

INSTACART



BC2407 Presentation Team 4

Denis Tan (U1810101D)

Jared Teo (U1810545C)

Jovan Teo (U1810708L)

Lee Mun Yee (U1810810A)

Ong Qing Quan (U1810077E)

US based online
grocery company

A crowdsourcing
platform with zero
inventory

A revenue of \$2
billion in 2017

Asset-light business
model with neither
warehousing &
physical stores

INSTACART INC

INSTACART'S BUSINESS MODEL

4 Main Stakeholders

Customers

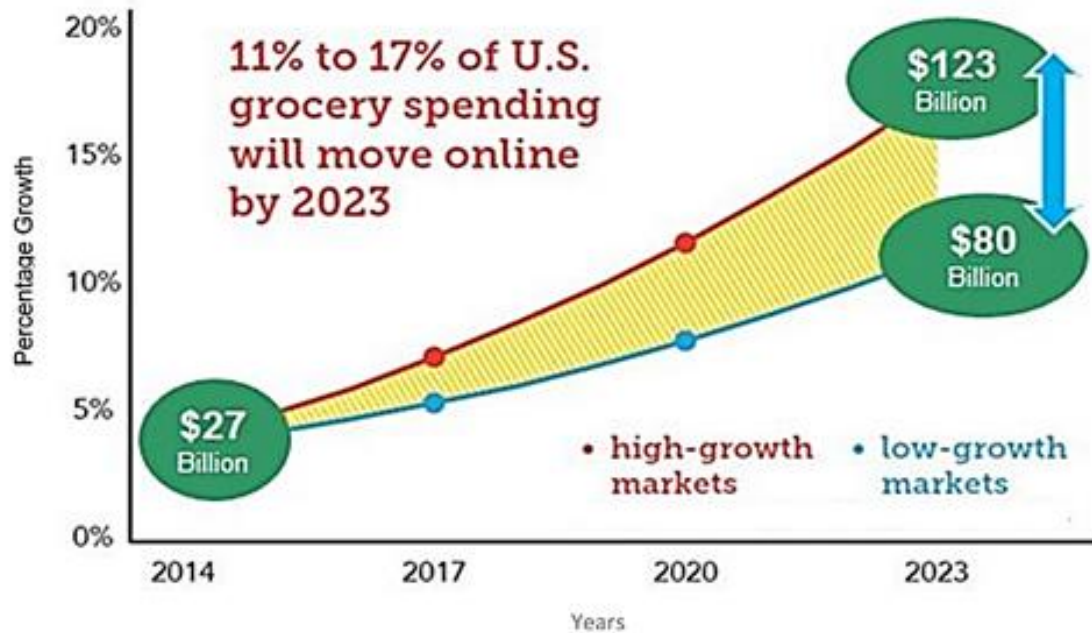
Stores

Products

Shoppers



Expected Growth for Online Grocery



Market Realist

Source: Bricks Meet Click Analysis, 2014

OPPORTUNITIES IN ONLINE GROCERY SECTOR

- Instacart's share of online grocery market is 14.5% as of 2019.
- Online Grocery has doubled its sales from 2016 to 2018.
- The sales is projected to grow by 17% to 19% each year over the next few years.
- 11% of people surveyed have purchased online in 2019.
- **Huge potential** in this market with increasing busy lifestyle.

Denis Tan

INSTACART'S BUSINESS PROBLEM

Threats

1. High reliance on sales volume due to its low profit margin (2%)
2. Fierce competition with strong competitors like Walmart

Instacart faces two problems:

1. How to generate higher sales volume to remain profitable?
2. How to retain its customers?

GENERAL APPROACH

Higher Sales

A recommendation system that makes product suggestion to customers based on past transaction records & users' activities

➤ Weightage-based recommendation system

Bundle sales promotion for items with high associative degree

Higher Customer Retention Rate

Better customer engagement & experience through recommendation system

Relevant bundling promotion for customers who want to buy products together at lower price

DESIRED OUTCOME

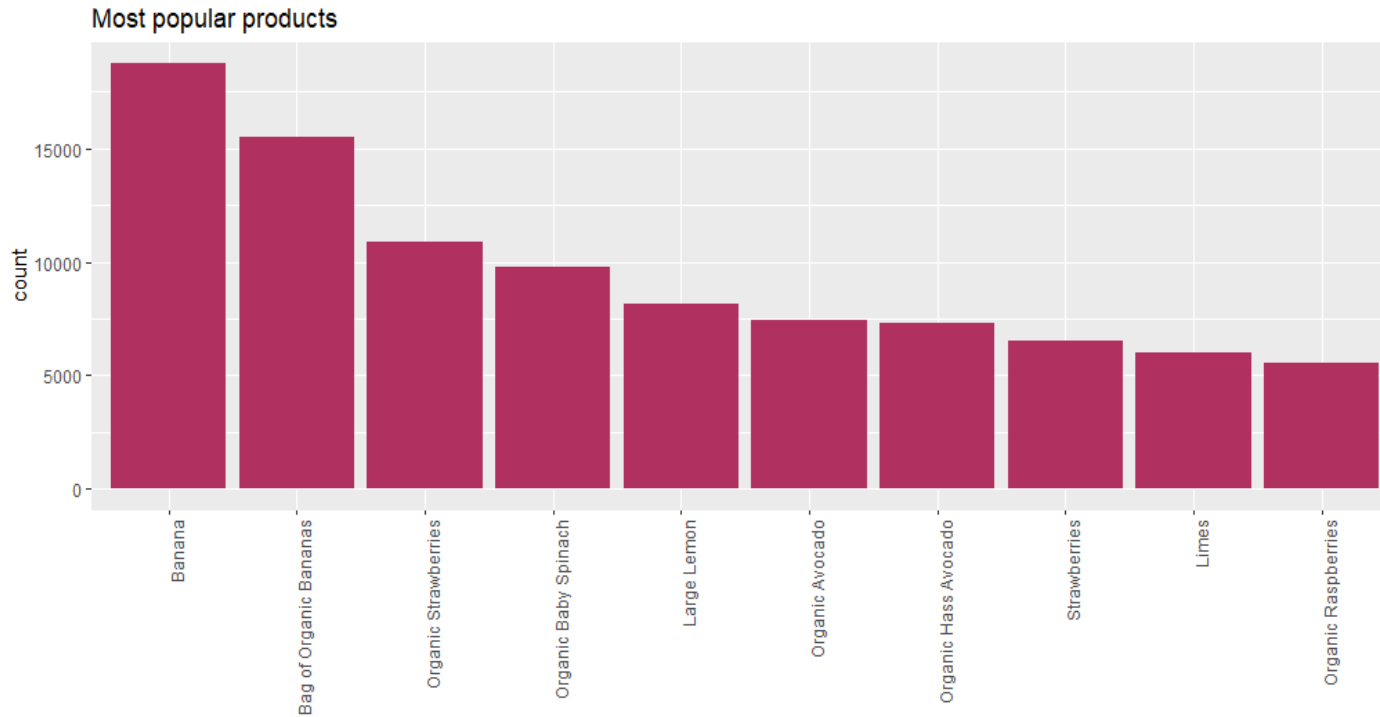
Higher Sales

- an increase in sales revenue of 3% or more

Higher Customer Retention Rate

- an increase of customer retention rate of 10% or more

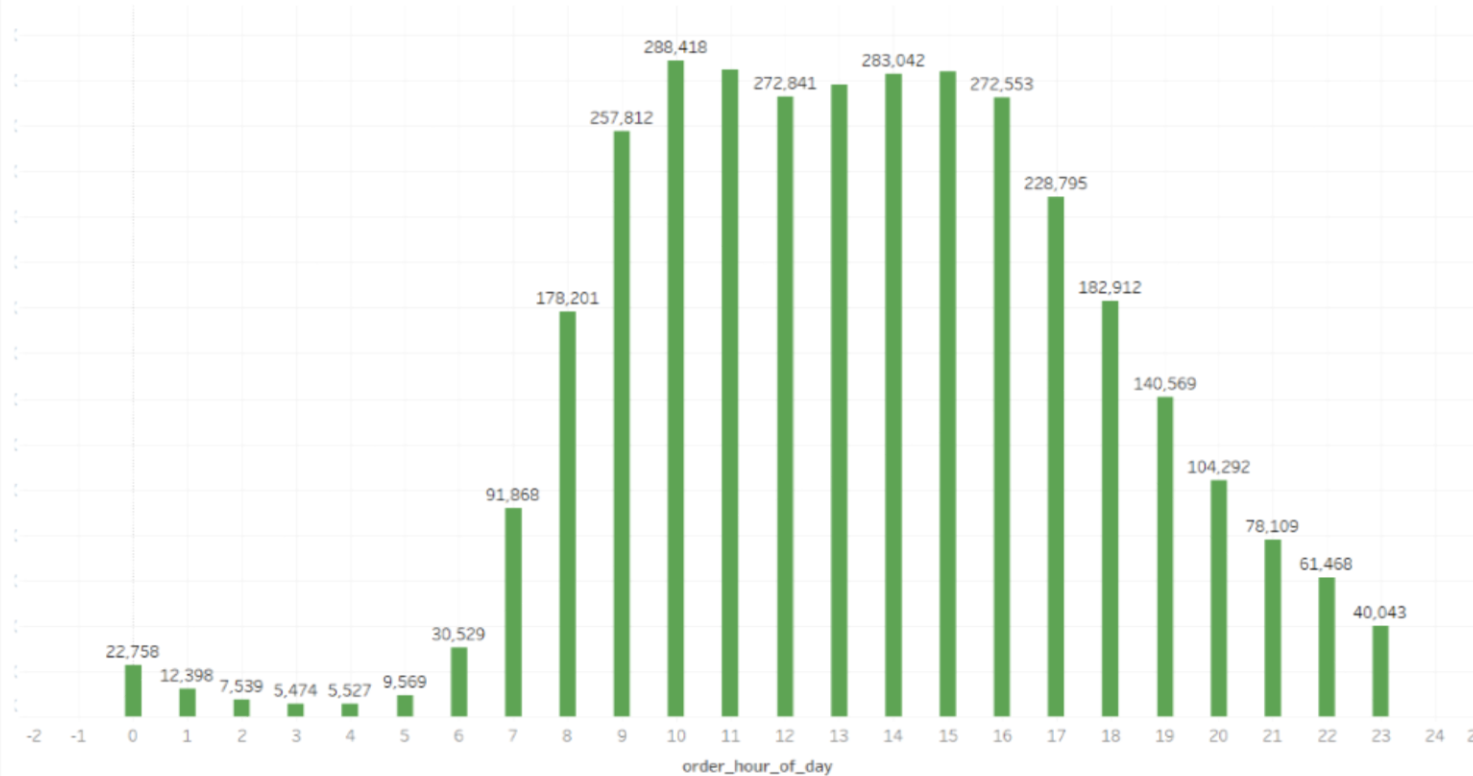
Customer Retention Rate = ((No. of Customers at End of Period – No. of Customers Acquired During Period) / No. of Customers at Start of Period)) X 100%



DATA EXPLORATION

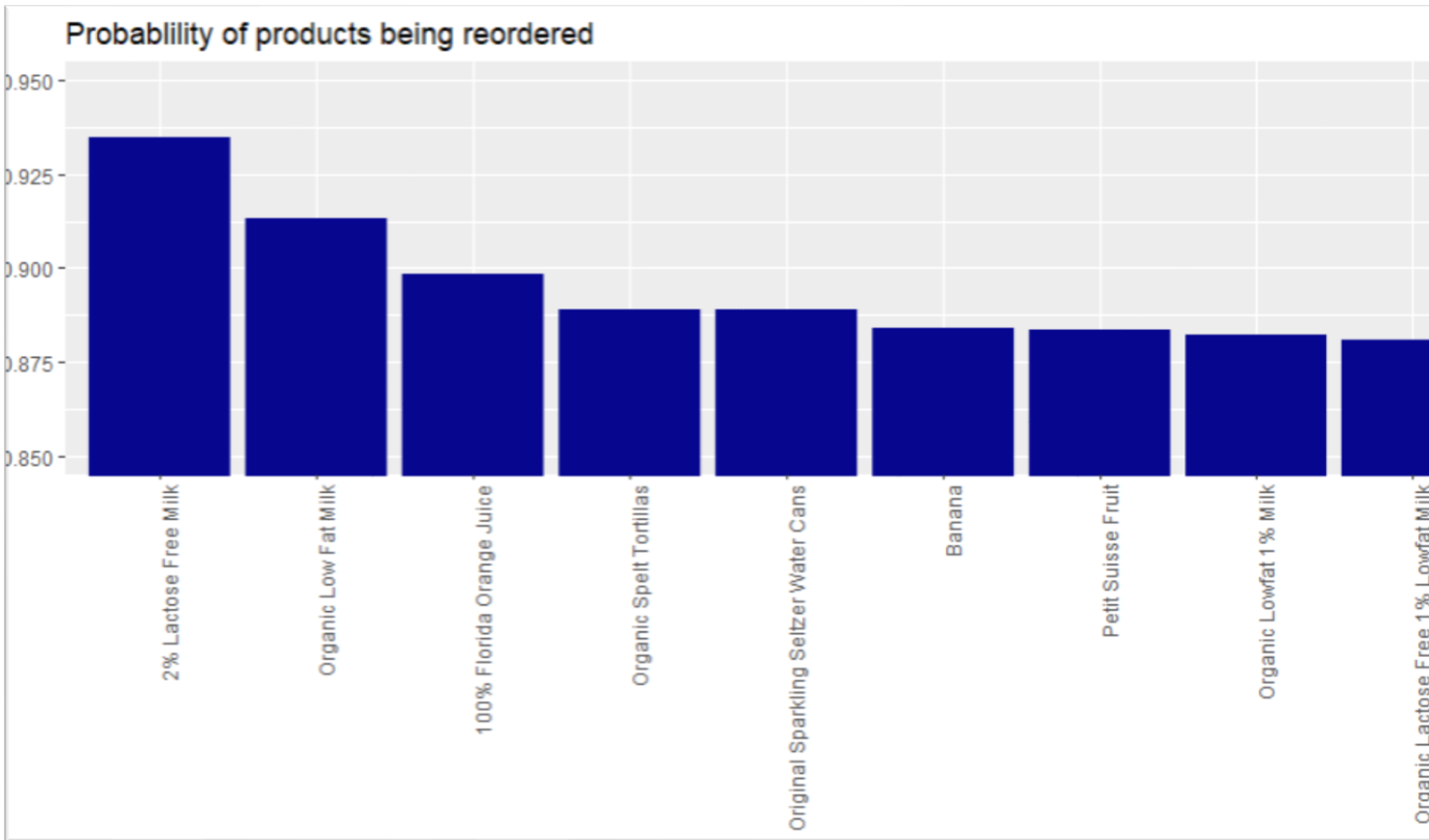
- Fruits and vegetables are most frequently ordered products.
- In fact, organic fruits and vegetables are seemingly popular amongst customers.
- Association rules and n rules and further data exploration could be performed to determine which organic products can be packaged and marketed together with these organic fruits and vegetables.

When do most people order in the day?



DATA EXPLORATION

- Aggregated between 10 a.m. to 5 p.m.
- Time our promotion and push notifications on promotions during high volume period each day
- Maximize the conversion of marketing efforts through the application to boost sales.



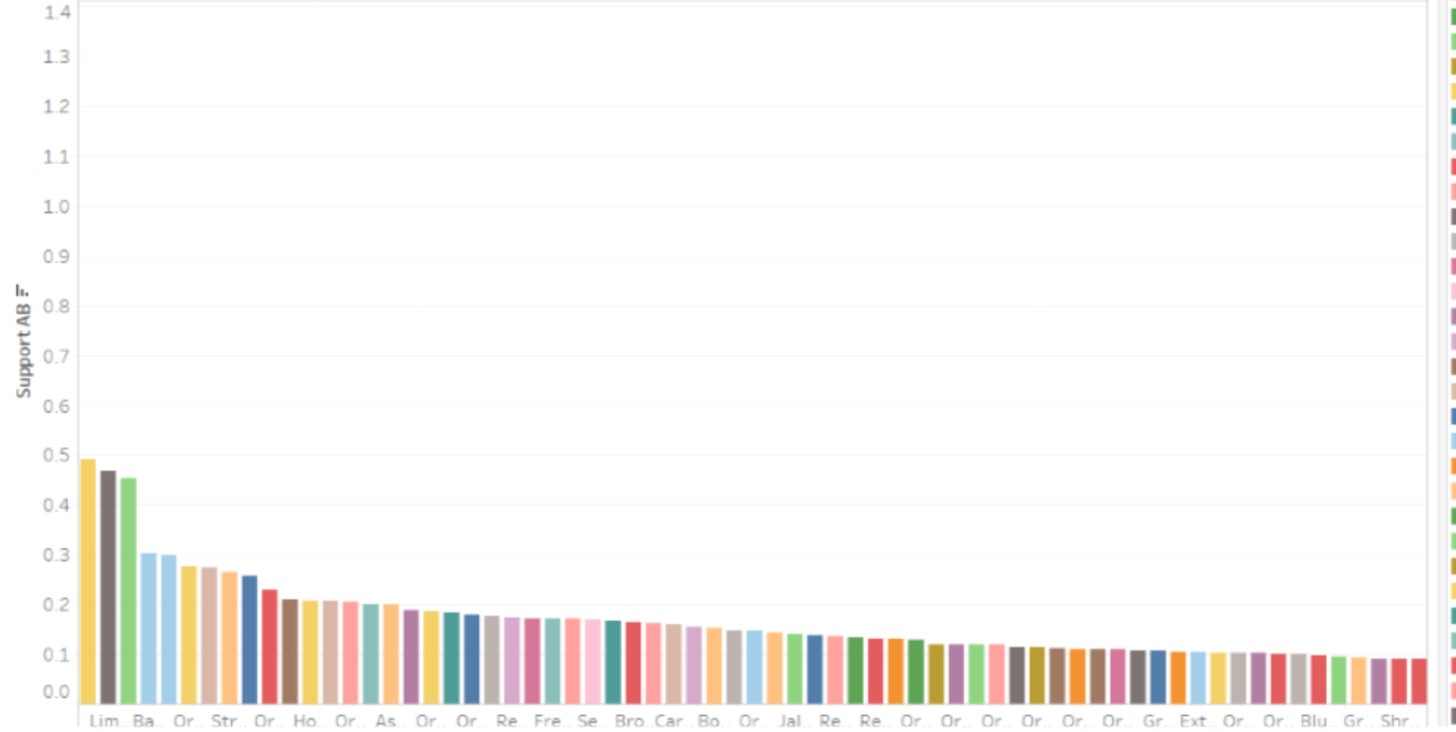
DATA EXPLORATION

- Focus on the products with high re-order rate
- Instacart can push these items as “Popular items other customers buy”
- Instacart can also easily narrow down the products to recommend to customers
- Explore the Association rules with the most reordered products as a baseline to bundle the products for promotion

Support Level for each pair of items

Item A / Item B

Organic Avocado



DATA EXPLORATION

- Sort the popularity of item pairs in descending order
- Instacart can focus on item pairs with high occurrence in our analytical model below or bundle the items together to be sold as a package.

PROPOSED SOLUTIONS TO INSTACART

Association Rule
Mining

```
graph TD; A[Association Rule Mining] --> B[Higher Sales]; A --> C[Higher Customer Retention Rate];
```

Higher Sales

Higher Customer
Retention Rate

ANALYTICAL SOLUTION: ASSOCIATION RULE

- Large data (over 3million rows)
- Python generators leveraged on the Apriori algorithm are used instead of python's built in Apriori Package.
- Rule comprise of only 2 items ($A \rightarrow B$) due to the wide variety of items sold in Instacart

Item A	Item added to the cart
Item B	Suggested item to be considered by customers.
Support AB	% simultaneous occurrence for Item A and B
Confidence AtoB	The confidence of the rule that A results in B
Lift	Ratio of the confidence of the rule and the expected confidence of the rule.

BIBLE FOR ASSOCIATION RULE

	itemA	itemB	freqAB	supportAB	freqA	supportA	freqB	supportB	confidenceAtoB	confidenceBtoA	lift
0	Oh My Yog! Pacific Coast Strawberry Trilayer Y...	Oh My Yog! Organic Wild Quebec Blueberry Cream...	860	0.028907	2856	0.095998	2271	0.076335	0.301120	0.378688	3.944745
2	Unsweetened Blackberry Water	Raspberry Essence Water	660	0.022184	3108	0.104468	2025	0.068066	0.212355	0.325926	3.119850
3	Organic Fiber & Protein Pear Blueberry & Spina...	Fiber & Protein Organic Pears, Raspberries, Bu...	606	0.020369	2782	0.093511	2167	0.072839	0.217829	0.279649	2.990560
4	Oh My Yog! Organic Wild Quebec Blueberry Cream...	Oh My Yog! Pacific Coast Strawberry Trilayer Y...	642	0.021579	2271	0.076335	2856	0.095998	0.282695	0.224790	2.944798
1	Oh My Yog! Pacific Coast Strawberry Trilayer Y...	Oh My Yog! Madagascar Vanilla Trilayer Yogyurt	640	0.021512	2856	0.095998	2567	0.086284	0.224090	0.249318	2.597119
5	Unsweetened Watermelon Water	Unsweet Peach Water	599	0.020134	4107	0.138048	1777	0.059730	0.145849	0.337085	2.441803

TOP 6 ASSOCIATION RULES ARRANGED
ACCORDING TO CONFIDENCE LEVEL AT
MINSUPPORT= 0.02

FINDINGS IN ASSOCIATION RULE

- Most rules with high confidence have lift > 1
- Items of similar name have high confidence association rule
- Useful insights derived from unique association rules from dissimilar items

	itemA	itemB	freqAB	supportAB	freqA	supportA	freqB	supportB	confidenceAtoB	confidenceBtoA	lift
45	Chocolate Peanut Butter	Coconut Chia Bar	970	0.032604	4649	0.156266	4646	0.156165	0.208647	0.208782	1.336069

ANALYTICAL SOLUTION TARGET

Increase user
experience

Increase
retention
rate

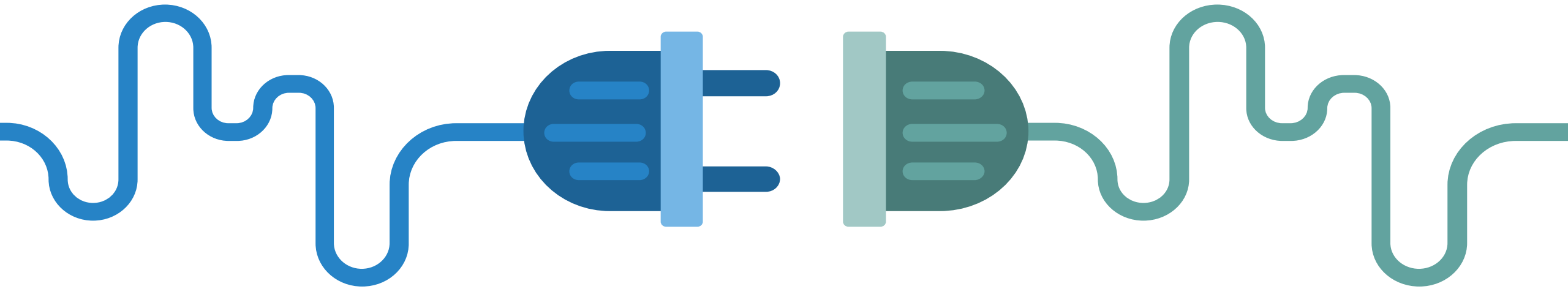
Higher
Profitability

Increase the chances
of customers buying
recommended items

Increase
sales

Higher
Profitability

ANALYTICAL SOLUTION PERFORMANCE MEASURES



Quantile Regression

- Improvement on Y-outcomes after implementation
- Y-outcomes regressed against time
- Y-outcomes stratified into different quantile range

Cost-performance Matrix

- Usability and sustainability measurement
- Benchmarked against similar companies and industry standards
- Cost analysis needed due to high cost of implementation

MODEL IMPROVEMENTS AND FUTURE ADVANCEMENTS



Semi-supervised
Machine learning



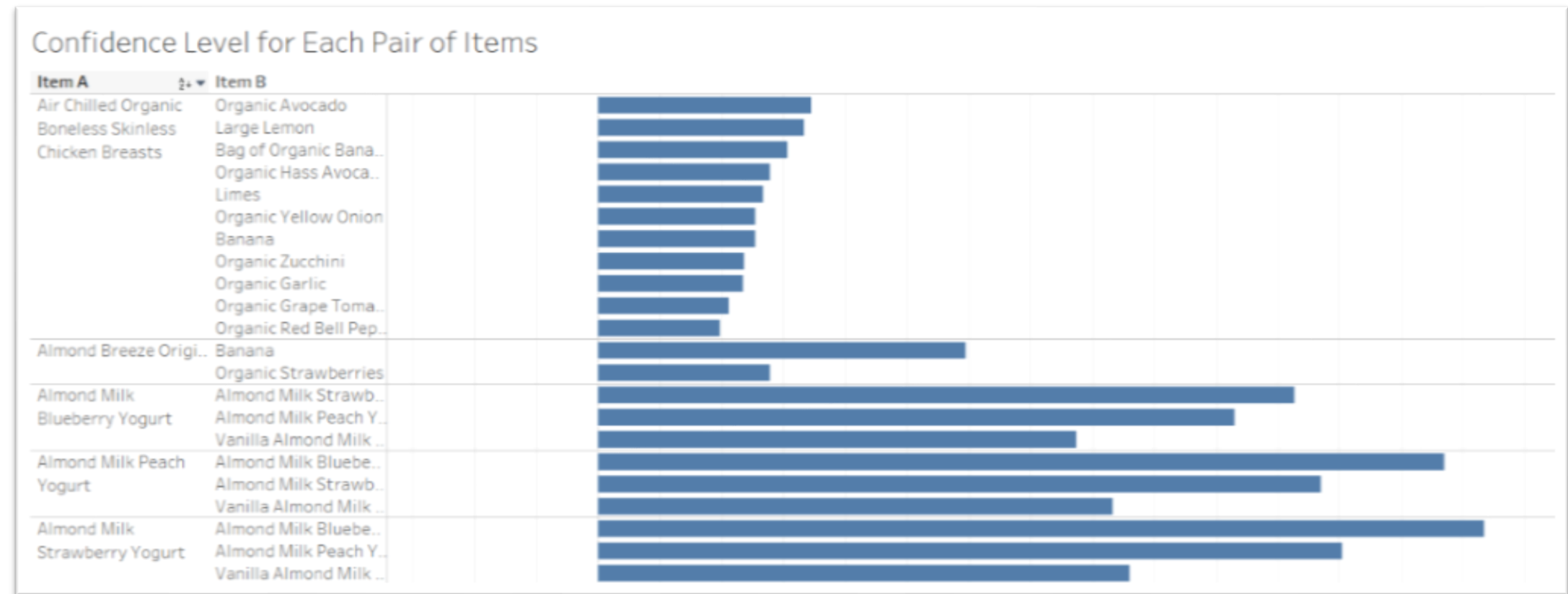
Threshold
adjustments



Item Grouping



Gibbs sampling



DEMONSTRATION OF ANALYTICAL SOLUTION

```
select itemA, itemB, supportAB, confidenceAtoB from project1.resultba2 where supportAB > 0  
and confidenceAtoB > 0 and itemA = "Coconut Chia Bar"  
order by confidenceAtoB desc;
```

itemA	itemB	supportAB	confidenceAtoB
Coconut Chia Bar	Chocolate Peanut Butter	0.030251481	0.193715024
Coconut Chia Bar	Acai Berry Chia Bar	0.029243098	0.187257856

DEMONSTRATION OF ANALYTICAL SOLUTION

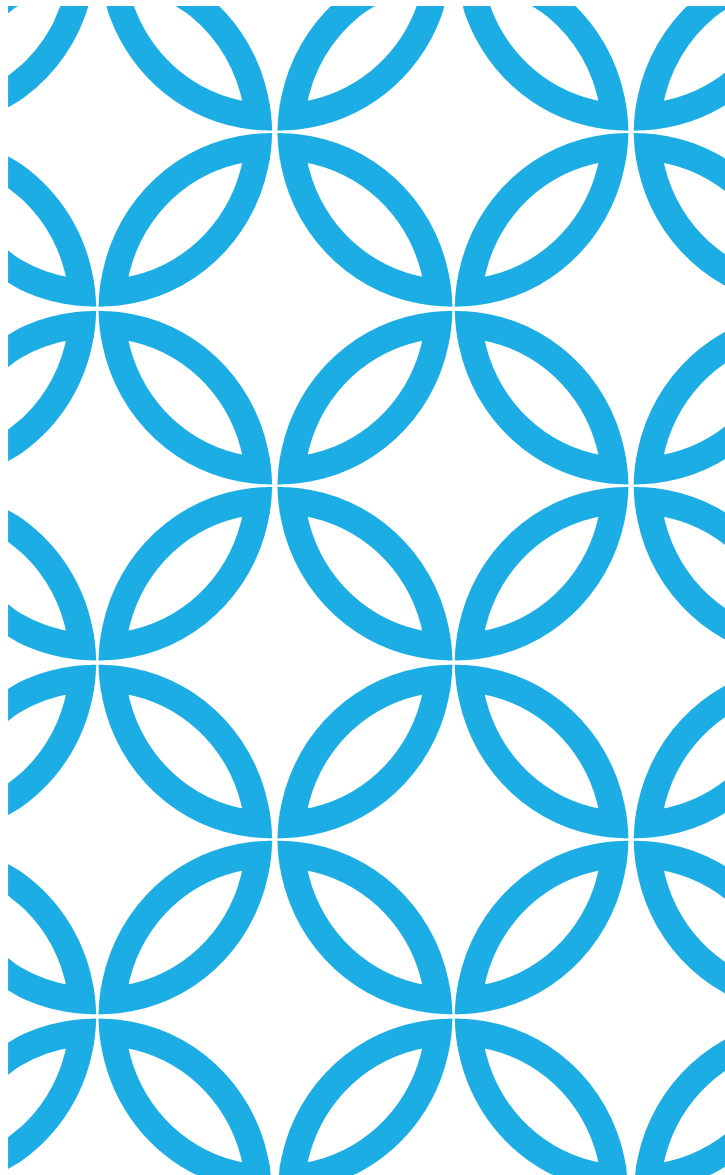
```
create view CartItems1( itemA, itemB, supportAB, confidenceAtoB)
as select itemA, itemB, supportAB, confidenceAtoB from project1.resultba2 where supportAB > 0
and confidenceAtoB > 0 and itemA= "Coconut Chia Bar" or itemA = "Chocolate Sea Salt" or
itemA= "Peanut Butter Bar"
order by confidenceAtoB desc;
```

```
select* from CartItems1 where itemB != "Coconut Chia Bar" and
itemB != "Chocolate Sea Salt" and itemB != "Chocolate Peanut Butter"
order by confidenceAtoB desc limit 3;
```

itemA	itemB	supportAB	confidenceAtoB
Chocolate Sea Salt	Coconut Chocolate Bar	0.028873358	0.198566805
Coconut Chia Bar	Acai Berry Chia Bar	0.029243098	0.187257856
Chocolate Sea Salt	Peanut Butter Egg White Protein Bar	0.021209649	0.145862228

CONCLUSION

- Increase sales and retention rate can be achieved via our model
- Expansion of our model and Instacart's database is needed for continual increased profits
- Variables such as Day and Time of purchase can be added
- Utilize the advantages of data science to differentiate themselves from others



THANK YOU!
