A photograph of a park path with trees and people, overlaid with a semi-transparent grey box containing text. The path is paved and lined with a black metal fence. There are several people walking on the path, including a person pushing a stroller. The trees have green and yellow leaves, suggesting autumn. In the background, there are brick buildings and a sign that says "PARK".

Mapping the Relationship Between Vegetation and Poverty & Race in Philadelphia

Qiu, Xinyi | CPLN 503 | Project 1 Data Preparation

CONTENTS

- INTRODUCTION**.....3
 - Background & Objectives.....4
 - Project Flow.....5

- PART I Data Preparation**.....6
 - Datasets Description.....7
 - Vegetation Visualization.....8
 - Population and Poverty Visualization.....9

- PART II Data Interpretation**.....10
 - Evaluating the Vegetation in Philadelphia.....11
 - Evaluate the score of average tree height.....11
 - Evaluate the score of tree canopy coverage.....14
 - Evaluate the accessibility of parks.....16
 - Sum the total score of the vegetation.....20
 - Explore the Relationship between Vegetation and Poverty or Race.....22
 - Examine the relationship between vegetation and poverty.....22
 - Examine the relationship between vegetation and race.....23

- PART III DATA AUTOMATION**.....24
 - Overall Model to Calculate the Score of Vegetation.....25
 - Park Accessibility Model.....26
 - Tree Height Model.....27
 - Tree Canopy Area Model.....28
 - Submodel of Vegetation.....29

- Conclusion**.....30

A photograph of a park path in autumn. The path is paved and covered with fallen yellow and orange leaves. On either side of the path are black metal fences with decorative scrollwork. In the background, there are large trees with green and yellowing leaves, and a multi-story brick building with white window frames. A few people are walking on the path, including a person pushing a stroller. The overall scene is bright and pleasant.

PROJECT INTRODUCTION

Project Background and Objectives

Background

Abundant research proved that access to greenspace varies highly based on income, level of education, gender, race, age and other socioeconomic and personal characteristics differences. A common finding is that income and level of education are positively correlated with a greater accessibility to greenspace, while low-income, racial minorities and other vulnerable groups in the city have less access to vegetation.

According to the reports, this situation is happening in Philadelphia, especially the lack of trees in poor neighborhood. We need to ensure if it is true and where exactly demand for more trees.

Objectives

- I. Identify the characteristics of poverty status in Philadelphia, and stress the importance on race differences;
- II. Identify the characteristics of tree canopy in Philadelphia, including the area and height;
- III. Explore the relationship between tree canopy and poverty in Philadelphia

A photograph of a park path with trees and a brick building in the background. The path is paved and covered with fallen leaves. There are several people walking on the path, including a person pushing a stroller. The trees are mostly green, with some showing yellow and orange autumn colors. A brick building is visible in the background. The text "PART I DATA PREPARATION" is overlaid on the image in a large, bold, black font.

PART I DATA PREPARATION

Project Data

Data Catalogue	Data Items	Data Source
Socioeconomic Data	Poverty Population	https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml
	White Poverty Population	
	Black Poverty Population	
	Black Population	
	White Population	
Vegetation Data	Tree Average Height	https://www.opendataphilly.org/dataset/ppr-tree-canopy
	Tree Canopy Coverage	
	Philadelphia PPR Park	https://www.opendataphilly.org/dataset/ppr-properties
Base Map	Philadelphia Census Tract Map	https://www.opendataphilly.org/dataset/census-tracts

Part I Data Preparation-Vegetation

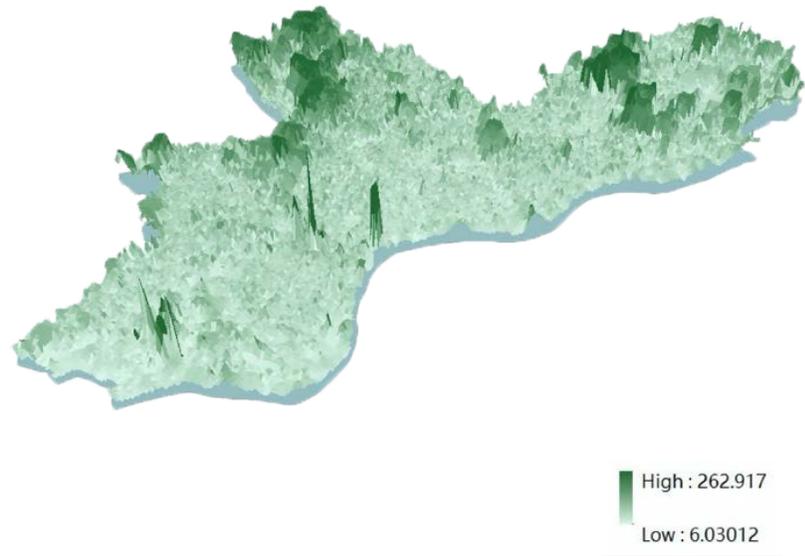


Figure 1.1 Average Tree Height in Philadelphia

Figure 1.2 Percentage of Tree Canopy Area by Census Tracts in Philadelphia

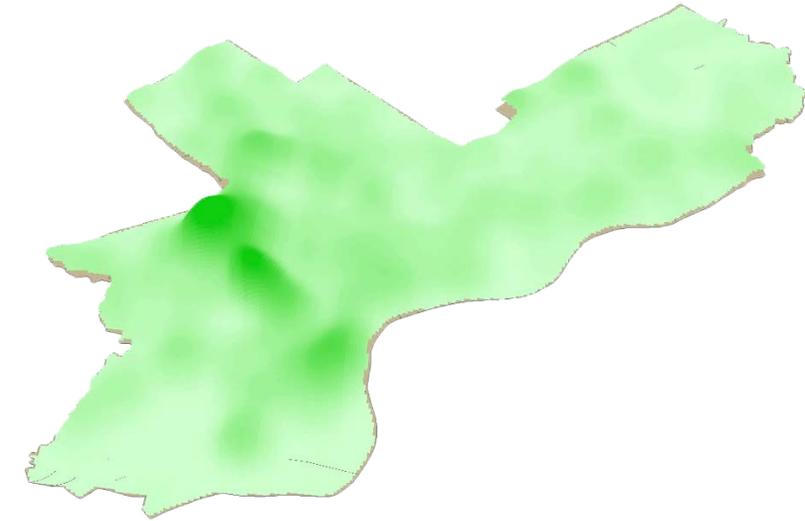
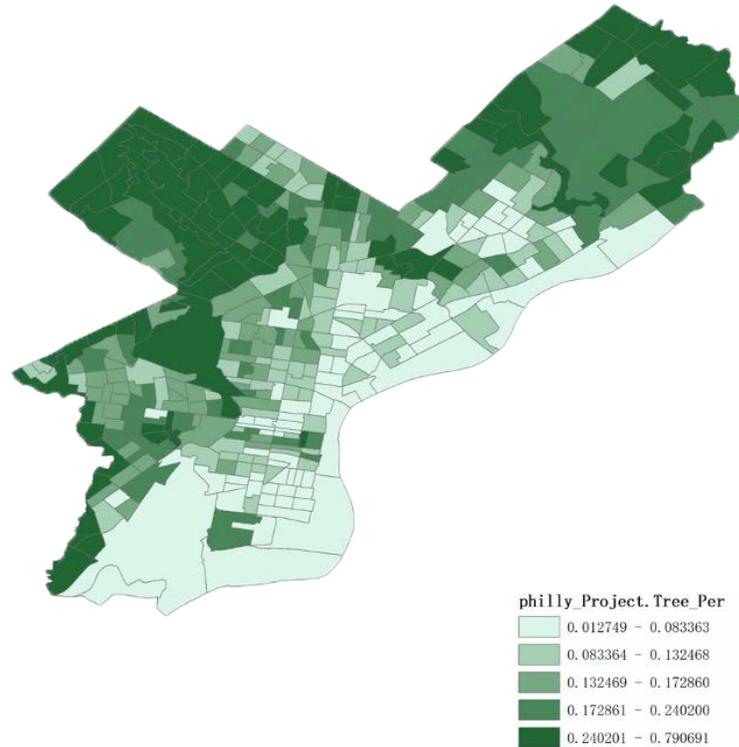


Figure 1.3 PPR Park Location in Philadelphia

Part I Data Preparation-Population & Poverty

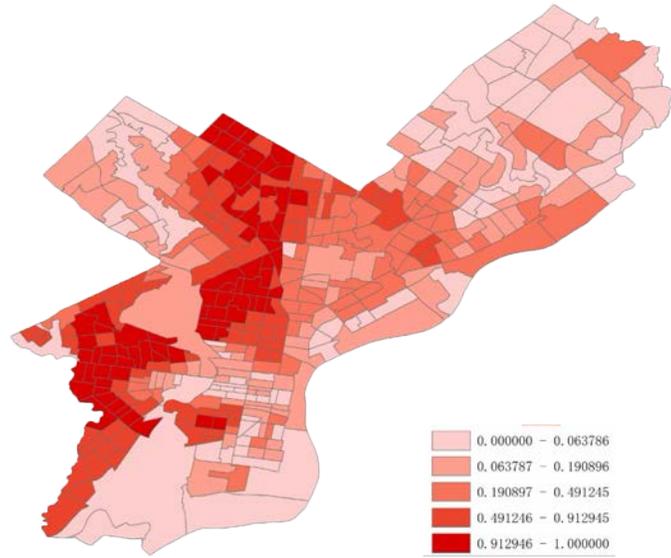


Figure 1.4 Percentage of Black Population by Census Tracts

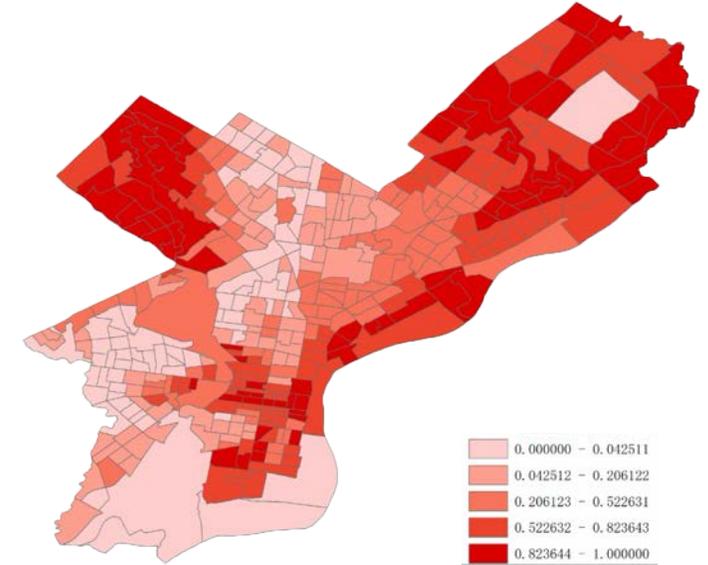


Figure 1.6 Percentage of White Population by Census Tracts

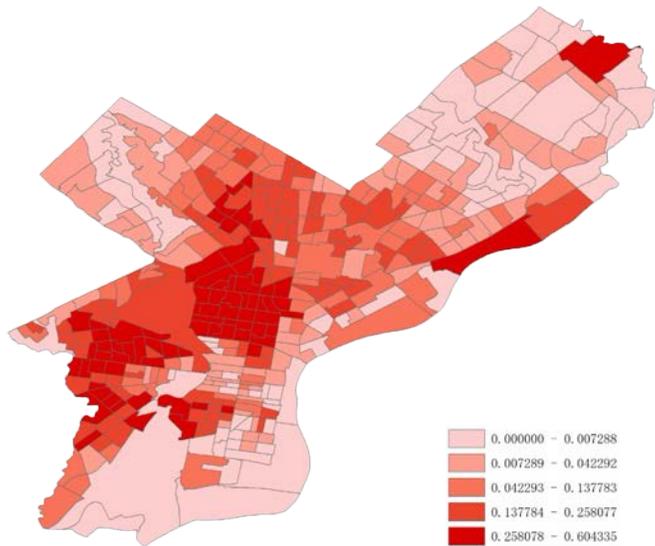


Figure 1.5 Black Population Poverty Rate by Census Tracts

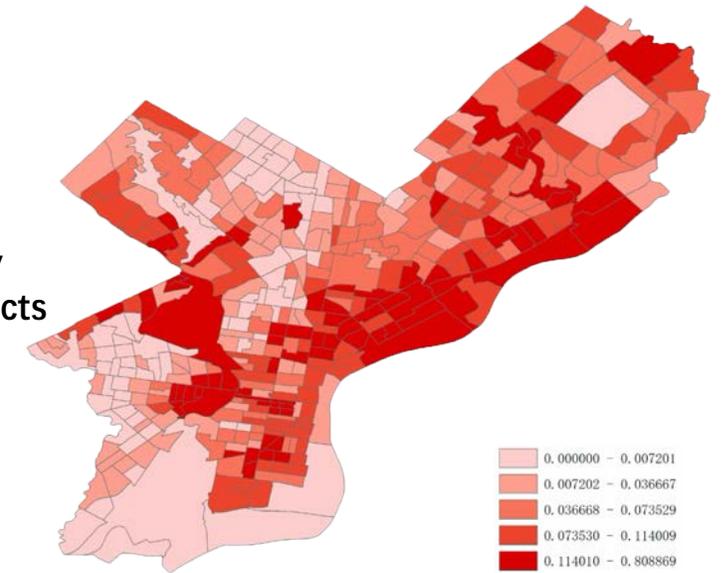
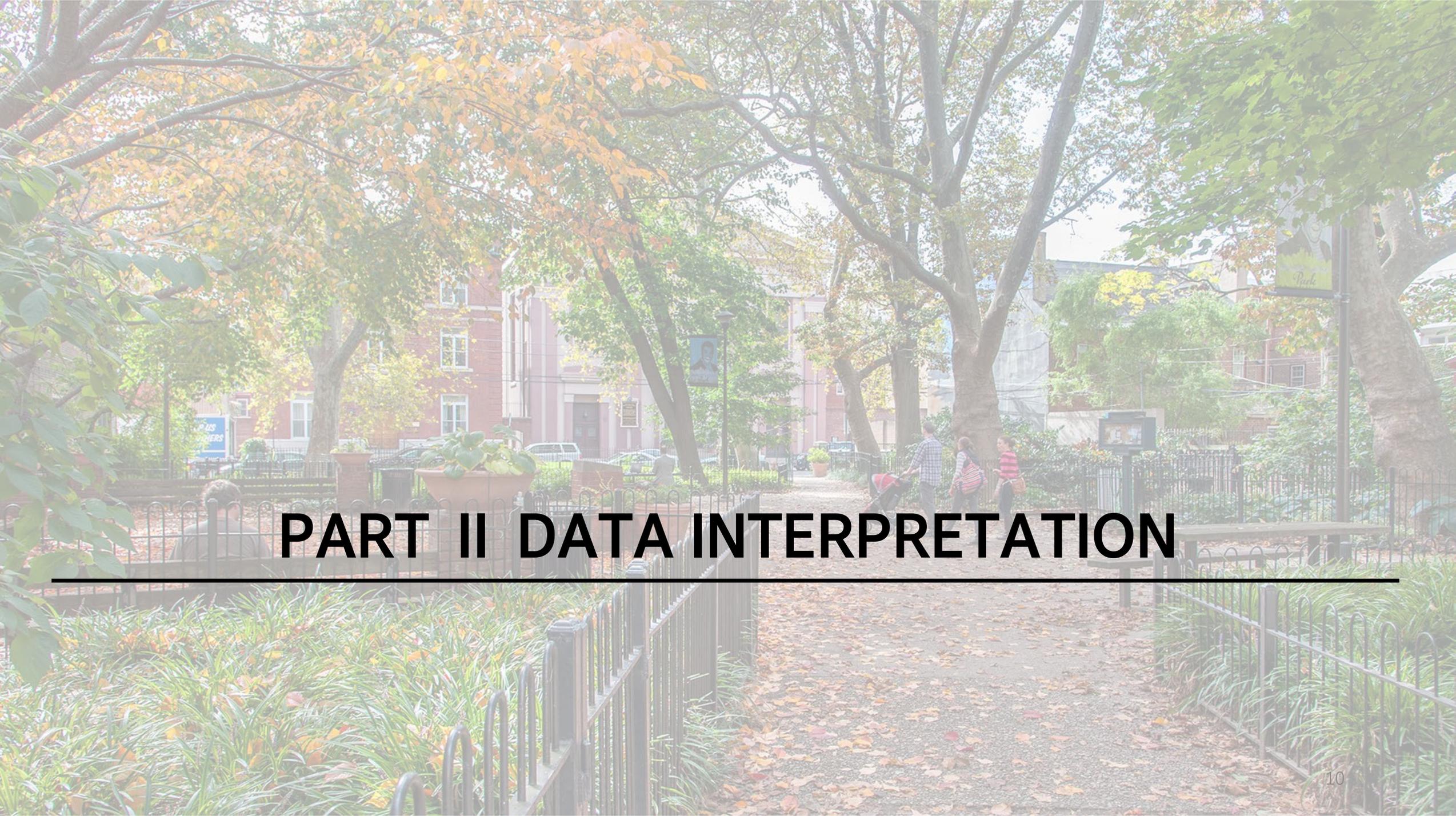


Figure 1.7 White Population Poverty Rate by Census Tracts

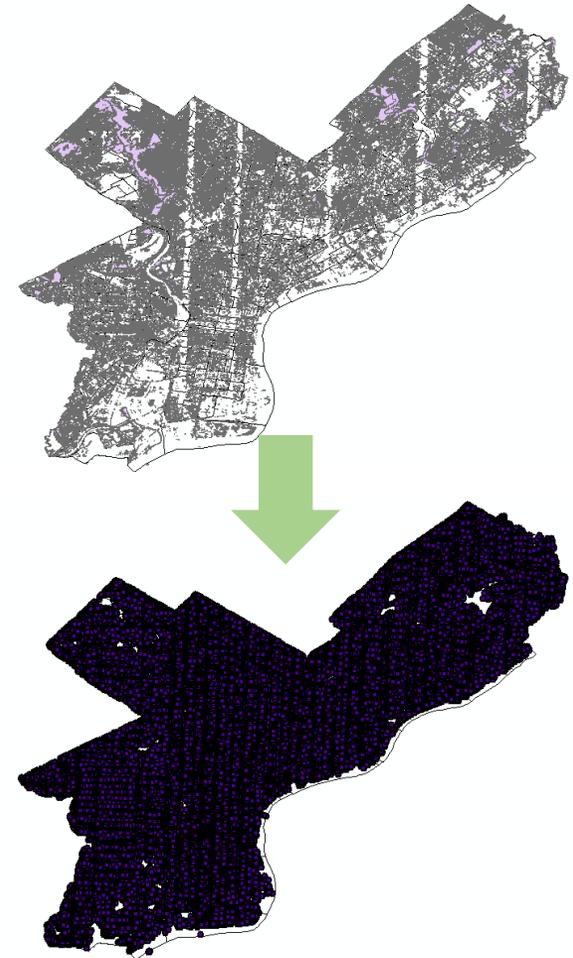
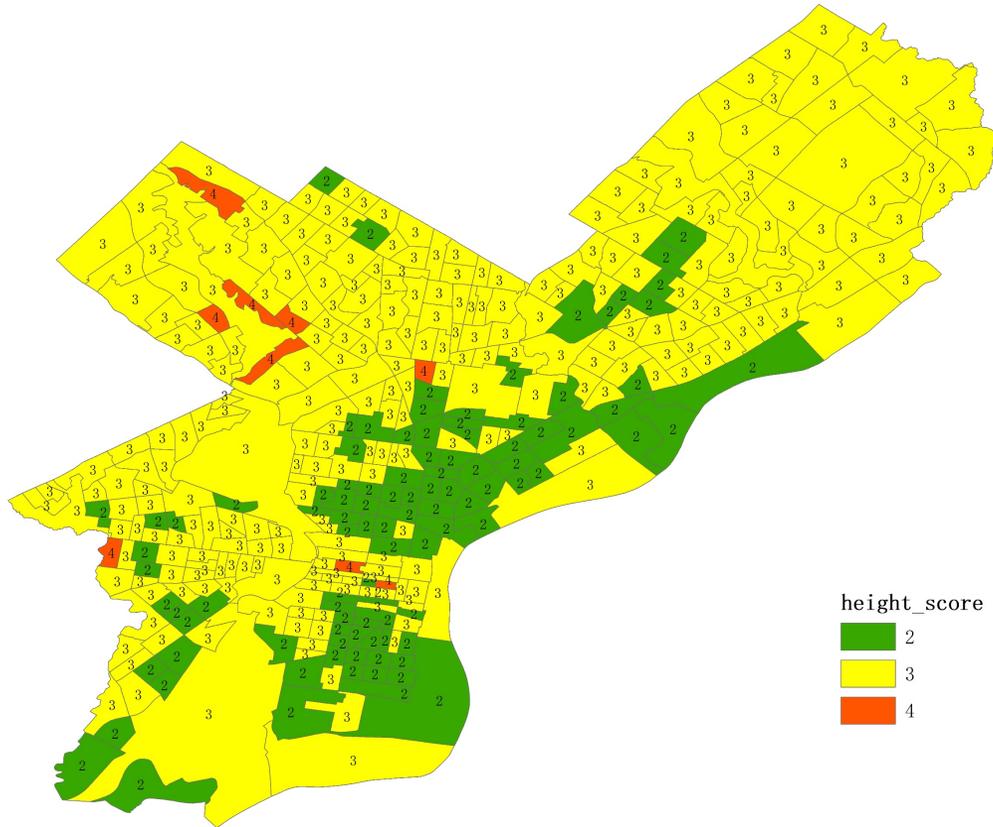


PART II DATA INTERPRETATION

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

1.-Evaluating tree height

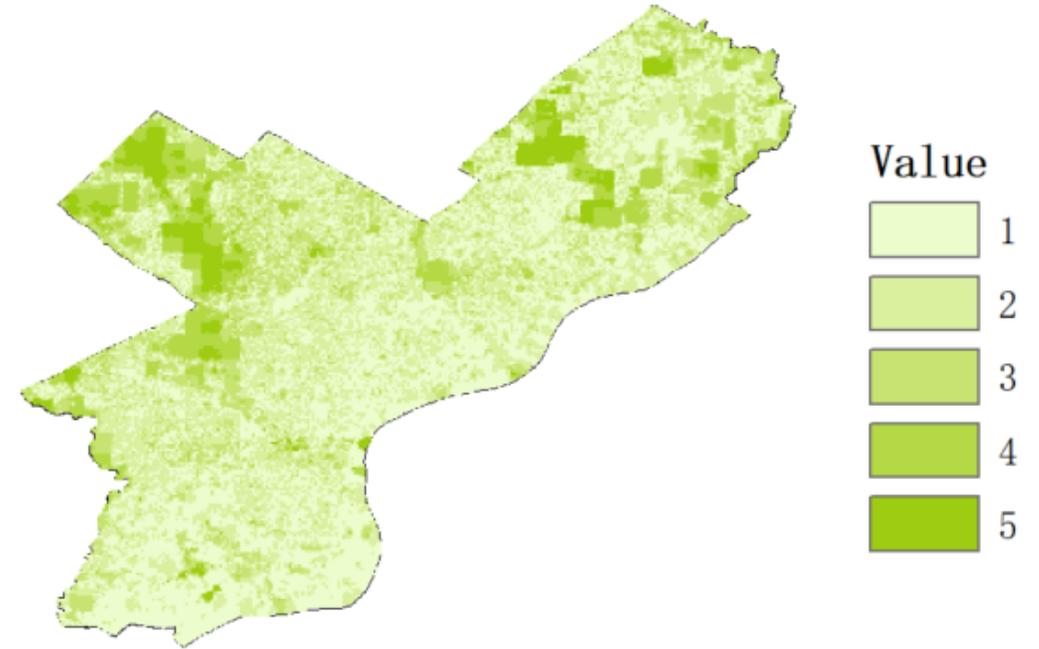
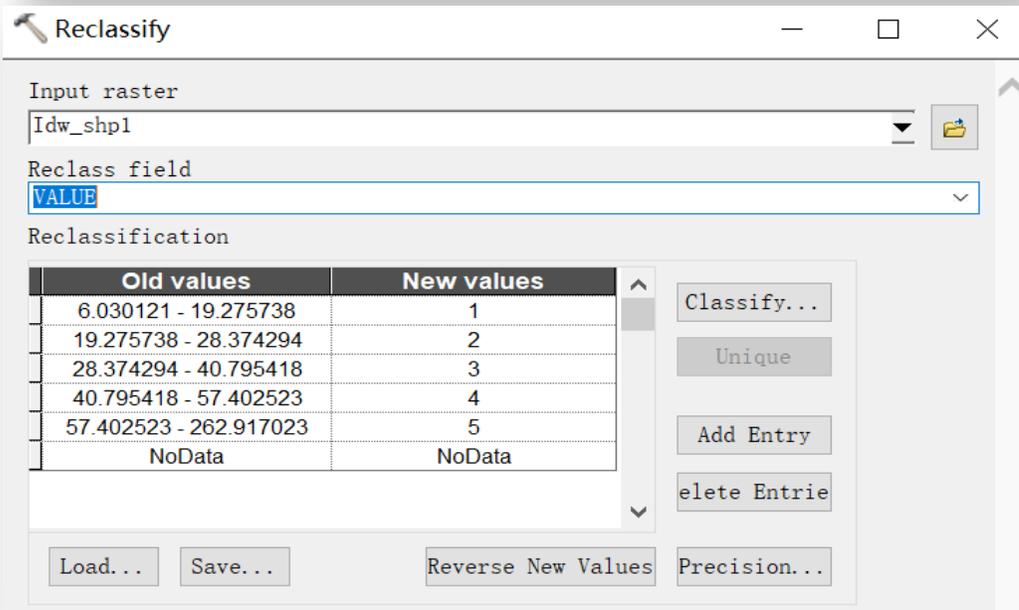
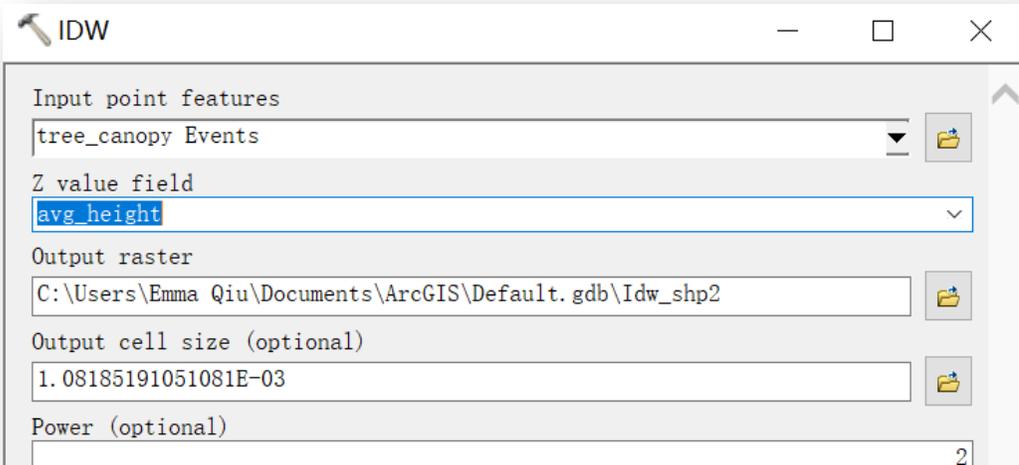
Tree height score map



In data preparation, centroid points of each tree canopy has been displayed

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

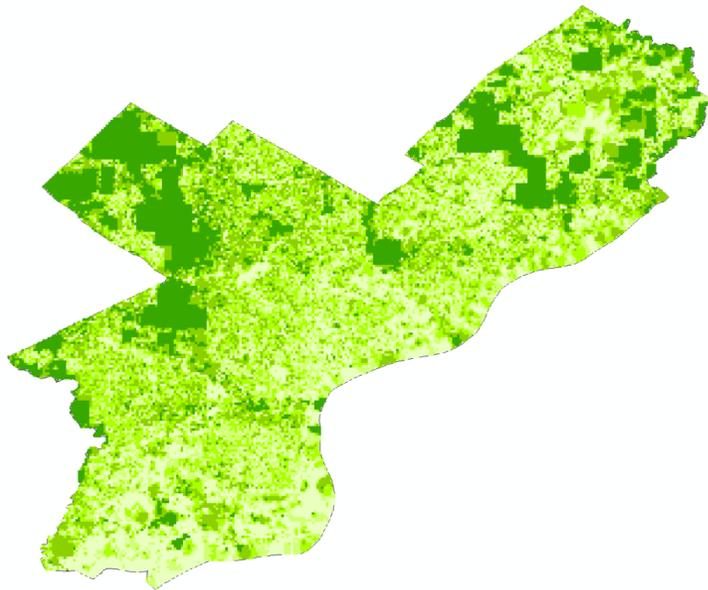
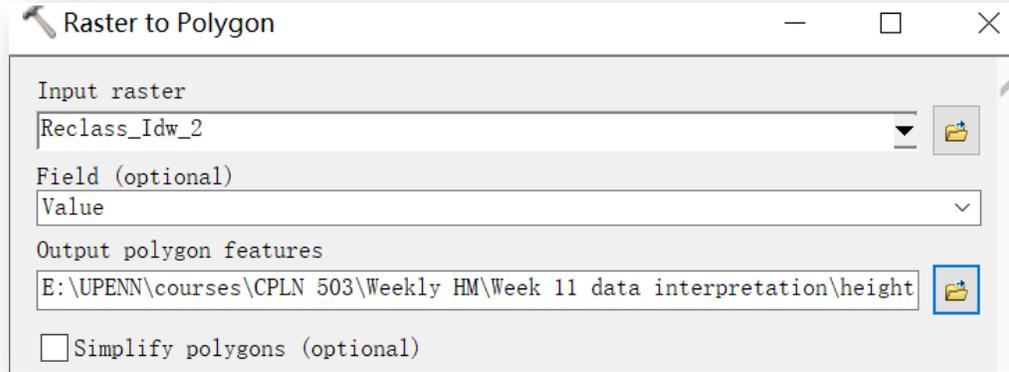
1.-Evaluating tree height



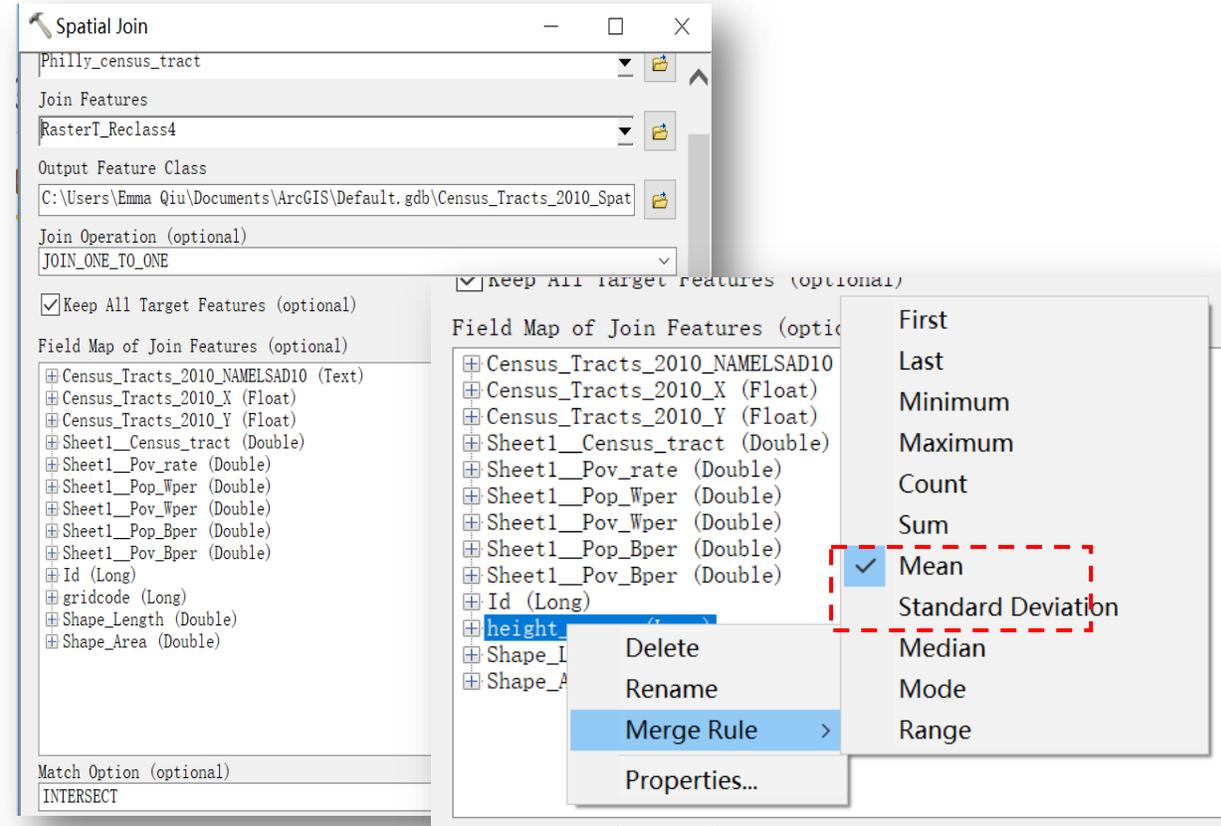
Since the centroids of tree canopy has been prepared in the data preparation report, use **Spatial analyst tools-interpolation-IDW** to interpolate the tree height within the city;
And then use **Spatial Analyst Tools-Reclass-Reclassify** to attribute the score to different classes by “Quantile” classification method

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

1.-Evaluating tree height



Use **Conversion Tool-From Raster-Raster to Polygon** to transform the IDW raster into polygon for further data interpretation

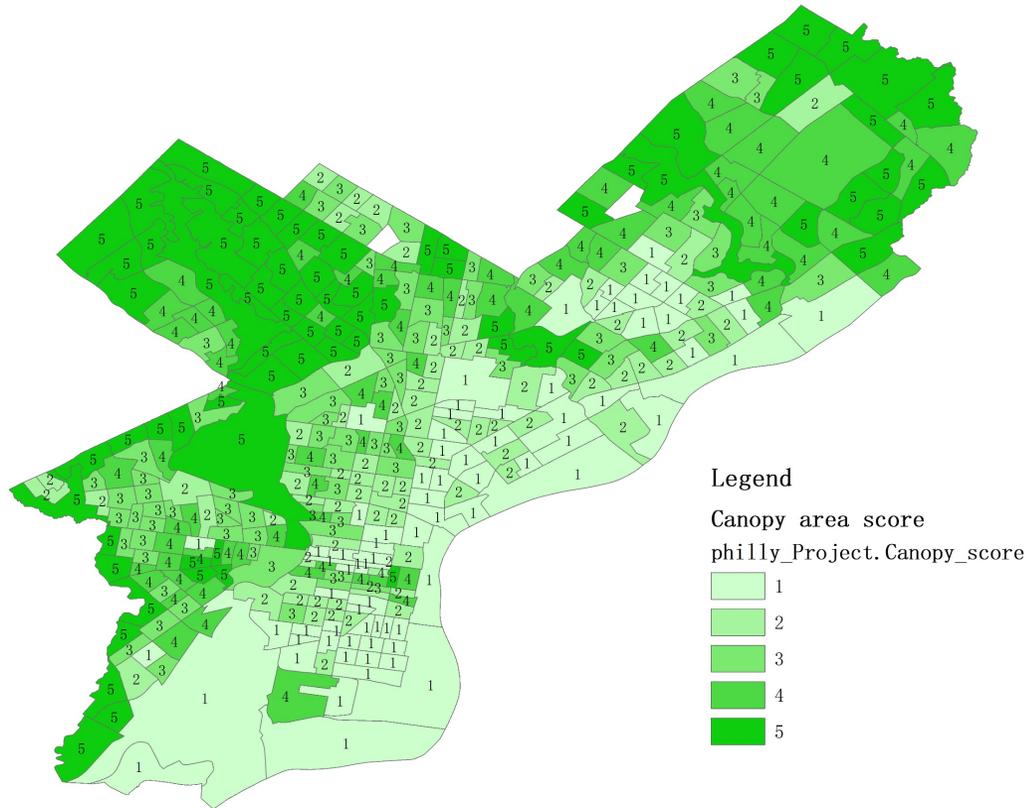


Use **Analysis Tools-Overlay-Spatial join** to join the reclassified interpolated tree height layer to the census tract layer. It should be noted that the merge rule must be Mean when doing the spatial join, so that we can get the average tree height score of each census tracts.

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

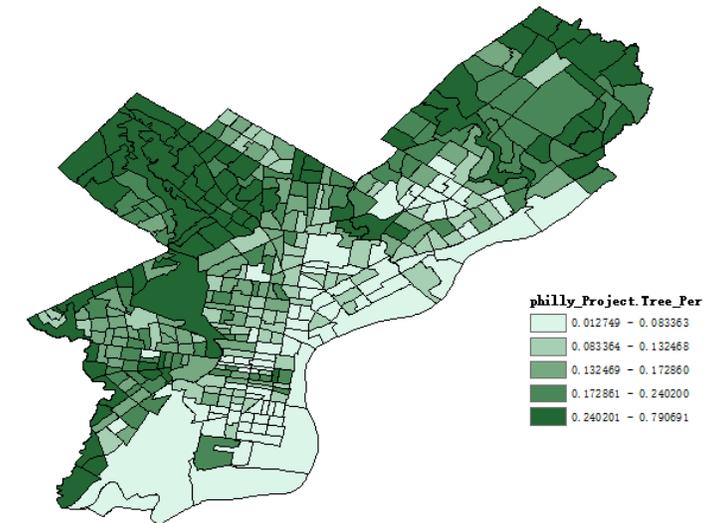
2.-Evaluating tree canopy coverage

Tree Canopy Covergae score map



In the data preparation report, I used Identity, Dissolve, Project, Calculate Geometry and Field Calculator tools to obtain the tree canopy coverage rate map.

What to do next is to assign the score to different Quantile intervals.



Part II Data Interpretation-Evaluating the vegetation in Philadelphia

2.-Evaluating tree canopy coverage

Select by Attributes

Enter a WHERE clause to select records in the table window.

Method : Create a new selection

philly_Project.Tree_Per

philly_Project.Canopy_score

philly_Project.Canopy_score =

1

2

3

4

5

```
SELECT * FROM philly_Project_Dissolve_Clip_identity_tree_ WHERE:  
philly_Project.Tree_Per >0.012749 AND philly_Project.Tree_Per <0.083363  
philly_Project.Tree_Per >0.083364 AND philly_Project.Tree_Per <0.132468  
philly_Project.Tree_Per >0.132469 AND philly_Project.Tree_Per <0.172860  
philly_Project.Tree_Per >0.172861 AND philly_Project.Tree_Per <0.240200  
philly_Project.Tree_Per >0.240201 AND philly_Project.Tree_Per <0.790691
```

Field Calculator

Parser
 VB Script Python

Fields:
philly_Project.Shape_Length
philly_Project.Shape_Area
philly_Project.Tree_Per
philly_Project.Canopy_score
Dissolve_Clip_identity_tree_.OBJECTID
Dissolve_Clip_identity_tree_.Census_
Dissolve_Clip_identity_tree_.Tree_Ar
Dissolve_Clip_identity_tree_.Shape_A

Type:
 Number
 String
 Date

Functions:
Abs ()
Atn ()
Cos ()
Exp ()
Fix ()
Int ()
Log ()
Sin ()
Sqr ()
Tan ()

Show Codeblock

philly_Project.Canopy_score =

1

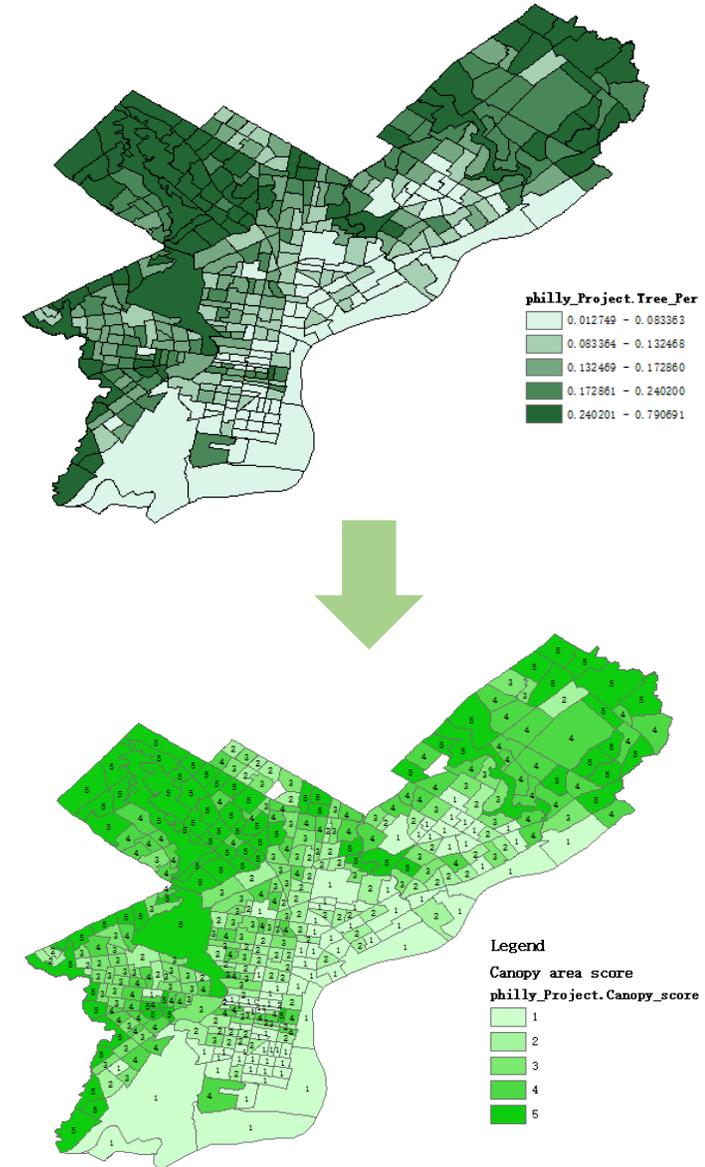
2

3

4

5

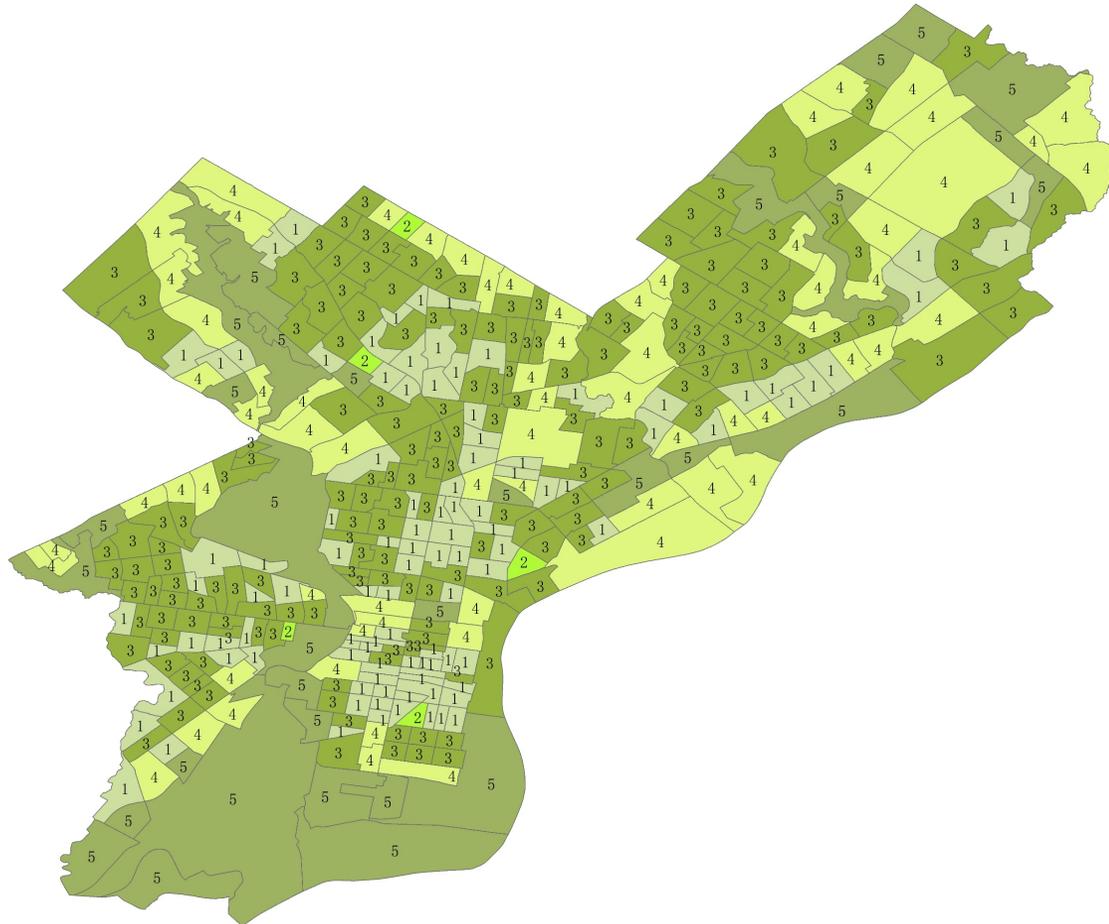
Use **Select by Attributes** to select census tracts in different tree canopy coverage intervals one by one and use **Field Calculator** to assign the score to the selected census tract from 1 to 5.



Part II Data Interpretation-Evaluating the vegetation in Philadelphia

3.-Evaluate the accessibility of parks in Philadelphia

Parks' accessibility score map

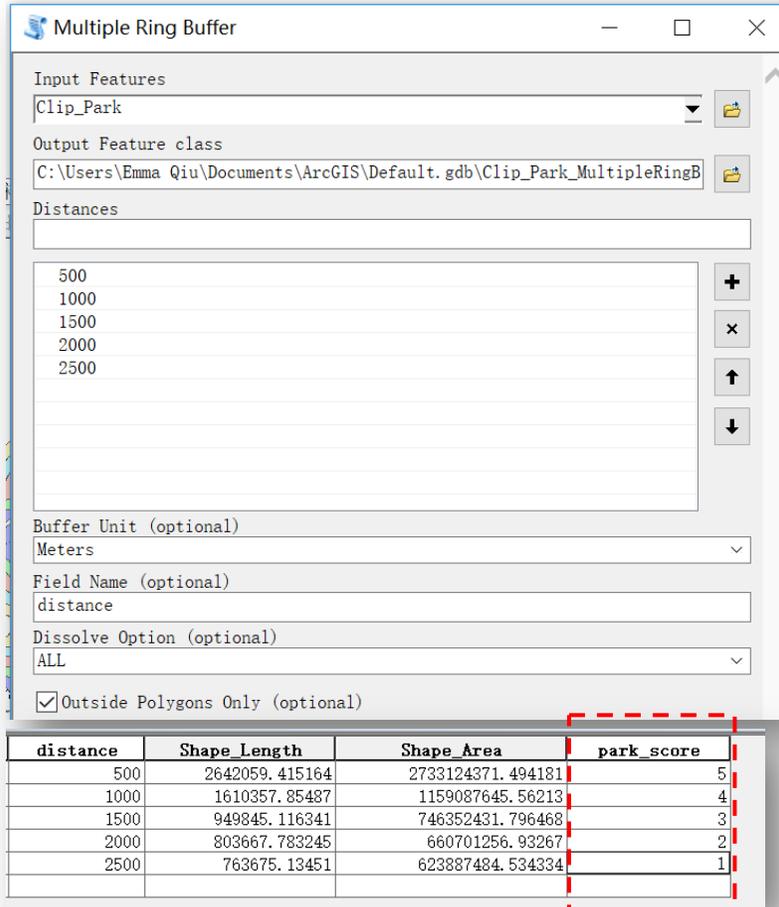


Assume that the less distance away from the park, the better living experience the residents will have. So evaluate the accessibility of parks by the distance from the boundary of the parks in Philadelphia

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

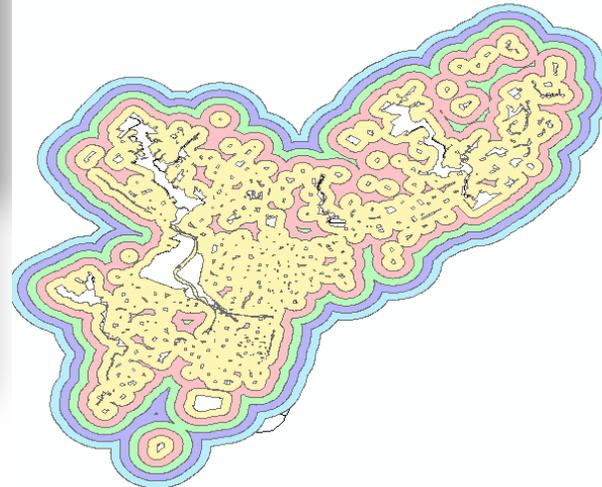
3.-Evaluate the accessibility of parks in Philadelphia

Step 1- BUFFER

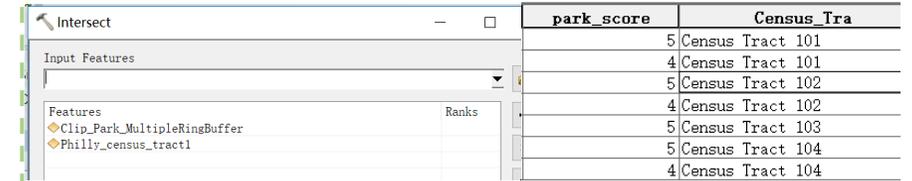


distance	Shape_Length	Shape_Area	park_score
500	2642059.415164	2733124371.494181	5
1000	1610357.85487	1159087645.56213	4
1500	949845.116341	746352431.796468	3
2000	803667.783245	660701256.93267	2
2500	763675.13451	623887484.534334	1

Use **Analysis Tools-Proximity-Multiple Ring Buffer** to find the areas which are 500 meters, 1000 meters, 1500 meters, 2000 meters and 2500 meters away from the park. The Multiple Ring Buffer helps to dissolve the circles by distance intervals instead of using dissolve and erase tool later when using buffer tool.

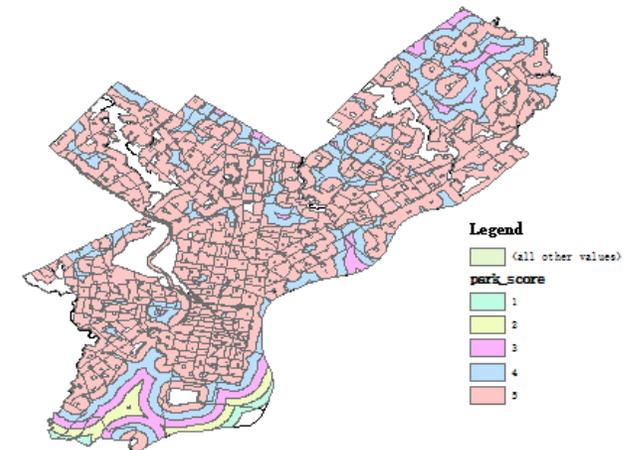


Step 2- INTERSECT



park_score	Census_Tra
5	Census Tract 101
4	Census Tract 101
5	Census Tract 102
4	Census Tract 102
5	Census Tract 103
5	Census Tract 104
4	Census Tract 104

Use **Analysis Tools-Overlay-Intersect**, we can intersect the buffer layer with the base layer and obtain a layer with combined attributes. Besides, the buffer are divided by the census tract boundary, so that I can calculate the sum score of each census tract



Part II Data Interpretation-Evaluating the vegetation in Philadelphia

3.-Evaluate the accessibility of parks in Philadelphia

Step 3- Dissolve

Input Features: Intersect_bufFercensus

Output Feature Class: E:\UPENN\courses\CPLN 603\Weekly HM\Week 11 data interpretation\access

Dissolve Field(s) (optional):

- FID
- FID_Clip_P
- distance
- Shape_Leng
- Shape_Area
- park_score
- FID_Census
- Census_Tra
- Census_T_1
- Census_T_2

Statistics Field(s) (optional):

Field	Statistic Type
park_score	SUM

Statistics of accessibility score

Field: SUM_park_s

Statistics:

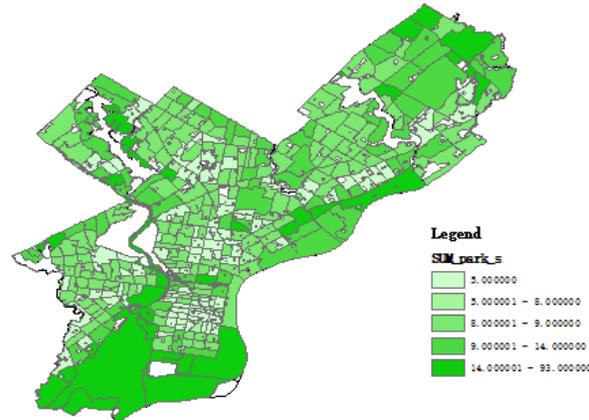
- Count: 384
- Minimum: 5
- Maximum: 93
- Sum: 3727
- Mean: 9.705729
- Standard Deviation: 6.388415
- Nulls: 0

Frequency Distribution

Score	Count
5	103
7	2
8	3
9	174
10	18
12	24
13	4
14	24
15	7
17	4
18	4
19	4
21	1
22	2
23	2
24	1
27	1
29	1
30	1
36	2
45	1
93	1

Use Data Management Tool-Generalization-Dissolve tool to dissolve the intersected layer by census tract and sum the park score by census tract.

Then use Quantile classification to define five level of parks' accessibility in Philadelphia



Step 4- Select by Attributes and assign the score

Select by Attributes

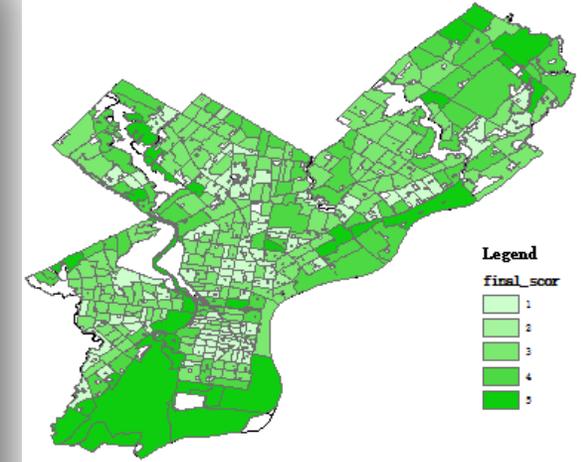
Enter a WHERE clause to select records in the table window.

Method: Create a new selection

Fields: "FID", "Census_Tra", "SUM_park_s"

Operators: =, <>, Like, >, >=, And, <, <=, Or, %, (), Not, Is, In, Null, Get Unique Values, Go To

SELECT * FROM accessibility score WHERE:
"SUM_park_s" = 5.000000



Use Select by Attributes and Field Calculator to assign the score from 1 to 5 to the sum scores' intervals.

Field Calculator

Parser: VB Script (selected), Python

Type: Number (selected), String, Date

Fields: FID, Shape, Census_Tra, SUM_park_s, final_score

Functions: Abs(), Atn(), Cos(), Exp(), Fix(), Int(), Log(), Sin(), Sqr(), Tan()

Show Codeblock

final_score =
1

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

3.-Evaluate the accessibility of parks in Philadelphia

Step 5- Join the table and Display

Join Data

Join lets you append additional data to this layer's attribute table so you can, for example, symbolize the layer's features using this data.

What do you want to join to this layer?

Join attributes from a table

1. Choose the field in this layer that the join will be based on:
NAMELSAD10
2. Choose the table to join to this layer, or load the table from disk:
accessibility score
 Show the attribute tables of layers in this list
3. Choose the field in the table to base the join on:
Census_Tra

Join Options

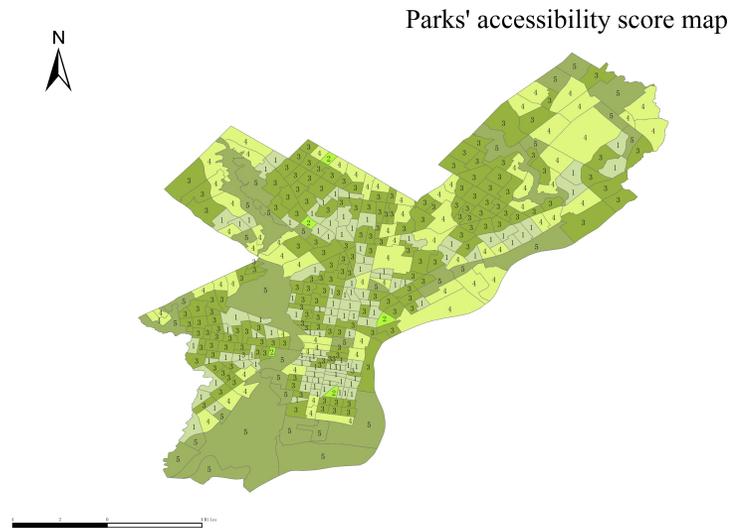
Keep all records
All records in the target table are shown in the resulting table. Unmatched records will contain null values for all fields being appended into the target table from the join table.

Keep only matching records
If a record in the target table doesn't have a match in the join table, that record is removed from the resulting target table.

Validate Join

[About joining data](#)

OK Cancel



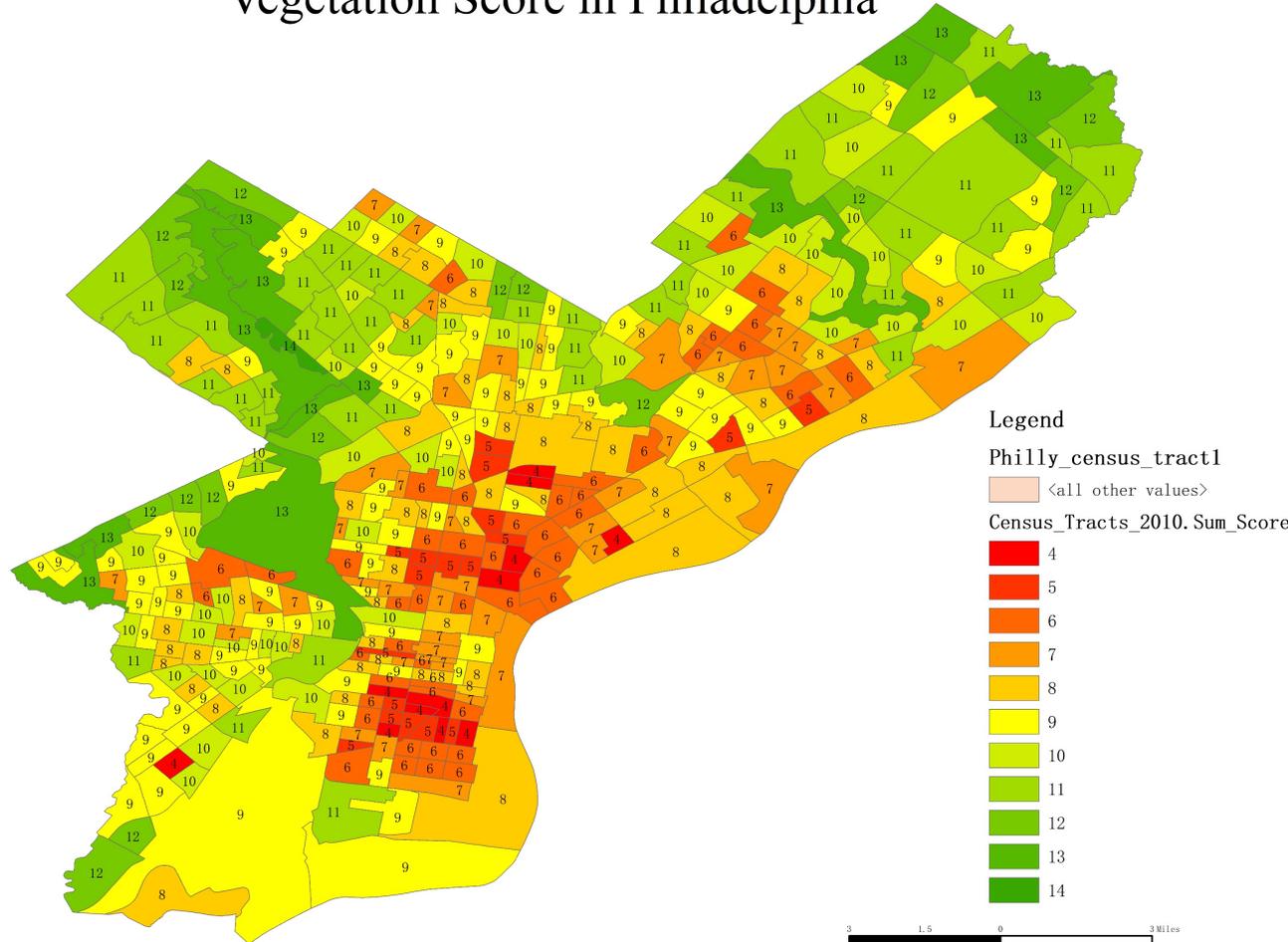
Join the table of the intersected layer with the base layer and display the distribution of parks' accessibility in the base map.

Part II Data Interpretation-Evaluating the vegetation in Philadelphia

4.-Sum the total score of the vegetation



Vegetation Score in Philadelphia

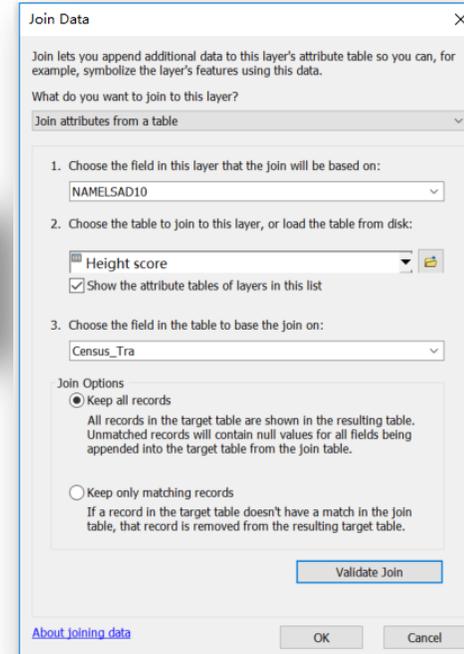


Vegetation Score= Tree Height Score+ Tree Canopy Area Score+ Park Accessibility Score

Step 1- Join the table

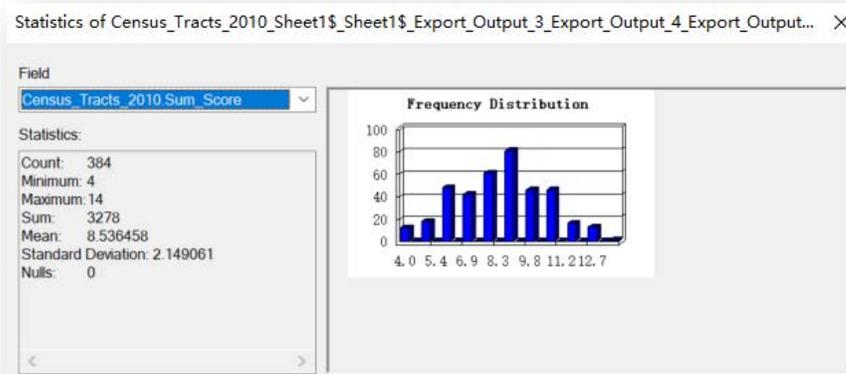
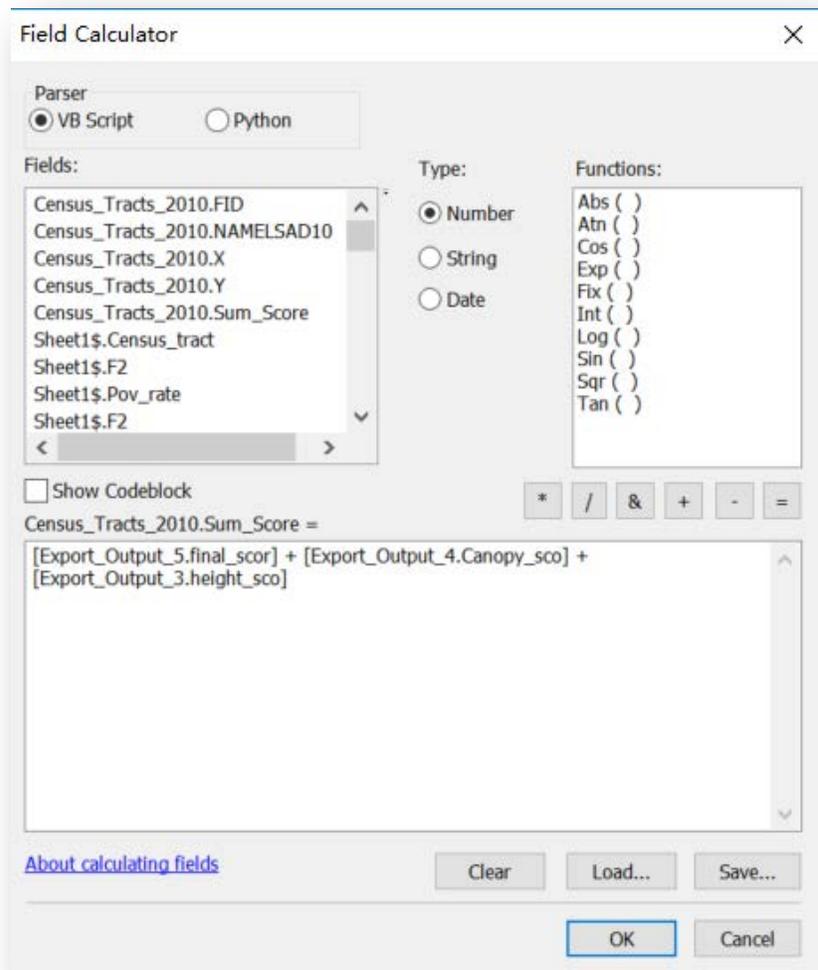
- Export_Output_2
- Height score
- Canopy Area score
- park accessibility score

Export all the score layers' data into dbf. File, and join them into the base layer to calculate the final score later



Part II Data Interpretation-Evaluating the vegetation in Philadelphia

4.-Sum the total score of the vegetation



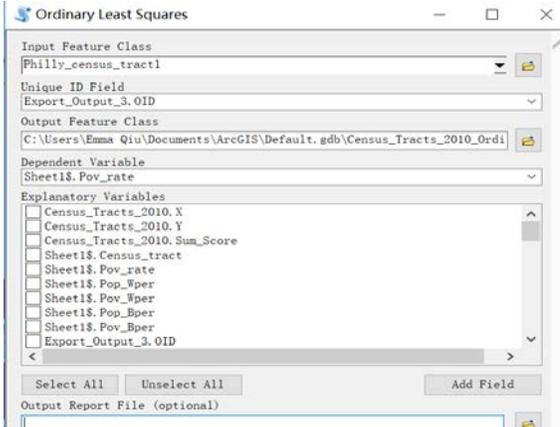
Step 2- Field Calculator

Use Field Calculator to calculate the total score of vegetation in Philadelphia. The Statistics shows that the minimum score is 4, and the maximum score is 14. Most census tracts are between 6 and 11 score interval.

	OID	Sum Score	Count Sum Score
	0	4	12
	1	5	18
	2	6	48
	3	7	42
	4	8	61
	5	9	81
	6	10	46
	7	11	46
	8	12	16
	9	13	13
	10	14	1

Part II Data Interpretation- Explore the relationship

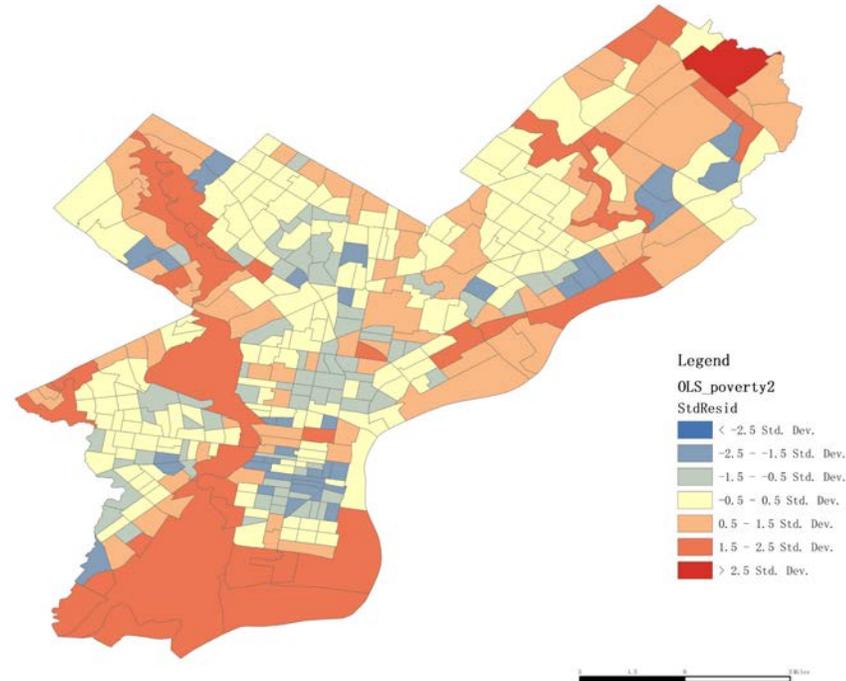
1.-examine the relationship between vegetation and poverty



Summary of OLS Results								
Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust SE	Robust_t	Robust_Pr [b]	
Intercept	3.146409	0.114434	27.495321	0.000000*	0.125077	25.155781	0.000000*	
EXPORT OUTPUT 4.POV RATE	-1.388129	0.382611	-3.628045	0.000336*	0.453509	-3.060863	0.002373*	

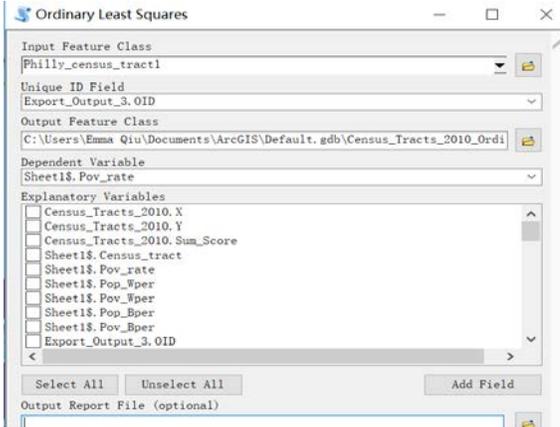
Input Features:		OLS Diagnostics	
Philly_census_tract1		Dependent Variable: EXPORT_OUTPUT_5.FINAL_SCORE	
Number of Observations:	384	Akaike's Information Criterion (AICc) [d]:	1253.891665
Multiple R-Squared [d]:	0.033310	Adjusted R-Squared [d]:	0.030779
Joint F-Statistic [e]:	13.162707	Prob(>F), (1,382) degrees of freedom:	0.000324*
Joint Wald Statistic [e]:	9.368880	Prob(>chi-squared), (1) degrees of freedom:	0.002207*
Koenker (BP) Statistic [f]:	2.717229	Prob(>chi-squared), (1) degrees of freedom:	0.099270
Jarque-Bera Statistic [g]:	10.300461	Prob(>chi-squared), (2) degrees of freedom:	0.005798*

Use Spatial Statistics Tools- Modeling Spatial Relationships- Ordinary Least Squares to build the relationship between vegetation and poverty. Although the R-square is small, the overall model is significant according to the Joint F-Statistic. And, the poverty rate is significantly related to the vegetation. According to the summary of OLS result, when the poverty rate is higher, there is supposed to be less vegetation.



Part II Data Interpretation- Explore the relationship

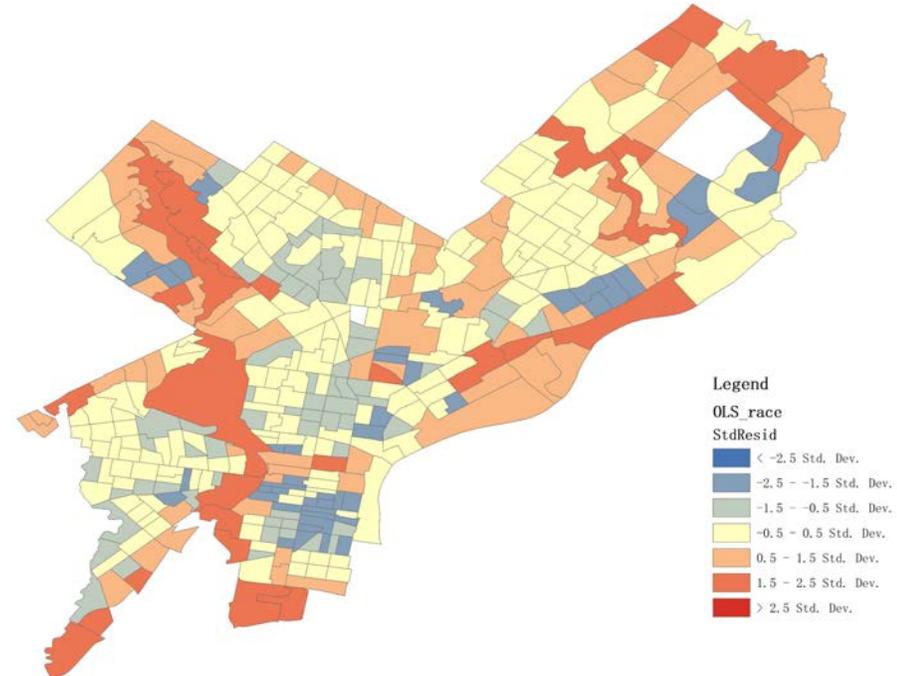
2.-examine the relationship between vegetation and race

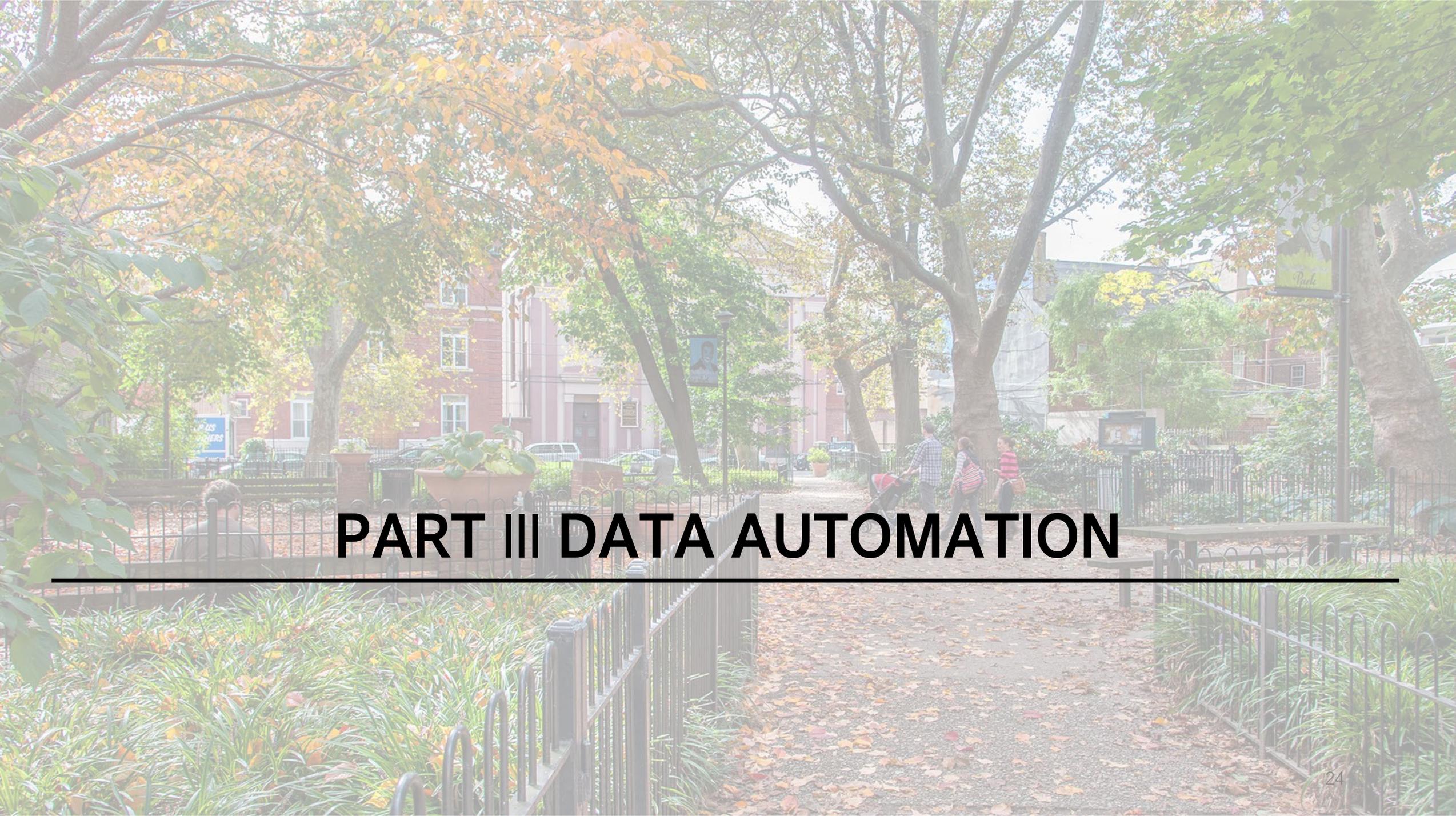


Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]
Intercept	2.963465	0.099030	29.924824	0.000000*	0.101981	29.058955	0.000000*
SHEET1\$.POP_	-0.430970	0.171894	-2.507183	0.012582*	0.161599	-2.666918	0.007982*

Input Features:	Philly_census_tract1	Dependent Variable:	EXPORT_OUTPUT_5.FINAL_S
Number of Observations:	377	Akaike's Information Criterion (AICc) [d]:	1225.777904
Multiple R-Squared [d]:	0.016486	Adjusted R-Squared [d]:	0.013864
Joint F-Statistic [e]:	6.285969	Prob(>F), (1,375) degrees of freedom:	0.012592*
Joint Wald Statistic [e]:	7.112450	Prob(>chi-squared), (1) degrees of freedom:	0.007655*
Koenker (BP) Statistic [f]:	6.075476	Prob(>chi-squared), (1) degrees of freedom:	0.013707*
Jarque-Bera Statistic [g]:	16.003797	Prob(>chi-squared), (2) degrees of freedom:	0.000335*

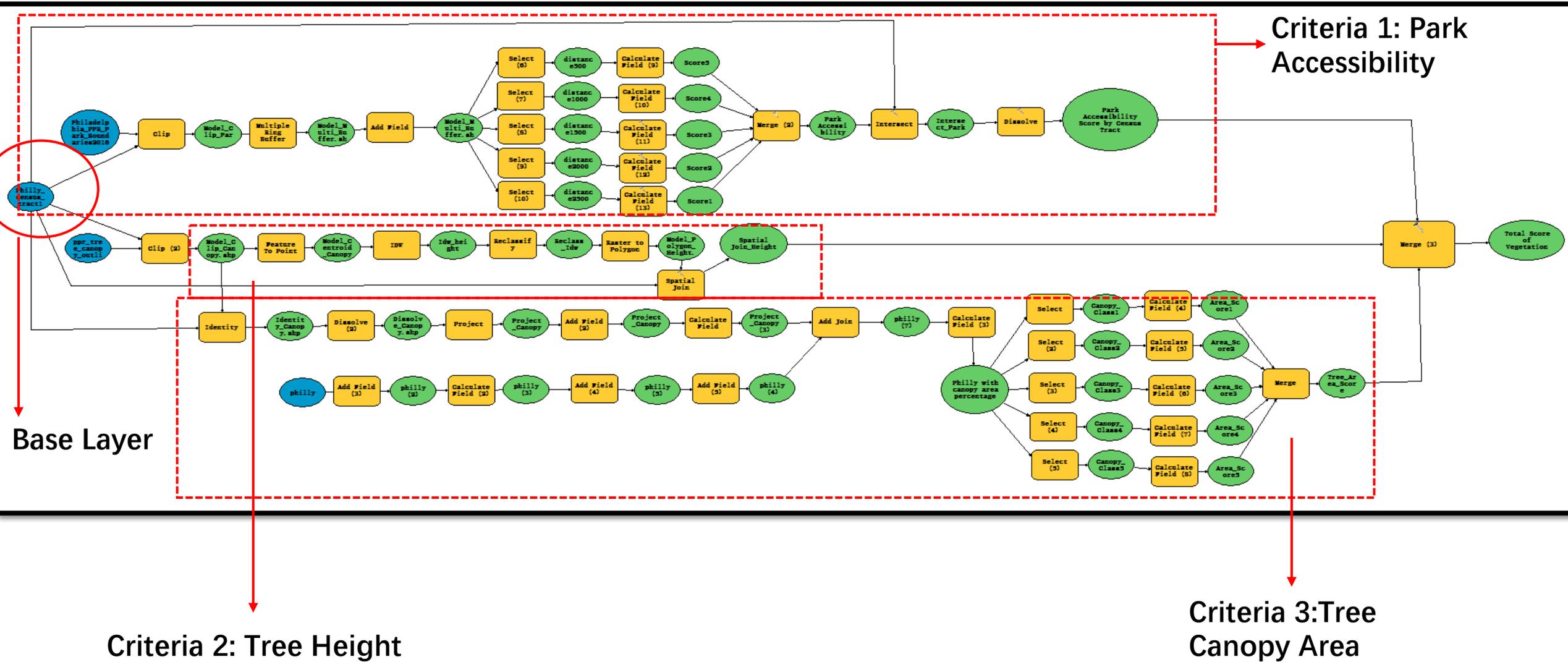
Use Spatial Statistics Tools- Modeling Spatial Relationships- Ordinary Least Squares to build the relationship between vegetation and race. Same as the previous explanation, although the R-square is small, the overall model is significant according to the Joint F-Statistic. And, the Black population percentage is significantly related to the vegetation. According to the summary of OLS result, when there are more Black race in the census tract, it is likely to have less vegetation.



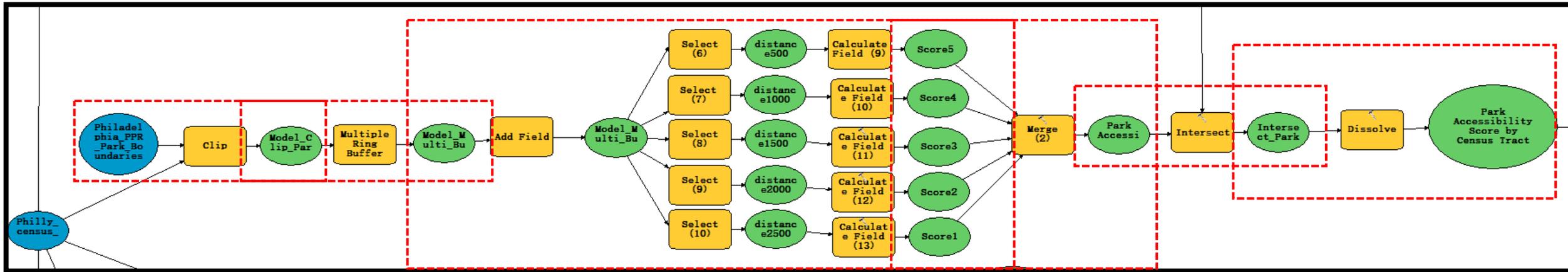
A photograph of a park path with trees and a building in the background. The path is paved and covered with fallen leaves. There are several people walking on the path, including a person pushing a stroller. The trees are mostly green, with some showing yellow and orange autumn colors. In the background, there is a large, multi-story brick building with many windows. A black metal fence runs along the path. A sign with a portrait and the word "Park" is visible on the right side of the path.

PART III DATA AUTOMATION

Overview



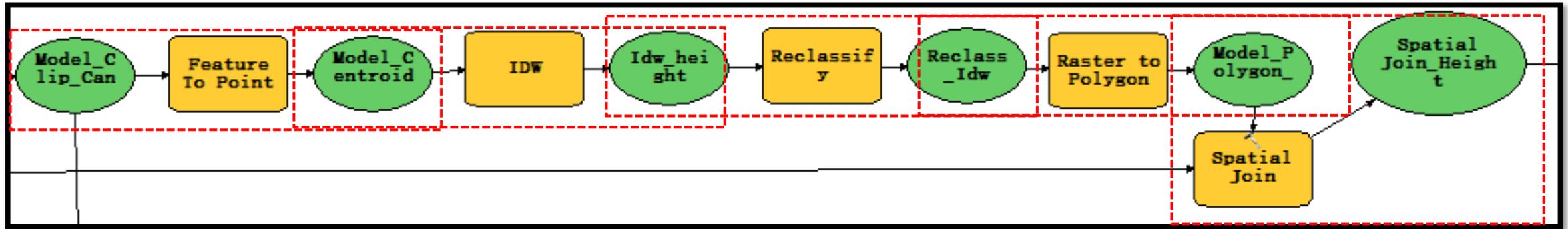
Criteria 1_ Park Accessibility



- First, clip the park within the boundaries of Philadelphia;
- Second, create multiple ring buffer;
- Third, after adding a field named “score” , use select and calculate field to assign the score from 1 to 5 to different distances;
- Fourth, merge the five buffers into one shapefile
- Fifth, intersect the merged shapefile with the base layer;
- Sixth, dissolve the park by census tracts, and sum the total score of each census tracts’ park accessibility

Distance from Park	Score
500 meters	1
1000 meters	2
1500 meters	3
2000 meters	4
2500 meters	5

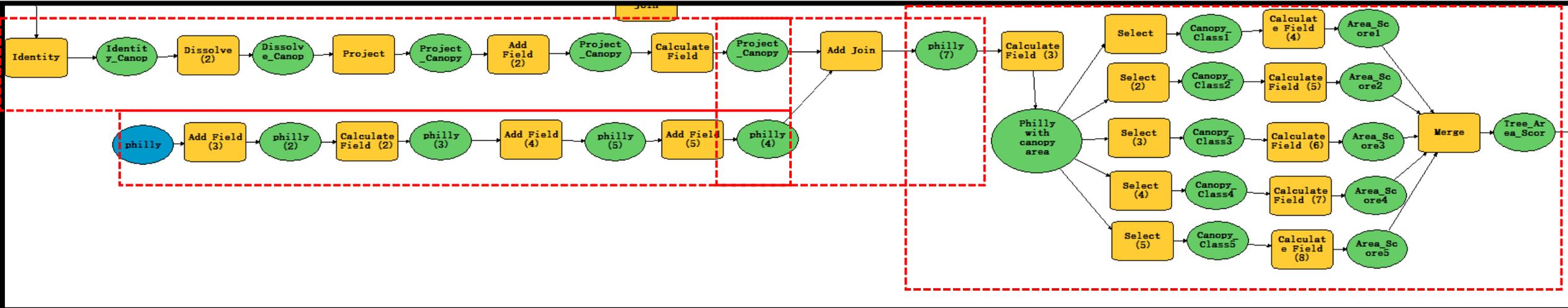
Criteria 2_ Tree Height



- After clipping the tree canopy outlines within Philadelphia, first, use feature to point to display centroid of each tree canopy;
- Second, use IDW to interpolate the tree height;
- Third, use reclassify to reclass the quantile classification to score 1 to 5;
- Fourth, transfer the raster to polygon for the further spatial join;
- Fifth, join the clipped canopy with the interpolated tree height shapefile to get the score of tree height in Philadelphia

Tree Height	Score
6.030121 – 19.275738	1
19.275738 – 28.374294	2
28.374294 – 40.795418	3
40.795418 – 57.402523	4
57.402523 – 262.917203	5

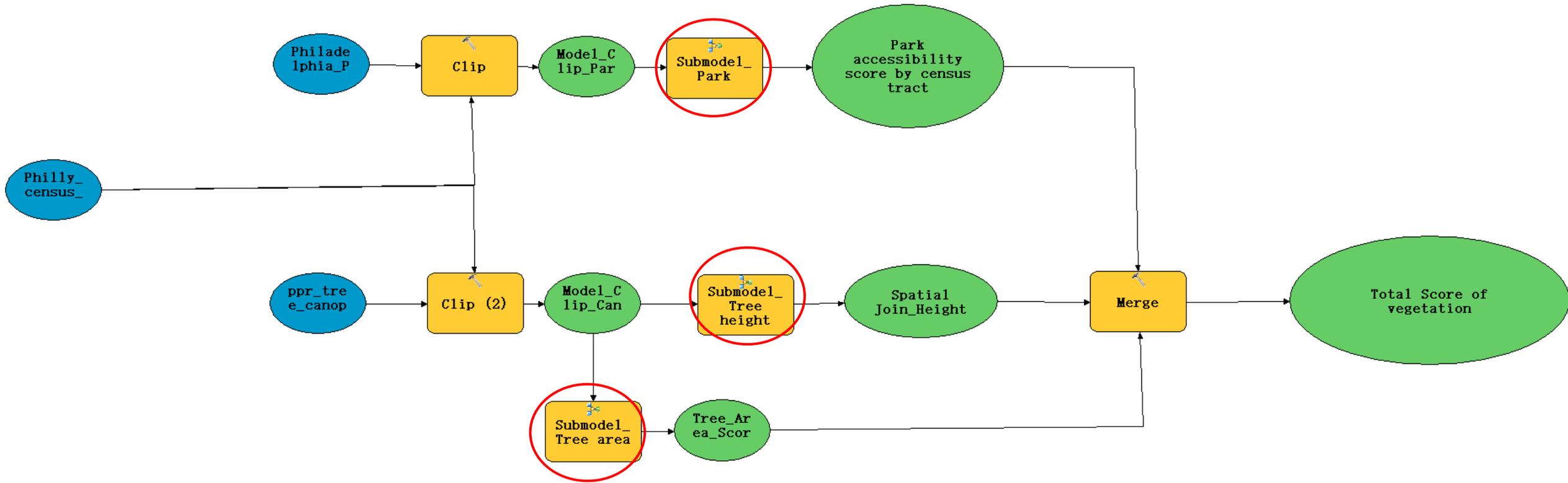
Criteria 3_ Tree Canopy Area



- Use the same base layer as the second criteria, the clipped tree canopy shapefile
- First, use identity, dissolve, project, add field and field calculator to calculate the sum area of each tree canopy in the census tracts;
- At the same time, calculate the area of each census tracts;
- Second, calculate the percentage of the area of tree canopy in different census tract;
- Third, use select and calculate field to assign the different score to different quantile level;
- Finally, merge the five score shapefiles into one, it is the tree area score shapefile.

Tree Canopy area %	Score
0.012749 – 0.083363	1
0.083364 – 0.132468	2
0.132469 – 0.172860	3
0.172861 – 0.240200	4
0.240201 – 0.790691	5

Submodel



- To make the model more clearly, create three sub models about three criteria

CONCLUSION

1. The legacy of structural racism is still influencing the living pattern in Philadelphia, where most of the Black are living in West Philadelphia, North and Upper North Philadelphia, while the White are living in the Central Philadelphia, Riverwards, Northwest Philadelphia, and Far Northeast Philadelphia.
2. Vegetation in Philadelphia is mostly in Northwest and Northeast Philadelphia, while respectively less vegetation in South Philadelphia.
3. The spatial distribution of vegetation in Philadelphia is significantly related to race and poverty. It is more likely to have less vegetation where has more Black population or higher poverty rate.