

# Make Kernel Density maps using ArcView and CrimeStat III

NICAR 2010

## Big Picture

Making sense of patterns of dots in GIS can be challenging. While datapoints can be readily placed on a map in ArcView, patterns can be difficult to see or to convey because there are so many datapoints or because dots are close to or on top of one another.

One way to summarize and display is to count the datapoints into another layer of polygons, like census tracts or ZIPs, and then shade the polys according to the results. This approach masks all data patterns below the poly level, and portrays the polys as homogenous. Also, a relationship between datapoints falling on either side of a polygon boundary, say on either side of a street, may be lost.

Kernel density maps are another way to summarize datapoint densities. They use very small rectangular features, often only hundreds of feet on a side, to accumulate datapoint counts. These counts are then smoothed among nearby rectangles. The end result of a map that looks like a topographic map describing the datapoint pattern.

CrimeStat is a free software, the development of which was funded by federal money. It was created to allow police to make maps of crime hotspots from incident-level crime data. But CrimeStat also can be used to perform a Kernel density analysis using any set of datapoints.

## Get ready

Go to the CrimeStat III web page at <http://www.icpsr.umich.edu/CRIMESTAT/>  
Download the program and manual. The link Download All Files gets them all. Unzip the archive. Also download the User Workbook & its sample files.

Check out the manual & workbook. The workbook & each manual chapter is a separate .pdf. It is a lot of material. Here's what I found of interest:

**Manual Chapter 1** – Installing the program and samples.

**Manual Chapter 2** – A quick overview. Focus on these the screens

Data setup | primary file

Data setup | reference file

Spatial modeling | interpolation

**Manual Chapter 8** – Kernel Density Interpolation. This chapter covers more in detail about how Kernel density works and the optional settings.

Don't worry about the complicated formulas. But the graphs, sample maps and other figures are invaluable.

**Workbook Chapter 6** – Kernel Density Estimation.

## The ingredients

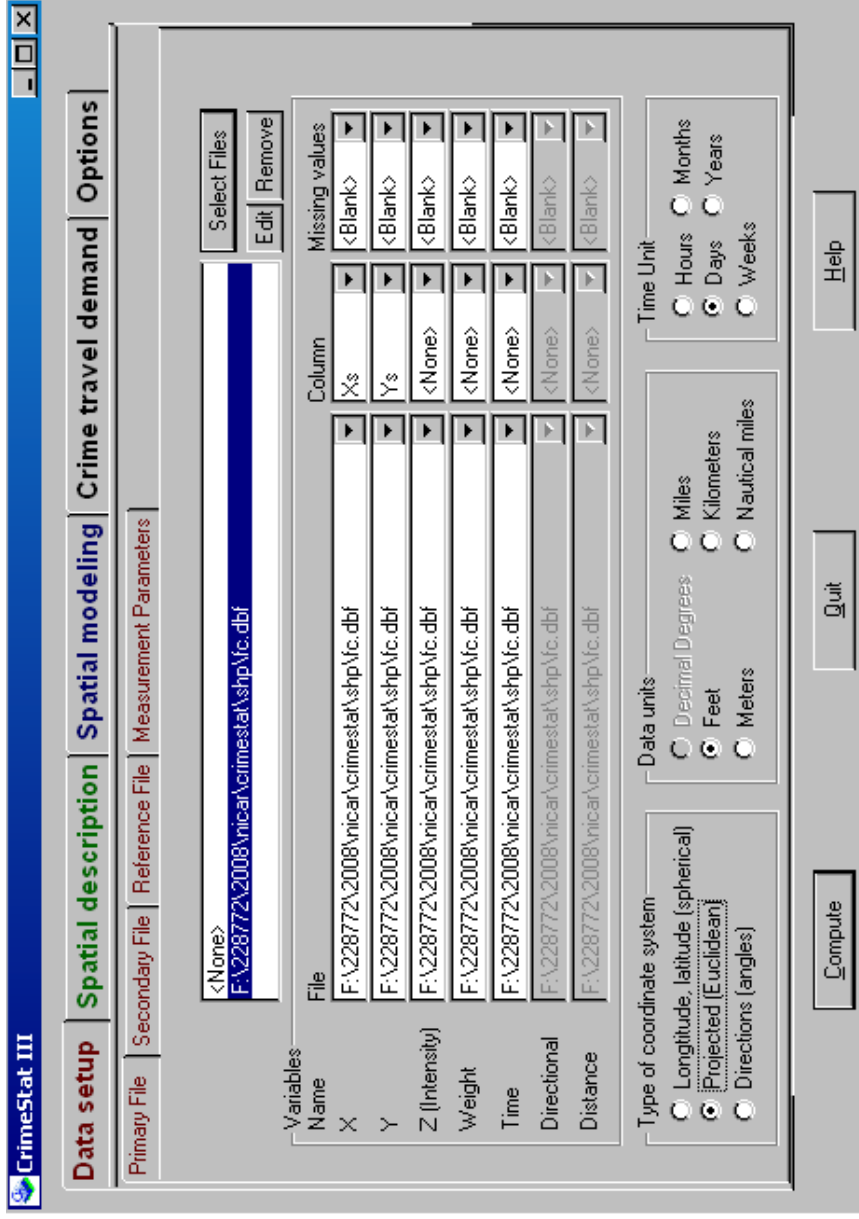
**Datapoint file.** This is the file of events you want to portray. You need a datapoint file with fields holding X and Y coordinates (like longitude & latitude). This can be the .dbf portion of a shapefile. In this example, I'm using fc.dbf from a shapefile of foreclosures I have added columns Xs and Ys and filled them with the coordinates of each point. The coordinate system for my sample files:

NAD\_1983\_StatePlane\_North\_Carolina\_FIPS\_3200\_Feet, so Xs and Ys are in feet.

**Other shapefiles,** for reference geography. I'm using a county outline and major roads.

**Grid bounds.** In ArcView, display your datapoint file and, make a note of the maximum and minimum coordinates that would bound a rectangle enclosing your data. If you activate the arrow tool, you can hover over the topmost datapoint and record the max Y value that is displayed in the ArcView status bar. Hover over the bottom datapoint to find the min Y. Do the same, on left and right sides to get X max and min.

Open the CrimeStat program.

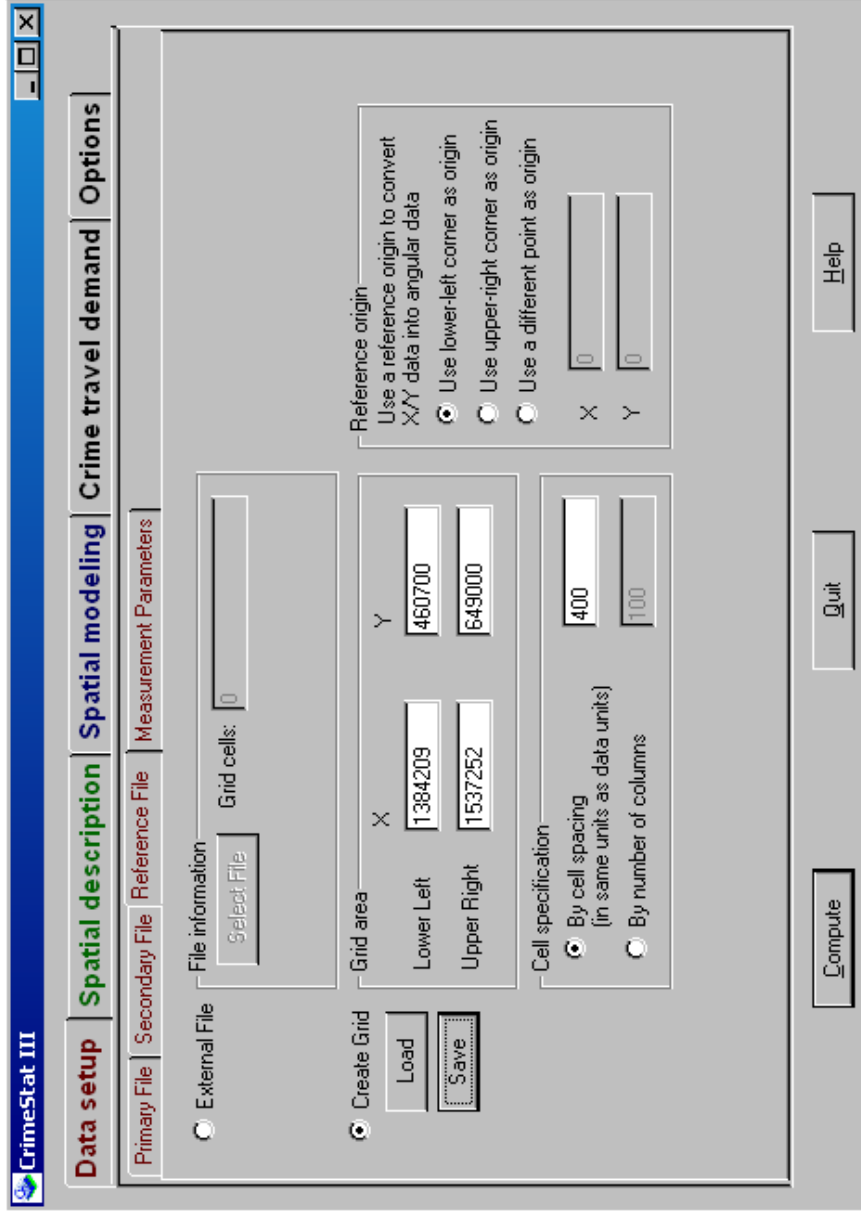


In the Data setup | Primary file tab:

Point to your datapoint file in the top input box.

Select the datapoint fields that hold X and Y coordinates

I changed type of coordinate system to Projected, and the Data units to feet

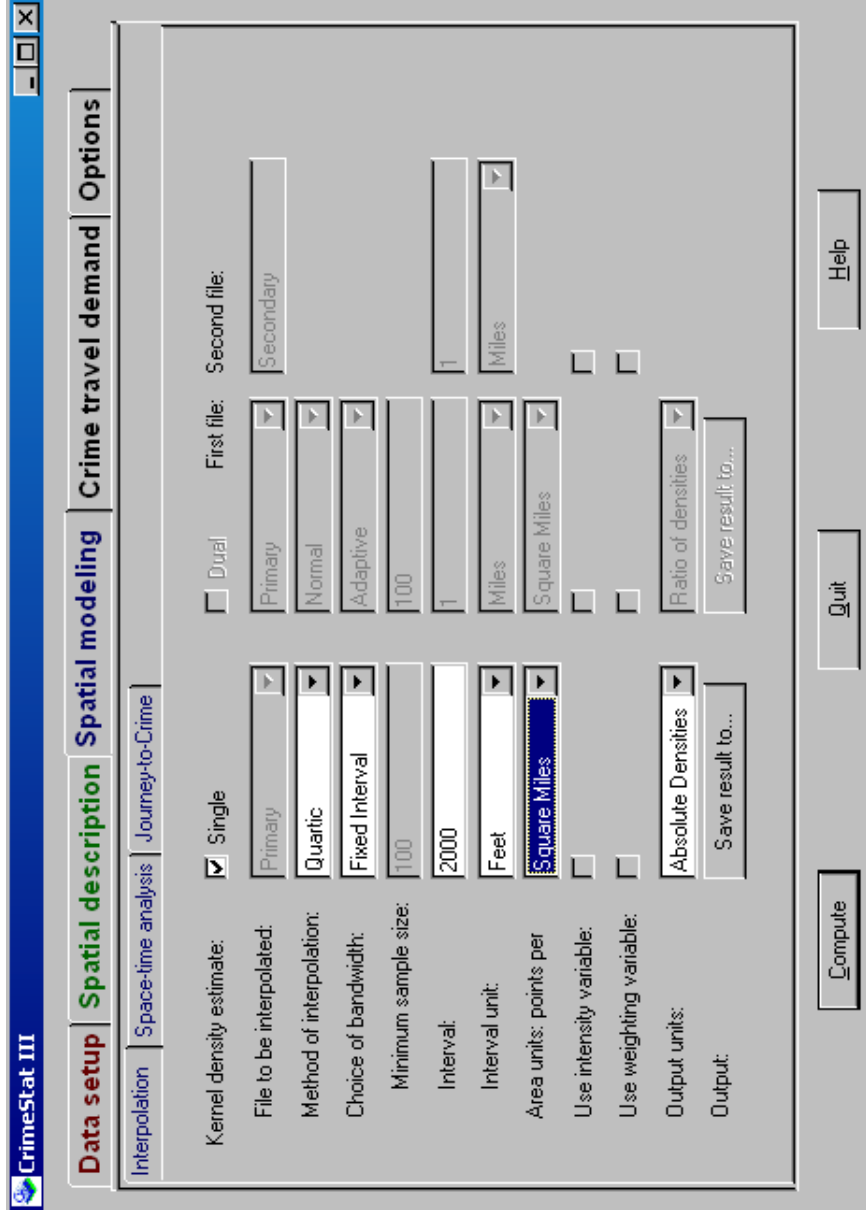


In the Data setup | Reference file tab:

Enter the X and Y coordinates in the Grid area to define your area of interest.

In Cell specification, choose By cell spacing, and pick a cell size. I'm using 400 feet, because smaller grids take longer to process. But a 250-foot grid cell would make a prettier map.

Under Create Grid, click Save, in case you want to use this grid again.



In the Spatial modeling | Interpolation tab:

Select Single.

Select Choice of bandwidth Fixed Interval

Set Interval to 2000

Set Output units to Absolute Densities.

Click Save Result to, and supply a name for a results shapefile.

You can save all settings for later reuse using Options | Save parameters

My run took about 18 minutes. It depends on your computer speed & number of grid cells.

Go to ArcView and add the results layer. (CrimeStat puts K in front of the name you supplied.)

ArcView will run faster if you reduce the number of cells. At minimum, I deleted grid cells that fall outside of the county. You could also delete all cells in which the attribute column named Z=0.

Create a thematic map using the results layer and the Z column as the value. The aim is to show areas of concentration, so I usually make a large low value range that is not shaded at all. Remove outlines from all ranges to make the map smoother.

**To do a map that represents a rate**, I repeated the above procedure with a universe file to make a grid of density of all single-family homes. Then I joined the foreclosure density to the sf home density, and created a column with the ratio of fc / sf. You can definitely delete cells where the sf.Z=0. You'll also want to set a minimum value for sf.Z for a cell to be display. This gets rid of large rates based on a small population. You can actually do this in one step in CrimeStat, however I think it's helpful to have the sf.Z value available to filter.

Most published legends I've seen just go High – Low and don't mention exact density. If you think you might need numbers, be sure to read the Chapter 8 section on Density Calculations.

