Your Name = [Ankila KUmari   *GIS 5555 Basic Spatial Analysis*

] internet_id = [kuma0389]

Time_Spent = [70 mins] (after-class)

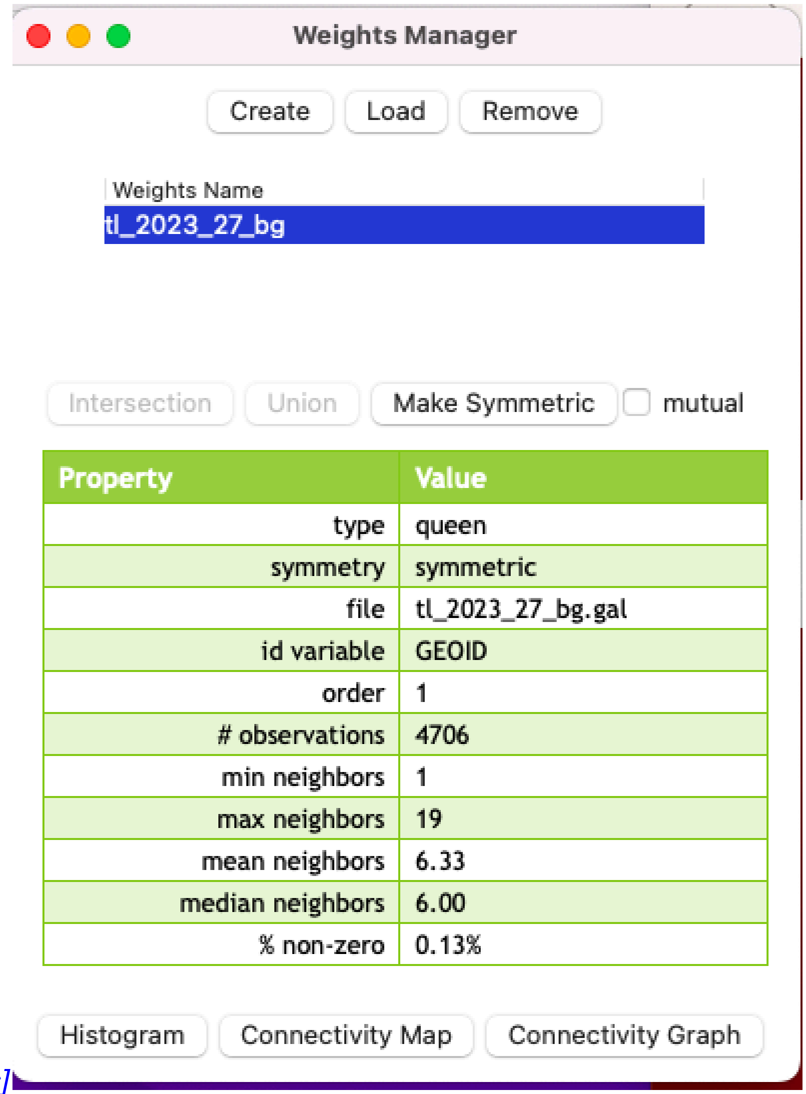
[Windows+Shift+S for screenshot of your analysis]

[Fill the above-listed info and then submit the completed document in Canvas (try to include all analysis results that can help reflect your workflow and thoughts, i.e., images, information about data, your statements, etc.)]

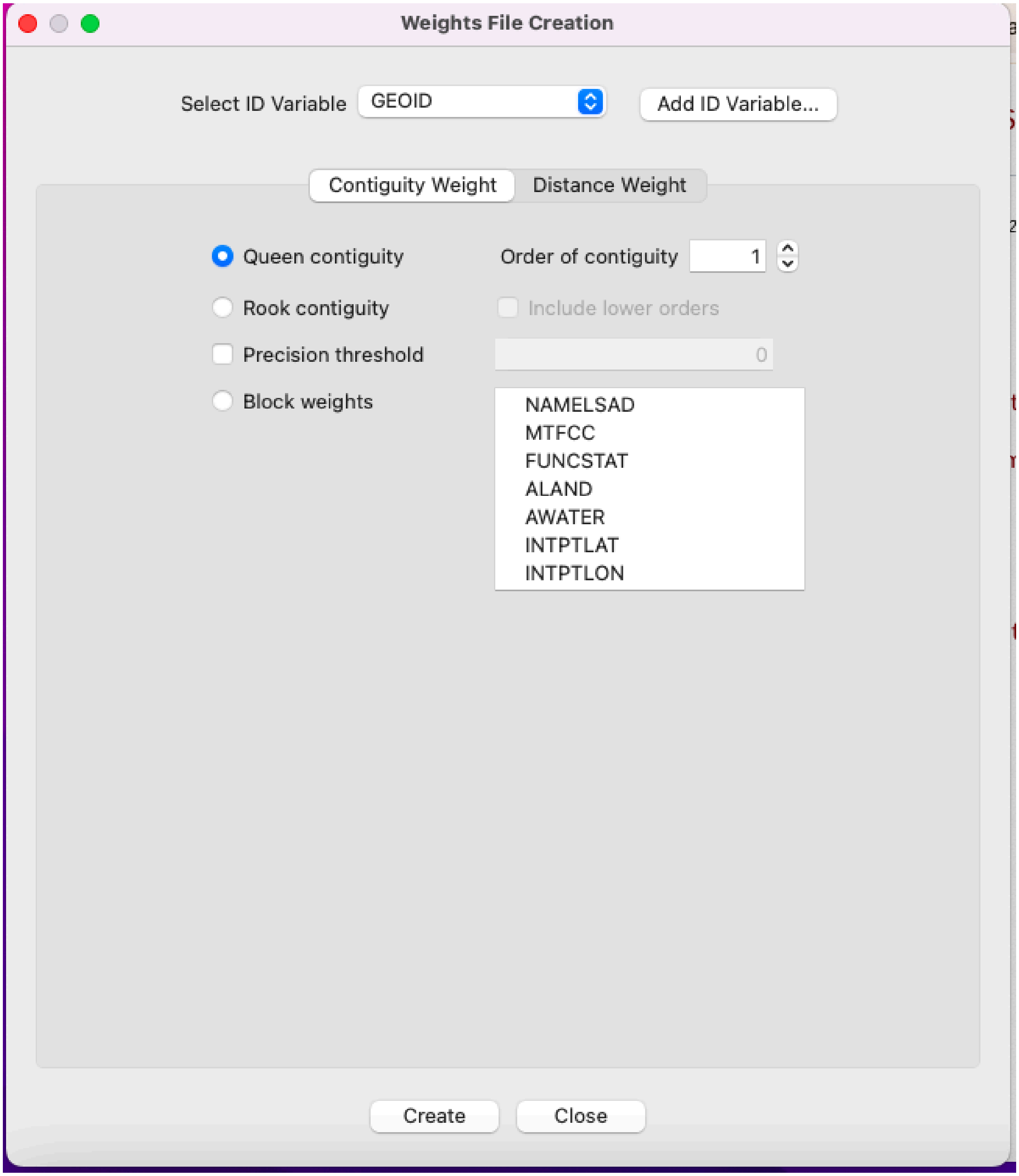
Assignment for Lab 3a

"Spatial Weights"

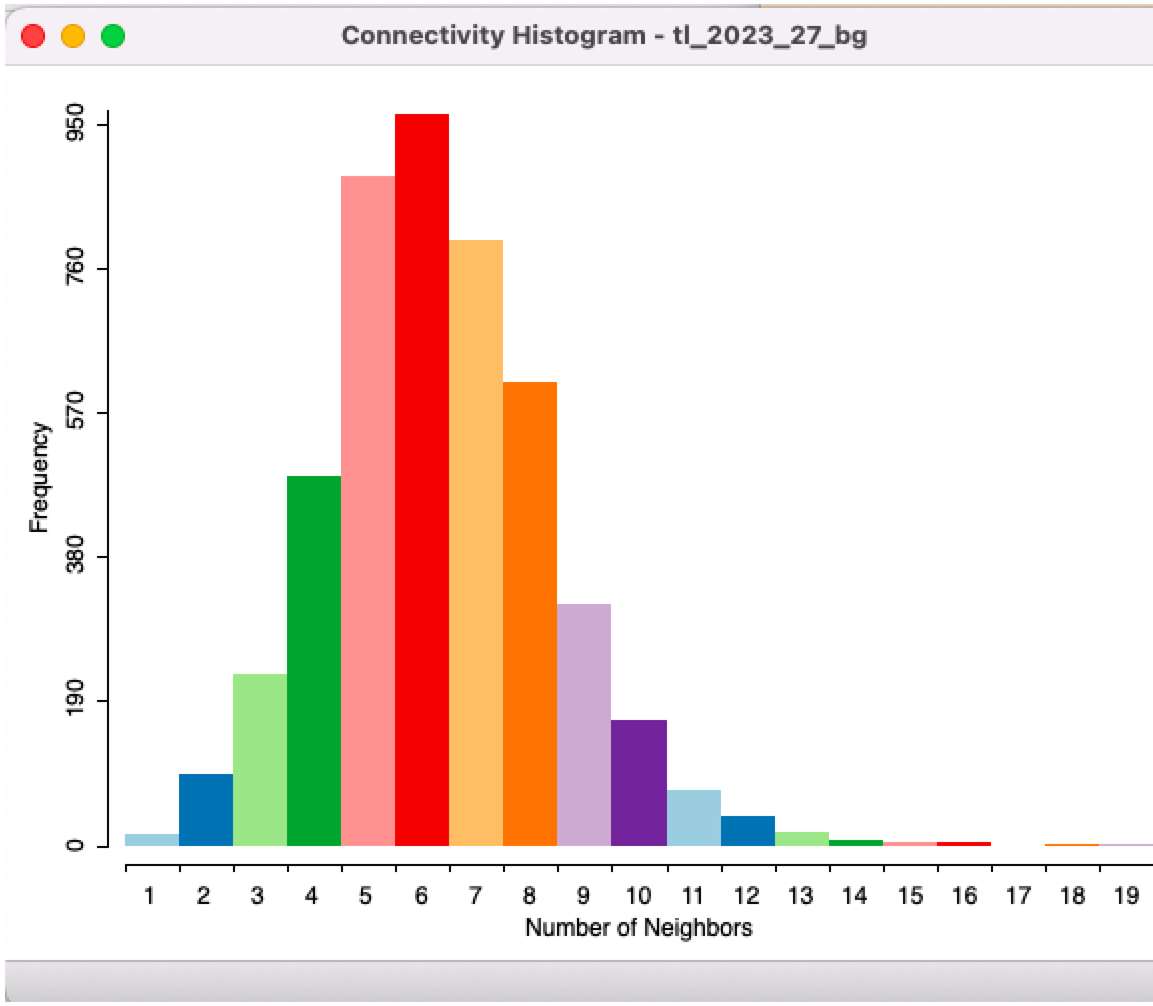
- ***Task 1 Contiguity weights and connectivity***
  - *Create the spatial weights for your data based on Queen contiguity, save it to the project file and show us the screenshot when you identify the stored weights information in the project file.*
  - *Select the category with the largest number of neighbors in your connectivity histogram and show the results in the linked map view.*
  - *Save the neighbor cardinality to your data table and create a chropleth map.*
  - *Choose one of the observation in the category you selected from connectivity histogram, checking its neighbors in the connectivity map or connectivity graph.*



[Your answers]


Weights File Creation
Select ID Variable
GEOID
Add ID Variable...
Contiguity Weight
Distance Weight
Queen contiguity
Order of contiguity
1
Rook contiguity
Include lower orders
Precision threshold
0
Block weights
NAMELSAD
MTFCC
FUNCSTAT
ALAND
AWATER
INTPTLAT
INTPTLON
Create
Close

1. *Xxx*



2.

[image screenshot & captions]

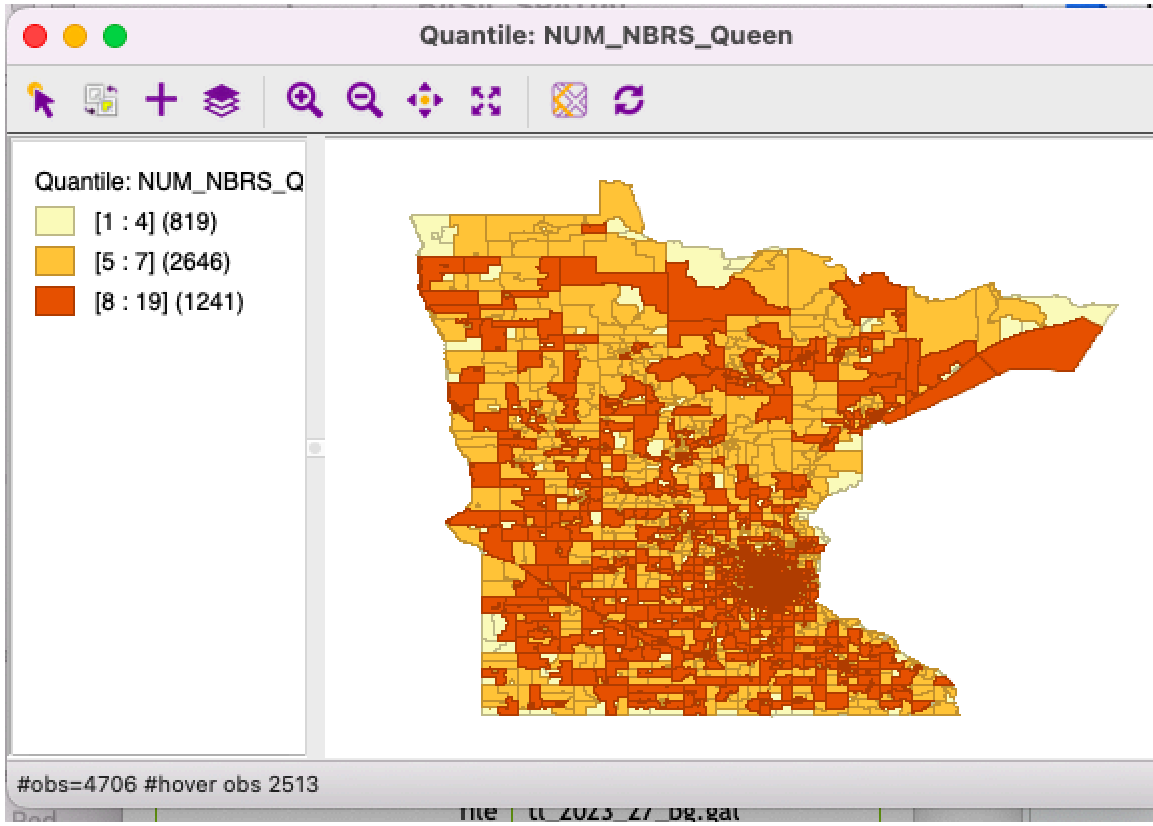
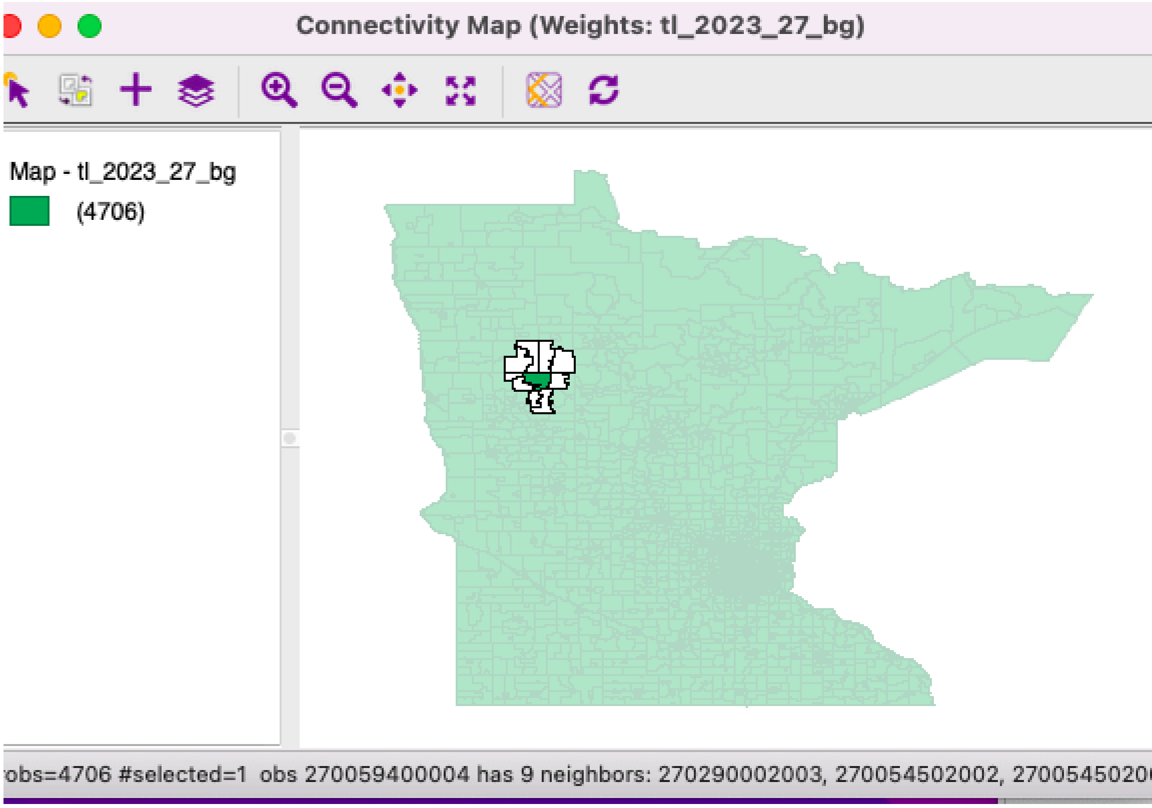


Table - tl_2023_27_bg

		NAMELSAD	MTFCC	FUNCSTAT	ALAND	AWATER	INTPTLAT	INTPTLON	NUM_NBRS_Q
1	1	Block Group 1	G503	S	1606303	0	+44.0521618	-092.4407842	
2	1	Block Group 1	G503	S	413740	0	+44.0470838	-092.4980347	
3	2	Block Group 2	G503	S	1263028	100320	+45.5537547	-094.1393030	
4	4	Block Group 4	G503	S	265750741	6924541	+46.0789140	-093.1547400	
5	2	Block Group 2	G503	S	76834729	1751919	+45.9390240	-093.1962301	
6	2	Block Group 2	G503	S	985937	118	+45.0164320	-093.1846115	
7	4	Block Group 4	G503	S	49567	0	+44.9468480	-093.0920156	
8	2	Block Group 2	G503	S	98152	35470	+44.9443374	-093.0904425	
9	7	Block Group 7	G503	S	8947231	26696	+45.1378675	-093.3745295	
10	5	Block Group 5	G503	S	1434608	6646	+45.1120213	-093.3894445	
11	3	Block Group 3	G503	S	1964816	6390	+44.8685741	-093.4424036	
12	4	Block Group 4	G503	S	866587	64279	+44.8888858	-093.4135066	
13	6	Block Group 6	G503	S	1261584	20796	+44.8817339	-093.4044825	
14	5	Block Group 5	G503	S	817192	6334	+44.8673421	-093.4022925	
15	6	Block Group 6	G503	S	1115752	0	+45.1263670	-093.3668735	
16	5	Block Group 5	G503	S	619286	0	+45.1454247	-093.2870982	
17	4	Block Group 4	G503	S	2864755	79459	+45.1327340	-093.2233885	
18	2	Block Group 2	G503	S	747652	58989	+45.1528015	-093.2589245	
19	4	Block Group 4	G503	S	1800029	1437411	+45.1466629	-093.0954501	
20	1	Block Group 1	G503	S	818130	1161	+45.1830634	-093.3602780	
21	2	Block Group 2	G503	S	1435787	0	+45.1798296	-093.3454438	

#row=4706

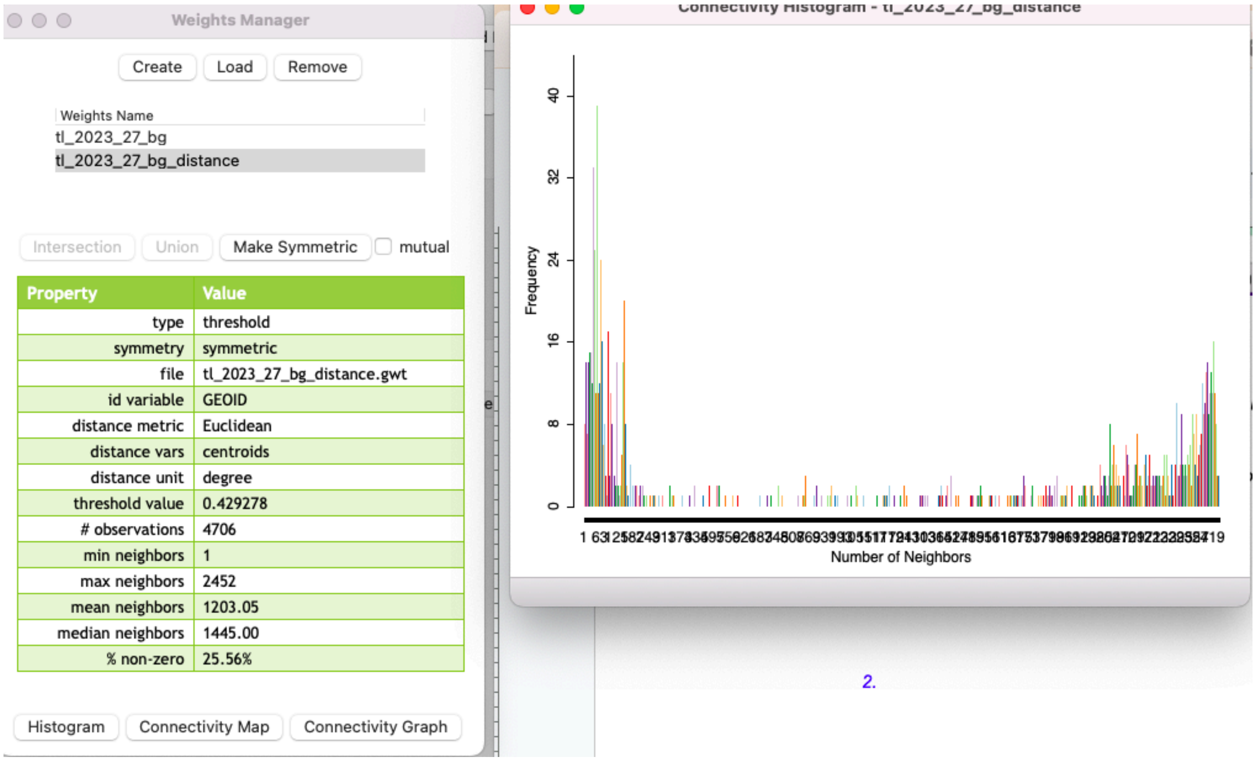
3. *(if necessary)*



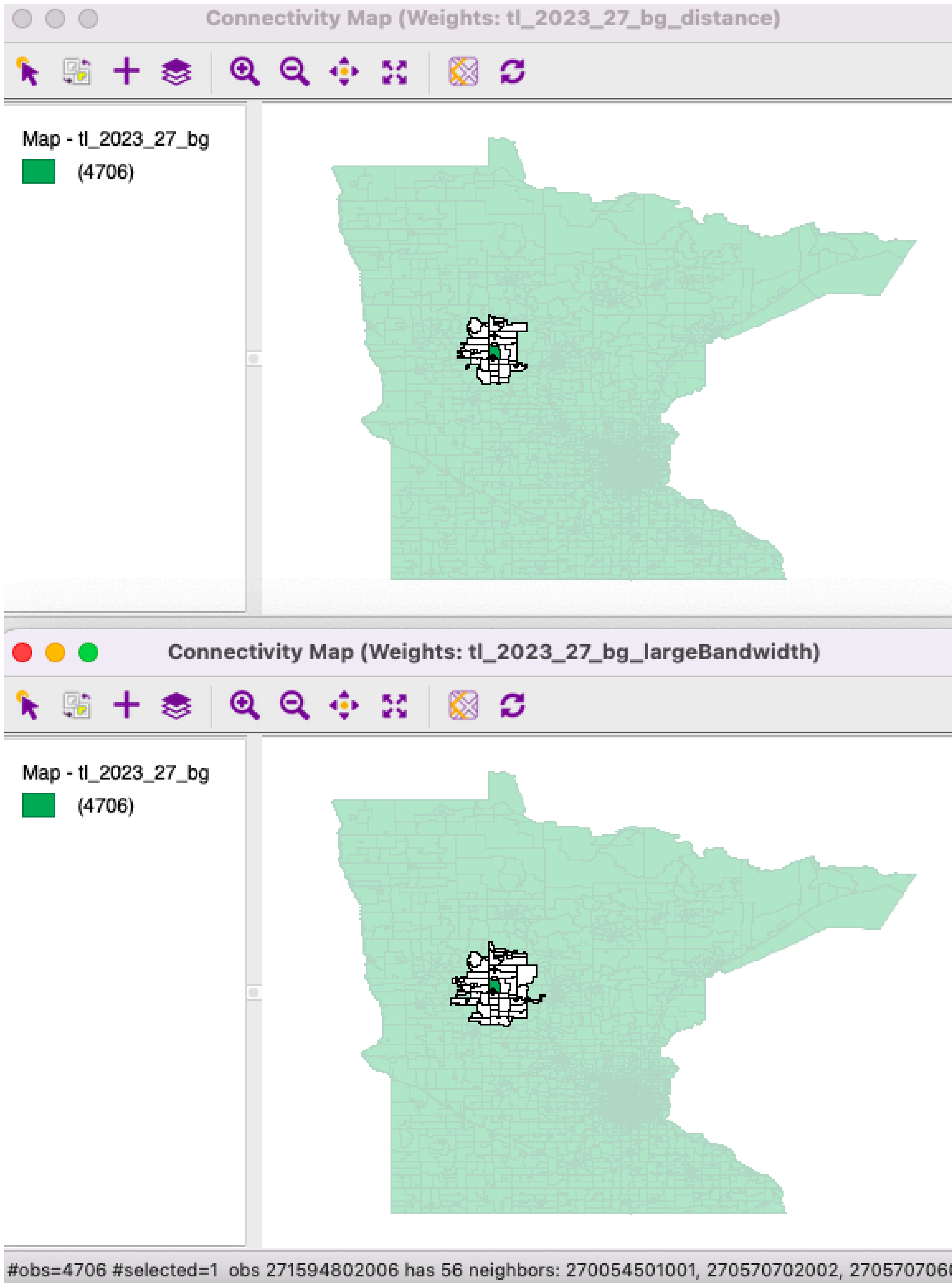
I created spatial weights using Queen Contiguity in GeoDa, with GEOID as the ID variable. The weights were saved to the project file, and identified the stored weights information in the project file. From the connectivity histogram, I selected the category with the number of neighbors and displayed the results in the linked map view. I saved the neighbor cardinality to the data table and created a choropleth map to visualize the distribution of neighbors. Finally, I checked the neighbors of one observation from the selected category using the connectivity map and graph.

- ***Task 2 Distance-band weights***
  - *Check the projection of your dataset, explain why you choose a certain distance metric to be used in your distance weights.*
  - *Save the distance-band weights to GWT file*
  - *Change your distance-band threshold (bandwidth) to a larger one and interpret the differences via a connectivity map.*

[Your answers]



2.

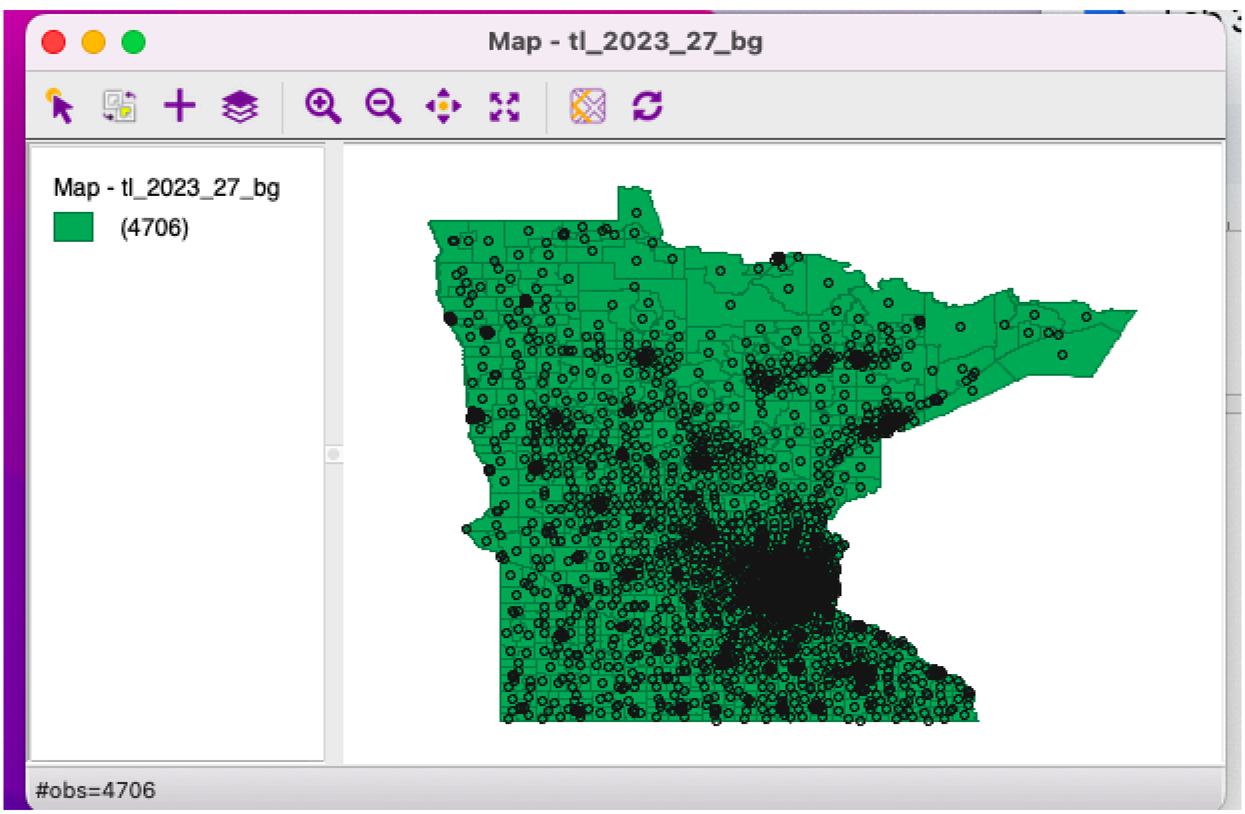

Connectivity Map (Weights: tl_2023_27_bg_distance)
Map - tl_2023_27_bg
(4706)
Connectivity Map (Weights: tl_2023_27_bg_largeBandwidth)
Map - tl_2023_27_bg
(4706)
#obs=4706 #selected=1 obs 271594802006 has 56 neighbors: 270054501001, 270570702002, 27057070

1.

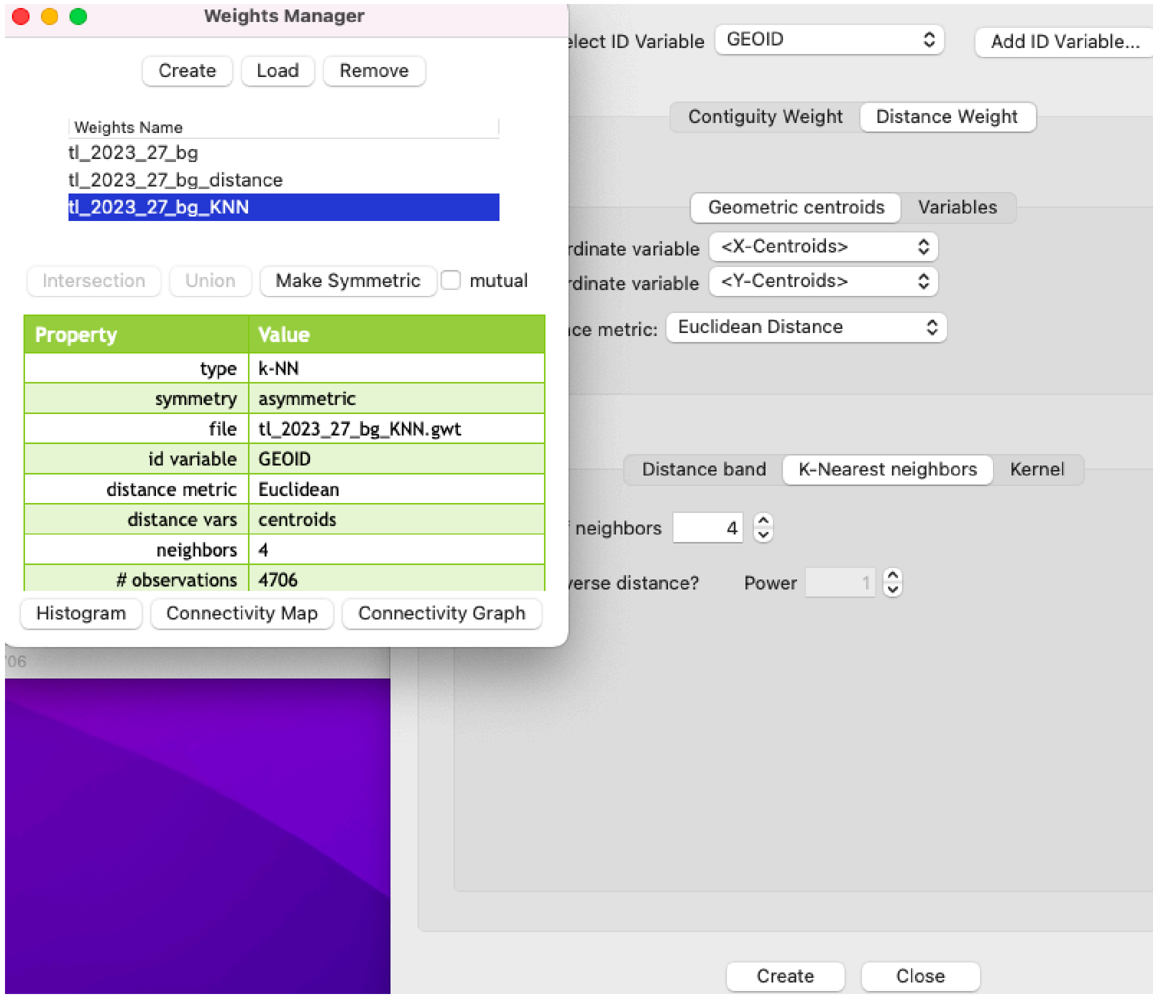
2. *I checked the projection of my dataset and chose **Euclidean distance** as the metric since the data was projected. I created distance band weights using **GEOID** as the ID variable and saved them as a GWT file. By increasing the distance band threshold, I observed more neighbors in the connectivity map, indicating higher spatial connectivity. This change was interpreted through a comparison of connectivity maps before and after adjusting the bandwidth.*

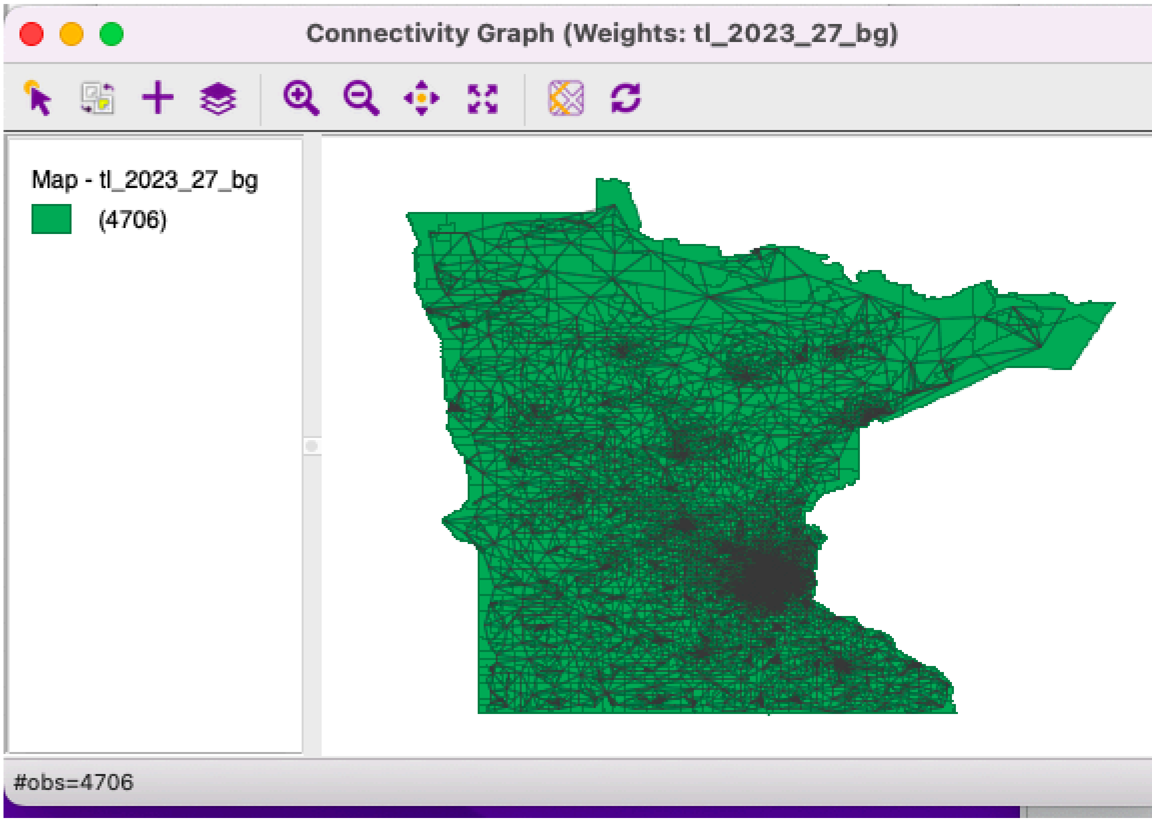
- ***Task 3 Generalizing the concept of contiguity***
  - *If your data is a point layer, create a rook contiguity map by explicitly call out the Thiessen polygon (do screenshots)*
    - *Create rook weights for your data*
    - *Display your connectivity graph on the Thiessen polygon*
  - *If your data is a polygon layer, derive the centroid points (do screenshots)*
    - *Create KNN weights for your data*
    - *Display your connectivity graph on the Point layer*

[Your answers]



1.


Weights Manager
Create
Load
Remove
Weights Name
tl_2023_27_bg
tl_2023_27_bg_distance
tl_2023_27_bg_KNN
Intersection
Union
Make Symmetric
mutual
Property
Value
type
k-NN
symmetry
asymmetric
file
tl_2023_27_bg_KNN.gwt
id variable
GEOID
distance metric
Euclidean
distance vars
centroids
neighbors
4
observations
4706
Histogram
Connectivity Map
Connectivity Graph
elect ID Variable
GEOID
Add ID Variable...
Contiguity Weight
Distance Weight
Geometric centroids
Variables
rdinate variable
<X-Centroids>
rdinate variable
<Y-Centroids>
ce metric:
Euclidean Distance
Distance band
K-Nearest neighbors
Kernel
neighbors
4
verse distance?
Power
1
Create
Close



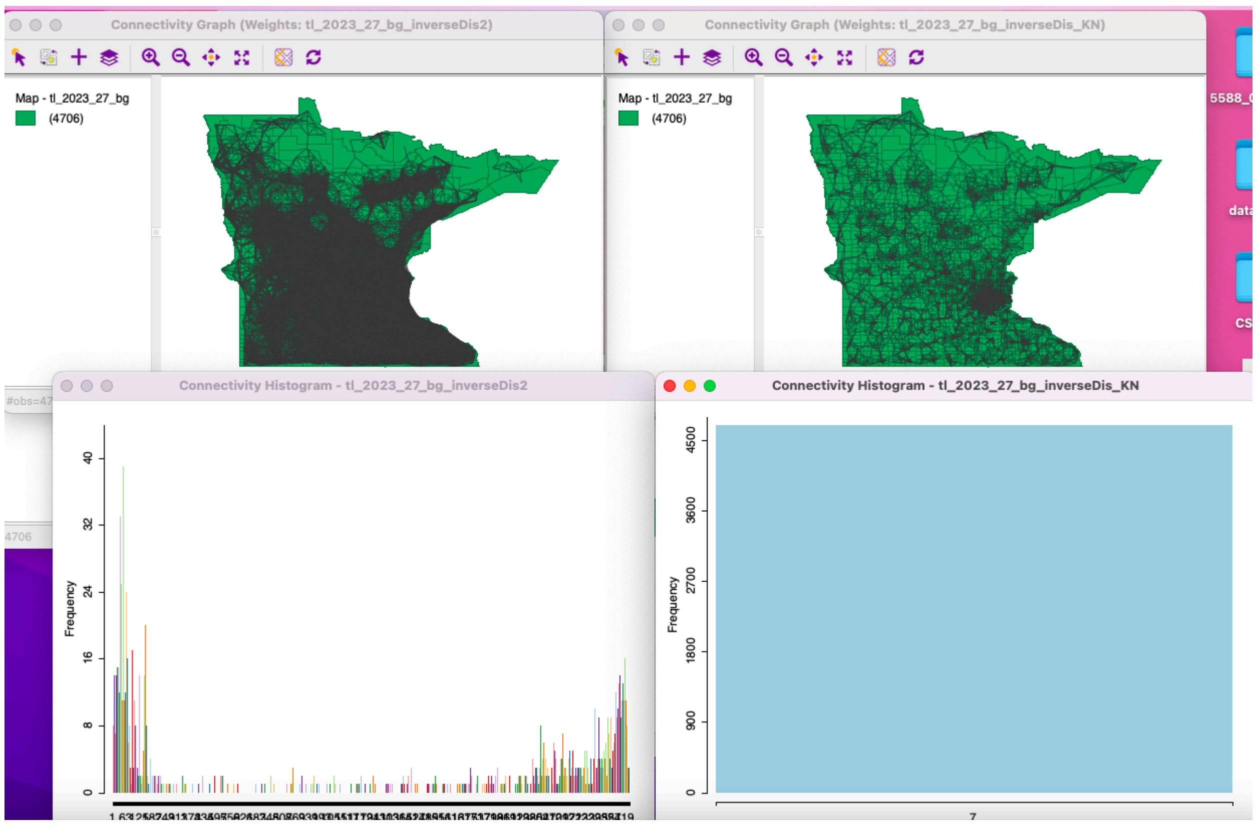
2.

Since my data was a polygon layer, I derived centroid points and created KNN weights using GEOID as the ID variable. I displayed the connectivity graph on the point layer, which showed the connections between the centroids based on the KNN criteria. This generalized the concept of contiguity by treating polygon centroids as points and applying distance based neighbor relationships.

- ***Task 4 Inverse distance weights***
  - *Derive and save the inverse distance weights based on distance band (use the power parameter to be 2, and make sure no isolate is in your spatial weights matrix)*
  - *Derive and save the inverse distance weights based on KNN (pick a proper K for your case)*
  - *Compare the connectivity graph/histograme for these two inverse distance weights, explain the differences. Open the corresponding weights file in txt reader, explain the differences.*

    [Your answers]

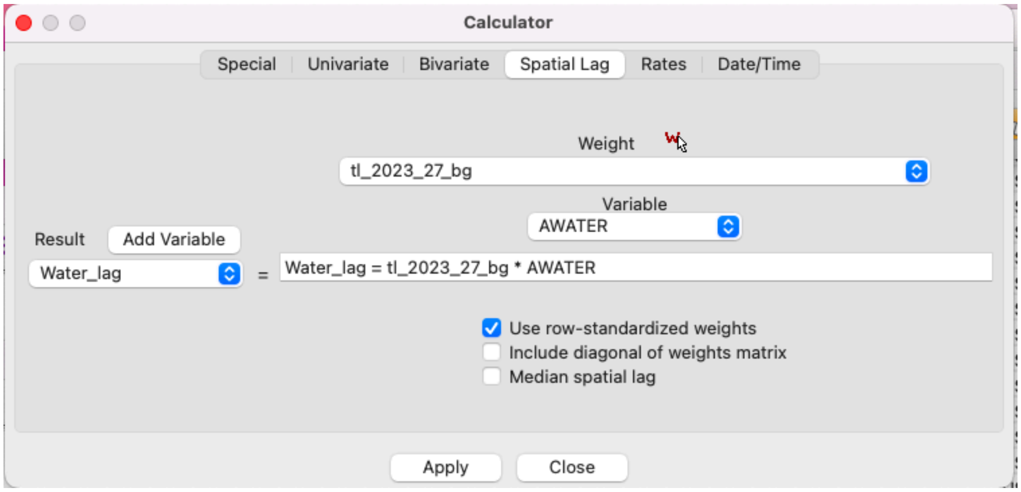
*I derived inverse distance weights using a **distance band** with a power parameter of 2 and ensured no isolates were present in the weights matrix. I also created inverse distance weights using **KNN** with K=7 and saved both sets of weights. Comparing the connectivity graphs, I found that the distance band weights resulted in varying numbers of neighbors, while KNN ensured each observation had exactly 7 neighbors. The GWT files showed differences in how neighbor IDs and distances were listed.*



1.

2.

- ➢ ***Task 5 Spatially lagged variable***
    - ○ *Add a spatially lagged variable based on your target variable using the row-standardized inverse distance spatial weights (pick your power and distance threshold).*



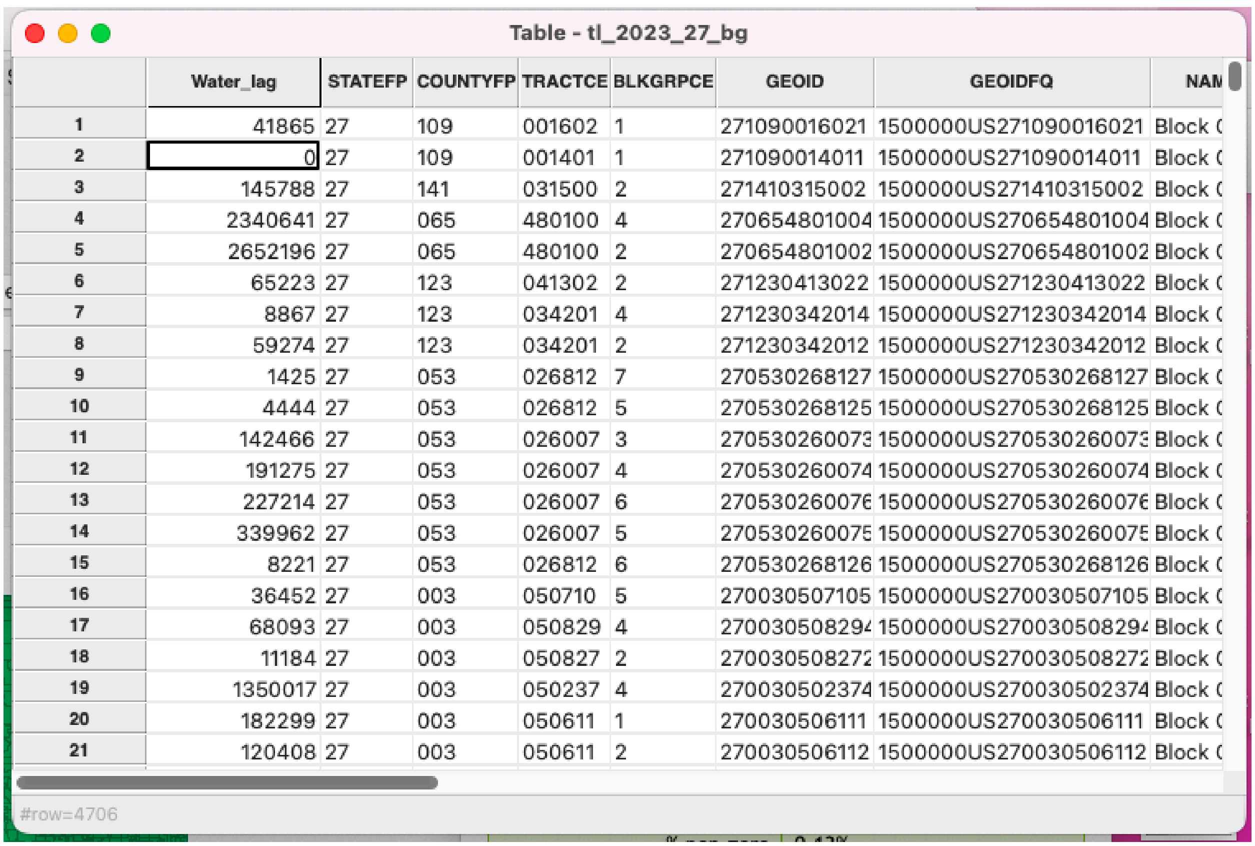
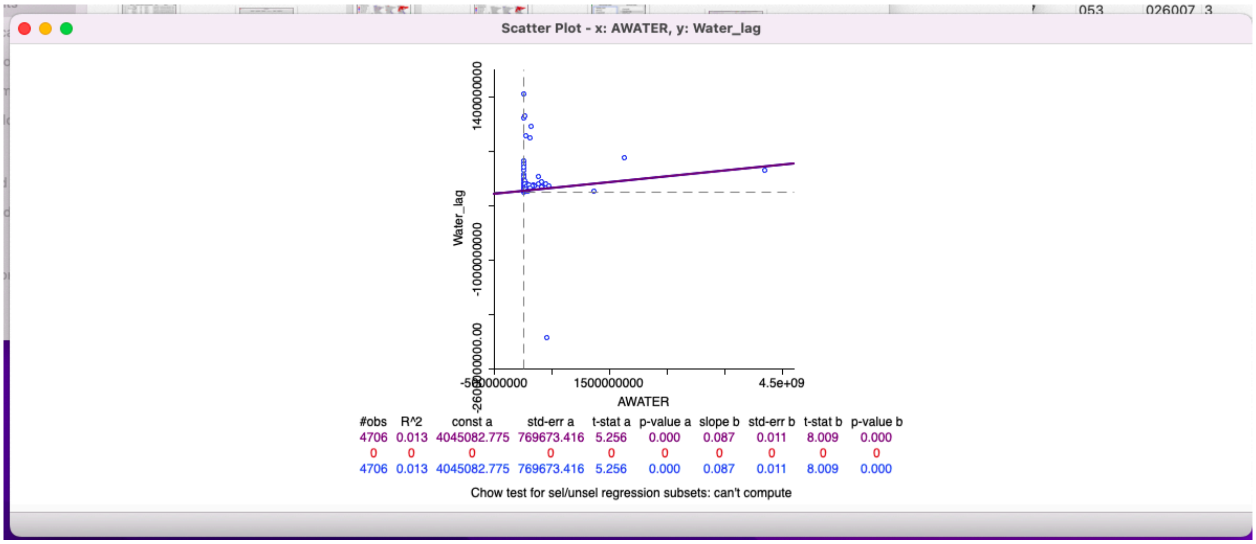


Table - tl_2023_27_bg

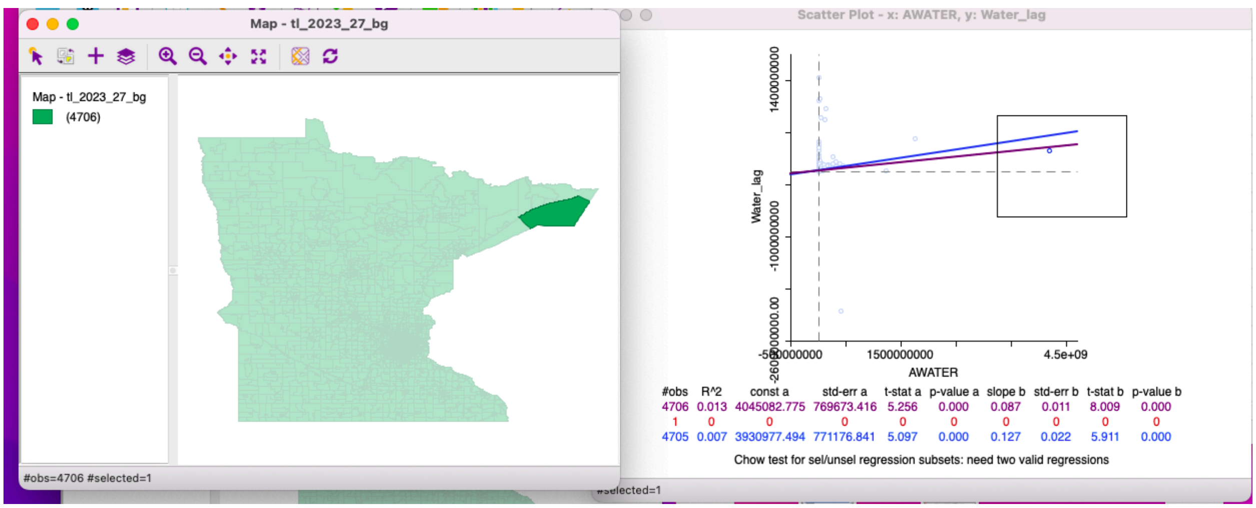
	Water_lag	STATEFP	COUNTYFP	TRACTCE	BLKGRPCE	GEOID	GEOIDFQ	NAM
1	41865	27	109	001602	1	271090016021	1500000US271090016021	Block (
2	0	27	109	001401	1	271090014011	1500000US271090014011	Block (
3	145788	27	141	031500	2	271410315002	1500000US271410315002	Block (
4	2340641	27	065	480100	4	270654801004	1500000US270654801004	Block (
5	2652196	27	065	480100	2	270654801002	1500000US270654801002	Block (
6	65223	27	123	041302	2	271230413022	1500000US271230413022	Block (
7	8867	27	123	034201	4	271230342014	1500000US271230342014	Block (
8	59274	27	123	034201	2	271230342012	1500000US271230342012	Block (
9	1425	27	053	026812	7	270530268127	1500000US270530268127	Block (
10	4444	27	053	026812	5	270530268125	1500000US270530268125	Block (
11	142466	27	053	026007	3	270530260073	1500000US270530260073	Block (
12	191275	27	053	026007	4	270530260074	1500000US270530260074	Block (
13	227214	27	053	026007	6	270530260076	1500000US270530260076	Block (
14	339962	27	053	026007	5	270530260075	1500000US270530260075	Block (
15	8221	27	053	026812	6	270530268126	1500000US270530268126	Block (
16	36452	27	003	050710	5	270030507105	1500000US270030507105	Block (
17	68093	27	003	050829	4	270030508294	1500000US270030508294	Block (
18	11184	27	003	050827	2	270030508272	1500000US270030508272	Block (
19	1350017	27	003	050237	4	270030502374	1500000US270030502374	Block (
20	182299	27	003	050611	1	270030506111	1500000US270030506111	Block (
21	120408	27	003	050611	2	270030506112	1500000US270030506112	Block (

#row=4706

- *Create a scatterplot with your target variable as the independent var. and the spatially lagged variable as the dependent var.*



Brush and select some observations in your scatterplot that seems to be far away from the linear fit. What does it imply with the reference of a choropleth map?



-

[Your answers]

1. *I added a spatially lagged variable inverse distance weights and GEOID as the ID variable. A scatterplot was created with the target variable as the independent variable(AWATER) and the spatially lagged variable (Water_lag)as the dependent variable. Observations far from the linear fit were brushed and selected, indicating potential outliers. These observations were cross referenced with a choropleth map to analyze their spatial distribution.*

2.