

Python Applications

Py4Econ in Ulaanbaatar, 2024

py4econ@gmail.com

Welcome everyone!

About the course

1. Basic programming know-hows
2. Elementary to Intermediate Python
 - Supplementary materials:
 - Python specific: Data basics, Functions/Class, Conditional/Loop
 - General programming: Command line interface (CLI), Code editor (VScode), Git/Github
 - **Main topics:** Data cleaning, SQL, Visualization/Dashboard, Webscraping, Machine learning (ML/DL), Automation (DevOps)

Organization

- Lecture sessions: 3 hours at 17 p.m on Saturday, 8 sessions through 10 Aug - 29 Sep.
 - At the same zoom link emailed to all of you.
- Supplementary materials will be uploaded on Monday. To be studied by the lecture on Saturday. I will keep you informed in due course.
- Homework (practice problems after each session). To be submitted by the next session via Github repo.
- Every session will be recorded fully and be shared on a Google drive, to stay there until Google goes bankrupt. All codes will be shared via Github. Facebook chat group?
- A Zoom session lasts 40 minutes. We regather after 5 minutes.
- Asking questions is highly encouraged!

Let's get to know each other :)

- I am
 - Your instructor: This course since Oct 2021; Python online contents since 2018 (Py4Econ Official YouTube channel).
 - Education: Bach. in Economics, IFE Mongolia, Master in Finance, Czechia
 - 4 years of banking experience in Mongolia
 - 2 years of experience in Macroeconomic consultancy in Czechia
 - Currently pursuing a Ph.D in Macroeconomics at the Bonn Graduate School of Economics, Germany
 - Data is everywhere, couldn't escape. Love coding.

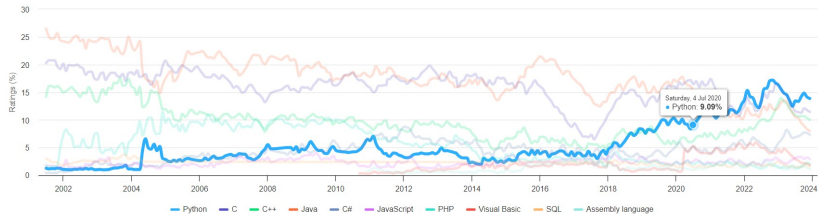
Week 1: Learning objectives

1. Background information
2. Computer basics (Folder structures, Path, Shell)
3. Python /Anaconda/
4. VScode, Jupyter notebook and other IDEs
5. Git & Github

Why Python?

TIOBE Programming Community Index

Source: www.tiobe.com

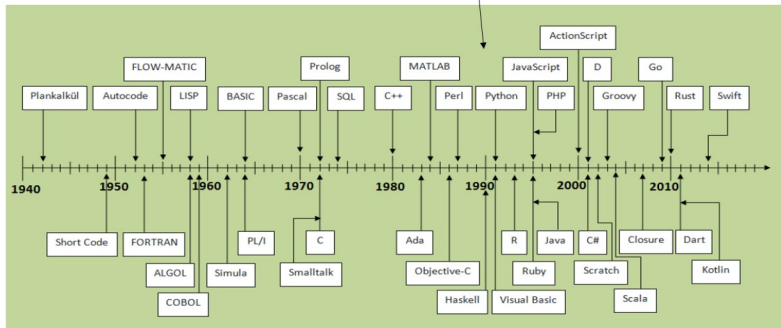


Popular applications of Python

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

Python was first released in 1991

Timeline Of Programming Languages :



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

Guide: 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)*

Communities and learning platforms

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

ChatGPT and Github Copilot

chatgpt.com/c/0924f306-063a-4bf1-85de-b17aebbee18c


 ChatGPT 3.5 ▾

 You

a python code for a logistic regression model to predict house price

 ChatGPT

Sure, here's a simple example of Python code using scikit-learn to implement a logistic regression model to predict house prices. However, logistic regression is typically used for binary classification tasks, such as predicting whether an email is spam or not. For house price prediction, you might want to consider using linear regression instead.

```
python  Copy code
```

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score

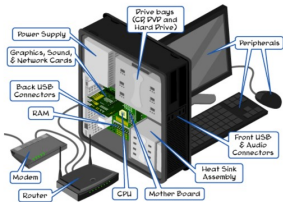
# Load your dataset
# Assuming your dataset has features (X) and target variable (y)
# X should be a DataFrame containing features like area, number of bedrooms, etc.
# y should be a Series containing the target variable, i.e., house prices
# Replace 'features.csv' and 'target.csv' with your actual file names or data sources
X = pd.read_csv('features.csv')
```

Message ChatGPT

ChatGPT can make mistakes. Check important info.

Computer basics

Processor

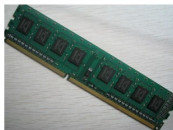


Data units

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

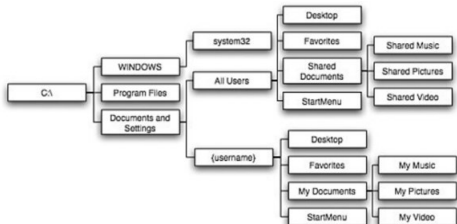
Mac, Linux,
Windows



Folder structure

Folder structure

Figure 4.7. Windows XP file structure.



Copy path:

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

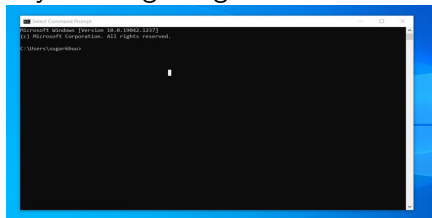
When naming folder & file:

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

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

Command line interpreters/Shells

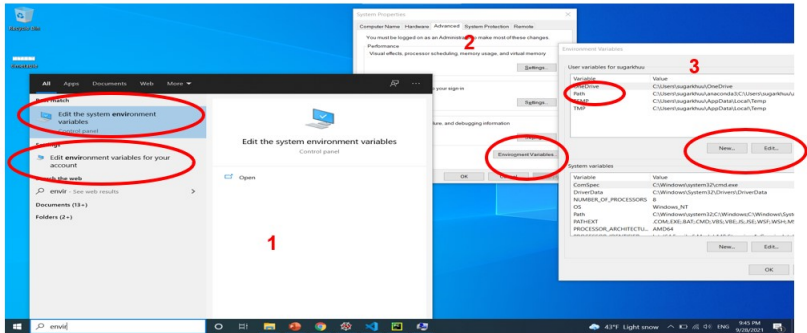
You are able to control your computer through commanding the OS from terminals (Win+CMD, Ctrl+T). More powerful and flexible than usual GUI way of doing things



- Windows: Command prompt, Powershell. More recently, Windows Terminal
 - Sample commands: cd, dir, mkdir, exit, cls
- Linux: Bash, Mac: Terminal (zsh)

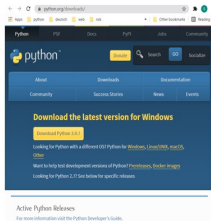
Path (environment variable)

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



Getting Python

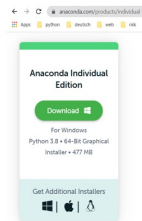
Standalone Python



<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>

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
```

Python basic concepts for today:

Detailed instructions in the supplementary materials next Monday.

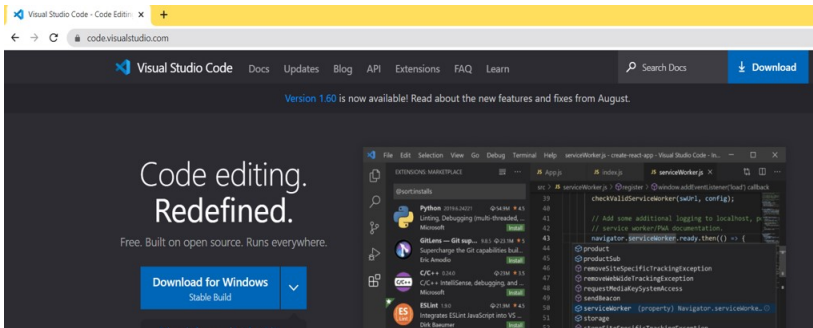
- Basic syntax
- Basic operators
- Packages (+ pip)

Possible to use Python in many environments

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

Using VS code (code editor) for Python

- Extensions
- Various functionalities: Navigation, Search/Replace, Version control, Code execution, Debugging
- Settings



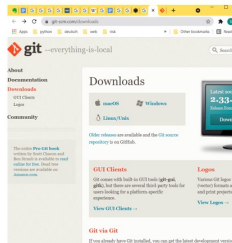
Git – Version control system (VCS)

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.

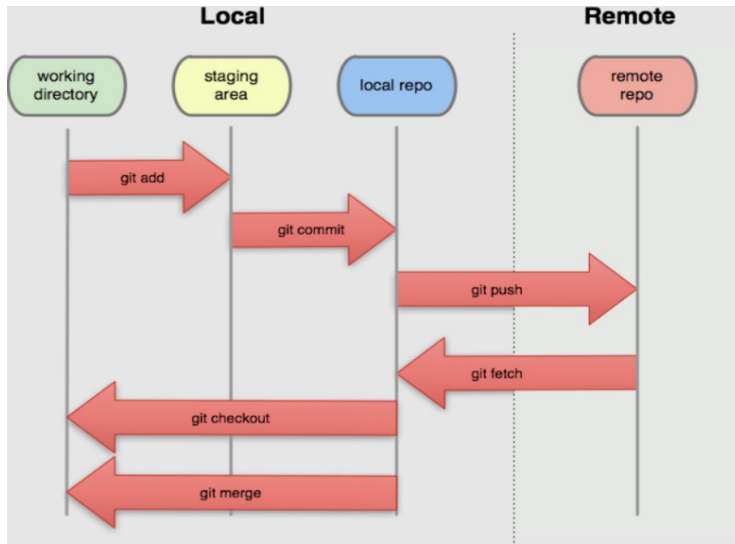
<https://www.geeksforgeeks.org/version-control-systems/>

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 web interface for the repository 'TheAlgorithms/Python'. The top navigation bar includes links for 'Why GitHub?', 'Team', 'Enterprise', 'Explore', 'Marketplace', and 'Pricing', along with a search bar and 'Sign in'/'Sign up' buttons. Below the repository name, there are tabs for 'Code', 'Issues', 'Pull requests', 'Actions', 'Projects', 'Wiki', 'Security', and 'Insights'. The main content area displays a list of files and their commit history. The 'About' section on the right provides information about the repository, including a list of all algorithms implemented in Python and a link to the README.

← → github.com/TheAlgorithms/Python ☆ ★ 5

Why GitHub? Team Enterprise Explore Marketplace Pricing

Search / Sign in Sign up

TheAlgorithms / Python Public

Sponsor Notifications Star 119k Fork 31.7k

<> Code Issues 8 Pull requests 33 Actions Projects Wiki Security Insights

master 15 branches 0 tags

Go to file Code +

File	Commit	Age
.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

Joe-Sin7h Fixed #4764 (#4779) ✓ 502103 4 days ago 2,180 commits

the-algorithms.com/

python education algorithm
practice interview sorting-algorithms
learn algo algorithm-competitions
sorts hacktoberfest
algorithms-implemented
community-driven searches

Readme

Mit License

Releases

No releases published

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

Git demonstration

- Configure user email and user name on the computer
- Create a new repo on Github and clone locally
- Make changes locally, compare and commit
- Make changes remotely and pull
- git log, git diff
- Repo access and rights

Homework

1. Task 1

2. Task 2

- Submit your result in a 'Homework' github repo
- Deadline: 1 week

Task 1

1. Create a new repository in your Github
2. Clone this repository to your local machine
3. Create a python file in the local repo (in your folder)
4. In the file, write a code which asks a question and receives the answer from the user
5. Commit and push your change
6. Create 3 more questions, and commit and push

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. Commit, push хоёрын ялгаа юу вэ?
6. Push хийхээс өмнө олон дахин commit хийж болох уу?
7. Commit хийхэд github repo-д access хэрэгтэй юу?

Thank you. See you next week!