🧠 Summary: Setting up a Data Science Environment

Goal:
Prepare your computer for the *Introductory Data Analysis Course* using **Python**, **Conda**, and essential libraries.

⚙️ 1. Install Miniconda or Anaconda

- **Conda** is both a **package** and **environment manager** that simplifies installing and managing Python libraries.
- **Steps:**
    1. Download from:
        - [Miniconda](#) or
        - [Anaconda](#)
    2. Run the installer and follow prompts.
    3. Add Conda to your system PATH (optional but useful).
    4. Restart your terminal.

🧩 2. Create and Activate a Virtual Environment

```
conda create -n IDA python=3.9
conda activate IDA
```

This isolates dependencies for your project.

📦 3. Install Required Libraries

```
conda install numpy pandas matplotlib seaborn scikit-learn jupyter scipy statsmodels
```

These libraries cover:

- **NumPy** → numerical computing
- **Pandas** → data manipulation
- **Matplotlib / Seaborn** → data visualization
- **Scikit-learn** → machine learning
- **SciPy / Statsmodels** → statistics and hypothesis testing
- **Jupyter** → interactive notebooks

🧪 4. Test Installation

```
conda activate IDA
mkdir ~/Desktop/IDA_exercises
cd ~/Desktop/IDA_exercises
jupyter notebook .
```

Then test imports inside Jupyter:

```
import numpy as np, pandas as pd, matplotlib.pyplot as plt, seaborn as sns
import sklearn
from scipy import stats
import statsmodels.api as sm
```

💡 5. Tips for Smooth Use

- Update regularly: `conda update --all`
- Close unused notebooks to free memory.
- Practice small projects to master tools.

✅ Outcome:

You'll have a ready-to-use **Python Data Science Environment** capable of data analysis, visualization, and modeling using Jupyter Notebook.

🧠 Data Science Environment Setup – Cheat Sheet

Step	Task	Command / Description	Purpose
1	**Install Conda**	Download & install: 🔗 [Miniconda](#) or [Anaconda](#)	Manages Python packages & environments
2	**Create environment**	`conda create -n IDA python=3.9`	Makes isolated environment for the course
3	**Activate environment**	`conda activate IDA`	Switch to the working environment
4	**Install libraries**	`conda install numpy pandas matplotlib seaborn scikit-learn jupyter scipy statsmodels`	Installs main Data Science libraries
5	**Test installation**	`python import numpy as np, pandas as pd, matplotlib.pyplot as plt, seaborn as sns import sklearn, statsmodels.api as sm, scipy`	Verifies all packages are correctly installed
6	**Launch Jupyter Notebook**	`jupyter notebook .`	Opens interactive workspace
7	**Organize workspace**	`mkdir ~/Desktop/IDA_exercisescd ~/Desktop/IDA_exercises`	Keeps all exercises and notebooks together
8	**Update environment**	`conda update --all`	Keeps packages up to date
9	**Close unused notebooks**	—	Saves memory and avoids kernel crashes

---

🧩 Main Libraries Overview

Library	Purpose
NumPy	Numerical computations, arrays
Pandas	Data manipulation & cleaning
Matplotlib / Seaborn	Data visualization
SciPy / Statsmodels	Statistical tests, modeling
Scikit-learn	Machine learning
Jupyter Notebook	Interactive data analysis workspace

✅ Result:
You now have a fully functional **Python Data Science Environment** ready for data analysis, visualization, and modeling.