

```
In [31]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

```
In [32]: df= pd.read_csv(r"C:\Users\del1\Downloads\Yahoo Finance_industry - Stock Market.csv")
```

```
In [33]: df.head(18)
```

```
Out[33]:
```

	Symbol	Name	Price	Change	market cap	PE_ratio
0	BRK-B	Berkshire Hathway Inc.	321.26	-1.38	709.361B	NaN
1	V	Visa Inc.	230.76	-0.25	483.304B	30.81
2	JPM	JPMorgan Chase & Co.	133.89	-2.16	392.325B	9.90
3	MA	Mastercard Incorporated	380.21	-3.18	359.338B	37.85
4	JPM-PD	JPMorgan Chase & Co.	24.57	-0.08	326.427B	1.82
5	JPM-PD	JPMorgan Chase & Co.	25.02	-0.09	328.207B	1.86
6	BAC-PC	Bank of America Corporation	24.38	-0.04	244.596B	6.94
7	BAC-PC	Bank of America Corporation	1141.76	0.78	325.77	3
8	BAC-PC	Bank of America Corporation	20.02	-0.03	270.350	5.70
9	BWL-PH	Bank of America Corporation	19.90	0.05	269.612B	5.66

```
In [34]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1258 entries, 0 to 1249
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype
---  --
 0   Symbol      1258 non-null   object
 1   Name        1258 non-null   object
 2   Price       1258 non-null   float64
 3   Change      1258 non-null   float64
 4   market cap  1258 non-null   object
 5   PE_ratio    1258 non-null   float64
dtypes: float64(2), object(4)
memory usage: 88.7+ KB
```

```
In [35]: df['PE_ratio'] = df['PE_ratio'].fillna(df['PE_ratio'].mean())
```

```
In [36]: df.drop_duplicates(inplace = True)
```

```
In [37]: df
```

```
Out[37]:
```

	Symbol	Name	Price	Change	market cap	PE_ratio
0	BRK-B	Berkshire Hathway Inc.	321.26	-1.38	709.361B	21.220227
1	V	Visa Inc.	230.76	-0.25	483.304B	30.810000
2	JPM	JPMorgan Chase & Co.	133.89	-2.16	392.325B	9.900000
3	MA	Mastercard Incorporated	380.21	-3.18	359.338B	37.850000
4	JPM-PD	JPMorgan Chase & Co.	24.57	-0.08	326.427B	1.820000
...
1239	BN	Brookfield Corporation	30.25	-0.95	49.475B	25.430000
1243	US	U.S. Bancorp	29.10	0.24	48.480B	7.940000
1244	PHC	The PNC Financial Services Group Inc.	111.18	1.82	48.538B	7.700000
1246	BBVA	Banco Bilbao Vizcaya Argentaria, S.A.	7.16	0.04	42.965B	6.970000
1249	TRV	The Travelers Companies, Inc.	181.84	-0.97	42.44B	15.000000

438 rows x 6 columns

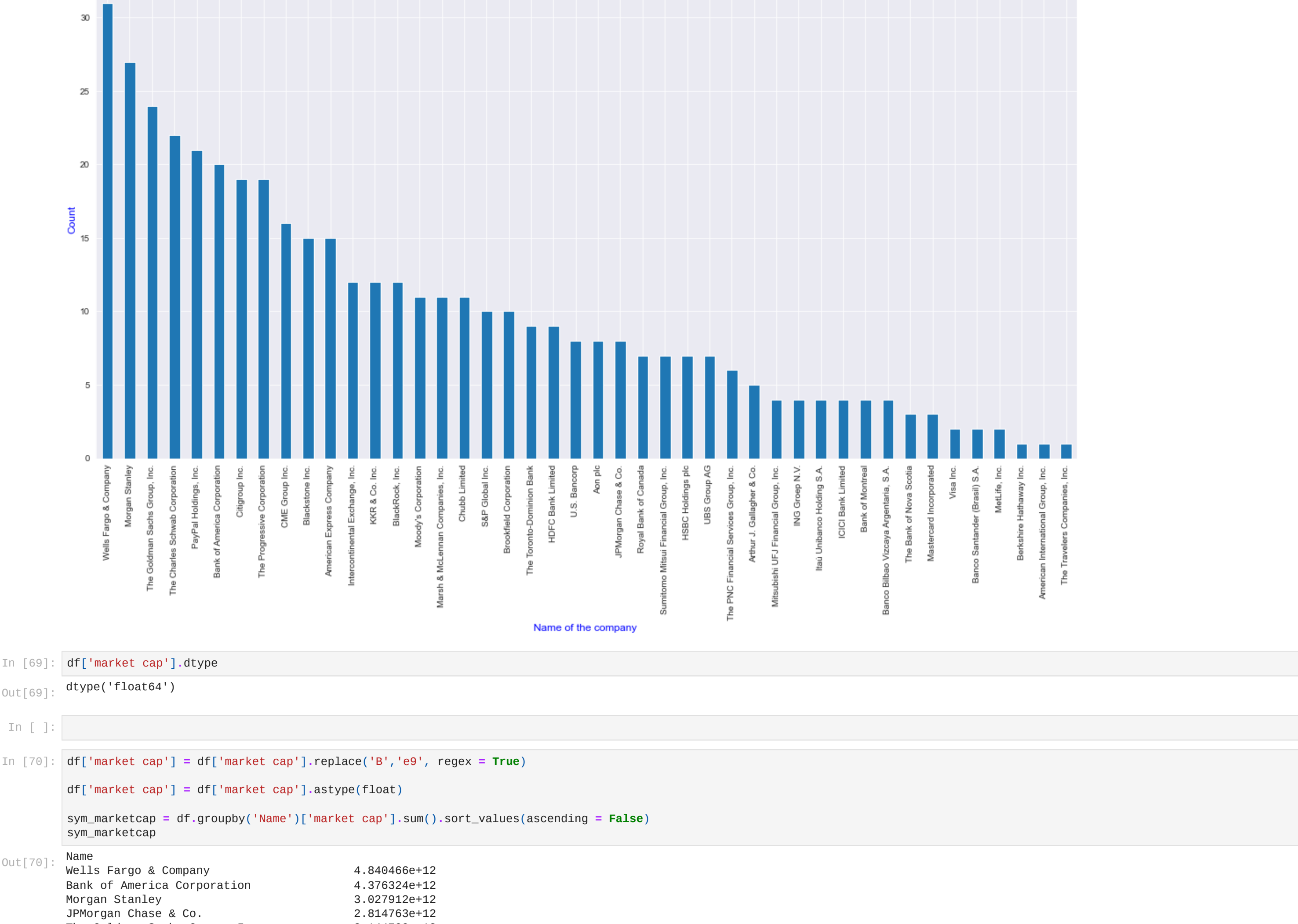
Eda

```
In [38]: df['Name'].value_counts()
```

```
Out[38]:
```

Wells Fargo & Company	31
Morgan Stanley	27
The Goldman Sachs Group, Inc.	24
The Charles Schwab Corporation	22
Paycom Holdings, Inc.	21
Bank of America Corporation	20
Citigroup Inc.	19
The Progressive Corporation	19
ONE Group Inc.	16
Blackstone Inc.	15
American Express Company	15
Intercontinental Exchange, Inc.	12
KKR & Co. Inc.	12
BlackRock, Inc.	12
SP Global Inc.	11
Marsh & McLennan Companies, Inc.	11
Chubb Limited	11
SAP Global Inc.	10
Brookfield Corporation	10
The Toronto-Dominion Bank	9
HSBC Bank Limited	9
U.S. Bancorp	8
Aon plc	8
JPMorgan Chase & Co.	8
Royal Bank of Canada	8
Sun Life of Canada	7
HSBC Holdings plc	7
USG Group AG	7
The PNC Financial Services Group, Inc.	6
Arthur J. Gallagher & Co.	6
Mitsubishi UFJ Financial Group, Inc.	4
ING Groep N.V.	4
Itau Unibanco Holding S.A.	4
ICICI Bank Limited	4
Bank of Montreal	4
Banco Bilbao Vizcaya Argentaria, S.A.	4
The Bank of Nova Scotia	4
Mastercard Incorporated	3
Visa Inc.	2
Banco Santander (Brasil) S.A.	2
MetLife, Inc.	2
Berkshire Hathway Inc.	1
American International Group, Inc.	1
The Travelers Companies, Inc.	1

Name: Name, dtype: object



```
In [69]: df['market cap'].dtype
Out[69]: dtype('float64')
```

```
In [ ]:
```

```
In [70]: df['market cap'] = df['market cap'].replace('B','e9', regex = True)
df['market cap'] = df['market cap'].astype(float)
sym_marketcap = df.groupby('Name')['market cap'].sum().sort_values(ascending = False)
sym_marketcap
```

```
Out[70]:
```

Name	market cap
Wells Fargo & Company	4.840466e+12
Bank of America Corporation	4.376324e+12
Morgan Stanley	3.027912e+12
JPMorgan Chase & Co.	3.847638e+12
The Goldman Sachs Group, Inc.	2.144789e+12
The Charles Schwab Corporation	1.931544e+12
Citigroup Inc.	1.889686e+12
American Express Company	1.858892e+12
Paycom Holdings, Inc.	1.858892e+12
The Progressive Corporation	1.584761e+12
Blackstone Inc.	1.483577e+12
BlackRock, Inc.	1.381366e+12
SP Global Inc.	1.148599e+12
HSBC Bank Limited	1.078386e+12
ONE Group Inc.	1.072564e+12
HSBC Holdings plc	1.058249e+12
The Toronto-Dominion Bank	1.087184e+12
Marsh & McLennan Companies, Inc.	9.815240e+11
Visa Inc.	9.667440e+11
Chubb Limited	9.094340e+11
Royal Bank of Canada	8.949340e+11
Intercontinental Exchange, Inc.	7.277630e+11
Berkshire Hathway Inc.	7.089310e+11
KKR & Co. Inc.	6.555466e+11
Moody's Corporation	6.269800e+11
Aon plc	5.479950e+11
Brookfield Corporation	4.952770e+11
USG Group AG	4.166860e+11
U.S. Bancorp	4.147160e+11
Sun Life of Canada	3.982740e+11
Sun Life of Canada	3.283680e+11
Mitsubishi UFJ Financial Group, Inc.	3.121880e+11
The PNC Financial Services Group, Inc.	2.731610e+11
Bank of Montreal	2.471120e+11
Arthur J. Gallagher & Co.	2.323400e+11
Itau Unibanco Holding S.A.	2.128930e+11
ING Groep N.V.	1.932240e+11
The Bank of Nova Scotia	1.754180e+11
Banco Santander (Brasil) S.A.	1.719830e+11
Banco Santander (Brasil) S.A.	9.080300e+10
MetLife, Inc.	8.487580e+10
American International Group, Inc.	4.211300e+10
The Travelers Companies, Inc.	4.224800e+10

Name: market cap, dtype: float64

