1. Download the twitter script from the following link & look at the script.

a. How does the program work?

Ans. The program is in python. First it install fresh packages, then it goes for collecting tweepy, once the downloading is done it finishes with the status done. Then it is stored in the directory and the installation is done successfully, then it ask for enter the search string.

b. How do you think you can use this code?

Ans. The code is asking for Access Token , Access Secret, as the security identifier of user account , consumer key , consumer secret then entering the searching string from the date to the data. Then it ask for number of tweets in order to retrieve for search string. Then the api is required. Then I is initialized to 0 ,then a for loop is run, then i+=1, followed by res.user.screen\_name, then returns the string representing the date then fetches res.text.replace after that the followers count and closing of the write portion and then printing Tweets retrieved.

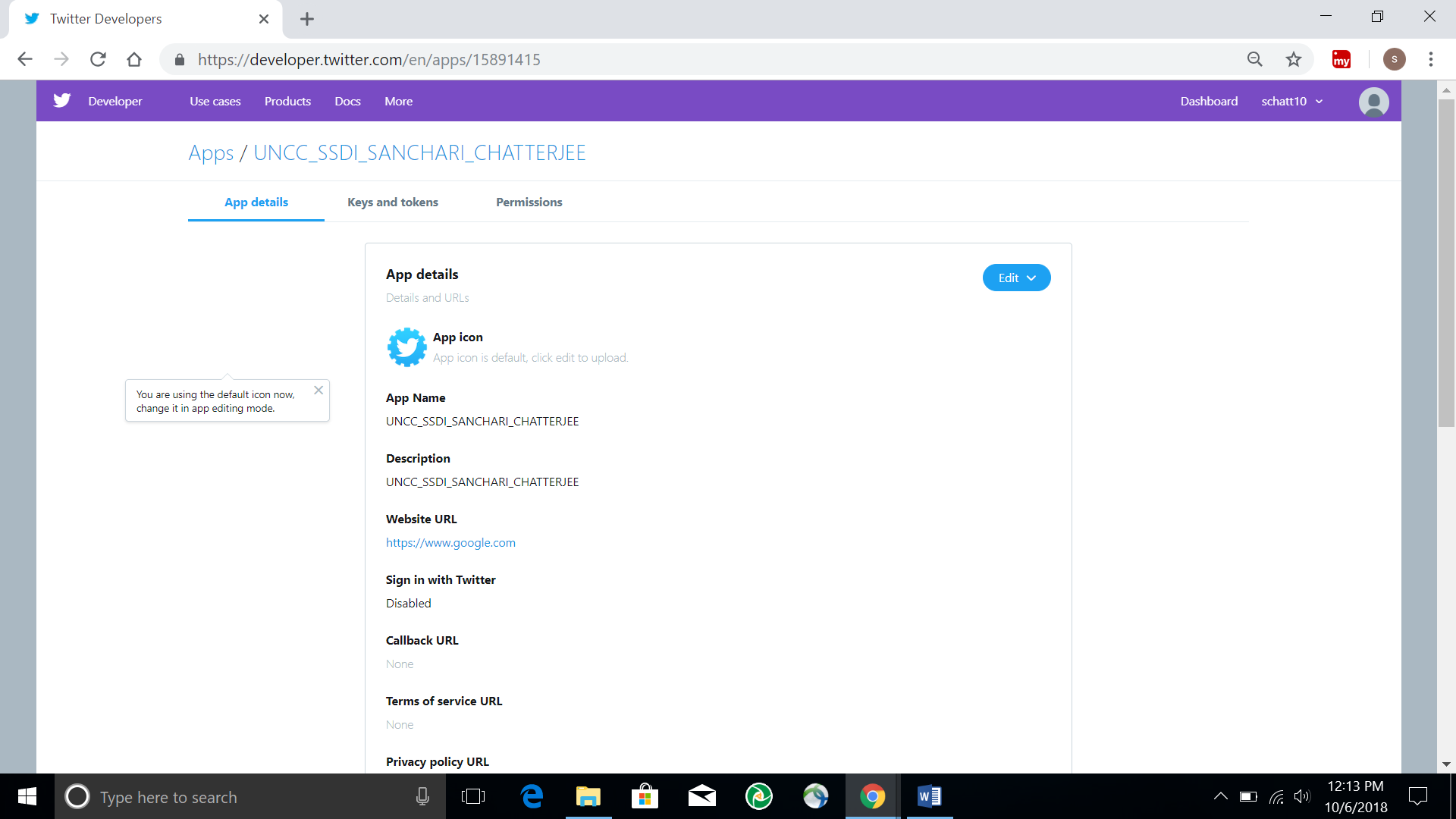
c. Can you think of different scenarios where this code could be used for data

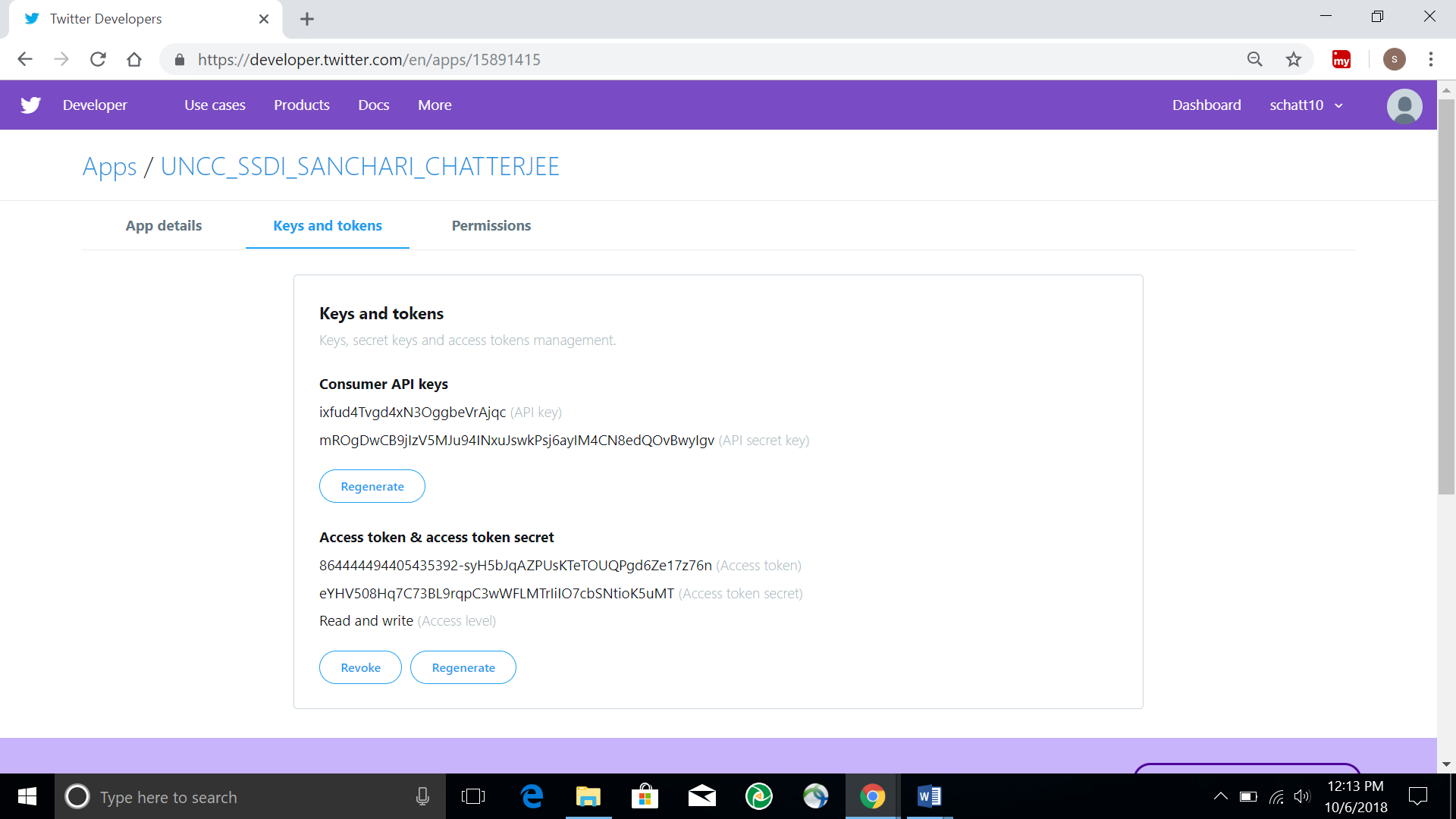
collection?

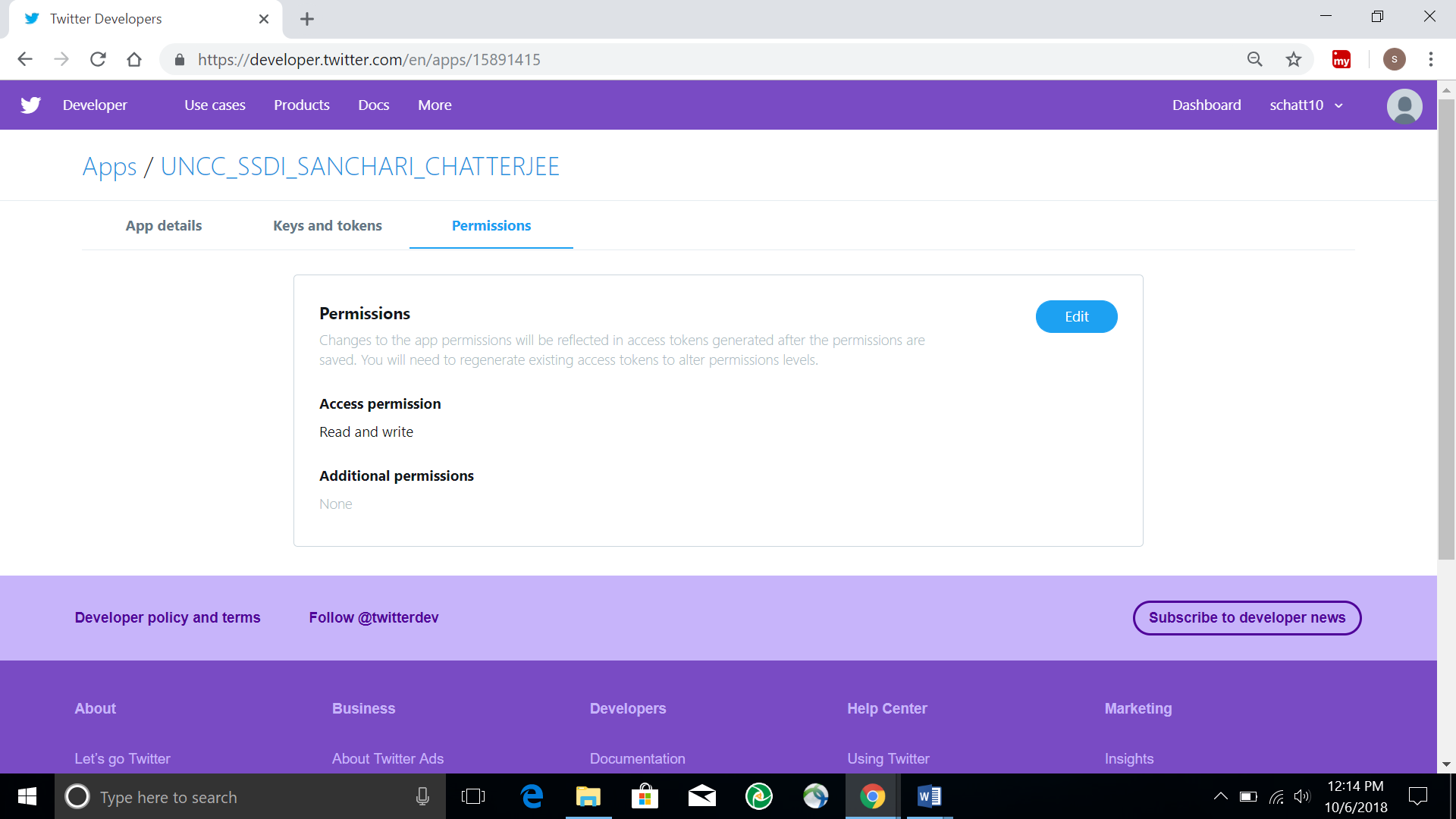
Ans. The scenarios are:

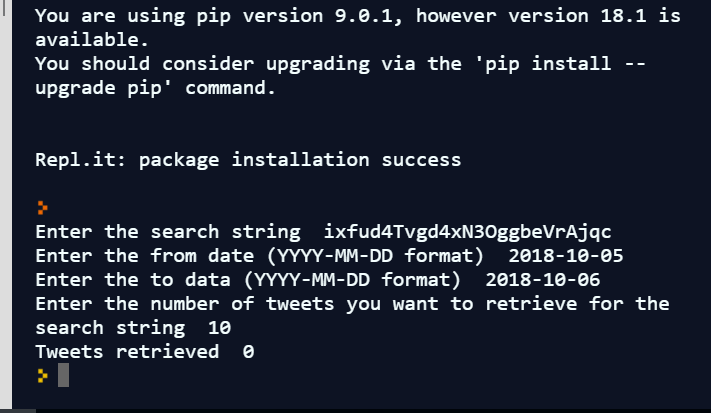
1. To mark some people and follow them whether those people are following us or not.

2. Tracking the account who is having enough followers and friends to follow them.









**Pseudo code:**

import sys

import operator

import string

// defining most\_Twitter\_Users():

// Taking the input file with the input location specification

INPUTFILE = input(C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Most\_Twitter\_Users.txt')

// Input filepath to be included with ‘.txt’ format

INPUT\_FILE\_PATH = INPUTFILE+ '.txt'

// Twitter API reading files

twitter=myFile.readlines()

// running the for loop finding the data

for data in twitter:

fileTemp = data.split()

if file\_Temp[0] in l:

l[file\_Temp[0]] +=1

else:

l[file\_Temp[0]] = 1

// Sorting the file based on the item generated

l = sorted(l.items(), key = operator.itemgetter(1), reverse = True)

// The output file to open for this twitter segment

outputFile = open(C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Most\_Recent\_Users.txt', 'w', encoding = "utf-8")

// Finding out the top 10 users tweeting the most for the entire timeline

outputFile.write("Top 10 users tweeting the most in the timeline durtion:")

for i in range (0,10):

outputFile.write("The user " + l[i][0] + " tweeted " + str(l[i][1]) + " times" + "\n")

// Closing the output file - outputFile.close

// Defining the most users per hour in the twitter

def mostUsersPerHour():

INPUTFILE = input(C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Most\_Twitter\_Users\_Per\_Hour.txt')

// Input filepath to be included with ‘.txt’ format

INPUT\_FILE\_PATH = INPUTFILE+ '.txt'

// Splitting the twitter text, swapping the position of index and sorting the user numbers

for dat in data:

file\_Temp = dat.split()

file\_Temp2 = file\_Temp[1].split(":")

twitter\_Temp = file\_Temp[0] + " " + file\_Temp2[1]

if twitter\_Temp in l:

l[twitter\_Temp]+=1

else:

l[twitter\_Temp]=1

l = sorted(l.items(), key = operator.itemgetter(1), reverse = True)

total\_Entries\_Print = 10\*length(l5)

//

outputFile = open('C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Most\_Twitter\_Users\_Per\_Hour.txt', 'w', encoding = 'latin-1')

outputFile.close

mostUsersPerHour()

// Defining the maximum followers in the twitter

def maxFollowers():

INPUTFILE = input('C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Max\_Twitter\_Follower.txt')

// Input filepath to be included with ‘.txt’ format

INPUT\_FILE\_PATH = INPUTFILE+ '.txt'

// Finding the data and splitting through the twitter text

for data in twitter:

fileTemp = dat.split()

if fileTemp[0] not in l:

l[fileTemp[0]] = int(fileTemp[-4])

// Finding the top 10 users who has the maximum number of followers

outputFile = open('C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Max\_Twitter\_Follower.txt', 'w', encoding = "latin-1")

outputFile.write("The top 10 users who has the maximum number of followers:")

// Fiding the top 10 users name with the following attributes

for i in range (0, 10):

outputFile.write(str(i+1) + ". Username: " + l[i][0] + " Total Number of Followers: " + str(l[i][1]))

outputFile.close

maxFollowers()

// Finding the maximum number of retweets in this twitter segment

def maxRetweet():

INPUTFILE = input('C:/Users/schatterjee/Desktop/UNCC SSDI Fall 2018 /TwitterTextAnalysis/Max\_Re\_Twitter.txt')

// Input filepath to be included with ‘.txt’ format

INPUT\_FILE\_PATH = INPUTFILE+ '.txt'

for dat in twitter:

file\_Temp = dat.split()

y = len(file\_Temp)-2

tweet = "\""

for x in range (4, y):

tweet += file\_Temp[x] + " "

tweet += " ::::;:::: " + file\_Temp[0]

if tweet not in l:

l[tweet] = int(file\_Temp[-1])

outputFile = open('C:/Users/VenksUV/Desktop/6112 - SSDI/TwitterTextAnalysis/MaxRetweets.txt', 'w', encoding = "utf-8")

outputFile.write("The top 10 tweets having the maximum number of retweets: ")

for x in range (0, 10):

outputFile.write(str(x+1) + ". Username: " + l[x][0].split()[-1] + " Tweet: " + l[x][0].split(":;:")[0] + "No of retweets: " + str(l[x][1])

outputFile.close

maxRetweet()