Bocconi

TEXT ANALYSIS WITH PYTHON

Lesson 1 – 6/8/2021

a.y. 2020-2021

Lesson content

Building a common ground

- Why NLP in today's world: its applications
- Preliminaries
- Introduction to Jupyter Notebook
- Brief recap of Python basics
- Pandas: the essentials

Attendance registration

To track your presence in class (wherever you are), please:

- either use the app on your smartphone or tablet
- or go to this web page → www.unibocconi.it/attendance

using

- your own yoU@B credentials
- today's six-digit code*

If you have problems with the app:

- try to log out and then log in again
- if the problem persists, notify our tutor via chat who will manually register your presence

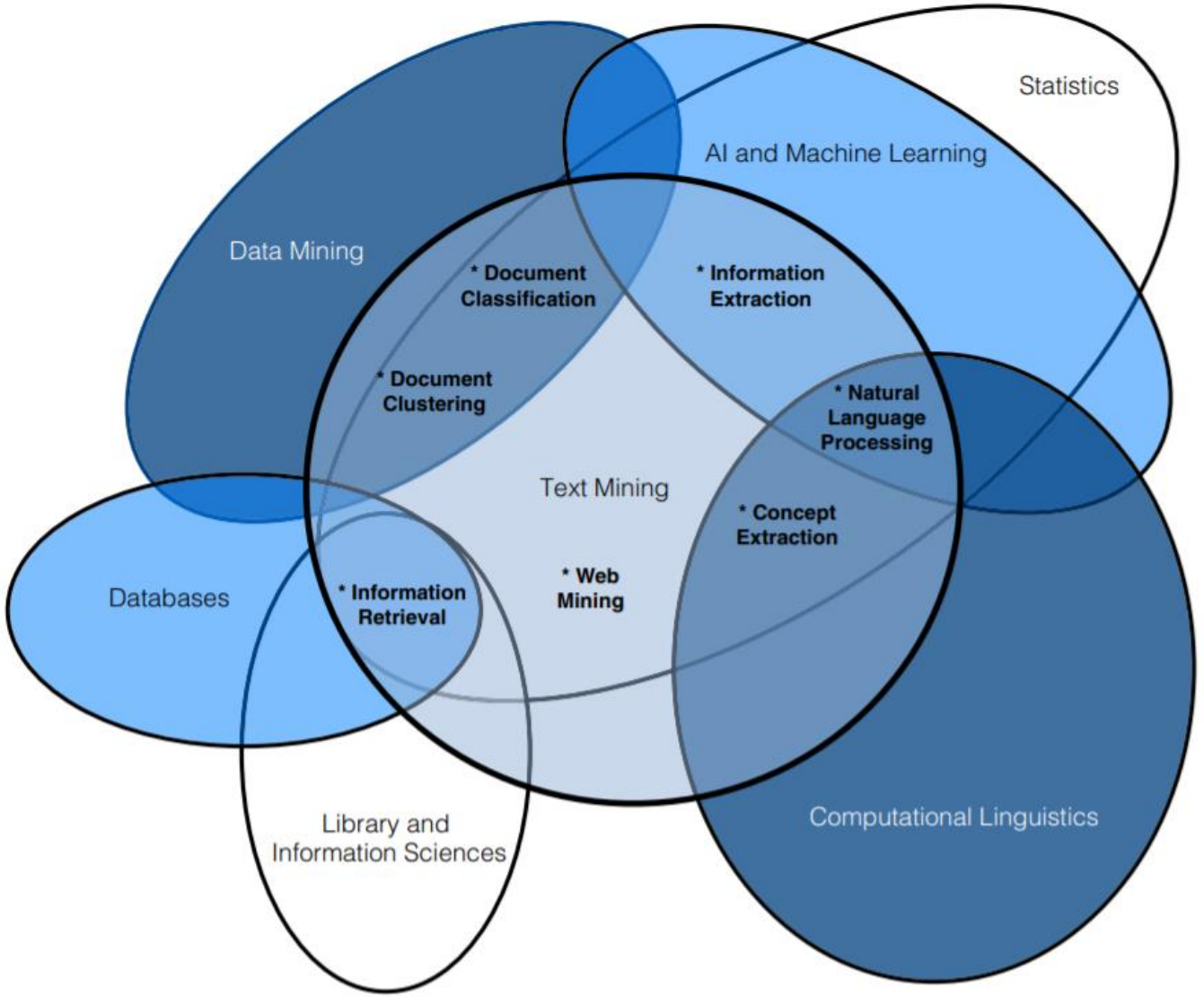
((*) It will be shown in the classroom and will remain active only for 10 minutes starting from the first of you who will register

WHY NLP TODAY?

The context

- More than 80% of the knowledge is made by not structured data = text
- Having such a wealth of material without knowing/studying it is equivalent to not having information available
- Information is of no value if it is not managed: big amount needs an automated approach
- Today 4 out of 5 companies still prefer to analyze (only) structured data

Text mining, its practice areas and the overlapping fields



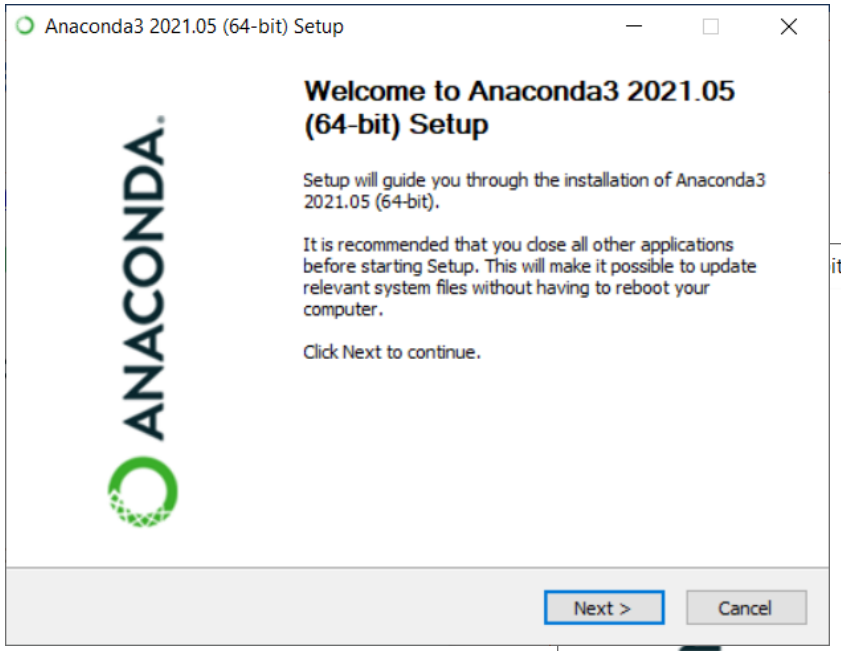
Practical Text Mining and Statistical Analysis for Non-Structured Text Data Applications
G. Miner, D. Delen, J. Elder, A. Fast, T. Hill, and R. Nisbet, Elsevier, January 2012

http://www.textanalyticsworld.com/wp-content/uploads/2012/03/PracticalTextMining_Excerpt.pdf

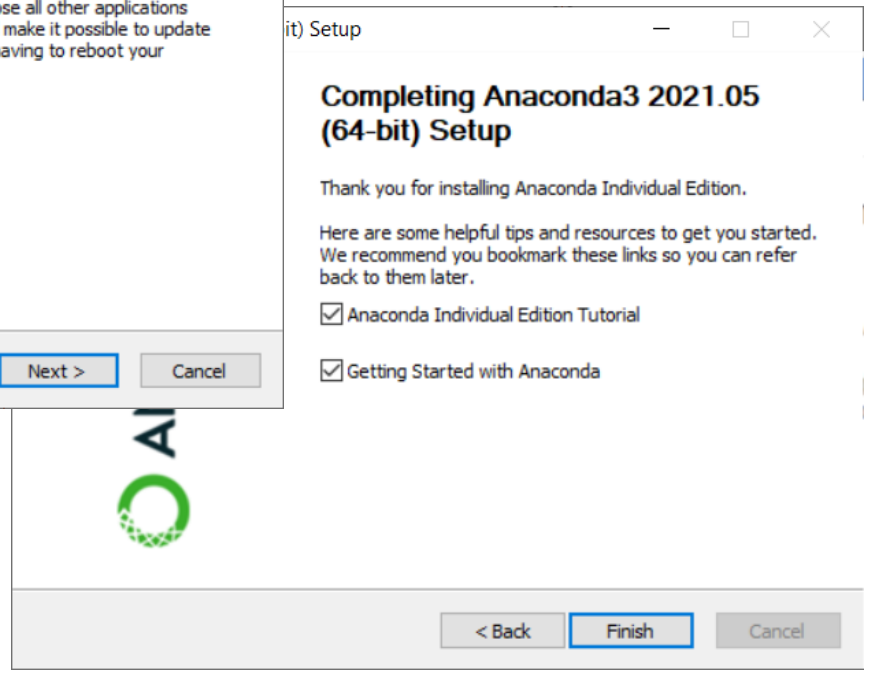
ANACONDA AND JUPYTER NOTEBOOKS

Install Anaconda

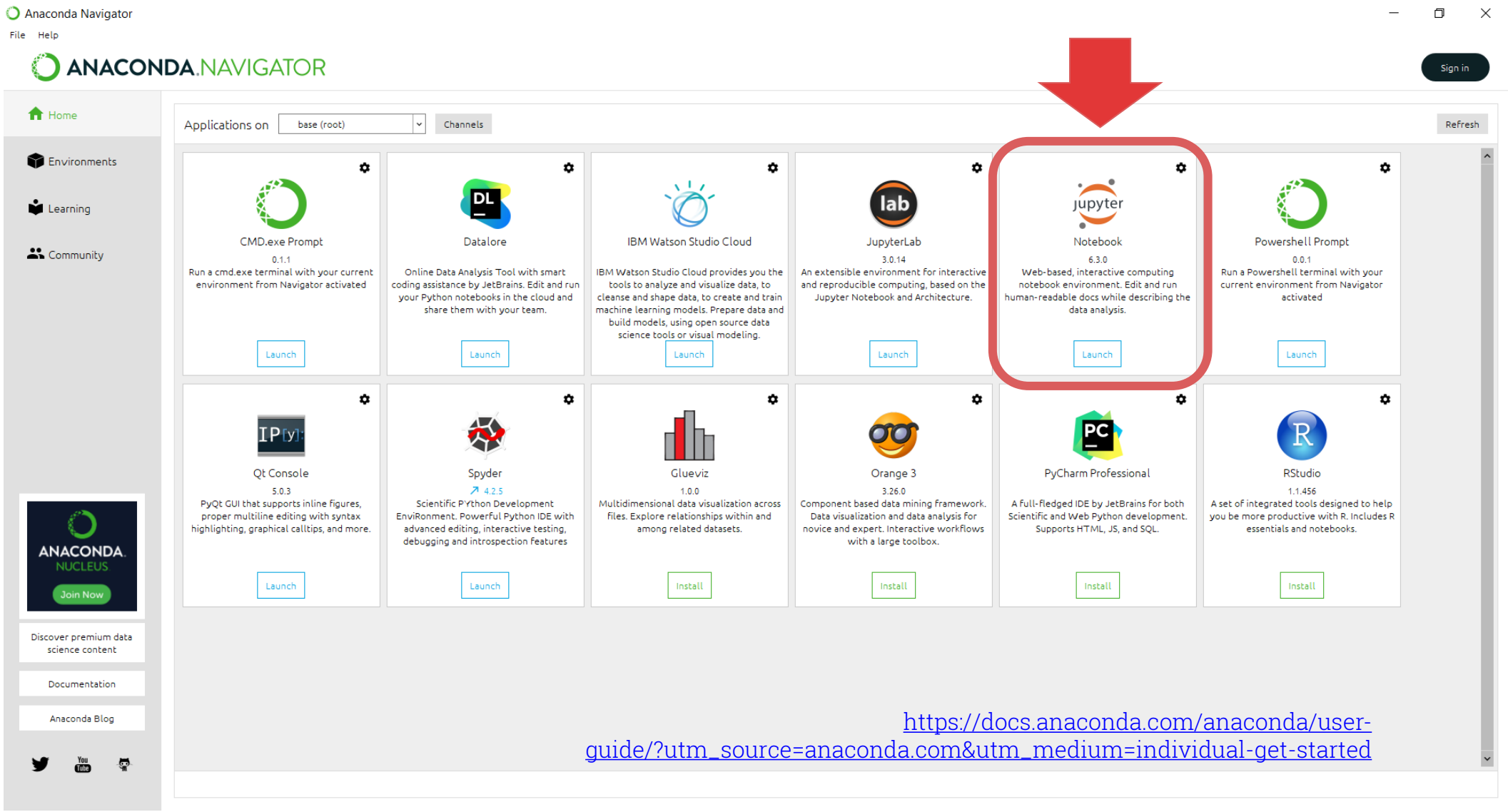
https://www.anaconda.com/products/individual-b



Accept all default settings



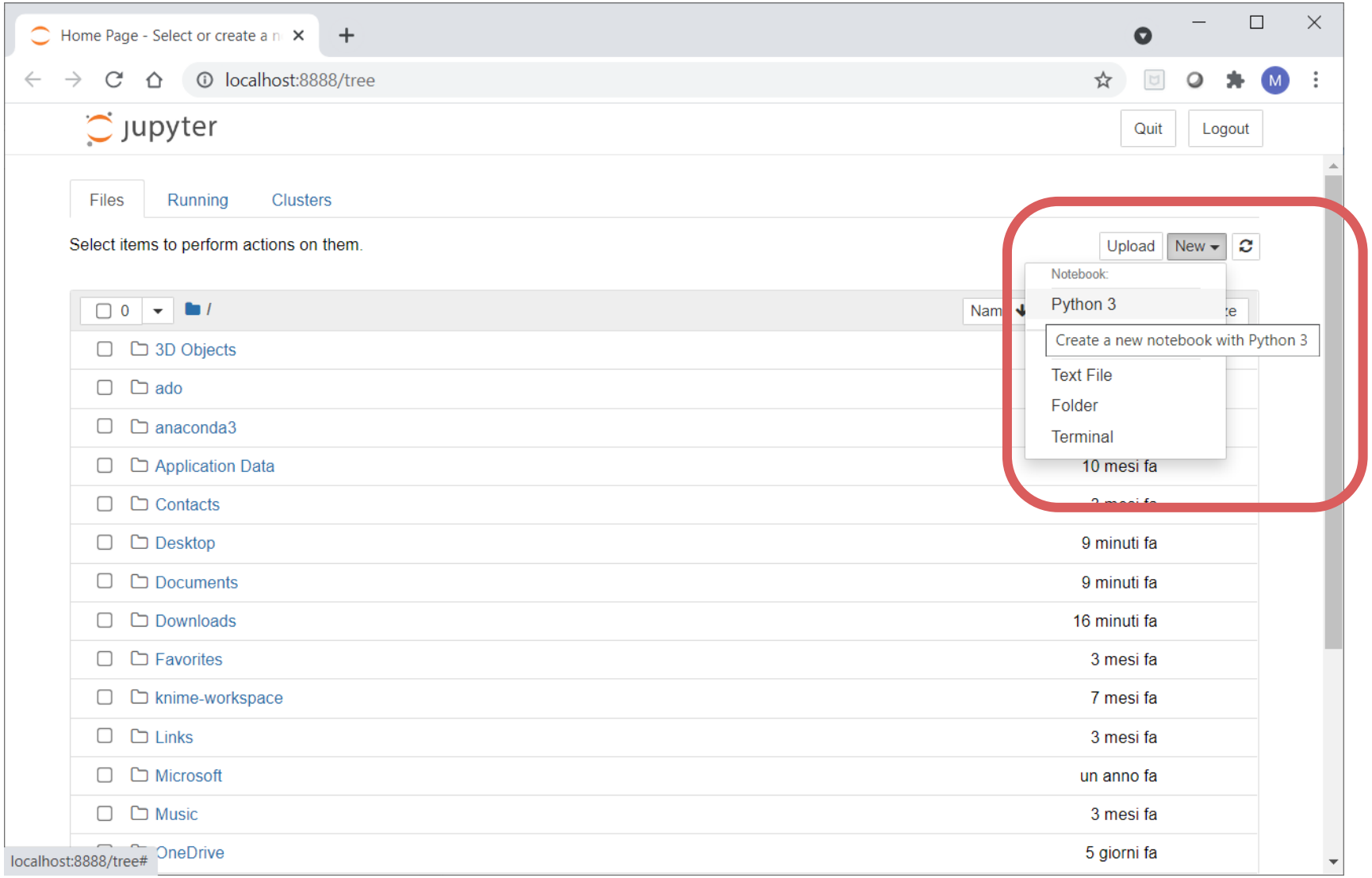
Anaconda Navigator



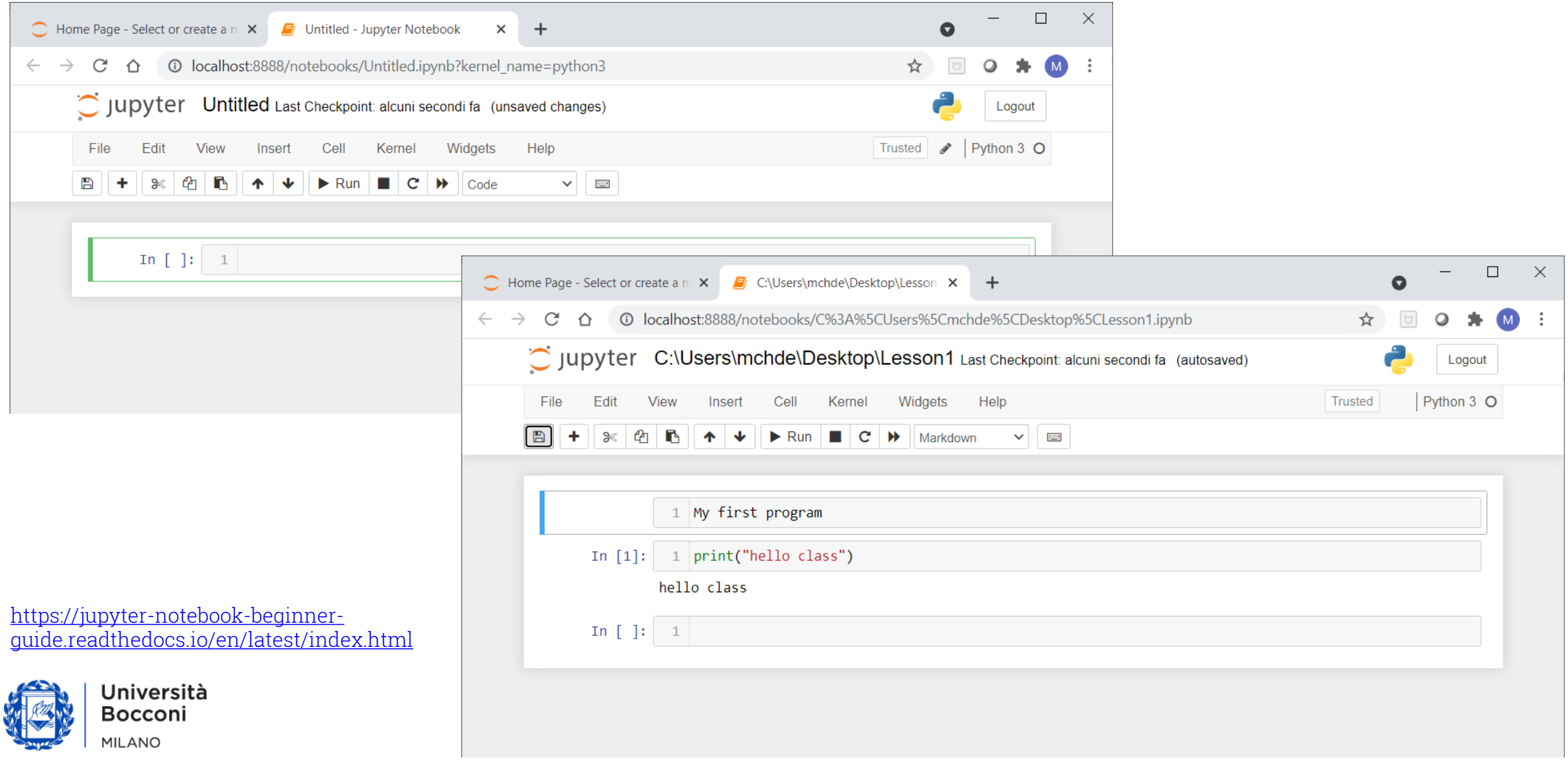
https://docs.anaconda.com/anaconda/user-guide/?utm_source=anaconda.com&utm_medium=individual-get-started

Notebook Dashboard

- It is mainly used to open notebook documents, and to manage the running kernels (visualize and shutdown).
- It has other features similar to a file manager, namely navigating folders and renaming/deleting files.



Create a new notebook

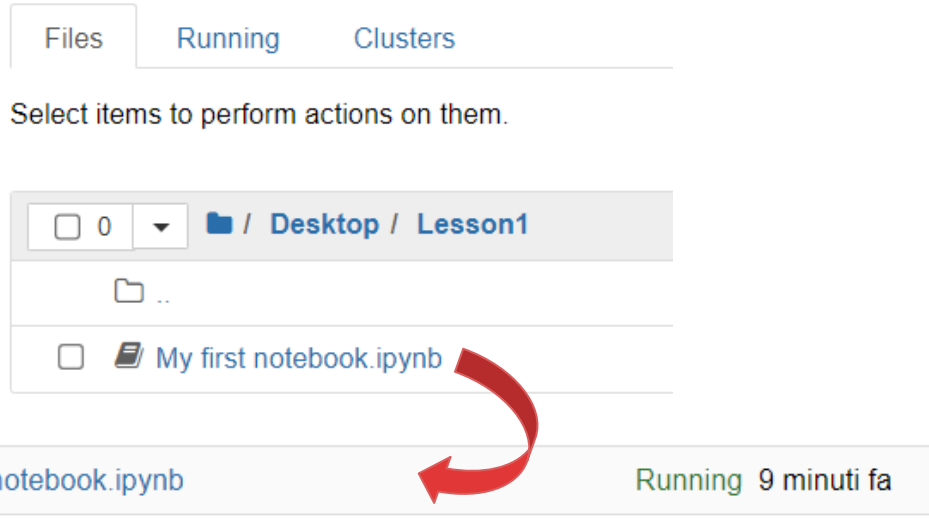


https://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/index.html

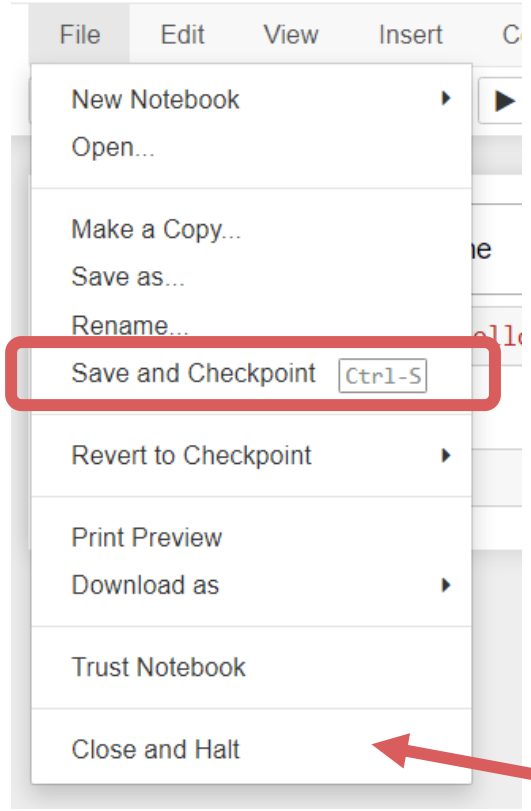
Università Bocconi MILANO

Open, save and download

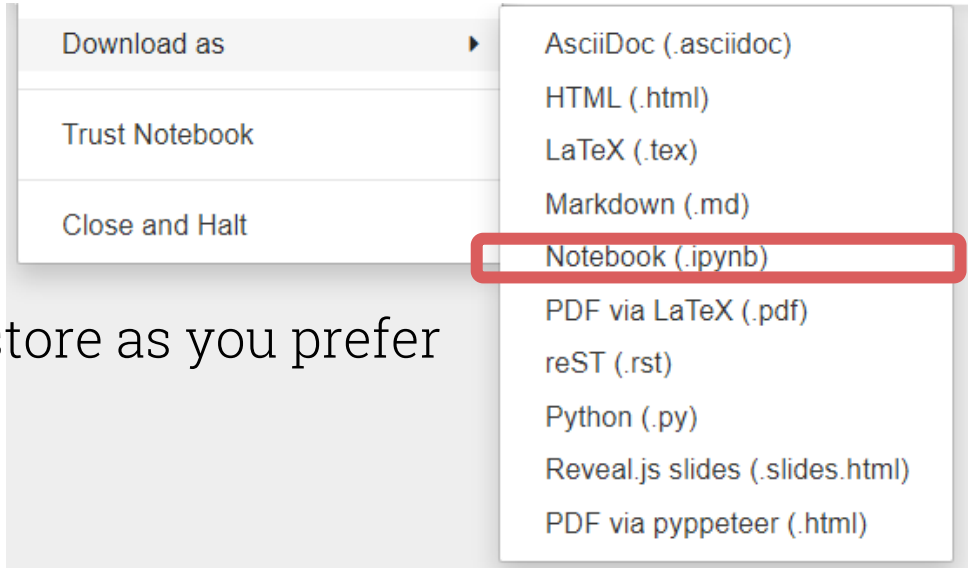
- Open a notebook from the Dashboard



- Save it



- Download and store as you prefer
- After Close it



AND NOW LET'S DO SOME CODING...