Big Data Visualization Tools and Techniques Applied to the Examination of flickr Picture Locations in Selected Cities

Tidewater Big Data Enthusiasts
Chuck Cartledge
Developer

Wednesday 27th January, 2016 at 11:13

Contents

List of Figures **ii**

1 Introduction **1**

2 Approach **1**

3 Other things that can be done with geographic data **6**
3.1 Replacing city names with self descriptions 6
3.2 Real time data 7
3.3 Showing changes in a system over time 7

4 Visualization tools **13**

5 Other sources of information **31**

6 Conclusion **32**

7 References **33**

A Fairbanks, AK **34**

B Orlando, FL **37**

C Honolulu, HI **40**

D Reno, NV **43**

E Tidewater, OR **46**

F San Antonio, TX **49**

G Chesapeake, VA **52**

H Hampton, VA **55**

I Norfolk, VA **58**

J Portsmouth, VA **61**

K Virginia Beach, VA **64**

List of Figures

1 Sample interactive map. 4
2 Notational representative ZCTAs 6
3 California map from "A More Perfect Union" by R. Luke DuBois. 8
4 Real time wind vectors for the continental US. 9
5 Real time wind vectors from the Weather Underground. 10
6 Real time wind vectors from the Weather Chanel. 11
7 Dynamic system change over time. 12
8 tableau home page. 14
9 tableau sample. 15
10 watson home page. 16
11 watson sample. 17
12 gephi home page. 18
13 gephi sample. 19
14 imagePlot home page. 20
15 imagePlot sample. 21
16 treeMap home page. 22
17 treeMap sample. 23
18 indieMapper home page. 24
19 indieMapper sample. 25
20 geoCommons home page. 27
21 geoCommons sample. 28

22 arcGIS home page. 29
23 arcGIS sample. 30
24 Simple plot Fairbanks, AK 34
25 ggplot plot Fairbanks, AK 36
26 Simple plot Orlando, FL 37
27 ggplot plot Orlando, FL 39
28 Simple plot Honolulu, HI 40
29 ggplot plot Honolulu, HI 42
30 Simple plot Reno, NV 43
31 ggplot plot Reno, NV 45
32 Simple plot Tidewater, OR 46
33 ggplot plot Tidewater, OR 48
34 Simple plot San Antonio, TX 49
35 ggplot plot San Antonio, TX 51
36 Simple plot Chesapeake, VA 52
37 ggplot plot Chesapeake, VA 54
38 Simple plot Hampton, VA 55
39 ggplot plot Hampton, VA 57
40 Simple plot Norfolk, VA 58
41 ggplot plot Norfolk, VA 60
42 Simple plot Portsmouth, VA 61
43 ggplot plot Portsmouth, VA 63
44 Simple plot Virginia Beach, VA 64
45 ggplot plot Virginia Beach, VA 66

1 Introduction

The Tidewater Big Data Enthusiasts (TBDE) meet-up was established in September 2015 as a social organization of people interested in all aspects of Big Data. Big data is not a well defined term, and seems to vary from source to source [1, 2, 3, 5, 7, 8, 10]. TBDE has taken the approach that any processing that cannot be done in a timely manner in a single machine falls under the Big Data umbrella. One of these areas is how to visualize data of different types in different ways. This report contains a few different ways to look at data from flickr, a picture sharing site. The number of pictures, their location, time of day, and time of date could be used as a surrogate for the how popular a location is when. (The reality is that we are only seeing what people uploaded to flickr.) Some of the ideas in this report are base on Yau's work [13, 14].

To set the "scene," we are going to assume three different actors in this Big Data exploration. They are:

- **Data visionary**: the visionary source or leader of the exploration. This is the one who says: how about we . . . , or what would happen if we . . . , or can we find a relationship between . . . A visionary has original ideas about what could be.
- **Data broker**: the source of the data. Any particular data exploration may involve more than one data broker. The data provided by the broker could be in any format, and could be very messy.
- **Data scientist**: the one who "connects the dots" between what the broker provides and what the visionary sees. The data scientist makes the vision real. A data scientist may engage in "data exploration" while working towards the visionary's goal. The classical steps in data exploration are:
  - Computing descriptive statistics as needed (mean, mode, median, and so on),
  - Understanding the overall shape and flow of the data (where does the data come from, how was it collected, are the same collection techniques used through out the life of the data collection process, etc.),
  - Dispersion (is the data skewed to the left, right, up, down, and why), and
  - Visualizing the data (telling a story with the data).

2 Approach

Based on ideas from Yau's book [14], various aspects of pictures from flickr were investigated. The general question of interest is: *Where and when are areas popular based on flickr pictures?* The question could be used (in a very limited sense), to help city planners and tourist operators focus their attentions.

flickr has an web based application program interface (API)[1] to allow non-human access to their database of pictures. One of the API methods is called: *flickr.photos.search* which takes as some of its arguments a latitude longitude bounding box. Given a bounding box, flickr will return a list of photos that were claimed to be created within that box. Many times, there is a lot of meta-data associated with an image that the person creating the image (with a camera, or smart-phone) is unaware is being attached to the image. This additional information is called exchangeable image file (exif)[2]. The exif data can include things about the camera settings, time when the picture was created, where the picture was created, etc. When a person uploads an image to flickr, the exif data may be uploaded as well. The *flickr.photos.search* API method as well as others, works based on the exif data associated with the image.

The intended users (city planners, and tour operators) generally don't think in terms of latitude longitude bounding boxes. They think in terms of cities, states, locales, etc. A prototype program was written to:

1. Take a city state name combination,
2. Identify the ZIP Codes associate with that city[3],
3. Identify the bounding box for each ZIP Code Tabulation Area[4],
4. Convert it into a latitude longitude bounding box,
5. Query flickr,
6. Perform simple analysis on the data,
7. "Bin" the latitude and longitudes of all pictures returned from Flickr (bin size = 10^{-3} degrees of latitude $\approx$ 111 meters),
8. Query Google for a list of named location in each bin (the number of API calls that can be made per day is limited, so there is no guarantee that all bins will be serviced),
9. Report the 25 most often named locations, and
10. Create the report you are reading.

By default, the program collects data from 200 pictures on:

- Chesapeake, VA

[1] `https://www.flickr.com/services/api/`
[2] `https://en.wikipedia.org/wiki/Exchangeable_image_file_format`
[3] `http://www.unitedstateszipcodes.org/zip-code-database/`
[4] `https://www.census.gov/geo/maps-data/data/cbf/cbf_zcta.html`

- Fairbanks, AK
- Hampton, VA
- Honolulu, HI
- Norfolk, VA
- Portsmouth, VA
- Reno, NV
- Virginia Beach, VA

For each city, three plots are created:

1. A simple scatter plot of each picture based on its latitude and longitude. The title for this plot shows the number of pictures that were evaluated. This number might be less than the number available from flickr. This plot is typical of "quick and dirty" data exploration, and may not be too helpful.

2. A time oriented plot of the pictures. Each picture has the date and time it was created. The date is converted to a day number and plotted on the X-axis. The time the image was created is converted to a 24 hour time and plotted on the Y-axis. The combination shows seasonal, and day time variations. The title for this plot shows the range of dates as reported by the exif data. In some cases the dates are obviously wrong (dates in the future, or probably dates from before 1970).

3. A geographic plot based on ggplot data showing where each picture was taken. This is useful to get a feel where things are happening.

For each city, a summary table of named locations is created using the gridded data. Each grid is passed to a Google API that returns all the named locations within that grid. This can be a time consuming process because the API limits how many calls can be made per day, and how fast the calls can be made. A record is kept of every named location returned by the API. After all the API calls are made, the records are sorted to identify the 25 most often returned names. The summary table lists these "often named" locations. Some names are very generic, while others are uniquely identifiable.

The program has the option to create an interactive image for external use (see Figure 1). The option is normally disabled because of the length of time it can take to start internal JavaScript processes. An image with a few hundred icons (as in the sample figure), is fairly fast. An image with thousands of icons can take a long time to load. Once the image is loaded, processing time to pan the image, zoom in and out, change aspect ratios, etc. is very fast.

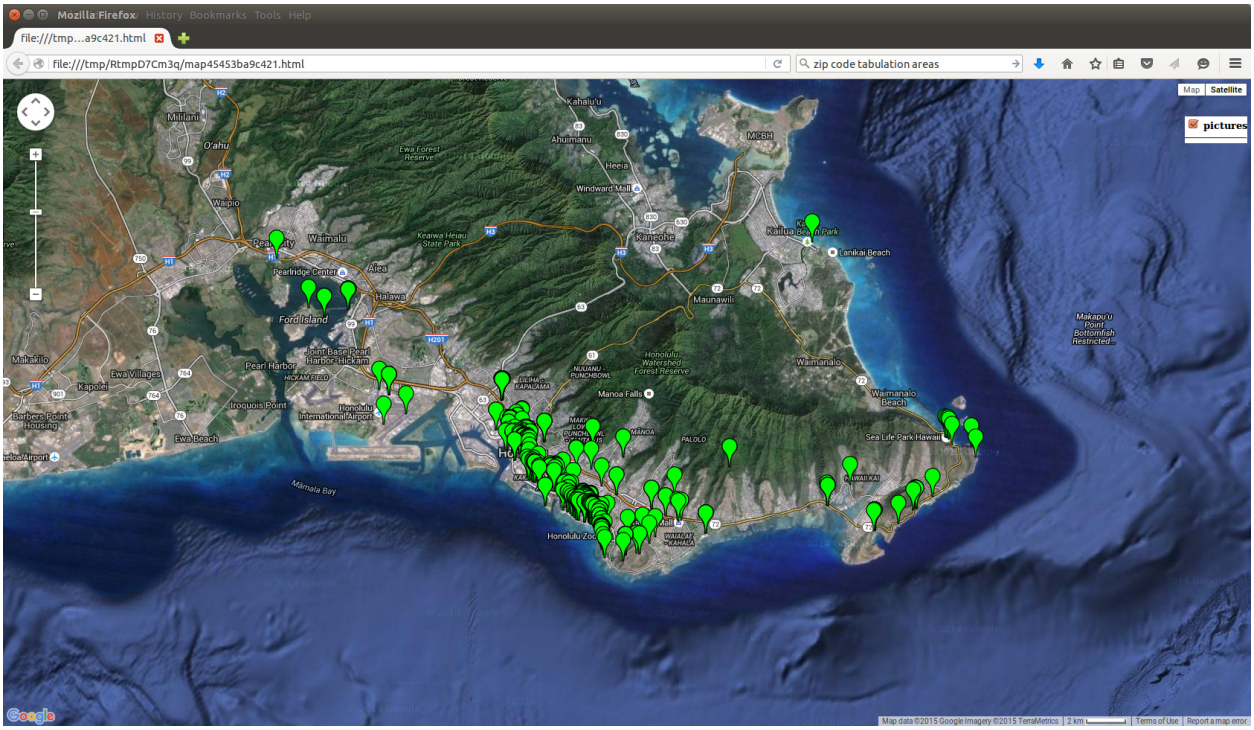
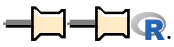


Figure 1: Sample interactive map. The interactive image can be VERY slow to load depending on the number of icons that are displayed.

The program is attached to this PDF report  .

Notes on using "binned" geographic data.

The data returned from flickr and Google is approximate because of the binning of the location data. Binning data is a "quick and easy" way to get a feel for how data looks. The technique works well when the data is "reasonable." Geographic data may not be reasonable. The geographic extents (its minimum and maximum values) in latitude and longitude is based on the extents of the ZIP Code Tabulation Areas (ZCTAs) associated with that city. If a ZCTA is rectangular then its extents accurately reflect the area it encompasses (see Figure 2a). If the ZCTA is roughly circular, then area in the binned extents are not part of the ZCTA but the binned extents will be included (see Figure 2b). An extreme case of binned data being too generous would be a fish hook shaped area, where the majority of the area is outside the ZCTA. Both the flickr and Google APIs use latitude and longitude extents. Both may return data that is outside the ZCTAs of interest. The collectDataFromFlickrWorker function in the flickr.R R script, could be improved by checking the location of each flickr image to see if it was inside a ZCTA rather than just blindly assuming it is correct[a].

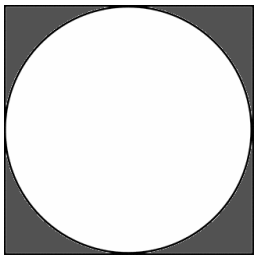
[a] Application of the point.in.polygon() function from the *sp* package may be appropriate.

Notes on flickr data.

flickr is the source of the data used in this report. People are the source of the data that flikr provides to us. Our software makes very few assumptions about the quality of the data. Detailed analysis of the data is beyond the purpose of this report. Very limited analysis is done, and is summarized in the title of the plot. These are a few things to watch for in the titles:

- the percentage can be greater than 100 because the total number of pictures available can decrease while downloading data,
- there have been cases where the positional information is inconsistent (or unreliable),
- pictures taken before 2000 may be questionable,
- pictures taken in the future are probably timestamped incorrectly.

(a) A rectangular representative.



(b) A circular representative.

Figure 2: Notational representative ZCTAs In the rectangular representation, the latitude and longitude extents accurately describe the ZCTA. In the cirular ZCTA, the latitude and longitude extents include areas in the corners that are not really part of the ZCTA.

3 Other things that can be done with geographic data

3.1 Replacing city names with self descriptions

> *"For a series of maps called 'A More Perfect Union' (2011), included in the Ringling exhibition, Mr. DuBois took a romantic census of the United States. Pushed by friends into joining an online dating site after a breakup, he became obsessed with the words people use to describe themselves in their profiles. He ultimately joined 21 dating sites as a straight man, gay man, straight woman and gay woman in every ZIP code in America, downloaded 19 million profiles, and algorithmically determined the most common word in each location, which he inserted in place of the city names on his maps. New York is 'Now,' Seattle is 'Heartbreak,' Atlanta is 'God.' Upper Peninsula of Michigan is dotted with 'Masochist,' 'Depression,' 'Futile,' 'Rustic,' 'Fairytale."'*
>
> H. M. Sheets [11]

3.2 Real time data

With a little creativity, it is possible to make something aesthetically pleasing out the mundane[5]. Wind can become a work of art (see Figure 4). Data taken from the National Weather Service (NWS) National Digital Forecast Database (NDFD)[6] is available for download. The gridded data is available hourly. Grids for the CONUS are currently available from NDFD at 5 kilometer spatial resolution. The spatial resolution for the grids for Hawaii and Guam is 2.5 kilometers; for Puerto Rico/the Virgin Islands is 1.25 kilometers; for Alaska, 6 kilometers. For the North Pacific Ocean Domain the spatial resolution is 10 km.

When you take the current conditions at each weather site, and imagine a speck of dust being blown past that site and on to another, then you can trace the wind as it moves across the country. Compare the image in Figure 4 to the ones from the Weather Underground[7], and the Weather Channel[8] (see Figures 5 and 6).

3.3 Showing changes in a system over time

There are times when it makes sense to display data dynamically, such as showing how a system changes or evolves over time. An example of this is showing how well a piece of digital data preserves itself and spreads across a network of computers [9] (see Figure 7). A full explanation of the system dynamics is in the figure's caption.

[5] http://hint.fm/wind/

[6] http://ndfd.weather.gov/technical.htm

[7] http://www.wunderground.com/maps/?

[8] http://www.weather.com/maps/ususcurrentwindsgusts

[9] http://www.youtube.com/v/SXDaNj_3nqw?hl=en&fs=1

Figure 3: California map from "A More Perfect Union" by R. Luke DuBois.

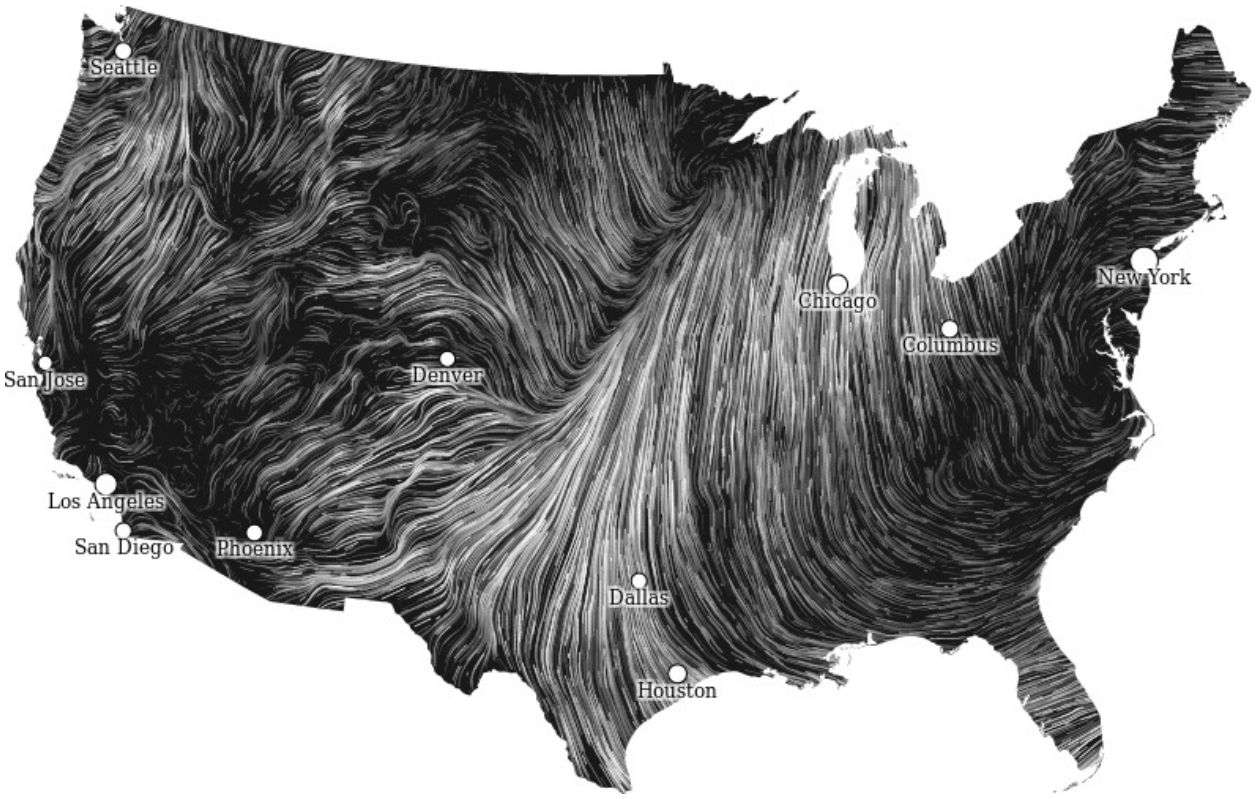


Figure 4: Real time wind vectors for the continental US. Real time image from `http://hint.fm/wind/`.

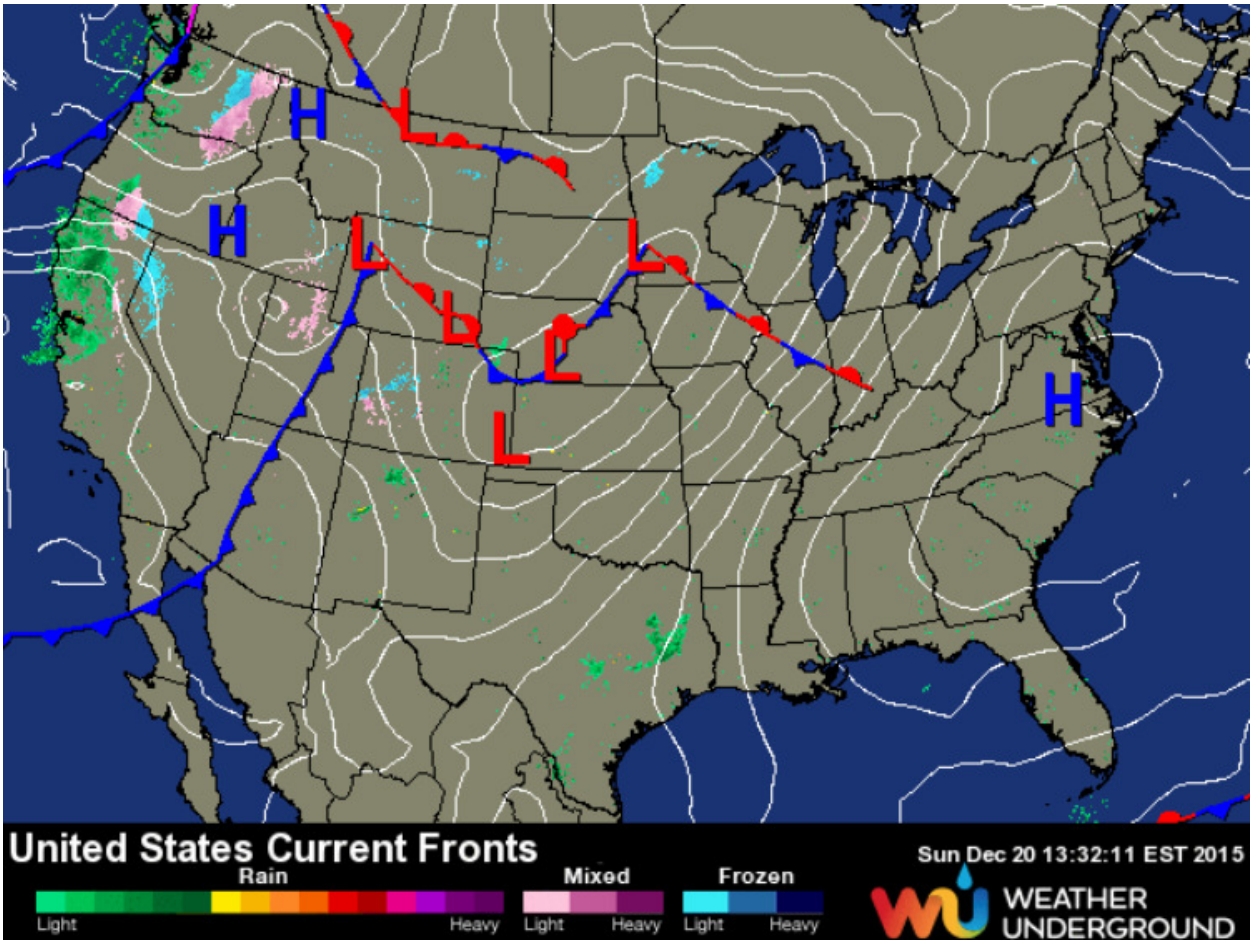


Figure 5: Real time wind vectors from the Weather Underground.

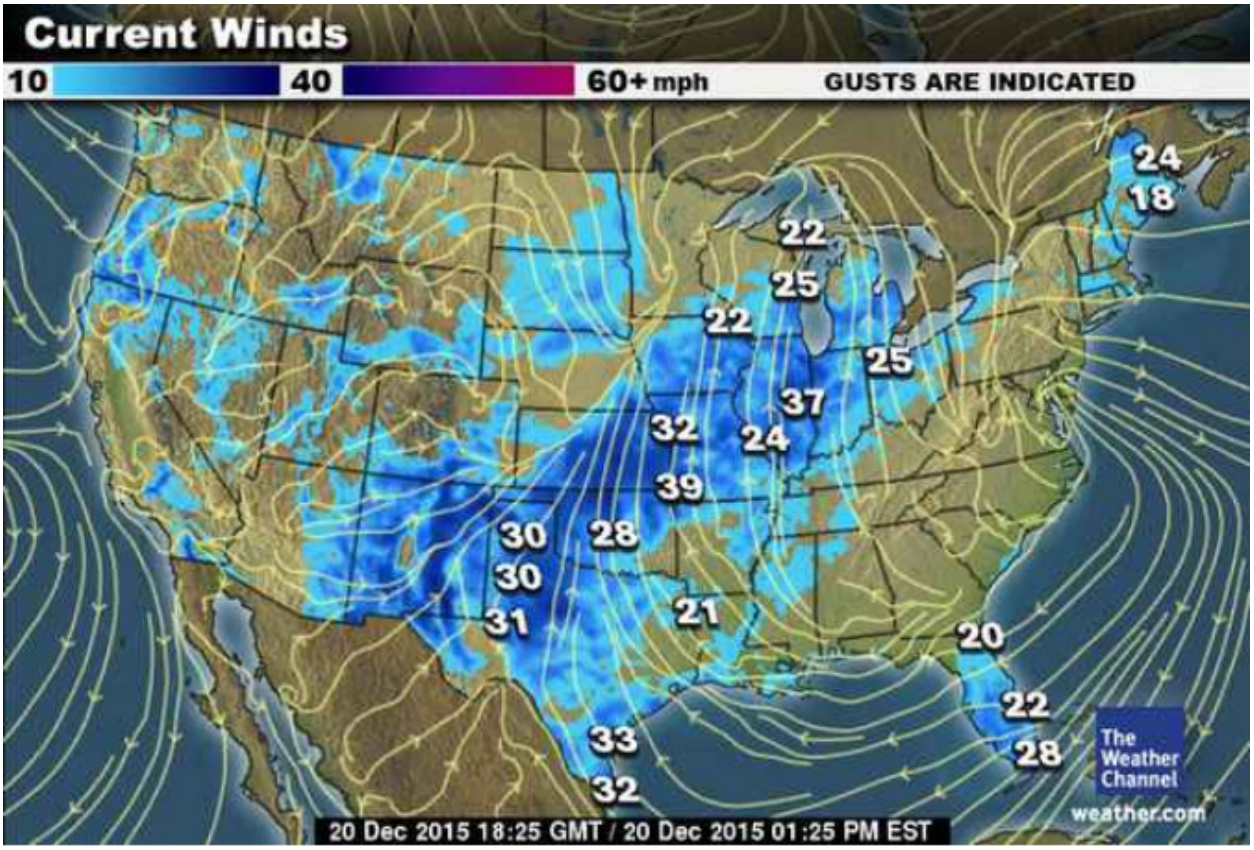


Figure 6: Real time wind vectors from the Weather Chanel.

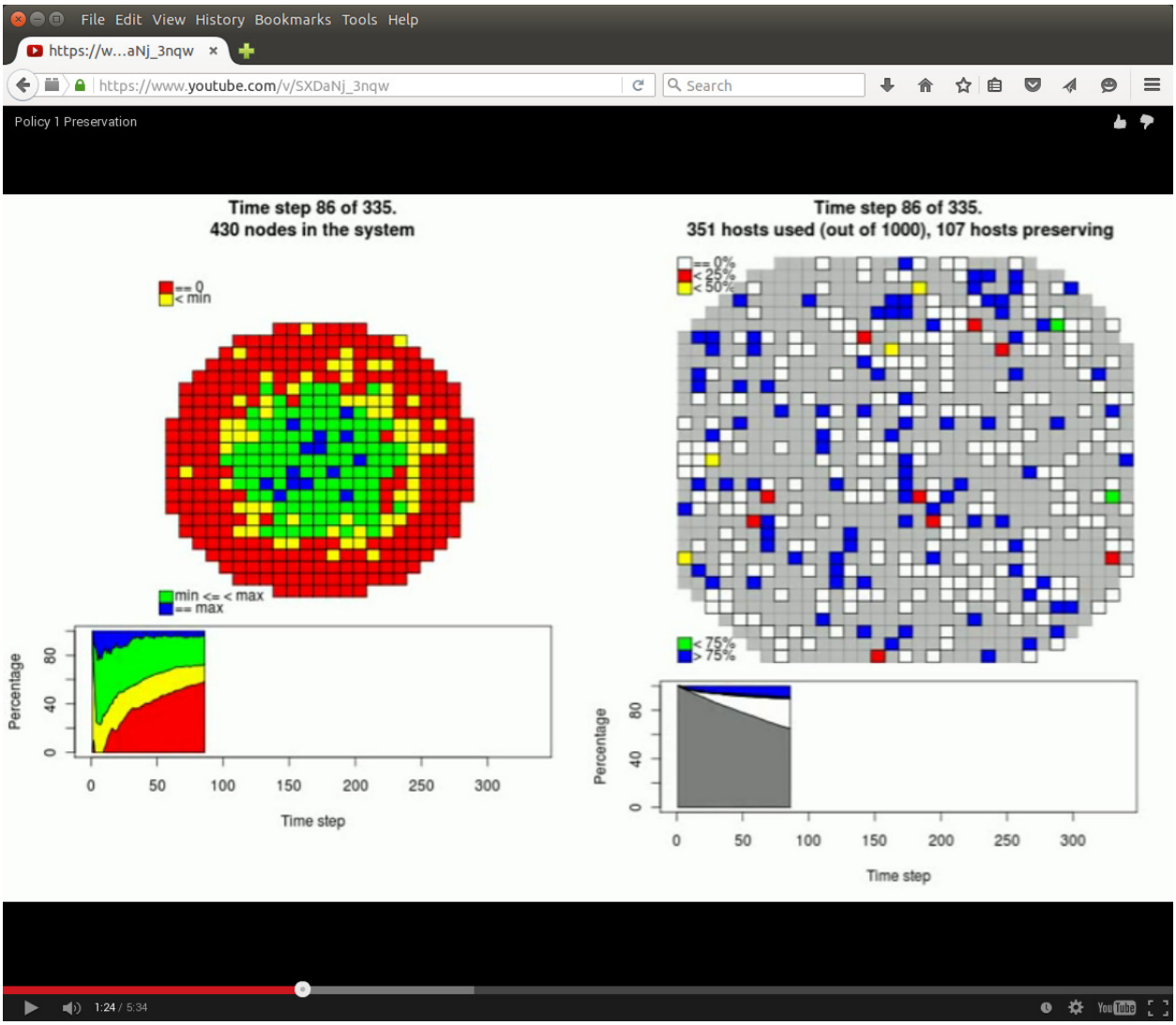


Figure 7: Dynamic system change over time. The left and right hand halves of the figure show different aspects of the system. The upper and lower halves of the halves show aggregate data about the system. In the left half, new "agent nodes" (pieces of software) are created at various time steps. Each node attempts to create copies of itself on different hosts (shown on the right half). A node has internal values for the minimum and maximum number of copies of itself it will create. It will attempt to create as many copies of itself as it can. The agent will sacrifice one of its copies to make room on a host for a different node's copy. Once it has created it's minimum number of copies, it will never go below that value. So, in the life of a node, it will have 0 or less then its minimum, it will work to create its maximum number, and it will sacrifice copies for the good of all nodes. Hosts (on the right half) have a finite storage capacity whose usage is shown as different colors for 0%, 25%, and so on. The histograms at the bottoms of the halves show the percentage of nodes, or hosts that are at different states. The full movie is available on YouTube `http://www.youtube.com/v/SXDaNj_3nqw?hl=en&fs=1`

4 Visualization tools

There are raft of data visualization tools. Each with their own strengths and weaknesses. Here is a partial list based on [14].

- **Microsoft Excel:** The familiar spreadsheet software is universal and has been around for decades. Its limited by the amount of data it can handle at once, and unless you know Visual Basic for Applications (VBA), the programming language built in to Excel, it can be a chore to reproduce charts for different datasets.

- **Google Spreadsheets:** Googles version of Microsoft Excel, but its simpler and online. You can quickly access your data across different machines and devices, and you can collaborate via built-in chat and real-time editing[10].

- **Tableau Software:** Tableau Software is the up-and-coming analysis software. If you want to dig deeper into your data than you can in Excel, without programming, this is a good place to look. It is pricey, but there are free options[11] (see Figures 8 and 9).

- **Many Eyes:** a project by IBM Research that enables you to upload datasets and explore your data via a wide variety of visualization tools, some traditional and others experimental. Many Eyes has been folded into IBM Watson Analytics[12] (see Figures 10 and 11). .

- **Gephi:** An open-source graphing software that enables you to interactively explore networks and hierarchy[13] (see Figures 12 and 13). .

- **ImagePlot:** enables you to explore large sets of images as if they were data points. Plot by color, time, or volume to, for example, explore the trends and changes of an artist or photo collection[14] (see Figures 14 and 15). .

- **Treemap:** Treemaps are useful for exploring hierarchical data in a small space[15] (see Figures 16 and 17). .

- **indiemapper:** Is a free service provided by cartography group Axis Maps. It enables you to create custom maps and map your own data, but it runs in the browser rather than as a desktop client[16] (see Figures 18 and 19). .

[10] `http://drive.google.com/`
[11] `http://tableausoftware.com`
[12] `http://www.ibm.com/analytics/us/en/technology/advanced-analytics/`
[13] `http://gephi.org`
[14] `http://lab.softwarestudies.com/p/imageplot.html`
[15] `http://www.cs.umd.edu/hcil/treemap/`
[16] `http://indiemapper.com/`

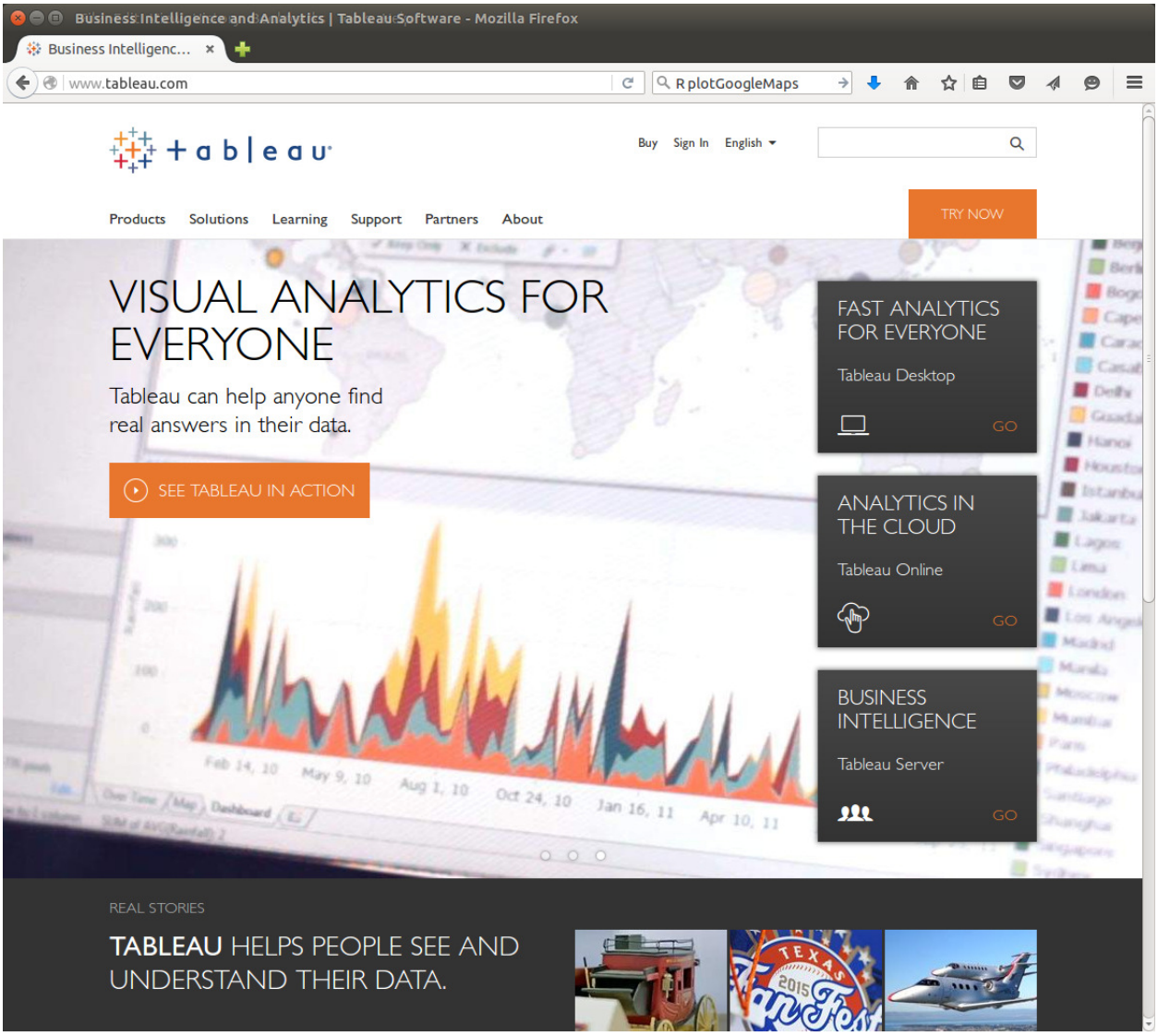


Figure 8: tableau home page.

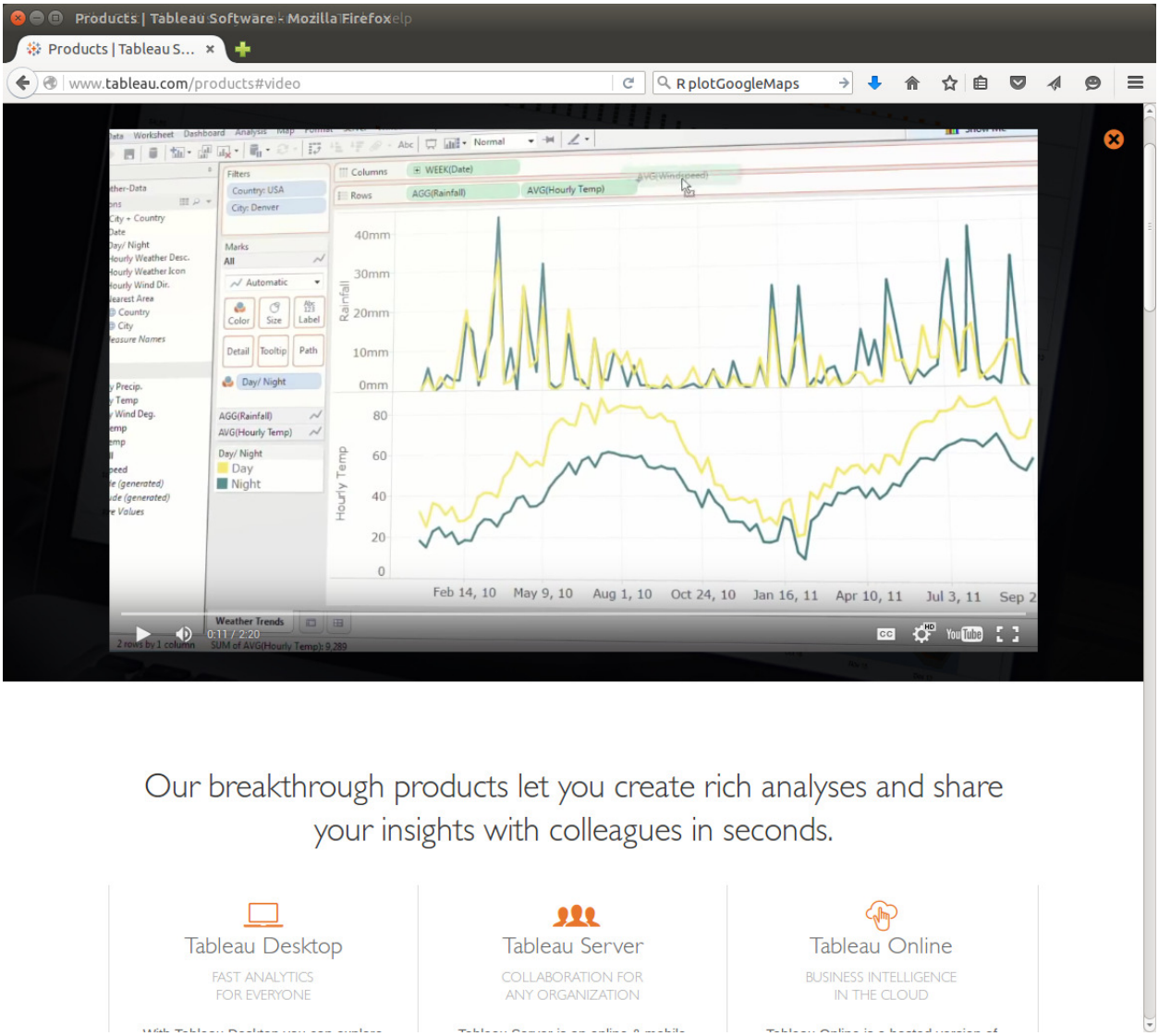


Figure 9: tableau sample.

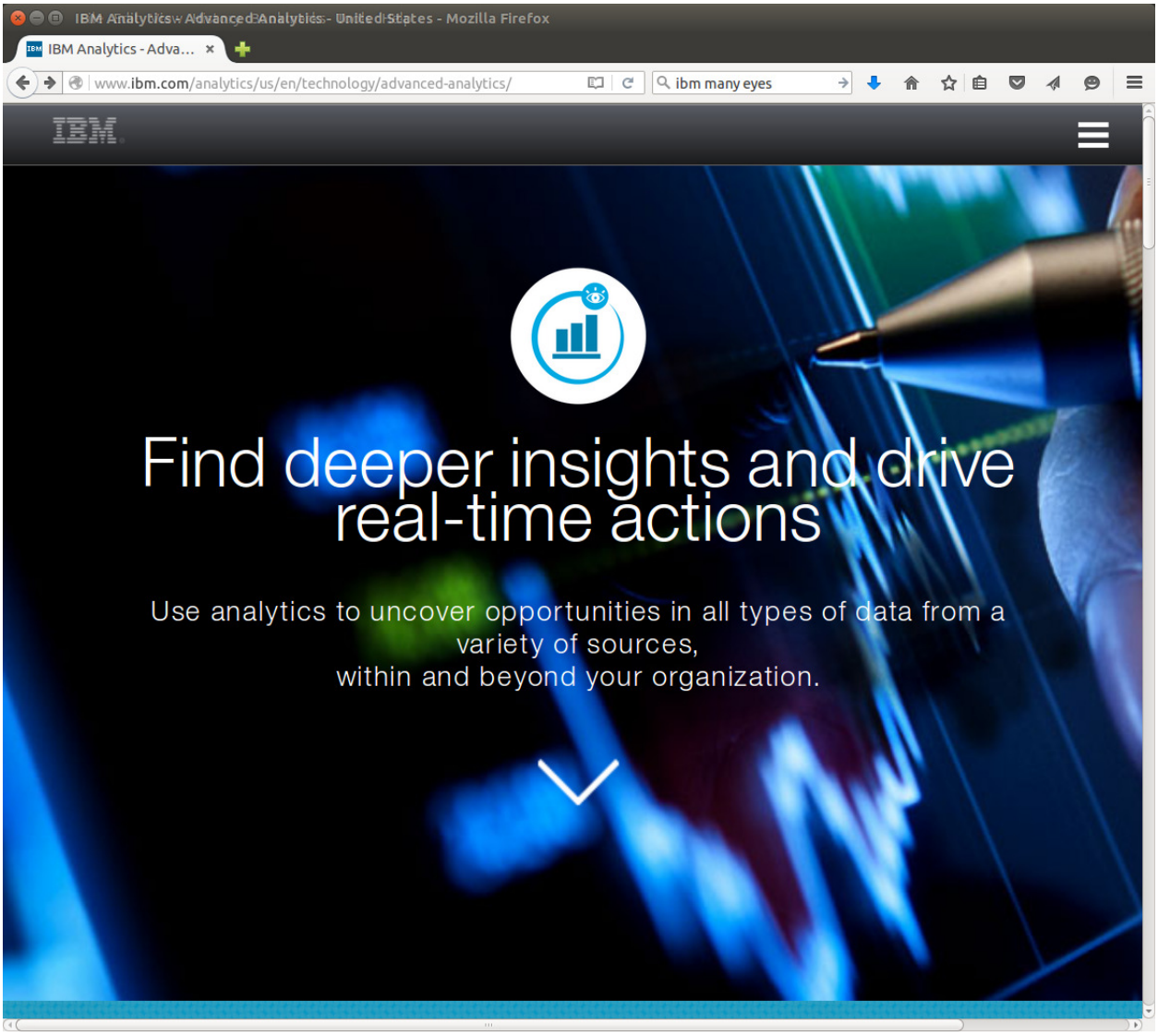


Figure 10: watson home page.

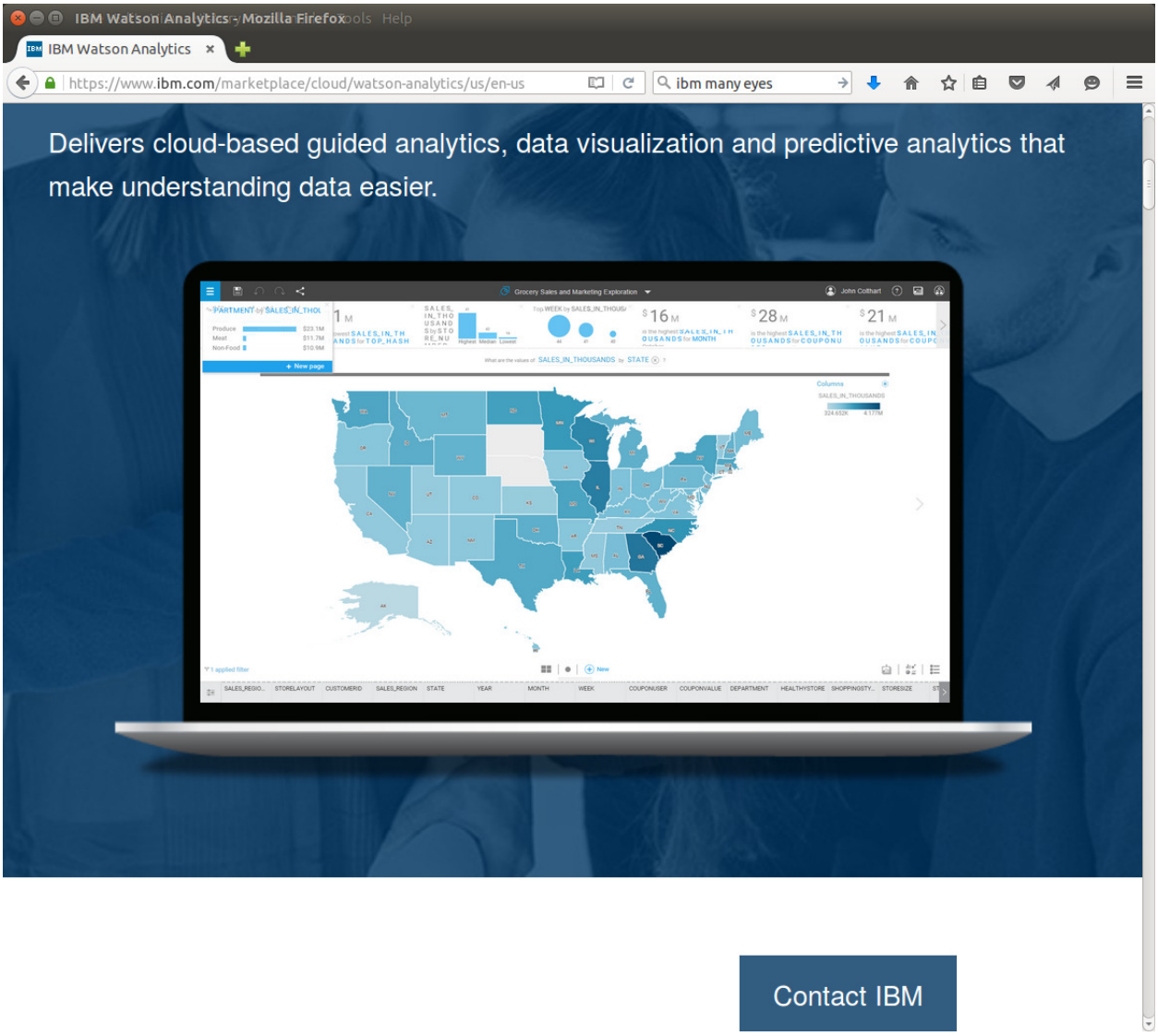


Figure 11: watson sample.

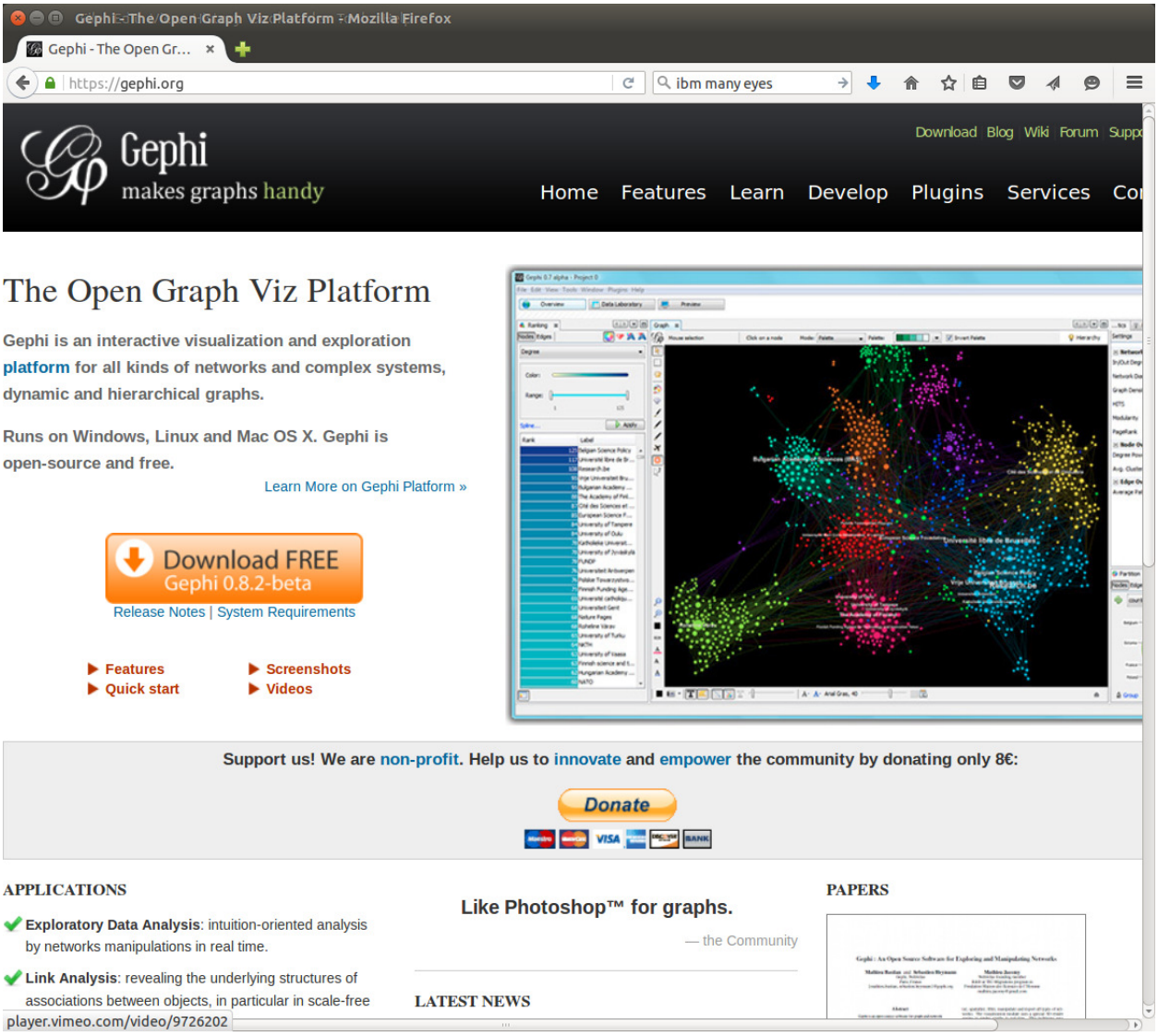


Figure 12: gephi home page.

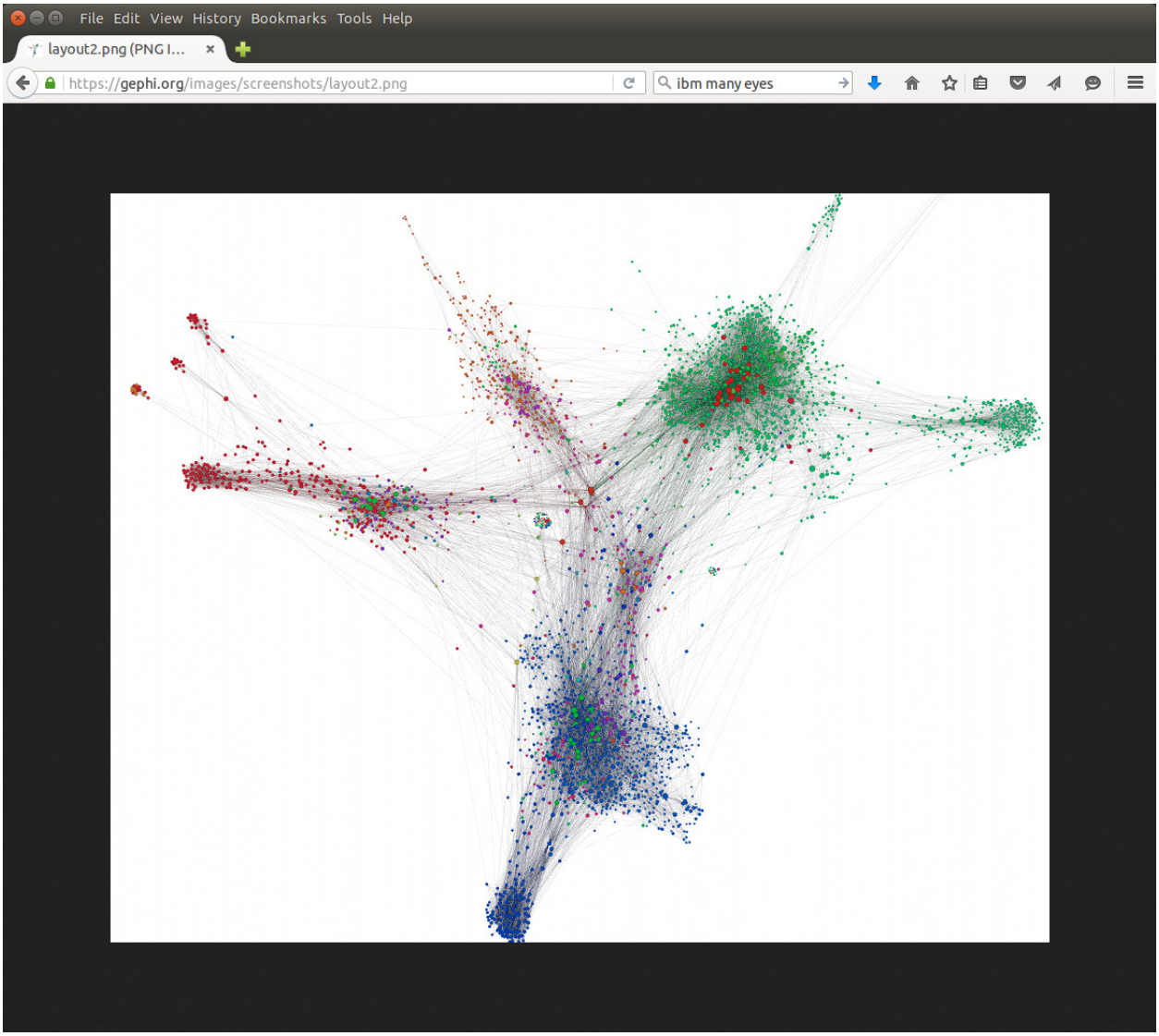


Figure 13: gephi sample.

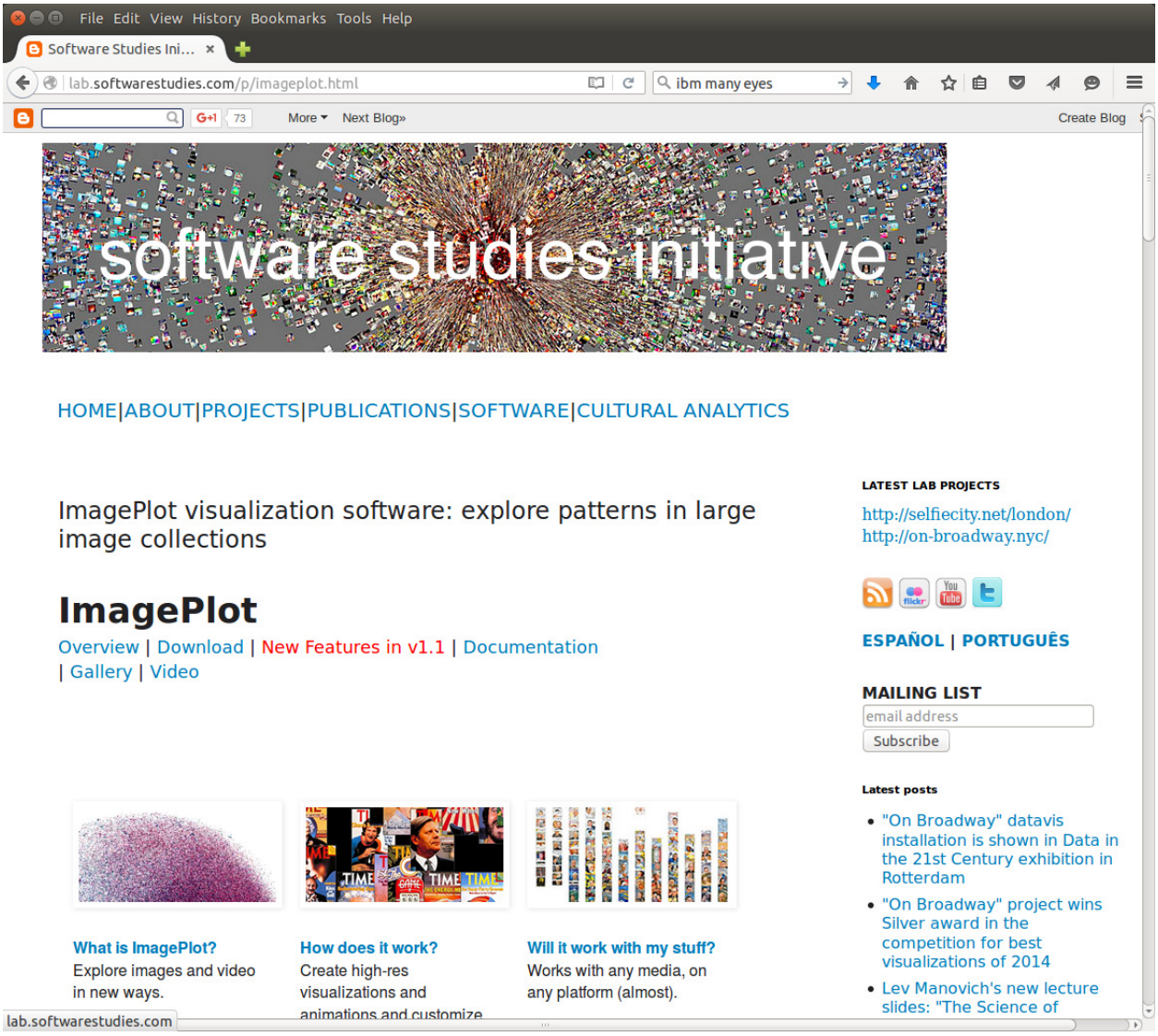


Figure 14: imagePlot home page.

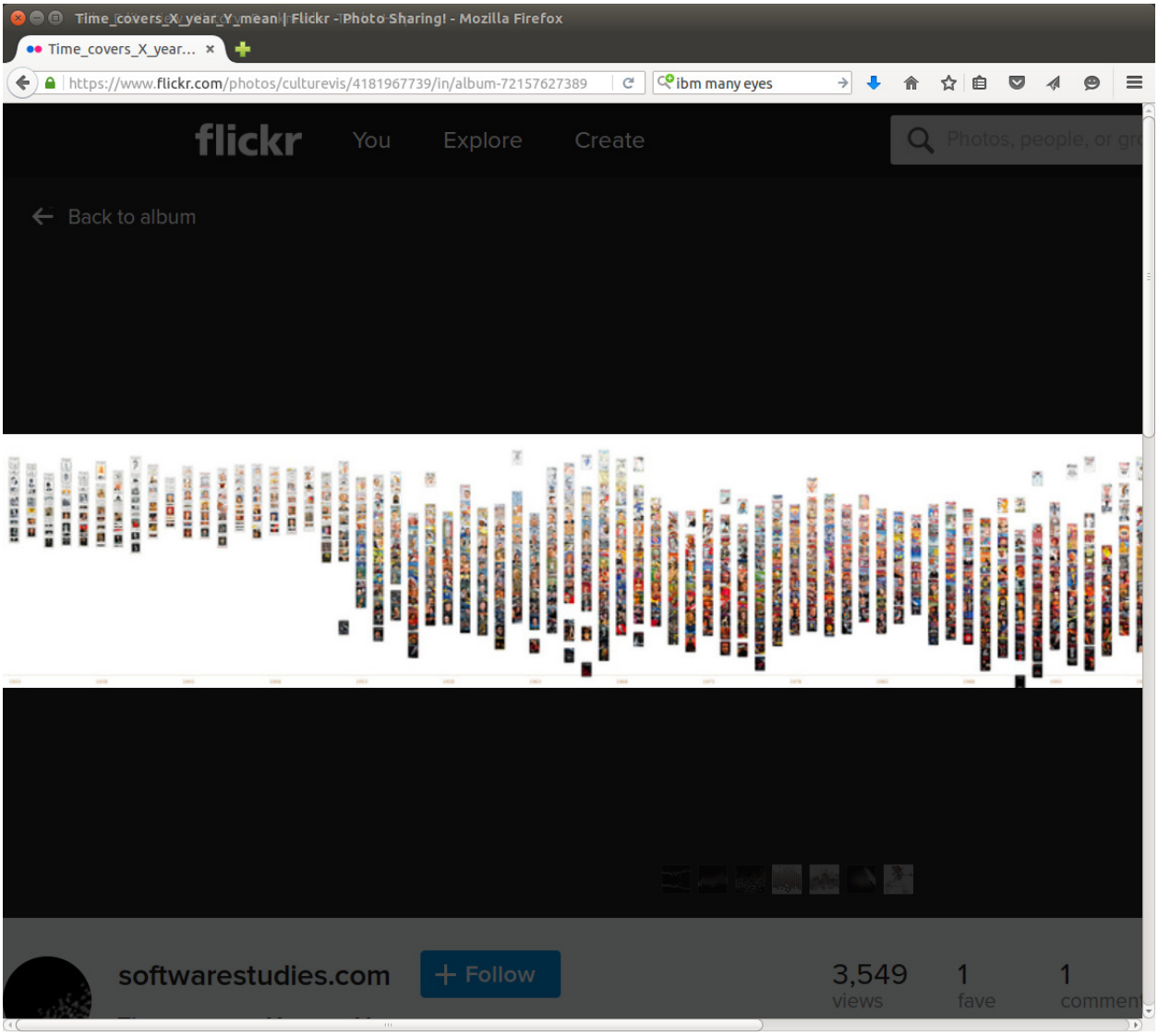


Figure 15: imagePlot sample.

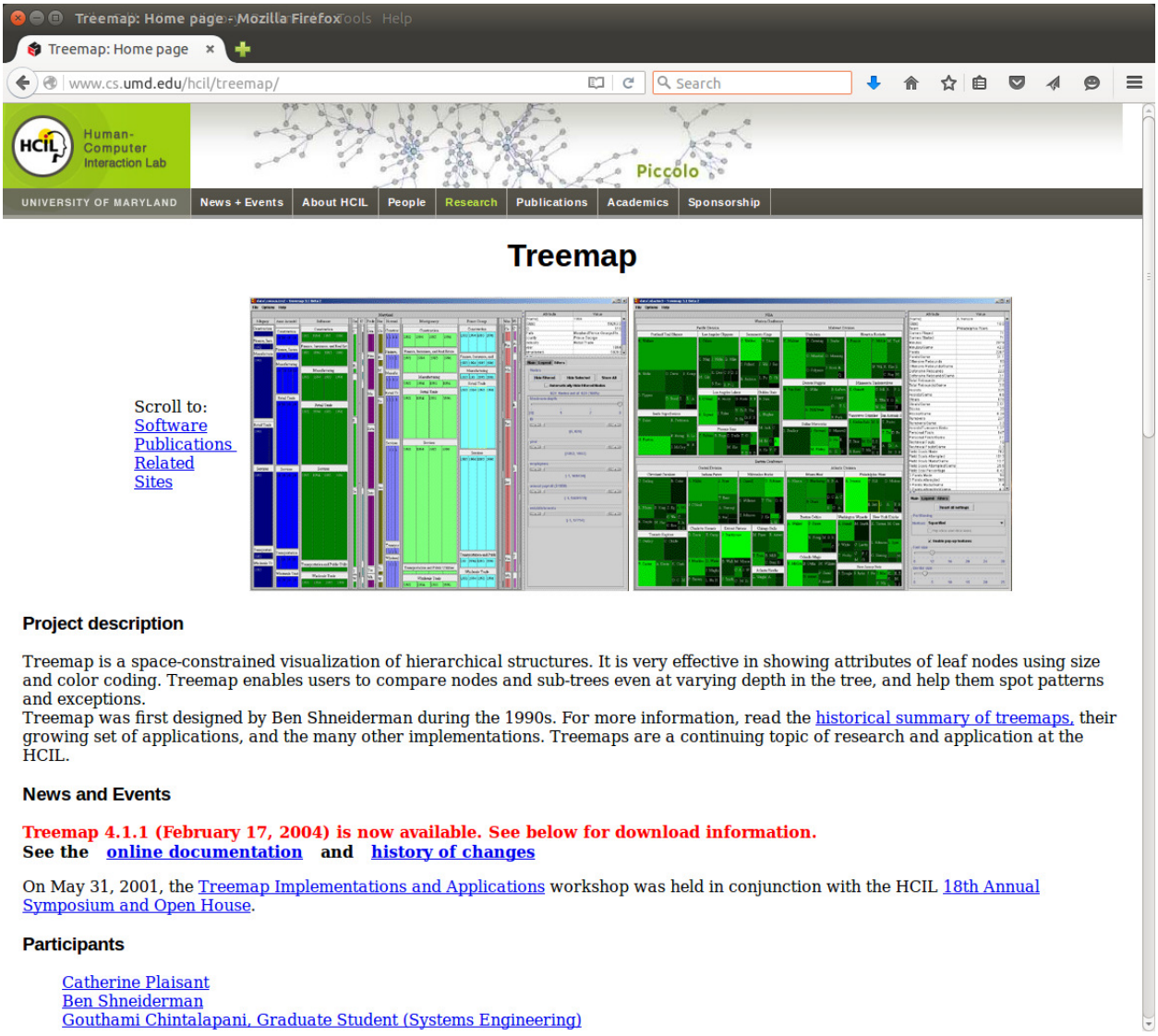


Figure 16: treeMap home page.

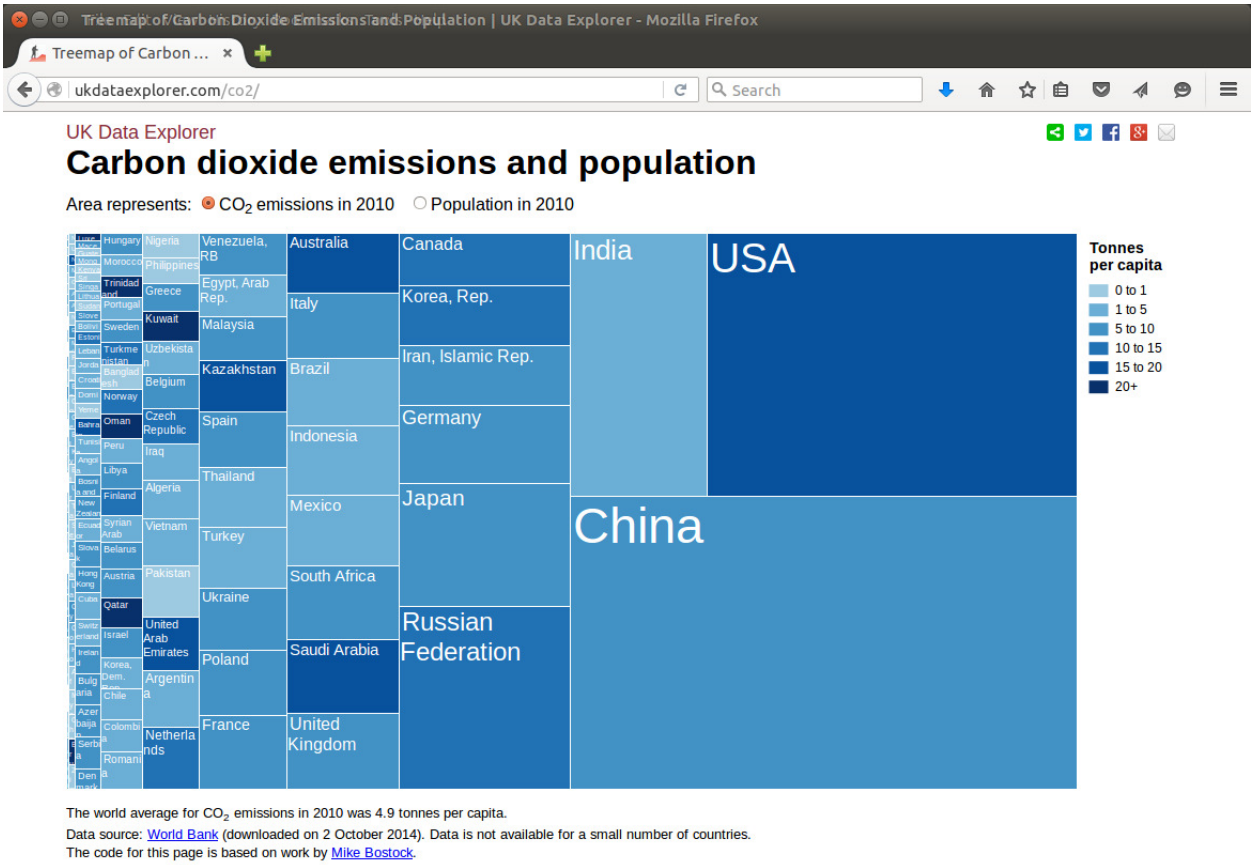


Figure 17: treeMap sample.

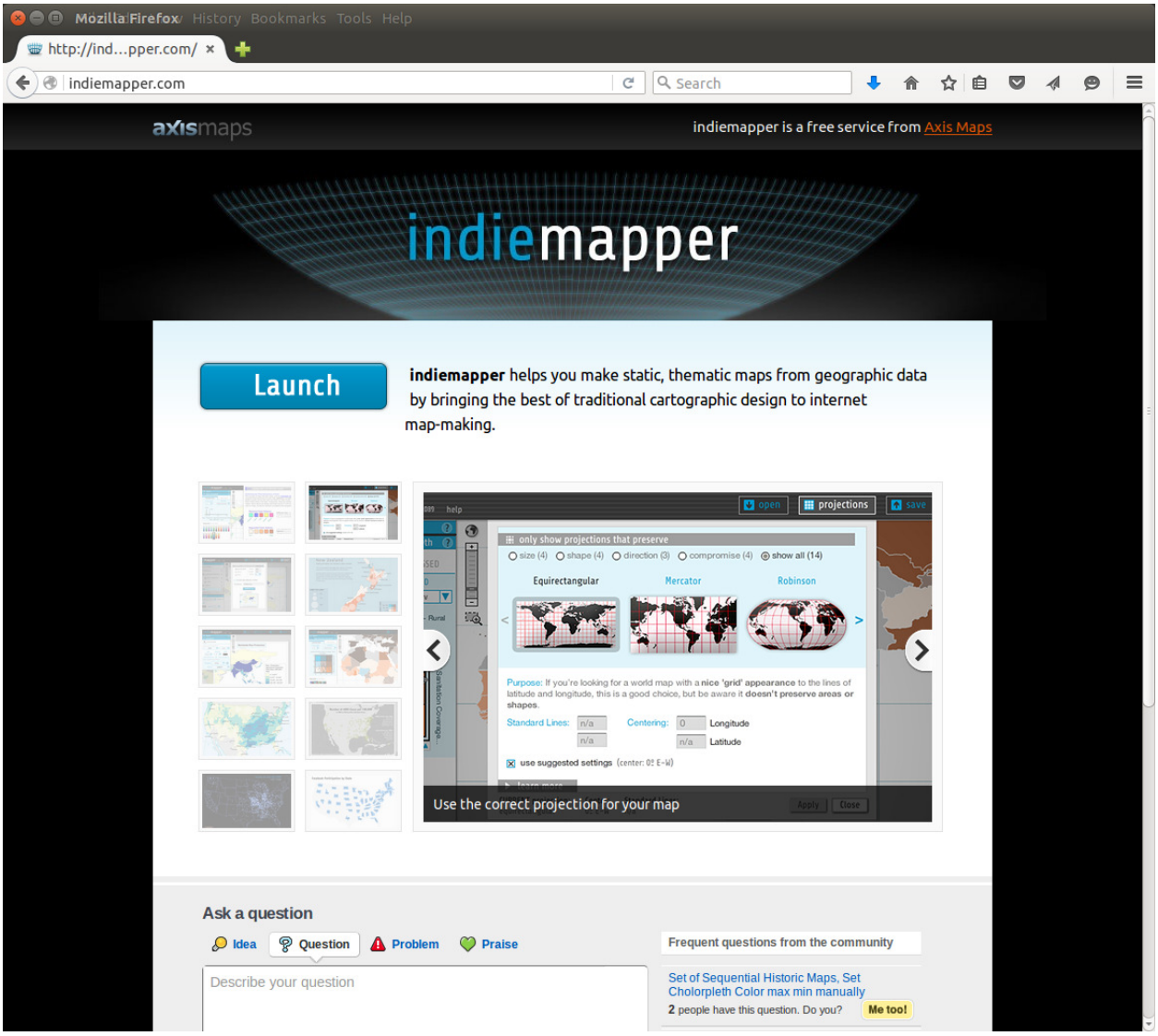


Figure 18: indieMapper home page.

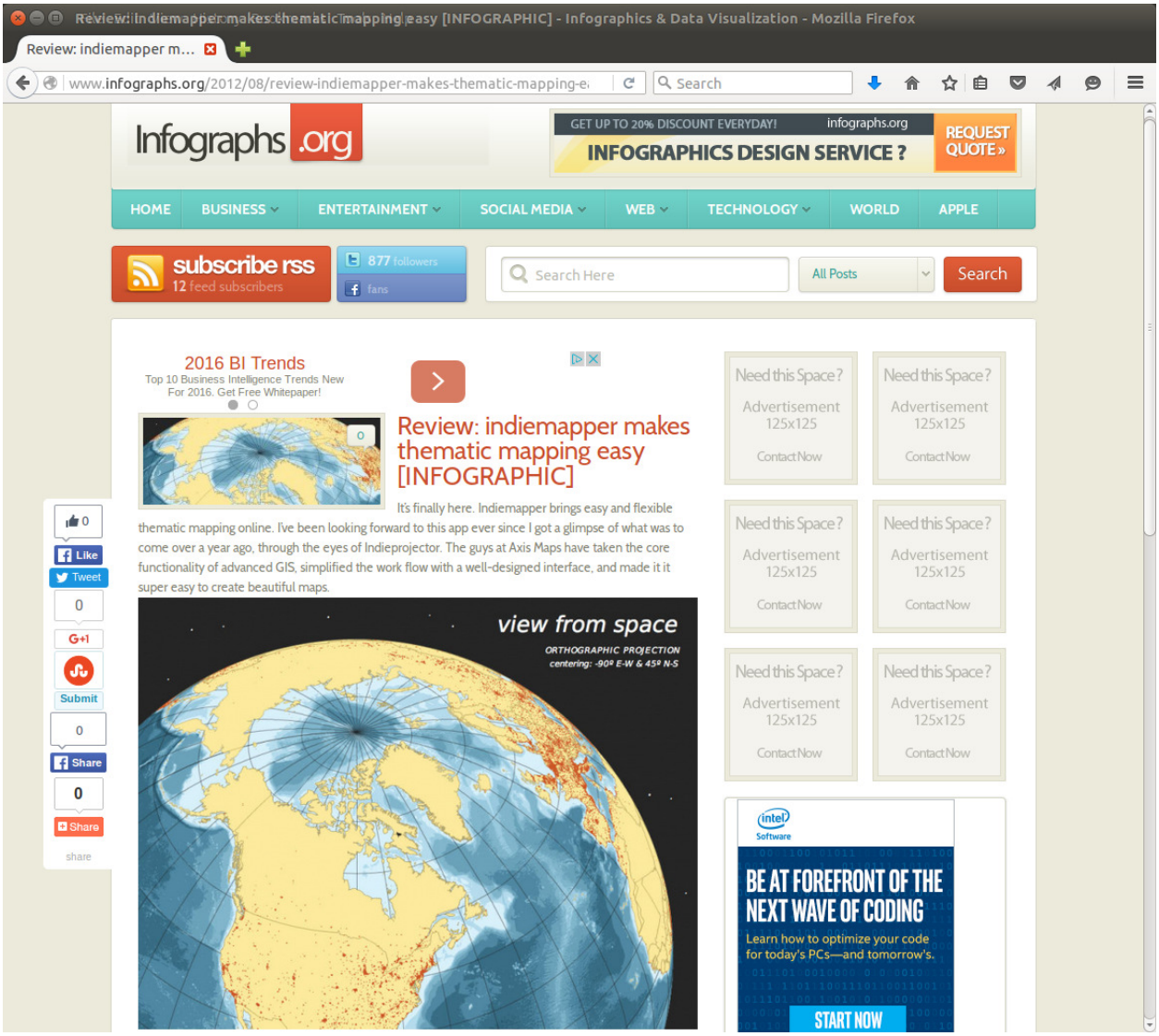


Figure 19: indieMapper sample.

- **GeoCommons:** GeoCommons is similar to indiemapper, but its more focused on exploration and analysis. You can upload your own data or draw from the GeoCommons database and then interact with points and areas[17] (see Figures 20 and 21). .
- **ArcGIS:** Before the previously mentioned mapping tools were available, ArcGIS was the primary mapping software for most people. Its a feature-rich platform that enables you to do just about anything with maps. For most though, the basic subset of features is enough, so to avoid the hefty cost of the software, its probably best to try the free options first, and if those arent enough, try ArcGIS[18] (see Figures 22 and 23).

[17] `http://geocommons.com/`
[18] `http://arcgis.com/`

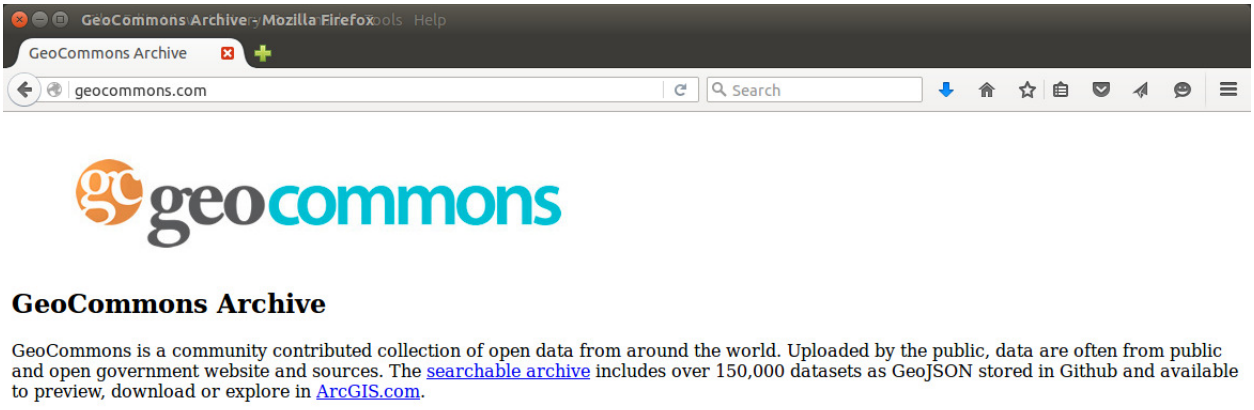


Figure 20: geoCommons home page.

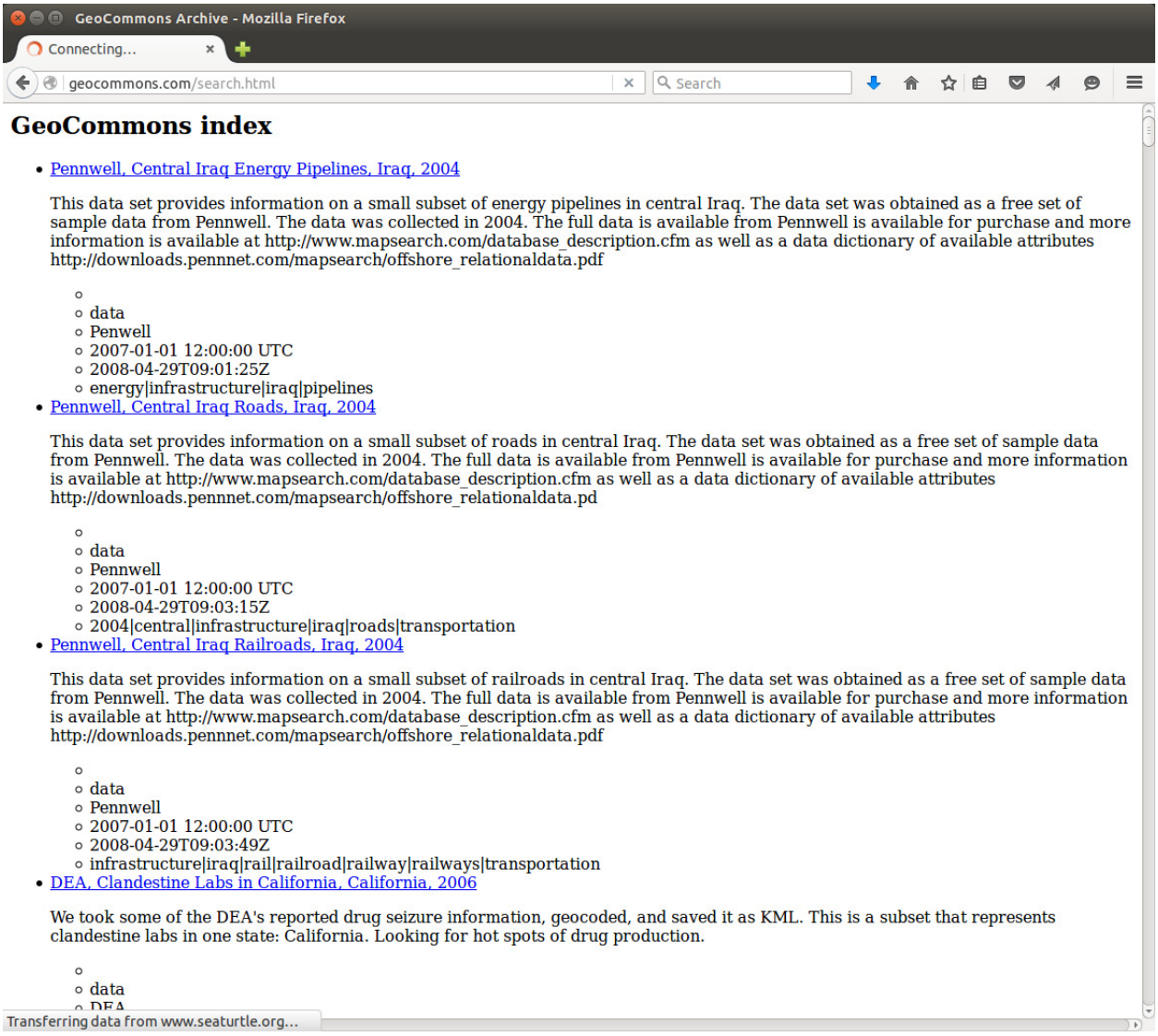


Figure 21: geoCommons sample.

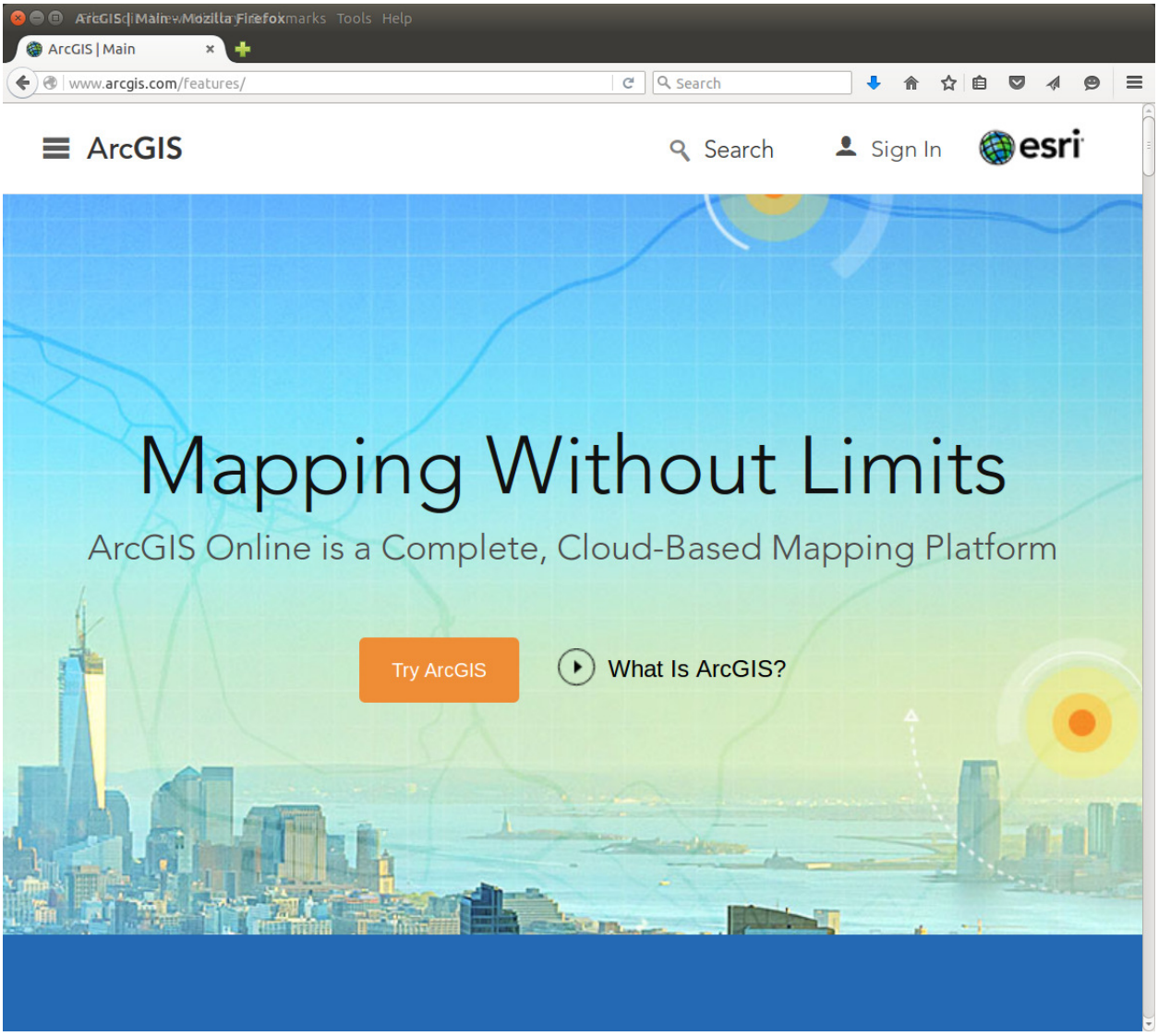


Figure 22: arcGIS home page.

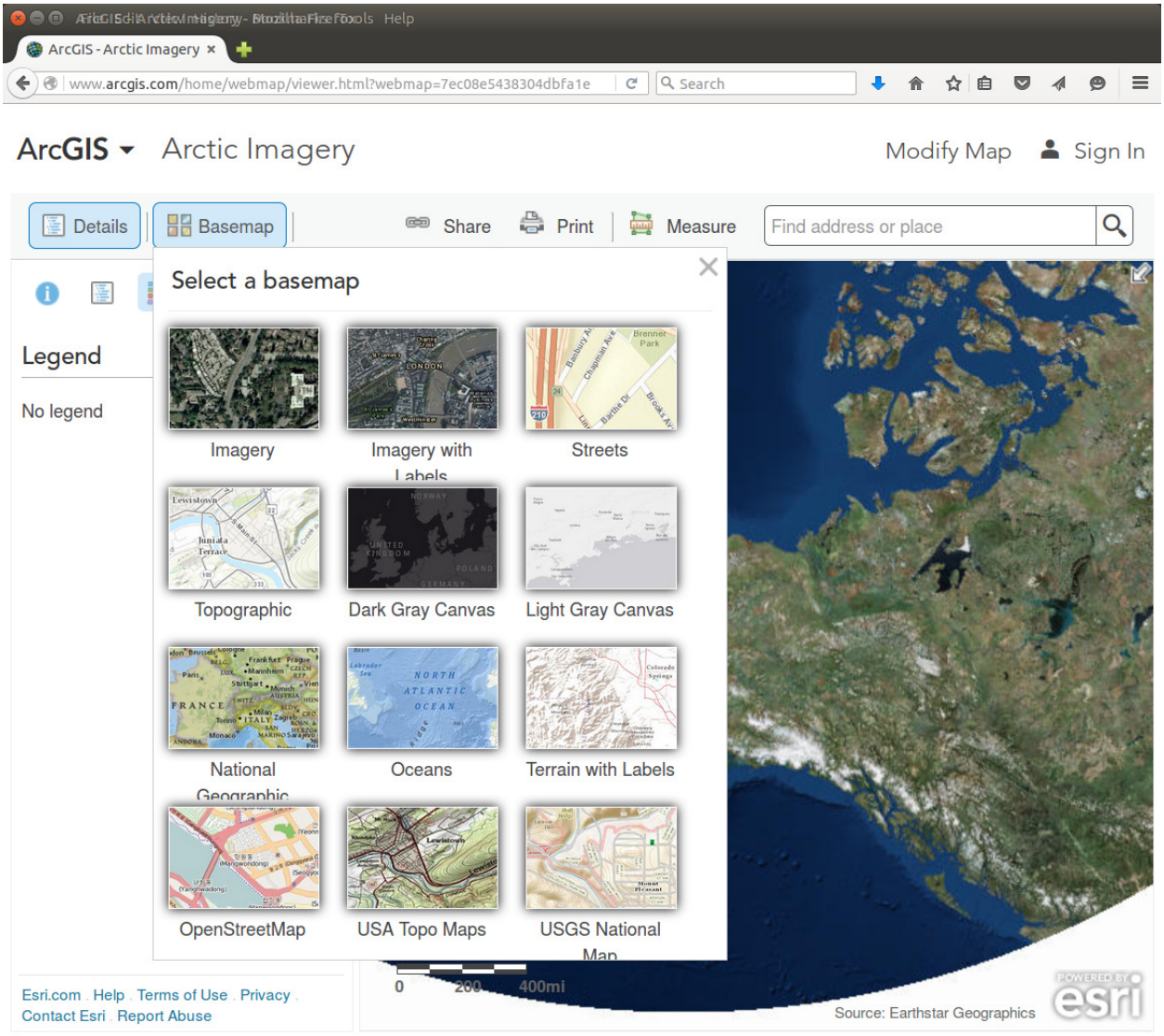


Figure 23: arcGIS sample.

5 Other sources of information

A collection of Big Data related sources that I have found useful.

- Books and papers
  1. **A Very Short History Of Big Data**, by Gil Press, Forbes Tech Blog[9]
  2. **Big data: A revolution that will transform how we live, work, and think**, by Viktor Mayer-Schönberger and Kenneth Cukier[8]
  3. **Data Points: Visualization That Means Something**, by Nathan Yau[14]
  4. **Information Graphics: A Comprehensive Illustrated Reference**, by Robert L. Harris[4]
  5. **The Pathologies of Big Data**, by Adam Jacobs in the Communications of the ACM[6]
  6. **The Visual Display of Quantitative Information**, by Edward R. Tufte [12]
- Internet sites
  1. **10 Cool Big Data Visualizations**: `http://www.mastersindatascience.org/blog/10-cool-big-data-visualizations/`
  2. **20 Inspiring Big Data Visualization Examples**: `http://www.keywebmetrics.com/2013/07/big-data-visualizations/`
  3. **Data is beautiful**: `http://imgur.com/r/dataisbeautiful`
  4. **Flowing Data**: `http://flowingdata.com/`
  5. **Mashup Guide**: `http://blog.mashupguide.net/`
  6. **Rich Blocks Poor Blocks**: `http://www.richblockspoorblocks.com/`
  7. **Run Keeper**: `https://runkeeper.com/developer/healthgraph/getting-started`

6 Conclusion

Data represents the "state of the world" at a particular point in time. Data visualization requires both statistics (at some level), and design knowledge. A good visualization will allow the data to tell a story that might not be visible in another form, or medium. Exposure to, and attempting to recreate data visualizations that are interesting work to increase the set of tools that can be applied to any visualization task.

The key to a good visualization is exposure to new ideas, and the correct tools to aid you in letting the data tell the story.

7 References

[1] Jules J Berman, *Principles of Big Data: Preparing, Sharing, and Analyzing Complex Information*, Newnes, 2013.

[2] Jim Gray, *The transaction concept: Virtues and limitations*, Very Large Databases, vol. 81, 1981, pp. 144–154.

[3] Christian Hagen, KHalid Khan, Marco Ciobo, and Jason Miller, *Big Data and the Creative Destruction of Today's Business Models*, http://www.atkearney.com/strategic-it/ideas-insights/article/-/asset_publisher/LCcgOeS4t85g/content/big-data-and-the-creative-destruction-of-today-s-business-models/10192, 2013.

[4] Robert L. Harris, *Information graphics: A comprehensive illustrated reference*, Oxford University Press, 1996.

[5] Joab Jackson, *The Big Promise of Big Data*, Business Software (2012).

[6] Adam Jacobs, *The Pathologies of Big Data*, Communications of the ACM **52** (2009), no. 8, 36 – 44.

[7] Doug Laney, *3d data management: Controlling data volume, velocity and variety*, META Group Research Note **6** (2001).

[8] Viktor Mayer-Schönberger and Kenneth Cukier, *Big data: A revolution that will transform how we live, work, and think*, Houghton Mifflin Harcourt, 2013.

[9] Gil Press, *A Very Short History Of Big Data*, Forbes Tech Blog http://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/ (2013).

[10] Eric Redmond and Jim R Wilson, *Seven databases in seven weeks*, Pragmatic Bookshelf, 2012.

[11] Hilarie M. Sheets, *Portraits from clips and bytes:r. luke dubois mines data to reveal art*, `http://www.nytimes.com/2014/01/12/arts/design/r-luke-dubois-mines-data-to-reveal-art.html`, 2014.

[12] Edward R. Tufte, *The Visual Display of Quantitative Information*, vol. 2, Graphics press Cheshire, CT, 1983.

[13] Nathan Yau, *Visualize this!*, John Wiley & Sons, 2012.

[14] ———, *Data points: Visualization that means something*, John Wiley & Sons, 2013.

A Fairbanks, AK

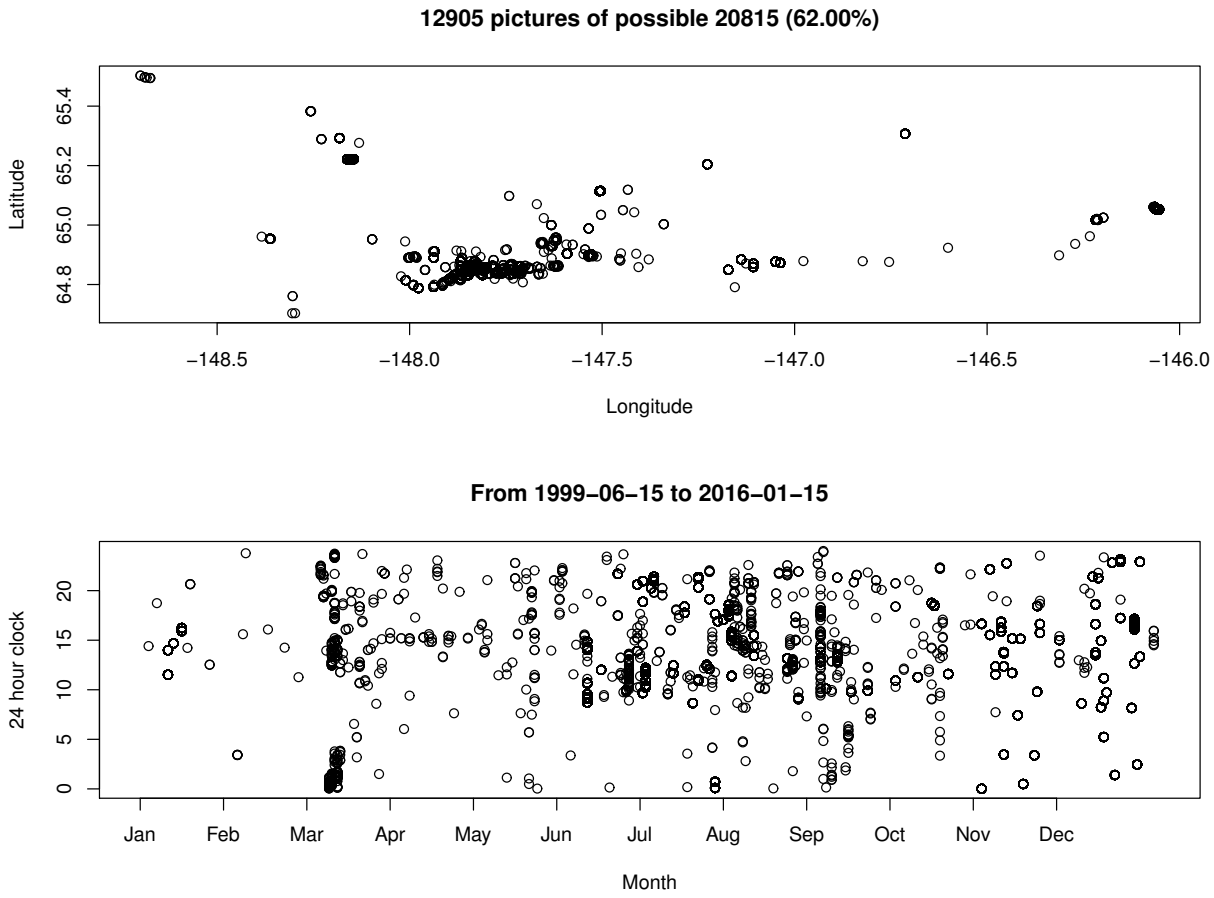


Figure 24: Simple plot Fairbanks, AK The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 1: Fairbanks, AK places of interest. The most frequently identified locations (586 out of 2,738) as per the Google API (489 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Arctic	151
2	College	127
3	Fairbanks	110
4	Downtown	20
5	Ester	20
6	Fox	13
7	Fudge Pot	8
8	Katn	8
9	District Court-Administrator	8
10	District Court-Jury Clerk	8
11	Lady Lee's Antiques Emporium	8
12	Livengood	8
13	District Court-Civil	8
14	Gold Rush Fine Jewelry	8
15	Two Rivers	8
16	River City Cafe and Espresso	8
17	Two Street Station	8
18	Festival Fairbanks	8
19	Alaska Airlines Inc	7
20	Cynde's Boutique	7
21	District Ct-Domestic Violence	7
22	Ebony and Ivory Gifts and Art	7
23	Alaska Weather Associates	7
24	Alaska Children's Proceedings	7
25	Gallery 49	7

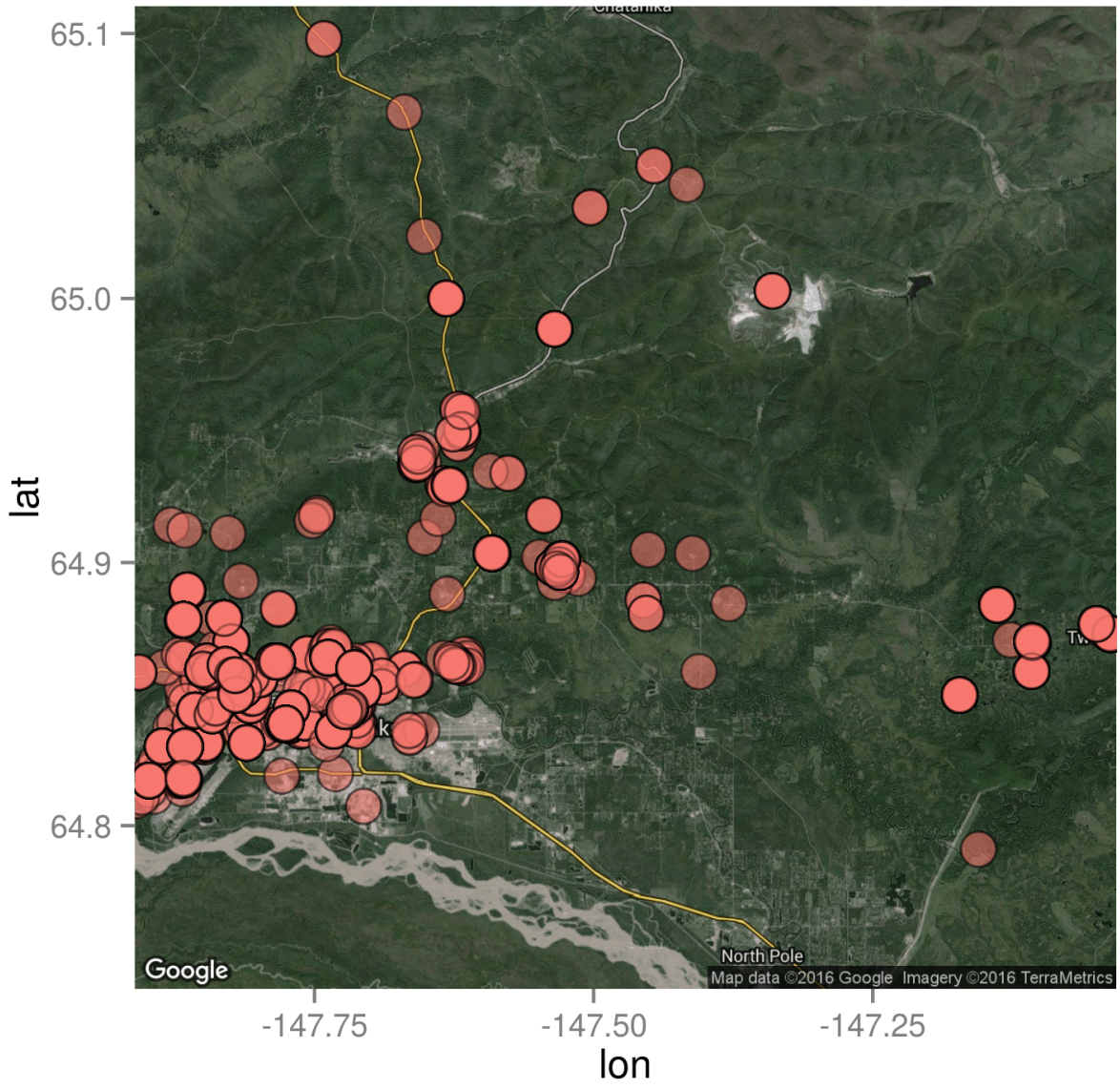


Figure 25: ggplot plot Fairbanks, AK The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

B Orlando, FL

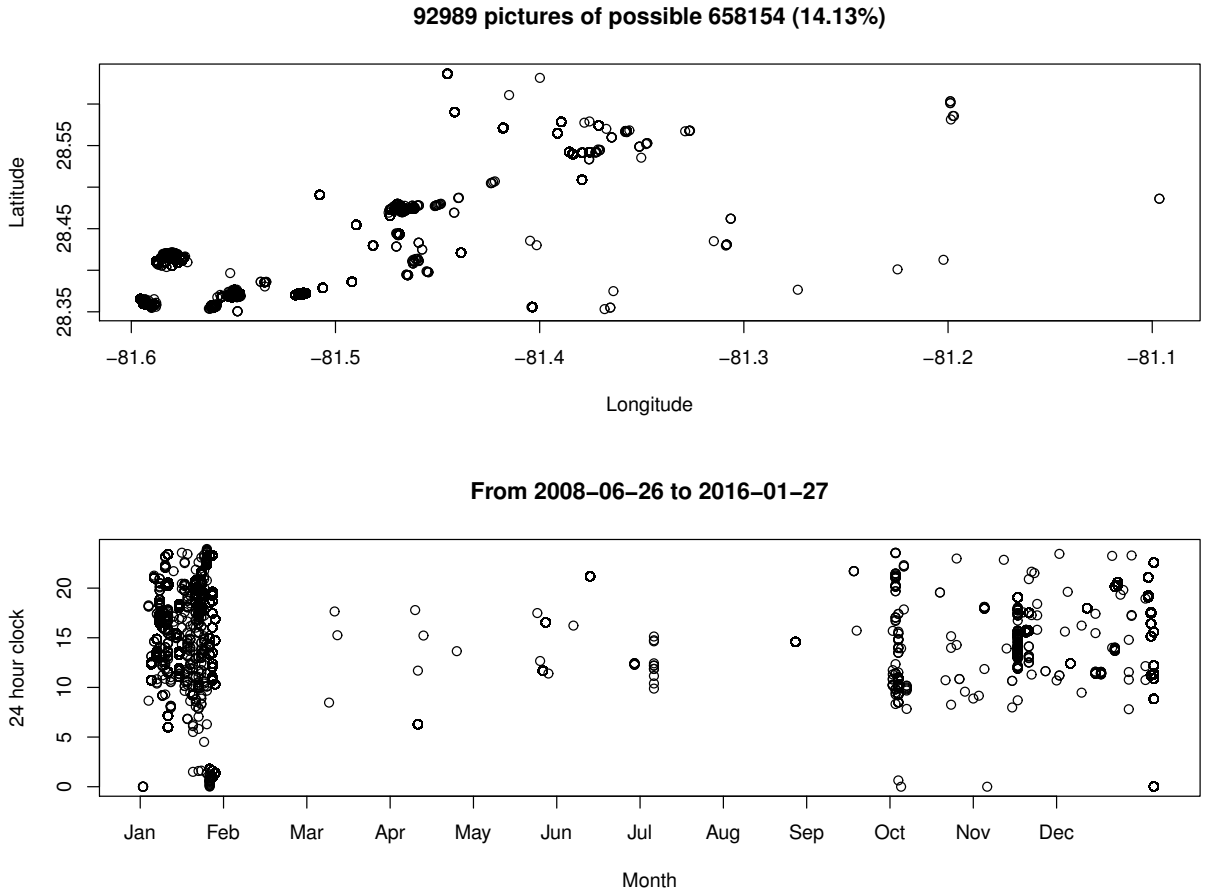


Figure 26: Simple plot Orlando, FL The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 2: Orlando, FL places of interest. The most frequently identified locations (414 out of 4,138) as per the Google API (378 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Bay Lake	136
2	Florida Center	53
3	Southwest Orlando	48
4	Americas	28
5	Lake Buena Vista	15
6	Northeast Orlando	13
7	Florida Center North	8
8	South Eola	8
9	Lake Eola Heights	7
10	Orlando	7
11	Southeast Orlando	7
12	Crystal Arts	6
13	Sir Mickey's	6
14	Tinker Bells Treasures	6
15	Main Street Electrical Parade	6
16	Sweet Spells	6
17	Casey's Corner	6
18	Seven Dwarfs' Mine	6
19	Island Supply Company	6
20	Liberty Square Portrait Gallery	6
21	Main Steet Bakery	6
22	Main Street Cinema	6
23	The Chapeau	6
24	Swiss Family Treehouse	6
25	The Crystal Palace	6

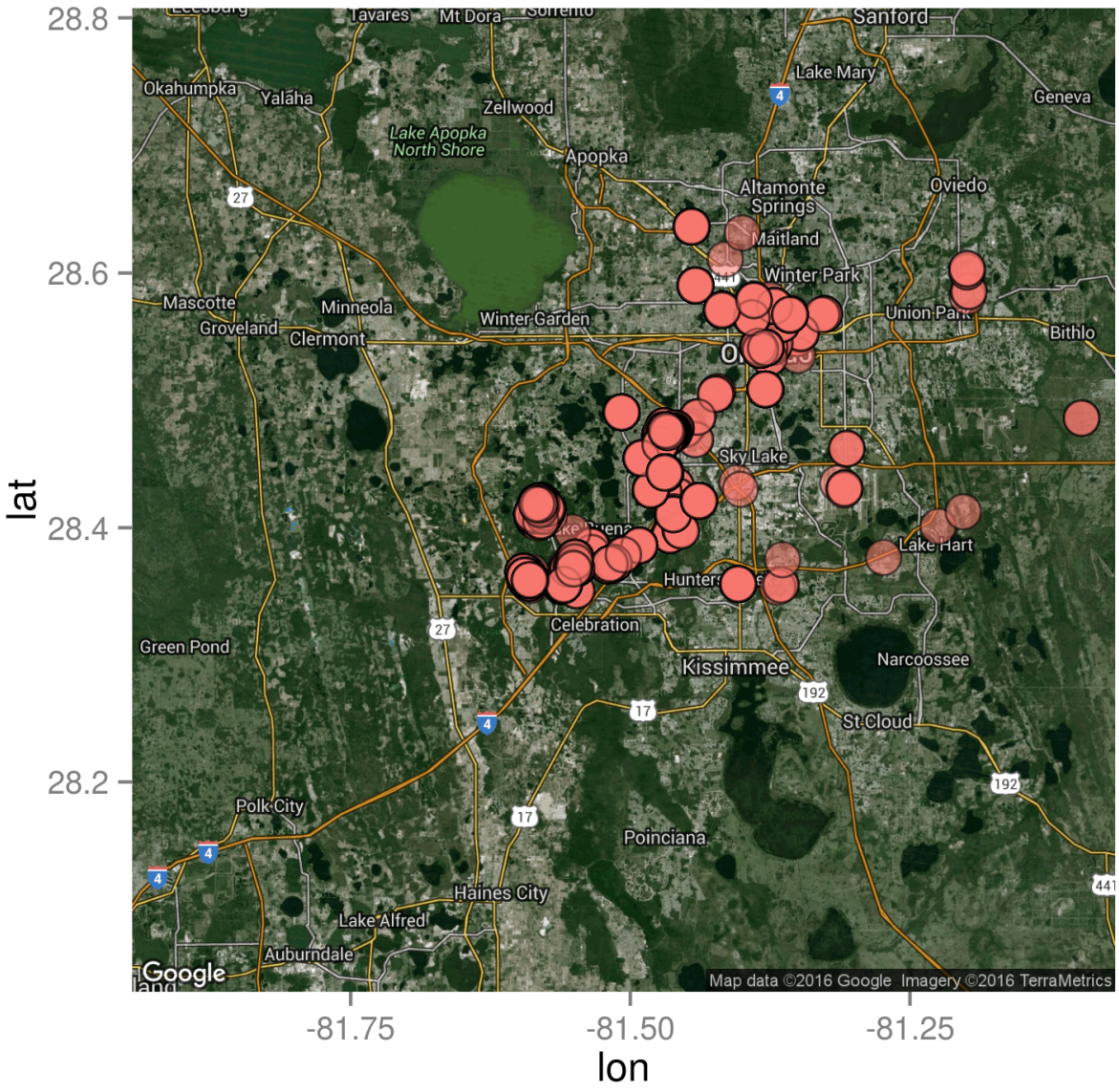


Figure 27: ggplot plot Orlando, FL The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

C Honolulu, HI

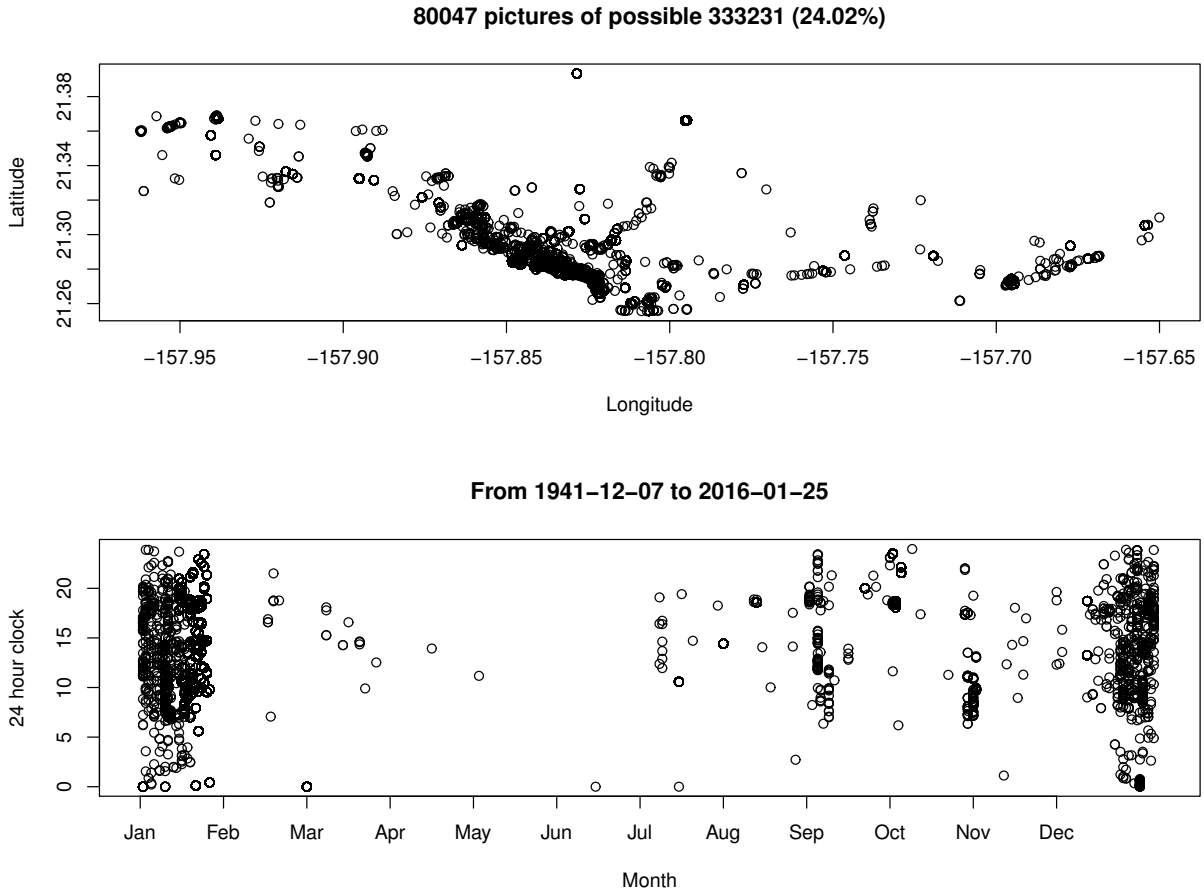


Figure 28: Simple plot Honolulu, HI The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 3: Honolulu, HI places of interest. The most frequently identified locations (1,105 out of 14,699) as per the Google API (1,013 API calls, 12 failed, success rate 98.8 per cent).

Ranking	**Name**	**Count**
1	Honolulu	409
2	Waikk	128
3	Downtown	78
4	Ala Moana	71
5	Diamond Head - Kapahulu - St. Louis	58
6	Hawaii Kai	48
7	Kaka'ako	48
8	McCully - Moiliili	39
9	Mnoa	35
10	Kalihi - Palama	30
11	Kuliouou - Kalani Iki	26
12	Makiki - Lower Punchbowl - Tantalus	19
13	Naval Station Pearl Harbor	16
14	Joint Base Pearl Harbor-Hickam	15
15	Kaimuki	12
16	Moanalua	11
17	Waialae - Kahala	10
18	Liliha - Kapalama	9
19	Nuuanu - Punchbowl	7
20	Brandy Melville	6
21	Sheraton Travel Services	6
22	Dakine Cellular	6
23	Maui Divers Jewelry	6
24	Shaka Tours	6
25	H-Zone	6

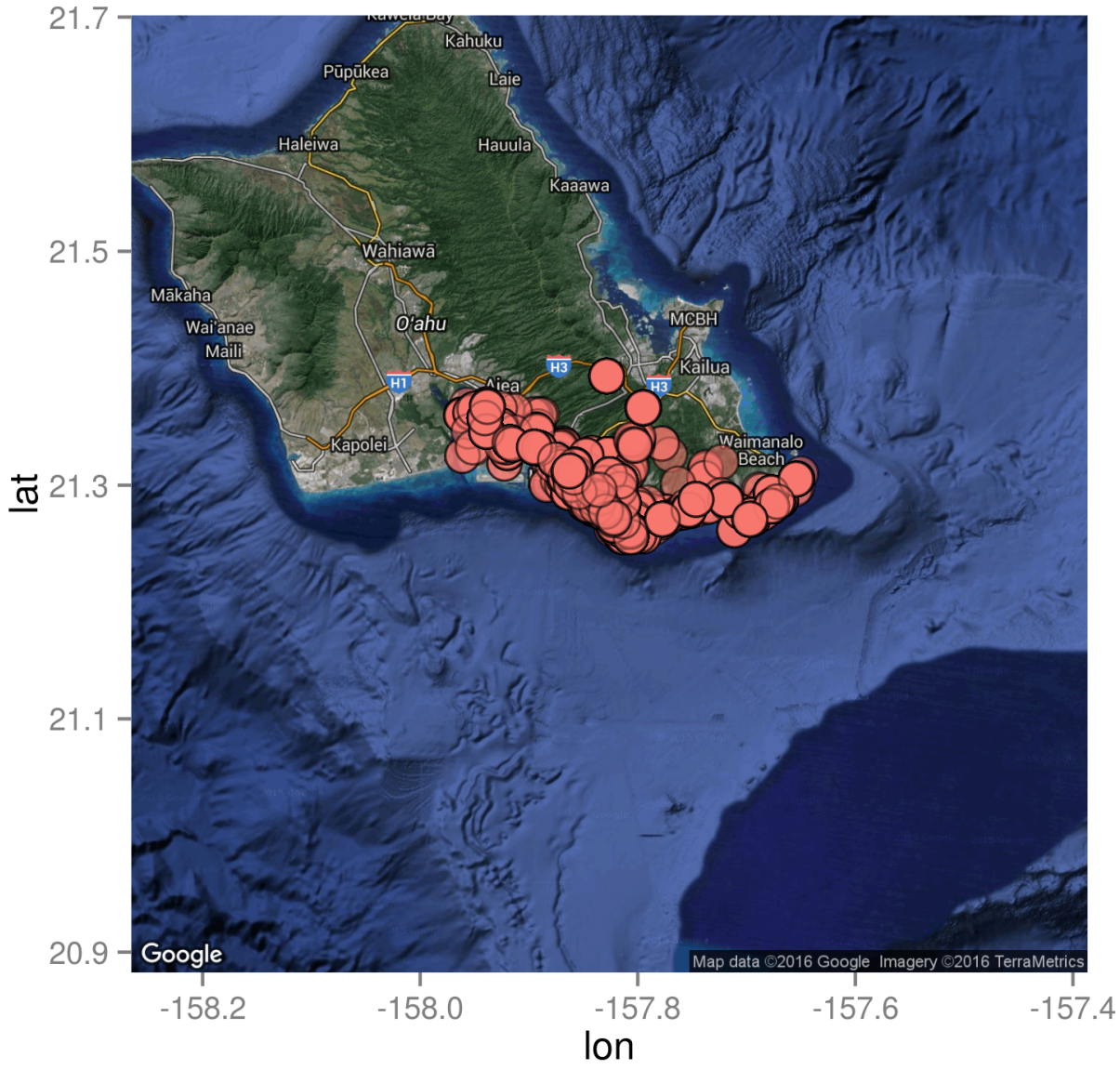


Figure 29: ggplot plot Honolulu, HI The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

D Reno, NV

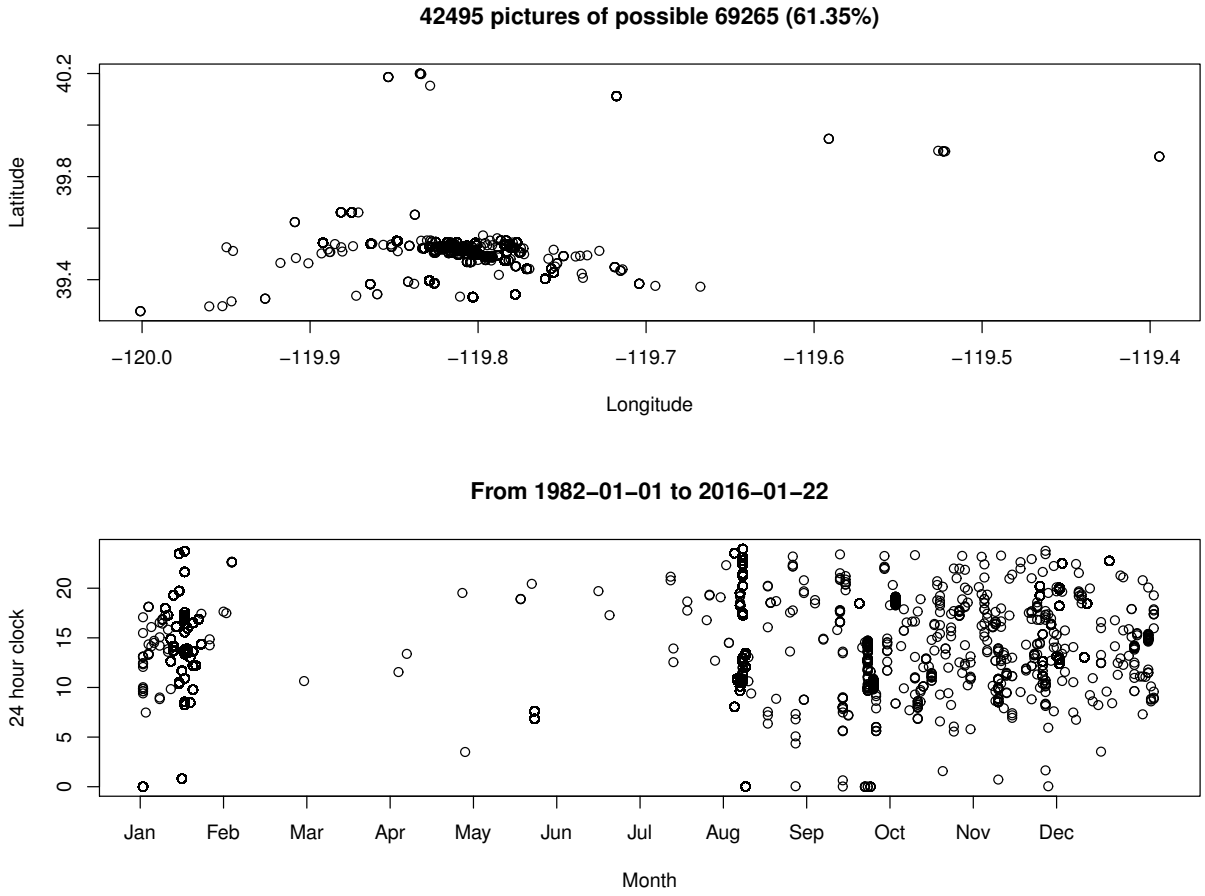


Figure 30: Simple plot Reno, NV The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 4: Reno, NV places of interest. The most frequently identified locations (409 out of 3,483) as per the Google API (338 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Reno	198
2	Americas	42
3	West University	28
4	Northeast Reno	19
5	Downtown	15
6	Northwest Reno	9
7	Sparks	7
8	Mira Loma	6
9	In Town Motel	6
10	W 4th St/Arlington Ave	6
11	Belvedere Towers	6
12	Lido Inn	6
13	Banyan Bar	5
14	Romanza	5
15	Oceano at the Peppermill	5
16	White Orchid	5
17	Peppermill Resort Spa Casino	5
18	EDGE Nightspot at the Peppermill	5
19	Bimini Steakhouse at the Peppermill	5
20	Mardi Gras Motor Lodge	5
21	Spa Toscana at the Peppermill	5
22	G4S Technology	4
23	Jessie Beck Elementary School	4
24	Bob's Cleaners	4
25	Chapel Tavern	4

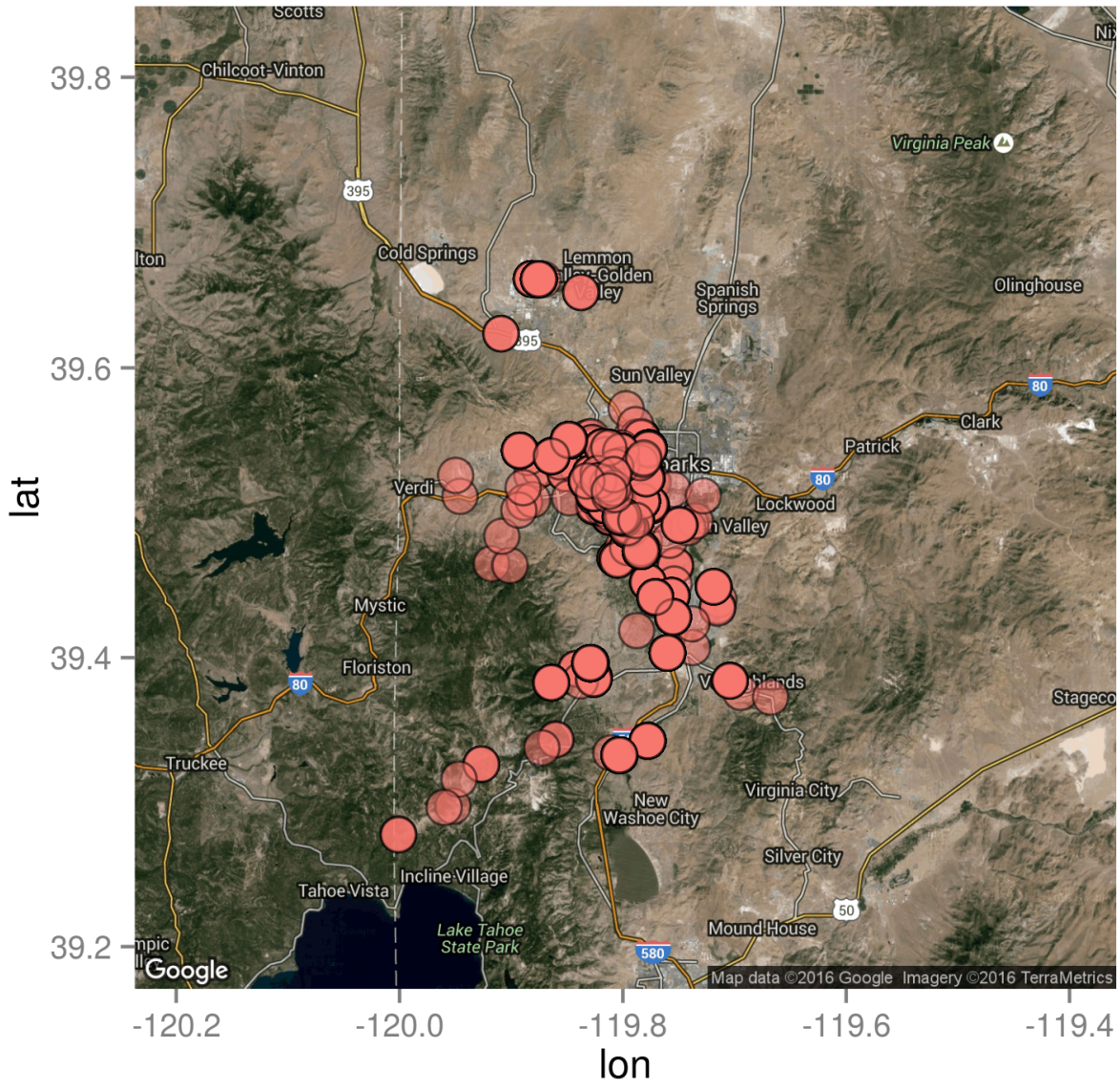


Figure 31: ggplot plot Reno, NV The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

E Tidewater, OR

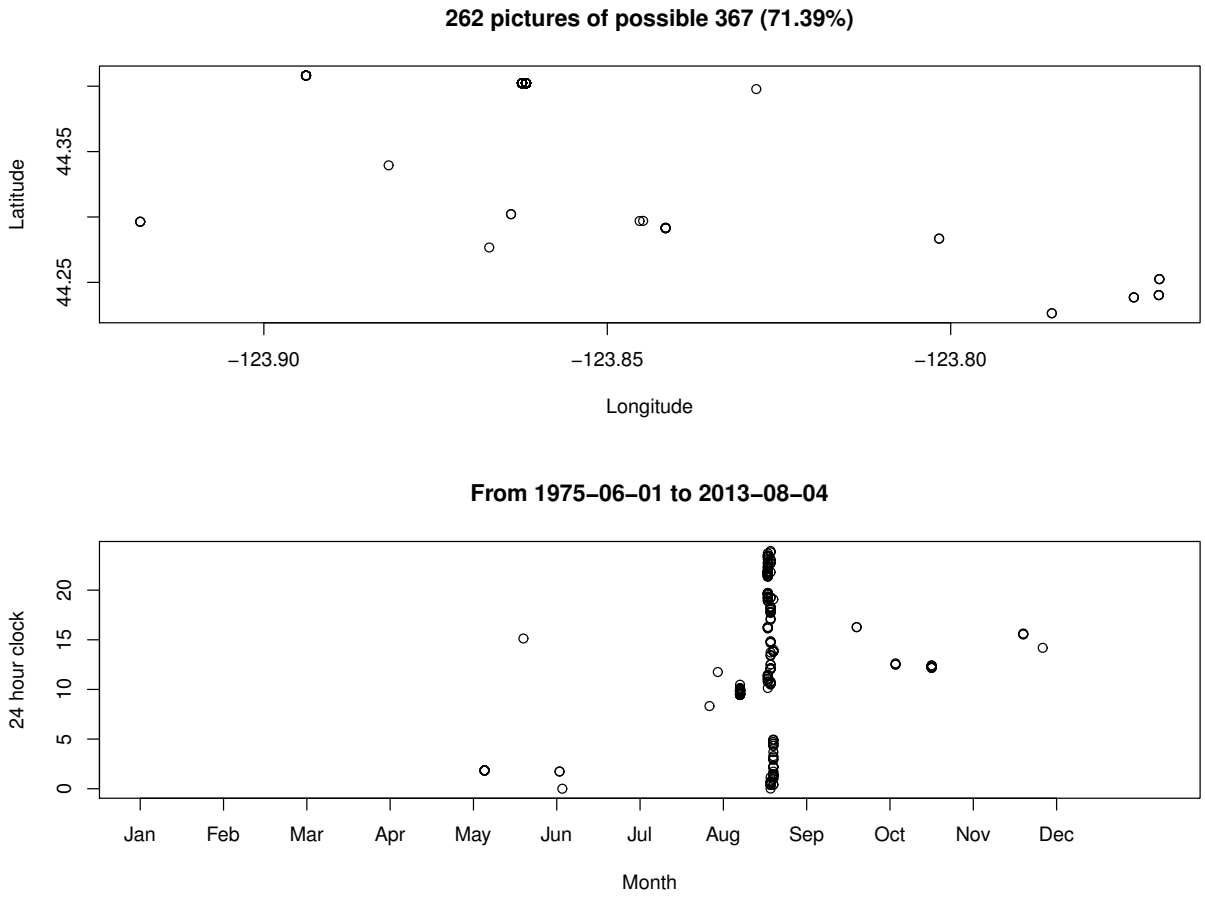


Figure 32: Simple plot Tidewater, OR The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 5: Tidewater, OR places of interest. The most frequently identified locations (17 out of 17) as per the Google API (16 API calls, 0 failed, success rate 100.0 per cent).

Ranking	Name	Count
1	Americas	16
2	Tradenet	1

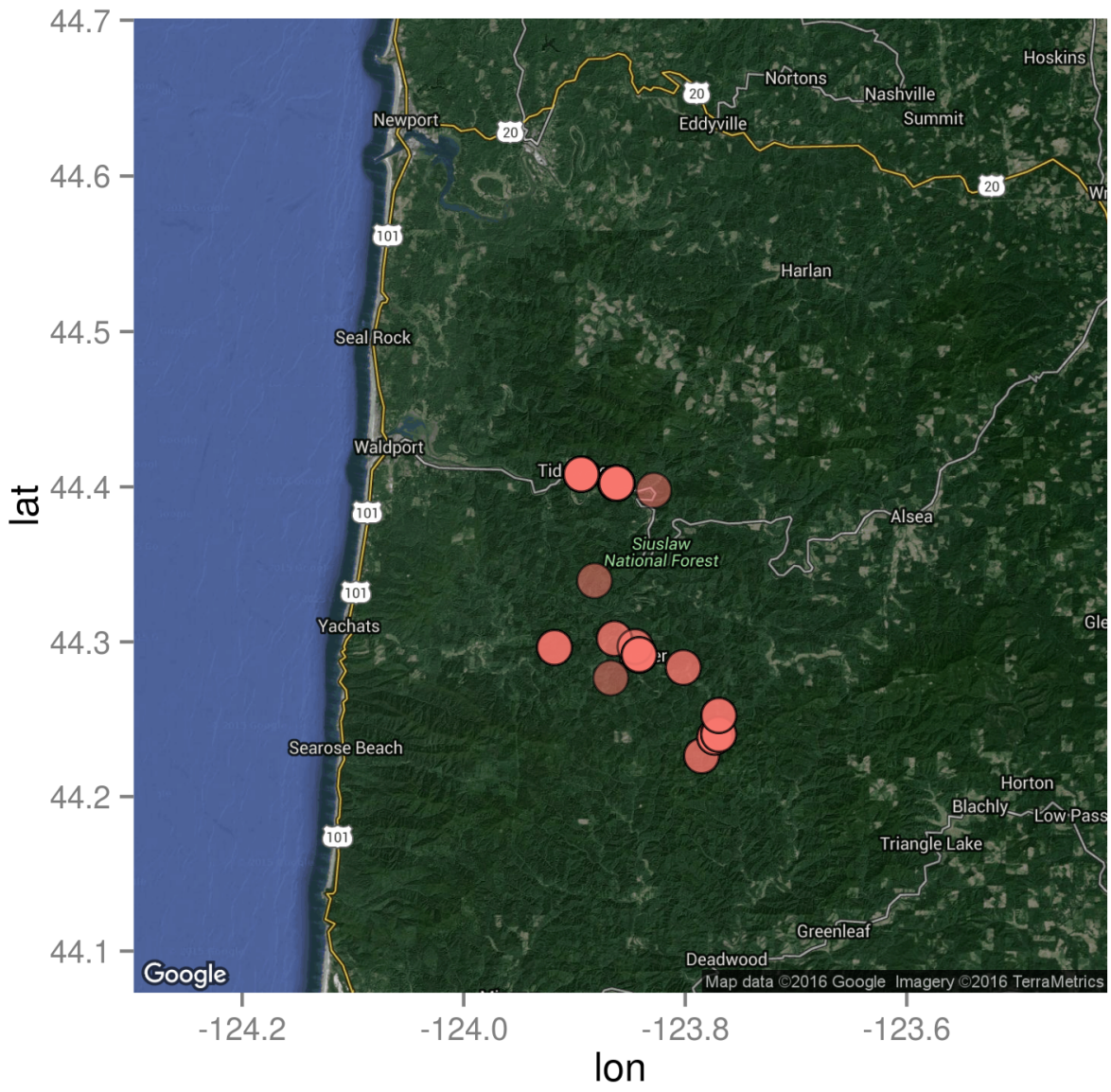


Figure 33: ggplot plot Tidewater, OR The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

F San Antonio, TX

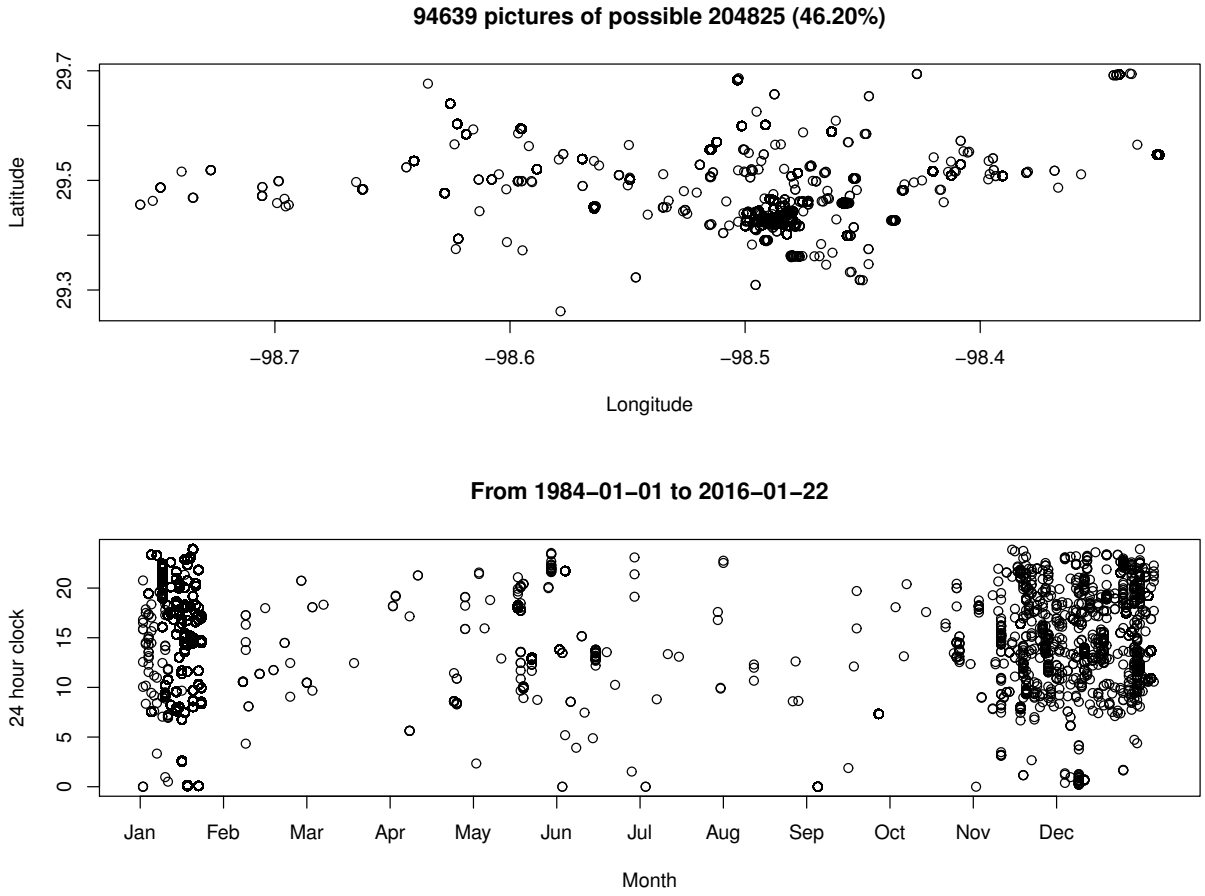


Figure 34: Simple plot San Antonio, TX The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 6: San Antonio, TX places of interest. The most frequently identified locations (543 out of 7,490) as per the Google API (579 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	San Antonio	85
2	Downtown	80
3	Midtown	33
4	Uptown	28
5	Northwest Side	26
6	Alamo Heights	26
7	Alamo Plaza	23
8	North Central	21
9	Tobin Hill	21
10	Hemisfair	20
11	Far West Side	19
12	East Pyron	15
13	Near East Side	15
14	Southside	15
15	Harlandale	14
16	Mahncke Park	13
17	La Villita	13
18	Main/Military Plaza	12
19	Inner West Side	12
20	Northeast Side	11
21	Americas	10
22	East Side	8
23	Far North Central	8
24	Shearer Hills / Ridgeview	8
25	Riverside	7

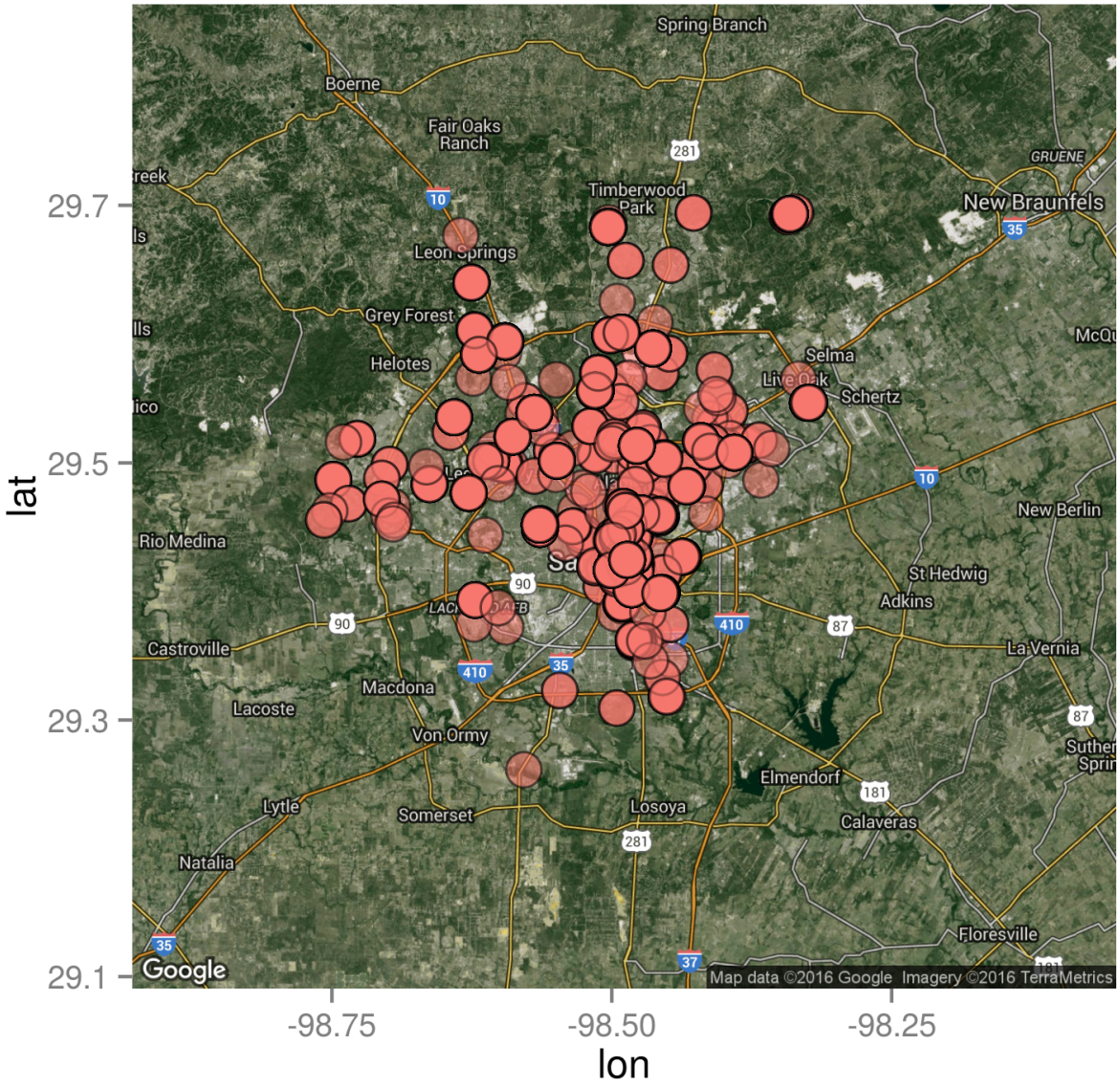


Figure 35: ggplot plot San Antonio, TX The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

G Chesapeake, VA

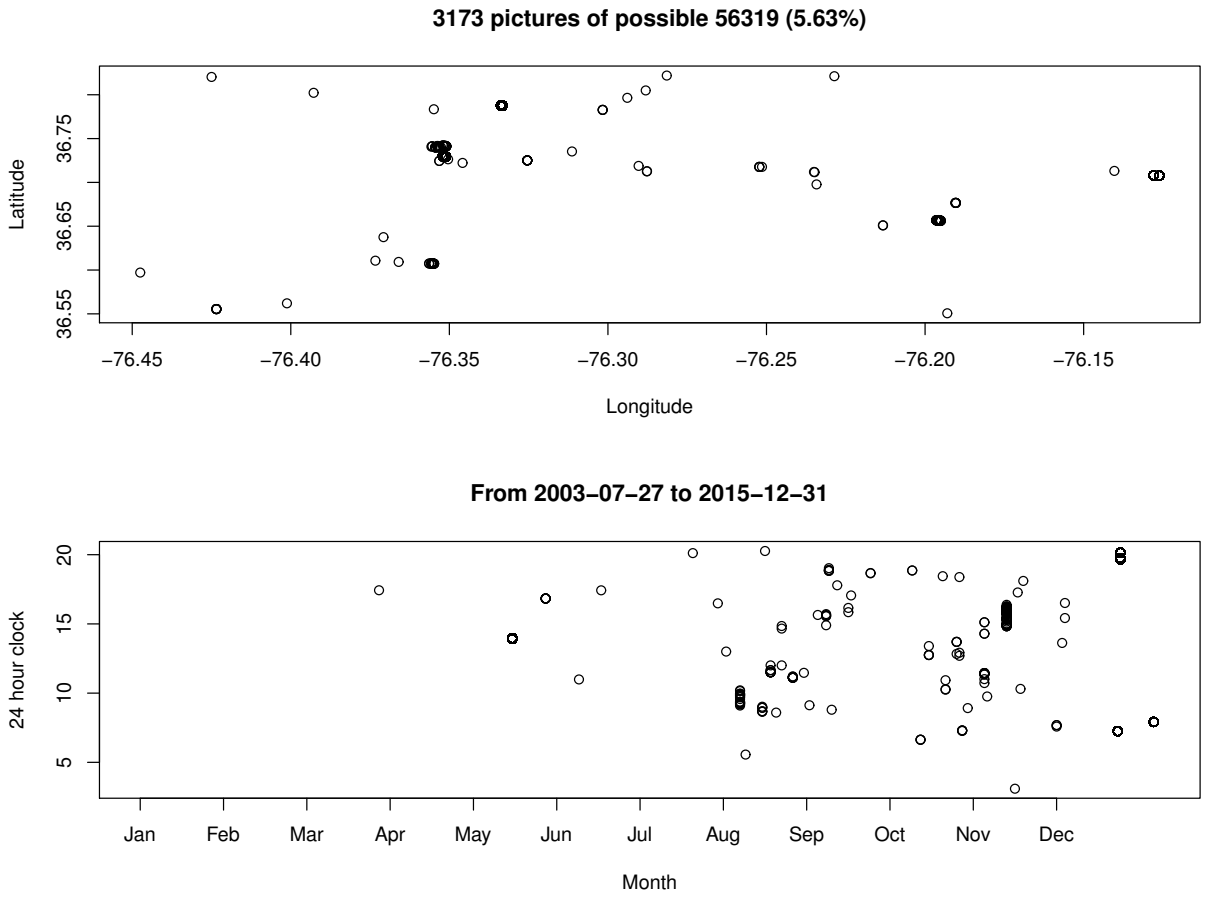


Figure 36: Simple plot Chesapeake, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 7: Chesapeake, VA places of interest. The most frequently identified locations (122 out of 195) as per the Google API (57 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Chesapeake	49
2	Pleasant Grove West	6
3	Great Bridge	5
4	Butts Station	5
5	Deep Creek North	5
6	Culpepper Landing	5
7	Pleasant Grove East	4
8	White's Old Mill Garden Center	4
9	Great Bridge East	4
10	Deep Creek South	4
11	White's Nursery and Greenhouses	3
12	South Norfolk	3
13	Brentwood Veterinary Clinic	3
14	J T Fisher Funeral Services	3
15	George Washington and Delwood	2
16	Chesapeake City Magistrate	2
17	Payment Alliance International, Inc.	2
18	Western Branch South	2
19	Chesapeake City Employee CU	2
20	Cottage Trails at Culpepper Landing	2
21	Shamrock Gardens Apartments	2
22	BayPort Credit Union	2
23	Laurel Avenue Church of Christ	1
24	Moyock	1
25	Great Bridge Congregational Church	1

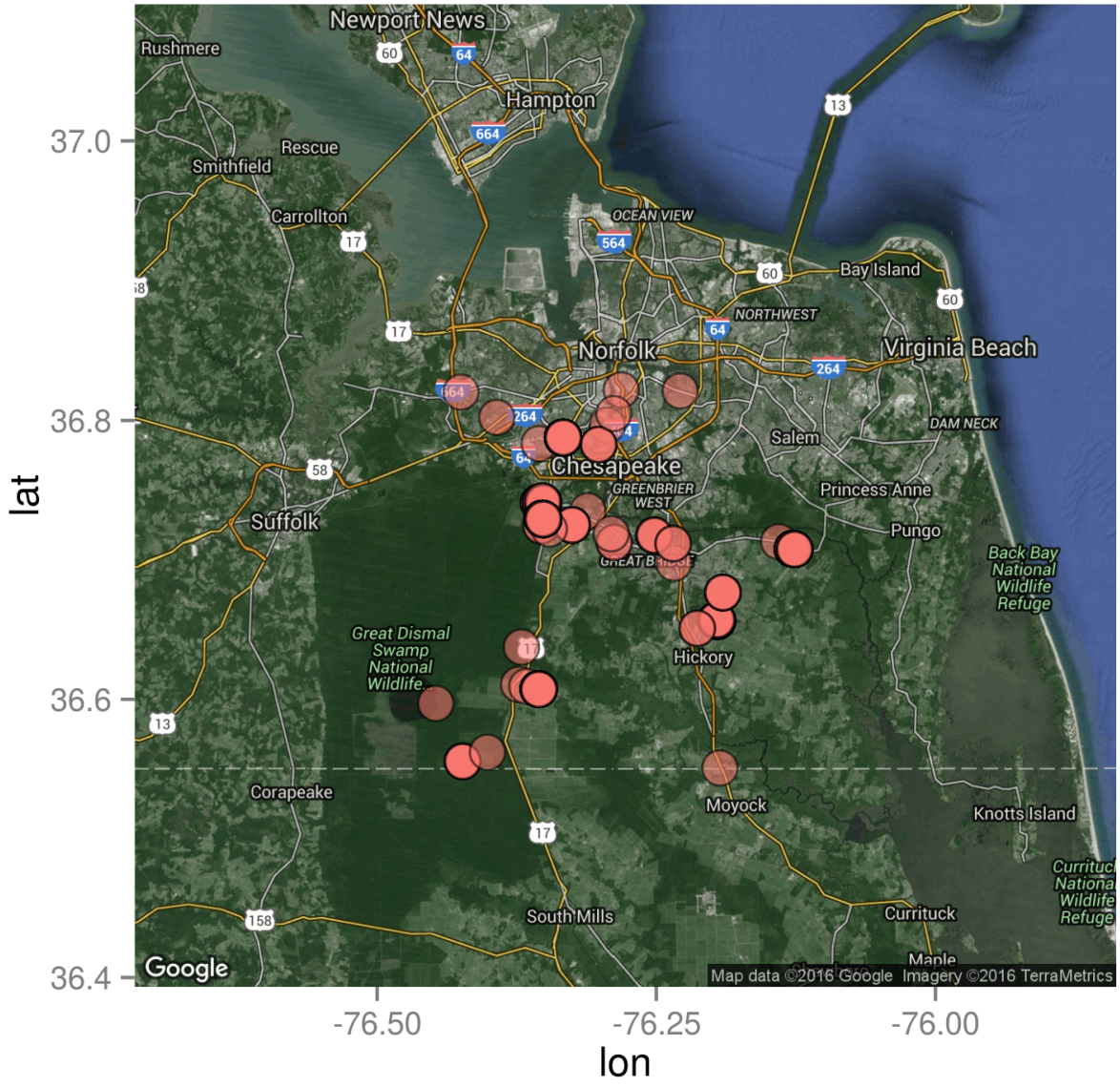


Figure 37: ggplot plot Chesapeake, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

H Hampton, VA

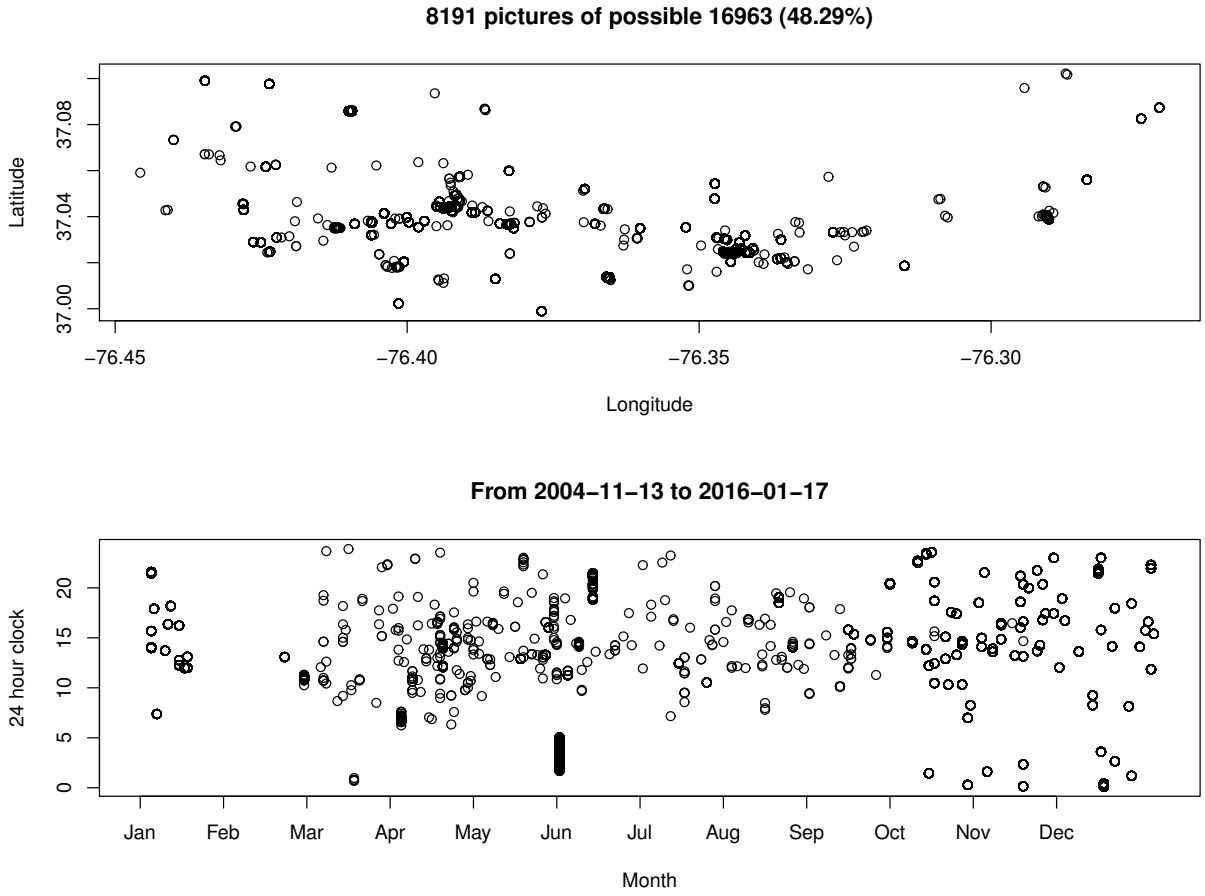


Figure 38: Simple plot Hampton, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 8: Hampton, VA places of interest. The most frequently identified locations (425 out of 2,502) as per the Google API (249 API calls, 0 failed, success rate 100.0 per cent).

Ranking	Name	Count
1	Hampton	161
2	Coliseum Central	47
3	Briarfield	35
4	Phoebus	25
5	Downtown Hampton	22
6	Downtown	18
7	Buckroe Beach	15
8	Hampton Roads Center	14
9	Farmington	11
10	Wythe	11
11	Victoria Boulevard Historic District	6
12	Pasture Point	6
13	Northampton	6
14	Mcw Lawyers: Walter Douglas	4
15	Willow Beach Real Estate, LLC	4
16	ATM 7ELEVEN, INC.	4
17	Young Nails	4
18	Chico's	4
19	Settlers Landing Parking Garage	4
20	Beauty for Ashes Contemporary School of Dance	4
21	Cricket Wireless	4
22	Clarks	4
23	Conch and Bucket	4
24	Citibank ATM	4
25	Dan Winters-Neil Pryde Sails	4

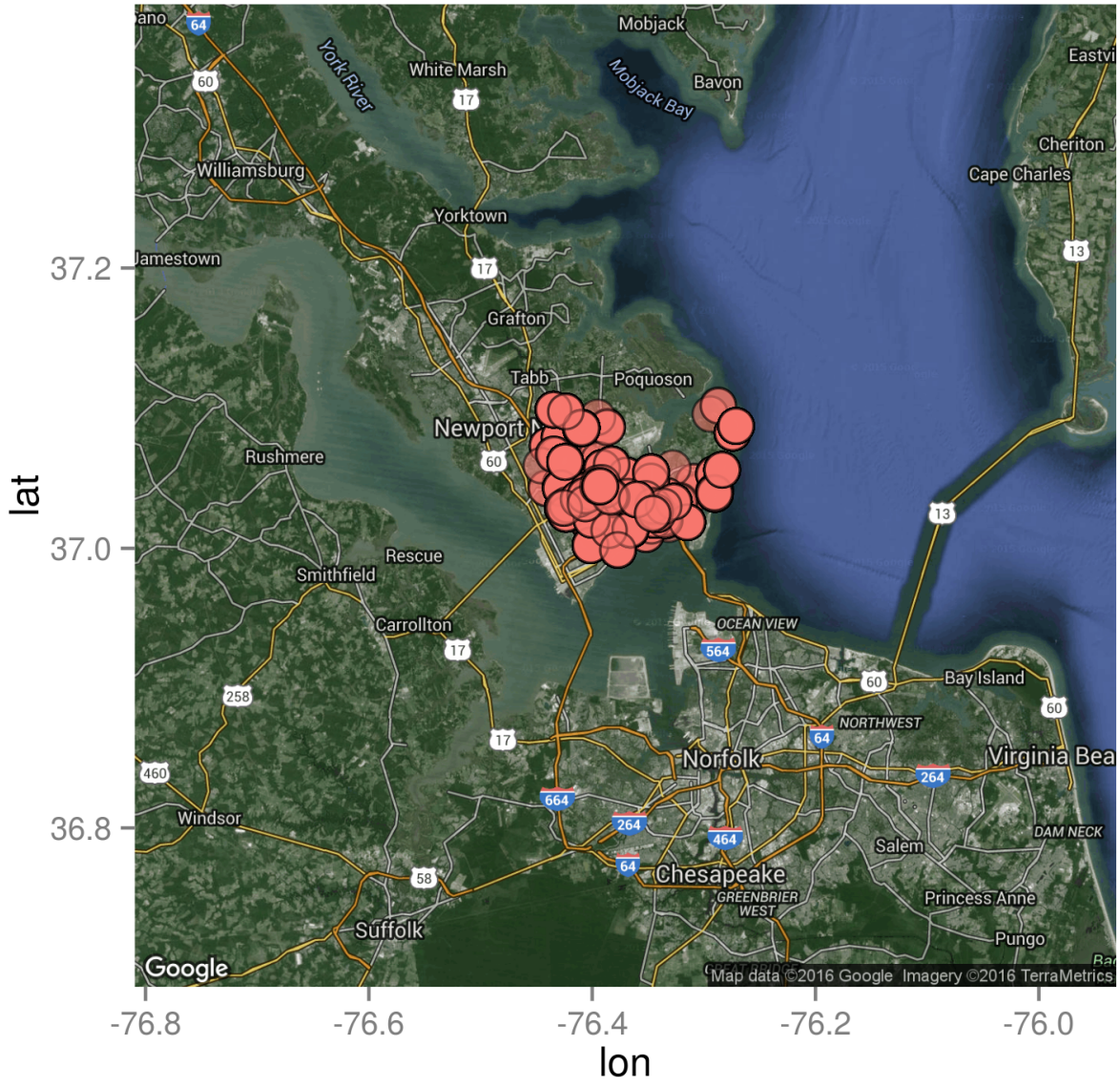


Figure 39: ggplot plot Hampton, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

I Norfolk, VA

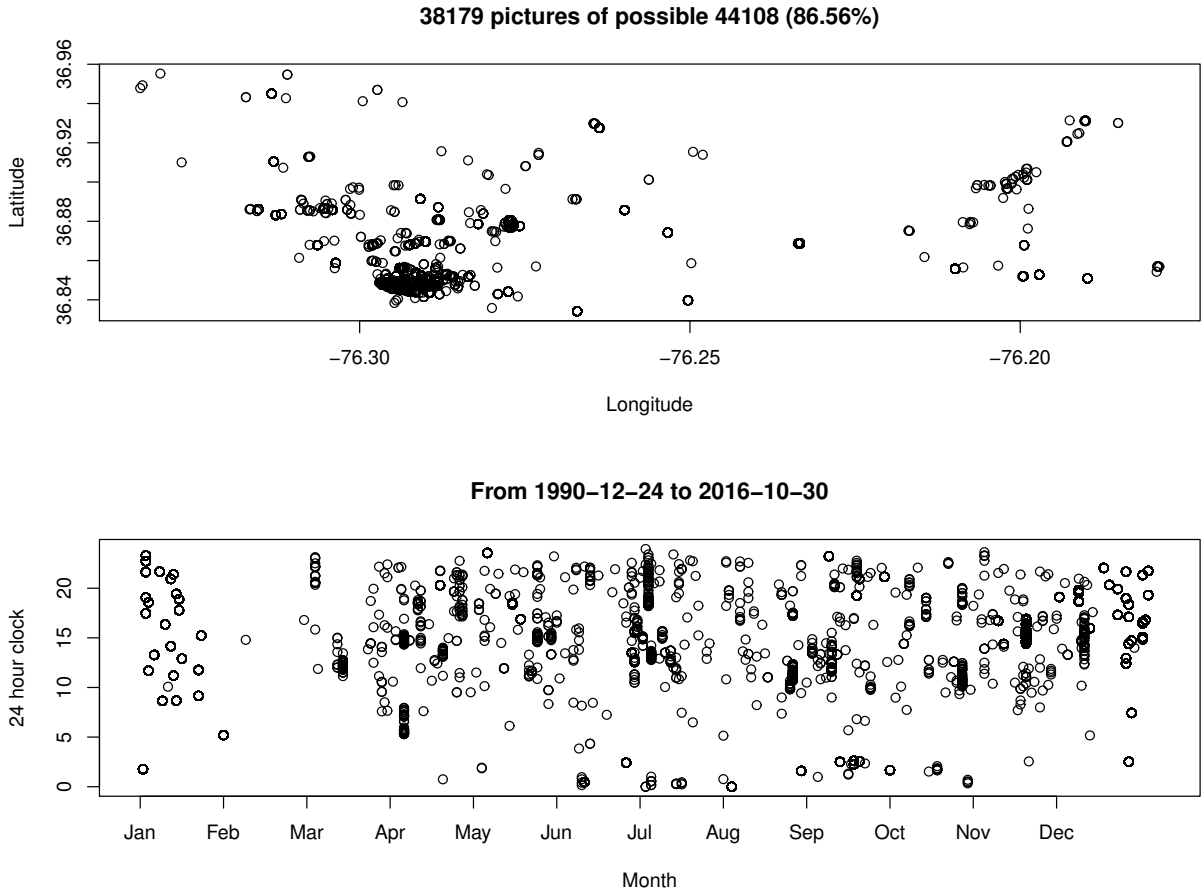


Figure 40: Simple plot Norfolk, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 9: Norfolk, VA places of interest. The most frequently identified locations (453 out of 5,531) as per the Google API (414 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Norfolk	184
2	Downtown	57
3	Ghent	31
4	West Freemason	29
5	Lamberts Point	16
6	Neon District	12
7	Larchmont/Edgewater	12
8	Highland Park	8
9	Park Place	8
10	Bryant C Mc Gann Attorney	6
11	Document Technologies Inc	6
12	The Vineyards Norfolk	6
13	Details On Granby	6
14	Pagoda and Oriental Garden	6
15	Kaufman and Canoles: Land Charles E	6
16	Moffatt and Nichol	6
17	Glen McClure Photography	6
18	Wilbanks Smith and Thomas Asset Management	6
19	CB Richard Ellis	6
20	US Attorney	6
21	Arlene F. Klinedinst	6
22	Littman's, Inc.	6
23	Famous Uncle Al's Hot Dogs	6
24	Wall Einhorn and Chernitzer, P.C. (WEC)	6
25	McPhillips, Roberts and Deans, PLC	6

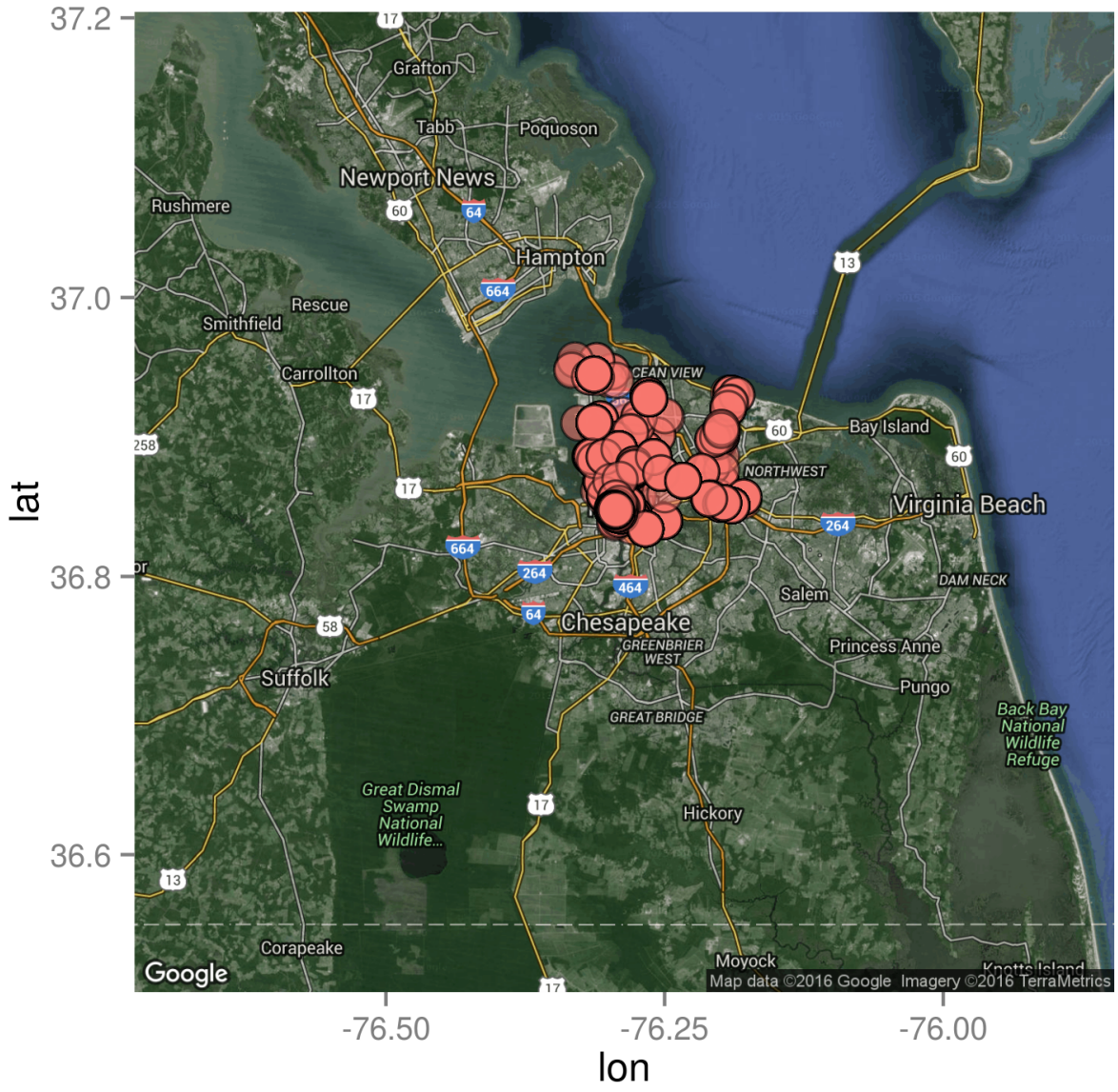


Figure 41: ggplot plot Norfolk, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

J Portsmouth, VA

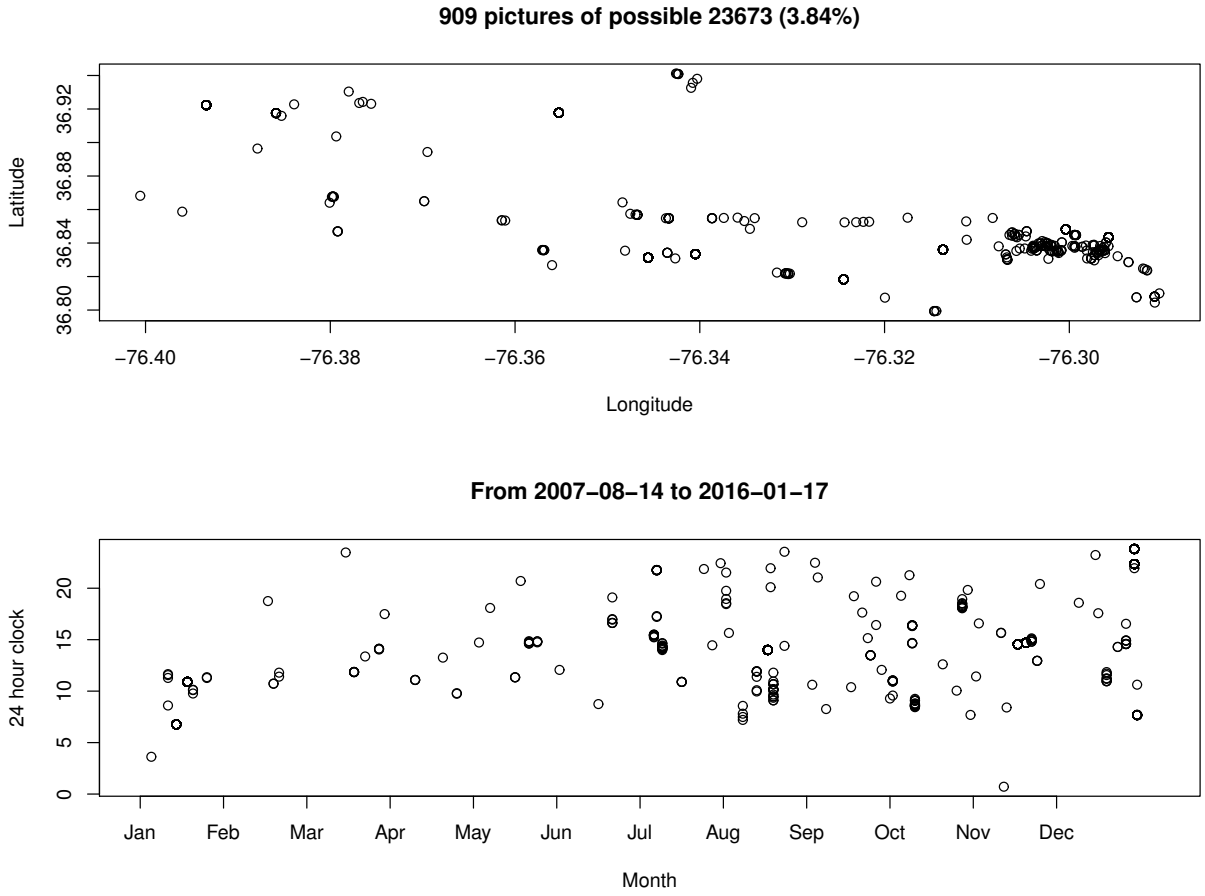


Figure 42: Simple plot Portsmouth, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 10: Portsmouth, VA places of interest. The most frequently identified locations (259 out of 1,468) as per the Google API (155 API calls, 0 failed, success rate 100.0 per cent).

Ranking	**Name**	**Count**
1	Portsmouth	116
2	Olde Towne	42
3	Midtown	9
4	Port Norfolk	7
5	Marshall W Butt Library	5
6	Massie Law Group	4
7	Riverview Gallery LLC	4
8	Queen Bees LLC	4
9	Griffin Pappas and Scarborough	4
10	Cycle Classics	4
11	Bangel, Bangel, and Bangel	4
12	Ciccotti Enterprises	4
13	Cornerstone Systems	4
14	Portsmouth Art and Cultural Center	4
15	Holistic Farmacy	4
16	Center For Community Development	4
17	The Bier Garden	4
18	Portsmouth Schools Federal CU	4
19	Trinity Episcopal Church	4
20	ATM Express	4
21	Portsmouth Olde Town Historic District	4
22	Monumental Ministries	4
23	Naval Medical Center Portsmouth	4
24	Griffin Pappas and Scarborough: Scarborough Boyd	4
25	Professional Building	4

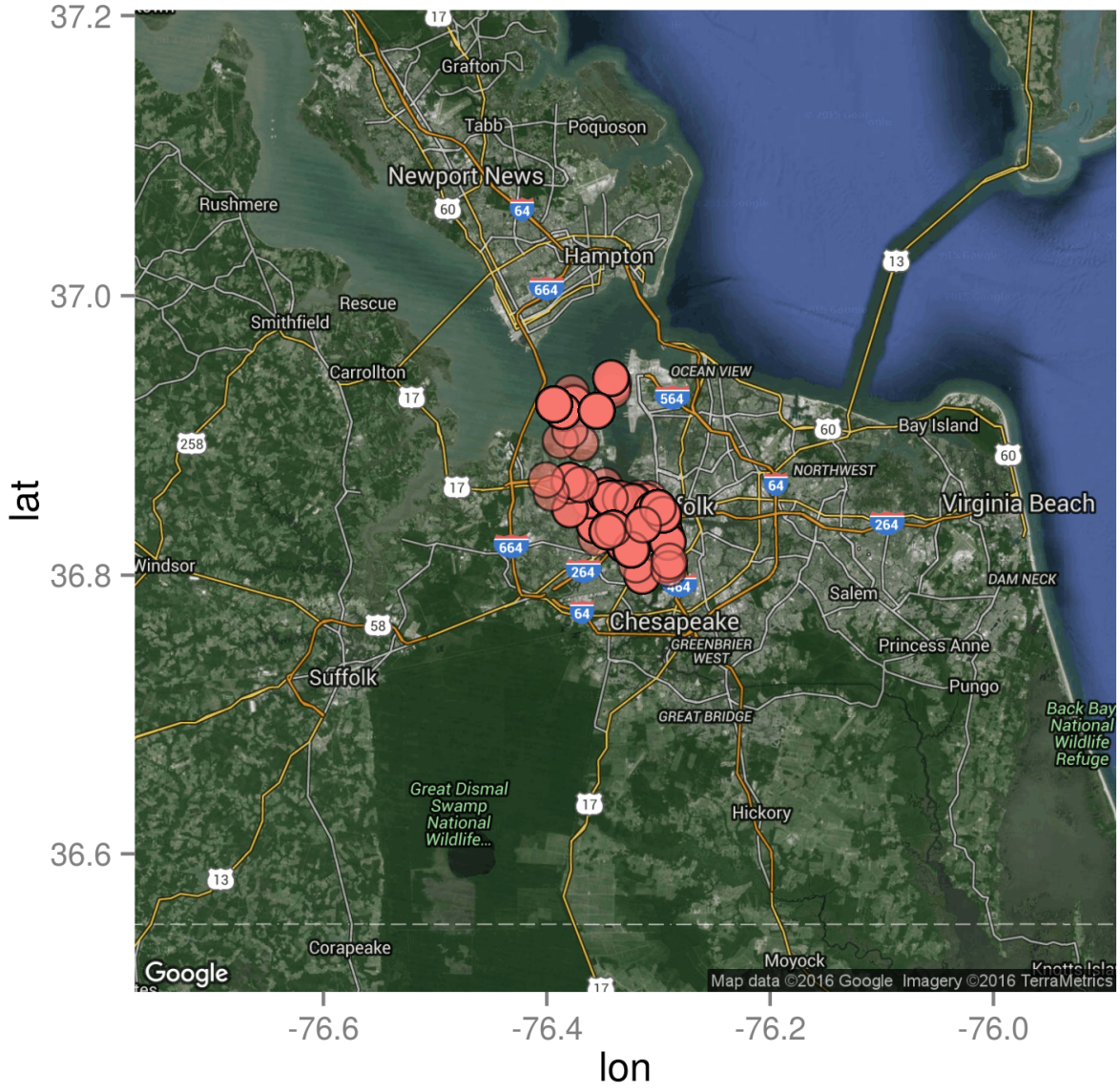


Figure 43: ggplot plot Portsmouth, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

K Virginia Beach, VA

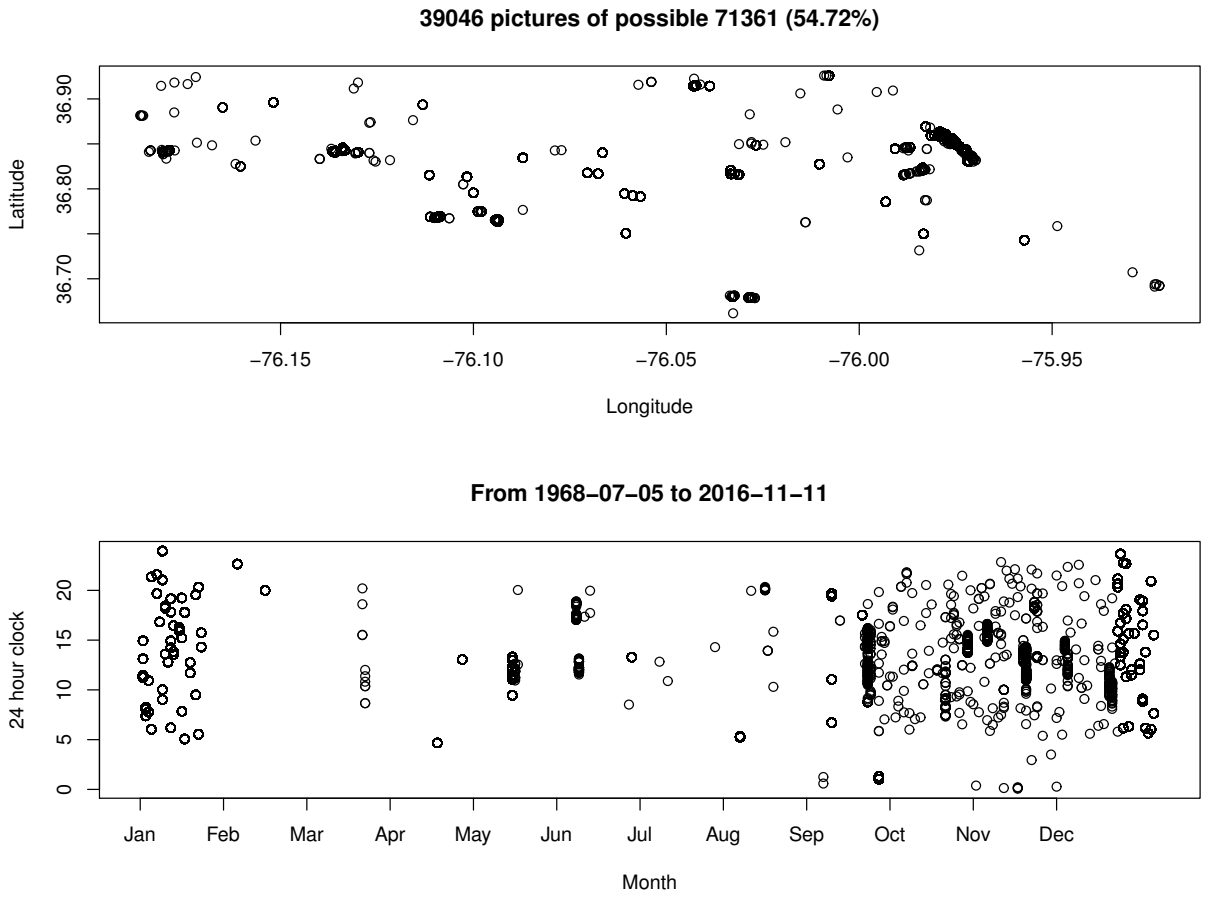


Figure 44: Simple plot Virginia Beach, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (position, time, and date) was parsed and plotted. Dates were converted to Julian day numbers, and time was converted to a 24 hour representation.

Table 11: Virginia Beach, VA places of interest. The most frequently identified locations (402 out of 2,839) as per the Google API (265 API calls, 0 failed, success rate 100.0 per cent).

Ranking	Name	Count
1	Virginia Beach	181
2	Northeast Virginia Beach	77
3	Northwest	31
4	Dam Neck	17
5	Oceana Naval Air Station	9
6	North Central	6
7	Sweet Frog	5
8	Sports Alley	5
9	Souvenir City	5
10	Sandbridge	5
11	Sunshine Paradise	5
12	24th Street Park	4
13	Blue Marlin Inn and Suites	4
14	Courtyard Virginia Beach Oceanfront/South	4
15	Hampton Inn VIRGINIA BEACH OCEANFRONT NORTH	4
16	Rudee Loop Lot	4
17	Hilton Virginia Beach Oceanfront	4
18	Top Gun Miniature Indoor Golf	4
19	Virginia Beach Convention and Visitors Bureau	4
20	Atlantic and 29th	4
21	Naval Aviation Monument Park	4
22	Cheries Bike and Blade Rental	4
23	Ocean Eddie's Seafood Restaurant	4
24	CP Shuckers Cafe and Raw Bar	4
25	31Ocean	4

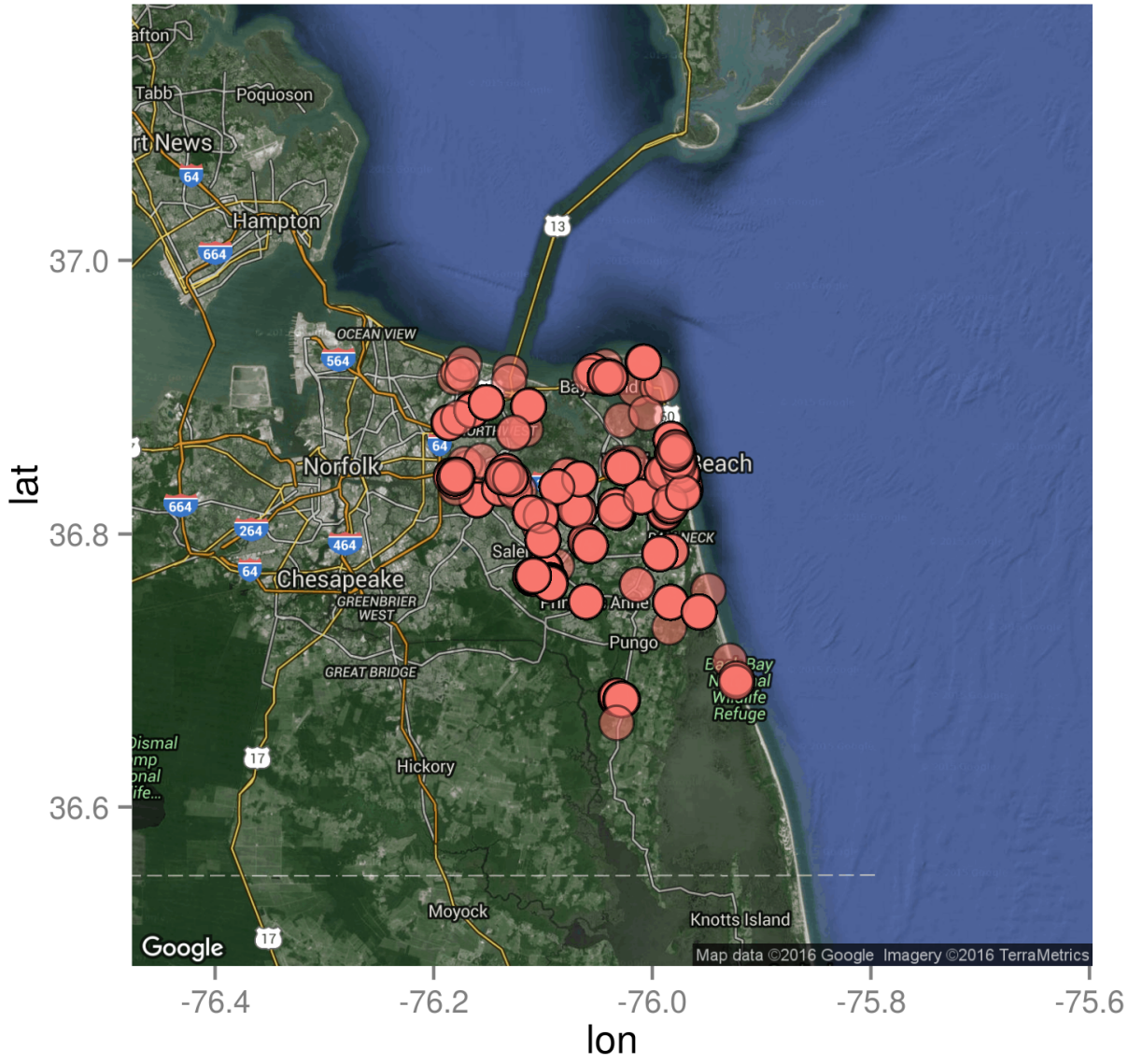


Figure 45: ggplot plot Virginia Beach, VA The plots show user submitted data associated with a picture uploaded to flickr. The user associated meta-data (latitude and longitude) was used to place the picture.

A place for notes for presentations

- http://www.youtube.com/v/SXDaNj_3nqw?hl=en&fs=1
- http://hint.fm/wind/
- http://www.wunderground.com/maps/?