

## Histograms and Density Plots

### Source

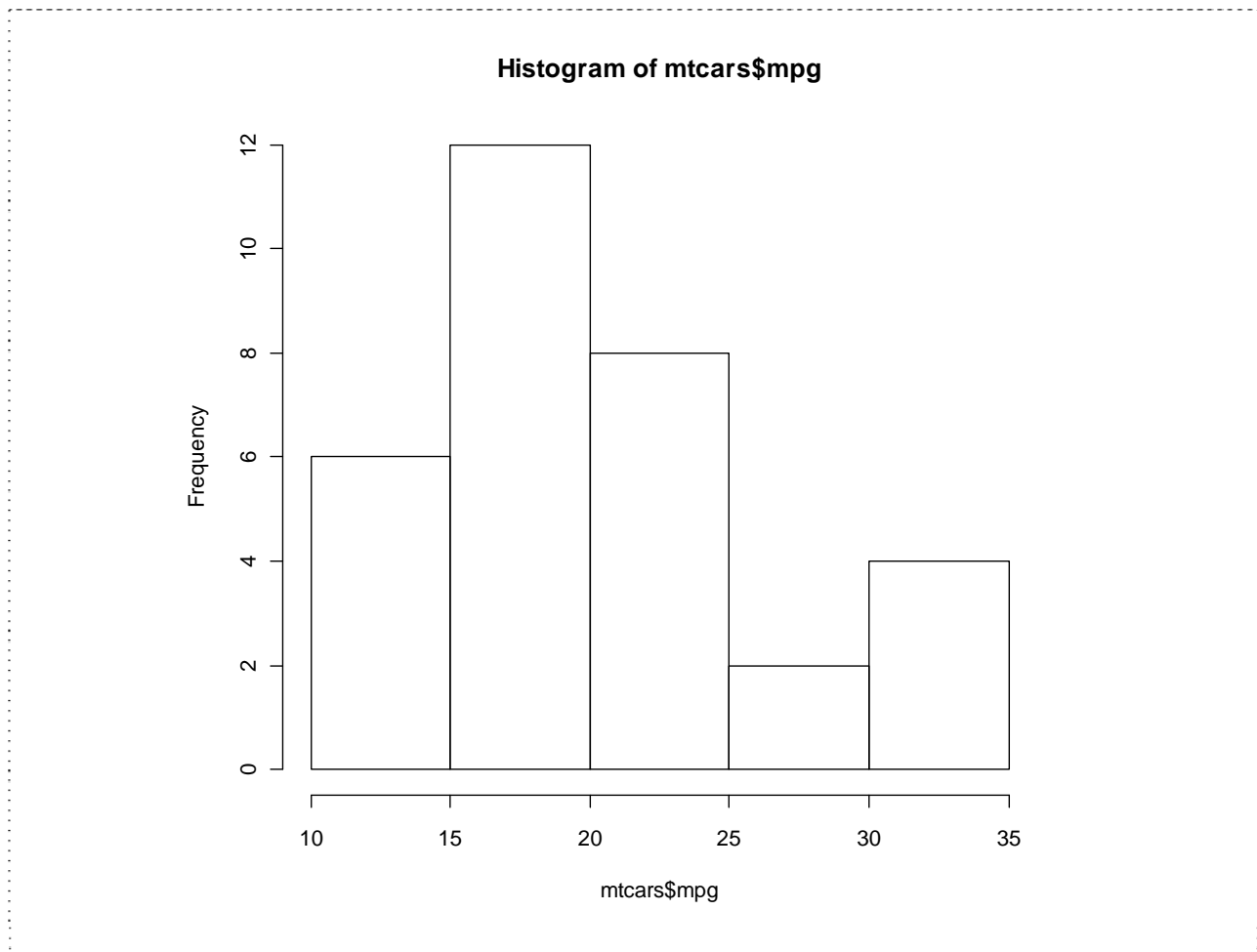
Reprinted with permission from Quick-R (<http://www.statmethods.net>).  
© Copyright 2008 Robert I. Kabacoff, Ph.D. All rights reserved.

### Histograms

You can create histograms with the function `hist(x)` where `x` is a numeric vector of values to be plotted. The option `freq=FALSE` plots probability densities instead of frequencies. The option `breaks=` controls the number of bins.

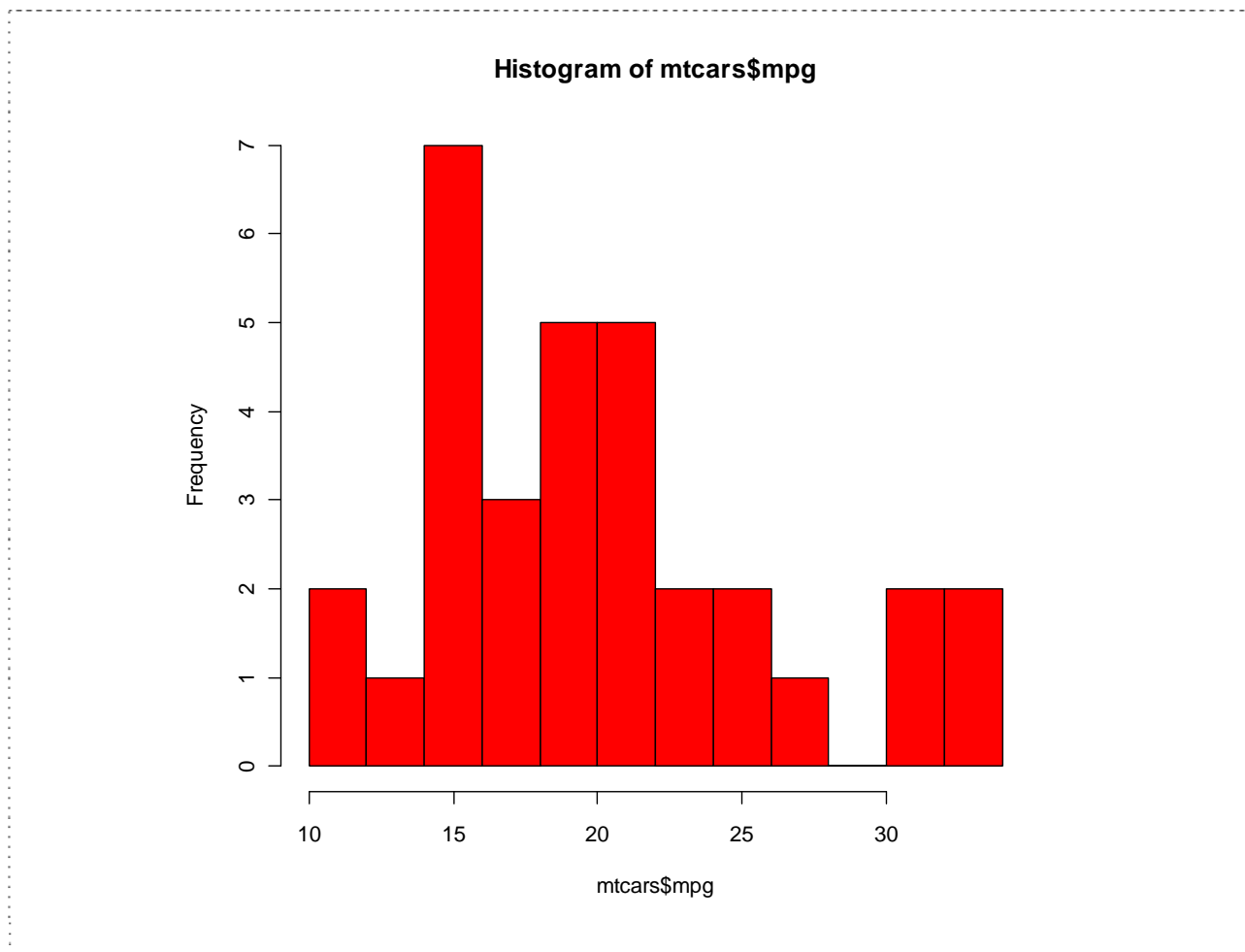
Simple Histogram:

```
hist(mtcars$mpg)
```



Colored Histogram with Different Number of Bins:

```
hist(mtcars$mpg, breaks=12, col="red")
```

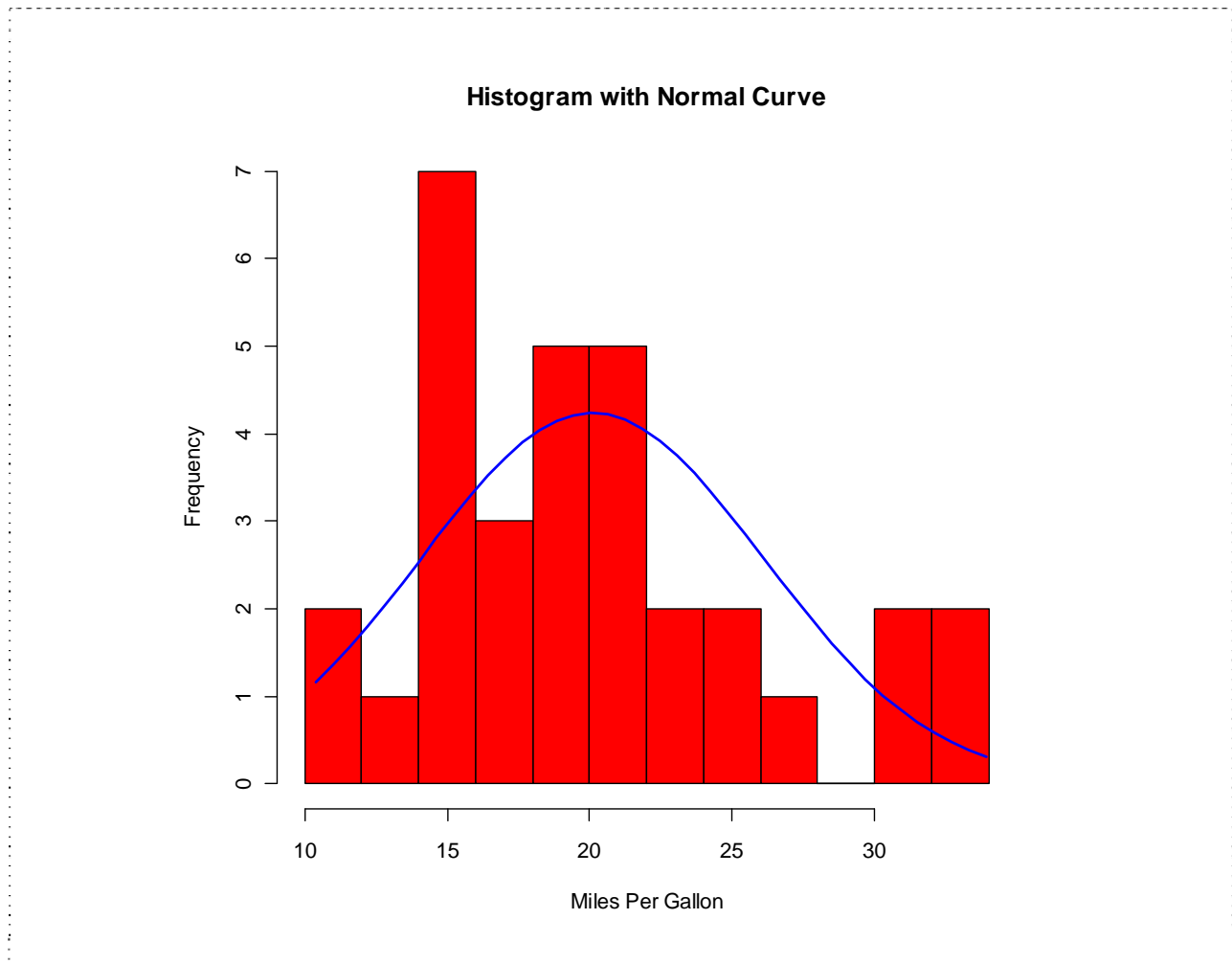


Add a Normal Curve (Thanks to Peter Dalgaard):

```
x <- mtcars$mpg

show <- function()
{
  h <- hist(
    x,
    breaks=10,
    col="red",
    xlab="Miles Per Gallon",
    main="Histogram with Normal Curve"
  )
  xfit <- seq(min(x),max(x),length=40)
  yfit<-dnorm(xfit,mean=mean(x),sd=sd(x))
  yfit <- yfit*diff(h$mids[1:2])*length(x)
  lines(xfit, yfit, col="blue", lwd=2)
}
```

show()



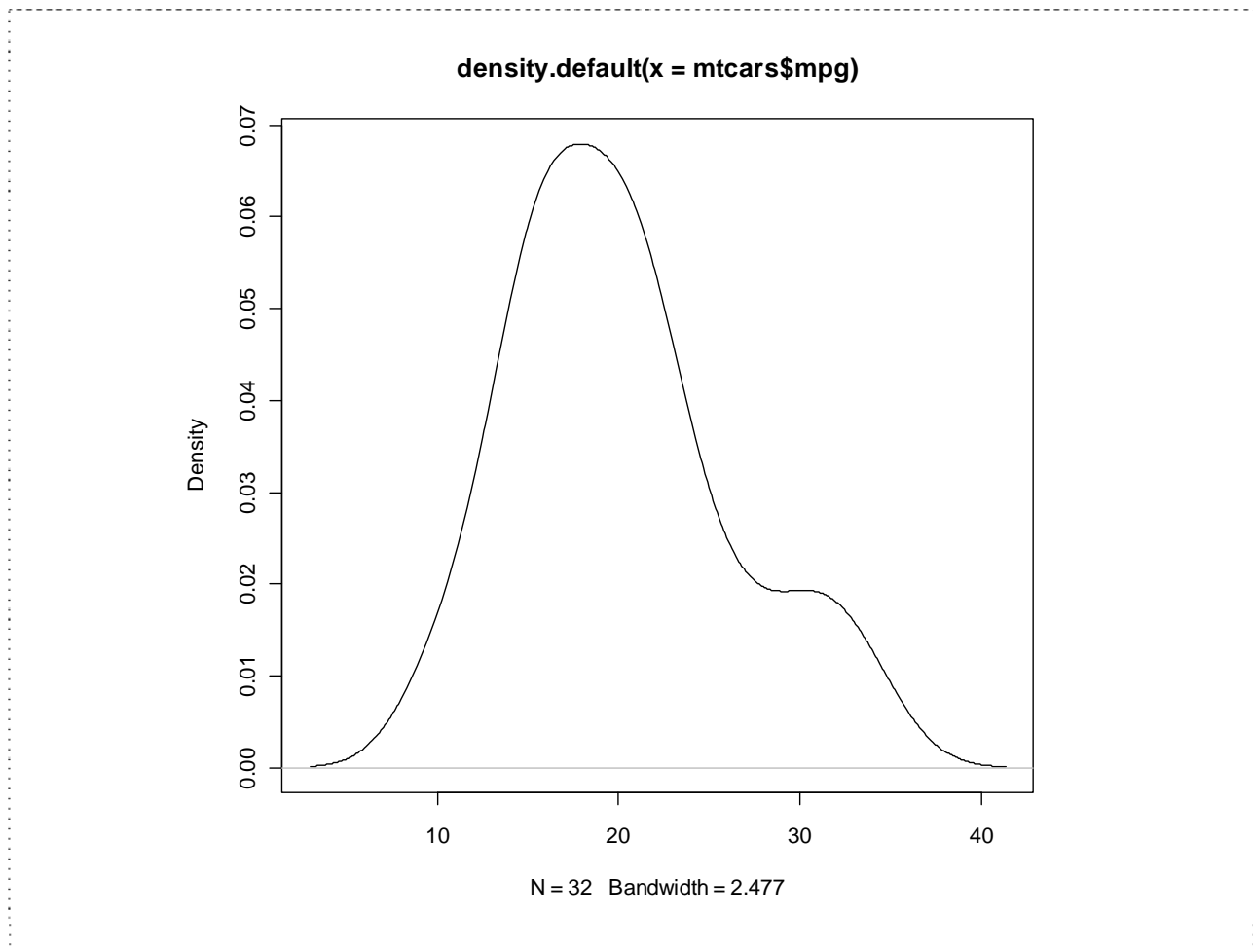
## Kernal Density Plots

Histograms can be a poor method for determining the shape of a distribution because it is so strongly affected by the number of bins used. Kernal density plots are usually a much more effective way to view the distribution of a variable. Create the plot using `plot(density(x))` where `x` is a numeric vector.

Kernal Density Plot:

```
# returns the density data
d <- density(mtcars$mpg)

# plots the results
plot(d)
```

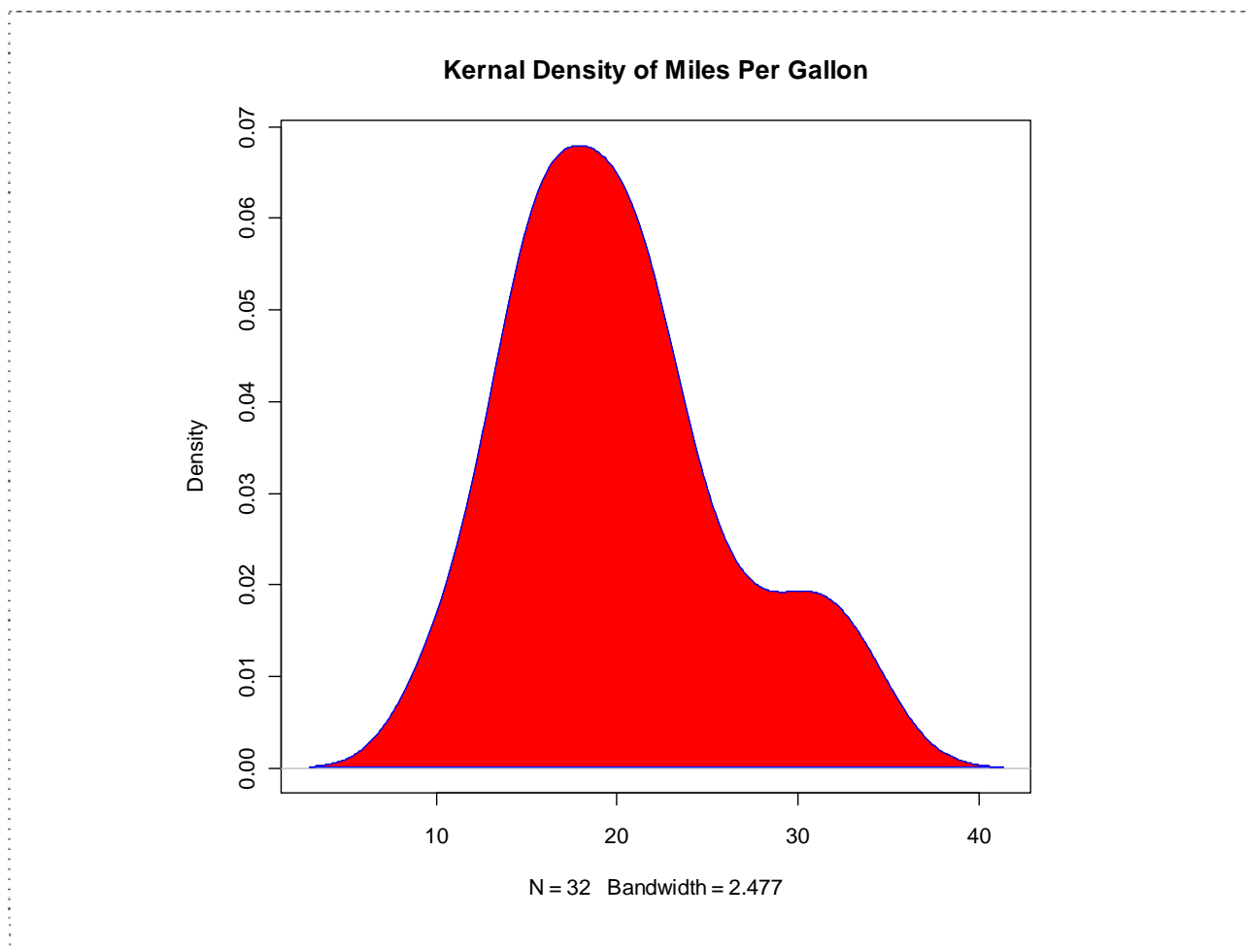


Filled Density Plot:

```
d <- density(mtcars$mpg)

show <- function()
{
  plot(d, main="Kernal Density of Miles Per Gallon")
  polygon(d, col="red", border="blue")
}

show()
```



### Comparing Groups VIA Kernal Density

---

The `sm.density.compare()` function in the `sm` package allows you to superimpose the kernel density plots of two or more groups. The format is `sm.density.compare(x, factor)` where `x` is a numeric vector and `factor` is the grouping variable.

Compare MPG distributions for cars with 4,6, or 8 cylinders:

```
attach(mtcars)

# create value labels
mtcars$cyl.f <- factor(
  cyl,
  levels= c(4,6,8),
  labels = c("4 cylinder", "6 cylinder", "8 cylinder")
)
# plot densities
show <- function()
{
  sm.density.compare(
    mpg,
```

```
mtcars$cyl.f,  
  xlab="Miles Per Gallon"  
)  
title(main="MPG Distribution by Car Cylinders")  
}
```

```
show()
```

