Measure of central tendency

Overview

This is the vector of numbers for which we are going to calculate the **measure of central tendency**, which is a set of measures classed as summary statistics.

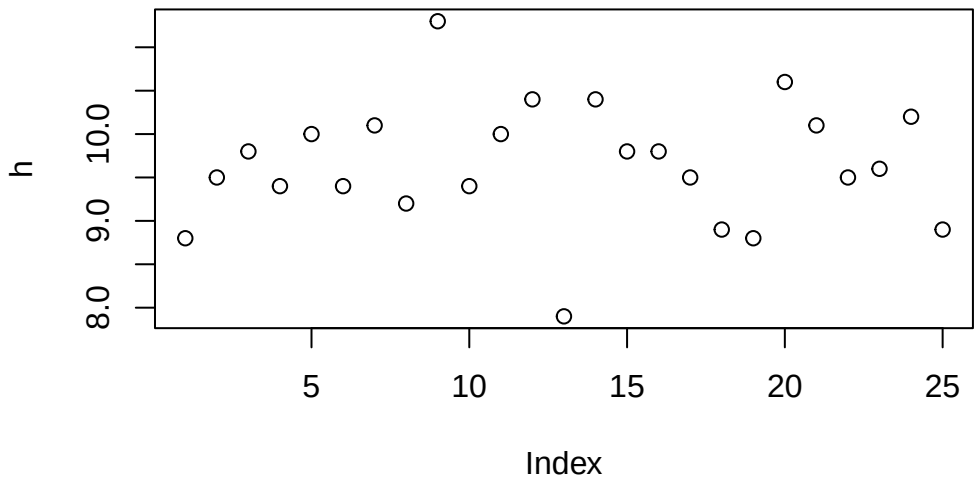
The vector of numbers:

```
h <- c(8.8,9.5,9.8,9.4,10,
       9.4,10.1,9.2,11.3,9.4,10,10.4,7.9,
       10.4,9.8,9.8,9.5,8.9,8.8,10.6,10.1,
       9.5,9.6,10.2,8.9)
```

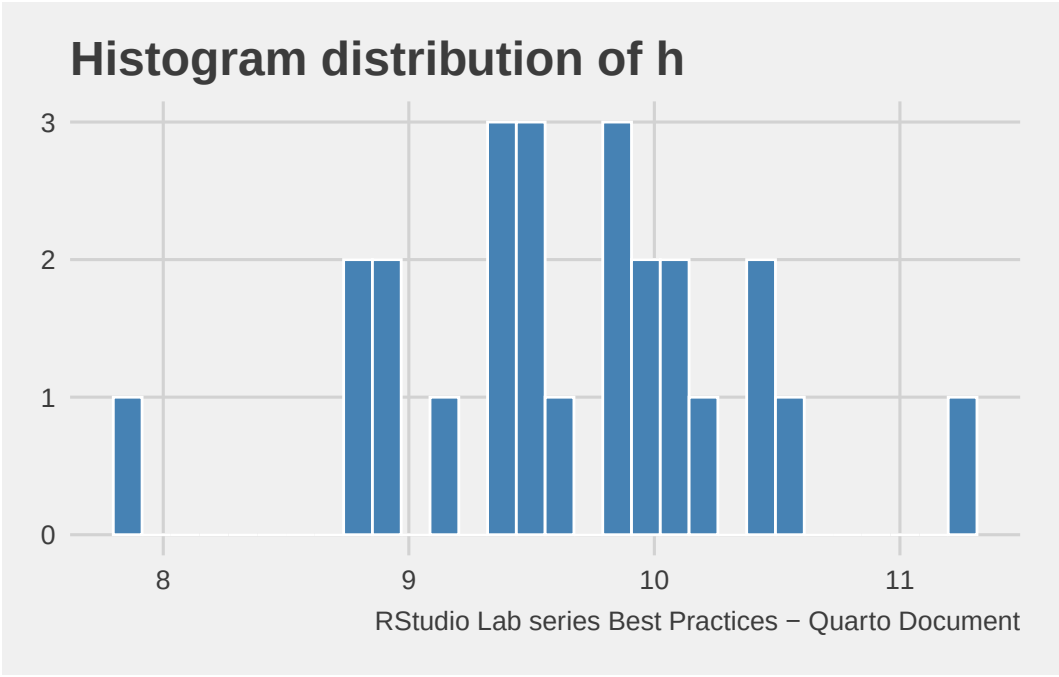
It has a length of:

```
[1] 25
```

Have a quick look at the distribution of `h` with a scatterplot using the `plot()` function:



Load the necessary library and make it a dataframe to produce a `histogram` with {ggplot2} package:



The `measure of central tendency` is given by the **mean**:

$$\mu = \frac{\sum_{i=1}^{n=25} x_i}{n}$$

In this case $n = 25$ and the mean is:

```
[1] 9.652
```

The **variance** is:

$$\sigma^2 = \frac{\sum_{i=1}^{n=25} (x_i - \hat{x})^2}{n}$$

```
[1] 0.4942667
```

The **standard deviation** is the square root of the variance:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n=25} (x_i - \hat{x})^2}{n}}$$

```
[1] 0.703041
```

The **coefficient of variation**:

$$c_v = \frac{\sigma}{\mu}$$

```
[1] 0.07283889
```

The **95% confidence interval** for the mean:

```
[1] 0.1406082
```

$$\alpha = 0.05$$

The degree of freedom:

```
[1] 24
```

```
[1] 2.063899
```

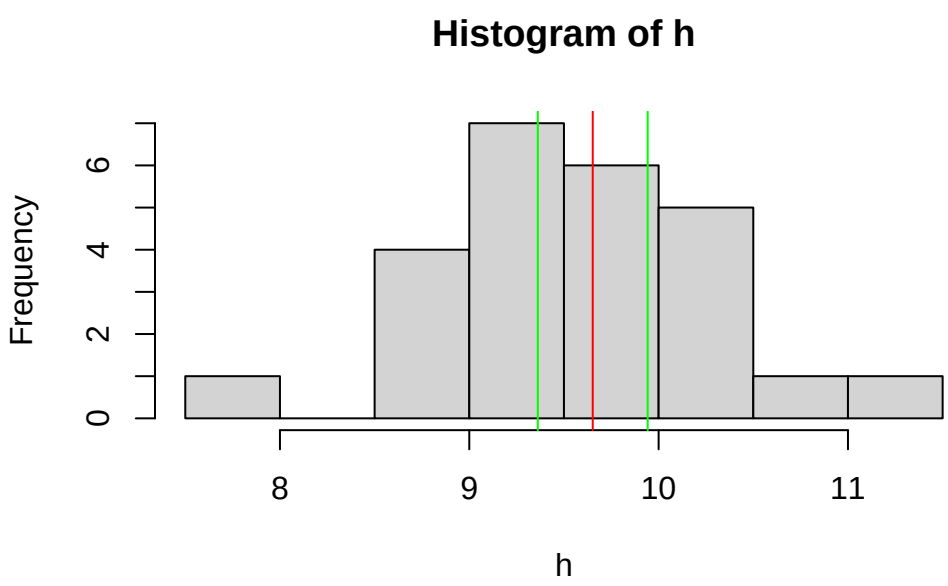
```
[1] 9.361799
[1] 9.942201
```

```
Call:
lm(formula = h ~ 1, data = hh)

Residuals:
   Min     1Q Median     3Q    Max
-1.752 -0.252 -0.052  0.448  1.648

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   9.6520     0.1406   68.64   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.703 on 24 degrees of freedom
```

Verify the values of the confidence intervals with `confint()` function:

```
               2.5 %   97.5 %
(Intercept) 9.361799 9.942201
```

The **histogram** releases the frequencies of the provided vector.


Histogram of h
Frequency
0
2
4
6
8
9
10
11
h