

Safe!

Reliable!

Convenient!



Easy!

Fun!

The data is a combination of multiple open-source data sets from Instacart. A customer data set was created and included for the purpose of this project and does not reflect the actual Instacart company.

[CUSTOMER DATA SET](#)

[DATA DICTIONARY](#)



Data wrangling

Data Merging

Deriving Variables

Grouping Data

Aggregating Data

Visualizations

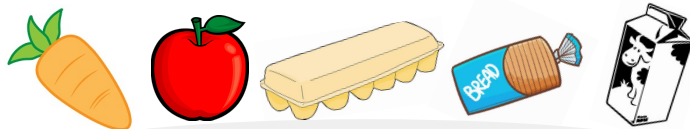


Consistency Checks

Wrangling Display

Derivation Display

Reporting



Shop from home!

BACKGROUND

KEY QUESTIONS

DATA

TOOLS - SKILLS



An online grocery company, instacart* is in need of exploratory data analysis to determine temporal sales patterns, customer purchasing habits, and department trends in order to create a targeted marketing campaign.

+ 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.

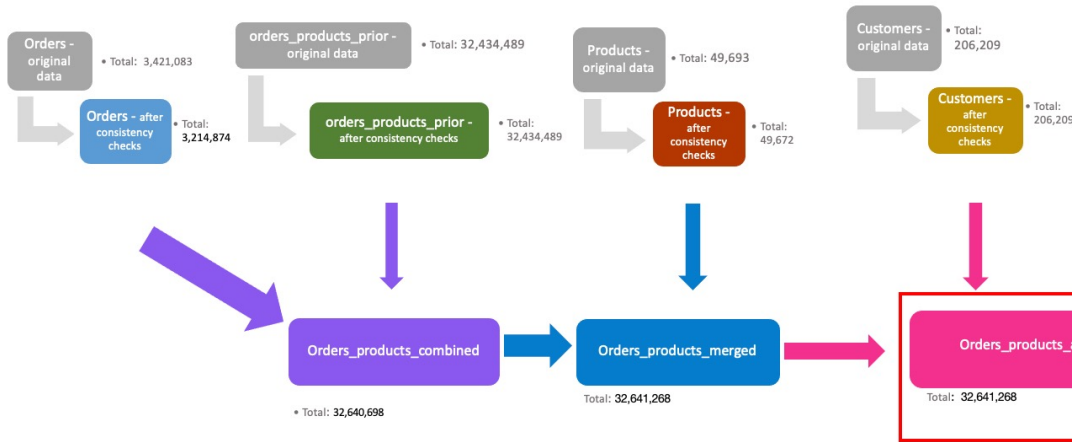
+ Are there certain types of products that are more popular than others? The marketing and sales teams want to know which departments have the highest frequency of product orders.

+ The marketing and sales teams are particularly interested in the different types of customers in their system and how their ordering behaviors differ.

* Instacart is a real company that has made their data available online. However, the contents of this project have been fabricated for the purpose of this Achievement.

InstaCart Analysis

Population flow

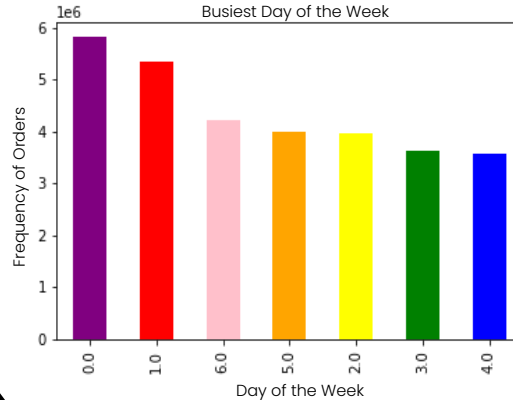
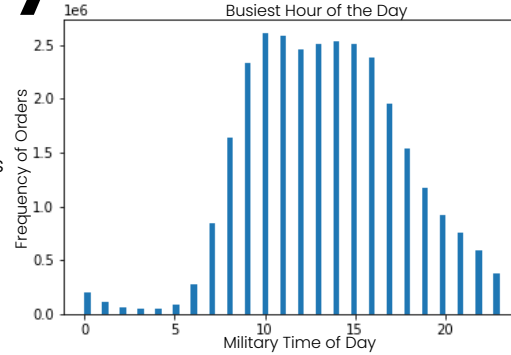


The flow chart above represents the original data and the process of exclusion and inclusion of data within the Jupyter notebook. The grey boxes represent original data. The second row of multiple colored boxes represents the data sets post-data checks and additions/subtractions including removal of missing values and duplicates for data integrity. Lastly, the third row is the merging of multiple data sets.

Analysis



By calculating the busiest hours of the day and busiest days of the week we can place targeted ads for those who shop more during this period and offer deals for those days when it is not so busy. The bar chart on the right indicate that orders increase close to 7 AM and eventual decline close to 7 PM.

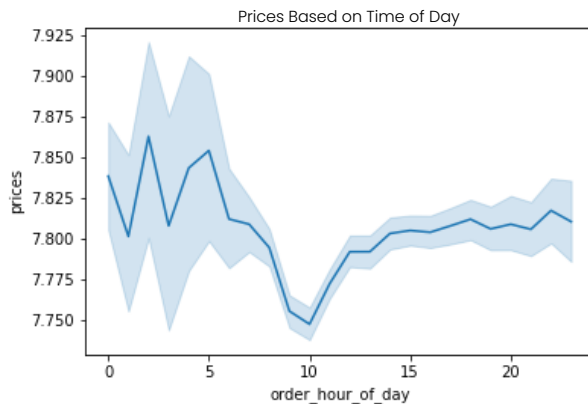


Note the bar chart on the left indicates the busiest days of the week being Saturdays and Sundays. Utilizing this data analysis we can offer greater rewards and offers during less busier days which are shown to be Wednesdays and Tuesdays.

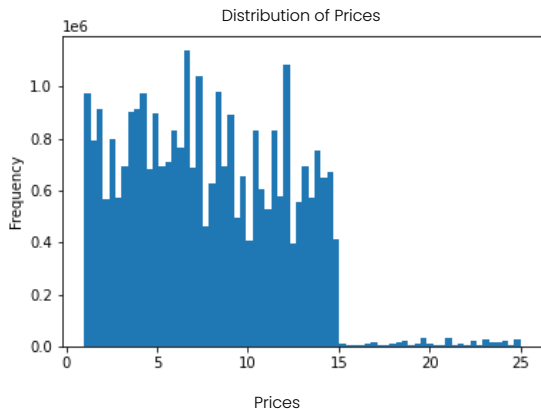
Days Of the Week: 0 = Saturday, 1 = Sunday, 2 = Monday, 3 = Tuesday, 4 = Wednesday, 5 = Thursday, 6 = Friday



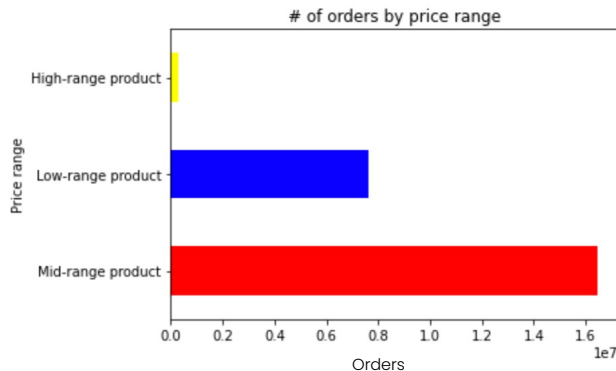
InstaCart Analysis



The line chart to the left indicates decreased spending at 10 & 11am. Increased spending is noted between 2 & 7am. Expenditures increase again in the late afternoon and evening. Based on the analysis, pop-up sales catering to evening and lunch hours should be offered to encourage sales. Increased sales are occurring during morning hours. Rebates/deals need to be offered for lunch, afternoon and evenings hours.



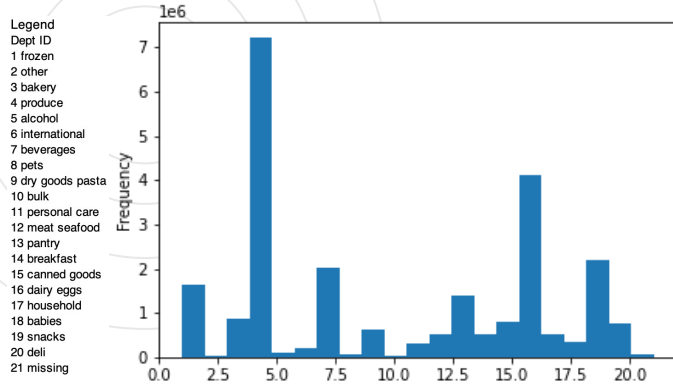
The histogram to the left indicates that most products bought are priced between \$1 and \$15.



The bar chart to the left indicates that the MID-RANGE (Cost of \$10-\$15) products have an increased amount or orders when compared to high range products, or items costing greater than \$15. High-range products are suffering in sales and need to be coupled with mid-low-range products for combo deals. Taking % rebates, sales off of High-Range products or offering alternatives when buying Mid-range products will increase sales.



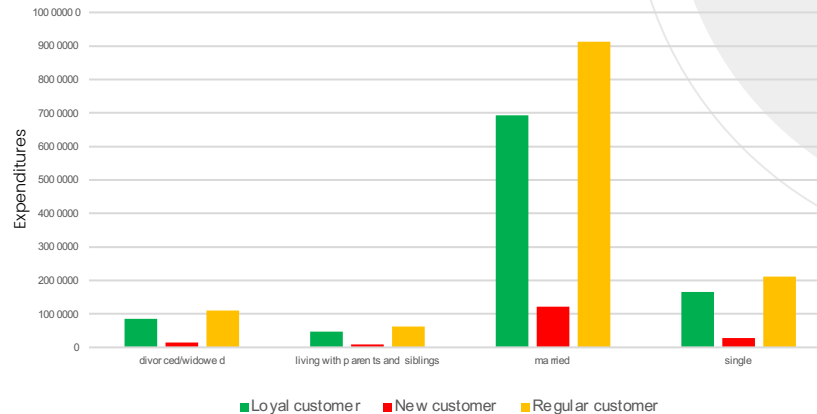
InstaCart Analysis



The histogram at left indicates that most popular department is produce followed closely by dairy eggs. We can also see the least popular departments that include bulk and frozen items. Bulk, pets, alcohol and international item sales are suffering, again coupling rebates when customers buy higher ordered items such as produce and dairy can increase sales and awareness that other items exist in the Instacart product offerings. People will buy items if you make them aware products do exist by offering rebates on coupled lows ordered items.

As the bar chart visualization to the right shows, there is a significant representation of married customers with kids who make up most the orders. Customers who are married with kids have higher expenditures with a higher cost of living due to increased individuals within the family circle. Increased ads need to be directed with offerings that apply to the other brand loyalty representation among the different section/regions of the U.S. market.

Marital Status & Spending Habits



Roadmap To Online Retail

Success =  instacart

1 Pop-Up flash Sales, Instant Rebates and double rewards need to be offered during decreased sales hours and days.

3 Lower ordered items need more revenue generation, targeting ads with sales on these items will increase sales.

5 Coupling rebates when customers buy higher ordered items such as produce and dairy can increase sales and awareness that other items that exist in the Instacart product offerings.

2 Taking % rebates, sales on of High-Range products or offering alternatives when buying Mid-range products will increase sales significantly.

4 People will buy items if you make them aware products do exist by offering rebates on coupled lower ordered items.

6 Customers who are married with kids have higher expenditures with a higher cost of living due to increased individuals within the family circle. Increased ads need to be directed with offerings that apply to the other brand loyalty representation among the different section/regions of the U.S. market.

A circular frame containing a desk scene. In the background, there is a potted plant with green leaves in a white pot with a lace-like top. To the right of the plant is a large, black, three-dimensional letter 'A'. Further right is a small, square wooden block with a circular design. The desk surface is light-colored wood. The text 'UNLEASH YOUR DATA' is overlaid on the scene, with 'UNLEASH YOUR' in white and 'DATA' in red. Below it, 'UNLEASH YOUR POTENTIAL!' is overlaid, with 'UNLEASH YOUR' in white and 'POTENTIAL!' in red.

UNLEASH YOUR **DATA**
UNLEASH YOUR **POTENTIAL!**

Thank you!