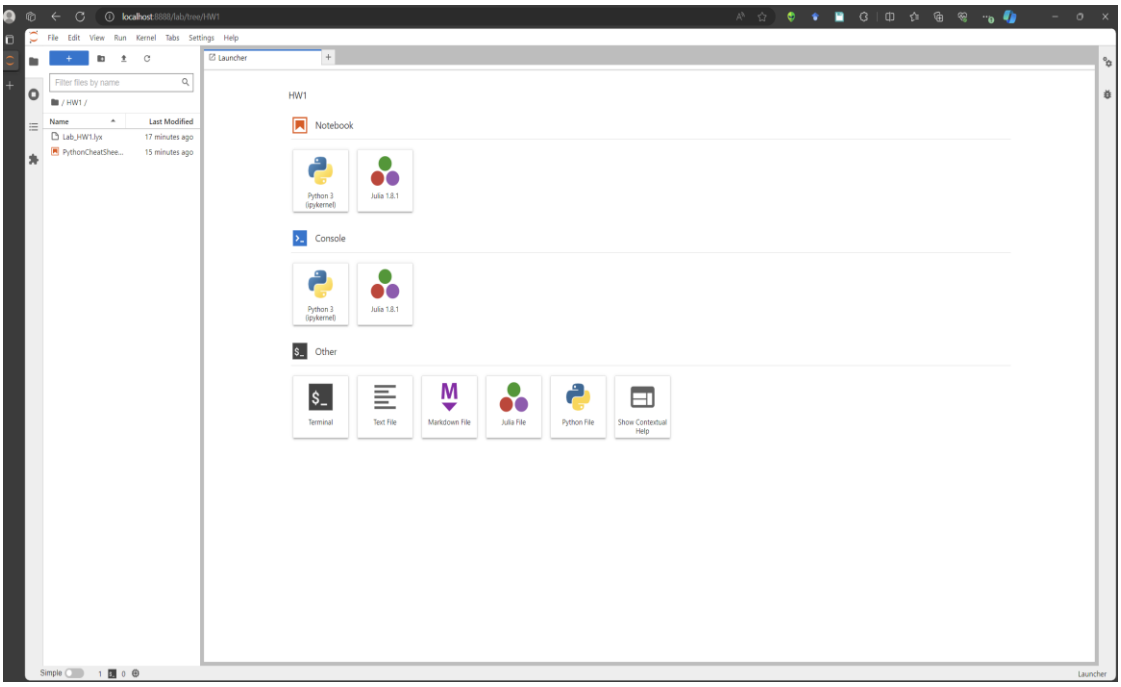
Lab Homework 1

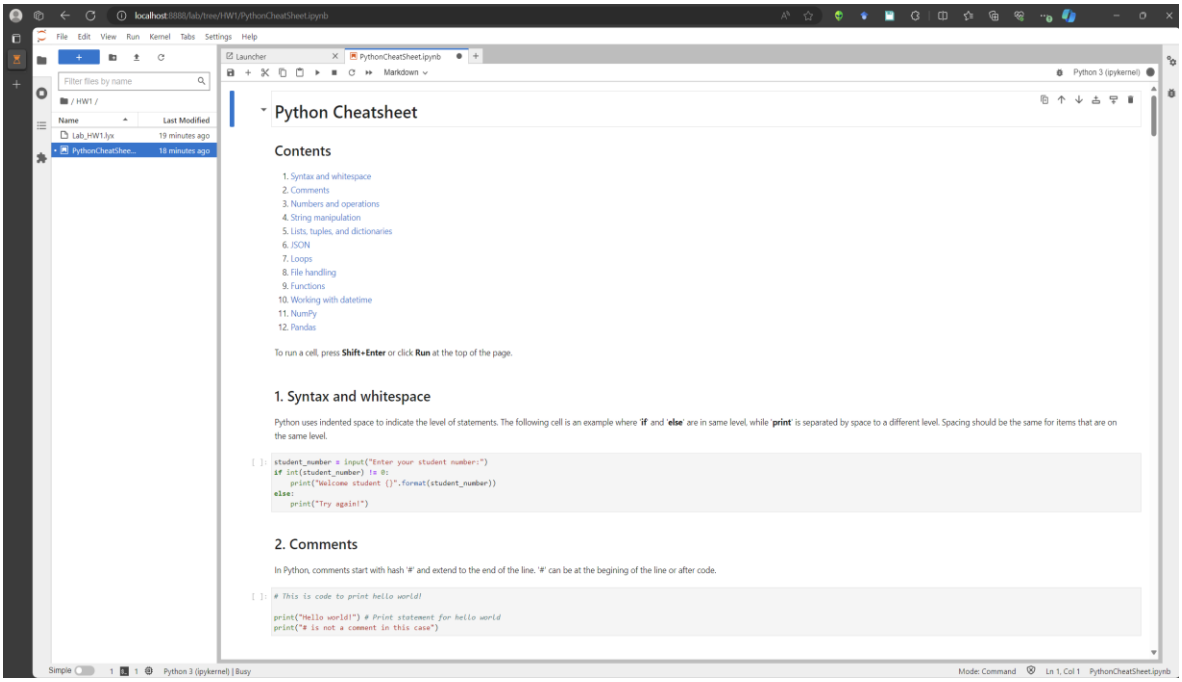
Submitted by : Manas Vishal (01971464)

Homework on introduction to jupyter notebook and importing datasets

Task 1 : Introduction to jupyter lab/notebooks

- We first start the jupyterlab session either via terminal or through a navigator like anaconda or jupyter. I personally prefer terminal paired with miniconda for low memory usage.
- We then upload the test notebook provided in the class using “Upload” button

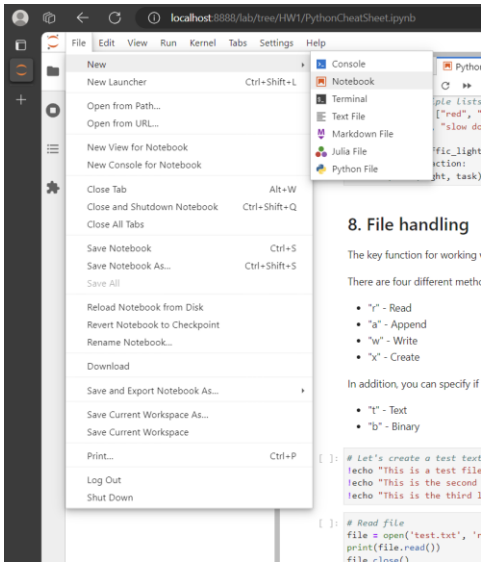




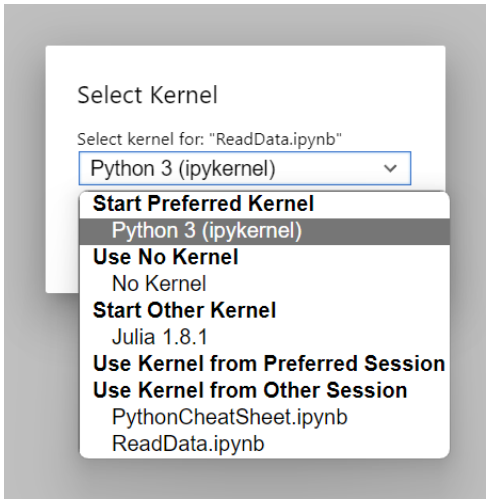
- We go through the notebook and execute the blocks of code by using Shift + Enter and later with the Run all cells function
- It is noted that Jupyter notebook also allows markdown for ease of presentation

Task 2 : Downloading data

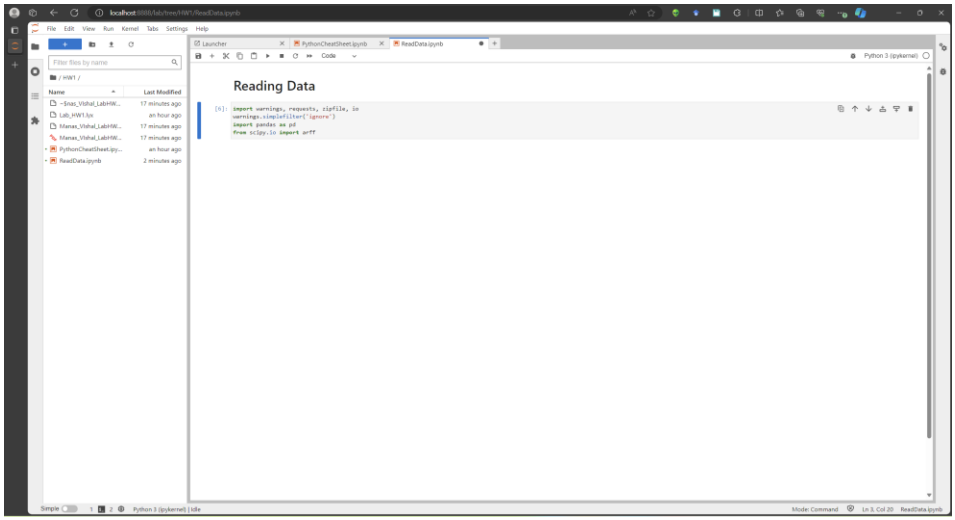
- A new notebook is initiated to download and read data for this task



- The kernel is chosen as Python 3 kernel



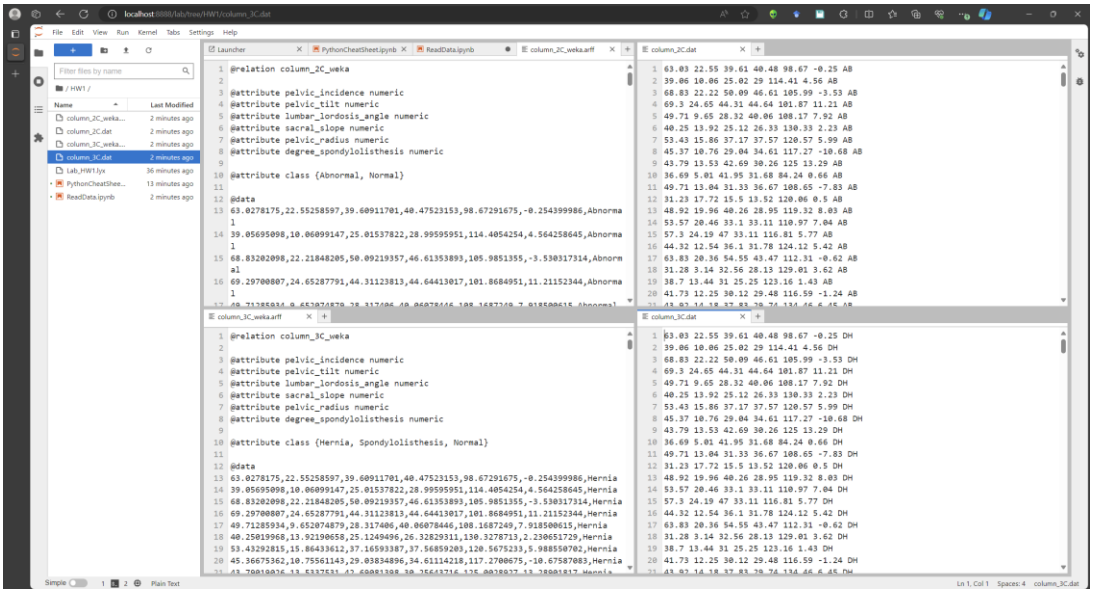
- We import the modules that are used to suppress the warnings and helps in file i/o.



- We now download the data from web using the following commands

```
[6]: f_zip = 'http://archive.ics.uci.edu/ml/machine-learning-databases/00212/vertebral_column_data.zip'
r=requests.get(f_zip, stream=True)
Vertebral_zip= zipfile.ZipFile(io.BytesIO(r.content))
Vertebral_zip.extractall()
```

- We examine the four data files downloaded. **Not all the datasets have column headings.**



- We now examine the data using pandas for a given data file

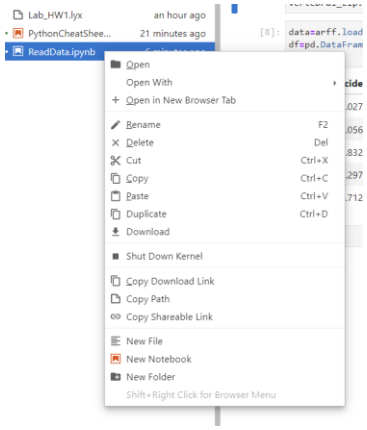
```
[8]: data=arff.loadarff('column_2c_weka.arff')
     df=pd.DataFrame(data[0])
     df.head()
```

[8]:

	pelvic_incidence	pelvic_tilt	lumbar_lordosis_angle	sacral_slope	pelvic_radius	degree_spondylolisthesis	class
0	63.027817	22.552586	39.609117	40.475232	98.672917	-0.254400	b'Abnormal'
1	39.056951	10.060991	25.015378	28.995960	114.405425	4.564259	b'Abnormal'
2	68.832021	22.218482	50.092194	46.613539	105.985135	-3.530317	b'Abnormal'
3	69.297008	24.652878	44.311238	44.644130	101.868495	11.211523	b'Abnormal'
4	49.712859	9.652075	28.317406	40.060784	108.168725	7.918501	b'Abnormal'

Task 3 : Downloading a notebook

- Right clicking on any file on the left pane of the browser gives us the context menu. From there, we can perform multiple functions



- We use the download function and then specify the path to download the file

Conclusion:

This lab homework taught us :

- How to use Anaconda and Spyder for jupyter notebook/lab
- How to create a new file and upload an existing file to the given jupyter session
- How to run python codes inside a jupyter notebook
- How to add markdown code inside a jupyter notebook
- How to download data and read it using different modules of python
- How to save any notebook to our local storage

In conclusion, after this lab I can work with online datasets in jupyter notebooks.