# Uploading your data to CSCS using Jupyter notebook

o **Install python 3.7:**

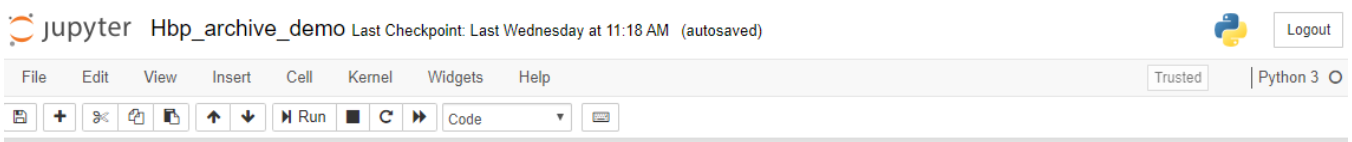https://www.anaconda.com/distribution/#download-section (select your appropriate operating system). Install for "Just me" to avoid regulatory issues with your institutional computer. Click next/continue until you see the installing bar (I.e. Approve all recommendations)

o **Starting a notebook:**

Download HBP_archive.ipynb to a folder. Find the anaconda prompt on your Windows/Mac system. (Windows users: Click your Windowsbutton. Type 'anaconda prompt' (also called 'anaconda powershell prompt') to open your command window.) In the prompt type `pip install hbp_archive`. When this operation is done, navigate to the folder where you stored your downloaded notebook. To do this type `cd path`, e.g. `cd C:\apps` and press enter. Now the path should be changed if it exists. Then type `jupyter notebook` and the notebook should open.

(If it still doesn't open, try copy-pasting the path you receive in your prompt window onto a browser path)

o **Running a notebook:**



o **"Running" a notebook effectively means clicking the first "cell" and pressing shift+enter.** If an asterix **(*)** appears on the side of the cell it means the operation is in progress, if a number **(1)** appears it means it is done with its operation. In the cells of the notebook there is **green** text (with **#** in the beginning of the line) telling you what to do and what information to put in.

o **Do not alter any text asides from where the arrows point.** If you alter or delete text by accident, download a new version of the Hbp_archive_demo.ipynb notebook.

o **If you have started an operation and wish to cancel it:** Go to the headers of the page and select the "Kernel" dropdown meny. Select "Restart & Clear output". This resets the whole notebook, and start again.

o **Enter your CSCS password when asked.** The notebook will stop if you do not enter your password, in this case you would need to restart to run cells again.

o   STEP-BY-STEP GUIDE FOR NON PYTHON USERS:

**Only change the piece of text that has an arrow on its side, if you remove other pieces of text the script will not work.**

1) **Run the top cell** (shift+enter).

```
In [ ]: from hbp_archive import Container, PublicContainer, Project, Archive
        import os
        from pathlib import Path, PurePosixPath

        ### Please see https://hbp-archive.readthedocs.io/en/latest/ for additional details
```

2) **Then specify the subproject you are working on and username before you run this cell** (shift+enter) **and enter your password in the slot that appears.**

```
In [2]: subproject = 'bp00sp05' #<---Put in your subprojectnumber
        usernumber = 'bp000222'#<-- Put in you usernumber


        print('Please enter your password:')
        SP_a = Project(subproject, usernumber)

        Please enter your password:
        Password: ........
```

3) **Next, specify the name of your container and path (apply double slashes) before you run the cell.**

```
user_container = 'test_IA' #<-- Specify the name of your container
container = SP_a.get_container(user_container)

## If you do not know the name of your container:
##SP_a.containers


path = 'C:\\apps\\testfolder2\\' #<-- Specify the folder you wish to upload (For windows-users apply double slashes).
                                 #E.g.  'C:\\apps\\testfolder2\\'

# r=root, d=directories,f = files
for r, d, f in os.walk(path):
    files = []
    if r == path:
        for file in f:
            files.append(os.path.join(r, file))
        container.upload(files, overwrite = True)
    else:
        p2 = Path(r)
        subdirname = p2.relative_to(path)
        subdirmultiple = subdirname.parts
        adjustpath = str(PurePosixPath(*subdirmultiple))
        adjustpath += '/'
        for file in f:
            files.append(os.path.join(r, file))
        container.upload(files, remote_directory=adjustpath, overwrite=True)
```

**4) If you made a mistake and want to reload your files you can delete all files in the container by running this last cell and redo step 3.**
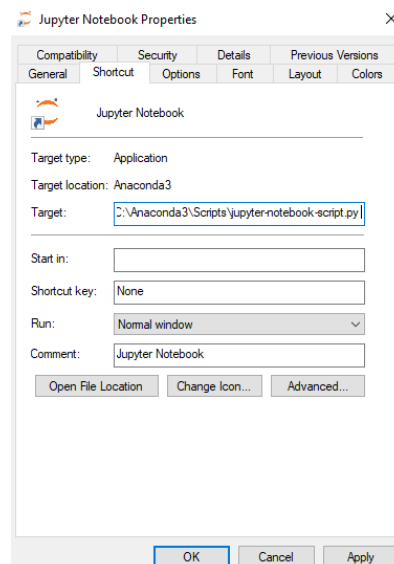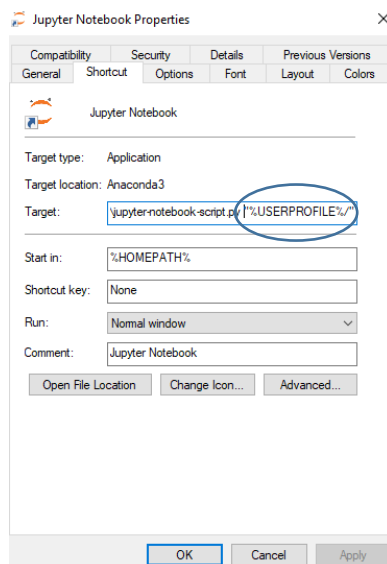
```
In [6]: ### To delete all files perform:

        all_files = container.list()
        for file in all_files:
            filename = str(file)
            container.delete(filename.strip('\'"'))
            print('Deleting {}'.format(file))
```

- **Jupyter notebook shortcut option for Windows users:** *(*This makes it possible to start the jupyter notebook in an easy manner (double-clicking the icon), without having to write in command promp).

**Note:** Install python on your c:\ drive for the shortcut steps to work.

1) Press your windows button and search for Jupyter notebook
2) Right-click on your notebook and select Open file location
3) When in the file location folder, copy the jupyter notebook app to your desktop. Now you have a copy you can modify.
4) Right-click the desktop jupyter notebook and select "Properties": Delete the contents of "Start in" (of the shortcut tab) and the text after jupyter-notebook-script.py in the "Target" field (see figures). Replace this last piece of text (often "%USERPROFILE%" with the path of the notebook you downloaded (Hbp_archive_demo.ipynb)