### Isolate Community start area and Community End area vectors by filtering out bad Time, Distance, Speed and Community area itself, data

import csv

with open("C:/Users/Akshay/Desktop/Taxi\_Trips.csv",encoding='utf-8',newline='') as csvfile:

b=csv.reader(csvfile)

next(b)

pcom=[]

dcom=[]

for row in b:

if row[4]=='' or row[5]=='' or float(row[4])==0.0 or float(row[5])==0.0 or float(row[4])>1000.0 or (float(row[5])/(float(row[4])/3600.0))>70.0 or row[8]=='' or row[9]=='':

continue

else:

pcom.append(int(row[8]))

dcom.append(int(row[9]))

### To create a list of tuples of community area pairs line by line

arr=[]

k=0

while k<len(pcom):

arr.append((pcom[k],dcom[k]))

k=k+1

### To count the pairwise frequencies

### To find the most frequent pairings(10,100, 1000…..)

from collections import Counter

c=Counter(arr)

c.most\_common(5)