

5 c. Use the diabetes data set from UCI and Pima Indians Diabetes data set for performing the following: Multiple Regression

Aim

Multiple regression is like linear regression, but with more than one independent value, meaning that we try to predict a value based on **two or more** variables.

Procedure

The Pandas module allows us to read csv files and return a DataFrame object.

Then make a list of the independent values and call this variable X.

Put the dependent values in a variable called y.

From the sklearn module we will use the LinearRegression() method to create a linear regression object.

This object has a method called fit() that takes the independent and dependent values as parameters and fills the regression object with data that describes the relationship.

We have a regression object that are ready to predict age values based on a person Glucose and BloodPressure

Program

```
import pandas as pd
from sklearn import linear_model
df = pd.read_csv (r'd:\\diabetes.csv')
print (df)
X = df[['Glucose', 'BloodPressure']]
y = df['Age']
regr = linear_model.LinearRegression()
regr.fit(X, y)
predictedage = regr.predict([[150, 13]])
print(predictedage)
```

Output

```
[28.77214401]
```