

INSTACART GROCERY BASKET ANALYSIS

Prepared by: Ambika Timilsena



Introduction



Overview:

A online grocery delivery store that operates its service through its app.

Objective:

The company already has good sales, they would like to uncover some insights regarding the behaviour of their customers spending and orders for better segmentation marketing strategy.

Role:

Data Analyst

a part of CareerFoundry Data Analytics Program

Primary Stakeholder:

CareerFoundry Data Analytics Course

Tools Used:

- ★ Python
- ★ Excel

Data Source

The dataset used for the analysis was from Instacart, containing over 30 million records. Using variables such as department ID, age, time of day of order, among other variables. The fictional customer data was from CareerFoundry.

Python Skills

- ★ Pandas
- ★ Seaborn
- ★ Matplotlib
- ★ Scipy
- ★ Numpy

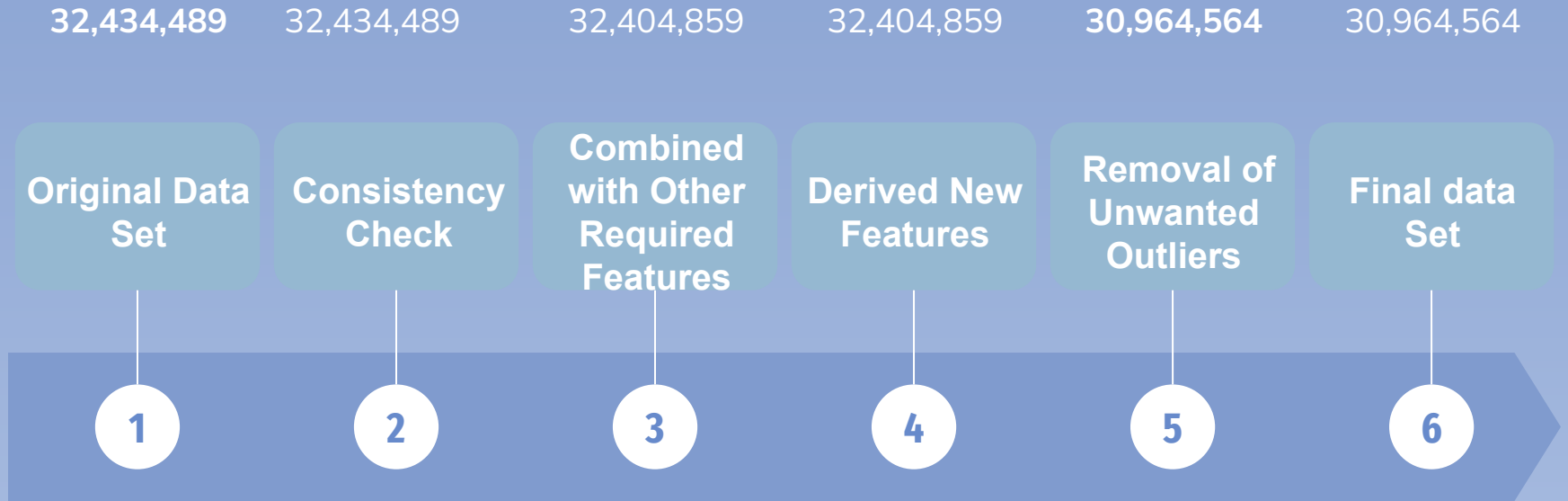


Key Questions to Explore

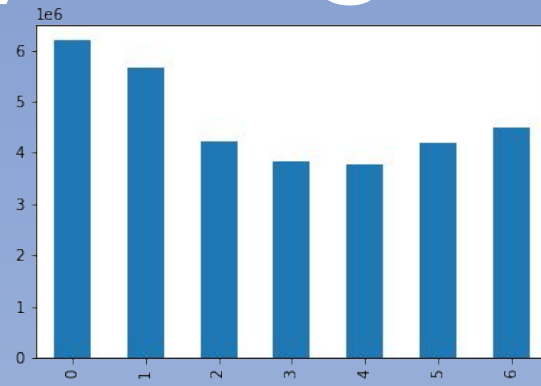
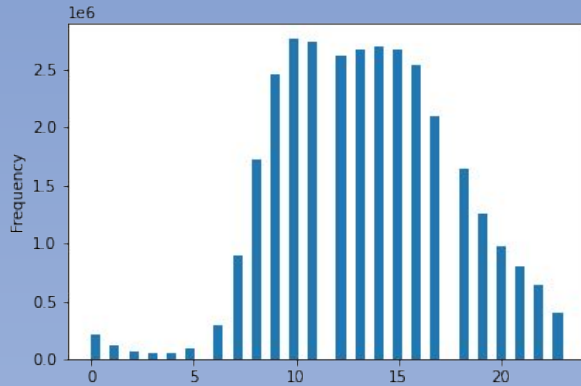
- ❖ What are the **busiest days** of the week and hours of the day?
- ❖ At what **time** do people spend the most money?
- ❖ Are there differences in ordering habits based on a customer's loyalty status?
- ❖ What purchasing differences are there between different customer profiles?



Data Preparation



Key Insights



9 AM
to
4 PM

Most Busy Hours

Most Busy Days

0: Saturday
1: Sunday

Weekends

Expensive Products

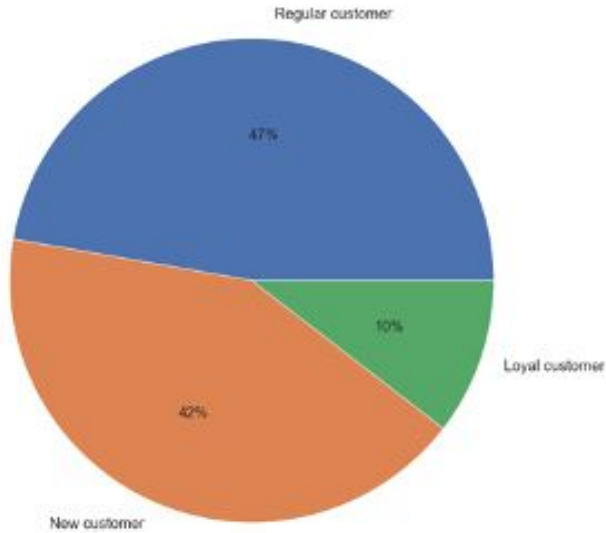
Purchased On

Weekends

Key Insights



Composition of Orders By Loyalty Flag

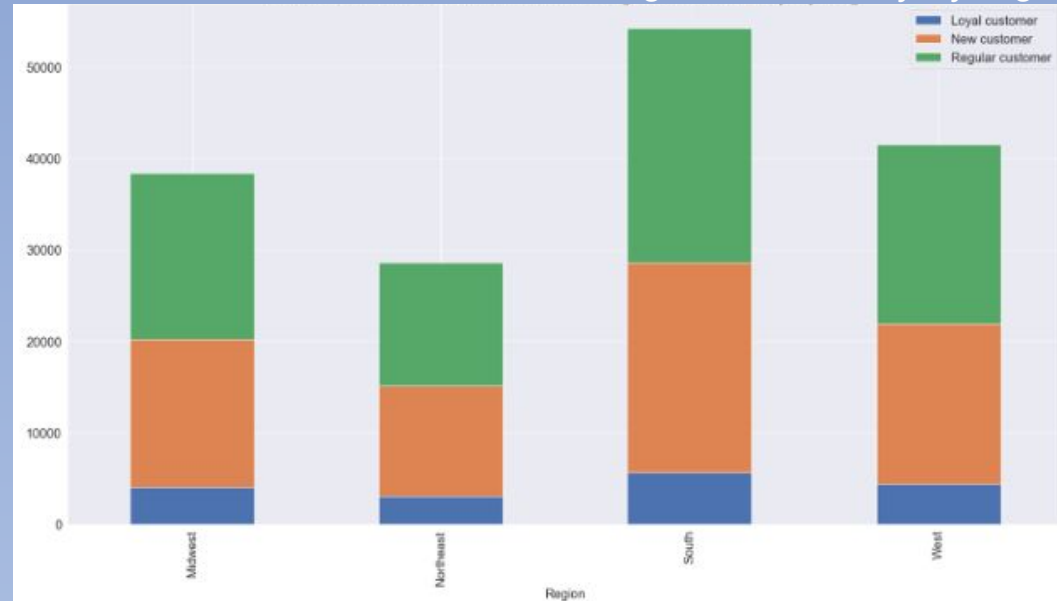


Loyal, Regular and New

South Region

Highest Customer

Distribution of Customers across US Regions based on Loyalty Flag



Regular Customers

Orders Made

47%

Challenges

The dataset contains over 30 million records which makes quite challenging for data cleaning, finding and treating missing values, formatting and standardizing the dataset for analysis.

Due to large data set, I often ran out of memory and system was quite slow.



Recommendations

- ✓ Ads should be scheduled from 4 pm till 9 am during weekdays when there are fewer orders.
- ✓ Customers should be targeted with ads based on their order history in order to encourage them to order more frequently and become loyal customers.
- ✓ As we have the highest composition of new customers and regular customers in ordering habits, they should be sent regular notifications to make orders in order to encourage them to continue ordering to become loyal customers.
- ✓ The top 5 products in term of highest frequency of products ordered are produce, dairy products, snacks, beverages and frozen. New ad campaign should also be targeted to the departments having lowest market sales like bulk, pets, alcohol.



Do You Have Any Questions?

Here are additional links with more recommendations and python code:



[Github Link](#)