BIG BASKET CASE

Team New York

Executive Summary - Situation

Who is Big Basket

One of the largest online grocery retailers in India

Value proposition is convenience for customers

Current Situation

- Customers tend to order a high number of items per order(+80)
- Order frequency varies significantly between different customers
- Same customer tends to order same items every times
- 30% of their customers ordering on a smartphone

Executive Summary – Problem & Objectives

Problem

- Mobile ordering is not well configured for mobile experience
  - UI is clunky and not personalized
  - Hard to find products and make orders with larger number of items
- Customers "forget" items and tend to place a second order shortly after
- Membership, order and item growth has begun to stagnate

Objectives

- Improve customer convenience through product recommendations
- Improve order profitability through improved order recommendations
- Increase profitability through reducing customer “micro" orders ("did you forget" recommendations)

Big Basket – Analytics Approach

Descriptive Analytics & Validation

Dataset

- March 2011 to Dec 2014
- 62,141 Observations
- 8,387 Orders
- 1,732 Distinct Items
- 106 Members

Customer Deep Dive

- How many items do people order
- Are orders typically the same by customer
- How frequently do customers order
- Who are our best customers

Order Deep Dive

- What were the biggest orders
- When did people order
- How many items did people have
- How have orders changed over time

Recommendation Systems

Generic Methods

- Based on popular historic information

Association Rules Ordered by

- confidence
- support
- lift

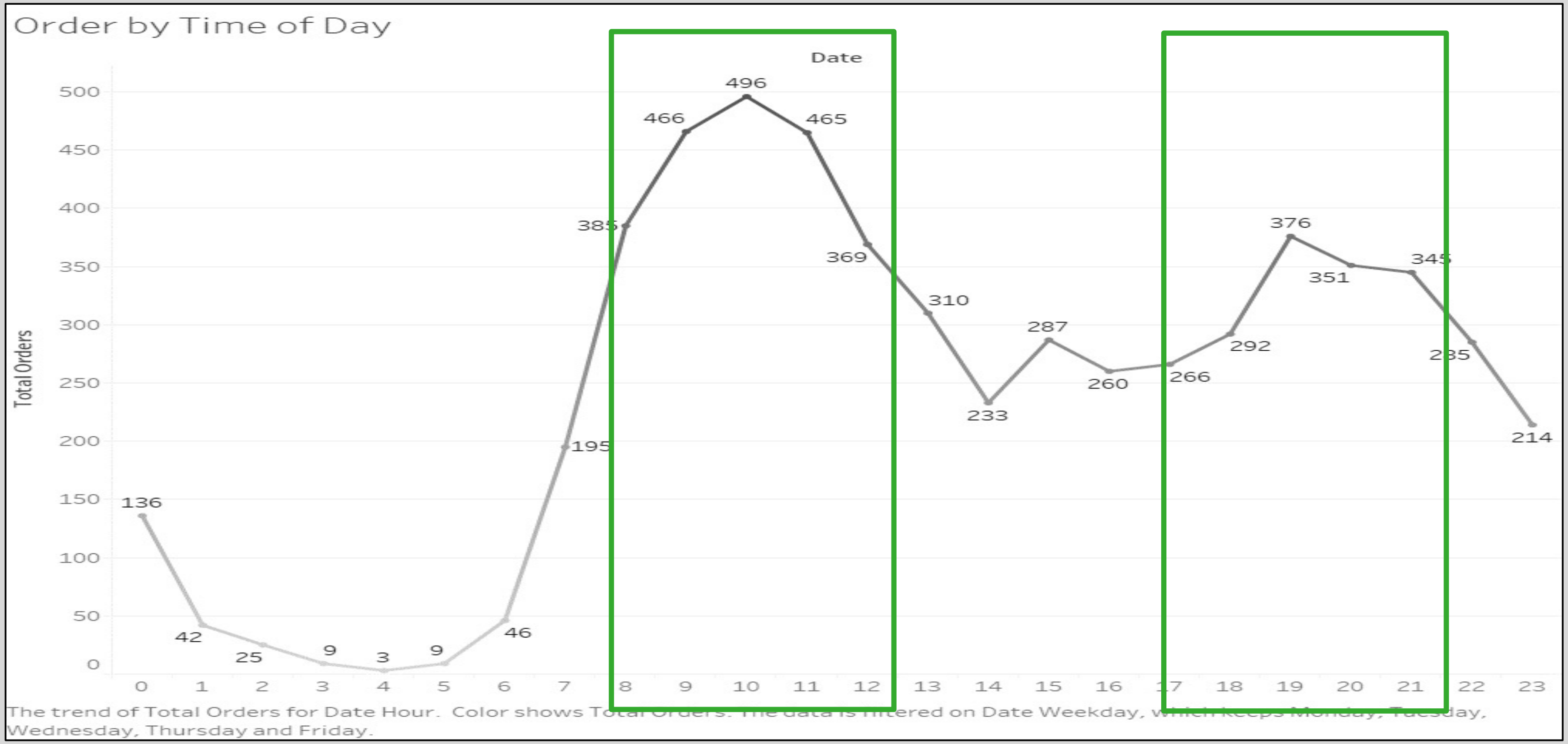
Final Results

ROI calculation

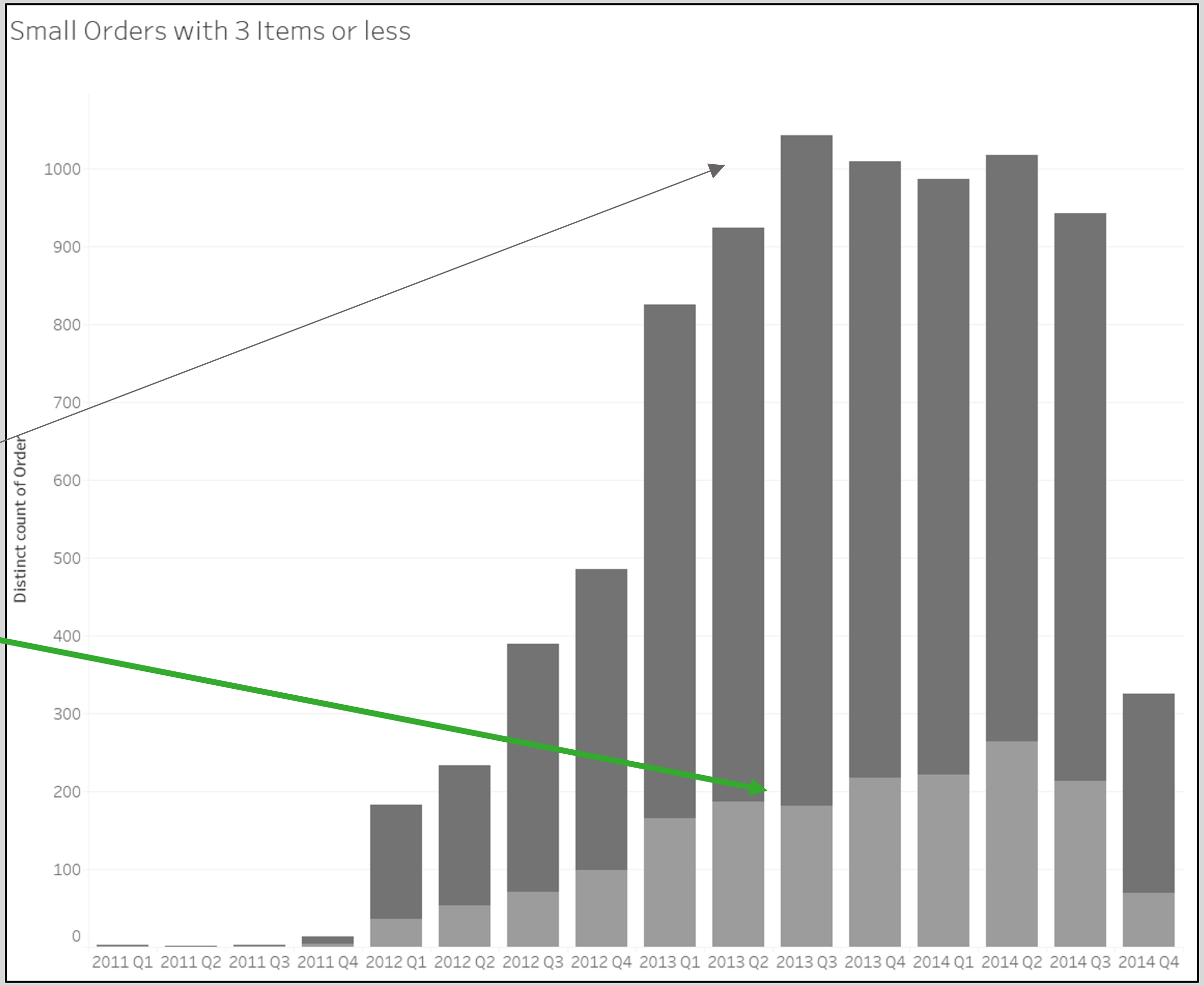
- Average order size
- Overall customer lifetime value

Most Customers Order During their Work Commute

Peak purchasing between
8am-noon
and
5pm-9pm



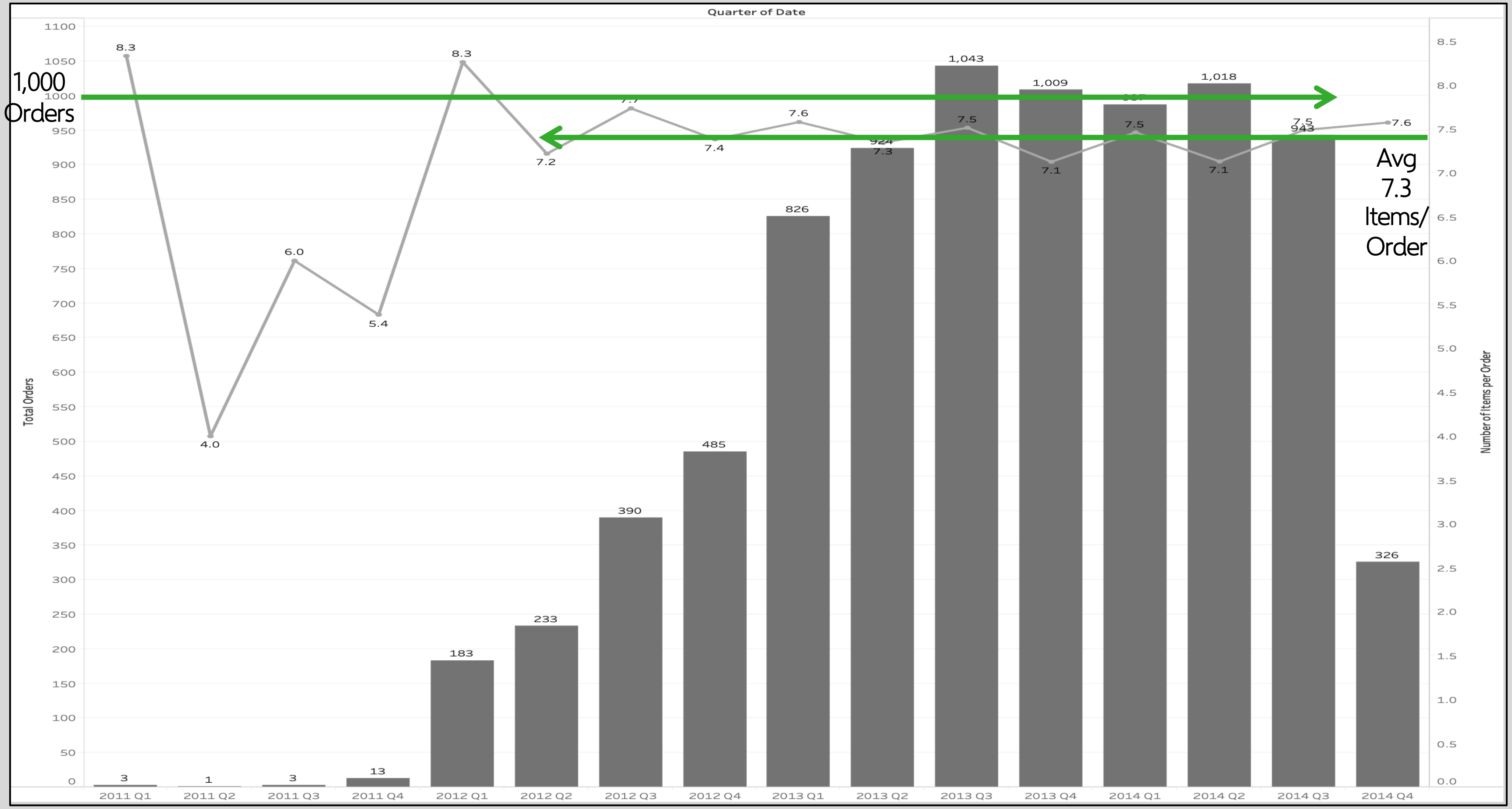
“Mini" Orders Account for 21% of Orders



While number of orders has increased to a plateau the % of “mini” orders (Orders<3 items) has remained constant

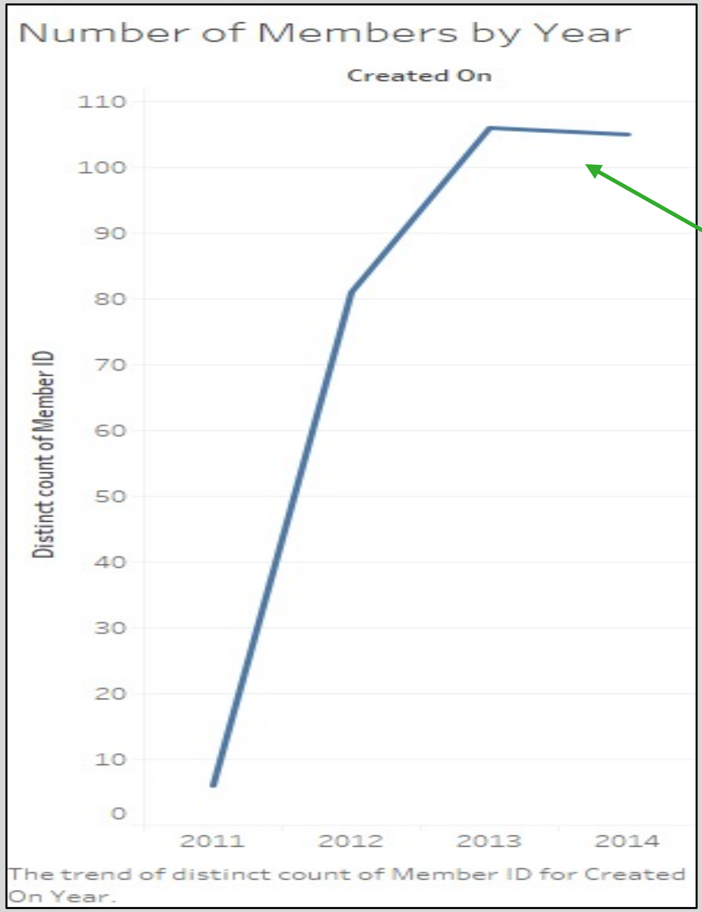
Total Orders & # of Items per Order Plateauing

Total number of orders has been stable around 1,000 orders / quarter



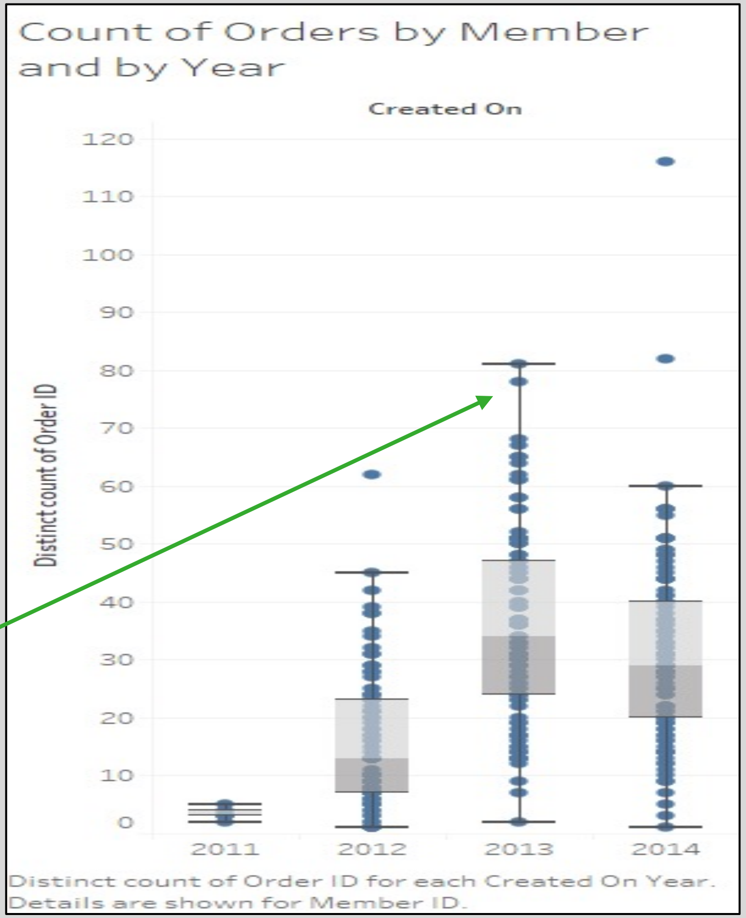
# of Items per Order has consistently been between 7.2 and 7.4 orders

Overall Membership is Plateauing as is # of Orders/Member



Unique number of members has increased consistently YoY except for 2014

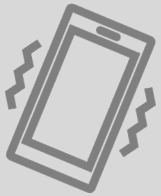
Number of orders per unique member by year has begun to plateau



WHERE WILL FUTURE GROWTH COME FROM?

Recommender Systems For Order Convenience & Growth

Customer Journey Through App



Opens and Searches In App

Fills Basket with Relevant Items (from rec. or not)

Begins to Check Out

PHASE I

Recommend Frequently Purchased Items and “Last Order”

PHASE II

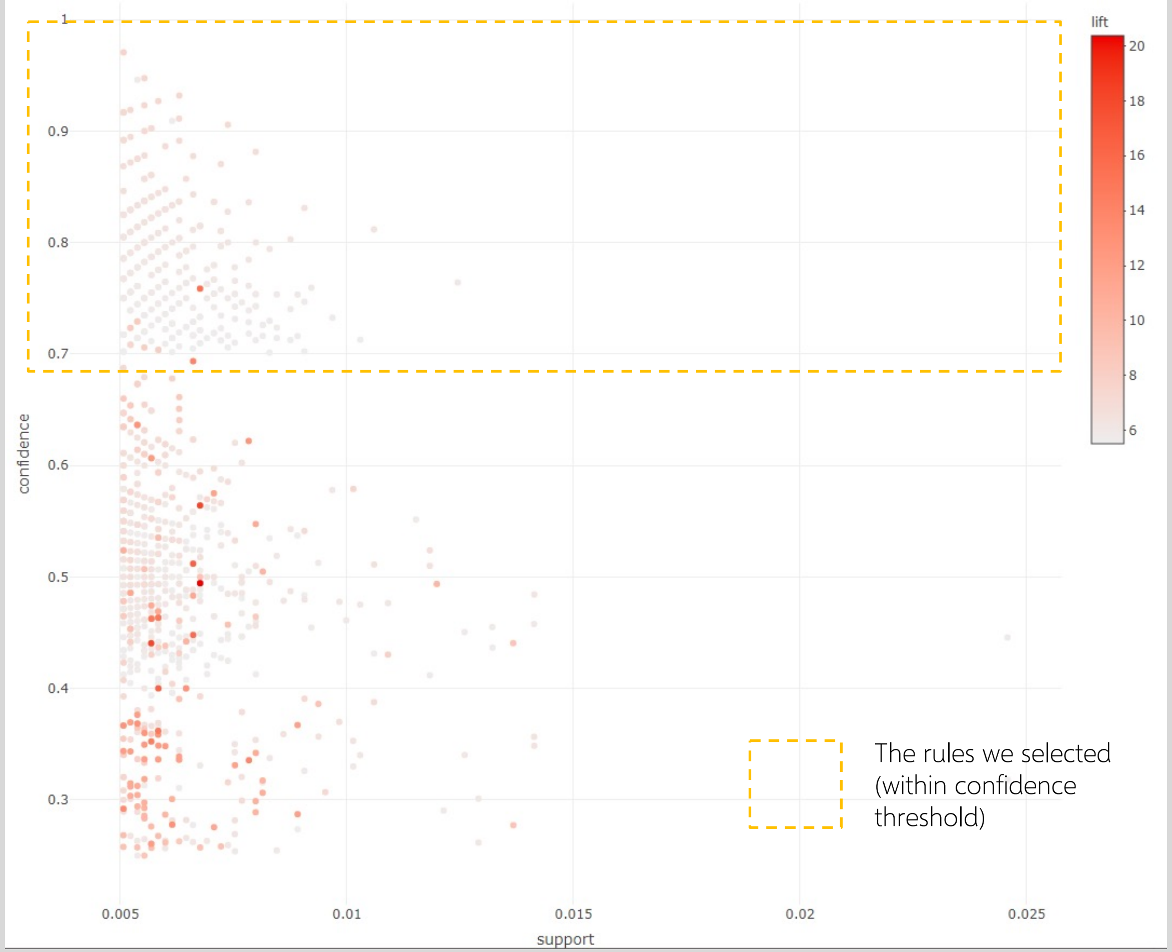
Generic Recommendations based on association rules

PHASE III

Recommendations based on association rules & “mini order” items

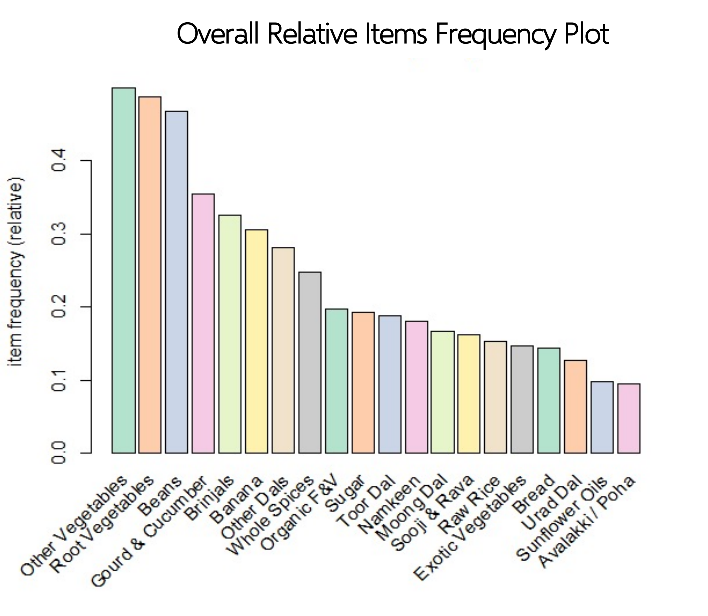
Building Recommender Models

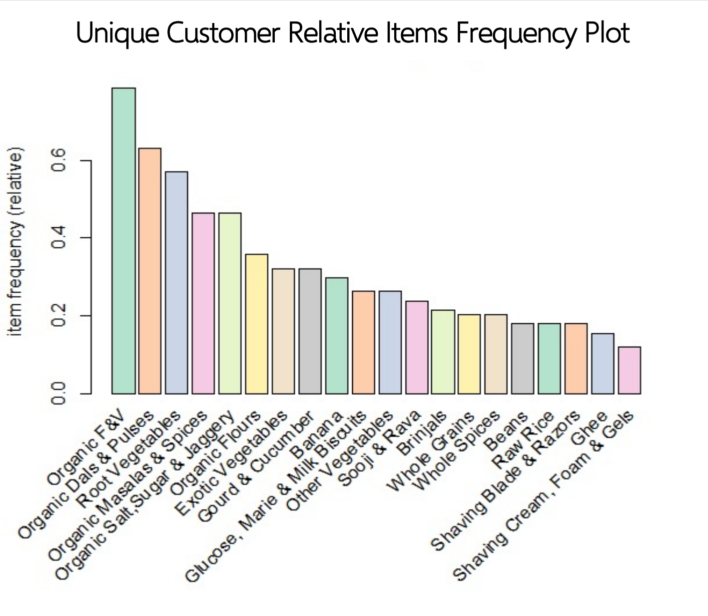
- Combined any orders placed by same member during same week
  - Anecdotal experience purchasing groceries 1-2 times a week
  - Knowledge that organics and vegetables (the most commonly purchased items among our customer sample) have an expiry of ~ 1 week or less.
- Modelling from the product description level not SKU
  - Created stronger rules
  - Opportunity to promote brands that are sponsored/partnered
- Built two association rules models
  - Based on overall orders of our entire sample dataset
  - Based on an example of an individual customer-level model to provide more personalized rules
  - Minimum Support 0.005, Confidence > 0.7, Lift > 1



Phase I: Recommend Items Based on Frequently Ordered Historical Items

- Brings historical/frequently purchased items to the forefront of the app
- Quick actions to repurchase previous baskets
  - Increase convenience and promote customer habitual behaviour
- First intercept point for promoting sponsor/partner branded items



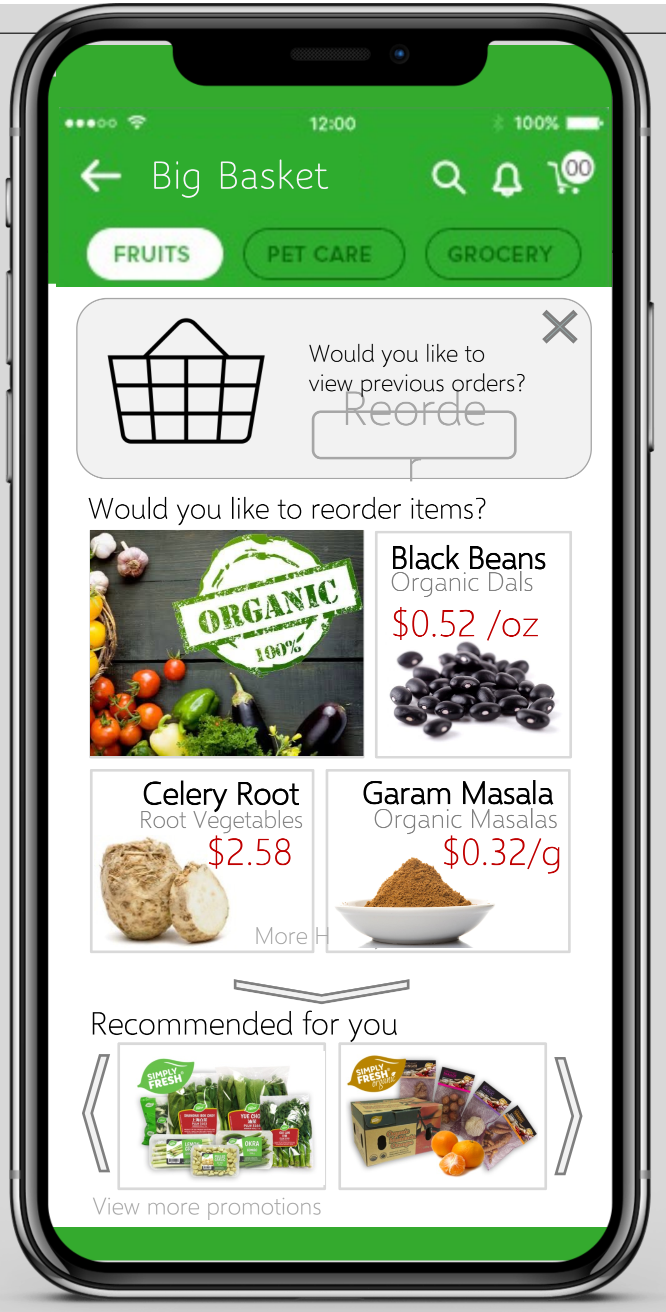


Strong Call to Action

Encourages Purchase

Populated with Individual Customer Preferences/ Previous Order

Reduces Search Time



Recommendations Based on Other Customers Frequent Items

Increases Basket Size

Phase I: Mock-Up

Phase II: Active Recommendations Based on Items Added to Customers Shopping Cart

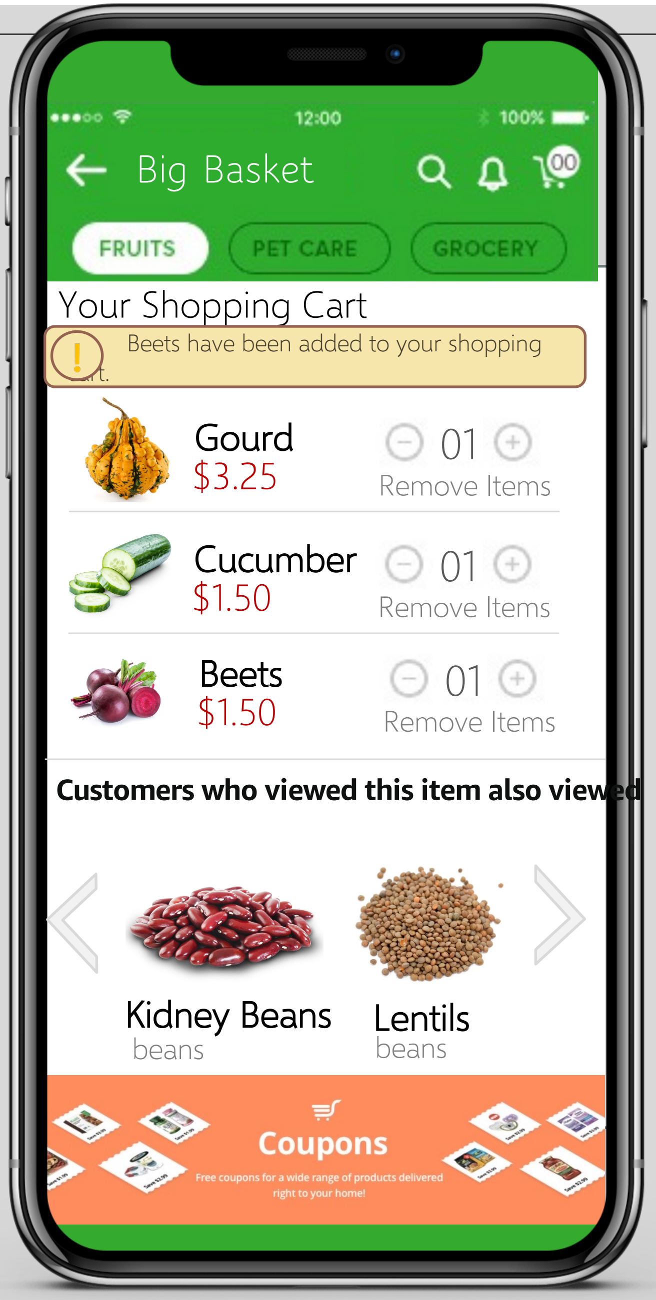
- Promotes larger carts by recommending items based on other customers
- Reduces customer search time in app
- Second intercept point for promoting sponsor/partner branded items

Sample: Ranked Associations Rules Based on All Customers

LHS		RHS	support	confidence	coverage	lift
{Gourd & Cucumber,Other Vegetables}	=>	{Beans}	0.17	0.71	0.25	1.51
{Beans,Gourd & Cucumber}	=>	{Other Vegetables}	0.17	0.75	0.23	1.51
{Gourd & Cucumber,Root Vegetables}	=>	{Other Vegetables}	0.16	0.74	0.22	1.48
{Gourd & Cucumber,Root Vegetables}	=>	{Beans}	0.15	0.70	0.22	1.50
{Beans,Brinjals}	=>	{Other Vegetables}	0.15	0.72	0.20	1.43
{Brinjals,Root Vegetables}	=>	{Other Vegetables}	0.14	0.71	0.20	1.42
{Brinjals,Gourd & Cucumber}	=>	{Other Vegetables}	0.13	0.74	0.18	1.49
{Gourd & Cucumber,Other Vegetables,Root Vegetables}	=>	{Beans}	0.12	0.74	0.16	1.59

Organized by highest lift and customer preferences (have they purchased before)

Reduces Search Time



Populated with popular items

Increases Basket Size

Phase II: Mock-Up

Phase III: Active Recommendations Based on Items Added to Customers Shopping Cart

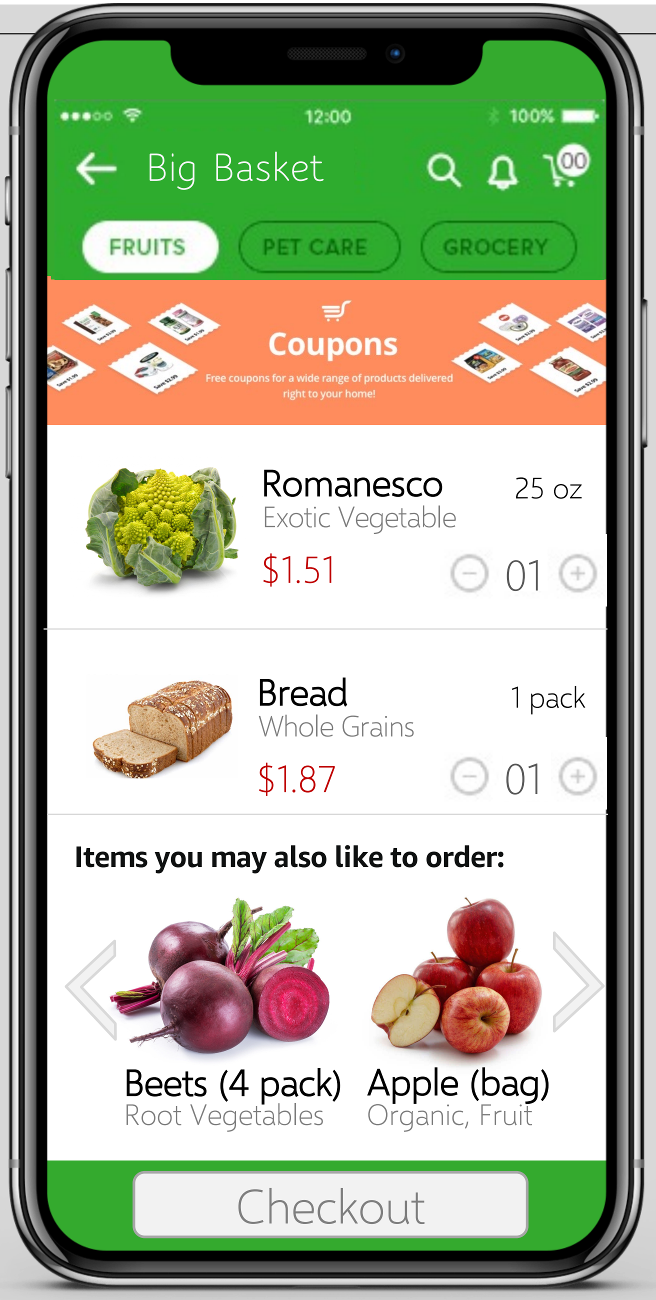
- Based on association rules unique to the customers
- Reduce/prevent micro-ordering (forgetting and item) with the "Did you forget" phase
- Final touch point to increase order size
- Promotes customer convenience (no one is happy if the customer forgets something and needs to place another order, we pay separate delivery costs and the customer is inconvenienced.

Sample: Ranked Associations Rules Based on Individual Customers

LHS		RHS	Support	Confidence	Coverage	Lift
{Exotic Vegetables,Whole Grains}	=>	{Brinjals}	0.06	0.56	0.11	2.59
{Exotic Vegetables,Whole Grains}	=>	{Other Vegetables}	0.06	0.56	0.11	2.12
{Exotic Vegetables,Whole Grains}	=>	{Organic Flours}	0.06	0.56	0.11	1.56
{Exotic Vegetables,Whole Grains}	=>	{Organic Masalas & Spices}	0.06	0.56	0.11	1.20

Populates with items based on individual customer association rules ranked by lift, support & confidence

Reduces “mini” orders



Phase III: Mock-Up

Estimated Results of Implementation

Customer Journey Through App

Opens and Logs Into App

PHASE I

Recommend Frequently Purchased Items and "Last Order"

Fills Basket with Relevant Items (from rec. or not)

PHASE II

Generic Recommendations based on association rules

Begins to Check Out

PHASE III

Recommendations based on association rules & "mini orders"

Return on Investment

Improved Convenience for Customers

Improved Order Profitability

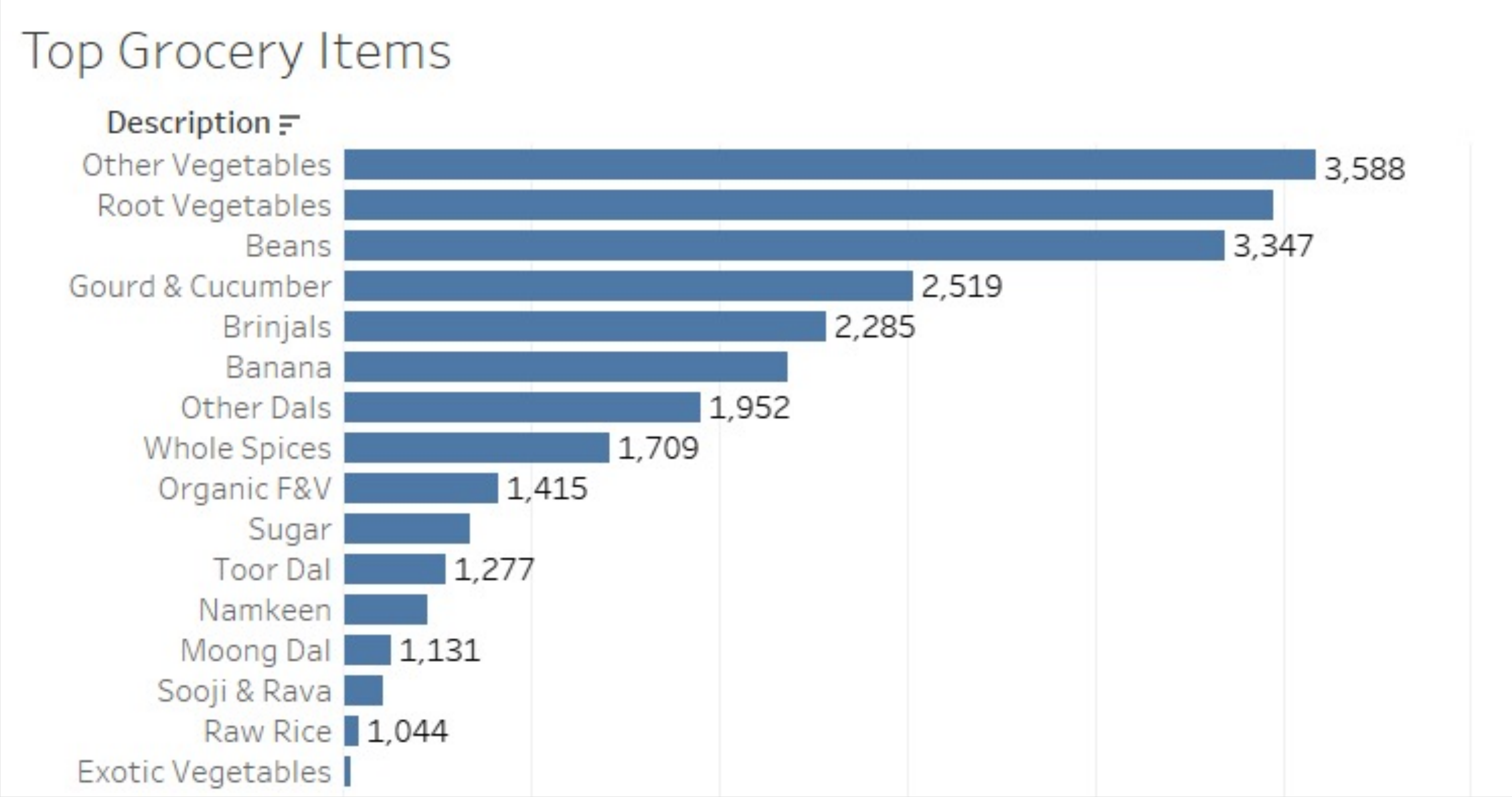
Reduced Supply Chain Costs

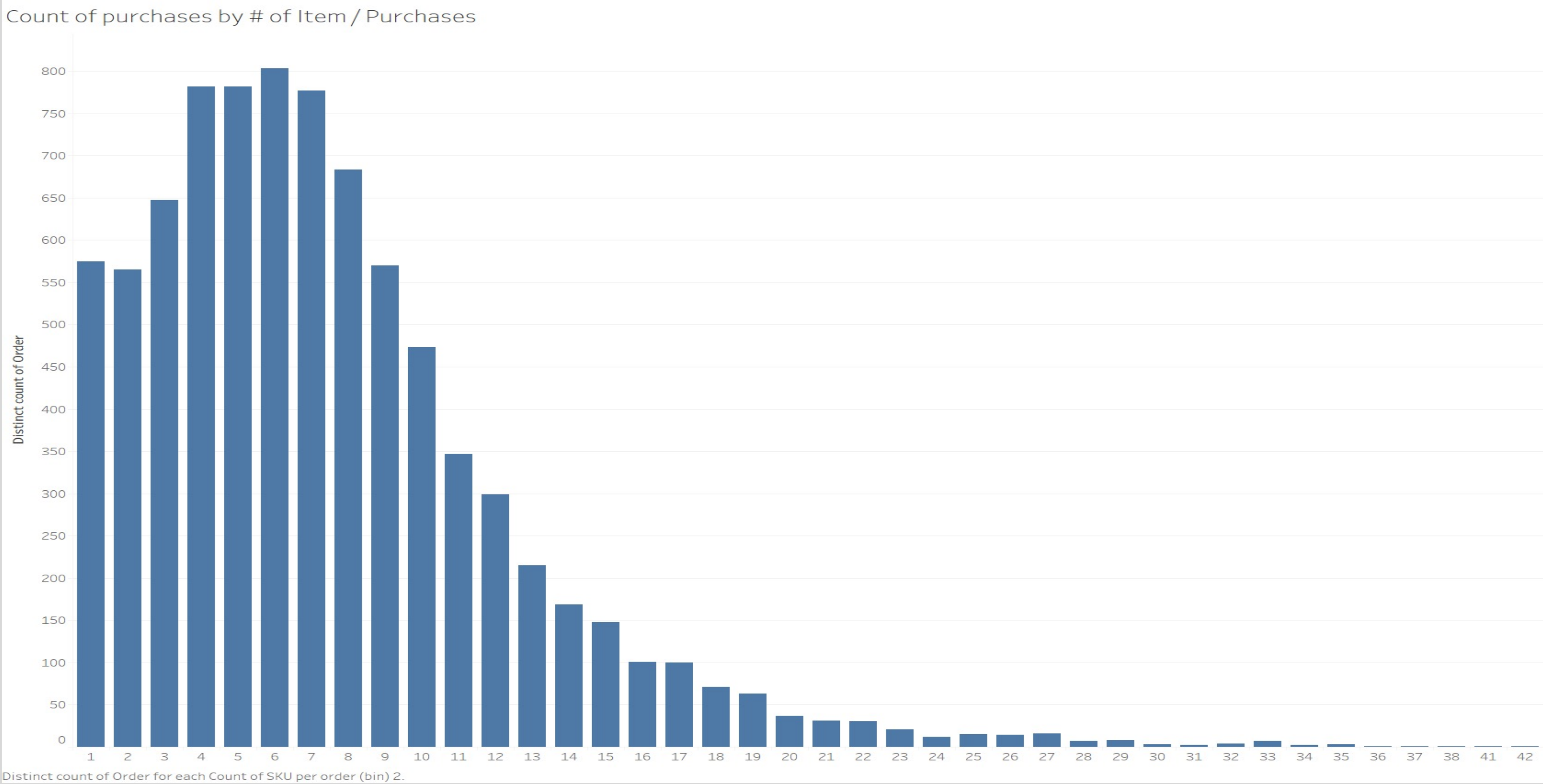
THANK YOU!

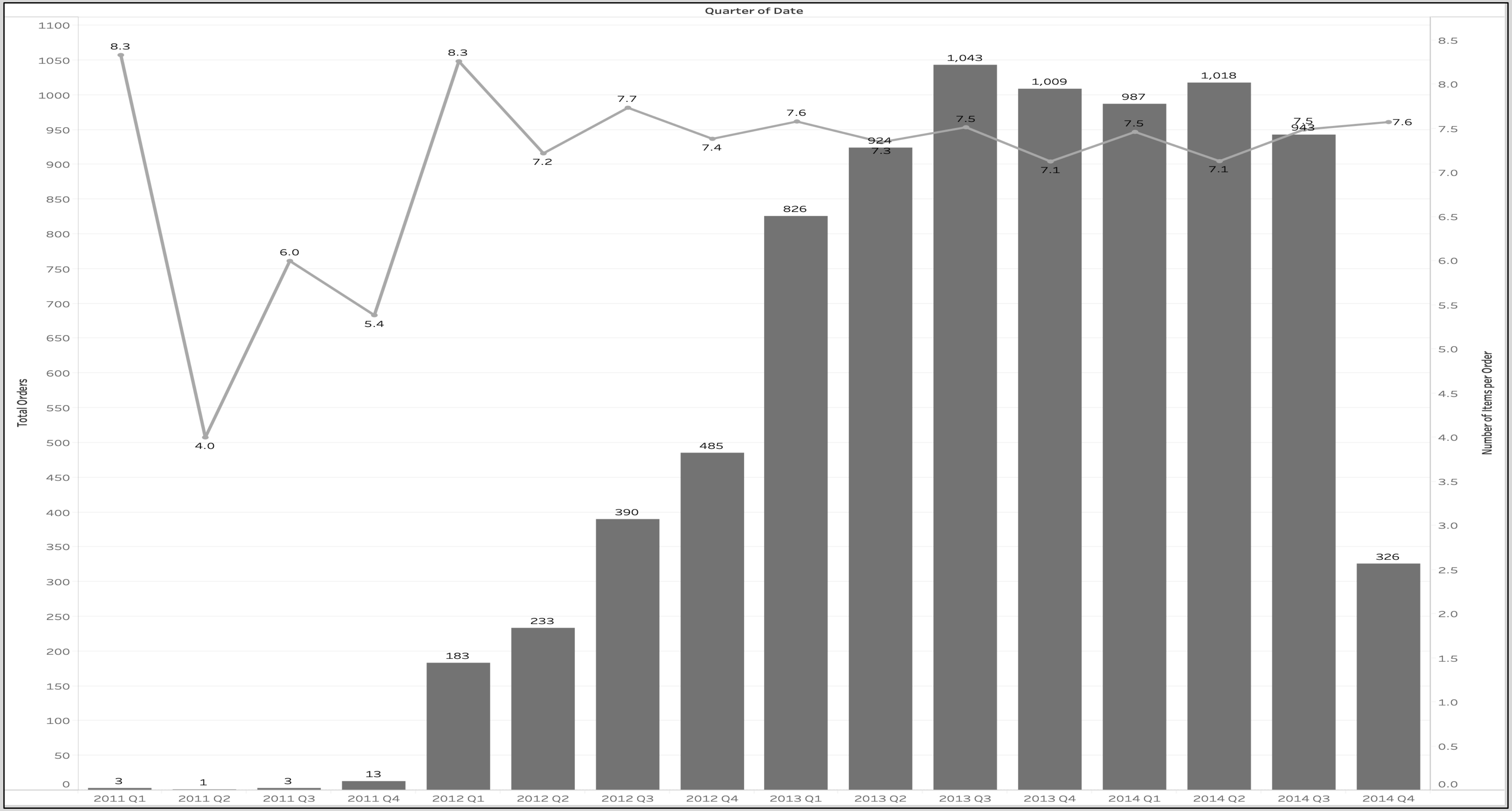
Team New York

APPENDIX

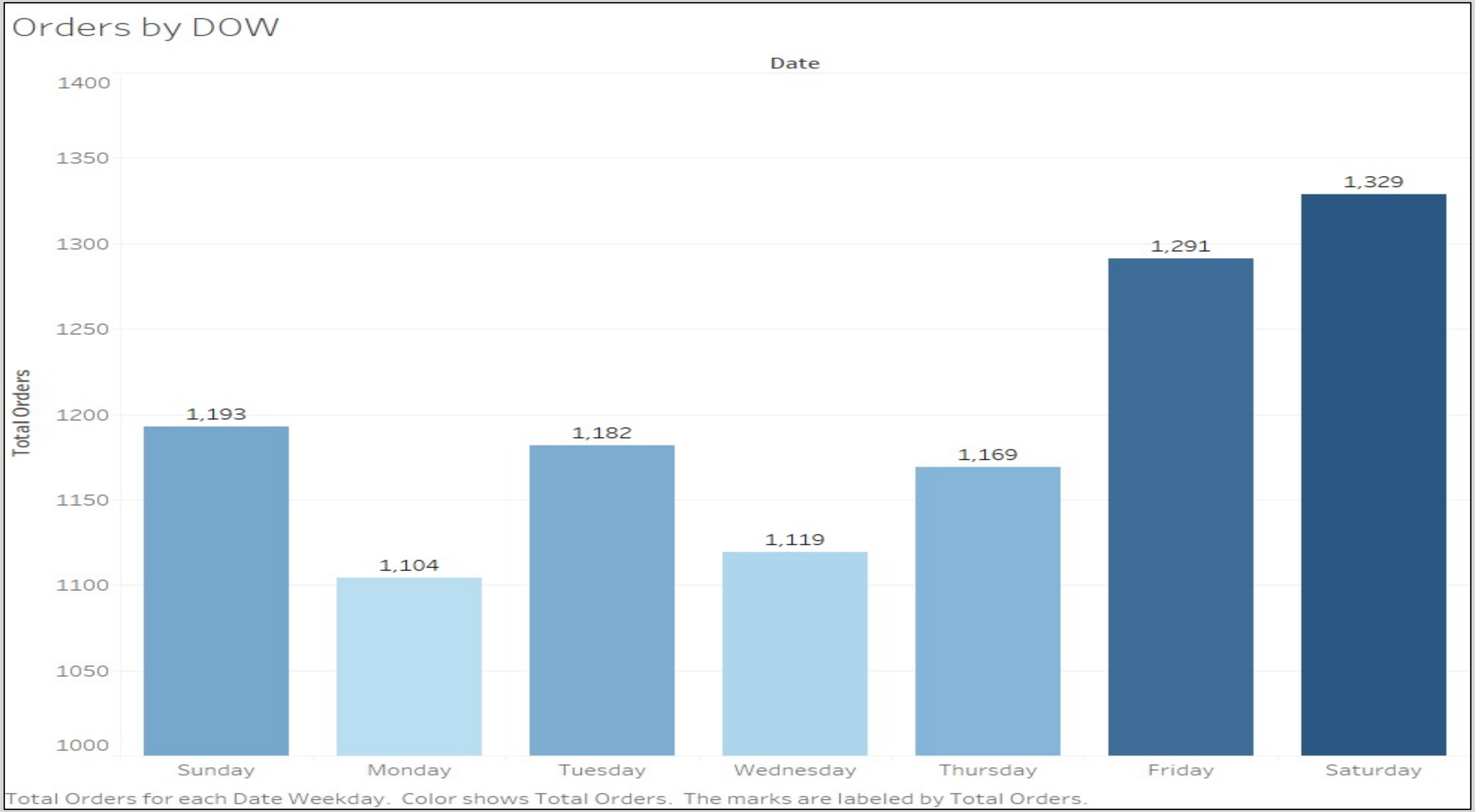
Most Frequently Ordered Items



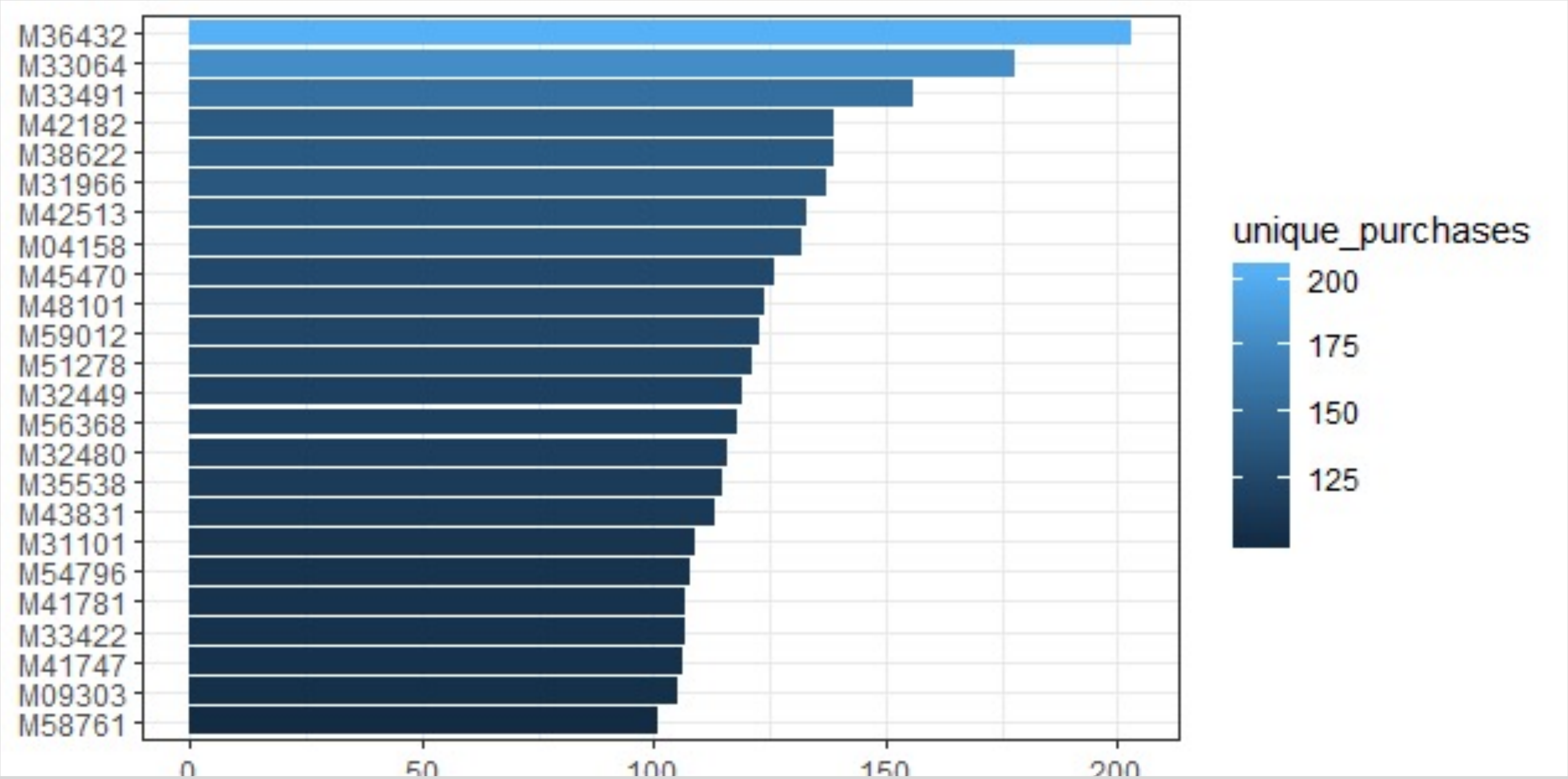

Count of purchases by # of Item / Purchases
Distinct count of Order
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 41 42
Distinct count of Order for each Count of SKU per order (bin) 2.



Quarter of Date
Total Orders
Number of Items per Order
2011 Q1
2011 Q2
2011 Q3
2011 Q4
2012 Q1
2012 Q2
2012 Q3
2012 Q4
2013 Q1
2013 Q2
2013 Q3
2013 Q4
2014 Q1
2014 Q2
2014 Q3
2014 Q4
3
1
3
13
183
233
390
485
826
924
1,043
1,009
987
1,018
943
326
8.3
4.0
6.0
5.4
8.3
7.2
7.7
7.4
7.6
7.3
7.5
7.1
7.5
7.1
7.5
7.6


Orders by DOW
Date
Total Orders
1400
1350
1300
1250
1200
1150
1100
1050
1000
1,193
1,104
1,182
1,119
1,169
1,291
1,329
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Total Orders for each Date Weekday. Color shows Total Orders. The marks are labeled by Total Orders.

Top Customers With Most Orders Per Week



Recommender System Parameters

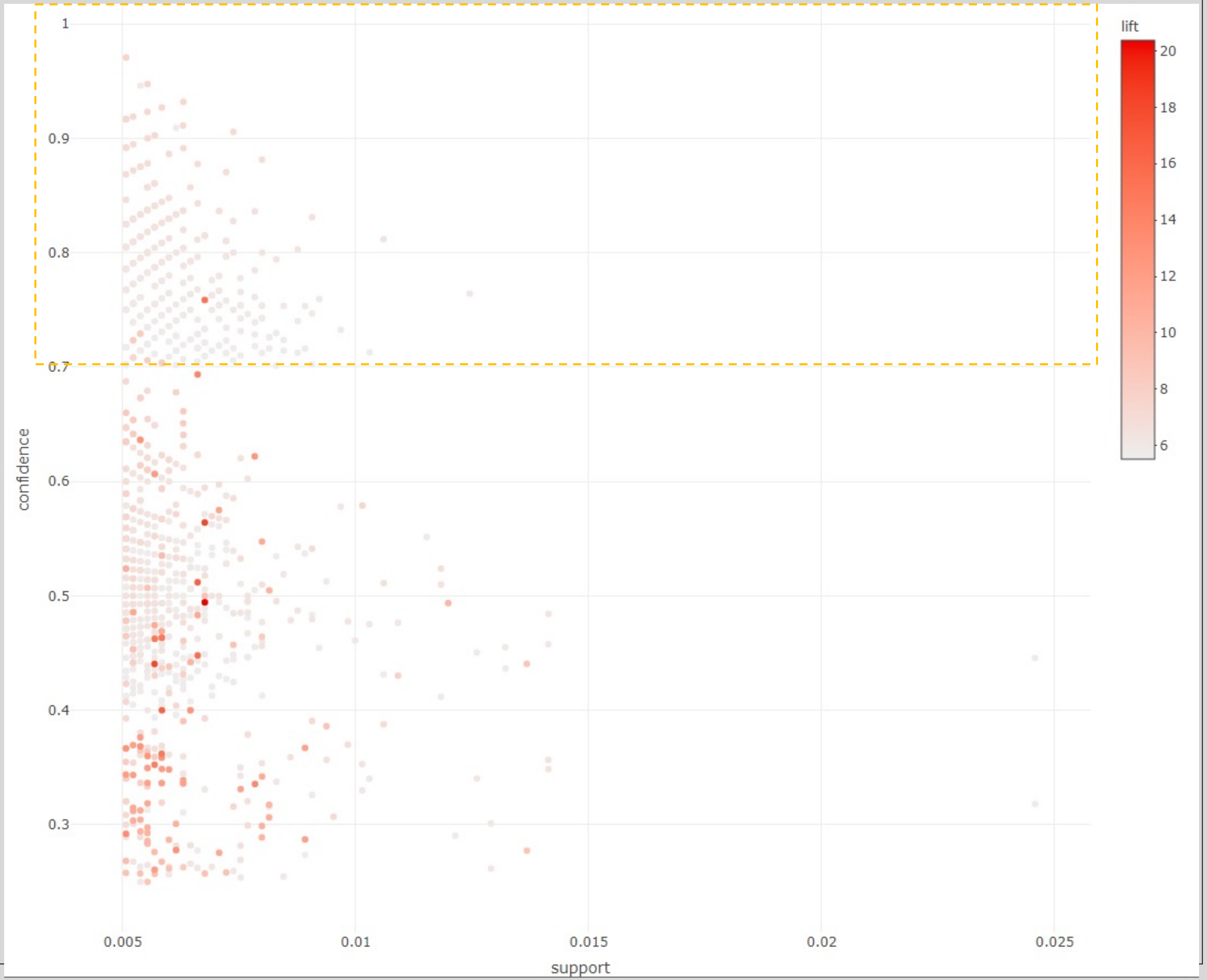
- Model Selection/Model Justification
  - Selected the apriori model for association rule mining because we felt the dataset was sufficiently large for this algorithm.
  - The apriori algorithm allows us to set support and confidence thresholds which we defined as a business decision. In contrast, other methods like Eclat, while quicker, utilize support as the sole parameter.
- How we built our model
  - To address the business problem of micro-orders, we opted to create our own customer order numbers by utilizing the year and week number combined with the customer number. What this created was a customer distinct order numbers aggregated at the weekly level to capture any associations between micro-orders and the principal order. This way, our association rules would not miss valuable associations from those single/small item baskets.
  - The one-week order threshold used in the model was due to: 1) anecdotal experience from grocery shopping being a weekly activity and 2) knowledge that organics and vegetables (the most commonly purchased items among our customer sample) have an expiry of ~ 1 week or less.
  - Instead of modelling from the SKU level of data, we opted to aggregate at the description level. This allows us to generalize more, creating stronger rules, and not be limited by brand (opportunity for us to promote brands that are sponsored/partnered in our company)
  - We built two association rule models: one model based on overall orders of our entire sample dataset and a second model based on an example of an individual customer-level model to provide more personalized rules (this in practice would be done at an individual or clustered individuals level depending on the company's capacity)

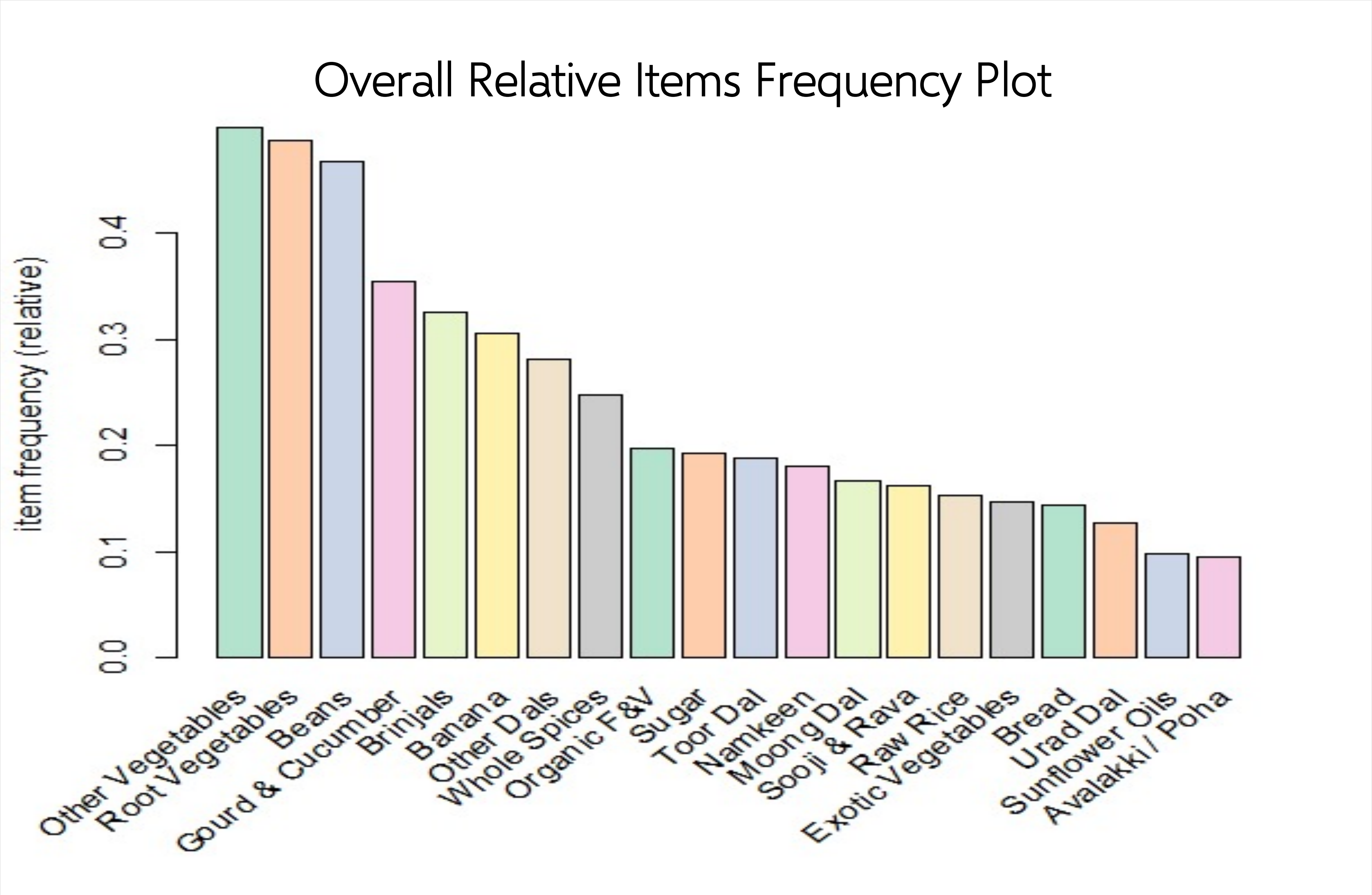
Recommender System Parameters

- Business decision for Recommender Systems (Support, Confidence and Lift)
  - Our model's primary parameter to excluding association rules was confidence and this was set to 0.70 due to trends we were seeing in the data. We wanted to be confident that the association rules of the item sets were not just due to the frequency of the items appearing the basket. This was especially important since the frequently purchased items of Vegetables, Fruits, Dals, etc. Not only appeared frequently together but were also frequently purchased on a stand-alone volumes basis.
  - We set a minimum support threshold of 0.005 just because going higher cut out quite a few rules.
  - Support, Confidence and Lift are all used in ranking the association rules based on what order they come up in our app.

Scatter Plot of Support vs confidence of our association rules

The rules we selected (within confidence threshold)




Overall Relative Items Frequency Plot
item frequency (relative)
0.0
0.1
0.2
0.3
0.4
Other Vegetables
Root Vegetables
Beans
Gourd & Cucumber
Brinjals
Banana
Other Dals
Whole Spices
Organic F&V
Sugar
Toor Dal
Namkeen
Moong Dal
Sooji & Rava
Raw Rice
Exotic Vegetables
Bread
Urad Dal
Sunflower Oils
Avalakki / Poha

Unique Customer Relative Items Frequency Plot