

Data Wrangling and Visualization using Python

Mohit Kumar

M.Sc. (Comp. Applications), Roll No: 20961

ICAR-IASRI, Library Avenue, New Delhi-110012

Abstract

Python is a general-purpose programming language that is becoming more and more popular for doing data analysis. The reason for this popularity is due to the available libraries that are growing constantly. Python's pandas library is used for data wrangling as it includes functions for loading, cleaning, aggregation, merging, joining etc.. For data visualization, matplotlib is low-level library. However seaborn is high level library that is built on the top of matplotlib that is used for visualization of data. Both matplotlib and seaborn provide static view of data. When there is need to make plots interactive and web application is required to share the plots across world-wide web, python's dash framework is suitable. By using dash, interactive web application can be built, without having knowledge of html and javascript. These application can be published to server or on cloud.

Key-words: *dash, matplotlib, pandas, python, seaborn.*