

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
data=pd.read_csv('/content/Financial Analytics data.csv')
data.head()
```

	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

Next steps:

[Generate code with data](#)[View recommended plots](#)[New interactive sheet](#)

```
data.shape
data.size
```

2440

```
data.describe()
```

	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

```
col=list(data)
print(col)
data.isnull().sum()
```

```
['S.No.', 'Name', 'Mar Cap - Crore', 'Sales Qtr - Crore', 'Unnamed: 4']
```

	0
S.No.	0
Name	0
Mar Cap - Crore	9
Sales Qtr - Crore	123
Unnamed: 4	394

```
#EDA
clean_data=data.drop(columns=['Unnamed: 4'])
clean_data.info()
clean_data.isnull().sum()
```

```

<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      365 non-null   float64
dtypes: float64(2), int64(1), object(1)
memory usage: 15.4+ KB

```

```

0
S.No.    0
Name     0
Mar Cap - Crore    9
Sales Qtr - Crore 123

```

```

clean_data=clean_data.dropna(subset=['Mar Cap - Crore' , 'Sales Qtr - Crore'])
print(clean_data)

```

```

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
..      ...
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

```

```
[365 rows x 4 columns]
```

```

#correlation
cor=clean_data[['Mar Cap - Crore' , 'Sales Qtr - Crore']].corr()
cor

```

```

Mar Cap - Crore  Sales Qtr - Crore
Mar Cap - Crore    1.000000    0.620702
Sales Qtr - Crore    0.620702    1.000000

```

Next steps: [Generate code with cor](#) [View recommended plots](#) [New interactive sheet](#)

```

# top companies by market capitalization and Quarterly sales
top_marketcap=clean_data[['Name','Mar Cap - Crore']].sort_values(by='Mar Cap - Crore',ascending=False).head()
top_quarterlysale=clean_data[['Name','Sales Qtr - Crore']].sort_values(by='Sales Qtr - Crore',ascending=False).head()

```

```
print(top_marketcap)
```

```

Name  Mar Cap - Crore
0  Reliance Inds.      583436.72
1         TCS          563709.84
2  HDFC Bank          482953.59
3         ITC          320985.27
4   H D F C          289497.37

```

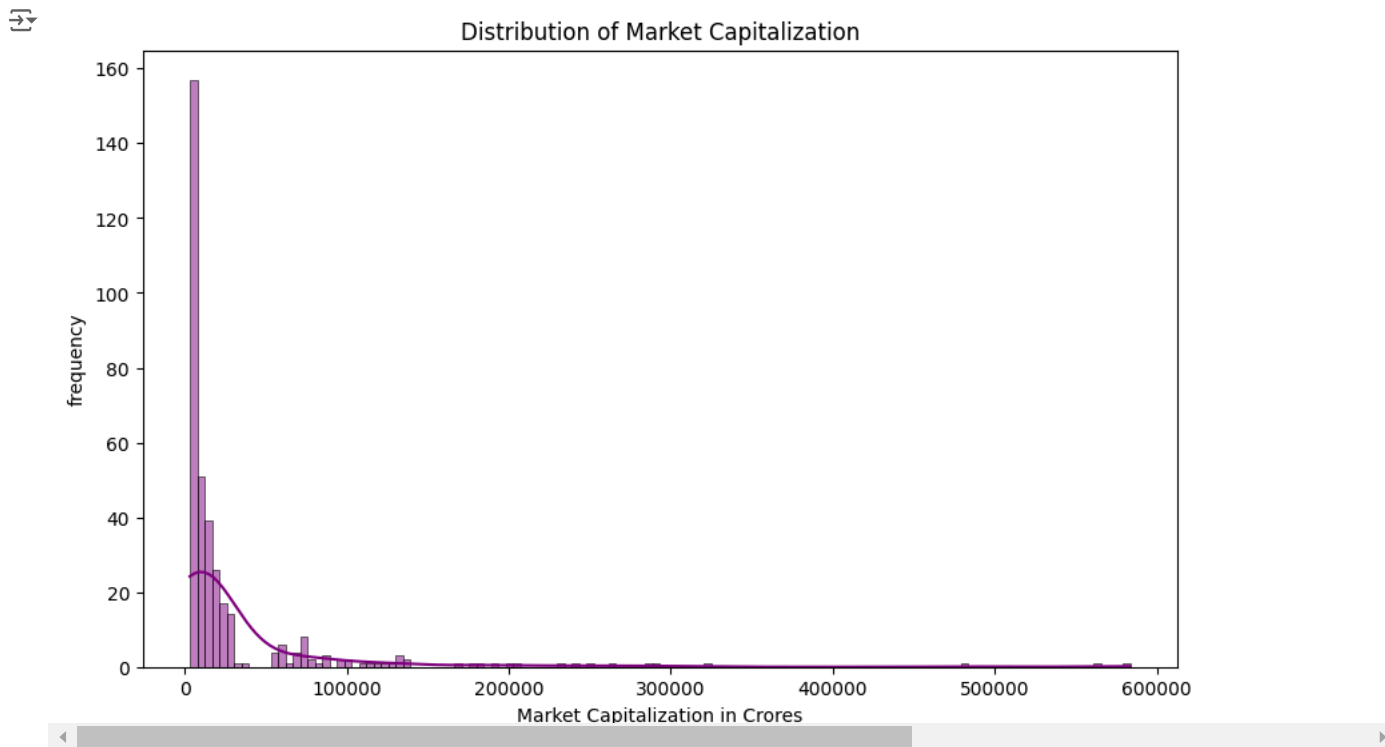
```
print(top_quarterlysale)
```

```

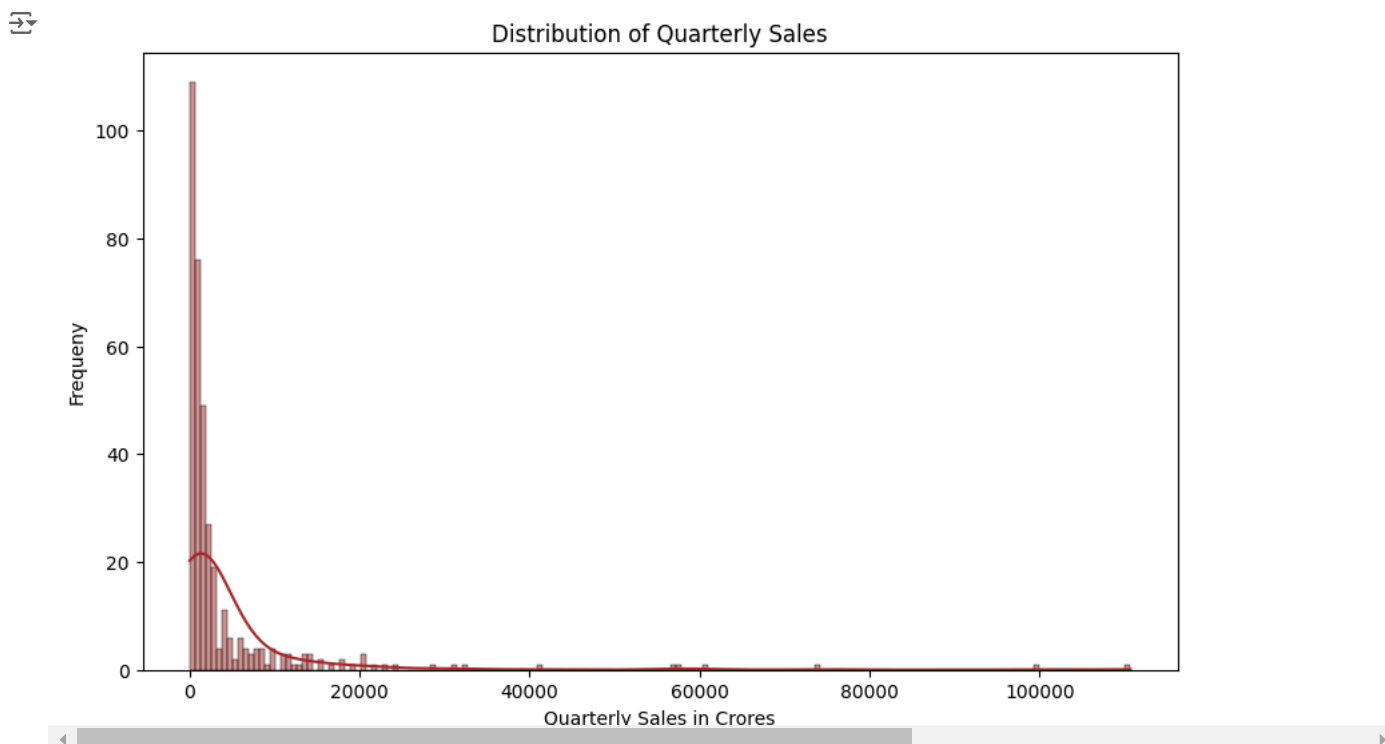
Name  Sales Qtr - Crore
14  I O C L          110666.93
0  Reliance Inds.      99810.00
23  Tata Motors       74156.07
27  B P C L           60616.36
54  H P C L           57474.25

```

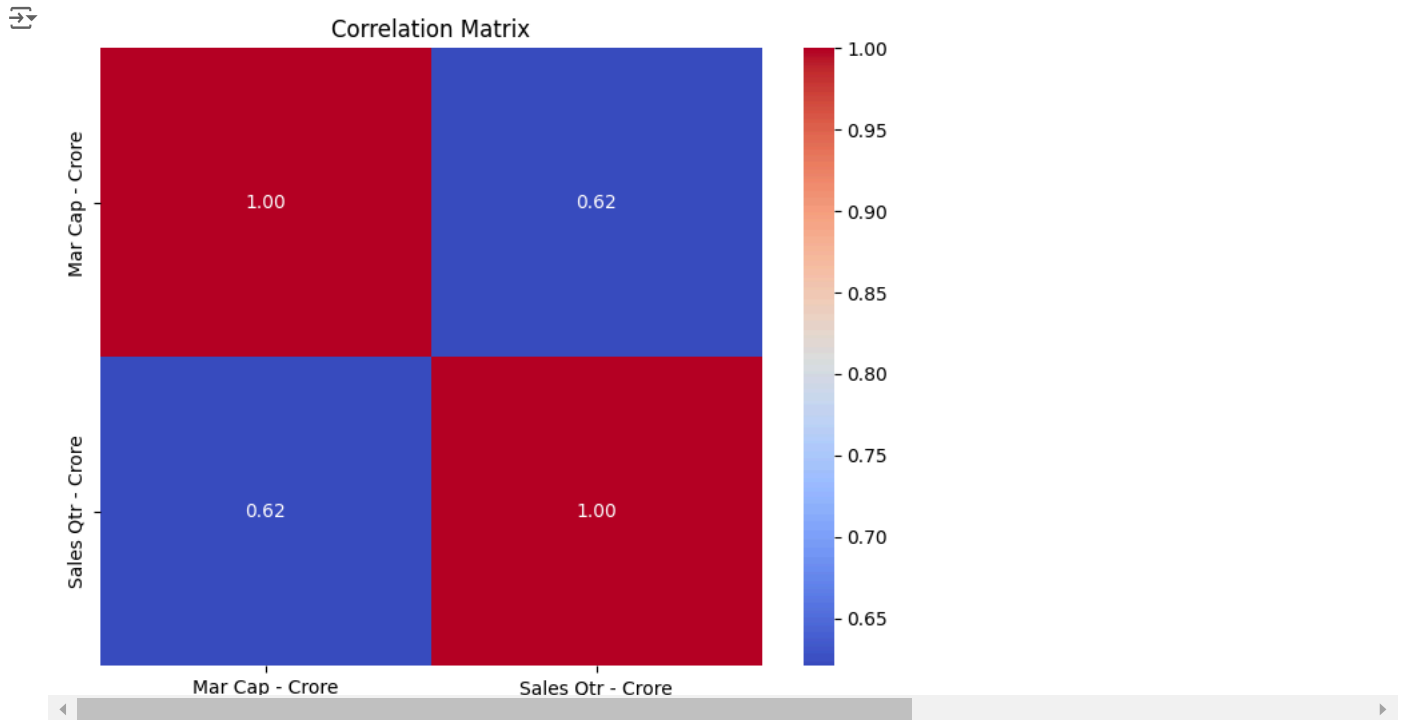
```
#visualization Market Capitalization
plt.figure(figsize=(10,6))
sns.histplot(clean_data['Mar Cap - Crore'],kde=True, color='purple')
plt.title('Distribution of Market Capitalization')
plt.xlabel('Market Capitalization in Crores')
plt.ylabel('frequency')
plt.show()
```



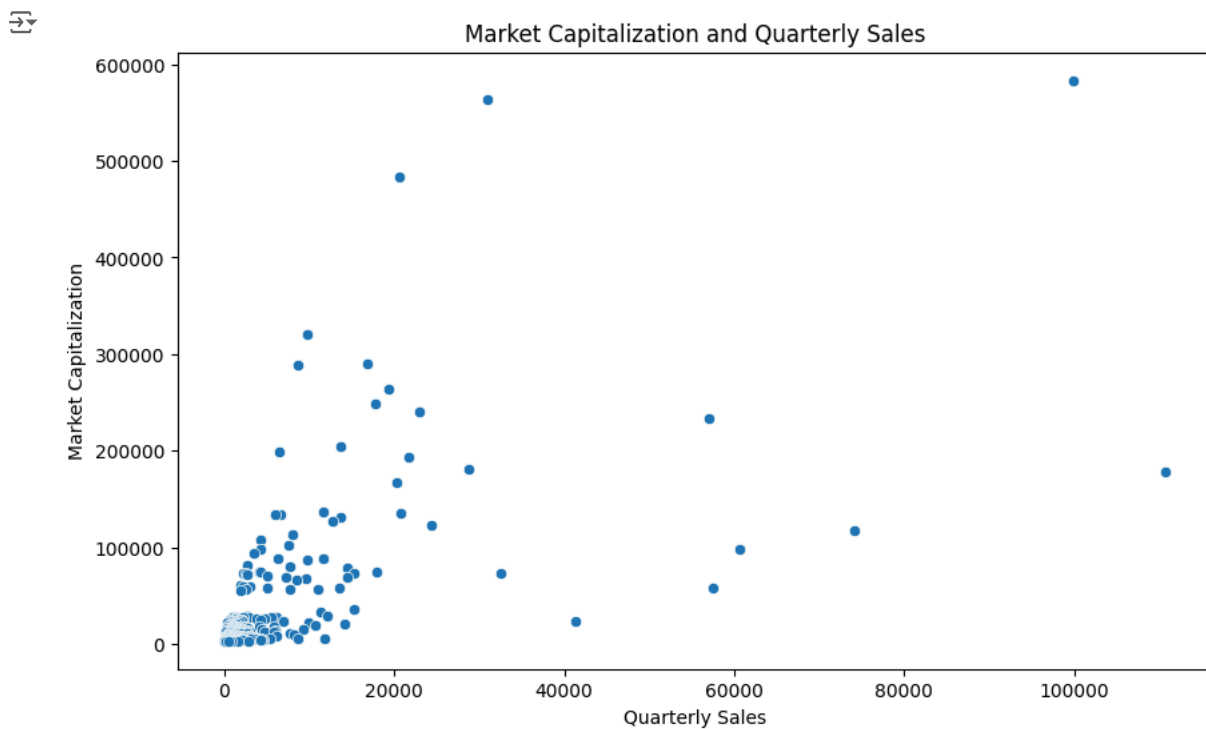
```
#visualization Quarterly Sales
plt.figure(figsize=(10,6))
sns.histplot(clean_data['Sales Qtr - Crore'],kde=True, color='brown')
plt.title('Distribution of Quarterly Sales')
plt.xlabel('Quarterly Sales in Crores')
plt.ylabel('Frequency')
plt.show()
```



```
#heatmap of correlation matrix
plt.figure(figsize=(8,6))
sns.heatmap(cor,annot=True,cmap='coolwarm',fmt='.2f')
plt.title('Correlation Matrix')
plt.show()
```



```
#scatterplot Market Capitalization and Quarterly Sales
plt.figure(figsize=(10,6))
sns.scatterplot(x='Sales Qtr - Crore',y='Mar Cap - Crore',data=clean_data)
plt.title('Market Capitalization and Quarterly Sales')
plt.xlabel('Quarterly Sales')
plt.ylabel('Market Capitalization')
plt.show()
```



Double-click (or enter) to edit

Double-click (or enter) to edit