**DESCRIPTIVE ANALYSIS FOR NUMERICAL COLUMNS**

Step 1 importing pandas.

Step 2: loading data set using pandas.

Step 3: identifying numerical columns in the data set.

Step 4: calculating mean, median, mode and standard deviation for numerical columns.

Step 4 a: Calculating metrics for Volume column.

Step 4 b: Calculating metrics for Avg Price.

Step 4 c: Calculating metrics for Total Sales Value.

Step 4 d: Calculating metrics for Discount Rate.

Step 4 e: Calculating metrics for Discount Amount.

Step 4 f: Calculating metrics for Total Sales Value Column.

**RESULTS:**

1.

Mean of the Volume column: 5.066666666666666

Median of the Volume column: 4.0

Mode of the Volume column: 0 3

Name: Volume, dtype: int64

Standard deviation of the Volume column: 4.231602391213926

2.

Mean of the Avg Price column: 10453.433333333332

Median of the Avg Price column: 1450.0

Mode of the Avg Price column: 0 400

1 450

2 500

3 1300

4 8100

Name: Avg Price, dtype: int64

Standard deviation of the Avg Price column: 18079.90483993645

3.

Mean of the Net Sales Value column: 30466.336131270138

Median of the Net Sales Value column: 4677.788059301853

Mode of the Net Sales Value column: 0 326.974801

1 330.822058

2 466.350839

3 485.317265

4 496.607104

...

445 162603.186379

446 163915.971000

447 163967.080070

448 166263.161348

449 179507.479049

Name: Net Sales Value, Length: 450, dtype: float64

Standard deviation of the Net Sales Value Amount (%) column: 46358.656624066934.

4.

Mean of the Discount Rate (%) column: 15.155241895330914

Median of the Discount Rate (%) column: 16.57776564929597

Mode of the Discount Rate (%) column: 0 5.007822

1 5.055218

2 5.059801

3 5.072124

4 5.084108

...

445 19.940610

446 19.965340

447 19.976879

448 19.982257

449 19.992407

Name: Discount Rate (%), Length: 450, dtype: float64

Standard deviation of the Discount Rate (%) column: 4.220602153325511

5.

Mean of the Discount Amount (%) column: 3346.499424342415

Median of the Discount Amount (%) column: 988.9337332382784

Mode of the Discount Amount (%) column: 0 69.177942

1 73.025199

2 93.649161

3 94.682735

4 102.705757

...

445 17900.983733

446 21153.498820

447 21496.675367

448 25328.224204

449 25738.022194

Name: Discount Amount, Length: 450, dtype: float64

Standard deviation of the Discount Amount (%) column: 4509.9029633539985

6.

Mean of the Total Sales Value column: 33812.83555555555

Median of the Total Sales Value column: 5700.0

Mode of the Total Sales Value column: 0 24300

Name: Total Sales Value, dtype: int64.

Standard deviation of the Total Sales Value Amount (%) column: 50535.07417255328.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NAME OF THE COLUMN** | **MEAN** | **MEDIAN** | **MODE** | **STANDARD DEVIATION** |
| Volume | 5.06 | 4.0 | 3 | 4.23 |
| Avg Price | 10453.43 | 1450 | 400,450,500,1300,8100  (repeated same number of times) | 18079.9 |
| Total Sales Value | 33812.835 | 5700.0 | 24300 | 50535.07 |
| Discount Rate | 15.155 | 16.578 | Can’t be determined (all the data points are unique) | 4.22060 |
| Discount Amount | 3346.495 | 988.934 | Can’t be determined (all the data points are unique) | 4509.903 |
| Net Sales Value | 30466.34 | 4677.79 | Can’t be determined (all the data points are unique) | 46358.66 |

INTERPRETATION:

The numerical columns are showing mean, median and mode values with respective data and from the results we can observe that the floating-point numerical columns such as Discount Rate and Discount Amount doesn’t contain the most repeated value since all the values are unique the mode is not containing a single value.  
(#since the Discount Rate column consists of floating data type values and continuous data, we can't able to find out the most repeated value particularly.)

**DATA VISUALIZATION**

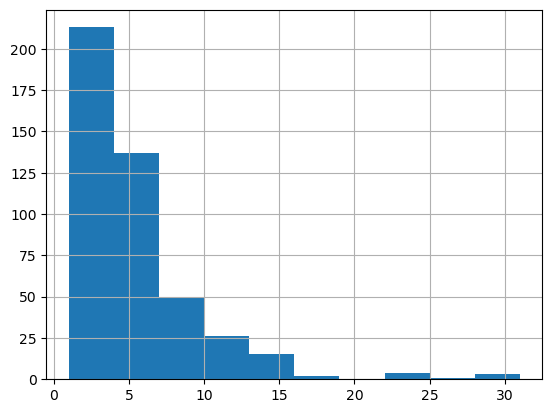
From the data set information ,there are 3 integer data types and 3 float data types there are totally 6 numerical columns

Identifying numerical columns

Plotting histograms

Step 1: histograms, skewnwss, kurtosis...

**Step 1 a. Performing histogram on Volume column and observing skewness and outliers.**



**Result:**

**Skewness:**

Skewness indicates the asymmetry of the distribution in a data set.

Here skewness is 2.73, we can get the inference that distribution is positively

skewed. In a positively skewed distribution,Mean  >  Median  >  Mode.

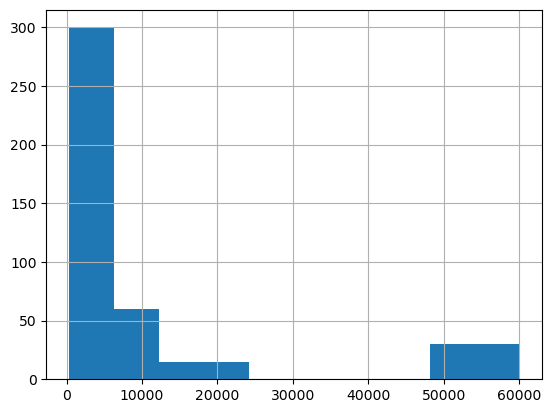
**Kurtosis:**

In general it describes the shape of the probability distribution curve. Here the

kurtosis value is 10.258, since the value is very high it indicates that there are

more outliers that is more number of tails.

**Step 1 b. performing histogram on Avg Price column and observing skewness and outliers**



**Result:**

**Skewness:**

Here the skewness is 1.908 it indicates that the distribution of the data is

positively skewed, means the tail on the right side of the distribution is longer

than the left side. In other words we can infer that there are more extreme

values on the right side of the distribution compared to the left side.

**Kurtosis:**

Here the kurtosis value is 2.07565 it is indicating that the distribution of the

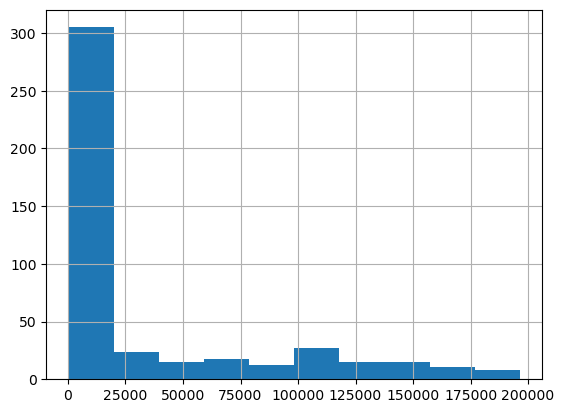
data has a higher peak and heavier tails than a normal distribution (k=3). This

means that the data is a leptokurtic curve in which , it is having more extreme

values and also more peaked than a normal distribution.

**Step 1 c.**

**Performing histogram on Total Sales Value column and observing skewness and outliers**



**Result:**

**Skewness:**

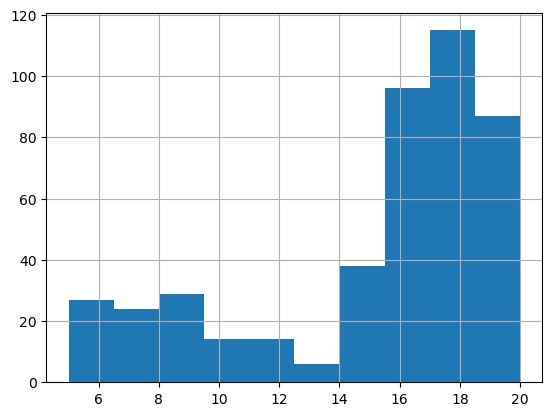
skewness value of 1.53 indicates that it is will be a positive skewness in the data distribution. The distribution is skewed towards the right, means it is having a longer tail on the right side and is concentrated towards the left side. The skewness is suggesting that the presence of outliers or extreme values on the right side of the distribution.

**Kurtosis:**

Here the kurtosis value is 1.025 and it is suggesting that the distribution of the data is close to a normal distribution, or mesokurtic. In other words ,the distribution is similar to a normal distribution in terms of its peakedness and the thickness. We can conclude that it's neither particularly peaked (leptokurtic) nor particularly flat (platykurtic).

**Step 1 d.**

**performing histogram on Discount Rate column and observing skewness and outliers**



**Result:**

**Skewness:**

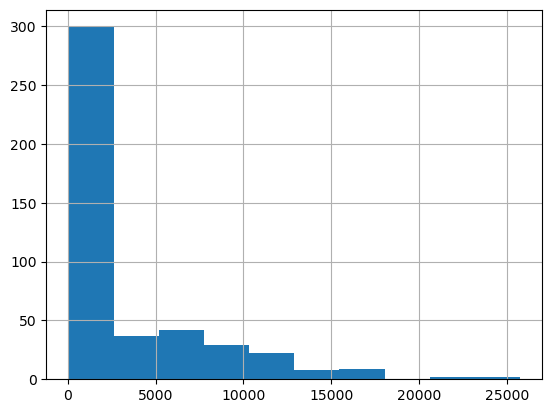
Here the skewness is -1.0623 is indicating a strong negative skewness in the data distribution. The distribution is significantly skewed towards left, in other words it is having a longer tail on the left side and is denser towards the right side.

**Kurtosis:**

Here the kurtosis value obtained is -0.1785 and it is suggesting that the distribution is platykurtic, means it is having thinner tails than compared a normal distribution. In other words, the tails of the distribution are lighter than the tails of a normal distribution.Also indicates that the distribution has lesser extreme values or outliers than a normal distribution.

**Step 1 e.**

**Performing histogram on Discount Amount column and observing skewness and outliers**



**Result:**

**Skewness:**

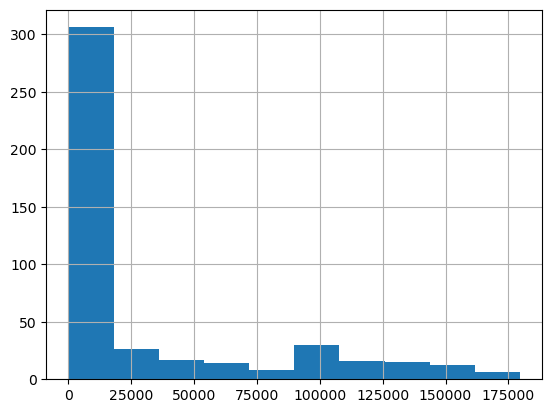
Here the skewness is 1.913 means the distribution is positively skewed. In a positively skewed distribution, the tail will be on the right side (the side which it contains larger data values) is longer than the other side (left side). In other words it is suggesting that there are more extreme values on the right side of the distribution.

**Kurtosis:**

The kurtosis value is 3.83 indicating that the distribution is leptokurtic. Leptokurtic distributions will be having heavier tails and a higher peak than compared to a normal distribution. In other words, there is an increment of extreme values or outliers in the distribution, and in resulting of a sharp and heavier tails.

**Step 1 f.**

**Performing histogram on Net Sales Value column and observing skewness and outliers.**



**Result:**

**Skewness:**

Here the skewness of 1.54 is indicating that the distribution is positively skewed. Positively skewed distributions are having a longer tail on the right side than compared to the left side. Also it is suggesting that there are more extreme values on the right side of the distribution.

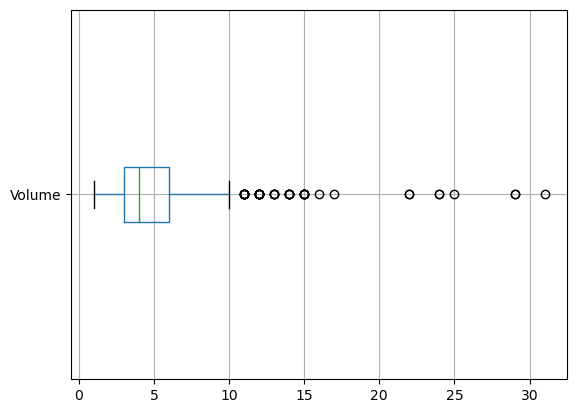
**Kurtosis:**

Here the kurtosis value is 1.012 suggesting that the distribution is mesokurtic. Mesokurtic distributions have kurtosis close to that of a normal distribution (where kurtosis = 3). In other words the distribution contains tails and a peak similar to normal distribution Mesokurtic distributions contains a bell-shaped curve that is similar to a normal distribution.

**Step 2:box plots for each of the numerical columns.**

**Step 2 a.**

**Performing Box plot on Volume column and observing outliers and inter quartile ranges.**



**Observations:**

1.Most of the data points lies on the left side.

2.The box plot is present over left side therefore the assumption is most of the data points lies on left side and thus it is positively skewed.

3.25th percentile is 3.0.

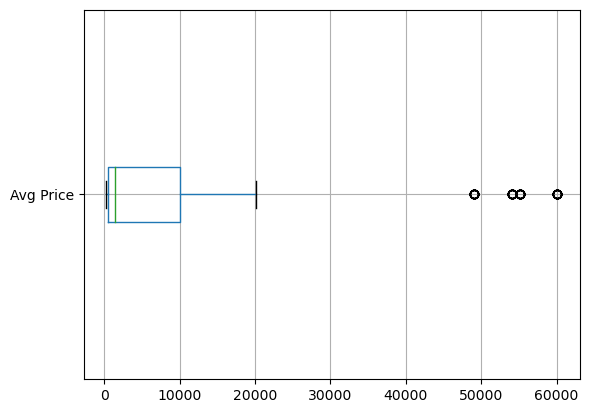
4.75th percentile that is Q3=6.0

5.Inter quartile range is 3.0.

6.Outliers are lying above upper whisker length.

**Step 2 b.**

**Performing Box plot on Avg Price column and observing outliers and inter quartile ranges.**

****

**Observations:**

1.Most of the data points lies on the left side.

2.The box plot is present over left side therefore the assumption is most of the data points lies on left side and thus it is positively skewed.

3.25th percentile is 465.

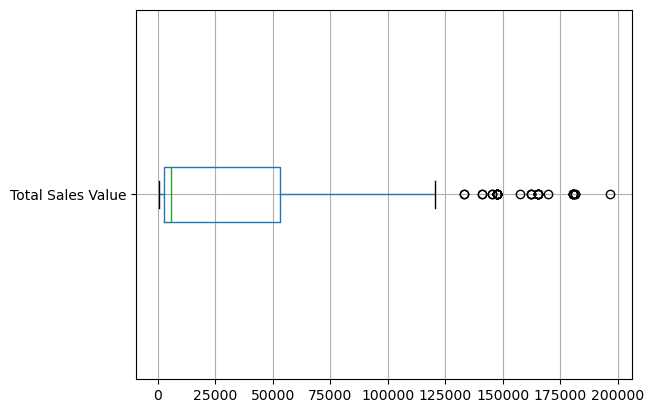
4.75th percentile that is Q3=10100

5.Inter quartile range is 9635.0.

6.Outliers are lying above upper whisker length.

**Step 2 c.**

**Performing Box plot on Total Sales Value column and observing outliers and inter quartile ranges.**



**Observations:**

1.Most of the data points lies on the left side.

2.The box plot is present over left side therefore the assumption is most of the data points lies on left side and thus it is positively skewed.

3.25th percentile is 2700.

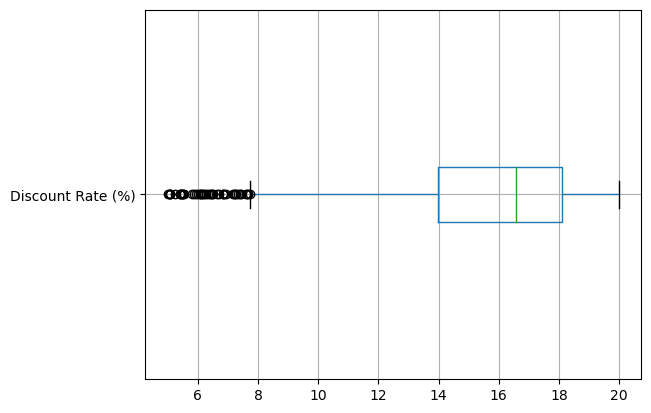
4.75th percentile that is Q3=53200

5.Inter quartile range is 50500.0.

6.Outliers are lying above upper whisker length.

**Step 2 d.**

**Performing Box plot on Discount Rate (%) column and observing outliers and inter quartile ranges**



**Observations:**

1.Most of the data points lies on the right side.

2.The box plot is present over right side therefore the assumption is most of the data points lies on right side and thus it is negatively skewed.

3.25th percentile is 13.965.

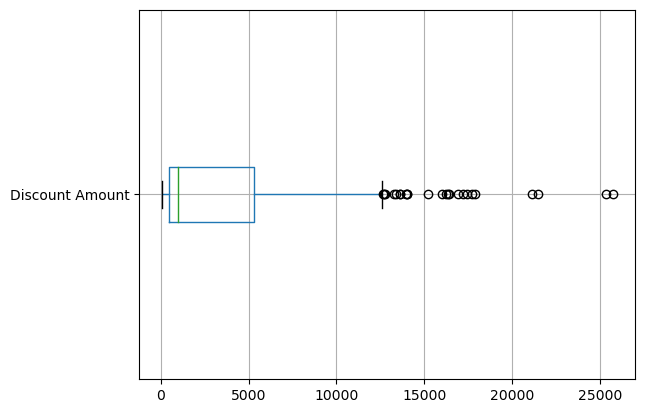
4.75th percentile that is Q3=18.115.

5.Inter quartile range is 4.15.

6.Outliers are lying less than(before) lower whisker length.

**Step 2 e.**

**Performing Box plot on Discount Amount column and observing outliers and inter quartile ranges.**



**Observations:**

1.Most of the data points lies on the left side.

2.The box plot is present over left side therefore the assumption is most of the data points lies on left side and thus it is positively skewed.

3.25th percentile is 460.5.

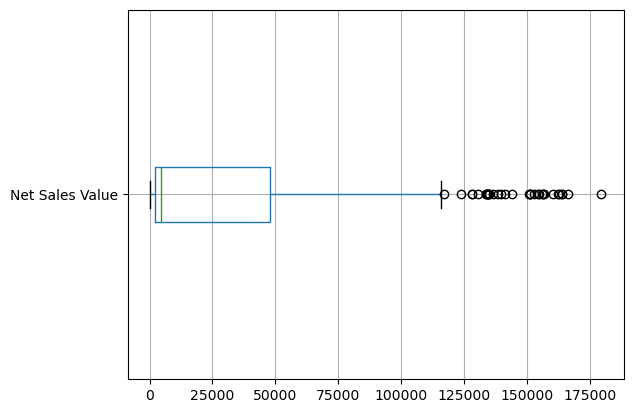
4.75th percentile that is Q3=5316.5.

5.Inter quartile range is 4856.04.

6.Outliers are lying above upper whisker length.

**Step 2 f.**

**Performing Box plot on Net Sales Value column and observing outliers and inter quartile ranges.**



**Observations:**

1.Most of the data points lies on the left side.

2.The box plot is present over left side therefore the assumption is most of the data points lies on left side and thus it is positively skewed.

3.25th percentile is 2202.21

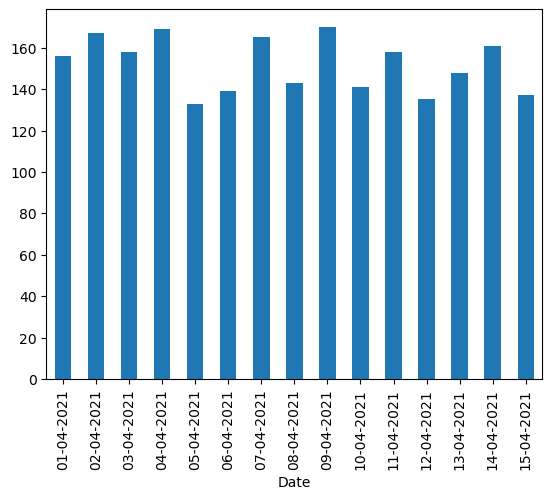
4.75th percentile that is Q3=47847.9

5.Inter quartile range is 45645.7.

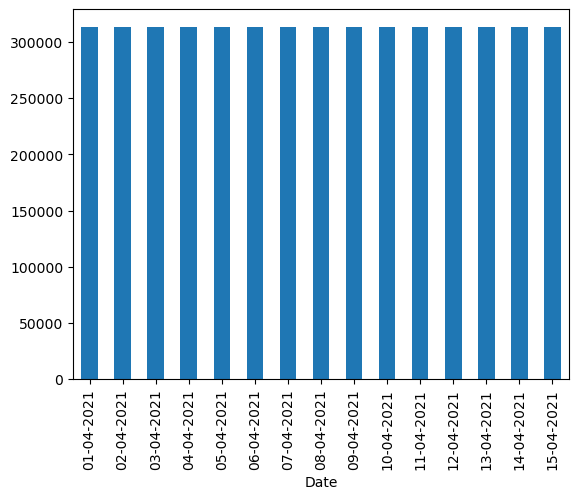
6.Outliers are lying above upper whisker length.

**Step 3: Bar Chart Analysis.**

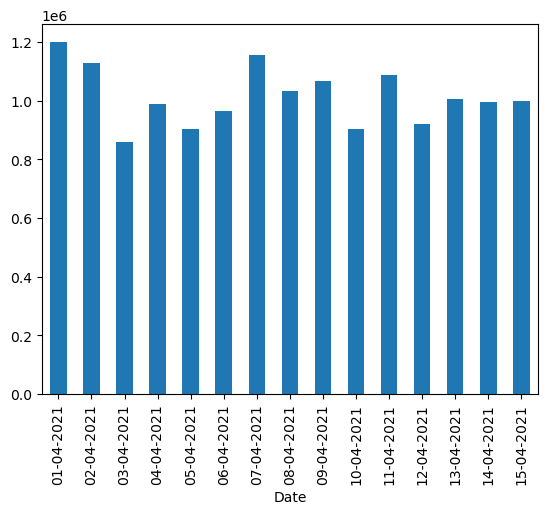
**Step 3 a: Bar Chart for Volume column.**



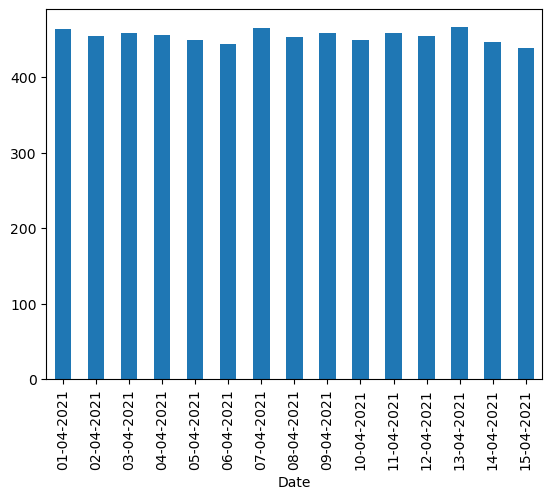
**Step 3 b: Building Bar Graph for "Avg Price" column.**



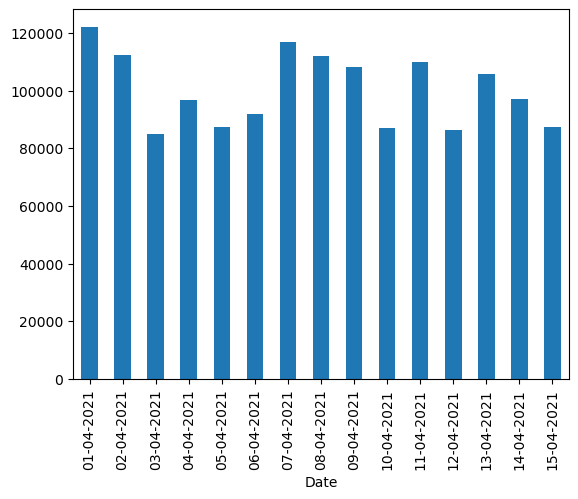
**Step 3 c: Building Bar Graph for "Total Sales Value" column**



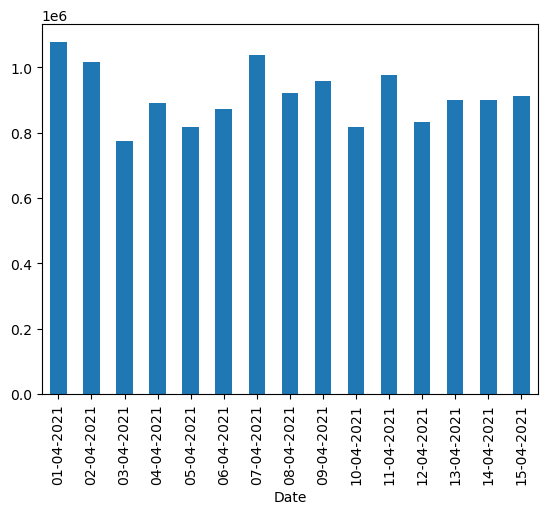
**Step 3 d: Building Bar Graph for "Discount Rate (%)" column**



**Step 3 e: Building Bar Graph for "Discount Amount" column**



**Step 3 f: Building Bar Graph for "Net Sales Value" column**

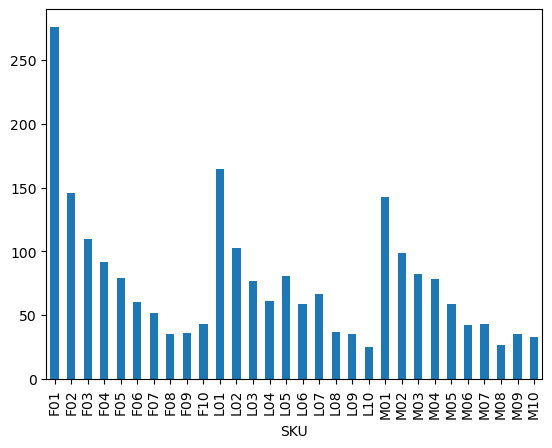


**Step 4: Bar Chart Analysis with categorical columns**

**Categorical Columns identified are "SKU","BU","Brand","Model".**

**step 4 a, Performing Bar Graph of “Volume” with categorical columns.**

**Step 4 a,1, “Volume” Numerical Column with ‘SKU’ Categorical Column.**

******

***Insights:***

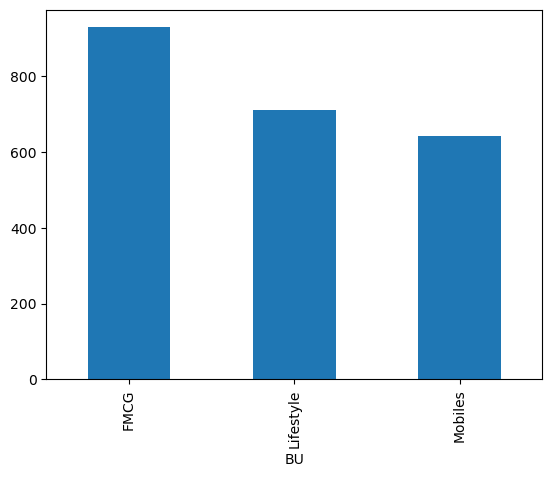
We can observe,

Highest frequency columns are F01,L01,M01.

Lowest frequency columns are F08,L10,M08.

**step 4,a,2**

**“Volume” Numerical Column with “BU” Categorical Column**



***Insights:***

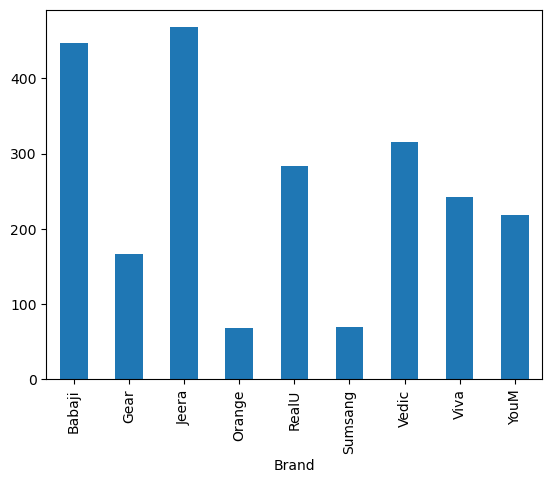
We can observe,

Highest frequency columns are FMCG.

Lowest frequency columns are Mobiles.

**Step 4,a,3**

**“Volume” Numerical Column with “Brand” Categorical Column.**

******

***Insights:***

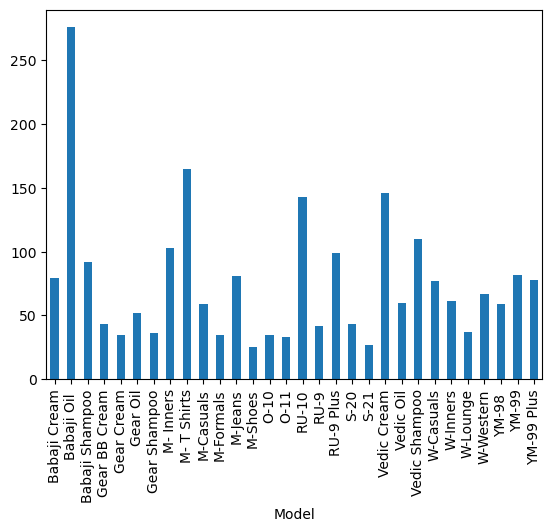
We can observe,

Highest frequency columns are Babaji and Jeera.

Lowest frequency columns are Orange and Samsung.

**step 4,a,4**

**“Volume” Numerical Column with “Model” Categorical Column.**



***Insights:***

We can observe,

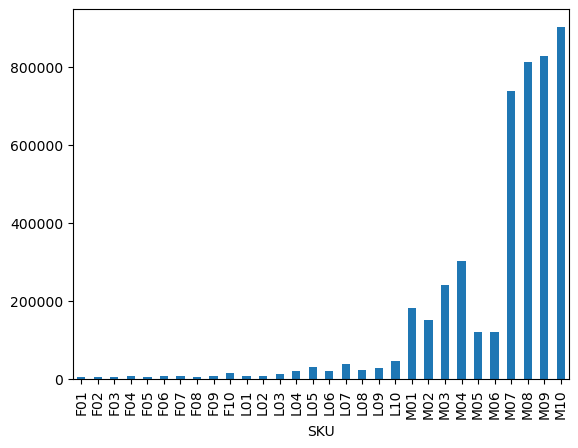
Highest frequency columns are Babaji Oil.

Lowest frequency columns are M-Shoes and S-21.

**Step 4 ,b**

**Performing Bar Graph of Avg Price with categorical columns**

**Step 4,b,1 “Avg Price” Numerical Column with “SKU” Categorical Column.**



***Insights:***

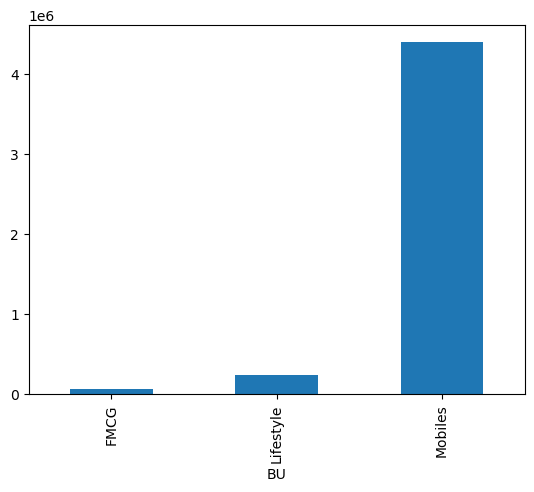
We can observe,

Highest frequency columns are MO10,MO9 and MO8.

Lowest frequency columns can be observed from FO1 to L10(approximately).

**Step 4,b,2**

**“Avg Price” Numerical Column with “BU” Categorical Column.**



***Insights:***

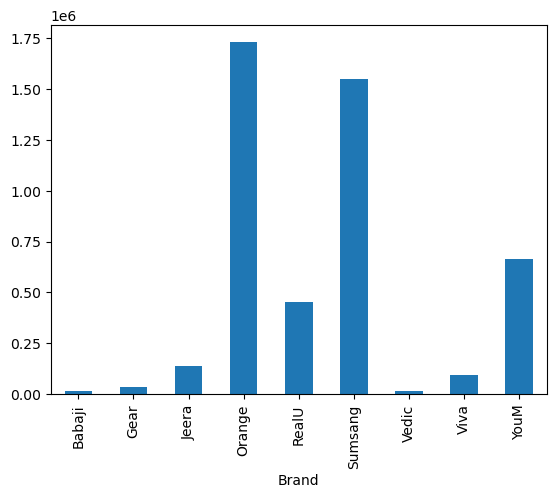
We can observe,

Highest frequency column is Mobiles.

Lowest frequency columns can be observed from FMCG.

**Step 4,b,3**

**“Avg Price” Numerical Column with “Brand” Categorical Column.**



.

Insights:

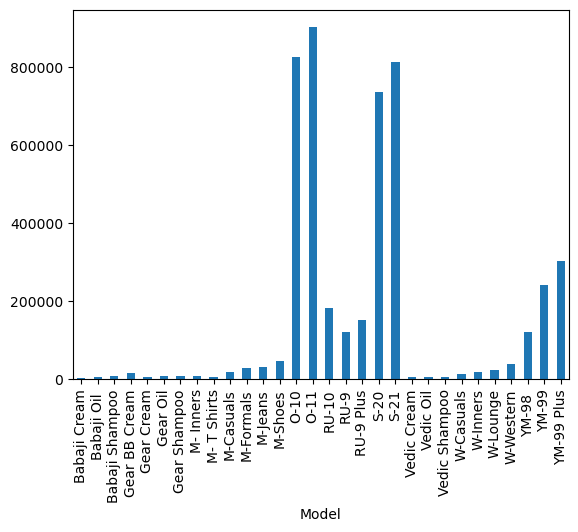
We can observe,

Highest frequency columns are Samsung and Orange.

Lowest frequency columns are Babaji, Vedic and Gear.

**Step 4,b,4**

**“Avg Price” Numerical Column with “Model” Categorical Column.**

****

**Insights:**

We can observe,

Highest frequency column is O11.

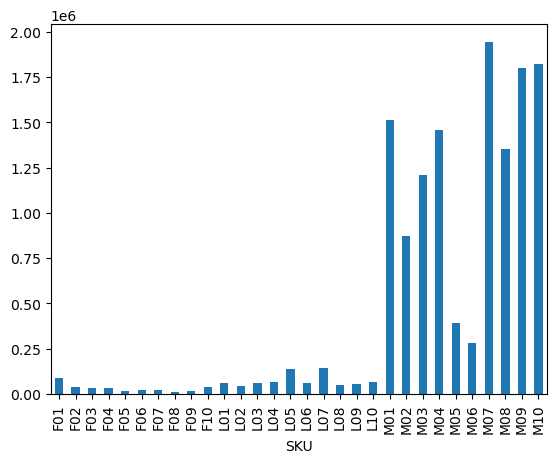
Lowest frequency columns are Babaji Cream, Vedic Cream, Gear Cream, M-T Shirts, Vedic Oil, Vedic Shampoo(approximately)…

**step 4 c,**

**Performing Bar Graph of Net Sales Value Numerical Column**

**with categorical columns.**

**step 4 c,1, “Net Sales Value” Numerical Column with “SKU” Categorical Column.**

**.**

**Insights:**

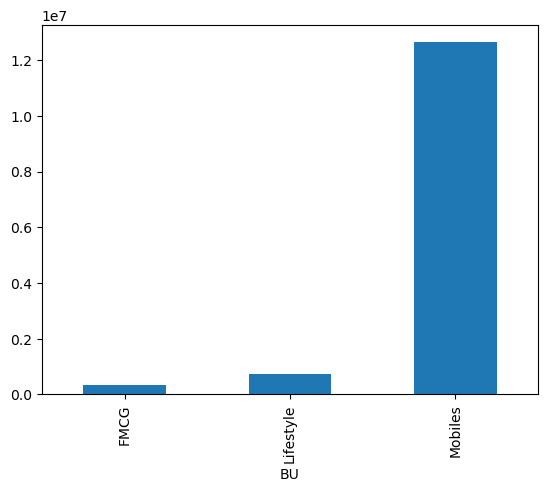
We can observe,

Highest frequency column is MO7.

Lowest frequency columns are FO8,FO9,FO5(approximately).

**Step 4,c,2**

**‘Net Sales Value’ Numerical Column with “BU” Categorical Column.**

****

**Insights:**

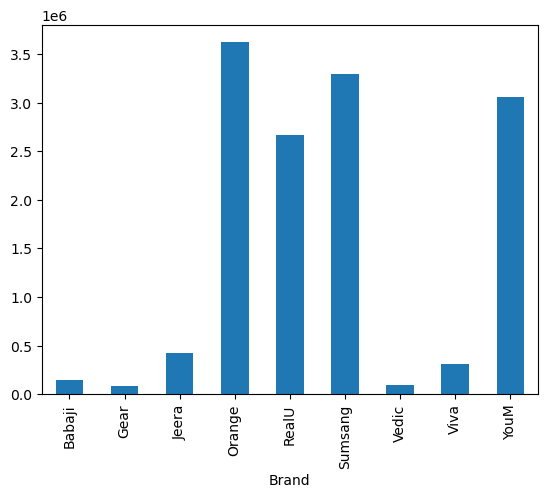
We can observe,

Highest frequency column is Mobiles.

Lowest frequency column is FMCG.

**Step 4,c,3**

**“Net Sales Value” Numerical Column with “Brand” Categorical Column.**



**Insights:**

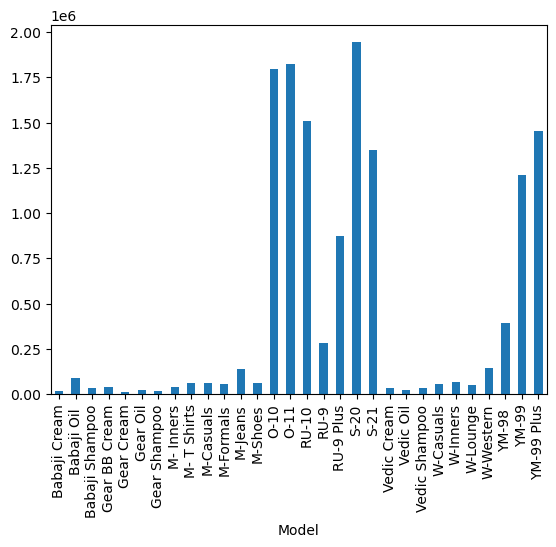
We can observe,

Highest frequency column is Orange.

Lowest frequency column is Vedic and Gear(approx.).

**Step 4,c,4**

**“Net Sales Value” Numerical Column with “Model” Categorical Column.**



**Insights:**

We can observe,

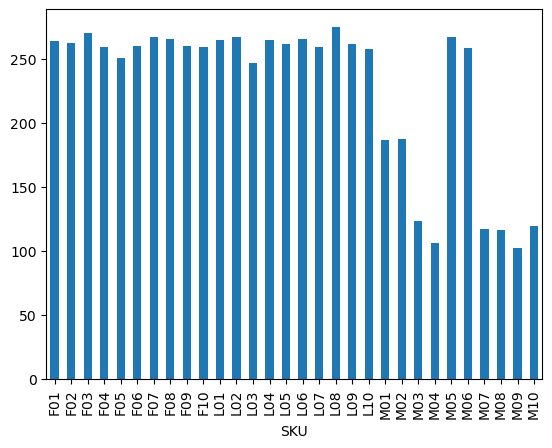
Highest frequency column is S-20.

Lowest frequency columns are Babaji Cream, Gear Cream and Gear Shampoo (approx.).

**step 4 d,**

**Performing Bar Graph of Discount Rate Numerical Column with with categorical columns.**

**step 4 d,1, “Discount Rate” Numerical Column with “SKU” Categorical Column.**



**Insights:**

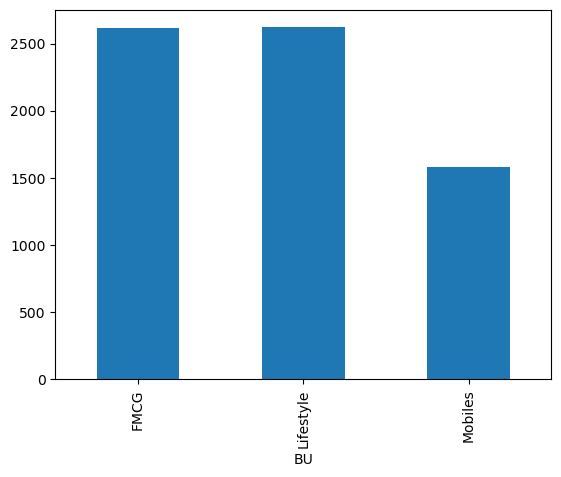
We can observe,

Highest frequency column is LO8.

Lowest frequency columns is MO4.

**Step 4,d,2**

**“Discount Rate (%)” Numerical Column with “BU” Categorical Column.**



**Insights:**

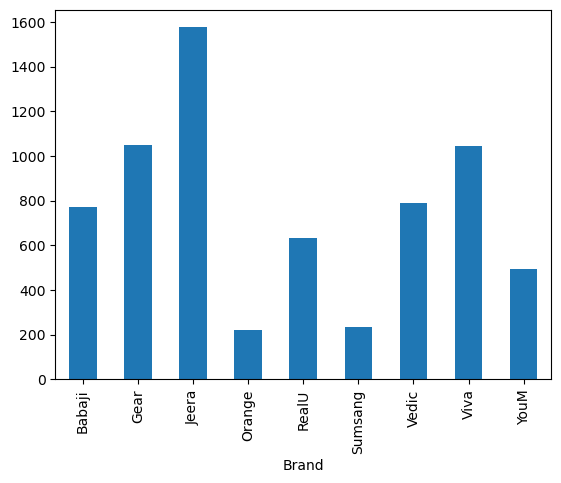
We can observe,

Highest frequency columns are FMCG and Life Style.

Lowest frequency columns is Mobiles.

**Step 4,d,3**

**“Discount Rate (%)” Numerical Column with “Brand” Categorical Column.**



**Insights:**

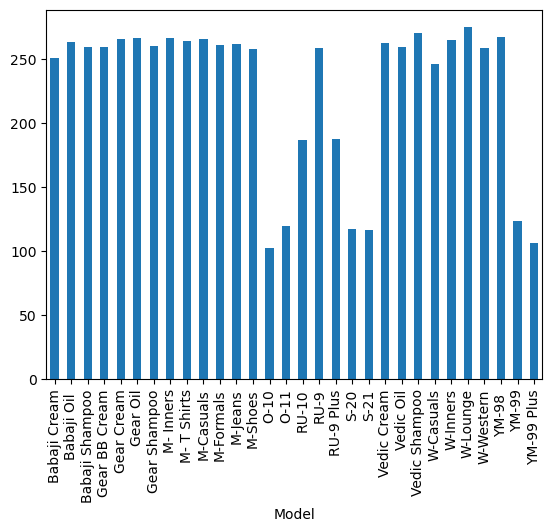
We can observe,

Highest frequency column is Jeera.

Lowest frequency column is Orange and Samsung.

**Step 4,d,4**

**“Discount Rate” Numerical Column with “Model” Categorical Column.**

****

**Insights:**

We can observe,

Highest frequency column is W-Lounge.

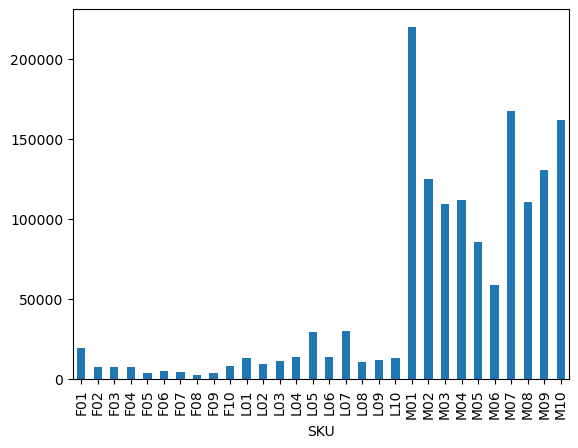
Lowest frequency column is O-10.

**Step 4 e,**

**Performing Bar Graph of Discount Amount Numerical Column.**

**with categorical columns**

**step 4 e,1, “Discount Amount” Numerical Column with “SKU” Categorical Column.**

****

**Insights:**

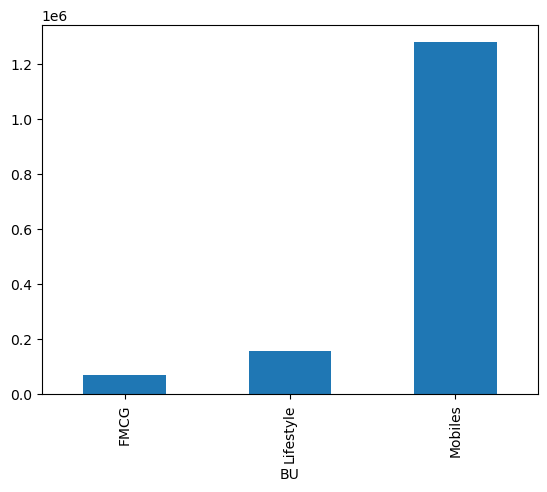
We can observe,

Highest frequency column is MO1.

Lowest frequency column is FO8.

**Step 4,e,2**

**“Discount Amount” Numerical Column with “BU” Categorical Column.**



**Insights:**

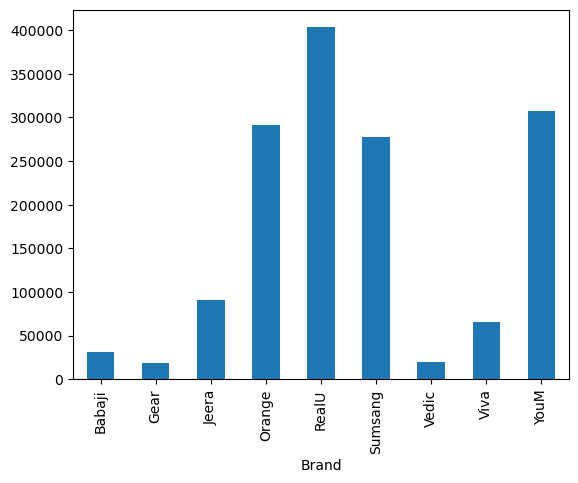
We can observe,

Highest frequency column is Mobiles.

Lowest frequency column is FMCG.

**Step 4,e,3**

**“Discount Amount” Numerical Column with “Brand” Categorical Column.**



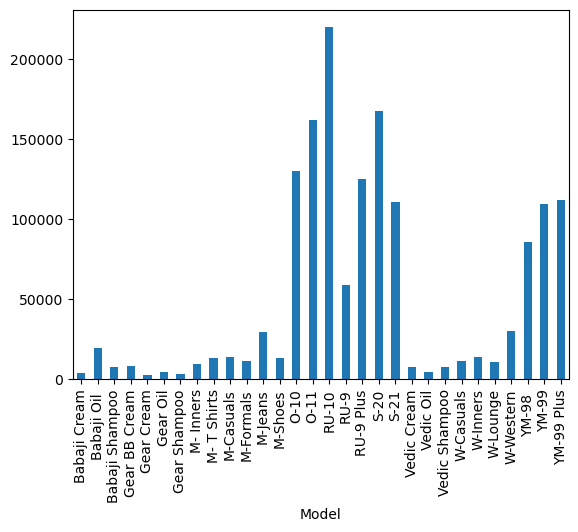
**Insights:**

We can observe,

Highest frequency column is RealU.

Lowest frequency column is Gear and Vedic(approximately).

**Step 4,e,4**

**“Discount Amount” Numerical Column with “Model” Categorical Column.**

**Insights:**

We can observe,

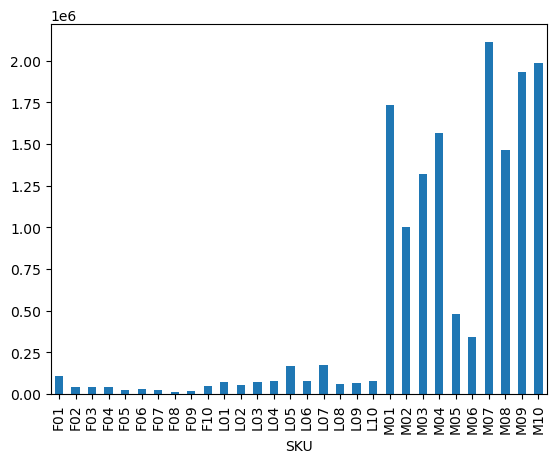
Highest frequency column is RU-10.

Lowest frequency column is Gear Cream, Gear Shampoo(approximately).

**Step 4 f,**

**Performing Bar Graph of Total Sales Value Numerical Column with categorical columns.**

**step 4 f,1, “Total Sales Value” Numerical Column with “SKU” Categorical Column.**



**Insights:**

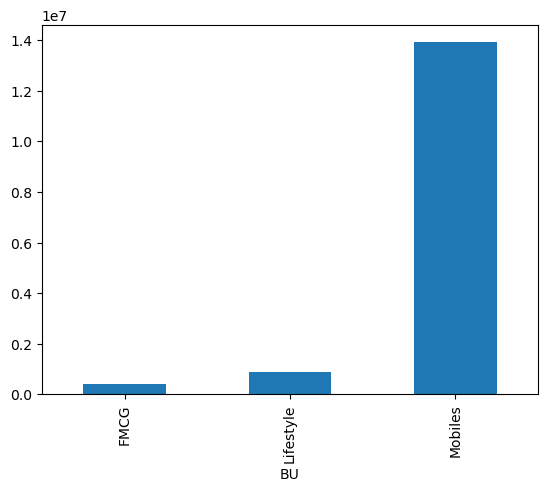
We can observe,

Highest frequency column is MO7.

Lowest frequency column is FO8,FO9(approximately).

**Step 4,f,2**

**“Total Sales Value” Numerical Column with “BU” Categorical Column.**



**Insights:**

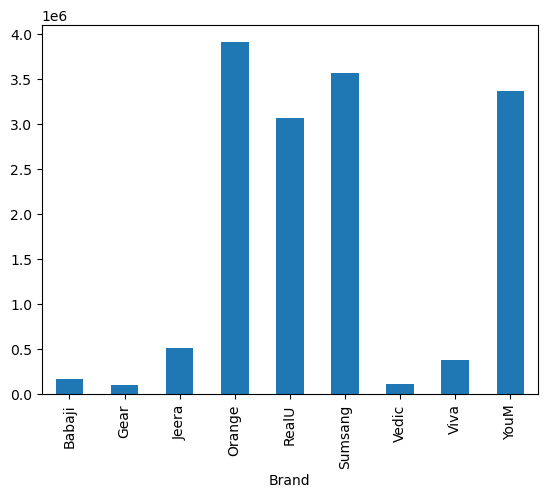
We can observe,

Highest frequency column is Models.

Lowest frequency column is FMCG(approximately).

**Step 4,f,3**

**“Total Sales Value” Numerical Column with “Brand” as Categorical Column.**



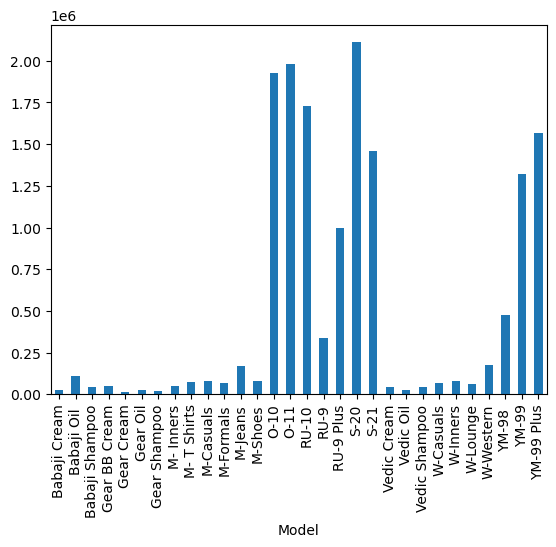
**Insights:**

We can observe,

Highest frequency column is Orange.

Lowest frequency column is Vedic and Gear(approximately).

**Step 4,f,4**

**“Total Sales Value” Numerical Column with “Model” Categorical Column.**

**Insights:**

We can observe,

Highest frequency column is S-20.

Lowest frequency columns are Gear Cream Gear Shampoo(approximately).

**STANDARDIZATION OF NUMERICAL VARIABLES**

* Importing pandas
* Importing scipy
* Step 1
* Normalization on Volume Column.
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* FINDING MIN VALUE.
* To find observations regarding distributions we are constructing histogram.
* step 2.
* Normalization on Avg Price Column.
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* To find observations regarding distributions we are constructing histogram.
* Normalization on Total Sales Value Column.
* Step 3 .
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* To find observations regarding distributions we are constructing histogram.
* Normalization on Discount Rate (%) Column.
* Step 4
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* To find observations regarding distributions we are constructing histogram
* Normalization on Discount Amount Column.
* Step 5
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* To find observations regarding distributions we are constructing histogram
* Normalization on Net Sales Value Column.
* Step 6.
* FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.
* To find observations regarding distributions we are constructing histogram.

**CONCEPT OF STANDARDIZSATION(Z-SCORE NORMALIZATION)**

Standardization, is also referred as z-score normalization, is a technique which is used machine learning also statistics for re-scaling data to have a mean of 0 and a standard deviation of 1. This process is helpful for comparing the data that have different units or scales, it also helps for machine learning model analysis. In other words, Z-Score Normalization will help the data to transform from its originality to around zero within the interval of -3 to +3 sigma.

The formula to apply standardization is ,

**Where the values obtained are called z-score,**

**x=actual data points**

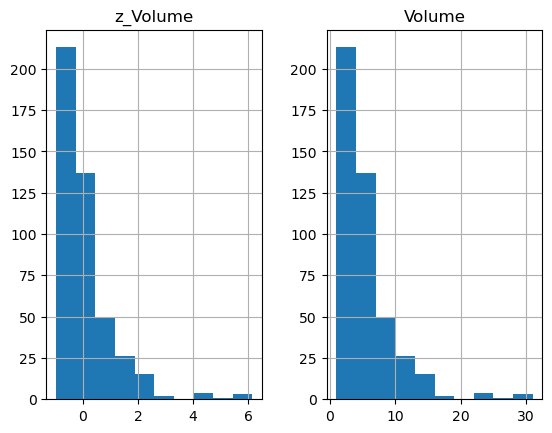
The mean (average value of the data set=**μ).**

The standard deviation is the amount of variation of the data set **(σ).**

In dealing with features of different units, then standardization will ensures that each feature contributes equally to the model or not.

**RESULTS:**

1.

****

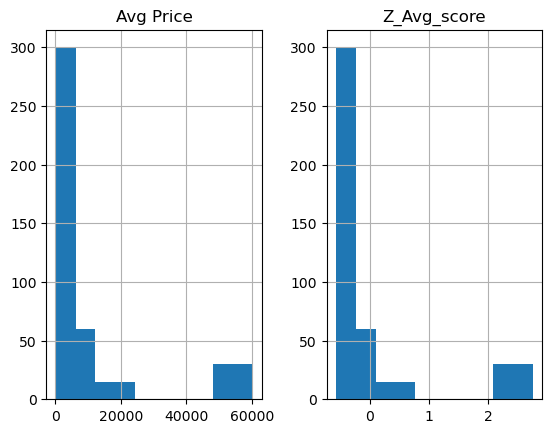
***Observations:***

*We can observe from the Z\_Volume graph that is after standardization the graph is looking like exceeding the range of 3 (that is the actual graph ratio for normal distribution is from -3 to +3 sigma) but it is exceeding more of the graph ratio(positively skewed since the outliers showing towards the right of graph).*

*So we can observe the outliers are lying above the 3 value in z-graph.*

*Outliers: It starts from value 4 in z-graph and the actual data graph where we can see the interval from around 22 to 32(approx.).*

2.



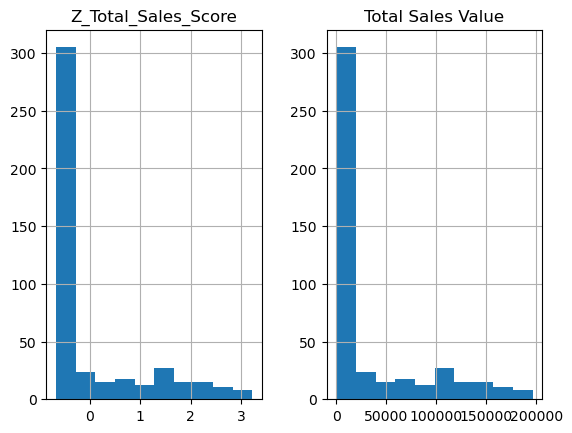
***Observations:***

*We can observe from the Z\_Avg\_score graph that is after standardization the graph is looking the range 2 - 3(that is the actual graph ratio for normal distribution is from -3 to +3 sigma) (positively skewed since the outliers showing towards the right of graph)..*

*So we can observe the outliers are lying below the 3 value in z-graph.*

*Outliers: It starts from value 2.1(approx.) in z-graph and the actual data graph where we can see the interval from around 43000-60000(approx.).*

3.



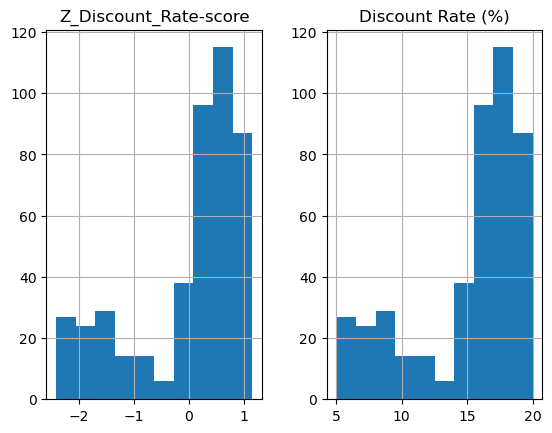
***Observations:***

*We can observe from the Z\_Total\_Sales\_score graph that is after standardization the graph is looking like the range 0- 3 (approx) (that is the actual graph ratio for normal distribution is from -3 to +3 sigma) (positively skewed since the outliers showing towards the right of graph)..*

*So we can observe the outliers are lying above 3 value also in z-graph.*

*Outliers: For finding the starting range to ending range of z-values and original data points we are going to calculate min and max values.*

*4.*



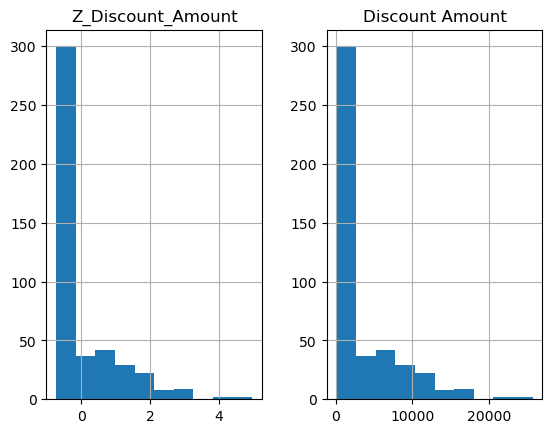
***Observations:***

*We can observe from the Z\_Discount\_Rate-score graph that is after standardization the graph is looking like the range -2.5 to above 1 (that is the actual graph ratio for normal distribution is from -3 to +3 sigma) (negatively skewed since the outliers showing towards the left of graph)..*

*So we can observe the outliers are lying below the 3 value in z-graph.*

*Starts from value -2.5(approx.) in z-graph and the actual data graph where we can see the interval from around 5-20(approx.).*

*5.*



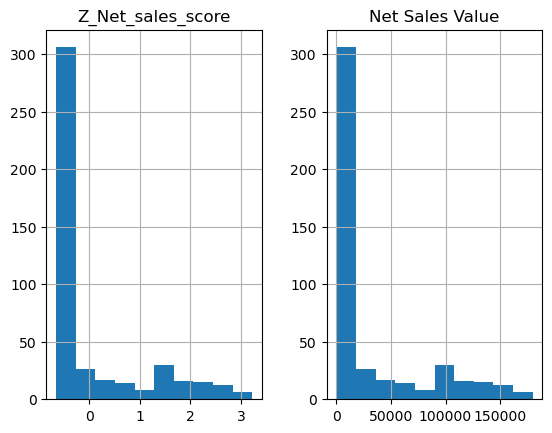
***Observations:***

*We can observe from the Z\_Discount\_Amount graph that is after standardization, the graph is looking like the range -0.5(approx.) to above 3 (that is the actual graph ratio for normal distribution is from -3 to +3 sigma) (positively skewed since the outliers showing towards the right of graph)..*

*So we can observe the outliers are lying close to value 4 in z-graph.*

*Starts from value -0.7(approx.) in z-graph and the in the actual data graph where we can see outliers the interval from above 20000(approx.).*

6.



***Observations:***

*We can observe from the Z\_Net\_Sales\_score graph that is after standardization the graph is looking like the range -0.5 to above 3 (that is the actual graph ratio for normal distribution is from -3 to +3 sigma) (positively skewed since the outliers showing towards the right of graph)..*

*So we can observe the outliers are also lying above the 3 value in z-graph.*

*Starts from value -0.7(approx.) in z-graph and the actual data graph where we can see the interval from around 0 to 17000(approx.).*

**KEYPOINTS:**

A normal distribution (or we can also call Gaussian distribution) falls under the range of +3 sigma to -3 sigma.

* Approximately 68.27% of the data falls in between the range of +1 sigma to -1 sigma from the center.
* Approximately 95.45% of the data falls in between the range of -2 sigma to +2 sigma from the center.
* Approximately 99.73% of the data falls in between the range of -3 sigma to +3 sigma from the center.

**INSIGHTS:**

**FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions.**

**1.**

**Maximum values.**

**distributions:**

z\_Volume 6.135311

Volume 31.000000

dtype: float64

**Minimum values.**

z\_Volume -0.962092

Volume 1.000000

dtype: float64

**2.**

**Maximum values.**

distributions: Avg Price 60100.000000

Z\_Avg\_score 2.745953

dtype: float64

**Minimum values.**

Avg Price 290.00000

Z\_Avg\_score -0.56214

dtype: float64

**3.**

**Maximum values.**

Z\_Total\_Sales\_Score 3.220894

Total Sales Value 196400.000000

dtype: float64

**Minimum values:**

Z\_Total\_Sales\_Score -0.661917

Total Sales Value 400.000000

dtype: float64

**4.**

**Maximum values.**

Z\_Discount\_Rate-score 1.147360

Discount Rate (%) 19.992407

dtype: float64

**Minimum values.**

Z\_Discount\_Rate-score -2.406935

Discount Rate (%) 5.007822

dtype: float64

**5.**

**Maximum values.**

*##FINDING MAX AND MIN VALUES in order to observe the ranges of the distributions*

df[["Z\_Discount\_Amount","Discount Amount"]].max()

Z\_Discount\_Amount 4.970494

Discount Amount 25738.022194

dtype: float64

**Minimum values.**

Z\_Discount\_Amount -0.727503

Discount Amount 69.177942

dtype: float64

**6.**

**Maximum values.**

Z\_Net\_sales\_score 3.218536

Net Sales Value 179507.479049

dtype: float64

**Minimum values.**

Z\_Net\_sales\_score -0.650858

Net Sales Value 326.974801

dtype: float64