RM03: SPATIAL ANALYSIS AND MODELLING

Supervision 3: Linking Big Data with QGIS

Haifeng Niu

hn303@cam.ac.uk

Yiqiao Chen

yc418@cam.ac.uk

Material for supervision1

- Sup3-exercises (11-12 March, 2020)
- Revision supervision (13 March, 2020)

- Cambridge Moodle: RM03
  https://www.vle.cam.ac.uk/course/view.php?id=179012#section-2
- Online:
  https://hn303.github.io/CamLandEc-RM03/

Preview dark color scheme

RM03: Spatial Analysis and Modelling

Welcome to 2020 lent term module RM03 : Spatial Analysis and Modelling.
This repo is created by **Haifeng Niu** and contributed by Heeseo Rain Kwon and Paul Scherer*. Meterials of supervision could be found here.

Course outline

Lectures	Topic	Lecturers
Lecture 1	Introduction: Concepts, theory and practice in spatial analysis using GIS and data science	(Elisabete A. Silva)
Lecture 2	Data types of data, data collection and processing: from census to new live data harvesting in a digital age of big data	(Elisabete A. Silva)
Lecture 3	GIS and Data Processing: vector/raster/image data sets	(Elisabete A. Silva)
Lecture 4	Spatial metrics & analysis: static and dynamic environments	(Elisabete A. Silva & José Reis)
Supervision 1	QGIS - data analysis [Slides] [Exercises] [Assignment]	(H. Niu, H. R. Kwon)
Lecture 5	Urban and Environmental Dynamic Modelling	(Elisabete A. Silva)
Lecture 6	Dynamic simulation models SA, MCA, ABM, CA, GA and NN: development, calibration, validation	(Elisabete Silva)

UNIVERSITY OF CAMBRIDGE

Supervision overview

Sensing flood/storm in England & Wales with Twitter data

Exercise Part 1 (30 mins)

- Introduction of Application Programming Interface (API)
- API Application from Twitter and tweets data collection
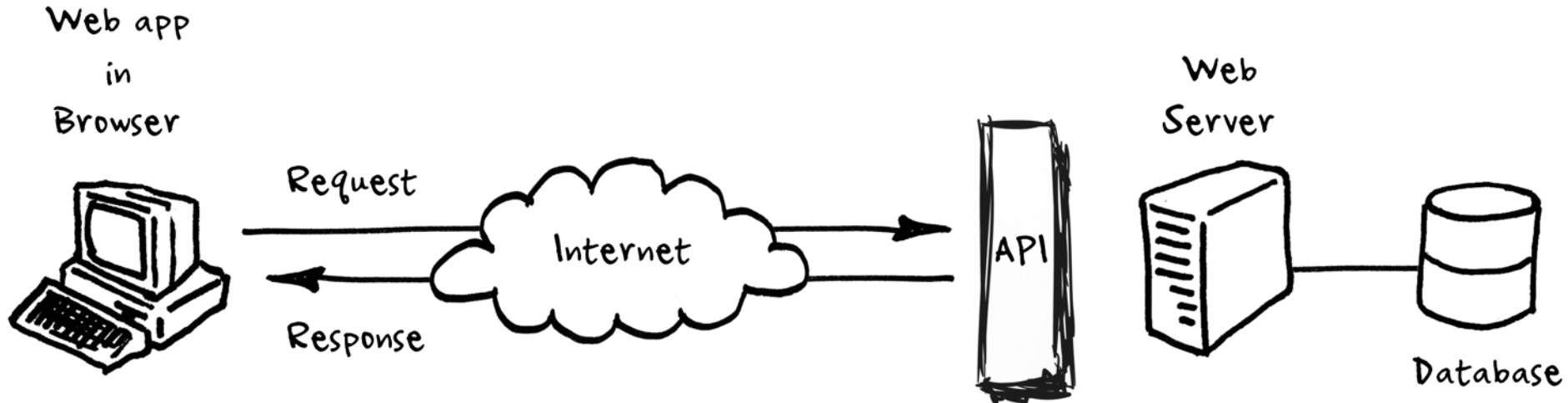- Data processing and Data mining (Sentiment analysis/ Discourse analysis)

Exercise Part 2 (50 mins)

- Import geotagged tweets to QGIS
- Visualization of location-based points on QGIS

UNIVERSITY OF CAMBRIDGE

Application Programming Interface (API)

- An **API** is a set of programming code that enables data transmission between one software product and another. It also contains the terms of this data exchange.

Web app in Browser — Request / Response — Internet — API — Web Server — Database

@Perry Eising

Twitter APIs for developers

Which API should I choose?

APIs	Description
Standard API	The included endpoints will let you perform the following: •Post, retrieve, and engage with Tweets and timelines •Post and receive direct messages •Manage and pull public account information •Create and manage lists •Follow, search and get users •Retrieve trends
Premium API	Advanced filtering functionality and scalable access to the the Search Tweets API, and real-time public account, engagement, and Tweet information via the webhook Account Activity API.
Enterprise API	Enterprise-level access to Twitter data, including real-time Tweets and public account information, historical Tweets, and Tweet insights.
Ads API	Programmatically integrate with the Twitter Ads platform.
Twitter for websites	Embed Twitter content and social actions to your site.
Twitter Developer Labs	Experimental endpoints being built around developer feedback.

Twitter APIs for developers

How to apply standard API from Twitter?

- Log in Twitter Developer Platform
- Create an new app
- Generate tokens

Keys and tokens
Keys, secret keys and access tokens management.

Consumer API keys   Regenerate

| API key: | coXtHK8eQraTJ6vDTZXus1DrW |
| API secret key: | 5mkF2ac563pAJ5Xaw8CuOCtTEmK8OeV0SCgRqjPvZQWtnKseFq |

Access token & access token secret   Revoke   Regenerate

We only show your access token and secret when you first generate it in order to make your account more secure. You can re-voke or regenerate them at any time, which will invalidate your existing tokens.

Access token:	xxxxxxxxxxxxxxxxxxxxxxxxxxxx	
Access token secret:	xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	*Last generated: Mar 3, 2020*
Access level:	Read, write, and Direct Messages	

# Twitter APIs for developers

Search result via API

```
{
  "statuses": [
    {
      "created_at": "Mon May 06 20:01:29 +0000 2019",
      "id": 1125490788736032770,
      "id_str": "1125490788736032770",
      "text": "Today's new update means that you can finally add Pizza Cat to your Retwee
      "truncated": true,
      "entities": {
        "hashtags": [],
        "symbols": [],
        "user_mentions": [],
        "urls": [
          {
            "url": "https://t.co/Rbc9TF2s5X",
            "expanded_url": "https://twitter.com/i/web/status/1125490788736032770",
            "display_url": "twitter.com/i/web/status/1…",
            "indices": [
              117,
              140
            ]
          }
        ]
      },
      "metadata": {
```
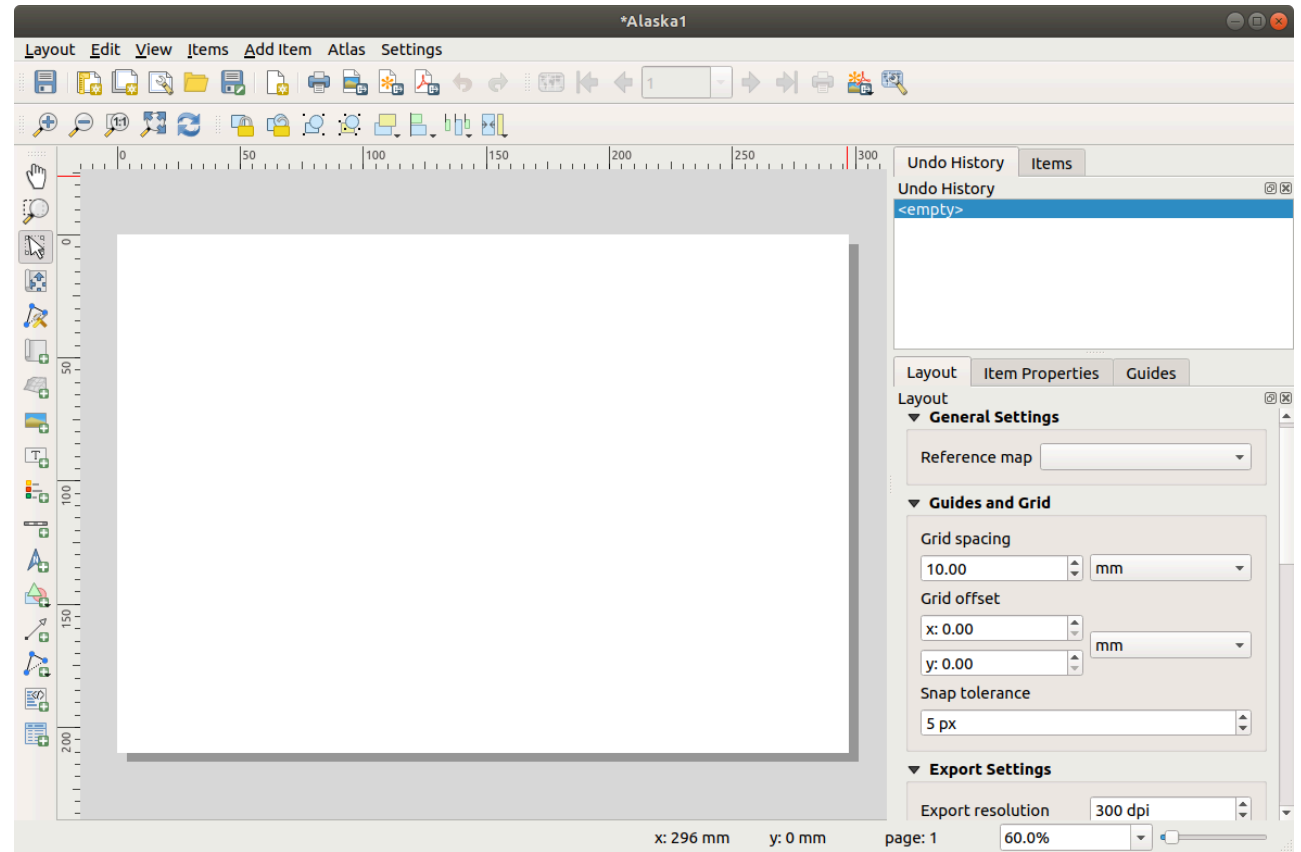
Collecting flood-related tweets in UK

- Sentiment analysis
    - AFINN package
- Discourse analysis
    - Word Cloud

UNIVERSITY OF CAMBRIDGE
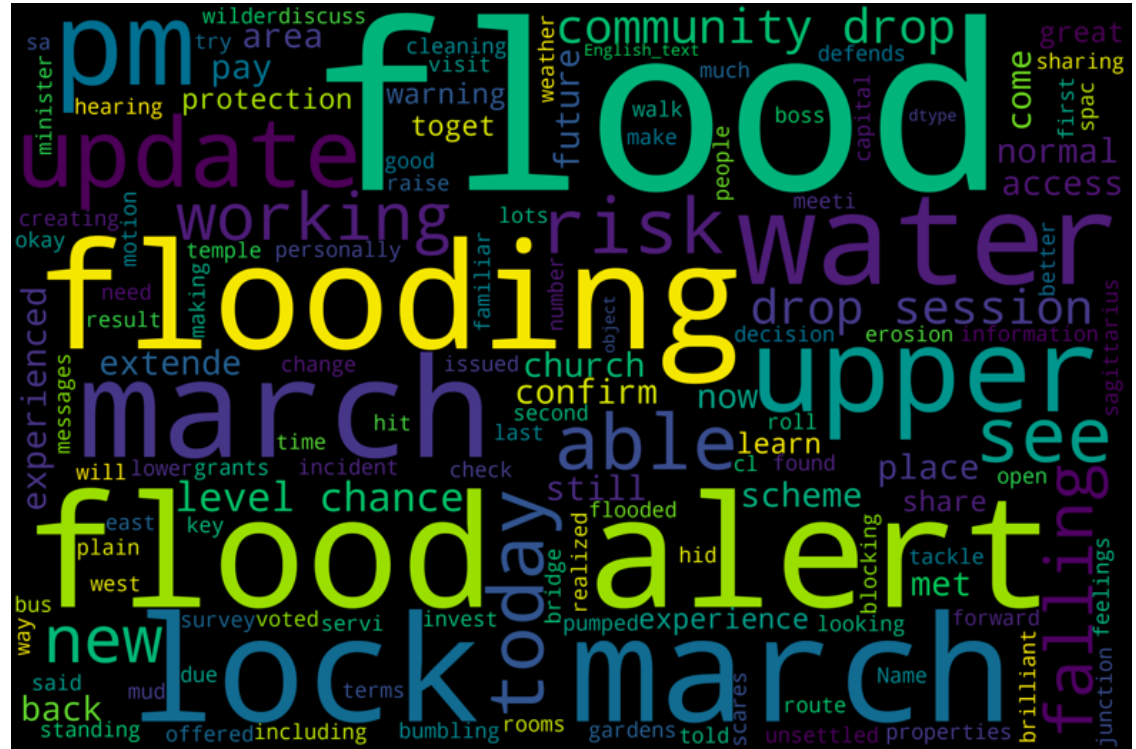
Making map with QGIS layout manager

- QGIS layout

UNIVERSITY OF
CAMBRIDGE

Online: https://hn303.github.io/CamLandEc-RM03/supervision3-exercises

CamLandEc-
RM03

Home
Supervision 1
Assignment 1
Assignment 1
Answers&Instruction
Supervision 2
Assignment 2
Assignment 2
Answers&Instruction
Supervision 3

Supervision 3 (12-13 March, 2020)

Instructions

1 Read through the instruction carefully. You may face problems if you overlook any of the steps.
2 The instruction for data collection via APIs is written in Google Colab, a free jupyter environment that requires no setup to use and runs python entirely in the cloud. You need log in with your **Google Account** to use this free platform. If you do not have Google account previously, you can try to log in with your Cambridge Email address (CRSid@cam.ac.uk). Know more about Google Colab, please check this link.
3 If you do not have **Twitter account**, please apply one via this Twitter Signup

Note: functions and filename are `highlighted` in this document.

Supervision overview

In this exercise, you will familiarise yourself with collecting data via Application programming interface(APIs), spatial visualization with geotagged tweets and creating a proper map on QGIS. The first two exercises will be practiced on Google Colab and the last exercise will be practiced on QGIS.

1. Collect Tweets via API

2. Content Analysis of Tweets

Please click this button below to move to Google Colab to start the first two exercises. Once open the colab, log in with your Google account and save a copy to your own Google Drive.

CO Open in Colab

3. Visualizaton of Geotagged Tweets

With geotagged location, social media can be used in mobility pattern identification, sentiment detection, emergency management and so on. In emergency management, social media paltform like Twitter can be used as crowdsourcing tool to collect real-time information in different effected areas. In this section, we will use geotagged tweets to identify the effected areas suffering floods or storms in the early spring 2020. Because of the limted time of supervision, we will use pre-collected data (data was collected in last week) to demonstrate how to process and visualize geotagged tweets.