from: 2012
GIS Tutorial for Crime
GORR & KURLAND
ESRI PRESS, REDLANDS CA
Ch 7
pp 127–149

OBJECTIVES

Identify spatial clusters
Use kernel density smoothing
Conduct Getis-Ord Gi* test for hot spot analysis

Chapter 7

Conducting hot spot analysis

Hot spot analysis is a well-accepted and widely used component of crime analysis. The objective is to automatically identify crime hot spots on a map. ArcMap has tools for this task that are available in ArcToolbox—in both the Spatial Statistics and Spatial Analyst tools. In this chapter, you learn how to put some key components of these tools to work.

Hot spot analysis concepts

Several studies have shown that very small parts of urban areas, hot spots, generate a very large number of crimes (for example, see Weisburd et al. 2004). Hot spots tend to persist, so identified hot spots make good targets for law enforcement and crime prevention. The objective of hot spot analysis is to automatically identify and estimate the boundaries of crime hot spots—spatial clusters of crime incidents for a certain period of time—to provide a "snapshot" of crime. Hot spots are also dynamic—they emerge, persist, decline, and re-emerge. The field officers' pin map in chapter 3 and corresponding animations in chapter 6 help identify such dynamic behavior because they include crime map layers from different time frames, providing a time context. You can see newer crimes in the context of recent but older crimes.

Spatial clustering and surface estimation are two approaches to automatically identifying hot spots. Spatial clustering involves searching for crime incidents that occur unusually close to each other. Surface estimation assumes there is an expected or average crime density (crimes per unit area) distributed over space that is smooth and continuous. The peaks of this surface define the hot spots. In this chapter, you learn how to use both clustering and smoothing techniques.

Spatial clustering

Some approaches to hot spot analysis use statistical methods, in particular, spatial statistics. As is the case with other advanced topics in this book, this introduction provides a brief description of underlying theories and methods and refers you to other resources for more in-depth coverage. You can find excellent documentation with additional references for the ArcToolbox Spatial Statistics toolbox in ArcMap by visiting ArcGIS Desktop Help.

Internet WWW keyword search **spatial statistics**

One simple test you use in this chapter is the Average Nearest Neighbor index. This Spatial Statistics tool can be used to test for spatial clustering. The index uses the ratio of the average distance between crime locations and their nearest-neighbor crime locations for a data sample that is divided by the same measure but for a random rearrangement of crime locations. If the ratio or index is significantly less than 1, it is evidence of spatial clustering. It is expected that crime would be clustered; for instance, many parts of an urban area have low or no crime rates (such as cemeteries, steep hillsides, and bodies of water) while other areas have high crime rates (such as public housing projects, rundown commercial areas, and gang territories). A test such as the Average Nearest Neighbor index is designed more to test how much clustering there is than to test whether or not clustering exists.

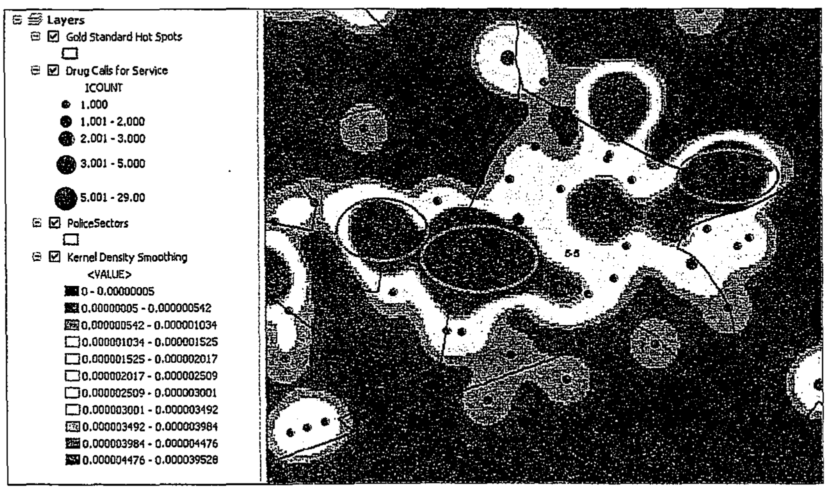
The Spatial Statistics toolbox also provides more sophisticated tests for clustering based on spatial autocorrelation, a measure of spatial clustering that uses aggregate data, such as the number of drug calls for service per block. Two such tests are Moran's I and Getis-Ord Gi* (pronounced G-i-star). Moran's I statistic, created by Patrick Moran, is a measure of spatial autocorrelation, or the degree to which a variable at a location has values similar to the same variable at nearby locations. The Getis-Ord Gi* test, named for its creators, Arthur Getis and Keith Ord, finds the locations and extents of point clusters in two-dimensional space. You will be using the Getis-Ord Gi* test for hot spot analysis in this chapter. A single block may have a high crime count, but it is not considered a hot spot unless there are other nearby blocks that also have high crime counts (that is, there is significant positive spatial autocorrelation).

Surface estimation using kernel density smoothing

Surface elevation is commonly used with rasters to illustrate crime density through surface estimation. A raster map is similar to a photograph and consists of a rectangular matrix of very small square grid cells, each with its own solid color. Raster maps can be used to portray surfaces that can then be used to depict varying levels of crime. You will create raster maps in this chapter that represent crime density—the expected or mean level of crime per unit area—through the use of surface elevation. The higher the expected crime count per unit area, the higher the surface peak. But rather than representing mountain peaks, the elevation in your raster maps will represent crime density peaks, or crime hot spots.

ArcGIS Desktop Help **"What is raster data?"**

Kernel density smoothing is a leading method of surface estimation. "Kernel" is a mathematical term for a kind of function that can be used for averaging. Kernel density smoothing assigns a bell-shaped surface (the kernel) over every crime location, with the highest point of the bell directly over the crime location. Then at each point on the map, the kernel density smoothing method simply sums up all the bell surfaces overhead. Nearby crimes contribute a lot to the sum while distant crimes add only a small amount. The more crimes clustered near each other, the higher the surface at that point. The result, the kernel density surface, changes smoothly over space. For example, in the figure, the clustered drug calls for service result in crime density peaks, shown in red.



This map shows drug calls for service for a three-month period, gold standard hot spots, and a kernel density smoothing surface.

ArcGIS Desktop Help **kernel density smoothing**

Hot spot analysis with raster maps amounts to estimating the crime surface over a jurisdiction, and then selecting a threshold elevation to yield a decision on what constitutes a hot spot. Any area (collection of raster cells) that is at or above the threshold is then considered a crime hot spot. The threshold in the figure map is the minimum crime density that corresponds to the boundary of the red areas. The red areas have crime density at the threshold or higher.

Calibrating kernel density smoothing for hot spots

The critical parameter for calibrating (or tuning) kernel density smoothing is the search radius (or radius of the bell surface placed over each crime). A relatively long radius results in "long rolling hills" for the surface, whereas a short radius results in sharp peaks. Of course, the search radius you need depends on the nature of the hot spots you want to identify.

Implementing hot spot methodology in practice, however, including kernel density smoothing, has led to a great deal of confusion, largely because there has been no "gold standard" approach to designating true hot spots for use in calibrating hot spot estimation methods. Some crime analysts use "rules of thumb" based on expert opinion for kernel density calibration, but there is no way to determine how such rules perform without having a target in place.

The approach used in this chapter to arrive at these gold standards, or assumed true hot spots, is to have crime analysts study crime pin maps and use their expert judgment to manually draw boundaries around the areas they perceive as hot spots. All that is needed is a good sample of expert hot spots; you do not have to exhaustively identify all hot spots in a dataset. Three gold standard hot spots are shown in the kernel density smoothing figure.

In kernel density smoothing, a search radius must be found that generates a crime density surface in which the peaks approximate the gold standard hot spot boundaries. This is not difficult, as you will see in the following exercises. The red areas of the kernel density smoothing surface in the kernel density smoothing figure do a reasonable job of approximating the gold standards identified there. In effect, this approach automates crime analysts' expert judgment by calibrating hot spot methods to replicate crime analysts' expertise as much as possible. The result is a hot spot method that saves time, is comprehensive, and provides consistent hot spots for police work. See "Empirical Calibration of Time Series Monitoring Methods Using Receiver Operating Characteristic Curves" (Cohen, Garman, and Gorr 2009) for an in-depth approach to calibrating methods for detecting the emergence of crime hot spots.

Conducting hot spot modeling

The workflow in the following exercises starts with testing for the existence of spatial clusters, moves on to kernel density estimation of hot spots calibrated by using expert judgment, and finishes with confirmation of the kernel density results by using the Getis-Ord Gi* hot spot analysis. You'll be using the tools that ArcMap provides in ArcToolbox, which simplifies the work. All you have to do is open a tool's parameter input form, select the inputs, and then run the tool to get the outputs.

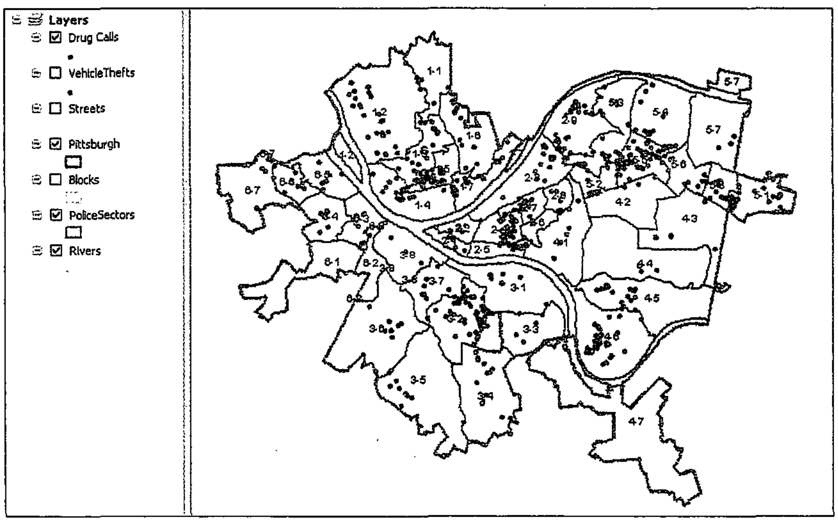
Tutorial 7-1 Lab: Hotspot Analysis 1

Testing for crime spatial clusters

The null hypothesis (or possibility that there is no significant clustering) is a random distribution of points. If crimes are spread out uniformly and randomly, with no spatial clustering, a good strategy would be to use randomized patrols. If, however, crimes are spatially clustered, a strategy of targeted patrols in clustered or hot spot areas makes sense. The nearest-neighbor test that you'll run nex makes a determination of spatial clustering versus the uniformly random, null hypothesis.

Open map document

1 **Open Tutorial7-1.mxd in ArcMap.** The starting map document opens with 911 drug calls for summer 2008 (June through August). You can see what appears to be quite a bit of spatial clustering for drug calls.



2 **Save your map to your chapter 7 folder in MyExercises.**

Create size-graduated point markers

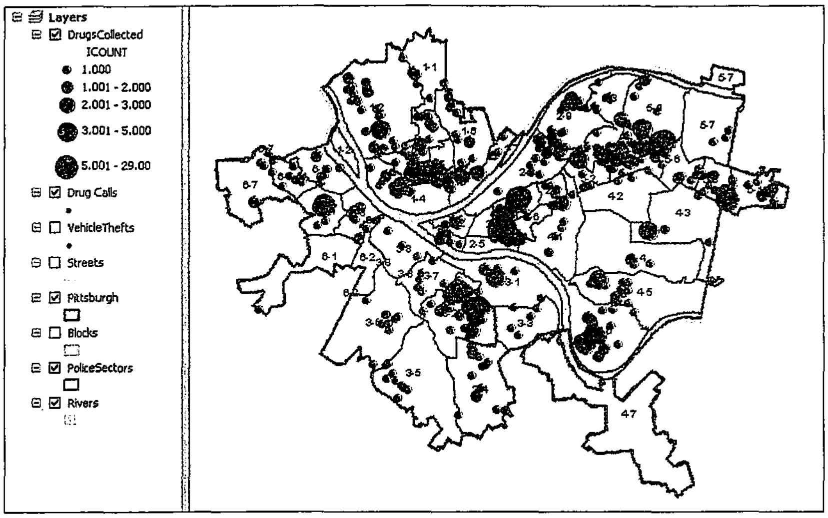
To make clustering look even more apparent on the map, you can start by counting the number of drug calls at each address, and then use size-graduated point markers to symbolize Drug Calls. The larger the point symbol, the more drug calls at the same location. Many drug-dealing locations have multiple calls for service.

1 **On the Standard toolbar, click ArcToolbox.**

2 **In the ArcToolbox window, expand the Spatial Statistics toolbox and then the Utilities toolset. Then double-click the Collect Events tool.**

3 **For Input Incident Features, select Drug Calls. For Output Weighted Point Feature Class, click the Browse button, go to the chapter 7 geodatabase in MyExercises, and type DrugsCollected for Name. Click OK > Close.** Leave the ArcToolbox window open.

4 **In the table of contents, right-click DrugsCollected, select Properties, and click the Symbology tab. Then click the Classify button and select Quantile for Classification Method. Click OK. Click the Template button and select Mars Red for color. Click OK twice.** Now, with size-graduated point markers, it is even more evident that there is a lot of clustering. Try zooming in on a few police sectors that have clusters for a closer look. Turn Streets on while zoomed in. Next, you'll examine whether the nearest-neighbor statistical test provides further confirmation.

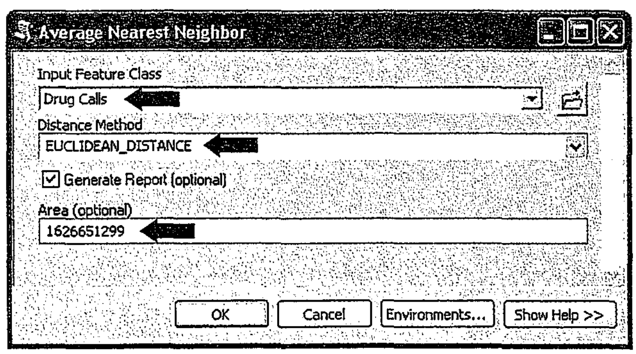


Test for spatial clustering

1 **In the Spatial Statistics toolbox, expand the Analyzing Patterns toolset and double-click the Average Nearest Neighbor tool.** The tool needs the total area of the study region (Pittsburgh). If you were to open the Pittsburgh layer's attribute table, you would see that its area is 1,626,651,299 sq ft. Note that to get a more realistic test, you could subtract the areas of steep hillsides, water areas, cemeteries, and other land types or uses where you'd expect no drug dealing. Then results would be less statistically significant, however, because the null hypothesis would have higher crime densities per the area reduction. That is, crime clusters would be less dense in relation to the null hypothesis.

2 **Enter selections as shown in the figure.**

3 **Click OK.**



4 **Click in the message window that appears after the tool runs and see what's inside.** You should get a warning message that there were bad records that the tool ignored, which is fine. These are simply records that did not get geocoded from police records to points on the map, because they had invalid street addresses or the street centerline map was incomplete or inaccurate. Scroll down in the Results window where you'll also get the following output of the test:

```
Average Nearest Neighbor Summary
Observed Mean Distance: 249.565425
Expected Mean Distance: 630.183834
Nearest Neighbor Ratio: 0.396020
z-score: -36.974592
p-value: 0.000000
```

The test results provide evidence that drug calls are heavily clustered in Pittsburgh because the observed mean distance to the nearest neighbors of drug calls is significantly less than the expected mean distance under the null hypothesis that drug calls are uniformly and randomly distributed. The important statistic is the z-score, a common measure used in statistics that is calculated by subtracting the estimated mean of the distribution at hand from a data observation, and then dividing the difference by the sampled standard deviation. If the z-score has a value that is less than -2 or -3, it is evidence that the clustering does not occur by chance alone. In this case, the z-score is -36.97, which indicates a high level of statistical significance to the clustering of the crime data. In other words, there is just a small possibility that the clustering happened by chance alone and is therefore a phenomenon with some causal basis.

5 **Close the Results window.**

YOUR TURN YOUR TURN YOUR TURN YOUR TURN YOUR

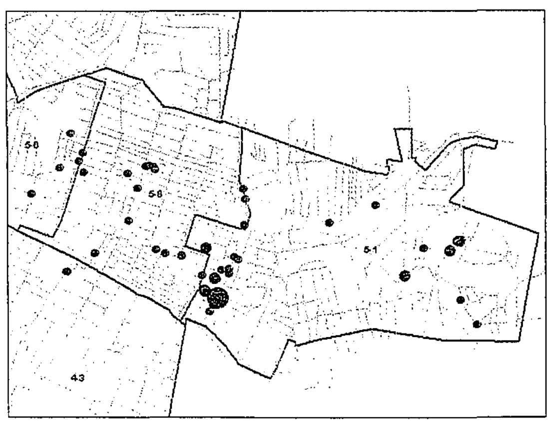
Turn Vehicle Thefts on and turn Drug Calls and DrugsCollected off. Do you think vehicle thefts are spatially clustered? It would be easier to answer this question by visual inspection if vehicle thefts had size-graduated point markers, so create **VehicleTheftsCollected** and save to the chapter 7 geodatabase in MyExercises. Test for spatial clustering by using the Average Nearest Neighbor statistic with Vehicle Thefts as the input. You should get a z-score of -13.4, which is highly significant but less so than the same test for drug calls.

Visually scan for drug-dealing clusters

In preparation for using smoothing methods to estimate hot spots, you can try exercising your judgment about hot spots. The objective in this exercise is to identify hot spots for directing patrols, so you'll want hot spots that are fairly small. You can draw ellipses around areas that you believe are drug-dealing hot spots. You will have to zoom in to see potential hot spots.

1 **Turn DrugsCollected on, turn any other point layers off, and turn the Pittsburgh layer off.**

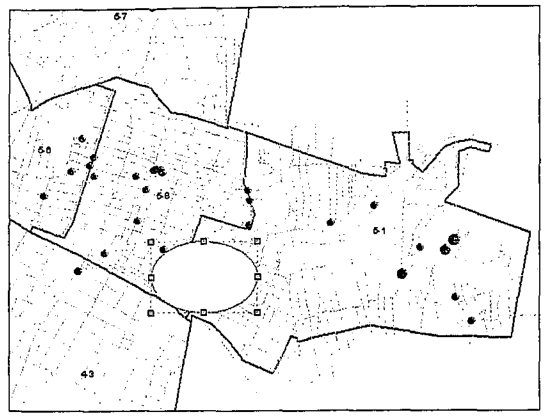
2 **Zoom in on the eastern portion of Pittsburgh and turn Streets on.**



3 **From the Menu bar, click Customize > Toolbars > Draw.**

4 **On the Draw toolbar, click the Draw Shapes arrow and select Ellipse. Then draw an ellipse roughly as shown in the figure, adjusting the ellipse's position, shape, and size by dragging its handles.**

5 **Right-click the ellipse, select Properties, and click the Symbol tab. Select No Color for the fill color, a dark blue for the outline color, and 2 for the outline width. Click OK.**



YOUR TURN

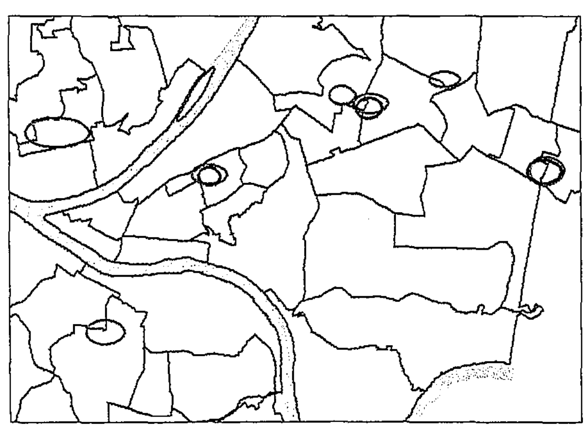
Zoom in on other parts of Pittsburgh, look for two more hot spots, and draw ellipses around them. Make sure no graphics are selected, right-click the Layers label at the top of the table of contents, and select Convert Graphics To Features. Then click the Browse button for "Output shapefile or feature class," select "File and Personal Geodatabase feature classes" from the "Save as type" menu, and save **MyExpertHotSpots** to the chapter 7 geodatabase in MyExercises. Select the "Automatically delete graphics after conversion" check box. Click OK > Yes.

Add sample of supplied hot spot ellipses

Now, you know how to create expert-drawn hot spots. To save time and prepare for the following exercises, you'll next add a layer that has a larger sample of expert hot spot ellipses.

1 **Click Add Data, go to the chapter 7 geodatabase in FinishedExercises, and add SuppliedHotSpots to the map.**

2 **Symbolize the added layer with No Color fill and a size 2 green outline.**



3 **Compare your ellipses with the supplied hot spot ellipses.** The three supplied ellipses with dark blue outlines represent the work you've done so far, and those with green outlines are the ones supplied for use in the following exercises. Perhaps you have ellipses that overlap the supplied ellipses. Neither selection is exhaustive of all potential hot spots, but you do not need to have all hot spots identified—just a good sampling of them. The objective is to train the kernel density smoothing method to approximate your judgment or someone else's. So in the next tutorial, you'll learn how to calibrate kernel density smoothing to produce a drug call surface that matches the SuppliedHotSpots layer.

4 **Save your map document.**

Tutorial 7-2

Using kernel density smoothing

The Spatial Analyst extension makes it simple to estimate density surfaces. All you need to do is set up the processing environment, and then supply a few inputs to a tool.

Activate Spatial Analyst and make environment settings

Spatial Analyst is one of several ArcMap extensions that users can purchase to add more functionality to their maps. With Spatial Analyst installed on your computer, you must activate it to be able to use it. You'll also need to change some overall settings, called Environment Settings, that affect all the tools in ArcToolbox.

1 **Open Tutorial7-2.mxd in ArcMap.**

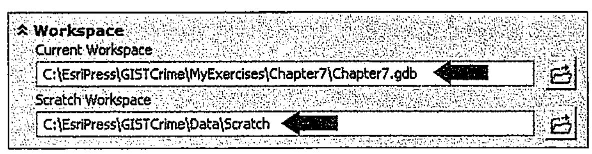
2 **Save your map document to your chapter 7 folder in MyExercises.**

3 **From the Menu bar, click Customize > Extensions. Select the Spatial Analyst check box to turn the extension on. Click Close.** There are no apparent changes, but now you can use Spatial Analyst.

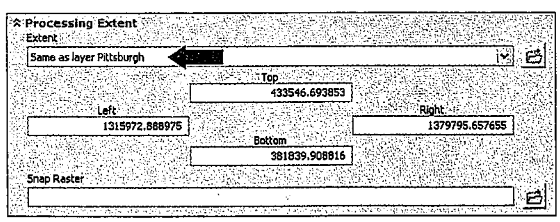
4 **If necessary, click ArcToolbox on the Standard toolbar to open the ArcToolbox window.**

5 **In the ArcToolbox window, right-click the ArcToolbox icon and select Environments. Lengthen the Environment Settings window so you can see all the possible settings.**

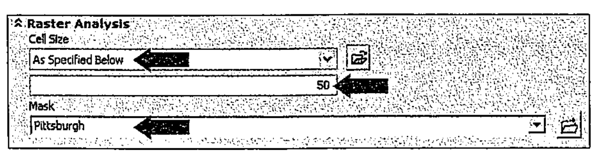
6 **Click Workspace and make selections as shown in the figure.**



7 **Click Workspace to close the panel. Then click Processing Extent and select "Same as layer Pittsburgh."** Selecting this layer supplies the coordinates of the bounding rectangle, or extent, of any new layer that you create by using Spatial Analyst.



8 **Click Processing Extent to close the panel. Then click Raster Analysis to expand that panel and enter selections as shown in the figure. When finished, close the Raster Analysis panel.** A small cell size such as the default "1" increases the smoothness of the appearance of the raster map when zoomed in, but it also markedly increases computation times and the requirements for hard-disk storage. Setting the cell size to 50 ft to a side is adequate. Raster maps are all rectangular in shape. Specifying a mask limits the display of such a map to a jurisdiction boundary, such as Pittsburgh. ArcMap assigns all raster cells outside the mask No Color, thus making other portions of the map invisible.



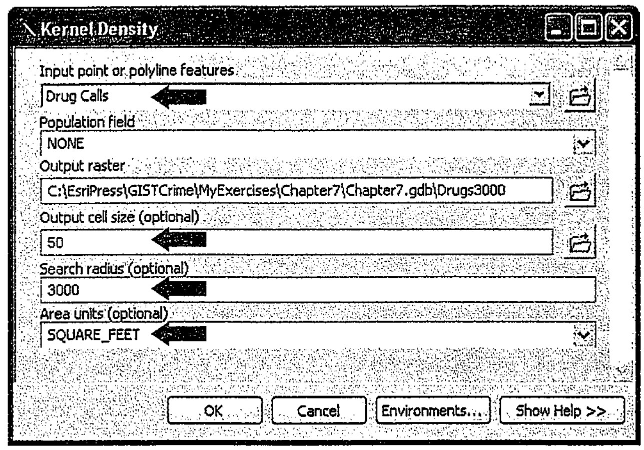
9 **Click OK to save settings and close the Environment Settings window.**

Apply kernel density smoothing

The critical parameter to vary in the kernel density method is the search radius. The search radius includes points within that distance of a grid cell used to estimate the expected, or "smoothed," drug calls per square foot. The search radius will be relatively small because the expert cluster ellipses are small—around 2,000 ft along the long axis and 1,000 ft on the short axis. So, try 3,000, 1,000, and 500 ft for the search radius and see if one of these radii generates peak areas that match the green ellipses.

1 **In ArcToolbox, expand the Spatial Analyst toolbox and then the Density toolset. Then double-click the Kernel Density tool.**

2 **Enter selections as shown in the figure.**



3 **Click OK.** This is the smoothest of the three surfaces you'll create, because it has the largest radius. In fact, it looks too smooth, as if it could use more peaks and valleys.

Symbolize density surface

Step 1 avoids a potential problem in symbolizing the raster map. You need to execute this step just once for a given map document.

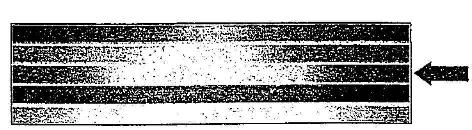
1 **From the Menu bar, click Customize > ArcMap Options. Click the Raster tab and then the Raster Catalog tab. Type 1000000 for "Maximum number of rasters for color matching." Then click OK.**

2 **Double-click Drugs3000 in the table of contents and click the Symbology tab.**

3 **Click Classify.**

4 **For Classification Method, select Standard Deviation, and for Classes, select 1/3 Std Dev. Click OK.**

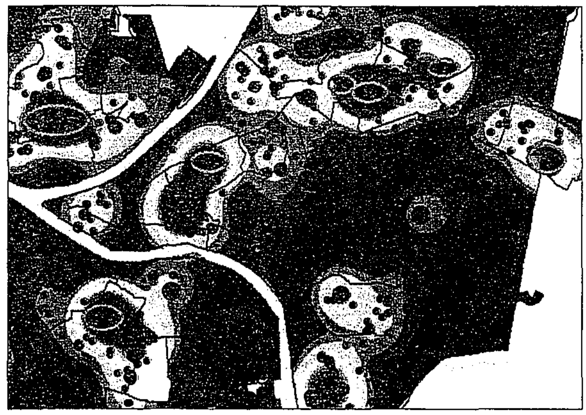
5 **Select the color ramp that ranges from red to yellow to blue.**



6 **Click the Symbol heading above the resulting color chips in the Symbology panel and select Flip Colors so that low densities are blue and high densities are red.**

7 **Click OK.**

8 **Turn the DrugsCollected layer on and off for comparison with the kernel density and supplied hot spots.**

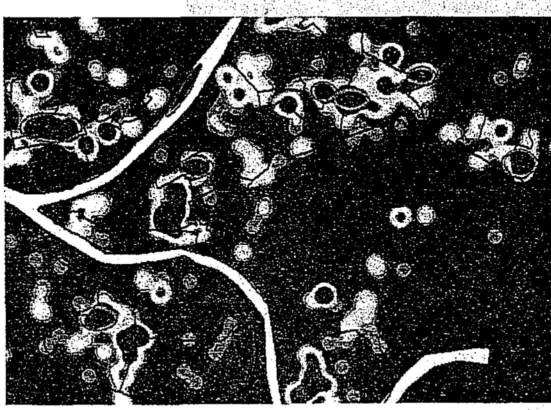


3,000 ft search radius.

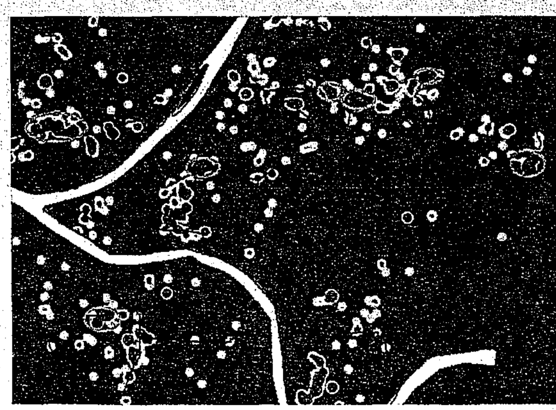
9 **Try zooming in on hot spots.** This level of smoothing, with a search radius equal to 3,000 ft, does a nice job of capturing large hot spots in the red-colored raster cells. It might serve well as a map for the public—say, with the use of neighborhoods instead of police sectors. You'll need smaller hot spots, however, to direct police patrols. Drug dealing and gang turfs are more on the order of several blocks, which is quite a bit smaller than these areas.

YOUR TURN YOUR TURN YOUR TURN YOUR TURN YOUR

Run kernel density smoothing two more times, once with a search radius of **1000** ft, called **Drugs1000**, and once with a search radius of **500** ft, called **Drugs500**. Save the new raster layers to the chapter 7 geodatabase in MyExercises. Symbolize these layers as you did for the 3,000 ft layer. The 500 ft radius appears to be too short, resulting in smaller hot spots than identified by the expert ellipses, while the 1,000 ft radius does a fair job of matching them. Zoom in on several hot spots in the 1,000 ft radius density map and see how well the density smoothing does at this scale. The results are reasonable and useful. If you have new data for drug calls in Pittsburgh, you can use smoothing with a 1,000 ft search radius and be comfortable that Spatial Analyst will automatically find all the hot spots for you. When you are finished investigating the various density maps for hot spots, save your map document.



1,000 ft search radius.



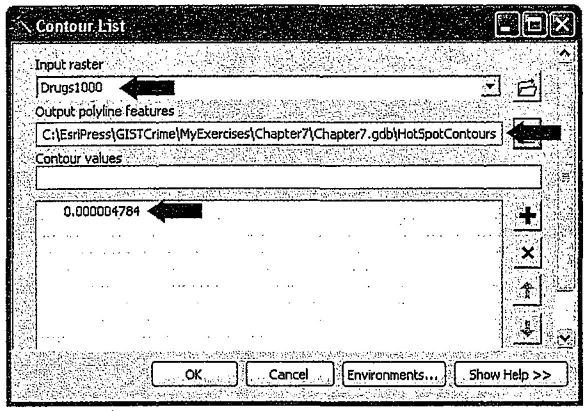
500 ft search radius.

Create contour polygons defining a drug hot spot

The blue to red color ramp you have been using has hot colors that correspond to the threshold levels needed to define hot spots. Part of the reason this choice of color ramps works is because you used standard deviations for break points in the color ramp, and the high standard deviations pertain to the very high values of the distribution. Next, you'll learn how to convert the boundaries of the hot spot cells of the 1,000 ft radius map into vector contour lines stored as polygons. The polygon format makes the results easier to use and display.

1 **In the ArcToolbox window, expand the Spatial Analyst toolbox and then the Surface toolset. Then double-click the Contour List tool.**

2 **Enter selections as shown in the figure.**

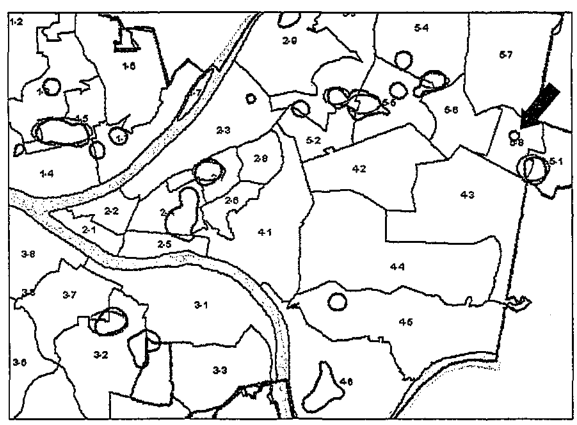


Hint: Type the value 0.000004784 in the text box directly under the Contour Values label, and then press the Add button to get the value entered as shown in the figure.

The contour value of 0.000004784 corresponds to the start of the second-highest range (reddish orange color chip) in the 1,000 ft search radius density map. So, the resulting contours will enclose hot spots that are reddish orange and red on the color ramp.

3 **Click OK.**

4 **Turn all density maps off and turn SuppliedHotSpots on.**



5 **Symbolize the resulting HotSpotContour layer with a red outline of width 2.** The result is a simple representation of drug hot spots that can be used for directing patrols. Notice that the contours from kernel density smoothing fit the supplied expert hot spots reasonably well. Also, some hot spots cross the boundaries of more than one police sector, suggesting that corresponding patrol units should coordinate patrols in the affected areas. Finally, kernel density smoothing found more hot spots than there were in the set of expert hot spots, but the latter dataset was intended only as a representative sample for training kernel density smoothing and not as the definitive set of hot spots.

6 **Save your map document.**

Edit hot spot contours

Suppose upon visual examination you feel that some of the areas defined by contours are too small or otherwise not drug hot spots. In that case, you can simply edit the contour map and delete unwanted polygons.

1 **In the table of contents, click List By Selection and make HotSpotContour the only selectable layer.**

2 **From the Menu bar, click Customize > Toolbars > Editor.**

3 **On the Editor toolbar, click Editor > Start Editing, Select HotSpotContour1000. Click OK.**

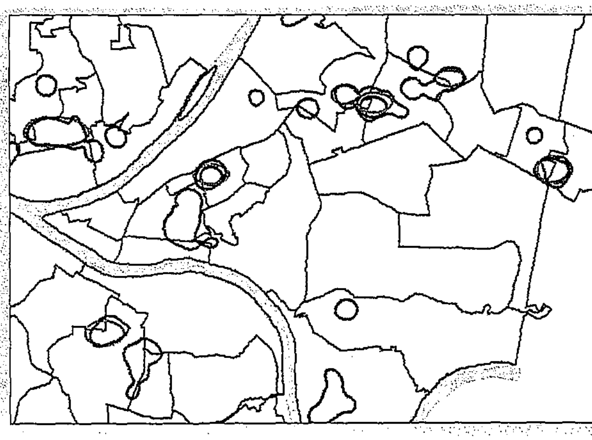
Next, you'll delete all polygons that are roughly the size of the contours indicated in HotSpotContour within Police Sector 5-3.

4 **On the Standard toolbar, click the Select Features tool and select a polygon you wish to delete by drawing a rectangle around it.**

5 **Open the HotSpotContour attribute table, right-click the row selector of the highlighted row (gray cell on the left side of the record), and click Delete Selected.** Both the record and the polygon disappear.

YOUR TURN YOUR TURN YOUR TURN YOUR TURN YOUR

Use the Select Features tool, pressing SHIFT to select one or two more small contour areas, and then delete them in the attribute table. When you are finished, click Editor > Stop Editing. Click Yes. Save your map document.



Tutorial 7-3

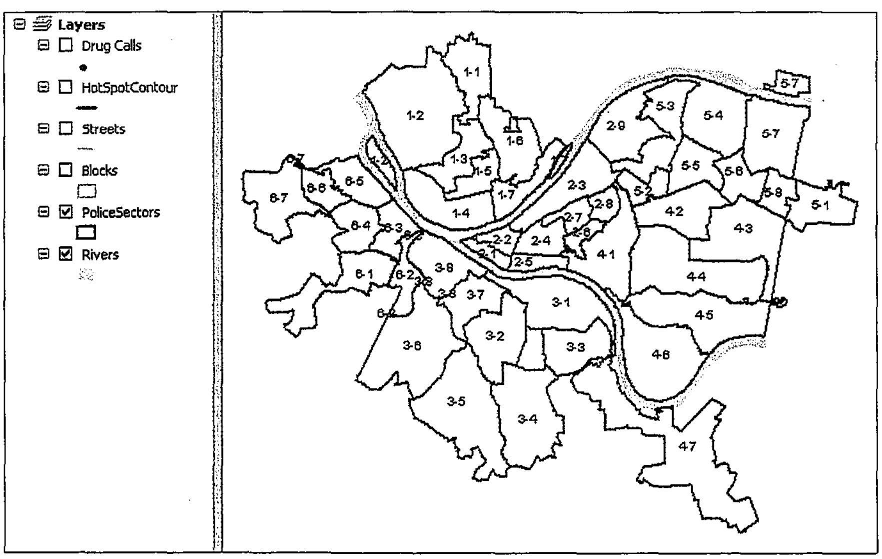
Conducting Getis-Ord Gi* test for hot spot analysis

The Getis-Ord Gi* test identifies clusters of points that have higher values than expected by chance. The test, which you learn how to use in the following exercises, requires data aggregated from individual crime point locations to crime counts for small areas represented by centroid points. Aggregation is required because the test expects there to be variations in data, so, for example, the data cannot consist of all 1's. If an area's crime count is high and its neighbors' crime counts are also high, the Getis-Ord Gi* test concludes that these areas are part of a hot spot. A block is a good-size area to use for much of crime analysis, including this test, because the hot spots used for directing patrols or other police interventions need to be for small areas, and crime problems related to the hot spots are often at the scale of blocks. Creating the needed aggregate data takes several steps.

Open map document

1 **Open Tutorial7-3.mxd in ArcMap.**

2 **In the table of contents, click List By Drawing Order.**



3 **Save your map document to your chapter 7 folder in MyExercises.**

Add XY centroid coordinates to a polygon attribute table

The centroid of a polygon is the point from which the polygon would balance if it were cut from cardboard and placed on a pencil point. (The case of a quarter-moon-shaped polygon is an exception because its centroid is not inside its boundary.) So, a centroid is a kind of center of a polygon that makes a handy map layer for displaying certain kinds of polygon information. In this exercise, you add centroid coordinates to the Blocks attribute table. You need to do this task just once to have the results available for future use.

1 **In the table of contents, right-click Blocks and select Open Attribute Table.**

2 **Click Table Options > Add Field. Type X for Name and select Double for Type. Click OK.**

3 **Repeat step 2 to create the Y field.**

4 **Right-click the X column heading and select Calculate Geometry. Click Yes. Select X Coordinate of Centroid and click OK.**

5 **Repeat step 4 to compute the Y Coordinate of Centroid.**

OBJECTID*	Shape*	BLKIDFP00	Shape_Length	Shape_Area	X	Y
1	Polygon	420032507001029	946.199601	41055.393258	1337035.571578	418516.163244
2	Polygon	420032107002003	1627.614639	88197.0056	1335744.758819	417237.087807
3	Polygon	420032507001033	924.436852	42069.51039	1336748.908627	418315.409402
4	Polygon	420032507001030	550.128088	18246.792107	1337290.261691	418562.555745
5	Polygon	420032107002012	1100.189362	63867.763186	1336630.172077	417079.809194
6	Polygon	420032107002046	1405.710049	78135.671335	1335996.603714	415873.945166

6 **Click Table Options > Export. Click the Browse button and go to the chapter 7 geodatabase in MyExercises. Select "File and Personal Geodatabase tables" from the "Save as type" menu, change Name to BlockCentroidsTable, and click Save. Click OK > No.**

7 **Close the table.**

Create block centroid map layer

Creating a feature layer for block center points, or block centroids, is another one-time task. The results can be used over the life of the crime mapping system or until the block map is revised.

1 **From the Menu bar, click Windows > Catalog.**

2 **Expand the Catalog tree to the contents of the chapter 7 geodatabase in MyExercises.**

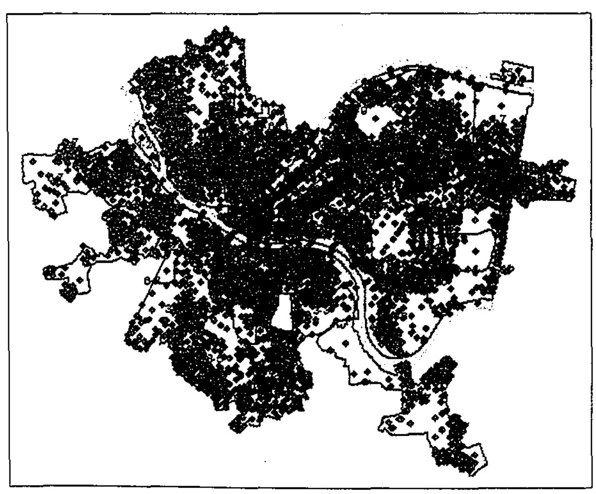
3 **Right-click BlockCentroidsTable and click Create Feature Class > From XY Table.**

4 **Click the Coordinate System of Input Coordinates button, click Import, and select Blocks. Click Add > OK.**

5 **Change Output Feature Class to BlockCentroids in the chapter 7 geodatabase. Click Save > OK.**

6 **Drag BlockCentroids from Catalog to the table of contents under Drug Calls.**

7 **Close the Catalog window.** Now you have a feature layer for block centroids. In the next exercise, you'll count the drug calls by block and attach the counts to the block centroids layer you just created.



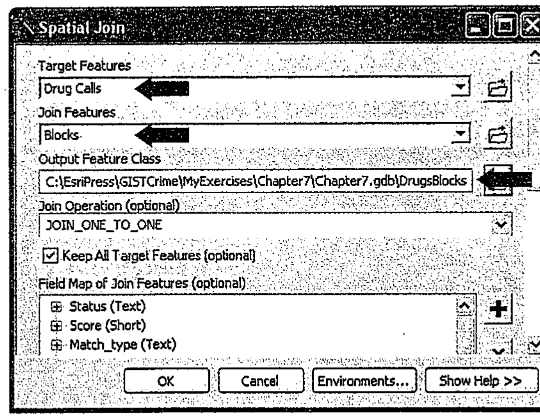
Overlay points with polygons

Every time you have new data for hot spot analysis, you'll need to perform this task and the remaining ones in this section. In this exercise, you use spatial overlay to assign a block ID to each crime point. Note that spatial overlay arbitrarily assigns crime points located at a street intersection to a single block, and thus does not double-count these points. There is no "correct answer" for assigning such points to blocks, so this is the best solution. Crime points with street addresses have a 20 ft offset from street centerlines, and thus fall within the correct blocks.

1 **In ArcToolbox, expand the Analysis toolbox and then the Overlay toolset. Then double-click the Spatial Join tool.**

2 **Enter selections as shown in the figure.**

3 **Click OK.** Now every drug call record in the DrugsBlocks layer includes the block ID, BLKIDFP00, for the block that it's in.



Aggregate drug calls by block

1 **In the table of contents, right-click DrugsBlocks and select Open Attribute Table.**

2 **Right-click the BLKIDFP00 column heading and select Summarize.**

3 **For item 3 of the Summarize dialog box, specify Output Table as DrugsAggregated and save to the chapter 7 geodatabase in MyExercises. Click OK > Yes.** The resulting table has the number of drug calls per block in Pittsburgh. The table in the figure shows rows 25–30.

OBJECTID *	BLKIDFP00	Count_BLKIDFP00
25	420030203001086	1
26	420030305001010	2
27	420030305002001	2
28	420030305002002	13
29	420030305002005	14
30	420030305002007	3

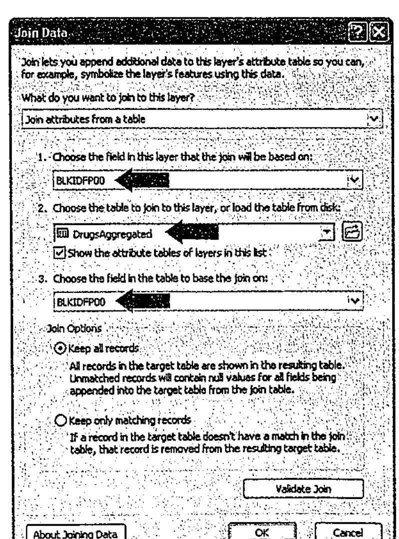
4 **Close the DrugsAggregated table.**

Join drug calls by block to a block centroids layer

1 **In the table of contents, right-click BlockCentroids. Then click Joins and Relates > Join.**

2 **Make selections as shown in the figure.**

3 **Click OK > Yes.**



Export feature layer with joined data to a new feature layer

If you open the BlockCentroids attribute table, you can see that there are many blocks with no drug calls and that the joined count attribute, Count_BLKIDFP00, has "NULL" for many values. You need to replace these nulls with a zero (0), but ArcMap will not let you calculate joined fields. The remedy is to export BlockCentroids to a new feature layer. That way, ArcMap includes the joined attributes as permanent attributes in the resulting feature layer, and you can replace the nulls with zeros.

1 **Remove DrugsBlocks from the table of contents.** This version of DrugsBlocks has individual drug call points. You can now replace this layer with the version that has drug counts per block, presented as block centroids.

2 **In the table of contents, right-click BlockCentroids. Then click Data > Export Data.**

3 **In the Export Data dialog box under Output Feature Classes, click the Browse button and go to the chapter 7 geodatabase in MyExercises. Type DrugsBlocksAggregated for Name, select "File and Personal Geodatabase feature class" from the "Save as type" menu, and click Save. Click OK > Yes.**

4 **Right-click DrugsBlocksAggregated and select Open Attribute Table.**

5 **Click Table Options > Select By Attributes.**

6 **In the top panel, or attributes panel, double-click "Cnt_BLKIDFP00", click the Is button, and then click Get Unique Values. In the values panel, double-click NULL to yield** `"Cnt _ BLKIDFP00" Is NULL`**. Click Apply.**

7 **Right-click the Cnt_BLKIDFP00 column heading in the attribute table and select Field Calculator.**

8 **In the Field Calculator dialog box, type 0 in the bottom panel and click OK.**

9 **Close the attribute table.**

10 **From the Menu bar, click Selection > Clear Selected Features.** You have now cleared the blocks with no drug calls from the map.

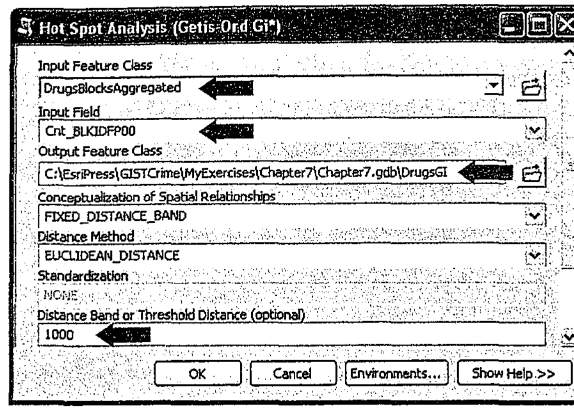
11 **Save your map document.**

Conduct Getis-Ord Gi test*

In this exercise, you can see how the hot spot analysis from kernel density smoothing compares to a sophisticated statistical approach, the Getis-Ord Gi* test. The key idea here is to match the search radius of both methods by using a 1,000 ft fixed band (search radius) for Getis-Ord Gi*, which is the same search radius chosen for kernel density smoothing. As with most statistical tests, the Getis-Ord Gi* test focuses on false positive rates. A false positive rate in this case is the fraction of blocks that are not true members of hot spots yet are falsely classified as hot spot areas by the Getis-Ord Gi* test. By tradition, statistical testing most often uses conservative false positive rates of 0.01 or 0.05 to establish decision rules for classifying potential hot spots.

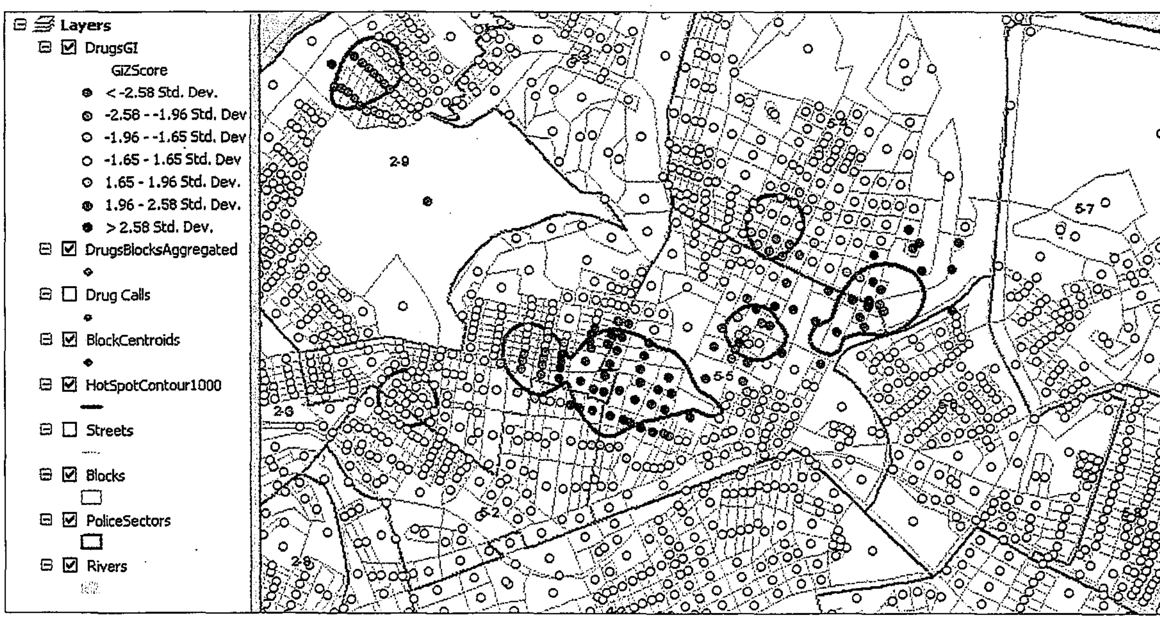
1 **In the ArcToolbox window, expand the Spatial Statistics toolbox and then the Mapping Clusters toolset. Then double-click the Hot Spot Analysis (Getis-Ord Gi*) tool.**

2 **Enter selections as shown in the figure.** The selection of a fixed band of 1,000 ft corresponds to the best-case kernel density map, which similarly weighs neighboring points by use of a 1,000 ft search radius.



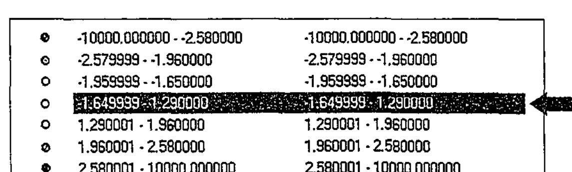
3 **Click OK.**

4 **After ArcMap adds the new DrugsGI layer to the table of contents, turn the HotSpotContour layer on and zoom in on several of the areas that have red and orange point markers.** The red and orange point markers indicate blocks that are significantly hotter than other blocks and, as you can see, correspond reasonably well to the HotSpotContour boundaries. The orange point markers correspond to a false positive error rate of 0.05 (z-value = 1.65 and higher) and the red point markers correspond to a 0.01 false positive error rate. This test result provides confirmation that the hot spots you found from kernel density smoothing are statistically significant. Next, you'll try a more lenient false positive error rate of 0.10 (z-value = 1.29 and higher) for orange point markers, which is closer to police preferences. Police organizations are willing to spend resources on a certain amount of false positives in order to find real positives, or hot spots.

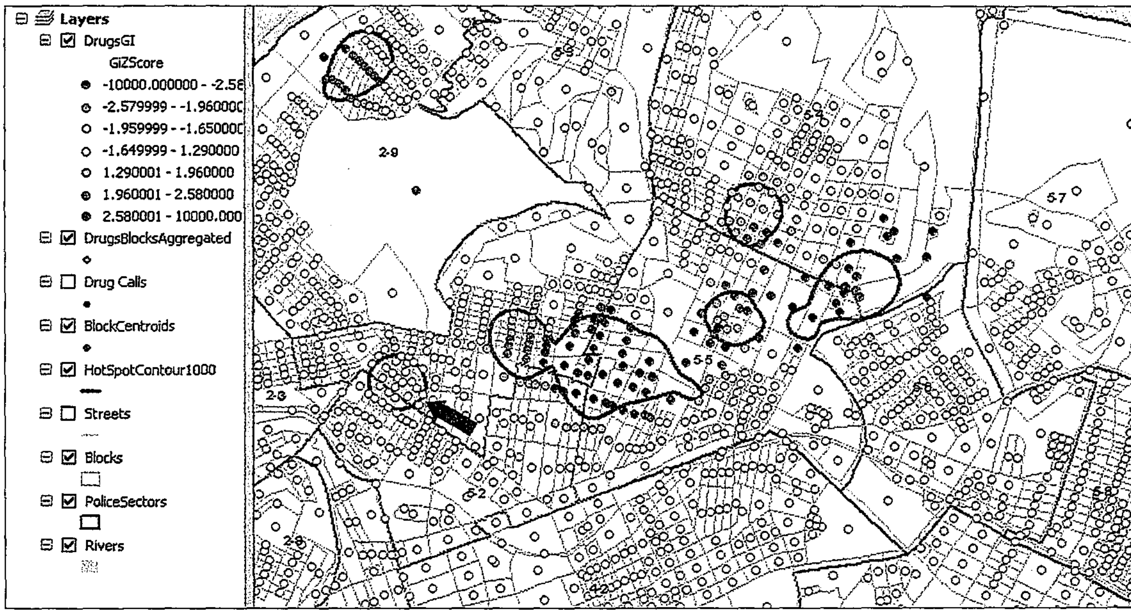


5 **In the table of contents, double-click DrugsGI and click the Symbology tab.**

6 **Click the Range -1.649999–1.65000 (yellow point marker), type 1.29, and press the TAB key.** Even though you type just one value, 1.29, ArcMap fills out the entire range and completes a revised set of categories. This makes the orange category correspond to a false positive rate of 0.10, so orange, orange-red, and red categories are now considered significant.



7 **Click OK.** There is some expansion of the number of blocks that are determined to be in hot spots. The Getis-Ord Gi* analysis considers the circular hot spot that was derived from the smoothing method, indicated in the figure, to be a false positive, whereas all the other smoothing hot spots are confirmed. Overall, the Getis-Ord Gi* test provides confirmation of using the judgmental approach to calibrate kernel density smoothing.



8 **Save your map document and exit ArcMap.**

Assignment 7-1

Assess impact of date duration on hot spots

For some policing purposes, the more recent the information, the better. So, the question becomes, how short can the time interval be (year, quarter, month, week) for hot spot analysis to be meaningful? Working in the opposite direction, the more that reliability of statistical methods and modeling (such as smoothing) is improved, the larger the sample size. So, you might conclude that for a given crime, there is a minimum threshold at which it makes no sense to attempt hot spot analysis short of that time interval.

For this assignment, investigate the impact of time intervals by using the drug call data applied to the exercises in this chapter. The hot spots for three summer months looked fine in the exercises, so try hot spots for a month and then a week to see how short a time interval is feasible for crime analysis.

Create new map document

Create a new map document called **Assignment7-1YourName.mxd** and a new file geodatabase called **Assignment7-1.gdb** and save them to your assignment 7-1 folder in MyAssignments. Include the following map layers in your map document:

- Streets, Rivers, Pittsburgh, and PoliceSectors from the Pittsburgh geodatabase in the Data folder
- CADDrugsSummer08 from the Police geodatabase in the Data folder and DrugsCollected from the chapter 7 geodatabase in FinishedExercises

Save all new map layers that you create to your assignment 7-1 geodatabase.

Symbolize layers

- Symbolize Streets, Rivers, Pittsburgh, and Tracts, using best practices.
- In the table of contents, rename DrugsCollected to **DrugsCollectedSummer08.** Symbolize this layer, using graduated point markers with a Circle 2 point marker in Mars Red and the quantile classification method with five classes. Save the results as a layer file named **DrugsCollected.lyr** to your assignment 7-1 folder for reuse when classifying collected drugs for other time periods.

Create definition queries

From CADDrugsSummer08, make two additional layers: **CADDrugsAugust08**, which has drug calls for August 2008, and **CADDrugsLastWeekAugust08**, which has drug calls for the last seven days of August 2008.

Conduct spatial analysis

Create a Word document called **Assignment7-1YourName.docx** and save it to your assignment 7-1 folder. Run the nearest-neighbor test to see if each kind of crime is spatially clustered and include the results in your document.

Create a kernel density map for each time duration—summer, month, and week. Use Pittsburgh as the mask, 50 ft for the grid cell size, and 1,500 ft for the search radius. Render these maps using the symbology used in this chapter: standard deviations with one-third increments and the blue to red color ramp.

Hint: You can save a lot of time by using collected offense points as the inputs to density estimation, instead of aggregating points to blocks or other polygons. Although it may be better to aggregate points to small polygons because that is a smoothing step in itself, collected points nonetheless provide a reasonable approximation for quick analysis. So, to do this, use the Collect Events tool, which is in the Spatial Statistics toolbox in the Utilities toolset. You will need to use this tool for each queried time period: summer, month, and week. Use your exported layer file, DrugsCollected, to keep the symbology consistent.

Zoom in on the North Side of Pittsburgh, the area north of the three rivers. Export each density map with the corresponding collected-drugs layers displayed. Add the maps to your Word document and label each with a caption such as "Figure 1. Density Map for Drug Calls for Service: Pittsburgh North Side, June–August 2008."

Comment on each map's hot spots. Consider only the hottest density interval, red. Besides the density map for the summer period, are the density maps for the two shorter periods useful? Explain. Does the nearest-neighbor test provide any guidelines?

What to turn in

Use a compression program to compress and save your assignment 7-1 folder to **Assignment7-1YourName.zip**. Turn in your compressed file.

Assignment 7-2

Examine broken-windows theory for simple and aggravated assaults

Broken-windows theory predicts that soft crimes will harden over time to serious crimes. Simple assault is a major leading indicator of serious violent crimes, and aggravated assault is the serious violent crime in this crime category. A simple assault is an intentional act by a person that causes apprehension in another of imminent harm or contact. It is an attempt to frighten another person. Aggravated assault is more serious and includes intent to do serious bodily harm—that is, to rob, rape, or kill. If simple assaults tend to harden into aggravated assaults, you'd expect for hot spots for simple assaults to contain or overlap hot spots for aggravated assaults. While not a complete test of the broken-windows theory, this expectation of connected hot spots is nevertheless predicted by the theory.

For this assignment, analyze these two types of crime, simple assaults and aggravated assaults, to see if there is a correlation between them.

Create new map document

Create a new map document called **Assignment7-2YourName.mxd** and a new file geodatabase called **Assignment7-2.gdb** and save them to your assignment 7-2 folder in MyAssignments. Add the following map layers:

- Streets, Rivers, Pittsburgh, and PoliceSectors from the Pittsburgh geodatabase in the Data folder
- Two copies of Offenses2008 from the Police geodatabase in the Data folder

Save all new map layers to your assignment 7-2 geodatabase.

Symbolize layers

- Symbolize Streets, Rivers, Pittsburgh, and Tracts using best practices.
- Turn both copies of Offenses2008 off in the table of contents.

Create definition queries

Limit each copy of Offenses2008 to June through August 2008. Add to your date query to make one copy **Simple Assaults: Summer 2008** and the other **Aggravated Assaults: Summer 2008.**

Conduct spatial analysis

Create a Word document called **Assignment7-2YourName.docx** and save it to your assignment 7-2 folder. Run the nearest-neighbor test to see if each kind of crime is spatially clustered and include the results in your document.

Create a kernel density map for each crime type, using Pittsburgh as the mask, 50 ft as the grid cell size, and 1,500 ft as the search radius. Render these maps using the symbology used in this chapter: standard deviations with one-third increments and the blue to red color ramp.

Hint: Use the same hint as in assignment 7-1: instead of aggregating points to blocks or other polygons, use the collected points as the input to kernel density smoothing.

Create a contour map for each density map using the start value of the highest-density class (red). Symbolize each contour map using a different line color and turn off all map layers except the contours and spatial context layers. Export the map (or better, a layout) as **Contours.jpg** to your assignment 7-2 folder and add it to your Word document. Comment on the evidence for broken windows in regard to simple and aggravated assaults. Is there a tendency for hot spots for simple assaults and hot spots for aggravated assaults to overlap?

What to turn in

Use a compression program to compress and save your assignment 7-2 folder to **Assignment7-2YourName.zip**. Turn in your compressed file.

References

Cohen, J., S. Garman, and W. L. Gorr. 2009. Empirical calibration of time series monitoring methods using receiver operating characteristic curves. *International Journal of Forecasting* 25:48–61.

Weisburd, D., S. Bushway, C. Lum, and S. Yang. 2004. Trajectories of crime at places: A longitudinal study of street segments in the City of Seattle. *Criminology* 42:283–321.