

PROBLEM STATEMENT

An automobile parts manufacturing company has collected data of transactions for 3 years. They do not have any in-house data science team; thus, they have hired you as their consultant. Your job is to use your magical data science skills to provide them with suitable insights about their data and their customers.

ORDERNUMBER:	Order Number	CUSTOMER NAME:	Customer
QUANTITYORDERED:	Quantity ordered	PHONE:	Phone of the customer
PRICEEACH:	Price of Each item	ADDRESSLINE1:	Address of customer
ORDERLINENUMBER:	order line	CITY:	City of customer
SALES:	Sales amount	POSTALCODE:	Postal Code of customer
ORDERDATE:	Order Date	COUNTRY:	Country customer
DAYS_SINCE_LASTORDER:	Days_Since_Lastorder	CONTACTLASTNAME:	Contact person customer
STATUS:	Status of order like Shipped or not	CONTACTFIRSTNAME:	Contact person customer
PRODUCTLINE:	Product line – CATEGORY	DEALSIZE:	Size of the deal based on Quantity and Item Price
MSRP:	Manufacturer's Suggested Retail Price	PRODUCTCODE:	Code of Product

DATA SUMMARY

- An automobile parts manufacturing company has collected data of transactions for 3 years.
- The data has 2747 entries (0 To 2746) of rows and 20 columns. The data has 1 datetime64 , 2 float64, 5 int64, and 12 Object data types.
- The dataset contains customer geographical information and transaction history.

The Sales_Data has 2747 rows and 20 columns

The size of Sales_Data 54940

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2747 entries, 0 to 2746
Data columns (total 20 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   ORDERNUMBER                          2747 non-null   int64
1   QUANTITYORDERED                      2747 non-null   int64
2   PRICEEACH                           2747 non-null   float64
3   ORDERLINENUMBER                      2747 non-null   int64
4   SALES                                2747 non-null   float64
5   ORDERDATE                           2747 non-null   datetime64[ns]
6   DAYS_SINCE_LASTORDER                 2747 non-null   int64
7   STATUS                               2747 non-null   object
8   PRODUCTLINE                          2747 non-null   object
9   MSRP                                 2747 non-null   int64
10  PRODUCTCODE                          2747 non-null   object
11  CUSTOMERNAME                         2747 non-null   object
12  PHONE                                2747 non-null   object
13  ADDRESSLINE1                         2747 non-null   object
14  CITY                                 2747 non-null   object
15  POSTALCODE                           2747 non-null   object
16  COUNTRY                              2747 non-null   object
17  CONTACTLASTNAME                      2747 non-null   object
18  CONTACTFIRSTNAME                     2747 non-null   object
19  DEALSIZE                             2747 non-null   object
dtypes: datetime64[ns](1), float64(2), int64(5), object(12)
memory usage: 429.3+ KB
```

DESCRIPTION OF THE DATA

- The company is into automobile part manufacture, and they have different product line like Classic car , Motorcycle, plane, train, ship, Bus truck, vintage cars etc.
- The data maintained each transactions entry as order number.
- Manufacturer's Suggested Retail Price(MSRP) for each product code is decided but we found that this is not matching with Price of Each item & is inconsistent with MSRP.

	count	unique	top	freq	first	last	mean	std	min	25%	50%	75%	max
ORDERNUMBER	2747	NaN	NaN	NaN	NaT	NaT	10259.8	91.8775	10100	10181	10264	10334.5	10425
QUANTITYORDERED	2747	NaN	NaN	NaN	NaT	NaT	35.103	9.76214	6	27	35	43	97
PRICEEACH	2747	NaN	NaN	NaN	NaT	NaT	101.099	42.0425	26.88	68.745	95.55	127.1	252.87
ORDERLINENUMBER	2747	NaN	NaN	NaN	NaT	NaT	6.49108	4.23054	1	3	6	9	18
SALES	2747	NaN	NaN	NaN	NaT	NaT	3553.05	1838.95	482.13	2204.35	3184.8	4503.09	14082.8
ORDERDATE	2747	246	2018-11-14 00:00:00	38	2018-01-06	2020-05-31	NaN	NaN	NaN	NaN	NaN	NaN	NaN
DAYS_SINCE_LASTORDER	2747	NaN	NaN	NaN	NaT	NaT	1757.09	819.281	42	1077	1761	2436.5	3562
STATUS	2747	6	Shipped	2541	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
PRODUCTLINE	2747	7	Classic Cars	949	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
MSRP	2747	NaN	NaN	NaN	NaT	NaT	100.692	40.1148	33	68	99	124	214
PRODUCTCODE	2747	109	S18_3232	51	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CUSTOMERNAME	2747	89	Euro Shopping Channel	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
PHONE	2747	88	(91) 555 94 44	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
ADDRESSLINE1	2747	89	C/ Moralzarzal, 86	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CITY	2747	71	Madrid	304	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
POSTALCODE	2747	73	28034	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
COUNTRY	2747	19	USA	928	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CONTACTLASTNAME	2747	76	Freyre	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
CONTACTFIRSTNAME	2747	72	Diego	259	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN
DEALSIZE	2747	3	Medium	1349	NaT	NaT	NaN	NaN	NaN	NaN	NaN	NaN	NaN

- No missing values in the dataset

ORDERNUMBER	0
QUANTITYORDERED	0
PRICEEACH	0
ORDERLINENUMBER	0
SALES	0
ORDERDATE	0
DAYS_SINCE_LASTORDER	0
STATUS	0
PRODUCTLINE	0
MSRP	0
PRODUCTCODE	0
CUSTOMERNAME	0
PHONE	0
ADDRESSLINE1	0
CITY	0
POSTALCODE	0
COUNTRY	0
CONTACTLASTNAME	0
CONTACTFIRSTNAME	0
DEALSIZE	0
dtype:	int64