

John M Warlop

Python for Infomatics

UCSD Winter 2018

Assignment #8



[Project #1 - Building a Google map from geocoded data](#)

[Running of program and visualizations\(added UCSD,USD & SDSU\)](#)

[Project #2 - Visualizing Networks and Interconnections](#)

[HTML Page Force.htm](#)

[SQL Database Spider.sqlite](#)

[Project #3 - Visualizing Mail Data](#)

[Directory Listing after Running Completing Project](#)

[Running gmain.py](#)

[Database content.sqlite](#)

[Running gmodel.py](#)

[Database index.sqlite](#)

[Running gbasic.py](#)

[Word Cloud from gword.py=>gword.jsonp & gword.html](#)

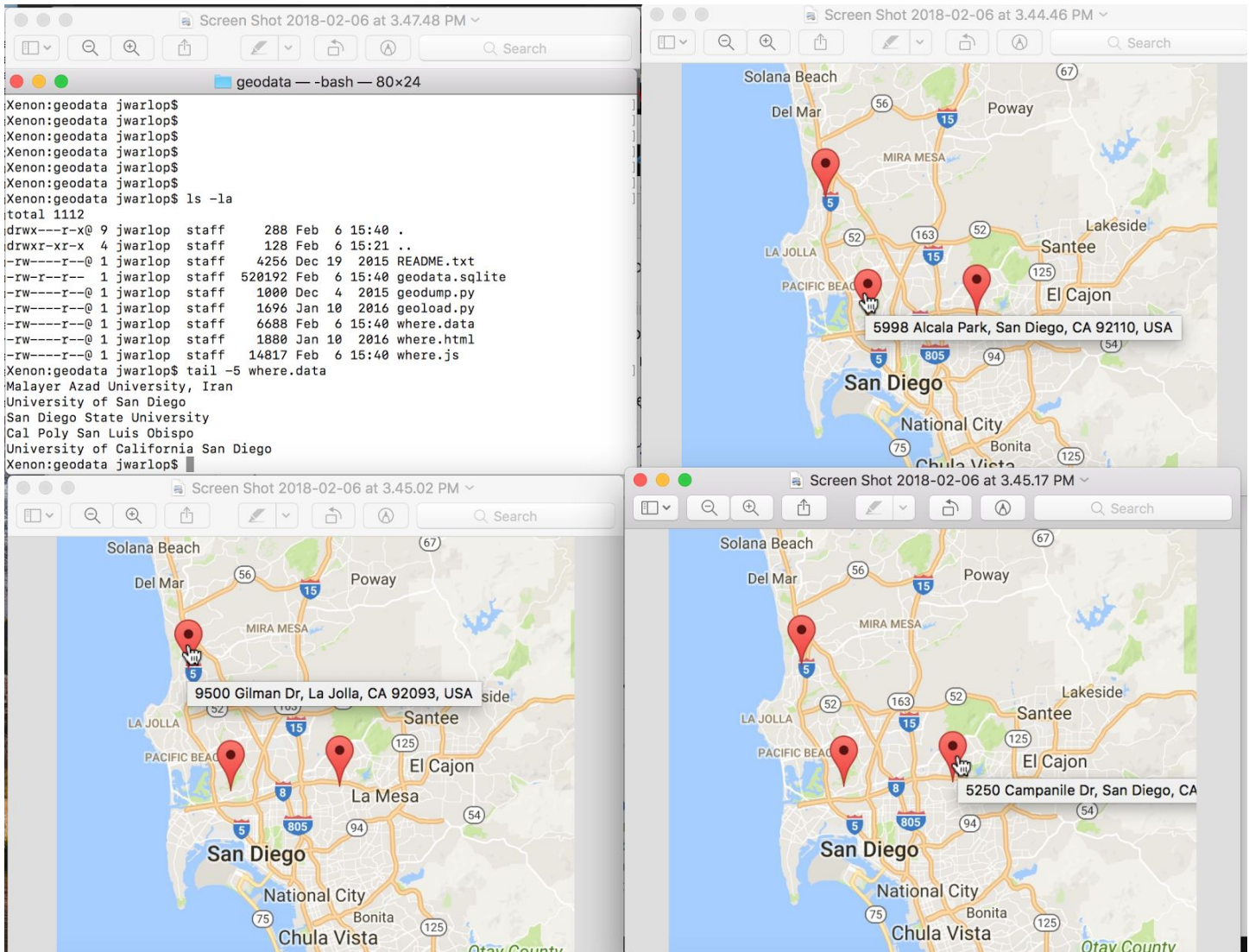
[Running gline.py](#)

[Gline.htm](#)

Project #1 - Building a Google map from geocoded data



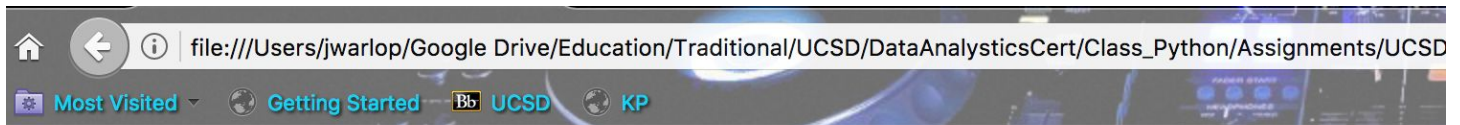
Running of program and visualizations(added UCSD,USD & SDSU)



Project #2 - Visualizing Networks and Interconnections



HTML Page Force.htm



Starting url: https://lti-tools.dr-chuck.com/tsugi/mod/python-data/data/known_by_Miranne.html



If you don't see a chart above, check the JavaScript console. You may need to use a different browser.

SQL Database Spider.sqlite

SQLite Manager - /Users/jwarlop/Google Drive/Education/Traditional/UCSD/DataAnalyticsCert/Class_Python/Assignments/UCSD_A08/pagerank/spider.sqlite

spider.sqlite

Master Table (1)

Tables (3)

Links

Pages

Webs

Views (0)

Indexes (2)

Triggers (0)

Directory

(Select Profile Database)

Go

Structure

Browse & Search

Execute SQL

DB Settings

TABLE

Pages

Search

Show All

Add

Duplicate

Edit

Delete

id	url	html	error	old_rank	new_rank
1	http://www.dr-chuck.c...	BLOB (Size: 8528)		1	1.876713277903755
2	http://www.dr-chuck.c...	BLOB (Size: 64793)		1	3.836963675828944
3	http://www.dr-chuck.c...			1	1
4	http://www.dr-chuck.c...	BLOB (Size: 2133)		1	1.244014550264551
5	http://www.dr-chuck.c...			1	1
6	http://www.dr-chuck.c...			1	1
7	http://www.dr-chuck.c...	BLOB (Size: 2426)		1	0.577347883597884
8	http://www.dr-chuck.c...	BLOB (Size: 76358)		1	3.056615981195535
9	http://www.dr-chuck.c...	BLOB (Size: 35565)		1	2.878123812723208
10	http://www.dr-chuck.c...			1	1
11	http://www.dr-chuck.c...			1	1
12	http://www.dr-chuck.c...	BLOB (Size: 37112)		1	2.878123812723208
13	http://www.dr-chuck.c...			1	1
14	http://www.dr-chuck.c...			1	1
15	http://www.dr-chuck.c...			1	1
16	http://www.dr-chuck.c...	BLOB (Size: 34949)		1	0.5487134668682289
17	http://www.dr-chuck.c...			1	1
18	http://www.dr-chuck.c...	BLOB (Size: 31851)		1	0.5487134668682289
19	http://www.dr-chuck.c...	BLOB (Size: 39189)		1	0.5487134668682289
20	http://www.dr-chuck.c...			1	1
21	http://www.dr-chuck.c...			1	1
22	http://www.dr-chuck.c...	BLOB (Size: 65461)		1	0.5487134668682289
23	http://www.dr-chuck.c...			1	1

<< < 1 to 66 of 66 > >>

SQLite 3.17.0 Gecko 52.6.0 0.8.3.1-signed.1-signed Shared Number of files in selected directory: 9 ET: 78 ms

Project #3 - Visualizing Mail Data



Directory Listing after Running Completing Project

```
gmane — -bash — 80x24
[Xenon:gmane jwarlop$
[Xenon:gmane jwarlop$
[Xenon:gmane jwarlop$ ls -la
total 1835928
drwx---r-x@ 19 jwarlop  staff          608 Feb  8 06:34 .
drwxr-xr-x  7 jwarlop  staff          224 Feb  7 05:59 ..
-rw----r--@ 1 jwarlop  staff          7495 Feb 15 2016 README.txt
-rw-r--r--  1 jwarlop  staff 669151232 Feb  8 06:01 content.sqlite
-rw----r--@ 1 jwarlop  staff       10877 Dec  4 2015 d3.layout.cloud.js
-rwx---r-x@ 1 jwarlop  staff     253776 Dec  4 2015 d3.v2.js
-rw----r--@ 1 jwarlop  staff         945 Feb 15 2016 gbasic.py
-rw----r--@ 1 jwarlop  staff         763 Dec  4 2015 gline.htm
-rw----r--@ 1 jwarlop  staff        3771 Dec  4 2015 gline.jsonp
-rw----r--@ 1 jwarlop  staff        1680 Feb 15 2016 gline.py
-rw----r--@ 1 jwarlop  staff        4508 Feb 15 2016 gmane.py
-rw----r--@ 1 jwarlop  staff        7884 Feb  8 2016 gmodel.py
-rw----r--@ 1 jwarlop  staff        1087 Dec  4 2015 gword.htm
-rw-r--r--  1 jwarlop  staff        2837 Feb  8 06:34 gword.js
-rw----r--@ 1 jwarlop  staff        2825 Dec  4 2015 gword.jsonp
-rw----r--@ 1 jwarlop  staff        1360 Feb 15 2016 gword.py
-rw----r--@ 1 jwarlop  staff        1663 Feb 15 2016 gyear.py
-rw-r--r--  1 jwarlop  staff 263839744 Feb  8 06:31 index.sqlite
-rw----r--@ 1 jwarlop  staff       12288 Feb 15 2016 mapping.sqlite
Xenon:gmane jwarlop$
```

Running gmane.py

```
gmane — Python gmane.py — 80x24
Xenon:gmane jwarlop$ ]
Xenon:gmane jwarlop$ ]
Xenon:gmane jwarlop$ python2.7 gmane.py ]
60408
How many messages:10
http://mbox.dr-chuck.net/sakai.devel/60409/60410 9042
    neal.caidin@apereo.org 2015-05-29T11:13:45-04:00 re: [i18n] joint session on
    i18n by sakai en espanol and
http://mbox.dr-chuck.net/sakai.devel/60410/60411 10629
    neal.caidin-wgvhjocuptadnm+yrofe0a@public.gmane.org 2015-05-29T12:54:12-04:0
    0 sakai 10.5 rc01 available!
http://mbox.dr-chuck.net/sakai.devel/60411/60412 5588
    tokiwa@yujitokiwa.jp 2015-05-29T23:58:34+09:00 [i18n] joint session on i18n
    by sakai en espanol and
http://mbox.dr-chuck.net/sakai.devel/60412/60413 13368
    seduerey@gmail.com 2015-06-01T01:21:18-07:00 re: adding some classes to morp
    heus portal?
http://mbox.dr-chuck.net/sakai.devel/60413/60414 7102
    agschmid-r/hdy37nxir2fbvcvol8/a@public.gmane.org 2015-06-01T09:58:36-04:00 5
    02 bad gateway on https://qa10-oracle.nightly.sakaiproject.org/
http://mbox.dr-chuck.net/sakai.devel/60414/60415 25538
    mcarro@entornosdeformacion.com 2015-06-01T09:59:04+02:00 adding some classes
    to morpheus portal?
http://mbox.dr-chuck.net/sakai.devel/60415/60416 10377
```

Database content.sqlite

SQLite Manager - /Users/jwarlop/Google Drive/Education/Traditional/UCSD/DataAnalyticsCert/Class_Python/Assignments/UCSD_A08/gmane/content.sqlite

Directory (Select Profile Database) Go

Structure Browse & Search Execute SQL DB Settings

content.sqlite

- Master Table (1)
- Tables (1)
 - Messages
- Views (0)
- Indexes (1)
- Triggers (0)

TABLE Messages Search Show All Add Duplicate Edit Delete

rowid	id	email	sent_at	subject	headers	body
1	1	ggolden@umich.edu	2005-12-08T23:34:3...	call for participation: ...	From news@gmane.o...	At toda
2	2	csev@umich.edu	2005-12-09T00:58:0...	report from the austi...	From news@gmane.o...	At one
3	3	kevin.carpenter@rsm...	2005-12-09T09:01:4...	cas and sakai 1.5	From news@gmane.o...	Sorry if
4	4	michael.feldstein@su...	2005-12-09T09:43:1...	re: lms/vle rants/com...	From news@gmane.o...	Yup, I t
5	5	john@caret.cam.ac.uk	2005-12-09T13:32:2...	re: lms/vle rants/com...	From news@gmane.o...	I shoulc
6	6	s-githens@northwest...	2005-12-09T13:32:3...	re: sakaiportallogin a...	From news@gmane.o...	I think s
7	7	john@caret.cam.ac.uk	2005-12-09T13:42:2...	re: lms/vle rants/com...	From news@gmane.o...	BTW so
8	8	s-githens@northwest...	2005-12-09T15:23:0...	re: sakaiportallogin a...	From news@gmane.o...	Yuji Shi
9	9	ys2n@virginia.edu	2005-12-09T17:54:1...	sakaiportallogin and ...	From news@gmane.o...	I am try
10	10	ys2n@virginia.edu	2005-12-09T21:16:3...	re: sakaiportallogin a...	From news@gmane.o...	The por
11	11	slt@columbia.edu	2005-12-09T23:16:4...	re: cas and sakai 1.5	From news@gmane.o...	Kevin C
12	12	dgcotton@ucdavis.edu	2005-12-10T20:55:5...	memory error with 2.1	From news@gmane.o...	I keep g
13	13	vyu2@yahoo.com	2005-12-10T23:34:4...	re: memory error with...	From news@gmane.o...	Sakai d
14	14	winstonw.tw@gmail.c...	2005-12-11T03:07:2...	re: sakai ldap authent...	From news@gmane.o...	Could s
15	15	ajpoland@iupui.edu	2005-12-11T12:08:15...	re: memory error with...	From news@gmane.o...	Dan, Al
16	16	caseyd1@stanford.edu	2005-12-12T07:02:3...	re: sakai 2.1 provider ...	From news@gmane.o...	Thanks
17	17	aaronz@vt.edu	2005-12-12T08:55:0...	sakai 2.1 provider exa...	From news@gmane.o...	I put up
18	18	j.higham@hull.ac.uk	2005-12-12T09:49:1...	re: sakai 2.1 - lone ze...	From news@gmane.o...	Thanks
19	19	aaronz@vt.edu	2005-12-12T10:13:38...	re: sakai 2.1 provider ...	From news@gmane.o...	I shoulc
20	20	jeffrey.beeman@asu....	2005-12-12T11:14:55...	re: sakai logo	From news@gmane.o...	Not sur
21	21	csev@umich.edu	2005-12-12T11:33:29...	sepp library discussio...	From news@gmane.o...	The gro
22	22	pgoldweic@northwes...	2005-12-12T13:16:12...	problems accessing t...	From news@gmane.o...	Hi, I did
23	23	pgoldweic@northwes...	2005-12-12T14:20:41...	re: problems accessi...	From news@gmane.o...	You mig

<< < 1 to 100 of 59847 > >>

SQLite 3.17.0 Gecko 52.6.0 0.8.3.1-signed.1-signed Shared Number of files in selected directory: 9 ET: 6 ms

Running gmodel.py

```
gmane — -bash — 80x24
[Xenon:gmane jwarlop$ python2.7 gmodel.py
Loaded allsenders 1802 and mapping 29 dns mapping 1
1 2005-12-08T23:34:30-06:00 ggolden22@mac.com
251 2005-12-22T10:03:20-08:00 tpamsler@ucdavis.edu
501 2006-01-12T11:17:34-05:00 lance@indiana.edu
751 2006-01-24T11:13:28-08:00 vrajgopalan@ucmerced.edu
1001 2006-02-02T08:27:30-07:00 john.ellis@rsmart.com
1251 2006-02-15T19:13:17-05:00 ggolden22@mac.com
1501 2006-02-23T16:49:00-05:00 csev@umich.edu
1751 2006-03-13T13:48:36-05:00 csev@umich.edu
2001 2006-03-27T13:34:32-06:00 swgithen@mtu.edu
2251 2006-04-05T12:38:43-04:00 ccount@mit.edu
2501 2006-04-17T13:28:54-07:00 jholtzman@berkeley.edu
2751 2006-04-25T14:26:03-07:00 vrajgopalan@ucmerced.edu
3001 2006-05-03T13:49:11-04:00 cheryl.wogahn@yale.edu
3251 2006-05-10T17:14:01+00:00 nuno@ufp.pt
3501 2006-05-17T10:37:49-04:00 ys2n@virginia.edu
3751 2006-05-23T14:52:13-04:00 azeckoski@unicon.net
4001 2006-06-06T12:51:32-04:00 ajpoland@indiana.edu
4251 2006-06-14T08:28:01-04:00 lance@indiana.edu
4501 2006-06-21T14:31:06-05:00 mustansar@rice.edu
4751 2006-06-29T15:52:15+01:00 antranig@cam.ac.uk
5001 2006-07-12T11:50:55-04:00 markjnorton@earthlink.net
5251 2006-07-20T17:24:14-07:00 sinou@etudes.org
```


Database index.sqlite

SQLite Manager - /Users/jwarlop/Google Drive/Education/Traditional/UCSD/DataAnalyticsCert/Class_Python/Assignments/UCSD_A08/gmane/index.sqlite

index.sqlite

Master Table (1)

Tables (4)

Messages

Replies

Senders

Subjects

Views (0)

Indexes (3)

Triggers (0)

Directory

(Select Profile Database)

Go

Structure

Browse & Search

Execute SQL

DB Settings

TABLE Messages

Search

Show All

Add

Duplicate

Edit

Delete

id	guid	sent_at	sender_id	subject_id	headers	body
1	<b7eaf76e-3366-...	2005-12-09 05:34:...	1	1	x[...]	x[...]
2	<bb7ce631-cf0-4...	2005-12-09 05:58:...	2	2	x[...]	x[...]
3	<16703710.113414...	2005-12-09 16:01:...	3	3	x[...]	x[...]
4	<9622225.113413...	2005-12-09 14:43:...	4	4	x[...]	x[...]
5	<22585576.11341...	2005-12-09 13:32:...	5	4	x[...]	x[...]
6	<29494755.11341...	2005-12-09 19:32:...	6	5	x[...]	x[...]
7	<20623058.11341...	2005-12-09 13:42:...	5	4	x[...]	x[...]
8	<14721465.113416...	2005-12-09 21:23:...	6	5	x[...]	x[...]
9	<31872472.113415...	2005-12-09 17:54:...	7	6	x[...]	x[...]
10	<1304695.113416...	2005-12-09 21:16:...	7	5	x[...]	x[...]
11	<3598962.113418...	2005-12-10 04:16:...	8	7	x[...]	x[...]
12	<590808.1134277...	2005-12-11 04:55:...	9	8	x[...]	x[...]
13	<27775220.11342...	2005-12-11 07:34:...	10	9	x[...]	x[...]
14	<18541911.113424...	2005-12-10 19:07:...	11	10	x[...]	x[...]
15	<20021180.11343...	2005-12-11 17:08:15	12	9	x[...]	x[...]
16	<28266638.11343...	2005-12-12 15:02:...	13	11	x[...]	x[...]
17	<1755944.113439...	2005-12-12 13:55:...	14	12	x[...]	x[...]
18	<10733551.11343...	2005-12-12 09:49:...	15	13	x[...]	x[...]
19	<13486694.11344...	2005-12-12 15:13:...	14	11	x[...]	x[...]
20	<25165005.11344...	2005-12-12 18:14:...	16	14	x[...]	x[...]
21	<d3dc4255-5936-...	2005-12-12 16:33:...	2	15	x[...]	x[...]
22	<6.0.1.1.2.2005121...	2005-12-12 19:16:12	17	16	x[...]	x[...]
23	<6.0.1.1.2.2005121...	2005-12-12 20:20:...	17	17	x[...]	x[...]

<<

<

1 to 100 of 59742

>

>>

SQLite 3.17.0

Gecko 52.6.0

0.8.3.1-signed.1-signed

Shared

Number of files in selected directory: 9

ET: 4 ms

Running gbasic.py

```
gmane — -bash — 80x24
58001 2014-11-05T18:49:11+05:30 ismail.n@uaeexchange.com
58251 2014-12-08T10:38:55-05:00 hedrick@rutgers.edu
58501 2015-01-16T09:16:13-05:00 enietzel@anisakai.com
58751 2015-02-18T08:14:09-05:00 neal.caidin@apereo.org
59001 2015-03-23T08:46:02-04:00 csev@umich.edu
59251 2015-04-20T16:24:19-04:00 bjones86@uwo.ca
59501 2015-05-07T16:13:35-04:00 jmezzanotte@anisakai.com
[Xenon:gmane jwarlop$ python2.7 gbasic.py
How many to dump? 5

Top 5 Email list participants
steve.swinsborg@swinsborg.com 3335
azeckoski@unicon.net 1907
ian@cam.ac.uk 1591
csev@umich.edu 1490
david.horwitz@uct.ac.za 1221

Top 5 Email list organizations
umich.edu 6805
gmail.com 5690
swinsborg.com 3335
cam.ac.uk 2626
uct.ac.za 2582
Xenon:gmane jwarlop$
```

Word Cloud from gword.py=>gword.jsonp & gword.html



Running gline.py

```
gmane — -bash — 80x24
[Xenon:gmane jwarlop$ python2.7 gline.py
Top 10 Organizations
['umich.edu', 'gmail.com', 'swinsborg.com', 'cam.ac.uk', 'uct.ac.za', 'indiana.edu', 'longsight.com', 'unicon.net', 'ox.ac.uk', 'berkeley.edu']
({'2008-09', 'umich.edu'): 119, ('2013-02', 'longsight.com'): 54, ('2007-11', 'cam.ac.uk'): 142, ('2013-02', 'umich.edu'): 59, ('2014-01', 'swinsborg.com'): 30, ('2009-12', 'total'): 170, ('2012-07', 'berkeley.edu'): 2, ('2009-09', 'berkeley.edu'): 6, ('2010-05', 'longsight.com'): 8, ('2009-12', 'cam.ac.uk'): 2, ('2014-10', 'umich.edu'): 2, ('2013-11', 'ox.ac.uk'): 12, ('2008-05', 'swinsborg.com'): 44, ('2012-12', 'swinsborg.com'): 35, ('2013-08', 'ox.ac.uk'): 2, ('2014-10', 'swinsborg.com'): 9, ('2012-03', 'berkeley.edu'): 1, ('2007-05', 'total'): 407, ('2009-12', 'ox.ac.uk'): 12, ('2008-03', 'berkeley.edu'): 19, ('2010-03', 'total'): 356, ('2010-01', 'swinsborg.com'): 39, ('2008-10', 'indiana.edu'): 27, ('2012-07', 'unicon.net'): 8, ('2011-06', 'gmail.com'): 89, ('2010-01', 'umich.edu'): 57, ('2013-08', 'unicon.net'): 10, ('2014-07', 'swinsborg.com'): 42, ('2011-12', 'cam.ac.uk'): 2, ('2010-05', 'total'): 222, ('2009-10', 'umich.edu'): 77, ('2015-01', 'uct.ac.za'): 1, ('2014-02', 'umich.edu'): 59, ('2006-09', 'uct.ac.za'): 26, ('2011-12', 'swinsborg.com'): 31, ('2014-01', 'ox.ac.uk'): 10, ('2009-08', 'indiana.edu'): 25, ('2012-02', 'indiana.edu'): 10, ('2008-02', 'indiana.edu'): 34, ('2007-11', 'swinsborg.com'): 22, ('2010-12', 'umich.edu'): 26, ('2007-03', 'umich.edu'): 93, ('2006-03', 'indiana.edu'): 43, ('2009-05', 'indiana.edu'): 16, ('2011-10', 'umich.edu'): 23, ('2011-08', 'swinsborg.com'): 37, ('2007-06', 'indiana.edu'): 39, ('2012-04', 'ox.ac.uk'): 20, ('2006-03', 'longsight.com'): 1, ('2010-03', 'longsight.com'): 14, ('2013-09', 'indiana.edu'): 12, ('2008-05', 'g
```


Gline.htm

