PCA On Stock Price

Group2: Yaxin Li, Blake Hillier, Joe Puhalla

January 2020

We used Python for data processing and modeling. We first pull the minutes' data of 100 stocks from Bloomberg. We wanted to plot the data on sigma coordinates. So, we selected 3 stocks out of 100 and then calculated the sigma of each combination by using SVD. We had 33 unique combinations in total.

The 100 stocks are:
'abt','adp','aep','aig','axp','ba','bac','bax','bdx','bk','bmy','bp','cat','cl','clx','cmi','cop','cpb','csx','cvs','cvx','d','de','dhr','duk','eix','emr','es','etn','etr','exc','fdx','fitb','gd','ge1','gis','glw','gpc','gww','hal','hban','hpq','hsy','hum','ibm1','iff','intc','ip','jnj1','jpm','k','klac','kmb','ko','kr','l','leg','lhx','lly','lnc','mkc','mlhr','mmm','mo','mrk','msi','mtb','nem','nsc','ntrs','omc','pcar','peg','pep','pfe','pg','pnw','ppg','ppl','rtn','shw','slb','so','spgi','tgt','tmo','tsn','txn','unp','usb','utx','vfc','wba','wfc','whr','wmb','wy','xel','xom','xrx'

After calculated the sigmas by using SVD over 100 stocks. We plotted the data on sigma coordinates. We plotted 33 data points in the 3D graph as follows.
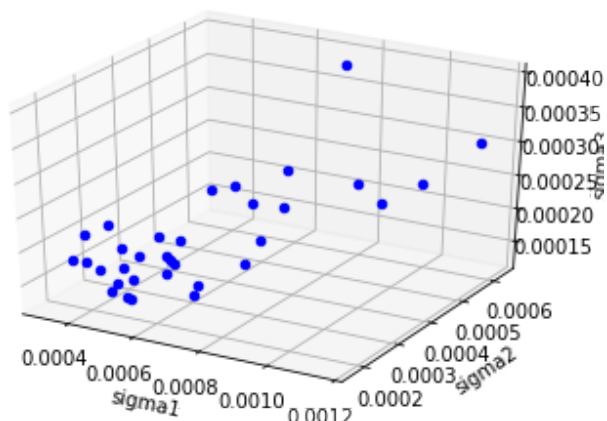
Then we applied the KMeans clustering and cluster the data points as following.