

Data Analytics Laboratory instructions

View and interact with Jupyter notebooks and learn about data mining techniques.

Prior Knowledge

Data Analytics lectures.

Software Requirements

The notebooks are available in the course GitHub repository:

<https://github.com/UppsalaIM/2IS239>

Binder is a hosted service that takes care of all the installation for you, and runs the notebooks on the cloud. For running on the cloud using Binder you will need only:

- A Web browser: Google Chrome or Mozilla Firefox.

For running on your computer you will need to install:

- Python 3.7 or newer
- Jupyter
- Extra Python packages: pandas, textmining3, wordcloud, matplotlib, mlxtend, graphviz, scikit-learn, okpy
- Graphviz binary installation.

Full installation instructions can be found in repository documentation (visit the URL given above).

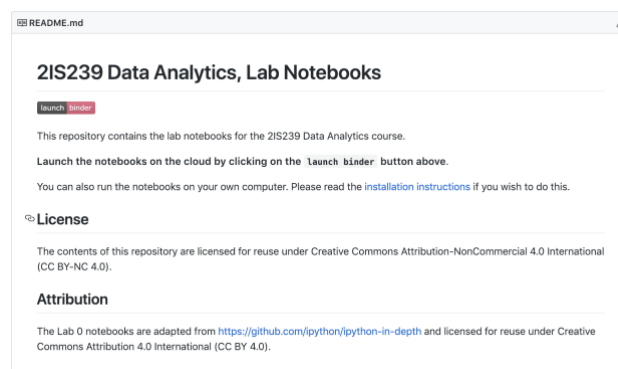
Starting the Labs

The laboratory sessions for this course are run in Jupyter.

Jupyter is an environment that is web based, and allows you to do interactive programming inside a Web browser. Jupyter allows you to view and create computational notebooks, which are like Web pages that contain cells that are static content and cells that are Python code that you can run and view the output in the browser.

1. Visit the GitHub source code repository for the Labs at <https://github.com/UppsalaIM/2IS239>

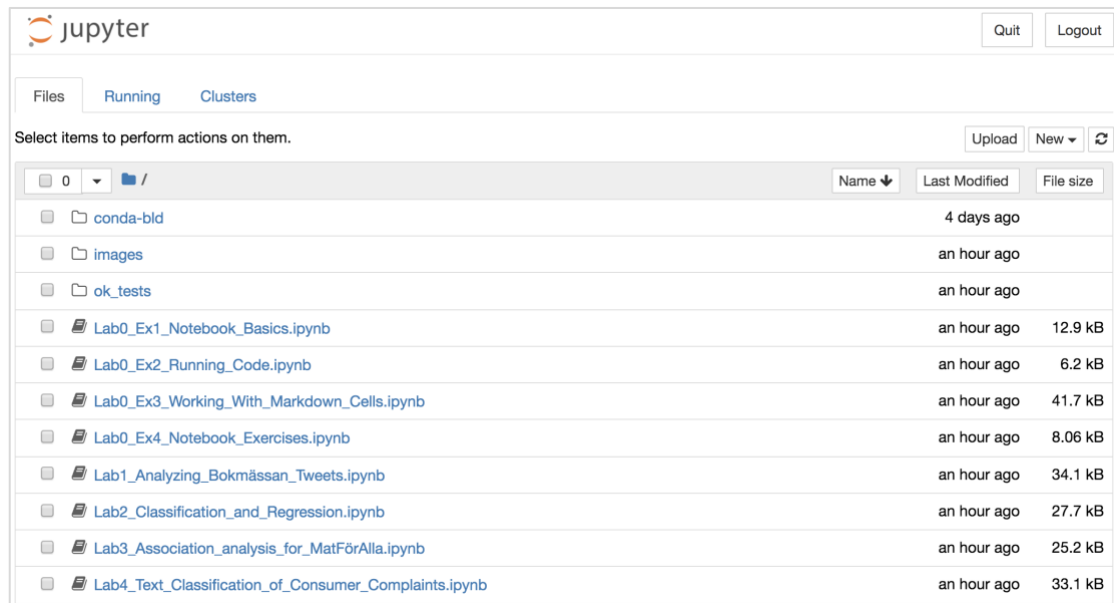
Scroll down and you should see some content that looks like this:



There are two ways to launch the notebooks repository. The simplest way is to use the **launch binder** button. If you installed Jupyter on to your own computer, launch Jupyter from the command-line with the command

```
jupyter notebook
```

or if you are using Anaconda launch Jupyter from the Anaconda Navigator app. Once launched, you should see the repository contents in the Jupyter dashboard like this:



To launch a specific notebook, click on any of the **.ipynb** files in the dashboard.

2. If you have not already done so, make sure to familiarize yourself with the Jupyter environment by doing the Lab 0 exercises 1-4

i.e. Open up, read and complete `Lab0_Ex1_Notebook_Basics.ipynb`,
`Lab0_Ex2...`, `Lab0_Ex3...` etc.
3. Open the Lab notebook corresponding to the session and complete the questions in the notebook.

e.g. During Lab 1, open `Lab1_Analyzing_Bökmässan_Tweets.ipynb`, etc.
4. When you have completed all the questions, remember to follow the instructions at the end of each notebook to download your completed lab as an HTML file and submit it in *Stadium*.

If you are stuck, do not hesitate to ask for help from the lecturer (David) or post in the discussion forum in Stadium!

Good luck!