



# InstaCart

## What will you order next?

michael fedell - 2019-06-10

# Some **motivation** for this project



## Shopping

Shopping is ubiquitous in our lives. With so much data, we can solve lots of problems



## Food

I am a personal lover of food and think it's exciting to explore new foods and ways to shop



## Solution

This application quantifies grocery shopping history and exposes hidden value to shoppers and suppliers



# Overview of Data



## Products

Specific product information like name, aisle, department, etc.



## Orders

Data for each order including attributes like date/time of order, size, and individual products



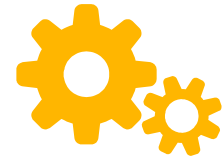
## Order Types

Archetype labels describing the type of order based on data-mined attributes



## Shoppers

Shopper profiles built on all historical orders in the dataset



# Modeling Techniques



## Clustering

Orders were clustered based on produce-level statistics and metadata. Gaussian Mixture Models were used to find 6 distinct neighborhoods of similar orders.



## Classification

After extensive model evaluation, a linear support vector classifier was chosen based on its light weight and relatively strong classification power.



## Factor Analysis

The model predicts on 52 features - this would not be ideal for user input. To remedy, features were mapped to 4 dimensions obtained by factor analysis



# Insights gained

- Shoppers are predictable
- Orders are not easily classified
- Significant pattern in temporality
- Online grocery shopping likely different than in-store



# Thanks for listening!

## Any Questions?

**– Michael Fedell**

MS in Analytics Candidate - Dec 2019  
michaelfedell@u.northwestern.edu