

Introduction to Python

Sugarkhuu Radnaa

Py4Econ in Ulaanbaatar

py4econ@gmail.com

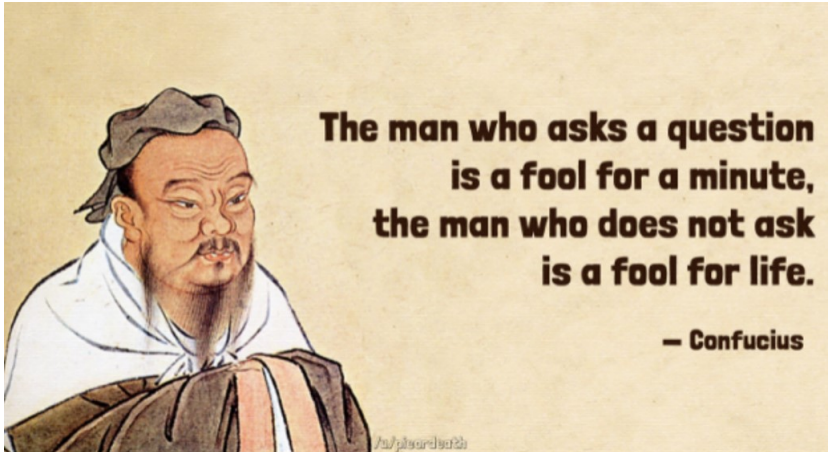
10 декабря 2021 г.

About the course

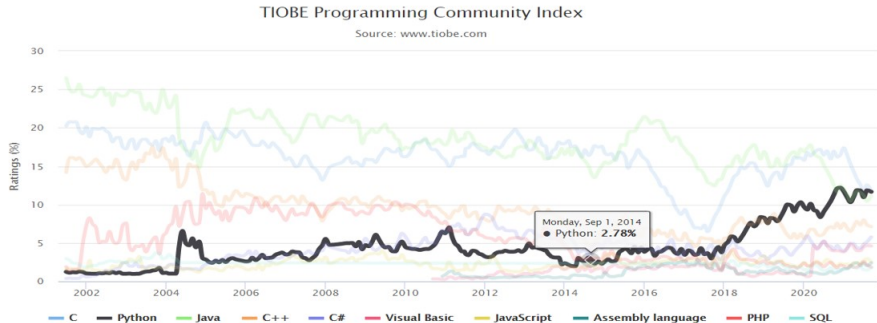
- ① Basic programming know-hows
- ② Elementary to Intermediate Python
 - Python topics covered: data structures, control flow, functions, modules, classes, files, regex, automation, data visualization, web scraping
 - Python topics not covered: data analysis, web development, ML AI
- ③ Final projects:
 - Excel manipulation
 - Data cleaning and visualization
 - Web scraping sending automated email

Week 1: Learning objectives

- 1 Overview of Python
- 2 Computer basics
- 3 Communities
- 4 Folder structures, Path, Shell
- 5 Python / Anaconda
- 6 Text editor, notebook, IDE
- 7 Git Github
- 8 Coding best practices



Why Python?



+

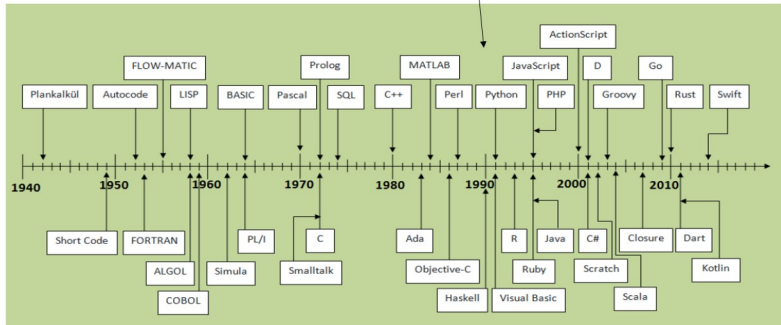
Beginner friendly

Popular applications of Python

- 1 Data science and data visualization
- 2 Machine learning AI
- 3 Scientific computing (incl. Financial modelling)
- 4 Web Game development
- 5 Desktop applications Software GUI
- 6 Automation

Python was first released in 1991

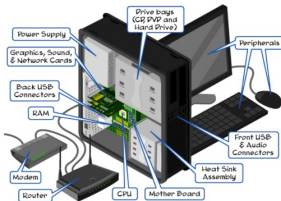
Timeline Of Programming Languages :



Source: <https://javaconceptsoftheday.com/history-of-programming-languages/>

Computer basics

Processor

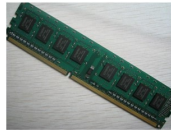


Data units

UNIT	ABBREVIATION	STORAGE
Bit	b	Binary Digit, Single 1 or 0
Nibble	-	4 bits
Byte/Octet	B	8 bits
Kilobyte	KB	1024 bytes
Megabyte	MB	1024 KB
Gigabyte	GB	1024 MB
Terabyte	TB	1024 GB
Petabyte	PB	1024 TB
Exabyte	EB	1024 PB
Zettabyte	ZB	1024 EB
Yottabyte	YB	1024 ZB

Storage units (www.byte-notes.com)

RAM



Hard disk



CPU (chip)



Operating systems

Mac, Linux,
Windows



Communities and learning platforms

- 1 Python - Official
- 2 Stack overflow - Forum
- 3 Medium - Blog
- 4 Towardsdatascience - Blog
- 5 Tutorialspoint
- 6 Geek for geeks
- 7 W3schools
- 8 Real Python
- 9 Programiz
- 10 Kaggle - Competition and Learning resource
- 11 DrivenData
- 12 TopCoder
- 13 DataHack

What is a programming language?

A programming language is a formal language comprising a set of strings that produce various kinds of machine code output.

Wikipedia

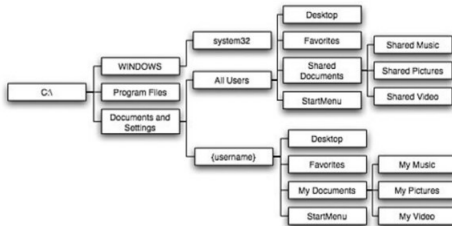
```
1  
2  
3 print("Hello Py4Econ!")  
4  
5
```

Must-have basics for a good programmer

- Data Structure and Algorithm
- A Version Control Tool (Git)
- One Text Editors (VScode)
- IDEs (Spyder or Pycharm)
- Database and SQL
- UNIX (Linux)
- An OOP Programming language (C++, Java or Python)
- One Scripting language (automation)
- Networking basics
- *Cloud Platform (AWS, GCP, or Azure)*
- *Containers (Docker and Kubernetes)*

Folder structure

Figure 4.7. Windows XP file structure.



Copy path:

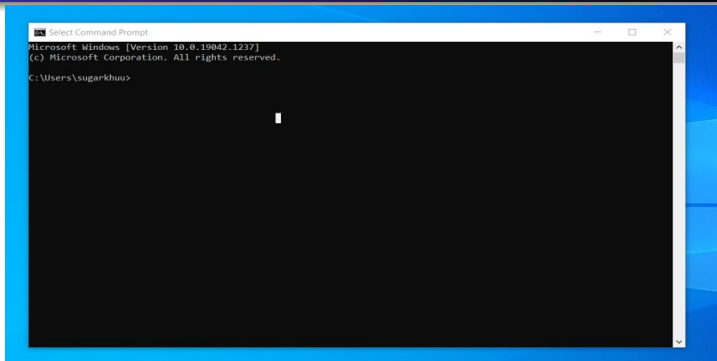
- Ctrl + L = folder path
- Shift + Right mouse > copy as path (a) = file path

When naming folder & file:

- Avoid spaces and uncommon characters
- Use either camel or snake cases

C:\Documents and Settings\sugarkhuu\My Documents\My Video

Command line interpreters/Shells



You are able to control your computer through commanding the OS from terminals. More powerful and flexible than usual GUI way of doing things

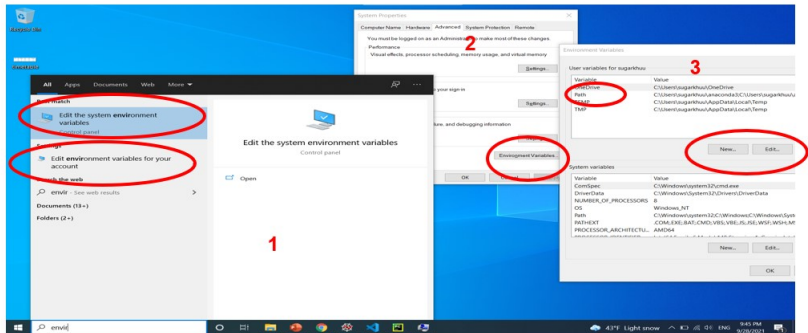
- Windows: Command prompt, Powershell. More recently, Windows Terminal
- Linux: Bash
- Mac: Terminal (zsh)

Common CMD/Bash commands

- `cd [cd]` – change directory. `/?` - help
- `dir [ls]` – directory content
- `copy [cp]` – copy file
- `ren [mv]` – rename file
- `del [rm]` – delete file
- `mkdir [mkdir]` – create new folder
- `exit [exit]` - close terminal
- `cls [ctrl+L]` – clear terminal

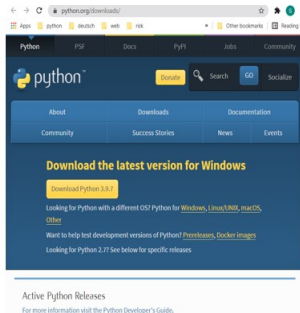
Path (environment variable)

The PATH variable makes it easy to run commonly used programs located in their own folders.



Getting Python

Standalone Python

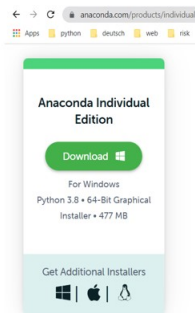


<https://www.python.org/downloads/>

Pre-installed all
packages useful
for data science



Anaconda (incl. python)



<https://www.anaconda.com/products/individual>
<https://www.datacamp.com/community/tutorials/installing-anaconda-windows>

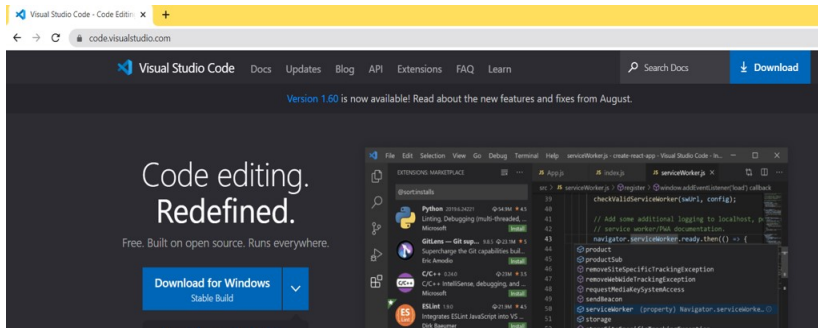
Possible to use Python in many environments

- 1 Terminal
- 2 IDEs – Spyder or Pycharm (IntelliJ), VScode
- 3 Notebook - (Jupyter notebook)

Python basic concepts for today:

- Basic syntax
- Basic operators
- Packages (+ pip)
- Docstring
- Read data

Using VS code (code editor) for Python



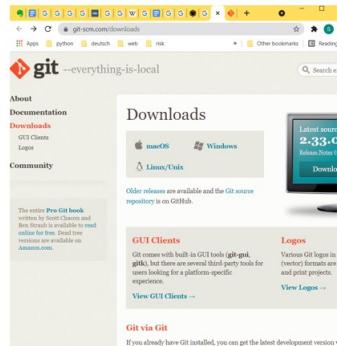
Git – Version control system (VCS)

Version control system (VCS):

Version control systems are a category of software tools that helps in recording changes made to files by keeping a track of modifications done to the code.

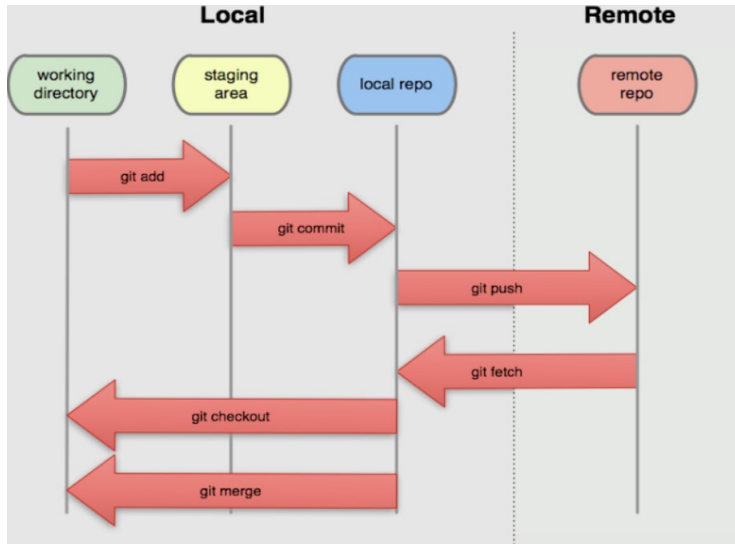
<https://www.geeksforgeeks.org/version-control-systems/>

Popular VCSs: Git, Subversion, Helix core, Microsoft TFS



<https://git-scm.com/downloads>

Git – Version control system (VCS)



Github – Share everything you want

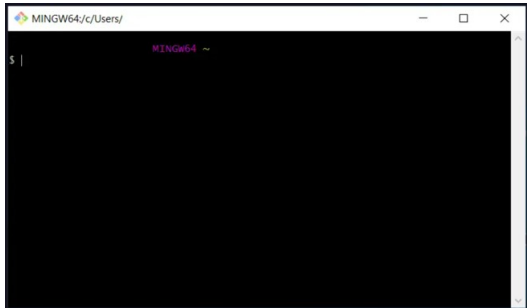
The screenshot shows the GitHub repository page for 'TheAlgorithms/Python'. The repository is public and has 15 branches and 0 tags. The commit history table lists the following entries:

Commit	Message	Time
Joe-Sin7h Fixed #4764 (#4779)	✓ 5482183 4 days ago 2,180 commits	
github	mypy --install-types --non-interactive . (#4530)	6 days ago
arithmetic_analysis	Pyupgrade to Python 3.9 (#4718)	21 days ago
backtracking	Pyupgrade to Python 3.9 (#4718)	21 days ago
bit_manipulation	Bit manipulation: get the bit at a given position (#4438)	4 months ago
blockchain	Pyupgrade to Python 3.9 (#4718)	21 days ago
boolean_algebra	Pyupgrade to Python 3.9 (#4718)	21 days ago
cellular_automata	Pyupgrade to Python 3.9 (#4718)	21 days ago
ciphers	from __future__ import annotations (#4763)	6 days ago
compression	Pyupgrade to Python 3.9 (#4718)	21 days ago
computer_vision	feat: CNN classification added to computer vision (#4350)	3 months ago

On the right, the 'About' section states 'All Algorithms implemented in Python' and lists tags: python, education, algorithm, practice, interview, sorting-algorithms, learn, algo, algorithm-competitions, sorts, hacktoberfest, algorithms-implemented, community-driven, searches. Below this is a 'Releases' section with the text 'No releases published'.

<https://code.visualstudio.com/>

A unix based commands on Windows (Emulator)



Coding best practices

- Description of the code
- Comments
- Simple and short variable name (convention)
- Indentation
- Commit little by little
- Logically consistent and well tested ...

Homework

① Task 1

② Task 2

Deadline: 18 December, 2021

Task 1

- 1 Create a new repository in your Github
- 2 Clone the previous repository to your local machine
- 3 Create a Jupyter notebook in the local repo (in your folder)
- 4 In the notebook, write a code which asks 5 separate questions and receives the answers from the user (we had an example with only one question in the lecture)
- 5 Commit and push your changes question by question one at a time
- 6 Make sure your code follows the good practice

Task 2

- 1 Нэг компьютерт хоёр үйлдлийн систем суулгаж болох уу?
- 2 Path-д программаа оруулаагүй бол яах вэ?
- 3 Фолдерийн нэр нь дундаа зайтай байвал фолдерийг танихад ямар асуудал үүсэх вэ?
- 4 Git, Github хоёрын ялгаа юу вэ?
- 5 Jupyter notebook IDE мөн үү?
- 6 Commit, push хоёрын ялгаа юу вэ?
- 7 Push хийхээс өмнө олон дахин commit хийж болох уу?
- 8 Commit хийхэд github repo-д access хэрэгтэй юу?

Thank you!