Pandas-Introduction

Pandas is an open-source library in Python that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the NumPy library of Python. Pandas is fast and it has high performance & productivity for users.

Getting Started

Installing Pandas

The first step of working in pandas is to ensure whether it is installed in the system or not. If not then we need to install it in our system using the pip command. Type the cmd command in the search box and locate the folder using the cd command where python-pip file has been installed. After locating it, type the command:

Pip install pandas

Importing Pandas

After the pandas have been installed into the system, you need to import the library. This module is generally imported as follows:

Import pandas as pd

Here, pd is referred to as an alias to the Pandas. However, it is not necessary to import the library using the alias, it just helps in writing less amount code every time a method or property is called Pandas Data Structures

Pandas generally provide two data structures for manipulating data, They are:

Series

DataFrame

Series

Pandas Series is a one-dimensional labeled array capable of holding data of any type (integer, string, float, python objects, etc.). The axis labels are collectively called indexes.

Pandas Series is nothing but a column in an Excel sheet. Labels need not be unique but must be a hashable type. The object supports both integer and label-based indexing and provides a host of methods for performing operations involving the index.

Conclusion

The generated graph images through ibm platform made us to understand about the air quality analysis through the tamilnadu state and it greatly helped for our easy prediction about air quality analysis and it's bad effects towards people in highly polluted area.