Plot Spatial Data

Plot spatial data on an interactive map

Product: IBM® SPSS® Modeler

Extension type: Output/Visualization

Table of Contents

Description....4
Reauirements....4
Installation....4
R Packages used....4
User interface....5
1. General settings....5
2. Plot settings....8
3. Advanced Settings....11
Result example....12
Important links....18
Learn....18
Get products and technologies....18
Discuss....18

Description

This IBM SPSS Modeler output node allows you to plot data containing latitude and longitude on an interactive map. You can use the same color for all points or use a legend column to specify a color code. This legend may be categorical or continuous. Several color palettes are available (sequential, divergent, qualitative or monochrome) covering all possible use of the node.

More precisely, this node generates an HTML file which can be saved to a specific directory and/or opened in the default browser on execution. This html page is an interactive map, that is to say you can move, zoom in and out, etc.

Note that you can obtain longitude and latitude from an address using a Geocoding Node.

Requirements

- SPSS Modeler v16.0 or later
- SPSS Modeler 'R essentials' plugin
- R v2.15.x

Installation

Close SPSS Modeler. Save the *.cfe* file in the CDB folder, located by default on Windows in "*C:\ProgramData\IBM\SPSS\Modeler\16\CDB*" or under your Modeler 16 installation directory.

Restart SPSS Modeler: the node will now appear in the output panel.

R Packages used

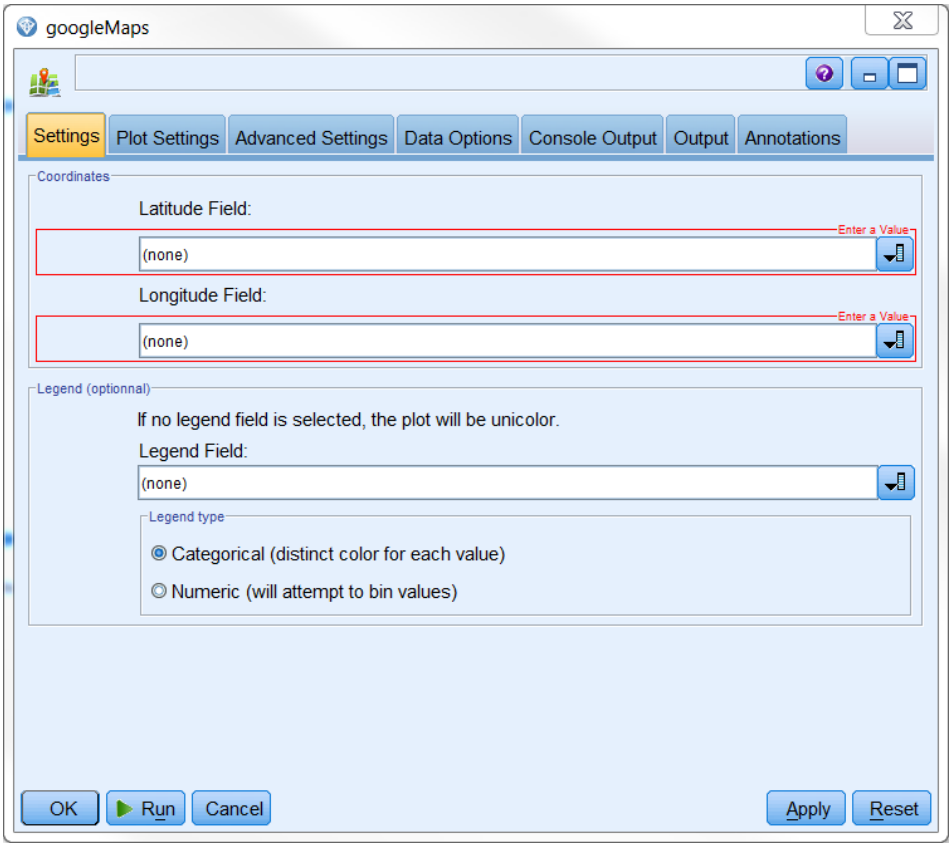
The R packages will be installed the first time the node is used as long as an Internet connection is available.

- plotGoogleMaps by Milan Kilibarda
- RColorBrewer by Erich Neuwirth

User interface

Double click on the node to open the user interface.

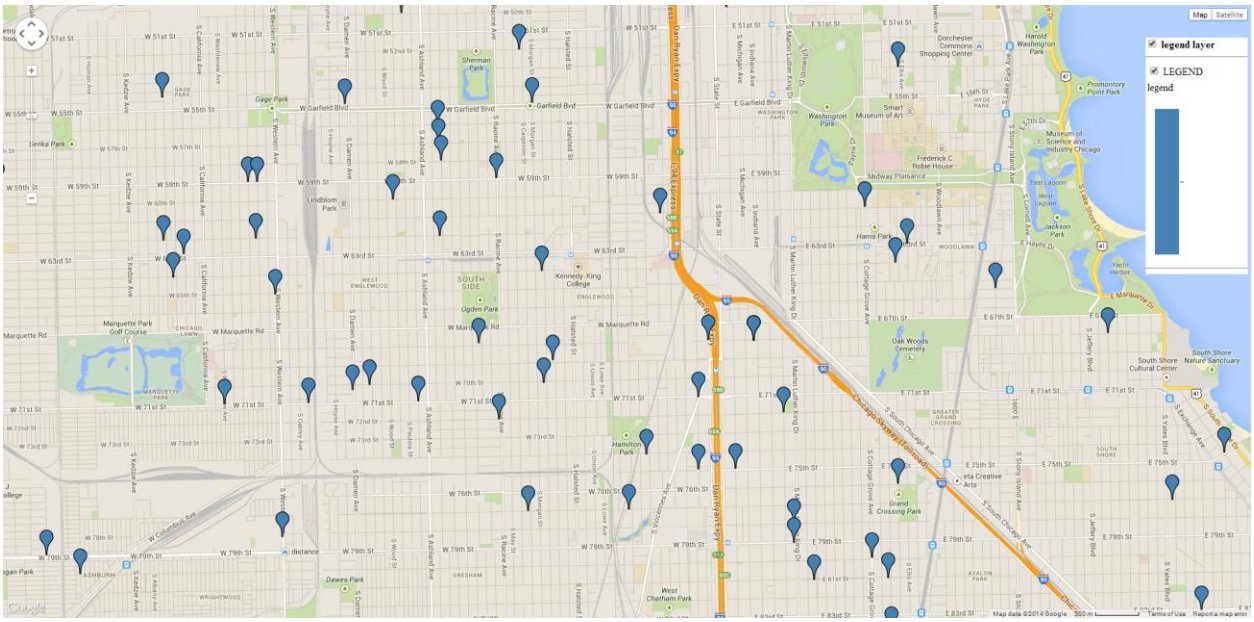
1. General settings



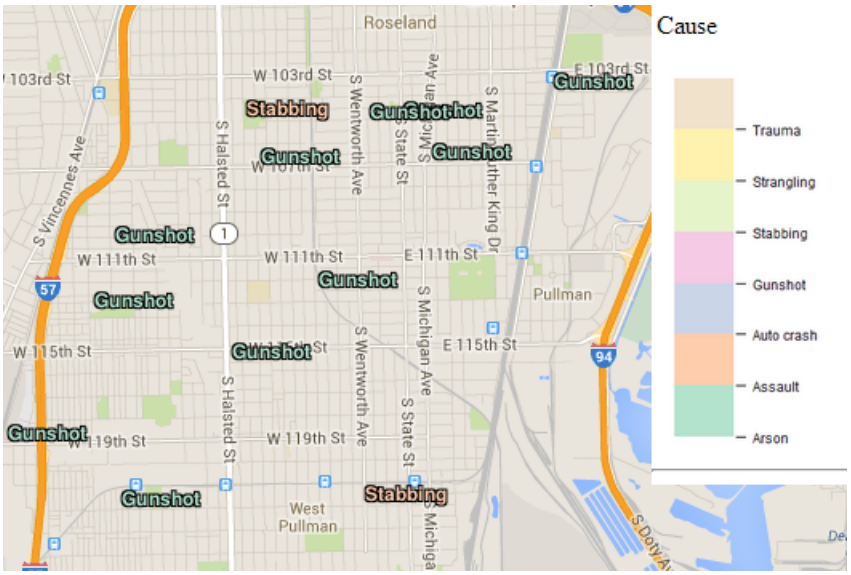
- Latitude and Longitude fields: specify which columns of your data must be used as Latitude and Longitude

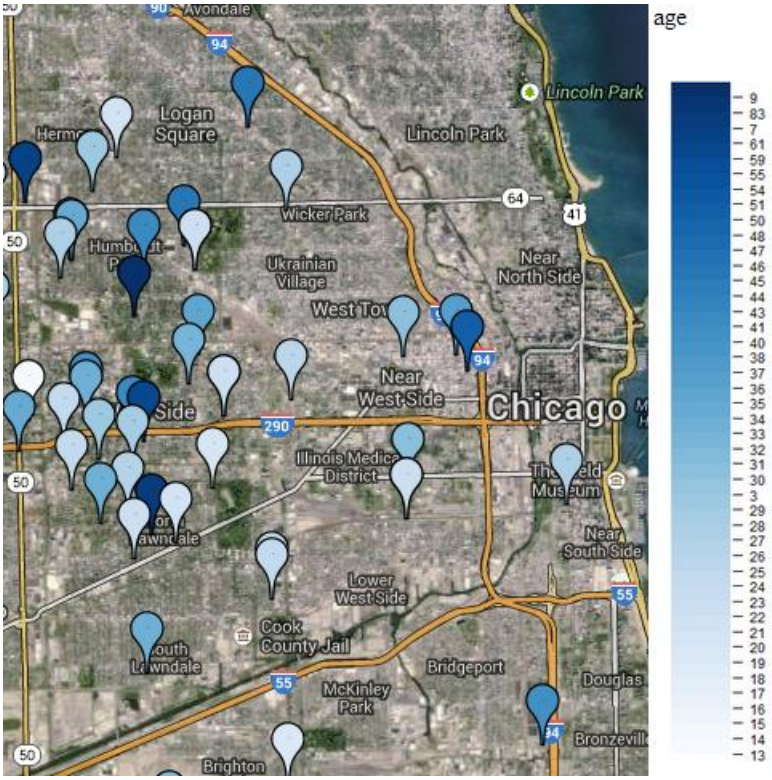
- Legend field: if you want to have a color-coded legend, specify which field must be used.

- If empty, the plot will be single color. The color depends on the chosen color palette (see *plot settings*). If the chosen color palette is monochrome, then it is used. If the chosen color palette is not monochrome, then the “mono-blue” color palette is used.
- Even if no legend field is selected, there will still be a one-color legend named “legend”.

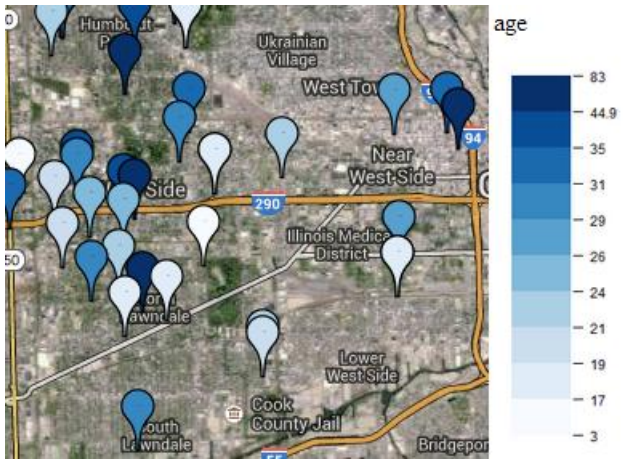


- Legend type: specify whether the legend is to be considered categorical or numeric
    - If no legend field is selected, the type doesn't matter
    - For a categorical legend, the values are treated like categories, each category having its own color. This is best used with a qualitative color palette. Note that you can select "categorical legend" even if the data is technically numeric, and every value will have its own color. Be aware however that the categories will be displayed in alphabetical order and not numerical order.

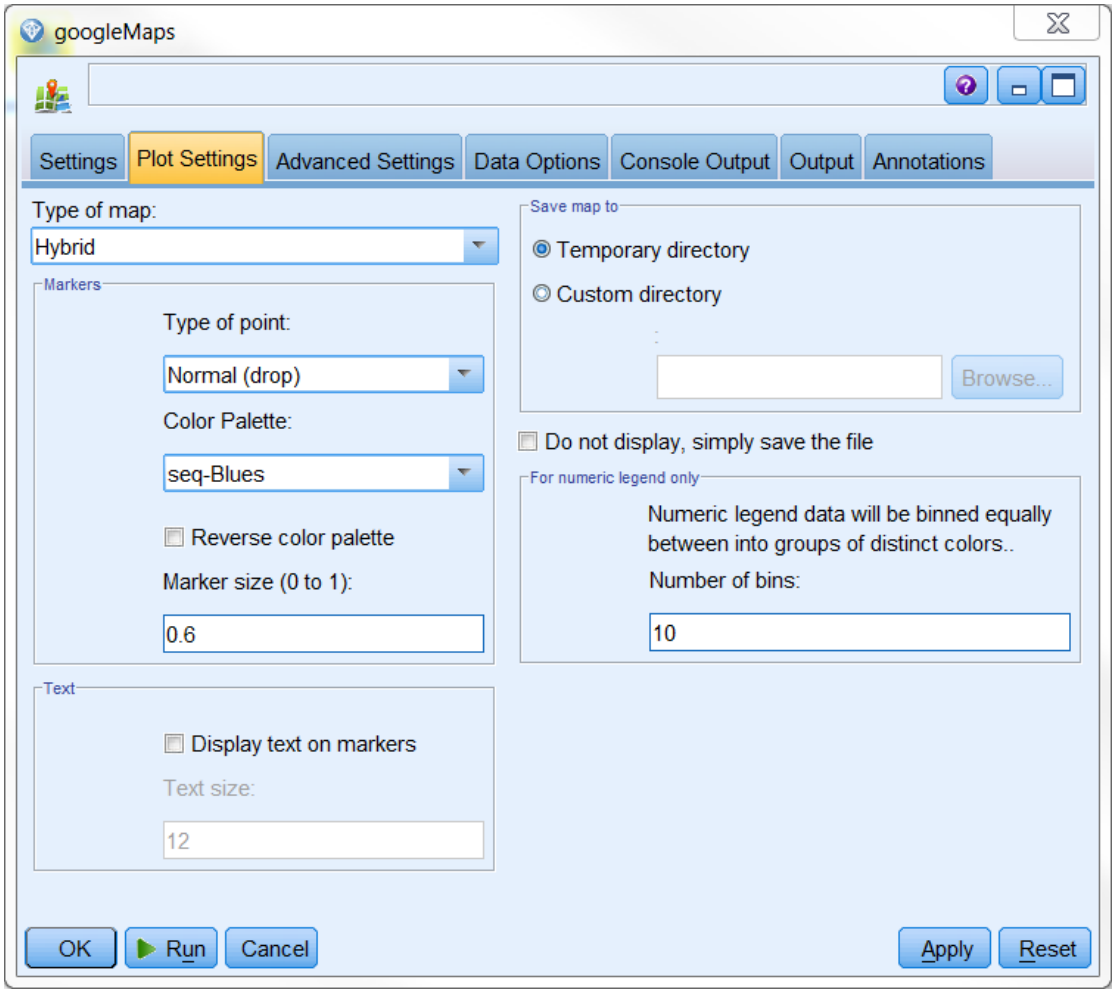




- For a numeric legend, the values will be binned, that is to say ranked according to the value then divided in groups of equal size. The default number of bins is ten, but you can specify another in the *plot settings* tab. The binning process only works with actual numeric data, so using “numeric legend”. If the legend is e.g. a string will result in an error.

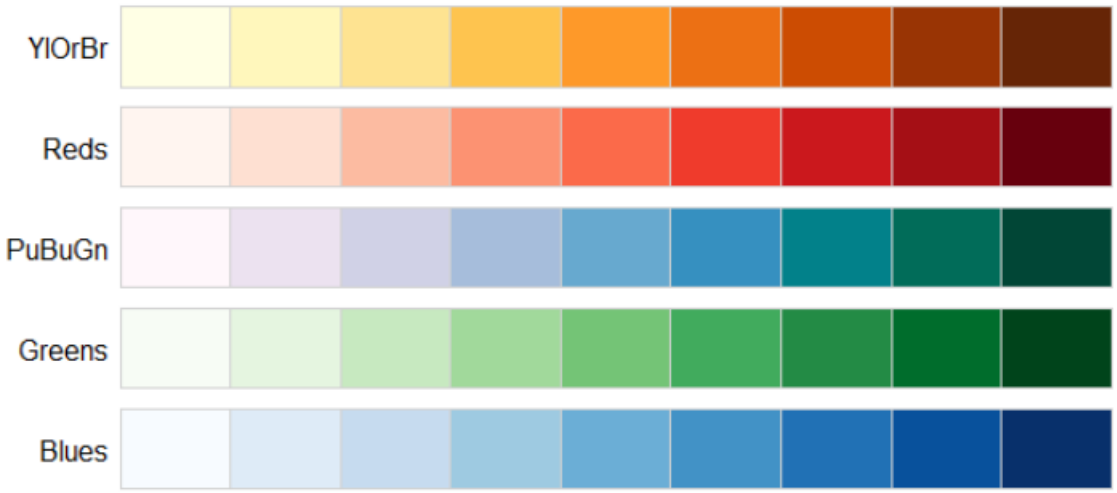


2. Plot settings

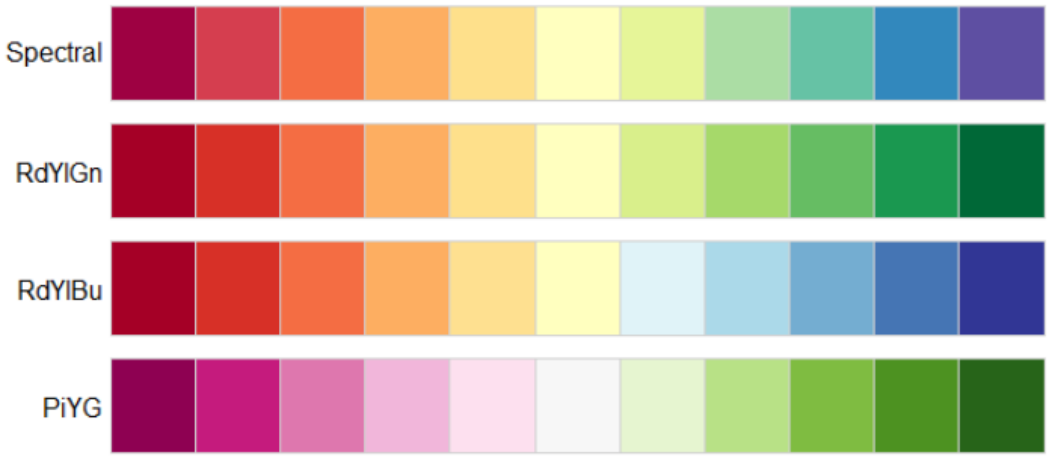


- Type of map:
  - Hybrid (default): satellite view + labels (roads, cities, etc.)
  - Satellite
  - Roadmap
  - Terrain: Roadmap + terrain information
- Type of point:
  - Normal 
  - Text-only 

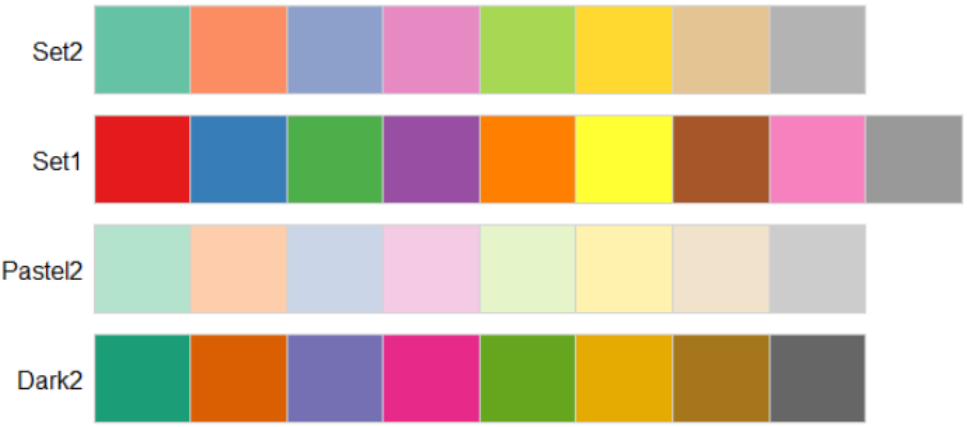
- Color palette: choose a color theme.
  - If you did not specify a legend field, you should select a “mono” palette.
  - There are several types of palettes. What follows is from the help of the RColorBrewer package.
    - Sequential palettes are suited to ordered data that progress from low to high. Lightness steps dominate the look of these schemes, with light colors for low data values to dark colors for high data values.



    - Diverging palettes put equal emphasis on mid-range critical values and extremes at both ends of the data range. The critical class or break in the middle of the legend is emphasized with light colors and low and high extremes are emphasized with dark colors that have contrasting hues.
    -



- Qualitative palettes do not imply magnitude differences between legend classes, and hues are used to create the primary visual differences between classes. Qualitative schemes are best suited to representing nominal or categorical data. The 9-colors palettes were hand-picked. If you need a lot of colors (typically more than 9), the color palettes will be programmatically expanded as needed but they will lose in color quality.



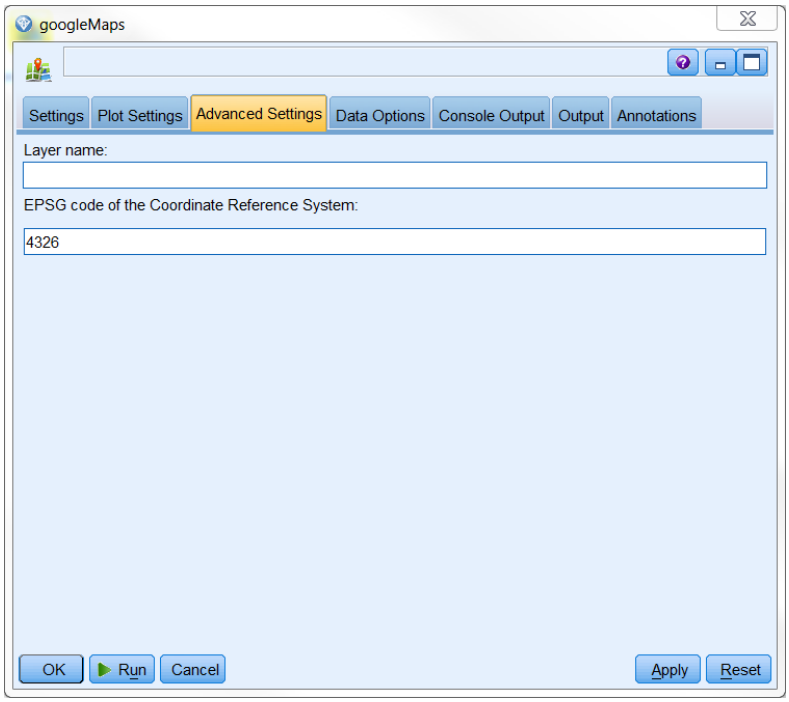
- Monochrome palette: hardly a 'palette', they consist of a single color. Available are: blue, red, green, orange and white.

Actual colors: steelblue, indianred, forestgreen, darkorange, floralwhite. These are more esthetic than their saturated counterparts.

- Reverse color palette: You can reverse the color palette. For example, you may feel that dark colors should represent low values in a sequential color palette, or you may want a divergent color palette to evolve from blue to red (use case: temperature)

- Marker size: set the marker size. Default is 0.6. This sets only the size of normal markers.Text-only markers are text, so they depend on the text size.
- Display text on marker: check if you want to display the legend label on the markers. Unchecked by default.
- Text size : set the size of texts, for “text-only” markers or text on markers in general. Default is 12.
- Save map: by default, the html file is written to temporary directory. If you want to keep the html file, you can specify a directory here. Note that they are actually 2 files: the .html and .png which contains the legend and which is saved along with the html.
- Do not display: if checked, the map won’t be opened in the default browser when the node is run. The map will still be generated and you will still be able to open the html file manually, but this will be difficult if you did not specify a custom directory to save the file.
- Number of bins: Specify the number of bins to use if a “numeric legend” is selected in the *settings* tab.

3. Advanced Settings



- Layer name: you can specify a custom layer name. It will be printed above the legend. This is really useful only for multi-layered map, a feature not yet available.

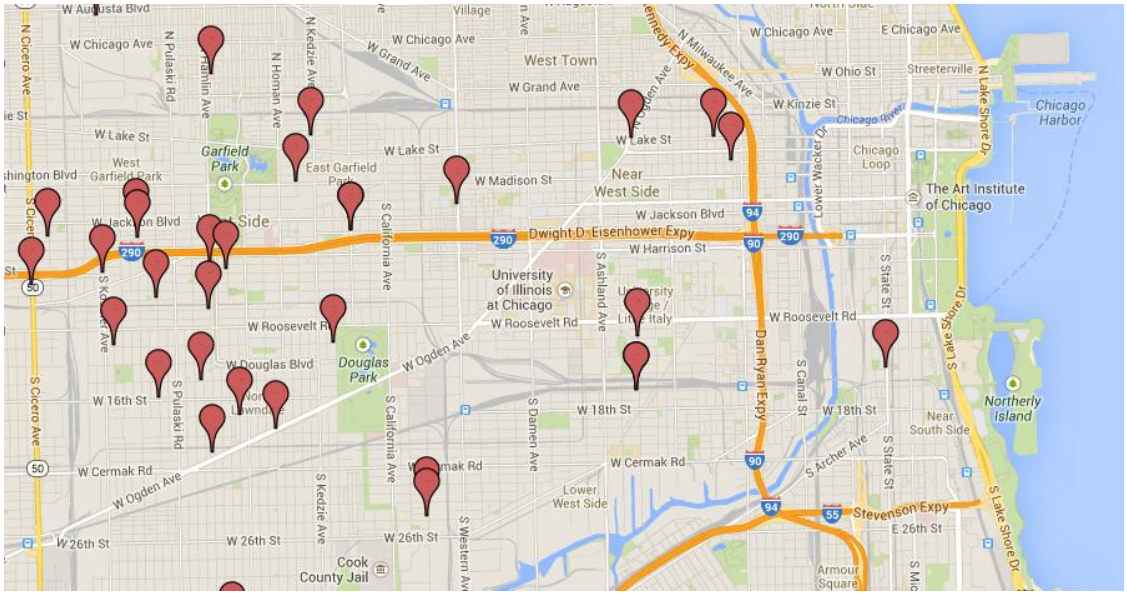
- EPSG code of the Coordinate Reference System (CRS)
    - Earth is a three-dimensional body, roughly spherical in shape, yet the vast majority of maps are flat (2-dimensional). A Coordinate Reference System (CRS) defines a method of projecting all or part of the Earth onto a 2D surface. (source: openstreetmaps.org). Every time you use spatial data (for this node, this means data with latitude and longitude), a specific CRS was used to generate the coordinates. To plot your data at the correct location, you must know the CRS. There is no way to determine it from the data, the CRS information is pure metadata.
    - All the common CRS are referenced by an EPSG code. If you know what CRS was used, a simple search will give you the EPSG code. If your points are not where they are supposed to be, the EPSG code of the CRS may be wrong. For example, US federal institutions use a specific CRS when sharing data, British institutions use another, most web APIs use EPSG 4326 or EPSG 3857.
    - The default is 4326, corresponding to the CRS WGS84 which is very common. This choice is motivated by the fact that the Google Maps service as well the Openstreetmaps service use this CRS. Therefore, this node is compatible by default with the Geocoding node and with any data using spatial data generated with the Google Maps API or the Openstreetmaps API.

Result example

These are some map outputs that you can easily reproduce. The data used is crime data gathered from the City of Chicago data portal. After removing sensitive information, casting the “age” to numeric and using a geocoding node to get Latitude and Longitude from the address, the table looks like this in SPSS:

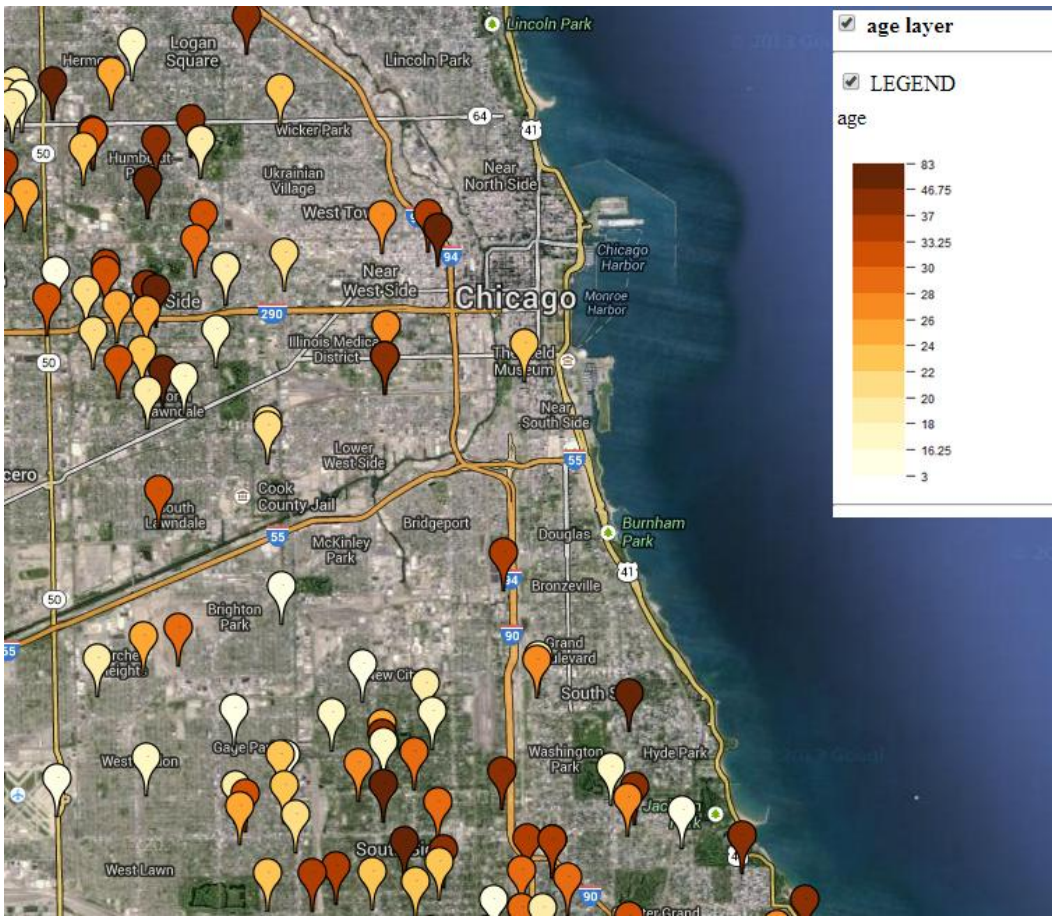
	Address (block of)	Locatison	Neighborhood	Gender	Cause	Story url	Location	age	Latitude	Longitude
1	100 N. Green St.	Sidewalk	Near West Side	Male	Gunshot	http://www.chicagobreakingnews.com/2009/03/murder-charges-new-years-day-chicago-near-west-side.html	100 N. Green St. Chicago	47	41.883	-87.649
2	11300 S. Racine Ave.	Alley	Morgan Park	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-shot-in-vehicle-on-far-south-side.html	11300 S. Racine Ave. Chicago	34	41.688	-87.652
3	2100 E. 68th St.	Street	South Shore	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-found-shot-to-death-in-vehicle.html	2100 E. 68th St. Chicago	40	41.772	-87.574
4	9200 S. Dobson Ave.		Burnside	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-shot-dead-south-side.html	9200 S. Dobson Ave. Chicago	20	41.727	-87.599
5	5900 S. Springfield Ave.	Sidewalk	West Lawn	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-shot-dead-on-south-side-1.html	5900 S. Springfield Ave. Chicago	18	41.786	-87.721
6	5000 S. Wabash Ave.	Street	Grand Blvd	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/shooting-south-side-homicide-bronzeville.html	5000 S. Wabash Ave. Chicago	28	41.804	-87.624
7	8900 S. Halsted St.	Street	Washington Heights	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-found-fatally-shot-in-trunk-of-vehicle.html	8900 S. Halsted St. Chicago	34	41.732	-87.644
8	6500 S. Western Ave.	Street	Chicago Lawn	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/south-side-shooting-homicide.html	6500 S. Western Ave. Chicago	21	41.775	-87.684
9	6200 S. Sacramento A...	Apartment	Chicago Lawn	Male	Stabbing	http://www.chicagobreakingnews.com/2009/01/rodney-bryant-greg-stabbed-sacramento.html	6200 S. Sacramento Ave. Chic...	19	41.781	-87.699
10	4600 S. Drake Ave.	Street	Brighton Park	Female	Strangling	http://www.chicagobreakingnews.com/2009/01/womans-body-found-on-southwest-side.html	4600 S. Drake Ave. Chicago	29	41.810	-87.712
11	3700 W. Lexington St.	Residen...	East Garfield Park	Male	Stabbing	http://www.chicagobreakingnews.com/2009/01/man-fatally-stabbed-in-west-garfield-park.html	3700 W. Lexington St. Chicago	54	41.872	-87.718
12	0 W. 15th St.	Residen...	Near South Side	Female	Gunshot	http://www.chicagobreakingnews.com/2009/01/woman-infant-found-slain-on-south-side.html	0 W. 15th St. Chicago	24	41.862	-87.627
13	0 W. 15th St.	Residen...	Near South Side	Female	Gunshot	http://www.chicagobreakingnews.com/2009/01/woman-infant-found-slain-on-south-side.html	0 W. 15th St. Chicago	$null$	41.862	-87.627
14	1100 W. 122nd St.	School	West Pullman	Male	Gunshot	http://www.chicagobreakingnews.com/2009/03/police-alerts-ask-for-public-help.html	1100 W. 122nd St. Chicago	33	41.672	-87.649
15	2000 W. 70th Pl.	Residen...	West Englewood	Male	Stabbing	http://www.chicagobreakingnews.com/2009/01/man-charged-in-brothers-killing.html	2000 W. 70th Pl. Chicago	35	41.766	-87.674
16	1600 S. Loomis St.	Residen...	Lower West Side	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-found-fatally-shot-in-pilsen-alley.html	1600 S. Loomis St. Chicago	43	41.860	-87.662
17	8100 S. Houston Ave.	Alley	South Chicago	Male	Gunshot		8100 S. Houston Ave. Chicago	47	41.748	-87.551
18	6800 S. Indiana Ave.	Apartment	Greater Grand Cro...	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/2-found-slain-in-south-side-apartment.html	6800 S. Indiana Ave. Chicago	36	41.771	-87.621
19	6800 S. Indiana Ave.	Apartment	Greater Grand Cro...	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/2-found-slain-in-south-side-apartment.html	6800 S. Indiana Ave. Chicago	51	41.771	-87.621
20	12200 S. State St.	Residen...	West Pullman	Male	Stabbing	http://www.chicagobreakingnews.com/2009/02/1m-bail-in-west-pullman-slaying.html	12200 S. State St. Chicago	61	41.672	-87.623
21	1400 W. Lake St.	Bar or ta...	Near West Side	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/3-wounded-on-near-west-side-street.html	1400 W. Lake St. Chicago	28	41.885	-87.662
22	8100 S. Eberhart Ave.	Sidewalk	Chatham	Male	Gunshot	http://www.chicagobreakingnews.com/2009/01/man-killed-in-south-side-street-robbery.html	8100 S. Eberhart Ave. Chicago	55	41.747	-87.613
23	3900 W. Argyle St.	Apartment	Albany Park	Female	Arson	http://www.chicagobreakingnews.com/2009/03/2nd-suspect-in-fatal-arson-released.html	3900 W. Argyle St. Chicago	7	41.972	-87.726
24	3900 W. Argyle St.	Apartment	Albany Park	Female	Arson	http://www.chicagobreakingnews.com/2009/03/2nd-suspect-in-fatal-arson-released.html	3900 W. Argyle St. Chicago	23	41.972	-87.726
25	8100 S. Drexel Ave.	Sidewalk	Chatham	Male	Stabbing	http://www.chicagobreakingnews.com/2009/02/4-women-charged-in-stabbing-slaying.html	8100 S. Drexel Ave. Chicago	30	41.747	-87.603
26	10300 S. Doty Ave.	Other	South Deering	Male	Gunshot		10300 S. Doty Ave. Chicago	50	41.704	-87.588
27	6200 S. Laflin St.	Street	West Englewood	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/man-fatally-shot-in-west-englewood-cops-say.html	6200 S. Laflin St. Chicago	48	41.781	-87.662
28	7200 S. Racine Ave.	Sidewalk	Englewood	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/georgia-man-slain-in-englewood-officials-say.html	7200 S. Racine Ave. Chicago	24	41.763	-87.654
29	5900 S. Campbell Ave.	Sidewalk	Chicago Lawn	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/teen-shot-marquette-park-campbell.html	5900 S. Campbell Ave. Chicago	16	41.786	-87.686
30	5200 W. Lake St.	Car wash	Austin	Male	Gunshot	http://www.chicagobreakingnews.com/2009/03/man-charged-in-double-homicide-at-carwash.html	5200 W. Lake St. Chicago	28	41.887	-87.756
31	5200 W. Lake St.	Car wash	Austin	Male	Gunshot	http://www.chicagobreakingnews.com/2009/03/man-charged-in-double-homicide-at-carwash.html	5200 W. Lake St. Chicago	34	41.887	-87.756
32	2400 W. Adams St.	Porch/h...	Near West Side	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/2-slain-in-separate-west-side-shootings.html	2400 W. Adams St. Chicago	22	41.879	-87.686
33	7500 S. Wabash Ave.	Street	Greater Grand Cro...	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/man-slain-in-south-side-shooting.html	7500 S. Wabash Ave. Chicago	19	41.758	-87.623
34	1900 S. Loomis St.	Sidewalk	Lower West Side	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/2-slain-in-separate-west-side-shootings.html	1900 S. Loomis St. Chicago	16	$null$	$null$
35	1500 S. Avers Ave.	Apartment	North Lawndale	Male	Stabbing	http://www.chicagobreakingnews.com/2009/02/3-chicagoans-slain-in-sunday-attacks.html	1500 S. Avers Ave. Chicago	23	41.861	-87.722
36	7800 S. Western Ave.	Not avail...	Ashburn	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/3-chicagoans-slain-in-sunday-attacks.html	7800 S. Western Ave. Chicago	21	41.752	-87.683
37	2900 W. 63rd St.	Sidewalk	Chicago Lawn	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/3-chicagoans-slain-in-sunday-attacks.html	2900 W. 63rd St. Chicago	33	41.779	-87.696
38	7200 S. Calumet Ave.	Alley	Greater Grand Cro...	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/man-found-slain-in-south-side-alley.html	7200 S. Calumet Ave. Chicago	30	41.764	-87.617
39	1500 W. 58th St.	Vacant ...	West Englewood	Male	Gunshot	http://www.chicagobreakingnews.com/2009/02/teen-found-shot-in-vacant-lot.html	1500 W. 58th St. Chicago	18	41.789	-87.662

You may want to simply plot the points by Latitude and Longitude, for example in red, over a simple roadmap.

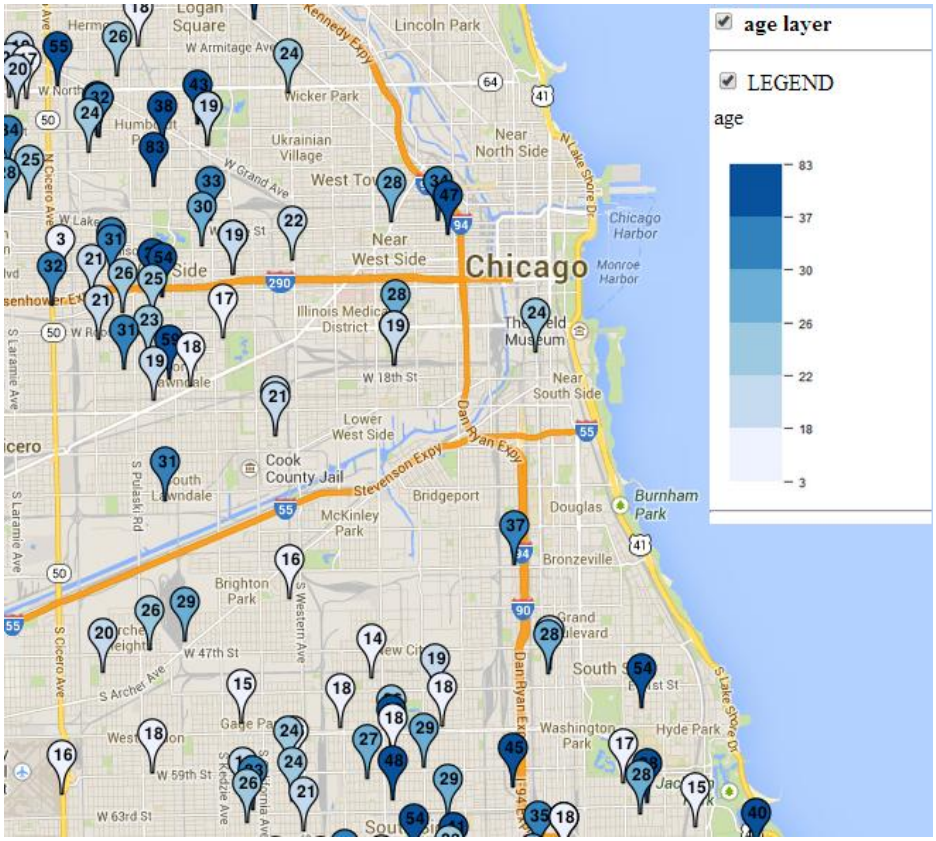


This node also allows you to use a color-coded legend.

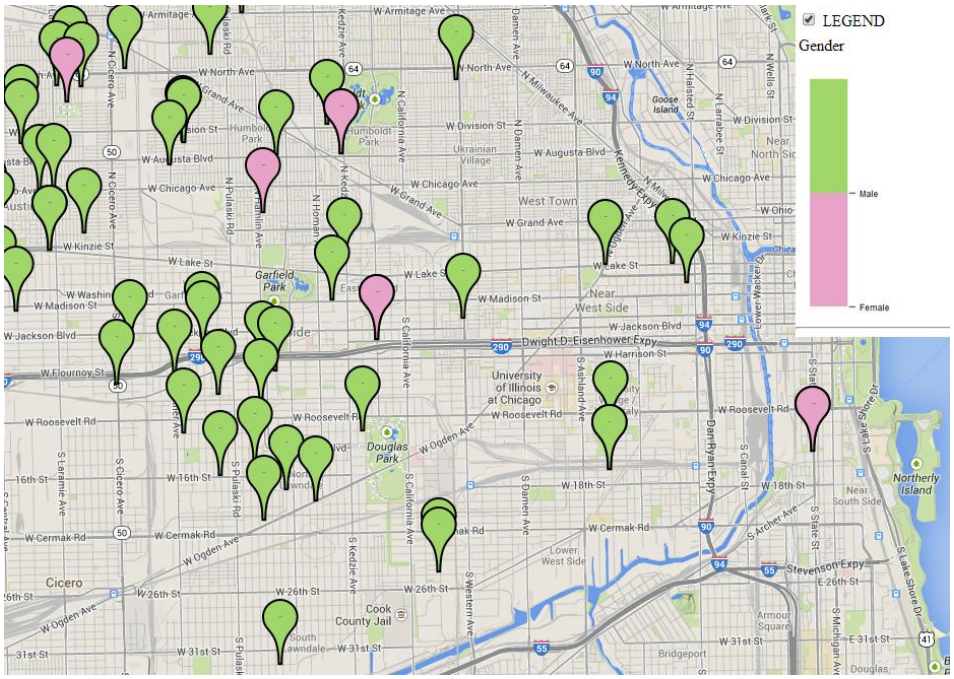
It can be continuous, like the ‘age’ field. Note that a numeric legend will be binned into a number of group that you can specify. Here it is with 12 bins on a hybrid map.



Here we used a sequential color palette. Several color palettes are available, fit for different uses. You can also choose the type of map background. You can also write the label on the dots if you wish. Let's plot the age again with less bins and a few different options.

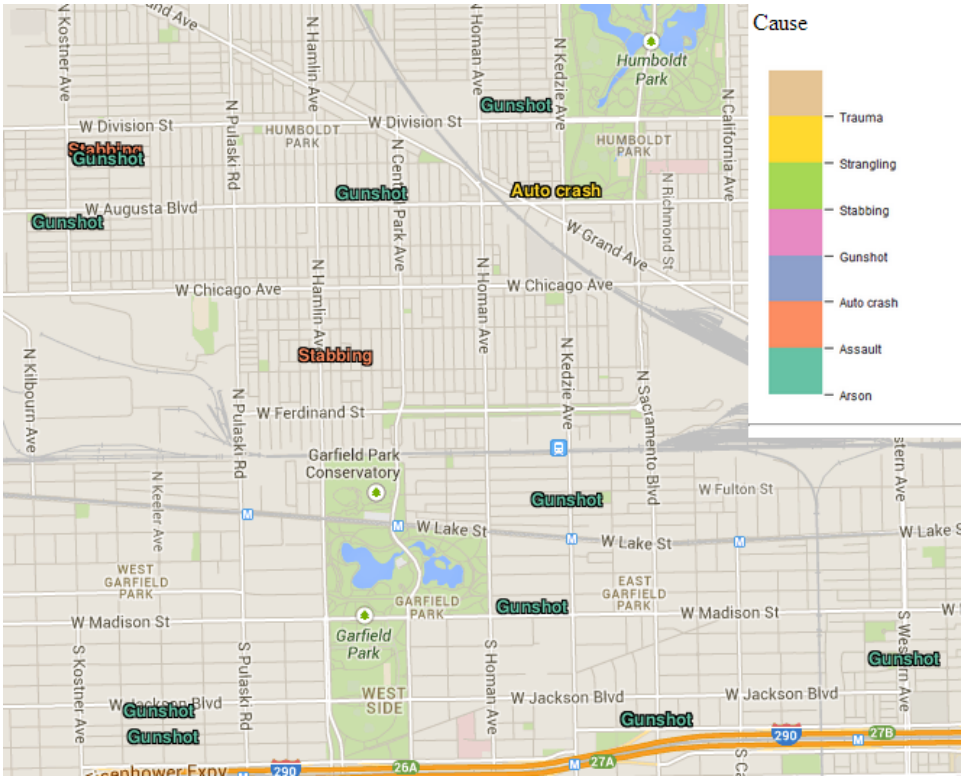


The legend can also be categorical (whether nominal or flag). In that case, each value will have its on color and no binning is attempted. You can also specify the size of the markers. To illustrate this, let's plot a flag field, for example the 'gender', with bigger markers.

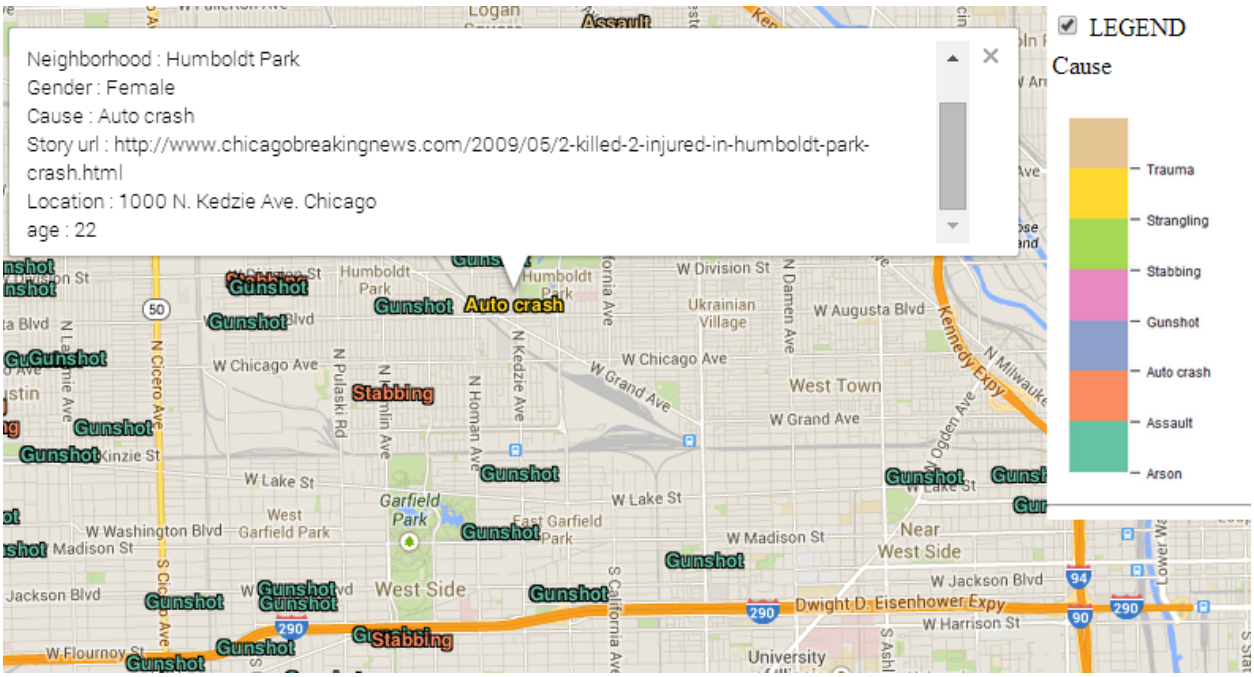


Finally, let's plot a categorical field with more than 2 possible values, for example the 'cause'. We will use the other type of markers available: text-only markers. You can set the size of the text just like you could set the size of the normal markers.

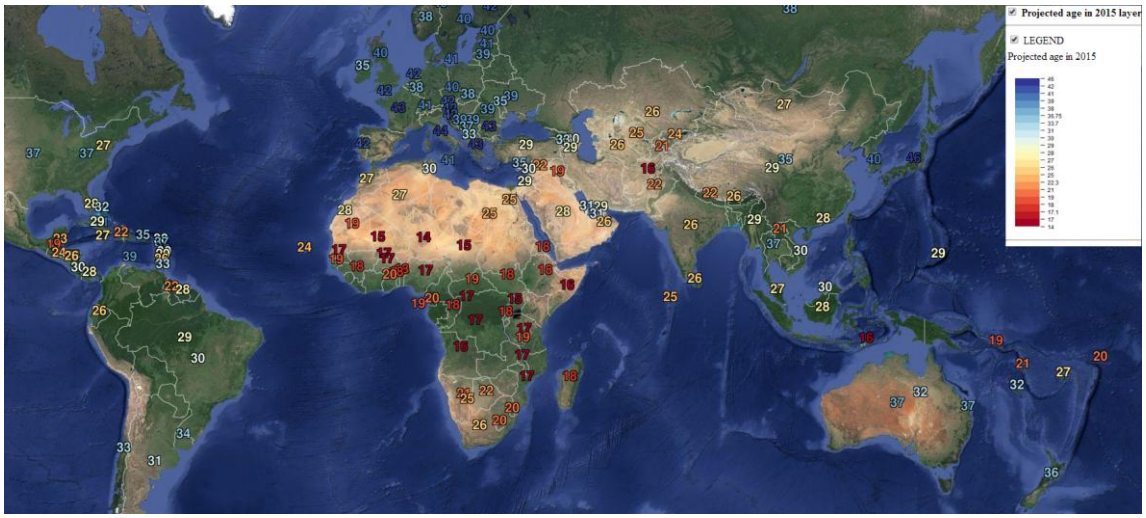
The most adapted color palette is here a qualitative one, designed to distinguish easily the different values.



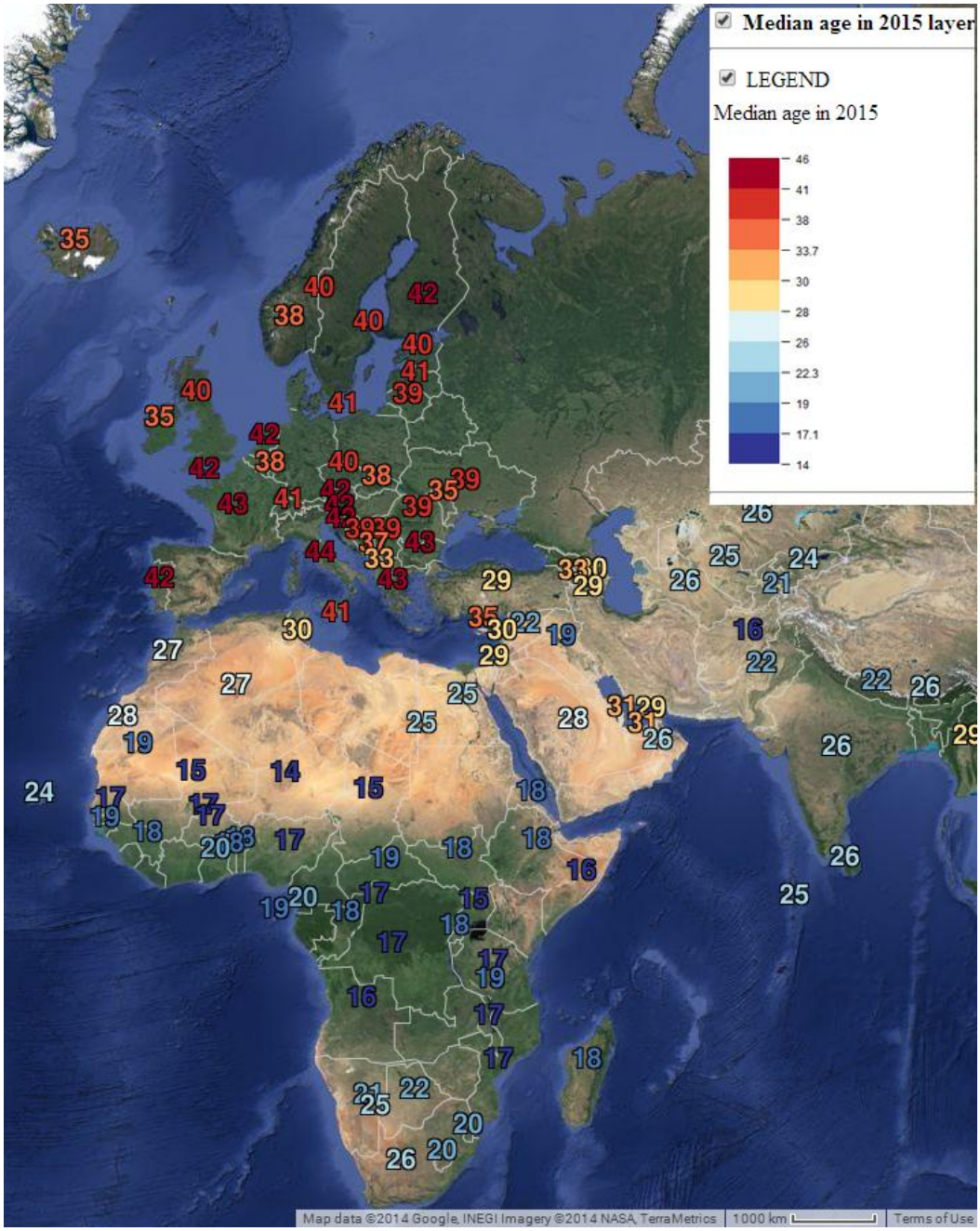
At any time, you can click on one of the markers to display the information of the corresponding row.



The text-only markers also work with a numeric legend. For this final picture, we will plot the median age in the world, using data from the UN and the geocoding node to get Longitude and Latitude from the country name. This is also the occasion to show the last type of color palette: divergent, designed to highlight the extreme values.



Finally, you may want to reverse the color palette if the color code is counter-intuitive. For example, plotting temperature with the above color code would result in high temperature in blue and low temperature in red. This is not an issue: just check “reverse color palette”.


Median age in 2015 layer
LEGEND
Median age in 2015
46
41
38
33.7
30
28
26
22.3
19
17.1
14
Map data ©2014 Google, INEGI Imagery ©2014 NASA, TerraMetrics
1000 km
Terms of Use

Important links

Learn

- Learn more about SPSS software.
- Visit developerWorks Business analytics for more technical analytics resources for developers.
- The Comprehensive R Archive Network is the main site for the R project and each R package. The help pages and manuals that are associated with `optimx`, `nlmrt`, and `Rcgmin` are detailed. Numerous references are provided.
- Read "Do I need to learn R?" (Catherine Dalzell, developerWorks, September 2013) to learn why R is a valuable tool for data analytics that was expressly designed to reflect the way that statisticians think and work.
- "Calling R from SPSS" describes how to use R code inside IBM SPSS Modeler 16.
- Read "Using Google maps API" to discover how to use Google Maps API with R.
- Read "Create new nodes for IBM SPSS Modeler 16 using R" to learn how to create new extensions easily.

Get products and technologies

- Download the R plug-in for SPSS plugin.
- Download the R 2.15.2 for Windows package.

Discuss

- Visit the IBM SPSS DevCentral developerWorks community to share tips and experiences with other IBM SPSS developers.
- Follow developerWorks on Twitter to be among the first to hear about new resources.