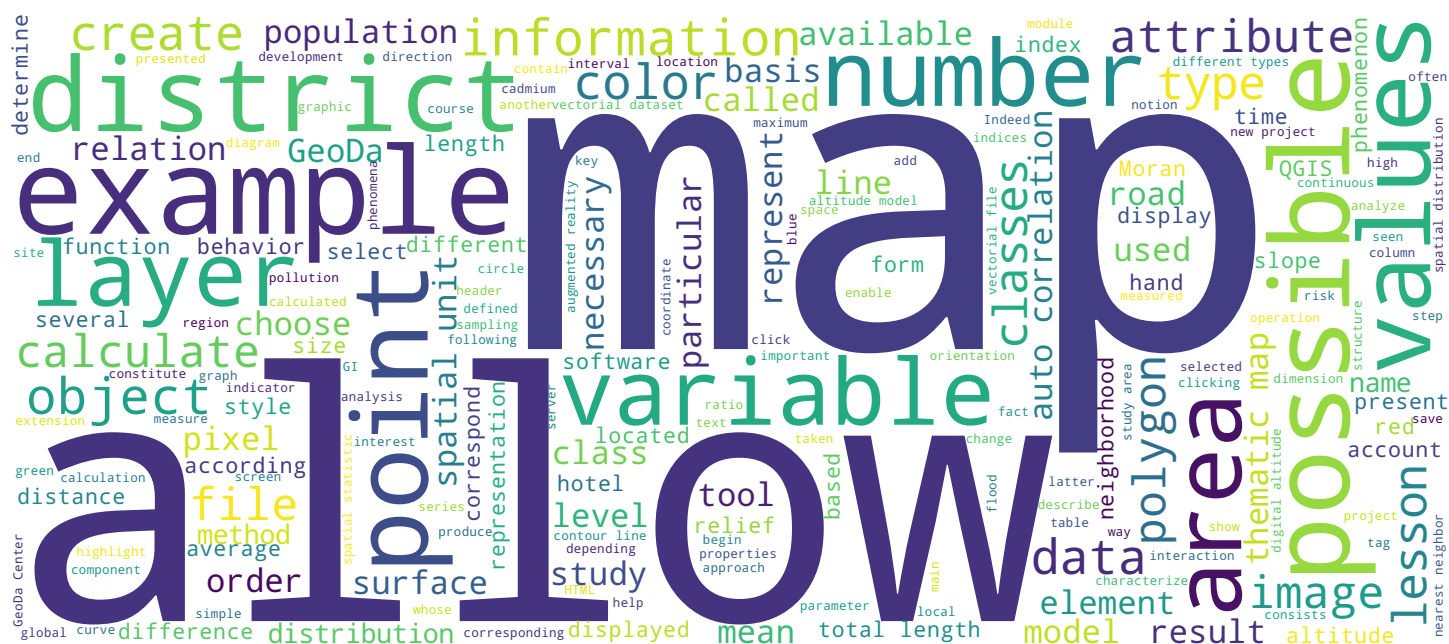




Analysis of Geographic Information: GeoDa

Stéphane Joost, Marc Soutter, Fernand Kouamé, Amadou Sall



Search MOOC



Video



Lesson Objectives

- Introduce GeoDa functions for:
 - Creating thematic maps
 - Dynamic and interactive exploratory analysis
 - Calculating spatial autocorrelation

An Introduction to Geographic Information Systems

The objective of this lesson is to present globally the main functions of the GeoDa open-source software which allows to map vectorial datasets and to analyze their attributes dynamically and interactively. After this lesson, you will be able to create a GeoDa project, import data and make thematic maps as well as linked dynamic graphics.

Notes

Summary



- Free and open-source program
- Prof Luc Anselin
- Exploratory analysis of geographical vector data
 - Thematic maps
 - Graphs and statistical diagrams
 - Spatial statistics
- Free open-source download :
<https://geodacenter.asu.edu/software/downloads>
(version 1.6.7)
- References:
 - <https://geodacenter.asu.edu/software/documentation>
 - Anselin, L., I. Syabri and Y. Kho (2005) GeoDa : An Introduction to Spatial Data Analysis. *Geographical Analysis* 38(1), 5-22

Brushing



An Introduction to Geographic Information Systems

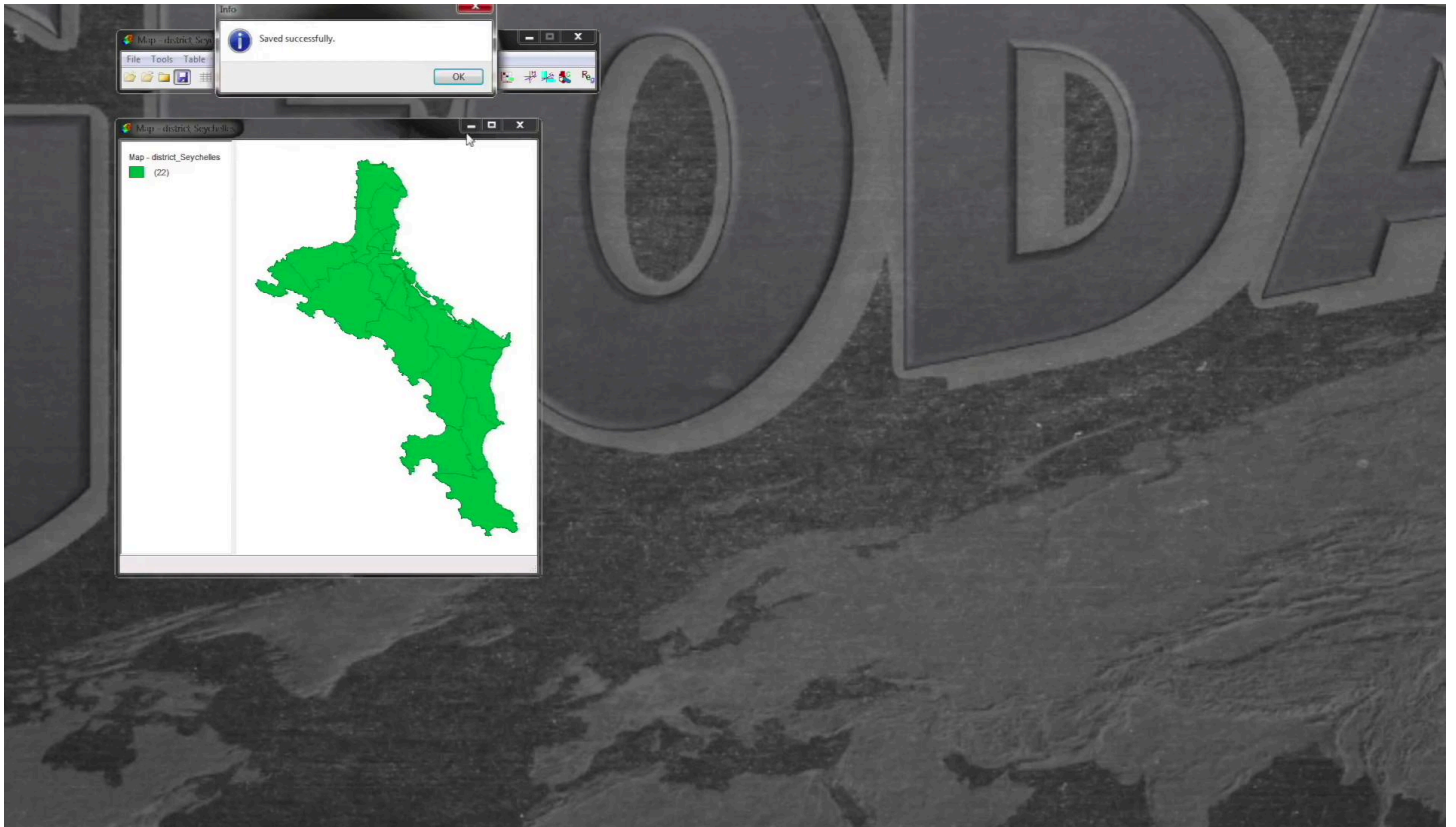
We will give you first some general information about GeoDa before explaining how to create a project and manage data. You will then learn how to make thematic maps before discussing the development of different types of graphics linked to the maps. The last part presents an introduction to the spatial statistics functions. GeoDa is a free software, developed initially by the Spatial Analysis Laboratory of the University of Illinois in the United States under the direction of Professor Luc Anselin. The software followed Luc Anselin at the Arizona State University and it is currently maintained and distributed by the GeoDa Center for spatial analysis. It is an exploratory analysis software which allows to dynamically analyze vectorial datasets by combining thematic maps, diagrams and spatial statistics. This software can be downloaded freely and free of charge on the GeoDa Center website. We specify that the course has been developed on the basis of the latest stable version, so the 1.6.7 version, it is available in English only. The main references are the website of the GeoDa Center on one hand where the documentation is available and the 2005 article in English which describes the software, but also the theoretical principles at its origin.

Notes

Summary



0m 47s



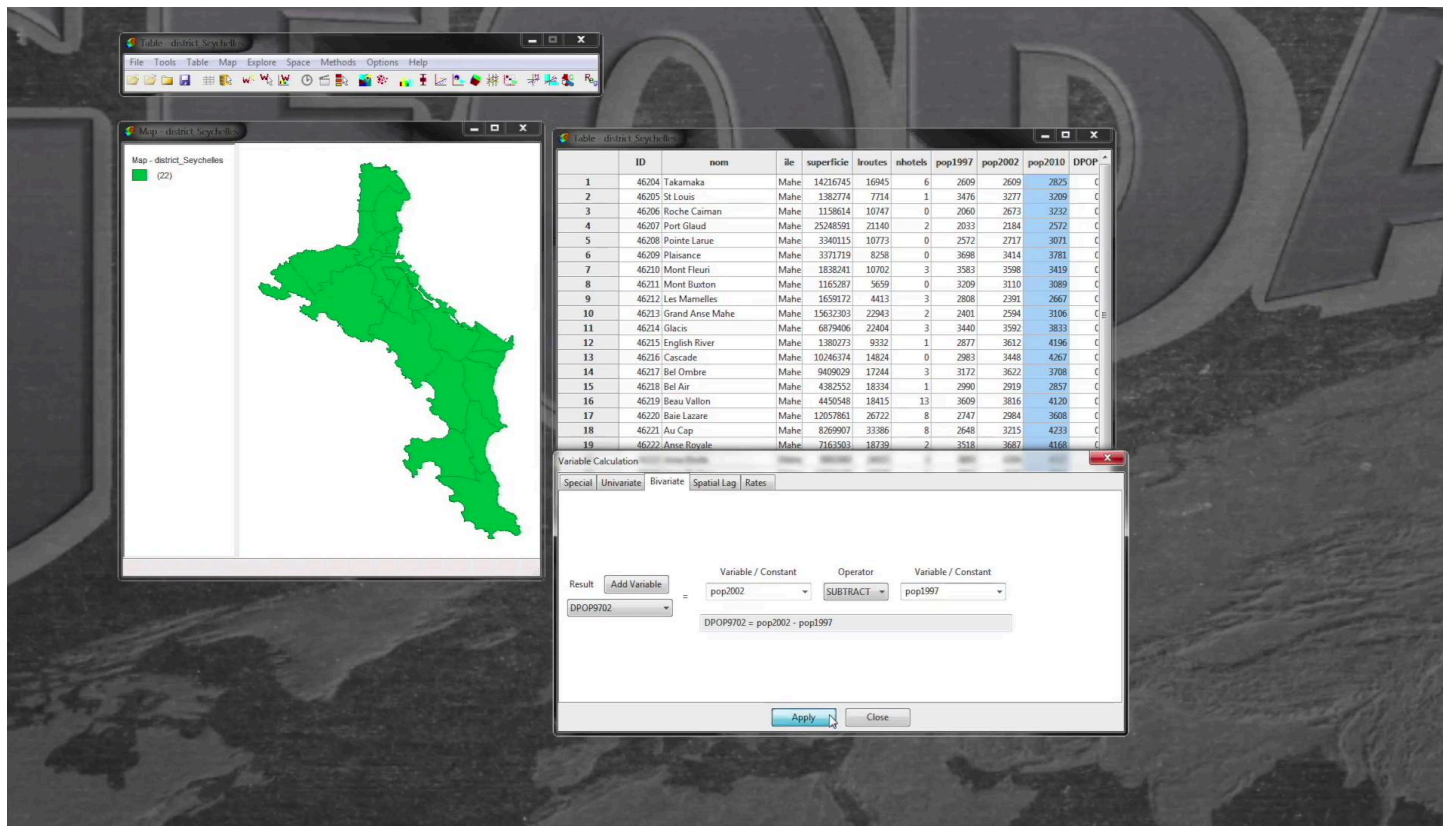
The GeoDa interface is very simple, it consists of a series of drop-down menus and a toolbar. We will begin by creating a new project in GeoDa and importing the vectorial dataset of the Seychelles' districts. To create a new project, we have to go to the File menu, and choose New Project. There are then different ways to import data into GeoDa. From the File submenu, we can import different file formats. All the standard formats of georeferenced vectorial files such as Shapefile, the JSON or the KML are available, but also table formats without spatial dimension such as DBF, CSV or Excel files, for example. It is also possible to connect directly to an Oracle, Esri, PostGIS or MySQL database. Finally, it is also possible to connect to a web server using a specifically formatted url to reach a Web Feature Service or WFS. In this example, we will import a Shape file. On the File tab, under Input File, we choose the Esri Shapefile format, and then we select the desired file. In our case, this is data on the districts of the Seychelles Islands. The map is displayed and from that moment, it is possible to save the project by clicking on the icon representing a floppy disk in the toolbar.

Notes

Summary



2m 11s

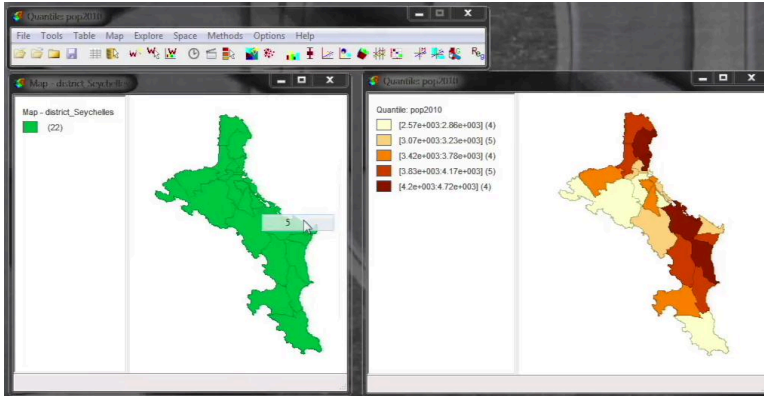


Note that in the Options menu, different zoom and selection tools are available. On the General toolbar, the first icons are the shortcuts allowing to create a project, close or save it. The fourth icon allows to open the attribute table associated with the vectorial file. For each district of Mahe Island, The available variables are the area, the total length of the roads crossing it, the number of hotels and the population in 1997, 2002 and 2010. When an attribute is selected in the table, it is directly selected on the map. We can also select spatial units according the values of an attribute. In the Table menu, go under Selection Tool, we can thus, for example, select all the districts whose population in 2010 was less than 3,000 inhabitants or carry out the reverse selection. It is also possible to add columns or perform calculations on attributes using the Variable Calculation tool in the Table menu. It is thus possible, for example, to calculate in a new column the difference in population between 1997 and 2002. To do this, under the Bivariate tab, we define a new variable, dpop9702, and we write the corresponding formula, so the subtraction of pop1997 to pop2002.

Notes

Summary





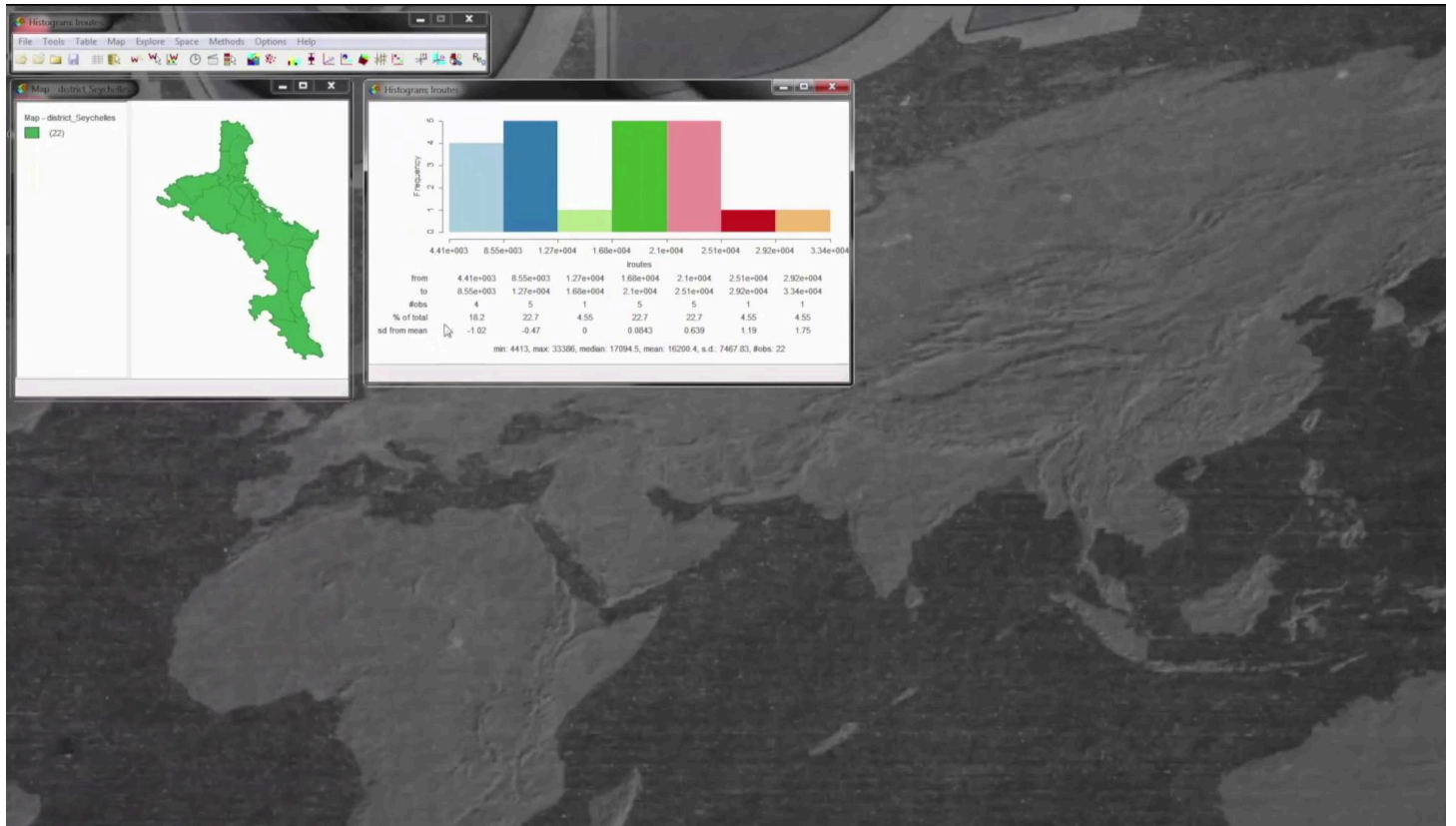
The resulting column is automatically added to the table. It is still possible to sort the values in ascending or descending order by double-clicking on the header of the column. A new double-click reverses the sorting order. This allows to identify the districts which have undergone the greatest loss of population or on the contrary the greatest growth between 1997 and 2002. Once the data are imported, it is possible to create thematic maps. This is the main function offered by GeoDa in order to explore a dataset and to analyze the spatial distribution of the phenomena studied. Different types of thematic maps can be quickly created in GeoDa. They correspond to several methods of discretization which will be presented in module four. The different options are available on the Map tab. It is possible, for example, to create a map in five quantiles on the 2010 population. We can see that the result contains five classes, with about the same number of districts per class. By clicking on one of the classes, all the districts of the latter are selected on the map. It is possible to change the color of a class by right-clicking on the class then selecting the Color For Category element.

Notes

Summary



5m 30s

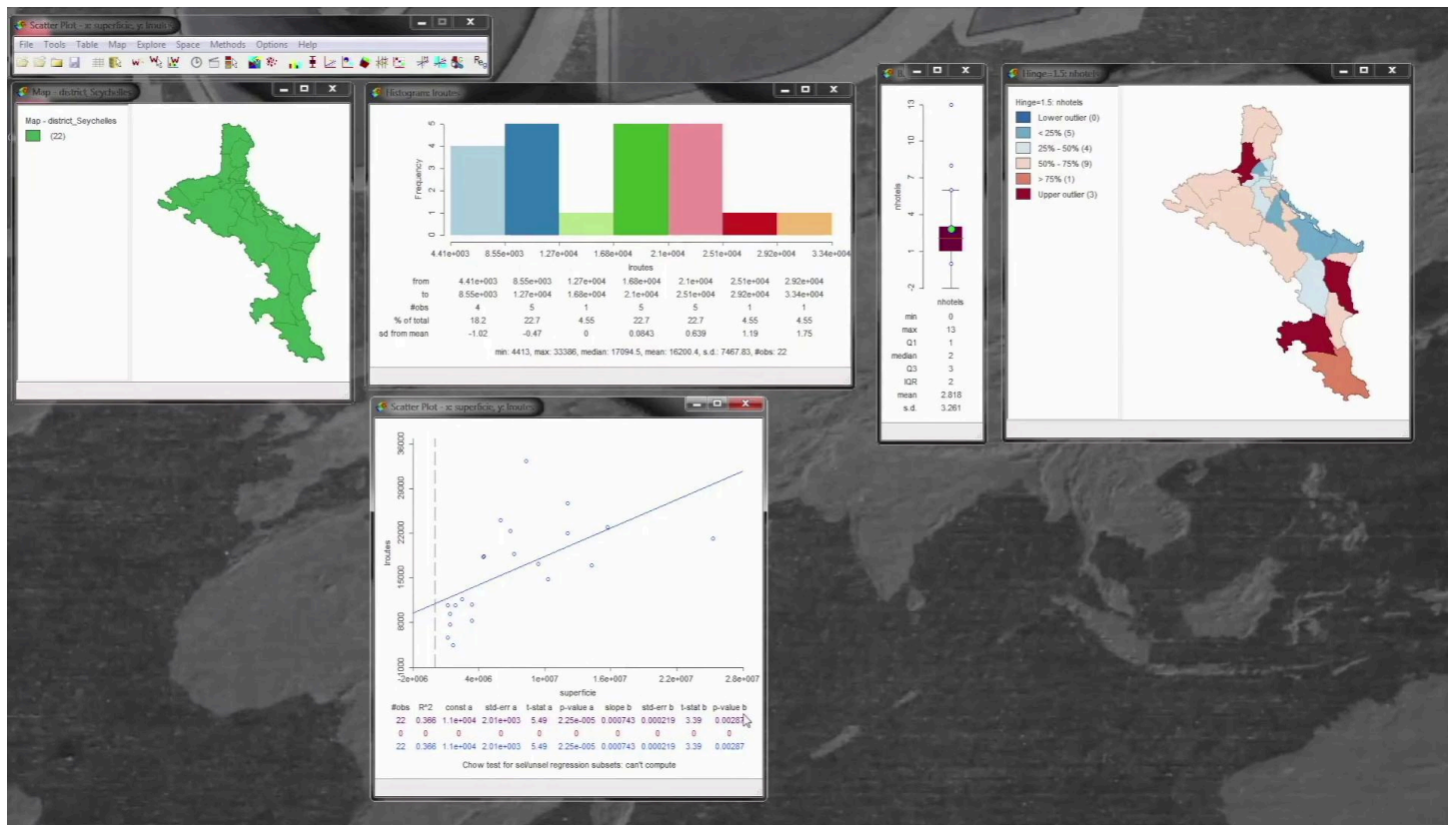


It is possible to choose other classifying methods such as natural breaks or equal intervals. It is also possible to highlight the distribution of values around the mean, for example by discretizing the data using the percentile method or the one based on the standard deviation which allows to characterize the behavior of the spatial units analyzed. Finally, the Cartogram tool allows to create bivariate maps by means of proportional circles. Here, for example, the circles are proportional to the population and the color to the fraction of hotels existing in the district. When a spatial unit is selected on a map, it is also automatically selected on all cartographic representations. GeoDa also offers the ability to view attributes through statistical graphs or diagrams. These tools allow for example to highlight spatial units which have one or more extreme values for a given attribute. It also allows to study the relation between different attributes previously selected. Different types of graphics can be created from the Explore tab. For example, we can create histograms which display the statistical distribution of an attribute as the length of roads here.

Notes

Summary



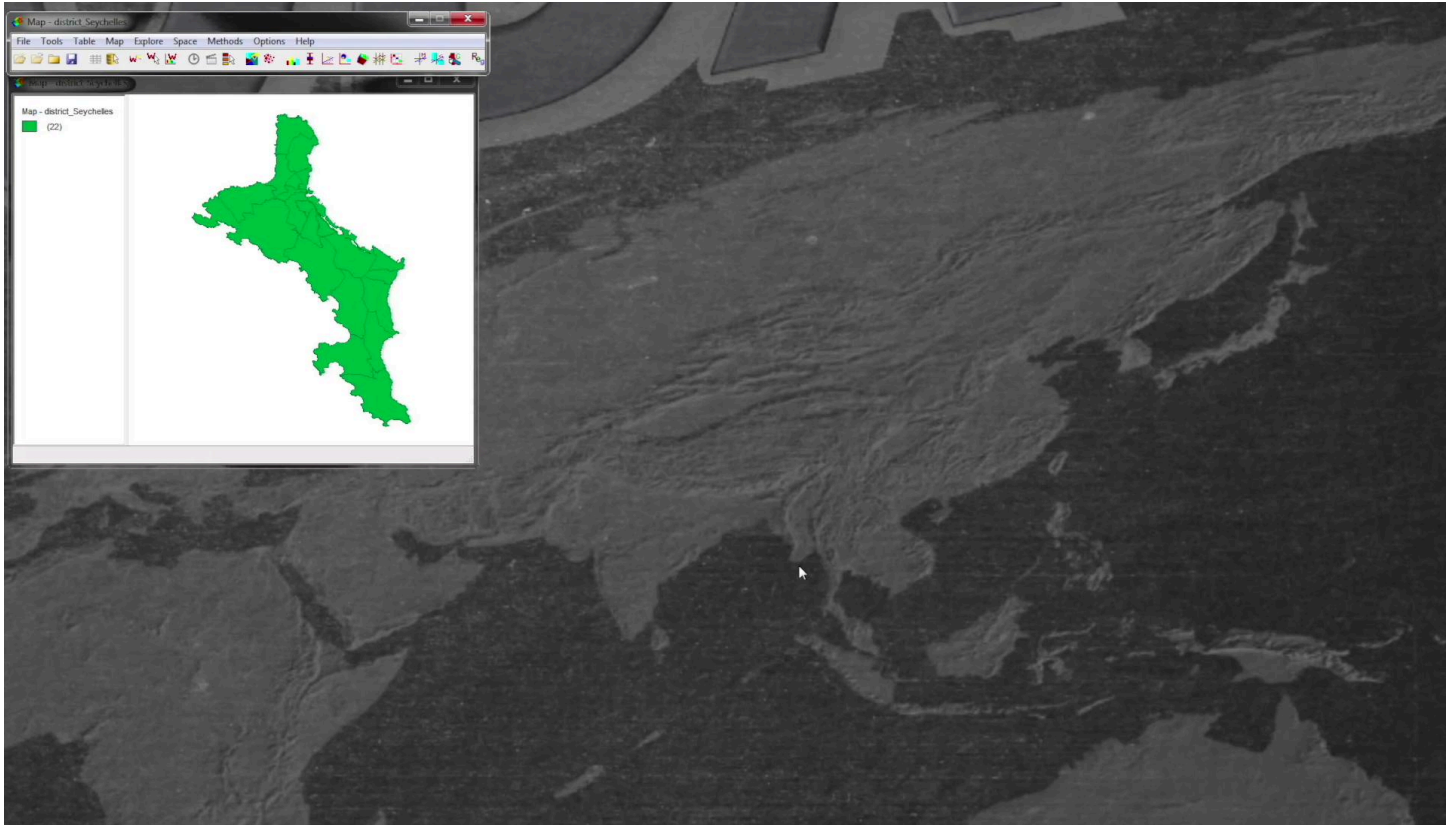


The statistical informations are automatically calculated for each class of the histogram. We can also generate box plots for example here on hotels by district. This allows us to highlight spatial units that have an atypical behavior. Through dynamic interaction or brushing, when we select a value on a map or a graph, the corresponding object is simultaneously selected on other types of representations, which allows to combine the different tools to analyze the data set. It is possible to create the corresponding box map from the Map tab. We can see that the classes of the box map correspond to the distribution presented in the box plot. Outliers, positives in particular, are easily identifiable. GeoDa also enables to study the relations between selected attributes using scatter plots. The example presented here is the relation between the area of the district and the total length of the roads crossing it. The regression parameters and the results of the statistical tests are directly displayed below the diagram. In that case, there is a positive and significant relation between the area of the districts and the total length of the roads.

Notes

Summary



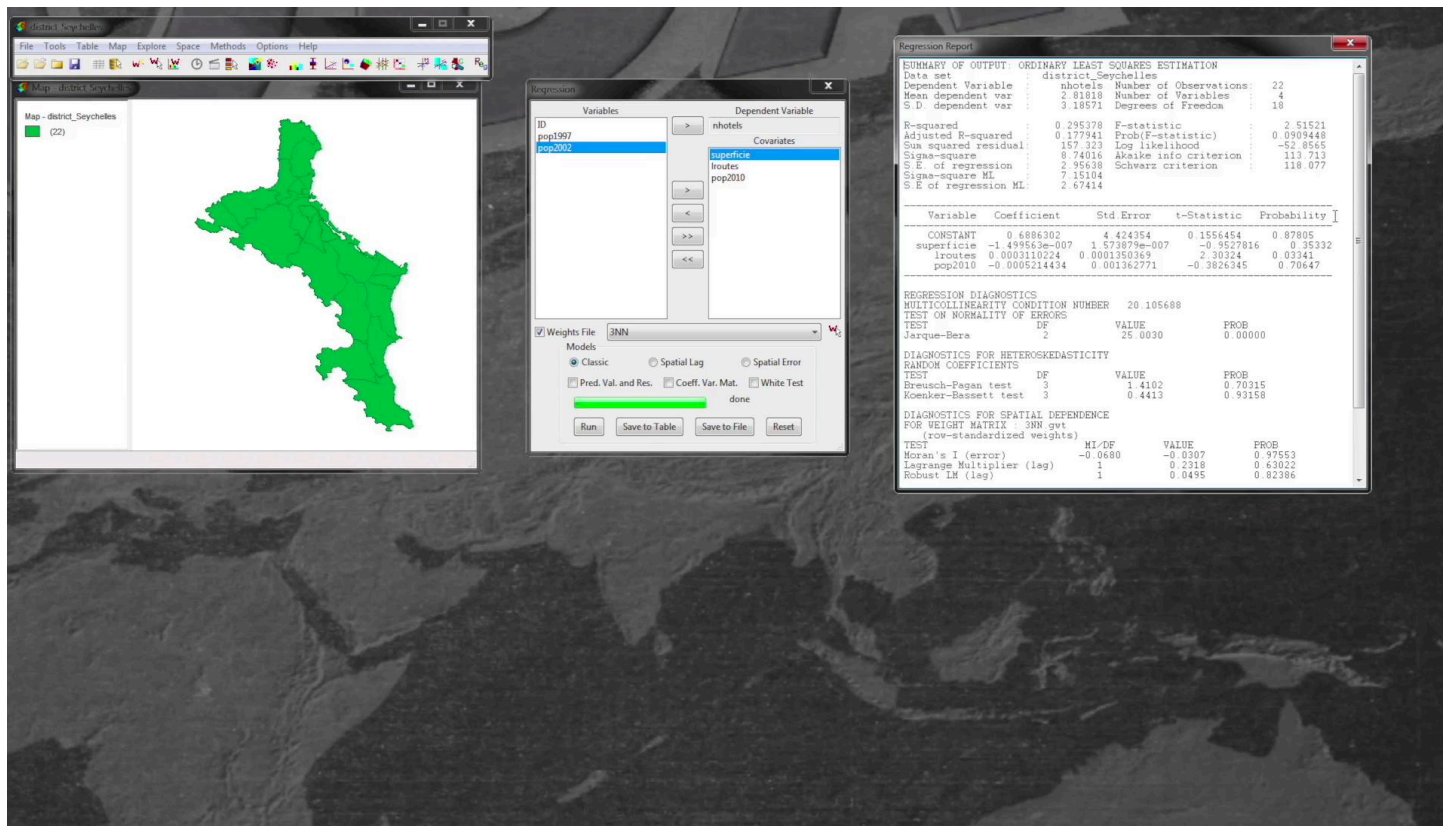


When we select only a subgroup of objects, the regression is calculated only on the basis of the selected points and it is displayed in red, the one obtained by excluding the selected points appears in blue. This allows to rapidly analyze the influence of a particular value on the relation. Beyond the dynamic exploratory cartography, GeoDa also includes spatial statistic functions. This is the calculation of the global and local auto-correlation as well as that of linear, spatial, united and multi-varied regressions. These notions will be taught in detail during module three. We will begin by illustrating how to calculate the global spatial auto-correlation on the variable of the total length of roads per district. Firstly, it is necessary to produce a weighting file which allows to take into account the values of the districts in a geographical neighborhood which we will specify later. To do this, in the Tools menu, choose Weights Manager and create a new weight file. For this, we must first indicate the column containing the unique identifier, and then choose a weighting diagram. We will consider here the three nearest neighbors. And on this basis, it is then possible to calculate the auto-correlation.

Notes

Summary





In the Space menu we choose the option Univariate Moran's I and we define the variable for which we want to calculate the auto-correlation on the basis of the values of the three nearest neighbors. For the length of roads, the overall Moran I is 0.23. To analyze the significance of Moran I, it is necessary to generate spatial configurations based on random permutations by right-clicking on the graph, then select Randomization, and then choose 999 permutations here for example. In our case, many spatial configurations produce higher Moran I than the observed situation highlighted by the yellow bar and whose Moran I is consequently not significant. It is also possible to calculate local or bivariate auto-correlation. The corresponding tools are in the same Space menu. Finally, GeoDa allows to calculate univariate and multi-varied spatial regressions with the help of the Regression tool in the Methods menu. For example, we can try to express the number of hotels depending on the population, on the total length of roads and on the area. The result gives us the regression parameters, the results of statistical tests and an estimate of the model error due to the spatial dependence between the samples.

Notes

Summary

