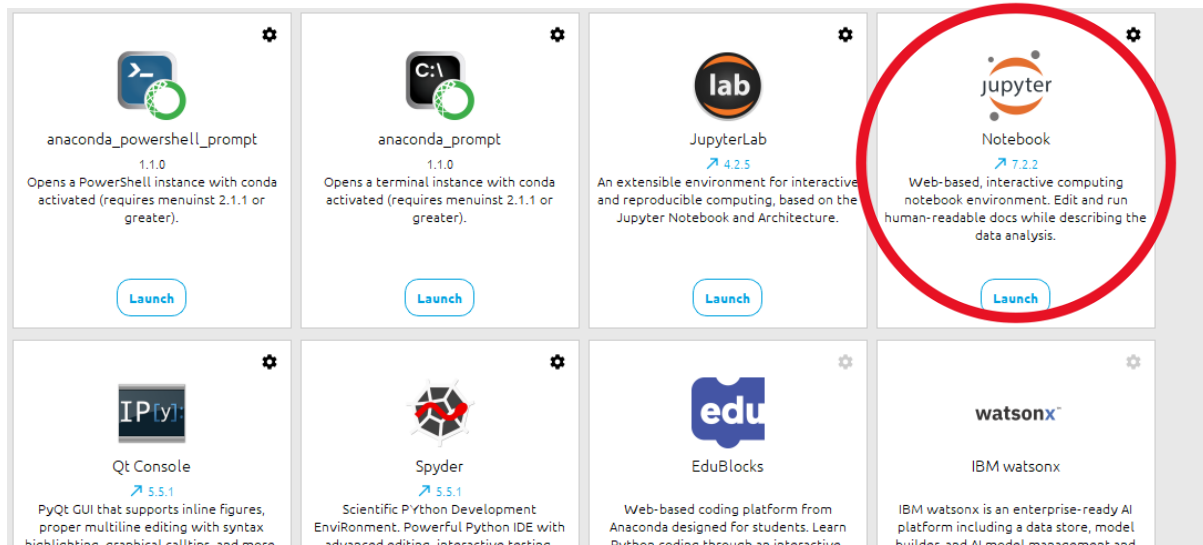


## EXERCISE #1: DATA VISUALISATION IN PYTHON

**Task:** For this exercise, we shall create graphs for gridded monthly rainfall for Kajiado county for the period 1980-2016.

After the successful installation of Anaconda, you are now ready to run your first exercise using Jupyter Notebook that provides an environment to run python codes. To begin using Anaconda, search for Anaconda Navigator from the Start Menu in Windows or from the Applications folder if you're on a Mac. You should be able to see Jupyter Notebook among the applications listed in Anaconda Navigator.



### What is a Jupyter Notebook?

A Jupyter Notebook is an open-source web application that allows users to share documents with text, live Python code, images and more. Using Notebooks is now a major part of the data science workflow at companies and organisations across the globe.

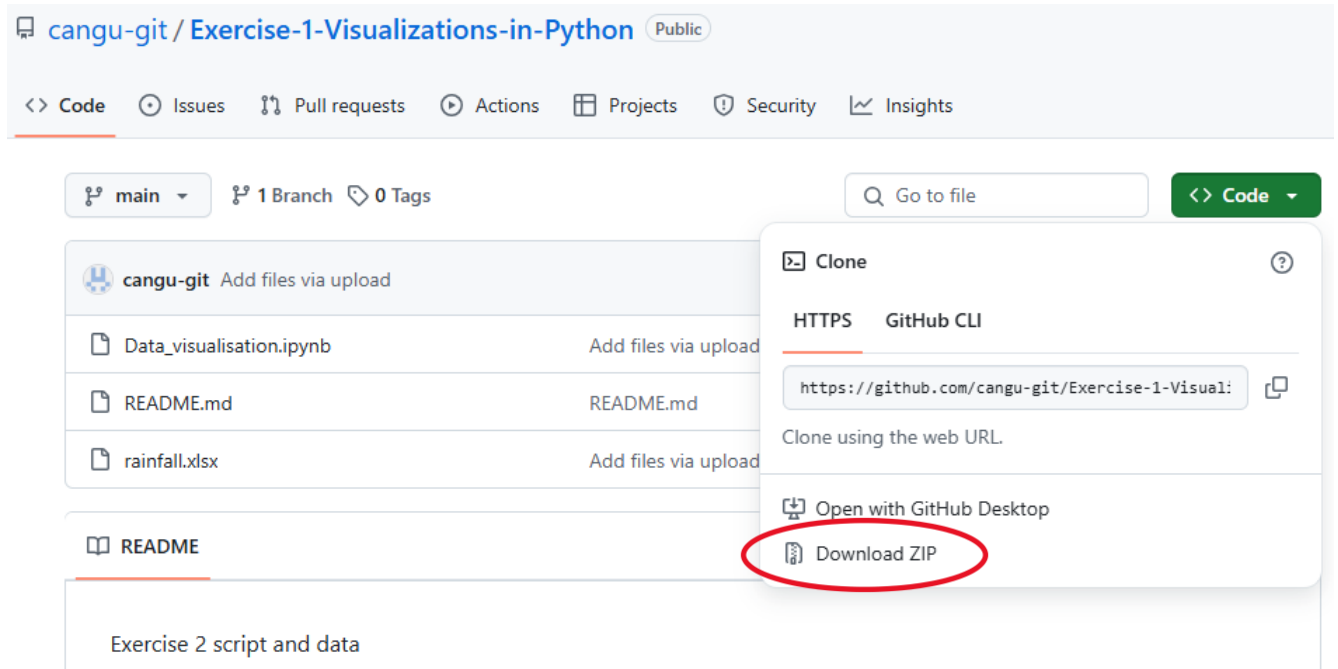
### Why use Python?

Python is an interpreted, object-oriented, high-level programming language that can be used for many tasks (e.g. data processing, visualisation, GIS, web design amongst others). Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very easy for beginners to start learning how to run and write Python code. Python's easy to learn syntax emphasizes readability and reduces the cost of program maintenance. The Python interpreter and the extensive standard library are available open source for all major platforms. Python comes installed on Macs and with Linux, but you'll need to install it yourself if you're using Windows (this was done during Anaconda installation).

For this exercise, you will be able to produce some visualisations by following the steps below.

## STEP 1

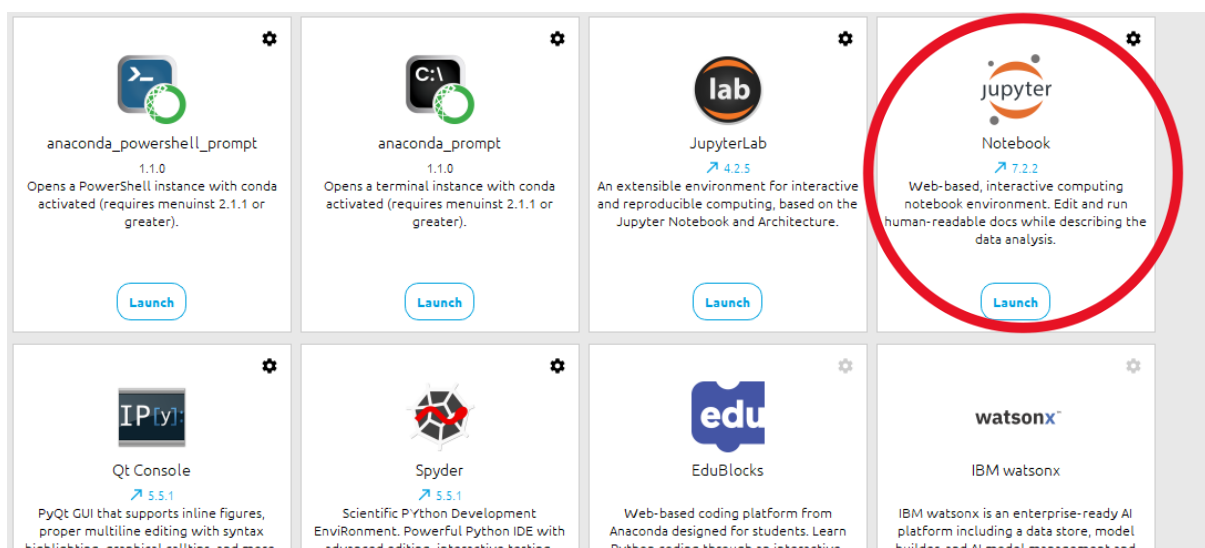
- 1.1. Download the 'Exercise-1-Visualizations-in-Python' from the GitHub repository <https://github.com/cangu-git/Exercise-1-Visualizations-in-Python>. To do this, click on 'Code' (green button) and then select 'Download ZIP'.



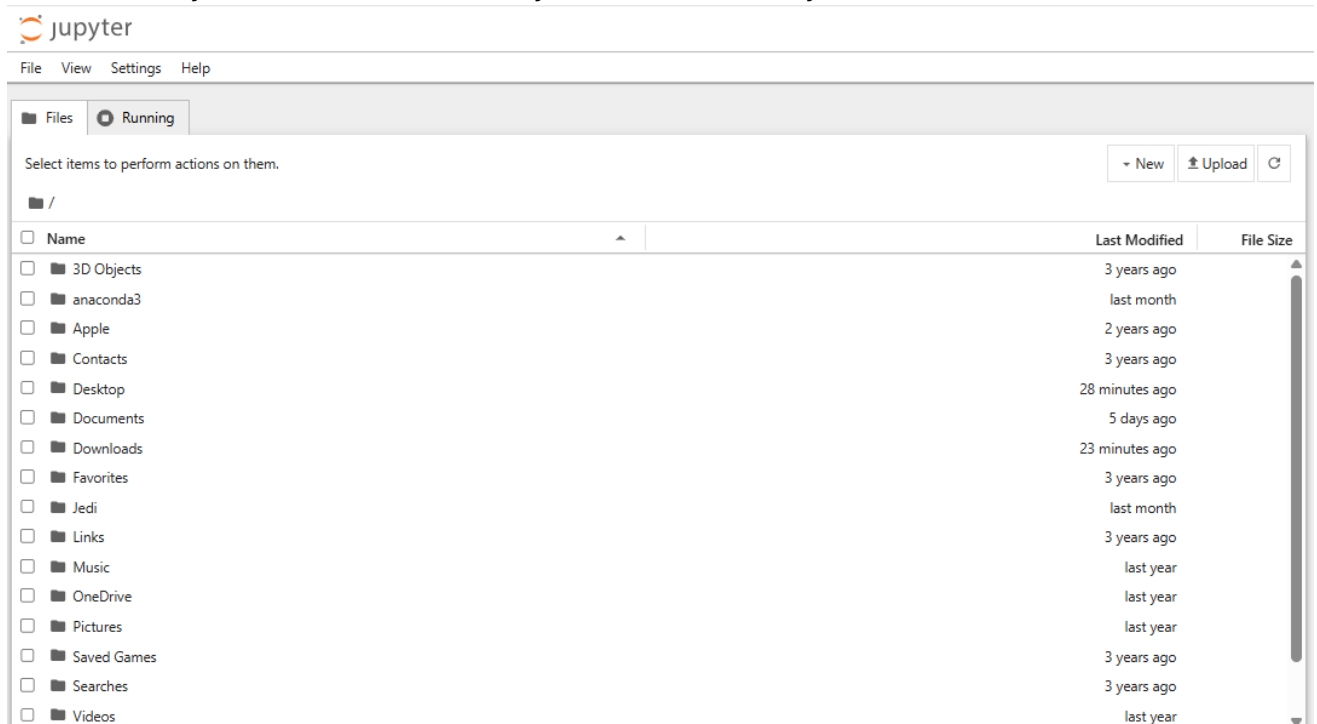
- 1.2. Unzip it and save the folder to a suitable location on your PC (preferably on your desktop). You should make a note of this directory. Do not move, edit or delete any of the files within the zipped folder.

## STEP 2

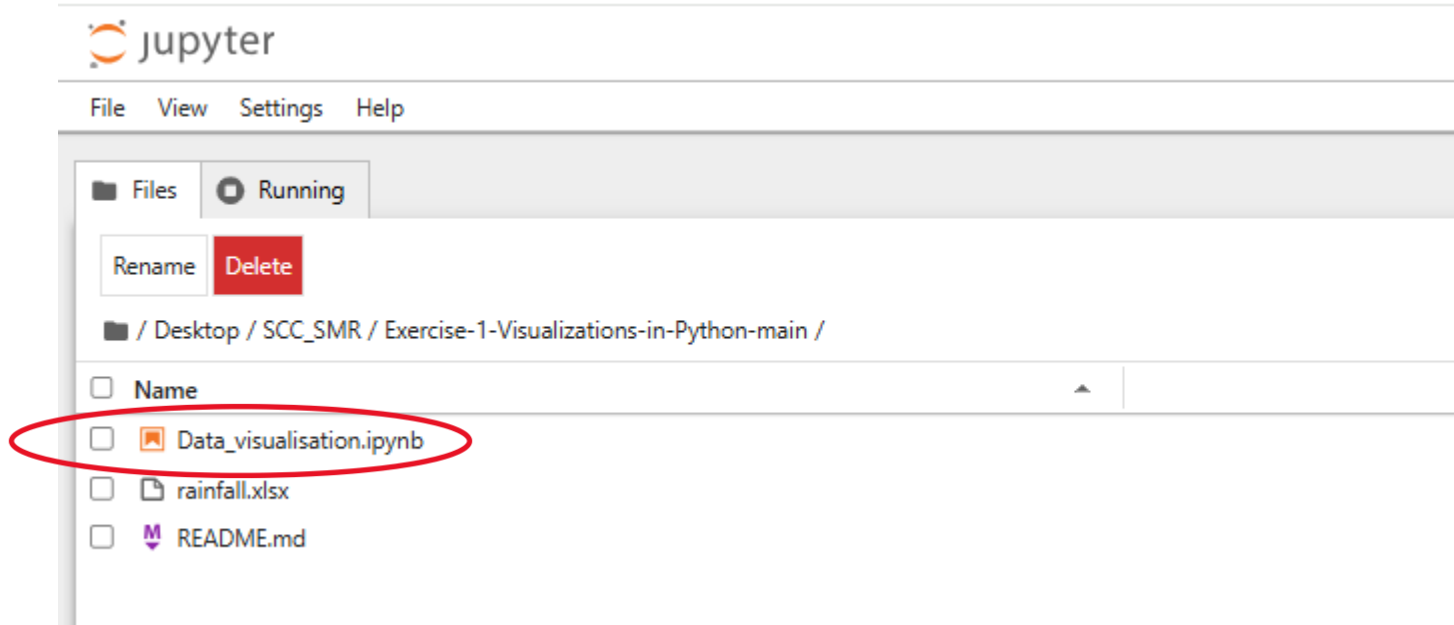
- 2.1. Open the Anaconda Navigator and launch the Jupyter Notebook.



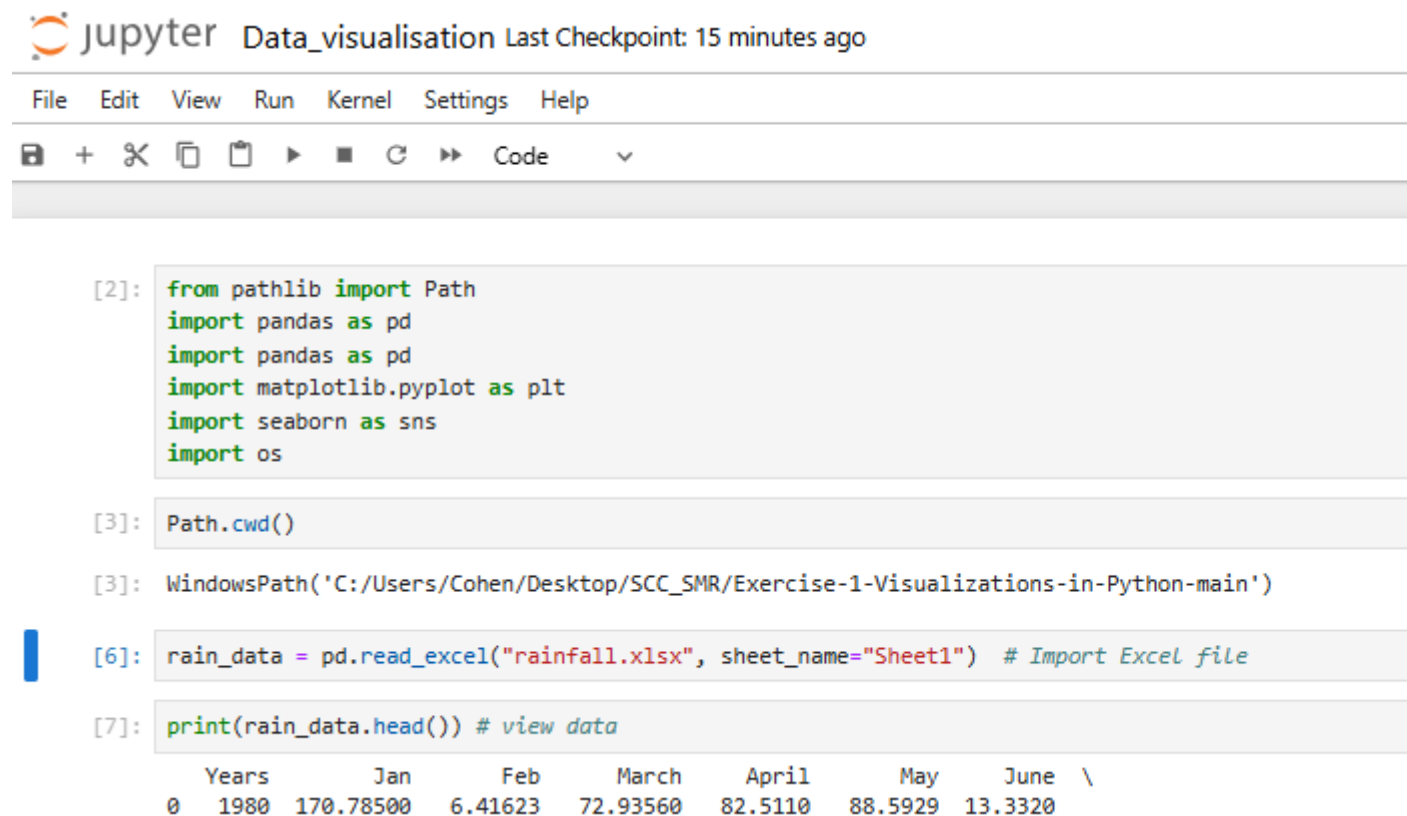
- 2.2. This will launch a web browser which will display the Jupyter Notebook interface where you can see all the files in your current directory.



- 2.3. Now navigate to the location where the 'Exercise-1-Visualizations-in-Python-main' folder is saved. Here, you will see the Data\_visualisation Jupyter Notebook (Data\_visualisation.ipynb) and the Rainfall Data (rainfall.xlsx).



- 2.4. Double click on “Data\_visualisation.ipynb” – this will open the Notebook as shown below:



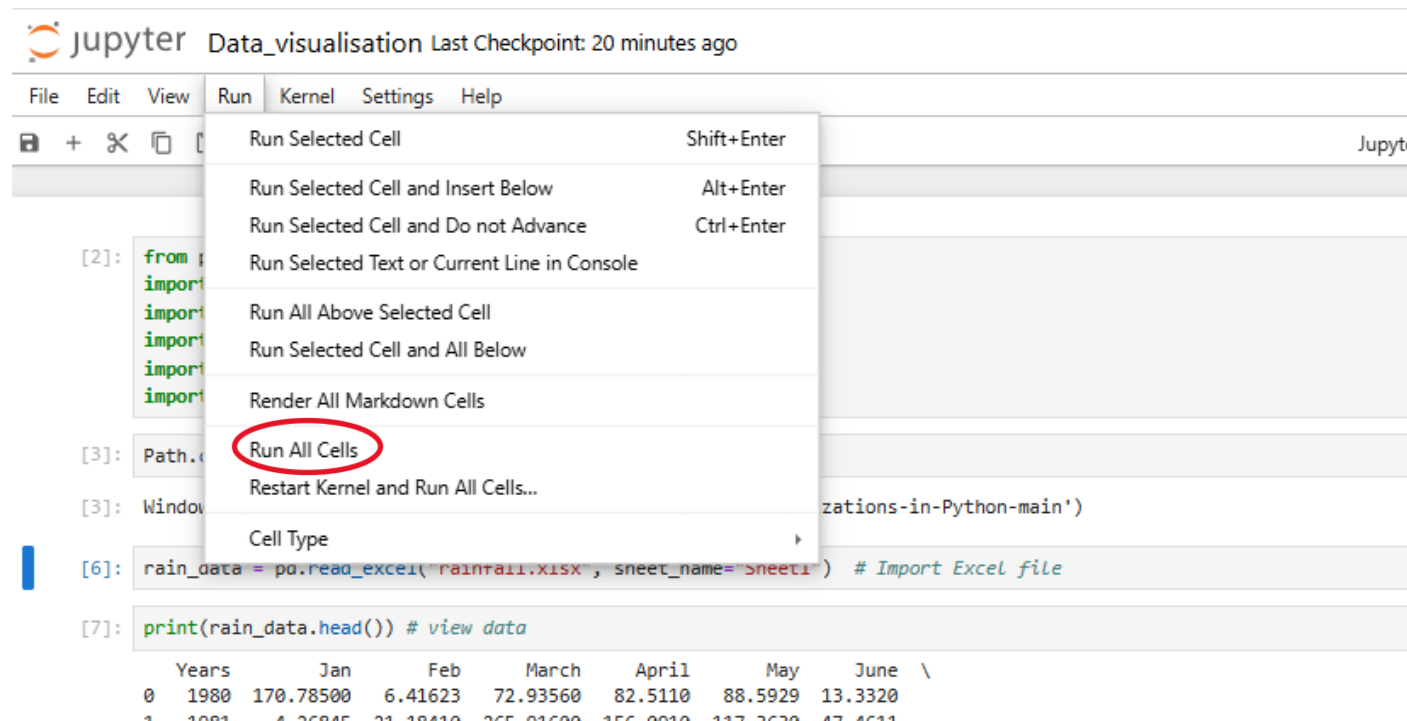
The image shows a Jupyter Notebook titled "Data\_visualisation" with a last checkpoint of 15 minutes ago. The interface includes a menu bar (File, Edit, View, Run, Kernel, Settings, Help) and a toolbar with icons for saving, adding, deleting, and running cells. The notebook contains several code cells:

- Cell [2]: Imports libraries: `from pathlib import Path`, `import pandas as pd`, `import matplotlib.pyplot as plt`, `import seaborn as sns`, and `import os`.
- Cell [3]: `Path.cwd()`
- Cell [3]: `WindowsPath('C:/Users/Cohen/Desktop/SCC_SMR/Exercise-1-Visualizations-in-Python-main')`
- Cell [6]: `rain_data = pd.read_excel("rainfall.xlsx", sheet_name="Sheet1")` # Import Excel file
- Cell [7]: `print(rain_data.head())` # view data

The output of cell [7] is a table showing the first five rows of data:

	Years	Jan	Feb	March	April	May	June	\
0	1980	170.78500	6.41623	72.93560	82.5110	88.5929	13.3320	

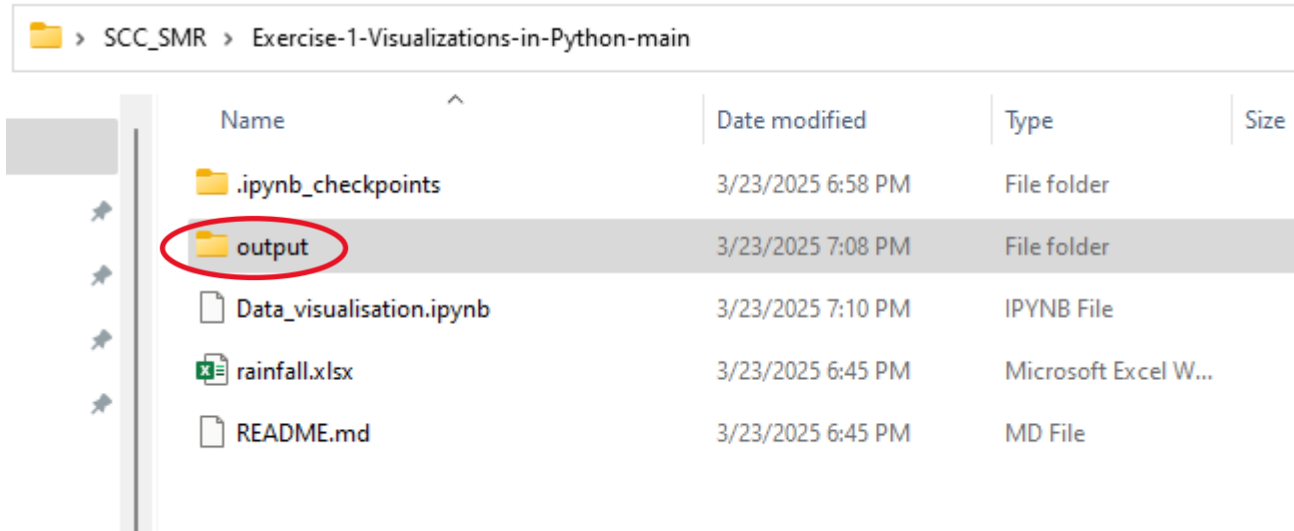
- 2.5. To run the script, navigate to ‘Run’ then select “Run All Cells”



The image shows the same Jupyter Notebook interface, but with the "Run" menu open. The "Run All Cells" option is highlighted with a red circle. The menu also includes options like "Run Selected Cell", "Run Selected Cell and Insert Below", "Run Selected Cell and Do not Advance", "Run Selected Text or Current Line in Console", "Run All Above Selected Cell", "Run Selected Cell and All Below", "Render All Markdown Cells", "Restart Kernel and Run All Cells...", and "Cell Type".

The code cells and their outputs are the same as in the previous image.

- 2.6. After the script has run successfully, go to the folder containing the script and the rainfall data on your desktop (or the location you saved it). You should see a new folder (output) created within this folder. Within it, there is a figure “January\_rainfall.png”.



The screenshot shows a Windows File Explorer window with the address bar displaying the path: > SCC\_SMR > Exercise-1-Visualizations-in-Python-main. The main area shows a list of files and folders. The 'output' folder is highlighted with a red circle. The table below represents the data shown in the screenshot.

Name	Date modified	Type	Size
.ipynb_checkpoints	3/23/2025 6:58 PM	File folder	
output	3/23/2025 7:08 PM	File folder	
Data_visualisation.ipynb	3/23/2025 7:10 PM	IPYNB File	
rainfall.xlsx	3/23/2025 6:45 PM	Microsoft Excel W...	
README.md	3/23/2025 6:45 PM	MD File	

### **STEP 3: Reproducing The Test**

Adapt the script and generate gridded rainfall graphs for the months of February-December for Kajiado County.

*(Hint: You only need to change the y value in the script to the appropriate month.)*