**INSTACART MARKET BASKET ANALYSIS**

List of downloaded the dataset

* Orders.csv
* Peoducts.csv
* Departments.csv
* Aisles.csv
* Order\_product\_train.csv
* Order\_product\_perior.csv

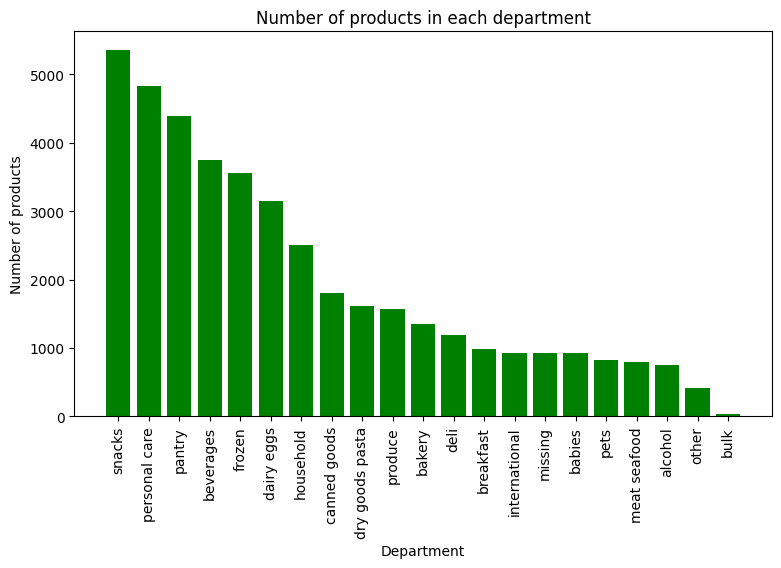
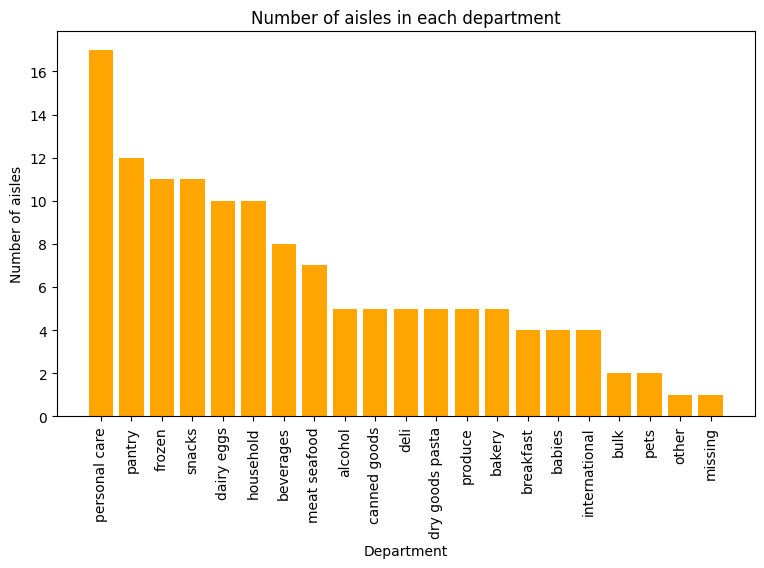
sample\_submission.csv file doesn’t contain any useful data. So I ignore that file.

**DATA PREPARATION**

* Imports libraries
* Upload datasets
* Identify any null values and replace with mean values.
* Merge all csv files into one file.

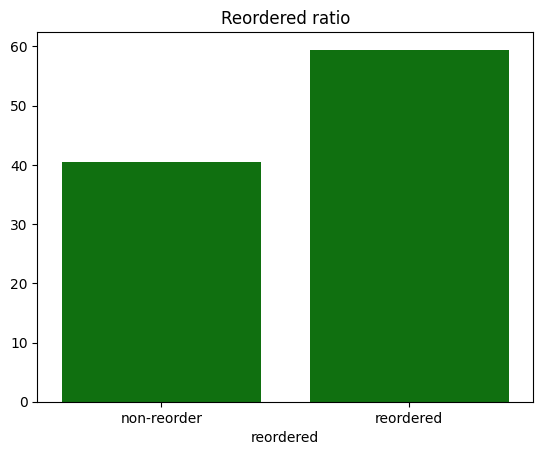
**EDA**

Number of aisles and products in each department:

****

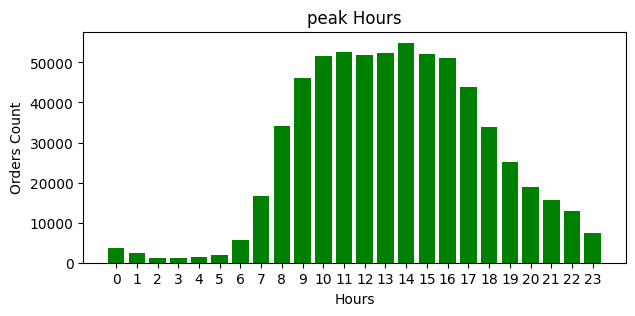
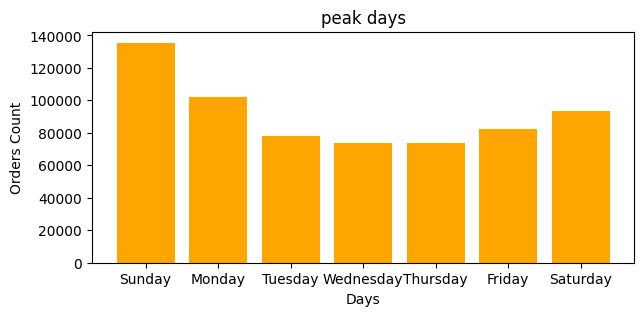
* There are total 21 departments.
* The snacks department has more products in it.
* In personal care department has more number aisle.

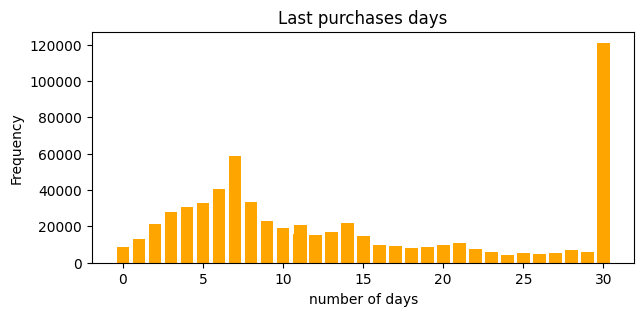
Reorder ratio:



* More than 59% of people reorded products.

Last purchases day and Peak day and hours:

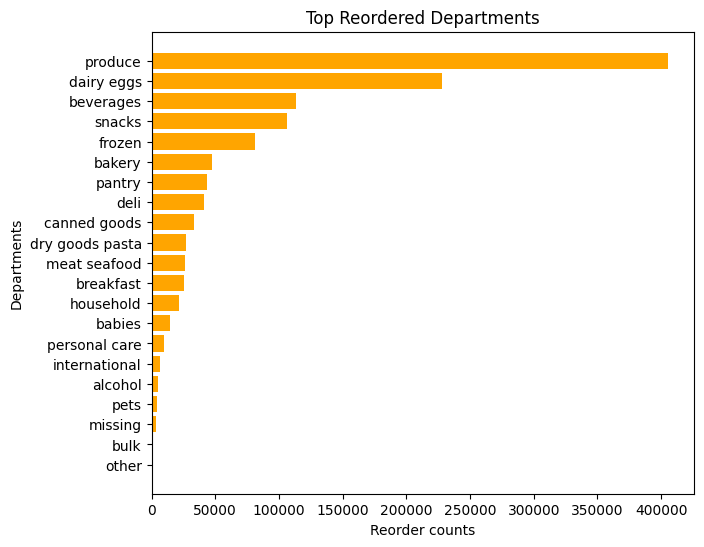
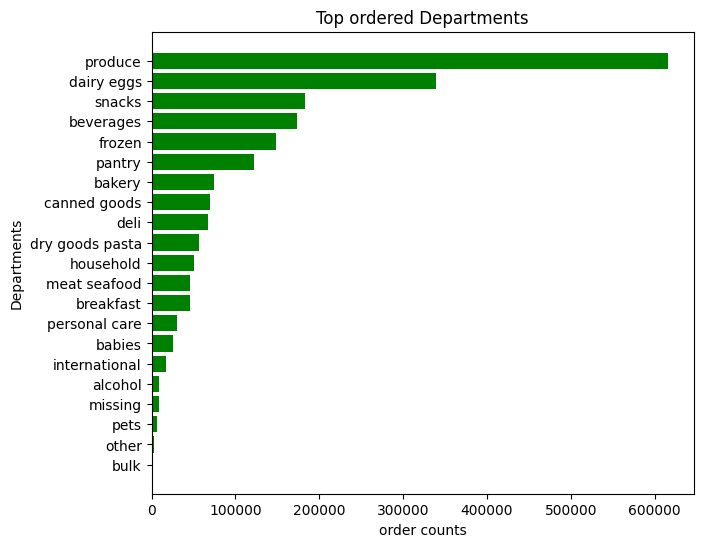
****

****

# **Insights**

* Sunday and Monday are the two days of peak orders of the week.
* 9am to 4pm is the peak hours of the days.
* Most of people order products in every 30 days.
* Some people also order products in weekly.

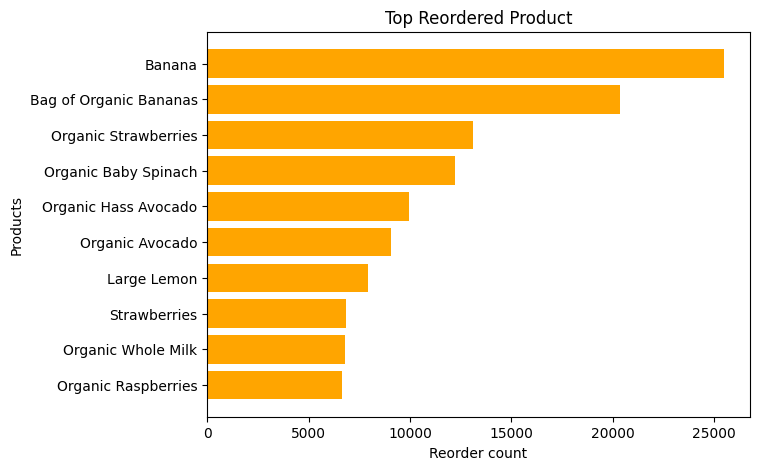
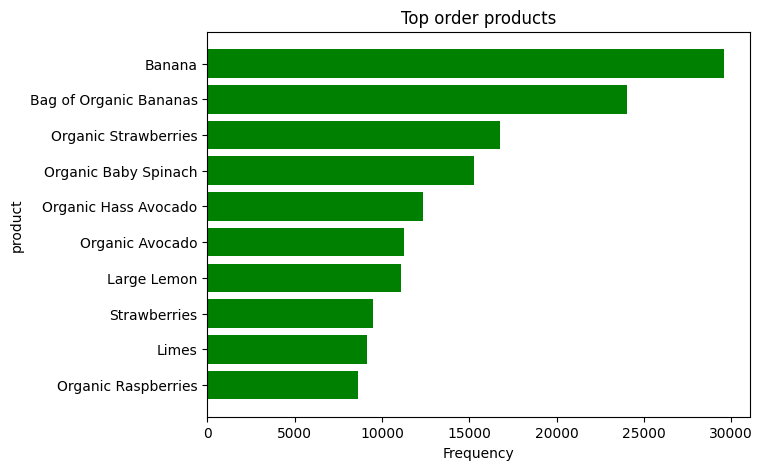
Departments:



# **Insights**

* Produce and dairy eggs took first and second place for both order and reorder departments.

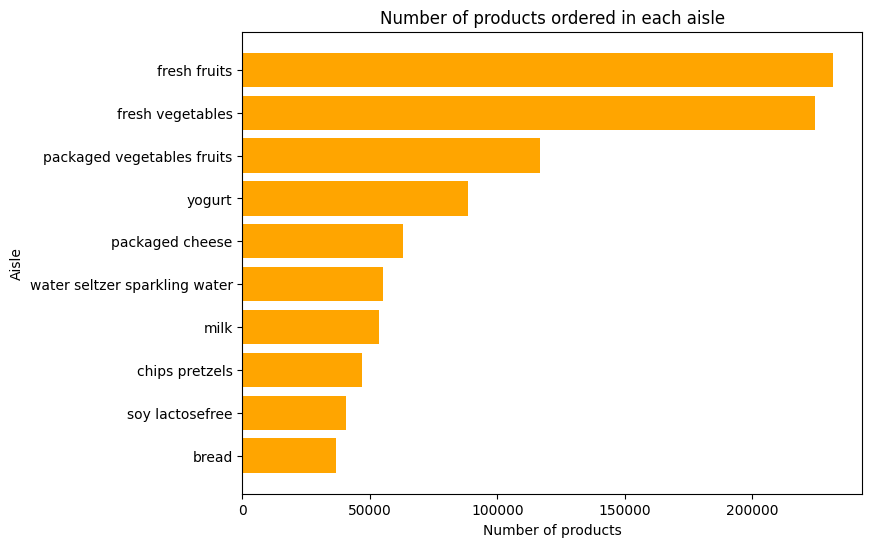
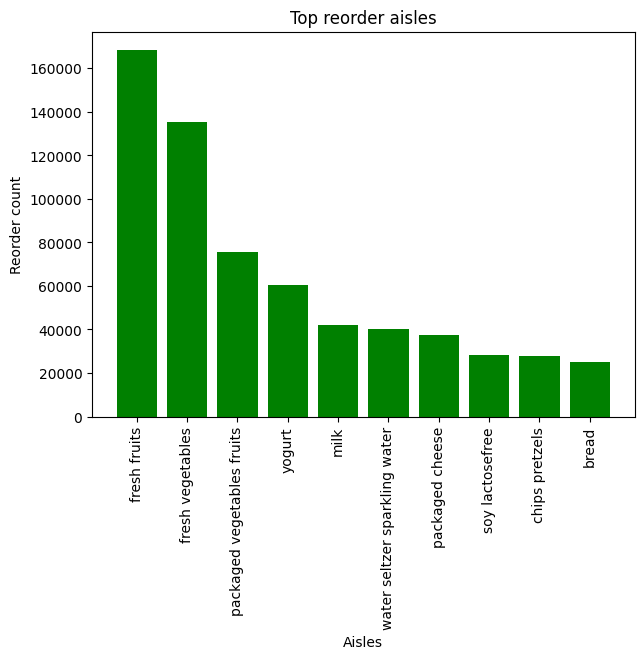
Products:



# **Insights**

* Banana and bag of organic bananas took first and second place for both order and reorder products.

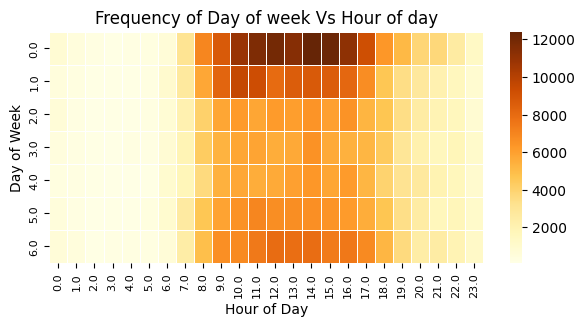
Aisles:



# **Insights**

* Fresh fruits and fresh vegetables took first and second place for both order and reorder products.
* These aisles are only produce department.

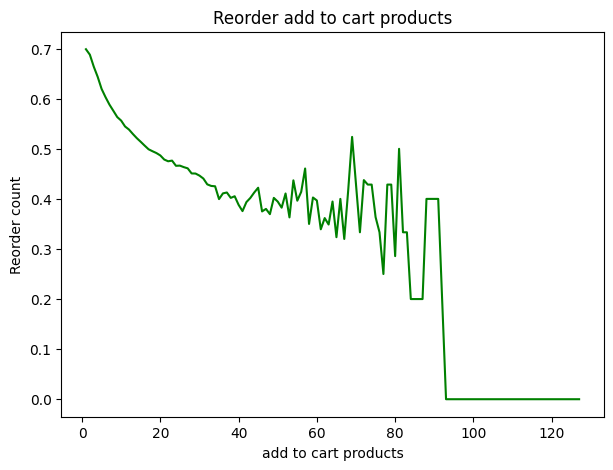
Peak order day and hours:



# **Insights**

* Sunday morning 8 to evening 5 o clock is the peak order peroid.
* Monday moring 9 to 11 only peak time after that orders counts start to decrease slightly.

Add to cart products:



Insights

* 70% of reorders are only for the first items in the add to cart.
* All items with an add to cart position below the 20th have a reorder rate below 40%.
* All 90th add to cart items have a 0% reorder rate.

### ****Recommendations****

* By placing personal care aisle around both Produce and dairy eggs two department aisles, can increase our sales.
* Wednesday and Thursday people are no interested to buy a products. So that time we can give some discounts, it helps increase our sales.
* Some people buy products in weekly. we target these customer, to give some special offer or voucher.
* Throught this we can retain our customers and our sales will increase.
* Push ads to customers based on fist 20 add to cart products.
* Because first 20 products only people recently add in cart, other after 20th products there is less chances of reorder.
* Sales will be increase through this way.

**Limitation**

* Dataset size was a problem to calculate rfm analysis.
* I used google colab notebook for analysis.
* Lot of time I have faced session crash issue, because of file size is too large.
* So I unable to encode it.

**Dashboard link:**

https://public.tableau.com/views/marketbasketanalysis\_17161310955880/Dashboard1?:language=en-GB&publish=yes&:sid=&:display\_count=n&:origin=viz\_share\_link

**Dashboard**

