



MICROSPECTRA SOFTWARE TECHNOLOGIES PVT. LTD.

**INTRODUCTION TO
DATA SCIENCE ARTIFICIAL INTELLIGENCE
MACHINE LEARNING DEEP LEARNING**

Topic:- Pandas with Python(lect no:-4)

Certification & Internship Program



INTRO TO DATA SCIENCE

WHAT IS PANDAS?

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.





INTRO TO DATA SCIENCE

Getting Started With Pandas

Installation

Q) Working With Conda?

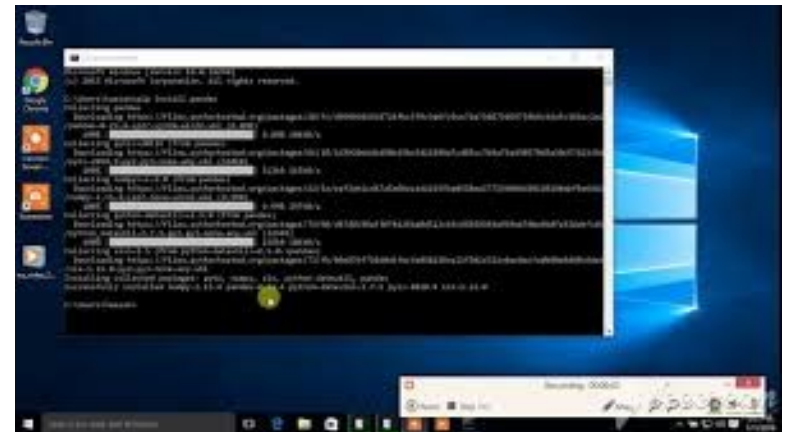
Pandas is part of the Anaconda distribution and can be installed with Anaconda or Miniconda:

conda install pandas

Q) Working with pip?

Pandas can be installed via pip from PyPI.

pip install pandas



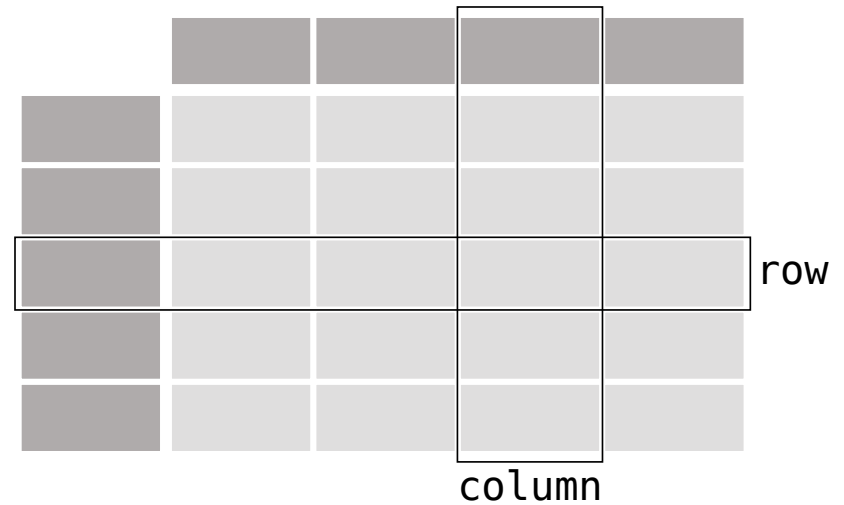


INTRO TO DATA SCIENCE

What kind of data does Pandas handle?

When working with tabular data, such as data stored in spreadsheets or databases, Pandas is the right tool for you. Pandas will help you to explore, clean and process your data. In Pandas, a data table is called a DataFrame.

DataFrame

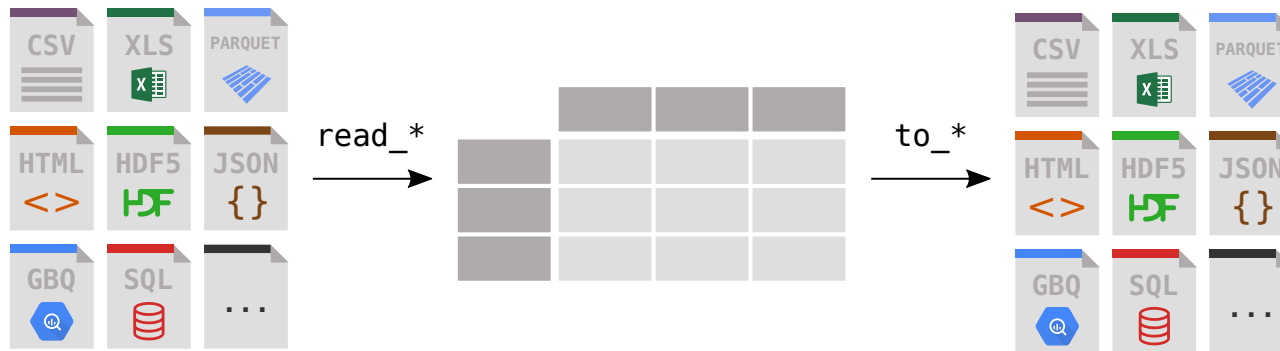




INTRO TO DATA SCIENCE

How do I read and write tabular data?

Pandas supports the integration with many file formats or data sources out of the box (csv, excel, sql, json, parquet,...). Importing data from each of these data sources is provided by function with the prefix `read_*`. Similarly, the `to_*` methods are used to store data.





ESSENTIALS OF DATA SCIENCE

How Do I Select A Subset of a Table?

Selecting or filtering specific rows and/or columns? Filtering the data on a condition? Methods for slicing, selecting, and extracting the data you need are available in Pandas.

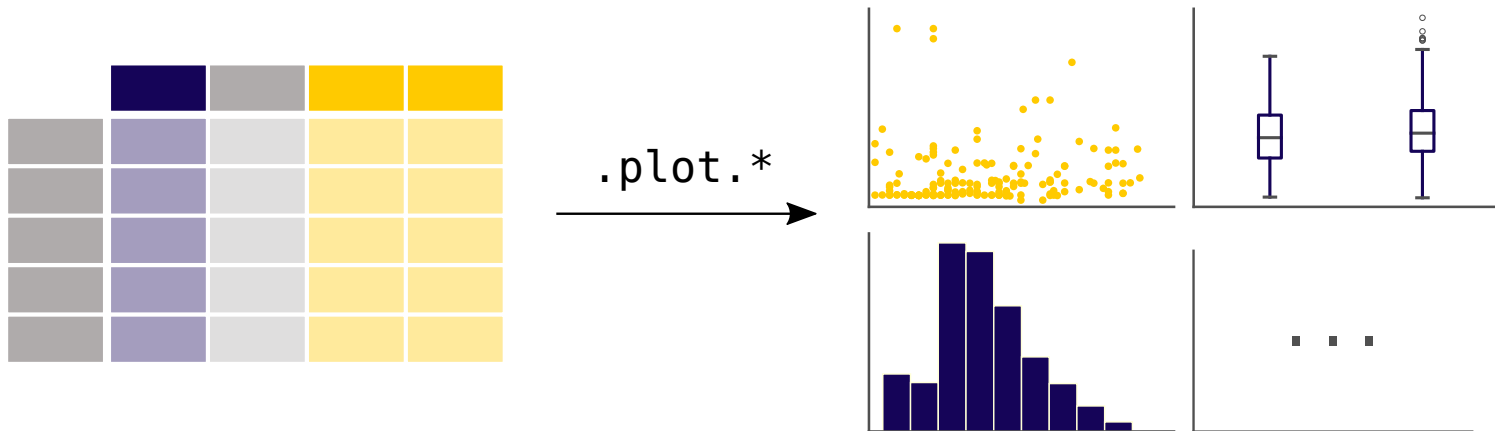




ESSENTIALS OF DATA SCIENCE

How To Create Plots In Pandas?

Pandas provides plotting your data out of the box, using the power of Matplotlib. You can pick the plot type (scatter, bar, boxplot,...) corresponding to your data.

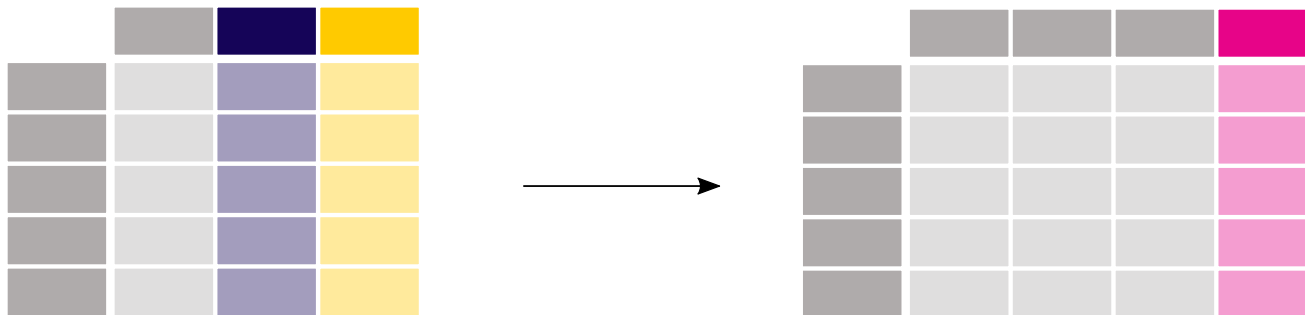




ESSENTIALS OF DATA SCIENCE

How to create new columns derived from existing columns??

There is no need to loop over all rows of your data table to do calculations. Data manipulations on a column work elementwise. Adding a column to a DataFrame based on existing data in other columns is straightforward.

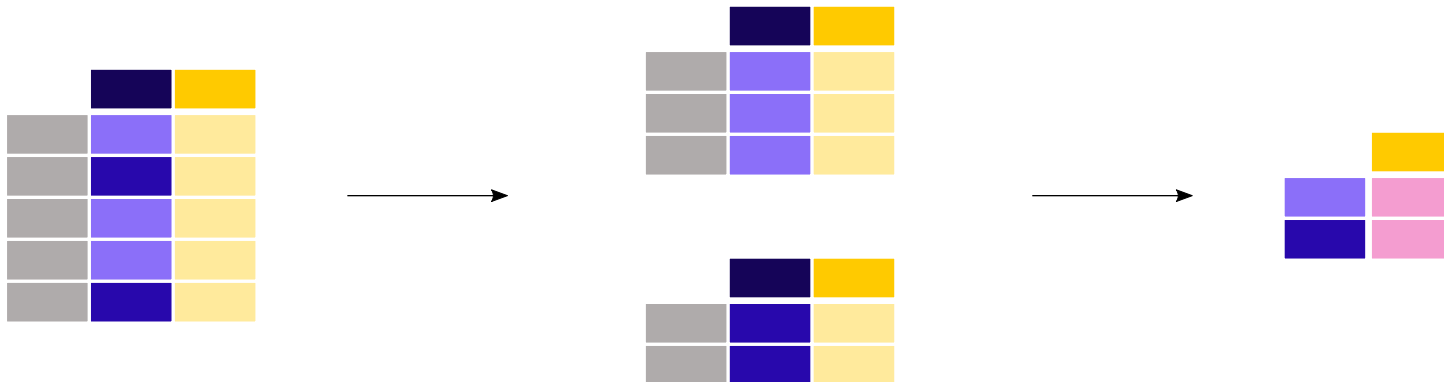




Essentials of Data Science

How to calculate summary statistics?

Basic statistics (mean, median, min, max, counts...) are easily calculable. These or custom aggregations can be applied on the entire data set, a sliding window of the data or grouped by categories. The latter is also known as the split-apply-combine approach.

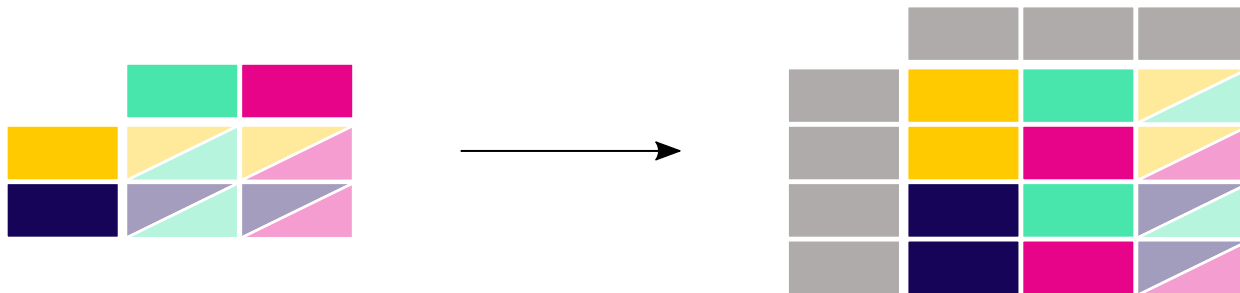




Essentials Of Data Science

How To Reshape The Layout Of Tables?

Change the structure of your data table in multiple ways. You can `melt()` your data table from wide to long/tidy form or `pivot()` from long to wide format. With aggregations built-in, a pivot table is created with a single command.

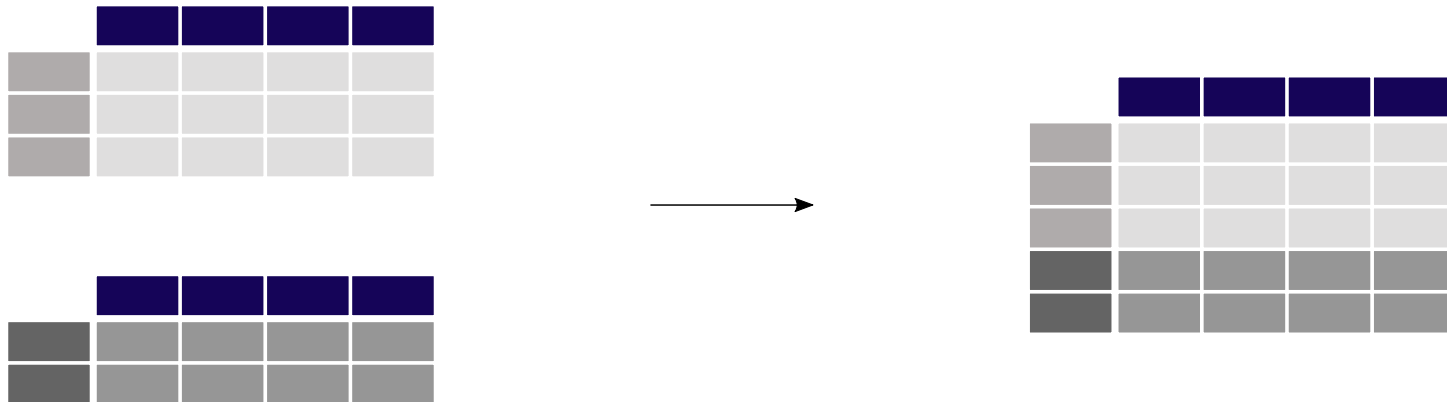




Essentials Of Data Science

How to combine data from multiple tables?

Multiple tables can be concatenated both column wise as row wise and database-like join/merge operations are provided to combine multiple tables of data.

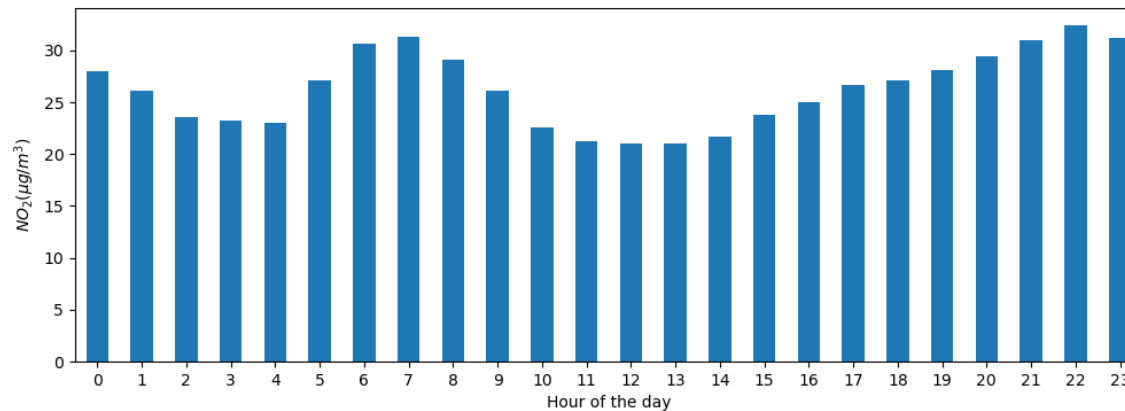




Essentials Of Data Science

How to handle time series data?

Pandas has great support for time series and has an extensive set of tools for working with dates, times, and time-indexed data.

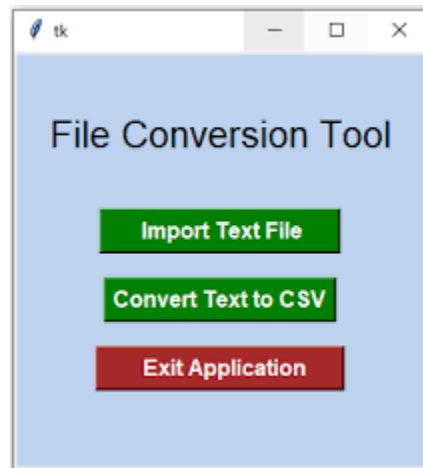




Essentials of Data Science

How to manipulate textual data?

Data sets do not only contain numerical data. Pandas provides a wide range of functions to cleaning textual data and extract useful information from it.





Essentials Of Data Science

What Other Tools You Can Master?

Currently working with other software for data manipulation in a tabular format? You're probably familiar to typical data operations and know what to do with your tabular data, but lacking the syntax to execute these operations. Get to know the pandas syntax by looking for equivalents from the software you already know:



The R programming language provides the data.frame data structure and multiple packages, such as tidyverse use and extend data.frames for convenient data handling functionalities similar to pandas.



Already familiar to SELECT, GROUP BY, JOIN,...? Most of these SQL manipulations do have equivalents in pandas.



The data set included in the STATA statistical software suite corresponds to the pandas data.frame. Many of the operations known from STATA have an equivalent in pandas.



The SAS statistical software suite also provides the data set corresponding to the pandas data.frame. Also vectorized operations, filtering, string processing operations,... from SAS have similar functions in pandas.

DATA VISUALIZATION THE ROLE OF THE COMPUTER

- Much faster at calculating than human
- Great for crunching numbers
- Great for doing many repetitive tasks quick
- Carrying out tasks that we gave it based on logical thinking



DATA VISUALIZATION

THE ROLE OF THE HUMAN

- We've developed to identify patterns
- Creativity
- Bringing in or remembering outside knowledge
- Understanding summary values and Images





DATA VISUALIZATION

PRESENTING DATA

- You can't just show spreadsheets with a bunch of number
- Sometimes even the statistical values aren't good enough summary or they don't really help bring the point across.
- Visualizations allow easy communication and understanding of the data.



DATA VISUALIZATION

INTERPRETING DATA

- Using our understanding of the system and reason
- Using outside knowledge
- Seeing(potential!) trends and patterns
- Considering the context of the data



DATA SCIENCE USING PYTHON

understand the language of data with
Microspectra Software Technologies Pvt. Ltd,
learn Data Science today and apply it tomorrow.



**Industry Authorised
Certification**



Online Training



Online Internship



1 on 1 Interaction



Online Assistant

ENROLL NOW

+91-9730298462

+91-8308848685

www.microspectra.in



GET CERTIFICATION
FROM MICROSOFT

MICROSOFT POWERED INTERNSHIP



SUCCESSFUL PROJECT
SUBMISSION LETTER



ONLINE PROGRAM
1 ON 1 INTERACTION

PLATFORM FOR SUMMER INTERNSHIP PROGRAM



24/7 ASSISTANCE

SELECT TECHNOLOGY

ANDROID DEVELOPMENT

PYTHON

FULL STACK DEVELOPMENT

CLOUD TECHNOLOGY

DATA SCIENCE

BLOCK CHAIN SPECIALIST

INTERNET OF THINGS

AI/ML/DL

DIGITAL MARKETING



+91-9730298462, +91-8308848685



WWW.MICROSPECTRA.IN



SUPPORT @MICROSPECTRA.IN