

# Milestone Report

**Problem statement** :- Maintaining the stock of the products has always been a challenge of every seller. Products should not be over-stocked as if you have an influx of product, you run the risk of not being able to sell all of it effectively nor it should be under-stocked where you run the risk of losing sales due to not having anything to sell.

## **About Dataset :**

Dataset used in the project is in the below mentioned link

<https://github.com/shivamkk32/CP-1>

The dataset is distributed in multiple csv file that are merged into one using inner join based on either product id or order id.

After merging the different csv to one the final csv file named "allProduct.csv" is used further in the project.

Github link :-

<https://github.com/shivamkk32/CP-1/blob/master/EDA-CP1.ipynb>

Major information used from the dataset to make sales forecasting are

1. The time at which the product was sold. Here in this project we are doing week-wise analysis to predict the sales for upcoming weeks. Using the time-series data the week information is extracted in which the product was sold.
2. Count of each product sold per week.
3. Actual selling price of the product.
4. Actual amount to be paid by the customer.

5. Percentage discount given on the product.
6. Average days to deliver the product.
7. Last 6 week sales trend of the product.

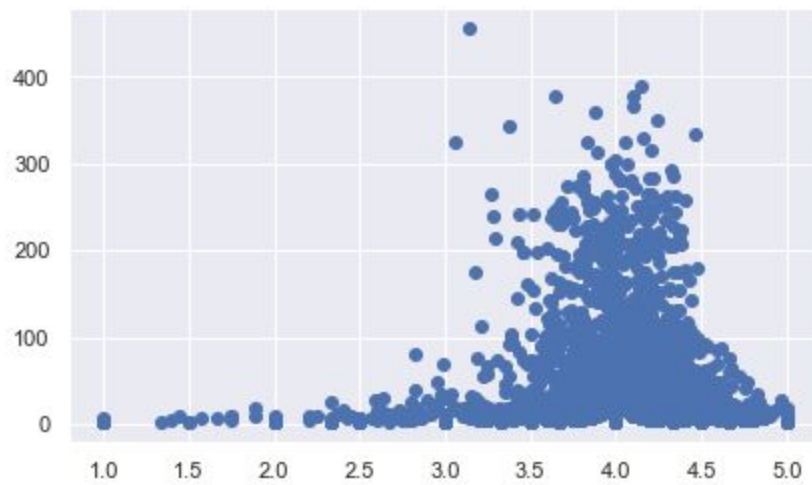
Sample :

Week	product_category_name_english	encodedCategory	count_perCategory	paymentValue	final_AmountPaid	Review_score	DiscountPercent	SaleRecord
52	office_furniture	57	4	299.237500	299.237500	3.000000	0.000000e+00	17.833333
52	perfumery	59	35	134.550571	132.842000	3.914286	1.269836e+00	83.666667
52	pet_shop	60	13	115.799231	115.799231	4.692308	0.000000e+00	24.166667
52	signaling_and_security	62	1	719.150000	719.150000	5.000000	0.000000e+00	0.666667
52	small_appliances	63	2	297.710000	297.710000	5.000000	1.909355e-14	9.333333
52	sports_leisure	65	60	142.936667	134.125667	4.033333	6.164269e+00	152.666667
52	stationery	66	54	107.755000	115.309815	4.092593	0.000000e+00	53.500000
52	telephony	68	58	55.685862	50.249828	4.379310	9.761965e+00	85.666667
52	toys	69	37	191.202703	186.727568	4.027027	2.340519e+00	141.333333
52	watches_gifts	70	54	202.143519	193.216111	3.888889	4.416371e+00	108.333333

Findings:

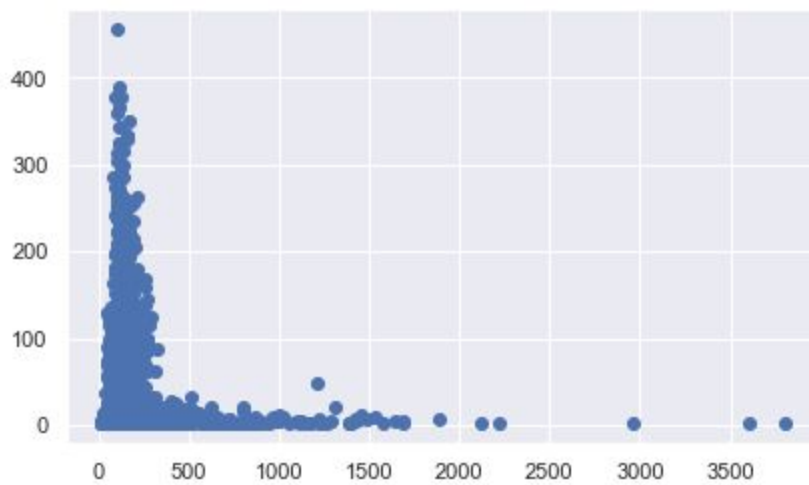
- 1.The review score of the items has positive correlation with the sales.

1. Price has a positive correlation with sales.



2. Price has a negative correlation with sales.

2. Price has a negative correlation with sales.



3. The discount given of the products are normally distributed.

