





CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Prof. Anis Koubaa

SEP 2024





INTERPRETED LANGUAGE

Python is a **interpreted**, **object-oriented**, **high-level**programming **scripting** language with **dynamic semantics**Developed in the late 1980s by Guido van Rossum



02

PYTHON IS SIMPLE

Python is simple and easy to learn syntax

Supports a large variety of standard libraries and packages

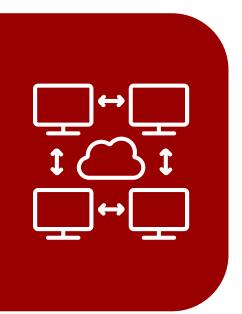
Increased productivity for developers - no segmentation fault



03

MULTI-PURPOSE LANGUAGE

Number 1 language for data science
Web development (back-end side - and recently front-end)
System programming (embedded)





Python Popularity

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

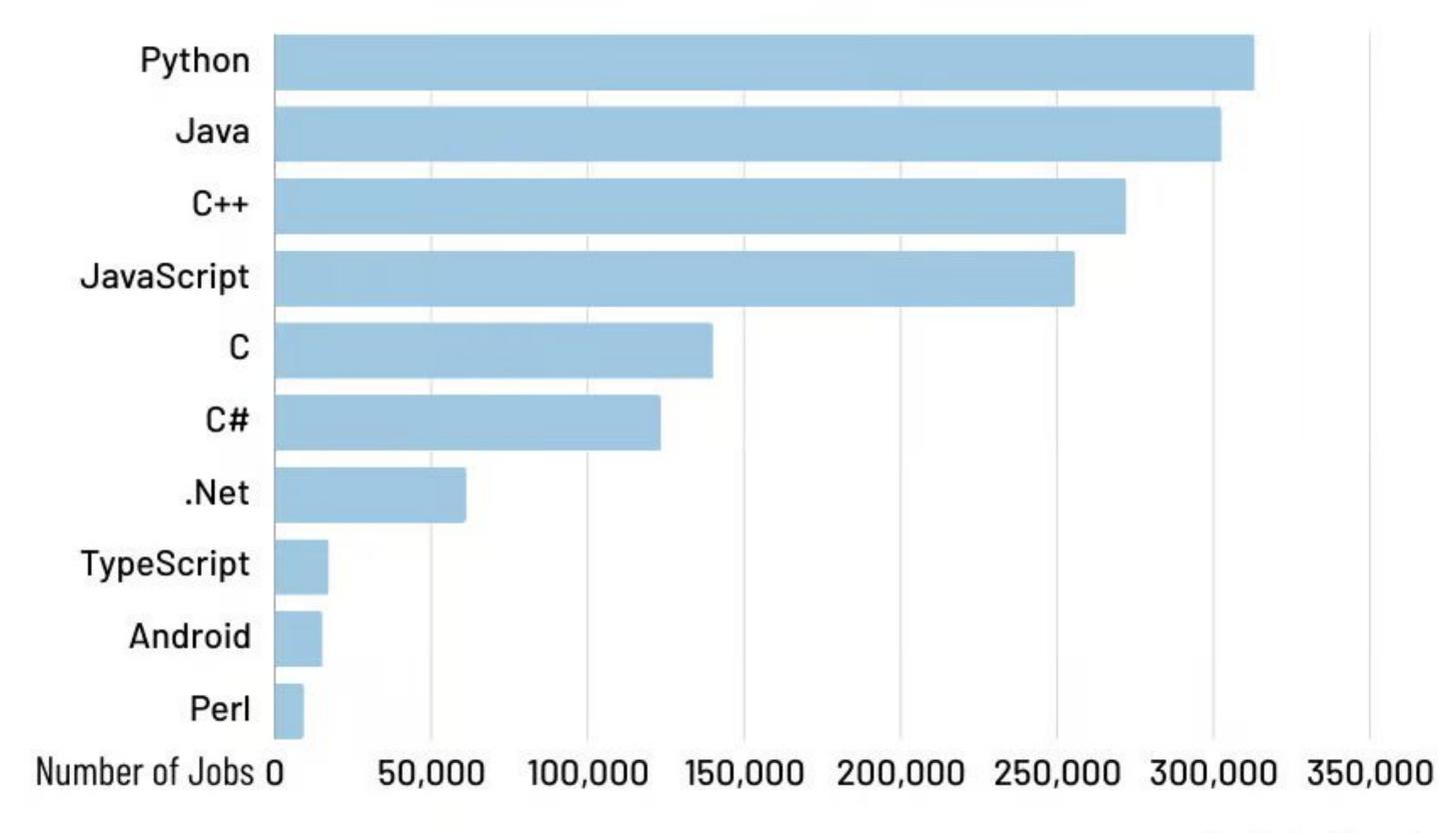
LECTURE 1
Introduction to Python Programming

Most in-demand programming languages of 2024

Based on LinkedIn job postings in the US

PYTHON POPULARITY

IS PYTHON THE MOST POPULAR?



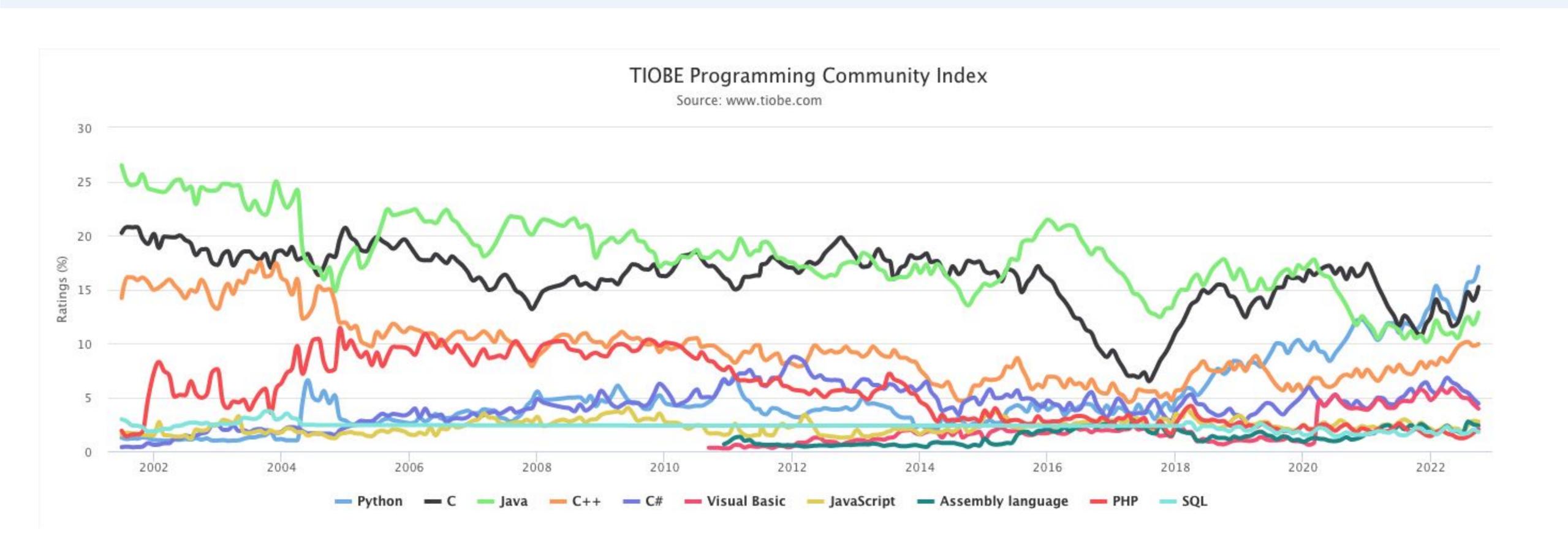
By: CodingNomads

PYTHON POPULARITY

IS PYTHON THE MOST POPULAR?

Oct 2022	Oct 2021	Change	Programming Language		Ratings	Change
1	1			Python	17.08%	+5.81%
2	2		9	С	15.21%	+4.05%
3	3		(4)	Java	12.84%	+2.38%
4	4		0	C++	9.92%	+2.42%
5	5		0	C#	4.42%	-0.84%
6	6		VB	Visual Basic	3.95%	-1.29%
7	7		JS	JavaScript	2.74%	+0.55%
8	10	^	ASM	Assembly language	2.39%	+0.33%
9	9		php	PHP	2.04%	-0.06%
10	8	•	SQL	sQL	1.78%	-0.39%
11	12	^	-GO	Go	1.27%	-0.01%
12	14	^	R	R	1.22%	+0.03%
13	29	*	0	Objective-C	1.21%	+0.76%
14	13	•	A	MATLAB	1.18%	-0.02%
15	17	^	8	Swift	1.05%	-0.06%
16	16		4	Ruby	0.88%	-0.24%
17	11	¥	4	Classic Visual Basic	0.87%	-0.96%
18	20	^	(3)	Delphi/Object Pascal	0.85%	-0.09%
19	18	•	1	Fortran	0.79%	-0.29%
20	26	*	®	Rust	0.70%	+0.17%

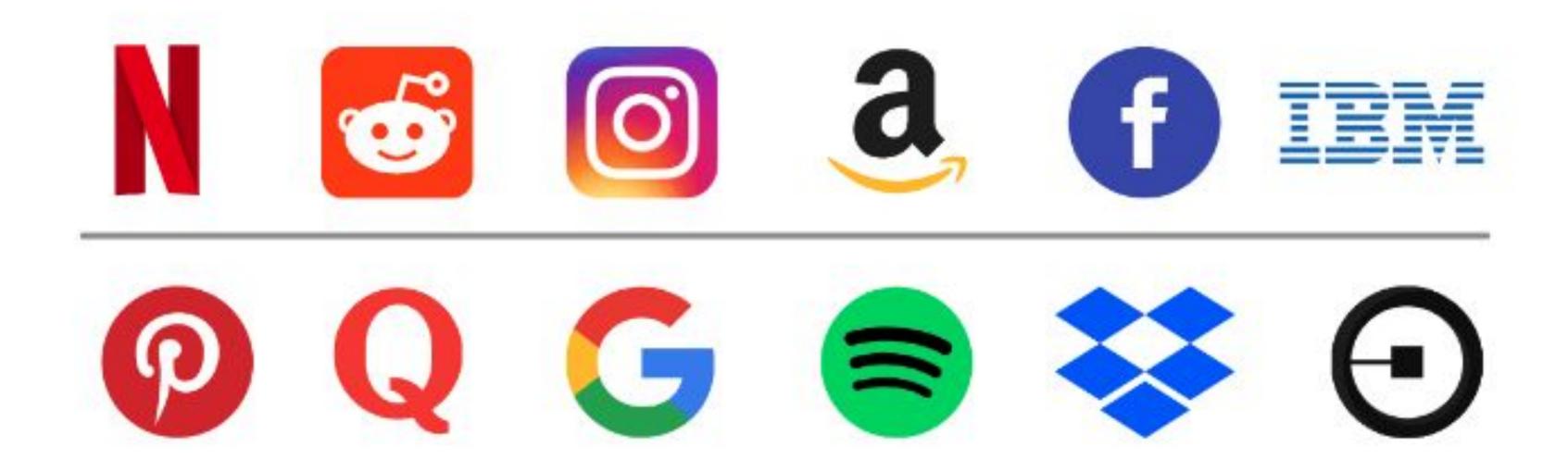
PROGRAMMING LANGUAGES TRENDS



PROGRAMMING LANGUAGES HISTORICAL RANKS

Programming Language	2022	2017	2012	2007	2002	1997	1992	1987
Python	1	5	8	7	12	28	-	6 7 3)
С	2	2	1	2	2	1	1	1
Java	3	1	2	1	1	16	-	-
C++	4	3	3	3	3	2	2	6
C#	5	4	5	8	13		-	673)
Visual Basic	6	15	_	22		-	2	_
JavaScript	7	8	10	9	8	23	÷	-
Assembly language	8	10	-	=	-	-	-	-
sqL	9	-	5 7 .0	7.	7		-	- -
PHP	10	7	6	5	6	_	_	_
Prolog	25	32	32	26	16	20	13	3
Lisp	31	31	13	16	14	9	5	2
Pascal	269	117	14	21	99	11	3	5
(Visual) Basic	772	_	7	4	4	3	6	4

BIG TECH APPS USING PYTHON





Install Python

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Python Installation Options

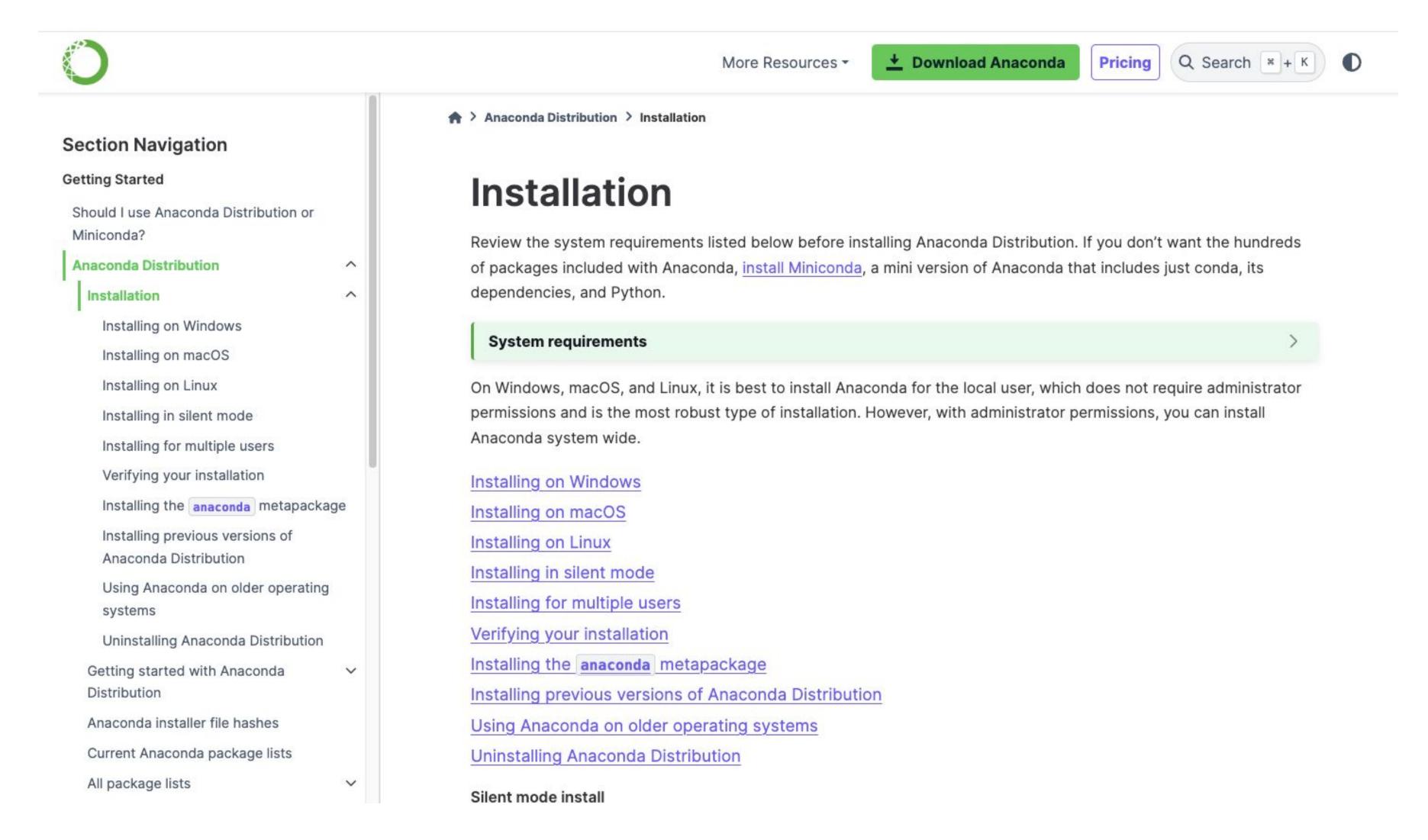
Official Python Distribution

- Availability: Windows, Mac, Linux
- Source: Download directly from <u>python.org</u>
- o Instructions: Follow installation steps for your OS, ensuring the addition of Python to the system PATH.

Anaconda Distribution (Recommended for Data Science)

- o Includes: Python, Conda (environment manager), and key pre-installed libraries (e.g., NumPy, Pandas, SciPy).
- Compatibility: Available for Windows, Mac, and Linux
- Download: Obtain from Anaconda.com
- o Advantages: Simplifies package management and deployment with isolated environments.

Python Installation Options



https://docs.anaconda.com/anaconda/install/

CS316: INTRODUCTION TO DATA SCIENCE INTRODUCTION TO DATA SCIENCE ANIS KOUBAA | 2024

1. Check Existing Python Installation

- Command: python --version Or python3 --version
- o **Purpose:** Verify if Python is already installed and check the current version.

1. Installing Python on Different OS

- Linux:
 - Update Package List: sudo apt-get update
 - Install Python:
 - Python 3: sudo apt-get install python 3
 - Pip (Package Manager): sudo apt-get install python3-pip

1. Installing Python on Different OS

- macOS:
 - Using Homebrew:
 - Install Homebrew (if not installed):

```
/bin/bash -c "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

Install Python:

brew install python

1. Installing Python on Different OS

- Windows:
 - Using Windows Subsystem for Linux (WSL):
 - **Enable WSL:**

```
wsl --install
```

- Install Python within WSL:
 - Update: sudo apt-get update
 - Install: sudo apt-get install python3



Python IDE

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Development Environments

1. IDLE

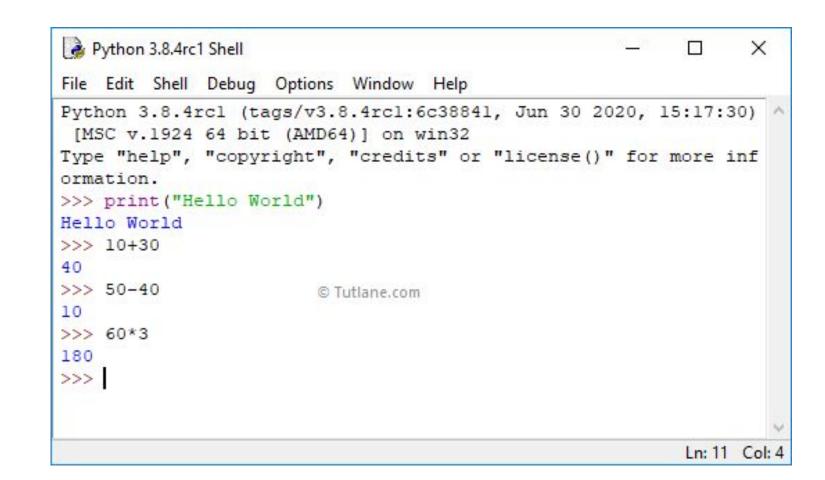
- Description: Integrated Development and Learning Environment, comes pre-installed with Python.
- Use Case: Suitable for beginners and small-scale projects.

2. PyCharm

- Versions: Community Edition (free) and Professional Edition.
- Features: Tailored for Python development with advanced debugging, code analysis, and refactoring tools.

3. Visual Studio Code (VS Code)

- Description: Free, open source, and highly customizable.
- Python Support: Enhanced through extensions like
 Python and Pylance, ideal for various development
 tasks.







Notebook Development Environments

1. Jupyter Notebook

- Description: Ideal for interactive Python sessions, especially in data science.
- Access: Available via Anaconda distribution or pip install notebook.

2. Google Colab

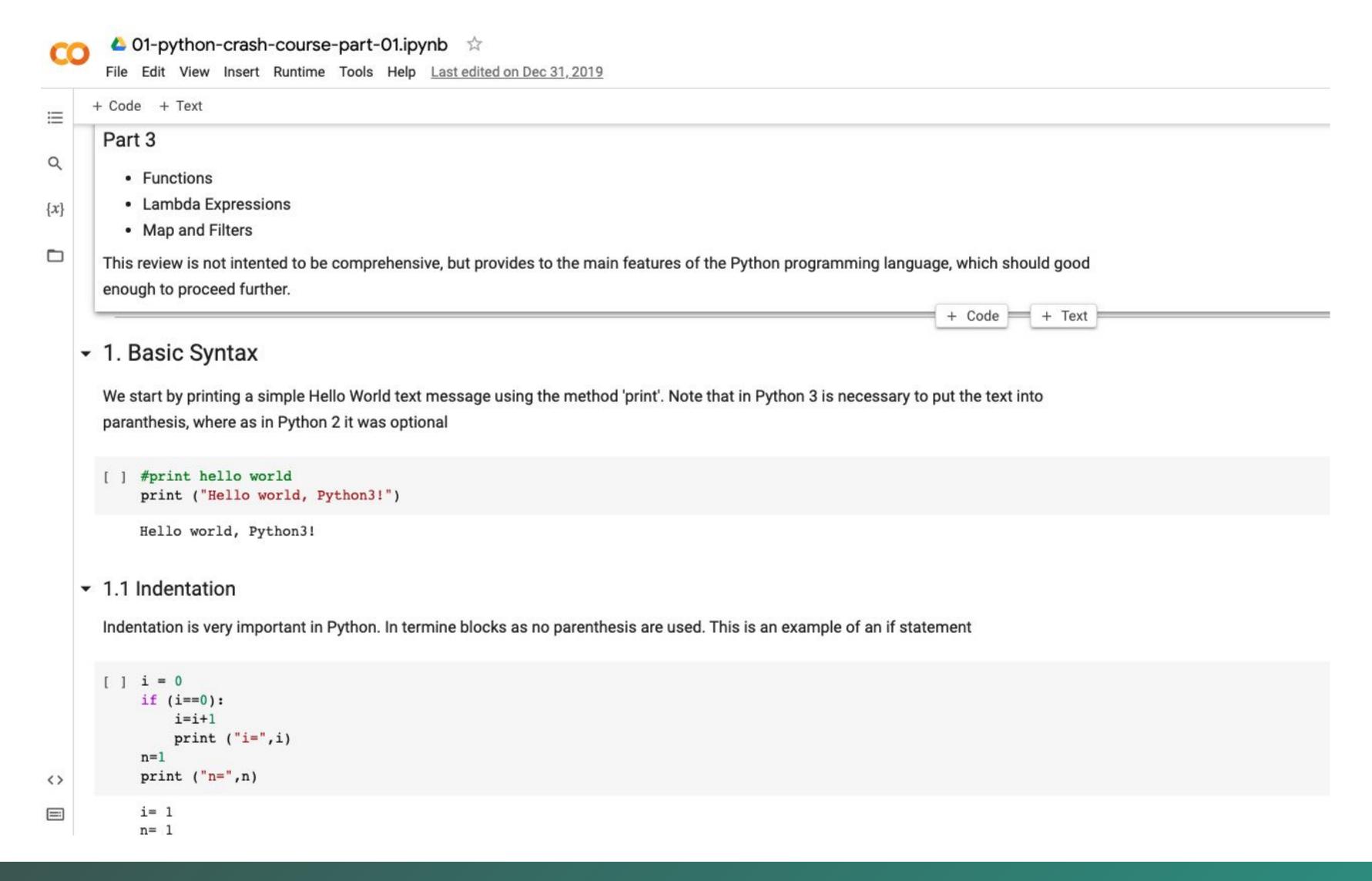
- Versions: Free cloud-based Jupyter notebook environment.
- Advantages: Provides GPU/TPU support, easy sharing, and no local setup required.





GOOGLE COLAB







Python Libraries

Prof. Anis Koubaa

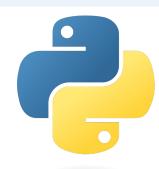


CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

PYTHON FOR DATA SCIENCE LIBRARIES/PACKAGES



DATA STRUCTURES





VISUALIZATION





SCIENTIFIC COMPUTING







STORAGE AND DATABASES









DEEP LEARNING















CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 2
Introduction to Python Programming
Virtual Envs, Setting-up a Python Package

Prof. Anis Koubaa

SEP 2024



Python Package Structure

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Python Module

1. Definition

A single file containing Python code, typically with a `.py` extension.

2. Example

`example.py`

3. Usage

- Functions, classes, and variables defined in the file can be imported into other Python scripts.
- Import Syntax: `import example`
- Example Code:

```
# In example.py
def greet():
    print("Hello, world!")

# In another script
import example
example.greet()
```



Python Package

2. Package

- **Definition**: A directory (folder) containing multiple modules, and an `__init__.py` file.
- Purpose: Organizes related modules under a single namespace.
- Example Structure:
 - `mypackage/`
 - `__init__.py` (can be empty)
 - `module1.py`
 - `module2.py`
- Import Syntax:
 - `from mypackage import module1`
 - `import mypackage.module2`



Python Libraries

3. Standard Library

- Description: A rich set of modules and packages included with Python by default.
- Purpose: Provides tools for various tasks such as file I/O, system calls, and web development.
- Examples: `os`, `sys`, `datetime`, `json`

4. Third-Party Packages

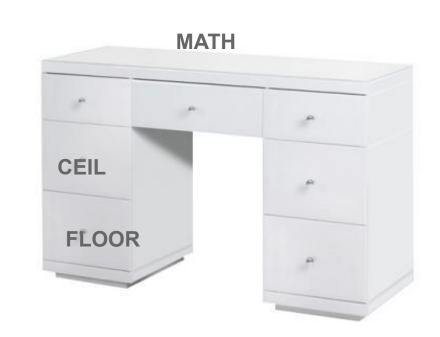
- Definition: External libraries developed by the community, available through package managers like `pip`.
- Purpose: Extend Python's functionality for specific use cases, such as data science or web development.
- Examples: `numpy`, `requests`, `pandas`

```
Copy code
python
# Importing standard library module
import datetime
# Importing third-party library module
import pandas as pd
def main():
   # Using datetime to get today's date
   today = datetime.date.today()
   # Using pandas to create a DataFrame with today's date
   df = pd.DataFrame({'Date': [today]})
   print(df)
if __name__ == "__main__":
   main()
```

PYTHON FILE STRUCTURE MODULE vs. PACKAGE

PACKAGE

A collection of Python module. It has a __init__.py file to distinguish it from a classical director



MODULE-1

A module is a file that contains a Python script in runtime for the code specified to the users.

MODULE-2

A module is a file that contains a Python script in runtime for the code specified to the users.



Python Virtual Environment

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Creating a Virtual Environment with pip

2. Creating a Virtual Environment with `pip`

- Steps:
 - Install `venv` (if not installed):



2. Create the Environment:



- 3. Activate the Environment:
 - Windows:



macOS/Linux:



4. Deactivate:



Creating a Virtual Environment with conda

- 3. Creating a Virtual Environment with `conda`
- Steps:
 - 1. Create the Environment:



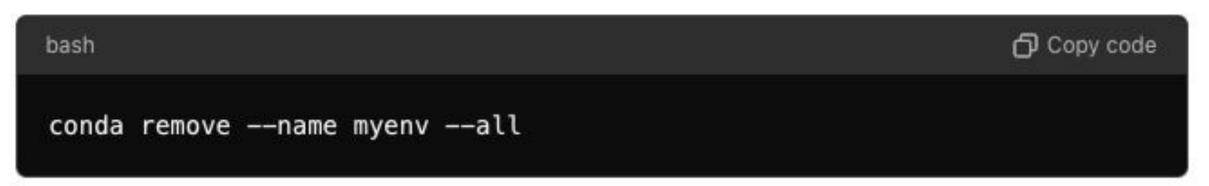
2. Activate the Environment:



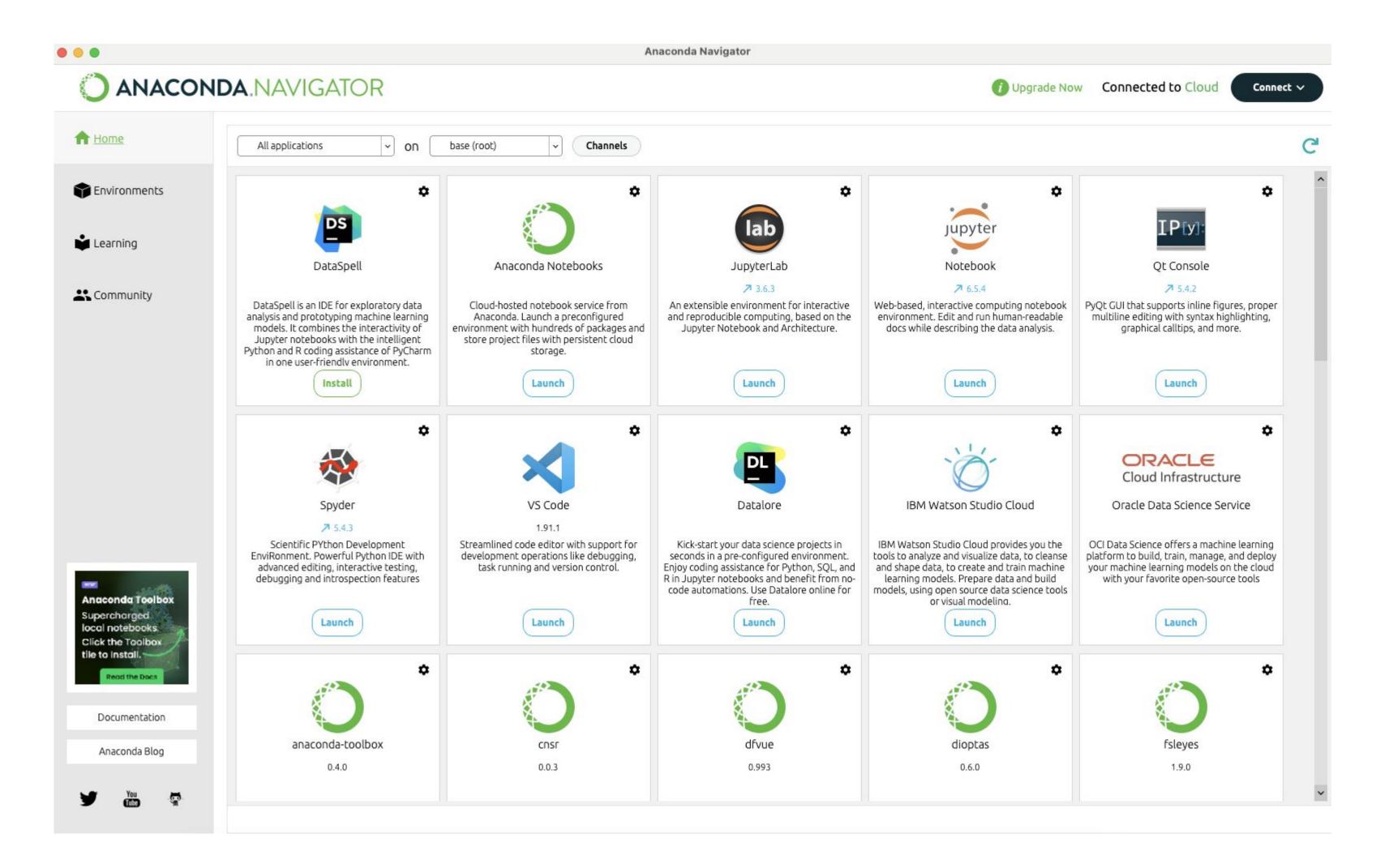
3. Deactivate:



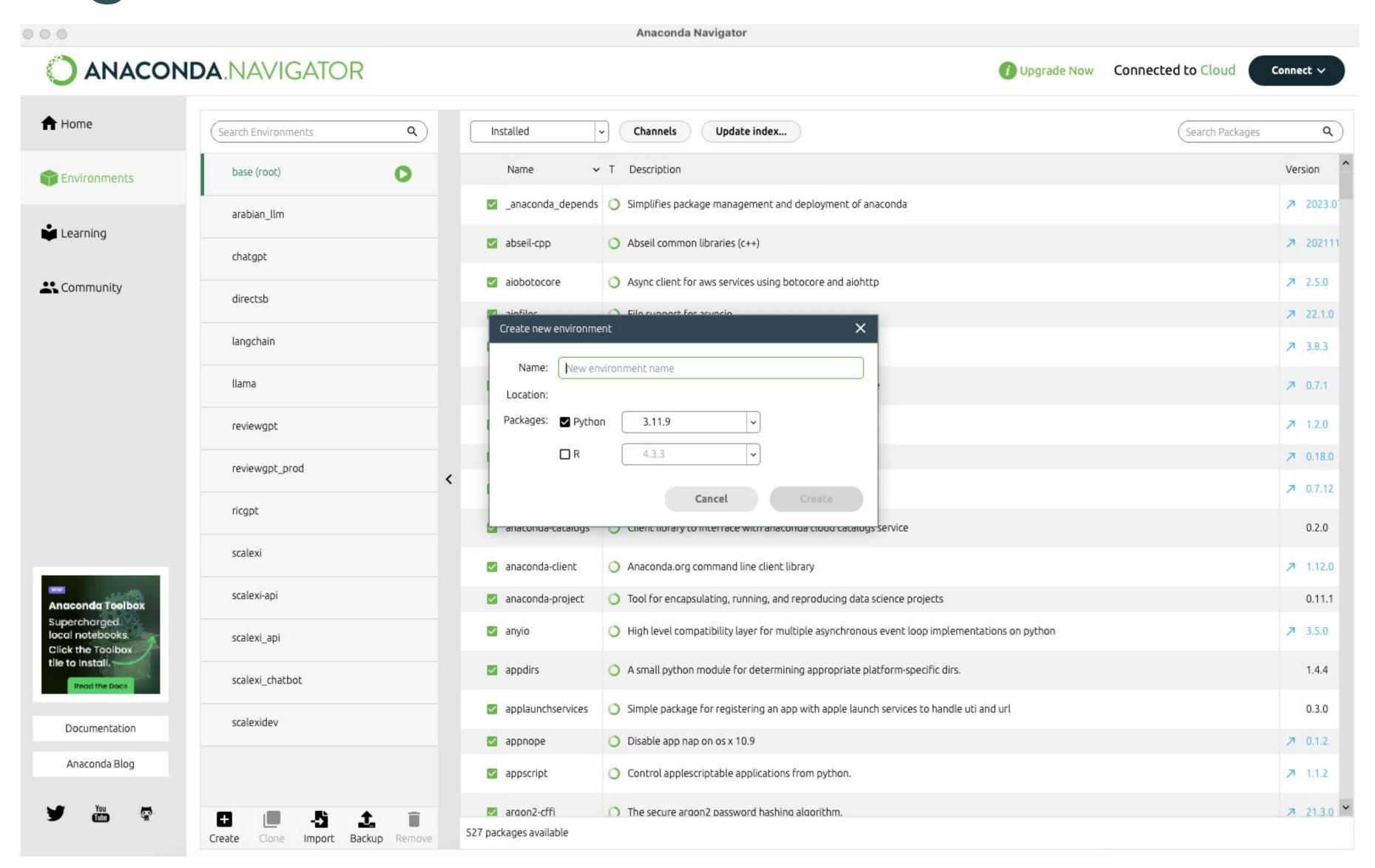
4. Remove the Environment:



Creating a Virtual Environment with conda



Creating a Virtual Environment with conda



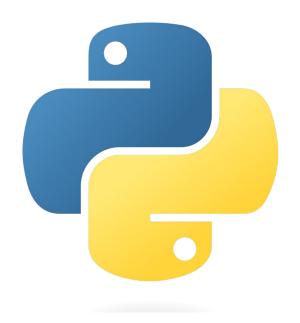
Managing Packages within the Environment

4. Managing Packages within the Environment

- Using `pip`:
 - Install a package: `pip install package_name`
 - List installed packages: `pip list`
- Using `conda`:
 - Install a package: `conda install package_name`
 - List installed packages: `conda list`

5. Considerations

- pip : Works with the standard Python distribution; great for most projects.
- conda: Ideal for data science projects with complex dependencies, especially when working with packages that require compiled binaries (e.g., `numpy`, `scipy`).



Setup a Python Package

1. Package Structure

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Python Package Structure

1. Package Structure

Structure Overview:

```
Copy code
plaintext
mypackage/
— mypackage/
      — __init__.py
     module1.py
     module2.py
 — tests/
    — test_module1.py
     — test_module2.py
 setup.py
   README.md
└─ LICENSE
```

- Key Files:
 - `__init__.py`: Marks the directory as a Python package.
 - `setup.py`: Script for packaging and distribution.
 - `README.md`: Documentation for the package.
 - `LICENSE`: License information.

Writing setup.py

2. Basic `setup.py` Script Example

```
python
                                                                      Copy code
from setuptools import setup, find_packages
setup(
   name="mypackage",
   version="0.1.0",
   author="Your Name",
   author_email="your.email@example.com",
   description="A short description of the package",
   long_description=open("README.md").read(),
   long_description_content_type="text/markdown",
   url="https://github.com/yourusername/mypackage",
   packages=find_packages(),
   classifiers=[
       "Programming Language :: Python :: 3",
       "License :: OSI Approved :: MIT License",
       "Operating System :: OS Independent",
   python_requires='>=3.6',
   install_requires=[
       "numpy>=1.21.0", # Numerical computations
       "pandas>=1.3.0", # Data manipulation and analysis
       "requests>=2.26.0",# HTTP library for making requests
       "matplotlib>=3.4.0" # Plotting and visualization
   entry_points={
       'console_scripts': [
            'mypackage-cli=mypackage.command_line:main',
```

1. Purpose

'setup.py' is the script used for packaging and distributing your Python project. It contains
metadata about the package and specifies dependencies.

3. Explanation

- name : The name of the package.
- · `version`: The current version of the package.
- `author` and `author_email`: Your details.
- 'description': A brief description of the package.
- `long_description`: Detailed description, often pulled from `README.md`.
- `url`: The URL of the project's repository or website.
- packages : Automatically find all packages to include in the distribution.
- `classifiers`: Metadata about the package (e.g., programming language, license).
- `python_requires`: Specifies the Python version compatibility.
- install_requires: Lists the dependencies that will be installed with the package.
 - Examples:
 - `numpy`: A library for numerical computations.
 - pandas : A library for data manipulation and analysis.
 - requests: A library for making HTTP requests.
 - `matplotlib`: A library for plotting and visualization.
- 'entry_points': Defines entry points for creating command-line scripts.

README

MyPackage

Overview

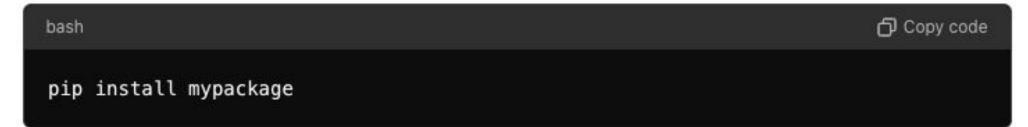
MyPackage is a simple Python package designed for educational purposes. It provides basic utility functions that demonstrate key concepts in Python, such as string manipulation, mathematical operations, and data handling.

Features

- String Utilities: Functions to manipulate and analyze strings.
- Math Utilities: Basic mathematical operations, including addition, subtraction, multiplication, and division.
- · Data Utilities: Simple data processing functions.

Installation

You can install MyPackage using `pip`:



Alternatively, you can clone the repository and install it manually:

```
git clone https://github.com/yourusername/mypackage.git
cd mypackage
pip install .
```

Usage

Below are some examples of how to use MyPackage:

1. String Utilities

```
from mypackage import string_utils

# Reverse a string
reversed_string = string_utils.reverse_string("hello")
print(reversed_string) # Output: "olleh"

# Count vowels in a string
vowel_count = string_utils.count_vowels("hello")
print(vowel_count) # Output: 2
```

README

Usage

Below are some examples of how to use MyPackage:

1. String Utilities

```
from mypackage import string_utils

# Reverse a string
reversed_string = string_utils.reverse_string("hello")
print(reversed_string) # Output: "olleh"

# Count vowels in a string
vowel_count = string_utils.count_vowels("hello")
print(vowel_count) # Output: 2
```

2. Math Utilities

```
python

from mypackage import math_utils

# Add two numbers
result = math_utils.add(5, 3)
print(result) # Output: 8

# Divide two numbers
result = math_utils.divide(10, 2)
print(result) # Output: 5.0
```

3. Data Utilities

```
from mypackage import data_utils

# Sum of a list of numbers
total = data_utils.sum_list([1, 2, 3, 4])
print(total) # Output: 10

# Find maximum in a list of numbers
maximum = data_utils.find_max([1, 2, 3, 4])
print(maximum) # Output: 4
```

Modules

1. string_utils

- reverse_string(s: str) -> str`: Reverses the given string `s`.
- `count_vowels(s: str) -> int`: Counts the number of vowels in the given string `s`.

2. math_utils

- `add(a: int, b: int) -> int`: Returns the sum of `a` and `b`.
- `subtract(a: int, b: int) -> int`: Returns the difference between `a` and `b`.
- `multiply(a: int, b: int) -> int`: Returns the product of `a` and `b`.
- 'divide(a: int, b: int) -> float': Returns the division of `a` by `b`.

3. data_utils

• `sum_list(numbers: list) -> int`: Refines the sum of all numbers in the list.

response to the response to the property of th

README

Contributing

Contributions are welcome! If you have any suggestions or find any bugs, please feel free to open an issue or submit a pull request.

Steps to Contribute:

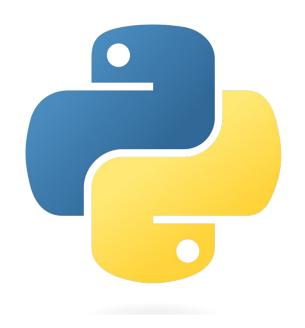
- Fork the repository.
- 2. Create a new branch (`git checkout -b feature-branch`).
- Make your changes.
- 4. Commit your changes ('git commit -m 'Add new feature').
- 5. Push to the branch ('git push origin feature-branch').
- 6. Open a pull request.

License

This project is licensed under the MIT License - see the LICENSE file for details.

Contact

For any questions or suggestions, feel free to contact me at your.email@example.com.



Setup a Python Package

2. PyPi Repository

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1 Introduction to Python Programming

Understanding PyPI

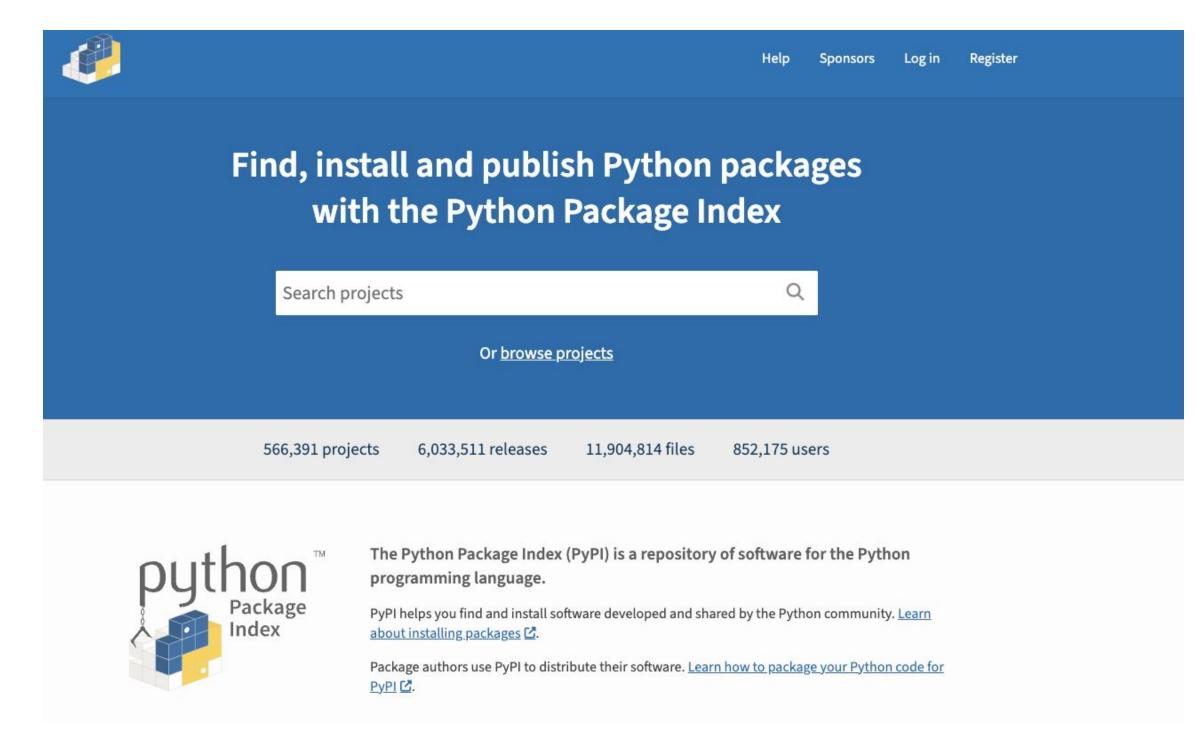
1. What is PyPI?

- Definition: The Python Package Index (PyPI) is the official third-party software repository for Python.
- Purpose: Allows developers to distribute their Python software and libraries to users.

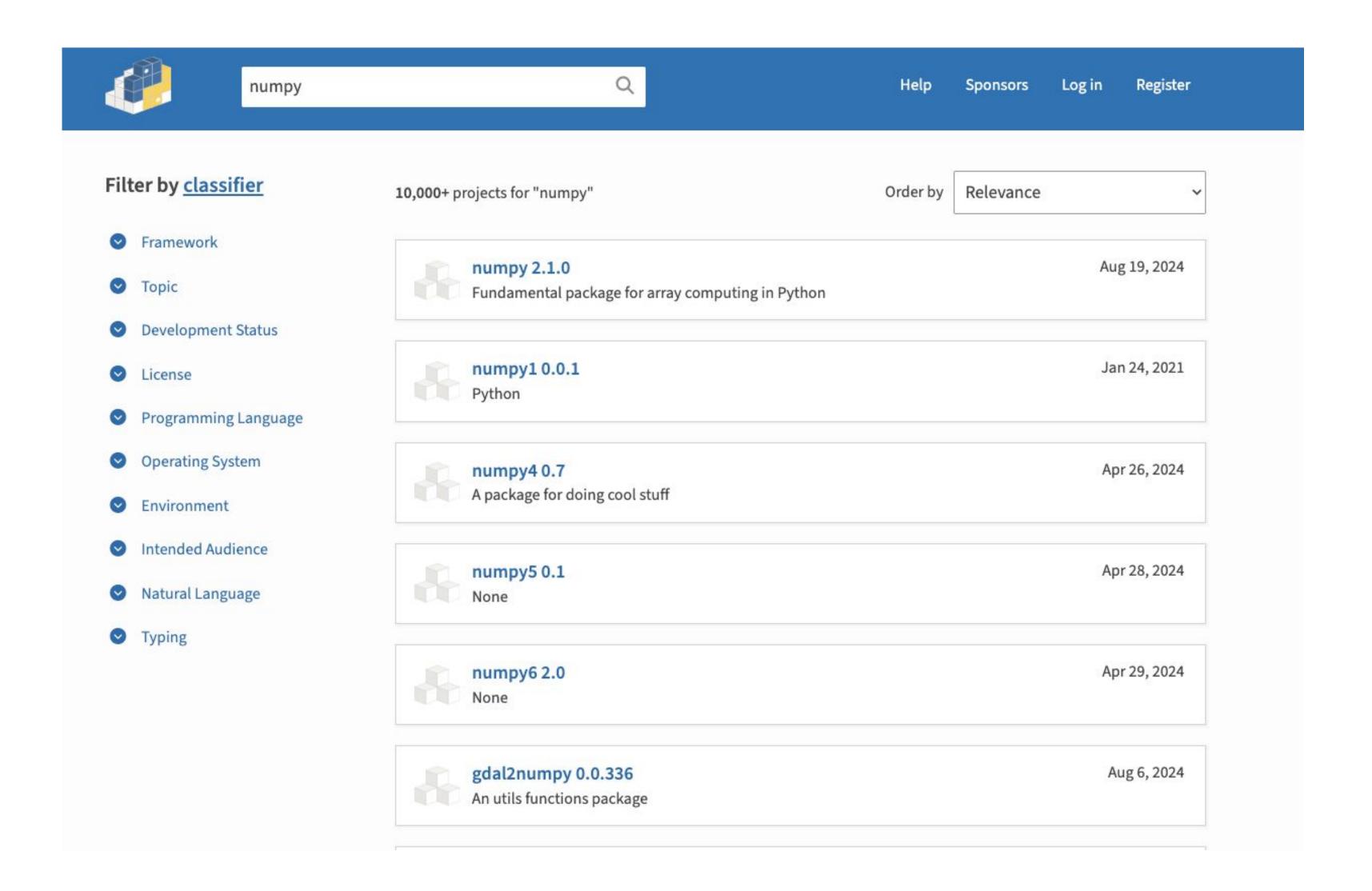
2. Key Features

- Package Management: Easy installation and updating of packages using `pip`.
- Searchable: Users can search for and discover packages based on functionality and popularity.
- Community-Driven: Contributions from developers around the world.

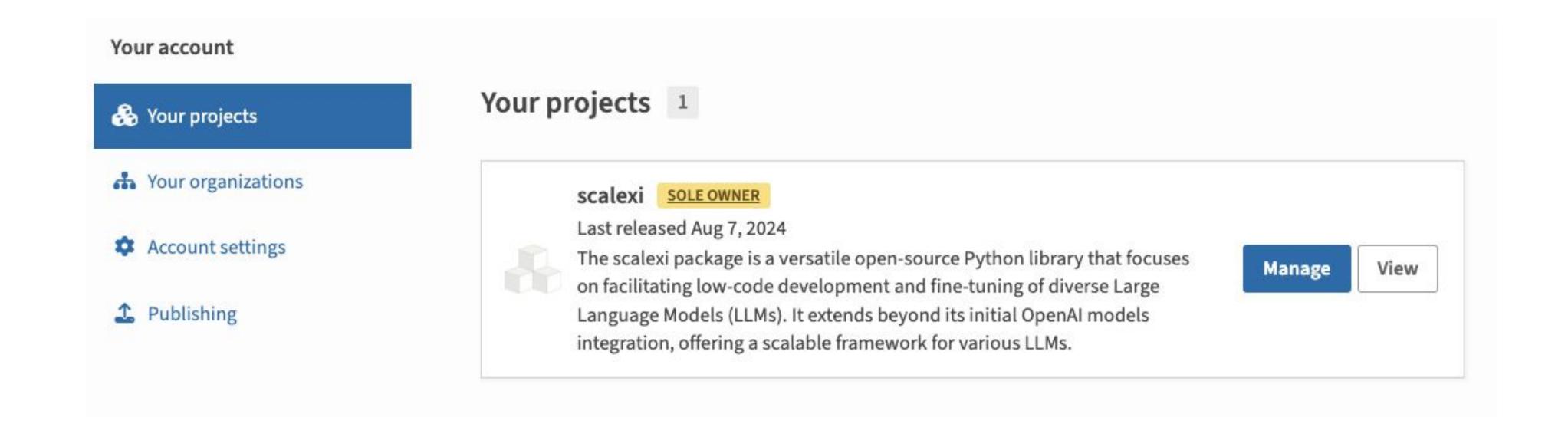
https://pypi.org/



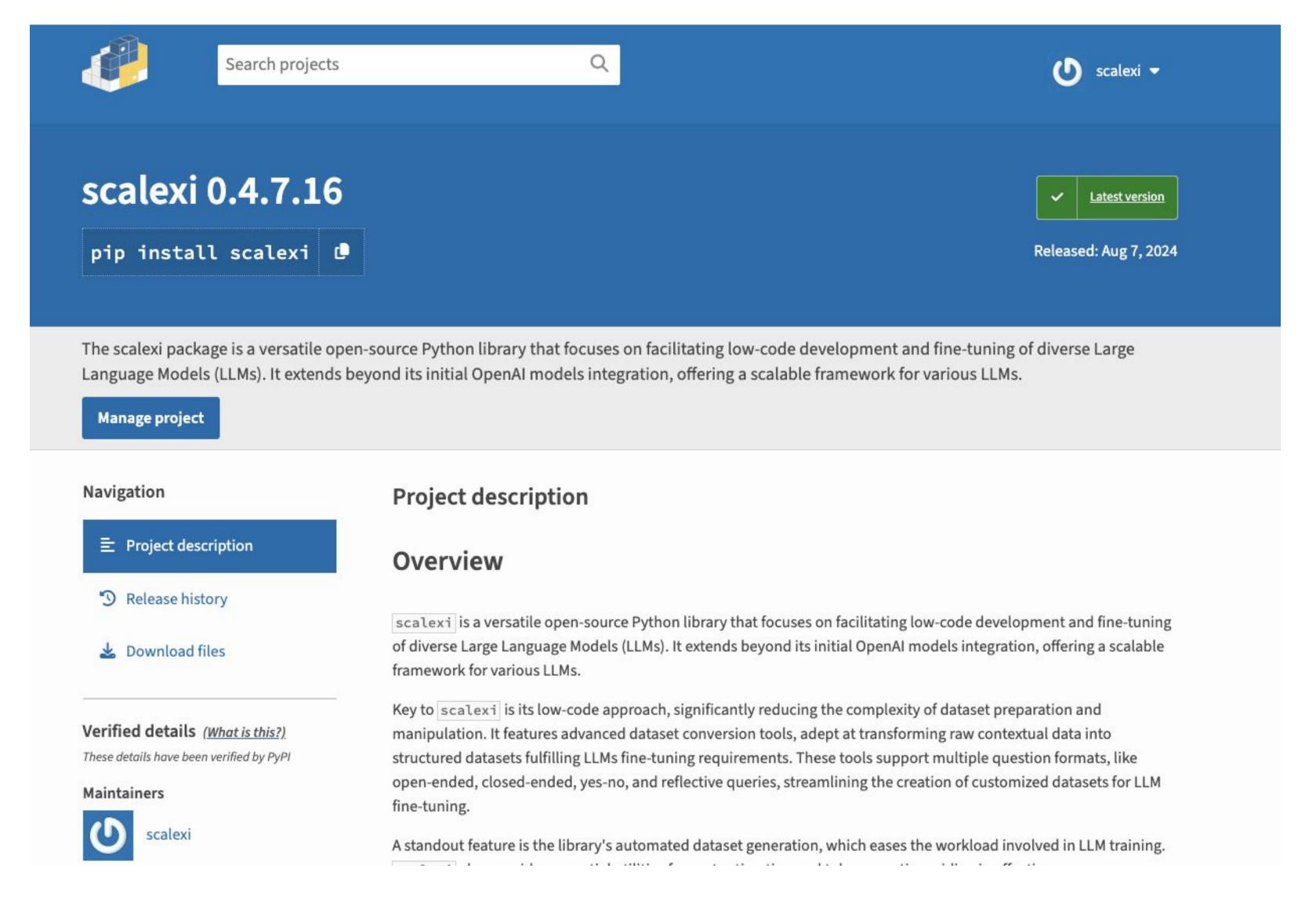
Understanding PyPI



Example



Example



Best Practices for PyPl

1. Versioning

- Use Semantic Versioning: Follow the `MAJOR.MINOR.PATCH` format.
- Example: `1.0.0` for initial stable release.

2. Documentation

- Include a `README.md`: Helps users understand the package's purpose and usage.
- Provide Examples: Demonstrate how to use your package.

3. Testing Before Upload

- Use TestPyPI: Upload to TestPyPI to ensure everything works correctly before the official release.
- Continuous Integration: Implement CI/CD pipelines to automate testing and deployment.

Setting Up PyPI Credentials for Uploading a Wheel

1. Create a PyPI Account

- Steps:
 - Go to pypi.org.
 - Click on "Register" and create a new account.
 - Verify your email address to activate your account.

2. Install `twine`

- Purpose: `twine` is the recommended tool for securely uploading your distribution packages to PyPI.
- Command:



Setting Up PyPl Credentials for Uploading a Wheel

3. Set Up PyPI Credentials

- Option 1: Use `.pypirc` File (Recommended)
 - Create a `.pypirc` file in your home directory (`~/.pypirc` on Linux/macOS or `%USERPROFILE%\.pypirc` on Windows).
 - 2. Add the following configuration:

```
[distutils]
index-servers =
   pypi

[pypi]
username = your-username
password = your-password
```

3. Explanation:

- The `username` and `password` are your PyPI credentials.
- This file allows you to avoid entering your credentials every time you upload a package.

Setting Up PyPI Credentials for Uploading a Wheel

3. Set Up PyPI Credentials

- Option 1: Use `.pypirc` File (Recommended)
 - Create a `.pypirc` file in your home directory (`~/.pypirc` on Linux/macOS or `%USERPROFILE%\.pypirc` on Windows).
 - 2. Add the following configuration:

```
[distutils]
index-servers =
   pypi

[pypi]
username = your-username
password = your-password
```

3. Explanation:

- The `username` and `password` are your PyPI credentials.
- This file allows you to avoid entering your credentials every time you upload a package.

- Option 2: Manual Entry
 - When using `twine` without a `.pypirc` file:



 Prompt: `twine` will prompt you for your PyPI username and password during the upload process.

4. Upload the Wheel to PyPI

Build the Wheel:



Upload the Package:



Outcome: Your package will be uploaded to PyPI and can be installed by others using `pip install your-package-name`.



3. Create a Wheel (Python Distribution)

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Python Wheel

1. What is a Wheel?

- Definition: A Wheel is a built distribution format for Python packages. It is the standard format for distributing Python libraries and applications and is designed to be faster and simpler to install compared to source distributions (`.tar.gz`).
- File Extension: `.whl`

2. Advantages of Using Wheels

- Faster Installation: Wheels are pre-compiled, which speeds up the installation process because they don't require building or compiling code on the target machine.
- Compatibility: Wheels ensure compatibility across different environments, as they bundle all the necessary files and metadata.
- Ease of Use: Installing a wheel is as simple as running `pip install package_name.whl`.

Python Package Structure for Wheel Distribution

1. Package Structure Overview

```
Copy code
plaintext
mypackage/
   mypackage/
                             # Initializes the package
      — __init__.py
     module1.py
                             # Example module 1
     module2.py
                             # Example module 2
   tests/
       test_module1.py
                             # Unit tests for module 1
       test_module2.py
                             # Unit tests for module 2
   dist/
                             # Distribution files, including .whl
       mypackage-0.1.0-py3-none-any.whl # The Wheel file
                             # Build artifacts created by setup.py
   build/
   mypackage.egg-info/
                             # Metadata about the package
                             # Script for packaging and distribution
   setup.py
                             # Package description and usage
    README.md
   LICENSE
                             # License information
```

Key Components

`mypackage/` Directory: Contains the actual Python package with modules.

`__init__.py`: Initializes the package and can define what is exposed to users.

`tests/` Directory: Contains unit tests for the package.

`dist/` Directory: Stores the distribution files, including the `.whl` file after building.

`build/` Directory: Contains temporary build artifacts created during the package build process.

`mypackage.egg-info/` Directory: Contains metadata about the package, including version, dependencies, and more.

`setup.py`: Script that defines the package details and is used to build the distribution files.

`README.md`: Provides a detailed description of the package, including installation and usage instructions.

`LICENSE`: Specifies the licensing terms under which the package is distributed.

Building the Package and Creating a Wheel

1. Building the Package

Command:



- Explanation:
 - 'sdist': Creates a source distribution.
 - bdist_wheel: Creates a binary distribution (wheel).

2. Resulting Files

- · The command generates the following files in the `dist/` directory:
 - `mypackage-0.1.0.tar.gz`: Source distribution.
 - `mypackage-0.1.0-py3-none-any.whl`: Wheel file.

1. Package Structure Overview

```
Copy code
plaintext
mypackage/
— mypackage/
                           # Initializes the package
    ___init__.py
   module1.py
                           # Example module 1
   module2.py
                           # Example module 2
 — tests/
   — test_module1.py
                           # Unit tests for module 1
   — test_module2.py
                           # Unit tests for module 2
                           # Distribution files, including .whl
 — dist/
   mypackage-0.1.0-py3-none-any.whl # The Wheel file
                           # Build artifacts created by setup.py
 — build∕
 — mypackage.egg-info/
                           # Metadata about the package
                           # Script for packaging and distribution
 setup.py
                           # Package description and usage
  README.md
 — LICENSE
                           # License information
```

Python Package Distribution Files

1. Generated Files After Building the Package

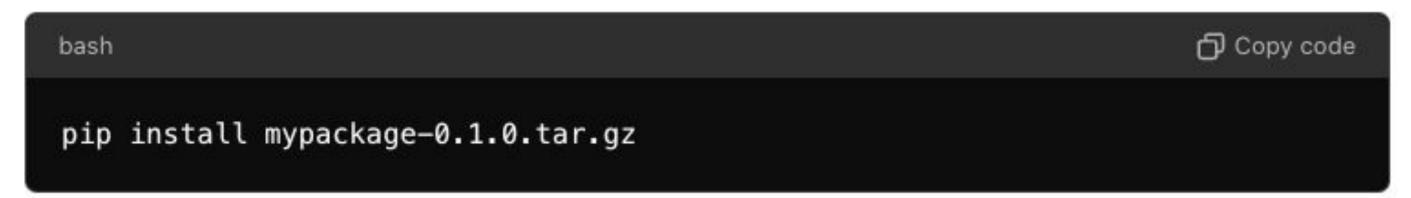
After running the build command (`python setup.py sdist bdist_wheel`), two key files are generated in the `dist/` directory:

- `mypackage-0.1.0.tar.gz`
 - Type: Source Distribution
 - Description: This is a compressed archive containing the source code of the package. It
 includes the Python files, `setup.py`, and any other resources needed to build the
 package.
 - Purpose: Useful for users who want to build the package from source, possibly in environments where pre-built wheels aren't available or applicable.
- mypackage=0.1.0-py3-none-any.whl`
 - Type: Wheel File
 - Description: This is a built distribution format, commonly referred to as a "wheel". It is a
 binary package format that contains everything needed to install the package quickly.
 - Purpose: Provides a fast and efficient installation method because the package is already built and does not require compiling on the user's system. The `py3` indicates compatibility with Python 3, and `none-any` signifies that this wheel is universal, meaning it can be installed on any platform without modification.

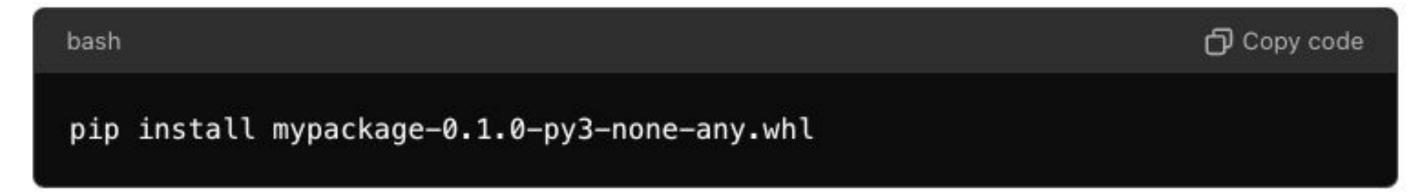
Using Wheels to Install Python Package Locally

2. Usage

Installing the Source Distribution:



Installing the Wheel File:



Installing Directly from the Source Directory:



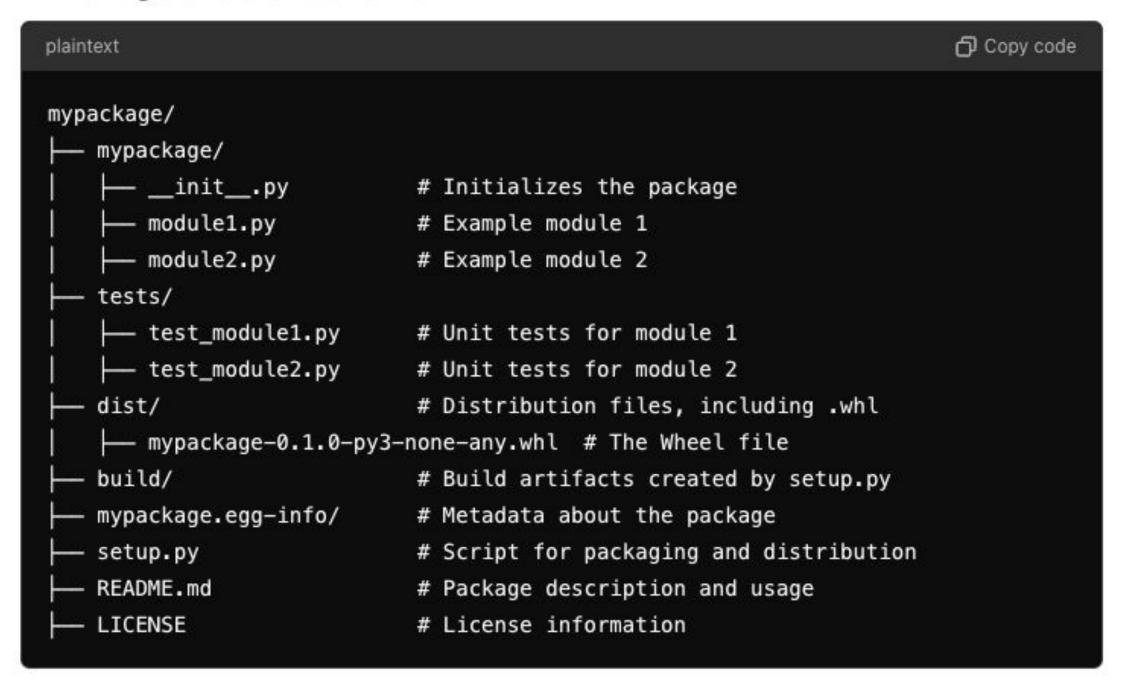
Explanation: The command `pip install .` installs the package directly from the
current directory where `setup.py` is located. This is useful during development or when
testing a package before distribution.

3. Best Practices

- Distribute Both: Offer both the source distribution (`.tar.gz`) and the wheel file (`.whl`) to
 give users flexibility based on their environment.
- Use Wheels When Possible: Recommend the wheel file for faster installations, especially for packages with complex dependencies.

Python Package Structure for Wheel Distribution

1. Package Structure Overview



2. Key Components

- `mypackage/` Directory: Contains the actual Python package with modules.
- `__init__.py`: Initializes the package and can define what is exposed to users.
- `tests/` Directory: Contains unit tests for the package.
- 'dist/' Directory: Stores the distribution files, including the '.whl' file after building.
- build/`Directory: Contains temporary build artifacts created during the package build process.
- `mypackage.egg-info/` Directory: Contains metadata about the package, including version, dependencies, and more.
- setup.py`: Script that defines the package details and is used to build the distribution files.
- `README.md`: Provides a detailed description of the package, including installation and usage instructions.
- `LICENSE`: Specifies the licensing terms under which the package is distributed.

Python Wheel

5. Uploading a Wheel to PyPI

- Steps:
 - 1. Build the Wheel:



2. Upload to PyPI using Twine:



- Explanation:
 - This process uploads the wheel to PyPI, making it available for others to install.

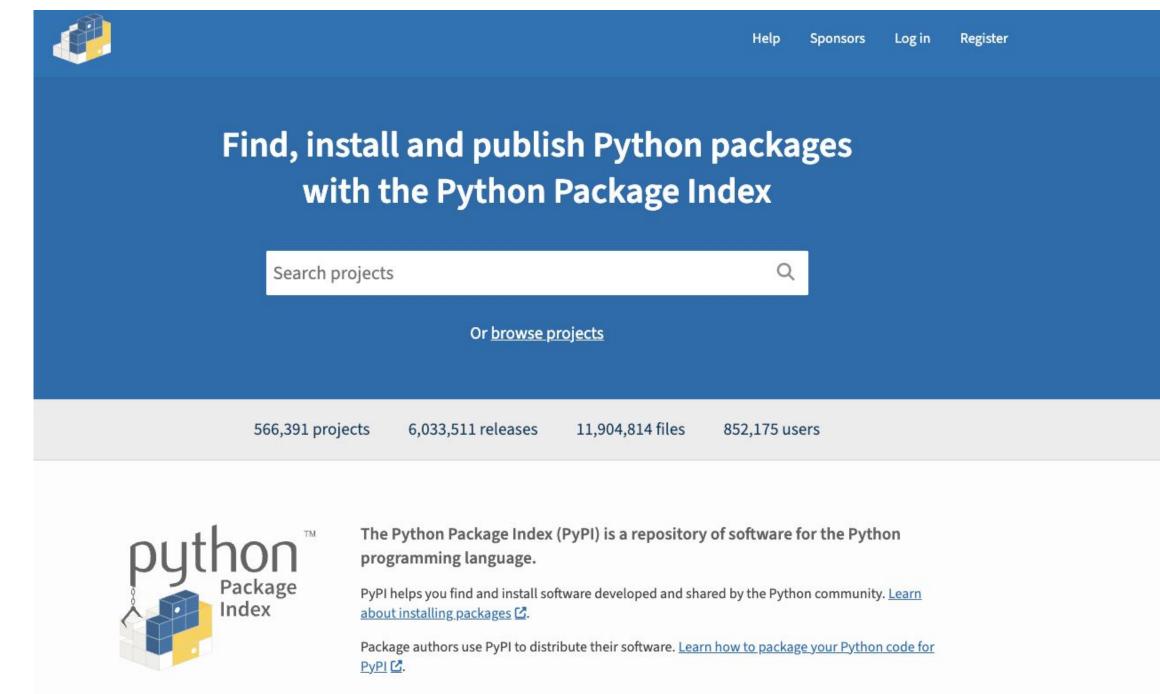
Relationship Between Distributions and PyPl

1. What is PyPI?

- Definition: The Python Package Index (PyPI) is the official repository where Python packages
 are uploaded and shared with the Python community.
- Purpose: Enables easy distribution, installation, and management of Python packages via `pip`.

2. Role of Distributions on PyPI

- Source Distribution (`.tar.gz`)
 - Purpose: Allows users to download and build the package from source.
 - Usage on PyPI: Uploaded to PyPI for environments that may require or prefer building from source, such as when specific platform optimizations or custom build steps are needed.
- Wheel File (`.whl`)
 - Purpose: Provides a pre-built, binary distribution of the package.
 - Usage on PyPI: Uploaded to PyPI for easy and fast installation across various environments without the need for compilation, making it the preferred distribution format for most users.



Uploading the Package to PyPl

1. Install `twine`

Command:



2. Upload to PyPI

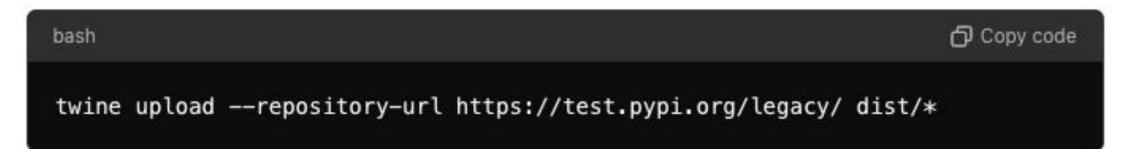
Command:



- · Explanation:
 - · Uploads the package files from the `dist/` directory to PyPI.
 - · You'll need your PyPI credentials.

3. Test on TestPyPI (Optional)

Command:



· Purpose: Test the upload process before uploading to the official PyPI.

Good Practice Define a Shell Script

```
upload_package() {
    echo "Removing old distribution..."
    rm -rf dist/*
    sleep 2 # Waits for 2 seconds

    echo "Building the distribution..."
    python setup.py sdist bdist_wheel
    sleep 2 # Waits for 2 seconds

    echo "Uploading the distribution..."
    pip install twine
    sleep 1
    twine upload dist/*.whl dist/*.tar.gz --verbose
}
```

• Availability: Once uploaded, the package becomes available to anyone using `pip`:

```
pip install package_name
```

• Example: Users can install your package directly from PyPI with a simple command.



Setup a Python Package

2. GitHub Repository

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

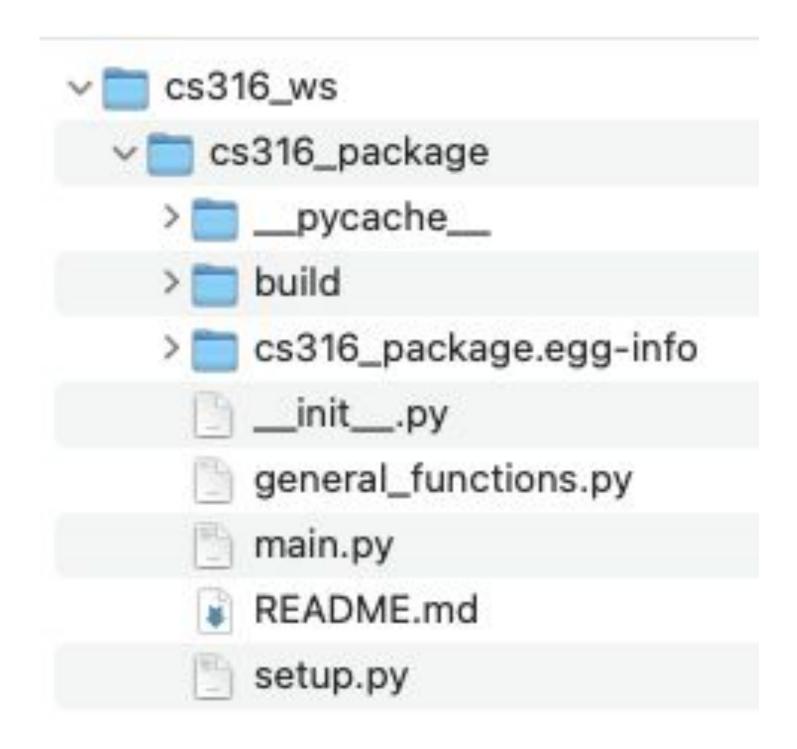
Setting Up a GitHub Account and Uploading a Python Package

1. Creating a GitHub Account

- Visit github.com.
- Click on "Sign up" and fill in your details.
- Verify your email to activate the account.
- Set up a new repository after logging in.

2. Setting Up a Python Package

- **Directory Structure:**
 - Root directory: `csi **
 - Package directory: `cs316_package/`
 - Essential files:
 - `__init__.py`: Marks the directory as a Python package.
 - `general_functions.py`: Contains general utility functions.
 - `main.py`: Main execution file.
 - `setup.py`: Used for package installation.
 - README.md`: Project description.



Setting Up a GitHub Account and Uploading a Python Package

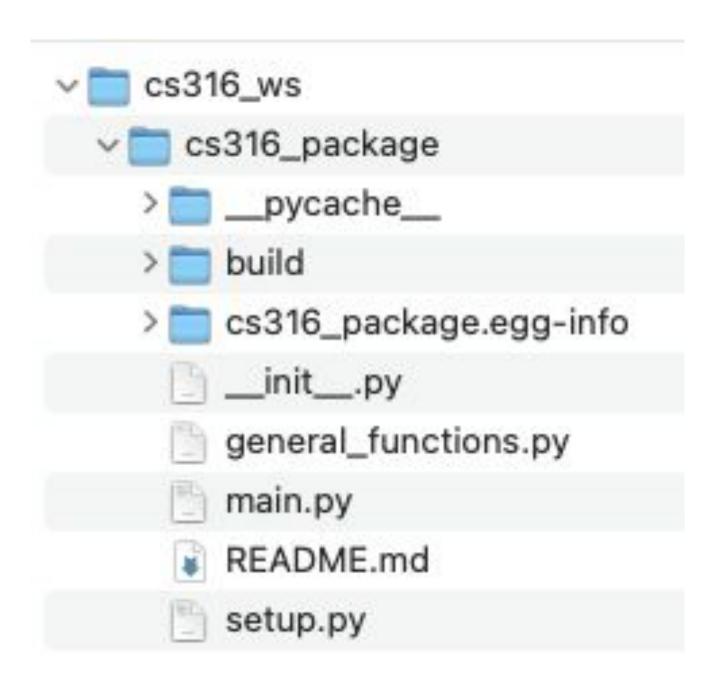
3. Uploading Package to GitHub

- · Initialize Local Repository:
 - `git init` inside `cs316_ws/`.
- · Add Files:
 - `git add .` to stage all files.
- Commit Changes:
 - `git commit -m "Initial commit"`
- · Push to GitHub:
 - Link repository:



Push changes:

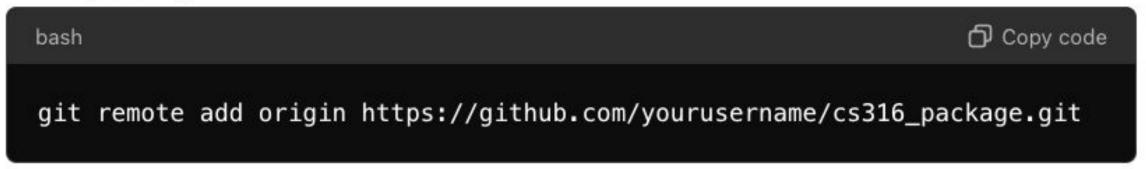




Setting Up a GitHub Account and Uploading a Python Package

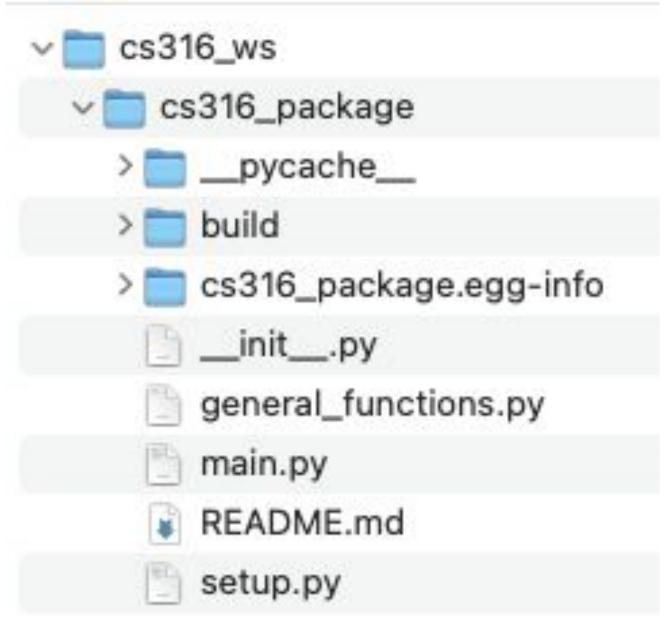
3. Uploading Package to GitHub

- · Initialize Local Repository:
 - `git init` inside `cs316_ws/`.
- · Add Files:
 - `git add .` to stage all files.
- Commit Changes:
 - `git commit -m "Initial commit"`
- · Push to GitHub:
 - · Link repository:



Push changes:





4. Verifying on GitHub

- Check the repository on GitHub to ensure all files are uploaded.
- Update `README.md` with project details.







CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Python Programming Language

Basic Syntax

Prof. Anis Koubaa

SEP 2024







CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Python Programming Language
Basic Syntax

Prof. Anis Koubaa

SEP 2024



Python Basic Syntax

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Python Basic Syntax

1. Introduction to Python Syntax

- Overview: Python syntax is designed to be simple and easy to read, making code more intuitive and accessible.
- Key Features:
 - Whitespace: Indentation is used to define code blocks (e.g., loops, functions).
 - Minimal Syntax: No need for semicolons or curly braces.
 - Readability: Emphasizes clear, human-readable code.

2. Code Example

```
python

# Define a function
def greet(name):
    print(f"Hello, {name}!")

# Call the function
greet("Alice")
```

- Explanation:
 - Indentation: The function body is indented to indicate it belongs to the `greet` function.
 - Function Call: The function is called with an argument (`"Alice"`), and the output is `Hello, Alice!`.

Basic Syntax Elements

1. Indentation

- Description: Python uses indentation to define blocks of code, unlike other languages that use braces `{}`.
- Standard Practice: Use four spaces per indentation level.
- Example:

```
Copy code
python
if x > 0:
    print("Positive number")
```

Explanation: The indented line belongs to the `if` statement block.

Basic Syntax Elements

2. Comments

- Description: Use the hash `#` symbol to write comments. Python ignores
 anything after `#` on a line.
- Usage: Comments are used for notes, documentation, and explanations within the code.
- Example:

```
# This is a comment

print("Hello, world!") # This prints a message
```

Basic Syntax Elements

3. Case Sensitivity

- Description: Python is case-sensitive, meaning that variables like `myVariable`, `MyVariable`, and `myvariable` are distinct.
- Example:

```
myVariable = 10
MyVariable = 20
print(myVariable) # Outputs: 10
print(MyVariable) # Outputs: 20
```

Explanation: Different capitalization creates different variables.



Python Variables and Data Types

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

LECTURE 1
Introduction to Python Programming

Variables and Data Types

1. Declaring Variables

- Description: Python automatically allocates memory when you assign a value to a variable no need for explicit declaration.
- Example:

```
python

x = 10  # Integer

y = 3.14  # Float

name = "Alice"  # String

is_valid = True  # Boolean
```

 Explanation: The variables `x`, `y`, `name`, and `is_valid` are automatically created with their respective data types based on the assigned values.

Variables and Data Types

2. Basic Data Types

- Integers: Whole numbers, positive or negative, without decimals.
 - Example: `x = 10`
- Floats: Decimal numbers, used for more precise calculations.
 - Example: `y = 3.14`
- Strings: Sequences of characters, enclosed in single or double quotes.
 - Example: `name = "Alice"`
- Booleans: Represent logical values, either `True` or `False`.
 - Example: `is_valid = True`

Non-Basic Python Data Types

Data Type	Description	Example	Characteristics
List	Ordered, mutable collection of items	`my_list = [1, 2, 3]`	Dynamic size, supports mixed data types, accessed by index
Tuple	Ordered, immutable collection of items	`my_tuple = (1, 2, 3)`	Fixed size, supports mixed data types, accessed by index
Set	Unordered collection of unique items	`my_set = {1, 2, 3}`	No duplicates, unordered, supports set operations
Dictionary	Unordered collection of key-value pairs	<pre>`my_dict = {'a': 1, 'b': 2}`</pre>	Keys must be unique, values accessed by keys
Complex	Numbers with real and imaginary parts	`z = 3 + 4j`	Precision depends on float representation for real/imaginary parts
Bytes	Immutable sequence of bytes	<pre>`my_bytes = b'hello'`</pre>	Used for binary data, immutable
Bytearray	Mutable sequence of bytes	<pre>`my_bytearray = bytearray(b'hello')`</pre>	Mutable version of `bytes`, can be modified in place
NoneType	Represents the absence of a value	`value = None`	Singleton, only one instance (`None`) exists

Finding the Type of a Variable with type()

1. Using `type()` Function

Purpose: `type()` is a built-in Python function used to determine the type of a variable.

2. Example Code

```
Copy code
python
# Define variables of different types
x = 10
                  # Integer
y = 3.14
                  # Float
name = "Alice"
                  # String
is_valid = True
                  # Boolean
# Use type() to find the type of each variable
print(type(x)) # Output: <class 'int'>
print(type(y)) # Output: <class 'float'>
print(type(name)) # Output: <class 'str'>
print(type(is_valid)) # Output: <class 'bool'>
```

3. Explanation

- The `type()` function returns the type of the variable passed as an argument.
- The output shows the class of the variable (`int`, `float`, `str`, `bool`), which corresponds to its data type.

Finding the Type of a Variable with type()

```
#String variable
        print ("--- STRING VARIABLES --- ")
        name ="Mohamed"
        print ("student name: ", name)
        print ("the type of the variable name is: ", type(name))
        #Boolean variable
        print ("--- BOOLEAN VARIABLES --- ")
        flag = True
        print ("Flag: ", flag)
        print ("the type of the variable flag is: ", type(flag))
    13
    14 #Hexadecimal Number
    15 print ("--- HEXADECIMAL VARIABLES --- ")
    16 \quad w = 0 \times fa2
    17 v=hex(547)
    18 print ("w: ", w)
    19 print ("v: ", v)
    20 print ("the type of the variable w is: ", type(w))
    21 print ("the type of the variable v is: ", type(v))
→ --- STRING VARIABLES ---
    student name: Mohamed
   the type of the variable name is: <class 'str'>
    --- BOOLEAN VARIABLES ---
   Flag: True
   the type of the variable flag is: <class 'bool'>
    --- HEXADECIMAL VARIABLES ---
    w: 4002
   v: 0x223
   the type of the variable w is: <class 'int'>
   the type of the variable v is: <class 'str'>
```



Type Conversion in Python

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Implicit Type Conversion

1. Implicit Type Conversion

- **Description**: Python automatically converts one data type to another when necessary.
- Example:

```
Copy code
python
           # Integer
x = 5
y = 3.2 # Float
z = x + y # Implicitly converts x to float
           # Output: 8.2 (float)
print(z)
```

Explanation: In the above example, the integer `x` is implicitly converted to a float during the addition.

Explicit Type Conversion

2. Explicit Type Conversion (Type Casting)

- Description: Converting one data type to another manually using built-in functions.
- Common Functions:
 - `int(x)`: Converts `x` to an integer.
 - `float(x)`: Converts `x` to a float.
 - `str(x)`: Converts `x` to a string.
 - 'bool(x)': Converts `x` to a boolean.

4. Considerations

- Loss of Data: Converting from a float to an integer truncates the decimal part.
- Errors: Converting incompatible types (e.g., `int("abc")`) will raise a `ValueError`.

3. Examples of Explicit Type Conversion

Integer to String:

```
python

num = 10
num_str = str(num)
print(num_str) # Output: "10" (string)
```

String to Integer:

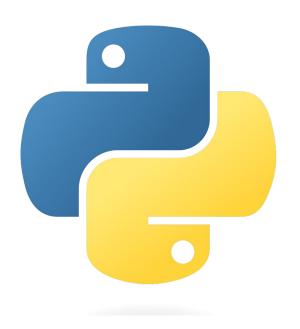
```
python

s = "123"
num = int(s)
print(num) # Output: 123 (integer)
```

Float to Integer:

```
python

pi = 3.14
int_pi = int(pi)
print(int_pi) # Output: 3 (integer, truncated)
```



Python Basic Operators

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Basic Operators in Python

Operator Type	Operator	Description	Example	Result
Arithmetic Operators	`+`	Addition	`5 + 3`	,8,
	·	Subtraction	`5 - 3`	`2`
	`*`	Multiplication	`5 * 3`	`15`
	٠١,	Division (returns a float)	`5 / 2`	`2.5`
	`%`	Modulus (remainder of division)	`5 % 2`	`1`
	`**`	Exponentiation (power)	`5 ** 3`	`125`
Comparison Operators	`==`	Equal to	`5 == 3`	`False`
	`!=`	Not equal to	`5 != 3`	`True`
	,>,	Greater than	`5 > 3`	`True`
	`<`	Less than	`5 < 3`	`False`
	`>=`	Greater than or equal to	`5 >= 3`	`True`
	`<=`	Less than or equal to	`5 <= 3`	`False`

Basic Operators in Python

```
a = 2
    b = 3
    add = a+b
 5 sub = a-b
   mult = a*b
    div = a/b
    mod = a%b
    e = a ** b
11
    print ("a = ",a)
    print ("b = ",b)
14 print ("add = a+b = ",add)
15 print ("sub = a-b = ", sub)
16 print ("mult = a*b = ", mult)
17 print ("div = a/b = ",div)
18 print ("mod = a%b = ", mod)
    print ("e = a**b = ",e)
20
21
a = 2
add = a+b = 5
sub = a-b = -1
mult = a*b = 6
mod = a\%b = 2
e = a**b = 8
```

Typing Constraints

- Type Compatibility: Python operators generally require operands to be of compatible types. For example:
 - Arithmetic Operators: Operands must typically be of numeric types (`int`, `float`), though some operators like `+` can also concatenate strings.
 - Example: "Hello" + " World" results in "Hello World"
 - Comparison Operators: Can compare numbers, strings, and other comparable types, but comparing incompatible types (e.g., `int` and `str`) will raise a `TypeError`.
 - Example: `5 > "3"` will result in a `TypeError`
- Implicit Type Conversion: Python may perform implicit type conversion (e.g., `int` to `float`) when necessary, but explicit conversion using functions like `int()`, `float()`, or `str()` is often required to avoid errors.



Simple Input/Output in Python

Prof. Anis Koubaa



INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Simple Input/Output in Python

1. Input: `input()` Function

- Purpose: Used to get user input from the keyboard.
- Syntax: `variable = input("Prompt message")`
- Example:

```
Copy code
python
name = input("Enter your name: ")
```

Explanation: The input provided by the user is stored as a string in the variable `name`.

Simple Input/Output in Python

2. Output: `print()` Function

- Purpose: Used to display output to the screen.
- Syntax: `print(value1, value2, ..., sep=' ', end='\n')`
- Example:

```
Copy code
python
print("Hello, World!")
print("Welcome,", name)
```

- Explanation:
 - The first `print()` displays the message `"Hello, World!"`.
 - The second `print()` uses the variable `name` to display a personalized welcome message.

Simple Input/Output in Python

3. Combined Example

```
# Getting user input
age = input("Enter your age: ")

# Displaying the input
print("You are", age, "years old.")
```

 Explanation: This code prompts the user for their age and then displays it back to them in a complete sentence.



Formatted Output in Python

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

1. Basic `print()`

- Description: Displays one or more values separated by a space (default).
- Example:



Output:

`Hello World!`

2. Using `sep` Parameter

- Description: Customizes the separator between multiple values.
- Example:

```
Copy code
python
print("Hello", "World", sep="-")
```

Output:

`Hello-World`

3. Using `end` Parameter

- **Description**: Customizes what is printed at the end of the output (default is a newline `\n`).
- Example:

```
Copy code
python
print("Hello", end=", ")
print("World!")
```

Output:

`Hello, World!`

4. String Formatting with `%` Operator

- Description: Uses placeholders (`%s`, `%d`, `%f`) to embed values in a string.
- Example:

```
Copy code
python
name = "Alice"
age = 30
print("My name is %s and I am %d years old." % (name, age))
```

Output:

`My name is Alice and I am 30 years old.`

5. `str.format()` Method

- Description: A more versatile way to format strings, using curly braces `{}` as placeholders.
- Example:

```
python

name = "Bob"
score = 95.5
print("Name: {}, Score: {:.1f}".format(name, score))
```

Output:

`Name: Bob, Score: 95.5`

6. f-Strings (Python 3.6+)

- Description: An easy and efficient way to embed expressions inside string literals using `{}` with an `f` prefix.
- Example:

```
Copy code
python
name = "Charlie"
height = 1.75
print(f"{name} is {height} meters tall.")
```

Output:

`Charlie is 1.75 meters tall.`



Math Functions in Python

Prof. Anis Koubaa



CS316 INTRODUCTION TO AI AND DATA SCIENCE

CHAPTER 2
Python for Data Science

Common Mathematical Functions in Python

Function	Description	Example	Result
`abs(x)`	Returns the absolute value of `x`	`abs(-5)`	`5`
`round(x, n)`	Rounds `x` to `n` decimal places	`round(3.14159, 2)`	`3.14`
`max(x1, x2,)`	Returns the largest value among arguments	`max(3, 7, 1)`	`7`
`min(x1, x2,)`	Returns the smallest value among arguments	`min(3, 7, 1)`	`1`
`pow(x, y)`	Returns `x` raised to the power of `y`	`pow(2, 3)`	`8`
`sum(iterable)`	Returns the sum of all items in an iterable	`sum([1, 2, 3, 4])`	`10`
`math.sqrt(x)`	Returns the square root of `x`	`math.sqrt(16)`	`4.0`
`math.exp(x)`	Returns `e` raised to the power of `x`	`math.exp(1)`	`2.718281828459045
`math.log(x, base)`	Returns the logarithm of `x` to the specified base (default is `e`)	`math.log(8, 2)`	`3.0`
`math.sin(x)`	Returns the sine of `x` (in radians)	`math.sin(math.pi / 2)`	`1.0`
`math.cos(x)`	Returns the cosine of `x` (in radians)	`math.cos(0)`	`1.0`

Basic Operators in Python

Operator Type	Operator	Description	Example	Result
Arithmetic Operators	`+`	Addition	`5 + 3`	,8,
	·	Subtraction	`5 - 3`	`2`
	`*`	Multiplication	`5 * 3`	`15`
	٠١,	Division (returns a float)	`5 / 2`	`2.5`
	`%`	Modulus (remainder of division)	`5 % 2`	`1`
	`**`	Exponentiation (power)	`5 ** 3`	`125`
Comparison Operators	`==`	Equal to	`5 == 3`	`False`
	`!=`	Not equal to	`5 != 3`	`True`
	,>,	Greater than	`5 > 3`	`True`
	`<`	Less than	`5 < 3`	`False`
	`>=`	Greater than or equal to	`5 >= 3`	`True`
	`<=`	Less than or equal to	`5 <= 3`	`False`

Common Mathematical Functions in Python

```
1 x = -2.56
 y = 3.54
    n = 3
    from math import *
    print ('the absolute value of x is : ', abs(x)) #we can use abs or fabs, math.fak
    print ('----')
    print ('the ceiling of x is : ', ceil(x))
    print ('the floor of x is : ', floor(x))
    print ('the ceiling of y is : ', ceil(y))
    print ('the floor of y is : ', floor(y))
    print ('----')
    print ('the log of y is: ', log(y)) #log only applies for positive values
    print ('the log10 of x is : ', log10(y)) #log10 only applies for positive values
    print ('the exponential of y is : ', exp(y))
    print ('the square root of y is : ', sqrt(y))
    print ('the power of x by y is : ', pow(x,n))
the absolute value of x is: 2.56
the ceiling of x is : -2
the floor of x is: -3
the ceiling of y is: 4
the floor of y is: 3
the log of y is: 1.264126727145683
the log10 of x is: 0.5490032620257879
the exponential of y is: 34.46691919085739
the square root of y is: 1.8814887722226779
the power of x by y is : -16.777216000000003
```

END OF LECTURE