MAT1011 Applied Statistics Lab (26-04-2022)

Examples (Given in lab class) :-

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
> # Example-3
> x=c(6,9,11,13,22,26,28,33,35)
> y=c(68,67,65,53,44,40,37,34,32)
> cor(x,y)
[1] -0.977326
>
> # Example-4
> x=c(1,2,3,4,5,6,7,8,9,10)
> y=c(10,12,16,28,25,36,41,49,40,50)
> cor(x,y)
[1] 0.9581574
>
> # Example-1
> # The average prices of stocks and bonds listed on the New York Stock
> # Exchange during the years 1950 through 1959 ar given as :
> # Year  1950  1951  1952  1953  1954  1955  1956  1957  1958  1959
> # Price
> # (stock)35.22 39.87 41.85 43.23 40.06 53.29 54.14 49.12 40.17 55.15
>
> # Price
> # (bond) 102.43 100.93 97.43 97.81 98.32 100.07 97.08 91.59 94.85 94.65
>
> X=c(35.22,39.87,41.85,43.23,40.06,53.29,54.14,49.12,40.17,55.15)
> Y=c(102.43,100.93,97.43,97.81,98.32,100.07,97.08,91.59,94.85,94.65)
> cor(X,Y)
[1] -0.4517697
>
> # Example-2
> # Calculate rank correlation coefficient of data :
> # x  80  78  75  75  68  57  60  59
> # y  110 111 114 114 114 116 115 117
> cor.test(x,y,method="spearman")

        Spearman's rank correlation rho

data:  x and y
S = 8, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.9515152
```

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
> # Calculate rank correlation coefficient to the following data :
> # X  48  33  40  9  16  16  65  24  16  57
> # Y  13  13  24  6  15  4   20  9   6   19
> X=c(48,33,40,9,16,16,65,24,16,57)
> Y=c(13,13,24,6,15,4,20,9,6,19)
> cor.test(X,Y,method="spearman")

        Spearman's rank correlation rho

data:  X and Y
S = 41.757, p-value = 0.01305
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.7469278
```

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
> # An accurate assessment of soil productivity is critical to national land-use planning
> # Unfortunately, as th author of article "Productivity Ratings Based on Soil Series)
> # argues, an acceptable soil productivity index is not so easy to come by.
> # One difficulty is that productivity is determined partly by which crop is planned,
> # and relationship between the yield of two different crops planted in same soil may
> # not be very strong. To illustrate, the article presents the accompanying data on corn
> # yield x and peanut yield y (mT/Ha) for eight different types of soil
>
> # x  2.4  3.4  4.6  3.7  2.2  3.3  4.0  2.1
> # y  1.33 2.12 1.80 1.65 2.00 1.76 2.11 1.63
> x=c(2.4,3.4,4.6,3.7,2.2,3.3,4.0,2.1)
> y=c(1.33,2.12,1.80,1.65,2.00,1.76,2.11,1.63)
> cor.test(x,y,method="spearman")

        Spearman's rank correlation rho

data:  x and y
S = 48, p-value = 0.2992
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.4285714
```

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
>
> # Following data pertain to the resistance in ohms
> # and failure times in minutes.
> x=c(43,29,44,33,33,47,34,31,48,34,46,37,36,39,36,47,28,40,42,33,46,28,48,45)
> y=c(32,20,45,35,22,46,28,26,37,33,47,30,36,33,21,44,26,45,39,25,36,25,45,36)
> cor(x,y)
[1] 0.8085055
> fit=lm(y~x)
> fit

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
     -5.518        1.019
```

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
> # A study was made by a retail merchant to determine
> # the relation between weekly advertising expenditures
> # Advertising Costs ($)        Sales($)
> #       40                        385
> #       20                        400
> #       25                        395
> #       20                        365
> #       30                        475
> #       30                        440
> #       40                        490
> #       20                        420
> #       50                        560
> #       40                        525
> #       25                        480
> #       50                        510
> x=c(40,20,25,20,30,30,40,20,50,40,25,50)
> y=c(385,400,395,365,475,440,490,420,560,525,480,510)
> cor(x,y)
[1] 0.7297485
> fit=lm(y~x)
> fit

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
    322.273        4.045
```

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
>
> # The following data are a result of an investigation
> # as to the effect of reaction temperature x on percent
> # conversion of a chemical process y. Fit a simple
> # regression, and use a lack-of-fit test to determine if
> # a model is adequate.
> # Observation    Temperature(x)   Conversion(y)\
> #      1              200              43
> #      2              250              78
> #      3              200              69
> #      4              250              73
> #      5              189.65           48
> #      6              260.35           78
> #      7              225              65
> #      8              225              74
> #      9              225              76
> #      10             225              79
> #      11             225              83
> #      12             225              81
>
> x=c(200,250,200,250,189.65,260.35,225,225,225,225,225,225)
> y=c(43,78,69,73,48,78,65,74,76,79,83,81)
> cor(x,y)
[1] 0.6795114
> fit=lm(x~y)
> fit

Call:
lm(formula = x ~ y)

Coefficients:
(Intercept)            y
    144.956        1.134
```

Questions on Spearman Correlation :–

Q.1] In a certain type of metal test specimen, the normal stress on a specimen is known to be functionally related to the shear resistance. The following is a set of coded experimental data on the two variables:

Normal Stress, x	Shear Resistance, y
26.8	26.5
25.4	27.3
28.9	24.2
23.6	27.1
27.7	23.6
23.9	25.9
24.7	26.3
28.1	22.5
26.9	21.7
27.4	21.4
22.6	25.8
25.6	24.9

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
> x=c(26.8,25.4,28.9,23.6,27.7,23.9,24.7,28.1,26.9,27.4,22.6,25.6)
> y=c(26.5,27.3,24.2,27.1,23.6,25.9,26.3,22.5,21.7,21.4,25.8,24.9)
> cor.test(x,y,method="spearman")

        Spearman's rank correlation rho

data:  x and y
S = 474, p-value = 0.02398
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.6573427
```

Q.2] The following is a portion of a classic data set called the “pilot plot data” in Fitting Equations to Data by Daniel and Wood, published in 1971. The response y is the acid content of material produced by titration, whereas the regressor x is the organic acid content produced by extraction and weighing.

y	x	y	x
76	123	70	109
62	55	37	48
66	100	82	138
58	75	88	164
88	159	43	28

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
> x=c(123,55,100,75,159,109,48,138,164,28)
> y=c(76,62,66,58,88,70,37,82,88,43)
> cor.test(x,y,method="spearman")

        Spearman's rank correlation rho

data:  x and y
S = 4.5129, p-value = 2.369e-06
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.9726489
```

Q.3] A study of the amount of rainfall and the quantity of air pollution removed produced the following data :

Daily Rainfall, x (0.01 cm)	Particulate Removed y ($\mu g/m^3$)
4.3	126
4.5	121
5.9	116
5.6	118
6.1	114
5.2	118
3.8	132
2.1	141
7.5	108

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
>
> x=c(4.3,4.5,5.9,5.6,6.1,5.2,3.8,2.1,7.5)
> y=c(126,121,116,118,114,118,132,141,108)
> cor.test(x,y,method="spearman")

        Spearman's rank correlation rho

data:  x and y
S = 239.5, p-value = 1.542e-08
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
-0.9958246
```

Question of Simple Correlation :–

Q.1] A study was made on the amount of converted sugar in a certain process at various temperatures. The data were coded and recorded as follows:

Temperature, x	Converted Sugar, y
1.0	8.1
1.1	7.8
1.2	8.5
1.3	9.8
1.4	9.5
1.5	8.9
1.6	8.6
1.7	10.2
1.8	9.3
1.9	9.2
2.0	10.5

```
> # KHAN MOHD OWAIS RAZA (20BCD7138)
> # MAT1011 Lab (26-04-2022)
>
> x=c(1.0,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2.0)
> y=c(8.1,7.8,8.5,9.8,9.5,8.9,8.6,10.2,9.3,9.2,10.5)
> cor(x,y)
[1] 0.7070264
```

Q.2] The grades of a class of 9 students on a midterm report (x) and on the final examination (y) are as follows:

x	77	50	71	72	81	94	96	99	67
y	82	66	78	34	47	85	99	99	68

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
> x=c(77,50,71,72,81,94,96,99,67)
> y=c(82,66,78,34,47,85,99,99,68)
> cor(x,y)
[1] 0.5610055
```

Q.3] A mathematics placement test is given to all entering freshmen at a small college. A student who receives a grade below 35 is denied admission to the regular mathematics course and placed in a remedial class. The placement test scores and the final grades for 20 students who took the regular course were recorded.

Placement Test	Course Grade
50	53
35	41
35	61
40	56
55	68
65	36
35	11
60	70
90	79
35	59
90	54
80	91
60	48
60	71
60	71
40	47
55	53
50	68
65	57
50	79

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
>
> # Placement Test = x
> # Course Grade = y
> x=c(50,35,35,40,55,65,35,60,90,35,90,80,60,60,60,40,55,50,65,50)
> y=c(53,41,61,56,68,36,11,70,79,59,54,91,48,71,71,47,53,68,57,79)
> cor(x,y)
[1] 0.4535134
```

Questions on regression :–

Q.1] A study was conducted at Virginia Tech to determine if certain static arm-strength measures have an influence on the "dynamic lift" characteristics of an individual. Twenty-five individuals were subjected to strength tests and then were asked to perform a weight-lifting test in which weight was dynamically lifted overhead. The data are given here.

Individual	Arm Strength, x	Dynamic Lift, y
1	17.3	71.7
2	19.3	48.3
3	19.5	88.3
4	19.7	75.0
5	22.9	91.7
6	23.1	100.0
7	26.4	73.3
8	26.8	65.0
9	27.6	75.0
10	28.1	88.3

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
>
> # Arm Strength = x
> # Dynamic Lift = y
> x=c(17.3,19.3,19.5,19.7,22.9,23.1,26.4,26.8,27.6,28.1)
> y=c(71.7,48.3,88.3,75.0,91.7,100.0,73.3,65.0,75.0,88.3)
> fit=lm(x~y)
> fit

Call:
lm(formula = x ~ y)

Coefficients:
(Intercept)            y
   19.44198      0.04672
```

Q.2] The following data were collected to determine the relationship between pressure and the corresponding scale reading for the purpose of calibration

Pressure, x (lb/sq in.)	Scale Reading, y
10	13
10	18
10	16
10	15
10	20
50	86
50	90
50	88
50	88
50	92

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
>
> # Pressure(lb/sq.in) = x
> # Scale Reading = y
> x=c(10,10,10,10,10,50,50,50,50,50)
> y=c(13,18,16,15,20,86,90,88,88,92)
> fit=lm(x~y)
> fit

Call:
lm(formula = x ~ y)

Coefficients:
(Intercept)            y
     1.0497       0.5504
```

Q.3] The thrust of an engine (y) is a function of exhaust temperature (x) in degree F when other important variables are held constant. Consider the following data.

y	x	y	x
4300	1760	4010	1665
4650	1652	3810	1550
3200	1485	4500	1700
3150	1390	3008	1270
4950	1820		

```
> # KHAN MOHD OWAIS RAZA 20BCD7138
> # MAT1011 Lab (26-04-2022)
>
> x=c(1760,1652,1485,1390,1820,1665,1550,1700,1270)
> y=c(4300,4650,3200,3150,4950,4010,3810,4500,3008)
> fit=lm(y~x)
> fit

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
  -1847.633        3.653
```