Importing Useful Libraries and Reading Data from CSV

```
In [33]: import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         from sklearn.preprocessing import StandardScaler
         from scipy.linalg import eigh
         from sklearn import decomposition
```

Data Loading and Understanding

```
In [7]: df=pd.read_csv("Data/secom.data",delimiter=" ",header=None)
```

```
In [8]: df.head()
```

Out[8]:

	0	1	2	3	4	5	6	7	8	
0	3030.93	2564.00	2187.7333	1411.1265	1.3602	100.0	97.6133	0.1242	1.5005	0.0162
1	3095.78	2465.14	2230.4222	1463.6606	0.8294	100.0	102.3433	0.1247	1.4966	-0.0005
2	2932.61	2559.94	2186.4111	1698.0172	1.5102	100.0	95.4878	0.1241	1.4436	0.0041
3	2988.72	2479.90	2199.0333	909.7926	1.3204	100.0	104.2367	0.1217	1.4882	-0.0124
4	3032.24	2502.87	2233.3667	1326.5200	1.5334	100.0	100.3967	0.1235	1.5031	-0.0037

5 rows × 590 columns

```
In [9]: df.describe()
```

Out[9]:

	0	1	2	3	4	5	
count	1561.000000	1560.000000	1553.000000	1553.000000	1553.000000	1553.0	1553.0(
mean	3014.452896	2495.850231	2200.547318	1396.376627	4.197013	100.0	101.112
std	73.621787	80.407705	29.513152	441.691640	56.355540	0.0	6.2372
min	2743.240000	2158.750000	2060.660000	0.000000	0.681500	100.0	82.131
25%	2966.260000	2452.247500	2181.044400	1081.875800	1.017700	100.0	97.920(
50%	3011.490000	2499.405000	2201.066700	1285.214400	1.316800	100.0	101.51:
75%	3056.650000	2538.822500	2218.055500	1591.223500	1.525700	100.0	104.58(
max	3356.350000	2846.440000	2315.266700	3715.041700	1114.536600	100.0	129.25:

8 rows × 590 columns

Data Insights:

- We can see there are NULL values in the data. So, we can use respective imputation methods to replace them
- If we feel the respective columns are not useful to complete task, we can remove them

Data Preprocessing

Imputation : NULL values are replaced with median imputation method

```
In [11]: impt_df=df.fillna(df.median())
```

In [12]: `impt_df.describe()`

Out[12]:

	0	1	2	3	4	5	
count	1567.000000	1567.000000	1567.000000	1567.000000	1567.000000	1567.0	1567.0(
mean	3014.441551	2495.866110	2200.551958	1395.383474	4.171281	100.0	101.116
std	73.480841	80.228143	29.380973	439.837330	56.103721	0.0	6.2093{
min	2743.240000	2158.750000	2060.660000	0.000000	0.681500	100.0	82.131
25%	2966.665000	2452.885000	2181.099950	1083.885800	1.017700	100.0	97.937{
50%	3011.490000	2499.405000	2201.066700	1285.214400	1.316800	100.0	101.51:
75%	3056.540000	2538.745000	2218.055500	1590.169900	1.518800	100.0	104.53(
max	3356.350000	2846.440000	2315.266700	3715.041700	1114.536600	100.0	129.25:

8 rows × 590 columns

Columns with only one unique value after imputation mean they do not contribute anything to our signal, we can safely ignore them

In [13]: `cols = impt_df.nunique()==1`

In [18]: `prun_data=impt_df.drop(list(impt_df.columns[cols]),axis=1)`

In [19]: `prun_data.shape`

Out[19]: `(1567, 474)`

Principle Component Analysis

Step1 : Standardization

In [22]:
```
std_data=StandardScaler().fit_transform(prun_data)
std_data.shape
```

Out[22]: `(1567, 474)`

Step2 : Covariance Matrix

X.X^T

```
In [23]: covar_matrix = np.matmul(std_data.T,std_data)
```

Udacity Feedback

No tests loaded

Hide Test Results

```
In [25]: covar_matrix.shape
```

```
Out[25]: (474, 474)
```

Step3 : Finding Eigen Values and Eigen Vectors

```
In [30]: ei_values,ei_vectors = eigh(covar_matrix)
```

```
In [31]: ei_vectors.shape
```

```
Out[31]: (474, 474)
```

```
In [32]:  ei_values
```

```
Out[32]: array([-3.19699635e-13, -1.23832267e-13, -5.32527121e-14, -3.06980482e-16,
                 4.70483139e-13,  2.94765182e-0                     284434e-04,
                 4.14830248e-04,  4.90169224e-04,       575e-04,  1.08672745e-03,
                 2.55230422e-03,  3.71502474e-03,  3.97943660e-03,  5.47744555e-03,
                 5.67069581e-03,  6.18012341e-03,  6.75068619e-03,  7.13477913e-03,
                 8.08436908e-03,  1.12135379e-02,  1.33919741e-02,  1.42546723e-02,
                 1.52665284e-02,  1.64840364e-02,  2.06802053e-02,  2.43945360e-02,
                 2.64216193e-02,  3.00232648e-02,  3.28949102e-02,  3.54097527e-02,
                 4.28791454e-02,  4.80209785e-02,  5.14296888e-02,  7.40364404e-02,
                 7.78820006e-02,  8.72572776e-02,  9.35095547e-02,  1.04211154e-01,
                 1.06873660e-01,  1.27039159e-01,  1.31054631e-01,  1.35328503e-01,
                 1.38382097e-01,  1.48672855e-01,  1.63268517e-01,  1.71126072e-01,
                 1.79155996e-01,  1.81481406e-01,  1.96681819e-01,  2.14337823e-01,
                 2.24625493e-01,  2.58788682e-01,  2.71803040e-01,  2.86571094e-01,
                 2.94291554e-01,  3.39726182e-01,  3.42768567e-01,  3.61098902e-01,
                 3.76792672e-01,  3.96494031e-01,  4.41745317e-01,  4.77610946e-01,
                 4.87615370e-01,  5.52788358e-01,  5.76609249e-01,  6.08846267e-01,
                 6.61908092e-01,  6.75621582e-01,  7.58663112e-01,  8.14248679e-01,
                 8.29838535e-01,  8.47849747e-01,  9.77476530e-01,  9.96601945e-01,
                 1.02699994e+00,  1.04795777e+00,  1.07165805e+00,  1.15620249e+00,
                 1.19376466e+00,  1.20755410e+00,  1.25885423e+00,  1.33925875e+00,
                 1.35632322e+00,  1.40399993e+00,  1.50005471e+00,  1.54140390e+00,
                 1.61989748e+00,  1.70199760e+00,  1.79560529e+00,  1.81358600e+00,
                 1.99380794e+00,  2.10951055e+00,  2.12930189e+00,  2.26397642e+00,
                 2.27895363e+00,  2.52286107e+00,  2.66833676e+00,  2.75724044e+00,
                 2.86665248e+00,  2.95324948e+00,  3.06081963e+00,  3.19199501e+00,
                 3.36570683e+00,  3.56620734e+00,  3.73026360e+00,  3.83114623e+00,
                 4.12725334e+00,  4.25516016e+00,  4.45133875e+00,  4.81608085e+00,
                 4.85287258e+00,  4.88884174e+00,  5.17103142e+00,  5.44584296e+00,
                 5.54778113e+00,  5.84795939e+00,  6.16439519e+00,  6.34725982e+00,
                 6.61685170e+00,  6.81110112e+00,  7.39981306e+00,  7.60438094e+00,
                 8.06273659e+00,  8.31278239e+00,  8.56868702e+00,  8.96890928e+00,
                 9.16406162e+00,  9.51541550e+00,  9.82459967e+00,  9.87692914e+00,
                 1.01433501e+01,  1.06790191e+01,  1.08970628e+01,  1.09712429e+01,
                 1.13451933e+01,  1.14287206e+01,  1.18229703e+01,  1.23179616e+01,
                 1.26551097e+01,  1.29047496e+01,  1.34593140e+01,  1.42227937e+01,
                 1.46253875e+01,  1.49938894e+01,  1.51050705e+01,  1.54676642e+01,
                 1.62973184e+01,  1.68831507e+01,  1.79104429e+01,  1.84324447e+01,
                 1.86660196e+01,  1.87849861e+01,  1.94423660e+01,  1.95435284e+01,
                 2.00871781e+01,  2.05762492e+01,  2.14803392e+01,  2.27702273e+01,
                 2.28616894e+01,  2.35587481e+01,  2.37214810e+01,  2.44554193e+01,
                 2.48026907e+01,  2.58626850e+01,  2.69076371e+01,  2.72635888e+01,
                 2.74125930e+01,  2.87048059e+01,  2.93249163e+01,  2.97314102e+01,
                 3.09822118e+01,  3.13109994e+01,  3.20351023e+01,  3.28416124e+01,
                 3.34560118e+01,  3.36377072e+01,  3.47046482e+01,  3.55301169e+01,
                 3.63593493e+01,  3.73547187e+01,  3.75244744e+01,  3.81955521e+01,
                 3.87995932e+01,  3.97855840e+01,  4.01376768e+01,  4.24989505e+01,
                 4.26685870e+01,  4.32474565e+01,  4.38973120e+01,  4.44815754e+01,
                 4.51378254e+01,  4.66315221e+01,  4.74353524e+01,  4.95299037e+01,
                 4.99377177e+01,  5.13641249e+01,  5.21278152e+01,  5.35500131e+01,
                 5.45930094e+01,  5.47155275e+01,  5.82259310e+01,  5.84905658e+01,
                 5.91222724e+01,  6.03703803e+01,  6.16784265e+01,  6.29535088e+01,
                 6.44512781e+01,  6.61323614e+01,  6.74786375e+01,  6.98449874e+01,
                 7.18976600e+01,  7.48836174e+01,  7.74445288e+01,  7.80537930e+01,
                 8.06476951e+01,  8.14171464e+01,  8.47745928e+01,  8.62061680e+01,
                 9.06357657e+01,  9.22651141e+01,  9.58483494e+01,  9.67350254e+01,
                 9.87241009e+01,  1.05437268e+02,  1.07549990e+02,  1.12389144e+02,
```

```
1.18538124e+02,   1.22240437e+02,   1.23416835e+02,   1.28171589e+02,
1.34115940e+02,   1.41095971e+02,   1.43598974e+02,   1.48030478e+02,
1.51723948e+02,   1.55446697e+02,   1.59970591e+02,   1.66407546e+02,
1.74133377e+02,   1.80981469e+02,   1.84959422e+02,   1.90728492e+02,
1.98668866e+02,   2.14967444e+02,   2.18098178e+02,   2.20744933e+02,
2.38924999e+02,   2.47491410e+02,   2.54553273e+02,   2.60395616e+02,
2.66945722e+02,   2.79249529e+02,   2.84782502e+02,   2.96708539e+02,
3.08751098e+02,   3.14387055e+02,   3.22026356e+02,   3.32542342e+02,
3.45081619e+02,   3.46883495e+02,   3.56267634e+02,   3.66444552e+02,
3.81476568e+02,   3.86268880e+02,   3.99332039e+02,   4.00690953e+02,
4.18221353e+02,   4.34171884e+02,   4.45874461e+02,   4.58783423e+02,
4.66575212e+02,   4.81838908e+02,   4.95499770e+02,   5.06616570e+02,
5.14735821e+02,   5.26024932e+02,   5.34009345e+02,   5.44459258e+02,
5.60148502e+02,   5.71119727e+02,   5.74266375e+02,   5.86029030e+02,
5.99210037e+02,   6.09343969e+02,   6.31918267e+02,   6.37383691e+02,
6.49432979e+02,   6.72720527e+02,   6.78409602e+02,   6.85072849e+02,
6.93944693e+02,   7.05317735e+02,   7.23040149e+02,   7.28585197e+02,
7.48084061e+02,   7.50330866e+02,   7.77222897e+02,   7.78506432e+02,
7.83773914e+02,   7.95486676e+02,   8.12828483e+02,   8.15644503e+02,
8.36875115e+02,   8.64240154e+02,   8.67703714e+02,   8.79586515e+02,
8.84710446e+02,   8.90477332e+02,   9.12961586e+02,   9.20398000e+02,
9.29109184e+02,   9.37127913e+02,   9.52382759e+02,   9.58316890e+02,
9.65574596e+02,   9.79924204e+02,   9.96424880e+02,   1.01608801e+03,
1.02822499e+03,   1.03907971e+03,   1.05644746e+03,   1.07562220e+03,
1.07891748e+03,   1.09191656e+03,   1.09654544e+03,   1.11933553e+03,
1.13367246e+03,   1.14306049e+03,   1.16412054e+03,   1.17310880e+03,
1.18245722e+03,   1.20220429e+03,   1.21842956e+03,   1.22580991e+03,
1.24253809e+03,   1.25395781e+03,   1.26953486e+03,   1.28284101e+03,
1.30453105e+03,   1.32023665e+03,   1.33061663e+03,   1.33728244e+03,
1.35718826e+03,   1.36673471e+03,   1.37716643e+03,   1.38565619e+03,
1.40159341e+03,   1.40759618e+03,   1.43515802e+03,   1.45147386e+03,
1.46593939e+03,   1.48208831e+03,   1.48980161e+03,   1.50619197e+03,
1.53529141e+03,   1.55073453e+03,   1.56278419e+03,   1.56845764e+03,
1.58316295e+03,   1.61726888e+03,   1.65284643e+03,   1.67076833e+03,
1.68740720e+03,   1.73051605e+03,   1.75683287e+03,   1.76640227e+03,
1.79094350e+03,   1.80579921e+03,   1.82008927e+03,   1.88844342e+03,
1.90507519e+03,   1.91315723e+03,   1.94475726e+03,   1.96436173e+03,
1.98634447e+03,   2.06367279e+03,   2.07048558e+03,   2.09589902e+03,
2.13352798e+03,   2.16654015e+03,   2.22819109e+03,   2.28977724e+03,
2.31142943e+03,   2.34911148e+03,   2.43622383e+03,   2.46244450e+03,
2.48644763e+03,   2.55037788e+03,   2.60313741e+03,   2.66625688e+03,
2.68571616e+03,   2.70932278e+03,   2.76067482e+03,   2.79577704e+03,
2.84004986e+03,   2.97107238e+03,   3.04608236e+03,   3.11988452e+03,
3.15017457e+03,   3.17068199e+03,   3.20043796e+03,   3.26728762e+03,
3.29968238e+03,   3.37328005e+03,   3.40675991e+03,   3.44077766e+03,
3.49402151e+03,   3.56155950e+03,   3.62825701e+03,   3.67597021e+03,
3.76955354e+03,   3.80981671e+03,   3.89779855e+03,   3.97145659e+03,
4.01135425e+03,   4.11016556e+03,   4.15302031e+03,   4.16167627e+03,
4.21835885e+03,   4.35489623e+03,   4.44848369e+03,   4.46962936e+03,
4.50752914e+03,   4.62047350e+03,   4.65393736e+03,   4.75591378e+03,
4.88271981e+03,   4.94303678e+03,   4.97044552e+03,   5.02476769e+03,
5.16696396e+03,   5.28940650e+03,   5.34171406e+03,   5.41132393e+03,
5.55828739e+03,   5.60452329e+03,   5.70797932e+03,   5.74641048e+03,
5.91060796e+03,   6.00412055e+03,   6.08321649e+03,   6.18995623e+03,
6.28467364e+03,   6.46594391e+03,   6.60378940e+03,   6.86980365e+03,
7.00799810e+03,   7.10535897e+03,   7.17679696e+03,   7.37888207e+03,
7.49987039e+03,   7.66920317e+03,   7.88821214e+03,   8.32165253e+03,
```

```
       8.45772687e+03,  8.56289390e+03,  8.86809112e+03,  9.38594567e+03,
       9.49150009e+03,  9.74201192e+03,  9.89149854e+03,  1.08501639e+04,
       1.20452264e+04,  1.33675702e+04,  1.36272520e+04,  1.46659280e+04,
       1.54265000e+04,  1.62706168e+04,  1.88973434e+04,  2.09723748e+04,
       2.70516926e+04,  4.12850355e+04])
```

In [34]: `pca = decomposition.PCA(std_data)`

In [35]: `pca.n_components = 474`
          `pca_data = pca.fit_transform(std_data)`

In [37]: `pca.explained_variance_`

```
Out[37]: array([2.63633688e+01, 1.72743886e+01, 1.33923210e+01, 1.20672691e+01,
        1.03899213e+01, 9.85089401e+00, 9.56321356e+00, 8.78184690e+00,
        8.53612401e+00, 7.69171544e+00, 6.92858488e+00, 6.31641050e+00,
        6.22095270e+00, 6.06098346e+00, 5.99357961e+00, 5.66289344e+00,
        5.46800377e+00, 5.40084730e+00, 5.31395436e+00, 5.03717250e+00,
        4.89732003e+00, 4.78918927e+00, 4.71192980e+00, 4.58288439e+00,
        4.53726627e+00, 4.47509457e+00, 4.38684780e+00, 4.21697918e+00,
        4.12895524e+00, 4.01320156e+00, 3.95271790e+00, 3.88455715e+00,
        3.83404888e+00, 3.77433459e+00, 3.66948307e+00, 3.64494209e+00,
        3.57887822e+00, 3.54935338e+00, 3.45550698e+00, 3.41105623e+00,
        3.37765422e+00, 3.29946613e+00, 3.20866391e+00, 3.17397543e+00,
        3.15647304e+00, 3.11795646e+00, 3.03698198e+00, 2.97186294e+00,
        2.95049394e+00, 2.87837110e+00, 2.85416945e+00, 2.84066647e+00,
        2.78090436e+00, 2.69371574e+00, 2.65751997e+00, 2.65199254e+00,
        2.62462680e+00, 2.56152890e+00, 2.53605147e+00, 2.48901568e+00,
        2.43283315e+00, 2.40712231e+00, 2.34736284e+00, 2.31689465e+00,
        2.27430364e+00, 2.23117593e+00, 2.19717603e+00, 2.17545333e+00,
        2.15407410e+00, 2.10707687e+00, 2.08639056e+00, 2.04370240e+00,
        2.02470114e+00, 2.01160573e+00, 1.99226342e+00, 1.94513561e+00,
        1.89723651e+00, 1.81356952e+00, 1.78529824e+00, 1.76288303e+00,
        1.73009117e+00, 1.71501671e+00, 1.70259060e+00, 1.66228443e+00,
        1.62859379e+00, 1.58776988e+00, 1.57244221e+00, 1.55569849e+00,
        1.50007119e+00, 1.47600857e+00, 1.46218215e+00, 1.42285510e+00,
        1.38348669e+00, 1.36240612e+00, 1.33837741e+00, 1.32214916e+00,
        1.31779872e+00, 1.26841920e+00, 1.25438169e+00, 1.24186287e+00,
        1.22168406e+00, 1.21652311e+00, 1.20590257e+00, 1.16225369e+00,
        1.15312848e+00, 1.14364208e+00, 1.12797080e+00, 1.12186007e+00,
        1.10505495e+00, 1.07752695e+00, 1.06690187e+00, 1.05545749e+00,
        1.03273875e+00, 1.01095974e+00, 1.00156938e+00, 9.97946484e-01,
        9.90251938e-01, 9.80390427e-01, 9.61808411e-01, 9.51342022e-01,
        9.46416545e-01, 9.36104335e-01, 9.26867089e-01, 9.16448288e-01,
        8.98848137e-01, 8.95014948e-01, 8.84837924e-01, 8.79416625e-01,
        8.72755244e-01, 8.66659169e-01, 8.53947916e-01, 8.49691333e-01,
        8.43062996e-01, 8.33033874e-01, 8.19183275e-01, 8.10686375e-01,
        8.00739345e-01, 7.93447057e-01, 7.82764948e-01, 7.78052080e-01,
        7.67691118e-01, 7.55081239e-01, 7.49111620e-01, 7.43371992e-01,
        7.29923681e-01, 7.23928776e-01, 7.14773647e-01, 7.00220588e-01,
        6.97264727e-01, 6.88963908e-01, 6.86859645e-01, 6.74615237e-01,
        6.63524721e-01, 6.56593225e-01, 6.48842920e-01, 6.36286641e-01,
        6.25749811e-01, 6.16586588e-01, 6.11952037e-01, 6.08162681e-01,
        5.98421400e-01, 5.93300884e-01, 5.87738186e-01, 5.82989518e-01,
        5.68631758e-01, 5.64949199e-01, 5.61677213e-01, 5.54089217e-01,
        5.51877493e-01, 5.34403011e-01, 5.20845788e-01, 5.19047563e-01,
        5.07973611e-01, 5.00494198e-01, 4.97130544e-01, 4.96310917e-01,
        4.79138484e-01, 4.77703743e-01, 4.65252361e-01, 4.61711462e-01,
        4.50394467e-01, 4.43131988e-01, 4.37466698e-01, 4.33211751e-01,
        4.29578880e-01, 4.14708160e-01, 4.07013851e-01, 4.03523797e-01,
        3.89108537e-01, 3.82637316e-01, 3.74220326e-01, 3.66709052e-01,
        3.64699698e-01, 3.57693807e-01, 3.47675133e-01, 3.41002136e-01,
        3.35903532e-01, 3.28694649e-01, 3.23509943e-01, 3.16411092e-01,
        3.07687681e-01, 2.97940748e-01, 2.92965149e-01, 2.84721878e-01,
        2.77248968e-01, 2.67063444e-01, 2.55869063e-01, 2.55001302e-01,
        2.46659566e-01, 2.43599341e-01, 2.34000352e-01, 2.27501682e-01,
        2.21509256e-01, 2.20358633e-01, 2.12351432e-01, 2.05636243e-01,
        2.00758017e-01, 1.97159066e-01, 1.89469054e-01, 1.81853449e-01,
        1.78320261e-01, 1.70463424e-01, 1.66280726e-01, 1.62549982e-01,
        1.58040492e-01, 1.52570242e-01, 1.40961004e-01, 1.39270867e-01,
```

```
1.37271676e-01, 1.26863899e-01, 1.21793418e-01, 1.18109465e-01,
1.15569265e-01, 1.11196282e-01, 1.06268301e-01, 1.02152657e-01,
9.92635358e-02, 9.68863013e-02, 9.51663331e-02, 9.32055388e-02,
9.00995983e-02, 8.56423630e-02, 8.18464811e-02, 7.88102394e-02,
7.80590276e-02, 7.56948431e-02, 7.17682911e-02, 6.86781547e-02,
6.73290344e-02, 6.30422100e-02, 6.17720469e-02, 6.12058425e-02,
5.89176974e-02, 5.78772450e-02, 5.50486386e-02, 5.41344781e-02,
5.19905150e-02, 5.14991668e-02, 4.98427797e-02, 4.94537221e-02,
4.78184019e-02, 4.59116603e-02, 4.46008859e-02, 4.30898068e-02,
4.22301158e-02, 4.11566271e-02, 4.02001972e-02, 3.93859684e-02,
3.85506898e-02, 3.77536861e-02, 3.73502975e-02, 3.71813097e-02,
3.49396727e-02, 3.48614364e-02, 3.41954106e-02, 3.32872383e-02,
3.27995689e-02, 3.18887086e-02, 3.16282910e-02, 3.02907742e-02,
2.97774726e-02, 2.88236433e-02, 2.84045820e-02, 2.80314892e-02,
2.76165112e-02, 2.72468627e-02, 2.71385380e-02, 2.56307004e-02,
2.54058646e-02, 2.47762408e-02, 2.43905186e-02, 2.39619888e-02,
2.38535879e-02, 2.32179753e-02, 2.26884527e-02, 2.21613335e-02,
2.14800173e-02, 2.13639922e-02, 2.09716554e-02, 2.04566426e-02,
1.99942525e-02, 1.97842987e-02, 1.89855749e-02, 1.87260002e-02,
1.83300165e-02, 1.75048487e-02, 1.74096991e-02, 1.71823992e-02,
1.65151245e-02, 1.58382444e-02, 1.56164874e-02, 1.51478168e-02,
1.50439004e-02, 1.45987799e-02, 1.45403750e-02, 1.37166917e-02,
1.31393673e-02, 1.28270613e-02, 1.24799032e-02, 1.24153039e-02,
1.19955211e-02, 1.19195527e-02, 1.17703989e-02, 1.14370644e-02,
1.07810669e-02, 1.04069721e-02, 9.87718022e-03, 9.64563891e-03,
9.57464202e-03, 9.33932788e-03, 9.08224374e-03, 8.59470884e-03,
8.24058087e-03, 8.08116839e-03, 7.86587583e-03, 7.54978947e-03,
7.29803359e-03, 7.24469561e-03, 7.00590222e-03, 6.95853311e-03,
6.81929702e-03, 6.47723504e-03, 6.30710673e-03, 6.27369072e-03,
6.07625511e-03, 5.85189120e-03, 5.72727285e-03, 5.47170308e-03,
5.30829016e-03, 5.14861851e-03, 4.85592653e-03, 4.72529569e-03,
4.34936214e-03, 4.22532037e-03, 4.05316719e-03, 3.93639539e-03,
3.73432911e-03, 3.54264440e-03, 3.47754978e-03, 3.30206348e-03,
3.12186573e-03, 3.09889692e-03, 3.07540284e-03, 2.84248962e-03,
2.71721594e-03, 2.63553853e-03, 2.44645353e-03, 2.38203295e-03,
2.27727161e-03, 2.14923808e-03, 2.03831099e-03, 1.95454638e-03,
1.88585535e-03, 1.83055714e-03, 1.76068993e-03, 1.70391875e-03,
1.61102239e-03, 1.45527051e-03, 1.44570653e-03, 1.35970747e-03,
1.34706932e-03, 1.27318514e-03, 1.15810090e-03, 1.14661896e-03,
1.08684393e-03, 1.03441729e-03, 9.84293678e-04, 9.57889339e-04,
8.96551677e-04, 8.66106784e-04, 8.55209931e-04, 8.03866050e-04,
7.71107344e-04, 7.62301823e-04, 7.38315765e-04, 6.84328257e-04,
6.69193980e-04, 6.55810943e-04, 6.36399709e-04, 6.24186801e-04,
5.41411077e-04, 5.29909665e-04, 5.19954457e-04, 4.84459203e-04,
4.31431406e-04, 4.22674388e-04, 3.88790720e-04, 3.68205140e-04,
3.52993843e-04, 3.11376354e-04, 3.04987833e-04, 2.82085132e-04,
2.53189036e-04, 2.40608348e-04, 2.30586783e-04, 2.18881588e-04,
2.16938813e-04, 1.87925641e-04, 1.82995590e-04, 1.73565160e-04,
1.65254586e-04, 1.43439012e-04, 1.36869619e-04, 1.25595031e-04,
1.15888510e-04, 1.14403574e-04, 1.09275908e-04, 1.04258312e-04,
9.49379664e-05, 8.83666012e-05, 8.64166689e-05, 8.36875035e-05,
8.11233456e-05, 6.82462706e-05, 6.65460752e-05, 5.97123593e-05,
5.57198452e-05, 4.97330783e-05, 4.72774204e-05, 3.28414360e-05,
3.06647372e-05, 2.73813189e-05, 2.26115918e-05, 2.10056898e-05,
1.91719443e-05, 1.68720430e-05, 1.55776092e-05, 1.32057505e-05,
1.05262046e-05, 9.74874096e-06, 9.10260045e-06, 8.55170758e-06,
7.16062443e-06, 5.16243236e-06, 4.55605308e-06, 4.31078301e-06,
```

```
3.94643896e-06, 3.62113398e-06, 3.49773023e-06, 2.54114726e-06,
2.37230188e-06, 1.62982390e-06, 6.43395130e-07, 4.04648215e-07,
3.13007168e-07, 2.64897987e-07, 8.19185402e-08, 6.60263612e-09,
1.88228466e-10, 5.19925774e-30, 1.12380342e-31, 1.12380342e-31,
1.12380342e-31, 1.12380342e-31])
```

Preserved variance ratio

In [38]:
```
each_pc_explained_variance_ratio=pca.explained_variance_ratio_
each_pc_explained_variance_ratio
```

Udacity Feedback

No tests loaded

Hide Test Results

```
Out[38]: array([5.55834276e-02, 3.64206008e-02, 2.82358113e-02, 2.54421271e-02,
                2.19056770e-02, 2.07692142e-02, 1.97435136e-02, 1.83468363e-02,
                1.79972080e-02, 1.62168921e-02, 1.46079305e-02, 1.33172566e-02,
                1.31159973e-02, 1.27787248e-02, 1.26366134e-02, 1.19394084e-02,
                1.15285112e-02, 1.13869213e-02, 1.12037198e-02, 1.06201645e-02,
                1.03253054e-02, 1.00973270e-02, 9.93443635e-03, 9.66236238e-03,
                9.56618303e-03, 9.43510281e-03, 9.24904700e-03, 8.89090309e-03,
                8.70531709e-03, 8.46126684e-03, 8.33374563e-03, 8.19003833e-03,
                8.08354882e-03, 7.95764968e-03, 7.73658511e-03, 7.68484394e-03,
                7.54555762e-03, 7.48330868e-03, 7.28544685e-03, 7.19172874e-03,
                7.12130533e-03, 6.95645683e-03, 6.76501322e-03, 6.69187746e-03,
                6.65497616e-03, 6.57376940e-03, 6.40304619e-03, 6.26575191e-03,
                6.22069840e-03, 6.06863762e-03, 6.01761187e-03, 5.98914275e-03,
                5.86314282e-03, 5.67931795e-03, 5.60300430e-03, 5.59135050e-03,
                5.53365371e-03, 5.40062073e-03, 5.34690518e-03, 5.24773689e-03,
                5.12928398e-03, 5.07507633e-03, 4.94908194e-03, 4.88484407e-03,
                4.79504697e-03, 4.70411831e-03, 4.63243433e-03, 4.58663509e-03,
                4.54156003e-03, 4.44247302e-03, 4.39885887e-03, 4.30885694e-03,
                4.26879548e-03, 4.24118565e-03, 4.20040514e-03, 4.10104282e-03,
                4.00005436e-03, 3.82365436e-03, 3.76404837e-03, 3.71678908e-03,
                3.64765210e-03, 3.61586972e-03, 3.58967104e-03, 3.50469118e-03,
                3.43365925e-03, 3.34758782e-03, 3.31527159e-03, 3.27996983e-03,
                3.16268755e-03, 3.11195494e-03, 3.08280388e-03, 2.99988838e-03,
                2.91688565e-03, 2.87244026e-03, 2.82177913e-03, 2.78756416e-03,
                2.77839187e-03, 2.67428217e-03, 2.64468607e-03, 2.61829190e-03,
                2.57574773e-03, 2.56486660e-03, 2.54247470e-03, 2.45044721e-03,
                2.43120802e-03, 2.41120729e-03, 2.37816660e-03, 2.36528300e-03,
                2.32985178e-03, 2.27181290e-03, 2.24941142e-03, 2.22528257e-03,
                2.17738332e-03, 2.13146536e-03, 2.11166711e-03, 2.10402876e-03,
                2.08780590e-03, 2.06701430e-03, 2.02783675e-03, 2.00576986e-03,
                1.99538519e-03, 1.97364335e-03, 1.95416793e-03, 1.93220136e-03,
                1.89509394e-03, 1.88701220e-03, 1.86555539e-03, 1.85412535e-03,
                1.84008077e-03, 1.82722806e-03, 1.80042818e-03, 1.79145378e-03,
                1.77747887e-03, 1.75633389e-03, 1.72713186e-03, 1.70921735e-03,
                1.68824545e-03, 1.67287070e-03, 1.65034898e-03, 1.64041257e-03,
                1.61856795e-03, 1.59198180e-03, 1.57939571e-03, 1.56729452e-03,
                1.53894066e-03, 1.52630125e-03, 1.50699896e-03, 1.47631589e-03,
                1.47008388e-03, 1.45258278e-03, 1.44814624e-03, 1.42233064e-03,
                1.39894786e-03, 1.38433378e-03, 1.36799336e-03, 1.34152023e-03,
                1.31930481e-03, 1.29998545e-03, 1.29021416e-03, 1.28222484e-03,
                1.26168673e-03, 1.25089085e-03, 1.23916269e-03, 1.22915079e-03,
                1.19887949e-03, 1.19111534e-03, 1.18421682e-03, 1.16821860e-03,
                1.16355550e-03, 1.12671303e-03, 1.09812954e-03, 1.09433824e-03,
                1.07099038e-03, 1.05522110e-03, 1.04812931e-03, 1.04640125e-03,
                1.01019560e-03, 1.00717066e-03, 9.80918680e-04, 9.73453196e-04,
                9.49592916e-04, 9.34281008e-04, 9.22336547e-04, 9.13365594e-04,
                9.05706201e-04, 8.74353395e-04, 8.58131034e-04, 8.50772750e-04,
                8.20380216e-04, 8.06736565e-04, 7.88990533e-04, 7.73154076e-04,
                7.68917637e-04, 7.54146710e-04, 7.33023755e-04, 7.18954687e-04,
                7.08205003e-04, 6.93006094e-04, 6.82074876e-04, 6.67107955e-04,
                6.48715878e-04, 6.28165852e-04, 6.17675505e-04, 6.00295738e-04,
                5.84540165e-04, 5.63065430e-04, 5.39463665e-04, 5.37634113e-04,
                5.20046745e-04, 5.13594694e-04, 4.93356587e-04, 4.79655061e-04,
                4.67020880e-04, 4.64594955e-04, 4.47712905e-04, 4.33554881e-04,
                4.23269833e-04, 4.15681955e-04, 3.99468654e-04, 3.83412231e-04,
                3.75963005e-04, 3.59397977e-04, 3.50579349e-04, 3.42713606e-04,
                3.33205984e-04, 3.21672737e-04, 2.97196305e-04, 2.93632890e-04,
```

```
2.89417878e-04, 2.67474555e-04, 2.56784164e-04, 2.49017072e-04,
2.43661420e-04, 2.34441604e-04, 2.22001759e-04, 2.15373170e-04,
2.09283100e-04, 2.04271038e-04, 2.00644729e-04, 1.96510672e-04,
1.89962237e-04, 1.80564787e-04, 1.72561708e-04, 1.66160223e-04,
1.64576399e-04, 1.59591851e-04, 1.51313273e-04, 1.44798158e-04,
1.41953729e-04, 1.32915567e-04, 1.30237608e-04, 1.29043847e-04,
1.24219617e-04, 1.22025970e-04, 1.16062254e-04, 1.14134877e-04,
1.09614634e-04, 1.08578696e-04, 1.05086439e-04, 1.04266166e-04,
1.00818325e-04, 9.67982304e-05, 9.40346484e-05, 9.08487522e-05,
8.90362155e-05, 8.67729167e-05, 8.47564197e-05, 8.30397337e-05,
8.12786672e-05, 7.95982977e-05, 7.87478100e-05, 7.83915232e-05,
7.36653493e-05, 7.35003990e-05, 7.20961781e-05, 7.01814254e-05,
6.91532436e-05, 6.72328238e-05, 6.66837701e-05, 6.38638054e-05,
6.27815817e-05, 6.07705678e-05, 5.98870364e-05, 5.91004230e-05,
5.82255008e-05, 5.74461494e-05, 5.72177621e-05, 5.40387001e-05,
5.35646657e-05, 5.22371932e-05, 5.14239525e-05, 5.05204581e-05,
5.02919102e-05, 4.89518111e-05, 4.78353877e-05, 4.67240316e-05,
4.52875730e-05, 4.50429505e-05, 4.42157639e-05, 4.31299324e-05,
4.21550484e-05, 4.17123905e-05, 4.00283945e-05, 3.94811180e-05,
3.86462426e-05, 3.69064931e-05, 3.67058838e-05, 3.62266541e-05,
3.48197999e-05, 3.33926942e-05, 3.29251509e-05, 3.19370253e-05,
3.17179325e-05, 3.07794589e-05, 3.06563205e-05, 2.89197009e-05,
2.77024942e-05, 2.70440414e-05, 2.63121076e-05, 2.61759092e-05,
2.52908566e-05, 2.51306881e-05, 2.48162183e-05, 2.41134297e-05,
2.27303519e-05, 2.19416262e-05, 2.08246350e-05, 2.03364629e-05,
2.01867760e-05, 1.96906495e-05, 1.91486240e-05, 1.81207258e-05,
1.73740971e-05, 1.70379985e-05, 1.65840847e-05, 1.59176613e-05,
1.53868697e-05, 1.52744142e-05, 1.47709521e-05, 1.46710811e-05,
1.43775215e-05, 1.36563323e-05, 1.32976409e-05, 1.32271880e-05,
1.28109229e-05, 1.23378834e-05, 1.20751433e-05, 1.15363106e-05,
1.11917777e-05, 1.08551326e-05, 1.02380330e-05, 9.96261644e-06,
9.17001381e-06, 8.90848931e-06, 8.54552871e-06, 8.29933193e-06,
7.87330381e-06, 7.46916375e-06, 7.33192097e-06, 6.96193298e-06,
6.58201156e-06, 6.53358507e-06, 6.48405113e-06, 5.99298661e-06,
5.72886479e-06, 5.55665956e-06, 5.15800062e-06, 5.02217896e-06,
4.80130452e-06, 4.53136396e-06, 4.29748992e-06, 4.12088409e-06,
3.97605879e-06, 3.85947035e-06, 3.71216525e-06, 3.59247124e-06,
3.39661245e-06, 3.06823168e-06, 3.04806736e-06, 2.86675053e-06,
2.84010479e-06, 2.68433048e-06, 2.44169165e-06, 2.41748361e-06,
2.29145643e-06, 2.18092230e-06, 2.07524375e-06, 2.01957395e-06,
1.89025218e-06, 1.82606343e-06, 1.80308896e-06, 1.69483766e-06,
1.62577057e-06, 1.60720538e-06, 1.55663418e-06, 1.44280917e-06,
1.41090069e-06, 1.38268445e-06, 1.34175861e-06, 1.31600943e-06,
1.14148854e-06, 1.11723944e-06, 1.09625030e-06, 1.02141359e-06,
9.09611990e-07, 8.91149058e-07, 8.19710144e-07, 7.76308366e-07,
7.44237502e-07, 6.56492922e-07, 6.43023631e-07, 5.94736532e-07,
5.33813208e-07, 5.07288609e-07, 4.86159559e-07, 4.61480815e-07,
4.57384749e-07, 3.96214587e-07, 3.85820272e-07, 3.65937546e-07,
3.48415880e-07, 3.02420832e-07, 2.88570198e-07, 2.64799327e-07,
2.44334502e-07, 2.41203725e-07, 2.30392769e-07, 2.19813879e-07,
2.00163250e-07, 1.86308458e-07, 1.82197302e-07, 1.76443243e-07,
1.71037080e-07, 1.43887592e-07, 1.40302971e-07, 1.25895049e-07,
1.17477399e-07, 1.04855149e-07, 9.96777422e-08, 6.92415145e-08,
6.46522534e-08, 5.77296312e-08, 4.76733374e-08, 4.42875206e-08,
4.04213280e-08, 3.55723120e-08, 3.28431818e-08, 2.78424538e-08,
2.21930110e-08, 2.05538390e-08, 1.91915433e-08, 1.80300637e-08,
1.50971620e-08, 1.08842572e-08, 9.60579237e-09, 9.08867517e-09,
```

```
       8.32050736e-09, 7.63464791e-09, 7.37446860e-09, 5.35764893e-09,
       5.00166237e-09, 3.43625276e-09, 1.46309761e-09, 8.32864475e-10,
       6.59931262e-10, 5.58499871e-10, 1.72713635e-10, 1.39207227e-11,
       3.96853050e-13, 1.09618983e-32, 2.36938028e-34, 2.36938028e-34,
       2.36938028e-34, 2.36938028e-34])
```

In [40]:
```python
cumulative_Variance = pca.explained_variance_ratio_.cumsum()
```

We can see that We should consider 200 - 230, approx optimal number of principal components Principal Components to explain more than 98% of variance using below graph

In [41]:
```python
plt.figure(1, figsize=(6, 4))
plt.clf()
plt.plot(cumulative_Variance, linewidth=2)
plt.axis('tight')
plt.grid(True)
plt.xlabel('n_components')
plt.ylabel('cumulativeVariance')
plt.show()
```
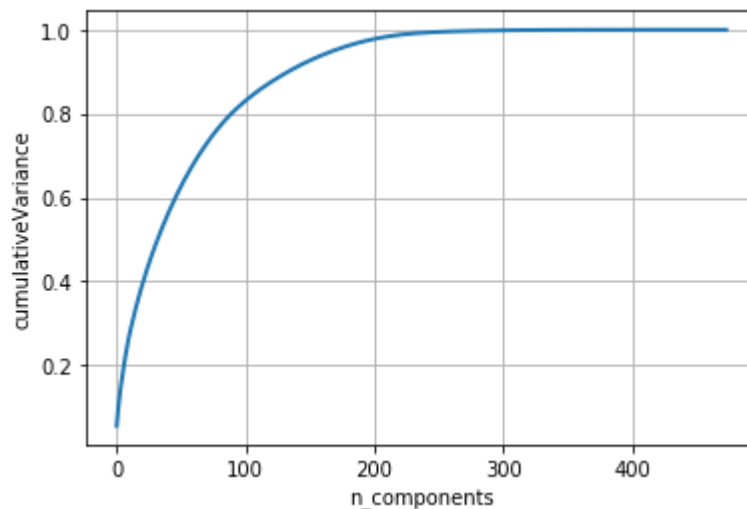

We can see around 98% of variance below 200-230 components

In [44]:
```python
pca.n_components=7
principal7Components = pca.fit_transform(std_data)
principalDF = pd.DataFrame(data=principal7Components,columns=['pcomp1','pcomp
2','pcomp3','pcomp4','pcomp5','pcomp6','pcomp7'])
```

Top 7 principal-components based values of their variances

In [45]: `principalDF`

Out[45]:

	pcomp1	pcomp2	pcomp3	pcomp4	pcomp5	pcomp6	pcomp7
0	-1.694695	2.925812	3.890342	-2.545239	-0.070348	-1.184549	-1.716559
1	-2.247194	0.876782	2.878214	-2.034221	-0.328795	-1.527402	-2.236821
2	0.442994	1.231955	1.146580	-0.371850	0.603738	2.164022	-2.154083
3	1.158528	5.131687	4.251766	-3.428451	3.023635	1.925283	-8.182085
4	0.753948	2.379104	2.543256	-0.033414	2.039298	3.268601	-4.900000
5	2.076392	2.875918	3.572881	-2.209836	1.779089	3.565122	-2.458103
6	-1.747077	4.526124	2.858190	-1.451715	0.475425	-1.430457	-2.957467
7	1.292039	2.012678	4.734057	-2.362956	0.374715	1.455374	-2.476023
8	-0.471772	19.143200	-3.488871	3.608310	0.417586	2.585036	-0.662857
9	1.303172	4.006341	5.713491	-2.966116	1.035043	2.495597	-2.476655
10	-2.083069	22.620039	-7.100650	6.414667	-2.238695	-1.268602	2.242799
11	0.735209	3.659814	4.836746	-2.406979	1.306302	2.116889	-2.212877
12	0.561499	1.502583	3.127285	-0.570487	1.310680	2.024431	-3.988153
13	-2.618774	2.687974	4.617412	-1.328211	-0.295889	-1.277196	-1.291162
14	-3.186040	32.201045	-11.994388	7.766983	0.588387	3.348226	2.470326
15	-2.980570	27.061397	-9.958615	5.956671	0.768443	1.754741	0.604205
16	-0.916633	0.769981	1.212576	-1.983483	1.273528	1.628671	-4.723774
17	-1.232625	3.851757	2.888540	-2.577332	1.657917	2.073207	-6.885032
18	-0.749024	0.223831	2.387614	-0.576839	-0.041625	0.556619	-3.552802
19	1.020533	2.615800	2.561091	-0.774736	1.347508	1.960637	-3.280060
20	0.153006	-0.613494	1.383259	-0.801180	0.404798	1.696359	-3.806064
21	-0.282477	0.818745	3.129836	-1.507462	0.758892	0.941241	-4.548035
22	-0.582044	0.337954	1.925557	-0.591190	0.218661	0.566078	-4.681425
23	-2.219011	25.077582	-6.981846	5.048530	0.225625	1.916324	0.185654
24	-3.235047	24.221556	-9.295245	4.397435	-0.342400	0.885021	0.482012
25	-3.543949	23.942219	-7.839054	5.193226	-0.974467	0.387509	0.876493
26	-1.775195	1.526452	1.956624	-1.990451	-0.749068	-3.613862	0.839460
27	-0.587920	2.883265	-0.440802	1.551498	0.051368	2.981275	-4.574252
28	-0.628810	3.647641	3.043362	-2.339473	1.206649	1.325085	-3.950534
29	-0.882877	2.084894	-0.595677	-0.128845	0.695389	1.009549	-2.654831
...	...	...	...	...	...	...	...
1537	-0.703425	-1.231339	0.891798	1.979978	-0.129514	0.427263	-2.811470

	pcomp1	pcomp2	pcomp3	pcomp4	pcomp5	pcomp6	pcomp7
1538	0.091720	-3.087884	-3.192241	5.471369	0.068783	3.973099	-2.149885
1539	2.564356	-2.983508	-2.726740	2.125188	1.092110	1.886720	1.657666
1540	-0.470646	-2.763350	-3.081914	2.322429	0.600848	0.821994	-0.973785
1541	-0.339324	-3.486483	-2.681056	1.859345	0.065839	3.177510	-1.637235
1542	-2.601628	-0.698569	-2.068284	1.004057	-2.183479	-3.642114	-1.820422
1543	-2.227351	-0.418138	0.136023	-0.399528	-1.425138	-1.282659	-2.956743
1544	-2.173570	-1.554030	-1.565042	0.977956	-1.454302	-2.460943	-1.048450
1545	1.555604	-1.098203	-1.772633	1.078833	1.690175	-0.074746	-0.500997
1546	-1.787294	-1.470350	-1.929721	0.879589	-0.884571	-1.094185	2.969880
1547	1.938959	-2.775672	-0.625056	2.159403	0.208924	-0.536735	0.434084
1548	1.913413	-0.302717	0.880078	-0.568864	2.056295	0.947159	-1.148983
1549	-0.107957	-4.074654	-0.826749	2.184694	-0.707332	-0.341727	1.203746
1550	-0.860093	-3.616090	-4.467705	3.361321	-0.141656	1.741185	-0.063746
1551	-0.671974	-1.450034	-1.091798	2.733379	-0.184463	-0.013225	-2.577605
1552	1.784591	-0.973232	1.846763	-2.748249	1.679823	1.936444	2.690457
1553	-2.012300	0.061228	0.937146	-1.167804	-0.769876	-1.277587	-2.888866
1554	1.032687	-1.518211	-1.900624	0.879672	0.810838	0.544758	-0.962426
1555	-0.394507	-3.160581	-3.301104	2.694920	-0.416933	3.781954	-2.987121
1556	0.902511	-3.448693	-1.525163	3.925534	1.053600	-0.794520	0.172042
1557	0.875302	-1.185432	-1.124881	0.600505	-0.322499	-0.501485	-0.016299
1558	-0.392312	-1.084178	-0.672872	-0.138581	0.533097	0.530556	-2.472355
1559	1.200491	-3.551510	-3.918440	3.862557	0.430796	4.072492	-1.845865
1560	-1.482319	-1.260658	1.881994	0.224500	0.058204	-1.214285	-2.121000
1561	2.229063	-2.341924	-0.451043	1.357294	1.253069	1.578701	-0.403294
1562	-1.190341	-3.735384	-2.611351	5.708814	-0.557646	1.874632	-1.194507
1563	-0.389805	0.484748	2.564024	-0.089495	1.314357	1.159761	-2.480222
1564	-1.169632	-1.773264	-1.618357	1.098914	-0.768666	-2.221571	0.520276
1565	-1.176656	-3.234535	-3.496680	3.074201	-0.007506	3.095718	-2.471983
1566	1.953314	-3.012189	-2.618094	3.718272	0.471644	3.758435	-2.638329

1567 rows × 7 columns

In [48]:
```python
plt.plot(ei_values[:10])
plt.show()
```
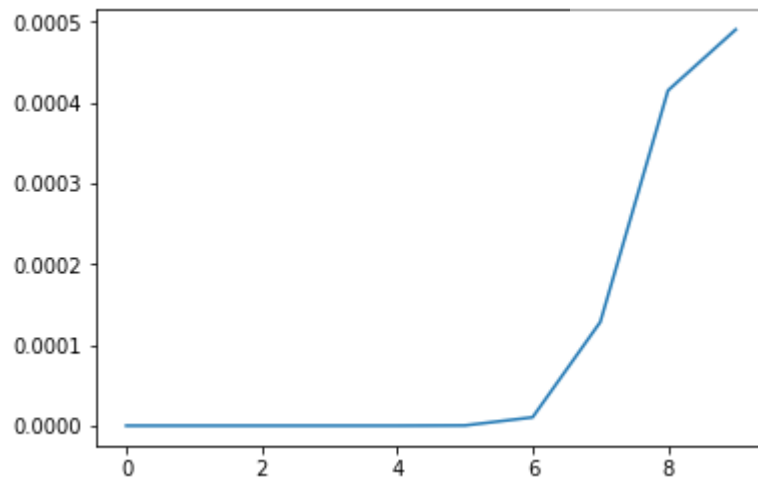
Saturation point is 5 from above graph of eigen values