Experiment-3

Applying correlation and simple linear regression model to real dataset

Aim: To understand the simple correlation and linear regression with computation and interpretation.

Code and Result:

```
# cars

# Import the inbuilt data set "cars"
```

speed	dist
4	2
4	10
7	4
7	22
8	16
9	10
10	18
10	26
10	34
11	17
11	28
12	14
12	20
12	24
12	28
13	26
13	34
13	34
13	46
14	26
14	36
14	60
14	80
15	20
15	26
15	54

16	32
16	40
17	32
17	40
17	50
18	42
18	56
18	76
18	84
19	36
19	46
19	68
20	32
20	48
20	52
20	56
20	64
22	66
23	54
24	70
24	92
24	93
24	120
25	85

correlation coefficient using Pearson's formula

Test for the association between paired samples

Visualize the samples

Linear Regression model of "speed" with respect to "dist"

Visualize linear regression line

Linear Regression model of "dist" with respect to "speed"

Practice problems:

1. The body Weight and BMI of 12 school-going children are given in the following table:

Weight	15	26	27	25	25.5	27	32	18	22	20	26	24
BMI	13.3	16.12	16.74	16	13.59	15.73	15.65	13.85	16.07	12.8	13.65	14.42

   i. Find the correlation coefficient between weight and BMI and visualize the scatter plot.
   ii. Find the rank correlation coefficient between weight and BMI and visualize the scatter plot.
   iii. Find linear regression line weight for BMI and visualize.
   iv. Find linear regression line BMI for weight and visualize.

2. The following table gives the weight (in 1000 lbs.) and highway fuel efficiency (in miles/gallon) for 13 cars.

Vehicle	**X**	**Y**
Chevrolet Camaro	3.545	30
Dodge Neon	2.6	2.6
Honda Accord	3.245	30
Lincoln Continental	3.93	24
Oldsmobile Aurora	3.995	26
Pontiac Grand Am	3.115	30
Mitsubishi Eclipse	3.235	33
BMW 3-Series	3.225	27
Honda Civic	2.44	37
Toyota Camry	3.24	32

Hyundai Accent	2.29	37
Mazda Protégé	2.5	34
Cadillac DeVille	4.02	26

Then, find the correlation coefficient, rank correlation coefficient, and linear regression lines.

3. The following table shows the trend of cinema admissions and TV sets growth in a locality from 1974-1980. Calculate the product-moment correlation coefficient between the two variables.

Year	1974	1975	1976	1977	1978	1979	1980
Admissions (in thousands)	13	12	9	9	8	6	6
No. of TV sets	54	53	57	61	67	72	70

4. Calculate Karl Pearson's coefficient of correlation from the following data

X	46	33	41	38	36	45	34	37	50	40
y	12	13	24	16	15	14	21	17	19	19

5. Twelve recruits were subjected to a selection test to ascertain their suitability for a particulartraining course. At the end of training, they were given a proficiency test. The marks scored by the recruits are recorded below.

Recruit	1	2	3	4	5	6	7	8	9	10	11	12
Selection Test Score	44	49	52	54	47	76	65	60	63	58	50	67
Proficiency Test Score	48	55	45	60	43	80	58	50	77	46	47	65

Calculate the rank correlation coefficient.