Project 4 - Clustering_Model

In [1]:

```
# Importing Modules...
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.decomposition import PCA
from sklearn.cluster import KMeans
from datetime import datetime
```

In [2]:

```
# Loading Data...
data = pd.read_csv("data_stocks.csv")
data.head()
```

Out[2]:

	DATE	SP500	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDA
0	1491226200	2363.6101	42.3300	143.6800	129.6300	82.040	1(
1	1491226260	2364.1001	42.3600	143.7000	130.3200	82.080	1(
2	1491226320	2362.6799	42.3100	143.6901	130.2250	82.030	1(
3	1491226380	2364.3101	42.3700	143.6400	130.0729	82.000	1(
4	1491226440	2364.8501	42.5378	143.6600	129.8800	82.035	1(

5 rows × 502 columns

In [3]:

```
# DATA FRAME columns
data.columns
```

Out[3]:

```
Index(['DATE', 'SP500', 'NASDAQ.AAL', 'NASDAQ.AAPL', 'NASDAQ.ADBE',
       'NASDAQ.ADI', 'NASDAQ.ADP', 'NASDAQ.ADSK', 'NASDAQ.AKAM', 'NASDAQ.ALX
N',
       ...
       'NYSE.WYN', 'NYSE.XEC', 'NYSE.XEL', 'NYSE.XL', 'NYSE.XOM', 'NYSE.XR
X',
       'NYSE.XYL', 'NYSE.YUM', 'NYSE.ZBH', 'NYSE.ZTS'],
      dtype='object', length=502)
```

In [4]:

```
#The summary statistics of the data dataframe
data.describe()
```

Out[4]:

	DATE	SP500	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.AD
count	4.126600e+04	41266.000000	41266.000000	41266.000000	41266.00000	41266.00000(
mean	1.497749e+09	2421.537882	47.708346	150.453566	141.31793	79.44687:
std	3.822211e+06	39.557135	3.259377	6.236826	6.91674	2.00028:
min	1.491226e+09	2329.139900	40.830000	140.160000	128.24000	74.80000(
25%	1.494432e+09	2390.860100	44.945400	144.640000	135.19500	78.03000(
50%	1.497638e+09	2430.149900	48.360000	149.945000	142.26000	79.41000(
75%	1.501090e+09	2448.820100	50.180000	155.065000	147.10000	80.58000(
max	1.504210e+09	2490.649900	54.475000	164.510000	155.33000	90.44000(

8 rows × 502 columns

In [5]:

```
#Return a tuple representing the dimensionality of data DataFrame.
data.shape
```

Out[5]:

```
(41266, 502)
```

In [6]:

```
#Check for any 'NAN' in the dataframe.
data.isnull().values.any()
```

Out[6]:

```
False
```

In [7]:

```
#Making a copy of the data dataframe
data_new = data.copy()
```

In [8]:

```
#Dropping DATE and SP500 columns ( For PCA and KMeans clustering )
data_new.drop(['DATE', 'SP500'], axis=1, inplace = True)
```

In [9]:

```
#Returns the first 5 rows of new dataframe
data_new.head()
```

Out[9]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.ADSK
0	42.3300	143.6800	129.6300	82.040	102.2300	85.2200
1	42.3600	143.7000	130.3200	82.080	102.1400	85.6500
2	42.3100	143.6901	130.2250	82.030	102.2125	85.5100
3	42.3700	143.6400	130.0729	82.000	102.1400	85.4872
4	42.5378	143.6600	129.8800	82.035	102.0600	85.7001

5 rows × 500 columns

In [10]:

```
#Return a tuple representing the dimensionality of data DataFrame.
data_new.shape
```

Out[10]:

```
(41266, 500)
```

In [11]:

```
#Return columns of data dataframe.
data_new.columns
```

Out[11]:

```
Index(['NASDAQ.AAL', 'NASDAQ.AAPL', 'NASDAQ.ADBE', 'NASDAQ.ADI', 'NASDAQ.AD
P',
       'NASDAQ.ADSK', 'NASDAQ.AKAM', 'NASDAQ.ALXN', 'NASDAQ.AMAT',
       'NASDAQ.AMD',
       ...
       'NYSE.WYN', 'NYSE.XEC', 'NYSE.XEL', 'NYSE.XL', 'NYSE.XOM', 'NYSE.XR
X',
       'NYSE.XYL', 'NYSE.YUM', 'NYSE.ZBH', 'NYSE.ZTS'],
      dtype='object', length=500)
```

In [12]:

```
#Creating an instance of PCA
pca = PCA(n_components=3)
```

In [13]:

```
#Fitting the pca object
pca.fit(data_new)
```

Out[13]:

```
PCA(copy=True, iterated_power='auto', n_components=3, random_state=None,
  svd_solver='auto', tol=0.0, whiten=False)
```

In [14]:

```
#Transforming the data_new dataframe
data_new_pca = pca.transform(data_new)
```

In [15]:

```
#Return a tuple representing the dimensionality of data_new_pca DataFrame.
data_new_pca.shape
```

Out[15]:

```
(41266, 3)
```

In [16]:

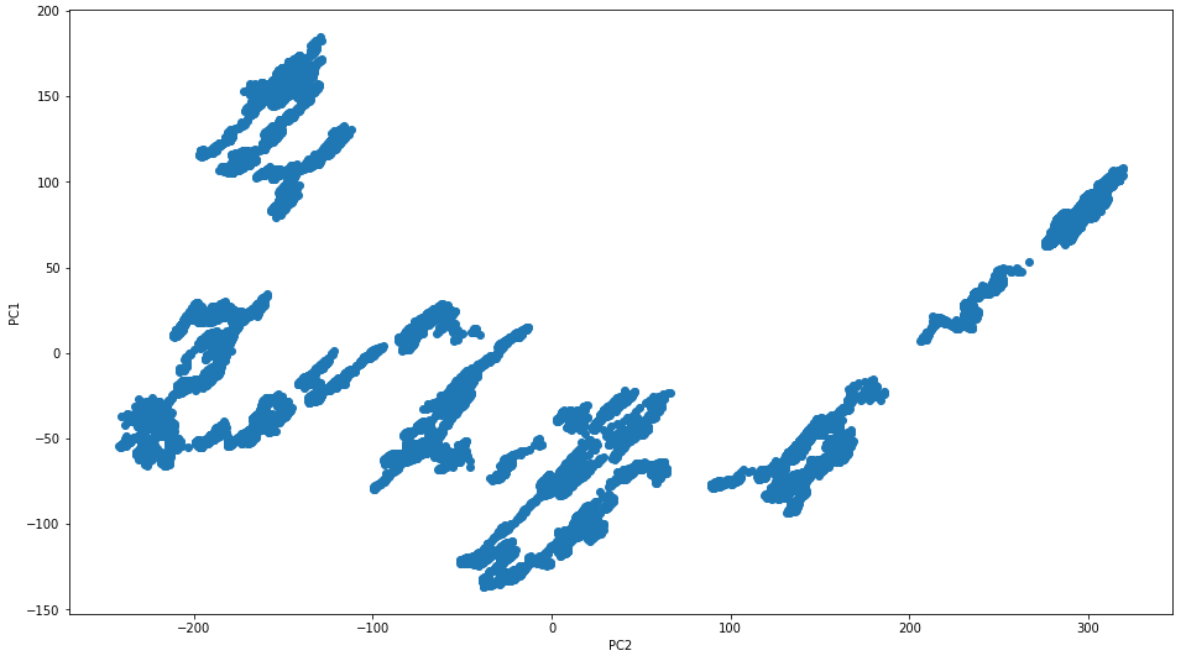
```
#Return a tuple representing the dimensionality of data_new_pca dataframe's 1st row.
data_new_pca[:1].shape
```

Out[16]:

```
(1, 3)
```

In [17]:

```
#Scatter Plot
plt.figure(figsize=(16,9))
plt.scatter(data_new_pca[:,0],data_new_pca[:,1])
plt.ylabel('PC1')
plt.xlabel('PC2')
plt.show()
```



In [18]:

```
#Returns explained variance array
pca.explained_variance_
```

Out[18]:

```
array([29552.09265369,  6861.74983656,  3733.21310127])
```

In [19]:

```
#Transformed data_new dataframe
data_new_pca
```

Out[19]:

```
array([[ 306.7014591 ,   83.2240155 ,   60.07857695],
       [ 303.07353165,   79.13615476,   60.77302358],
       [ 301.07978706,   78.29811525,   58.19185706],
       ...,
       [-195.96140814,  118.43969002,  -59.58416424],
       [-195.00245926,  118.98135157,  -59.20880889],
       [-196.53458095,  117.44181957,  -60.4663573 ]])
```

In [20]:

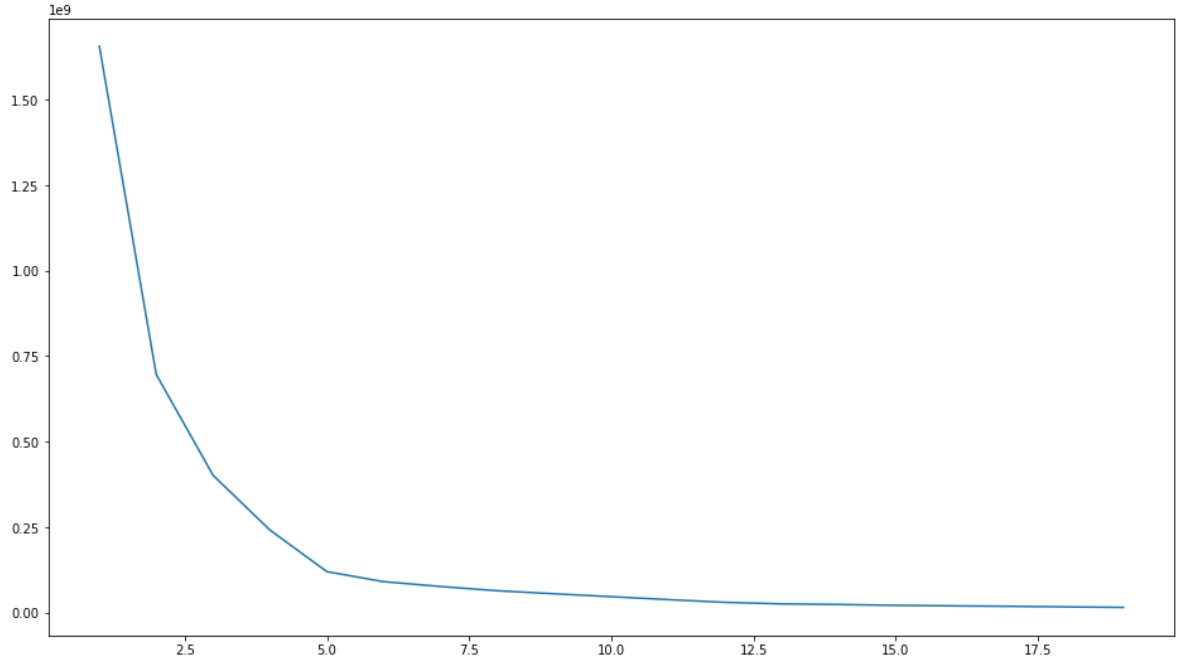
```
#Finding k and intertia for KMeans clustering using elbow method
k = []
inertia = []
for i in range(1,20):
    k_means = KMeans(n_clusters = i)
    k_means.fit(data_new_pca)
    k.append(i)
    inertia.append(k_means.inertia_)
#Inertia List Data
inertia
```

Out[20]:

```
[1656668248.9838114,
 696465712.1408997,
 401139261.74254686,
 240947420.6309432,
 119415116.52315424,
 90075099.94494633,
 75991875.22458388,
 63542419.934407495,
 54713380.12923427,
 46213902.5253189,
 37590426.483729444,
 29645281.050814178,
 25109501.366533708,
 23438071.18626793,
 20844857.468325835,
 19312934.113680866,
 18068094.545340694,
 16408778.414630968,
 14856678.051683592]
```

In [21]:

```
#Plot to find number of clusters (Elbow Method)
plt.figure(figsize=(16,9))
plt.plot(k,inertia)
plt.show()
```



In [22]:

```
#Initializing and fitting KMeans
km = KMeans(n_clusters = 5)
km.fit(data_new)
```

Out[22]:

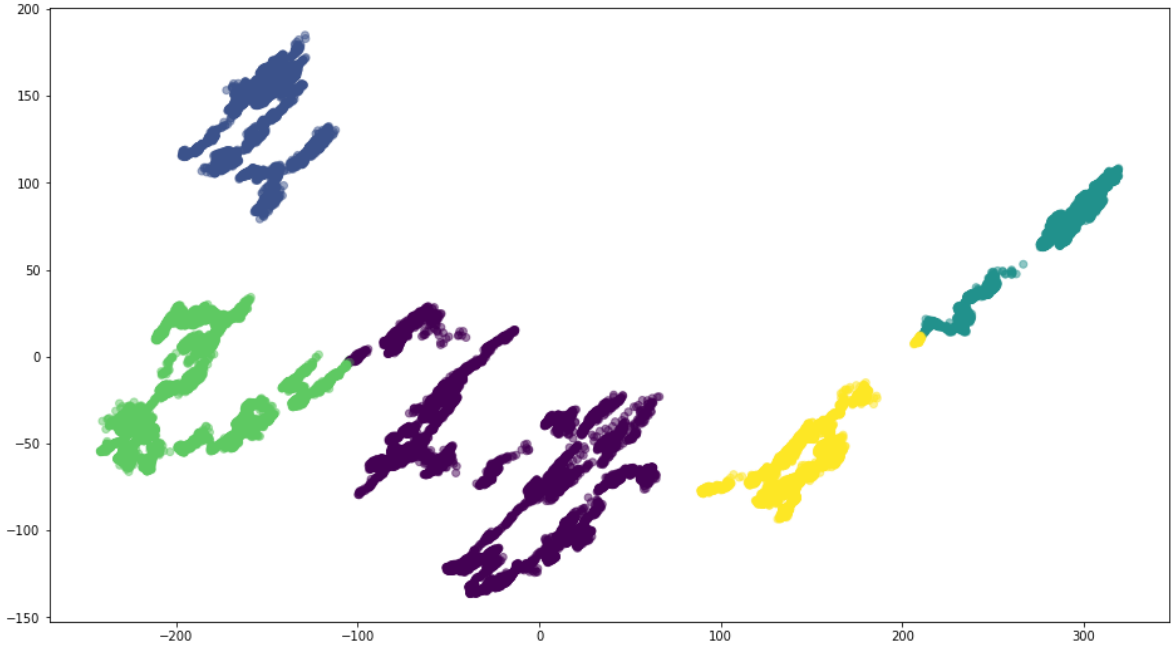
```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
    n_clusters=5, n_init=10, n_jobs=None, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

In [23]:

```
#Predicted values using KMeans
y_predict = km.predict(data_new)
```

In [24]:

```
#Scatter Plot
x = data_new_pca[:,0]
y = data_new_pca[:,1]
plt.figure(figsize=(16,9))
plt.scatter(x, y, c = y_predict, alpha=0.5)
plt.show()
```



In [25]:

```
#Adding 'Y_PREDICT' column in data_new dataframe
data_new['Y_PREDICT'] = y_predict
```

In [26]:

```
#Returns the first 5 rows of data_new dataframe
data_new.head()
```

Out[26]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.ADSK
0	42.3300	143.6800	129.6300	82.040	102.2300	85.2200
1	42.3600	143.7000	130.3200	82.080	102.1400	85.6500
2	42.3100	143.6901	130.2250	82.030	102.2125	85.5100
3	42.3700	143.6400	130.0729	82.000	102.1400	85.4872
4	42.5378	143.6600	129.8800	82.035	102.0600	85.7001

5 rows × 501 columns

In [27]:

```
#Returns 'Y_PREDICT' column containing counts of unique values in data_new dataframe.
data_new['Y_PREDICT'].value_counts()
```

Out[27]:

```
0    12029
3     8905
2     6976
4     6709
1     6647
Name: Y_PREDICT, dtype: int64
```

1. There are various stocks for which we have collected a data set, which all stocks are apparently simillar in performance.

In [28]:

```
#Read CSV (comma-separated) file into DataFrame
stocks= pd.read_csv('data_stocks.csv')
```

In [29]:

```
#Returns the first 5 rows of stocks dataframe
stocks.head()
```

Out[29]:

	DATE	SP500	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDA
0	1491226200	2363.6101	42.3300	143.6800	129.6300	82.040	1(
1	1491226260	2364.1001	42.3600	143.7000	130.3200	82.080	1(
2	1491226320	2362.6799	42.3100	143.6901	130.2250	82.030	1(
3	1491226380	2364.3101	42.3700	143.6400	130.0729	82.000	1(
4	1491226440	2364.8501	42.5378	143.6600	129.8800	82.035	1(

5 rows × 502 columns

In [30]:

```
#Adding a new column 'NEW_DATE' in stocks dataframe
stocks['NEW_DATE'] = pd.to_datetime(stocks['DATE'],unit='s')
```

In [31]:

```
#Creating a list of columns from stocks dataframe
cols = stocks.columns.tolist()
```

In [32]:

```
#Making 'NEW_DATE' as first column
cols = cols[-1:] + cols[:-1]
```

In [33]:

```
#cols list data
cols
```

Out[33]:

```
['NEW_DATE',
 'DATE',
 'SP500',
 'NASDAQ.AAL',
 'NASDAQ.AAPL',
 'NASDAQ.ADBE',
 'NASDAQ.ADI',
 'NASDAQ.ADP',
 'NASDAQ.ADSK',
 'NASDAQ.AKAM',
 'NASDAQ.ALXN',
 'NASDAQ.AMAT',
 'NASDAQ.AMD',
 'NASDAQ.AMGN',
 'NASDAQ.AMZN',
 'NASDAQ.ATVI',
 'NASDAQ.AVGO',
 'NASDAQ.BBBY',
```

In [34]:

```
#Removing 'DATE' and 'SP500' columns
cols.remove('DATE')
cols.remove('SP500')
```

In [35]:

```
#cols list data
cols
```

Out[35]:

```
['NEW_DATE',
 'NASDAQ.AAL',
 'NASDAQ.AAPL',
 'NASDAQ.ADBE',
 'NASDAQ.ADI',
 'NASDAQ.ADP',
 'NASDAQ.ADSK',
 'NASDAQ.AKAM',
 'NASDAQ.ALXN',
 'NASDAQ.AMAT',
 'NASDAQ.AMD',
 'NASDAQ.AMGN',
 'NASDAQ.AMZN',
 'NASDAQ.ATVI',
 'NASDAQ.AVGO',
 'NASDAQ.BBBY',
 'NASDAQ.BIIB',
 'NASDAQ.CA',
```

In [36]:

```
#Removing 'DATE' and 'SP500' from stocks dataframe
stocks.drop(columns=['DATE','SP500'],axis=1,inplace=True)
```

In [37]:

```
#Returns the first 5 rows of stocks dataframe
stocks.head()
```

Out[37]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.ADSK
0	42.3300	143.6800	129.6300	82.040	102.2300	85.2200
1	42.3600	143.7000	130.3200	82.080	102.1400	85.6500
2	42.3100	143.6901	130.2250	82.030	102.2125	85.5100
3	42.3700	143.6400	130.0729	82.000	102.1400	85.4872
4	42.5378	143.6600	129.8800	82.035	102.0600	85.7001

5 rows × 501 columns

In [38]:

```
#Creating a df objects which has cols list column data from stocks dataframe
df = stocks[cols]
```

In [39]:

```
#Returns the first 5 rows of df dataframe
df.head()
```

Out[39]:

	NEW_DATE	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NA
0	2017-04-03 13:30:00	42.3300	143.6800	129.6300	82.040	102.2300	
1	2017-04-03 13:31:00	42.3600	143.7000	130.3200	82.080	102.1400	
2	2017-04-03 13:32:00	42.3100	143.6901	130.2250	82.030	102.2125	
3	2017-04-03 13:33:00	42.3700	143.6400	130.0729	82.000	102.1400	
4	2017-04-03 13:34:00	42.5378	143.6600	129.8800	82.035	102.0600	

5 rows × 501 columns

In [40]:

```
#Return a tuple representing the dimensionality of df DataFrame.
df.shape
```

Out[40]:

```
(41266, 501)
```

In [41]:

```
#Setting NEW_DATE as index
df.set_index('NEW_DATE',inplace=True)
```

In [42]:

```
#Returns the first 5 rows of df dataframe
df.head()
```

Out[42]:

NEW_DATE	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASD
2017-04-03 13:30:00	42.3300	143.6800	129.6300	82.040	102.2300	
2017-04-03 13:31:00	42.3600	143.7000	130.3200	82.080	102.1400	
2017-04-03 13:32:00	42.3100	143.6901	130.2250	82.030	102.2125	
2017-04-03 13:33:00	42.3700	143.6400	130.0729	82.000	102.1400	
2017-04-03 13:34:00	42.5378	143.6600	129.8800	82.035	102.0600	

5 rows × 500 columns

In [43]:

```
#Creating transpose of the df dataframe
df_transpose = df.transpose()
```

In [44]:

```
#Returns the first 5 rows of df_transpose dataframe
df_transpose.head()
```

Out[44]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.AAL	42.33	42.36	42.3100	42.3700	42.5378	42.5399	42.470	42.47
NASDAQ.AAPL	143.68	143.70	143.6901	143.6400	143.6600	143.7800	143.864	143.81
NASDAQ.ADBE	129.63	130.32	130.2250	130.0729	129.8800	130.0700	130.180	130.14
NASDAQ.ADI	82.04	82.08	82.0300	82.0000	82.0350	82.0400	82.120	82.19
NASDAQ.ADP	102.23	102.14	102.2125	102.1400	102.0600	102.0400	102.330	102.37

5 rows × 41266 columns

In [45]:

```
#Creating an instance of PCA
pca_new = PCA(n_components=3)
```

In [46]:

```
#Fitting and tranforming the df_transpose dataframe
df_transpose_pca = pca_new.fit_transform(df_transpose)
```

In [47]:

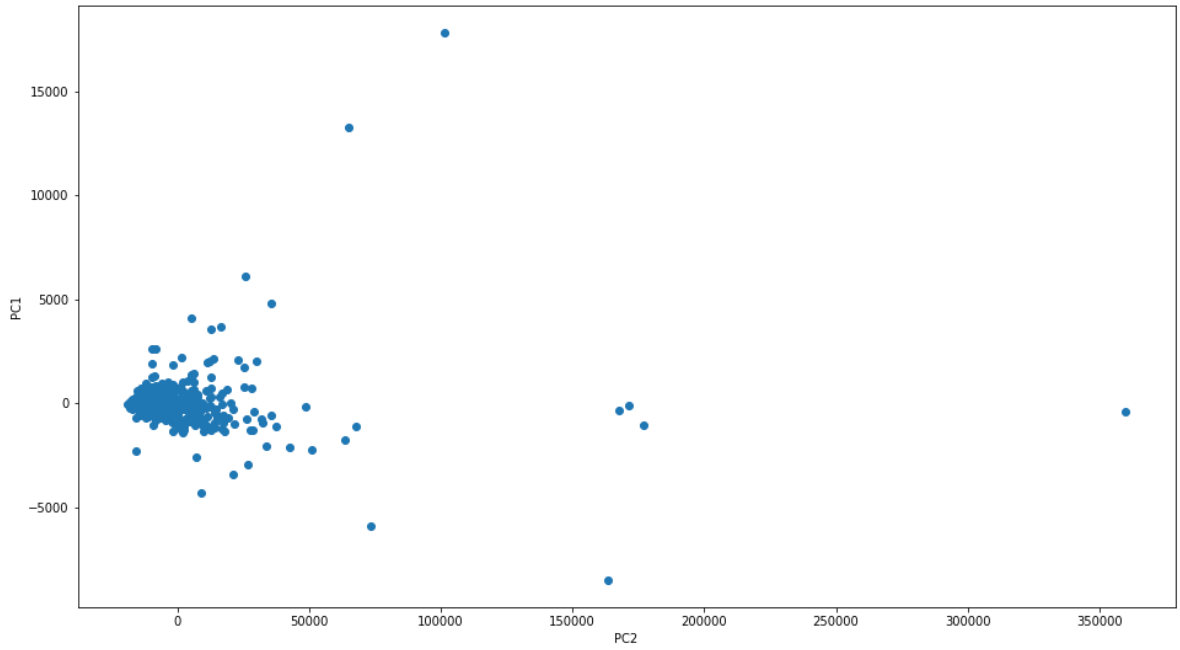
```
#Return a tuple representing the dimensionality of df_transpose_pca DataFrame.
df_transpose_pca.shape
```

Out[47]:

```
(500, 3)
```

In [48]:

```
#Scatter Plot
plt.figure(figsize=(16,9))
plt.scatter(df_transpose_pca[:,0],df_transpose_pca[:,1])
plt.ylabel('PC1')
plt.xlabel('PC2')
plt.show()
```



In [49]:

```
#Returns explained variance array
pca_new.explained_variance_
```

Out[49]:

```
array([6.83663380e+08, 1.88968326e+06, 3.64618121e+05])
```

In [50]:

```
#Finding k_new and intertia_new list data for KMeans clustering using elbow method
k_new = []
inertia_new = []
for i in range(2,10):
    km_new=KMeans(n_clusters=i)
    km_new.fit(df_transpose)
    k_new.append(i)
    inertia_new.append(km_new.inertia_)
```

In [51]:

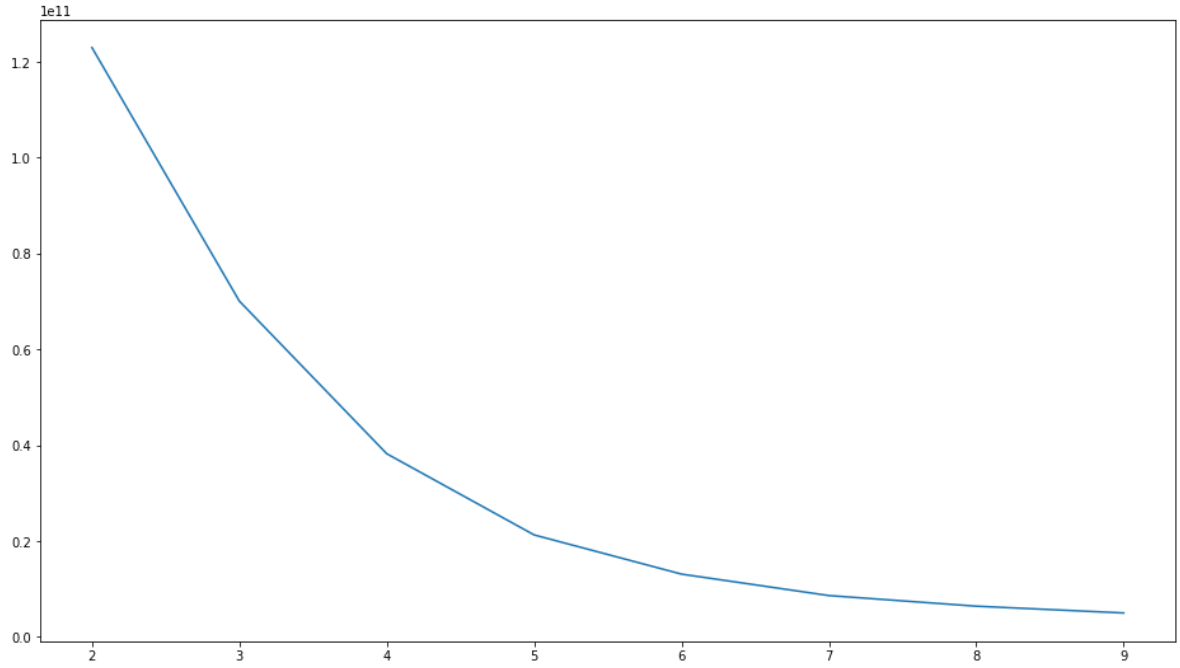
```
#inertia_new list data
inertia_new
```

Out[51]:

```
[123006116341.59377,
 70077213281.75235,
 38215217146.89998,
 21266645730.415653,
 13080591570.448448,
 8602418066.634865,
 6386405324.411578,
 4970342004.526426]
```

In [52]:

```
#Plot to find number of clusters (Elbow Method)
plt.figure(figsize=(16,9))
plt.plot(k_new,inertia_new)
plt.show()
```



In [53]:

```
#Initializing and fitting KMeans
km = KMeans(n_clusters = 6)
km.fit(df_transpose)
```

Out[53]:

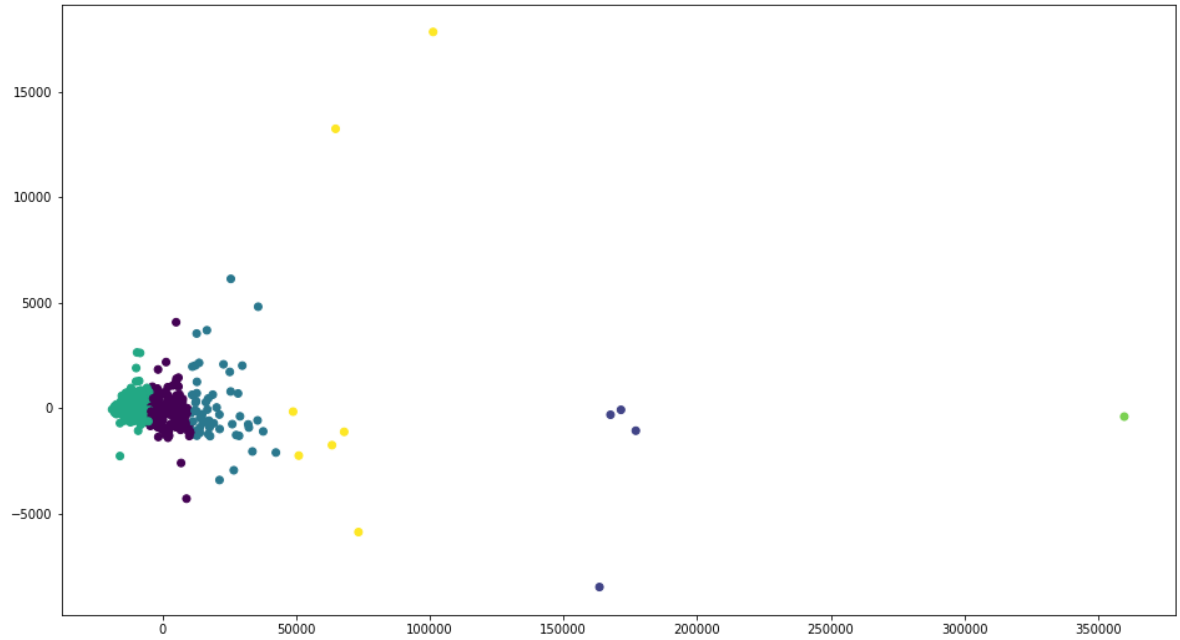
```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
    n_clusters=6, n_init=10, n_jobs=None, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

In [54]:

```
#Predicted values using KMeans
y_predict_new = km.predict(df_transpose)
```

In [55]:

```
#Scatter Plot
plt.figure(figsize=(16,9))
plt.scatter(df_transpose_pca[:,0],df_transpose_pca[:,1],c=y_predict_new)
plt.show()
```



In [56]:

```
#Adding 'y_predict_new' values in df_transpose dataframe creating 'Y_PREDICT' column
df_transpose['Y_PREDICT'] = y_predict_new
```

In [57]:

```
#Returns 'Y_PREDICT' column containing counts of unique values in df_transpose dataframe.
df_transpose['Y_PREDICT'].value_counts()
```

Out[57]:

```
3    263
0    169
2     56
5      7
1      4
4      1
Name: Y_PREDICT, dtype: int64
```

In [58]:

```
#Returns the first 5 rows of df_transpose dataframe
df_transpose.head()
```

Out[58]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.AAL	42.33	42.36	42.3100	42.3700	42.5378	42.5399	42.470	42.47
NASDAQ.AAPL	143.68	143.70	143.6901	143.6400	143.6600	143.7800	143.864	143.81
NASDAQ.ADBE	129.63	130.32	130.2250	130.0729	129.8800	130.0700	130.180	130.14
NASDAQ.ADI	82.04	82.08	82.0300	82.0000	82.0350	82.0400	82.120	82.19
NASDAQ.ADP	102.23	102.14	102.2125	102.1400	102.0600	102.0400	102.330	102.37

5 rows × 41267 columns

In [59]:

```
#Apparently similar performing stocks of Type-1 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==0]
```

Out[59]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.ADBE	129.6300	130.3200	130.2250	130.0729	129.8800	130.0700	130.1800	130.140
NASDAQ.ADI	82.0400	82.0800	82.0300	82.0000	82.0350	82.0400	82.1200	82.190
NASDAQ.ADP	102.2300	102.1400	102.2125	102.1400	102.0600	102.0400	102.3300	102.370
NASDAQ.ADSK	85.2200	85.6500	85.5100	85.4872	85.7001	85.9200	85.9120	85.820
NASDAQ.ALXN	121.5200	121.4800	121.9300	121.4400	121.6000	121.7000	121.6300	121.610
NASDAQ.CBOE	81.0300	81.2100	81.2100	81.1300	81.1200	81.2200	81.1600	81.160
NASDAQ.CELG	124.8900	124.9900	125.0000	124.7300	124.8300	125.0800	125.1300	125.170
NASDAQ.CME	119.3850	118.8100	118.8300	118.6800	118.9350	118.7800	118.9800	118.960
NASDAQ.CTAS	126.8600	126.8600	126.9576	126.7000	126.8900	126.9500	126.8400	126.720
NASDAQ.CTXS	83.5400	83.9400	83.9400	83.7064	84.0000	83.8743	83.9100	83.810
NASDAQ.DLTR	78.4600	78.6200	78.6311	78.7060	78.5850	78.5700	78.1750	78.175
NASDAQ.EA	89.9000	89.7300	89.2400	89.1700	88.9500	89.4500	89.4150	89.720
NASDAQ.EXPE	126.3400	126.3900	126.5300	126.5400	126.4500	126.4800	126.2900	126.300
NASDAQ.FFIV	142.7300	142.5700	142.4600	142.2200	142.3000	142.3528	142.3300	142.380
NASDAQ.FISV	115.5000	115.3300	115.1600	115.0700	115.0300	115.0700	115.2740	115.180
NASDAQ.HAS	99.9200	99.8100	99.7944	99.7600	99.9600	100.1700	100.1400	100.020
NASDAQ.INCY	134.9300	135.6100	135.8999	135.6200	136.5800	136.7000	137.2500	137.650
NASDAQ.INTU	116.0500	115.7900	115.6800	115.4700	115.4612	115.4500	115.6900	115.900
NASDAQ.JBHT	91.9500	91.9500	92.4900	92.4900	92.4900	92.1300	92.1000	92.380
NASDAQ.KHC	90.8100	90.9400	91.0000	91.0000	90.8600	90.8300	90.8600	90.770
NASDAQ.KLAC	95.3400	95.0700	94.8100	95.0900	95.1300	95.4100	95.3100	95.175
NASDAQ.MAR	94.2250	94.3200	94.5200	94.4550	94.4100	94.3500	94.3600	94.200
NASDAQ.MCHP	73.9400	74.1790	74.1100	74.3300	74.2800	74.4000	74.4600	74.190
NASDAQ.NTRS	87.4100	86.7100	86.6400	86.5000	86.6500	86.7400	86.5200	86.560
NASDAQ.NVDA	108.9800	109.2100	108.7300	108.8352	108.9300	108.9130	108.9500	109.080
NASDAQ.SNI	78.4500	79.1500	79.0900	79.1900	79.1500	79.3200	79.2400	78.980
NASDAQ.SNPS	72.4100	72.4359	72.3300	72.2900	72.3300	72.2700	72.4100	72.400
NASDAQ.SRCL	82.8900	83.1450	82.9400	82.9400	82.8200	82.8700	82.8700	82.870
NASDAQ.SWKS	98.1500	98.4200	98.4000	98.1648	98.2000	98.4700	98.2900	98.320
NASDAQ.TROW	68.2500	68.4500	68.2900	68.1900	68.2800	68.4400	68.4000	68.380
...	...	...	...	...	...	...	...	..
NYSE.RJF	76.2900	76.2900	76.3000	76.1500	76.3400	76.3600	76.3200	76.440
NYSE.RL	81.9800	81.6700	81.6700	81.9400	81.9400	81.9000	81.9300	81.860

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NYSE.SJM	131.1500	130.7550	130.9500	131.1600	131.1500	131.1443	131.1443	131.250
NYSE.SLG	106.5700	106.5700	106.3600	106.5800	106.8200	106.7450	107.0000	107.000
NYSE.SPGI	131.2700	131.2700	131.2500	131.3200	131.4900	131.4600	131.4600	131.440
NYSE.SRE	110.4800	109.5800	109.8300	110.3400	109.8700	110.0100	110.5000	110.170
NYSE.STT	79.6700	79.7800	79.7100	79.4500	79.7200	79.6600	79.6100	79.625
NYSE.SWK	133.1200	133.3700	133.3050	133.1700	133.3400	133.2084	133.2800	133.250
NYSE.SYK	131.6500	131.5650	131.3500	131.6100	131.5900	131.8500	131.8200	131.880
NYSE.TAP	95.8200	95.9000	95.8200	95.8050	95.8200	95.7400	95.8500	95.940
NYSE.TEL	74.6500	74.6500	74.5200	74.6600	74.6600	74.7100	74.6900	74.710
NYSE.TIF	95.4100	95.4200	95.4900	95.1900	95.1900	95.4400	95.4000	95.245
NYSE.TMK	77.1500	77.3699	77.3699	76.9300	76.9300	76.9300	76.9300	77.240
NYSE.TRV	120.5700	120.5700	120.8050	120.4550	120.6500	120.5926	120.6200	120.670
NYSE.TSO	81.3300	81.0100	80.6700	80.9200	80.8000	80.7900	80.8250	80.830
NYSE.TWX	97.9300	97.8900	97.9800	97.9200	97.8100	97.9250	97.9200	97.960
NYSE.UHS	124.6600	124.8200	124.8200	124.8200	124.8200	124.3900	124.0200	124.070
NYSE.UNP	106.2700	106.2700	106.3800	106.3300	106.5600	106.5900	106.7600	106.540
NYSE.UPS	107.6500	107.6500	107.3100	107.4900	107.4800	107.3952	107.4300	107.470
NYSE.URI	124.5901	124.5901	124.7100	124.4889	125.2449	125.1900	125.2280	125.230
NYSE.UTX	112.3600	112.3600	112.1300	112.0100	112.2300	112.1500	112.1700	112.150
NYSE.V	89.3500	89.3500	89.1600	89.0600	89.1000	89.2000	89.3750	89.450
NYSE.VAR	91.1300	91.2100	91.0800	91.0150	91.0000	91.0700	91.0500	90.915
NYSE.VMC	120.1300	120.1300	120.3368	120.3100	120.3600	120.3700	120.5975	120.510
NYSE.VNO	100.3500	100.0300	100.3900	100.3900	100.1100	100.1100	100.3550	100.110
NYSE.WMT	72.1000	72.1800	72.3192	72.4300	72.4800	72.5100	72.3503	72.270
NYSE.WYN	84.3700	84.3700	84.5850	84.4600	84.4700	84.4825	84.5000	84.520
NYSE.XEC	119.0350	119.0350	119.2600	119.2600	119.6100	119.5800	119.3200	119.110
NYSE.XOM	82.0300	82.0300	82.0200	82.0200	82.0300	82.0300	82.0566	82.000
NYSE.ZBH	122.0000	121.7700	121.7000	121.7000	121.6950	121.8500	122.1000	122.280

169 rows × 41267 columns

In [60]:

```
#Apparently similar performing stocks of Type-2 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==1]
```

Out[60]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-0 0 13:37:0
NASDAQ.AMZN	888.55	887.1173	887.5110	886.2700	886.5800	886.8000	887.31	888.599
NASDAQ.GOOG	829.41	833.1000	835.9184	833.6200	835.3200	835.1400	835.30	835.800
NASDAQ.GOOGL	847.83	852.1700	854.2800	852.7800	854.7500	853.9700	854.25	854.825
NASDAQ.ISRG	766.47	766.5400	767.0300	766.5853	766.5853	766.5853	766.61	767.265

4 rows × 41267 columns

In [61]:

```
#Apparently similar performing stocks of Type-3 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==2]
```

Out[61]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.AAPL	143.680	143.7000	143.6901	143.6400	143.6600	143.7800	143.8640	143.8100
NASDAQ.AMGN	164.630	164.6800	164.9050	164.7600	164.8500	164.9225	164.8200	164.7269
NASDAQ.AVGO	219.110	219.9800	219.3900	219.3000	219.1800	219.4900	219.7200	219.8700
NASDAQ.BIIB	274.080	273.9900	274.2750	273.5900	273.5400	274.1300	273.8700	274.0500
NASDAQ.COST	167.740	167.7760	168.0000	168.2000	168.0400	168.0950	168.0825	168.0100
NASDAQ.FB	141.940	141.8625	141.9100	141.8700	141.7800	141.9286	142.1500	142.0600
NASDAQ.HSIC	170.220	170.8150	170.8150	170.8150	170.8150	169.9750	169.9700	169.9700
NASDAQ.IDXX	155.070	155.4837	155.7100	155.5000	155.3500	155.3800	155.3411	155.4265
NASDAQ.ILMN	171.000	171.2950	171.0540	170.8200	170.8384	170.7150	171.3700	171.2700
NASDAQ.LRCX	128.530	128.6500	128.3100	128.2101	128.2200	128.5102	128.4301	128.4300
NASDAQ.NFLX	146.750	147.4008	147.0500	147.0196	147.2600	147.2492	147.2400	147.4250
NASDAQ.ORLY	266.000	263.7750	263.0850	262.6900	260.6800	260.5600	257.4500	259.2300
NASDAQ.ULTA	285.510	285.9800	285.3650	284.6600	284.9000	284.9000	285.7600	285.7400
NYSE.ADS	249.015	249.0150	249.6400	249.8000	249.9200	249.9300	249.9400	249.7300
NYSE.AGN	238.720	238.7200	238.8700	238.8700	239.3000	239.5300	239.6500	239.6200
NYSE.AMG	164.980	164.9800	164.5750	164.2100	164.2000	163.6500	163.6050	164.2650
NYSE.ANTM	165.925	165.9250	166.6000	166.2600	166.5550	166.9000	166.9900	167.0250
NYSE.AVB	184.000	184.0000	183.3950	183.3700	183.0400	183.0400	182.6950	182.7150
NYSE.AYI	207.650	207.6500	208.7500	208.9900	208.4300	207.6200	207.8800	207.8300
NYSE.BA	177.350	177.3500	177.4100	177.1900	177.0200	177.1900	177.3601	177.3550
NYSE.BCR	249.000	248.3400	247.9800	247.5100	248.4400	248.6500	248.7800	248.7800
NYSE.BDX	184.000	184.4000	184.5700	184.4000	184.4000	184.3900	184.4600	184.3700
NYSE.BRK.B	166.760	167.0000	166.8100	166.5800	166.6300	166.7076	166.7450	166.6900
NYSE.CI	146.700	146.9200	147.7400	147.5900	147.9000	147.9600	147.9400	147.9500
NYSE.CMI	151.410	151.5900	151.4500	151.1500	151.3800	151.6400	151.6700	151.5500
NYSE.COO	199.780	200.8600	200.8600	199.7050	199.5100	199.5000	199.9700	199.8000
NYSE.ESS	231.630	231.4500	231.4500	231.1700	231.1700	231.1700	231.1700	231.1700
NYSE.FDX	196.590	196.5900	196.5800	196.4850	196.2900	196.3889	196.5600	196.7400
NYSE.GD	187.900	187.9000	188.0400	187.8300	187.9000	187.7400	187.8100	187.8000
NYSE.GS	229.720	229.9400	229.4600	229.2900	229.5400	229.8300	229.7300	229.5000
NYSE.GWW	233.690	233.6900	234.2800	233.7500	233.8950	233.7800	233.8700	233.8200
NYSE.HD	146.800	146.9600	147.0200	146.9100	147.0000	146.9800	146.7600	146.8000
NYSE.HUM	207.140	207.1400	207.1400	207.1300	207.1300	207.9299	207.5900	207.6400

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NYSE.IBM	173.560	173.5600	173.6500	173.7800	173.6800	173.5800	173.6100	173.7388
NYSE.LLL	164.810	165.2900	165.2900	165.2900	165.2900	165.4900	165.3000	165.4950
NYSE.LMT	267.730	267.6000	268.2800	268.1700	268.0900	268.1800	268.0647	267.9800
NYSE.MCK	148.080	148.2200	148.4800	148.6300	148.1100	148.1100	148.0000	148.0900
NYSE.MHK	229.920	229.9200	231.0800	230.2100	230.6300	230.6900	230.8400	230.8400
NYSE.MLM	218.200	218.2600	218.2600	217.5500	217.5350	217.7792	218.0300	218.1100
NYSE.MMM	192.080	192.0800	192.0800	191.7700	191.8800	191.7700	191.8700	191.8330
NYSE.MTB	155.270	155.6900	155.5350	155.0400	155.4200	155.3536	155.3000	155.2100
NYSE.NOC	238.455	238.4550	238.1600	238.3900	238.4630	238.3500	238.3800	238.0700
NYSE.PH	160.000	160.8600	160.7400	160.7200	160.8100	160.9900	161.0500	160.7550
NYSE.PSA	219.000	218.4600	218.3300	217.5150	217.7300	217.7900	217.7770	218.2700
NYSE.PXD	186.040	185.9300	186.3000	186.3600	186.1850	186.5250	186.1700	185.8700
NYSE.ROK	155.420	155.9900	156.2300	156.0700	155.8500	156.0000	156.0000	155.9800
NYSE.ROP	207.130	207.1000	207.0912	206.6400	206.3800	206.1300	206.4500	206.4500
NYSE.RTN	152.200	152.2000	152.5300	152.5700	152.7600	152.6100	152.7500	152.6500
NYSE.SNA	168.460	168.4600	169.0800	168.6550	168.6550	168.6550	168.7100	168.6100
NYSE.SPG	172.460	172.4600	173.0450	172.6300	172.6700	172.6500	172.8100	172.7700
NYSE.STZ	162.070	162.0800	162.0500	161.6300	161.8550	161.9150	162.0500	161.9200
NYSE.TDG	219.440	219.4400	218.6800	218.6700	219.7900	220.4000	220.4200	220.7400
NYSE.TMO	153.980	153.9800	153.9550	153.7200	153.9800	153.9500	153.9700	154.2200
NYSE.UNH	164.620	164.3900	165.3000	165.1600	165.2400	165.6700	165.4200	165.2900
NYSE.WAT	156.390	157.1900	157.0700	156.9450	157.1600	157.1650	157.0400	157.4000
NYSE.WHR	171.830	171.5400	171.5500	171.2600	171.1800	171.1800	171.9650	171.9400

56 rows × 41267 columns

In [62]:

```
#Apparently similar performing stocks of Type-4 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==3]
```

Out[62]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.AAL	42.3300	42.3600	42.3100	42.3700	42.5378	42.5399	42.4700	42.4700
NASDAQ.AKAM	59.7600	59.8400	59.7950	59.6200	59.6200	59.6100	59.5400	59.4100
NASDAQ.AMAT	38.9900	39.0100	38.9100	38.8400	38.9300	38.9500	38.9600	38.9600
NASDAQ.AMD	14.6100	14.7100	14.6400	14.6300	14.6700	14.6800	14.6563	14.6840
NASDAQ.ATVI	49.8500	49.9400	49.8600	49.9150	49.9150	49.9700	50.0000	50.0200
NASDAQ.BBBY	39.4300	39.6800	39.6000	39.5700	39.5500	39.5800	39.4900	39.5100
NASDAQ.CA	31.7800	31.7800	31.7650	31.8300	31.8300	31.8300	31.7900	31.8100
NASDAQ.CERN	58.8200	58.4950	58.4700	58.4200	58.6000	58.5200	58.5900	58.6100
NASDAQ.CHRW	77.7250	77.9400	77.8150	77.9500	78.0500	78.0700	77.9700	78.0700
NASDAQ.CINF	72.4300	72.0400	72.0550	72.1400	72.2150	72.3000	72.4600	72.5300
NASDAQ.CMCSA	37.4700	37.5400	37.6100	37.6200	37.6250	37.6900	37.7650	37.7150
NASDAQ.CSCO	33.7400	33.8800	33.9000	33.8499	33.8400	33.8800	33.9000	33.8900
NASDAQ.CSX	46.6450	46.6100	46.8850	46.7000	46.8562	46.9664	46.9400	46.9200
NASDAQ.CTSH	59.7200	59.8300	59.9050	59.7000	59.5000	59.5300	59.4500	59.4400
NASDAQ.DISCA	29.0200	29.1400	29.1500	29.1250	29.1400	29.2300	29.2100	29.1800
NASDAQ.DISCK	28.3700	28.4000	28.3750	28.3500	28.4100	28.4000	28.4300	28.4100
NASDAQ.DISH	63.4900	63.4600	63.5250	63.5300	63.6344	63.6300	63.7000	63.5400
NASDAQ.EBAY	33.3975	33.3950	33.4100	33.3350	33.4000	33.4300	33.4650	33.5300
NASDAQ.ESRX	66.0400	66.0000	65.9200	66.0300	65.9197	65.7688	65.8600	65.9150
NASDAQ.ETFC	34.8600	34.8500	34.9300	34.7050	34.8500	34.8100	34.9200	34.8770
NASDAQ.EXPD	56.1700	56.3700	56.4900	56.5500	56.5026	56.5000	56.6500	56.6700
NASDAQ.FAST	51.4500	51.6550	51.7800	51.7800	51.8400	51.9052	51.8950	51.7600
NASDAQ.FITB	25.5400	25.5100	25.5000	25.4100	25.4800	25.4800	25.4900	25.4900
NASDAQ.FLIR	36.3100	36.2200	36.3700	36.3400	36.3400	36.4000	36.3400	36.3000
NASDAQ.FOX	31.9000	31.7900	31.8500	31.7900	31.8000	31.8300	31.8400	31.8000
NASDAQ.FOXA	32.4800	32.4289	32.4900	32.4300	32.4300	32.4500	32.4600	32.4300
NASDAQ.GILD	67.9000	67.8300	67.9000	67.7900	67.8200	67.7900	67.7700	67.6800
NASDAQ.GRMN	51.1900	51.1900	51.3800	51.2200	51.3100	51.3100	51.2200	51.2900
NASDAQ.GT	36.0700	36.0200	35.9190	35.8100	35.7700	35.7750	35.8600	35.8400
NASDAQ.HBAN	13.5200	13.4800	13.4700	13.4400	13.4600	13.4500	13.4650	13.4700
...	...	...	...	...	...	...	...	...
NYSE.SYY	52.0600	51.8650	51.8600	51.8900	51.9200	51.9400	51.9700	51.9800
NYSE.T	41.5800	41.5800	41.5600	41.5400	41.4975	41.4989	41.5200	41.5500

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NYSE.TGT	55.2200	55.1900	55.1300	55.0200	55.0800	55.0470	54.9050	54.9500
NYSE.TJX	79.0000	79.0000	78.8100	78.8500	78.7600	78.7600	78.6500	78.6900
NYSE.TSN	62.0000	61.9550	61.9600	62.0300	61.9300	62.1000	61.9900	62.0200
NYSE.TSS	53.4800	53.7100	53.6200	53.6200	53.4800	53.4150	53.4900	53.4500
NYSE.TXT	47.5600	47.5200	47.6000	47.5688	47.6082	47.6082	47.5600	47.5600
NYSE.UA	18.6100	18.5100	18.4397	18.3700	18.3850	18.4866	18.4625	18.3860
NYSE.UAA	19.8901	19.8901	19.7900	19.7875	19.7324	19.8200	19.8000	19.7224
NYSE.UAL	70.6900	70.7700	70.6200	70.6200	70.7800	70.8300	70.7500	70.6800
NYSE.UDR	36.2600	36.2600	36.1900	36.0000	36.0100	36.0100	35.9300	36.0300
NYSE.UNM	47.0400	47.0400	46.9800	46.8097	47.0200	47.0300	47.0300	46.9800
NYSE.USB	51.6200	51.5800	51.4600	51.3800	51.4347	51.4450	51.4900	51.5200
NYSE.VFC	55.2100	55.1200	55.0900	55.2100	55.3200	55.4500	55.4400	55.3200
NYSE.VLO	66.5900	66.3550	66.2417	66.1700	66.1800	66.1300	66.1800	66.1900
NYSE.VTR	65.0700	65.0400	65.0800	64.9500	64.8400	65.0900	64.9200	64.9600
NYSE.VZ	48.7300	48.7520	48.7200	48.6950	48.6300	48.6363	48.6600	48.6900
NYSE.WEC	60.4600	59.9300	60.0800	60.3900	60.1800	60.2800	60.3400	60.2400
NYSE.WFC	55.7000	55.8100	55.6400	55.6400	55.6800	55.6400	55.6600	55.6450
NYSE.WM	73.0800	72.9500	72.9800	73.0900	72.9900	73.0340	73.0201	73.0200
NYSE.WMB	29.6200	29.6400	29.6700	29.7900	29.7100	29.6900	29.6300	29.6200
NYSE.WRK	52.1600	52.2450	52.2000	52.1400	52.1400	52.1400	52.3500	52.2900
NYSE.WU	20.3400	20.3200	20.2600	20.2000	20.2200	20.2447	20.2600	20.2700
NYSE.WY	34.0000	34.0700	34.1100	33.9700	34.0000	34.0800	34.1100	34.0300
NYSE.XEL	44.4000	44.1100	44.0900	44.2500	44.1100	44.1500	44.2000	44.1200
NYSE.XL	39.8800	39.8800	39.9800	39.9900	39.9600	39.9600	39.9800	39.9500
NYSE.XRX	7.3600	7.3800	7.3600	7.3500	7.3600	7.3600	7.3650	7.3600
NYSE.XYL	50.2200	50.2200	50.1200	50.1600	50.2000	50.2000	50.3300	50.3000
NYSE.YUM	63.8600	63.7400	63.7500	63.8800	63.9100	63.8400	63.8300	63.8300
NYSE.ZTS	53.3500	53.3500	53.3650	53.3800	53.2400	53.2200	53.3300	53.3700

263 rows × 41267 columns

In [63]:

```
#Apparently similar performing stocks of Type-5 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==4]
```

Out[63]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.PCLN	1776.26	1779.4606	1777.55	1776.4301	1774.08	1775.0	1776.585	1777.01

1 rows × 41267 columns

In [64]:

```
#Apparently similar performing stocks of Type-6 are following:
df_transpose.loc[df_transpose['Y_PREDICT']==5]
```

Out[64]:

NEW_DATE	2017-04-03 13:30:00	2017-04-03 13:31:00	2017-04-03 13:32:00	2017-04-03 13:33:00	2017-04-03 13:34:00	2017-04-03 13:35:00	2017-04-03 13:36:00	2017-04-03 13:37:00
NASDAQ.CHTR	330.73	330.73	330.730	330.730	330.7300	330.7300	330.730	330.7300
NASDAQ.EQIX	400.94	400.40	401.985	401.500	399.7704	399.7704	401.015	400.7900
NASDAQ.REGN	388.00	390.16	389.890	390.375	390.5144	390.7350	390.105	390.1239
NYSE.AZO	720.50	719.46	715.010	712.270	708.0000	704.3200	703.500	704.1500
NYSE.BLK	385.20	385.20	385.200	383.330	384.1050	383.9800	384.041	384.0200
NYSE.CMG	449.35	449.93	451.680	450.390	450.1000	449.9025	450.810	451.4662
NYSE.SHW	309.90	311.50	311.500	311.500	311.5000	311.5000	311.500	311.5000

7 rows × 41267 columns

2. How many Unique patterns that exist in the historical stock data set, based on fluctuations in price.

In [65]:

```
#There are 5 unique patterns that exists in historical stock data set, based on fluctuation
#Pattern-1 stocks (based on fluctuations in price):
data_new.loc[data_new['Y_PREDICT']==0]
```

Out[65]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AI
13685	46.2700	154.9000	139.1600	82.2800	99.3700	113.2
13686	46.3115	154.6320	139.3760	81.7300	99.4200	113.6
13687	46.1100	154.0000	139.4900	81.6400	99.4400	113.6
13688	46.0650	154.1800	139.5500	81.7360	99.5200	113.8
13689	46.1500	154.2000	139.2500	81.6500	99.5300	113.8
13690	46.1300	154.0800	139.4632	81.6800	99.5100	113.8
13691	46.0600	154.1300	139.4700	81.4950	99.2100	114.1
13692	46.1600	154.3103	139.2200	81.5300	99.2600	113.9
13693	46.1800	154.3900	139.2000	81.6000	99.2400	113.8
13694	46.0700	154.2600	139.1400	81.4800	99.2600	113.5
13695	46.1600	154.2400	139.2600	81.6600	99.2700	113.6
13696	46.1900	154.2470	139.3300	81.5500	99.2400	113.5
13697	46.0297	154.2200	139.4238	81.3200	99.2750	113.5
13698	46.0600	154.2240	139.2000	81.2300	99.2100	113.3
13699	46.0500	154.0653	139.0000	81.2200	99.2262	113.2
13700	46.0960	153.9800	138.9900	81.2100	99.3000	113.2
13701	46.0600	154.0000	139.0000	81.0600	99.2100	113.2
13702	46.0300	153.9749	138.9200	80.9500	99.2200	113.0
13703	46.0001	153.9400	138.9400	81.0188	99.2080	113.0
13704	46.0300	153.9350	139.0000	81.3500	99.1700	113.1
13705	46.0900	153.9100	138.9200	81.2650	99.1300	113.1
13706	46.0800	153.9400	138.8800	81.1900	99.1300	113.0
13707	46.0800	153.8700	138.8600	81.1000	99.0900	112.9
13708	46.1600	153.8400	138.8150	81.0141	99.1600	112.9
13709	46.2300	153.8200	138.7300	81.0500	99.1600	112.8
13710	46.3000	153.9400	138.7412	81.1000	99.0900	112.8
13711	46.3700	153.8600	138.8400	80.9300	99.1300	112.9
13712	46.4000	153.8700	138.9300	80.9400	99.1500	113.0
13713	46.3900	153.8805	138.9100	81.0200	99.1800	113.1
13714	46.3700	153.9800	138.9200	81.0100	99.2200	113.1
...	...	...	...	...	...	
25684	52.9450	143.8000	142.2450	77.9800	104.6800	104.0

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AD
25685	52.9700	143.7295	142.3000	77.9800	104.7500	103.9
25686	52.9900	143.8250	142.3500	78.0299	104.7832	104.0
25687	53.0201	143.8700	142.4000	78.0400	104.8000	103.9
25688	53.0400	144.0000	142.4300	78.1200	104.8600	104.0
25689	53.0800	143.9900	142.3900	78.0900	104.8500	103.9
25690	53.1400	144.0200	142.3015	78.0600	104.8600	103.9
25691	53.0350	144.0100	142.3500	77.9900	104.8500	103.8
25692	52.9900	143.9300	142.2650	77.9800	104.8000	103.7
25693	53.0900	143.9800	142.2900	78.0200	104.7700	103.9
25694	53.0899	144.0100	142.4100	77.9900	104.7200	103.9
25701	53.0200	144.0200	142.3400	77.9300	104.8650	103.9
25702	52.9500	144.0000	142.3700	77.9000	104.8900	103.8
25703	52.9650	143.9300	142.2900	77.8300	104.8800	103.7
25704	52.9400	143.9900	142.2500	77.8700	104.8800	103.8
25705	52.9300	144.0100	142.2100	77.8500	104.8700	103.8
25706	52.8950	143.9400	142.1300	77.8400	104.8400	104.0
25707	52.8850	143.9498	142.1600	77.8650	104.8100	103.9
25708	52.8850	143.9200	142.0600	77.8464	104.8500	104.0
25709	52.9105	143.9550	142.1300	77.8900	104.8800	104.0
25710	52.9300	143.9000	142.0250	77.8300	104.8800	103.9
25711	52.9400	143.8100	142.0400	77.8000	104.9100	103.9
25712	52.9600	143.8900	142.0800	77.8200	104.9800	103.8
25713	52.9650	143.8800	142.0800	77.8300	104.9900	103.8
25714	52.9250	143.8600	142.0100	77.8300	105.0450	103.8
25715	52.8750	143.8700	141.9500	77.7800	105.0400	103.7
25716	52.8800	143.8300	141.8710	77.7200	105.0050	103.6
25717	52.8700	143.8053	141.8250	77.7350	104.9700	103.5
25718	52.8300	143.8050	141.8820	77.7600	104.9000	103.5
25721	52.7400	143.8950	141.9400	77.8400	104.9200	103.7

12029 rows × 501 columns

In [66]:

```
#Pattern-2 stocks (based on fluctuations in price):
data_new.loc[data_new['Y_PREDICT']==1]
```

Out[66]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AD
34619	49.7400	159.2500	146.6800	78.5000	108.5100	107.7
34620	49.6400	159.7100	146.8600	78.3600	108.6100	106.9
34621	49.5271	159.8600	146.5600	78.1400	108.8600	106.6
34622	49.4950	159.6400	146.6600	78.2100	108.6500	107.0
34623	49.2800	159.8600	146.7000	78.2300	108.6700	107.3
34624	49.3494	159.9114	146.5900	78.2850	108.7700	107.4
34625	49.4400	160.1298	146.7800	78.2900	108.4542	107.2
34626	49.3800	160.2000	146.7000	78.0000	108.4800	107.2
34627	49.3400	159.9000	146.6810	77.8400	108.6800	107.0
34628	49.4600	159.8600	147.0800	77.8500	108.5200	107.1
34629	49.3750	159.9800	147.1700	77.8800	108.5400	107.1
34630	49.4500	160.0505	147.4800	77.8900	108.3800	107.5
34631	49.4700	159.5700	147.1200	77.6700	108.2700	107.2
34632	49.4200	159.3000	146.7300	77.5600	108.2400	106.8
34633	49.3600	159.4121	146.6758	77.4575	108.1300	107.0
34634	49.3100	159.6200	146.6500	77.4900	108.1200	106.8
34635	49.3800	159.3700	146.6500	77.4200	108.1300	106.6
34636	49.3600	159.6500	146.8550	77.4600	108.0900	106.8
34637	49.4200	159.6841	146.8640	77.4900	108.1650	106.8
34638	49.4500	159.8400	146.9000	77.5104	108.2600	106.9
34639	49.4750	159.8100	146.9200	77.5200	108.2400	106.8
34640	49.4201	159.6938	146.7500	77.5400	108.4050	106.6
34641	49.4200	159.5175	146.5600	77.5800	108.5450	106.5
34642	49.4500	159.6516	146.6900	77.6400	108.6100	106.7
34643	49.5200	159.7400	146.7900	77.6400	108.6850	106.7
34644	49.5600	159.8728	147.0000	77.7000	108.8500	107.0
34645	49.6050	159.9800	146.8000	77.7400	108.9700	106.9
34646	49.5722	160.1800	147.0100	77.7200	109.0400	106.9
34647	49.5600	160.5500	147.0700	77.7200	109.2100	107.0
34648	49.6450	160.4400	147.0700	77.7600	109.2900	107.1
...	...	...	...	...	...	
41236	44.7200	164.0700	155.0800	83.8000	106.4600	114.4
41237	44.7150	164.0750	155.0700	83.8000	106.4600	114.4
41238	44.7850	164.1500	155.1200	83.8050	106.5000	114.4

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AD
41239	44.8000	164.1300	155.1300	83.6900	106.4900	114.5
41240	44.7800	164.1400	155.1300	83.6600	106.5100	114.5
41241	44.7800	164.1450	155.1100	83.7900	106.5400	114.5
41242	44.7900	164.2050	155.1000	83.7500	106.6000	114.5
41243	44.7800	164.2350	155.1300	83.7100	106.6300	114.5
41244	44.8000	164.2900	155.1600	83.7100	106.6550	114.5
41245	44.7850	164.2800	155.1600	83.6500	106.6500	114.5
41246	44.7800	164.2600	155.1200	83.7200	106.6600	114.5
41247	44.7800	164.2634	155.1100	83.7600	106.6900	114.5
41248	44.8000	164.2400	155.0800	83.7850	106.7050	114.6
41249	44.8300	164.2700	155.1300	83.8200	106.7800	114.6
41250	44.8300	164.2000	155.1100	83.8100	106.7400	114.6
41251	44.7550	164.1350	155.0900	83.7300	106.6600	114.5
41252	44.7300	164.1800	155.0600	83.7300	106.6200	114.5
41253	44.7150	164.2300	155.0700	83.7299	106.6500	114.4
41254	44.7450	164.2100	155.0700	83.7100	106.6500	114.5
41255	44.7400	164.1850	155.0600	83.7100	106.6500	114.4
41256	44.7200	164.1600	155.0600	83.7100	106.6200	114.5
41257	44.7300	164.0600	155.0100	83.7000	106.5600	114.5
41258	44.7600	164.1900	155.1200	83.6300	106.6200	114.5
41259	44.7500	164.2000	155.0900	83.6900	106.6000	114.5
41260	44.7100	164.1400	155.0400	83.6600	106.6300	114.4
41261	44.7200	164.1100	155.0900	83.6700	106.5650	114.4
41262	44.7300	164.1200	155.1600	83.6500	106.5900	114.5
41263	44.7400	164.0100	155.0650	83.6200	106.5200	114.4
41264	44.7100	163.8800	154.9600	83.5800	106.4000	114.3
41265	44.7400	163.9800	155.1600	83.6900	106.4700	114.4

6647 rows × 501 columns

In [67]:

```
#Pattern-3 stocks (based on fluctuations in price):
data_new.loc[data_new['Y_PREDICT']==2]
```

Out[67]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.ADS
0	42.3300	143.6800	129.6300	82.0400	102.2300	85.22
1	42.3600	143.7000	130.3200	82.0800	102.1400	85.65
2	42.3100	143.6901	130.2250	82.0300	102.2125	85.51
3	42.3700	143.6400	130.0729	82.0000	102.1400	85.48
4	42.5378	143.6600	129.8800	82.0350	102.0600	85.70
5	42.5399	143.7800	130.0700	82.0400	102.0400	85.92
6	42.4700	143.8640	130.1800	82.1200	102.3300	85.91
7	42.4700	143.8100	130.1400	82.1900	102.3700	85.82
8	42.3900	143.8150	130.1000	82.2300	102.3800	85.88
9	42.3300	143.8000	130.2100	82.1650	102.3300	85.86
10	42.4000	143.8900	130.1400	82.2100	102.2000	85.89
11	42.2900	143.9700	130.3700	82.1500	102.2800	85.96
12	42.2900	143.9199	130.4599	82.1450	102.2700	85.95
13	42.3900	144.0239	130.6300	82.1600	102.3100	86.05
14	42.4197	144.0500	130.6500	82.2100	102.3200	86.13
15	42.4300	144.0638	130.6950	82.1500	102.2800	86.12
16	42.4400	144.0200	130.5700	82.1400	102.3000	86.30
17	42.4000	144.0200	130.4750	82.1143	102.2100	86.38
18	42.3900	143.9700	130.3700	82.0400	102.1700	86.50
19	42.4400	144.0300	130.3800	82.1500	102.1800	86.60
20	42.3800	143.9800	130.2900	82.0800	102.1500	86.65
21	42.3400	143.9000	130.3500	82.0800	102.1900	86.69
22	42.2950	143.9100	130.4400	82.0100	102.2200	86.87
23	42.2800	143.9100	130.4700	81.9800	102.2100	87.08
24	42.2500	143.8600	130.4700	82.0100	102.2100	86.82
25	42.2600	143.8600	130.2800	81.9500	102.1700	86.68
26	42.2100	143.8935	130.2600	81.8700	102.1900	86.61
27	42.2000	143.9300	130.2200	81.8500	102.2000	86.49
28	42.1700	143.9500	130.1700	81.7700	102.2500	86.45
29	42.1400	143.8700	130.1800	81.8400	102.2400	86.38
...	...	...	...	...	...	
6946	43.6700	143.6950	133.3200	77.8610	105.1800	90.50
6947	43.6401	143.6658	133.3000	77.8500	105.1850	90.51
6948	43.6900	143.7500	133.3800	77.8800	105.2000	90.55

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.ADS
6949	43.6950	143.7500	133.4000	77.8900	105.2200	90.57
6950	43.7150	143.7700	133.3960	77.8700	105.2100	90.61
6951	43.6750	143.7900	133.4100	77.8400	105.2200	90.60
6952	43.7100	143.7900	133.4190	77.8300	105.2200	90.63
6953	43.6878	143.7950	133.4200	77.8600	105.2100	90.65
6954	43.6950	143.8000	133.4200	77.8700	105.2300	90.63
6955	43.7250	143.7770	133.4400	77.8450	105.2322	90.54
6956	43.7350	143.7828	133.5100	77.8800	105.2500	90.56
6957	43.7400	143.7900	133.4700	77.8300	105.2650	90.63
6958	43.7700	143.7700	133.4300	77.8400	105.2600	90.59
6959	43.7550	143.7400	133.4250	77.8000	105.2450	90.59
6960	43.7550	143.7835	133.4700	77.8500	105.2500	90.64
6961	43.8200	143.7950	133.4700	77.8500	105.2500	90.65
6962	43.8500	143.8000	133.4600	77.8600	105.2500	90.74
6963	43.8450	143.7700	133.4565	77.8600	105.2600	90.74
6964	43.8000	143.7700	133.4580	77.7900	105.2400	90.69
6965	43.7950	143.7800	133.4500	77.7900	105.2300	90.72
6966	43.8200	143.7900	133.4300	77.7800	105.2100	90.70
6967	43.8300	143.7650	133.3900	77.7900	105.2100	90.71
6968	43.8350	143.7800	133.3950	77.8100	105.2100	90.69
6969	43.7550	143.7000	133.3300	77.7600	105.1700	90.64
6970	43.7600	143.7150	133.3500	77.7500	105.1700	90.63
6971	43.7544	143.7100	133.3700	77.7800	105.1700	90.64
6972	43.7700	143.7143	133.3600	77.8000	105.1800	90.61
6974	43.7800	143.7400	133.3450	77.7950	105.1400	90.61
6986	43.7900	143.5500	133.1900	77.5700	105.0200	90.52
6987	43.8250	143.5800	133.1700	77.6000	105.0400	90.53

6976 rows × 501 columns

In [68]:

```
#Pattern-4 stocks (based on fluctuations in price):
data_new.loc[data_new['Y_PREDICT']==3]
```

Out[68]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AI
25695	53.1200	144.0400	142.3800	77.9900	104.7300	103.9
25696	53.0900	143.9500	142.2900	77.9500	104.7100	103.9
25697	53.0100	143.9800	142.2900	77.9350	104.7256	103.8
25698	53.0400	144.0400	142.2700	77.9300	104.7300	103.9
25699	53.0750	144.1050	142.3950	78.0200	104.8000	104.0
25700	53.0500	144.0500	142.3603	77.9600	104.8300	104.0
25719	52.8350	143.8300	141.8800	77.7900	104.9300	103.6
25720	52.8000	143.8100	141.8800	77.7800	104.9000	103.6
25722	52.7750	143.9400	142.0700	77.8700	104.9000	103.7
25723	52.8000	143.8900	141.9400	77.8300	104.8400	103.6
25724	52.8400	143.9200	142.0000	77.8400	104.9100	103.7
25725	52.8200	143.9000	142.0800	77.8700	104.9900	103.8
25726	52.8500	143.8800	142.0800	77.8600	105.0300	103.7
25727	52.8800	143.9001	142.0500	77.8900	104.9400	103.7
25728	52.8500	143.8300	141.9900	77.8700	104.8800	103.7
25729	52.8500	143.8399	141.9800	77.8300	104.9500	103.6
25730	52.8800	143.8700	142.0600	77.9000	104.9200	103.6
25731	52.8950	143.8600	142.1000	77.9500	104.9200	103.6
25732	52.8400	143.7100	141.9052	77.8400	104.9000	103.4
25733	52.8200	143.7500	141.9400	77.9200	104.9400	103.4
25734	52.8900	143.7400	142.0000	77.9300	104.9600	103.3
25735	52.8250	143.7201	141.9100	77.8900	104.9950	103.3
25736	52.8000	143.7300	141.8900	77.8450	104.9800	103.3
25737	52.8050	143.7400	141.9000	77.8600	105.0300	103.3
25738	52.7700	143.7400	141.8222	77.8057	105.1032	103.2
25739	52.8200	143.7450	141.8900	77.8100	105.1250	103.2
25740	52.8600	143.7600	141.9200	77.8900	105.1429	103.2
25741	52.9200	143.7600	142.0000	77.9200	105.1500	103.3
25742	52.9500	143.8100	141.9500	77.9200	105.1000	103.3
25743	52.9500	143.8200	142.0400	77.9500	105.1350	103.4
...	...	...	...	...	...	
34589	50.1400	160.3278	147.9600	78.9700	108.4800	108.5
34590	50.1000	160.3032	147.9000	78.9600	108.5500	108.4
34591	50.0400	160.1200	147.8097	78.8800	108.4350	108.3

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AD
34592	50.0400	159.9744	147.6000	78.8500	108.4200	108.1
34593	50.0900	160.0694	147.7300	78.8600	108.3800	108.1
34594	50.0018	159.6700	147.5900	78.8100	108.3150	108.0
34595	49.9700	159.6900	147.4850	78.8073	108.2600	108.0
34596	49.9950	160.0200	147.6800	78.8700	108.3400	108.1
34597	49.9900	160.0600	147.6100	78.9100	108.3600	108.1
34598	49.9938	160.1790	147.5900	78.8800	108.2850	108.1
34599	50.0350	160.2000	147.6000	78.9500	108.2600	108.2
34600	50.0000	160.1295	147.5900	78.9700	108.2200	108.2
34601	50.0000	160.0734	147.5890	78.9400	108.3900	108.2
34602	49.9450	160.1560	147.6000	78.9500	108.4100	108.3
34603	49.9000	159.9380	147.5300	78.8200	108.3700	108.2
34604	49.9900	160.1300	147.6100	78.8900	108.3900	108.4
34605	49.9600	159.9850	147.5900	78.8900	108.1050	108.2
34606	49.9500	160.0600	147.6000	78.8400	108.2550	108.2
34607	49.9500	160.1100	147.5900	78.8900	108.2900	108.2
34608	49.9600	160.0800	147.7600	78.9300	108.4300	108.3
34609	49.9500	160.0000	147.6400	78.8750	108.4900	108.1
34610	49.9550	159.8800	147.5800	78.8700	108.3900	108.1
34611	49.9500	159.9301	147.6200	78.8700	108.4250	108.2
34612	49.9500	159.8900	147.6500	78.8500	108.3650	108.2
34613	49.9750	159.8800	147.7800	78.9000	108.3200	108.3
34614	49.9850	159.8500	147.7700	78.9400	108.3400	108.2
34615	49.9800	159.8850	147.7300	78.9150	108.3300	108.2
34616	50.0000	159.9900	147.8450	78.9200	108.5600	108.3
34617	50.0100	159.9700	147.7900	78.9700	108.4800	108.3
34618	50.0000	160.0800	147.8100	78.9700	108.5100	108.3

8905 rows × 501 columns

In [69]:

```
#Pattern-5 stocks (based on fluctuations in price):
data_new.loc[data_new['Y_PREDICT']==4]
```

Out[69]:

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AI
6973	43.7500	143.7200	133.3600	77.7800	105.150	90.6
6975	43.7700	143.7437	133.3500	77.7700	105.150	90.6
6976	43.7750	143.7600	133.3450	77.7500	105.160	90.6
6977	43.7622	143.7700	133.3400	77.7300	105.150	90.6
6978	43.8050	143.7500	133.3450	77.7400	105.135	90.6
6979	43.8100	143.7500	133.3500	77.7200	105.120	90.6
6980	43.8056	143.7200	133.3500	77.7200	105.120	90.6
6981	43.7950	143.6600	133.3400	77.6800	105.100	90.5
6982	43.8250	143.6999	133.3100	77.6900	105.110	90.6
6983	43.8300	143.7154	133.3200	77.7100	105.130	90.6
6984	43.7900	143.6259	133.2700	77.6750	105.080	90.5
6985	43.7700	143.5600	133.2200	77.6200	105.040	90.5
6988	43.8750	143.5900	133.1900	77.6000	105.090	90.5
6989	43.8700	143.6538	133.2100	77.6400	105.120	90.5
6990	43.9050	143.6501	133.2100	77.6100	105.115	90.5
6991	43.9100	143.6950	133.2400	77.6700	105.120	90.5
6992	43.9000	143.7000	133.2400	77.6700	105.130	90.6
6993	43.9150	143.6900	133.2400	77.7200	105.135	90.6
6994	43.8700	143.6922	133.2300	77.7100	105.125	90.5
6995	43.9050	143.7100	133.2400	77.6974	105.140	90.6
6996	43.8950	143.6900	133.2200	77.6900	105.150	90.5
6997	43.8744	143.6720	133.2100	77.6700	105.130	90.5
6998	43.8450	143.7200	133.1900	77.6700	105.105	90.5
6999	43.8200	143.7500	133.1950	77.6700	105.090	90.5
7000	43.8650	143.7254	133.2200	77.7050	105.120	90.5
7001	43.8600	143.7284	133.2200	77.7000	105.110	90.5
7002	43.8671	143.7500	133.2500	77.6600	105.105	90.5
7003	43.8600	143.7200	133.2500	77.6700	105.105	90.5
7004	43.8500	143.7500	133.2546	77.6800	105.130	90.5
7005	43.9500	143.7600	133.2708	77.6900	105.130	90.5
...	...	...	...	...	...	
13655	46.1400	153.7700	138.8950	81.9400	99.390	113.1
13656	46.1300	153.7750	138.9000	81.9400	99.370	113.2
13657	46.1400	153.8200	138.8950	81.9450	99.380	113.2

	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDAQ.ADP	NASDAQ.AD
13658	46.1200	153.7893	138.9050	81.9700	99.380	113.1
13659	46.1400	153.8000	138.9200	81.9800	99.420	113.1
13660	46.1750	153.8250	138.9400	82.0300	99.440	113.2
13661	46.1550	153.7966	138.9350	82.0200	99.440	113.1
13662	46.1790	153.8539	138.9700	82.0400	99.455	113.3
13663	46.1700	153.8750	138.9900	82.0600	99.485	113.3
13664	46.1510	153.8500	138.9800	82.0426	99.470	113.2
13665	46.1600	153.8942	139.0300	82.0497	99.460	113.2
13666	46.1700	153.9500	139.0400	82.1000	99.480	113.2
13667	46.1550	153.9400	139.0025	82.0500	99.475	113.1
13668	46.1450	153.9201	139.0050	82.0499	99.455	113.1
13669	46.1300	153.9400	138.9900	82.0700	99.430	113.2
13670	46.1253	153.9100	139.0100	82.0400	99.460	113.2
13671	46.1200	153.9300	138.9958	82.0200	99.460	113.3
13672	46.1300	153.9400	138.9800	82.0500	99.480	113.3
13673	46.1300	153.9250	138.9650	82.0700	99.500	113.2
13674	46.1600	153.9200	138.9800	82.0600	99.530	113.2
13675	46.1750	153.9500	138.9800	82.0900	99.530	113.3
13676	46.1700	153.9100	138.9300	82.0950	99.535	113.4
13677	46.1900	153.9521	138.9647	82.1300	99.520	113.4
13678	46.2000	153.9900	138.9700	82.1300	99.510	113.5
13679	46.2300	153.9900	138.9501	82.1200	99.500	113.5
13680	46.2050	154.0000	138.9250	82.1100	99.510	113.5
13681	46.2000	153.9700	138.9200	82.1300	99.500	113.5
13682	46.1900	153.9400	138.9000	82.1400	99.455	113.4
13683	46.2000	153.9000	138.8500	82.0800	99.430	113.3
13684	46.2100	153.9900	138.8600	82.0900	99.450	113.3

6709 rows × 501 columns

3. Identify which all stocks are moving together and which all stocks are different from each other.

In [70]:

```
#Read CSV (comma-separated) file into DataFrame
df_new = pd.read_csv('data_stocks.csv')
```

In [71]:

```
#Returns the first 5 rows of df_new dataframe
df_new.head()
```

Out[71]:

	DATE	SP500	NASDAQ.AAL	NASDAQ.AAPL	NASDAQ.ADBE	NASDAQ.ADI	NASDA
0	1491226200	2363.6101	42.3300	143.6800	129.6300	82.040	1(
1	1491226260	2364.1001	42.3600	143.7000	130.3200	82.080	1(
2	1491226320	2362.6799	42.3100	143.6901	130.2250	82.030	1(
3	1491226380	2364.3101	42.3700	143.6400	130.0729	82.000	1(
4	1491226440	2364.8501	42.5378	143.6600	129.8800	82.035	1(

5 rows × 502 columns

In [72]:

```
#Returns a tuple representing the dimensionality of df_new dataframe.
df_new.shape
```

Out[72]:

```
(41266, 502)
```

In [73]:

```
#Removing 'DATE' and 'SP500' columns from df_new dataframe
df_new.drop(columns=['DATE','SP500'],inplace=True,axis=1)
```

In [74]:

```
#Listing all the df_new dataframe columns
category_cols = df_new.columns
```

In [75]:

```
#Creating the columns with the difference of the previous row
for cat in category_cols:
    df_new["DIFF_"+ cat] = df_new[cat] - df_new[cat].shift(periods=1)
```

In [76]:

```
#Returns a tuple representing the dimensionality of df_new dataframe.
df_new.shape
```

Out[76]:

```
(41266, 1000)
```

In [77]:

```
#Removing the category_cols list columns from df_new dataframe
df_new.drop(category_cols,axis=1,inplace=True)
```

In [78]:

```
#Returns a tuple representing the dimensionality of df_new dataframe.
df_new.shape
```

Out[78]:

```
(41266, 500)
```

In [79]:

```
#Returns the first 5 rows of df_new dataframe
df_new.head()
```

Out[79]:

	DIFF_NASDAQ.AAL	DIFF_NASDAQ.AAPL	DIFF_NASDAQ.ADBE	DIFF_NASDAQ.ADI	DIFF_NAS
0	NaN	NaN	NaN	NaN	
1	0.0300	0.0200	0.6900	0.040	
2	-0.0500	-0.0099	-0.0950	-0.050	
3	0.0600	-0.0501	-0.1521	-0.030	
4	0.1678	0.0200	-0.1929	0.035	

5 rows × 500 columns

In [80]:

```
#Removing the rows which containd NaN
df_new.dropna(inplace=True)
```

In [81]:

```
#Returns the first 5 rows of df_new dataframe
df_new.head()
```

Out[81]:

	DIFF_NASDAQ.AAL	DIFF_NASDAQ.AAPL	DIFF_NASDAQ.ADBE	DIFF_NASDAQ.ADI	DIFF_NAS
1	0.0300	0.0200	0.6900	0.040	
2	-0.0500	-0.0099	-0.0950	-0.050	
3	0.0600	-0.0501	-0.1521	-0.030	
4	0.1678	0.0200	-0.1929	0.035	
5	0.0021	0.1200	0.1900	0.005	

5 rows × 500 columns

In [82]:

```
#Computes pairwise correlation of columns of df_new dataframe
df_new_corr = df_new.corr()
```

In [83]:

```
#Pairwise correlation dataframe of columns of df_new dataframe
df_new_corr
```

Out[83]:

	DIFF_NASDAQ.AAL	DIFF_NASDAQ.AAPL	DIFF_NASDAQ.ADBE	DIFF_NA
DIFF_NASDAQ.AAL	1.000000	0.174426	0.217080	
DIFF_NASDAQ.AAPL	0.174426	1.000000	0.418750	
DIFF_NASDAQ.ADBE	0.217080	0.418750	1.000000	
DIFF_NASDAQ.ADI	0.178742	0.340841	0.339840	
DIFF_NASDAQ.ADP	0.123543	0.134411	0.139378	
DIFF_NASDAQ.ADSK	0.176964	0.307407	0.405820	
DIFF_NASDAQ.AKAM	0.157573	0.216343	0.158988	
DIFF_NASDAQ.ALXN	0.001607	0.099733	0.128689	
DIFF_NASDAQ.AMAT	0.201468	0.426848	0.436495	
DIFF_NASDAQ.AMD	0.169037	0.312322	0.364518	
DIFF_NASDAQ.AMGN	0.211207	0.162813	0.191355	
DIFF_NASDAQ.AMZN	0.192244	0.429087	0.402510	
DIFF_NASDAQ.ATVI	0.181312	0.374506	0.505197	
DIFF_NASDAQ.AVGO	0.200875	0.434704	0.426062	
DIFF_NASDAQ.BBBY	0.089416	0.053547	-0.028159	
DIFF_NASDAQ.BIIB	0.139348	0.149576	0.183678	
DIFF_NASDAQ.CA	0.107593	0.160929	0.343338	
DIFF_NASDAQ.CBOE	0.083853	0.115148	0.128055	
DIFF_NASDAQ.CELG	0.211386	0.222359	0.251950	
DIFF_NASDAQ.CERN	0.128233	0.195906	0.269431	
DIFF_NASDAQ.CHRW	0.090437	0.083287	0.077045	
DIFF_NASDAQ.CHTR	0.111289	0.145371	0.147149	
DIFF_NASDAQ.CINF	0.101418	0.143724	0.100843	
DIFF_NASDAQ.CMCSA	0.125269	0.188820	0.183305	
DIFF_NASDAQ.CME	0.188858	0.137984	0.146153	
DIFF_NASDAQ.COST	0.058841	0.103562	0.092467	
DIFF_NASDAQ.CSCO	0.131301	0.187948	0.281916	
DIFF_NASDAQ.CSX	0.262975	0.213635	0.248895	
DIFF_NASDAQ.CTAS	0.145357	0.128471	0.145221	
DIFF_NASDAQ.CTSH	0.154751	0.296954	0.338683	
...	...	...	...	
DIFF_NYSE.USB	0.204524	0.192999	0.162756	
DIFF_NYSE.UTX	0.137926	0.102929	0.114241	
DIFF_NYSE.V	0.205370	0.327865	0.364027	

	DIFF_NASDAQ.AAL	DIFF_NASDAQ.AAPL	DIFF_NASDAQ.ADBE	DIFF_NA
DIFF_NYSE.VAR	0.107984	0.113776	0.136005	
DIFF_NYSE.VFC	0.105153	0.087001	0.068844	
DIFF_NYSE.VLO	0.155292	0.080561	0.030746	
DIFF_NYSE.VMC	0.092086	0.116134	0.110280	
DIFF_NYSE.VNO	0.032999	0.018841	0.007895	
DIFF_NYSE.VTR	0.009274	-0.008628	0.009295	
DIFF_NYSE.VZ	0.036339	0.054172	0.013495	
DIFF_NYSE.WAT	0.080984	0.100164	0.120478	
DIFF_NYSE.WEC	-0.000818	-0.017592	0.002429	
DIFF_NYSE.WFC	0.183643	0.165340	0.110756	
DIFF_NYSE.WHR	0.114159	0.088352	0.120773	
DIFF_NYSE.WM	0.102859	0.131910	0.161231	
DIFF_NYSE.WMB	0.129089	0.065624	0.047237	
DIFF_NYSE.WMT	0.039521	0.102954	0.038606	
DIFF_NYSE.WRK	0.157969	0.139555	0.150271	
DIFF_NYSE.WU	0.120176	0.144781	0.137194	
DIFF_NYSE.WY	0.139937	0.121996	0.128927	
DIFF_NYSE.WYN	0.170238	0.161219	0.166740	
DIFF_NYSE.XEC	0.040174	0.029748	-0.022685	
DIFF_NYSE.XEL	-0.023792	-0.044910	-0.015124	
DIFF_NYSE.XL	0.043705	0.072842	0.055081	
DIFF_NYSE.XOM	0.077091	0.029744	-0.037262	
DIFF_NYSE.XRX	0.041886	0.078002	0.035395	
DIFF_NYSE.XYL	0.146242	0.170238	0.162710	
DIFF_NYSE.YUM	0.139152	0.140320	0.175532	
DIFF_NYSE.ZBH	0.255040	0.116605	0.119406	
DIFF_NYSE.ZTS	0.091314	0.117178	0.149031	

500 rows × 500 columns