In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
coor=sns.color_palette()
import sklearn.metrics as metrics
import warnings
warnings.filterwarnings("ignore")
```

In [2]:
```
Default=pd.read_csv('Financial Analytics data.csv')
Default.head()
```

Out[2]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore	Unnamed: 4
0	1	Reliance Inds.	583436.72	99810.00	NaN
1	2	TCS	563709.84	30904.00	NaN
2	3	HDFC Bank	482953.59	20581.27	NaN
3	4	ITC	320985.27	9772.02	NaN
4	5	H D F C	289497.37	16840.51	NaN

In [3]:
```
Default.shape
```

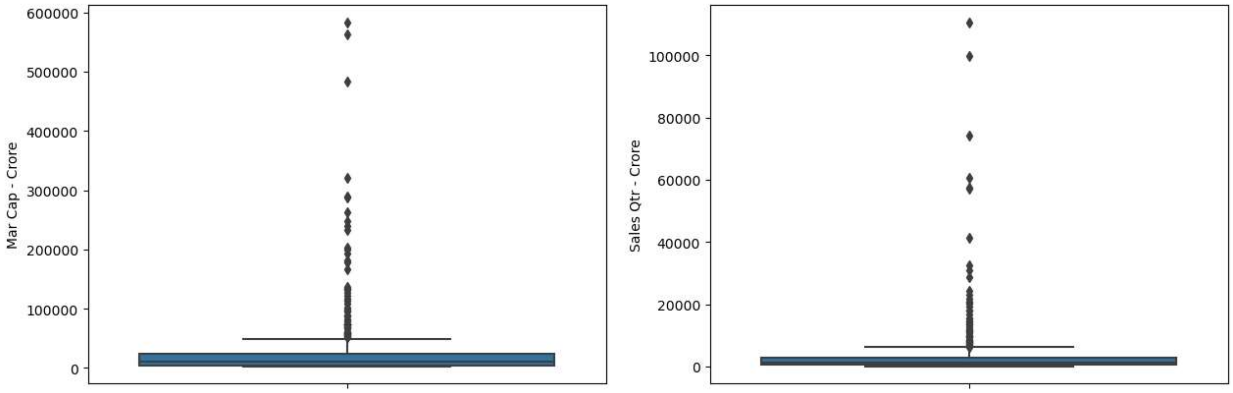
Out[3]:
```
(488, 5)
```

In [4]:
```
Default.describe()
```

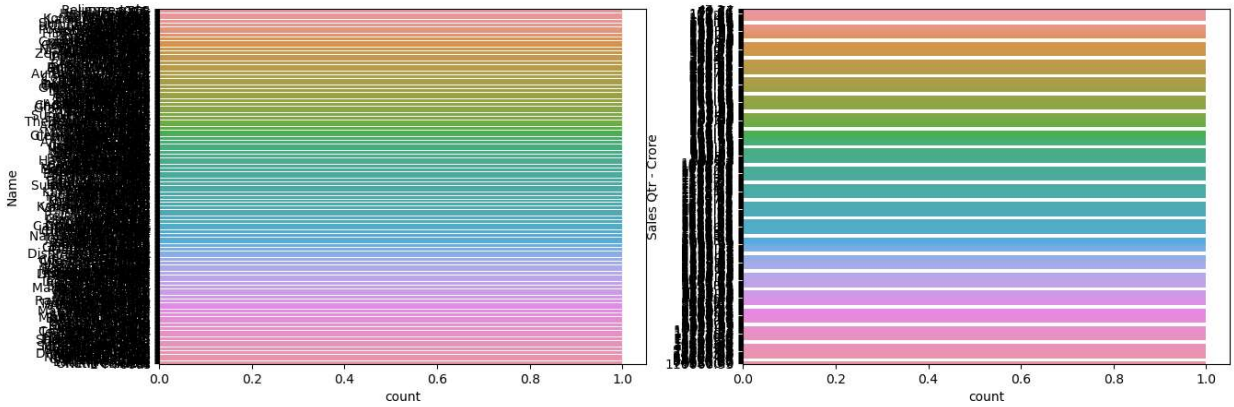
Out[4]:

	S.No.	Mar Cap - Crore	Sales Qtr - Crore	Unnamed: 4
count	488.000000	479.000000	365.000000	94.000000
mean	251.508197	28043.857119	4395.976849	1523.870106
std	145.884078	59464.615831	11092.206185	1800.008836
min	1.000000	3017.070000	47.240000	0.000000
25%	122.750000	4843.575000	593.740000	407.167500
50%	252.500000	9885.050000	1278.300000	702.325000
75%	378.250000	23549.900000	2840.750000	2234.815000
max	500.000000	583436.720000	110666.930000	7757.060000

In [5]:
```
plt.figure(figsize=(15,5))
plt.subplot(1,2,1)
sns.boxplot(y=Default['Mar Cap - Crore'])
plt.subplot(1,2,2)
sns.boxplot(y=Default['Sales Qtr - Crore'])
plt.show()
```



In [8]:
```
plt.figure(figsize=(15,5))
plt.subplot(1,2,1)
sns.countplot(y=Default['Name'])
plt.subplot(1,2,2)
sns.countplot(y=Default['Sales Qtr - Crore'])
plt.show()
```



In [9]:
```
Default["Name"].value_counts()
```

Out[9]:
```
Name
Reliance Inds.      1
Dishman Carbogen    1
Timken India        1
GE Power            1
Guj Alkalies        1
                   ..
Tata Global         1
Reliance Nip.Lif    1
Apollo Hospitals    1
MphasiS             1
L T Foods           1
Name: count, Length: 488, dtype: int64
```

In [11]:
```
Default["Sales Qtr - Crore"].value_counts()
```

Out[11]:
```
Sales Qtr - Crore
99810.00    1
584.42      1
299.80      1
1484.24     1
201.50      1
           ..
464.17      1
627.03      1
1730.39     1
394.00      1
2840.75     1
Name: count, Length: 365, dtype: int64
```

In [12]:
```
Default["Name"].value_counts(normalize=True)
```

Out[12]:
```
Name
Reliance Inds.      0.002049
Dishman Carbogen    0.002049
Timken India        0.002049
GE Power            0.002049
Guj Alkalies        0.002049
                      ...
Tata Global         0.002049
Reliance Nip.Lif    0.002049
Apollo Hospitals    0.002049
MphasiS             0.002049
L T Foods           0.002049
Name: proportion, Length: 488, dtype: float64
```

In [13]:
```
Default["Sales Qtr - Crore"].value_counts(normalize=True)
```

Out[13]:
```
Sales Qtr - Crore
99810.00    0.00274
584.42      0.00274
299.80      0.00274
1484.24     0.00274
201.50      0.00274
             ...
464.17      0.00274
627.03      0.00274
1730.39     0.00274
394.00      0.00274
2840.75     0.00274
Name: proportion, Length: 365, dtype: float64
```

In [19]:
```
pd.crosstab(Default['Name'],Default['Sales Qtr - Crore'],normalize='index').round(2)
```

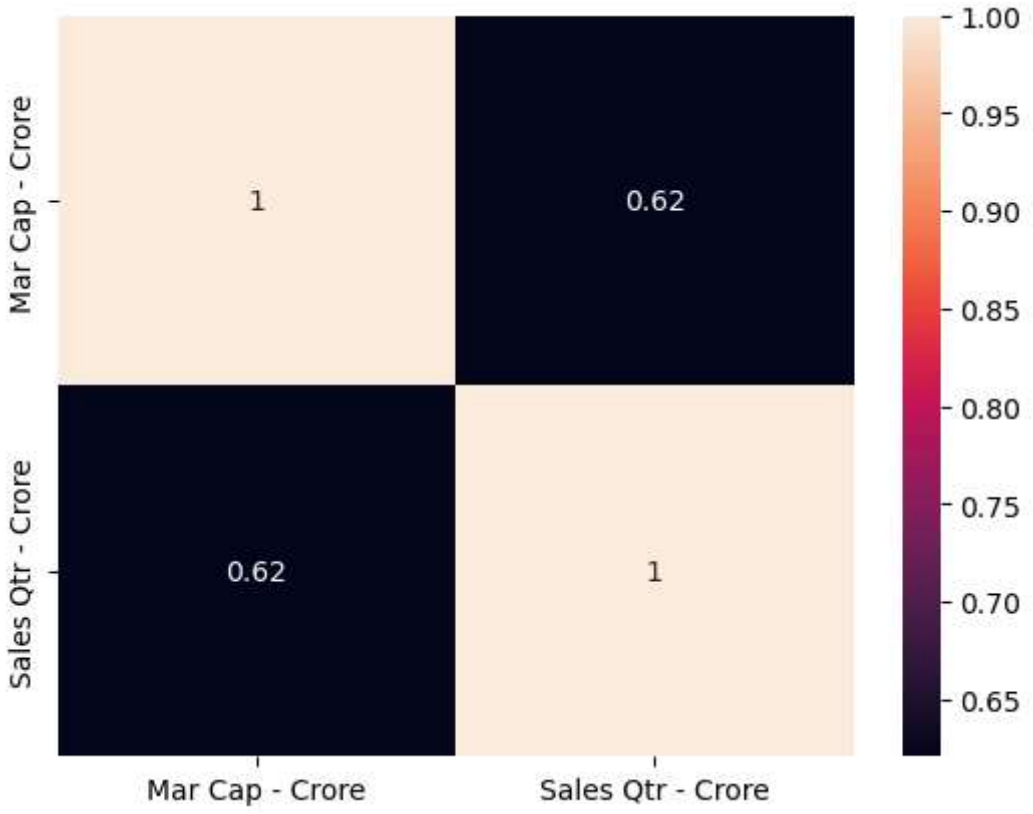
Out[19]:

Sales Qtr - Crore	47.24	69.77	70.64	77.84	102.14	112.05	132.40	138.65	148.42	162.68	...	28747
Name												
3M India	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
AIA Engg.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
APL Apollo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
AU Small Finance	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
Abbott India	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
...	...	...	...	...	...	...	...	...	...	...	...	
Yes Bank	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
Zee Entertainmen	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
Zensar Tech.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	
Zydus Wellness	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	...	
eClerx Services	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	

365 rows × 365 columns

In [20]:
```
sns.heatmap(Default[['Mar Cap - Crore','Sales Qtr - Crore']].corr(),annot=True)
```

Out[20]:
```
<Axes: >
```



In [21]:
```
Default.isnull().sum()
```

Out[21]:
```
S.No.                  0
Name                   0
Mar Cap - Crore        9
Sales Qtr - Crore    123
Unnamed: 4           394
dtype: int64
```

In [23]:
```
Q1,Q3=Default['Mar Cap - Crore'].quantile([.25,.75])
IQR=Q3-Q1
LL=Q1-1.5*(IQR)
UL=Q3+1.5*(IQR)
```

In [24]:
```
UL
```

Out[24]:
```
51609.387500000004
```

In [30]:
```
df=Default[Default['Mar Cap - Crore']>UL]
df
```

Out[30]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore	Unnamed: 4
0	1	Reliance Inds.	583436.72	99810.00	NaN
1	2	TCS	563709.84	30904.00	NaN
2	3	HDFC Bank	482953.59	20581.27	NaN
3	4	ITC	320985.27	9772.02	NaN
4	5	H D F C	289497.37	16840.51	NaN
...	...	...	...	...	...
58	59	Hindalco Inds.	55854.68	11022.81	NaN
59	60	Zee Entertainmen	54817.89	1838.07	NaN
60	61	Cairn India	53528.57	NaN	2149.36
61	62	Indiabulls Hous.	52781.67	NaN	3115.89
62	63	Ambuja Cem.	52361.46	NaN	6170.71

63 rows × 5 columns

In [31]:
```
df['Name'].count()
```

Out[31]:
```
63
```

In [27]:
```
df['Name'].value_counts(normalize=True)
```

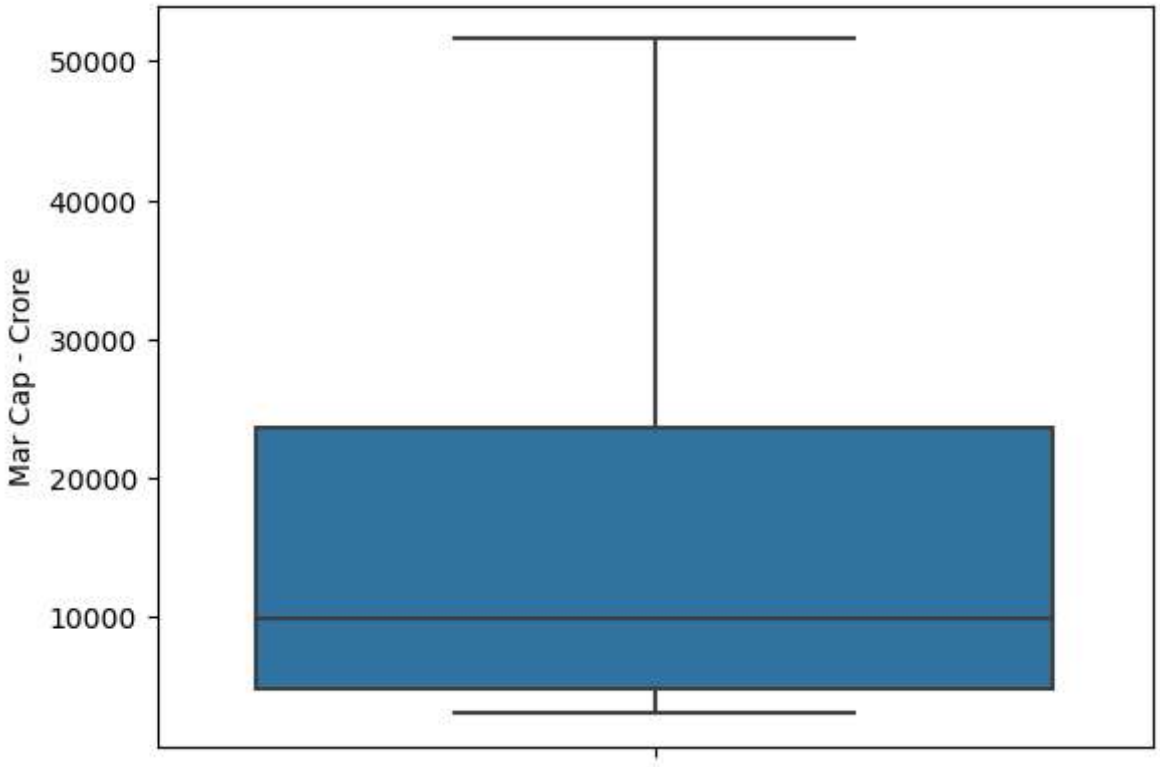
Out[27]:
```
Name
Reliance Inds.      0.015873
SBI Life Insuran    0.015873
Bajaj Finserv       0.015873
GAIL (India)        0.015873
Avenue Super.       0.015873
                      ...
Power Grid Corpn    0.015873
B P C L             0.015873
IndusInd Bank       0.015873
Bajaj Fin.          0.015873
Ambuja Cem.         0.015873
Name: proportion, Length: 63, dtype: float64
```

In [28]:
```
df['Name'].value_counts()
```

Out[28]:
```
Name
Reliance Inds.      1
SBI Life Insuran    1
Bajaj Finserv       1
GAIL (India)        1
Avenue Super.       1
                   ..
Power Grid Corpn    1
B P C L             1
IndusInd Bank       1
Bajaj Fin.          1
Ambuja Cem.         1
Name: count, Length: 63, dtype: int64
```

In [32]:
```
Default['Mar Cap - Crore']=np.where(Default['Mar Cap - Crore']>UL,UL,Default['Mar Cap
```

In [34]:
```
sns.boxplot(y=Default['Mar Cap - Crore'])
plt.show()
```



In [35]:
```
Default=pd.get_dummies(Default,drop_first=True)
```

In [36]:
```
Default.head()
```

Out[36]:

	S.No.	Mar Cap - Crore	Sales Qtr - Crore	Unnamed: 4	Name_A B B	Name_ACC	Name_AIA Engg.	Name_APL Apollo	Name_AU Small Finance	N
0	1	51609.3875	99810.00	NaN	False	False	False	False	False	
1	2	51609.3875	30904.00	NaN	False	False	False	False	False	
2	3	51609.3875	20581.27	NaN	False	False	False	False	False	
3	4	51609.3875	9772.02	NaN	False	False	False	False	False	
4	5	51609.3875	16840.51	NaN	False	False	False	False	False	

5 rows × 491 columns

In [ ]:

In [ ]: