

A background image showing two hands, one from the left and one from the right, reaching towards each other and forming a heart shape with their fingers. The hands are wearing white and grey ribbed sleeves. The background is a soft, out-of-focus green and yellow bokeh.

Lab 5: Search & Sort

SUNY Korea - Francois Rameau

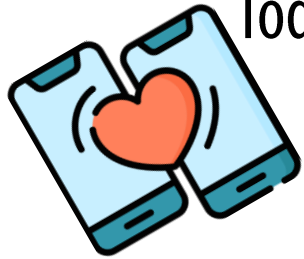
Spring 2023

GitHub Classroom



Lab 5 - Search and sort

Finding love with Python



Today you are an entrepreneur, you just opened your online dating platform

Your role is to find the ideal partner to our two clients



Pamela



George

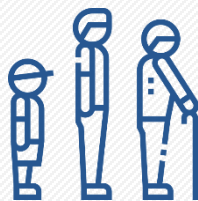
Our database

For each user you stored the following

Name



Age



In year

Height



In cm

Income



In k\$

Each person is stored in a list

```
user = ["Joe", 25, 175, 50]
```

Name

Age

Height

income

Our database

Let's assume, you have already collected data from male and female that you saved as CSV files (comma-separated values)



list_females.csv

Peyton, 26, 157, 49
 Leah, 27, 163, 54
 Audrey, 23, 156, 50
 Savannah, 24, 168, 57
 Alyssa, 29, 165, 59
 Aubrey, 22, 160, 52
 Kylie, 31, 162, 61
 Bella, 25, 158, 55
 Claire, 26, 157, 50
 Stella, 30, 171, 61
 Maya, 24, 163, 56
 Gabriella, 28, 167, 60
 Skylar, 21, 155, 48
 Samantha, 23, 162, 55
 Allison, 26, 160, 52
 Caroline, 30, 169, 61
 Elena, 25, 158, 55
 Isabelle, 27, 163, 54
 Ariana, 32, 154, 57
 Genesis, 20, 172, 47
 Ruby, 26, 157, 49
 Reagan, 31, 162, 61
 Clara, 22, 160, 52
 Naomi, 23, 156, 50
 Khloe, 29, 165, 59
 Annabelle, 24, 168, 57
 Alexa, 26, 157, 50



list_males.csv

Samuel, 27, 168, 40
 Derek, 34, 175, 52
 Frank, 40, 188, 70
 Tyler, 28, 172, 48
 Raymond, 37, 180, 65
 Shawn, 35, 178, 60
 Travis, 26, 170, 42
 Gregory, 38, 183, 67
 Jeffery, 33, 174, 52
 Trevor, 32, 174, 52
 Douglas, 39, 183, 67
 Austin, 29, 169, 45
 Lee, 31, 173, 48
 Cody, 27, 169, 45
 Bryan, 37, 180, 65
 Luis, 33, 176, 52
 Carlos, 35, 178, 60
 Alan, 38, 183, 67
 Mario, 42, 188, 70
 Juan, 32, 174, 52
 Jeremiah, 26, 171, 45
 Scott, 34, 175, 52
 Alexander, 29, 171, 45
 Victor, 36, 180, 65
 Joel, 31, 173, 48
 Jeff, 40, 187, 75
 Johnny, 28, 172, 30

1. Open the database

A function to read the database is already implemented

```
candidates_list = read_database('list_males.csv')  
candidates_list = candidates_list[1:]
```

← Just to remove the header

The database is stored in a nested list

```
[[ 'Douglas', 39, 183, 67 ],  
 [ 'Austin', 29, 169, 45 ],  
 [ 'Lee', 31, 173, 48 ],  
 [ 'Cody', 27, 169, 45 ],  
 [ 'Bryan', 37, 180, 65 ],  
 [ 'Luis', 33, 176, 52 ],  
 [ 'Carlos', 35, 178, 60 ],  
 [ 'Alan', 38, 183, 67 ],  
 [ 'Mario', 42, 188, 70 ],  
 [ 'Juan', 32, 174, 52 ] ]
```

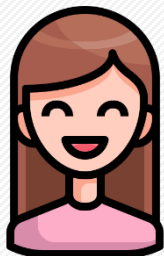
← candidates_list[2]

2. The clients

The goal of this lab is to find a good match to our two clients
We will define a good match by “a partner with similar features”

Close age, close height and close income

For instance



Pamela

[25, 157, 65]



Henry

[65, 190, 55]



Peter

[24, 175, 60]

Peter has a closer profile than Henry then he is a better match for Pamela

3. Euclidean distance

TO DO:

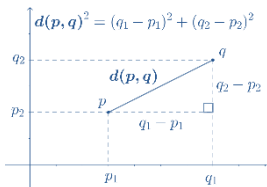
Implement a function to compute the Euclidean distance between a client (Pamela or George) and one person in the dataset

Euclidean distance

$$d(p, q) = \sqrt{\sum_{i=1}^N (p_i - q_i)^2}$$

In our case

$$\begin{aligned} & d(\text{person1}, \text{person2}) \\ &= \sqrt{(\text{age1} - \text{age2})^2 + (\text{height1} - \text{height2})^2 + (\text{income1} - \text{income2})^2} \end{aligned}$$



A small distance represent a good match!
Try to implement it with a **FOR** loop in the function
EuclideanDistance()

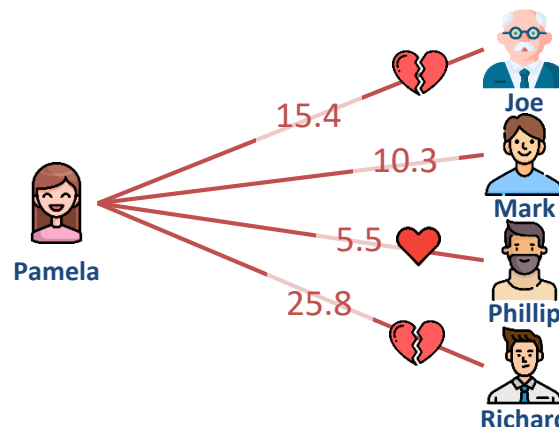
4. Compute the distance between the client and the candidates

TO DO:

Compute all the distances between Pamela (or George) and all the candidates in the database (male or female)

You will implement this in
`ComputeAllDistances()`

→ This function will return the list of distances between one client (Pamela or George) and a database.

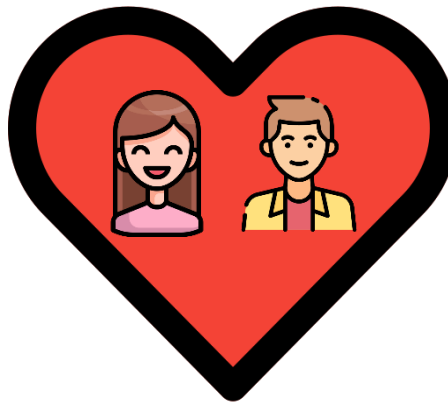


5. Find the best match

TO DO:

Complete the function `FindMin()` that returns the minimum value in a list and its corresponding index

Using the index returned by your function, display the name and characteristic of the best match for our client



By the way, you just implemented a naïve nearest neighbor search

6. Bonus

- What can you say about the complexity of this search strategy?
- How would you return the top 5 best match? (try to implement it)
- Can we return all the results given a range over each criterion (no need to implement)



An illustration of a person in a blue shirt and dark pants standing next to a whiteboard. The whiteboard contains various charts, including a bar chart, a pie chart, and a line graph. To the right of the person is a laptop displaying code. Above the laptop, there are icons for 'FAIL' (a red circle with a dot) and 'PASS' (a green circle with a dot). There are also icons for a red 'X' and a green checkmark. The background is light blue with some floating squares and a gear icon.

Novelty today THE UNIT TEST!

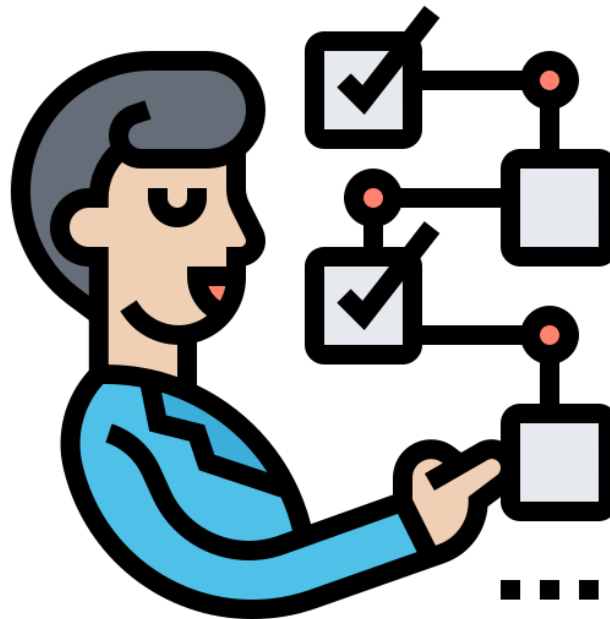
```
// Validate the input parameters  
if (ppErrorText == NULL || pdwSize == NULL)  
{  
    ERROR_INVALID_PARAMETER;  
}
```

```
// Validate the input parameters  
if (ppErrorText == NULL || pdwSize  
{  
    VALID_PARAMETER;  
}
```

Unit test

! Terminology alert !

A **unit test** is a type of software testing technique that focuses on testing individual components or functions of a program in isolation to ensure they perform as expected. This helps developers **identify and fix errors** early in the development process.



Unit test

Note that you now have:

YourHomework.py

AND

YourHomework_test.py ←

This file contains the
unit tests

Do not touch



Do not modify the unit test
by any mean

Unit test



WARNING 2



Only work on the origin *.py file I am providing you!

```
lab5_random_py_file.py  
lab5Search_test.py  
lab5Search.py
```

Do not do that!!



Work only here



Unit test

- Does it change the way I push my work on Github?
 - Not at all, push and commit as before, **follow Lab0 very carefully**
- Why do we do unit test?
 - Because it will test every single piece of your code to make sure it is working as expected. It saves time for the TA and ensure a fair evaluation
- What unit test does for me?
 - It will directly tell you if you are code is right or wrong such that you can correct a function that you have not coded properly!

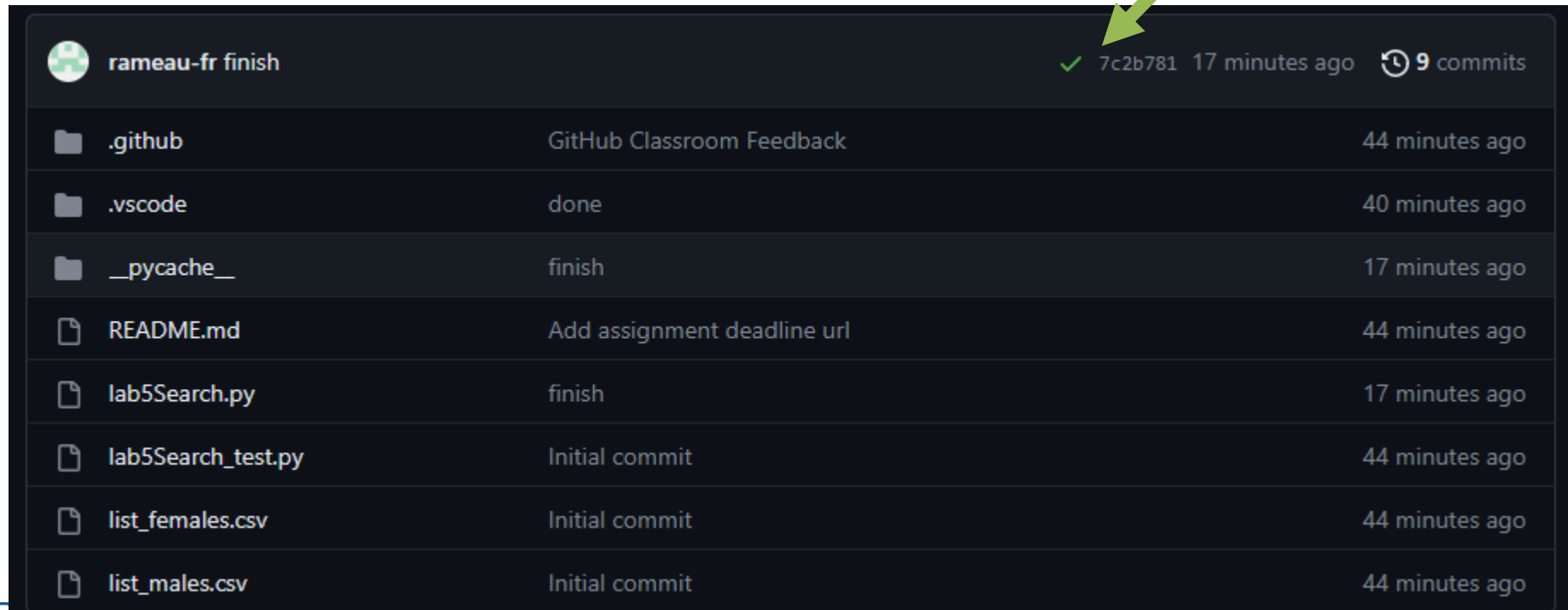
Unit test











Let's push a code on GitHub to see what is happening!!

Two possibilities

1. Your code is working as intended

You have this little green check that means that all your code is working as intended
You can click on it!



 rameau-fr finish	✓ 7c2b781 17 minutes ago	 9 commits
 .github	GitHub Classroom Feedback	44 minutes ago
 .vscode	done	40 minutes ago
 __pycache__	finish	17 minutes ago
 README.md	Add assignment deadline url	44 minutes ago
 lab5Search.py	finish	17 minutes ago
 lab5Search_test.py	Initial commit	44 minutes ago
 list_females.csv	Initial commit	44 minutes ago
 list_males.csv	Initial commit	44 minutes ago

Unit test

```
✓ Run education/autograding@v1
12 Ran 1 test in 0.000s
13
14 OK
15 ['David', 35, 185, 60]
16
17 ✓ test_euclidean
18
19 🚩 test_ComputeAllDistances
20
21
22 .
23 -----
24 Ran 1 test in 0.000s
25
26 OK
27 ['David', 35, 185, 60]
28
29 ✓ test_ComputeAllDistances
30
31 🚩 test_FindMin
32
33
34 .
35 -----
36 Ran 1 test in 0.001s
37
38 OK
39 ['David', 35, 185, 60]
40
41 ✓ test_FindMin
42
43
44 ::***::
45
46 All tests passed
47
48 🌟🌟💖💎👤💎💖🌟🌟💖💎👤💎💖🌟
49
50 Points 30/30
```

You will directly see
which tests has been
successful or failing

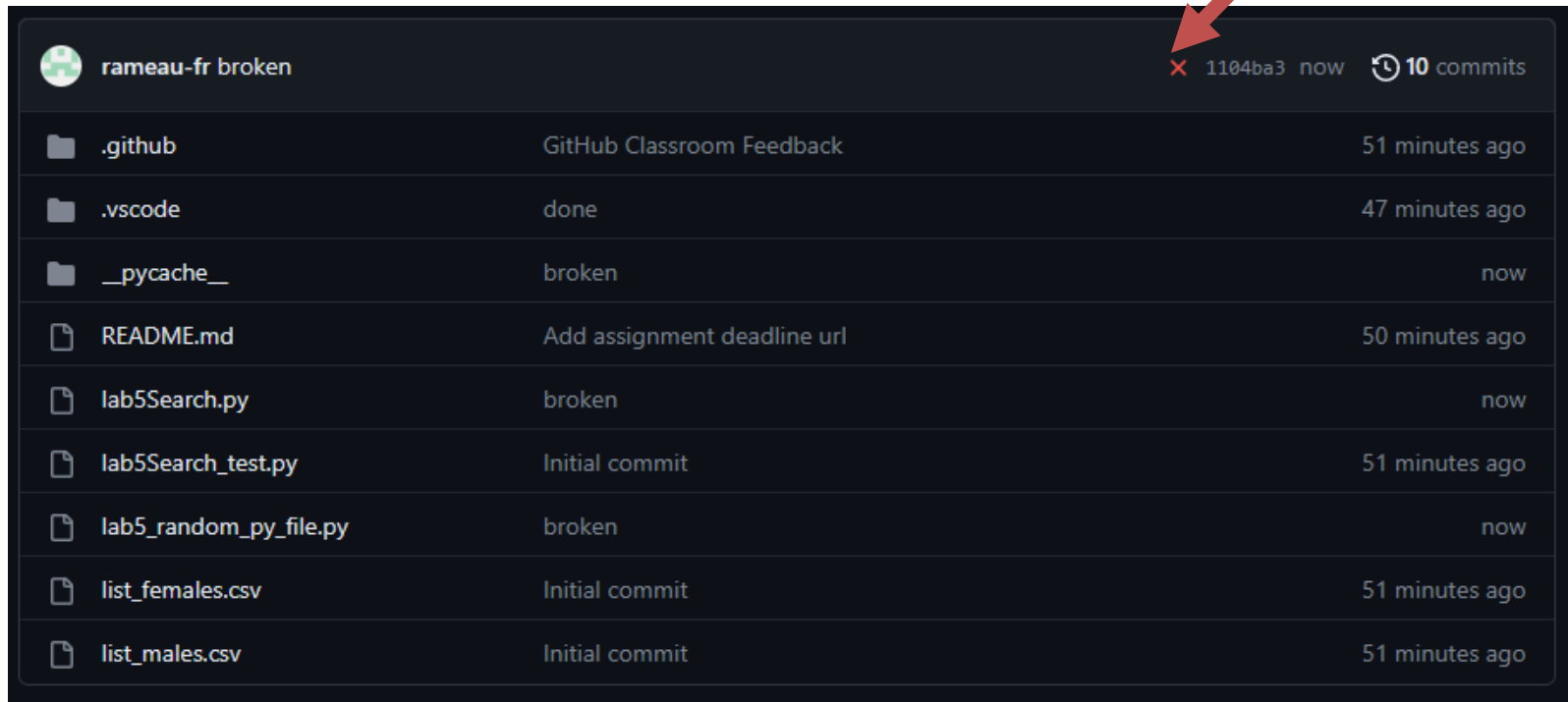
Unit test













Let's push a code on GitHub to see what is happening!!

Two possibilities

2. Your code is NOT working as intended

You have this little red cross that means that all your code is NOT working as intended
You can click on it!



 rameau-fr broken	 1104ba3 now  10 commits
 .github	GitHub Classroom Feedback 51 minutes ago
 .vscode	done 47 minutes ago
 __pycache__	broken now
 README.md	Add assignment deadline url 50 minutes ago
 lab5Search.py	broken now
 lab5Search_test.py	Initial commit 51 minutes ago
 lab5_random_py_file.py	broken now
 list_females.csv	Initial commit 51 minutes ago
 list_males.csv	Initial commit 51 minutes ago

Unit test

```
Run education/autograding@v1
41 Ran 1 test in 0.001s
42
43 FAILED (failures=1)
44 ['David', 35, 185, 60]
45
46 ✖ test_ComputeAllDistances
47 ::error::Error: Exit with code: 1 and signal: null
48 📄 test_FindMin
49
50
51 .
52 -----
53 Ran 1 test in 0.000s
54
55 OK
56 ['David', 35, 185, 60]
57
58 ✔ test_FindMin
59
60
61 ::***::
62 Points 20/30
63 Warning: The `set-output` command is deprecated and
https://github.blog/changelog/2022-10-11-github-act
```

You will directly see
which tests has been
successful or failing
AND
Your grade! (of course
the TA will ALWAYS
double check for
assignments)

Unit test

What do I see as an instructor

Students total 61
61 Rostered 0 Added students


Accepted assignments 1
1 Students

Assignment submissions 1
0 Submitted 1 Not submitted


Passed students 1
1/1 Passed

Unlinked accounts ▾ Accepted ▾ Submitted ▾ Passing ▾ Sort by: ▾


Classroom roster




Francois Rameau Submitted

@rameau-fr [Latest commit April 18, 2023 13:05](#) ✓  2 commits


30/30

 Repository


 Feedback

OR


Classroom roster




Francois Rameau Submitted

@rameau-fr [Latest commit April 18, 2023 13:50](#) ✓  4 commits

20/30

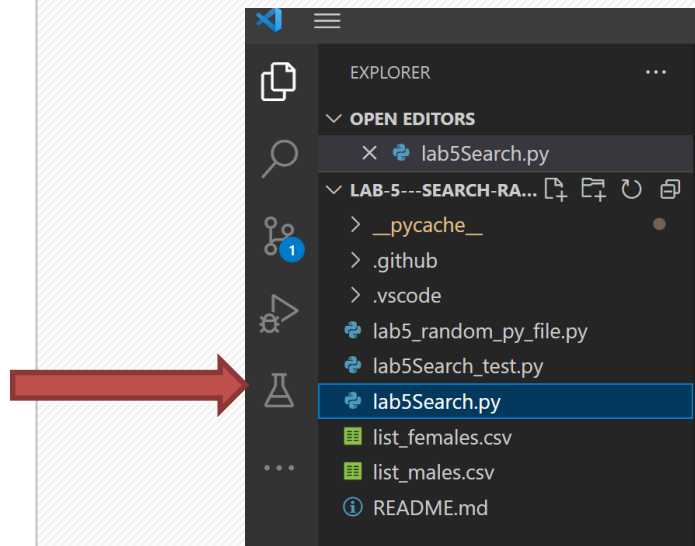
 Repository

 Feedback

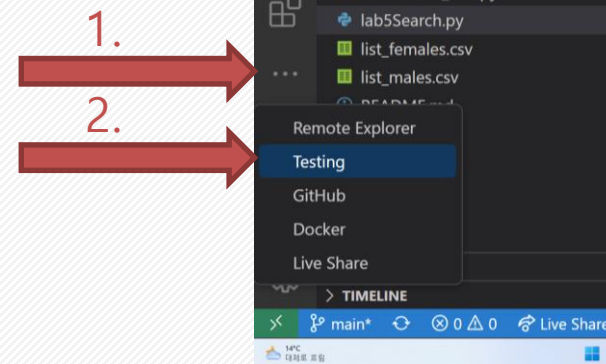
Unit test

How to run unit test locally
In 5 clicks only!

Step 1: click on Testing



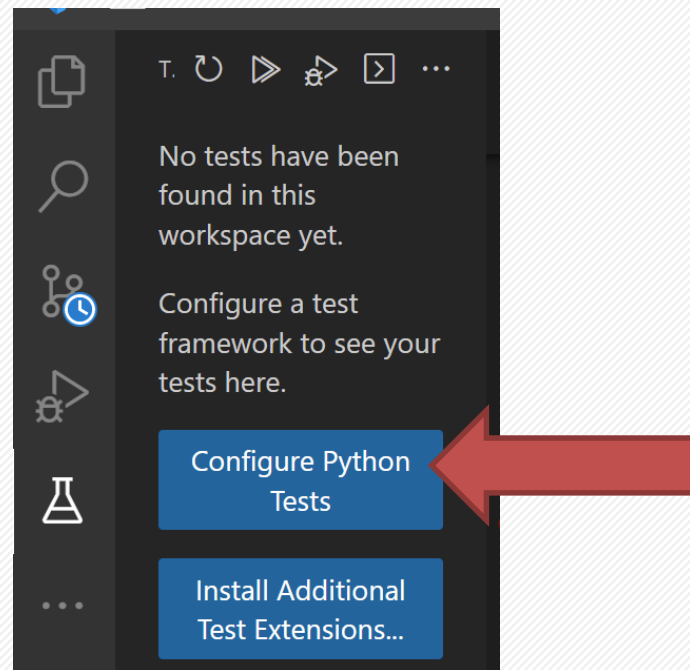
OR



Unit test

How to run unit test locally

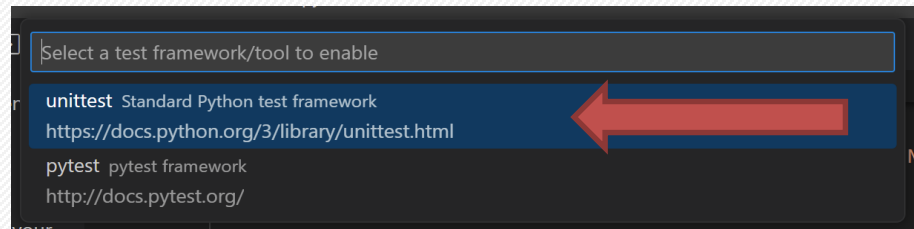
Step 2: configure



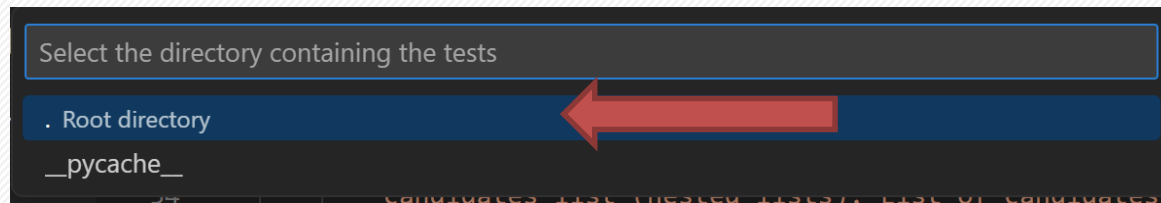
Unit test

How to run unit test locally

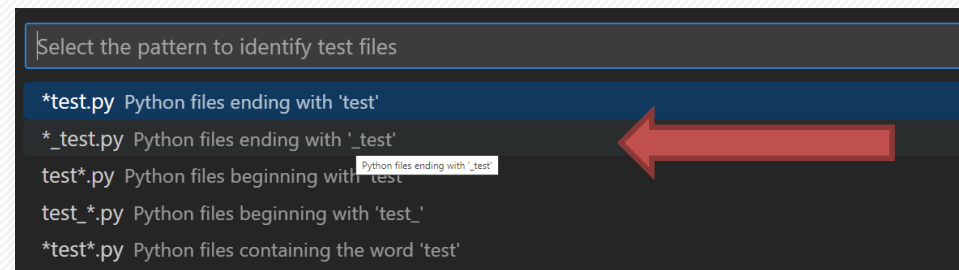
Step 3: Select unittest



Step 4: .Root directory

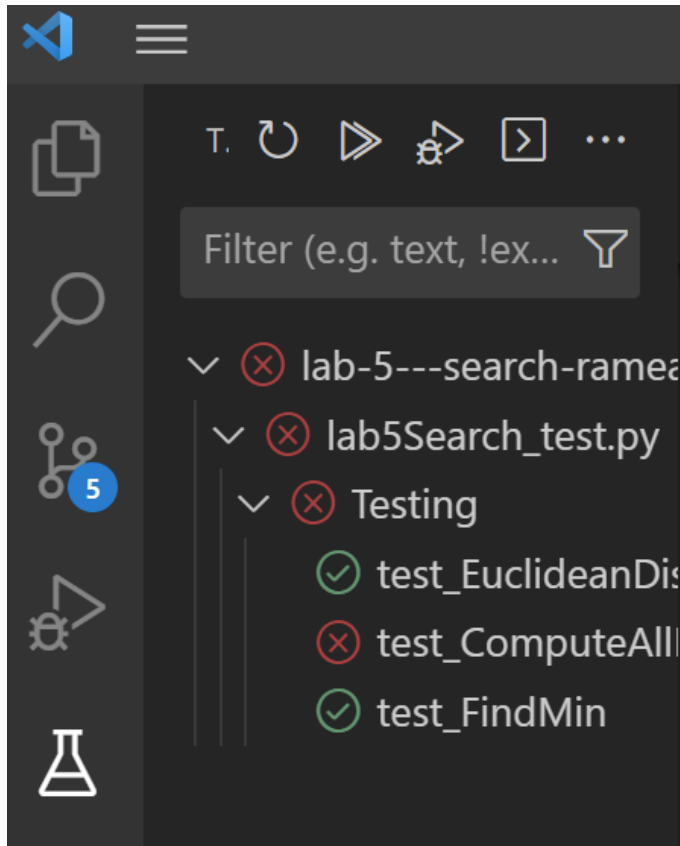


Step 5: *_test.py



Unit test

How to run unit test locally



Now you are ready to go if you want to test your code locally

Feedback via pull request

You now have a branch feedback in your repository, we will use it to send you feedback directly (you will receive an email each time we are trying to reach you out via this mean)

