|  |
| --- |
|  |

**Air Quality**

**Problem Statement:**

In this problem we are going to find the emissions of carbon by using carbon dioxide and temperature.

**REQUIREMENTS:**

The dataset contains few years data in which attributes are:

* Year
* Temperature
* Carbon Dioxide
* Emission of fossil fuel burning(carbon)

**Steps for executing this model:**

**1)** Importing the **libraries** which are required for this model is:

* import matplotlib.pyplot as plt
* import numpy as np
* import pandas as pd
* %matplotlib inline

**2)** Read the excel sheet by using **Pandas** with the below command

* a = pd.read\_excel('CO2.xls')

**3)** Here we are importing **train\_test\_split** class from **sklearn.cross\_validation** package. The scikit-learn version should be 0.19.0, if the scikit-learn version is greater than 0.19.0 then downgrade the version by using the below command

* pip install scikit-learn==0.19.0

**4)** We will split the dataset into four parts by using below command

* X\_train, X\_test, y\_train, y\_test = train\_test\_split (x, y, test\_size=0.2)
* Here the training contains 80% of data and testing contains 20% of data

**5)** We are using **LinearRegression** model which is imported from **sklearn.linear\_model** package, then we create an object for the Linear Regressionclass and train the dataset which we have taken.

**6)** After training we test the model by using x\_test and y\_test and we got an accuracy of **96.92%.**

**7)** We can change the data in x\_test to predict the data in the future and compare that value with our own data present in the y\_test.