DAY-5(FEB-16) U.Hasini

22A81A0658

SOC\_2

COURSE:ADVANCED PYTHON

**NOTES:TOPICS COVERED**

**#Linear Regression :**

* + - [Linear regression](https://www.geeksforgeeks.org/ml-linear-regression/) is a statistical method that is used to predict a continuous dependent variable(target variable) based on one or more independent variables(predictor variables).
    - This technique assumes a linear relationship between the dependent and independent variables, which implies that the dependent variable changes proportionally with changes in the independent variables.
    - In other words, linear regression is used to determine the extent to which one or more variables can predict the value of the dependent variable.

**#Types of Linear Regression:**

There are two main types of linear regression:

* **#Simple linear regression:** This involves predicting a dependent variable based on a single independent variable.
* **#Multiple linear regression:** This involves predicting a dependent variable based on multiple independent variables.

#**Linear Regression formula**--->y=a0+a1\*x

->a1-->(meanof(x\*y))-(meanof(x)\*meanof(y)))/meanof(x^2)- (meanof(x)^2)

->a0=mean(y)-a1\*meanof(x)

#To check duplicate values.

#Based on index value try to check the performance.

#Min value and Max value.

#Studied hrs-how many hours count.

#To get all the unique values.

#To know how many students studies in each hour.

#Instead of linear regression-#Ridge.

**#Ridge:**

* + Ridge regression is a statistical regularization technique. It corrects for overfitting on training data in machine learning models.
  + Ridge regression—also known as L2 regularization—is one of several types of regularization for [linear regression](https://www.ibm.com/topics/linear-regression) models.

**#Ml-Unsupervised Learning:**

* Unsupervised learning in artificial intelligence is a type of machine learning that learns from data without human supervision.
* Unlike supervised learning, unsupervised machine learning models are given unlabeled data and allowed to discover patterns and insights without any explicit guidance or instruction.

**#K means clustering:**

* K-Means Clustering is an Unsupervised Machine Learning algorithm, which groups the unlabeled dataset into different clusters

**#countplot:**

**->seaborn.countplot()** method is used to Show the counts of observations in each categorical bin using bars.

***->Syntax :****seaborn.countplot(x=None, y=None, hue=None, data=None, order=None, hue\_order=None, orient=None, color=None, palette=None, saturation=0.75, dodge=True, ax=None, \*\*kwargs)*

**#violinplot:**

* + - A violin plot plays a similar activity that is pursued through whisker or box plot do.
    - As it shows several quantitative data across one or more variables.
    - It can be an effective and attractive way to show multiple data at several units.
    - A “wide-form” Data Frame helps to maintain each numeric column which can be plotted on the graph.
    - It is possible to use NumPy or Python objects, but pandas objects are preferable because the associated names will be used to annotate the axes.

**Syntax:***seaborn.violinplot(x=None, y=None, hue=None, data=None, order=None, hue\_order=None, bw=’scott’, cut=2, scale=’area’, scale\_hue=True, gridsize=100, width=0.8, inner=’box’, split=False, dodge=True, orient=None, linewidth=None, color=None, palette=None, saturation=0.75, ax=None, \*\*kwargs)***Parameters:*****x, y, hue:****Inputs for plotting long-form data.****data:****Dataset for plotting.****scale:****The method used to scale the width of each violin.*

**Returns:** This method returns the Axes object with the plot drawn onto it.