Investigate the R implements of mean, median, standard deviation, variance, correlation, and covariance.

Mean

It is calculated by taking the sum of the values and dividing with the number of values in a data series.

The function mean() is used to calculate this in R.

Syntax

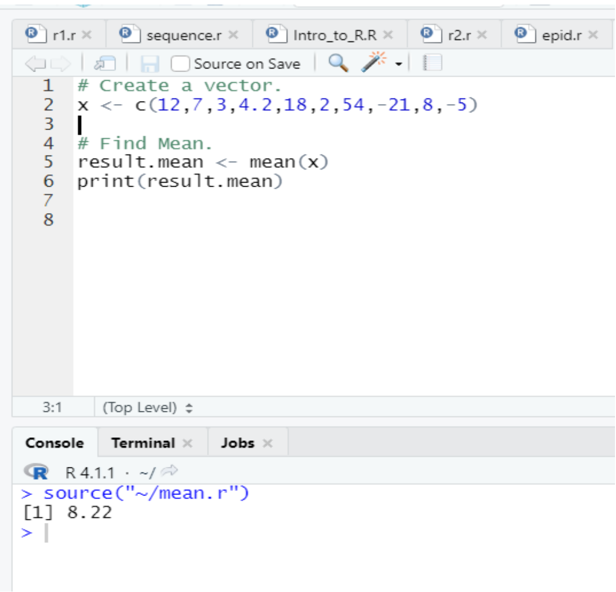
The basic syntax for calculating mean in R is −

mean(x, trim = 0, na.rm = FALSE, ...)

Following is the description of the parameters used −

- x is the input vector.
- trim is used to drop some observations from both end of the sorted vector.
- na.rm is used to remove the missing values from the input vector.

Example :



```
# Create a vector.
x <- c(12,7,3,4.2,18,2,54,-21,8,-5)

# Find Mean.
result.mean <- mean(x)
print(result.mean)
```

```
R 4.1.1 · ~/
> source("~/mean.r")
[1] 8.22
>
```

Median

The middle most value in a data series is called the median. The median() function is used in R to calculate this value.

Syntax

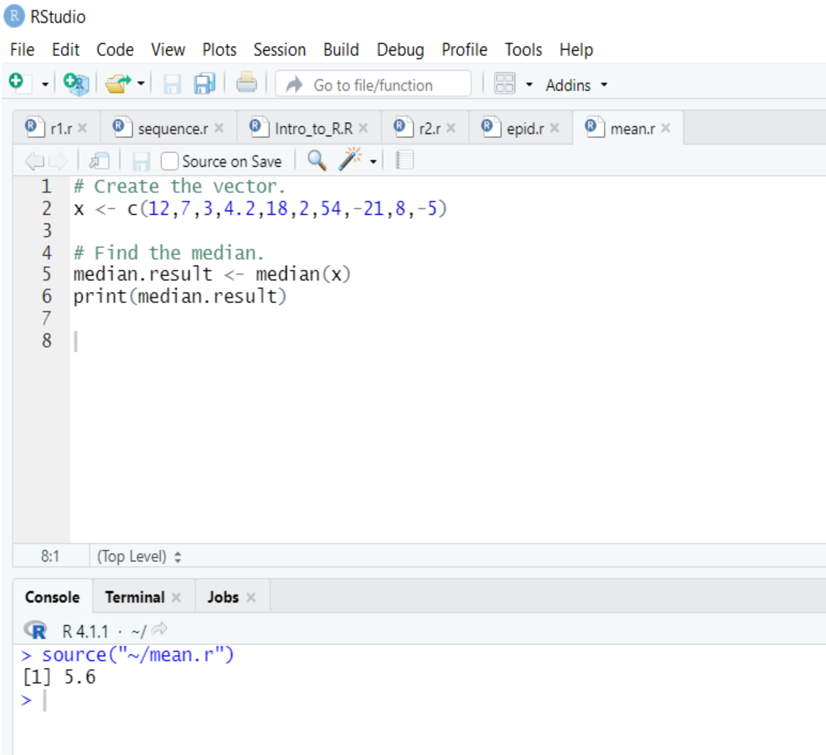
The basic syntax for calculating median in R is −

median(x, na.rm = FALSE)

Following is the description of the parameters used −

- x is the input vector.
- na.rm is used to remove the missing values from the input vector.

EXAMPLE



```
# Create the vector.
x <- c(12,7,3,4.2,18,2,54,-21,8,-5)

# Find the median.
median.result <- median(x)
print(median.result)

```

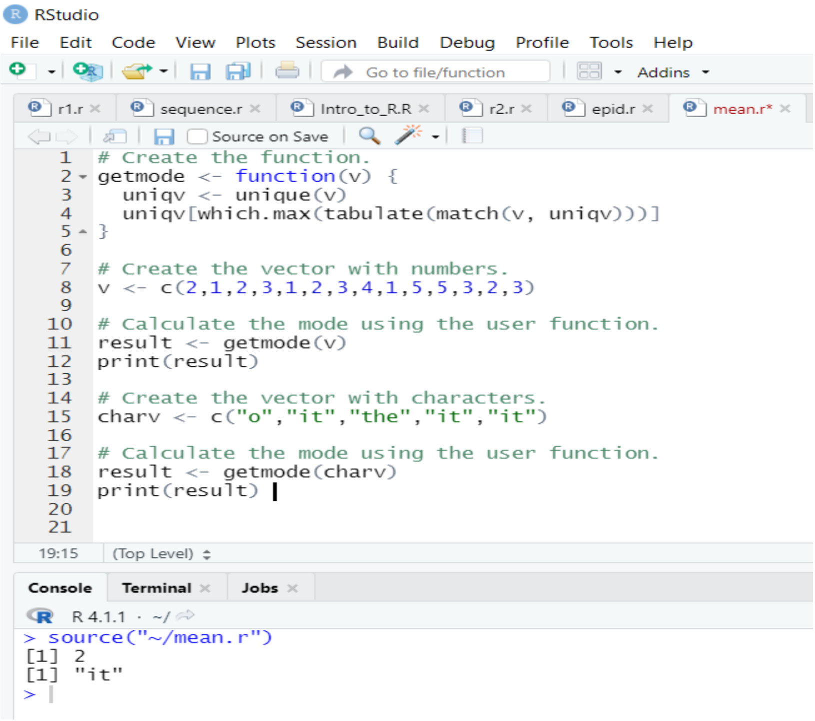
```
> source("~/mean.r")
[1] 5.6
>
```

Mode

The mode is the value that has highest number of occurrences in a set of data. Unike mean and median, mode can have both numeric and character data.

R does not have a standard in-built function to calculate mode. So we create a user function to calculate mode of a data set in R. This function takes the vector as input and gives the mode value as output.

EXAMPLE:



```
# Create the function.
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}

# Create the vector with numbers.
v <- c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)

# Calculate the mode using the user function.
result <- getmode(v)
print(result)

# Create the vector with characters.
charv <- c("o","it","the","it","it")

# Calculate the mode using the user function.
result <- getmode(charv)
print(result)
```

```
> source("~/mean.r")
[1] 2
[1] "it"
>
```

Standard Deviation

sd() Function

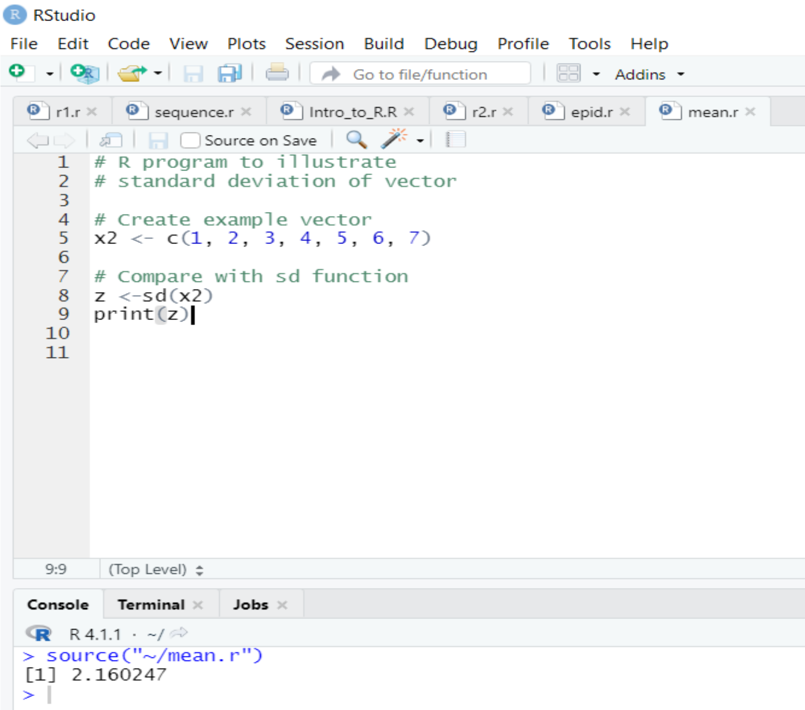
sd() function is used to compute the standard deviation of given values in R. It is the square root of its variance.

Syntax: sd(x)

Parameters:

x: numeric vector

EXAMPLE



RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
r1.r sequence.r Intro_to_R.R r2.r epid.r mean.r
Source on Save
9:9 (Top Level)
Console Terminal Jobs
R 4.1.1 · ~/

```
# R program to illustrate
# standard deviation of vector

# Create example vector
x2 <- c(1, 2, 3, 4, 5, 6, 7)

# Compare with sd function
z <-sd(x2)
print(z)
```

```
> source("~/mean.r")
[1] 2.160247
>
```

VARIANCE

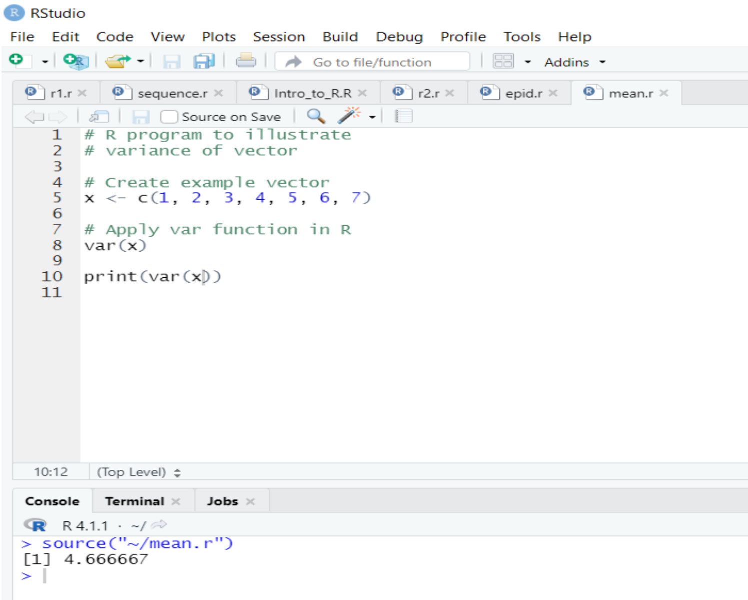
var() function in R Language computes the sample variance of a vector. It is the measure of how much value is away from the mean value.

Syntax: var(x)

Parameters:

x : numeric vector

EXAMPLE:



RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
r1.r sequence.r Intro_to_R.R r2.r epid.r mean.r
Source on Save
10:12 (Top Level)
Console Terminal Jobs
R 4.1.1 · ~/

```
# R program to illustrate
# variance of vector

# Create example vector
x <- c(1, 2, 3, 4, 5, 6, 7)

# Apply var function in R
var(x)

print(var(x))
```

```
> source("~/mean.r")
[1] 4.666667
>
```

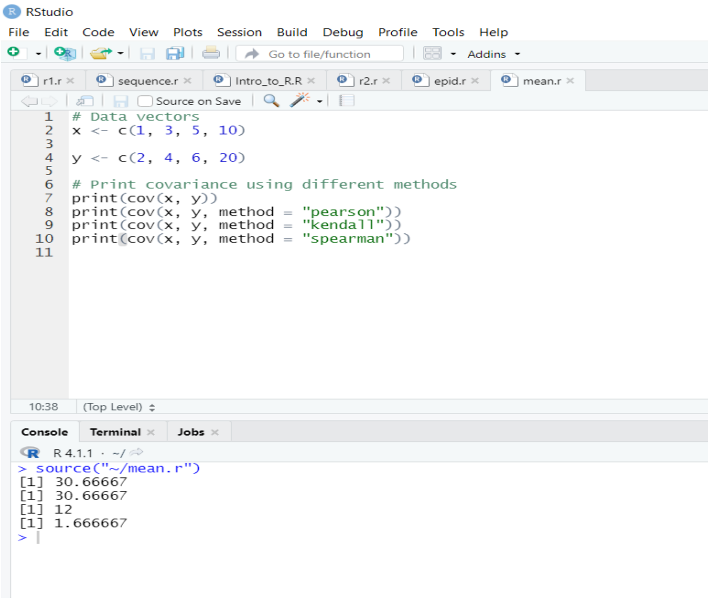
Covariance

In R programming, covariance can be measured using cov() function. Covariance is a statistical term used to measures the direction of the linear relationship between the data vectors.

Syntax: cov(x, y, method)

where,

- x and y represents the data vectors
- method defines the type of method to be used to compute covariance. Default is "pearson".



```
# Data vectors
x <- c(1, 3, 5, 10)

y <- c(2, 4, 6, 20)

# Print covariance using different methods
print(cov(x, y))
print(cov(x, y, method = "pearson"))
print(cov(x, y, method = "kendall"))
print(cov(x, y, method = "spearman"))
```

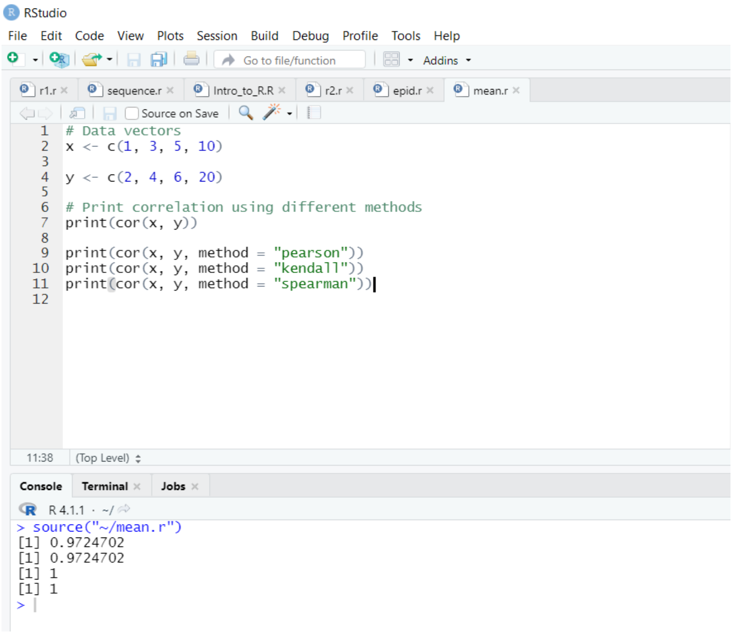
```
> source("~/mean.r")
[1] 30.66667
[1] 30.66667
[1] 12
[1] 1.666667
>
```

Correlation

Syntax: cor(x, y, method)

where,

- x and y represents the data vectors
- method defines the type of method to be used to compute covariance. Default is "pearson".



```
# Data vectors
x <- c(1, 3, 5, 10)

y <- c(2, 4, 6, 20)

# Print correlation using different methods
print(cor(x, y))

print(cor(x, y, method = "pearson"))
print(cor(x, y, method = "kendall"))
print(cor(x, y, method = "spearman"))

```

```
> source("~/mean.r")
[1] 0.9724702
[1] 0.9724702
[1] 1
[1] 1
> 
```