TASK-1

1.DESCRIPTIVE ANALYSIS FOR NUMERICAL COLUMNS

```
---------- RESTART: C:\python\sprint3\jobsclm
Mean_volume                    5.066667
Median_volume                  4.000000
Mode volume                    3.000000
standard devation of voume     4.231602
Name: Volume, dtype: float64
Mean avg price                    10453.433333
Median avg price                   1450.000000
Mode avg price                      400.000000
standard devation of avg price    18079.904840
Name: Avg Price, dtype: float64
Mean total sales                    33812.835556
Median total sales                   5700.000000
Mode total sales                    24300.000000
standard devation of total sales    50535.074173
Name: Total Sales Value, dtype: float64
Mean of discount rate                  15.155242
Median discount rate                   16.577766
Mode discount rate                      5.007822
standard devation of discount rate      4.220602
Name: Discount Rate (%), dtype: float64
Mean discount amount                    3346.499424
Median discount amount                   988.933733
Mode discount amount                      69.177942
standard devation of discount amount    4509.902963
Name: Discount Amount, dtype: float64
Mean net sales value                30466.336131
Median net sales value               4677.788059
Mode net sales value                  326.974801
standard devation of sales_value    46358.656624
Name: Net Sales Value, dtype: float64
|
```

DATA VISULATION

HISTOGRAM

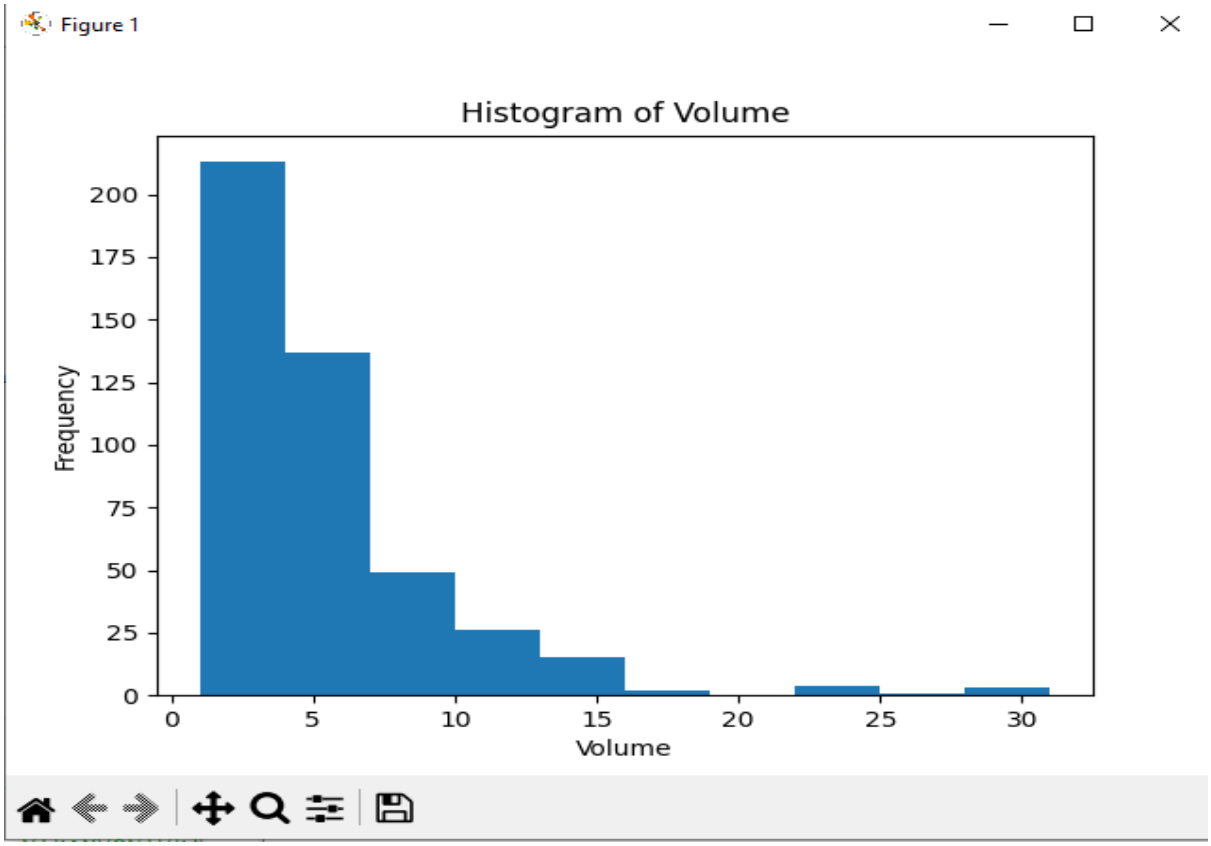


Figure 1
Histogram of Volume
Frequency
200
175
150
125
100
75
50
25
0
0
5
10
15
20
25
30
Volume

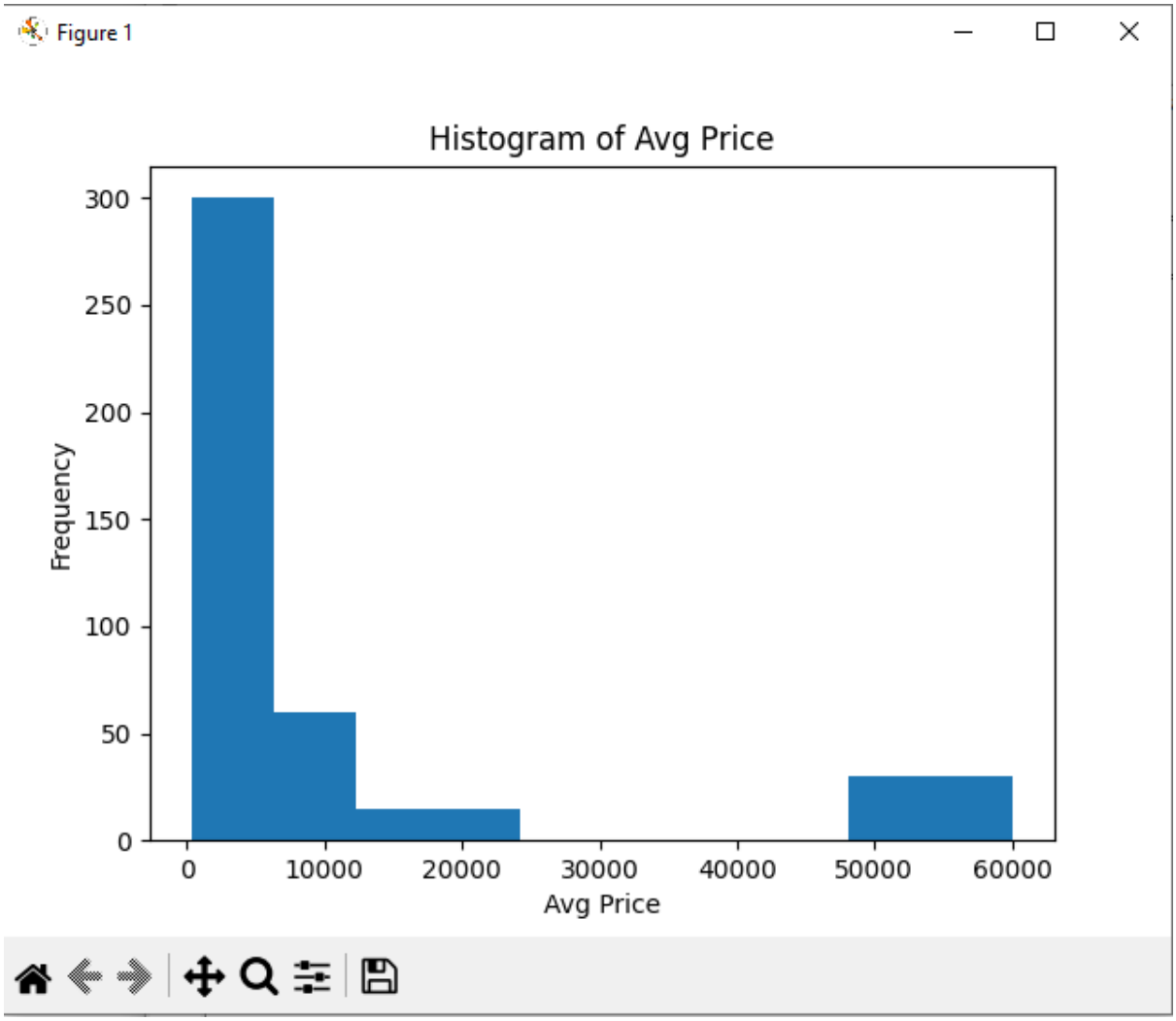


Figure 1
Histogram of Avg Price
Frequency
300
250
200
150
100
50
0
0
10000
20000
30000
40000
50000
60000
Avg Price

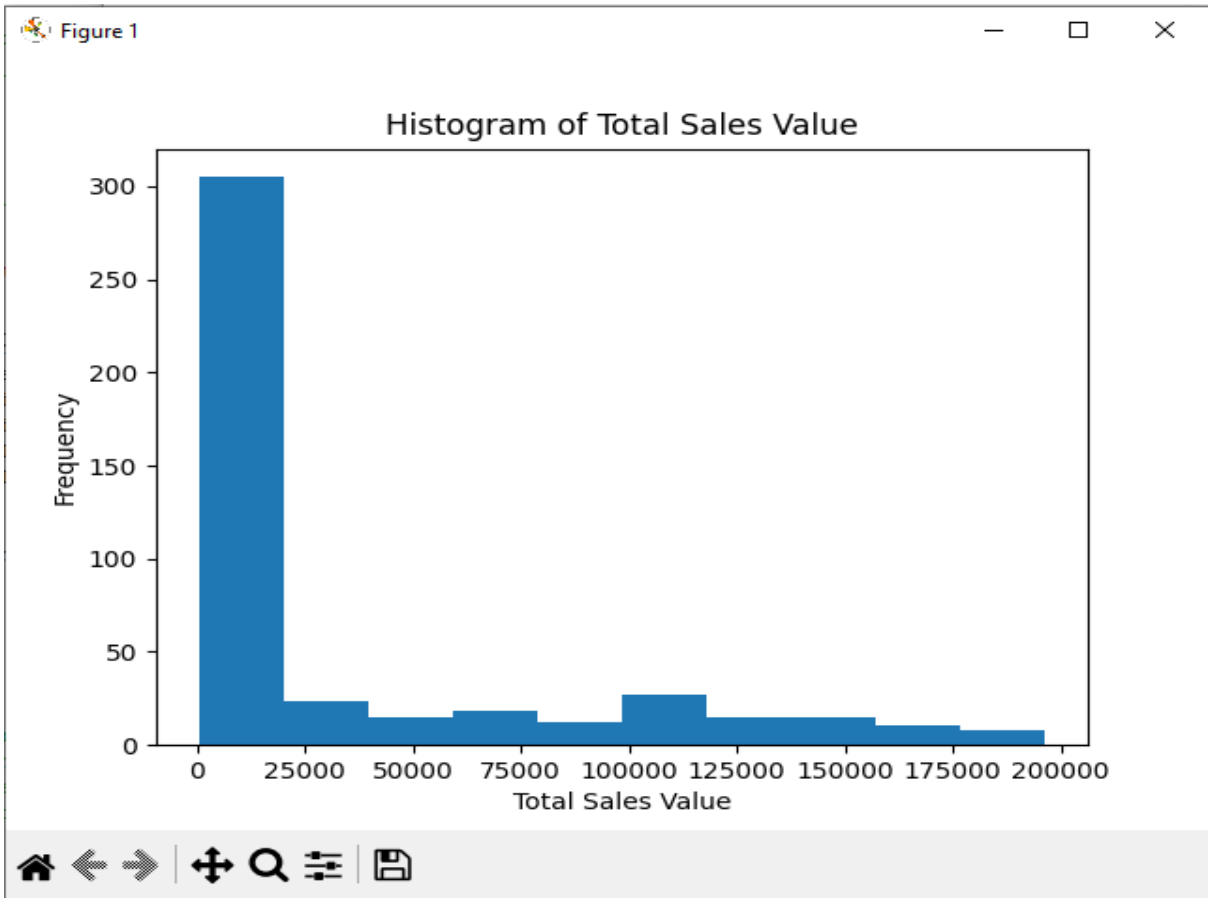


Figure 1
Histogram of Total Sales Value
Frequency
Total Sales Value
0
50
100
150
200
250
300
0
25000
50000
75000
100000
125000
150000
175000
200000

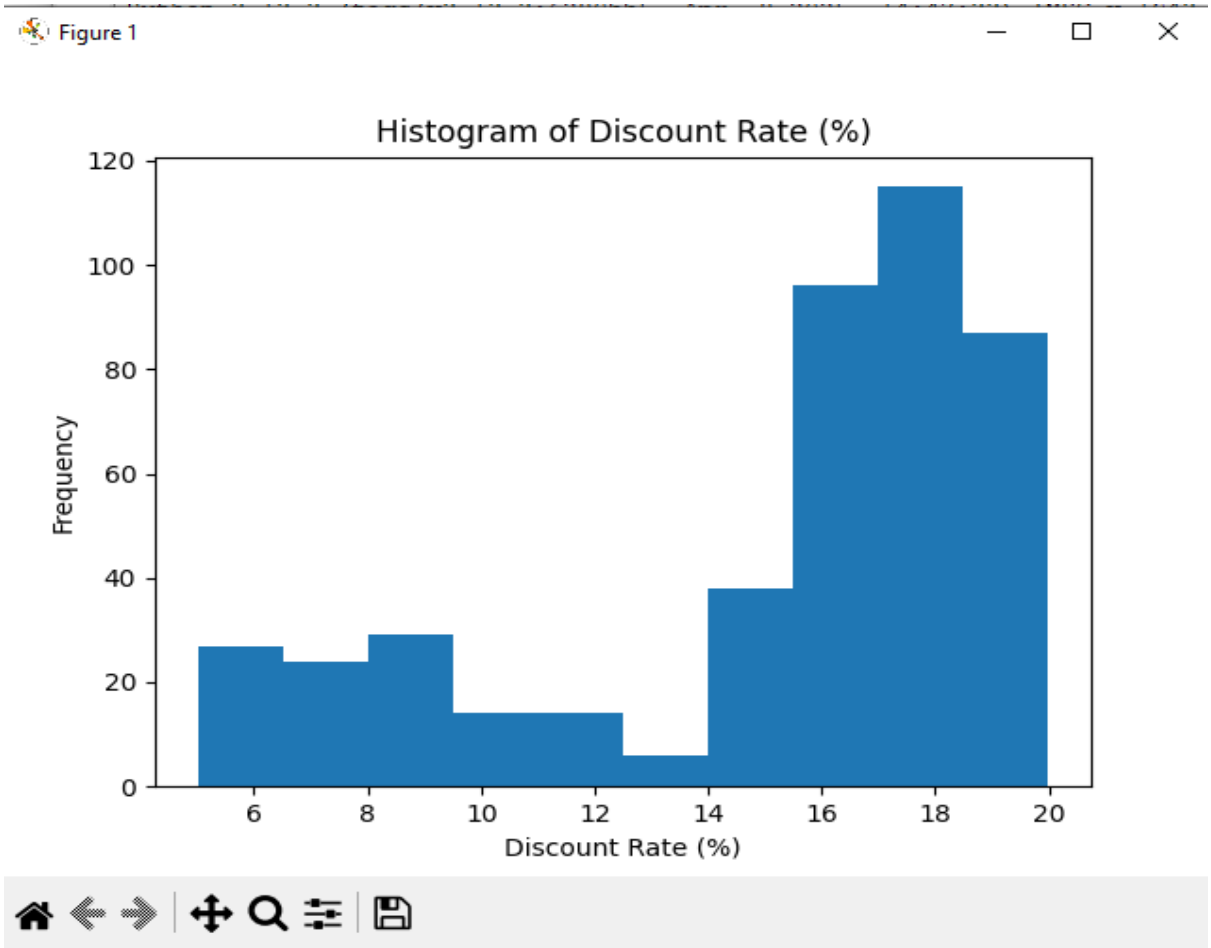


Figure 1
Histogram of Discount Rate (%)
Frequency
Discount Rate (%)
0
20
40
60
80
100
120
6
8
10
12
14
16
18
20

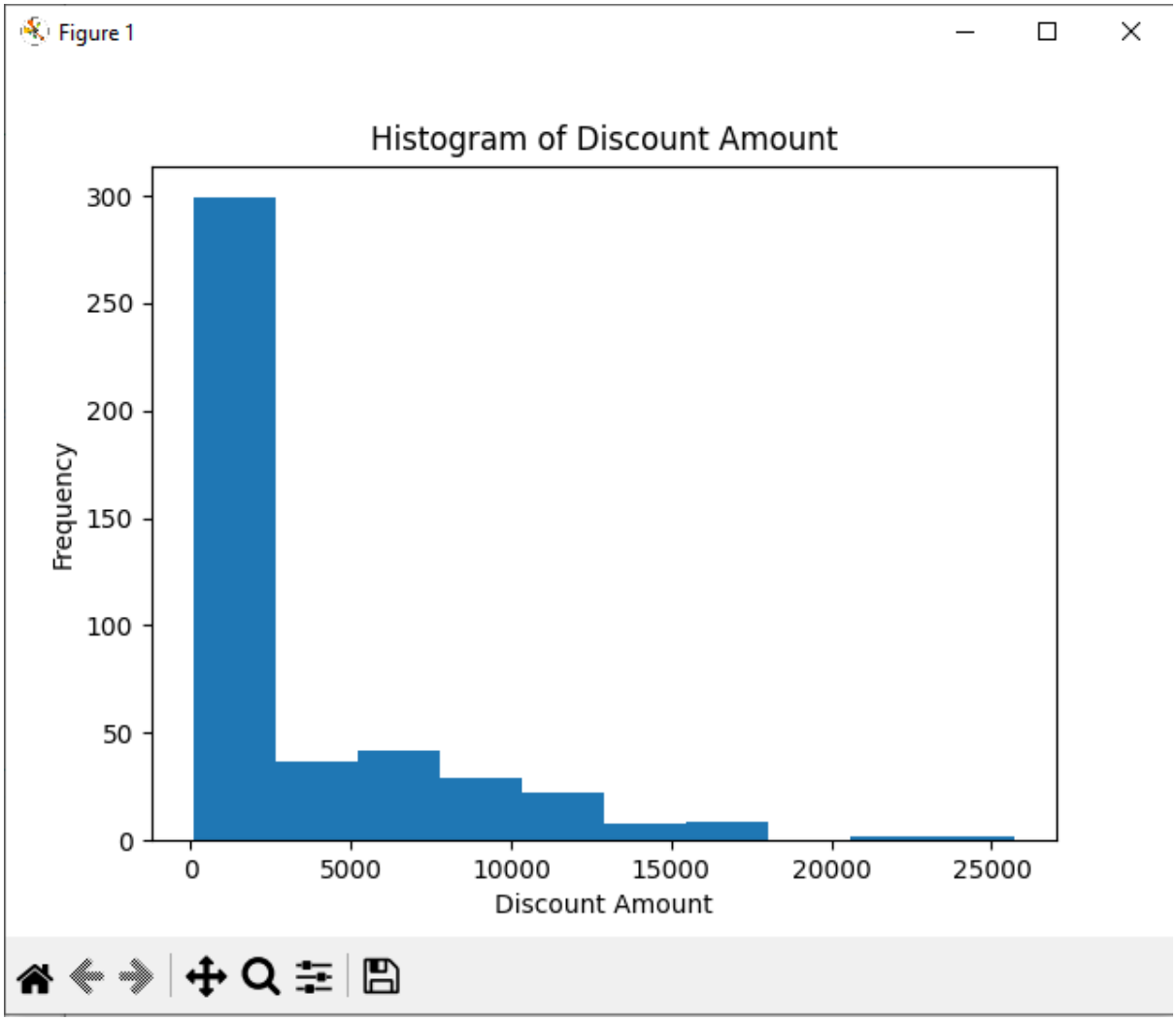


Figure 1
Histogram of Discount Amount
Frequency
300
250
200
150
100
50
0
0
5000
10000
15000
20000
25000
Discount Amount

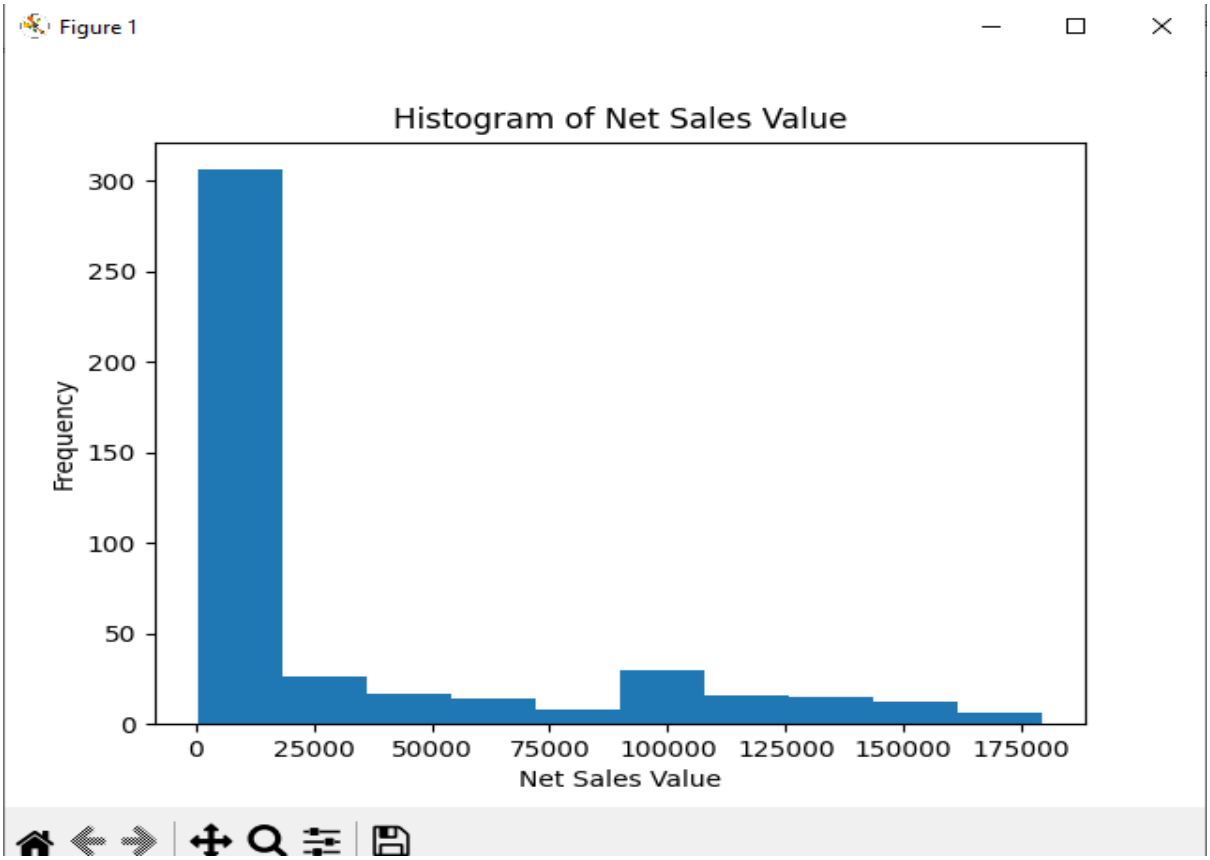


Figure 1
Histogram of Net Sales Value
Frequency
300
250
200
150
100
50
0
0
25000
50000
75000
100000
125000
150000
175000
Net Sales Value

BOX PLOT

```
number of values: 450
number of outliers:
           Date        Day  ... Discount Amount Net Sales Value
40   2021-04-02     Friday  ...      2283.816576     9316.183424
70   2021-04-03   Saturday  ...      1636.828489     7963.171511
100  2021-04-04     Sunday  ...      1793.227484     8206.772516
190  2021-04-07  Wednesday  ...      1600.081938     7199.918062
250  2021-04-09     Friday  ...      2155.607868    10244.392132
260  2021-04-09     Friday  ...      1378.138849     5821.861151
280  2021-04-10   Saturday  ...      1858.433334     7741.566666
310  2021-04-11     Sunday  ...      1600.588711     7199.411289
400  2021-04-14  Wednesday  ...      1744.927347     9855.072653
431  2021-04-15   Thursday  ...      1018.232342     4081.767658

[10 rows x 13 columns]
number of values: 450
number of outliers:
```

Squeezed text (63 lines).

```
number of values: 450
number of outliers:
 Empty DataFrame
Columns: [Date, Day, SKU, City, Volume, BU, Brand, Model, Avg Price, Total Sales
 Value, Discount Rate (%), Discount Amount, Net Sales Value]
Index: []
number of values: 450
number of outliers:
 Empty DataFrame
Columns: [Date, Day, SKU, City, Volume, BU, Brand, Model, Avg Price, Total Sales
 Value, Discount Rate (%), Discount Amount, Net Sales Value]
Index: []
number of values: 450
number of outliers:
           Date       Day  ... Discount Amount Net Sales Value
0    2021-04-01  Thursday  ...     21153.498820   160346.501180
210  2021-04-08  Thursday  ...     25328.224204   144071.775796
240  2021-04-09    Friday  ...     25738.022194   155761.977806
300  2021-04-11    Sunday  ...     21496.675367   123703.324633
```

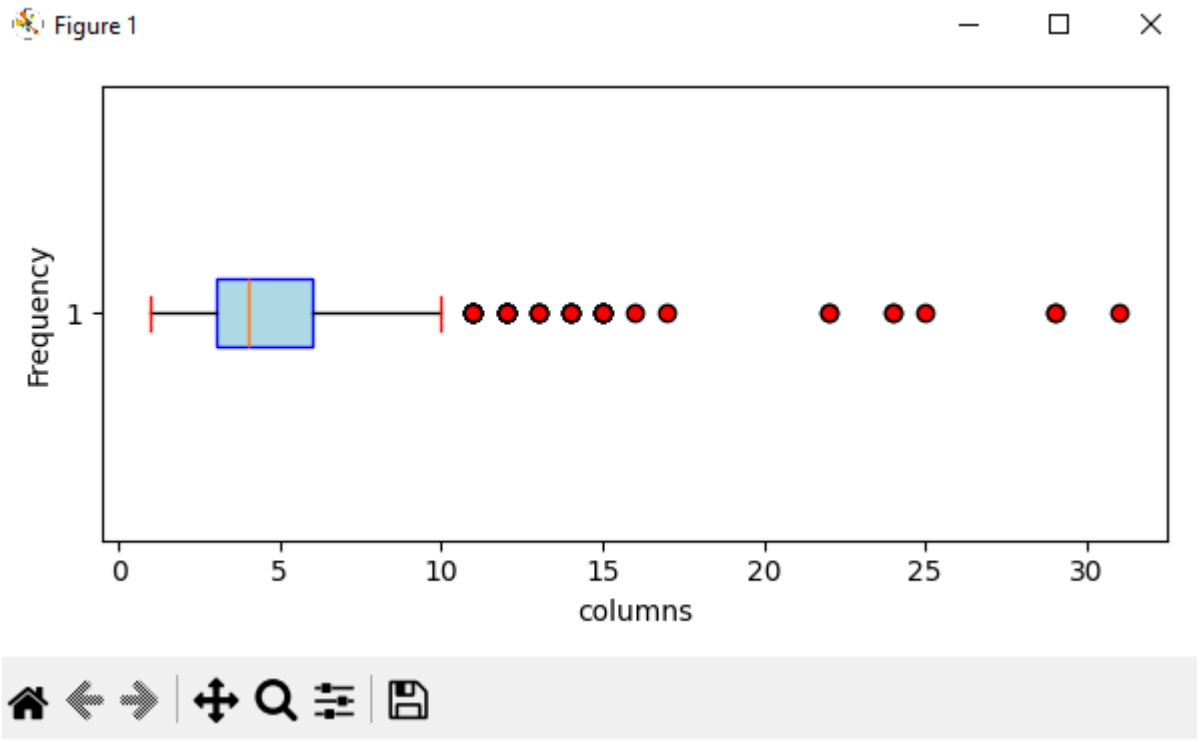


Figure 1
Frequency
1
0
5
10
15
20
25
30
columns

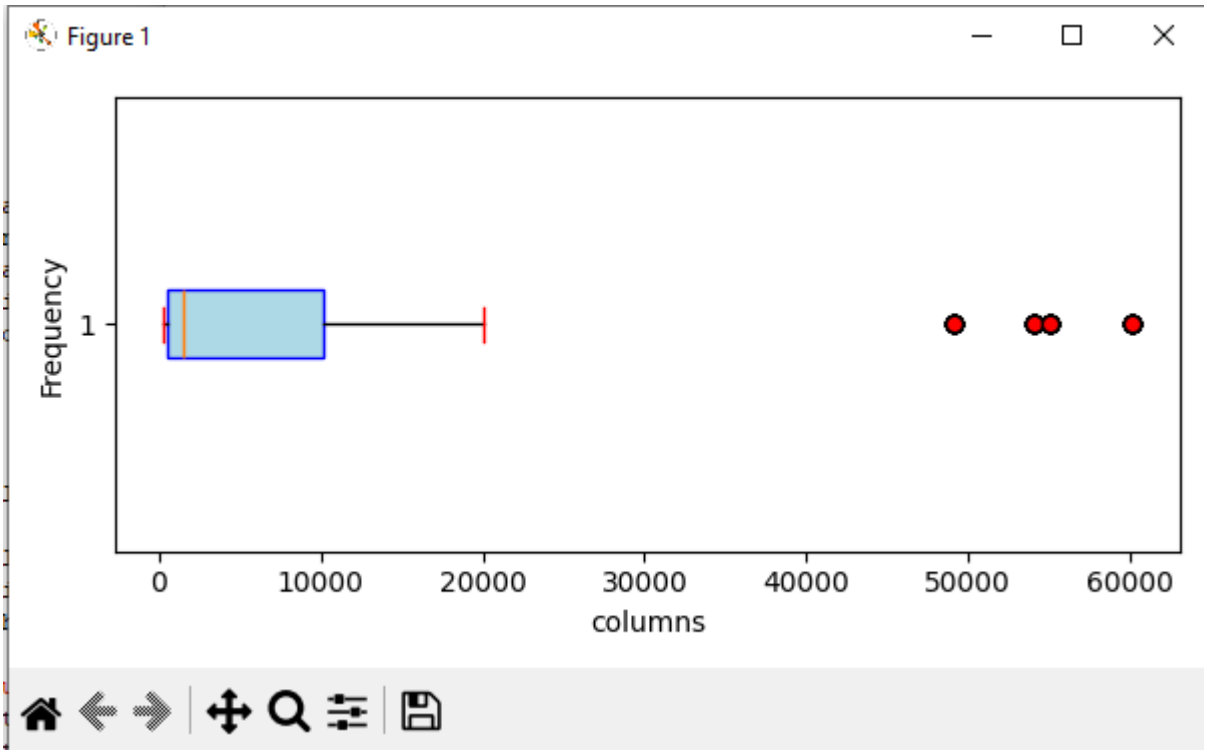


Figure 1
Frequency
1
0
10000
20000
30000
40000
50000
60000
columns

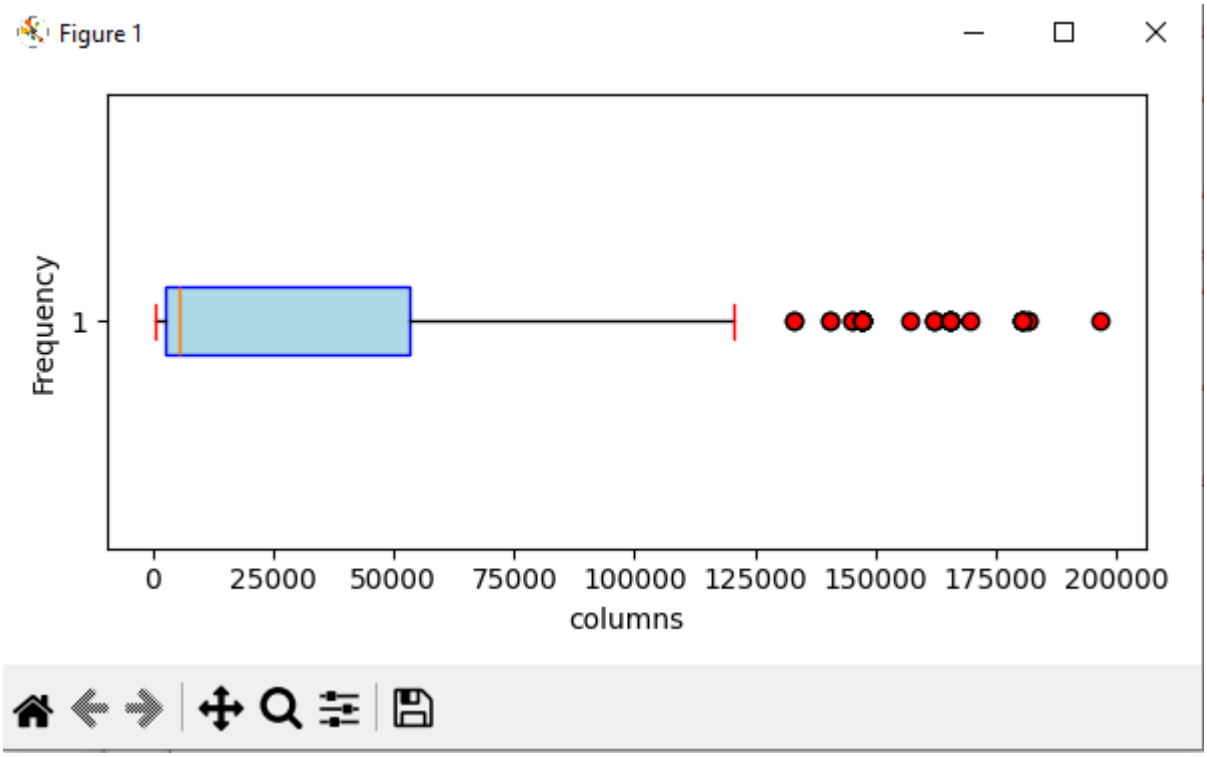


Figure 1
Frequency
1
0
25000
50000
75000
100000
125000
150000
175000
200000
columns

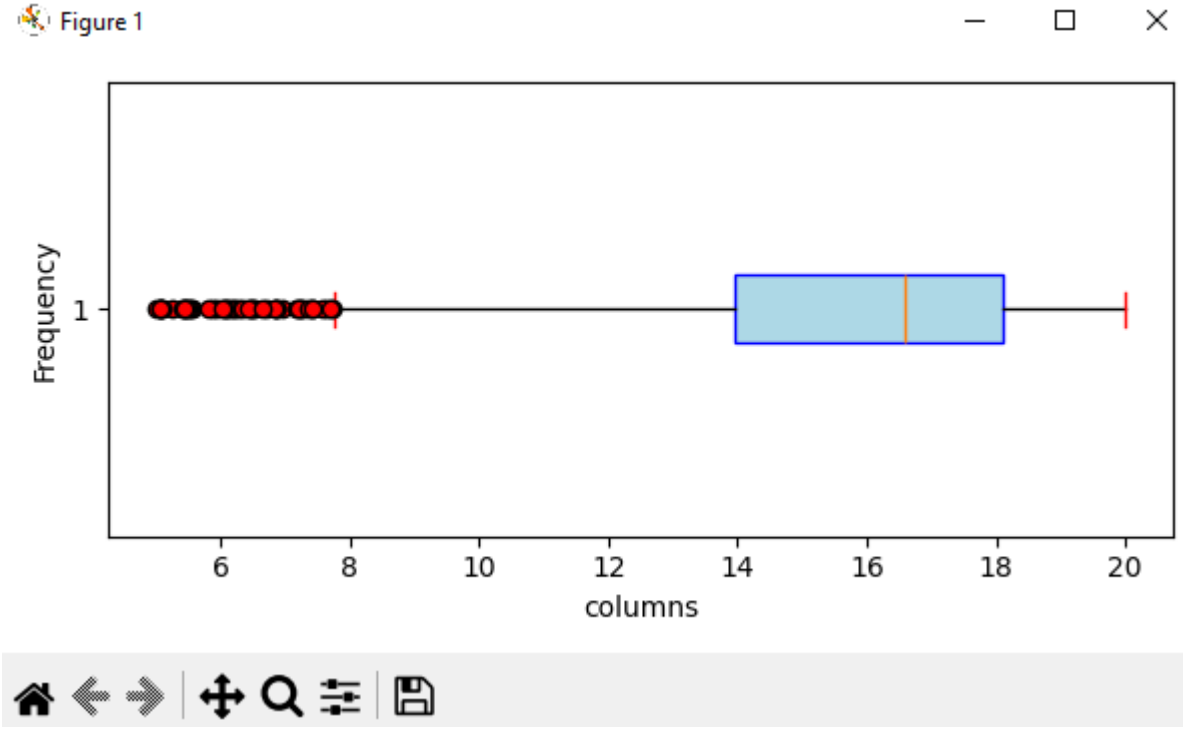
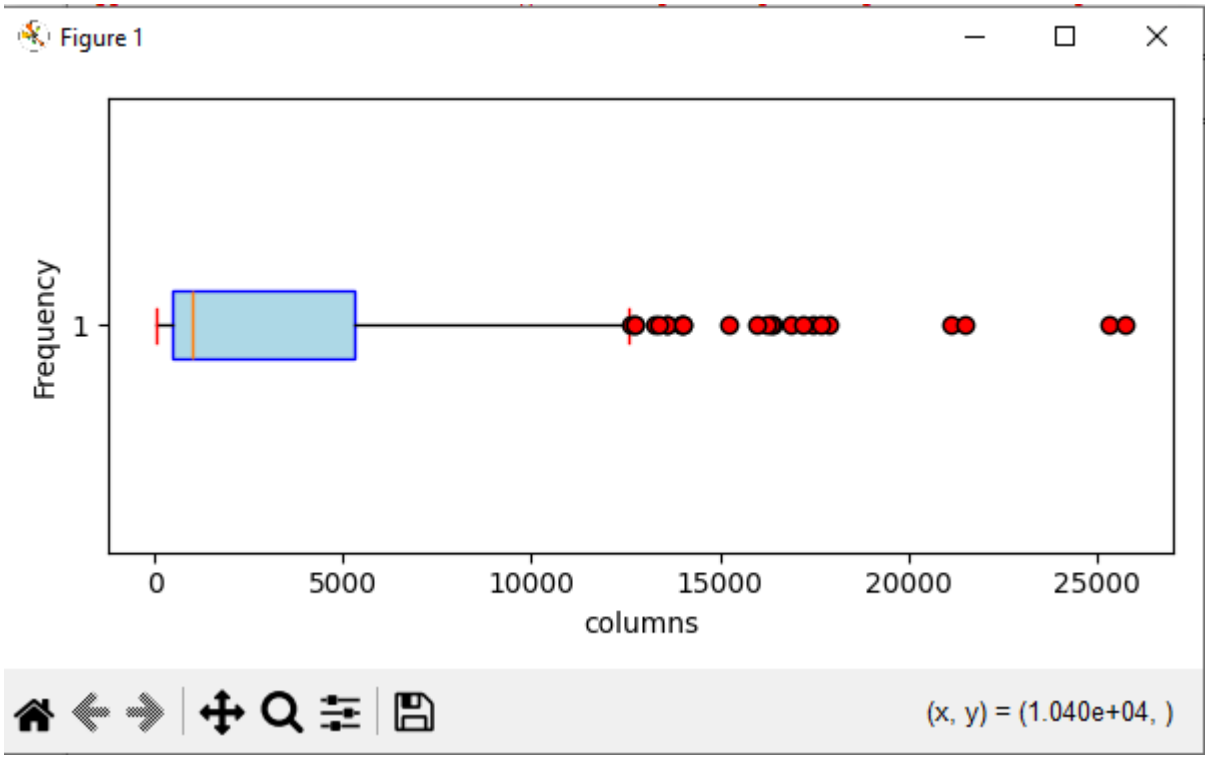
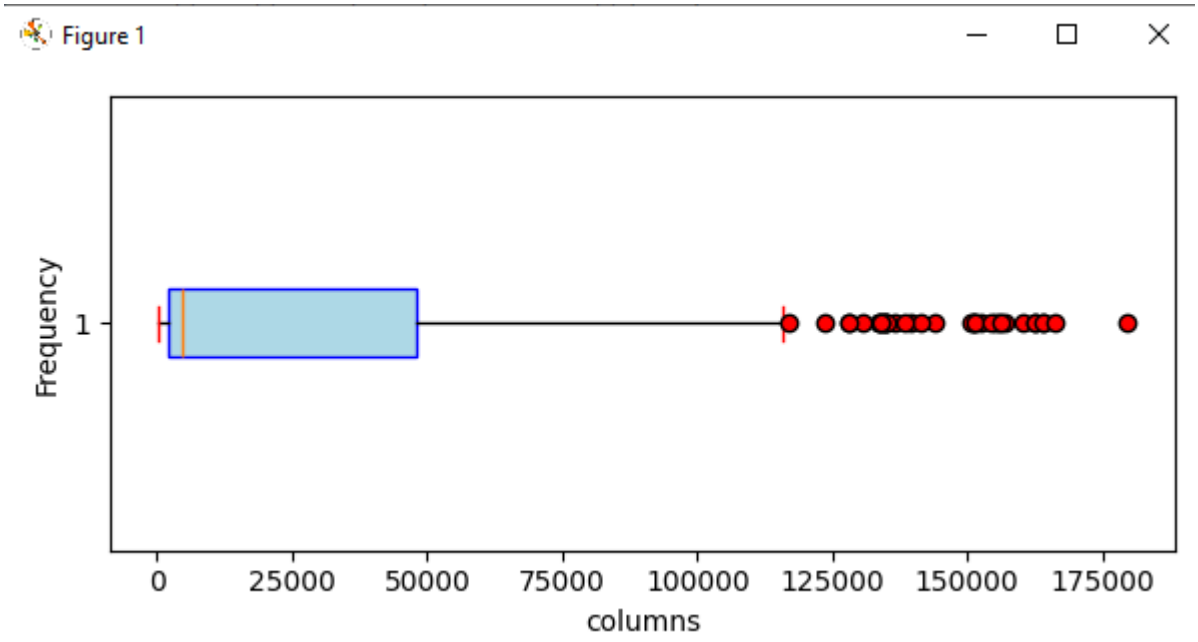


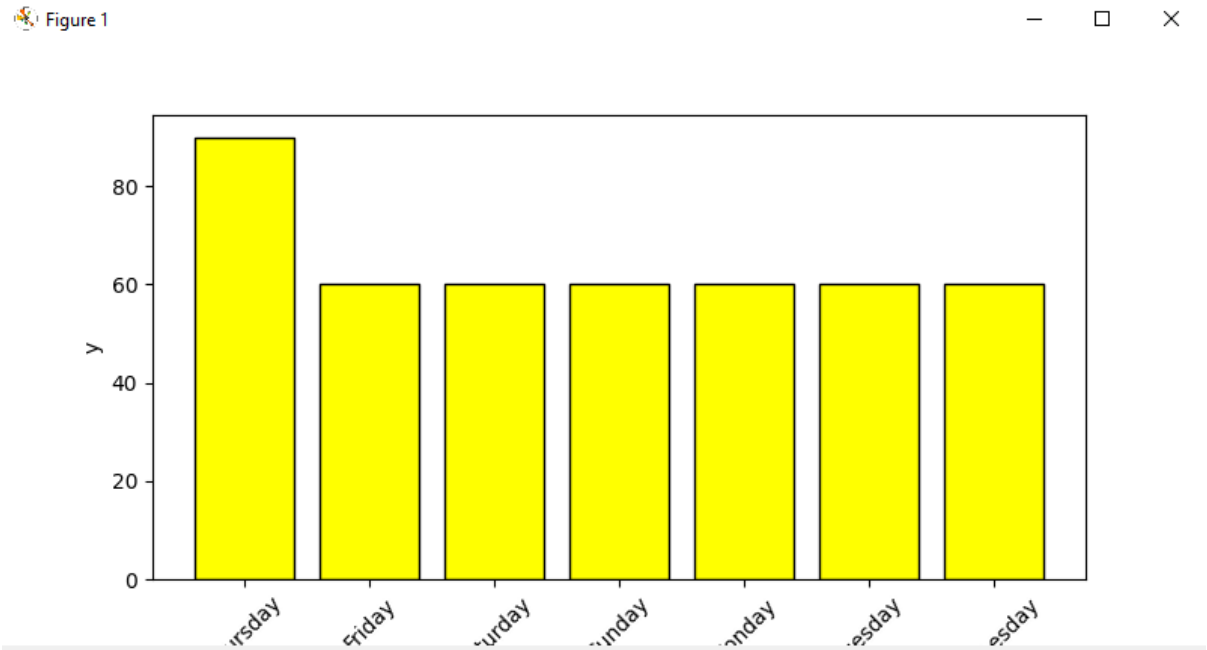
Figure 1
Frequency
1
6
8
10
12
14
16
18
20
columns



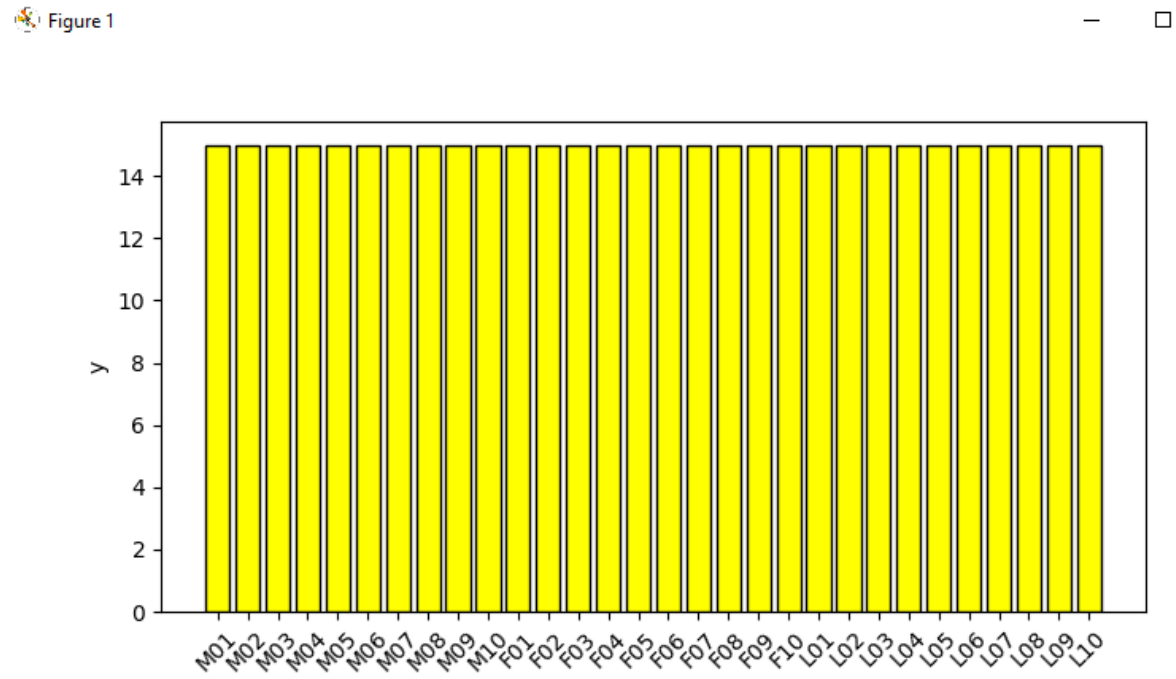


BAR CHART

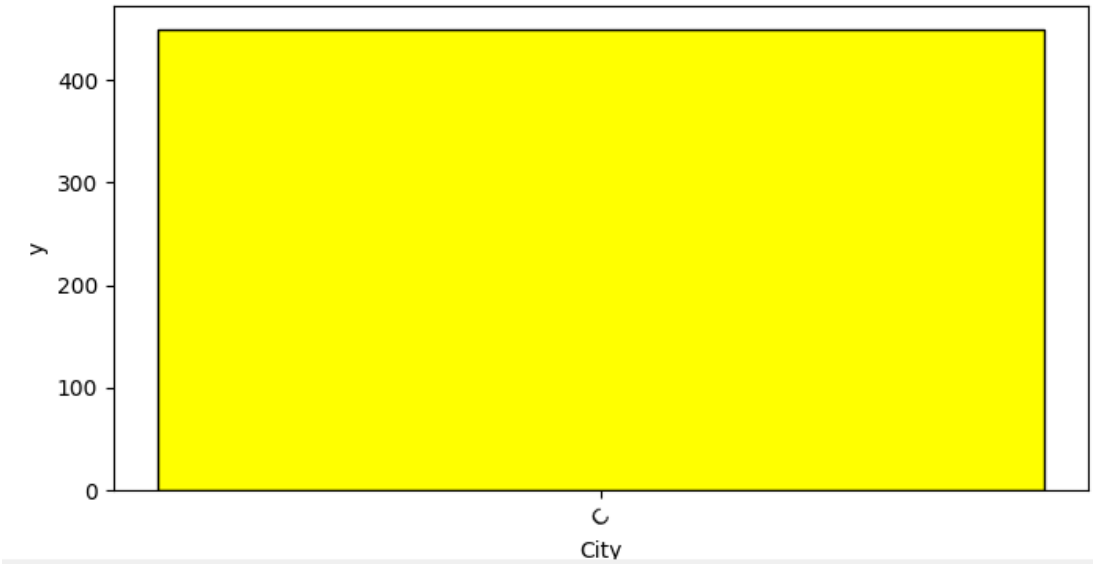
BAR CHAR OF COLUMN DAY



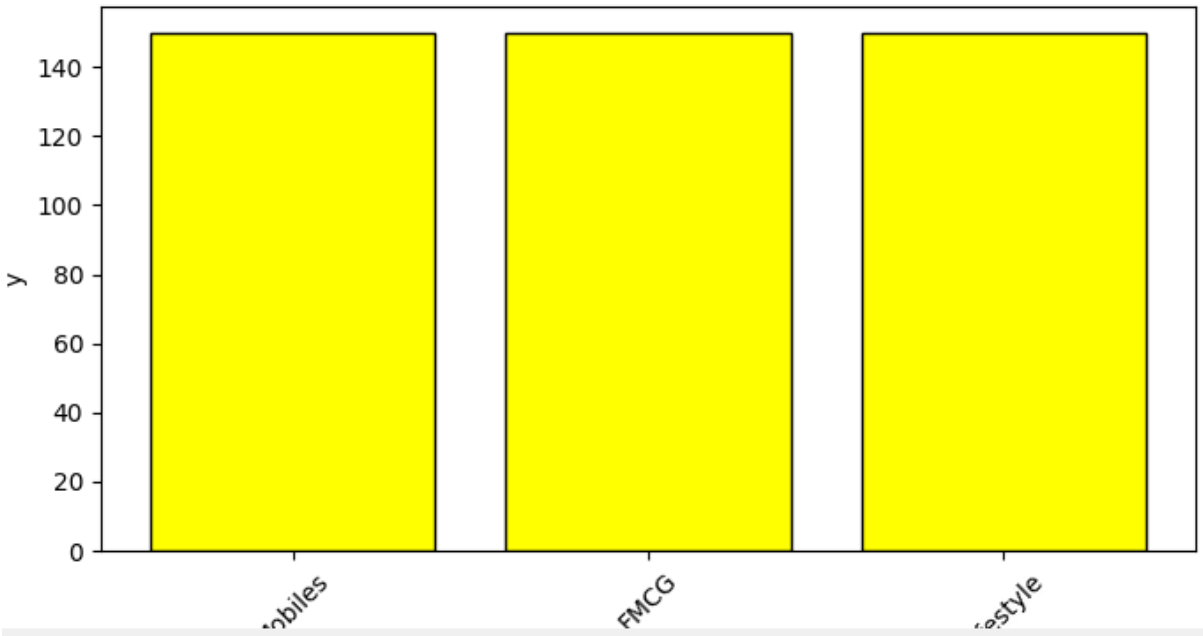
BARCHART OF COLUMN SKU



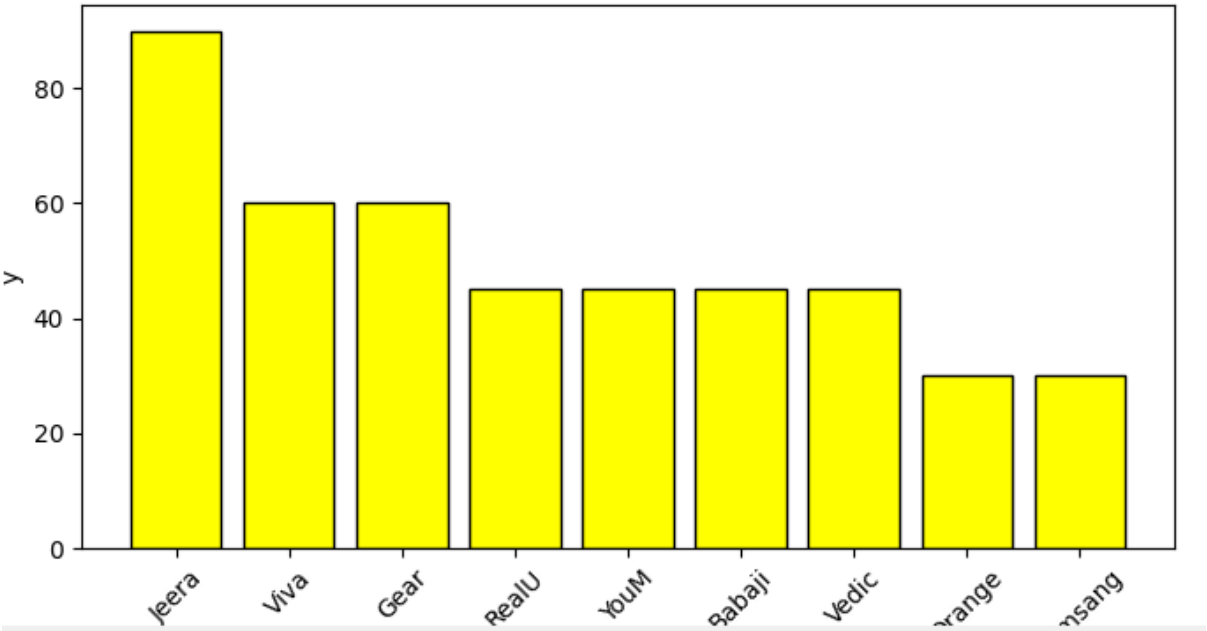
BARCHART OF COLUMN CITY



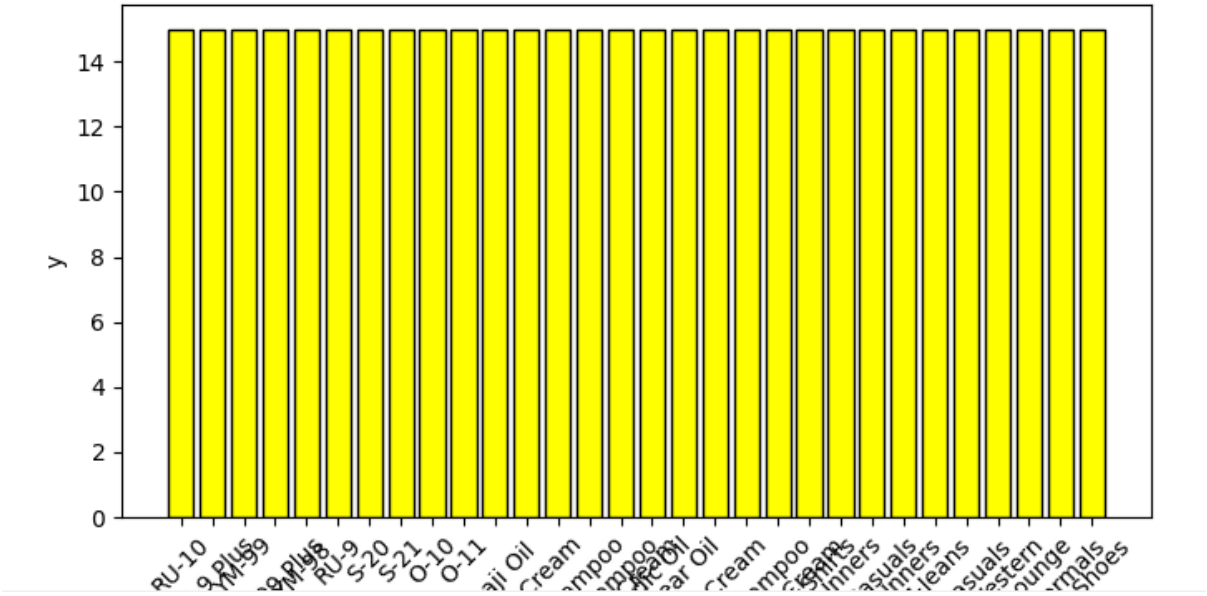
BARCHART OF COLUMN BU



BARCHART OF COLUMN BRAND



BARCHART OF COLUMN MODEL



TASK-2

```
---------- RESTART: C:\python\sprint3\jobsimulation1.py ----------
4.983333333333333
t-test -8.422963179596133
t_critical value -1.7958848187036696
The average battery life is less than 5.5 hours,The null hypothesis is rejected.
```