Getting Started

Dependencies and imports

Flickrapi https://pypi.org/project/flickrapi/

json (Javascript object notation)

Pandas (my best friend)

Numpy (arrays for python)

Re (regex for python)

Flickr

You will need to obtain a Flickr key and secret to interact with the API. Flickr has lots of documentation at Flickr Developer Guide, a link to get your own key can be found under requesting keys.

The two methods used by FlickrResearch are flickr.photos.search and flickr.people.getinfo.

Pandas

A must have for data analysis and manipulation. FlickrResearch uses the functionality of Pandas for data collection, storage, and manipulation. For more info check out Pandas. You need two use two pandas functions alongside FlickrResearch functions, read_csv and to_csv.

Hemisphere dependency

The recursive bounding box method is hardcoded to the northwestern hemisphere. No offense to any other hemisphere, but the northwestern hemisphere is where I live, work, and research. The recursiveBBox method can be modified for other hemispheres.

Bounding Boxes

Geographic bounding boxes consist of 4 coordinates. A longitude and latitude for the lower left corner, and a longitude and latitude for the upper right corner. The Flickr API needs this parameter to know where you want to pull data from.

Example of a bounding box.

-66.976,47.292

-70.931, 42.907

The constructor

The constructor takes in parameters needed for a Flickr API call to flickr.photos.search. Parameters included are minimum date taken, maximum date taken, hasGeo, extras, Flickr API key, Flickr API secret.

Example of a Constructor call

```
apiCall = FlickrResearch('2021-01-01','2021-12-31', '1', 'geo,date_taken,tags',
                         '123456ABC','sample8675309')
```

FlickrResearch will return a "Key error" when there is an error in the parameters being fed into the constructor. If you get this error, check the constructor for errors before you dive into other possibilities.

Many of the methods described in this program use information from the constructor, and you need to use it to change the date parameters of your calls.

Pass those data frames around!

Many functions take a data frame in as a parameter and return a different data frame. Successive steps in data cleaning and processing can be accomplished by passing data frames through successive functions. For example, the code shown below calls the data for a bounding box, drops duplicates, and appends user home location information before writing out a csv file.

```
DF2 =apiCall.recursiveBBox(-7.791,61.37,-6.102,62.524)
DF3 = apiCall.dropDuplicates(DF2)
DF4 = apiCall.appendUserInfo(DF3)
DF4.to_csv('Faroe2021_Loc.csv')
```

Column structure

Indexing is used to access a specific row and column in pandas. The methods in FlickrResearch rely on the order of the columns and need the following structure with incoming csv files.

	0	1	2	3	4	5	6	7	8
	A	B	C	D	E	F	G	H	I
	id	accuracy	latitude	longitude	owner	datetaken	title	tags	
317	38870345	16	46.86313	-67.998	10092281(	5/27/2004 19:30	Broken Ra	railroad bl
318	6022587	16	46.86081	-68.0281	10092281(	8/8/2004 12:37	Cloud	Lomo
319	5724600	16	46.85662	-68.0036	10092281(	7/18/2004 17:49	Neglectec	grate negl
320	5320122	16	46.70683	-68.0006	10092281(	7/18/2004 18:26	Harvest Bl	maine pot
321	5320121	16	46.86081	-68.0281	10092281(	8/23/2004 17:14	Inside Loc	window di
572	3.28E+09	15	42.90943	-70.8108	10329027(	7/1/2004 0:00	Hanson @	new music
1031	1.75E+09	15	44.35273	-68.2243	10695339(	8/4/2004 15:09	Cadillac Mountain - A
1032	1.75E+09	15	44.36157	-68.2305	10695339(	8/4/2004 15:36	View of Frenchman E
1028	2.15E+09	16	44.826	-68.7706	10713496(	6/27/2004 10:11	Picture 27	flowers fo
600	1.23E+09	15	43.66505	-70.1907	10790808(	12/25/2004 15:04	Parents' H	maine

Files created by the recursiveBBox are output in this format. As long as you don't alter the order of the columns you are good to go. Feel free to rename the columns.