DATA ANALYTICS WITH COGNOS

PRODUCT SALES ANALYSIS

SUBMITTED BY:

GOKUL P -410121243010

PRAVEEN KUMAR D -410121243041

SHANMUGAM S-410121243054

MANIKANDAN K-410121243025

PRADEEPKUMAR R-410121243039

PHASE-4: DEVOLOPMENT PART 2

PROJECT

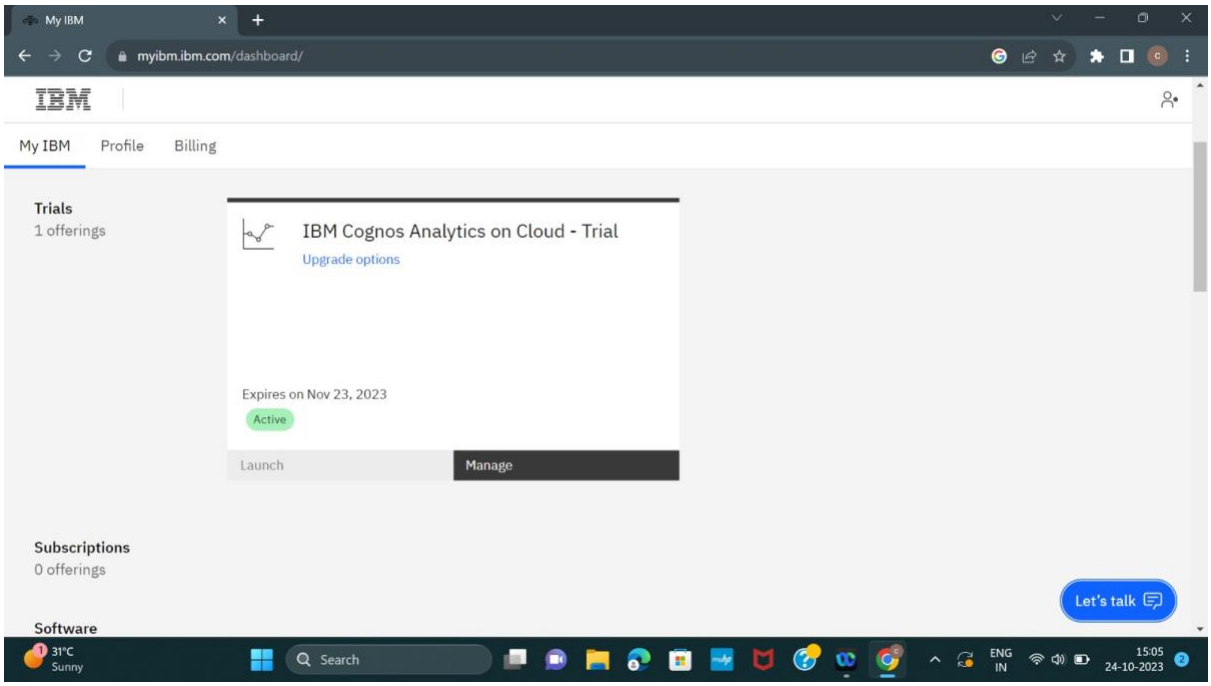
The aim of this project is to demonstrate the data analysis skills I've learned thus far and to apply them to real-world scenarios. As such, this project asks and answers real-world questions about real-world sales data. For instance, "What was the best month for sales?", or, "Which time of the day should we display advertisements to maximize the likelihood of customers' purchasing products?"

INTRODUCTION

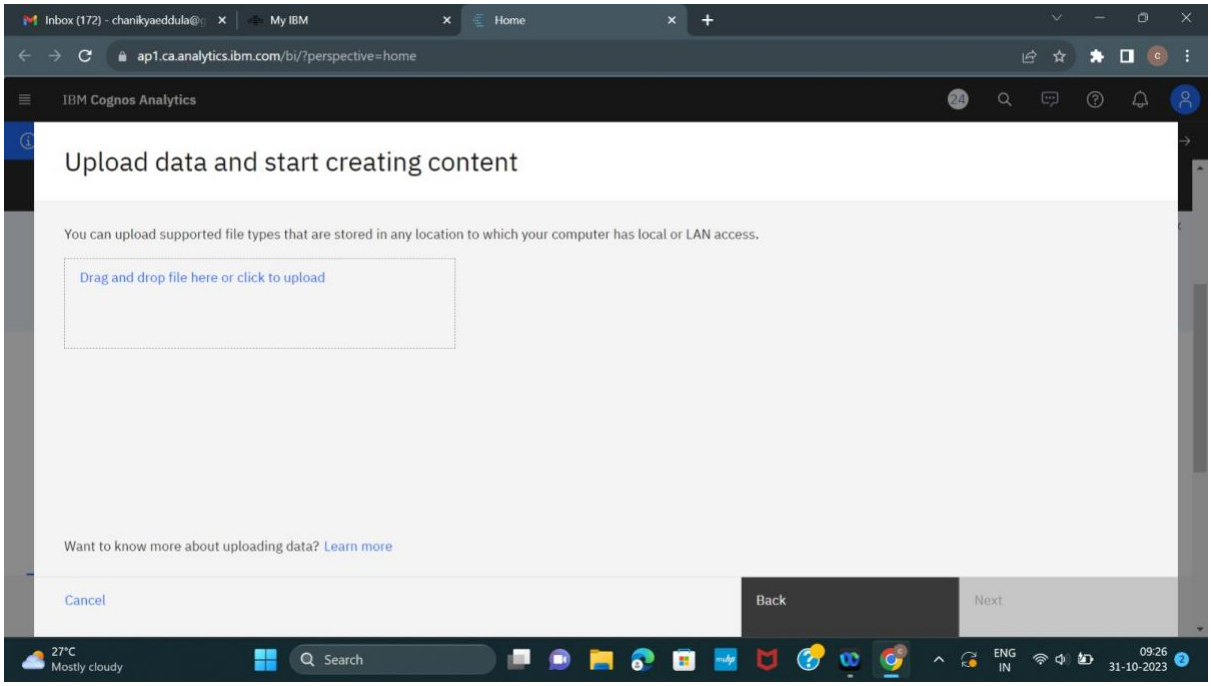
Product sales analysis of top-selling products is a critical aspect of any business's strategic planning and decision-making process. It involves the systematic examination of sales data and related information to gain valuable insights into the performance of a company's most popular and

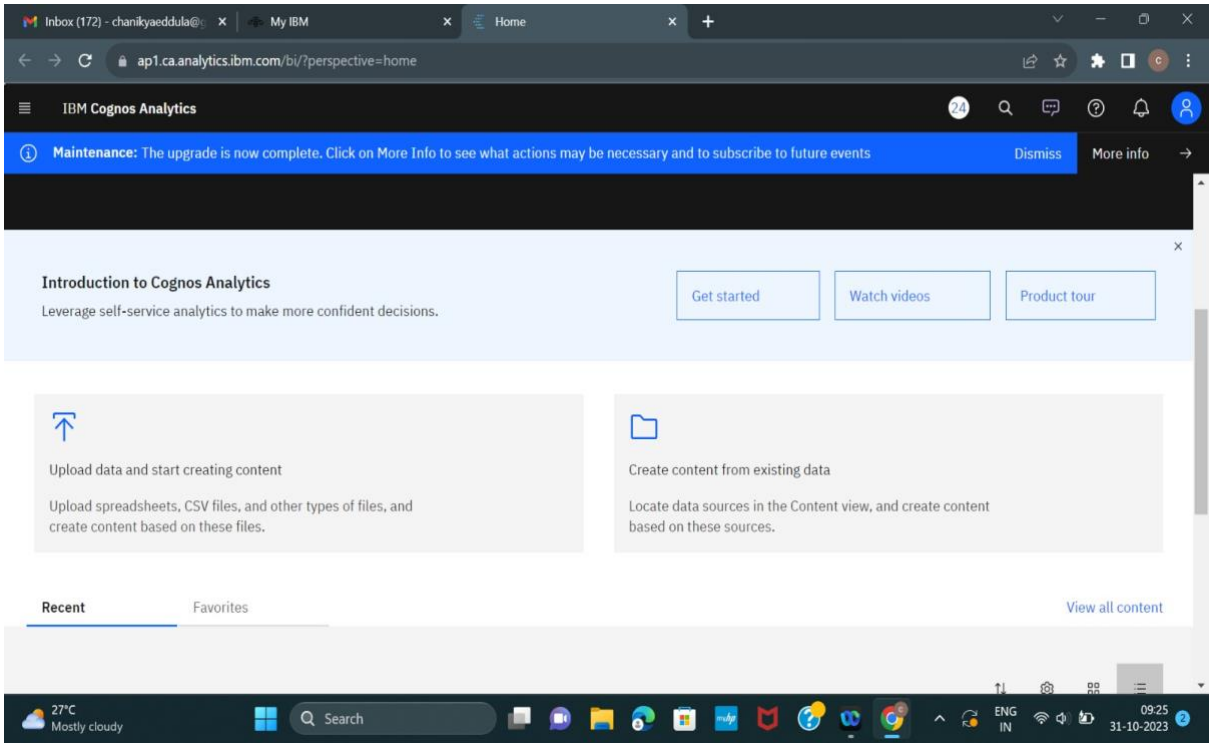
profitable products. This analysis helps businesses understand various aspects of their top-selling `products, such as their market share, customer preferences, sales trends, and overall impact on the company's bottom line.

Step1: login IBM cognos account on cloud

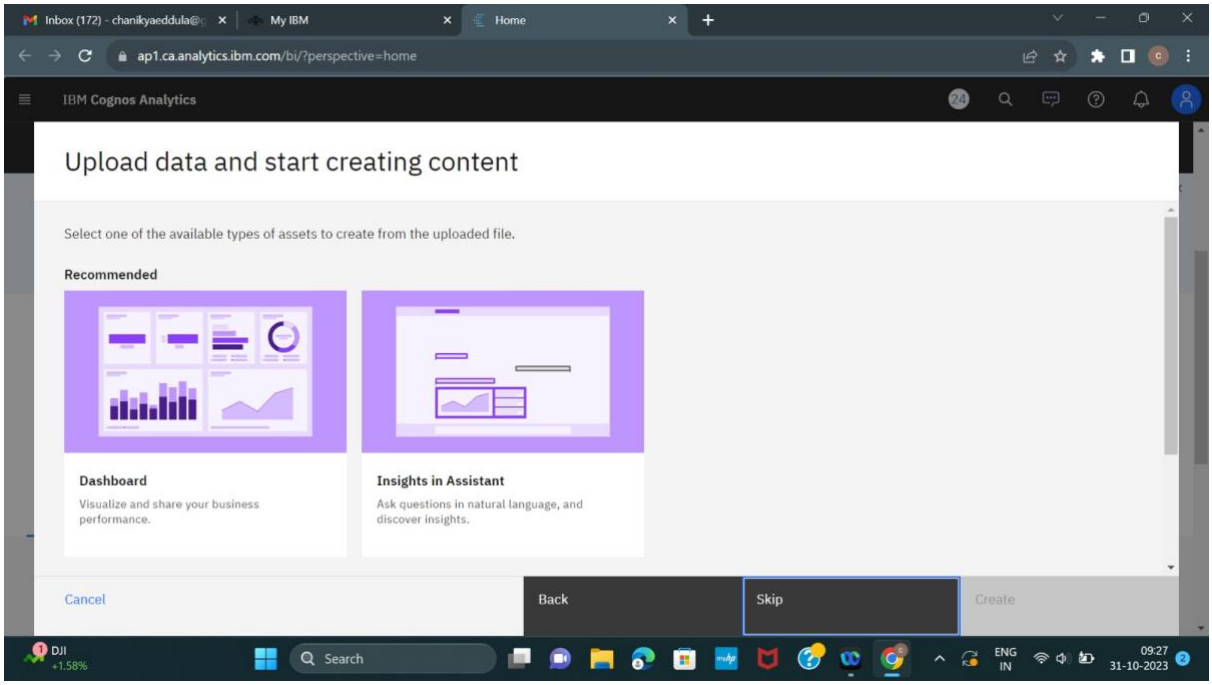


Step 2: Upload data in csv file

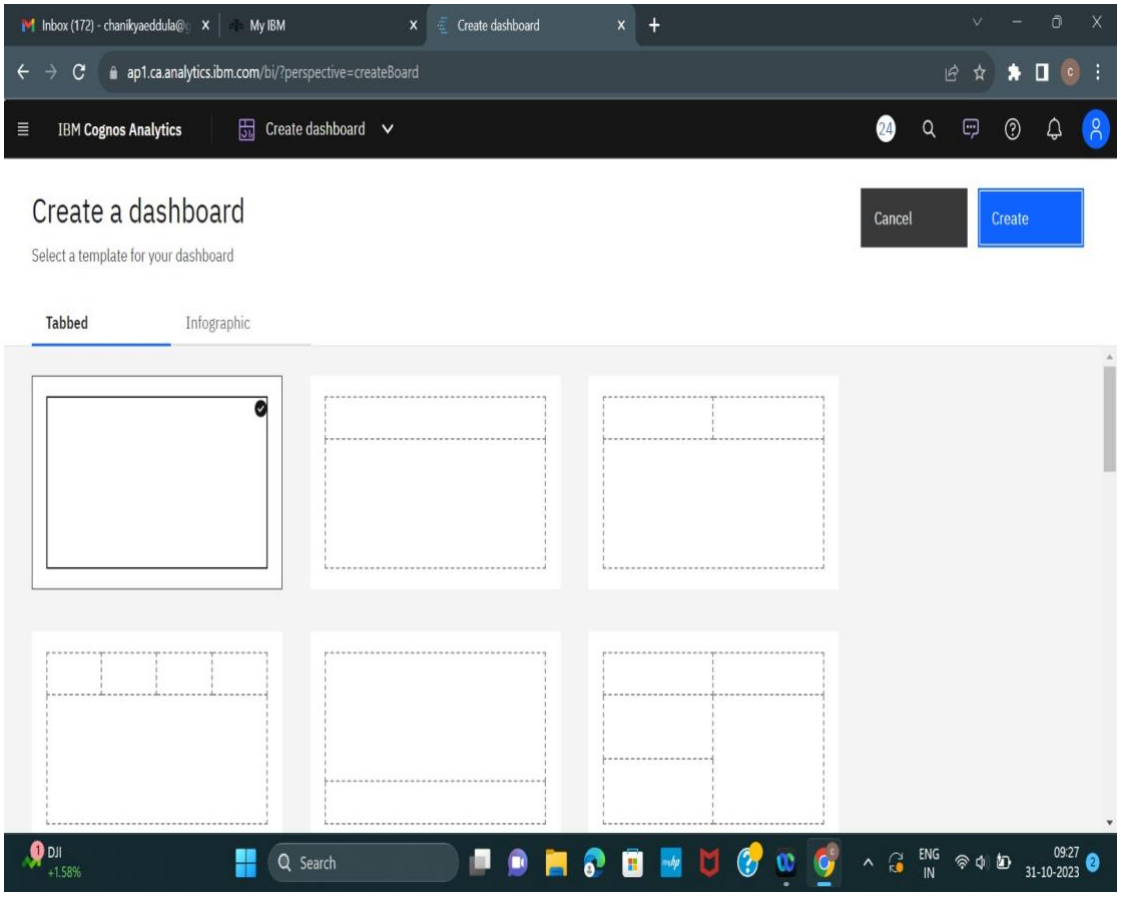




Step 3:drag and drop the file

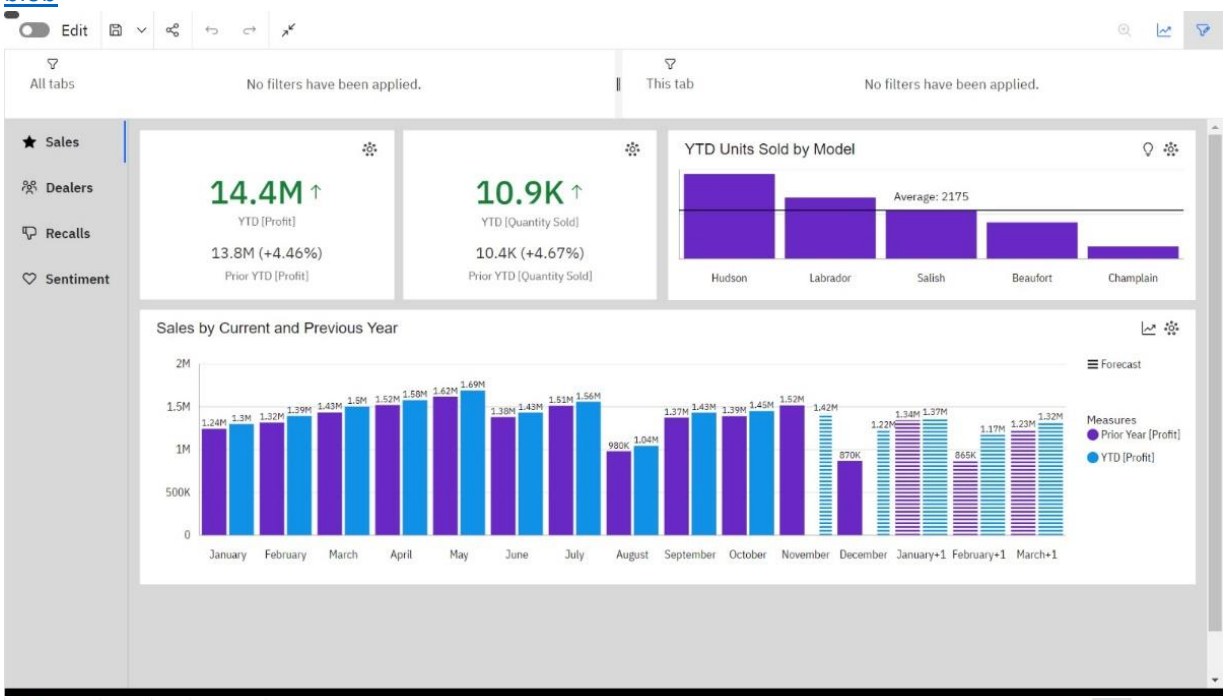


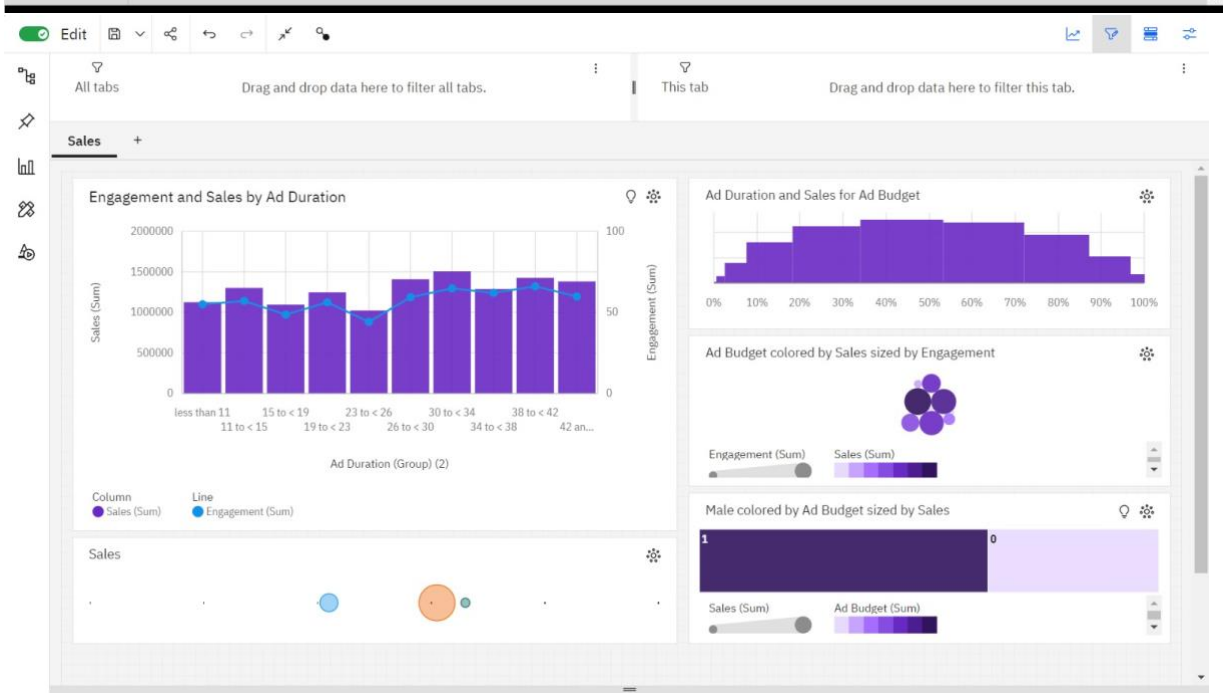
step 4: creat a dash board template

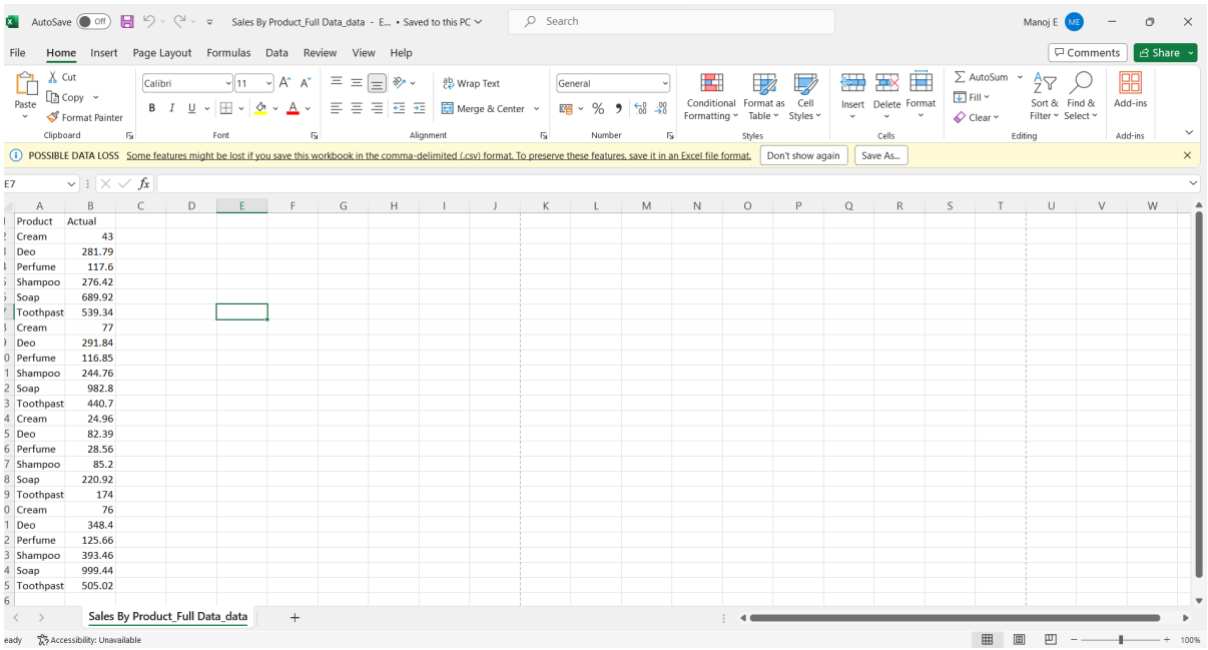


Dash board of sales analysis of product

[blob]

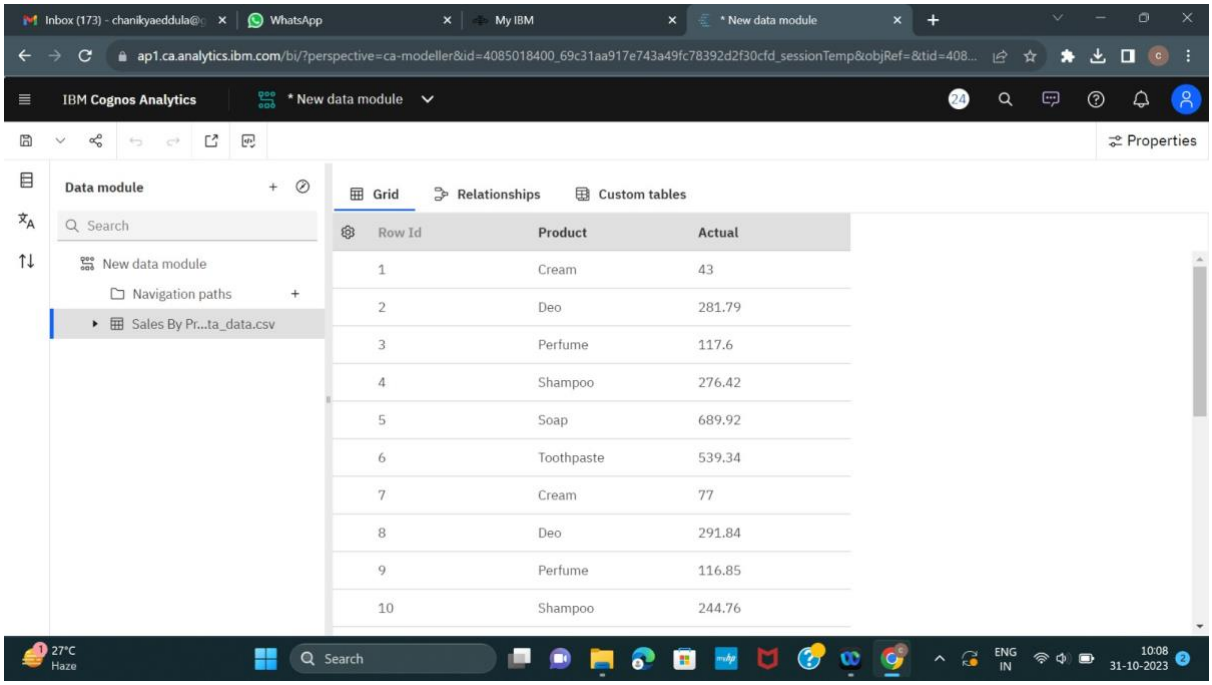

Edit
All tabs
No filters have been applied.
This tab
No filters have been applied.
Sales
Dealers
Recalls
Sentiment
14.4M ↑
YTD [Profit]
13.8M (+4.46%)
Prior YTD [Profit]
10.9K ↑
YTD [Quantity Sold]
10.4K (+4.67%)
Prior YTD [Quantity Sold]
YTD Units Sold by Model
Average: 2175
Hudson
Labrador
Salish
Beaufort
Champlain
Sales by Current and Previous Year
Forecast
Measures
Prior Year [Profit]
YTD [Profit]
January
February
March
April
May
June
July
August
September
October
November
December
January+1
February+1
March+1


Edit
All tabs
Drag and drop data here to filter all tabs.
This tab
Drag and drop data here to filter this tab.
Sales
Engagement and Sales by Ad Duration
Sales (Sum)
Engagement (Sum)
Ad Duration (Group) (2)
Column
Sales (Sum)
Line
Engagement (Sum)
Ad Duration and Sales for Ad Budget
Ad Budget colored by Sales sized by Engagement
Engagement (Sum)
Sales (Sum)
Male colored by Ad Budget sized by Sales
Sales (Sum)
Ad Budget (Sum)
Sales



Product	Actual
Cream	43
Deo	281.79
Perfume	117.6
Shampoo	276.42
Soap	689.92
Toothpast	539.34
Cream	77
Deo	291.84
Perfume	116.85
Shampoo	244.76
Soap	982.8
Toothpast	440.7
Cream	24.96
Deo	82.39
Perfume	28.56
Shampoo	85.2
Soap	220.92
Toothpast	174
Cream	76
Deo	348.4
Perfume	125.66
Shampoo	393.46
Soap	999.44
Toothpast	505.02

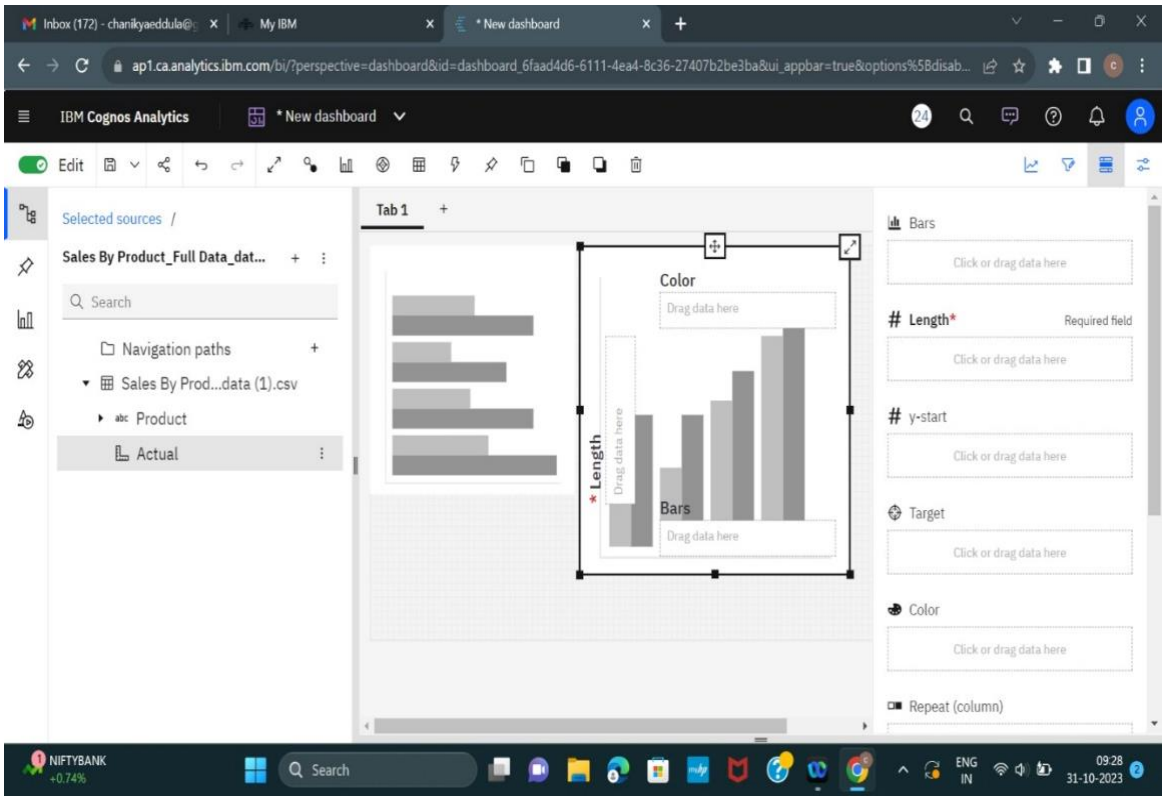
step 5: to explore data module of sales analysis



Row Id	Product	Actual
1	Cream	43
2	Deo	281.79
3	Perfume	117.6
4	Shampoo	276.42
5	Soap	689.92
6	Toothpaste	539.34
7	Cream	77
8	Deo	291.84
9	Perfume	116.85
10	Shampoo	244.76

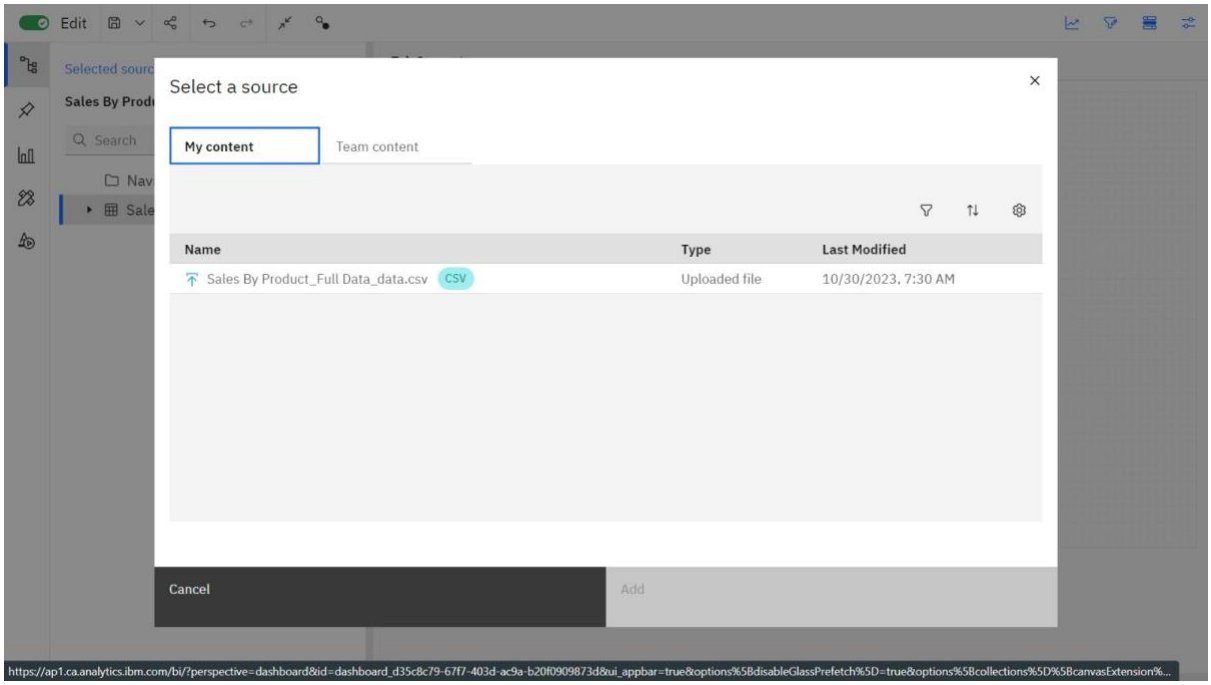
Step 6: here we to create visualizations of sales

Visualization of bar and column



after add the file in my contact source

one month to the next.



Exploring the Data

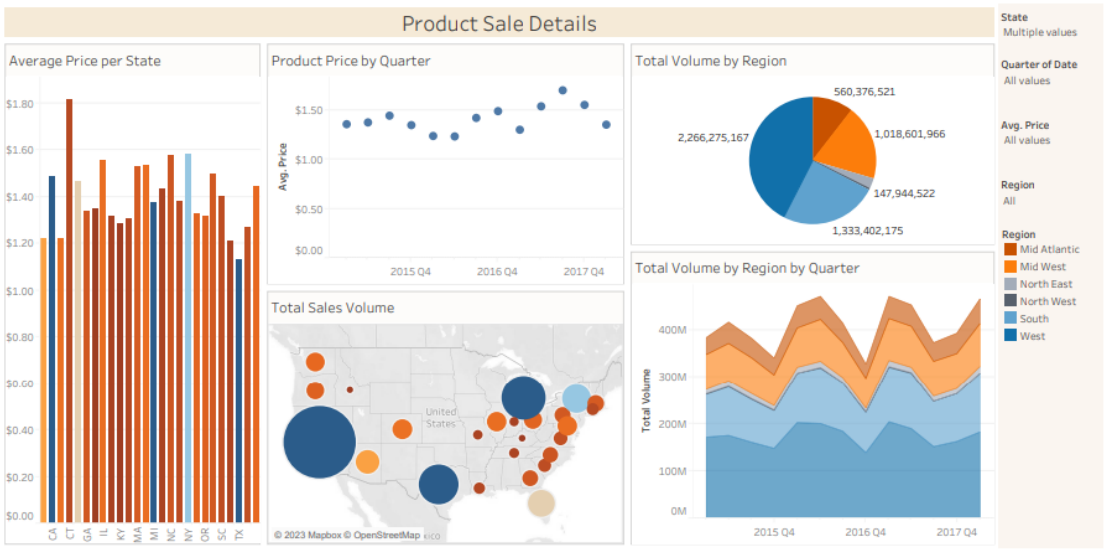
Question 1: What was the best month for sales? How much was earned that month?

Out[14]:

	Order ID	Product	Quantity Ordered	Price Each	Sales	Order Date	Months	Purchase Address
0	176558	USB-C Charging Cable	2	11.950000	23.900000	19/04/19 08:46	Apr	917 1st St, Dallas, TX 75001
1	176559	Bose SoundSport Headphones	1	99.989998	99.989998	07/04/19 22:30	Jul	682 Chestnut St, Boston, MA 02215
2	176560	Google Phone	1	600.000000	600.000000	12/04/19 14:38	Dec	669 Spruce St, Los Angeles, CA 90001
3	176560	Wired Headphones	1	11.990000	11.990000	12/04/19 14:38	Dec	669 Spruce St, Los Angeles, CA 90001
4	176561	Wired Headphones	1	11.990000	11.990000	30/04/19 09:27	Apr	333 8th St, Los Angeles, CA 90001

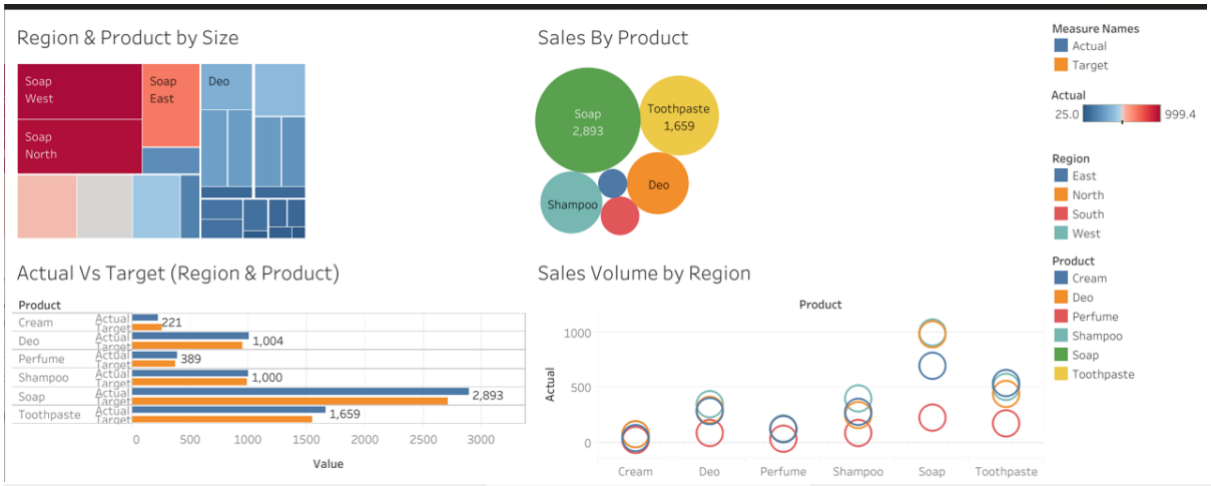
In [15]:
```
#Third, grouping the data by month and calculating the total sales (and quantities) amount for each month
sales_per_month = df.groupby(['Months']).sum()[['Quantity Ordered', 'Sales']]

#sorting the data in alphabetical order by month name
#specifying the sorting order
sort_order = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']

#sorting the data
sales_per_month = sales_per_month.reindex(sort_order)

#to present the sales in USD and display them in a reader-friendly manner
sales_per_month_USD = sales_per_month.copy()
sales_per_month_USD['Sales'] = sales_per_month_USD['Sales'].apply(lambda sale: '${:,.2f}'.format(sale))

#displaying the results
print('The following table displays the total sales amount (and quantities ordered) for each month:')
sales_per_month_USD
```

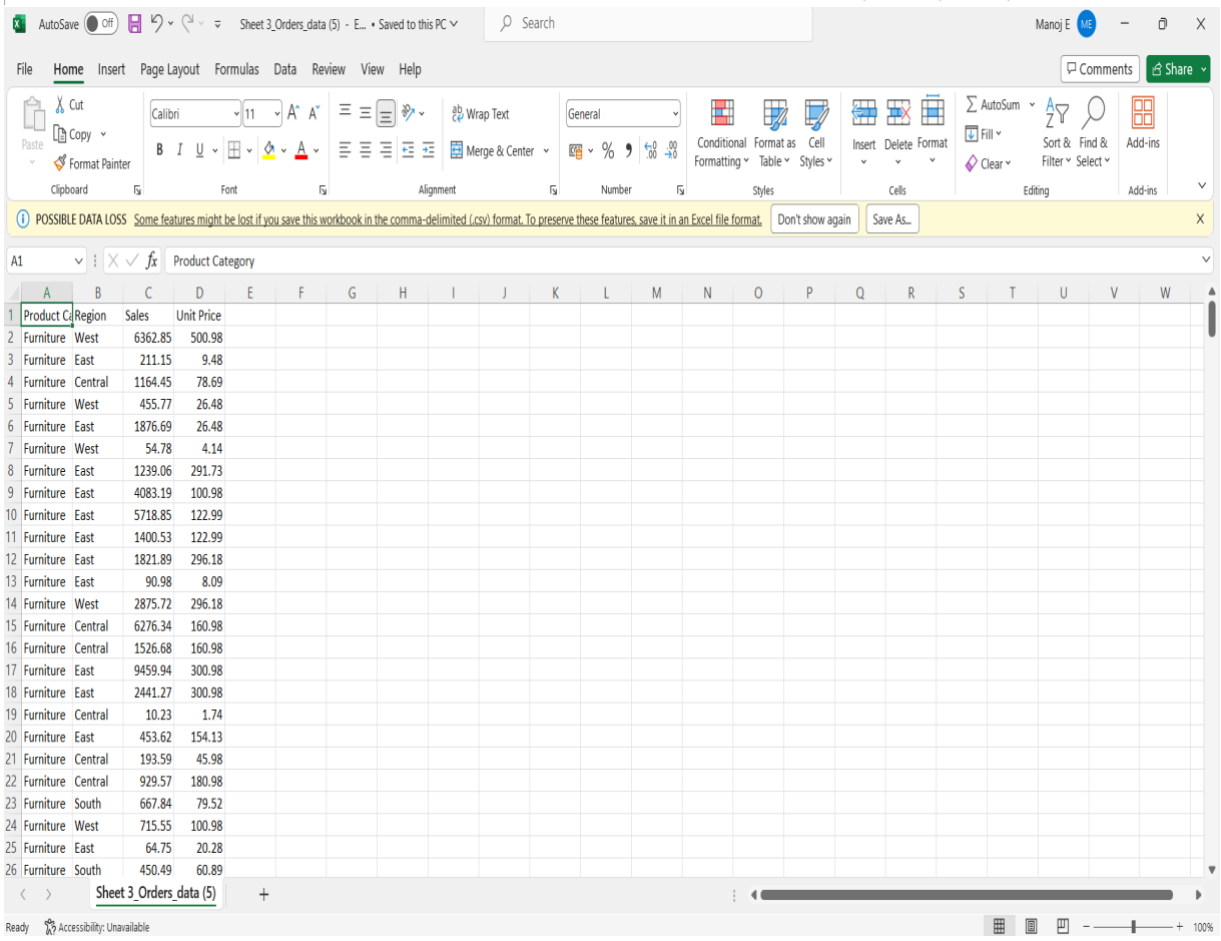
```
The following table displays the total sales amount (and quantities ordered) for each month:
```



Rise in Online Sales

While this year there has been a decline in overall sales, we can see a **shift towards online sales**. In **2020**, online sales rose by **72%** , while in store sales plummeted resulting in the closure of 2 stores.

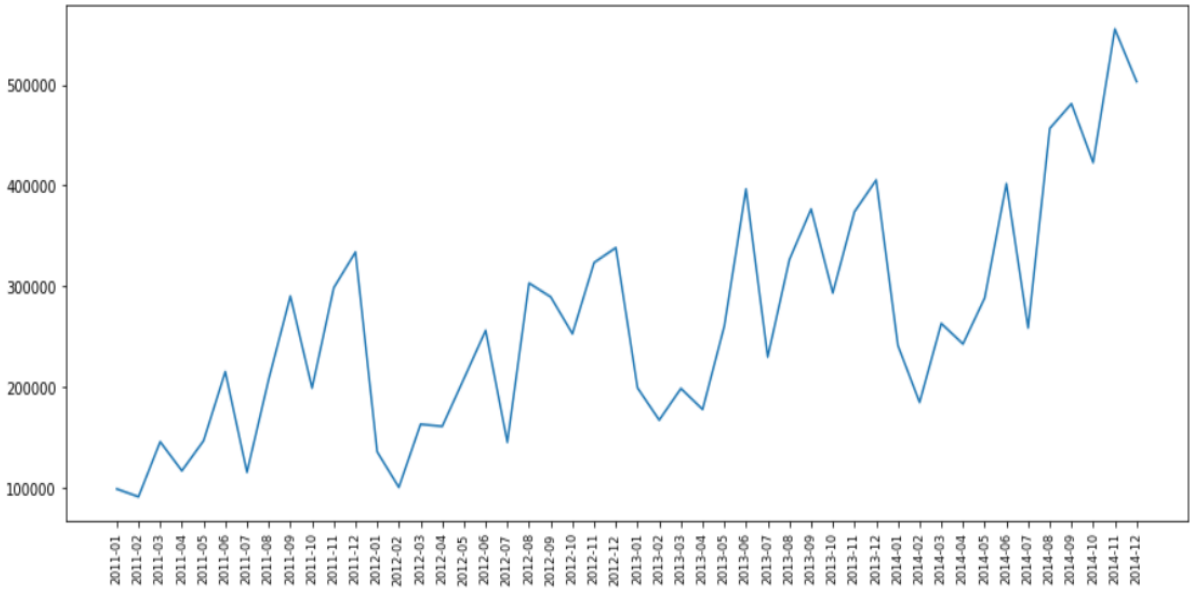




	A	B	C	D
1	Product Ca	Region	Sales	Unit Price
2	Furniture	West	6362.85	500.98
3	Furniture	East	211.15	9.48
4	Furniture	Central	1164.45	78.69
5	Furniture	West	455.77	26.48
6	Furniture	East	1876.69	26.48
7	Furniture	West	54.78	4.14
8	Furniture	East	1239.06	291.73
9	Furniture	East	4083.19	100.98
10	Furniture	East	5718.85	122.99
11	Furniture	East	1400.53	122.99
12	Furniture	East	1821.89	296.18
13	Furniture	East	90.98	8.09
14	Furniture	West	2875.72	296.18
15	Furniture	Central	6276.34	160.98
16	Furniture	Central	1526.68	160.98
17	Furniture	East	9459.94	300.98
18	Furniture	East	2441.27	300.98
19	Furniture	Central	10.23	1.74
20	Furniture	East	453.62	154.13
21	Furniture	Central	193.59	45.98
22	Furniture	Central	929.57	180.98
23	Furniture	South	667.84	79.52
24	Furniture	West	715.55	100.98
25	Furniture	East	64.75	20.28
26	Furniture	South	450.49	60.89

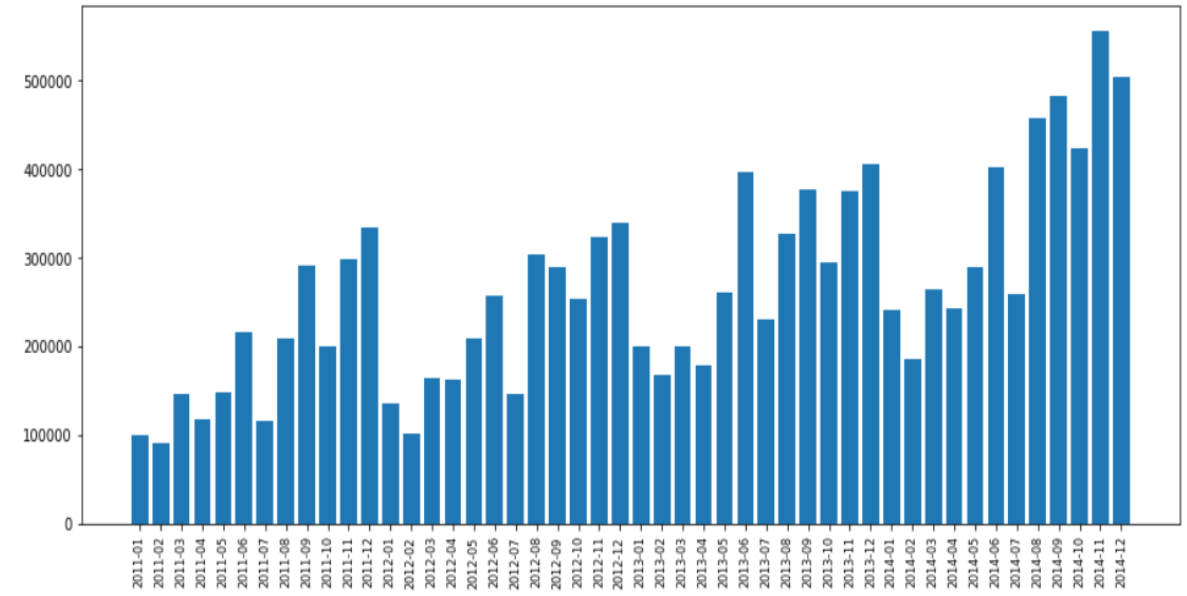
VISUALIZE SALES TRAND BY MONTHS

```
plt.figure(figsize=(15,6))
plt.plot(sales_by_month['month_year'],sales_by_month['sales'])
plt.xticks(rotation='vertical',size=8)
plt.show()
```



]:

```
plt.figure(figsize=(15,6))
plt.bar(sales_by_month['month_year'],sales_by_month['sales'])
plt.xticks(rotation='vertical',size=8)
plt.show()
```



DISPLAY MOST SELLING PRODCUTS

In [22]:
```
products_sales = pd.DataFrame(sales.groupby('product_name').sum()['sales'])
products_sales = products_sales.sort_values('sales',ascending=False)
```

TOP 10 MOST SALES PRODUCTS

In [23]:
```
products_sales[:10]
```

Out[23]:

product_name	sales
Apple Smart Phone, Full Size	86935.7786
Cisco Smart Phone, Full Size	76441.5306
Motorola Smart Phone, Full Size	73156.3030
Nokia Smart Phone, Full Size	71904.5555
Canon imageCLASS 2200 Advanced Copier	61599.8240
Hon Executive Leather Armchair, Adjustable	58193.4841
Office Star Executive Leather Armchair, Adjustable	50661.6840
Harbour Creations Executive Leather Armchair, Adjustable	50121.5160
Samsung Smart Phone, Cordless	48653.4600
Nokia Smart Phone, with Caller ID	47877.7857

:
```
products_by_quantity = pd.DataFrame(sales.groupby('product_name').sum()['quantity'])
products_by_quantity_sorted = products_by_quantity.sort_values('quantity',ascending=False)
```

TOP 10 MOST QUANTITY SELLING PRODUCTS ITEMS

:
```
products_by_quantity_sorted[:10]
```

:

product_name	quantity
Staples	876
Cardinal Index Tab, Clear	337
Eldon File Cart, Single Width	321
Rogers File Cart, Single Width	262
Sanford Pencil Sharpener, Water Color	259
Stockwell Paper Clips, Assorted Sizes	253
Avery Index Tab, Clear	252
Ibico Index Tab, Clear	251
Smead File Cart, Single Width	250
Stanley Pencil Sharpener, Water Color	242

VISUALIZE MOST USED SHIP MODS

In [14]:
```
sales.isnull().sum()
```

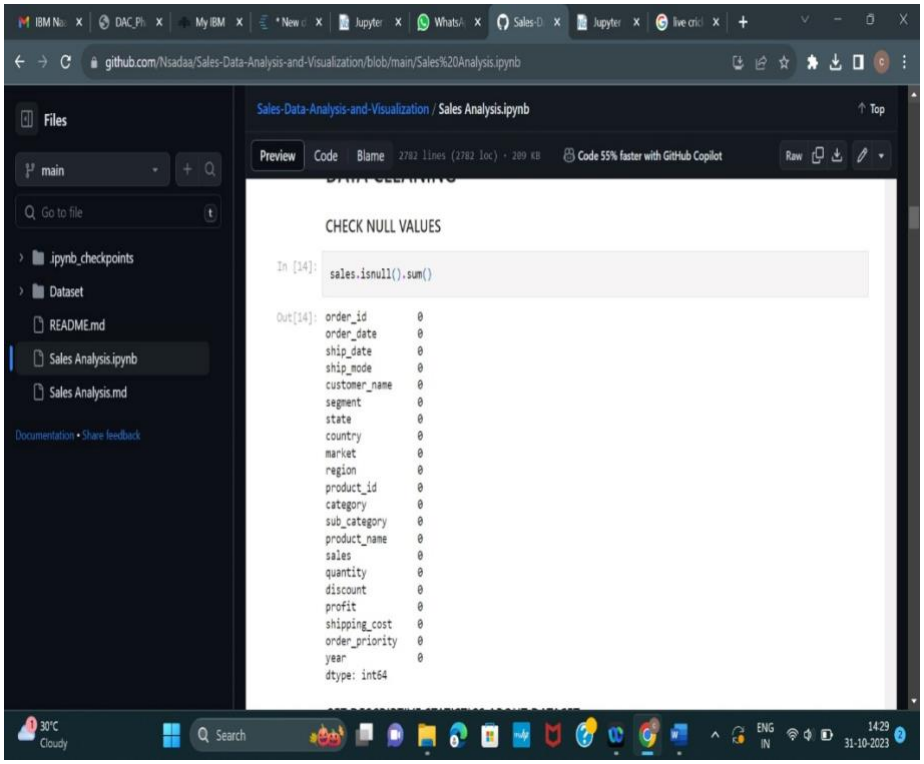
Out[14]:
```
order_id          0
order_date        0
ship_date         0
ship_mode         0
customer_name     0
segment           0
state             0
country           0
market            0
region            0
product_id        0
category          0
sub_category      0
product_name      0
sales             0
quantity          0
discount          0
profit            0
shipping_cost     0
order_priority    0
year              0
dtype: int64
```

GET INFORMATIONS ABOUT DATASET

[13]:
```
sales.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51290 entries, 0 to 51289
Data columns (total 21 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   order_id        51290 non-null  object
 1   order_date      51290 non-null  datetime64[ns]
 2   ship_date       51290 non-null  datetime64[ns]
 3   ship_mode       51290 non-null  object
 4   customer_name   51290 non-null  object
 5   segment         51290 non-null  object
 6   state           51290 non-null  object
 7   country         51290 non-null  object
 8   market          51290 non-null  object
 9   region          51290 non-null  object
 10  product_id      51290 non-null  object
 11  category        51290 non-null  object
 12  sub_category    51290 non-null  object
 13  product_name    51290 non-null  object
 14  sales           51290 non-null  float64
 15  quantity        51290 non-null  int64
 16  discount        51290 non-null  float64
 17  profit          51290 non-null  float64
 18  shipping_cost   51290 non-null  float64
 19  order_priority  51290 non-null  object
 20  year            51290 non-null  int64
dtypes: datetime64[ns](2), float64(4), int64(2), object(13)
memory usage: 8.2+ MB
```

DATA CLEANING



GETTING KNOW ABOUT DATSET SHAPE & COLUMNS

In [11]:
```
sales.shape
```

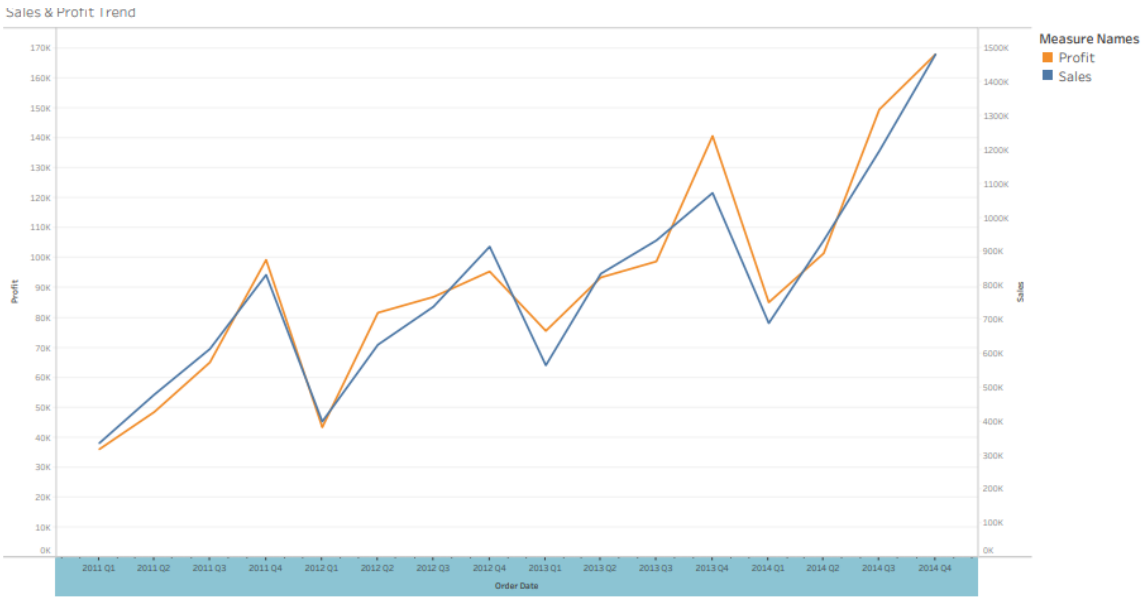
Out[11]: (51290, 21)

In [12]:
```
for columns in sales.columns:
        print(columns)
```

```
order_id
order_date
ship_date
ship_mode
customer_name
segment
state
country
market
region
product_id
category
sub_category
product_name
sales
quantity
discount
profit
shipping_cost
order_priority
year
```

GET INFORMATIONS ABOUT DATASET

Sales and profit trend



Conclusion : In this document we have seen using IBM cognos to design interactive dashboards and reports that display such as top-selling products , sales trends ,and customer preferences and from the visualizations ,such as identifying poducts with the highest sales , peak sales periods.