

Big Data Analytics

02: Introduction to Python

Instructor: Oleh Tymchuk

#02: Agenda

- What is Python?
- Python Interpreter
- IDEs (Jupyter Notebook, Google Colab)
- Python Practice

What is Python?

What is Python?

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics.

Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance.

Python supports modules and packages, which encourages program modularity and code reuse.

The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

Main areas where Python scripts are used

- Data analysis and machine learning
- Web and Internet Development
- Automation / Scripting
- Software testing and prototyping
- Everyday tasks

TIOBE Index / Feb 2025

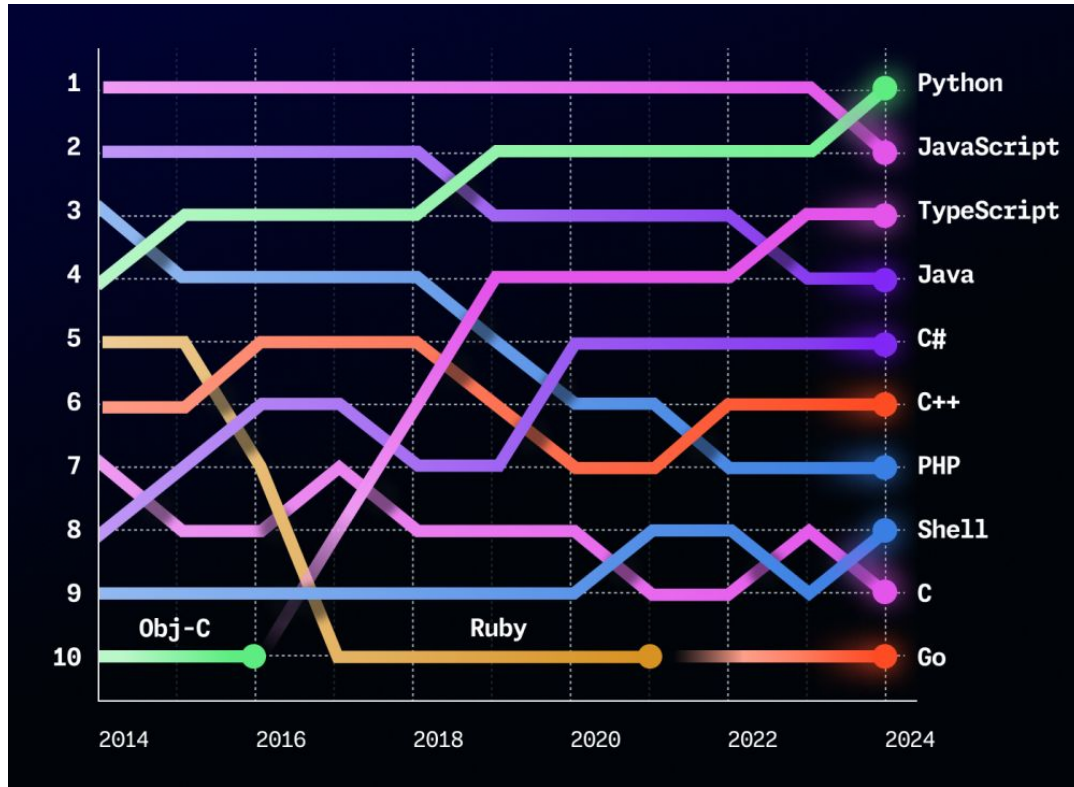
Feb 2025	Feb 2024	Change	Programming Language	
1	1			Python
2	3	▲		C++
3	4	▲		Java
4	2	▼		C
5	5			C#
6	6			JavaScript
7	7			SQL
8	8			Go
9	12	▲		Delphi/Object Pascal
10	9	▼		Visual Basic

<https://www.tiobe.com/tiobe-index/>

TIOBE Index / Very Long Term History

Programming Language	2025	2020	2015	2010	2005	2000	1995	1990	1985
Python	1	3	7	7	7	24	23	-	-
C++	2	4	4	4	3	2	1	2	13
C	3	2	1	2	1	1	2	1	1
Java	4	1	2	1	2	3	-	-	-
C#	5	5	5	6	9	9	-	-	-
JavaScript	6	7	8	9	10	7	-	-	-
Go	7	16	36	184	-	-	-	-	-
Visual Basic	8	19	234	-	-	-	-	-	-
SQL	9	9	-	-	100	-	-	-	-
Fortran	10	30	31	25	15	18	5	8	11
PHP	13	8	6	3	5	29	-	-	-
Ada	25	35	30	26	16	17	7	4	3
Lisp	28	31	18	17	14	16	6	3	2
Objective-C	35	10	3	23	40	-	-	-	-
(Visual) Basic	-	-	77	5	4	4	3	5	4

GitHub Octoverse / 2024



<https://github.blog/news-insights/octoverse/octoverse-2024/>

Top Python Libraries for Data Analysis



Pandas

Data analysis and manipulation tool



Scrapy

Web scraping and data extraction framework



NumPy

Numerical computing and array processing library



Scikit-learn

Machine learning library



Matplotlib

Data visualization and plotting library



Tensorflow

Framework for machine learning and deep learning model development



SeaBorn

Statistical data visualization library

NLTK

Natural Language Toolkit

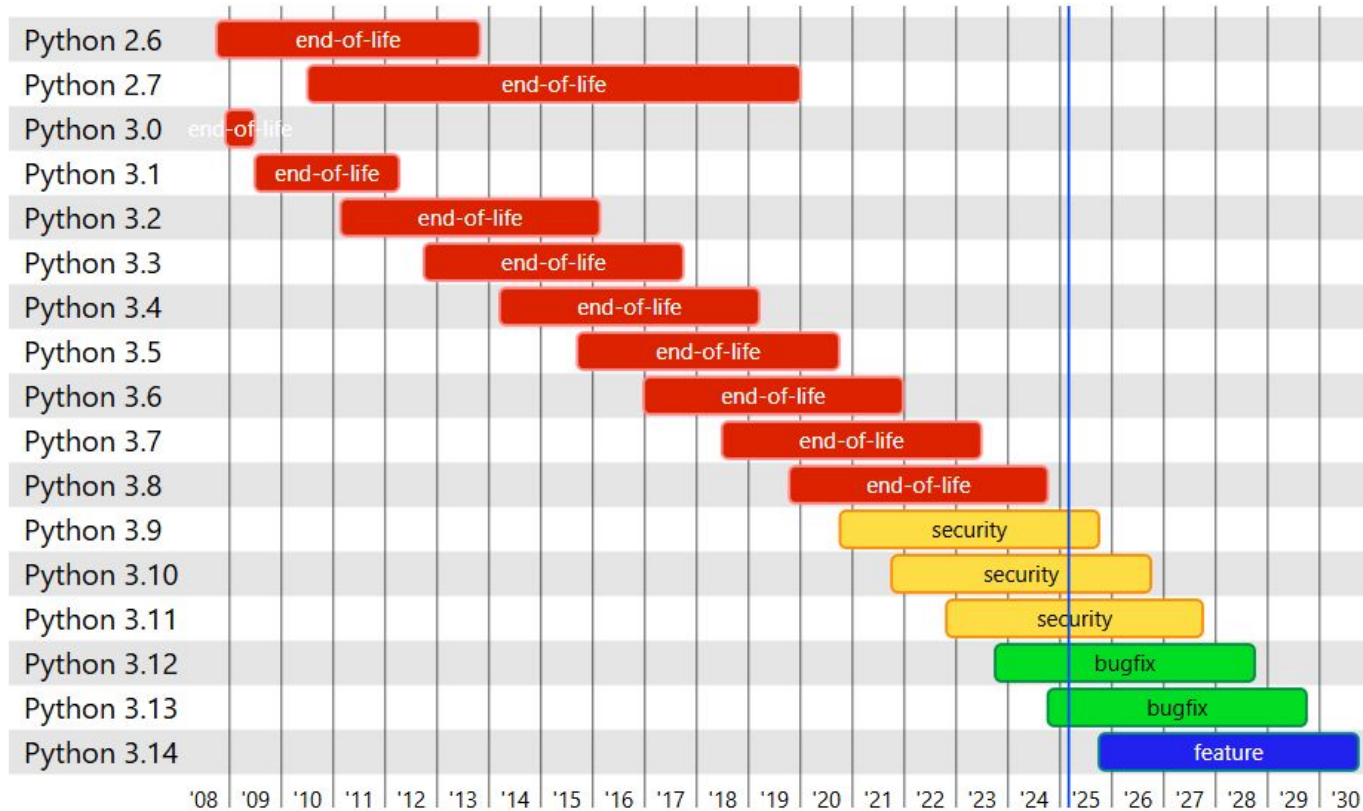
Library for working with human language data

Python Interpreter

Python Releases

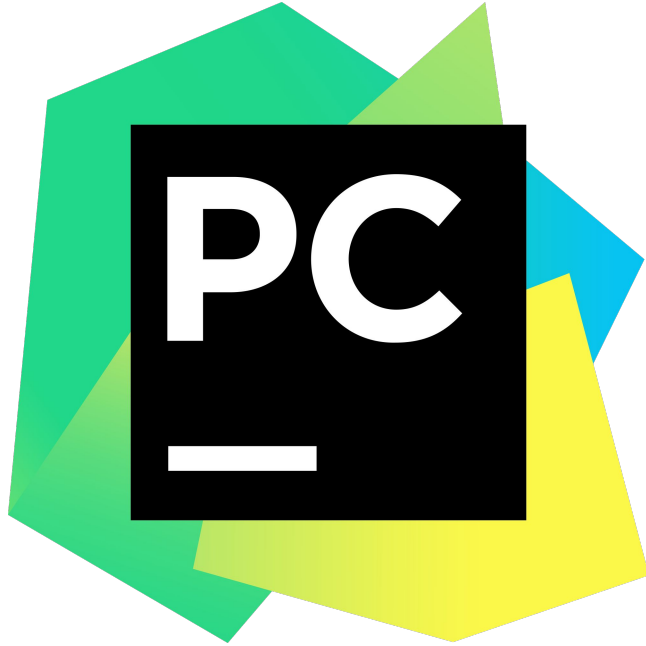
Python version	Maintenance status	First released	End of support
3.14	pre-release	2025-10-01 (planned)	2030-10
3.13	bugfix	2024-10-07	2029-10
3.12	bugfix	2023-10-02	2028-10
3.11	security	2022-10-24	2027-10
3.10	security	2021-10-04	2026-10
3.9	security	2020-10-05	2025-10

Status of Python versions

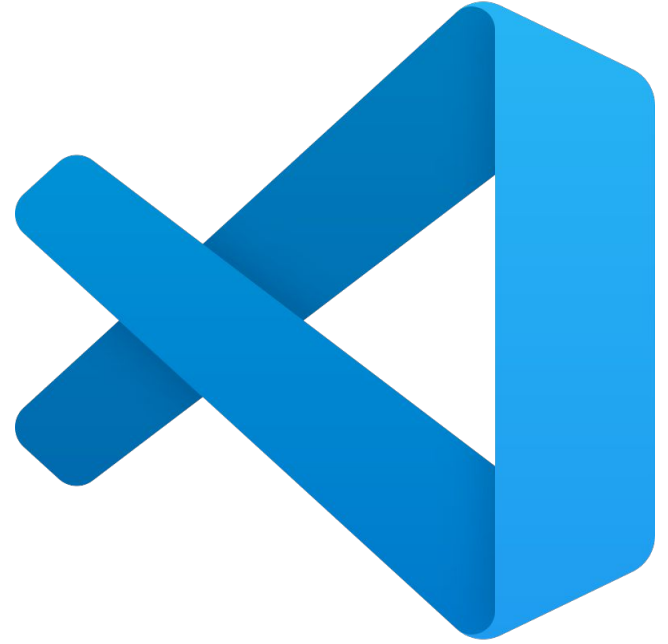


IDEs

For Developers



[PyCharm](#)



[Visual Studio Code](#)

For Data Analysts



[Jupyter](#)



[Google Colab](#)

Using Google Colab and Saving Results to GitHub

1. Open Google Colab

Go to [Google Colab](#).

You can log in with your Google account.

2. Create or Open a Notebook

To create a new notebook, click on File > New notebook.

To open an existing notebook, click on File > Open notebook and choose the file from:

Google Drive, GitHub, or upload from your local computer.

3. Writing and Running Code

In each code cell, write your Python code and press Shift + Enter to run it.

You can also add text cells for explanations or markdown.

Python Practice

Useful Links

Python

<https://docs.python.org/3/>

<https://peps.python.org/pep-0008/>

IDEs

<https://jupyter.org/>

<https://colab.research.google.com/>

Others...

https://colab.research.google.com/notebooks/basic_features_overview.ipynb

https://colab.research.google.com/notebooks/markdown_guide.ipynb