Module 07 Practical

Kyle Lago

HLTH 661

20658108

25-June-2023

1. We can identify clustering within the log population per square km map of Kitchener. While I’m not familiar with the map of Kitchener, there are two clusters on the map where the log population is much higher than other areas comparatively. I would imagine that these are highly populated residential or urban areas like a downtown or university campus.
2. A spatial weight matrix allows us to capture and quantify the spatial relationships between two geographical units. It helps us understand spatial patterns and perform spatial analysis. For example, it can help us identify spatial autocorrelation between neighboring geographic units.
3. The main difference between queen contiguity and rook contiguity is how they consider their neighboring polygons. Queen contiguity considers two polygons neighbors if they share any point along their boundaries, where as rook only considers this if they share an entire edge.

A screenshot of a computer

Description automatically generated with medium confidence

1. Morans I = 0.368
2. The pseudo p-value is 0.005000
3. The pseudo p-value is 0.00100
4. Both of these p-values that we have generated indicate that there is evidence of spatial autocorrelation for this map. The permutation values are simply the amount of times that the data values are randomly shuffled to create random patterns and each time a new Moran’s I is calculated. It compares the new Moran’s I to the original Moran’s I (0.368), and the p-value tells us if our original Moran’s I is significant.

The difference between the p-values yielded by the “999 permutations” and “199 permutations” is just the amount of random shuffling. The former is shuffled 800 more times, and this reduces the likelihood of obtaining extreme Moran’s I value by chance. Thus, this provides us with a more reliable p value. This is like running statistical tests where you have significance levels of p=0.05 and p=0.01. In the “999 permutation” scenario, the chances of us finding the Moran’s I of 0.368 by chance alone is less than 1%, so we conclude that is significant and that there is spatial autocorrelation in our data.



|  |  |  |
| --- | --- | --- |
| 999 Permutations | Morans I | P Values |
| 1st order contiguities | 0.368 | 0.001 |
| 2nd order contiguities | 0.158 | 0.001 |
| 3rd order contiguities | -0.000 | 0.365000 |

10. As we move from 1st order contiguities to 2nd order, our Morans I value goes down. This suggests that our previous positive autocorrelation has now reduced. This is because as we increase the level of our orders (i.e. from 1st to 2nd) we are telling GeoDa to include more neighboring polygons. In the first order, we are only looking at the neighbours, but in the second order we are looking at the neighbors of the neighbors as well.

This trend continues into the 3rd order contiguity where we lose significance entirely and Morans I indicates that there is no longer any spatial autocorrelation. The significant positive autocorrelation that we observed at the 1st and 2nd order does not exist at the 3rd order.

This intuitively makes sense with our Kitchener data. 3rd order contiguities would include more relationships between distant polygons that do not have any relationship with some of the clusters we noticed in the immediate visualization and introduces noise to our data.

1. A screenshot of a computer

   Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

12 & 13. By using a smaller areal units (dissemination area versus census tract) we see our Morans I increase positively for each nth order contiguity. Of note is the Moran’s I of 0.067 which remains significant at 3rd order contiguity with a p-value of 0.001. This contrasts with what we saw with the census tract map. This is because we are using a smaller areal unit, and this increases our sensitivity in small geographical areas. These smaller areas are more likely to share boundaries with each other as well. As they are more likely to share more boundaries, our queen contiguity will include more neighboring polygons, especially at the first order where we have our highest Moran’s I (0.402).

This can be seen roughly in the following snippets from our projects. In the large DA cluster, there are far more polygons with neighbors than in the CT cluster. This demonstrates that using smaller areal units can be beneficial for spatial scale when looking at high populations in small geographic boundaries.

Figure 1. DA Cluster Snippet


*Figure 1. DA Cluster Snippet*

*A picture containing map

Description automatically generated with medium confidence*

*Figure 2. CT Cluster Snippet*

14*.* One could imagine the usefulness of this finer spatial scale when trying to determine the outbreak location of a disease in a major urban area. Moran’s I would be extremely helpful in quantifying spatial autocorrelation between potential disease sites. For example, studying COVID-19 cases in a densely populated city. Calculating Moran’s I would allow them to find out whether there is positive, none, or negative spatial correlation of COVID-19 in the city. If they identified positive autocorrelation in certain areas, it can inform them to target these areas more consistently with interventions.