**Super Store Case Study Report**

Objective: To clean and analyze the super store dataset and draw meaningful insights.

Tech Stack Used:

Python Libraries:

1. Pandas
2. NumPy
3. Seaborn
4. Matplotlib

Procedure

1. Step 1.

Import the Libraries:

A picture containing screenshot, text, multimedia software, software

Description automatically generated

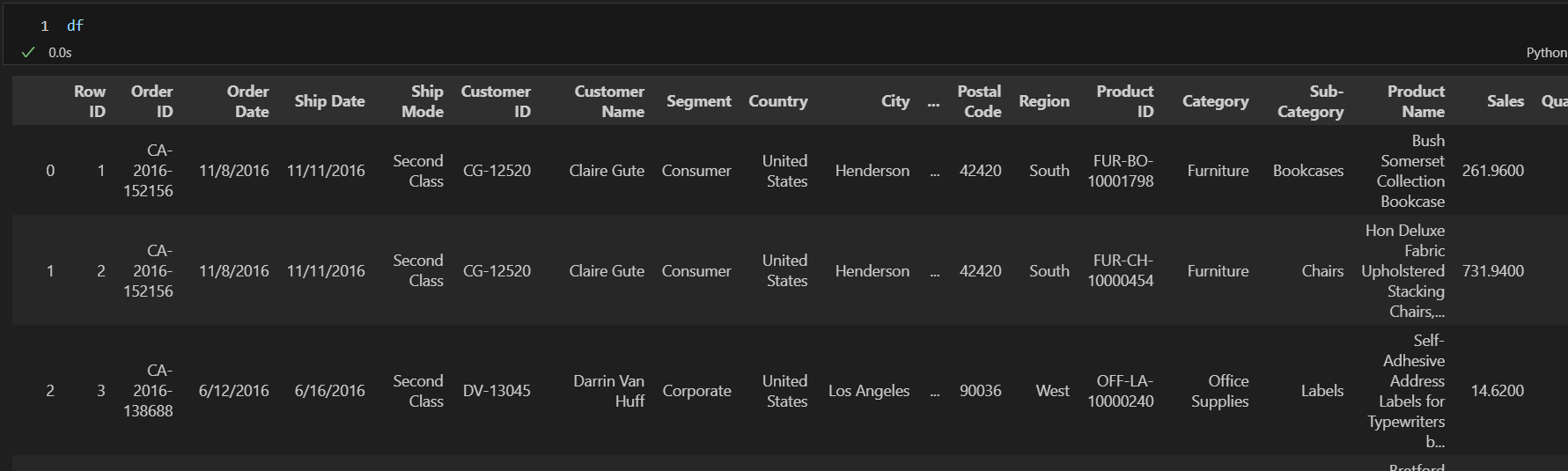
1. Step 2.

Load dataset:



1. Step 3.

Going Through:



Dataset size:

A picture containing screenshot, multimedia software

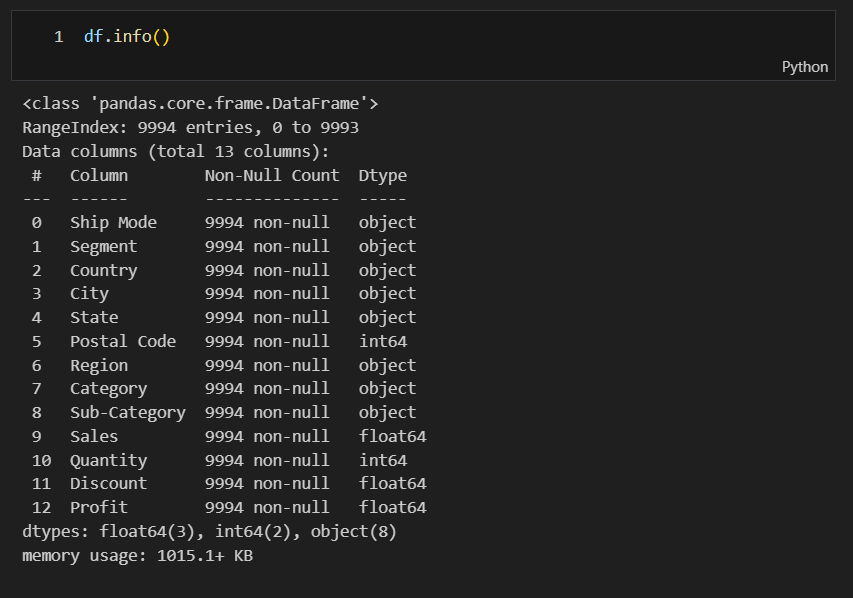
Description automatically generated

Column names:

A screen shot of a computer

Description automatically generated with medium confidence

Column data types:



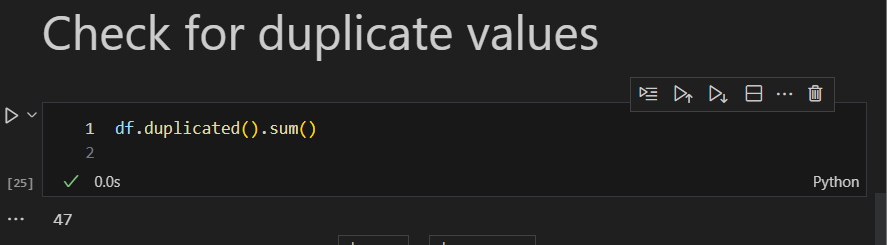
1. Step 4.

Data Cleaning:

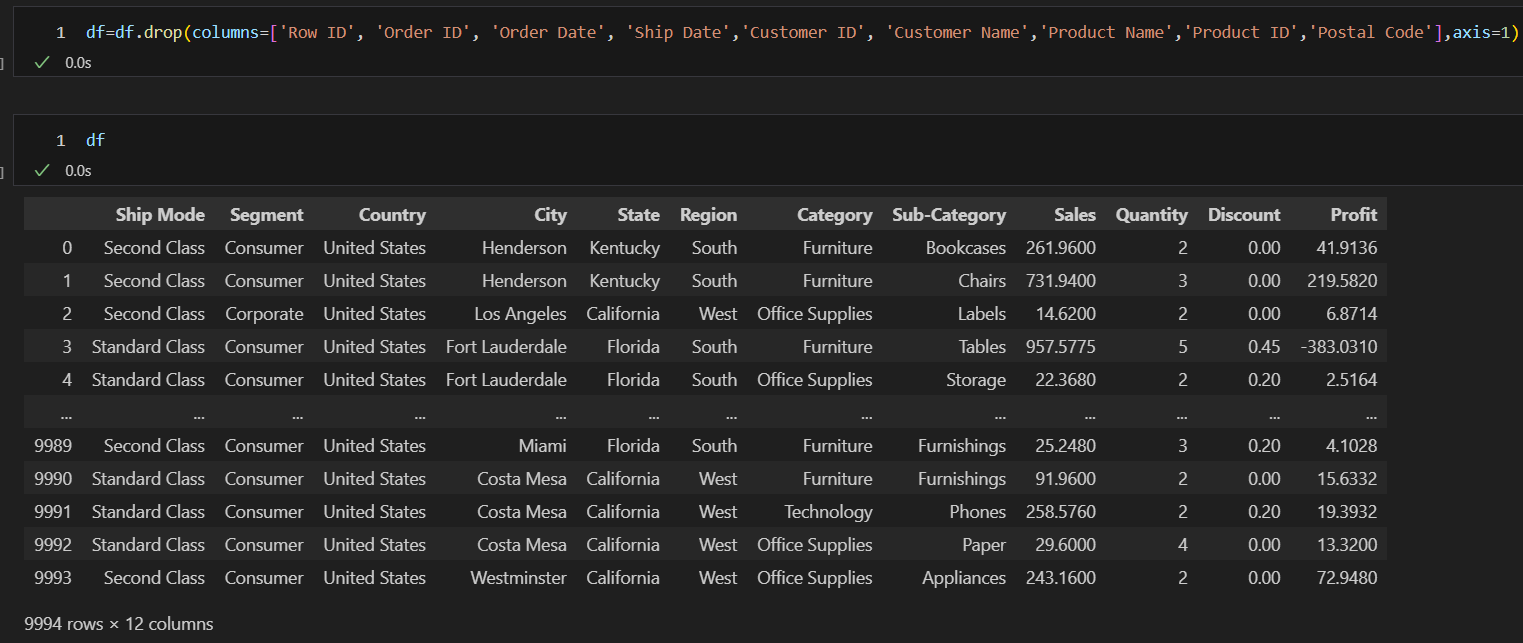
Check for null values:



Check for Duplicate Values



Dropping Redundant Columns:



A picture containing screenshot, text, multimedia software

Description automatically generated

Here the number of columns was reduced from 21 to 13.

1. Step 5.

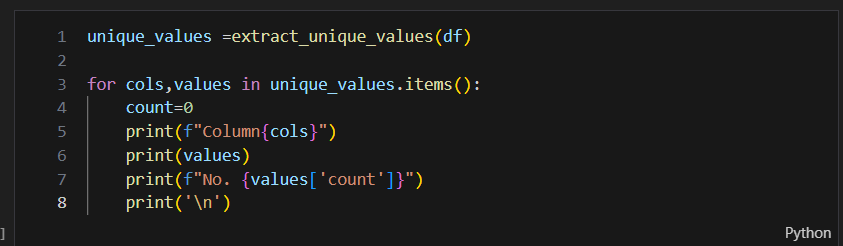
**Exploratory Data Analysis:**

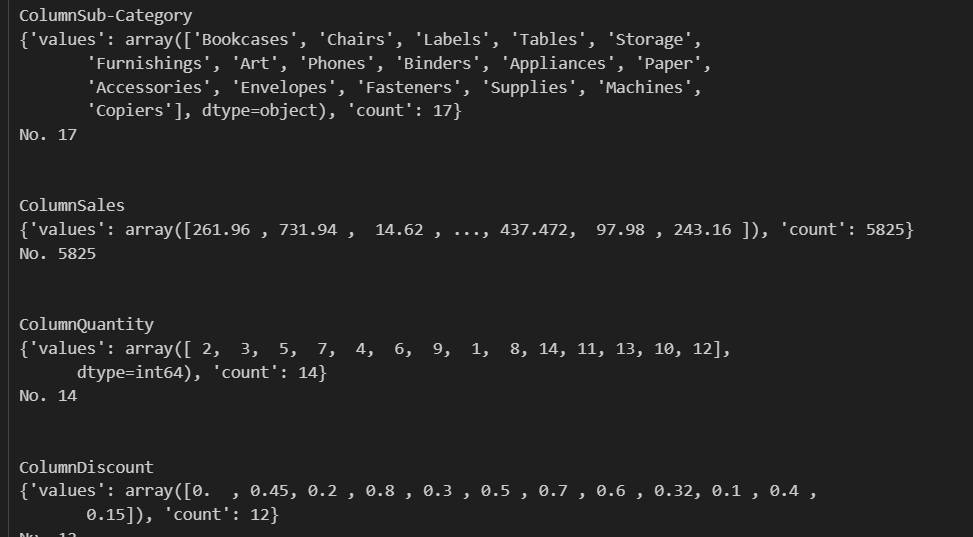
Checking for unique values in each column

This is a function that goes through each column and returns the number of unique values in the column.

A picture containing text, screenshot, software, multimedia software

Description automatically generated





Number of Total items:

A screen shot of a computer

Description automatically generated with medium confidence

Number of Total items in each category:

A screen shot of a computer

Description automatically generated with medium confidence

Number of items in subcategory:

A screenshot of a computer

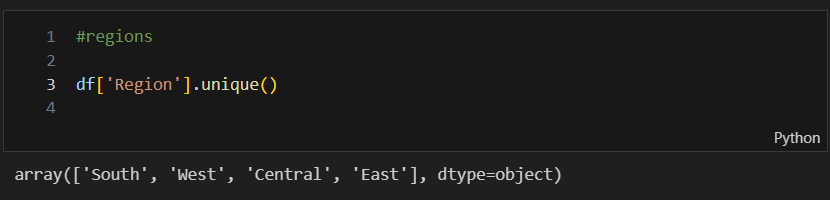
Description automatically generated

Unique Country:

A black screen with white text

Description automatically generated with low confidence

Unique Regions in the Country:



Statistical Summary:

A screenshot of a computer

Description automatically generated

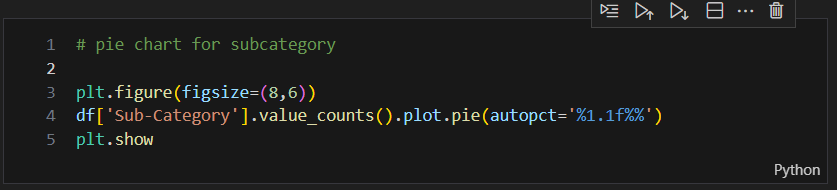
Correlation Between numeric Attributes:

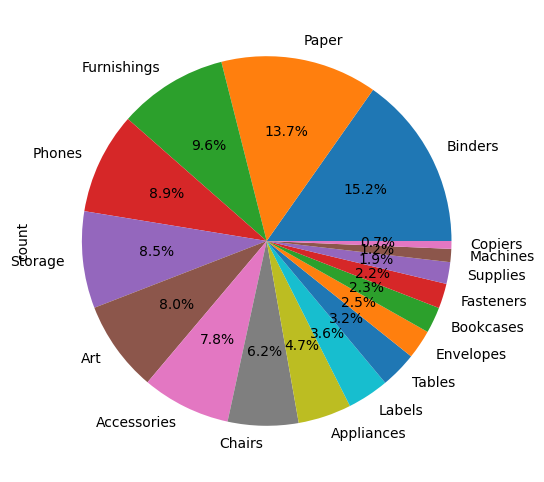
A screenshot of a computer

Description automatically generated with medium confidence

Visualization:

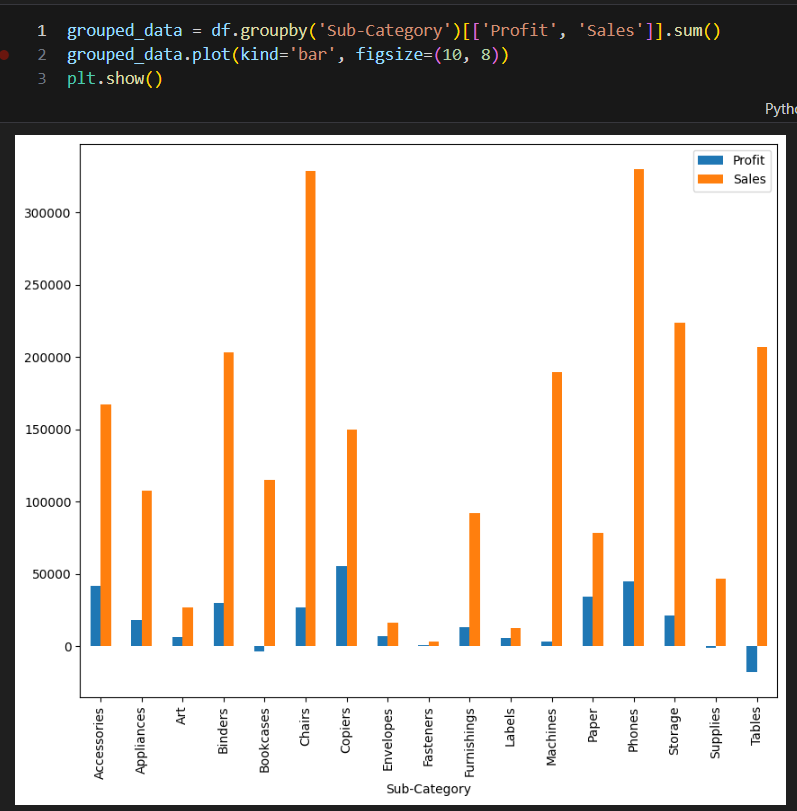
Pie chart for Sub category:



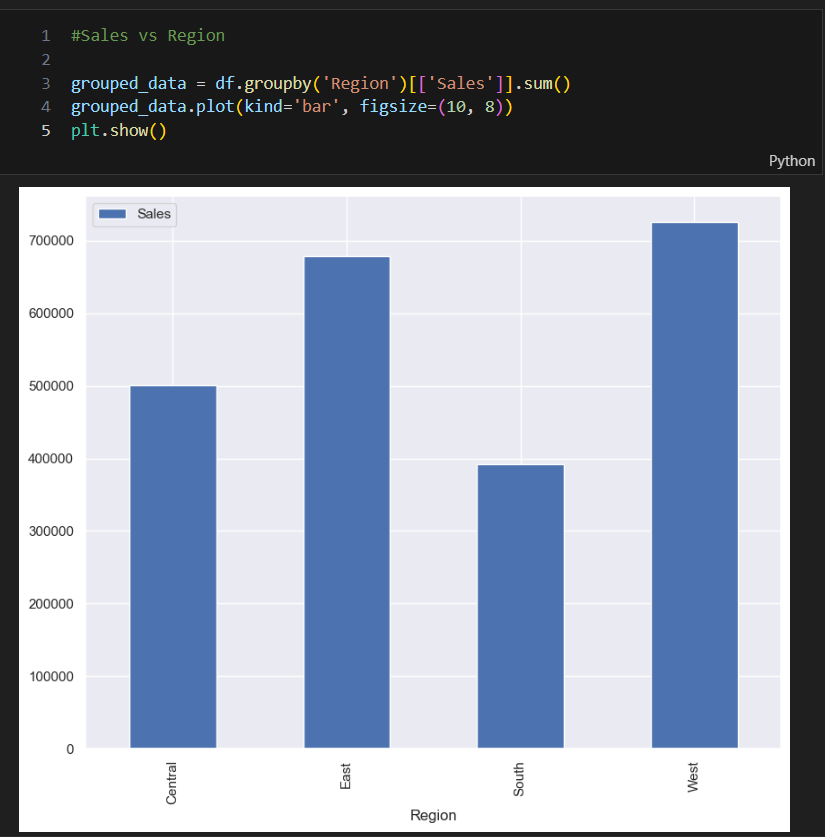


This shows Paper and Blinder contribute the most with respect to count.

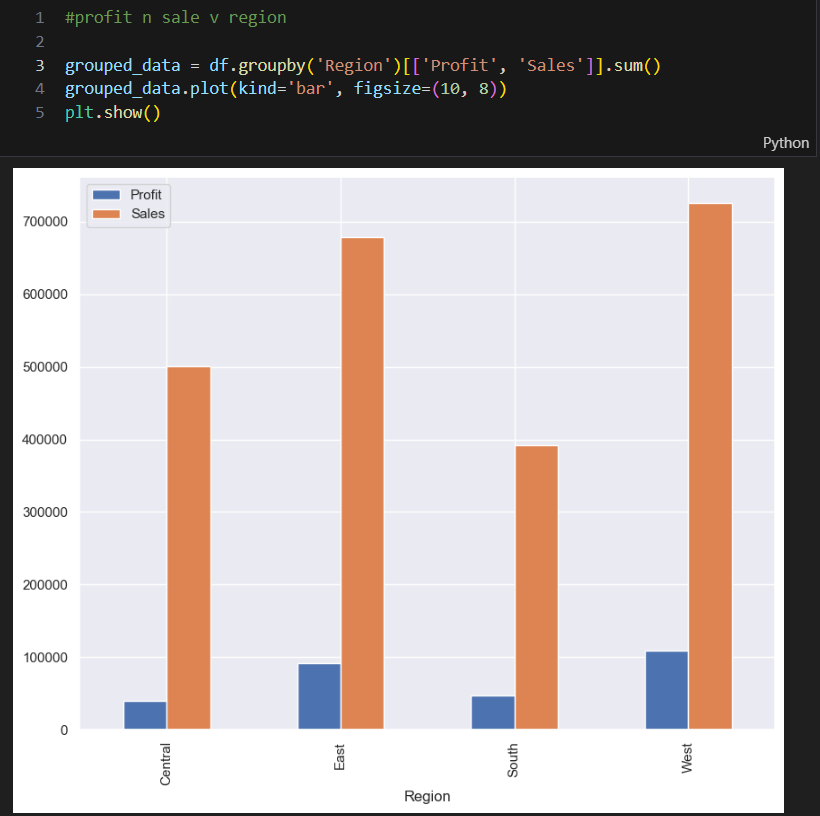
Profit vs Sales for Subcategory



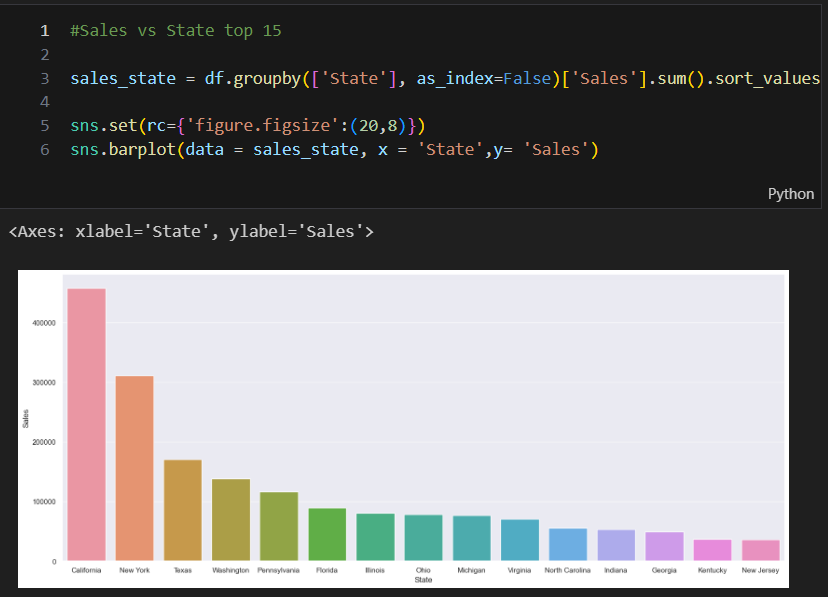
Sales vs Region



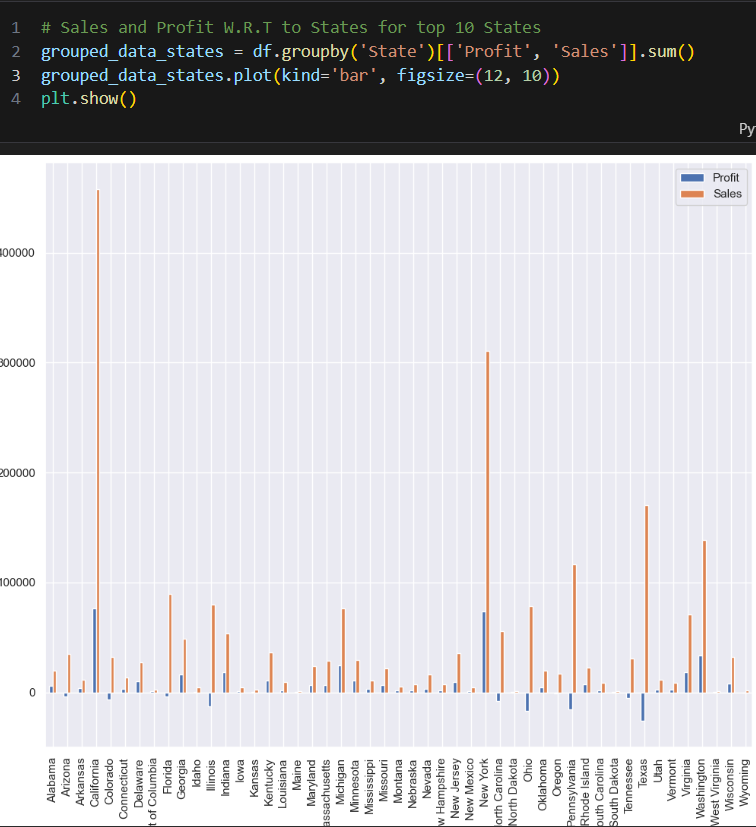
Profit and Sale for each Region



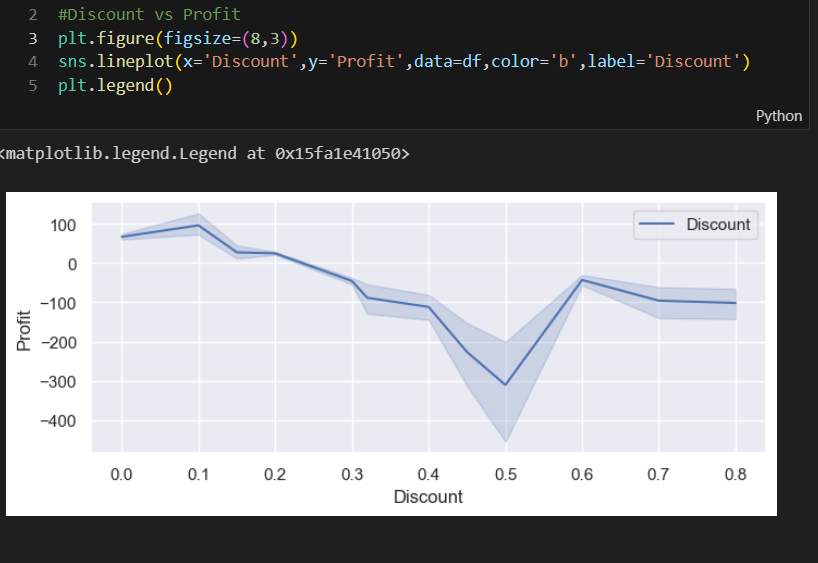
Sales vs State for top 15



Sales and Profit WRT States for top 10



Discount vs Profit



Relation Between Different Numerical Attributes and subcategory



Conclusion

1)In terms of sales and profit table are the non-performing assets their net profit is negative.

2) Here you can see that sales in California state are the highest followed by New York but in terms of profit New York and California is almost the same

3) Here we can conclude that New York is a more profitable state with respect to its Sales.

4) He the most profitable item is copiers with respect to its sales although the sales for chairs are substantially high, but profit isn’t.

5) There are some items that are good in sales but do not produce any profit.