import os

path = "C:\Users\econ12\Documents\Python ancestry scraping"

file\_names = os.listdir(path)

OUTFILENAME = 'records.txt'

#edit this to be a dict you lazy fuck ## many the dict value a list!

info = []

zips = []

cities = []

counties = []

states = []

fullrecords = []

for file in file\_names:

with open(file, 'r') as readfile:

listreadfile = list(readfile)

for i in range(len(listreadfile)):

if listreadfile[i].split().isnumeric():

#checks the first string of the row

info.append(list(readfile))

# info is a list of the rows with the information we want

for i in range(len(info))

zips.append(info[i].split()[0])

cities.append(info[i].split()[1])

counties.append(info[i].split()[2])

states.append(info[i].split()[3])

fullrecords.append(zips[i])

fullrecords.append(cities[i])

fullrecords.append(counties[i])

fullrecords.append(states[i])

#fullrecords has full records

with open(OUTFILENAME, 'a') as outfile:

for i in range(len(fullrecords)/4): #check that /4 works!

record\_id = 1

zip = fullrecords[i]

city = fullrecords[i+1]

county = fullrecords[i+2]

state = fullrecords[i+3]

record = '\t'.join([record\_id, zip, city, county, state])

outfile.write(record + '\n')

record\_id += 1

UPDATED to a dict

import os

path = "C:\Users\econ12\Documents\Python ancestry scraping"

file\_names = os.listdir(path)

OUTFILENAME = 'records.txt'

counter = 1

fullrecords = {}

for file in file\_names:

with open(file, 'r') as readfile:

listreadfile = list(readfile)

for i in range(len(listreadfile)):

if listreadfile[i].split().isnumeric():

#checks the first string of the row for a number

recordlist = listreadfile[i].replace("," , "").split()

fullrecords[counter] = recordlist[:4]

#fullrecords has full records

counter += 1

with open(OUTFILENAME, 'a') as outfile:

for i in range(1,len(fullrecords)+1):

record = '\t'.join([fullrecords[i], i])

outfile.write(record + '\n')