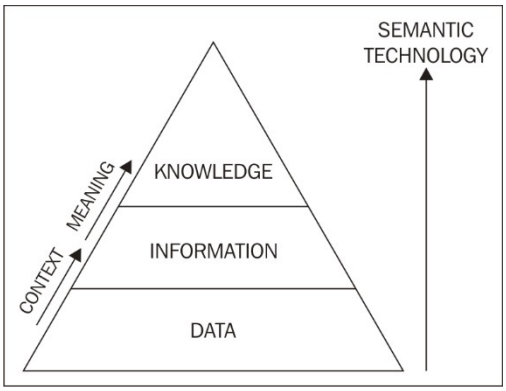
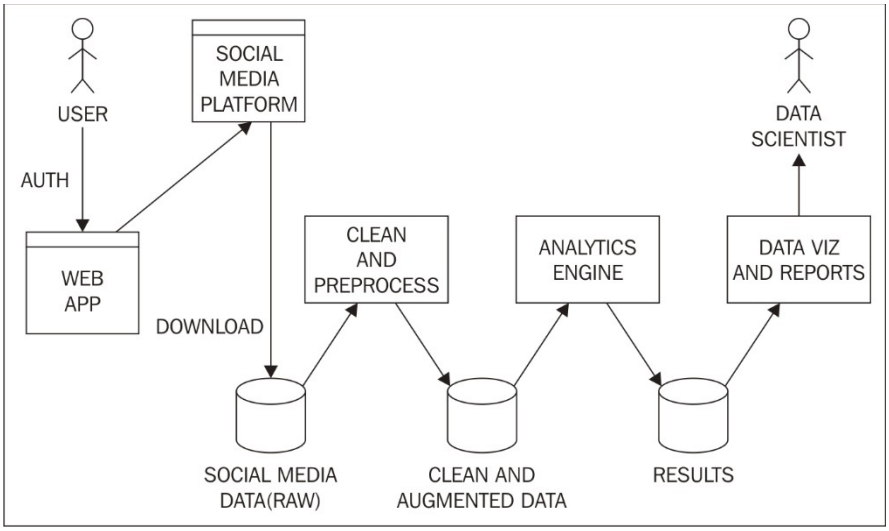
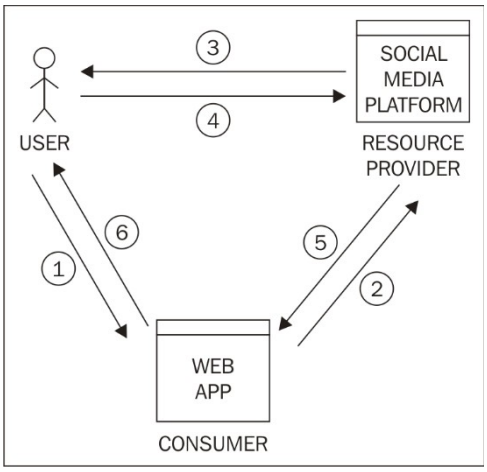
Chapter 1: Social Media, Social Data, and Python







Intro to Text Mining

Test OAuth

Details | Settings | Keys and Access Tokens | Permissions

Application Settings

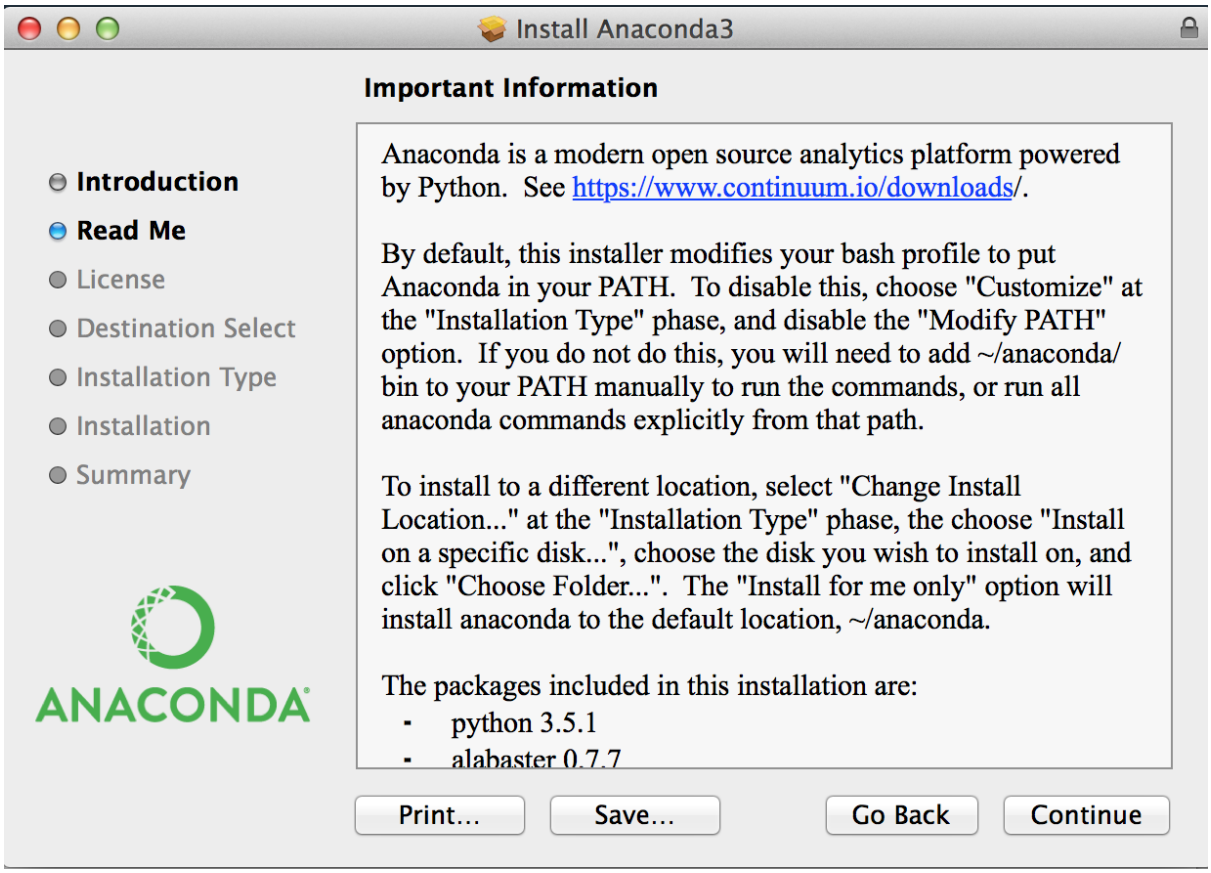
Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

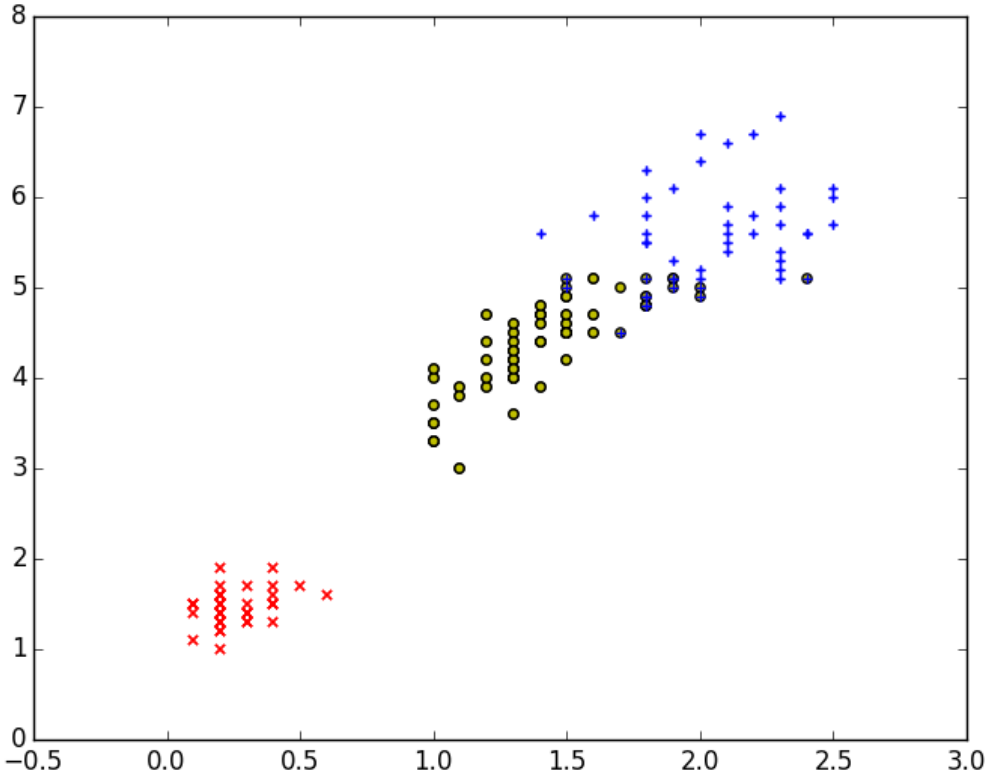
Consumer Key (API Key)	[redacted]
Consumer Secret (API Secret)	[redacted]
Access Level	Read-only (modify app permissions)
Owner	marcobonzanini
Owner ID	19018614

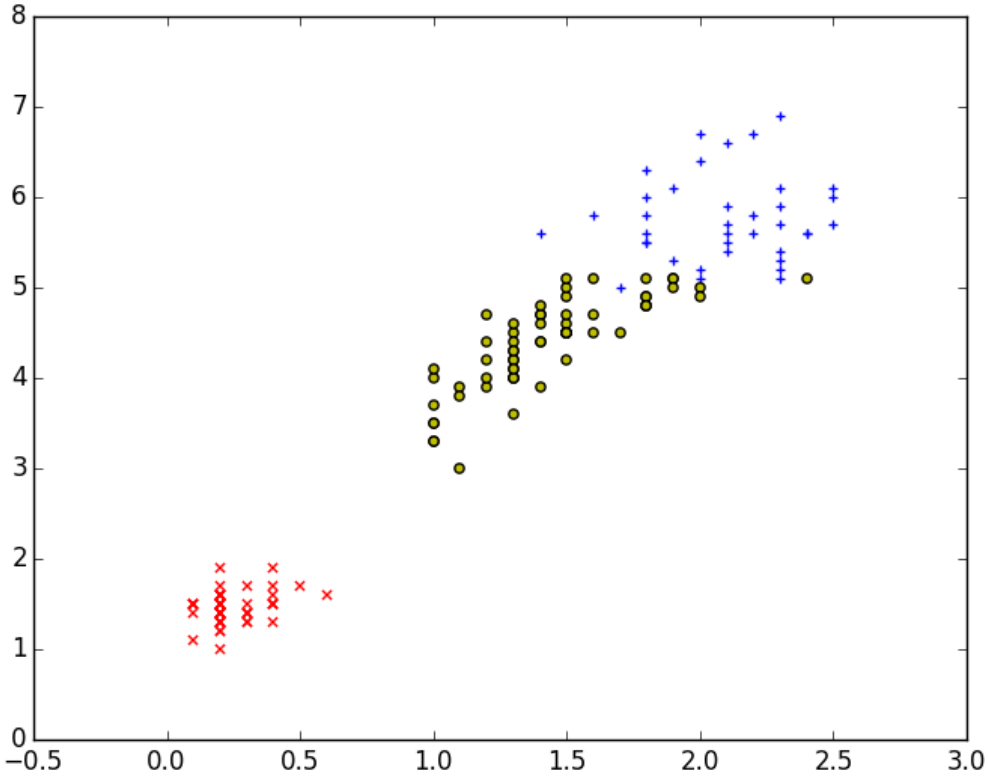
Application Actions

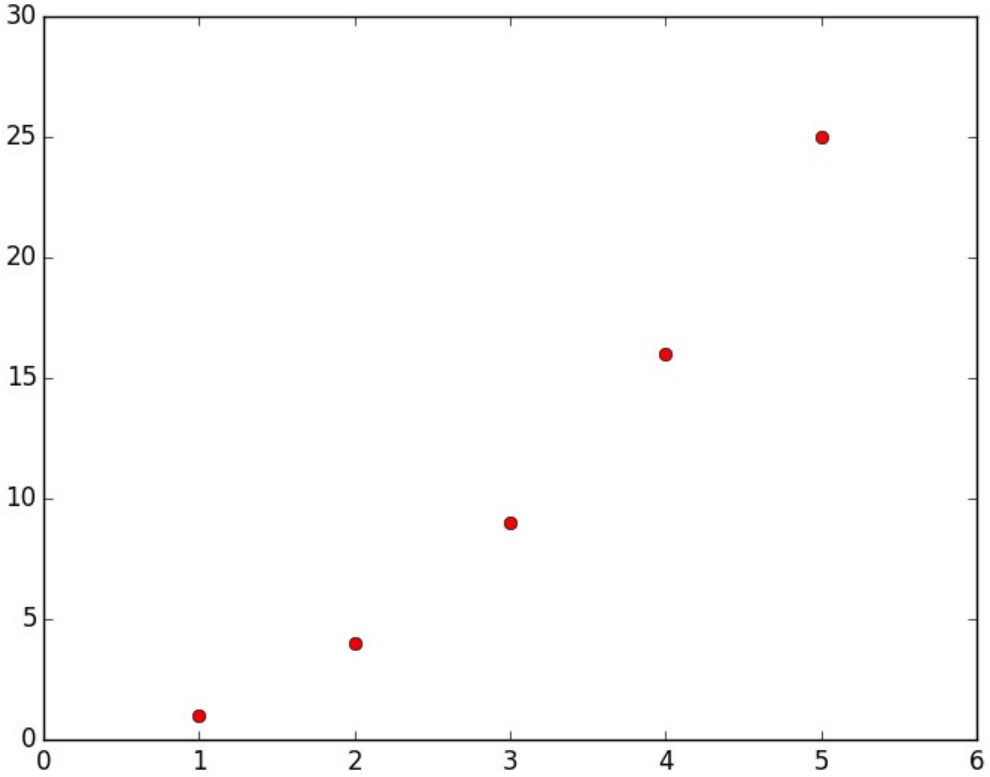
Regenerate Consumer Key and Secret | Change App Permissions

Your Access Token




8
7
6
5
4
3
2
1
0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0


8
7
6
5
4
3
2
1
0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0

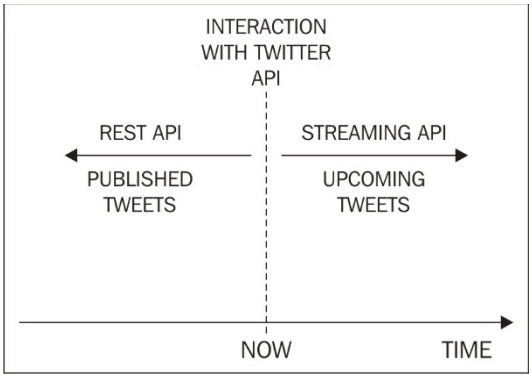


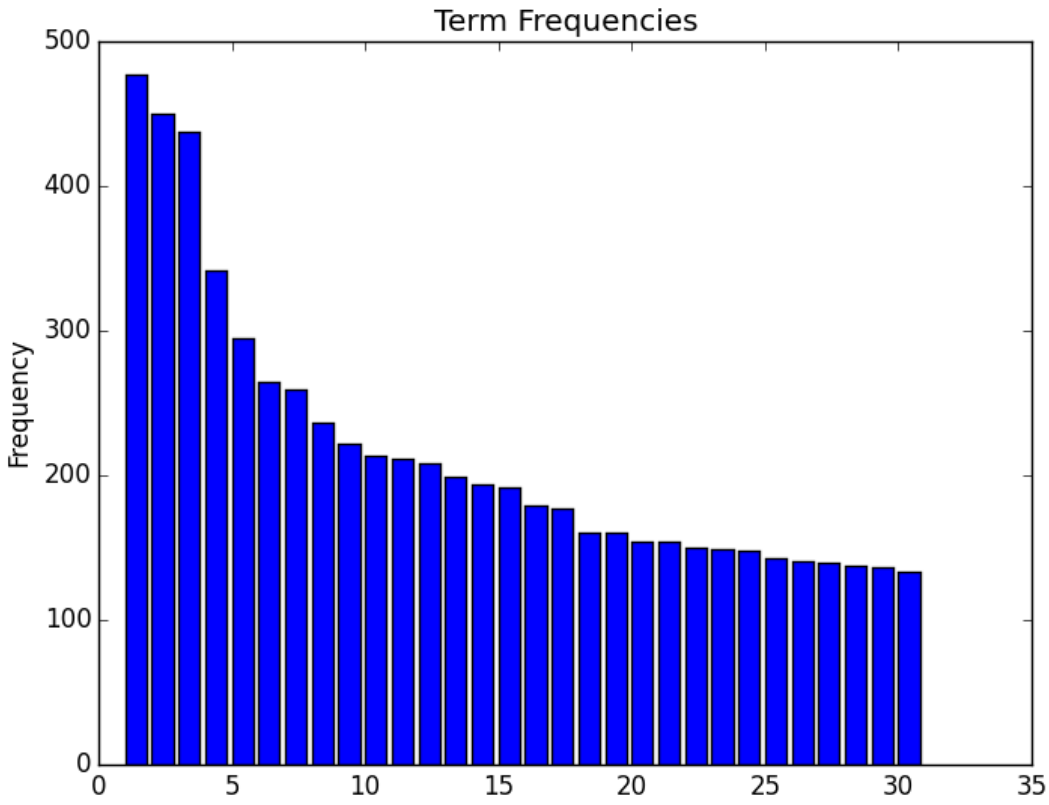
jsonviewer.stack.hu

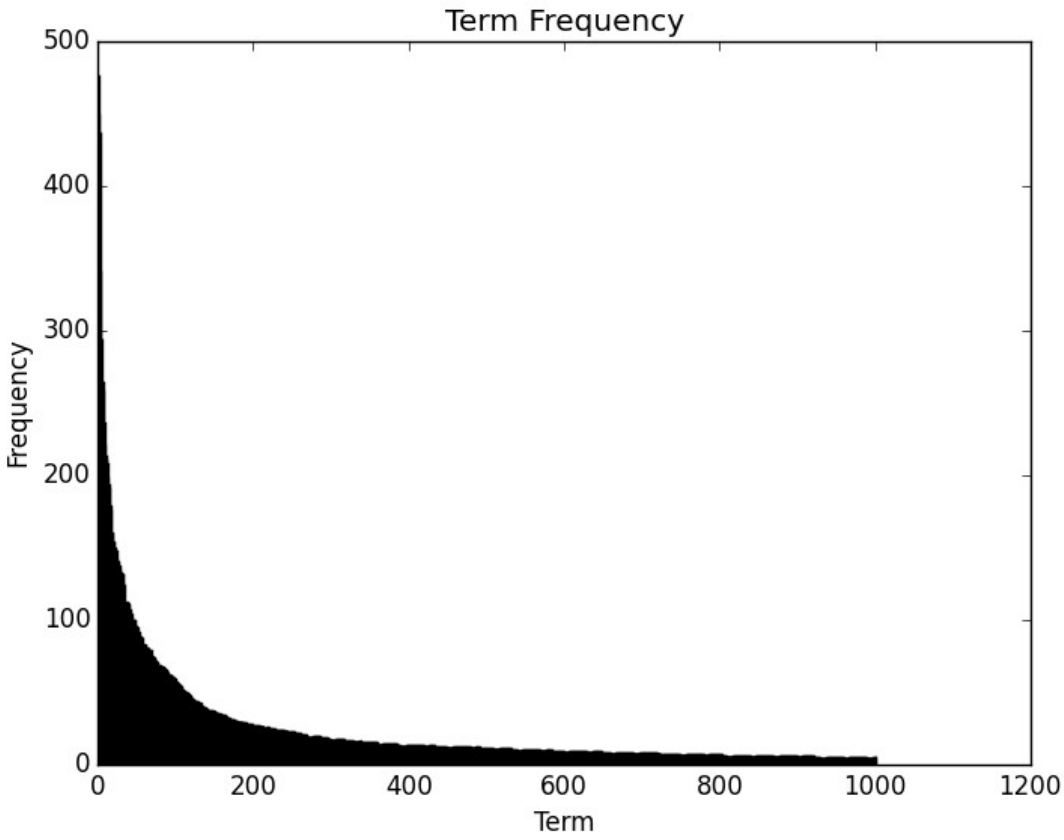
Viewer | Text

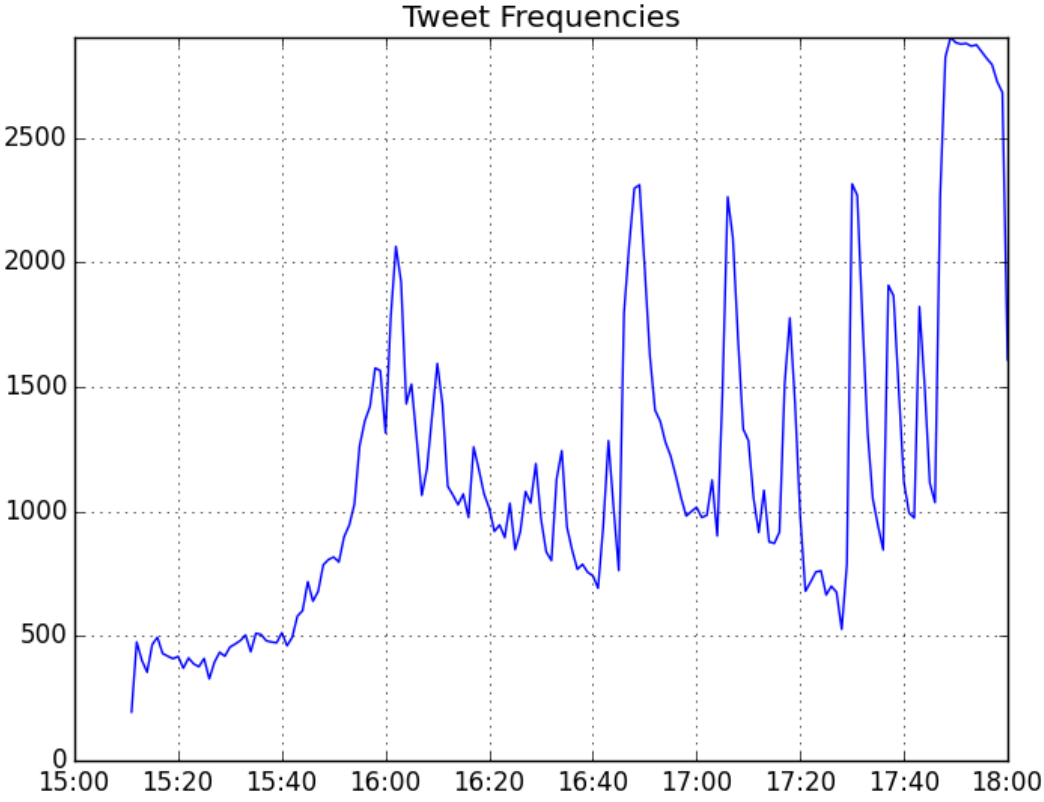
- JSON
  - user_id : "1"
  - name : "Marco"
  - likes
    - 0 : "Python"
    - 1 : "Data Mining"

Chapter 2: #MiningTwitter – Hashtags, Topics, and Time Series

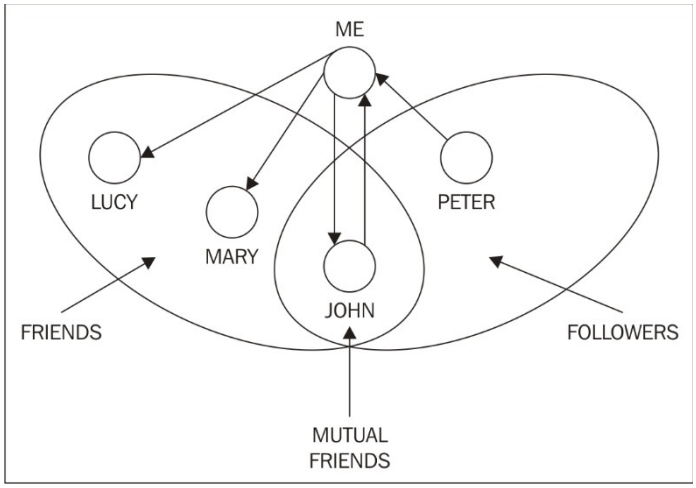




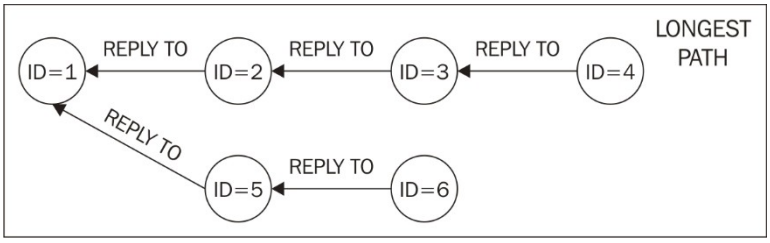

Term Frequency
Frequency
Term
0
100
200
300
400
500
0
200
400
600
800
1000
1200

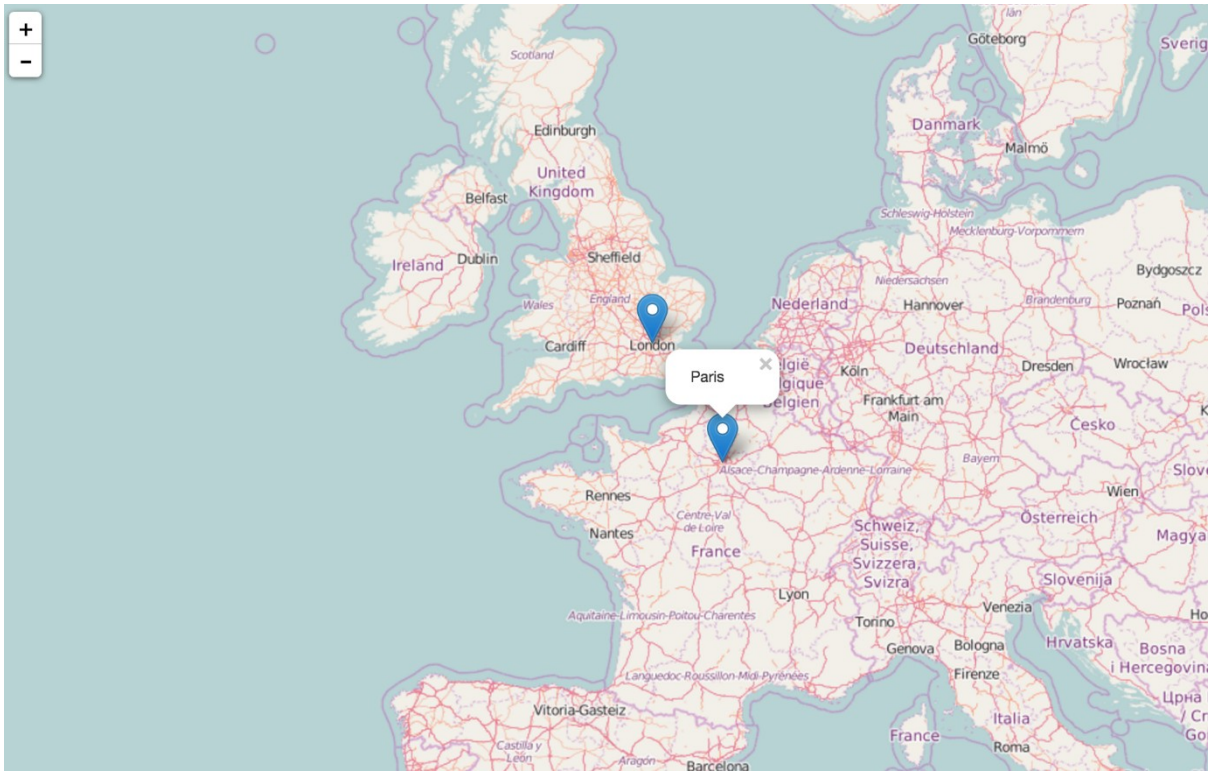

Tweet Frequencies
0
500
1000
1500
2000
2500
15:00
15:20
15:40
16:00
16:20
16:40
17:00
17:20
17:40
18:00

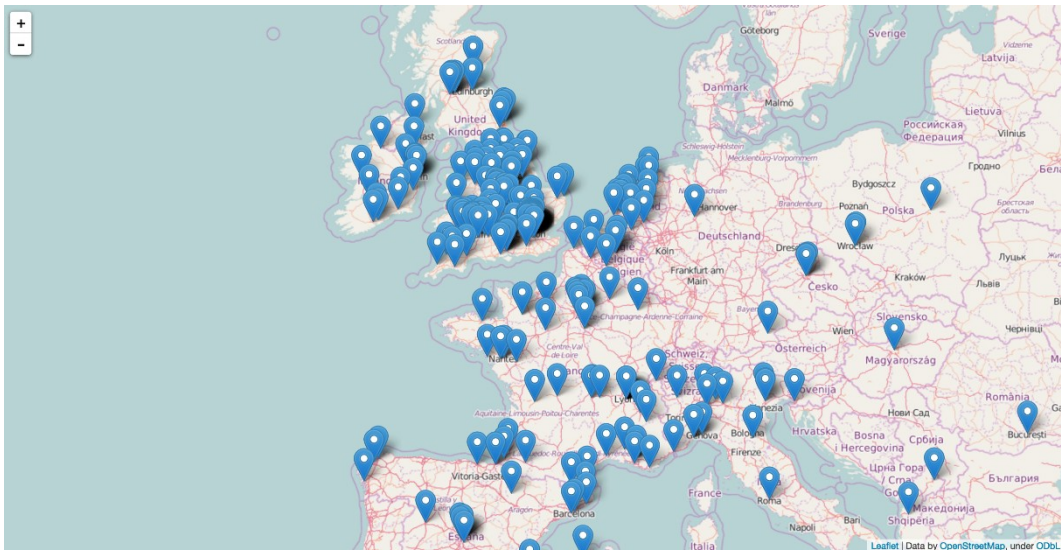
Chapter 3: Users, Followers, and Communities on Twitter

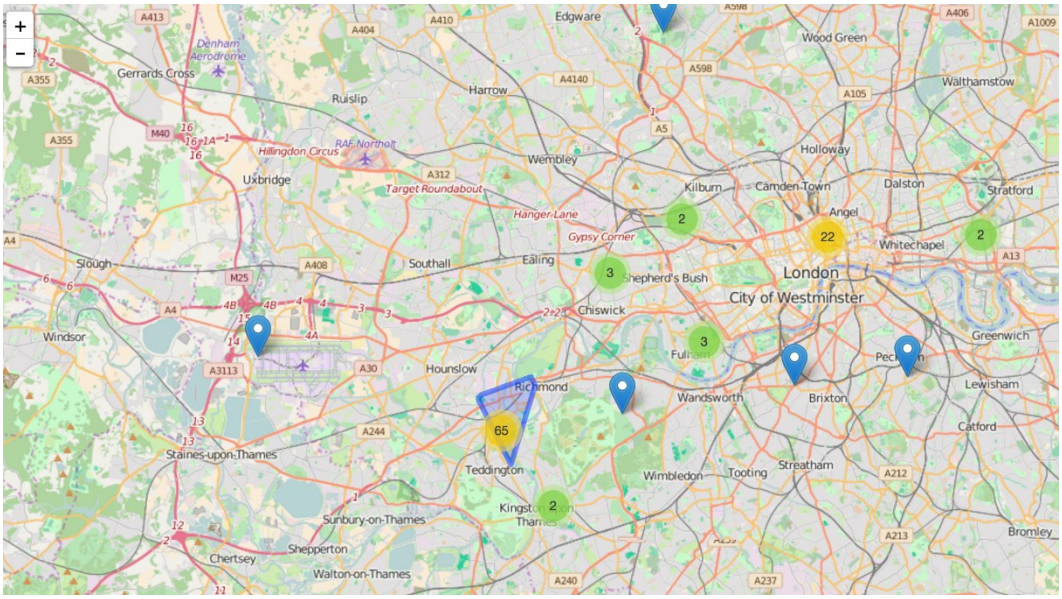


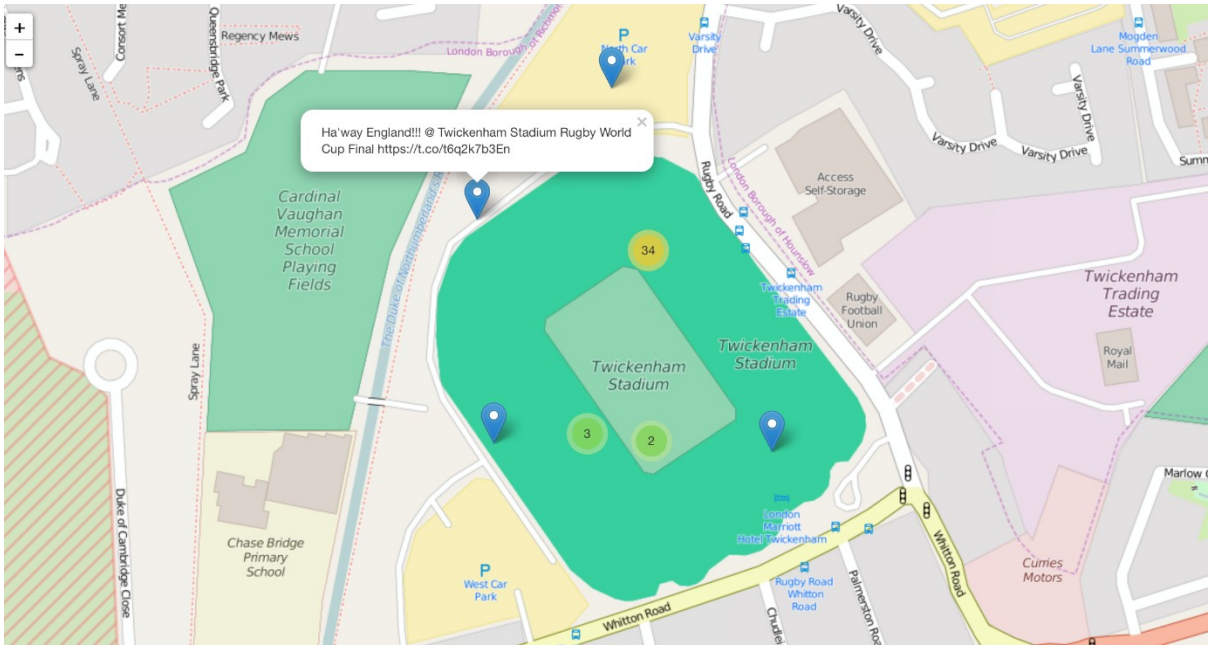

```
# ---------- Cluster 0
PhD Student #FutureOfNews, #TextMining, Topic Detection & Tracking,
Summarization, Information Retrieval, mostly with Breaking News
Data
PhD candidate - How and what can search engines learn from their
users?
PhD candidate. Information Retrieval. Data Mining. Open Source.
PhD Candidate | Semantic Search in eDiscovery | Information
Retrieval, Text Mining
# ---------- Cluster 1
Co-fundador de Viz Analytics. Diplomado en Ingeniería Informatica
de Gestión y con un Master Universitario en Inteligencia
Artificial.
Lenguas, comunicación y nuevos medios. No tuitear en caso de
incendio.
Lic. en Administración mención informática, promotor del software
libre y las oportunidades que brinda.
# ---------- Cluster 2
I am part time web-developer and I love python
boiling down to Python and Infosec nowadays. Any advices?
Python Padawan, Django noob.
Sherlock of #Data, Love #machinelearning #python
# ---------- Cluster 3
Data Scientist
Data scientist in the making
Data scientist, interested in text analysis, social networks,
artificial and natural intelligence, data, complexity, illusions...
Data Scientist
Wanna be a Data Scientist: using data for good
```



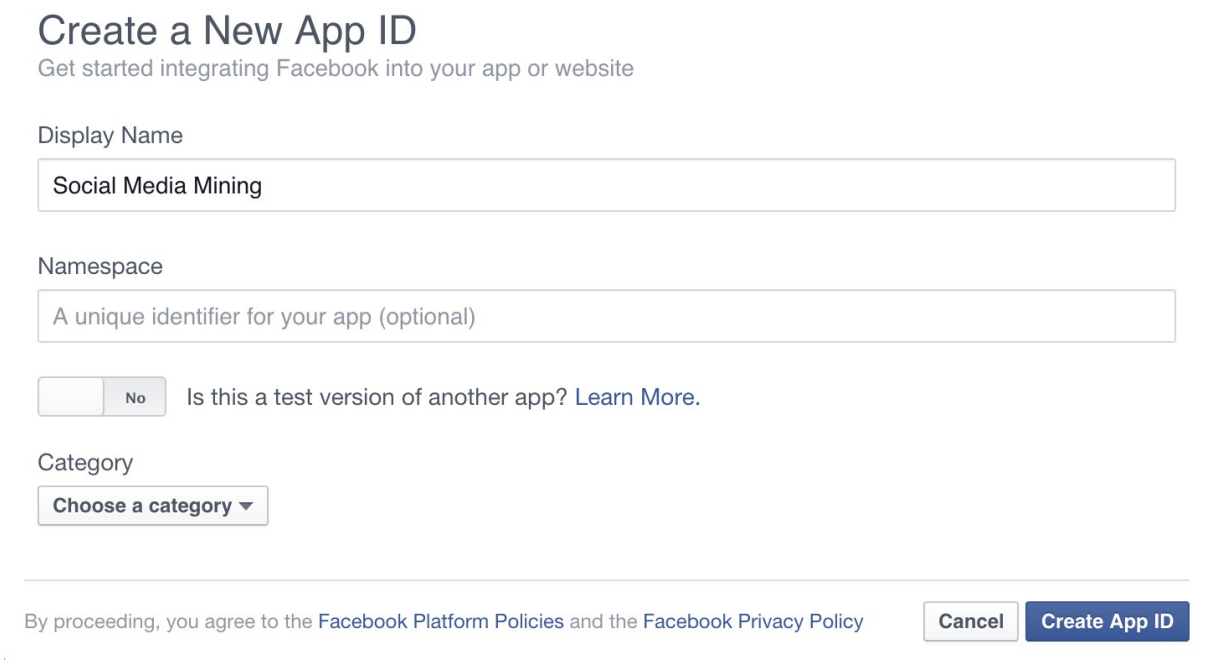

Paris
Edinburgh
United Kingdom
Belfast
Ireland
Dublin
Sheffield
Cardiff
London
Nederland
Hannover
Deutschland
Köln
Dresden
Wrocław
Frankfurt am Main
Česko
Rennes
Nantes
France
Schweiz, Suisse, Svizzera, Svizra
Österreich
Wien
Slovenija
Lyon
Venezia
Torino
Genova
Bologna
Hrvatska
Firenze
Vitoria-Gasteiz
Barcelona
Italia
Roma

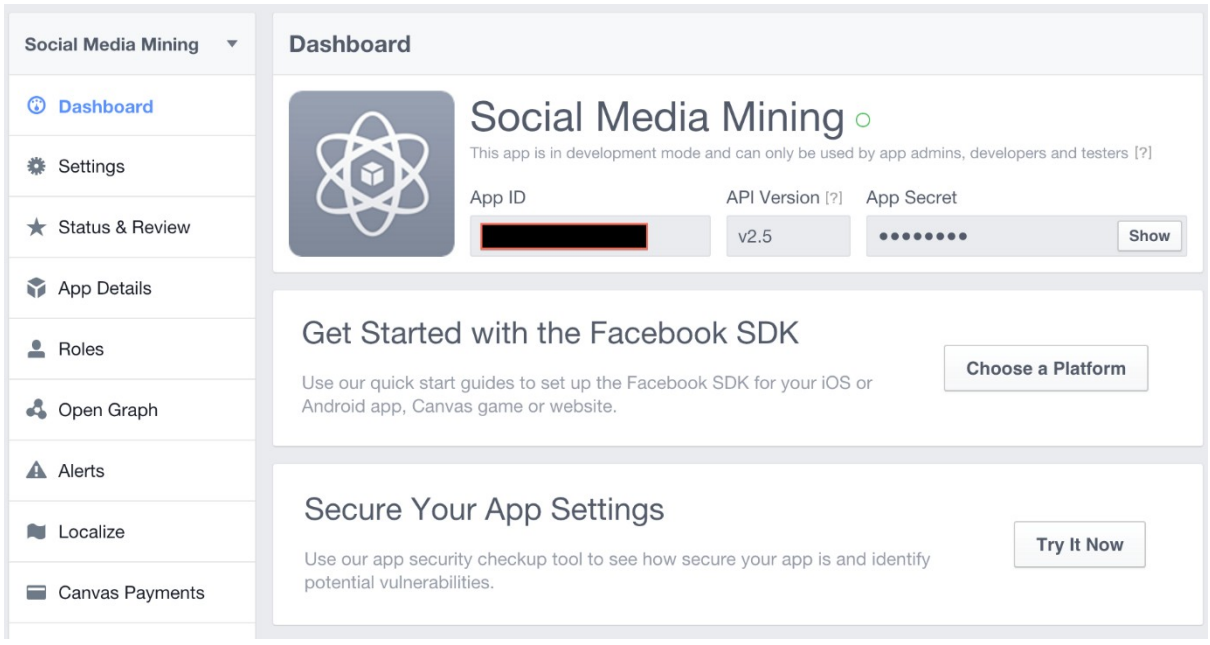

Leaflet | Data by OpenStreetMap, under ODbL


Edgware
Wood Green
Walthamstow
Gerrards Cross
Denham Aerodrome
Ruislip
Harrow
Hillingdon Circus
RAF Northolt
Uxbridge
Wembley
Holloway
Kilburn
Camden Town
Dalston
Stratford
Target Roundabout
Hanger Lane
Gypsy Corner
Angel
Whitechapel
Slough
Southall
Ealing
Shepherd's Bush
London
City of Westminster
Chiswick
Windsor
Greenwich
Hounslow
Richmond
Wandsworth
Brixton
Lewisham
Catford
Staines-upon-Thames
Teddington
Wimbledon
Tooting
Streatham
Sunbury-on-Thames
Shepperton
Chertsey
Walton-on-Thames
Bromley
65
2
3
22


Ha'way England!!! @ Twickenham Stadium Rugby World Cup Final https://t.co/t6q2k7b3En
Regency Mews
Queensbridge Park
Spray Lane
Cardinal Vaughan Memorial School Playing Fields
Twickenham Stadium
Twickenham Stadium
Access Self-Storage
Rugby Road
Twickenham Trading Estate
Rugby Football Union
Royal Mail
Varsity Drive
Mogden Lane Summerwood Road
Chase Bridge Primary School
Duke of Cambridge Close
West Car Park
Whitton Road
Palmerston Road
Curries Motors
34
3
2

Chapter 4: Posts, Pages, and User Interactions on Facebook





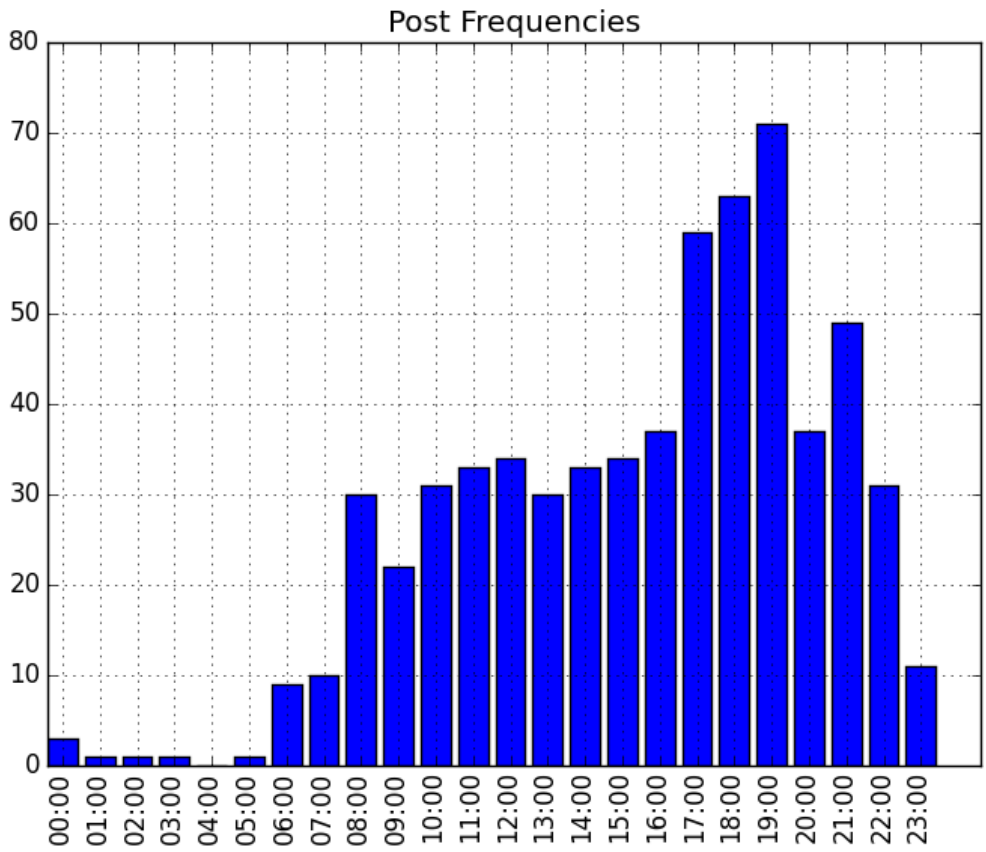
Select Permissions

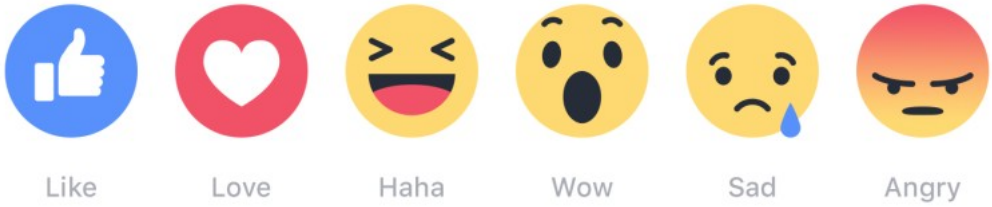
User Data Permissions | Extended Permissions

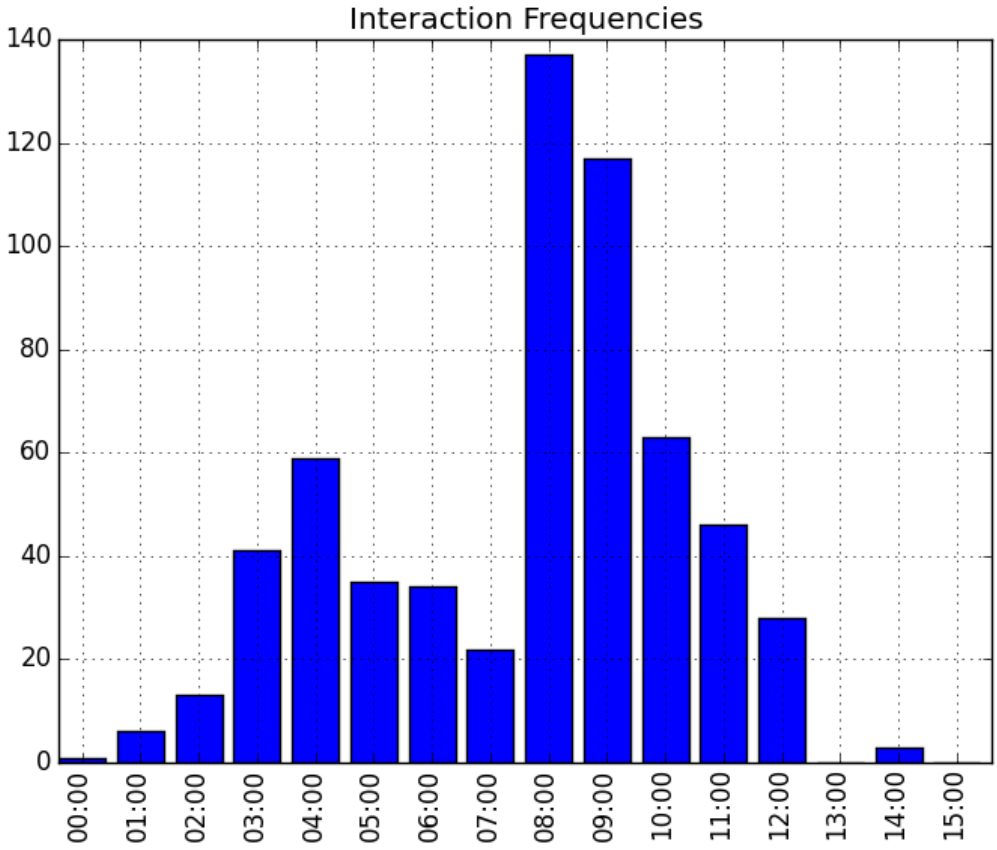
- [x] user_about_me
- [ ] user_actions.books
- [ ] user_actions.fitness
- [ ] user_actions.music
- [ ] user_actions.news
- [ ] user_actions.video
- [ ] user_birthday
- [ ] user_education_history
- [ ] user_events
- [x] user_friends
- [ ] user_games_activity
- [x] user_hometown
- [x] user_likes
- [x] user_location
- [ ] user_managed_groups
- [ ] user_photos
- [ ] user_posts
- [ ] user_relationship_details
- [x] user_relationships
- [ ] user_religion_politics
- [x] user_status
- [x] user_tagged_places
- [ ] user_videos
- [ ] user_website
- [ ] user_work_history

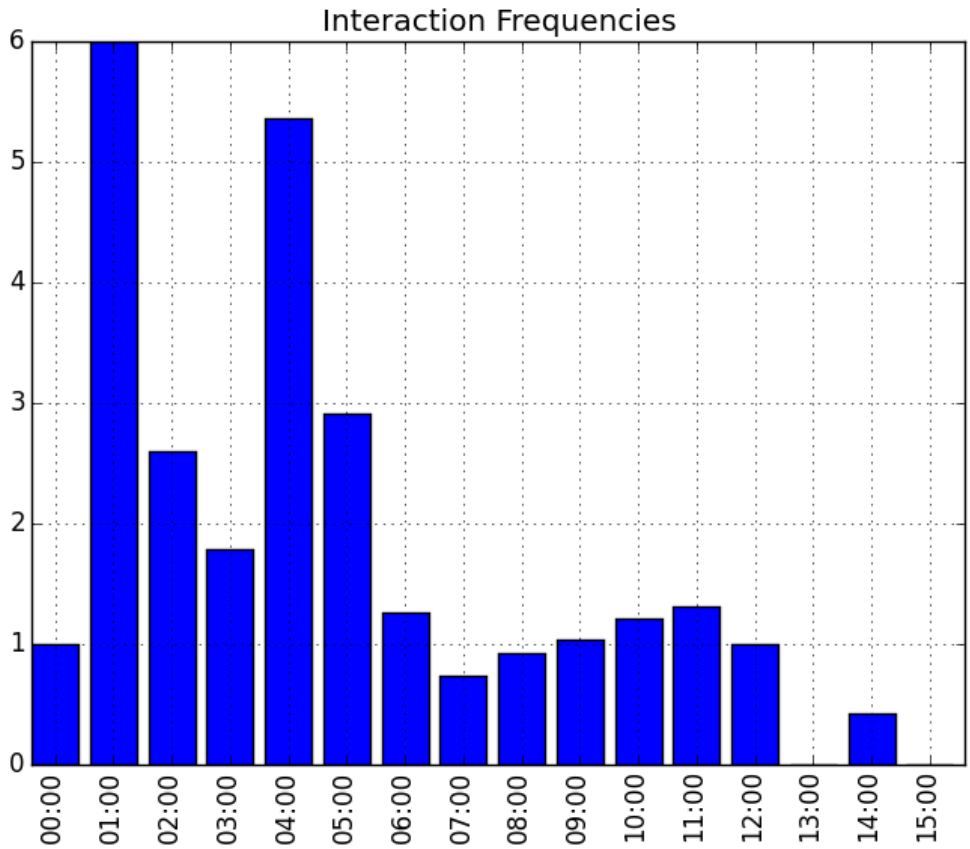
Public profile included by default.

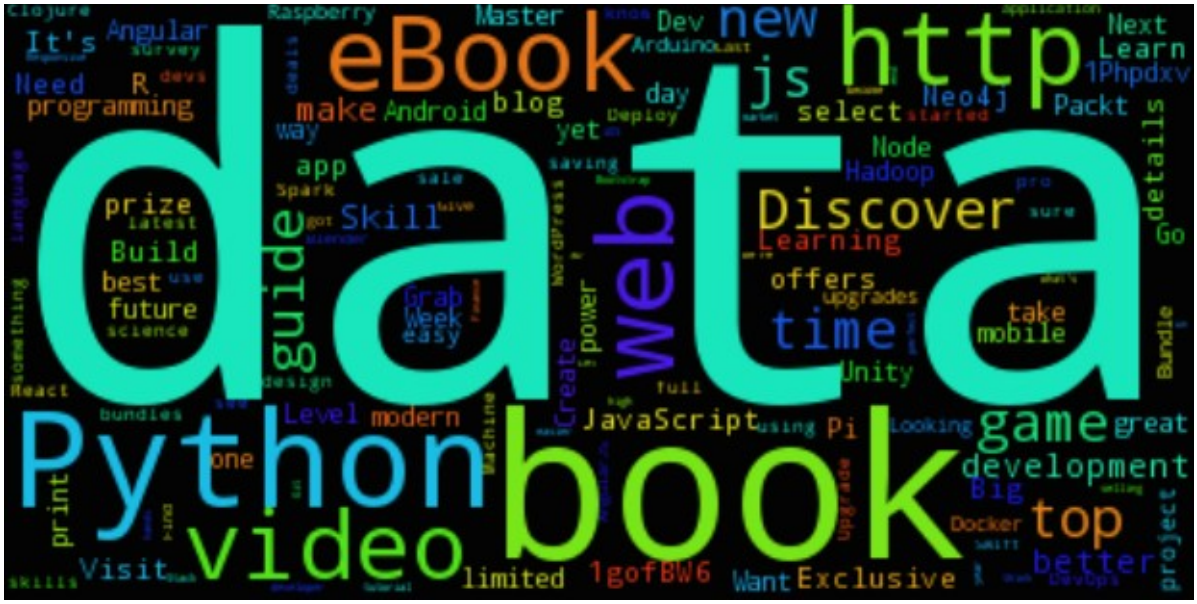
Get Access Token | Clear | Cancel



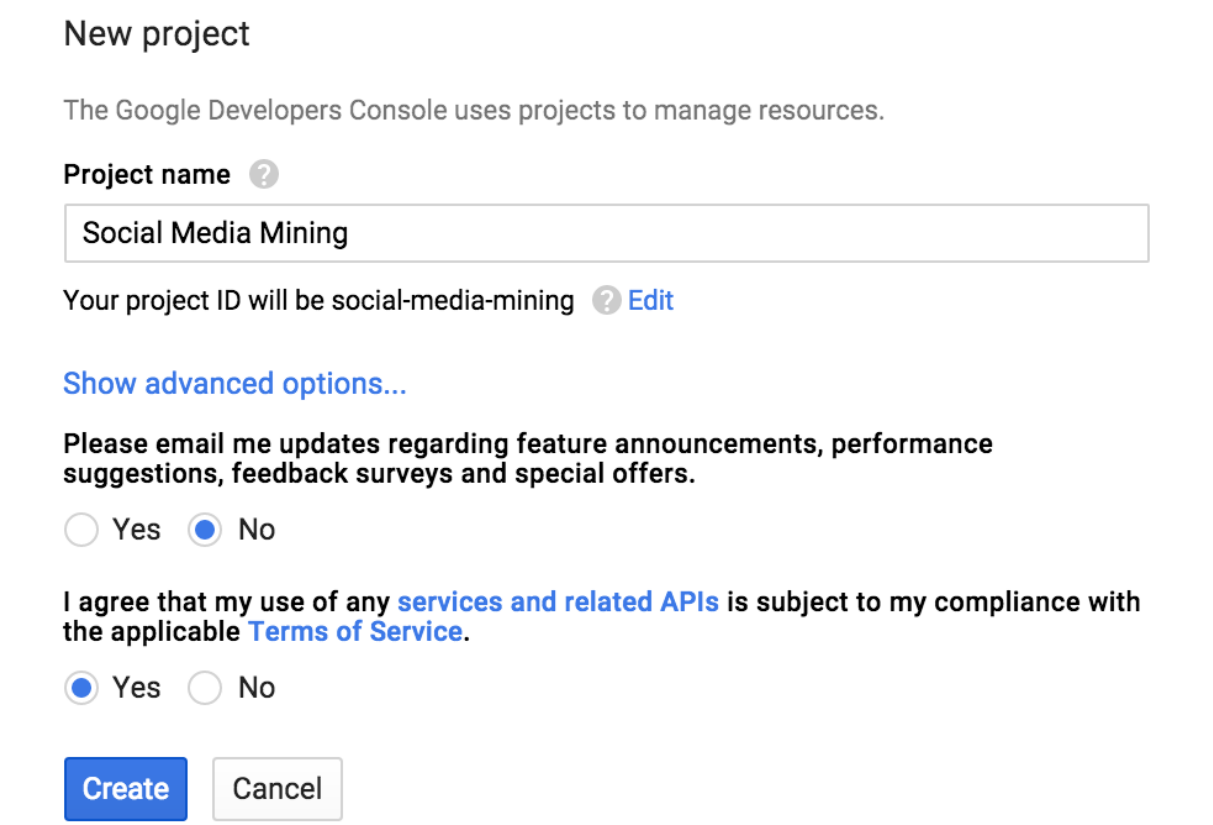

Like
Love
Haha
Wow
Sad
Angry

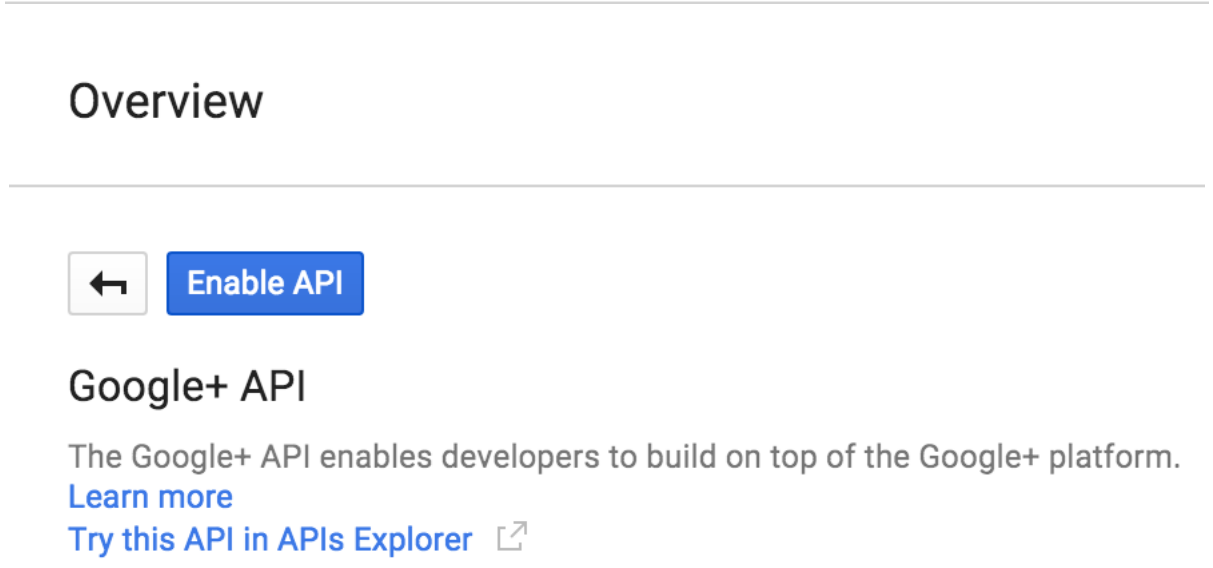

Interaction Frequencies
140
120
100
80
60
40
20
0
00:00
01:00
02:00
03:00
04:00
05:00
06:00
07:00
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00

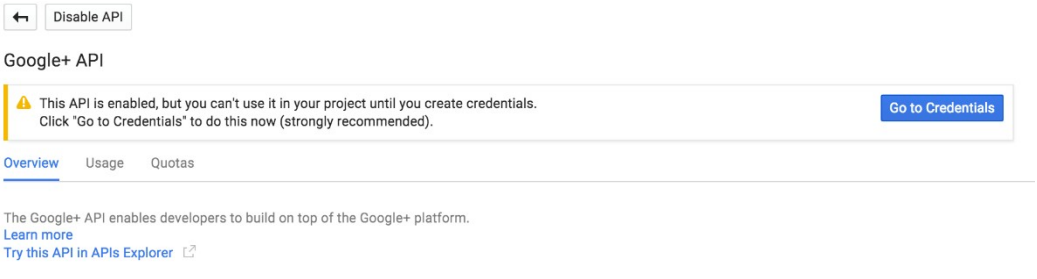

Interaction Frequencies
0
1
2
3
4
5
6
00:00
01:00
02:00
03:00
04:00
05:00
06:00
07:00
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00

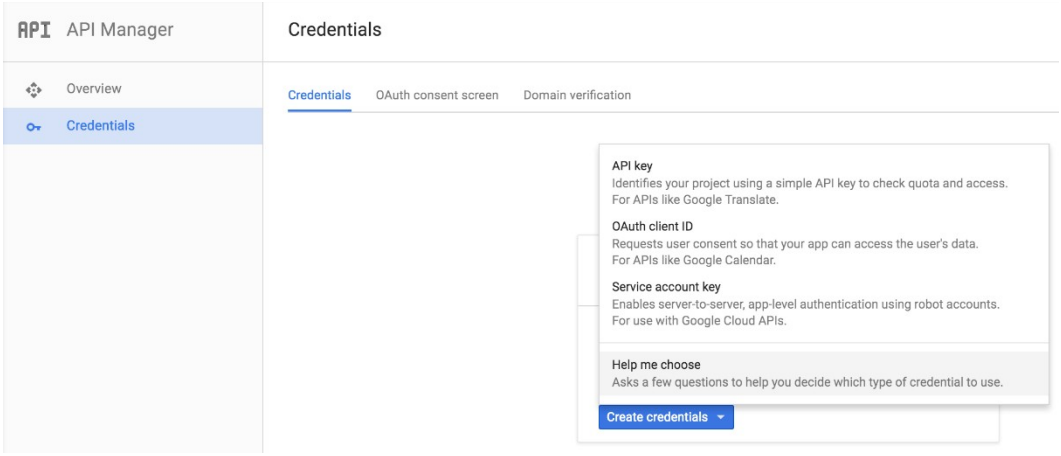

data
book
Python
video
eBook
http
web
Discover
time
game
top
development
Skill
guide
Learning
offers
JavaScript
Exclusive
Want
limited
Visit
print
Build
best
future
prize
programming
Need
Angular
It's
make
Android
blog
app
select
Node
Hadoop
Packt
Next
Learn
Go
take
mobile
Unity
Level
modern
one
great
Big
Docker
better
project
Create
power
Master
Dev
new
js
day
yet
way
Neo4j
1Phpdxv
details
Raspberry
Arduino
Spark
Grab
Week
easy
Bundle
React
bundles
Pi
Looking
1gofBW6

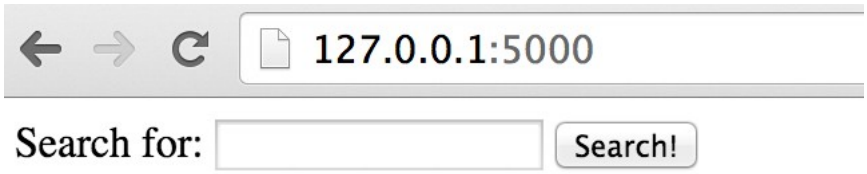
Chapter 5: Topic Analysis on Google+

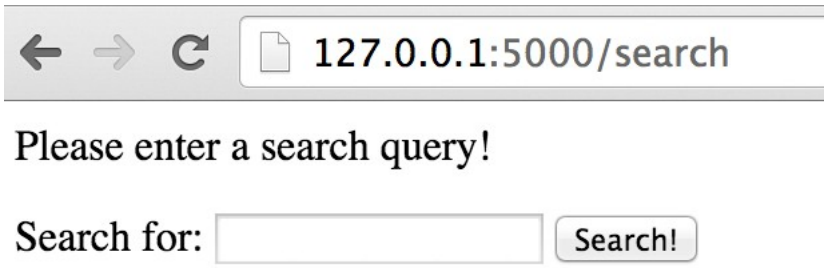





Disable API
Google+ API
This API is enabled, but you can't use it in your project until you create credentials.
Click "Go to Credentials" to do this now (strongly recommended).
Go to Credentials
Overview
Usage
Quotas
The Google+ API enables developers to build on top of the Google+ platform.
Learn more
Try this API in APIs Explorer


API Manager
Overview
Credentials
Credentials
OAuth consent screen
Domain verification
API key
Identifies your project using a simple API key to check quota and access.
For APIs like Google Translate.
OAuth client ID
Requests user consent so that your app can access the user's data.
For APIs like Google Calendar.
Service account key
Enables server-to-server, app-level authentication using robot accounts.
For use with Google Cloud APIs.
Help me choose
Asks a few questions to help you decide which type of credential to use.
Create credentials


127.0.0.1:5000
Search for:
Search!

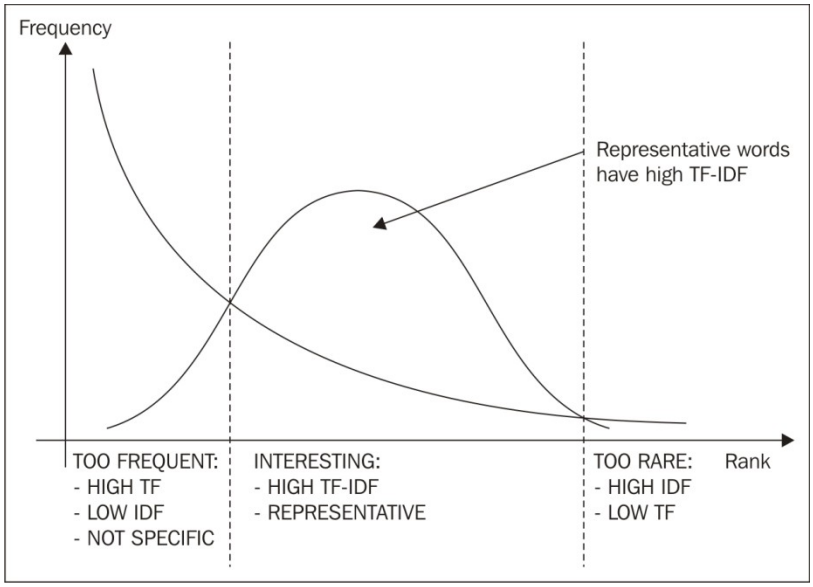

127.0.0.1:5000/search
Please enter a search query!
Search for:
Search!

Searching for **packt**:

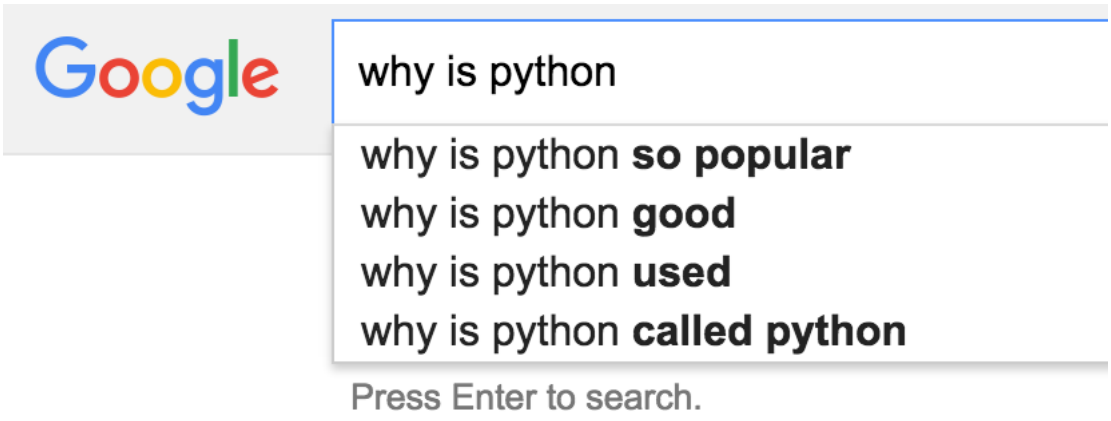
Packt Publishing (page)

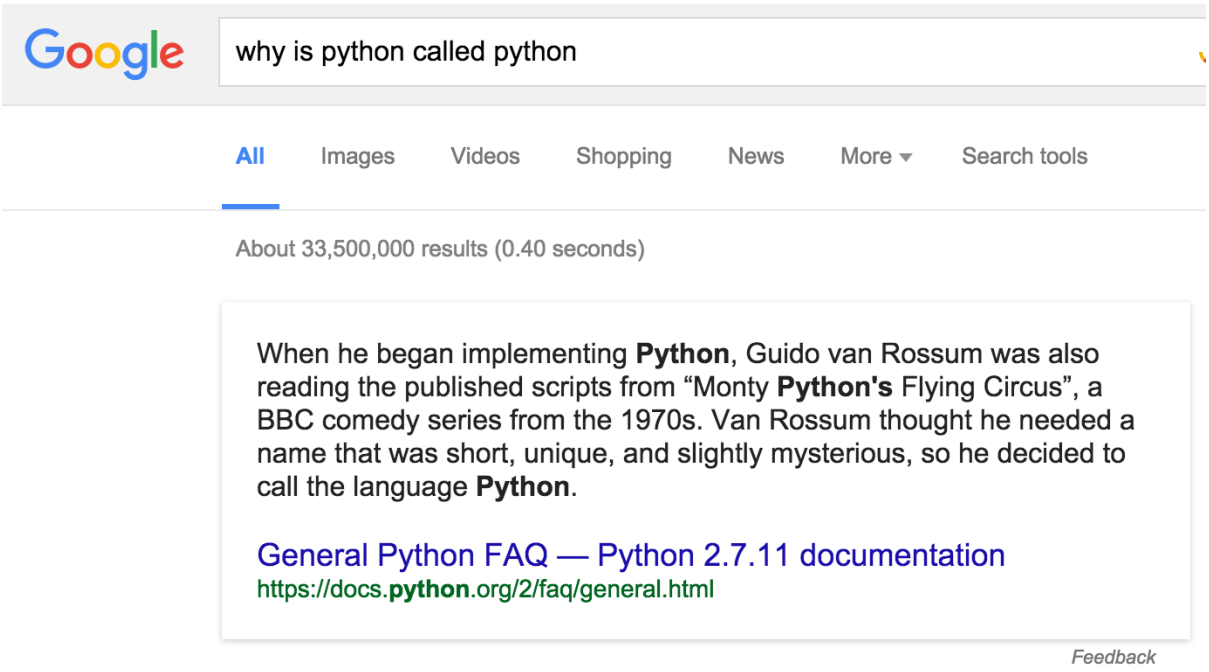
Packt Video (page)

Packt Authors (page)



Chapter 6: Questions and Answers on Stack Exchange





stackapps

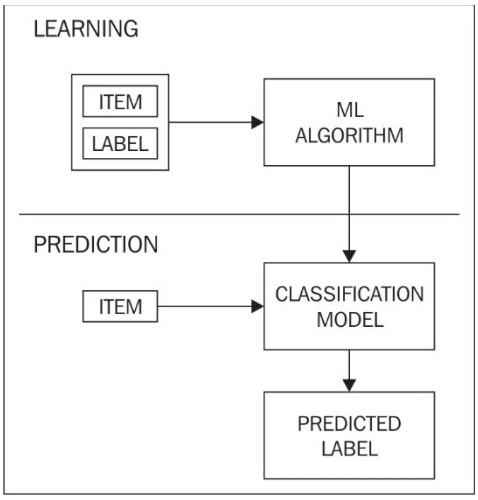
Register Your V2.0 Application

Application Name

Social Media Mining

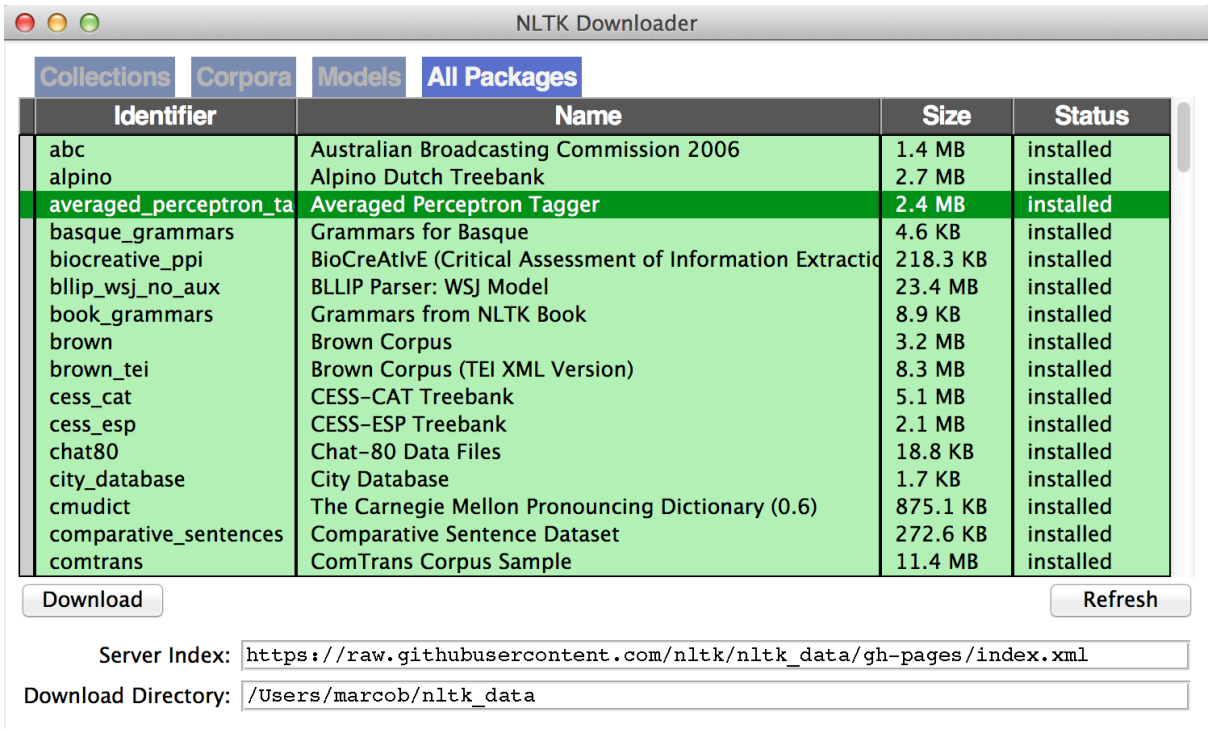
Description

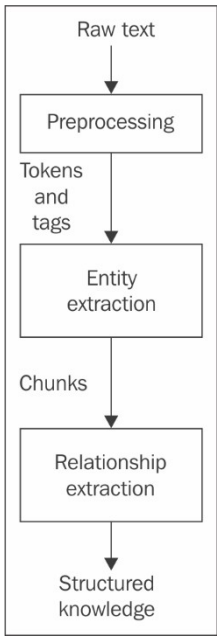
Test application to showcase examples on the book "Mastering Social Media Mining with Python"



Chapter 7: Blogs, RSS, Wikipedia and Natural Language Processing

```
>>> wikipedia.summary('London', sentences=2)
'London /ˈlʌndən/ is the capital and most populous city of England
and the United Kingdom. Standing on the River Thames in the south
east of Great Britain, London has been a major settlement for two
millennia.'
```

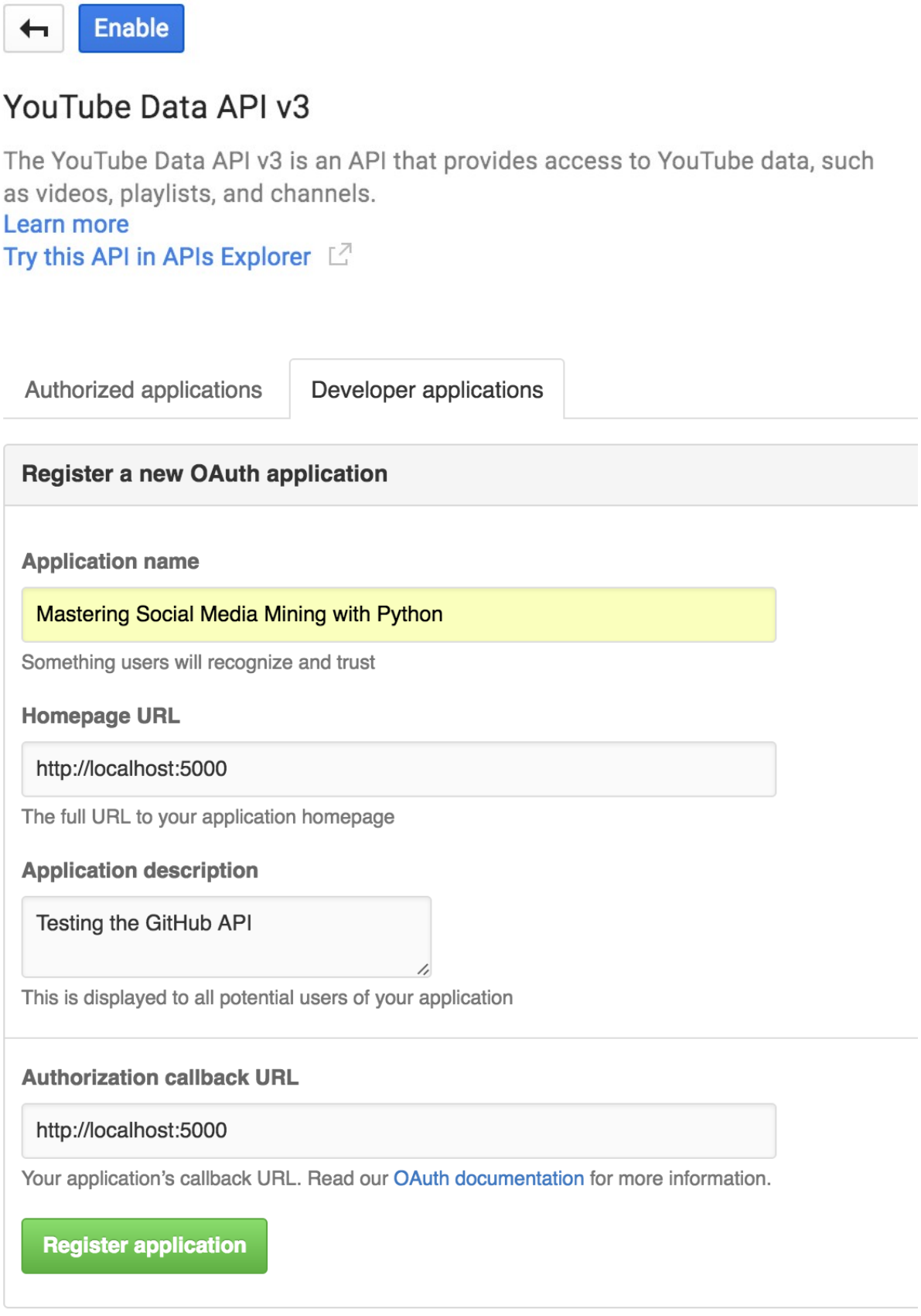




```
-----
Description of London
London /ˈlʌndən/ is the capital and most populous city of England and
the United Kingdom. Standing on the River Thames in the south east of
Great Britain, London has been a major settlement for two millennia.
-----
Noun phrases in description:
London
/ˈlʌndən/
capital
# more nouns ...
-----
Named entities in description:
London: London /ˈlʌndən/ is the capital and most populous city of
England and the United Kingdom.
England: England /ˈɪŋgl ənd/ is a country that is part of the United #
# more entities ...
```

London, /ˈlʌndən/, capital, city, England, United Kingdom, Standing, River Thames, south east, Great Britain, London, settlement, and millennia.

Chapter 8: Mining All the Data!



Complete the following form to gain access to the Yelp API.

Your website URL:

http://127.0.0.1:5000

How you will use the API:

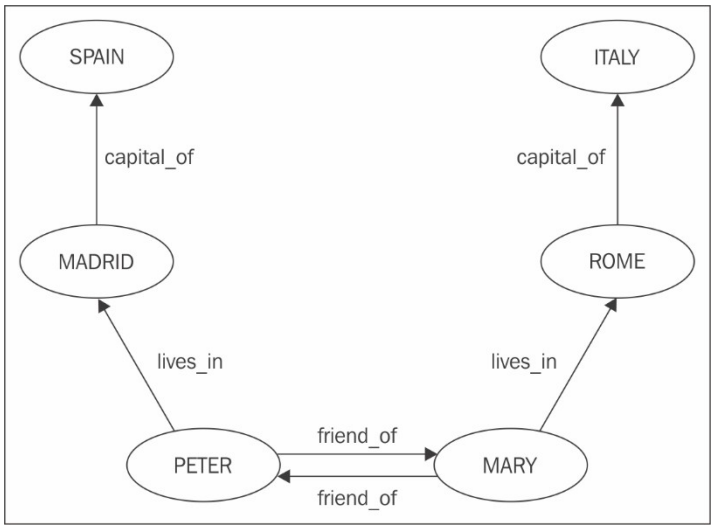
Testing Yelp API capabilities

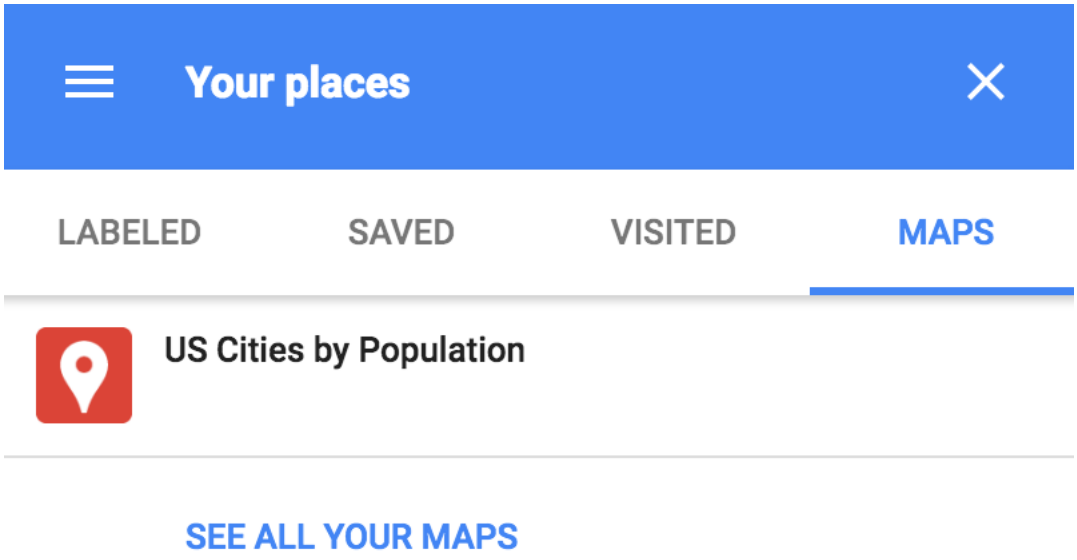
Industry:

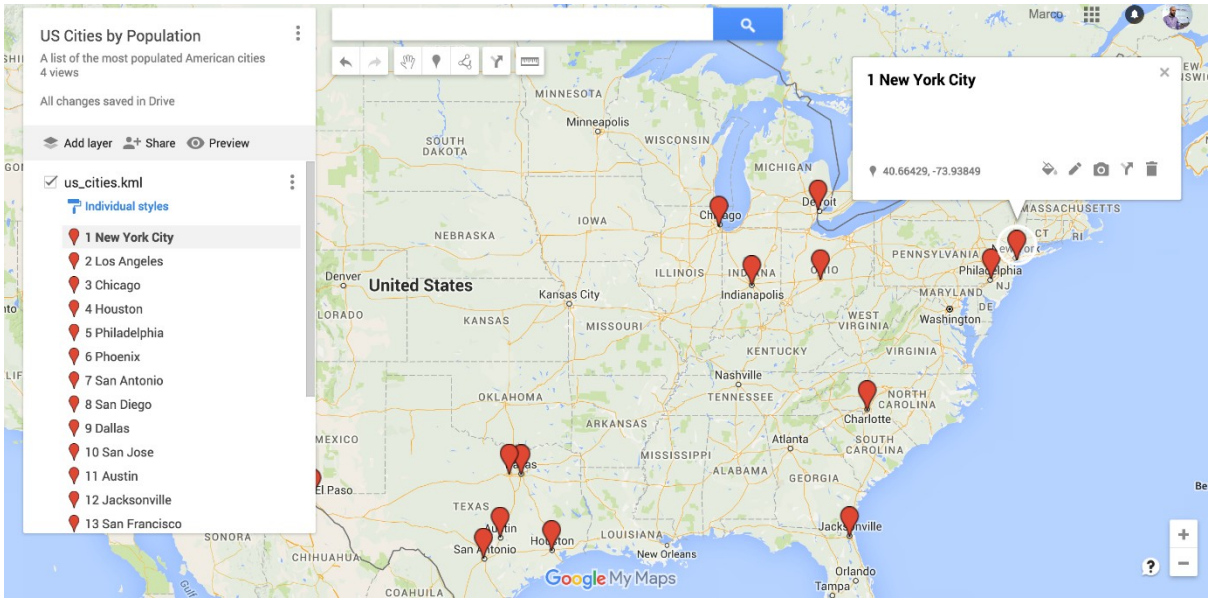
Other

Publishing

Chapter 9: Linked Data and the Semantic Web






US Cities by Population
A list of the most populated American cities
4 views
All changes saved in Drive
Add layer
Share
Preview
us_cities.kml
Individual styles
1 New York City
2 Los Angeles
3 Chicago
4 Houston
5 Philadelphia
6 Phoenix
7 San Antonio
8 San Diego
9 Dallas
10 San Jose
11 Austin
12 Jacksonville
13 San Francisco
1 New York City
40.66429, -73.93849
United States
Minneapolis
Chicago
Indianapolis
Kansas City
Denver
Nashville
Charlotte
Atlanta
New Orleans
Orlando
Tampa
Washington
Google My Maps