Processing of Raw Stock Data

We obtained the historical raw data for a year 2009/2010 from the site: <http://pages.swcp.com/stocks/>

We need to process this data before we can provide it as input to our clustering program. We filter closing prices of last day of every month from one year historical data. We executed the following steps to come up with an 11-point representation of a stock.

* + We generated returns for each stock using r(t+1) =  (s(t+1) - s(t))/s(t) generating 11 data points for each stock
  + Returns were normalized using (r(t) - mu)/std where “mu” is mean and std is standard deviation

We have already processed the raw data for you and is available on the resource page. Please look for the file “stocks.txt”. Please note that the file stocks.txt does not have all the stocks in S&P500. We had to filter out certain stocks as the data was not complete for those stocks.