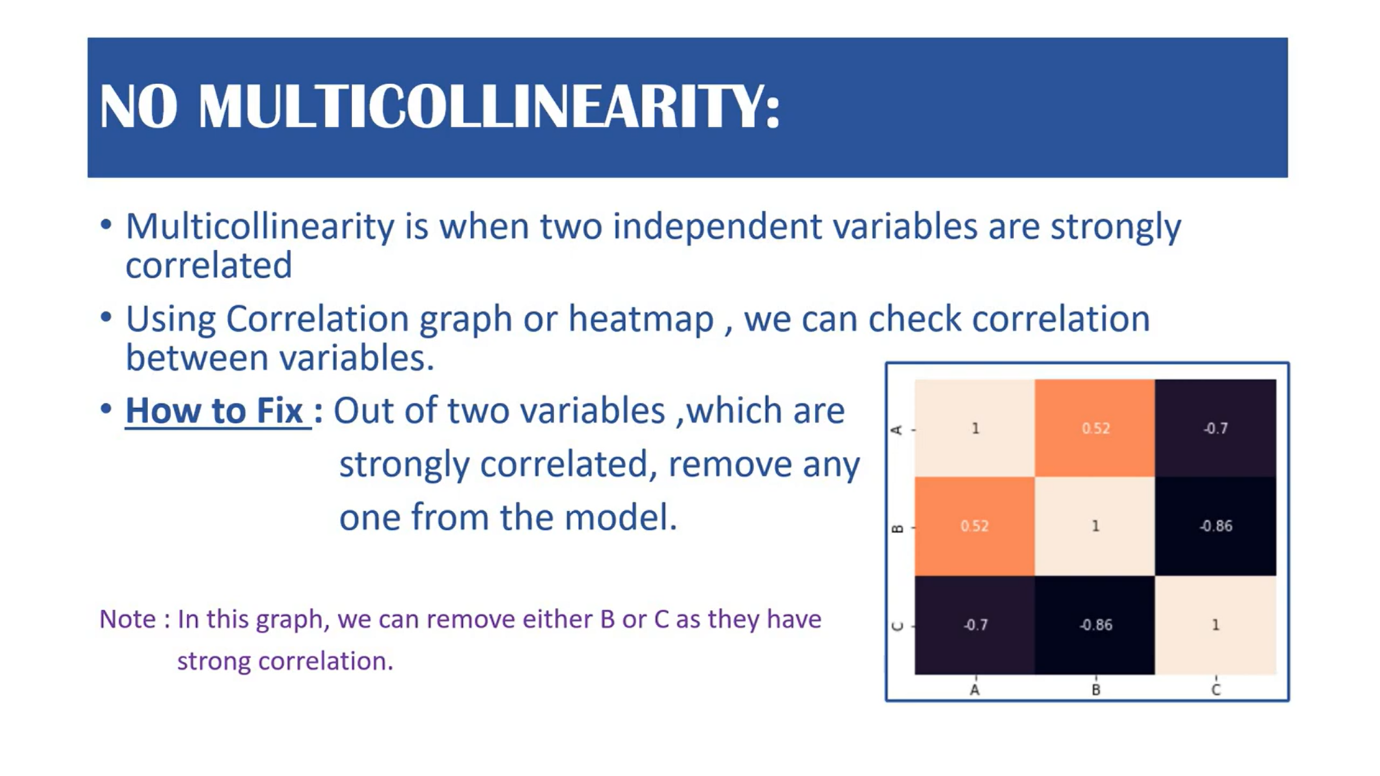
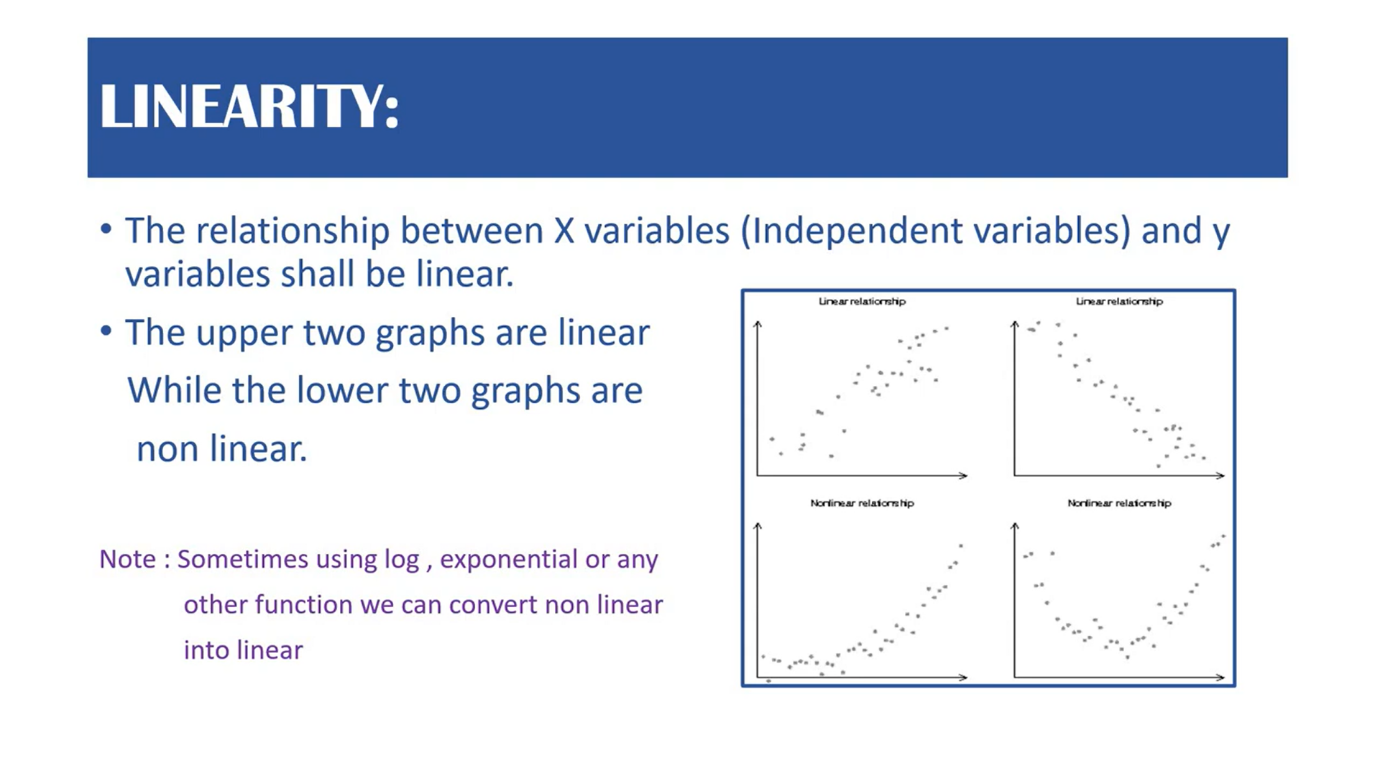
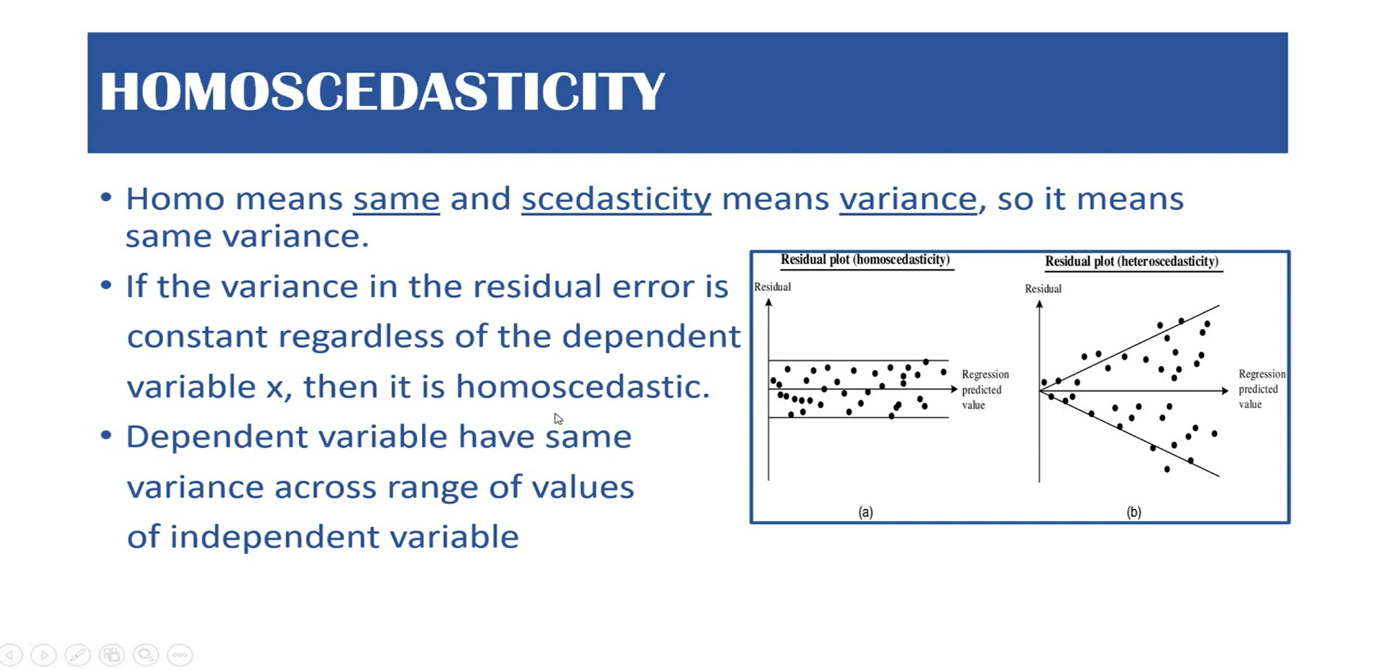
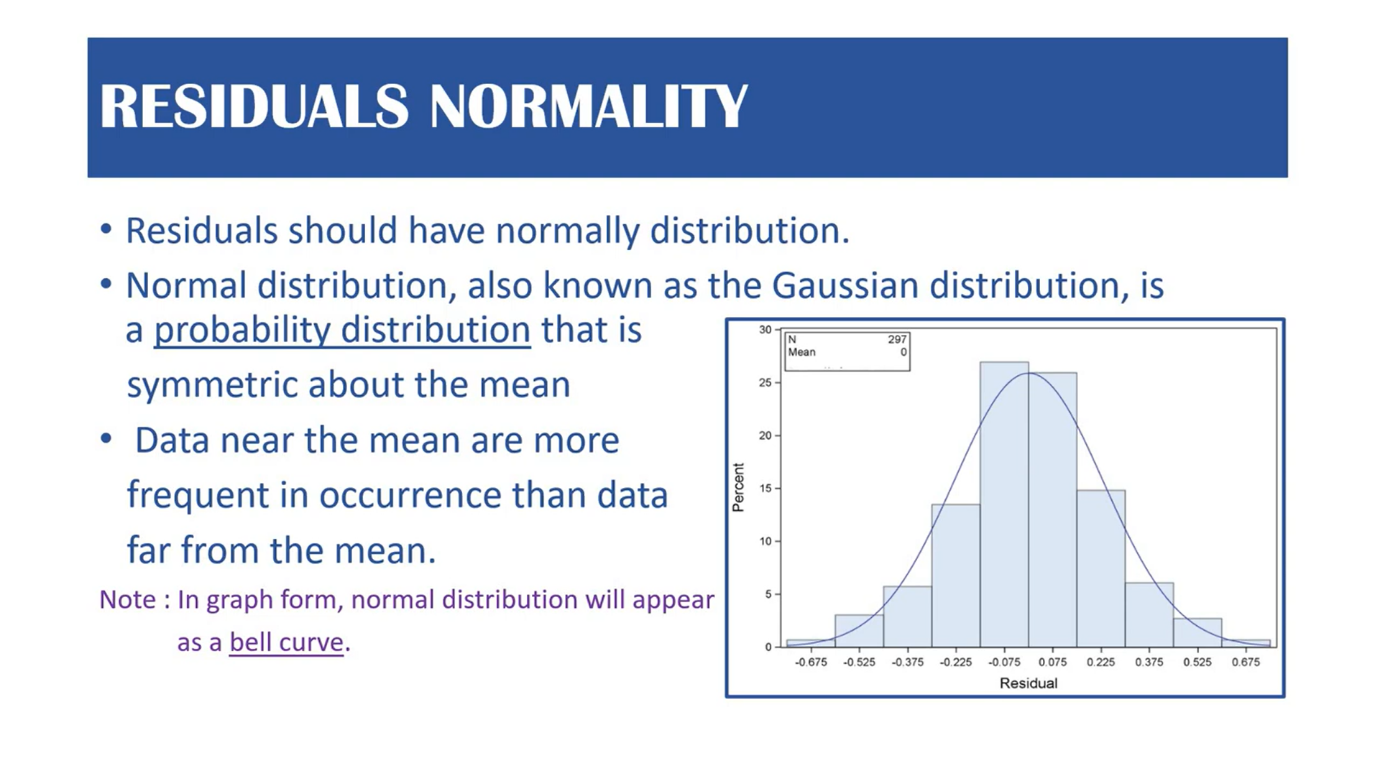
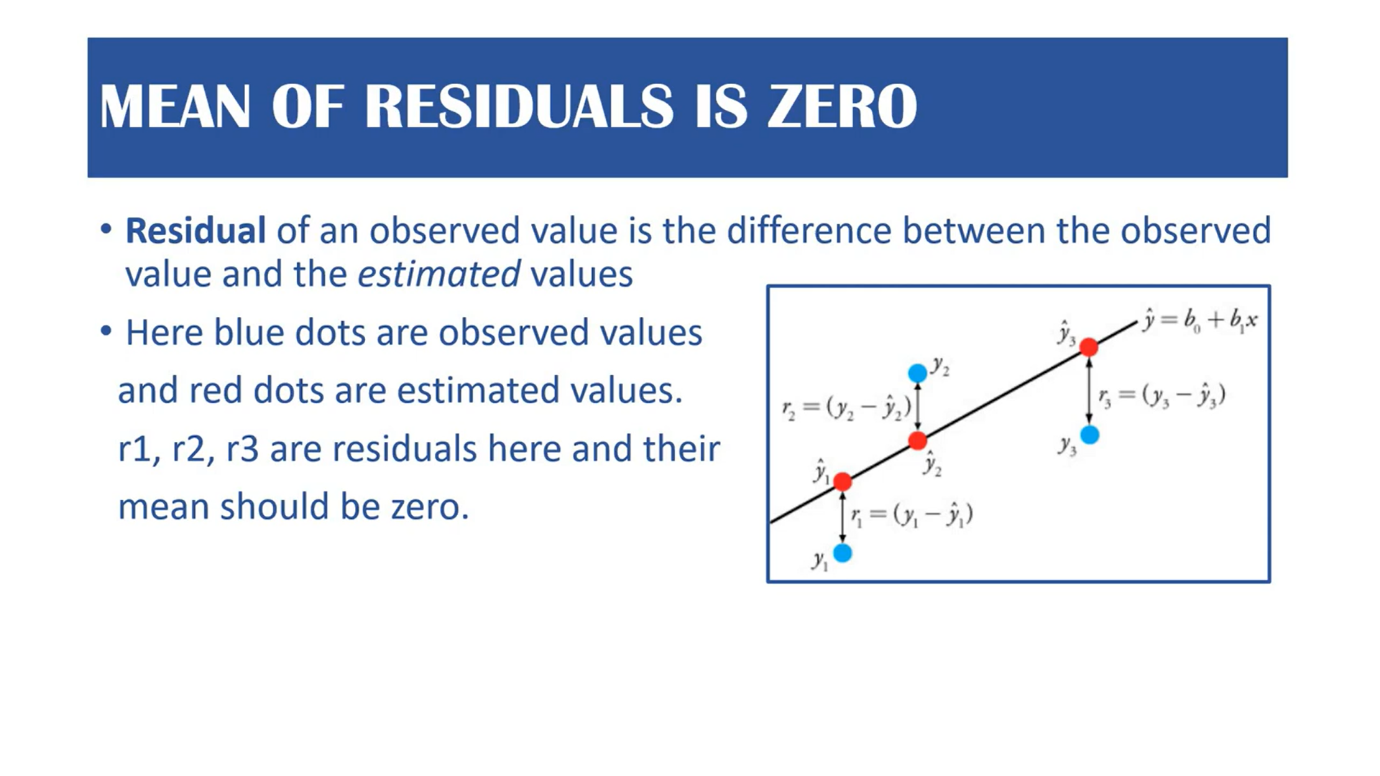
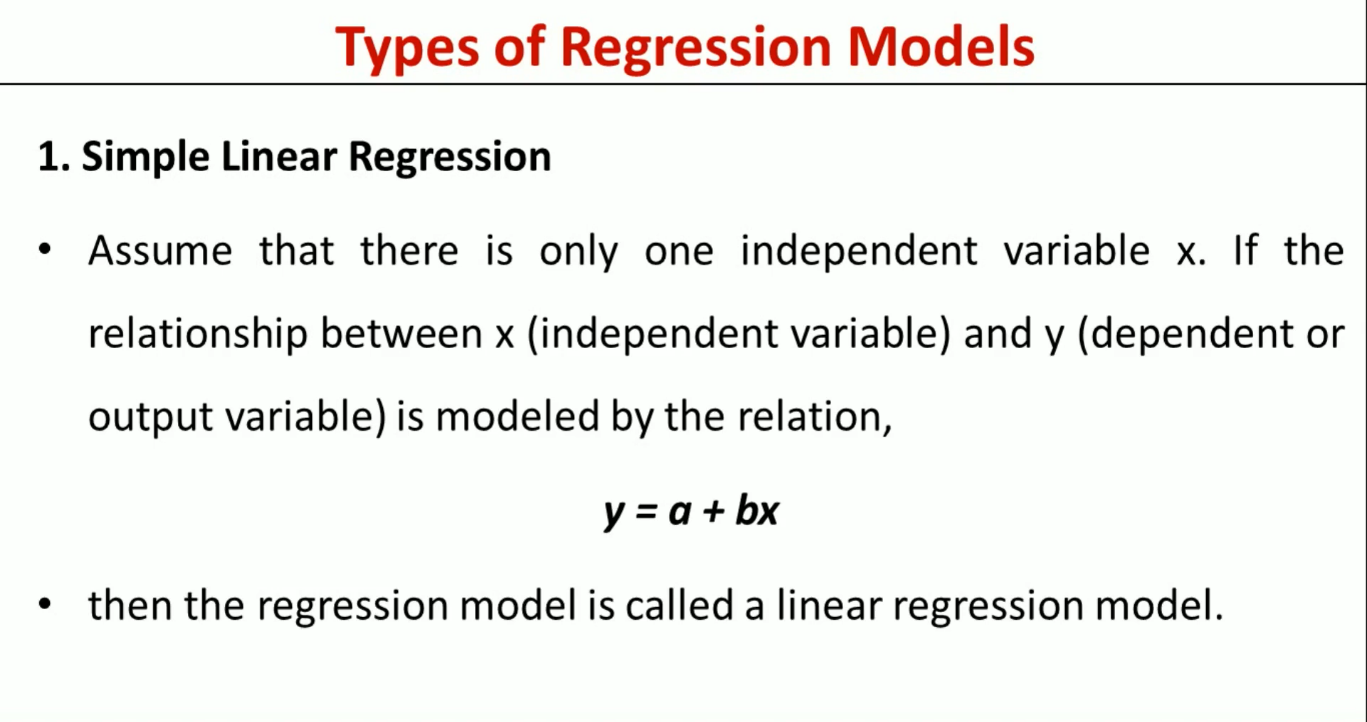
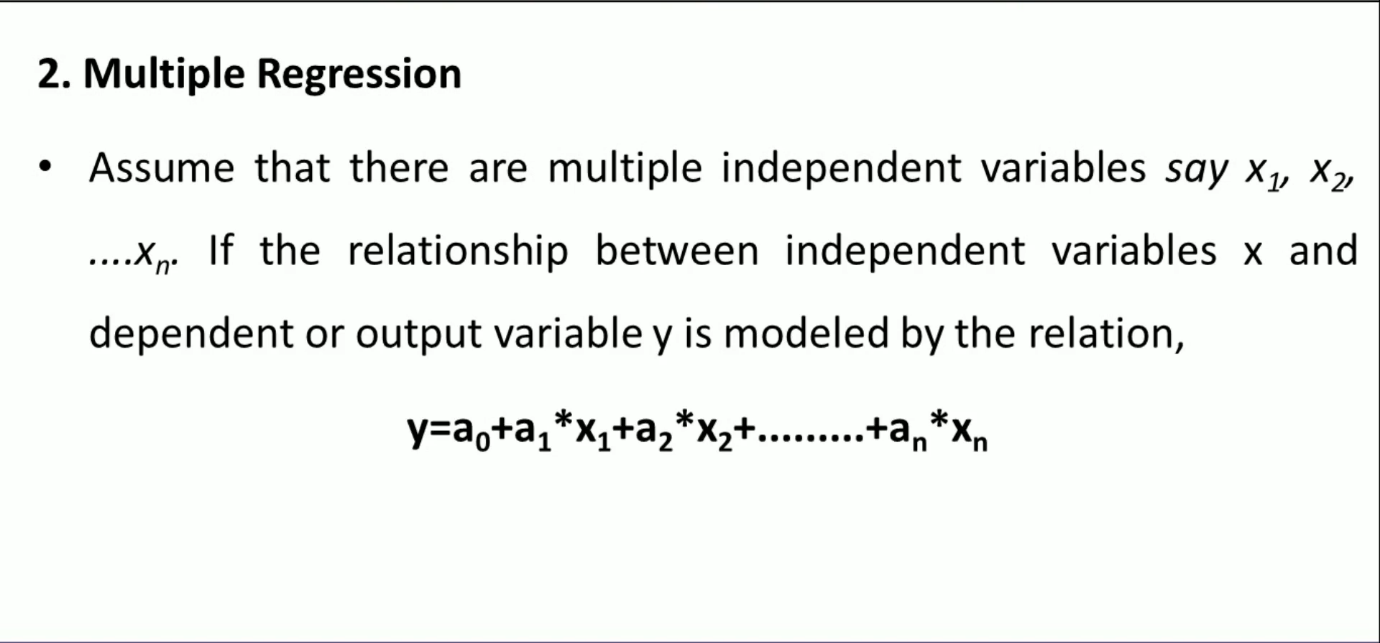
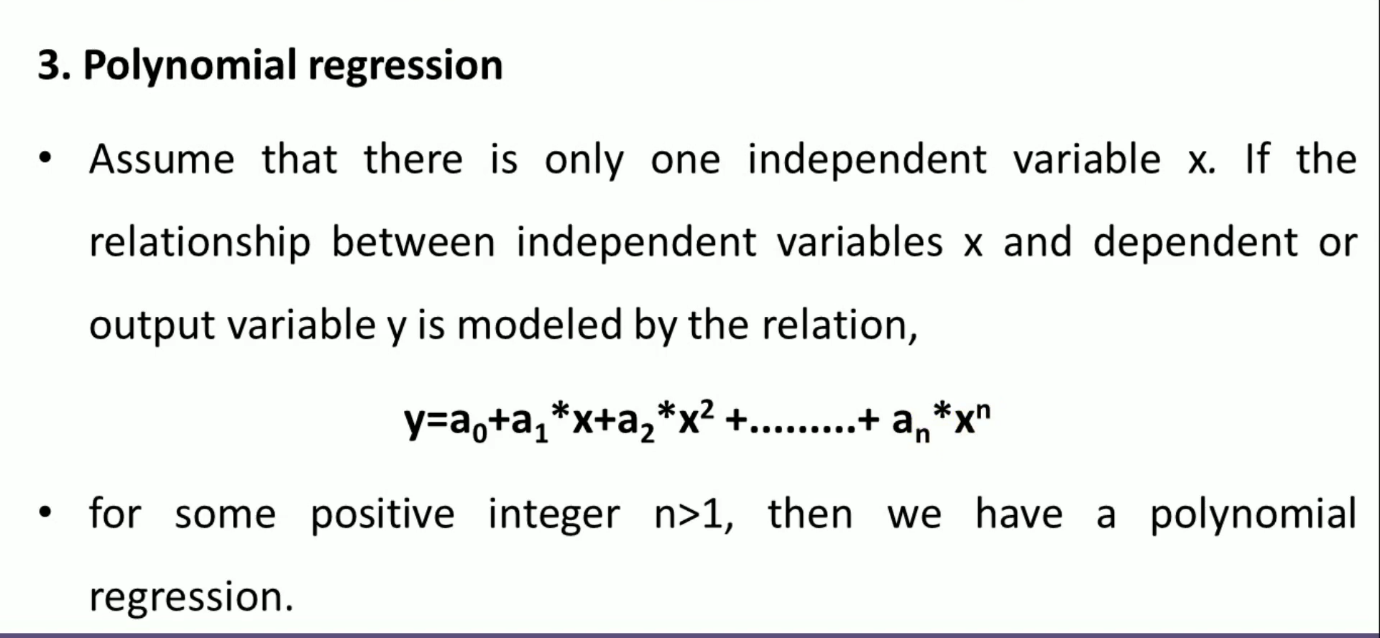
LINEAR REGRESSION

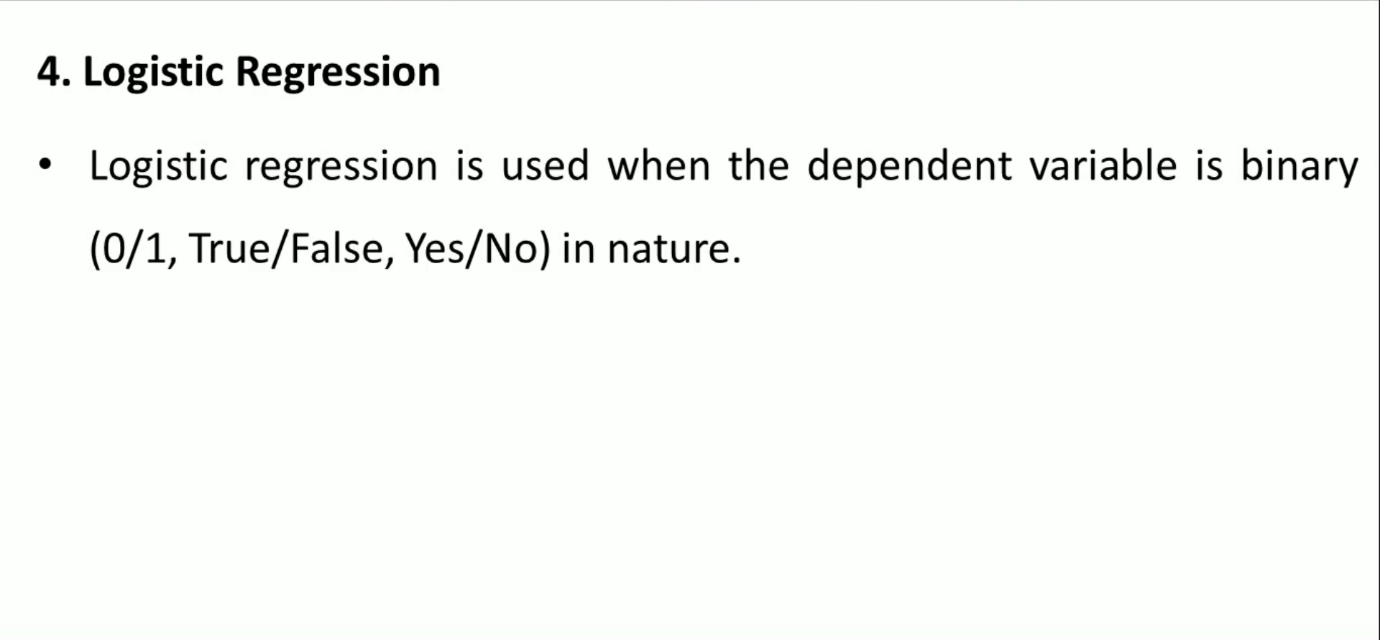
Assumption for linear regression:-

*  strongly correlated means near 1 and -1
* 









**Metrics**

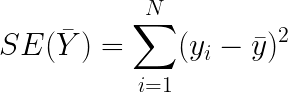
### **R² Coefficient of determination**

R² Coefficient of determination actually works as a post metric, meaning it’s a metric that’s calculated using other metrics.

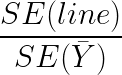
The point of even calculating this coefficient is to answer the question **“How much (what %) of the total variation in Y(target) is explained by the variation in X(regression line)”**

This is calculated using the sum of squared errors. Let’s go through the formulation to understand it better.

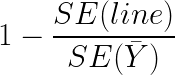
Total variation in Y (Variance of Y):



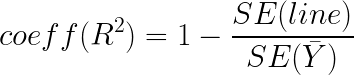
Percentage of variation described the regression line:



Subsequently, the percentage of variation described the regression line:



Finally, we have our formula for the coefficient of determination, which can tell us how good or bad the fit of the regression line is:

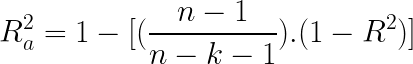


This coefficient can be implemented simply using NumPy arrays in Python.

### **Adjusted R²**

The Vanilla R² method suffers from some demons, like misleading the researcher into believing that the model is improving when the score is increasing but in reality, the learning is not happening. This can happen when a model overfits the data, in that case the variance explained will be 100% but the learning hasn’t happened. To rectify this, R² is adjusted with the number of independent variables.

Adjusted R² is always lower than R², as it adjusts for the increasing predictors and only shows improvement if there is a real improvement.

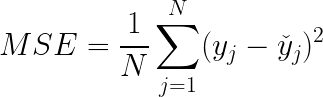


Where:

* n = number of observations
* k = number of independent variables
* Ra² = adjusted R²

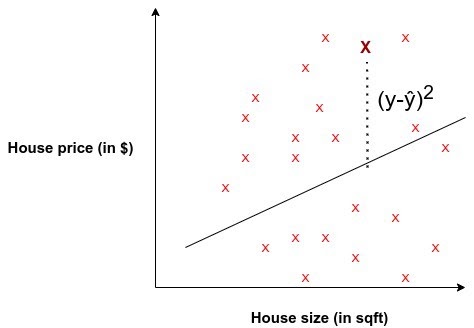
### **Mean Squared Error (MSE)**

Mean squared error is perhaps the most popular metric used for regression problems. It essentially finds the average of the squared difference between the target value and the value predicted by the regression model.



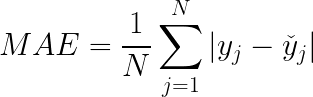
Where:

* y\_j: ground-truth value
* y\_hat: predicted value from the regression model
* N: number of datums



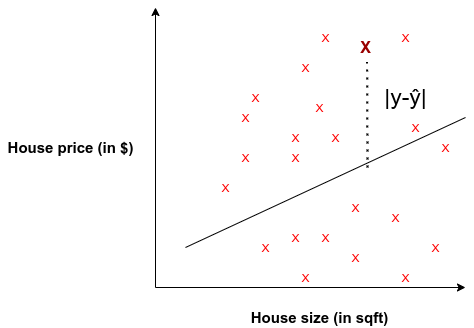
### **Mean Absolute Error (MAE)**

Mean Absolute Error is the average of the difference between the ground truth and the predicted values. Mathematically, its represented as :



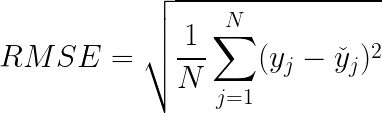
Where:

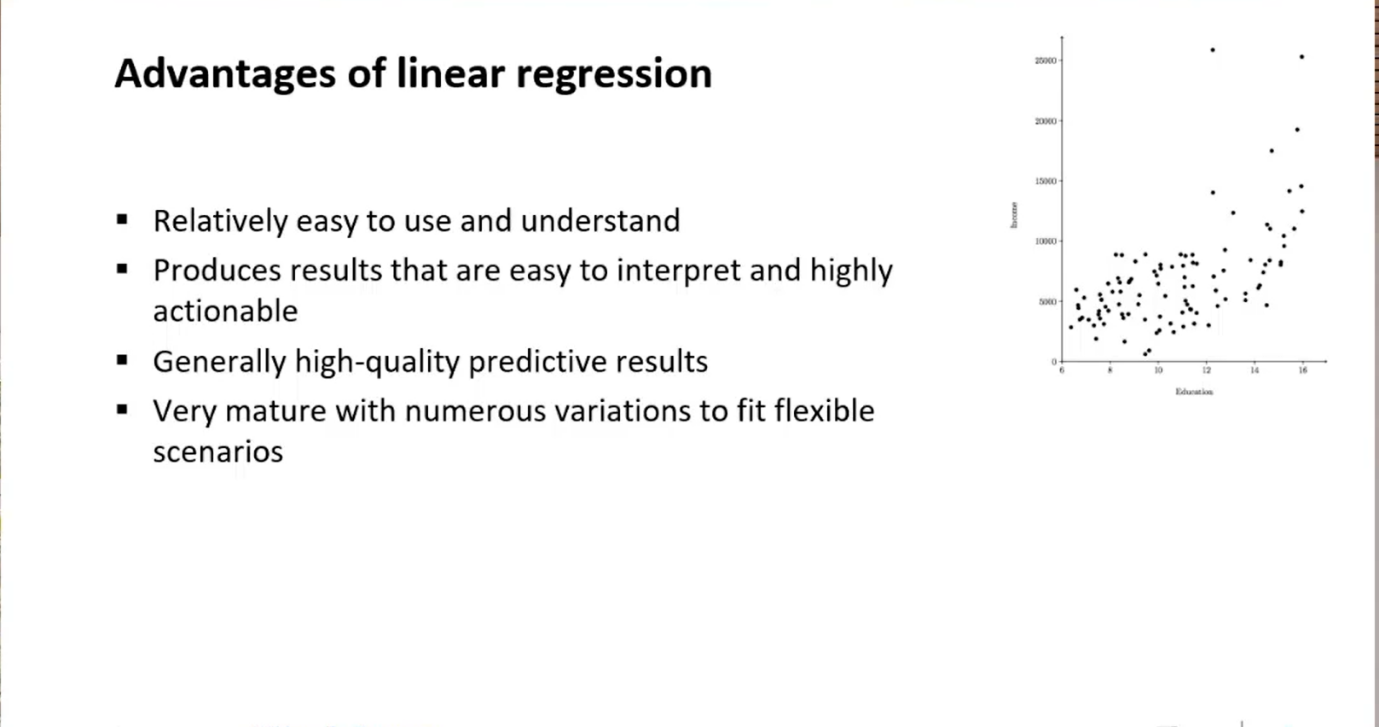
* y\_j: ground-truth value
* y\_hat: predicted value from the regression model
* N: number of datums

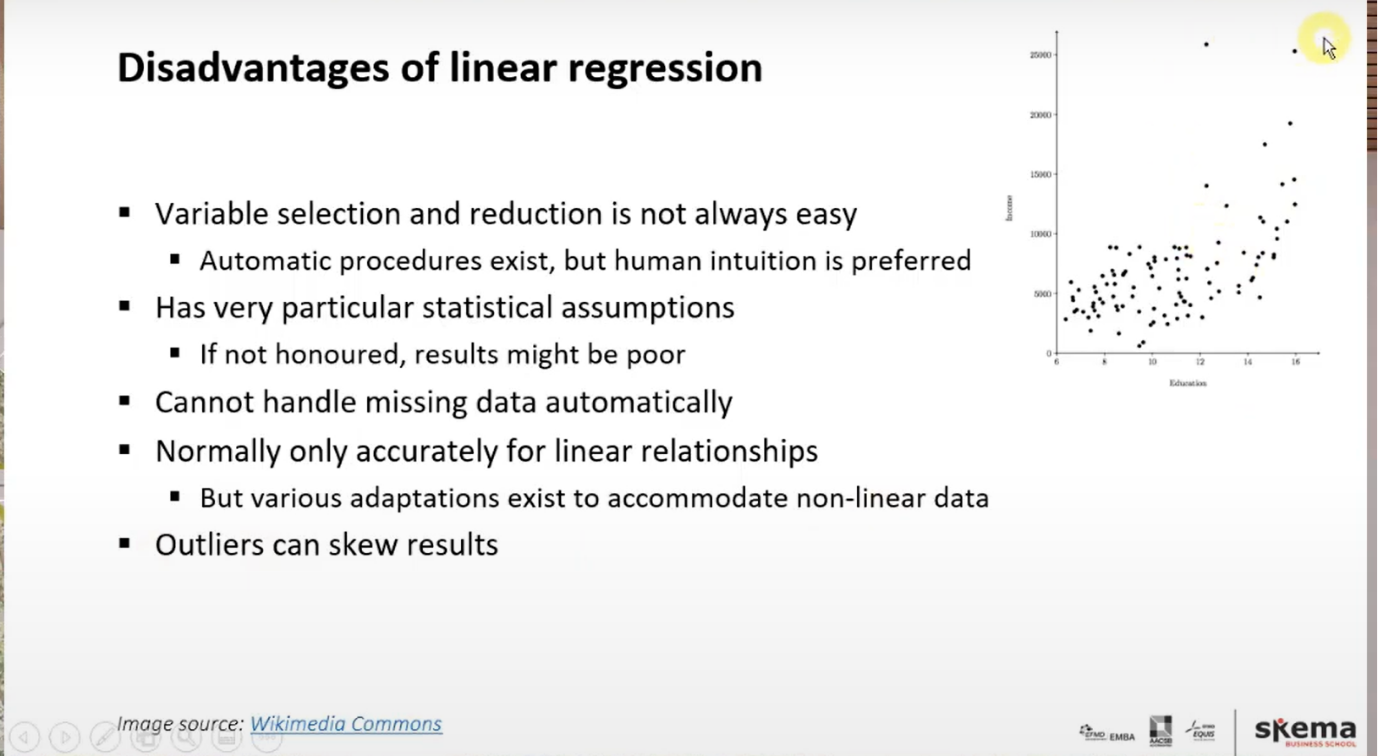


### **Root Mean Squared Error (RMSE)**

Root Mean Squared Error corresponds to the square root of the average of the squared difference between the target value and the value predicted by the regression model. Basically, sqrt(MSE). Mathematically it can be represented as:







### **Gradient Descent:**

* Gradient descent is used to minimize the MSE by calculating the gradient of the cost function.
* A regression model uses gradient descent to update the coefficients of the line by reducing the cost function.
* It is done by a random selection of values of coefficient and then iteratively update the values to reach the minimum cost function.