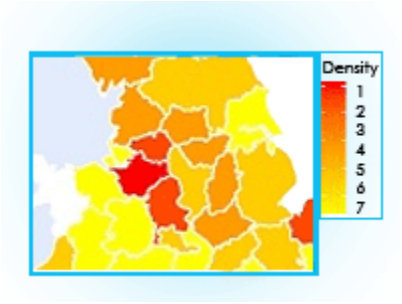
Plot heatmaps

Visualize your spatial data with density heatmaps



Product: IBM® SPSS® Modeler

Extension type: Output/Visualization

Table of Contents

Description

This SPSS Modeler 'output' node allows you to plot spatial data on a density heatmap. Simply install the node, select the coordinate's fields and choose the visual parameters that suit you best. Then start visualizing simply and easily your data directly in IBM® SPSS® Modeler.

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.

- 'ggmap' package created by David Kahle and Hadley Wickham,
  available on http://cran.r-project.org/web/packages/ggmap/

- 'ggplot2' package created by Hadley Wickham and Winston Chang,
  available on http://cran.r-project.org/web/packages/ggplot2/

Tutorial

1.Import and prepare your data.

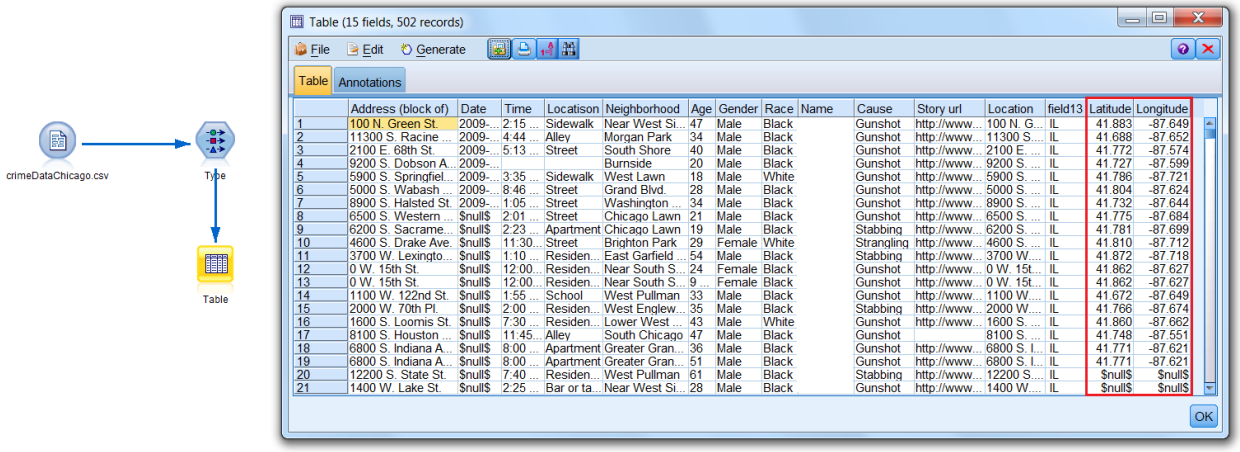
For example, we imported in SPSS Modeler a .csv file with Chicago crime data (this official dataset is available for free on https://data.cityofchicago.org/ and a sample of it is given with this tutorial).

"Latitude" and "Longitude" columns are required for the 'Google Heatmaps' node to work.

NB: *If your data only has an "address" column you can install the 'Geocoding' node which transforms addresses into coordinates.*

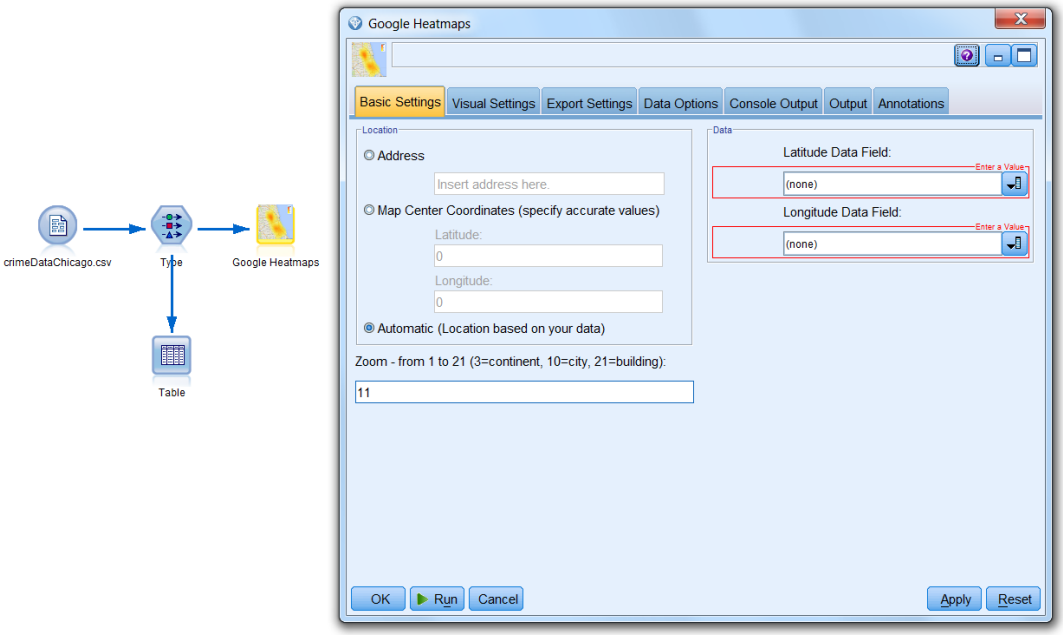
Your coordinates must be typed as "Continuous" in Modeler; you can do this with a 'Type' node.

NB: *The heatmaps node automatically discards rows containing $null$ in 'Latitude' and/or 'Longitude'.*



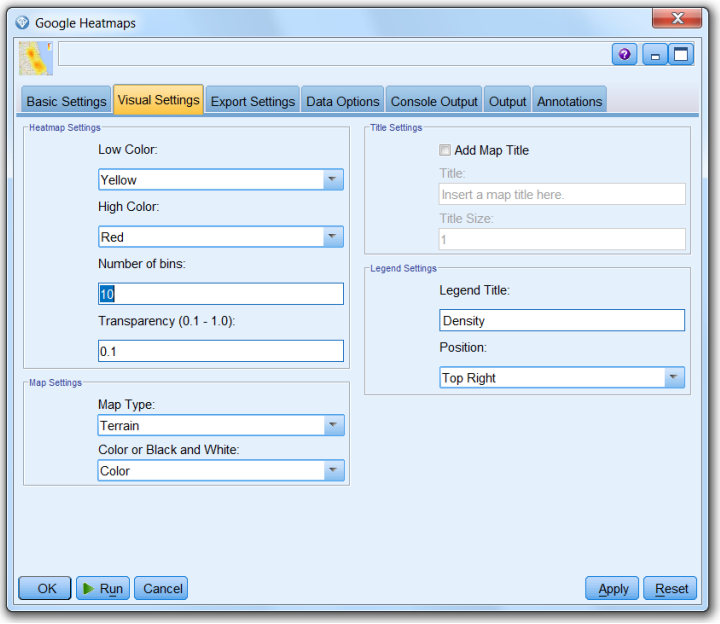
2. Connect a 'Google Heatmaps' node to the last node of your stream, with your data prepared.

Double click on the 'Google Heatmaps' node.

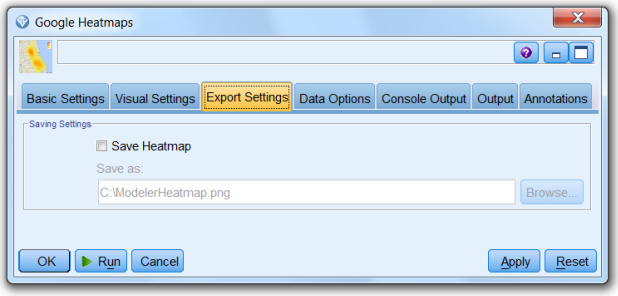


Now you have to fill the fields within three tabs:

- Tab ‘Basic Settings’:
  - ‘Location’. You have three possible choices:
    - ‘Address’. Example: “Chicago”
    - ‘Map Center Coordinates’. The map will be centered on these coordinates; for example Latitude = 41.882598 and Longitude = -87.630901 which are the coordinates of Chicago (information available on Wikipedia.org)
    - ‘Automatic’. The location will be based on your data.
  - ‘Zoom’. Choose an integer between 1 and 21. 3 is scaled for a continent; 10 for a city; and 21 for a building.
  - ‘Latitude Data Field’ and ‘Longitude Data Field’. Select your Latitude and Longitude columns.



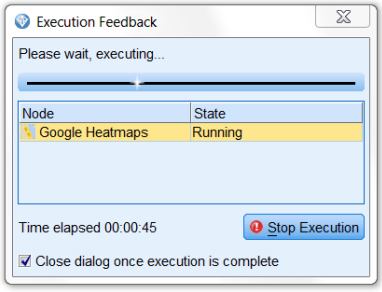
- Tab ‘Visual Settings’:
  - ‘Low Color’ and ‘High Color’. Choose which colors you want you heatmap plotted.
  - ‘Number of bins’. With a high number of bins (20+) the heatmap will be smoother; with a low number of bins (8-) you will be able to more see specific zones on the map.
  - ‘Transparency’. Choose a real number between 0.1 and 1.
  - ‘Map Settings’. Several types of maps are available: Terrain; Satellite; Roadmap; Hybrid (GoogleMaps); Watercolor; Toner (Stamen Maps). It can be color or Black and White.
  - ‘Title Settings’. You can choose whether to add or not a map title and specify its size.
  - ‘Legend Settings’. Choose the legend title and its position.



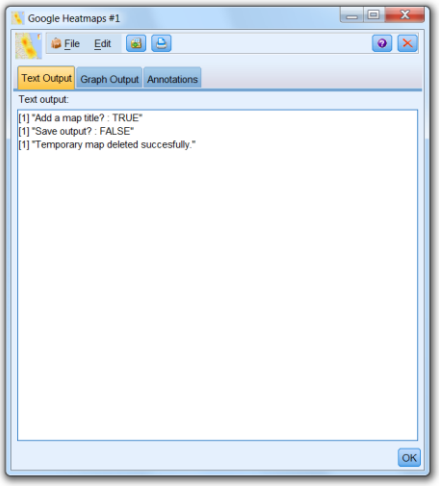
- Tab 'Export Settings': Choose whether to save or not your heatmap as a png file and specify the folder and filename you prefer. NB: We recommend to save the heatmap only once you are satisfied with its general look (see steps 3 to 5).

After filling the fields, click on 'Run'.

3. The process should last approximately 1 minute; so don't be worried if it is taking some time to finish.

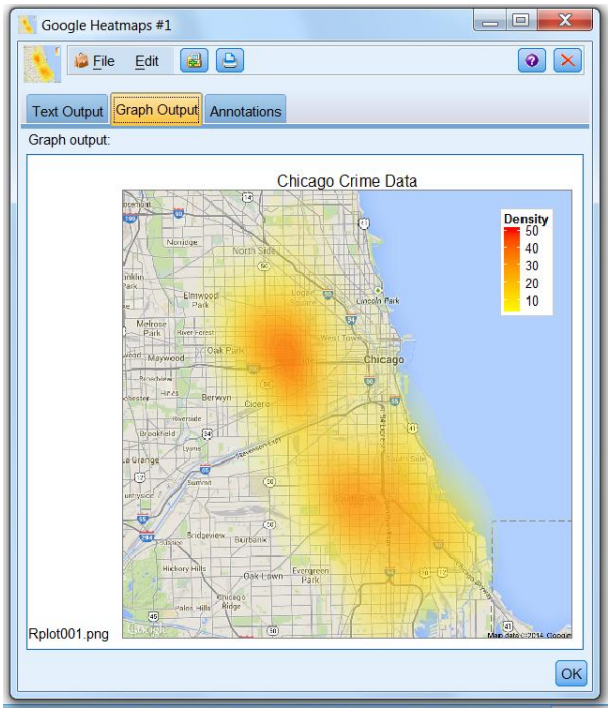


After it is finished, you will see an output window appear:



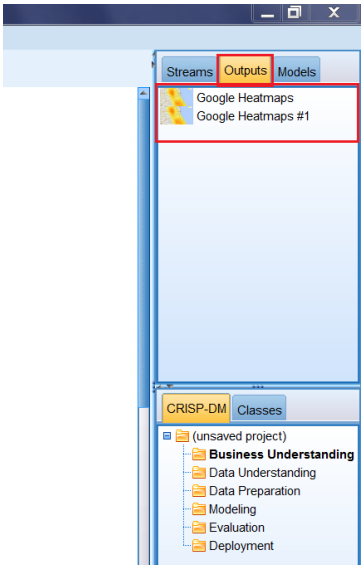
The open tab will be 'Text Ouput' which allows you to know if some operations worked correctly; adding or not a title to the map, saving or not the ouput and whether the temporary file has been deleted.

4. Now go to the ‘Graph Output’ tab to see your creation:

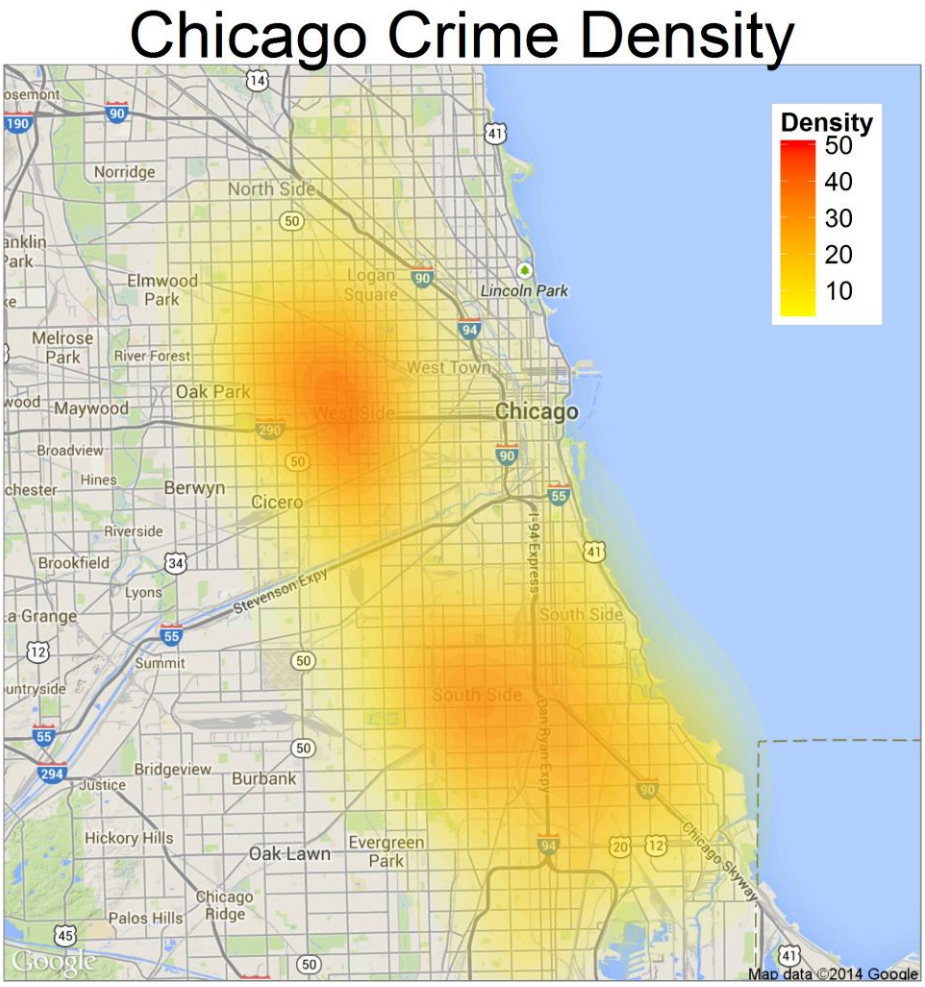


This view is only meant as a preview with a bad picture definition. You will be able to get the heatmap in high definition by saving it through the ‘Export Settings’ tab.

Check if you are satisfied with the colors, the titles and the global look of your heatmap. You can try different visual settings; their previews will all be available in the Ouput tab of the main SPSS window:

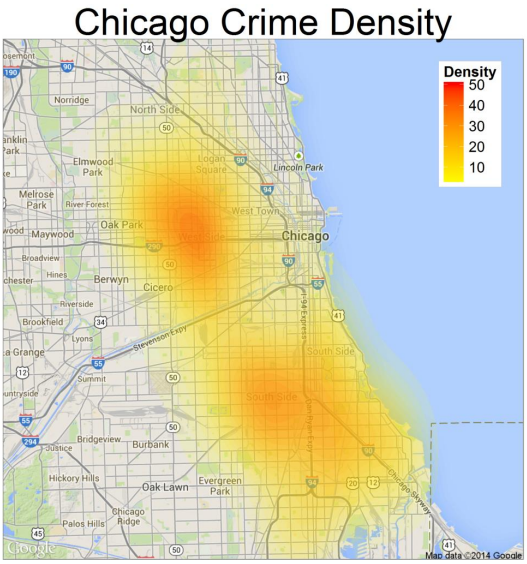


5. Once you are satisfied with the look of your preview, go back to the node and save your heatmap using the 'Export Settings' tab. Here is an example of what you can get with a few clicks:



Result example

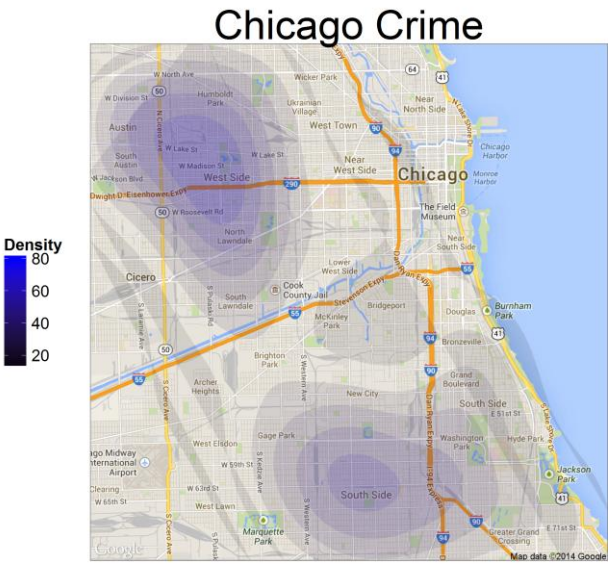
Here are a few examples of heatmap plotting of Chicago crime data, with the chosen parameters.



Location: Automatic / Zoom = 11

Low Color: Yellow / High Color: Red / Number of bins = 20 / Transparency = 0.1

Map: Terrain, Color / Map Title Size = 2

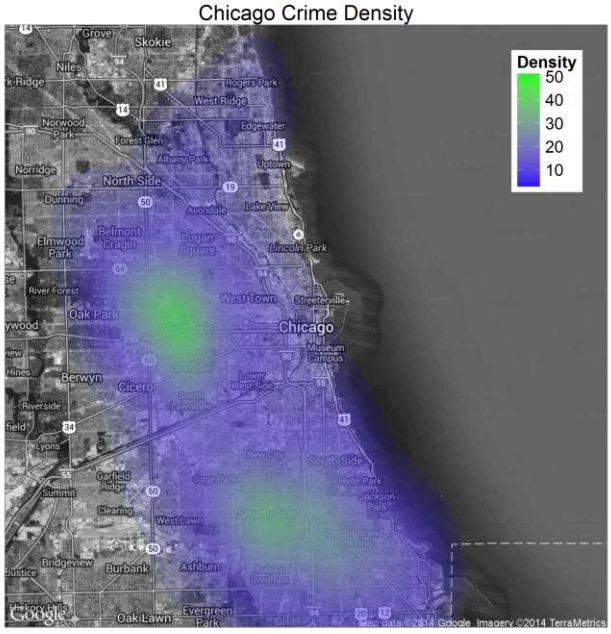


Location: Automatic / Zoom = 12

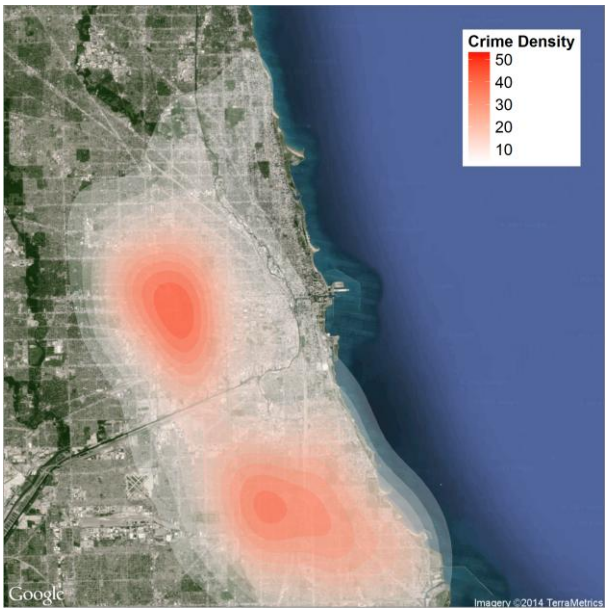
Low Color: Black / High Color: Blue / Number of bins = 8 / Transparency = 0.1

Map: Roadmap, Color / Map Title Size = 2

NB: Sometimes plotting bugs appear when zooming so much that a lot of points are outside the view. We plan on fixing this for the next version of the node. One option is to select the data before plotting.



Location: Address = Chicago / Zoom = 11
Low Color: Blue / High Color: Green / Number of bins = 25 / Transparency = 0.1
Map: Hybrid, B&W / Map Title Size = 1



Location: Lat = 41.882598 ; Lon = -87.630901 / Zoom = 11
Low Color: White / High Color: Red / Number of bins = 12 / Transparency = 0.25
Map: Satellite, Color / No Map Title

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.