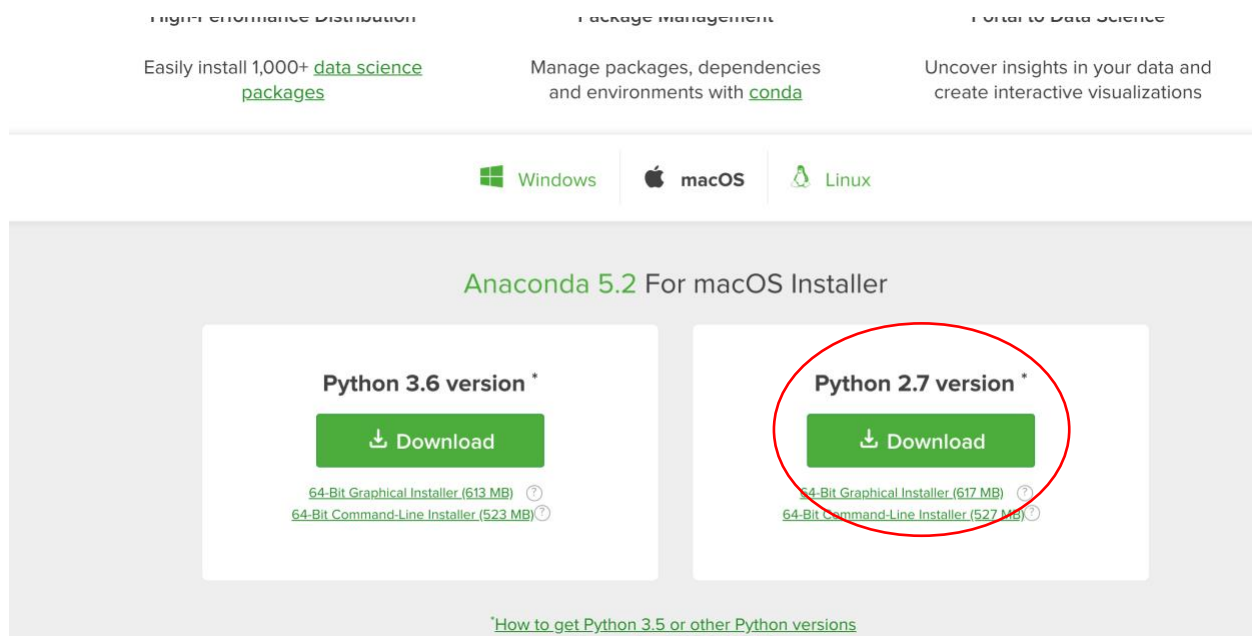
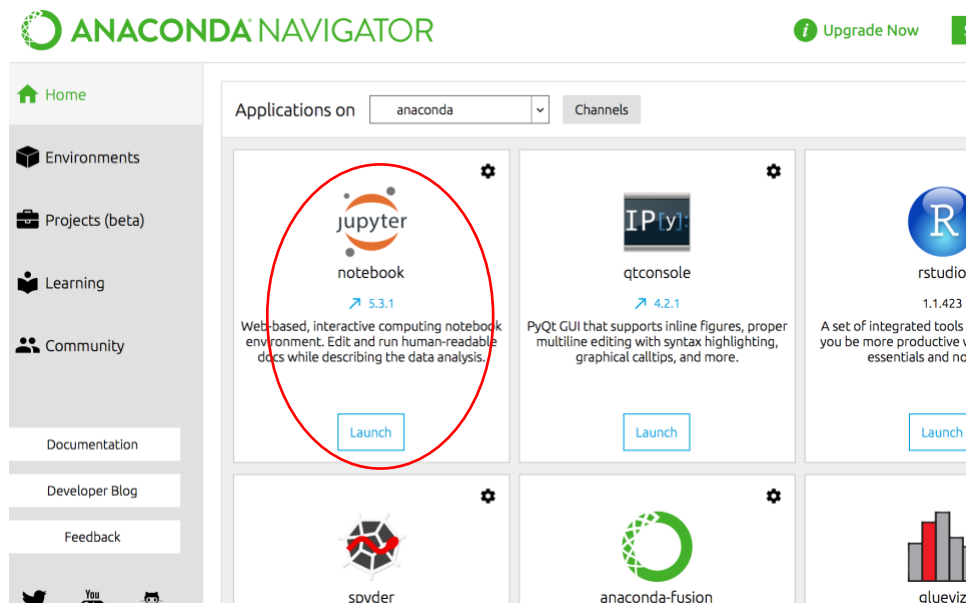


1. Download the anaconda python 2.7 distribution, which will create a separate python environment. There is no need to uninstall python 3.6

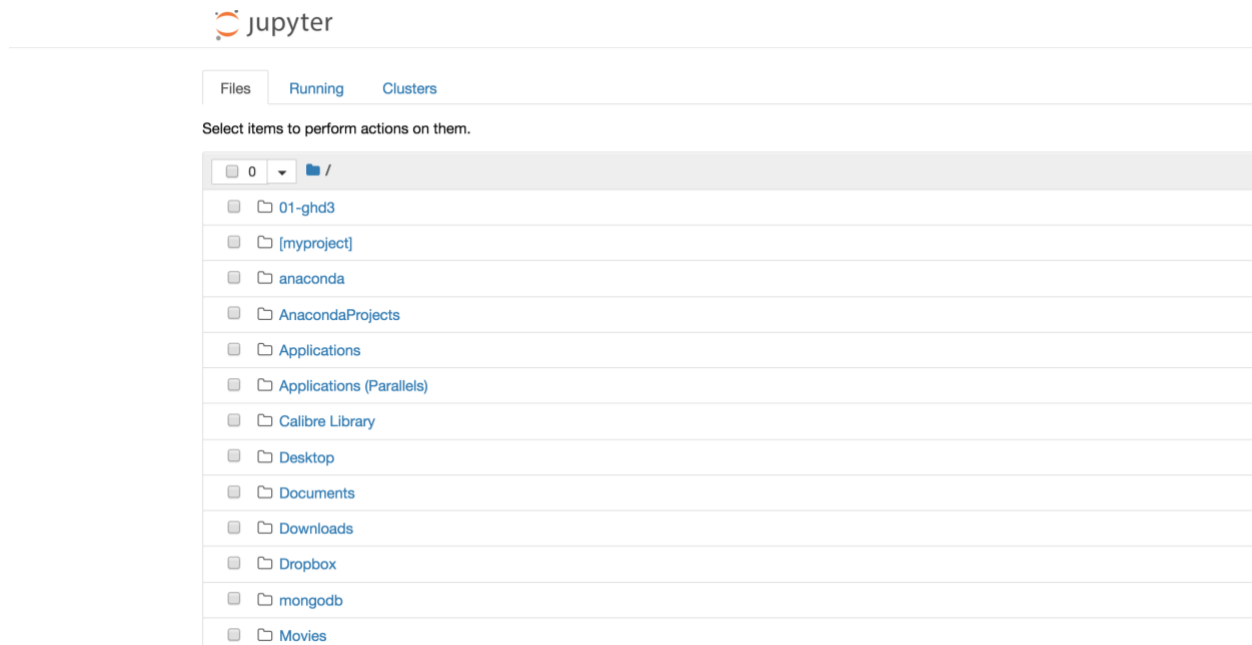
<https://www.anaconda.com/download/#macos>



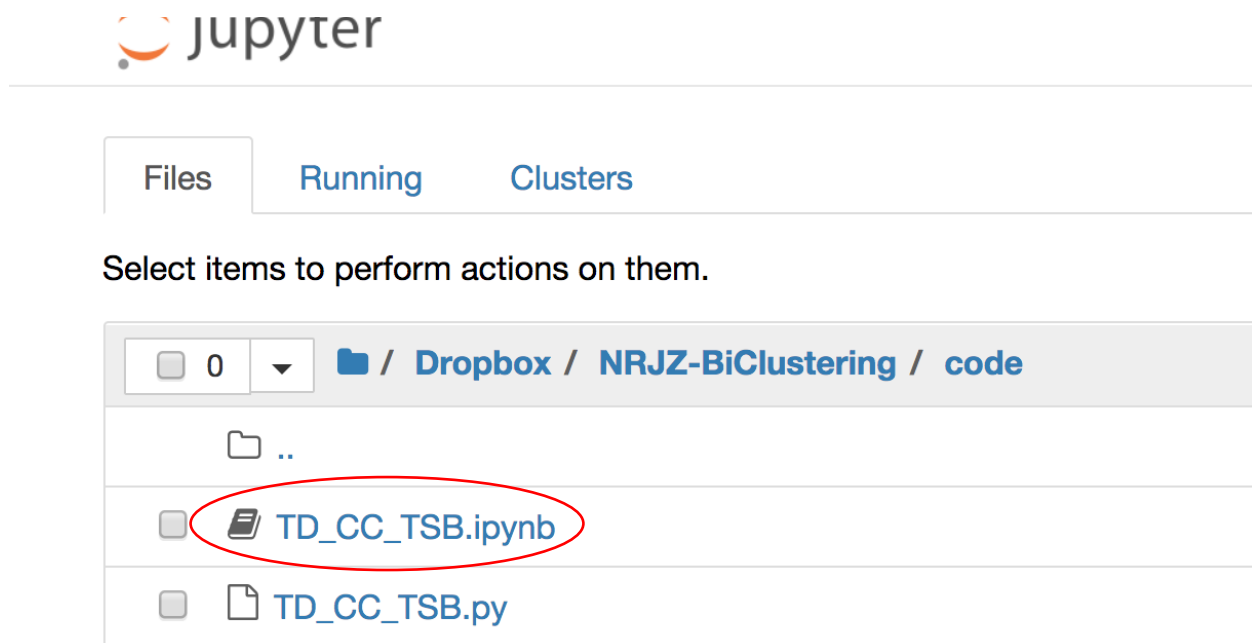
2. After download the anaconda and install it. Click Anaconda-Navigator icon. Then click the launch Jupyter.



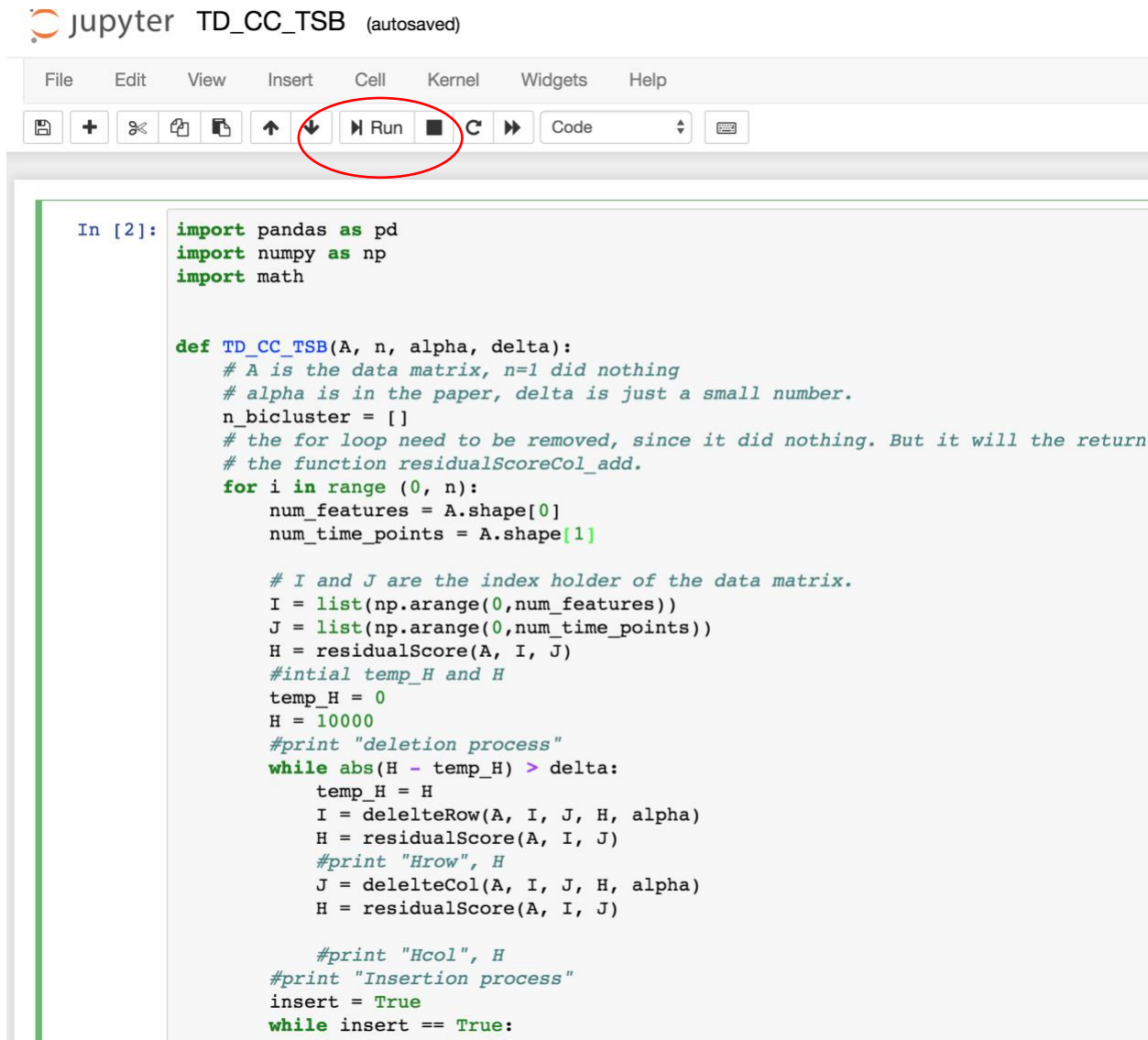
3. An interface will show up in the internet browser. Then navigate to the our dropbox share folder



4. Click the TD\_CC\_TSB.ipynb rather than the TD\_CC\_TSB.py. The .py file need to be run from command line and is good for cluster computing.



5. Click run or use the keyboard shift + enter. It will run automatically, since I have put two days data files in the same folder as the code. So there is no need to set up file directory.



Jupyter TD\_CC\_TSB (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Run




















```
In [2]: import pandas as pd
import numpy as np
import math

def TD_CC_TSB(A, n, alpha, delta):
    # A is the data matrix, n=1 did nothing
    # alpha is in the paper, delta is just a small number.
    n_biclust = []
    # the for loop need to be removed, since it did nothing. But it will the return
    # the function residualScoreCol_add.
    for i in range(0, n):
        num_features = A.shape[0]
        num_time_points = A.shape[1]

        # I and J are the index holder of the data matrix.
        I = list(np.arange(0, num_features))
        J = list(np.arange(0, num_time_points))
        H = residualScore(A, I, J)
        #intial temp_H and H
        temp_H = 0
        H = 10000
        #print "deletion process"
        while abs(H - temp_H) > delta:
            temp_H = H
            I = deleteRow(A, I, J, H, alpha)
            H = residualScore(A, I, J)
            #print "Hrow", H
            J = deleteCol(A, I, J, H, alpha)
            H = residualScore(A, I, J)

        #print "Hcol", H
        #print "Insertion process"
        insert = True
        while insert == True:
```

6. After the program finish. The follow files will be in the same folder. The files with names starting with date is the identified bicluster. The date\_bicluster file will record the date, the number of biclusters and the number stocks explained.

	20130102_bicluster1.csv	✓	39 KB	Comm...t (.csv)
	20130102_bicluster2.csv	✓	66 KB	Comm...t (.csv)
	20130102_bicluster3.csv	✓	91 KB	Comm...t (.csv)
	20130102_bicluster4.csv	✓	58 KB	Comm...t (.csv)
	20130102_bicluster5.csv	✓	43 KB	Comm...t (.csv)
	20130102_bicluster6.csv	✓	31 KB	Comm...t (.csv)
	20130102_bicluster7.csv	✓	23 KB	Comm...t (.csv)
	20130102_bicluster8.csv	✓	15 KB	Comm...t (.csv)
	20130102_bicluster9.csv	✓	14 KB	Comm...t (.csv)
	20130103_bicluster1.csv	✓	43 KB	Comm...t (.csv)
	20130103_bicluster2.csv	✓	80 KB	Comm...t (.csv)
	20130103_bicluster3.csv	✓	79 KB	Comm...t (.csv)
	20130103_bicluster4.csv	✓	57 KB	Comm...t (.csv)
	20130103_bicluster5.csv	✓	23 KB	Comm...t (.csv)
	20130103_bicluster6.csv	✓	39 KB	Comm...t (.csv)
	20130103_bicluster7.csv	✓	31 KB	Comm...t (.csv)
	20130103_bicluster8.csv	✓	19 KB	Comm...t (.csv)
	20130103_bicluster9.csv	✓	15 KB	Comm...t (.csv)
	date_bicluster.csv	✓	39 bytes	Comm...t (.csv)