Jonathan Engelbert

02/27/2017

Global Pattern Analysis of Crime Data in Seattle in 2010

The distribution of crime incidence across Seattle for graffiti, assault and credit card fraud all seems to be clustered, particularly around downtown (*maps 1,2,3*). The following inferential statistics suggest that rather than being random, these patterns are a product of spatial processes. They focus on the presence or absence of each of these types of crimes, rather than attributes; this analyses looks at the *count* of each type of crime across the city.

Moran's I

This tools analysis patterns at a global scale, assessing the correlation among nearby locations. In this case, the Moran's I tool returns very high positive z-scores (some as high as +10) and very low p-values (the majority expressing 99% confidence) when analyzing the distribution of all three types of crime (tables 1,3,5).

This confirms that there is clustering and that this clustering is not the result of chance, consequently rejecting the null hypothesis of complete spatial randomness. These results encourage the investigation of the spatial processes driving these clusterings.

Getis-Ord General G

This tool measures the degrees of clustering for high or low values, which in this case, is comparing high/low counts of crime incidence in the given neighborhoods.

Like the previous tool, it returns high positive z-scores and low p-values, in this case indicating that the distribution of high values is more clustered than it would be expected – in other words, some areas have more incidence of crimes than it would be expected if this patterns were random -, and that the level of confidence that these patterns are not random is high, respectively.

Incremental Spatial Auto Correlation

This tool assessed how clustering patterns are distributed based on distance. The distance band used for the analysis here was one mile, with increments of one mile (incremental analysis 1,2,3).

Credit card fraud is the outlier here, as both assault and graffiti incidents are more continuous, and tapper off gradually as one moves away from the city core. Credit card fraud on the other hand, drops sharply on the first mile, and then almost completely rebounds on the second. This reveals that at two miles, the spatial processes operating these clusters are more intense. This is somewhat expected when looking at the data points distribution before the analysis, as they seem to be more evenly distributed, and some intense cluster appear in north Seattle.

Conclusions

All crimes happen more often around the city core, with assault being overwhelmingly more prevalent in Downtown, Belltown and Capitol Hill – although high incidences are seem south of the city core as well.

Most reported graffiti offenses occurred in the Capitol Hill area, but the overall pattern of clustering around it is very similar. Credit Card Fraud seems to be a bit more evenly distributed across the same areas as other crimes, with a few clusters occurring away from each other, around the city core, and in north Seattle.

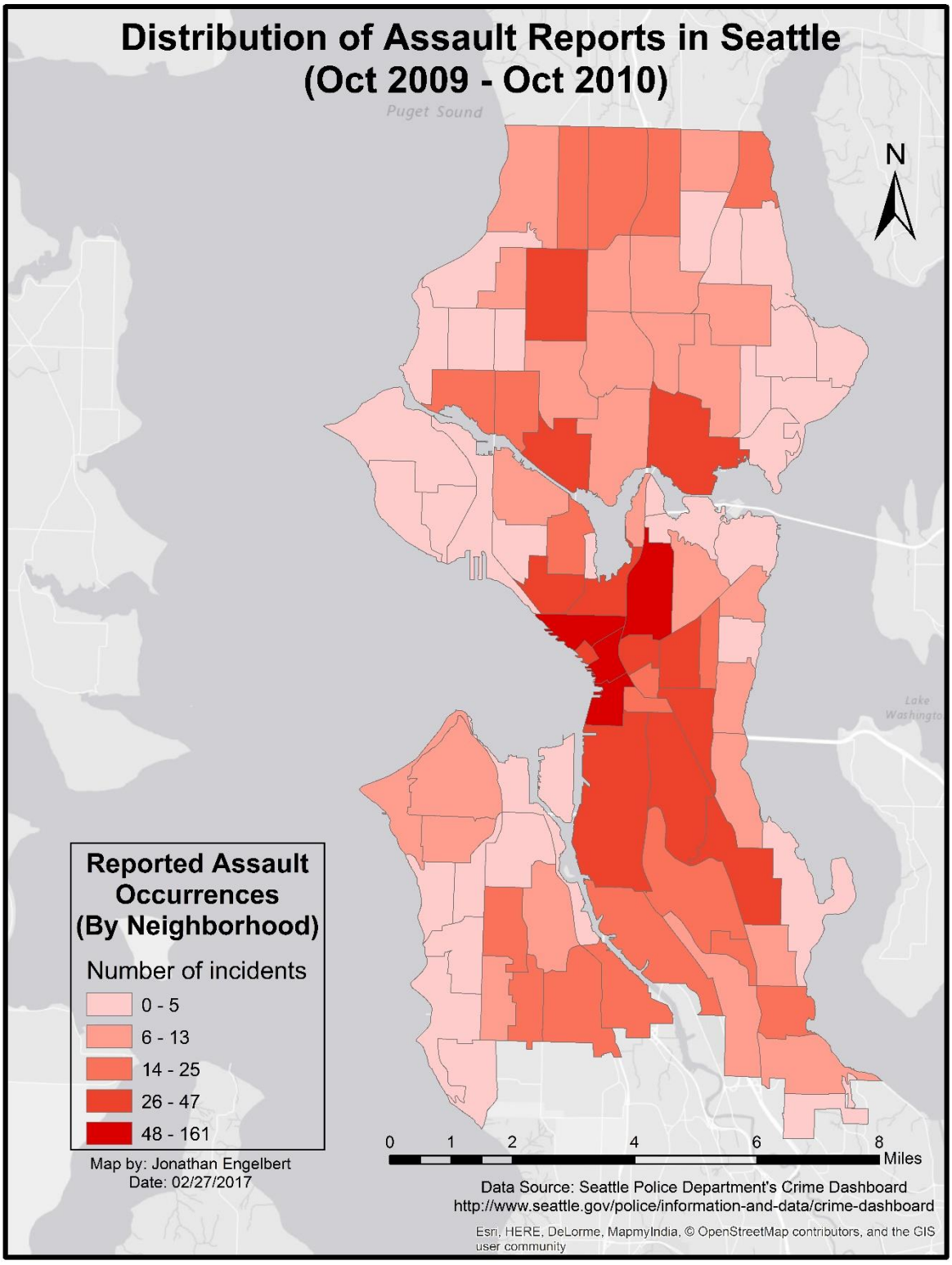
Notes on standardization and validity of results

Because we are assuming we have a complete set of all reported crime incidents, we may also confidently draw assumptions regarding clustering; the maps show that there is a clustering of crime incidence for all three types in the city core. Also, because the data is assumed to be complete, we do not need to standardize by rows for example; we are confident that the data collected is complete and not the result of a bias data collection.

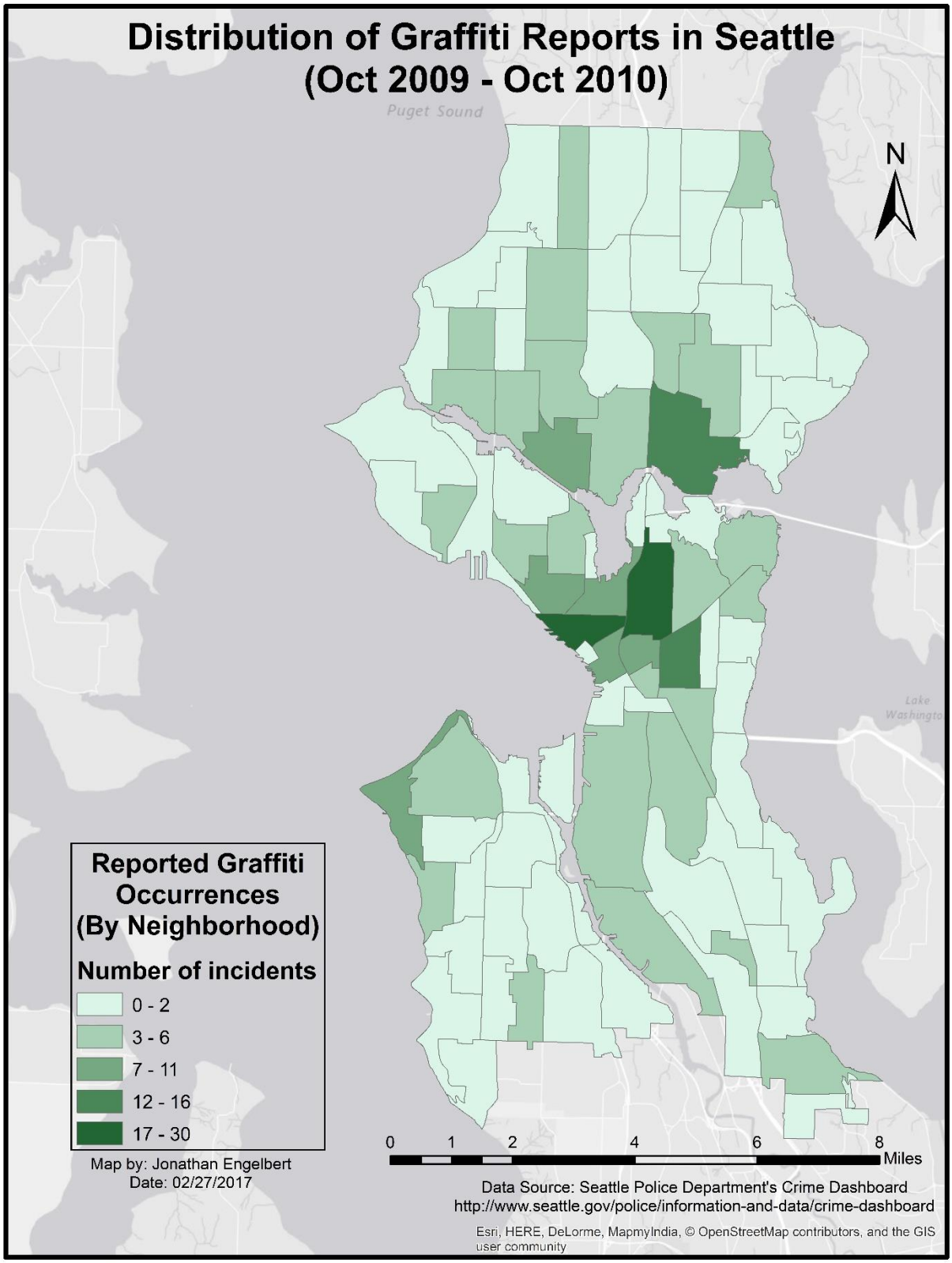
Finally, the distance band used for the analysis – that of one mile - resulted in some features having no neighbors, which could potentially invalidate the analysis.

Maps

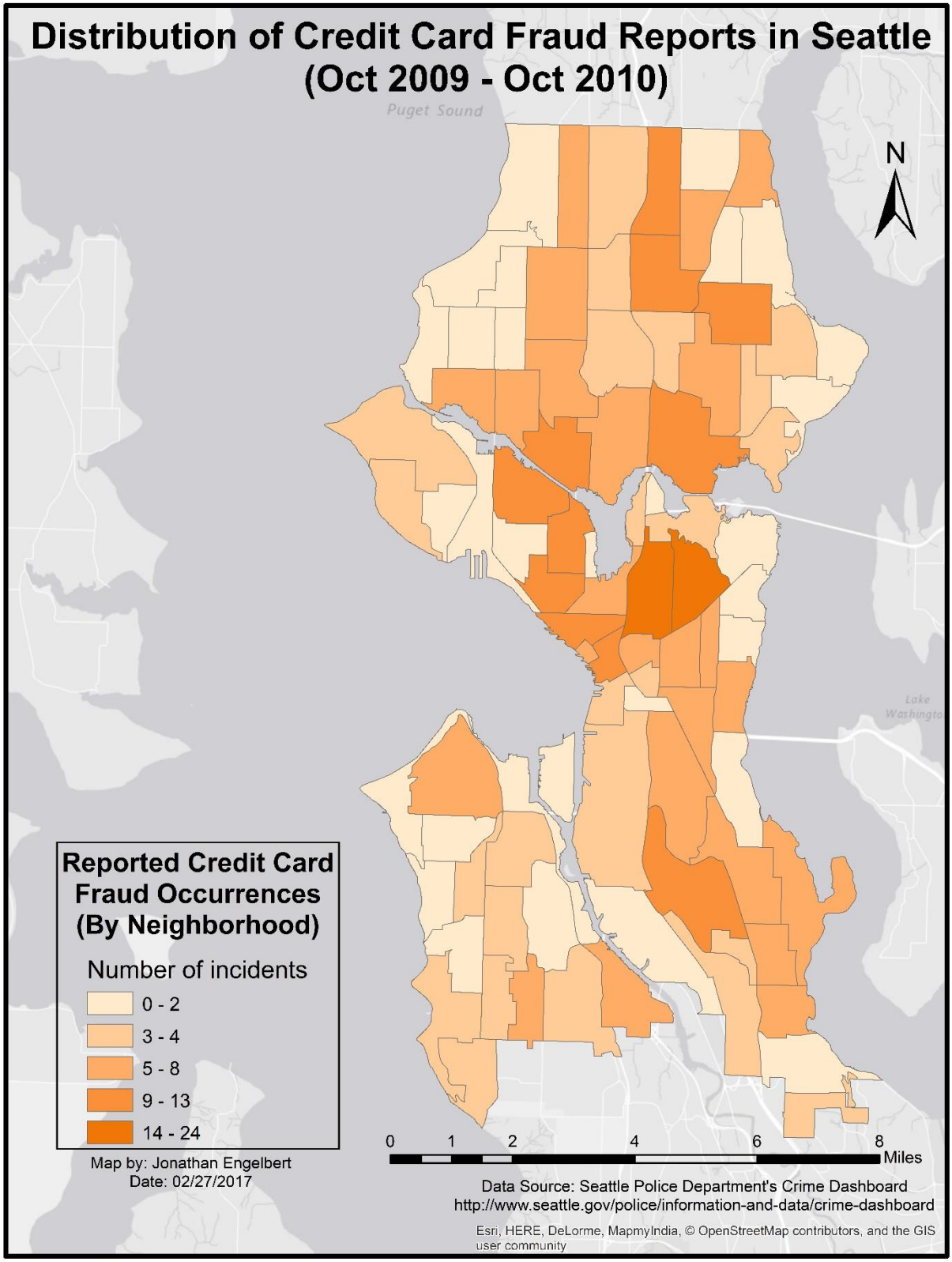
Map #1


Distribution of Assault Reports in Seattle
(Oct 2009 - Oct 2010)
Puget Sound
N
Lake Washington
Reported Assault Occurrences (By Neighborhood)
Number of incidents
0 - 5
6 - 13
14 - 25
26 - 47
48 - 161
Map by: Jonathan Engelbert
Date: 02/27/2017
0 1 2 4 6 8
Miles
Data Source: Seattle Police Department's Crime Dashboard
http://www.seattle.gov/police/information-and-data/crime-dashboard
Esri, HERE, DeLorme, MapmyIndia, © OpenStreetMap contributors, and the GIS user community

Map #2

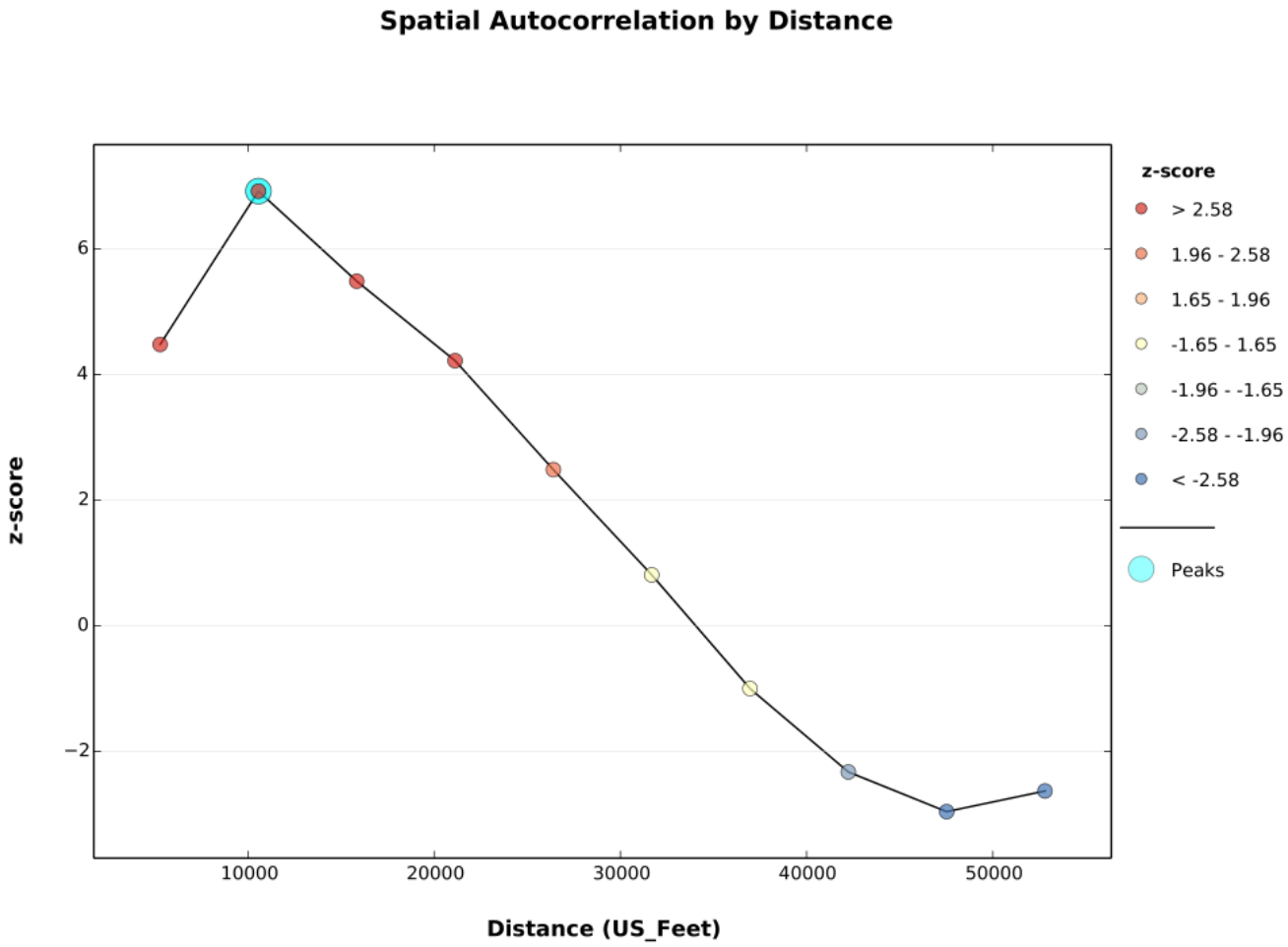

Distribution of Graffiti Reports in Seattle
(Oct 2009 - Oct 2010)
Puget Sound
N
Lake Washington
Reported Graffiti Occurrences (By Neighborhood)
Number of incidents
0 - 2
3 - 6
7 - 11
12 - 16
17 - 30
Map by: Jonathan Engelbert
Date: 02/27/2017
0 1 2 4 6 8
Miles
Data Source: Seattle Police Department's Crime Dashboard
http://www.seattle.gov/police/information-and-data/crime-dashboard
Esri, HERE, DeLorme, MapmyIndia, © OpenStreetMap contributors, and the GIS user community

Map #3


Distribution of Credit Card Fraud Reports in Seattle
(Oct 2009 - Oct 2010)
Puget Sound
N
Reported Credit Card Fraud Occurrences (By Neighborhood)
Number of incidents
0 - 2
3 - 4
5 - 8
9 - 13
14 - 24
Map by: Jonathan Engelbert
Date: 02/27/2017
0 1 2 4 6 8
Miles
Data Source: Seattle Police Department's Crime Dashboard
http://www.seattle.gov/police/information-and-data/crime-dashboard
Esri, HERE, DeLorme, MapmyIndia, © OpenStreetMap contributors, and the GIS user community

Incremental Analysis

Incremental Analysis #1: Assault



Global Moran's I Summary by Distance

Distance	Moran's Index	Expected Index	Variance	z-score	p-value
5280.00*	0.409689	-0.011905	0.008861	4.478795	0.000008
10560.00	0.294141	-0.011236	0.001948	6.919105	0.000000
15840.00	0.149843	-0.011236	0.000862	5.485184	0.000000
21120.00	0.080104	-0.011236	0.000469	4.219857	0.000024
26400.00	0.029723	-0.011236	0.000272	2.485599	0.012933
31680.00	-0.000864	-0.011236	0.000163	0.811195	0.417254
36960.00	-0.021437	-0.011236	0.000104	-0.998741	0.317920
42240.00	-0.030849	-0.011236	0.000071	-2.325928	0.020022
47520.00	-0.031725	-0.011236	0.000048	-2.957839	0.003098
52800.00	-0.026346	-0.011236	0.000033	-2.628692	0.008571

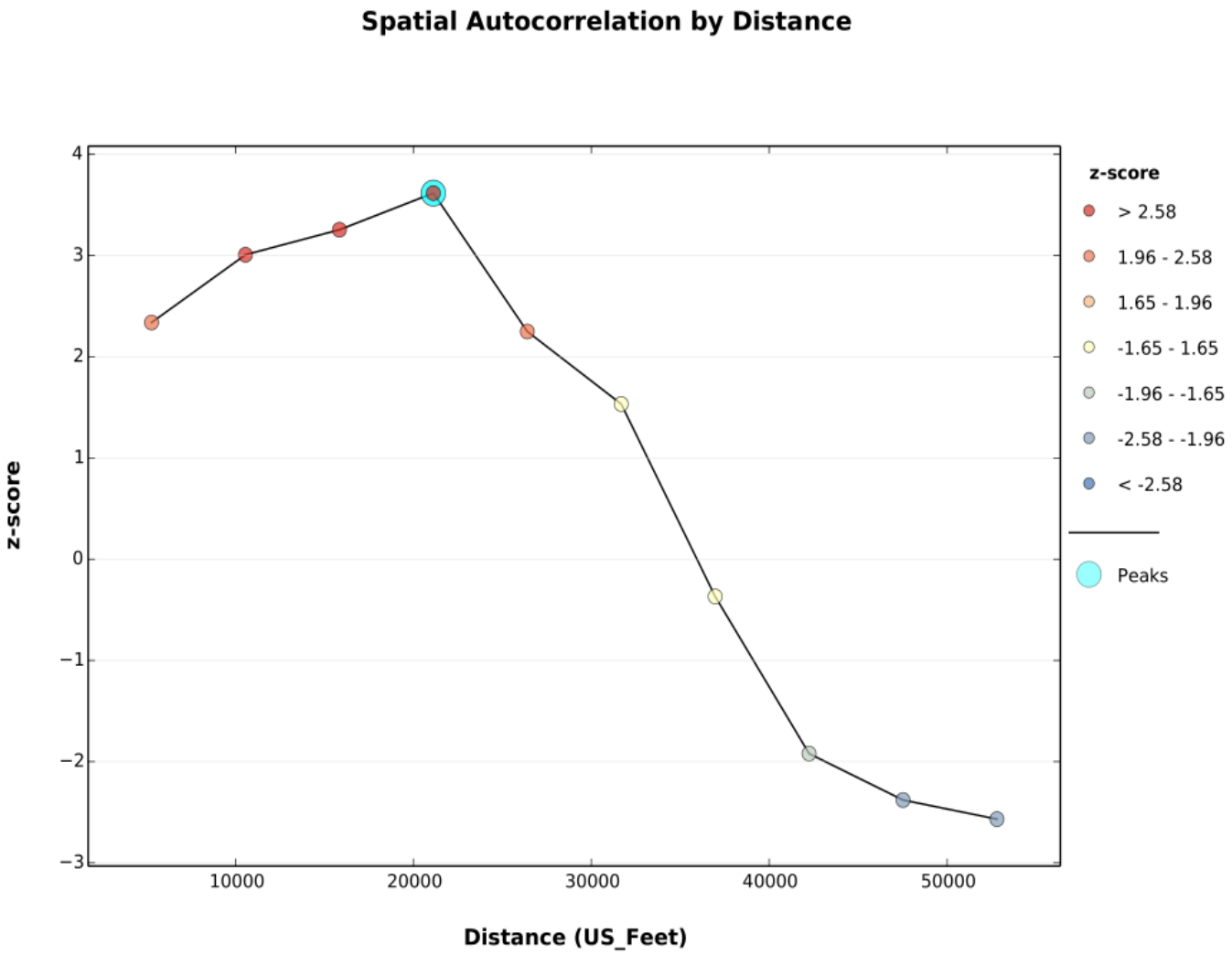
First Peak (Distance, Value): 10560.00, 6.919105

Max Peak (Distance, Value): 10560.00, 6.919105

Distance measured in US_Feet

* At least one distance increment resulted in features with no neighbors which may invalidate the significance of the corresponding results.

Incremental Analysis #2: Graffiti



Global Moran's I Summary by Distance

Distance	Moran's Index	Expected Index	Variance	z-score	p-value
5280.00*	0.209850	-0.011905	0.008990	2.338859	0.019343
10560.00	0.123250	-0.011236	0.001999	3.008102	0.002629
15840.00	0.085603	-0.011236	0.000885	3.255479	0.001132
21120.00	0.068066	-0.011236	0.000481	3.617276	0.000298
26400.00	0.026316	-0.011236	0.000279	2.250110	0.024442
31680.00	0.008620	-0.011236	0.000168	1.533557	0.125139
36960.00	-0.015043	-0.011236	0.000107	-0.368125	0.712780
42240.00	-0.027626	-0.011236	0.000073	-1.920599	0.054782
47520.00	-0.027912	-0.011236	0.000049	-2.379364	0.017343
52800.00	-0.026171	-0.011236	0.000034	-2.568036	0.010228

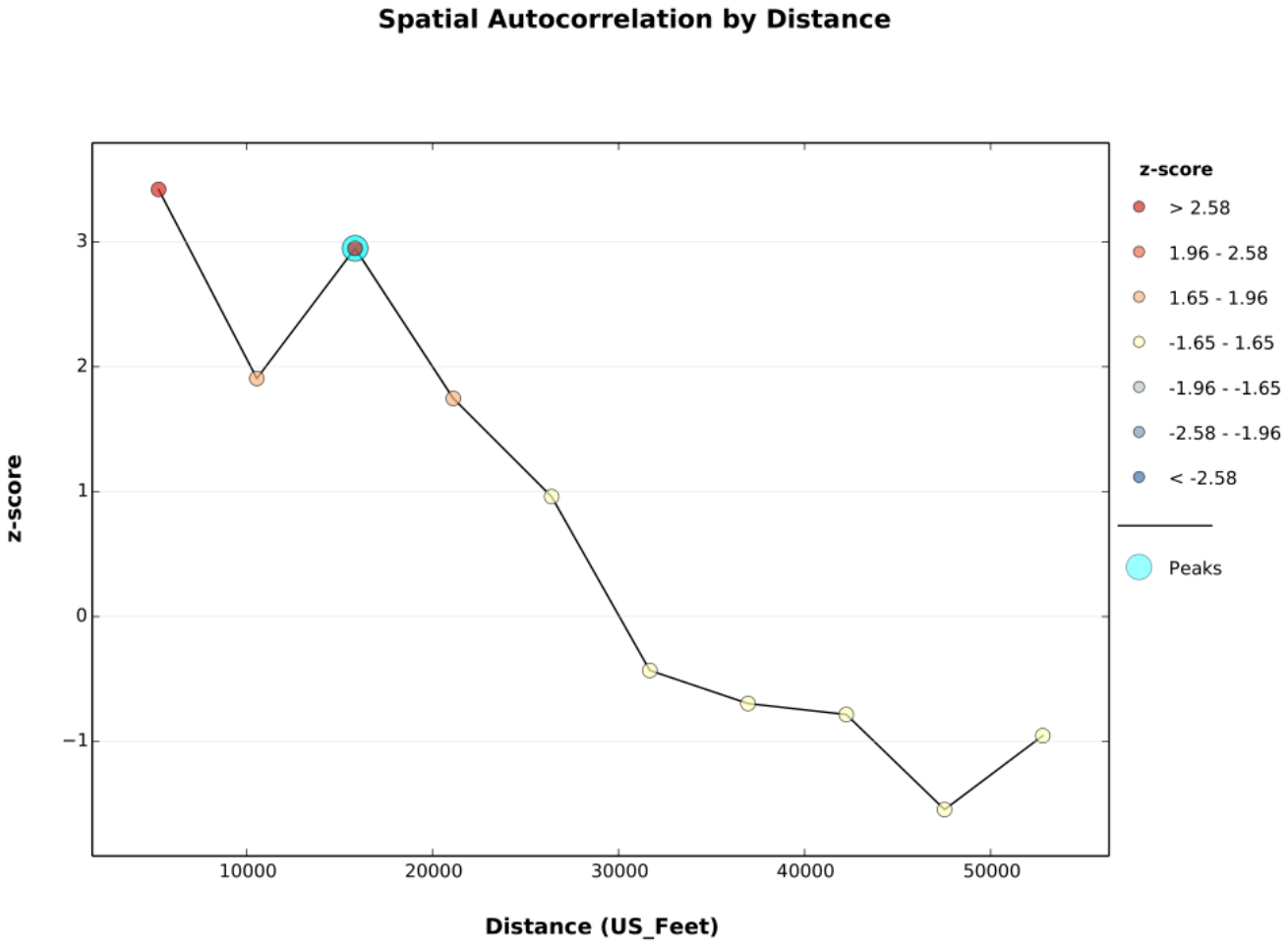
First Peak (Distance, Value): 21120.00, 3.617276

Max Peak (Distance, Value): 21120.00, 3.617276

Distance measured in US_Feet

* At least one distance increment resulted in features with no neighbors which may invalidate the significance of the corresponding results.

Incremental Analysis #3: Credit Card Fraud



Global Moran's I Summary by Distance

Distance	Moran's Index	Expected Index	Variance	z-score	p-value
5280.00*	0.329774	-0.011905	0.009982	3.419816	0.000627
10560.00	0.078021	-0.011236	0.002193	1.906187	0.056626
15840.00	0.080598	-0.011236	0.000970	2.947981	0.003199
21120.00	0.028845	-0.011236	0.000527	1.746370	0.080747
26400.00	0.005536	-0.011236	0.000305	0.960215	0.336947
31680.00	-0.017080	-0.011236	0.000184	-0.431349	0.666215
36960.00	-0.018763	-0.011236	0.000117	-0.696241	0.486278
42240.00	-0.018223	-0.011236	0.000079	-0.784172	0.432939
47520.00	-0.022531	-0.011236	0.000053	-1.544892	0.122372
52800.00	-0.017021	-0.011236	0.000037	-0.953719	0.340226

First Peak (Distance, Value): 15840.00, 2.947981

Max Peak (Distance, Value): 15840.00, 2.947981

Distance measured in US_Feet

* At least one distance increment resulted in features with no neighbors which may invalidate the significance of the corresponding results.

Moran's And Getis-Ord Results

Table 1

Assault	Moran's I	Z score	P-value	Dispersed, Random or Clustered
contiguity edges only	0.5134	8.4489	0.0001	Clustered
contiguity edges and corners	0.4874	8.2858	0.0001	Clustered
Inverse distance	0.1235	10.4030	0.0001	Clustered
Inverse distance squared	0.4185	10.4914	0.0001	Clustered

Table 2

Assault	General G	Z score	P-value	Clustering of High or Low
contiguity edges only	0.1162	8.6383	0.0001	High clustering
contiguity edges and corners	0.1189	8.2199	0.0001	High clustering
Inverse distance	0.0000	9.1724	0.0001	High clustering
Inverse distance squared	0.0000	8.8959	0.0001	High clustering

Table3

Graffiti	Moran's I	Z score	P-value	Dispersed, Random or Clustered
contiguity edges only	0.4282	6.9871	0.0001	Clustered
contiguity edges and corners	0.4053	6.8349	0.0001	Clustered
Inverse distance	0.2846	3.7409	0.0002	Clustered
Inverse distance squared	0.2708	3.0053	0.0026	Clustered

Table 4

Graffiti	General G	Z score	P-value	Clustering of
				High or Low
contiguity edges only	0.1167	7.6275	0.0001	High clustering
contiguity edges and corners	0.1177	7.0711	0.0001	High clustering
Inverse distance	0.0000	4.2790	0.0001	High clustering
Inverse distance squared	0.0000	3.5489	0.0001	High clustering

Table 5

Credit Card Fraud	Moran's I	Z score	P-value	Dispersed, Random
				or Clustered
contiguity edges only	0.2324	3.7000	0.0001	Clustered
contiguity edges and corners	0.2149	3.5440	0.0003	Clustered
Inverse distance	0.2694	3.3924	0.0006	Clustered
Inverse distance squared	0.3212	3.3860	0.0007	Clustered

Table 6

Credit Card Fraud	General G	Z score	P-value	Clustering of
				High or Low
contiguity edges only	0.0690	4.9877	0.0001	High clustering
contiguity edges and corners	0.0717	4.6370	0.0001	High clustering
Inverse distance	0.0000	3.0781	0.0020	High clustering
Inverse distance squared	0.0000	2.5757	7.0000	High clustering