In [ ]:
```
#Key Metrices- 1.Comparing the performanance of different companies based on their market cap.
#2. company with maximum Sales Qtr.
#3. Top company based on market cap.
#4. company having min. sales QTR
#5. relation between market cap and sales Qtr
#6. Top 10 companies based on marketcap
#7.TOP 10 companies based on sales Qtr
#8. Bottom 10 companies based on marketcap
#9.Bottom 10 companies based on sales Qtr
#10.which companies are most in number based on  large cap,mid cap and small cap
#11. Count of large cap,mid cap and small cap companies
```

In [46]:
```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
%matplotlib inline
import seaborn as sns
```

In [47]:
```
df=pd.read_excel('Financial Analytics dataset.xlsx')
```

In [48]:
```
df
```

Out[48]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore
0	1	Reliance Inds.	583436.72	99810.00
1	2	TCS	563709.84	30904.00
2	3	HDFC Bank	482953.59	20581.27
3	4	ITC	320985.27	9772.02
4	5	H D F C	289497.37	16840.51
...	...	...	...	...
483	496	Lak. Vilas Bank	3029.57	790.17
484	497	NOCIL	3026.26	249.27
485	498	Orient Cement	3024.32	511.53
486	499	Natl.Fertilizer	3017.07	2840.75
487	500	L T Foods	NaN	NaN

488 rows × 4 columns

In [49]:
```
df.shape  # This dataset has 488 rows and 5 columns
```

Out[49]:
```
(488, 4)
```

In [50]:
```
df.head(10)     # It will give first 10 rows of dataset.
```

Out[50]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore
0	1	Reliance Inds.	583436.72	99810.00
1	2	TCS	563709.84	30904.00
2	3	HDFC Bank	482953.59	20581.27
3	4	ITC	320985.27	9772.02
4	5	H D F C	289497.37	16840.51
5	6	Hind. Unilever	288265.26	8590.00
6	7	Maruti Suzuki	263493.81	19283.20
7	8	Infosys	248320.35	17794.00
8	9	O N G C	239981.50	22995.88
9	10	St Bk of India	232763.33	57014.08

In [51]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 488 entries, 0 to 487
Data columns (total 4 columns):
 #   Column             Non-Null Count  Dtype  
---  ------             --------------  -----  
 0   S.No.              488 non-null    int64  
 1   Name               488 non-null    object 
 2   Mar Cap - Crore    479 non-null    float64
 3   Sales Qtr - Crore  459 non-null    float64
dtypes: float64(2), int64(1), object(1)
memory usage: 15.4+ KB
```

In [52]:
```
df.dtypes
```

Out[52]:
```
S.No.                 int64
Name                 object
Mar Cap - Crore     float64
Sales Qtr - Crore   float64
dtype: object
```

In [53]:
```
# Data Cleaning
```

In [54]:
```
df.isnull().sum()     # To find any null value present in dataset
```

Out[54]:
```
S.No.                0
Name                 0
Mar Cap - Crore      9
Sales Qtr - Crore    29
dtype: int64
```

In [55]:
```
df['Sales Qtr - Crore']=df['Sales Qtr - Crore'].fillna(0)
df['Mar Cap - Crore']=df['Mar Cap - Crore'].fillna(0)
```

In [56]:
```
df
```

Out[56]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore
0	1	Reliance Inds.	583436.72	99810.00
1	2	TCS	563709.84	30904.00
2	3	HDFC Bank	482953.59	20581.27
3	4	ITC	320985.27	9772.02
4	5	H D F C	289497.37	16840.51
...	...	...	...	...
483	496	Lak. Vilas Bank	3029.57	790.17
484	497	NOCIL	3026.26	249.27
485	498	Orient Cement	3024.32	511.53
486	499	Natl.Fertilizer	3017.07	2840.75
487	500	L T Foods	0.00	0.00

488 rows × 4 columns

In [57]:
```
df.isnull().sum()
```

Out[57]:
```
S.No.                0
Name                 0
Mar Cap - Crore      0
Sales Qtr - Crore    0
dtype: int64
```

In [58]:
```
df['Sales Qtr - Crore'].replace(0,np.nan,inplace=True)
df['Mar Cap - Crore'].replace(0,np.nan,inplace=True)
```

In [59]:
```
df['Sales Qtr - Crore']=df['Sales Qtr - Crore'].fillna(df['Sales Qtr - Crore'].mode()[0])
df['Mar Cap - Crore']=df['Mar Cap - Crore'].fillna(df['Mar Cap - Crore'].mode()[0])
```

In [60]:
```
df.tail(20)
```

Out[60]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore
468	481	Puravankara	3274.90	440.09
469	482	Tejas Networks	3209.89	229.87
470	483	ITD Cem	3192.67	472.98
471	484	Hathway Cable	3189.10	138.65
472	485	Dhanuka Agritech	3188.62	221.51
473	486	Mahindra Logis.	3187.51	835.06
474	487	Heritage Foods	3185.45	581.74
475	488	Mah. Seamless	3164.73	563.66
476	489	Navneet Educat.	3148.36	174.41
477	490	Firstsour.Solu.	3139.94	887.24
478	491	Kaveri Seed Co.	3125.83	70.64
479	492	Star Ferro Cem.	3115.98	393.49
480	493	Deepak Fert.	3079.06	1644.92
481	494	Va Tech Wabag	3041.93	460.89
482	495	Prime Focus	3031.50	609.61
483	496	Lak. Vilas Bank	3029.57	790.17
484	497	NOCIL	3026.26	249.27
485	498	Orient Cement	3024.32	511.53
486	499	Natl.Fertilizer	3017.07	2840.75
487	500	L T Foods	3017.07	19.42

In [61]:
```
df['scale']=['large cap' if x>=100000   else 'mid cap' if 100000>=x>30000  else 'small cap' for x in df['Mar Cap - Crore']]
```

In [ ]:
```

```

In [62]:
```
df.tail(40)
```

Out[62]:

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore	scale
448	461	JK Tyre & Indust	3577.98	2123.24	small cap
449	462	Deepak Nitrite	3531.90	371.14	small cap
450	463	Heidelberg Cem.	3531.77	491.23	small cap
451	464	Amber Enterp.	3529.87	19.42	small cap
452	465	Sharda Cropchem	3528.07	325.46	small cap
453	466	Dixon Technolog.	3526.80	677.80	small cap
454	467	Himatsing. Seide	3511.08	595.61	small cap
455	468	La Opala RG	3510.93	69.77	small cap
456	469	H F C L	3482.71	714.42	small cap
457	470	Reliance Home	3470.60	403.00	small cap
458	471	Rupa & Co	3460.91	277.96	small cap
459	472	Hind.Construct.	3452.57	19.42	small cap
460	473	Ent.Network	3380.99	148.42	small cap
461	474	Supreme Petroch.	3377.57	725.02	small cap
462	475	MAS FINANC SER	3376.20	112.05	small cap
463	476	Thyrocare Tech.	3374.38	77.84	small cap
464	477	Prakash Inds.	3336.05	725.97	small cap
465	478	Repco Home Fin	3331.08	277.48	small cap
466	479	Sonata Software	3329.58	766.75	small cap
467	480	Central Dep. Ser	3316.31	47.24	small cap
468	481	Puravankara	3274.90	440.09	small cap
469	482	Tejas Networks	3209.89	229.87	small cap
470	483	ITD Cem	3192.67	472.98	small cap
471	484	Hathway Cable	3189.10	138.65	small cap

	S.No.	Name	Mar Cap - Crore	Sales Qtr - Crore	scale
472	485	Dhanuka Agritech	3188.62	221.51	small cap
473	486	Mahindra Logis.	3187.51	835.06	small cap
474	487	Heritage Foods	3185.45	581.74	small cap
475	488	Mah. Seamless	3164.73	563.66	small cap
476	489	Navneet Educat.	3148.36	174.41	small cap
477	490	Firstsour.Solu.	3139.94	887.24	small cap
478	491	Kaveri Seed Co.	3125.83	70.64	small cap
479	492	Star Ferro Cem.	3115.98	393.49	small cap
480	493	Deepak Fert.	3079.06	1644.92	small cap
481	494	Va Tech Wabag	3041.93	460.89	small cap
482	495	Prime Focus	3031.50	609.61	small cap
483	496	Lak. Vilas Bank	3029.57	790.17	small cap
484	497	NOCIL	3026.26	249.27	small cap
485	498	Orient Cement	3024.32	511.53	small cap
486	499	Natl.Fertilizer	3017.07	2840.75	small cap
487	500	L T Foods	3017.07	19.42	small cap

In [63]:
```
df.groupby('Name')['scale'].count().head(10)
```

Out[63]:
```
Name
3M India            1
A B B               1
ACC                 1
AIA Engg.           1
APL Apollo          1
AU Small Finance    1
Abbott India        1
Adani Enterp.       1
Adani Ports         1
Adani Power         1
Name: scale, dtype: int64
```

In [64]:
```
df['Sales Qtr - Crore'].value_counts()
```

Out[64]:
```
Sales Qtr - Crore
19.42       31
99810.00     1
431.21       1
584.42       1
619.93       1
            ..
5797.20      1
1321.50      1
2283.72      1
2276.54      1
2840.75      1
Name: count, Length: 458, dtype: int64
```

In [65]:
```
df['Mar Cap - Crore'].value_counts()
```

Out[65]:
```
Mar Cap - Crore
3017.07      10
583436.72     1
5495.76       1
5498.45       1
5502.94       1
             ..
16655.58      1
16683.97      1
16728.78      1
17097.54      1
3024.32       1
Name: count, Length: 479, dtype: int64
```

In [66]:
```
df.columns
```

Out[66]:
```
Index(['S.No.', 'Name', 'Mar Cap - Crore', 'Sales Qtr - Crore', 'scale'], dtype='object')
```

In [67]:
```
df.describe()
```

Out[67]:

	S.No.	Mar Cap - Crore	Sales Qtr - Crore
count	488.000000	488.000000	488.000000
mean	251.508197	27582.297520	3582.700697
std	145.884078	59008.934601	9728.843354
min	1.000000	3017.070000	19.420000
25%	122.750000	4643.832500	459.357500
50%	252.500000	9460.155000	982.300000
75%	378.250000	23400.815000	2580.797500
max	500.000000	583436.720000	110666.930000

In [68]:
```
# Analyzing Data
```

In [69]:
```
import cufflinks as cf
```

In [70]:
```
cf.go_offline()
```

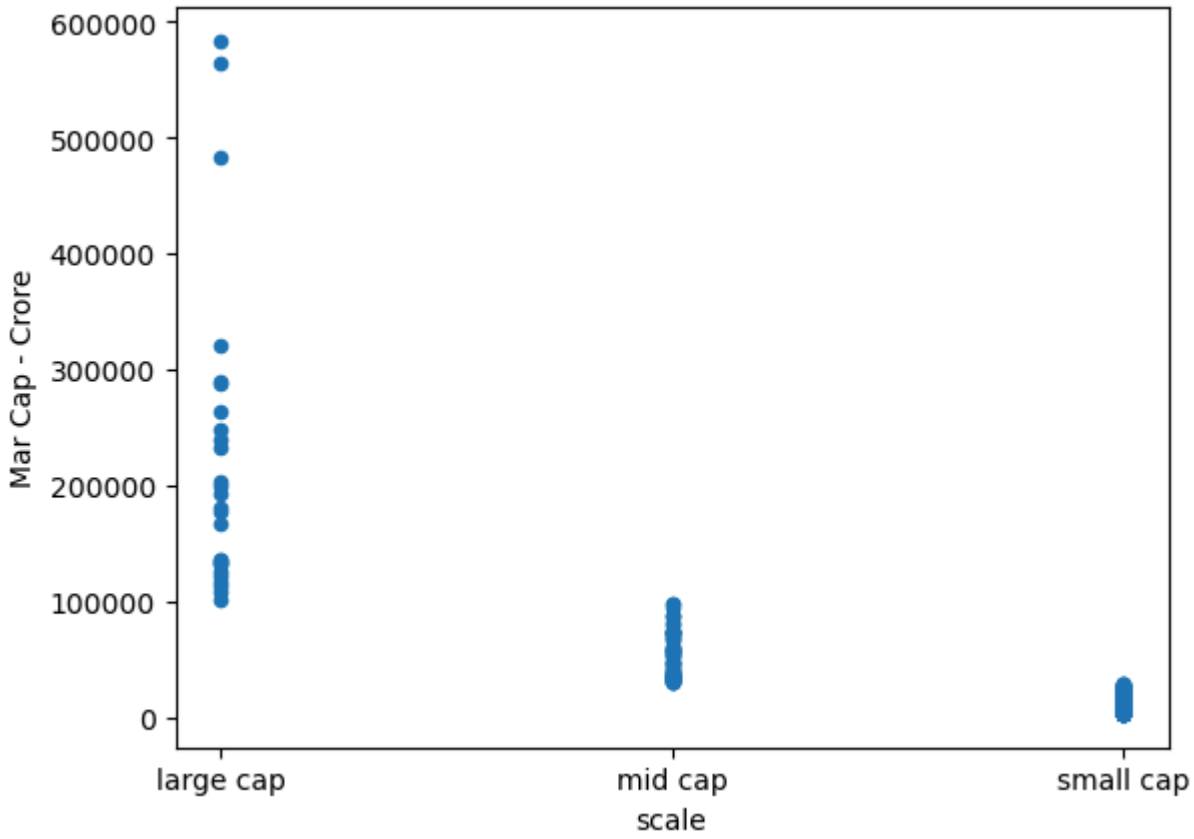
In [ ]:
```

```

In [71]:
```
#1.Comparing the performanance of different companies based on their market cap.

df.plot.scatter(x= 'scale',y='Mar Cap - Crore')
```

Out[71]:
```
<Axes: xlabel='scale', ylabel='Mar Cap - Crore'>
```



In [109…
```
#10.which companies are most in number based on  large cap,mid cap and small cap
df.groupby('Name')['scale'].sum().sort_values(ascending=False)
```

Out[109]:
```
Name
3M India            small cap
MphasiS             small cap
Natco Pharma        small cap
Narayana Hrudaya    small cap
NOCIL               small cap
                      ...
Coal India          large cap
H D F C             large cap
Hind.Zinc           large cap
I O C L             large cap
O N G C             large cap
Name: scale, Length: 488, dtype: object
```

In [73]:
```
# it means most companies are small cap.
```

In [74]:
```
#2Q. company with maximum Sales Qtr.
```

In [75]:
```
df.groupby('Name')['Sales Qtr - Crore'].sum().sort_values(ascending=False).head(3)
```

Out[75]:
```
Name
I O C L           110666.93
Reliance Inds.     99810.00
Tata Motors        74156.07
Name: Sales Qtr - Crore, dtype: float64
```

In [76]:
```
# IOCL is the company having maximum sales Quaterly.
```

In [77]:
```
#3Q. Top company based on market cap.
```

In [78]:
```
df.groupby('Name')['Mar Cap - Crore'].sum().sort_values(ascending=False).head(3)
```

Out[78]:
```
Name
Reliance Inds.    583436.72
TCS               563709.84
HDFC Bank         482953.59
Name: Mar Cap - Crore, dtype: float64
```

In [79]:
```
# Reliance INds. is the top company based on market capitalisation.
```

In [80]:
```
#4Q. company having min. sales QTR
```

In [81]:
```
df.groupby('Name')['Sales Qtr - Crore'].sum().sort_values(ascending=True).head(32)
```

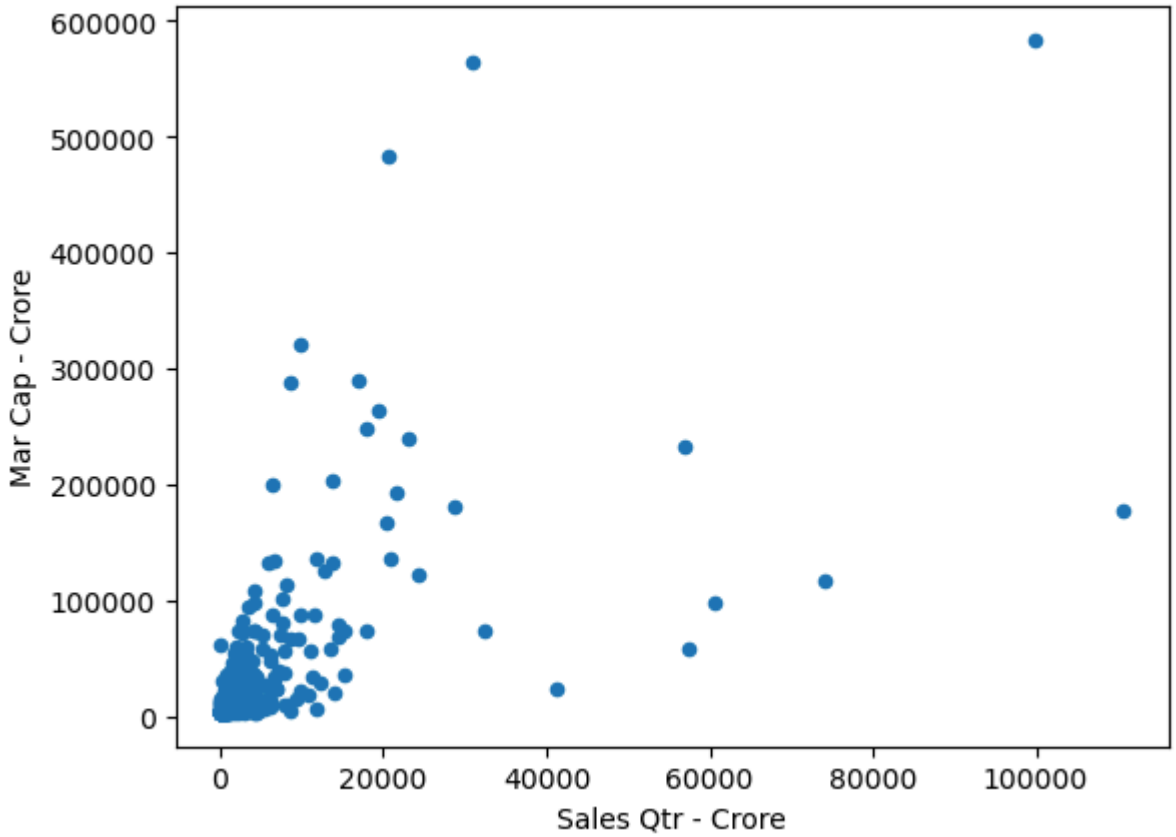
Out[81]:
```
Name
BSE                 19.42
ISGEC Heavy         19.42
Stand.Chart.PLC     19.42
Endurance Tech.     19.42
Shoppers St.        19.42
SPARC               19.42
Bajaj Corp          19.42
Info Edg.(India)    19.42
JP Associates       19.42
JP Power Ven.       19.42
Sundaram Clayton    19.42
Jindal Saw          19.42
Bharti Infra.       19.42
L T Foods           19.42
Bombay Burmah       19.42
Linde India         19.42
MMTC                19.42
Colgate-Palm.       19.42
Mahanagar Gas       19.42
Natl. Aluminium     19.42
Mahindra CIE        19.42
Prism Cement        19.42
Swan Energy         19.42
Max Financial       19.42
Force Motors        19.42
Gayatri Projects    19.42
Hind.Construct.     19.42
GE Power            19.42
HMT                 19.42
Ujjivan Fin.Ser.    19.42
Amber Enterp.       19.42
Tata Inv.Corpn.     47.02
Name: Sales Qtr - Crore, dtype: float64
```

In [82]:
```
# there are some companies whose value is taken as mode of sales quater. data
# we can consider them or can consider Tata Inv.Corpn. as a  company having min. sales QTR
```

In [83]:
```
#5Q . relation between market cap and sales Qtr
```

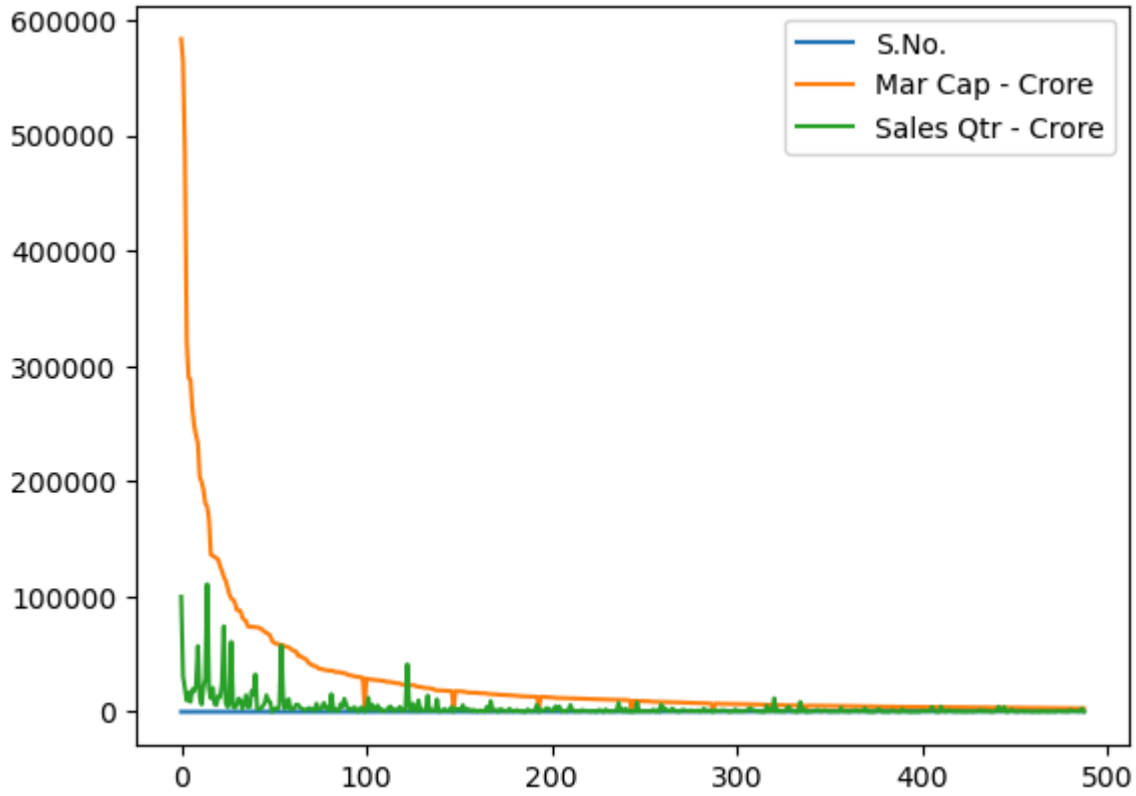
In [111…
```
df.plot(x='Sales Qtr - Crore',y='Mar Cap - Crore',kind='scatter')
```

Out[111]: <Axes: xlabel='Sales Qtr - Crore', ylabel='Mar Cap - Crore'>



In [85]:
```
df.plot()
```

Out[85]: <Axes: >



In [86]:
```
#6Q. Top 10 companies based on marketcap
```

In [87]:
```
df.groupby('Name')['Mar Cap - Crore'].sum().sort_values(ascending=False).head(10)
```

Out[87]:
```
Name
Reliance Inds.    583436.72
TCS               563709.84
HDFC Bank         482953.59
ITC               320985.27
H D F C           289497.37
Hind. Unilever    288265.26
Maruti Suzuki     263493.81
Infosys           248320.35
O N G C           239981.50
St Bk of India    232763.33
Name: Mar Cap - Crore, dtype: float64
```

In [88]:
```
#7.TOP 10 companies based on sales Qtr
```

In [89]:
```
df.groupby('Name')['Sales Qtr - Crore'].sum().sort_values(ascending=False).head(10)
```

Out[89]:
```
Name
I O C L            110666.93
Reliance Inds.      99810.00
Tata Motors         74156.07
B P C L             60616.36
H P C L             57474.25
St Bk of India      57014.08
Rajesh Exports      41304.84
Tata Steel          32464.14
TCS                 30904.00
Larsen & Toubro     28747.45
Name: Sales Qtr - Crore, dtype: float64
```

In [90]:
```
#8. Bottom 10 companies based on marketcap
```

In [91]:
```
df.groupby('Name')['Mar Cap - Crore'].sum().sort_values(ascending=True).iloc[10:22]
```

Out[91]:
```
Name
Orient Cement      3024.32
NOCIL              3026.26
Lak. Vilas Bank    3029.57
Prime Focus        3031.50
Va Tech Wabag      3041.93
Deepak Fert.       3079.06
Star Ferro Cem.    3115.98
Kaveri Seed Co.    3125.83
Firstsour.Solu.    3139.94
Navneet Educat.    3148.36
Mah. Seamless      3164.73
Heritage Foods     3185.45
Name: Mar Cap - Crore, dtype: float64
```

In [92]:
```
# only considered those Bottom 10 companies whose data is given in dataset.
```

In [93]:
```
#9.Bottom 10 companies based on sales Qtr
```

In [94]:
```
df.groupby('Name')['Sales Qtr - Crore'].sum().sort_values(ascending=True).iloc[31:43]
```

Out[94]:
```
Name
Tata Inv.Corpn.      47.02
Central Dep. Ser     47.24
Multi Comm. Exc.     60.97
Forbes & Co          63.93
Indian Energy Ex     64.75
La Opala RG          69.77
Kaveri Seed Co.      70.64
CARE Ratings         74.82
Thyrocare Tech.      77.84
NESCO                80.62
ICRA                 82.87
I T D C             102.14
Name: Sales Qtr - Crore, dtype: float64
```

In [95]:
```
# only considered those Bottom 10 companies whose data is given in dataset.
```

In [96]:
```
#11. Count of large cap ,mid cap and small cap companies
```

In [97]:
```
df.groupby('scale')['Name'].count()
```

Out[97]:
```
scale
large cap     27
mid cap       71
small cap    390
Name: Name, dtype: int64
```

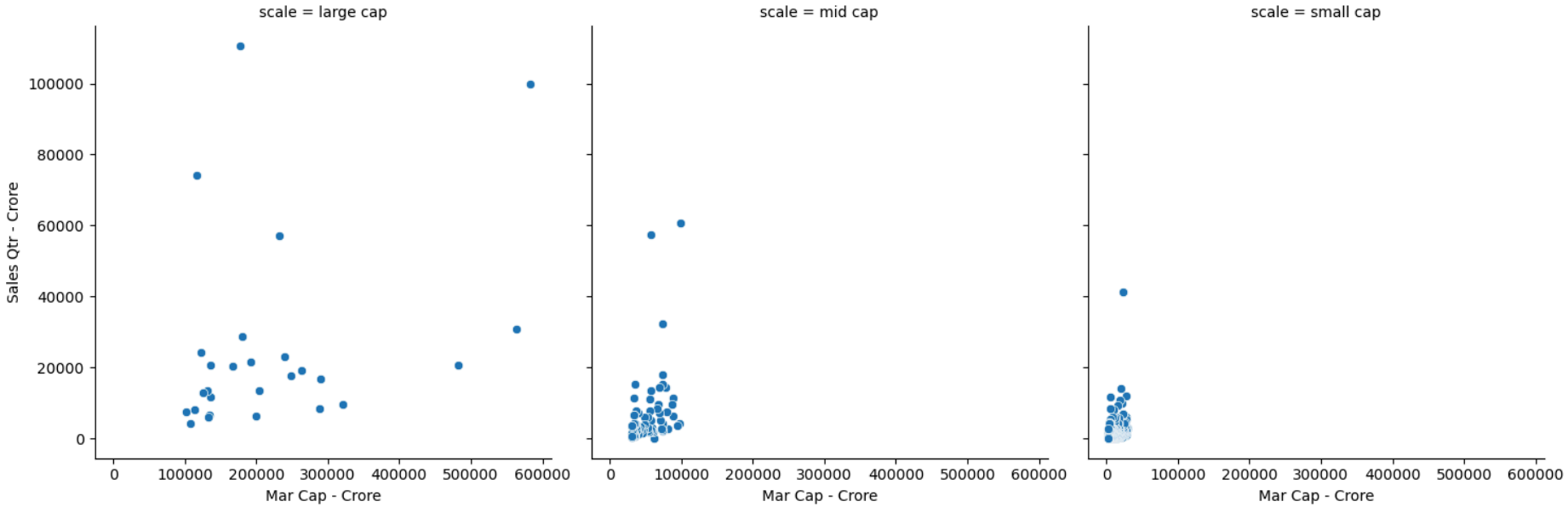
In [ ]:
```

```

In [98]:
```
sns.relplot(x='Mar Cap - Crore',y='Sales Qtr - Crore',data=df,col='scale')
plt.show()
```

```
D:\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning:

The figure layout has changed to tight
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: