

TUTORIAL 5 | FEATURE DISTANCE ANALYSIS

Zoom: **Friday 06.03, 6:30-7:30pm (extra tutorial this week)**

<https://virginiatech.zoom.us/j/2981092726>

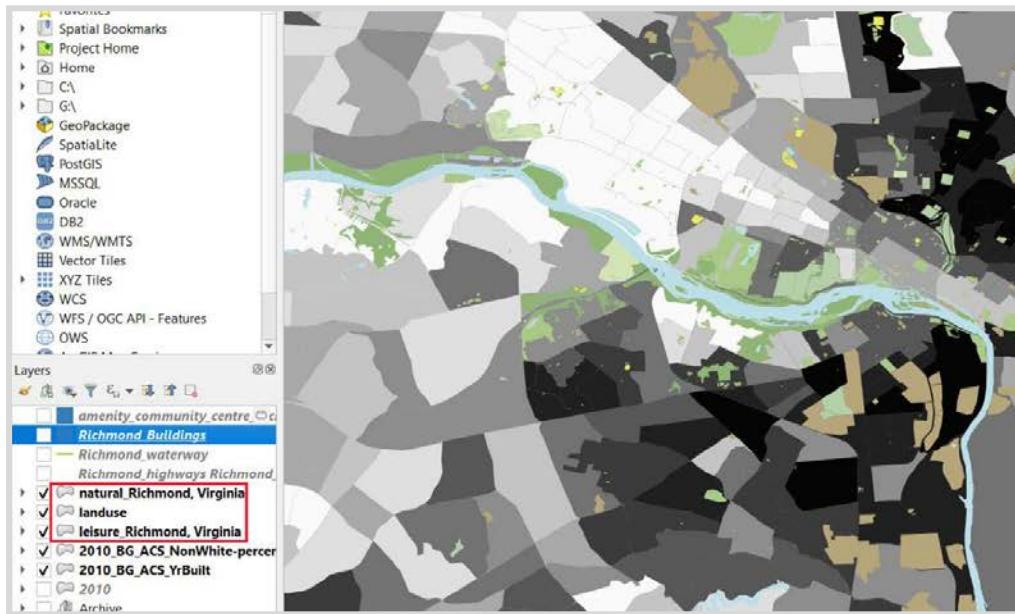
Goals

- Compare census data with urban amenities.
- Filter features with function equations.
- Find the centerpoint of features.
- Create a distance matrix to analyze the relationship between data sets.
- Compare buffers with Isochrones (map distance with experienced distance)

NOTE: as you finish more tutorials, the directions will become more general, relying on skills and processes that you've practiced in previous tutorials. Please refer to earlier tutorial 1 for questions about loading OSM data.

Step 1: Download OSM data for your city

1a Using the “QuickOSM” plugin, download “leisure”, “natural”, and “landuse” categories for your Appalachian city.

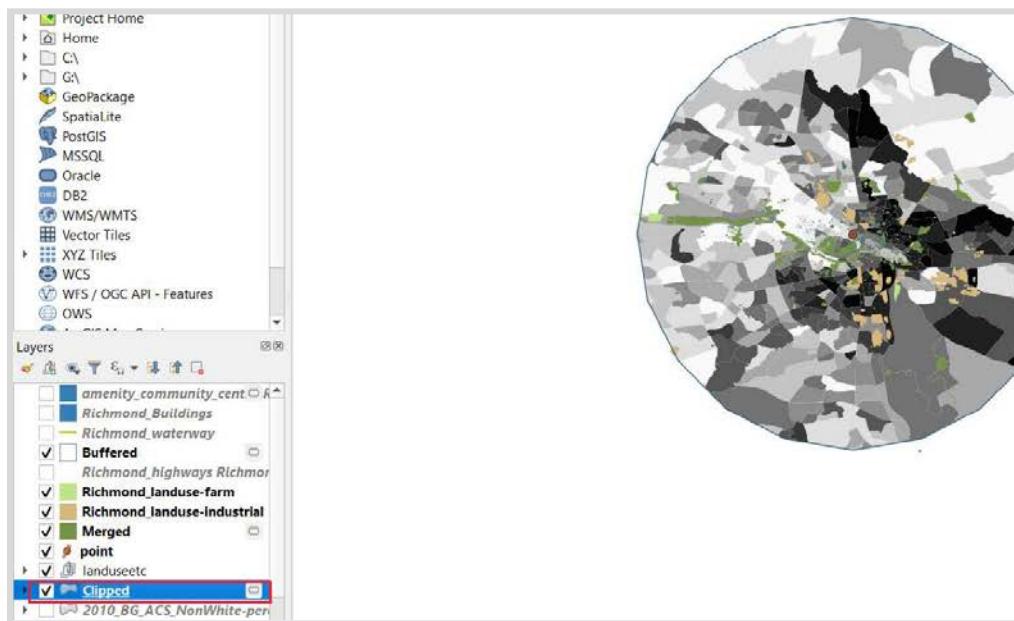


Step 2: Buffer census data

2a Create a **Shapefile layer**, and draw a **point** in the middle of your city. Note: check that the new vector layer uses the same **CRS** as your project.

2b **Buffer** around the point. Use the buffer to **Clip the Nonwhite population census layer**.

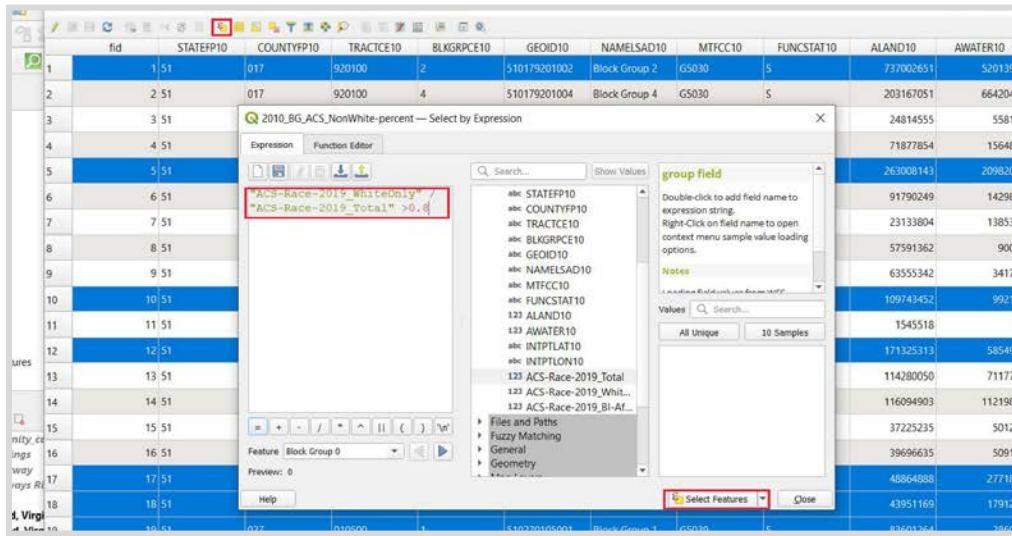
Note: you may need to re-style your layer after clipping. Make sure you save the scratch layer as a permanent layer before you do this.



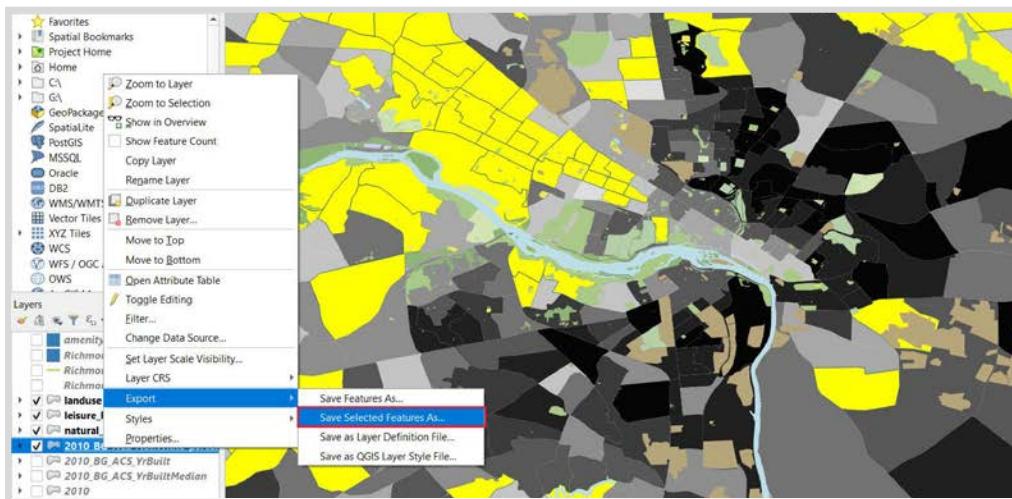
Step 3: Search by feature and export the highest 20% of nonwhite and white census blocks

To compare the location of urban features with population characteristics, we need to separate out the census block groups with the highest percentage of non-white population and the lowest percentage. We'll use 20% as a cut-off: block groups with under 20% white (or, over 80% non-white), and under 20% non-white (over 80% white).

3a First, open the **Attribute Table** of your census layer, and select features using “**Select by Expression**”, the Sigma symbol in the upper toolbar. Search by expression has a similar interface to the Graduated Symbology function creator we've used in previous tutorials. In this case, we'll find the ratio of White population by dividing White by Total; then, add a greater than symbol and 0.8 (eg **census blocks where white population is greater than 0.8 of total**).



3b Check in your map that the right polygons have **highlighted**. Then, right click on the census layer and **Export > Save Selected Features As...** to save your selection as a new layer.



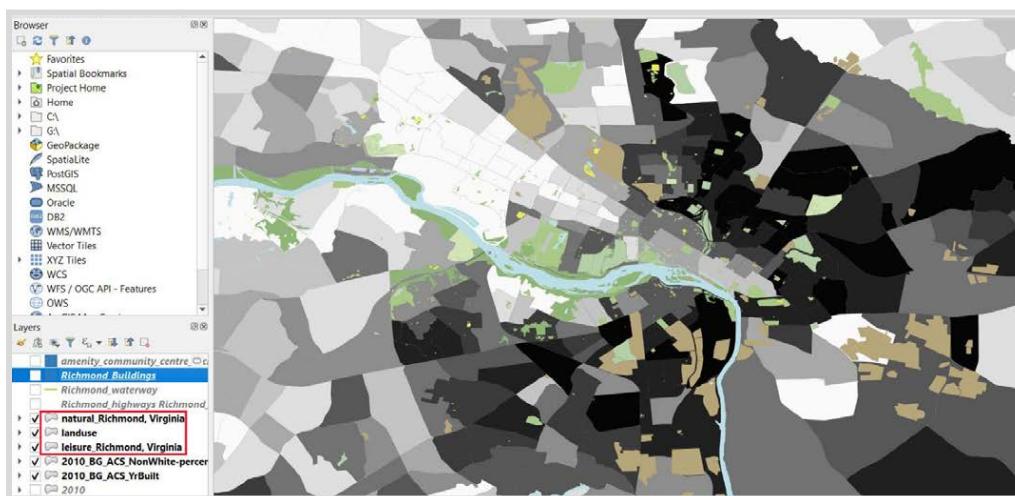
3c Repeat 3a-3b for white population **less than 0.2 of total** (eg non-white is great than 80%). **Export** selection as another new layer.

Step 4: Search by feature and export your chosen urban feature as a separate layer.

4a Style the urban feature layers (leisure, natural, and landuse) to find adjacencies with concentrations of nonwhite or white population areas.

In my example, I looked at areas of greenspace and of industry. Industrial areas can be found entirely in the landuse layer (I **combined** “industrial”, “quarry”, “landfill”, “construction”, and “brownfield”: **Layer Properties > Symbology > Categorized by “landuse”**, selecting the fields, right clicking, and “**Merge**”). Natural areas, on the other hand, I found spread across the three layers. In landuse I **Merged** the categorized “reservoir”, “recreation_ground”, “grass”, “forest”, “flowerbed”, “cemetery”, and “meadow”; in leisure, “park”, “pitch”, “playground”, “garden”; and in natural, most of the categories qualify as greenspace.

When I **styled the layers** accordingly, I could immediately see that the industrial areas fell mostly in nonwhite neighborhoods, and green spaces tended to be surrounded by majority white neighborhoods. It’s important to do this kind of investigation **before** you choose which features to export, so you have a visual sense of what adjacencies might exist. Here, the industrial areas are colored brown and greenspace green:



4b Once you've found a pattern with some chosen urban features, go to their attribute tables and select them using “Select by Expression”. You can use “OR” to select multiple types of features at the same time.

2	r1236828	1236828	relation	water	NULL	NULL	NULL	NULL	NULL
3	r1236830	1236830							NULL
4	r1656327	1656327							NULL
5	r3551641	3551641							NULL
6	r4117093	4117093							NULL
7	r6036457	6036457							NULL
8	r6243498	6243498							NULL
9	r6243499	6243499							NULL
10	r7119708	7119708							NULL
11	r7602313	7602313							NULL
12	r9294512	9294512							NULL
13	r10966141	10966141							NULL
14	r11271935	11271935							NULL
15	r11275266	11275266							NULL
16	r12607906	12607906							NULL
17	r12674673	12674673							NULL
18	r12809295	12809295							NULL
19	r12824295	12824295	relation	water	NULL	NULL	NULL	NULL	NULL

Q natural_Richmond_Virginia — Select by Expression

Expression Function Editor

"natural" = 'scrub' OR "natural" = 'wetland' OR "natural" = 'wood'

operator OR

Returns 1 when condition a or b is true.

Syntax

Examples

- * 4 = 2+2 OR 1 = 1 → 1
- * 4 = 2+2 OR 1 = 2 → 1
- * 4 = 2 OR 1 = 2 → 0

Feature Preview: 0

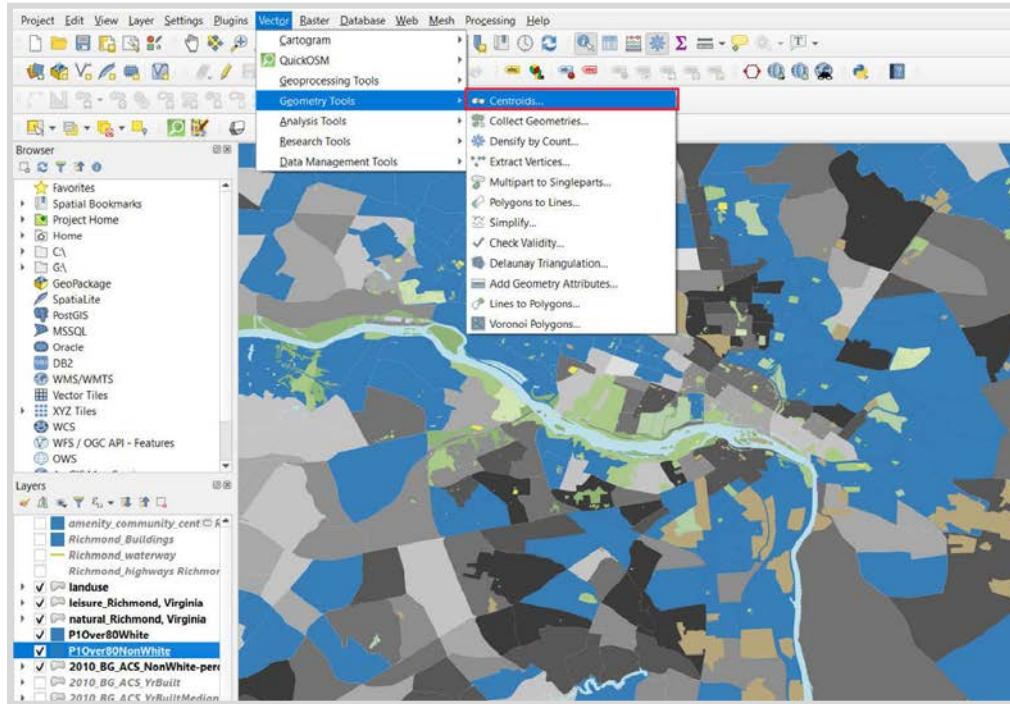
Help Selected Features Close

4c When you've selected the features you want, right click on the layer and "Export Selection As..." for each layer.

If you've **exported features from multiple layers** (like I did with greenspace features from Leisure, Land Use, and Natural layers), merge them with **Vector > Data Management Tools > Merge Vector Layers...** So that you have one combined layer for each aspect you want to compare (eg one Greenspace layer and one Industrial layer).

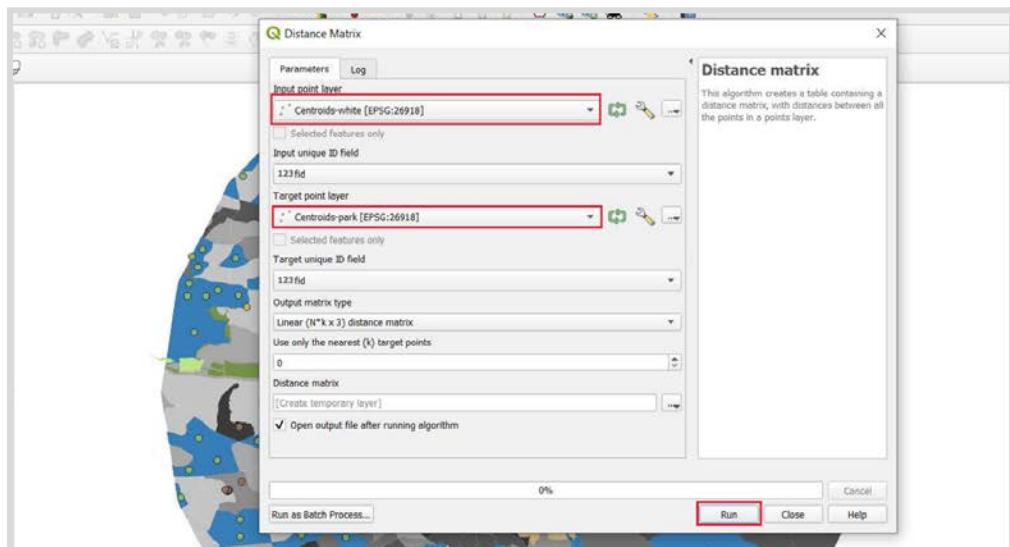
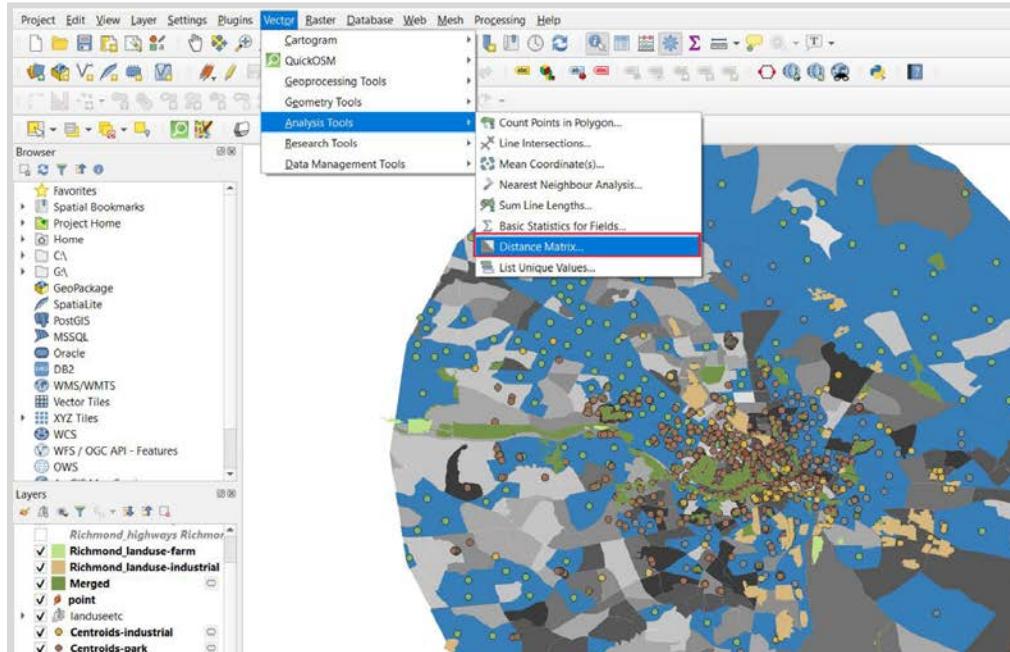
Step 5: Create centerpoints for the highest and lowest percentage of white population areas, and for the merged urban features.

5a Vector > Geometry Tools > Centroids. Select each merged layer.



Step 6: Create a Distance Matrix for each combination of population and urban characteristics.

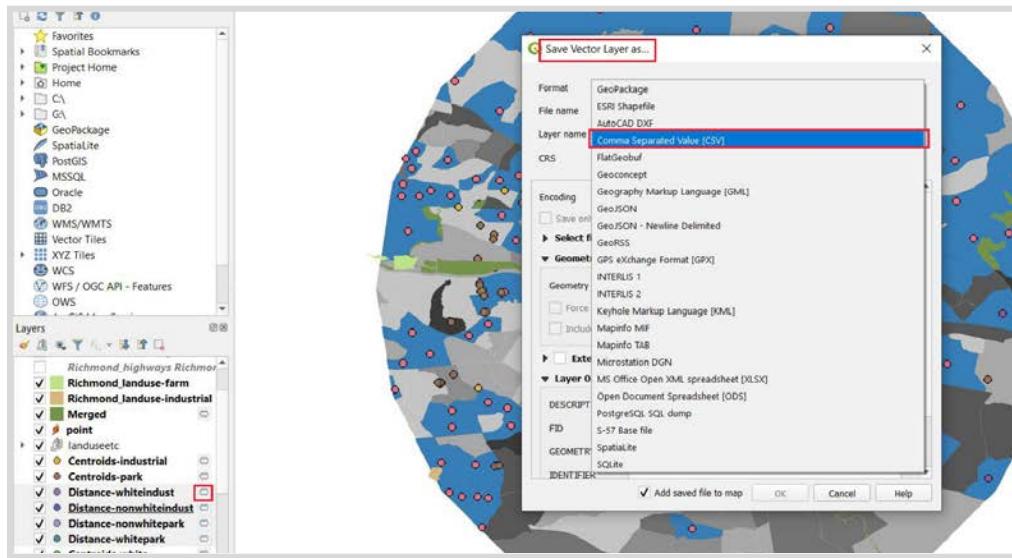
6a Open Vector > Analysis Tools > Distance Matrix. Input either the 80% White or 80% NonWhite centroid layer, and Target on of your urban feature centroid layers (eg. Greenspace or Industrial). Click “Run” to generate the new Matrix layer.



Repeat with each combination of centroid layers (eg White-Park, White-Industrial, Nonwhite-Park, Nonwhite-Industrial).

Step 7: Export the Distance Matrices as .csv files to calculate their averages.

7a Right click on each distance matrix layer (or click the scratch layer box symbol highlighted below), and Export As... Instead of "Geopackage", export as a "Comma Separated Value (CSV)".



7b Open the .csv files in Google Sheets. Use the "Average" formula to calculate the average of each "Distance" column (=AVERAGE(C:C)).

	A	B	C	D	E	F	G
1	InputID	TargetID	Distance				
2	437	313	0.111085		0.160339		
3	437	303	0.109487				
4	437	298	0.109518				
5	437	297	0.10962				
6	437	296	0.10981				
7	437	295	0.109907				
8	437	241	0.141558				
9	437	240	0.139469				

=AVERAGE(C:C)

Dist_whitepark

	A	B	C	D	E	F	G
1	InputID	TargetID	Distance				
2	445	313	0.15354		0.075106		
3	445	303	0.151361				
4	445	298	0.15198				
5	445	297	0.151829				
6	445	296	0.151784				
7	445	295	0.151739				
8	445	241	0.107583				
9	445	240	0.109635				

=AVERAGE(C:C)

Dist_nonwhitepark

	A	B	C	D	E	F	G
1	InputID	TargetID	Distance				
2	437	544	0.190509		0.148928		
3	437	543	0.195538				
4	437	542	0.20563				
5	437	423	0.240912				
6	437	419	0.153011				
7	437	390	0.195273				
8	437	389	0.21478				
9	437	377	0.225567				

=AVERAGE(C:C)

Dist_whiteindust

	A	B	C	D	E	F	G
1	InputID	TargetID	Distance				
2	445	544	0.090912		0.086082		
3	445	543	0.08961				
4	445	542	0.072264				
5	445	423	0.164765				
6	445	419	0.159664				
7	445	390	0.088251				
8	445	389	0.033107				
9	445	377	0.086564				

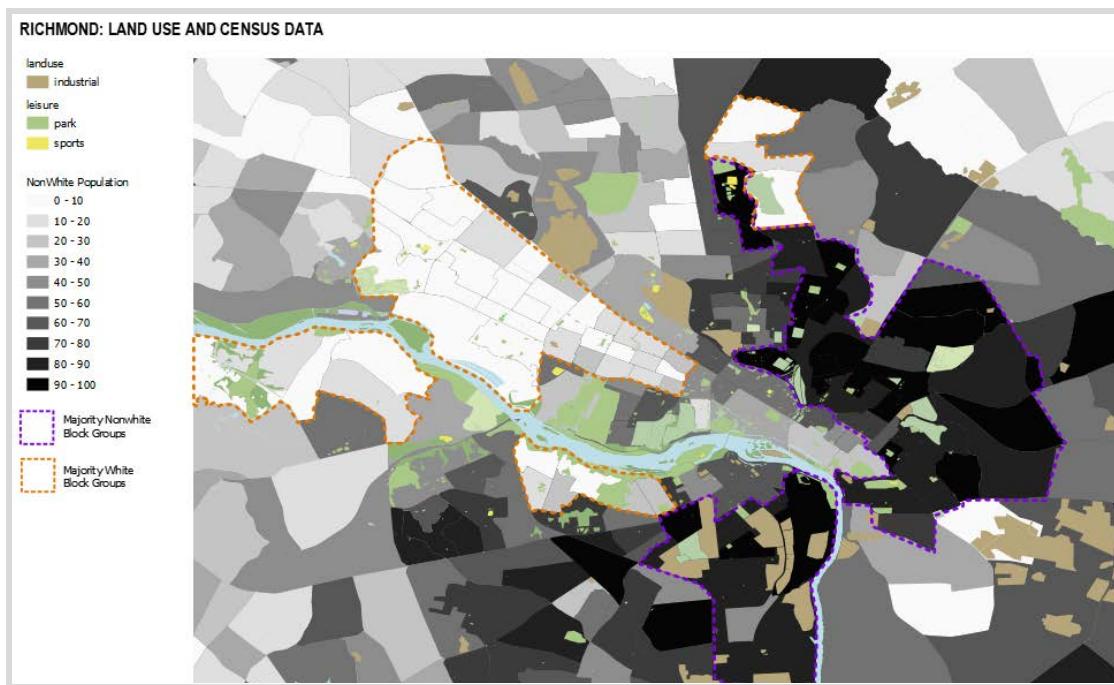
=AVERAGE(C:C)

Dist_nonwhiteindust

With my Distance Matrix calculations, I found that the White areas are more than twice as likely to be near Parks as NonWhite. I also found that White areas are more likely to be near Industrial (probably because there are lot more majority White block groups in Richmond); however, they're LESS than half as likely to be near industrial areas as Nonwhite areas. From this, I conclude that nonwhite neighborhoods in Richmond are disproportionately near industrial areas versus greenspace.

Step 8: Create a map layout for the new map.

Draw lines to indicate the areas where population characteristics correspond with urban features.



Step 9: Export map as pdf with legend.

Please be sure to include your distance matrix calculations.

- No Bonus -