05 – Measuring Geographic Distributions

J.S. Onésimo Sandoval
Sociology
Saint Louis University

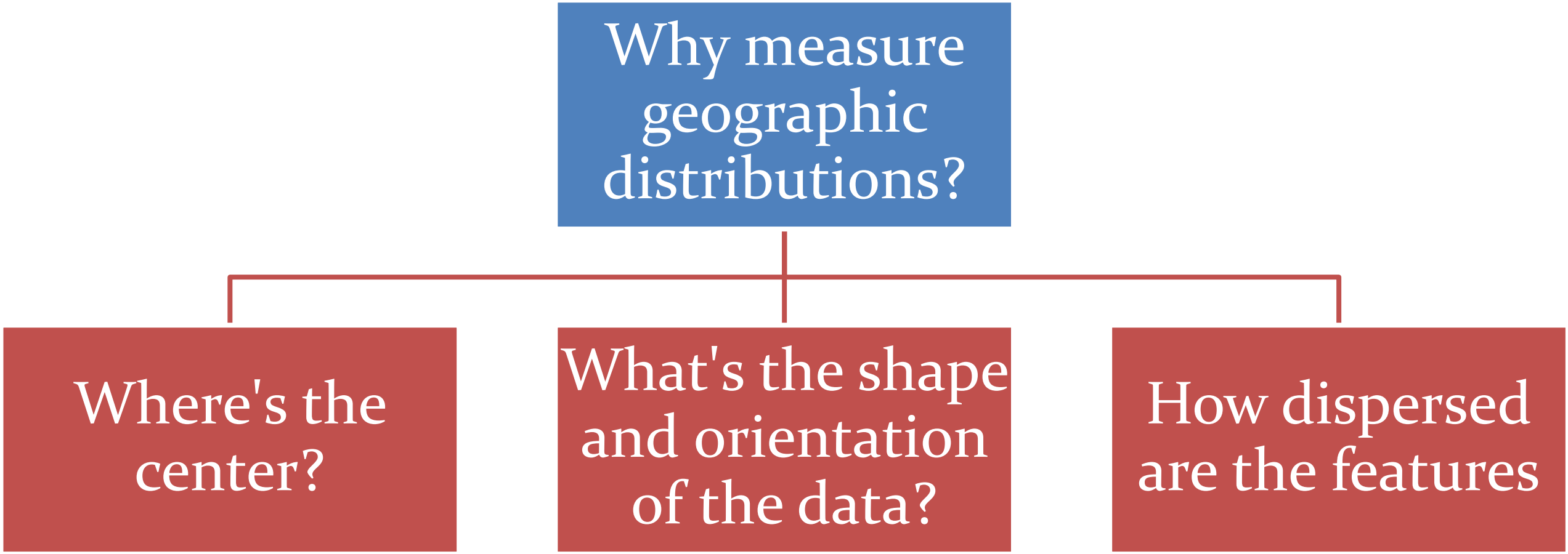
Outline

- Why Geographic Distribution?
- Mean Center
- Median Center
- Central Feature
- Standard Distance
- Standard Deviational Ellipse
- Polygons and Distribution
- Lab Exercise

Although basic science is directed at the discovery of general principles, the ultimate value of such knowledge, apart from simple curiosity, lies in our ability to apply it to local conditions and, thus, determine specific outcomes. Although such science may itself be placeless, the application of scientific knowledge in policy inevitably requires explicit attention to ***spatial variation****, particularly when the basis of policy is local.*

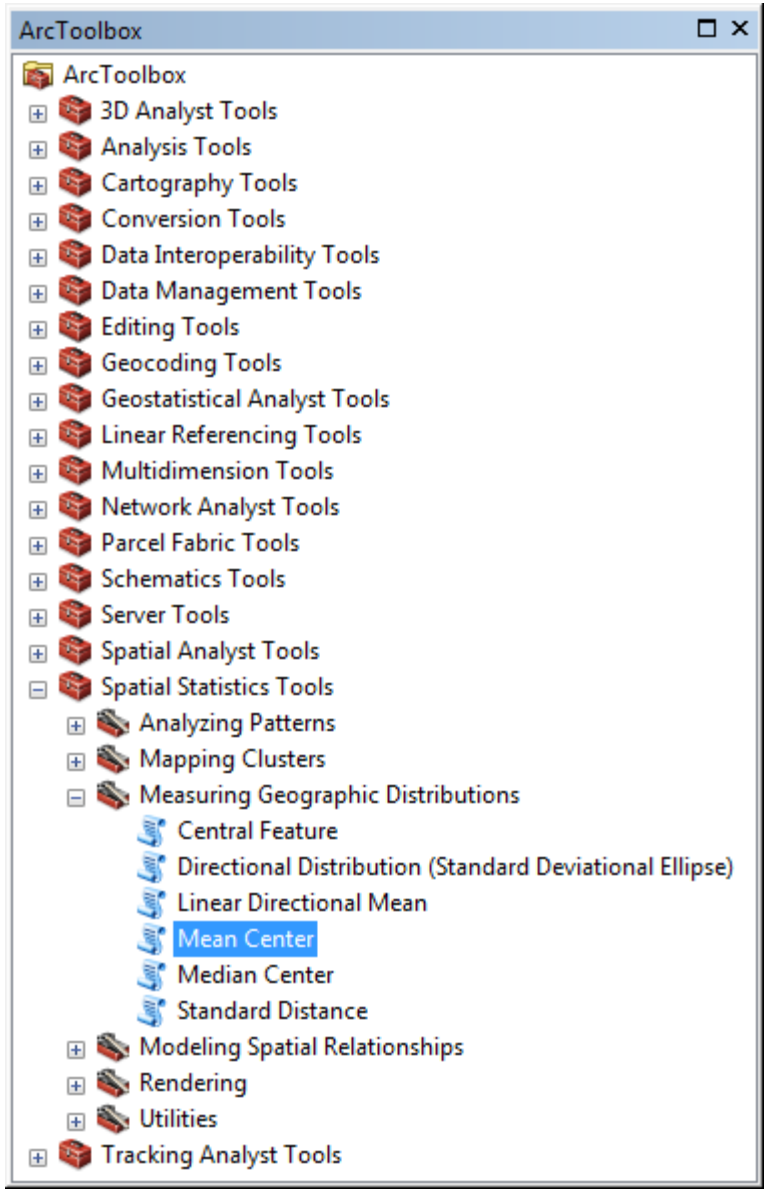
(Goodchild et al. 2000, p. 142)

Why Geographic Distribution?


Why measure geographic distributions?
Where's the center?
What's the shape and orientation of the data?
How dispersed are the features

Descriptive Spatial Statistics

- Goals
  - To describe the *distribution* of a set of features
  - To objectively measure *compactness*
  - To objectively measure *orientation* and *direction*



Three Kinds of Center

Center	What it represents	What it's good for
Mean	The *average x-coordinate and average y-coordinate* for all features in the study area	Tracking changes or comparing distribution
Median	The x,y coordinate having *the shortest distance* to all features in the study area	Finding the most accessible location
Central Feature	The feature having the *shortest total distance* to all other features in the study area	Finding the most accessible feature

Mean Center

Mean Center

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\bar{X}$ = the mean x-coordinate

$\sum_i X_i$ = sum the x-coordinates values (X) of the features,

(i represents each indiviual observation or feature)

n = divide the sum by the number of features

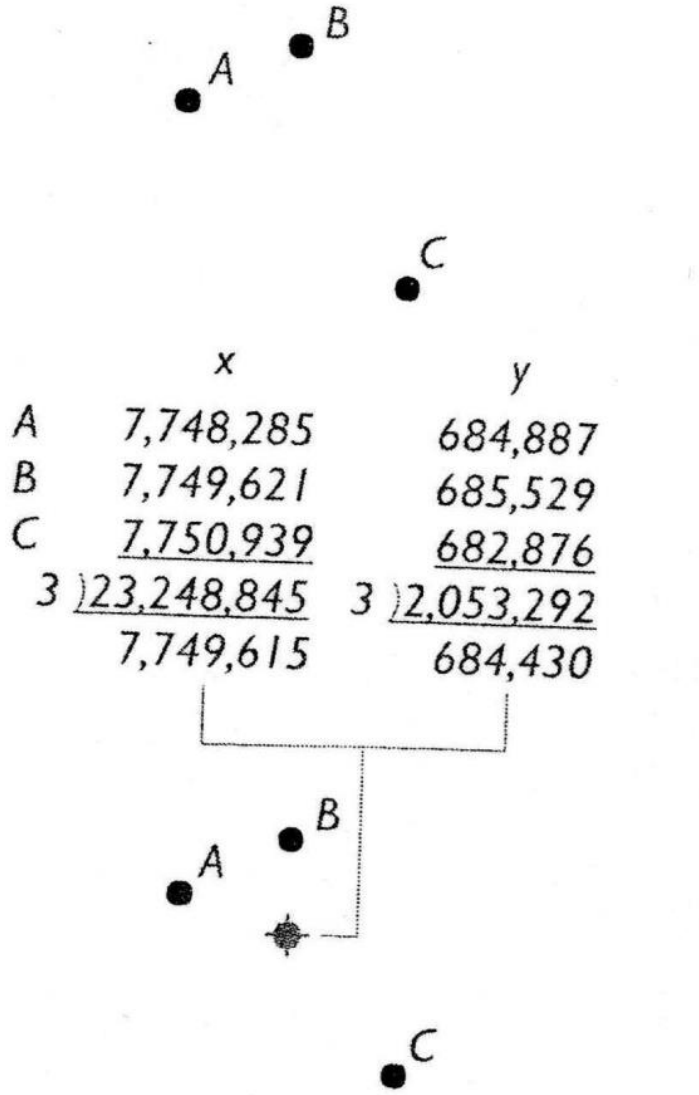
$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

$\bar{Y}$ = the mean y-coordinate

$\sum_i Y_i$ = sum the y-coordinates values (y) of the features,

(i represents each indiviual observation or feature)

n = divide the sum by the number of features

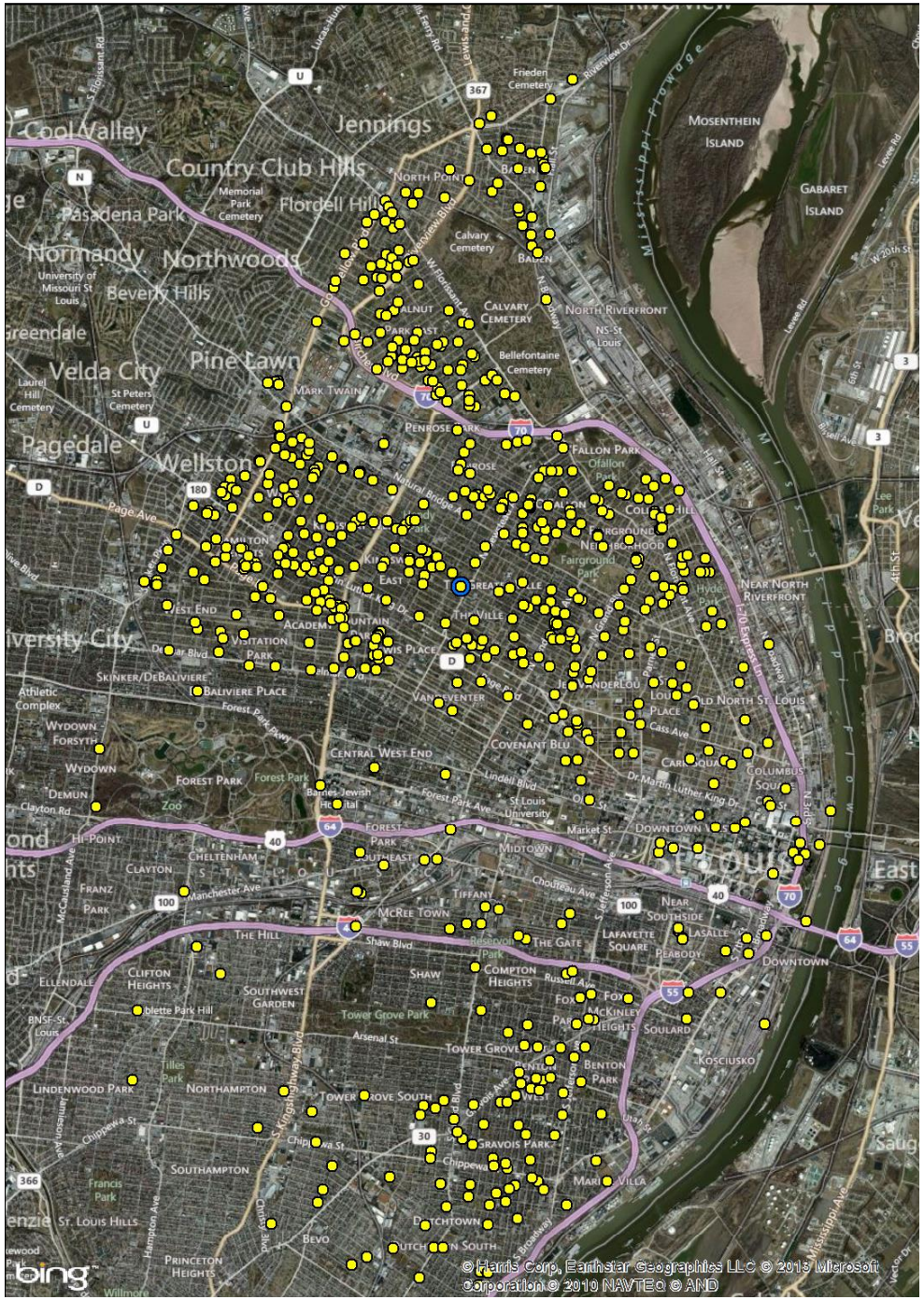


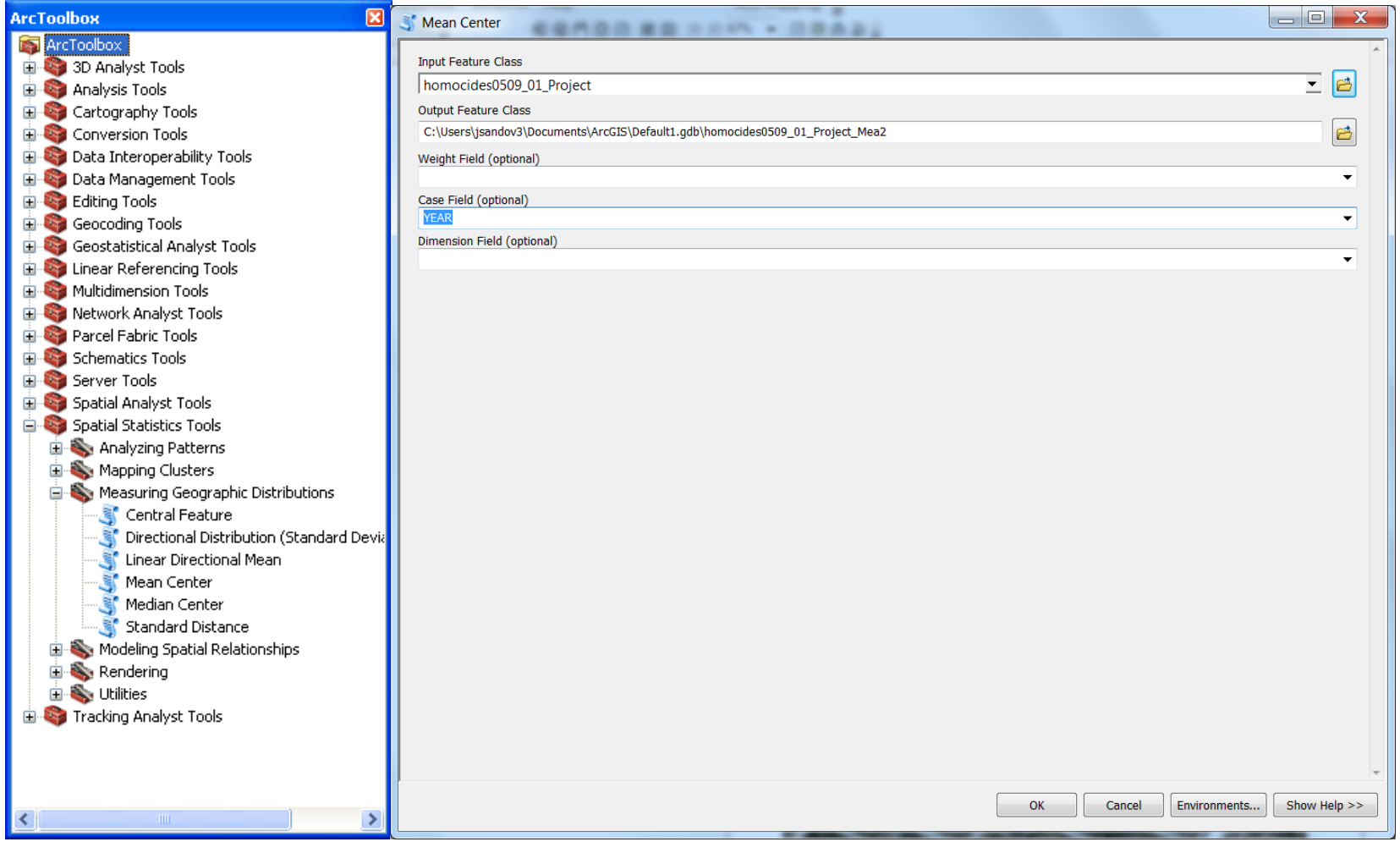
Weighted Mean Center

$$\bar{X}_w = \frac{\sum_{i=1}^{n} W_i X_i}{\sum_{i=1}^{n} W_i} \qquad \bar{Y}_w = \frac{\sum_{i=1}^{n} W_i Y_i}{\sum_{i=1}^{n} W_i}$$

W_i = is the weight of feature i

Homicides in Saint Louis 2005-2009 (n=780)



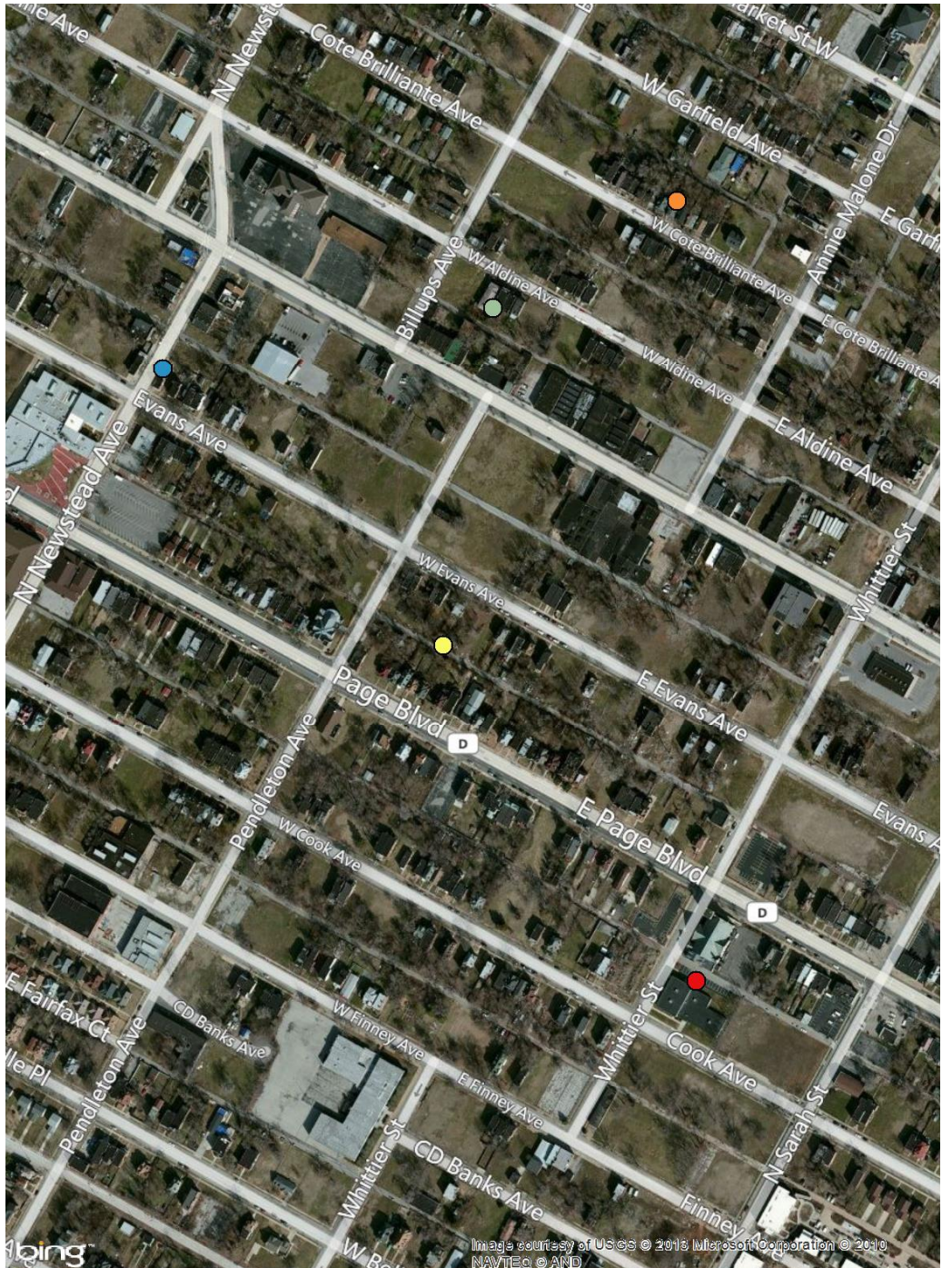

ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Devia
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools
Mean Center
Input Feature Class
homocides0509_01_Project
Output Feature Class
C:\Users\jsandov3\Documents\ArcGIS\Default1.gdb\homocides0509_01_Project_Mea2
Weight Field (optional)
Case Field (optional)
YEAR
Dimension Field (optional)
OK
Cancel
Environments...
Show Help >>

Mean Center

Legend

YEAR

- 2005
- 2006
- 2007
- 2008
- 2009



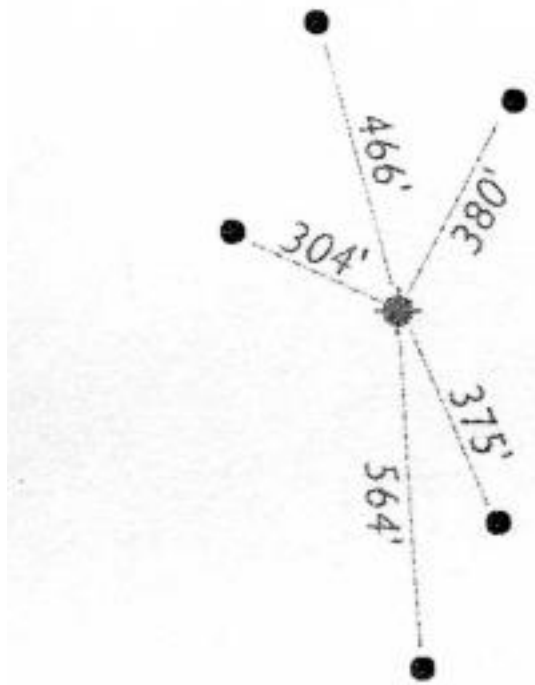
Median Center

Median Center

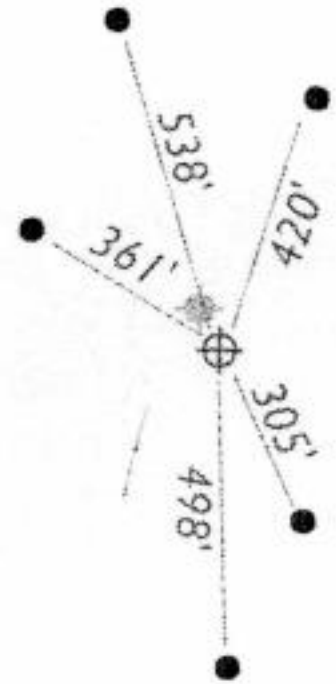
- The x,y coordinate having *the shortest distance* to all features in the study area

- Note: There is no single equation for calculating the exact median center. The software approximates the center by interactively calculating the mean center, summing the distances from it to each feature, offsetting the center slightly and summing the distance again. Since the calculation attempts to minimize the total distance from all features, the center will gravitate toward the areas with the most features.

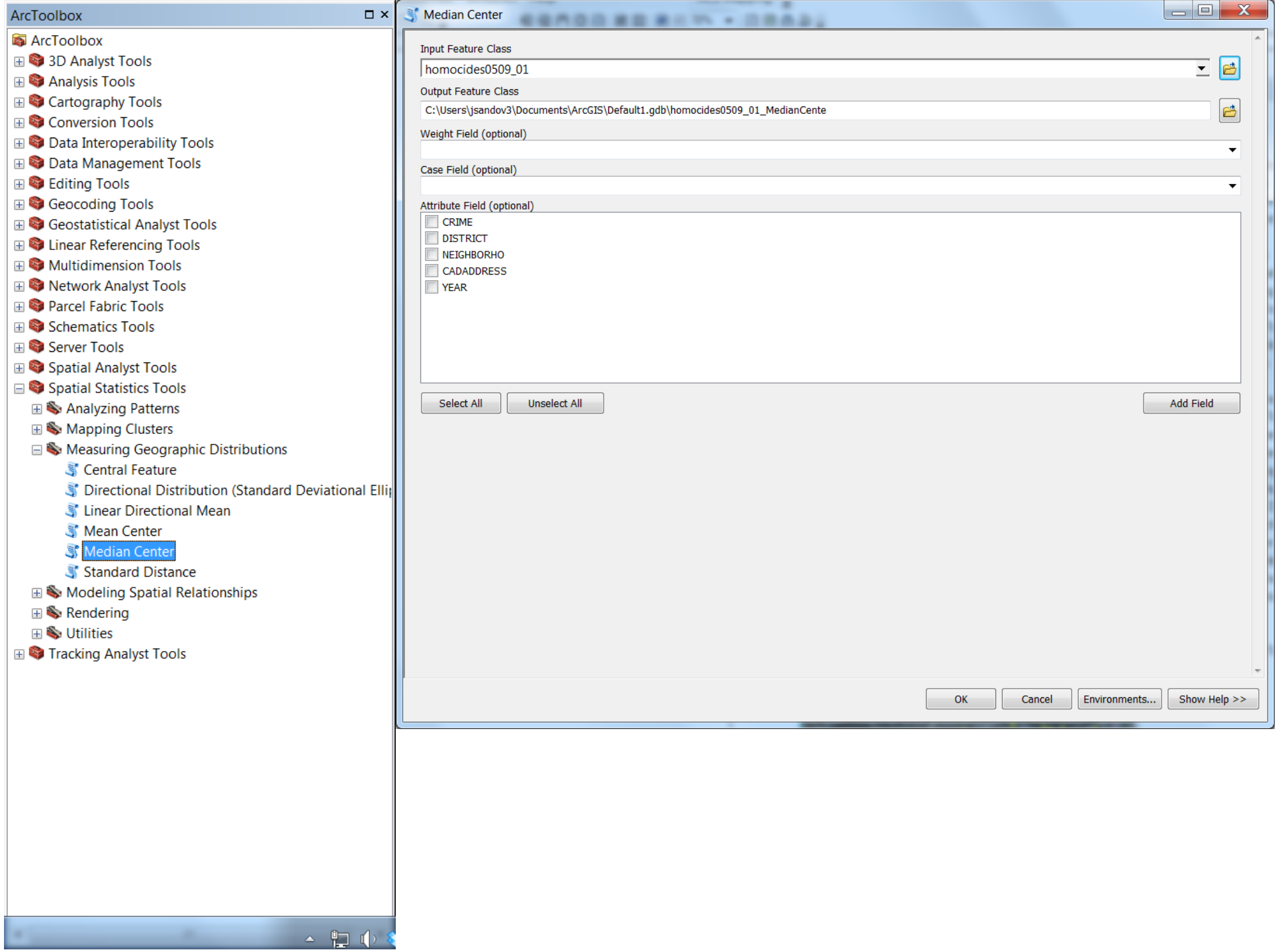
Conceptualization of Median Center



Total distance to mean center = 2,089 ft.

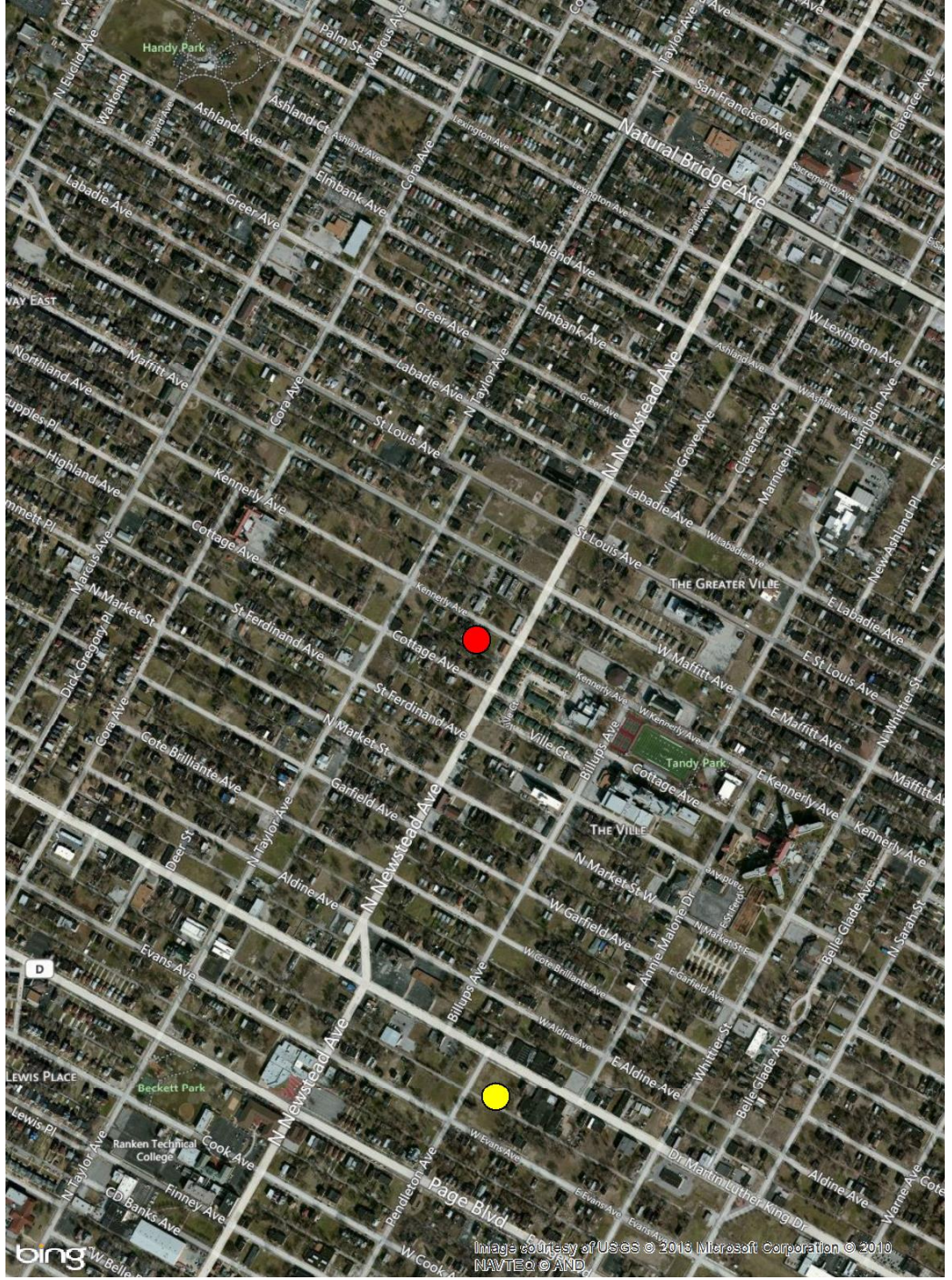


Total distance to offset = 2,122 ft.


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellip
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools
Median Center
Input Feature Class
homocides0509_01
Output Feature Class
C:\Users\jsandov3\Documents\ArcGIS\Default1.gdb\homocides0509_01_MedianCente
Weight Field (optional)
Case Field (optional)
Attribute Field (optional)
CRIME
DISTRICT
NEIGHBORHO
CADADDRESS
YEAR
Select All
Unselect All
Add Field
OK
Cancel
Environments...
Show Help >>

Legend

- Median0509
- Mean0509

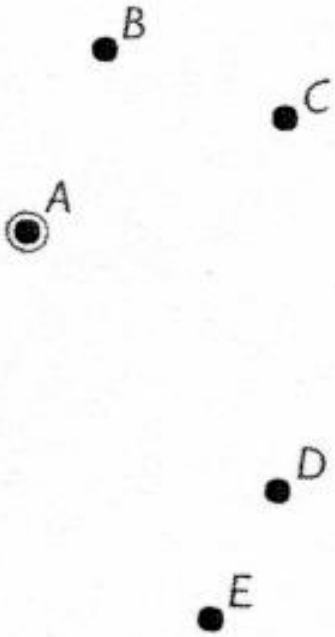


Central feature

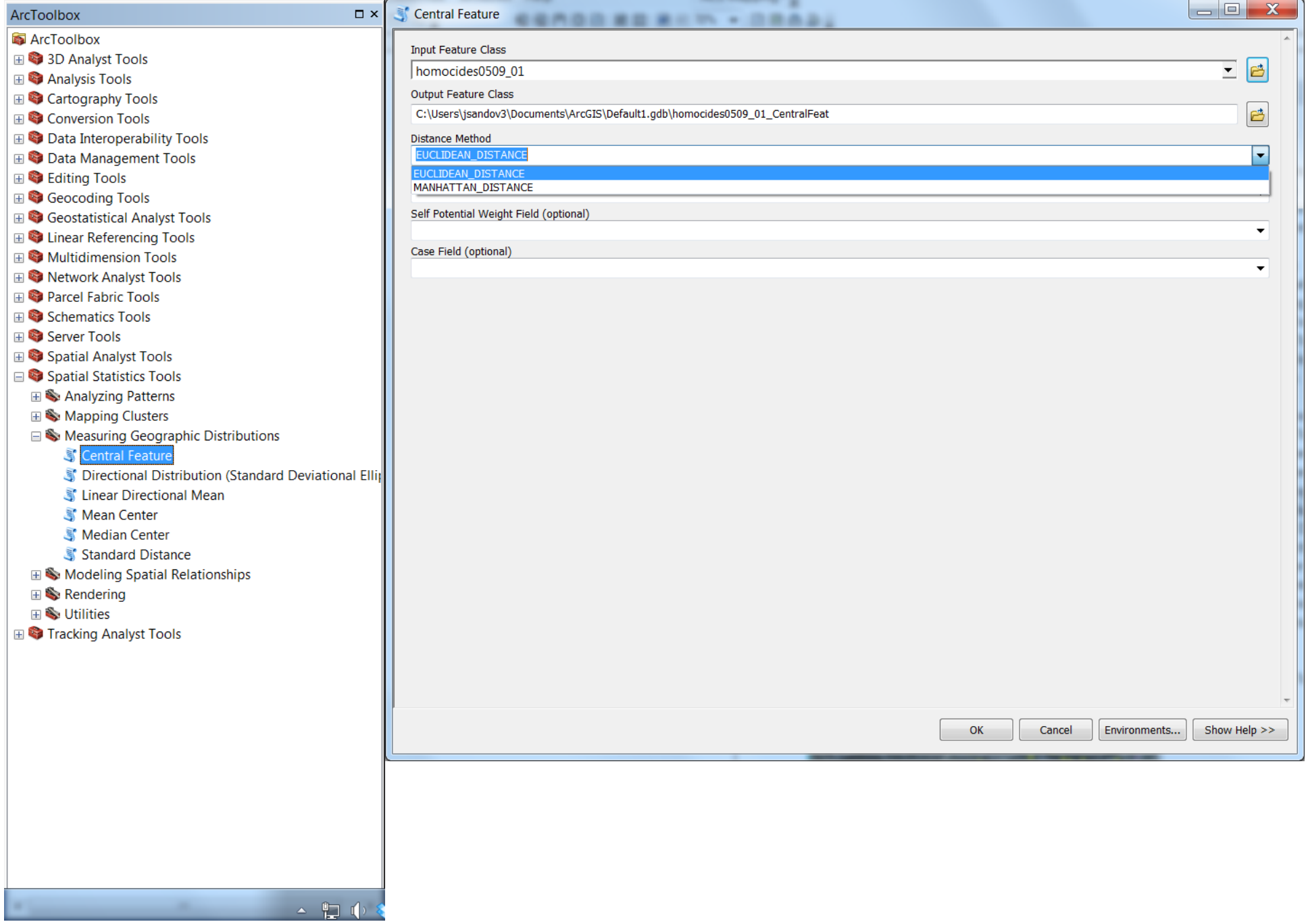
Central Feature

- The feature having the *shortest total distance* to all other features in the study area.

  Note: The software totals the distance from each feature to every feature. The feature with the lowest total distance to all other features is the central feature.

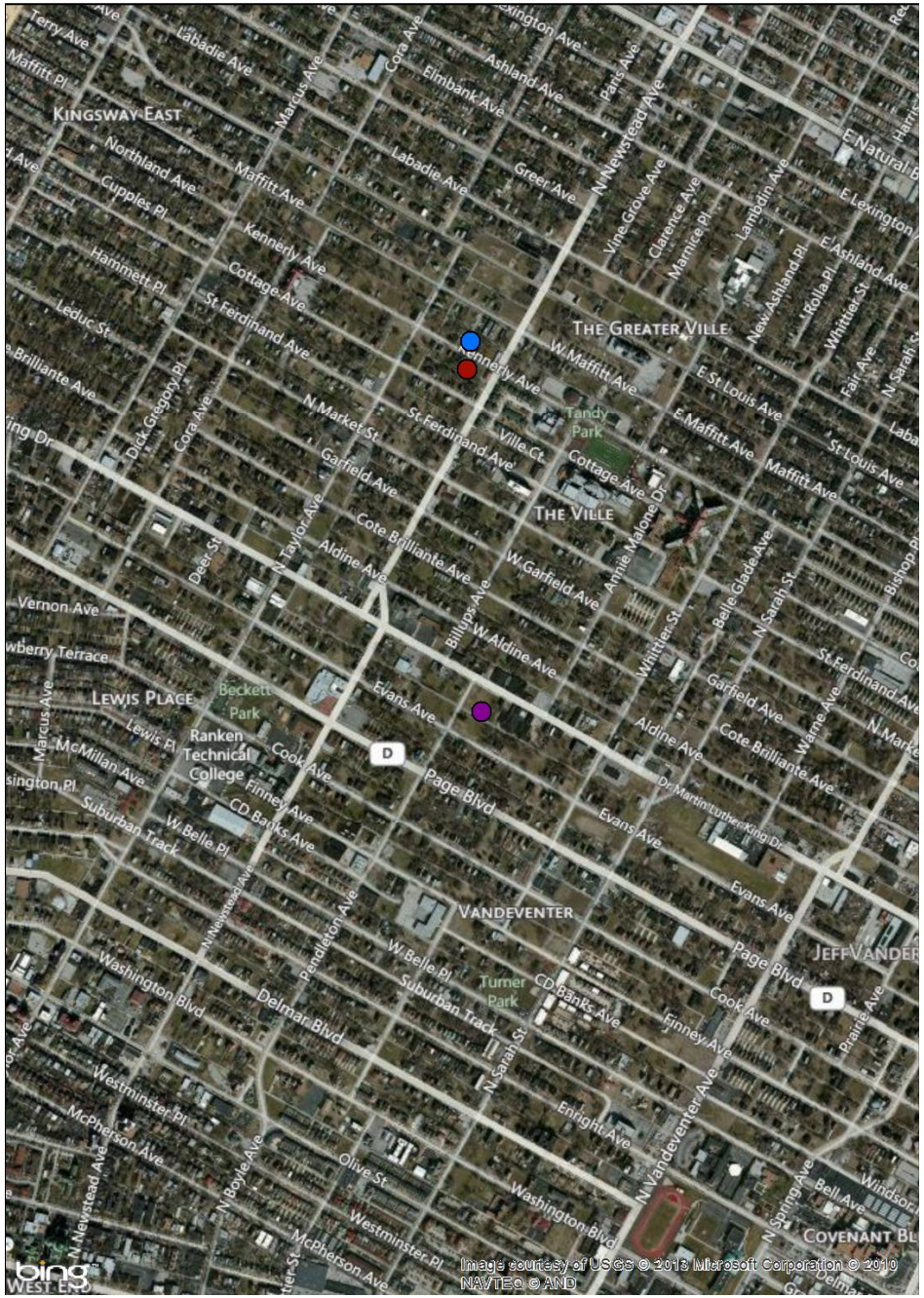


	A	B	C	D	E	SUM
A	0'	352'	517'	640'	759'	**2268'**
B	352'	0'	354'	840'	1023'	**2569'**
C	517'	354'	0'	658'	901'	**2430'**
D	640'	840'	658'	0'	262'	**2400'**
E	759'	1023'	901'	262'	0'	**2945'**


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellip
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools
Central Feature
Input Feature Class
homocides0509_01
Output Feature Class
C:\Users\jsandov3\Documents\ArcGIS\Default1.gdb\homocides0509_01_CentralFeat
Distance Method
EUCLIDEAN_DISTANCE
EUCLIDEAN_DISTANCE
MANHATTAN_DISTANCE
Self Potential Weight Field (optional)
Case Field (optional)
OK
Cancel
Environments...
Show Help >>

- This is for five years

- We can compute these statistics for year to track change.

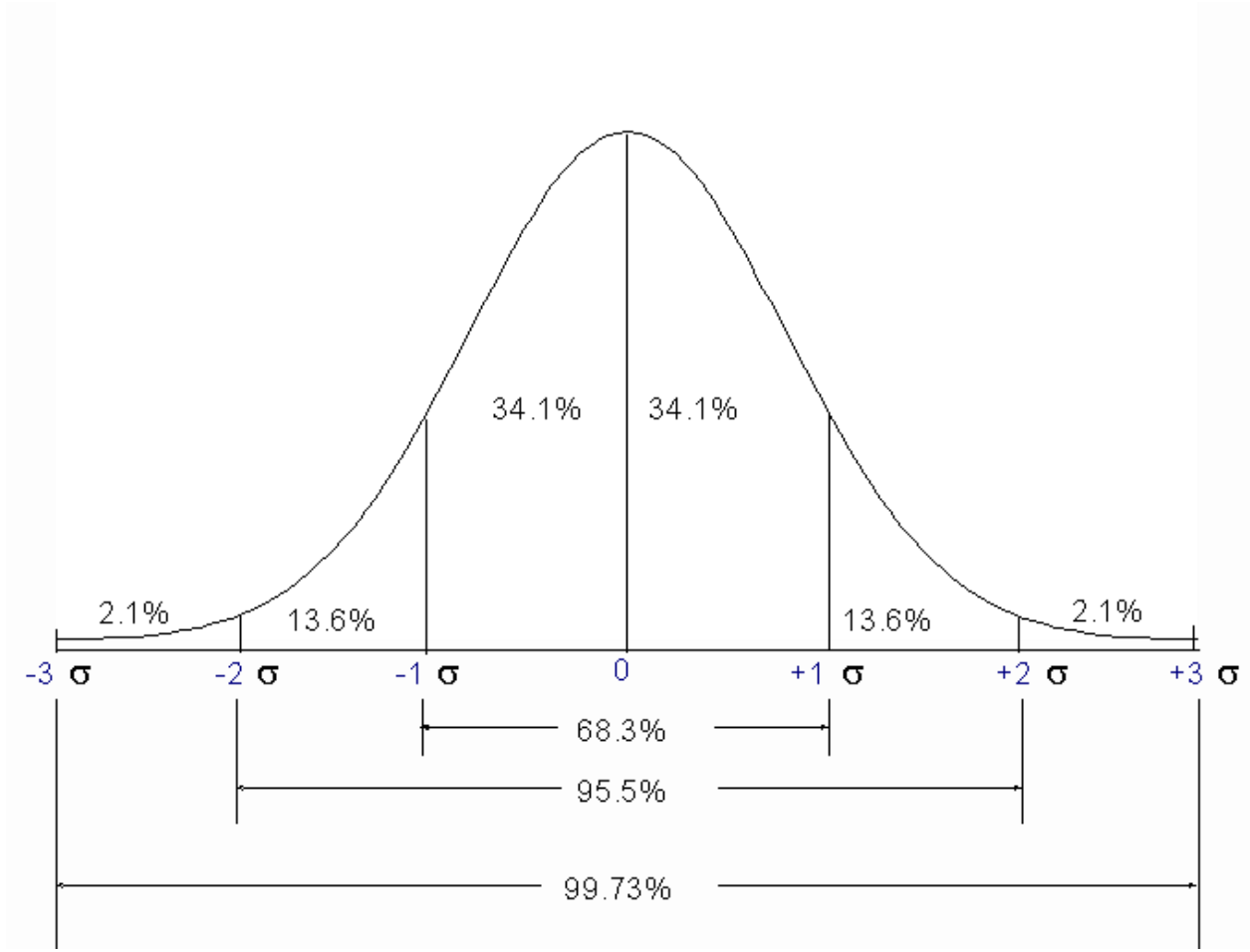


Legend

- central_feature0509e
- median_center0509
- mean_center0509

Standard Distance

Percentages Under the Normal Curve or "The 68-95-99 rule"

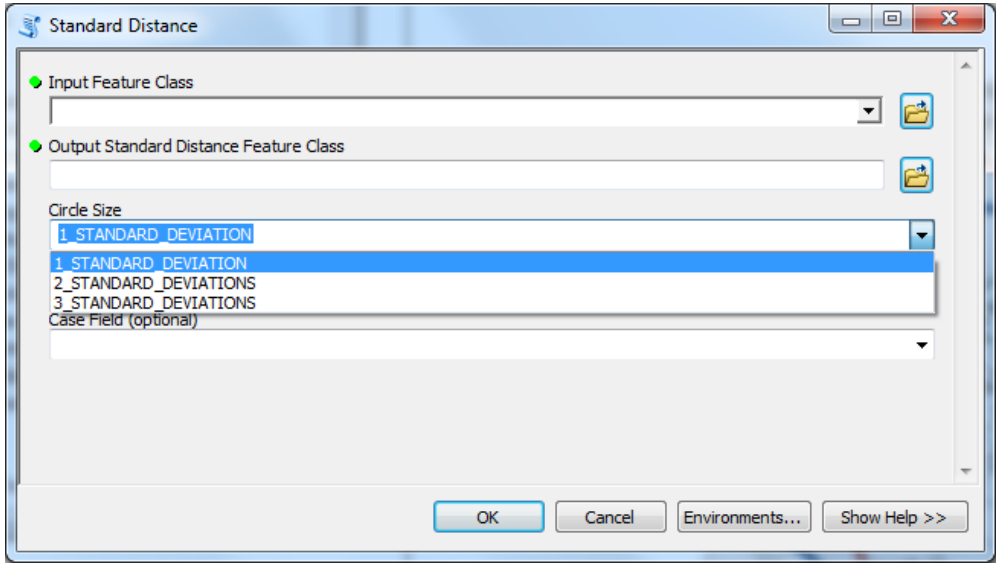


Measuring Compactness (Circle or Ellipses Size)

- A one standard deviation polygon
will cover approximately 68 percent
of the features;

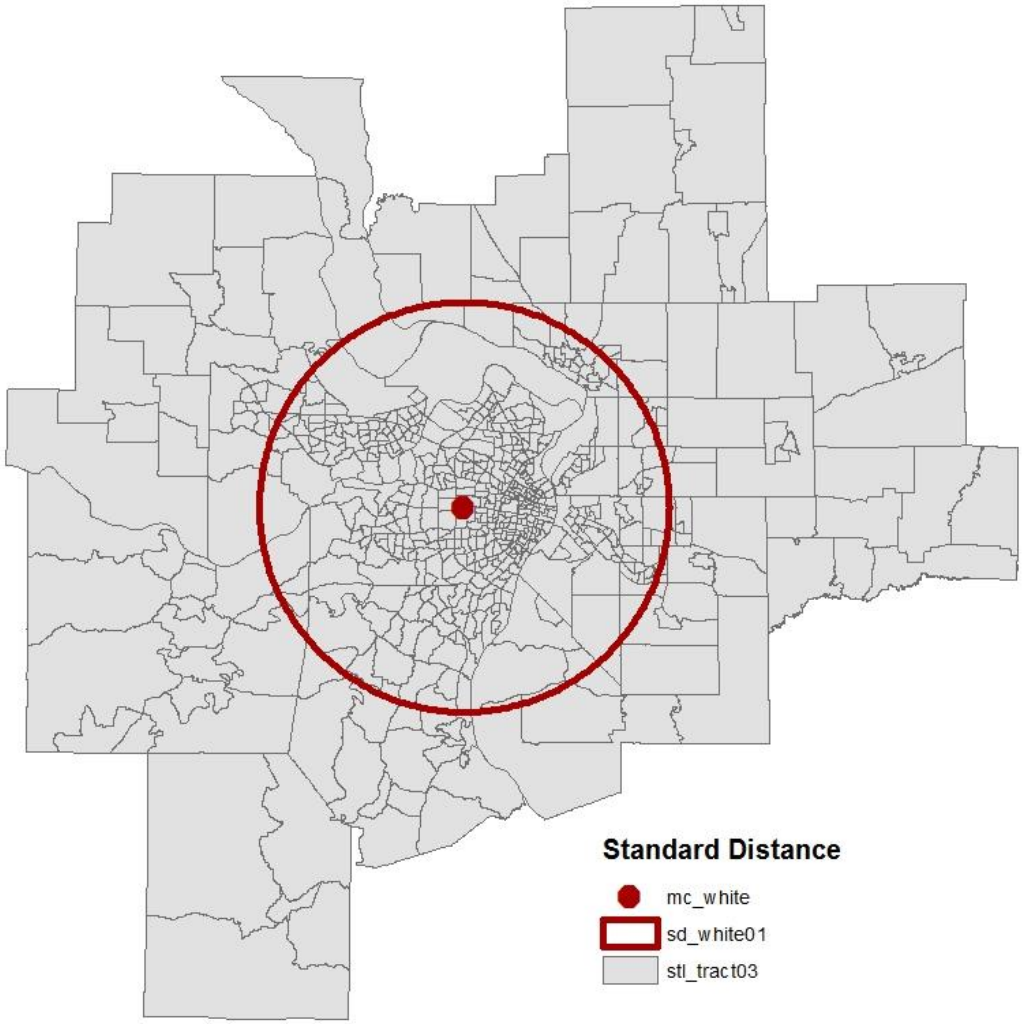
- A two standard deviation polygon
will contain approximately 95 percent
of the features;

- A three standard deviation polygon
will cover approximately 99 percent
of the features in the cluster.



The standard distance measure the extent to which the distances between the mean center and the features vary from the average distance.

The greater the standard distance, the more the distances vary from the average, and the more widely dispersed the features around the center.

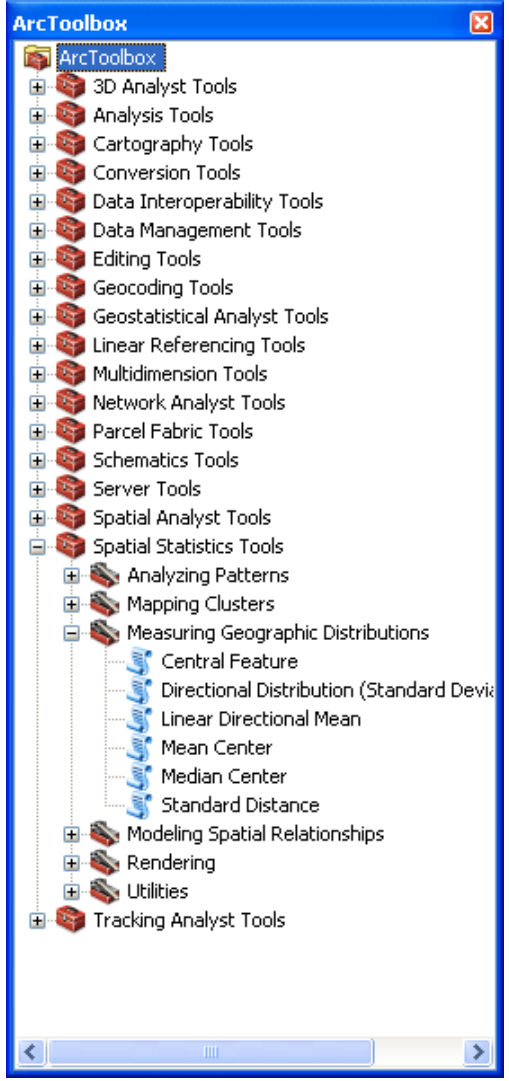


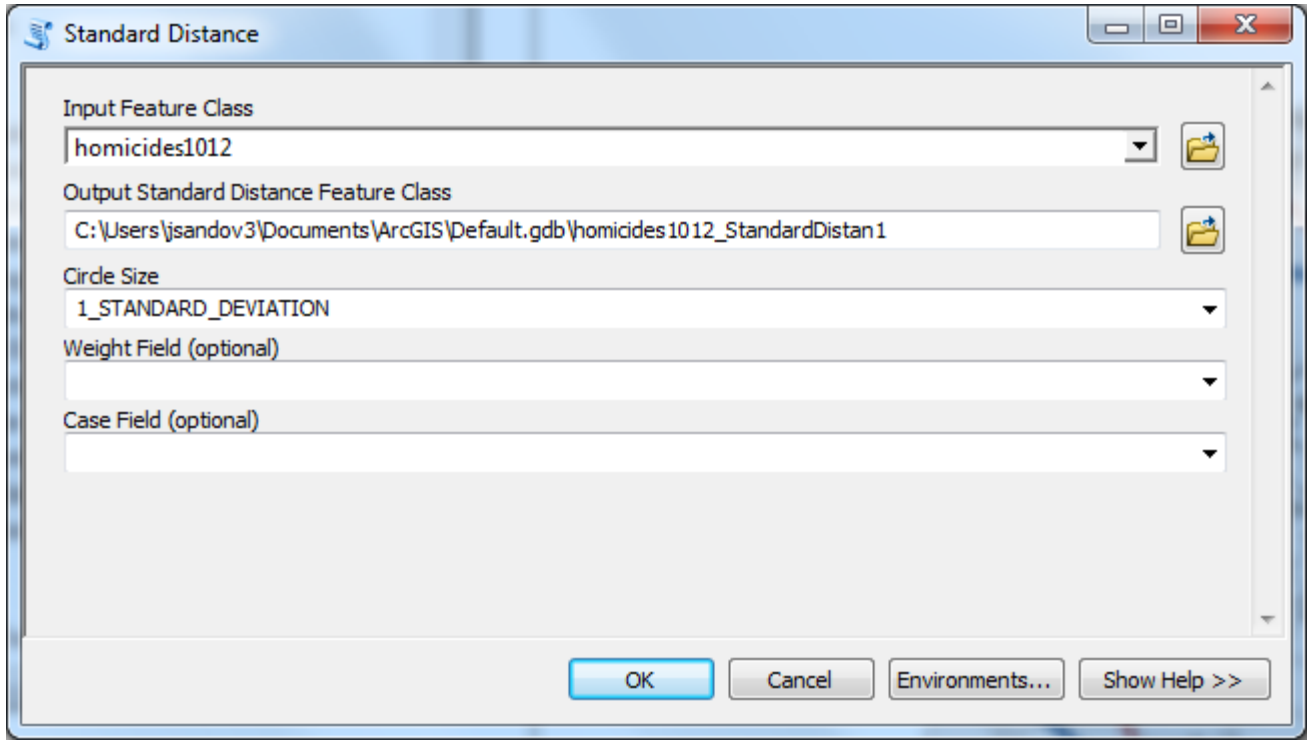
Standard Distance (SD)

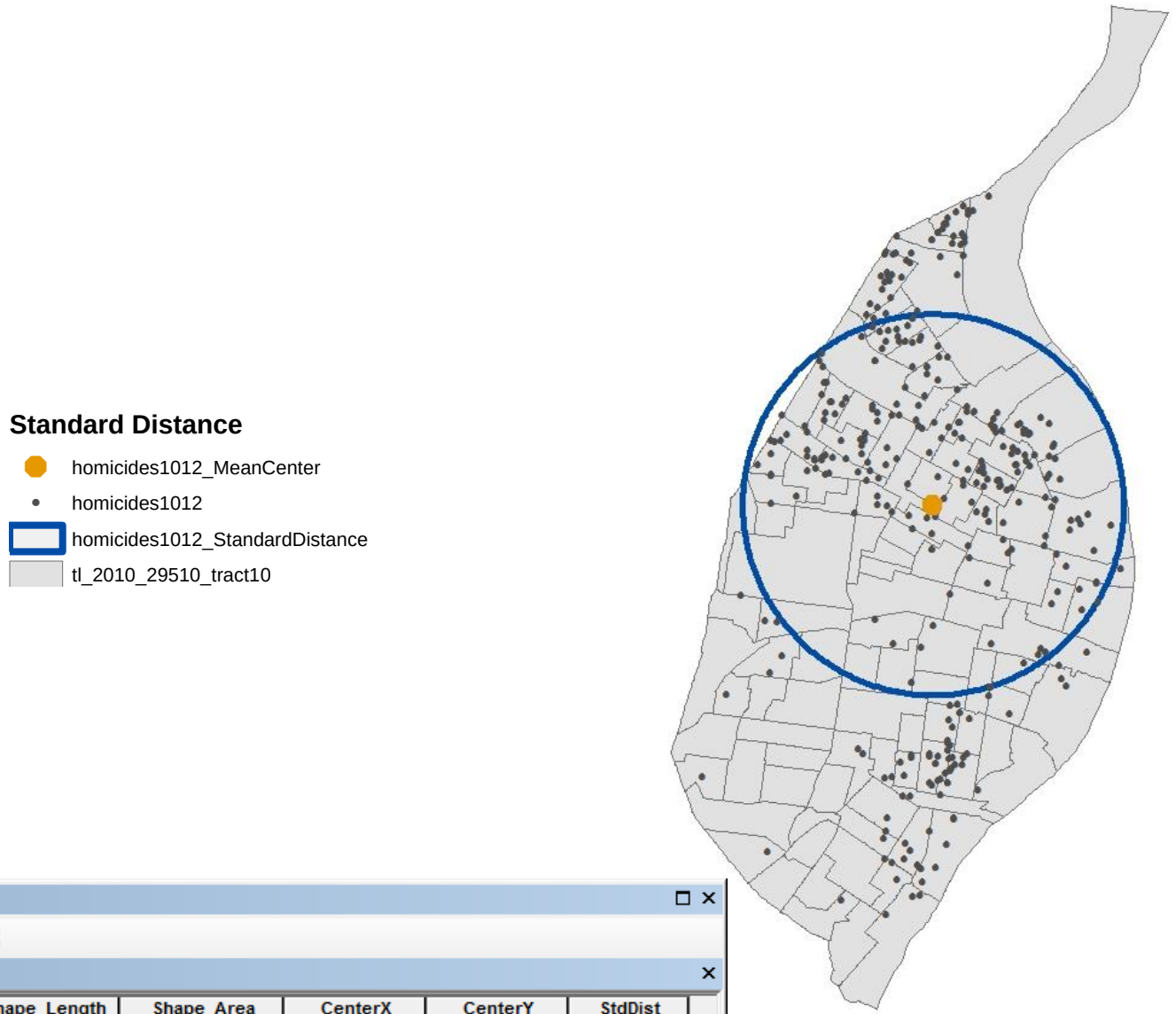
The SD measures the extent to which the distances between the mean center and the features vary from the average distance

$$SD = \sqrt{\frac{\sum_i (X_i - \bar{X})^2}{n} + \frac{\sum_i (Y_i - \bar{Y})^2}{n}}$$

	X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})^2$	$(Y - \bar{Y})^2$	
A	7,748,285	684,887	-1,330	457	1,768,900	208,849	
B	7,749,621	685,529	6	1,099	36	1,207,801	
C	7,750,939	682,876	1,324	-1554	1,752,976	2,414,916	
Total					3,521,912/3	3,831,566/3	
					1,173,971	1,277,189	2,451,159
						SD	$\sqrt{2451159}$
							1,565.6


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Devia
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools


Standard Distance
Input Feature Class
homicides1012
Output Standard Distance Feature Class
C:\Users\jsandov3\Documents\ArcGIS\Default.gdb\homicides1012_StandardDistan1
Circle Size
1_STANDARD_DEVIATION
Weight Field (optional)
Case Field (optional)
OK
Cancel
Environments...
Show Help >>



Table

homicides1012_StandardDistan

OBJECTID *	Shape *	Shape_Length	Shape_Area	CenterX	CenterY	StdDist
1	Polygon	32134.705477	82172738.198793	272673.374791	313048.000929	5114.462139

1 (0 out of 1 Selected)

homicides1012_StandardDistan

Standard Deviational Ellipse

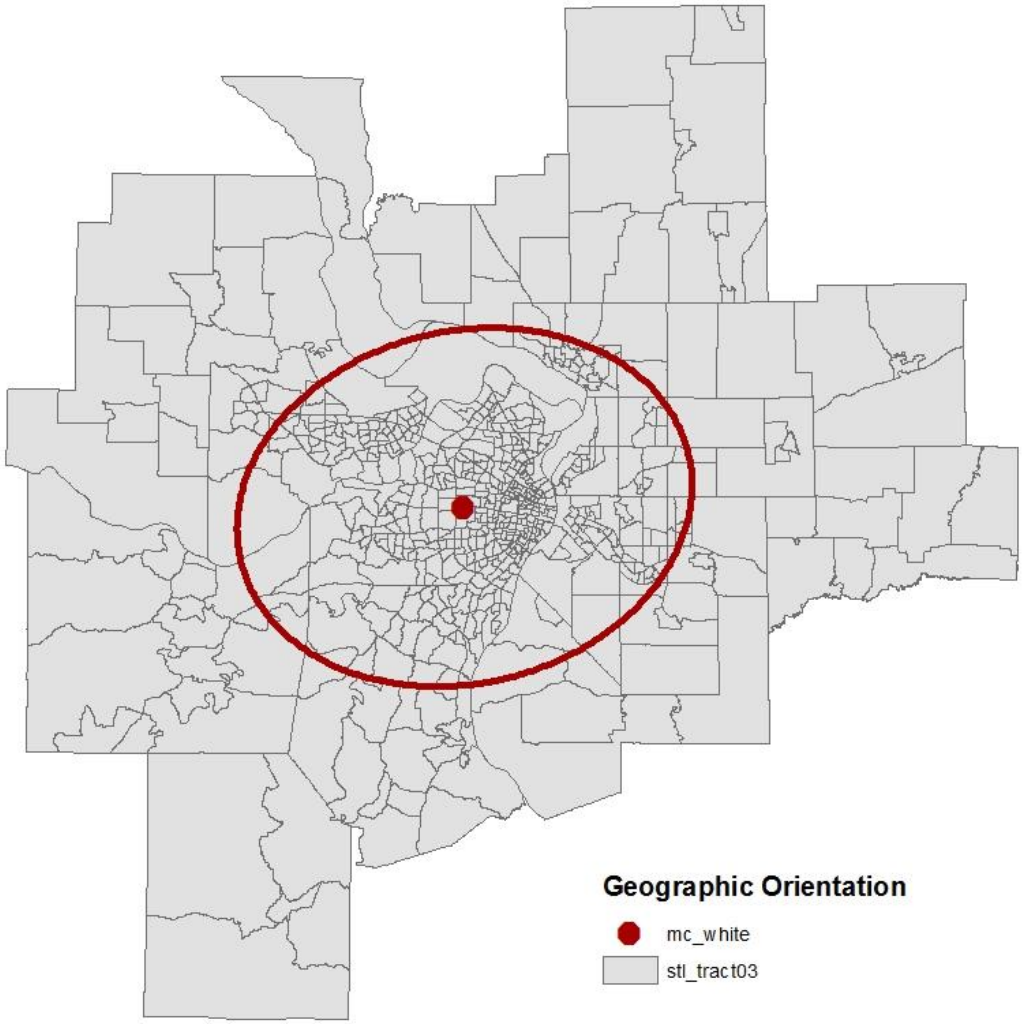
The standard deviational ellipse measures orientation and direction.

Typically, it provides a more accurate understanding of your data.

Compare multiple features.

1 standard deviational ellipse shows concentration

2 standard deviational ellipses will encompass most of your feature of interest. Note: helpful if you don't know the area boundaries.

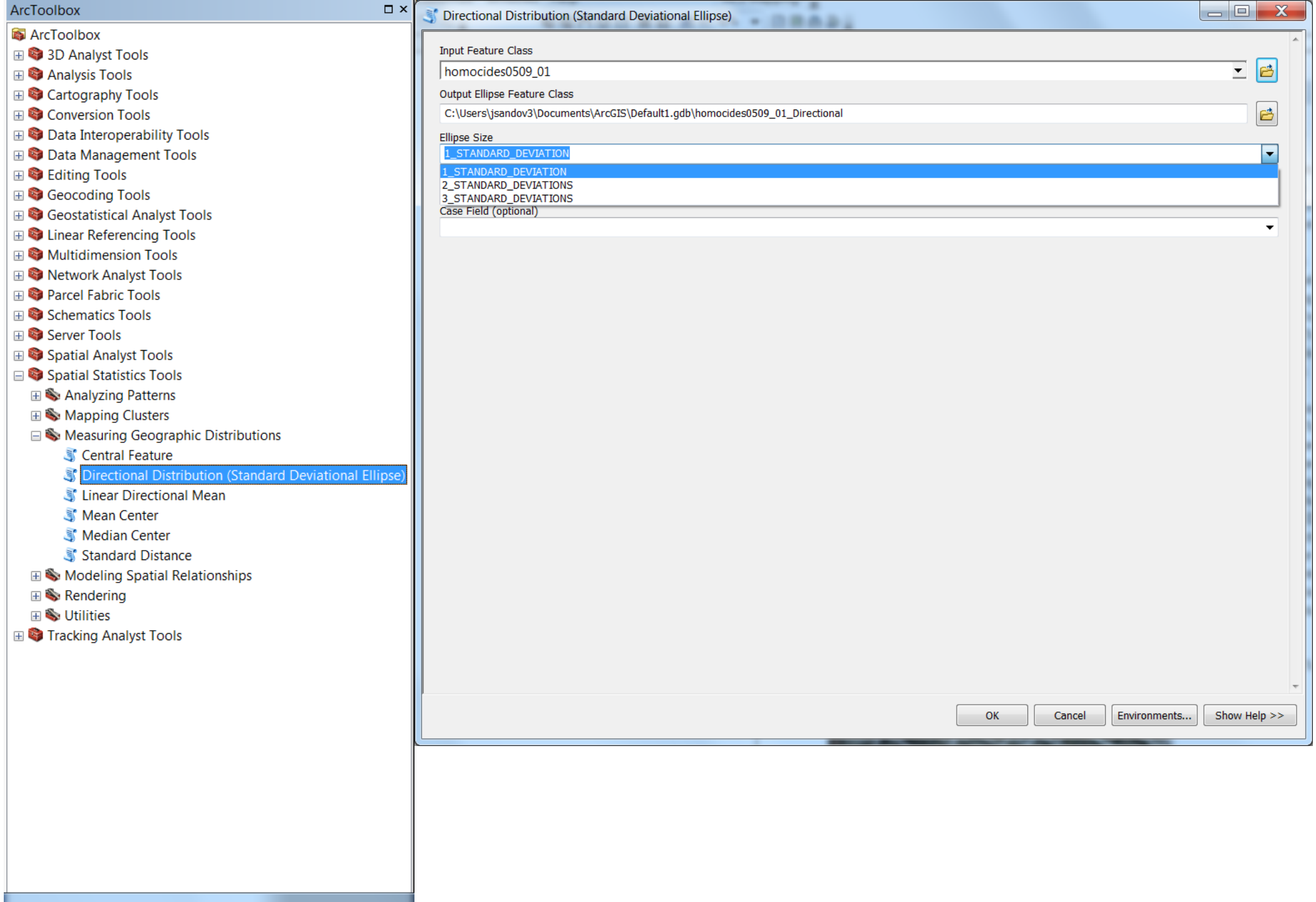


Standard Deviational Ellipse (SDE)

$$SD_X = \sqrt{\frac{\sum_i^n (X_i - \bar{X})^2}{n}}$$

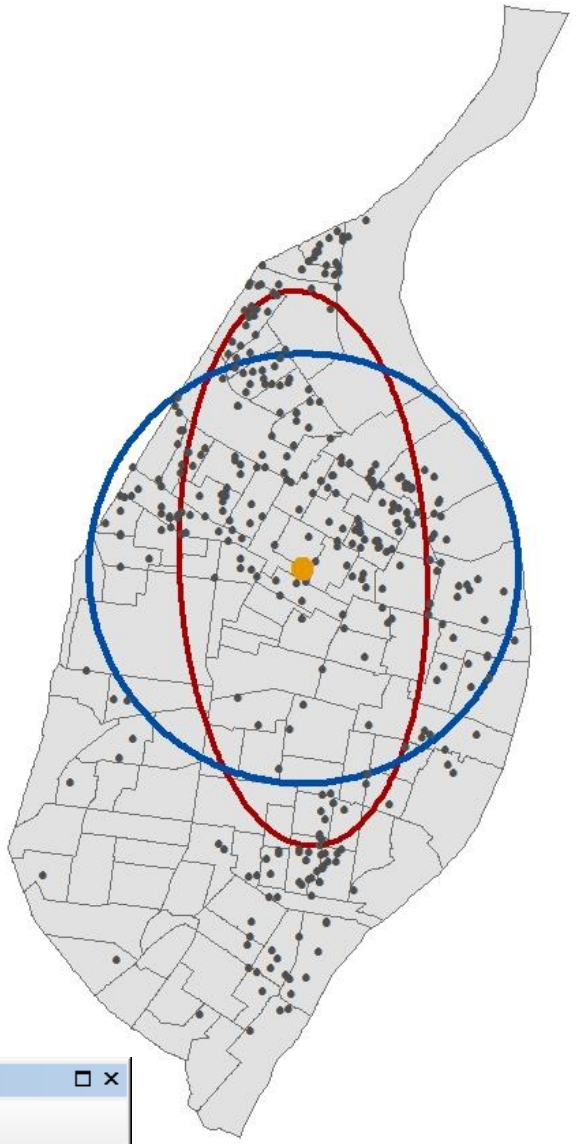
$$SD_Y = \sqrt{\frac{\sum_i^n (Y_i - \bar{Y})^2}{n}}$$

- SDE is defined by 3 parameters
  - Angle of rotation
    - Trignometric function. Tries to find the best fit of both axes to minimize the distance of the features to the axes.
  - Dispersion along major axis
    - The major axis defines the direction of maximum spread of the distribution
  - Dispersion along minor axis
    - The minor axis is perpendicular to it and defines the minimum spread
  - Note: since the standard deviation is measured in each direction from the mean center, the total length of each axis is twice its standard deviation.


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellipse)
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools
Directional Distribution (Standard Deviational Ellipse)
Input Feature Class
homocides0509_01
Output Ellipse Feature Class
C:\Users\jsandov3\Documents\ArcGIS\Default1.gdb\homocides0509_01_Directional
Ellipse Size
1_STANDARD_DEVIATION
1_STANDARD_DEVIATION
2_STANDARD_DEVIATIONS
3_STANDARD_DEVIATIONS
Case Field (optional)
OK
Cancel
Environments...
Show Help >>

Standard Distance & Directional Deviation

- homicides1012_MeanCenter
- homicides1012
- homicides1012_StandardDistance
- tl_2010_29510_tract10
- homicides1012_DirectionalDis



Table

homicides1012_DirectionalDis

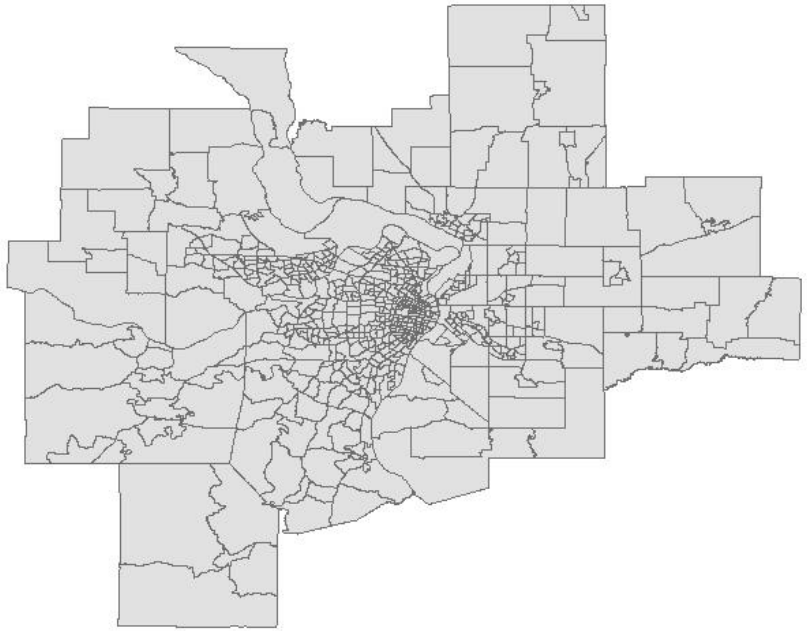
OBJECTID *	Shape *	Shape_Length	Shape_Area	CenterX	CenterY	XStdDist	YStdDist	Rotation
[redacted]	Polygon	31118.253969	61130753.745656	272673.374791	313048.000929	6605.792327	2946.006395	177.73699

1 (0 out of 1 Selected)

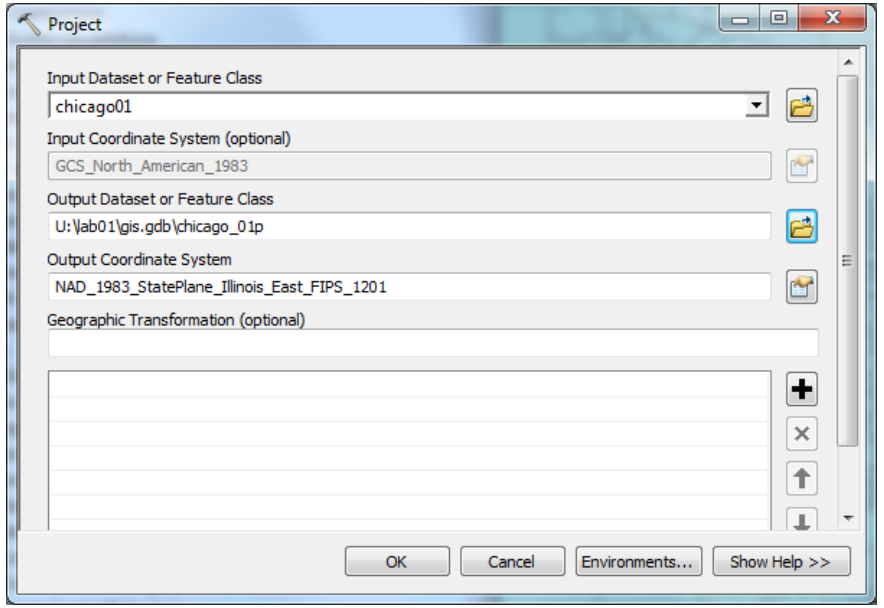
homicides1012_DirectionalDis

Polygons and Distribution

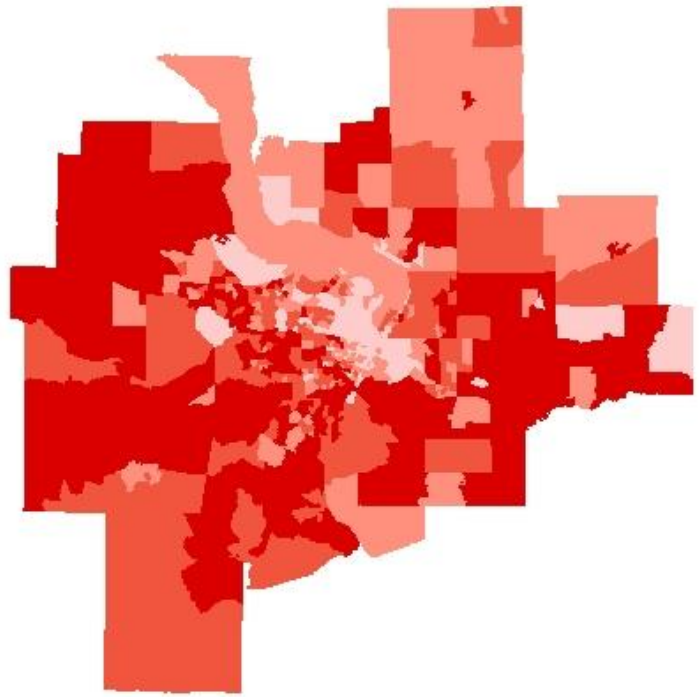
Let's work with Saint Louis



•Project shapefile NAD 83 Meters
•Arc Tool Box → Data Management Tools → Projections and Transformations → Project → select your file → Projected Coordinated System → State Plane → NAD 1983 (Meters) →Illinois East
Note:
NAD_1983_StatePlane_Missouri_East_FIPS_2401for Saint Louis
NAD_1983_StatePlane_Missouri_West_FIPS_2403 for Kansas City



Dispersion of Whites and Blacks

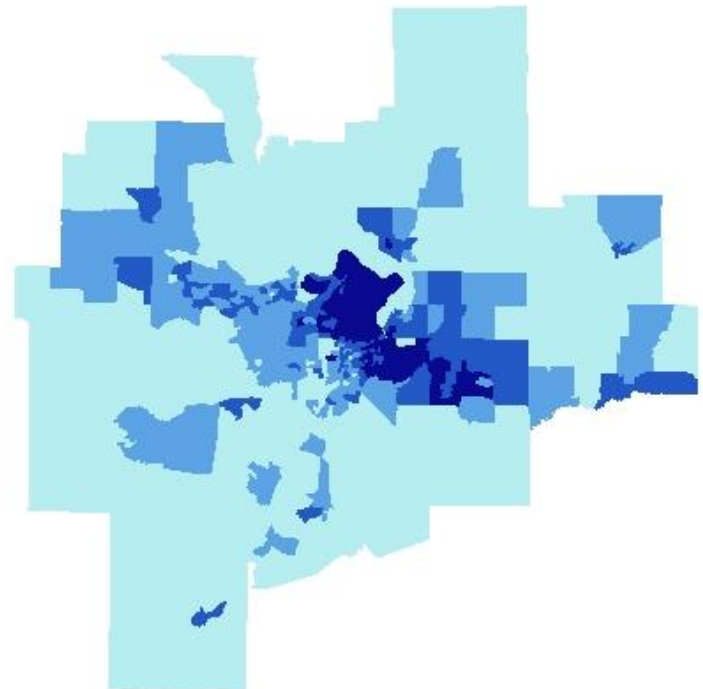


Distribution in Numbers

stl_tract02

V30 - White Non-Hispanic

- 1 - 1869
- 1870 - 3362
- 3363 - 4975
- 4976 - 12147



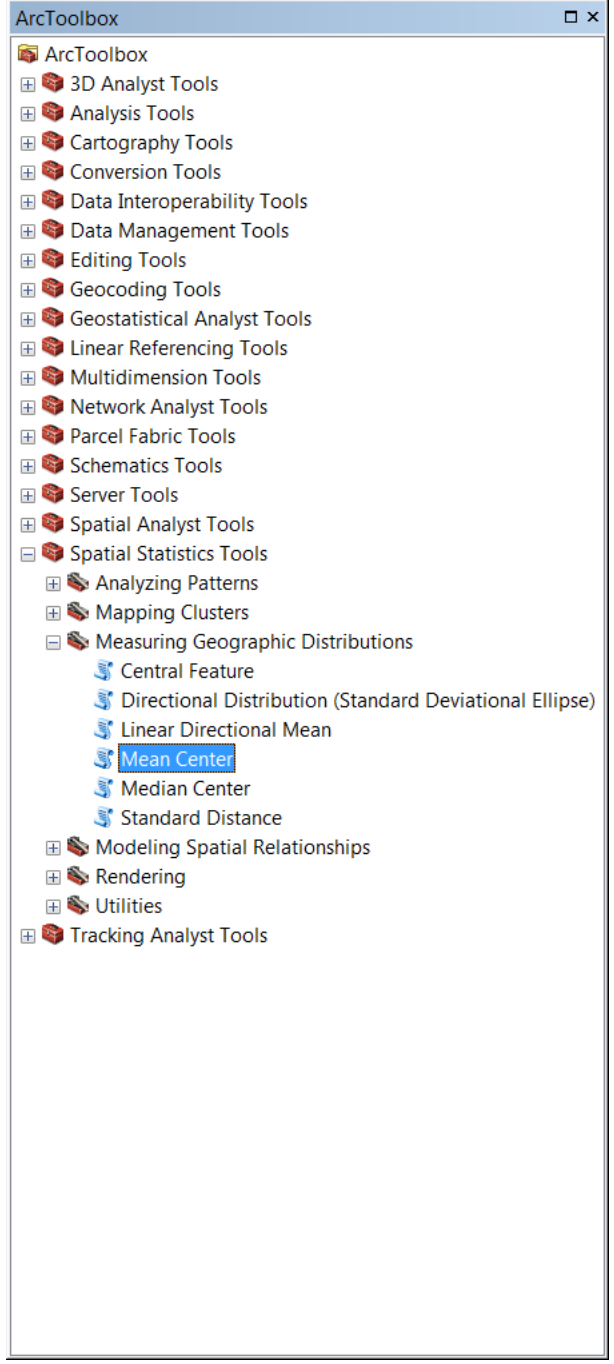
Distribution in Numbers

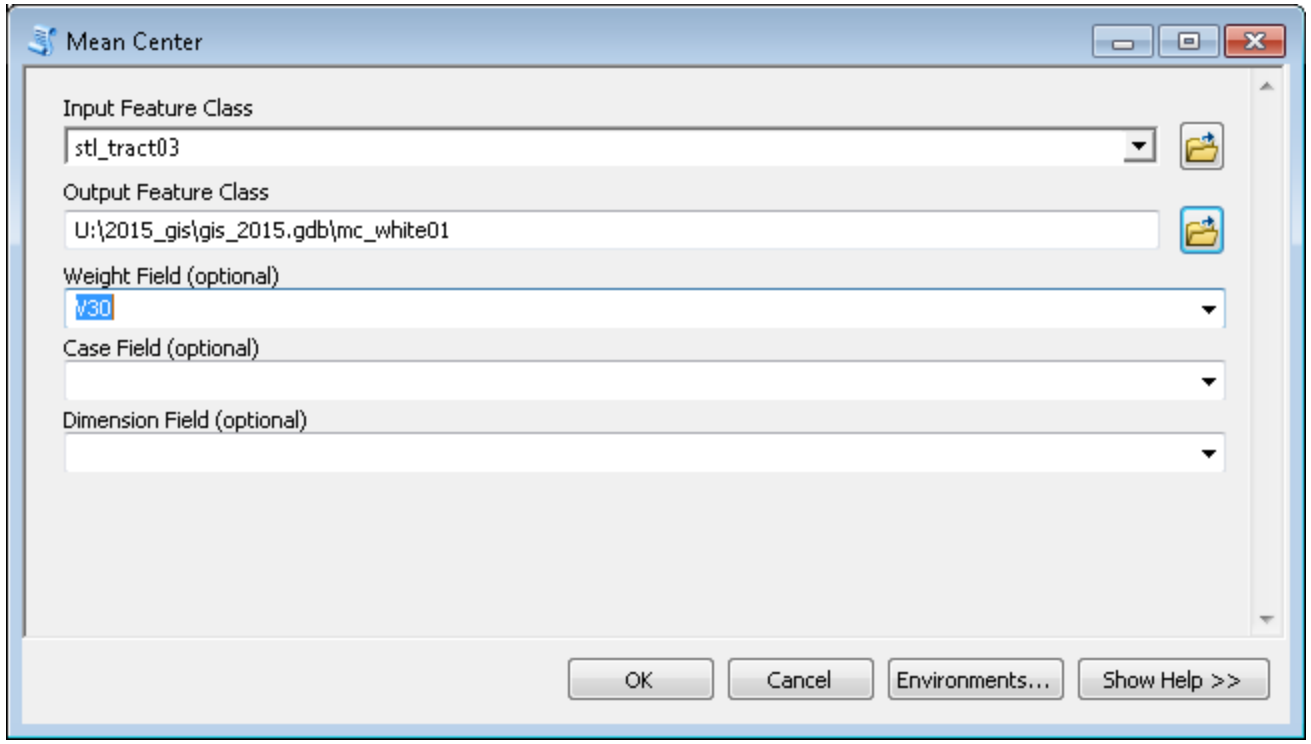
stl_tract02

V31-Black Non-Hispanic

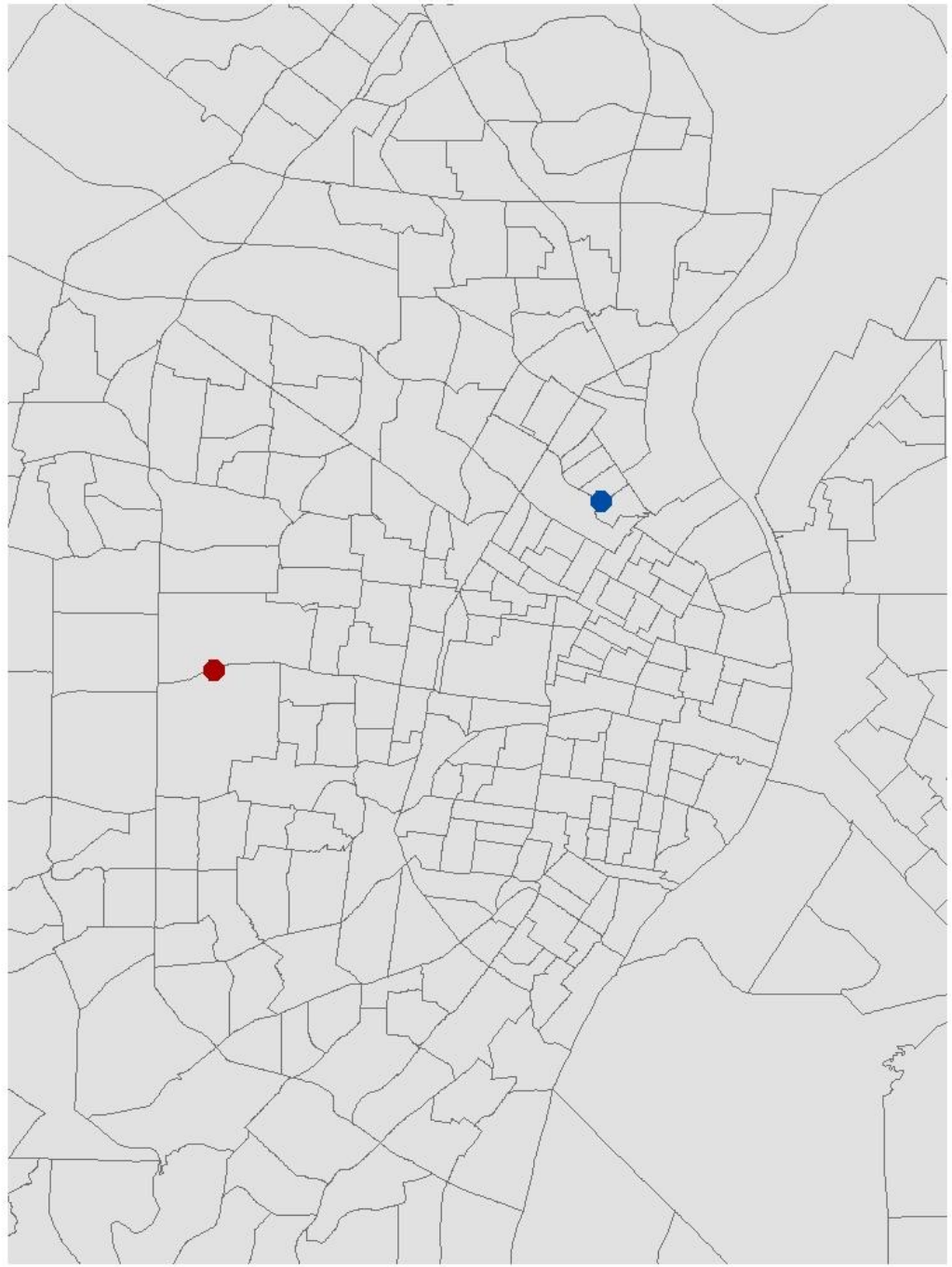
- 0 - 60
- 61 - 215
- 216 - 1094
- 1095 - 7659

Mean Center

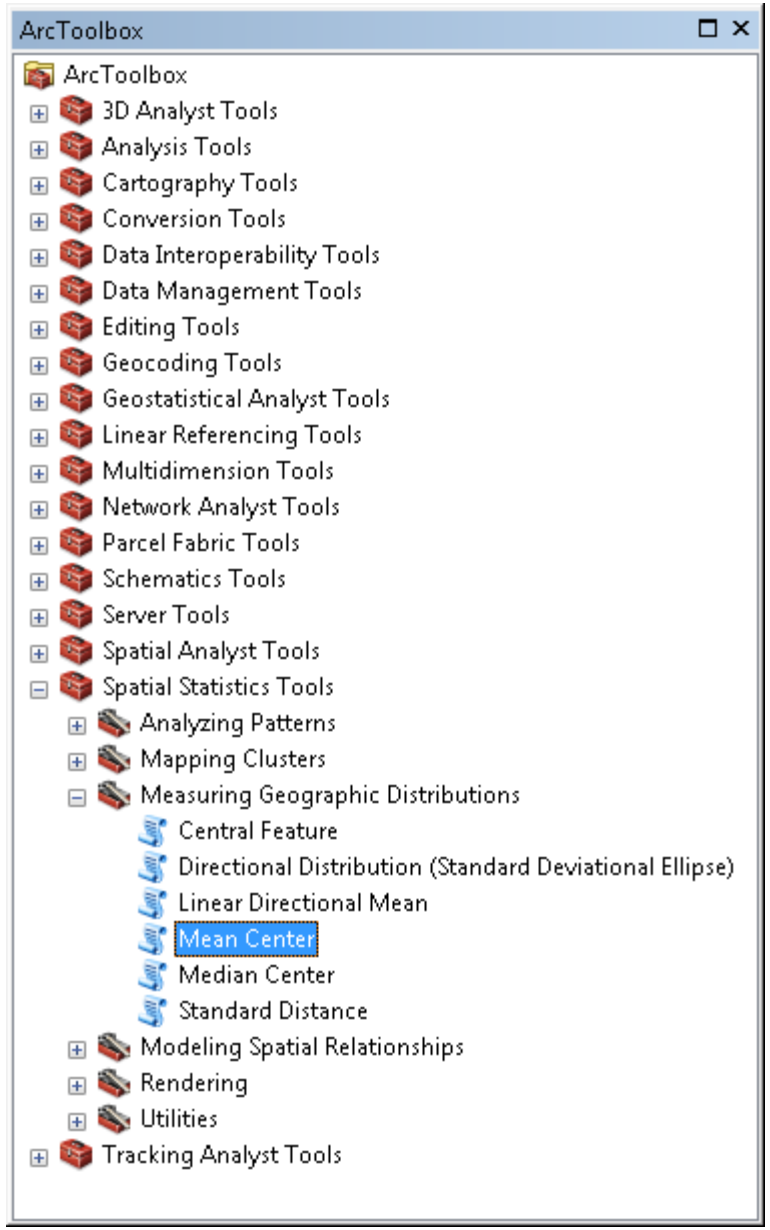

ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellipse)
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools

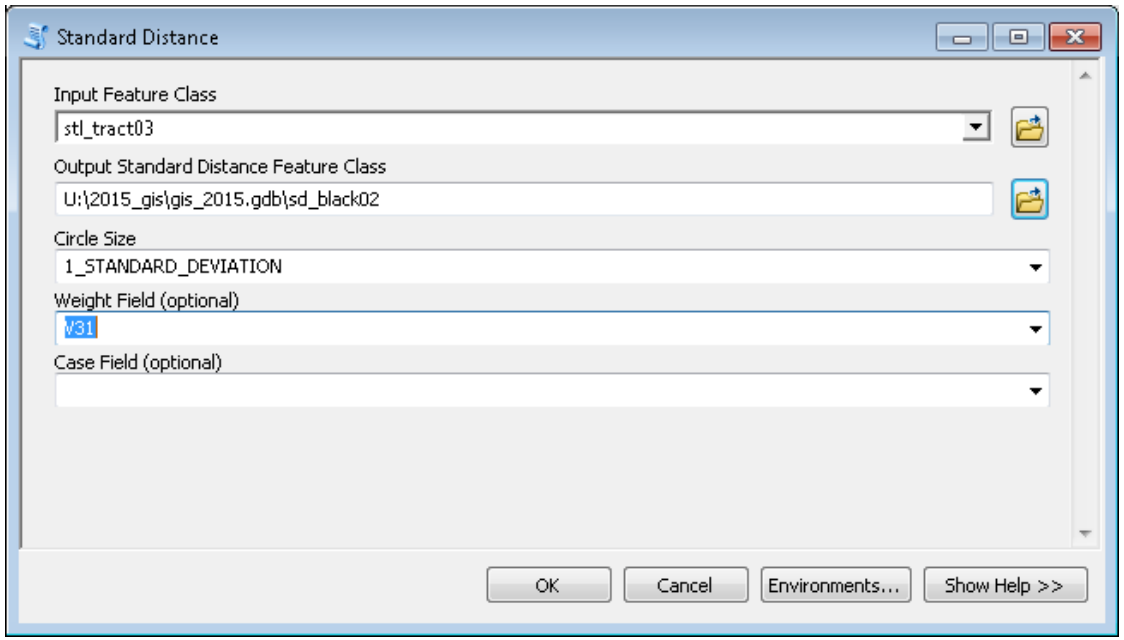

Mean Center
Input Feature Class
stl_tract03
Output Feature Class
U:\2015_gis\gis_2015.gdb\mc_white01
Weight Field (optional)
V30
Case Field (optional)
Dimension Field (optional)
OK
Cancel
Environments...
Show Help >>

Mean Center



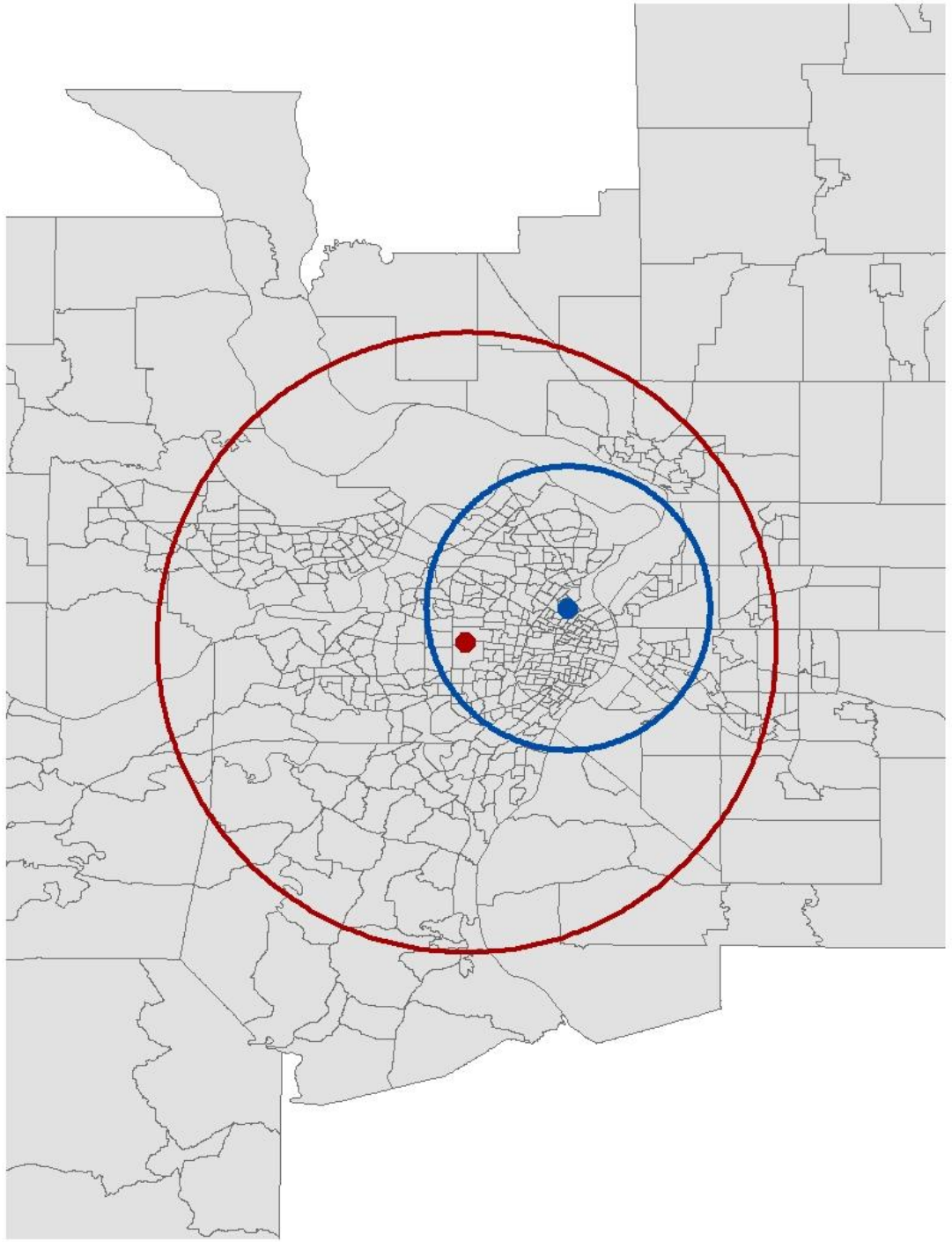
Standard Distance


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellipse)
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools

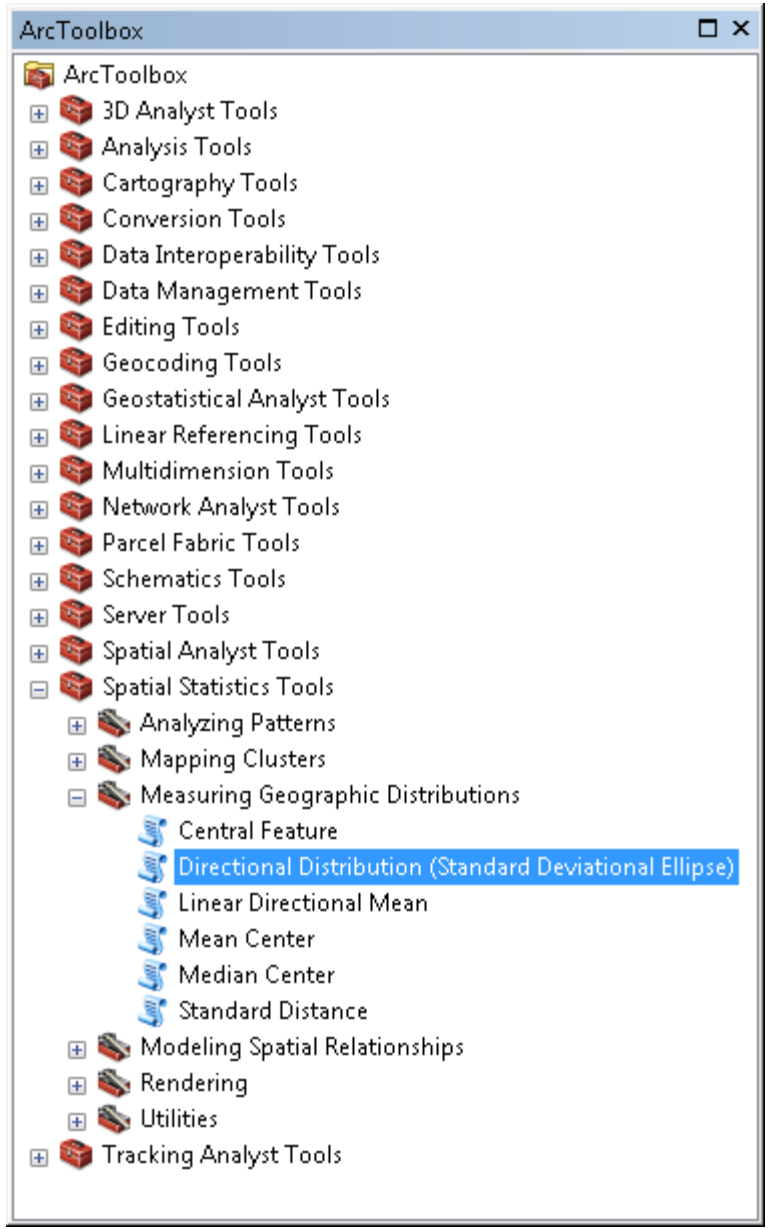

Standard Distance
Input Feature Class
stl_tract03
Output Standard Distance Feature Class
U:\2015_gis\gis_2015.gdb\sd_black02
Circle Size
1_STANDARD_DEVIATION
Weight Field (optional)
V31
Case Field (optional)
OK
Cancel
Environments...
Show Help >>

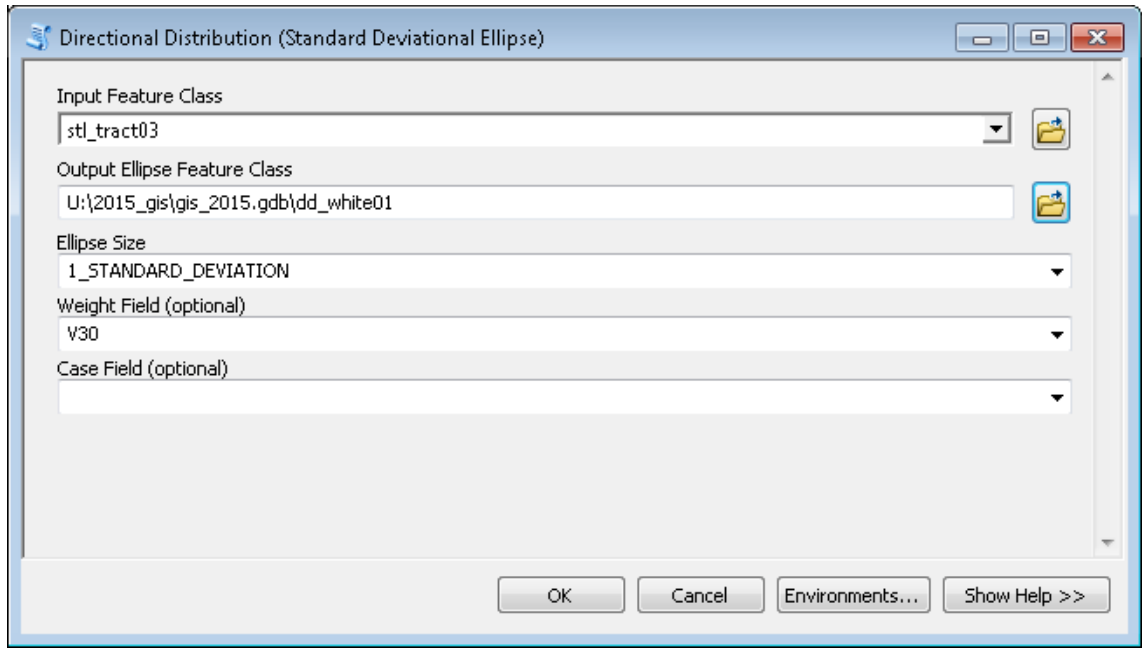
Standard Distance
Mean Center

- mc_black
- mc_white
- sd_black
- sd_white
- stl_tract02

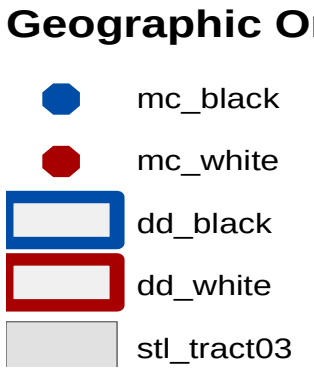


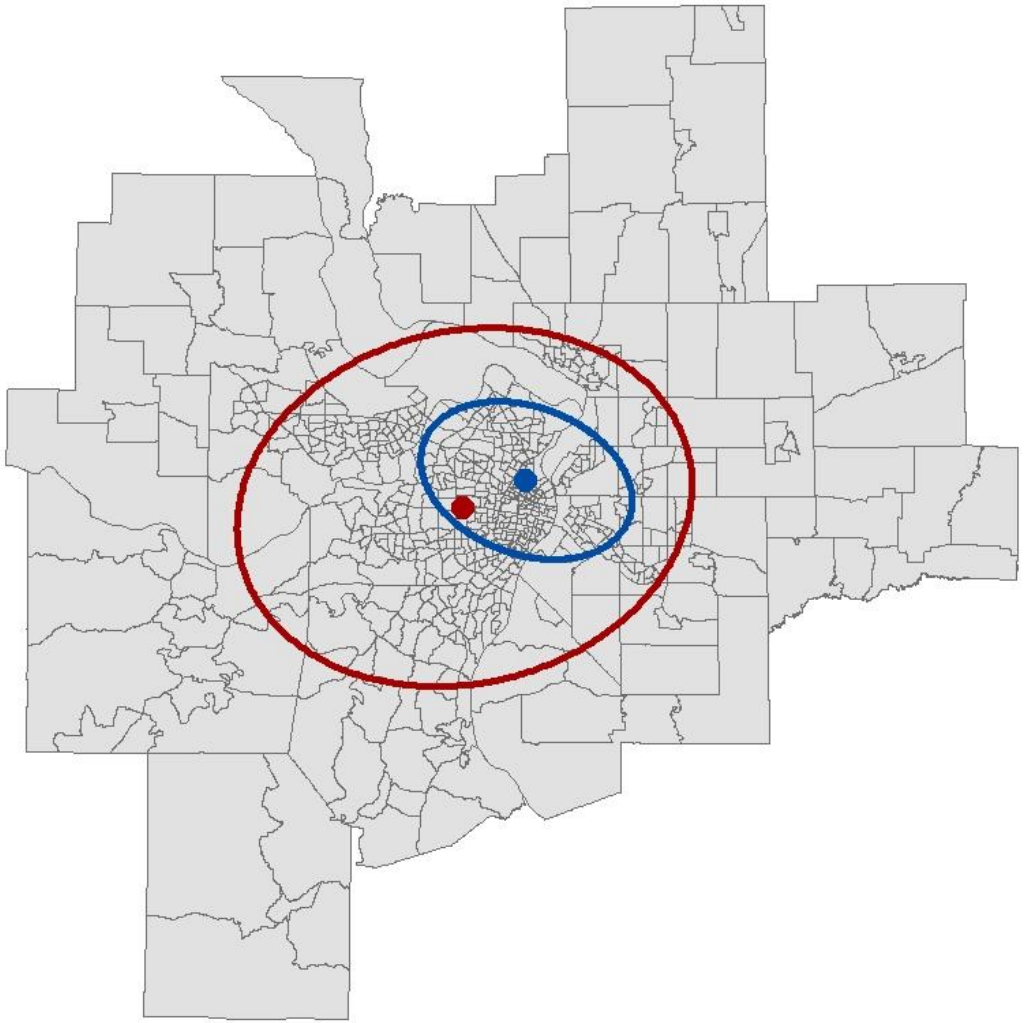
Geographic Orientation


ArcToolbox
ArcToolbox
3D Analyst Tools
Analysis Tools
Cartography Tools
Conversion Tools
Data Interoperability Tools
Data Management Tools
Editing Tools
Geocoding Tools
Geostatistical Analyst Tools
Linear Referencing Tools
Multidimension Tools
Network Analyst Tools
Parcel Fabric Tools
Schematics Tools
Server Tools
Spatial Analyst Tools
Spatial Statistics Tools
Analyzing Patterns
Mapping Clusters
Measuring Geographic Distributions
Central Feature
Directional Distribution (Standard Deviational Ellipse)
Linear Directional Mean
Mean Center
Median Center
Standard Distance
Modeling Spatial Relationships
Rendering
Utilities
Tracking Analyst Tools


Directional Distribution (Standard Deviational Ellipse)
Input Feature Class
stl_tract03
Output Ellipse Feature Class
U:\2015_gis\gis_2015.gdb\dd_white01
Ellipse Size
1_STANDARD_DEVIATION
Weight Field (optional)
V30
Case Field (optional)
OK
Cancel
Environments...
Show Help >>

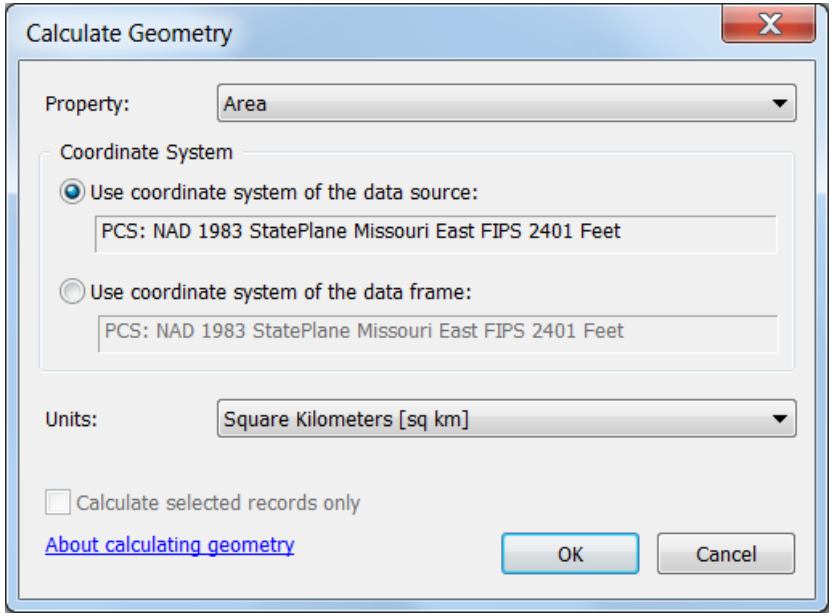
Geographic Orientation

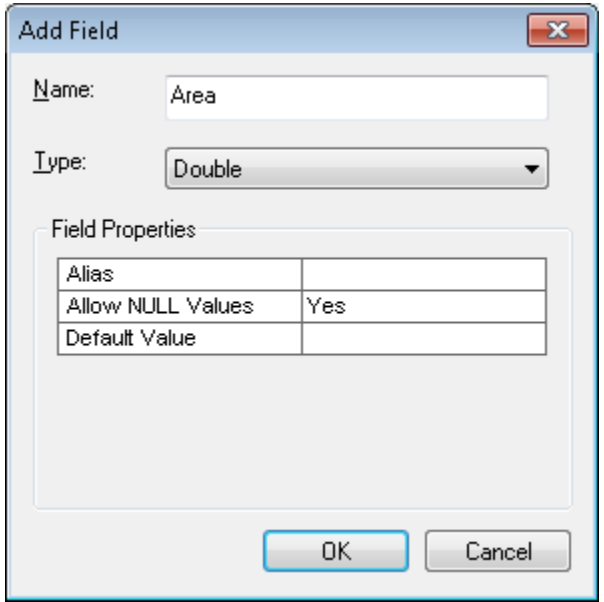




Computing the Area of the Ellipses

- Once you create ellipses, you need to create a new variable called “Area.” Then go to Calculate Geometry and select square kilometers.





Review your Attribute Table

The White Population in Saint Louis

Table

dd_white

OBJECTID *	Shape *	Shape_Length	Shape_Area	CenterX	CenterY	XStdDist	YStdDist	Rotation	Area
1	Polygon	251172.782695	4887037997.608633	259991.614537	311432.36744	34484.695695	45112.192906	76.591806	4887.037998

The Black Population in Saint Louis

Table

dd_black

OBJECTID *	Shape *	Shape_Length	Shape_Area	CenterX	CenterY	XStdDist	YStdDist	Rotation	Area
1	Polygon	113824.687371	972178244.204456	272116.942565	316684.859691	21470.617437	14413.81939	110.696539	972.178244

Create a Summary Table

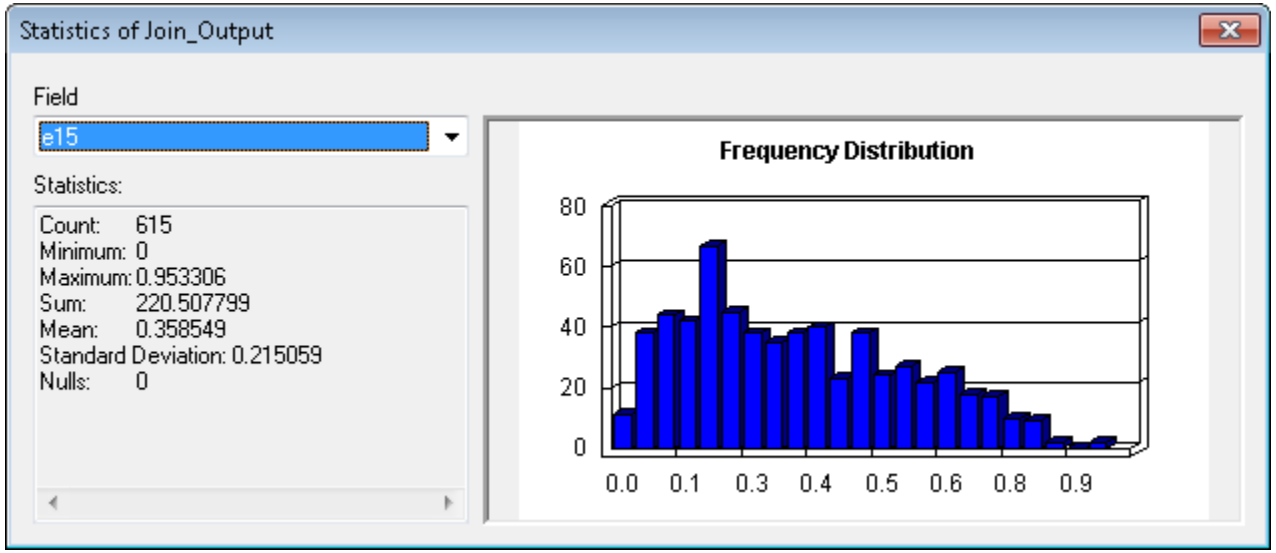
Table 1. Descriptive Spatial Statistics for the Saint Louis MSA

Variable	Name	White	Black	Latino
Mean Center	Mean Center - X Coordinate	259991.6	272116.9	264673.9
	Mean Center - Y Coordinate	311432.4	316684.9	313349.2
Directional Distribution	Angle Rotation	76.6	110.7	93.7
	Y Axis length (km)	45.1	14.4	24.2
	X Axis length (km)	34.5	21.5	38.5
	Area of Ellipse (km^2)	4887.0	972.2	2923.3
Standard Distance	Standard Distance (km)	40.2	18.3	32.1
	Area (km^2)	5064.5	1050.4	3245.3

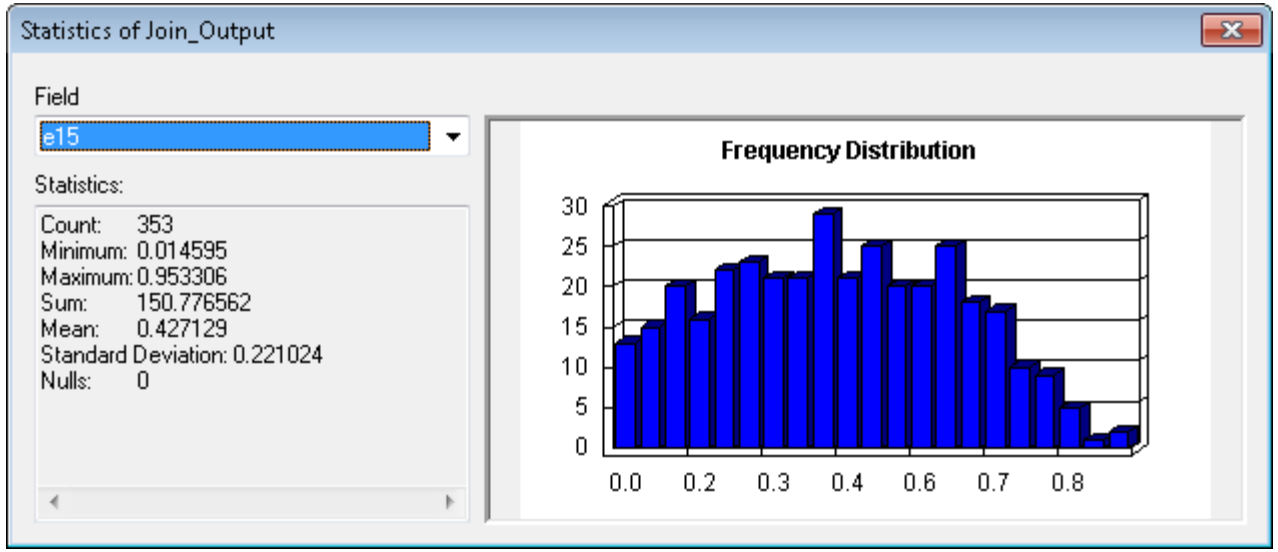
Source: U.S. Census Bureau, 2010 Census, Table Created by Author

Interval-Ratio Variables

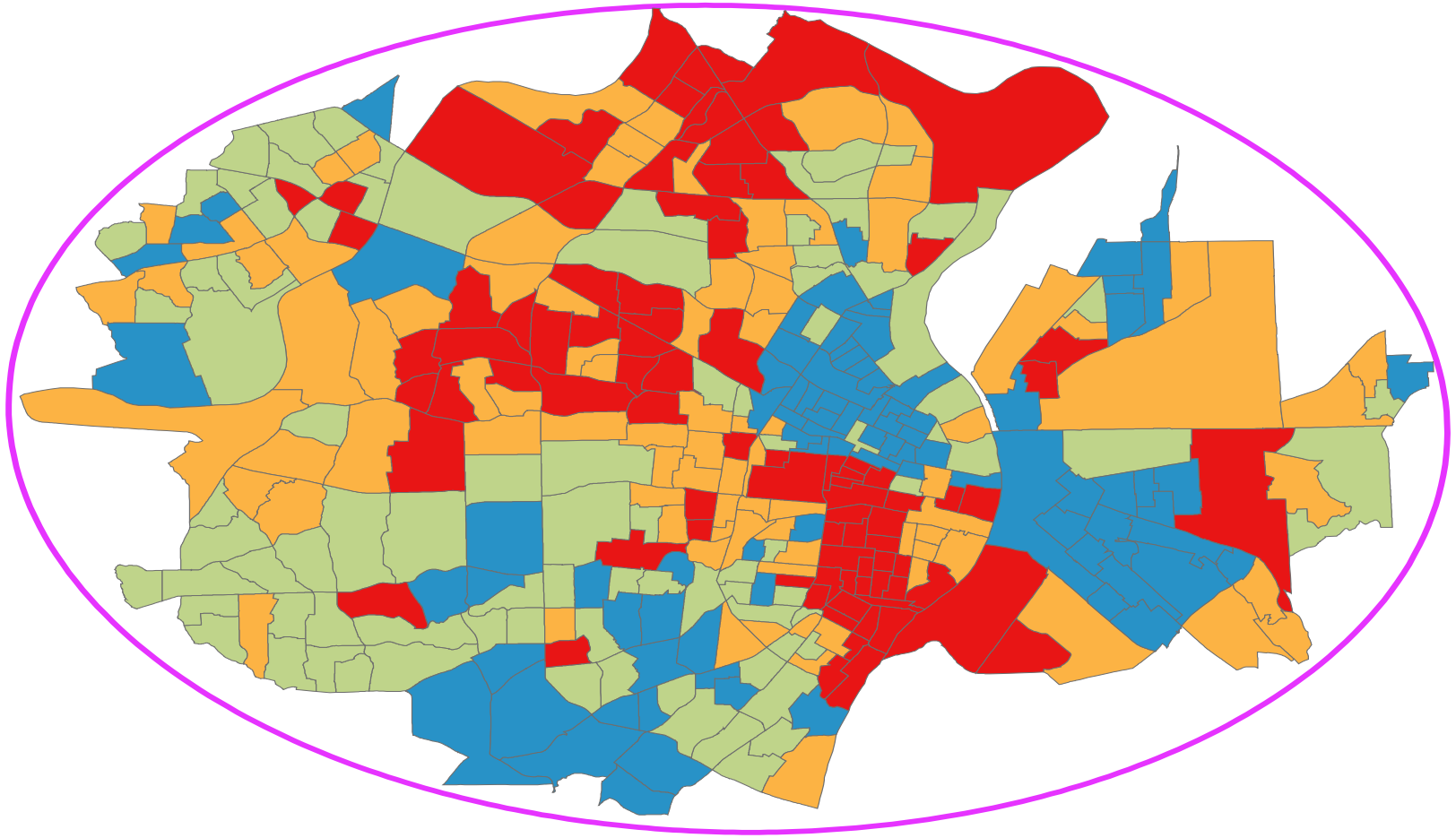
For example, the total diversity score for STL was 220 points



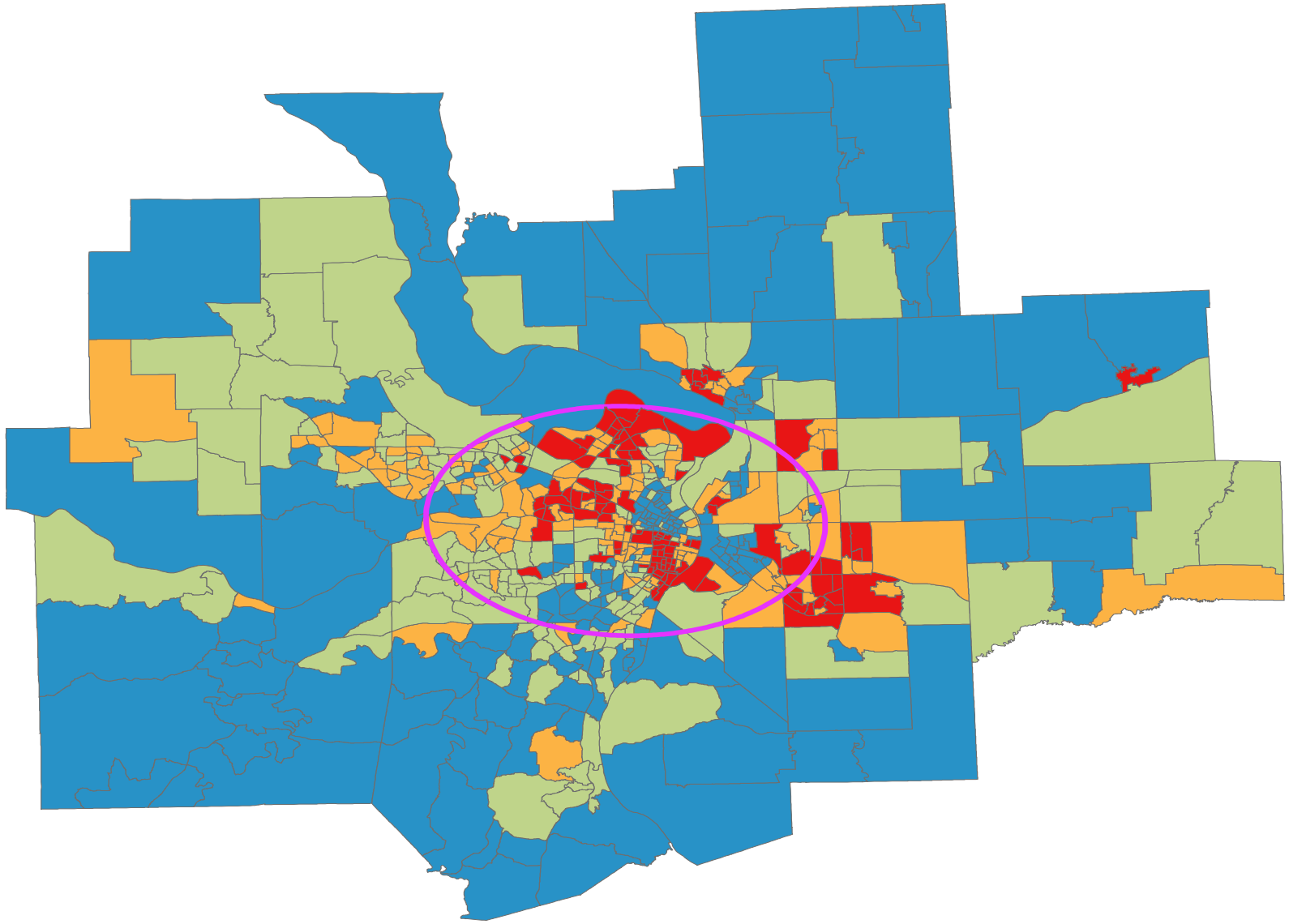
1 SDE contains 150 of the 220 points or 68%



So, tracts have high diversity will have the most weight in calculating the SDE



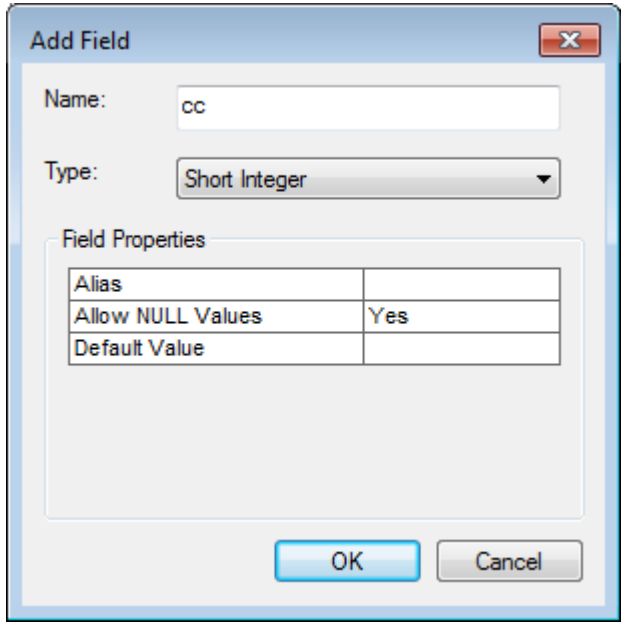
Thus the majority of highly diverse tracts should be in SDE

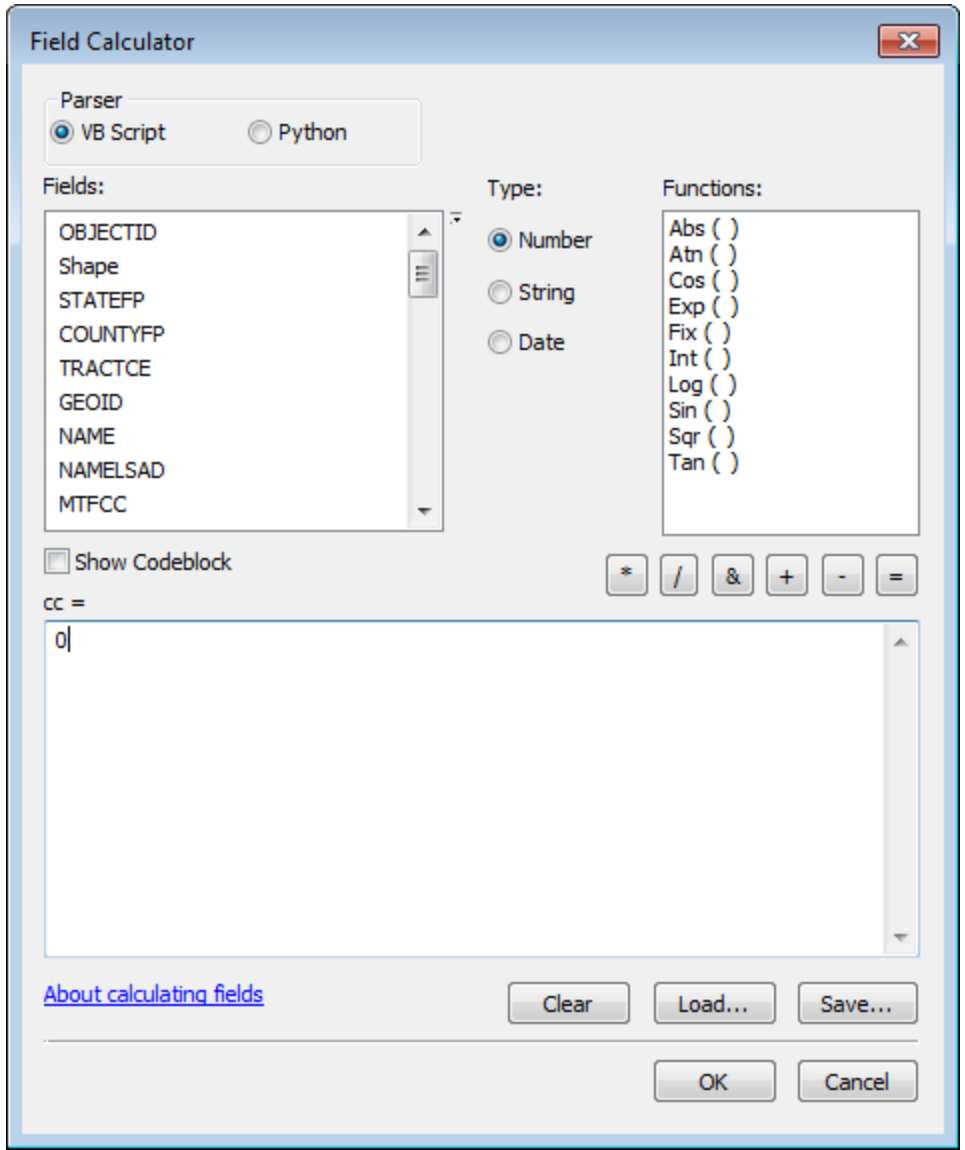


Creating Spatial Variables

Dummy Variables

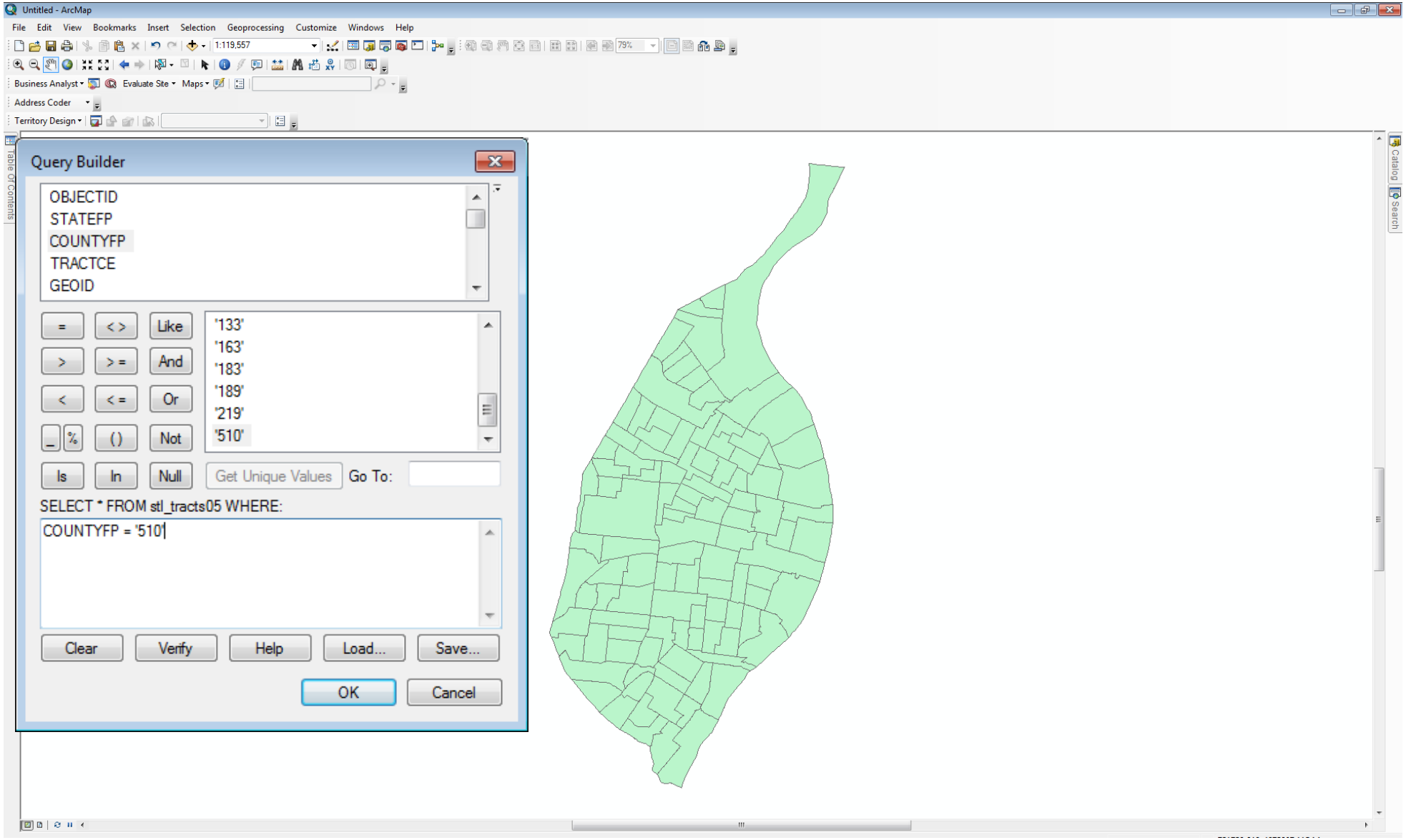
1. Creating a Categorical Variable – Dummy 1=yes and 0=no
2. Open attribute table – add a field – cc
3. Open field calculary for cc and input 0





4. Open Query Builder

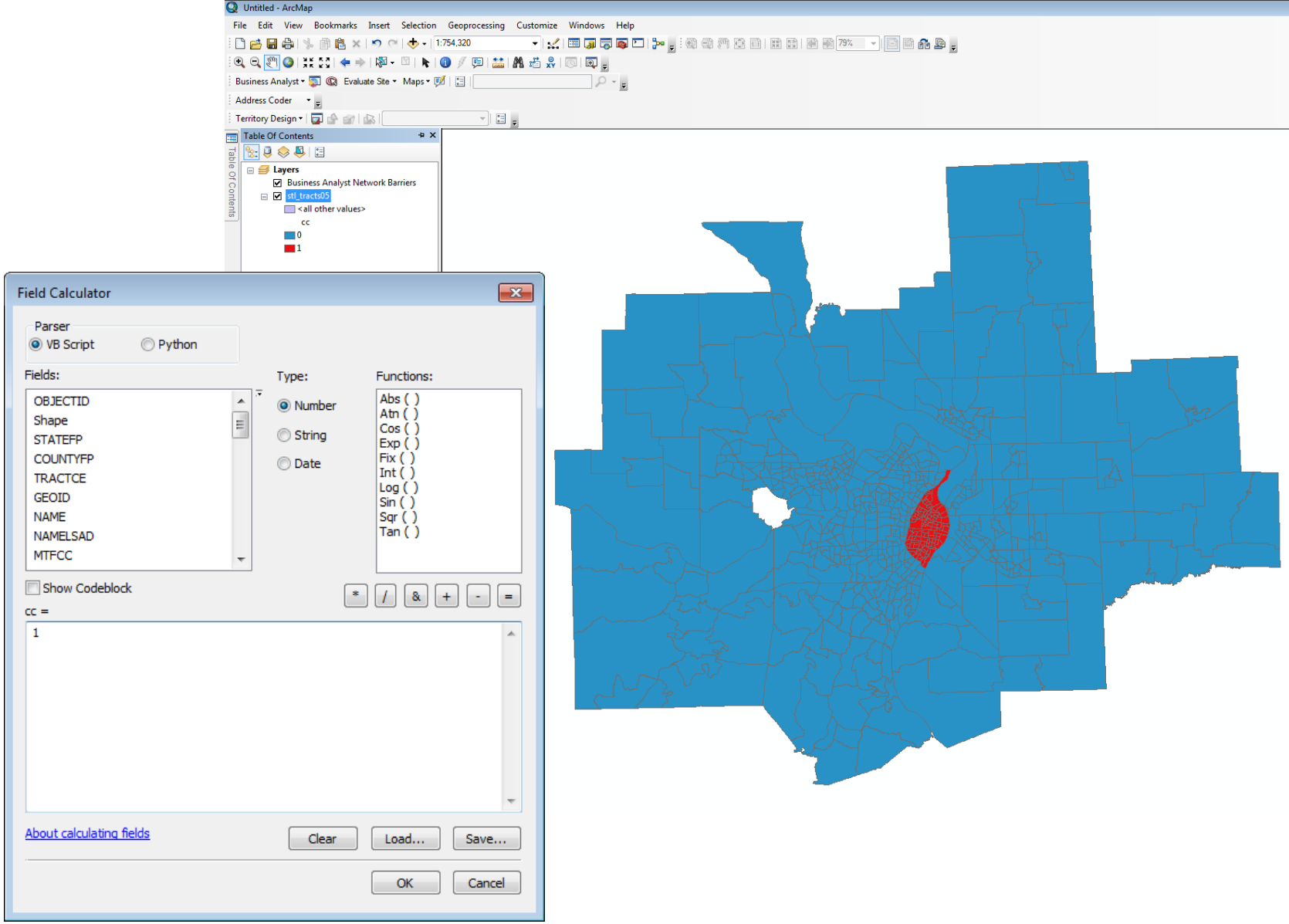
5. Select COUNTYFP=510, which is St. Louis City



6. Open Field Calculator for cc

7. Input the value 1 for Yes.

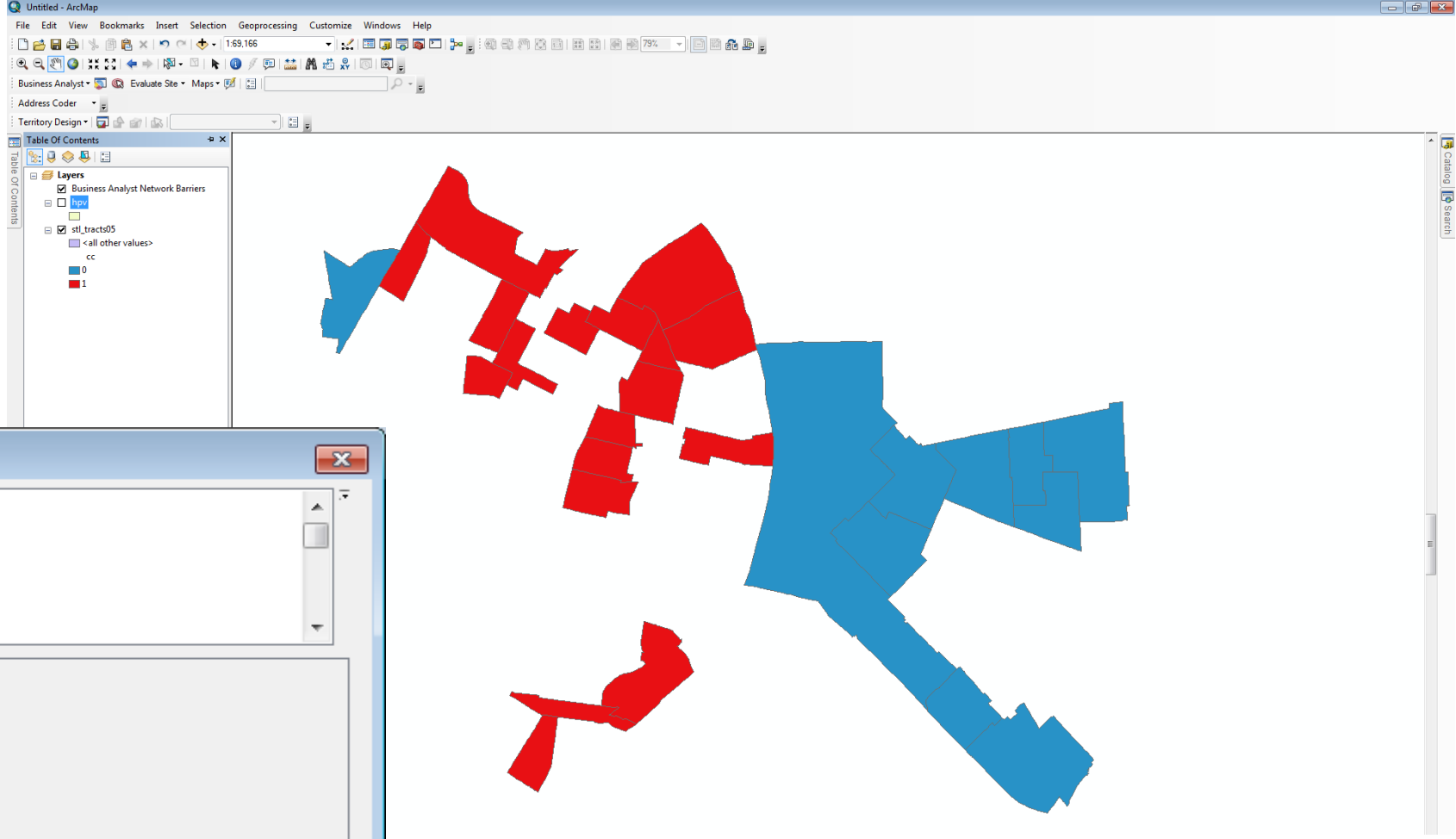
8. Make a map that is red for the city tracts and blue for non-city tracts

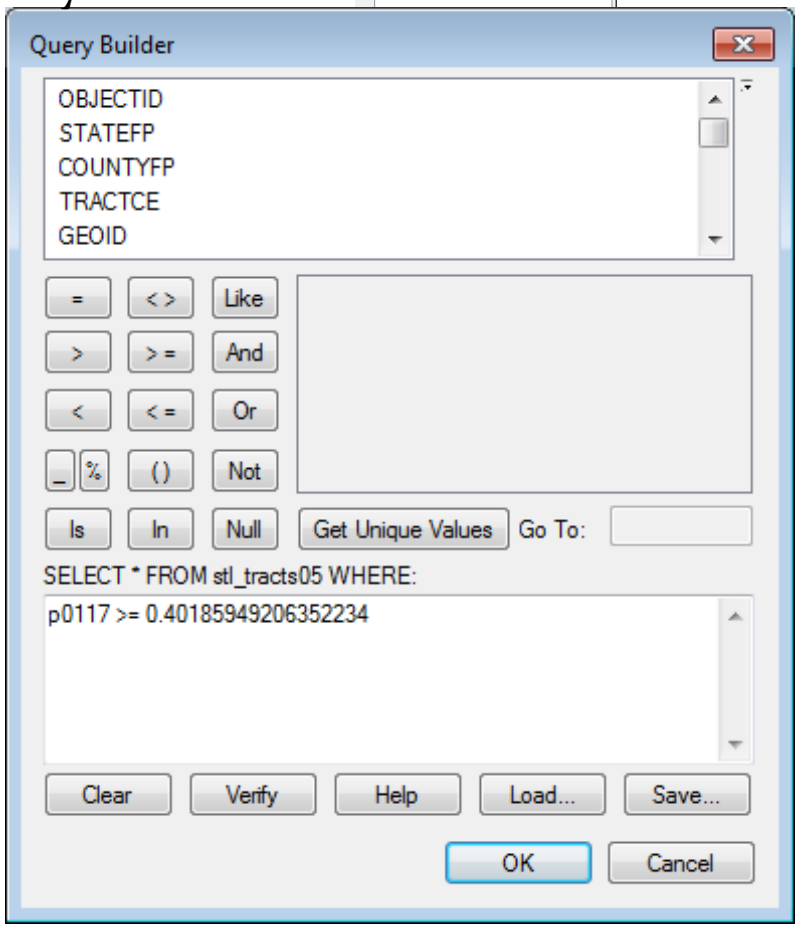


High Poverty Neighborhoods

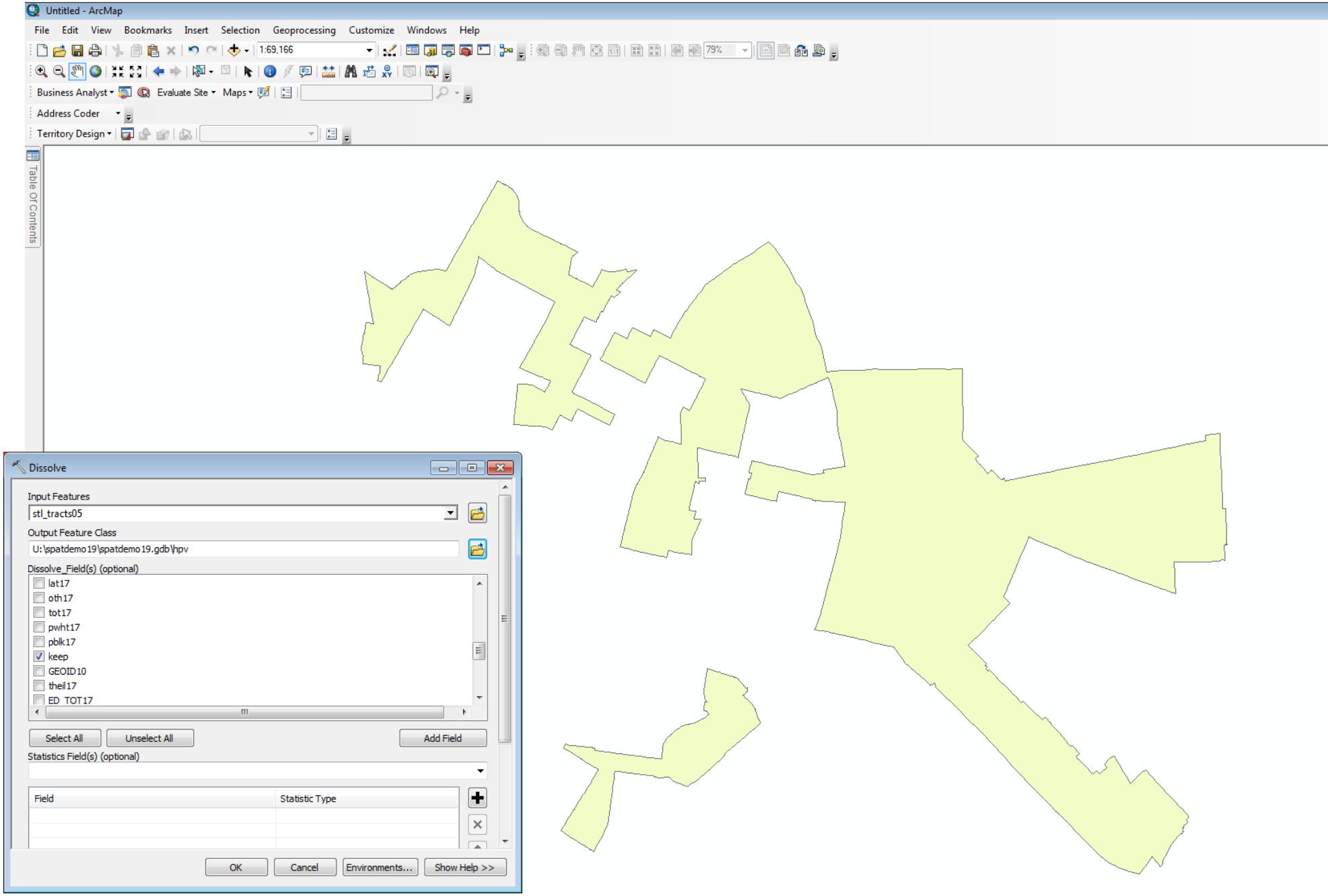
1. Creating a Categorical Variable – Dummy 1=yes and 0=no

2. Open field calculary for P0117>=.40





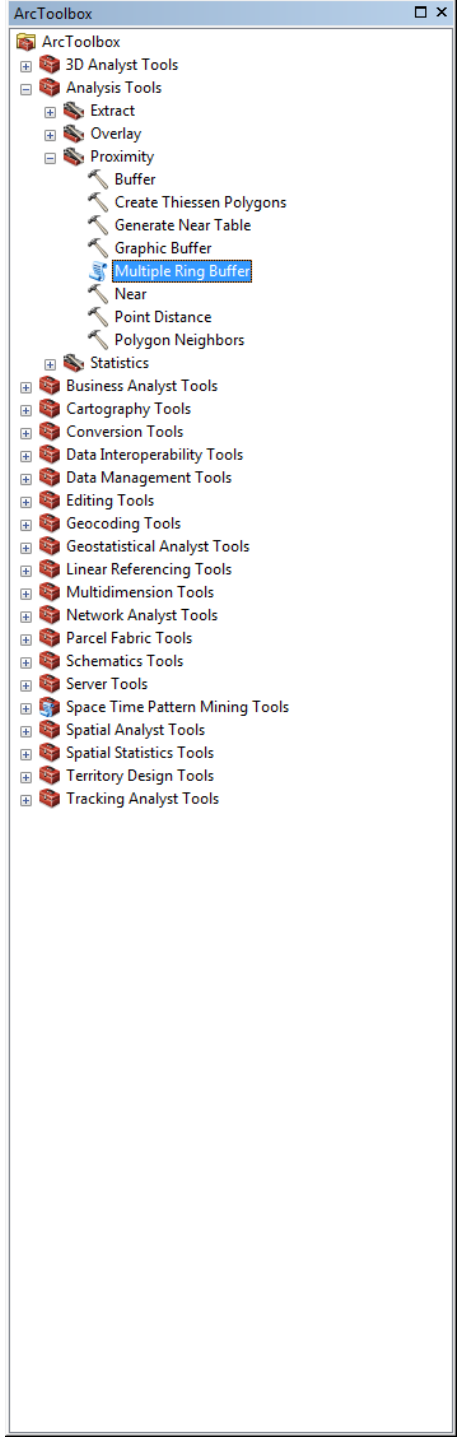
3. Dissolve based on common variable, e.g., keep

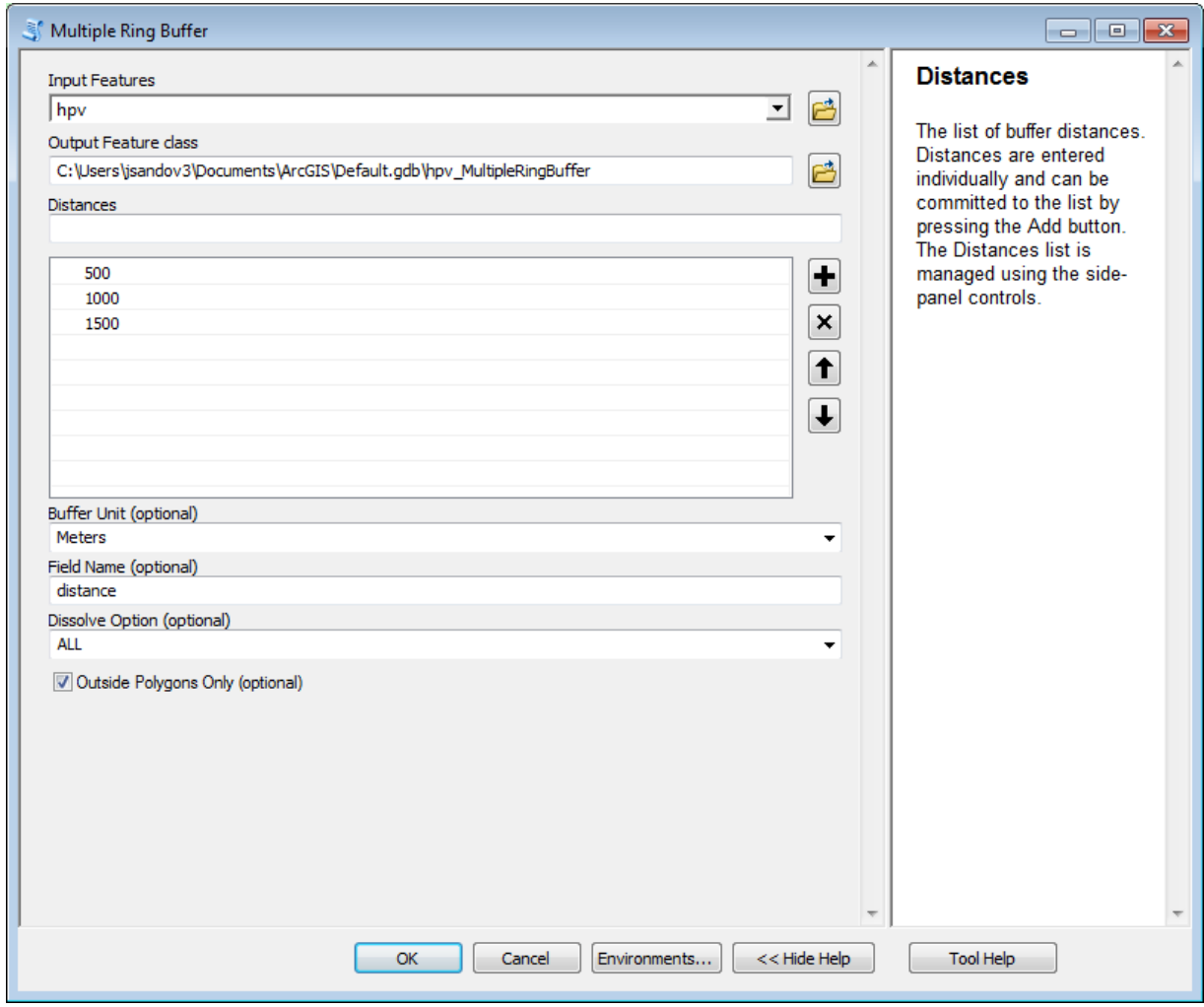


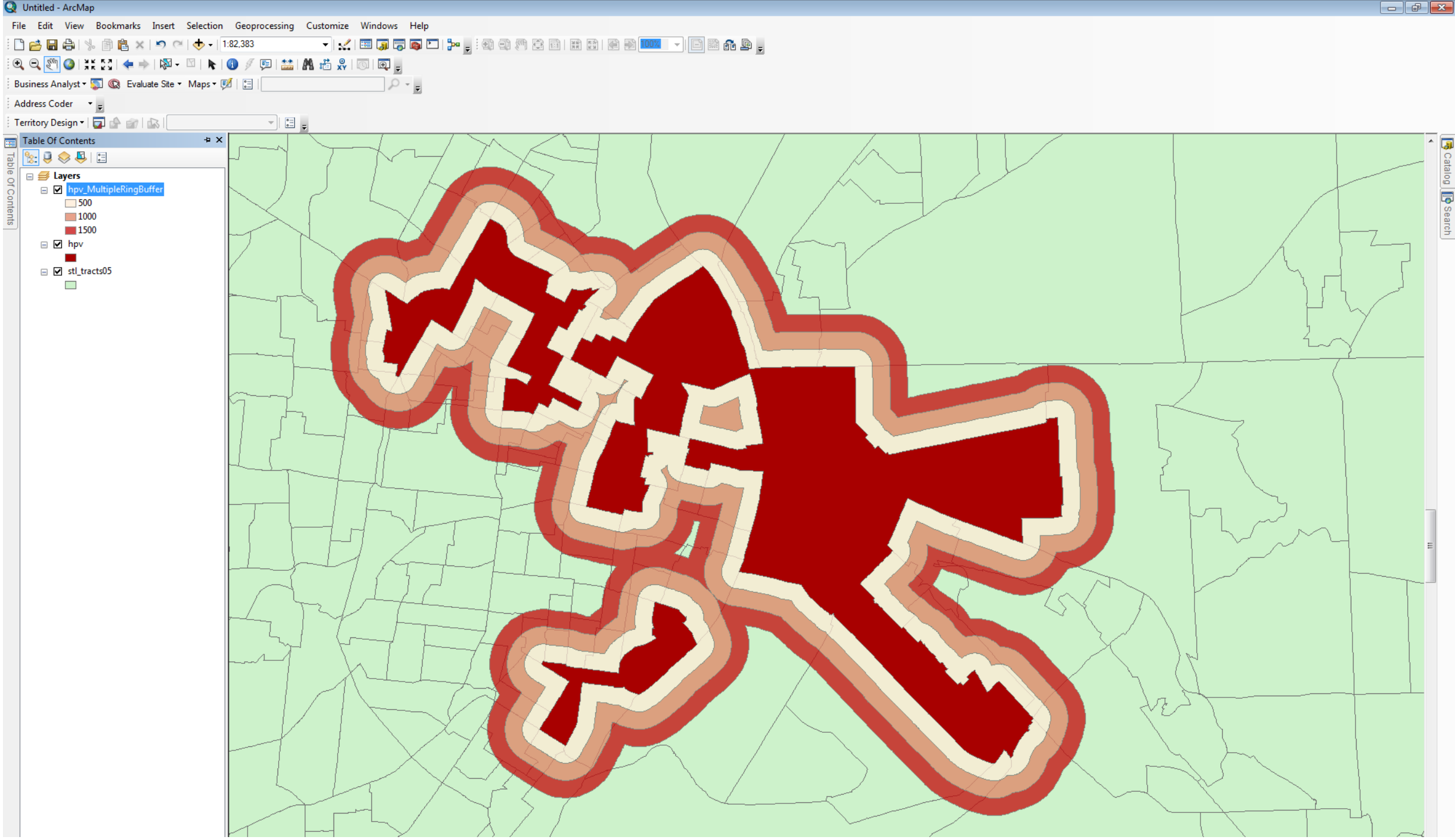
Creating Spatial Variables

Ordinal Variables

1. Open up Analysis Tools
2. Open up Proximity
3. Select Multiple Ring Buffer
4. In our example, we will make a
   500 meter buffer
   1000 meter buffer
   1500 meter buffer

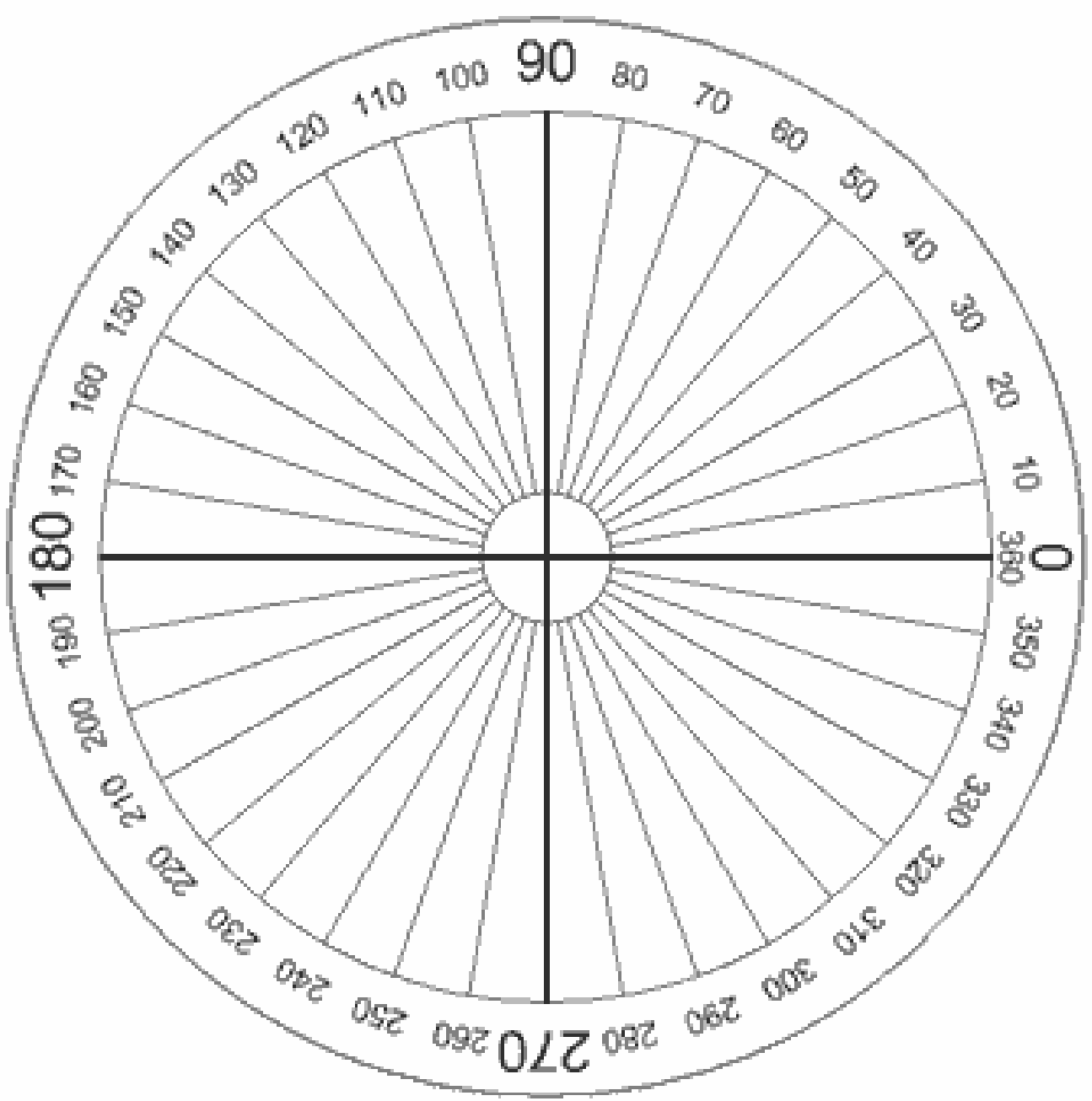





Untitled - ArcMap
File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help
1:82,383
Business Analyst Evaluate Site Maps
Address Coder
Territory Design
Table Of Contents
Layers
hpv_MultipleRingBuffer
500
1000
1500
hpv
stl_tracts05
Table Of Contents
Catalog
Search

Lab

- Compute Basic Spatial Statistics
- Compute Area
- Make Summary Table


90
80
70
60
50
40
30
20
10
0
360
350
340
330
320
310
300
290
280
270
260
250
240
230
220
210
200
190
180
170
160
150
140
130
120
110
100