

Exploratory Data Analysis - Retail (Level - Beginner)

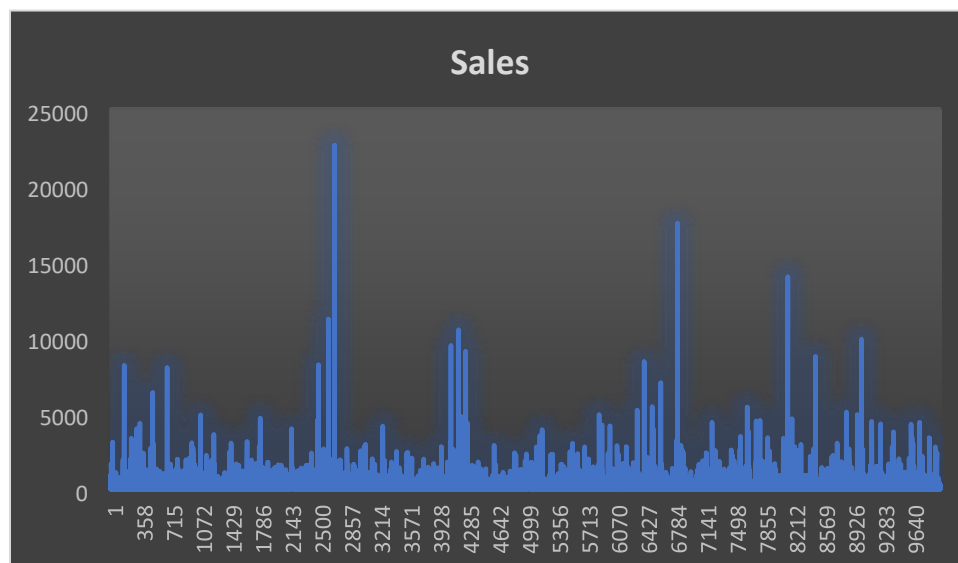
Task: Perform 'Exploratory Data Analysis' on dataset 'SampleSuperstore'

Exploratory data analysis for the dataset SampleSuperstore is done using Microsoft Excel

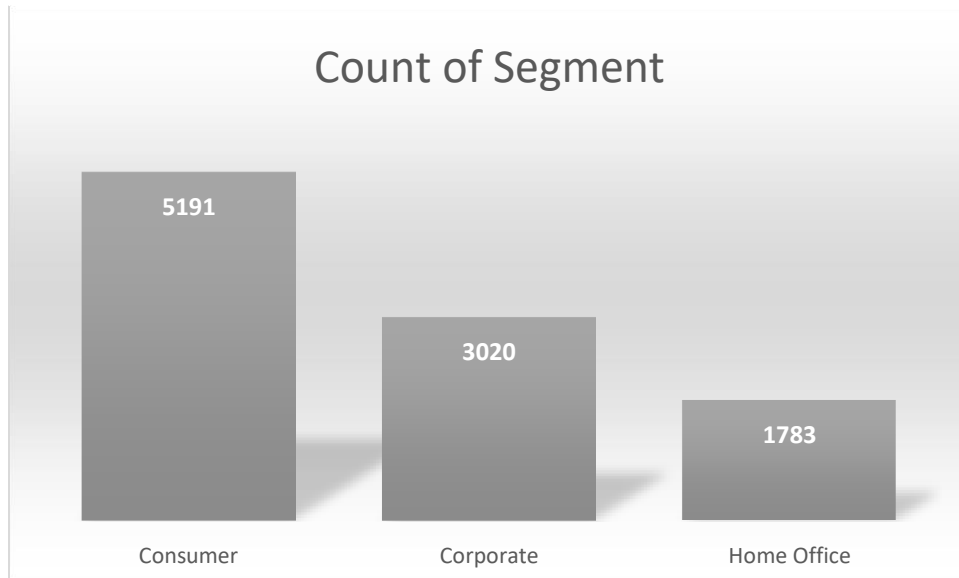
- The bar chart for the distribution of profit can be visualized as follows.



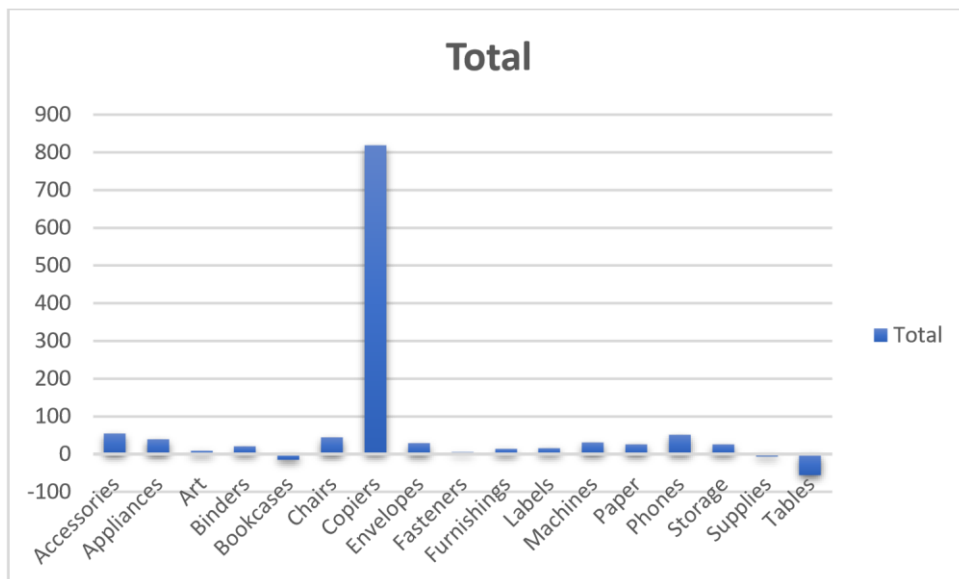
- The bar chart for the distribution of sales can be visualized as follows.



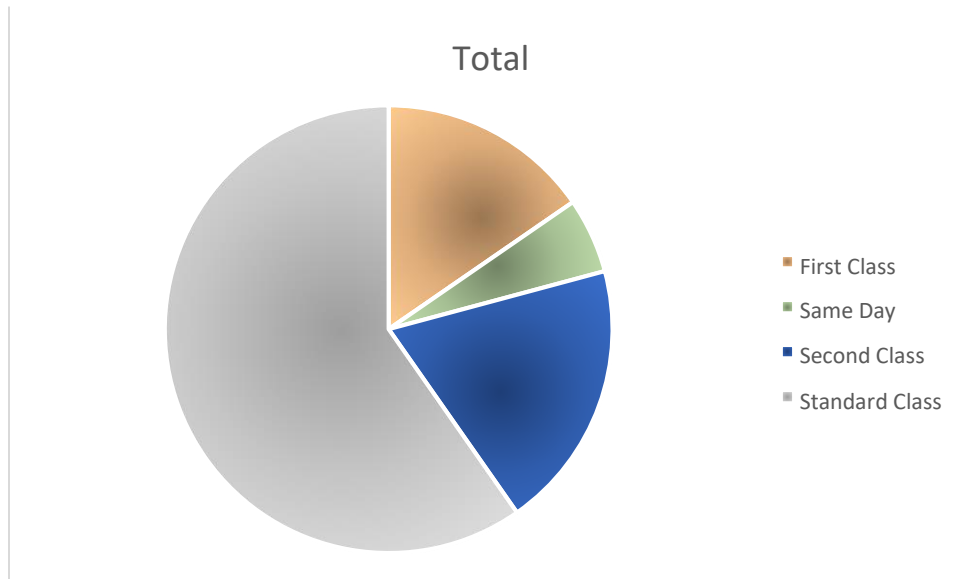
- The frequency of the variable segment among the dataset can be visualized as follows



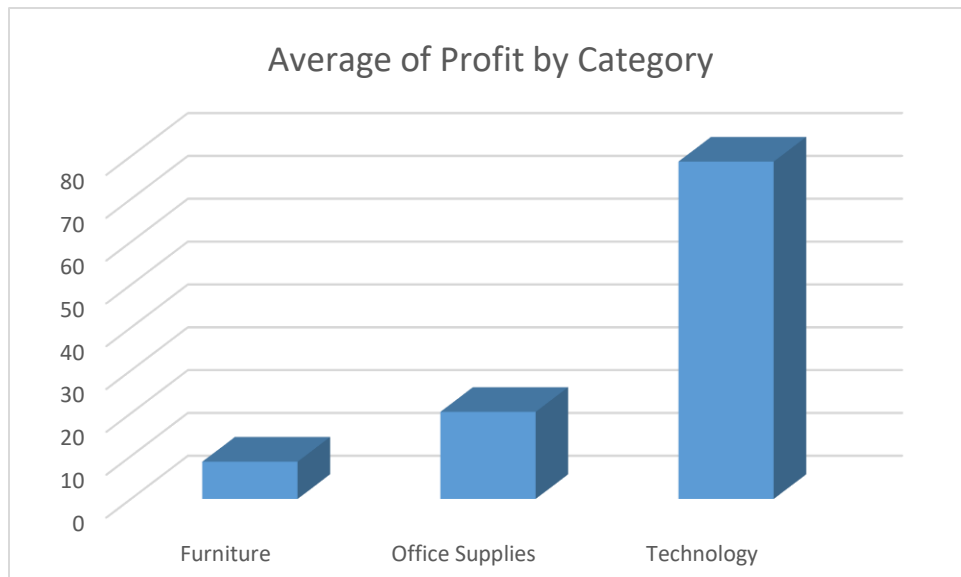
- The frequency of the variable Category can be visualized as follows



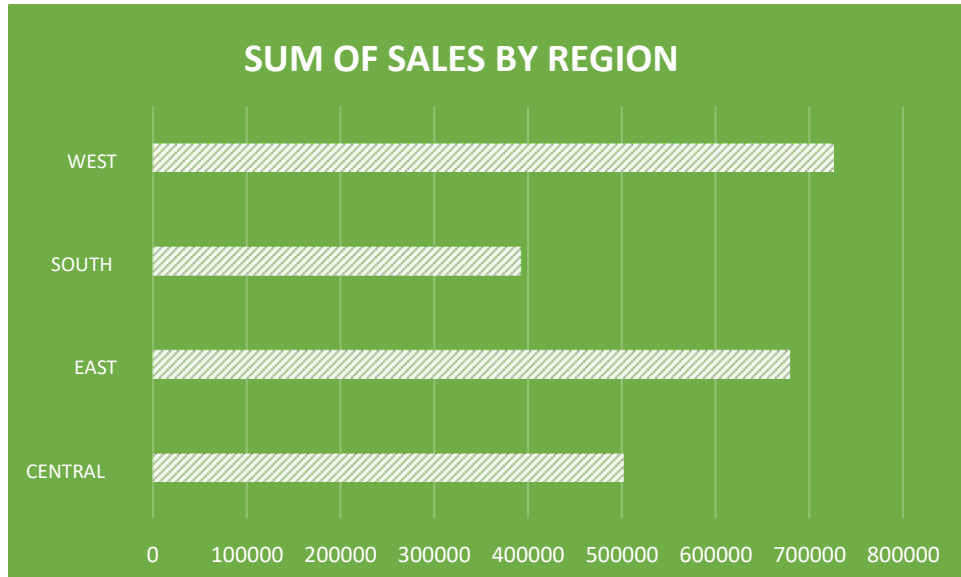
- The frequency of the variable ship mode can be visualized as follows.



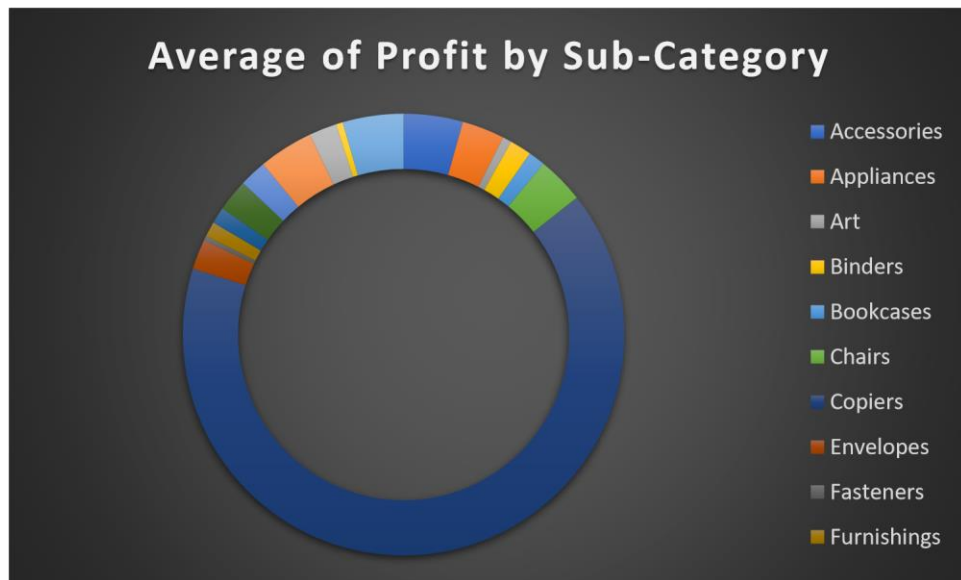
- The average of profit by category can be visualized as follows.



- The sum of sales by region can be visualized as follows



- The average of profit distributed among the different Sub-category can be visualized as follows.



Pivot table and pivot chart

We can use pivot table and pivot chart to visualize and compare more variables.

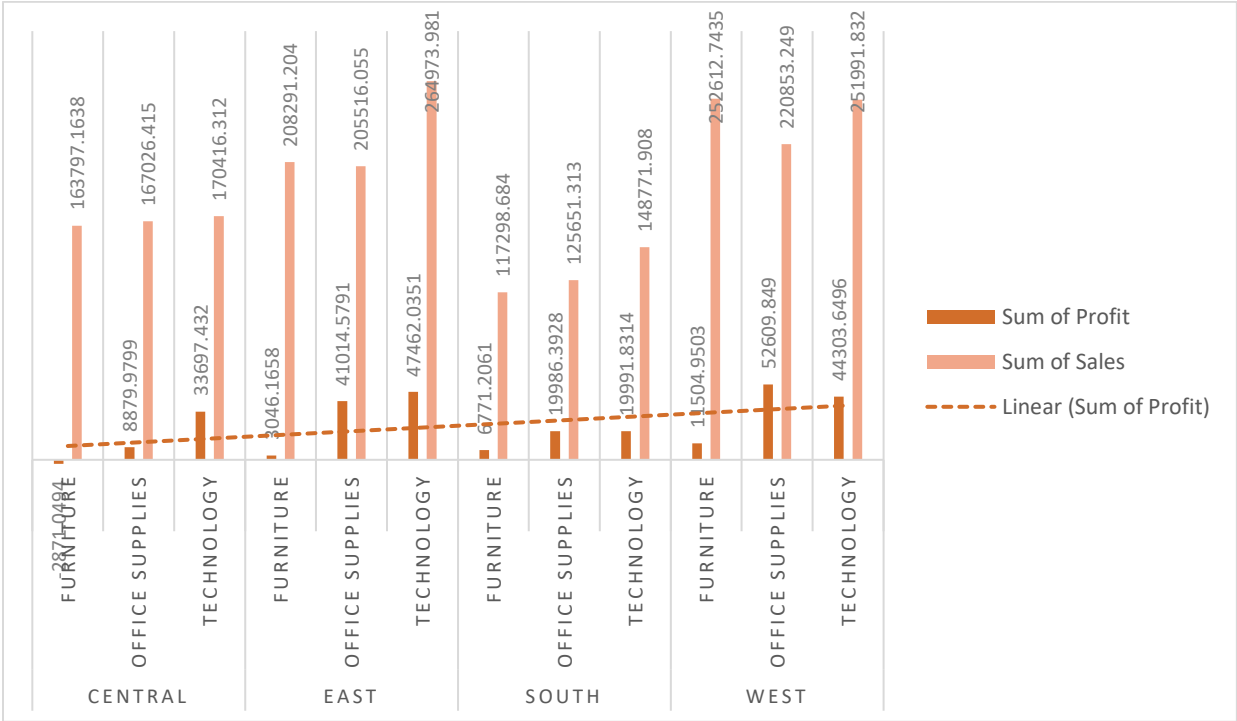
- Here is the cross tabulation of ship mode, quantity, sales , profit and discount

Row Labels	Sum of Quantity	Sum of Sales	Sum of Profit	Sum of Discount
First Class	5693	351428.4229	48969.8399	253.17
Same Day	1960	128363.125	15891.7589	82.75
Second Class	7423	459193.5694	57446.6354	270.15
Standard Class	22797	1358215.743	164088.7875	955.02
Grand Total	37873	2297200.86	286397.0217	1561.09



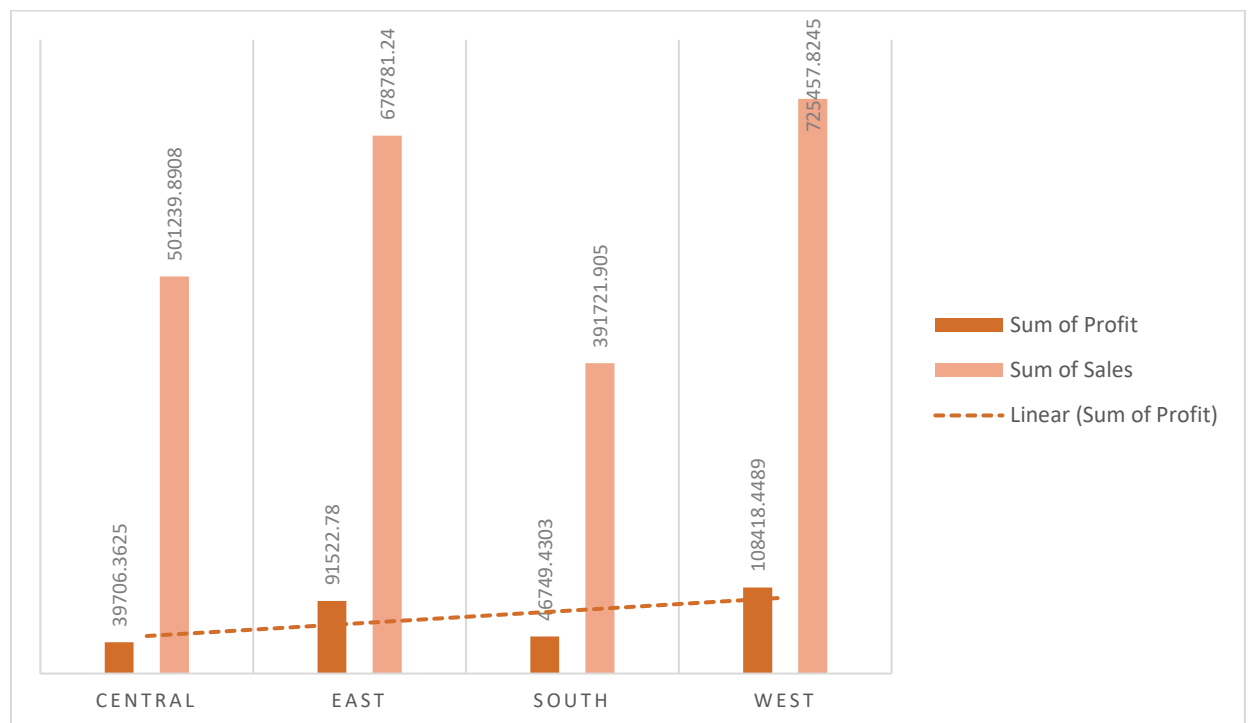
- Here is the cross tabulation of region, category, sales and profit.

Row Labels	Sum of Profit	Sum of Sales
Central	39706.3625	501239.8908
Furniture	-2871.0494	163797.1638
Office Supplies	8879.9799	167026.415
Technology	33697.432	170416.312
East	91522.78	678781.24
Furniture	3046.1658	208291.204
Office Supplies	41014.5791	205516.055
Technology	47462.0351	264973.981
South	46749.4303	391721.905
Furniture	6771.2061	117298.684
Office Supplies	19986.3928	125651.313
Technology	19991.8314	148771.908
West	108418.4489	725457.8245
Furniture	11504.9503	252612.7435
Office Supplies	52609.849	220853.249
Technology	44303.6496	251991.832
Grand Total	286397.0217	2297200.86



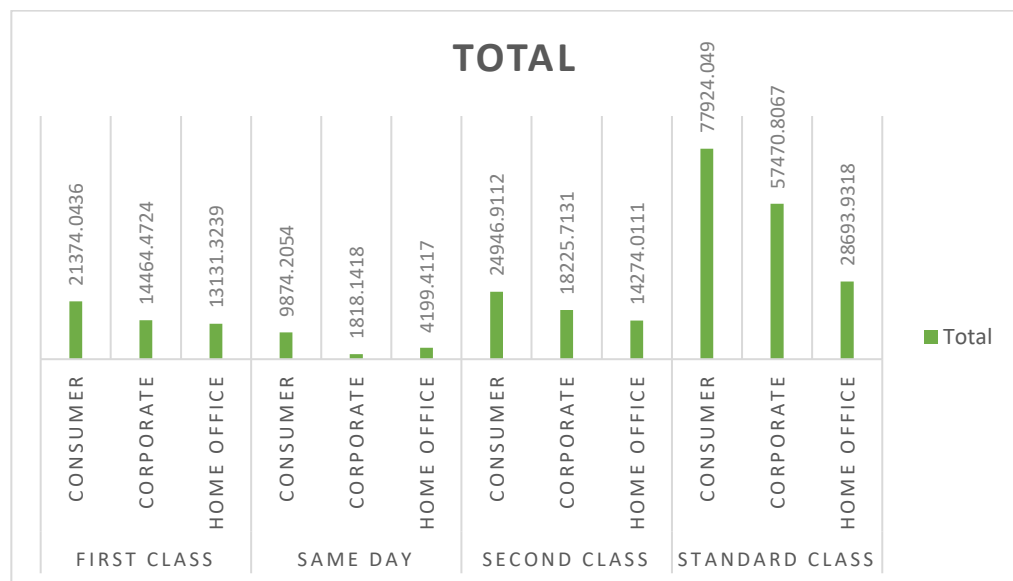
Here is the simple visualization of the above pivot chart leaving out the category

Row Labels	Sum of	
	Profit	Sum of Sales
Central	39706.3625	501239.8908
East	91522.78	678781.24
South	46749.4303	391721.905
West	108418.4489	725457.8245
Grand Total	286397.0217	2297200.86



- Here is the pivot table of segment, ship mode and profit

Row Labels	Sum of Profit
First Class	48969.8399
Consumer	21374.0436
Corporate	14464.4724
Home Office	13131.3239
Same Day	15891.7589
Consumer	9874.2054
Corporate	1818.1418
Home Office	4199.4117
Second Class	57446.6354
Consumer	24946.9112
Corporate	18225.7131
Home Office	14274.0111
Standard Class	164088.7875
Consumer	77924.049
Corporate	57470.8067
Home Office	28693.9318
Grand Total	286397.0217



Here is the short summary of the above table which shows only segment and profit.

Row Labels	Sum of Profit
First Class	48969.8399
Same Day	15891.7589
Second Class	57446.6354
Standard Class	164088.7875
Grand Total	286397.0217

