

JB INSTITUTE OF ENGINEERING AND TECHNOLOGY

Department of AI&D, CSE (DS)

Value-Added Course on Python for Data Science - Report

Date: 25.09.2025

Module Number: 02

Key Topics Covered: Pandas Library Operations on CSV Files

Resource Person(s): Mr. Sagar NV, Dr. Sudheer Reddy Bandi

Duration: 2h

Mode of Delivery: Face to Face

Target Audience: AI&DS, CSD Students

Number of Participants: 65

Venue: First floor, Lab 105

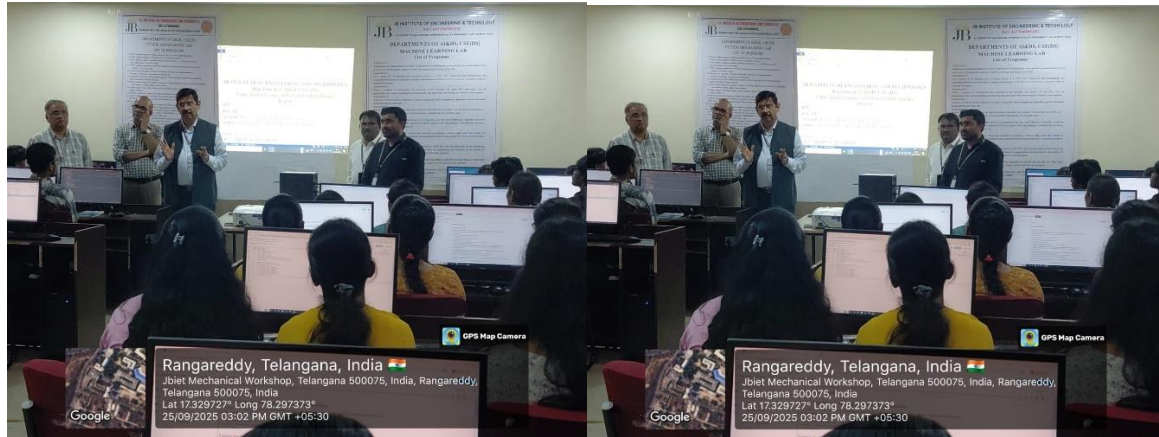


Image: Principal, Dean CD, Dean CS addressing the gathering.

Objectives of the Course Module:

- To enable students to work with real-world data stored in CSV format using Pandas.
- To teach students how to read, write, and export CSV files using Pandas functions.
- To provide hands-on experience with data exploration techniques including viewing, describing, and summarizing datasets.
- To demonstrate advanced Pandas operations such as filtering, sorting, grouping, and aggregation on CSV data.
- To develop skills in data cleaning and preprocessing including handling missing values, duplicates, and data type conversions.

Session Content:

The second module built upon the foundational knowledge from Module 01, focusing specifically on practical applications of the Pandas library for working with CSV files. The session began with importing CSV files using `pd.read_csv()` and exploring various parameters for handling different file formats and encoding issues.

Students learned essential data exploration techniques including `head()`, `tail()`, `info()`, `describe()`, and `shape` to understand dataset structure and characteristics. The instructors demonstrated data selection methods using `loc[]` and `iloc[]` for accessing specific rows and columns.

Advanced operations covered included:

- Filtering data based on single and multiple conditions
- Sorting datasets using `sort_values()`
- Grouping operations with `groupby()` for aggregated analysis

- Handling missing data using `isnull()`, `fillna()`, and `dropna()`
- Removing duplicate entries and resetting indices
- Creating new columns through data transformation
- Merging and concatenating multiple CSV files

Practical exercises involved working with real datasets where students performed data cleaning, exploratory analysis, and generated summary statistics. The hands-on approach ensured that participants could apply these operations to solve practical data analysis problems.

Expected Learning Outcomes of the Module:

By the end of the module, students will be able to:

- Successfully import and export data from CSV files using appropriate Pandas functions.
- Perform comprehensive data exploration and generate descriptive statistics from datasets.
- Apply filtering and conditional selection to extract relevant subsets of data.
- Execute sorting and ranking operations on single and multiple columns.
- Utilize groupby operations to perform aggregated analysis across different categories.
- Identify and handle missing values, duplicates, and inconsistent data effectively.
- Transform and manipulate datasets through column operations and calculations.

Summary:

The meeting was addressed by the Principal Dr P.C. Krishnamachary, Dr I Sudharashan Kumar (Dean College and Development), and Dr Venkata Krishna (Dean Computer Sciences). Further, they elaborated the importance and significance of the course. This module provided an intensive hands-on session focused on practical data manipulation using Pandas with CSV files. Led by Mr. Sagar NV and Dr. Sudheer Reddy Bandi, the 2-hour session equipped 65 students from AI&DS and CSD departments with essential skills for real-world data analysis. The session successfully transitioned students from basic Python and Pandas concepts to practical applications involving actual datasets. Students gained confidence in handling CSV files, performing data cleaning operations, and extracting meaningful insights through various Pandas functions.

Signature of the HoD