Instacart Grocery Basket

An online grocery store that operates through an app.

	Objective To uncover information about Instacart's sales patterns. Performed an initial and exploratory analysis of their data to derive insights and suggested strategies for better segmentation based on the provided criteria.

	Data ● "The Instacart Online Grocery Shopping Data 2017" Accessed from here

	Limitations ● Customer information limited to age, gender, income, and marital status. ● Dataset only contains data from 2017.

	Skills Data Cleaning: Wrangling and subsetting. Data consistency checks. Combining and exporting data. Deriving new variables. Grouping data and aggregating variables. Used Python for visualization and excel for reporting.

|	
	Tools MS Excel, Python, Anaconda, Pandas, Numpy, Jupyter, Seaborn, Matplotlib. Data wrangling and visualization In Python and final report in Excel.

	●Check the Python scripts here ●Check the report here

Initial Analysis

- Scanned dataset for irregularities using Python.
- Noted steps followed during data wrangling and cleaning to facilitate better process understanding among peers.
- Merged dataframes to identify trends and patterns.
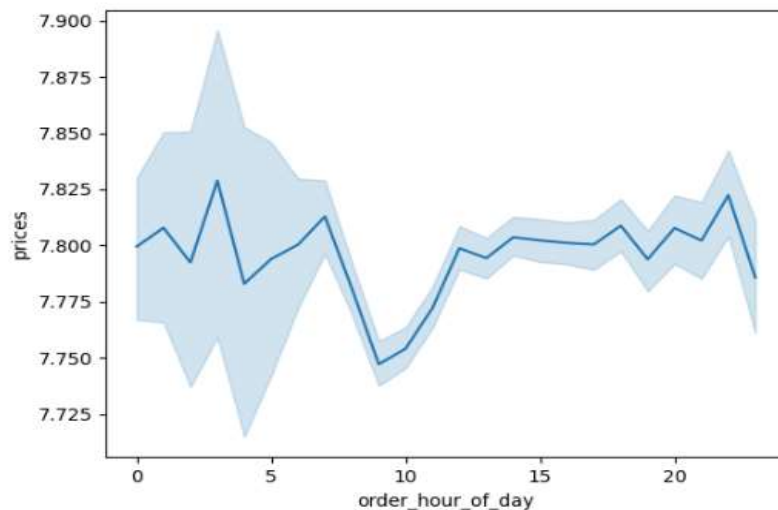
instacart

Population flow
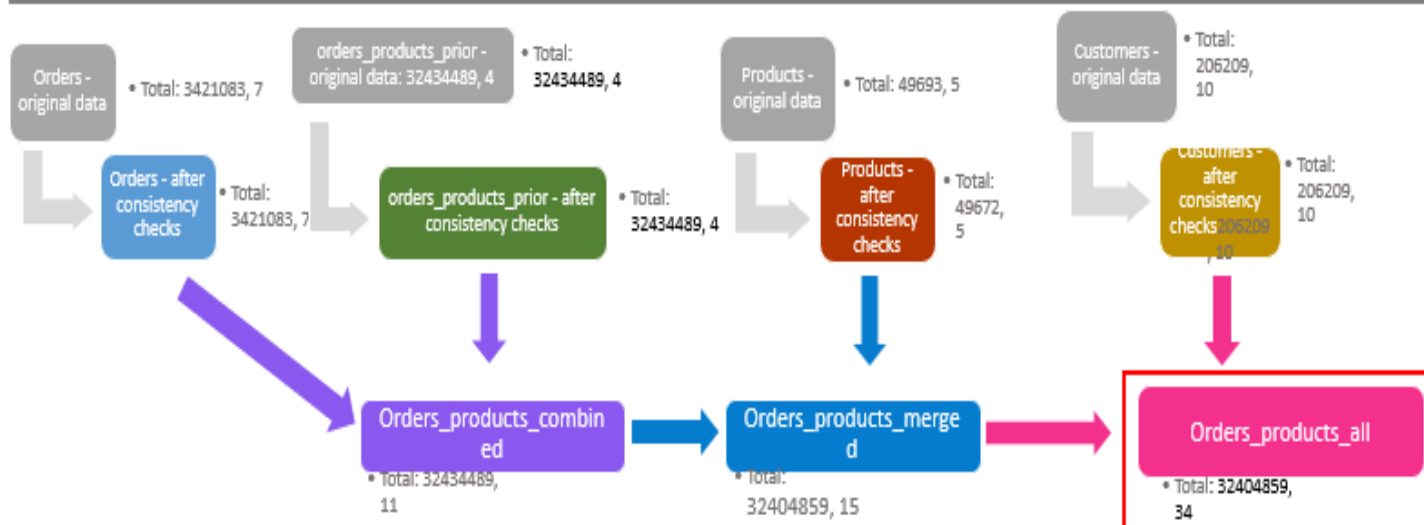
Customer Profiling and Behavioral Analysis

- Created crosstabs from the merged dataframes to understand correlation between variables.
- Used Seaborn, Matplotlib to create charts that highlight the trends across regions and customer profiles.
- Regular customers are more than both Loyal customers and new customers, but Loyal customers return to Instacart most often with an average of 5.9. Regular customers' return to Instacart average 11.6, while that of new customers is the longest with 15.23.
- The most popular products are produce-4, then dairy eggs-16 and then snacks-19. The most frequently ordered products is the snack-19.
- The time of the day with high spendings are between 12am to 6am, while the weekend is the busiest time of the week for Instacart orders and the hours of the day with most sales are from 9am to 5pm.
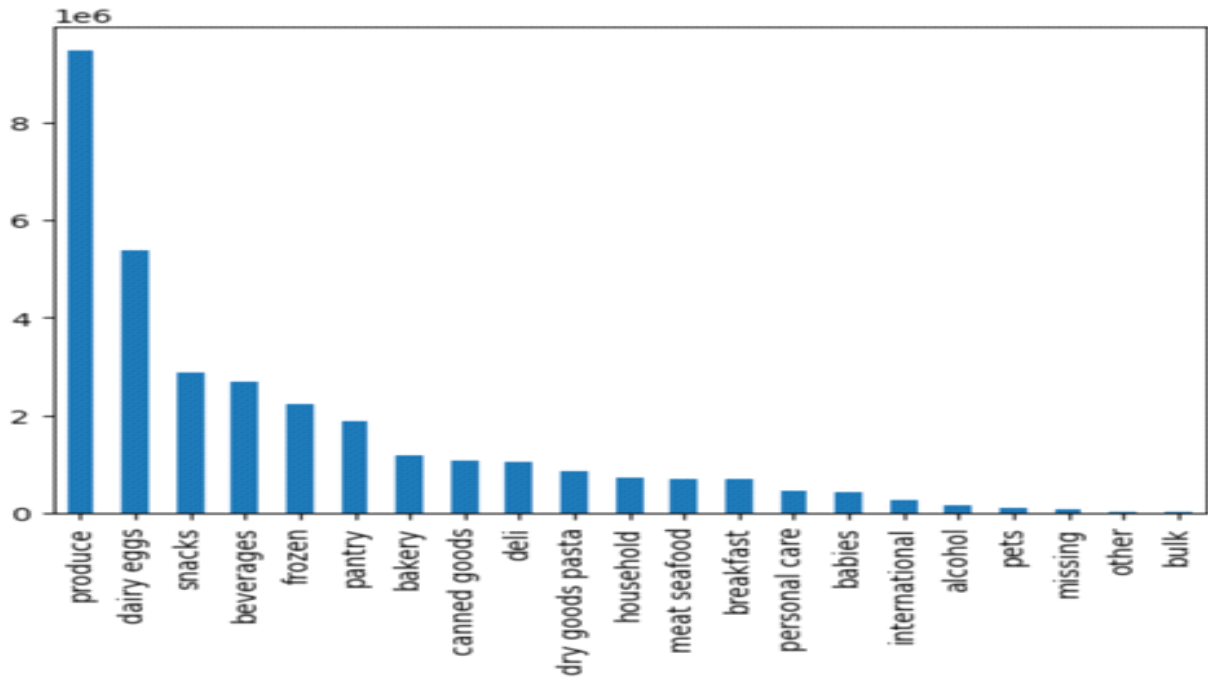
Spending power vs age.
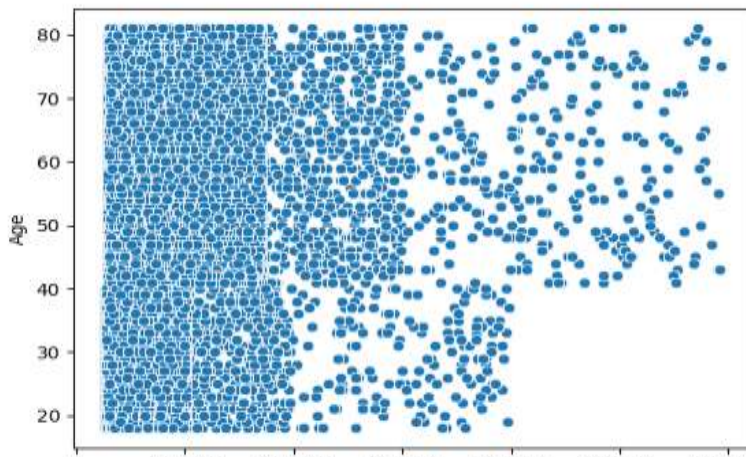
The scatterplot shows that most customers earn less than $200, 000 and below and it cuts across all the ages (18 to 81 years). Most of the customers that earn below $400, 000 are ages 40 years and below. While the customers that earn above $400,000 are all above 40 years.

The prices peak at about 3am and the lowest prices are at about 9am.

The most popular products are produced, then dairy eggs.
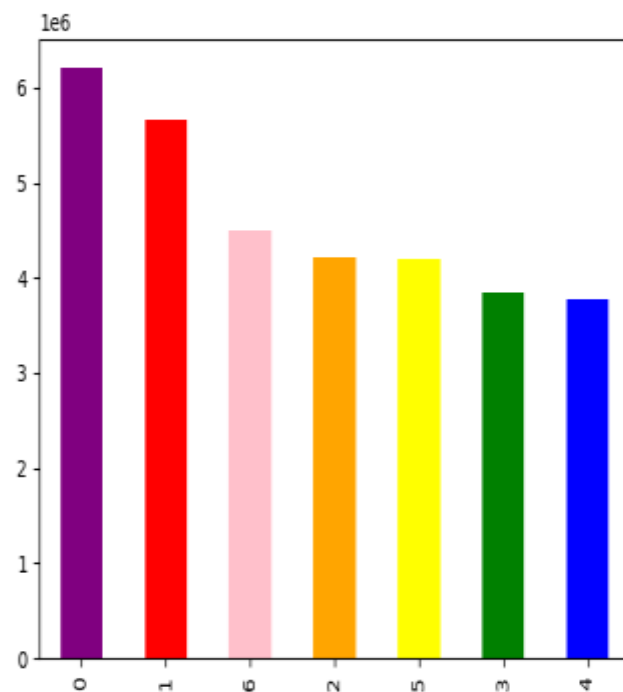
0- Sunday
1- Monday
6-Fiday
2-Tuesday
5- Friday
3- Wednesday
4-Thursday

The weekend is the busiest time of the week for Instacart orders.

RECOMMENDATIONS

ADVERTISEMENTS: I will recommend that ads at times should be fixed between the hours 11pm to 6am on Tuesdays, Wednesdays, and Thursdays.

PRODUCT ENDORSEMENTS: The ads should focus on fast-moving items such as dairy eggs, beverages, produce, bakery, and deli while the inclusion of slow -moving goods like international, pantry, and personal care to boost product awareness.

SALES STRATEGY: Offering special discounts and rewards through a point system tailored to members' interests maintain loyal buyers and converts non-frequent buyers' status to a higher level.

TARGET MARKET: The south region generates the highest traffic activities overall, while some least popular segments like babies, breakfast, dry goods pasta, and households can possibly grow well if positioned properly.

PRICING STRATEGY: Most customers have low to mid-range incomes, thus spending power may likely stay within these ranges to fit their budget. Investigating competitors' pricing schemes and conducting surveys on consumers' household expenditures can disclose key information for the effective implementation of pricing tactics.

NEW PRODUCT DEVELOPMENT: An extensive marketing research analysis on high-end goods to expand the services on the top-notched brands to new target audiences. Existing high-income members may be interested in the premium brands and an exclusive offer on executive packages can spark new interests and eventually extend business in that area.