IMPORTING LIBRARIES

In [147…
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import scipy

warnings.filterwarnings("ignore")
''' Warnings are provided to warn the developer of situations that aren't necessari
 Usually, a warning occurs when there is some obsolete of certain programming eleme

%matplotlib inline
# Matplotlib Inline command is a magic command that makes the plots generated by ma
#the IPython shell that we are running and not in a separate output window.
```

IMPORTING AND READING THE DATASET

In [79]:
```
df=pd.read_csv("/content/AAPL.csv")
```

QUICK OVERVIEW OF THE DATASET

In [80]:
```
df.head(10) #top rows
```

Out[80]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	1980-12-12	0.128348	0.128906	0.128348	0.128348	0.100178	469033600
1	1980-12-15	0.122210	0.122210	0.121652	0.121652	0.094952	175884800
2	1980-12-16	0.113281	0.113281	0.112723	0.112723	0.087983	105728000
3	1980-12-17	0.115513	0.116071	0.115513	0.115513	0.090160	86441600
4	1980-12-18	0.118862	0.119420	0.118862	0.118862	0.092774	73449600
5	1980-12-19	0.126116	0.126674	0.126116	0.126116	0.098436	48630400
6	1980-12-22	0.132254	0.132813	0.132254	0.132254	0.103227	37363200
7	1980-12-23	0.137835	0.138393	0.137835	0.137835	0.107583	46950400
8	1980-12-24	0.145089	0.145647	0.145089	0.145089	0.113245	48003200
9	1980-12-26	0.158482	0.159040	0.158482	0.158482	0.123699	55574400

In [81]:
```
df.shape #dimension of the dataset
```

Out[81]:
```
(10468, 7)
```

CHECKING DATATYPES OF THE COLUMNS

In [82]:
```
df.info() #showing information about the dataset
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10468 entries, 0 to 10467
Data columns (total 7 columns):
 #   Column     Non-Null Count  Dtype  
---  ------     --------------  -----  
 0   Date       10468 non-null  object 
 1   Open       10468 non-null  float64
 2   High       10468 non-null  float64
 3   Low        10468 non-null  float64
 4   Close      10468 non-null  float64
 5   Adj Close  10468 non-null  float64
 6   Volume     10468 non-null  int64  
dtypes: float64(5), int64(1), object(1)
memory usage: 572.6+ KB
```

- PARTITIONING NUMERICAL AND CATEGOCIAL COLUMNS

In [83]:
```
numerical_columns=[column for column in df.columns if df[column].dtype != 'object'
categorical_columns=[column for column in df.columns if df[column].dtype == 'objec

#printing the columns

print('{} column(s) is/are numerical columns, and these columns are: {}'.format(le
print('{} column(s) is/are categorical columns, and these columns are: {}'.format(
```

```
6 column(s) is/are numerical columns, and these columns are: ['Open', 'High', 'Lo
w', 'Close', 'Adj Close', 'Volume']
1 column(s) is/are categorical columns, and these columns are: ['Date']
```

DATA CLEANING

- CHECKING FOR NULL VALUES

In [84]:
```
df.isnull().sum()
```

Out[84]:
```
Date         0
Open         0
High         0
Low          0
Close        0
Adj Close    0
Volume       0
dtype: int64
```

there are no null values

- CHANGING DATATYPES

In [85]:
```
# CHANGING THE DATATYPE OF DATE COLUMN FROM OBJECT TO DATE-TIME
df['Date']=pd.to_datetime(df['Date'])
```

In [86]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10468 entries, 0 to 10467
Data columns (total 7 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Date       10468 non-null  datetime64[ns]
 1   Open       10468 non-null  float64
 2   High       10468 non-null  float64
 3   Low        10468 non-null  float64
 4   Close      10468 non-null  float64
 5   Adj Close  10468 non-null  float64
 6   Volume     10468 non-null  int64
dtypes: datetime64[ns](1), float64(5), int64(1)
memory usage: 572.6 KB
```

In [87]:
```
df.head(10)
```

Out[87]:

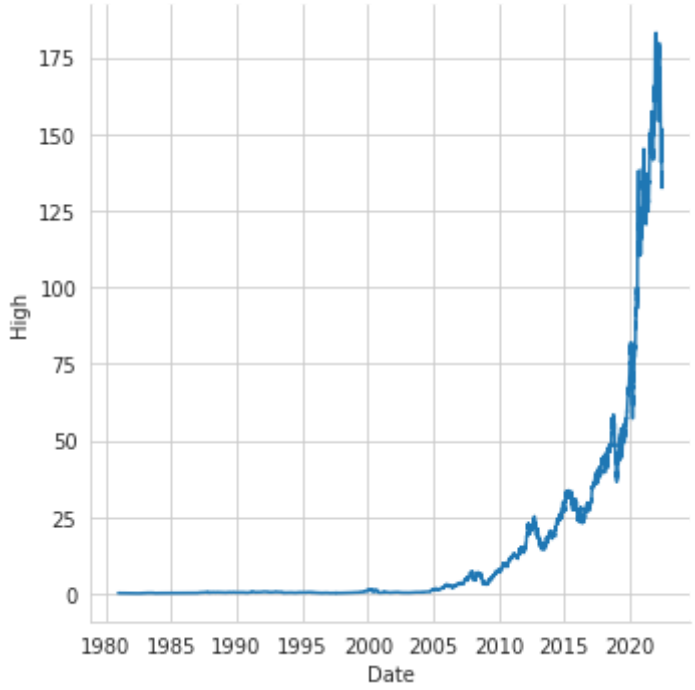
	Date	Open	High	Low	Close	Adj Close	Volume
0	1980-12-12	0.128348	0.128906	0.128348	0.128348	0.100178	469033600
1	1980-12-15	0.122210	0.122210	0.121652	0.121652	0.094952	175884800
2	1980-12-16	0.113281	0.113281	0.112723	0.112723	0.087983	105728000
3	1980-12-17	0.115513	0.116071	0.115513	0.115513	0.090160	86441600
4	1980-12-18	0.118862	0.119420	0.118862	0.118862	0.092774	73449600
5	1980-12-19	0.126116	0.126674	0.126116	0.126116	0.098436	48630400
6	1980-12-22	0.132254	0.132813	0.132254	0.132254	0.103227	37363200
7	1980-12-23	0.137835	0.138393	0.137835	0.137835	0.107583	46950400
8	1980-12-24	0.145089	0.145647	0.145089	0.145089	0.113245	48003200
9	1980-12-26	0.158482	0.159040	0.158482	0.158482	0.123699	55574400

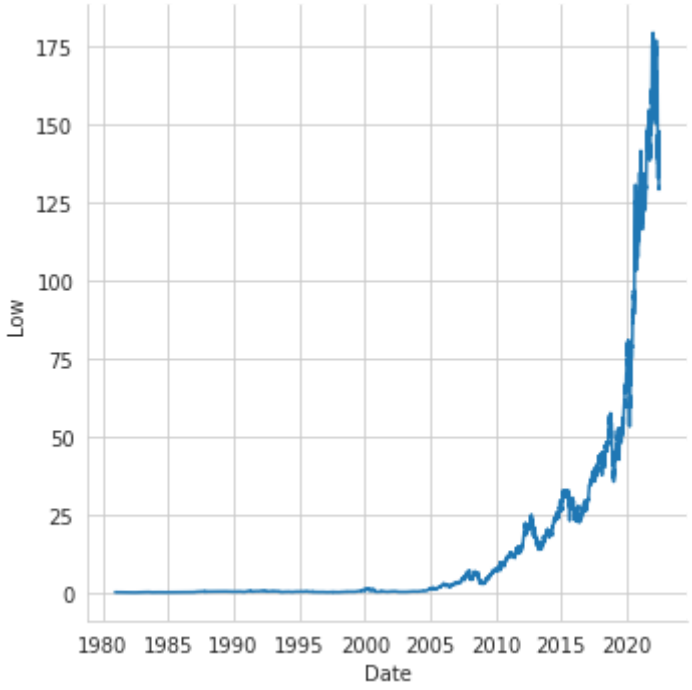
DATA VISUALIZATION

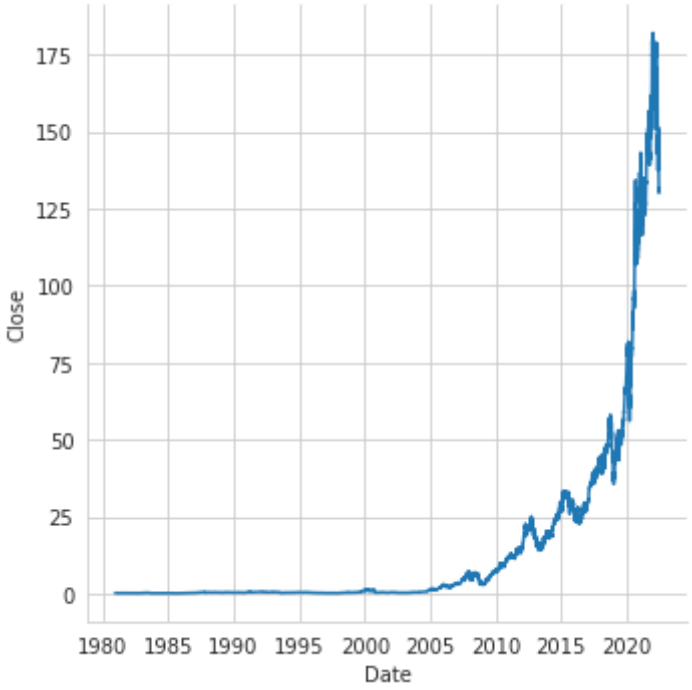
- HERE WE ARE VISUALIZING HOW EACH STOCK'S ATTRIBUTES ARE CHANGING OVER THE YEARS

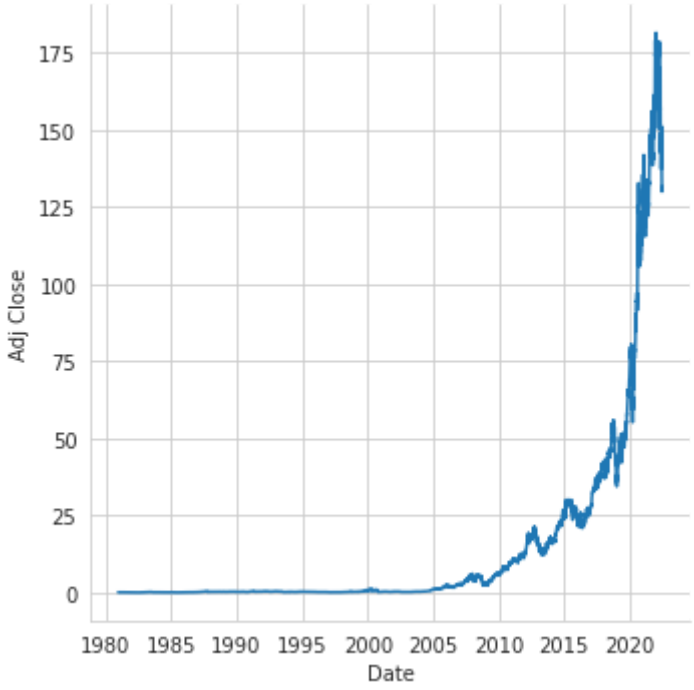
In [88]:
```
for i in df.columns:
    if i !='Date':
       sns.relplot(data=df, x="Date", y=i, kind="line")
```

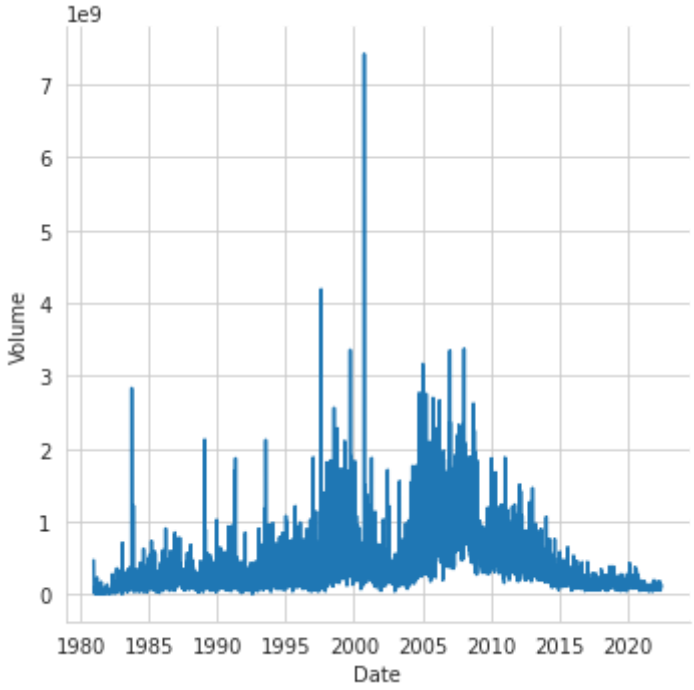

175
150
125
100
75
50
25
0
Open
1980
1985
1990
1995
2000
2005
2010
2015
2020
Date


175
150
125
100
75
50
25
0
High
1980
1985
1990
1995
2000
2005
2010
2015
2020
Date


175
150
125
100
75
50
25
0
Low
1980
1985
1990
1995
2000
2005
2010
2015
2020
Date


175
150
125
100
75
50
25
0
Close
1980
1985
1990
1995
2000
2005
2010
2015
2020
Date

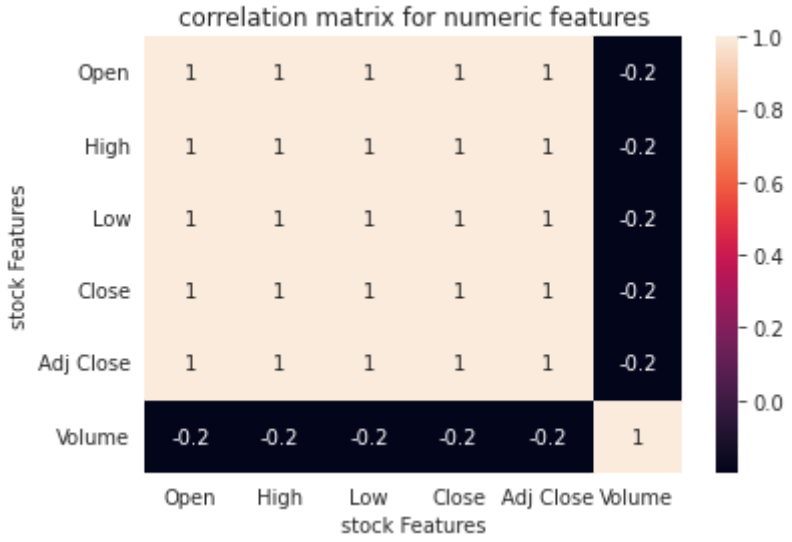




- CORRELATION MATRIX

In [89]:
```
correlation_matrix=df.corr(method='pearson')
sns.heatmap(correlation_matrix,annot=True)
plt.title("correlation matrix for numeric features")
plt.xlabel("stock Features")
plt.ylabel("stock Features")
```

Out[89]:
```
Text(33.0, 0.5, 'stock Features')
```



From the above heatmap, you can see a large number of 1s and values close to 1. This means those variables show high positive correlations and are interrelated. This might be possible because of the comparatively very small difference between those values. However, in-stock market this small value is what makes the difference. So, what we are going to do, is add new variables to our dataset that represents the difference between those values.

In [105…

```
df['open-high'] = df['Open']-df['High']
df['open-low'] = df['Open'] - df['Low']
df['close-high'] = df['Close']-df['High']
df['close-low'] = df['Close'] - df['Low']
df['high-low'] = df['High'] - df['Low']
df['open-close'] = df['Open'] - df['Close']
#we also creat a year column
df['Year'] = df['Date'].apply(lambda x: x.strftime('%Y'))
df.head()
```

Out[105]:

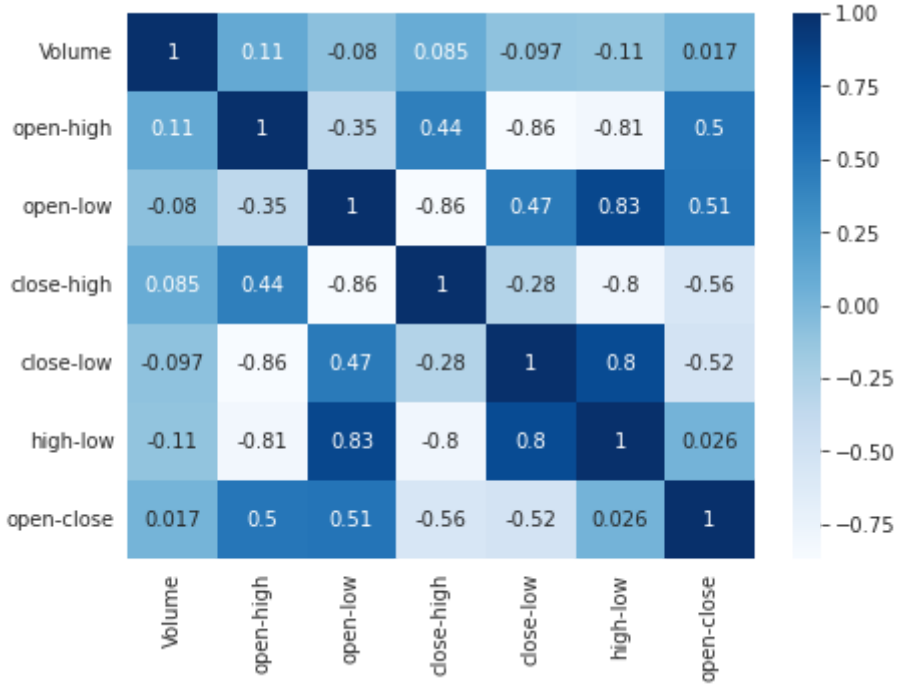
	Date	Open	High	Low	Close	Adj Close	Volume	open-high	open-low	clo h
0	1980-12-12	0.128348	0.128906	0.128348	0.128348	0.100178	469033600	-0.000558	0.000000	-0.000
1	1980-12-15	0.122210	0.122210	0.121652	0.121652	0.094952	175884800	0.000000	0.000558	-0.000
2	1980-12-16	0.113281	0.113281	0.112723	0.112723	0.087983	105728000	0.000000	0.000558	-0.000
3	1980-12-17	0.115513	0.116071	0.115513	0.115513	0.090160	86441600	-0.000558	0.000000	-0.000
4	1980-12-18	0.118862	0.119420	0.118862	0.118862	0.092774	73449600	-0.000558	0.000000	-0.000

We have added 6 new entries to our dataset. We can now create a heatmap with these new values and see their inter-correlation.

From the previous heatmap, we could realize that those values aren't very useful. So for more clarity, we will drop those values and take only the new values for creating the heatmap.

In [91]:
```
df2 = df.copy()
df2 = df2.drop(['Open','High','Low','Close', 'Adj Close'],axis=1)
plt.figure(figsize=(7,5))
sns.heatmap(df2.corr(),cmap='Blues',annot=True)
```

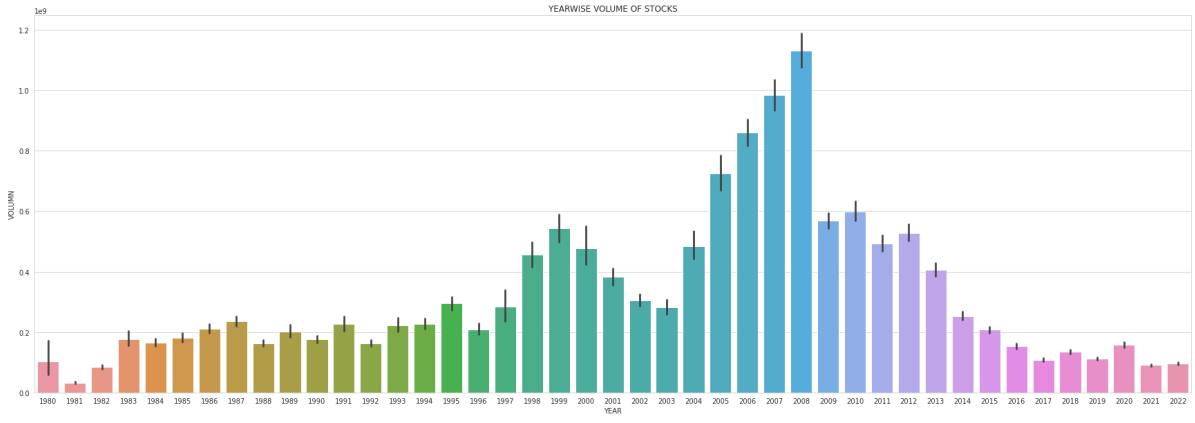
Out[91]:
```
<AxesSubplot:>
```



Now this looks more informative.

In [92]:
```
fig, ax = plt.subplots(figsize=(30, 10))
sns.barplot(data=df, x="Year", y="Volume",ax=ax)
plt.xlabel('YEAR')
plt.ylabel('VOLUMN')
plt.title('YEARWISE VOLUME OF STOCKS')
```

Out[92]:
```
Text(0.5, 1.0, 'YEARWISE VOLUME OF STOCKS')
```



INSIGHTS USING GROUP BY

- HIGHEST VALUES

In [130…
```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].max().reset_i
```

Out[130]:

	Year	Volume	Open	High	Low	Close	Adj Close
0	1980	469033600	0.160714	0.161272	0.160714	0.160714	0.125441
1	1981	231302400	0.154018	0.155134	0.154018	0.154018	0.120214
2	1982	357078400	0.151228	0.155692	0.147321	0.151228	0.118037
3	1983	2832345600	0.280134	0.282366	0.273996	0.280134	0.218651
4	1984	626796800	0.148438	0.153460	0.145089	0.148438	0.115859
5	1985	733980800	0.136719	0.138951	0.136719	0.136719	0.106712
6	1986	904131200	0.194754	0.195871	0.190290	0.195313	0.152446
7	1987	777280000	0.531250	0.533482	0.515625	0.529018	0.414073
8	1988	680960000	0.420759	0.426339	0.415179	0.421875	0.332239
9	1989	2127171200	0.444196	0.449777	0.433036	0.443080	0.352256
10	1990	933732800	0.424107	0.426339	0.417411	0.422991	0.339942
11	1991	1868372800	0.647321	0.654018	0.625000	0.649554	0.526456
12	1992	848355200	0.625000	0.625000	0.609375	0.623884	0.510264
13	1993	2120596800	0.571429	0.582589	0.569196	0.580357	0.477911
14	1994	980134400	0.386161	0.390625	0.380580	0.386161	0.325469
15	1995	1211963200	0.437500	0.447545	0.434152	0.440848	0.374739
16	1996	1026625600	0.310268	0.316964	0.303571	0.312500	0.267193
17	1997	4190480000	0.256696	0.263951	0.253348	0.260603	0.222820
18	1998	2561350400	0.388393	0.390625	0.377232	0.383929	0.328266
19	1999	3357558400	1.040737	1.053571	1.017857	1.051897	0.899391
20	2000	7421640800	1.271763	1.342634	1.250000	1.287388	1.100740
21	2001	1873670400	0.476786	0.484286	0.460000	0.474821	0.405980
22	2002	1709467200	0.463036	0.467321	0.453214	0.466250	0.398652
23	2003	1555708000	0.443750	0.446607	0.438929	0.443214	0.378956
24	2004	2768427200	1.231250	1.242321	1.203750	1.222143	1.044954
25	2005	3164716800	2.673929	2.695000	2.662857	2.677857	2.289616
26	2006	2663768800	3.321429	3.327143	3.252143	3.278929	2.803543
27	2007	3349298400	7.163929	7.248571	7.064286	7.136786	6.102081
28	2008	3372969600	7.116786	7.152143	6.881786	6.961786	5.952452
29	2009	1831634000	7.611786	7.641071	7.520000	7.558571	6.462714
30	2010	1867110000	11.650714	11.666429	11.610714	11.623929	9.938669
31	2011	1880998000	15.062857	15.239286	14.856786	15.080000	12.893674
32	2012	1506120000	25.086071	25.181070	24.984644	25.075001	21.531624
33	2013	1460852400	20.451786	20.540714	20.228930	20.360357	18.004278
34	2014	1065523200	29.817499	29.937500	29.512501	29.750000	26.860983
35	2015	648825200	33.615002	33.634998	32.849998	33.250000	30.160797

	Year	Volume	Open	High	Low	Close	Adj Close
36	2016	533478800	29.545000	29.672501	29.362499	29.562500	27.608709
37	2017	447940000	43.777500	44.299999	43.715000	44.105000	42.067352
38	2018	384986800	57.695000	58.367500	57.445000	58.017502	55.972595
39	2019	365248800	72.779999	73.492500	72.379997	73.412498	72.139954
40	2020	426510000	138.050003	138.789993	134.339996	136.690002	135.480576
41	2021	195432700	181.119995	182.130005	178.529999	180.330002	179.836319
42	2022	182602000	182.630005	182.940002	179.119995	182.009995	181.511703

- LOWEST VALUES

In [129…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].min().reset_i
```

Out[129]:

	Year	Volume	Open	High	Low	Close	Adj Close
0	1980	35750400	0.113281	0.113281	0.112723	0.112723	0.087983
1	1981	0	0.064174	0.064732	0.063616	0.063616	0.049654
2	1982	13596800	0.049665	0.049665	0.049107	0.049107	0.038329
3	1983	1388800	0.079799	0.085938	0.077009	0.079799	0.062285
4	1984	27686400	0.101004	0.101563	0.097656	0.097656	0.076223
5	1985	21504000	0.065290	0.065848	0.064732	0.064732	0.050525
6	1986	50176000	0.098214	0.099330	0.097098	0.098772	0.077094
7	1987	56638400	0.180246	0.183594	0.179129	0.182478	0.142428
8	1988	29937600	0.319196	0.328125	0.316964	0.322545	0.255164
9	1989	27854400	0.301339	0.308036	0.290179	0.301339	0.239045
10	1990	44273600	0.225446	0.236607	0.216518	0.223214	0.179931
11	1991	17987200	0.366071	0.375000	0.359375	0.367188	0.298363
12	1992	2396800	0.375000	0.390625	0.370536	0.386161	0.317320
13	1993	39592000	0.202009	0.205357	0.196429	0.202009	0.167797
14	1994	62272000	0.224330	0.233259	0.219866	0.224330	0.188420
15	1995	39390400	0.285714	0.289063	0.282366	0.284598	0.243336
16	1996	42291200	0.155134	0.155134	0.142857	0.150670	0.128826
17	1997	41316800	0.114955	0.118304	0.113839	0.115513	0.098766
18	1998	111115200	0.121652	0.145089	0.120536	0.141741	0.121191
19	1999	123681600	0.287946	0.299107	0.285714	0.287388	0.245722
20	2000	70828800	0.246094	0.261161	0.243304	0.250000	0.213754
21	2001	50629600	0.258929	0.272321	0.257813	0.265625	0.227114
22	2002	39340000	0.241786	0.247321	0.238571	0.242679	0.207495
23	2003	76098400	0.231964	0.235536	0.227143	0.234286	0.200319
24	2004	141635200	0.382500	0.388393	0.378214	0.380000	0.324907
25	2005	176383200	1.139107	1.159107	1.117857	1.130179	0.966323
26	2006	194770800	1.847500	1.888929	1.791429	1.809643	1.547278
27	2007	352466800	2.964286	3.016786	2.925000	2.973929	2.542763
28	2008	271334000	2.926071	3.004286	2.826429	2.874643	2.457871
29	2009	286454000	2.835357	2.928571	2.792857	2.792857	2.387943
30	2010	157494400	6.870357	7.000000	6.794643	6.858929	5.864507
31	2011	179662000	11.310000	11.346429	11.089286	11.261429	9.628724
32	2012	175753200	14.621429	14.732143	14.607143	14.686786	12.557464
33	2013	167554800	13.856071	14.271429	13.753571	13.947500	12.101386
34	2014	57918400	17.684999	17.911785	17.626785	17.849285	15.783781
35	2015	52185600	23.717501	26.757500	23.000000	25.780001	23.572208

	Year	Volume	Open	High	Low	Close	Adj Close
36	2016	45903600	22.500000	22.917500	22.367500	22.584999	20.978708
37	2017	56106800	28.950001	29.082500	28.690001	29.004999	27.227137
38	2018	50055600	37.037498	37.887501	36.647499	36.707500	35.537251
39	2019	45448000	35.994999	36.430000	35.500000	35.547501	34.414238
40	2020	46691300	57.020000	57.125000	53.152500	56.092499	55.250996
41	2021	41000000	119.029999	120.400002	116.209999	116.360001	115.502800
42	2022	53950200	130.070007	132.389999	129.039993	130.059998	130.059998

- MEAN VALUES

In [128…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].mean().reset_
```

Out[128]:

	Year	Volume	Open	High	Low	Close	Adj Close
0	1980	1.034501e+08	0.136075	0.136461	0.135903	0.135903	0.106075
1	1981	3.239899e+07	0.108883	0.109265	0.108548	0.108548	0.084724
2	1982	8.444667e+07	0.085446	0.086665	0.084630	0.085459	0.066703
3	1983	1.759408e+08	0.167274	0.171085	0.163449	0.167274	0.130561
4	1984	1.659250e+08	0.119956	0.122294	0.117642	0.119651	0.093390
5	1985	1.805249e+08	0.091057	0.092023	0.089937	0.090233	0.070429
6	1986	2.107637e+08	0.144586	0.147048	0.142207	0.144913	0.113108
7	1987	2.362502e+08	0.347817	0.355806	0.339820	0.347751	0.271952
8	1988	1.632134e+08	0.371034	0.376663	0.365092	0.370884	0.292086
9	1989	2.020072e+08	0.371563	0.377679	0.365881	0.371953	0.295725
10	1990	1.755017e+08	0.334839	0.341244	0.328725	0.335373	0.269527
11	1991	2.266706e+08	0.468318	0.477012	0.459883	0.468701	0.380760
12	1992	1.619603e+08	0.489318	0.496445	0.481831	0.489309	0.400988
13	1993	2.231341e+08	0.366636	0.373019	0.359684	0.366310	0.302930
14	1994	2.268091e+08	0.304040	0.309923	0.298332	0.304291	0.255578
15	1995	2.947085e+08	0.362706	0.368455	0.356325	0.361966	0.307642
16	1996	2.094261e+08	0.223641	0.226978	0.218790	0.222495	0.190238
17	1997	2.844402e+08	0.161005	0.163930	0.157402	0.160418	0.137160
18	1998	4.571198e+08	0.272430	0.279160	0.265864	0.272900	0.233335
19	1999	5.440584e+08	0.514817	0.527677	0.502688	0.515805	0.441023
20	2000	4.773873e+08	0.817692	0.843950	0.792435	0.816283	0.697937
21	2001	3.816847e+08	0.360094	0.370828	0.350399	0.361061	0.308714
22	2002	3.056108e+08	0.341572	0.348608	0.334219	0.341776	0.292224
23	2003	2.826597e+08	0.330746	0.336491	0.325125	0.331156	0.283145
24	2004	4.833400e+08	0.632526	0.643383	0.623654	0.634410	0.542432
25	2005	7.238134e+08	1.665373	1.689345	1.641536	1.667001	1.425316
26	2006	8.593584e+08	2.535272	2.569254	2.493227	2.528951	2.162299
27	2007	9.840478e+08	4.585324	4.645374	4.506589	4.581211	3.917017
28	2008	1.130360e+09	5.082634	5.182524	4.959204	5.070679	4.335521
29	2009	5.684670e+08	5.236396	5.303417	5.177317	5.243362	4.483168
30	2010	5.993053e+08	9.284201	9.370315	9.173129	9.280088	7.934642
31	2011	4.922990e+08	13.002194	13.122270	12.867775	13.000154	11.115365
32	2012	5.278568e+08	20.594740	20.779479	20.354327	20.573204	17.633141
33	2013	4.064348e+08	16.897433	17.058533	16.723111	16.879817	14.758661
34	2014	2.526109e+08	23.054958	23.253166	22.868885	23.066164	20.655213
35	2015	2.073976e+08	30.043938	30.311131	29.715774	30.010000	27.332785

	Year	Volume	Open	High	Low	Close	Adj Close
36	2016	1.536901e+08	26.126944	26.356796	25.922569	26.151002	24.311881
37	2017	1.085383e+08	37.611225	37.851514	37.371912	37.637769	35.650773
38	2018	1.360803e+08	47.277859	47.748526	46.795877	47.263357	45.450381
39	2019	1.121228e+08	51.967272	52.457927	51.567907	52.063988	50.857012
40	2020	1.575646e+08	95.267668	96.661769	93.912796	95.347075	94.202858
41	2021	9.052463e+07	140.861866	142.347818	139.474087	140.989365	140.227270
42	2022	9.599775e+07	161.348620	163.562328	159.031982	161.285949	161.060847

- MEDIAN VALUES

In [127…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].median().rese
```

Out[127]:

	Year	Volume	Open	High	Low	Close	Adj Close
0	1980	6.888000e+07	0.132254	0.132813	0.132254	0.132254	0.103227
1	1981	2.632000e+07	0.111049	0.111607	0.110491	0.110491	0.086241
2	1982	6.608000e+07	0.079799	0.079799	0.079241	0.079241	0.061849
3	1983	1.339968e+08	0.172991	0.177455	0.165179	0.170201	0.132846
4	1984	1.398432e+08	0.118862	0.121094	0.116629	0.118304	0.092339
5	1985	1.455328e+08	0.088170	0.089286	0.086217	0.087612	0.068383
6	1986	1.807904e+08	0.149554	0.153460	0.147879	0.150670	0.117601
7	1987	2.038400e+08	0.343750	0.351563	0.334821	0.343750	0.268713
8	1988	1.456672e+08	0.368304	0.373884	0.361607	0.368304	0.289597
9	1989	1.599920e+08	0.376116	0.382254	0.368304	0.376674	0.298110
10	1990	1.534624e+08	0.339286	0.343750	0.334821	0.340402	0.274465
11	1991	1.908704e+08	0.455357	0.462054	0.446429	0.453125	0.369911
12	1992	1.409912e+08	0.504464	0.513393	0.498326	0.505580	0.414533
13	1993	1.770048e+08	0.345982	0.348214	0.337054	0.339286	0.280589
14	1994	1.879360e+08	0.304129	0.310268	0.300223	0.304688	0.256095
15	1995	2.523528e+08	0.358817	0.364397	0.353795	0.358259	0.304527
16	1996	1.604288e+08	0.221540	0.223772	0.216518	0.220424	0.188466
17	1997	1.854048e+08	0.155134	0.157366	0.150670	0.154018	0.131688
18	1998	3.724000e+08	0.272879	0.275391	0.267299	0.272042	0.232600
19	1999	4.332552e+08	0.430246	0.441964	0.423828	0.439732	0.375978
20	2000	3.859828e+08	0.928292	0.953125	0.896484	0.922154	0.788458
21	2001	3.261412e+08	0.355603	0.367746	0.348303	0.356965	0.305211
22	2002	2.698024e+08	0.309196	0.317411	0.301250	0.308839	0.264064
23	2003	2.377872e+08	0.345625	0.351429	0.339286	0.344286	0.294370
24	2004	3.855936e+08	0.547947	0.558482	0.542053	0.550357	0.470565
25	2005	5.945674e+08	1.534822	1.558035	1.517857	1.531964	1.309857
26	2006	7.842772e+08	2.477857	2.530000	2.429286	2.472143	2.113727
27	2007	8.998108e+08	4.427143	4.470714	4.339286	4.425000	3.783455
28	2008	1.041174e+09	5.287857	5.389643	5.208929	5.299286	4.530986
29	2009	5.186734e+08	5.076964	5.131429	4.996964	5.085178	4.347919
30	2010	5.502196e+08	9.176250	9.277679	9.085714	9.180893	7.849827
31	2011	4.401838e+08	12.761429	12.856607	12.647143	12.750893	10.902240
32	2012	4.875598e+08	20.759642	21.011071	20.517857	20.803215	17.804374
33	2013	3.591070e+08	16.466072	16.633572	16.274822	16.466964	14.304721
34	2014	2.292350e+08	23.488751	23.687500	23.291249	23.475001	20.999386
35	2015	1.893452e+08	30.198751	30.396249	29.837499	30.075000	27.469740

	Year	Volume	Open	High	Low	Close	Adj Close
36	2016	1.349624e+08	26.373750	26.570000	26.206249	26.437500	24.453612
37	2017	9.763800e+07	38.112499	38.465000	37.782501	38.185001	36.101116
38	2018	1.266524e+08	46.572498	46.849998	46.235001	46.512501	44.778759
39	2019	1.035100e+08	50.785000	51.178751	50.398750	50.753749	49.570383
40	2020	1.380234e+08	91.279999	92.970001	90.910004	91.632500	90.502144
41	2021	8.563055e+07	140.655006	142.225006	138.739998	141.305001	140.660843
42	2022	9.107065e+07	164.154998	166.340004	162.035004	164.150002	163.899521

- MODE VALUES

In [140…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].agg([statisti
```

Out[140]:

	Year	Volume mode	Open mode	High mode	Low mode	Close mode	Adj Close mode
0	1980	469033600	0.128348	0.128906	0.128348	0.128348	0.100178
1	1981	35728000	0.114955	0.084821	0.084263	0.084263	0.065769
2	1982	33331200	0.082031	0.082031	0.083705	0.083705	0.065334
3	1983	189772800	0.095982	0.111607	0.188616	0.090960	0.070996
4	1984	168492800	0.114955	0.114955	0.109375	0.110491	0.086241
5	1985	118294400	0.068080	0.068080	0.089286	0.068080	0.053138
6	1986	246400000	0.160714	0.161830	0.106027	0.160714	0.125441
7	1987	239680000	0.361607	0.354911	0.345982	0.350446	0.363406
8	1988	143360000	0.372768	0.352679	0.343750	0.354911	0.292986
9	1989	135374400	0.399554	0.401786	0.397321	0.399554	0.318406
10	1990	182470400	0.299107	0.310268	0.366071	0.303571	0.283435
11	1991	188563200	0.457589	0.462054	0.433036	0.450893	0.361752
12	1992	144558400	0.511161	0.569196	0.562500	0.399554	0.327445
13	1993	177004800	0.491071	0.488839	0.491071	0.254464	0.405268
14	1994	298793600	0.299107	0.303571	0.267857	0.299107	0.272648
15	1995	332875200	0.341518	0.390625	0.339286	0.341518	0.289502
16	1996	139294400	0.222098	0.223214	0.214286	0.220982	0.188944
17	1997	222432000	0.151786	0.152902	0.149554	0.148438	0.126917
18	1998	718110400	0.316964	0.243304	0.238839	0.313616	0.268147
19	1999	952884800	0.413504	0.418527	0.299107	0.415179	0.354986
20	2000	535796800	0.901786	0.267857	0.250000	0.330357	0.282461
21	2001	324889600	0.340402	0.410714	0.366071	0.348214	0.297729
22	2002	529496800	0.283929	0.271250	0.268036	0.262500	0.224442
23	2003	228256000	0.256429	0.383929	0.256250	0.258036	0.220625
24	2004	144642400	0.407857	0.990179	0.412500	0.410714	0.351168
25	2005	691992000	1.447143	1.330357	1.517857	1.526786	1.305430
26	2006	807234400	2.687500	2.603214	2.381429	2.650000	2.265797
27	2007	1238319600	3.111071	3.078571	3.121429	3.181071	2.719872
28	2008	1079178800	6.117857	6.428571	5.214286	4.643214	3.970031
29	2009	746015200	3.230714	3.335000	3.131071	3.155714	2.698192
30	2010	493729600	8.931429	7.571429	9.625000	7.643214	6.535084
31	2011	445138400	11.957143	12.857143	11.853571	12.265714	10.487404
32	2012	302220800	20.053572	20.500000	19.241785	14.686786	12.557464
33	2013	560518000	15.858929	16.314285	18.507143	16.145714	14.008640
34	2014	200760000	19.028214	19.375000	18.892857	21.154642	18.818619

	Year	Volume mode	Open mode	High mode	Low mode	Close mode	Adj Close mode
35	2015	212818400	29.010000	28.200001	27.200001	31.650000	29.496325
36	2016	270597600	24.000000	25.032499	23.147499	24.334999	24.459383
37	2017	115127600	30.000000	40.000000	30.400000	30.000000	28.161146
38	2018	102223600	42.540001	44.847500	43.020000	43.750000	41.372166
39	2019	146118800	42.427502	39.532501	50.875000	49.250000	38.221371
40	2020	135480400	113.910004	80.860001	110.000000	79.807503	89.904625
41	2021	143301900	127.820000	144.300003	129.470001	132.029999	126.129395
42	2022	104487900	177.830002	174.139999	177.710007	172.190002	171.718597

- STANDARD DEVIATION

In [141…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].agg([np.std])
```

Out[141]:

	Year	Volume	Open	High	Low	Close	Adj Close
		std	std	std	std	std	std
0	1980	1.161565e+08	0.017162	0.017160	0.017160	0.017160	0.013394
1	1981	2.656779e+07	0.021831	0.021857	0.021817	0.021817	0.017029
2	1982	6.393526e+07	0.026687	0.028016	0.025740	0.027002	0.021076
3	1983	2.063877e+08	0.054684	0.055022	0.053978	0.054631	0.042640
4	1984	9.712616e+07	0.009464	0.009805	0.009043	0.009696	0.007568
5	1985	1.225049e+08	0.020420	0.020658	0.020007	0.020196	0.015764
6	1986	1.236627e+08	0.024725	0.024980	0.024279	0.024717	0.019292
7	1987	1.298807e+08	0.076883	0.077332	0.075344	0.075806	0.059474
8	1988	8.836423e+07	0.022420	0.022350	0.022385	0.022295	0.017333
9	1989	1.794046e+08	0.036650	0.037596	0.036306	0.037186	0.029893
10	1990	9.959380e+07	0.040937	0.040822	0.041371	0.041410	0.033266
11	1991	1.879740e+08	0.062809	0.064787	0.061204	0.063419	0.051093
12	1992	8.538428e+07	0.067111	0.067426	0.066527	0.066601	0.053995
13	1993	1.837278e+08	0.117782	0.119079	0.117086	0.118274	0.096873
14	1994	1.402649e+08	0.039210	0.039464	0.039208	0.039187	0.033441
15	1995	1.793976e+08	0.035617	0.036628	0.035782	0.036480	0.030795
16	1996	1.567978e+08	0.028847	0.029617	0.028233	0.029202	0.024968
17	1997	4.237077e+08	0.025441	0.026371	0.024620	0.025548	0.021844
18	1998	3.282436e+08	0.060531	0.061435	0.058428	0.059337	0.050734
19	1999	3.836842e+08	0.193520	0.198935	0.188320	0.194463	0.166270
20	2000	5.302301e+08	0.298238	0.308891	0.291374	0.301159	0.257496
21	2001	2.124995e+08	0.044791	0.045331	0.043080	0.044325	0.037899
22	2002	1.591937e+08	0.076609	0.077988	0.075050	0.076939	0.065784
23	2003	1.900182e+08	0.059021	0.059864	0.057877	0.059084	0.050518
24	2004	3.525949e+08	0.236537	0.238967	0.231699	0.235365	0.201242
25	2005	4.624646e+08	0.406458	0.409911	0.403735	0.408118	0.348948
26	2006	3.590202e+08	0.338334	0.339767	0.337781	0.340239	0.290910
27	2007	4.221247e+08	1.263317	1.282736	1.226940	1.258777	1.076277
28	2008	4.549178e+08	1.206796	1.201830	1.205704	1.204722	1.030059
29	2009	2.135105e+08	1.430512	1.425560	1.422396	1.421818	1.215680
30	2010	2.602546e+08	1.339099	1.337319	1.349294	1.341302	1.146837
31	2011	2.186541e+08	0.927495	0.941109	0.912776	0.925852	0.791620
32	2012	2.318070e+08	2.411801	2.417580	2.378964	2.392032	2.059878
33	2013	1.770416e+08	1.607185	1.591152	1.609774	1.603135	1.475523
34	2014	1.176648e+08	3.345616	3.389360	3.298561	3.342823	3.114685

	Year	Volume	Open	High	Low	Close	Adj Close
		std	std	std	std	std	std
35	2015	8.467721e+07	1.922002	1.866414	1.979353	1.920888	1.695854
36	2016	6.837414e+07	1.896533	1.888594	1.920181	1.910186	1.885187
37	2017	4.515806e+07	3.686197	3.705652	3.619850	3.655298	3.617986
38	2018	5.850376e+07	5.126908	5.146255	5.109996	5.148465	5.054042
39	2019	4.329647e+07	8.593396	8.621376	8.576434	8.634744	8.645282
40	2020	6.983072e+07	22.011907	22.094637	21.579913	21.809814	21.759932
41	2021	2.919806e+07	14.493334	14.755514	14.446238	14.659580	14.800573
42	2022	2.320847e+07	12.521892	12.266973	12.660937	12.549698	12.422967

- SKEWNESS

In [144…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].skew().reset_
```

Out[144]:

	Year	Volume	Open	High	Low	Close	Adj Close
0	1980	3.012705	0.225182	0.213476	0.213491	0.213491	0.213514
1	1981	3.470329	0.071604	0.073677	0.075309	0.075309	0.075310
2	1982	1.801751	1.004342	1.021864	0.968825	0.992652	0.992653
3	1983	9.256316	0.068791	0.048793	0.074156	0.070257	0.070257
4	1984	1.539477	0.686170	0.649342	0.542215	0.623630	0.623615
5	1985	1.506238	0.858277	0.836260	0.878043	0.854131	0.854128
6	1986	2.151787	-0.248122	-0.276063	-0.226547	-0.246106	-0.246108
7	1987	1.251298	0.308554	0.300508	0.326305	0.330933	0.330789
8	1988	1.919526	0.218870	0.181898	0.300118	0.275459	0.300971
9	1989	5.964389	-0.215559	-0.183560	-0.213107	-0.196201	-0.181898
10	1990	2.989685	-0.358820	-0.341846	-0.429440	-0.370302	-0.349700
11	1991	5.099360	0.889793	0.906236	0.858218	0.900131	0.878047
12	1992	2.829742	-0.082389	-0.110463	-0.096567	-0.105444	-0.122354
13	1993	5.741373	0.150253	0.154080	0.162437	0.161639	0.157641
14	1994	2.209241	-0.015973	-0.001040	-0.030629	-0.027688	0.007140
15	1995	2.019704	0.029395	0.052592	0.025819	0.032115	0.055771
16	1996	2.575446	0.736250	0.784663	0.578879	0.704268	0.704272
17	1997	6.082422	0.835049	0.913353	0.811554	0.864538	0.864536
18	1998	2.796565	-0.363776	-0.326733	-0.394114	-0.349525	-0.349525
19	1999	2.878222	0.917528	0.904799	0.891193	0.898427	0.898427
20	2000	9.494427	-0.823891	-0.818271	-0.810722	-0.797099	-0.797099
21	2001	2.441088	0.175734	0.208182	0.164211	0.200103	0.200103
22	2002	4.060306	0.178904	0.167008	0.203166	0.190904	0.190905
23	2003	3.952730	-0.153339	-0.176168	-0.148985	-0.157276	-0.157276
24	2004	3.008060	1.243956	1.227797	1.226280	1.210603	1.210603
25	2005	2.334423	1.123856	1.105520	1.121527	1.104321	1.104321
26	2006	1.839736	0.366458	0.356010	0.364183	0.341012	0.341013
27	2007	1.747202	0.434355	0.423553	0.452427	0.435470	0.435470
28	2008	1.220549	-0.269989	-0.287244	-0.266207	-0.275571	-0.275571
29	2009	1.702020	-0.032509	-0.023356	-0.040642	-0.040501	-0.040501
30	2010	1.568410	0.112561	0.087765	0.145475	0.114885	0.114885
31	2011	2.034517	0.271619	0.273624	0.276960	0.274307	0.274307
32	2012	1.209950	-0.569664	-0.620660	-0.499425	-0.558474	-0.552136
33	2013	1.970707	0.435271	0.448657	0.438960	0.436560	0.484059
34	2014	2.550451	0.131316	0.132476	0.120213	0.122300	0.129885
35	2015	1.811139	-0.245077	-0.166116	-0.301634	-0.220034	-0.267945

	Year	Volume	Open	High	Low	Close	Adj Close
36	2016	2.058890	0.066456	0.042413	0.056307	0.036501	0.078373
37	2017	2.943399	-0.362762	-0.359753	-0.333390	-0.333044	-0.306951
38	2018	1.510052	0.439883	0.499236	0.375300	0.430733	0.445746
39	2019	1.965866	0.542817	0.545804	0.550608	0.566809	0.565928
40	2020	1.486194	0.106241	0.118275	0.078155	0.092926	0.096675
41	2021	1.017808	0.768861	0.802776	0.720637	0.747964	0.749668
42	2022	1.264083	-0.721966	-0.787363	-0.652626	-0.702777	-0.702603

- KURTOSIS

In [148…

```
df.groupby('Year')['Volume','Open','High','Low','Close','Adj Close'].agg([scipy.st
```

Out[148]:

	Year	Volume kurtosis	Open kurtosis	High kurtosis	Low kurtosis	Close kurtosis	Adj Close kurtosis
0	1980	5.909853	-1.438974	-1.428083	-1.428085	-1.428085	-1.428092
1	1981	17.335070	-1.050793	-1.050422	-1.046641	-1.046641	-1.046639
2	1982	3.366209	-0.131853	-0.152297	-0.176050	-0.179023	-0.179020
3	1983	109.527623	-1.140870	-1.146092	-1.138206	-1.137760	-1.137760
4	1984	3.341868	0.337347	0.239215	0.157337	0.288825	0.288794
5	1985	2.505340	-0.272144	-0.304339	-0.205147	-0.255538	-0.255537
6	1986	6.798238	-0.801260	-0.793956	-0.793170	-0.806044	-0.806048
7	1987	1.979777	-0.267096	-0.232962	-0.304516	-0.257190	-0.258767
8	1988	6.035966	-0.685288	-0.703633	-0.736866	-0.717382	-0.729771
9	1989	53.711517	-0.953342	-0.952304	-1.001050	-0.973718	-1.006655
10	1990	15.436411	-0.472100	-0.556264	-0.354515	-0.490923	-0.523231
11	1991	37.005928	0.306219	0.268804	0.233148	0.326887	0.324822
12	1992	16.417727	-1.349952	-1.360192	-1.370161	-1.371087	-1.373269
13	1993	47.950062	-1.622429	-1.612851	-1.622183	-1.611245	-1.610639
14	1994	6.296206	-0.597403	-0.635526	-0.692912	-0.650118	-0.687272
15	1995	5.132363	-0.831756	-0.738624	-0.857659	-0.782649	-0.787695
16	1996	7.927926	1.002924	1.148504	0.960652	1.190684	1.190701
17	1997	46.076538	0.665721	0.802427	0.530757	0.638959	0.638956
18	1998	11.785240	-0.799062	-0.894162	-0.747615	-0.831772	-0.831774
19	1999	12.877575	-0.235389	-0.283506	-0.309238	-0.287890	-0.287890
20	2000	116.110357	-0.811118	-0.813839	-0.841717	-0.841319	-0.841319
21	2001	10.553750	-0.301227	-0.428029	-0.359835	-0.365530	-0.365533
22	2002	27.932650	-1.739608	-1.749818	-1.730187	-1.729392	-1.729392
23	2003	20.001513	-1.416809	-1.419593	-1.399244	-1.395173	-1.395171
24	2004	12.359575	0.276816	0.225236	0.225296	0.175437	0.175436
25	2005	6.719199	0.246986	0.176022	0.254710	0.185263	0.185263
26	2006	4.779200	-0.494427	-0.538883	-0.546216	-0.555302	-0.555301
27	2007	5.323861	-1.043742	-1.064759	-0.997855	-1.035631	-1.035631
28	2008	2.316423	-1.306798	-1.311143	-1.329931	-1.323260	-1.323260
29	2009	5.277289	-1.334437	-1.332150	-1.334862	-1.327644	-1.327644
30	2010	3.831291	-0.895418	-0.912684	-0.898003	-0.897819	-0.897819
31	2011	6.935024	-1.188153	-1.224951	-1.149853	-1.186143	-1.186142
32	2012	1.754772	0.032346	0.133339	-0.033847	0.062522	0.100368
33	2013	6.241224	-0.764879	-0.784005	-0.754495	-0.778898	-0.749310
34	2014	11.261329	-1.128316	-1.136363	-1.138539	-1.138806	-1.140575

	Year	Volume kurtosis	Open kurtosis	High kurtosis	Low kurtosis	Close kurtosis	Adj Close kurtosis
35	2015	4.886658	-1.000260	-1.300645	-0.944306	-1.226869	-1.133459
36	2016	5.953769	-1.284405	-1.296980	-1.278520	-1.291167	-1.324219
37	2017	14.306797	-0.251850	-0.268850	-0.240958	-0.258794	-0.308908
38	2018	3.049823	-0.917747	-0.940046	-0.895199	-0.925985	-0.955381
39	2019	5.979527	-0.329565	-0.315268	-0.314443	-0.302246	-0.311371
40	2020	2.222071	-1.394260	-1.427698	-1.377215	-1.401041	-1.405907
41	2021	0.882969	0.202839	0.232662	0.096757	0.143717	0.124964
42	2022	2.042568	-0.403897	-0.282617	-0.531761	-0.407847	-0.398334