

# DSC540-T302 Data Preparation

## Project Milestone 4

Saravanan Janarthanan

---

```
In [1]: import sqlite3  
import pandas as pd  
import numpy as np
```

### Project Requirement

- Load each dataset into SQL Lite as an individual table and then you must join the datasets together in Python into 1 dataset.
- Once all the data is merged together in your database, create 5 visualizations that demonstrate the data you have cleansed.

### Read, Load and transform data for all the three formats

- Read **Flat File** , clean or transform and load into a sqllite table
  - Read website contents by **webscraping** it, parse or find the content , clean it up or transform and load into a sqllite table
  - Use the **API** call to recieve webservice response , clean if needed , transform and load into a sqllite table
  - Read the data in these 3 tables by joining them using SQL query and load them in a dataframe
  - Use the merged data in Dataframe and choose column / features to build visulaization
- 
- 

## 1. FLAT File

---

---

### Read Flat File, clean or transform and load into a sqllite table

#### 1.1 Read and Load Flat file into Database

```
In [2]: # Load the data file in a dataframe  
sp500_df = pd.read_csv('SNP_500_Performance_data.csv')  
sp500_df.head()
```

Out[2]:

	Symbol	Name	Market Cap	P/E	EPS	Net Income	Beta	Dividend	Div Yield	Extended Hr Last Price
0	A	Agilent Technologies	4.223508e+10	26.32	5.37	1240000000	1.11	0.922	0.64%	144.12
1	AAL	American Airlines Gp	9.009457e+09	5.24	2.64	822000000	1.57	NaN	0.00%	13.75
2	AAPL	Apple Inc	2.620000e+12	26.30	6.42	96995000000	1.27	0.960	0.57%	169.47
3	ABBV	Abbvie Inc	3.010000e+11	15.11	11.11	4863000000	0.58	5.990	3.52%	169.99
4	ABNB	Airbnb Inc Cl A	1.050000e+11	36.85	4.31	4792000000	1.25	NaN	0.00%	161.77

5 rows × 30 columns

## 1.2 Change header of file dataset

- View the header names
- identify the ones that needs to be changed
- Retrieve the current headers in a list
- Perform the changes to the identified headers in the list
- Set the new headers using the modified list values

In [3]:

```
header_nms = sp500_df.columns.tolist()
# Change 4th column (index 3) header from "P/E" to "Price Earnings"
# Change 12th column (index 11) header from "%Chg (Ext)" to "Extended Hr %Change"
# Change 18th column (index 17) header from "20D Rel Str" to "20D Relative Strength"

header_nms[3] = "Price Earnings Ratio"
header_nms[11] = "Extended Hr %Change"
header_nms[17] = "20D Relative Strength"

sp500_df.columns = header_nms
```

## 1.3 Convert float number to int64 to remove exponential listing or displaying

- Column 'Market Cap' displays in exponential form that impacts readability.
- Convert that column values to integer 64 base to display the whole numbers

In [4]:

```
sp500_df["Market Cap"] = sp500_df["Market Cap"].astype('int64')
```

## 1.4 Replace special character, replace with a no character or null character

- Column 'Volume' has '?' character in between the values
- Use the String replace method to replace the '?' character with ''

In [5]:

```
sp500_df['Volume'] = sp500_df['Volume'].str.replace('?', '')
```

## 1.5 Persist the Flat file data's dataframe in Database

```
In [6]: # Fill blanks with '_' in all header names for easy querying
sp500_df.columns = sp500_df.columns.str.replace(' ', '_')

# connect to a database, creates one if not in memory or connects to a new one
conn = sqlite3.connect('SJ_Project.db')
try:
    # Save the DataFrame to database
    sp500_df.to_sql('FlatFile', conn, if_exists='replace', index=False)
    print("DataFrame written to SQLite successfully.")
# Catch any exception and print error message
except Exception as e:
    print(f"An error occurred: {e}")

# Verify the loaded Flatfile data by reading from table
try:
    df_read = pd.read_sql('SELECT Symbol, Name, Market_Cap FROM FlatFile limit 10', conn)
    print(df_read)
except Exception as e:
    print(f"An error occurred while reading from table FlatFile: {e}")

# Close the connection
conn.close()
```

DataFrame written to SQLite successfully.

	Symbol	Name	Market_Cap
0	A	Agilent Technologies	42235084000
1	AAL	American Airlines Gp	9009457000
2	AAPL	Apple Inc	2620000000000
3	ABBV	Abbvie Inc	301000000000
4	ABNB	Airbnb Inc Cl A	105000000000
5	ABT	Abbott Laboratories	193000000000
6	ACGL	Arch Capital Grp Ltd	35601260000
7	ACN	Accenture Plc	223000000000
8	ADBE	Adobe Systems Inc	217000000000
9	ADI	Analog Devices	96583040000

---

---

## 2. Web Scrape Data

---

---

**Read website contents by webscraping it, parse or find the content , clean it up or transform and load into a sqlite table**

### 2.1 Read and Load web scraped data into Database

```
In [7]: # Load the BeautifulSoup
from bs4 import BeautifulSoup
import requests
```

```
In [8]: # Url to load the web scrape data
web_url = 'https://en.wikipedia.org/wiki/List_of_S%26P_500_companies#S&P_500_component'

# Fetch the HTML content of the webpage
response = requests.get(web_url)

# Check if the request was successful (status code 200)
if response.status_code == 200:
    # Parse the HTML content using BeautifulSoup
    soup = BeautifulSoup(response.content, 'html.parser')

# scrape the html data
tables = soup.find_all('table')
wiki_tables = soup.find("table", {"class": "wikitable sortable"})
th_elements = wiki_tables.find_all('th')
tables_hdr = [cell.text.strip() for cell in th_elements]
```

```
In [9]: # Retrive all the TR tags
rows = wiki_tables.find_all('tr')

# Declare a List to hold all the row data of the table
temp_lst = []

for row in rows:
    # filter all the TD, tags
    cells = row.find_all(['td'])

    # strip other values and retrieve only the tag text values
    row_data = [cell.text.strip() for cell in cells]
    temp_lst.append(row_data)
```

```
In [10]: # Load the parsed data in a dataframe
web_df = pd.DataFrame(temp_lst, columns=tables_hdr)
web_df.head()
```

Out[10]:

	Symbol	Security	GICS Sector	GICS Sub-Industry	Headquarters Location	Date added	CIK	Founded
0	None	None	None	None	None	None	None	None
1	MMM	3M	Industrials	Industrial Conglomerates	Saint Paul, Minnesota	1957-03-04	0000066740	1902
2	AOS	A. O. Smith	Industrials	Building Products	Milwaukee, Wisconsin	2017-07-26	0000091142	1916
3	ABT	Abbott	Health Care	Health Care Equipment	North Chicago, Illinois	1957-03-04	0000001800	1888
4	ABBV	AbbVie	Health Care	Biotechnology	North Chicago, Illinois	2012-12-31	0001551152	2013 (1888)

## 2.2 Replace the Web scraped data frame column headers

- Replace 'Security' with 'Company Name'
- Replace 'CIK' with 'Central Index Key'
- Replace 'Founded' with 'Year Founded'

```
In [11]: # Define new column header
tables_hdr[0] = 'WS_Symbol'
tables_hdr[1] = 'Company Name'
tables_hdr[6] = 'Central Index Key'
tables_hdr[7] = 'Year Founded'
```

```
In [12]: # Fill blanks with '_' in all header names for easy querying
tables_hdr = [x.replace(' ', '_') for x in tables_hdr]
```

```
In [13]: # update the column headers
web_df.columns = tables_hdr
```

## 2.3 Remove bad Data in Web scraped dataframe

Remove rows that has no value

- Remove the first row that has no value 'None'

```
In [14]: # Drop row at index position 0
web_df = web_df.drop(web_df.index[0])
```

## 2.4 Format Data in Web scraped dataframe

Year founded column has year value. some of the rows has other info appended

- Trim those value appended after year value. Year is 4 char in length, hence slice the characters after 4 characters

```
In [15]: web_df['Year_Founded'] = web_df['Year_Founded'].str[:4]
```

## 2.5 Persist the web scraped data's dataframe in Database

```
In [16]: # connect to a database, creates one if not in memory or connects to a new one
conn = sqlite3.connect('SJ_Project.db')
try:
    # Save the DataFrame to database
    web_df.to_sql('WebScrapedData', conn, if_exists='replace', index=False)
    print("DataFrame written to SQLite successfully.")
# Catch any exception and print error message
except Exception as e:
    print(f"An error occurred: {e}")

# Verify by reading from table
try:
    df_read = pd.read_sql('SELECT WS_Symbol, Company_Name, Year_Founded FROM WebScrapedData', conn)
    print(df_read)
except Exception as e:
    print(f"An error occurred while reading from table FlatFile: {e}")

# Close the connection
conn.close()
```

```
DataFrame written to SQLite successfully.
   WS_Symbol          Company_Name Year_Founded
0      MMM                  3M        1902
1      AOS            A. O. Smith     1916
2      ABT             Abbott       1888
3     ABBV            AbbVie       2013
4     ACN           Accenture     1989
5    ADBE            Adobe Inc.    1982
6     AMD  Advanced Micro Devices    1969
7     AES        AES Corporation    1981
8     AFL              Aflac       1955
9      A  Agilent Technologies    1999
```

---



---

### 3. Read API data

---



---

**Use the API call to receive webservice response , clean if needed , transform and load into a sqlite table**

#### 3.1 Read and Load API queried data into Database

In [17]:

```
import json

with open('PolygonKey.json') as f:
    keys = json.load(f)
    Polygon_key = keys['Polygon_key']
```

In [18]:

```
# url to access the api, using 5/15 date data
poly_url = "https://api.polygon.io/v2/aggs/grouped/locale/us/market/stocks/2024-05-15"
# define the url load params
url_param = {}
url_param['apikey'] = Polygon_key
url_param['adjusted'] = 'false'

poly_data = ""

# use the response module to get the JSON data
try:
    response = requests.get(url=poly_url, params=url_param)
    json_res_data = response.json()
    # check if the response is good, status attribute has value 'OK', if so Load the data
    # that contains stock data
    if json_res_data['status'] == 'OK':
        poly_df = pd.DataFrame(json_res_data['results'])
    else:
        err_msg = "Error Occured while searching title " + title_name
        if 'Error' in list(json_res_data.keys()):
            err_msg = err_msg + " Error Message received : " + json_res_data['Error']
        print(err_msg)
except Exception as e:
    print(f"Error while retrieving title details from OMDB : {e}")
    print(e)
```

```
In [19]: # print the Dataframe  
poly_df
```

```
Out[19]:
```

	T	v	vw	o	c	h	I	t	
<b>0</b>	BPTH	136171.0	2.3038	2.400	2.24000	2.680000	2.0800	1715803200000	91
<b>1</b>	POLA	201286.0	0.5142	0.507	0.53690	0.537000	0.4800	1715803200000	45
<b>2</b>	BNOX	38878.0	0.9249	0.925	0.92005	0.948999	0.8978	1715803200000	13
<b>3</b>	IAT	174090.0	43.5808	43.650	43.65000	43.950000	43.3300	1715803200000	219
<b>4</b>	RUM	3120835.0	6.9658	6.940	7.15000	7.250000	6.5600	1715803200000	2422
...	...	...	...	...	...	...	...	...	...
<b>10495</b>	ZTST	0.0	NaN	40.170	40.17000	40.170000	40.1700	1715803200000	N
<b>10496</b>	ZBZX	0.0	NaN	25.000	25.00000	25.000000	25.0000	1715803200000	N
<b>10497</b>	ZVZZT	211707.0	18.7798	20.250	39.85000	39.850000	11.0000	1715803200000	141
<b>10498</b>	ZTEST	0.0	NaN	8662.955	8662.95500	8662.955000	8662.9550	1715803200000	N
<b>10499</b>	ZIEXT	1000.0	1.0000	1.000	1.00000	1.000000	1.0000	1715803200000	

10500 rows × 9 columns

### 3.2 Change the header names in API queried data

- Change header name 'T' as 'Symbol' in column 1 (index 0)
- Change 'v' as 'Volume' in column 2 (index 1)
- Change 'c' as 'Closing Price' in column 5 (index 4)
- Change 'h' as 'Highest Price' in column 6 (index 5)
- Change 'l' as 'Lowest Price' in column 7 (index 6)
- Change 't' as 'Timestamp' in column 8 (index 7)
- Change 'vw' as 'Vol Wt Avg Price' in column 3 (index 2)
- Change 'o' as 'Open Price' in column 4 (index 3)
- Change 'n' as 'Number of Transactions' in column 9 (index 8)

```
In [20]: # Retrieve the header names to change them  
header_nms = poly_df.columns.tolist()
```

```
# Map the new names to the right indexes as intended  
header_nms[0] = 'API_Symbol'  
header_nms[1] = 'API_Volume'  
header_nms[2] = 'Vol_Wt_Avg_Price'  
header_nms[3] = 'Open_Price'  
header_nms[4] = 'Closing_Price'  
header_nms[5] = 'Highest_Price'  
header_nms[6] = 'Lowest_Price'  
header_nms[7] = 'Timestamp'  
header_nms[8] = 'Number_of_Transactions'
```

```
In [21]: # Fill blanks with '_' in all header names for easy querying
header_nms = [x.replace(' ', '_') for x in header_nms]
```

```
In [22]: # Update the new headernames to dataframe
poly_df.columns = header_nms
poly_df.head()
```

Out[22]:

	API_Symbol	API_Volume	Vol_Wt_Avg_Price	Open_Price	Closing_Price	Highest_Price	Lowest_Price
0	BPTH	136171.0	2.3038	2.400	2.24000	2.680000	2.0800
1	POLA	201286.0	0.5142	0.507	0.53690	0.537000	0.4800
2	BNOX	38878.0	0.9249	0.925	0.92005	0.948999	0.8978
3	IAT	174090.0	43.5808	43.650	43.65000	43.950000	43.3300
4	RUM	3120835.0	6.9658	6.940	7.15000	7.250000	6.5600

### 3.3 Convert 'Volume' and "Number of Transactions" to intger values

```
In [23]: import math

# fill the misssing values with 0 to avoid any conversion data format issues
poly_df['Number_of_Transactions'] = poly_df['Number_of_Transactions'].fillna(0)

# use astype to change the data type
poly_df['API_Volume'] = poly_df['API_Volume'].round().astype(int)
poly_df['Number_of_Transactions'] = poly_df['Number_of_Transactions'].round().astype(int)
```

### 3.4 Fill the average price of misssing values in 'Vol Wt Avg Price' with average price of 'Open price and 'Closing Price'

- Create a dummy column to fill the avg price for all columns
- Then using fillna method copy the new column values with avg price into na rows of 'Vol Wt Avg Price' column
- Drop the new column

```
In [24]: # create a new temp column to store the average price using mean value
poly_df['new_avg_price'] = poly_df[['Open_Price', 'Closing_Price']].mean(axis=1)
poly_df['new_avg_price'] = poly_df['new_avg_price'].round(2) # round the values to 2 decimal places

# Copy the average price in the temp column to the null or NA rows values for 'Vol Wt Avg Price'
poly_df['Vol_Wt_Avg_Price'] = poly_df['Vol_Wt_Avg_Price'].fillna(poly_df['new_avg_price'])
poly_df = poly_df.drop(columns=['new_avg_price'])
```

### 3.5 Convert 'Timestamp' field with milli second values to a timezone field with date and time

```
In [25]: # Convert timestamp in millis to regulat date, time format in UTC timezone format and
import pytz
poly_df['Timestamp'] = pd.to_datetime(poly_df['Timestamp'], unit='ms')
poly_df['Timestamp'] = poly_df['Timestamp'].dt.tz_localize('UTC')
poly_df['Timestamp'] = poly_df['Timestamp'].dt.tz_convert('US/Eastern')
```

### 3.6 Persist the web scraped data's dataframe in Database

```
In [26]: # connect to a database, creates one if not in memory or connects to a new one
conn = sqlite3.connect('SJ_Project.db')

try:
    # Save the DataFrame to database
    poly_df.to_sql('APIData', conn, if_exists='replace', index=False)
    print("DataFrame written to SQLite successfully.")

# Catch any exception and prijt error message
except Exception as e:
    print(f"An error occurred: {e}")

# Verify by reading from table
try:
    df_read = pd.read_sql('SELECT API_Symbol, API_Volume, Timestamp FROM APIData limit 10')
    print(df_read)
except Exception as e:
    print(f"An error occurred while reading from table FlatFile: {e}")

# Close the connection
conn.close()
```

DataFrame written to SQLite successfully.

	API_Symbol	API_Volume	Timestamp
0	BPTH	136171	2024-05-15 16:00:00-04:00
1	POLA	201286	2024-05-15 16:00:00-04:00
2	BNOX	38878	2024-05-15 16:00:00-04:00
3	IAT	174090	2024-05-15 16:00:00-04:00
4	RUM	3120835	2024-05-15 16:00:00-04:00
5	AJX	116684	2024-05-15 16:00:00-04:00
6	INMD	1058873	2024-05-15 16:00:00-04:00
7	BALT	73089	2024-05-15 16:00:00-04:00
8	ALTI	150409	2024-05-15 16:00:00-04:00
9	CRON	2600970	2024-05-15 16:00:00-04:00

```
In [27]: query = "SELECT * FROM FlatFile \
LEFT JOIN WebScrapedData ON FlatFile.Symbol = WebScrapedData.WS_Symbol \
LEFT JOIN APIData ON FlatFile.Symbol = APIData.API_Symbol"
```

```
In [28]: # connect to a database, creates one if not in memory or connects to a new one
conn = sqlite3.connect('SJ_Project.db')

try:
    df_combined = pd.read_sql(query, conn)
except Exception as e:
    print(f"An error occurred while reading from table FlatFile: {e}")

# Close the connection
conn.close()
```

```
In [29]: # Display the merged dataframe data  
df_combined
```

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividend
0	A	Agilent Technologies	42235084000	26.32	5.37	12400000000	1.11	0.922
1	AAL	American Airlines Gp	9009457000	5.24	2.64	8220000000	1.57	NaN
2	AAPL	Apple Inc	2620000000000	26.30	6.42	96995000000	1.27	0.960
3	ABBV	Abbvie Inc	301000000000	15.11	11.11	48630000000	0.58	5.990
4	ABNB	Airbnb Inc Cl A	1050000000000	36.85	4.31	4792000000	1.25	NaN
...	...	...	...	...	...	...	...	...
498	XYL	Xylem Inc	31092520000	34.58	3.68	609000000	1.10	1.350
499	YUM	Yum! Brands	38289832000	26.36	5.17	1597000000	1.10	2.485
500	ZBH	Zimmer Biomet Holdings	26161348000	16.87	7.56	1024000000	1.01	0.960
501	ZBRA	Zebra Technologies	14920015000	33.29	8.80	296000000	1.79	NaN
502	ZTS	Zoetis Inc Cl A	75857568000	31.11	5.32	2344000000	0.85	1.557

503 rows × 47 columns

```
In [30]: df_combined.isnull().sum()
```

```

Out[30]:      Symbol      0
              Name       0
              Market_Cap  0
              Price_Earnings_Ratio 0
              EPS        0
              Net_Income   0
              Beta        2
              Dividend    99
              Div_Yield    0
              Extended_Hr_Last_Price 0
              Extended_Hr_Change_ 0
              Extended_Hr_%Change 0
              Time        0
              Last_Price   0
              Change       0
              %Chg        0
              Opinion      1
              20D_Relative_Strength 0
              20D_His_Vol   0
              20D_Avg_Vol   0
              52W_Low       0
              52W_High      0
              Wtd_Alpha    0
              YTD_%Chg     0
              1M_%Chg      0
              3M_%Chg      0
              52W_%Chg     2
              Sector       0
              ISIN         0
              Volume       0
              WS_Symbol    3
              Company_Name 3
              GICS_Sector   3
              GICS_Sub-Industry 3
              Headquarters_Location 3
              Date_added   3
              Central_Index_Key 3
              Year_Founded 3
              API_Symbol   1
              API_Volume   1
              Vol_Wt_Avg_Price 1
              Open_Price   1
              Closing_Price 1
              Highest_Price 1
              Lowest_Price  1
              Timestamp    1
              Number_of_Transactions 1
              dtype: int64

```

## Visualization on combined Data

### **Visulization 1**

**Column 'Sector' from Flat file dataset to display the pie chart**

- Group the sector values and count of stick symbols or companies

- Since there are high number of sectors listed, top 5 records are used for building the pie chart

```
In [31]: # Use the sector value counts - Number of records against each sector category
sector_data = df_combined.Sector.value_counts()

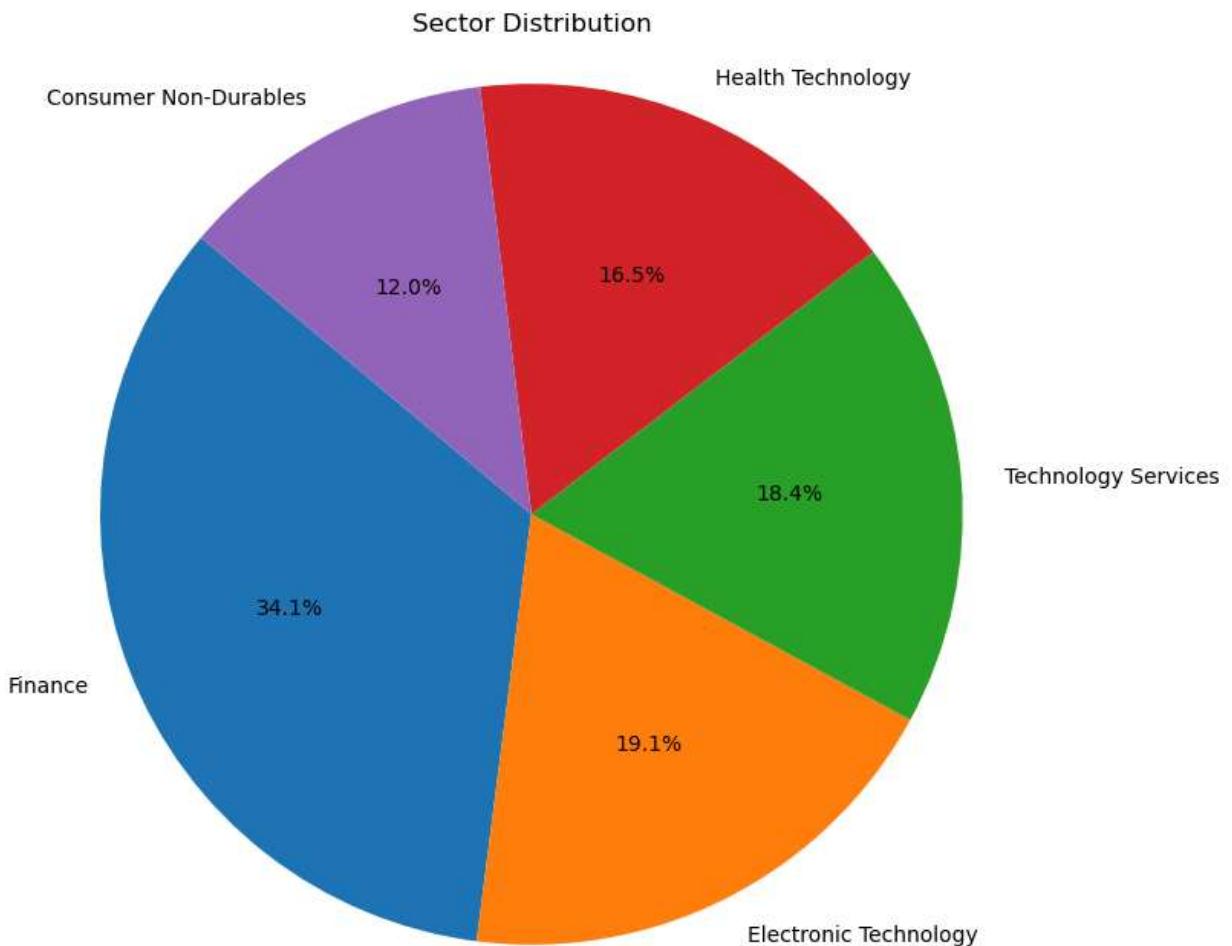
# assign the value count data into category index and its counts
sec_idx = sector_data.index
sec_values = sector_data.values
```

```
In [32]: import matplotlib.pyplot as plt

# choose top 5 count records
labels = sec_idx[:5]
sizes = sec_values[:5]

# Create a pie chart
plt.figure(figsize=(8, 8))
# plot the pie with the first category pie line at 140 degrees
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=140)
plt.axis('equal') # Equal aspect ratio ensures that pie is drawn as a circle
plt.title('Sector Distribution')

# Show the plot
plt.show()
```



**Pie Chart shows that Finance sector S&P 500 companies are more in the top 5 sectors**

---

## Visulization 2

**Use 'Headquarters Location' from web scraped dataset and 'Net Income' from Flat File dataset to build bar chart to show number of companies**

- Group the headquarters location data and choose the top 5 locations
- Use those 5 locations and aggregate the net income values (summing the values)
- Aggregated values are re-scaled to billions for better chart scales

```
In [33]: # Retrieve the head quarters unique values and their counts
HQ_Loc = df_combined['Headquarters_Location'].value_counts()

In [34]: # choose the top five Locations
Top5_HQ = HQ_Loc.index[:5].tolist()

In [35]: # Step 3: Filter the DataFrame to filter the top 5 Locations
filtered_df = df_combined[df_combined['Headquarters_Location'].isin(Top5_HQ)]

In [36]: # Aggregate the net income for the top 5 locations from new dataframe
aggregated_df = filtered_df.groupby('Headquarters_Location')['Net_Income'].sum().reset_index()

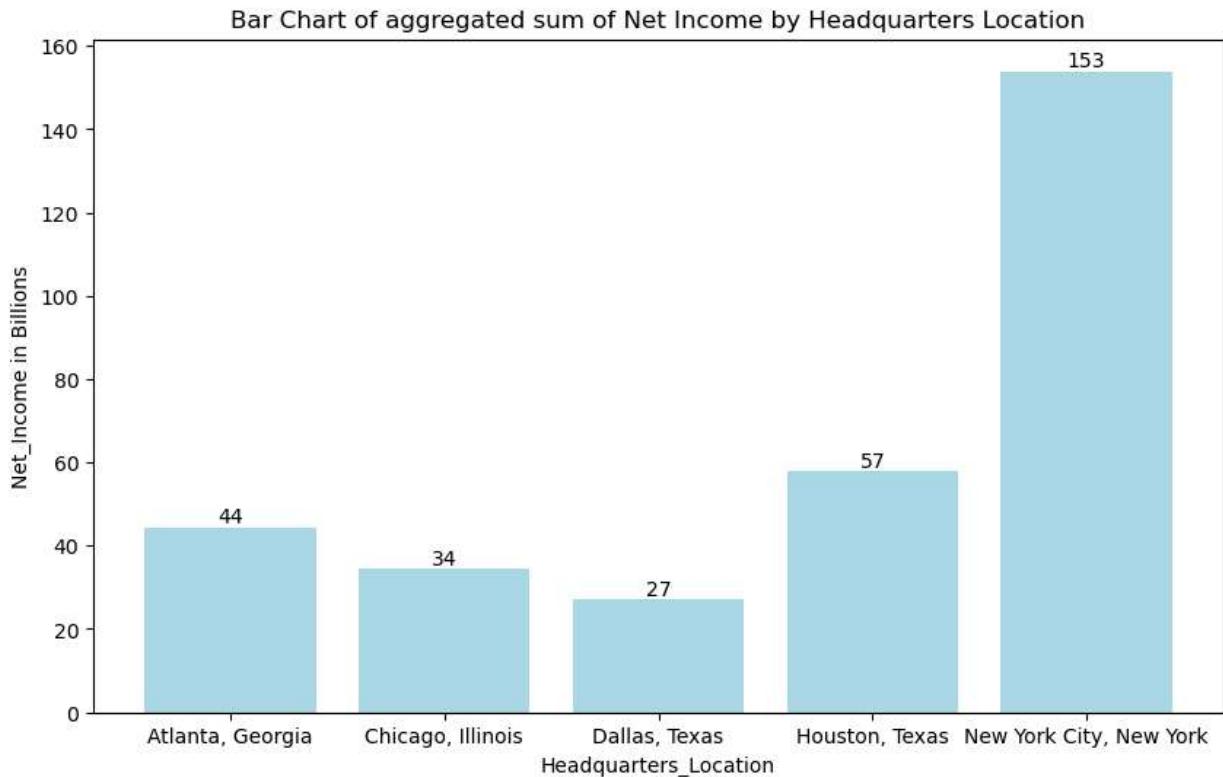
In [37]: # Convert the net ccom values to billions
aggregated_df['Net_Income'] = aggregated_df['Net_Income'] / 1000000000

In [38]: # Plot the bar chart
plt.figure(figsize=(10, 6))
bars = plt.bar(aggregated_df['Headquarters_Location'], aggregated_df['Net_Income'], color='blue')

# Add titles and labels
plt.title('Bar Chart of aggregated sum of Net Income by Headquarters Location')
plt.xlabel('Headquarters_Location')
plt.ylabel('Net_Income in Billions')

for bar in bars:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, yval, int(yval), ha='center', va='bottom')

# Show the plot
plt.show()
```



**Bar chart shows New york as head quarters for max number of S&P500 compnaies**

---

### Visulization 3

**\*Using Column 'Year Founded' from webscraped dataset and bin those as a decade and then use the count of stock symbols or companies to show the magnitude of compaines founded in the decade**

This visualization shows whihc decade had more S&P 500 compaines were founded

- Use 'Year Founded' column as a seperate dataframe
- Bin the year founded data into a equal width , 10 years , groups
- find the count of records for each decade and use the same for visualization

```
In [39]: # Create a year founded feature dataframe
df_temp = df_combined['Year_Founded']
# drop null records
df_temp = df_temp.dropna()
# convert to integer value
df_temp = df_temp.astype(int)
```

```
In [40]: # Bin the years to decades
df_decade = (df_temp // 10) * 10
```

```
In [41]: decades_num_of_companies = df_decade.value_counts().sort_index()
# Plot the bar chart
plt.figure(figsize=(15, 6))
color = (0.2, 0.4, 0.2, 0.6 )
```

```

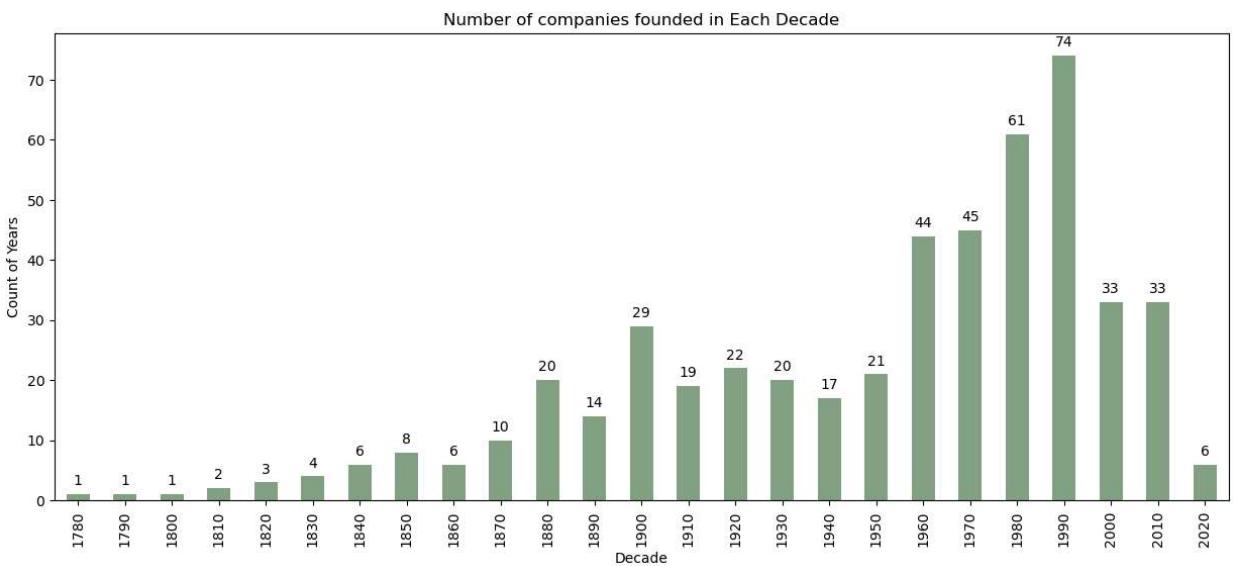
#bars = plt.bar(decades_num_of_companies.index, decades_num_of_companies.values, color
ax = decades_num_of_companies.plot(kind='bar', color=color)

# Add titles and Labels
plt.title('Number of companies founded in Each Decade')
plt.xlabel('Decade')
plt.ylabel('Count of Years')

# Annotate each bar with its value
for i, v in enumerate(decades_num_of_companies):
    ax.text(i, v + 1, str(v), ha='center', va='bottom')

# Show the plot
plt.show()

```



Bart chart shows 1990 decade had max number of S&P 500 companies were founded

#### Visulization 4

**Use 'EPS' and 'P/E ratio' from Flat File and 'GICS Sector' from web scraping dataset to show correlation between EPS and P/E ration and highlight the sector that shows spread**

- The High EPS and high P/E ratio shows a god performance of the stock
- Using the sector highlighting helps to identify the sector growth as well individual stock symbol performce within a sector

```

In [42]: import seaborn as sns

# Use the Price earning ratio and EPS (Earnings Per share) values to print the scatter
x = df_combined['Price_Earnings_Ratio']
y = df_combined['EPS']

# Create a scatter plot
plt.figure(figsize=(10, 8))

```

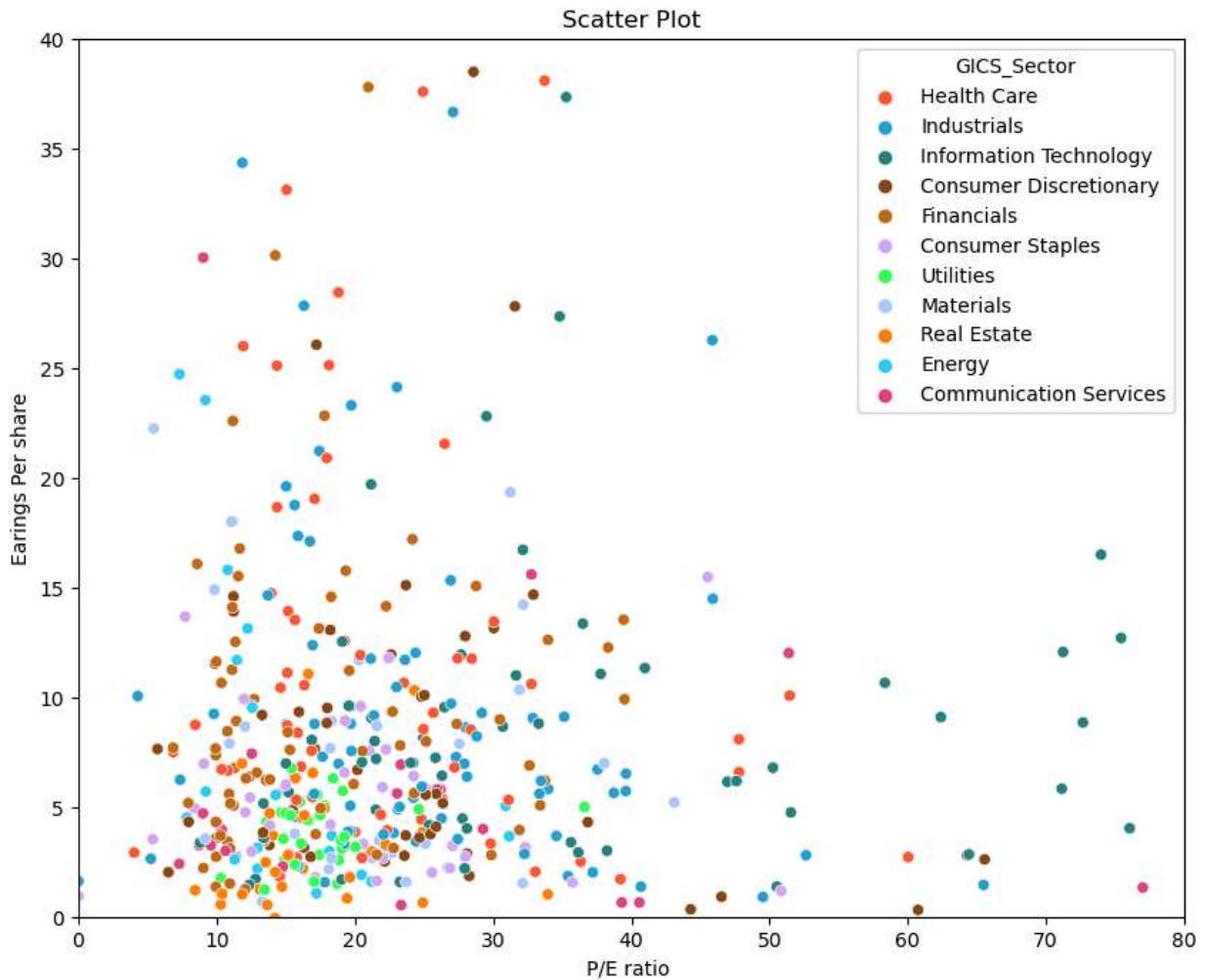
```

#custom_palette = sns.color_palette(["#9b59b6", "#3498db", "#95a5a6", '#ff6347', '#ff4380'])
#custom_palette = sns.color_palette(['#FF5733', '#33FF57', '#3357FF', '#FF33F6', '#573380'])
custom_palette = sns.color_palette(['#FF5733', '#27a2d4', '#248280', '#874914', '#c46a7b'])
sns.scatterplot(data=df_combined, x='Price_Earnings_Ratio', y='EPS', hue='GICS_Sector')
sns.scatterplot(data=df_combined, x='Price_Earnings_Ratio', y='EPS', hue='GICS_Sector')

plt.title('Scatter Plot')
plt.xlabel('P/E ratio ')
plt.ylabel('Earnings Per share')
plt.xlim(0, 80)
plt.ylim(0, 40)

# Show the plot
plt.show()

```



### Visulization 5

**Use the 'Volume' from Flat file dataset and 'API\_Volume' from API dataset to compare volume values between these three datasets**

Flat file volume has a postfix M for millions and K for 1000s , hence convert he same to whole number.

```
In [43]: # Define a method to convert values to 1000s or millions based on prefix
def convertValuePostFixToNumber(value):
    retval = float(value[:-1])
    postfix = value[-1:]

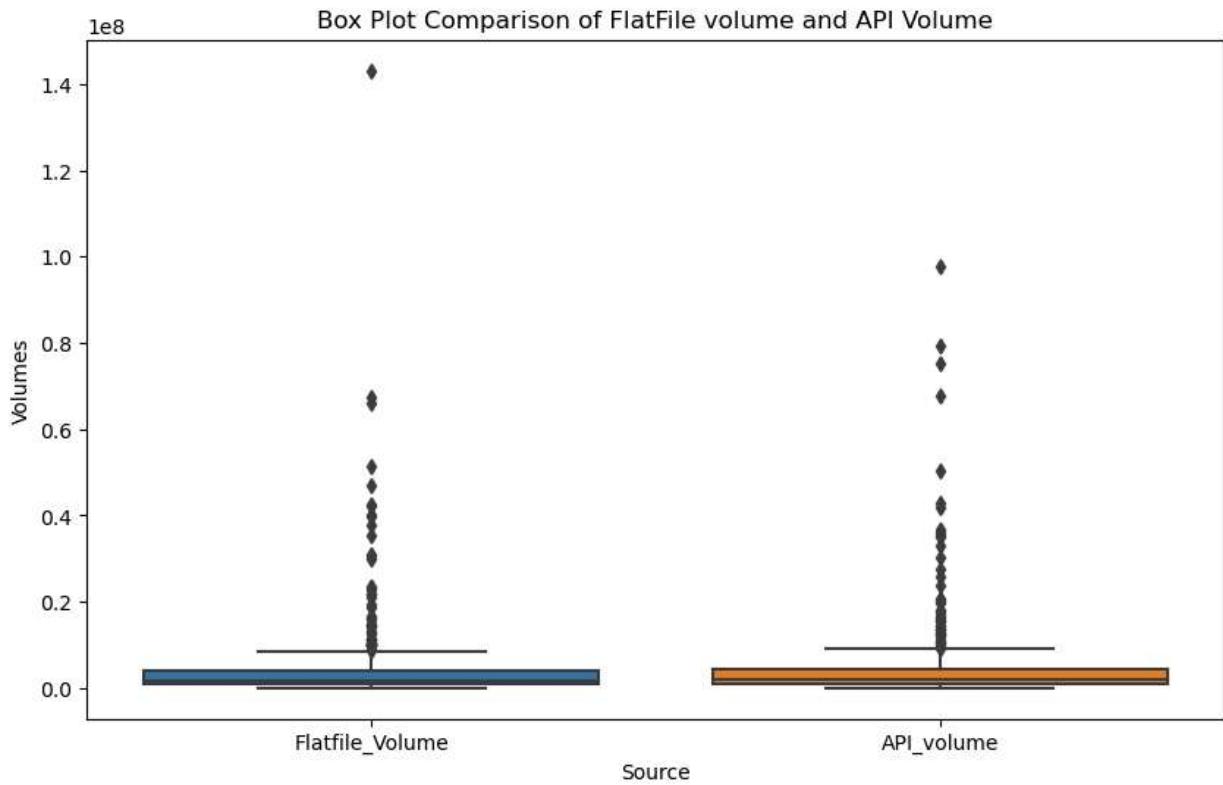
    if postfix == 'M':
        return retval*1000000
    elif postfix == 'K':
        return retval*1000
    else:
        return 0

# Create a new Dataframe to hold both volume features
df_vol = pd.DataFrame()
# Create a new column to Load the merged data Flat file volume column
df_vol['Flatfile_Volume'] = df_combined['Volume']
# Create a new column to Load the merged API volume column
df_vol['API_volume'] = df_combined['API_Volume']
# Drop any null values
df_vol = df_vol.dropna()
# Convert post values to respective whole numbers
df_vol['Flatfile_Volume'] = [convertValuePostFixToNumber(x) for x in df_vol['Flatfile_Volume']]
# Convert DataFrame to Long format for seaborn, melt the DataFrame to Long format using
# This format is required by seaborn for creating box plots with multiple columns.
df_melted = df_vol.melt(var_name='Column', value_name='Value')

# Plot the box plot
plt.figure(figsize=(10, 6))
sns.boxplot(x='Column', y='Value', data=df_melted)

# Add titles and Labels
plt.title('Box Plot Comparison of FlatFile volume and API Volume')
plt.xlabel('Source')
plt.ylabel('Volumes')

# Show the plot
plt.show()
```



Above box plot comparison shows volumes are almost simillar

```
In [45]: # Set display options to display all rows and columns
pd.set_option('display.max_rows', None) # None means unlimited
pd.set_option('display.max_columns', None) # None means show all columns
```

```
In [46]: # print the Dataframe
df_combined
```

Out[46]:

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
0	A	Agilent Technologies	42235084000	26.32	5.37	1240000000	1.11	0.1
1	AAL	American Airlines Gp	9009457000	5.24	2.64	822000000	1.57	N
2	AAPL	Apple Inc	2620000000000	26.30	6.42	96995000000	1.27	0.1
3	ABBV	Abbvie Inc	301000000000	15.11	11.11	4863000000	0.58	5.0
4	ABNB	Airbnb Inc Cl A	105000000000	36.85	4.31	4792000000	1.25	N
5	ABT	Abbott Laboratories	193000000000	24.80	4.44	5723000000	0.74	2.0
6	ACGL	Arch Capital Grp Ltd	35601260000	10.78	8.45	4443000000	0.57	N
7	ACN	Accenture Plc	223000000000	27.68	11.94	6871560000	1.19	4.0
8	ADBE	Adobe Systems Inc	217000000000	36.49	13.35	5428000000	1.28	N
9	ADI	Analog Devices	96583040000	21.19	9.06	3314580000	1.17	3.0
10	ADM	Archer Daniels Midland	32355018000	9.07	6.97	3483000000	0.77	1.0
11	ADP	Automatic Data Procs	101000000000	28.02	8.62	3412000000	0.78	5.0
12	ADSK	Autodesk Inc	52349884000	51.54	4.75	906000000	1.43	N
13	AEE	Ameren Corp	19378014000	16.68	4.37	1152000000	0.44	2.0
14	AEP	American Electric Power Company	44207232000	16.02	5.24	2208100000	0.51	3.0

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>15</b>	AES	The Aes Corp	12851409000	10.31	1.76	249000000	1.08	0.31
<b>16</b>	AFL	Aflac Inc	49151352000	13.56	6.22	4659000000	0.91	1.11
<b>17</b>	AIG	American International Group	52916936000	11.30	6.78	3643000000	1.03	1.21
<b>18</b>	AIZ	Assurant Inc	9348244000	11.56	15.51	642500000	0.50	2.11
<b>19</b>	AJG	Arthur J. Gallagher & Company	53263752000	27.37	8.78	969500000	0.68	2.01
<b>20</b>	AKAM	Akamai Technologies	16098545000	25.36	4.18	547630000	0.72	N/A
<b>21</b>	ALB	Albemarle Corp	14508461000	5.45	22.24	1573480000	1.62	1.11
<b>22</b>	ALGN	Align Technology	23851522000	47.79	6.59	445050000	1.65	N/A
<b>23</b>	ALL	Allstate Corp	45463240000	188.16	0.91	-188000000	0.52	3.31
<b>24</b>	ALLE	Allegion Plc	11584270000	18.82	6.96	540400000	1.14	1.01
<b>25</b>	AMAT	Applied Materials	173000000000	24.96	8.15	6856000000	1.59	1.01
<b>26</b>	AMCR	Amcor Plc	13311609000	13.29	0.69	1048000000	0.78	0.31
<b>27</b>	AMD	Adv Micro Devices	275000000000	85.04	1.95	854000000	1.64	N/A
<b>28</b>	AME	Ametek Inc	42131448000	28.11	6.38	1313190000	1.22	1.11
<b>29</b>	AMGN	Amgen Inc	145000000000	14.37	18.65	6717000000	0.58	8.01

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>30</b>	AMP	Ameriprise Financial Services	43564048000	14.24	30.12	2556000000	1.36	5.1%
<b>31</b>	AMT	American Tower Corp	88686160000	19.28	3.18	1483300000	0.67	6.1%
<b>32</b>	AMZN	Amazon.com Inc	1920000000000	64.29	2.80	30425000000	1.16	N/A
<b>33</b>	ANET	Arista Networks Inc	93039880000	46.95	6.15	2087320000	1.09	N/A
<b>34</b>	ANSS	Ansys Inc	29748840000	50.25	6.79	500410000	1.16	N/A
<b>35</b>	AON	AON Plc	63255080000	22.26	14.14	2564000000	0.86	2.2%
<b>36</b>	AOS	Smith A.O. Corp	12885451000	22.81	3.82	556600000	1.25	1.1%
<b>37</b>	APA	Apa Corp	10787011000	7.84	4.52	2855000000	3.27	1.1%
<b>38</b>	APD	Air Products and Chemicals	53058800000	20.28	11.69	2300200000	0.81	7.1%
<b>39</b>	APH	Amphenol Corp	70410976000	38.24	3.01	1928000000	1.30	0.9%
<b>40</b>	APTV	Aptiv Plc	20377302000	15.49	4.86	2938000000	1.90	N/A
<b>41</b>	ARE	Alexandria Real Estate Equities	21795644000	13.66	0.54	103640000	1.05	5.1%
<b>42</b>	ATO	Atmos Energy Corp	17532134000	18.43	6.30	885860000	0.66	3.1%
<b>43</b>	AVB	Avalonbay Communities	25746680000	16.95	6.55	928830000	0.92	6.1%
<b>44</b>	AVGO	Broadcom Ltd	621000000000	35.29	37.33	14082000000	1.25	19.1%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
45	AVY	Avery Dennison Corp	17562216000	27.56	7.88	503000000	0.90	3.00
46	AWK	American Water Works	23090302000	24.62	4.89	944000000	0.63	2.00
47	AXON	Axon Enterprise Inc	23385980000	109.97	2.79	174230000	0.93	N/A
48	AXP	American Express Company	161000000000	19.59	11.21	8374000000	1.23	2.00
49	AZO	Autozone	53715432000	21.72	142.02	2528430000	0.70	N/A
50	BA	Boeing Company	112000000000	0.00	-5.82	-2222000000	1.52	N/A
51	BAC	Bank of America Corp	293000000000	10.80	3.42	26515000000	1.39	0.00
52	BALL	Ball Corporation	20980724000	22.68	2.91	707000000	0.89	0.00
53	BAX	Baxter International Inc	21643584000	15.79	2.70	2656000000	0.62	1.00
54	BBWI	Bath & Body Works Inc	10230565000	13.81	3.27	878000000	1.86	0.00
55	BBY	Best Buy Company	17146482000	12.38	6.38	1241000000	1.49	3.00
56	BDX	Becton Dickinson and Company	71119008000	20.39	11.92	1484000000	0.43	3.00
57	BEN	Franklin Resources	14243392000	9.96	2.73	882800000	1.40	1.00
58	BF.B	Brown Forman Inc Cl B	24089272000	25.58	2.01	783000000	0.77	0.00
59	BG	Bunge Ltd	15325646000	7.72	13.67	2243000000	0.64	2.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
60	BIIB	Biogen Inc	29634744000	13.94	14.73	1161100000	-0.02	N/A
61	BIO	Bio-Rad Laboratories	9256654000	27.43	11.77	-637320000	0.87	N/A
62	BK	Bank of New York Mellon Corp	42717900000	11.22	5.06	32860000000	1.12	1.12
63	BKNG	Booking Holdings Inc	123000000000	22.94	153.54	4289000000	1.40	8.00
64	BKR	Baker Hughes Company	34340228000	21.20	1.60	1943000000	1.42	0.40
65	BLDR	Builders Firstsource	24892832000	13.71	14.64	1540560000	2.04	N/A
66	BLK	Blackrock Inc	119000000000	20.97	37.78	5502000000	1.37	20.00
67	BMY	Bristol-Myers Squibb Company	104000000000	6.85	7.50	8025000000	0.39	2.00
68	BR	Broadridge Financial Solutions Llc	23878272000	27.32	7.27	630600000	0.99	3.00
69	BRK.B	Berkshire Hathaway Cl B	905000000000	24.16	17.19	96223000000	0.89	N/A
70	BRO	Brown & Brown	24296788000	29.87	2.81	870500000	0.78	0.40
71	BSX	Boston Scientific Corp	101000000000	33.05	2.05	1591000000	0.78	N/A
72	BWA	Borgwarner Inc	7952686000	7.97	4.32	625000000	1.31	0.40
73	BX	Blackstone Inc	91174280000	31.90	3.95	1390880000	1.51	3.00
74	BXP	Boston Properties	9632625000	8.45	1.21	190220000	1.15	3.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
75	C	Citigroup Inc	118000000000	10.90	5.59	9228000000	1.52	2.10
76	CAG	Conagra Brands Inc	14861176000	11.43	2.75	683600000	0.42	1.00
77	CAH	Cardinal Health	26736170000	16.10	6.84	261000000	0.68	2.00
78	CARR	Carrier Global Corp	50801816000	20.46	2.73	1349000000	1.39	0.00
79	CAT	Caterpillar Inc	189000000000	17.43	21.21	10335000000	1.17	5.00
80	CB	Chubb Ltd	103000000000	11.17	22.58	9028000000	0.62	3.00
81	CBOE	CBOE Global Markets Inc	19102950000	23.28	7.80	761400000	0.60	2.00
82	CBRE	CBRE Group Inc	29389616000	24.86	3.84	986000000	1.42	N/A
83	CCI	Crown Castle Inc	43790924000	13.42	3.46	1502000000	0.73	6.00
84	CCL	Carnival Corp	16969478000	44.29	0.34	-75000000	2.54	N/A
85	CDNS	Cadence Design Sys	83855056000	76.05	4.03	1041140000	1.04	N/A
86	CDW	CDW Corp	34250324000	26.47	9.54	1104300000	1.09	2.00
87	CE	Celanese Corp	17710292000	18.31	8.92	1960000000	1.36	2.00
88	CEG	Constellation Energy Corp	60843564000	36.62	5.00	1623000000	0.70	1.00
89	CF	Cf Industries Holdings	15977359000	10.92	7.89	1525000000	1.03	1.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
90	CFG	Citizens Financial Group Inc/Ri	15577416000	10.96	3.11	1608000000	1.38	1.10
91	CHD	Church & Dwight Company	24878308000	32.31	3.16	755600000	0.53	1.00
92	CHRW	C.H. Robinson Ww	8327633000	22.26	3.22	325130000	0.72	2.20
93	CHTR	Charter Communications Inc	38846236000	9.03	30.02	4557000000	0.96	N/A
94	CI	The Cigna Group	106000000000	14.36	25.10	5164000000	0.56	5.10
95	CINF	Cincinnati Financial	19037930000	19.91	6.04	1843000000	0.63	3.00
96	CL	Colgate-Palmolive Company	72236128000	27.06	3.23	2300000000	0.42	1.10
97	CLX	Clorox Company	18112030000	25.14	5.83	149000000	0.43	4.00
98	CMA	Comerica Inc	6974274000	6.85	7.70	881000000	1.27	2.00
99	CMCSA	Comcast Corp A	162000000000	10.36	3.97	15389000000	0.98	1.00
100	CME	CME Group Inc	75987744000	22.72	9.34	3226200000	0.49	9.00
101	CMG	Chipotle Mexican Grill	79669488000	63.65	44.87	1228740000	1.28	N/A
102	CMI	Cummins Inc	41119328000	15.04	19.60	735000000	1.04	6.00
103	CMS	Cms Energy Corp	17513946000	19.06	3.11	887000000	0.38	1.00
104	CNC	Centene Corp	38930456000	10.82	6.66	2702000000	0.38	N/A

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
105	CNP	Centerpoint Energy Inc	17756518000	18.73	1.50	917000000	0.92	0.00
106	COF	Capital One Financial Corp	54593860000	11.35	12.52	4887000000	1.48	2.00
107	COO	Cooper Companies	19597342000	29.82	3.33	294200000	0.94	0.00
108	COP	Conocophillips	157000000000	15.01	8.78	10957000000	1.24	3.00
109	COR	Cencora Inc	48518008000	19.27	12.56	1745290000	0.45	1.00
110	COST	Costco Wholesale	317000000000	45.53	15.48	6292000000	0.77	4.00
111	CPAY	Corpay Inc	22107638000	19.34	15.76	981890000	1.23	N/A
112	CPB	Campbell Soup Company	12907860000	15.22	2.89	858000000	0.24	1.00
113	CPRT	Copart Inc	54372824000	40.69	1.37	1237740000	1.22	N/A
114	CPT	Camden Property Trust	10516990000	14.06	3.70	403310000	0.83	4.00
115	CRL	Charles River Laboratories Intl	13037435000	23.54	10.65	474620000	1.40	N/A
116	CRM	Salesforce Inc	293000000000	47.60	6.18	4136000000	1.28	0.00
117	CSCO	Cisco Systems Inc	197000000000	13.40	3.59	12613000000	0.87	1.00
118	CSGP	Costar Group Inc	38696752000	86.61	1.06	374700000	0.84	N/A
119	CSX	CSX Corp	71060624000	19.63	1.84	3715000000	1.18	0.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
120	CTAS	Cintas Corp	68292632000	45.89	14.48	1348010000	1.27	5.10
121	CTLT	Catalent Inc	10256882000	0.00	-0.56	-2560000000	1.20	N/A
122	CTRA	Coterra Energy Inc	21292308000	12.95	2.19	1625000000	0.21	0.10
123	CTSH	Cognizant Tech Sol	35411500000	15.74	4.55	2126000000	1.06	1.10
124	CTVA	Corteva Inc	39731328000	21.25	2.68	735000000	0.77	0.00
125	CVS	CVS Corp	93855200000	8.46	8.74	8344000000	0.52	2.20
126	CVX	Chevron Corp	300000000000	12.24	13.13	21369000000	1.12	6.00
127	CZR	Caesars Entertainment Inc	9281433000	46.52	0.91	786000000	2.81	N/A
128	D	Dominion Energy Inc	40681892000	18.86	2.58	1994000000	0.61	2.00
129	DAL	Delta Air Lines Inc	29631502000	7.35	6.24	4609000000	1.43	0.00
130	DAY	Dayforce Inc	9495504000	101.92	0.63	54800000	1.30	N/A
131	DD	Dupont Denemours Inc	31540044000	21.77	3.48	423000000	1.40	1.10
132	DE	Deere & Company	115000000000	11.86	34.34	10166000000	1.04	5.00
133	DECK	Deckers Outdoor Corp	22585008000	31.58	27.80	516820000	1.01	N/A
134	DFS	Discover Financial Services	31536112000	11.11	11.25	2940000000	1.45	2.20

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
135	DG	Dollar General Corp	35048512000	21.04	7.56	1661270000	0.42	2.1%
136	DGX	Quest Diagnostics Inc	14507245000	15.10	8.71	854000000	0.91	2.1%
137	DHI	D.R. Horton	52599632000	11.23	13.90	4745700000	1.64	1.1%
138	DHR	Danaher Corp	181000000000	28.39	8.52	4764000000	0.83	1.1%
139	DIS	Walt Disney Company	217000000000	29.27	4.00	2354000000	1.42	0.0%
140	DLR	Digital Realty Trust	44533124000	21.23	2.97	948840000	0.54	4.1%
141	DLTR	Dollar Tree Inc	28492560000	21.99	5.90	-998400000	0.90	N/A
142	DOC	Healthpeak Properties Inc	13219420000	10.30	0.56	306010000	1.02	1.1%
143	DOV	Dover Corp	24156198000	19.70	8.79	1056830000	1.27	2.0%
144	DOW	Dow Inc	41919872000	26.72	2.24	589000000	1.31	2.1%
145	DPZ	Domino's Pizza Inc	17177658000	32.91	14.67	519120000	0.86	5.5%
146	DRI	Darden Restaurants	18661780000	17.97	8.82	981900000	1.25	5.5%
147	DTE	Dte Energy Company	22575518000	19.12	5.73	1397000000	0.64	3.0%
148	DUK	Duke Energy Corp	73990544000	17.29	5.56	2841000000	0.46	4.1%
149	DVA	Davita Healthcare Partners Inc	11717596000	15.87	8.38	691540000	1.04	N/A

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
150	DVN	Devon Energy Corp	33826448000	9.22	5.70	3747000000	2.19	2.10
151	DXCM	Dexcom Inc	53559600000	88.30	1.51	541500000	1.20	N/A
152	EA	Electronic Arts Inc	34833032000	23.05	5.62	802000000	0.76	0.10
153	EBAY	Ebay Inc	26967080000	16.32	3.15	2767000000	1.33	1.00
154	ECL	Ecolab Inc	65116456000	43.11	5.21	1372300000	1.10	2.00
155	ED	Consolidated Edison Company	30974972000	17.82	5.06	2519000000	0.35	3.00
156	EFX	Equifax Inc	31821770000	37.56	6.71	545300000	1.51	1.00
157	EG	Everest Re Group	16479081000	5.76	65.84	2517000000	0.56	6.00
158	EIX	Edison International	27243038000	14.67	4.76	1407000000	0.96	3.00
159	EL	Estee Lauder Companies	51774248000	94.32	1.53	1006000000	1.07	2.00
160	ELV	Elevance Health Inc	118000000000	15.06	33.11	5987000000	0.78	6.00
161	EMN	Eastman Chemical Company	11740306000	15.57	6.40	894000000	1.54	3.00
162	EMR	Emerson Electric Company	65379432000	23.12	4.89	13219000000	1.37	2.00
163	ENPH	Enphase Energy Inc	15241662000	35.63	3.39	438940000	1.47	N/A
164	EOG	Eog Resources	78194864000	11.47	11.69	7594000000	1.40	4.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
165	EPAM	Epam Systems Inc	15287141000	30.70	8.66	417080000	1.43	N/A
166	EQIX	Equinix Inc	74221664000	24.32	10.31	969180000	0.58	15.00
167	EQR	Equity Residential	23452640000	16.25	2.20	835440000	0.84	2.00
168	EQT	Eqt Corp	16066776000	15.91	2.31	1735230000	1.12	0.00
169	ES	Eversource Energy	20521036000	13.61	4.33	-442240000	0.58	2.00
170	ESS	Essex Property Trust	15260482000	15.70	6.31	405830000	0.80	9.00
171	ETN	Eaton Corp	132000000000	35.14	9.11	3218000000	1.07	3.00
172	ETR	Entergy Corp	22226318000	15.41	6.77	2362310000	0.71	4.00
173	ETSY	Etsy Inc	7699610000	23.58	2.78	307570000	2.01	N/A
174	EVRG	Evergy Inc	11992741000	14.84	3.55	731300000	0.57	2.00
175	EW	Edwards Lifesciences Corp	55798632000	36.33	2.51	1402400000	1.05	N/A
176	EXC	Exelon Corp	37322788000	15.64	2.38	2328000000	0.54	1.00
177	EXPD	Expeditors Intl	16649969000	23.23	5.00	752880000	0.99	1.00
178	EXPE	Expedia Group Inc	17961394000	17.11	7.64	797000000	1.86	N/A
179	EXR	Extra Space Storage Inc	30964240000	17.80	4.94	803200000	0.74	7.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
180	F	Ford Motor Company	52769844000	6.51	2.03	4347000000	1.63	0.10
181	FANG	Diamondback Energy	36438432000	11.16	18.01	3143000000	1.93	8.00
182	FAST	Fastenal Company	43850208000	37.21	2.02	1155000000	1.05	1.00
183	FCX	Freeport-Mcmoran Inc	70960224000	32.14	1.53	1848000000	2.04	0.00
184	FDS	Factset Research Systems Inc	16510709000	28.77	15.06	468170000	0.72	3.00
185	FDX	Fedex Corp	67308072000	15.89	17.34	3972000000	1.21	5.00
186	FE	Firstenergy Corp	21875362000	14.89	2.57	1102000000	0.48	1.00
187	FFIV	F5 Inc	11315449000	19.67	9.58	394950000	1.11	N/A
188	FI	Fiserv Inc	92659520000	20.55	7.54	3068000000	0.91	N/A
189	FICO	Fair Isaac and Company	30608956000	73.98	16.49	429380000	1.21	N/A
190	FIS	Fidelity National Information Services	41759196000	15.23	4.72	-6654000000	0.97	1.00
191	FITB	Fifth Third Bancorp	24047136000	9.91	3.56	2349000000	1.27	1.00
192	FMC	FMC Corp	7331751000	15.66	3.78	1321500000	0.87	2.00
193	FOX	Fox Corp Cl B	13510222000	8.74	3.25	1239000000	0.85	0.00
194	FOXA	Fox Corp Cl A	14730654000	9.58	3.25	1239000000	0.78	0.00

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
195	FRT	Federal Realty Investment Trust	8268688000	15.17	2.80	236990000	1.21	4.
196	FSLR	First Solar Inc	18688836000	21.43	8.00	830780000	1.41	N
197	FTNT	Fortinet Inc	54106528000	50.53	1.37	1147800000	1.08	N
198	FTV	Fortive Corp	29610792000	24.42	3.43	865800000	1.13	0.
199	GD	General Dynamics Corp	80987952000	24.39	12.02	3315000000	0.68	5.
200	GE	General Electric Company	171000000000	52.64	2.80	9480000000	1.29	0.
201	GEHC	Ge Healthcare Technologies Inc	40285436000	22.32	3.94	1568000000	0.94	0.
202	GEN	Gen Digital Inc	14107556000	12.81	1.72	1349000000	0.85	0.
203	GILD	Gilead Sciences Inc	86632464000	10.35	6.72	5665000000	0.19	3.
204	GIS	General Mills	39633744000	15.20	4.63	2593900000	0.15	2.
205	GL	Globe Life Inc	10198312000	10.34	10.65	970760000	0.69	0.
206	GLW	Corning Inc	27833156000	19.04	1.70	581000000	1.08	1.
207	GM	General Motors Company	51060568000	5.72	7.64	10127000000	1.51	0.
208	GNRC	Generac Holdings Inc	8100757000	24.23	5.42	203090000	1.45	N
209	GOOG	Alphabet Cl C	1910000000000	26.20	5.80	73795000000	1.05	N

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>210</b>	GOOGL	Alphabet Cl A	19000000000000	25.95	5.80	73795000000	1.05	11.1%
<b>211</b>	GPC	Genuine Parts Company	20912058000	15.95	9.33	1316520000	0.94	3.3%
<b>212</b>	GPN	Global Payments Inc	32120302000	12.71	9.91	986230000	0.96	1.1%
<b>213</b>	GRMN	Garmin Ltd	28292860000	25.88	5.60	1289640000	0.98	2.0%
<b>214</b>	GS	Goldman Sachs Group	132000000000	17.80	22.82	8516000000	1.43	10.1%
<b>215</b>	GWW	W.W. Grainger	49861636000	27.10	36.65	1829000000	1.14	7.1%
<b>216</b>	HAL	Halliburton Company	36555304000	12.95	3.14	2638000000	2.00	0.0%
<b>217</b>	HAS	Hasbro Inc	7830588000	22.16	2.52	-1489300000	0.74	2.1%
<b>218</b>	HBAN	Huntington Bchs	19720000000	9.94	1.36	1951000000	1.11	0.0%
<b>219</b>	HCA	Hca Holdings Inc	86962656000	17.09	19.03	5242000000	1.65	2.1%
<b>220</b>	HD	Home Depot	355000000000	23.69	15.10	15143000000	0.98	8.8%
<b>221</b>	HES	Hess Corp	48361084000	30.91	5.05	1382000000	1.27	1.1%
<b>222</b>	HIG	Hartford Financial Services Group	30693150000	11.43	8.91	2504000000	0.89	1.1%
<b>223</b>	HII	Huntington Ingalls Industries	11331342000	16.74	17.10	681000000	0.60	5.0%
<b>224</b>	HLT	Hilton Inc	53379960000	33.57	6.22	1141000000	1.25	0.0%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>225</b>	HOLX	Hologic Inc	18304402000	20.02	3.86	456000000	1.00	N/A
<b>226</b>	HON	Honeywell International Inc	128000000000	21.38	9.17	5658000000	1.02	4.1%
<b>227</b>	HPE	Hewlett Packard Enterprise Comp	23400000000	12.37	1.45	2025000000	1.21	0.6%
<b>228</b>	HPQ	HP Inc	28549110000	8.80	3.37	3263000000	1.05	1.1%
<b>229</b>	HRL	Hormel Foods Corp	19005192000	21.58	1.63	793570000	0.27	1.1%
<b>230</b>	HSIC	Henry Schein Inc	9319254000	16.20	4.50	416000000	0.87	N/A
<b>231</b>	HST	Host Marriott Financial Trust	14269008000	10.40	1.04	740000000	1.31	0.0%
<b>232</b>	HSY	Hershey Foods Corp	39513272000	20.43	9.59	1861790000	0.34	4.1%
<b>233</b>	HUBB	Hubbell Inc B	22366744000	26.93	15.32	759800000	0.94	4.1%
<b>234</b>	HUM	Humana Inc	37777656000	11.93	25.99	2489000000	0.45	2.1%
<b>235</b>	HWM	Howmet Aerospace Inc	27514984000	35.42	1.85	765000000	1.43	0.0%
<b>236</b>	IBM	International Business Machines	173000000000	19.56	9.61	7502000000	0.70	6.1%
<b>237</b>	ICE	Intercontinental Exchange	78946568000	24.28	5.63	2368000000	1.02	1.1%
<b>238</b>	IDXX	Idexx Laboratories	43149876000	51.45	10.07	845040000	1.26	N/A
<b>239</b>	IEX	Idex Corp	18118434000	28.81	8.22	596100000	0.97	2.1%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
240	IFF	International Flavors & Fragrances	21278284000	25.11	3.34	-2567000000	1.23	2.1%
241	ILMN	Illumina Inc	20299474000	144.83	0.87	-1161000000	1.19	N/A
242	INCY	Incyte Corp	12402816000	20.52	2.68	597600000	0.65	N/A
243	INTC	Intel Corp	164000000000	99.32	0.40	1689000000	1.00	0.0%
244	INTU	Intuit Inc	178000000000	58.34	10.65	2384000000	1.21	3.3%
245	INVH	Invitation Homes Inc	21296102000	19.43	0.84	519470000	0.94	1.1%
246	IP	International Paper Company	13427468000	18.11	2.17	288000000	0.99	1.0%
247	IPG	Interpublic Group of Companies	11914689000	10.64	3.00	1098400000	1.13	1.1%
248	IQV	Iqvia Holdings Inc	43823176000	25.68	9.29	1358000000	1.50	N/A
249	IR	Ingersoll Rand Inc	38249768000	32.27	2.86	778700000	1.44	0.0%
250	IRM	Iron Mountain Inc	22984584000	19.63	1.79	184230000	0.94	2.2%
251	ISRG	Intuitive Surg Inc	138000000000	83.83	4.52	1798000000	1.36	N/A
252	IT	Gartner Inc	36673028000	40.98	11.33	882470000	1.29	N/A
253	ITW	Illinois Tool Works Inc	78552008000	26.96	9.71	2957000000	1.13	5.5%
254	IVZ	Invesco Plc	7479247000	10.95	1.51	-96900000	1.44	1.1%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>255</b>	J	Jacobs Engineering Group Inc	18586296000	19.75	7.55	665780000	0.73	1.1%
<b>256</b>	JBHT	J B Hunt Transport	20177196000	27.91	6.97	728290000	1.11	1.1%
<b>257</b>	JBL	Jabil Circuit	17916384000	16.88	8.05	818000000	1.40	0.8%
<b>258</b>	JCI	Johnson Controls Intl	44425484000	19.28	3.34	1849000000	1.24	1.1%
<b>259</b>	JKHY	Jack Henry & Assoc	12389744000	33.41	5.07	366650000	0.66	2.1%
<b>260</b>	JNJ	Johnson & Johnson	367000000000	14.62	10.43	35153000000	0.53	4.1%
<b>261</b>	JNPR	Juniper Networks	12024804000	23.25	1.59	310200000	0.98	0.6%
<b>262</b>	JPM	JP Morgan Chase & Company	569000000000	11.67	16.77	49552000000	1.15	4.1%
<b>263</b>	K	Kellanova	19634056000	13.84	4.16	951000000	0.40	2.1%
<b>264</b>	KDP	Keurig Dr Pepper Inc	42696176000	17.14	1.79	2181000000	0.66	0.6%
<b>265</b>	KEY	Keycorp	13967729000	12.06	1.25	967000000	1.29	0.6%
<b>266</b>	KEYS	Keysight Technologies Inc	27445438000	21.54	7.19	1057000000	1.05	N/A
<b>267</b>	KHC	Kraft Heinz Company	45051544000	12.54	2.97	2855000000	0.67	1.1%
<b>268</b>	KIM	Kimco Realty Corp	12478202000	11.85	1.02	654270000	1.45	1.1%
<b>269</b>	KLAC	K L A-Tencor Corp	92343184000	29.52	22.79	3387280000	1.34	5.1%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
270	KMB	Kimberly-Clark Corp	42555604000	19.25	6.57	1764000000	0.40	4.1%
271	KMI	Kinder Morgan	40969624000	17.22	1.07	2391000000	0.90	1.1%
272	KMX	Carmax Inc	12867270000	28.11	2.87	484760000	1.61	N/A
273	KO	Coca-Cola Company	257000000000	22.04	2.69	10714000000	0.59	1.2%
274	KR	Kroger Company	41359940000	12.07	4.76	2164000000	0.48	1.1%
275	KVUE	Kenvue Inc	38770284000	0.00	0.94	1664000000	NaN	0.0%
276	L	Loews Corp	17033930000	12.12	6.30	1434000000	0.83	0.0%
277	LDOS	Leidos Holdings Inc	17386342000	17.66	7.29	199000000	0.70	1.1%
278	LEN	Lennar Corp	46678072000	11.22	14.59	3938510000	1.53	1.1%
279	LH	Laboratory Corp of America Holdings	17766964000	15.17	13.92	418000000	1.01	2.2%
280	LHX	L3Harris Technologies Inc	39724968000	16.94	12.37	1227000000	0.65	4.0%
281	LIN	Linde Plc	224000000000	32.18	14.21	6199000000	0.93	5.5%
282	LKQ	LKQ Corp	13730961000	13.35	3.83	936000000	1.28	1.1%
283	LLY	Eli Lilly and Company	745000000000	121.53	6.32	5240400000	0.34	4.0%
284	LMT	Lockheed Martin Corp	1100000000000	16.31	27.83	6920000000	0.48	12.0%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
285	LNT	Alliant Energy Corp	12510485000	17.42	2.82	703000000	0.55	1.10
286	LOW	Lowe's Companies	137000000000	18.22	13.06	7726000000	1.13	4.00
287	LRCX	Lam Research Corp	127000000000	34.81	27.34	4510930000	1.51	7.00
288	LULU	Lululemon Athletica	44968476000	27.98	12.78	1550190000	1.30	N/A
289	LUV	Southwest Airlines Company	16909458000	17.82	1.57	465000000	1.15	0.00
290	LVS	Las Vegas Sands	39952324000	28.27	1.86	1221000000	1.16	0.00
291	LW	Lamb Weston Holdings Inc	11517998000	14.82	5.73	1008900000	0.72	1.00
292	LYB	Lyondellbasell Industries NV	34003520000	12.06	8.66	2121000000	1.19	5.00
293	LYV	Live Nation Entertainment	23892416000	77.00	1.33	315840000	1.30	N/A
294	MA	Mastercard Inc	445000000000	38.32	12.26	11195000000	1.08	1.00
295	MAA	Mid-America Apartment Communities	14834604000	13.83	4.71	552810000	0.80	5.00
296	MAR	Marriot Int Cl A	73378408000	24.81	10.03	3083000000	1.62	1.00
297	MAS	Masco Corp	16962142000	19.69	3.89	908000000	1.25	1.00
298	MCD	McDonald's Corp	193000000000	22.62	11.94	8468800000	0.71	6.00
299	MCHP	Microchip Technology	47322832000	15.21	5.67	2237700000	1.65	1.00

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>300</b>	MCK	Mckesson Corp	70796064000	18.81	28.43	35600000000	0.45	2.1%
<b>301</b>	MCO	Moody's Corp	71625776000	39.51	9.91	16070000000	1.27	3.1%
<b>302</b>	MDLZ	Mondelez Intl Inc	92527320000	20.61	3.31	49590000000	0.56	1.1%
<b>303</b>	MDT	Medtronic Inc	112000000000	15.77	5.32	37580000000	0.76	2.1%
<b>304</b>	MET	Metlife Inc	53214272000	9.94	7.36	15780000000	1.04	2.1%
<b>305</b>	META	Meta Platforms Inc	1340000000000	32.77	15.59	39098000000	1.20	0.0%
<b>306</b>	MGM	MGM Resorts International	14750754000	16.80	2.73	1142180000	2.19	N/A
<b>307</b>	MHK	Mohawk Industries	7794332000	13.29	9.19	-439520000	1.34	N/A
<b>308</b>	MKC	Mccormick & Company	20243846000	27.99	2.73	680600000	0.74	2.1%
<b>309</b>	MKTX	Marketaress Holdings	8535069000	32.63	6.89	258060000	0.98	2.1%
<b>310</b>	MLM	Martin Marietta Materials	38293156000	31.25	19.33	1168900000	0.96	2.1%
<b>311</b>	MMC	Marsh & McLennan Companies	101000000000	25.16	7.98	3756000000	0.90	2.1%
<b>312</b>	MMM	3M Company	50372456000	9.80	9.24	-6995000000	1.02	6.1%
<b>313</b>	MNST	Monster Beverage Cp	58171556000	35.74	1.56	1630990000	0.72	N/A
<b>314</b>	MO	Altria Group	73448192000	8.39	4.95	8130000000	0.65	3.1%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
315	MOH	Molina Healthcare Inc	22072688000	17.97	20.89	1091000000	0.47	N/A
316	MOS	Mosaic Company	10599652000	9.18	3.57	1164900000	1.50	0.00
317	MPC	Marathon Petroleum Corp	78960432000	9.19	23.53	9681000000	1.52	3.00
318	MPWR	Monolithic Power Sys	31839372000	72.67	8.85	427370000	1.24	4.00
319	MRK	Merck & Company	324000000000	85.11	1.50	365000000	0.38	3.00
320	MRNA	Moderna Inc	39390696000	0.00	-4.27	-4714000000	1.62	N/A
321	MRO	Marathon Oil Corp	16967152000	11.30	2.61	1554000000	2.22	0.00
322	MS	Morgan Stanley	151000000000	16.93	5.45	9087000000	1.43	3.00
323	MSCI	MSCI Inc	42970300000	39.44	13.53	1148590000	1.06	5.00
324	MSFT	Microsoft Corp	3160000000000	37.78	11.06	72361000000	0.89	2.00
325	MSI	Motorola Solutions	58525128000	31.66	10.99	1709000000	0.90	3.00
326	MTB	M&T Bank Corp	23112000000	8.59	16.07	2741000000	0.80	5.00
327	MTCH	Match Group Inc	9034684000	14.79	2.28	651540000	1.41	N/A
328	MTD	Mettler-Toledo International	27822364000	33.73	38.08	788780000	1.15	N/A
329	MU	Micron Technology	1370000000000	0.00	-3.63	-5833000000	1.27	0.00

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>330</b>	NCLH	Norwegian Cruise Ord	8040661000	60.74	0.31	166180000	2.60	1.1%
<b>331</b>	NDAQ	Nasdaq Inc	35685840000	21.58	2.83	10590000000	0.93	0.4%
<b>332</b>	NDSN	Nordson Corp	15506468000	29.18	9.28	487490000	1.00	2.1%
<b>333</b>	NEE	Nextera Energy	132000000000	20.04	3.18	7310000000	0.51	1.1%
<b>334</b>	NEM	Newmont Mining Corp	45708004000	23.72	1.59	-2494000000	0.50	1.1%
<b>335</b>	NFLX	Netflix Inc	275000000000	51.39	12.01	5407990000	1.22	1.1%
<b>336</b>	NI	NiSource Inc	12186231000	16.98	1.60	714300000	0.49	1.0%
<b>337</b>	NKE	Nike Inc	135000000000	24.67	3.61	5070000000	1.08	1.1%
<b>338</b>	NOC	Northrop Grumman Corp	67356896000	19.73	23.29	2056000000	0.35	7.1%
<b>339</b>	NOW	Servicenow Inc	161000000000	192.27	3.94	1731000000	0.97	1.1%
<b>340</b>	NRG	NRG Energy	15626192000	14.98	4.72	-202000000	1.13	1.1%
<b>341</b>	NSC	Norfolk Southern Corp	61561136000	21.15	11.75	1827000000	1.30	5.1%
<b>342</b>	NTAP	Netapp Inc	21679902000	21.52	4.87	1274000000	1.28	2.1%
<b>343</b>	NTRS	Northern Trust Corp	17704768000	13.85	6.25	1107300000	1.13	3.1%
<b>344</b>	NUE	Nucor Corp	48228780000	11.07	17.99	4524800000	1.61	2.1%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>	
<b>345</b>	NVDA	Nvidia Corp	2200000000000		71.23	12.06	29760000000	1.73	0.
<b>346</b>	NVR	NVR Inc	25290224000		16.82	463.25	1591610000	1.13	N
<b>347</b>	NWS	News Corp Cl B	15111809000		40.58	0.65	149000000	1.35	0.
<b>348</b>	NWSA	News Corp Cl A	14654914000		39.31	0.65	149000000	1.33	0.
<b>349</b>	NXPI	Nxp Semiconductors	61809184000		19.09	12.53	2797000000	1.59	4.0
<b>350</b>	O	Realty Income Corp	45648184000		13.12	1.26	872310000	0.92	3.1
<b>351</b>	ODFL	Old Dominion Freight Line Inc	48883048000		38.73	5.64	1239500000	1.02	0.
<b>352</b>	OKE	Oneok Inc	46486928000		14.27	5.55	2659000000	1.65	3.1
<b>353</b>	OMC	Omnicom Group Inc	18264664000		12.54	7.43	1391400000	0.97	2.1
<b>354</b>	ON	On Semiconductor	29855984000		13.39	5.16	2183700000	1.80	N
<b>355</b>	ORCL	Oracle Corp	343000000000		27.78	4.47	8503000000	1.01	1.1
<b>356</b>	ORLY	O'Reilly Automotive	65966220000		28.58	38.48	2346580000	0.86	N
<b>357</b>	OTIS	Otis Worldwide Corp	39641332000		27.45	3.54	1406000000	1.01	1.1
<b>358</b>	OXY	Occidental Petroleum Corp	61387704000		18.33	3.69	4696000000	1.64	0.
<b>359</b>	PANW	Palo Alto Networks Inc	86917136000		101.61	2.61	439700000	1.19	N

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
360	PARA	Paramount Global Cl B	7810246000	23.34	0.53	-608000000	1.77	0.
361	PAYC	Paycom Software Inc	11411786000	33.98	5.82	340790000	1.17	1.
362	PAYX	Paychex Inc	44052276000	26.26	4.48	1557300000	0.93	3.
363	PCAR	Paccar Inc	64039380000	12.53	9.62	4600800000	0.92	4.
364	PCG	Pacific Gas & Electric Company	43897076000	13.47	1.23	2256000000	1.32	0.
365	PEG	Public Service Enterprise Group Inc	33166004000	19.00	3.48	2563000000	0.58	2.
366	PEP	Pepsico Inc	232000000000	22.26	7.62	9074000000	0.53	5.
367	PFE	Pfizer Inc	151000000000	14.56	1.83	2119000000	0.61	1.
368	PFG	Principal Financial Group Inc	19933204000	12.90	6.56	623200000	1.25	2.
369	PG	Procter & Gamble Company	367000000000	24.25	6.41	14653000000	0.44	3.
370	PGR	Progressive Corp	124000000000	33.75	6.20	3902400000	0.35	1.
371	PH	Parker-Hannifin Corp	72766656000	23.04	24.12	2082940000	1.52	5.
372	PHM	Pultegroup	24548456000	9.87	11.53	2602370000	1.56	0.
373	PKG	Packaging Corp of America	16844318000	21.59	8.69	765200000	0.78	5.
374	PLD	Prologis Inc	115000000000	22.23	3.29	3059210000	1.00	3.

Symbol		Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
375	PM	Philip Morris International Inc	139000000000	14.99	6.01	7813000000	0.61	5.
376	PNC	PNC Bank	62603044000	11.13	14.10	5578000000	1.18	6.
377	PNR	Pentair Ltd	13873958000	22.08	3.75	622700000	1.20	0.
378	PNW	Pinnacle West Capital Corp	8252949000	16.56	4.41	501560000	0.49	3.
379	PODD	Insulet Corp	11537790000	60.02	2.72	206300000	0.98	N
380	POOL	Pool Corp	15250078000	30.05	13.15	523230000	1.00	4.
381	PPG	PPG Industries	32724590000	18.22	7.67	1270000000	1.28	2.
382	PPL	PPL Corp	20077554000	17.05	1.60	740000000	0.81	0.
383	PRU	Prudential Financial Inc	41539888000	9.97	11.62	2488000000	1.37	5.
384	PSA	Public Storage	49645004000	16.62	11.06	2148330000	0.54	12.
385	PSX	Phillips 66	73050952000	10.80	15.80	7015000000	1.37	4.
386	PTC	Ptc Inc	22150596000	64.42	2.84	245540000	1.17	N
387	PWR	Quanta Services	38361136000	39.62	6.52	744690000	1.11	0.
388	PXD	Pioneer Natural Resources Company	63727680000	12.98	20.79	4894000000	1.31	12.
389	PYPL	Paypal Holdings	69823992000	17.16	3.76	4246000000	1.45	N

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividend_Yield
390	QCOM	Qualcomm Inc	191000000000	24.22	7.01	7232000000	1.41	3.1%
391	QRVO	Qorvo Inc	11111708000	28.11	4.02	103150000	1.56	N/A
392	RCL	Royal Caribbean Cruises Ltd	34913476000	20.19	6.69	1697000000	2.52	N/A
393	REG	Regency Centers Corp	10864711000	14.24	2.04	364560000	1.16	2.0%
394	REGN	Regeneron Pharmaceuticals	103000000000	24.93	37.58	3953600000	0.11	N/A
395	RF	Regions Financial Corp	18515110000	9.04	2.22	2074000000	1.21	0.0%
396	RHI	Robert Half Inc	7916440000	19.49	3.87	411150000	1.26	1.1%
397	RJF	Raymond James Financial	26816198000	15.15	8.41	1739000000	1.05	1.1%
398	RL	Ralph Lauren Corp	10992074000	18.00	9.51	522700000	1.53	3.1%
399	RMD	Resmed Inc	27764520000	27.21	6.80	897560000	0.67	1.1%
400	ROK	Rockwell Automation Inc	31998670000	23.62	11.70	1387400000	1.43	4.0%
401	ROL	Rollins Inc	21983352000	49.52	0.90	434960000	0.63	0.0%
402	ROP	Roper Industries	58114020000	32.15	16.71	1384200000	0.99	2.0%
403	ROST	Ross Stores Inc	46979752000	25.13	5.56	1874520000	1.04	1.1%
404	RSG	Republic Services	59272712000	33.35	5.60	1731000000	0.68	2.0%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>405</b>	RTX	Rtx Corp	135000000000	19.67	5.05	3195000000	0.88	2.1%
<b>406</b>	RVTY	Revvity Inc	12634138000	21.87	4.65	693090000	1.09	0.1%
<b>407</b>	SBAC	SBA Communications	23070460000	16.34	4.61	501810000	0.56	3.3%
<b>408</b>	SBUX	Starbucks Corp	98331560000	23.68	3.70	4124500000	0.97	2.2%
<b>409</b>	SCHW	The Charles Schwab Corp	127000000000	22.80	3.13	5067000000	1.01	1.0%
<b>410</b>	SHW	Sherwin-Williams Company	84444208000	31.90	10.34	2388800000	1.15	2.2%
<b>411</b>	SJM	J.M. Smucker Company	12445951000	11.97	9.92	-91300000	0.21	4.4%
<b>412</b>	SLB	Schlumberger N.V.	78549544000	18.15	2.99	4203000000	1.60	0.0%
<b>413</b>	SMCI	Super Micro Computer	53025596000	75.43	12.70	640000000	1.28	N/A
<b>414</b>	SNA	Snap-On Inc	15468854000	15.64	18.75	1011100000	0.99	6.1%
<b>415</b>	SNPS	Synopsys Inc	87756128000	62.39	9.09	1229890000	1.06	N/A
<b>416</b>	SO	Southern Company	76525800000	19.24	3.64	3976000000	0.49	2.1%
<b>417</b>	SPG	Simon Property Group	48933728000	11.83	6.98	2283130000	1.65	7.0%
<b>418</b>	SPGI	S&P Global Inc	138000000000	33.97	12.61	2626000000	1.15	3.3%
<b>419</b>	SRE	Sempra	44579220000	15.37	4.61	3075000000	0.71	2.2%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>420</b>	STE	Steris Corp	21094812000	24.96	8.55	107030000	0.79	2.1%
<b>421</b>	STLD	Steel Dynamics Inc	23260858000	9.85	14.90	2450880000	1.42	1.0%
<b>422</b>	STT	State Street Corp	22887356000	9.92	7.66	1944000000	1.52	2.1%
<b>423</b>	STX	Seagate Technology Hldgs Plc	18853896000	0.00	-1.08	-529000000	1.04	2.0%
<b>424</b>	STZ	Constellation Brands Inc	48562480000	22.44	11.78	-71000000	1.75	0.1%
<b>425</b>	SWK	Stanley Black & Decker Inc	14674440000	65.48	1.45	-281700000	1.34	3.3%
<b>426</b>	SWKS	Skyworks Solutions	16900638000	15.01	6.98	982800000	1.37	2.0%
<b>427</b>	SYF	Synchrony Financial	16794520000	7.95	5.18	2238000000	1.61	0.0%
<b>428</b>	SYK	Stryker Corp	134000000000	32.77	10.60	3165000000	0.89	3.0%
<b>429</b>	SYY	Sysco Corp	38168624000	18.23	4.20	1770120000	1.15	2.0%
<b>430</b>	T	AT&T Inc	125000000000	7.29	2.41	144000000000	0.57	1.0%
<b>431</b>	TAP	Molson Coors Brewing Company	14527259000	12.41	5.43	948900000	0.84	1.0%
<b>432</b>	TDG	Transdigm Group Inc	68688968000	45.86	26.26	1298000000	1.42	N/A
<b>433</b>	TDY	Teledyne Technologies Inc	19751244000	21.17	19.69	885700000	1.02	N/A
<b>434</b>	TECH	Bio-Techne Cp	10530291000	39.21	1.71	285260000	1.22	0.0%

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
435	TEL	Te Connectivity Ltd	44957216000	20.52	7.04	1910000000	1.39	2.1%
436	TER	Teradyne Inc	16409519000	36.16	2.93	448750000	1.56	0.1%
437	TFC	Truist Financial Corp	50834568000	10.28	3.69	-1091000000	1.08	2.0%
438	TFX	Teleflex Inc	10210084000	15.67	13.52	356330000	1.11	1.1%
439	TGT	Target Corp	79304496000	19.29	8.93	4138000000	1.14	4.1%
440	TJX	TJX Companies	110000000000	25.47	3.76	4474000000	0.86	1.1%
441	TMO	Thermo Fisher Scientific Inc	221000000000	26.49	21.54	5995000000	0.80	1.1%
442	TMUS	T-Mobile US	191000000000	23.32	6.93	8317000000	0.49	1.1%
443	TPR	Tapestry Inc	9846682000	10.03	4.29	936000000	1.57	1.1%
444	TRGP	Targa Resources	25807384000	31.16	3.67	828200000	2.21	2.1%
445	TRMB	Trimble Navigation	15097466000	27.95	2.20	311300000	1.50	N/A
446	TROW	T Rowe Price Group	26085836000	15.30	7.60	1788700000	1.40	4.1%
447	TRV	The Travelers Companies Inc	52902904000	17.40	13.13	2991000000	0.58	3.1%
448	TSCO	Tractor Supply Company	27251750000	25.03	10.09	1107230000	0.84	4.1%
449	TSLA	Tesla Inc	525000000000	65.56	2.61	149990000000	2.39	N/A

<b>Symbol</b>		<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>450</b>	TSN	Tyson Foods	21042348000	50.84	1.17	-648000000	0.77	1.1%
<b>451</b>	TT	Trane Technologies Plc	69182032000	32.87	9.05	20239000000	1.05	3.0%
<b>452</b>	TTWO	Take-Two Interactive	25765304000	96.51	1.55	-11247000000	0.74	N/A
<b>453</b>	TXN	Texas Instruments	152000000000	24.04	7.01	6510000000	1.04	5.0%
<b>454</b>	TXT	Textron Inc	18594204000	17.05	5.60	921000000	1.30	0.1%
<b>455</b>	TYL	Tyler Technologies	17619820000	71.14	5.83	165920000	0.81	N/A
<b>456</b>	UAL	United Airlines Holdings Inc	14167442000	4.30	10.05	2618000000	1.55	N/A
<b>457</b>	UBER	Uber Technologies Inc	161000000000	87.12	0.86	1887000000	1.36	N/A
<b>458</b>	UDR	United Dominion Realty Trust	12019968000	14.73	1.36	444350000	0.80	1.1%
<b>459</b>	UHS	Universal Health Services	11658138000	16.34	10.55	717800000	1.24	0.1%
<b>460</b>	ULTA	Ulta Beauty Inc	21467638000	17.22	26.05	1291010000	1.30	N/A
<b>461</b>	UNH	Unitedhealth Group Inc	420000000000	18.13	25.12	22381000000	0.54	7.7%
<b>462</b>	UNP	Union Pacific Corp	148000000000	23.01	10.46	6379000000	1.07	5.5%
<b>463</b>	UPS	United Parcel Service	129000000000	17.08	8.78	6708000000	1.05	6.6%
<b>464</b>	URI	United Rentals	47480448000	17.00	40.82	2424000000	1.84	6.1%

	<b>Symbol</b>	<b>Name</b>	<b>Market_Cap</b>	<b>Price_Earnings_Ratio</b>	<b>EPS</b>	<b>Net_Income</b>	<b>Beta</b>	<b>Dividend_Yield</b>
<b>465</b>	USB	U.S. Bancorp	67046464000	9.94	4.32	5429000000	1.05	1.1%
<b>466</b>	V	Visa Inc	509000000000	30.48	8.99	17273000000	0.96	1.1%
<b>467</b>	VFC	V.F. Corp	5350108000	11.34	1.22	118580000	1.47	0.1%
<b>468</b>	VICI	Vici Properties Inc	30584776000	13.55	2.48	2513540000	0.90	1.1%
<b>469</b>	VLO	Valero Energy Corp	60973876000	7.31	24.71	8835000000	1.54	4.4%
<b>470</b>	VLTO	Veralto Corporation	21501562000	0.00	1.62	839000000	NaN	0.0%
<b>471</b>	VMC	Vulcan Materials Company	35733280000	38.04	6.99	933200000	0.83	1.1%
<b>472</b>	VRSK	Verisk Analytics Inc	32642730000	39.60	5.72	614600000	0.78	1.1%
<b>473</b>	VRSN	Verisign Inc	18951038000	25.83	7.24	817600000	0.88	N/A
<b>474</b>	VRTX	Vertex Pharmaceutical	105000000000	30.06	13.44	3619600000	0.35	N/A
<b>475</b>	VTR	Ventas Inc	17410472000	14.22	-0.07	-40970000	1.26	1.1%
<b>476</b>	VTRS	Viatris Inc	14156119000	4.03	2.92	54700000	1.13	0.0%
<b>477</b>	VZ	Verizon Communications Inc	177000000000	9.01	4.71	11614000000	0.41	2.1%
<b>478</b>	WAB	Wabtec Corp	26638500000	24.85	5.93	815000000	1.40	0.0%
<b>479</b>	WAT	Waters Corp	19888652000	28.45	11.75	642230000	0.92	N/A

	Symbol	Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
480	WBA	Walgreens Boots Alliance	16503699000	5.38	3.53	-3080000000	0.73	1.1%
481	WBD	Discovery Inc Series A	20298196000	0.00	-0.66	-3126000000	1.51	N/A
482	WDC	Western Digital Cp	23823622000	0.00	-6.61	-1706000000	1.57	N/A
483	WEC	Wisconsin Energy Corp	25323120000	17.50	4.63	1332900000	0.40	3.3%
484	WELL	Welltower Inc	54234476000	24.91	0.64	340090000	1.08	2.2%
485	WFC	Wells Fargo & Company	203000000000	10.98	5.16	19142000000	1.19	1.1%
486	WM	Waste Management	83040384000	33.46	6.19	2304000000	0.72	2.2%
487	WMB	Williams Companies	47798816000	20.48	1.91	3179000000	1.05	1.1%
488	WMT	Walmart Inc	482000000000	26.88	2.21	15511000000	0.49	0.1%
489	WRB	W.R. Berkley Corp	22386468000	17.53	4.94	1381360000	0.57	1.1%
490	WRK	Westrock Company	12542850000	18.29	2.67	-1649000000	1.10	1.1%
491	WST	West Pharmaceutical Services	28844968000	47.79	8.08	593400000	1.04	0.1%
492	WTW	Willis Towers Watson Public Ltd	27766202000	18.28	14.57	1055000000	0.73	3.3%
493	WY	Weyerhaeuser Company	25237452000	33.96	1.02	839000000	1.43	0.1%
494	WYNN	Wynn Resorts Ltd	12043902000	25.88	4.10	729990000	1.93	1.1%

Symbol		Name	Market_Cap	Price_Earnings_Ratio	EPS	Net_Income	Beta	Dividends
495	XEL	XCEL Energy Inc	29717500000	16.10	3.34	1771000000	0.36	2.
496	XOM	Exxon Mobil Corp	482000000000	12.58	9.52	36010000000	0.96	3.
497	XRAY	Dentsply Sirona Inc	6699899000	17.44	1.83	-132000000	0.98	0.
498	XYL	Xylem Inc	31092520000	34.58	3.68	609000000	1.10	1.
499	YUM	Yum! Brands	38289832000	26.36	5.17	1597000000	1.10	2.
500	ZBH	Zimmer Biomet Holdings	26161348000	16.87	7.56	1024000000	1.01	0.
501	ZBRA	Zebra Technologies	14920015000	33.29	8.80	296000000	1.79	N
502	ZTS	Zoetis Inc Cl A	75857568000	31.11	5.32	2344000000	0.85	1.

## The following changes were made to the data

### Flat File data

- The 4th, 12th and 18th column header were renamed, , Changed "P/E" to "Price Earnings", FChanged "%Chg (Ext)" to "Extended Hr %Change" and changed "20D Rel Str" to "20D Relative Strength"
- Changed float numbers to int64 to eliminate exponential notation in the 'Market Cap' column.
- Values in the 'Volume' column contained a special character '?' within the values. Adjustments were made to eliminate this special character.

### Web Scrapped Data

3 headers of the web content table was changed for better context and clarity.

- Changed 'Security' to 'Company Name'
- Changed 'CIK' to 'Central Index Key'
- Changed 'Founded' to 'Year Founded'

## **API data**

9 headers of the web content received was changed for better context and clarity.

- Change header name 'T' as 'Symbol' in column 1
- Change 'v' as 'Volume' in column 2
- Change 'c' as 'Closing Price' in column 5
- Change 'h' as 'Highest Price' in column 6
- Change 'l' as 'Lowest Price' in column 7
- Change 't' as 'Timestamp' in column 8
- Change 'vw' as 'Vol Wt Avg Price' in column 3
- Change 'o' as 'Open Price' in column 4
- Change 'n' as 'Number of Transactions' in column 9

## **Legal or regulatory guidelines**

All three datasets are available in public and are bound to legal and regulatory compliance. These data are obtained from stock exchanges and respective company websites. These information does not influence any financial impact as these are post transaction data and are available through many other sources.

## **No risks created based on the transformations done?**

Following transformations done are to convert the datatype or infer values within the existing values or cosmetic changes

### **Flat File**

- Convert float number to int64 to remove exponential listing or displaying (no value changed)
- Replace special character, replace with a no character or null character (just removed a

### **Web Scrapped Data**

- Remove rows that has no value (removed a row that had no data . might be due to html structure and not a data)

## **API data**

- Convert 'Volume' and "Number of Transactions" to integer values
- Fill the average price of missing values in 'Vol Wt Avg Price' with average price of 'Open price' and 'Closing Price'
- Convert 'Timestamp' field with millisecond values to a timezone field with date and time

## **Assumptions in cleaning/transforming the data**

No assumptions were made on the data and used as is.

Filling the average value in API dataset using high and low price mean value assumed that number of transactions were same between those high and low prices of stock for that day.

## Data Source verification

Data were sources from public domain and like Flat file from S&P500 web site, web page in Wikipedia and API data from a Polygon website. Any contradicted data in Flat file from S&P 500 website or Wikipedia can or will be verified by respective stock company or by public people using those data and feedback provided to rectify the same API data is a feed from stock exchanges and can be verified across other financial or stock websites, to retain subscriptions API data will be free from any corrupted data or human biased data

## Source of Data

Flat file is sourced from S&P website - <https://www.spglobal.com/en> Web page is scraped from [https://en.wikipedia.org/wiki/List\\_of\\_S%26P\\_500\\_companies#S&P\\_500\\_component\\_stocks](https://en.wikipedia.org/wiki/List_of_S%26P_500_companies#S&P_500_component_stocks) and API data from <https://polygon.io/>

## Ethical Implication

As the data is used as it is and no stock buy or sell recommendations are predicted or suggested , there are no ethical implication