

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
**Федеральное государственное автономное образовательное
учреждение высшего образования
«Севастопольский государственный университет»**

Институт радиоэлектроники и интеллектуальных технических систем
Кафедра «Информатика и управление в технических системах»



МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к выполнению лабораторной работы
«Инструментальные средства для работы с данными»

по дисциплине
«Обработка данных в автоматизированных системах»

*для студентов очной формы обучения направления 27.03.04
«Управление в технических системах» (профиль подготовки
«Интеллектуальные робототехнические системы»)*

Севастополь — 2023

УДК 004.6

Методические указания к выполнению лабораторной работы «Инструментальные средства для работы с данными» по дисциплине «Обработка данных в автоматизированных системах» для студентов очной формы обучения направления 27.03.04 «Управление в технических системах» (профиль подготовки «Интеллектуальные робототехнические системы») / Сост. Альчаков В.В. — Севастополь: Изд-во ФГАОУ ВО «Севастопольский государственный университет», 2023. — с. 42.

Методические указания:

рассмотрены и рекомендованы к изданию решением кафедры «Информатика и управление в технических системах», протокол № X от XX.XX.2023 г.;

допущены учебно-методическим центром ФГАОУ ВО «Севастопольский государственный университет» в качестве методических указаний.

Рецензент:

Крамарь В.А., д-р техн. наук., профессор, профессор кафедры «Информатика и управление в технических системах» СевГУ.

© СевГУ, 2023

© Альчаков В.В., 2023

Содержание

Цель работы.....	4
Порядок выполнения и задание на работу.....	4
Требования к отчёту по лабораторной работе.....	4
Основные теоретические сведения.....	5
Установка пакета Anaconda.....	5
Установка Python, PyCharm.....	14
Работа с блокнотом Jupyter Notebook.....	21
Создание проекта в PyCharm.....	26
Установка пакетов с помощью менеджера пакетов pip.....	30
Библиотека math.....	33
Работа с файлами.....	35
Варианты задания.....	36
Пример выполнения работы.....	37
Контрольные вопросы.....	40
Список использованных источников.....	40
Полезные ссылки.....	41

Цель работы

Изучить современные программные инструментальные средства, используемые для обработки данных. Научиться выполнять установку и настройку рабочего окружения. Освоить работу с блокнотами Jupyter Notebook и с IDE PyCharm. Научиться устанавливать дополнительные библиотеки.

Порядок выполнения и задание на работу

1. Ознакомиться с основными современными программными средствами для анализа данных.
2. Выполнить установку пакета Anaconda.
3. Изучить структуру блокнота Jupyter Notebook.
4. Изучить горячие клавиши для работы с блокнотом Jupyter Notebook.
5. Научиться сохранять и загружать блокноты Jupyter Notebook с диска.
6. Изучить возможности пакета Anaconda для администрирования пакетов.
7. Выполнить установку Python и IDE PyCharm.
8. Изучить структуру проекта PyCharm.
9. Научиться сохранять и загружать проекты PyCharm с диска.
10. Изучить возможности PyCharm для администрирования пакетов.
11. Изучить утилиту pip для администрирования пакетов.
12. Используя Jupyter Notebook и IDE PyCharm написать скрипт для расчета значения математической функции в заданном диапазоне изменения аргумента. Математическая функция должна быть реализована с помощью инструкции def.

Требования к отчёту по лабораторной работе

Отчёт о выполненной лабораторной работе должен содержать:

- титульный лист;
- цель лабораторной работы;
- основные положения;
- вывод по работе.

Отчёт составляется каждым обучающимся индивидуально и должен соответствовать варианту задания, назначенного преподавателем.

В отчёте к данной лабораторной работе необходимо привести описание структуры блокнота Jupyter Notebook и проекта PyCharm, горячие клавиши для работы с блокнотом Jupyter Notebook, сравнительную характеристику двух инструментальных средств для анализа данных (Jupyter Notebook vs PyCharm). Также в отчёте должны быть представлены форматы команд, используемые для установки дополнительных библиотек. В выводе по работе сделать выбор инструментального средства и обосновать своё решение.

Основные теоретические сведения

Описать основные инструменты (Excel, R, Python, Statistica, Tableau, RapidMiner, Orange, Power BI), применяемые для анализа данных. Сделать акцент на Python. Дать общую характеристику Anaconda и PyCharm.

Установка пакета Anaconda

Инструкция по установке для ОС Linux

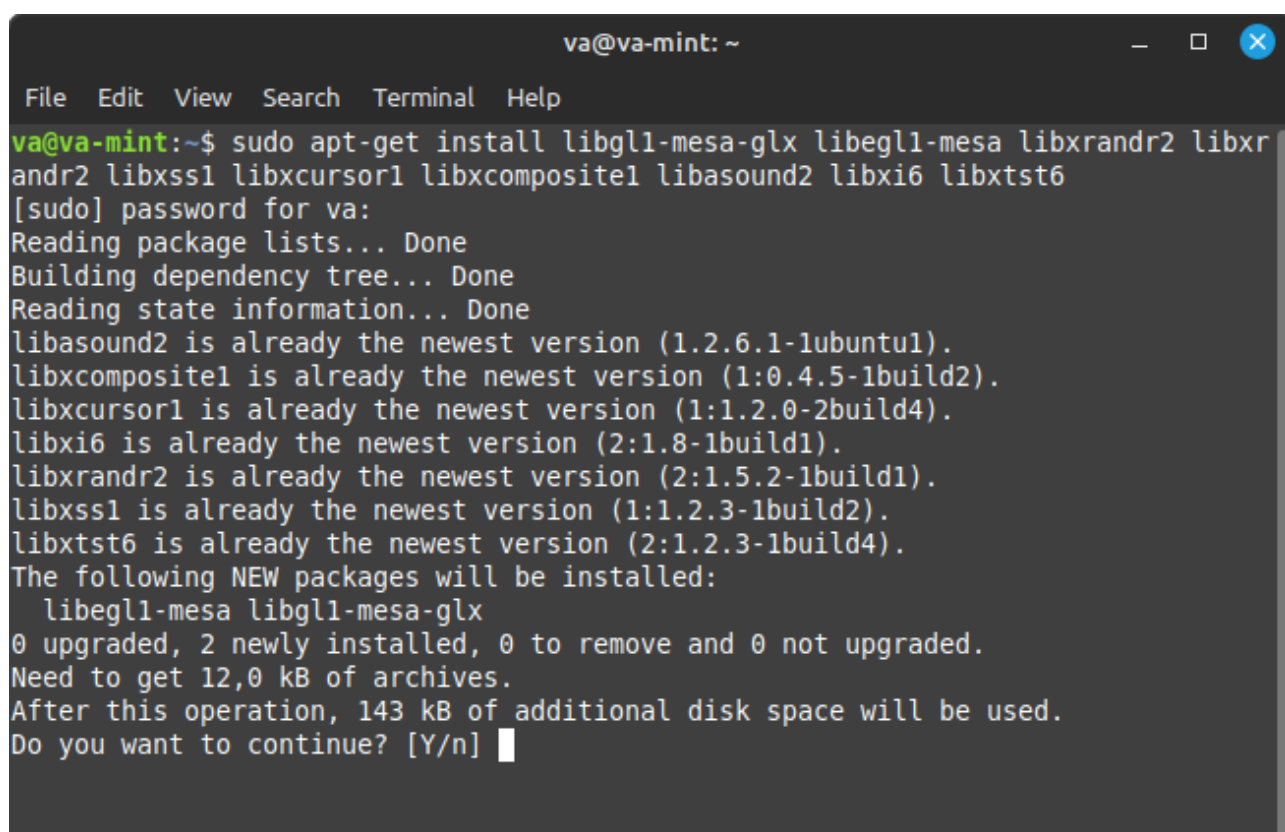
Актуальная инструкция установки пакета Anaconda для операционной системы семейства Linux доступна по ссылке <https://docs.anaconda.com/anaconda/install/linux/>

Далее процесс установки будет описан на примере ОС Linux Mint 21.1 Cinnamon Edition.

Шаг 1. Открыть окно терминала и выполнить установку зависимостей с помощью команды

```
sudo apt-get install libgl1-mesa-glx libegl1-mesa libxrandr2 libxrandr2 libxss1 libxcursor1 libxcomposite1 libasound2 libxi6 libxtst6
```

В ходе установки подтвердить запрос программы установщика, введя символ **Y** в окне терминала.



```
va@va-mint: ~  
File Edit View Search Terminal Help  
va@va-mint:~$ sudo apt-get install libgl1-mesa-glx libegl1-mesa libxrandr2 libx  
andr2 libxss1 libxcursor1 libxcomposite1 libasound2 libxi6 libxtst6  
[sudo] password for va:  
Reading package lists... Done  
Building dependency tree... Done  
Reading state information... Done  
libasound2 is already the newest version (1.2.6.1-1ubuntu1).  
libxcomposite1 is already the newest version (1:0.4.5-1build2).  
libxcursor1 is already the newest version (1:1.2.0-2build4).  
libxi6 is already the newest version (2:1.8-1build1).  
libxrandr2 is already the newest version (2:1.5.2-1build1).  
libxss1 is already the newest version (1:1.2.3-1build2).  
libxtst6 is already the newest version (2:1.2.3-1build4).  
The following NEW packages will be installed:  
  libegl1-mesa libgl1-mesa-glx  
0 upgraded, 2 newly installed, 0 to remove and 0 not upgraded.  
Need to get 12,0 kB of archives.  
After this operation, 143 kB of additional disk space will be used.  
Do you want to continue? [Y/n]
```

Рис. 1 — Установка зависимостей

Шаг 2. Дождаться окончания процесса установки зависимостей, после чего перейти на страницу загрузки пакета Anaconda для Linux по ссылке <https://www.anaconda.com/download/#linux>

На странице загрузке нажать кнопку **Download** (Рис. 2)

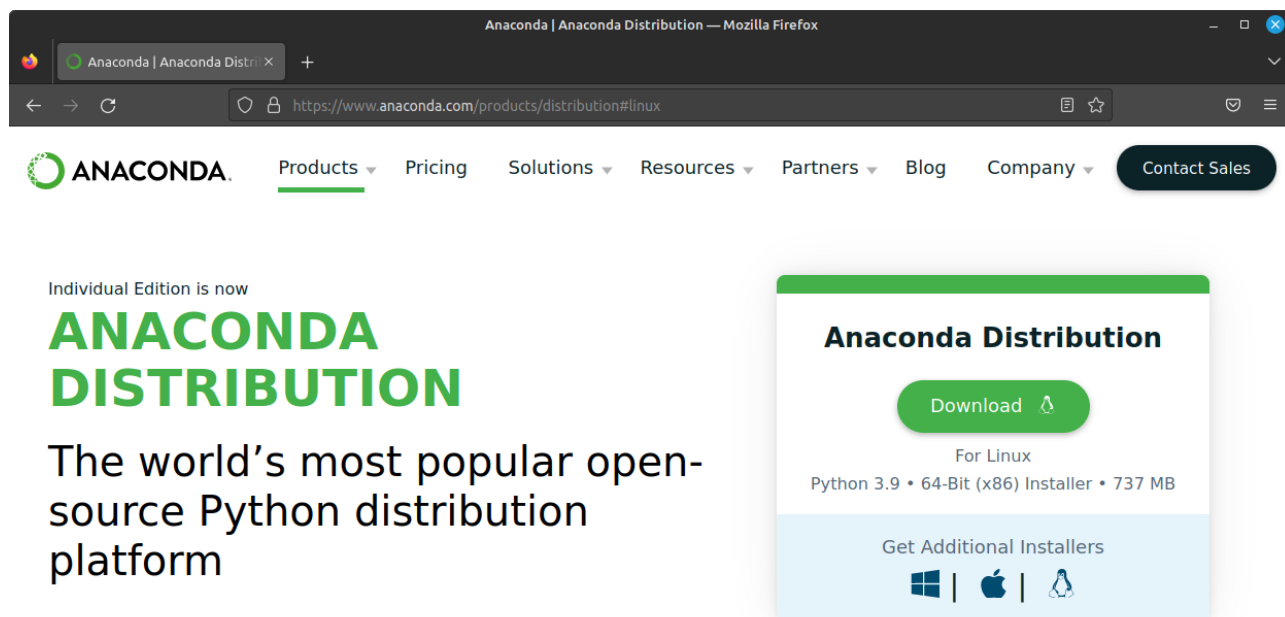


Рис. 2 — Страница загрузки пакета Anaconda

Шаг 3. По завершению загрузки, перейти в папку Downloads и открыть содержимое папки в окне терминала. Для этого вызвать контекстное меню и выбрать опцию **Open in Terminal** (Рис. 3). После чего выполнить команду `bash Anaconda3-2022.10-Linux-x86_64.sh`

Внимание: имя файла установщика может отличаться и должно соответствовать актуальной версии пакета. Содержимое папки может быть получено с помощью команды `ls` (Рис. 4).

После запуска команды необходимо нажать кнопку **Enter** и следовать дальнейшим указаниям установщика (ознакомиться с содержанием лицензионного соглашения). Дождаться окончания установки пакета.

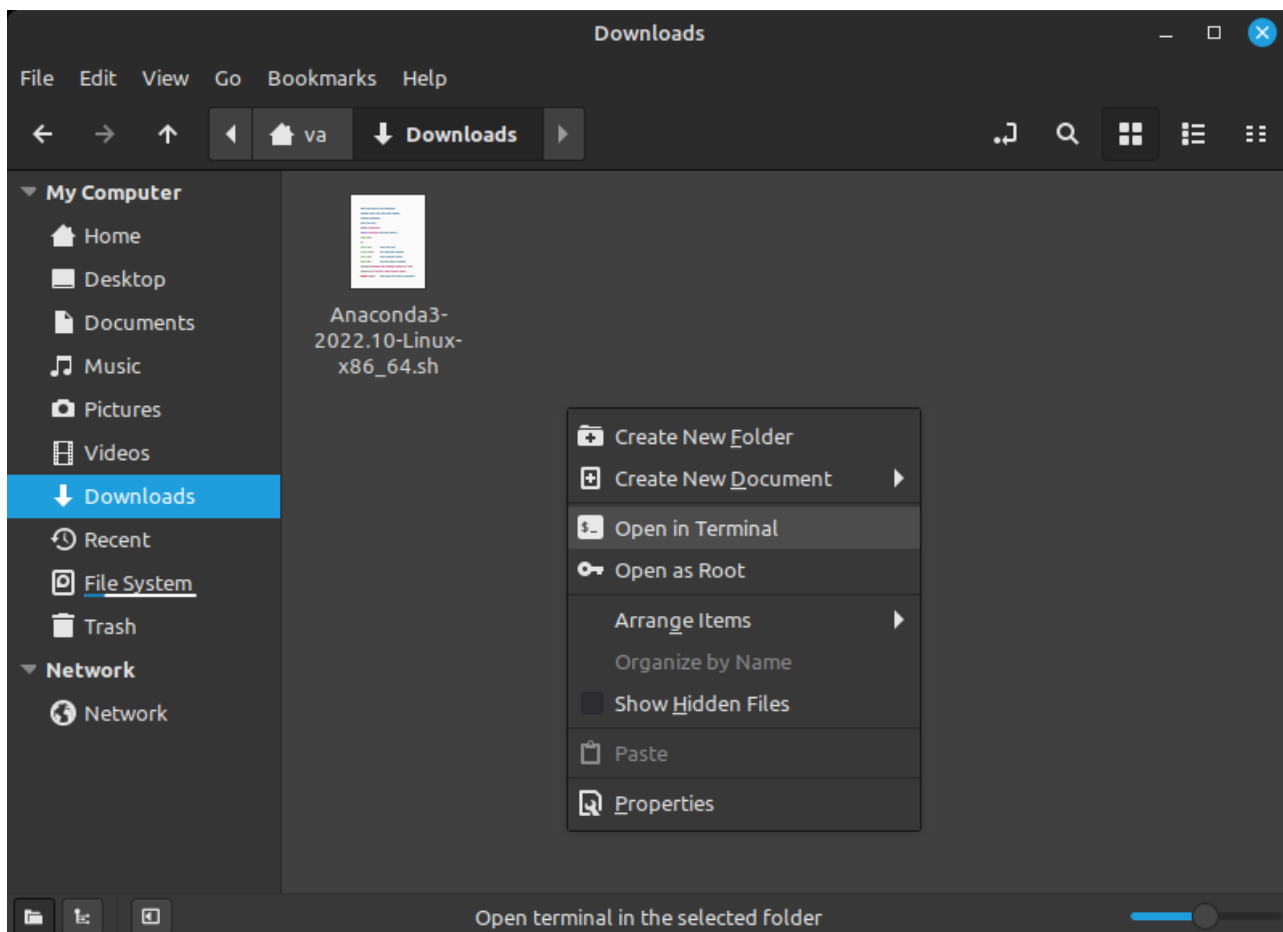


Рис. 3 — Открытие папки в окне терминала

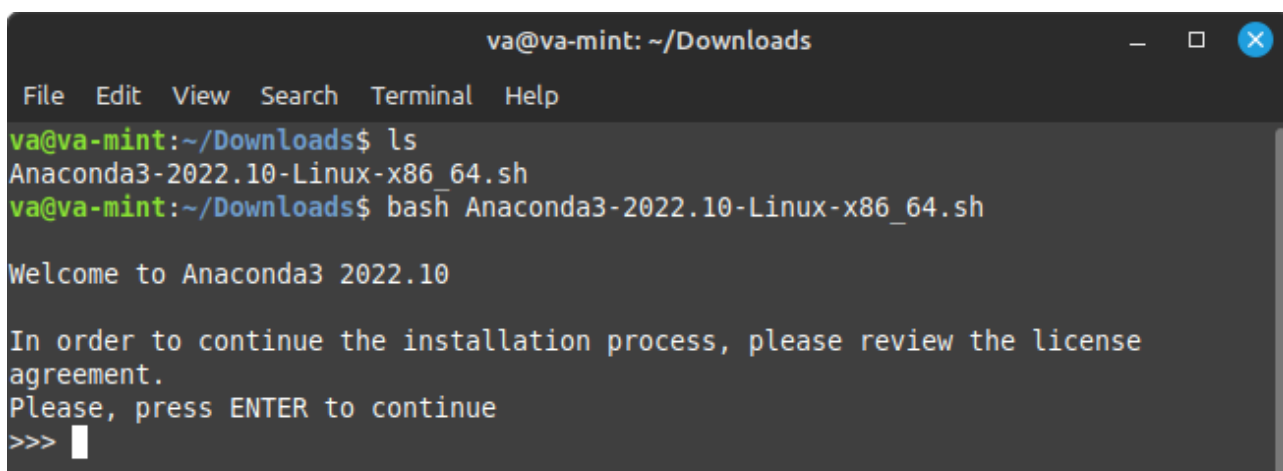


Рис. 4 — Запуск скрипта установки пакета Anaconda

Шаг 4. По окончании установки может быть запущена оболочка, которая используется для удобства запуска отдельных программ пакета — Anaconda Navigator (далее навигатор) (Рис. 5). Для запуска навигатора в окне терминала необходимо выполнить команду `anaconda-navigator`

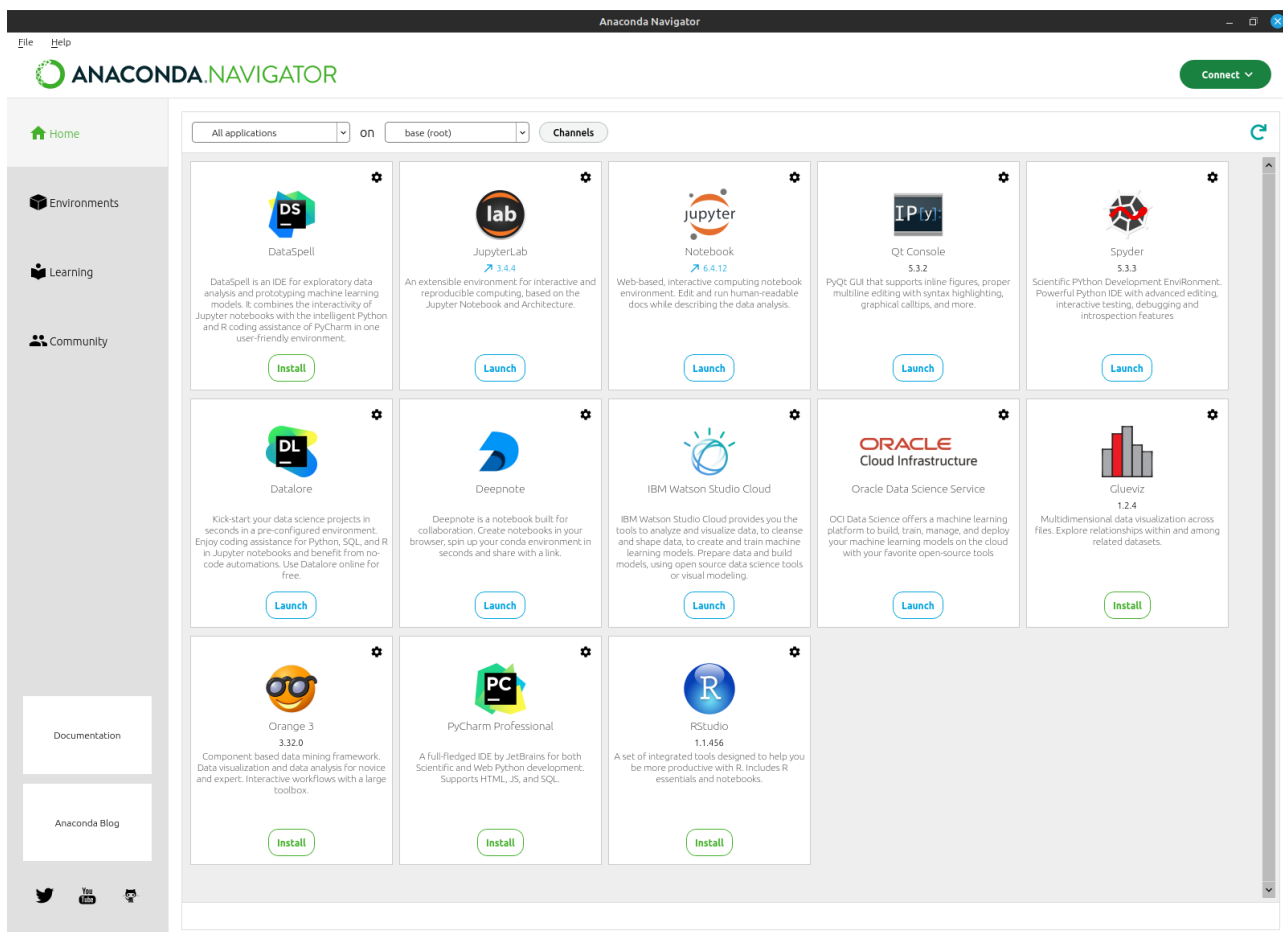


Рис. 5 — Окно Anaconda Navigator

Как сказано выше, навигатор может быть использован для быстрого доступа к инструментам, входящим в состав пакета Anaconda, в том числе и к Jupyter Notebook. Для запуска выбранного инструмента необходимо нажать кнопку **Launch** в соответствующей панели (Рис. 6)

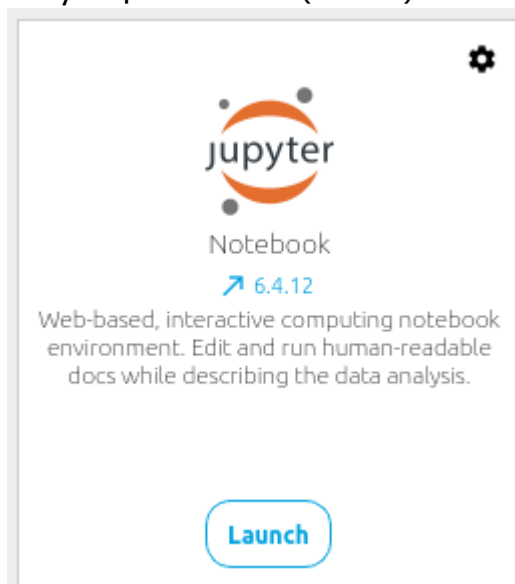


Рис. 6 — Панель запуска Jupyter Notebook

Инструкция по установке для ОС Windows

Актуальная версия установочного пакета Anaconda для операционной системы семейства Windows доступна по ссылке <https://www.anaconda.com/products/distribution>

Шаг 1. Скачать установочный пакет, нажав на кнопку Download на странице загрузки

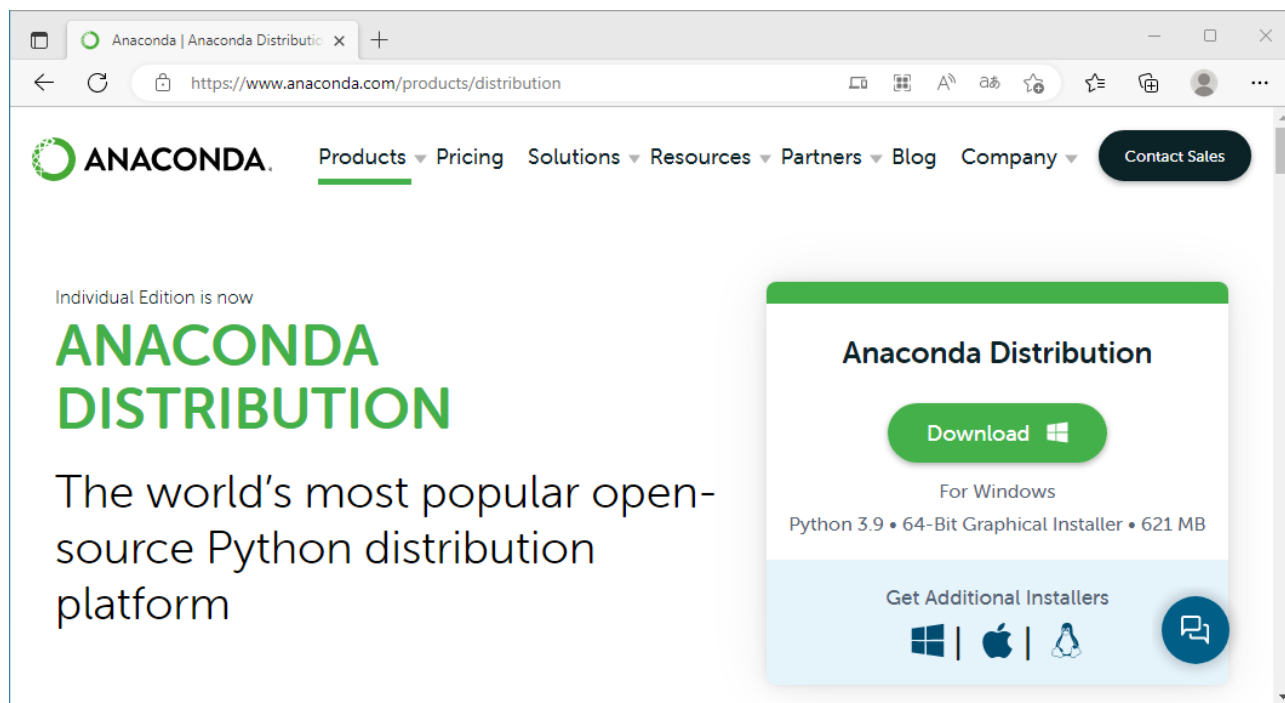


Рис. 7 — Страница загрузки пакета Anaconda для ОС Windows

Шаг 2. По окончании загрузки запустить файл **Anaconda3-2022.10-Windows-x86_64.exe** на исполнение.

Внимание: в зависимости от версии пакета имя скачанного файла может отличаться.

Далее необходимо следовать указаниям программы-установщика. Параметры установки рекомендуется оставить без изменения.

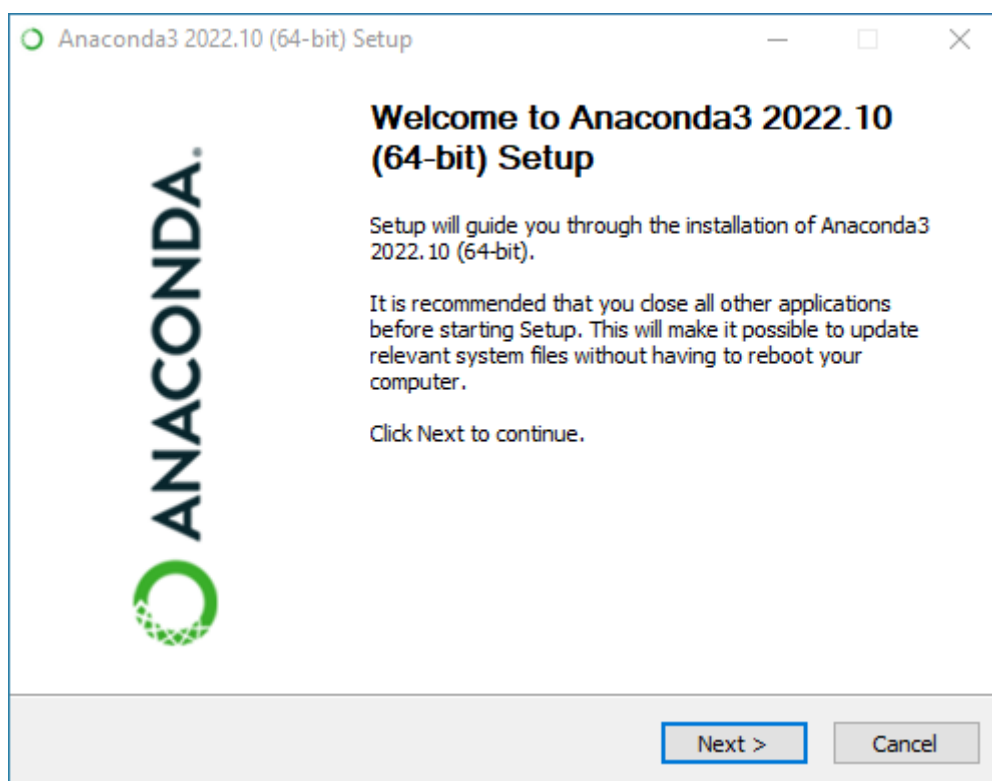


Рис. 8 — Страница приветствия программы-установщика Anaconda

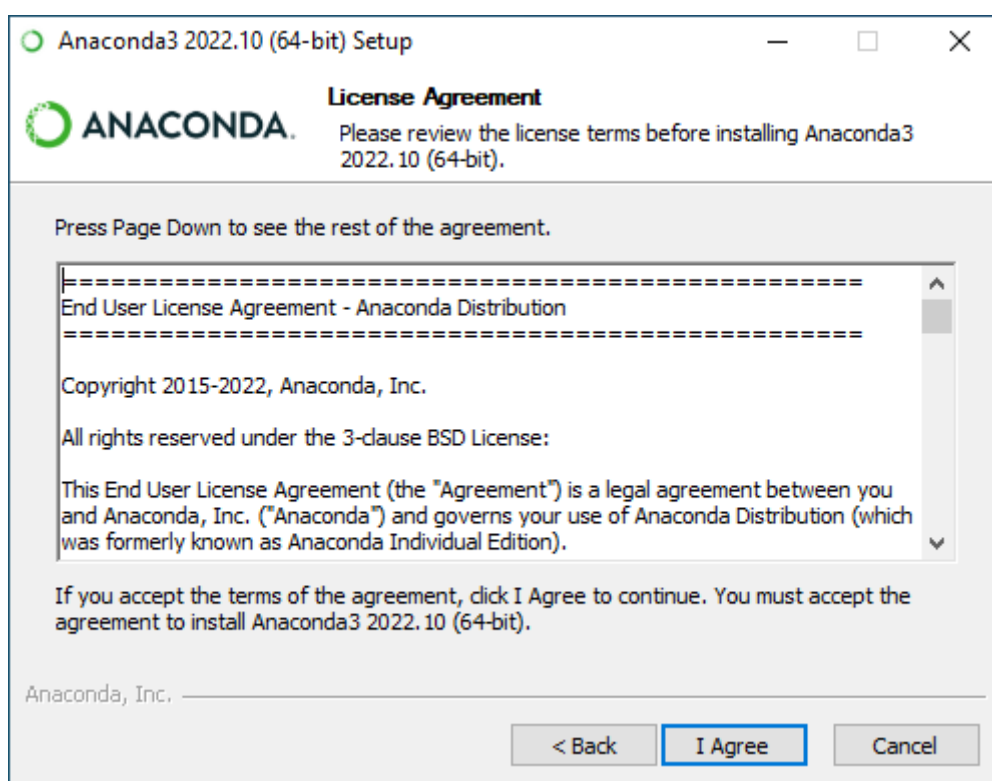


Рис. 9 — Страница лицензионного соглашения Anaconda

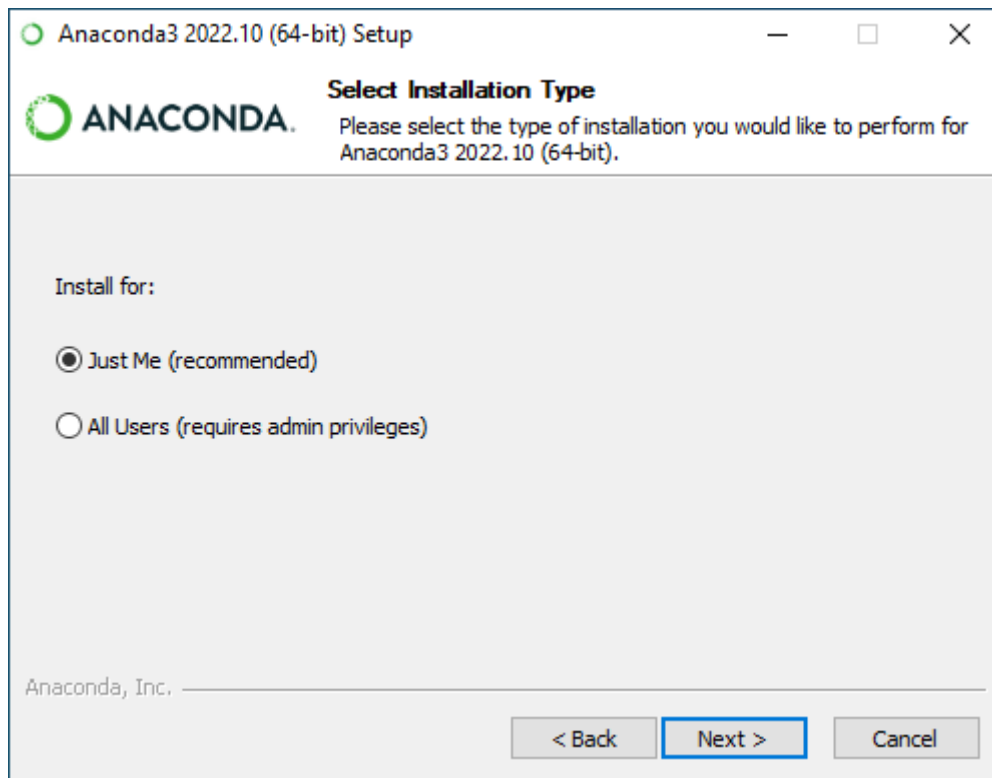


Рис. 10 — Страница конфигурации параметров установки Anaconda

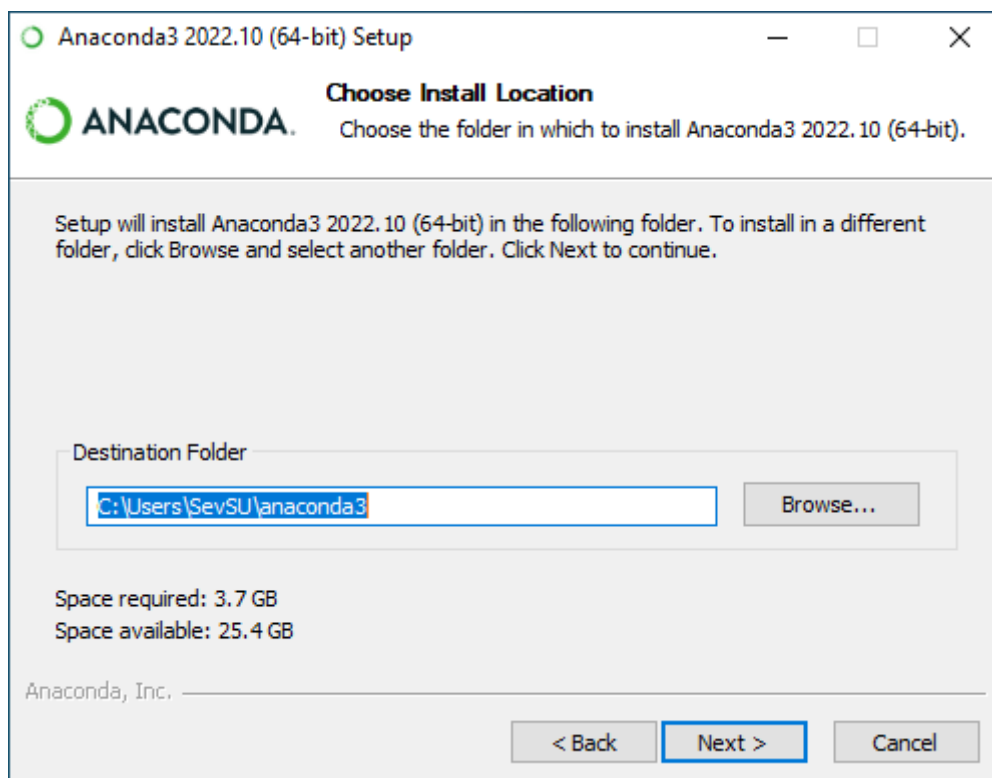


Рис. 11 — Страница задания папки для установки Anaconda

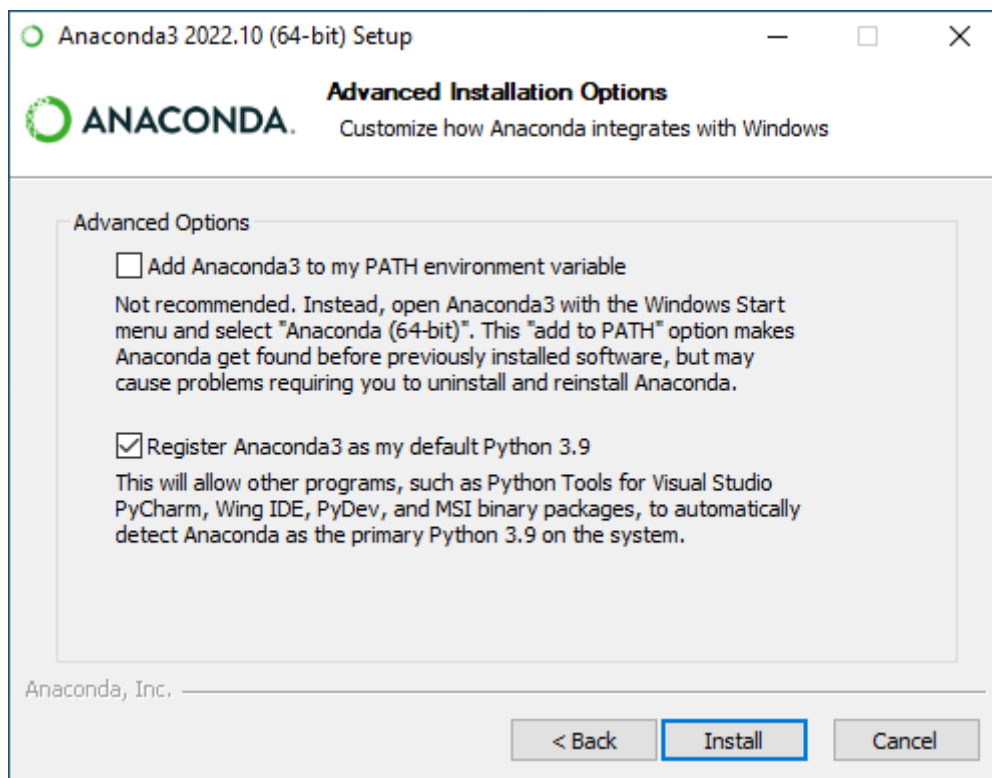


Рис. 12 — Страница конфигурации параметров установки Anaconda

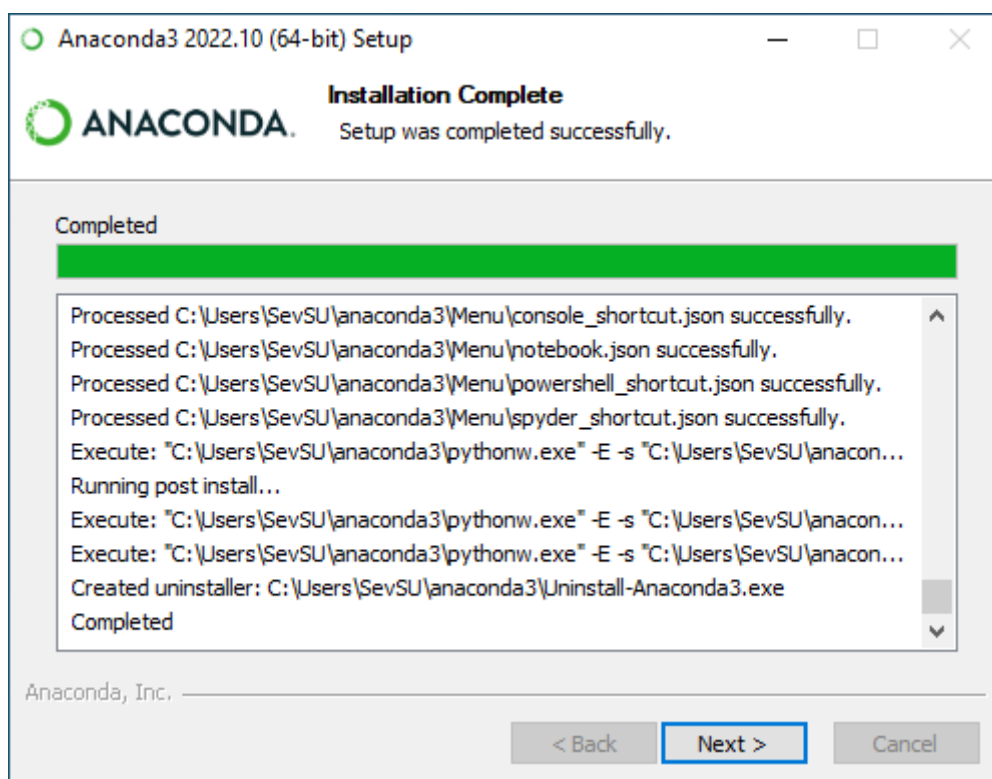


Рис. 13 — Страница прогресса установки Anaconda

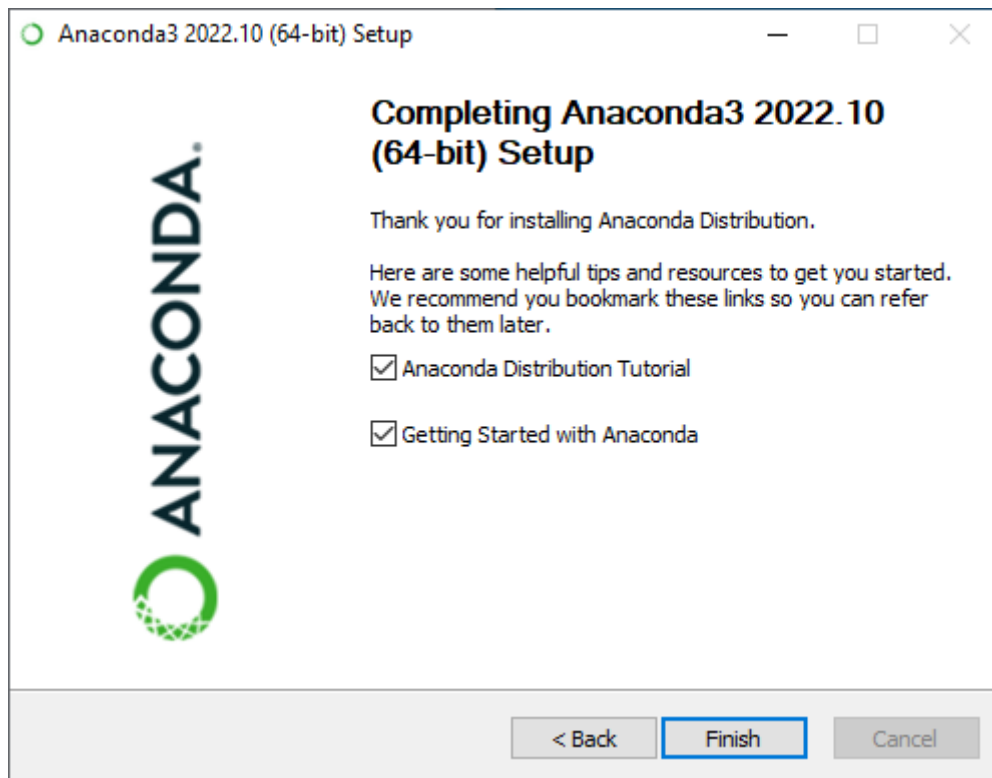


Рис. 14 — Завершающая страница программы-установщика Anaconda

Шаг 3. Для запуска менеджера Anaconda можно воспользоваться системным меню ОС Windows

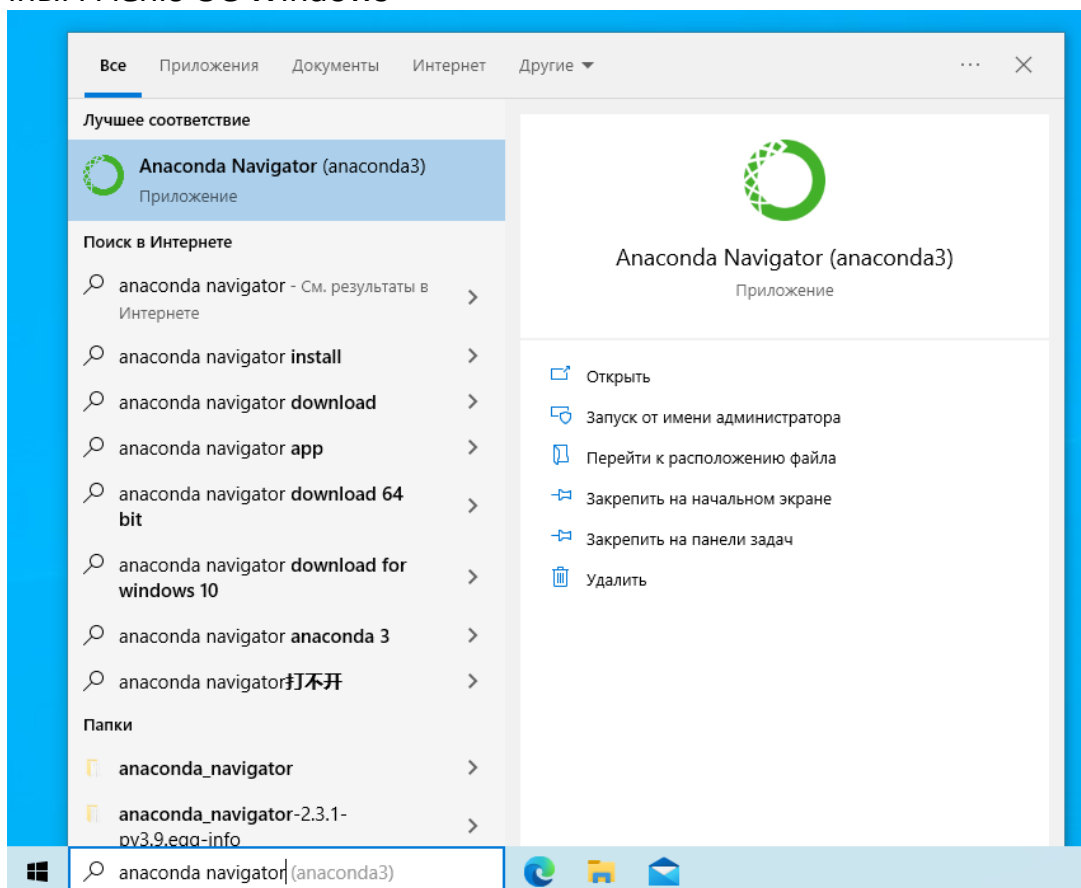


Рис. 15 — Запуск менеджера Anaconda Navigator из системного меню Windows

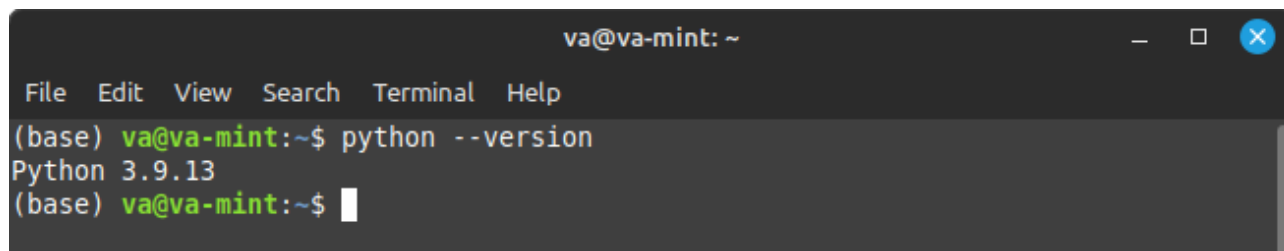
Установка Python, PyCharm

Инструкция по установке для ОС Linux

Для дистрибутива Linux Mint Python входит в набор предустановленного программного обеспечения. Для того чтобы узнать версию установленного Python нужно запустить окно терминала и ввести команду

```
python --version
```

В окне терминала будет выведена информация о текущей версии как это показано на рис. 16.



```
va@va-mint: ~  
File Edit View Search Terminal Help  
(base) va@va-mint:~$ python --version  
Python 3.9.13  
(base) va@va-mint:~$
```

Рис. 16 — Проверка версии Python

Для установки Python на другие дистрибутивы Linux рекомендуется обратиться к справочной системе на сайте разработчика ОС.

Установка IDE PyCharm может быть выполнена несколькими способами. Самый простой заключается в установке с помощью встроенного менеджера программ (Software Manager). Для этого необходимо запустить Software Manager и в строке поиска ввести PyCharm. Если соответствующий пакет будет найден, то дальнейшая установка не вызовет никаких затруднений. Если пакет отсутствует, то необходимо выполнить дополнительные действия.

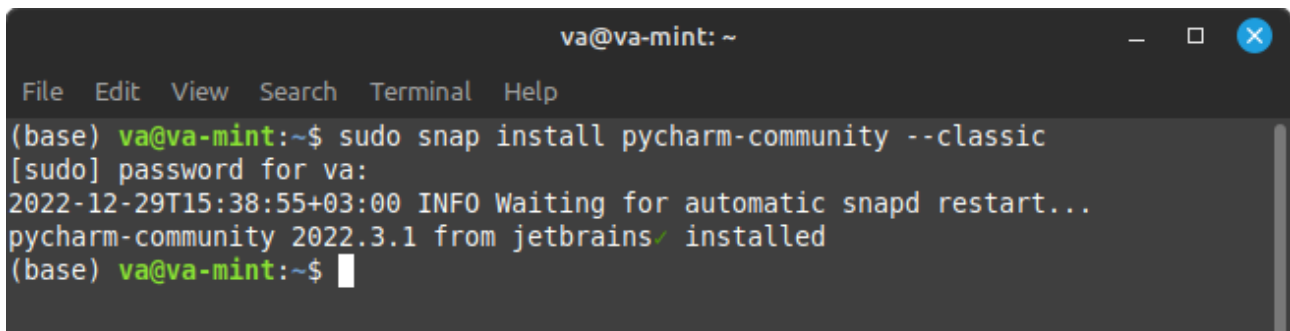
Шаг 1. Включить поддержку установки snap пакетов. Для этого выполнить последовательность команд, приведённую ниже. Последняя команда выполнит перезагрузку системы после завершения конфигурации системных настроек.

```
sudo mv /etc/apt/preferences.d/nosnap.pref ~/Documents/nosnap.backup  
sudo apt update  
sudo apt install snapd  
sudo reboot
```

Шаг 2. После перезагрузки системы выполнить команду и дождаться окончания установки (Рис. 8)

```
sudo snap install pycharm-community --classic
```

Внимание: устанавливать необходимо PyCharm Community Edition, поскольку данная версия IDE является бесплатной и не требует лицензирования для использования.



```
va@va-mint: ~  
File Edit View Search Terminal Help  
(base) va@va-mint:~$ sudo snap install pycharm-community --classic  
[sudo] password for va:  
2022-12-29T15:38:55+03:00 INFO Waiting for automatic snapd restart...  
pycharm-community 2022.3.1 from jetbrains✓ installed  
(base) va@va-mint:~$
```

Рис. 17 — Установка PyCharm с помощью snap пакета

После окончания установки PyCharm будет доступен через стандартное меню операционной системы (Рис. 9)

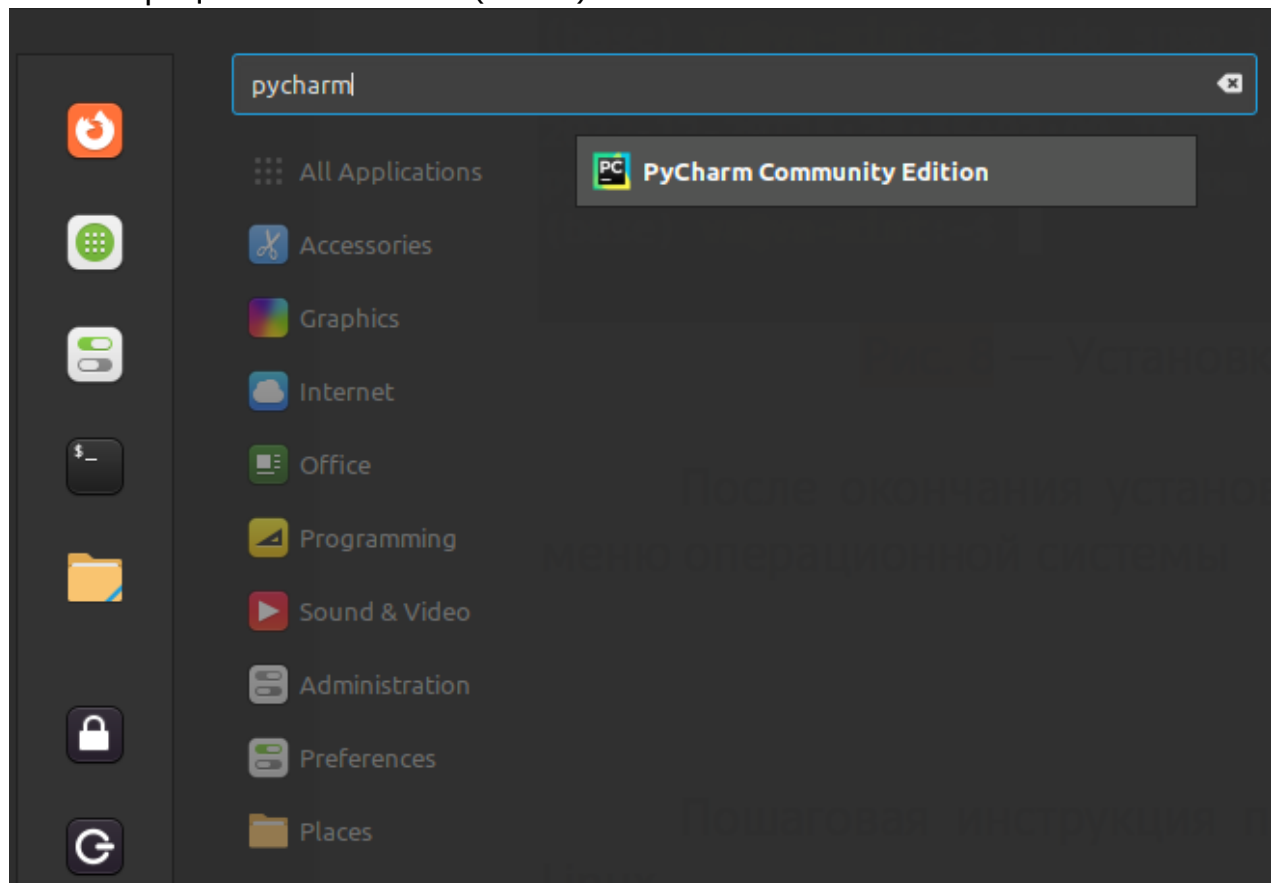


Рис. 18 — Результат установки PyCharm

Инструкция по установке для ОС Windows

Установочный пакет для актуальной версии языка Python доступен по ссылке <https://www.python.org/downloads/>

Шаг 1. Скачать установочный пакет, нажав на кнопку **Download** на странице загрузки

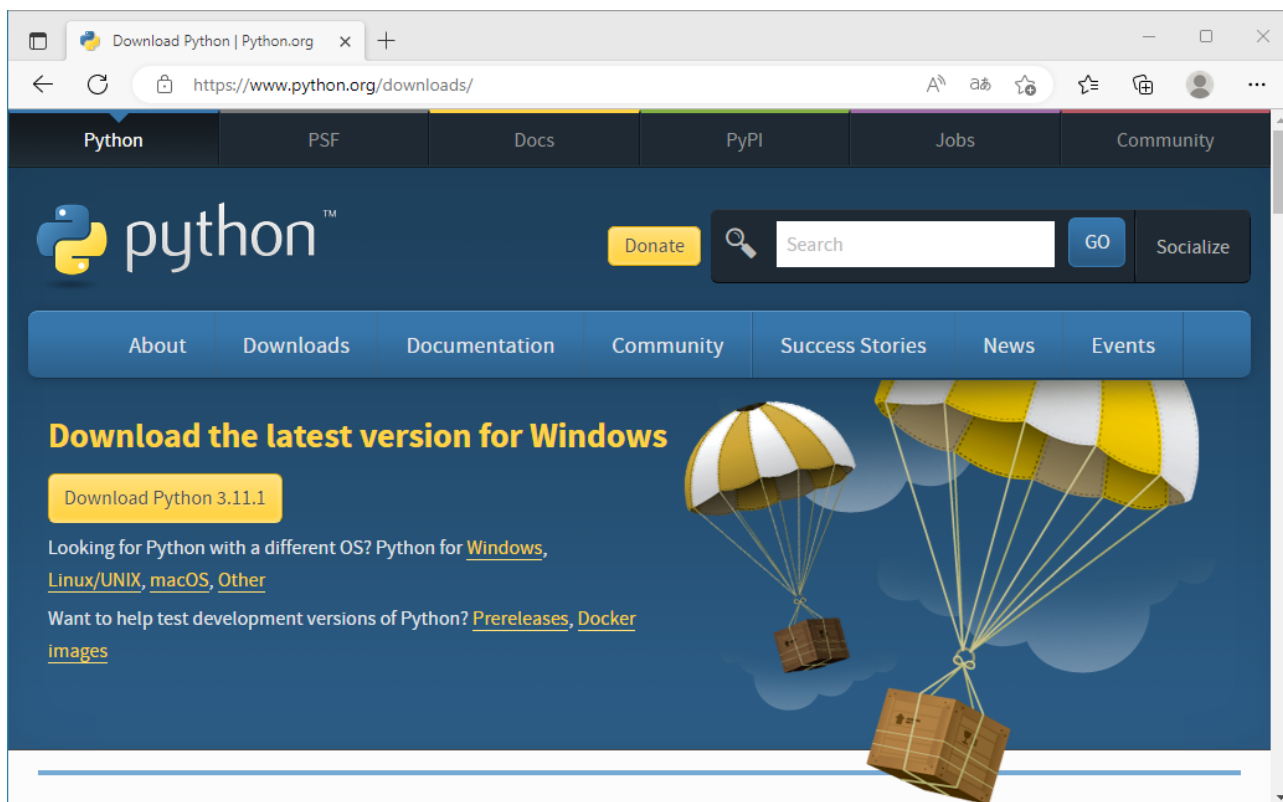


Рис. 19 — Страница загрузки установочного пакета Python для ОС Windows

Шаг 2. По окончании загрузки запустить файл **python-3.11.1-amd64.exe** на исполнение.

Внимание: в зависимости от версии пакета имя скачанного файла может отличаться.

В появившемся окне программы установщика отметить опцию **Add python.exe to PATH** для добавления пути к Python в системную переменную, содержащую список директорий, в которых ОС будет искать исполняемый файл при вызове команды из консоли. Нажать кнопку **Install Now**.

Далее следовать указаниям программы-установщика.

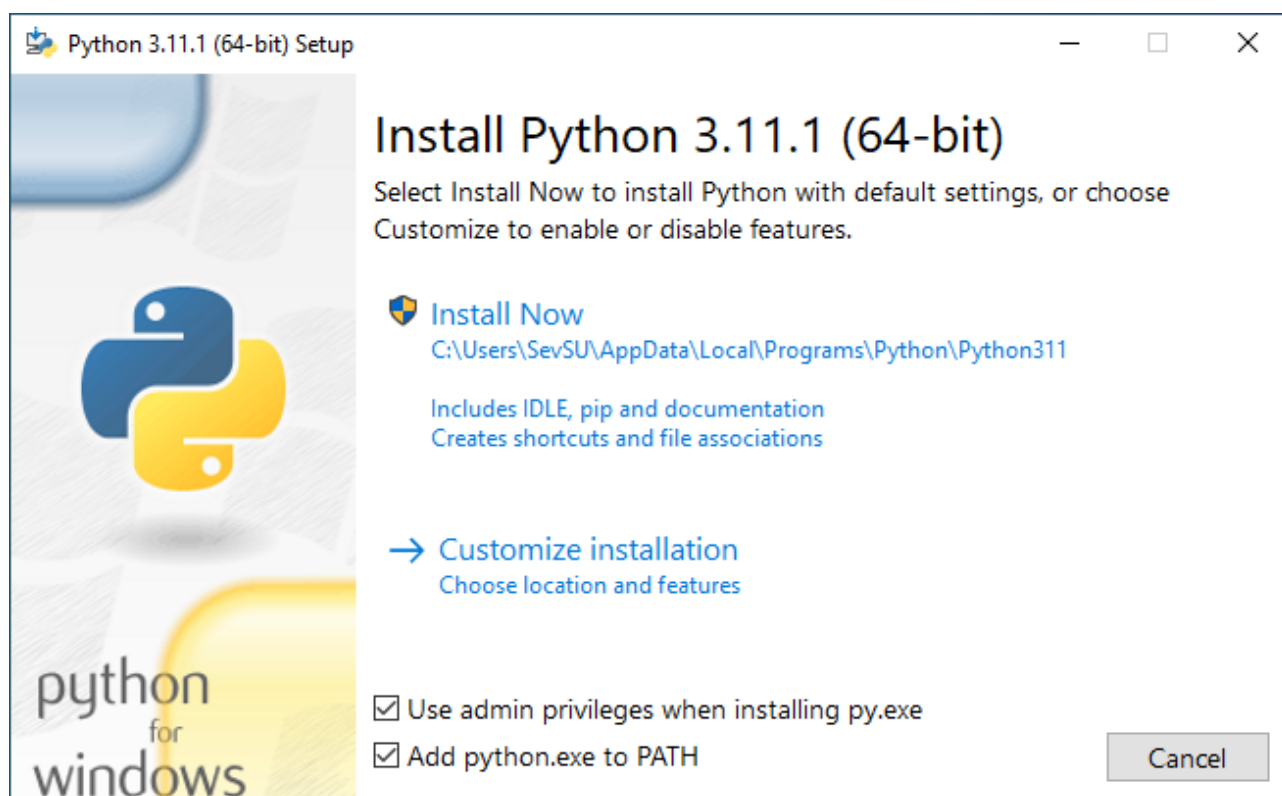


Рис. 20 — Начальная страница программы установщика Python

Шаг 3. Дождаться окончания процесса установки. На последнем шаге установки, нажать на кнопку **Disable path length limit** после чего окно установщика можно закрыть.

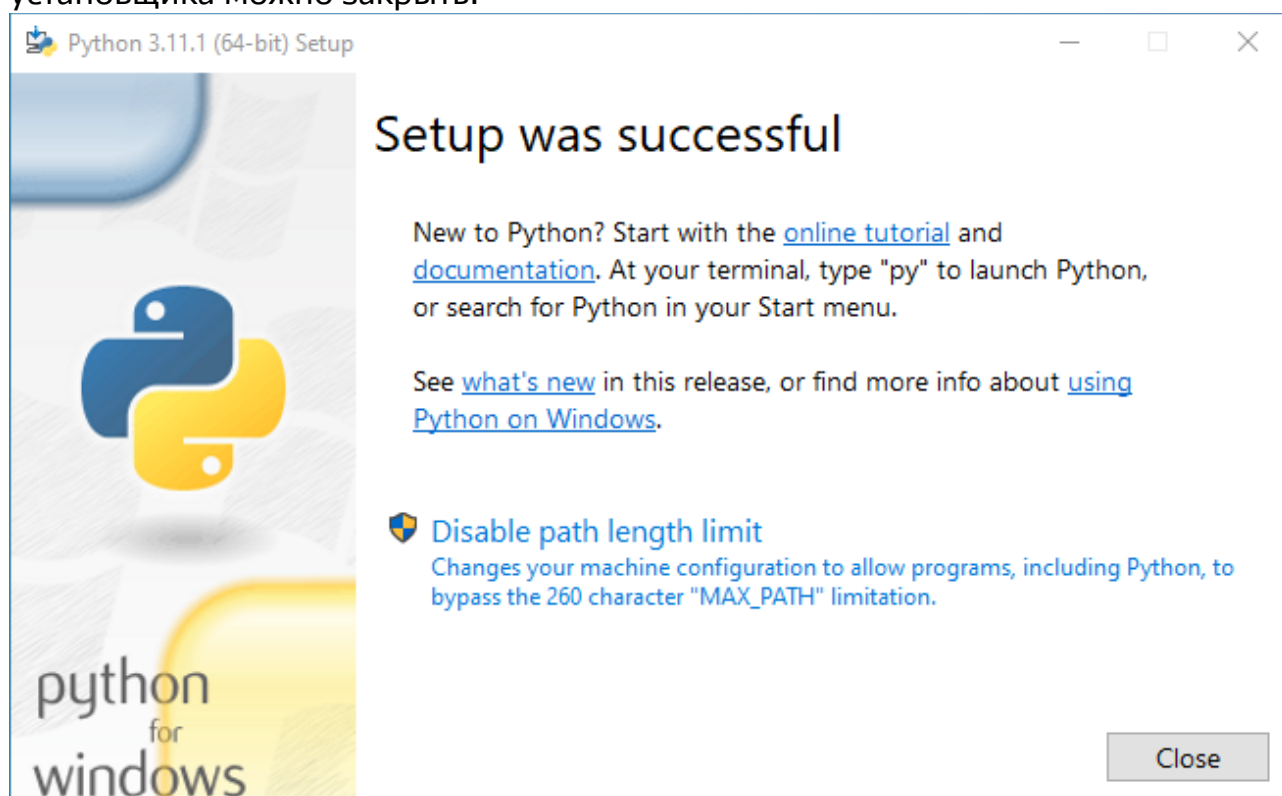
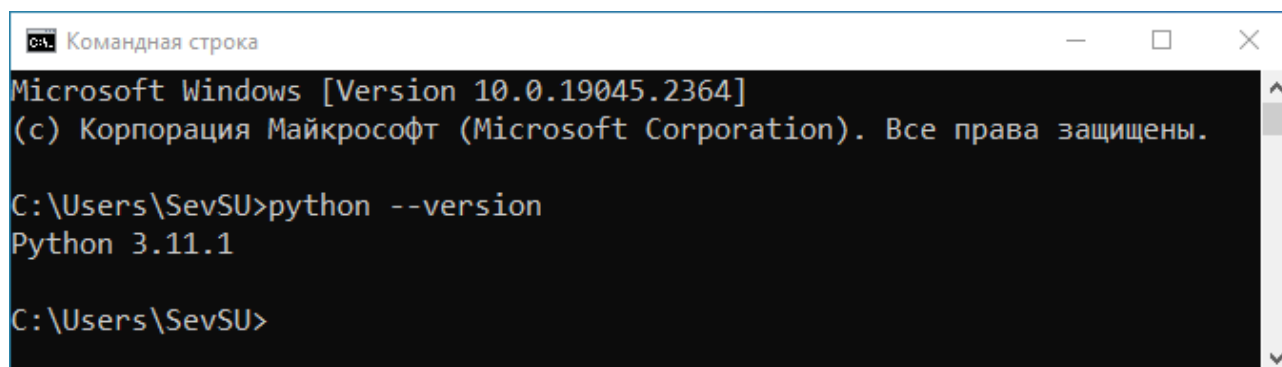


Рис. 21 — Завершающая страницы программы-установщика Python

Шаг 4. Проверить установленную версию Python можно с помощью командной строки Windows открыв окно терминала и введя команду `python --version`



```
Командная строка
Microsoft Windows [Version 10.0.19045.2364]
(c) Корпорация Майкрософт (Microsoft Corporation). Все права защищены.

C:\Users\SevSU>python --version
Python 3.11.1

C:\Users\SevSU>
```

Рис. 22 — Проверка версии Python

Для установки IDE PyCharm необходимо посетить страницу установки <https://www.jetbrains.com/pycharm/download/#section=windows> и загрузить актуальную версию программы-установщика.

Шаг 1. Скачать установочный пакет, нажав на кнопку **Download** на странице загрузки. Скачивать рекомендуется Community версию, поскольку данная версия продукта является бесплатной и не требует процедуры лицензирования.

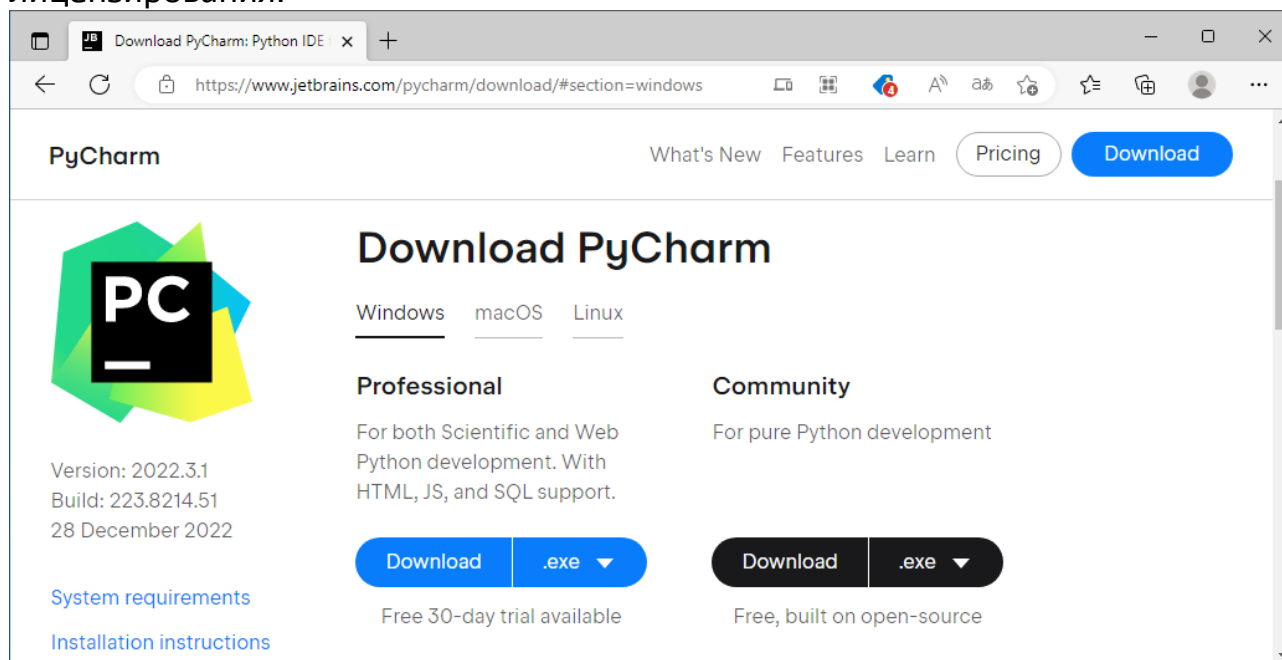


Рис. 23 — Страница загрузки PyCharm

Шаг 23. По окончании загрузки запустить файл **pycharm-community-2022.3.1.exe** на исполнение.

Внимание: в зависимости от версии пакета имя скачанного файла может отличаться.

В появившемся окне программы-установщика

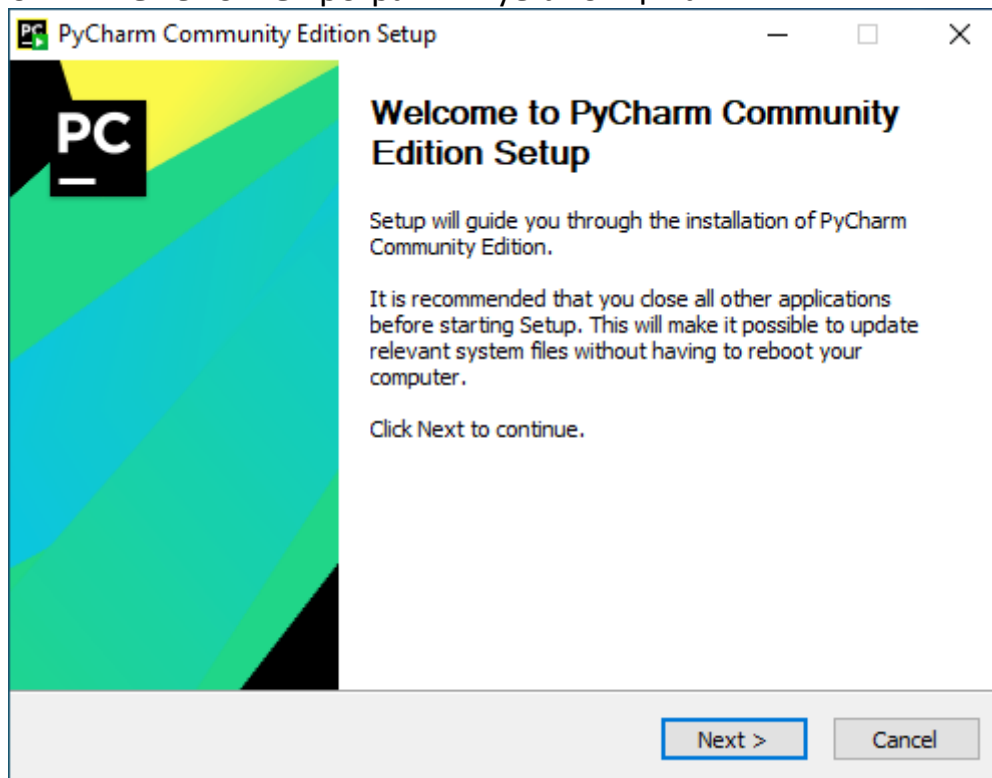


Рис. 24 — Стартовое окно программ-установщика PyCharm

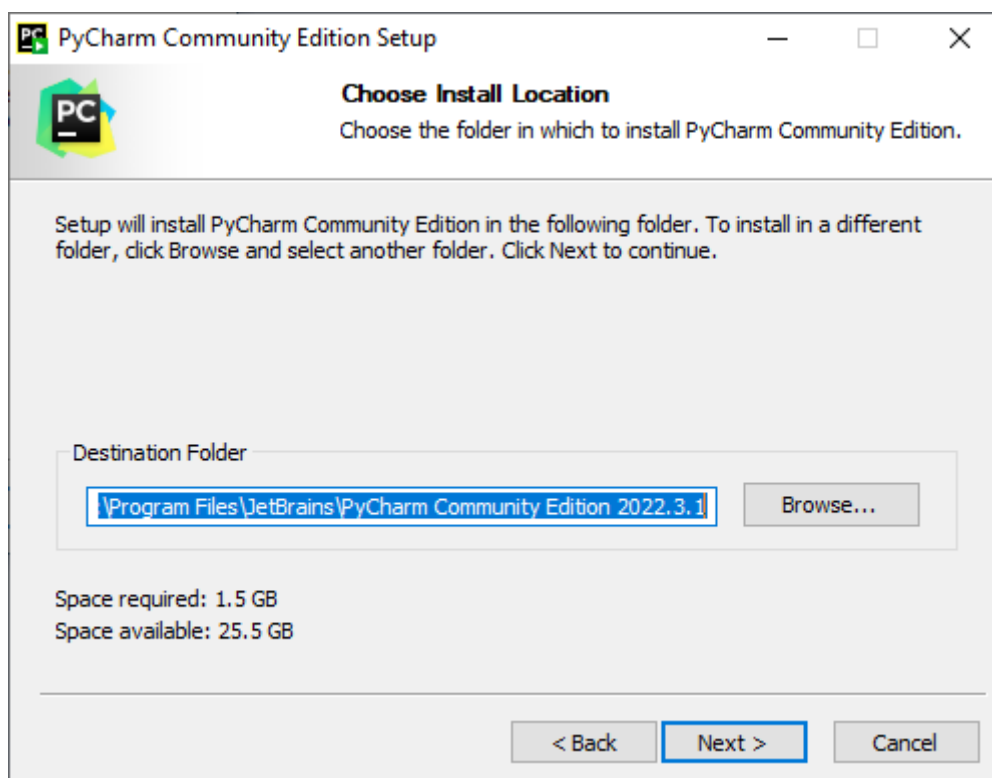


Рис. 25 — Выбор папки установки PyCharm

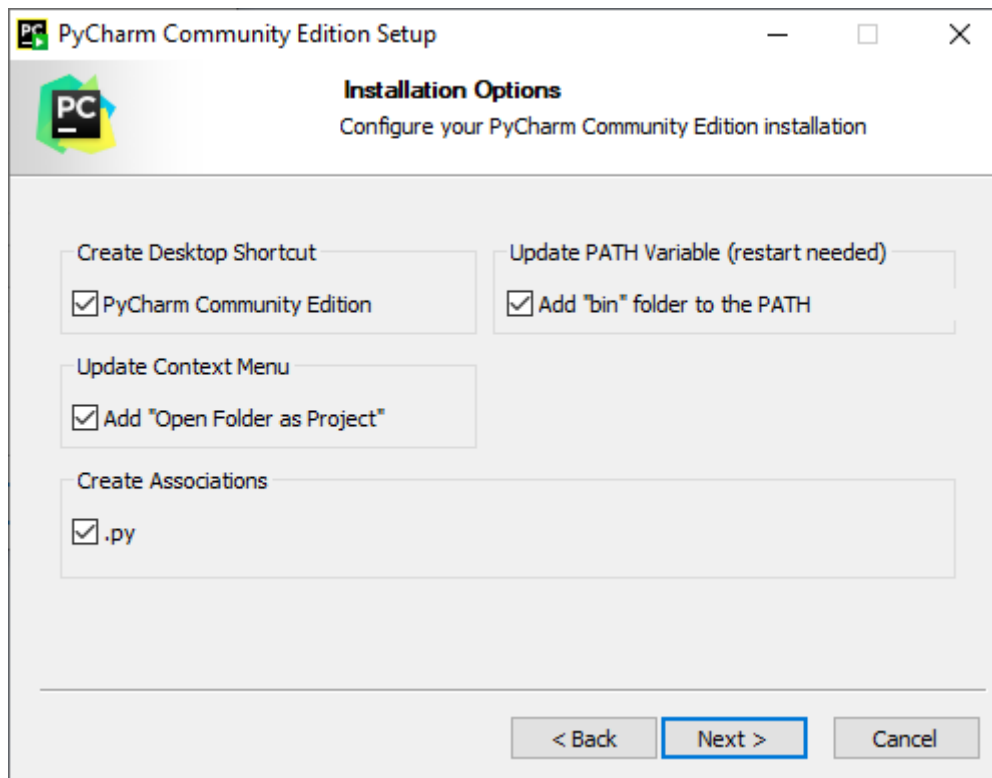


Рис. 26 — Окно настройки установщика PyCharm

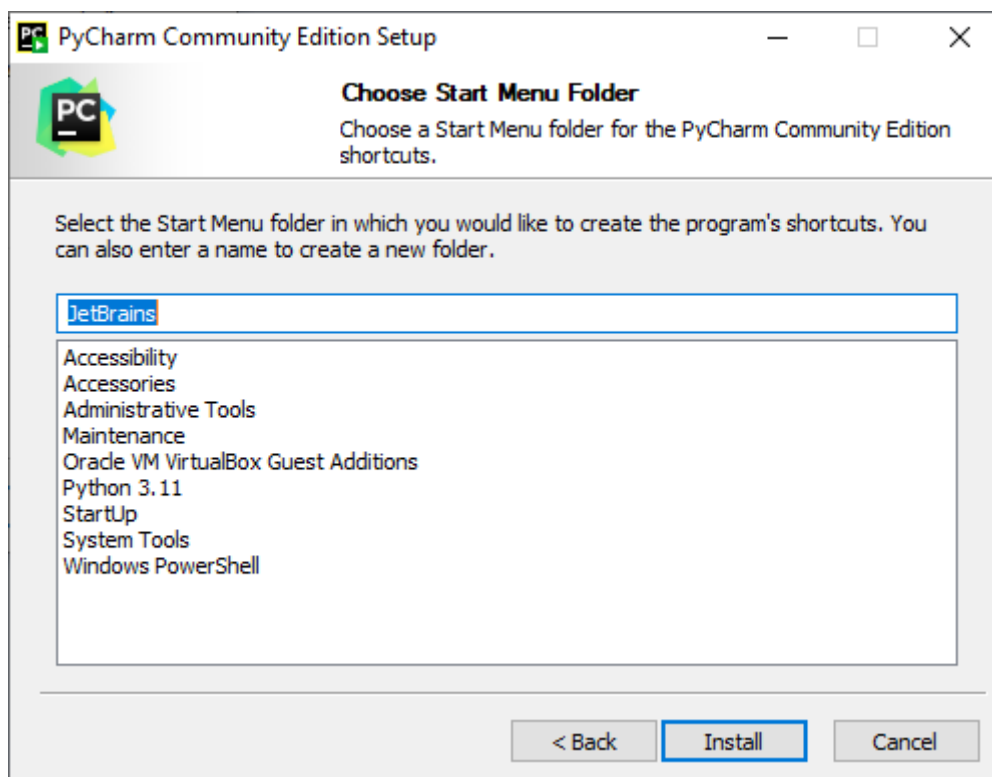


Рис. 27 — Задание имени программы в меню Пуск

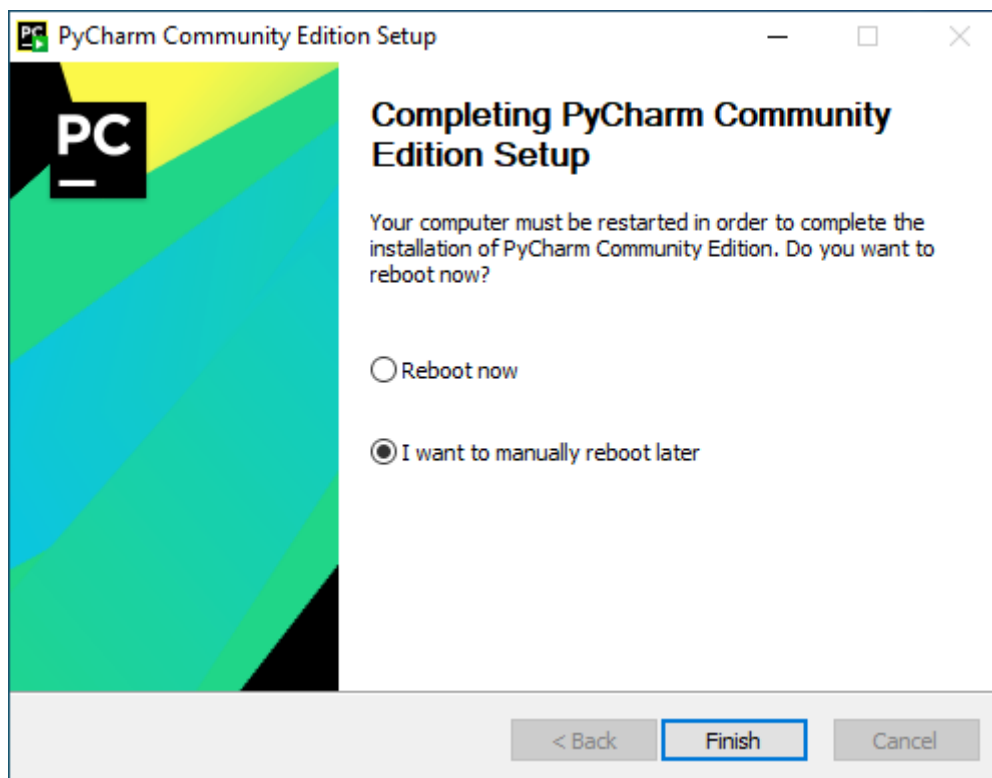


Рис. 28 — Окончание установки PyCharm

После завершения установки рекомендуется выполнить перезагрузку компьютера.

Работа с блокнотом Jupyter Notebook

Работа со скриптами на языке Python возможна в нескольких режимах, среди которых создание текста программ в обычном текстовом редакторе и последующий запуск через командную строку, или использование оболочки IPython (сокращение от «интерактивный Python») в качестве продвинутого интерпретатора Python. Если язык Python представляет собой механизм решения практически любой задачи в области науки о данных, то оболочку IPython можно рассматривать как интерактивную панель управления. Использование оболочки IPython тесно связано с проектом Jupyter, представляющим своеобразный блокнот (текстовый редактор) для браузера, удобный для разработки, совместной работы и использованию ресурсов, а также для публикации научных результатов.

Блокнот Jupyter представляет собой веб-приложение для командной оболочки IPython, позволяющее создавать скрипты на языке Python и предоставляющее большой набор возможностей динамической визуализации. Помимо выполнения операторов, блокнот Jupyter позволяет пользователю вставлять форматированный текст в блокнот, добавлять статические и динамические визуализации, математические формулы, виджеты JavaScript и

многое другое. Блокноты Jupyter могут быть сохранены таким образом, чтобы другие пользователи могли открывать их в своих системах.

Блокнот Jupyter должен иметь возможность взаимодействовать с основным ядром системы. Для этого предназначены специальные протоколы интерактивных вычислений. В ядре Jupyter для Python в качестве такого ядра используется IPython.

Запуск Jupyter Notebook проще всего выполнить из панели Anaconda. Для этого нужно запустить Anaconda Navigator, после чего нажать кнопку Launch, как это показано на рис. 6.

На большинстве платформ Jupyter Notebook открывается в окне браузера, установленного в системе по умолчанию. На рис. 29 показано как выглядит окно начального запуска, в котором можно выбрать блокнот для запуска или создать новый блокнот.

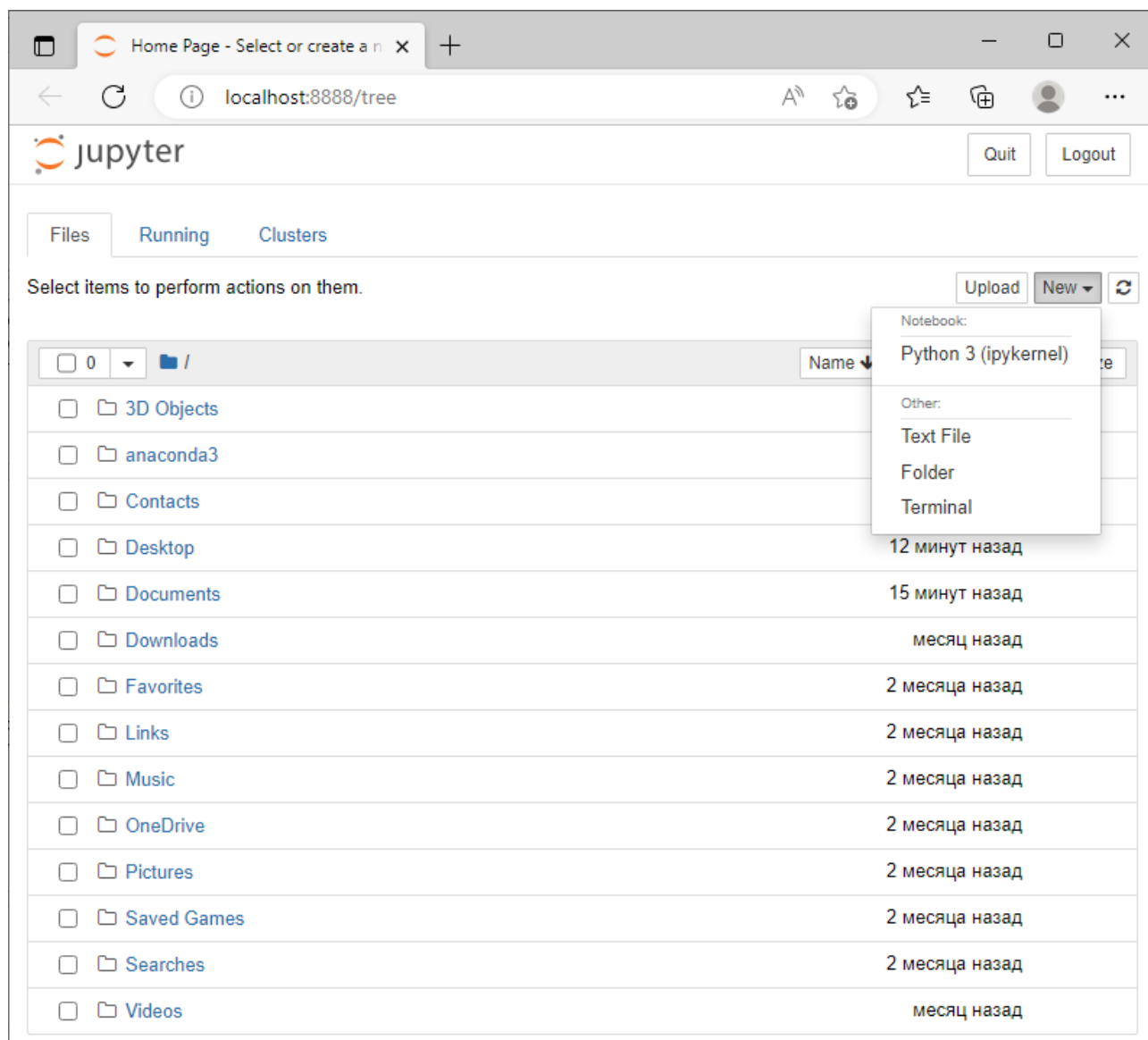


Рис. 29 — Окно начального запуска Jupyter Notebook

Для создания нового блокнота необходимо в комбобоксе **New** выбрать опцию **Python 3 (ipykernel)** в результате чего откроется новое окно браузера с новым блокнотом, как это показано на рис. 30.

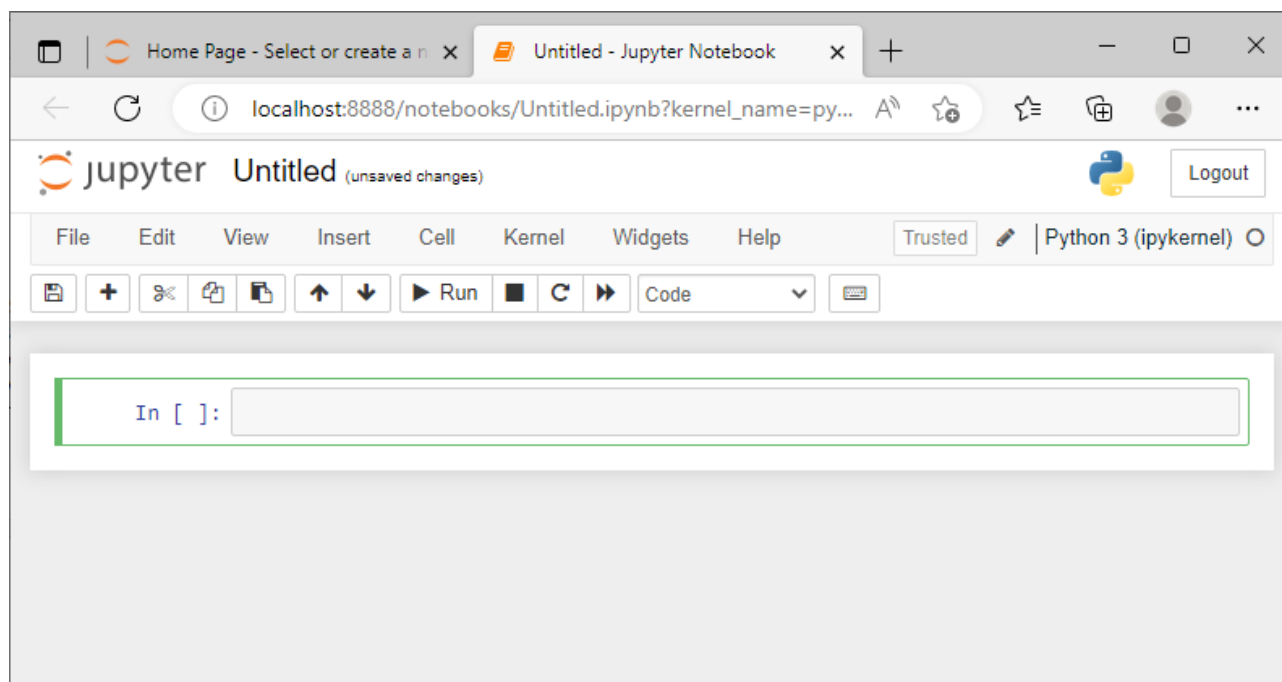


Рис. 30 — Окно с новым блокнотом Jupyter Notebook

Вновь созданный блокнот рекомендуется переименовать. Для этого нужно отредактировать имя по умолчанию **Untitled** в верхней части окна, слева от логотипа Jupyter. Под логотипом находятся пункты меню блокнота и панель инструментов, предназначенные для выполнения манипуляций с ячейками блокнота. Ниже панели инструментов располагается рабочая область с ячейками, предназначенные для ввода команд (ячейки In[]) и вывода результатов (ячейки Out[]). Блокноты Jupyter сохраняются с расширением *.ipynb* и располагаются в каталоге *C:\Users\[Имя_пользователя]*.

Для начала ввода команды, необходимо кликнуть по ячейке, после чего ввести команду или последовательность команд. Для запуска содержимого ячейки на выполнение необходимо нажать комбинацию клавиш **Shift + Enter**.

В результате, если код ячейки не содержит ошибок, то он будет выполнен и автоматически создастся новая ячейка для ввода следующей последовательности команд, так как это показано на рис. 31. При необходимости можно вернуться к предыдущей ячейке и запустить код, находящийся в ячейке, на выполнение еще раз. Если необходимо выполнить запуск на выполнение всех ячеек, входящих в блокнот, то для этого можно воспользоваться пунктом меню **Kernel** → **Restart & Run All**. В этом случае осуществляется перезапуск ядра и ячейки блокнота выполняются по порядку, начиная с первой.

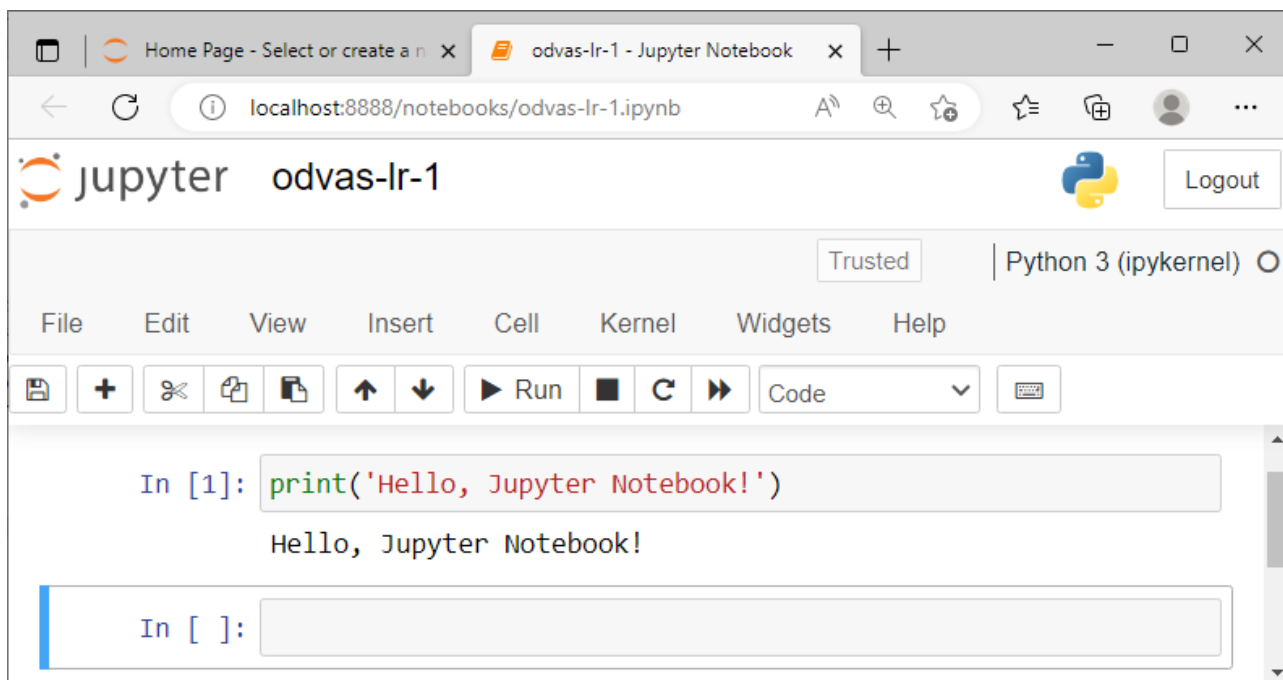


Рис. 31 — Ввод команды в ячейку Jupyter Notebook

Как видно на рис. 31, вновь созданная ячейка обведена серым прямоугольником с левой границей окрашенной в синий цвет. Это означает, что данная ячейка выделена, и с ней можно выполнять определенные манипуляции. Например, данную ячейку можно удалить вместе с ее содержимым, используя сочетание горячих клавиш **D + D**. Если граница прямоугольника окрашена в зеленый цвет, это означает, что ячейка находится в режиме редактирования и готова для ввода новой последовательности команд (см. рис. 30). Для выделения ячейки, необходимо выполнить щелчок левой кнопкой мыши слева от надписи **In [N]**, где N – номер интересующей ячейки.

Для работы с блокнотами Jupyter Notebook рекомендуется выучить набор сочетаний горячих клавиш, что в значительной степени ускорит работу с блокнотом. Горячие клавиши могут быть из стандартного набора, а также могут быть назначены пользователем самостоятельно и по своему усмотрению. Для вывода текущих настроек сочетаний горячих клавиш необходимо перейти в пункт меню **Help** → **Keyboard Shortcuts**.

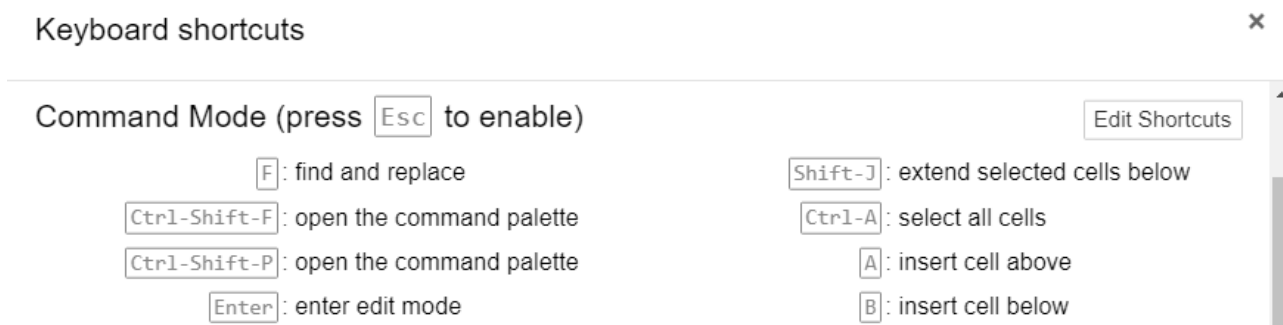


Рис. 32 — Окно с информацией о сочетаниях горячих клавиш

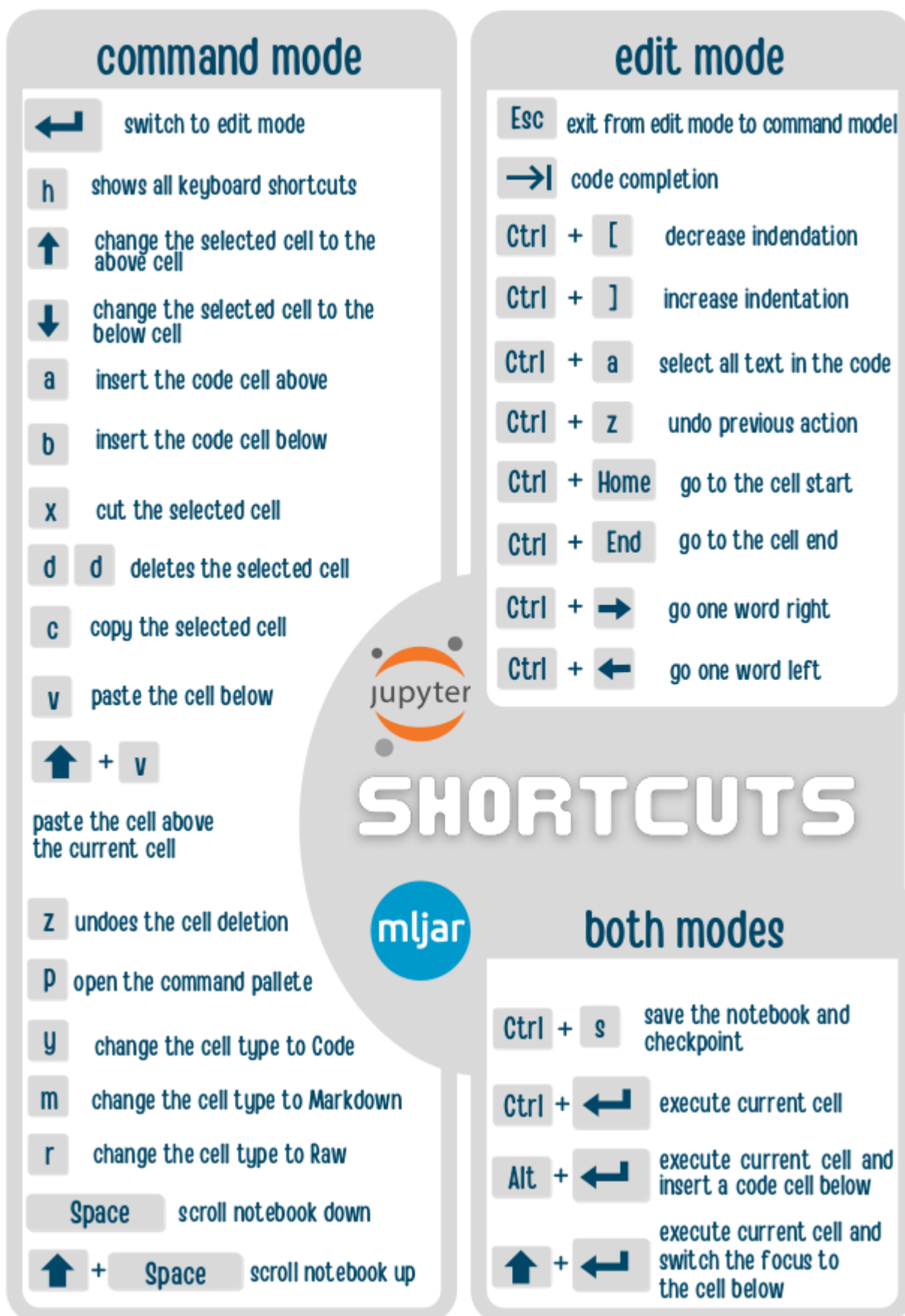


Рис. 33 — Основные сочетания горячих клавиш Jupyter Notebook

Создание проекта в PyCharm

Для создания нового проекта с помощью IDE PyCharm необходимо запустить среду разработки, в появившемся окне нажать на кнопку **New Project** как это показано на рис. 34.

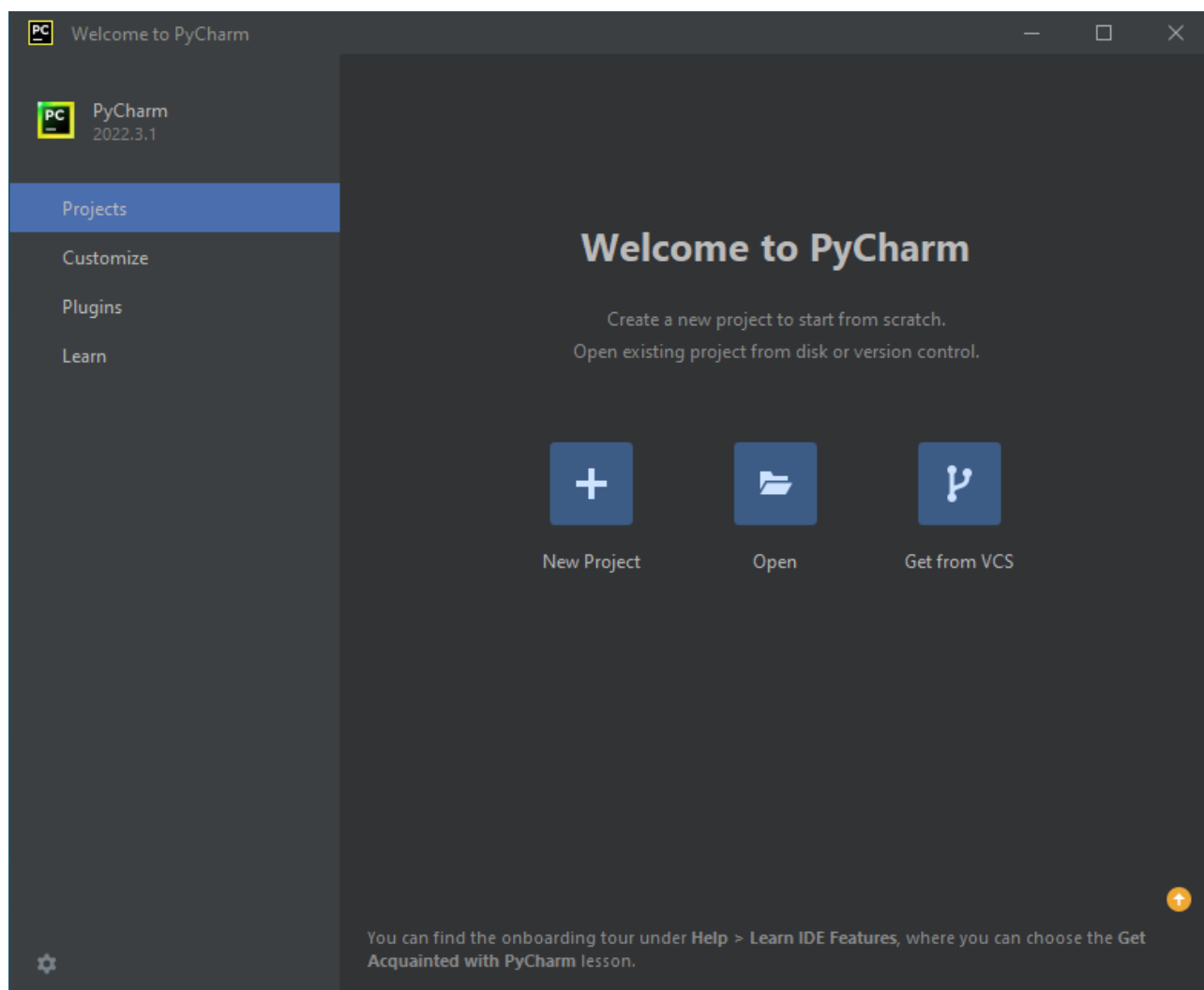


Рис. 34 — Окно приветствия PyCharm

На следующем шаге необходимо выполнить конфигурацию будущего проекта. Окно с настройками проекта представлено на рис. 35. Прежде всего в поле **Location** задаётся путь к папке проекта. В разделе **Python Interpreter** необходимо определить тип окружения, которое будет использоваться для проекта. Это может быть общесистемное окружение (Pipenv) или виртуальное окружение (Virtualenv). Выбор типа окружения определяется тем, из какого хранилища будут подключаться библиотеки. В случае общесистемного окружения все проекты будут использовать одни и те же версии библиотек. Если при создании проекта выбрать виртуальное окружение, тогда проект может использовать свои собственные версии библиотек, независимо от того,

установлена библиотека в общесистемном окружении или нет. Интерпретатор Python при импорте модуля или пакета ищет его сначала в пользовательском окружении, затем — в общесистемном. Перед закрытием окна настроек необходимо убедиться в том, что в поле **Base interpreter** прописан действующий путь к интерпретатору Python.

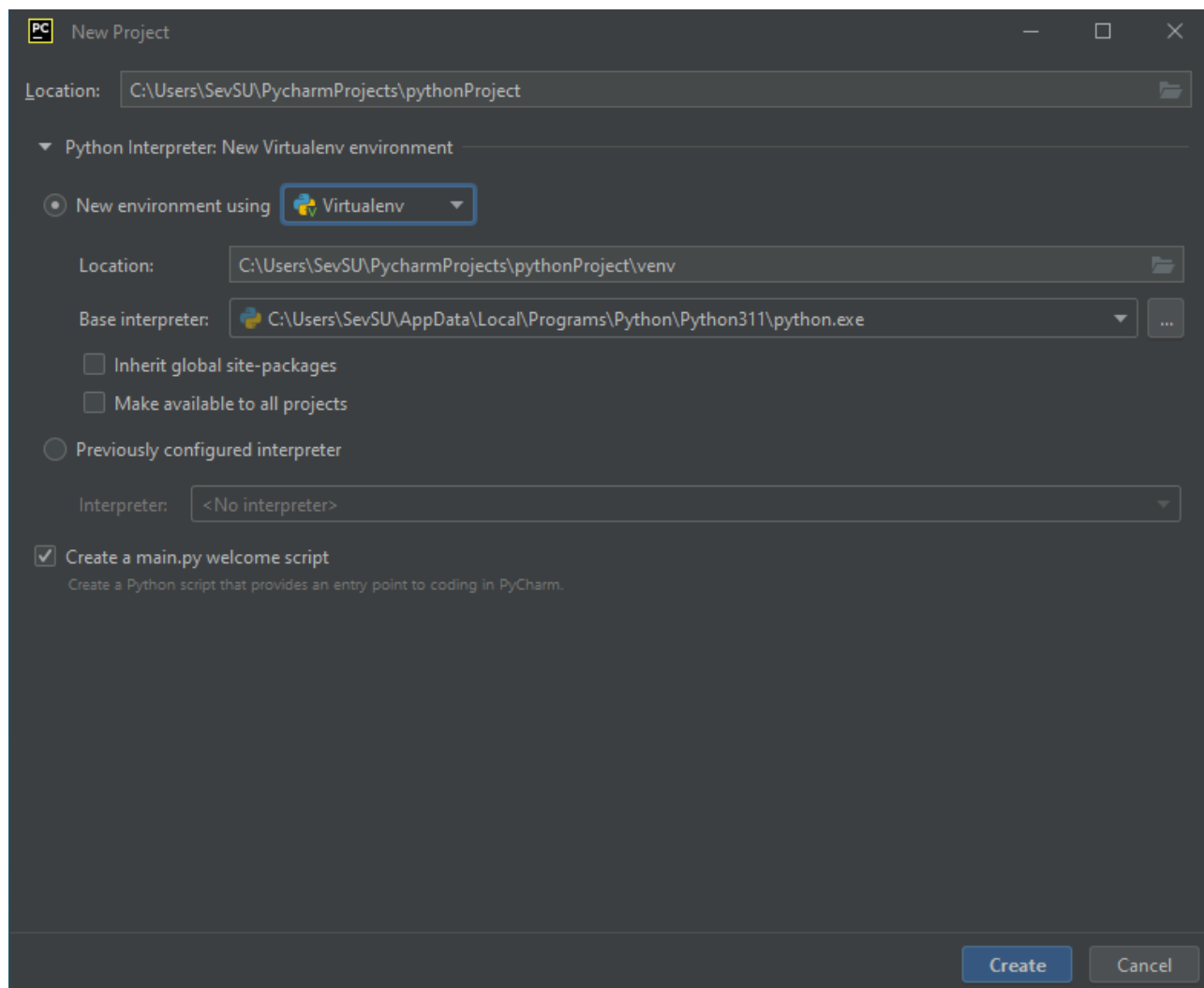


Рис. 35 — Окно настроек проекта PyCharm

После нажатия на кнопку **Create** окна настроек проекта, открывается основное окно разработки, которое представлено на рис. 36.

Окно условно разделено на три области: в левой части окна представлено дерево, отражающее структуру файлов и папок проекта, в правой части расположен редактор кода, внизу находятся вспомогательные закладки для отображения окна терминала, консольного вывода и других системных утилит. Для запуска скрипта на выполнения необходимо нажать на кнопку с зелёной стрелкой в правой части панели инструментов.

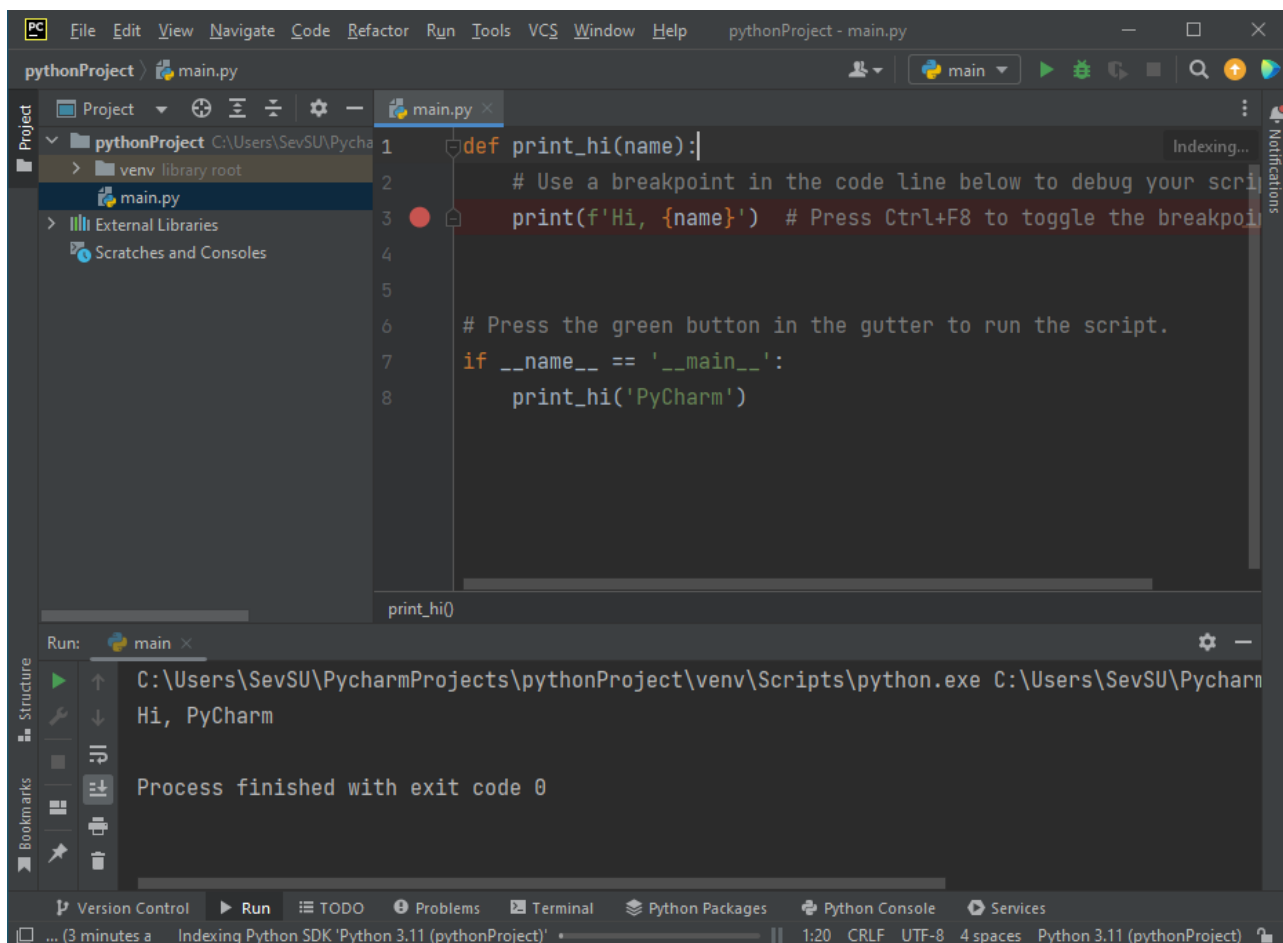


Рис. 36 — Основное окно разработки PyCharm

Поскольку при создании проекта использовано виртуальное окружение, все библиотеки должны быть установлены заново. Для добавления библиотеки необходимо открыть окно настроек (пункт меню **File** → **Settings** или сочетание горячих клавиш Ctrl+Alt+S) и в дереве настроек выбрать пункт **Project** → **Python Interpreter** как это показано на рис. 37. В появившемся списке будут перечислены установленные в виртуальное окружение библиотеки и их версии. При необходимости добавить библиотеку, нужно нажать на кнопку + и в появившемся окне выбрать необходимую библиотеку. После чего нажать кнопку **Install Package** и дождаться окончания установки. На рис. 38 представлен пример установки библиотеки NumPy, а на рис. 39 пример подключения библиотеки к коду с помощью команды `import` и вывод версии библиотеки в консоль с помощью функции `print`. При необходимости можно указать версию библиотеки, которую необходимо установить, в противном случае будет установлена последняя актуальная версия библиотеки.

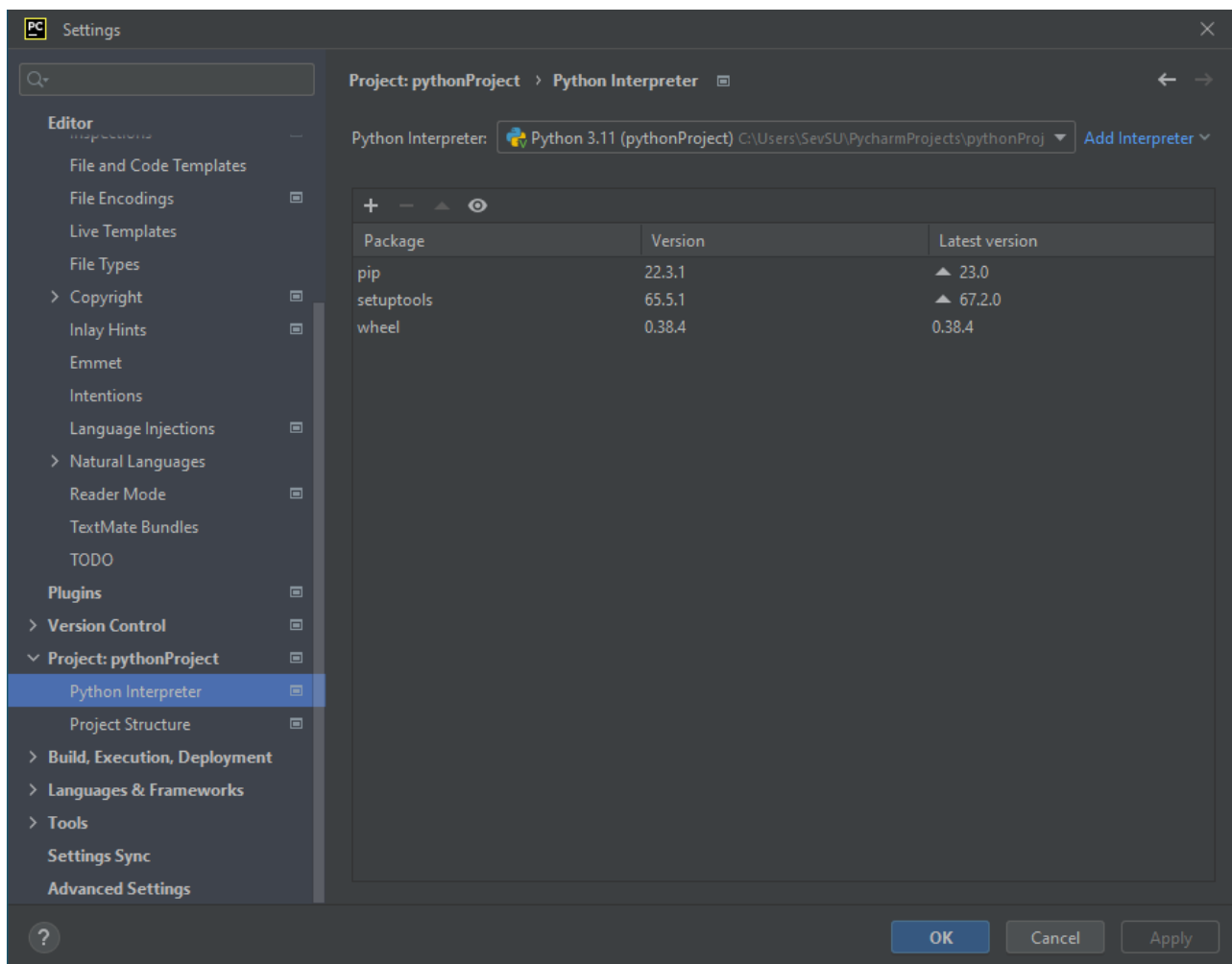


Рис. 37 — Окно настроек проекта

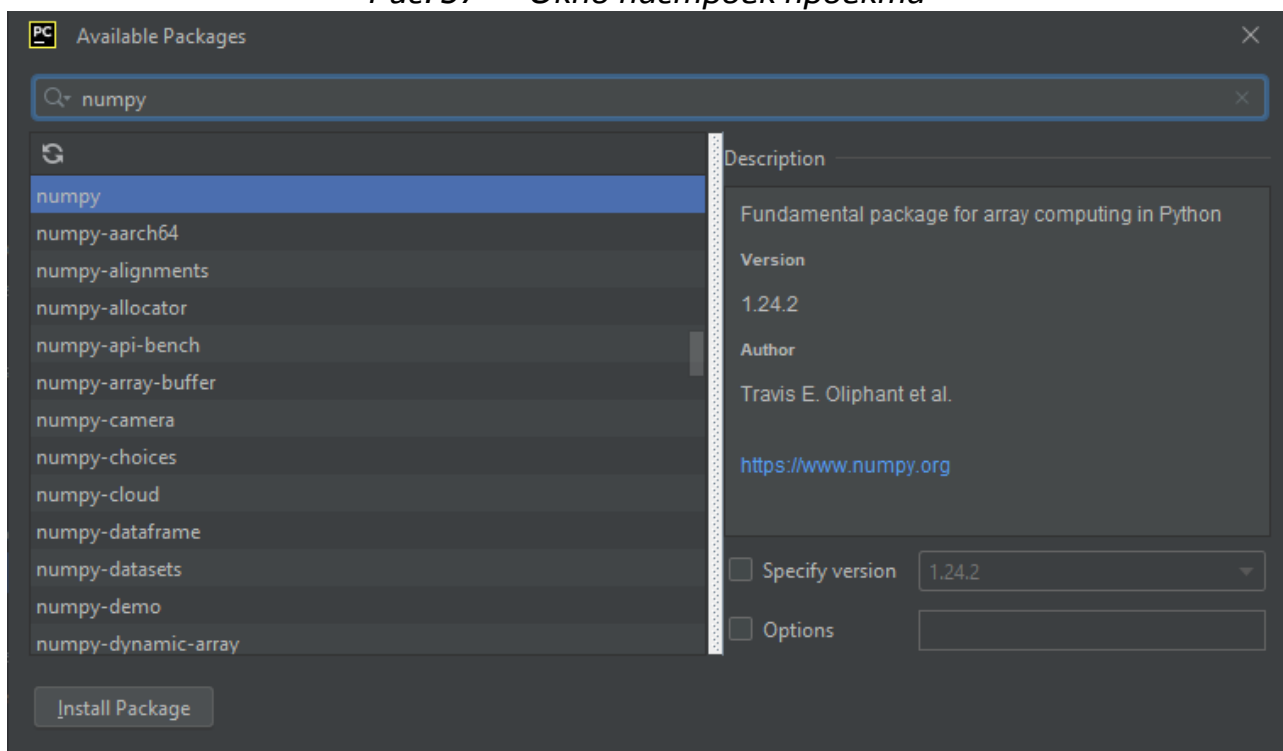


Рис. 38 — Окно установки библиотека

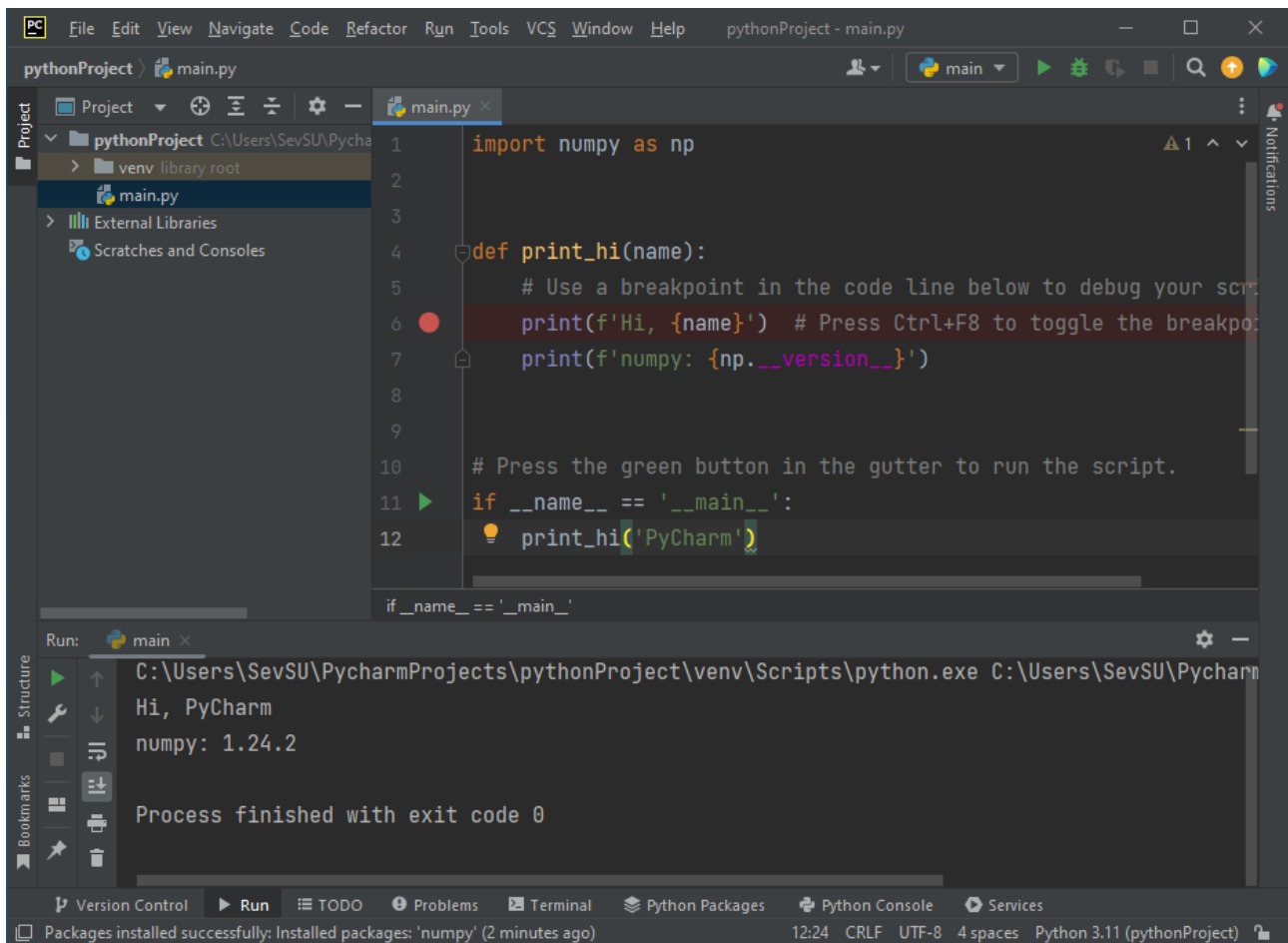


Рис. 39 — Подключение библиотеки и вывод версии библиотеки в консоль

Установка пакетов с помощью менеджера пакетов pip

Чистый Python, как язык программирования, достаточно плохо адаптирован для решения задач анализа данных. Чтобы превратить его в мощный инструмент анализа, необходимо использовать специализированные библиотеки, такие как NumPy, Pandas, Matplotlib и др. В случае, если разработка ведётся в пакете Anaconda, то пользователю не придётся выполнять никаких дополнительных действий по конфигурации окружения, но если используется связка Python + PyCharm, об установке пакетов придётся позаботиться самостоятельно.

Установка пакетов Python осуществляется с помощью менеджера пакетов **pip**, входящего в стандартный пакет установки Python. Работа с **pip** осуществляется в режиме командной строки. Основной формат команды для манипулирования с пакетом имеет вид

pip [команда] [имя_библиотеки]

Основные команды менеджера пакетов **pip** приведены в табл. 1. Полное описание формата команд для различных операционных систем можно найти по ссылке <https://pip.pypa.io/en/stable/getting-started/>

Табл. 1 — Основные команды менеджера пакетов *pip*

№	Команда	Назначение	Пример
1	install	Установка библиотеки	<code>pip install libname</code>
2	install --upgrade	Обновление библиотеки	<code>pip install --upgrade libname</code>
3	uninstall	Удаление библиотеки	<code>pip uninstall libname</code>

Если требуется выполнить установку нескольких пакетов одновременно, то можно воспользоваться следующим приёмом. Создаётся файл `requirements.txt` в котором все необходимые для установки библиотеки перечисляются в каждой строке файла. После создания файла, для установки пакета библиотек необходимо выполнить команду

`pip install -r requirements.txt`

Для того чтобы правильно указать формат команды или имя библиотеки, рекомендуется обратиться к ресурсу <https://pypi.org/>

Пример. Выполним установки библиотеки `xgboost`. Перед установкой библиотеки убедимся в том, что менеджер пакетов *pip* установлен в системе, так как это показано на рис. 40.

```

Командная строка
Microsoft Windows [Version 10.0.19045.2486]
(c) Корпорация Майкрософт (Microsoft Corporation). Все права защищены.

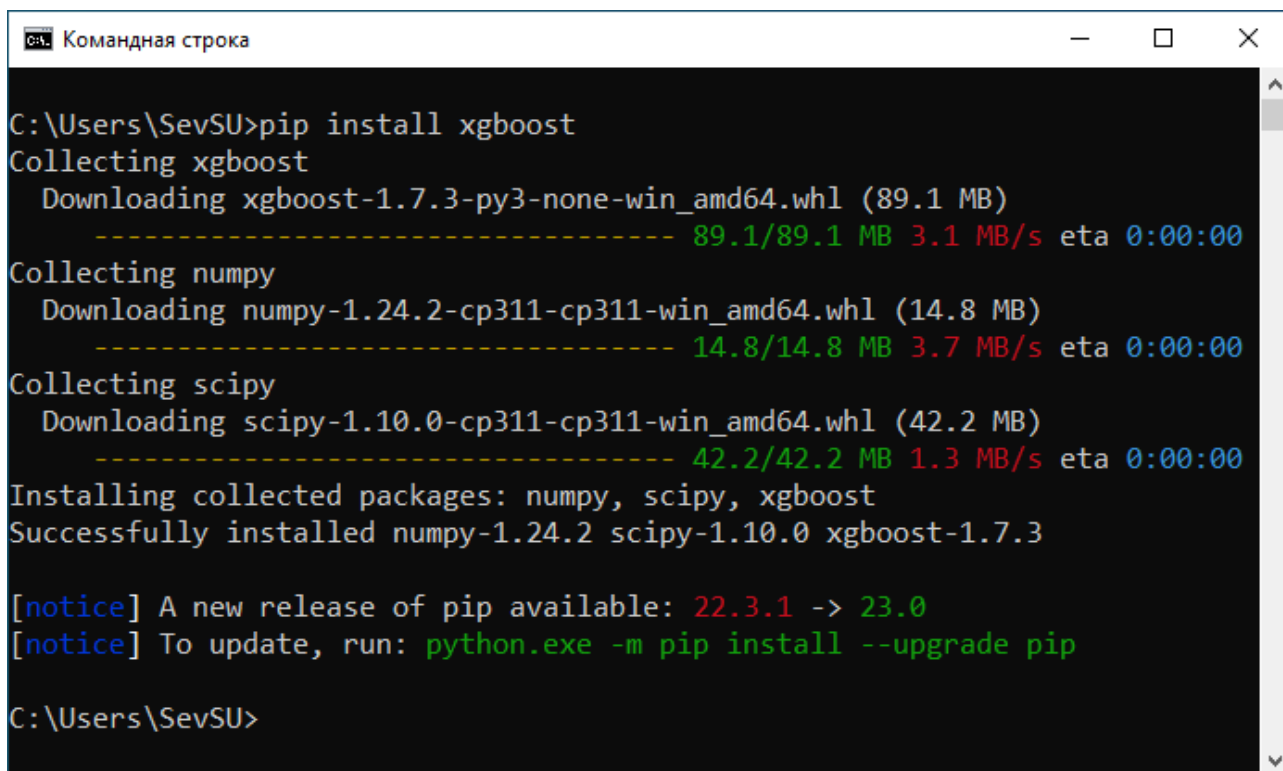
C:\Users\SevSU>pip --version
pip 22.3.1 from C:\Users\SevSU\AppData\Local\Programs\Python\Python311\Lib\
site-packages\pip (python 3.11)

C:\Users\SevSU>

```

Рис. 40 — Проверка установки менеджера пакетов *pip*

Как видно из результатов работы команды, в системе установлен менеджер пакетов версии 22.3.1, следовательно, можно выполнить установку библиотеки. Результат выполнения команды установки приведён на рис. 41.



```
C:\Users\SevSU>pip install xgboost
Collecting xgboost
  Downloading xgboost-1.7.3-py3-none-win_amd64.whl (89.1 MB)
----- 89.1/89.1 MB 3.1 MB/s eta 0:00:00
Collecting numpy
  Downloading numpy-1.24.2-cp311-cp311-win_amd64.whl (14.8 MB)
----- 14.8/14.8 MB 3.7 MB/s eta 0:00:00
Collecting scipy
  Downloading scipy-1.10.0-cp311-cp311-win_amd64.whl (42.2 MB)
----- 42.2/42.2 MB 1.3 MB/s eta 0:00:00
Installing collected packages: numpy, scipy, xgboost
Successfully installed numpy-1.24.2 scipy-1.10.0 xgboost-1.7.3

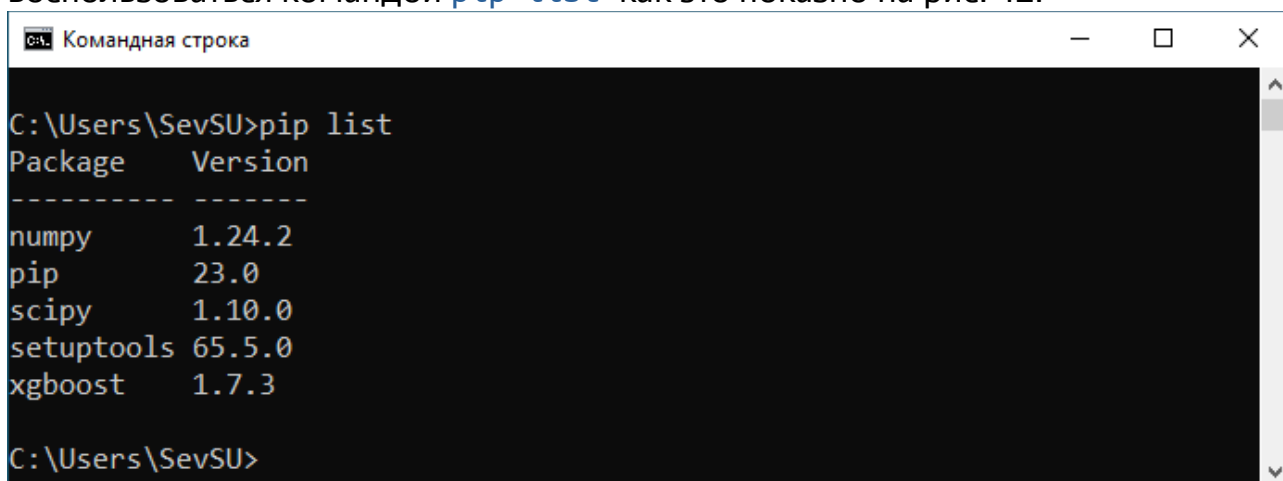
[notice] A new release of pip available: 22.3.1 -> 23.0
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\SevSU>
```

Рис. 41 — Установка библиотеки xgboost

На рис. 41 видно, что при установке библиотеки xgboost были установлены зависимости — библиотеки, которые необходимы для нормальной работы устанавливаемой библиотеки. Также, пользователю предлагается обновиться до актуальной версии менеджера пакетов.

Для просмотра всех библиотек, установленных в системе, можно воспользоваться командой `pip list` как это показано на рис. 42.



```
C:\Users\SevSU>pip list
Package      Version
-----
numpy        1.24.2
pip          23.0
scipy        1.10.0
setuptools   65.5.0
xgboost      1.7.3

C:\Users\SevSU>
```

Рис. 42 — Получение списка установленных библиотек

Помимо перечня библиотек, пользователю доступна информация о версии установленной библиотеки. При необходимости, библиотека может быть обновлена с помощью команды `upgrade`.

Библиотека math

В языке Python существует ряд стандартных библиотечных модулей, которые могут быть использованы для расширения базовых функций языка. К таким библиотечным модулям относится модуль `math` который экспортирует стандартные средства из библиотеки математических функций C для применения в Python. В табл. 2 и табл. 3 перечислены математические функции и константы, экспортируемые этим модулем. Для получения детального описания той или иной функции библиотеки необходимо импортировать модуль `math` и вызвать функцию `help(math.имя_функции)`. В результате будут выведено описание функции, сведения об аргументах и примечания к её применению. Подробнее о модуле `math` см. в руководстве по библиотеке Python <https://docs.python.org/3/library/math.html>

Табл. 2 — Математические функции модуля `math`

Функция	Описание
<code>math.acos()</code>	Returns the arc cosine of a number
<code>math.acosh()</code>	Returns the inverse hyperbolic cosine of a number
<code>math.asin()</code>	Returns the arc sine of a number
<code>math.asinh()</code>	Returns the inverse hyperbolic sine of a number
<code>math.atan()</code>	Returns the arc tangent of a number in radians
<code>math.atan2()</code>	Returns the arc tangent of y/x in radians
<code>math.atanh()</code>	Returns the inverse hyperbolic tangent of a number
<code>math.ceil()</code>	Rounds a number up to the nearest integer
<code>math.comb()</code>	Returns the number of ways to choose k items from n items without repetition and order
<code>math.copysign()</code>	Returns a float consisting of the value of the first parameter and the sign of the second parameter
<code>math.cos()</code>	Returns the cosine of a number
<code>math.cosh()</code>	Returns the hyperbolic cosine of a number
<code>math.degrees()</code>	Converts an angle from radians to degrees
<code>math.dist()</code>	Returns the Euclidean distance between two points (p and q), where p and q are the coordinates of that point
<code>math.erf()</code>	Returns the error function of a number
<code>math.erfc()</code>	Returns the complementary error function of a number
<code>math.exp()</code>	Returns E raised to the power of x
<code>math.expm1()</code>	Returns $e^x - 1$
<code>math.fabs()</code>	Returns the absolute value of a number
<code>math.factorial()</code>	Returns the factorial of a number

Табл. 2 (продолжение) — Математические функции модуля *math*

Функция	Описание
<code>math.floor()</code>	Rounds a number down to the nearest integer
<code>math.fmod()</code>	Returns the remainder of x/y
<code>math.frexp()</code>	Returns the mantissa and the exponent, of a specified number
<code>math.fsum()</code>	Returns the sum of all items in any iterable (tuples, arrays, lists, etc.)
<code>math.gamma()</code>	Returns the gamma function at x
<code>math.gcd()</code>	Returns the greatest common divisor of two integers
<code>math.hypot()</code>	Returns the Euclidean norm
<code>math.isclose()</code>	Checks whether two values are close to each other, or not
<code>math.isfinite()</code>	Checks whether a number is finite or not
<code>math.isinf()</code>	Checks whether a number is infinite or not
<code>math.isnan()</code>	Checks whether a value is NaN (not a number) or not
<code>math.isqrt()</code>	Rounds a square root number downwards to the nearest integer
<code>math.ldexp()</code>	Returns the inverse of <code>math.frexp()</code> which is $x * (2^{**i})$ of the given numbers x and i
<code>math.lgamma()</code>	Returns the log gamma value of x
<code>math.log()</code>	Returns the natural logarithm of a number, or the logarithm of number to base
<code>math.log10()</code>	Returns the base-10 logarithm of x
<code>math.log1p()</code>	Returns the natural logarithm of 1+x
<code>math.log2()</code>	Returns the base-2 logarithm of x
<code>math.perm()</code>	Returns the number of ways to choose k items from n items with order and without repetition
<code>math.pow()</code>	Returns the value of x to the power of y
<code>math.prod()</code>	Returns the product of all the elements in an iterable
<code>math.radians()</code>	onverts a degree value into radians
<code>math.remainder()</code>	Returns the closest value that can make numerator completely divisible by the denominator
<code>math.sin()</code>	Returns the sine of a number
<code>math.sinh()</code>	Returns the hyperbolic sine of a number
<code>math.sqrt()</code>	Returns the square root of a number
<code>math.tan()</code>	Returns the tangent of a number
<code>math.tanh()</code>	Returns the hyperbolic tangent of a number
<code>math.trunc()</code>	Returns the truncated integer parts of a number

Табл. 3 — Математические константы модуля math

Константа	Описание
math.e	Returns Euler's number (2.7182...)
math.inf	Returns a floating-point positive infinity
math.nan	Returns a floating-point NaN (Not a Number) value
math.pi	Returns PI (3.1415...)
math.tau	Returns tau (6.2831...)

Работа с файлами

При работе с фалами чаще всего встречаются две задачи: считывание информации из файла и запись информации в файл. В Python эти задачи решаются достаточно просто за счёт использования специальных файловых объектов. Независимо от типа операции с фалом, вначале необходимо создать файловый объект, после чего файл нужно открыть. Для этого существует специальная функция `open()`. Первым обязательным аргументом функции является полное имя файла, с которым планируется работа, следующим необязательным параметром является ключ, определяющий спецификатор доступа. Возможные значения ключа приведены в табл. 4. Созданный в результате вызова функции `open` файловый объект имеет специальные методы для выполнения основных действий — чтения `read()` и записи `write()`. После того, как работа с файловым объектом завершена, необходимо вызвать метод `close()`.

Табл. 4 — Значения ключей для доступа к файлу

Ключ	Значение
r	Режим чтения содержимого файла. Используется по умолчанию, если режим доступа явно не указан
w	Режим записи в файл. При использовании ключа w содержимое файла уничтожается (если файл уже был создан ранее)
x	Режим создания файла с возможностью записи данных в файл. Если файл уже существует, возникает ошибка
a	Режим записи в файл. Если файл существует, то новое содержимое дописывается в конец файла
b	Режим доступа к бинарному файлу
t	Режим доступа к текстовому файлу. Используется по умолчанию, если явно режим доступа не задан

Варианты задания

В соответствии с вариантом задания вычислить значение математической функции. Начальное и конечное значения аргумента, а также шаг изменения аргумента использовать в соответствии со значениями, представленными в табл. 5. Для каждой функции определить максимальное и минимальное значения. Результат табулирования функции вывести в текстовый файл, при выводе вещественных значений ограничиться двумя знаками после запятой для значения аргумента и тремя знаками для значения функции.

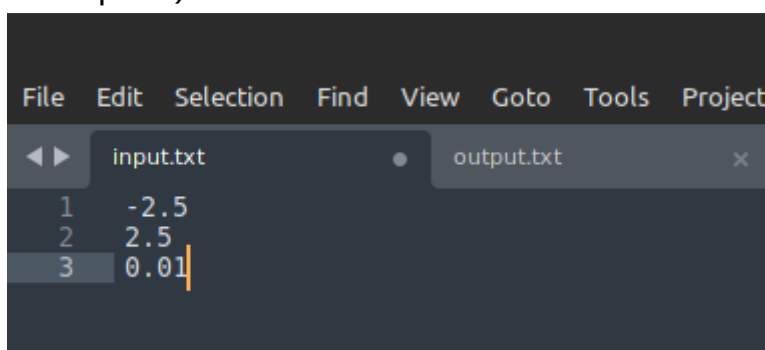
Табл. 5 — Варианты задания

№	Функция	x_0	x_f	h_x
1	$y(x) = 3x^3 - 15x^2 - 12x + 8$	-5	15	0,01
2	$y(x) = x^2 \sin^3(3x)$	-3	2,5	0,001
3	$y(x) = x^3 \sin^5(4x)$	-1,6	3,2	0,001
4	$y(x) = (1 + 6 \sin^2(x))^{1/2} \cos(3x)$	-5	12	0,001
5	$y(x) = 2x^4 - 12x^3 - 4x^2 + 2$	-5	10	0,01
6	$y(x) = \cos(2,5x) \sin(3x) - 0,2$	-7	7	0,01
7	$y(x) = 4^{-(x-1)^2} \cos(2x)$	-10	20	0,1
8	$y(x) = \cos(2,5x) \sin(3x) - 0,2$	-3,2	8	0,001
9	$y(x) = x(\sin(3x) + \cos(2x))$	-3,5	4,5	0,01
10	$y(x) = x^2(\sin^2(3x) - \cos^2(2x))$	-3,5	5,5	0,1

Пример выполнения работы

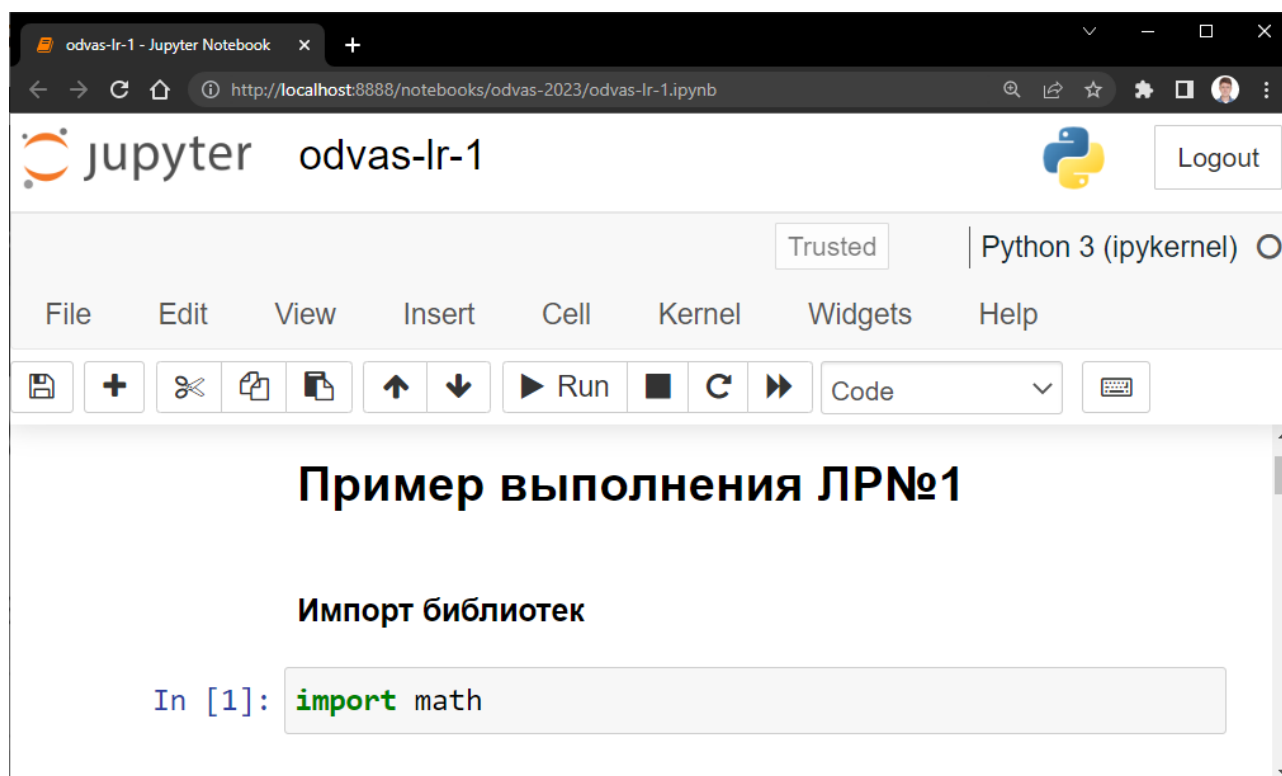
Вычислить значение функции $y(x)=x^3 \sin(|4x|)$ на интервале от $x_0=-2,5$ до $x_f=2,5$ с шагом $h_x=0,01$. Значения x_0, x_f, h_x прочитать из входного файла. Результат табулирования функции вывести в текстовый файл, при выводе вещественных значений ограничиться двумя знаками после запятой для значения аргумента и тремя знаками для значения функции. Математическую функцию определить с помощью блока **def**.

Входной файл данных имеет вид (каждое новое значение параметра начинается с новой строки)



```
File Edit Selection Find View Goto Tools Project
input.txt output.txt x
1 -2.5
2 2.5
3 0.01
```

Реализация задания выполнена в Jupyter Notebook. Ячейка [1] предназначена для импорта библиотеки `math`.



Далее осуществляется чтение исходных данных из файла

Чтение исходных данных

```
In [2]: in_file_name = 'input.txt' # Имя входного файла
in_file = open(in_file_name, 'r') # Открытие файла
X0 = in_file.readline().strip() # Читаем X0
Xf = in_file.readline().strip() # Читаем Xf
hX = in_file.readline().strip() # Читаем hX
# Выводим считанные значения параметров
print(f'X0 = {X0}; Xf = {Xf}; hX = {hX}')
in_file.close() # Закрываем файл
```

```
X0 = -2.5; Xf = 2.5; hX = 0.01
```

Поскольку данные хранятся в текстовом виде, при чтении из файла необходимо удалить служебные символы, такие как возврат каретки (перевод на новую строку). Для этого можно воспользоваться функцией `strip()`. Для вывода результата в консоль рекомендуется использовать строки с форматированием, которые сочетают в себе статический и динамический текст. Переменные, которые должны быть выведены в строке помещаются в фигурные скобки.

Следующее действие — необходимо определить пользовательскую функцию. Для этого используется служебное слово **def**.

Определяем пользовательскую функцию

```
In [3]: def my_func(X):
        return math.pow(X,3)*math.sin(math.fabs(4*X))
```

Далее происходит преобразование входных параметров из текстового формата в числовой. При этом могут быть использованы те же самые переменные. Также осуществляется расчёт общего количество точек, для которых должно быть вычислено значение функции.

Табулирование функции и запись в файл

```
In [4]: # Преобразование в числовой формат
Xf = float(Xf)
X0 = float(X0)
hX = float(hX)
N = math.ceil((Xf - X0)/hX) + 1
print(f'Кол-во точек: {N}')
```

```
Кол-во точек: 501
```

В ячейке [5] происходит основная процедура табулирования функции. При этом используется функция `range()` для генерации набора целых чисел в диапазоне от 0 до 500. В результате чего по заданному значению изменения шага аргумента и текущему номеру отсчёта i осуществляется вычисление текущего значения аргумента. Инициализация переменных для поиска минимального и максимального значения функции производится с помощью библиотеки `math`. Для вывода в файл используется форматирование строки и ограничение количества знаков после запятой. Это обеспечивается за счёт применения конструкции `{X:.2f}` которая означает вывод вещественного числа с двумя знаками после запятой. Для вывода значения функции используется конструкция `{Y:.4f}` что соответствует четырём знакам после запятой. Поиск максимума и минимума осуществляется с помощью оператора `if`.

```
In [5]: out_file_name = 'output.txt' # Имя входного файла
out_file = open(out_file_name, 'w') # Открытие файла
# Инициализируем переменные для максимума и минимума
Ymin = math.inf
Ymax = -math.inf
for i in range(N):
    X = X0 + i*hX # Вычисление аргумента
    Y = my_func(X) # Вычисление функции
    # Вывод в файл
    out_file.write(f'{X:.2f}\t{Y:.4f}\n')
    # Поиск минимума и максимума
    if (Y < Ymin):
        Ymin = Y

    if (Y > Ymax):
        Ymax = Y

out_file.close()
```

```
In [6]: print(f'Минимальное значение: {Ymin:.4f}')
print(f'Максимальное значение: {Ymax:.4f}')
```

Минимальное значение: -8.5003

Максимальное значение: 8.5003

В результате работы программы, формируется выходной файл `output.txt` который имеет вид

1	-2.50	8.5003
2	-2.49	7.8740
3	-2.48	7.2486
4	-2.47	6.6254
5	-2.46	6.0053

Контрольные вопросы

1. Назовите современные программные инструментальные средства для анализа данных. Дайте их краткую характеристику.
2. Что из себя представляет блокнот Jupyter Notebook?
3. Сравните возможности Jupyter Notebook и PyCharm. Какой из этих инструментов вы выбираете и почему?
4. Перечислите основные команды / горячие клавиши Jupyter Notebook.
5. Что такое пакеты?
6. Как узнать, какие пакеты установлены?
7. Как установить необходимые пакеты?
8. Что такое зависимость пакетов?
9. Что такое общесистемное и виртуальное окружение проекта Python?
10. Как использовать Jupyter Notebook для написания скрипта на языке Python?
11. Как создаётся проект Python в IDE PyCharm?
12. Как осуществляется вывод в файл с помощью Python?
13. Назовите основные правила по написанию кода на Python, перечислите основные синтаксические конструкции.

Список использованных источников

1. **Маккини У. Python и анализ данных / У. Маккини** — Пер. с англ. — М.: ДМК Прес, 2020. — 540 с.
2. **Плас Дж. В. Python для сложных задач: наука о данных и машинное обучение / Дж. В. Плас** — Пер. с англ. — СПб.: Питер, 2022. — 576 с.
3. **Грас Дж. Data Science. Наука о данных с нуля / Дж. Грас** — Пер. с англ. — 2-е издание — СПб.: БХВ-Петербург, 2022. — 416 с.
4. **Луц М. Изучаем Python. Том I / М. Луц** — Пер. с англ. — 5-е издание — СПб.: ООО «Диалектика», 2020. — 720 с.
5. **Луц М. Изучаем Python. Том II / М. Луц** — Пер. с англ. — 5-е издание — СПб.: ООО «Диалектика», 2020. — 832 с.
6. **Рашка С. Python и машинное обучение: машинное обучение с использованием Python, scikit-learn и TensorFlow / С. Рашка, В. Мирджалили** — Пер. с англ. — 3-е издание — СПб.: ООО «Диалектика», 2020. — 848 с.
7. **Брюс П. Практическая статистика для специалистов Data Science / П. Брюс, Э. Брюс, П. Гадек** — Пер. с англ. — 2-е издание — СПб.: БХВ-Петербург, 2022. — 352 с.
8. **Васильев А. Программирование на Python в примерах и задачах / А. Васильев** — М.: Эксмо, 2021. — 616 с.
9. **Луц М. Python. Карманный справочник / М. Луц** — Пер. с англ. — 5-е издание — СПб.: ООО «Диалектика», 2020. — 320 с.

Полезные ссылки

1. <https://linuxmint.com/> — Страница загрузки дистрибутива Linux Mint.
2. <https://www.python.org/> — Страница загрузки установщика интерпретатора Python.
3. <https://www.jetbrains.com/pycharm/> — Страница загрузки установщика IDE PyCharm.
4. <https://www.anaconda.com/> — Страница загрузки установщика среды Anaconda.
5. <https://pypi.org/> — Каталог библиотек Python (Python Package Index).
6. <https://docs.python.org/3/library/functions.html> — Документация по функциям Python.
7. <https://docs.python.org/3/library/math.html> — Документация по библиотеке math.

Учебное издание

Василий Викторович Альчаков

ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА ДЛЯ РАБОТЫ С ДАННЫМИ
Методические указания к выполнению лабораторной работы

Оригинал-макет и вёрстка В.В. Альчаков

© СевГУ, 2023

© Альчаков В.В., 2023