

Practical Setup: Conda, IDE, Jupyter & ML Libraries

1. What You Will Learn Today

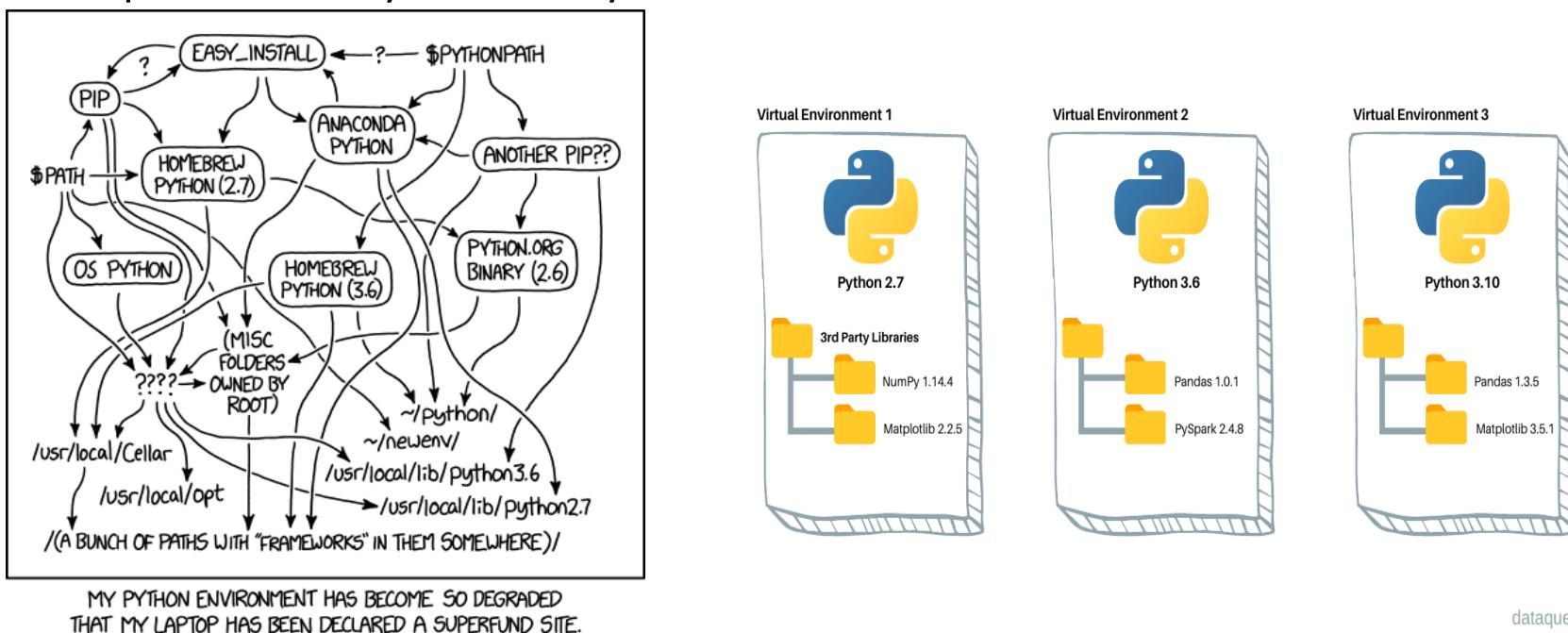
- Install Conda
- Create and activate a virtual environment
- Install Jupyter Notebook
- Install required ML libraries
 - NumPy
 - Matplotlib
 - Scipy
 - Scikit-learn
- Run your notebook

2. Why do we need Conda?

- Different projects require different library versions
- Installing everything globally causes conflicts
- Conda keeps projects separate and safe

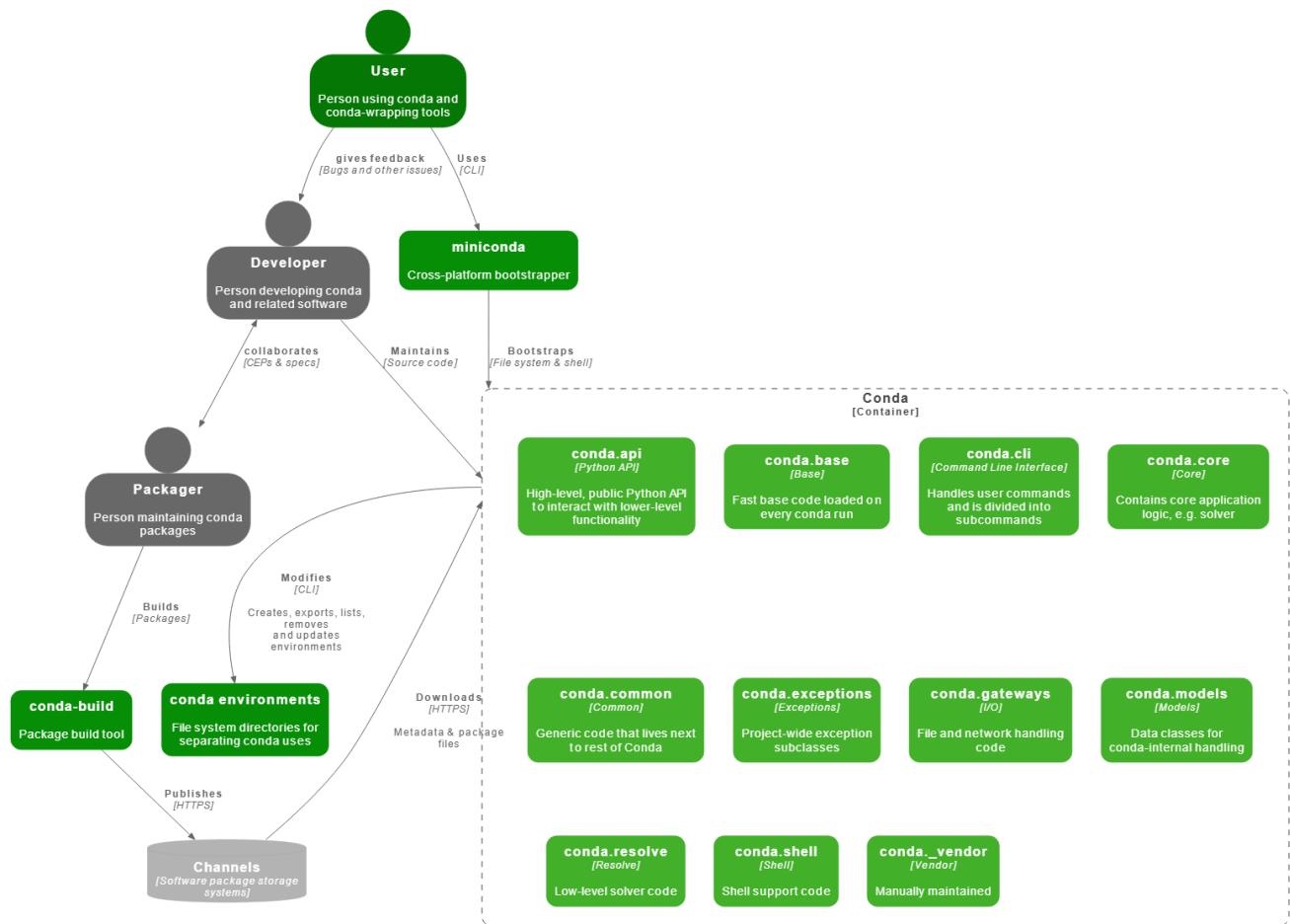
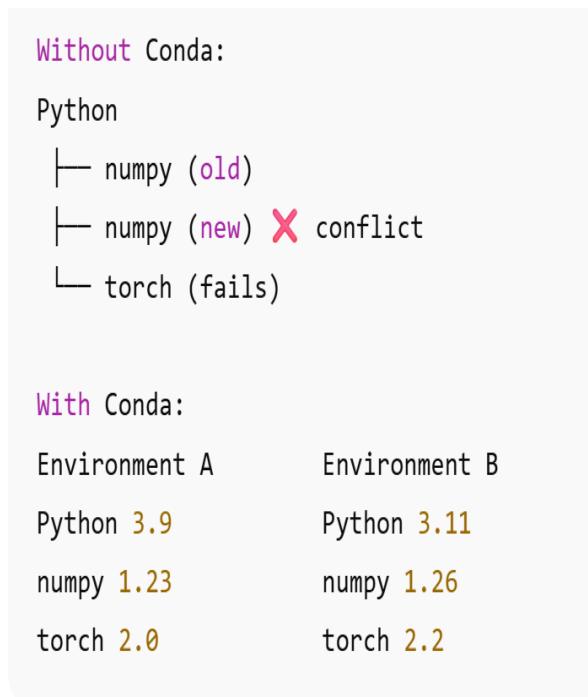
Key idea:

One computer → many isolated Python environments



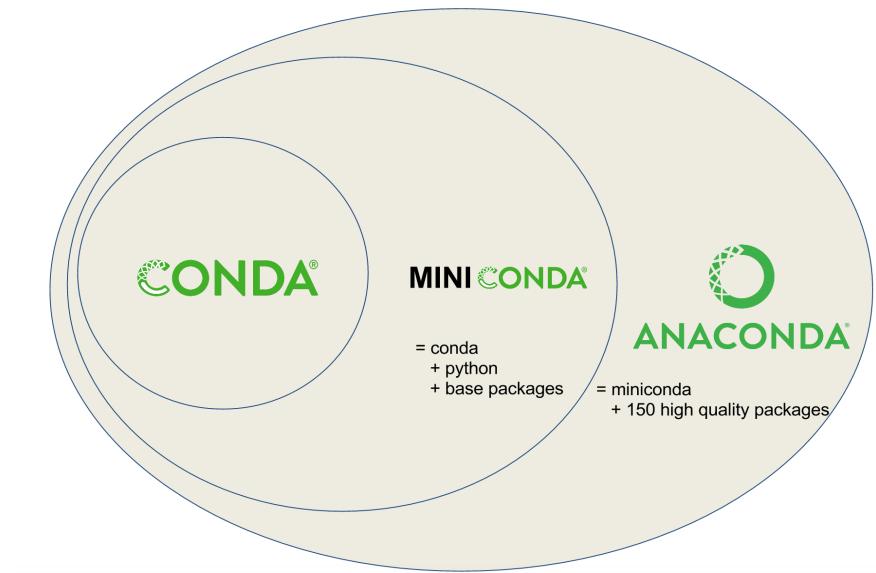
3. Without vs With Conda

Installing Conda, Jupyter Notebook, and Libraries:



4. Install Miniconda (Windows)

- Download Miniconda (Python 3, 64-bit) from:
- <https://docs.conda.io/en/latest/miniconda.html>
- Install and restart terminal
- Add Conda to PATH



- Or Search for “Anaconda Prompt” Run:

```
conda --version
```

5. Install Miniconda (macOS/Linux)

- Download the installer
- Follow standard install steps
- Open Terminal
- Check installation:

```
conda --version
```

6. Create and activate a new environment

Create an environment called **ml**:

- `conda create -n ml python=3.10`

Activate it:

- `conda activate ml`
- You should now see (ml) in the terminal
- `conda deactivate`

7. Integrated Development Environment (IDE)

- IDE Choices:
 - PyCharm:
 - Robust features for development
 - Easy configuration of project interpreters (link to conda env)
 - VS Code:
 - Lightweight and popular
 - Great Python extensions and terminal integration
 - Notebooks (e.g. Jupyter, Colab):
 - Interactive and useful for exploration

Important: final submissions must be .py files

8. Recap and key takeaways

- Use virtual environments to avoid dependency conflicts.
- Miniconda is a lightweight tool for managing these environments.
- pip remains your go-to tool for package installation within environments.
- Configure your IDE (PyCharm or VS Code) correctly for your projects.
- **Always submit a .py file for coursework.**

9. Creating a Conda Environment from environment.yml

What is **environment.yml**?

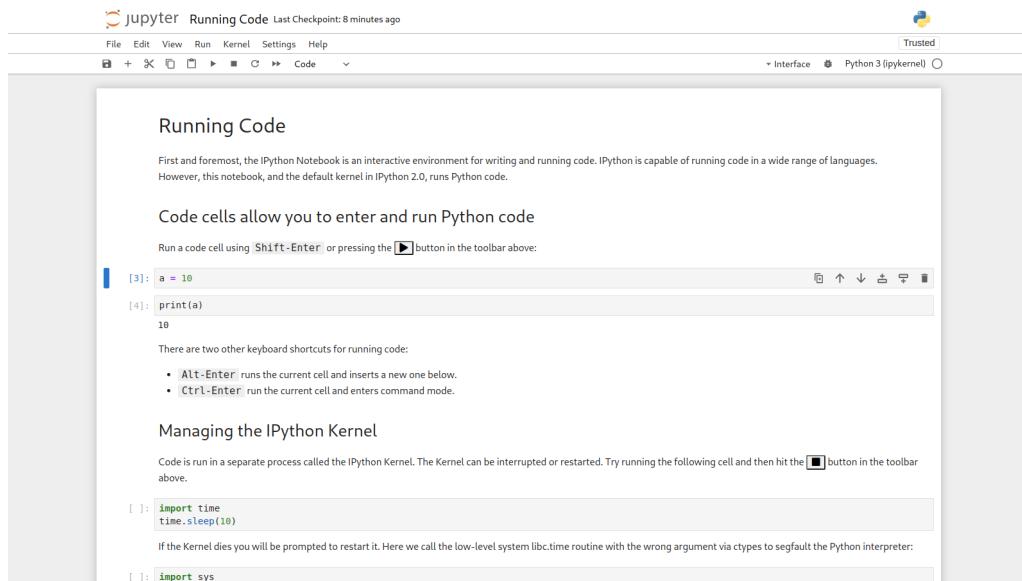
A file that defines a reproducible software environment, including:
environment name, Python version, required packages (Conda + pip)

How to Create the Environment

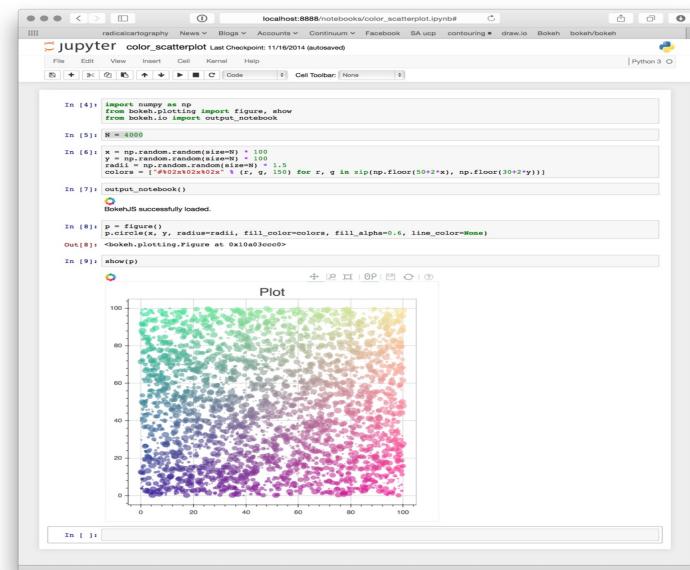
- Open **Anaconda Prompt** and navigate to the folder containing environment.yml `cd` path/to/project
- Create the environment: `conda env create -f environment.yml`
- Activate it: `conda activate coursework_env`

10. What is Jupyter Notebook

- Interactive coding environment
- Code + text + figures in one place



The screenshot shows the IPython Notebook interface. The main window title is "jupyter Running Code Last Checkpoint: 8 minutes ago". The top menu includes File, Edit, View, Run, Kernel, Settings, Help, and a "Trusted" status indicator. The toolbar has buttons for Interface, Python 3 (pykernel), and Cell Toolbar. The content area is titled "Running Code" and contains a section about the IPython Notebook being an interactive environment for writing and running code. It shows a code cell with the command `a = 10` and its output `10`. Below this, it lists keyboard shortcuts: `Alt-Enter` runs the current cell and inserts a new one below, and `Ctrl-Enter` runs the current cell and enters command mode. The "Managing the IPython Kernel" section shows a code cell with `import time` and `time.sleep(10)`, with a note that if the kernel dies, it will prompt to restart. The bottom cell is partially visible with `import sys`.



The screenshot shows the Jupyter Notebook interface. The title bar says "jupyter color_scatterplot Last Checkpoint: 11/16/2014 (autosaved)" and the URL "localhost:8888/notebooks/color_scatterplot.ipynb#". The top menu includes File, Edit, View, Insert, Cell, Kernel, Help, and a "Cell Toolbar" dropdown. The content area shows a series of code cells. Cell 4 imports numpy and bokeh, and defines N=4000. Cell 5 generates random data x and y. Cell 6 creates a scatter plot with radius=10, fill_alpha=0.6, and line_color=None. Cell 7 runs the notebook. Cell 8 shows the plot, which is a scatter plot of 4000 points colored by their coordinates. The plot has axes from 0 to 100.

11. Install and launch Jupyter Notebook

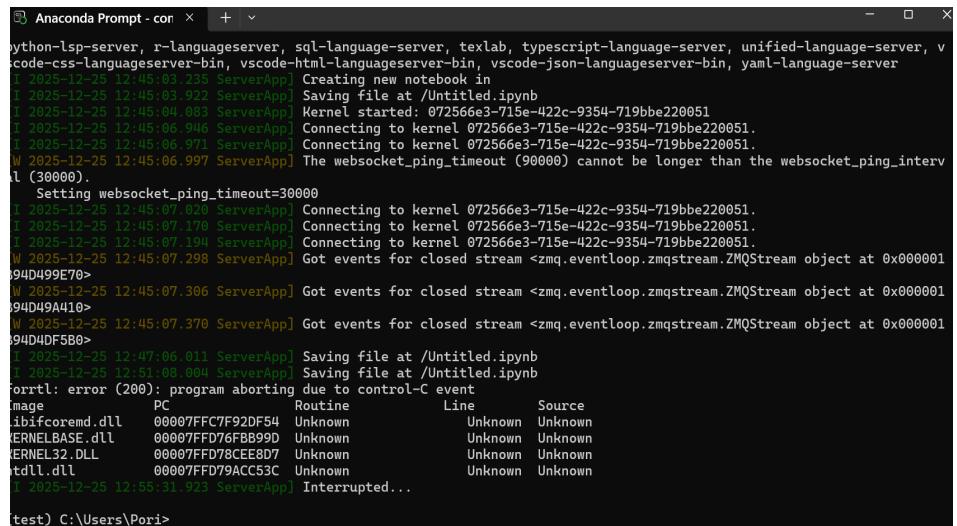
Make sure the **ml** environment is active:

- `pip install notebook`
- `jupyter notebook`

Browser opens automatically
Runs locally on your machine

- Shortcut to stop Jupyter Notebook

Ctrl + C



The screenshot shows an Anaconda Prompt window with the following text:

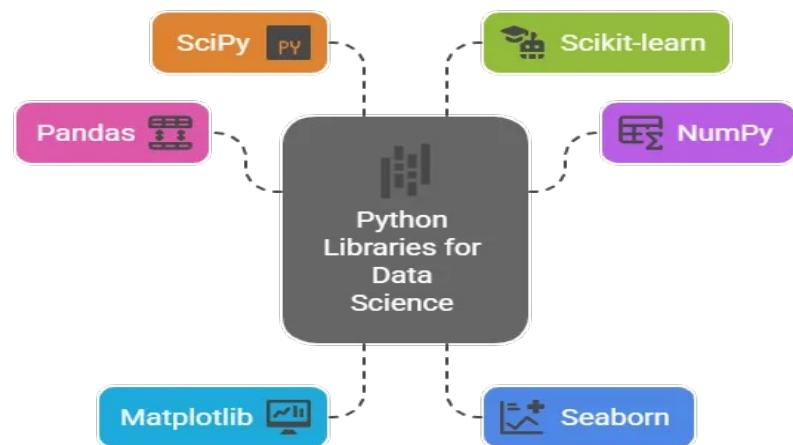
```
python-lsp-server, r-languageserver, sql-language-server, texlab, typescript-language-server, unified-language-server, vscode-css-languageserver-bin, vscode-html-languageserver-bin, vscode-json-languageserver-bin, yaml-language-server
I 2025-12-25 12:45:03.235 ServerApp] Creating new notebook in
I 2025-12-25 12:45:03.922 ServerApp] Saving file at /Untitled.ipynb
I 2025-12-25 12:45:04.083 ServerApp] Kernel started: 072566e3-715e-422c-9354-719bbe220051
I 2025-12-25 12:45:06.946 ServerApp] Connecting to kernel 072566e3-715e-422c-9354-719bbe220051.
I 2025-12-25 12:45:06.971 ServerApp] Connecting to kernel 072566e3-715e-422c-9354-719bbe220051.
W 2025-12-25 12:45:06.997 ServerApp] The websocket_ping_timeout (90000) cannot be longer than the websocket_ping_interval (30000).
Setting websocket_ping_timeout=30000
I 2025-12-25 12:45:07.020 ServerApp] Connecting to kernel 072566e3-715e-422c-9354-719bbe220051.
I 2025-12-25 12:45:07.170 ServerApp] Connecting to kernel 072566e3-715e-422c-9354-719bbe220051.
I 2025-12-25 12:45:07.194 ServerApp] Connecting to kernel 072566e3-715e-422c-9354-719bbe220051.
W 2025-12-25 12:45:07.298 ServerApp] Got events for closed stream <zmq.eventloop.zmqstream.ZMQStream object at 0x000001894D49E70>
W 2025-12-25 12:45:07.306 ServerApp] Got events for closed stream <zmq.eventloop.zmqstream.ZMQStream object at 0x000001894D49410>
W 2025-12-25 12:45:07.370 ServerApp] Got events for closed stream <zmq.eventloop.zmqstream.ZMQStream object at 0x000001894D4DF5B0>
I 2025-12-25 12:47:06.011 ServerApp] Saving file at /Untitled.ipynb
I 2025-12-25 12:51:08.004 ServerApp] Saving file at /Untitled.ipynb
forrtl: error (200): program aborting due to control-C event
Image PC Routine Line Source
libifcoremd.dll 00007FFC7F920F54 Unknown Unknown Unknown
KERNELBASE.dll 00007FFD76FB990 Unknown Unknown Unknown
KERNEL32.DLL 00007FFD78CEEB07 Unknown Unknown Unknown
stdll.dll 00007FFD79ACC53C Unknown Unknown Unknown
I 2025-12-25 12:55:31.923 ServerApp] Interrupted...
(test) C:\Users\Pori>
```

12. Install Required ML Libraries

- `pip install numpy matplotlib scipy scikit-learn`

What Each Library Is For

- NumPy → arrays & maths
- Matplotlib → plots
- Scipy → numerical integration & linear algebra & signal and image processing
- Scikit-learn → ML algorithms
- Jupyter → interactive interface



13. Verify Installation in Jupyter

- Create a new notebook and run:
- import numpy
- import matplotlib.pyplot as plt
- Import scipy
- import sklearn
- print('Setup successful')

14. Common Problems & Fixes

- Command not found → restart terminal
- Module not found → activate environment again
- Wrong Python → check (ml) is active

15. Final Checklist

- Conda installed
- Environment (ml) created and activated
- Jupyter opens correctly
- Libraries import without errors

16. Simple code

```
from sklearn.datasets import load_iris
iris = load_iris()
X = iris.data
y = iris.target
feature_names = iris.feature_names
target_names = iris.target_names
print("Feature names:", feature_names)
print("Target names:", target_names)
print("\nType of X is:", type(X))
print("\nFirst 5 rows of X:\n", X[:5])
```