GIS for Designers
VT | A+D | FA22

TUTORIAL 6 | FEATURE DISTANCE ANALYSIS

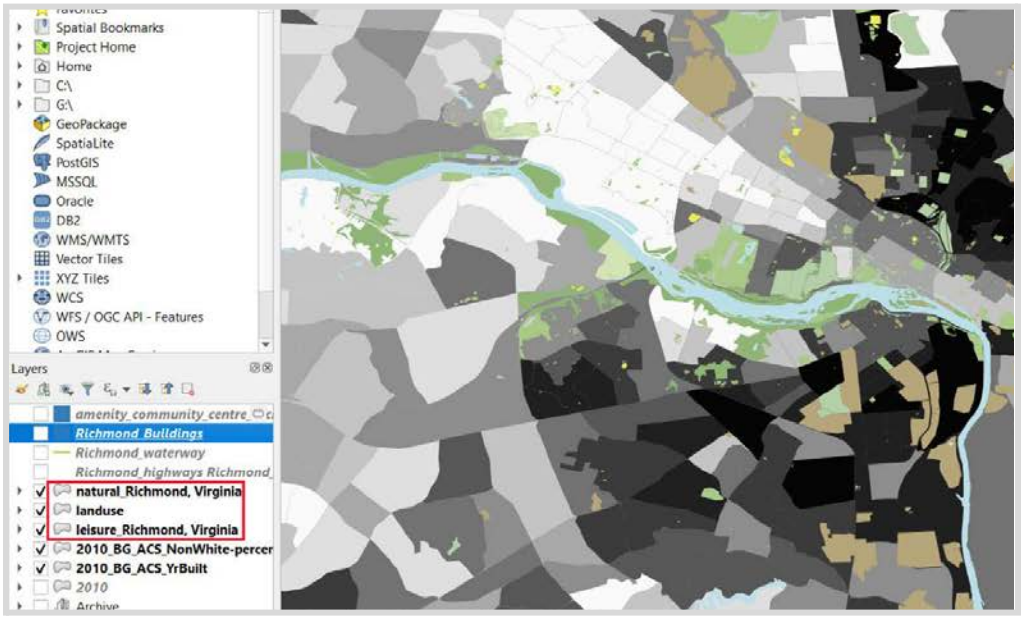
Goals

- 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. Note: you will only be using the "multipolygon" layers, so you can delete the point and line layers that OSM downloads.



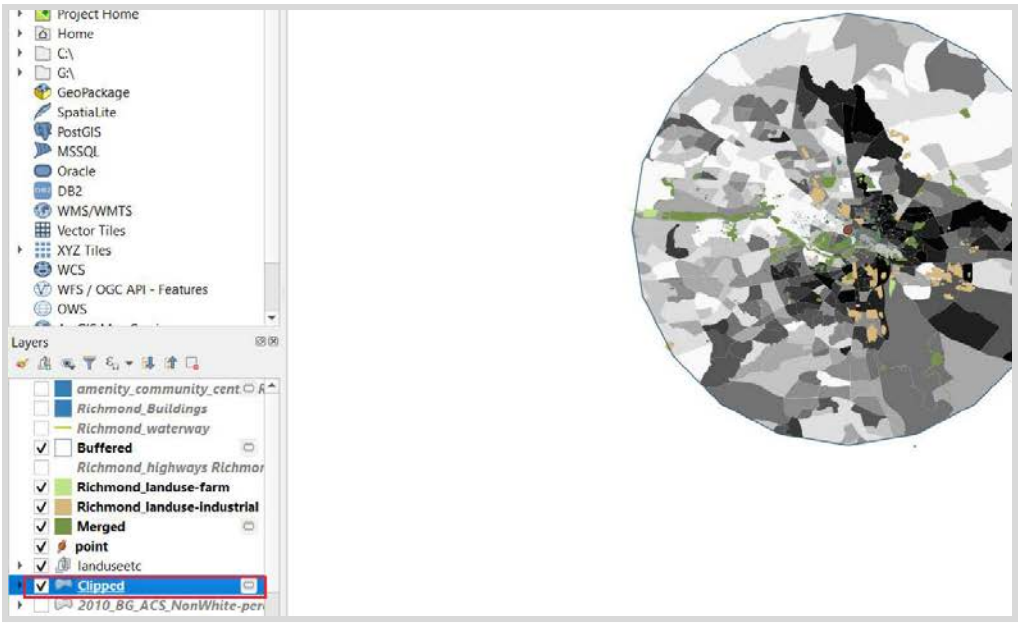
Step 2: Buffer census data

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 **Clip your Nonwhite population census layer by either:**

> **2b-option 1 Buffer** around the point. Use the buffer to **Clip the Nonwhite population census laye**r.
>
> (see instructions on Steps 4-6 of an old tutorial here)
>
> **2b-option 2 Alternatively, download the municipal boundary** of your city and use this to clip the census layer, like in Tutorial 2. You can check the correct “admin-level” to use by going to https://www.openstreetmap.org and using the identify tool. It will probably be admin-level 6 in the OSM search.

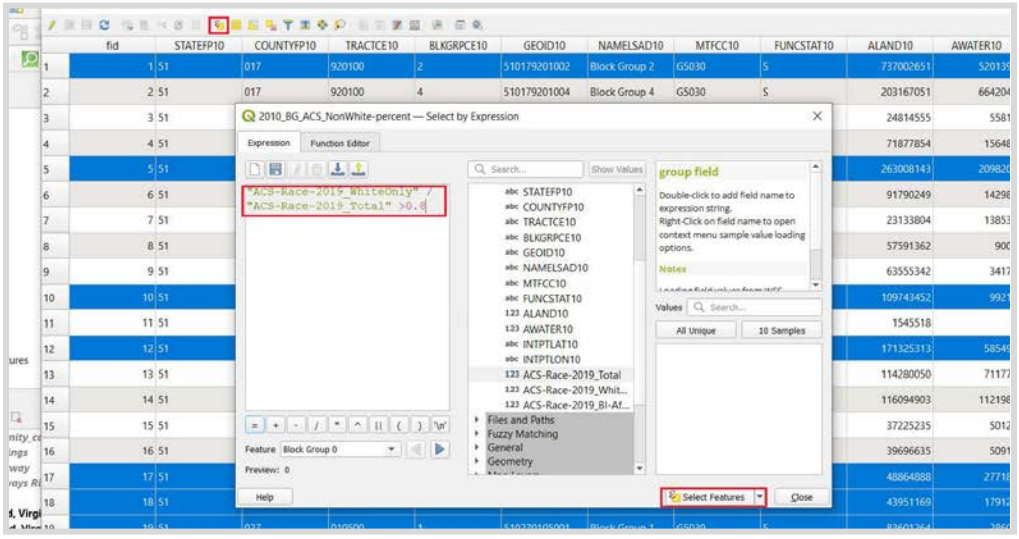
Note: you may need to re-style your layer after clipping. Make sure you save the scratch layer as a permanent layer before you do this.



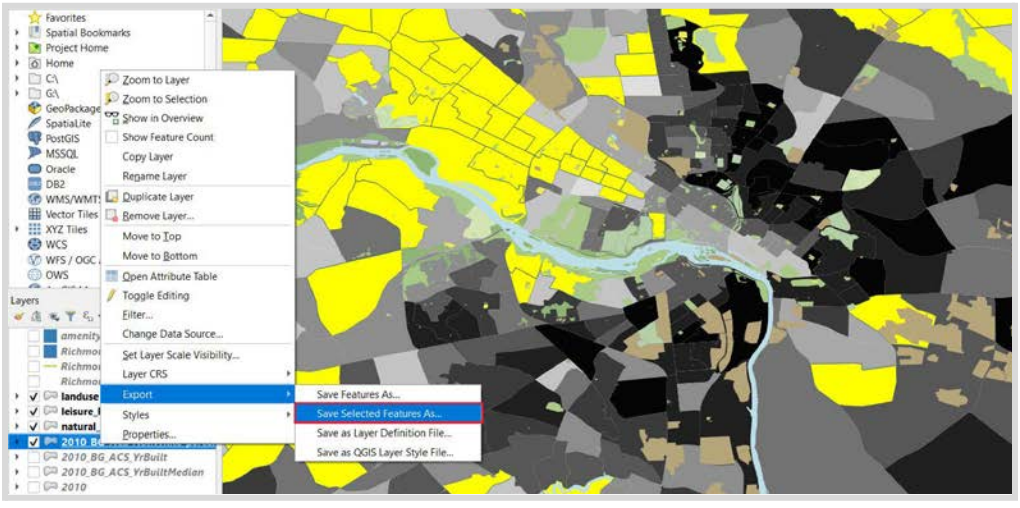
Step 3: Search by feature and export the highest 20% of nonwhite and white census blocks

> 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.

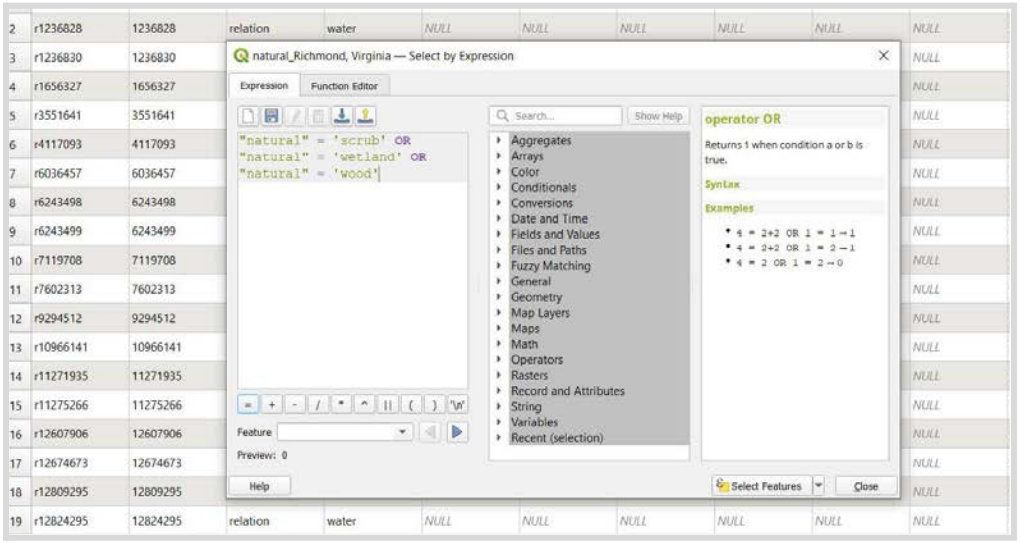
Step 4: Search by feature and export your chosen urban feature as a separate layer.

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.

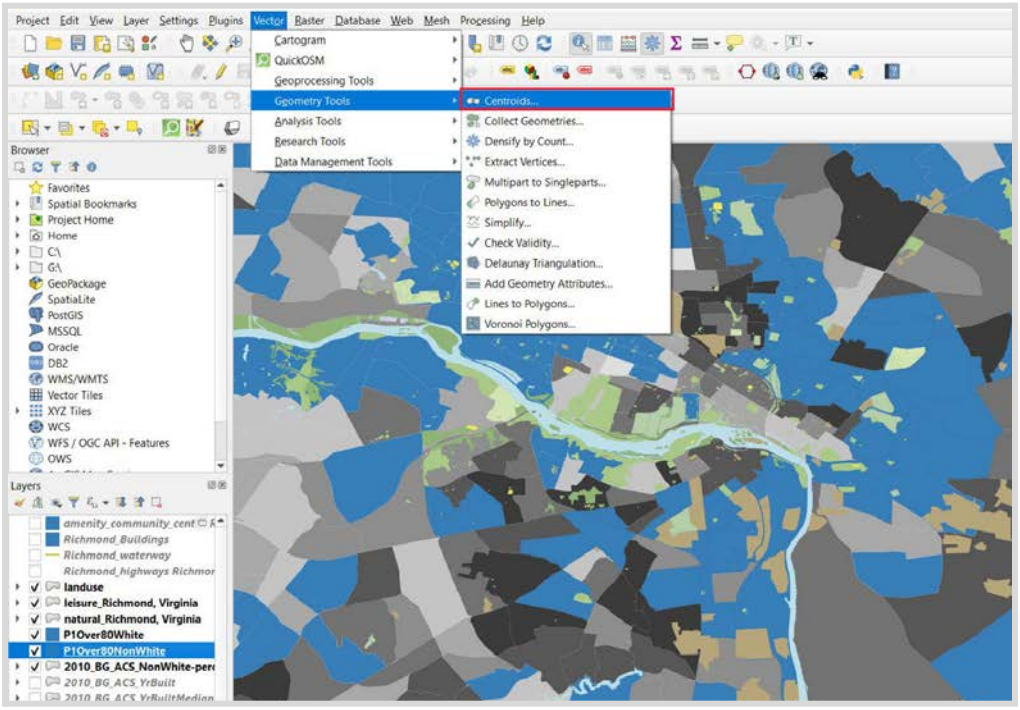


4c When you've selected the features you want, right click on the layer and "Export Selection As…" for each layer.

> 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).

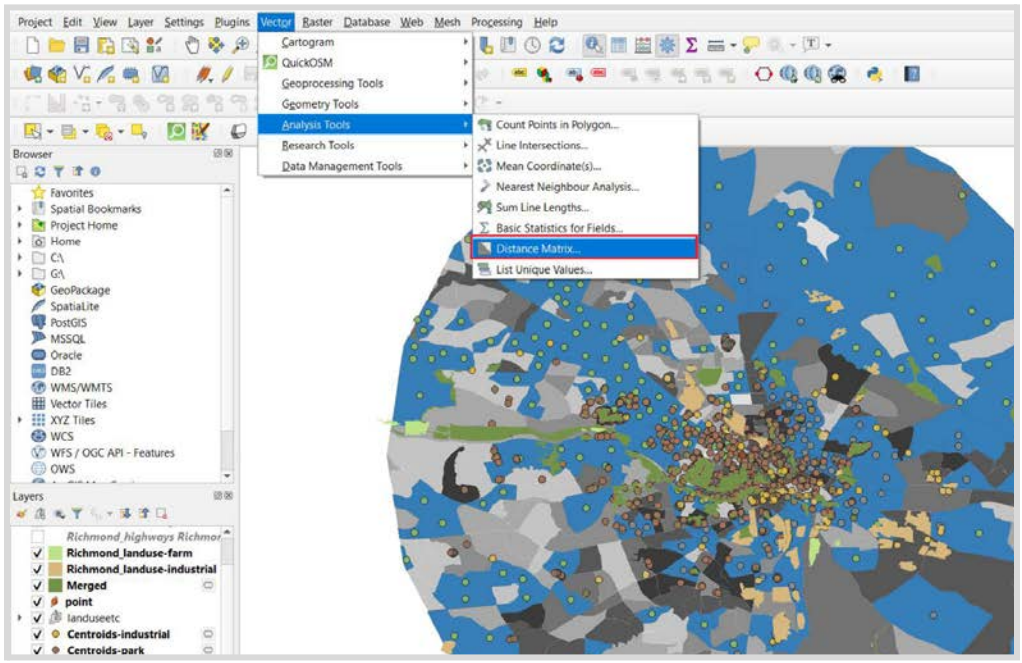
Step 5: Create centerpoints for the highest and lowest percentage of white population areas, and for the merged urban features.

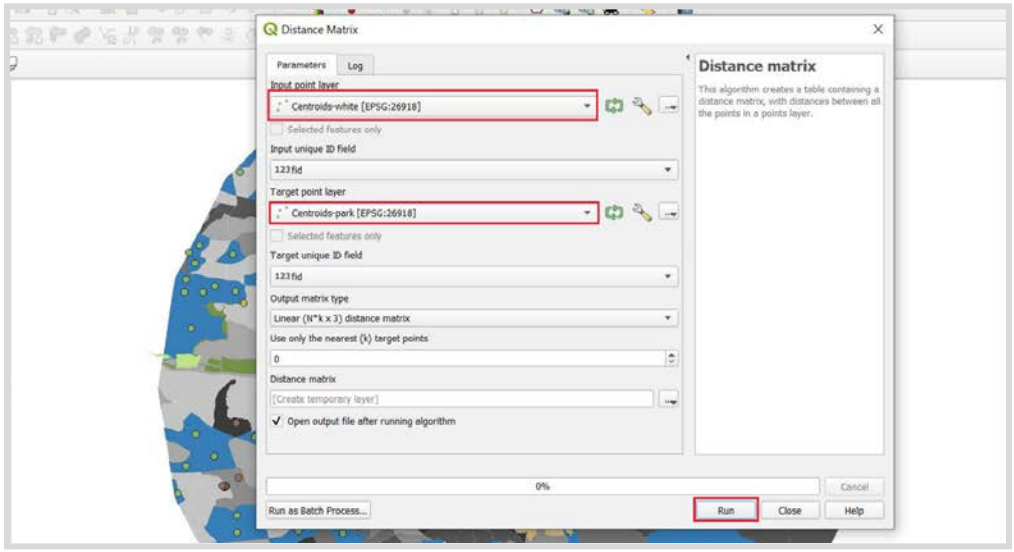
5a Vector > Geometry Tools > Centroids. Select each merged layer.



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.

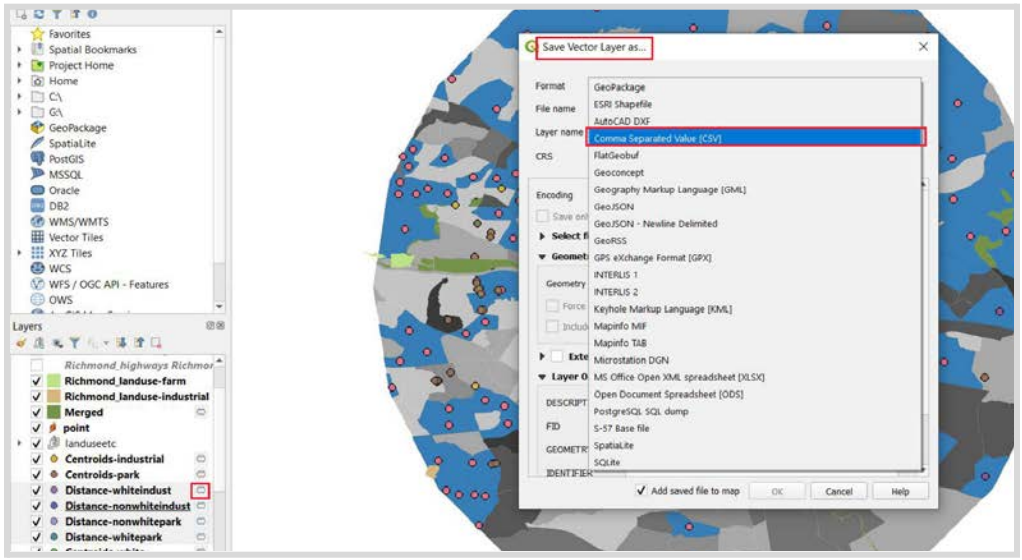




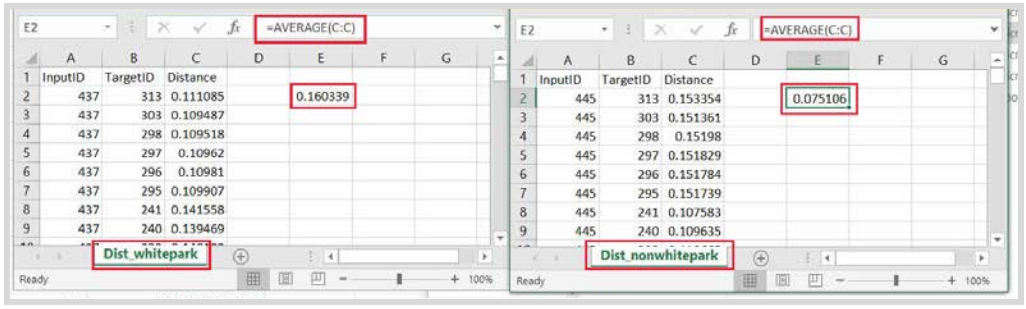
Repeat with each combination of centroid layers (eg White-Park, White-Industrial, Nonwhite-Park, Nonwhite-Industrial).

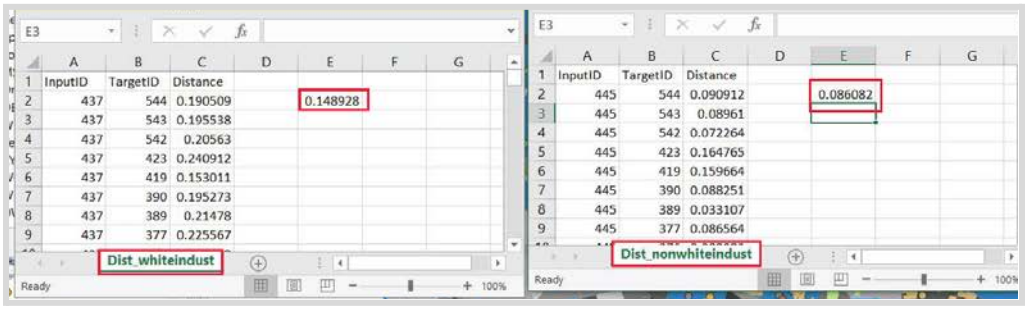
Step 7: Export the Distance Matrices as .csv files to calculate their averages.

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)).

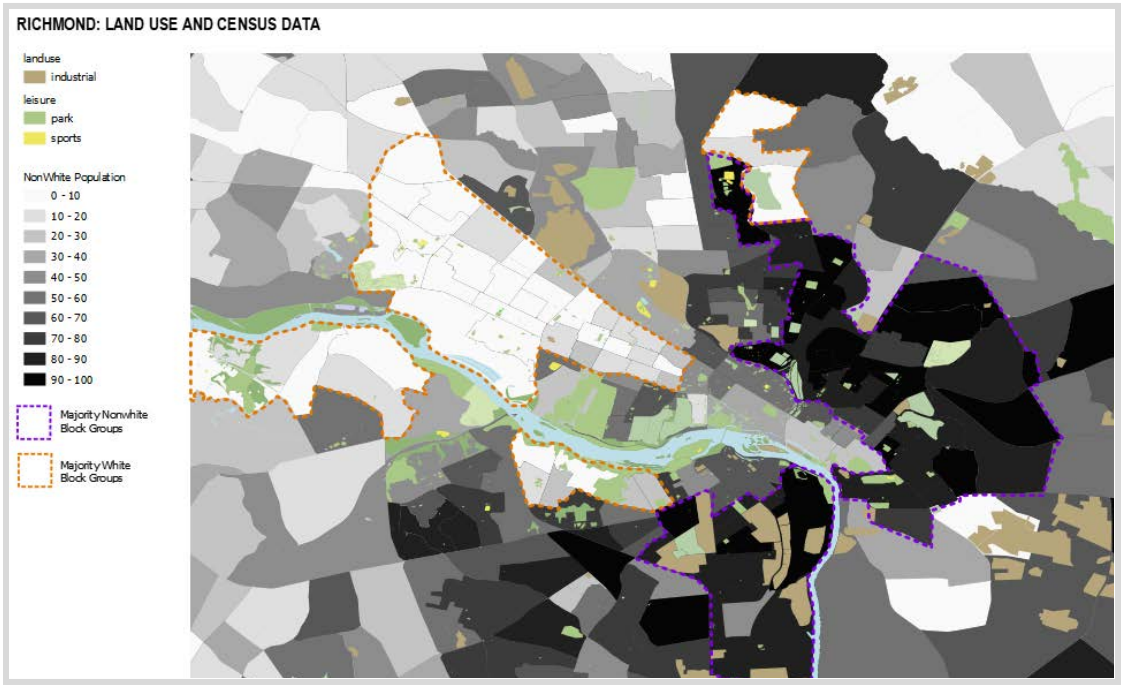




With my Distance Matrix calculations, I found that the White areas are more than twice as likely to be near Parks as NonWhite. I also found that White areas are more likely to be near Industrial (probably because there area lot more majority White block groups in Richmond); however, they’re LESS than 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.

In your final tutorial submission for 4-6, please be sure to include your distance matrix calculations, both screenshots of your spreadsheet tables and your final average distance numbers.

- No Bonus -