#Program Code (Part 2)

```
import pandas as pd
import numpy as np
import networkx as nx
import matplotlib.pyplot as plt



#Task 1
data=pd.read_csv(r"/Users/shivamsharma/Downloads/stock_data.csv")
data.sort_values(by=['Name'])



#Task 2(Identify, Sort unique names in alphabetical order, count them and list  to do first and last 5)
uniqueNames=df['Name'].unique()
no_of_uniques = uniqueNames.size
print(no_of_uniques)
arr=list(df.Name.unique())
arr.sort()
print("First five names:",arr[:5])
print("Last five names:", arr[-5:])



#Task 3
import pandas as pd
from datetime import datetime
df=pd.read_csv(r"/Users/shivamsharma/Downloads/stock_data.csv")
df["date"]=pd.to_datetime(df["date"])



Filtered_names=[]
Removed_names=[]
for name in df["Name"].unique():
    new_df=df[df["Name"]==name]
    First_date=datetime.strptime("Jan 1 2014", "%b %d %Y")
    Last_date=datetime.strptime("Dec 31 2017","%b %d %Y")
    if new_df["date"].iloc[0] > First_date or new_df["date"].iloc[-1] < Last_date:
        Filtered_names.append(name)
    else:
        Removed_names.append(name)
print("Filtered Names:", Filtered_names)
print("Removed Nmaes:", Removed_names)
```

```
print("Number of Names Left:", len(Removed_names))


print(invalid_stocks)
print(len(valid_stocks))
#Task 4
date=[]
for i in Filtered_names :
    print("for",i)
    print(df[df["Name"]==i]["date"])
    for j in (df[df["Name"]==i]["date"]):
        date.append(str(j))
        print(date[:5])
        print(date[-5:])



# Shape
df=pd.read_csv(r"/Users/shivamsharma/Downloads/stock_data.csv")
print('Shape:',df.shape)
df.head()


#Task 5
names_unique = list(df['Name'].unique())
df1=df[df['Name']==names_unique[0]][['date','close']]
for names in names_unique[1:]:
    df2=df[df['Name']==names][['date','close']]
    df1=pd.merge(df1,df2,how="outer",on="date")
names_unique.insert(0,'date')
df1.columns= names_unique
df1


from sklearn.decomposition import PCA
filtered_data=data[data['Name'].isin(valid_stock_names)].filter(items=['Name','date'])
dateGroup=filtered_data.groupby(by=['date']).count().reset_index()

filtered_dates=dateGroup[dateGroup['Name']==valid_stock_names.shape[0]]

filtered_dates=filtered_dates[(filtered_dates['date']>'2014-01-01')&(filtered_dates['date']<'201
7-12-31')]['date'].to_numpy()
print(len(filtered_dates))
new_dataframe={'date':filtered_dates}
filtered_df=data[data['date'].isin(filtered_dates)&data['Name'].isin(valid_stock_names)].filter(it
ems=['Name','date','close'])#.sort_values(by='date')
print(filtered_df)
filtered_data_array=filtered_df.to_numpy()
```

```
print(filtered_data_array)

for row in filtered_data_array:
      if row[0] in updated_dataframe:
        new_dataframe[row[0]]=np.append(new_dataframe[row[0]],row[2])
      else:
        updated_dataframe[row[0]]=np.array(row[2])


new_df=pd.DataFrame.from_dict(updated_dataframe)
print(new_df)


print(new_df.shape)
closing_values=new_df.filter(items=valid_stock_names)
closing_values_arr=closing_values.to_numpy()
current_close_values=closing_values_arr[1:]

previous_close_values=closing_values_arr[0:closing_values_arr.shape[0]-1]
close_value_ratio=np.array(np.multiply((current_close_values),(1/previous_close_values)))

##print(close_value_ratio.shape)
new_updated_df=pd.DataFrame(data=close_value_ratio,columns=valid_stock_names,index
=filtered_dates[1:])
print('closing values')
print(new_updated_df)




#Task 7 #Task 8 #Task 9 #Task 10
from pandas.core.frame import DataFrame
x= new_updated_df
pca = PCA()
principalcomponents = pca.fit_transform(x)
principalDF =pd.DataFrame(data = principalcomponents)
print('principal componenets')
print(principalDF)


x_values=range(1,21)
y_values=pca.explained_variance_ratio_[0:20]
plt.plot(x_values,y_values)
plt.xlabel('Component number')
plt.ylabel('explained variance')
plt.title('explained variance vs component number')
print('plotting graph')
plt.show()
```

```
print(y_values)
#cum_variance=np.cumsum(pca.explained_variance_ratio_)
#print(cum_variance)

def pca_analysis(dataframe:DataFrame):
    pca=PCA()
    pca.fit(dataframe)
    x_values=range(1,21)
    y_values=pca.explained_variance_ratio_[0:20]
    plt.plot(x_values,y_values)
    plt.xlabel('Component number')
    plt.ylabel('explained variance')
    plt.title('explained variance vs component number')
    plt.show()
    cum_variance=np.cumsum(pca.explained_variance_ratio_)
    plt.plot(range(1,cum_variance.shape[0]+1),cum_variance)
    plt.xlabel('Component number')
    plt.ylabel('cumulative variance')
    plt.title('cumulative variance vs component number')
    plt.show()
pca_analysis(new_updated_df)
normalised_df=pd.DataFrame((new_updated_df-new_updated_df.mean())/new_updated_df.
std())

pca_analysis(normalised_df)

normalised_df
```

Outputs:-

#Note:- ALL OUTPUTS ARE PRINTED BELOW AS THE SEQUENCE IN CODE

	date	open	high	low	close	volume	Name	
71611	2013-02-08	45.07	45.350	45.000	45.08	1824755	A	
72454	2016-06-15	45.11	45.380	44.900	44.94	2029194	A	

72453	2016-06-14	44.70	44.945	44.560	44.91	1624087	A
72452	2016-06-13	45.32	45.520	44.905	44.92	1272833	A
72451	2016-06-10	45.73	45.780	45.190	45.36	1261698	A
...	...	...	...	...	...	...	...
618197	2014-10-03	36.87	37.725	36.870	37.67	5948543	ZTS
618196	2014-10-02	36.58	36.970	36.460	36.91	3670179	ZTS
618195	2014-10-01	36.87	37.010	36.230	36.45	5508285	ZTS
618202	2014-10-10	37.02	37.300	36.640	36.68	3689339	ZTS
619039	2018-02-07	72.70	75.000	72.690	73.86	4534912	ZTS

619040 rows × 7 columns

In [9]:

505
First five names: ['A', 'AAL', 'AAP', 'AAPL', 'ABBV']
Last five names: ['XYL', 'YUM', 'ZBH', 'ZION', 'ZTS']

Filtered Names: ['APTV', 'BHF', 'BHGE', 'CFG', 'CSRA', 'DWDP', 'DXC', 'EVHC', 'FTV', 'GOOG', 'HLT', 'HPE', 'HPQ', 'INFO', 'KHC', 'NAVI', 'PYPL', 'QRVO', 'SYF', 'UA', 'WLTW', 'WRK']
Removed Nmaes: ['AAL', 'AAPL', 'AAP', 'ABBV', 'ABC', 'ABT', 'ACN', 'ADBE', 'ADI', 'ADM', 'ADP', 'ADSK', 'ADS', 'AEE', 'AEP', 'AES', 'AET', 'AFL', 'AGN', 'AIG', 'AIV', 'AIZ', 'AJG', 'AKAM', 'ALB',

'ALGN', 'ALK', 'ALLE', 'ALL', 'ALXN', 'AMAT', 'AMD', 'AME', 'AMGN', 'AMG', 'AMP', 'AMT', 'AMZN', 'ANDV', 'ANSS', 'ANTM', 'AON', 'AOS', 'APA', 'APC', 'APD', 'APH', 'ARE', 'ARNC', 'ATVI', 'AVB', 'AVGO', 'AVY', 'AWK', 'AXP', 'AYI', 'AZO', 'A', 'BAC', 'BAX', 'BA', 'BBT', 'BBY', 'BDX', 'BEN', 'BF.B', 'BIIB', 'BK', 'BLK', 'BLL', 'BMY', 'BRK.B', 'BSX', 'BWA', 'BXP', 'CAG', 'CAH', 'CAT', 'CA', 'CBG', 'CBOE', 'CBS', 'CB', 'CCI', 'CCL', 'CDNS', 'CELG', 'CERN', 'CF', 'CHD', 'CHK', 'CHRW', 'CHTR', 'CINF', 'CI', 'CLX', 'CL', 'CMA', 'CMCSA', 'CME', 'CMG', 'CMI', 'CMS', 'CNC', 'CNP', 'COF', 'COG', 'COL', 'COO', 'COP', 'COST', 'COTY', 'CPB', 'CRM', 'CSCO', 'CSX', 'CTAS', 'CTL', 'CTSH', 'CTXS', 'CVS', 'CVX', 'CXO', 'C', 'DAL', 'DE', 'DFS', 'DGX', 'DG', 'DHI', 'DHR', 'DISCA', 'DISCK', 'DISH', 'DIS', 'DLR', 'DLTR', 'DOV', 'DPS', 'DRE', 'DRI', 'DTE', 'DUK', 'DVA', 'DVN', 'D', 'EA', 'EBAY', 'ECL', 'ED', 'EFX', 'EIX', 'EL', 'EMN', 'EMR', 'EOG', 'EQIX', 'EQR', 'EQT', 'ESRX', 'ESS', 'ES', 'ETFC', 'ETN', 'ETR', 'EW', 'EXC', 'EXPD', 'EXPE', 'EXR', 'FAST', 'FBHS', 'FB', 'FCX', 'FDX', 'FE', 'FFIV', 'FISV', 'FIS', 'FITB', 'FLIR', 'FLR', 'FLS', 'FL', 'FMC', 'FOXA', 'FOX', 'FRT', 'FTI', 'F', 'GD', 'GE', 'GGP', 'GILD', 'GIS', 'GLW', 'GM', 'GOOGL', 'GPC', 'GPN', 'GPS', 'GRMN', 'GS', 'GT', 'GWW', 'HAL', 'HAS', 'HBAN', 'HBI', 'HCA', 'HCN', 'HCP', 'HD', 'HES', 'HIG', 'HII', 'HOG', 'HOLX', 'HON', 'HP', 'HRB', 'HRL', 'HRS', 'HSIC', 'HST', 'HSY', 'HUM', 'IBM', 'ICE', 'IDXX', 'IFF', 'ILMN', 'INCY', 'INTC', 'INTU', 'IPG', 'IP', 'IQV', 'IRM', 'IR', 'ISRG', 'ITW', 'IT', 'IVZ', 'JBHT', 'JCI', 'JEC', 'JNJ', 'JNPR', 'JPM', 'JWN', 'KEY', 'KIM', 'KLAC', 'KMB', 'KMI', 'KMX', 'KORS', 'KO', 'KR', 'KSS', 'KSU', 'K', 'LB', 'LEG', 'LEN', 'LH', 'LKQ', 'LLL', 'LLY', 'LMT', 'LNC', 'LNT', 'LOW', 'LRCX', 'LUK', 'LUV', 'LYB', 'L', 'MAA', 'MAC', 'MAR', 'MAS', 'MAT', 'MA', 'MCD', 'MCHP', 'MCK', 'MCO', 'MDLZ', 'MDT', 'MET', 'MGM', 'MHK', 'MKC', 'MLM', 'MMC', 'MMM', 'MNST', 'MON', 'MOS', 'MO', 'MPC', 'MRK', 'MRO', 'MSFT', 'MSI', 'MS', 'MTB', 'MTD', 'MU', 'MYL', 'M', 'NBL', 'NCLH', 'NDAQ', 'NEE', 'NEM', 'NFLX', 'NFX', 'NI', 'NKE', 'NLSN', 'NOC', 'NOV', 'NRG', 'NSC', 'NTAP', 'NTRS', 'NUE', 'NVDA', 'NWL', 'NWSA', 'NWS', 'OKE', 'OMC', 'ORCL', 'ORLY', 'OXY', 'O', 'PAYX', 'PBCT', 'PCAR', 'PCG', 'PCLN', 'PDCO', 'PEG', 'PEP', 'PFE', 'PFG', 'PGR', 'PG', 'PHM', 'PH', 'PKG', 'PKI', 'PLD', 'PM', 'PNC', 'PNR', 'PNW', 'PPG', 'PPL', 'PRGO', 'PRU', 'PSA', 'PSX', 'PVH', 'PWR', 'PXD', 'PX', 'QCOM', 'RCL', 'REGN', 'REG', 'RE', 'RF', 'RHI', 'RHT', 'RJF', 'RL', 'RMD', 'ROK', 'ROP', 'ROST', 'RRC', 'RSG', 'RTN', 'SBAC', 'SBUX', 'SCG', 'SCHW', 'SEE', 'SHW', 'SIG', 'SJM', 'SLB', 'SLG', 'SNA', 'SNI', 'SNPS', 'SO', 'SPGI', 'SPG', 'SRCL', 'SRE', 'STI', 'STT', 'STX', 'STZ', 'SWKS', 'SWK', 'SYK', 'SYMC', 'SYY', 'TAP', 'TDG', 'TEL', 'TGT', 'TIF', 'TJX', 'TMK', 'TMO', 'TPR', 'TRIP', 'TROW', 'TRV', 'TSCO', 'TSN', 'TSS', 'TWX', 'TXN', 'TXT', 'T', 'UAA', 'UAL', 'UDR', 'UHS', 'ULTA', 'UNH', 'UNM', 'UNP', 'UPS', 'URI', 'USB', 'UTX', 'VAR', 'VFC', 'VIAB', 'VLO', 'VMC', 'VNO', 'VRSK', 'VRSN', 'VRTX', 'VTR', 'VZ', 'V', 'WAT', 'WBA', 'WDC', 'WEC', 'WFC', 'WHR', 'WMB', 'WMT', 'WM', 'WU', 'WYNN', 'WYN', 'WY', 'XEC', 'XEL', 'XLNX', 'XL', 'XOM', 'XRAY', 'XRX', 'XYL', 'YUM', 'ZBH', 'ZION', 'ZTS']
Number of Names Left: 483

	Name	isValid
48	APTV	False
67	BHF	False
68	BHGE	False
93	CFG	False
120	CSRA	False
151	DWDP	False
152	DXC	False
172	EVHC	False
198	FTV	False

206 GOOG False
226 HLT False
231 HPE False
232 HPQ False
246 INFO False
267 KHC False
328 NAVI False
387 PYPL False
389 QRVO False
430 SYF False
453 UA False
484 WLTW False
488 WRK False
483

Shape: (619040, 7)

Out[21]:

	date	open	high	low	close	volume	Name
0	2013-02-08	15.07	15.12	14.63	14.75	8407500	AAL
1	2013-02-11	14.89	15.01	14.26	14.46	8882000	AAL
2	2013-02-12	14.45	14.51	14.10	14.27	8126000	AAL
3	2013-02-13	14.30	14.94	14.25	14.66	10259500	AAL
4	2013-02-14	14.94	14.96	13.16	13.99	31879900	AAL

date	AAL	AAPL	AAP	ABBV	ABC	ABT	ACN	ADBE	ADI	...	XLNX	XL	XOM	XRAY	XRX	XYL	YUM	ZBH	ZION	ZTS	
0	2013-02-08	14.75	67.8542	78.90	36.25	46.89	34.41	73.31	39.12	45.70	...	37.51	28.24	88.61	42.87	31.84	27.09	65.30	75.85	24.14	33.05
1	2013-02-11	14.46	68.5614	78.39	35.85	46.76	34.26	73.07	38.64	46.08	...	37.46	28.31	88.28	42.84	31.96	27.46	64.55	75.65	24.21	33.26
2	2013-02-12	14.27	66.8428	78.60	35.42	46.96	34.30	73.37	38.89	46.27	...	37.58	28.41	88.46	42.87	31.84	27.95	64.75	75.44	24.49	33.74
3	2013-02-13	14.66	66.7156	78.97	35.27	46.64	34.46	73.56	38.81	46.26	...	37.80	28.42	88.67	43.08	32.00	28.26	64.41	76.00	24.74	33.55

4	2013-02-14	13.99	66.6556	78.84	36.57	46.777	34.70	73.13	38.61	46.54	...	38.44	28.22	88.52	42.91	32.12	28.47	63.89	76.34	24.63	33.27
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1254	2018-02-01	53.88	167.7800	117.29	116.34	99.29	62.18	160.46	199.38	91.65	...	72.49	36.79	89.07	60.73	32.75	74.84	83.98	128.19	54.98	77.82
1255	2018-02-02	52.10	160.5000	113.93	115.17	96.02	61.69	156.90	195.64	89.11	...	70.64	38.25	84.53	60.06	31.63	75.66	82.63	125.79	54.15	76.78
1256	2018-02-05	49.76	156.4900	109.86	109.51	91.90	58.73	151.83	190.27	85.40	...	66.97	37.68	79.72	58.54	31.38	72.66	79.80	123.18	51.65	73.83
1257	2018-	51.18	163.0300	112.20	111.20	91.54	58.86	154.69	194.47	88.54	...	68.99	37.34	78.35	58.46	30.85	71.33	80.58	122.30	52.52	73.27

	2018-02-07	51.40	159.5400	109.93	113.62	94.22	58.67	155.15	192.34	85.35	...	66.97	42.00	76.94	58.30	31.18	71.79	80.13	120.78	54.02	73.86
1258																									

1259 rows × 506 columns

	Name	date	close
226	AAL	2014-01-02	25.360
227	AAL	2014-01-03	26.540
228	AAL	2014-01-06	27.030
229	AAL	2014-01-07	26.905
230	AAL	2014-01-08	27.630
...	...	...	...
619009	ZTS	2017-12-22	71.990
619010	ZTS	2017-12-26	72.340
619011	ZTS	2017-12-27	72.450
619012	ZTS	2017-12-28	72.390
619013	ZTS	2017-12-29	72.040

[480102 rows x 3 columns]

```
[['AAL' '2014-01-02' 25.36]
 ['AAL' '2014-01-03' 26.54]
 ['AAL' '2014-01-06' 27.03]
 ...
 ['ZTS' '2017-12-27' 72.45]
 ['ZTS' '2017-12-28' 72.39]
 ['ZTS' '2017-12-29' 72.04]]
```

```
 date     AAL     AAPL     AAP   ABBV    ABC    ABT     ACN  \
0    2014-01-02  25.360   79.0185  109.74  51.98  69.89  38.23   81.13
1    2014-01-03  26.540   77.2828  112.88  52.30  69.94  38.64   81.40
2    2014-01-06  27.030   77.7042  111.80  50.39  69.69  39.15   80.54
3    2014-01-07  26.905   77.1481  113.18  50.49  70.45  38.85   81.52
4    2014-01-08  27.630   77.6371  112.30  50.36  71.14  39.20   82.15
..         ...     ...       ...     ...    ...    ...    ...     ...
989  2017-12-22  52.590  175.0100  100.55  98.21  92.46  56.93  153.89
990  2017-12-26  52.850  170.5700  101.96  97.75  93.25  57.00  152.99
991  2017-12-27  52.400  170.6000   99.77  98.09  92.60  57.47  153.32
992  2017-12-28  52.460  171.0800   99.71  97.79  92.59  57.46  153.57
993  2017-12-29  52.030  169.2300   99.69  96.71  91.82  57.07  153.09

      ADBE    ADI  ...   XLNX    XL    XOM   XRAY    XRX    XYL    YUM  \
0    59.29  49.28  ...  45.97  31.25   99.75  47.96  47.64  34.16  75.09
1    59.16  49.61  ...  45.62  30.81   99.51  48.19  47.96  34.47  75.56
2    58.12  49.33  ...  45.42  30.37   99.66  47.90  48.36  34.41  75.50
3    58.97  49.59  ...  45.52  30.37  101.07  48.64  48.76  34.51  76.56
4    58.90  49.71  ...  45.91  30.39  100.74  48.73  48.32  34.49  76.53
..     ...    ...  ...    ...    ...     ...    ...    ...    ...    ...
989  175.00  88.85  ...  67.92  35.17   83.97  65.82  29.58  67.58  82.40
990  174.44  88.63  ...  67.60  35.24   83.98  66.26  29.41  67.50  82.19
991  175.36  89.10  ...  67.91  35.20   83.90  66.03  29.48  68.23  82.40
992  175.55  89.38  ...  68.50  35.37   84.02  66.43  29.45  68.25  82.67
993  175.24  89.03  ...  67.42  35.16   83.64  65.83  29.15  68.20  81.61

      ZBH   ZION    ZTS
0    92.24  29.65  32.36
1    92.64  29.86  32.05
2    93.24  29.65  31.98
3    95.10  29.74  32.10
4    97.43  30.00  31.74
..     ...    ...    ...
989  120.12  51.33  71.99
990  119.96  50.86  72.34
991  120.14  50.71  72.45
992  121.75  51.34  72.39
993  120.67  50.83  72.04

[994 rows x 484 columns]
```

(994, 484)
closing values

	A	AAL	AAP	AAPL	ABBV	ABC	\
2014-01-03	1.012631	1.046530	1.028613	0.978034	1.006156	1.000715	
2014-01-06	0.995081	1.018463	0.990432	1.005453	0.963480	0.996426	
2014-01-07	1.014301	0.995376	1.012343	0.992843	1.001985	1.010905	
2014-01-08	1.016362	1.026947	0.992225	1.006338	0.997425	1.009794	
2014-01-09	1.000343	1.064785	1.011131	0.987228	1.017077	1.003374	
...	...	...	...	...	...	...	
2017-12-22	0.997482	0.996211	1.004195	1.000000	1.003064	0.994300	
2017-12-26	0.998515	1.004944	1.014023	0.974630	0.995316	1.008544	
2017-12-27	1.000743	0.991485	0.978521	1.000176	1.003478	0.993029	
2017-12-28	1.002229	1.001145	0.999399	1.002814	0.996942	0.999892	
2017-12-29	0.992884	0.991803	0.999799	0.989186	0.988956	0.991684	

	ABT	ACN	ADBE	ADI	...	XL	XLNX	\
2014-01-03	1.010725	1.003328	0.997807	1.006696	...	0.985920	0.992386	
2014-01-06	1.013199	0.989435	0.982421	0.994356	...	0.985719	0.995616	
2014-01-07	0.992337	1.012168	1.014625	1.005271	...	1.000000	1.002202	
2014-01-08	1.009009	1.007728	0.998813	1.002420	...	1.000659	1.008568	
2014-01-09	1.001786	1.009738	1.003226	0.996178	...	1.000329	0.997168	
...	...	...	...	...	...	...	...	
2017-12-22	1.000000	0.997990	1.002521	1.002256	...	1.003137	0.993418	
2017-12-26	1.001230	0.994152	0.996800	0.997524	...	1.001990	0.995289	
2017-12-27	1.008246	1.002157	1.005274	1.005303	...	0.998865	1.004586	
2017-12-28	0.999826	1.001631	1.001083	1.003143	...	1.004830	1.008688	
2017-12-29	0.993213	0.996874	0.998234	0.996084	...	0.994063	0.984234	

	XOM	XRAY	XRX	XYL	YUM	ZBH	\
2014-01-03	0.997594	1.004796	1.006717	1.009075	1.006259	1.004337	
2014-01-06	1.001507	0.993982	1.008340	0.998259	0.999206	1.006477	
2014-01-07	1.014148	1.015449	1.008271	1.002906	1.014040	1.019949	
2014-01-08	0.996735	1.001850	0.990976	0.999420	0.999608	1.024501	
2014-01-09	0.990272	1.003694	0.997517	1.004059	0.980661	0.990147	
...	...	...	...	...	...	...	
2017-12-22	1.001431	1.004885	0.995624	0.997491	0.998788	1.001417	
2017-12-26	1.000119	1.006685	0.994253	0.998816	0.997451	0.998668	
2017-12-27	0.999047	0.996529	1.002380	1.010815	1.002555	1.001501	
2017-12-28	1.001430	1.006058	0.998982	1.000293	1.003277	1.013401	
2017-12-29	0.995477	0.990968	0.989813	0.999267	0.987178	0.991129	

	ZION	ZTS
2014-01-03	1.007083	0.990420
2014-01-06	0.992967	0.997816
2014-01-07	1.003035	1.003752
2014-01-08	1.008742	0.988785
2014-01-09	1.007333	1.006931
...	...	...
2017-12-22	0.997474	0.995988

2017-12-26 0.990844 1.004862
2017-12-27 0.997051 1.001521
2017-12-28 1.012424 0.999172
2017-12-29 0.990066 0.995165

[993 rows x 483 columns]

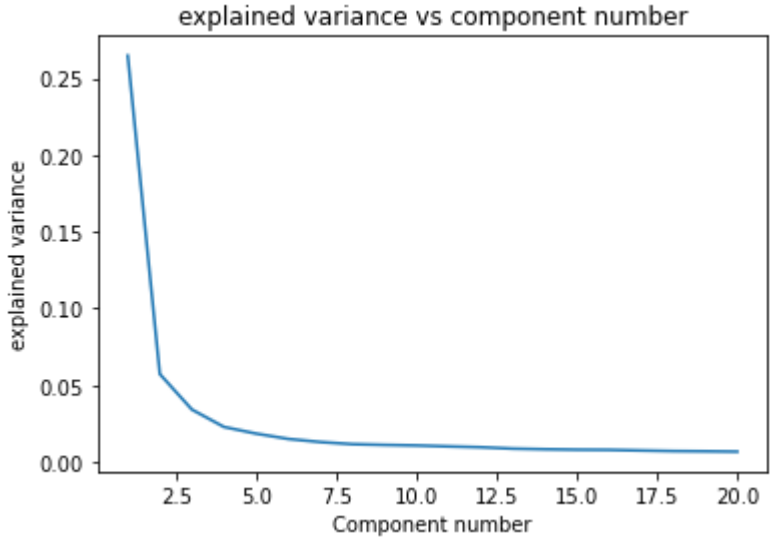
ncipal componenets

	0	1	2	3	4	5	6	\
0	-0.021483	-0.031666	0.025927	0.026893	0.001018	0.005838	-0.020973	
1	0.087005	0.020140	-0.010365	-0.023034	-0.022075	0.017957	-0.036388	
2	-0.138037	-0.022034	-0.043514	-0.055666	0.005071	0.017548	0.032574	
3	-0.034844	-0.040742	0.088260	-0.090818	0.032866	0.017027	-0.018246	
4	-0.040344	-0.061981	0.012722	0.043641	0.050495	0.061980	-0.022384	
..	...	...	...	...	...	...	...	
988	0.000283	0.018264	-0.055224	0.030221	0.005921	0.005422	-0.005312	
989	-0.029212	0.080591	-0.033625	0.062237	0.086519	-0.017535	0.000986	
990	0.018784	-0.040848	-0.011801	-0.043092	-0.048400	0.026628	0.003240	
991	-0.046583	0.023081	-0.023947	-0.001904	-0.015394	0.013047	-0.026996	
992	0.129948	0.020990	-0.015985	0.011557	-0.015644	0.017888	-0.007506	

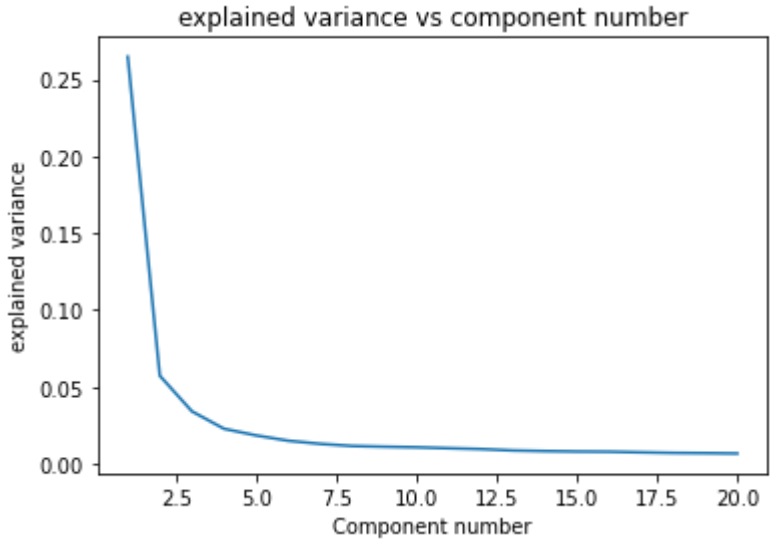
	7	8	9	...	473	474	475	\
0	0.022395	-0.013605	-0.013686	...	0.000903	0.001968	0.000053	
1	0.027984	-0.019011	0.023106	...	-0.000005	-0.000116	-0.001070	
2	0.016575	-0.016842	0.002946	...	-0.006188	0.001798	-0.001630	
3	0.010623	-0.003406	-0.028702	...	-0.000547	0.004850	0.000361	
4	0.016281	0.000676	-0.053828	...	0.001236	0.000087	-0.002313	
..	...	...	...	...	...	...	...	
988	-0.008588	0.003565	-0.017590	...	0.004748	0.004206	0.002363	
989	-0.020608	-0.000704	0.020202	...	0.003180	0.001026	-0.000315	
990	-0.008693	0.003176	-0.004485	...	0.000532	-0.000420	0.004172	
991	0.002249	-0.018126	-0.012762	...	0.004337	0.003022	-0.007049	
992	-0.021501	0.008250	-0.017794	...	-0.001750	0.003208	-0.000234	

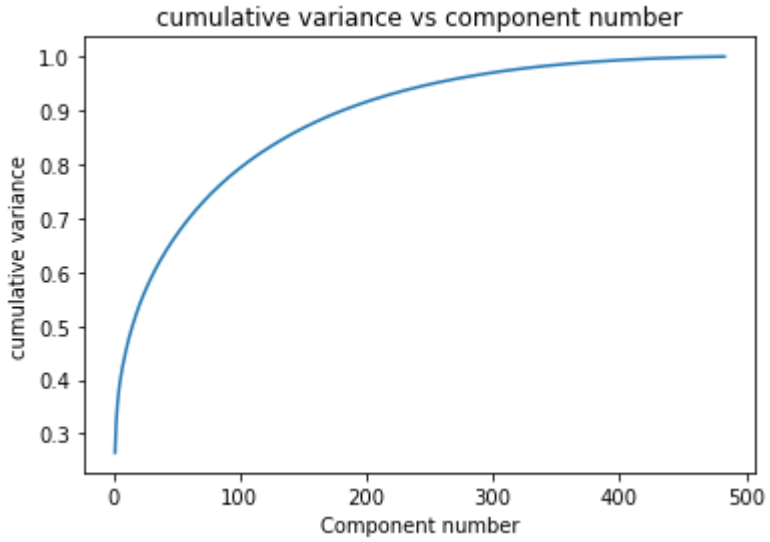
	476	477	478	479	480	481	482
0	-0.003230	0.001099	0.004300	-0.001307	0.000333	0.000494	-0.000301
1	-0.001131	-0.001388	-0.001915	-0.002788	-0.001083	-0.000038	0.001016
2	-0.000252	0.002106	-0.002996	-0.005596	-0.001928	0.001446	0.000153
3	-0.003724	0.000029	-0.001700	-0.003681	-0.000290	0.001052	0.002019
4	0.002415	-0.000581	0.002913	-0.004018	0.001961	0.002754	-0.003115

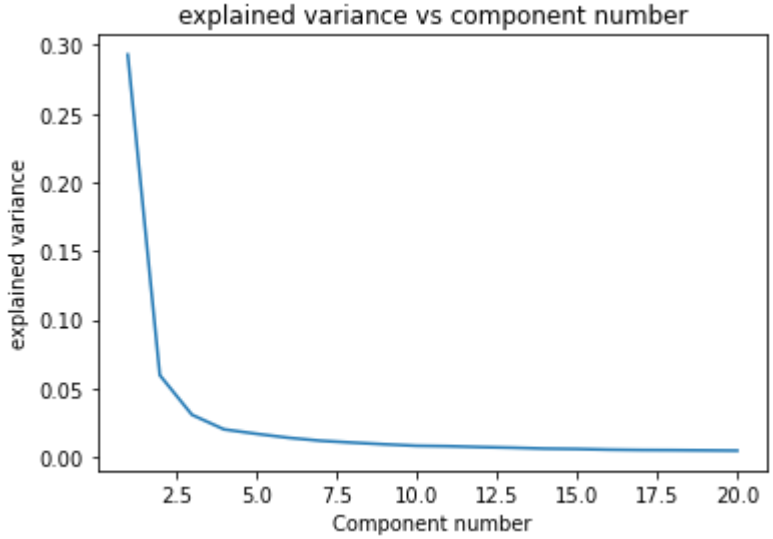
```
..       ...       ...       ...       ...       ...       ...       ...
988  0.002084 -0.000515  0.001713 -0.002017 -0.000891 -0.000106  0.000699
989 -0.000747  0.000707  0.001151  0.001408 -0.001897  0.000306  0.002511
990  0.000866 -0.003666 -0.002408  0.001042  0.001282 -0.003492  0.000223
991 -0.003281 -0.000775  0.001438  0.000477 -0.001191  0.002681  0.000712
992 -0.000372  0.002453 -0.000241  0.001755 -0.001062  0.001784 -0.001069

[993 rows x 483 columns]
plotting graph
```

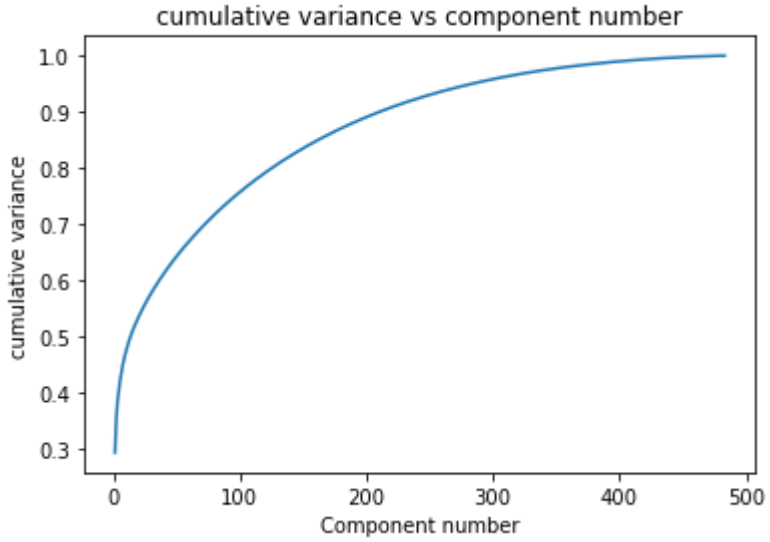



```
[0.26488994 0.05716809 0.03401231 0.02266639 0.01832904 0.0148591
 0.01282072 0.01144004 0.01095745 0.01057907 0.00996888 0.00943521
 0.00852503 0.00811363 0.00782795 0.00770922 0.00730028 0.00693095
 0.00677864 0.00655948]
```




cumulative variance vs component number
cumulative variance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
100
200
300
400
500
Component number


explained variance vs component number
explained variance
0.30
0.25
0.20
0.15
0.10
0.05
0.00
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Component number


cumulative variance vs component number
cumulative variance
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
100
200
300
400
500
Component number

Out[45]:

	A	AAL	AAP	AAPL	ABBV	ABC	ABT	ACN	ADBE	ADI	...	XL	XLNX	XOM	XRAY	XRX	XYL	YUM	ZBH	ZION	ZTS
2014-01-03	0.77	2.00	1.45	-1.59	0.32	0.02	0.83	0.22	-0.22	0.38	...	-1.19	-0.52	-0.20	0.37	0.40	0.64	0.34	0.29	0.38	-0.73
	7079	0207	7375	4400	2936	1822	5662	0664	9470	4758		7830	7581	2633	1563	1182	1829	7526	0258	4577	2962
2014-01-06	-0.32	0.76	-0.49	0.31	-2.23	-0.26	1.03	-0.97	-1.24	-0.40	...	-1.21	-0.31	0.14	-0.54	0.49	-0.19	-0.06	0.44	-0.46	-0.21
	9776	7835	3882	9994	3291	3844	7472	3537	3657	9185		4709	7714	3261	0314	3527	5008	0180	6658	3510	6120
2014-01-07	0.88	-0.24	0.62	-0.56	0.07	0.70	-0.66	0.98	0.88	0.29	...	-0.01	0.11	1.26	1.26	0.48	0.16	0.79	1.43	0.14	0.19
	2379	6281	5904	0407	3060	8003	1941	9936	8430	2746		5946	0239	3805	2099	9602	4528	7276	5711	1411	8746
2014-01-08	1.01	1.14	-0.40	0.38	-0.20	0.62	0.69	0.60	-0.15	0.10	...	0.03	0.52	-0.27	0.12	-0.49	-0.10	-0.03	1.76	0.48	-0.84
	2374	0497	7722	1839	3300	6381	5732	6447	8396	1796		9333	3914	6190	3198	4269	5169	6930	3810	4302	7239
2014	0.00	2.80	0.56	-0.95	0.97	0.19	0.10	0.78	0.13	-0.29	...	0.01	-0.21	-0.84	0.27	-0.12	0.25	-1.13	-0.74	0.39	0.42
	2066	2547	3935	2497	7073	8834	6558	0061	5737	7319		7516	6835	0347	8625	2225	3741	2145	6704	9640	0907

-01-09																				
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2017-12-22	-0.17	-0.20	0.20	-0.06	0.13	-0.40	-0.03	-0.23	0.08	0.09	...	0.24	-0.46	0.13	0.37	-0.22	-0.25	-0.08	0.07	-0.19	-0.34
2017-12-26	-0.11	0.17	0.71	-1.83	-0.32	0.54	0.06	-0.56	-0.29	-0.20	...	0.15	-0.33	0.02	0.53	-0.30	-0.15	-0.16	-0.12	-0.59	0.27
2017-12-27	0.02	-0.41	-1.10	-0.04	0.16	-0.48	0.63	0.12	0.26	0.29	...	-0.11	0.26	-0.07	-0.32	0.15	0.77	0.13	0.08	-0.21	0.04
2017-12-28	0.12	0.00	-0.03	0.13	-0.22	-0.03	-0.05	0.07	-0.00	0.15	...	0.38	0.53	0.13	0.47	-0.03	-0.03	0.17	0.95	0.70	-0.12

2017-12-29	-0.46	-0.40	-0.01	-0.81	-0.70	-0.57	-0.59	-0.33	-0.19	-0.29	...	-0.51	-1.05	-0.38	-0.79	-0.56	-0.11	-0.75	-0.67	-0.63	-0.40
	-0.83	-0.31	-0.51	-0.57	-0.73	-0.95	-0.27	-0.09	-0.67	-0.80		-0.43	-0.73	-0.97	-0.44	-0.04	-0.70	-0.54	-0.48	-0.78	-0.13
	-0.48	-0.92	-0.69	-0.46	-0.28	-0.98	-0.08	-0.90	-0.18	-0.01		-0.21	-0.63	-0.23	-0.95	-0.58	-0.12	-0.49	-0.98	-0.06	-0.73

993 rows × 483 columns

In [ ]: