REPORT

VISUALISATIONS AND RECOMMENDATIONS

Prepared by: Iryna Smologonova

July 2022

Project Name: **Instacart Grocery Basket Analysis**

Date: July 6, 2022

Analyst Name: Iryna Smologonova

Objective:

Instacart, an online grocery store that operates through an app. Instacart already has very good sales, but they want to uncover more information about their sales patterns. My task is to perform an initial data and exploratory analysis of some of their data in order to derive insights and suggest strategies for better segmentation based on the provided criteria.

Context:

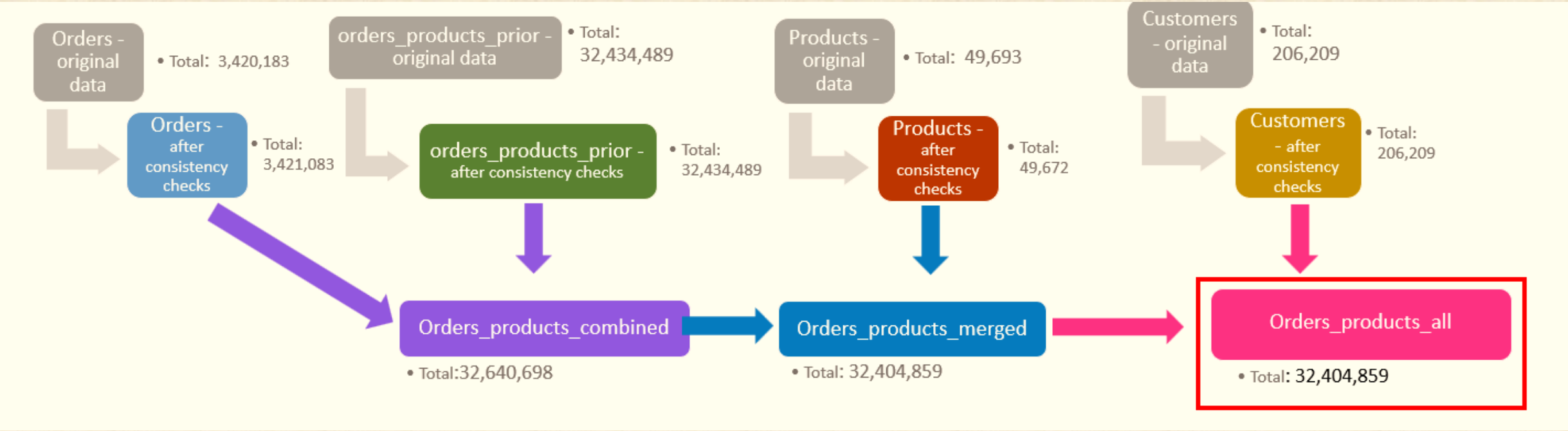
The Instacart stakeholders are most interested in the variety of customers in their database along with their purchasing behaviors. They assume they can't target everyone using the same methods, and they're considering a targeted marketing strategy. They want to target different customers with applicable marketing campaigns to see whether they have an effect on the sale of their products. My analysis will inform what this strategy might look like to ensure Instacart targets the right customer profiles with the appropriate products.

Data Sources:

"The Instacart Online Grocery Shopping Dataset 2017" accessed from https://www.instacart.com/datasets/grocery-shopping-2017 on June 7, 2022".

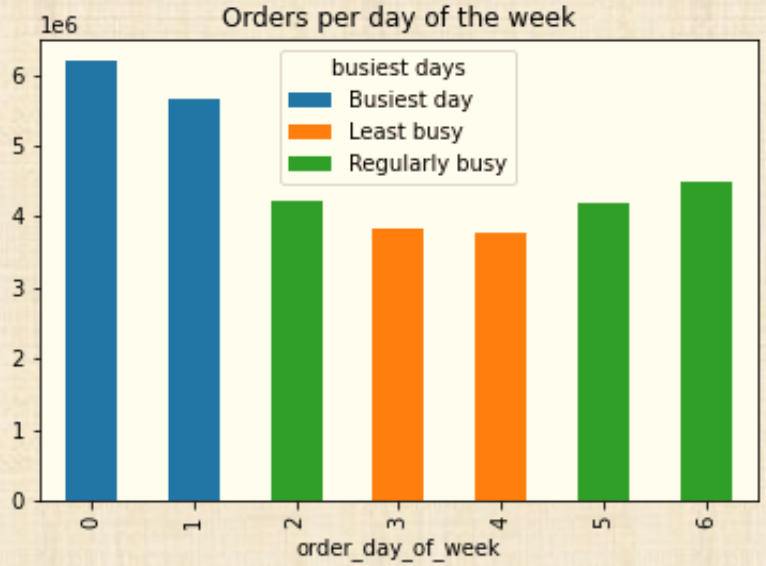
All data used has been taken from Instacart open-source data sets. The customer data set is fictional and used solely for the purpose of this analysis.

Population flow



Visualizations & Recommendations

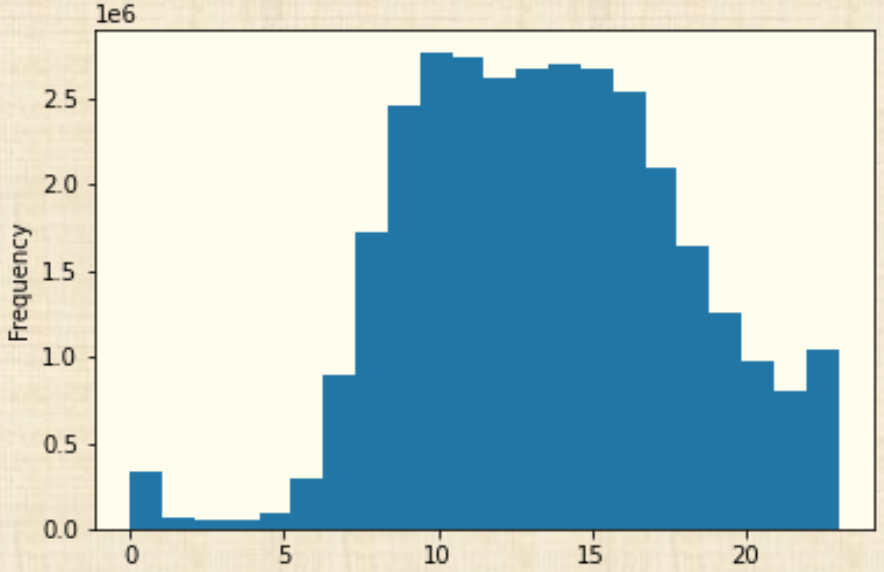
Key Question 1: The sales team needs to know what the busiest days of the week and hours of the day are (i.e., the days and times with the most orders) in order to schedule ads at times when there are fewer orders.



The bar chart displays the number of orders made for every day of the week

0 = Saturday, 1 = Sunday, 2 = Monday, 3 = Tuesday, 4 = Wednesday, 5 = Thursday, 6 = Friday

Answer: The most orders are made over the weekends Saturday and Sunday. The least busy days are the middle of the week: Tuesday and Wednesday



The histogram displays the number of orders made hourly (24-hour format)

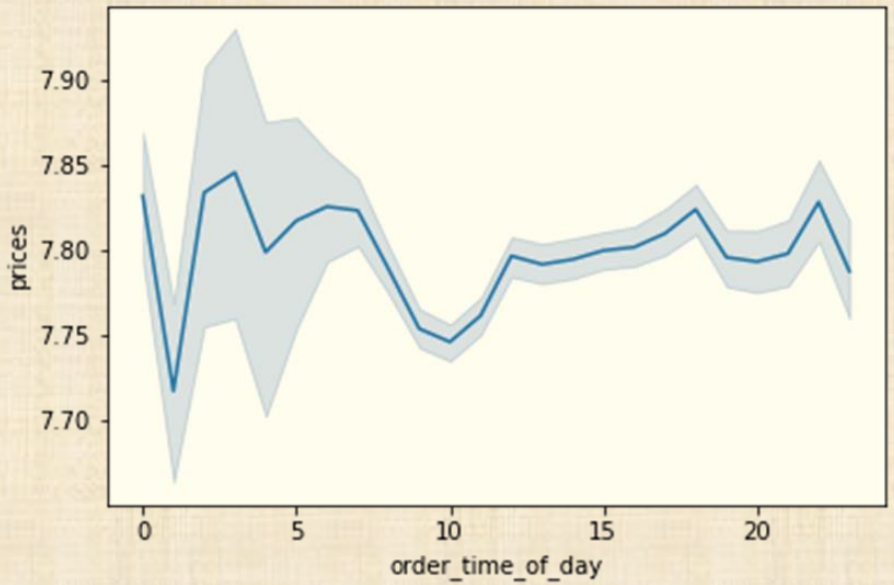
Answer: The busiest hours of the day are correlated with 9-5 work hours and started to drop down after 4 pm

Recommendation:

Ads should be running on Monday-Wednesday after 6pm with a target to increase number of orders on Tuesday and Wednesday between 6 pm and 10 pm.

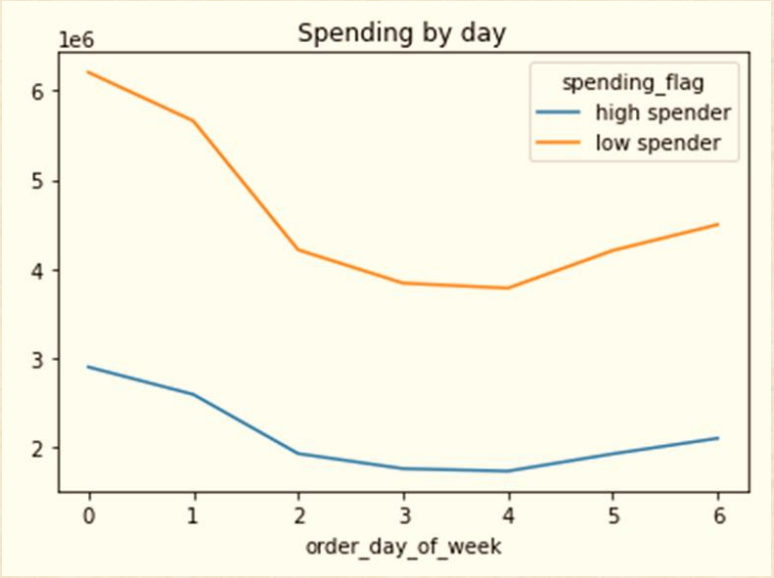
Visualizations & Recommendations

Key Question 2: The sales team also wants 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 line chart displays the average prices of goods purchased for each hour of the day

Answer: The most expensive products are ordered between 2 am and 7am, 6 pm and 10 pm.



The line chart shows the number of orders made by 'high spender' price >=$8) and 'lower spender' (price <$8).
0 = Saturday, 1 = Sunday, 2 = Monday, 3 = Tuesday, 4 = Wednesday, 5 = Thursday, 6 = Friday

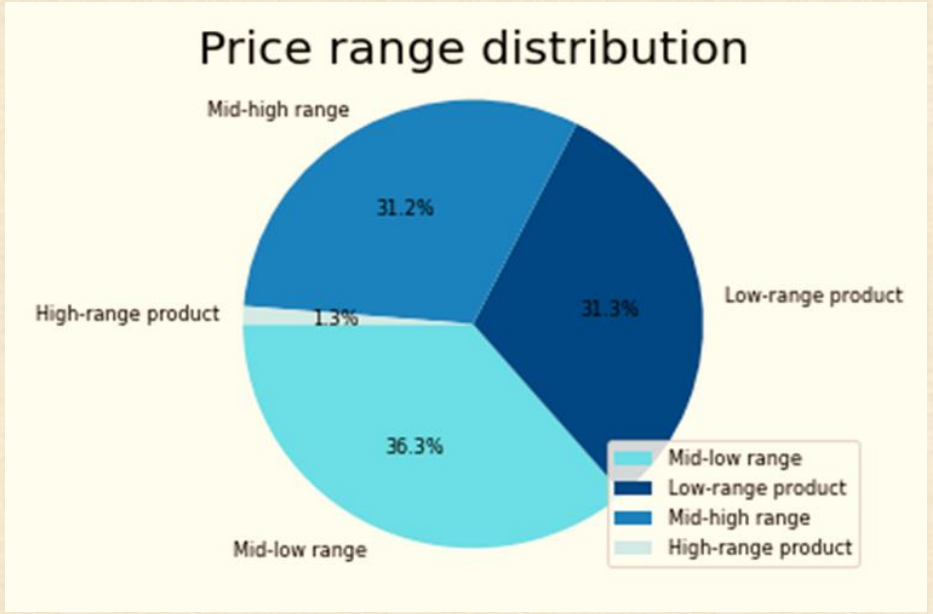
Answer: The main customer base consists from 'lower spenders' and there are fewer 'high spenders'

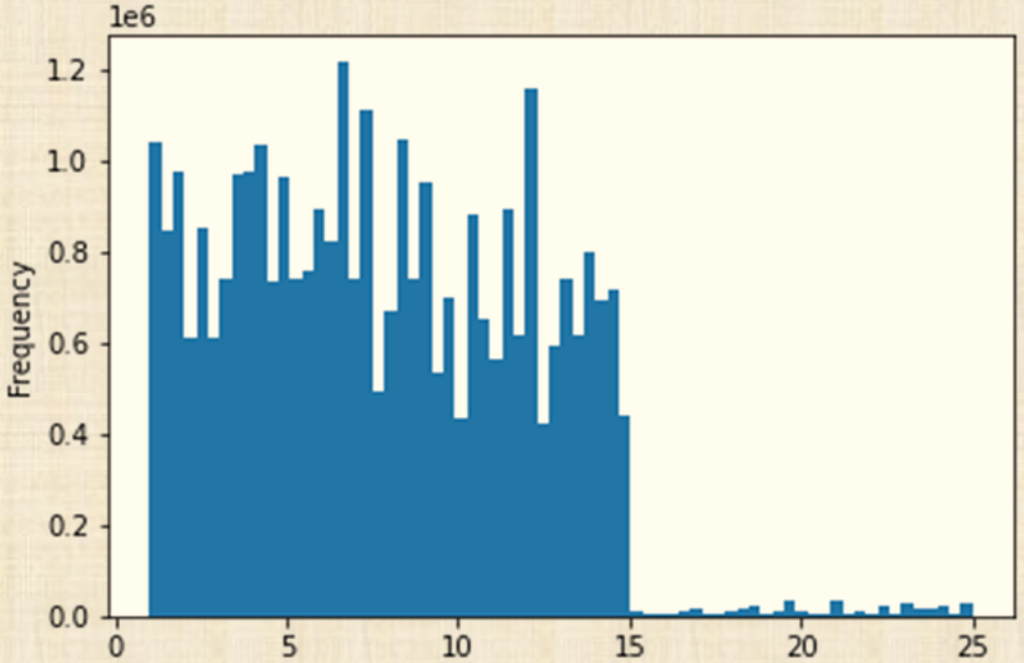
Recommendation:

Higher priced products can be promoted between 2 am and 7 am, but the main focus of ads on lower spenders at time slot between 6 pm and 10 pm.

Visualizations & Recommendations

Key Question 3: **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 pie chart displays the distribution of products in different price range.

The prices have been grouped into high range product (>$15), mid-high range product (<=$15 & >$10), mid-low range product (<=$10 & >$5), and low-range product (<=$5).

The histogram displays the number of products in goods' prices.

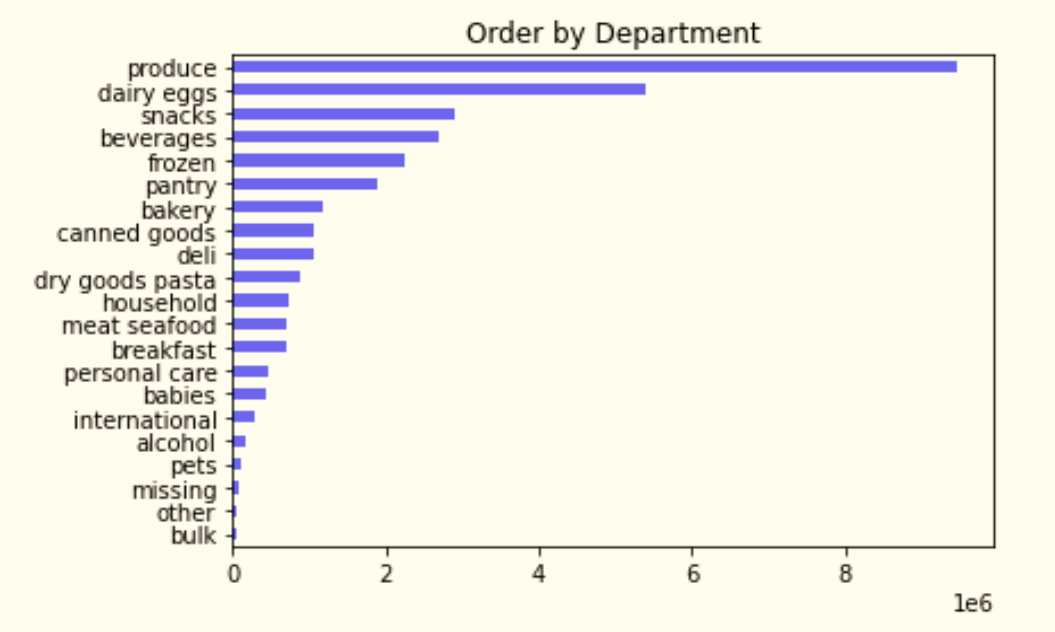
Answer: Mid-high and mid-low range products are the most popular products. High-range products are the least popular products.

Recommendation:

As most products are mid-high and mid-low range ($5-$15), I would recommend finding a suitable price in this range that most items can be sold at.

Visualizations & Recommendations

Key Question 4: 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 bar chart shows the number of goods purchased from each department

Answer: The most popular products are from produce, dairy eggs, snacks, and beverages departments. The fewest orders come from the bulk, pets and alcohol departments.

Recommendations:

- Promotion could be focused on BOGO including the products from the bottom
- Improve logistic or assortment of the bottom product
- Take into consideration increasing margin of the top-5 categories

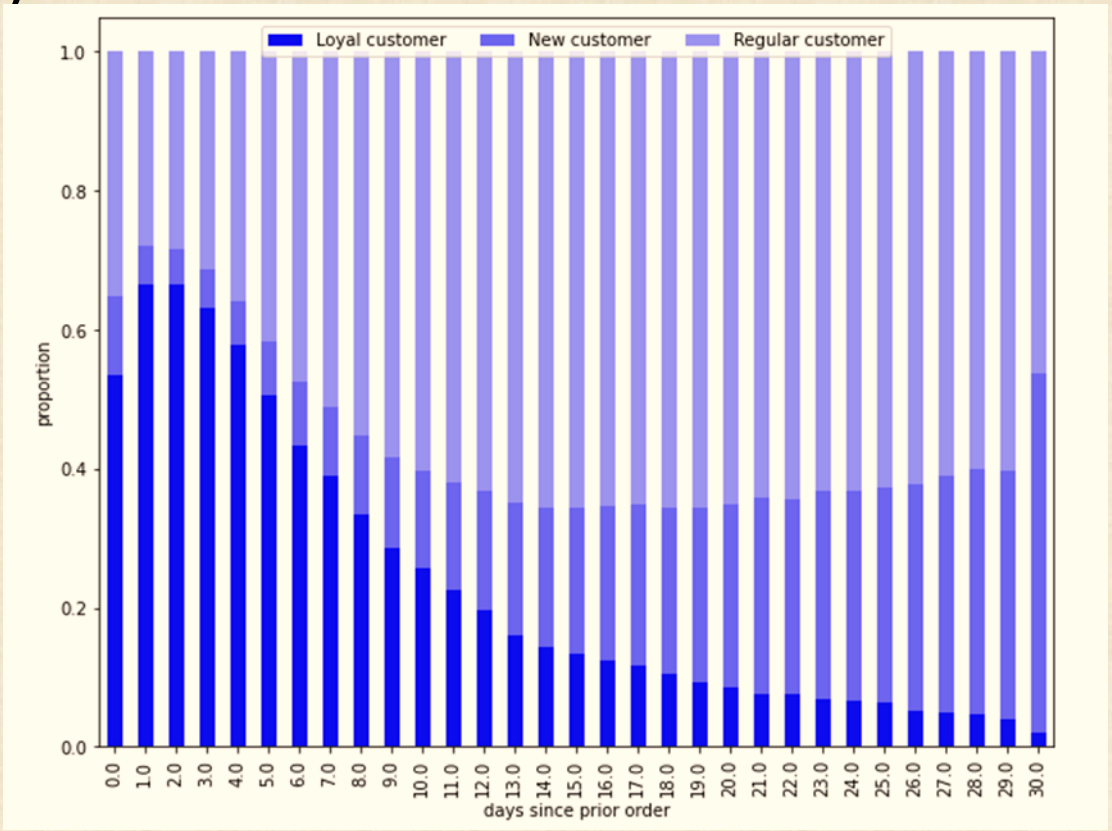
Visualizations & Recommendations

Key Question 5: What's the distribution among users in regards to their brand loyalty (i.e., how often do they return to Instacart)?

The pie chart shows the percentage of customers in different loyal categories
Answer: 31.7% of clients are loyal customers placed > 40 orders, 49% - regular customers placed 11-40 orders, and 19.3% - new customers placed < 10 orders

loyalty_flag	days_since_prior_order mean
Loyal customer	6.027344
New customer	17.984375
Regular customer	12.156250

Loyal customers on average order every 6 days, where New and Regular customers every 18 and 12 respectively.



The bar chart shows the number of days when the customer made the repeating order
Answer: Customers usually return to Instacart within 2 weeks. Majority of New and Regular customers return after two weeks.

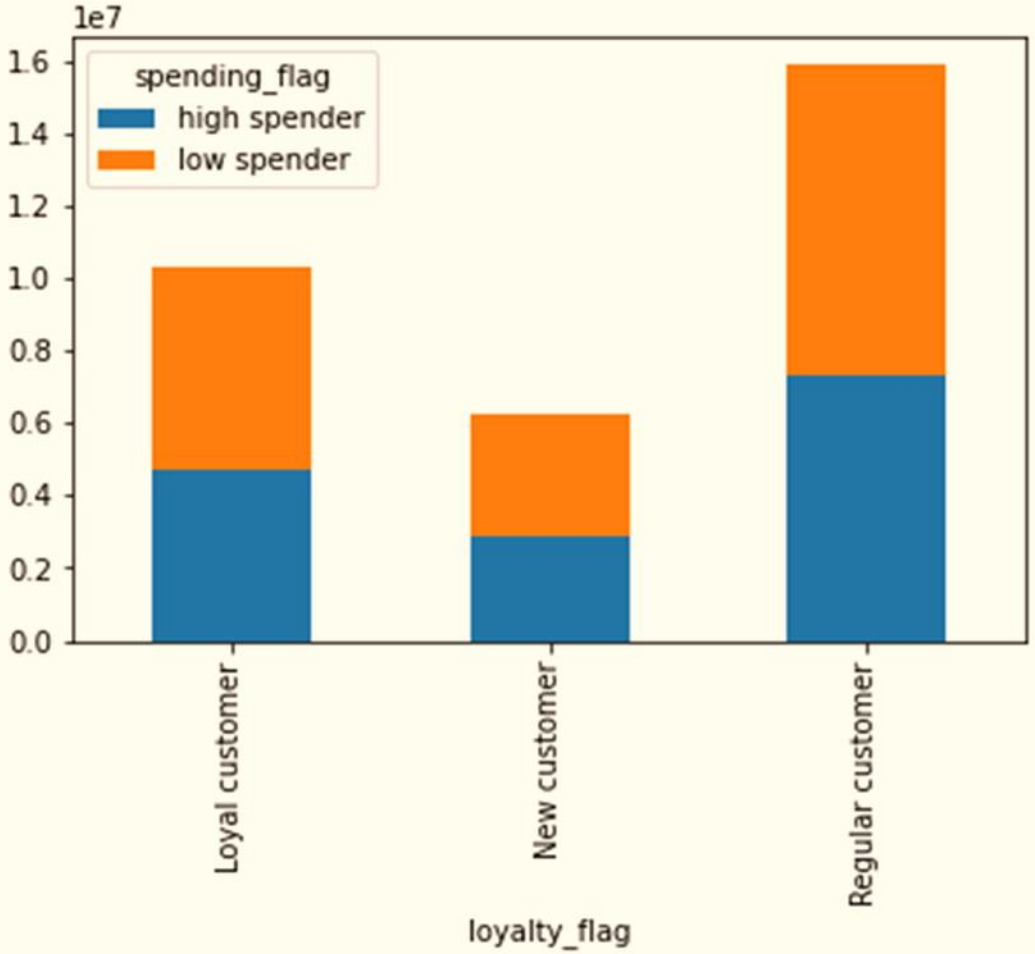
Recommendations:

- Target regular customers with orders 30+ to reach min 40 orders per year and transfer them into loyal category.

- Promotion could be focused on regular and new customers who placed the order after 2 weeks to encourage them to place the next order within 10 days, for example giving them a promotional coupon expire in 10 days.

Visualizations & Recommendations

Key Question 6: Are there differences in ordering habits based on a customer's loyalty status?



The bar chart shows the total orders in different loyal categories

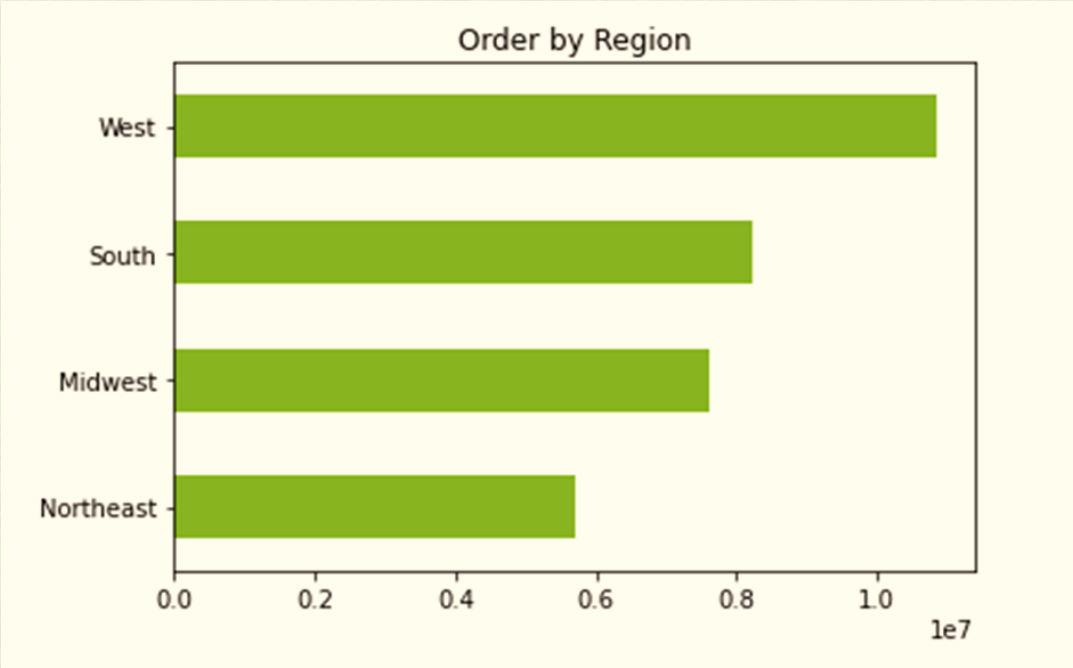
Answer: Regular customers order the most while new customers order the least

Recommendation:

Targeting new customers could increase sales

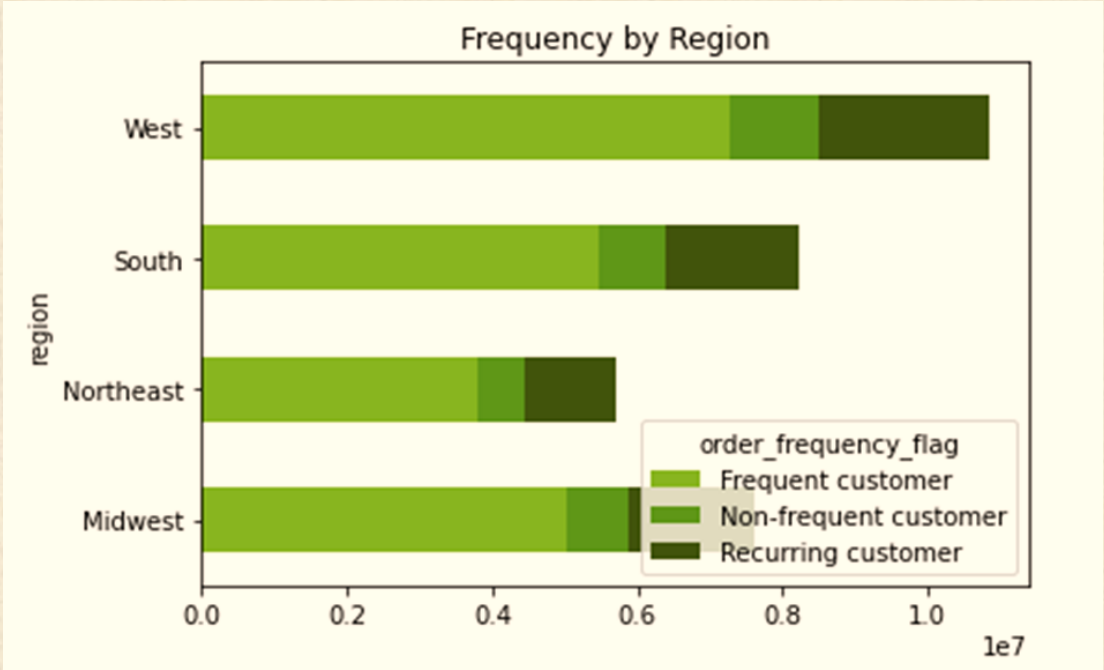
Visualizations & Recommendations

Key Question 7: Are there differences in ordering habits based on a customer's region?



The bar chart shows the total orders by region

Answer: The West has generated the most Instacart orders and the Northeast the fewest orders compared to the other regions that is correlated with the US population



The bar chart shows the orders by region come from frequent , non-frequent and recurring customers.

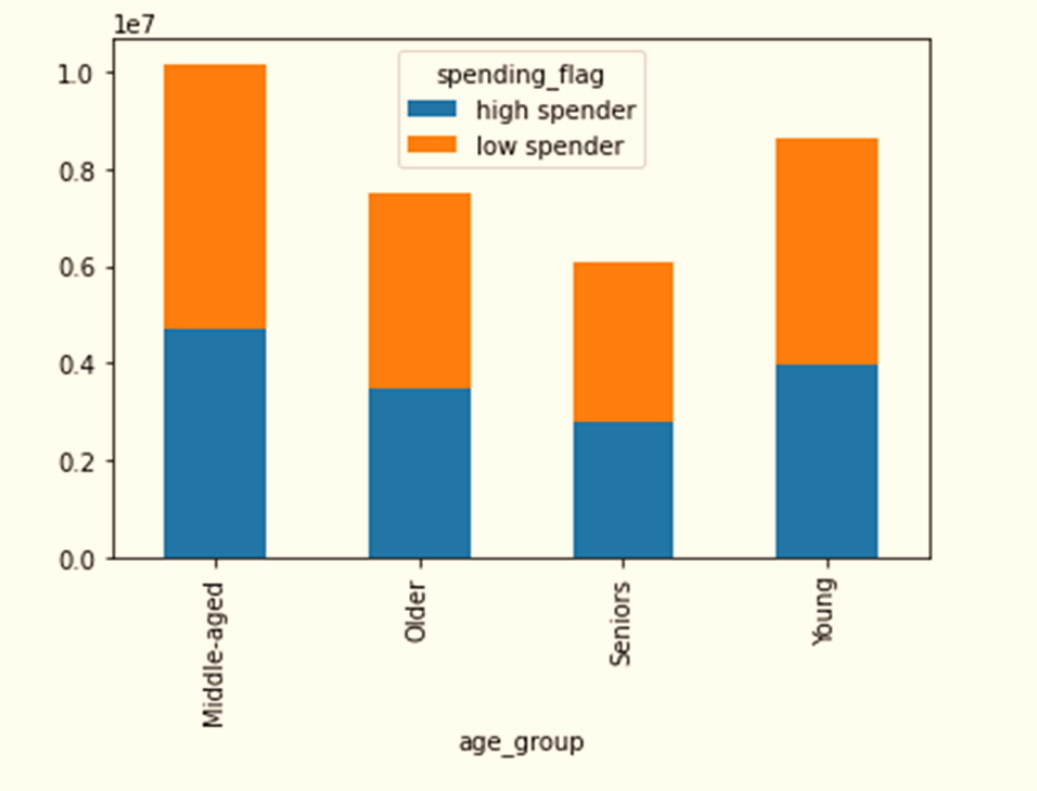
Answer: Frequent customers have the highest share in orders, Recurring the lowest

Recommendation:

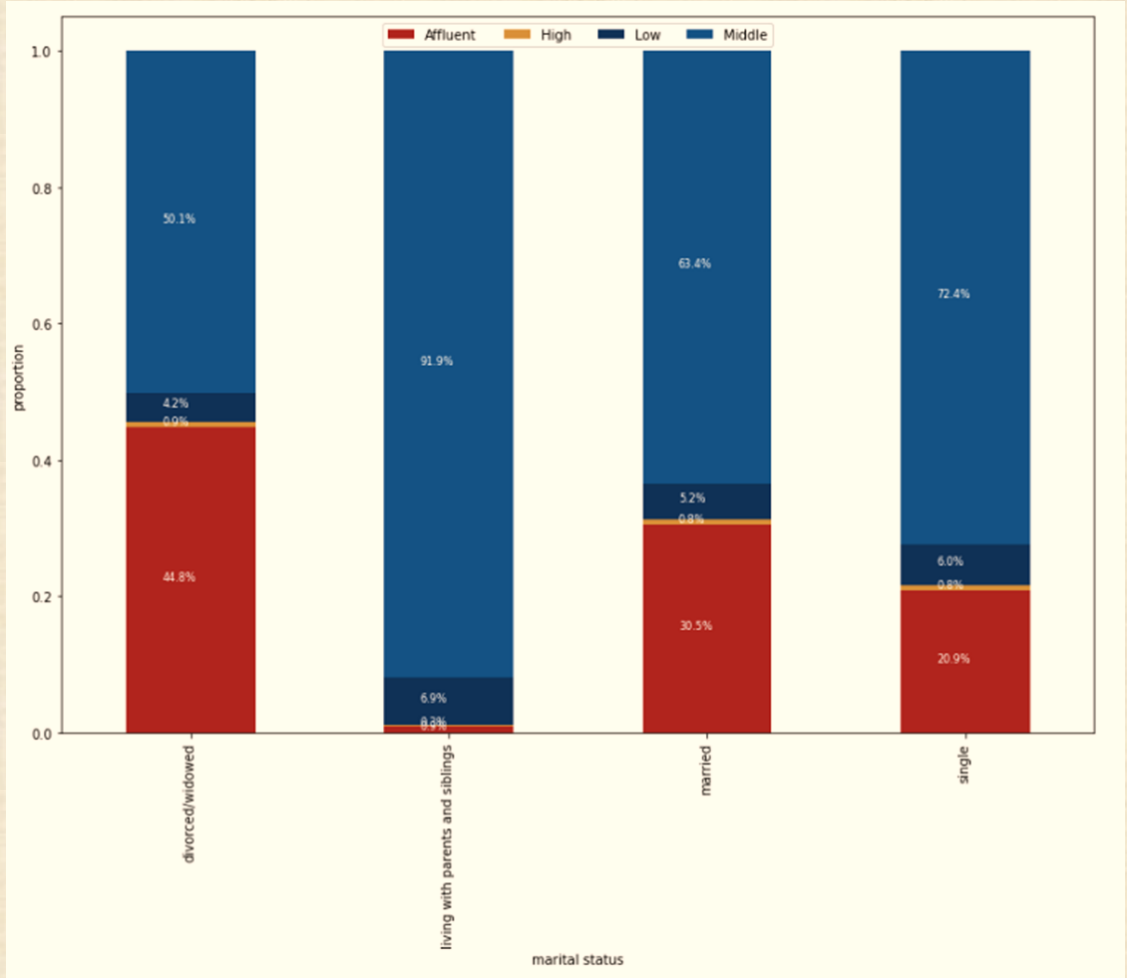
Targeting Recurring customers from West and South regions could boost sales.

Visualizations & Recommendations

Key Question 8: Is there a connection between age and family status in terms of ordering habits?



The bar chart shows spending by age group.
Young is up to 35, Middle-aged between 36 and 55, Older 56-70 and Seniors > 70
Answer: The middle-aged group spends slightly more than the other groups but generally ordering habits are similar

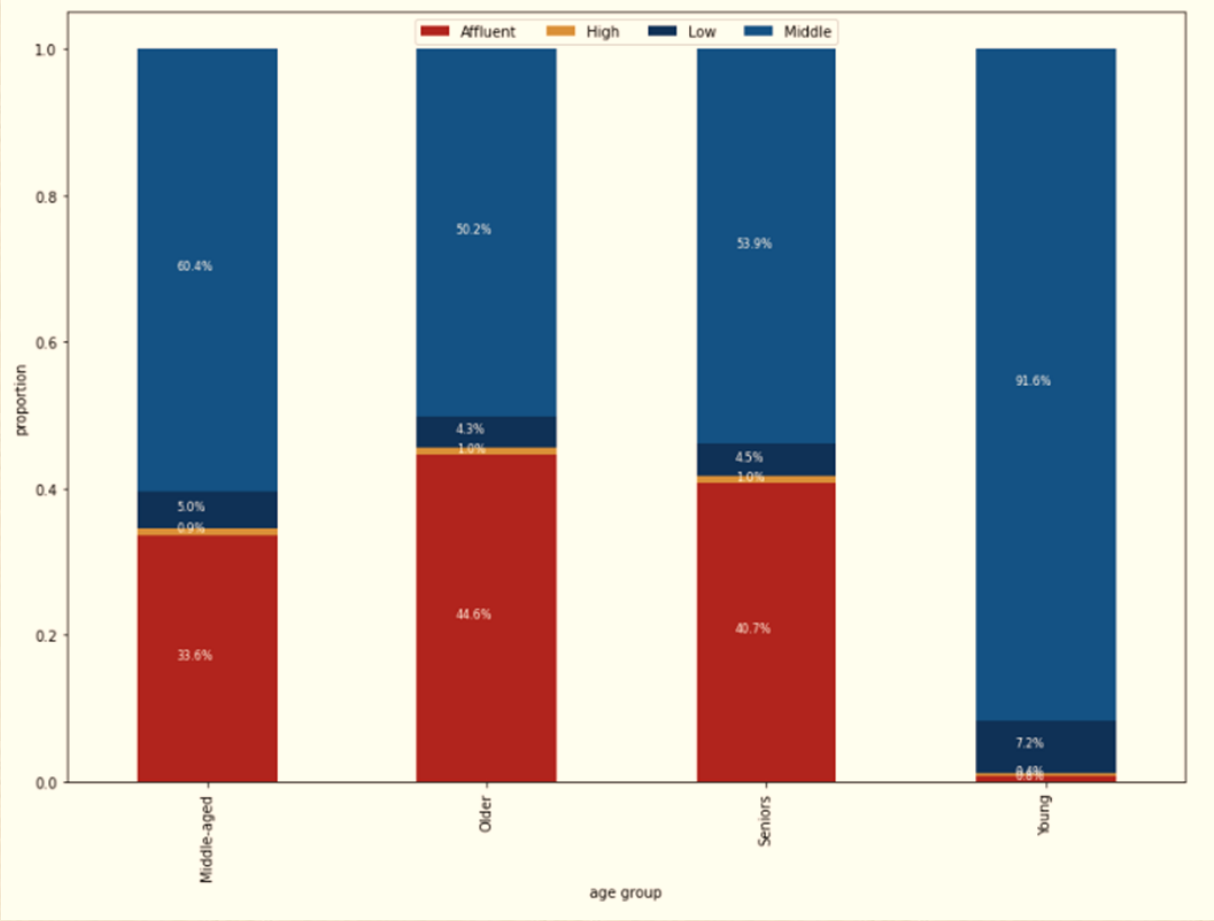


The bar chart shows income level by marital status.
Low income <$40K, Middle between $40K and $120K, Affluent between $120K and $200K, High income level >$200K
Answer: The highest proportion of Affluent customers in the divorced/widowed group, >90% the customers with the middle income level in the group 'living with parents and siblings'

Recommendations:
- Special promotions with different niche products to target each age group.
- Targeting divorced affluent and middle income level of the group 'living with parents and siblings' could increase sales with special promotions to increase $ per order.

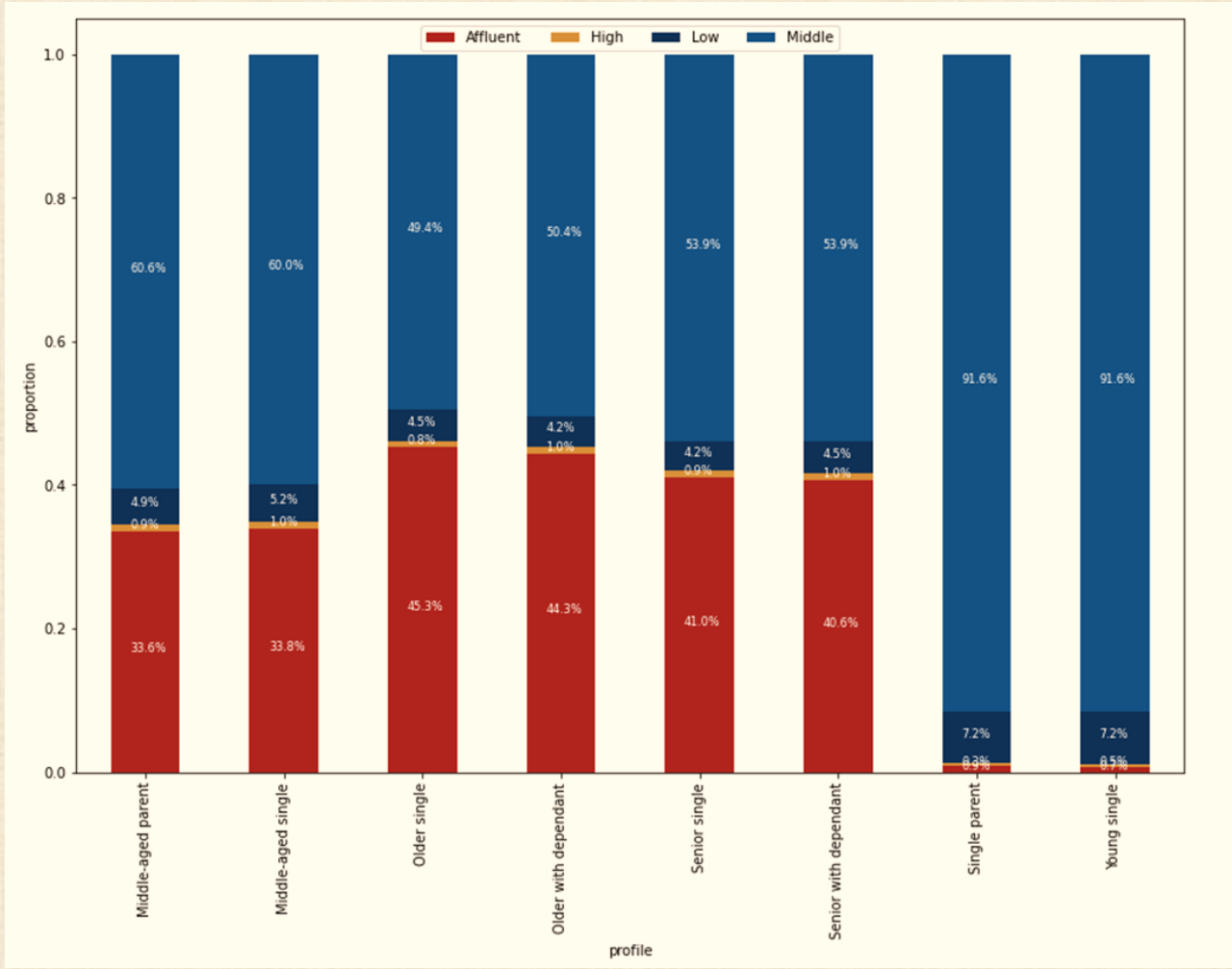
Visualizations & Recommendations

Key Question 9: What different classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?



The bar chart shows the proportion of income level category by age group.

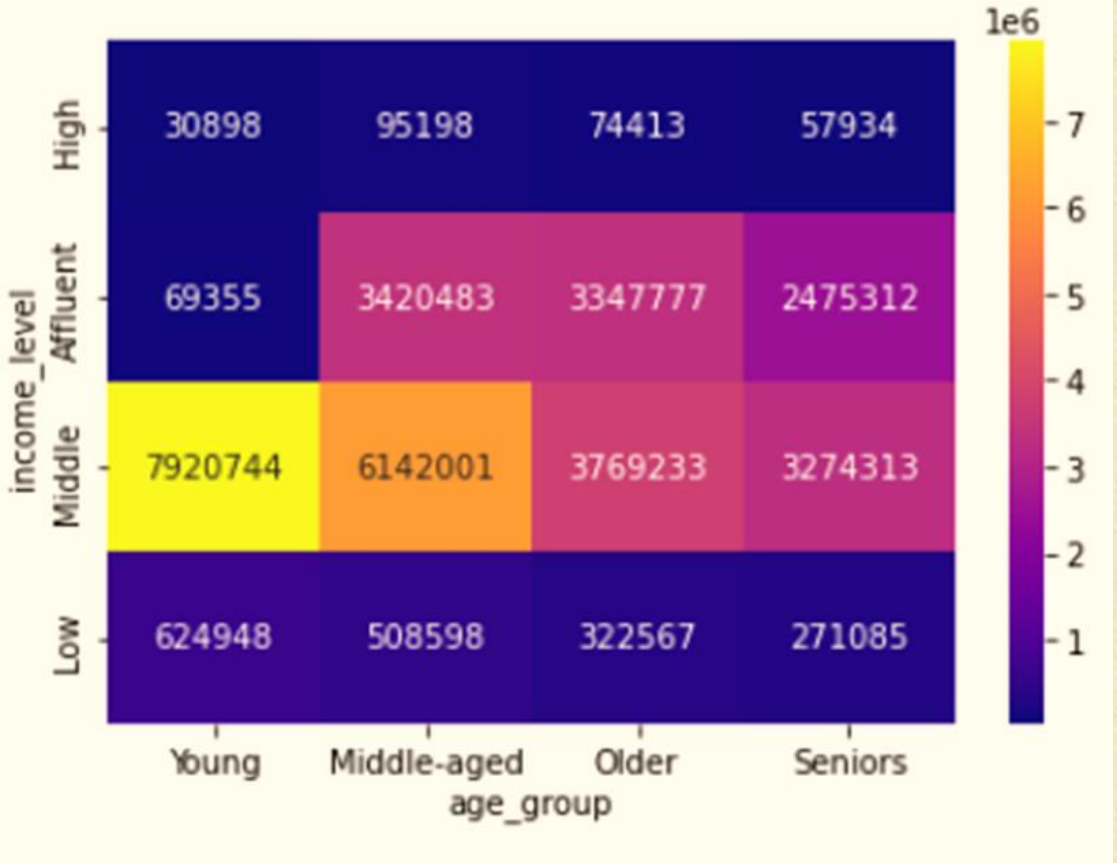
Answer: The highest share of Affluent customers are in Older and Seniors age group, but almost all Young people have Middle income level



The bar chart shows the proportion of income level category by customer profile.
Answer: The highest share of Affluent customers are in Older Single and Older with dependents profile, but 92% Young single and Single parents have Middle income level

Visualizations & Recommendations

Key Question 9: What different classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?



The heat map shows the number of orders made by customers by age group and income level

Young is up to 35, Middle-aged between 36 and 55, Older 56-70 and Seniors > 70

Low income <$40K, Middle between $40K and $120K, Affluent between $120K and $200K, High income level >$200K

Answer: The Middle and Affluent customers across all age groups make the major number of orders , Young and Middle-aged customers with the Middle income level are the core of the customer base.

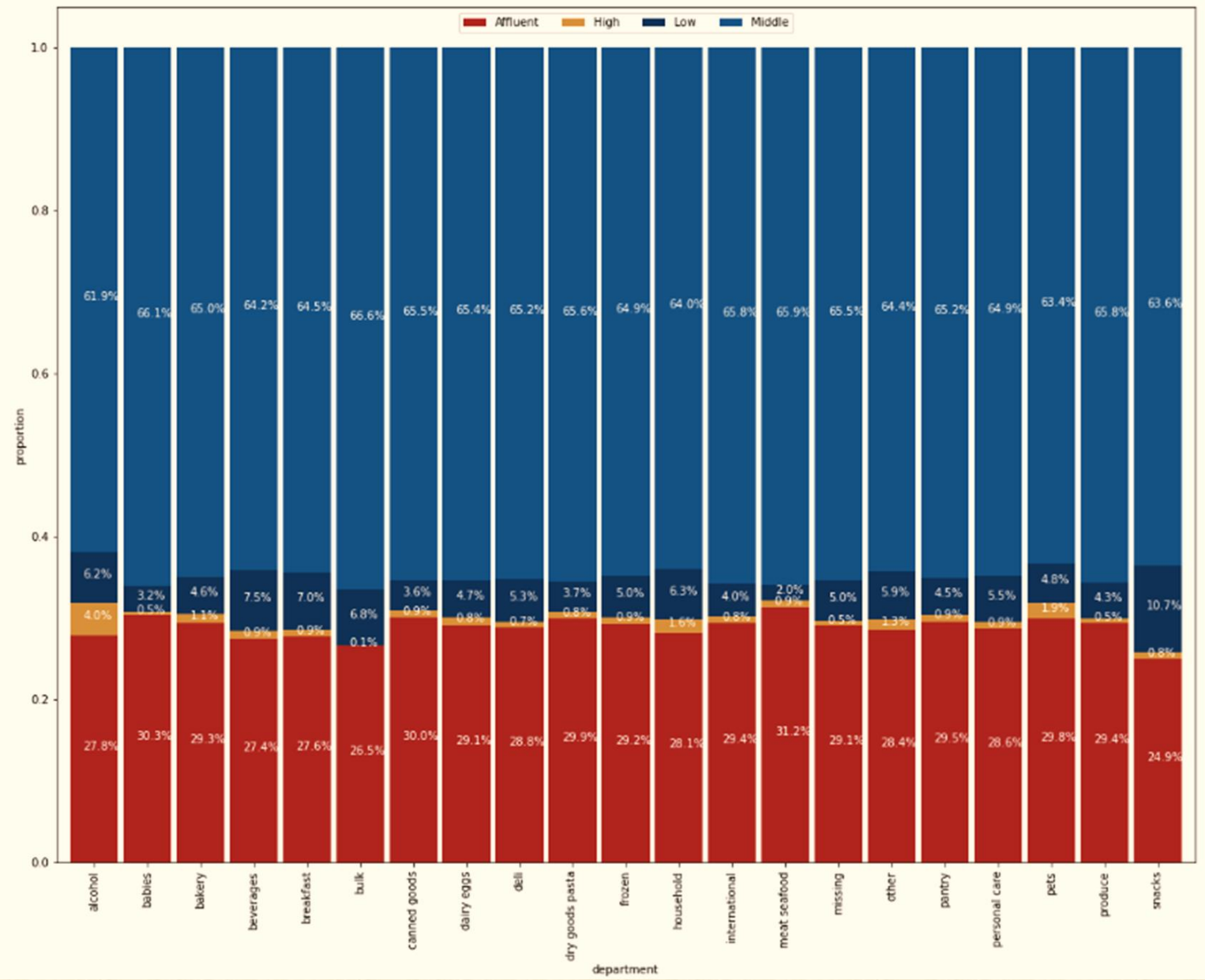
Recommendation:

To boost sales

- Focus on middle level income from customer profile: young single and single parent

- Affluent customers have more income to spend - focus on Older Single and Older with dependents to promote items for the customers and their dependents

Visualizations & Recommendations



Key Question 10: What differences can you find in ordering habits of different customer profiles? Consider the price of orders, the frequency of orders, the products customers are ordering, and anything else you can think of.

The bar chart shows proportion of orders from department by income level

Answer:

- Affluent customers prefer products from meat seafood and can goods.

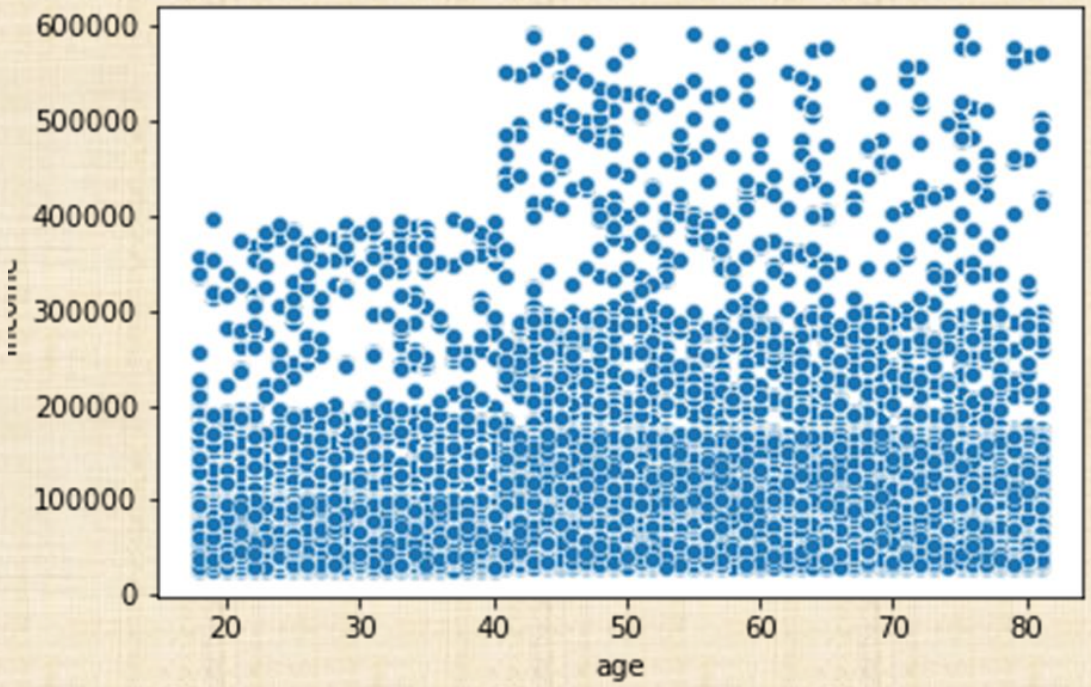
- High income level customers prefer products from alcohol and pets department.

- There is no clear preference for Middle income customers.

- Low income customers prefer products from snacks, beverages and breakfast departments.

Visualizations & Recommendations

Key Question 10: What differences can you find in ordering habits of different customer profiles? Consider the price of orders, the frequency of orders, the products customers are ordering, and anything else you can think of.



The chart shows income level by age, but's it's difficult to see percentages.

Answer: Customers aged up to 40 typically have less potential spending capability through income than those aged 40+

Recommendations:

- Target younger customers with low income level lower price products from snacks, beverages and breakfast departments

- Promote to the customers 40+ high range products from meat seafood, can goods, alcohol and pets departments due to they have more potential capability buy group of products