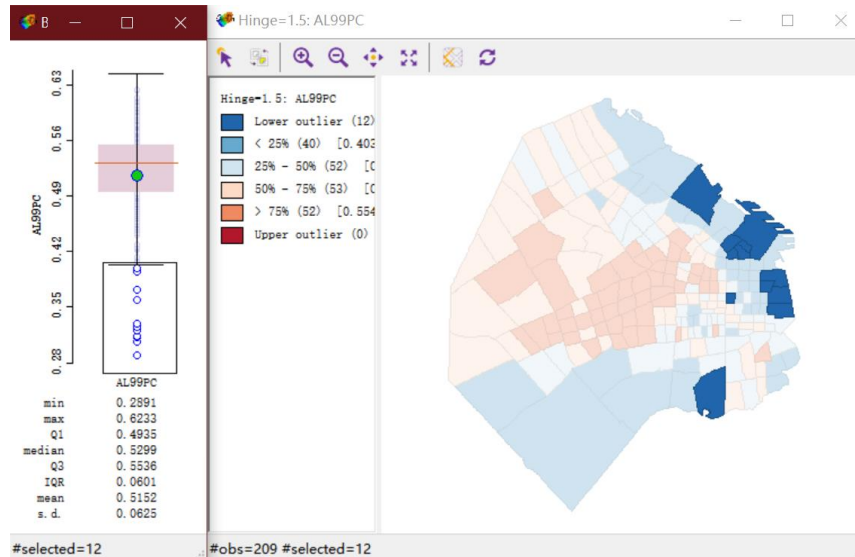


Lab2 Report

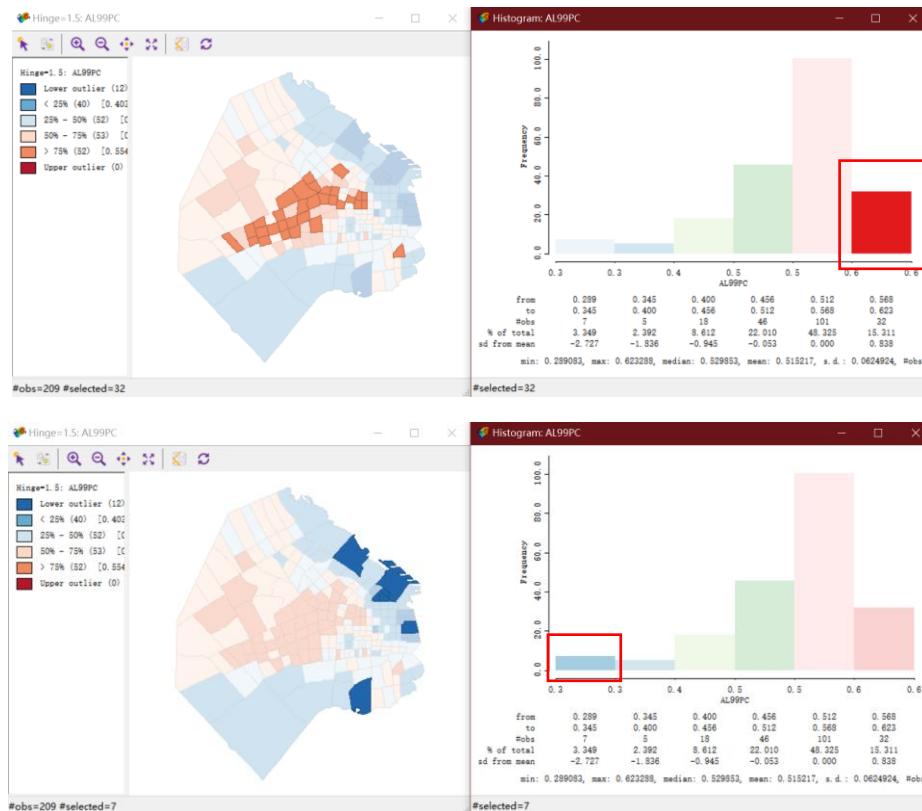
JIANG Yuhan 18106651x

Assignment1:

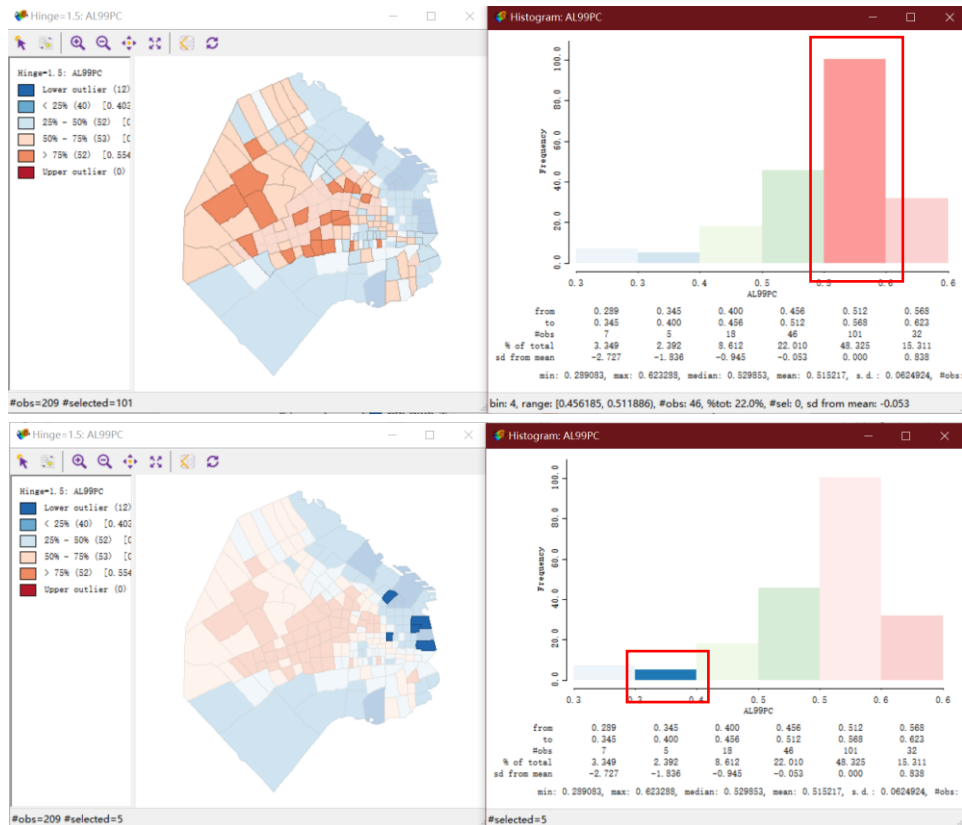
1. AL99PC



[1] The map pattern of AL99PC variable



From the box map of AL99PC variable, we can see that the place where AL99PC value is the highest is concentrated in the center while where the lowest value of AL99PC is concentrated in the eastern edge.



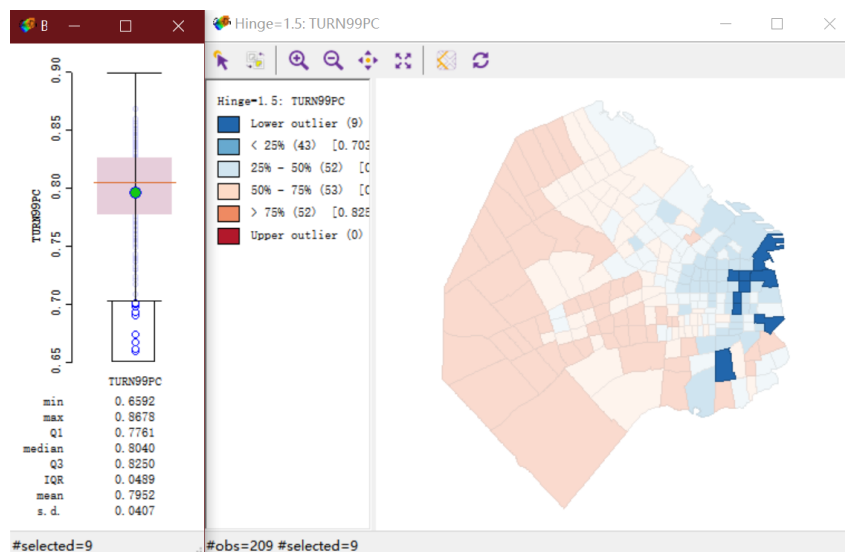
Additionally, the highest frequency of AL99PC corresponds to the value around 0.55 while the lowest frequency of AL99PC corresponds to the value around 0.35.

In conclusion, the lower AL99PC value is concentrated in the marginal area except for the western edge. The higher AL99PC value is concentrated in the central area and the western edge. Also, the closer to the center, the larger the value. On the contrary, the closer to the eastern edge, the smaller the value.

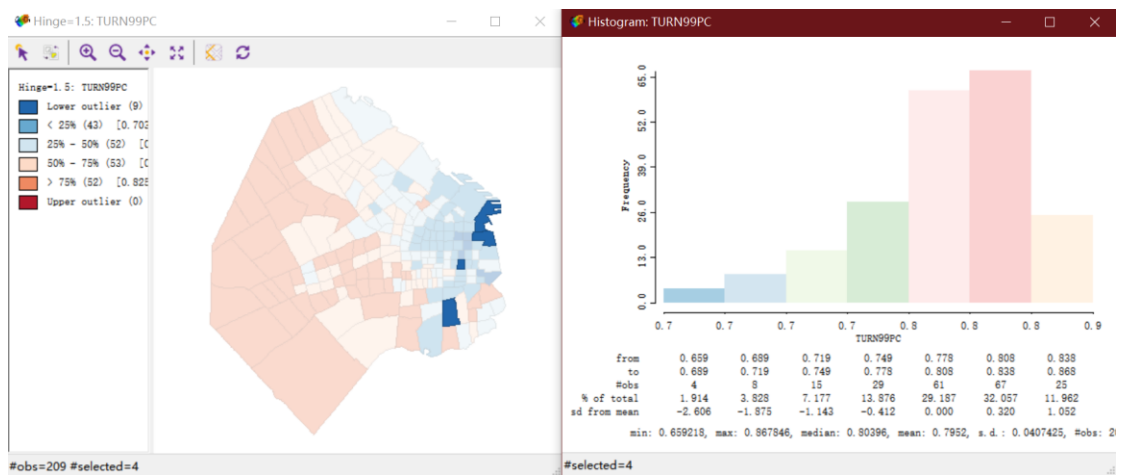
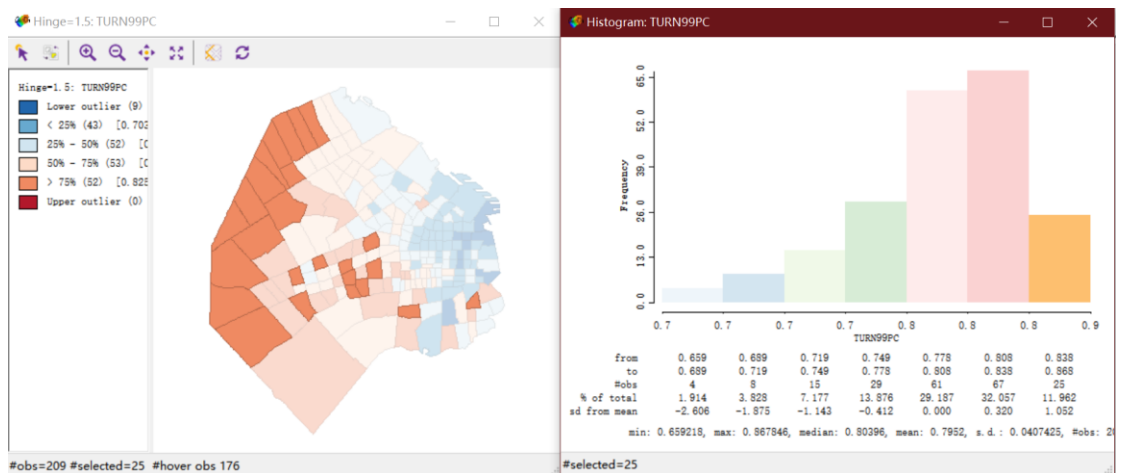
[2] The spatial pattern of the negative outliers

The negative outliers are almost at the eastern area.

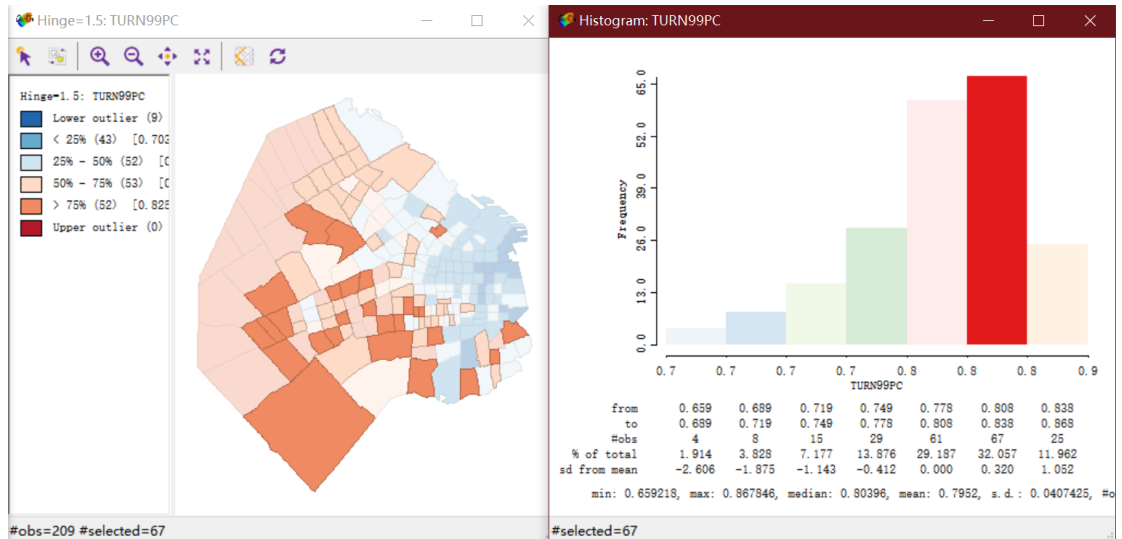
2. TUREN99PC

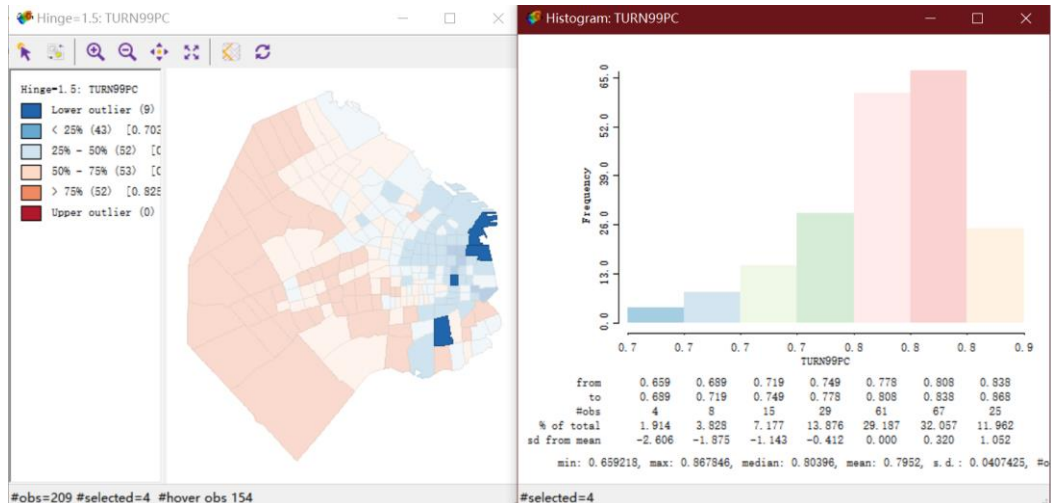


[1] The map pattern of TURN99PC variable



The box map of TURN99PC variable tells us that that the highest value of TURN99PC is in the west while the lowest value is in the east.



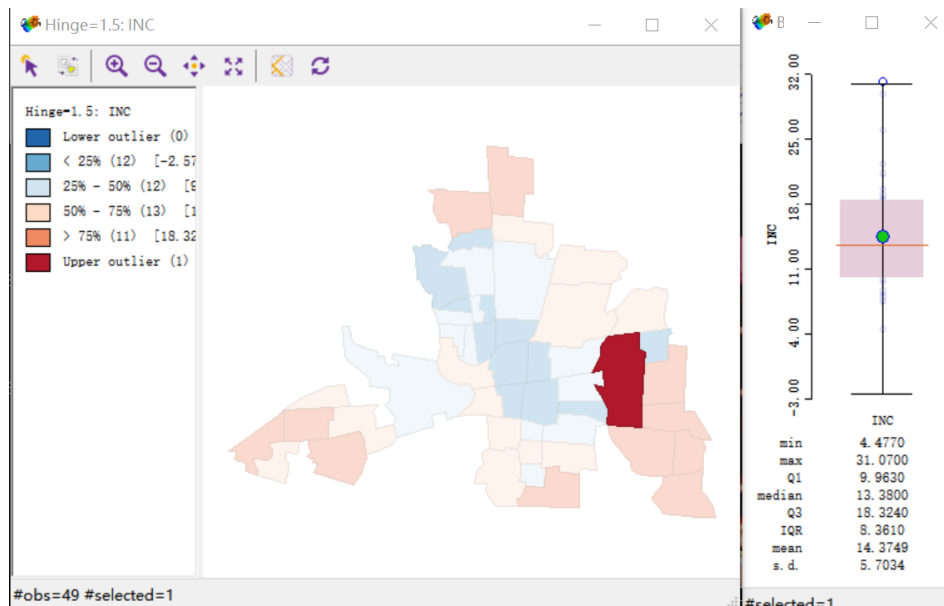


The highest frequency of TURN99PC corresponds to the value around 0.82 while the lowest frequency of TURN99PC just corresponds to the lowest value around 0.67. Additionally, the highest frequency area is in the intermediate area. Now, we can draw the conclusion that, the closer to the western edge, the bigger the value of TURN99PC. On the contrary, the closer to the eastern edge, the smaller the value of TURN99PC.

- [2] The spatial pattern of the negative outliers
The negative outliers is in the eastern area.

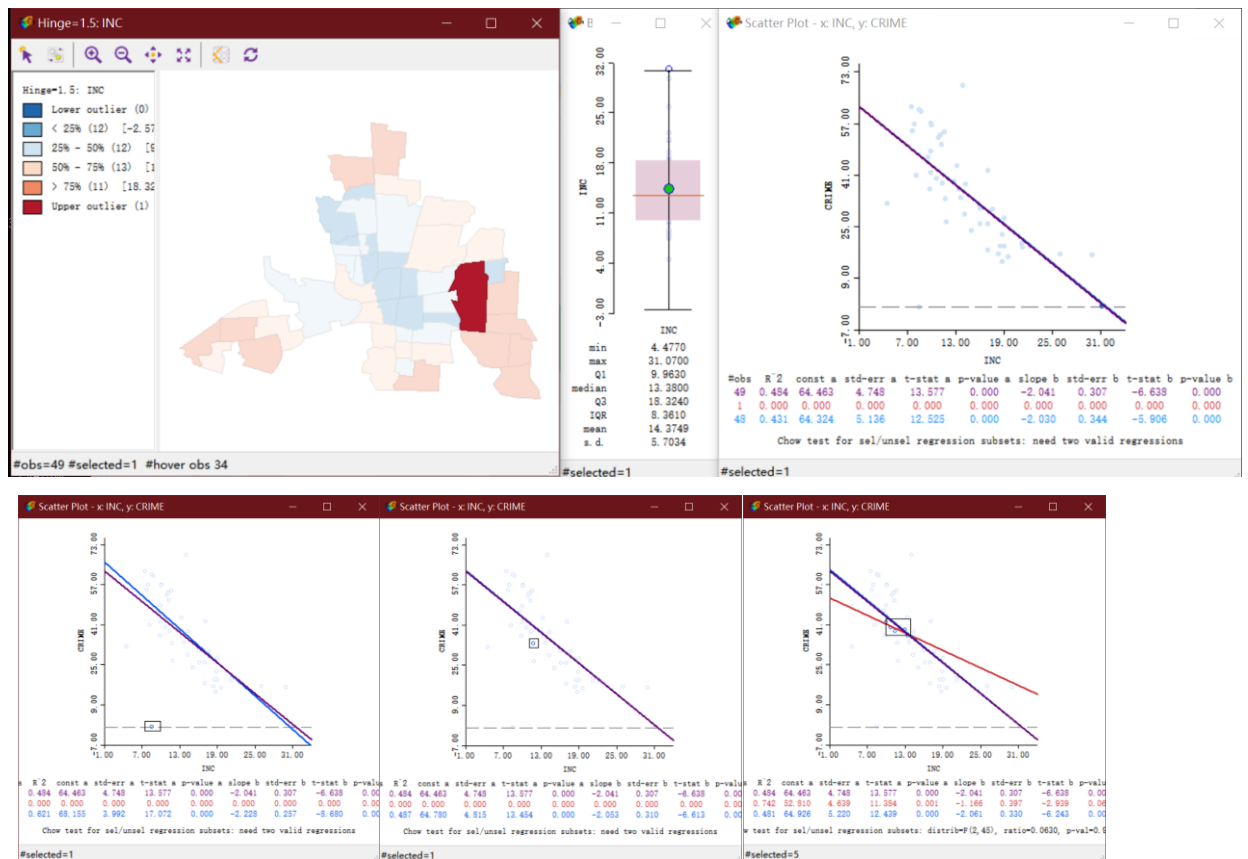
Assignment2:

1. Boxplot of INC variable



- [1] From the picture above, I think the selected observation is a potential outlier. In the box plot, the selected observation which is corresponds to the blue point on the top is just at the line of max. So, there is great possibility for this point to become an outlier.

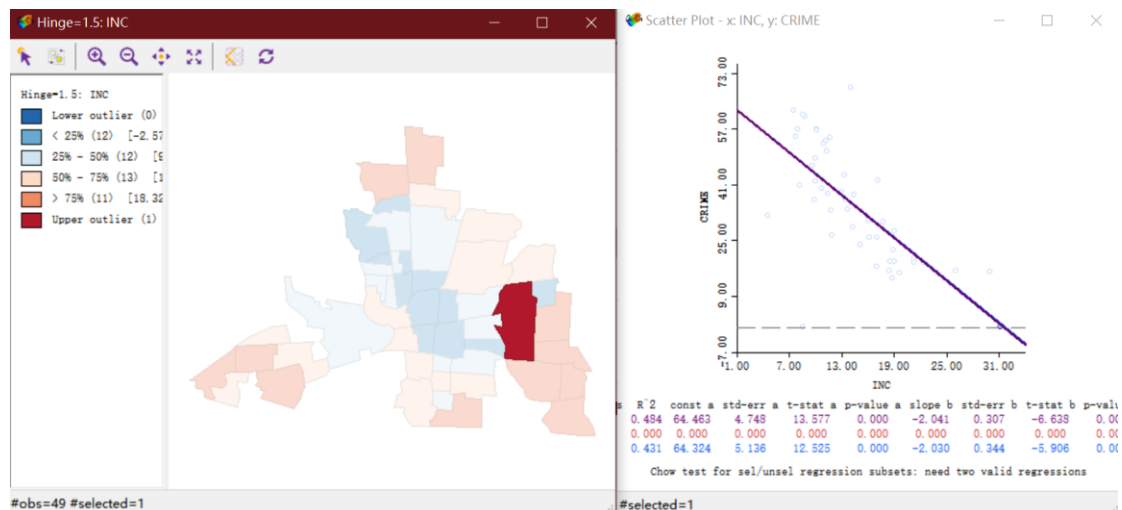
2. Scatter plot of INC variable



[1] choose three cases to compare. However, I don't think the observation with the highest **INC** value is a potential outlier. In the first picture, when I remove the point with low **INC** value, which is far from the line, the new line in purple is obviously off the old line which means this point has strong effect to the whole set of data. In other words, the feature of this point is obviously different from other points. The point which is far from the line maybe the outlier. This conclusion can be verified from the latter two cases.

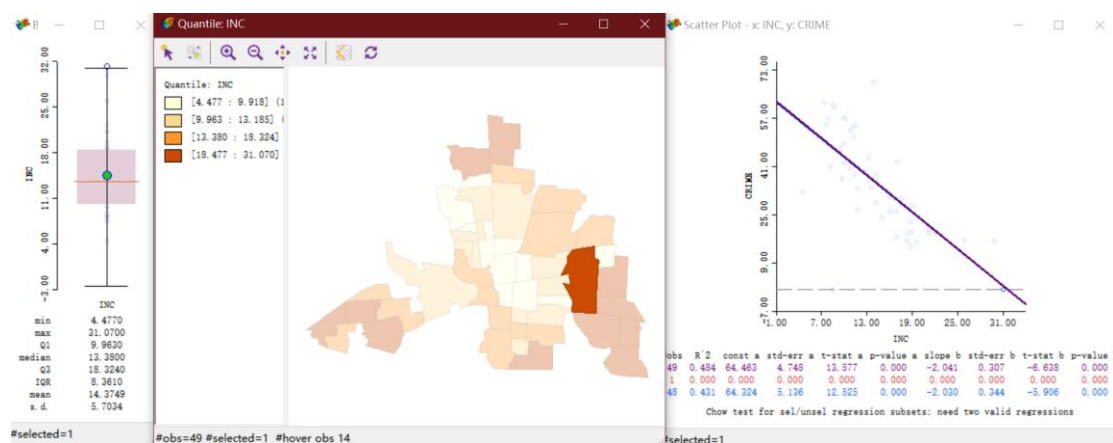
When I remove the points, which are around the line, the new line and the old line is almost coincide, which means those points I removed have the same characters with the points left.

So, we can draw the conclusion that the outlier is the point that far from the line. On the contrary, the closer the point is to the line, the more consistent it is with the characteristics of the group.



As the picture shown above, when I select the point with the highest INC value, there is little change to the line. Additionally, the point is almost in the line. So, in my opinion, the observations with the highest INC is not a potential outlier in this case.

3. Quantile map with INC variable



- [1] If the observation is a potential outlier in the map

I think the observation with the highest INC is not a potential outlier. Firstly, in the quantile map, the area of the four graded areas is approximately equal which means the classification is reasonable. So, the observation with the highest INC value is within acceptable limits. Additionally, combining the analysis in question2, the observation is less likely to be an outlier.

- [2] whether there is any spatial outlier in the map or not

I think there is no outlier in this map. Firstly, the smaller values are all above the min which are not likely to be outlier. Secondly, the bigger values are also not likely to be outlier as we analyzed before. As for the points far away from the line in the scatter plot, it maybe due to the value of Crime. We can not rule out that there is potential outlier in Crime. As a result, I hold the view that there is no outlier in INN.