



ANALYSIS REPORT

Assignment on Order data set



FEBRUARY 7, 2022
DATA SCIENCE LAB PVT.LTD.
Pune

Methodology

Our overall purpose in this analysis report is to analyze data and discover insights. There are 8 variables and 66 records in the "Order" data set that have no missing values.

Data set consists 10 customer's information related to order products history. The "Number of Delivery days" variable is calculated using the Excel command "INT," which is the difference between the order delivery date and the order date. Tableau Public and Python (Jupyter notebook) were used to do exploratory data analysis (EDA).

Analysis

The following links will take you to a graphical representation and a cross table:

Dashboard 1: [Click here](#)

Dashboard 2: [click here](#)

Python Work: [Click here](#)

Note: For a better visual experience, access the link on a laptop or computer.

Data types of each variable and missing values:

#	Column	Non-Null Count	Dtype
0	CustomerID	66 non-null	int64
1	OrderID	66 non-null	object
2	OrderDate	66 non-null	datetime64[ns]
3	OrderDeliveryDate	66 non-null	datetime64[ns]
4	CategoryCode	66 non-null	object
5	ProductCode	66 non-null	object
6	ProductQuantity	66 non-null	int64
7	PriceTotal	66 non-null	float64
8	Number_of_Delivery_days	66 non-null	int64

There are 66 records in all, none of which have a missing value (null value) for each column.

Summary Statistics:

	ProductQuantity	Number_of_Delivery_days	PriceTotal
count	66.000000	66.000000	66.000000
mean	1.454545	2.651515	141.593939
std	0.897549	2.222213	172.606655
min	1.000000	1.000000	21.800000
25%	1.000000	1.250000	47.825000
50%	1.000000	2.000000	94.850000
75%	2.000000	2.000000	146.825000
max	6.000000	8.000000	1032.000000

	OrderID	CategoryCode	ProductCode
count	66	66	66
unique	19	38	44
top	ONO114287752	S3-7421350	PID1531682
freq	24	5	4

In the provided data set, the following insights are found:

1. The loyal customers think highly of your brand, products, or services, which increases the chance that they will repeatedly buy from you.

Customer ID:202595 and 264981 are loyal customers in our database.

Customer ID:202595 purchased items in four different months.

Customer ID:264981 purchased items in two different months.

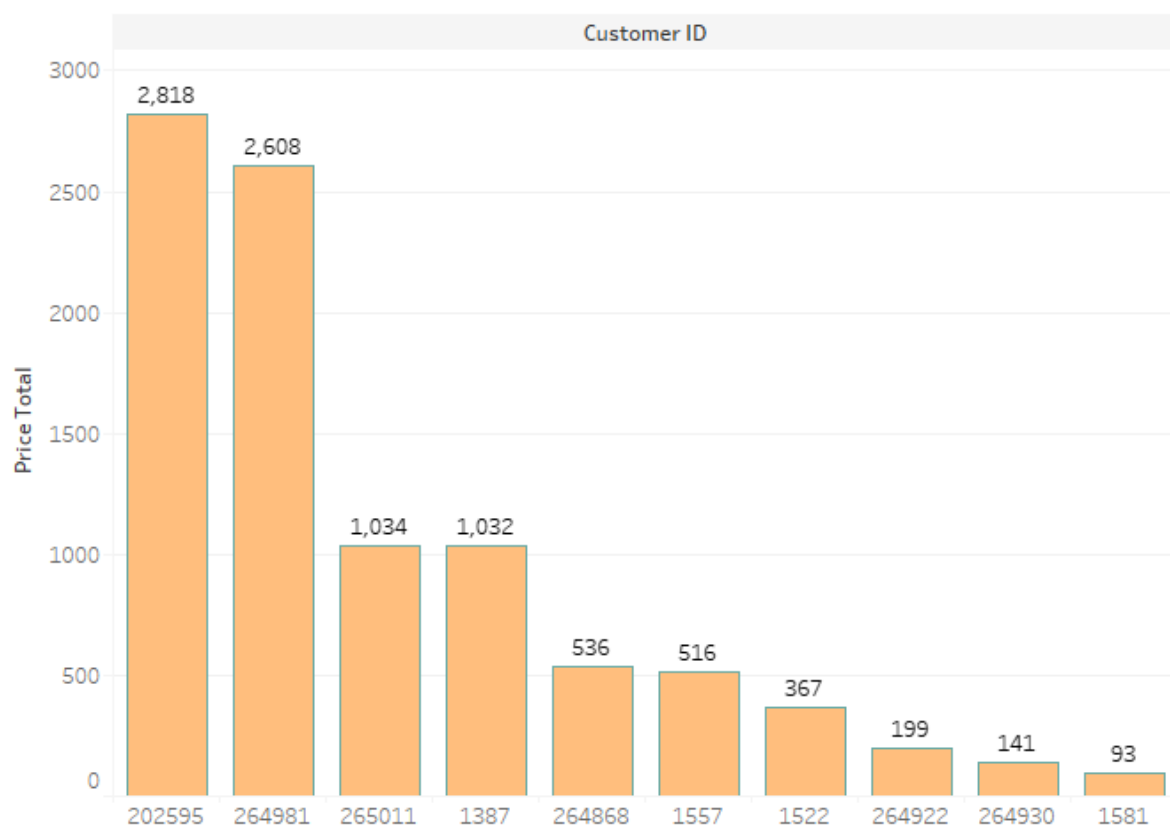
Above two customers out of ten have placed orders in different months, while the remaining eight have only placed in a month.

To attract customers, it would be advisable to provide an offer or provide a discount to that eight customers.

2. The overall amount of money spent by each consumer is shown in the table below.

Customer ID	
202595	2,818
264981	2,608
265011	1,034
1387	1,032
264868	536
1557	516
1522	367
264922	199
264930	141
1581	93

Customers' Expenditure



Customer ID: 202595 spent the most money, while customer ID: 1581 spent the least money.

3. There were 38 different category codes under which customers had placed their orders.

The majority of orders come from the following categorization codes:

Category code frequency

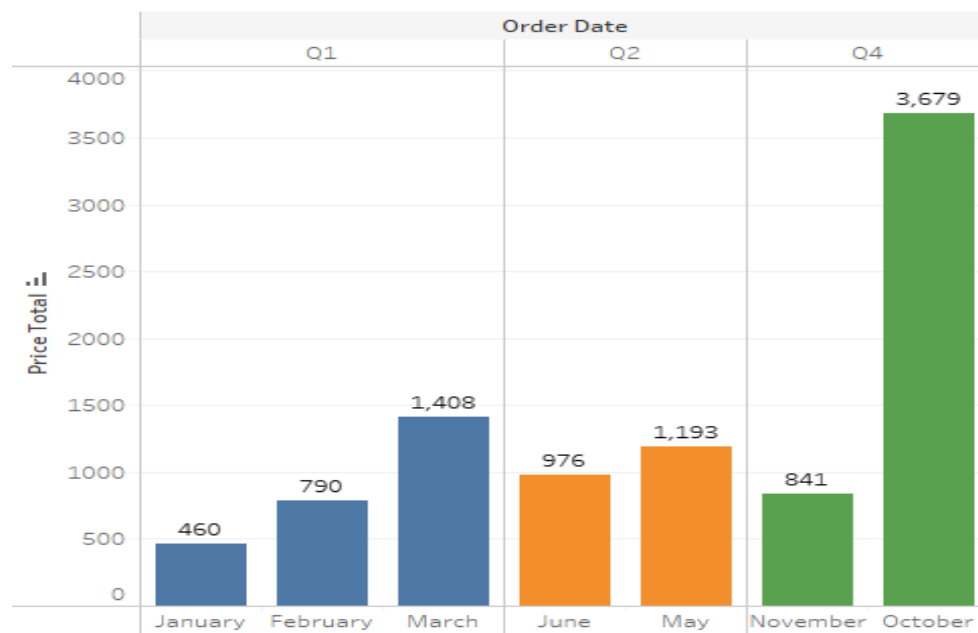
Category Code	
S3-7421350	5
S3-7861408	4
S3-7381341	4
S3-9261616	3
S3-9261613	3
S3-7861415	3
S3-7501359	3
W3-104519007710	2
S3-7701380	2
S3-7381345	2
S3-7381343	2
S3-7381342	2
S3-7261314	2
S3-7261313	2
S3-1760373	2
S3-1640353	2
S3-1600338	2

4. There were 44 different products which customers had placed their orders.

The following are the most frequently ordered products:

Product Code	
PID3134303	6
PID3134281	6
PID3134230	6
PID1255545	6
PID3110798	5

5. Revenue on a quarterly and monthly basis:

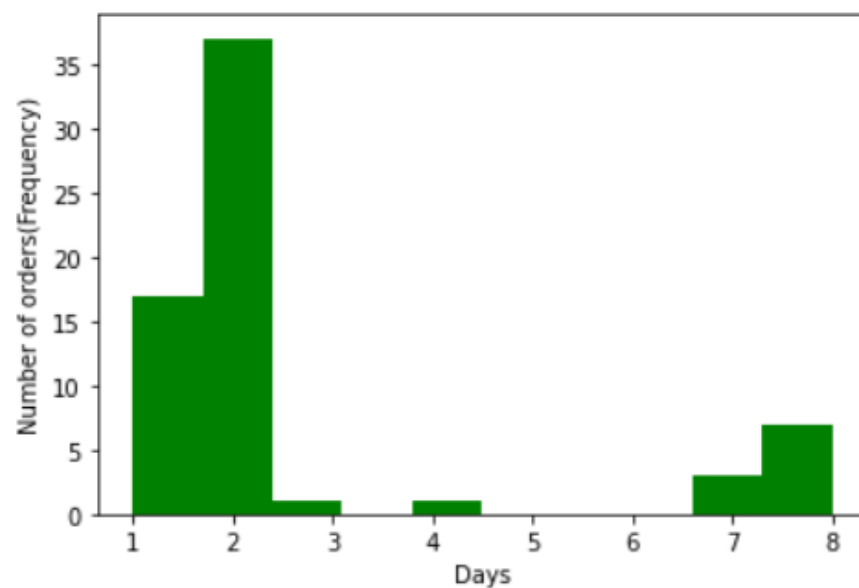


The above graph shows that there were no orders in the months of April, July, August, September, and October. As a result, quarter 3 will be absent of any orders.

In each quarter, total revenue from customers:

Quarter ..	
Q1	2,657
Q2	2,169
Q4	4,519

6. Delivery time:



Cumulative percentage table:

Days	Frequency	Cumulative Frequency	Cumulative Percentage
1	17	17	25.76%
2	37	54	81.82%
3	1	55	83.34%
4	1	56	84.85%
5	0	56	84.85%
6	0	56	84.85%
7	3	59	89.4%
8	7	66	100%

- It has happened seven times in the past for a product to take eight days to arrive.
- There is an 82 percent chance that the product will be delivered in two days.

In past, the following items took an additional eight days to arrive.

Category and product

Category Code	
S3-4540834	PID1284421
S3-7381341	PID3110798
S3-7421350	PID1531682
S3-7701380	PID1124319
S3-7861408	PID3134281
S3-9261613	PID2652785
S3-9261616	PID1338645