

Importing dependencies

```
In [2]:  
import numpy as np  
import pandas as pd  
import matplotlib.pyplot as plt  
  
import seaborn as sns  
import scipy as sp  
  
  
import warnings  
warnings.filterwarnings("ignore")
```

Importing Datasets

```
In [3]: df1 = pd.read_csv('D:/Downloads_D/siddhardhan/Scouto/Table_1.csv')
```

Importing Table 1

```
In [4]: df1.head()
```

Out[4]:

	Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets	M
0	Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marketing	1.679534e+06	2467.40	34.186904	7.793277	3.957021	9.7%
1	Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1.358569e+06	3626.70	41.892356	37.740680	25.932382	19.3%
2	HDFC Bank Ltd	HDFCBANK	Private Banks	8.206363e+05	1486.50	25.779258	16.435200	1.883433	20.4%
3	Infosys Ltd	INFY	IT Services & Consulting	7.775010e+05	1853.05	40.178853	27.135305	19.417895	18.8%
4	ICICI Bank Ltd	ICICIBANK	Private Banks	5.002790e+05	710.75	27.212283	12.382897	1.243333	11.3%

◀ ▶

```
In [5]: df1.shape
```

Out[5]: (2014, 15)

```
In [ ]:
```

Importing Table 2

```
In [6]:
```


Firstly, screen the companies and visualise them according to sub-sectors using a pie chart or a bar charts.

In [12]:

```
# Firstly checking if there is any null string or value in the 'sub-sectors' feature
# Secondly check for any 0 values or not
```

In [13]:

```
# so there were 67 null values
df1['Sub-Sector'].isnull().sum()
```

Out[13]: 67

In [14]:

```
df1[df1['Sub-Sector'] == '0']
```

Out[14]:

		Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Retu
1214		Cosmo Ferrites Ltd	COSMOFE	0	614.612700	485.40	3.414515e+02	1.079137e+01	2.26045
1420		Cressanda Solutions Ltd	CRESSAN	0	391.007820	13.47	NaN	0.000000e+00	0.00000
1478		Saraswati Commercial (India) Ltd	ZSARACOM	0	332.342317	3235.00	1.557368e+01	9.719439e+00	9.13801
1504		Ikab Securities and Investment Ltd	IKAB	0	313.728012	883.50	1.166275e+02	2.767490e+01	2.62055
1587		MK Exim (India) Ltd	MKEXIM	0	271.402477	105.85	3.379857e+01	2.351735e+01	1.83752
1630		BLS Infotech Ltd	BLSINFOTE	0	247.735264	5.38	8.930000e+18	6.820000e-17	6.79000
1698		Swiss Military Consumer Goods Ltd	SWISSMLTRY	0	221.659726	23.05	1.108299e+04	4.926108e-01	4.65657
1823		Kaiser Corporation Ltd	KACL	0	177.332837	35.35	1.886520e+02	1.115065e+01	3.92484
1856		Diligent Industries Ltd	DILIGENT	0	167.537400	72.00	3.490362e+02	3.953871e+00	9.86235
1886		Gayatri BioOrganics Ltd	GAYATRIBI	0	157.182343	19.70	-3.143647e+02	NaN	-1.76056
1914		Alka India Ltd	ALKA	0	150.000000	2.85	-4.545455e+02	-1.559546e+00	-1.34255
1943		Banas Finance Ltd	BANASFN	0	143.504644	54.20	1.229688e+01	1.339916e+01	1.18942

In [15]: `df1[df1['Sub-Sector'].isnull()]`

Out[15]:

		Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets
94		Adani Wilmar Ltd	AWL	NaN	49361.793420	391.25	67.837275	24.797022	5.795124
136		NaN	NaN	NaN	NaN	NaN	NaN	NaN	1.000000
276		NaN	NaN	NaN	NaN	NaN	NaN	NaN	2.000000
416		NaN	NaN	NaN	NaN	NaN	NaN	NaN	3.000000
452		Sindhu Trade Links Ltd	SINDHUTRAD	NaN	6355.316455	120.00	101.118798	9.698550	4.766618
...
1940		Markolines Traffic Controls Ltd	MTCL	NaN	144.261776	72.15	33.706022	19.270599	4.327166
1942		Shahlon Silk Industries Ltd	SHAHLON	NaN	143.776945	16.10	41.434278	3.809628	1.220306
1956		NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.000000
1957		Ravinder Heights Ltd	RVHL	NaN	139.039193	23.55	-29.519999	-1.740287	-1.484797
1988		Veer Global Infraconstruction Ltd	VGIL	NaN	131.542082	171.00	730.789347	1.413983	0.529334

67 rows × 15 columns

Feature Engineering

In [16]: `quotient = 79 / 2014`

`percent = quotient * 100`

`print("Only", percent, "% of data has either 0 values or null values")`

Only 3.922542204568024 % of data has either 0 values or null values

In [17]:

`# So 79 observations have either 0 values or null values, so we drop these observations
can not affect the data`

In [18]:

`df1.dropna(subset=['Sub-Sector'], inplace=True)
df1= df1[df1['Sub-Sector'] != 0]`

In [19]:

`df1['Sub-Sector'].isnull().sum()`

Out[19]: 0

In [20]: # So now there is no null values in the Sub-Sector variable

In [21]: df1['Sub-Sector'].value_counts()

```
Out[21]: Pharmaceuticals      104
Textiles                  86
Auto Parts                 84
Industrial Machinery       76
Commodity Chemicals        73
...
Mining - Copper             1
Payment Infrastructure       1
Mining - Iron Ore            1
Soft Drinks                  1
Metals - Iron                  1
Name: Sub-Sector, Length: 137, dtype: int64
```

In [22]: df1[df1['Sub-Sector'] == 'Pharmaceuticals']

Out[22]:

		Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity
	23	Sun Pharmaceutical Industries Ltd	SUNPHARMA	Pharmaceuticals	218723.375900	913.80	75.322636	5.889831
	54	Cipla Ltd	CIPLA	Pharmaceuticals	84594.451670	1050.30	35.176310	13.883780
	71	Dr Reddy's Laboratories Ltd	DRREDDY	Pharmaceuticals	67080.243220	4004.95	34.371922	11.742302
	88	Gland Pharma Ltd	GLAND	Pharmaceuticals	54222.363130	3282.65	54.387156	20.880090
	97	Torrent Pharmaceuticals Ltd	TORNTPHARM	Pharmaceuticals	48644.763090	2763.60	38.857369	23.486570

	1882	Bharat Immunologicals and Biologicals Corporat...	BIBCL	Pharmaceuticals	158.686500	36.60	-8.757533	-66.483214
	1884	Tyche Industries Ltd	TYCHE	Pharmaceuticals	158.003750	154.90	7.908096	23.078256
	1905	Alpa Laboratories Ltd	ALPA	Pharmaceuticals	152.228741	73.20	21.026069	6.961204
	1923	Panchsheel Organics Ltd	PANCHSHEEL	Pharmaceuticals	147.853992	144.50	29.570798	13.333333
	1955	Medico Remedies Ltd	MEDICO	Pharmaceuticals	139.828040	84.25	53.780015	8.342692

104 rows × 15 columns

In [23]:

```
l1 = []
for i in df1['Sub-Sector'].unique():
    l1.append(i)
l1
#Len(l)
```

Out[23]:

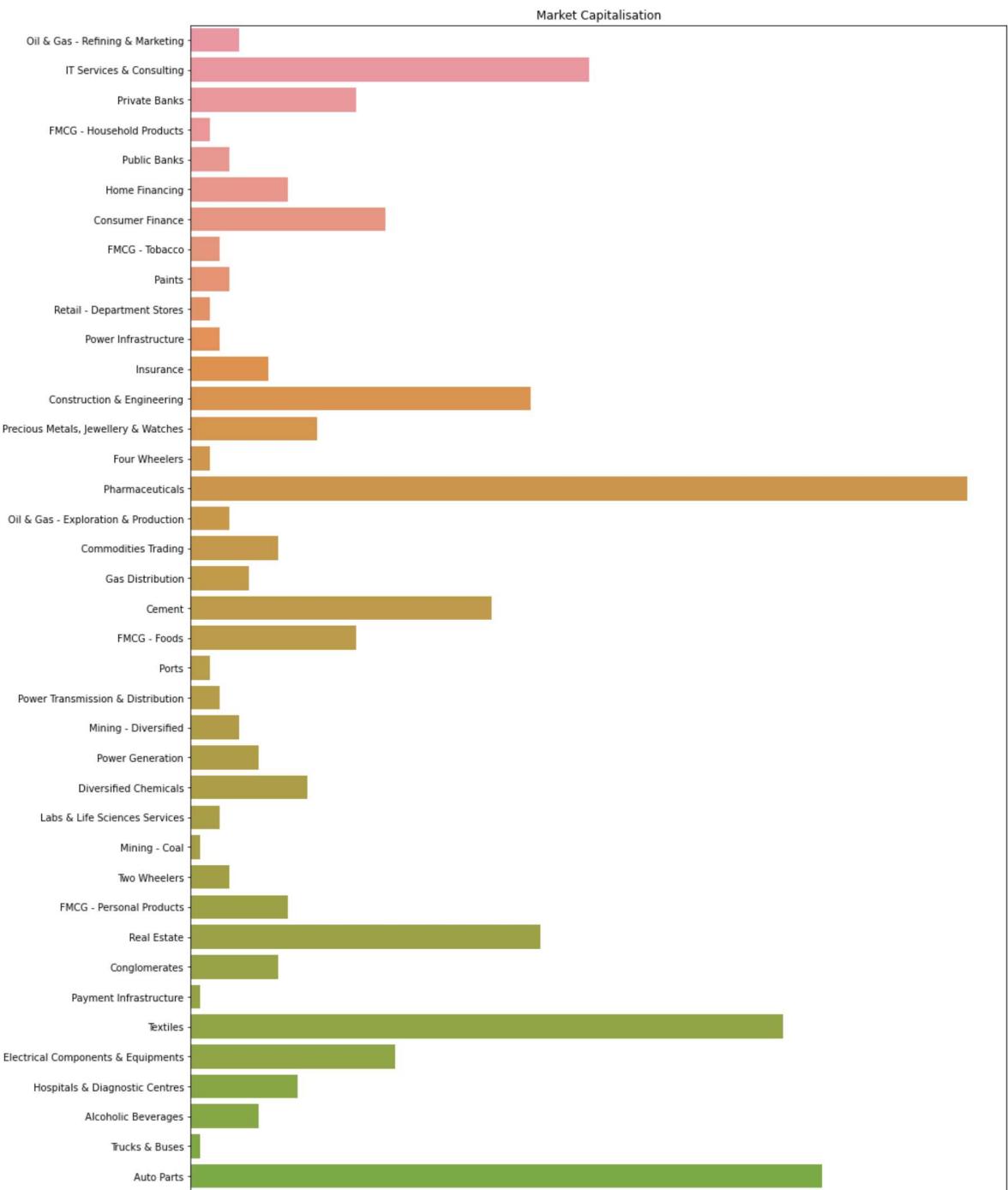
```
['Oil & Gas - Refining & Marketing',
 'IT Services & Consulting',
 'Private Banks',
 'FMCG - Household Products',
 'Public Banks',
 'Home Financing',
 'Consumer Finance',
 'Telecom Services',
 'FMCG - Tobacco',
 'Paints',
 'Renewable Energy',
 'Retail - Department Stores',
 'Power Infrastructure',
 'Insurance',
 'Construction & Engineering',
 'Precious Metals, Jewellery & Watches',
 'Four Wheelers',
 'Pharmaceuticals',
 'Oil & Gas - Exploration & Production',
 'Commodities Trading',
 'Gas Distribution',
 'Cement',
 'FMCG - Foods',
 'Iron & Steel',
 'Ports',
 'Power Transmission & Distribution',
 'Metals - Diversified',
 'Mining - Diversified',
 'Metals - Aluminium',
 'Power Generation',
 'Diversified Chemicals',
 'Labs & Life Sciences Services',
 'Mining - Coal',
 'Two Wheelers',
 'FMCG - Personal Products',
 'Real Estate',
 'Conglomerates',
 'Payment Infrastructure',
 'Textiles',
 'Business Support Services',
 'Airlines',
 'Electrical Components & Equipments',
 'Tea & Coffee',
 'Hospitals & Diagnostic Centres',
 'Alcoholic Beverages',
 'Trucks & Buses',
 'Auto Parts',
 'Online Services',
 'Fertilizers & Agro Chemicals',
 'Telecom Infrastructure',
 'Asset Management',
 'Diversified Financials',
 'Electronic Equipments',
 'Software Services',
 'Aerospace & Defense Equipments',
 'Apparel & Accessories',
```

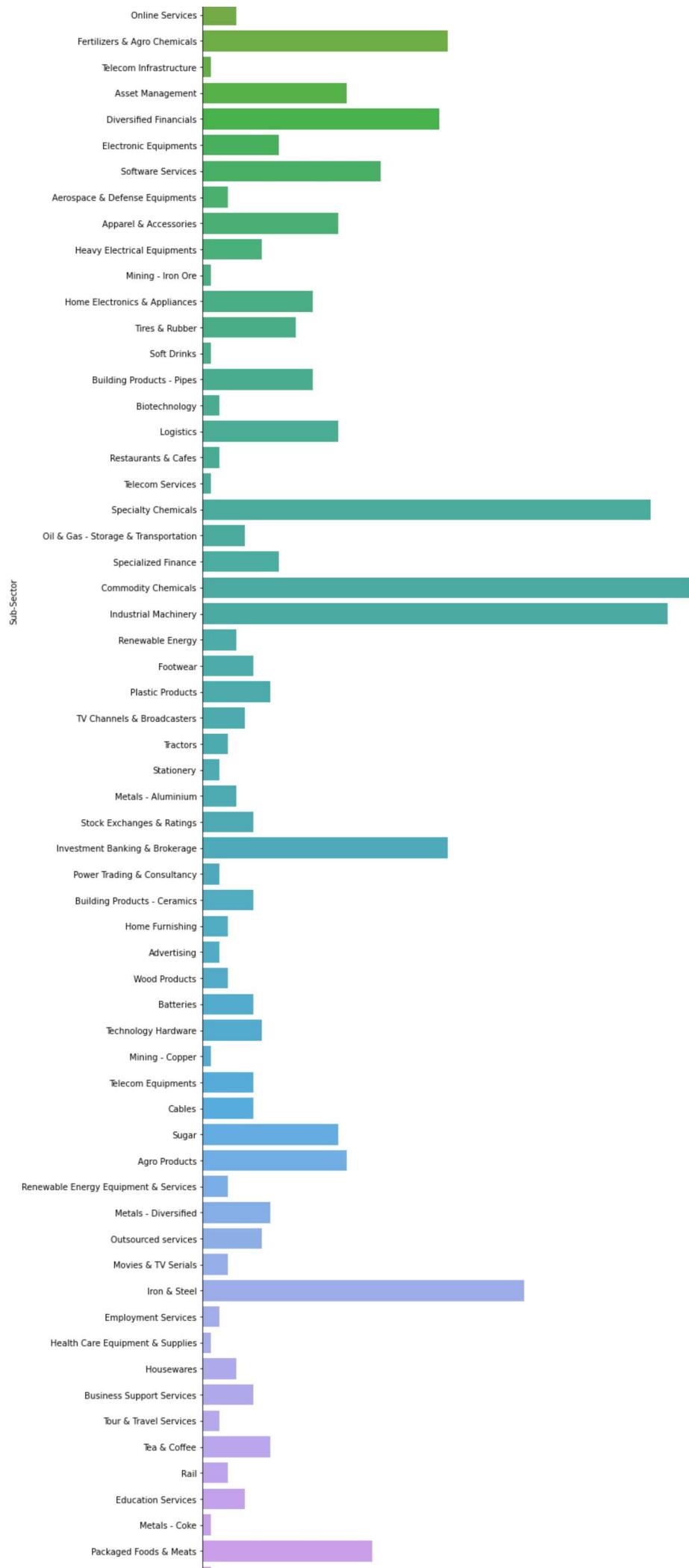
'Retail - Apparel',
'Heavy Electrical Equipments',
'Mining - Iron Ore',
'Home Electronics & Appliances',
'Tires & Rubber',
'Soft Drinks',
'Building Products - Pipes',
'Biotechnology',
'Logistics',
'Restaurants & Cafes',
'Specialty Chemicals',
'Oil & Gas - Storage & Transportation',
'Specialized Finance',
'Commodity Chemicals',
'Cycles',
'Packaged Foods & Meats',
'Industrial Machinery',
'Footwear',
'Hotels, Resorts & Cruise Lines',
'Plastic Products',
'TV Channels & Broadcasters',
'Tractors',
'Stationery',
'Stock Exchanges & Ratings',
'Investment Banking & Brokerage',
'Power Trading & Consultancy',
'Equity',
'Building Products - Ceramics',
'Home Furnishing',
'Advertising',
'Wood Products',
'Batteries',
'Technology Hardware',
'Mining - Copper',
'Theatres',
'Telecom Equipments',
'Cables',
'Sugar',
'Agro Products',
'Renewable Energy Equipment & Services',
'Outsourced services',
'Movies & TV Serials',
'Employment Services',
'Debt',
'Health Care Equipment & Supplies',
'Housewares',
'Theme Parks & Gaming',
'Tour & Travel Services',
'Infra REIT',
'Roads',
'Rail',
'Education Services',
'Metals - Coke',
'Packaging',
'Shipbuilding',
'Gold',
'Paper Products',
'Building Products - Granite',
'Building Products - Laminates',
'Heavy Machinery',
'Mining - Manganese',
'Cable & D2H',
'Seeds',
'Oil & Gas - Equipment & Services',

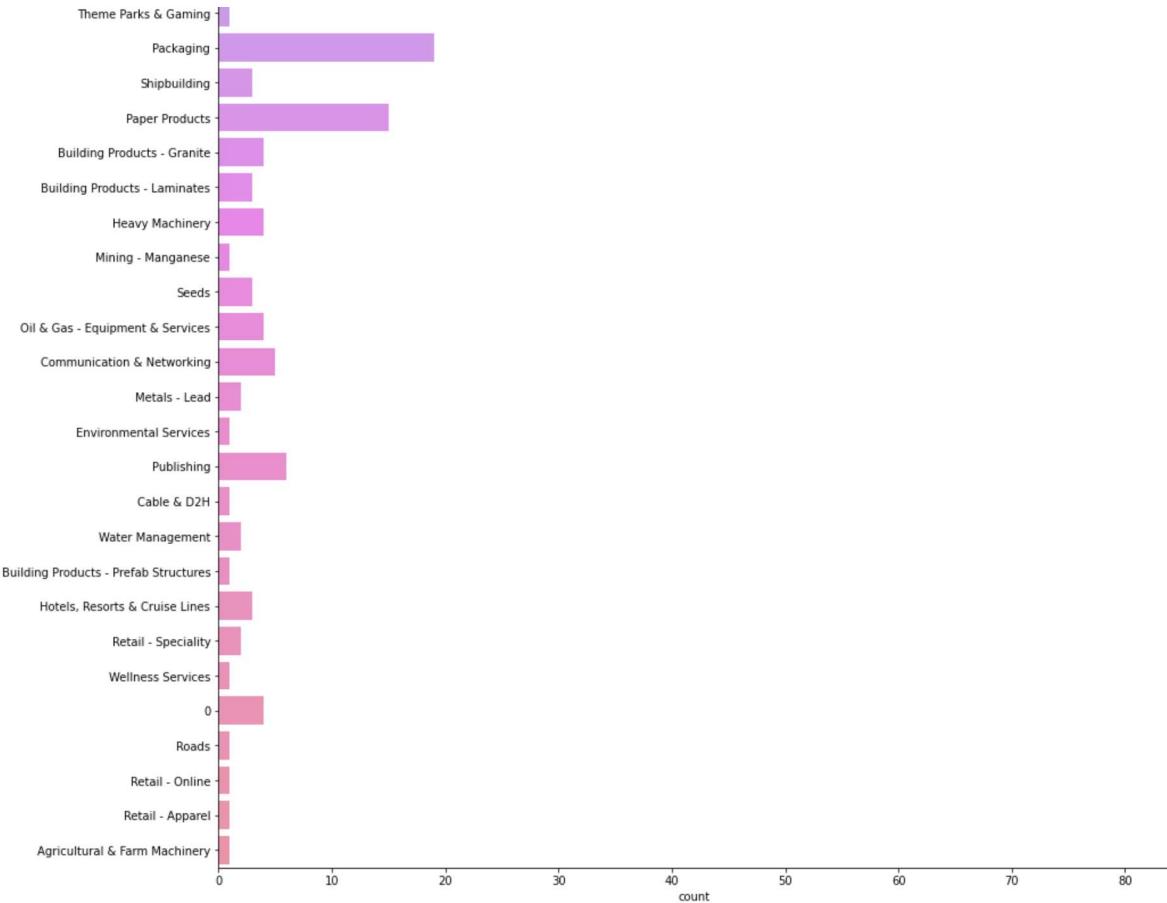
```
'Communication & Networking',  
'Animation',  
'Metals - Lead',  
'Environmental Services',  
'Agricultural & Farm Machinery',  
'Publishing',  
'Water Management',  
'Building Products - Prefab Structures',  
'Radio',  
'Dredging',  
'Retail - Speciality',  
'Wellness Services',  
'0',  
'Airports',  
'Three Wheelers',  
'Retail - Online',  
'Metals - Iron']
```

In [40]:

```
ax = plt.subplots(figsize=(15, 70))  
sns.countplot(y = "Sub-Sector", data = df1).set_title('Market Capitalisation');
```







In [41]: `# The above graph shows that 'Pharmaceuticals' were the highest sub sector having th`

In [25]: `df1.isnull().sum()`

Out[25]:

Name	0
Ticker	0
Sub-Sector	0
Market Cap	0
Close Price	0
PE Ratio	39
Return on Equity	123
Return on Assets	44
Net Profit Margin	43
5Y Avg EBITDA Margin	108
EBITDA Margin	43
1Y Return vs Nifty	0
5Y CAGR	377
Debt to Equity	160
Long Term Debt to Equity	162

`dtype: int64`

In []:

In []:

In []:

Task 2

Secondly, visualise the companies by segregating them according to Market Cap in three categories: Large Cap(>20,000 crore), Mid Cap (5,000 to 20,000 crore) & Small Cap(<5,000).

In [26]:

```
df1.head()
```

Out[26]:

	Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets	M
0	Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marketing	1.679534e+06	2467.40	34.186904	7.793277	3.957021	9.7%
1	Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1.358569e+06	3626.70	41.892356	37.740680	25.932382	19.3%
2	HDFC Bank Ltd	HDFCBANK	Private Banks	8.206363e+05	1486.50	25.779258	16.435200	1.883433	20.4%
3	Infosys Ltd	INFY	IT Services & Consulting	7.775010e+05	1853.05	40.178853	27.135305	19.417895	18.8%
4	ICICI Bank Ltd	ICICIBANK	Private Banks	5.002790e+05	710.75	27.212283	12.382897	1.243333	11.3%

In [27]:

```
df1['Market Cap']
```

Out[27]:

```
0      1.679534e+06
1      1.358569e+06
2      8.206363e+05
3      7.775010e+05
4      5.002790e+05
      ...
2009    1.256577e+02
2010    1.255488e+02
2011    1.255056e+02
2012    1.253262e+02
2013    1.249950e+02
Name: Market Cap, Length: 1947, dtype: float64
```

In [28]:

```
conditions = [
    (df1['Market Cap'] < 5000),
    (df1['Market Cap'] >= 5000) & (df1['Market Cap'] <= 20000),
    (df1['Market Cap'] > 20000)
]

values = ['Small Cap', 'Mid Cap', 'Large Cap']

df1['Categorised Cap'] = np.select(conditions, values)

df1
```

Out[28]:

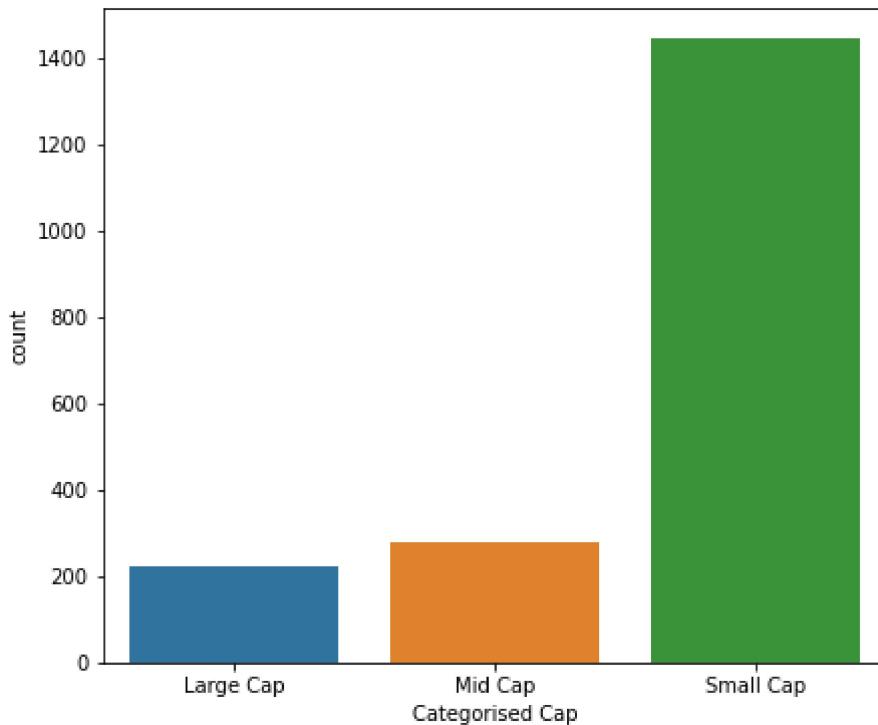
	Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets
0	Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marketing	1.679534e+06	2467.40	34.186904	7.793277	3.957021
1	Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1.358569e+06	3626.70	41.892356	37.740680	25.932382
2	HDFC Bank Ltd	HDFCBANK	Private Banks	8.206363e+05	1486.50	25.779258	16.435200	1.883433
3	Infosys Ltd	INFY	IT Services & Consulting	7.775010e+05	1853.05	40.178853	27.135305	19.417895
4	ICICI Bank Ltd	ICICIBANK	Private Banks	5.002790e+05	710.75	27.212283	12.382897	1.243333
...
2009	Galaxy Bearings Ltd	GALXBRG	Heavy Machinery	1.256577e+02	385.05	15.986985	18.976340	14.784163
2010	Signet Industries Ltd	SIGIND	Plastic Products	1.255488e+02	43.00	8.999914	7.117710	2.017514
2011	Suraj Ltd	SURAJLTD	Iron & Steel	1.255056e+02	64.95	91.609935	1.548810	0.767980
2012	Polson Ltd	POLSON	Commodity Chemicals	1.253262e+02	10440.00	14.692403	8.487562	4.800090
2013	LKP Finance Ltd	LKPFIN	Investment Banking & Brokerage	1.249950e+02	100.70	2.061602	30.499522	24.118384

1947 rows × 16 columns

In [29]: `df1['Categorised Cap'].value_counts()`Out[29]:

```
Small Cap    1445
Mid Cap      278
Large Cap    224
Name: Categorised Cap, dtype: int64
```

In [30]: `plt.figure(figsize = (7,6))
sns.countplot(df1['Categorised Cap'])`Out[30]: `<AxesSubplot:xlabel='Categorised Cap', ylabel='count'>`



In [39]: `# The above graph shows that 'Small Cap' Companies are in high numbers.`

In []:

Task 3

Thirdly, pick 10 random companies from the entire group and visualise the following through a line chart

Find the Intrinsic Value of the company based on 3 cases of growth (g): Assume 3 Cases for g (Growth) : Good (15% Growth) ; Bad (-5% Growth); Best (25%Growth) Visualise these 10 companies on a line chart for all 3 cases of growth going forward against its current market cap.

$$\diamond V = (\text{EPS} (8.5 + 2g) 6) / Y$$

In []:

In [31]:

```
df1['g'] = df2['5Y Historical EPS Growth']
```

In [32]:

```
df1
```

Out[32]:

	Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets
0	Reliance Industries Ltd	RELIANCE	Oil & Gas - Refining & Marketing	1.679534e+06	2467.40	34.186904	7.793277	3.957021
1	Tata Consultancy Services Ltd	TCS	IT Services & Consulting	1.358569e+06	3626.70	41.892356	37.740680	25.932382

		Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on Assets
2	HDFC Bank Ltd	HDFCBANK		Private Banks	8.206363e+05	1486.50	25.779258	16.435200	1.883433
3	Infosys Ltd	INFY		IT Services & Consulting	7.775010e+05	1853.05	40.178853	27.135305	19.417895
4	ICICI Bank Ltd	ICICIBANK		Private Banks	5.002790e+05	710.75	27.212283	12.382897	1.243333
...
2009	Galaxy Bearings Ltd	GALXBRG		Heavy Machinery	1.256577e+02	385.05	15.986985	18.976340	14.784163
2010	Signet Industries Ltd	SIGIND		Plastic Products	1.255488e+02	43.00	8.999914	7.117710	2.017514
2011	Suraj Ltd	SURAJLTD	Iron & Steel	1.255056e+02	64.95	91.609935	1.548810	0.767980	
2012	Polson Ltd	POLSON	Commodity Chemicals	1.253262e+02	10440.00	14.692403	8.487562	4.800090	
2013	LKP Finance Ltd	LKPFIN	Investment Banking & Brokerage	1.249950e+02	100.70	2.061602	30.499522	24.118384	

1947 rows × 17 columns

In [33]: `df1['g'].isnull().sum()`

Out[33]: 655

```
In [34]: quotient = 655 / 1947
percent = quotient * 100
print("Only", percent, "% of data has either 0 values or null values")
```

Only 33.64149974319466 % of data has either 0 values or null values

In [35]: `# There is 33% of data having null values in 'g'`In [36]: `df1.dropna(subset=['g'], inplace=True)`

```
In [37]: # After dropping nan values again checking for null values
df1['g'].isnull().sum()
```

Out[37]: 0

In []:

In [38]: df1.sample(n = 10)

Out[38]:

		Name	Ticker	Sub-Sector	Market Cap	Close Price	PE Ratio	Return on Equity	Return on A
350		Blue Star Ltd	BLUESTARCO	Home Electronics & Appliances	9465.247343	978.75	94.322345	11.999785	2.87
1833		Basant Agro Tech (India) Ltd	BASANTGL	Commodity Chemicals	174.457938	19.05	19.847319	6.736148	3.18
444		Jubilant Pharmova Ltd	JUBLPHARMA	Pharmaceuticals	6485.544589	449.65	7.759035	16.159259	7.93
75		ICICI Lombard General Insurance Company Ltd	ICICIGI	Insurance	62313.554770	1240.55	42.302690	21.710691	3.86
1460		Shankar Lal Rampal Dye-Chem Ltd	SRD	Commodities Trading	348.405728	155.40	45.662612	16.572546	12.19
1135		Bliss GVS Pharma Ltd	BLISSGVS	Pharmaceuticals	760.995948	72.65	11.107808	8.681603	6.50
134		Honeywell Automation India Ltd	HONAUT	Electronic Equipments	34847.006180	39557.45	75.747775	19.341114	12.41
186		Kansai Nerolac Paints Ltd	KANSAINER	Paints	25644.494880	472.15	48.409587	13.460278	10.17
1049		Uniphos Enterprises Ltd	UNIENTER	Fertilizers & Agro Chemicals	940.259879	136.90	38.456437	1.155340	1.13
925		Kewal Kiran Clothing Ltd	KKCL	Apparel & Accessories	1284.677350	205.85	64.882694	4.506299	3.24



In []:

I am not able to interpret the 3rd task as there is no primary key to which I can combine the three tables. If you have time, can we have a google meet to discuss this task?

In []:



In []:

In []: