NAME – ANSH GOEL
REG NO. – 20BCE1798
COURSE NAME – FOUNDATION OF DATA ANALYTICS (FDA)
COURSE CODE: 3505
DATE – 6TH OCTOBER, 2022

LAB-9

Q. Take a dataset and choose two attributes. Make a sample of 20 from them. Calculate the correlation using R and verify it using manual calculations.

```
> data(mtcars)
> head(mtcars)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
> mtcars
                     mpg cyl  disp  hp drat    wt  qsec vs am gear carb
Mazda RX4           21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag       21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
Datsun 710          22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive      21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout   18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
Valiant             18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
Duster 360          14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
Merc 240D           24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
Merc 230            22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
Merc 280            19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
Merc 280C           17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
Merc 450SE          16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
Toyota Corona       21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
```

```
> df<-data.frame(mtcars)
> df
                     mpg cyl  disp  hp drat    wt  qsec vs am gear carb
Mazda RX4           21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag       21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
Datsun 710          22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive      21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout   18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
Valiant             18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
Duster 360          14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
Merc 240D           24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
Merc 230            22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
Merc 280            19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
Merc 280C           17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
Merc 450SE          16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
Toyota Corona       21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
AMC Javelin         15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
Camaro Z28          13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
Pontiac Firebird    19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
Fiat X1-9           27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
Porsche 914-2       26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
Lotus Europa        30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
Ford Pantera L      15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
Ferrari Dino        19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
Maserati Bora       15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
```

```
> df_20<-df[sample(nrow(df),size=20),]
> df_20
                     mpg cyl  disp  hp drat    wt  qsec vs am gear carb
Merc 450SLC         15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
Hornet Sportabout   18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
Chrysler Imperial   14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
Fiat X1-9           27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
Merc 280            19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
Duster 360          14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
Toyota Corolla      33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
Datsun 710          22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
Merc 450SL          17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
Honda Civic         30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
Ferrari Dino        19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
Hornet 4 Drive      21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
Ford Pantera L      15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
Lotus Europa        30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
Dodge Challenger    15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
Cadillac Fleetwood  10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
Fiat 128            32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
Merc 240D           24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
Merc 280C           17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
> x<-df_20$mpg
> y<-df_20$cyl
```

```
> x
 [1] 15.2 18.7 14.7 27.3 19.2 14.3 33.9 22.8 17.3 30.4 19.7 21.4 15.8 30.4 15.5 10.4 32.4 24.4 17.8 10.4
> y
 [1] 8 8 8 4 6 8 4 4 8 4 6 6 8 4 8 8 4 4 6 8
>
> cor(x, y, method = c("pearson"))
[1] -0.8977234
> cor(x, y, method = c("kendall"))
[1] -0.8089641
> cor(x, y, method = c("spearman"))
[1] -0.917594
>
```

Manual Calculations:

Manual calculations

Page No.
Date : / /

x	y	dx $(x-\bar{x})$	dy $(y-\bar{y})$	dxdy
15.2	8	-5.4	1.8	-9.72
18.7	8	-1.9	1.8	-3.42
14.7	8	5.9	1.8	10.62
27.3	4	6.7	-2.2	-14.74
19.2	6	-1.4	-0.2	0.28
14.3	8	-6.3	1.8	-11.34
33.9	4	13.3	-2.2	-29.26
22.8	4	2.2	-2.2	-4.84
17.3	8	-3.3	1.8	-5.94
30.4	4	9.8	-2.2	-21.56
19.7	6	-0.9	-0.2	0.18
21.4	6	0.8	-0.2	-.16
15.8	8	-4.8	1.8	8.64
30.4	4	9.8	-2.2	-21.56
15.5	8	-5.1	1.8	-9.18
10.4	8	-10.2	1.8	-18.36
32.4	4	11.8	-2.2	-25.96
24.4	4	3.8	-2.2	-8.36
17.8	6	-2.8	-0.2	.56
10.4	8	-10.2	1.8	-18.36

$\bar{x} = 20.6$

$\bar{y} = 6.2$

$\Sigma dx^2 = 958.92$

$\Sigma dy^2 = 63.2$

$\Sigma dxdy = -221$

Asc order

x	d^2	d	Rank of y	Rank of x	dy^2	dx^2
1·5	121	-11	16	5	3·24	24·16
1·5	36	-6	16	10	3·24	3·61
3	144	-12	16	4	3·24	3481
4	144	12	4	16	4·84	44·89
5	225	15	9·5	11	0·04	1·96
6	169	-13	16	3	3·24	39·69
7	256	16	4	20	4·84	176·89
8	100	10	4	14	4·84	4·48
9	64	-8	16	8	3·24	10·89
10	182·25	13·5	4	12·5	4·84	96·04
11	6·25	2·5	9·5	11	0·04	0·81
12	12·25	3·5	9·5	13	0·04	0·64
13	81	9	16	7	3·24	23·04
14	182·25	13·5	4	17·5	4·84	96·04
15	100	-10	16	6	2·24	26·01
16	210·25	-14·5	16	1·5	3·24	104·00
17·5	225	15	4	19	4·84	129·24
17·5	121	11	4	15	4·84	1441
19	0·25	-0·5	9·5	9	0·04	7·84
20	210·25	-14·5	16	1·5	3·24	104·04

Page No.
Date : / /

$\Sigma d^2 = 2367.$

Correlation using pearson's method

$$r = \frac{\Sigma dx\,dy}{\sqrt{dx^2}\sqrt{\Sigma dy^2}} = \frac{-221}{\sqrt{458.92}\sqrt{65.2}}$$

$$= -0.8977234012$$

correlation using spearman's method.

$$r = 1 - 6\left[\frac{\Sigma D_i^2 + \frac{1}{12}(m_1^3 - m_1) + \frac{1}{12}(m_2^3 - m_2) + \frac{1}{12}(m_3^3 - m_3) \cdots}{n^3 - n}\right]$$

$$= 1 - 6\left[\frac{2367 + \frac{1}{12}(7^3-7) + \frac{1}{12}(9^3-9) + \frac{1}{12}(4^3-4) + \frac{1}{12}(2^3-2)}{20(399)}\right]$$

$$= -0.87$$

Rank → n Kendall's Method.

$$\tau = \frac{\Sigma C - \Sigma D}{\Sigma C + \Sigma D} \Rightarrow -0.809$$

Result:

The manual calculations and R answers match each other.