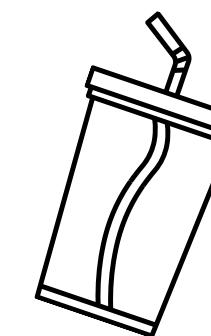
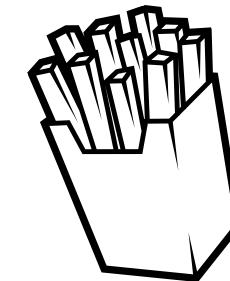
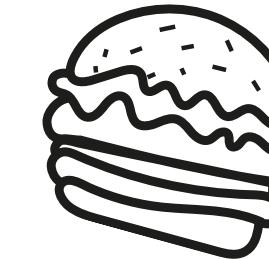


FAST FOOD COMPANIES: A DEEP PERFORMANCE ANALYSIS



Group 12: Julia, Lorenzo, Nicolás , Ricardo, Sassine

Data Preprocessing

- Before delving into the questions, some raw analysis to the datasets need to be done.
 - Check for missing values
 - Imputing missing values

Total Missing Values	
BRK-A	3.287
DNUT	239
DPZ	1.489
LKN CY	397
MCD	4.314
PZZA	2.309
QSR	0
SBUX	2.392
WEN	3.292
YUM	2.003

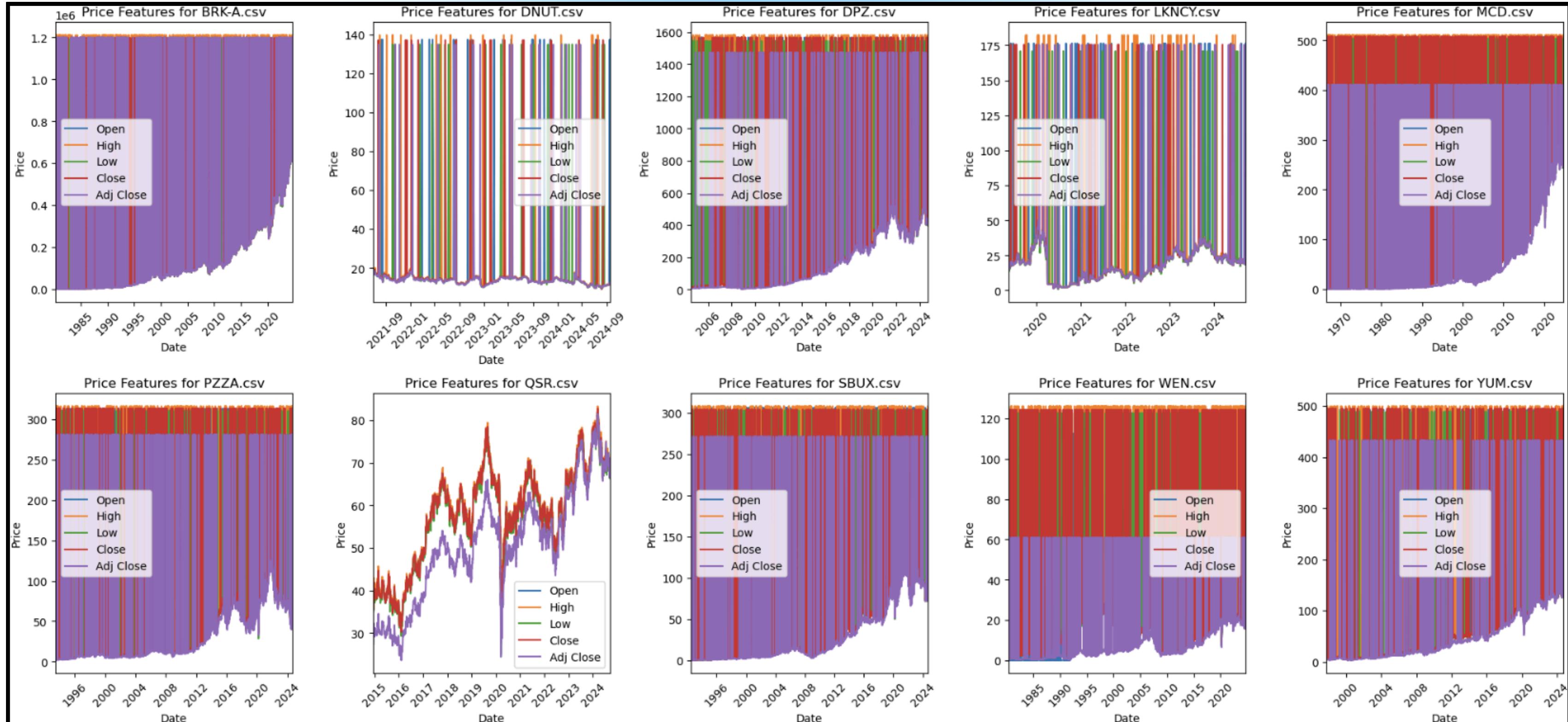
```
# Here we use linear interpolation to impute missing values in each column
for col in variables:
    df[col] = df[col].interpolate(method="linear", limit_direction="both")
```



Total Missing Values	
BRK-A	0
DNUT	0
DPZ	0
LKN CY	0
MCD	0
PZZA	0
QSR	0
SBUX	0
WEN	0
YUM	0

Outlier Removal

- Plotting to check for outliers

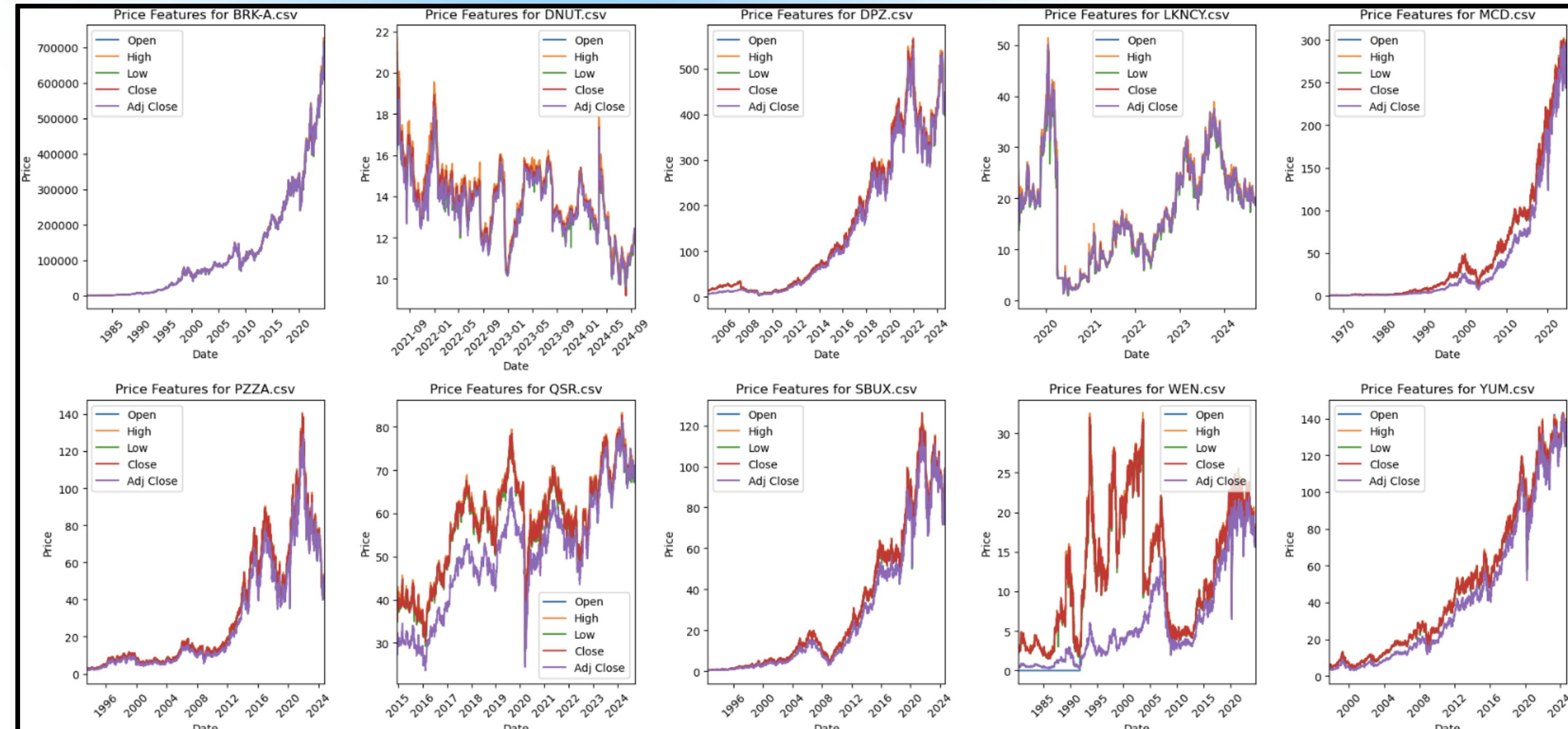


Outlier Removal

- Removing outliers with Median Absolute Deviation (MAD) for each specified column within a rolling window

- Window = 20
- Threshold= 5

- Finally imputing again



Question 1

- We ran through the files using a for loop to display information about their data types. The information is shown in an organised manner.
- Given the pre-processing, the “Date” column has been changed from the previous object to a Datetime dictionary.

```

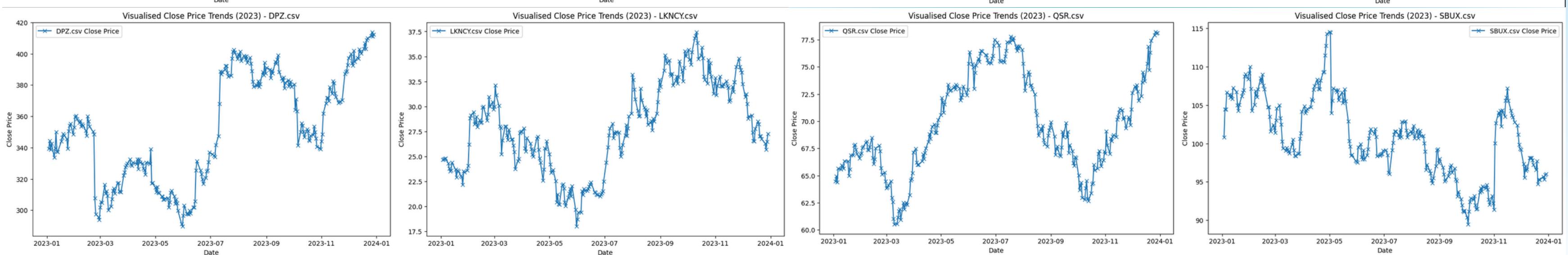
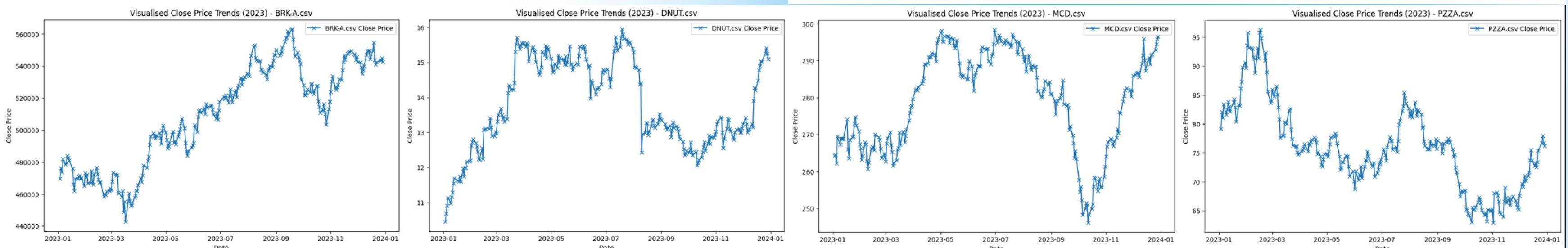
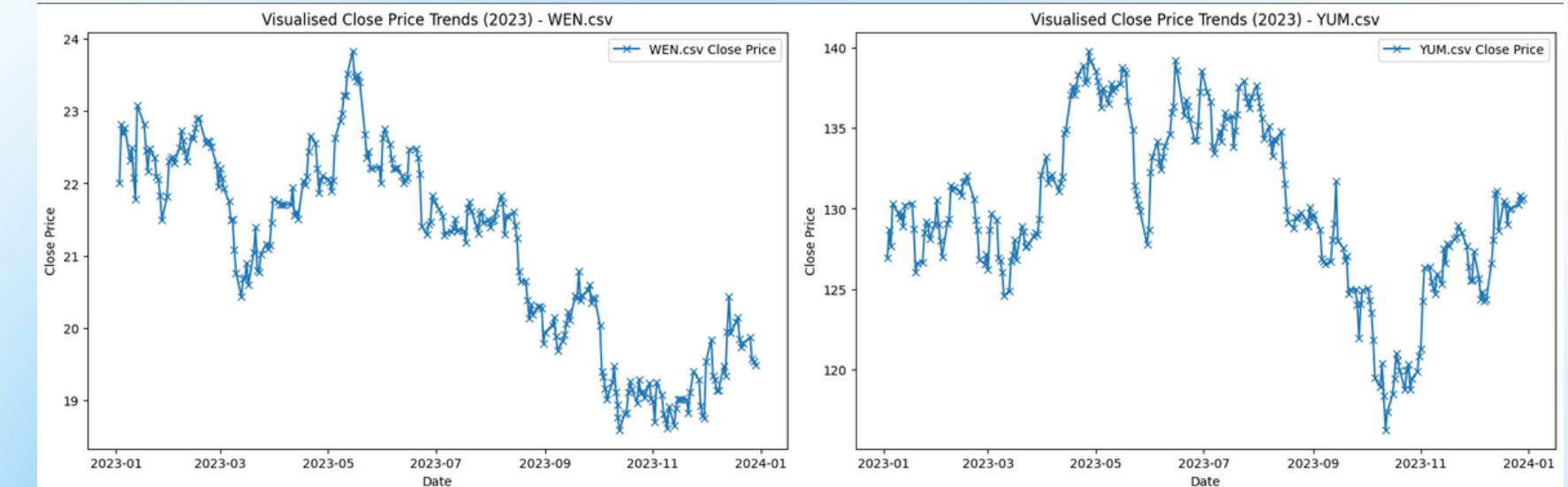
Information for the file 'BRK-A.csv'  Information for the file 'DPZ.csv'  -----
Information for the file 'MCD.csv'  Information for the file 'QSR.csv'  Information for the file 'WEN.csv'
Total rows: 11227  Total rows: 5088  Total rows: 14657  Total rows: 2459  Total rows: 11192
Total columns: 7  Total columns: 7  Total columns: 7  Total columns: 7  Total columns: 7
Date      datetime64[ns]  Date      datetime64[ns]  Date      datetime64[ns]  Date      datetime64[ns]
Open      float64        Open      float64        Open      float64        Open      float64
High      float64        High     float64        High     float64        High     float64
Low       float64        Low      float64        Low      float64        Low      float64
Close     float64        Close    float64        Close    float64        Close   float64
Adj Close float64        Adj Close float64        Adj Close float64        Adj Close float64
Volume    float64        Volume   float64        Volume  int64
dtype: object  dtype: object  dtype: object  dtype: object  dtype: object
-----  -----  -----  -----  -----
Information for the file 'DNUT.csv'  Information for the file 'LKNCY.csv'  Information for the file 'PZZA.csv'  Information for the file 'SBUX.csv'  Information for the file 'YUM.csv'
Total rows: 815  Total rows: 1350  Total rows: 7883  Total rows: 8122  Total rows: 6801
Total columns: 7  Total columns: 7  Total columns: 7  Total columns: 7  Total columns: 7
Date      datetime64[ns]  Date      datetime64[ns]  Date      datetime64[ns]  Date      datetime64[ns]
Open      float64        Open      float64        Open      float64        Open      float64
High      float64        High     float64        High     float64        High     float64
Low       float64        Low      float64        Low      float64        Low      float64
Close     float64        Close    float64        Close    float64        Close   float64
Adj Close float64        Adj Close float64        Adj Close float64        Adj Close float64
Volume    float64        Volume   float64        Volume  float64
dtype: object  dtype: object  dtype: object  dtype: object  dtype: object
-----  -----  -----  -----  -----

```



Question 2

- 2023 columns for “Closing Price”.
- Useful to see an upward trend towards the middle of the year and a low towards the latter months.



Question 3

Company	Date (Most Recent)	Max Close Price
BRK-A	2024-09-03	715910.0
YUM	2024-04-29	143.19
QSR	2024-03-13	82.75
MCD	2024-01-19	300.53
DPZ	2021-12-31	564.33
PZZA	2021-11-04	140.01
SBUX	2021-07-26	126.06
DNUT	2021-07-01	21.0
LKNKY	2020-01-17	50.02
WEN	1993-09-13	32.0



Question 4



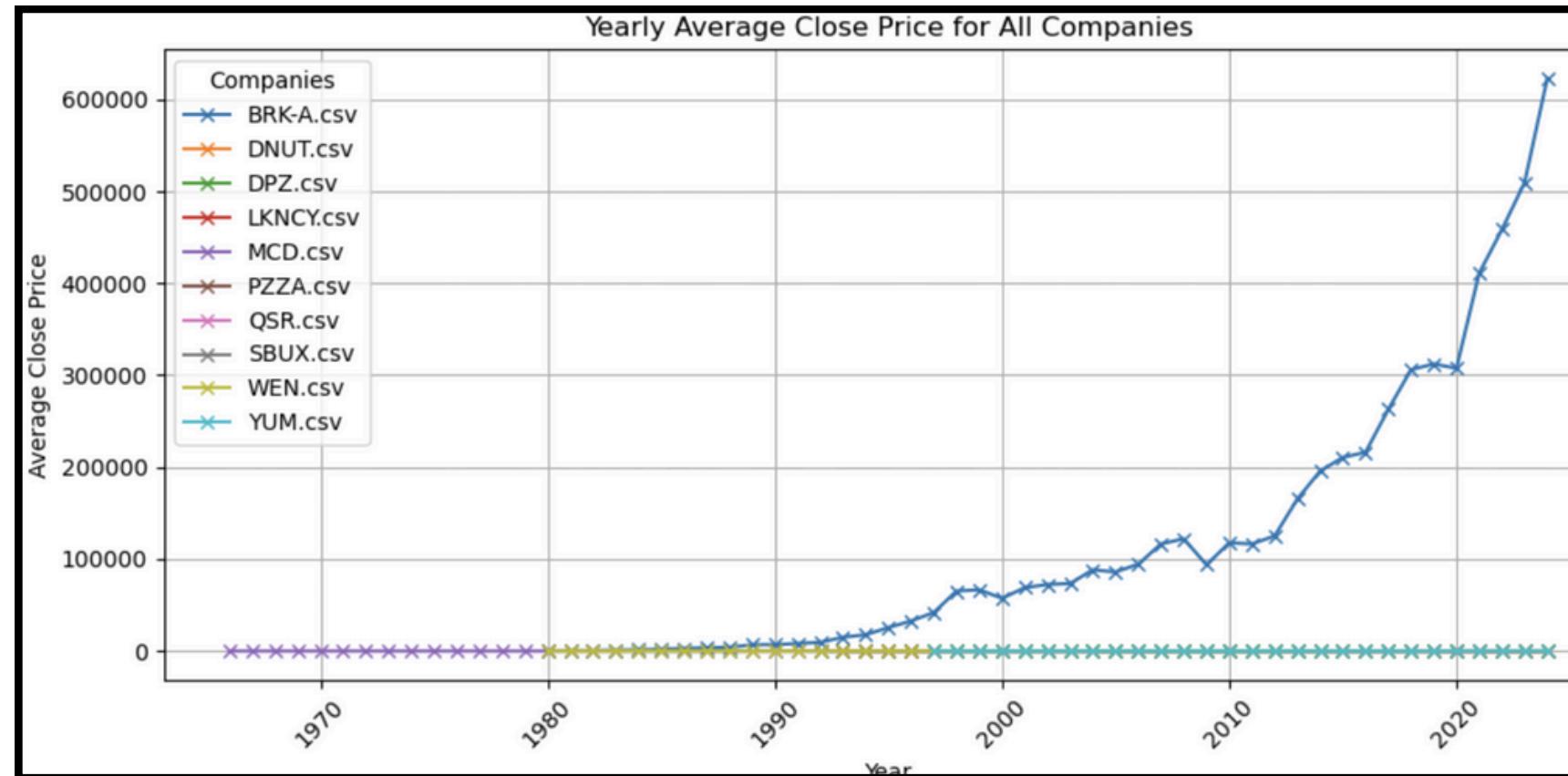
- Companies selected
 - DPZ
 - MCD
 - PZZA

- Line chart
 - time-series data
 - reveals patterns
 - clear comparison

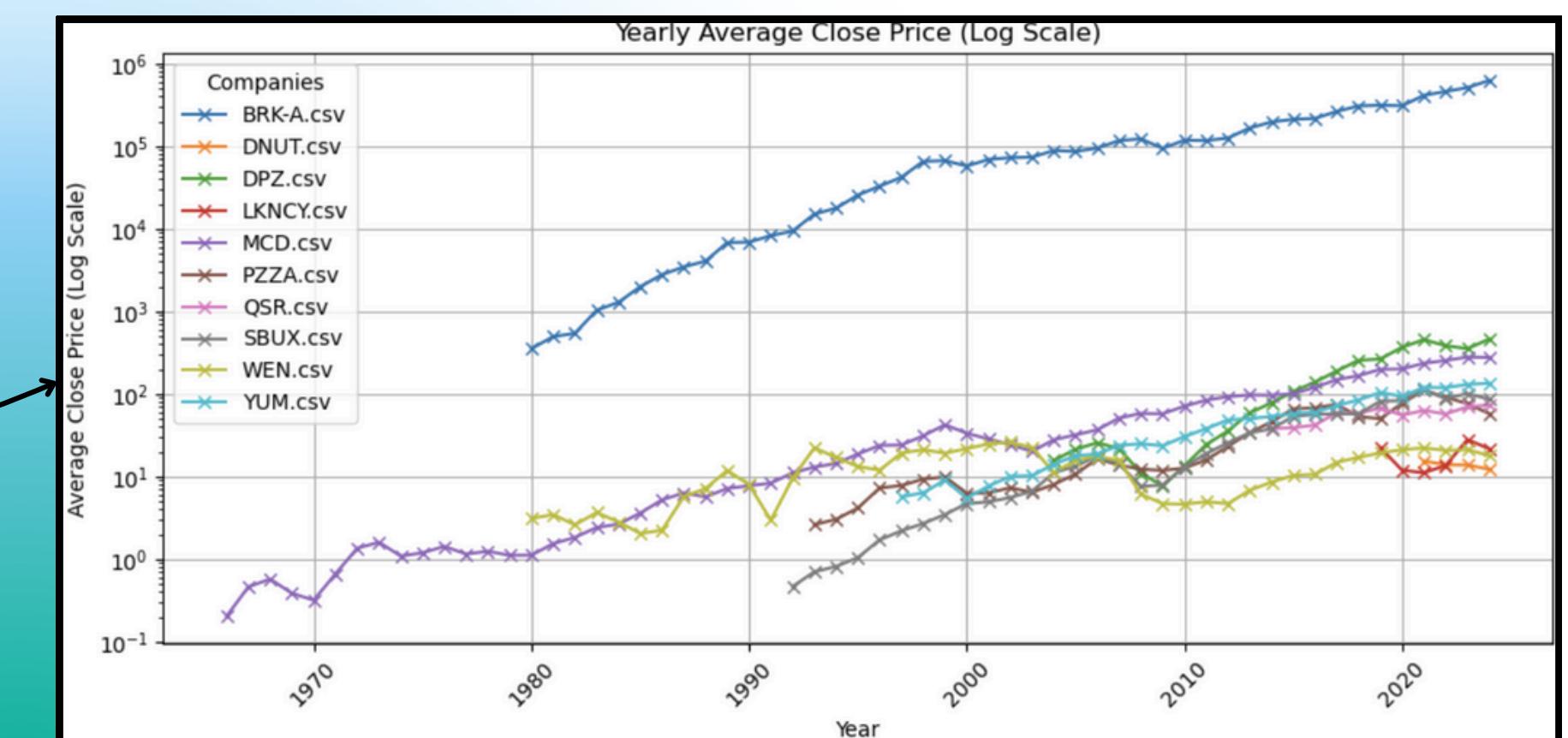


Question 5

- Yearly average of the Close price.

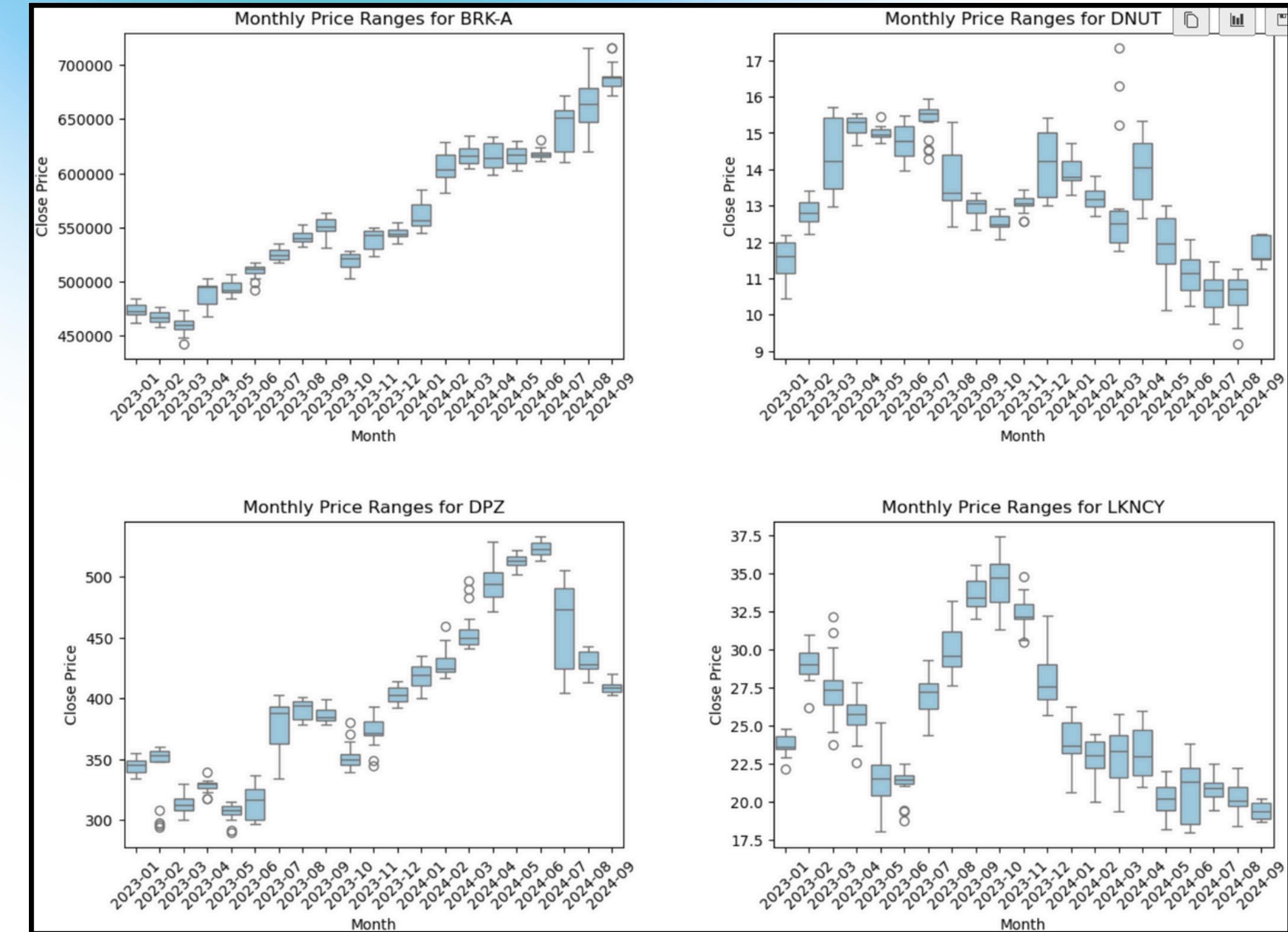


- Yearly average of the Close price using logarithmic scale

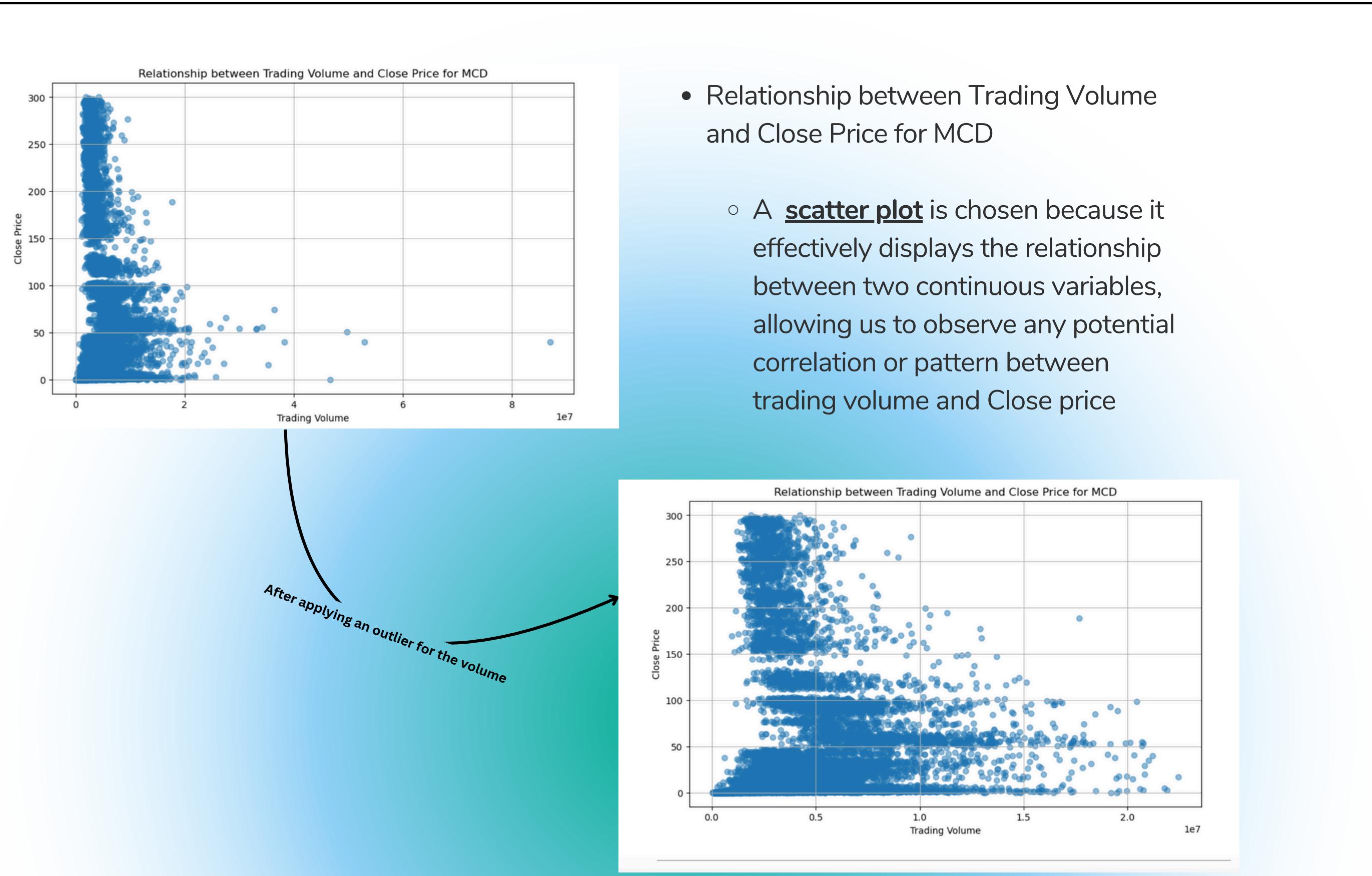


Question 6

- Range of prices for each month
 - A box plot is ideal for this exercise because it effectively visualizes the distribution of monthly average Close prices by showing the median, quartiles, and range of values, while also highlighting any outliers.



Question 7



Question 8

Identify the month with the highest total trading volume

	Company	Month	Total Volume
0	BRK-A.csv	2010-02	11.125.200
1	DNUT.csv	2021-07	91.481.800
2	DPZ.csv	2010-03	33.716.300
3	LKNKY.csv	2020-06	1.398.382.800
4	MCD.csv	2006-10	426.365.450
5	PZZA.csv	1999-12	130.818.800
6	QSR.csv	2020-03	110.310.400
7	SBUX.csv	1999-07	1.259.891.600
8	WEN.csv	2009-06	292.713.350
9	YUM.csv	2015-10	229.313.585

Approach 1

Approach 2

	Company	Month	Total Volume
0	BRK-A.csv	02	33.274.850
1	DNUT.csv	07	160.412.250
2	DPZ.csv	10	353.095.600
3	LKNKY.csv	06	1.636.695.000
4	MCD.csv	10	5.919.338.775
5	PZZA.csv	12	530.247.950
6	QSR.csv	03	382.737.500
7	SBUX.csv	07	11.766.899.350
8	WEN.csv	06	1.721.038.200
9	YUM.csv	10	2.995.973.010

Question 9

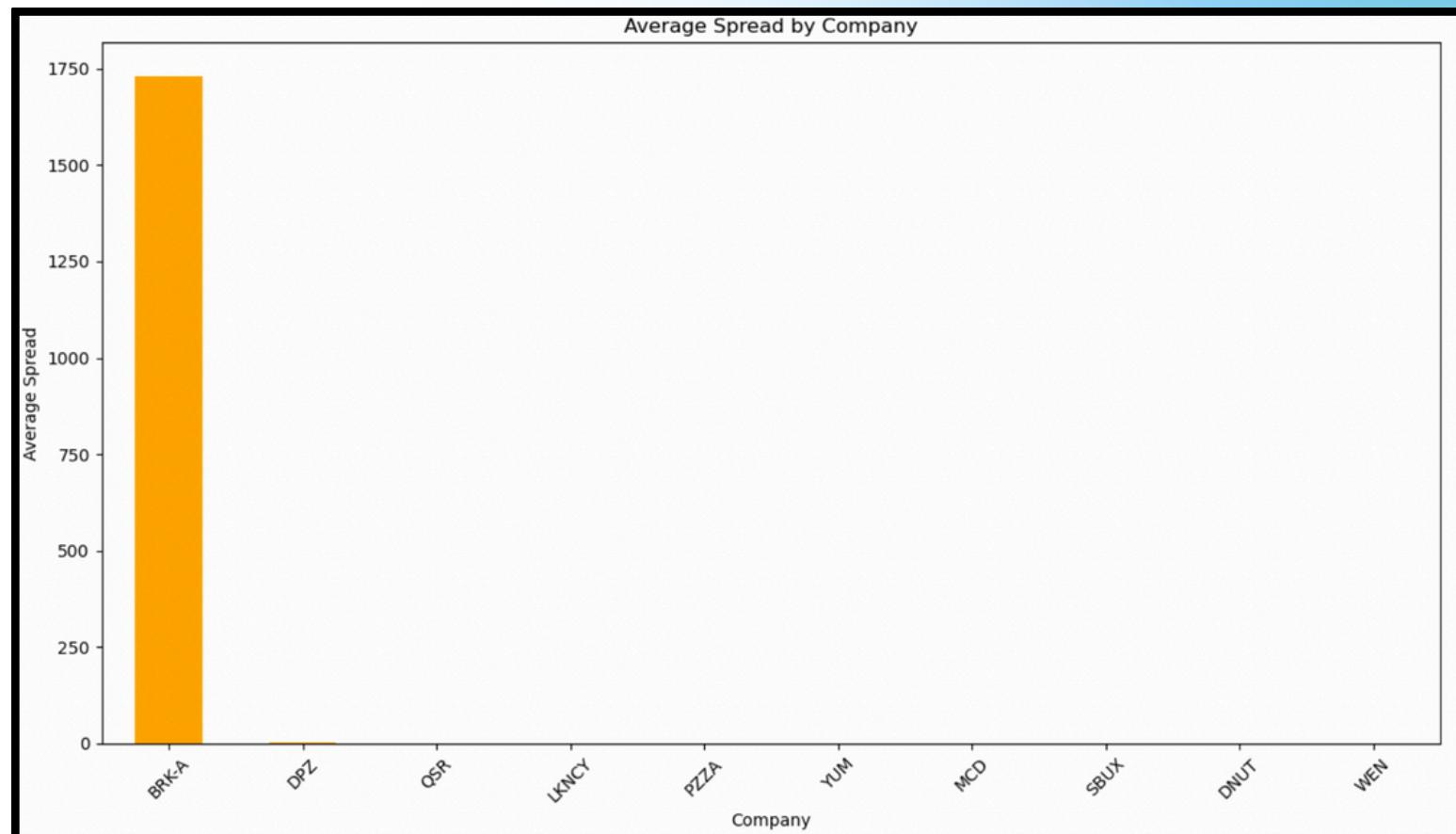
- Combined dataset for each year
- The exercise requires creating separate datasets for each year, so it is necessary to include a "Year" column. This column is crucial for correctly splitting the combined dataset into individual yearly datasets

		Date	Open	High	Low	Close	Adj Close	Volume	Year_Month	Year	Month	Company
Year												
1966	0	1966-07-05	0.00	0.27	0.27	0.27	0.12	388800.0	1966-07	1966	07	MCD
	1	1966-07-06	0.00	0.28	0.27	0.27	0.12	692550.0	1966-07	1966	07	MCD
	2	1966-07-07	0.00	0.29	0.27	0.27	0.12	1858950.0	1966-07	1966	07	MCD
	3	1966-07-08	0.00	0.28	0.27	0.28	0.12	1239300.0	1966-07	1966	07	MCD
	4	1966-07-11	0.00	0.28	0.27	0.28	0.12	656100.0	1966-07	1966	07	MCD
...												
2024	1807	2024-09-13	133.48	133.95	132.93	133.65	133.65	1811500.0	2024-09	2024	09	YUM
	1808	2024-09-16	134.43	135.48	133.79	134.56	134.56	1929300.0	2024-09	2024	09	YUM
	1809	2024-09-17	134.59	135.21	132.16	132.35	132.35	2009600.0	2024-09	2024	09	YUM
	1810	2024-09-18	132.40	133.10	130.27	130.60	131.72	2186200.0	2024-09	2024	09	YUM
	1811	2024-09-19	132.12	132.51	129.97	131.09	131.09	2616500.0	2024-09	2024	09	YUM



Question 10

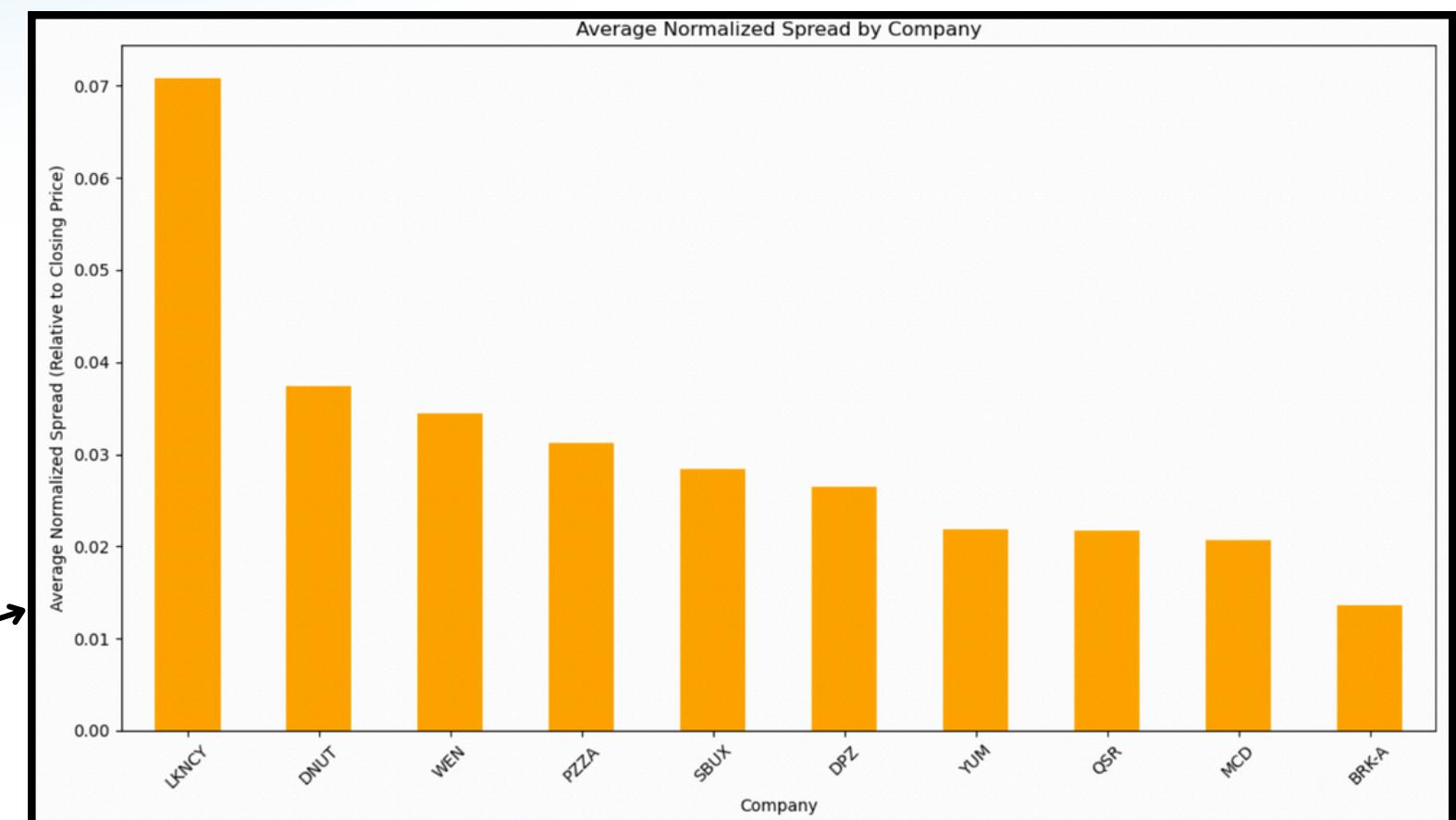
- Average spread (high - low) for each company



Average Spread by Company:

Company	Average Spread
BRK-A	1730.868709
DPZ	3.693502
QSR	1.202290
LKNCY	1.103296
PZZA	0.929258
YUM	0.888315
MCD	0.810959
SBUX	0.610409
DNUT	0.522675
WEN	0.338711

- Average normalized spread (spread / close) for each company



Thank you!