

# Correlation and Regression

## 1.) Aim:

To find the rank correlation coefficient of the given data using Karl Pearson's method.

## Data:

selection=c(44,49,52,54,47,76,65,60,63,58,50,67)

proficiency=c(48,55,45,60,43,58,50,77,46,  
47,65)

## Syntax:

cor.test(selection,proficiency)

## Output:

Pearson's product-moment correlation

data: selection and proficiency

$t = 3.9475$ ,  $df = 10$ , p-value 0.002741

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.3741265 0.9353910

sample estimates:

cor

0.7804552

## Conclusion:

The rank using Karl Pearson's method is found to be 0.7804552.

## 2).Aim:

To find the coefficient of correlation using Spearman's method.

## Data:

statistics=c(55,40,70,60,62,73,65,65,20,35,46,50)

computer=c(35,32,65,50,63,45,50,65,70,72,72,40)

## Syntax:

```
cor.test(statistics,computer,method= "spearman")
```

## Output:

Spearman's rank correlation rho

data: statistics and computer

S = 344.41, p-value 0.5243

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.2042266

## Conclusion:

The rank using Spearman correlation method is found to be -0.2042266.

## 3). Aim:

To estimate the score in Zoology using regression.

## Data:

```
> x=c(34,37,36,32,32,36,35,34,29,35)
```

```
> y=c(37,37,34,34,33,40,39,37,36,35)
```

## Syntax:

```
> fit=lm(x~y)
```

## Output:

Call:

`lm(formula = x ~ y)`

Coefficients:

(Intercept)	y
18.9167	0.4167

## Conclusion:

The equation of the line of regression is

$$X = 18.9167 + 0.4167Y.$$

The required score of the student in Zoology is 30.58333.

## 4.) Aim:

To find the products for the given table in the form of multiple regression.

## Data:

```
> Y=c(110,80,70,120,150,90,70,120)
```

```
> X1=c(30,40,20,50,60,40,20,60)
```

```
> X2=c(11,10,7,15,19,12,8,14)
```

## Syntax:

```
> input_data=data.frame(Y,X1,X2)
```

```
> fit=lm(Y~X1+X2)
```

## Output:

	Y	X1	X2
1	110	30	11
2	80	40	10
3	70	20	7
4	120	50	15
5	150	60	19
6	90	40	12
7	70	20	8
8	120	60	14

Call:

```
lm(formula = Y ~ X1 + X2, data = input_data)
```

Coefficients:

(Intercept)	X1	X2
16.8314	-0.2442	7.8488

## Conclusion:

The equation of the multiple regression model is  
 $16.8314 - 0.2442X_1 + 7.8488X_2$ .