

Jupyter Notebooks and Papermill

Code Explorers Club Session

15-08-2024

What Are Jupyter Notebooks?

- Interactive environment for documents with text, code, and output.
- Often used in data science, machine learning, and research.
- Supports many programming languages (e.g., Python, R, Julia, etc.)

How Do Jupyter Notebooks Work?

- Code and text are written in cells.
- Each cell can be executed individually.
- A kernel (IPython for Python) runs the code and returns the output directly within the notebook.

```
In [1]: a = 10  
b = 20  
a + b
```

```
Out[1]: 30
```

This is a text cell!

Why Are Jupyter Notebooks Useful?

- Ideal for exploratory data analysis (EDA) and iterative development.
- Allows for combining code, results, visualizations, and explanations all in one place.
- Also supports interactive widgets.
- Helps in documenting workflows for others to follow.

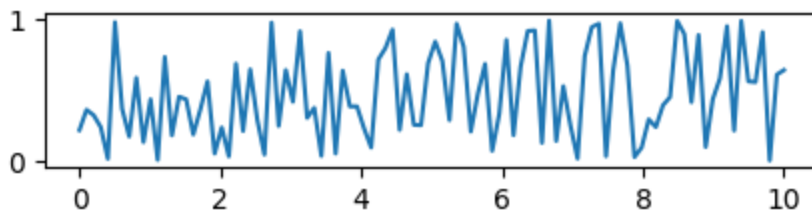
```
In [2]: import numpy as np  
import matplotlib.pyplot as plt
```

Data

```
In [3]: x = np.linspace(0, 10, 100)  
y = np.random.random(100)
```

Plotting data here

```
In [4]: plt.figure(figsize=(5,1))  
plt.plot(x, y);
```



Why Are Jupyter Notebooks Not Used for Making Workflows?

- Difficult to version control.
- Hard to parameterize (introducing parameters that can be adjusted or changed without altering the core code itself).
- Notebooks can become disorganized if you mess with execution order, leading to confusion and/or inconsistent results.