

Understanding crime hotspot maps

Melissa Burgess

The distribution of crime across a region is not random. A number of factors influence where crime occurs, including the physical and social characteristics of the place and the people using the place. Crime mapping can show us where the high crime areas are and help to provide an understanding of the factors that affect the distribution and frequency of crime. This knowledge can help improve crime prevention policies and programs. For example, it can help us to anticipate at-risk places, times and people; direct law enforcement resources; allocate victim services; design the most suitable crime prevention strategies; and so forth. This brief provides a description of how the Bureau's Local Government Area crime hotspot maps are produced and how they should be interpreted.

Data for crime mapping

The Bureau obtains crime data from the NSW Police Force's Computerised Operational Policing System (COPS). COPS contains information on all recorded criminal incidents in NSW, including the:

- Type of offence;
- Date and time of the incident;
- Location of the incident;
- Type of premises where the incident occurred;
- Whether drugs or alcohol were involved;
- Whether a weapon was used;
- The age and gender of the offender;
- The age and gender of the victim, and more.

Before we produce crime maps, the crime data needs to be geocoded. Geocoding involves assigning a geographic reference (longitude and latitude) to the incident. Geocodes are determined using street addresses, place names and other location information recorded in COPS.

Incidents that have a full address, including street number, street, suburb and postcode, can be geocoded very accurately. So too can those incidents with a location or landmark name (e.g. Hyde Park or the Myer Centre). Incidents occurring at landmarks that cover a large area (e.g. parks, schools and shopping malls) are geocoded to the centre of the landmark. Incidents where the police have not recorded a street number or place name are geocoded to the centre of the street. Incidents without a street or suburb recorded are not geocoded

at all and are therefore not used for crime mapping. Recorded criminal incidents that take place in correctional, detention or remand centres are not represented on the maps.

Key points:

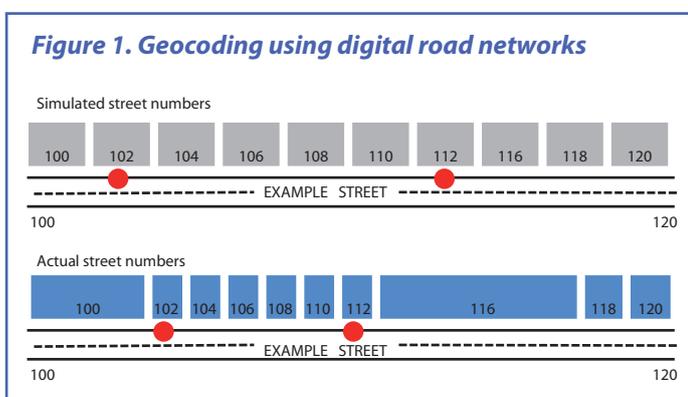
Crime maps show us where crime occurs. This helps us to understand why crime occurs in certain areas and to develop and implement strategies which reduce crime.

Data for our crime maps comes from the NSW Police Force's Computerised Operational Policing System (COPS). Each crime incident is geocoded and mapped according to the location information recorded in COPS. Care is taken to match incidents to an exact address or landmark, however sometimes this information is not available and incidents are geocoded to the centre of a street. The geocoding accuracies for all our crime maps are publically available.

The accuracy of crime maps

There are several points to consider about the geocoding process. Firstly, the quality of geocodes is dependent upon the accuracy and completeness of the information stored in COPS. If police do not record the location of an incident correctly or fail to provide all necessary location information, then poor geocodes will be produced. The Bureau routinely assesses the quality of crime incident location information, the resulting accuracy of geocodes and how incidents geocoded to the centre of the street could affect the resulting crime maps. Data improvements are made when appropriate and feedback is provided to the police.

Secondly, geocoding is also subject to data quality issues inherent in the geocoding software and in the datasets we use for geocoding (e.g. datasets on the road network). A section of a residential street from house number 100 to house number 120 is shown in Figure 1 to illustrate how an inaccurate digital road network can affect how incidents are geocoded. The top figure shows that the computer simulates the placement of street numbers along the street section by assuming that all house fronts are of equal size and that houses are evenly distributed. However, in reality, houses are of varying size with some street numbers occupying a much larger proportion of the street than others (shown in the bottom of the figure). The red points on the image show the location of two crime incidents occurring at house numbers 102 and 112. We can see how the geocodes can vary according to the on-ground distances between houses in the digital road network.



We use top-range and regularly updated commercial Geographic Information Systems and digital spatial data to ensure the best quality geocodes are obtained. However, while we take such precautions, maps should be viewed with these potential problems in mind.

For all Local Government Area (LGA) hotspot maps, the Bureau provides a breakdown of the number and percent of incidents (by offence type) that are geocoded to a landmark, street number, street centre, or not at all. Information on the accuracy of our geocoding for all crime maps, by LGA, is publically available from our website. Table 1 below shows an example of how information on geocoding accuracy is displayed. Some offence types are likely to have better geocoding than others. For example, the majority of break and enter-dwelling offences occur on premises that can be easily identified (like residential premises with a unique street address). This is reflected in the geocoding results, which indicate that 97 per cent of break and enter dwelling incidents are geocoded to a street address or a landmark and thus can be very accurately mapped. A high level of confidence can be placed on the hotspot map for break and enter dwelling incidents.

In contrast, many robbery incidents occur in public settings where the location may not be easily described. In these instances police are likely to record the street name without any additional location information. Table 1 shows that 52 per cent of robbery incidents were geocoded to a landmark or full street address, 36 per cent were geocoded to the street centre and 12 per cent were not geocoded or mapped at all. The resulting hotspot map should therefore be analysed with a degree of caution.

Table 1. Example of geocoding results for all incidents in an LGA, by offence type

Offence type		Geocoded to a landmark	Geocoded to a street number	Geocoded to the street centre	Not geocoded	Total
Assault - domestic violence related	No.	56	940	79	24	1,099
	%	5	86	7	2	100
Assault - non-domestic violence related	No.	392	506	313	111	1,322
	%	30	38	24	8	100
Robbery	No.	72	49	85	27	233
	%	31	21	36	12	100
Break and enter - dwelling	No.	16	1,375	36	8	1,435
	%	1	96	3	1	100
Break and enter - non-dwelling	No.	176	117	25	3	321
	%	55	36	8	1	100
Motor vehicle theft	No.	130	509	98	22	759
	%	17	67	13	3	100
Steal from motor vehicle	No.	267	552	99	40	958
	%	28	58	10	4	100
Steal from person	No.	97	28	32	21	178
	%	54	16	18	12	100
Steal from dwelling	No.	8	506	17	8	539
	%	1	94	3	1	100
Malicious damage to property	No.	790	2,315	340	108	3,553
	%	22	65	10	3	100

Point crime maps

The first step in hotspot mapping involves the creation of point crime maps. Point crime maps show sites of crime illustrated by dots or other symbols on a map (see Figure 2 - top). A variation of a point map is the repeat victimisation map, which shows multiple incidents occurring at the same location. Higher rates of repeat victimisation are usually displayed by an increasingly larger marker symbol (usually a graduated circle). As point maps show the specific location of incidents, they are particularly effective when examining crime in a small localised area, as shown in Figure 2 - bottom.

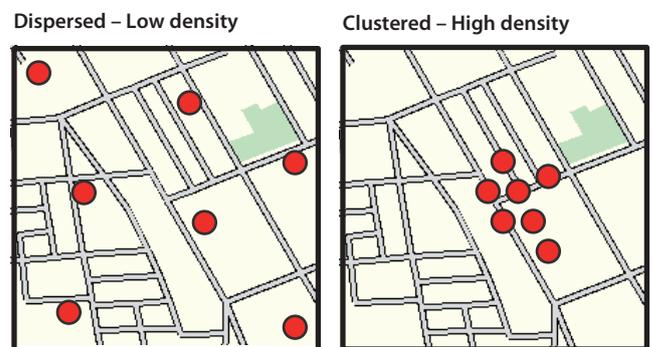
The Bureau does not make point maps publically available for several reasons. Firstly, by pinpointing the exact places where incidents occurred, point maps are highly sensitive in nature, having the potential to identify the offenders, victims and premises involved. Releasing point maps could result in a serious breach of privacy for those individuals, companies or other parties. Secondly, because crime locations are not always accurately recorded, point maps can sometimes be quite

misleading. Lastly, when there are a large number of incidents or when the map covers a large area, the points become overlapping, cluttered and illegible. It is therefore difficult to distinguish patterns in the distribution of crime or identify problematic hotspots using point maps.

Density and hotspot maps

Density mapping is used to identify crime hotspots. Crime density analyses provide a measurement of the number of crime incidents within a specified area. This number, referred to as a 'density score', provides an indication of the level of clustering and dispersion of crime incidents. When incidents are clustered within a small area, the area will have a high density score. When incidents are dispersed they are distributed over a large area, which will have a low density score (See Figure 3).

Figure 3. Clustered and dispersed crime incidents



Areas with a very high density score relative to crime concentrations across NSW are considered to be crime hotspots. Crime hotspot maps for LGAs across NSW are made publically available by the Bureau. An example of a hotspot map is shown in Figure 4. This map shows hotspots of steal from motor vehicle incidents in the Canterbury LGA, 2009. The crime hotspots are

Key points:

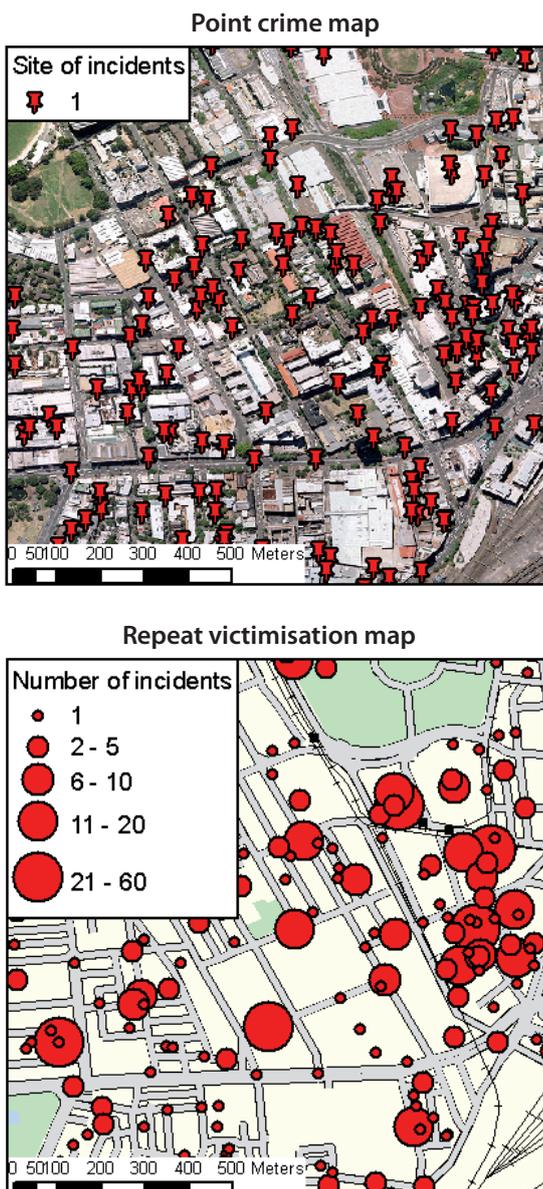
The Bureau uses point crime maps to create hotspot maps for LGAs in NSW. Hotspot maps illustrate areas of high crime density relative to crime concentrations across NSW. The hotspots indicate areas with a high level of clustering of recorded criminal incidents for the selected offence. Briefly, they are calculated as the sum of the number of all weighted incidents per 50m by 50m. Generally, less than one per cent of the land area of NSW has a high density of crime and will be classed as a hotspot.

Three levels of hotspots are indicated on our hotspot maps.

Red areas are high density hotspots, orange areas are medium density hotspot, while yellow-orange areas are low density hotspots. Despite the fact that there are high, medium and low density hotspots, all hotspots indicate areas of high crime clustering relative to all areas of NSW.

LGAs with crime hotspots do not necessarily have a high count of incidents relative to other LGAs. This is because hotspots reflect the density of incidents in specific areas, not the number of incidents in the entire LGA. Hotspots are not adjusted for the number of people residing in or visiting the LGA and so do not necessarily reflect areas where people have a higher than average risk of victimisation.

Figure 2. Example point and repeat victimisation maps



coloured from yellow-orange to red to indicate the intensity of the hotspot. While all the hotspots have a substantially high concentration of crime (compared to NSW averages) the red areas indicates the highest density hotspots, the dark orange areas indicate medium density hotspots and the yellow-orange areas indicate low density hotspots.

The technical bit: Creating the crime density and hotspot maps

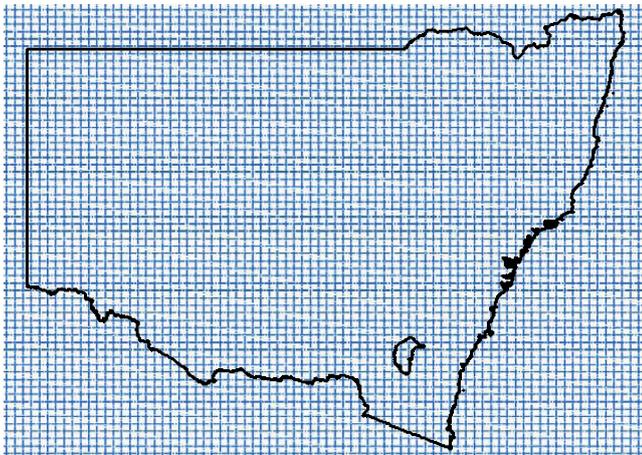
We use the 'quartic kernel density estimation' function in ESRI's ArcMap software to determine the density scores for any given area. The kernel density estimation function is a complex mathematical equation that will not be described here. However, we will illustrate the general process so that our hotspots maps can be better understood.

Calculating crime density

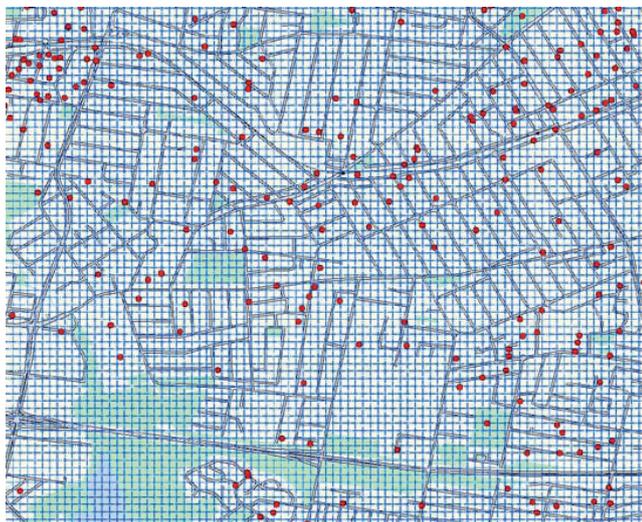
1. A rectangular grid is generated over a point crime map of NSW (see Figure 5). The grid has square cells that are 50m by 50m (2,500m²) in size.

Figure 5. Grid used to create hotspot maps

A 50m by 50m grid is generated over the whole of NSW



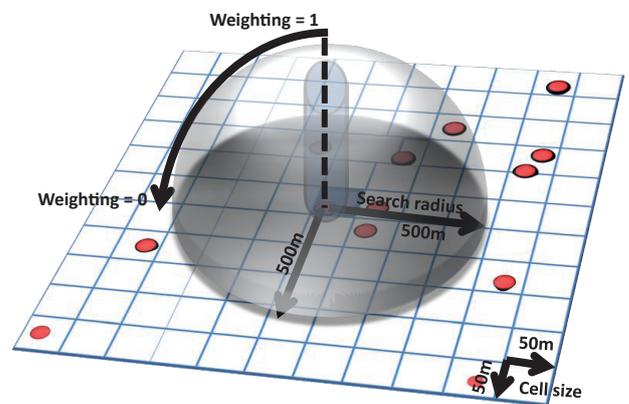
Detail of 50m by 50m grid over the point map



2. Crime density is most simply described as the number of crime incidents within a certain area. For example, if there were two robberies within one of these 50m by 50m areas (2,500m²), then the crime density would be 2/2,500, which equals 0.0008 incidents per 50m by 50m. However, crimes occurring within close proximity to one another are often related. For this reason we also take into account the number of crimes occurring in neighbouring areas.

To do this, a 'moving three-dimensional window' is placed over each 50m by 50m grid cell, referred to as the destination grid cell. This moving window can be imagined as a hemisphere with a radius of 500m (see Figure 6). All incidents within the 500m search radius of this window are used to allocate a density score to the destination grid cell.

Figure 6. Visualisation of the moving three-dimensional kernel density estimation window



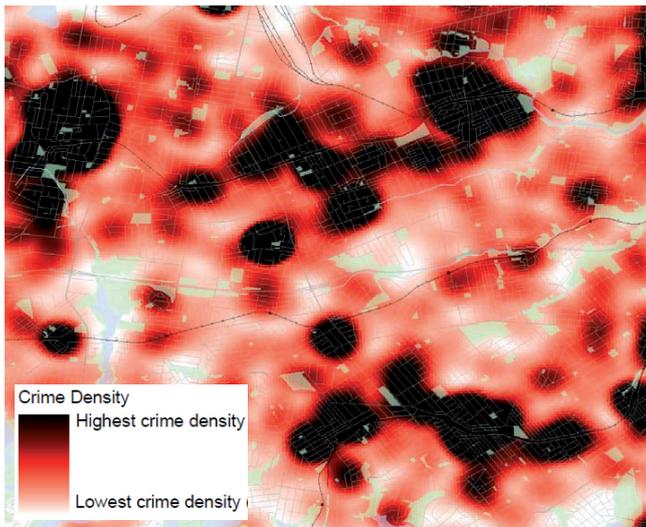
Each incident is weighted according to how close it is located to the destination grid cell. The closer the incidents are to the centre of the destination grid cell, the higher they are weighted and the more they will contribute to the overall density score. Incidents at the centre of the destination cell are weighted '1', while incidents outside the search radius are weighted '0'.

3. A density score is allocated to the destination grid cell by summing the count of weighted crime incidents that are located within 500m of the centre of the kernel window. The density score is therefore a fractional number that represents the sum of the weighted crime counts per 50m by 50m. This value can be difficult to interpret due to the complex methods used to obtain it. Nevertheless, it reflects the most widely used and recognised indication of the density of incidents per area unit.
4. The moving window is repositioned over the centre of every 50m by 50m grid cell in turn, until all the cells in NSW have been covered. Each of the hundreds of thousands of 50m by 50m grid cells are therefore assigned a density score.

Visualising crime density

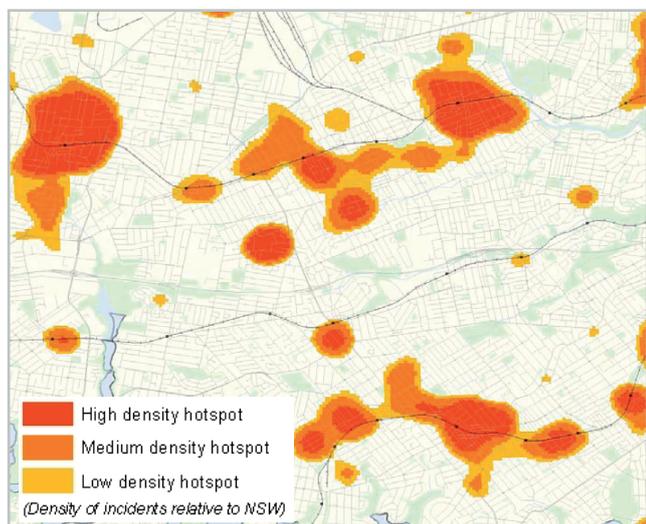
- The grid cells are then shaded according to their density scores to create a density map with a smooth surface. The surface illustrates areas of high and low crime concentration. Areas with no criminal incidents are transparent (not displayed); areas with a small number of incidents are shaded lightly and the shading becomes progressively darker as the density of criminal incidents increases. See Figure 7.

Figure 7. Example kernel density map



While density maps like the one above are very effective at showing crime concentrations in a realistic and smooth manner, we classify these maps to show only those pertinent hotspots of crime across NSW. The final result will look similar to Figure 8.

Figure 8. Example hotspot map

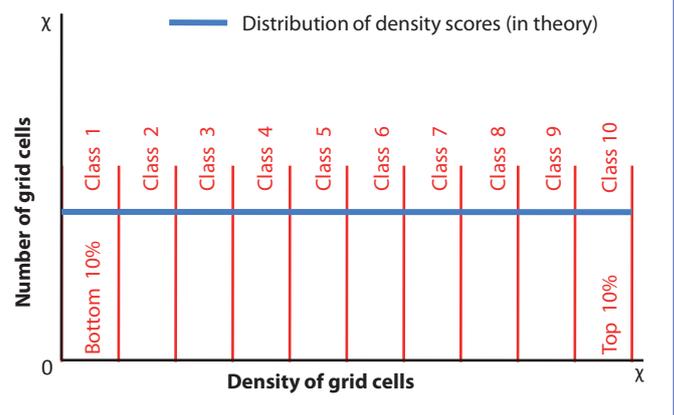


Classification of crime densities

- To classify the kernel density scores, the hundreds of thousands of density scores are listed in ascending order. The density scores are then categorized into 10 groups.

In theory, each of the 10 groups should have an equal number of density scores (grid cells) and therefore represent 1/10th of the sample or 1/10th the area of NSW. The first class (Class 1) will comprise of the bottom 10 per cent of the data and the last class (Class 10) will comprise of the top 10 per cent of the data. This is shown in Figure 9.

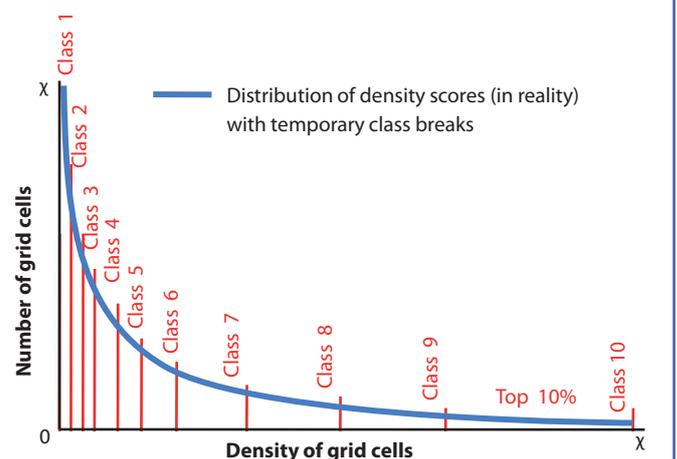
Figure 9. Distribution of density scores in theory



According to this method of classification, the top class represents those areas of NSW with the top 10 per cent of density scores. In other words, it is the 10 per cent of NSW with the highest crime density (weighted count of crimes per 50m by 50m).

- However, in reality, the distribution of crime is not linear and each unit of land area is not as equally likely to have a low crime density as a high crime density (as suggested by Figure 9). Instead, most of NSW will have no crime at all or very little crime and will therefore have a low crime density. Only a small proportion of the land area of NSW has a high crime density relative to crime concentrations throughout the state. This means that the distribution of crime is non-linear and skewed, as shown in Figure 10.

Figure 10. Distribution of density scores in reality, with temporary class breaks



8. In Figure 10, the values of the class breaks have been adjusted to ensure an equal number of density scores (grid cells) in each grouping. This ensures that Class 10 will consist of the top 10 per cent of the data.

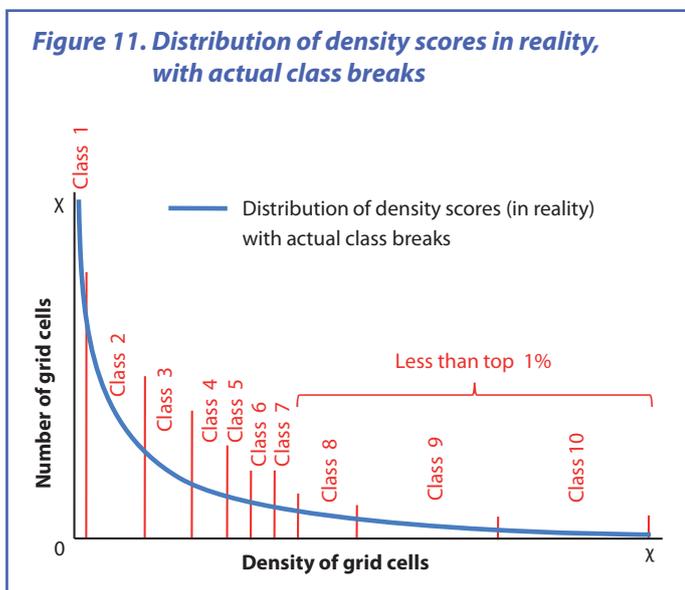
Yet in practice, crime densities cannot be broken into an equal number of groups. As noted earlier, most of the land area of NSW has no crime at all or very little crime. In fact, as much as 99 per cent of the 50m by 50m grid cells in the sample could have a density score of zero or very close to zero.

As a result, the classification method does not break cells into equal groups. Class 1 will contain all of those grid cells with a density score of zero or very close to zero. Thus Class 1 can contain as much as 99 per cent of the sample.

Because the crime distribution is such that there is much land area with similarly low crime density, the second class will also contain a large proportion of the remaining data. As much as 50 per cent of the remaining data could be classified into this group, meaning that 99.5 per cent of NSW could fall into Classes 1 and 2.

The classification process continues along the sorted dataset in this manner, assigning each density score into one of the 10 groups by attempting to make the remaining groups as equally sized as possible, yet ensuring no scores with the same or similar value are split up.

In reality, the top three classes generally contain less than 1 per cent of the data in total and therefore less than 1 per cent of the land area of NSW, see Figure 11.



Identifying crime hotspots

9. To show those areas of NSW with the highest crime density (crime hotspots) those 50m by 50m grid cells in the top three classes are coloured. The grid cells representing those areas in the top class (Class 10, which contains those cells with the highest density scores) are coloured red. The grid cells in Classes 9 and 8 (areas with the next highest crime densities) are coloured dark orange and yellow-orange respectively. Remember that the colours represent crime density (number of weighted incidents per 50m by 50m).

Key points:

The hotspot maps show areas with a high concentration of crime compared to other parts of NSW. That is, the density or number of weighted incidents per 50m by 50m.

If a suburb or LGA contains a hotspot, it does not mean that the suburb or LGA has a high count of incidents relative to other suburbs or LGAs in NSW. Counts of incidents per LGA or suburb do not consider the amount of land area in which these crime incidents occur, like the density calculations used in the creation of the hotspots do.

Hotspots are not adjusted for the number of people residing in or visiting the LGA and so do not necessarily reflect areas where people have a higher than average risk of victimisation.

Crimes in those hotspots

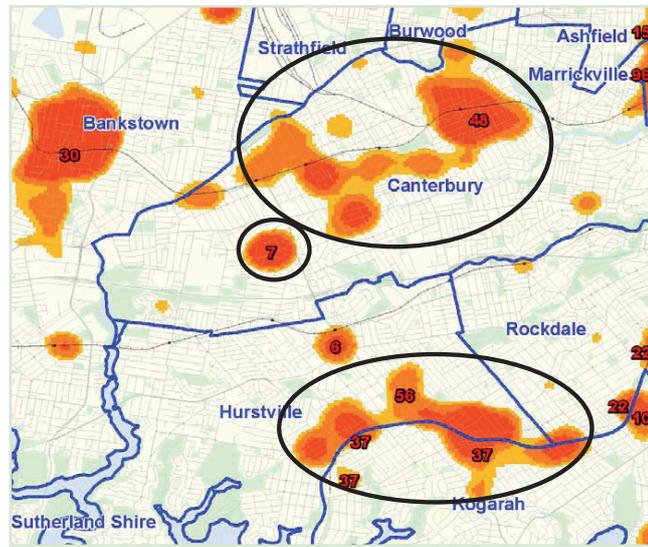
10. The final step in the creation of the hotspot maps is to provide an illustration of the proportion of crimes in an LGA that are occurring in the hotspots. The LGA boundaries are represented by the blue lines. The number next to each hotspot indicates the percentage of incidents in the LGA that are located in that particular hotspot. The hotspot is determined as the area comprising Class 8 (the yellow-orange area and class with the third highest density scores).

The number of incidents within the hotspot can be calculated by looking at the LGA hotspot map in conjunction with the crime data tables for the corresponding LGA available from the Bureau's website. Figure 12 shows a circular hotspot that contains 7 per cent of the steal from motor vehicle incidents in Canterbury LGA for 2009. There were 814 steal from motor vehicle incidents in Canterbury in 2009. Thus, 57 incidents occurred within the area defined by this hotspot in 2009. There is another large sprawling hotspot in Canterbury LGA that contains 48 per cent of the steal from motor vehicle incidents for 2009. A total of 391 incidents occurred within the area defined by this hotspot in 2009 (i.e. the area bounded by yellow-orange cells).

Several percentages (one for each LGA) are provided for hotspots that cross LGA boundaries. Figure 12 shows a large hotspot crossing between Hurstville, Kogarah and Rockdale LGAs. You can see that 56 per cent of the steal from motor vehicle incidents in Hurstville in 2009 took place within the portion of the hotspot located within Hurstville LGA. There were 415 steal from motor vehicle incidents in Hurstville LGA during 2009 and thus it can be determined that there were 232 incidents within this section of the hotspot. For Kogarah, there were 274 steal from motor vehicle incidents in the LGA in 2009. Figure 12 shows us that 37 per cent of these incidents were located within the Kogarah section of the hotspot, which equates to a total of 101 incidents. You will notice that the '37' is written three times. This is because the section of the hotspot within Kogarah LGA is split by the LGA boundaries into three parts. The 37 per cent provides a total for those three parts of the hotspot within Kogarah LGA.

The Rockdale section of the hotspot has no per cent value because fewer than 5 per cent of incidents occurred in that hotspot.

Figure 12. Example hotspot maps with hotspot percentages



Key points:

Remember that hotspots are determined by crime density and not crime counts. The unit of analysis is weighted counts of crime incidents per 50m by 50m. Thus a hotspot can be created from very few incidents if they are concentrated in a small area.

Also it is important to note that although the hotspot boundaries have been clearly defined for illustrative purposes, crime density varies gradually in reality and that incidents within 500 metres from the hotspot boundaries were also used in the calculations for identifying hotspots (as noted above).

Why is my LGA all red?

A large proportion of the land within some LGAs will be covered by hotspots. These LGAs will have a high crime density relative to crime concentrations across the state. Often these LGAs will also have a large residential population and/or number of visitors to the area. It is important to remember that even though the crime density is high, the risk of victimisation may not be high once these populations are taken into account.

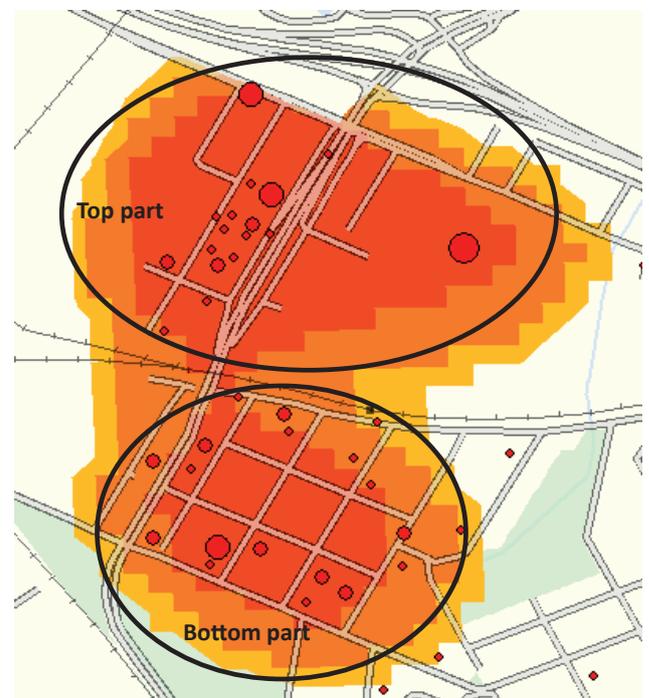
My LGA has the same number of crimes as another LGA, yet my LGA has hotspots and the other LGA does not. Why?

Again, the presence of hotspots represents the level of crime clustering occurring in an area. Many of the incidents in your LGA may be clustered into a small area, which will have a high crime density. The same number of incidents may be dispersed over a much wider area in another LGA, in which case the crime density would be lower and no hotspot formed.

There is a very large hotspot in my LGA. Do all the incidents occur in the centre of that hotspot?

Not necessarily. This is demonstrated in the hotspot in Figure 13, which has two primary parts. The crime incidents in the top part of the hotspot occur in two main areas, while the crime incidents in the bottom part of the hotspot occur around the edges of this part of the hotspot. In many other hotspots, the crime incidents will occur in the centre of the hotspot.

Figure 13. Example hotspot maps with crime incidents



Frequently asked questions

What if there are no hotspots in my LGA?

Areas with low crime density relative to crime concentrations across NSW (i.e. in the bottom 7 of 10 Classes) are not shown in the LGA hotspots maps. Therefore there are several LGA maps for selected offences that do not show hotspots because there are no areas in the LGA with a high density of crime.

There are low crime counts in my LGA, so why is there a crime hotspot?

The hotspot maps reflect crime density rather than incident numbers. LGAs with high numbers of incidents compared to other LGAs will not have a crime hotspot if its crime incidents do not concentrate in any particular area. Likewise, LGAs with low numbers of criminal incidents may have crime hotspots if the crime incidents in question concentrate in a specific area.