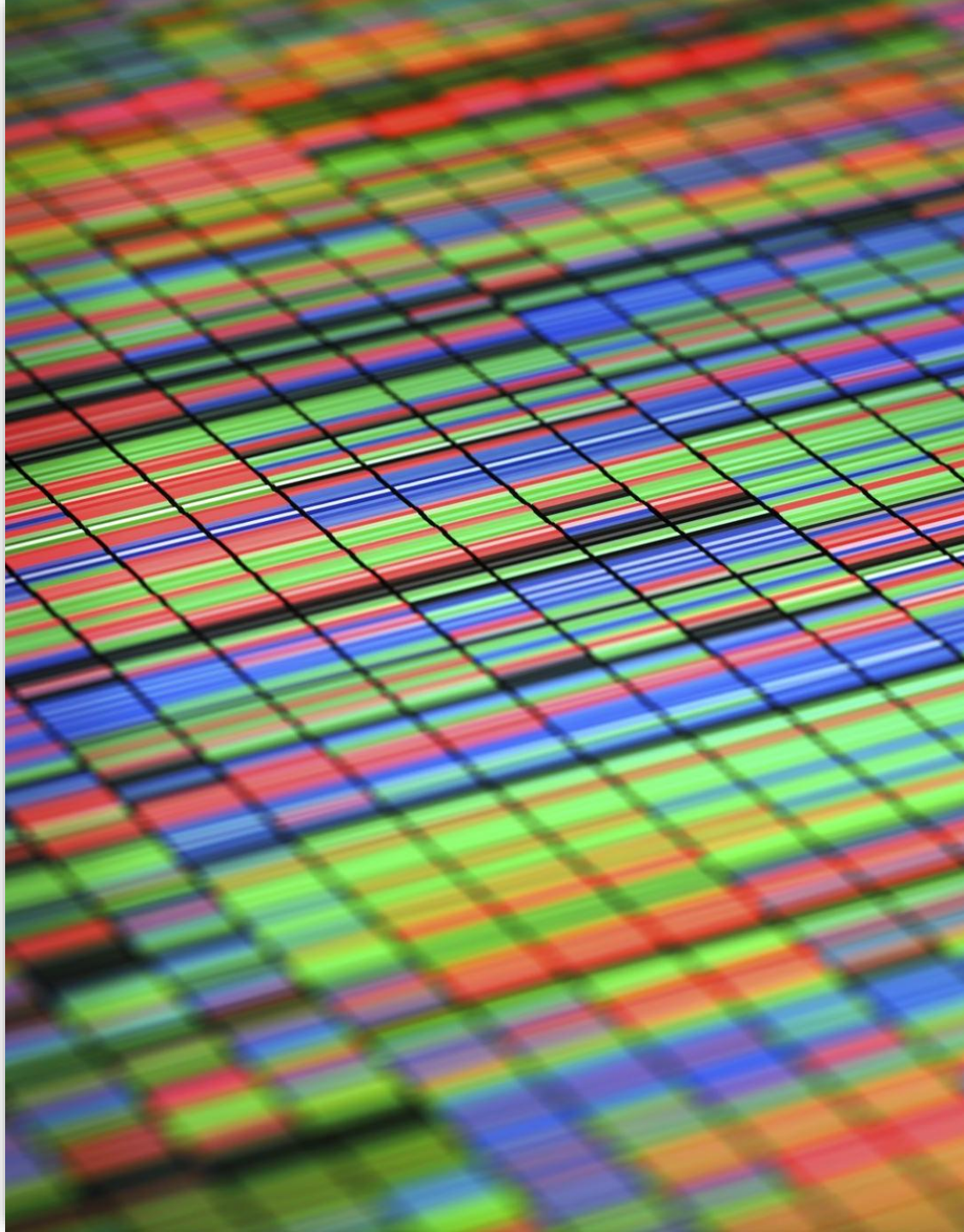
Instacart Case Study

ANDY CEN

Data Analytics Student

Key Skills Applied

- Python
- Data Cleaning
- Data Wrangling
- Data Subsets
- Merging Data frames
- Deriving new variables
- Data Aggregation & Grouping
- Data Visualization



Contents

Project Goals

instacart Grocery Basket Analysis

Goal: The Instacart stakeholders are most interested in the variety of customers in their database along with their purchasing behaviors. They want to target different customers with applicable marketing campaigns to see whether they have an effect on the sale of their products. This analysis will inform what this strategy might look like to ensure Instacart targets the right customer profiles with the appropriate products.

Motivation: Instacart is the leading grocery delivery platform in the US. Instacart stakeholders aim to understand customer diversity and purchasing behaviors to develop a targeted marketing strategy.

Objective: Perform an initial data and exploratory analysis to derive insights and strategies for better segmentation.

Defining the Key Questions & Objectives

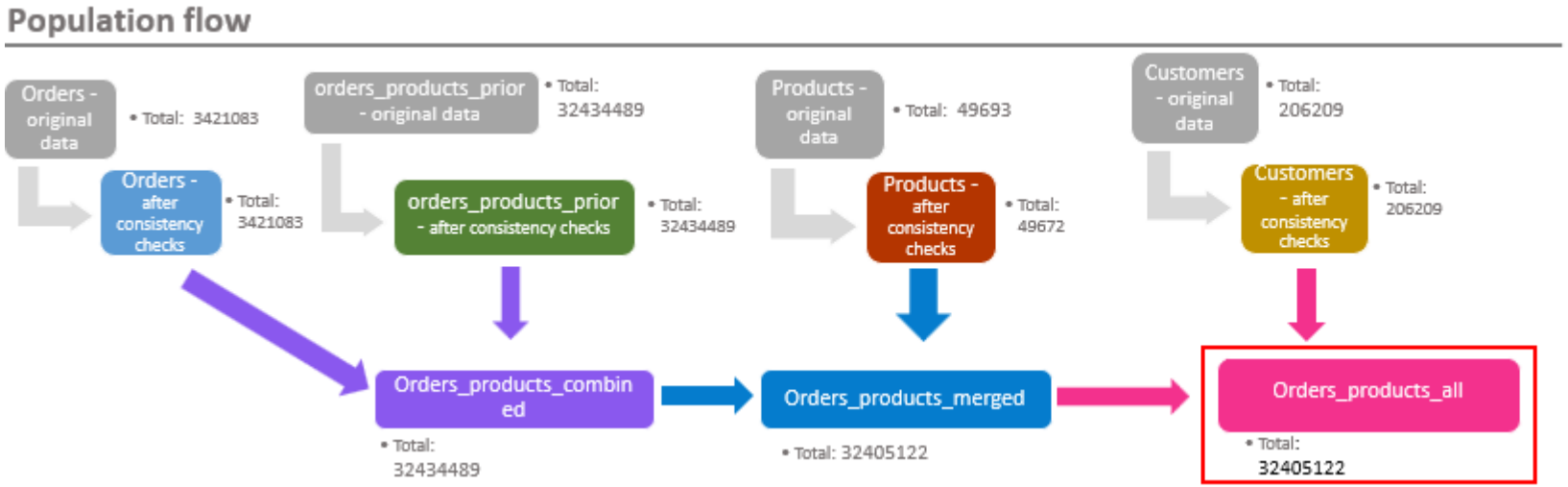
- The sales team needs to know what the *busiest days of the week* and *hours of the day* are (i.e., the days and times with the most orders) in order to schedule ads at times when there are *fewer orders*.
- They also want to know whether there are particular times of the day when people *spend the most money*, as this might inform the type of products they advertise at these times.
- Instacart has a lot of products with different price tags. Marketing and sales want to use *simpler price range groupings* to help direct their efforts.
- Are there certain types of products that are more *popular* than others? The marketing and sales teams want to know which departments have the *highest frequency of product orders*.

- What's the distribution among users in regards to their *brand loyalty* (i.e., how often do they return to Instacart)?
- Are there differences in ordering habits based on a customer's loyalty status?
- Are there differences in ordering habits based on a customer's region?
- What different classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?
- What differences can you find in ordering habits of different customer profiles?
    - Considering the price of orders, the frequency of orders, the products customers are ordering, and etc.

Sourcing and Preparing the Data

- Customers Dataset from CareerFoundry
- Instacart Dataset
- Instacart Data Dictionary
- : “The Instacart Online Grocery Shopping Dataset 2017”, Accessed from www.instacart.com/datasets/grocery-shopping-2017 via Kaggle on April 9.

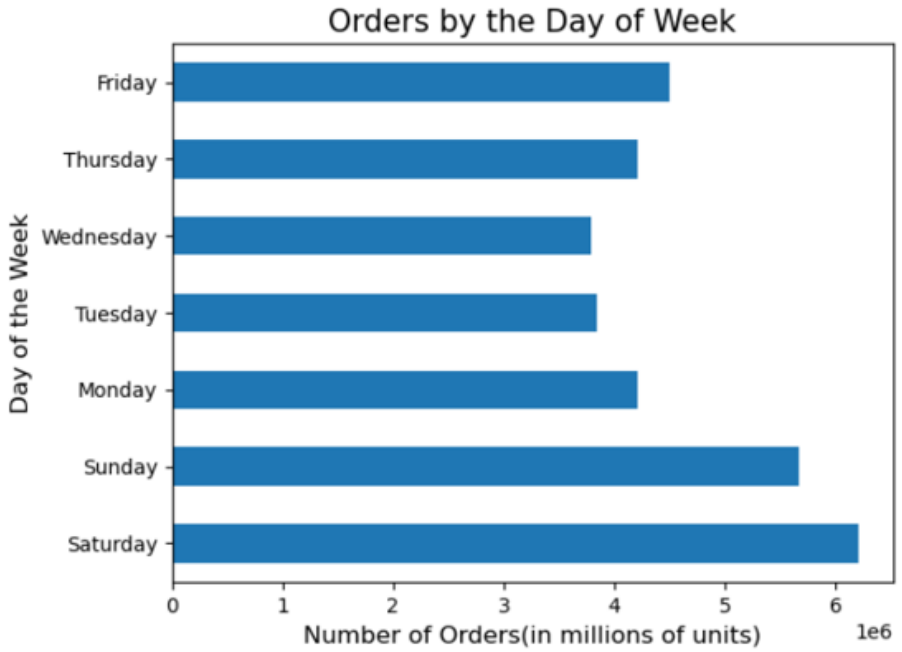
This data was cleaned, wrangled, transformed, integrated, and summarized in Excel and Python.

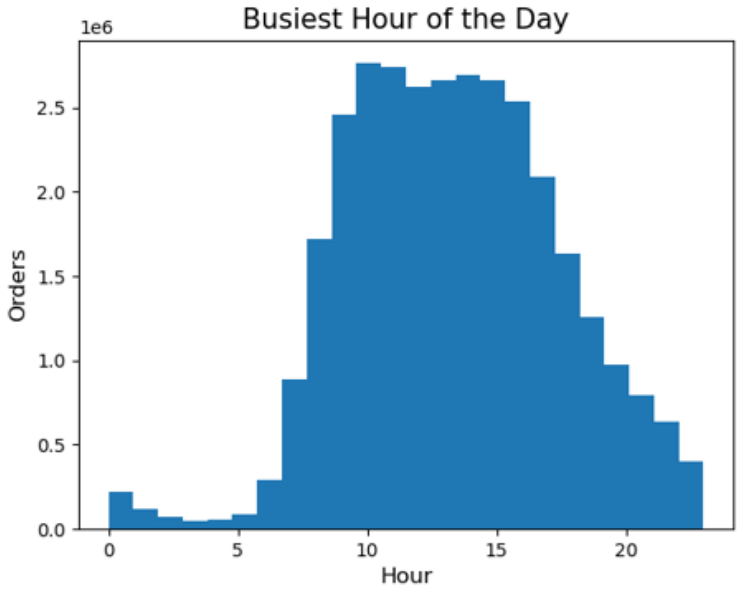


Exploratory and Visual Analysis

The sales team needs to know what the busiest days of the week and hours of the day are (i.e., the days and times with the most orders) in order to schedule ads at times when there are fewer orders.

Saturday and Sunday seem to have the most orders of the week. Wedneday and Tuesday have the least orders of the week. The busiest hours are from around hour 9 to hour 16. The least busiest orders are from hour 0 to hour 5.

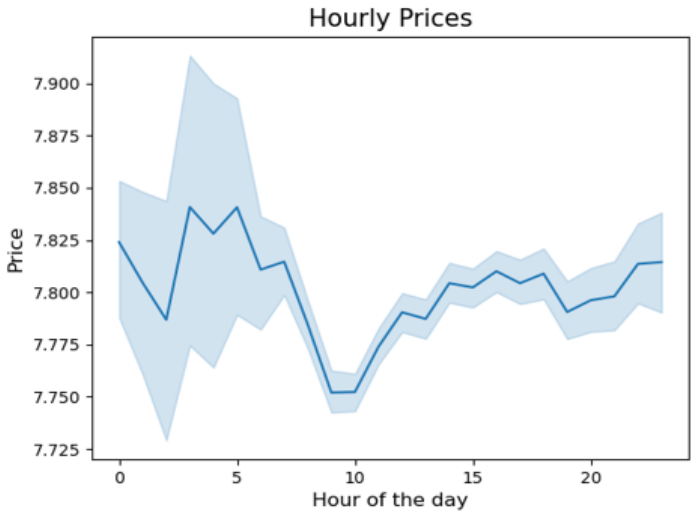


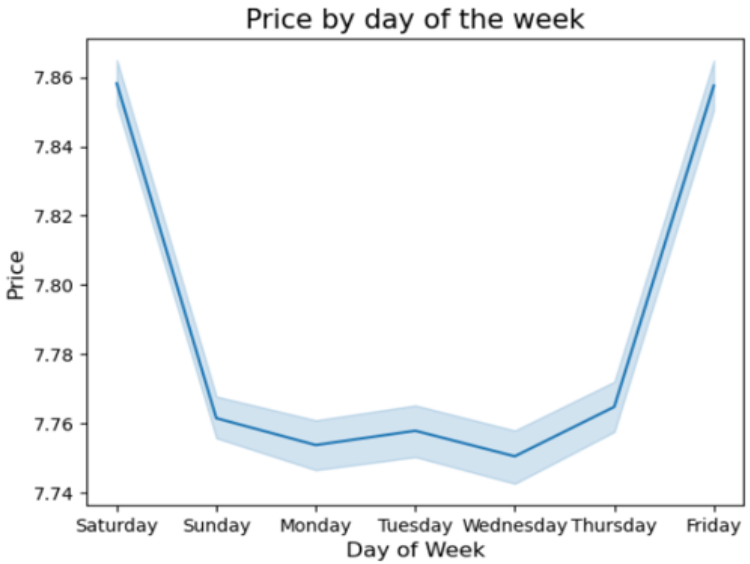


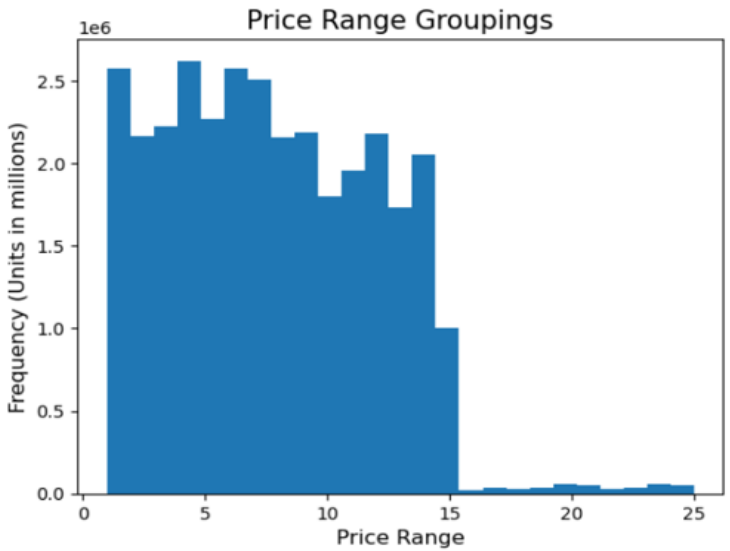
Visualizations

They want to know whether there are particular times of the day when people spend the most money, as this might inform the type of products they advertise at these times.

The hourly price at around hour 3-6 is when the prices are the most expensive. At hour 8-11, the prices will be the least expensive. Friday and Saturday's prices are the highest compared to the rest of the week.





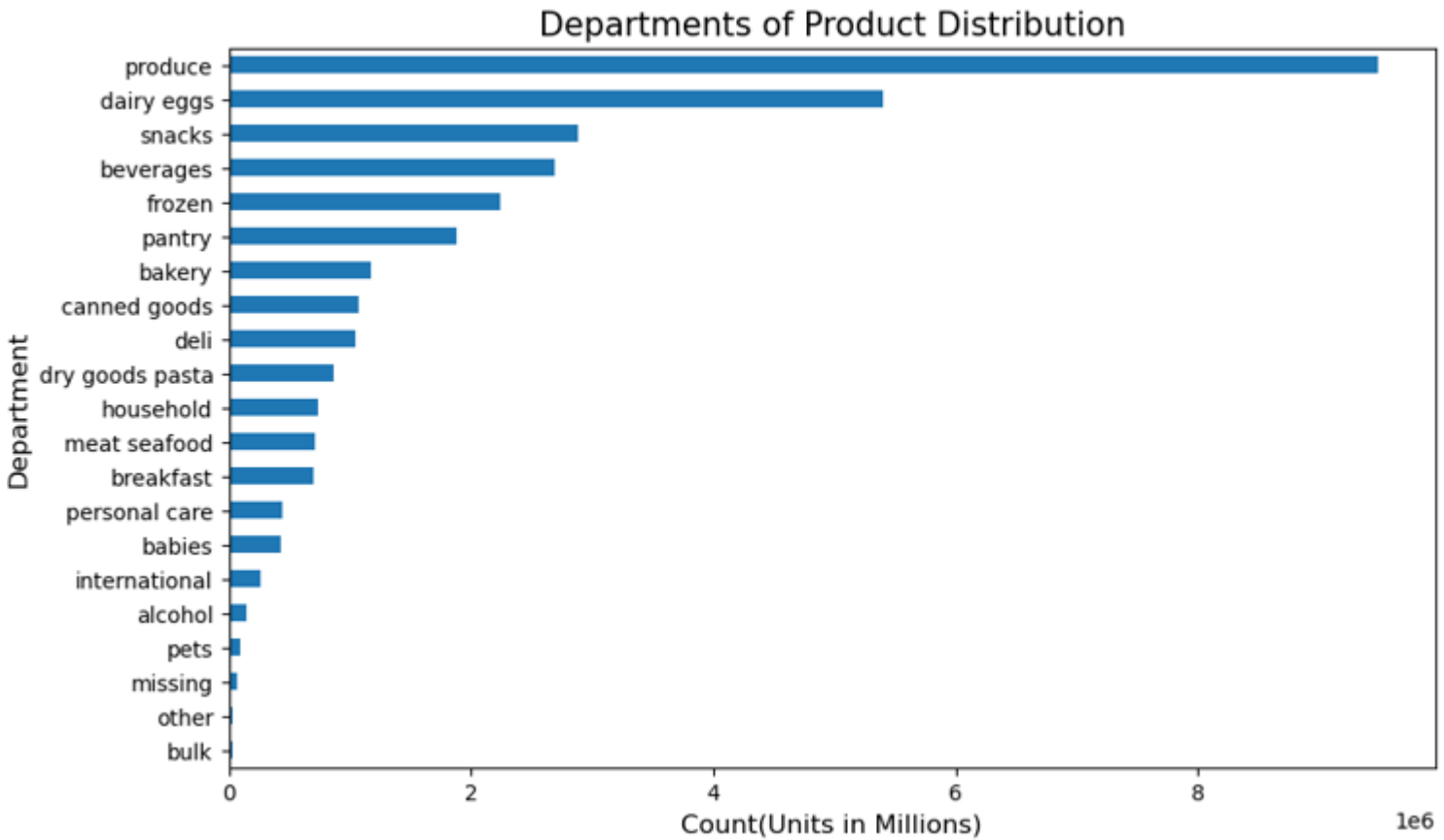


Instacart has a lot of products with different price tags. Marketing and sales want to use simpler price range groupings to help direct their efforts.

The most common prices range from 0 to 15. The least common price ranges from 16 and beyond.

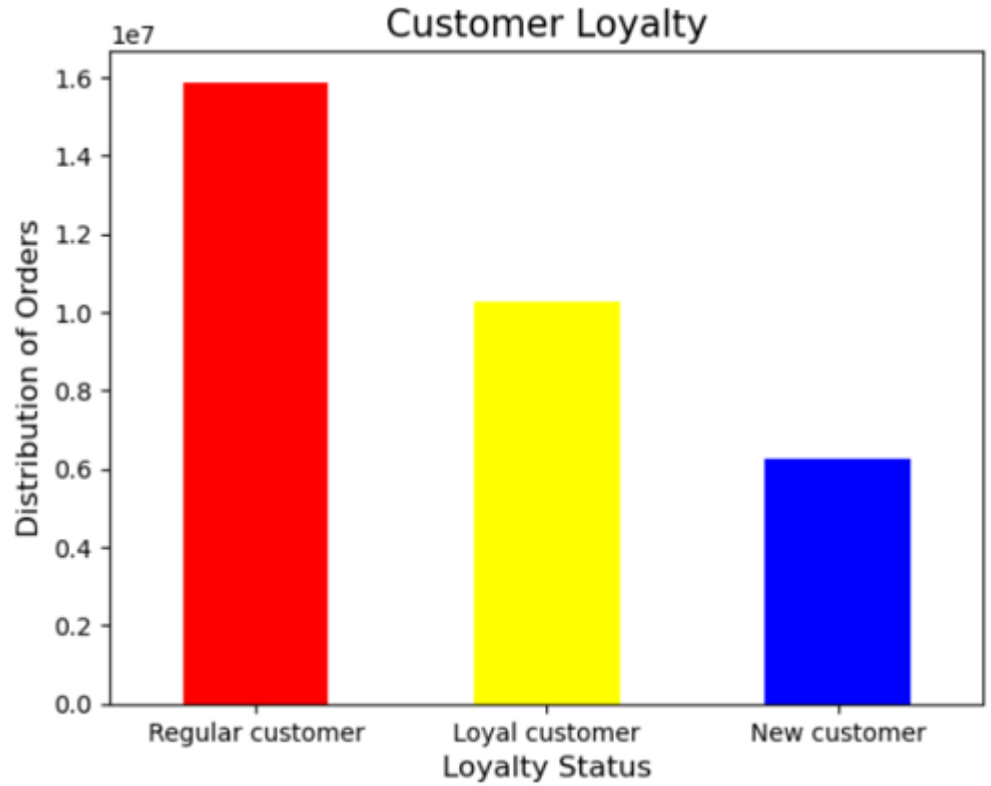
Are there certain types of products that are more popular than others? The marketing and sales teams want to know which departments have the highest frequency of product orders.

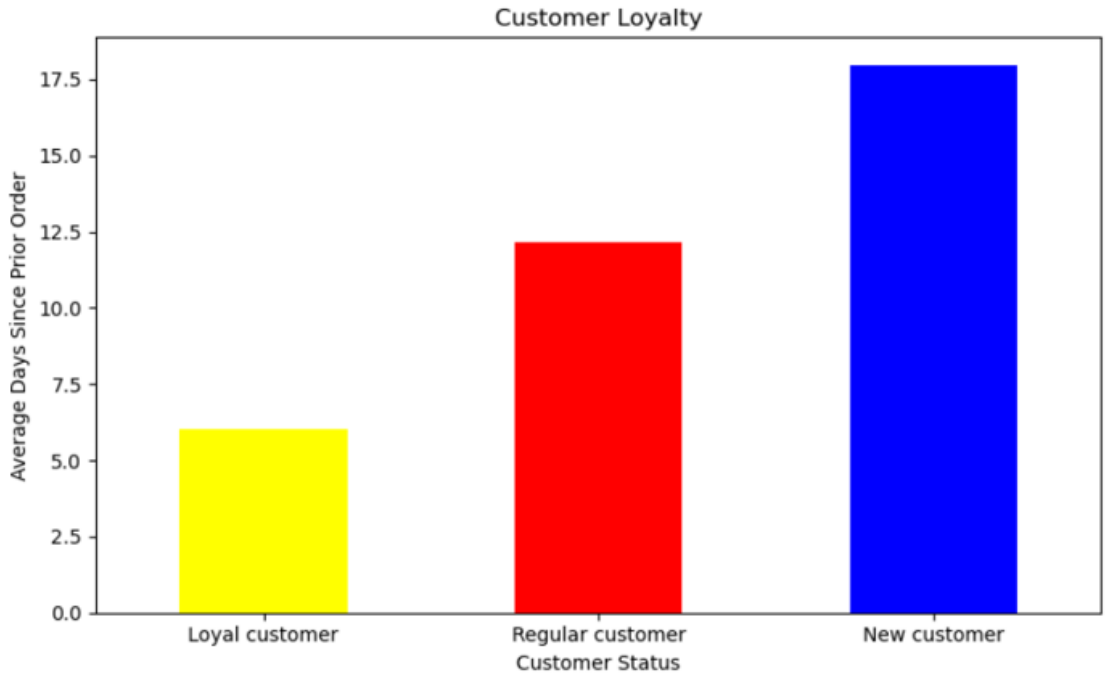
Produce, dairy eggs, and snacks are the most popular products.



Are there differences in ordering habits based on a customer's loyalty status?

Regular customers have the highest distribution of orders. New Customers have the lowest distribution of orders.



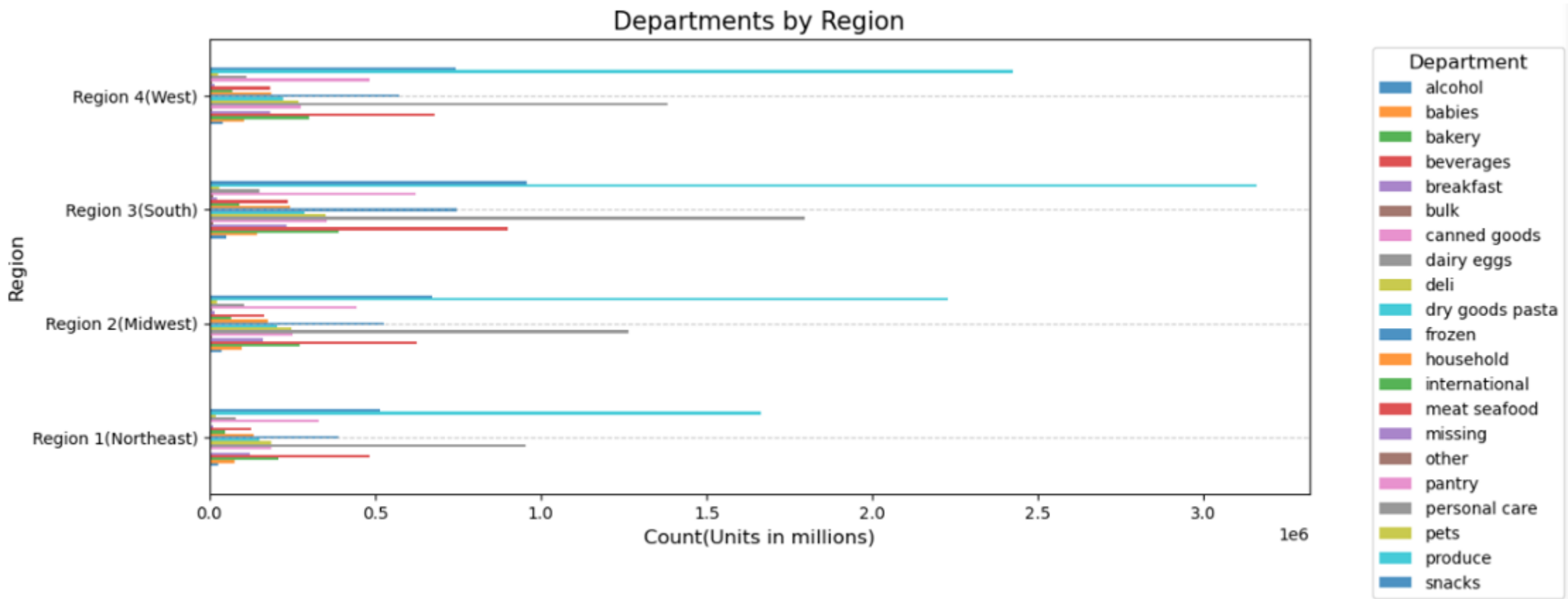


What's the distribution among users in regards to their brand loyalty (i.e., how often do they return to Instacart)?

Loyal Customers return at an average of around 6 days. Regular Customers return at an average of 12 days. New Customers return at an average of 18 days.

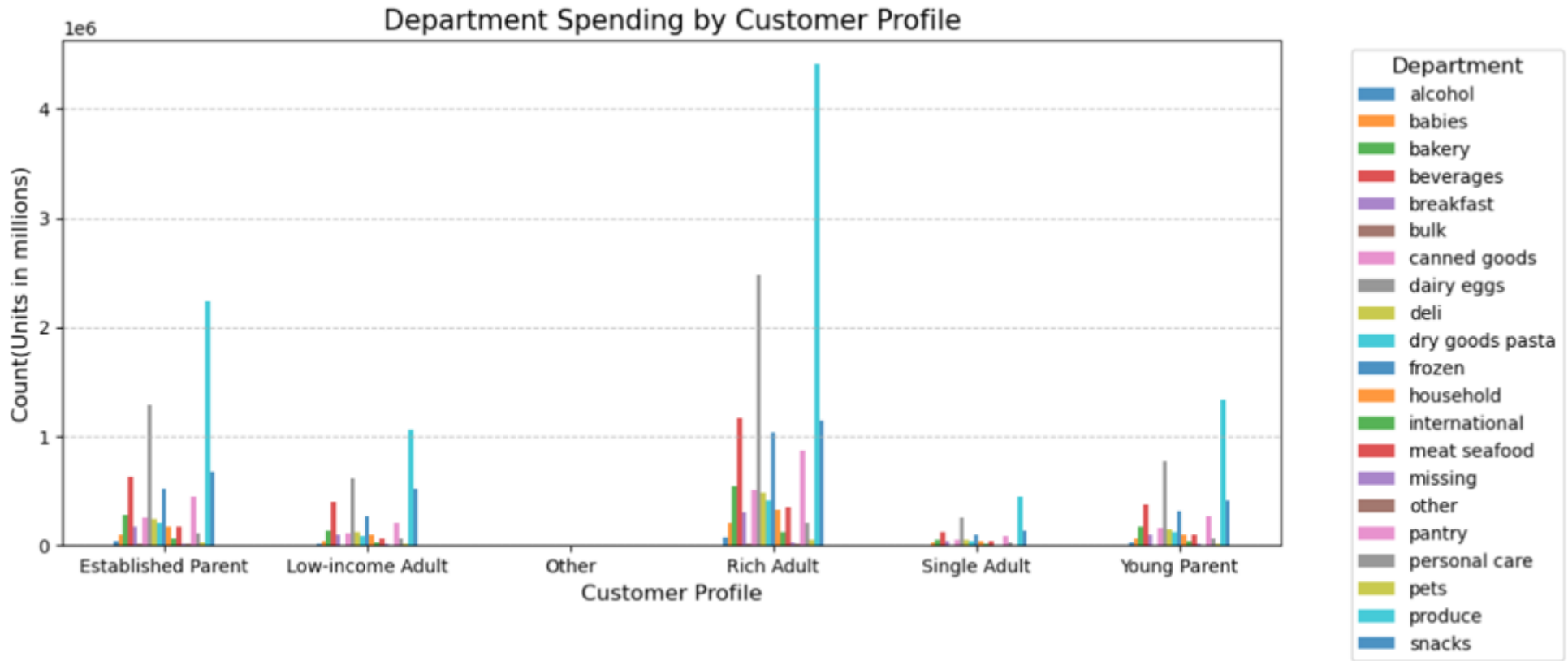
Are there differences in ordering habits based on a customer's region?

There is little to no difference other than the population of each region in ordering habits based on region. Produce seem to accumulate the highest number of orders followed by dairy eggs.



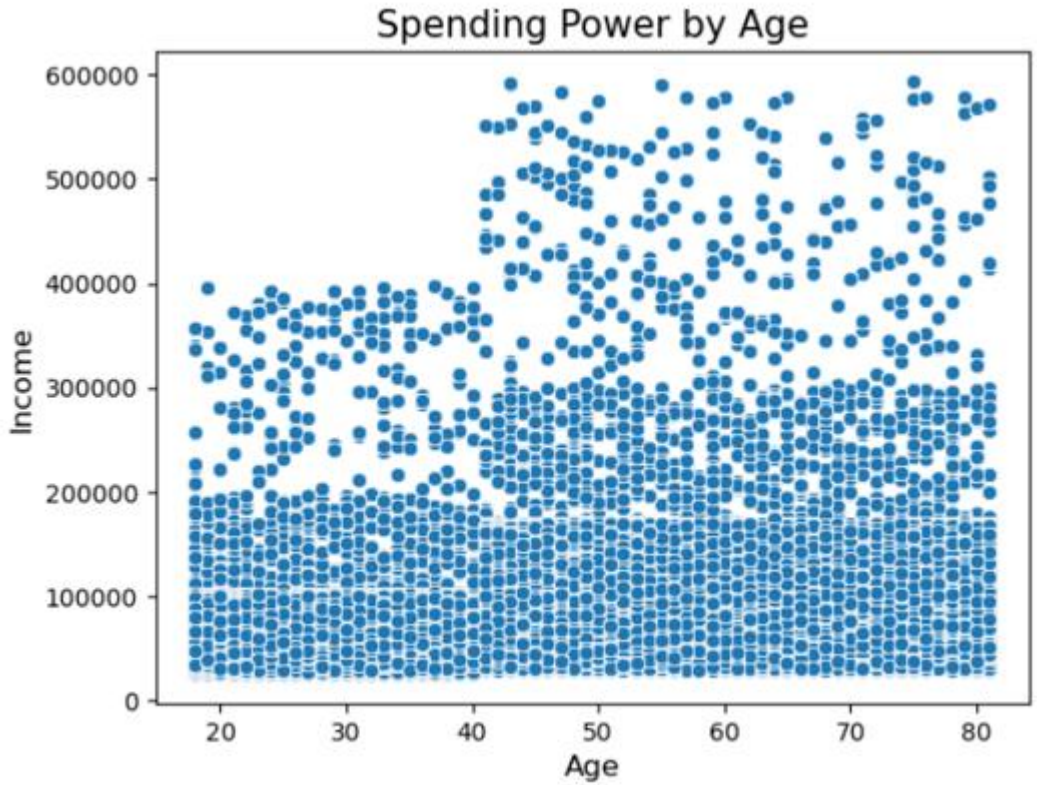
What differences can you find in ordering habits of different customer profiles? Considering the price of orders, the frequency of orders, the products customers are ordering, and etc.

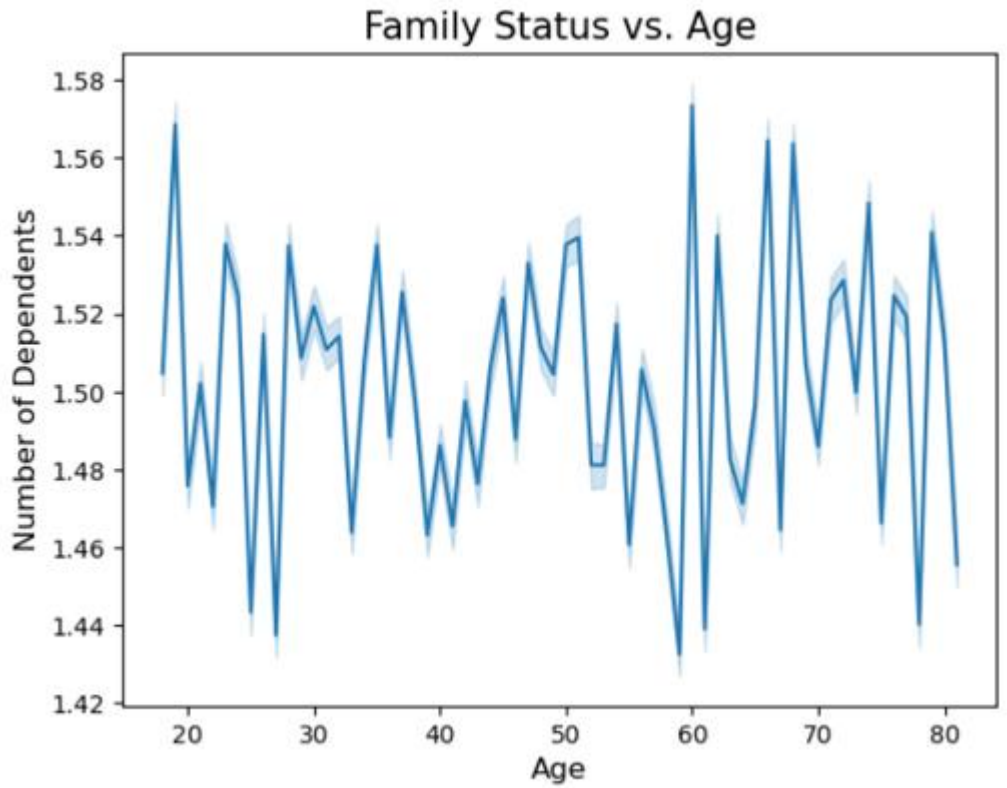
Other than the population of each customer profile, the order habits of customer profiles are still very similar with produce and dairy eggs having the highest amount of orders.



What different classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?

People over the age of 40 have a higher range of income compared to those under 40.
There is no correlation between the number of dependents and age.





Key Insights & Recommendations

Order Patterns

- Highest orders on Saturday & Sunday
- Lowest orders on Tuesday & Wednesday
- Peak hours: 9 AM – 4 PM
- Off-peak: 12 AM – 5 AM
    - *Promo Tip: Weekday-only deals & off-hour discounts*

Pricing

- Minimal price variance (~10¢)
- Most products priced between $0–$15
    - *Strategy: Focus inventory & promotions on $0–$15 range*

Product Trends

- Most popular: Produce, Dairy & Eggs, Snacks
- Least popular: Pet items, Alcohol, International foods
    - *Opportunity: Promote underperforming categories*

Customer Behavior & Growth Strategies

Retention & Loyalty

- Return time: Loyal (6 days) → Regular (12 days) → New (18 days)
- Regulars = most orders
  - *Strategy: Offer post-order discounts & launch a points-based loyalty program*

Region & Demographics

- Low order volume in Northeast & Midwest
- Age 40+ = higher income, higher spending
  - *Strategy: Boost regional marketing & target older customers*

Buying Patterns

- Families shop staples in bulk
  - *Opportunity: Highlight bulk/family-sized items in promotions*

Thank You!

Connect with me:

andycen7@gmail.com

LinkedIn

GitHub

(415) 806-3581