Introduction to Python

Columbia Biostatistics Computing Club
Fall 2025

Originally By: Amy Pitts and Muhire Honorine Kwizera
Adapted By: Kate Colvin

What is Python?

- In short it is another programming language.
  - We have talked a lot about R in the last couple of meetings which is a different programming language. You can do basically everything that you do in R with python.

- Python can be used for more than just data cleaning, statistics, visualizations
  - It can be used for web development
  - It can be used for desktop/app development
  - It is a general purpose programming language!
  - Machine Learning!

Why learn Python?

According to the TIOBE index Python is the most used programming language this year and last year! (up from #2 in 2020)

Oct 2025	Oct 2024	Change	Programming Language	Ratings	Change
1	1		Python	24.45%	+2.55%
2	4	↑	C	9.29%	+0.91%
3	2	↓	C++	8.84%	-2.77%
4	3	↓	Java	8.35%	-2.15%
5	5		C#	6.94%	+1.32%
13	17	⇈	R	1.52%	+0.43%

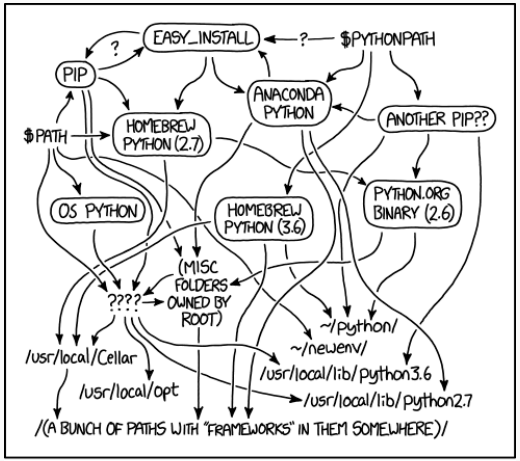
Source: https://www.tiobe.com/tiobe-index/

Python Compared to R

- One is not better than the other!
  - It just all depend on what you are trying to do and what job you might want after the program you are in now!
- R is much stronger for most statistical computing
  - More/better packages available for statistics and data visualization
  - Widely used for academic research across different disciplines
- Python has a much larger user base and is more versatile
  - Most tech companies use Python
  - Python has a **vast** collection of libraries for everything from image processing, language translation, music generation, etc.
  - Python is better for larger scale programming (>1000 lines) (easier object-oriented programming)
  - Nearly all AI/ML work is done in Python
- It is good to be familiar with both languages!

Where to Write/Run Python Code
(just a few examples)

- Using a text editor and running python through your terminal
  - There are a lot of types of text editors (**Atom, Vim, Visual Studio Code, Notepad++**)
- **Jupyter Notebook**
  - https://jupyter.org/
  - Similar to R Studio (However, need to download more things).
- **Google Colab Notebook**
  - There is an example of this on the computing club website in resources.
    - https://colab.research.google.com/notebooks/intro.ipynb
  - This is the most like R Studio, no need to download anything!
  - We will also be using Google Colab in our coding demonstration.

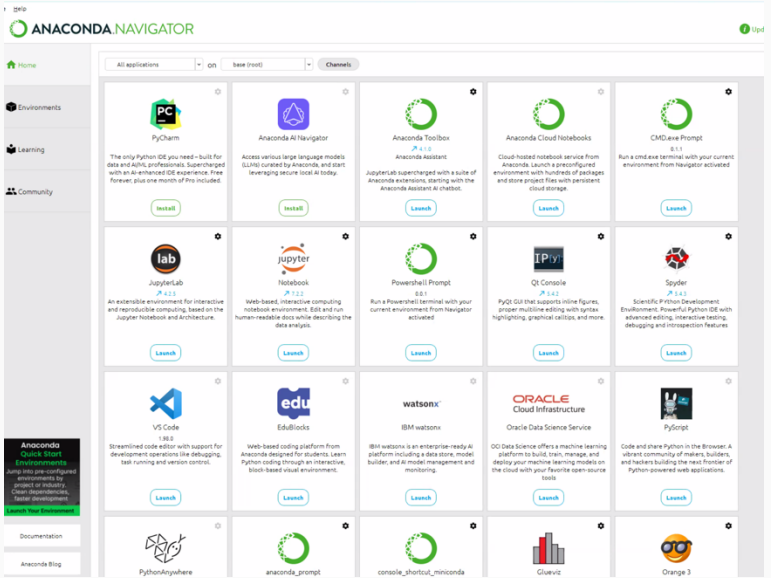


MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

Many of these (Python, VS Code, Jupyter Notebook), are bundled in one free download by Anaconda (anaconda.com/download)

How to Download Python

- A lot of computers come with Python, but they are likely out of date and missing libraries you may need
- If you're new to Python, I would **download it through Anaconda**: anaconda.com/download
  - Python
  - Environments such as VS Code, Jupyter Notebook, Spyder, PyCharm
  - **Many** critical libraries (including Pandas, NumPy, SciPy, Matplotlib, Seaborn, Scikit-learn etc.)
- Here is how to check your python version: https://phoenixnap.com/kb/check-python-version
- Download Python directly: https://wiki.python.org/moin/BeginnersGuide/Download



Common Libraries

(Just like R we need to install and call libraries)

- numpy -> library for some optimized data structures
  - `import numpy as np`
- pandas -> library for nice data frames (like tibbles)
  - `import pandas as pd`
  - https://pandas.pydata.org/
- seaborn -> library for plotting
  - `import seaborn as sns`
- sklearn -> library popular for machine learning
  - `import sklearn as sk`
  - https://scikit-learn.org/stable/index.html

List of Python Tutorials:

- TutorialsPoint
  - https://www.tutorialspoint.com/python/index.htm
  - This is a great online resource! (Runs python through terminal)
- Keras
  - https://keras.io/about/
  - This has a lot of code examples. It is an open source machine learning platform!
  - Intro for researchers: https://keras.io/getting_started/intro_to_keras_for_researchers/
- Python Tutorial
  - https://www.learnpython.org/
- DataCamp
  - datacamp.com/?utm_source=learnpython_com&utm_campaign=learnpython_tutorials
- A Video Tutorial
  - https://www.youtube.com/watch?v=rfscVS0vtbw

Additional Resources:

Introduction to Python using practical applications:

- https://automatetheboringstuff.com/

Learning platforms with free basic Python classes

- https://www.udemy.com/topic/python/free/
- https://www.learnpython.org/
- https://www.udacity.com/course/introduction-to-python--ud1110

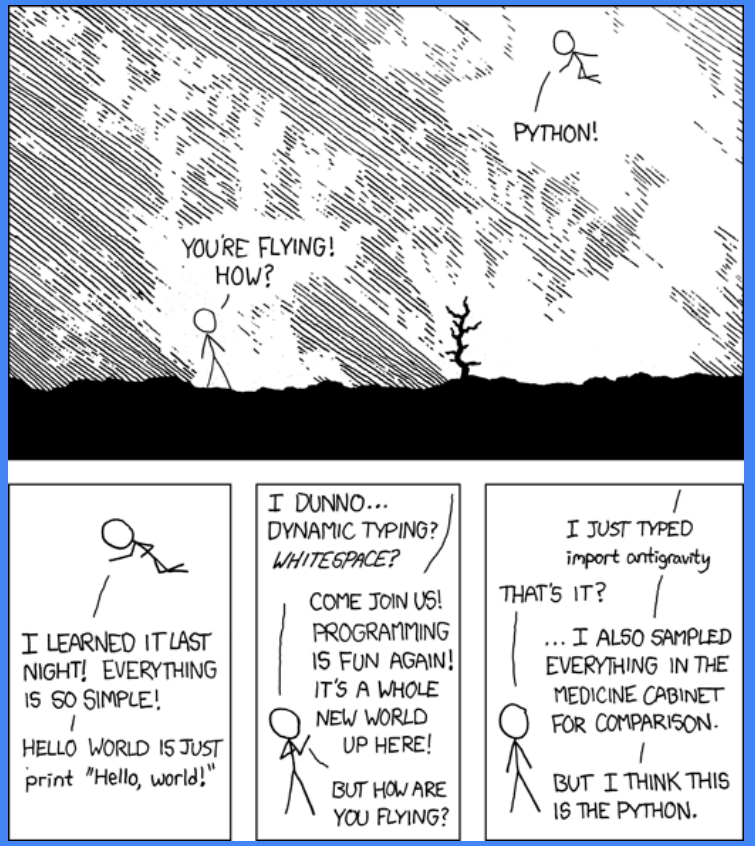
Free online Python books and lists of free Python books:

- https://pythonbooks.org/free-books/
- https://github.com/pamoroso/free-python-books/blob/master/README.md
- https://python.swaroopch.com/
- https://thepythoncodingbook.com/

Lists of Online Tutorials:

- https://wiki.python.org/moin/BeginnersGuide/Programmers
- https://gitconnected.com/learn/python
- https://hackr.io/tutorials/learn-python

Questions before the Demo?



What is a Terminal?

- This in itself could be an entire computing club presentation!
- In simple terms it is a command line system that allows you quick access to all files, the ability to run those files, and control your operating system.
  - There is a lot of things you can do with your terminal!
- To run a file you edited in a text editor you will need to go to that file's directory and then run it using python.
  - Example command: python3 hello.py
  - Tutorial: https://www.datacamp.com/community/tutorials/running-a-python-script