What is Regression Analysis?

Regression analysis is a statistical method used to model the relationship between a dependent variable and one or more independent variables. It helps in predicting outcomes, identifying trends, and understanding the impact of various factors on a given outcome.

Swipe for more

Why Do We Use Regression Analysis?

Regression analysis is used to:

- Predict future values based on historical data.
- Understand relationships between variables.
- Identify important factors affecting the outcome.
- Make informed decisions based on data trends.

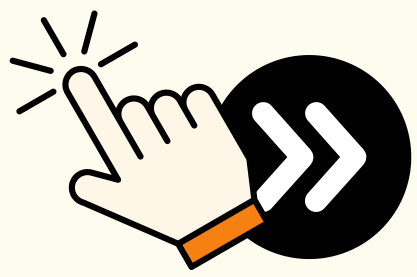
Swipe for more

What Do We Get from Using Regression Analysis?

By using regression analysis, we gain:

- Accurate predictions of outcomes.
- Insights into the strength and nature of relationships between variables.
- A better understanding of the data and the ability to make data-driven decisions.

Swipe for more



Types of Regression Analysis in Python

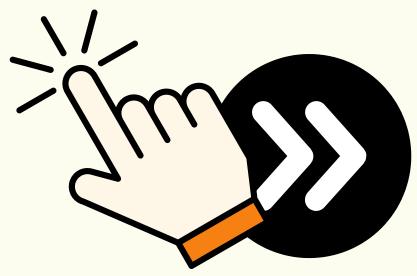
Swipe for more

1. Linear Regression

Definition:

Models the relationship between a dependent variable and one or more independent variables using a straight line.

Swipe for more



Use Case:

Used when the relationship between variables is linear. Commonly used in finance, economics, and biology for predicting continuous outcomes.

Example:

Predicting housing prices based on square footage.

Swipe for more

2. Logistic Regression

Definition:

Used for binary classification, modeling the probability that a given input belongs to a particular category.

Swipe for more

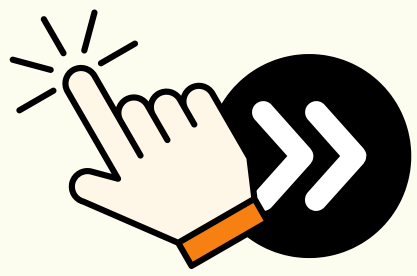
Use Case:

Useful for classification problems such as predicting whether an email is spam or not, or whether a customer will make a purchase.

Example:

Predicting if a customer will buy a product (yes/no)

Swipe for more



3. Polynomial Regression

Definition:

Extends linear regression by fitting a polynomial curve to the data, allowing for a non-linear relationship between variables.

Swipe for more

Use Case:

Effective when the data shows a curvilinear relationship. Used in scenarios like modeling the growth of a population or sales trends over time.

Example:

Modeling the growth rate of a species population over time.

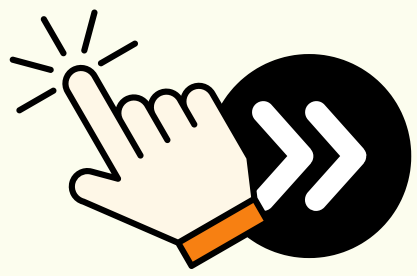
Swipe for more

4. Ridge Regression

Definition:

A type of linear regression that includes a regularization term to prevent overfitting by penalizing large coefficients.

Swipe for more



Use Case:

Ideal when there are many features, and multicollinearity (correlation between independent variables) is a concern.

Example:

Predicting stock prices with multiple correlated financial indicators.

Swipe for more

5. Lasso Regression

Definition:

Similar to ridge regression but can shrink some coefficients to zero, effectively selecting a subset of features.

Swipe for more

Use Case:

Useful for feature selection when dealing with high-dimensional data, reducing the number of variables in the model.

Example:

Selecting the most important factors affecting house prices.

Swipe for more