PANDAS WORKSHOP

LEVERAGING PYTHON TO YOUR NEEDS

Data Science and Informatics
Marielle Doenges

DSI
DATA SCIENCE & INFORMATICS

Github Repository

01

ABOUT US

How you can get involved with the DSI club at UF

02

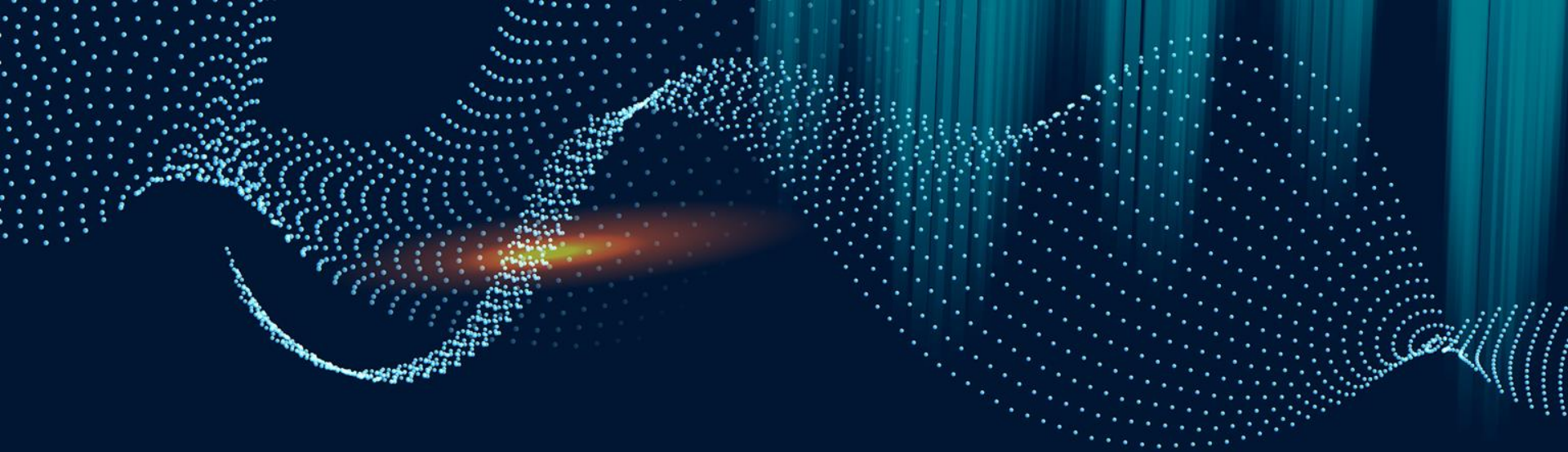
WHY PANDAS

Applications of Pandas in the real world and its origin.

03

PRACTICAL

Interactive tutorial with Python code in Jupyter Notebook



01 ABOUT US

How you can get involved

DSI discord

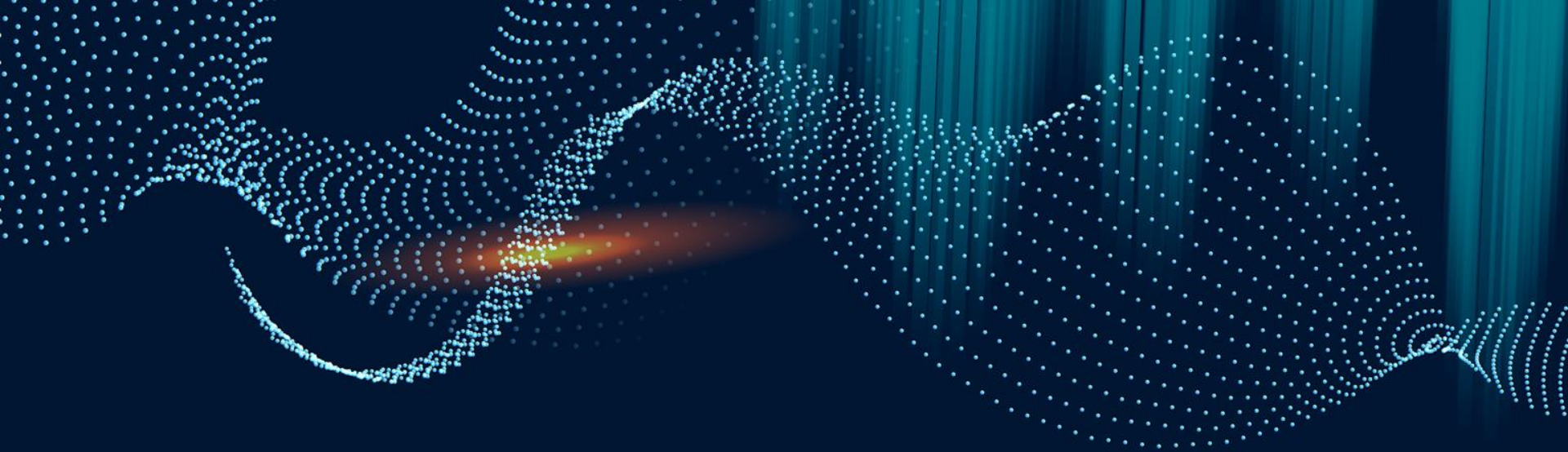
WHY JOIN DSI

Apply Data Science Skills
- Workshops
- Project Opportunities

Career Building
- Resume Building
- Career and Interview Prep
- Guest Speakers from Different Industries

Community
- Interact with Peers
- Socials

DSI
DATA SCIENCE & INFORMATICS



02 WHY PANDAS

Applications and history

BENEFITS OF PANDAS

Rich Functionality:
Built-in functions for common data operations, reducing the need for custom code.
Integrates well with other popular Python libraries (e.g., Matplotlib for plotting).

Versatility:
Handles data of different formats: CSV, Excel, SQL, and many others.
Supports various data operations: filtering, aggregation, transformation.

Efficiency:
Optimized for performance, making it quick even on large datasets.
Uses memory efficiently with its DataFrame structure.

Flexibility:
Can reshape and pivot data with ease.
Supports merging and joining datasets.

Image generated with DALLE3

POPULARITY

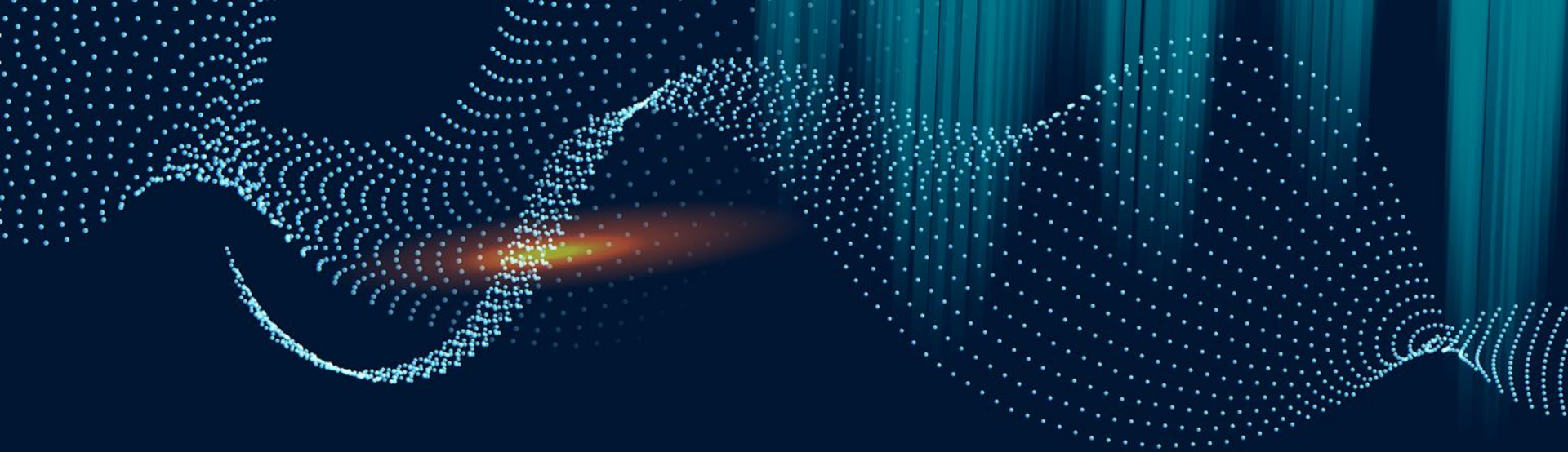
Origin:
"Pandas" stands for "Panel Data" – a term for data that is observed over time for the same subjects.

Growing Popularity:
One of the most preferred and used data analysis tools in Python.
Large community: means better support, more resources, and continuous development.

Wide Adoption:
Used in finance, academia, marketing, …
Companies like Google, JP Morgan, and Dropbox use Pandas.

Open-Source & Python Integration:
Integrates well with other libraries and tools.
Open-source means it's free to use and benefits from community contributions.

Image generated with DALLE3

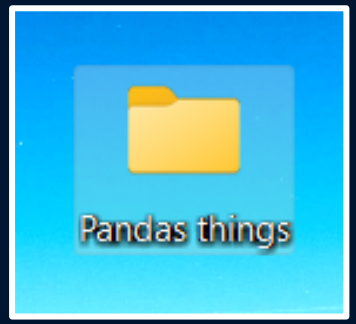


03 TUTORIAL

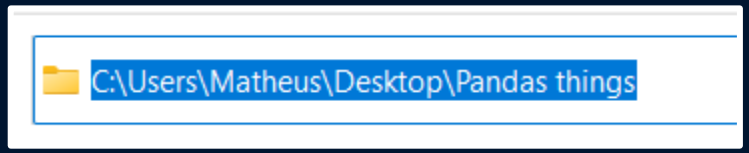
What you came here for!

Create a folder for the Workshop

(It can be anywhere you want within your computer)



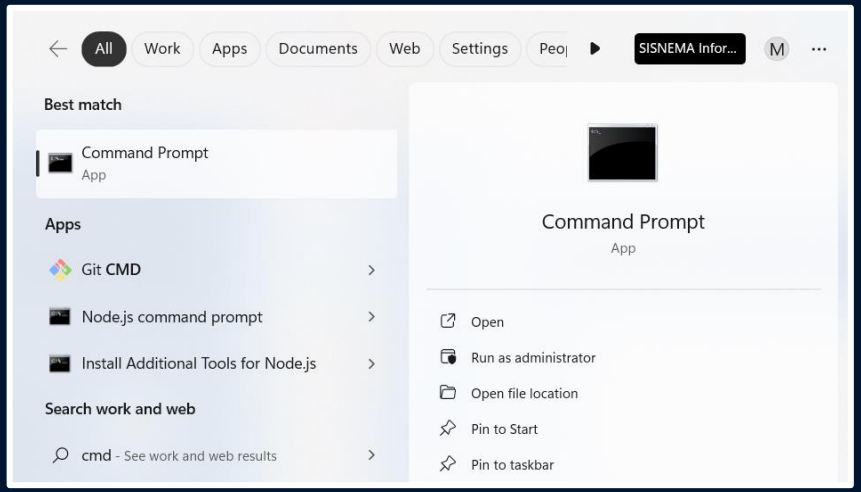
Folder created in desktop



Double click the folder and copy its file path

Launch the Terminal

We promise we won't delete your graphics card

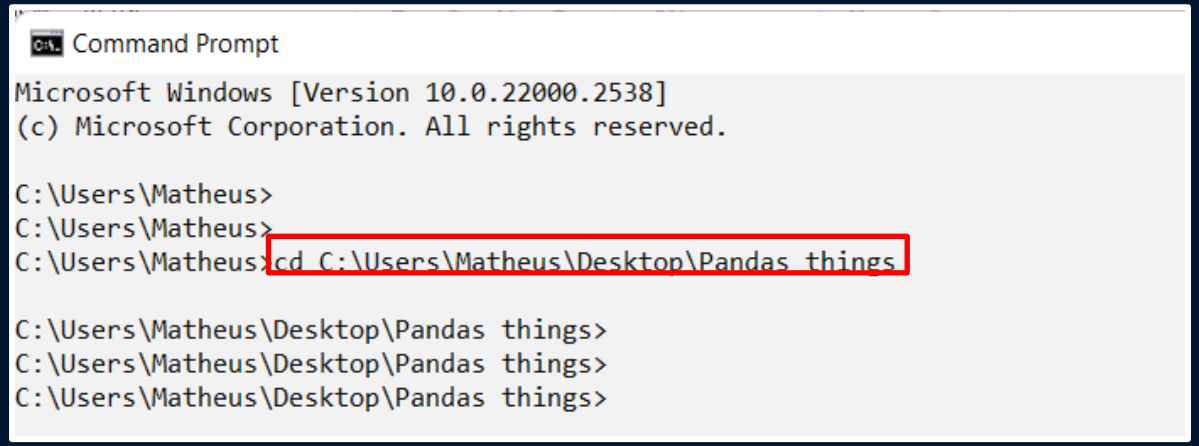


Open your respective OS terminal

How your terminal should look

Run the following commands (1)

This will tell the computer where you want to save the workshop material



“cd <paste the file path you copied before>”

Run the following commands (2)

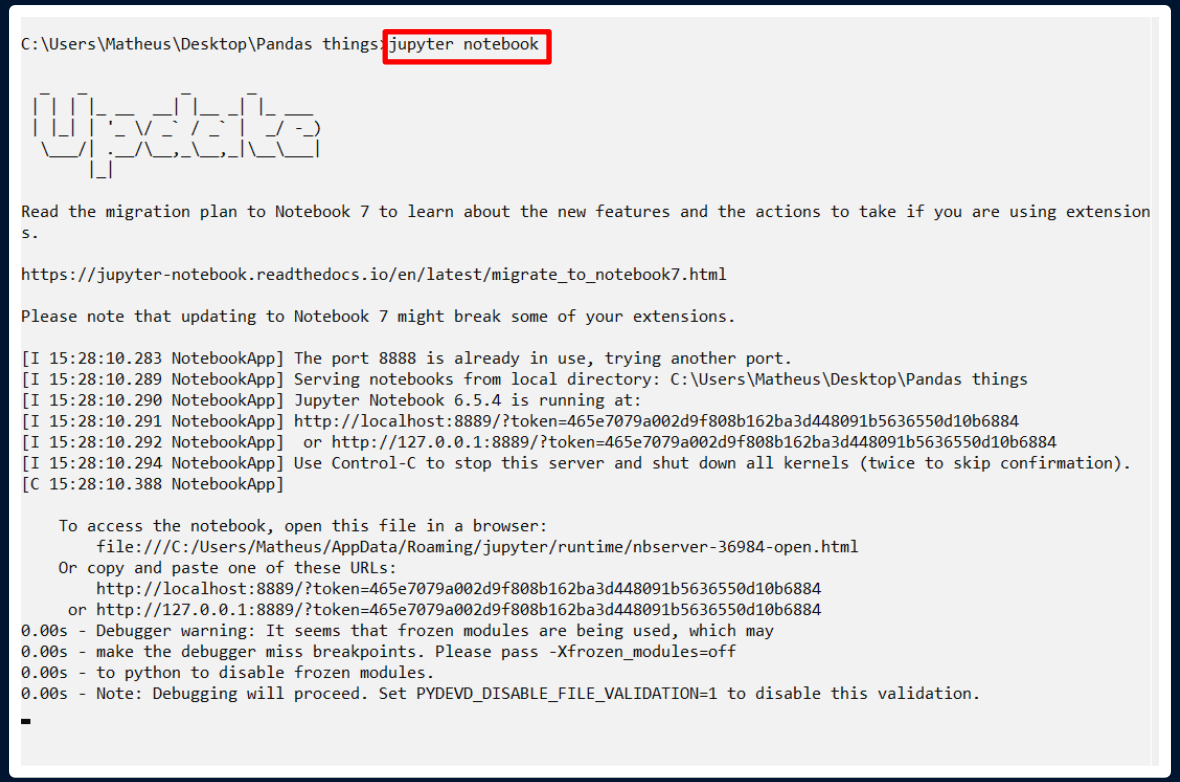
This will download the materials to your newly created folder

```
C:\Users\Matheus\Desktop\Pandas things>git clone https://github.com/matheusmaldaner/WorkshopArchive
Cloning into 'WorkshopArchive'...
remote: Enumerating objects: 83, done.
remote: Counting objects: 100% (34/34), done.
remote: Compressing objects: 100% (19/19), done.
remote: Total 83 (delta 20), reused 14 (delta 14), pack-reused 49
Receiving objects: 100% (83/83), 1.93 MiB | 3.13 MiB/s, done.
Resolving deltas: 100% (28/28), done.
```

“git clone <link to github repository>”

Run the following commands (3)

This will start the Jupyter notebook. Remember to keep the terminal open!



“jupyter notebook”

THANKS!

Do you have any questions?

- marielledoenges@ufl.edu
- @uf_dsi

Save the Repository :)