

Feb 22



IT & Software

Requirements and Installation Guide

IMPORTANT - ORGANISATION SYSTEM RESTRICTIONS

Depending on your organisation's systems there may be restrictions in place for certain software or tools. We have where possible provided a selection of options and solutions for each tool required to ensure you have full access to the range of software required.

You may need to check with your employer if there are any specific steps you need to take to access these tools or if there are suitable alternatives they can provide.

Contents

Now please navigate to the relevant section depending on your apprenticeship programme:

1. [All programmes - Tools needed for all programmes](#)
2. [L3 - Data Citizen or Data-Driven Journalism](#)
3. [L4 - Data Analyst](#)
4. [L4 - Digital Business Analyst or Digital Business Lead](#)
5. [L7 - AI Apprenticeships \(MLOps, DataOps and Advanced Data Science\)](#)



All programmes



Required Tools used for all programmes

This is the complete checklist of tools you will need to use. Instructions for each tool are provided in the following pages.

- 1. Microsoft Teams**
- 2. EDUKATE.AI**
- 3. Aptem**
- 4. Archive Files**

Microsoft Teams (MS Teams)



MS Teams



Microsoft Teams is a unified communication and collaboration platform that combines persistent workplace chat, video meetings, file storage (including collaboration on files), and application integration.

The screenshot shows a Jupyter Notebook titled "Lecture Notes Dealing with Time" running in a Firefox browser. The notebook is divided into three sections: "Working with Pandas Date and Time", "Working with Timedeltas", and "Working with Pandas Timedelta".

```
import pandas as pd

pd.to_datetime("14th of October, 2018")
pd.Timestamp(year=2018, month=10, day=14, hour=12, minute=0, second=30)
dates = pd.date_range("2014-07-04", "2014-08-04", "2014-08-04", "2015-07-04",
                    "2015-08-04")
series = pd.Series([10, 4, 14, 30], index=dates)
pd.date_range("2015-07-04", "2015-07-10")
pd.date_range("2018 Oct 1", periods=10, freq="W")
```

Working with Timedeltas

- Timedeltas represent time between two Timestamps
- They are returned when doing arithmetic on Timestamps
- Or you can create them from scratch

```
import pandas as pd

diff = pd.Timestamp("2019-07-07") - pd.Timestamp("2019-07-05")

diff.components
# Components (days=2, hours=0, minutes=0, seconds=0, milliseconds=0, microseconds=0)
diff.days
# 2
```

Working with Pandas Timedelta

```
import pandas as pd

diff = pd.Timestamp("2019-07-07") - pd.Timestamp("2019-07-05")

diff.components
# Components (days=2, hours=0, minutes=0, seconds=0, milliseconds=0, microseconds=0)
diff.days
# 2
```

The Jupyter Notebook interface shows the following code cells and their outputs:

```
In [18]: 1 date_as_string = "14th of October, 2018"

In [19]: 1 pd.to_datetime(date_as_string)
Out[19]: Timestamp('2018-10-14 00:00:00')
```

```
In [20]: 1 pd.Timestamp(year=2018, month=10, day=14)
Out[20]: Timestamp('2018-10-14 00:00:00')
```

```
In [21]: 1 dates = ['2014-07-04', '2014-08-04', '2015-07-04', '2015-08-04']

In [22]: 1 pd.to_datetime(dates)
Out[22]: DatetimeIndex(['2014-07-04', '2014-08-04', '2015-07-04', '2015-08-04'],
                        dtype='datetime64[ns]', freq=None)

In [ ]: 1
```



MS Teams - Setup

Before a videoconference

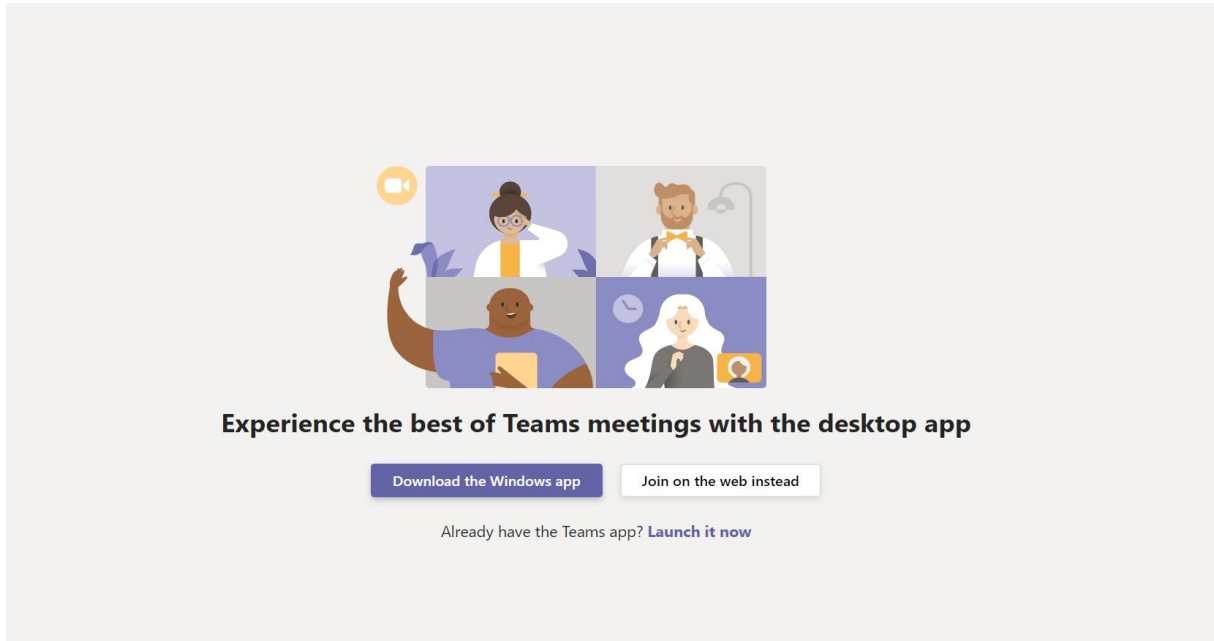
- You will need a computer, tablet, or smartphone with speaker or headphones. You will have the opportunity to check your audio immediately upon joining a meeting.
- You will receive an official calendar invite for the live event from Cambridge Spark a few days prior to your first live session .

To join the videoconference

- At the start time of your meeting, click on the link '*Join Microsoft Teams Meeting*' in the invitation. The following screen will appear:

MS Teams - Setup

To access a live meeting, you can either use the **web browser** (if you don't have a personal or work Microsoft 365 account) or **download the Windows app** (if you have a personal or work Microsoft 365 account).



Experience the best of Teams meetings with the desktop app

[Download the Windows app](#) [Join on the web instead](#)

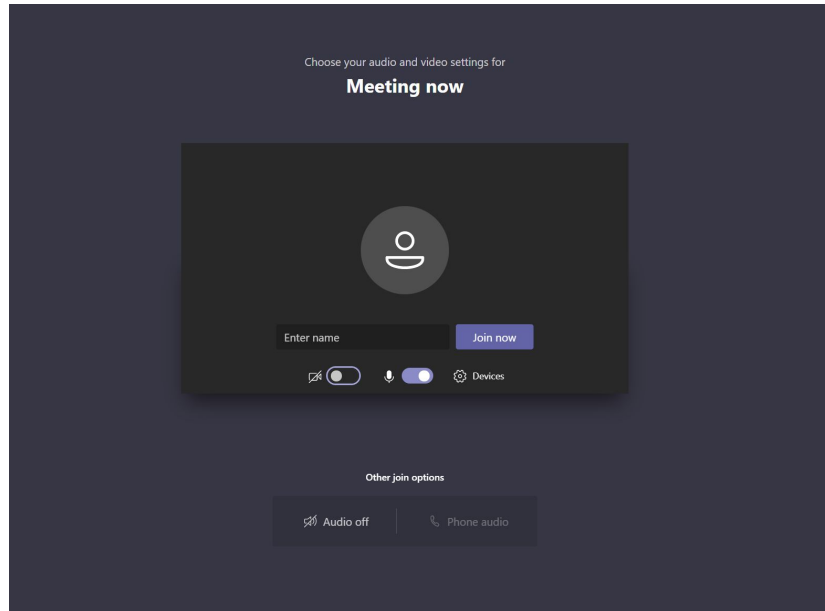
Already have the Teams app? [Launch it now](#)



MS Teams - Setup

Web Browser

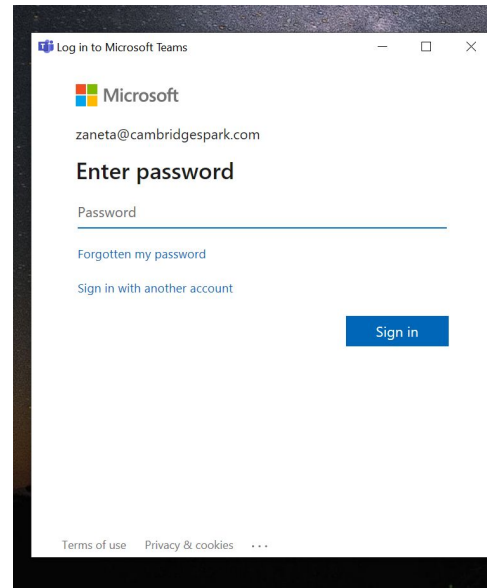
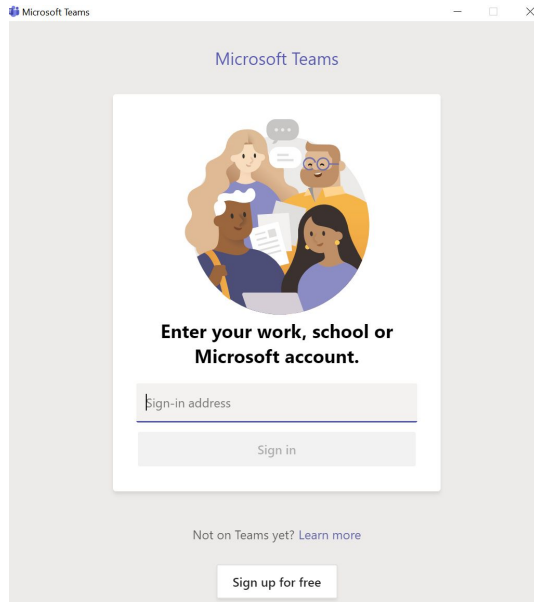
- Click the 'web browser' option, it will take you directly to the meeting, you will be asked to enter your name, after that you can join the session.



MS Teams - Setup

Download the Windows App

- Click the 'download the Windows app' option, once the software has been downloaded, please install it.
- Open the app and sign in with your Microsoft 365 username and password.
- You will then be able to join the meeting.



MS Teams - Setup

Additional functions

Using the icons in the middle of the MS Teams screen, you can:

- Mute/Unmute your microphone
- Turn on/off camera
- View Participant list – opens a pop-out screen where you can see all participants and access the **meeting's chat**
- Share your screen

MS Teams - Instructions for the live remote session

During the session

- When you enter the broadcasting, immediately **mute your microphone** whilst the teacher is speaking. This is very important for sound quality! But we encourage you to put your camera on during sessions, especially at the start and whilst raising questions in session.

How to ask questions during the session

- Use the dedicated live session to ask questions, tutors will be answering them as they pop up. You can use it for 1:1 questions as well (e.g. screen share, call) during the session. If one question is trending the tutors might leave it for the teacher to answer. The tutor's support on teams is available only during the live remote session.
- For any Assignment and follow-up questions - can you please use KATE support functionality or the dedicated knowledge base.

EDUKATE.AI



EDUKATE.AI



EDUKATE.AI is our proprietary tool for hands-on practical coding projects where learners can apply their new skills and receive immediate and personalised feedback.

It is browser based so learners will need to access it via their web browser (Chrome is recommended).

Learners need access to:

<https://app.edukate.ai/>

Please ensure there are no firewall restrictions on that URL.

Your user account will be created by Cambridge Spark separately.

A screenshot of the EDUKATE.AI web application interface. The header shows the logo and navigation links for 'Assignments', 'Projects', 'Help', and a user profile for 'Tony Clark'. The main content area is titled 'All Assignments' and is divided into two sections: 'Tutorial' and 'Preparation'. The 'Preparation' section contains four assignment cards. The first card, 'Maths Fundamentals', shows a progress of 9.2% and a latest submission of Mar 12, 2020 4:41 PM. The second card, 'Programming Fundamentals', shows a progress of 11.8% and a latest submission of Apr 9, 2020 3:18 PM. The third card, 'pre1-maths', shows 0.0% progress and 'No Activity Yet'. The fourth card, 'pre2-programming', also shows 0.0% progress and 'No Activity Yet'. Each card has a 'View Assignment' button at the bottom.

All Programmes



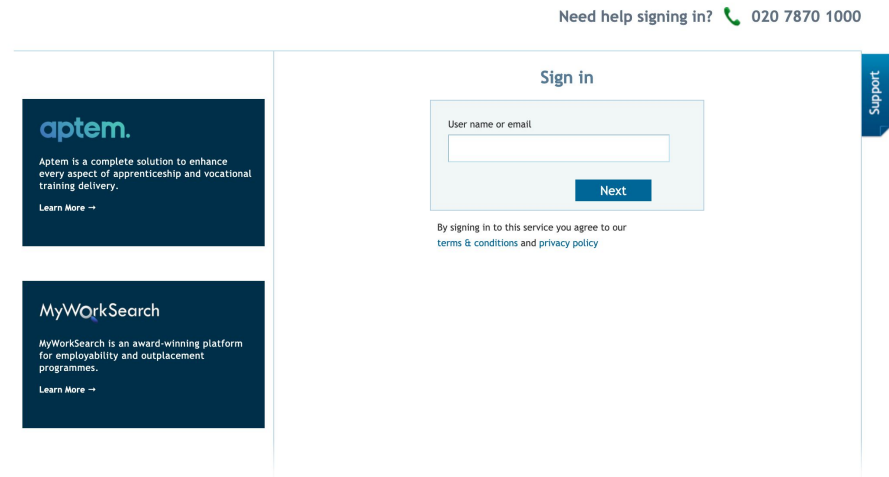
Aptem



Aptem is a complete end-to-end apprenticeship and vocational training delivery platform. It is where learners can access all information relating to their programme and learning and is also where they will build their portfolios and track the number of hours they are learning (off the job training)

Learners need access to:

<https://cambridgespark.aptem.co.uk/Users/Account/LogOn>



The screenshot shows the Aptem login interface. At the top right, there is a link for help: "Need help signing in? 020 7870 1000". The main heading is "Sign in". Below this is a form with a label "User name or email" and an input field. A "Next" button is positioned below the input field. Underneath the form, a small text line reads: "By signing in to this service you agree to our terms & conditions and privacy policy". On the right side of the page, there is a vertical "Support" button. The left sidebar contains two promotional cards: one for "aptem." with the text "Aptem is a complete solution to enhance every aspect of apprenticeship and vocational training delivery." and a "Learn More" link; and another for "MyWorkSearch" with the text "MyWorkSearch is an award-winning platform for employability and outplacement programmes." and a "Learn More" link.



Archive Files

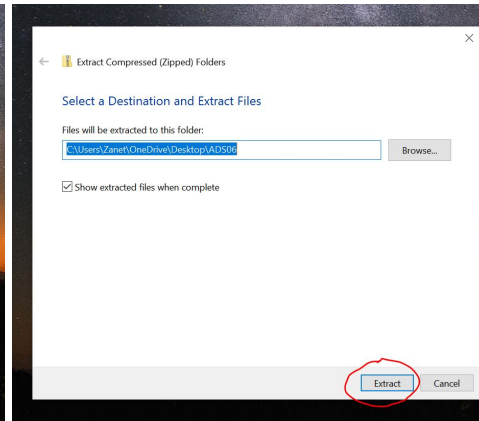
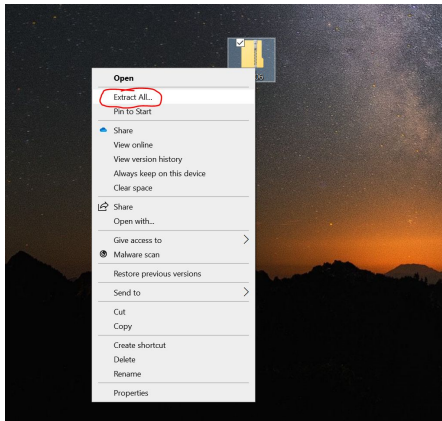


Archive Files

An **archive file** is a file that is composed of one or more computer files along with metadata. Archive files are used to collect multiple data files together into a single file for easier portability and storage, or simply to compress files to use less storage space. Filename extensions used to distinguish different types of archives include zip, rar, 7z, and tar. Our study materials are available in archive files.

How to open an archive file (windows):

1. Download the zip file to your machine
2. Right click the file, select 'Extract All' and then click 'Extract'
3. A new folder will be automatically created in the location of your file. You can now open the content in the folder.



L3 Apprenticeship Data Citizen Data Driven Journalism



This is a guide to installing our recommended set of programs that you will need on your computer during:

- 1. The L3 Data Citizen Certificate**
- 2. The L3 Data-Driven Journalism Certificate**

Computer

- A suitable computer with enough RAM and space Memory (GB) to run Microsoft Office Suite

Large or extra screen

- You will need a large screen during workshops, or if using a laptop a second screen is highly recommended

Hardware Specifications



Hardware Specifications

Your computer will need to be able to run Microsoft Office suite. Microsoft Excel's minimum system requirements are as follows:

Computer and processor

Windows OS: 1.6 GHz or faster, 2-core

macOS: Intel processor

Memory

Windows OS: 4 GB RAM; 2 GB RAM (32-bit)

macOS: 4 GB RAM

Hard disk

Windows OS: 4 GB of available disk space

macOS: 10 GB of available disk space; HFS+ hard disk format (also known as Mac OS Extended) or APFS

Updates may require additional storage over time.

Tools Specifications



Required Tools

This is the complete checklist of tools you will need to use. Instructions for each tool are provided in the following pages.

1. **Microsoft Teams** (for cohorts starting from September 2021)
2. **EDUKATE.AI**
3. **Aptem**
4. **Archive Files**
5. **Power BI or Tableau**
6. **Microsoft Office**



Power BI & Tableau



Power BI & Tableau

NOTE: You do not need both of these tools but should use the one that is accessible to you or your organisation.

Power BI

- You must have a computer with Windows as an operating system as Power BI Desktop is only available through Windows
- download and install Power BI [here](#)
- To use the Power BI Pro (Web) service a work or school email address is required to sign up, i.e. outlook or Gmail email address domains will not suffice

Tableau

- download and install Tableau Desktop or Tableau Public (it is recommended to use Tableau 2019.2 or higher)

Tableau Desktop - a licensed version of the product, obtain a 14-day free trial licence by downloading the application from [here](#) and registering it via the form that pops up when you first open Tableau Desktop. If you do not have a licence assigned to you and have already used your free trial, follow the instructions for Tableau Public below.

Tableau Public - a free version of Tableau Desktop. You can download Tableau Public [here](#). Note that you cannot save workbooks locally in the Tableau Public desktop application, but you can publish them to Tableau Public on the web to preserve them.

Additional - if your 'My Tableau Repository' is located on a network drive, you may experience slower performance. You can find out where your repository is located by clicking 'File -> Repository Location'. If this is on a network drive, it is recommended to copy this folder to a local drive (e.g. C:\) and re-point Tableau to this repository location (via File -> Repository Location')

Microsoft Office



Microsoft Office

Please make sure you have the latest versions for your required operating systems.

Microsoft Office:

- Word
- Excel desktop app
- Powerpoint

L4 Apprenticeship Data Analyst



This is a guide to installing our recommended set of programs that you will need on your computer during your L4 Data Analyst apprenticeship.

Computer

- A suitable computer with enough RAM (16GB or more)

Large or extra screen

- You will need a large screen during workshops, or if using a laptop a second screen is highly recommended

Hardware Specifications



Hardware Specifications

There are very few hardware specifications that we recommend for apprenticeships

We recommend that you have

- a hard drive capacity of at least 100GB +
- CORES, RAM as below

	IDEAL	ADEQUATE	MINIMUM
CORES	4	2	2
RAM	16 GB	16 GB	8 GB

Tools Specifications



Required Tools

This is the complete checklist of tools you will need to use. Instructions for each tool are provided in the following pages.

1. **Microsoft Teams** (for cohorts starting from September 2021)
2. **EDUKATE.AI**
3. **Aptem**
4. **Archive Files**
5. **Anaconda Python 3.x**
6. **Power BI / Tableau**
7. **Additional Software and Tools**

Anaconda Python 3.x



Anaconda

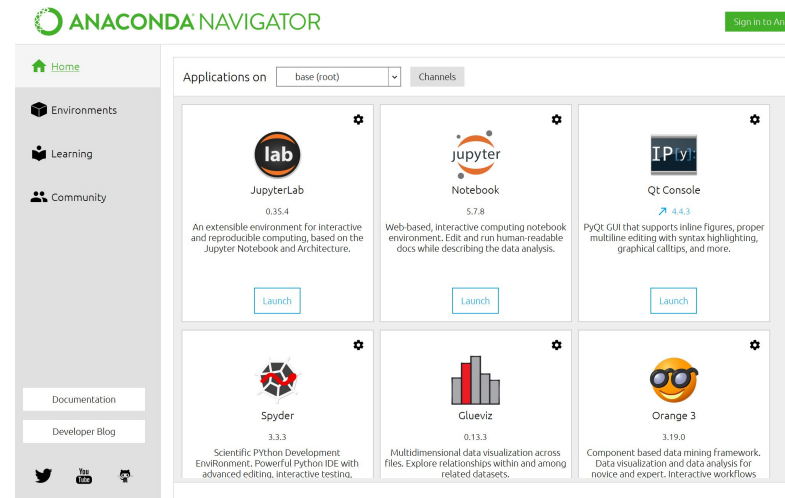


During the duration of your programme, you will need an interactive data science environment/IDE where you can run Python for work-based projects*. For this purpose, we suggest using **Anaconda** platform - please see setup instructions in the following slides.

If you cannot access/download Anaconda, please contact your Apprenticeship or IT team to find out what tools you can access within your organisation.

Anaconda provides the tools needed to easily:

- Collect data from files, databases, and data lakes
- Manage environments with Conda
- Share, collaborate on, and reproduce projects
- Deploy projects into production with the single click of a button



*learning materials and assignments can be accessed via KLOUD on EDUKATE.ai



Anaconda Python 3.x - Installation

Installation of **Anaconda Distribution** will ensure your computer has access to **Python** and many of the **packages** commonly used for data science, analysis and visualisation.

[Download](#) the **Python 3.x version** for your operating system.

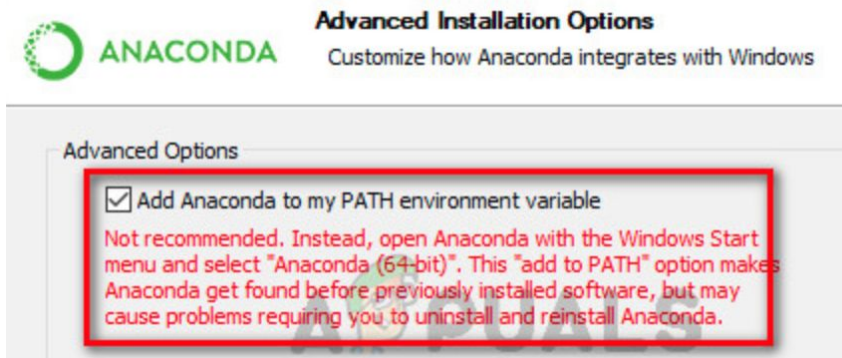
Windows | macOS | Linux

Anaconda 2019.07 for macOS Installer

<p>Python 3.7 version</p> <p>Download</p> <p>64-Bit Graphical Installer (653 MB) 64-Bit Command Line Installer (435 MB)</p>	<p>Python 2.7 version</p> <p>Download</p> <p>64-Bit Graphical Installer (634 MB) 64-Bit Command Line Installer (408 MB)</p>
--	--

Anaconda Python 3.x - Installation

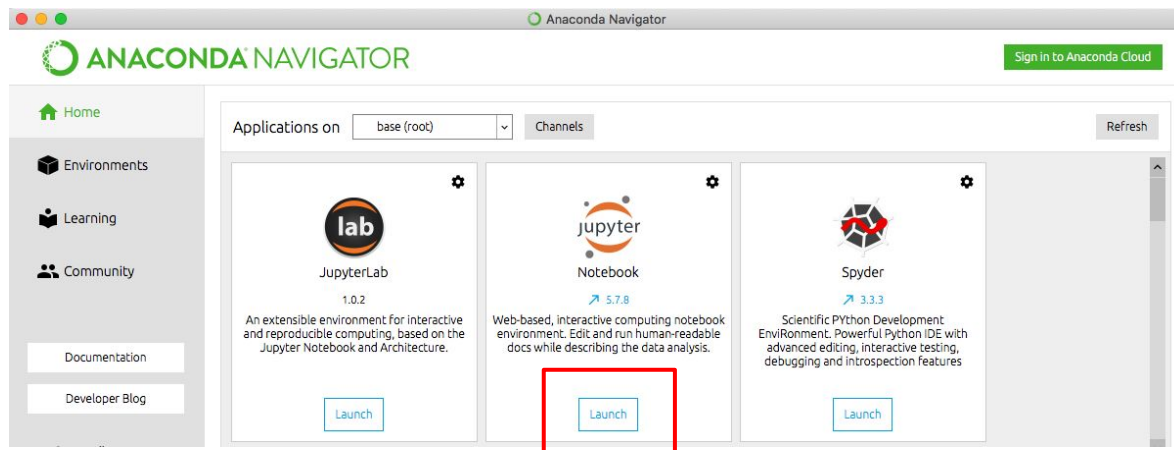
Windows users: During installation, we recommend that the "**Add Anaconda to my PATH environment variable**" checkbox is ticked (ignore the warning).



Anaconda Python 3.x - Installation

This is dealt with automatically for installations on other operating systems. The defaults are fine for the remaining settings.

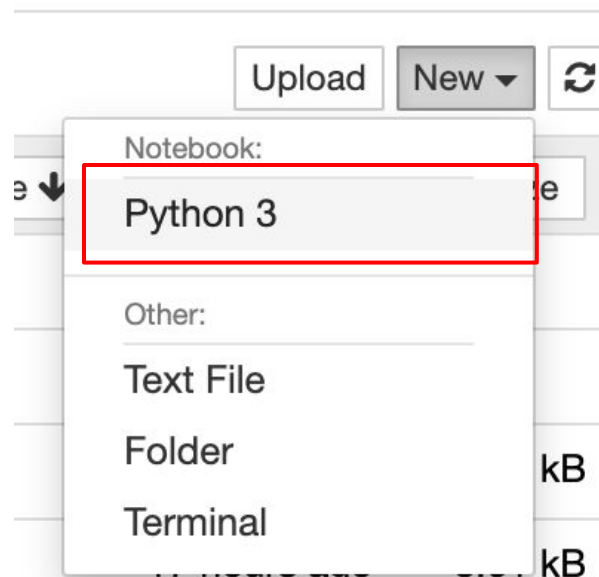
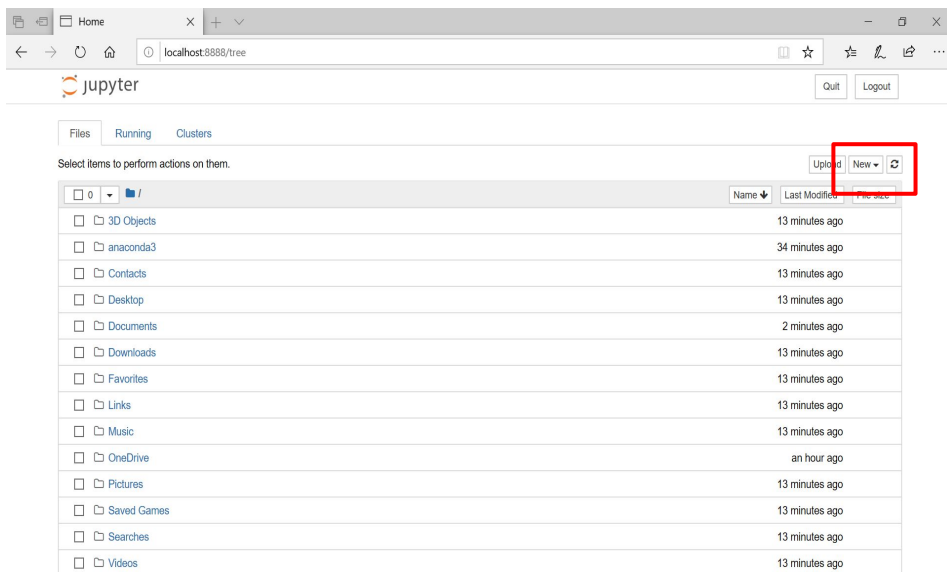
You should then be able to access [Anaconda Navigator](#), from which you will be able to launch **Jupyter Notebook**. Click on the **Launch** button in Jupyter Notebook. This will start a new tab in your default browser.



Anaconda Python 3.x - Installation

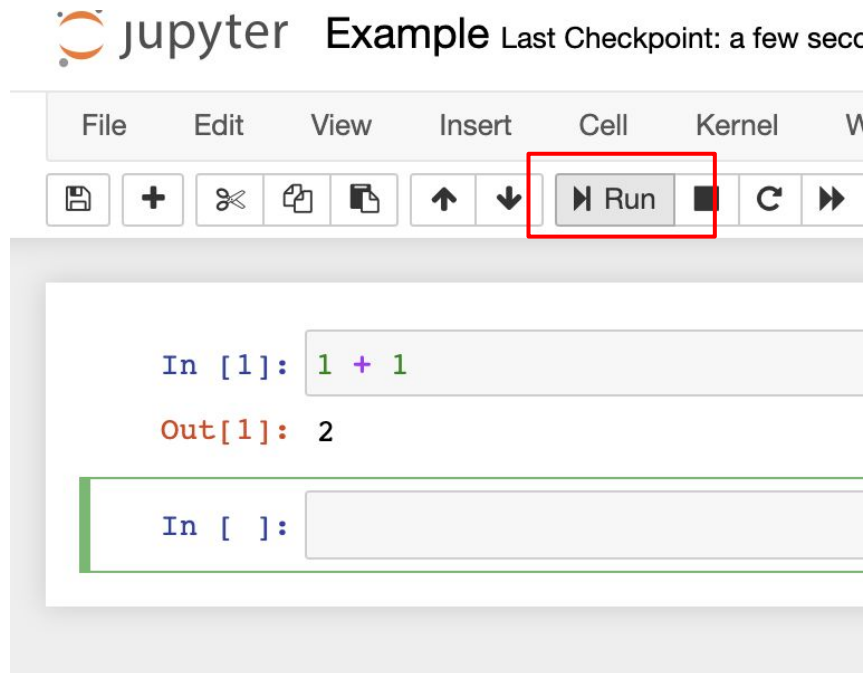
After clicking on the Launch button in Jupyter Notebook. This will start a new tab in your default browser. You can then navigate to where you stored your notebooks, data, or where you want to create new ones.

Make sure Chrome is your default browser, this tends not to work with Internet Explorer.



Anaconda Python 3.x - Installation

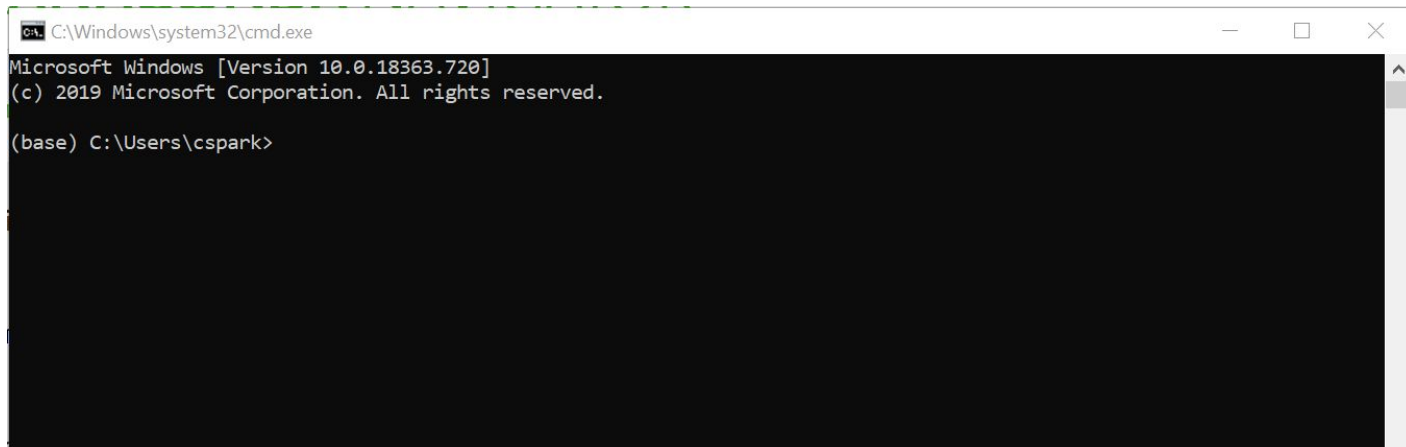
When you click on the new button, you will be able to create a new notebook. Type `1 + 1` in a cell and execute it by clicking on the run button.



The screenshot displays the Jupyter Notebook interface. At the top, the Jupyter logo is followed by the text "jupyter Example" and "Last Checkpoint: a few seconds ago". Below this is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, and Windows. Underneath the menu bar is a toolbar containing icons for saving, creating a new notebook, cutting, copying, pasting, moving up, moving down, running (highlighted with a red box), and other actions. The main area shows a code cell with the input `In [1]: 1 + 1` and the output `Out[1]: 2`. Below this is an empty code cell with the prompt `In []:`.

Anaconda Python 3.x - Installation (Advanced Notes)

With Jupyter, you will only be able to see files that are on your main disk - external harddrives or virtual drives such as OneDrive will not show by default. In order to access those, you will need to start Jupyter from the command prompt. To do so, click on Launch for the CMD.exe Prompt (or find "Anaconda Prompt" in your applications):



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.18363.720]
(c) 2019 Microsoft Corporation. All rights reserved.

(base) C:\Users\cspark>
```

Anaconda Python 3.x - Installation (Advanced Notes)

Use `cd path/to/your/folder/` to navigate to the folder of your choice. Then run `jupyter notebook` to start Jupyter from this folder. Your notebooks will be connected to this prompt, so make sure to keep it open.

```
Anaconda Prompt (anaconda3) - jupyter notebook
(base) C:\Users\cspark>cd Documents
(base) C:\Users\cspark\Documents>jupyter notebook
[I 14:26:56.963 NotebookApp] The port 8888 is already in use, trying another port.
[I 14:26:57.016 NotebookApp] JupyterLab extension loaded from C:\Users\cspark\anaconda3\lib\site-packages\jupyterlab
[I 14:26:57.016 NotebookApp] JupyterLab application directory is C:\Users\cspark\anaconda3\share\jupyter\lab
[I 14:26:57.032 NotebookApp] Serving notebooks from local directory: C:\Users\cspark\Documents
[I 14:26:57.032 NotebookApp] The Jupyter Notebook is running at:
[I 14:26:57.032 NotebookApp] http://localhost:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
[I 14:26:57.032 NotebookApp] or http://127.0.0.1:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
[I 14:26:57.032 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 14:26:57.055 NotebookApp]

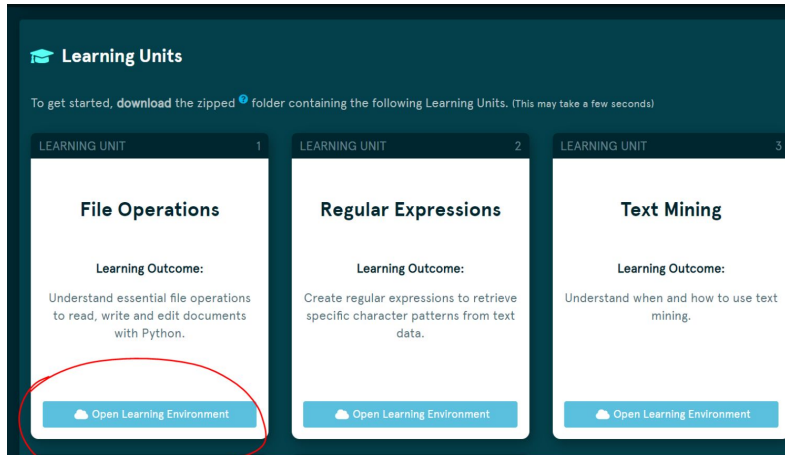
To access the notebook, open this file in a browser:
    file:///C:/Users/cspark/AppData/Roaming/jupyter/runtime/nbserver-1496-open.html
Or copy and paste one of these URLs:
    http://localhost:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
    or http://127.0.0.1:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
```

EDUKATE.AI



We have developed a cloud-based learning environment for our EDUKATE.AI platform called KLOUD.

On EDUKATE. AI, via the **Open Learning Environment** button, you will be able to access all learning units and assignments. Clicking on the button will spin a new Jupyter environment, hosted on our servers and pre-loaded with all the libraries you will need. This means that you can go through the materials directly in the browser, without local setup needed.



The screenshot shows the 'Learning Units' page. At the top, there is a header 'Learning Units' with a graduation cap icon. Below it, a message says: 'To get started, download the zipped folder containing the following Learning Units. (This may take a few seconds)'. There are three learning unit cards:

- Learning Unit 1: File Operations**
Learning Outcome: Understand essential file operations to read, write and edit documents with Python.
Button: Open Learning Environment
- Learning Unit 2: Regular Expressions**
Learning Outcome: Create regular expressions to retrieve specific character patterns from text data.
Button: Open Learning Environment
- Learning Unit 3: Text Mining**
Learning Outcome: Understand when and how to use text mining.
Button: Open Learning Environment

The 'Open Learning Environment' button for the first card is circled in red.



The screenshot shows the KLOUD BETA file explorer interface. At the top, there are tabs for 'Files', 'Running', and 'Clusters'. Below the tabs, there is a message: 'Select items to perform actions on them.' To the right of this message are 'Upload' and 'New' buttons. Below the message is a table of files and folders:

	Name	Last Modified	File size
<input type="checkbox"/>	/		
<input type="checkbox"/>	additional_materials	7 minutes ago	
<input type="checkbox"/>	lessons	7 minutes ago	
<input type="checkbox"/>	practicals	7 minutes ago	

Power BI & Tableau



Power BI & Tableau

NOTE: You do not need both of these tools but should use the one that is accessible to you or your organisation.

Power BI

- You must have a computer with Windows as an operating system as Power BI Desktop is only available through Windows
- download and install Power BI [here](#)
- To use the Power BI Pro (Web) service a work or school email address is required to sign up, i.e. outlook or Gmail email address domains will not suffice

Tableau

- download and install Tableau Desktop or Tableau Public (it is recommended to use Tableau 2019.2 or higher)

Tableau Desktop - a licensed version of the product, obtain a 14-day free trial licence by downloading the application from [here](#) and registering it via the form that pops up when you first open Tableau Desktop. If you do not have a licence assigned to you and have already used your free trial, follow the instructions for Tableau Public below.

Tableau Public - a free version of Tableau Desktop. You can download Tableau Public [here](#). Note that you cannot save workbooks locally in the Tableau Public desktop application, but you can publish them to Tableau Public on the web to preserve them.

Additional - if your 'My Tableau Repository' is located on a network drive, you may experience slower performance. You can find out where your repository is located by clicking 'File -> Repository Location'. If this is on a network drive, it is recommended to copy this folder to a local drive (e.g. C:\) and re-point Tableau to this repository location (via File -> Repository Location')

Additional Software & Tools

Additional Software and Tools (L4/Foundations)

The following list includes the libraries and tools that are needed for the Level 4 and Foundational Modules - M01-M13. Please go to each of these websites and download the latest versions for your required operating systems.

Please note there may be additional download requirements and this list is subject to change.

Python Libraries:

python=3.8 ([Windows download info](#)) ([Mac Download info](#))
bokeh==2.2.2 ([Download info](#))
pandas==1.1.3 ([Download info](#))
pandas-bokeh==0.5 ([Download info](#))
matplotlib==3.3.2 ([Download info](#))
seaborn==0.11.0 ([Download info](#))
panel==0.10.1 ([Download info](#))
Neo4js ([Download info](#))
Java 11 ([Download info](#))
postgresql-12 ([Download info](#))
ipython-sql ([Download info](#))
psycopg2 ([Download info](#))

py2neo ([Download info](#))
holidays=0.10.1 ([Download info](#))
python-graphviz=0.13.2 ([Download info](#))
pandas-datareader=0.8.1 ([Download info](#))
xlrd=1.2.0 ([Download info](#))
plotly=4.9.0 ([Download info](#))
numpy==1.19.1 ([Download info](#))
sympy==1.6.2 ([Download info](#))
scipy==1.5.0 ([Download info](#))
scikit-learn>=0.20.0 ([Download info](#))
ipython==7.15.0 ([Download info](#))
ipywidgets==7.5.1 ([Download info](#))
tqdm==4.50.0 ([Download info](#))

L4 Digital Business Analyst (Digital Business Lead) Apprenticeship



This is a guide to installing our recommended set of programs that you will need on your computer during:

- 1. Digital Business Lead**
- 2. Digital Development**
- 3. Digital Business Analyst**

Computer

- A suitable computer with enough RAM and space Memory (GB) to run Microsoft Office Suite

Large or extra screen

- You will need a large screen during workshops, or if using a laptop a second screen is highly recommended

Hardware Specifications



Hardware Specifications

Your computer will need to be able to run Microsoft Office suite. Microsoft Excel's minimum system requirements are as follows:

Computer and processor

Windows OS: 1.6 GHz or faster, 2-core

macOS: Intel processor

Memory

Windows OS: 4 GB RAM; 2 GB RAM (32-bit)

macOS: 4 GB RAM

Hard disk

Windows OS: 4 GB of available disk space

macOS: 10 GB of available disk space; HFS+ hard disk format (also known as Mac OS Extended) or APFS

Updates may require additional storage over time.

Tools Specifications



Required Tools

This is the complete checklist of tools you will need to use. Instructions for each tool are provided in the following pages.

- 1. Microsoft Teams** (for cohorts starting from September 2021)
- 2. EDUKATE.AI**
- 3. Aptem**
- 4. Archive Files**
- 5. Microsoft Office**
- 6. Process mapping tools e.g. Microsoft Visio**
- 7. Miro**
- 8. Access to video sharing platforms e.g. YouTube and TedTalks**

Microsoft Office



Microsoft Office

Please make sure you have the latest versions for your required operating systems.

Microsoft Office:

- Word
- Excel
- Powerpoint

Process Mapping Tools



Process Mapping Tools

Some modules will require you to use process mapping tools to develop visual representations of current and future business processes. The modules don't favour a particular tool and you can use the one used by your organisation however if you don't currently use one any of the following will be sufficient:

- Microsoft Visio
- Signavio
- Lucidchart (free version)
- Draw.io (free version)

Miro



Miro

Some modules will require you to access a miro board during the live workshop to complete interactive activities.

A miro board is an interactive digital whiteboard that will allow you to draw, upload files and add notes to a shared space as you complete activities as a class or in breakout sessions. If you don't already have a miro profile you can set one up for free

<https://help.miro.com/hc/en-us>

L7 Apprenticeships AI (all pathways)



This is a guide to installing our recommended set of programs that you will need on your computer during your L7 AI & Data Specialist apprenticeship.

Computer

- A suitable computer with enough RAM (16GB or more)

Large or extra screen

- You will need a large screen during workshops, or if using a laptop a second screen is highly recommended

Hardware Specifications



Hardware Specifications

There are very few hardware specifications that we recommend for apprenticeships

We recommend that you have

- a hard drive capacity of at least 100GB +
- CORES, RAM as below

	IDEAL	ADEQUATE	MINIMUM
CORES	4	2	2
RAM	16 GB	16 GB	8 GB

Tools Specifications



Required Tools

This is the complete checklist of tools you will need to use. Instructions for each tool are provided in the following pages.

1. **Microsoft Teams** (for cohorts starting from September 2021)
2. **EDUKATE.AI**
3. **Aptem**
4. **Archive Files**
5. **Anaconda Python 3.x**
6. **Microsoft VS Code & Git**
7. **Additional Software and Tools**

Anaconda Python 3.x



Anaconda

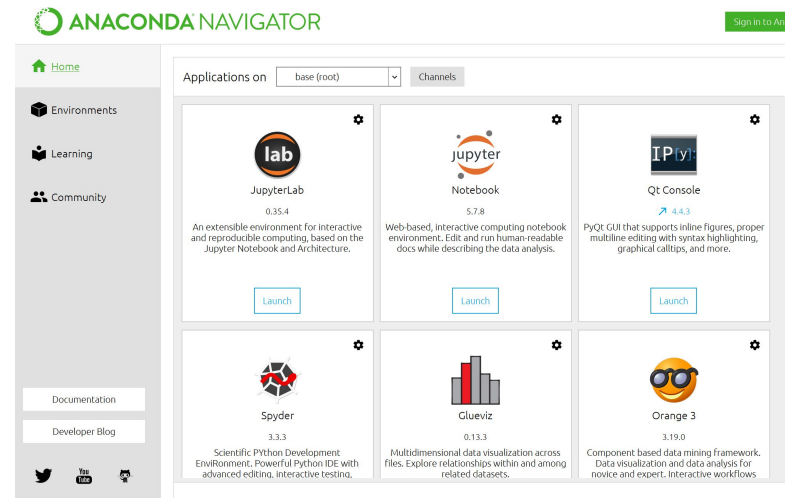


During the duration of your programme, you will need an interactive data science environment/IDE where you can run Python for work-based projects*. For this purpose, we suggest using **Anaconda** platform - please see setup instructions in the following slides.

If you cannot access/download Anaconda, please contact your Apprenticeship or IT team to find out what tools you can access within your organisation.

Anaconda provides the tools needed to easily:

- Collect data from files, databases, and data lakes
- Manage environments with Conda
- Share, collaborate on, and reproduce projects
- Deploy projects into production with the single click of a button






*learning materials and assignments can be accessed via KLOUD on EDUKATE.ai



Anaconda Python 3.x - Installation

Installation of **Anaconda Distribution** will ensure your computer has access to **Python** and many of the **packages** commonly used for data science, analysis and visualisation.

[Download](#) the **Python 3.x version** for your operating system.

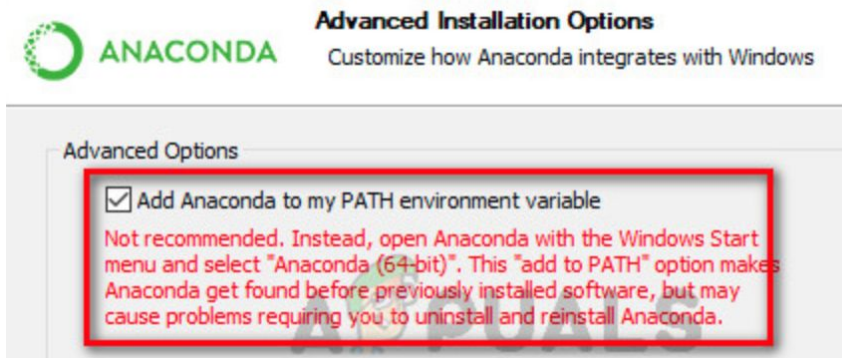
 Windows |  macOS |  Linux

Anaconda 2019.07 for macOS Installer

<h4>Python 3.7 version</h4> <p>Download</p> <p>64-Bit Graphical Installer (653 MB) 64-Bit Command Line Installer (435 MB)</p>	<h4>Python 2.7 version</h4> <p>Download</p> <p>64-Bit Graphical Installer (634 MB) 64-Bit Command Line Installer (408 MB)</p>
---	---

Anaconda Python 3.x - Installation

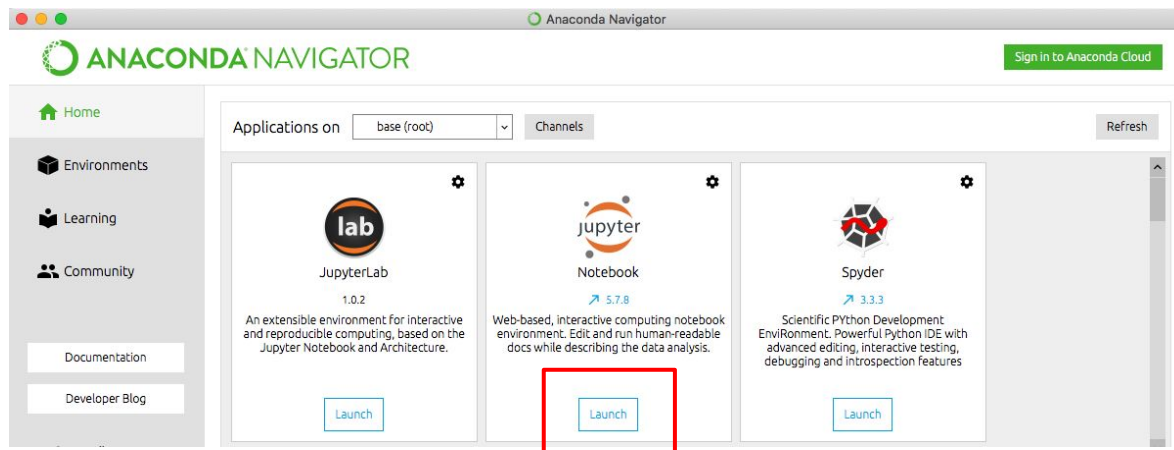
Windows users: During installation, we recommend that the "**Add Anaconda to my PATH environment variable**" checkbox is ticked (ignore the warning).



Anaconda Python 3.x - Installation

This is dealt with automatically for installations on other operating systems. The defaults are fine for the remaining settings.

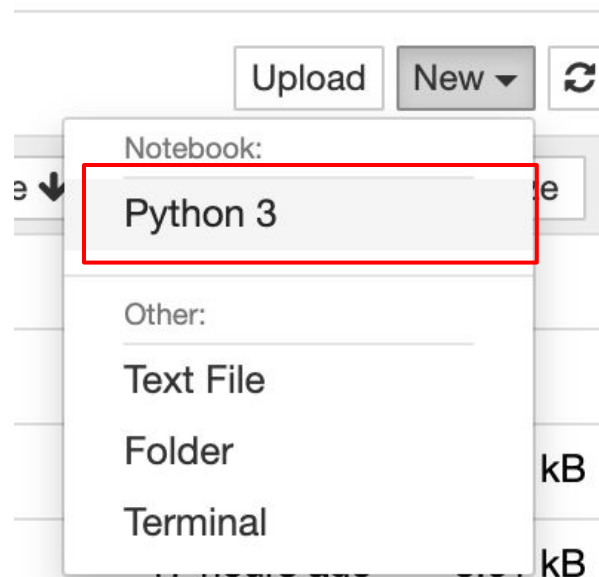
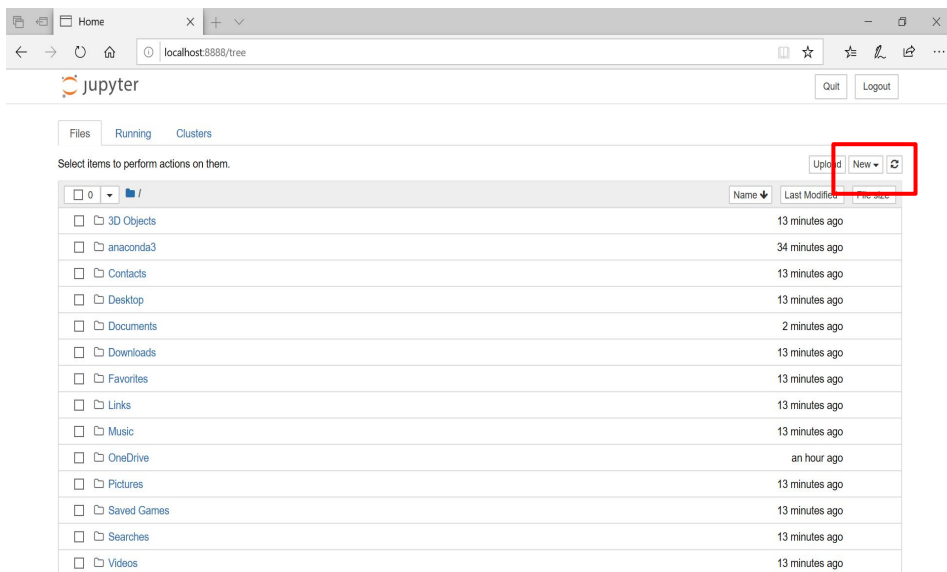
You should then be able to access [Anaconda Navigator](#), from which you will be able to launch **Jupyter Notebook**. Click on the **Launch** button in Jupyter Notebook. This will start a new tab in your default browser.



Anaconda Python 3.x - Installation

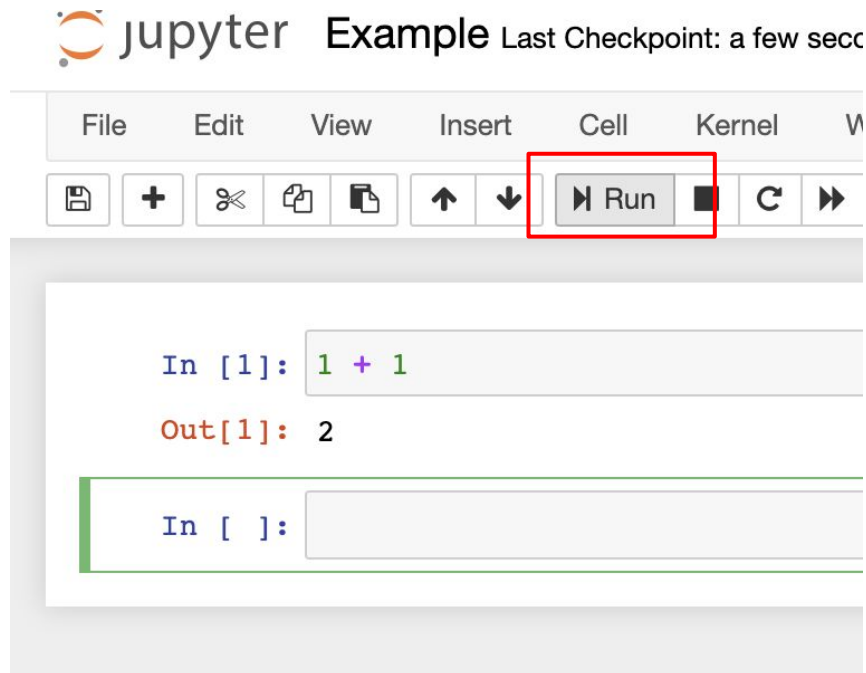
After clicking on the Launch button in Jupyter Notebook. This will start a new tab in your default browser. You can then navigate to where you stored your notebooks, data, or where you want to create new ones.

Make sure Chrome is your default browser, this tends not to work with Internet Explorer.



Anaconda Python 3.x - Installation

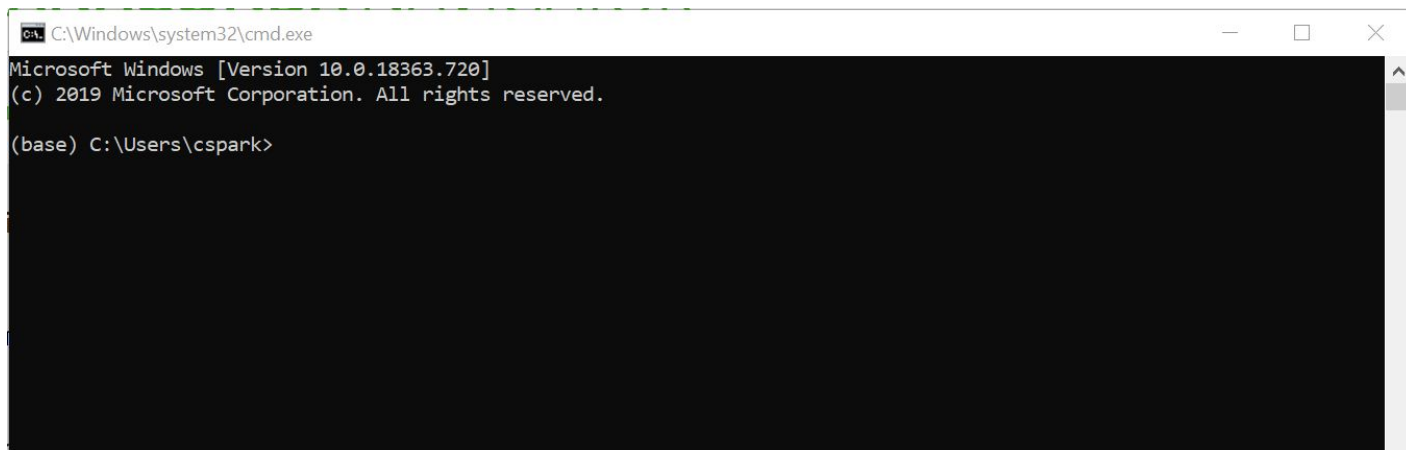
When you click on the new button, you will be able to create a new notebook. Type `1 + 1` in a cell and execute it by clicking on the run button.



The screenshot displays the Jupyter Notebook interface. At the top, the Jupyter logo is followed by the text "jupyter Example" and "Last Checkpoint: a few seconds ago". Below this is a menu bar with options: File, Edit, View, Insert, Cell, Kernel, and Windows. A toolbar contains various icons for file operations (save, new, copy, paste, undo, redo) and execution (run, stop, refresh, next). The "Run" button, which features a play icon and the text "Run", is highlighted with a red rectangular box. Below the toolbar, a code cell is shown with the input "In [1]: 1 + 1" and the output "Out[1]: 2". A new, empty code cell is visible below, labeled "In []:".

Anaconda Python 3.x - Installation (Advanced Notes)

With Jupyter, you will only be able to see files that are on your main disk - external hard drives or virtual drives such as OneDrive will not show by default. In order to access those, you will need to start Jupyter from the command prompt. To do so, click on Launch for the CMD.exe Prompt (or find "Anaconda Prompt" in your applications):



```
C:\Windows\system32\cmd.exe
Microsoft Windows [Version 10.0.18363.720]
(c) 2019 Microsoft Corporation. All rights reserved.

(base) C:\Users\cspark>
```

Anaconda Python 3.x - Installation (Advanced Notes)

Use `cd path/to/your/folder/` to navigate to the folder of your choice. Then run `jupyter notebook` to start Jupyter from this folder. Your notebooks will be connected to this prompt, so make sure to keep it open.

```
Anaconda Prompt (anaconda3) - jupyter notebook
(base) C:\Users\cspark>cd Documents
(base) C:\Users\cspark\Documents>jupyter notebook
[I 14:26:56.963 NotebookApp] The port 8888 is already in use, trying another port.
[I 14:26:57.016 NotebookApp] JupyterLab extension loaded from C:\Users\cspark\anaconda3\lib\site-packages\jupyterlab
[I 14:26:57.016 NotebookApp] JupyterLab application directory is C:\Users\cspark\anaconda3\share\jupyter\lab
[I 14:26:57.032 NotebookApp] Serving notebooks from local directory: C:\Users\cspark\Documents
[I 14:26:57.032 NotebookApp] The Jupyter Notebook is running at:
[I 14:26:57.032 NotebookApp] http://localhost:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
[I 14:26:57.032 NotebookApp] or http://127.0.0.1:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
[I 14:26:57.032 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 14:26:57.055 NotebookApp]

To access the notebook, open this file in a browser:
file:///C:/Users/cspark/AppData/Roaming/jupyter/runtime/nbserver-1496-open.html
Or copy and paste one of these URLs:
http://localhost:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
or http://127.0.0.1:8889/?token=88ae85a96ba8fe0508bb6e3267b0012da7c64d2e55d8095c
```

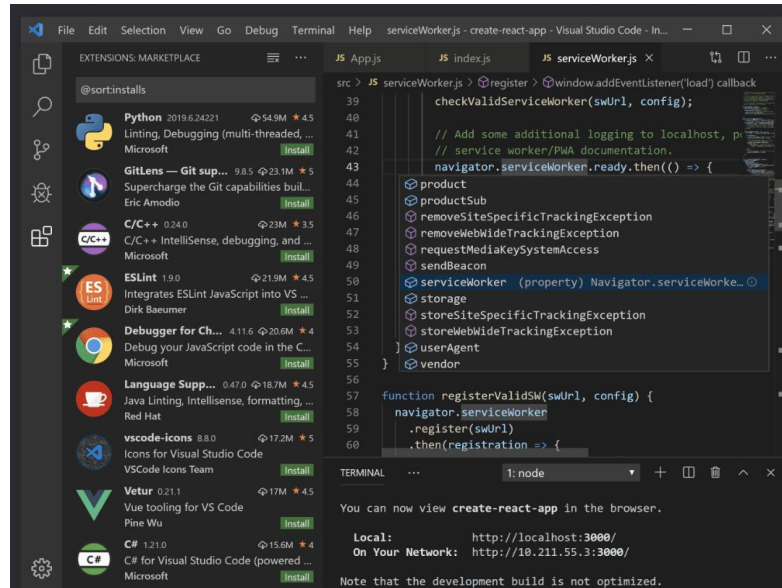
Microsoft VS Code



Microsoft VS Code

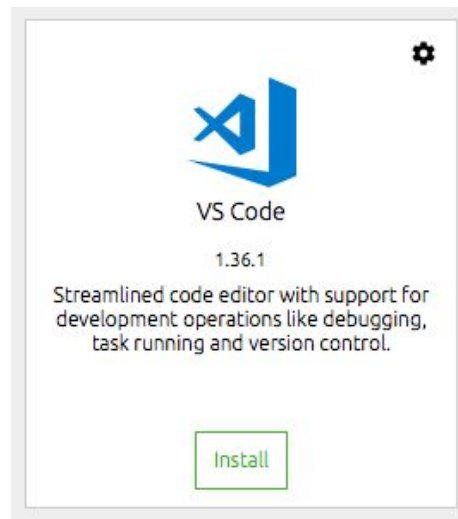
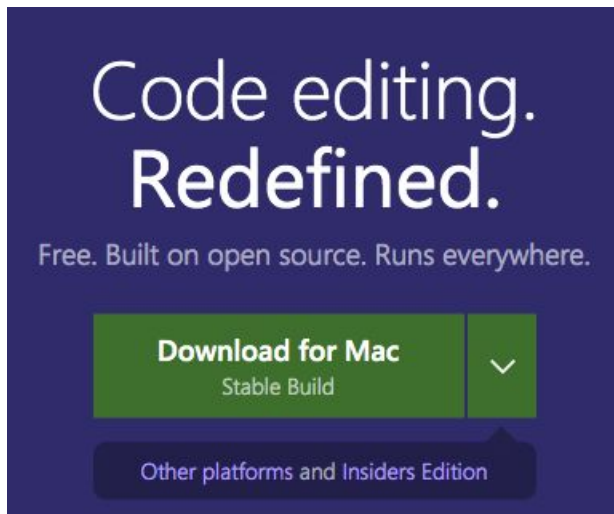


Visual Studio Code is a source-code editor developed by Microsoft for Windows, Linux and macOS. It includes support for debugging, embedded Git control and GitHub, syntax highlighting, intelligent code completion, snippets, and code refactoring.



Microsoft VS Code

Microsoft VS Code is a free, cross-platform code editor. It's available [here](#) (choose the **Stable** version for your operating system), or in **Anaconda Navigator**.

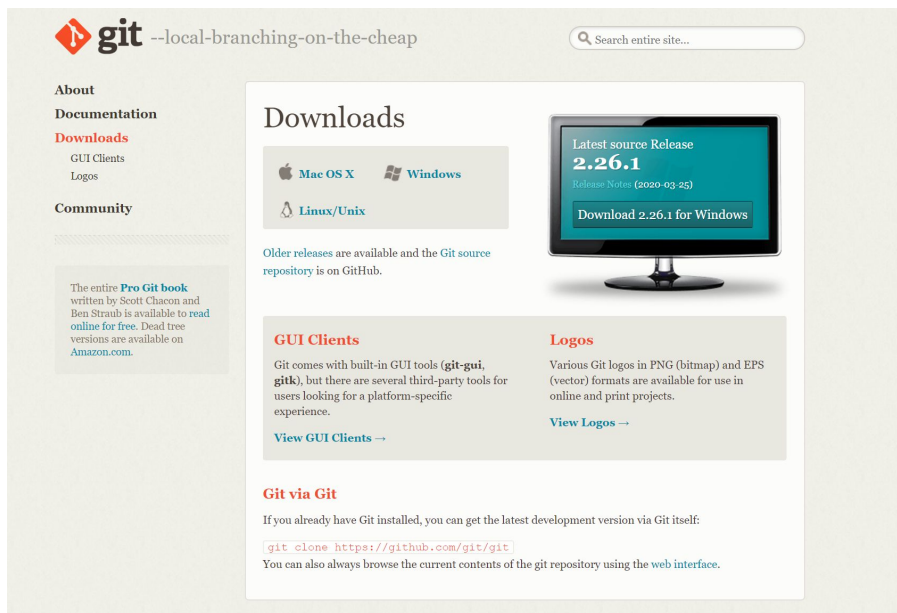


Git



Git is a **version control system** used by many systems for **software development and collaboration**.

[Download Git](#) for your operating system, then follow the installation instructions.



The screenshot shows the Git website's Downloads page. At the top left is the Git logo and the tagline "--local-branching-on-the-cheap". A search bar is located at the top right. On the left side, there is a navigation menu with links for "About", "Documentation", "Downloads" (highlighted), "GUI Clients", "Logos", and "Community". Below the menu, there is a text box mentioning the "Pro Git book" by Scott Chacon and Ben Straub. The main content area is titled "Downloads" and features a "Latest source Release" section with the version number "2.26.1" and a "Download 2.26.1 for Windows" button. Below this, there are sections for "GUI Clients" and "Logos". At the bottom, there is a "Git via Git" section with a code snippet for cloning the repository.

git --local-branching-on-the-cheap

Search entire site...

About

Documentation

Downloads

GUI Clients

Logos

Community

The entire **Pro Git book** written by Scott Chacon and Ben Straub is available to read online for free. Dead tree versions are available on Amazon.com.

Downloads

Mac OS X Windows

Linux/Unix

Latest source Release
2.26.1
Release Notes (2020-03-25)
Download 2.26.1 for Windows

Older releases are available and the Git source repository is on GitHub.

GUI Clients

Git comes with built-in GUI tools (**git-gui**, **gitk**), but there are several third-party tools for users looking for a platform-specific experience.

[View GUI Clients →](#)

Logos

Various Git logos in PNG (bitmap) and EPS (vector) formats are available for use in online and print projects.

[View Logos →](#)

Git via Git

If you already have Git installed, you can get the latest development version via Git itself:

```
git clone https://github.com/git/git
```

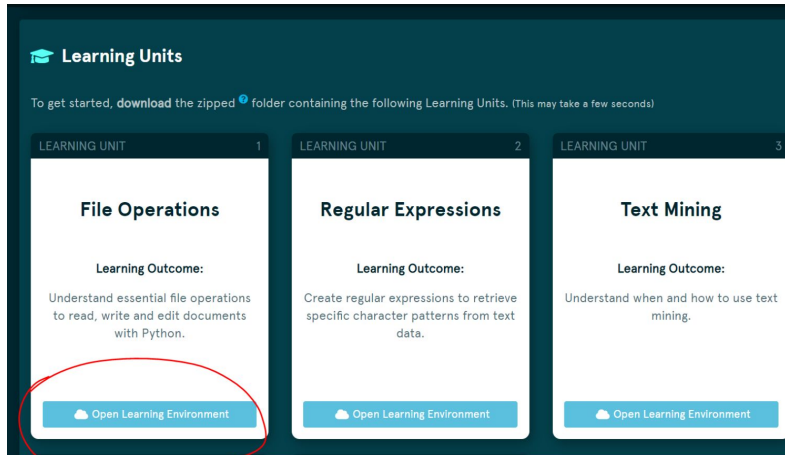
You can also always browse the current contents of the git repository using the web interface.

EDUKATE.AI



We have developed a cloud-based learning environment for our EDUKATE.AI platform called KLOUD.

On EDUKATE. AI, via the **Open Learning Environment** button, you will be able to access all learning units and assignments. Clicking on the button will spin a new Jupyter environment, hosted on our servers and pre-loaded with all the libraries you will need. This means that you can go through the materials directly in the browser, without local setup needed.



The screenshot shows the 'Learning Units' page. At the top, there is a header 'Learning Units' with a graduation cap icon. Below it, a message says: 'To get started, download the zipped folder containing the following Learning Units. (This may take a few seconds)'. There are three learning unit cards:

- Learning Unit 1: File Operations**
Learning Outcome: Understand essential file operations to read, write and edit documents with Python.
Button: Open Learning Environment (circled in red)
- Learning Unit 2: Regular Expressions**
Learning Outcome: Create regular expressions to retrieve specific character patterns from text data.
Button: Open Learning Environment
- Learning Unit 3: Text Mining**
Learning Outcome: Understand when and how to use text mining.
Button: Open Learning Environment



The screenshot shows the KLOUD BETA file explorer interface. It has tabs for 'Files', 'Running', and 'Clusters'. Below the tabs, there is a message: 'Select items to perform actions on them.' and buttons for 'Upload' and 'New'. A table lists the files and folders:

	Name	Last Modified	File size
<input type="checkbox"/>	/		
<input type="checkbox"/>	additional_materials	7 minutes ago	
<input type="checkbox"/>	lessons	7 minutes ago	
<input type="checkbox"/>	practicals	7 minutes ago	

Additional Software & Tools

Additional Software and Tools (L7/Advanced)

The following list includes all the libraries and tools that are needed for the Level 7 Modules (some of which you may already have installed). Please ensure you have downloaded the latest versions of each for your required operating systems.

Please note there may be additional download requirements and this list is subject to change.

python=3.8 ([Windows download info](#)) ([Mac Download info](#))
bokeh==2.2.2 ([Download info](#))
pandas==1.1.3 ([Download info](#))
pandas-bokeh==0.5 ([Download info](#))
matplotlib==3.3.2 ([Download info](#))
seaborn==0.11.0 ([Download info](#))
panel==0.10.1 ([Download info](#))
git version 2.21.0 ([Mac Download](#))
git version 2.28.0.msysgit.0 ([Windows](#))
Neo4js ([Download info](#))
Java 11 ([Download info](#))
postgresql-12 ([Download info](#))
ipython-sql ([Download info](#))
psycpg2 ([Download info](#))
Sqlalchemy ([Download info](#))
py2neo ([Download info](#))

numpy==1.19.1 ([Download info](#))
sympy==1.6.2 ([Download info](#))
scipy==1.5.0 ([Download info](#))
scikit-learn>=0.20.0 ([Download info](#))
ipython==7.15.0 ([Download info](#))
ipywidgets==7.5.1 ([Download info](#))
tqdm==4.50.0 ([Download info](#))
holidays=0.10.1 ([Download info](#))
fbprophet=0.6 ([Download info](#))
python-graphviz=0.13.2 ([Download info](#))
pandas-datareader=0.8.1 ([Download info](#))
xlrd=1.2.0 ([Download info](#))
plotly=4.9.0 ([Download info](#))
hypothesis ([Download info](#))
spacy ([Download info](#))
mock ([Download info](#))



Solutions for Data Science

Development | Assessment | Attraction

cambridgespark.com