**Lab 2: ESDA in GeoDa**

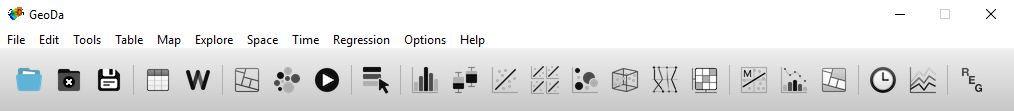
**Objectives**

* To conduct ESDA with Geoda.

Geoda is free software for spatial data analysis was originally developed by Luc Anselin. Geoda can be downloaded from <http://geodacenter.github.io/download.html>.

**1. Create a new GeoDa project and load data**

Launch GeoDa. File → New → ESRI Shapefile (\*.shp) (Figure 1). Select buenosaires.shp. For the meaning of the attribute table, open buenosaires\_metadata.html in the buenosaires folder.



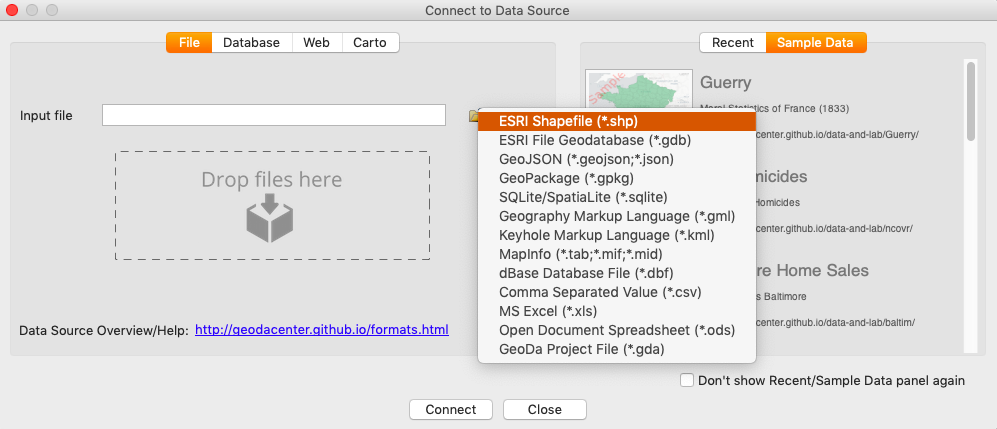


Figure 1 Create New Project

**2. Choropleth mapping**

Percentile Mapping: a percentile map shows the spatial pattern of a variable using its percentile values.

* To show the data, select Map → Themeless Map. By right click on a map, select Change Current Map Type → Percentile. Then, choose the variable APR99PC (the electoral results for the center right party APR, “Action por la Republica”).

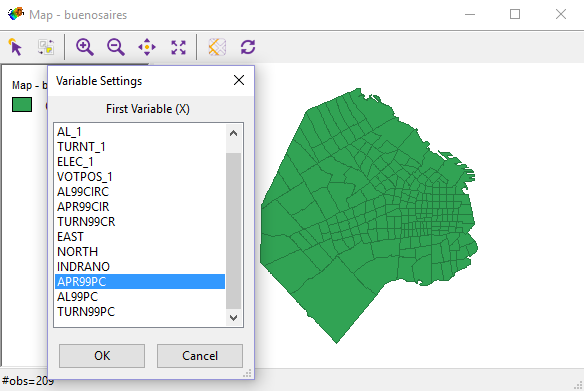
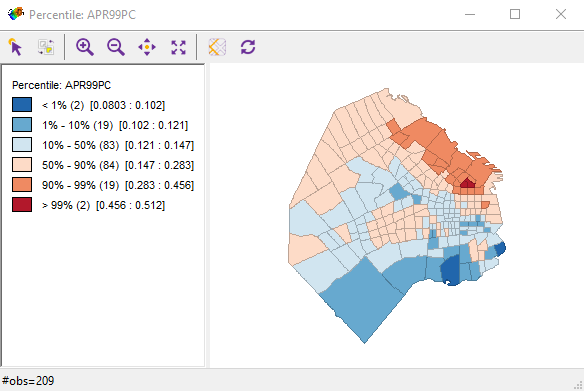
 

Figure 2 Interface for choropleth mapping and Percentile map

Box Map: a box map corresponds to boxplot of EDA. It shows the spatial pattern of a variable with its quartile values.

* By right click on a map, select Change Current Map Type → Box Map → Hinge=1.5. Then, select APR99PC variable. Note: If you continue from previous percentile map, GeoDa will default to the APR99PC variable.
* Open a boxplot to compare them. Select Explore → Box Plot and then, select APR99PC variable again.
* In boxplot, highlight positive outlier as Figure 3. Observe the spatial pattern in the box map.

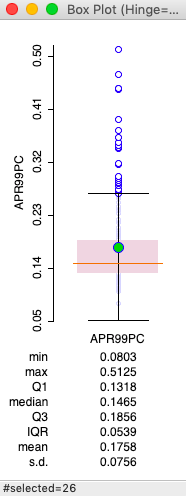
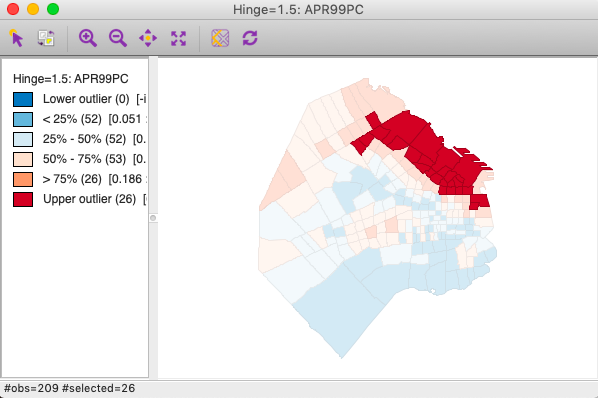


Figure 3 Box map and box plot

**Assignment I**

* For the variable AL99PC (the electoral results for the centrist “Alianza”), Make a box map and box plot as Figure 3. But highlight negative outliers (not positive outliers). Include these in your lab report.
* Briefly discuss about [1] the map pattern of the AL99PC variable and [2] the spatial pattern of the negative outliers.
* For the variable TURN99PC (the vote turnout), make a box map and box plot as Figure 3. But highlight negative outliers. Include these in your lab report.
* Briefly discuss about [1] the map pattern of the TURN99PC variable and [2] the spatial pattern of the negative outliers.

**3. Brushing and Linking**

For the rest of this lab, we will use the shapefile of Columbus. You need to close the previous opened buenosaires.shp by click File → Close. Then create a new project using Colmubus.shp. For the information of the data, open columbus.html in the columbus folder.

In order to browse the brushing and linking functions, we will make a choropleth map, open the attribute table, and make graphs.

* Quantile map: choose Map → Quantile, select the number of classes as 4, then choose CRIME for variable.
* Histogram: choose Explore → Histogram and choose HOVAL for variable. Press OK button.
* Parallel coordinate plot: choose Explore → Parallel Coordinate Plot and choose CRIME, HOVAL, INC for variables. Press OK button.
* Open table: Click the table button  on GeoDa.

For linking, select polygons which are in the 4th range (refer Figure 4). Hint: Press Shift button in the key board to select several polygons.

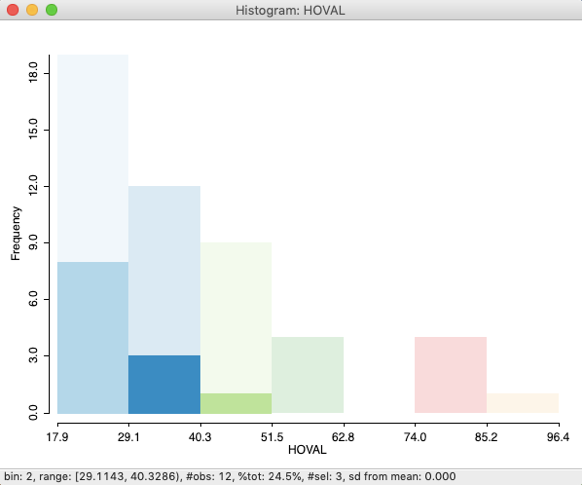
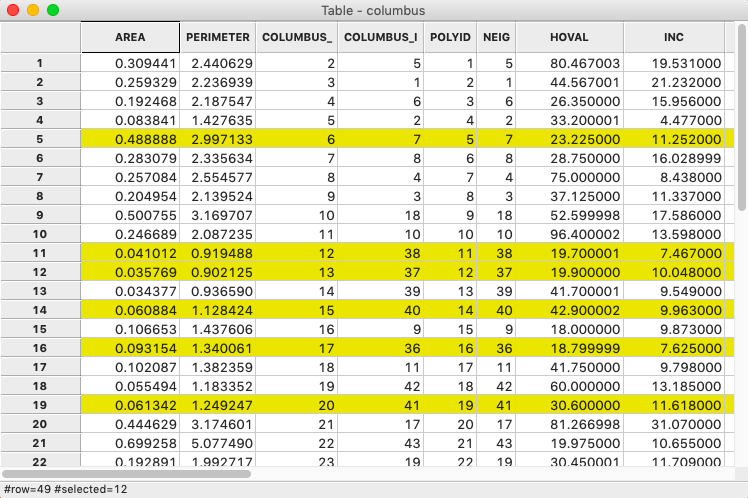
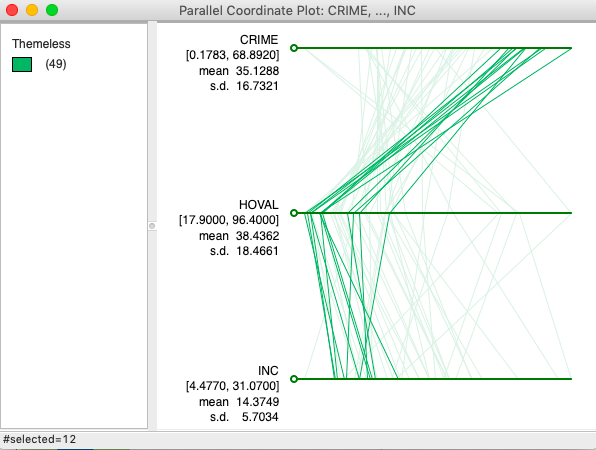
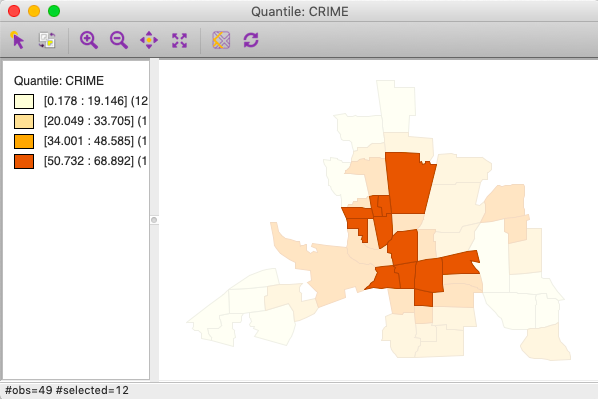


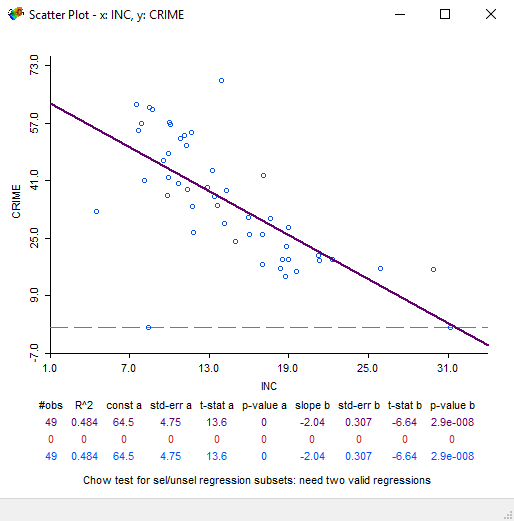
Figure 4 Maps, tables, and graphs in GeoDa

For brushing, move mouse pointer on the map. Then, holding <Ctrl> key, make a rectangle with the mouse. Move the mouse pointer on the map freely and see the dynamic change of selection.

**4. More plots**

Scatterplot: GeoDa provides dynamic functions on a scatterplot

* Choose Explore → Scatter plot and choose CRIME (Y variable) and INC (X variable). Press OK button. Adjust the window size to properly display all the statistics (Figure 5(a)).
* Select potential outliers on the left bottom. GeoDa provide the change of regression line and its coefficient value. See the dynamic change. Figure 5(b) shows the results by selecting two potential outliers in the lower left side (highlighted in the selection box).

(a) 

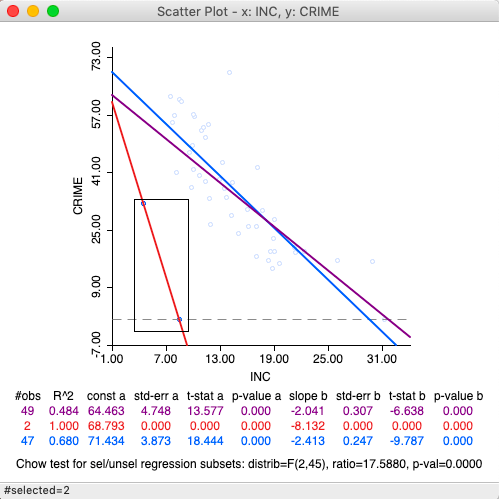
(b) 

Figure 5 Scatter plot and dynamic functions

3D Scatter Plot: Using 3D scatter plot, the relationship between 3 variables can be investigated.

* Choose Explore → 3D Scatter plot and choose HOVAL (X variable), INC (Y variable), and CRIME (Z variable). Press OK (Figure 6).

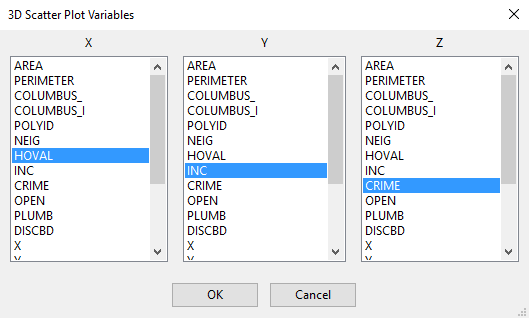


Figure 6 Choose variables

* Using a mouse, you can change the perspective of the 3D scatterplot (Figure 7).

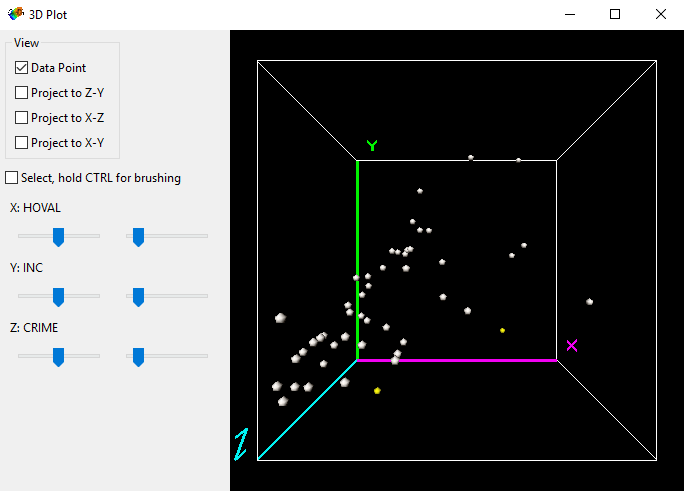
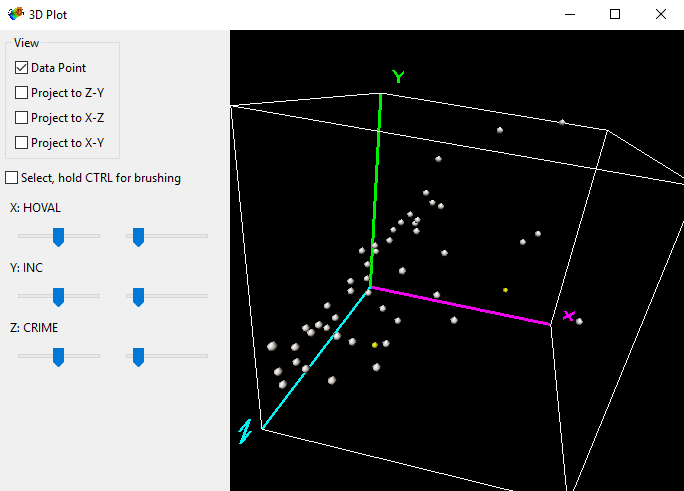
 

Figure 7 3D-plots

**Assignment II**

* Make a boxplot of INC variable and selected the observation that has the highest INC value. Include this boxplot in your lab report.
* Take a look at the boxplot. Is the selected observation a potential outlier?
* Take a look at the scatterplot of Figure 5 (a). Is the observation with the highest INC value a potential outlier?
* Make a quantile map with INC variable and select the observation with highest INC value. Include this quantile map in your lab report. Briefly discuss [1] if the observation is a potential outlier in the map and [2] whether there is any spatial outlier in the map or not.