Linear Regression

Regression

Regression is a well-known statistical technique to model the predictive relationship between several independent variables (DVs) and one dependent variable.

Regression

- Linear Regression.
- Multiple regression.

Introduction to Linear Regression

Linear regression may be defined as the statistical model that analyzes the linear relationship between a dependent variable with given set of independent variables.

Mathematically the relationship can be represented with the help of following equation :-

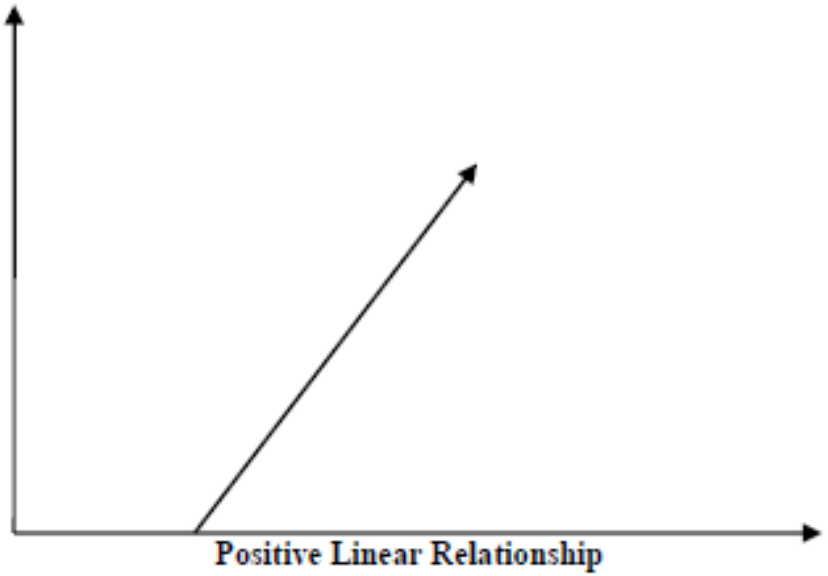
$$Y=mX+b$$

Here,

- Y is the dependent variable we are trying to predict.
- X is the independent variable we are using to make predictions.
- m is coefficient and b is intercept

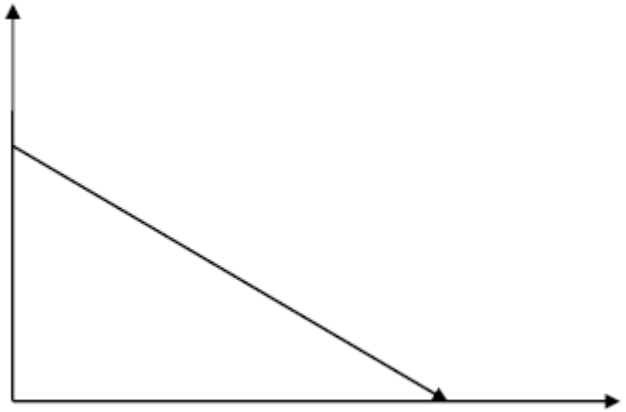
Positive Linear Relationship

- A linear relationship will be called positive if both independent and dependent variable increases. It can be understood with the help of following graph –



Negative Linear relationship

A linear relationship will be called Negative if independent increases and dependent variable decreases. It can be understood with the help of following graph −



Negative Linear Relationship

Linear Regression

The general mathematical equation for a linear regression is:

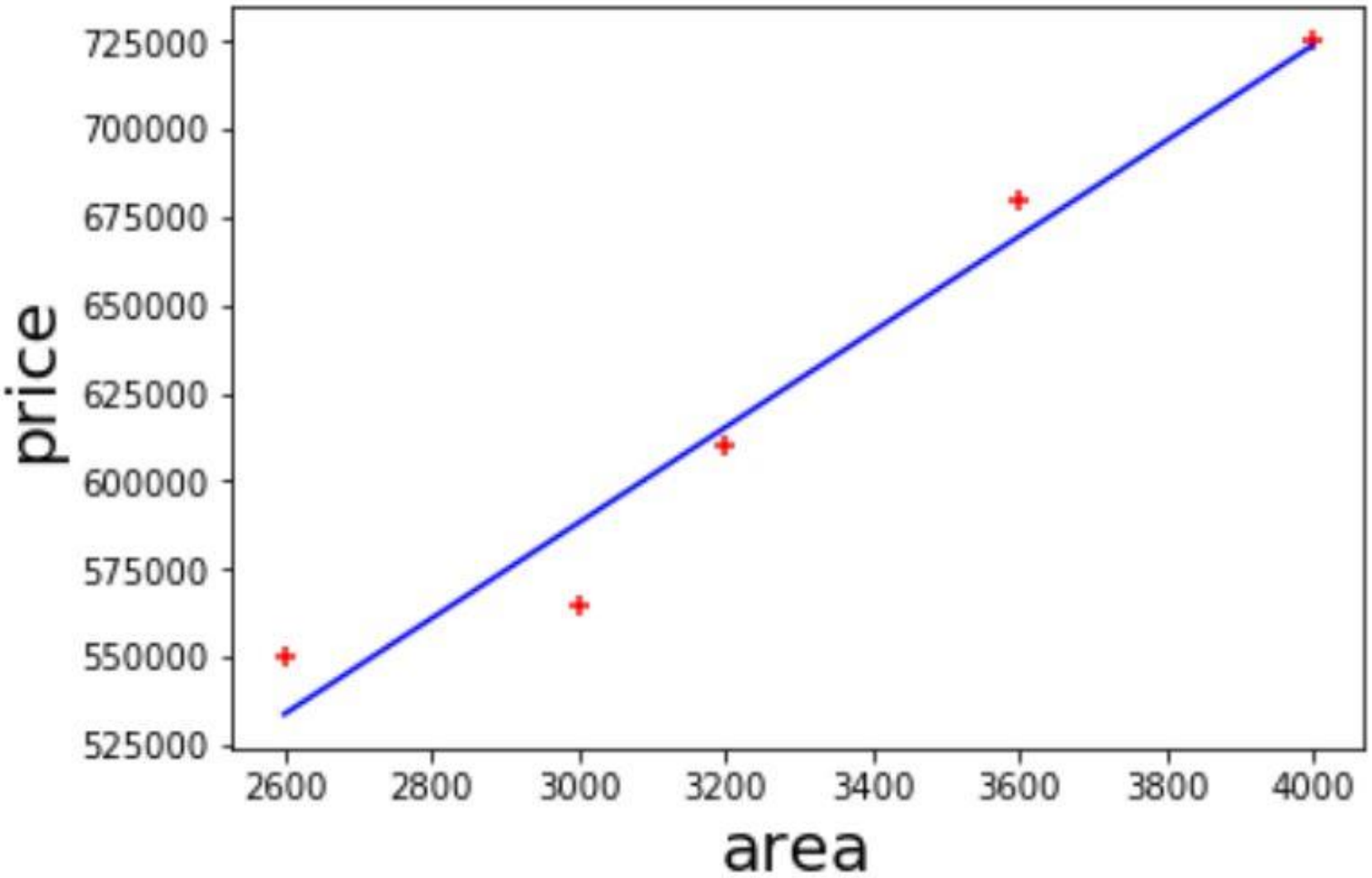
$$y = mx+b$$

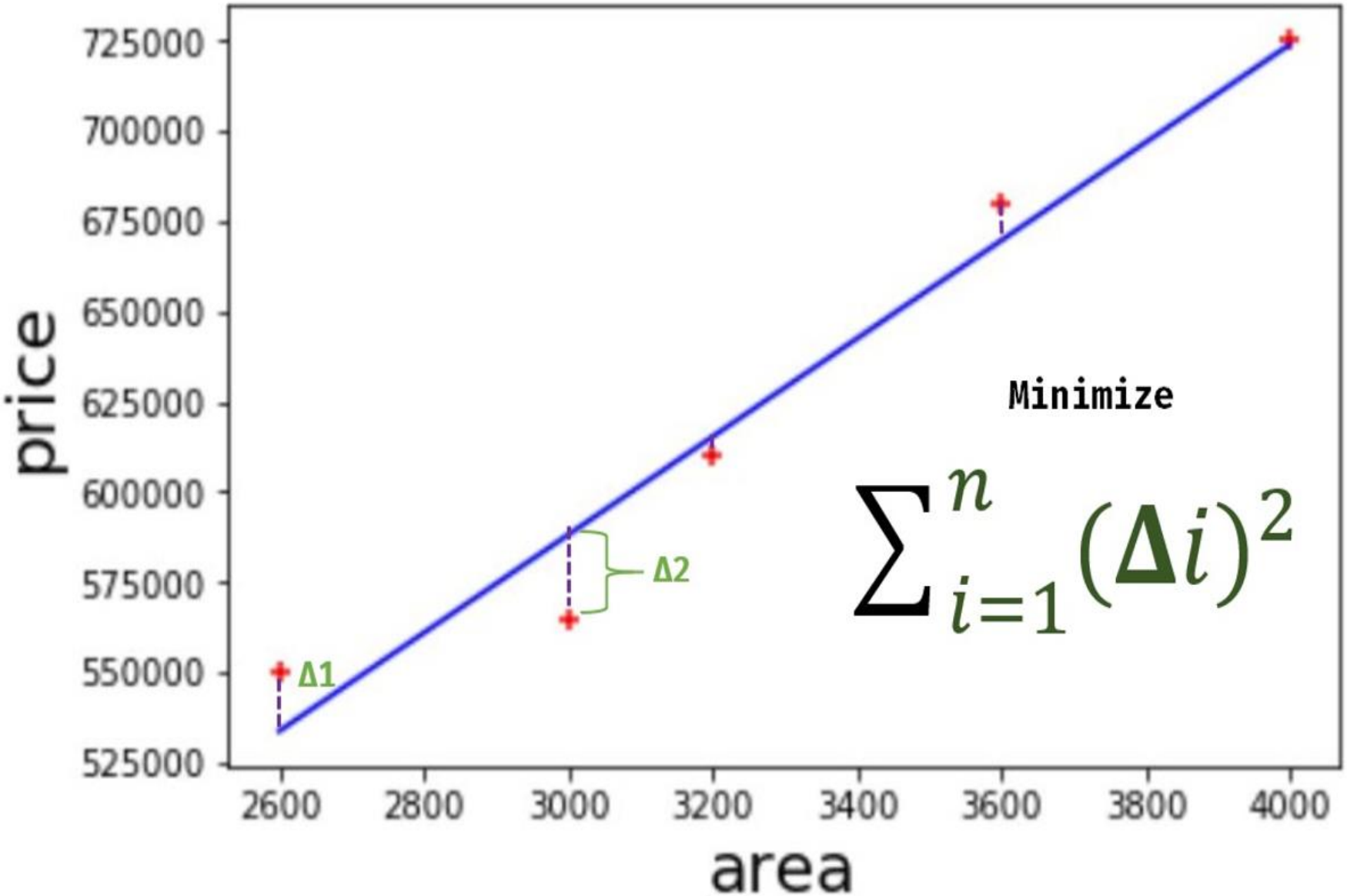
- Y is dependent variable
- X is Independent variable
- **m is coefficient**
- **b is intercept**

Predicting home price in monroe, new jersey (USA)

area	price
2600	550000
3000	565000
3200	610000
3600	680000
4000	725000

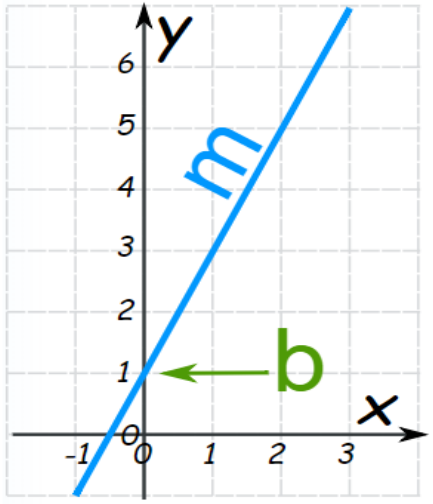
Problem Statement: Given above data build a machine learning model that can predict home prices based on square feet area


725000
700000
675000
650000
625000
600000
575000
550000
525000
price
2600
2800
3000
3200
3400
3600
3800
4000
area


725000
700000
675000
650000
625000
600000
575000
550000
525000
price
2600
2800
3000
3200
3400
3600
3800
4000
area
Δ1
Δ2
Minimize
$\sum_{i=1}^{n} (\Delta i)^2$

Home prices can be presented as following equation,

home price = m * (area) + b



price = m * area + b

$$y = mx + b$$

Slope (or Gradient) Y Intercept

Reference: https://www.mathsisfun.com/algebra/linear-equations.html

(1)Predict price of a home with area = 3300 sqr ft

(2)Predict price of a home with area = 5000 sqr ft

Exercise

- Predict canada's per capita income in year 2020. Consider **canada_per_capita_income.csv** file. Using this build a regression model and predict the per capita income for Canadian citizens in year 2020