**# Filename: kmeansStu.py**

```
from numpy import vstack
from scipy.cluster.vq import kmeans,vq
list1=[88.0,64.0,96.0,85.0]
list2=[92.0,99.0,95.0,94.0]
list3=[91.0,87.0,99.0,95.0]
list4=[78.0,99.0,97.0,81.0]
list5=[88.0,78.0,98.0,84.0]
list6=[100.0,95.0,100.0,92.0]
data = vstack((list1,list2,list3,list4,list5,list6))
centroids,_ = kmeans(data,2)
result,_= vq(data,centroids)
print(result)          #    “print result” in Python 2.x
```

**# Filename: kmeansDJI.py**

```
from scipy.cluster.vq import kmeans,vq
from matplotlib.finance import quotes_historical_yahoo_ochl
from datetime import datetime
start = datetime(2016,7,1)
end = datetime(2016,9,30)
listDji = ['AXP','BA','CAT','CSCO','CVX','DD','DIS','GE','GS','HD','IBM',
'INTC','JNJ','JPM','KO','MCD','MMM','MRK','MSFT','NKE','PFE','PG','T','TRV',
'UNH','UTX','V','VZ','WMT','XOM']
quotes = [ [0 for col in range(90)] for row in range(30)]
listTemp = [ [0 for col in range(90)] for row in range(30)]
for i in range(30):
      quotes[i] = quotes_historical_yahoo_ochl(listDji[i], start, end)
days = len(quotes[0])
for i in range(30):
      for j in range(days-1):
            if (quotes[i][j][2] and quotes[i][j+1][2] and (quotes[i][j+1][2] >= quotes[i][j][2])):
                  listTemp[i][j] = 1.0
            else:
                  listTemp[i][j] = -1.0
data = vstack(listTemp)
centroids,_ = kmeans(data,4)      #float or double is supported
result,_= vq(data,centroids)
print(result)       #    “print result” in Python 2.x
```

**# Filename: closeMeansKO.py**

```
import time
```

```
from matplotlib.finance import quotes_historical_yahoo_ochl
from datetime import date
from datetime import datetime
import pandas as pd
import matplotlib.pyplot as plt

today = date.today()
start = (today.year-1, today.month, today.day)
quotes = quotes_historical_yahoo_ochl('KO', start, today)
fields = ['date','open','close','high','low','volume']

list1 = []
for i in range(0,len(quotes)):
    x = date.fromordinal(int(quotes[i][0]))
    y = datetime.strftime(x,'%Y-%m-%d')
    list1.append(y)
#print(list1)
quoteskodf = pd.DataFrame(quotes, index = list1, columns = fields)
quoteskodf = quoteskodf.drop(['date'], axis = 1)
#print(quotesdf)
listtemp = []
for i in range(0,len(quoteskodf)):
    temp = time.strptime(quoteskodf.index[i],"%Y-%m-%d")
    listtemp.append(temp.tm_mon)

print(listtemp)     #   “print listtemp” in Python 2.x
tempkodf = quoteskodf.copy()
tempkodf['month'] = listtemp
closeMeansKO = tempkodf.groupby('month').mean().close
listKO = []
for i in range(1,13):
    listKO.append(closeMeansKO[i])
listKOIndex = closeMeansKO.index
plt.plot(listKOIndex,listKO,'g*')
plt.show()
```

**# Filename: quotesdfbar.py**

```
from matplotlib.finance import quotes_historical_yahoo_ochl
from datetime import date
from datetime import datetime
import pandas as pd

start = datetime(2015,10,1)
end = datetime(2015,10,31)
```

```
quotesINTC = quotes_historical_yahoo_ochl('INTC', start, end)
quotesGE = quotes_historical_yahoo_ochl('GE', start, end)
fields = ['date','open','close','high','low','volume']
#quotesdf = pd.DataFrame(quotes, columns = fields)
#quotesdf = pd.DataFrame(quotes, index = range(1,len(quotes)+1),columns = fields)
list1 = []
for i in range(0,len(quotesINTC)):
    x = date.fromordinal(int(quotesINTC[i][0]))
    y = datetime.strftime(x,'%Y-%m-%d')
    list1.append(y)
print(list1)            #   “print list1”   in Python 2.x

list2 = []
for i in range(0,len(quotesGE)):
    x = date.fromordinal(int(quotesGE[i][0]))
    y = datetime.strftime(x,'%Y-%m-%d')
    list2.append(y)
print(list2)      #   “print list2”   in Python 2.x

quotesdfINTC = pd.DataFrame(quotesINTC, index = list1, columns = fields)
quotesdfINTC = quotesdfINTC.drop(['date'], axis = 1)
quotesdfGE = pd.DataFrame(quotesGE, index = list1, columns = fields)
quotesdfGE = quotesdfGE.drop(['date'], axis = 1)
#print(quotesdf)

quotesdf = pd.DataFrame()
quotesdf['closeINTC'] = quotesdfINTC.close
quotesdf['closeGE'] = quotesdfGE.close
#quotesdf.plot(kind='bar')
#quotesdf.plot(kind='barh')
#quotesdf.plot(kind='scatter',x='closeINTC',color='g',y='closeGE')
quotesdf.plot(kind='kde')
plt.show()
#quotesdf.boxplot()
#quotesdfGE.close.plot()
quotesdf.to_excel('kkk.xls')
```