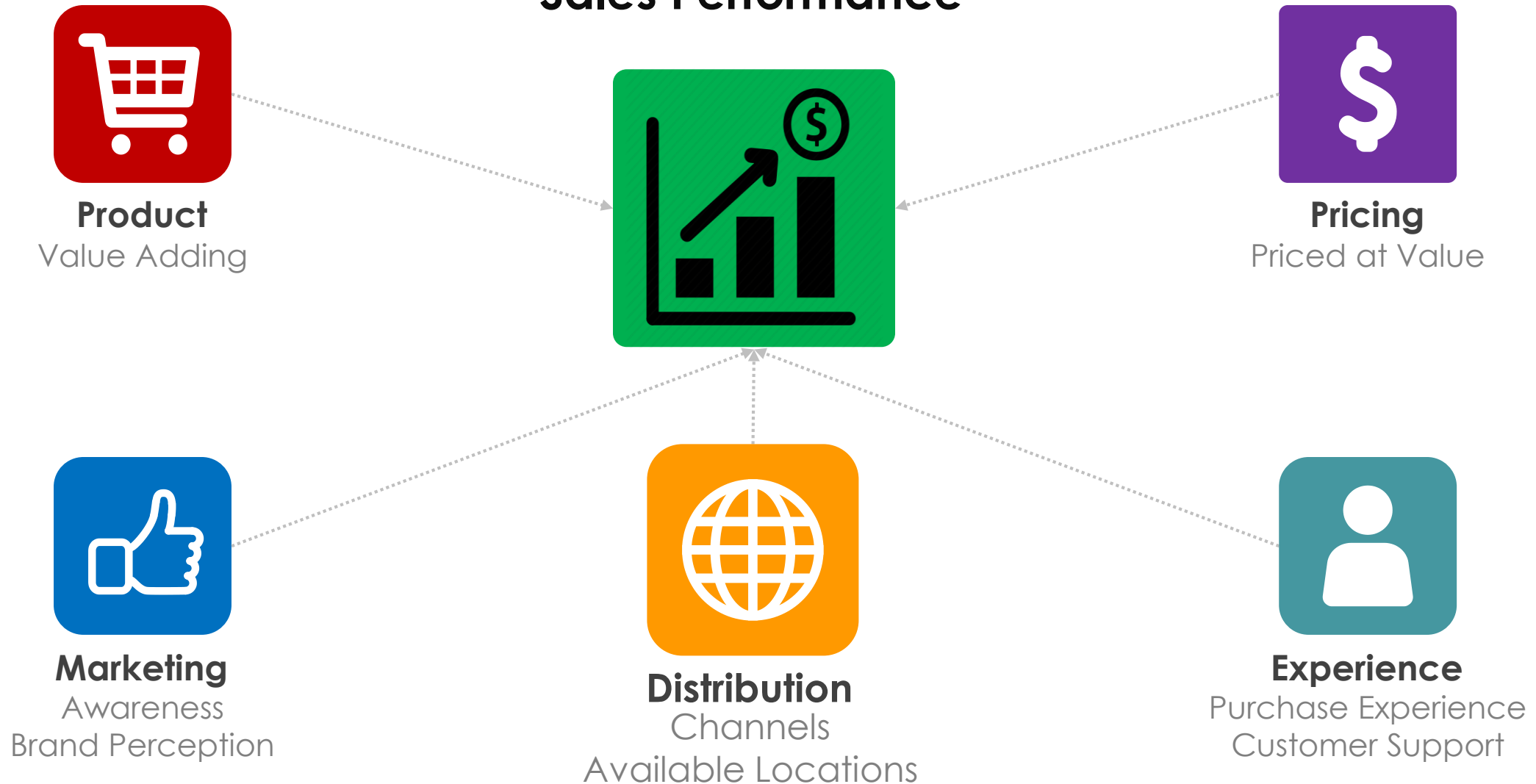
- Commercial Motivation
- Available Resources
 - Data Overview
 - Feature Engineering
 - Target Criterion
- Analysis
 - Modeling
 - Results
- Commercial Implications
- Extensions

Commercial Motivation – Maximizing Revenue



PHILIP MORRIS
INTERNATIONAL

Sales Performance



Available Resources – Data Overview



PHILIP MORRIS
INTERNATIONAL



POS Sales Volume



903 Points of Sale Locations
(anonymized by store code)



Hourly Sales Volume Numbers



August 2015 to June 2017



Surrounding Amenities



Queried in Q2 2017



89 different types of amenities within 500m radius
of a POS (ex. shops, café, bakery, schools etc.)



Individual Amenity Information
(ex. Address, Operating Hours, Contact etc.)



Rating scores for each applicable amenity
Sample reviews for each applicable amenity

Available Resources – Feature Engineering



POS Sales Volume



Sales Volume Aggregation
Sales per Point Distribution (SPPD)



Opening and Closing Dates



High Performing POS



Surrounding Amenities



Count of establishments for each amenity category

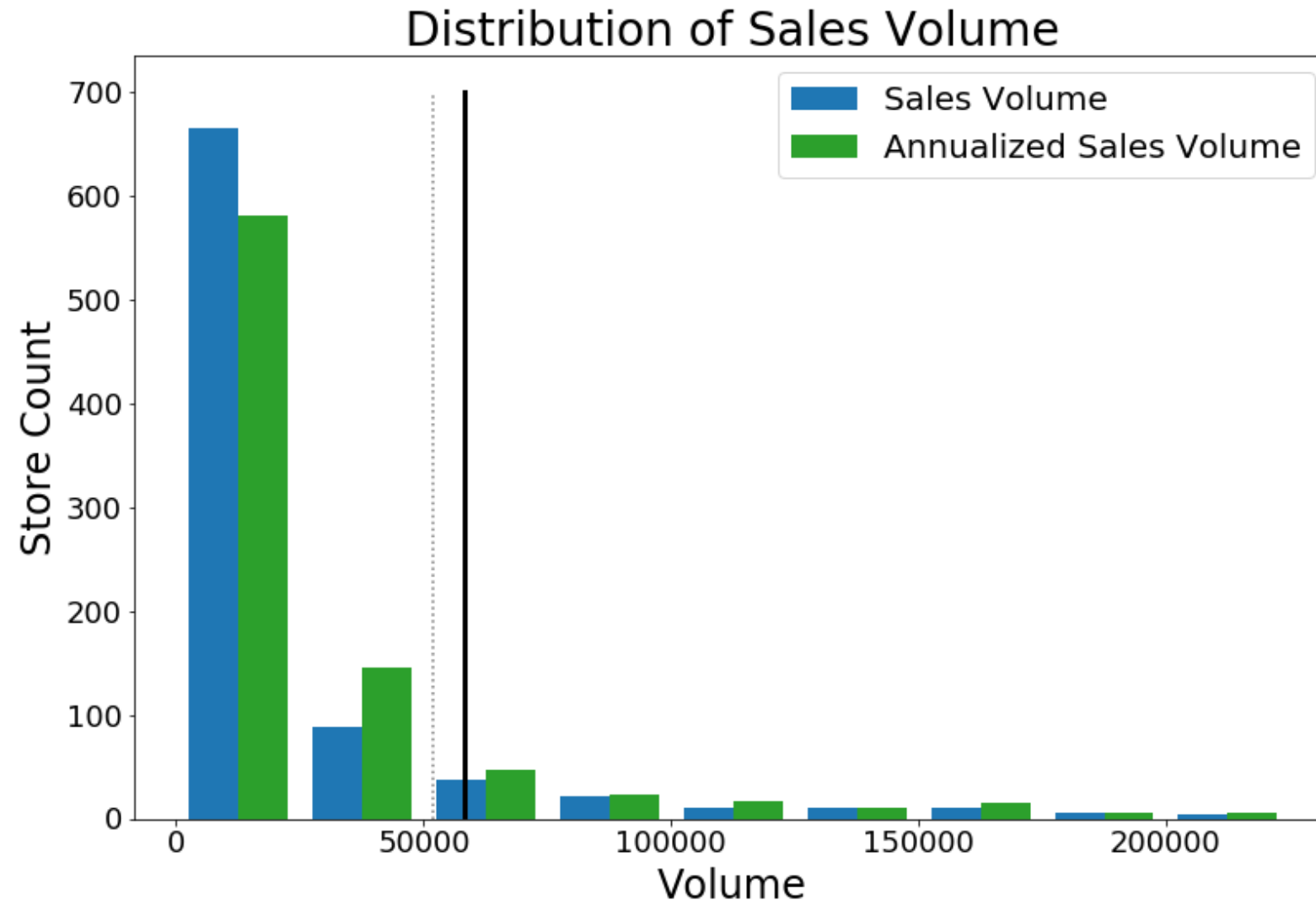


Total number of reviews for each amenity category



Average individual rating for each amenity category

Available Resources – Target Criterion



SPPD = Total Sales/POS Count

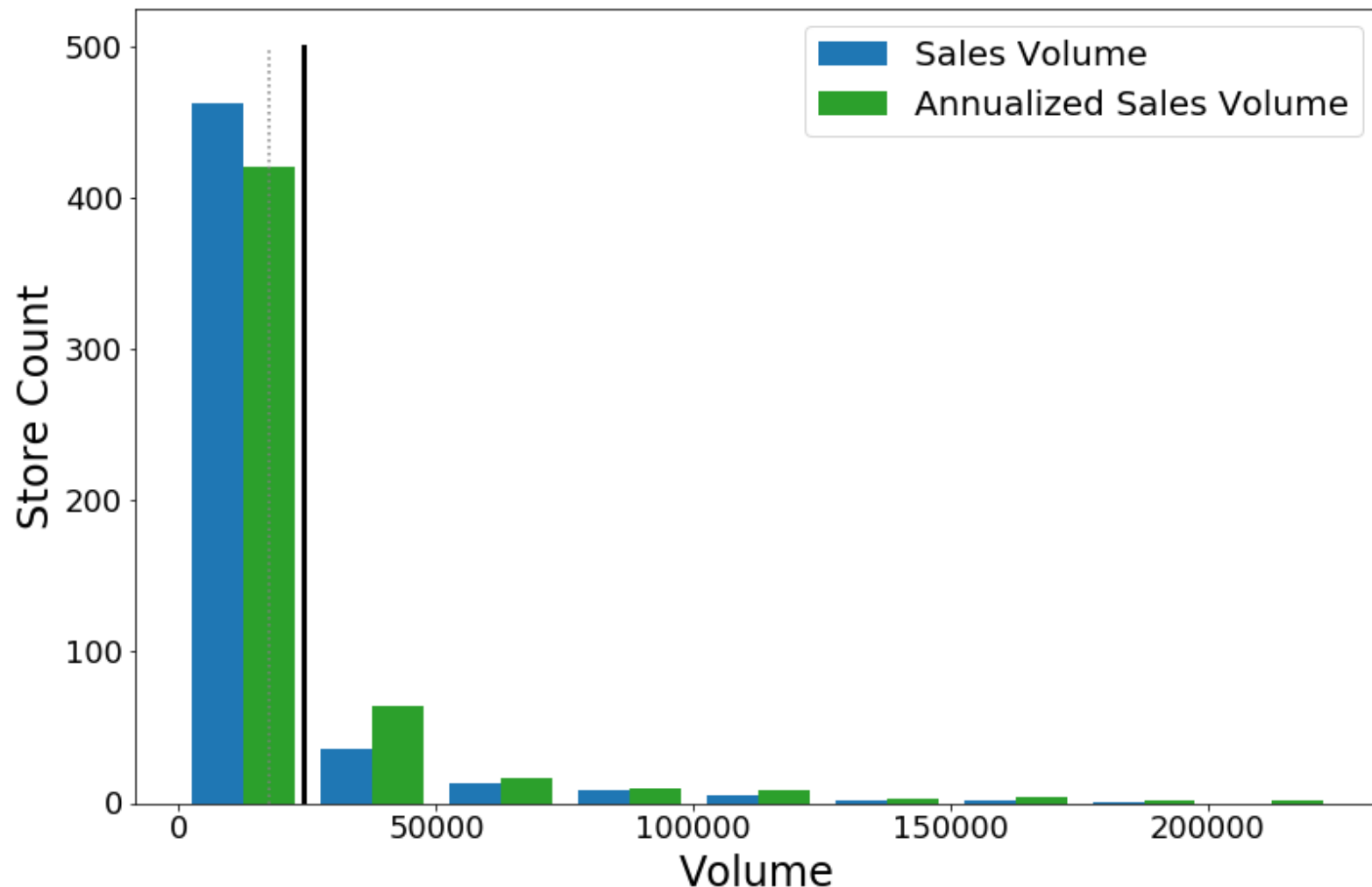
SPPD → Sales Efficiency

Stores with sales > SPPD are driving the sales pace in the channel

Sale data from June 25 2016 to June 25 2017
Total POS count: 903
156 POSs (17%) = High Performers

Available Resources - Treating Missing Data

Distribution of Sales Volume with Amenities Data



Amenities data missing for:
364 stores

Possible reasons/correlations:

- Dates?
- Location?
- Sales number?
- Hidden (latent) reasons

Remedies?

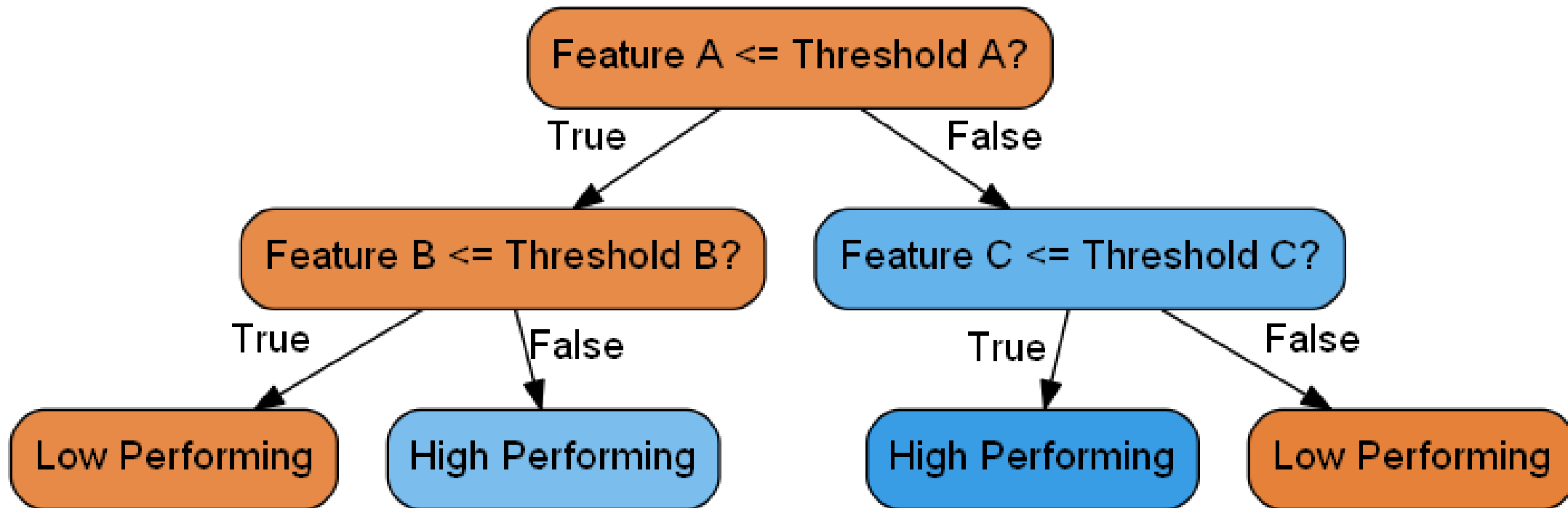
- Hot-deck
- Regression/Substitution
- Multiple imputation

Assumption:

- Missing **Completely** at Random
- Missing entries are dropped from training

**Sale data from June 25 2016 to June 25 2017
Total POS count: 539
48 POSs (9%) = High Performers*

Decision Trees



Analysis - Modeling



PHILIP MORRIS
INTERNATIONAL

Input Data



Performance
Category



Number of Amenities



Average Amenity
Rating



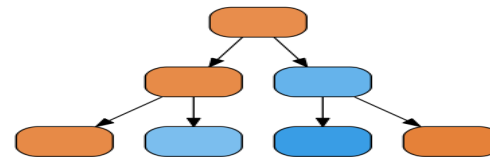
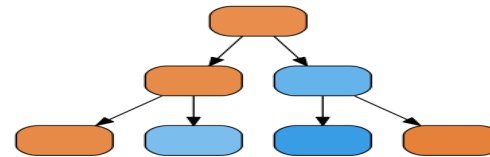
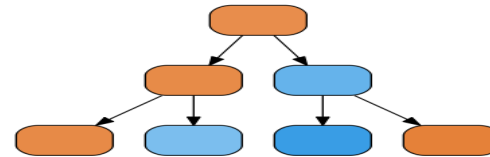
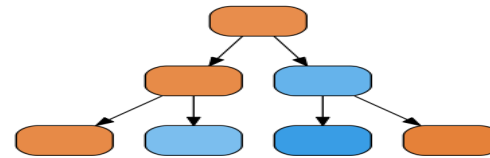
Number of Reviews

Random Forest

Gradient Boosting

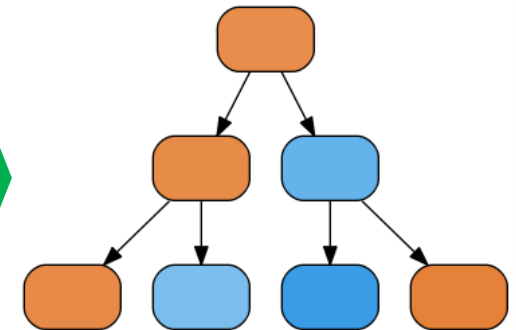
AdaBoost

Models



Precision

Selected Model



Analysis - Results



PHILIP MORRIS
INTERNATIONAL

Key Attributes

Random Forest

(Accuracy: 0.93)
(Precision: 0.22)
(AUC: 0.57)

Café Reviews

(>880)

Courthouses

(<0)

AdaBoost

(Accuracy: 0.91)
(Precision: 0.09)
(AUC: 0.5)

Shoe Store Ratings

(>3.8)

Gradient Boosting

(Accuracy: 0.93)
(Precision: 0.2219)
(AUC: 0.59)

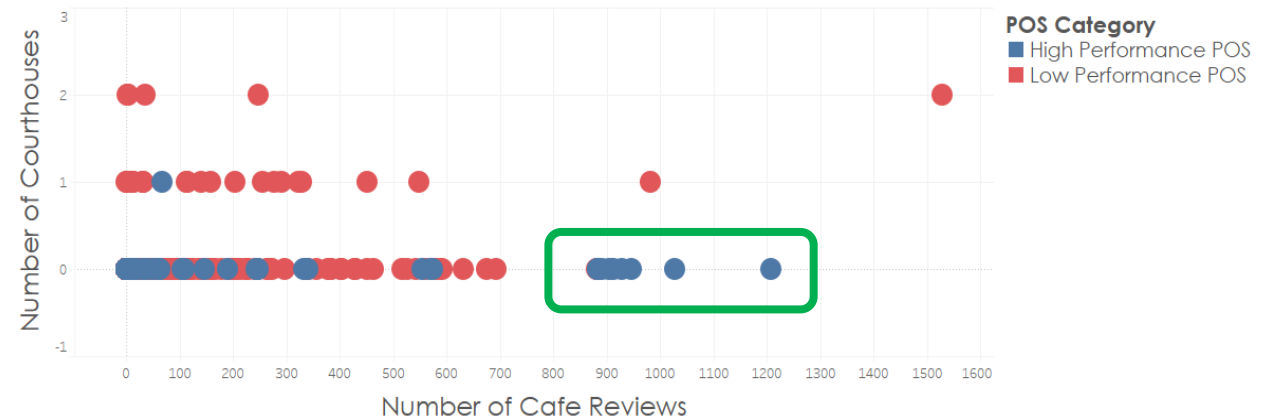
Gas Station Reviews

(>25)

Gym Reviews

(>110)

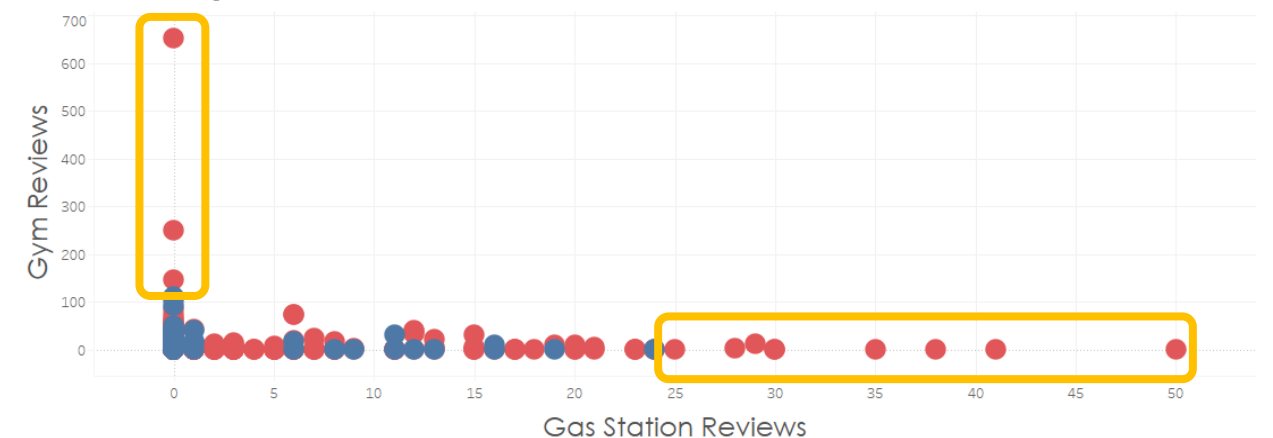
Cafe Review Counts vs Court Houses



Average Shoe Store Ratings



Gym Review Counts and Gas Station Review Counts



Commercial Implications



PHILIP MORRIS
INTERNATIONAL

When looking to introduce product to a new physical POS location:



Cafés

Cafés with **high** traffic
No courthouses nearby



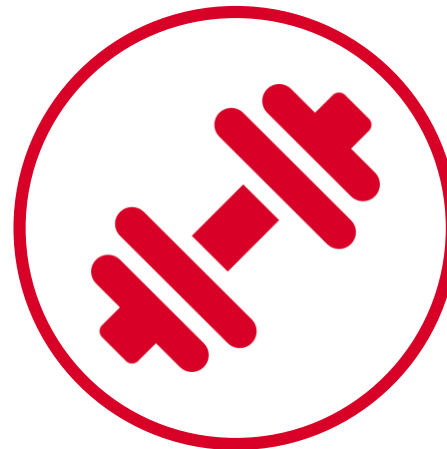
Gas Stations

Avoid busy gas stations



Shoe Stores

Avoid highly reputable
shoe stores.



Gyms

Avoid popular gyms
Make sure not too much foot traffic

- Recover missing amenities data
- Incorporating Additional Data:
 - POS information – ex. staff, space, address, type
 - Consumer/census data
 - Information from other distribution channels (ex. online sales)
 - Retail pricing
- Create ensemble/voting classifier for better prediction