David Eccles School of Business
THE UNIVERSITY OF UTAH

Swire Coca-Cola

Predicting Customer Success

Stephen Smart, Alvaro Gonzalez,
Brian Bombasi, Addison Farber

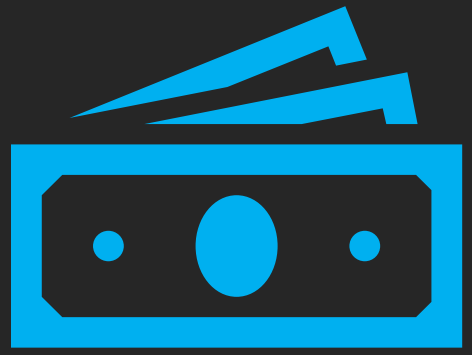
How can we identify successful customers?

Profitability

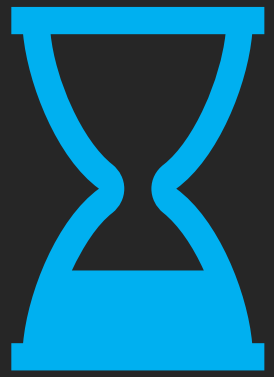
Longevity

Number of transactions

Target Variable



Projection profit in 3 years



Longevity & Activity in last 90 days

Profit in 3 years = *Profit per transaction * Expected Number of Transactions *3*

$$\text{Profit in 3 years} = \frac{Total\ Profit}{\#transactions} * \frac{\#transactions}{MaxPost\ date - MinPostdate} * 3$$

$$\text{Profit in 3 years} = \frac{Total\ Profit}{Max\ Date - Min\ Date} * 3$$

Profit in 3 years = *Profit per transaction * Expected Number of Transactions *3*

$$\text{Profit in 3 years} = \frac{Total\ Profit}{\#transactions} * \frac{\#transactions}{MaxPost\ date - MinPostdate} * 3$$

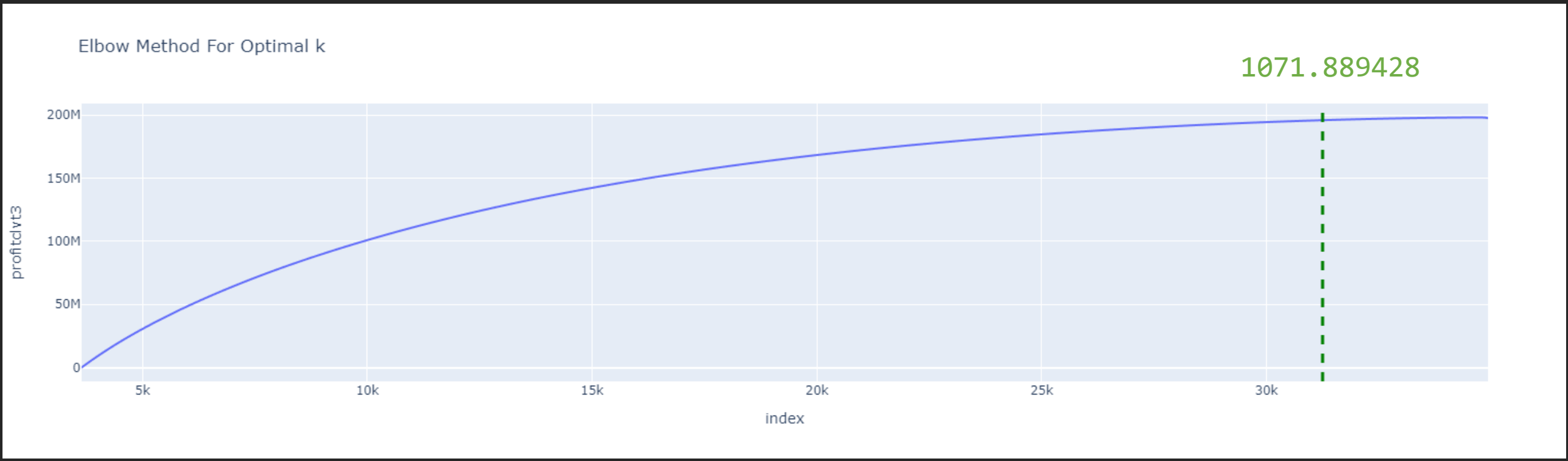
$$\text{Profit in 3 years} = \frac{Total\ Profit}{Max\ Date - Min\ Date} * 3$$

Profit in 3 years = *Profit per transaction * Expected Number of Transactions *3*

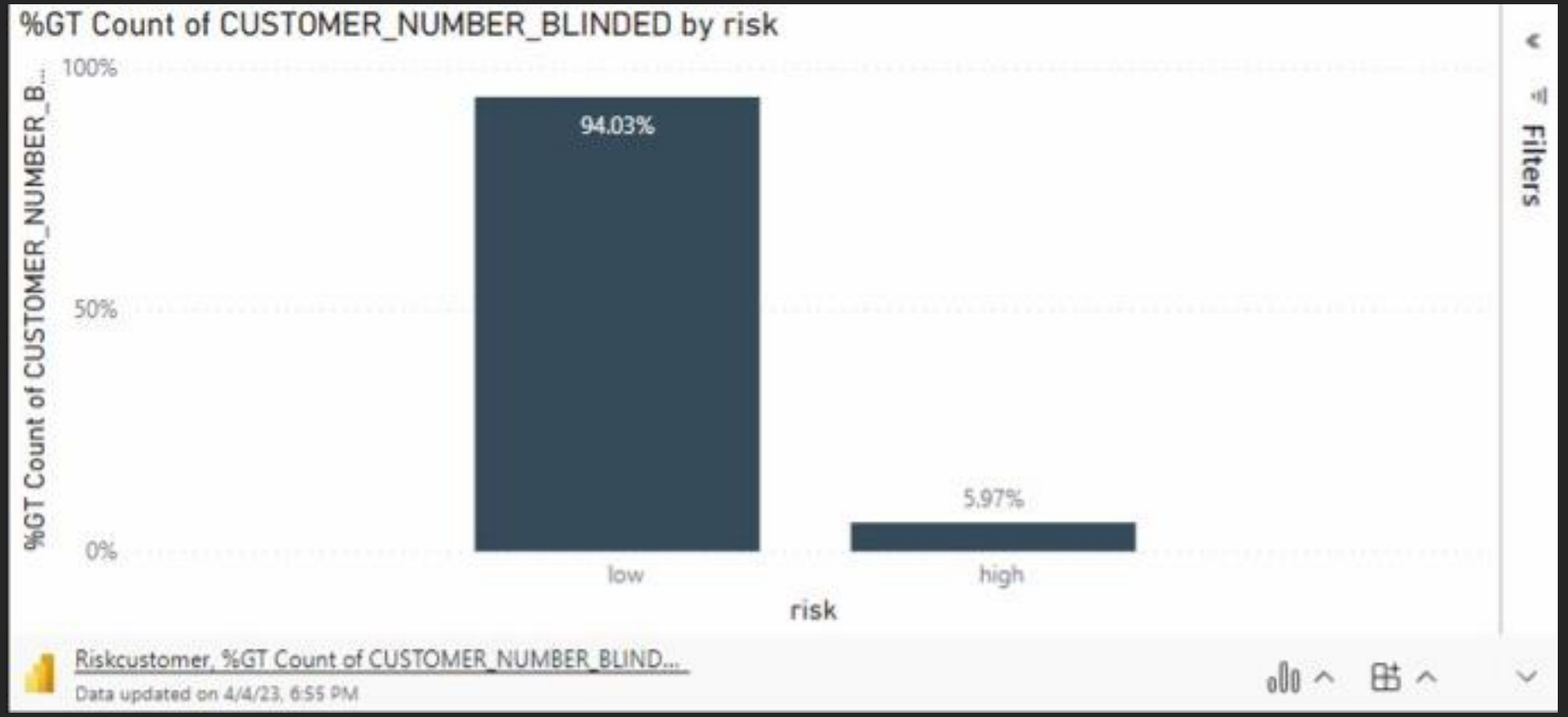
$$\text{Profit in 3 years} = \frac{Total\ Profit}{\#transactions} * \frac{\#transactions}{MaxPost\ date - MinPostdate} * 3$$

$$\text{Profit in 3 years} = \frac{Total\ Profit}{Max\ Date - Min\ Date} * 3$$

What is a good Profit in 3 years?



Imbalanced data set



Should we offer a deal ?

Profit in 3 yrs Median
4369

High Success Median
5669

Low Success Median
636

Average net benefit of Customer Success

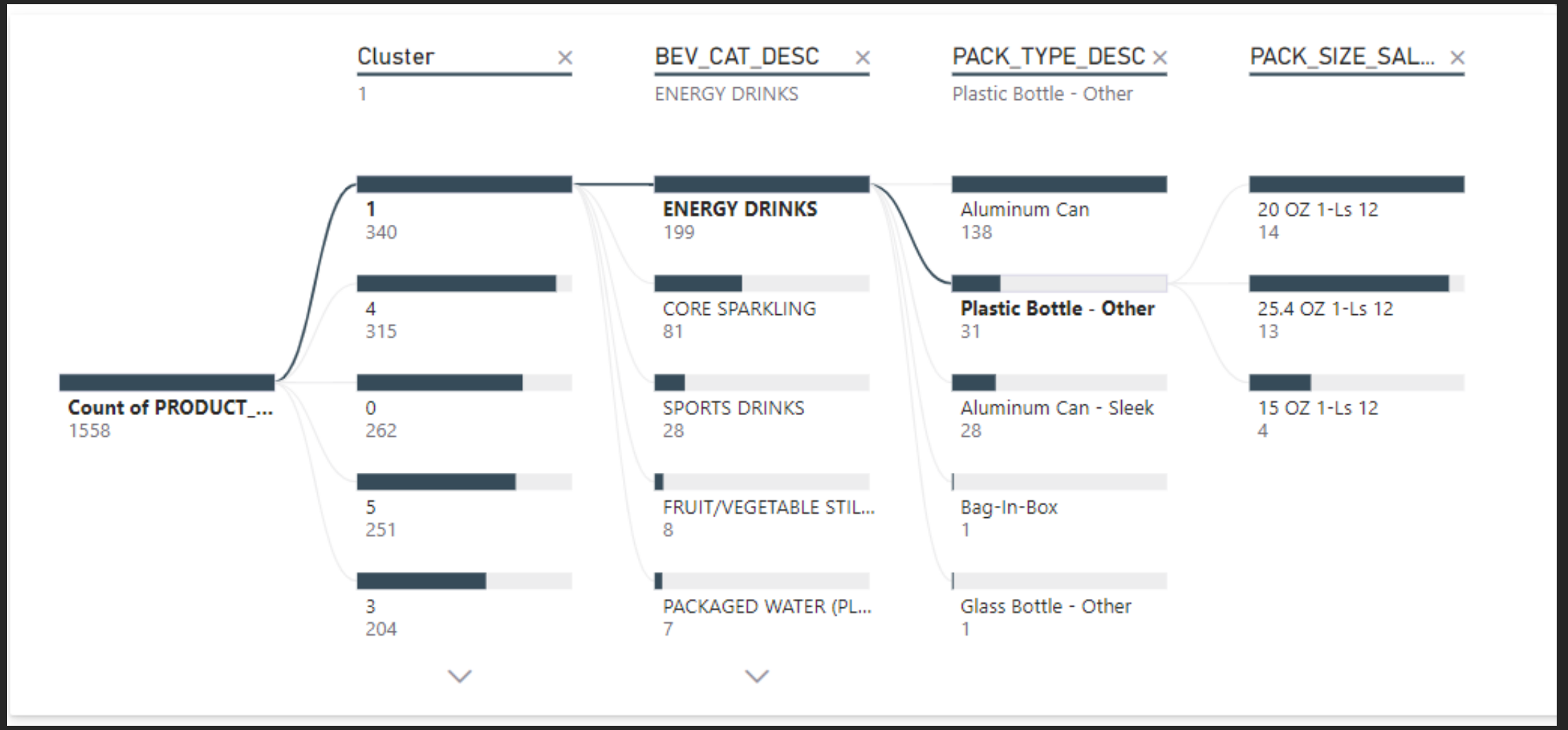
	Predicted low	Predicted High
Actual low	0	-3733
Actual High	-1300	1300

Profit in 3 yrs Median 4369

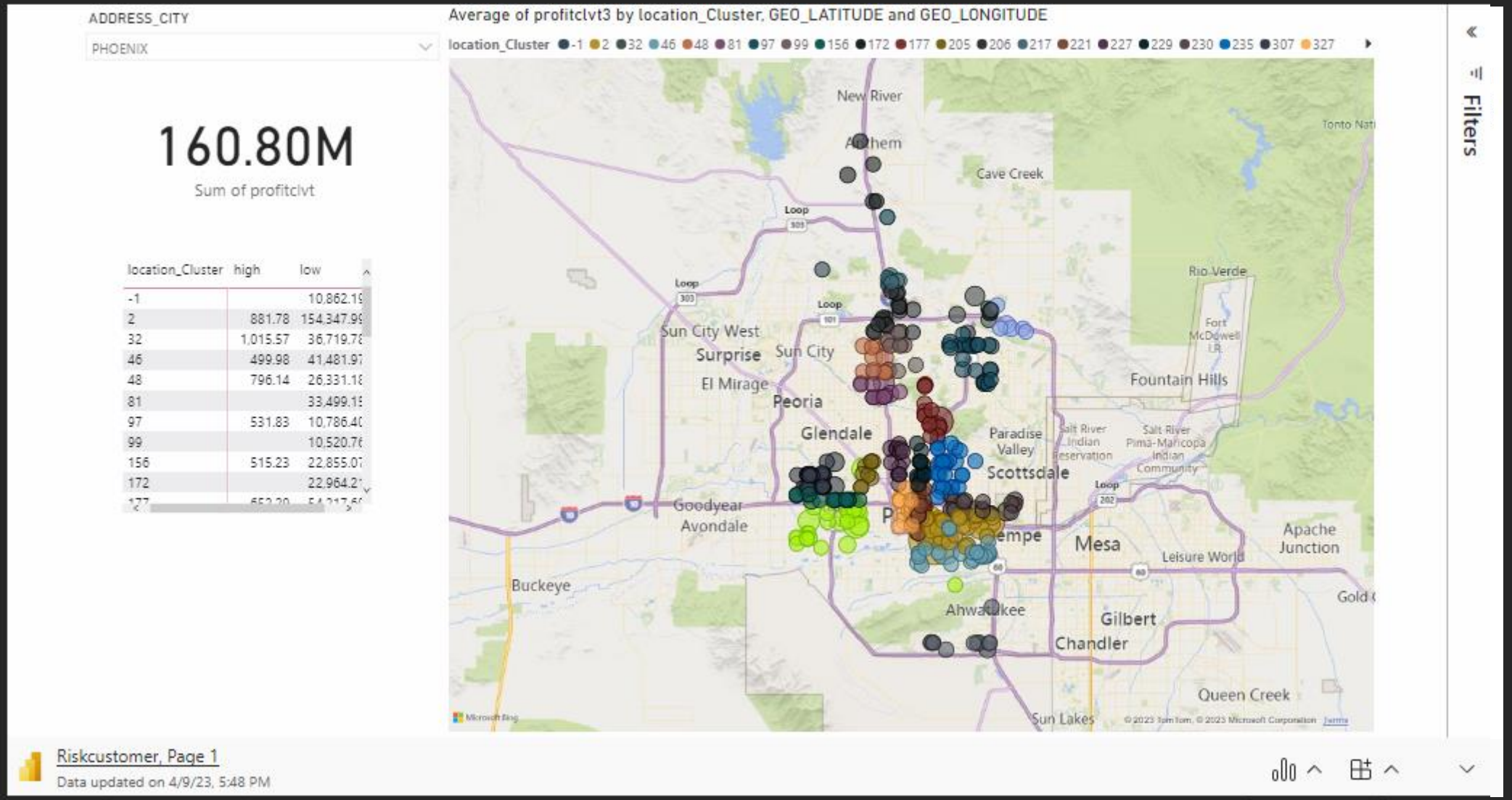
SVM increases in average 1087

Logistic Regression weighted increases 1064

Product Clustering



Location Clustering



Future Analysis
+
Recommendations

1) Improved methods of modeling customer success

1) Improved methods of modeling customer success

2) Discount Risk Mitigation

Improved Modeling Method

Current Challanges

- Sales predictions are not necessarily based off previous sales

Improved Modeling Method

Current Challanges

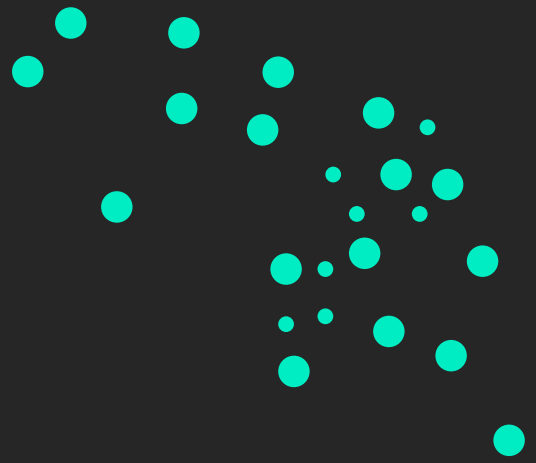
- Sales predictions are not necessarily based off previous sales
- We are actually asking for a prediction of the B2B partner's success

Improved Modeling Method

Current Challanges

- Sales predictions are not necessarily based off previous sales
- We are actually asking for a prediction of the B2B partner's success
- Location plays a much great factor than initially realized

Feature engineering



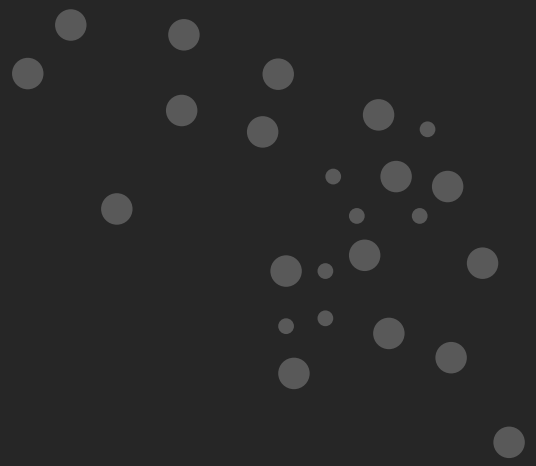
Clustering

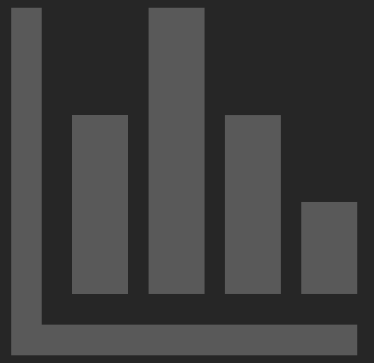
Census Data

Yelp Data

Google Analytics

Features engineering

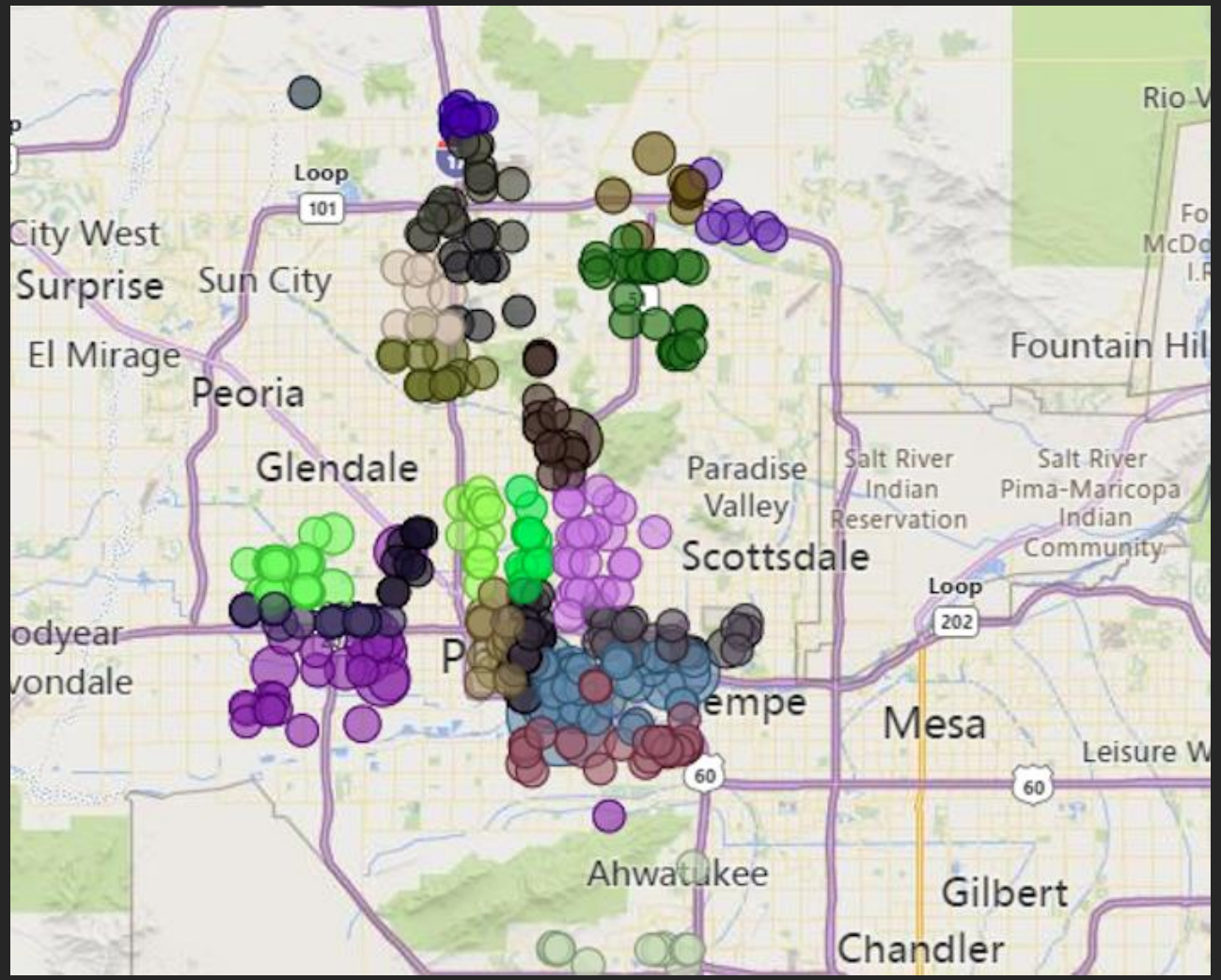




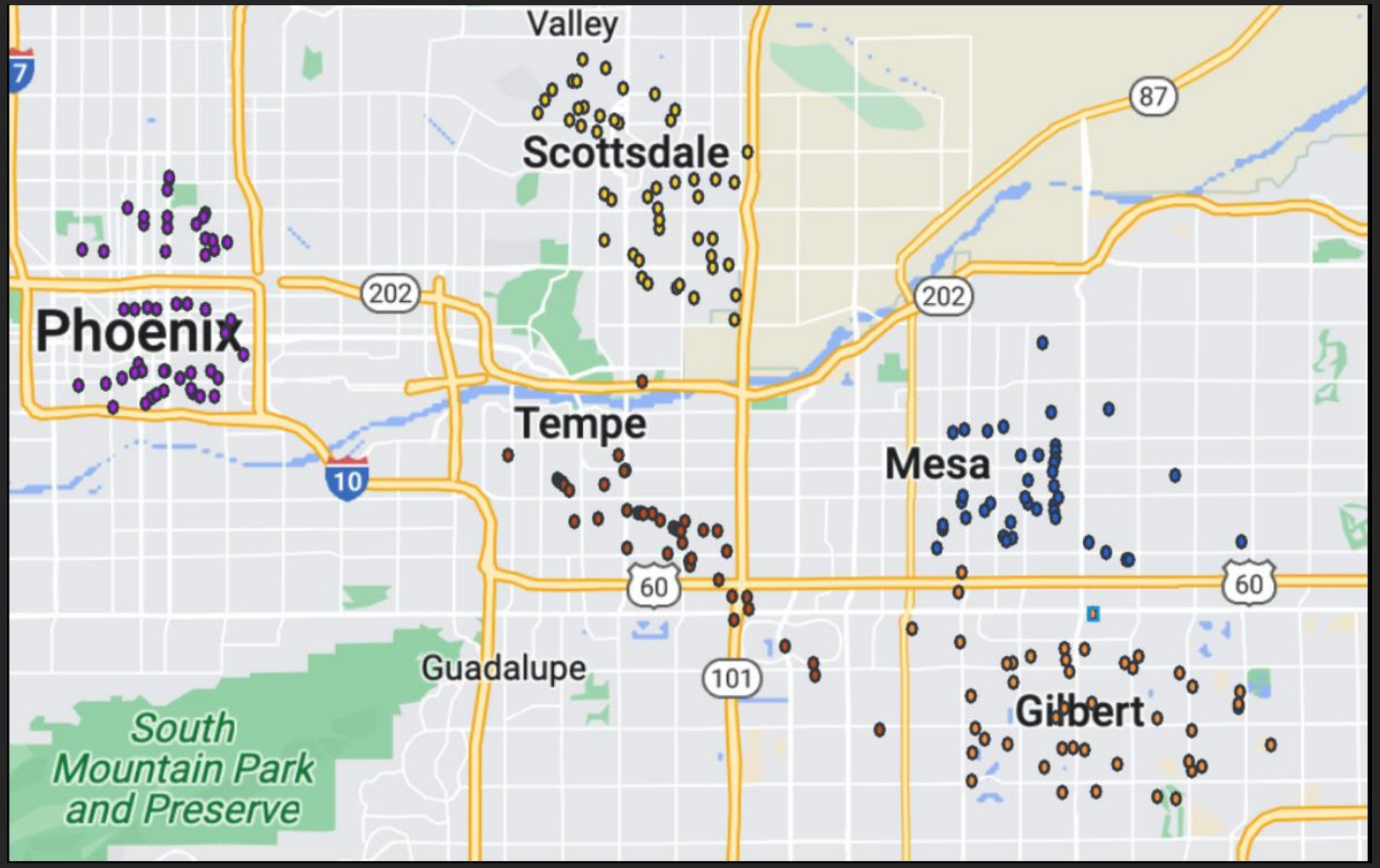
Clustering Census Data Yelp Data **Google Analytics**

This strategy provides a tailored analysis for each location and the assigned sales rep

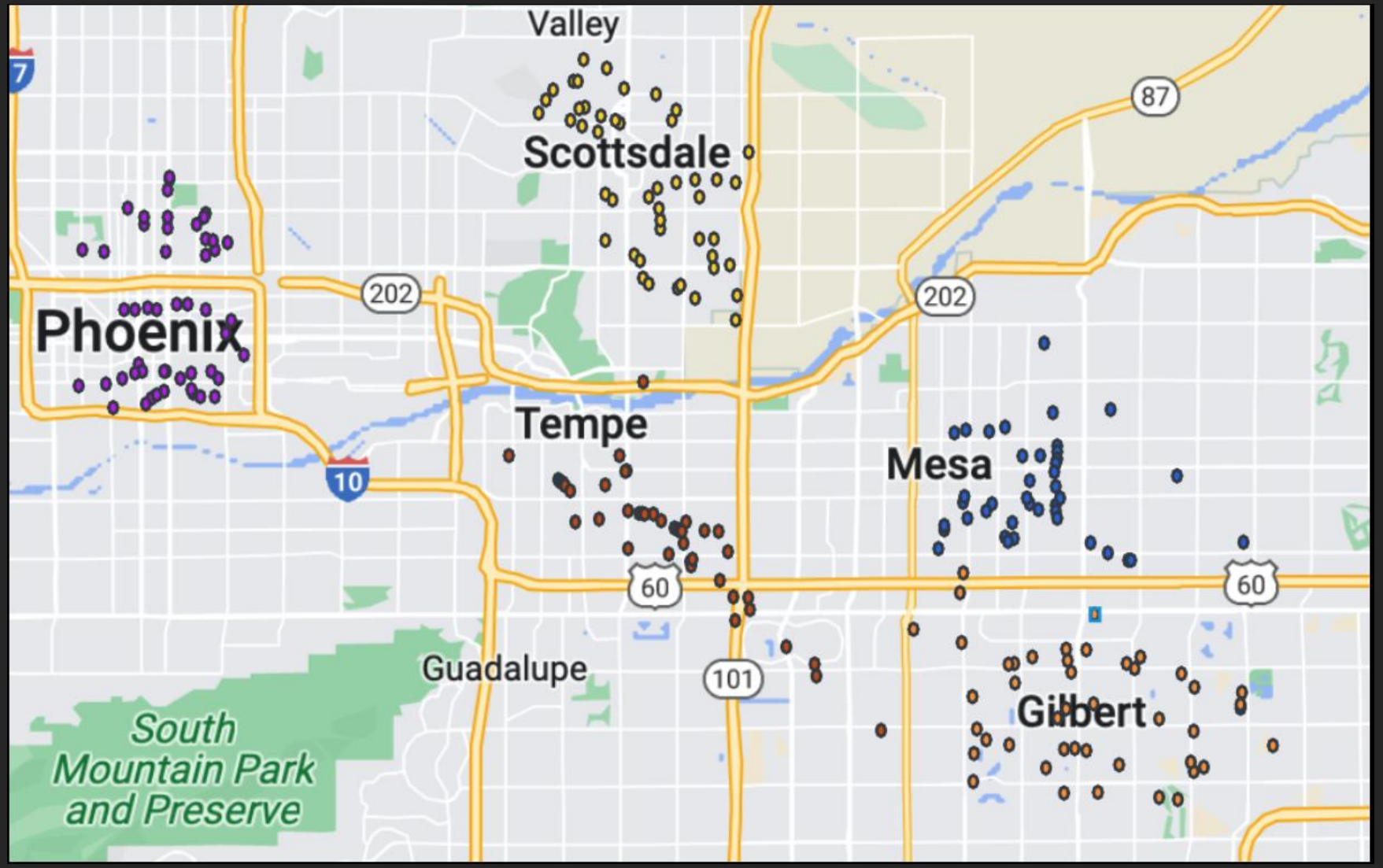
Improved Modeling Method



Improved Modeling Method

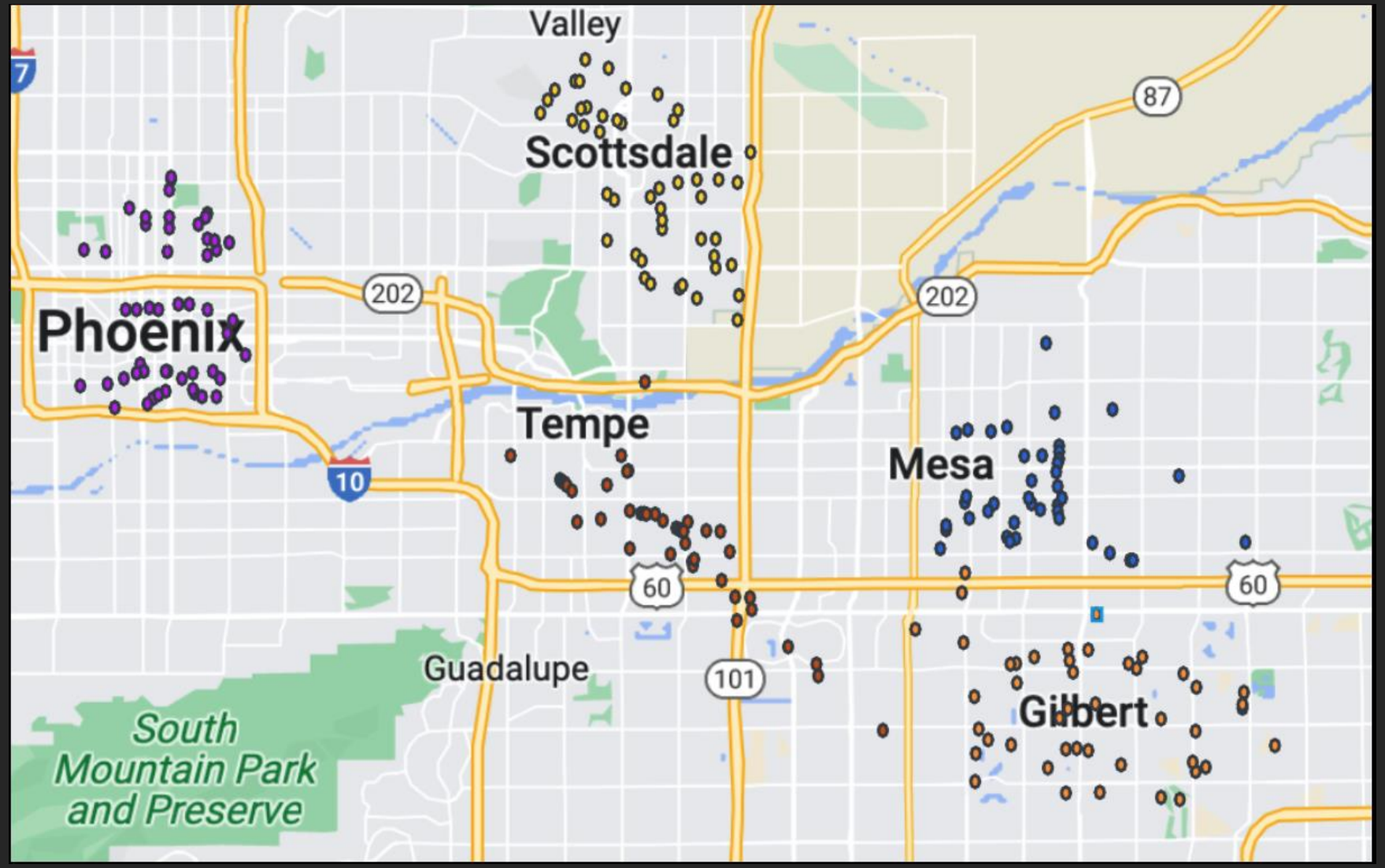


Improved Modeling Method



Mostly businesses: Customers + working professionals

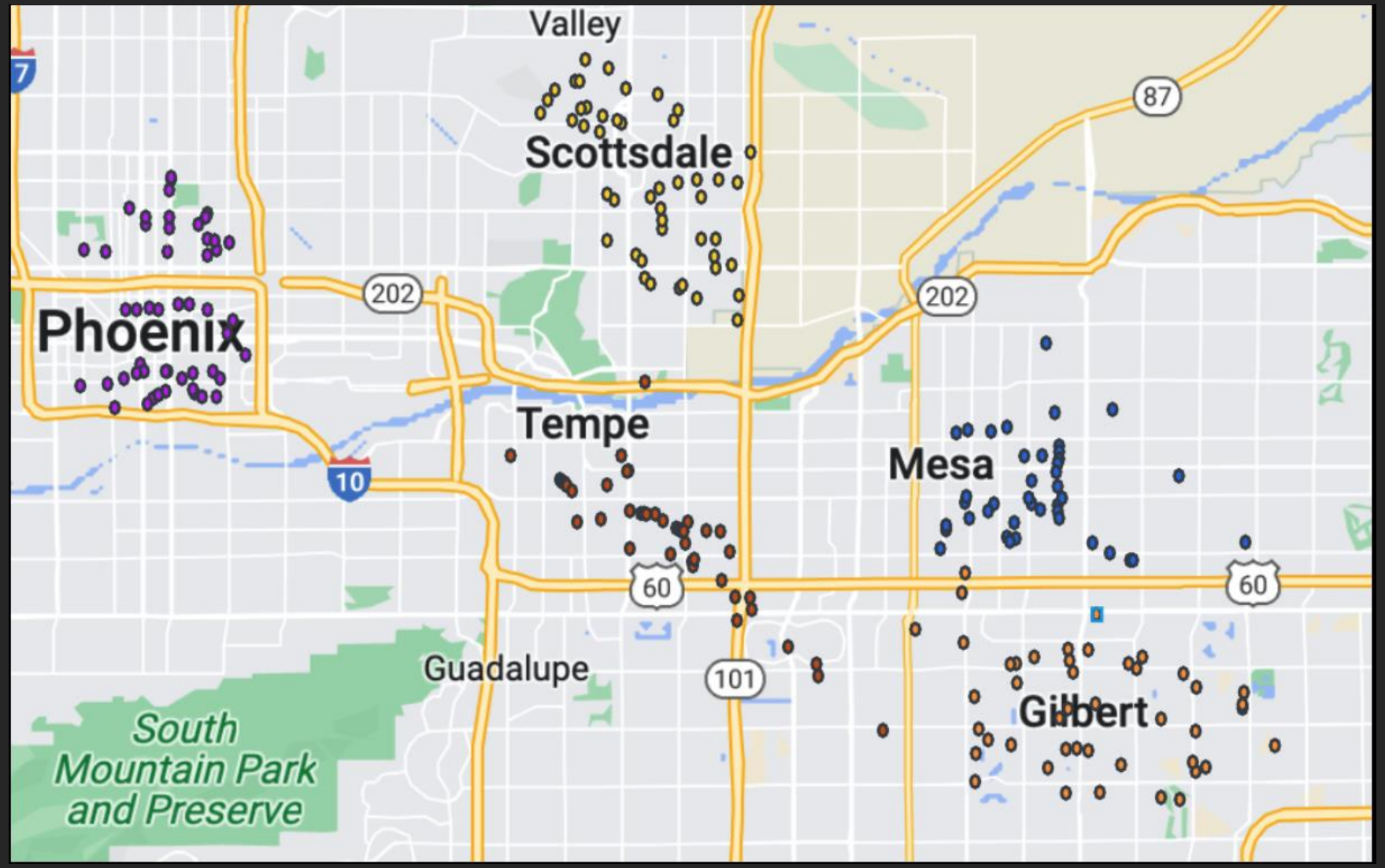
Improved Modeling Method



Mostly businesses: Customers + working professionals

Very wealthy area. Customers would be affluent

Improved Modeling Method

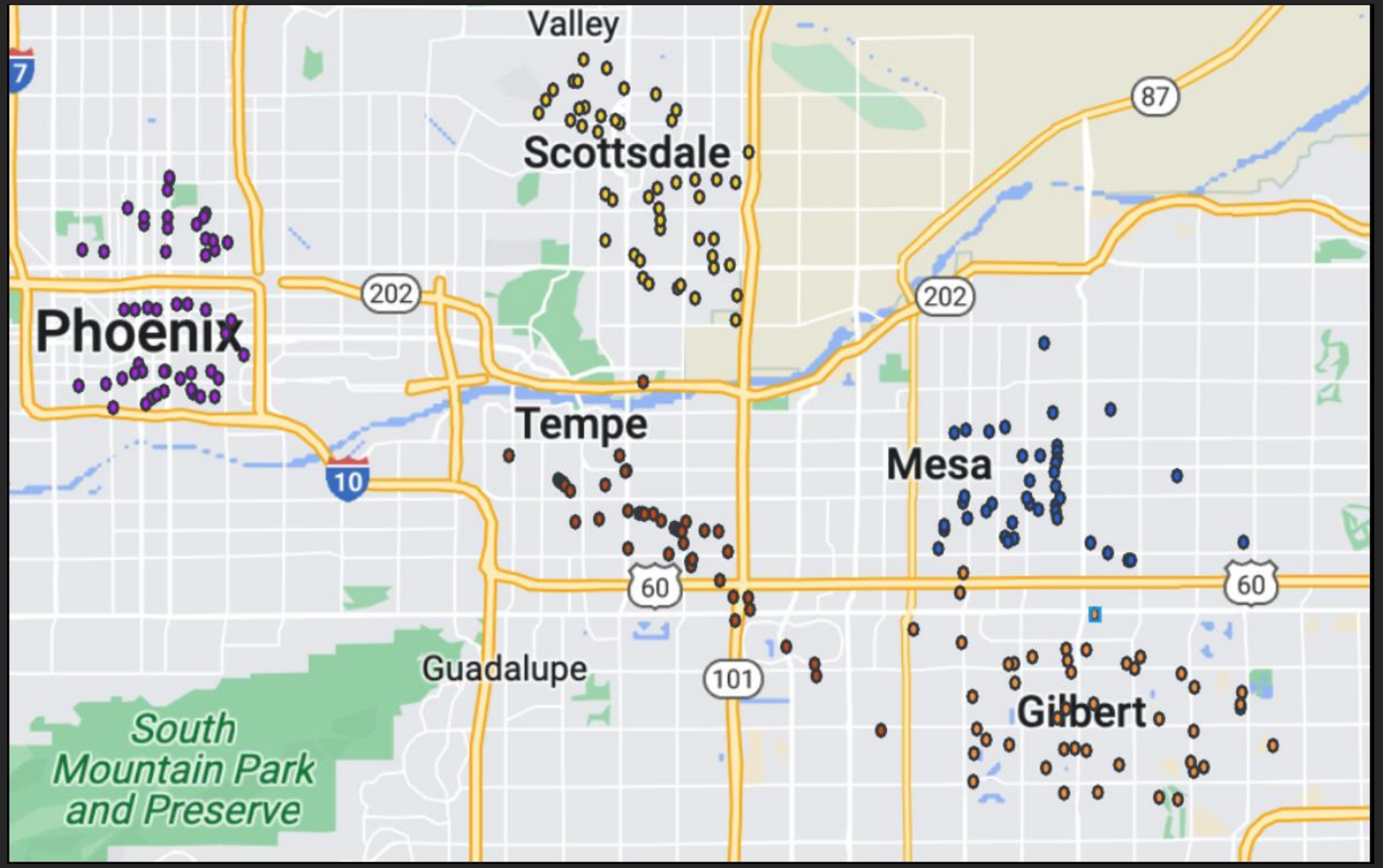


Mostly businesses: Customers + working professionals

Very wealthy area. Customers would be affluent

Customers are middle class full of suburbs

Improved Modeling Method



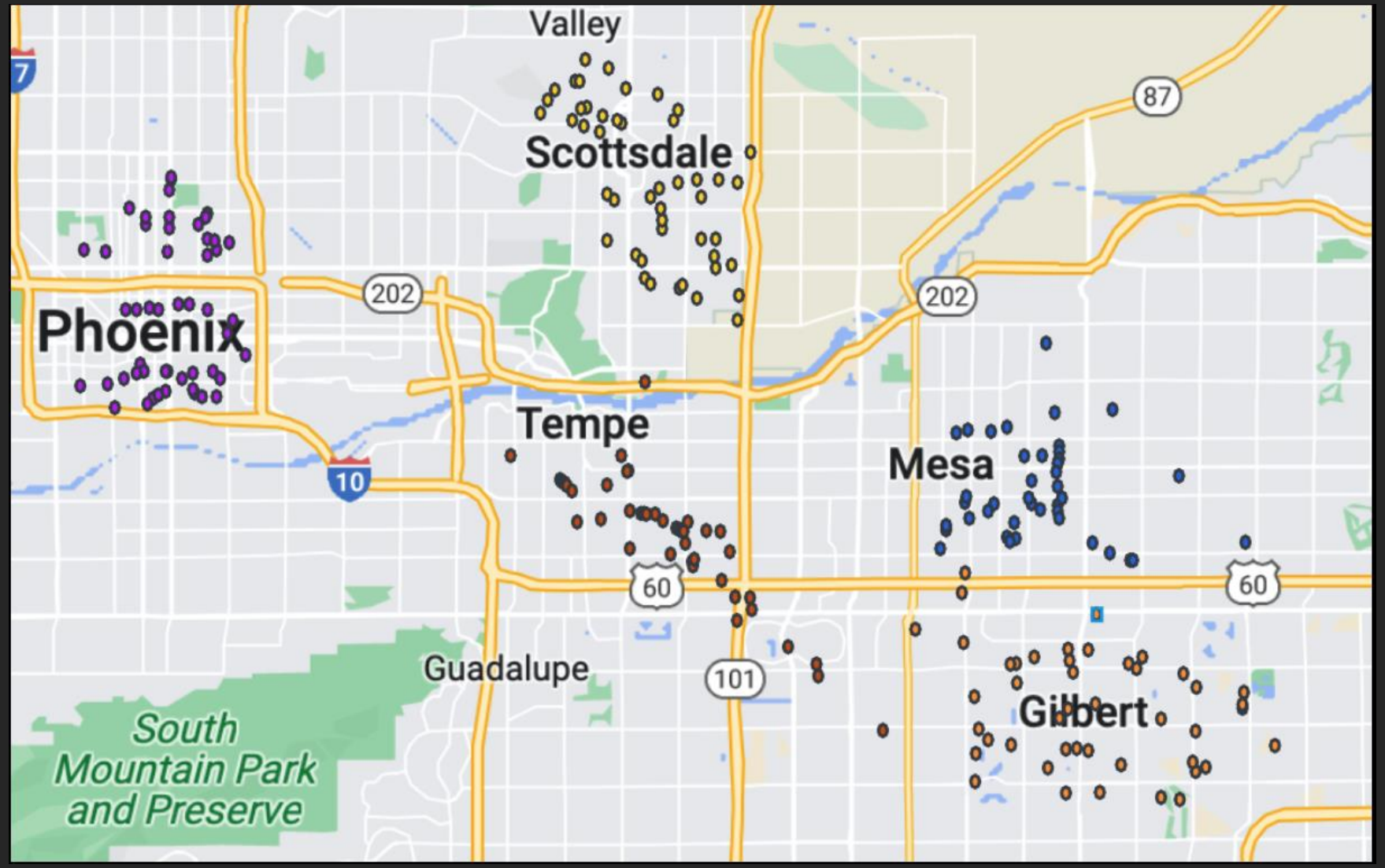
Mostly businesses: Customers + working professionals

Very wealthy area. Customers would be affluent

Customers are middle class full of suburbs

Popular tourist area as well as a large amount of college students and bars

Improved Modeling Method



Mostly businesses: Customers + working professionals

Very wealthy area. Customers would be affluent

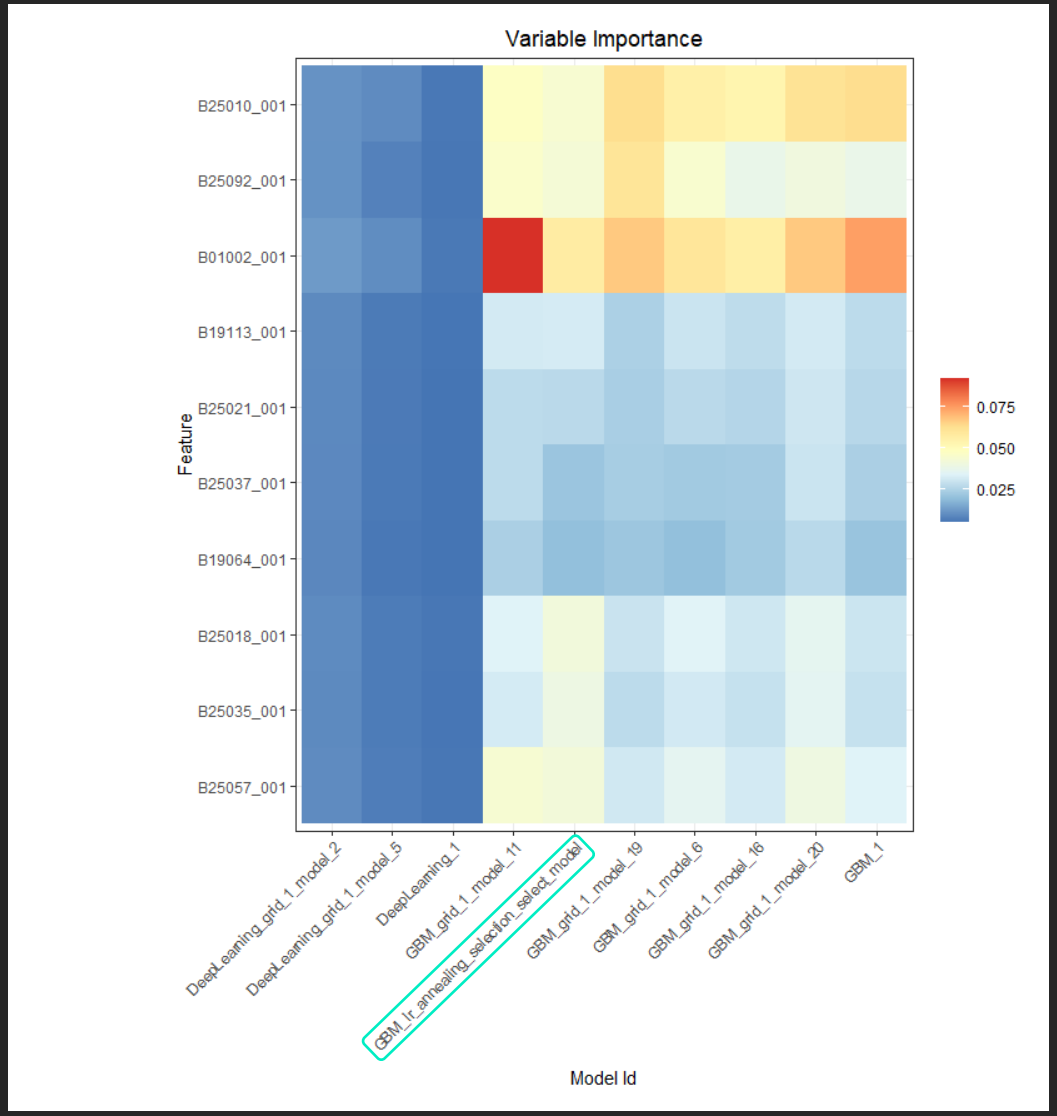
Customers are middle class full of suburbs

Popular tourist area as well as a large amount of college students and bars

Upper-Middle Class. Full of suburbs

Total Profit for EATING/DRINKING with Auto ML

Census Data



- Avg. Household Size
- Median Monthly Owner Costs
- Median Age
- Median Family Income
- Median # of Rooms
- *Median Year Structure by rent/own*
- Aggregate Interest/Dividends/Rental Income
- Median # of Rooms
- Median Year Structure Built
- Lower Contract Rent Quartile ($)

Swires's Target Market Demographic

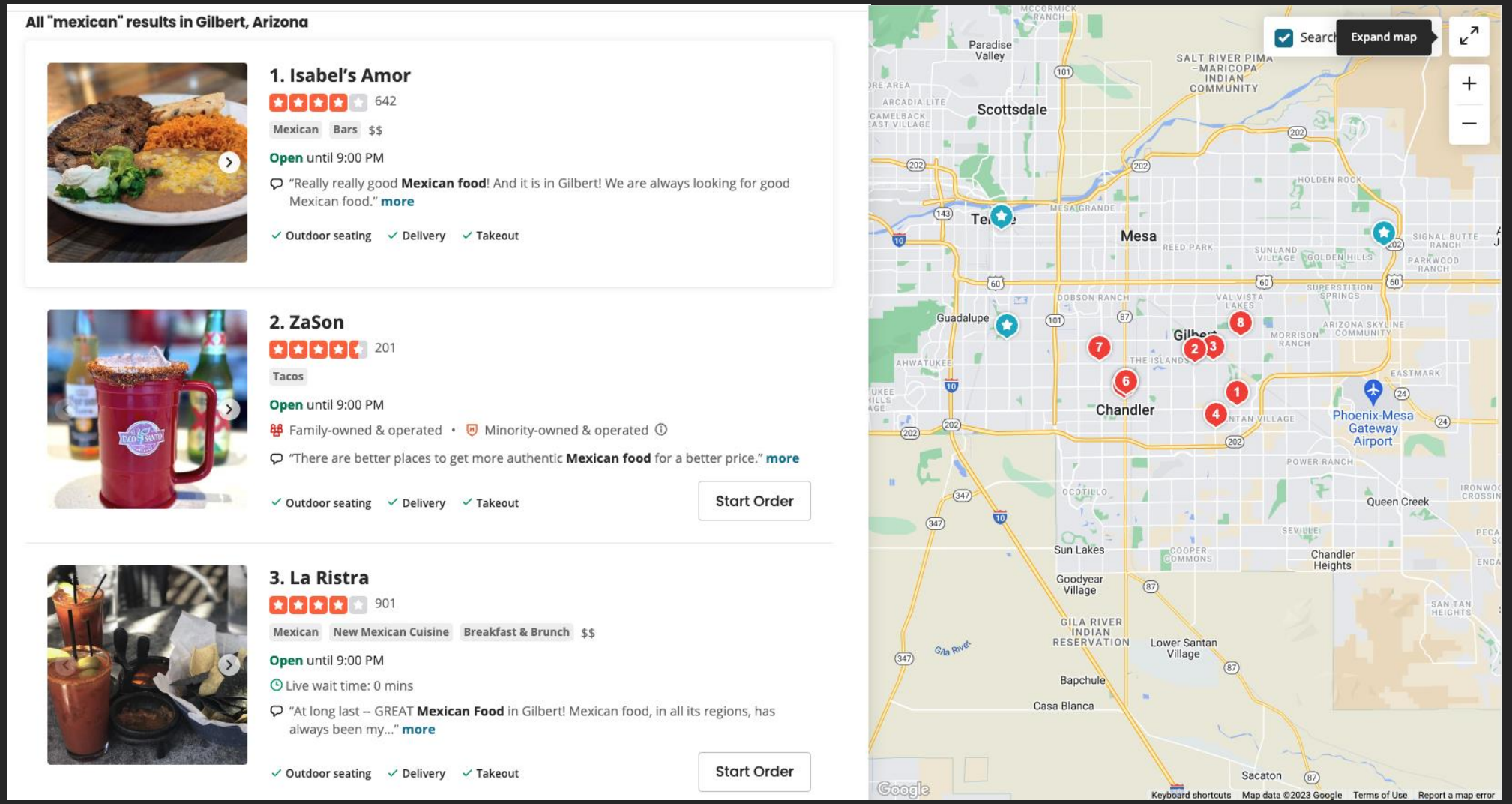
Potential B2B Partner's Demographic

Census Data

AGE	AGE	**75% MATCH**
GENDER	GENDER	**40% MATCH**
MEDIUM INCOME	MEDIUM INCOME	**90% MATCH**

Improved Modeling Method

Yelp Data



Improved Modeling Method

Yelp Data

RESPONSE 200 TRY IT


```
{
  "businesses": [
    {
      "id": "niu9aP1wJ9vru00CJw3mpA",
      "alias": "dominos-pizza-fort-lewis",
      "name": "Domino's Pizza",
      "image_url": "https://s3-media1.fl.yelpcdn.com/bpho
      "is_closed": false,
      "url": "https://www.yelp.com/biz/dominos-pizza-fort
      "review_count": 37,
      "categories": [
        {
          "alias": "pizza",
          "title": "Pizza"
        },
        {
          "alias": "chicken_wings",
          "title": "Chicken Wings"
        },
        {
          "alias": "sandwiches",
          "title": "Sandwiches"
        }
      ],
      "rating": 2,
      "coordinates": {
        "latitude": 47.085672,
        "longitude": -122.571218
      },
      "transactions": [],
      "price": "$",
      "location": {
        "address1": "3168 2nd Division Dr",
        "address2": "",
        "address3": "",
        "city": "Fort Lewis",
        "zip_code": "98433",
        "country": "US",
        "state": "WA"
```


SEO

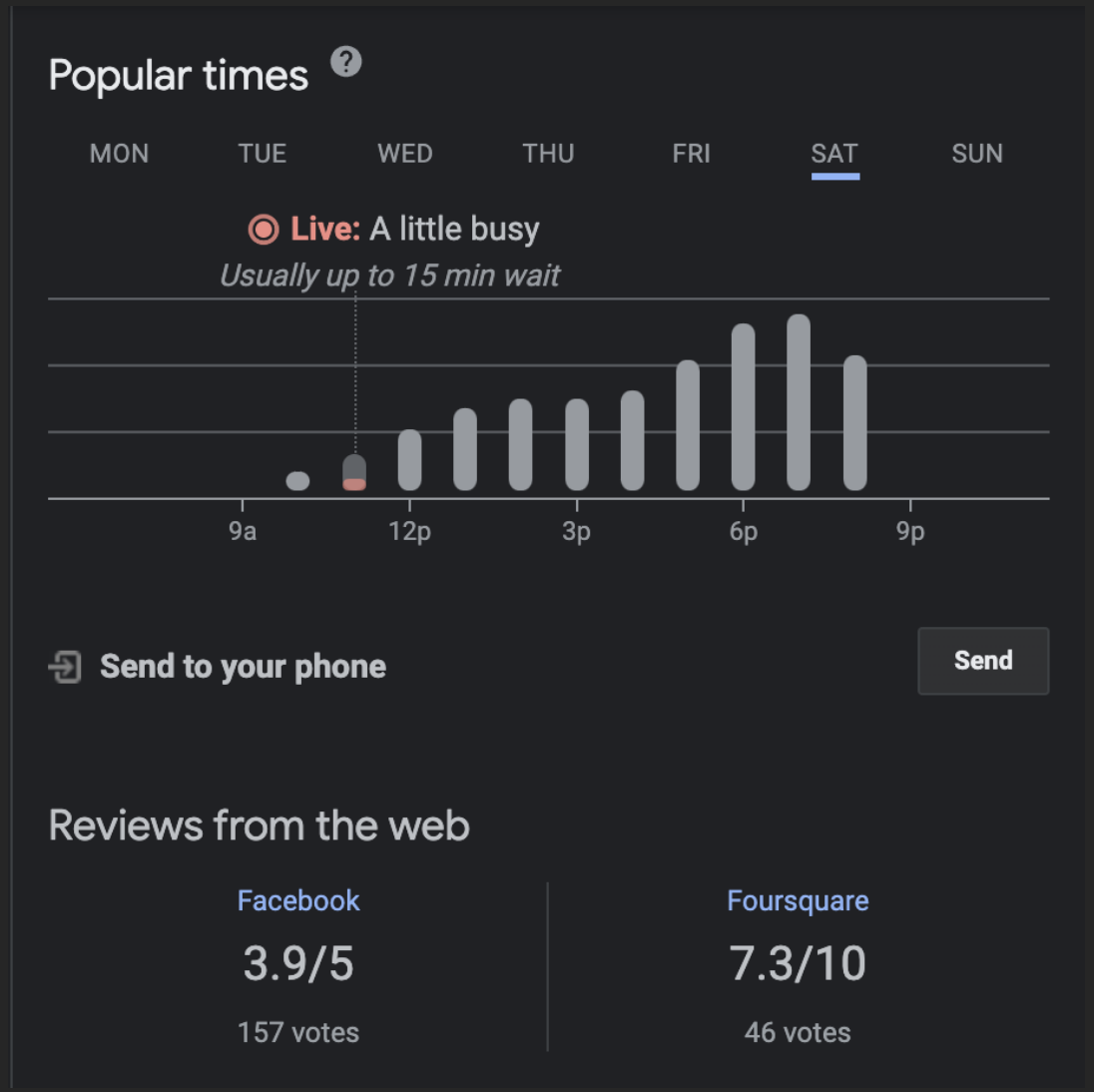
- What is a business doing to promote itself, rank higher is searches?

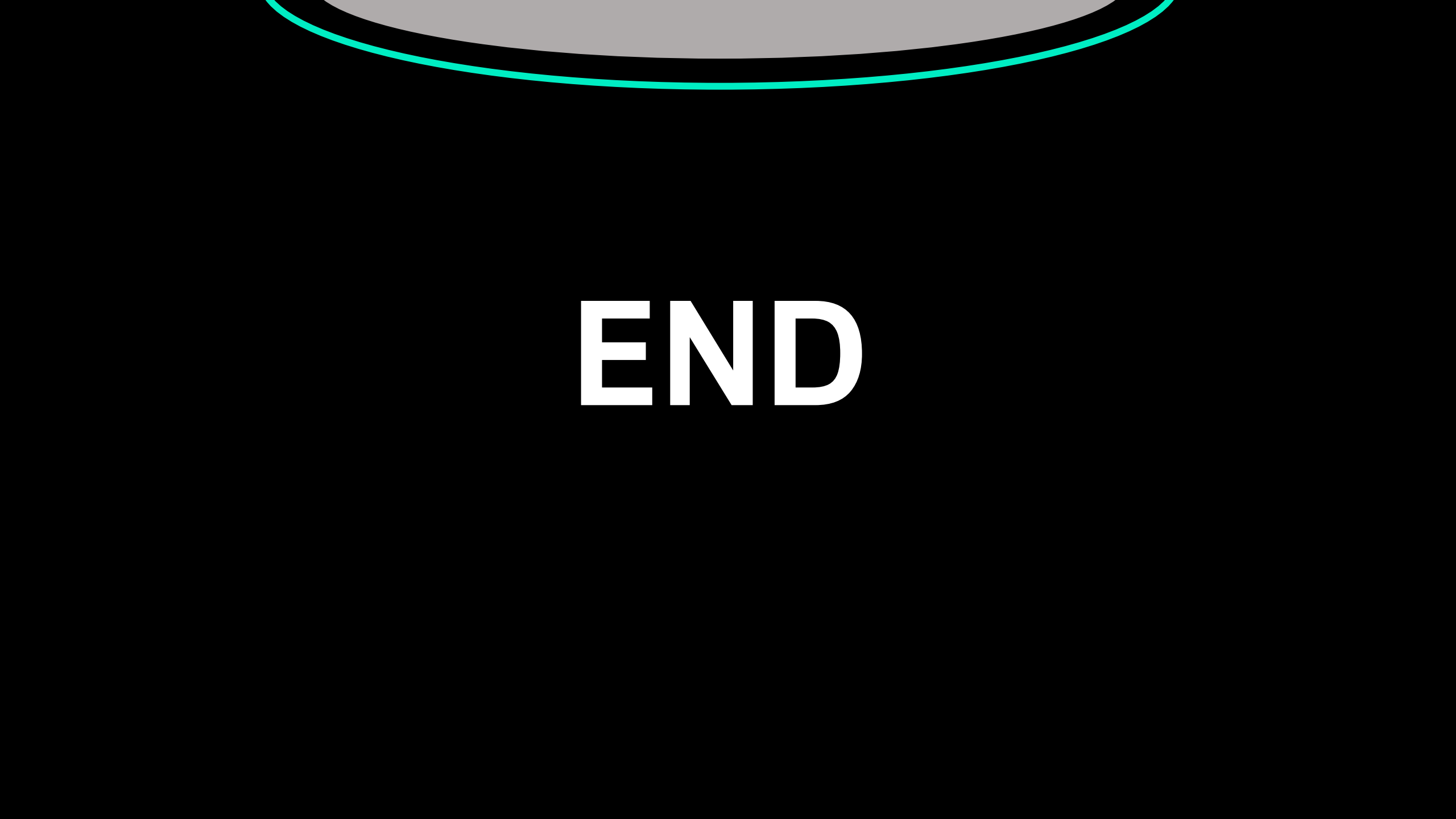
Reputation

- How is the business perceived by the public?
- How often are they getting reviews?

Improved Modeling Method

Google Analytics




END

Q + A