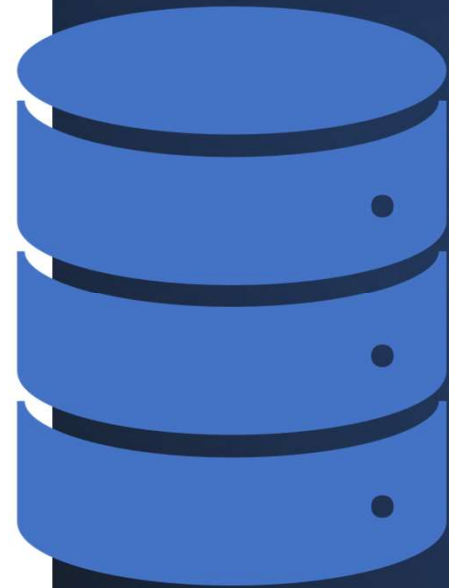


Analyzing Amazon Sales data

Gaurav Kumar Gautam

Objective

- Development of a predictive model for predicting sales.
- Perform ETL (Extract-Transform-Load) on dataset.
- Develop a dashboard by using Tableau.



What I have learnt from this data

01

Better understand
and optimize
revenue generation
in future

02

Maximize
forecasting accuracy

03

Make current sales
experience our top
priority

Data Preprocessing:

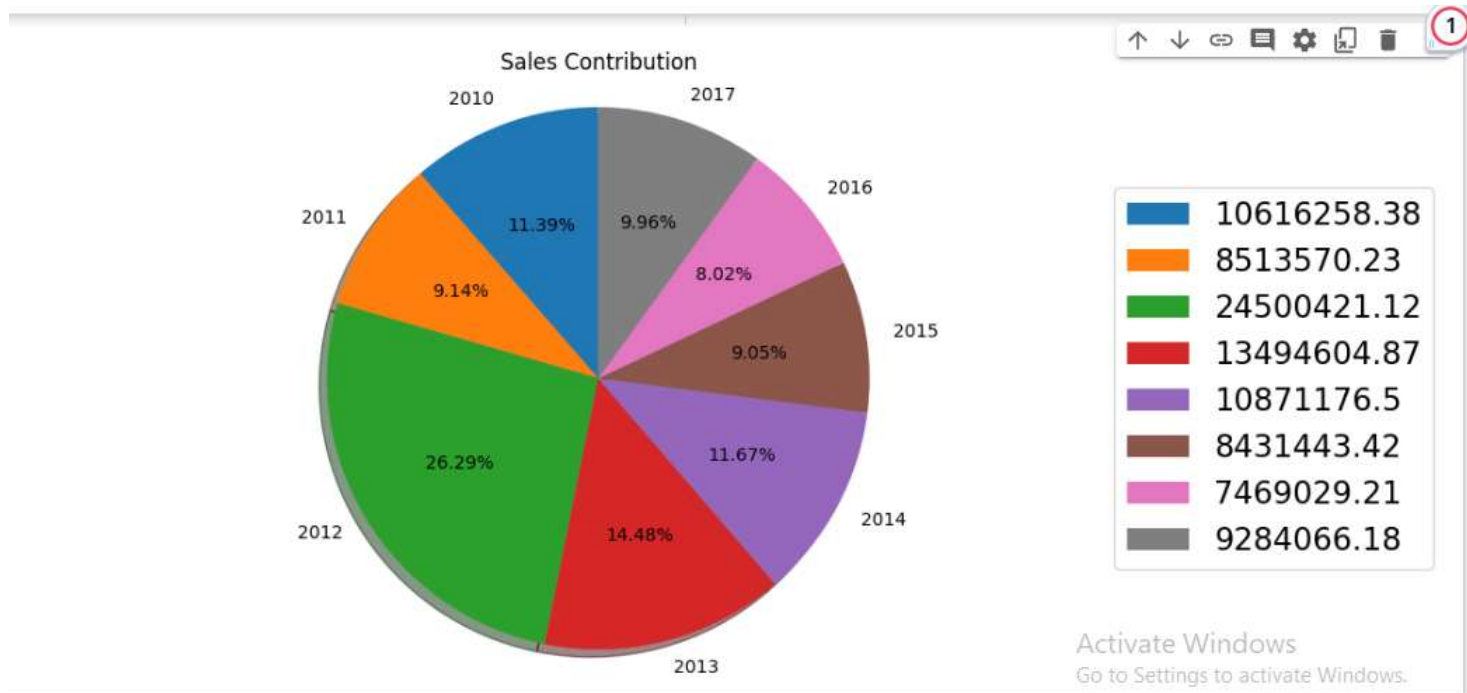
- Importing necessary libraries for data analysis such as Pandas, NumPy, Matplotlib & Seaborn, etc.
- Using `pd.read_csv()` function stores the data in pandas data frame named data.
- Using `data.columns` showing columns present in the data frame.
- `info()` function shows basic information about the data frame like the null value count of each column and their data type
- Changing the data type of different columns for model training and analysis
- Using `describe` function on the data frame for getting basic stats of the numerical dataset
- Adding an extra column to the data frame which contains only month, year, and month with year
- Using `isnull().sum()` check out the total null value in all the columns of data frame

Exploratory data Analysis

- Checking Outliers in the data frame by using Box Plot
 - Box Plot for Total Profit: Here we detect outliers in the specified column using the Z-score method and found 7 outliers
- Creating a bar chart for Total Revenue and Order Month: where it showcases the number of orders purchased in a particular month.
- Calculating the total revenue for each group with respect to Item Type and then sorting them in descending order.
- Calculating the total profit for each group with respect to Item Type and then sorting them in descending order
- Calculating correlation of the 'Total Revenue', 'Total Cost', and 'Total Profit' columns present in data frame

Predictive Analytics

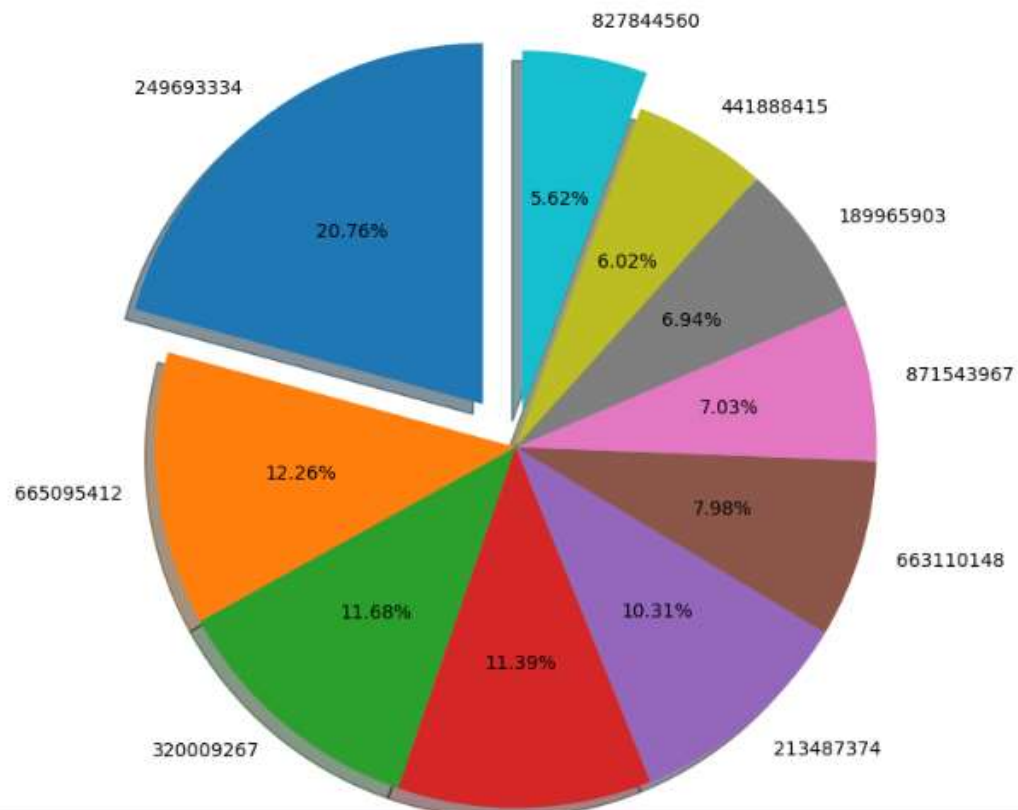
- Label Encoding of Item Type, Sales Channel, and Order Priority for model training.
- Dropping columns Region, Country, Order Date Month Year, Order ID, and Ship Date
- Plotting heat map to know correlations between each column
- Plotting bar plot to know yearly sales
- Plotting line plot to know the total cost of shipping each month
- Plotting Piechart to know sales over 3 years



Sales percentage in different years



Yearly sales profit



Activ
Go to

Total Profit on each order ID