

Problem Statement:

This dataset contains year-wise data of sales Year-wise sales analysis quarter-wise sales analysis monthly-wise sales analysis Find key metrics and factors and show the meaningful relationships between attributes. Do your own research and come up with your findings

Data set:

[illegible]

```
In [6]: 1 pd.read_excel('CUSTOMERS.xls')
```

Out[6]:

	Address Number	Business Family	Business Unit	Customer	Customer Number	Customer Type	Division	Line of Business	Phone	Region Code	Regional Sales Mgr	Search Type	Unnamed: 12
0	10000500	R3	1	City Supermarket	10000000	G2	2		816-455-8733	4	S16	C	NaN
1	10000463	R3	1	A Supermarket	10000463	G1	1		816-455-8733	5	S19	C	NaN
2	10000455	R3	1	Caribbean Supermarket	10000455	G2	2		816-455-8733	1	S16	C	NaN
3	10000456	R1	1	A&B Shop	10000456	G3	1		816-455-8733	0	S2	C	NaN
4	10000457	G2	1	A&O Shop	10000457	G1	1		816-455-8733	5	S1	C	division
...													
879	10027560	R2	1	Zlog Shop	10027560	G2	2		816-455-8733	2	S1	C	NaN
880	10027572	R3	1	ZipUp Com Shop	10027572	G2	2		816-455-8733	3	S1	C	NaN
881	10027575	R3	1	Zitel Shop	10027575	G2	2		816-455-8733	2	S1	C	NaN
882	10027583	R2	1	zNET Shop	10027583	G2	2		816-455-8733	4	S9	C	NaN

```
In [5]: 1 pd.read_excel('CUSTOMERADDRESS.xls')
```

Out[5]:

	Address Number	City	Country	Customer Address 1	Customer Address 2	Customer Address 3	Customer Address 4	State	Zip Code
0	1000471	Jeddah	AU	Al Thalia Street	PO Box 11605			NaN	11443
1	10004255	Seoul	AU	3F Shinwoo Building	195-12 Poi-Dong Kangnam-KU			NaN	135-260
2	10007117	Taipei	AU	8F No.19 Chang Chun Road				NaN	104
3	10010923	Seocho-GU	AU	231 Yangjae Dong				NaN	137-938
4	10013011	112 Reyjavik	AU	Vagnhofda 23				NaN	
...									
450	10027560	Odessa	US	3356 Kermit Highway				TX	79764
481	10027572	Elma	US	2210 Bowen Road				NY	14059
452	10027575	Dallas	US	10400 Plano Road				TX	75238
453	10027583	Morton	US	Attention: Charlene Hoyer	500 North Morton Avenue	PO Box 474		IL	61550-0474
454	10027629	Dearborn	US	Central Accounting Service	PO Box 6005			MI	48121

455 rows x 9 columns

The dataset consists of 3 excel sheets . Description of data: File1: Sales Data ----- it contains sales of all the items and sales amount File2: Customer Address----- it contains address of the customers File3: Customers----- it contains information of the customers.

Here ,

- 1) Cust-key is primary key and address number is foreign key.
- 2) Date key = invoice date = promised delivery date .
- 3) Discount Amount = Sales Amount Based on List Price – Sales Amount
- 4) item ---- names of items
- 5) List Price --- sum of price of all items in list based on its MRP

6) Sales Amount---- it is the price given by the customer

7) Sales Amount Based on List Price= List Price * Sales Quantity

8) Sales Cost Amount---- it is an actual price paid by amazon to buy particular item.

9) Sales Margin Amount----Sales Amount- Sales Cost Amount

10) Sales Price ----- price of one item

11) Sales Quantity---- number of items bought



SALES ANALYSIS

186.19M

Sales Amount

77.75M

Sales Margin

2

Country

658

Items

65.28K

Customers

sales by year

2.78K✓

Goal: 1.34K (+106.67%)

Sales by Quarter

2.78K!

Goal: 3.41K (-18.49%)

avg Sales

202.76K

20100

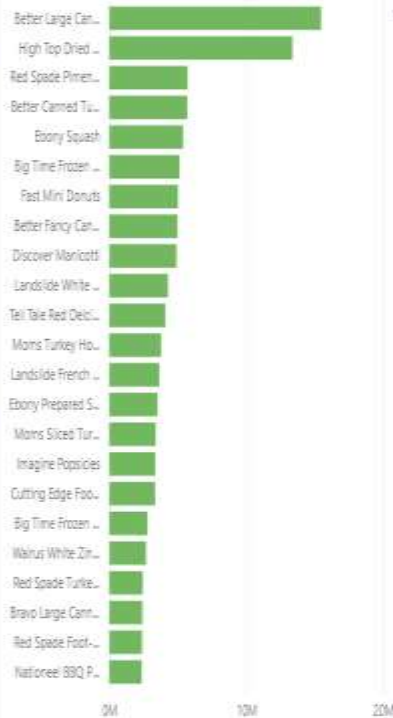
555.38K

Year

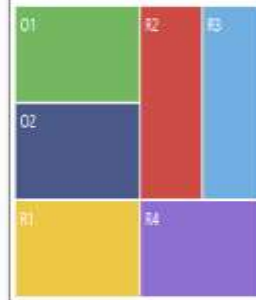
2017

2019

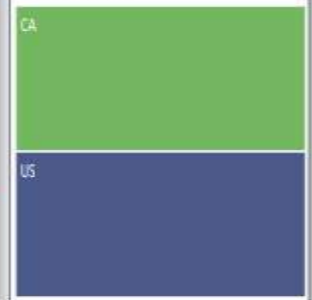
Sales Amount by Item



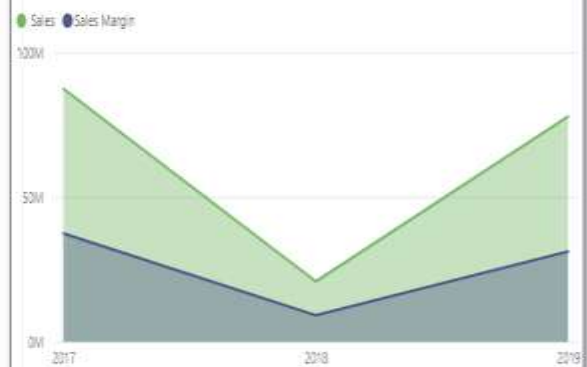
Sales by Business Family



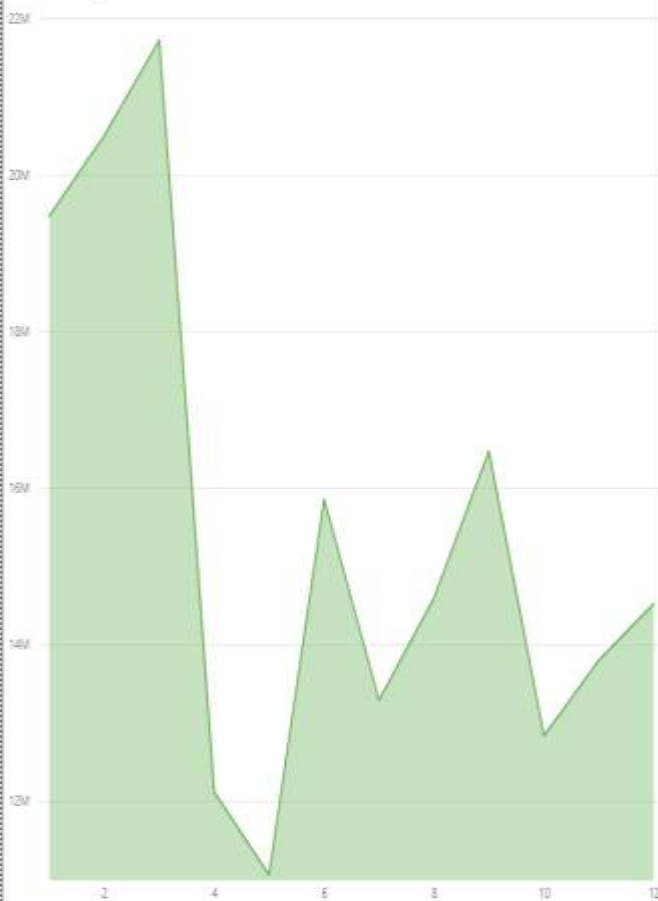
Sales by Country



Sales and Sales Margin by Year



Sales by Month



Sales by Quarter

