

Installation Instructions

Windows

Install Anaconda (Recommended)

Anaconda is free open-source software designed and built for data science and machine learning. It has packages for Python and R, as well as programmes that let you write and run code in them (called “integrated development environments” or IDEs) – for Python this is the “Spyder” IDE and for R, you may have come across “R-Studio”.

We strongly urge that you install Python via Anaconda, because it will automatically come packaged with all the individual modules/packages that you need to get started with this course (and in your later data science work). If you do not use Anaconda, you will need to install individual packages yourself, which can be tricky.

Anaconda is free to install, and you do not need to create a user account if you don’t want to. There is a detailed guide, including installation instruction videos here: (<https://www.anaconda.com/docs/getting-started/anaconda/install#windows-installation>). The installation slightly differs for Mac and for Windows users.

Alternatively, follow these few steps:

- Go to the Anaconda Downloads website here: <https://www.anaconda.com/download/success>
- There are two options: “Distribution Installers” and “Miniconda Installers”, you want the “Distribution Installers”
- Click the button underneath “Distribution Installers” that says “Windows” and download the “64-Bit Graphical Installer”
- Once the Download has completed, double click on the downloaded file
- This should open up a window that will guide you through the installation process

The installation process should take about 15 minutes, and will take up about 5GB of space on your laptop.

Once the installation is complete, you should find an application called “Anaconda Navigator” on your computer. If you do not, please try to double click the installer again and follow the instructions again. Make sure that you select the options that we showed above.

Mac

Install Anaconda (Recommended)

Anaconda is free open-source software designed and built for data science and machine learning. It has packages for Python and R, as well as programmes that let you write and run code in them (called “integrated development environments” or IDEs) – for Python this is the “Spyder” IDE and for R, you may have come across “R-Studio”.

We strongly urge that you install Python via Anaconda, because it will automatically come packaged with all the individual modules/packages that you need to get started with this course (and in your later data science work). If you do not use Anaconda, you will need to install individual packages yourself, which can be tricky.

Anaconda is free to install, and you do not need to create a user account if you don’t want to. There is a detailed guide, including installation instruction videos here: (<https://www.anaconda.com/docs/getting-started/anaconda/install#windows-installation>). The installation slightly differs for Mac and for Windows users.

Alternatively, follow these few steps:

- Go to the Anaconda Downloads website here: <https://www.anaconda.com/download/success>
- There are two options: “Distribution Installers” and “Miniconda Installers”, you want the “Distribution Installers”
- Click the button underneath “Distribution Installers” that says “Mac” and download the “64-Bit Graphical Installer”
 - Make sure that you use the correct version (“Apple Silicon” or “Intel”)
 - If your Mac is from 2020 onwards, or is an M1/M2/M3/M4 Mac, then it will be “Apple Silicon”
- Once the Download has completed, double click on the downloaded file
- This should open up a window that will guide you through the installation process

The installation process should not more than 15 minutes, and will take up about 5GB of space on your laptop.

Once the installation is complete, you should find an application called “Anaconda Navigator” on your computer. If you do not, please try to double click the installer again and follow the instructions again. Make sure that you select the options that we showed above.

Using Jupyter Notebooks

When working with python, you will see two file extensions going around “.py” and “.ipynb”.

“.py” files are pure Python files. These can be run using an IDE, or from a command line, and will carry out the instructions in the code on your computer’s processor. The downside is that you can’t interact with the programme while it runs! In this course, we will not be working with “.py” files.

Instead, everyone in data science uses Jupyter Notebooks to run their code. These are given by the “.ipynb” file extension. These are a way to structure a project so that you can have code, text and results all in one place! This is great for distribution, for experimentation, and for when you come back to check on the code.

In this course, we will show you how to use “.ipynb” files.

Opening the Jupyter Notebook Viewer

- Go to the Anaconda Navigator.
- This will show you all of the programmes that are packaged with Anaconda that you can use.
- Scroll down until you can see “JupyterLab” and click “Launch”
 - Note, this is NOT the same as “Notebook”
- After you click “Launch”, a window in your internet browser will open
 - This can take up to a minute, so be patient and do not click “Launch” multiple times
- The window that opens up contains the “JupyterLab” which is the environment we will be working in that allows us to run Jupyter Notebooks!
- On the left hand side of the screen, you can see all of your folders
- Navigate to where you have saved the Course Content
 - E.g. Click on Documents if you have saved the Course Content into the Documents folder
- If you click on an “.ipynb” file, which is a Jupyter Notebook, you will be able to open and view it

- When you click on an “.ipynb” file, you will be asked to “Select a Kernel”
 - The “kernel” is the version of Python that you want to use to run the Notebook
 - Make sure to pick “Python [conda env:base]” to avoid later problems!

You can also start a new notebook for yourself!

- Once you are on the JupyterLab page that opens up from the Anaconda Navigator, you should see a page like this
- This shows all of the versions of Python that you are able to open
- Make sure you click on “Python [conda env:base]” under “Notebook”
 - This will open up a Jupyter Notebook using the conda env:base kernel which you can then write code in

Running code in Jupyter Notebooks

- Jupyter Notebooks are a way for you to interactively run your code
- If you have some code in a notebook, you can run it by using the buttons at the top of the screen, or alternatively by using some key combinations
- The notebook is divided into “cells”
 - Each of these can contain code, outputs or plain text
- When you are inside a “cell”, by using “Ctrl + Enter”, the code inside that cell will be executed
- You can delete or move cells using the buttons on the right hand side of the cell
- You can add a new cell below your current one using the “+” symbol next to the save symbol on the top of the JupyterLab window
- Cells can also contain plain text that is not code
- This is useful for describing in words what you are trying to do. It is good practice to be as descriptive as possible in these cells
- You can change a cell from “Code” to plain text (which is called “Markdown”) using the Drop Down Menu at the top of the panel

Checklist

Before the course, make sure that you have done the following. It will mean that we can get on with the data science and precision medicine, rather than focusing on Python installation.

- Install the Anaconda distribution (as described above)
- Make sure that you are able to open the “Anaconda Navigator” application and that it opens as expected
- Make sure that you can open the “JupyterLab” application from “Anaconda Navigator” and that it opens as described here

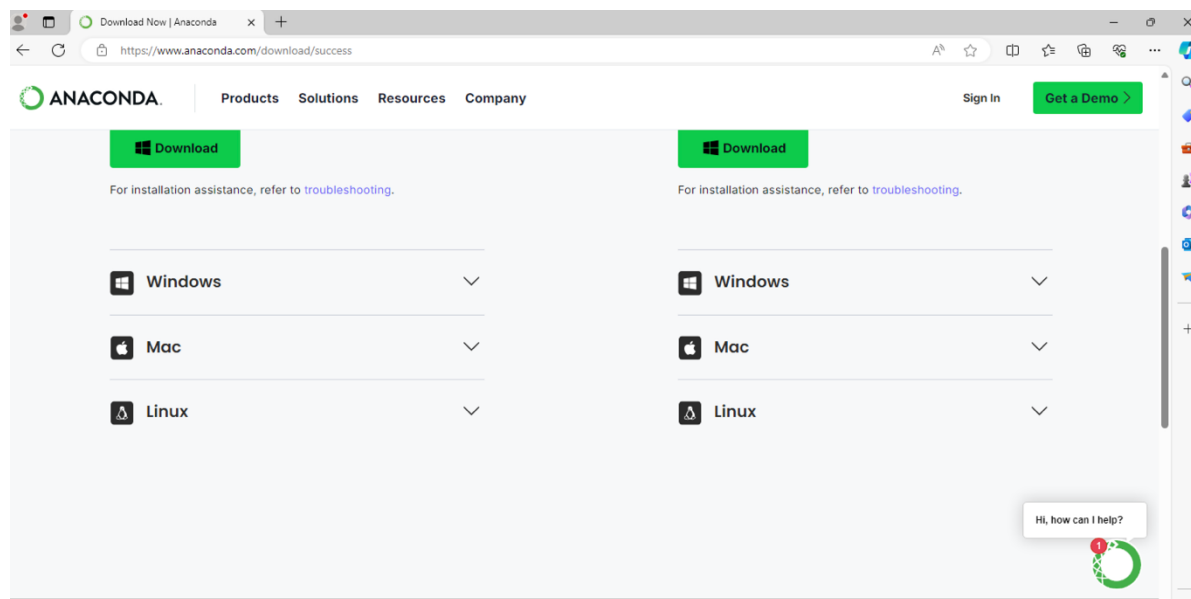
If you can do all of these things, you will be well set for the course!

Do not worry about all of the features in Anaconda/JupyterLab! We will gradually work our way through these as we do the course!

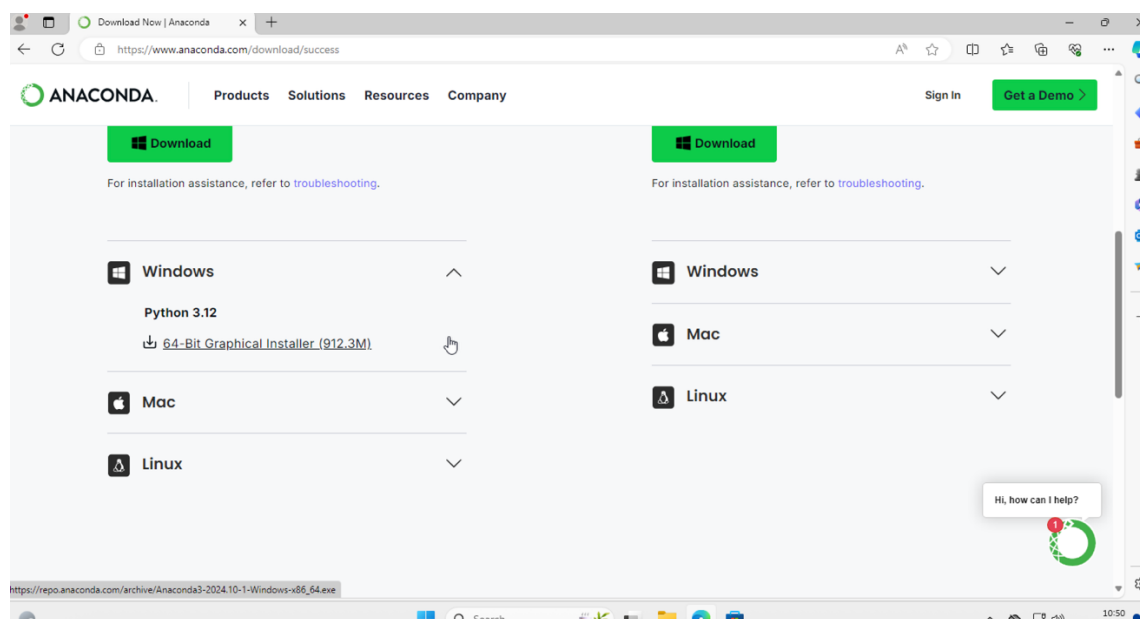
Photos

These photos show the steps of installing Anaconda/Jupyter notebooks. If stuck, you can look through these to work out the options that you need to pick!

1. When trying to download the Anaconda distribution, you should see a page like this

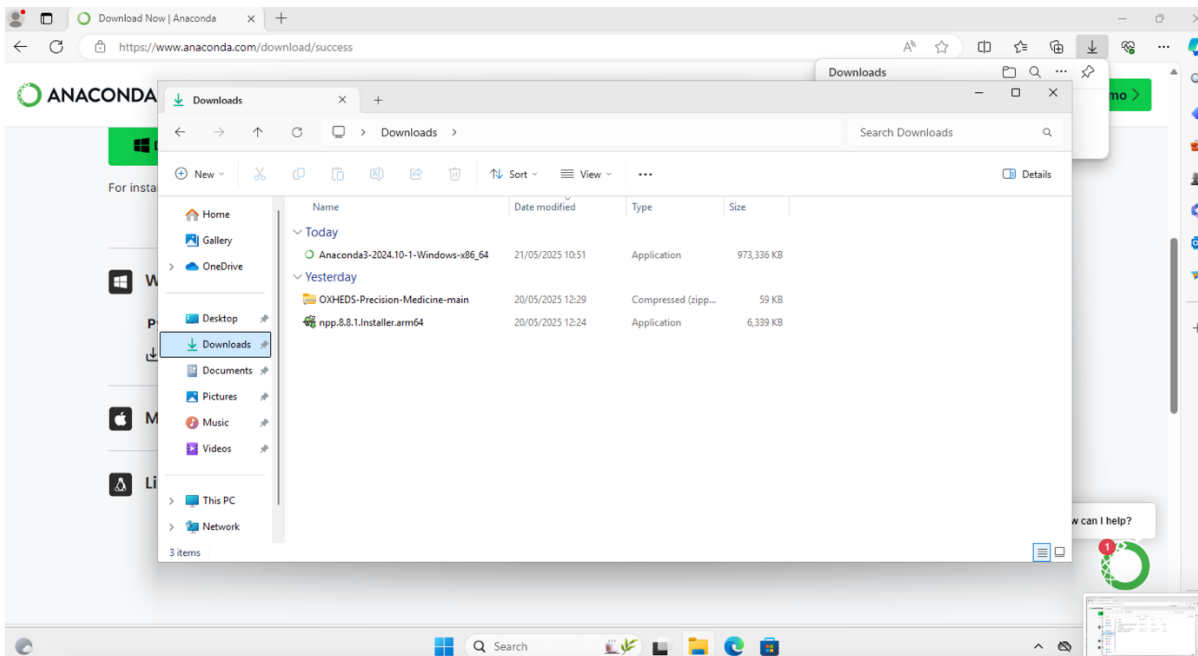


2. Make sure to click the correct option for your computer (e.g. Windows, Mac etc)



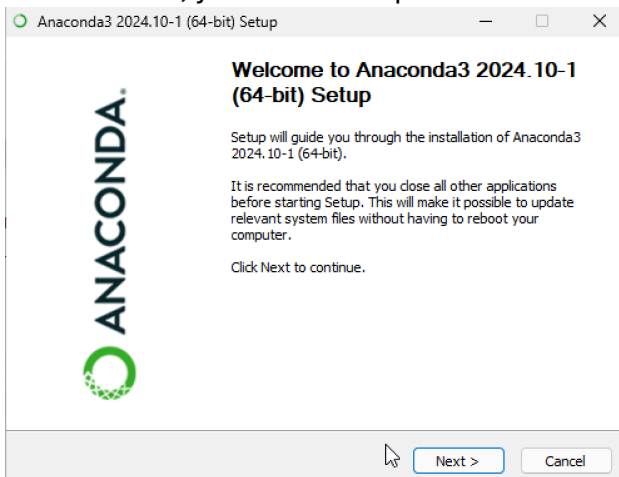
Here, I am picking the option for a Windows computer with 64-bit architecture. You can find the architecture of your computer in the device settings (instructions: <https://tinyurl.com/bdfjt5wm> for Windows, or <https://shorturl.at/cFD8M> for mac)

3. Once you download the Anaconda distribution, you will be able to find it in your “Downloads” folder. Double click on this installer.



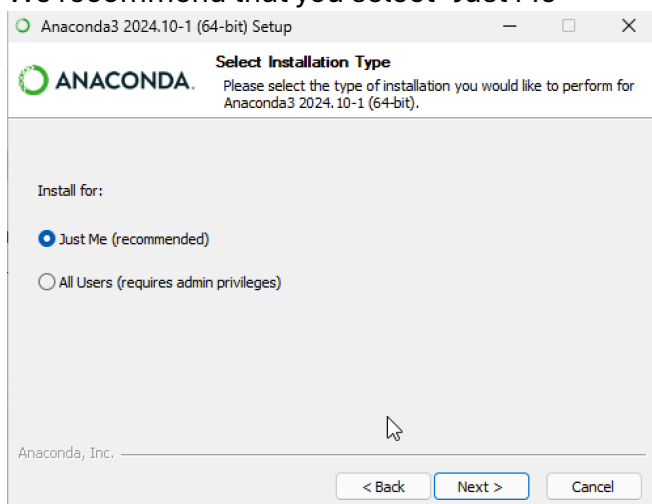
- This will open up a window that will talk you through the installation process. This will differ on Windows and Mac, so just follow the instructions that it gives.

On Windows, you should be presented with a menu like this

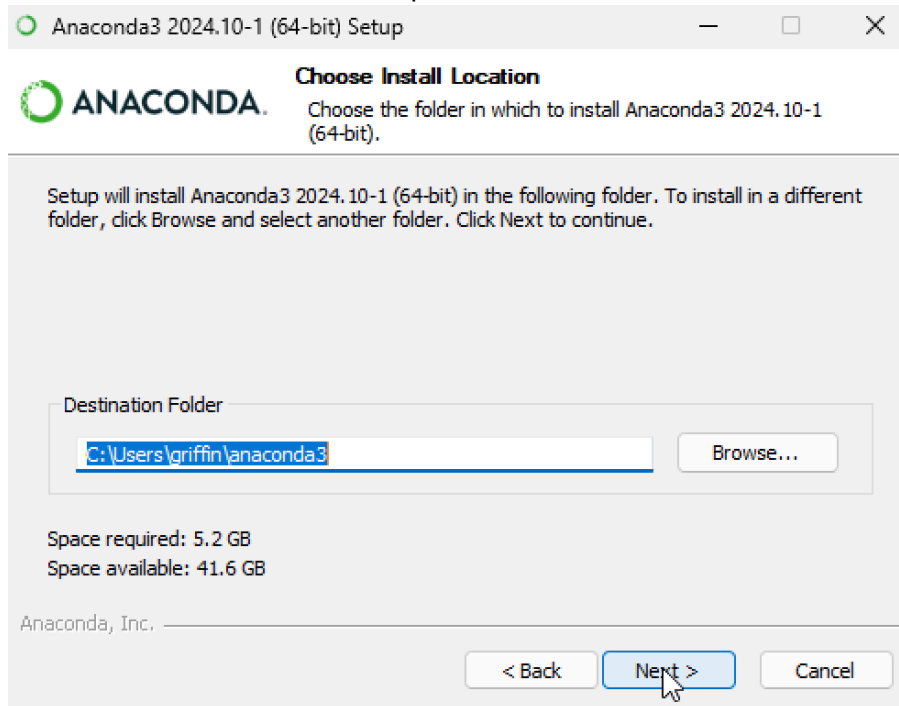


Click "Next", accept the "License Agreement" that comes up next

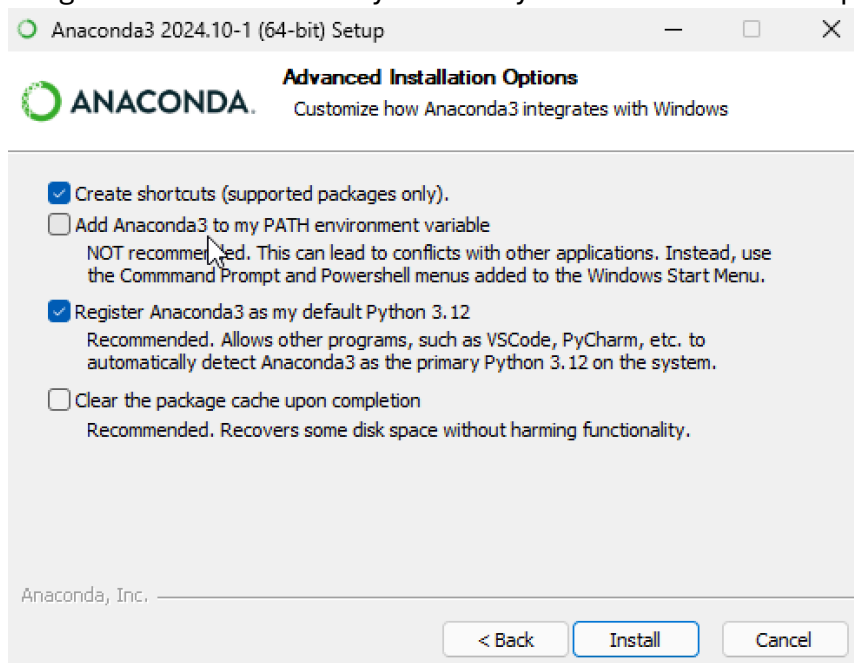
- You will be asked to "Select Installation Type"
We recommend that you select "Just Me"



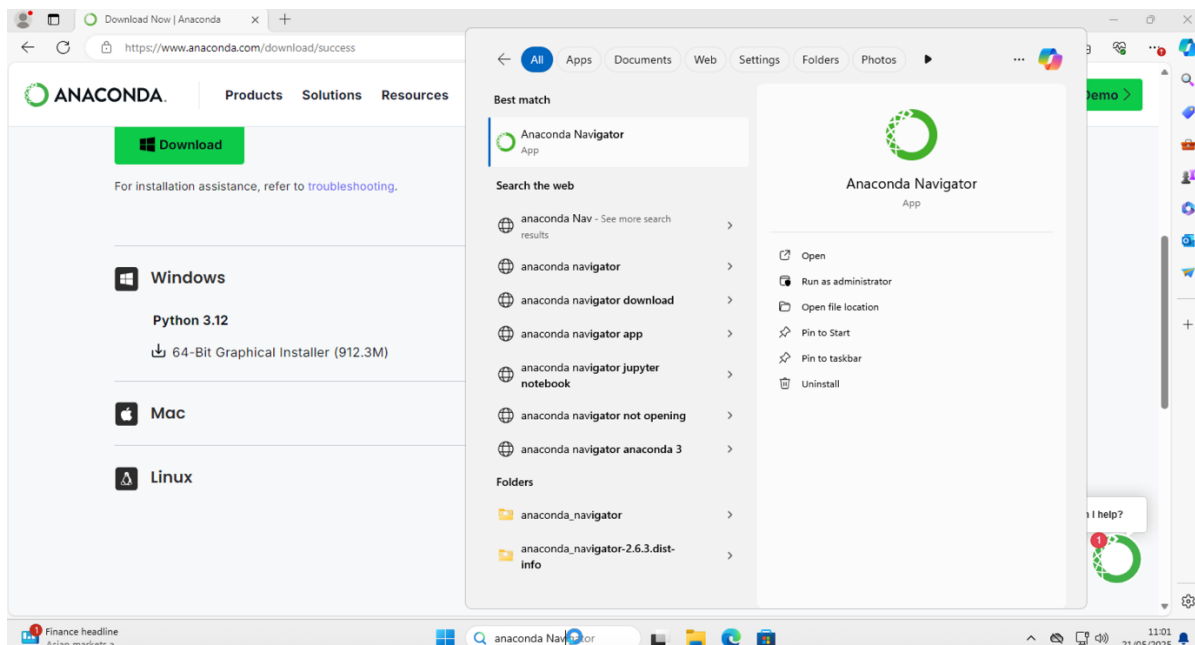
6. When asked to choose an install location, just select the location that is already suggested. You can move this to wherever you want, but there is no reason to not take the default. The installation should take up about 5 GB.



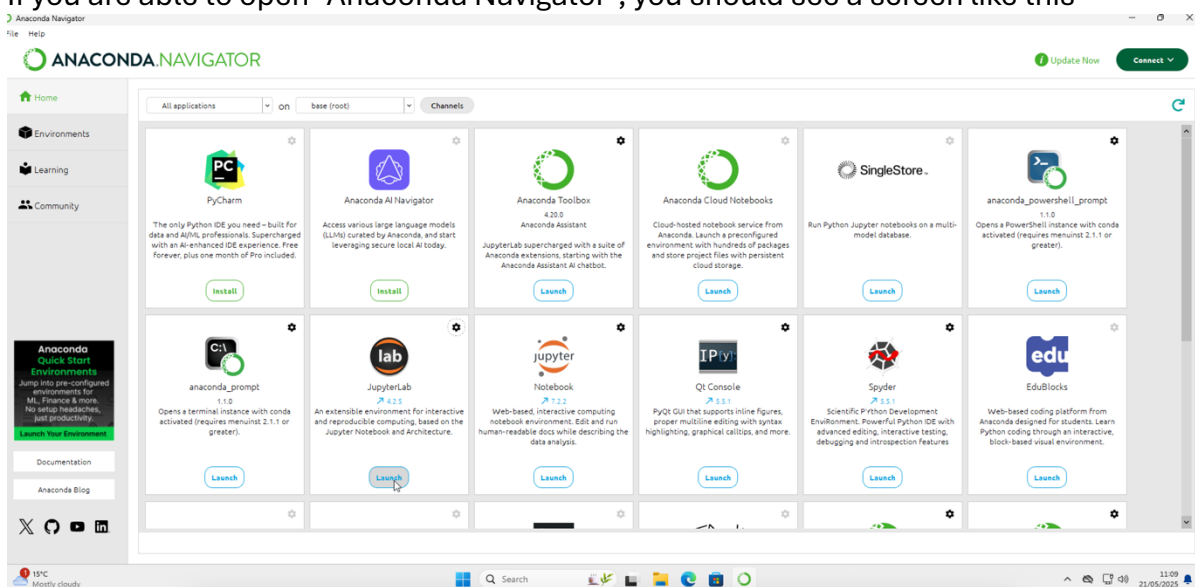
7. Under “Advanced Installation Options”, we suggest that you click “Create Shortcuts” and “Register Anaconda3 as my default Python 3.12”. The other options are not needed



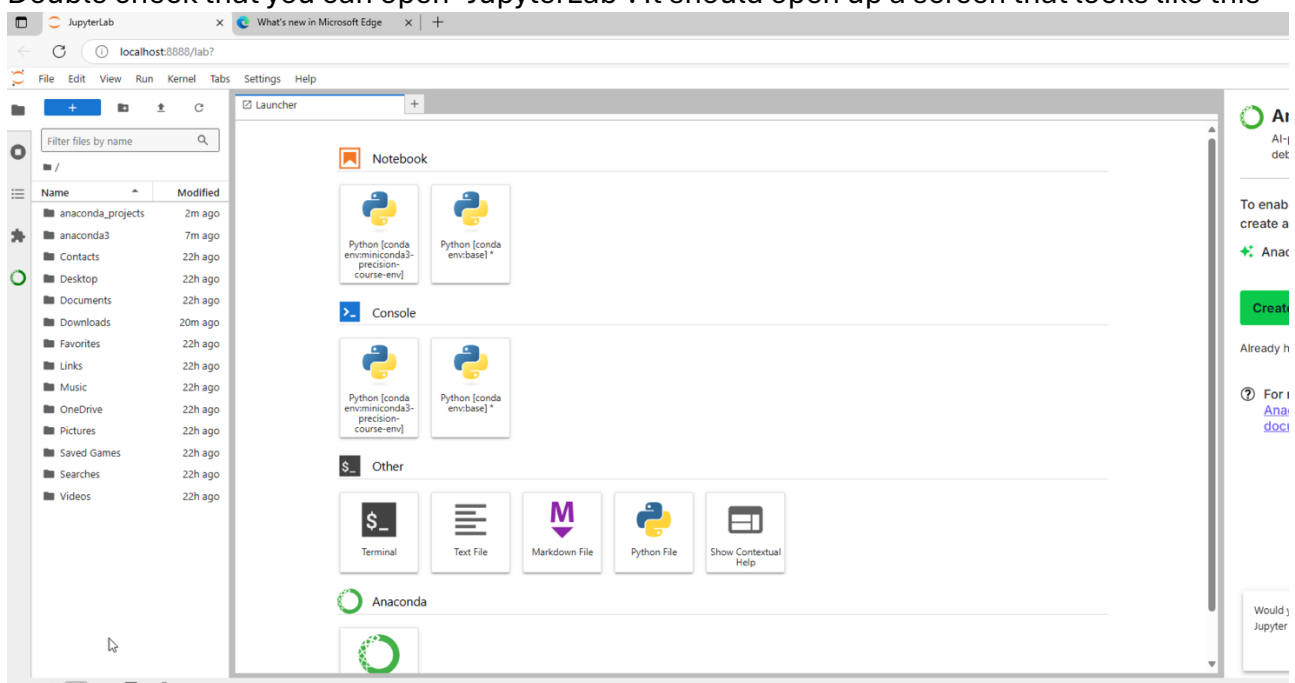
8. Once the installation is complete, you should be able to load the “Anaconda Navigator”
On Mac, go to the “Launchpad” and look for “Anaconda Navigator”
On Windows, search for “Anaconda Navigator” in the search bar.
Open up the application. It can take up to a minute to load, so be patient



9. If you are able to open “Anaconda Navigator”, you should see a screen like this



10. Double check that you can open “JupyterLab”. It should open up a screen that looks like this



If you have gotten to this step, your set up should be complete, and ready for our course!

If you have any questions, please contact us at
Griffin.Farrow@nds.ox.ac.uk