**In Local:**

**Instructions to run each of the following files:**

1. **Crawler:**

* Open terminal and ‘cd’ to the directory that contains the crawler file‘AroundTheWorldLocal.py’
* Run the crawler using the following command:

*python AroundTheWorldLocal.py*

1. **FindMBROfImages**

* Open terminal and ‘cd’ to the directory ‘findMBRs’
* Run the crawler using the following command:

*python get\_location.py*  and then

*python add\_coord.py*

The resulting json can be found under the folder DataForTiles with the name ‘data\_xxx\_new’.json

* Now import this json to MongoDB as collection with MBRs for each image. This is needed in order to implement the next step

1. **ExtractDataFromMongoDB**

* Open terminal and ‘cd’ to the directory that contains the extraction code file‘extractDataMongo.py’
* Run the code using the following command:

*Python tran\_photos\_from\_db.py*

**Pre-requisites:**

1. **Crawler:**

* Please make sure you have this version of python in your machine: *Python 3.7.0*
* Please install the following python packages:

flickrapi, urllib,PIL, json, os, time, datetime, xml.dom

* Please replace the Flickr API access key in line number **191** of file *‘*AroundTheWorldLocal.py’ with your own access key. Alternatively, you can continue using mine
* In order to run the crawler on a certain list (for example, list of cities), please have the file present in the directory from which the code is run.

‘citiesInTheWorld.txt’ is the list of all cities in the world

‘countriesinTheWorld.txt’ is the list of all countries in the world

‘countiesInTheUS.txt’ is the list of all counties in the United States

* If you want to append to the current json, please have the existing json file corresponding to your list choice (data\_City.json in case of cities) in the directory from which the code is run. If you want to start afresh, this step is not necessary. The crawler will create a new json file.
* Make sure you have these files in the directory from which the code is run;

*flickr\_lib.py*

*getLocation.py*

1. **FindMBROfImages**

* Please install the following Python Packages:

pip install -r requests==2.22.0

1. **ExtractDataFromMongoDB**

* Please install Mongo DB and Studio 3T in your machine
* Please change port number and host in *line # 32* according to your setup
* Please change database name in *line # 35* according to your setup
* Please change collection name *line # 36* according to your setup
* Please make sure you have imported the ‘data\_xxx\_new’.json (data\_city\_new.json in case of cities) as collection to Studio 3T

**On Cluster:**

**Instructions to run each of the following files:**

1. **Crawler:**

* ‘cd’ to the directory that contains the crawler file‘AroundTheWorldLocal.py’
* Run the crawler using the following command:

*python AroundTheWorldLocal.py*

**Pre-requisites:**

1. **Crawler:**

* Please make sure you have this version of python on the cluster: *Python 2.7.5*
* Please install the following python packages:
* flickrapi, urllib,PIL, json, os, time, datetime, xml.dom
* Please replace the Flickr API access key in line number **191** of file *‘*AroundTheWorldLocal.py’ with your own access key. Alternatively, you can continue using mine
* In order to run the crawler on a certain list (for example, list of cities), please have the file present in the directory from which the code is run.

‘citiesInTheWorld.txt’ is the list of all cities in the world

‘countriesinTheWorld.txt’ is the list of all countries in the world

‘countiesInTheUS.txt’ is the list of all counties in the United States

* If you want to append to the current json, please have the existing json file corresponding to your list choice (data\_City.json in case of cities) in the directory from which the code is run. If you want to start afresh, this step is not necessary. The crawler will create a new json file.
* Make sure you have these files in the directory from which the code is run;

*flickr\_lib.py*

*getLocation.py*