Individual Task-III

CORRELATION EXPLORATION

1.Introduction

Correlation analysis is a statistical method used to examine the relationship between two variables. It shows how changes in one variable affect another. Understanding correlations is useful in daily life, academics, and business decisions.

In this task, we explore the relationship between study hours per day and academic scores. The goal is to calculate the correlation coefficient, visualize the relationship, and interpret the results.

2.Dataset:

The data set was collected from a small group of students.

Student	Study hours per day	Score(%)
John	2	60
Ram	3	70
David	4	80
Ganesh	5	85
Anu	6	90

Explanation:

- Study Hours per Day (X): Number of hours a student studies daily.
- Score (%)(Y): Academic performance measured as percentage marks.

This dataset is small but sufficient to illustrate correlation analysis.

3. Correlation Calculation:

The Pearson correlation coefficient (r) is used to measure the strength and direction of a linear relationship between two variables.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

Where,

x-Study hours

y-Scores

n-Number of Students

Step 1:Prepare Sums

X	Y	XY	X^2	Y^2
2	60	120	4	3600
3	70	210	9	4900
4	80	320	16	6400
5	85	425	25	7225
6	90	540	36	8100

- $\sum X = 2+3+4+5+6 = 20$
- $\sum Y = 60+70+80+85+90 = 385$
- $\sum XY = 120+210+320+425+540 = 1615$
- $\sum X^2 = 4+9+16+25+36 = 90$
- $\sum Y^2 = 3600+4900+6400+7225+8100 = 30225$
- $n = 5$

Apply Formula:

$$r = \frac{5(1615) - (20)(385)}{\sqrt{[5(90) - 20^2][5(30,225) - 385^2]}}$$

$$r = \frac{375}{\sqrt{50 \cdot 2,900}}$$

$$r = \frac{375}{\sqrt{145,000}}$$

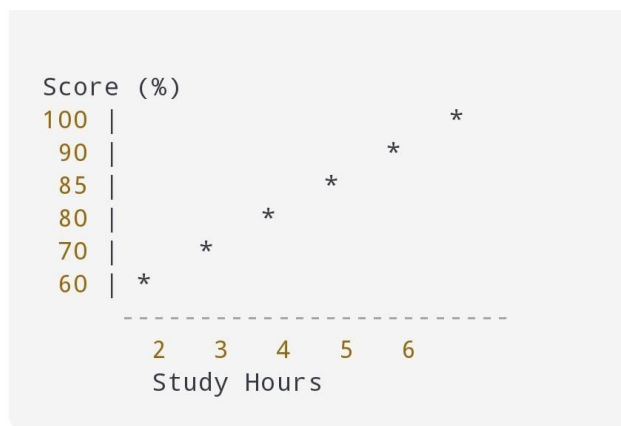$$r = \frac{375}{380.79} \approx 0.984$$

4.Interpretation:

- $r \approx 0.98$ indicates a strong positive correlation between study hours and academic scores.

- Explanation: As the number of study hours increases, students' scores also increase.
- A correlation close to +1 means the relationship is very strong and nearly linear.

Practical implication: Students who dedicate more time to studying are likely to perform better academically.

5. Graphical Representation

A scatter plot can help visualize the relationship:

```
Score (%)
100 |                         *
 90 |                   *
 85 |              *
 80 |          *
 70 |       *
 60 |  *

     - - - - - - - - - - - - - - - - - - -
       2    3    4    5    6
       Study Hours
```

- Points form an upward sloping line.
- Trendline indicates positive linear correlation.

6. Applications of Correlation Analysis:

Academic Planning: Helps students understand how study habits impact performance.

Goal Setting: Students can set study targets to achieve desired scores.

Education Research: Teachers can identify patterns in study time and outcomes.

Predictive Insights: Educators can predict scores based on study hours.

7. Conclusion:

Correlation analysis is a useful tool for exploring relationships between variables.

For this dataset:

- Strong positive correlation ($r \approx 0.98$) was observed between study hours and scores.
- Increasing study hours is likely to improve academic performance.

This exercise demonstrates how even small datasets from daily life can reveal meaningful patterns and guide decisions.