

```
In [1]: import pandas as pd
import numpy as np
from matplotlib import pyplot
import seaborn as sns
import missingno as msno
```

```
In [2]: data = pd.read_csv("file:///C:/Users/Vasu%20Prasad/OneDrive/Documents/MCA/Internships/Unified%20Mentors%20Inte
```

```
In [3]: data
```

```
Out[3]:
```

	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
...
483	496	Lak. Vilas Bank	3029.57	790.17	NaN
484	497	NOCIL	3026.26	249.27	NaN
485	498	Orient Cement	3024.32	511.53	NaN
486	499	Natl.Fertilizer	3017.07	2840.75	NaN
487	500	L T Foods	NaN	NaN	NaN

488 rows × 5 columns

```
In [4]: pd.concat([data.head(),data.tail()])
```

```
Out[4]:
```

	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
483	496	Lak. Vilas Bank	3029.57	790.17	NaN
484	497	NOCIL	3026.26	249.27	NaN
485	498	Orient Cement	3024.32	511.53	NaN
486	499	Natl.Fertilizer	3017.07	2840.75	NaN
487	500	L T Foods	NaN	NaN	NaN

```
In [5]: data.describe()
```

```
Out[5]:
```

	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 [6]: data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 488 entries, 0 to 487
Data columns (total 5 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 
4   Unnamed: 4            94 non-null    float64 
dtypes: float64(3), int64(1), object(1)
memory usage: 19.2+ KB
```

In [7]: data.shape

Out[7]: (488, 5)

In [8]: data.size

Out[8]: 2440

In [9]: data.index

Out[9]: RangeIndex(start=0, stop=488, step=1)

In [10]: data.columns

Out[10]: Index(['S.No.', 'Name', 'Mar Cap - Crore', 'Sales Qtr - Crore', 'Unnamed: 4'], dtype='object')

```
In [17]: data.nunique().reset_index().rename(columns = {0:"Count"})
```

```
Out[17]:
```

	index	Count
0	S.No.	488
1	Name	488
2	Mar Cap - Crore	479
3	Sales Qtr - Crore	365
4	Unnamed: 4	94

```
In [18]: data.duplicated().sum()
```

```
Out[18]: 0
```

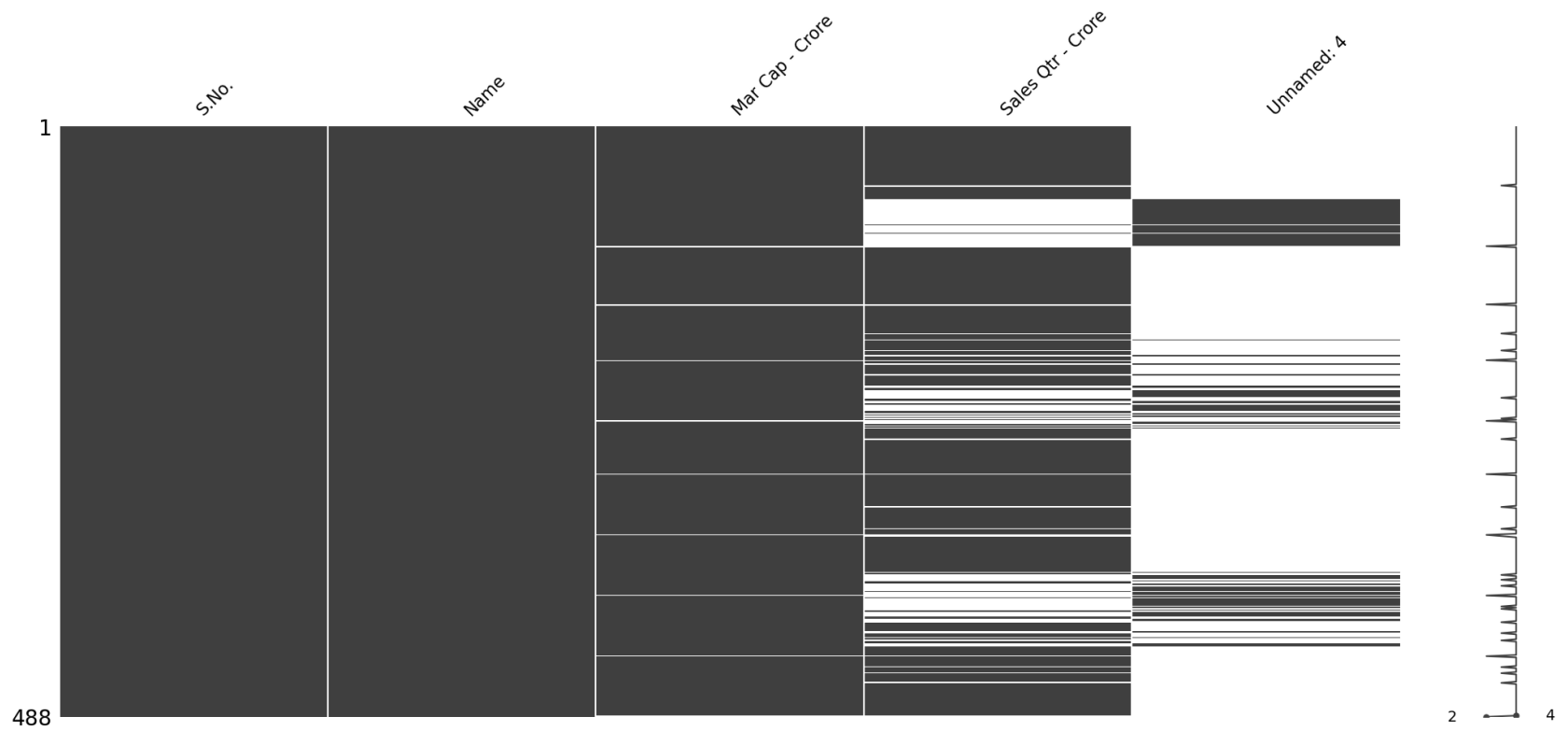
```
In [19]: data.isnull().sum().reset_index().rename(columns = {0:"Count"})
```

```
Out[19]:
```

	index	Count
0	S.No.	0
1	Name	0
2	Mar Cap - Crore	9
3	Sales Qtr - Crore	123
4	Unnamed: 4	394

```
In [23]: msno.matrix(data)
```

```
Out[23]: <Axes: >
```



```
data["Name"].unique()
```

```
array(['Reliance Inds.', 'TCS', 'HDFC Bank', 'ITC', 'H D F C',  
      'Hind. Unilever', 'Maruti Suzuki', 'Infosys', 'O N G C',  
      'St Bk of India', 'ICICI Bank', 'Kotak Mah. Bank', 'Coal India',  
      'Larsen & Toubro', 'I O C L', 'Bharti Airtel', 'Axis Bank', 'NTPC',  
      'Sun Pharma.Inds.', 'Hind.Zinc', 'Wipro', 'HCL Technologies',  
      'Vedanta', 'Tata Motors', 'UltraTech Cem.', 'Asian Paints',  
      'Power Grid Corpn', 'B P C L', 'IndusInd Bank', 'Bajaj Fin.',  
      'Bajaj Auto', 'M & M', 'HDFC Stand. Life', 'Adani Ports',  
      'Bajaj Finserv', 'GAIL (India)', 'Avenue Super.', 'Titan Company',  
      'JSW Steel', 'Grasim Inds', 'Tata Steel', 'Eicher Motors',  
      'Nestle India', 'Godrej Consumer', 'Yes Bank', 'Hero Motocorp',  
      'Motherson Sumi', 'SBI Life Insuran', 'General Insuranc',  
      'Bharti Infra.', 'Dabur India', 'Bosch', 'Shree Cement',  
      'New India Assura', 'H P C L', 'ICICI Pru Life', 'Britannia Inds.',  
      'Tech Mahindra', 'Hindalco Inds.', 'Zee Entertainmen',  
      'Cairn India', 'Indiabulls Hous.', 'Ambuja Cem.',  
      'Interglobe Aviat', 'Cipla', 'Piramal Enterp.', 'United Spirits',  
      'Pidilite Inds.', 'Siemens', 'Cadila Health.', 'NMDC', 'DLF',  
      'Marico', 'Ashok Leyland', 'Bharat Electron', 'ICICI Lombard',  
      'Himachal Pradesh', 'ONGC', 'Aditya Birla Gro', 'Dr. Reddy's Labs']
```

```
data.columns
```

```
Index(['S.No.', 'Name', 'Mar Cap - Crore', 'Sales Qtr - Crore', 'Unnamed: 4'], dtype='object')
```

```
data.drop("Unnamed: 4",axis = 1,inplace = True)
```

In [30]: data

Out[30]:

	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 [31]: data.dropna(inplace = True)

In [32]: data.reset_index(drop = True,inplace = True)

In [33]:

data

Out[33]:

	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
...
360	495	Prime Focus	3031.50	609.61
361	496	Lak. Vilas Bank	3029.57	790.17
362	497	NOCIL	3026.26	249.27
363	498	Orient Cement	3024.32	511.53
364	499	Natl.Fertilizer	3017.07	2840.75

365 rows × 4 columns

In []: