

```

#This code is written in python, an open-source high-level programming
language.
#Python can be downloaded at http://www.python.org

from LocalFuncs import ReadDataFile, WriteDataFile
# LocalFuncs is a personal package of convenience functions. My current
version
# can be downloaded from http://www.umat.maine.edu/~r.springuel/000CCFE8-80000018/
from Pycluster import treecluster, clustercentroids
# Pycluster can be downloaded from http://bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster/software.htm
# It also has an internal dependancy on Numeric, which can be downloaded
# from http://sourceforge.net/project/showfiles.php?group\_id=1369&package\_id=1351
from pylab import figure, show
# pylab is the default interface for matplotlib which can be downloaded
# from http://matplotlib.sourceforge.net/
from ClusterPlot import poptree, fadetree
# ClusterPlot can be downloaded from http://www.umat.maine.edu/~r.springuel/000CCFE8-80000018/
# It also had internal dependancies on matplotlib and numpy.
# numpy can be downloaded from http://sourceforge.net/project/showfiles.php?group\_id=1369&package\_id=175103

##Read in the data file
# The file name is formatted for a specific location on a computer running
# windows. If running Mac OS X or Linux, it is possible to use relative
# references to the working directory.
D = ReadDataFile('c:/datafiles/BCAccel.dat')

##Strip away all information that Pycluster deems irrelevant.
data = []
for i in range(len(D)-1):
    data.append(D[i+1][1:])

##Prepare the array that will recieve the clustering results.
Results = [D[0]]
Results[0][0] = 'Cluster'
Results[0].insert(1, 'N')

##Perform the clustering and then format the output.
tree = treecluster(data, method='a')
a = range(1, 50)
for i in a:
    cut = tree.cut(i)

```

```
cdata,cmask = clustercentroids(data,clusterid=cut)
for j in range(len(cdata)):
    c = list(cdata[j])
    c.insert(0,j+1)
    c.insert(1,list(cut).count(j))
    Results.append(c)
Results.append([])
print i
```

##Create the output, first as a comma seperated value spreadsheet, and then as

##a couple of graphs.

As before, the file name is formatted for a specific location on a computer

running windows.

```
WriteDataFile('c:/datafiles/BCAccelResults.csv',Results,',')
```

#These two graphs did not appear in the paper, but were used for preliminary analysis. The final versions generated for the paper had more complicated code

#associated with their generation.

```
figure(1)
```

```
fadetree(tree)
```

```
figure(2)
```

```
poptree(tree,lowerlimit=10)
```

```
show()
```