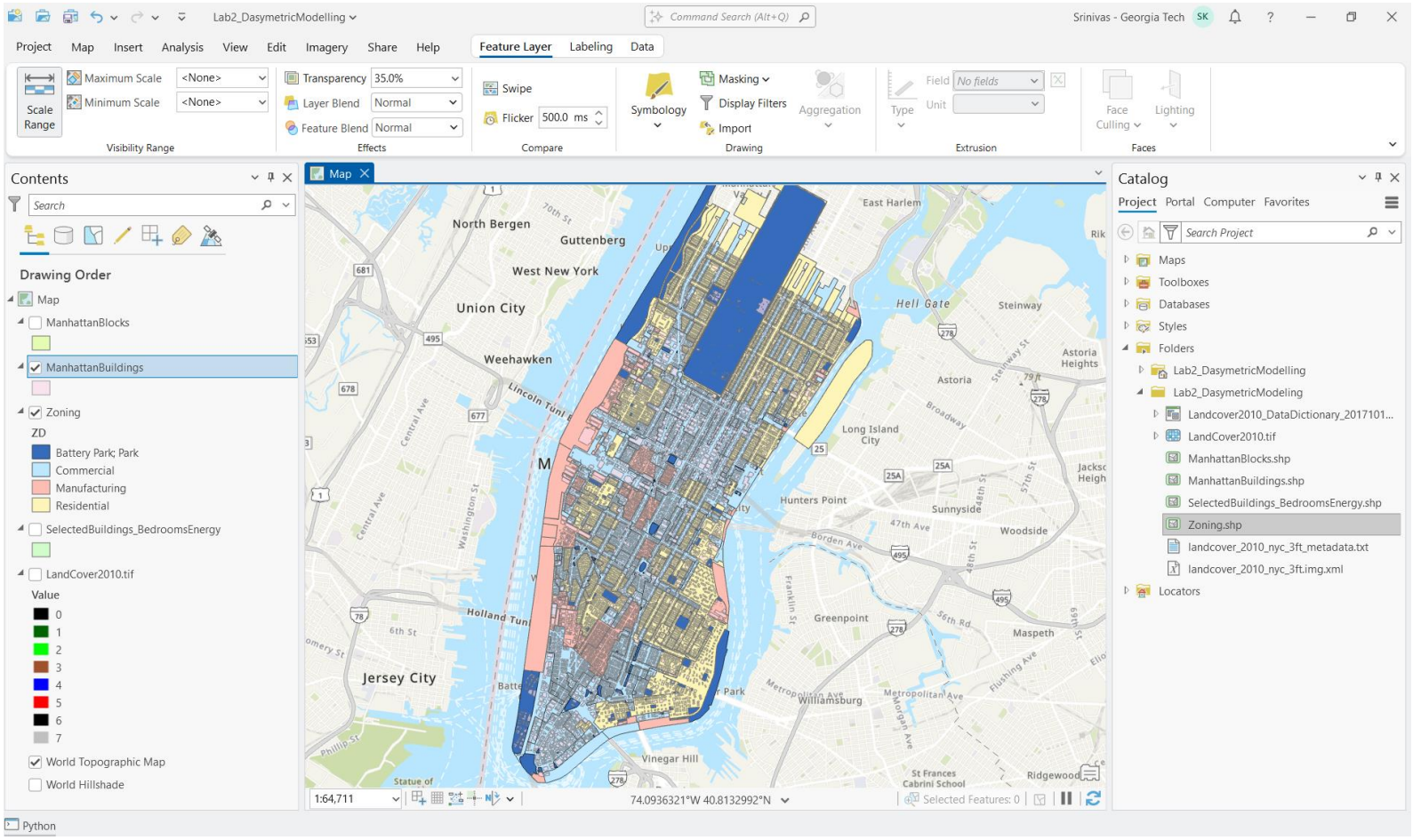
Lab Report

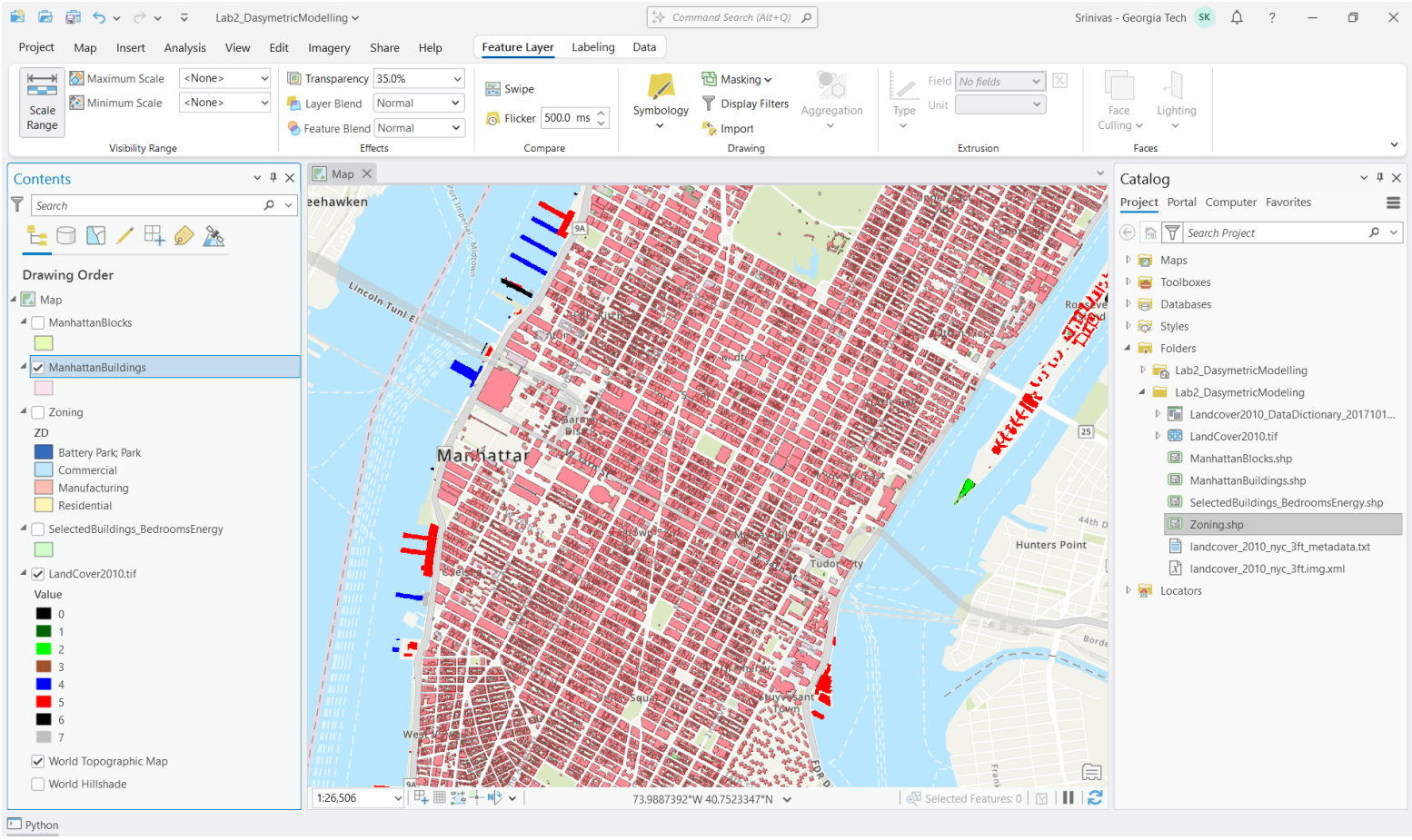
Srinivas Konreddy

01/23/2024

1a: Which dataset better aligns with the building footprints? (Have your answer be evident in the screencaps). (2 pts)

The Zoning data aligns better with the Manhattan Building outline shapefile than the LandCover dataset.





1b: Why might this not be a surprising answer? (Think of how the data was made). (3 pts)

This is not a surprising answer as the zoning data set offers a more comprehensive classification of the building outlines with respect to the background of the city. It clearly differentiates between the types of buildings in the study.

2: Given the zoning data and the population counts, do you think it is a good idea to remove buildings that are not zoned in residential areas for our analysis, why or why not? (5 pts)

In general, it is a good idea to remove buildings that are zoned in residential areas for our analysis as they do not contribute to the population density or the other parameters that we are focused on. In addition, for large file sizes it increases the processing time and the margin of error in calculation which may lower the precision of the results that are computed. However, there may be instances where buildings are incorrectly zoned or are multi-use buildings which still house residents. Depending on the requirements and the objective, this decision can be made.

3a: We just selected a method called ‘had their center in’. What is a polygon center called? (2 pts)
Polygon centers are called centroids

3b. How are these computed? (3 pts)
It is calculated as the center of the rectangular bounding box that encapsulates the polygon.

3c. Name three other encapsulating boundaries you can use to bound a spatial feature. (6 pts)
Spatial envelope (smallest rectangle that encapsulates a polygon), smallest encapsulating circle and a convex hull

3d: Are all centers INSIDE the buildings? Why or why not? (3 pts)
The centers can lie outside the buildings if the rectangular bounding box that surrounds the polygon has their center outside the polygon, but in this case we have selected to have their centers in while joining.

4a: Block ID is a FIPS code. What does FIPS stand for? (2 pts)
Federal Information Processing Standards

4b: Parse this block ID into spatial components including state, county, tract, block group and block (sometimes extra 0's are added): 130440025007003. (3 pts)
State:13
County (FIPS Code):044
Tract:002500
Block Group:7
Block:003

5: What is the before people per square mile and what is the after people per square mile for block: 360610063005000 (Census pop: 589)? (4 pts)
Old Density: 92292.34755532074
New Density: 193256.85141222988

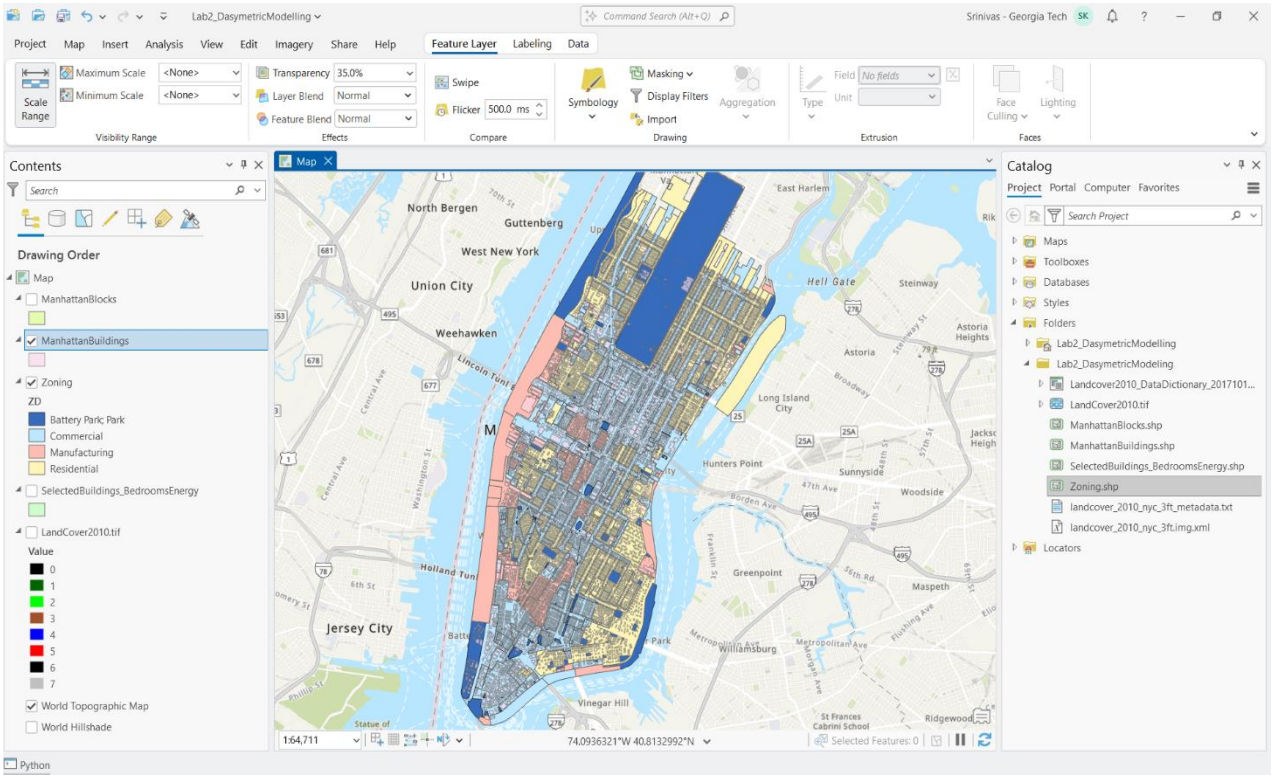
56: What is the total population of the columns BldgPop_Area and BldgPop_Volume? What was the original total population of the area? Have you preserved the number of people? What's the numeric difference? (5 pts)

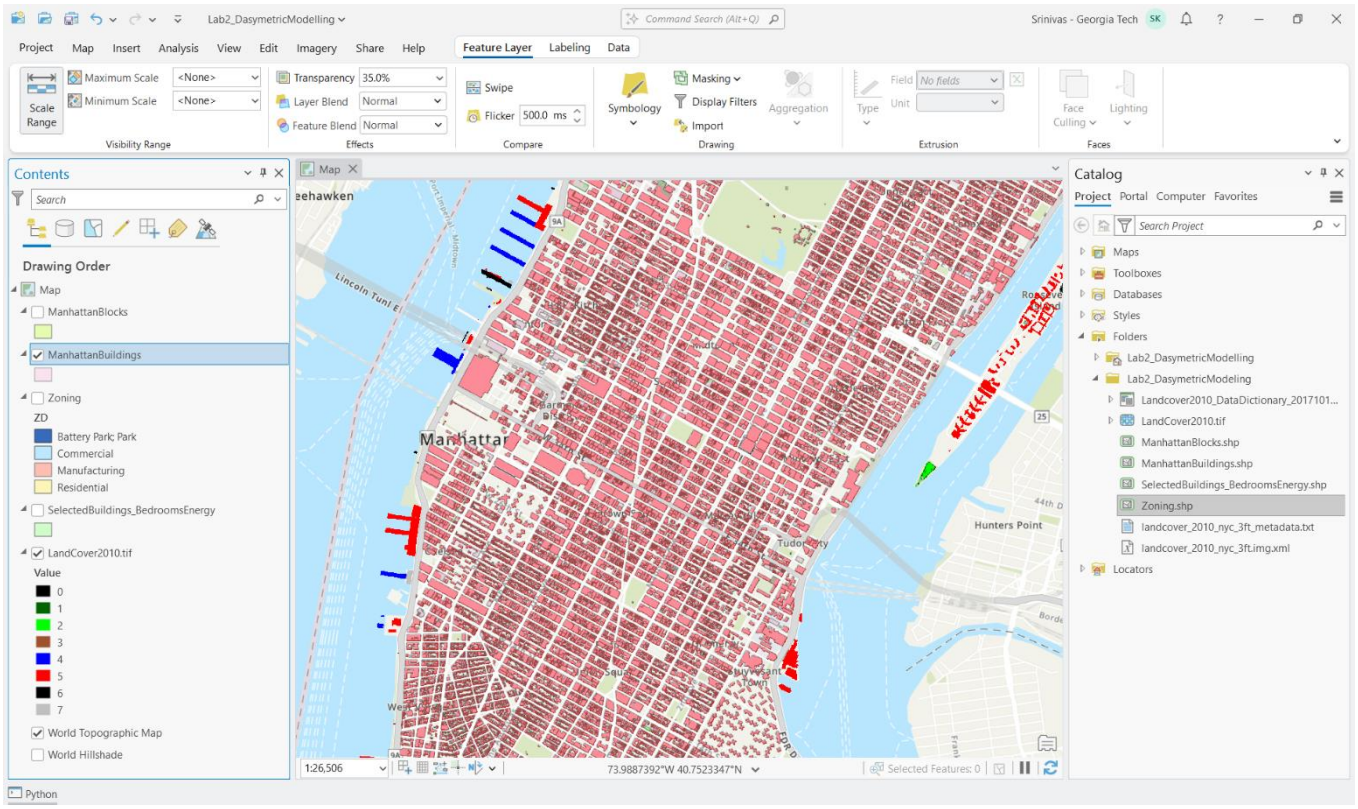
The total population of the columns are as follows:
BldgPop_Area: 880269
BldgPop_Volume: 879092

The original population of the area is: 883867

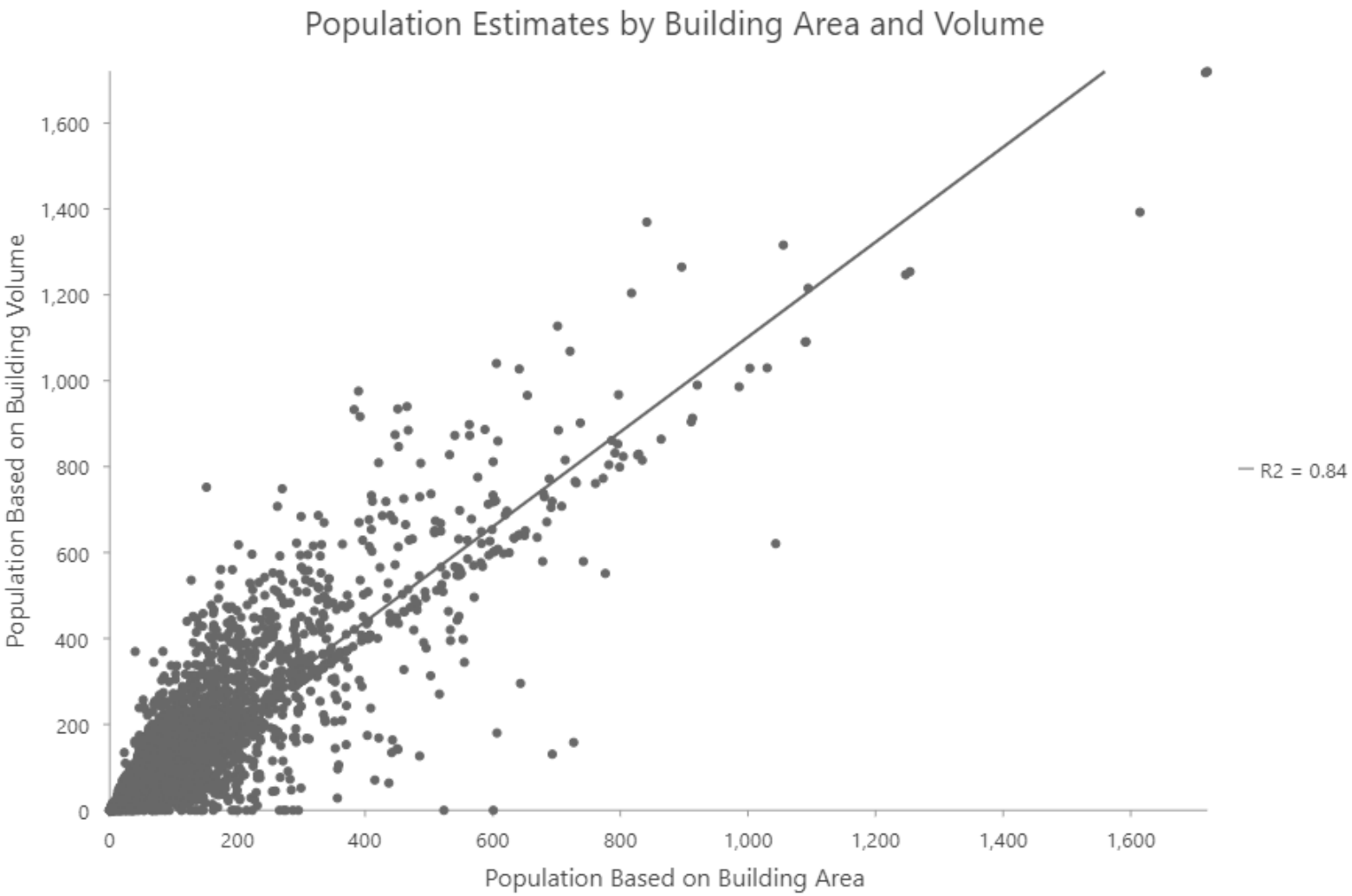
In this case the population has not been preserved, the numeric difference with respect to the estimates based on area is 3607 while the numeric difference with respect to the estimates based on volume is 4784.

Screen Cap 1: Take two screencaps here, one with the building footprint outlines on top of the zoning data, and one with the building footprint outlines on top of the land cover data. (It can be zoomed in).



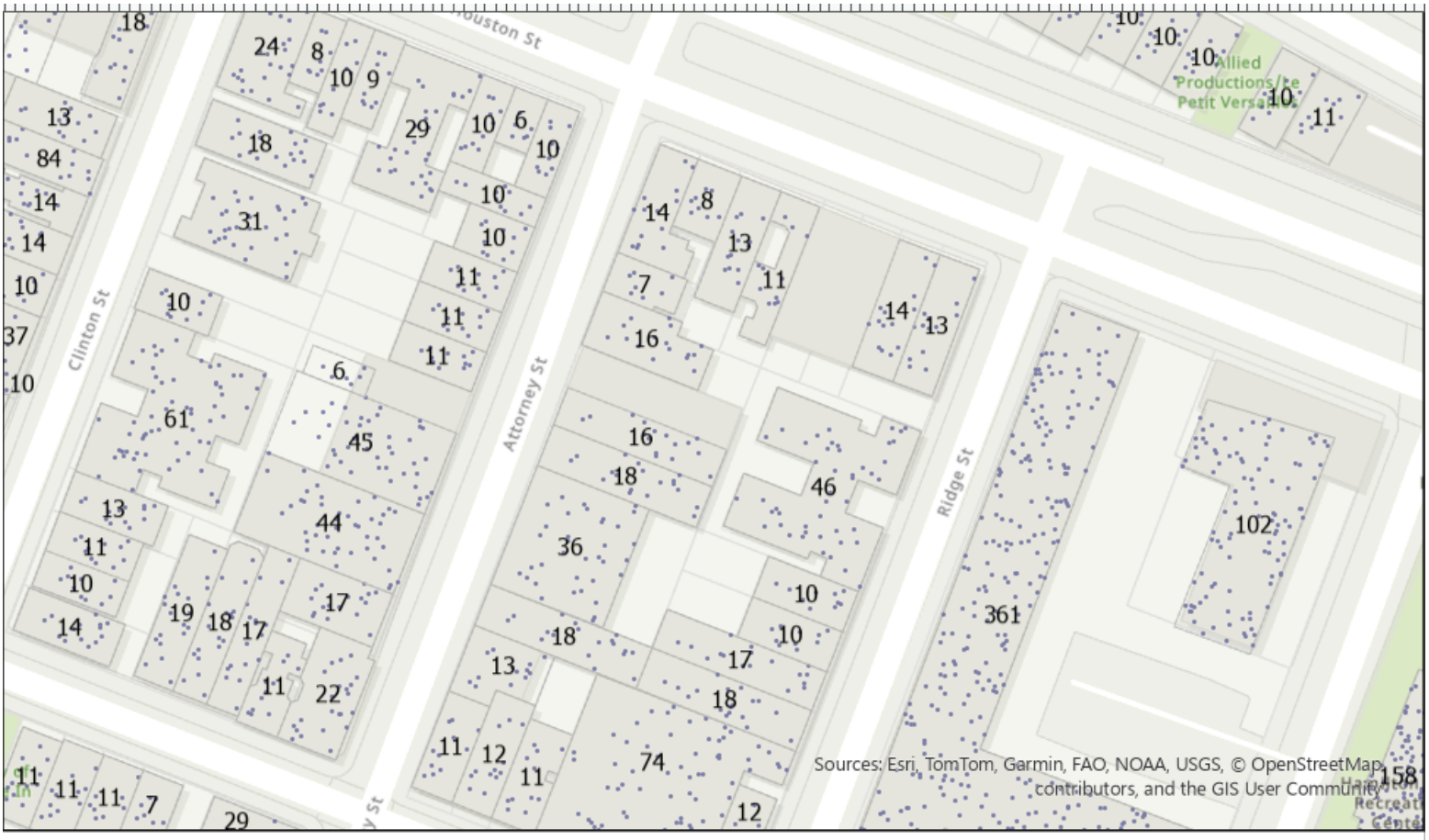


15 pts Graphic 1: Make a scatterplot comparing the BldgPop_Area (X axis) to BldgPop_Volume (Y axis). Make it nice by giving it axis labels on the x and y axis as well as a title. Remember that you are looking at number of people assigned to a building based on area vs volume. You can create the scatterplot anywhere you'd like. Make sure all the text is readable and of professional quality. Give it a caption that describes the trend! In the caption, say what each point represents (e.g. does it represent a particular unit of analysis?). *Give it a caption that describes the trend! For full points, In the caption, say what each point represents.*



The scatter plot illustrates the relationship between the population estimates based on building area and building volume. Each building is represented as a point and a linear trend line is included to show the co-relation between the two variables. With a high R^2 value of 0.84, this suggests that population estimates based on area and volume have high co-relation, with larger areas and volumes having higher population estimates.

15 pts Graphic 2: Make a nice map of BldgPop_Area OR BldgPop_Volume (hopefully we have some of both examples!). Use dot density as symbology. Zoom in on an area you like and label the polygons with the number of people in them. *Give it a caption and a legend.* Make sure your labels are in "people" numbers--integers, i.e., there should not be ½ a person.



Legend

1 Dot = 1 Person

- BldgPop_Area

By using BldgPop_Area as my classification feature for the map, I have visualized the distribution of the population density with respect the building area in Manhattan. Each dot represents a single person, with each polygon representing the total number of people living in the building rounded off to the nearest whole number.

10 pts Table 1: Use the buildings population to fill out the table below (Hint: this is a special spatial join and then a dissolve, like you did above.). USE THE FIELD ZD.

POPULATION		ZONING (field ZD)
BldgPop	BldgPopSec	
481414	54.6%	R Residential Districts
368014	41.8%	C Commercial Districts
19769	2.24%	M Manufacturing Districts
9705	1.10%	B Battery Park City
1367	0.15%	P Areas designated as PARK, BALL FIELD, PLAYGROUND and PUBLIC SPACES in NYC GIS Zoning Features.