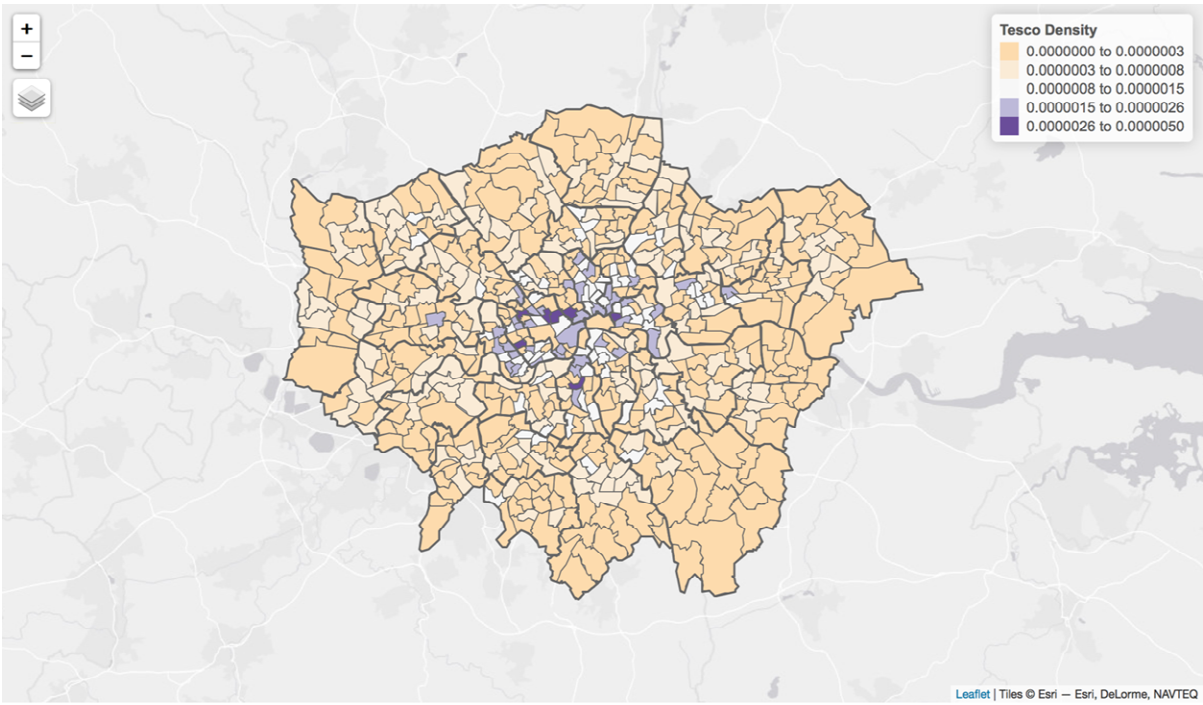
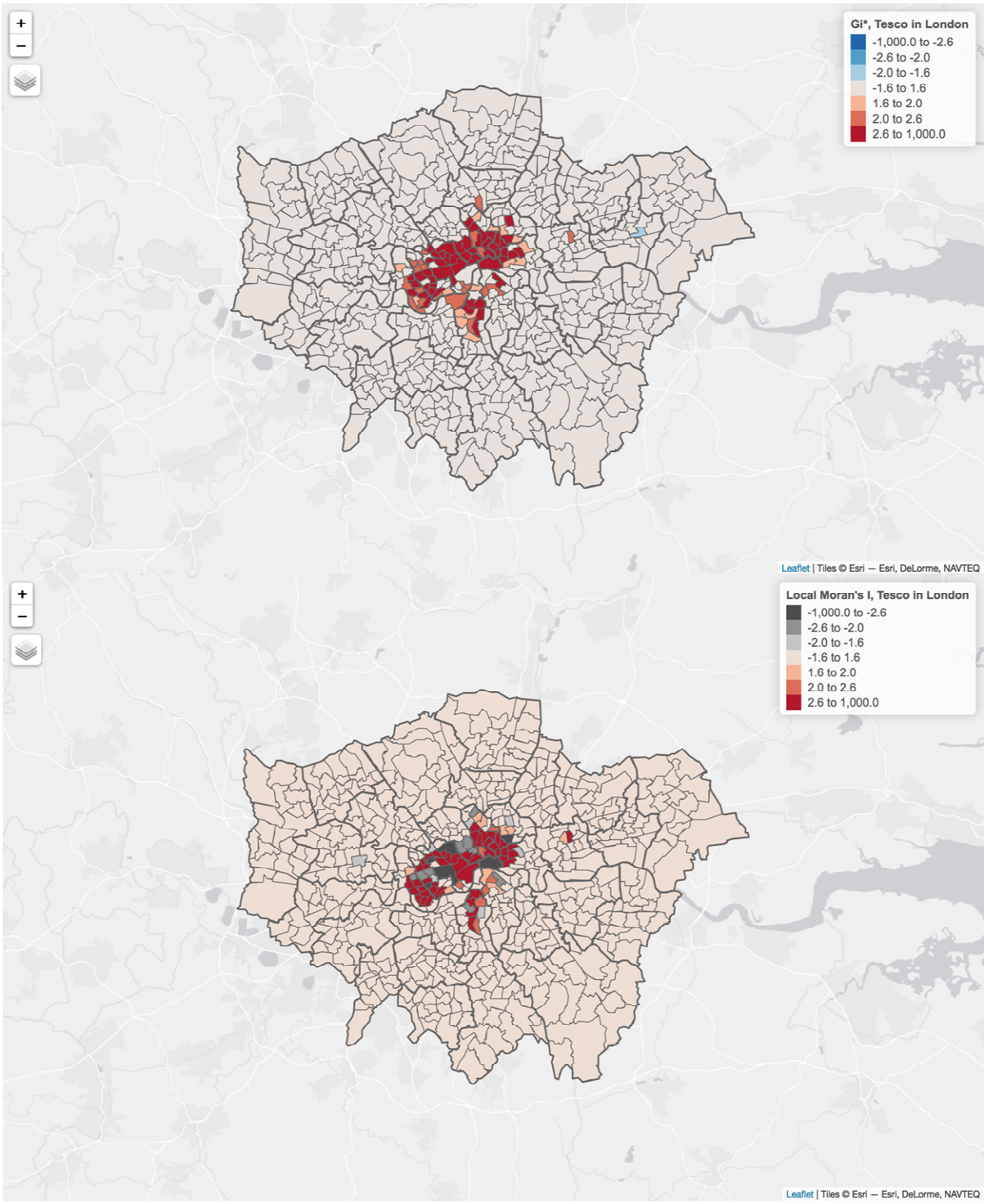
Part 1 - 2 Maps + 600 Word Commentary

Map Making Procedure

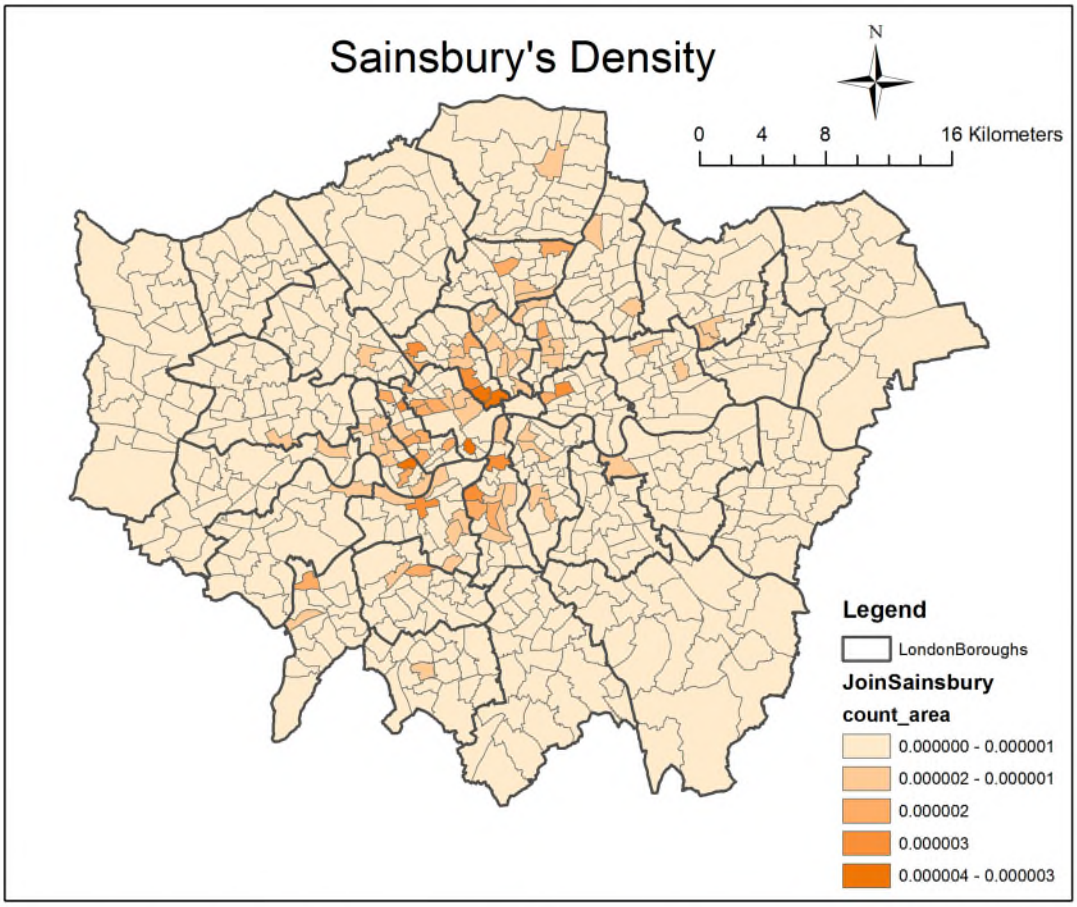
Two maps are made in this part: one is for the map of the density of Tesco supermarkets in different wards of London and its spatial autocorrelation created by R Studio, and the other is the same kind of map produced by ArcGIS but for Sainsbury's to make a comparison.

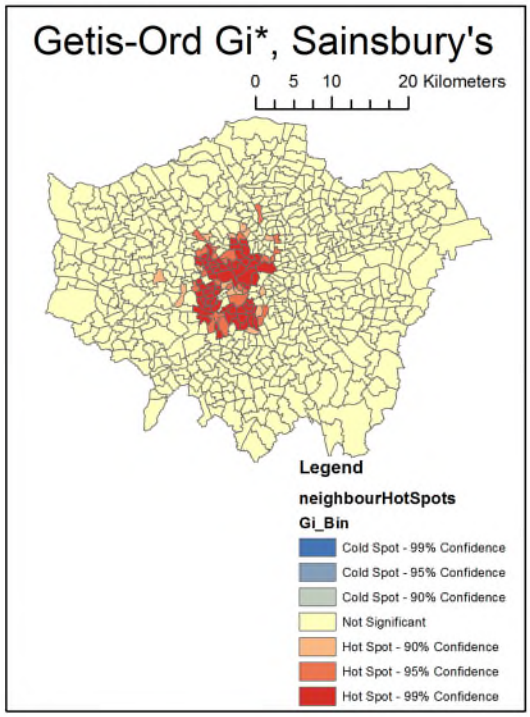
The data of the postcodes of Tesco and Sainsbury's are from allinlondon.co.uk, and matched with the coordination of the Great London Postcodes from doogal.co.uk to get a csv. When dragging the location csv of Sainsbury's into ArcMap to display XY data, we have to choose WGS 1984 as the coordinate system at first as many spatial data from nongovernment websites are based on WGS 1984 and export Events as layers. To match the coordination of the wards map as the base map, we need to use "Project" tool to transfer. As I haven't download the OSTN02 data file from OS and needn't very accurate transformation, I change the default Geographic Transformation from OSGB_1936_TO_WGS_1984_7 to OSGB_1936_TO_WGS_1984_Petroleum which has an accuracy of +/- 5m, and in R, we also need string like BNG = "+init=epsg:27700" to make every data's coordinate system matched. For ArcGIS, I choose to use spatial join to count the number of Sainsbury's falls within every wards, and open attribute table of the joined layer to add a new field to calculate the density by Join_Count/Shape_Area, which is a little complex compared with R, which only need to type three lines by using GISTool package. Then to calculate the Local Morans I and Getis-Ord Gi* in R, the codes provided in practical 6 are reusable, while in ArcGIS, the toolbox model built in practical 6 for ArcGIS fans is also reusable. Both Local Moran's I of Tesco and Sainsbury's show that there are many clusters (which are very similar with the neighbours) while also many outliners (which are very different with the neighbours) in the center of London. However, it was probably caused by MAUP. The number of supermarkets in wards are small resulting in the appearance of outliners easily. If compared by boroughs, the result maybe more accurate.

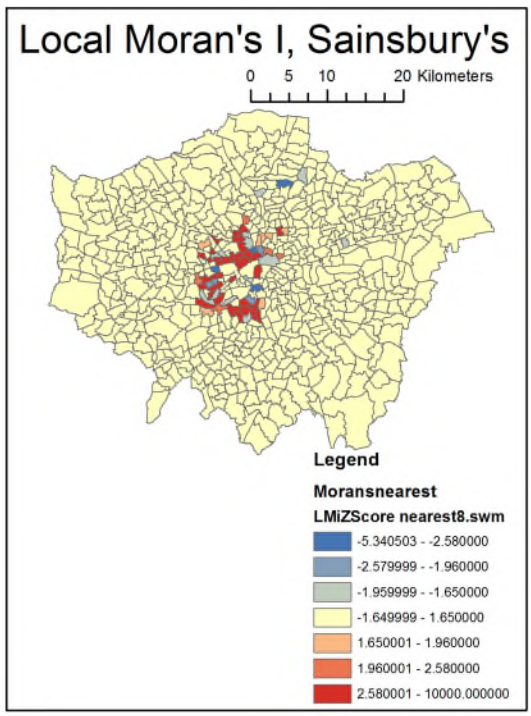




Map made by R Studio







Map made by ArcGIS

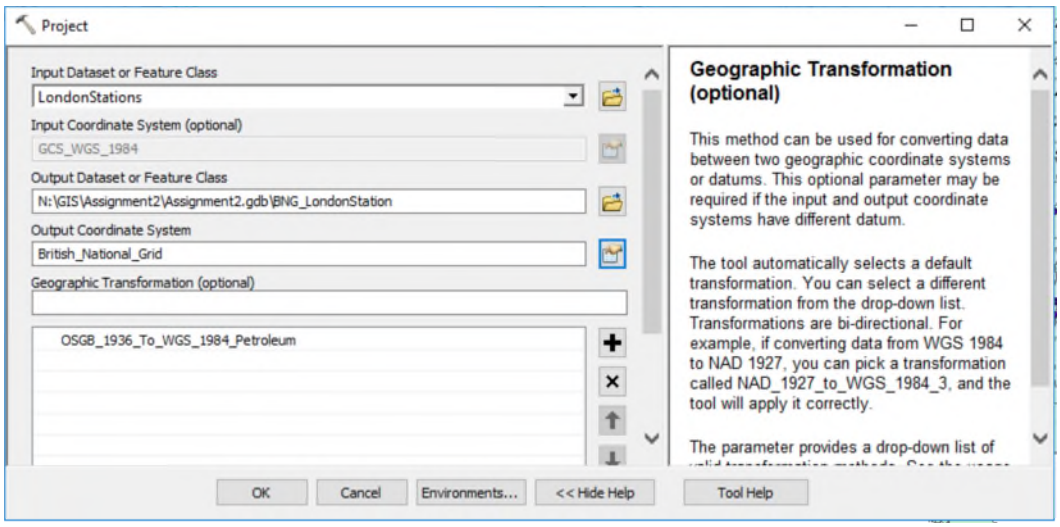
ArcGIS vs R Studio

Based on my own experience through using both ArcGIS and R Studio, I think ArcGIS is much more user-friendly and convenient for programming novices as it has systematically pre-created tools and technical support. Even if you are new for ArcGIS, you can speculate the function of each button from its name and icon, and you can find how to use this function and its principle very detailed by just searching in ArcGIS Resource website which provide the complete instructions in many different kinds of language. While for R Studio, you may find it's really hard to learn it systematically as it has countless packages and the code and logic within them are always different. On the other hand, R Studio has many advantages, for example, it's more customized, scalable and open as it's free. If you want to do more advanced and complex analyze beyond ArcGIS's frame, you can write it in R Studio by yourself and there are so many hobby friends on the Internet to talk and share codes. Also, I think ArcGIS is more suitable to make traditional static map and the layout view is quite handy for typesetting, whilst R Studio is expert in making interactive maps. As we become more familiar with packages and codes, I think we will use R Studio more as it's freer and typing codes sometimes is quicker than clicking mouse!

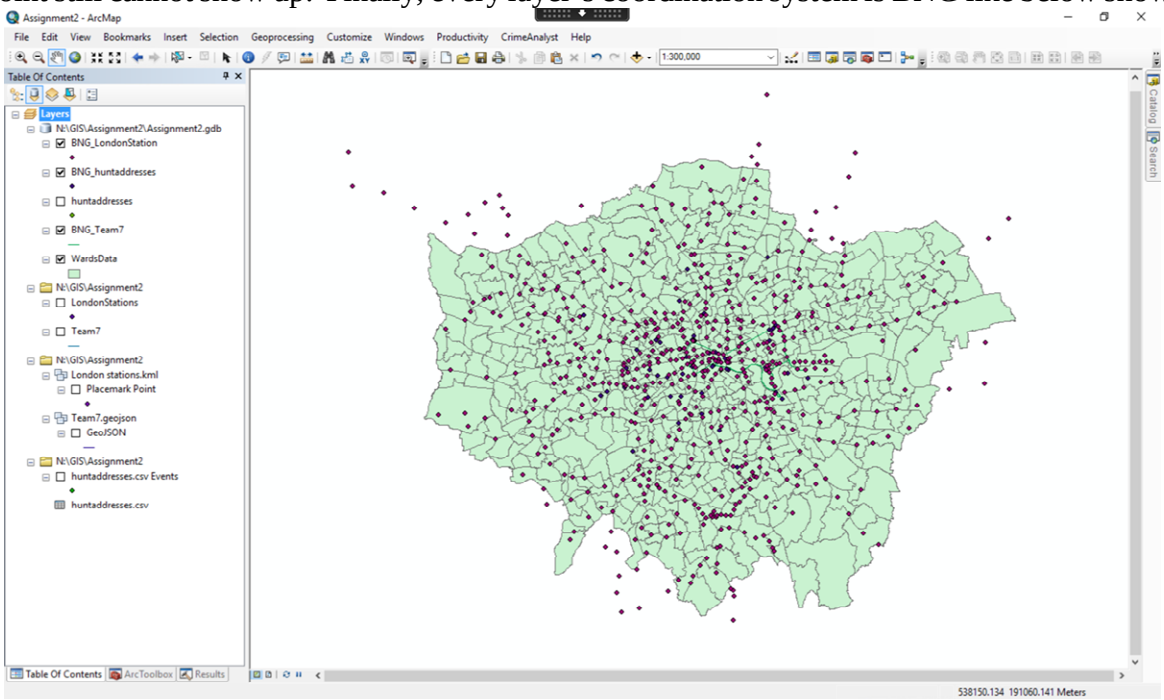
Part 2 - Spatial Analysis Challenge

I choose to use ArcGIS because when I read these questions, I can immediately figure out how to solve them by ArcGIS, but for R, it may take a long time to search for the suitable packages and codes.

First of all, drag LondonWardsNew.shp into ArcMap as the base map and check the coordination system of it is BNG, and merge the city of London wards into a single zone as appendix says. Then, drag Placemark Point in London stations.kml into the map and export data as shapefile. When check the coordination system of it, find it was WGS 1984, so use "project" tool to transfer as below shows.



The same procedure to deal with the hunt address and Team7 trace, though there is a mistake in huntaddresses.csv. The location of Platform 9 & ¾ is wrong! Change it in the csv before "display XY data" because if changing it in the layer attribute table through "edit" directly, the point still cannot show up. Finally, every layer's coordination system is BNG like below shows.



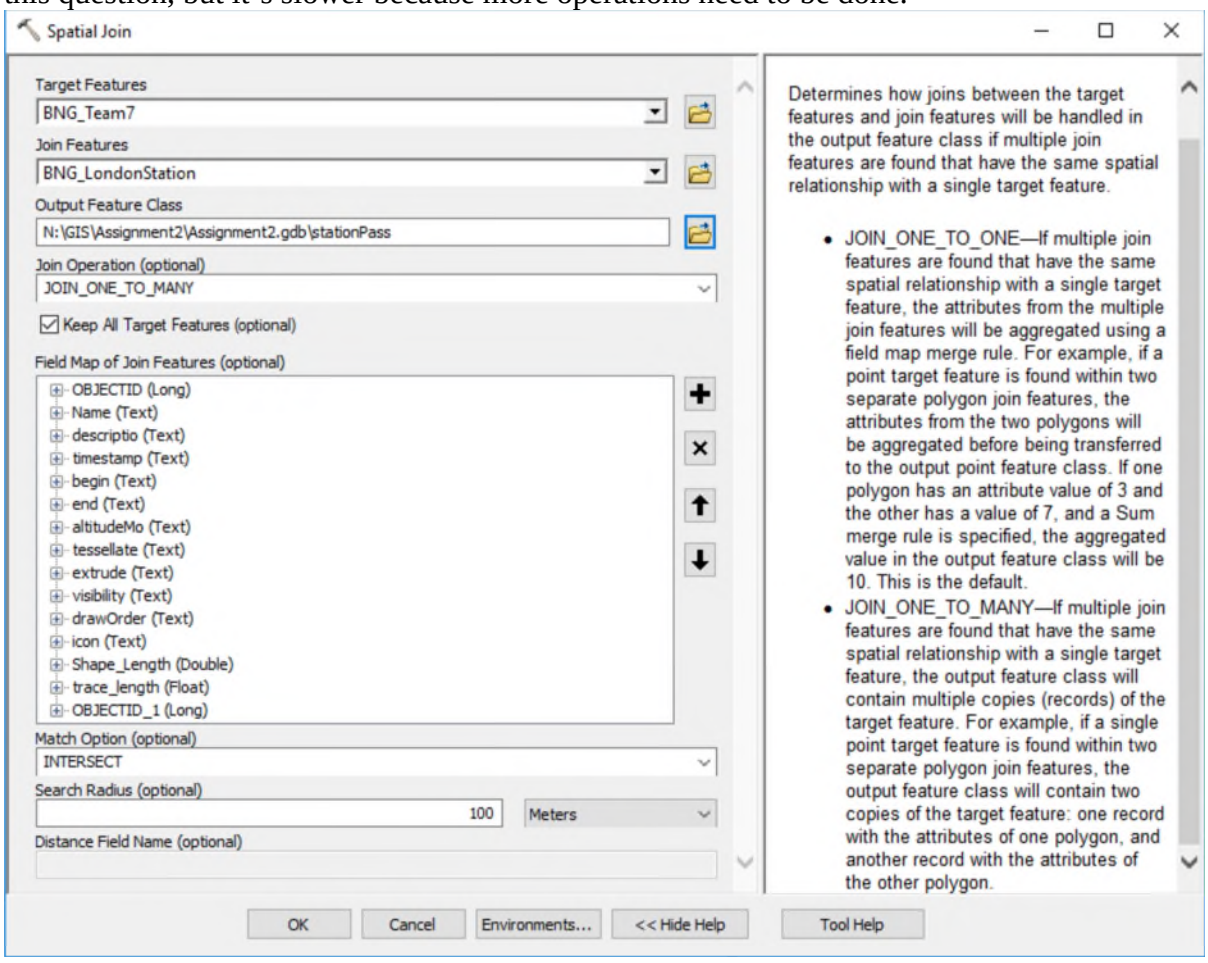
Question 1: to answer this question, right click BNG_Team7 layer and choose “open attribute table”, and in the table, we will find the trace length has been calculated automatically as “Shape_Length” shows when doing the projection transformation. But for double check, add a new field in the table called “trace_length” and right click the header and choose “Calculate Geometry” to calculate the length, and the result is the same – 46603.95m.

Table

BNG_Team7

end	altitudeMo	tessellate	extrude	visibility	drawOrder	icon	Shape_Length	trace_length
		1	0	-1			46603.953812	46603.95

Question 2: to get answer more quickly, use “Spatial Join” tool in ArcToolbox -> Analysis Tools -> Overlay, set “Search Radius” as 100 meters so that it can include the stations within 100 meters from the trace. We can set “Join Operation” as JOIN_ONE_TO_MANY, if we want the information of every joined station, while we can also choose JOIN_ONE_TO_ONE, if we only need to know the counts. We can also use “buffer” and then “intersect” to answer this question, but it’s slower because more operations need to be done.



Table

station_pass

OBJECTID_1 *	Shape *	Join_Count	TARGET_FID	JOIN_FID	OBJECTID	Name	descriptio	t
1	Polyline	1	1	1	1	Casa team 7	2016 treasure hunt	
2	Polyline	1	1	51	1	Casa team 7	2016 treasure hunt	
3	Polyline	1	1	52	1	Casa team 7	2016 treasure hunt	
4	Polyline	1	1	92	1	Casa team 7	2016 treasure hunt	
5	Polyline	1	1	93	1	Casa team 7	2016 treasure hunt	
6	Polyline	1	1	145	1	Casa team 7	2016 treasure hunt	
7	Polyline	1	1	150	1	Casa team 7	2016 treasure hunt	
8	Polyline	1	1	273	1	Casa team 7	2016 treasure hunt	
9	Polyline	1	1	317	1	Casa team 7	2016 treasure hunt	
10	Polyline	1	1	332	1	Casa team 7	2016 treasure hunt	
11	Polyline	1	1	339	1	Casa team 7	2016 treasure hunt	
12	Polyline	1	1	351	1	Casa team 7	2016 treasure hunt	
13	Polyline	1	1	365	1	Casa team 7	2016 treasure hunt	
14	Polyline	1	1	373	1	Casa team 7	2016 treasure hunt	
15	Polyline	1	1	421	1	Casa team 7	2016 treasure hunt	
16	Polyline	1	1	427	1	Casa team 7	2016 treasure hunt	
17	Polyline	1	1	491	1	Casa team 7	2016 treasure hunt	
18	Polyline	1	1	506	1	Casa team 7	2016 treasure hunt	
19	Polyline	1	1	513	1	Casa team 7	2016 treasure hunt	
20	Polyline	1	1	519	1	Casa team 7	2016 treasure hunt	
21	Polyline	1	1	520	1	Casa team 7	2016 treasure hunt	
22	Polyline	1	1	604	1	Casa team 7	2016 treasure hunt	
23	Polyline	1	1	612	1	Casa team 7	2016 treasure hunt	
24	Polyline	1	1	618	1	Casa team 7	2016 treasure hunt	

1 (0 out of 24 Selected)

station_pass

The Attribute Table When choosing JOIN_ONE_TO_MANY

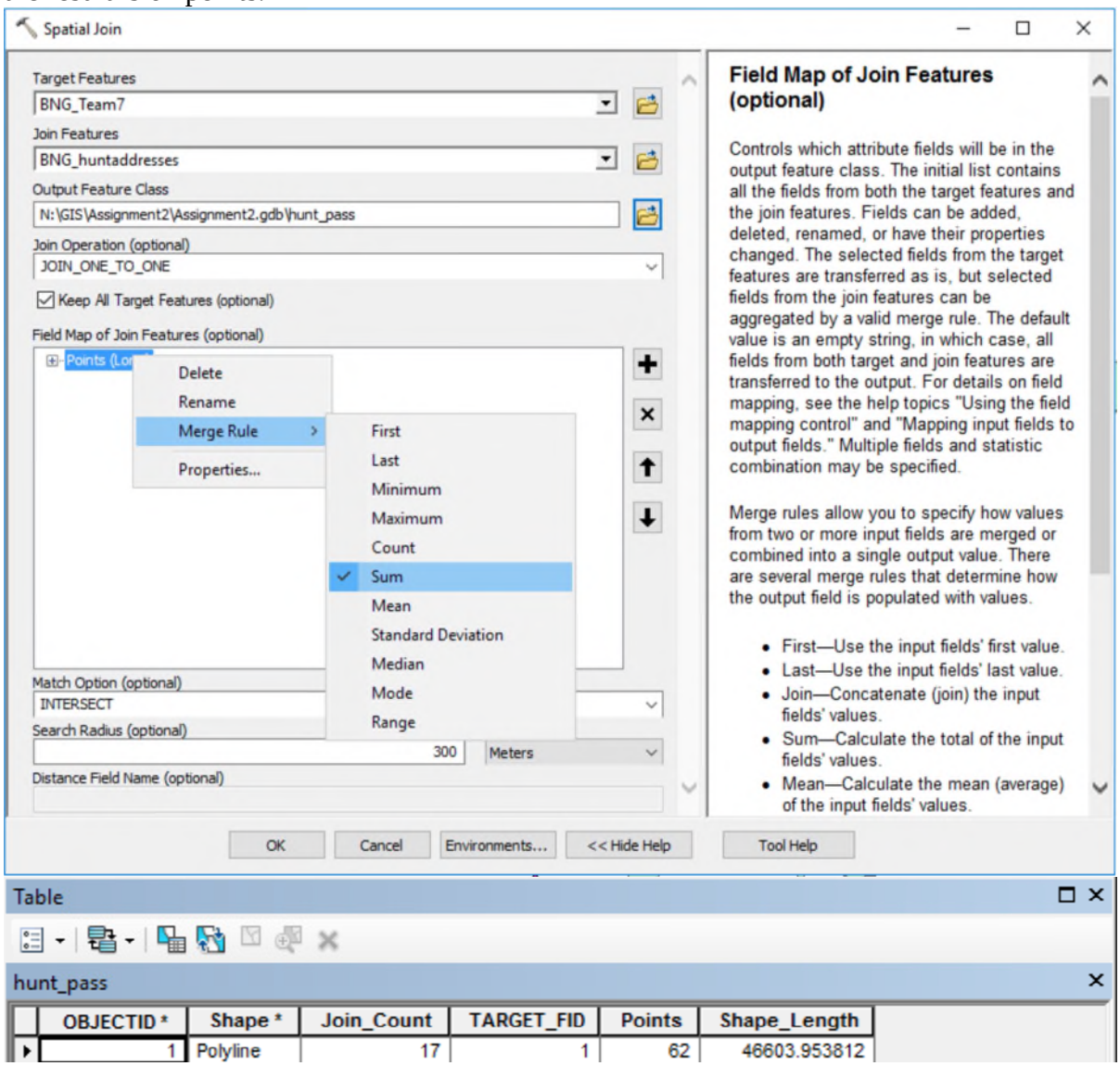
Table

station_onepass

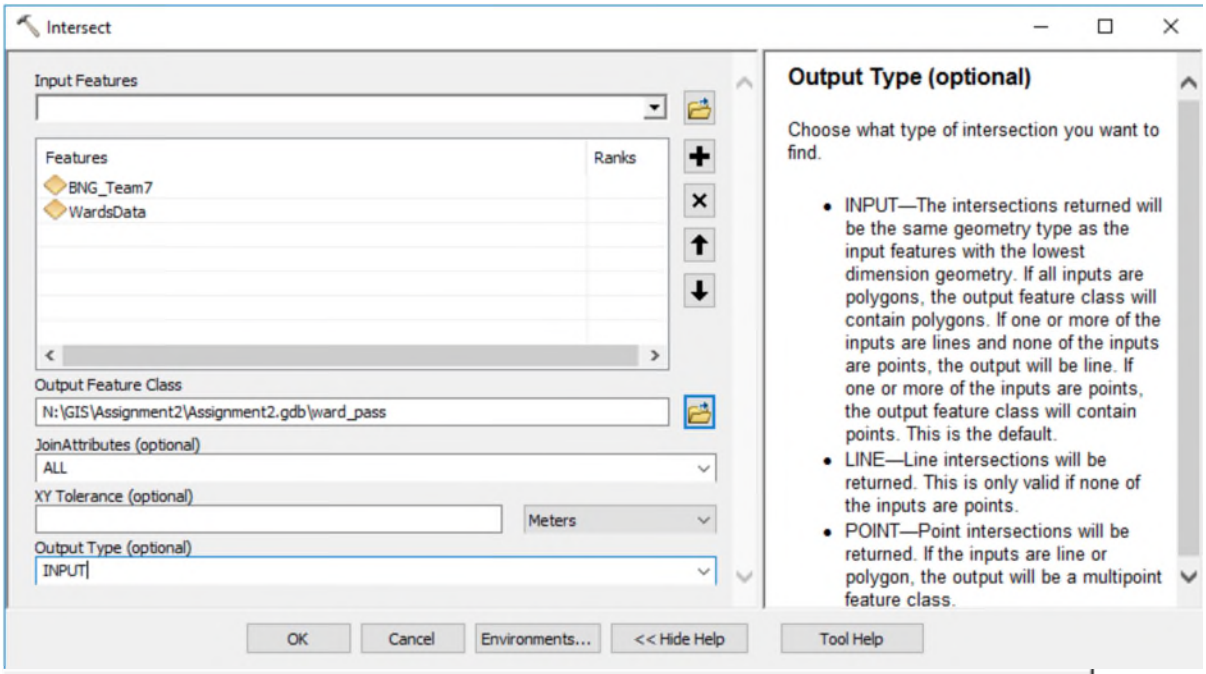
OBJECTID_1 *	Shape *	Join_Count	TARGET_FID	OBJECTID	Name	descriptio	timestamp
1	Polyline	24	1	1	Casa team 7	2016 treasure hunt	

The Attribute Table When choosing JOIN_ONE_TO_ONE

Question 3: The briefest and quickest way is using “spatial join” and setting like below, and the result is 62 points.



Question 4: use “intersect” in processing, then open the intersected layer’s attribute table, use sort ascending from low to high to male life expectancy column, and the first line shows Weavers had the lowest while city of London had the highest.


Intersect
Input Features
Features
Ranks
BNG_Team7
WardsData
Output Feature Class
N:\GIS\Assignment2\Assignment2.gdb\ward_pass
JoinAttributes (optional)
ALL
XY Tolerance (optional)
Meters
Output Type (optional)
INPUT
Output Type (optional)
Choose what type of intersection you want to find.
INPUT—The intersections returned will be the same geometry type as the input features with the lowest dimension geometry. If all inputs are polygons, the output feature class will contain polygons. If one or more of the inputs are lines and none of the inputs are points, the output will be line. If one or more of the inputs are points, the output feature class will contain points. This is the default.
LINE—Line intersections will be returned. This is only valid if none of the inputs are points.
POINT—Point intersections will be returned. If the inputs are line or polygon, the output will be a multipoint feature class.
OK
Cancel
Environments...
<< Hide Help
Tool Help

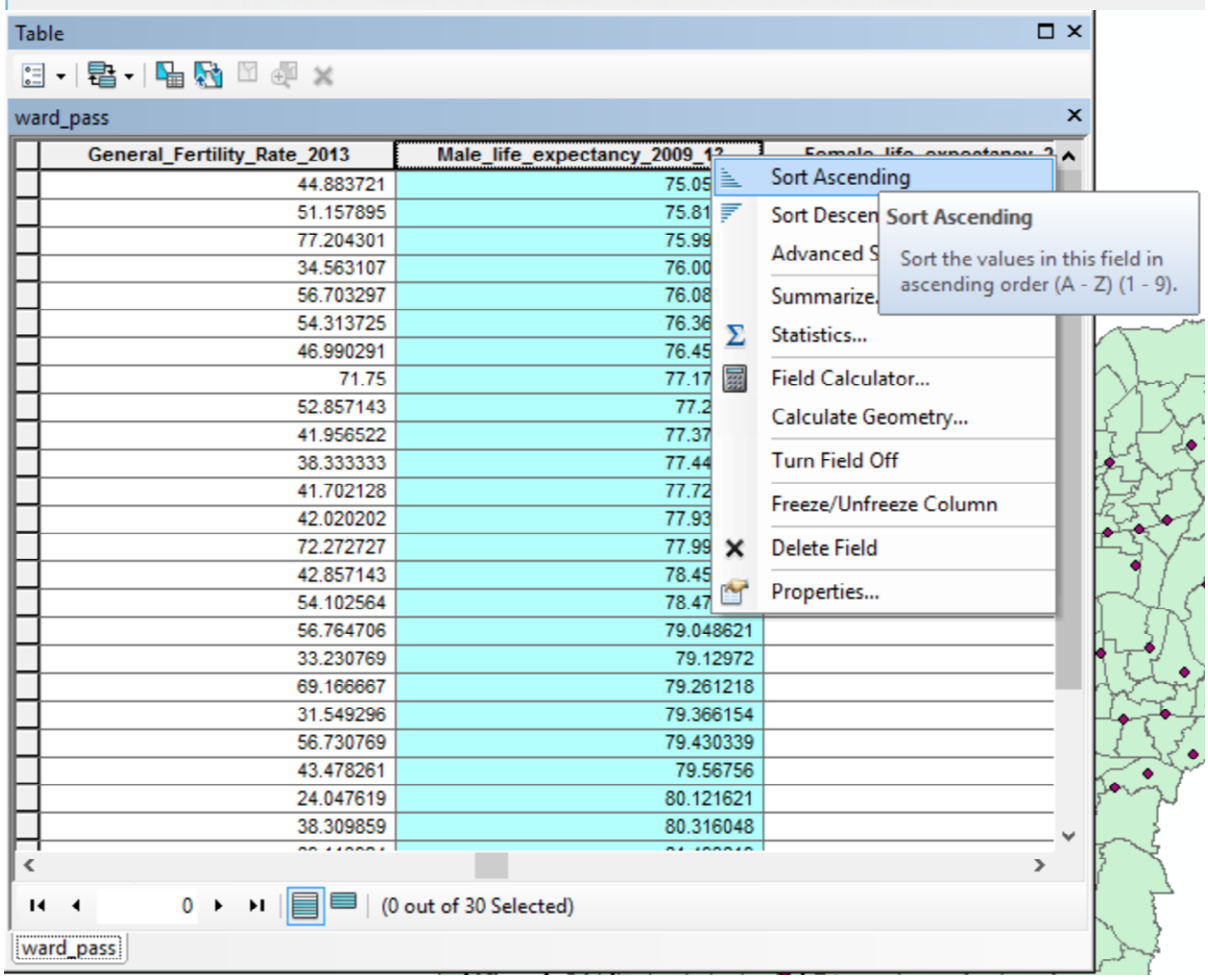


Table
ward_pass
General_Fertility_Rate_2013
Male_life_expectancy_2009_1?
44.883721
75.05
51.157895
75.81
77.204301
75.99
34.563107
76.00
56.703297
76.08
54.313725
76.36
46.990291
76.45
71.75
77.17
52.857143
77.2
41.956522
77.37
38.333333
77.44
41.702128
77.72
42.020202
77.93
72.272727
77.99
42.857143
78.45
54.102564
78.47
56.764706
79.048621
33.230769
79.12972
69.166667
79.261218
31.549296
79.366154
56.730769
79.430339
43.478261
79.56756
24.047619
80.121621
38.309859
80.316048
Sort Ascending
Sort Descen
Advanced S
Summarize...
Statistics...
Field Calculator...
Calculate Geometry...
Turn Field Off
Freeze/Unfreeze Column
Delete Field
Properties...
Sort Ascending
Sort the values in this field in ascending order (A - Z) (1 - 9).
0
(0 out of 30 Selected)
ward_pass

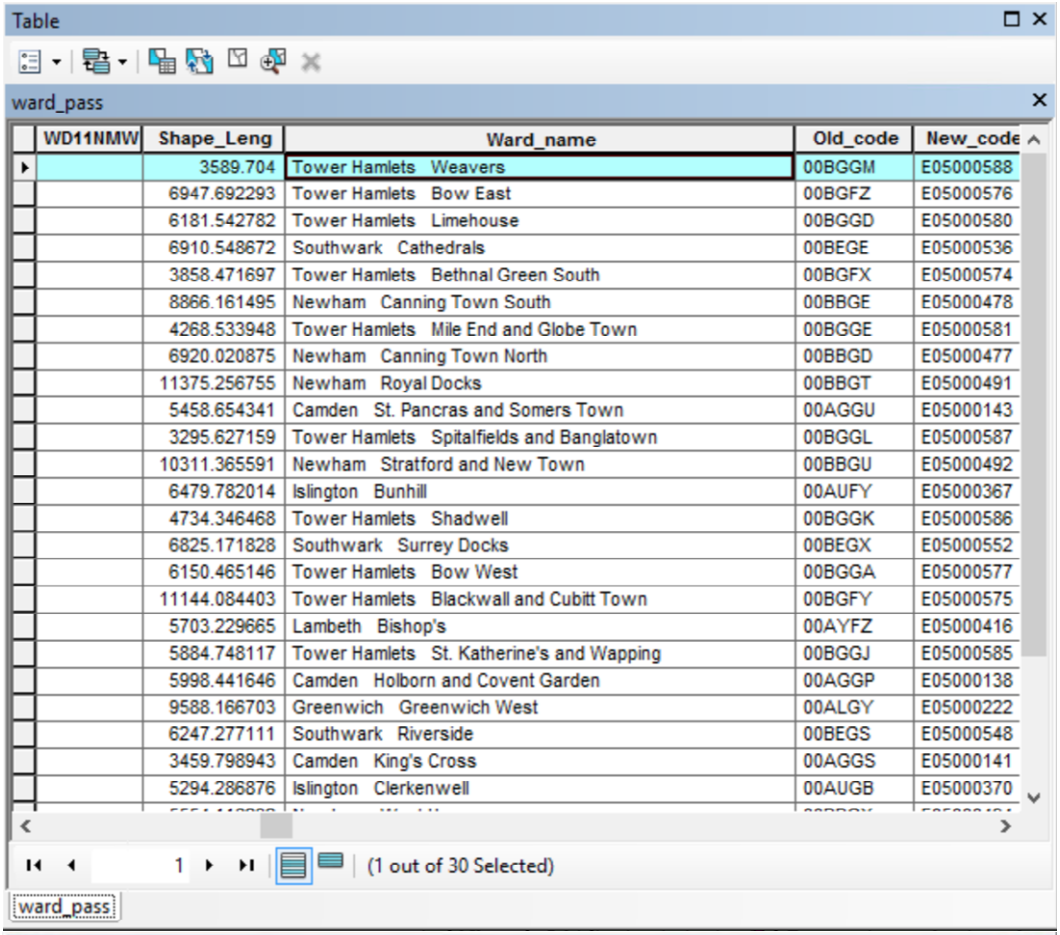


Table
ward_pass
WD11NMW
Shape_Leng
Ward_name
Old_code
New_code
3589.704 Tower Hamlets Weavers 00BGGM E05000588
6947.692293 Tower Hamlets Bow East 00BGFZ E05000576
6181.542782 Tower Hamlets Limehouse 00BGGD E05000580
6910.548672 Southwark Cathedrals 00BEGE E05000536
3858.471697 Tower Hamlets Bethnal Green South 00BGFX E05000574
8866.161495 Newham Canning Town South 00BBGE E05000478
4268.533948 Tower Hamlets Mile End and Globe Town 00BGGE E05000581
6920.020875 Newham Canning Town North 00BBGD E05000477
11375.256755 Newham Royal Docks 00BBGT E05000491
5458.654341 Camden St. Pancras and Somers Town 00AGGU E05000143
3295.627159 Tower Hamlets Spitalfields and Banglatown 00BGGL E05000587
10311.365591 Newham Stratford and New Town 00BBGU E05000492
6479.782014 Islington Bunhill 00AUFY E05000367
4734.346468 Tower Hamlets Shadwell 00BGGK E05000586
6825.171828 Southwark Surrey Docks 00BEGX E05000552
6150.465146 Tower Hamlets Bow West 00BGGA E05000577
11144.084403 Tower Hamlets Blackwall and Cubitt Town 00BGFY E05000575
5703.229665 Lambeth Bishop's 00AYFZ E05000416
5884.748117 Tower Hamlets St. Katherine's and Wapping 00BGGJ E05000585
5998.441646 Camden Holborn and Covent Garden 00AGGP E05000138
9588.166703 Greenwich Greenwich West 00ALGY E05000222
6247.277111 Southwark Riverside 00BEGS E05000548
3459.798943 Camden King's Cross 00AGGS E05000141
5294.286876 Islington Clerkenwell 00AUGB E05000370
1
(1 out of 30 Selected)
ward_pass

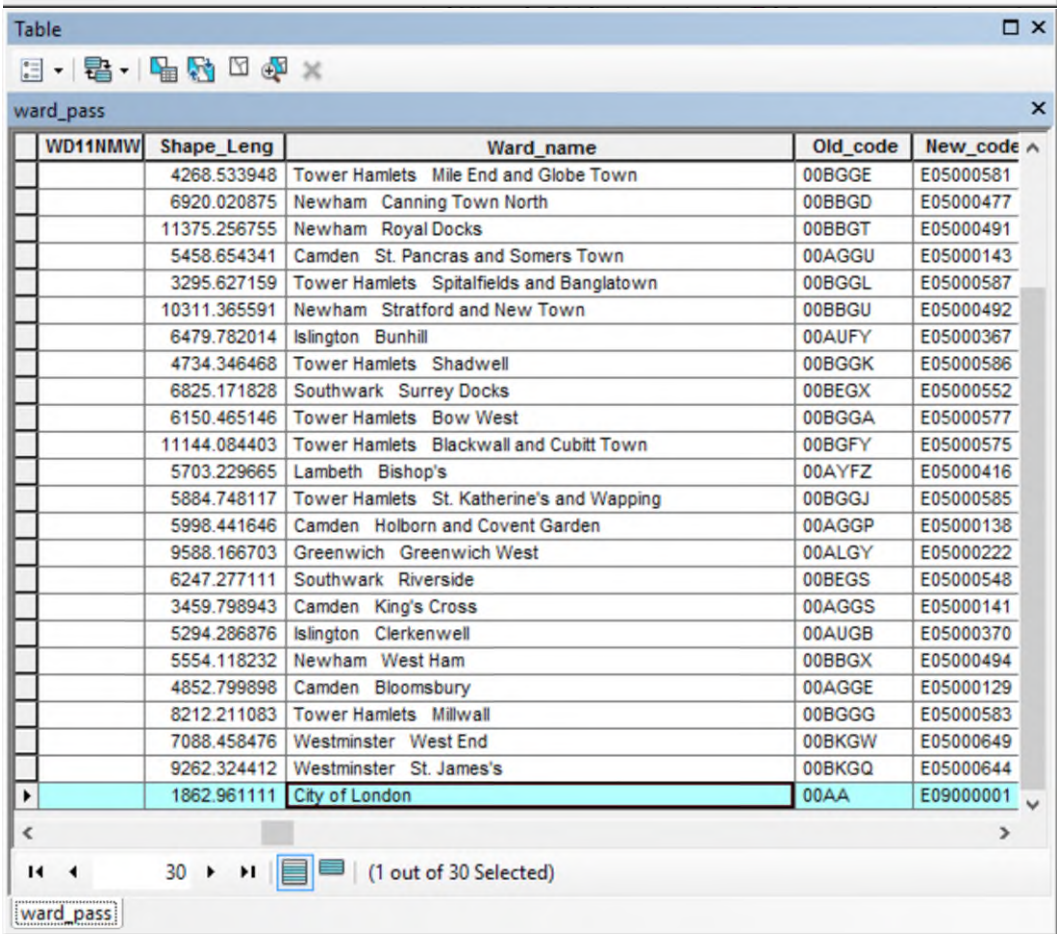
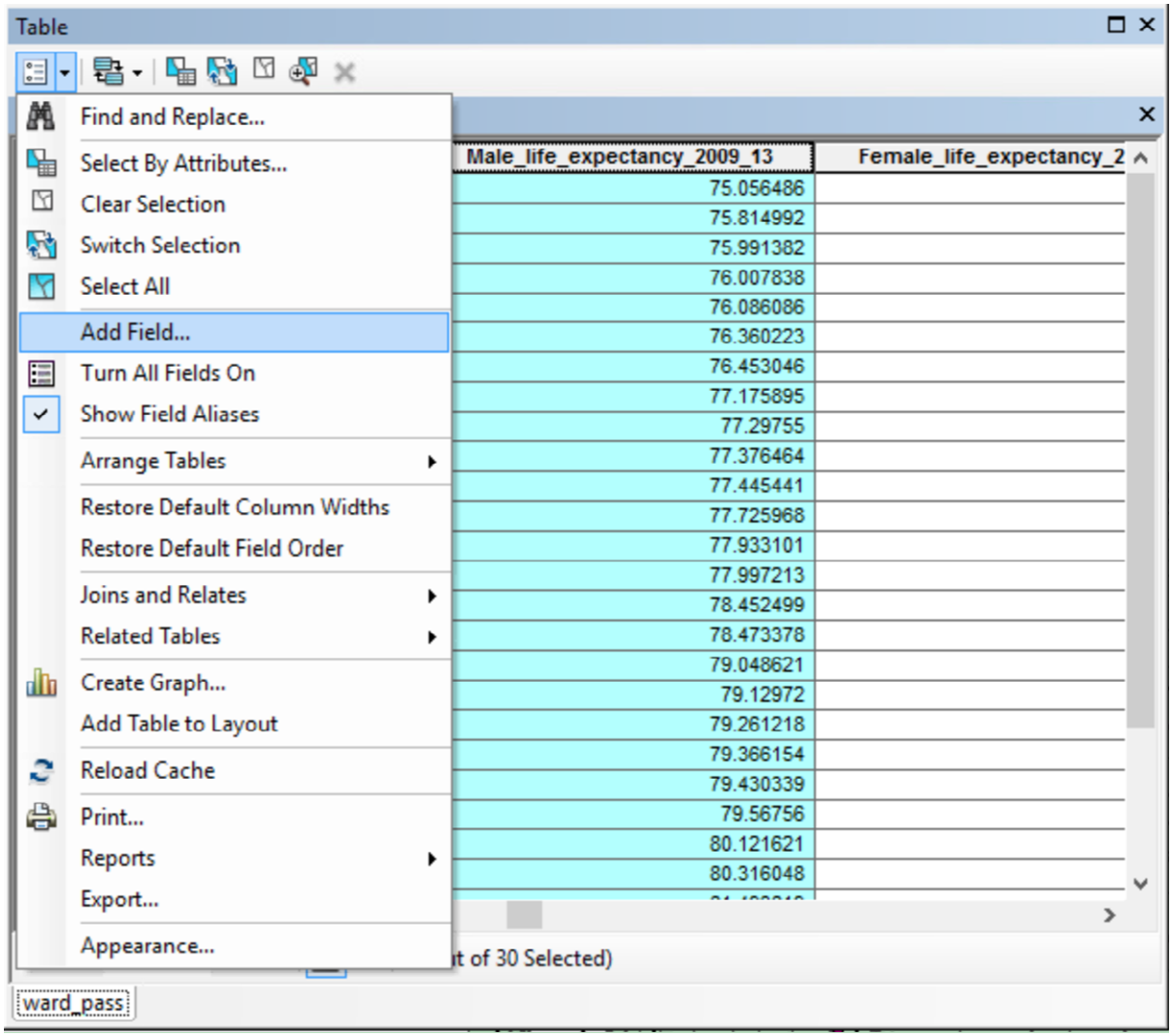


Table
ward_pass
WD11NMW
Shape_Leng
Ward_name
Old_code
New_code
4268.533948 Tower Hamlets Mile End and Globe Town 00BGGE E05000581
6920.020875 Newham Canning Town North 00BBGD E05000477
11375.256755 Newham Royal Docks 00BBGT E05000491
5458.654341 Camden St. Pancras and Somers Town 00AGGU E05000143
3295.627159 Tower Hamlets Spitalfields and Banglatown 00BGGL E05000587
10311.365591 Newham Stratford and New Town 00BBGU E05000492
6479.782014 Islington Bunhill 00AUFY E05000367
4734.346468 Tower Hamlets Shadwell 00BGGK E05000586
6825.171828 Southwark Surrey Docks 00BEGX E05000552
6150.465146 Tower Hamlets Bow West 00BGGA E05000577
11144.084403 Tower Hamlets Blackwall and Cubitt Town 00BGFY E05000575
5703.229665 Lambeth Bishop's 00AYFZ E05000416
5884.748117 Tower Hamlets St. Katherine's and Wapping 00BGGJ E05000585
5998.441646 Camden Holborn and Covent Garden 00AGGP E05000138
9588.166703 Greenwich Greenwich West 00ALGY E05000222
6247.277111 Southwark Riverside 00BEGS E05000548
3459.798943 Camden King's Cross 00AGGS E05000141
5294.286876 Islington Clerkenwell 00AUGB E05000370
5554.118232 Newham West Ham 00BBGX E05000494
4852.799898 Camden Bloomsbury 00AGGE E05000129
8212.211083 Tower Hamlets Millwall 00BGGG E05000583
7088.458476 Westminster West End 00BKGW E05000649
9262.324412 Westminster St. James's 00BKGQ E05000644
1862.961111 City of London 00AA E09000001
30
(1 out of 30 Selected)
ward_pass

Question 5: add new field in intersected layer, and calculate the average of male and female's life expectancy, then view the statistic table and find the mean of all wards' life expectancy is 81.43.



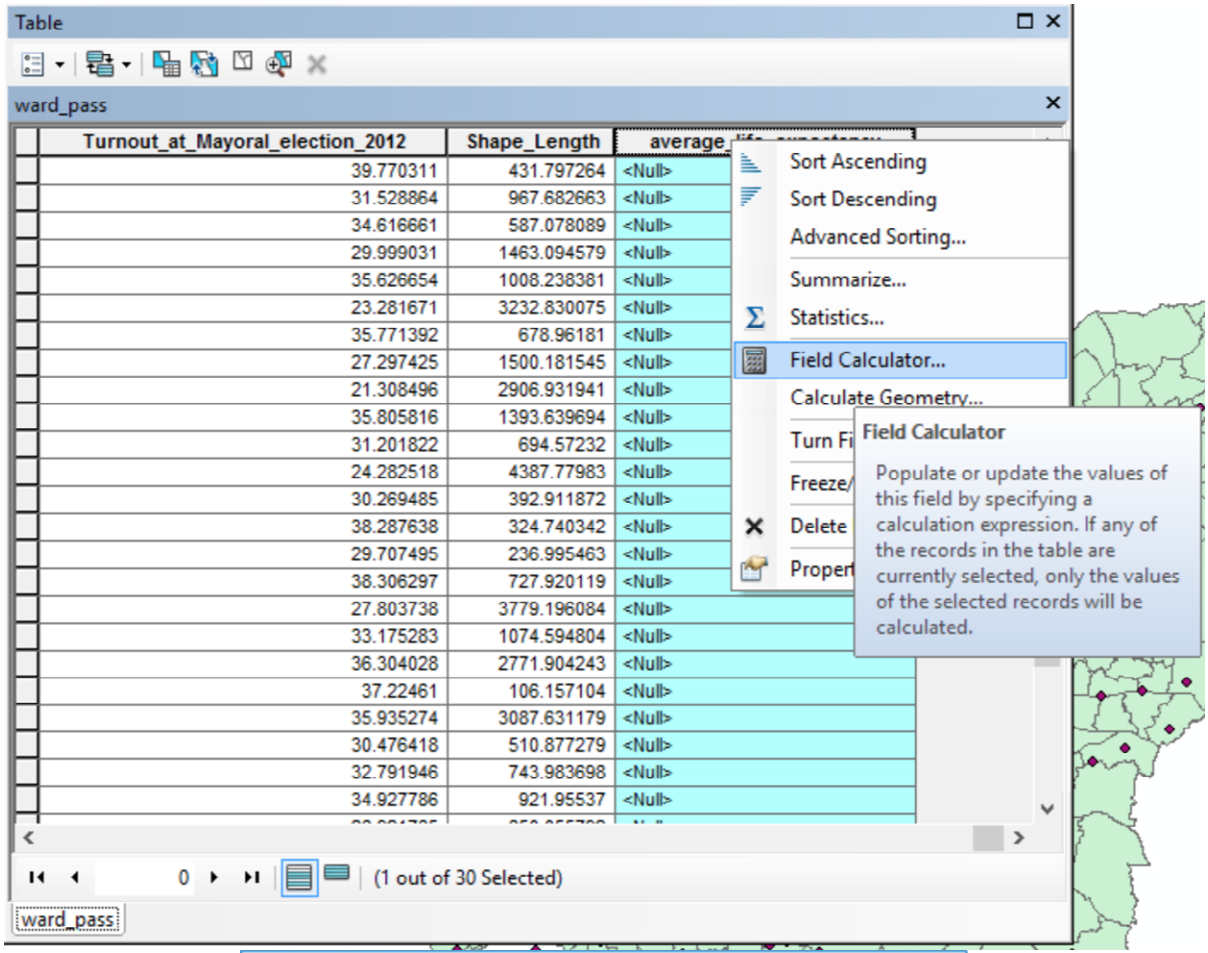
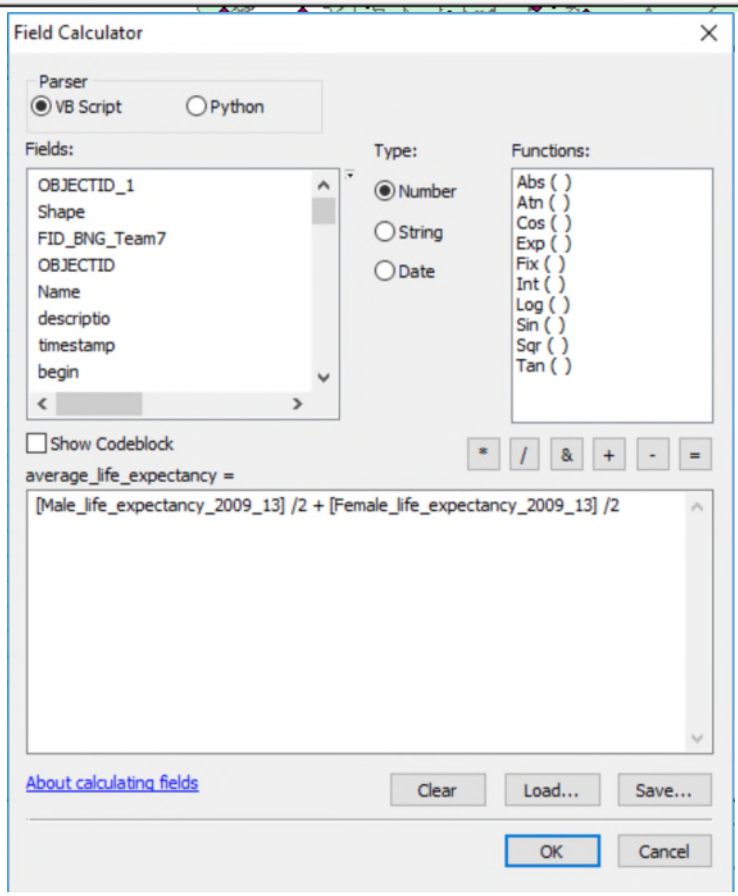
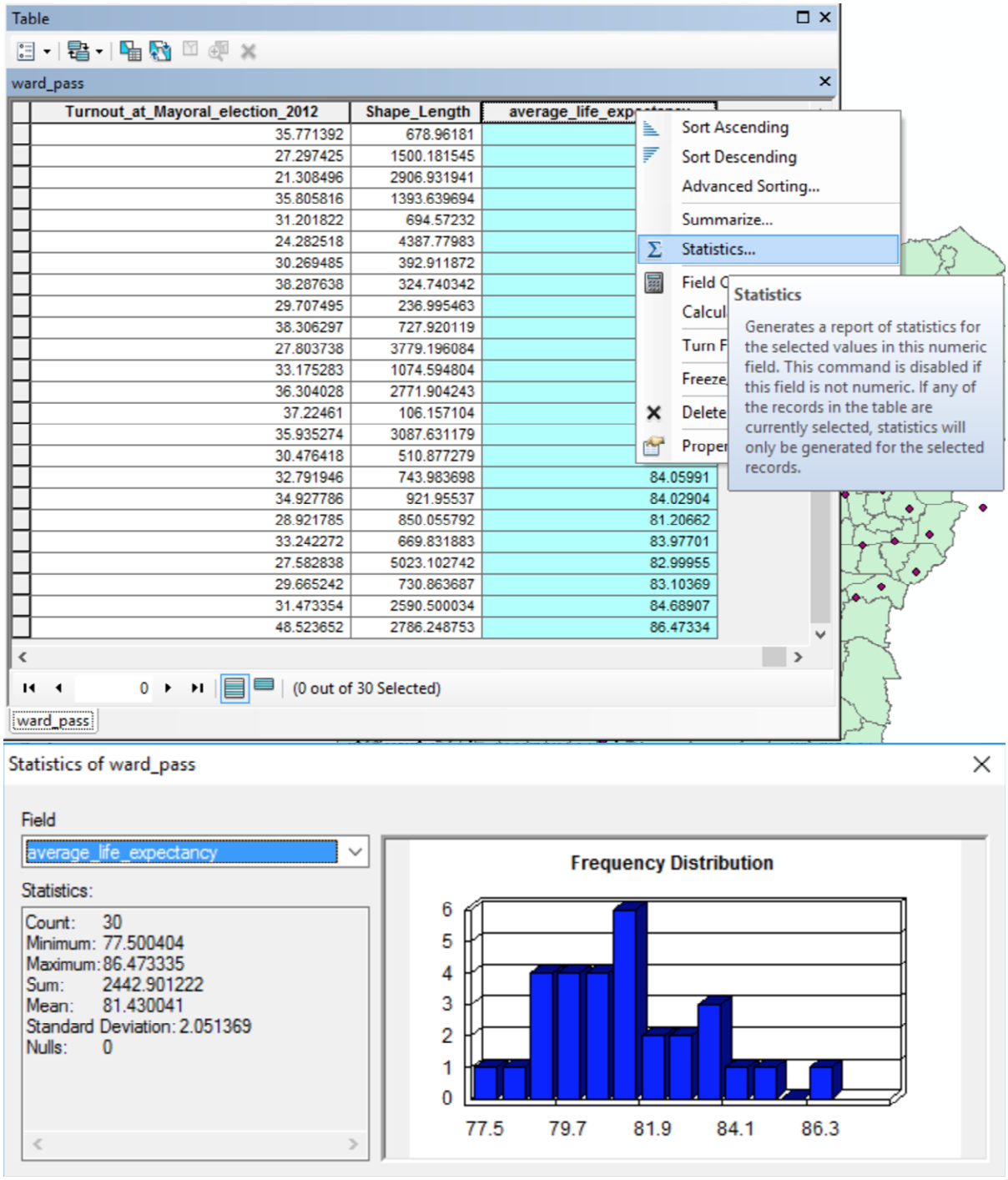
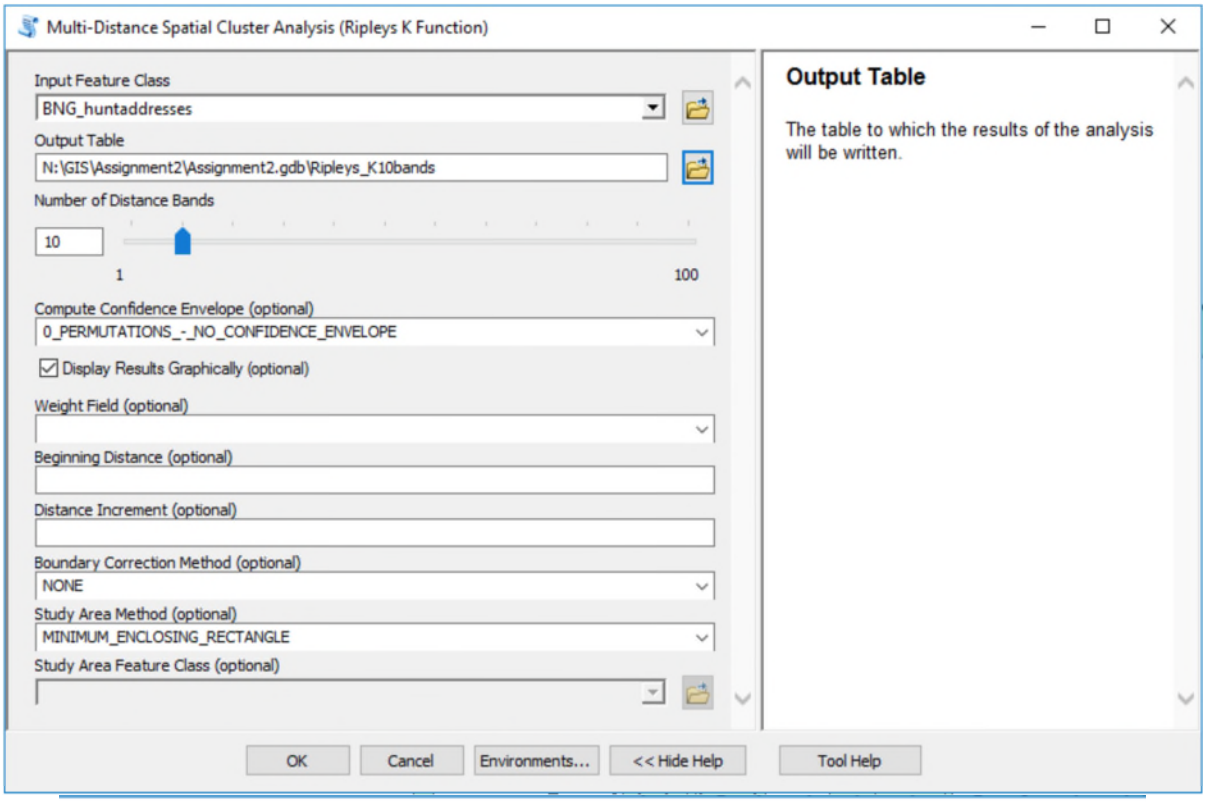


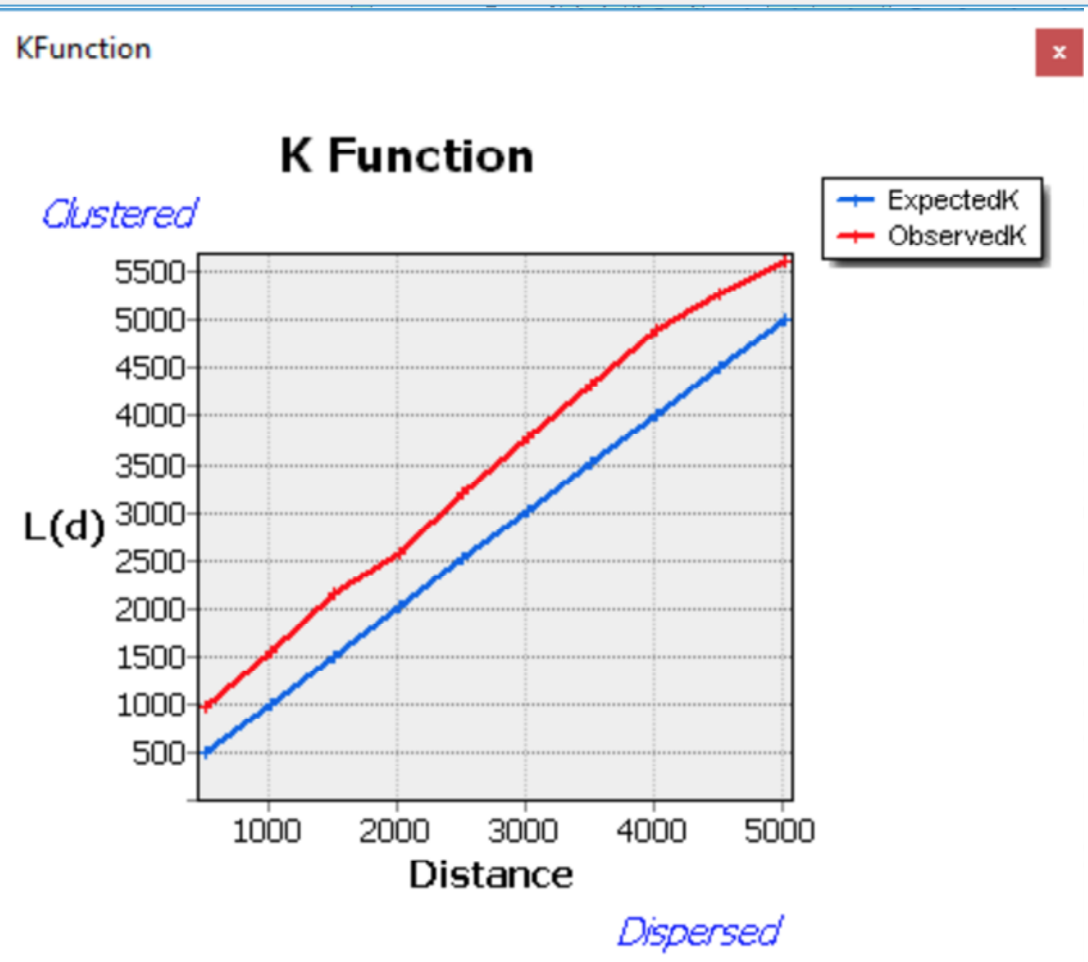
Table
ward_pass
Turnout_at_Mayoral_election_2012
Shape_Length
average
39.770311 431.797264 <Null>
31.528864 967.682663 <Null>
34.616661 587.078089 <Null>
29.999031 1463.094579 <Null>
35.626654 1008.238381 <Null>
23.281671 3232.830075 <Null>
35.771392 678.96181 <Null>
27.297425 1500.181545 <Null>
21.308496 2906.931941 <Null>
35.805816 1393.639694 <Null>
31.201822 694.57232 <Null>
24.282518 4387.77983 <Null>
30.269485 392.911872 <Null>
38.287638 324.740342 <Null>
29.707495 236.995463 <Null>
38.306297 727.920119 <Null>
27.803738 3779.196084 <Null>
33.175283 1074.594804 <Null>
36.304028 2771.904243 <Null>
37.22461 106.157104 <Null>
35.935274 3087.631179 <Null>
30.476418 510.877279 <Null>
32.791946 743.983698 <Null>
34.927786 921.95537 <Null>
Sort Ascending
Sort Descending
Advanced Sorting...
Summarize...
Statistics...
Field Calculator...
Calculate Geometry...
Turn Fi
Freeze/
Delete
Propert
Field Calculator
Populate or update the values of this field by specifying a calculation expression. If any of the records in the table are currently selected, only the values of the selected records will be calculated.
0
(1 out of 30 Selected)
ward_pass


Field Calculator
Parser
VB Script
Python
Fields:
OBJECTID_1
Shape
FID_BNG_Team7
OBJECTID
Name
descriptio
timestamp
begin
Type:
Number
String
Date
Functions:
Abs ( )
Atn ( )
Cos ( )
Exp ( )
Fix ( )
Int ( )
Log ( )
Sin ( )
Sqr ( )
Tan ( )
Show Codeblock
* / & + - =
average_life_expectancy =
[Male_life_expectancy_2009_13] /2 + [Female_life_expectancy_2009_13] /2
About calculating fields
Clear
Load...
Save...
OK
Cancel

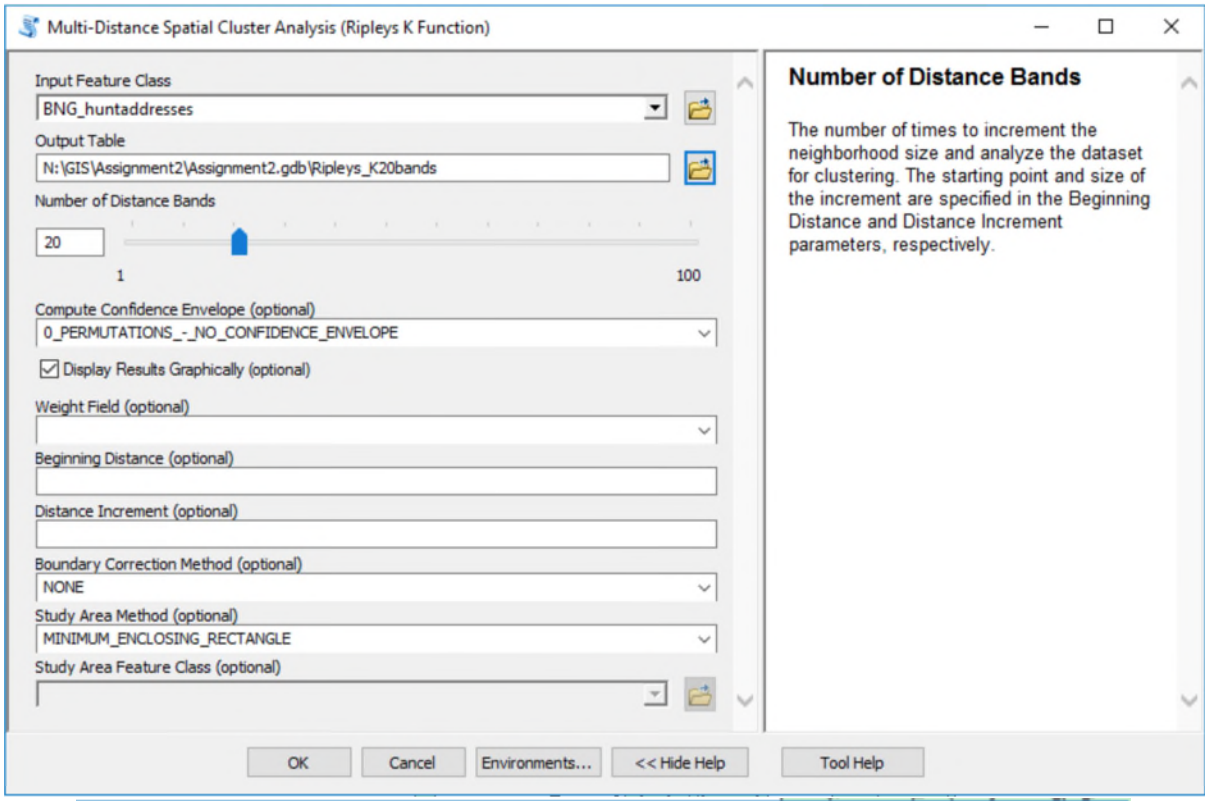


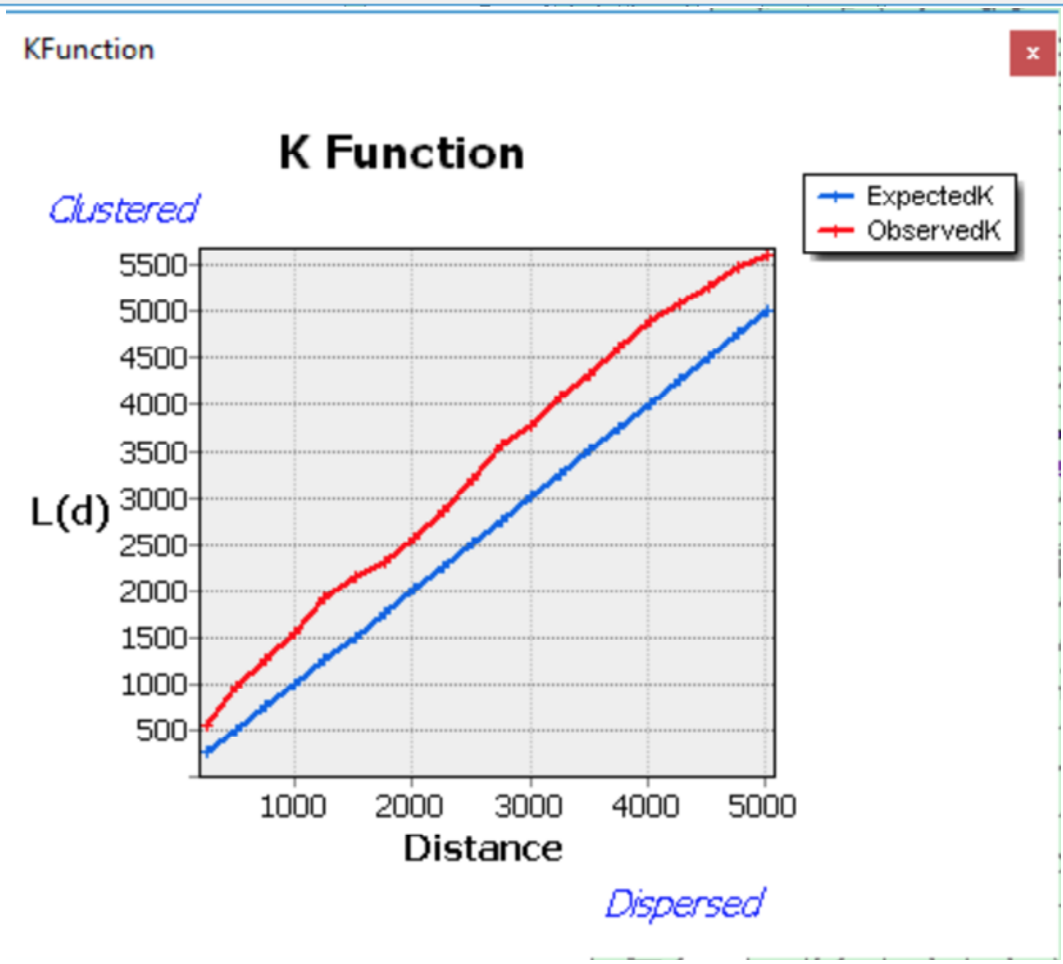
Question 6: use “Multi-Distance Spatial Cluster Analysis (Ripley's K Function)” in ArcToolbox -> Spatial Statistic Tools -> analyzing patterns. The observed K is always above the Expected K means the distribution of hunt addresses is clustering.





Ripley's K Function in 10 Distance Bands





Ripley's K Function in 20 Distance Bands

Uncertainty: for raw data: there may be some mistakes in data and the survey data and the trace recorded by GPS may also have deviation; for representation: the transformation in different coordinate systems will result in inaccuracy; for analysis: for Ripley's K, setting different number of distance bands, compute confidence envelope, study area method and so on will all influence the result. With changing the number of distance bands from 10 to 20, the clustering shown on the table becomes weaker in close distance.

Part 3 - Mini Project. 1800 word report

When riding a bike in London, we will find there are so many roads without the cycleway. You have to ride your bicycle very carefully because cars are probably just following you or passing you in very close distance. Whether the length of the cycleway has the relationship with the cycle accidents happened in every ward in London? It's the research question of this mini project. We set H_0 = The density of the length of the cycleway in every ward in London has no significant influence to the density of cycle accidents in wards, while H_1= The density of the length of the cycleway in every ward in London has significant influence to the density of cycle accidents in wards.

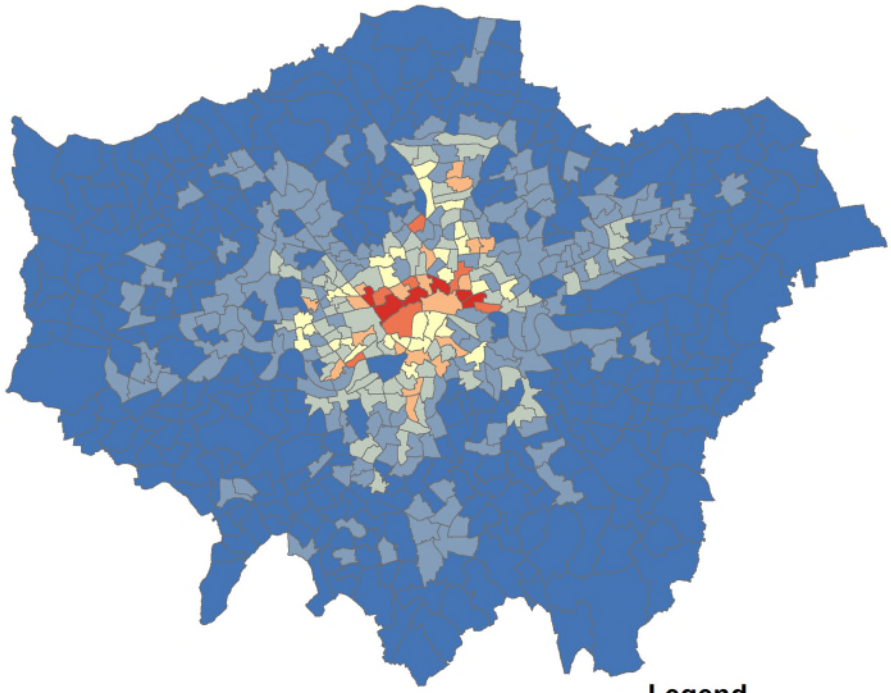
Data Source: The data for the coordination of every traffic accident happened in 2017 is obtained from http://content.tfl.gov.uk/2017-data-casualty.csv. The shapefile of all kinds of roads in London is downloaded from http://download.geofabrik.de/europe/great-britain/england/greater-london.html, and the original source is from OpenStreetMap. Ordnace Survey also provides open road map, but unfortunately, it doesn't contain cycleway in its classification system. Though the data from OpenStreetMap are donated by personal users which may have bigger deviation than the data from the government or consortium, it may be the only shapefile containing the cycleway which is open for the public. The London ward boundaries are downloaded from moodle but can also be acquired from London Datastore and many other websites easily.

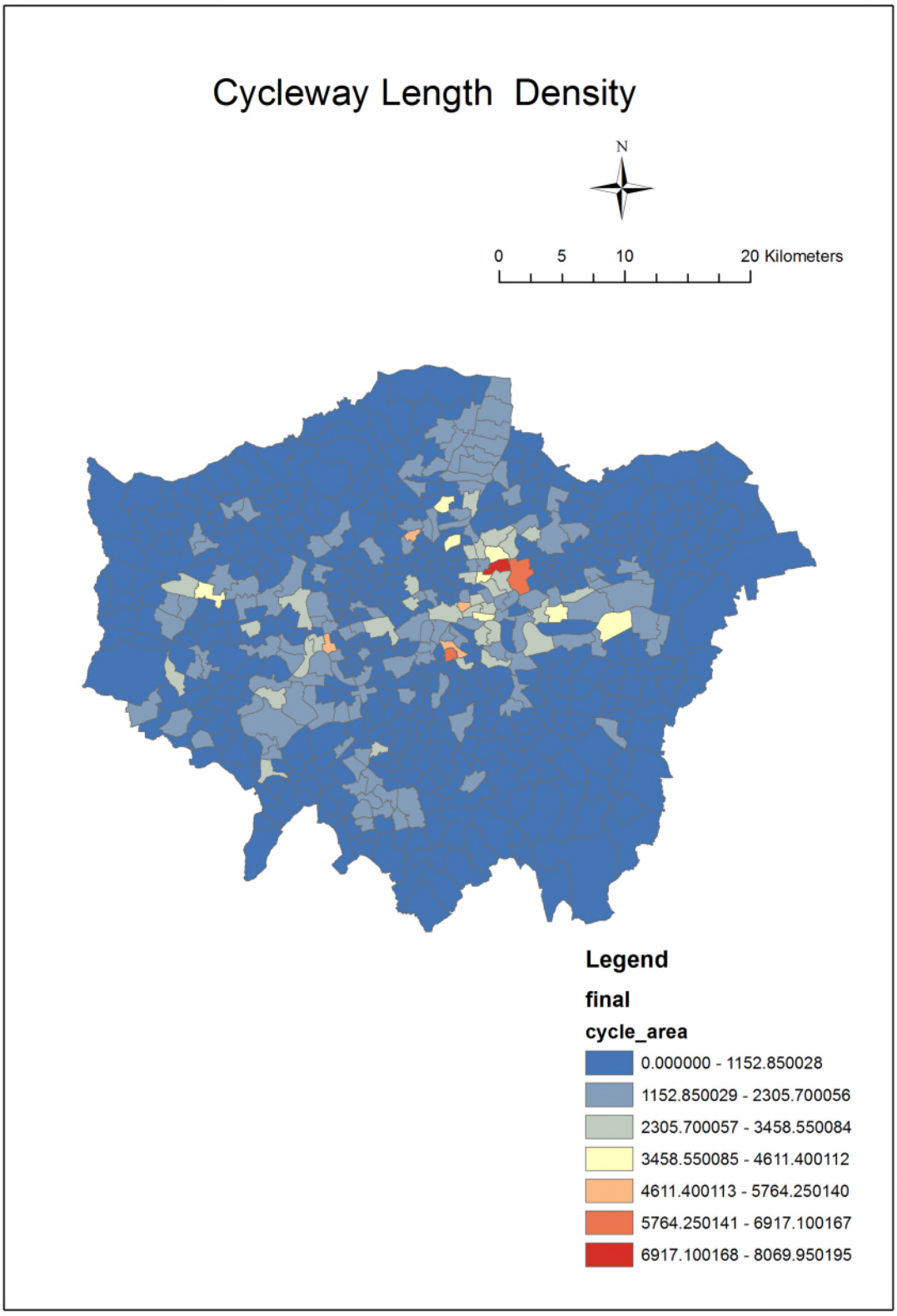
Software Choosing: It's a simple linear regression question and ArcGIS already has the "Ordinary Least Squares" tool in ArcToolbox -> Spatial Statistics Tools -> Modeling Spatial Relationships which can solve this question well and produce statistic file automatically. Therefore, I choose to use ArcGIS to do this research.

Raw Data Transformation: To reduce the burden of ArcMap, select only the cycle accidents which caused fatal or serious casualties by using filter in Excel, and then drag the new csv into ArcMap to display the XY events, and export as a shapefile of points. The coordination system of this CSV is British National Grid which is the same as the ward boundaries, so we needn't to transfer the projections of the cycle accident points. While the coordination system for OpenStreetMap is World Geodetic System, we select the cycleway by attribute first and export it as a new cycleway shapefile, then transfer it into BNG to reduce operation time.

To count the number of cycle accidents occurred in every wards by using "spatial join", choose "JOIN_ONE_TO_ONE", and a new shapefile "join_accident" will be output which contain the joint_count column. To calculate the total length of cycleway in every ward, we use "intersect" tool first to cut the cycleways which across the ward boundary. Later, join the intersected shapefile with "join_accident" by using "spatial join" again to gain the "final" shapefile which has the sum of the cycleway length and the count of the cycle accidents in every wards. Add new fields in "final" attribute table to calculate the cycleway length density and cycle accident density by "cycleway length/shape area*1000000" and "joint count/shape area*1000000" separately. Right click "final" to find Layer Properties -> Symbology -> Quantities to make the density map as below shows.

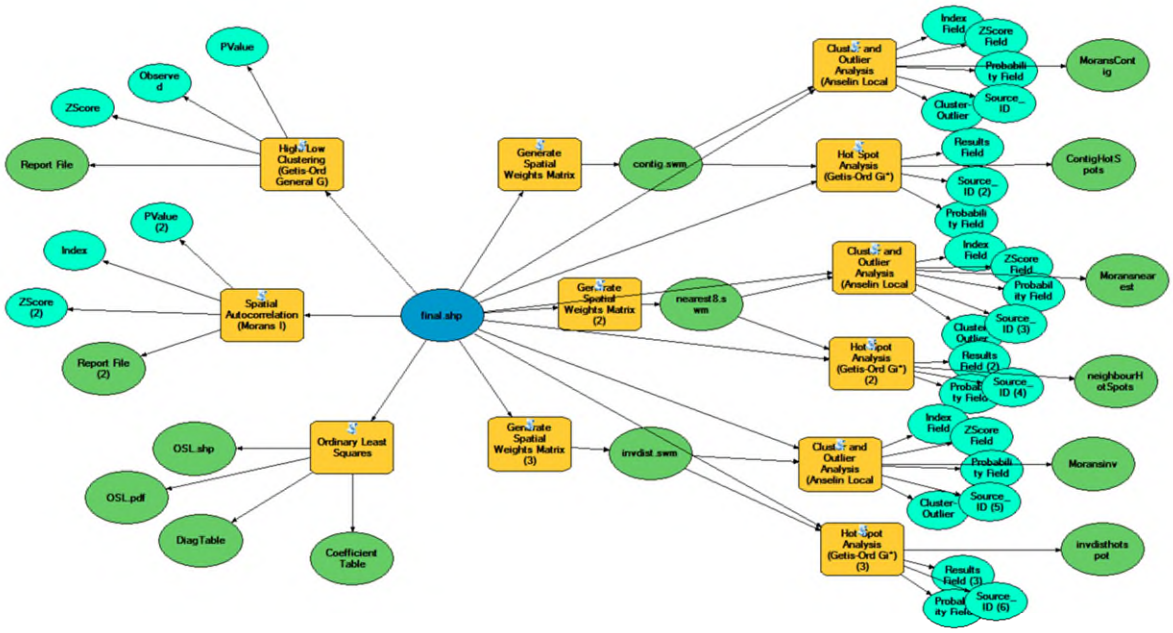
Cycle Accident Density

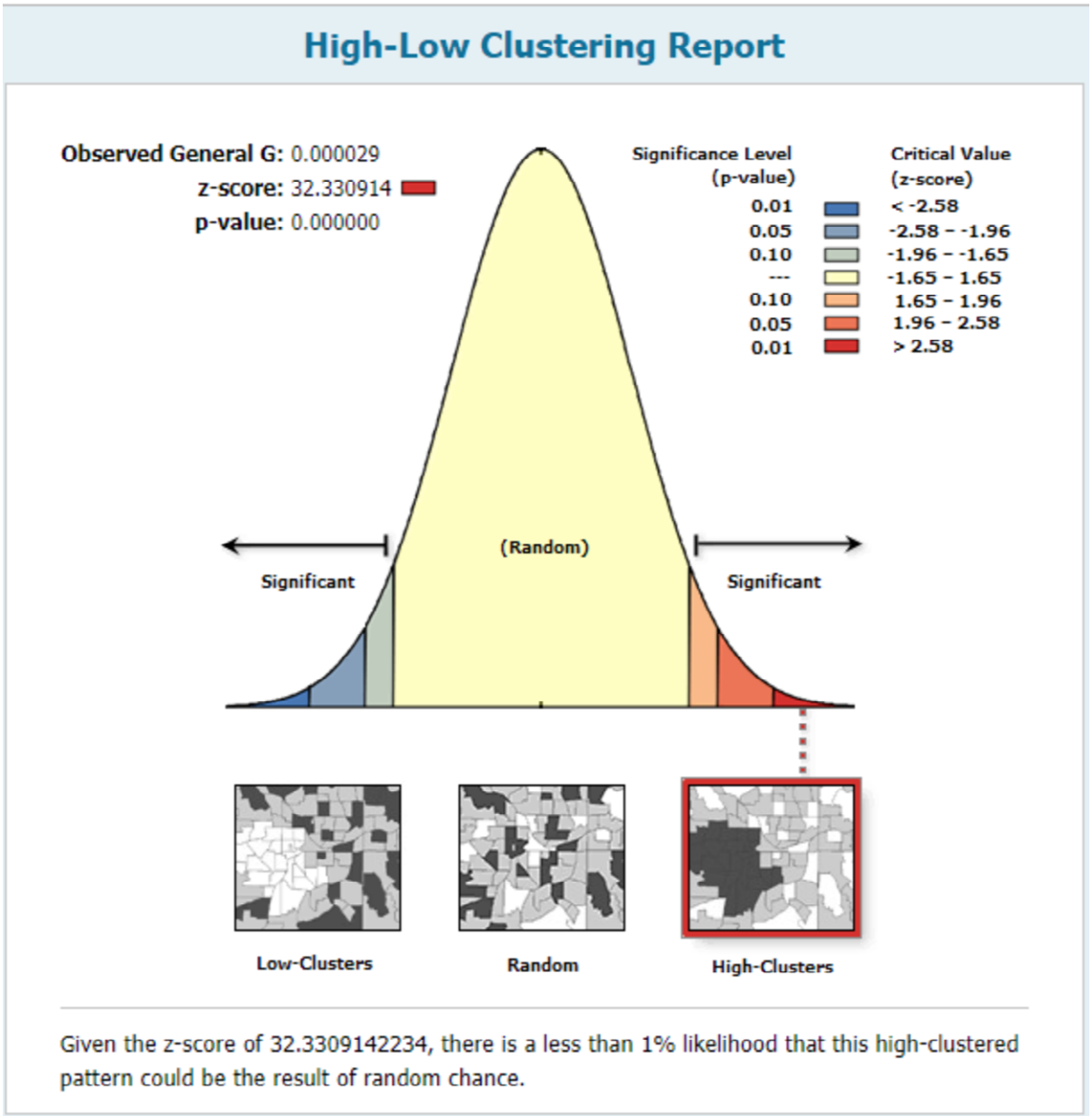




We can see both the densities of the cycleway length and cycle accidents have clustering and there are outliners in the center of London from these two maps intuitively. To prove our suppose, we can use the tool model build in practical 6 to analyze the Local Moran's I and Getis-Ord Gi* which can reflect the clustering in local area, also add General G and Moran's

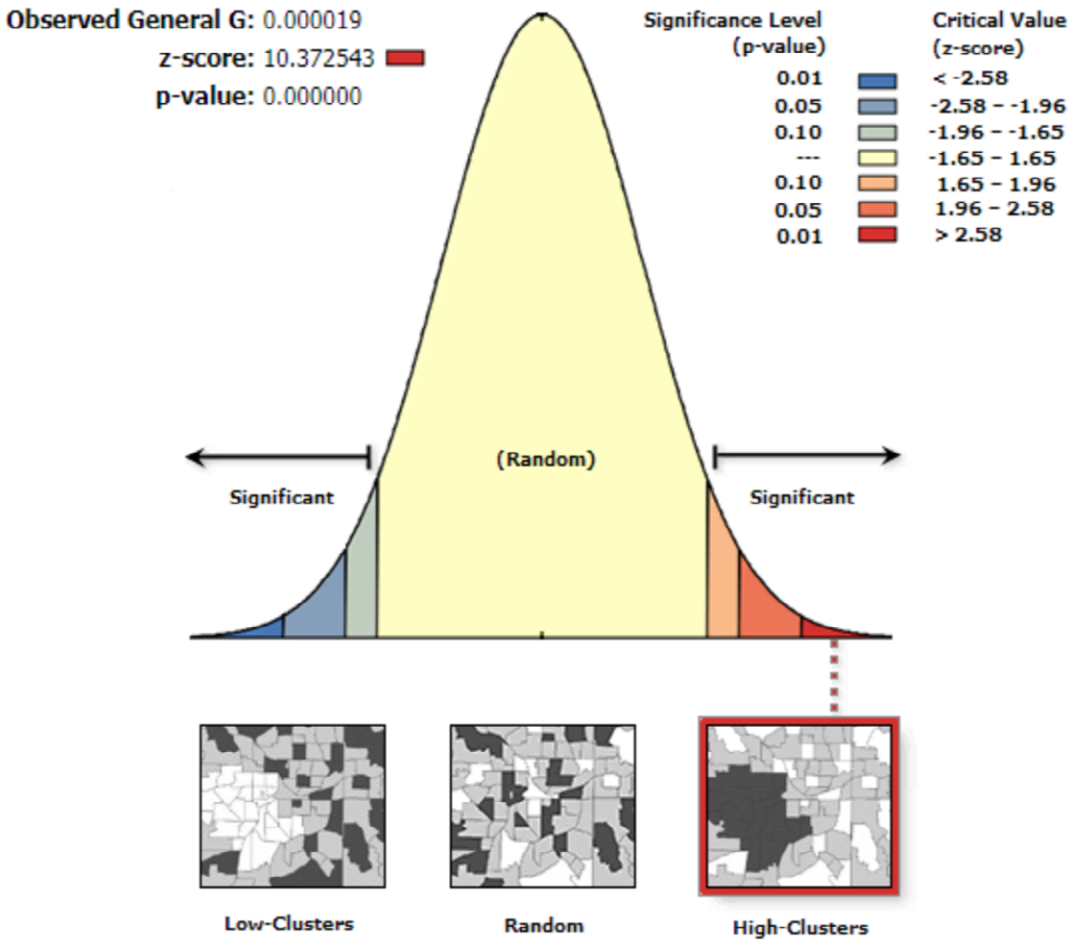
I to calculate the clustering in global area, and Ordinary Least Squares to analyze the linear relationship between the density of cycle accidents and cycleway length. After click “Run Entire Model”, all analysis will run automatically. Moreover, the model is reusable, and we can copy and change it easily. For cycle accident density clustering analysis, the Generate Spatial Weights Matrix is more suitable to be inverse distance. The influence of distance may be bigger than contiguity or neighbours, as the map above shows the further from the center, the lower the value is. But for cycleway length, the matric may be more related to neighbours. For convenience, we only compare these two densities by using inverse distance matrix. Both the p-value for high low clustering for these two densities are less than 0.01, which means the high value of them are clustering at a significant level of 0.01. As the z-score for cycle accident density is higher than the cycleway length, it’s more possible for cycle accident density has high-clustered pattern. Both the p-value for spatial autocorrelation for these two densities are also less than 0.01, which means there are clustering at a significant level of 0.01. As the z-score for cycle accident density is higher than the cycleway length, it’s more possible for cycle accident density has clustering pattern.





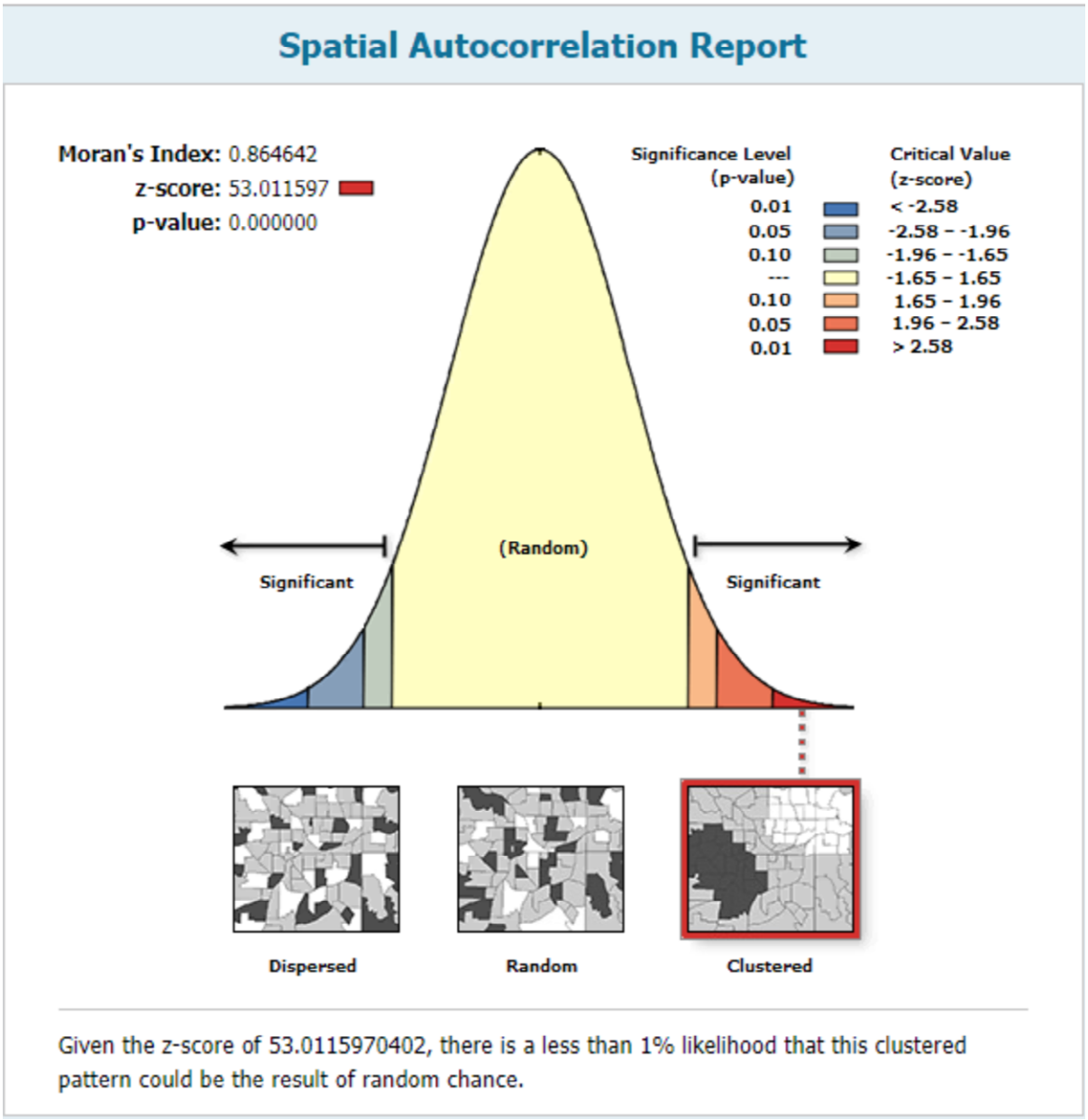
High-Low Clustering Report for Cycle Accident Density

High-Low Clustering Report

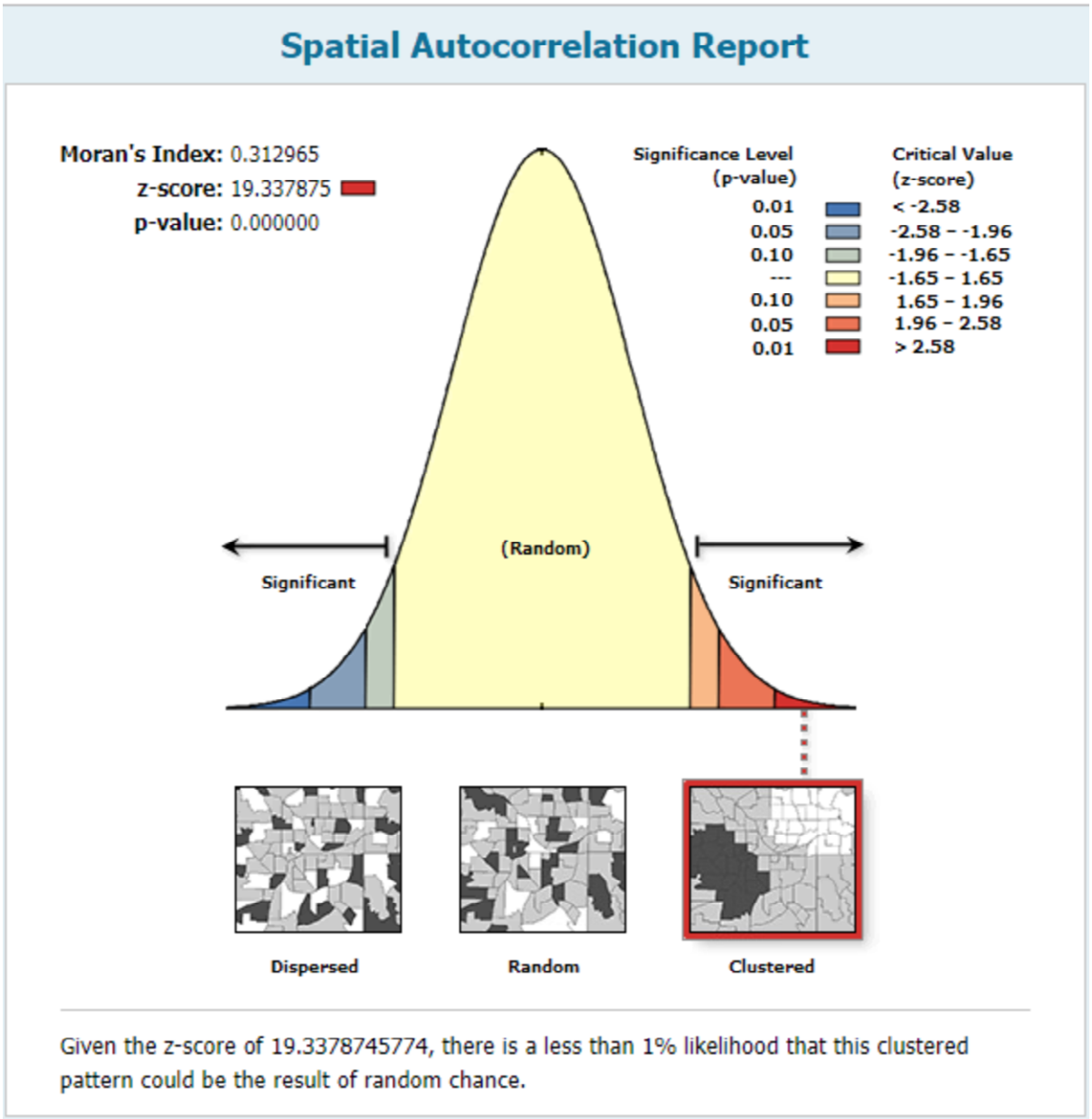


Given the z-score of 10.3725431283, there is a less than 1% likelihood that this high-clustered pattern could be the result of random chance.

High-Low Clustering Report for Cycleway Length Density



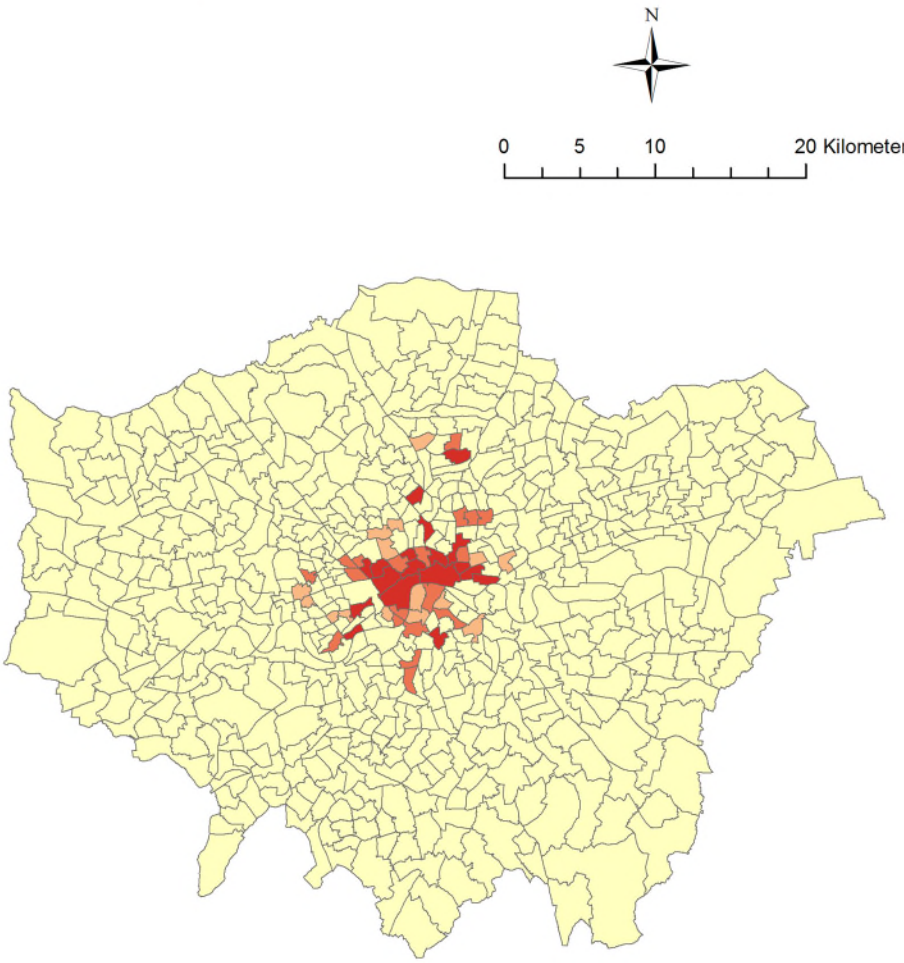
Spatial Autocorrelation Report for Cycle Accident Density



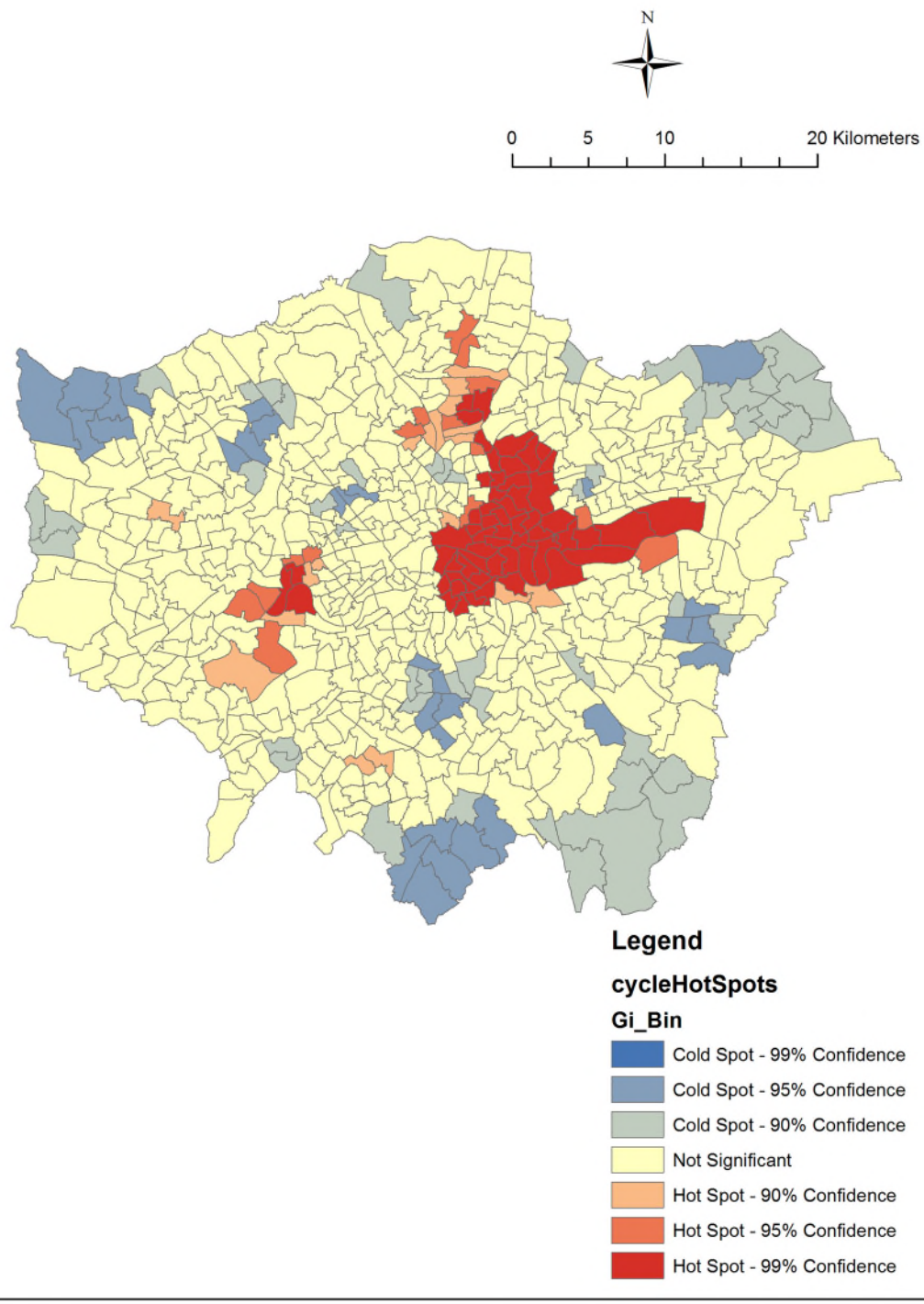
Spatial Autocorrelation Report for Cycleway Length Density

The Getis-Ord Gi* map for the cycle accident density below shows there have hot spots in the center of London in high confidence, while the map for the cycleway length density shows the hot spots may exist in the center bias to eastern of Local in high confidence, and the cold spots on the corner of London in high confidence. The Moran's I map shows both the cycle accident density and the cycleway length density has local high-high clustering in the center of London in high confidence, whilst the cycleway length density also has low-low clustering on the corner of London. Through comparing clustering pattern of these two density, we find there has many similarities but also many differences.

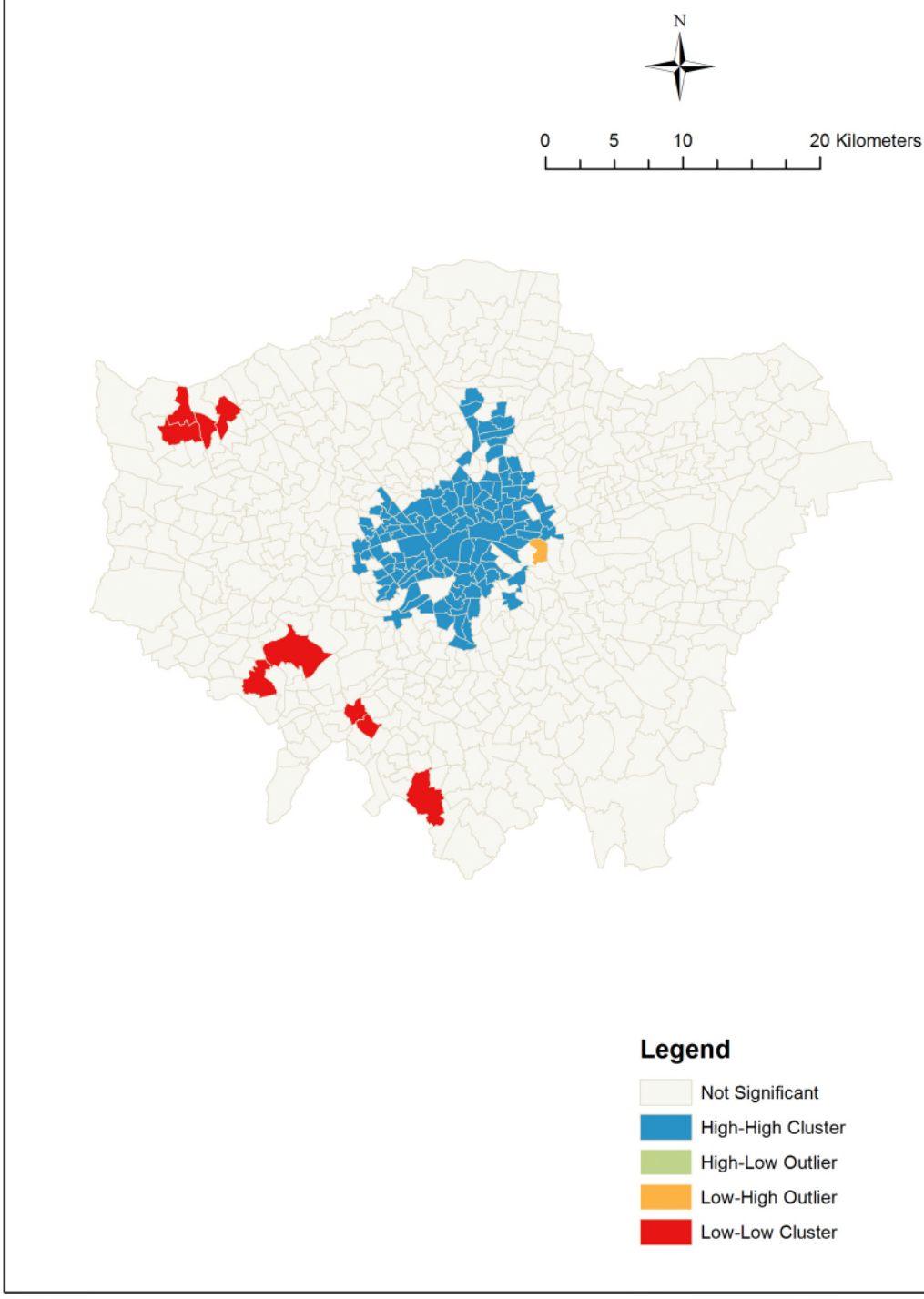
Getis-Ord Gi*, Cycle Accident in London

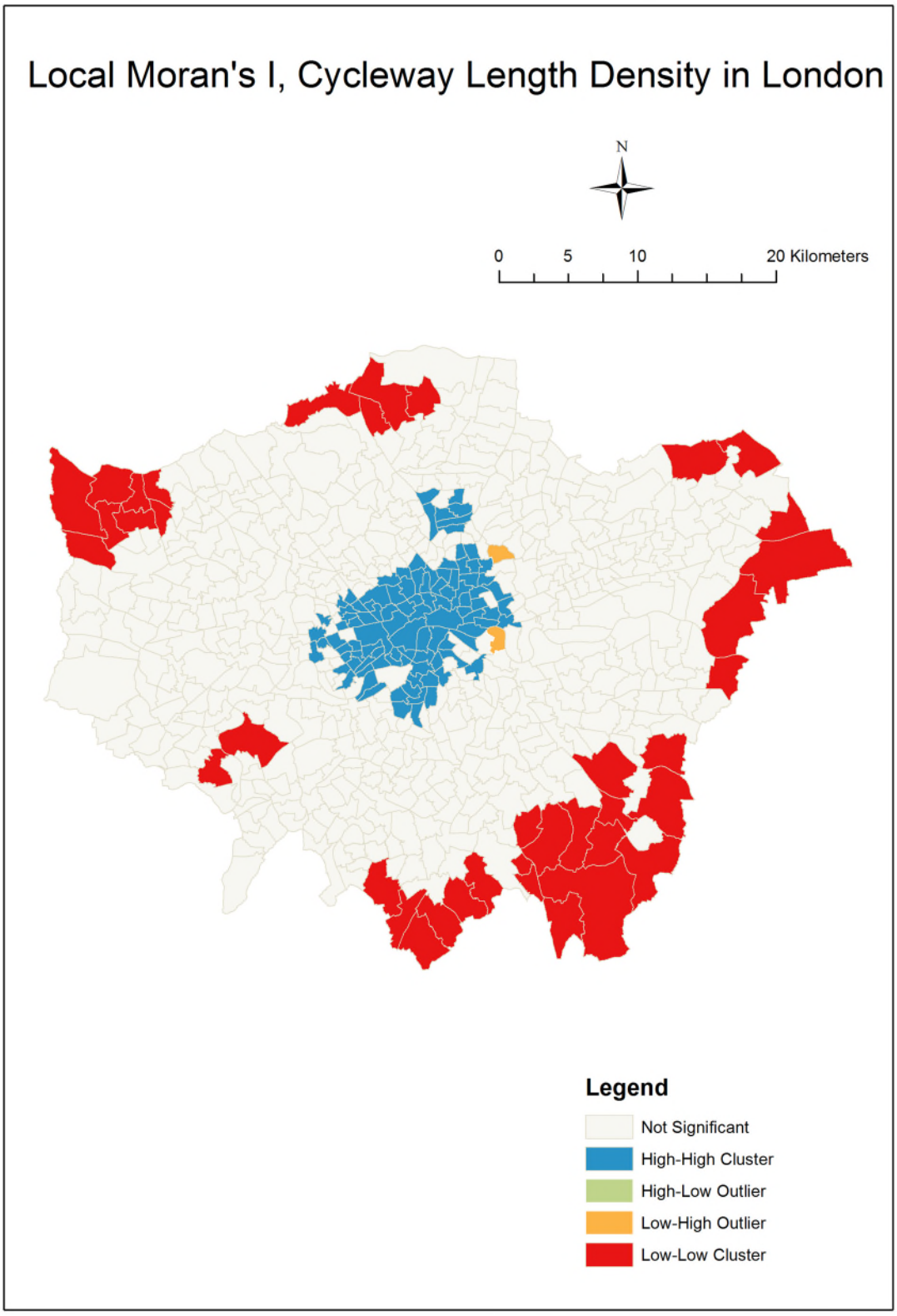


Getis-Ord Gi*, Cycleway Length Density in London



Local Moran's I, Cycle Accident in London





To analyze the linear relationship between the cycle accident density and the cycleway length density in every ward of London, the result of Ordinary Least Squares Analysis shows that the p-value is 0.000002 < 0.01, which means the liner relationship exists in 0.01 significant level.

However, adjusted R-Squared is only 0.035, which means the cycleway length density can explain about 3.5% of the density of cycle accident. The coefficient of the equation is 0.00551, which is different from our mind, as we usually think the increase of the length of the cycleway will reduce the bicycle accidents.

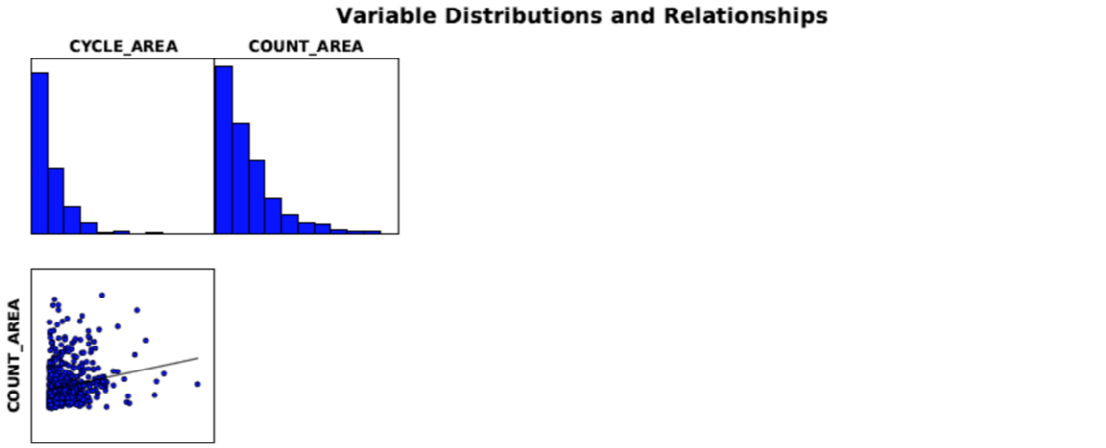
Summary of OLS Results - Model Variables

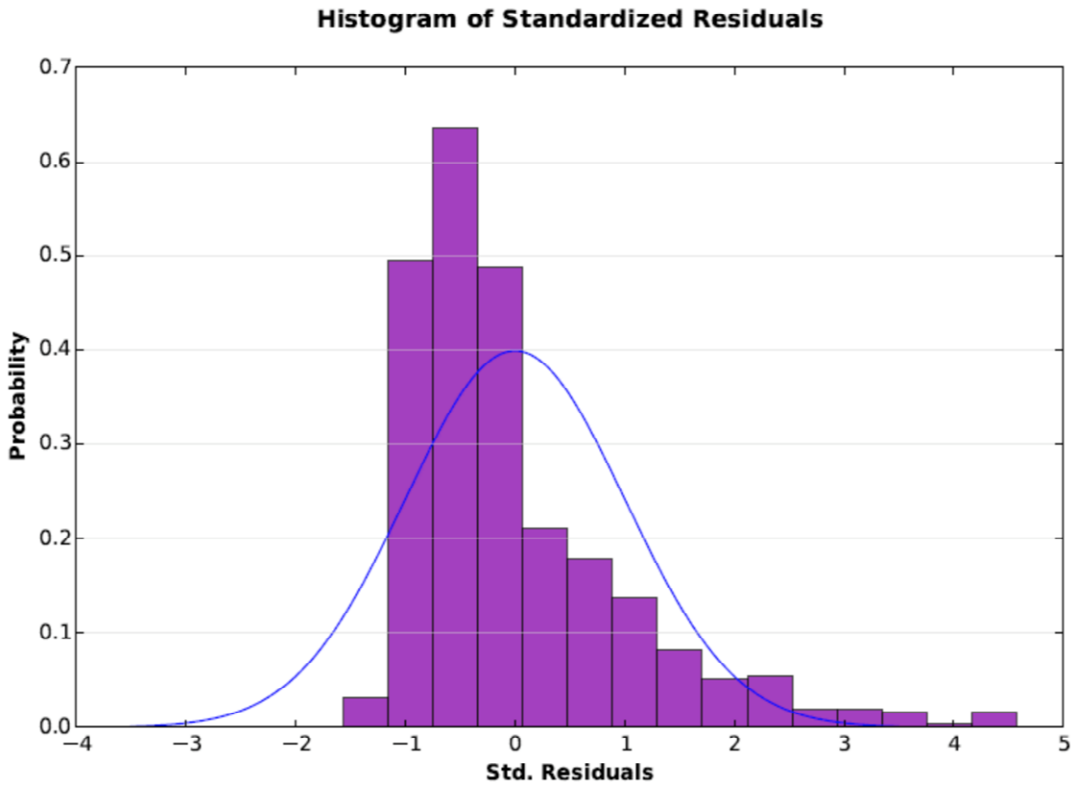
Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]
Intercept	27.472072	1.480488	18.556088	0.000000*	1.497003	18.351378	0.000000*
CYCLE_AREA	0.005510	0.001133	4.863954	0.000002*	0.001353	4.072077	0.000059*

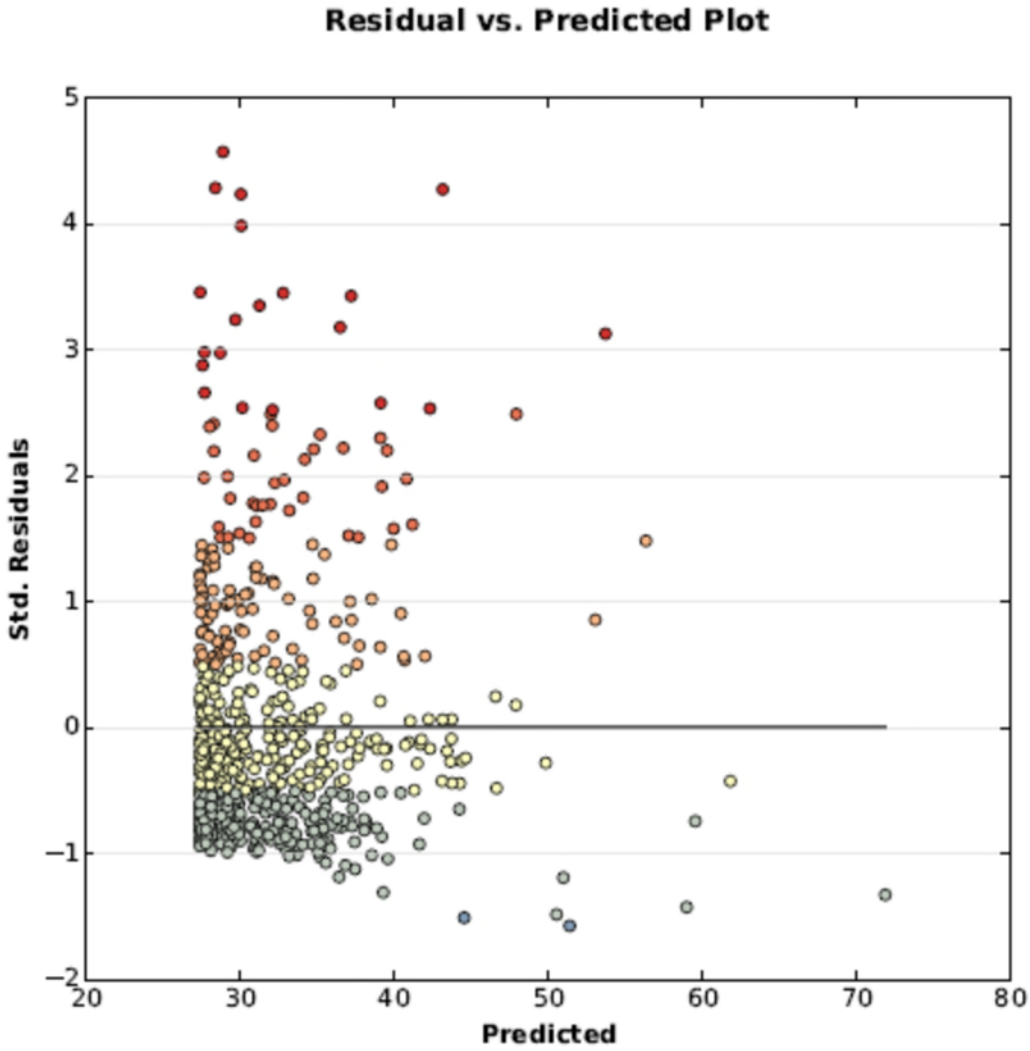
OLS Diagnostics

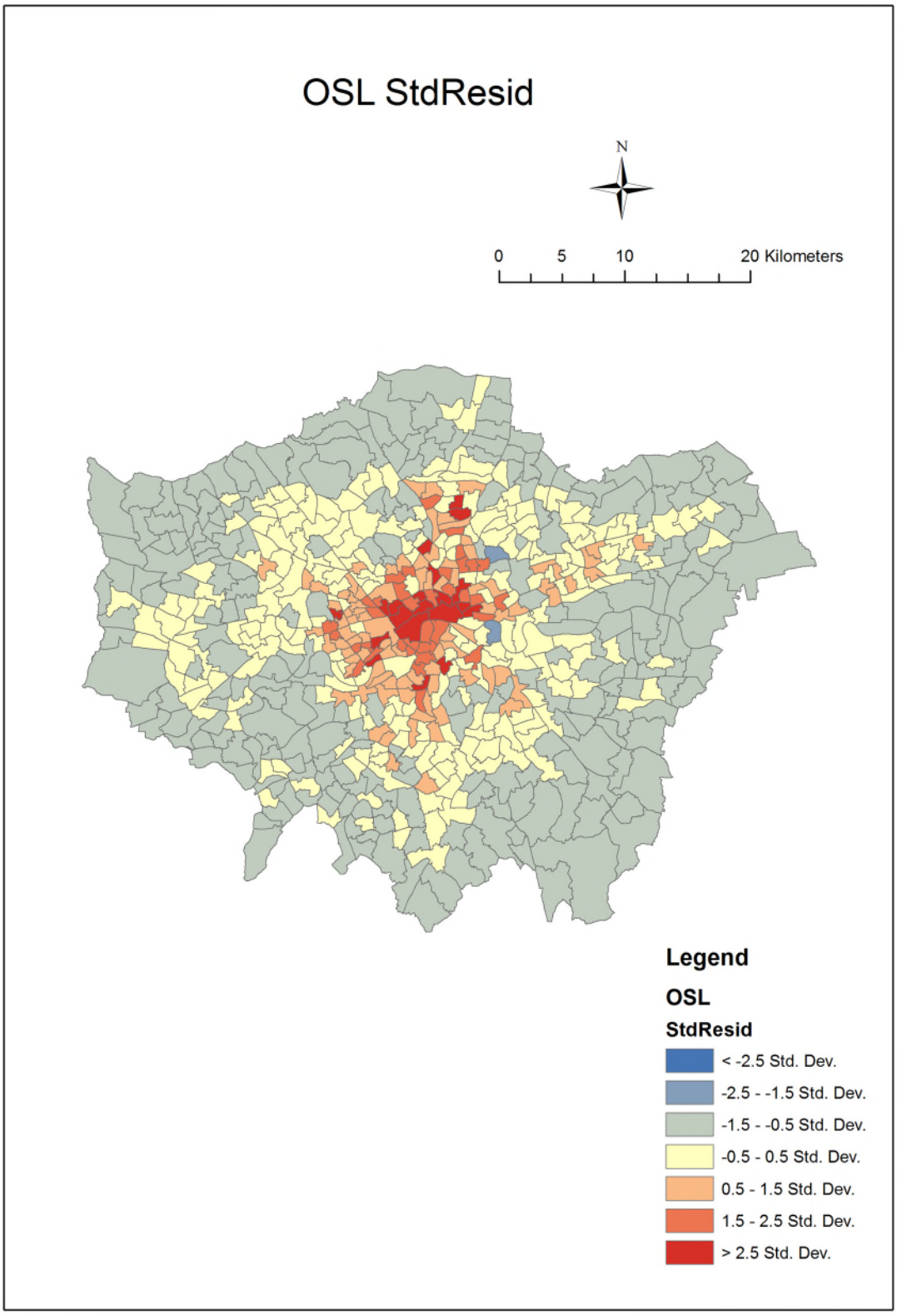
Input Features:	final	Dependent Variable:	COUNT_AREA
Number of Observations:	625	Akaike's Information Criterion (AICc) [d]:	5935.776293
Multiple R-Squared [d]:	0.036585	Adjusted R-Squared [d]:	0.035039
Joint F-Statistic [e]:	23.658050	Prob(>F), (1,623) degrees of freedom:	0.000001*
Joint Wald Statistic [e]:	16.581811	Prob(>chi-squared), (1) degrees of freedom:	0.000047*
Koenker (BP) Statistic [f]:	3.224451	Prob(>chi-squared), (1) degrees of freedom:	0.072546
Jarque-Bera Statistic [g]:	553.823020	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

All of Variable Distributions and Relationships Plot, Histogram of Standardized Residuals and Residual vs Plot show that the relationship between them may not be linear, but more like logarithmic or exponential.


Variable Distributions and Relationships
CYCLE_AREA
COUNT_AREA
COUNT_AREA


Histogram of Standardized Residuals
Probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
−4
−3
−2
−1
0
1
2
3
4
5
Std. Residuals


Residual vs. Predicted Plot
Std. Residuals
Predicted
5
4
3
2
1
0
−1
−2
20
30
40
50
60
70
80



The Std. Residual map also reflect that there are many residuals cannot be explained by cycleway length density, especially in the center of London.

In further research, we can try to use nonlinear regression to analyze the relationship between the cycleway length density and the cycle accident density. We can attempt to build multiple regression model as well, as the car flowrate and bicycle flowrate, and many other variables may contribute to the cycle accident density. Spatial difference may also exist between inner London and outer London, as the residual map shows the Std. Dev of residuals in inner London are much larger than outer London. The Getis-Ord Gi* map for cycleway length density in every ward of London also seems very interesting, as the hot spots are not in the center, but inclined to the east. The reasons behind the map can be analyzed in the future.

Also, there exist many uncertainty in this research. For data, the cycleway data from OSM probably has bias and pretty different from the real world as no one will draw the cycleway systematically for no money. The people decisions in real world are also much more complex than data model, if this road has no cycleway, the cyclist may choose another way which has cycleway though the trace will be longer, as a result, the influence of whether a road has a cycleway is not so important as our imagination. What's more, when you ride in a road without cycleway, you will really feel dangerous and may because you feel more nervous and became more careful, the probability of accident decrease. We also omitted the slight casualties when analyze the data. If we consider them together with fatal and serious casualties, they result may change a lot. The occurring of cycle accident may also been casual, cannot be explained by other variables well to some extent actually. Also, for MAUP, the choice of using wards as boundaries may not be the best decision either.

To conclude, it's still important to improve the cycleway for cyclists though it has low relationship with road safety, as it's a very natural-friendly and healthy ways for transportation. Wishing in the future, more data can be open to public so that the analysis for the real world can be more easily.