

BioNumerics Tutorial:

Geographical mapping of data

1 Aim

In many research projects, especially epidemiological, biological data is closely linked to geographical data. Geographical information provided as city names, postal or zip codes, or geographical coordinates can be plotted from within BioNumerics on GoogleTM Maps. The data can be plotted individually or as stacked bar graphs or pie charts, using different colors according to database groups. In this tutorial you will learn how to plot entries in your BioNumerics database and how to change the layout of the plots.

2 Preparing the database


The **DemoBase Connected** will be used in this tutorial and can be downloaded directly from the *BioNumerics Startup* window or restored from the back-up file available on our website:

1. To download the database directly from the *BioNumerics Startup* window, click the **Download example databases** link, located in the lower right corner of the *BioNumerics Startup* window. Select **DemoBase Connected** from the list and select **Database > Download**. Confirm the download action.
2. To restore the database from the back-up file, first download the file `DemoBase_Connected.bnbk` from <http://www.applied-maths.com/download/sample-data>, under 'DemoBase Connected'.

In the *BioNumerics Startup* window, press the  button, select **Restore database**, browse for the downloaded file and select **Create copy**. Specify a name and click **<OK>**.



In contrast to other browsers, some versions of Internet Explorer rename the `DemoBase_Connected.bnbk` database backup file into `DemoBase_Connected.zip`. If this happens, you should manually remove the `.zip` file extension and replace with `.bnbk`. A warning will appear ("If you change a file name extension, the file might become unusable."), but you can safely confirm this action. Keep in mind that Windows might not display the `.zip` file extension if the option "Hide extensions for known file types" is checked in your Windows folder options.

1. In the *BioNumerics Startup* window, double-click on the **DemoBase Connected** database to open it.
2. Select **File > Install / remove plugins...** () in the *Main* window to open the *Plugins* dialog box.
3. Select the *Geographical module* from the list in the *Utilities tab* and press the **<Activate>** button.
4. The program will ask to confirm the installation of the plugin. Confirm twice.

The plugin is now preceded with a green V-sign in the *Plugins* dialog box (see Figure 1).

5. Press **<Exit>** to close the *Plugins* dialog box.
6. Close and reopen the database with administrator privileges to activate the features of the *Geographical mapping plugin*.
7. Confirm the update of the database.

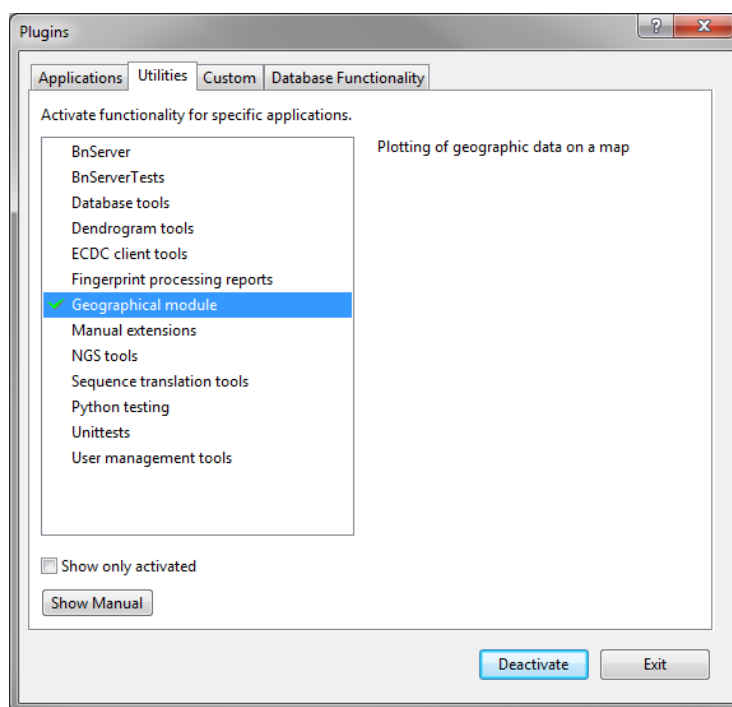



Figure 1: Installing the Geographical plugin.

3 Importing geographical data

As an example, we will import geographical data from a tab-delimited text file in this database. The text file can be downloaded from our website: go to <http://www.applied-maths.com/download/sample-data> and click on "Geo data file".

1. In the *Main* window, select **File > Import...** (, **Ctrl+I**).
2. Select **Import fields (text file)** under **Entry information data** and press **<Import>**.
3. Browse for the `Geo_information.txt` text file and press **<Next>** to continue.

The next dialog lists the columns present in the selected text file as **Source**, their linked **Source type** and the **Destination** component they are associated with (currently all set to **<None>**).

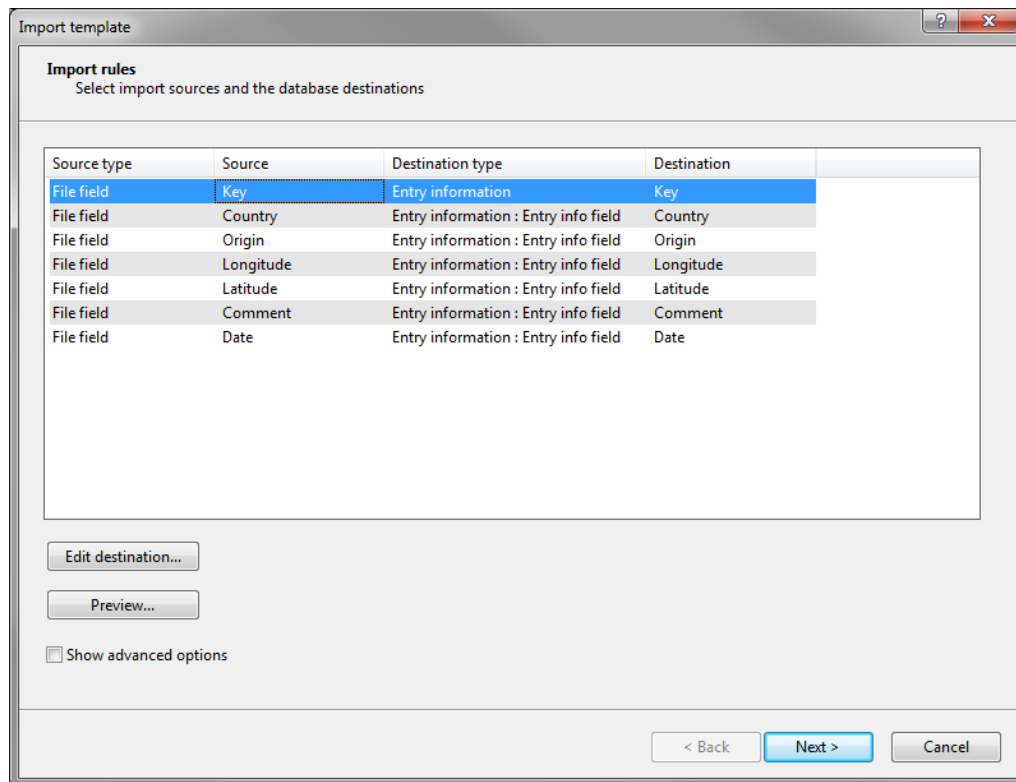
4. Select **Key** from the list. Select **<Edit destination>** and select **Key** as the BioNumerics corresponding field. Press **<OK>**.
5. Select all the other rows (Country, Origin, Longitude, Latitude, Comment and Date) using the **Shift-key**. Select **<Edit destination>**, select **Entry info field** from the list and press **<OK>**. Confirm the suggested information field names and the creation of these fields in the database.

The grid is updated (see Figure 2).

6. Press **<Next>** to proceed with the import.
7. Specify that the **Key** information should be used as linked information field and press **<Finish>**.
8. Enter a template name (e.g. "Import geo information") and press **<OK>**.

The import template is now listed in the *Import template* wizard page and is automatically selected.

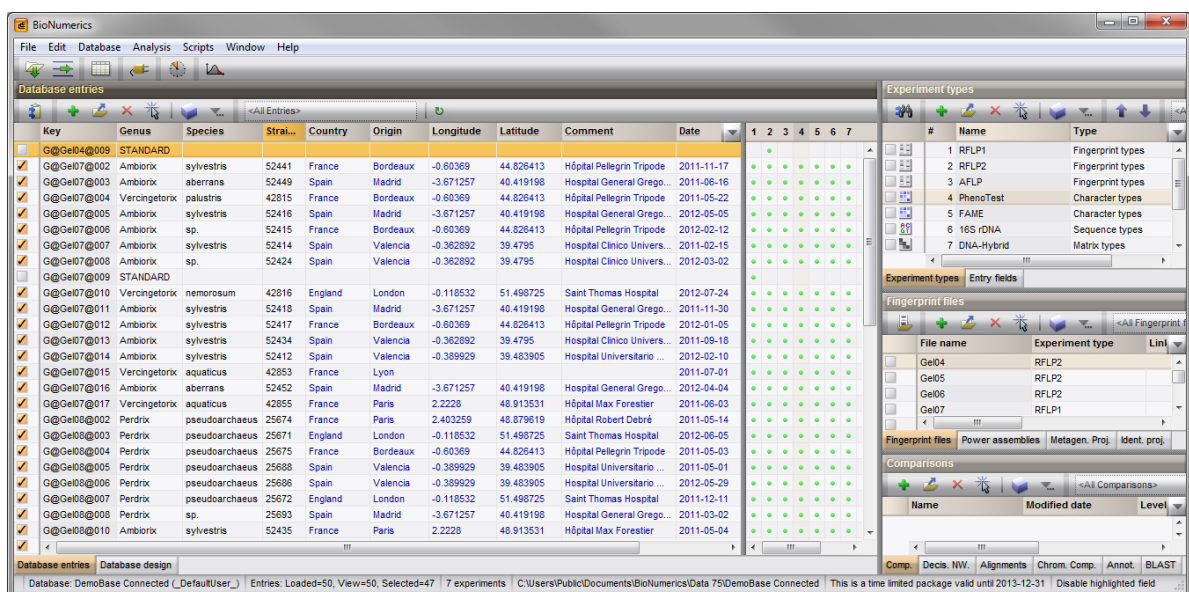
9. Press **<Next>** to proceed with the import.

Figure 2: The *Import rules* dialog box.

From the database links that are updated, verify that the geographical information will be updated for 47 entries.

10. Press <*Finish*> to import the data.

For all entries, except the ones defined as STANDARD, the database field information for the columns *Country*, *Origin*, *Longitude*, *Latitude*, *Comment* and *Date* is now updated (see Figure 3).

Figure 3: The *Main* window after import of the data.

Both the imported *Longitude* and *Latitude* columns contain numeric information. To allow geographical

binning on this information, the field type of these columns should be updated:

11. Right-click on the entry field header **Longitude** and select **Field properties** from the pop-up menu that appears. This opens the *Database field properties* dialog box. Change the *Content type* to *Number* and press <OK> to close the dialog.
12. Repeat the instructions above to change the content type of the **Latitude** field to *Number*.

Geographical binning can also be applied to time information i.e. date information. To use this feature, the field properties for the **Date** field should be changed:

13. Right-click the entry field header **Date**, select **Field properties** from the pop-up menu, and change the *Content type* in the the *Database field properties* dialog box to *Date*. Press <OK> to close the dialog.

In BioNumerics it is possible to add an information field to the database containing the information derived from two or more other information fields. As an example we will create a new field containing the location information from the **Origin** and **Country** fields.

14. In the *Main* window, select **Edit > Information fields > Add information field...** or highlight the *Entry fields* panel and select **Edit > Create new object...** (+).
15. In the *Create new entry information field* dialog box that pops up, enter e.g. “Address” as **Name** of the new information field.
16. Check the option **Calculate field content from other fields** and press the <Edit> button to display the *Calculated database field settings* dialog box.
17. Press <Add information field...>.
18. In the *Add information field* dialog box, select the **Origin** field from the list and press <OK>.

Back in the *Calculated database field settings* dialog box, [ORIGIN] is now displayed as part of the “formula” to calculate the field content.

19. Press <Add fixed text...>.
20. In the *Add fixed text* dialog box, enter “,” (i.e. a comma, followed by a space) and press <OK>.
21. Press <Add information field...> again.
22. In the *Add information field* dialog box, select the **Country** field from the list and press <OK>.

The *Calculated database field settings* dialog box now looks like in Figure 4.

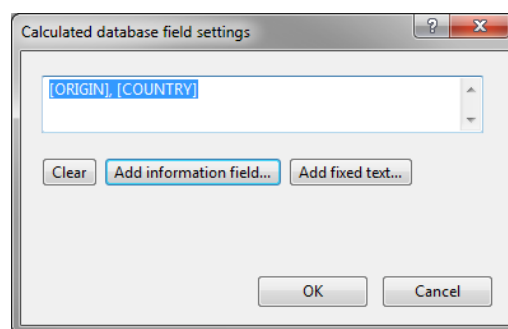


Figure 4: The *Calculated database field settings* dialog box, showing the calculated field settings for the Address field.

23. Press <OK> to accept the calculated field settings and press <OK> once more in the *Create new entry information field* dialog box.

The new calculated field **Address** now contains the information from the **Origin** and the **Country** fields, separated with a comma and a space (e.g. "Bordeaux, France").

4 Plot in geographical window

1. Select **Edit > Select all (Ctrl+A)** to select all entries and unselect the entries defined as STANDARD.
2. Select **Analysis > Geo chart...** (🌐).
3. Leave the **Pies** selected as graph type to plot the data in circular pie charts.

In the **DemoBase connected** coordinate information is available for most, but not all entries.

4. Select both the **Text address description** and the **Latitude and longitude** information to be used as location data (see Figure 5).

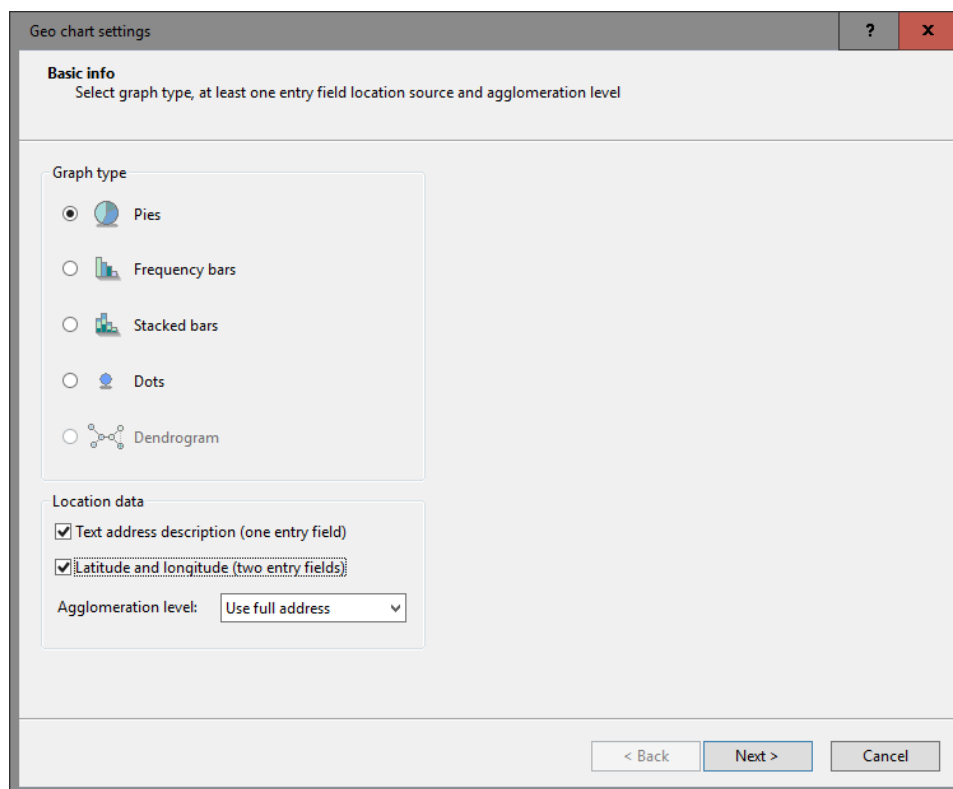


Figure 5: Geo chart settings: step 1.

Selecting these two options the entries will be mapped based on the provided coordinate information. If no coordinate information is available, the text address descriptions (e.g. city and/or country names) will be used.

5. Press **<Next>** to proceed.

The **Full address field** refers to the database information field that contains location information. In our database, city and country information is present in the information field **Address** (see 3).

6. Select **Address** from the list of **Full address fields**. This information will be used for the mapping of entries that do not have coordinate values.
7. Press **<Next>** to proceed to the next step.

8. Select the *Latitude* and *Longitude* fields from the list and proceed to the next step (see Figure 6).

Figure 6: Geo chart settings: step 3.

9. Select *Genus* as *Categories field* and press <Next> to go to the final step.

10. Specify “Europe” as hint to narrow down the search and press <Finish> (see Figure 7).

The entries are shown in pie charts and are colored based on the *Genus* groups (see Figure 8). The groups and group colors are displayed in the *Categories panel*. The information that is used to map the entries is displayed in the *Locations panel*. For three entries the *Address* information (**Lyon, France** in this case) is used since no latitude and longitude information is available for these entries. All other selected entries are mapped based on their coordinates.

11. Use the scroll wheel of the mouse to zoom in and out on the position of the mouse pointer.
12. To enlarge or shrink the bar graphs or pies, use *View > Scale up* (🔍) or *View > Scale down* (🔍), respectively.
13. Select the *Map* view, by clicking on the *Map* option in the right upper part of the *Map panel*.

This map view displays all geographical location information.

14. Zoom in on the city Valencia.

Because two different locations in Valencia are stored in the information fields *Latitude* and *Longitude* in the database, two pies appear in the city Valencia, corresponding to two different hospital locations (see Figure 9).

15. Select *Edit > Modify settings* (⚙️).
16. Select **Stacked bars** as graph type to plot the data in vertical bars.

Geo chart settings ? X

Geolocation hint and language
Give a geocoding hint and select the address language.

Location hint
Optional text to be appended to the full address, to narrow down the location and reduce the number of geocoding results (e.g. 'Europe')

Europe

Language
Address language: English

< Back Finish Cancel

Figure 7: Geo chart settings: step 5.

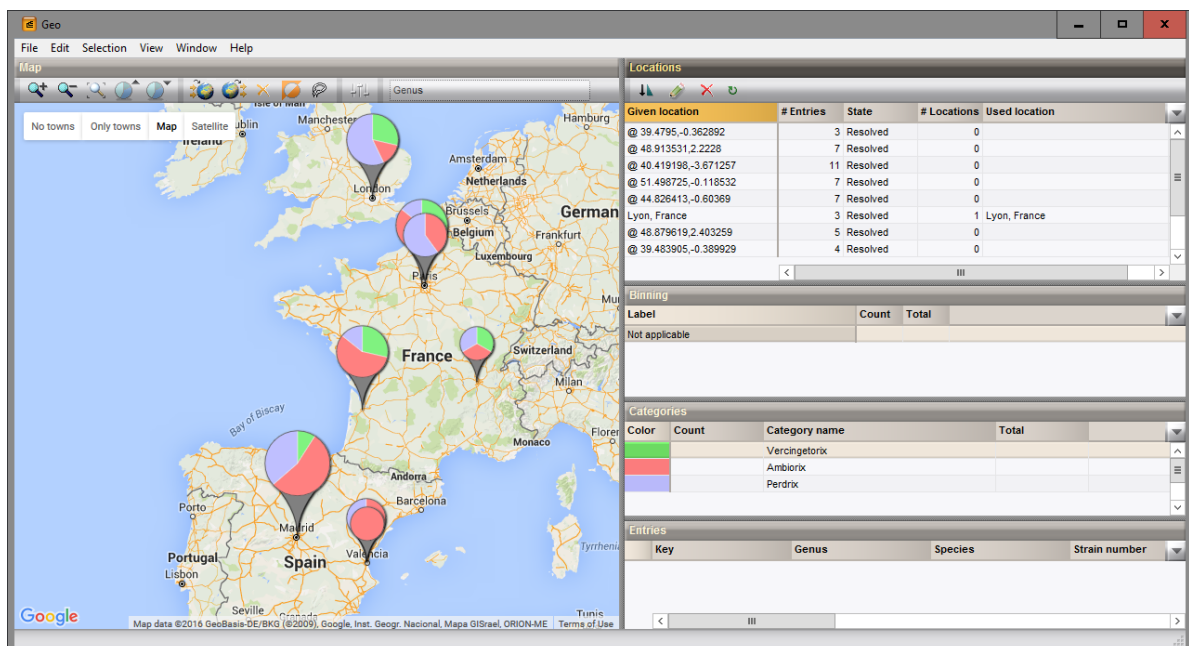


Figure 8: The Geo window, displaying the information as pie charts.

17. Make sure only the option *Text address description* is selected and press <Next>.
18. Select **Address** from the list of *Full address fields* and press <Next>.
19. Select *Species* as category field and press <Next>.

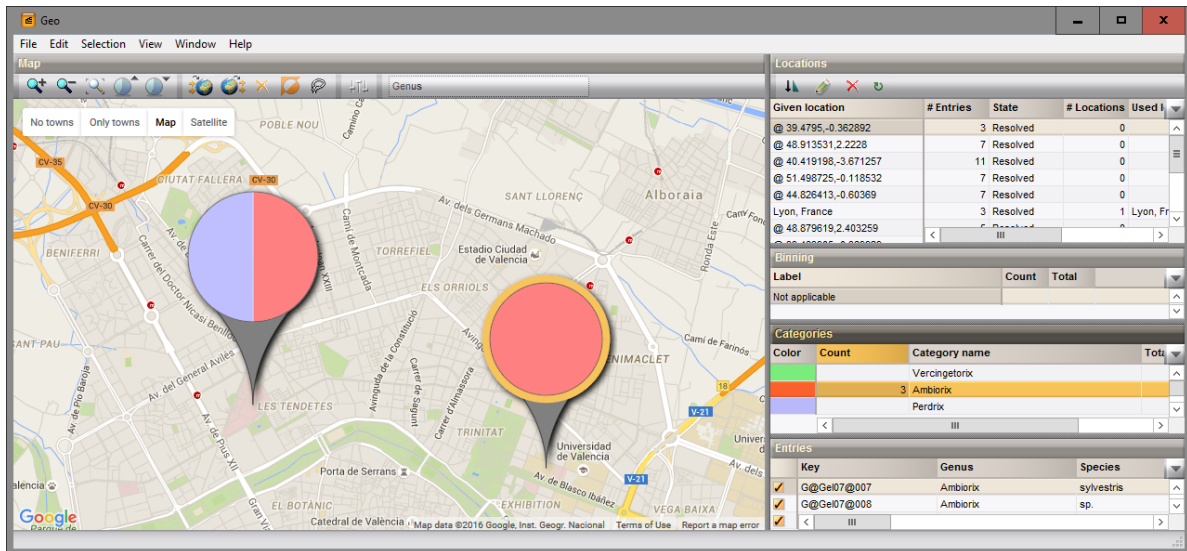


Figure 9: Two different location in the city Valencia.

Since the *Stacked bars* option was selected in the first step, a binning can be specified in the next step:

20. Select *Date* as *Binning field* and *Quarter* as grouping (see Figure 10). Press <Next>.

The screenshot shows the 'Geo chart settings' dialog box. The 'Binning field' section is active, showing a dropdown menu set to 'Date'. Below it, the 'Date grouping Span' is set to 'Quarter'. The dialog has 'Back', 'Next >', and 'Cancel' buttons at the bottom.

Figure 10: Binning data.

21. Specify "Europe" as hint to narrow down the search and press <Finish>.

The entries are shown in bars and are binned based on the date which is set to "Quarter". The bars are colored based on the *Species* groups and the height of the bars is proportional to the number of entries that have an identical location and bin (see Figure 11).

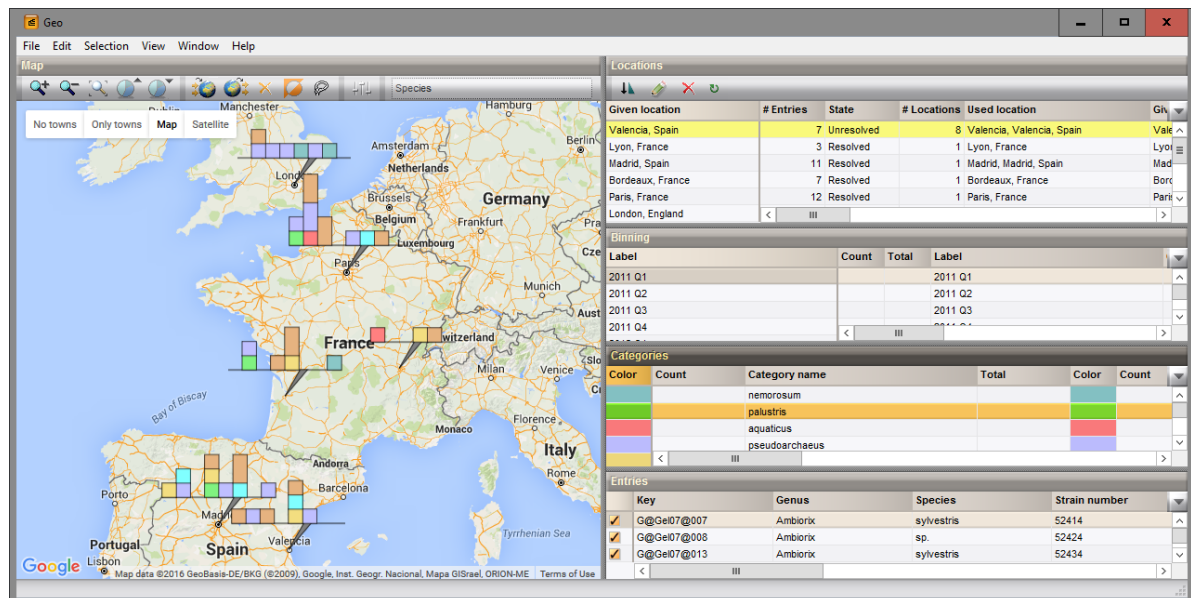


Figure 11: The Geo window.