```
# -*- coding: utf-8 -*-
"""
Created on Mon May 15 11:44:05 2017

@author: Anonymized for peer review
"""

from twitter import *
import re
import datetime
import csv


news_username_list=[]
#Create this five .txt (seperate rows for each case) editor files and put
    them in one folder. Specify path here:
stadt_list=open("path.txt").read().splitlines()
bundesland_list=open("path.txt").read().splitlines()
plz_list=open("path.txt").read().splitlines()
red_sitz_list=open("path.txt").read().splitlines()
news_username_list_with_at=open("/path.txt").read().splitlines()
for u in news_username_list_with_at:
    username=u.replace("@","")
    news_username_list.append(username)

#Specify which newspapers from the list shall be scraped. a=0/b=1 are the
    first two; a=len(news_username_list-3)/b=len(news_username_list) are the
    last 3
a=0
b=1

current_examples_list=news_username_list[a:b]
current_redaktionssitz_list=red_sitz_list[a:b]

#Generate Twtter Application in your account and enter the respective keys
    here:
consumer_key = "consumer key"
consumer_secret = "consumer secret"
access_key = "access key"
access_secret = "access secret"

current_goal=len(current_examples_list)
current_counter=0

while current_counter<current_goal:



    twittername=current_examples_list[current_counter]
    source_account=twitter.users.show(screen_name=twittername)

    source_location=current_redaktionssitz_list[current_counter]

    print("Started with: "+current_examples_list[current_counter]+" time: "+
        str(datetime.datetime.now()))

    bl_counter=0
    bl_goal=len(stadt_list)
    quellbundesland="n.a."
    while bl_counter<bl_goal:
```

```
        if source_location==stadt_list[bl_counter]:
            quellbundesland=bundesland_list[bl_counter]
        bl_counter=bl_counter+1

    plz_counter=0
    plz_goal=len(stadt_list)
    quellplz="n.a."
    while plz_counter<plz_goal:
        if source_location==stadt_list[plz_counter]:
            quellplz=plz_list[plz_counter]
        plz_counter=plz_counter+1

    full_id_list=[]
    follower_list=[]
    location_list=[]
    strip_location_list=[]

    twitter=Twitter(auth = OAuth(access_key, access_secret, consumer_key,
        consumer_secret),retry=True)

    returned_cursor=-1
    while returned_cursor is not 0:
        follower_ids=twitter.followers.ids(screen_name=current_examples_list
            [current_counter], cursor=returned_cursor)
        current_ids=follower_ids["ids"]
        full_id_list=full_id_list+current_ids
        current_next_cursor = follower_ids["next_cursor"]
        print(current_next_cursor)
        returned_cursor=current_next_cursor

    newspaper_n_active_followers=0
    newspaper_n_loc_followers=0
    newspaper_n_loc_identified_followers=0

    loc_matching=0
    loc_non_matching=0
    bl_matching=0
    bl_non_matching=0
    plz_1000_greater=0
    plz_1000_smaller=0
    plz_10000_greater=0
    plz_10000_smaller=0

    for user_id in full_id_list:
        try:
            user_account=twitter.users.show(id=user_id)
        except:
            pass
        newspaper_n_active_followers=newspaper_n_active_followers+1
        if user_account["location"]!="":
            loc=user_account["location"]
            newspaper_n_loc_followers=newspaper_n_loc_followers+1
            user_geo_strippy=[]
            strippy_counter=0
            strippy_goal=len(stadt_list)
            while strippy_counter<strippy_goal:
                stadtstring=stadt_list[strippy_counter]
                user_geo_strippy=re.findall(r'\b'+stadtstring+r'\b',loc,
                    flags=re.I)
                if len(user_geo_strippy)>0:
```

```
                    strippy_counter=strippy_goal
                else:
                    strippy_counter=strippy_counter+1
            if len(user_geo_strippy)>0:
                loc_id=user_geo_strippy[0]
                newspaper_n_loc_identified_followers=
                    newspaper_n_loc_identified_followers+1

                user_plz_counter=0
                user_plz_goal=len(stadt_list)
                user_plz=quellplz
                while user_plz_counter<user_plz_goal:
                    if loc_id.lower()==stadt_list[user_plz_counter].lower():
                        user_plz=plz_list[user_plz_counter]
                    user_plz_counter=user_plz_counter+1

                user_bl_counter=0
                user_bl_goal=len(stadt_list)
                user_bl="n.a."
                while user_bl_counter<user_bl_goal:
                    if loc_id.lower()==stadt_list[user_bl_counter].lower():
                        user_bl=bundesland_list[user_bl_counter]
                    user_bl_counter=user_bl_counter+1


                if source_location.lower()==loc_id.lower():
                    loc_matching=loc_matching+1
                else:
                    loc_non_matching=loc_non_matching+1

                if quellbundesland==user_bl:
                    bl_matching=bl_matching+1
                else:
                    bl_non_matching=bl_non_matching+1

                plz_diff=abs(int(user_plz)-int(quellplz))
                if plz_diff>=1000:
                    plz_1000_greater=plz_1000_greater+1
                else:
                    plz_1000_smaller=plz_1000_smaller+1
                if plz_diff>=10000:
                    plz_10000_greater=plz_10000_greater+1
                else:
                    plz_10000_smaller=plz_10000_smaller+1



    if source_account["description"]!="":
        source_text=source_account["description"]
        source_text=source_text.replace('"',"'")
    else:
        source_text="n.a."

    coded_at=str(datetime.datetime.now())
    source_name=source_account["name"]
    source_username=source_account["screen_name"]
    n_tweets=source_account["statuses_count"]
    created_at=source_account["created_at"]
    user_id=source_account["id_str"]
```

```
#path must match an existing folder (output folder)
    with open("path.csv",mode="a",encoding="utf-8")as fo:
        writer=csv.writer(fo)
#erase after first use to avoid multiple headline rows

            writer.writerow(("couded_at","source_name","source_username","so
            urce_text","source_location","source_bl","source_plz","n_followe
            rs_active","n_tweets","created_at","user_id","n_loc_follower","n
            _loc_follower_identified","loc_matching","loc_non_matching","bl_
            matching","bl_non_matching","plz_1000_greater","plz_1000_smaller
            ","plz_10000_greater","plz_10000_smaller"))
        writer.writerow((coded_at,source_name,source_username,source_text,
            source_location,quellbundesland,quellplz,
            newspaper_n_active_followers,n_tweets,created_at,user_id,
            newspaper_n_loc_followers,newspaper_n_loc_identified_followers,
            loc_matching,loc_non_matching,bl_matching,bl_non_matching,
            plz_1000_greater,plz_1000_smaller,plz_10000_greater,
            plz_10000_smaller))

    print("Done with: "+current_examples_list[current_counter]+" time: "+str
        (datetime.datetime.now()))
    current_counter=current_counter+1

print("done")
```