Experiment 13:

Illustrate the Correlation and Covariance analysis using R

Aim: To analyze data using Correlation and Covariance measures in R

Exp. 13: Correlation and Covariance using R

hen dealing with more than one variable, we need to test t
ationships with each other

vo simple measures

- Correlation
- Covariance

Exp. 13: Correlation using R

orrelation is used to evaluate the association between tw

ore variables

is computed as:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

If $r_{xy} = 0$ No Correlation

$r_{xy} > 0$ Positive Correlation

$r_{xy} < 0$ Negative Correlation

x and y are two vectors of length n

$\bar{x}$ and $\bar{y}$ corresponds to the means of x and y, respectively.

Exp. 13: Correlation using R

ax:

r(x, y , method)

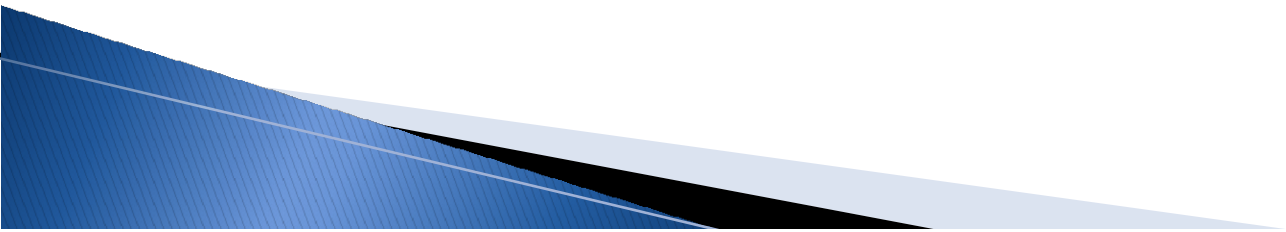
- **x** is a numeric vector, matrix or data frame
- **y is** NULL (default) or a vector, matrix or data frame
  compatible dimensions to x
- **method** ="pearson" or "kendall" or "spearm
  #Default:pearson

Correlation range is [-1,+1]

Exp. 13: Correlation using R

```
or=cor(airquality$Wind,airquality$Temp)
nt(pcor)
0.4579879                                    #Negative Correlation

w=cor(women$height,women$weight)
nt(corw)
.9954948                                     # Positive Correlation
```



Exp. 13: Correlation using R

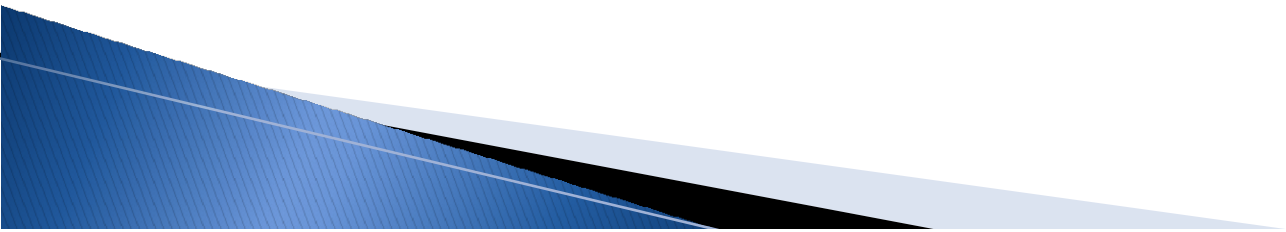
```
or=cor(iris[1:4])
nt(icor)                                    #Correlation Matrix
             Sepal.Length Sepal.Width Petal.Length Petal.Width
Length    1.0000000   -0.1175698    0.8717538   0.8179411
Width    -0.1175698    1.0000000   -0.4284401  -0.3661259
Length    0.8717538   -0.4284401    1.0000000   0.9628654
Width     0.8179411   -0.3661259    0.9628654   1.0000000
```

Exp. 13: Covariance using R

ovariance of two variables x and y in a data set measures how
o are linearly related.

positive covariance would indicate a positive linear relation
tween the variables, and a negative covariance would indicate
posite.

ovariance range is $[-\infty,+\infty]$



Exp. 13: Covariance using R

ovariance of two variables x and y in a data set is computed as

Population Covariance Formula

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{N}$$

Sample Covariance

$$Cov(x,y) = \frac{\Sigma(x_i-\bar{x})(y_i-y)}{N-1}$$

Exp. 13: Covariance using R

tax:

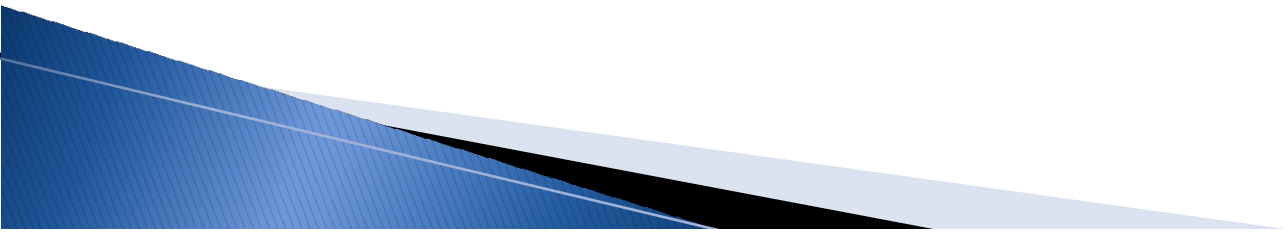
v(x, y , method)

- **x** is a numeric vector, matrix or data frame
- **y is** NULL (default) or a vector, matrix or data frame
  compatible dimensions to x
- **method** ="pearson" or "kendall" or "spearman"
  #Default:pearson

Exp. 13: Covariance using R

```
ov=cov(airquality$Wind,airquality$Temp)
nt(acov)
15.27214

cov=cov(women$height,women$weight)
nt(wcov)
9
```



Exp. 13: Covariance using R

```
ov=cov(iris[1:4])
nt(icov)                                    #Covariance Matrix
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Length	0.6856935	-0.0424340	1.2743154	0.5162707
Width	-0.0424340	0.1899794	-0.3296564	-0.1216394
Length	1.2743154	-0.3296564	3.1162779	1.2956094
Width	0.5162707	-0.1216394	1.2956094	0.5810063

onal values are variance of each attribute