

**STATCON**

# Spatial Statistics with JMP

---

A JMP Addin for Spatial Statistics and Maps

**Sebastian Hoffmeister**

**26.03.2015**

**Addin Version: 0.9**

This is the documentation for a JMP-Add-in about spatial statistics with JMP. The add-in allows you to geocode addresses, import shape-files as maps, calculate distance matrices and solve traveling salesmen problems (TSP).

## Table of Contents

Requirements and Support .....	2
Geocoding .....	3
Shape File Import .....	5
Solving TSPs.....	8
Literature and Links .....	11
Known Issues and Planned Improvements .....	12

## Requirements and Support

This add-in makes use of JMP's R-interface. To use most of the functions a working installation of R (tested on R version 3.1.1 and JMP version 11.2.0 with Windows 7) is required. Further the R-packages **ggmaps** and **TSP** are required.

R can be downloaded freely from <http://www.r-project.org>. For questions about installation of R and it's packages refer to:

- R-Mailing Lists: <http://www.r-project.org/mail.html>
- R FAQs: <http://cran.r-project.org/doc/FAQ/R-FAQ.html#R%20Bugs>
- Mail to [consult@statcon.de](mailto:consult@statcon.de)

Installation of the add-In itself should be fairly simple. Just open the provided add-in-file in JMP (**File** → **Open** → **Select downloaded com.statcon.sda-file** → **Ok**).

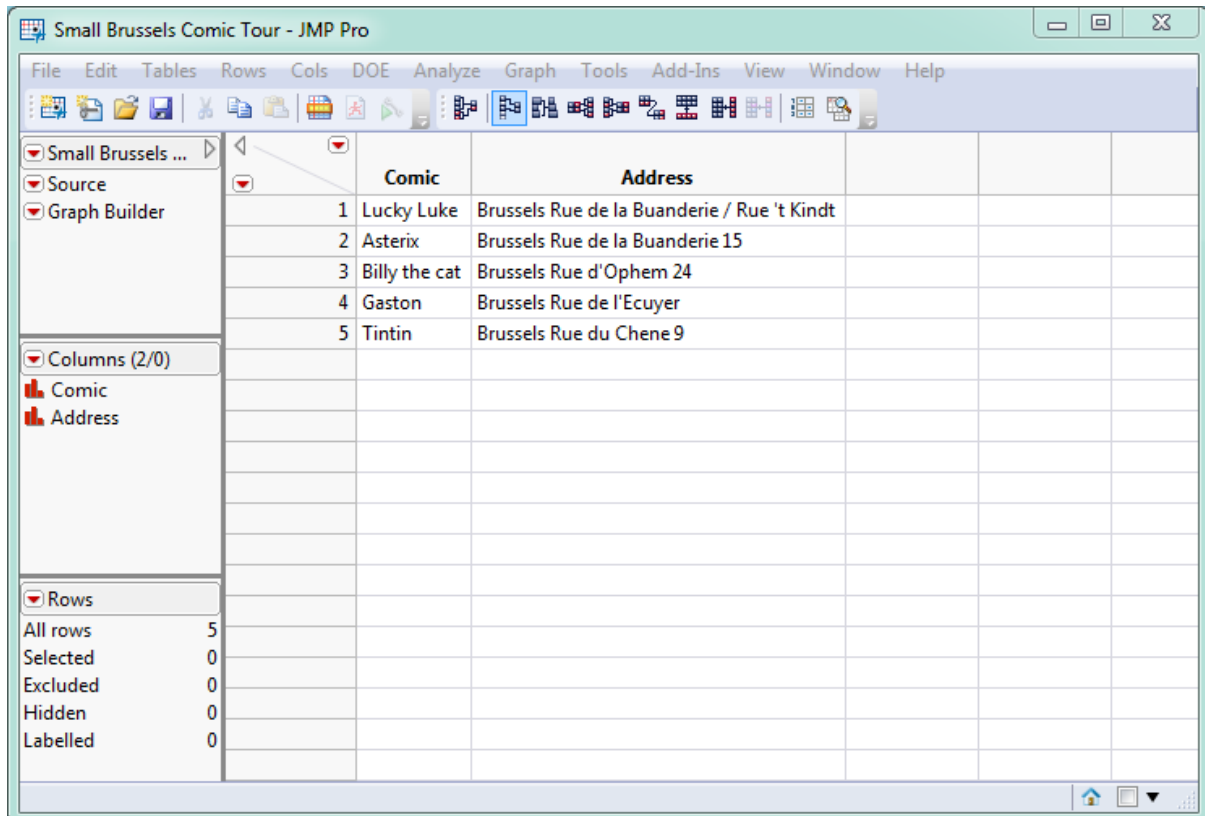
For all problems when installing the JMP-Add-In don't hesitate to ask in the JMP User Community forums (<https://community.jmp.com/welcome>) or mail to [consult@statcon.de](mailto:consult@statcon.de).

## Geocoding

Geocoding is the process of acquiring geo-coordinates for a given address. This feature requires a working internet-connection. The geocoding feature uses the google maps API. Therefore by using this feature you are accepting the googlemaps API's [terms of use](#).

### Example: Brussels Comic Tour


To geocode addresses using the add-in, create a data set with one column containing the relevant addresses as plain text.



The screenshot shows the JMP Pro interface with a data table titled 'Small Brussels Comic Tour'. The table has two columns: 'Comic' and 'Address'. The data is as follows:

	Comic	Address
1	Lucky Luke	Brussels Rue de la Buanderie / Rue 't Kindt
2	Asterix	Brussels Rue de la Buanderie 15
3	Billy the cat	Brussels Rue d'Ophem 24
4	Gaston	Brussels Rue de l'Ecuyer
5	Tintin	Brussels Rue du Chene 9

Figure 1 Selected Addresses of the Brussels Comic Tour

To start the geocoding-process select the add-Ins-menu in the main menu. Navigate to **Spatial Data Analysis** →  **Geocode**.

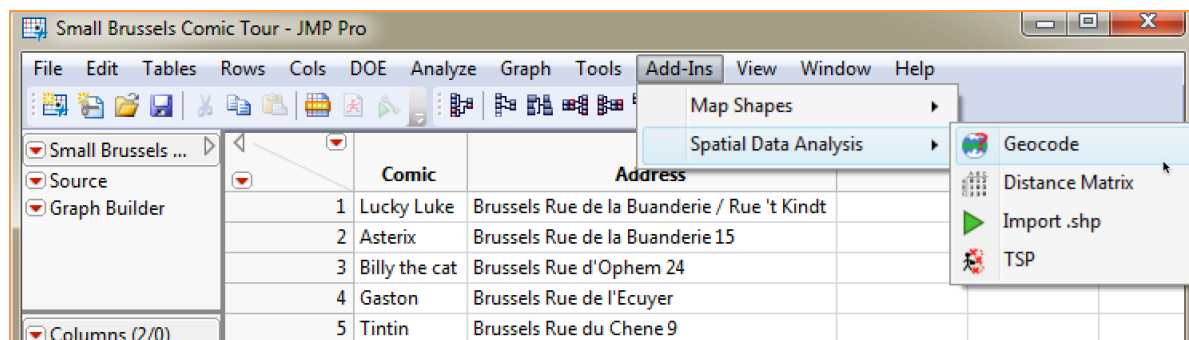
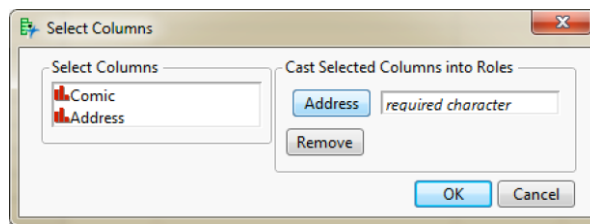


Figure 2 The geocoding menu

The dialog asks for the column containing the address data.



Be aware that the addresses are sent to the google maps API. To get the right geo-coordinates you need to provide the addresses in adequate form. E.g. geocoding “Main Street” will probably not be helpful, as there are multiple “Main Streets” in different towns all over the world. Use addresses in a way how you would use them when trying to find the address in google maps itself (e.g. “5<sup>th</sup> Avenue New York”).

Press **Ok** in the dialog and see how two new columns – representing latitudes and longitudes – will appear in your dataset. For large lists this may take a while. The google maps API restricts the number of addresses being geocoded to 2000 a day.

	Comic	Address	lat	lon
1	Lucky Luke	Brussels Rue de la Buanderie / Rue 't Kindt	50.8468° N	4.34174° E
2	Asterix	Brussels Rue de la Buanderie 15	50.8466° N	4.34222° E
3	Billy the cat	Brussels Rue d'Ophem 24	50.8535° N	4.34494° E
4	Gaston	Brussels Rue de l'Ecuyer	50.849° N	4.35424° E
5	Tintin	Brussels Rue du Chene 9	50.8446° N	4.35011° E

Figure 3 Geocoded Brussels Comic Tour Addresses

## Shape File Import

The add-in tries to ease the process of importing custom (ESRI) shape-files to make all kinds of maps available for JMP. An in depth explanation of how maps are organized in JMP is given at:

[http://www.jmp.com/support/help/Examples\\_of\\_Creating\\_Maps.shtml](http://www.jmp.com/support/help/Examples_of_Creating_Maps.shtml)

To import a shape-file into JMP, first download the relevant files from the internet. There are lots of sources around there. The following list is just a small excerpt of all available data.

### *Some Sources for shape files on the internet:*

1. **DIVA-GIS:** <http://www.diva-gis.org/gdata>  
(Download administrative areas, inland water, roads, railroads, elevation, land cover, population, climate and gazetteer for countries)  
Problem: Roads maps have non-closed shapes!
2. **Mapzen:** <https://mapzen.com/metro-extracts/>
3. **OpenStreetMap Data Extracts:** <http://download.geofabrik.de>  
(Country wise data from Open Street Map published by GEOFABRIK)
4. **PhilGIS:** <http://www.philgis.org/freegisdata.htm>  
(Shapefiles for the Philippines – Country, Cities, ...)

### *Example German Administrative Areas*

Here we will use a shape file from DIVA-GIS containing administrative areas in Germany (<http://www.diva-gis.org/gdata>). The downloaded data looks like **Figure 4**.

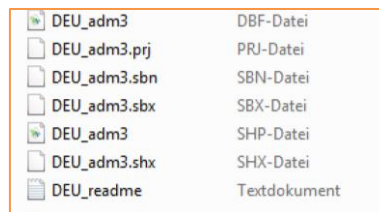
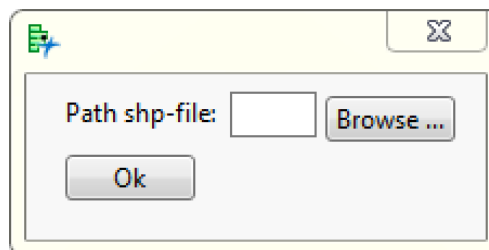
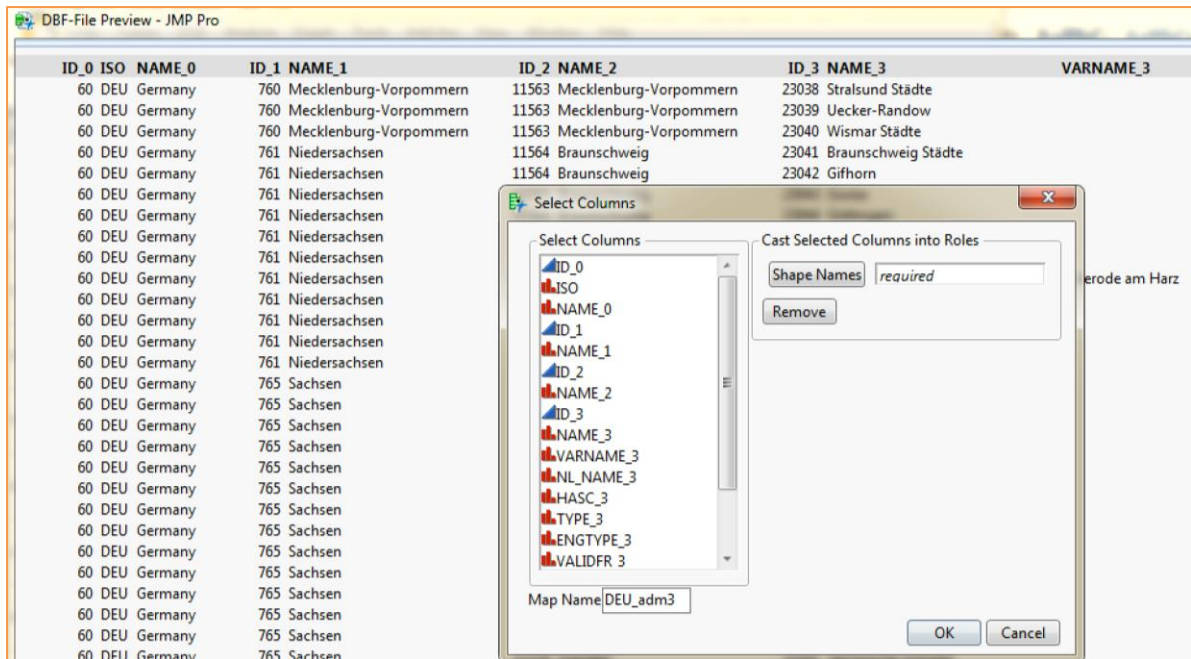


Figure 4 Shape File Data from div-gis.org

To import the map to JMP go the Add-In menu and select **Spatial Data Analysis** → **Import .shp**.



The dialog only asks for the location of the shape-file to be imported. Select the file using the Browse-button or paste the complete path to the text-field and press **Ok**. JMP will then load the corresponding .shp and .dbf and display a dialog and a preview of the .dbf-file in the background.

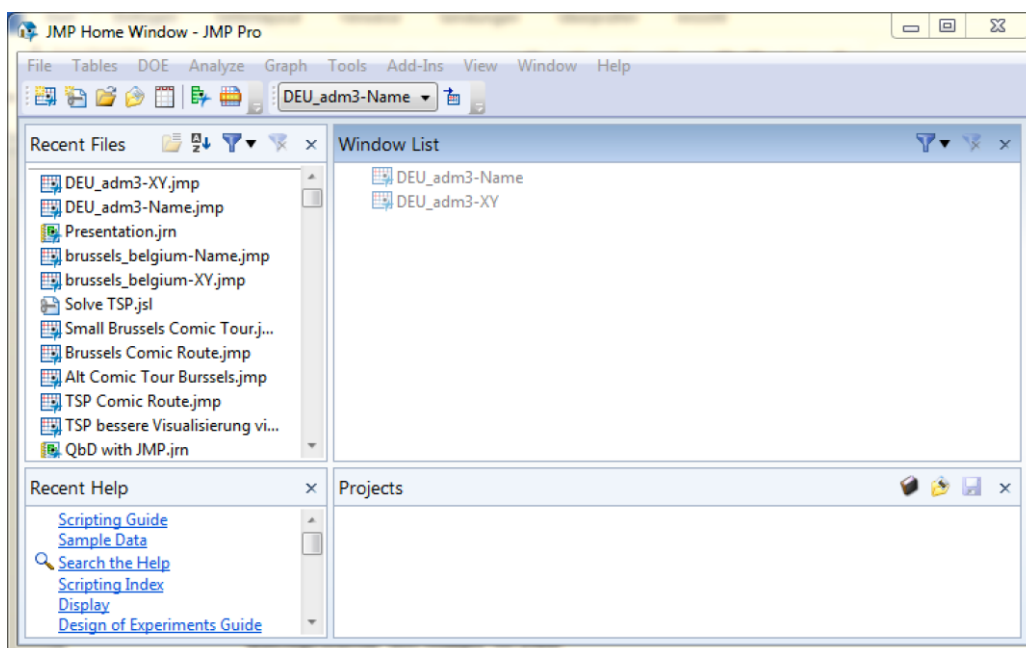


In the dialog select the column containing the names of the different shapes of your map as **Shape Names**. In this example the column containing the shape names is NAME\_3.

**Special Tip**

JMP does not need an individual name for each shape. Actually no shape names are required at all. Of course if no shape-names are provided the map cannot be used in the typical way, as you will not be able to give a colour to shapes that have no name in the graph builder. The map might still be very useful as a background map.

The text-field Map Name allows you to specify the name of the map. This will name will appear in all JMP dialogs using maps after the import. After selecting the right settings confirm with **Ok**.



The add-in will do all the other work for you and create the required data sets. These datasets are invisible, but they might be opened using double-click on them in the main window. To validate that the import worked correctly open the **DEU\_adm3-Name** file. Use the graph builder and drag-and-drop the **name\_3** column to the **Map Shape** area of the graph builder.

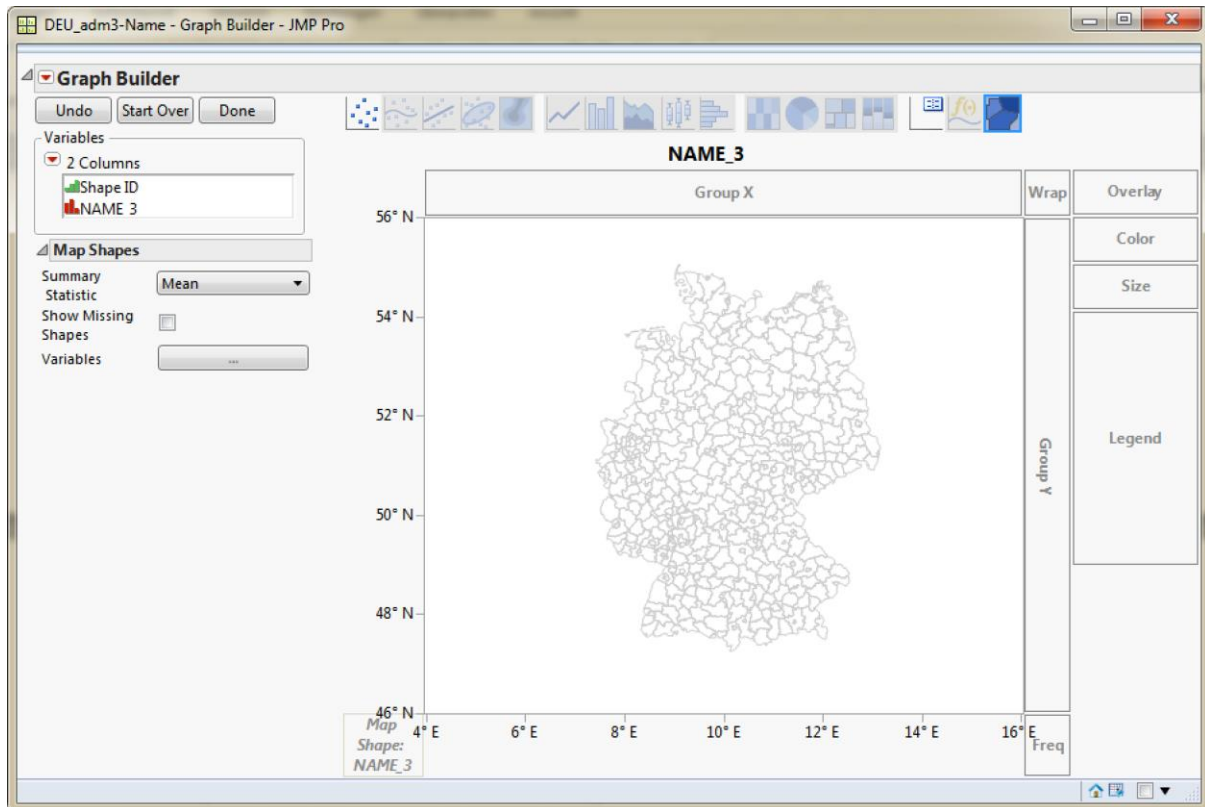


Figure 5 A new map is available in JMP

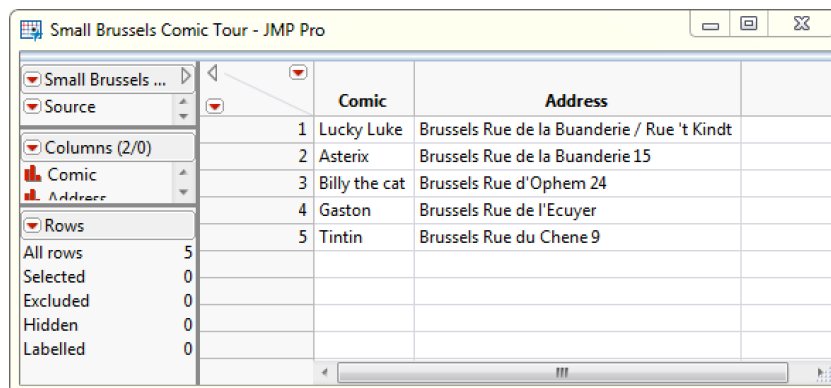
The add-in will import foreign shape files and do the data management to get the desired properties to use them in JMP. Of course there can't be a guarantee that it works due to the many different shape-file-formats and geo-coordinate-systems. If the import of a file does not work, please ask in the JMP community forums for help or send an email to [sebastian.hoffmeister@statcon.de](mailto:sebastian.hoffmeister@statcon.de).




## Solving TSPs

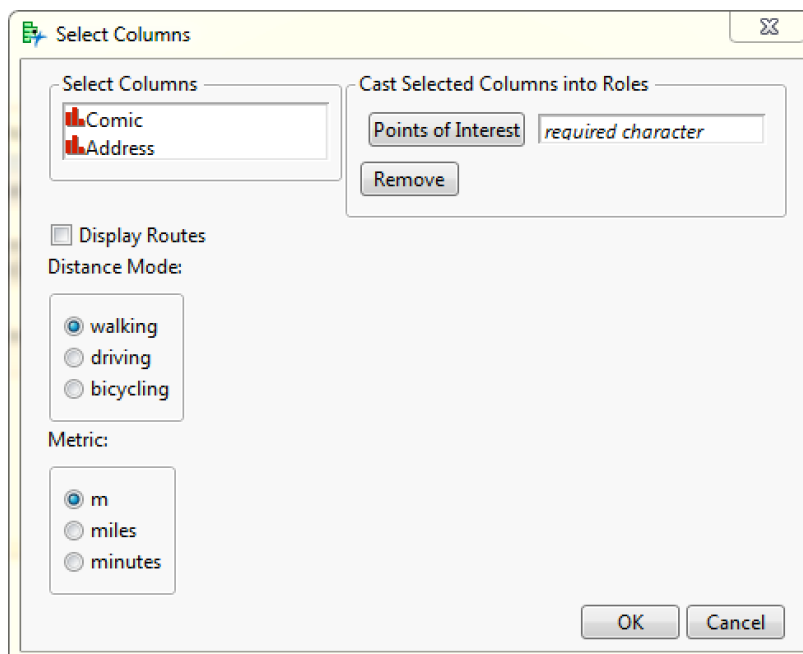
A travelling salesperson problem is the problem of trying to visit a number of  $k$  given cities in the shortest possible time/distance. While this task might sound simple it is a challenging problem as one has to calculate all possible routes to be sure to find the fastest route. While it might be possible to calculate all routes for a relatively small number of cities the problem becomes very computer intensive for larger  $k$ . The R-package TSP provides a couple of heuristic algorithms to solve this problem in reasonable time. This add-in uses the algorithms of the R-package and presents the solution in an interactive way in JMP.

Let's continue with the previous example of the [Brussel's Comic tour](#). The problem is in which order we want to visit the five selected sights in Brussels without wasting any time?



	Comic	Address
1	Lucky Luke	Brussels Rue de la Buanderie / Rue 't Kindt
2	Asterix	Brussels Rue de la Buanderie 15
3	Billy the cat	Brussels Rue d'Ophem 24
4	Gaston	Brussels Rue de l'Ecuyer
5	Tintin	Brussels Rue du Chene 9

The add-in provides a new function called TSP to solve this problem for you. Go to the **main menu** → **Add-Ins** → **Spatial Data Analysis** →  **TSP**.



Just add the column containing the addresses to the **Points of Interest**. You might actually get different results when optimizing for times (minutes) compared to distances (meters, miles) as some longer routes might still be faster and vice versa.

If you are only interested in the right order in which you should visit the **Points of Interest** do not use the checkbox **Display Routes**. Display routes will provide a data set with the exact route that is chosen by google maps. Of course getting that data might take some time depending on the number of Points of Interest.

	Comic	Address	lat	lon	Order of Visits
1	Gaston	Brussels Rue de l'Ecuyer	50.849° N	4.35424° E	1
2	Billy the cat	Brussels Rue d'Ophem 24	50.8535° N	4.34494° E	2
3	Asterix	Brussels Rue de la Buanderie 15	50.8466° N	4.34222° E	3
4	Lucky Luke	Brussels Rue de la Buanderie / Rue 't Kindt	50.8468° N	4.34174° E	4
5	Tintin	Brussels Rue du Chene 9	50.8446° N	4.35011° E	5

Figure 6 TSP Solution in the Last Column

If you are interested in a precise routing select the checkbox display routes and you will get an additional dataset containing the exact routing.

	leg	Label	lon	Label 2	lat
1	1	startLon	4.3542432° E	startLat	50.848997° N
2	1	endLon	4.3529252° E	endLat	50.849506° N
3	2	startLon	4.3529252° E	startLat	50.849506° N
4	2	endLon	4.3503548° E	endLat	50.850907° N
5	3	startLon	4.3503548° E	startLat	50.850907° N
6	3	endLon	4.3474243° E	endLat	50.850317° N
7	4	startLon	4.3474243° E	startLat	50.850317° N
8	4	endLon	4.3471621° E	endLat	50.850568° N
9	5	startLon	4.3471621° E	startLat	50.850568° N
10	5	endLon	4.34513° E	endLat	50.85291° N
11	6	startLon	4.34513° E	startLat	50.85291° N
12	6	endLon	4.3450837° E	endLat	50.853549° N
13	1	startLon	4.3450837° E	startLat	50.853549° N
14	1	endLon	4.34513° E	endLat	50.85291° N
15	2	startLon	4.34513° E	startLat	50.85291° N
16	2	endLon	4.3455428° E	endLat	50.852824° N
17	3	startLon	4.3455428° E	startLat	50.852824° N
18	3	endLon	4.3453139° E	endLat	50.85255° N
19	4	startLon	4.3453139° E	startLat	50.85255° N

The attached script "Graph Builder" will visualize the route chosen by the google maps API.



## Literature and Links

1. **Brussels Comic Route.** URL: [http://en.wikipedia.org/wiki/Brussels'\\_Comic\\_Book\\_Route](http://en.wikipedia.org/wiki/Brussels'_Comic_Book_Route)
2. **David Kahle and Hadley Wickham (2013).** *ggmap: A package for spatial visualization with Google Maps and OpenStreetMap.* R package version 2.3. URL: <http://CRAN.R-project.org/package=ggmap>
3. **Diva-gis: Free Spatial Data,** URL: <http://www.diva-gis.org>
4. **Mapzen: An open-source mapping lab.** URL: <https://mapzen.com>
5. **Michael Hahsler and Kurt Hornik (2015).** *TSP: Traveling Salesperson Problem(TSP).* R package version 1.0-10. <http://CRAN.R-project.org/package=TSP>
6. **Michael Hahsler, and Kurt Hornik (2007),** *TSP - Infrastructure for the traveling salesperson problem.* Journal of Statistical Software 23/2. URL: <http://www.jstatsoft.org/v23/i02/>.
7. **R Core Team (2014).** *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

## Known Issues and Planned Improvements

1. Currently the add-in is not tested on mac.
2. When importing shape files they are currently imported the expected windows-path. Actually it should probably check the JMP-version and OS to make sure that the right path is used.
3. Some shapefiles available on the net **don't have closed shapes**. This happens when the list of coordinates specifying a shape does not start and end with the same coordinates. To use maps in JMP closed shape-files are required. Currently there is no efficient way to correct shape files with open shapes.
4. There are **multiple standards for geo-coordinates**. This is a problem when importing shape files. If the shapesfiles to be imported use a different coordinate system than JMP expects the resulting map will not work in JMP. Currently there is no solution in the add-in to detect (and warn) if the coordinates are not ok. Neither is there a solution to transform coordinates into the correct coordinate system.
5. In future versions an animated way to display the route of a TSP might be included (e.g. via the bubble plot).
6. **The TSP solver sometimes does not find the optimal solution** even for simple problems - **Solution:** Currently the script runs multiple algorithms to solve the TSP and selects the shortest route. Running the script multiple times will help finding the overall best route. In future versions the algorithms will automatically be used multiple times to provide more stable results.
7. Currently it is not possible to select a **starting point for a route in the TSP**.

Please help us to develop the add-in and to set the right priorities! Send emails with feedback to [consult@statcon.de](mailto:consult@statcon.de) or [Sebastian.hoffmeister@statcon.de](mailto:Sebastian.hoffmeister@statcon.de).