How can I create kernel density plots in R with examples?

Authored by

stats writer

July 2, 2024

Kernel density plots, also known as density plots, are graphical representations that show the distribution of data over a continuous interval. These plots are useful for visualizing the shape and spread of a dataset, and can help identify patterns and outliers.

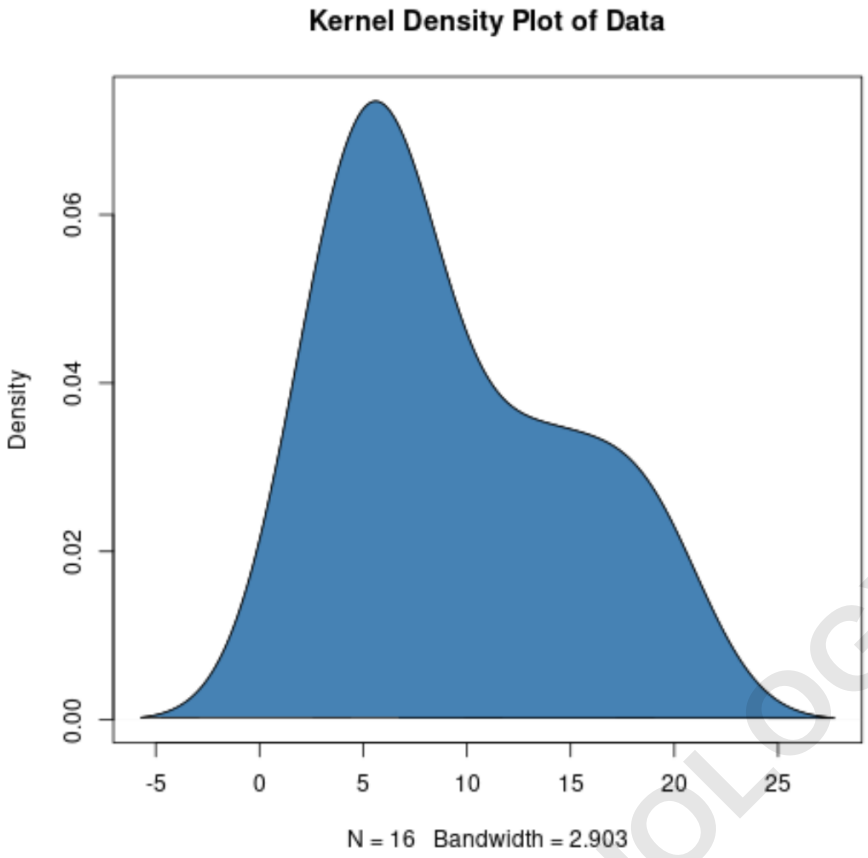
To create a kernel density plot in R, you can use the "density" function from the base graphics package or the "ggplot" function from the ggplot2 package. Both methods require a vector of data points as input.

For example, if you have a dataset of heights for a population, you can use the density function to create a plot that shows the distribution of heights. This can be done with the following code:

heights

Create Kernel Density Plots in R (With Examples)

A kernel density plot is a type of plot that displays the distribution of values in a dataset using one continuous curve.



A kernel density plot is similar to a , but it's even better at displaying the shape of a distribution since it isn't affected by the number of bins used in the histogram.

We can use the following methods to create a kernel density plot in R:

Method 1: Create One Kernel Density Plot

```
#define kernel density
kd <- density(data)
```

```
#create kernel density plot
plot(kd)
```

Method 2: Create a Filled-In Kernel Density Plot

```
#define kernel density
kd <- density(data)

#create kernel density plot
plot(kd)

#fill in kernel density plot with specific color
polygon(kd, col='blue', border='black')
```

Method 3: Create Multiple Kernel Density Plots

```
#plot first kernel density plot
kd1 <- density(data1)
plot(kd1, col='blue')

#plot second kernel density plot
kd2 <- density(data2)
lines(kd2, col='red')

#plot third kernel density plot
```

```
kd3 <- density(data3)
lines(kd3, col='purple')
```

...

The following examples show how to use each method in practice.

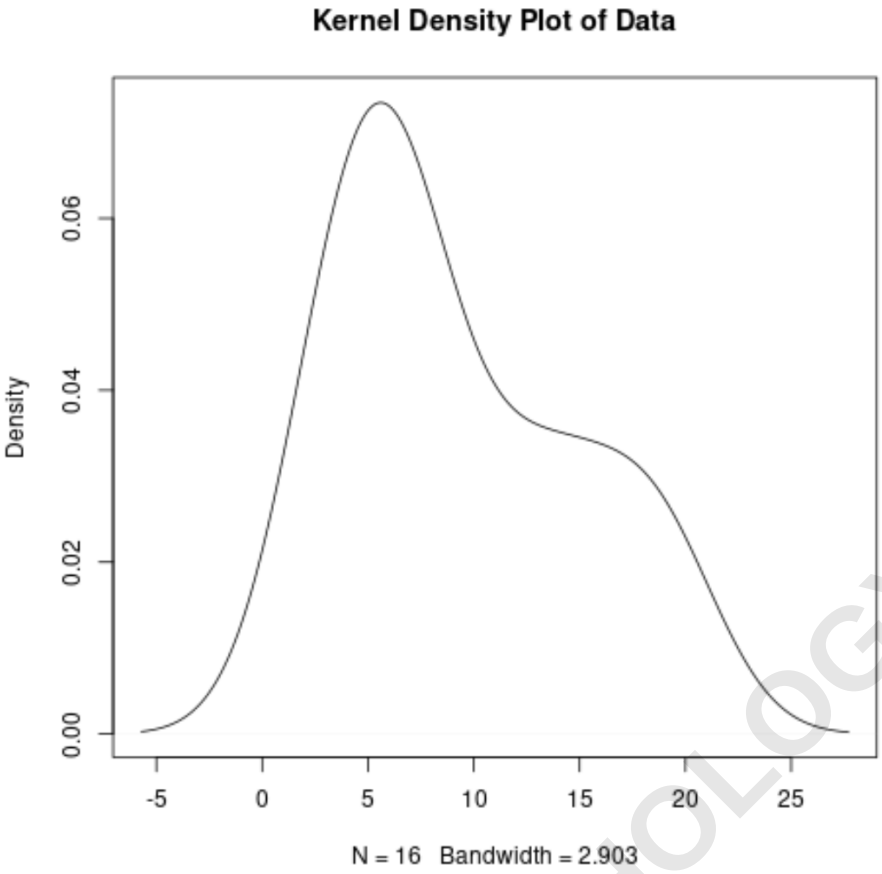
Method 1: Create One Kernel Density Plot

The following code shows how to create a kernel density plot for one dataset in R:

```
#create data
data <- c(3, 3, 4, 4, 5, 6, 7, 7, 7, 8, 12, 13, 14, 17, 19, 19)

#define kernel density
kd <- density(data)

#create kernel density plot
plot(kd, main='Kernel Density Plot of Data')
```



The x-axis shows the values of the dataset and the y-axis shows the relative frequency of each value. The highest points in the plot show where the values occur most often.

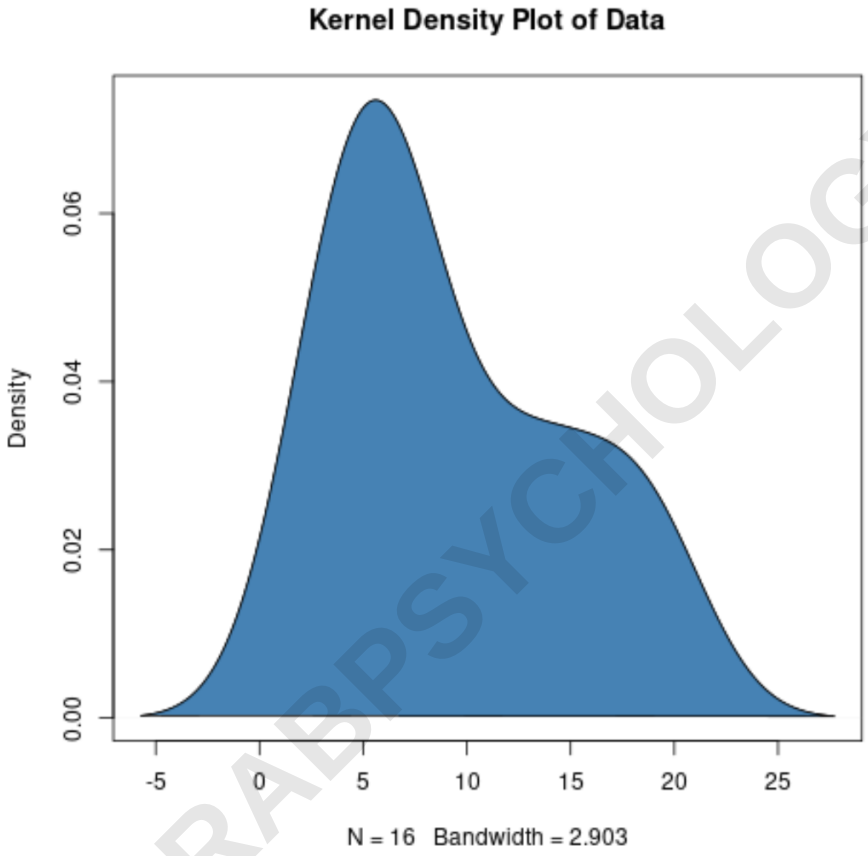
Method 2: Create a Filled-In Kernel Density Plot

```
#create data
data <- c(3, 3, 4, 4, 5, 6, 7, 7, 7, 8, 12, 13, 14, 17, 19, 19)

#define kernel density
kd <- density(data)
```

```
#create kernel density plot
plot(kd)

#add color
polygon(kd, col='steelblue', border='black')
```



Method 3: Create Multiple Kernel Density Plots

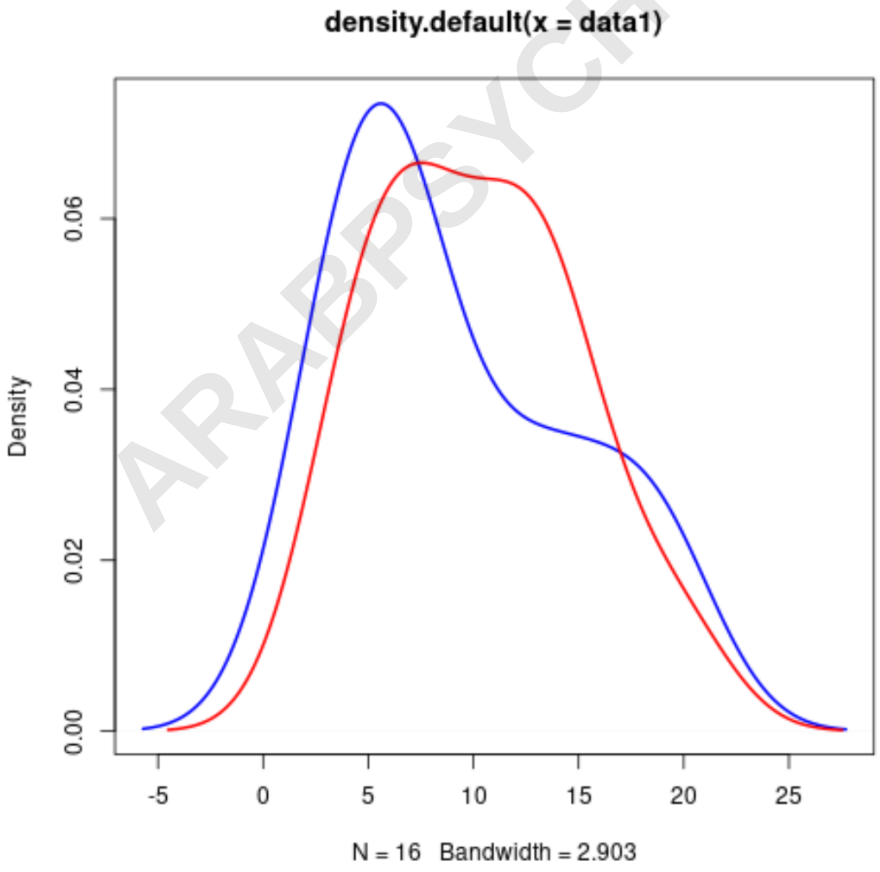
The following code shows how to create multiple kernel density plots in one plot in R:

```
#create datasets
```

```
data1 <- c(3, 3, 4, 4, 5, 6, 7, 7, 7, 8, 12, 13, 14, 17, 19, 19)
data2 <- c(12, 3, 14, 14, 4, 5, 6, 10, 14, 7, 7, 8, 10, 12, 17, 20)

#plot first kernel density plot
kd1 <- density(data1)
plot(kd1, col='blue', lwd=2)

#plot second kernel density plot
kd2 <- density(data2)
lines(kd2, col='red', lwd=2)
```



Note that we can use similar syntax to create as many kernel density plots in one chart as we'd like.

Additional Resources

The following tutorials explain how to create other common plots in R: