Data Mining and Data Science Laboratory instructions

Read all these instructions before you get started.

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

The laboratory sessions for this course are run in Jupyter. It 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.

Prior Knowledge
Data Mining and Data Science lectures.

Software Requirements
The notebooks and related information are available in the course GitHub repository: https://github.com/UppsalaIM/2IS063

OBS! Due to the current COVID-19 situation meaning we must do these labs by an entirely digital/remote way of working, it is more difficult than usual for your teachers to provide technical support to run the labs.

We recommend you do the labs on the cloud (see below) as it requires no setting up if you already have a computer with a Web browser.

If you want to run the labs on your own machines or on the lab remote server, we cannot spend much time helping you debug your set up since each person's own computers will have different underlying configurations.

On the cloud using Binder
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 we will need only:

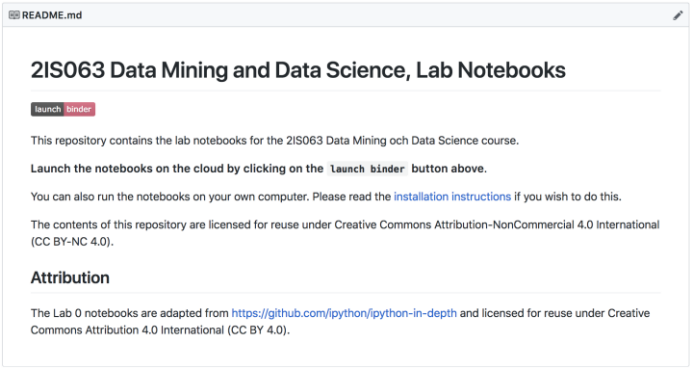
- A Web browser: Google Chrome or Mozilla Firefox.

On your own computer
Full installation instructions can be found in the `INSTALL.md` file contained in the GitHub repository. To do the labs you *do not need* to use your own computer.

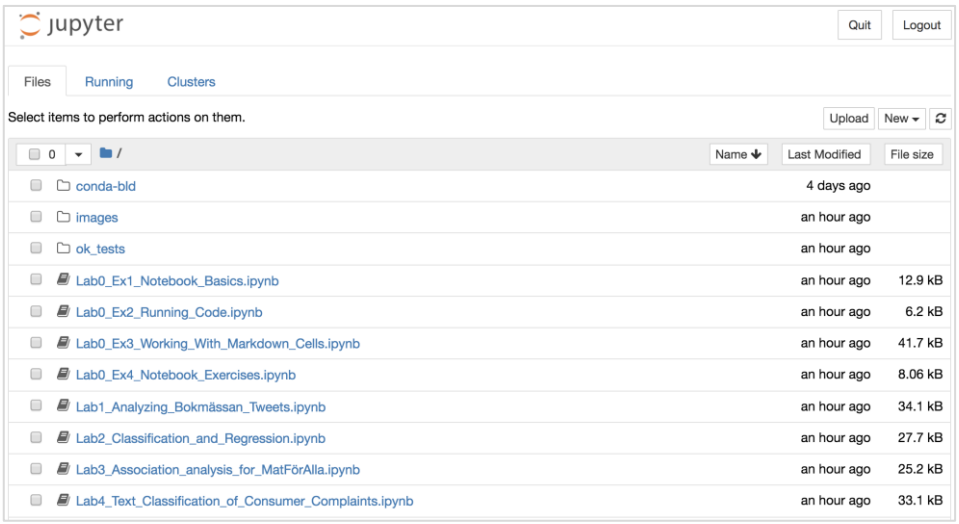
Getting started using Binder
An alternative way to run the notebooks is open the notebooks repository on the cloud using Binder. To do this:

1. Visit the GitHub source code repository for the Labs at https://github.com/UppsalaIM/2IS063

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



2. Click on the **`launch binder`** button and wait. Once launched, you should see the repository contents in the Jupyter dashboard.



Do the exercises for each lab notebook

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

1. If you have not already done so, make sure to familiarize yourself with the Jupyter environment by doing the Lab 0 exercises 1-3.

   i.e. Open up, read and complete `Lab0_Ex1_Notebook_Basics.ipynb`, `Lab0_Ex2...`, `Lab0_Ex3...` etc.

2. Once you are familiar with the Jupyter environment, open the Lab notebook corresponding to the session and complete the questions in the notebook.

   e.g. During Lab 1, open `Lab1_Analyzing_Bökmassan_Tweets.ipynb`, etc.

3. When you are finished, submit your answers/notes to Studium, as indicated in at the end of each notebook.

If you are stuck, do not hesitate to ask for help from one of the teachers! Good luck!