Data Science

Especialization

A Mathematic Driven Approach

Quick Recap

Functions

```python
# sample_function.py > ...
1    import json
2
3    def greet():
4        print('Hello user')
5
6    def main():
7        greet()
8
9    if __name__ == '__main__':
10       main()
```

Write Once, Run Everywhere
- With functions you can write a repetitive block of code and utilize it as many times as you please

More readability
- Functions provide a less conglomerated code, making it easier to read

Unique Qualities
- Functions put on the table elements such as recursion that provide a smaller O(n)
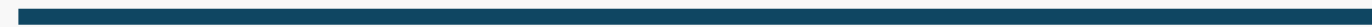
Let's Try It

The OOP Paradigm

What is OOP?

Object-oriented programming (OOP) is a paradigm that organizes code into reusable objects containing data (attributes) and behavior (functions), promoting modularity, inheritance, and encapsulation.
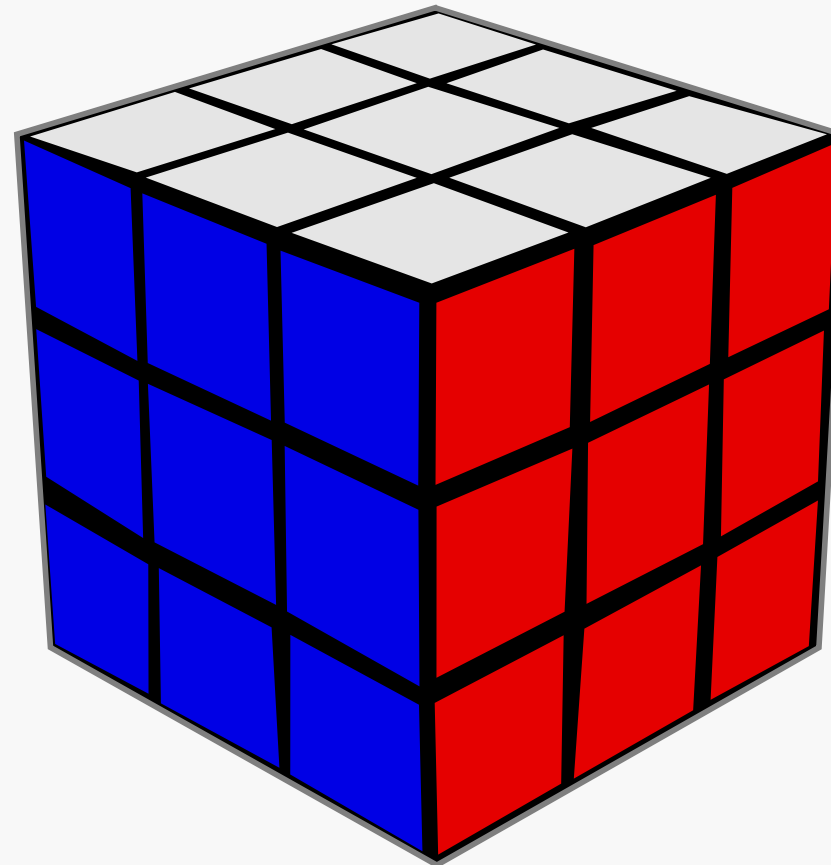
The Pilars of OOP

Abstraction

Simplifying complex objects by showing only essential details.

Inheritance

Allowing a class to derive from another, reusing attributes and methods.

Encapsulation

Protecting data by controlling access through methods.

Polymorphism

Enabling methods to have different behaviors depending on the context.

Let's Try It

Now, let's get to business

Data Science

What is Data Science?

●●●●●

Data science is an interdisciplinary field that uses statistical methods, algorithms, and tools to extract insights from data, enabling better decision-making, predictions, and solutions to complex problems.

●●●●●

What does it do?

Fraud Detection

Identify fraudulent transactions using machine learning models.

Recommendation Systems

Suggest products, movies, or content based on user behavior.

Healthcare Analytics

Predict diseases, optimize treatments, and improve patient care.

Customer Segmentation

Analyze customer data to target personalized marketing campaigns.

It Sounds Amazing

It isn't...

Data Science algorithms may have bias

Does it really affect us?

• • • • •

Yes, algorithmic bias in data science can affect us by reinforcing stereotypes, creating unfair outcomes, and impacting decision-making in areas like hiring, lending, healthcare, and law enforcement, leading to societal inequalities.

• • • • •

Examples of Coded Bias

COMPAS Algorithm

The tool used for risk assessments in sentencing was found to disproportionately flag Black defendants as high-risk,.

Amazon Hiring Tool

An AI system used for hiring was found to favor male candidates over females due to being trained on past hiring data, which was biased toward men.
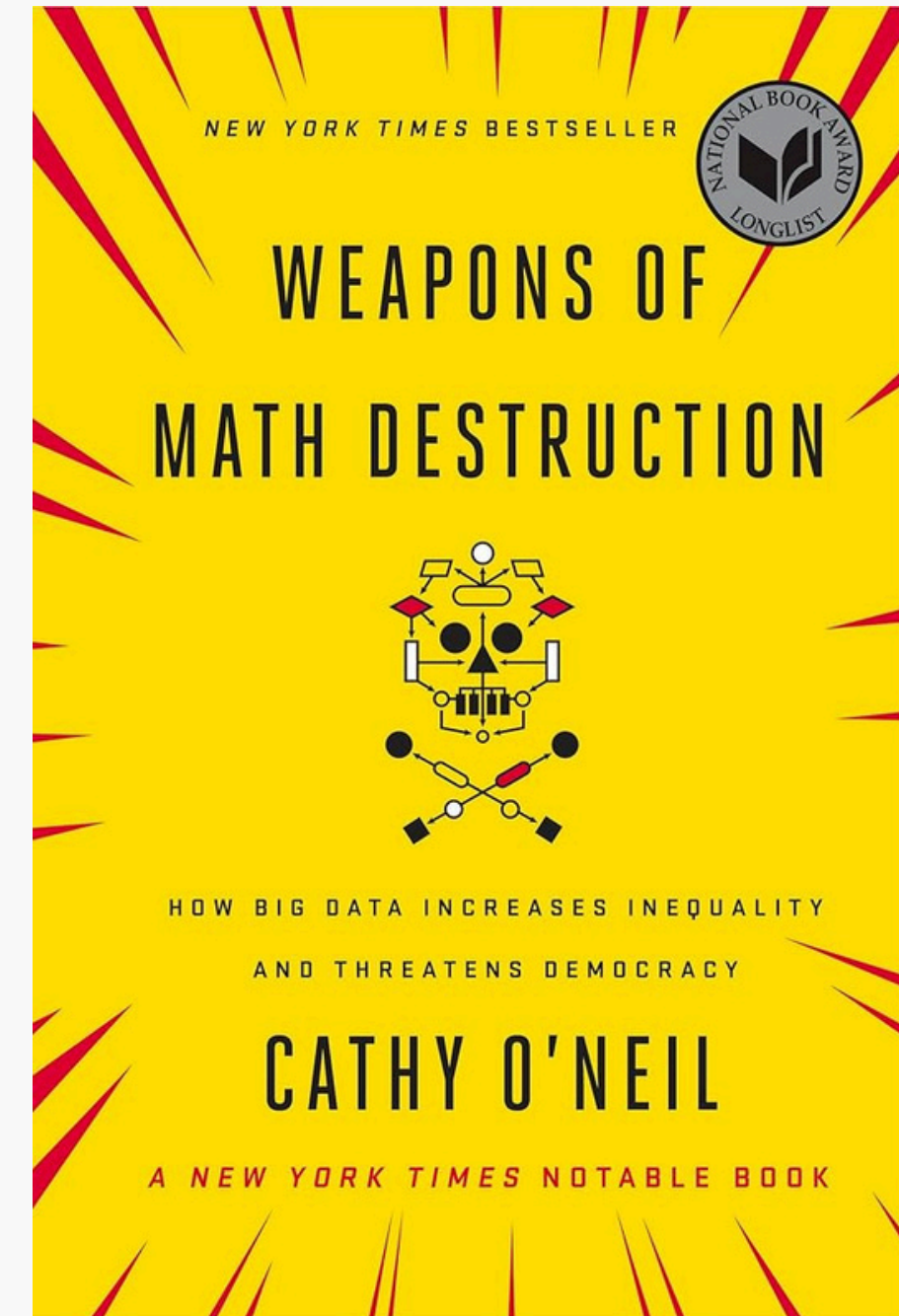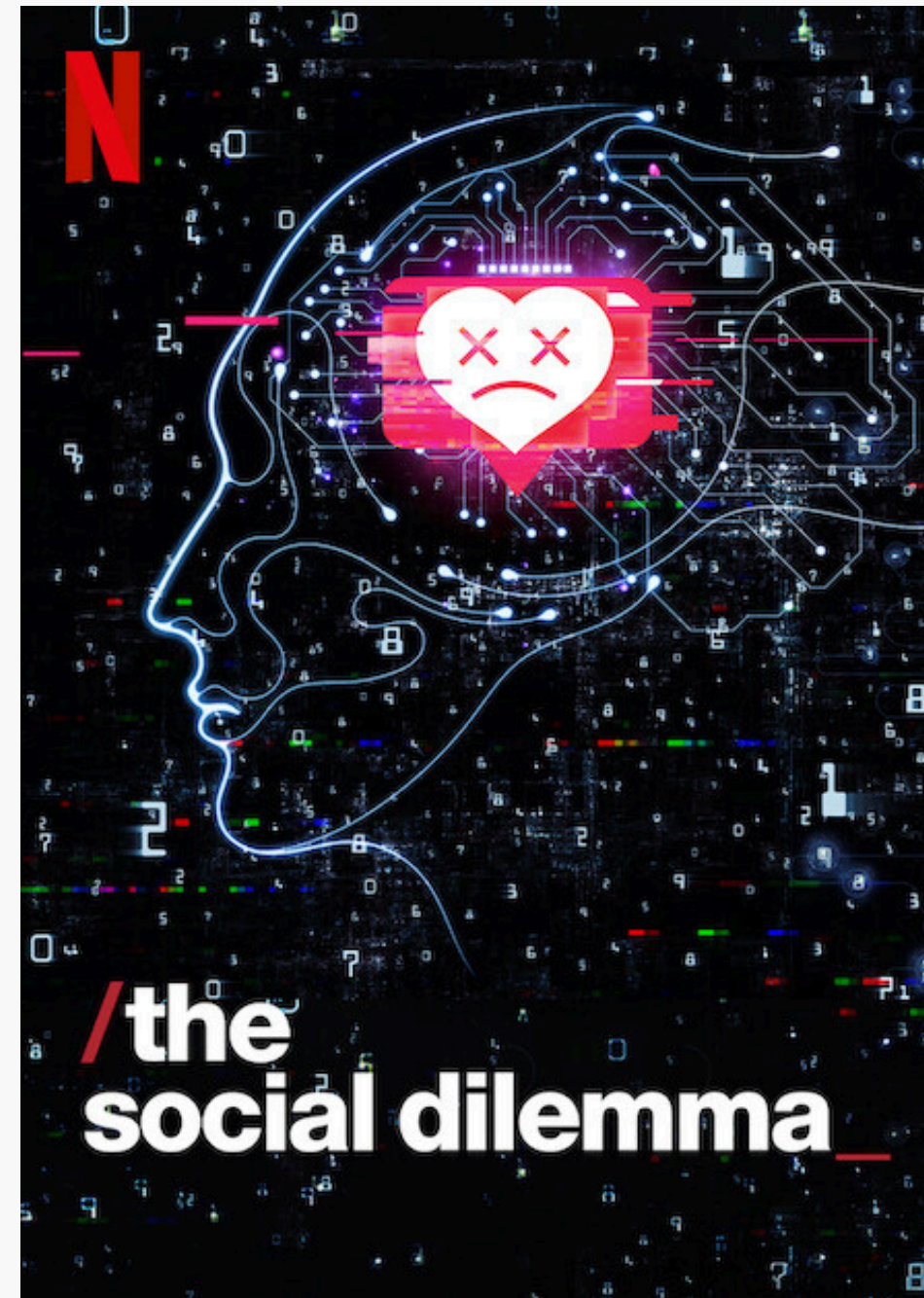
Google Photos (2015)

Google Photos' image recognition algorithm mistakenly labeled Black people as gorillas, highlighting racial bias in the training data

Apple Card (2019)

The Apple Card's credit algorithm was criticized for offering lower credit limits to women compared to men with similar financial profiles, causing concerns about gender bias.

You Should Check This Out

CODED BIAS

A SHALINI KANTAYYA FILM

OFFICIAL SELECTION 2020
sundance
film festival

N

/the
social dilemma

NEW YORK TIMES BESTSELLER

NATIONAL BOOK AWARD LONGLIST

**WEAPONS OF
MATH DESTRUCTION**

HOW BIG DATA INCREASES INEQUALITY
AND THREATENS DEMOCRACY

CATHY O'NEIL

A *NEW YORK TIMES* NOTABLE BOOK

But it can be better

It's up to us to deliver secure, well-rounded solutions

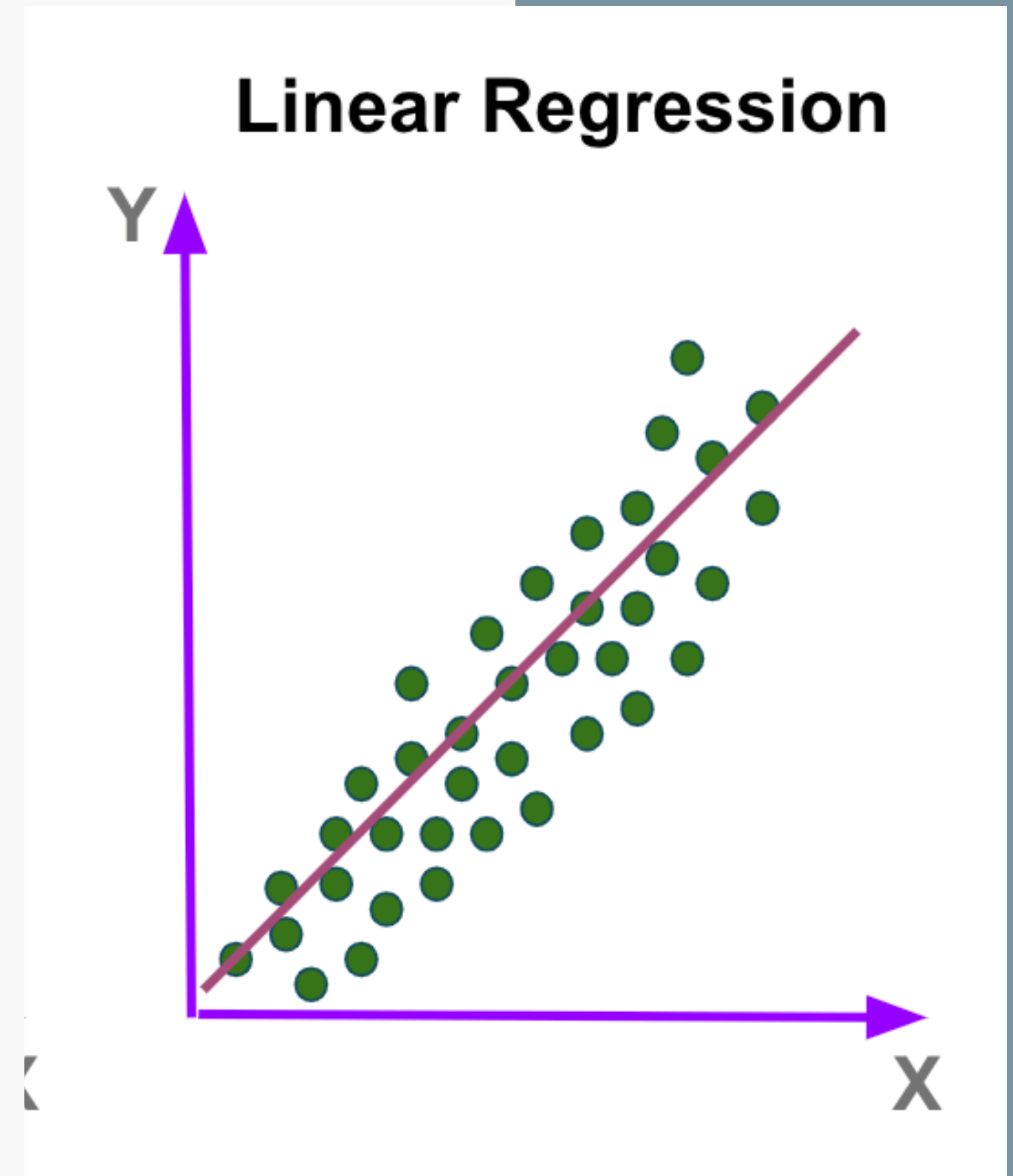*So let's make the difference*

The Linear Regression

The world's most powerful algorithm

What is it?

**A simple correlation between two variables**

Linear regression is a statistical method that models the relationship between a dependent variable and one or more independent variables, aiming to find the best-fitting line to predict outcomes.
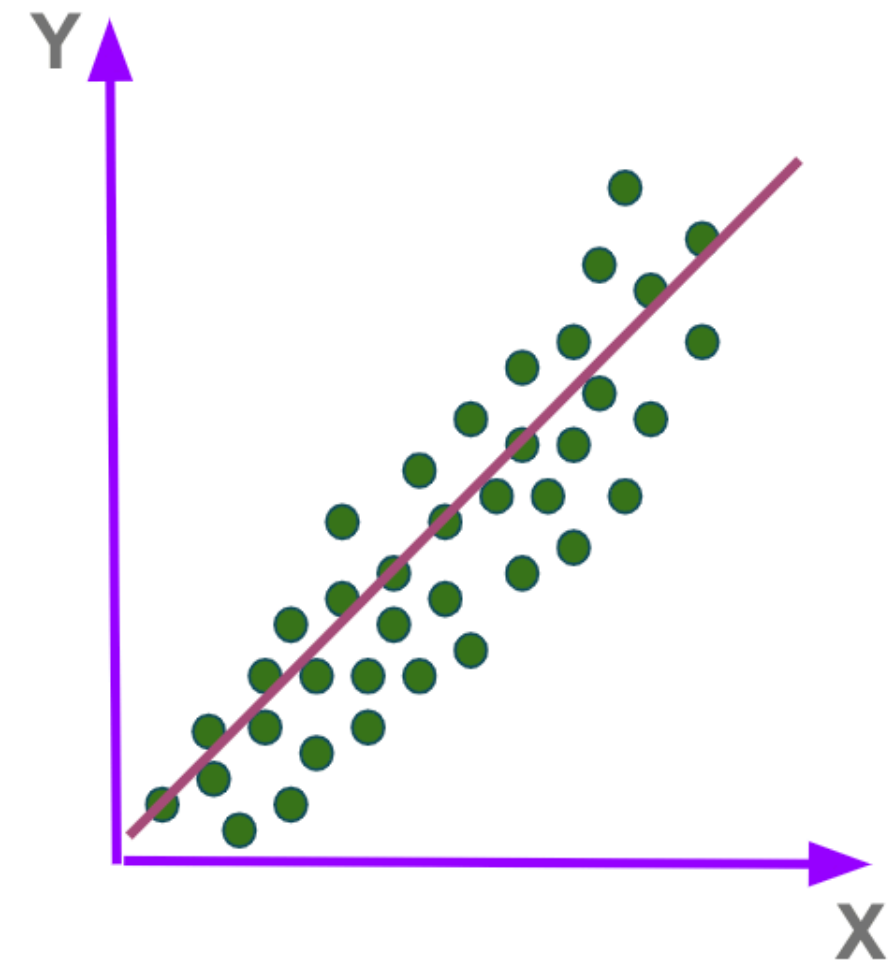
How does it work?

Linear regression works by finding the line (or hyperplane) that best fits the data points, minimizing the difference between the predicted and actual values. This is done through the least squares method.

Correlation
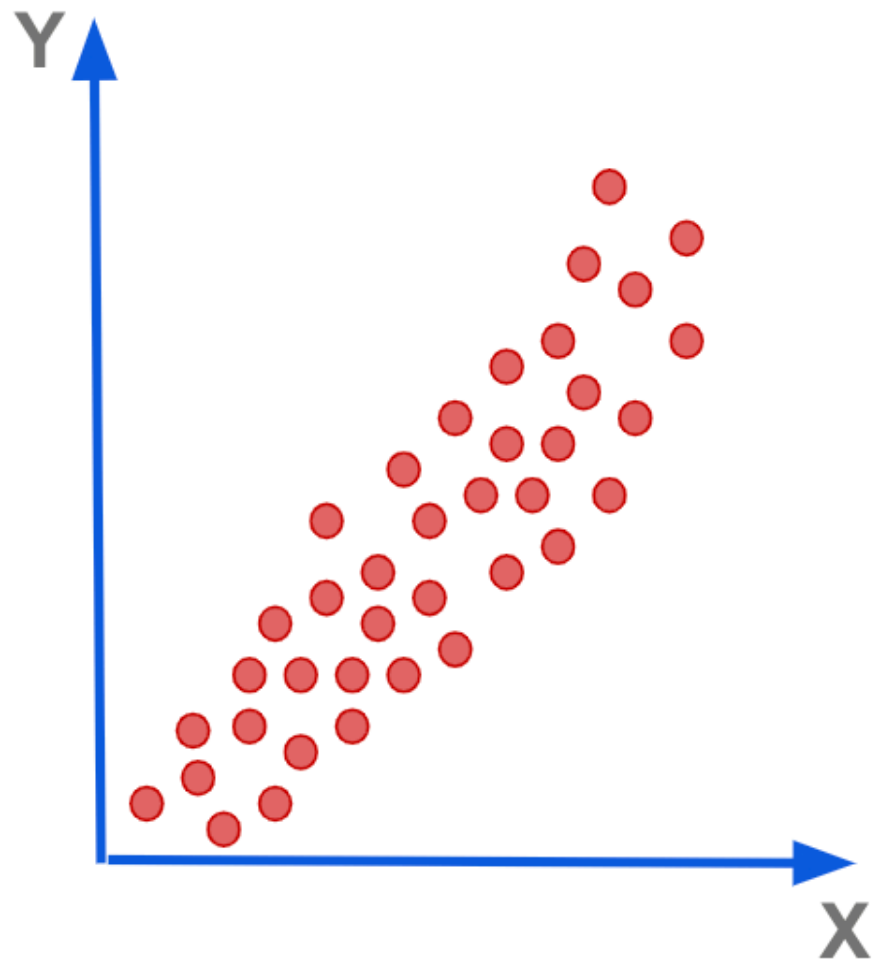
Linear Regression

How can we do it?

Using This:

$$m = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

Let's Try It

(Manually)

How can we assure it's quality?

Using This:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

Let's Try It

(Manually)

Now, Let's Code It

Can we make it better?

With This:

Let's See It

(Additional Instalation Required)

Thank you