Advanced apportionment settings

Various methods are provided to more accurately retrieve data. This is important when a trade area cuts across a demographic layer. Traditional selection methods use an apportionment method based on area alone. This process can compromise accuracy when equally distributed data across geographies is unavailable. In most cases, population or housing is distributed in irregular fashion across geographic boundaries, such as ZIP Codes, counties, or provinces.

Business Analyst offers the following data apportionment methods. You have the ability to modify these settings to reach your desired data retrieval approach.

- Block Apportionment: Used as the most accurate way to apportion and extract data at any level; however, processing time is increased because each block point is reviewed for demographic percentage weighting.
- Cascading Centroid: Used for faster retrieval of data, usually on larger geographic areas. Only the geographic centroids for the levels you choose are returned in your analysis. These centroids contain the exact same demographic attributes as their corresponding polygons.
- Hybrid: A merging of the first two that is used for high accuracy at lower levels (Block Apportionment) and faster processing at larger geographic levels (Cascading Centroid). The use of Cascading Centroid instead of Block Apportionment at the larger levels returns minimal statistical differences; thus, the overall accuracy is not compromised using this combination approach.

BDS Performance Indexes

Business Analyst allows you to create custom BDS performance indexes that are optimized for your custom data. Performance indexes allow you to aggregate data and create summary reports efficiently. These performance indexes can be created during Custom Data Setup or you can use the Build Index button, at any time, in Advanced Tolerance Settings. If the Build Index button is active, no indexes currently exist for your custom layer. Custom BDS performance indexes can only be used with the standard BDS geographic levels as a base (block groups, census tracts, ZIP Codes, and so forth). Each performance index is copied to the location where your custom BDS layer is saved. All Esri Data levels have BDS performance indexes created by default, thus the Rebuild Index button will be inactive.

Distance threshold settings

The data apportionment threshold area settings in Business Analyst are shown in squared area units, such as square miles or kilometers. Extensive research was done to identify data retrieval performance drop-offs by the geographic size of an area, the calculation method, and the data hierarchy. Because of this, default area thresholds are set for the hybrid, block apportionment, and cascading centroid methods. You can change these settings for your own custom methodology.

This section highlights the Data tab options and the default distance settings. Because most trade areas are irregularly shaped, the thresholds are based on the maximum distance of height or width.

To determine the distance threshold, a polygon is drawn around the outermost points of a trade area. The greater value between the width and height of the polygon is used as the distance threshold in the image above. The width is 80 miles and the height is 55 miles. In this example, the 80 mile width is the greatest value. The 80 mile distance falls under the 101 mile thresholds meaning block points are retrieved from the trade area.

1. In the Output Data Format section, you can change your output data format. You can choose from Shapefile, Personal Geodatabase and File Geodatabase.
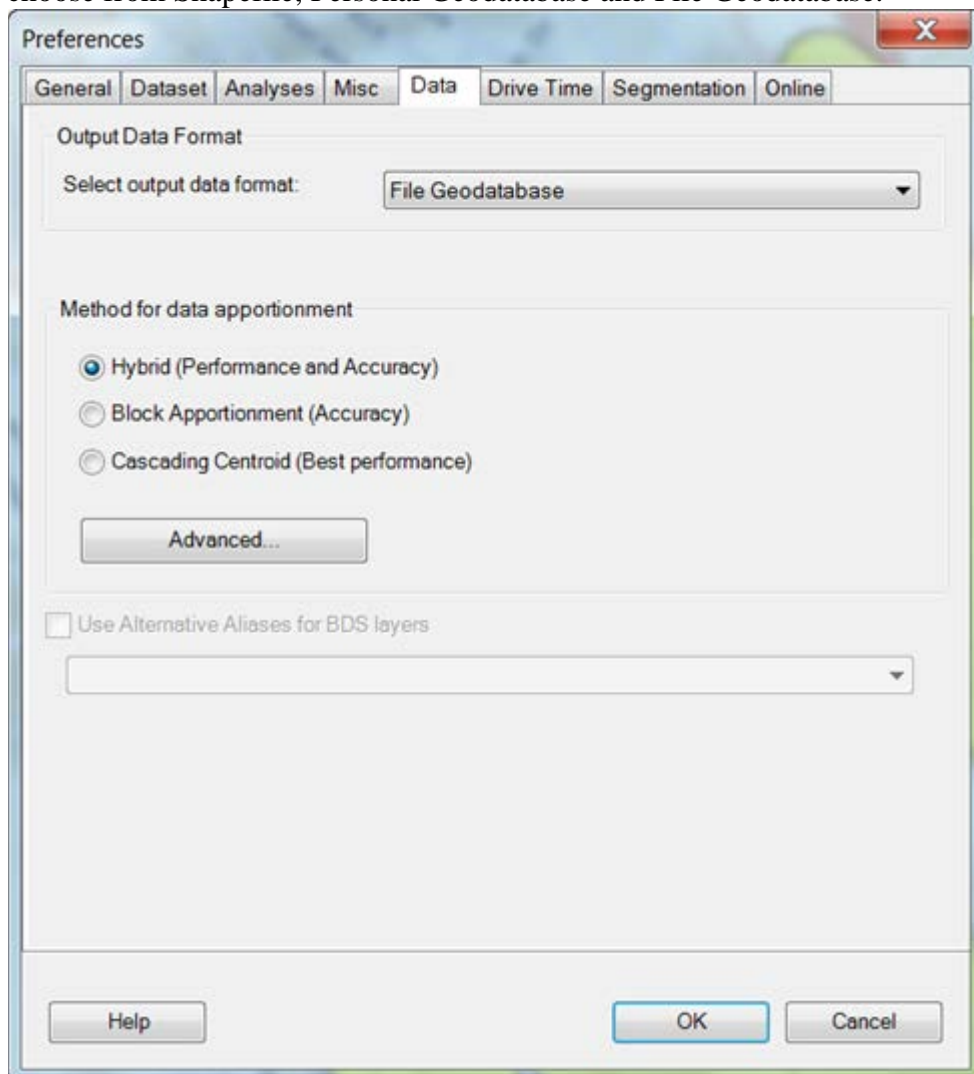

2. Click the Use Alternate Aliases for BDS layer check box to display values set in the Alternate Variable Description created in the Custom Data Setup wizard under Custom Data Parameters.

You can add or subtract levels by clicking the plus or minus buttons. Click on any calculation method to change the data calculation method to BlockApportionment (block apportionment) or CentroidsInPolygon (cascading centroid). Click on a data layer to change the data layer. For example, you can change BlockGroups to ZIP Codes. The counties layer is the largest boundary option for data retrieval.
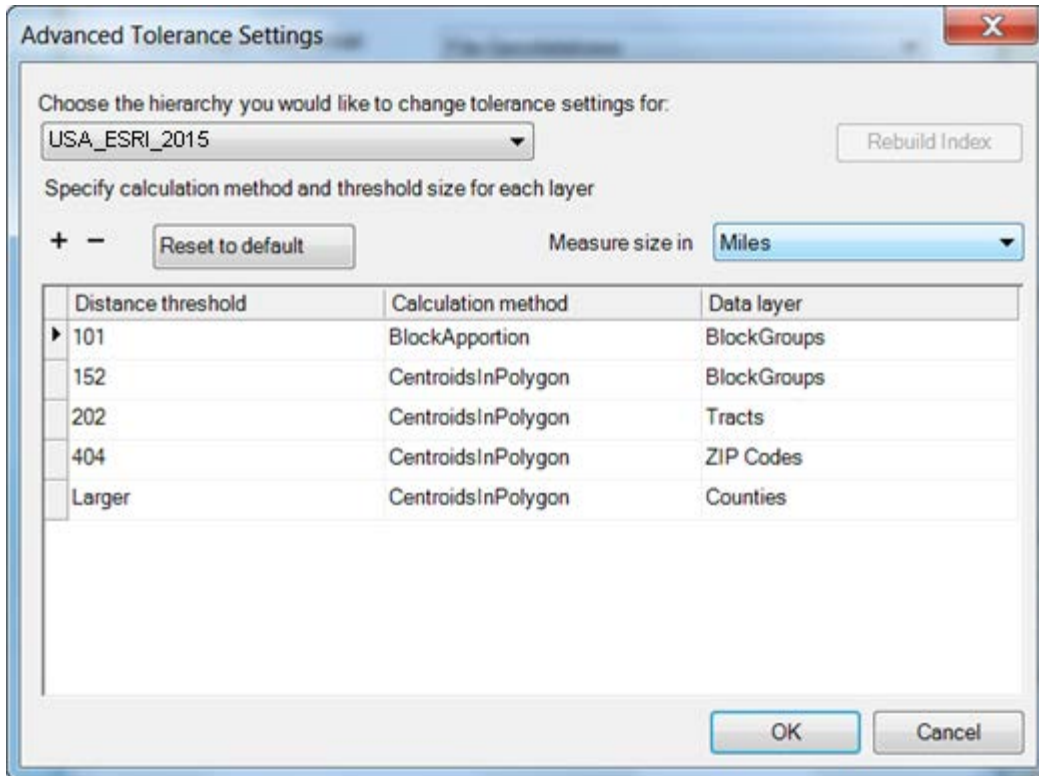
The image below shows a distance reference for the default hybrid data apportionment settings.
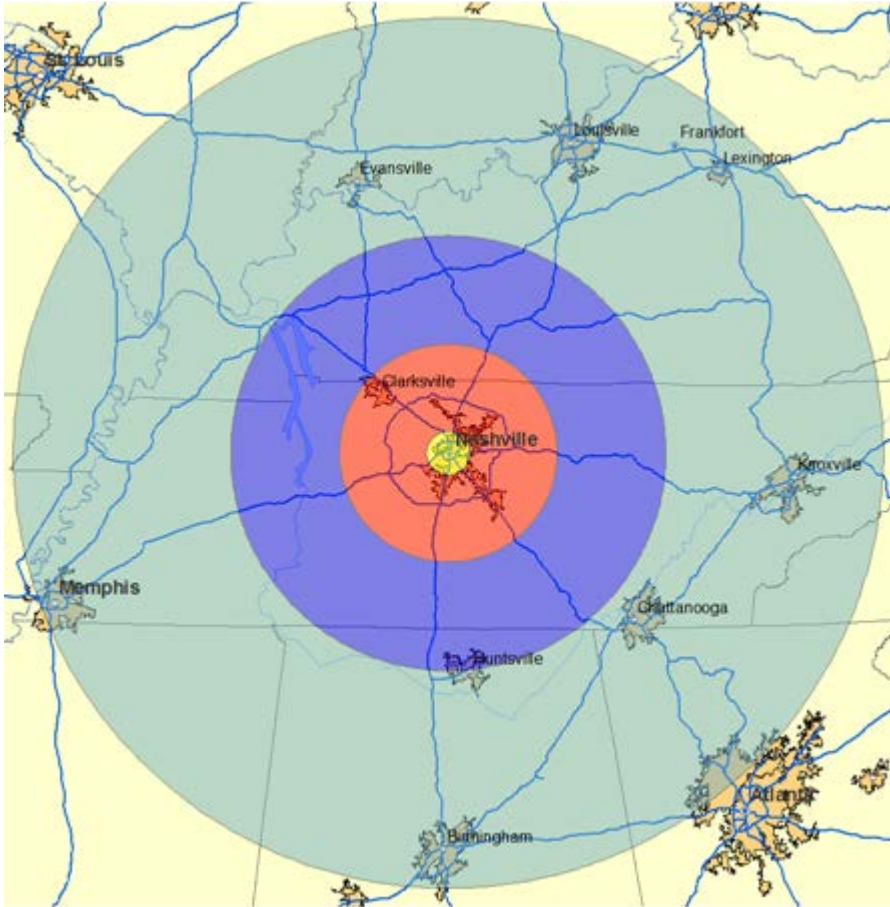
The numbers in the distance Advanced Tolerance Settings dialog box image above correspond with the image below.

101 miles - Block Group - block apportionment is indicated in the yellow area.

152 miles - Block Group Centroids is indicated in the red area.

202 miles - Census Tract Centroid is indicated in the dark blue area.

404 miles - ZIP Code Centroids is indicated in the light blue area.

More than 400 (or larger as indicated in the Advanced Tolerance Setting dialog box image above) - County Centroids is indicated in the off white area.

The distance threshold is calculated using the maximum value of a trade area's height and width. Unlike the example above, many trade areas are not perfect circles but are irregular shapes. The example below shows how a maximum value is calculated for an irregularly shaped trade area.

To determine the distance threshold, a polygon is drawn around the outermost points of a trade area. The greater value between width and height of the polygon is used as the

distance threshold figure. In the example below, 80 is the greatest value. The 80 mile distance falls under the 101 mile thresholds meaning block points are retrieved from the trade area.
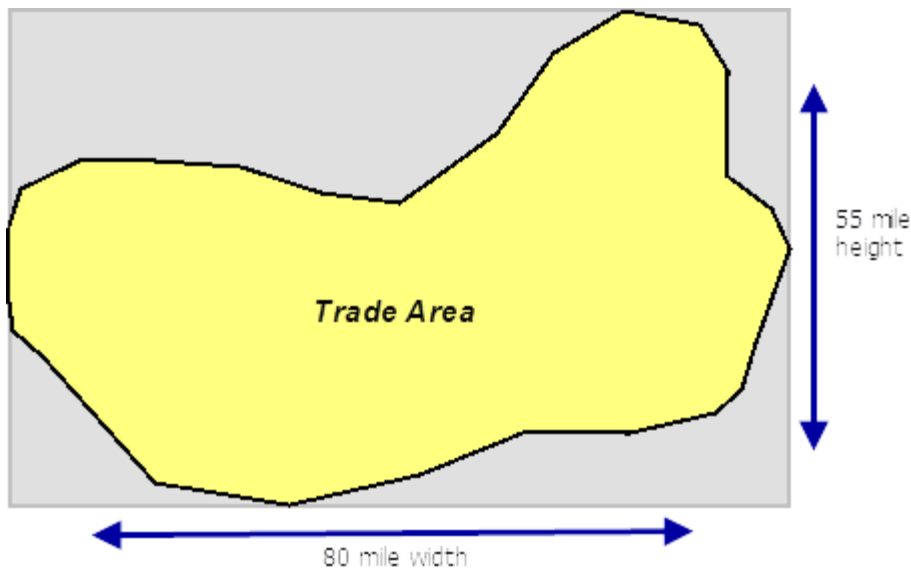
Block Apportionment versus Cascading Centroid

This section explains the block apportionment and cascading centroid methods when small areas of geography are used. In this example, both methods are used in the exact same area to retrieve data using Spatial Overlay from ZIP Code boundaries.

Use the image below for reference to the following notes:

- Imagine that you want to create a population report on this trade area (using the Spatial Overlay tool). You can see that the trade area is an irregular shape and cuts across multiple boundaries.
- ZIP Code centroids are shown here and will be used in the data retrieval process.
- ZIP Code centroids are shown in red. These represent the geographic center points for all ZIP Code boundaries and contain all demographic attributes from their corresponding polygons.

    Note:

    Each demographic boundary level included in Business Analyst contains a centroid. Each of these centroids contains the same demographic attributes as its corresponding boundary layer (for example, block groups, census tracts, DMAs, FSAs, provinces, states, and so on). These centroids are used directly in the cascading centroid apportionment method.

- The block points shown as green dots (derived from census block boundaries to form block groups) are the smallest unit of demographic data in Business Analyst. These demographic attributes are used for weighting all variables and include population, total households, housing units, and business counts.
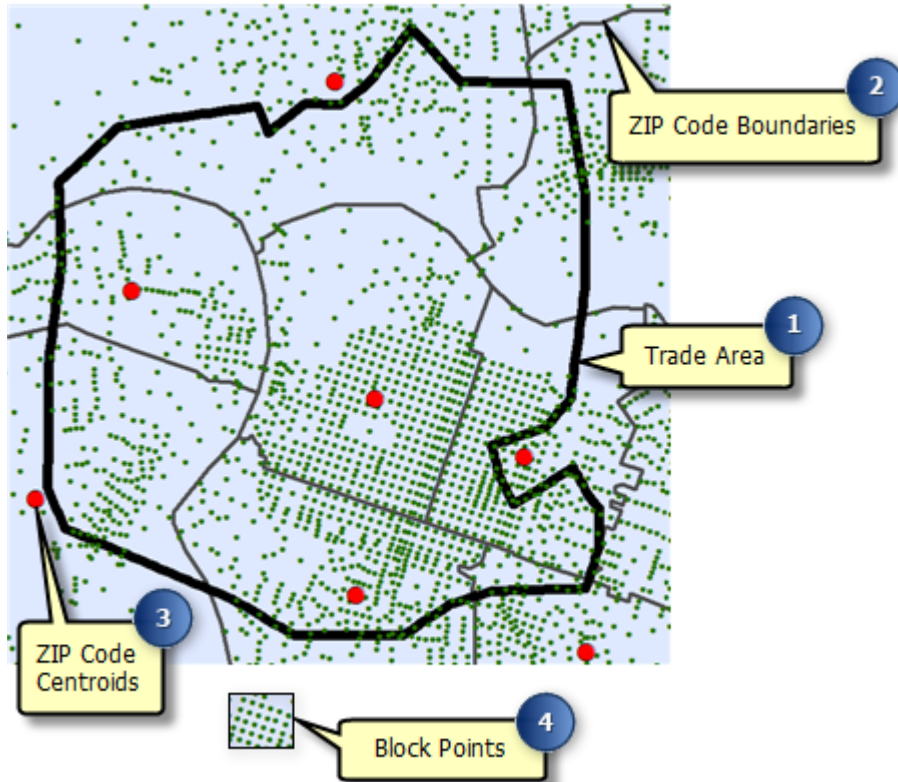
The example below shows the block apportionment data apportionment method. ZIP Codes are used in this case. This highlights that block apportionment is more accurate, especially in smaller geographic areas.

The weighted population value for the 16199 ZIP Code is 75% because 75% of the ZIP Code's population contained in the block points are within the trade area (or 12,115/16,119*100). This scenario can be applied to all other boundaries where the area is subdivided by a trade area.

In the image below, the boxes show the total population added from the block points contained only within the trade area. The block points contained outside the trade area are not included in the analysis. In this example, the figure is more accurately shown as 14.042 instead of 21,676 for the entire ZIP Code. This weighted population value is 65%.

Note:

ZIP Codes are used in this example, but you can use any demographic boundary layer. Block groups are effectively used for smaller geographic areas and are set as the block apportionment default layer.
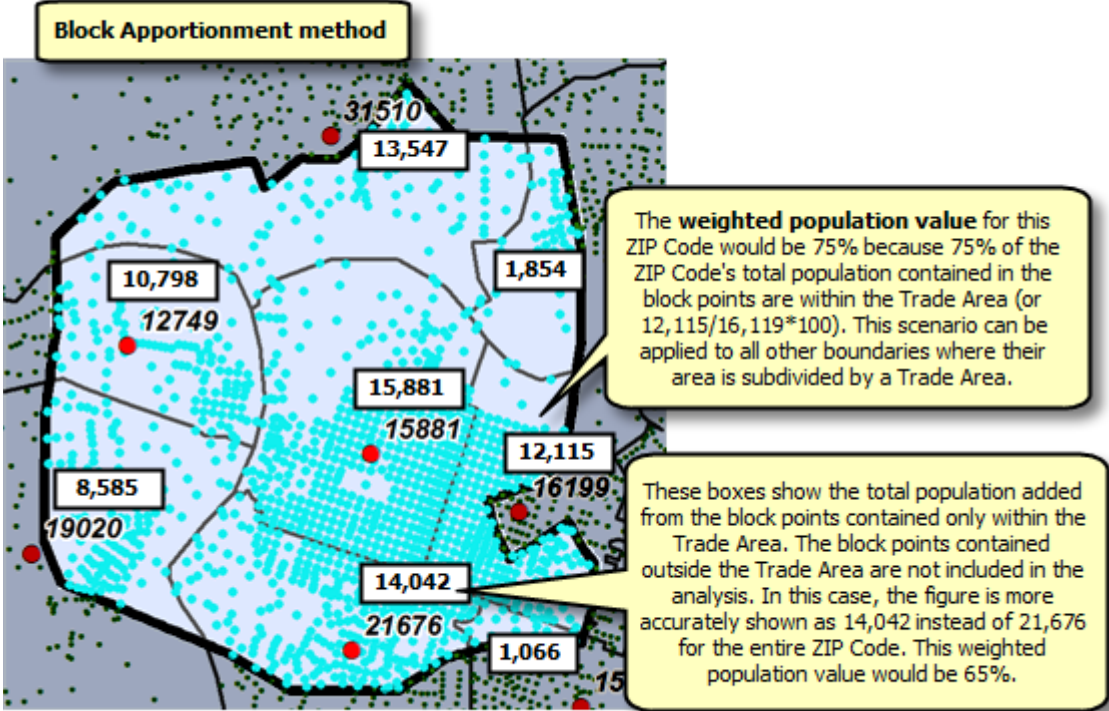
Here is an attribute table output using block apportionment. You can see that the block points included in the trade area are summed to equal the total population figure. So 13,547 + 10,798 + 8,585 + 14,042 + 15,881 + 1,854 + 12,115 + 1,066 = 77,888.

The 2000 Total Population field shows the 2000 Total Population, or 77,888 people, summed from each block point within the trade area. The count field shows the total number of geographies with block points included in the data retrieval. The block apportionment method returns data from nine different ZIP Codes while the cascading centroid method only returns three ZIP Code boundaries.

Block Apportionment output

Attributes of Block Apportionment - ZIP Codes

FID	Shape	ID	2000 Total Population	COUNT
0	Polygon	0	77888	9

Record: 1 / 1   Show: All Selected

This field shows the 2000 Total Population, or 77,888 people, summed from each block point within the Trade Area.

This field shows the total number of geographies with block points that are included in the data retrieval. The Block Apportionment method returned data from 9 different ZIP Codes, while the Cascading Centroid method only returned 3 ZIP Code boundaries.
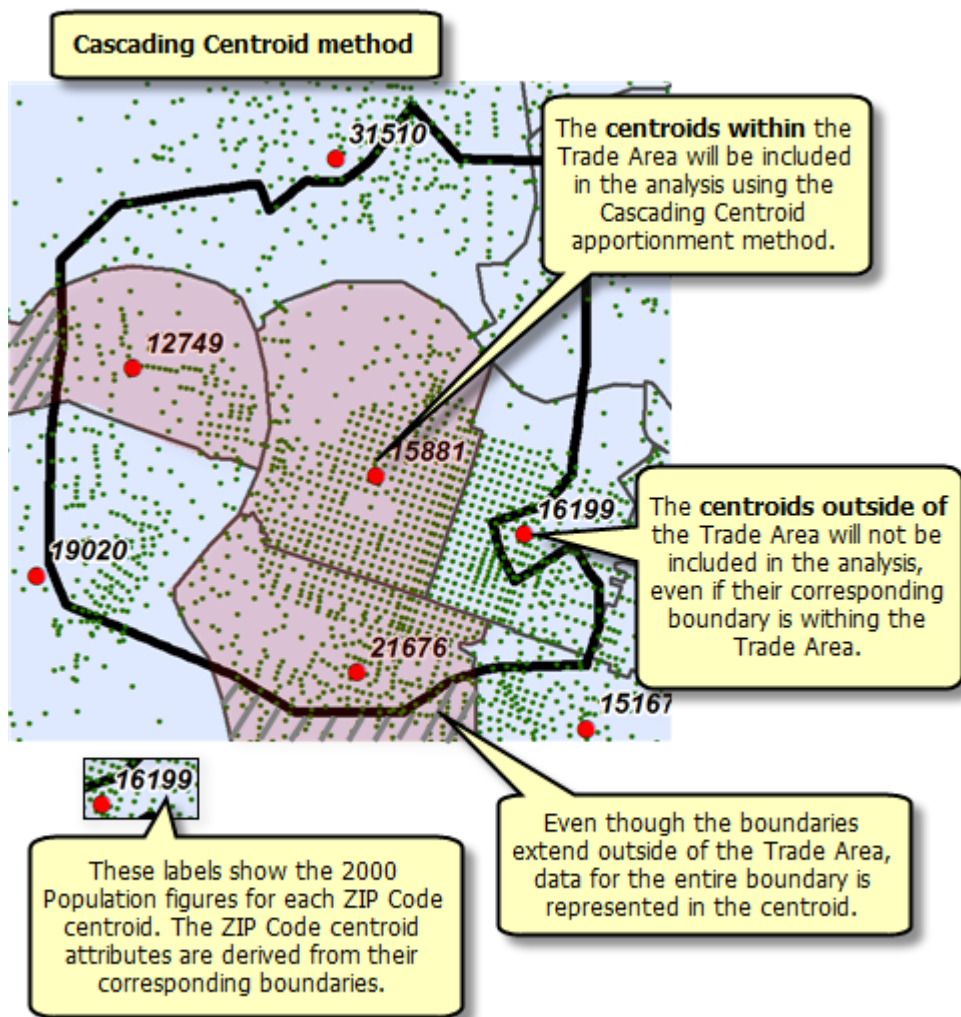
The example below shows the cascading centroid data apportionment method using ZIP Codes. This highlights that cascading centroids can be less accurate, especially in smaller geographic areas.

The centroids within the trade area, as shown for the 15881 ZIP Code, will be included in the analysis using the cascading centroid apportionment method. The centroids outside of the trade area, as shown for ZIP Code 16199, will not be included in the analysis even if the corresponding boundary is within the trade area. Even though the boundaries extend outside of the trade area, as shown in purple for 21676 ZIP Code, data for the entire boundary is represented in the centroid.

The labels show the 2000 Population figures for each ZIP Code centroid. The ZIP Code centroid attributes are derived from their corresponding boundaries.

Cascading Centroid method

The **centroids within** the Trade Area will be included in the analysis using the Cascading Centroid apportionment method.

The **centroids outside of** the Trade Area will not be included in the analysis, even if their corresponding boundary is withing the Trade Area.

These labels show the 2000 Population figures for each ZIP Code centroid. The ZIP Code centroid attributes are derived from their corresponding boundaries.

Even though the boundaries extend outside of the Trade Area, data for the entire boundary is represented in the centroid.

Here is an attribute table output using cascading centroid. You can see that the ZIP Code centroids included in the trade area are summed to equal the total population figure. So 12,749 + 15,881 + 21,676 = 50,306.

The 2000 Total Population field shows the 2000 Total Population, of 50,306 people, summed from each ZIP Code centroid within the trade area. The count field shows the total number of geographies with centroid points that are included in the data retrieval.

Cascading Centroid output

Attributes of Cascading Centroid - ZIP Codes

FID	Shape	ID	2000 Total Population	COUNT
0	Polygon	0	50306	3

Record: 14 ◄ 1 ► ►I Show: All Selected ▼

This field shows the 2000 Total Population, or 50,306 people, summed from each ZIP Code centroid within the Trade Area.

This field shows the total number of geographies with centroid points that are included in the data retrieval.