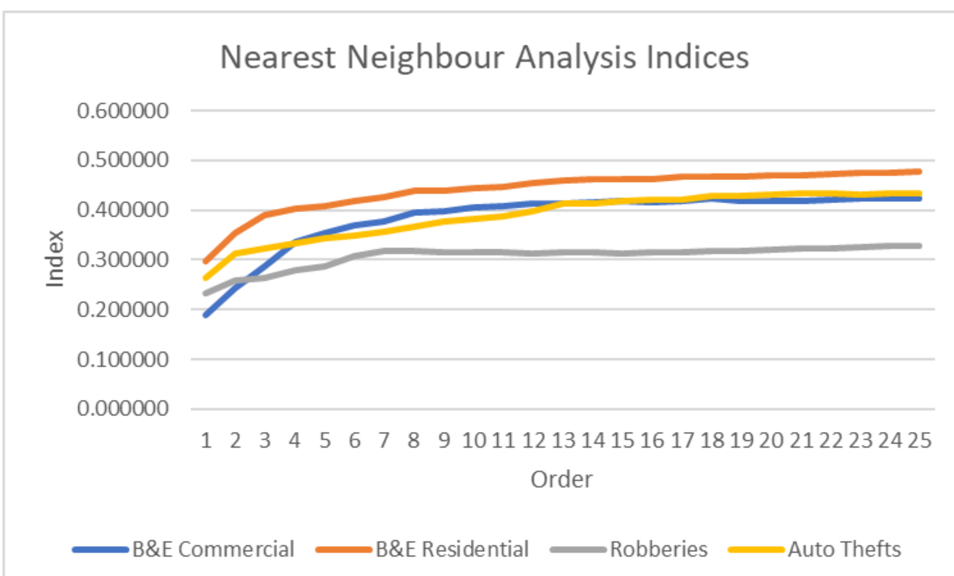
CRIME PATTERNS AND SPATIAL CLUSTERING IN OTTAWA-NEPEAN

Lab 3, GEOB 479

Nigel Tan 17568163
March 2020

1. First order crimes seem to show a relatively even amount of spatial aggregation between each crime type. From the small difference in index value at this point, there is a relatively high amount of spatial aggregation. Increasing the order sees patterns begin to emerge, such as robberies dropping away from the other crime types with a much lower index starting around the 6th order. Around the 12th order, auto thefts and commercial break & enters almost overlap, showing extremely high spatial aggregation. This could be because cars are more difficult to steal when parked in one's home, likely within a garage, meaning that car thefts would occur more when parked in open areas, which are often spatially clustered such as parking lots. Commercial areas are likely to be more clustered as well due to zoning restrictions such as downtown areas, while residential zones may be more randomly interspersed.



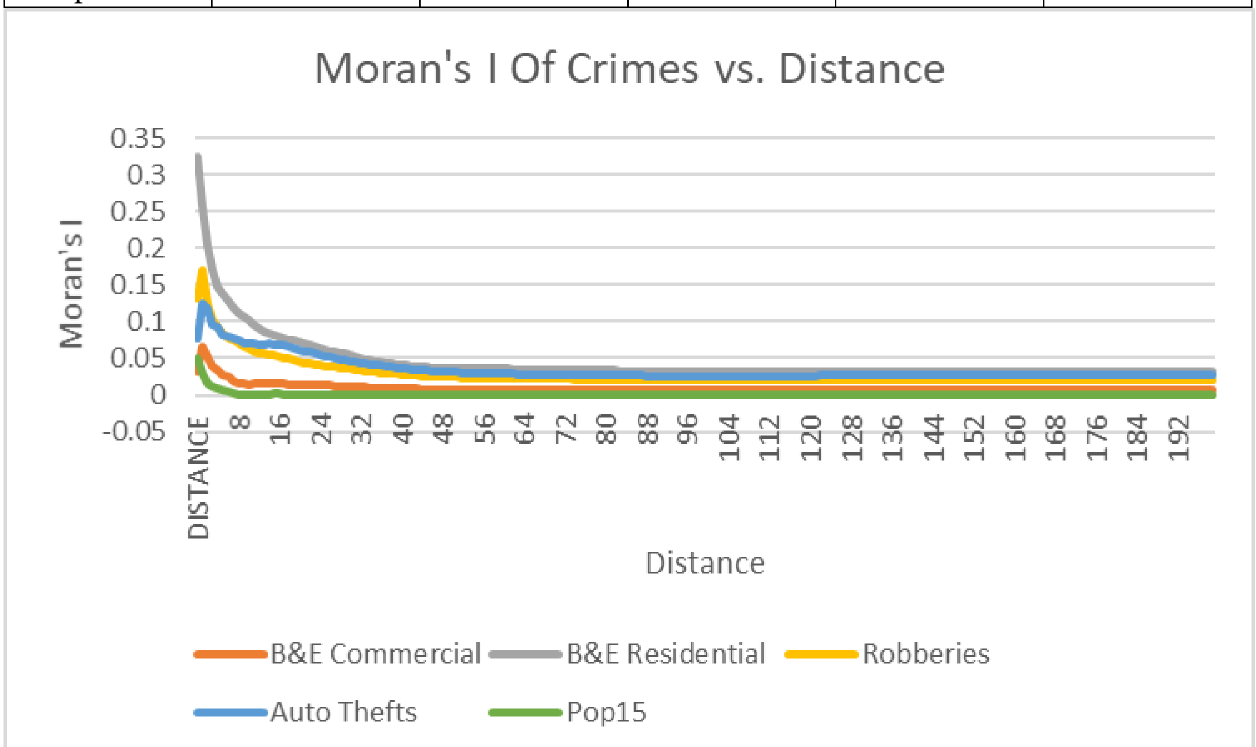
2. The correlograms show that intensity starts out spatially correlated to a certain extent for all crimes, but quickly drop to random distribution in an exponential manner.

Considering that all crime types had a much higher initial Moran's I than the pop15 default value, it can be said that the crimes are all spatially autocorrelated in terms of intensity of occurrence. A Moran's I correlogram analysis differs from the nearest neighbour analysis in that the value ranges are inverted, a nearest neighbour analysis shows 0 being perfectly related and 1 being completely randomly distributed while Moran's I goes from 1 being perfect autocorrelation to 0 being random. Due to urban land use zoning bylaws, perfect randomness is usually impossible and there will always be some degree of spatial clustering simply due to the fact that urban zoning forces given property and land use types to be clustered. As a result, middle values are quite common and represent a realistic zero value.

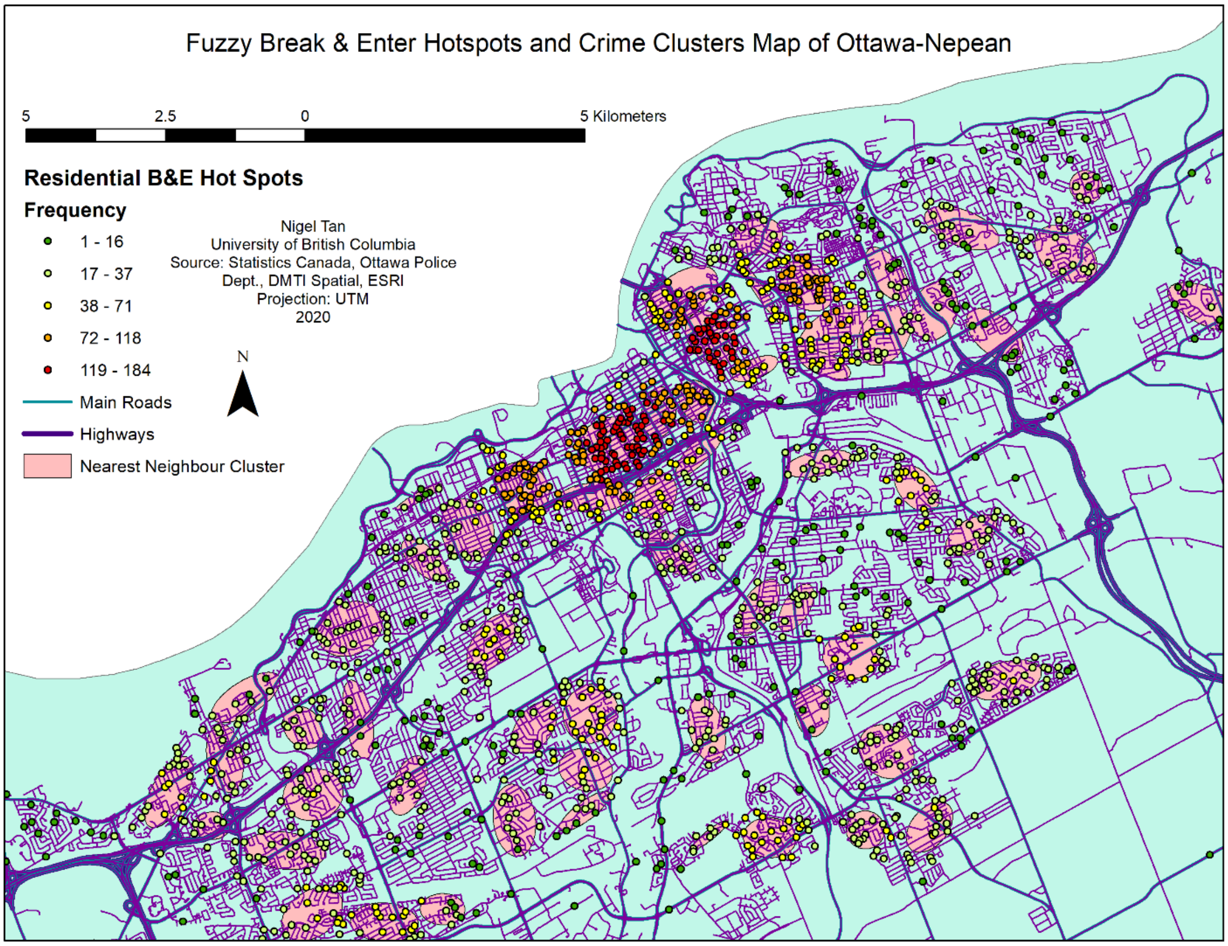
Moran's I Results

The Moran's results show the degree to which intensity of crimes is spatially autocorrelated. This is done through analyzing the pattern of crime distribution and ranges from 0 to 1, going from completely random to perfectly spatially autocorrelated. The graph below plots the Moran's I of each crime type against a given distance class, with each increasing distance class number indicating a greater distance away from the dissemination area. Here, the Pop15, which is population over the age of 15, shows almost perfect randomness each moving a certain distance away. The most spatially autocorrelated crime type is B & E residential, meaning that residential break and enters have the highest correlation of intensity. As with all the other values, the spatial autocorrelation of the intensity drops drastically and becomes a random distribution, with all crime types becoming completely random by the 100th distance class.

Correlograms	B&E Residential	B&E Commercial	Robberies	Vehicle Thefts	Pop15
Moran's I	0.032	0.032	0.0274	0.0264	-0.000031
Random I	-0.000754	-0.000754	-0.000754	-0.000754	-0.000754
Standard Error	0.00108	0.001081	0.001083	0.001081	0.001076
Normality Significance	30.645	30.706	19.851	25.156	0.671
Random Significance	30.675	30.736	20.288	25.253	0.677
Sample Size	1328	1328	1328	1328	1328



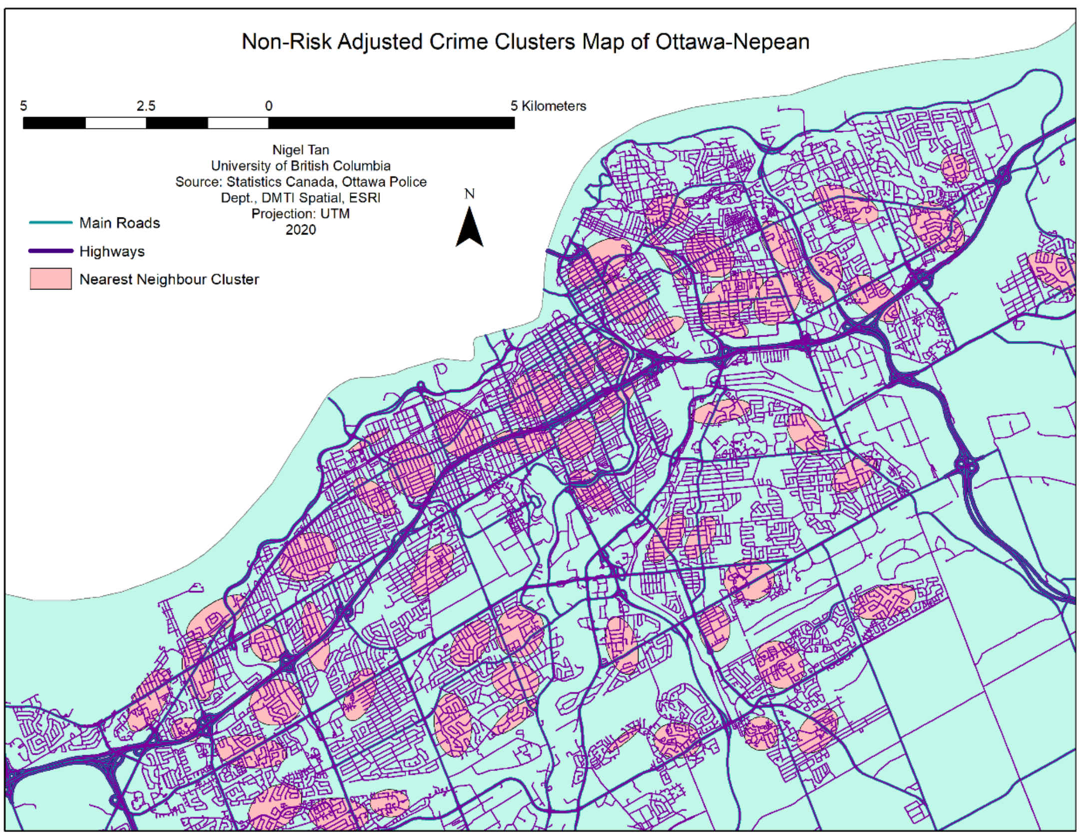
Fuzzy Hotspot and Nearest Neighbour Analysis

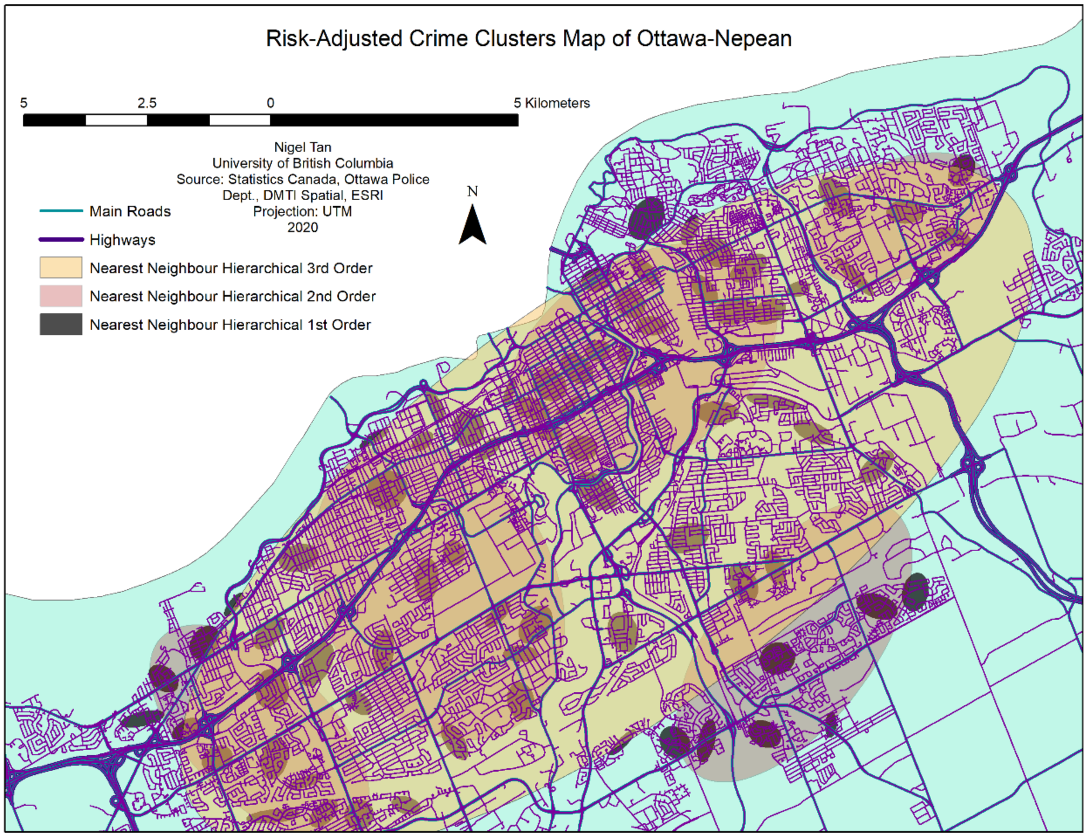


The fuzzy mode residential break and enter clusters are displayed as point data across the map. Broken into five classes, the points vary in colour based on even intensity, with a range of 1-184 events occurring at a single point and with colours ranging from green to red. Points that occur closer together tend to be redder, indicating spatial clustering. On the other hand, the nearest neighbour hierarchical analysis outlines entire hotspot areas without identifying individual points. This is more useful for instantly identifying clustered areas, but is more susceptible to generalization and may suffer from MAUP.

Comparing the results of the two, the map looks quite similar. All red fuzzy point are entirely contained within a nearest neighbour cluster. Orange points are mostly contained or in the immediate vicinity of clusters. However, several green points lie far outside the nearest cluster, indicating that there are too few events to be considered spatially clustered. Together, these tools are useful to identifying particular geographic areas that are high in a particular type of crime.

Non Risk-Adjusted vs Risk Adjusted Clusters





The above two maps compare normal nearest neighbour clustering to risk-adjusted clustering. Risk-adjusted clustering accounts for the population over the age of 15 and normalizes the data, creating a population weighted cluster. What this means is that the crime frequency rate is divided by the population. Thus, if an area has a high crime rate and high population, the individual risk rate is low, whereas an area with a high crime rate and low population would have a relatively high individual risk. The risk adjusted clusters are separated into three orders. The first order clusters are based on crime count and distance away, similar to the non risk-adjusted clusters. Second order clusters take into account the centre of the first order clusters, while third order clusters take into account all crime instances.

Comparing just the first order clusters, clusters are about the same size but differ in location and shape. This shows the weight that population has upon determining the cluster, as the risk-adjusted clusters represent risk level, rather than raw count data. This makes the risk-adjusted clusters far more valuable as they are normalized and standardized.

Knox Index

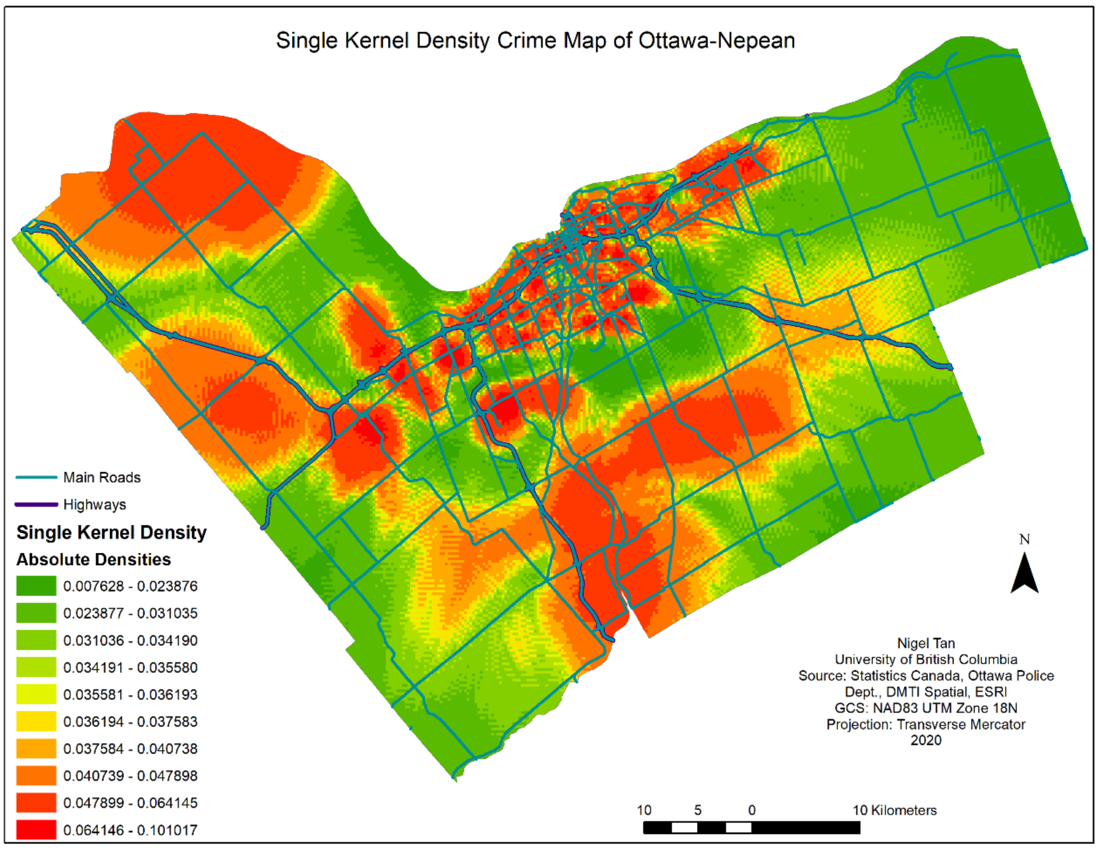
The Knox Index measures clustering while taking the dimension of time into account. The Knox results 99 simulations of a sample size of 2152. The results are divided up into four possible categories shown below:

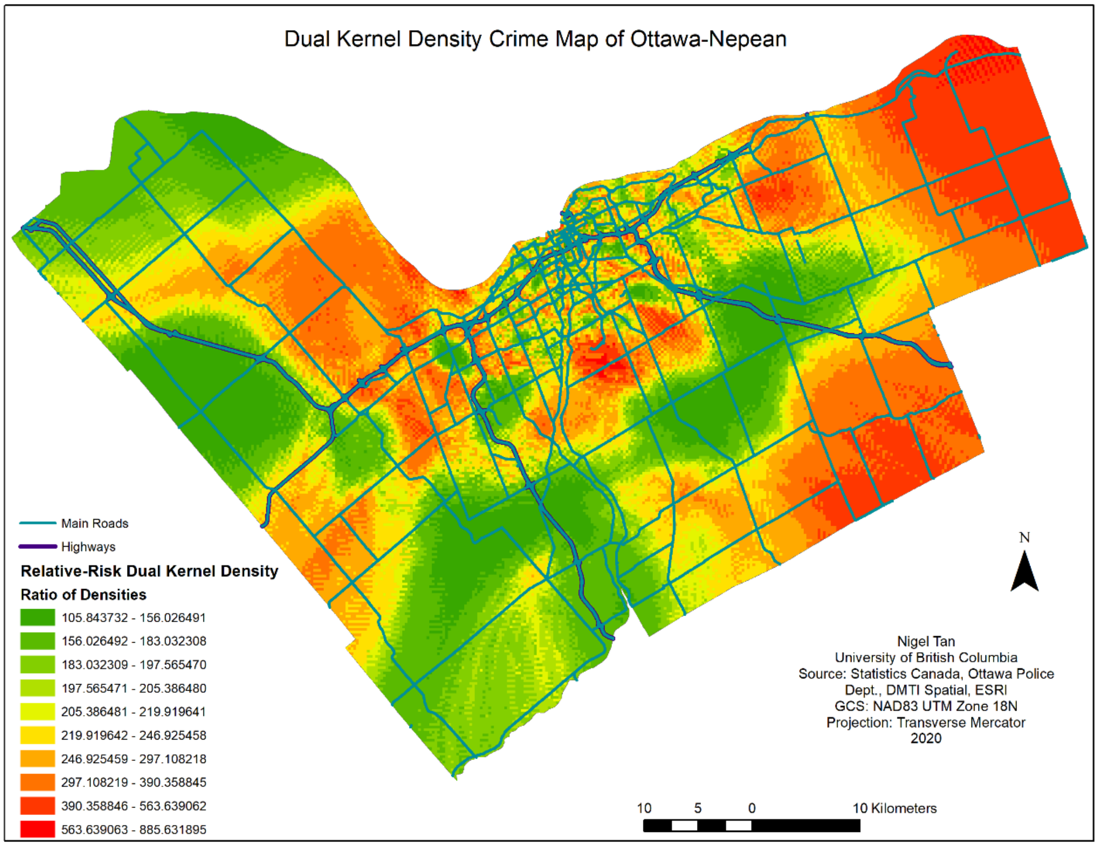
	Spatially Close	Spatially Not Close
Temporally Close	50289	1535276
Temporally Not Close	65608	210763

The time span considered to be close is within 12 hours, while the distance limit is 5km. Overall, the results show that car thefts often occur at the same time of day, regardless of location, as the number of spatially not close yet temporally close car thefts is by far the highest. This implies that auto thefts are not very spatially autocorrelated, but extremely temporally autocorrelated. Results for temporally not close and spatially not close are also much higher than all spatially close results combined, meaning that spatial clustering is extremely minimal and may likely be random.

The results come with a Chi-square value of 114.16713, which is extraordinarily high compared to the maximum expected Chi-square of 7.45710. This means that the results are extremely statistically significant, and results are not at all randomly distributed. It can be considered that car thefts are very tightly correlated with one another temporally.

Kernel Density Estimation





The two maps above show kernel density surfaces displaying crime intensity models. The first, the Single Kernel Density Map, shows the expected intensity value for crimes to occur in each given area. These maps used a 250m grid resolution, meaning that it identifies crime occurrence density in each 250m cell. Both Single and Dual Kernel maps were made using residential break and enter data.

The Dual Kernel map is somewhat similar to the risk-adjusted cluster map, as it normalizes the data against population over the age of 15. This relative risk map then shows individual risk rather than raw crime intensity counts, making it more useful for policy assessment.

Both types of kernel density maps are similar to hotspot analysis maps, but instead cover the entire study extent instead of isolated points. This is useful as it allows for crime estimation at any point over the grid and makes it easy to interpolate and extrapolate data regarding potential expected crimes.

Knox Index: Interaction of Space and Time

```
Sample size ...........: 2152
Measurement type ......: Direct
Input units .... ......: Meters
Time units ............: Hours
Simulation runs .......: 99
"Close" time ........:          12.00000 hours
"Close" distance ....:        5000.00000 m
                    | Close in space(1)  | Not close in space(0) |
---------------------+--------------------+----------------------+-----------------
Close in time(1)     |            502829 |             1535276 |      2038105
Not close in time(0) |             65608 |              210763 |       276371
---------------------+--------------------+----------------------+-----------------
                    |            568437 |             1746039 |      2314476
Expected:
                    | Close in space(1)  | Not close in space(0) |
---------------------+--------------------+----------------------+-----------------
Close in time(1)     |       500560.08007 |        1537544.91993 | 2038105.00000
Not close in time(0) |        67876.91993 |         208494.08007 | 276371.00000
---------------------+--------------------+----------------------+-----------------
                    |       568437.00000 |        1746039.00000 | 2314476.00000
Chi-square ..........:          114.16713
P value of Chi-square:           0.00010
Distribution of simulated index (percentile):
Percentile     Chi-square
---------- ---------------
      min        0.00620
      0.5        0.00620
      1.0        0.00620
      2.5        0.01035
```

5.0	0.02721
10.0	0.05128
90.0	3.84157
95.0	5.16858
97.5	6.12610
99.0	7.45710
99.5	7.45710
max	7.45710