

LINEAR REGRESSION (LEAST SQUARES METHOD) – DOCUMENTATION

This program performs **Simple Linear Regression** using the **Least Squares Method** to find the best-fit straight line for a given set of data points.

It supports:

- Multiple test cases
- Input from **input.txt**
- Output to **console and output.txt**
- Clear reporting of input, computations, and final regression equation

OBJECTIVE

To compute the best-fit linear equation:

$$y=a + bx$$

from a given set of experimental or numerical data using the least squares regression technique, which minimizes the total squared error.

CONCEPT

For n data points (x_i, y_i) the regression line:

$$y=a + bx$$

is determined by minimizing:

$$\sum (y_i - (a + bx_i))^2$$

This leads to two closed-form formulas:

Slope (b)

$$b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$$

Intercept (a)

$$a = \frac{\sum y - b \sum x}{n}$$

These values define the best-fit straight line.

PROGRAM FEATURES

- Reads all inputs from input.txt
- Writes results to both console and output.txt
- Handles multiple test cases
- Prints:
 - Number of data points
 - All x and y values
 - Computed intercept (a)
 - Computed slope (b)
 - Final regression line equation

INPUT FORMAT (input.txt)

T

n

x1 x2 x3 ... xn

y1 y2 y3 ... yn

(repeat for T test cases)

Where:

- t = number of test cases
- n = number of data points in the test case
- Next line = n values of x
- Next line = n values of y

EXAMPLE INPUT

3

5

1 2 3 4 5

2 4 5 4 5

4

2 4 6 8

3 5 7 9

6

5 10 15 20 25 30

12 18 26 33 40 48

OUTPUT (Written to output.txt and displayed on console)

For each test case, the program prints:

- Test Case Number
- Total points
- x[] array
- y[] array
- Intercept (a)
- Slope (b)
- Regression Equation

Example:

Test Case #1

Number of points: 5

x values: 1 2 3 4 5

y values: 2 4 5 4 5

Computed Intercept (a): 2.2

Computed Slope (b): 0.6

Line Equation: $y = 2.2 + 0.6x$

ALGORITHM (Least Squares Method)

- Read number of data points n
- Read arrays x[n] and y[n]
- Compute required sums:
 - $\sum x$
 - $\sum y$
 - $\sum xy$
 - $\sum x^2$
- Apply formulas:
 - Compute b (slope)
 - Compute a (intercept)
- Display and store:
 - Input values
 - Calculated coefficients
 - Final regression line

