

Predicting Grocery Customer Behaviors and Market Basket Analysis

Henry Shin



About Me

- Data Science Student at Flatiron School
- University of Texas at Austin graduate
 - Bachelors in Economics
- Passionate about sports.
 - Fantasy Football nerd and Dallas Stars fanatic
- Entrepreneurial experience in meal preparation business
 - Bento's Kitchen (Food Prep Startup)



Agenda

1. Objective
2. Data Overview
3. Analysis and Methodology
4. Modeling
5. Recommendations
6. Future Considerations

Objective

- **Stakeholder:**

- The executives at **GroceriesToGo**: a startup firm with online grocery shopping services

- **Business Problem:**

- Predict if a customer will reorder a product.

- **1%-3%:** margins of traditional grocery store

- **2%:** margins of most profitable online grocery stores



Bottom Line

- Our best model predicted **75 out of 100** reordering customers precisely
 - Why Precision metric?



Data Utilized

2017 Instacart Open Sourced

- 206K Unique Customers
- 3.4M Unique Orders
- 50K Total Products Ordered

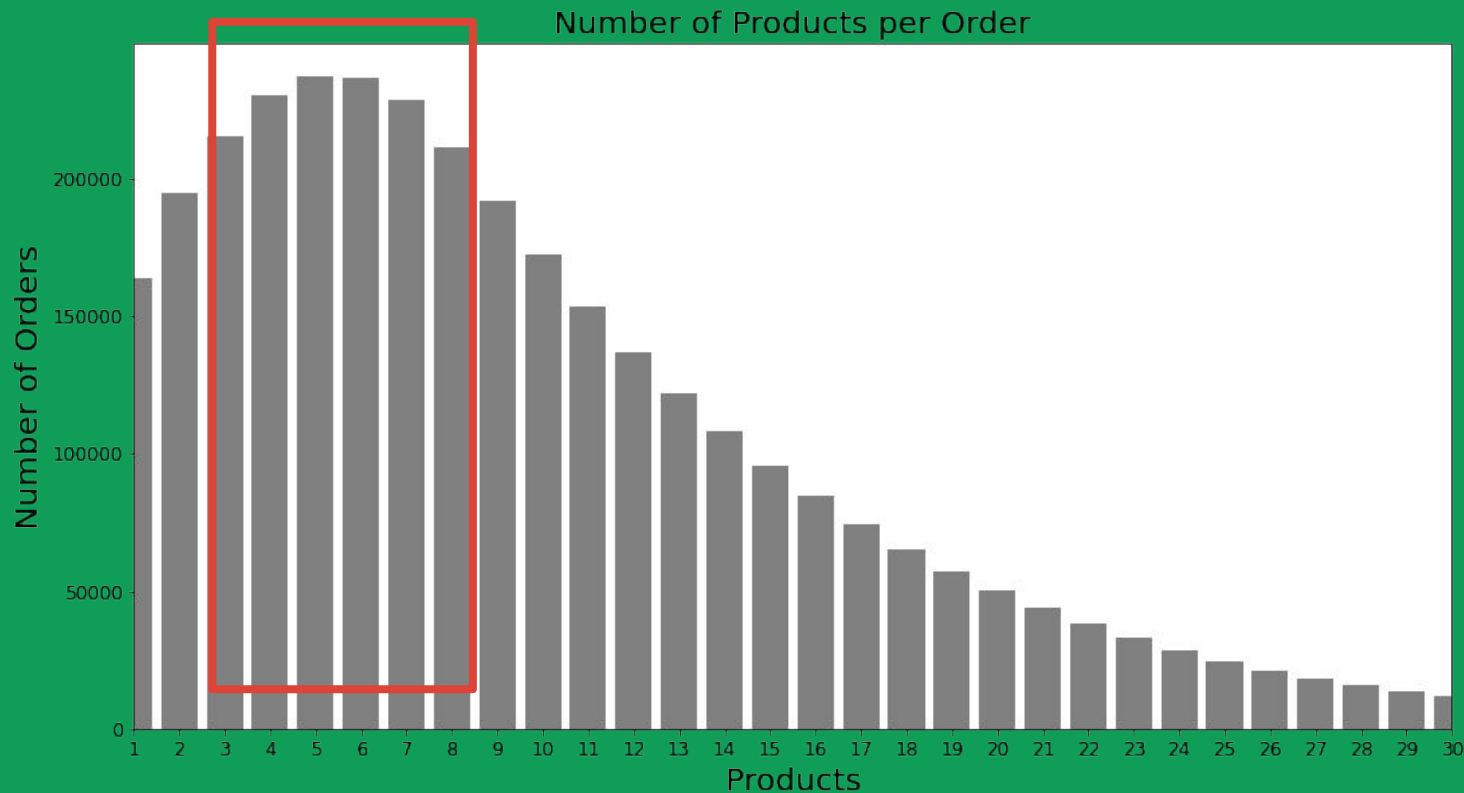
Limitations

- Class Imbalances
- User Data is anonymized
 - No demographic Data (gender, age, income)
 - Products does not include the Brand Names

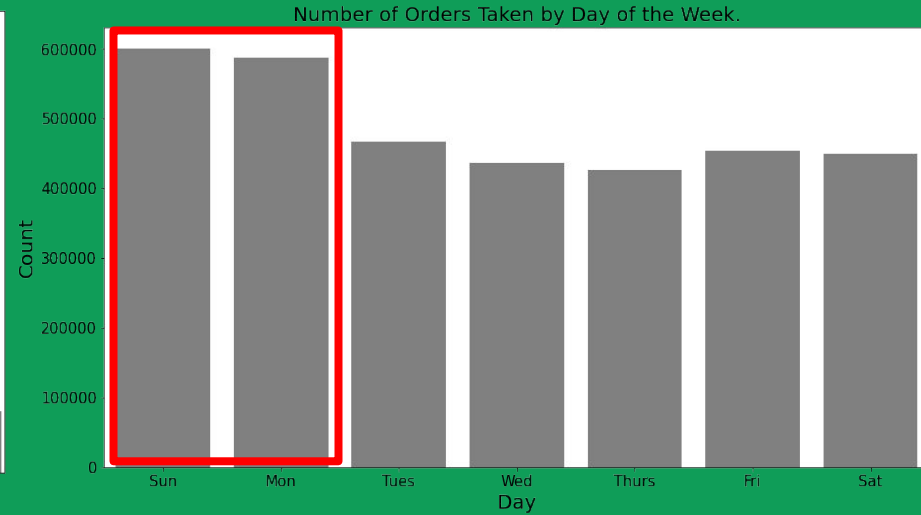
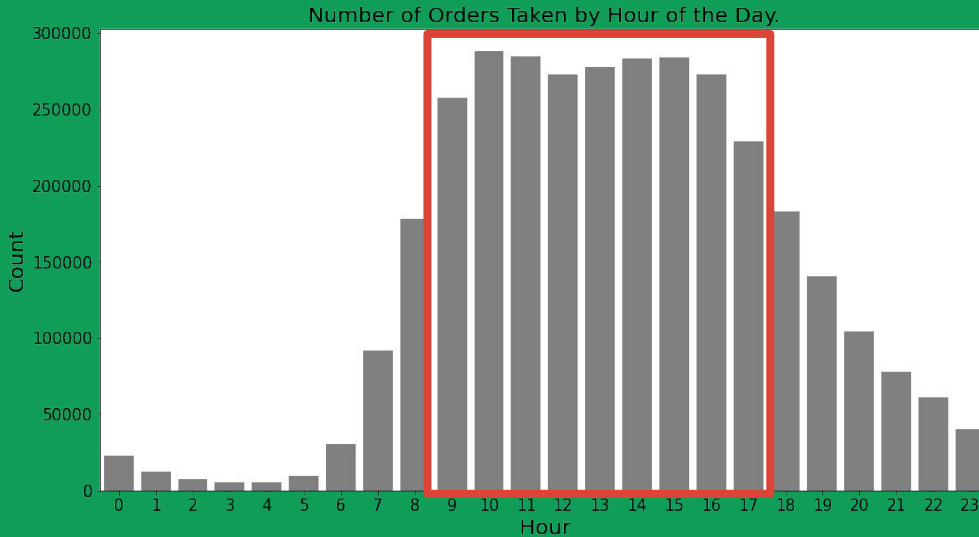


Exploratory Data Analysis

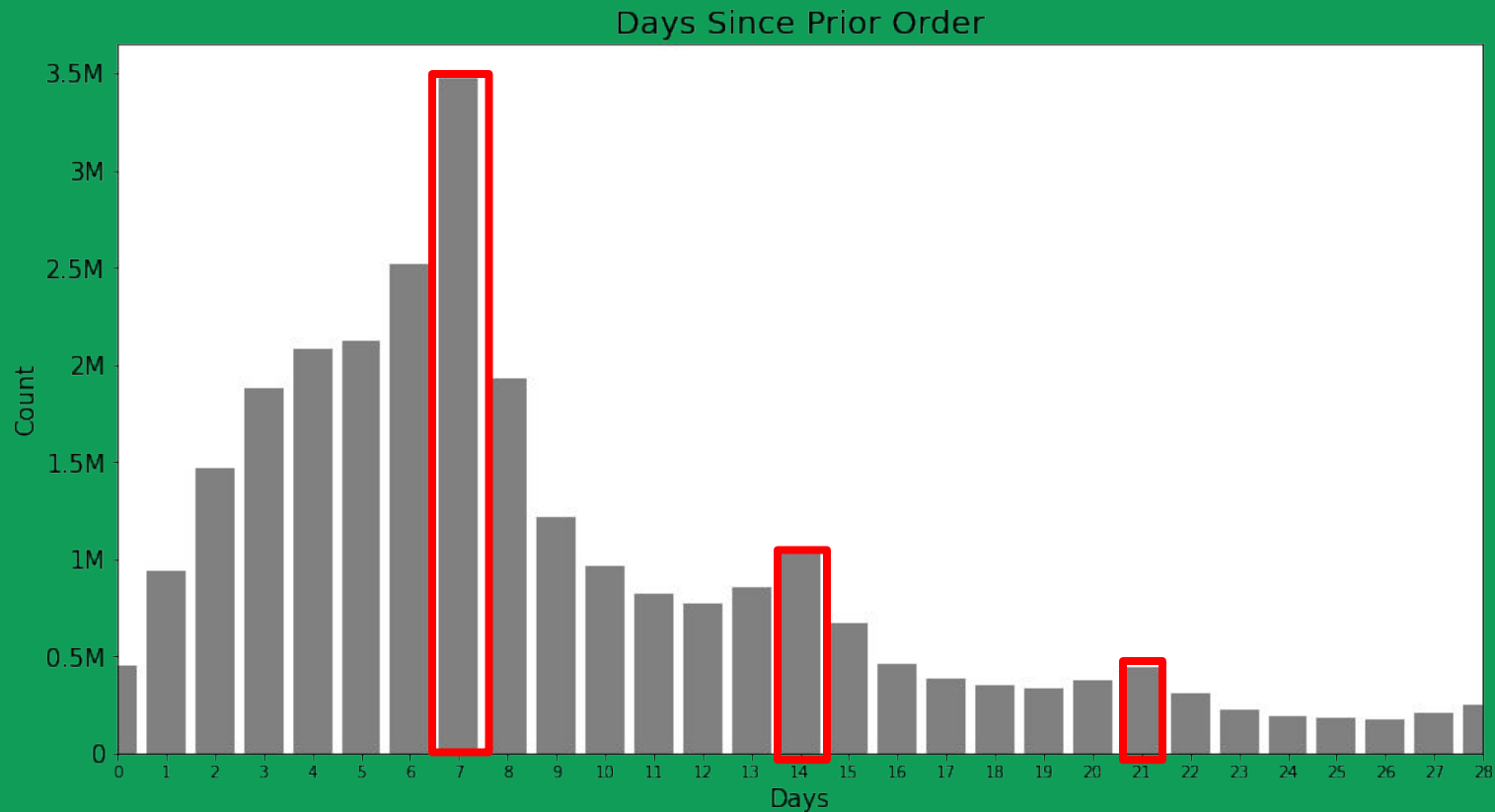
Basket Size Trends



Customer Behavior Time Trends



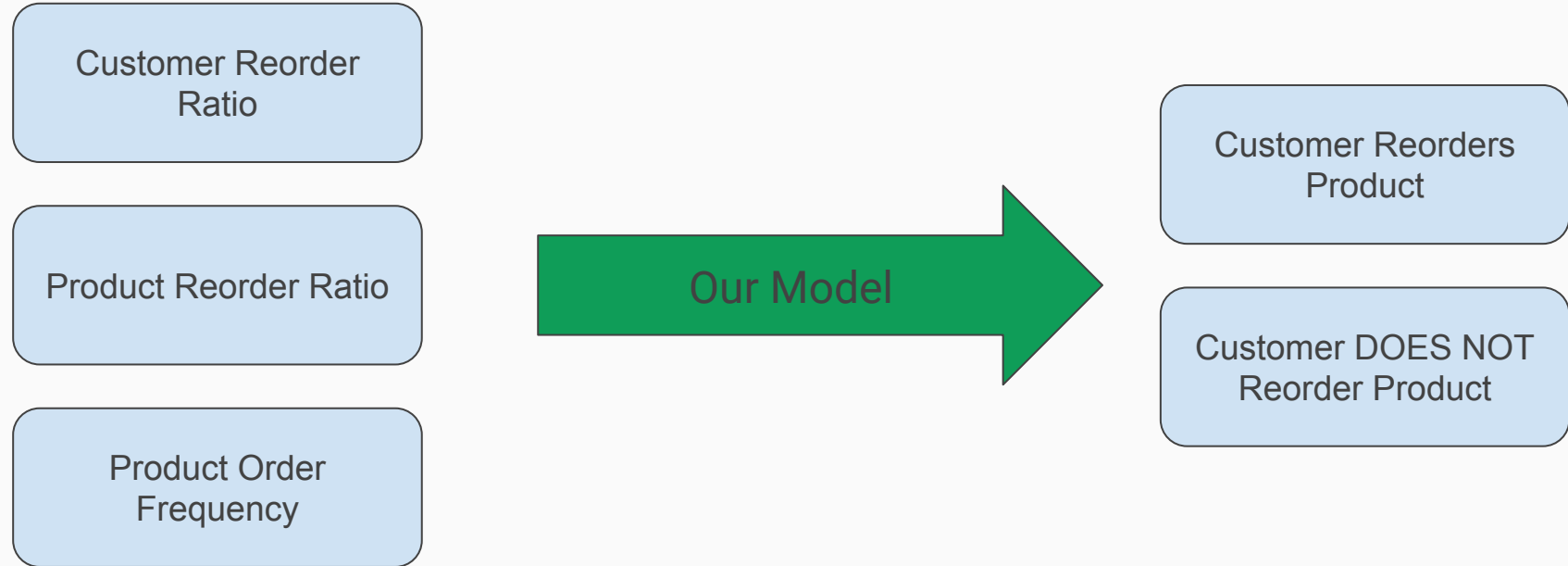
Weekly Trends in Reordering



Modeling Process

Modeling Process

Most Important Features

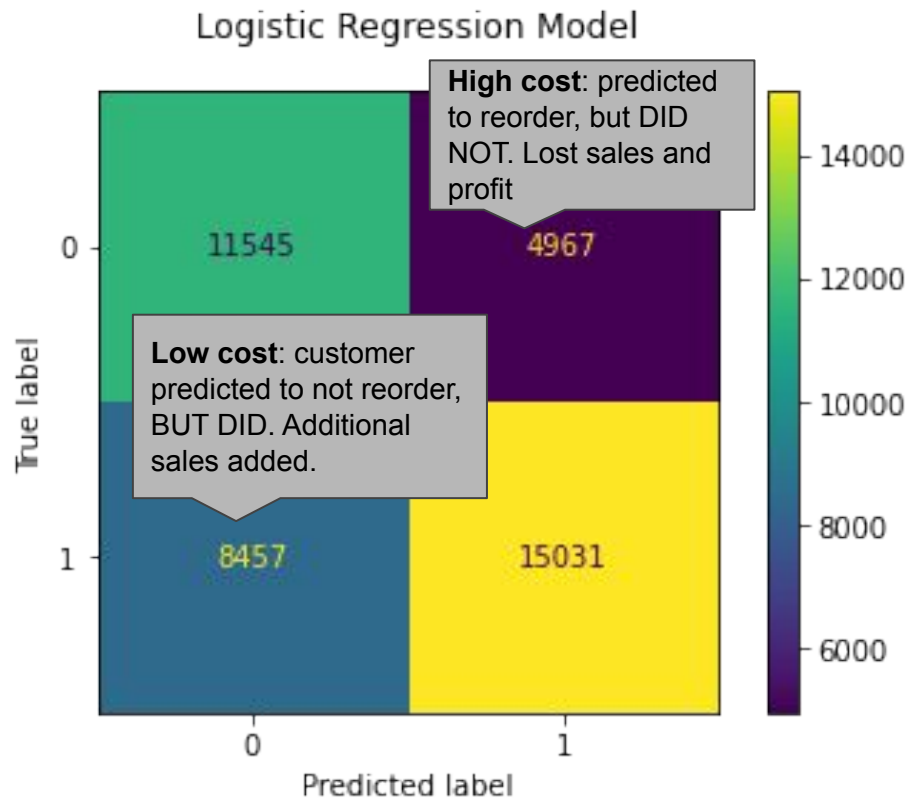


Models

- Our Best Model identified 75 out of 100 reordering customers precisely

| Model | Precision Score |
|----------------|-----------------|
| Baseline Model | 59% |
| First Model | 69% |
| Best Model | 75% |

Best Model Confusion Matrix: 75% Precision





Recommendations

- Target the reordering customers with coupons and ads on most popular products overall and personal most popular ordered products
- Create a loyalty or points program to retain and inform customers.
- Use these models to complement GroceriesToGo POS
 - Utilize insights to optimize supply chain and operations

Future Insights

- What are the significant differences in traditional vs online grocery store data and trends?
 - Pre COVID, Peak COVID, and Post-Quarantine COVID grocery trends
- Seasonal products and trends analysis.
 - Example: Pumpkins for Halloween
 - Larger basket sizes for major holidays
- **Advanced Market Basket Analysis** to give insights which products are often bought together
 - **Clustering techniques**
 - **Recommendation App**





Thanks! Questions?

Pictures: www.unsplash.com

Contact

<https://www.linkedin.com/in/henryhshin/>

<https://github.com/henryshin15>

hjshin386@yahoo.com

Appendix

<https://www.posnation.com/blog/grocery-store-profit-margins>

<https://www.statista.com/statistics/1178365/online-grocery-profit-margin-worldwide/>