Demographics of sentiment driven users on Twitter

The write up includes the files which we made during development of project.
Demographics and sentiments are extracted using some APIs.
The files description and codes are attached.

1.readmongo.py

A dictionary is created where all the words of database are stored.
We implemented a dictionary where we stored each word of database and its frequency.
The specific professions are compared to this dictionary and output.txt has output attached.
By comparison various demographics such as profession, gender are extracted using dictionary.

2. writedatabase.py

We extracted tweets using Twitter Stream API. We calculated the sentiments of each tweet using textblob and stored those tweets who polarity was non-zero. For each such tweet we then extracted the details of the user like screen name, location, follower's count, following, verification status, etc using twitter Rest API and all the details were stored in the mongodb database.

3. Record_db.json

This file includes the stored database of mongodb in json format. One of the record is:

```
{
"_id" : ObjectId("599c3986fa90c42a64f70a9e"),
"lang" : "pt",
"verified" : false,
"screen_name" : "lucasangelo_",
"friends_count" : 795,
"tweet" : "RT @guilhermemendss: OBRIGADOOOO OBRIGADOOOO VALEU MESMO MUITOO
OBRIGADO THANKSSS THANKSSS SO MUCH GRACIAS MUITAS GRACIAS ARIGATOU OBG O… ",
"followers_count" : 3200,
"location" : "Brasília, Brasil",
"statuses_count" : 109694,
"user_id" : "187531766",
"description" : "Antigamente eu sabia exatamente o que fazer"
}
```

4. output.txt

This file contains the number of people having gender and common profession. This file is output of readmongo.py.

References:

https://dev.twitter.com/

https://stackoverflow.com/

http://precog.iiitd.edu.in/teaching/psosmonnptel/

https://www.codecademy.com/