

ВЕБ-СКРАПИНГ (СБОР ДАННЫХ)

Лабораторная работа

Основная цель

В разработанной вами игре (угадай слово) очень скудный список предметов (блюд). Ваша цель — получить информацию из другого источника и добавить её на свой сайт. На этот раз мы берём информацию со страницы местных рецептов с большим количеством разнообразных блюд. Ручной сбор данных может занять месяцы, поэтому мы воспользуемся методом, называемым веб-скрапингом. В этой части мы подготовим все необходимые элементы перед получением информации с других сайтов (это мы сделаем в следующей лабораторной работе).

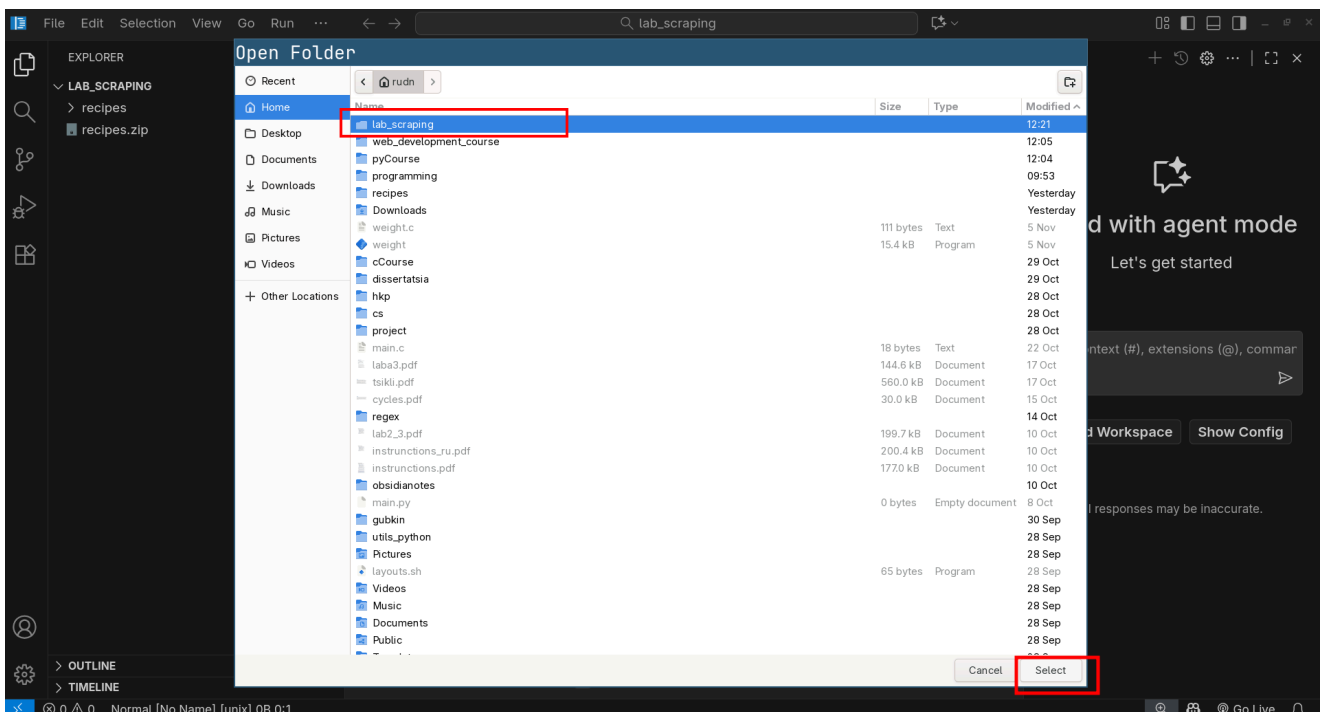
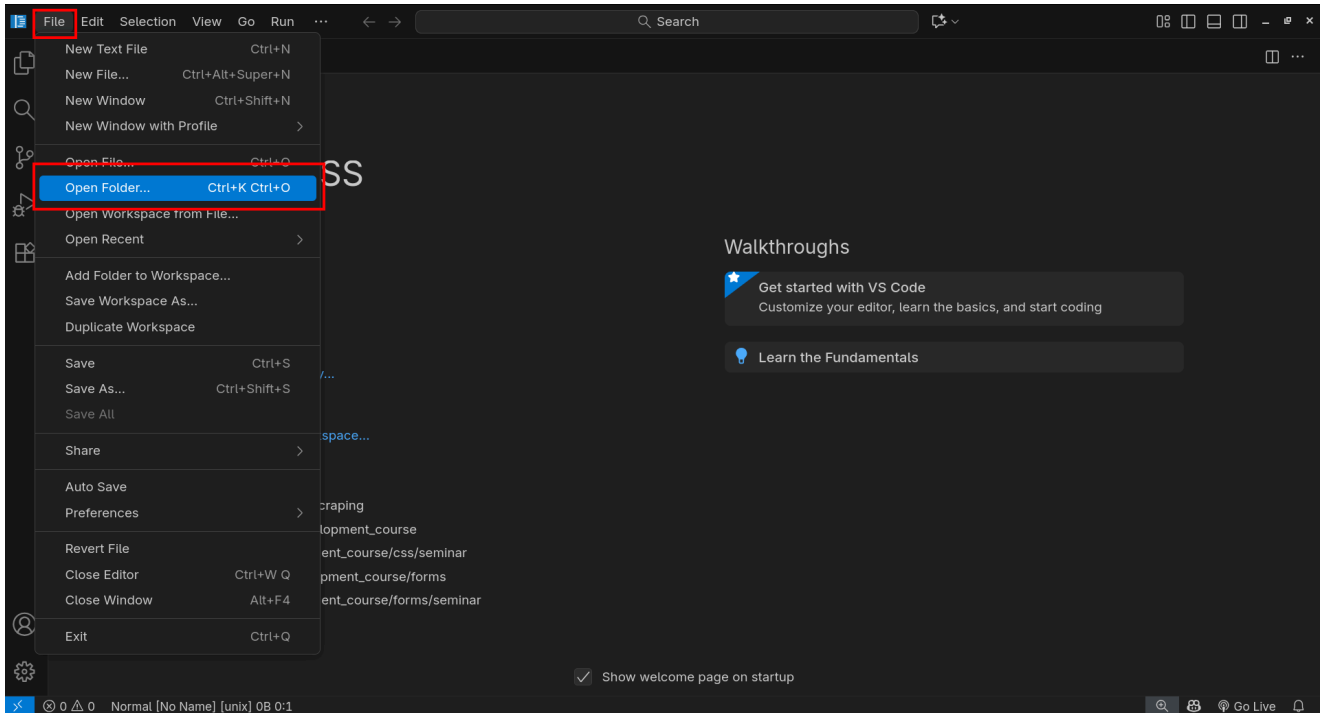
Важная информация перед началом работы

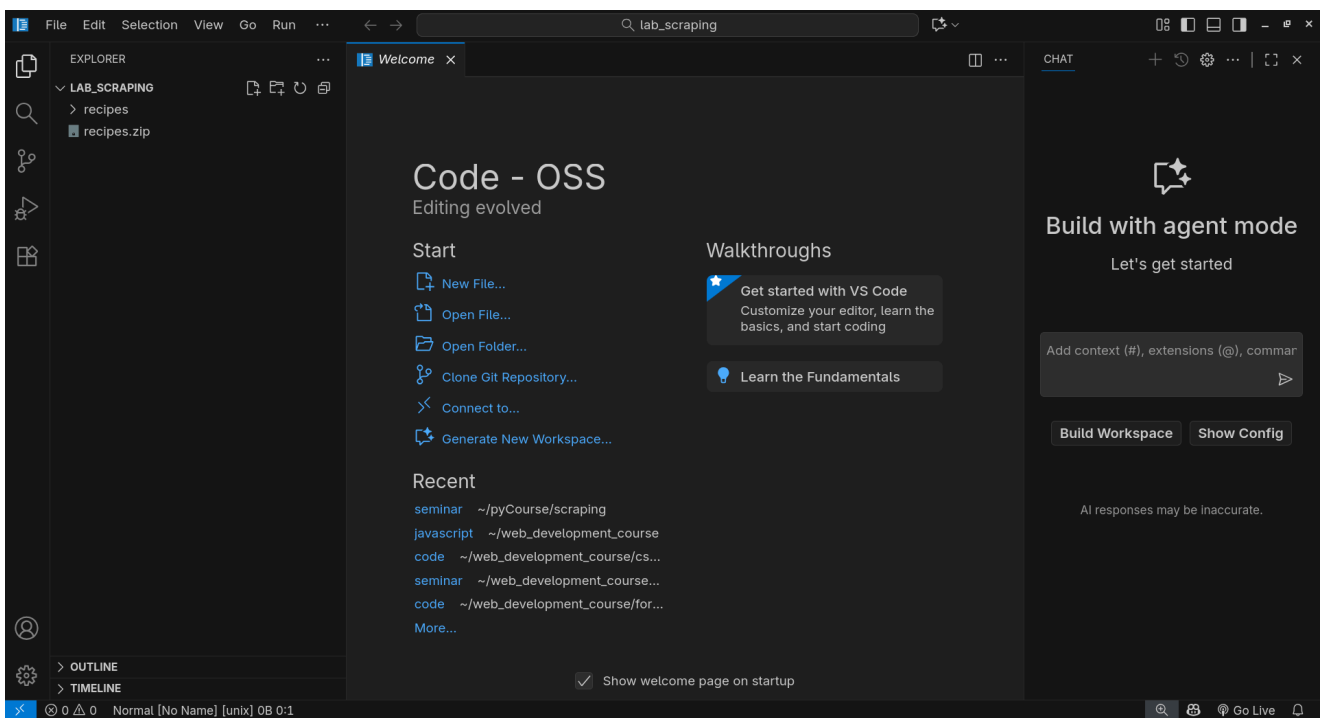
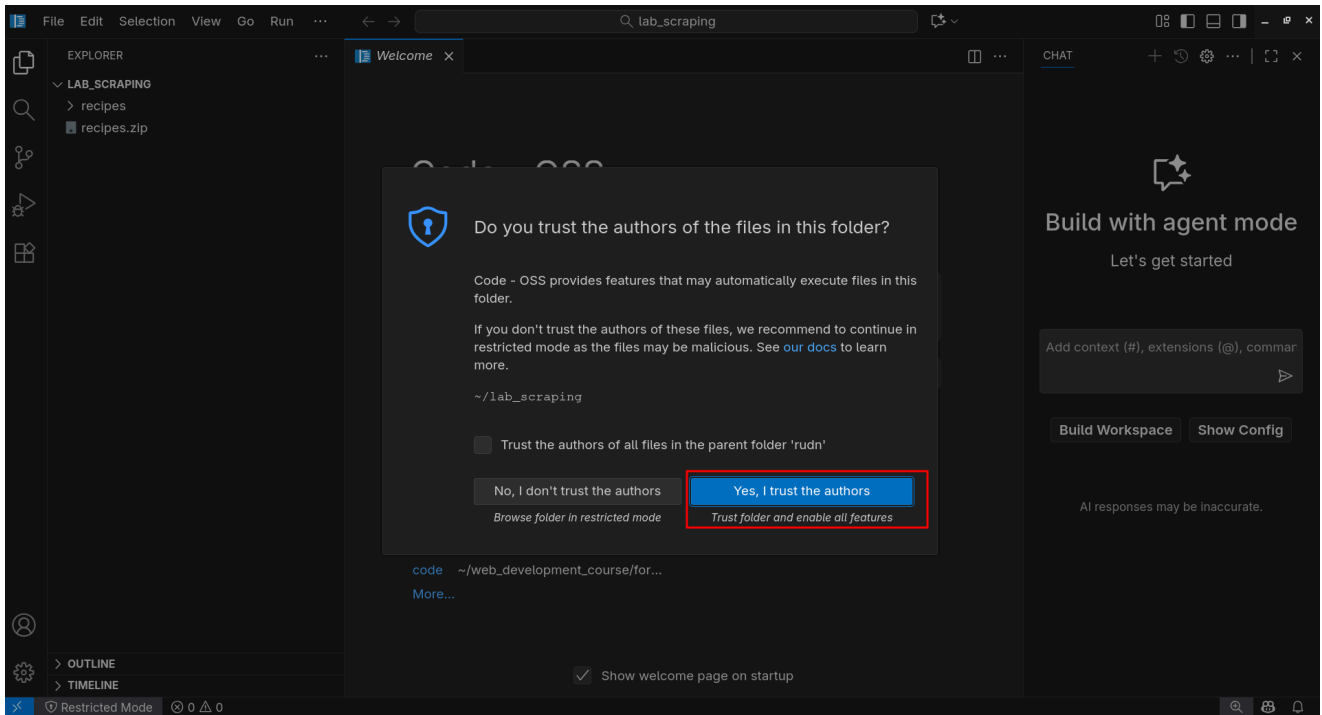
- **Веб-скрапинг имеет некоторые этические аспекты. Публичное использование информации с других сайтов без разрешения может привести к неудобствам или юридическим проблемам. Поэтому я рекомендую использовать общедоступную информацию, предназначенную для свободного использования.**
- **На сайтах установлено специальное программное обеспечение, блокирующее веб-скрапинг. Поэтому перед началом работы убедитесь, что сайт, который вы собираетесь скрапить, разрешает такую практику.**
- **Практически все сайты имеют защиту от DDoS-атак, что означает ограничение на количество запросов, которые вы можете сделать за определённый промежуток времени (до блокировки вашего IP-адреса). Поэтому использование библиотек, таких как `request`, следует обдумывать и ограничивать несколькими попытками (не используйте их циклами). Также рекомендуется по возможности добавлять паузы между запросами.**
- **Используйте веб-скрапинг в основном для автоматизации скучных задач или при ограниченном объёме информации. Не создавайте точные копии сайта и не используйте всю доступную информацию.**
- **Эта лаборатория использует общедоступную и ограниченную информацию в образовательных целях и Локальная среда, то есть мы отправляем запросы не к реальным веб-сайтам в интернете, а к локальному веб-сайту, работающему на вашем компьютере.**

После прочтения всей этой информации мы можем начать

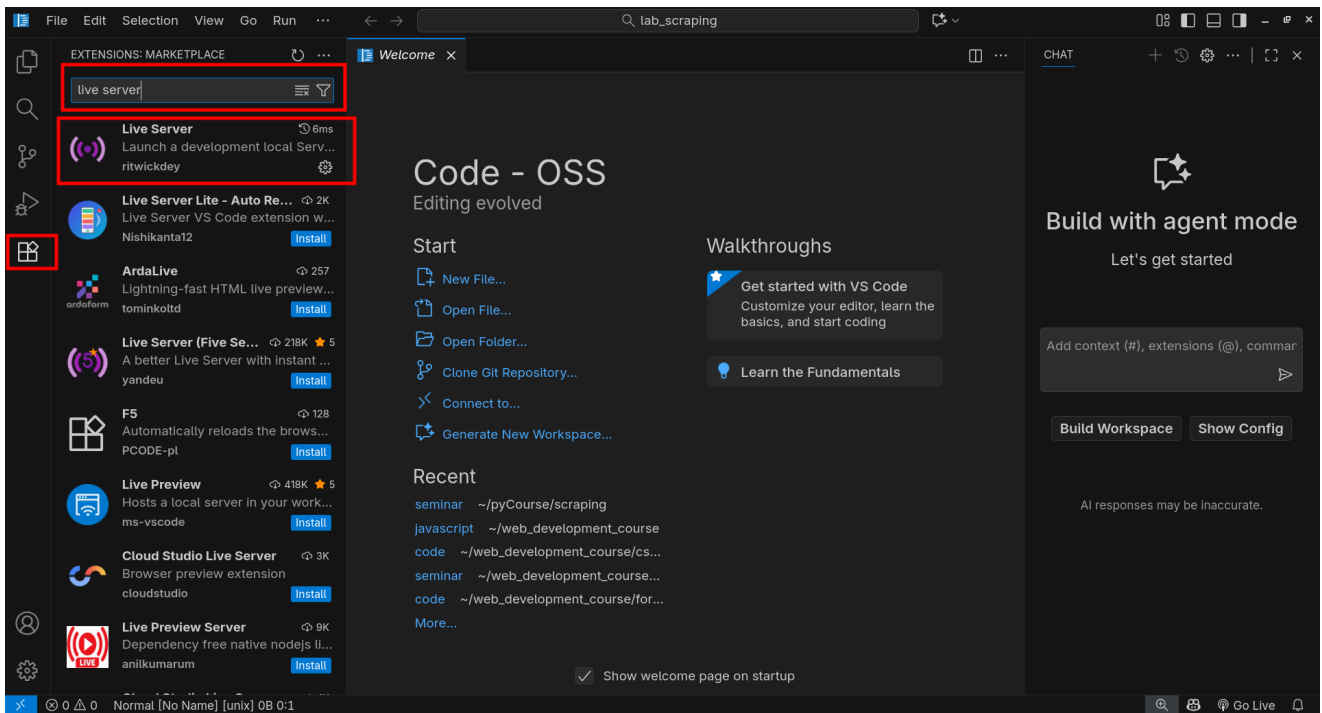
НАСТРОЙКА РАБОЧЕЙ ПАПКИ

1. Создайте новую папку с именем `lab_scraping`.
2. Скопируйте файл `recipes.zip` в папку `lab_scraping`.
3. Распакуйте файл `recipes.zip` в папку `lab_scraping`. Это создаст новую папку с именем `recipes`.
4. Откройте `vscode`, затем нажмите `file` → `open folder` и выберите папку `lab_scraping`. Эта папка откроется в `vscode`. Если `vscode` предложит вам доверять авторам, нажмите `yes`.

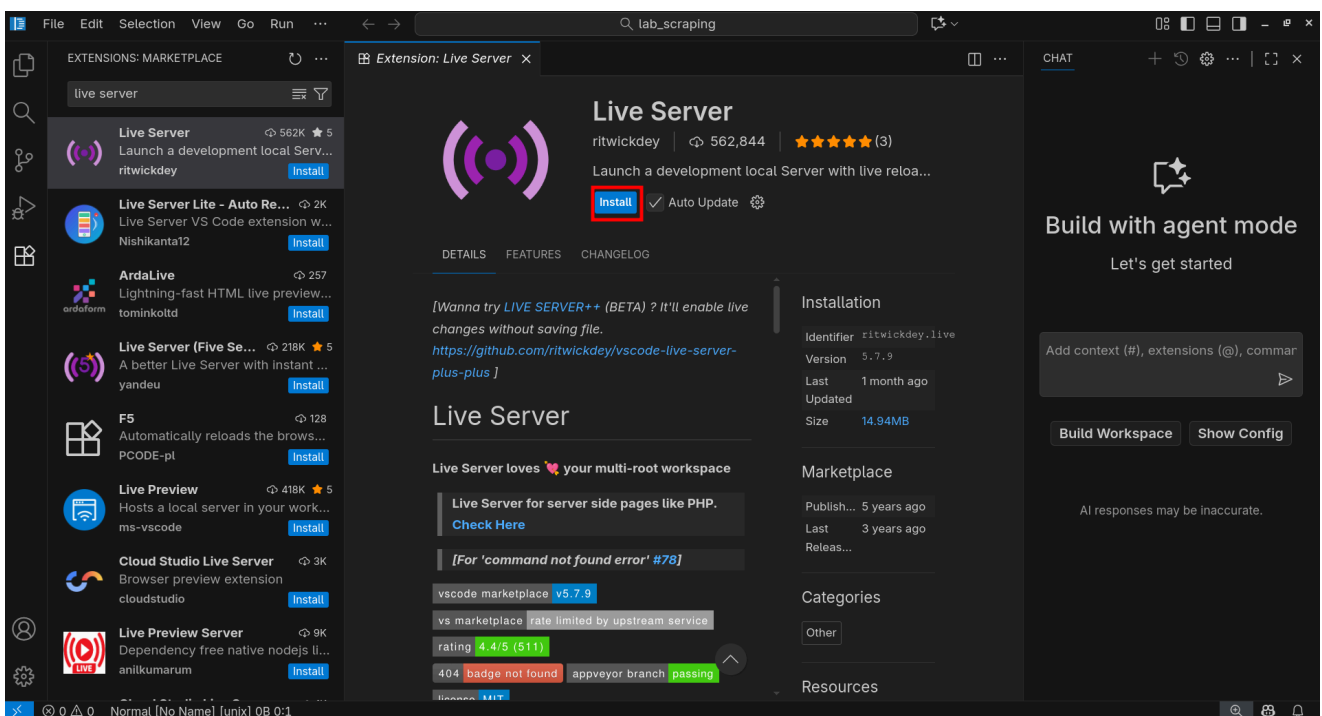




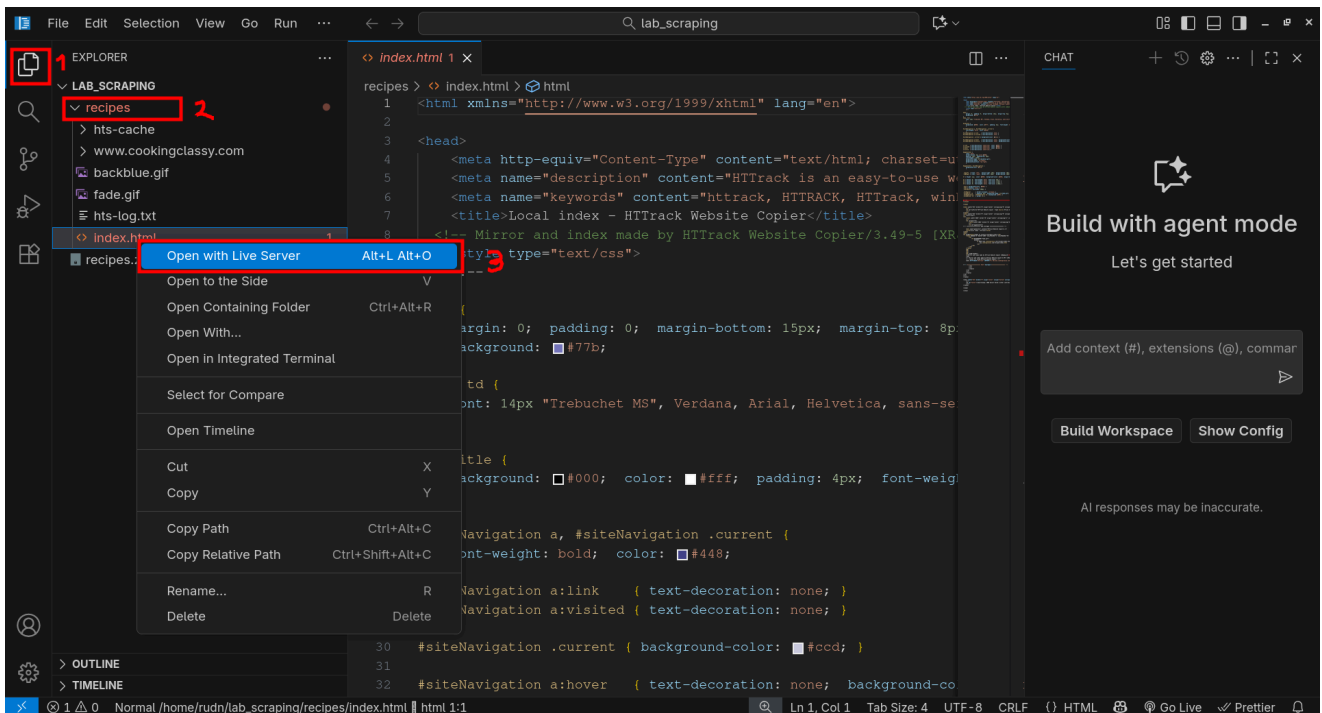
5. Нажмите кнопку «Расширения», затем в строке поиска напишите «live server». Затем выберите расширение live server (как показано на рисунке).



6. Нажмите «Установить» и дождитесь установки расширения.



5. В «vscode» перейдите в папку «recipes», чтобы отобразить её содержимое. Затем найдите в папке «recipes» файл «index.html», щёлкните по нему правой кнопкой мыши и выберите «Открыть с помощью Live Server». Всё это нужно сделать после установки расширения «live server», иначе оно не заработает.

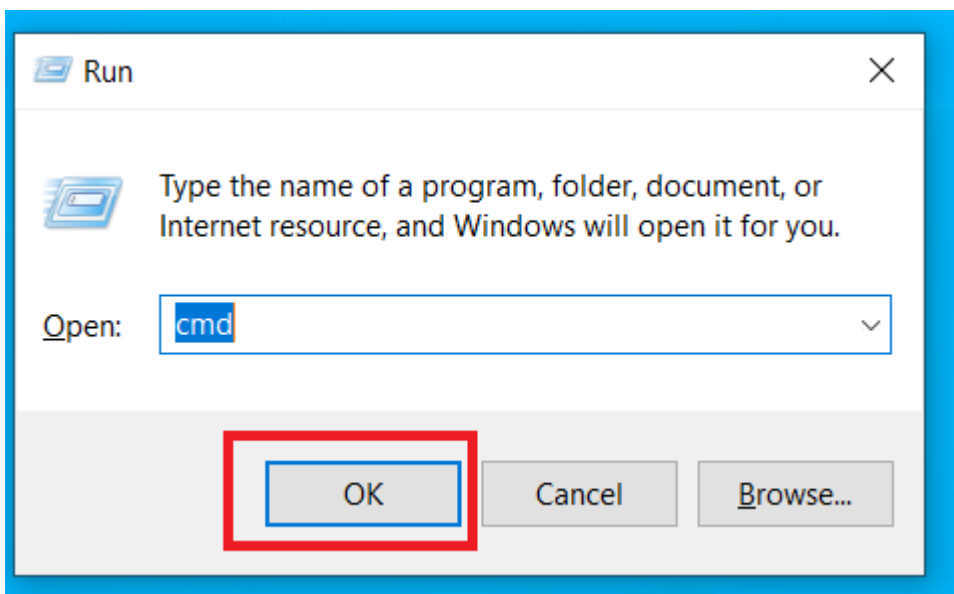


После этого откроется браузер со страницей, с которой мы будем скачивать рецепты. Эта веб-страница развёрнута в автономном режиме, то есть вся информация берётся из папки `recipes`, которая находится на вашем локальном компьютере.

УСТАНОВКА PYTHON

6. Убедитесь, что в вашей системе установлен Python:

- нажмите клавишу Windows, а затем клавишу `r` (`win + r`)
- откройте терминал, введя `cmd`, а затем нажмите `Enter`



- введите в терминале `python --version`

```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.18363.592]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\RUDN>python --version
Python 3.13.9
номер версии

C:\Users\RUDN>
```

Если терминал выдаёт версию, значит, Python установлен. Если нет, перейдите к шагу 7

- Используйте установщик, предоставленный TUIS. Если он не работает, скачайте его с сайта <https://www.python.org/downloads/>, выбрав свою операционную систему.

python

Download the latest source release

Download Python 3.14.0

Looking for Python with a different OS? Python for [Windows](#), [Linux](#), [Unix](#), [macOS](#), [android](#), [other](#)

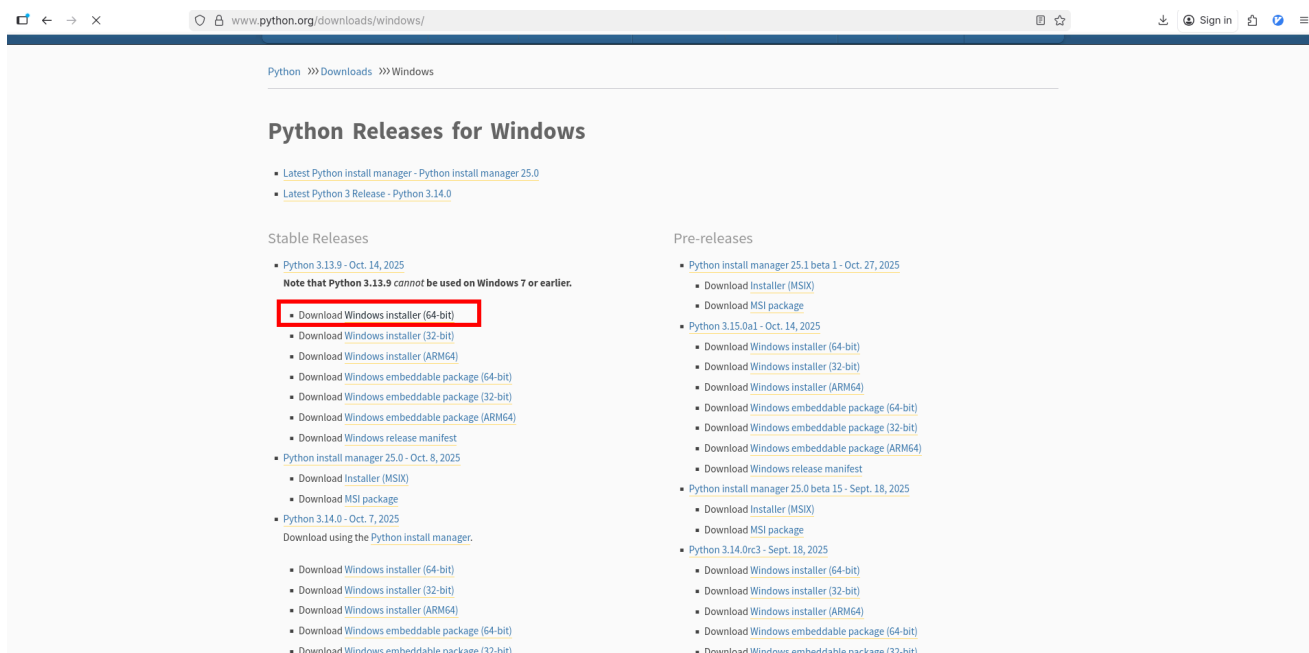
Want to help test development versions of Python 3.15? [Pre-releases](#), [Docker images](#)

Active Python Releases

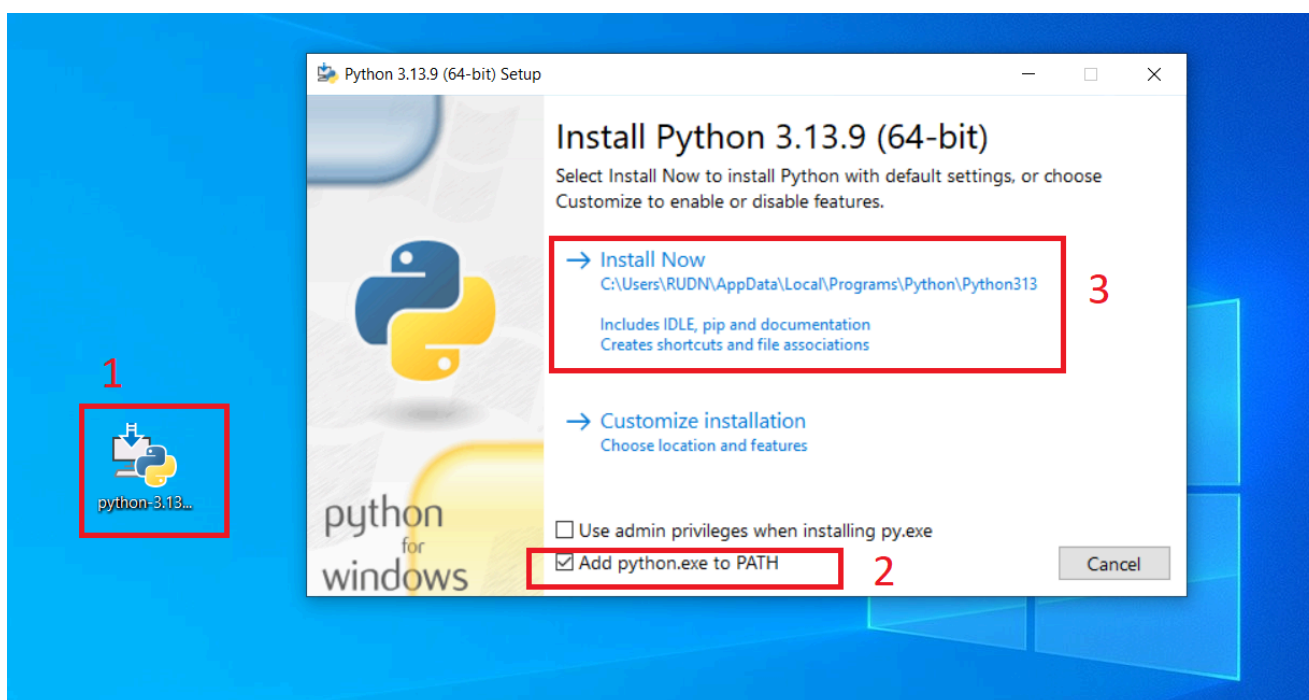
For more information visit the [Python Developer's Guide](#).

Python version	Maintenance status	First released	End of support	Release schedule
3.15	pre-release	2026-10-07 (planned)	2031-10	PEP 790
3.14	bugfix	2025-10-07	2030-10	PEP 745
3.13	bugfix	2024-10-07	2029-10	PEP 719
3.12	security	2023-10-02	2028-10	PEP 693
3.11	security	2022-10-24	2027-10	PEP 664
3.10	security	2021-10-04	2026-10	PEP 619
3.9	end of life, last release was 3.9.25	2020-10-05	2025-10-31	PEP 596

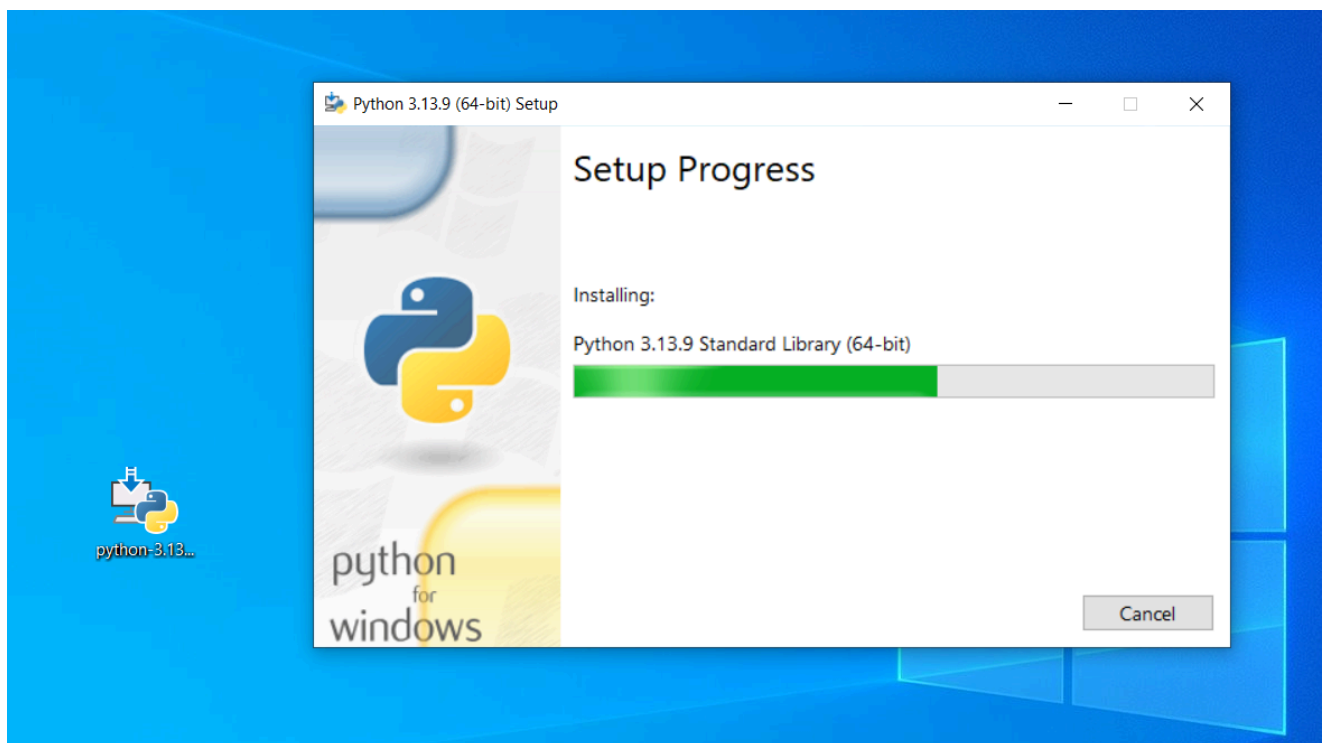
Looking for a specific release?



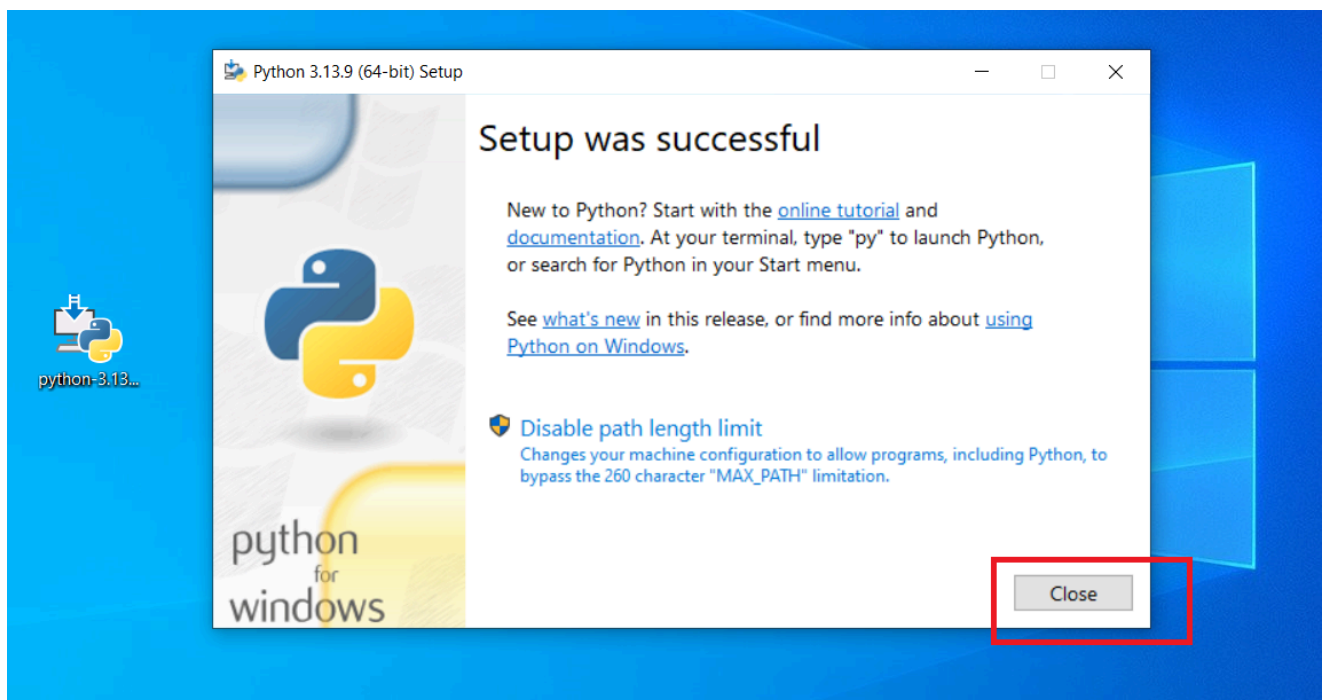
8. Дважды щёлкните по установщику Python, затем установите флажок «Добавить Python в переменную окружения PATH» и нажмите «Установить сейчас».



Python должен начать установку в вашей операционной системе.

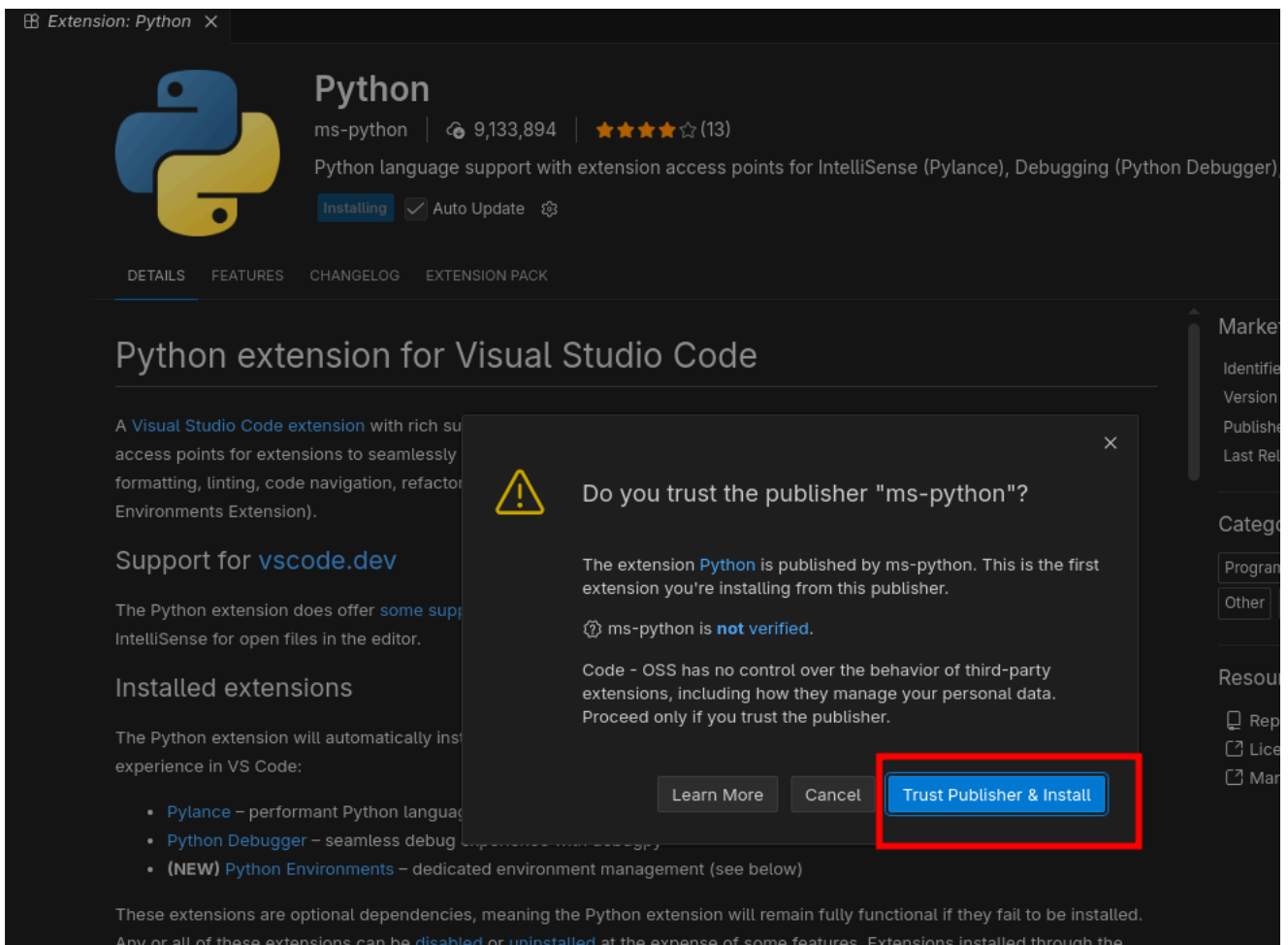
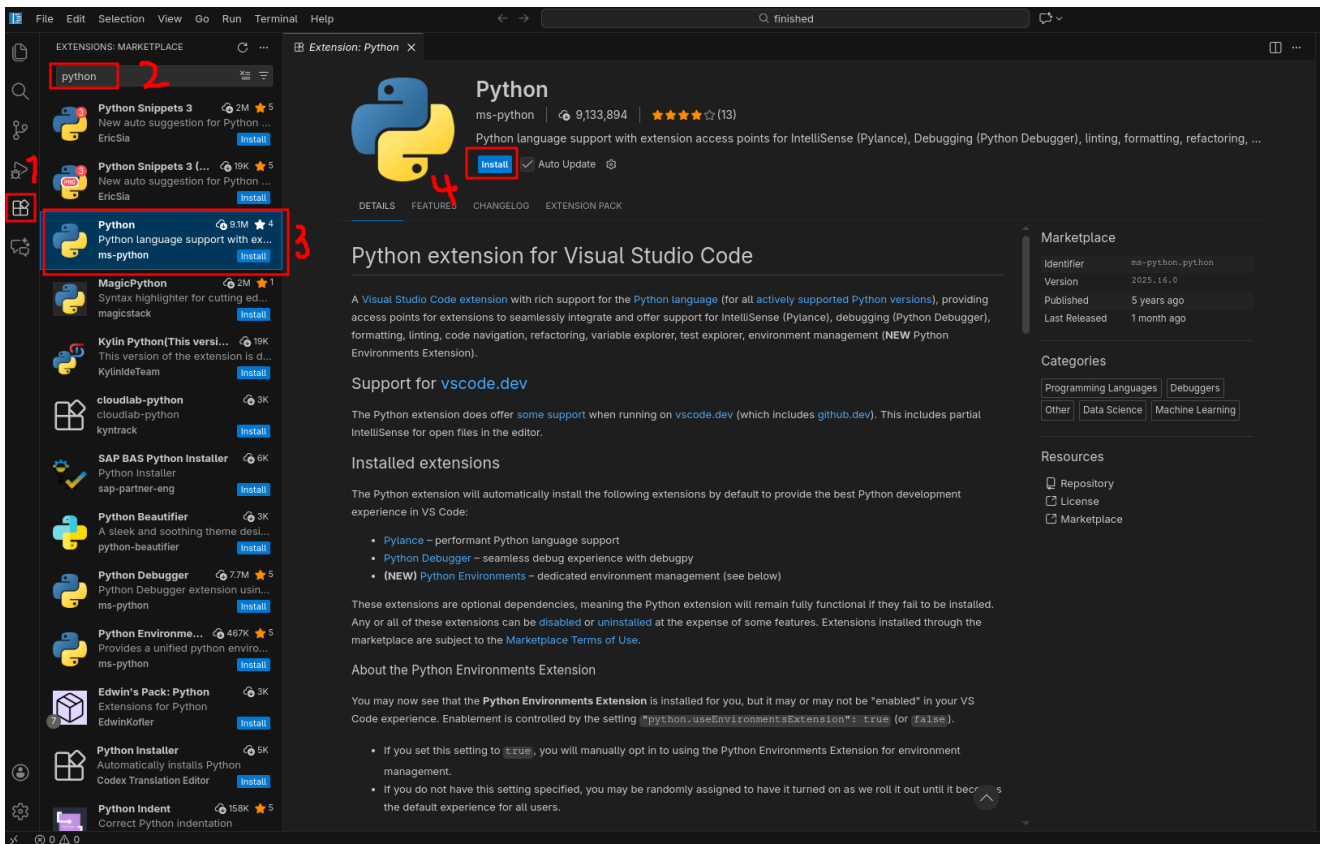


9. После завершения установки нажмите кнопку «Заккрыть» и повторите шаг 6 (убедитесь, что Python установлен в вашей системе).

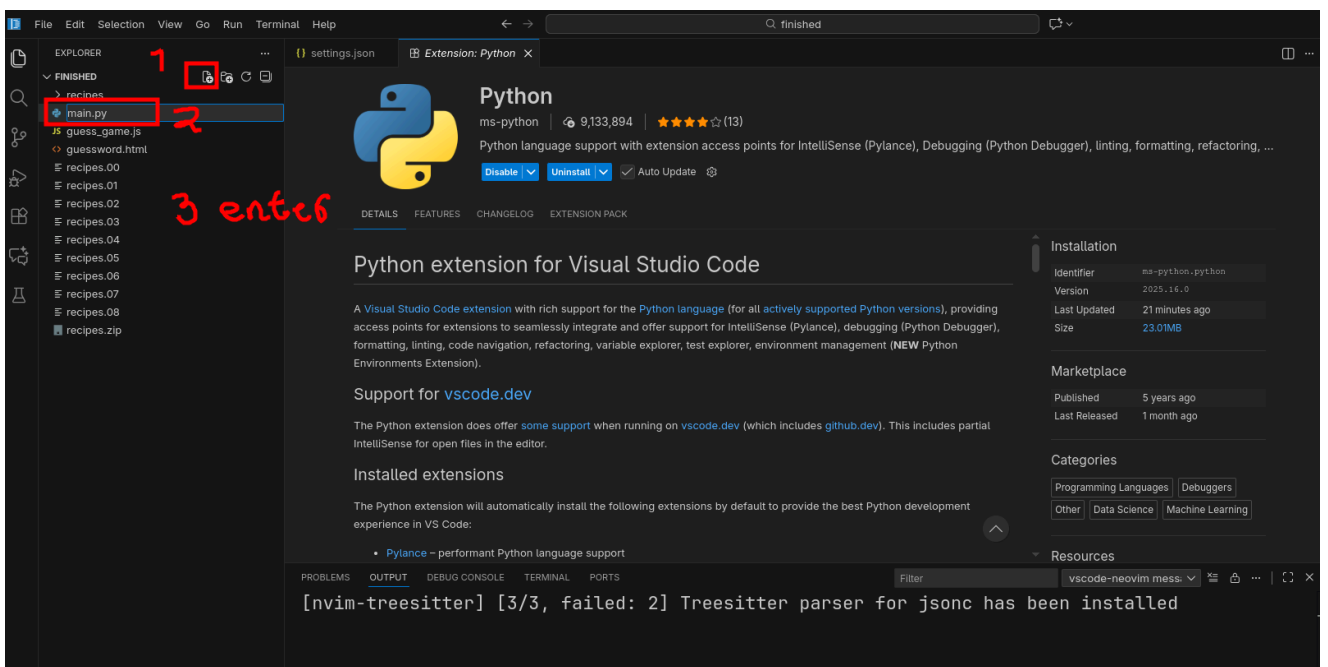


ПРОВЕРКА РАБОТЫ PYTHON В VSCODE

10. В VSCODE перейдите в раздел «Расширения», затем введите в строке поиска `python`, затем выберите расширение «python» (показано на рисунке) без каких-либо дополнительных настроек и установите его.



11. Вернитесь в проводник и создайте новый файл Python с именем `main.py`, нажав кнопку «Создать файл», затем указав имя файла и его расширение (не забудьте расширение «py») и нажав Enter.

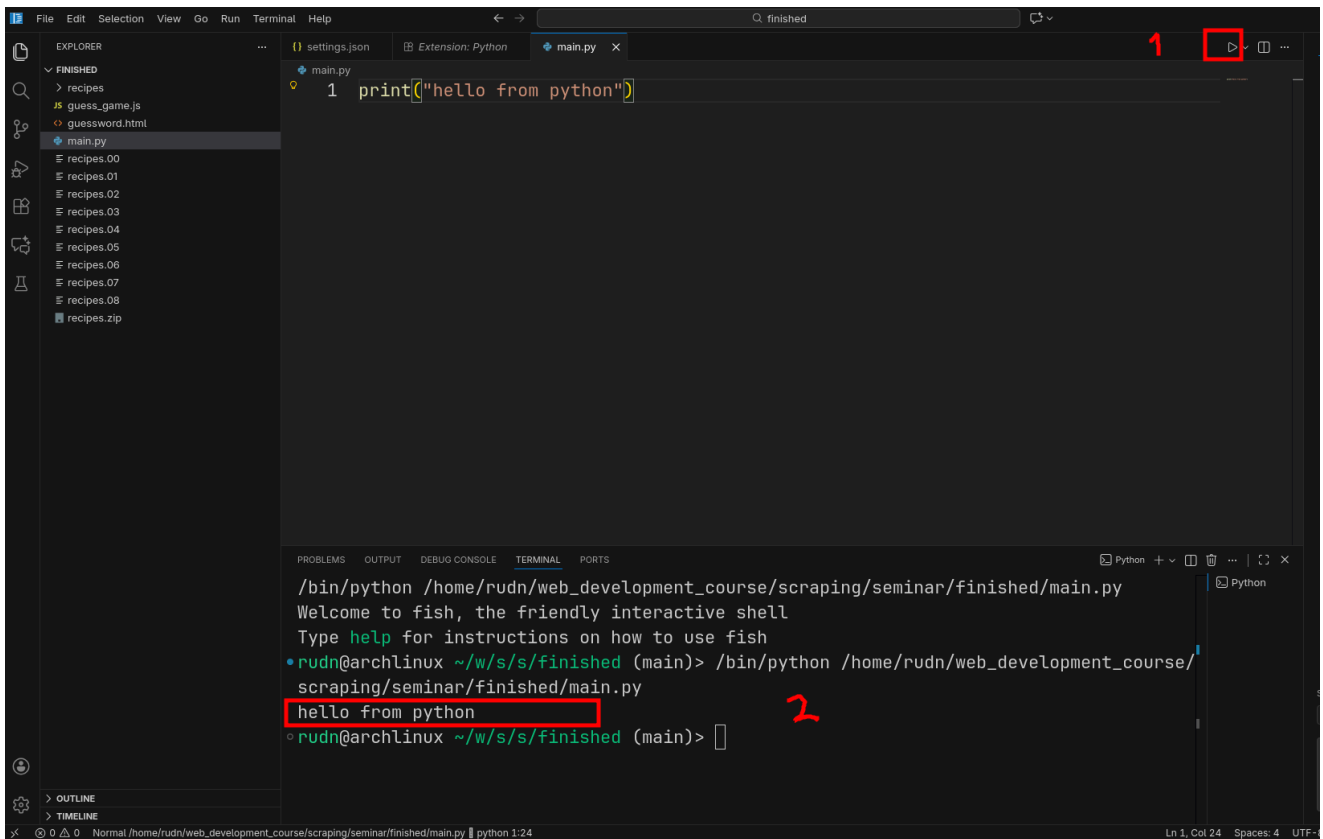


12. Выберите файл `main.py` , щёлкнув по нему, и введите в него следующую строку кода:

```
Python

1 print("hello from python")
```

13. Сохраните файл, нажав `CTRL+s` , затем нажмите кнопку «Run python file» и убедитесь, что в консоли выведено сообщение «hello from python».



ЗАГРУЗИТЕ СВОИ РЕЗУЛЬТАТЫ В TUIS

14. Прикрепите к TUIS следующие скриншоты:

- Работаящий сайт с рецептами (шаг 5)
- Установка версии Python (шаг 6)
- Работаящий Python в VScode (шаг 13) python working on vscode (step 13)