

INSTACART GROCERY ANALYSIS

Overview:

To help Instacart (an online grocery store) uncover more information about sales patterns through analysis of their customers and their purchasing behaviors.



Purpose and Context:

This is an achievement project for my Data Immersion course on Career Foundry focused on performing exploratory analysis to derive insights and suggest strategies for better segmentation.

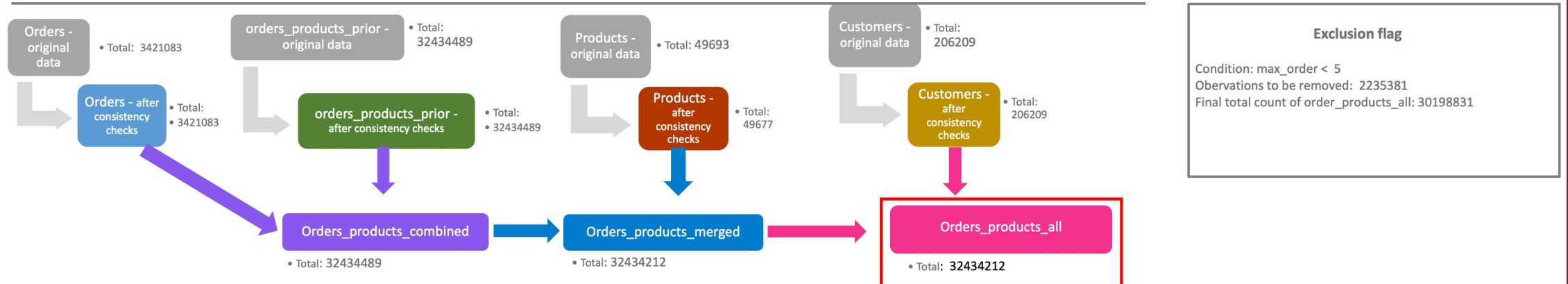
Objective:

To determine the best way to use a targeted marketing strategy to see if it has an effect on the sale of their products.



THE DATA:

Population flow



1. Started off by installing Anaconda/ the required Python libraries, created a notebook and imported the libraries.

2. Downloaded the data and imported it into notebook as a pandas dataframe.

3. Wrangled the data by changing data types of identifier variables and renaming columns to make more sense.

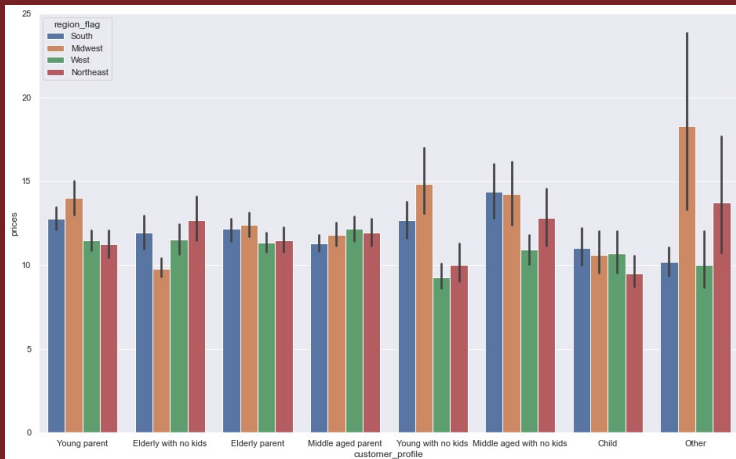
4. Accessed values and determined their meaning using a data dictionary.

5. Executed data consistency checks by fixing mixed type variables and dealing with missing values/duplicates.

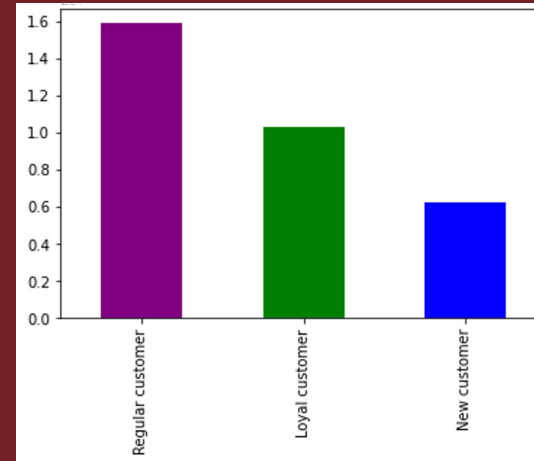
6. Concatenated, appended and merged data frames so I could export the data using the pickle format.

THE JOURNEY:

Derived the customer profile by age, marital status and number of dependents.

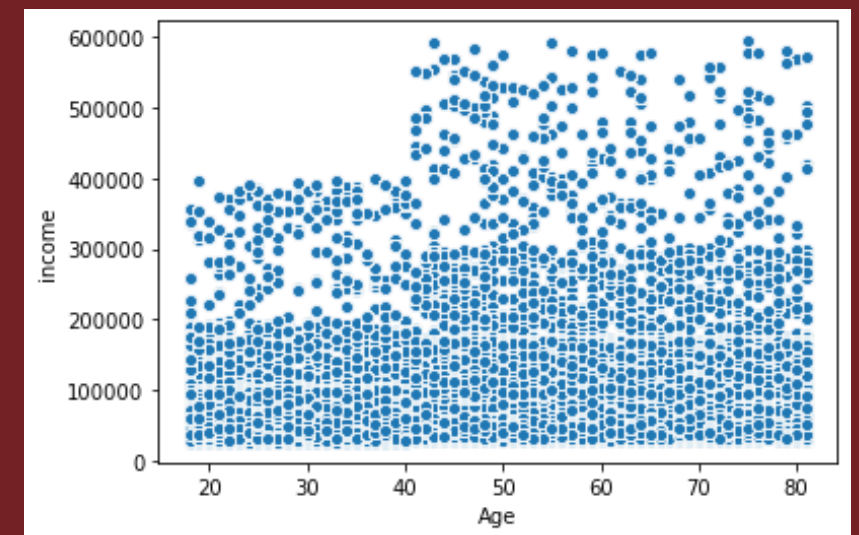


Created different flags such as the region flag and the spending flag to group customer information by region and by how much they spend.



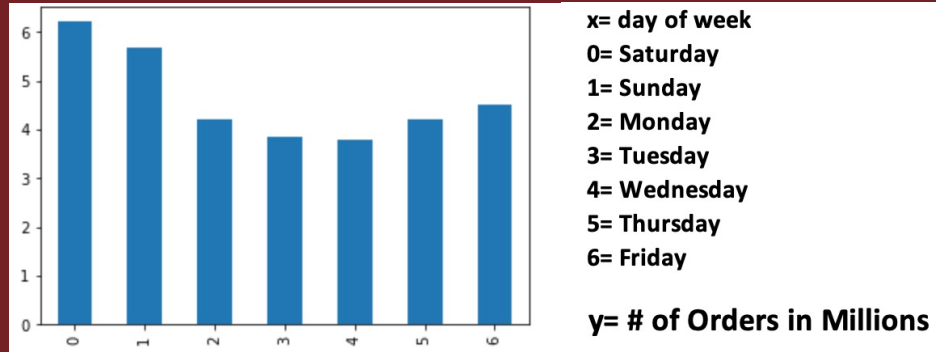
	median_days_since		
	mean	min	max
Loyalty_flag			
New customer	18.4598	0	30
Regular customer	10.6133	0	30
Loyal customer	5.17098	0	14

Utilized the loyalty flag to determine how much a new/regular/loyal customer spends and how often they shop.

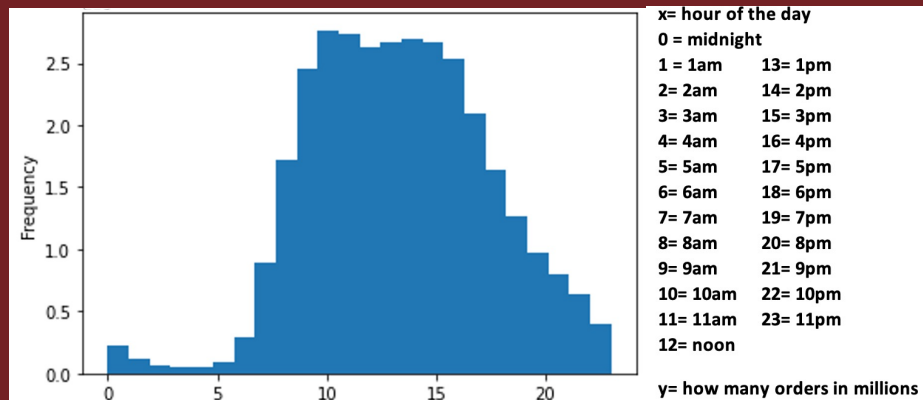


Compared the average income of customers by age and could see a big jump in average income around the age of 40.

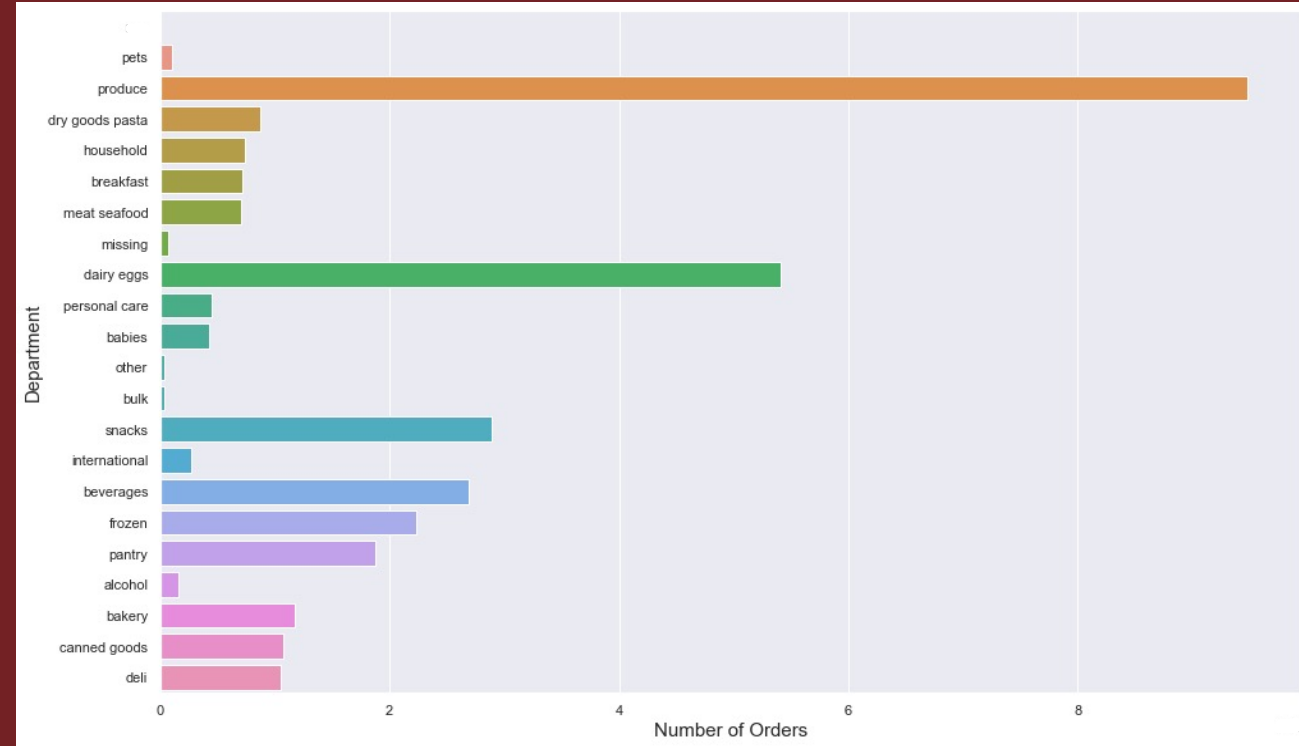
THE CONCLUSIONS:



Used the `groupby()` function to group the number of orders by day of the week as well as by hour of the day.



Determined Friday through Sunday to be the busiest days of the week and 9am to 4pm to be the busiest hours.



The most popular products in terms of number of items are produce and dairy/eggs (Over 9 million orders of produce and 5 million of dairy eggs).

THE LINKS

- Coding and the Charts it Created:

<https://github.com/SeanWard1979/Python-Code/blob/main/4-10%20Coding%20Etiquette%20%26%20Excel%20Reporting.ipynb>

- Instacart Grocery Basket Analysis Project:

<https://github.com/SeanWard1979/Python-Code/blob/main/Instacart%20Grocery%20Analysis-%20Sean%20Ward.xlsx>

Data sets:

- “The Instacart Online Grocery Shopping Dataset 2017”, Accessed from

<https://www.instacart.com/datasets/grocery-shopping-2017>