

CLOSE ADS

OSE

SAMSUNG

Samsung TV almanız için birçok sebep var

Üstün görüntü ve ses kalitesine sahip Samsung TV'lerdeki 25.000 TL'ye varan Para Yüklü Kart fırsatını kaçırmayın.

17 TILAK

Büyüye 1 Numara

25.000 TL'ye varan Para Yüklü Kart

KE

Kampanya Tanıtımı

*Gösteriler temsilidir. OI belirtilmiştir. Kampanya şartları arasında geçerli w.samsung.com/tyjette st-onale-2023/ Bir num 2023 yılında yapılan alışverişte televizyon g brezindeki satış adı göstermektedir.

Heatmap Using Kernel Density Estimation

Monday, November 20, 2017

AUTHOR geomatics

Follow @ideageo

visualization method to display event density or occurrence. Heatmap clustering points where more points in an area will have higher value point in the same area. Therefore, with a heatmap we can see a event's occurrence. For example criminal events in a city can be map, so one knows how the crime events spread around the city. Such called a crime map.

map is quite easy and straight forward. There are many software and used to produce a heatmap like QGIS, ArcGIS, Crimestat, [Google table](#) need to upload point dataset, setting some parameters and the result the purpose of this post is not just to produce a heatmap with some an that to explore and trying to get more understanding what is going

on behind the screen for data and parameters that have been specified. In another term, how the algorithm is working in producing a heatmap.

QGIS is an open source GIS software that can be used to produce a heatmap from a set of data point with [Heatmap Plugin](#). The plugin is using Kernel Density Estimation algorithm for creating a heatmap. Because of that I will discuss how this algorithm(Kernel Density Estimation) is applied to process an input point dataset into a heatmap.

Creating A Heatmap in QGIS

To get an insight for what we will discuss, firstly let's create a heatmap of theft from vehicle incident across Vancouver, Canada.

1. Download crime data of Vancouver city 2016 from [Vancouver city data catalogue](#). For convenient I provided the data [here](#). (Actually the data consist of several crime events such as mischief, theft from vehicle, break and enter residential and other theft). I just selected theft from vehicle crime.
2. Add the data into QGIS map window as in figure 1.

BEYMEN COM

GIVE

ALİŞVER

Islak Süpürmeyle Tanı

SEARCH



CLOSE ADS

OSE

SAMSUNG

Samsung TV almanız için birçok sebep var

Üstün görüntü ve ses kalitesine sahip Samsung TV'lerdeki 25.000 TL'ye varan Para Yüklü Kart fırsatını kaçırmayın.

17 TILAK

Büyüye 1 Numara

25.000 TL'ye varan Para Yüklü Kart

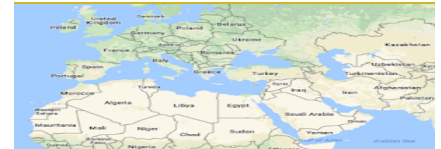
KE

Kampanya Tanıtımı

*Gösteriler temsilidir. OI belirtilmiştir. Kampanya şartları arasında geçerli w.samsung.com/tyjette st-onale-2023/ Bir num 2023 yılında yapılan alışverişte televizyon g brezindeki satış adı göstermektedir.

Subscribe to Our Newsletter!

POPULAR POSTS



- 1 This is How to Add Google Maps Layers in QGIS 3
- 2 How to Create Contour Lines in QGIS
- 3 Adding Free Satellite Imagery Layer in QGIS
- 4 Build Your Own Flight Tracking Application with Python and Open Air Traffic Data
- 5 How to Create DEM and Elevation Profile in QGIS

Reklam kaldırıldı. [Ayrıntılar](#)

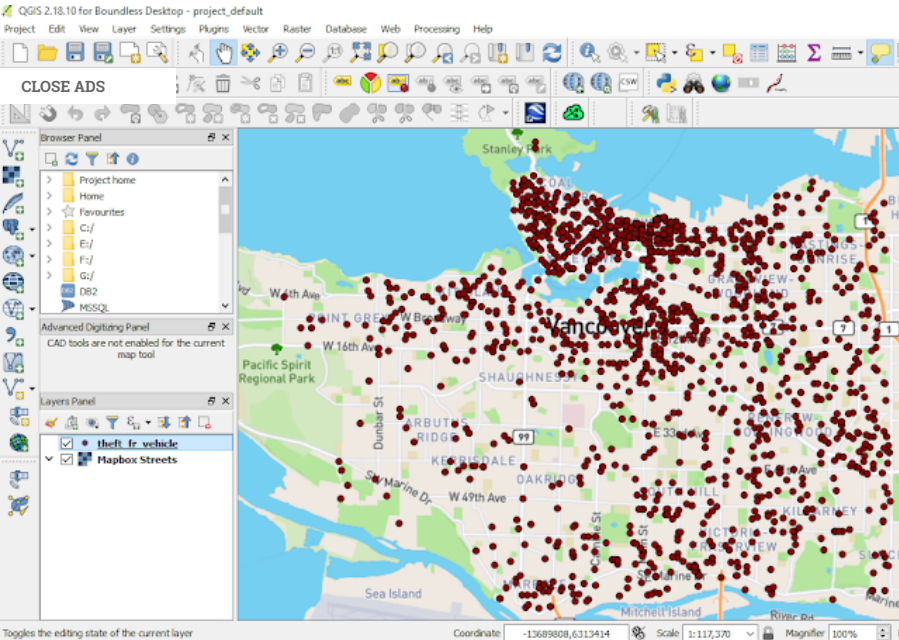


Figure 1. Theft from vehicle accident in Vancouver city 2016

3. Select Heatmap plugin. It can be found in menu Raster >> Heatmap >> Heatmap...(See figure 2)

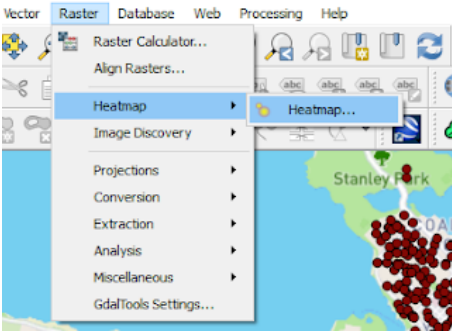
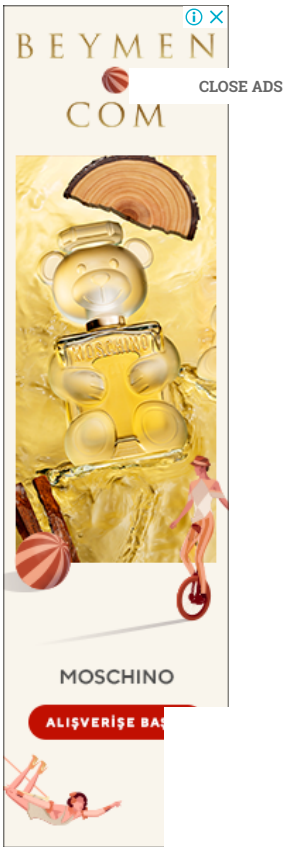
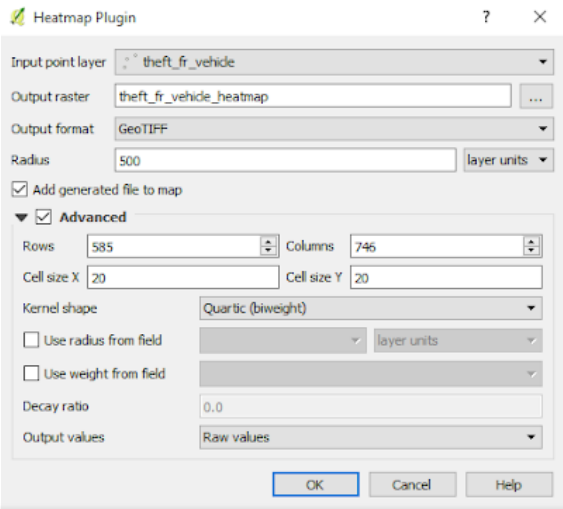


Figure 2. Heatmap plugin

4. The heatmap plugin window will appear as in figure 3. Make sure the Input point layer is the theft_fr_vehicle. Give a name for Output raster. I named it theft_fr_vehicle_heatmap. Then set Radius 500. Clicked Advanced option, set Cell size X 20 and Cell size Y 20. Leave the Kernel shape as Quartic(biweight) and the other parameter as default.



LABEL	
Tutorial	(75)
QGIS	(52)
Geoanalytics	(15)
Geonews	(10)
Google Maps API	(7)
KML	(2)





Figure 5. Theft from vehicle heatmap

6. Hmm...it looks ugly, doesn't it? So let's make it more colorful. Right click the heatmap layer. In Layer Properties window as in figure 6, choose Style. Change Render type to Singleband pseudocolor. Then choose your favorite color in the Color option.

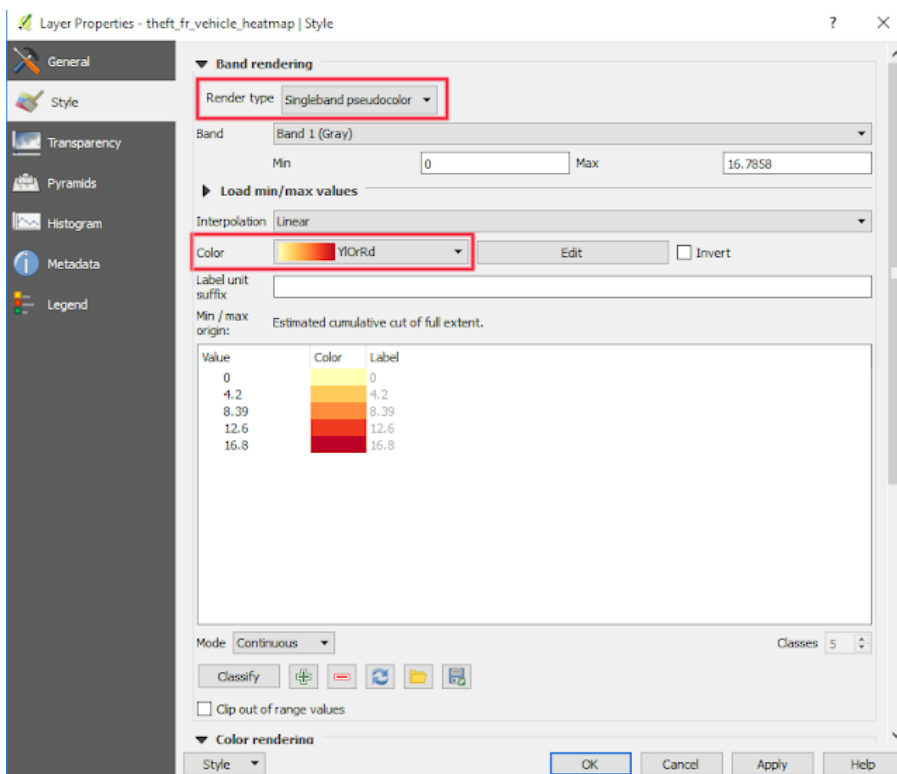


Figure 6. Change heatmap color style

8. Next goes to figure 7. Still in Layer Properties, here we change the blending mode. Choose Darken (You can experiment with other blending option). Click apply to see the result immediately in the thumbnail preview. Finally push OK button when done. The result of heatmap with the new style can be seen in the figure 8.

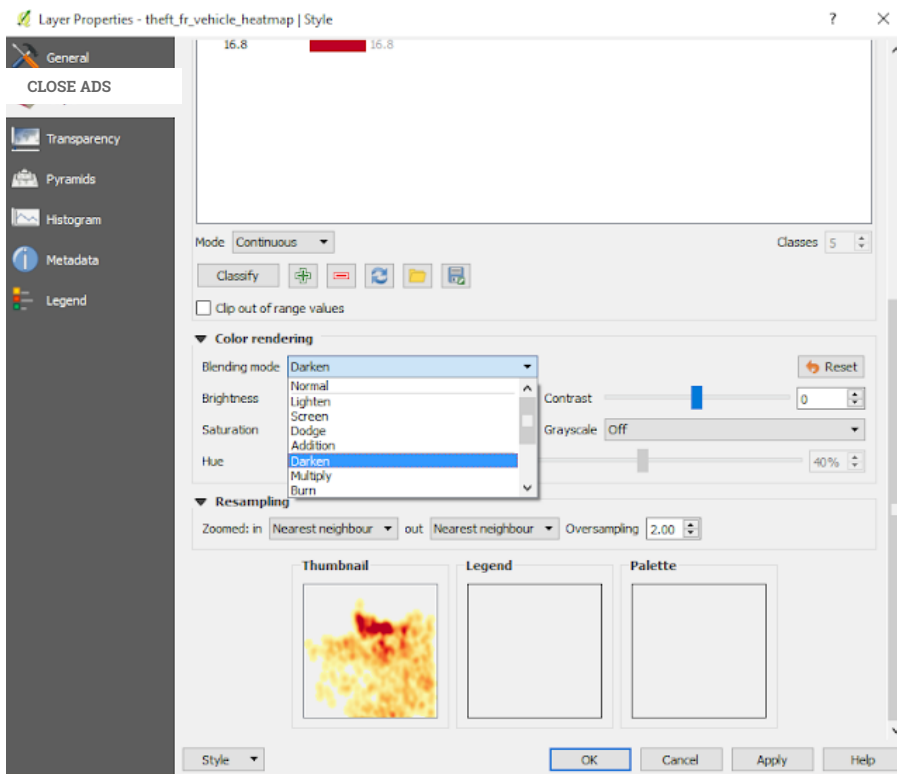


Figure 7. Change Heatmap color blend

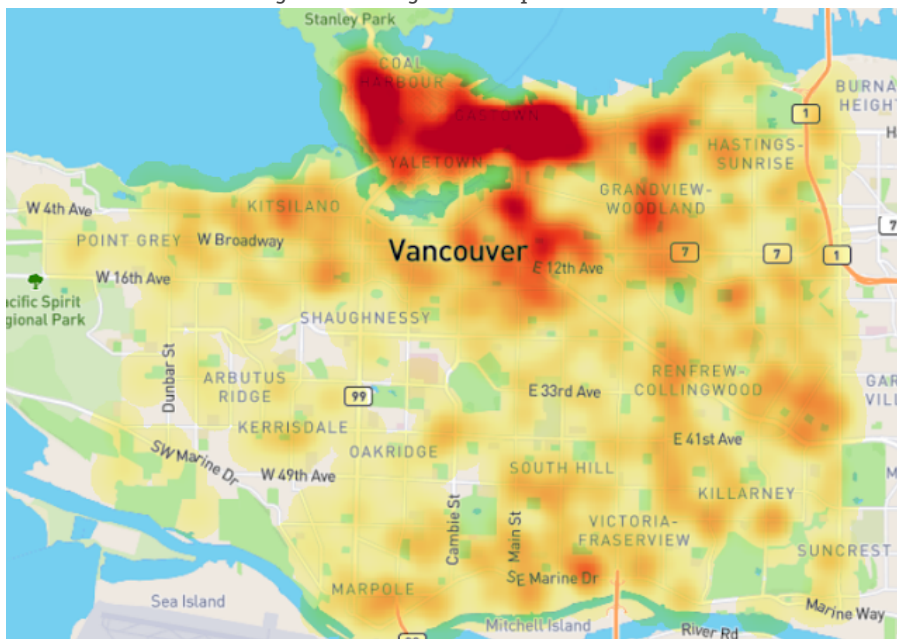


Figure 8. Theft from vehicle heatmap with singleband pseudocolor

It's done. The theft from vehicle heatmap of Vancouver city shows us how the incidents are spread in the city. In the north of the city can be seen the color is darker which indicates high density value. So take care when you park a car there :).

Kernel Density Estimation Algorithm

parameters to get a more meaningful heatmap cause we understand how the result

comes in
CLOSE ADS

CLOSE ADS

Kernel Shape

Estimation is predicting an unknown value at a location from reference points. Actually in predicting the unknown value we interpolate the value from known points value. Therefore we also called KDE as Kernel Density Interpolation. In estimation a point value, KDE uses a Probability Density Function (PDF). What is PDF and how could it be used to estimate an unknown value? To answer this question, I will give an illustration. Assume there is a road segment and I want to know how long it is. So I measure it with a tape, first measurement (forward) I got 100.251 m. Then backward I got 100.310 m? Hmm...why they are different. I am curious and take the third measurement, got the figure 100.279 m. How could this happen? The answer is Error. In every measurement there is an error. Nobody knows a real measurement value except God. But what could I do to get a value closest to the real value? Take more measurement!!! More measurement, more number of values that close the real value, thus it has more probability to get the real value which is my expected value. But I don't want to waste my time measuring the road 1000 times.....

The same problem already happened hundred years ago. In the 17th century, Galileo made an astronomical observation and noted that the errors from the observation were symmetric and small errors occurred more frequently than large errors. This led to several hypothesized of distribution errors. Later in 1809 Gauss developed the distribution normal function and showed that errors were fitted well by this distribution. (Read the history [here](#)). The function of normal distribution, which is also called Gauss distribution can be seen in equation 1. If this function were plotted with some defined parameters, we can see how it looks like in figure 9.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Equation 1. Distribution Normal function

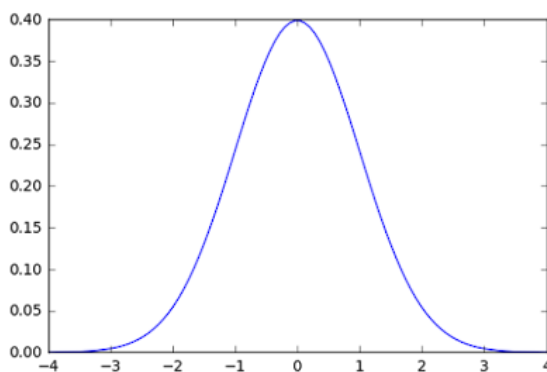


Figure 9. Normal distribution curve

The normal distribution curve looks like a bell, and simply in KDE we call this as Kernel Shape. If we look at heatmap plugin, there are some Kernel shapes available, there are: Quartic, Triangular, Uniform and Epanechnikov. But there are more kernel shapes available like Cosine, Gaussian, Tricube, etc. Some of these shapes can be seen in figure 10.

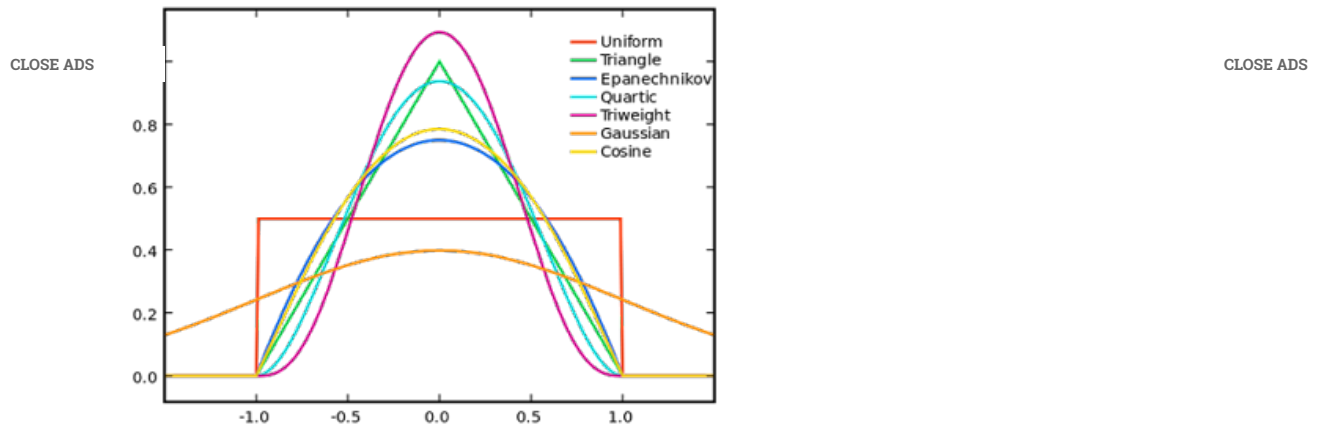


Figure 10. Kernel Shapes (Source: <https://en.wikipedia.org/>)

Next question. How those shapes are used in estimation of unknown point? As we already knew, the density value is defined by a Kernel shape with a function. To apply it for a heatmap, the function must be transformed into a "geospatial sense" form, because it will be used to predict an unknown value at a point from a point with known value. Here the distance between known to unknown point is a parameter which must be involved in the function. But where we will put it in the function? For example, in the normal distribution function in equation 1, we can see there is $(x-\mu)$ term. This term is the distance between a value (x) with the expected value (μ). This term can be translated as the distance between known to unknown point (d). Then σ is standard deviation. Standard deviation in statistics measures variation of dataset values. Actually it is the average distance of x to μ . So this term in KDE is translated into bandwidth (h). For Gaussian, h remains as standard deviation but for other kernels h is radius. Hence the equation of KDE with Gaussian Kernel shape has the form as in equation 2, with the visual illustration can be seen in figure 11. If we standardize the h parameter into 1 (100%), the equation 2 can be simplified into equation 3.

$$P(x) = \frac{1}{h\sqrt{2\pi}} e^{-d^2/2h^2}$$

Equation 2. Gaussian KDE function

$$P(x) = \frac{1}{\sqrt{2\pi}} e^{-d^2/2}$$

Equation 3. Standardize Gaussian KDE Function

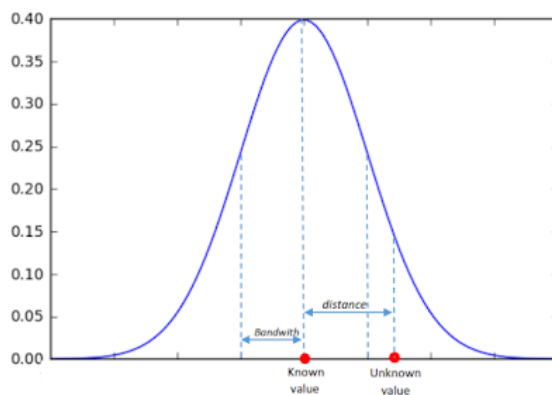

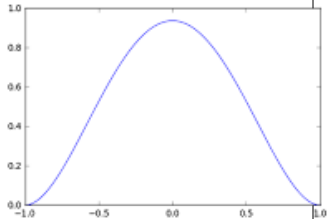
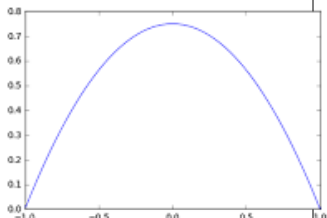
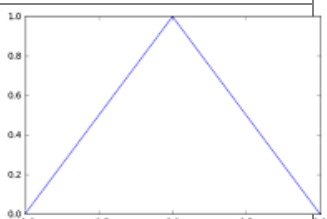
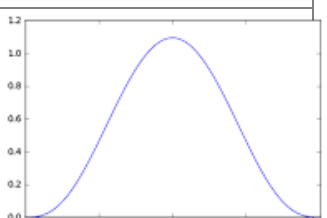
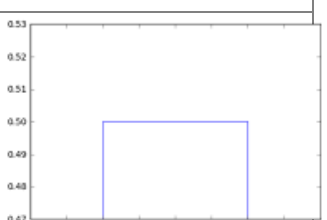


Figure 11. Gaussian Kernel Shape

From figure 11 we can observe that the density value of unknown point will decrease smoothly following the Gaussian Probability Density Function (PDF). It means the unknown value will deviate largely as it moves away from a known point. Other Kernel Shape functions are summarized in the following table.

Kernel	Function	Shape
Gaussian	$P(x) = \frac{1}{\sqrt{2\pi}}e^{-d^2/2}$	
Quartic	$P(x) = \frac{15}{16}(1 - d^2)^2$	
Epanechnikov	$P(x) = \frac{3}{4}(1 - d^2)$	
Triangular	$P(x) = (1 - d)$	
Triweight	$P(x) = \frac{35}{32}(1 - d^2)^3$	
Uniform	$P(x) = \frac{1}{2}$	

How to Choose A "Right" Kernel Shape


This is the hardest part in this post. How to choose a right kernel shape? Honestly I don't know the exact answer. But after some research I found a hint in determining a kernel shape. Based on Crimestat III workbook, there are two questions to answer in deciding a

good option. Otherwise if the accident will have the same effect throughout a specified radius the Uniform kernel would be a good option. So understanding the two questions above for a case will give a better decision in choosing a kernel shape and related parameters.

The following table gives kernel shape option for some types of crime. Although it is for crime mapping, hopefully it could give the idea in defining a kernel shape for other cases. The table was taken from Crimestat III workbook.

Incident Type	Interval (mile)	Interpolation Method	Reasoning
Residential burglaries	1	Moderately dispersed: quartic or uniform	Some burglars choose particular houses, but most cruise neighborhoods looking for likely targets. A housebreak in any part of a neighborhood transfers risk to the rest of the neighborhood.
Domestic violence	0.1	Tightly focused: negative exponential	Domestic violence occurs among specific individuals and families. Incidents at one location do not have much chance of being contagious in the surrounding area.
Commercial robberies	2	Focused: triangular or negative exponential	A commercial robber probably chooses to strike in a commercial area, and then looks for preferred targets (banks, convenience stores) within that area. The wide area may thus be at some risk, but the brunt of the weight should remain with the particular target that has already been struck.
Thefts from vehicles	0.25	Dispersed: Uniform	If a parking lot experiences a lot of thefts from vehicles, your GIS will probably geocode them at the center of the parcel. This method ensures that the risk disperses evenly across the parcel and part of the surrounding area (which probably makes sense)—but not too far, since we know that parking lots tend to be hot spots for specific reasons.

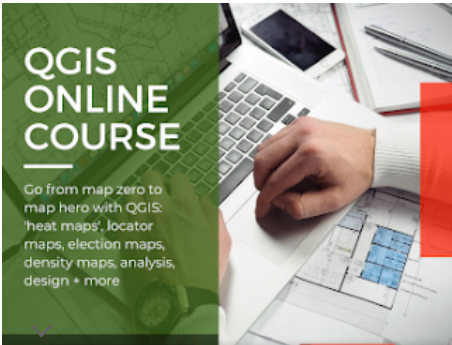
I hope this post could give you an understanding what is a Kernel Density Estimation and its application in creating a heatmap. In the [next post](#), I give a numerical example how this algorithm is used to estimate an unknown value from a reference point. So you will get a complete understanding from theory and its implementation.

 Buy me a coffee

Related Posts

CLOSE ADS

CLOSE ADS



-->

CLOSE ADS

CLOSE ADS

