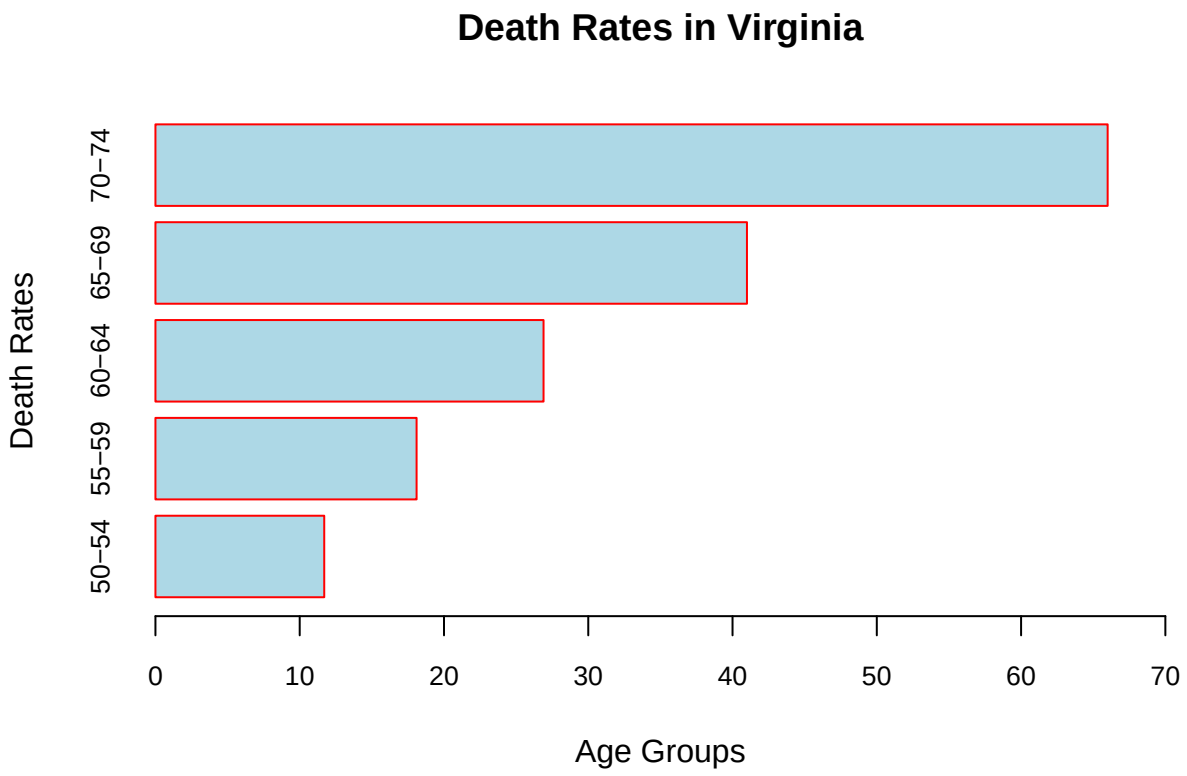
self-practice

```
gd <- as.data.frame(VADeaths)
gd
```

```
##       Rural Male Rural Female Urban Male Urban Female
## 50-54       11.7          8.7       15.4          8.4
## 55-59       18.1         11.7       24.3         13.6
## 60-64       26.9         20.3       37.0         19.3
## 65-69       41.0         30.9       54.6         35.1
## 70-74       66.0         54.3       71.1         50.0
```

```
# changing axis lenght and font-size
barplot(
    gd$`Rural Male`,
    horiz = TRUE,
    names.arg = c("50-54", "55-59", "60-64", "65-69", "70-74"),
    main = "Death Rates in Virginia",
    xlab = "Age Groups",
    ylab = "Death Rates",
    col = "lightblue",
    border = "red",
    beside = TRUE,
    xlim = c(0, 70),
    cex.axis = 0.8,
    cex.names = 0.8
)
```

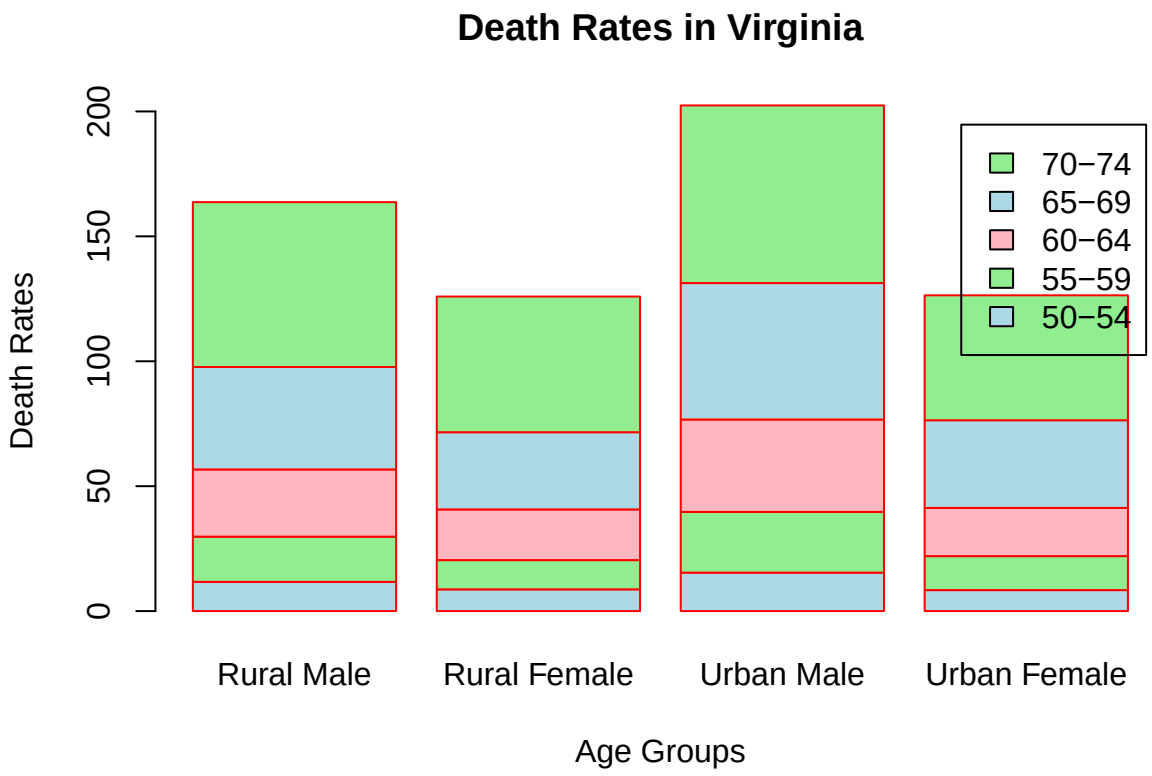


Sub divided bar plot

```
gdm <- as.matrix(gd)
gdm
```

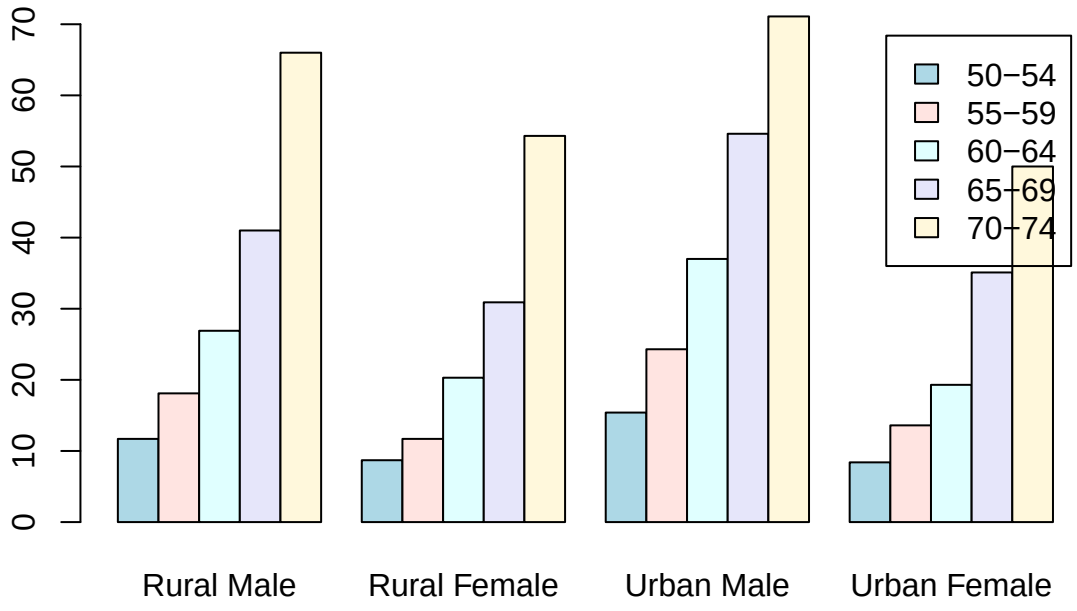
```
##       Rural Male Rural Female Urban Male Urban Female
## 50-54       11.7          8.7       15.4          8.4
## 55-59       18.1         11.7       24.3         13.6
## 60-64       26.9         20.3       37.0         19.3
## 65-69       41.0         30.9       54.6         35.1
## 70-74       66.0         54.3       71.1         50.0
```

```
barplot(
    gdm,
    main = "Death Rates in Virginia",
    xlab = "Age Groups",
    ylab = "Death Rates",
    col = c("lightblue", "lightgreen", "lightpink"),
    border = "red",
    legend = rownames(gd),
)
```



Multiple / Grouped bar diagram

```
barplot(
    gdm,
    col = c("lightblue", "mistyrose", "lightcyan", "lavender", "cornsilk"),
    legend  = rownames(gdm),
    beside = T
)
```

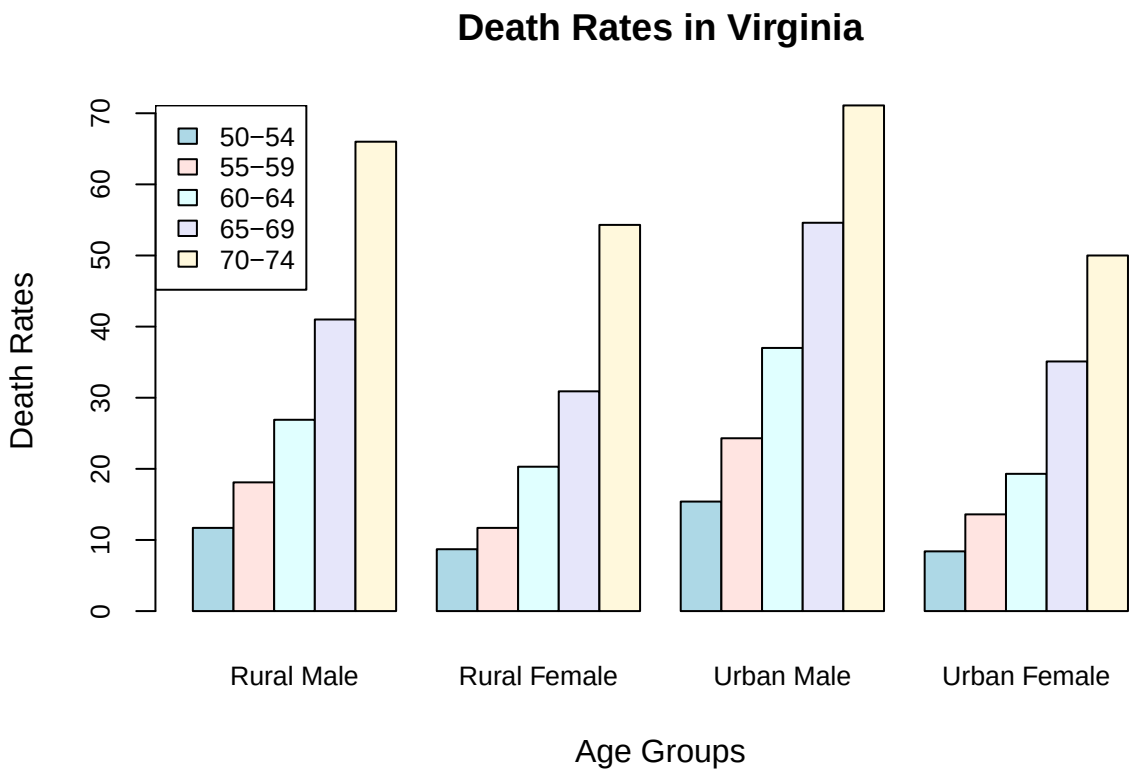



```
#  NOTE: Adding beside = TRUE will produce the multiple bar chart
```

Multiple / Group Bar diagram with change in legend values

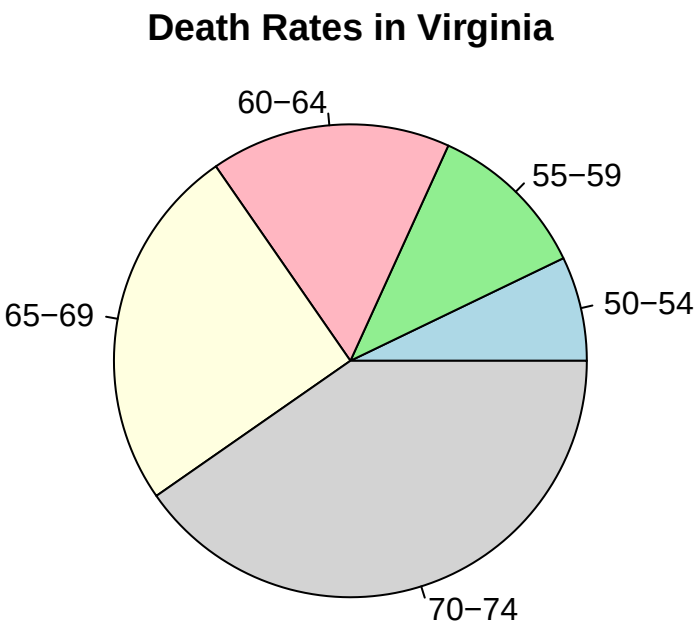
```
gdm <- as.matrix(VADeaths)

my_colors <- c("lightblue", "mistyrose", "lightcyan", "lavender", "cornsilk")
barplot(
    gdm,
    col = my_colors,
    beside = TRUE,
    main = "Death Rates in Virginia",
    xlab = "Age Groups",
    ylab = "Death Rates",
    cex.axis = 0.8,
    cex.names = 0.8
)
legend(
    "topleft",
    legend = rownames(gdm),
    fill = my_colors,
    cex = 0.8,
    lty = 0
)
```



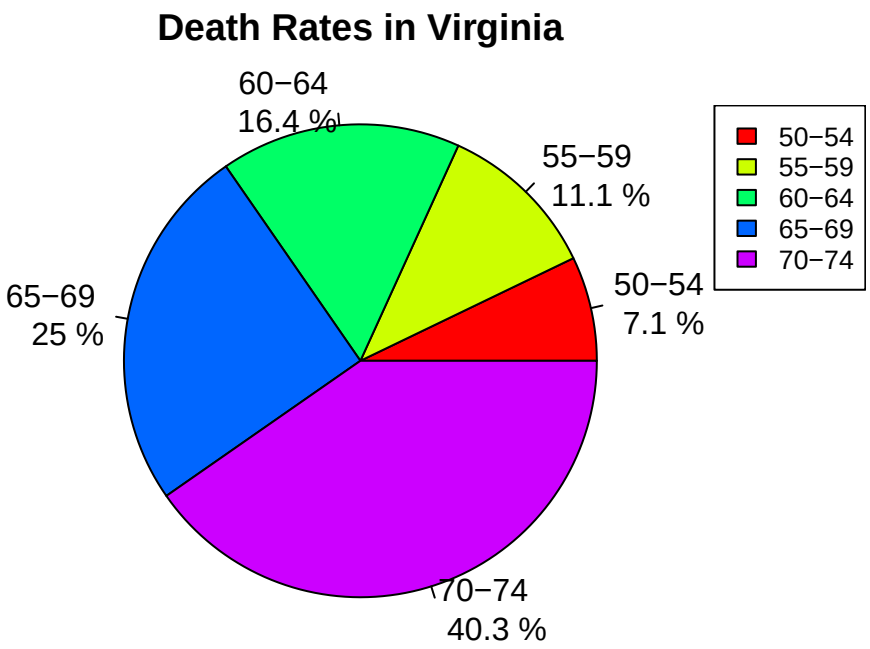
Pie chart

```
gd <- as.data.frame(VADeaths)
pie(
    gd$`Rural Male`,
    main = "Death Rates in Virginia",
    labels = rownames(gd),
    radius = 1,
    col = c("lightblue", "lightgreen", "lightpink", "lightyellow", "lightgray"),
)
```



Pie chart with percentage

```
gd$piepercent <- round(gd$`Rural Male` / sum(gd$`Rural Male`) * 100, 1)
pie(
    gd$`Rural Male`,
    main = "Death Rates in Virginia",
    labels = paste(rownames(gd), "\n", gd$piepercent, "%"),
    radius = 1,
    col = rainbow(length(gd$`Rural Male`)),
)
# How to place a legend on the topright corner?
# Pie chart with percentage and legend
legend(
    "topright",
    legend = rownames(gd),
    fill = rainbow(length(gd$`Rural Male`)),
    cex = 0.8,
    lty = 0
)
```

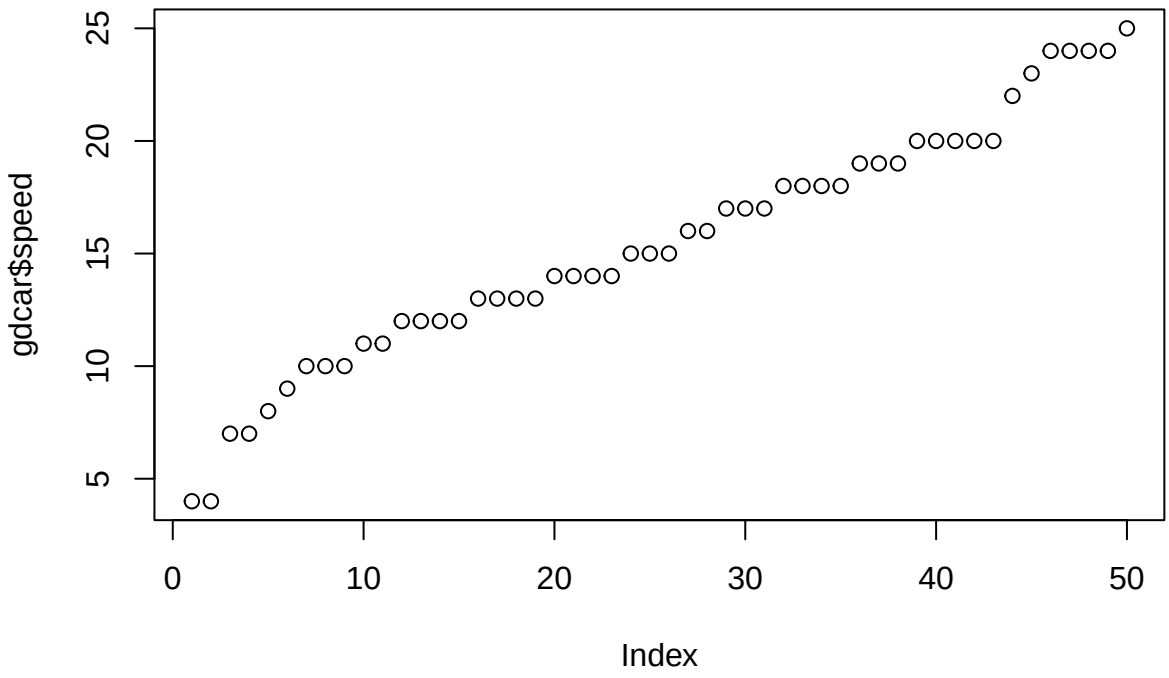


Historgram

```
gdcar = as.data.frame(cars)
str(gdcar)
```

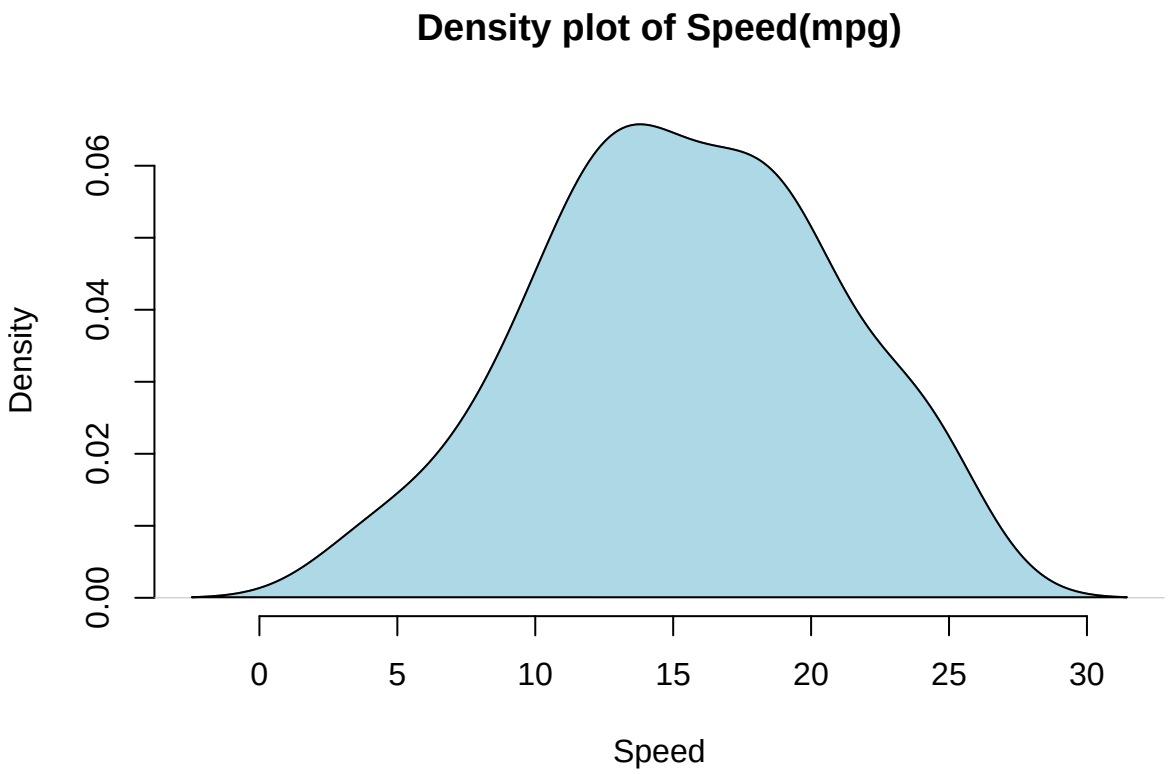
```
## 'data.frame':    50 obs. of  2 variables:
##  $ speed: num  4 4 7 7 8 9 10 10 10 11 ...
##  $ dist : num  2 10 4 22 16 10 18 26 34 17 ...
```

```
plot(gdcar$speed)
```



Density plot with polygon fill: Speed variable

```
dens <- density(cars$speed)
plot(
    dens,
    main = "Density plot of Speed(mpg)",
    xlab = "Speed",
    ylab = "Density",
    col = "steelblue",
    frame = FALSE
)
polygon(dens, col = "lightblue", border = "black")
```



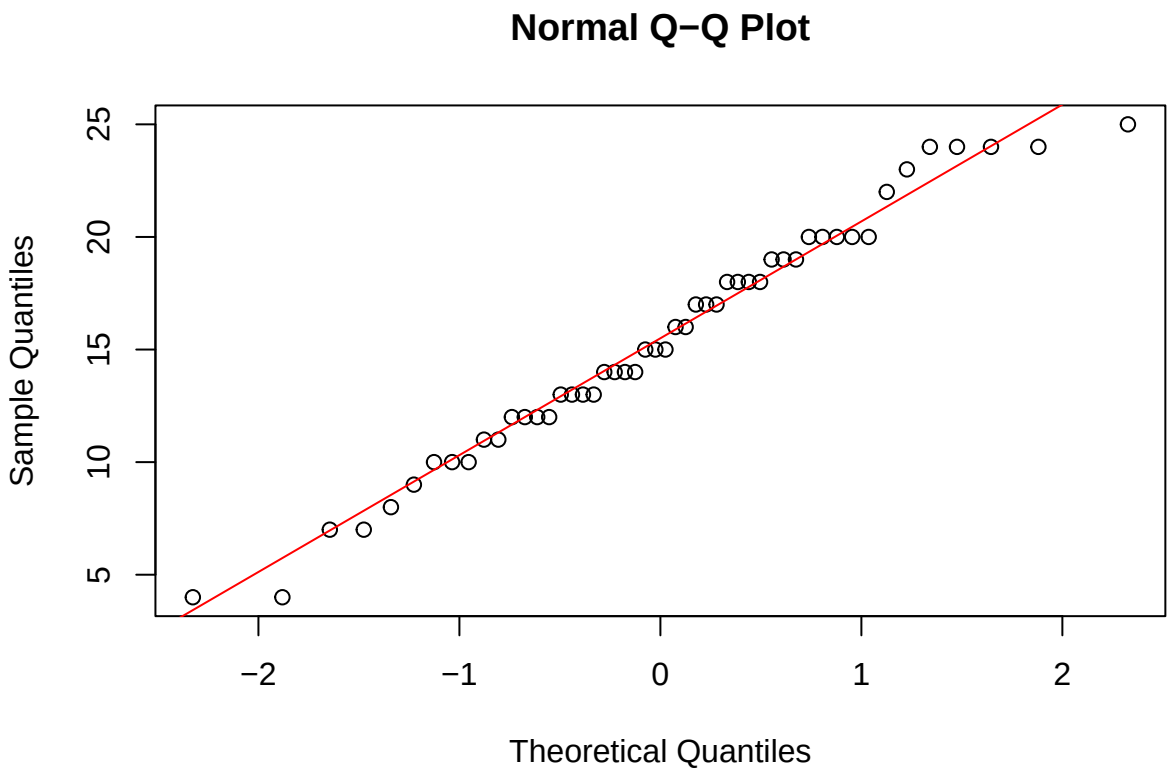
Q-Q plot

A Q-Q plot (quantile-quantile plot) is a graphical tool to help assess if a dataset follows a given distribution.

Always use this plot to access normality

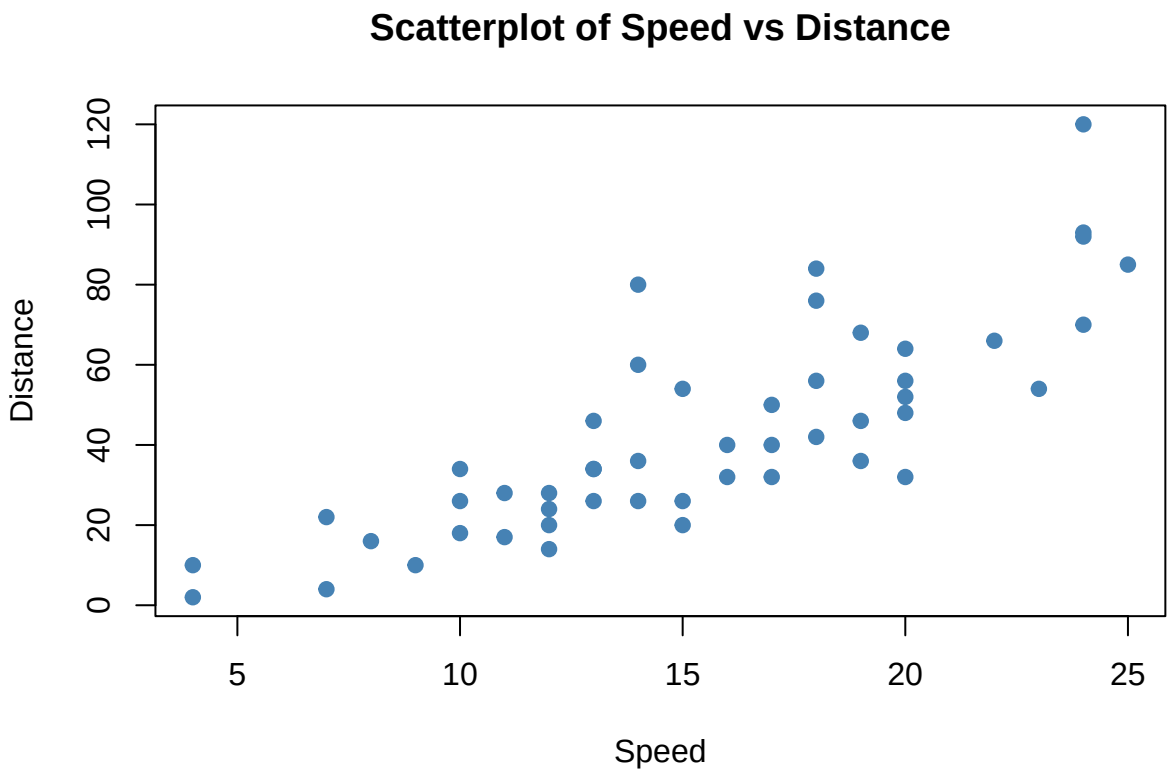
Q-Q plot of speed variable

```
qqnorm(gdcar$speed)
qqline(gdcar$speed, col = "red")
```



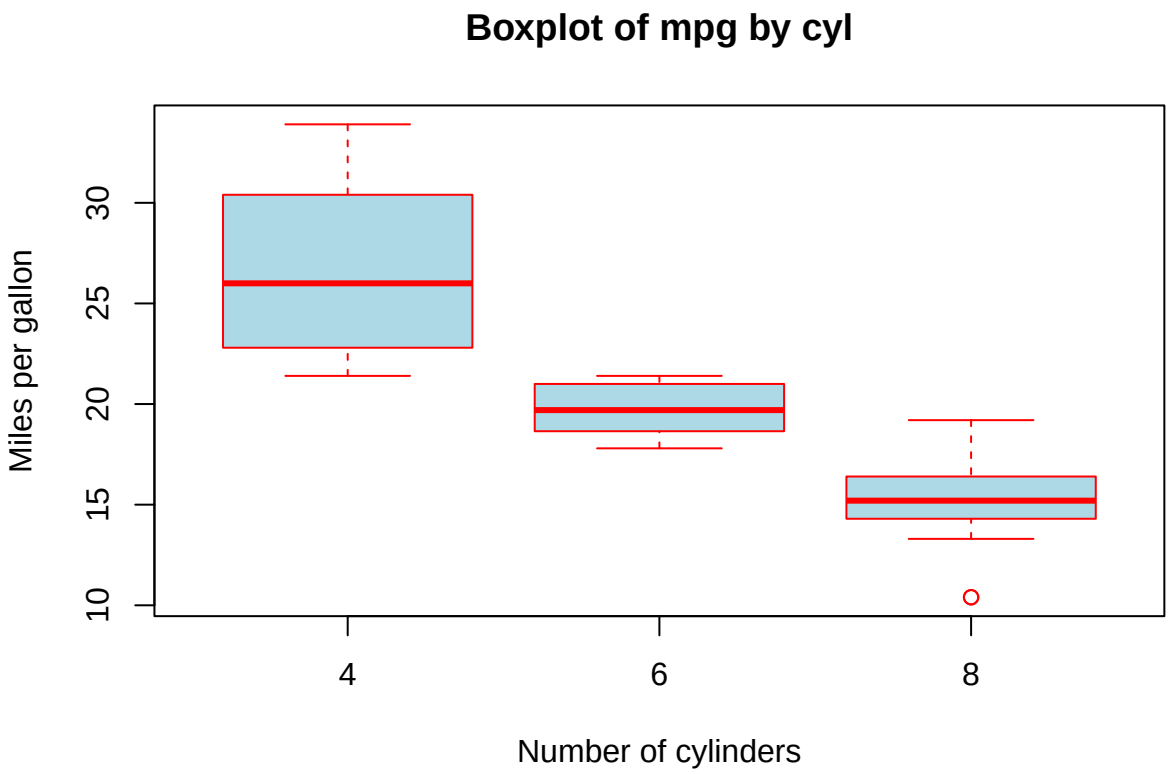
Scatter plot

```
# Scatterplot
plot(
    gdcar$speed,
    gdcar$dist,
    main = "Scatterplot of Speed vs Distance",
    xlab = "Speed",
    ylab = "Distance",
    col = "steelblue",
    pch = 19
)
```



Boxplot ## A boxplot is a standardized way of displaying the distribution of data based on a five-number summary (“minimum”, first quartile (Q1), median, third quartile (Q3), and “maximum”).

```
boxplot(
    mpg~cyl,
    data=mtcars,
    main = "Boxplot of mpg by cyl",
    xlab = "Number of cylinders",
    ylab = "Miles per gallon",
    col = "lightblue",
    border = "red",
    horizontal = FALSE,
    notch = FALSE
)
```


Boxplot of mpg by cyl
Miles per gallon
10
15
20
25
30
4
6
8
Number of cylinders