

# CORRELATION

## DESCRIPTIVE STATISTICS

---

*prepared by:*

Gyro A. Madrona

Electronics Engineer

# TOPIC OUTLINE

Covariance

Correlation Coefficient



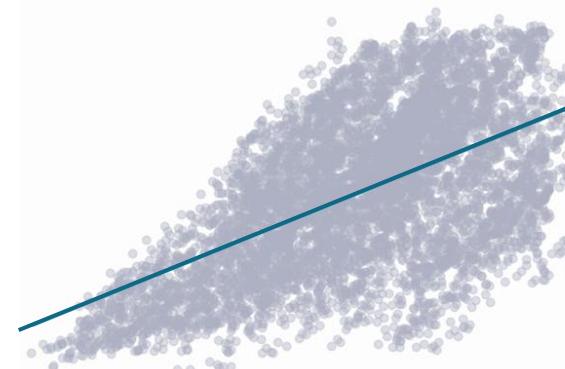
# COVARIANCE



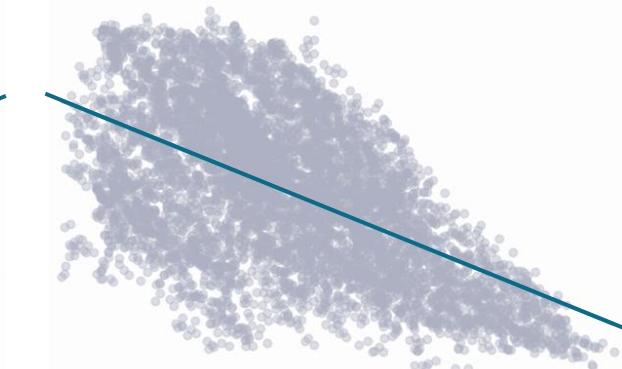
# COVARIANCE

Covariance is a statistical measure that quantifies the relationship between two random variables ( $X, Y$ ).

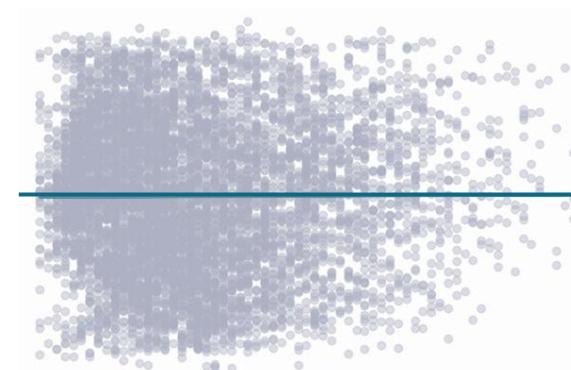
Scatter Plot



**Covariance > 0**



**Covariance < 0**



**Covariance = 0**

# COVARIANCE

Covariance is a statistical measure that quantifies the relationship between two random variables ( $X, Y$ ).

## Population Covariance

$$\sigma_{xy} = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{N}$$

## Sample Covariance

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

# COVARIANCE

---

Covariance is a statistical measure that quantifies the relationship between two random variables ( $X, Y$ ).

The `df.cov()` method is used to compute the covariance matrix of a DataFrame.

## EXERCISE

The given dataset contains five observations of current (A) and corresponding power (W) measurements. Does **current** and **power** consumption have a positive, negative, or no linear relationship?

solution

Device	
Current	Power
2	100
3.5	200
1.8	90
4.2	210
2.7	110



# CORRELATION COEFFICIENT

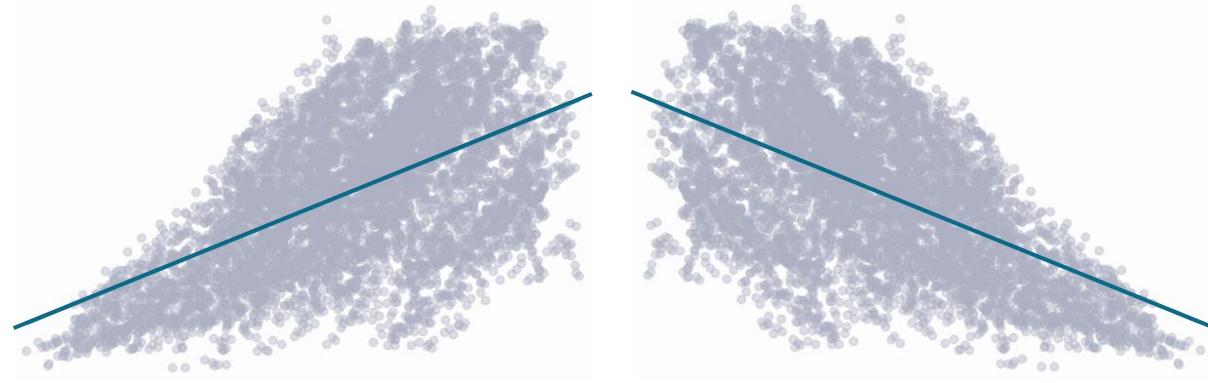
# CORRELATION COEFFICIENT

Correlation coefficient adjusts covariance, so that the relationship between the two variables becomes easy and intuitive to interpret.

It ranges from **-1** to **+1**:

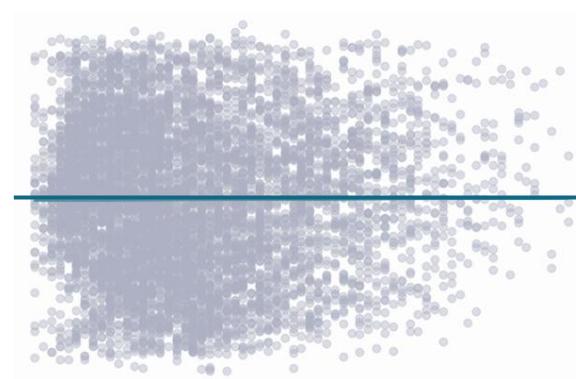
- +1** indicates perfect positive correlation
- 1** indicates perfect negative correlation
- 0** indicates no linear relationship

Scatter Plot



**Correlation > 0**

**Correlation < 0**



**Correlation = 0**

# CORRELATION COEFFICIENT

Correlation coefficient adjusts covariance, so that the relationship between the two variables becomes easy and intuitive to interpret.

It ranges from **-1** to **+1**:

- +1** indicates perfect positive correlation
- 1** indicates perfect negative correlation
- 0** indicates no linear relationship

## Population Correlation Coefficient

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

## Sample Correlation Coefficient

$$r = \frac{s_{xy}}{s_x s_y}$$

# CORRELATION COEFFICIENT

Correlation coefficient adjusts covariance, so that the relationship between the two variables becomes easy and intuitive to interpret.

It ranges from **-1** to **+1**:

- +1** indicates perfect positive correlation
- 1** indicates perfect negative correlation
- 0** indicates no linear relationship

The `df.corr()` method is used to compute the correlation matrix of a DataFrame.

# EXERCISE

Determine if each scenario suggests a positive, negative, or no correlation:

- |  |                |
|--|----------------|
| 1. Ice cream sales and umbrella sales in a city. | negative       |
| 2. Hours spent studying and exam scores.         | positive       |
| 3. A person's shoe size and their IQ.            | no correlation |
| 4. Age of a used car and its resale value.       | negative       |



## EXERCISE

---

The given dataset contains five observations of current (A) and corresponding power (W) measurements. Does **current** and **power** consumption have a positive, negative, or no linear relationship?

solution

Device	
Current	Power
2	100
3.5	200
1.8	90
4.2	210
2.7	110

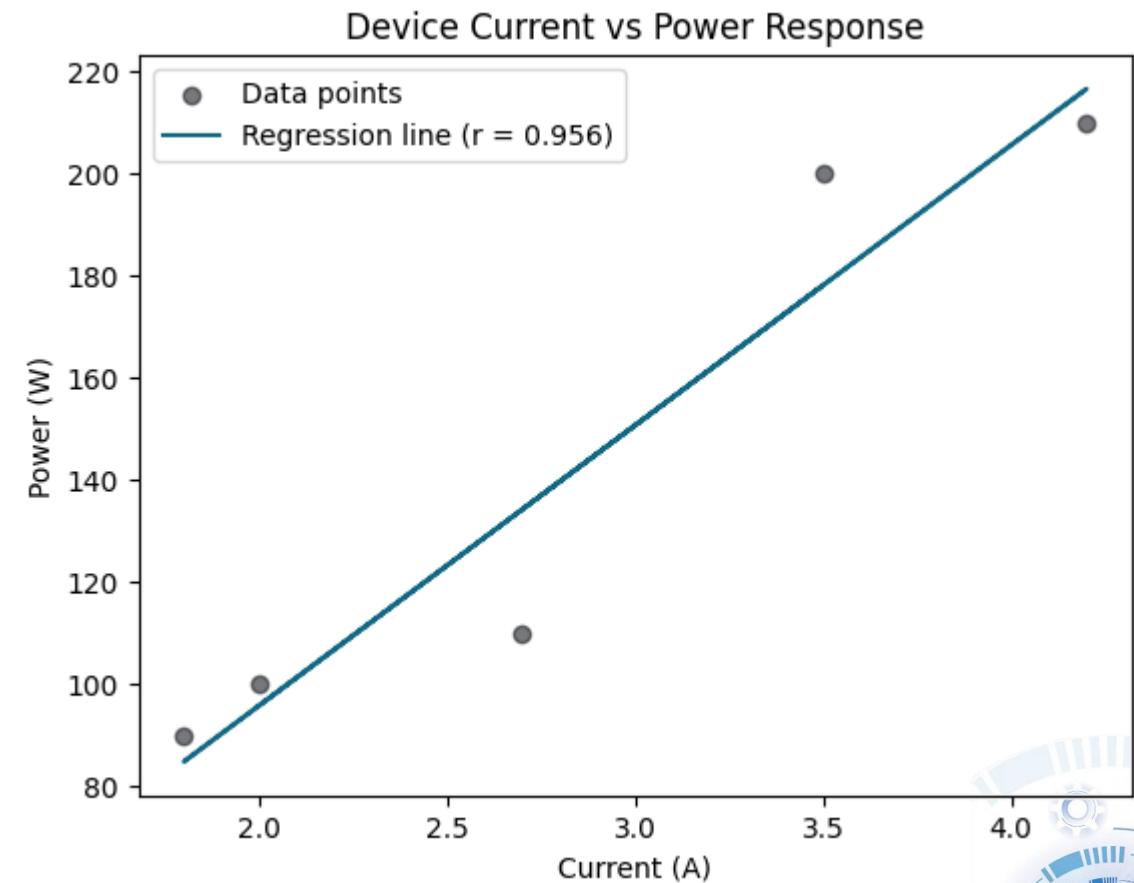


## EXERCISE

The given dataset contains five observations of current (A) and corresponding power (W) measurements. Does **current** and **power** consumption have a positive, negative, or no linear relationship?

Device	
Current	Power
2	100
3.5	200
1.8	90
4.2	210
2.7	110

solution



# LABORATORY