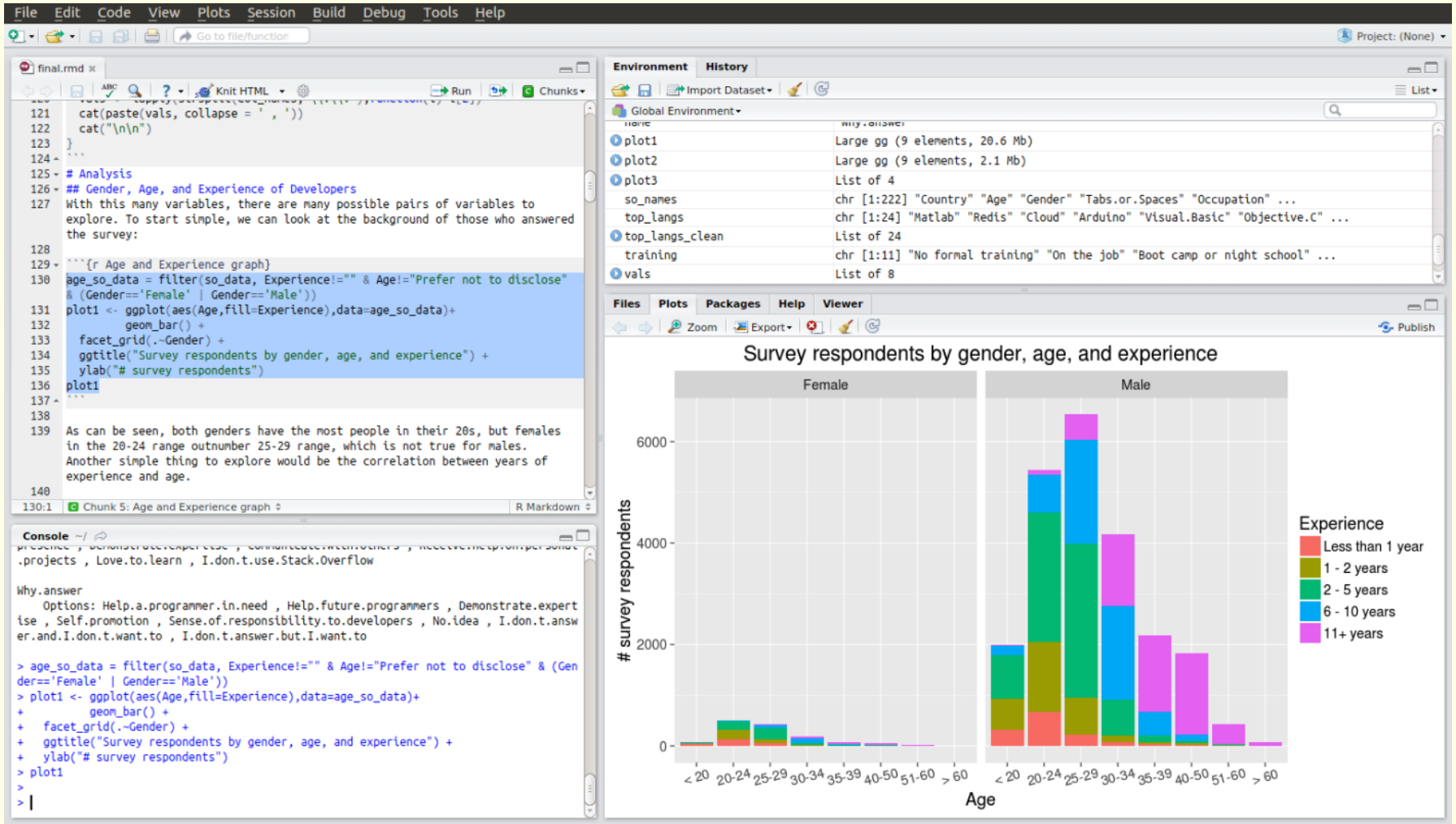
Anglia Ruskin University
School of Computing and Information Science

Principles of Data Mining and Machine Learning (MOD 007892) Practicals

Introduction to Anaconda, Jupyter Notebook

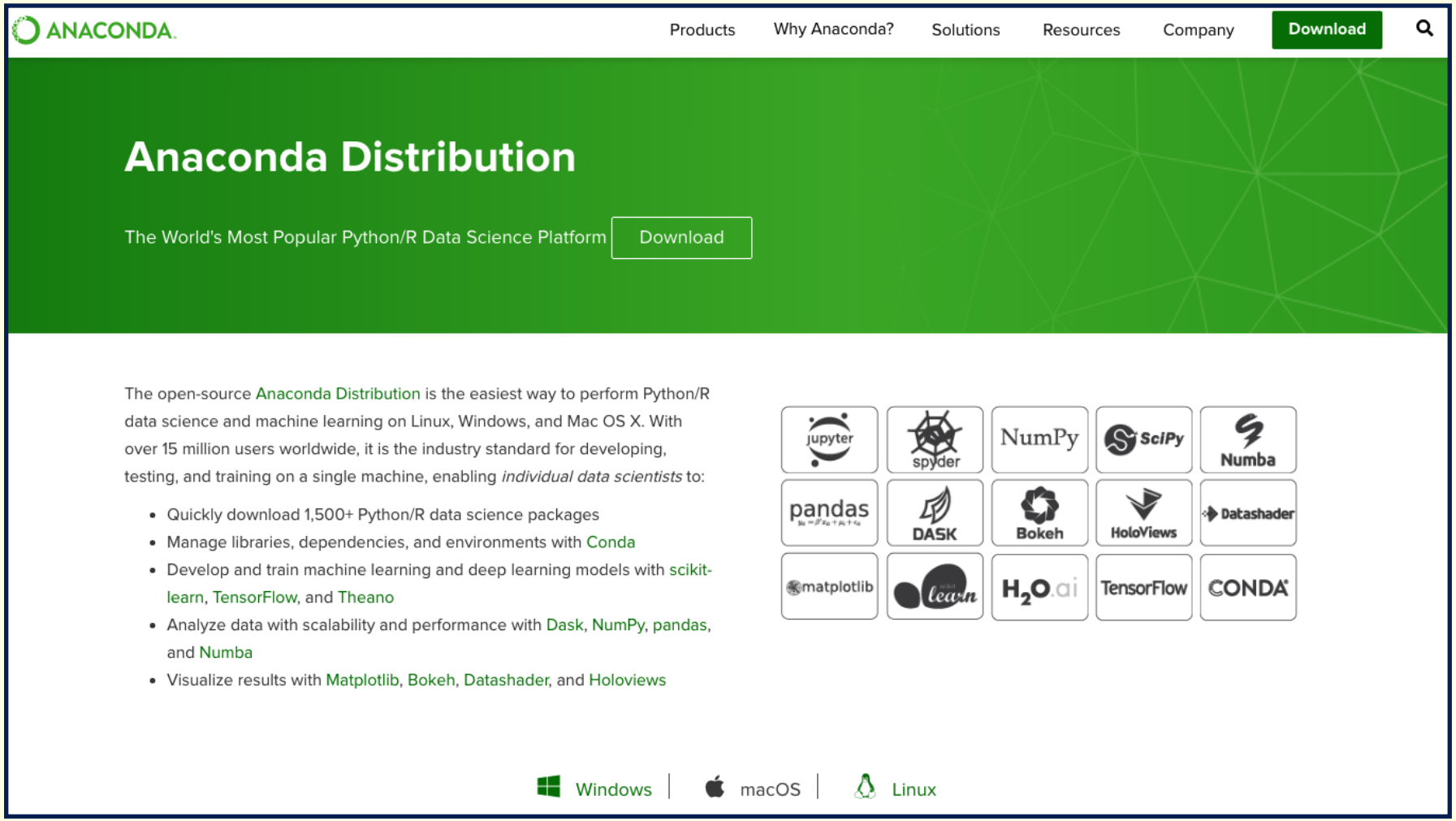
IPython Interactive Computing



- A powerful interactive framework
- A **Kernel** for **Jupyter**
- Tools for scientific and parallel computing

Source: https://ipython.org

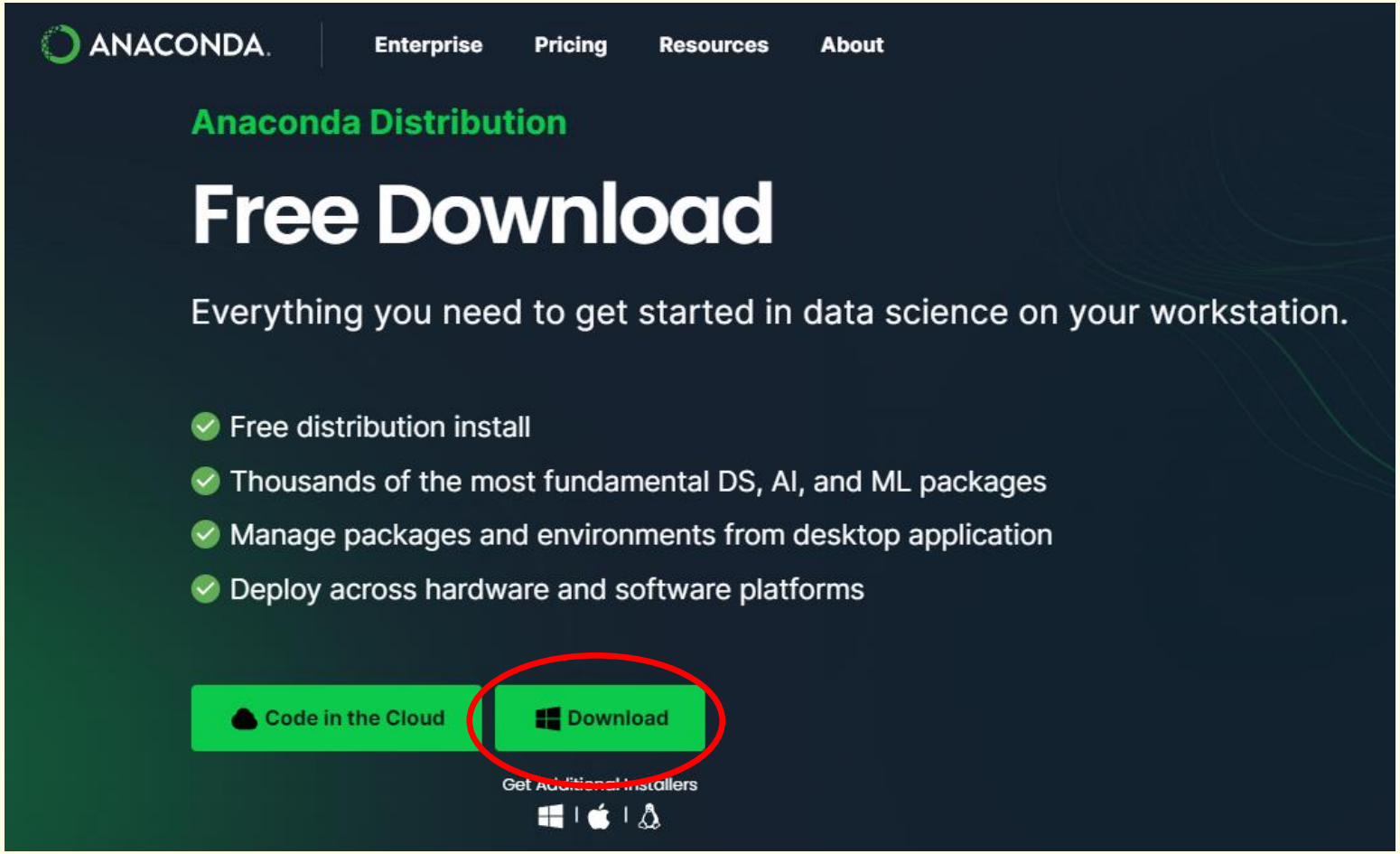
Anaconda installation



The open-source **Anaconda** is a way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X.

Source: https://www.anaconda.com/download

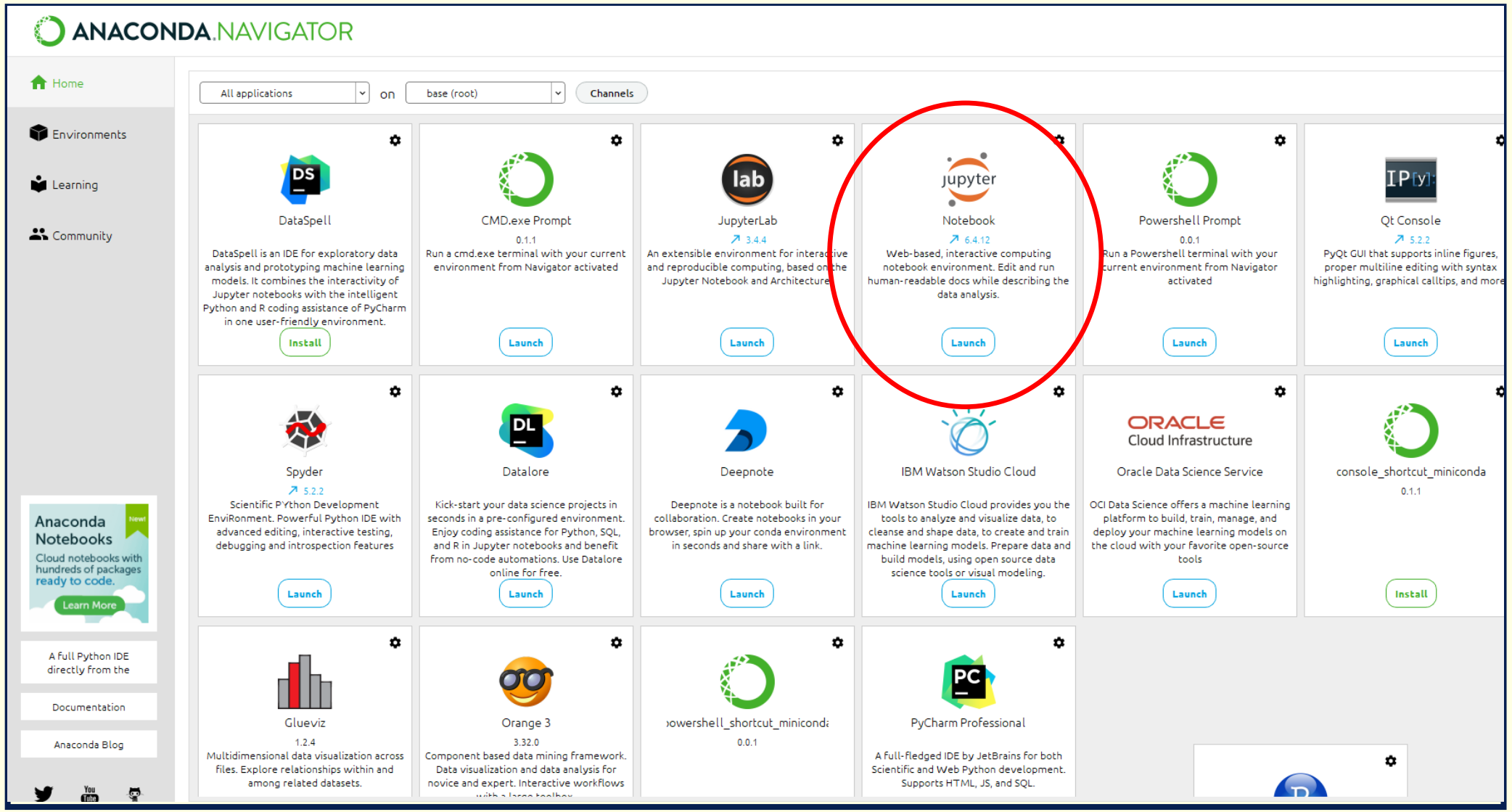
Anaconda and Python-3 installation



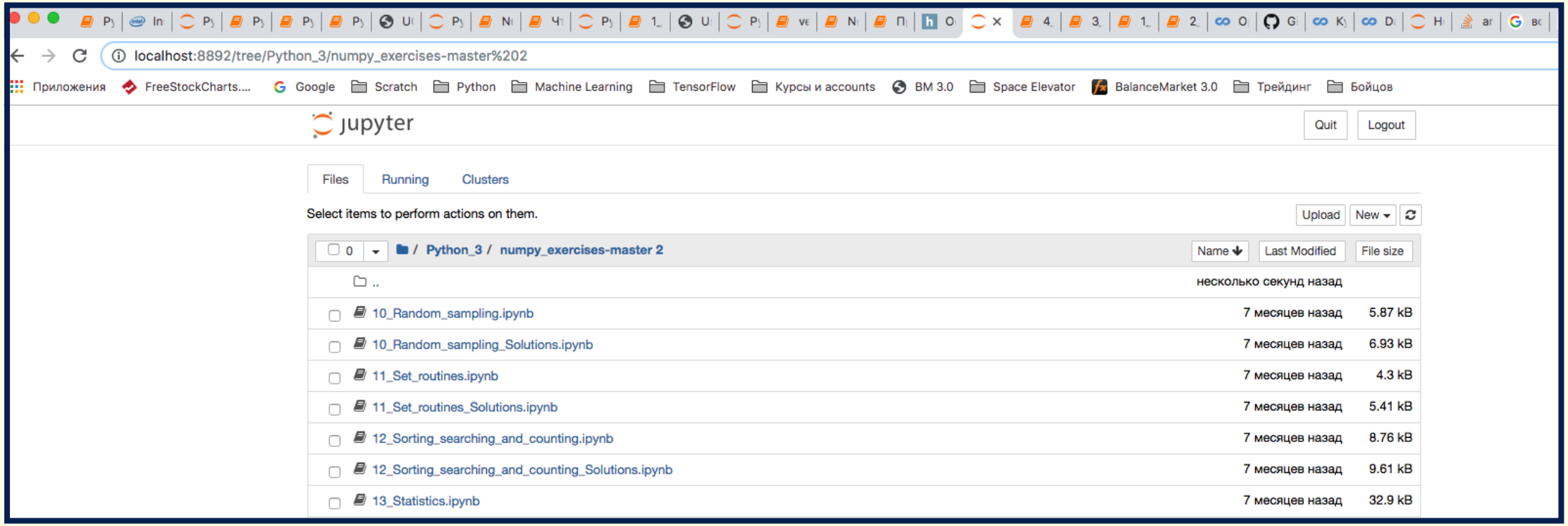
The open-source **Anaconda** is the frame to perform Python/R data science and machine learning tools on Linux, Windows, and Mac OS X

Source: https://www.anaconda.com/download

Anaconda installation



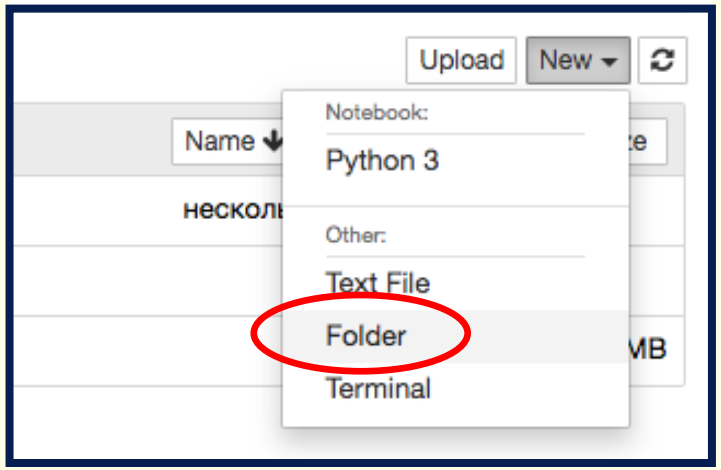
Introduction to Jupyter Notebook



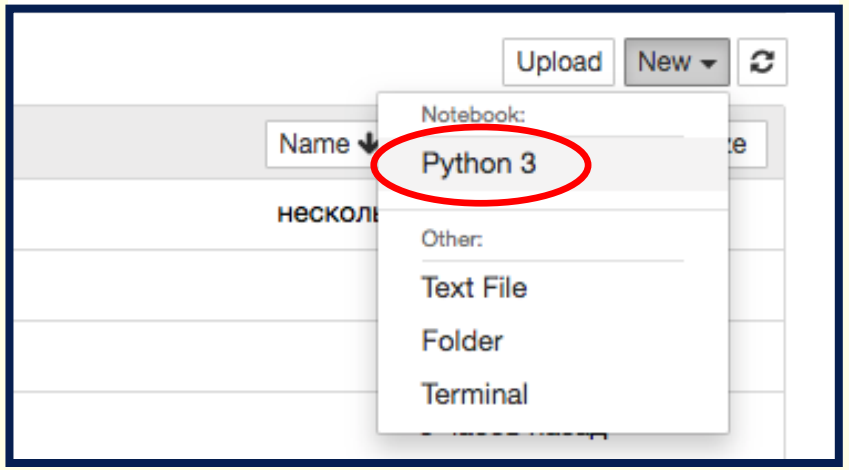
As a result, Jupyter Notebook will be launched in your browser

Introduction to Jupyter Notebook

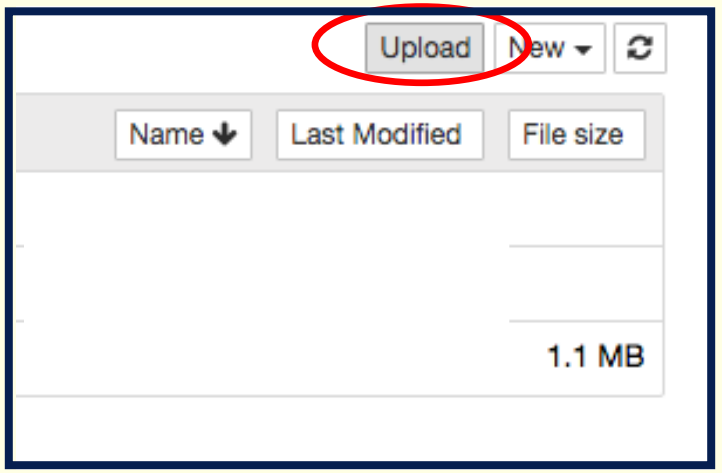
Create a new Folder :



Create a new Jupyter Notebook :



or upload a new Notebook :

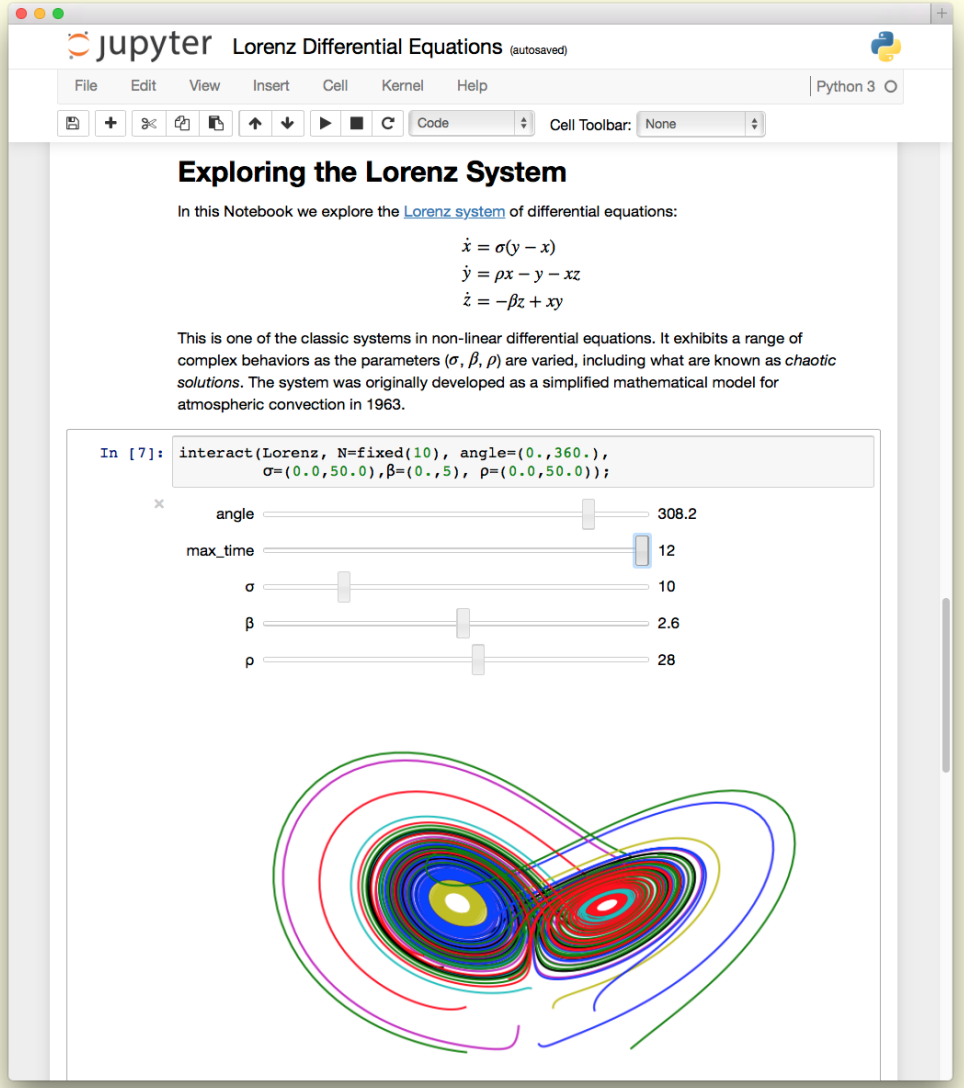


Introduction to Jupyter Notebook

"The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modelling, data visualization, machine learning, and much more."

- The analysis environment for multiple computing languages such as (Python, and R, etc.)
- Supports multiple content types: code, narrative text, images, movies, etc.

Try it in your browser | Install the Notebook

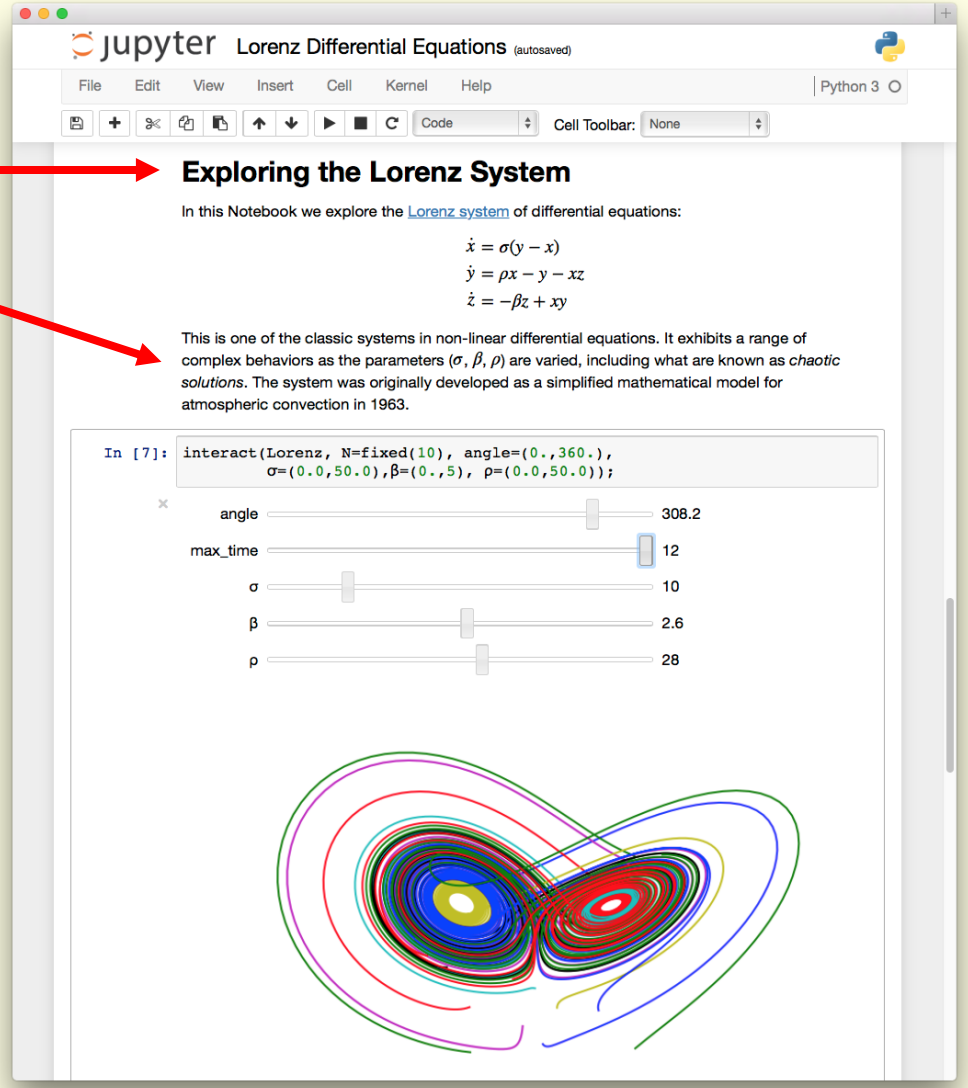


Source: http://jupyter.org/

Introduction to Jupyter Notebook

- HTML & Markdown

It has a header and some text explaining (commenting) what we are about to do.



Source: http://jupyter.org/

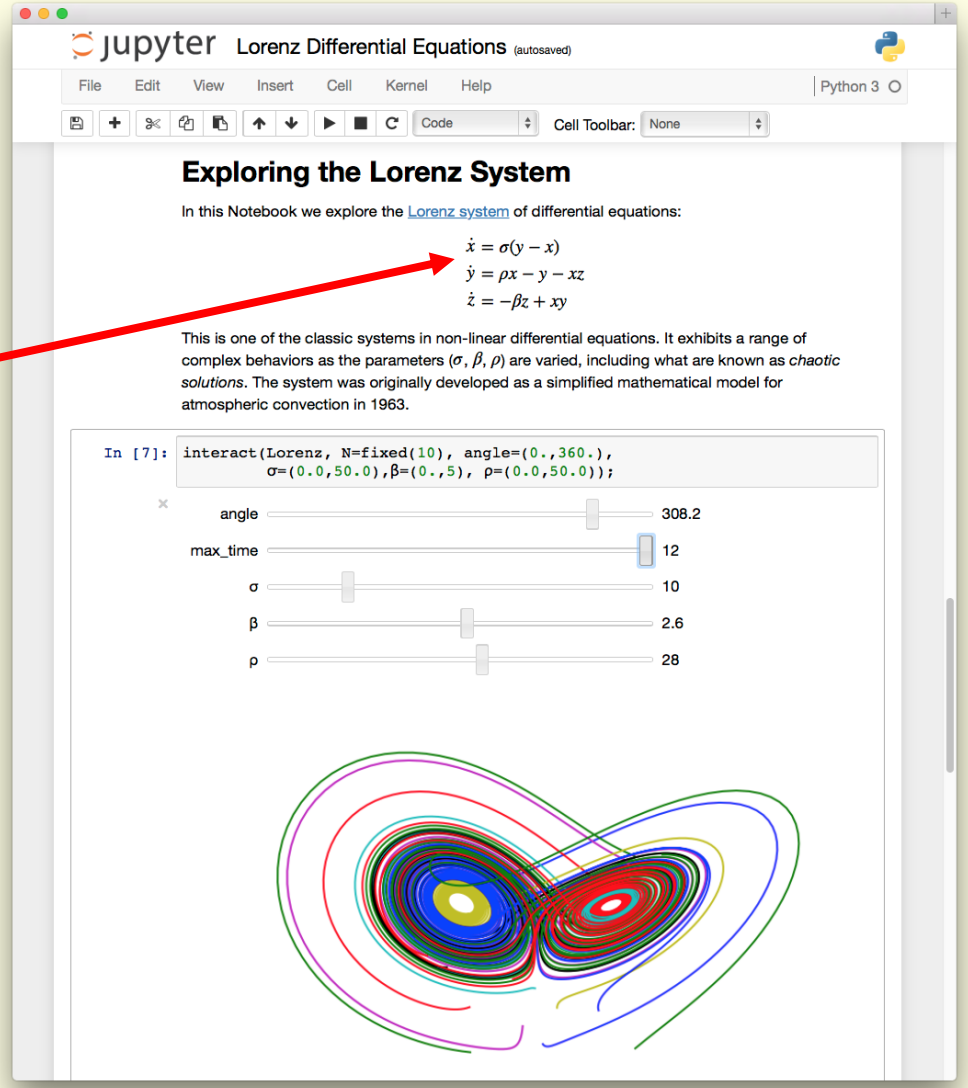
Introduction to Jupyter Notebook

- **HTML & Markdown**

It has a header and some text explaining (commenting) what we are going to do.

- **LaTeX (equations)**

Then, we could add some Latex format equations



Source: http://jupyter.org/

Introduction to Jupyter Notebook

- ## HTML & Markdown

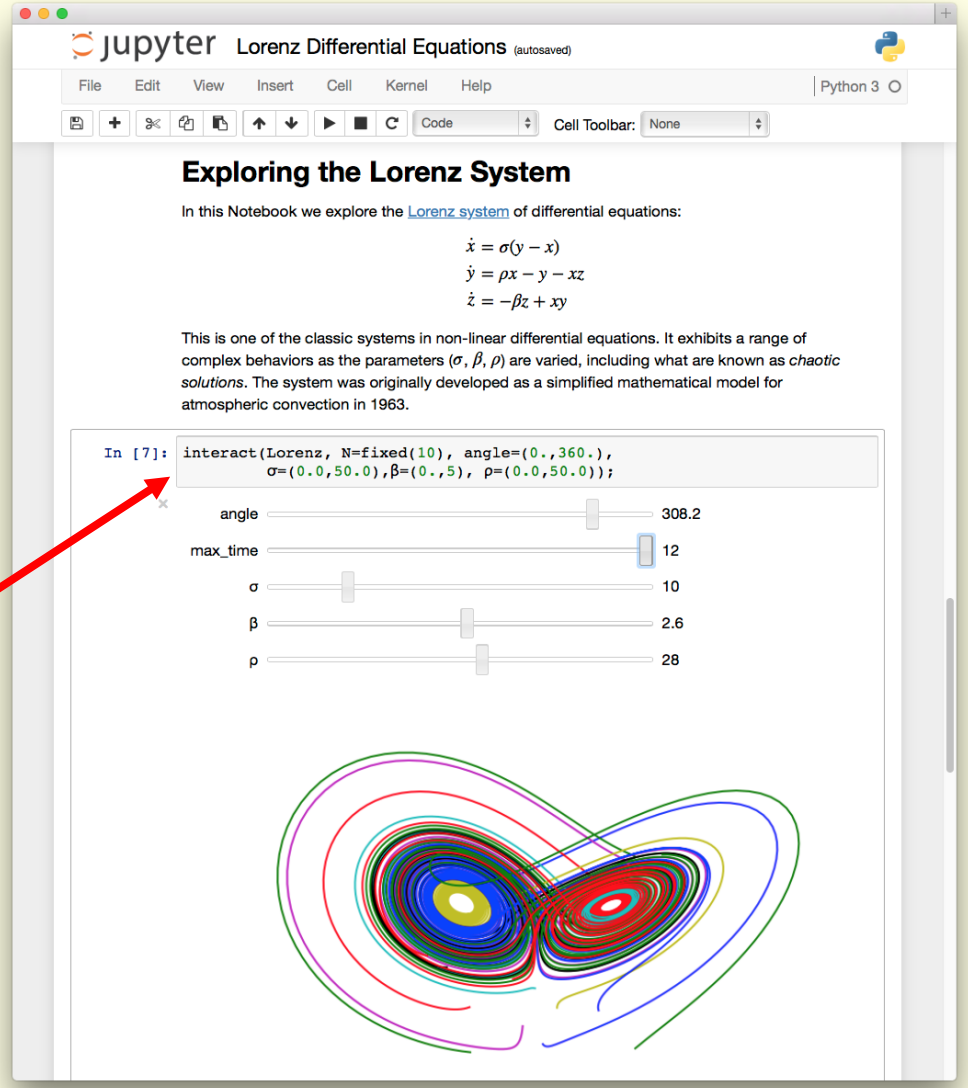
It has a header and some text explaining (commenting) what we are about to do.

- ## LaTeX (equations)

It is then followed by some Latex equations.

- ## Code

- There is some code which can define variables, modify them, etc.
- It connects programming with storytelling.
- Notebook could be saved to a .ipynb file which can be converted to html, pdf etc. to create websites or presentations.



Source: http://jupyter.org/

Introduction to Jupyter Notebook

- ## HTML & Markdown

It has a header and some text explaining (commenting) what we are about to do.

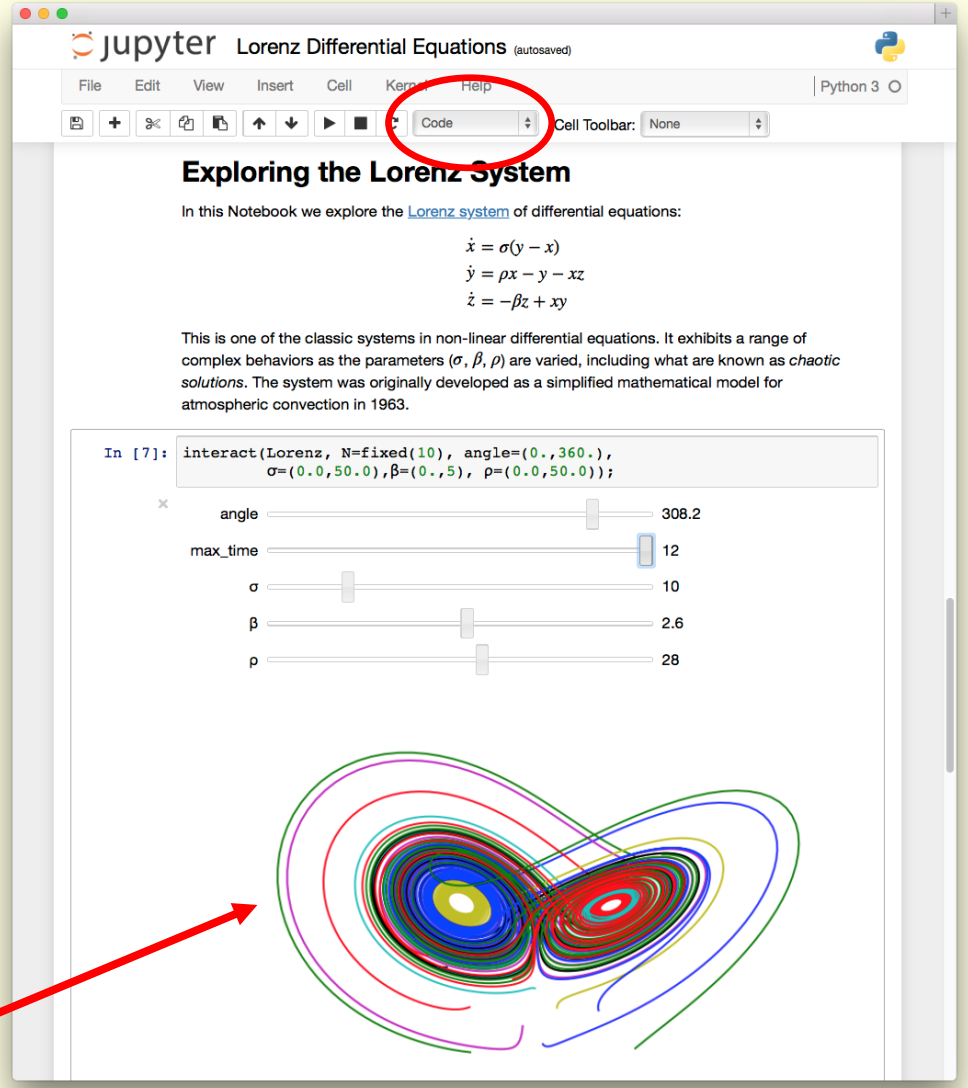
- ## LaTeX (equations)

It is then followed by some Latex equations.

- ## Code

- There is some code, which can define variables, modify them, etc.
- It connects programming with storytelling.
- The work is saved to a .ipynb file which can be converted to html, pdf etc. to create websites or presentations.

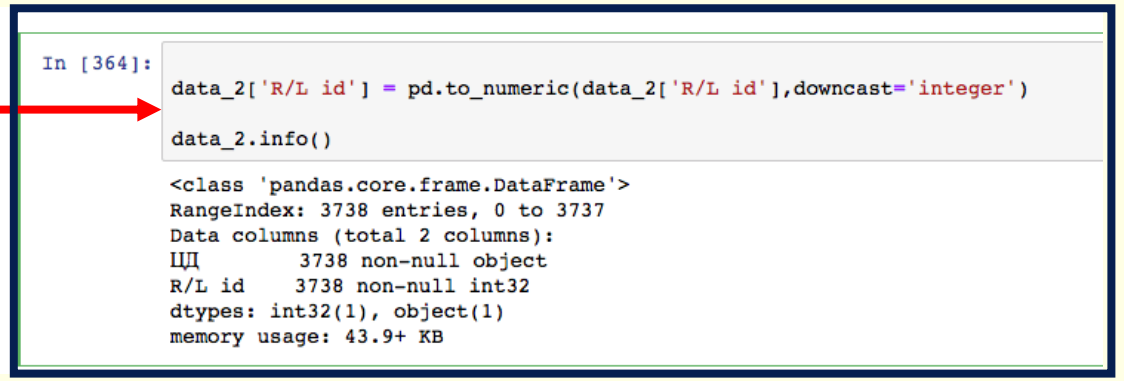
- ## Images & Movies



Source: http://jupyter.org/

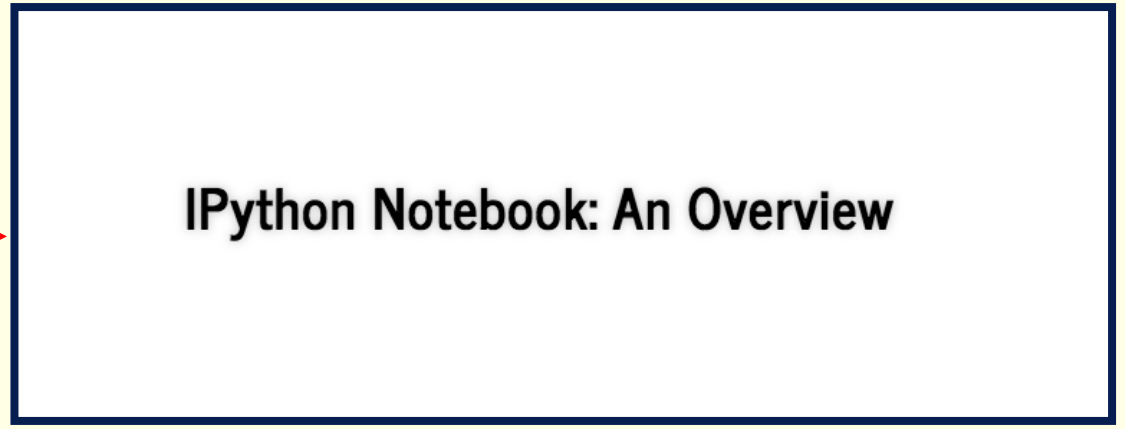
Introduction to Jupyter Notebook

Code is divided into cells to control execution



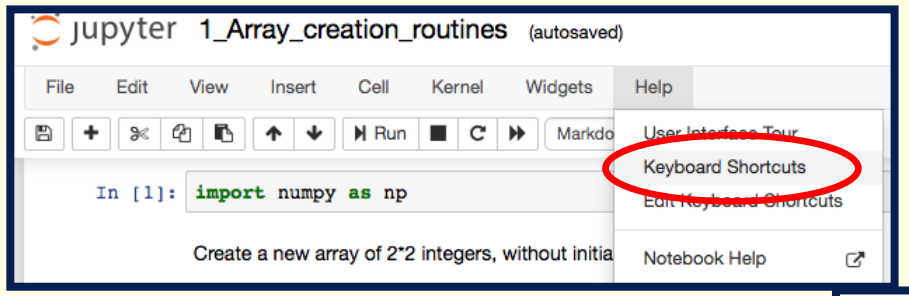
```
In [364]:
data_2['R/L id'] = pd.to_numeric(data_2['R/L id'],downcast='integer')

data_2.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3738 entries, 0 to 3737
Data columns (total 2 columns):
ЦД         3738 non-null object
R/L id     3738 non-null int32
dtypes: int32(1), object(1)
memory usage: 43.9+ KB
```

The cells might be used as a presentation.
For example: Jupyter interactive presentation from Ben Zaitlen.
http://quasiben.github.io/dfwmeetup_2014/#/



Principles of Data Mining and Machine Learning (MOD 007892) (Practicals)

Introduction to Jupyter Notebook



Jupyter Keyboard Hotkeys

Jupyter Keyboard Shortcuts

Keyboard shortcuts

The Jupyter Notebook has two different keyboard input modes. **Edit mode** allows you to type code/text into a cell and is indicated by a green cell border. **Command mode** binds the keyboard to notebook level actions and is indicated by a grey cell border with a blue left margin.

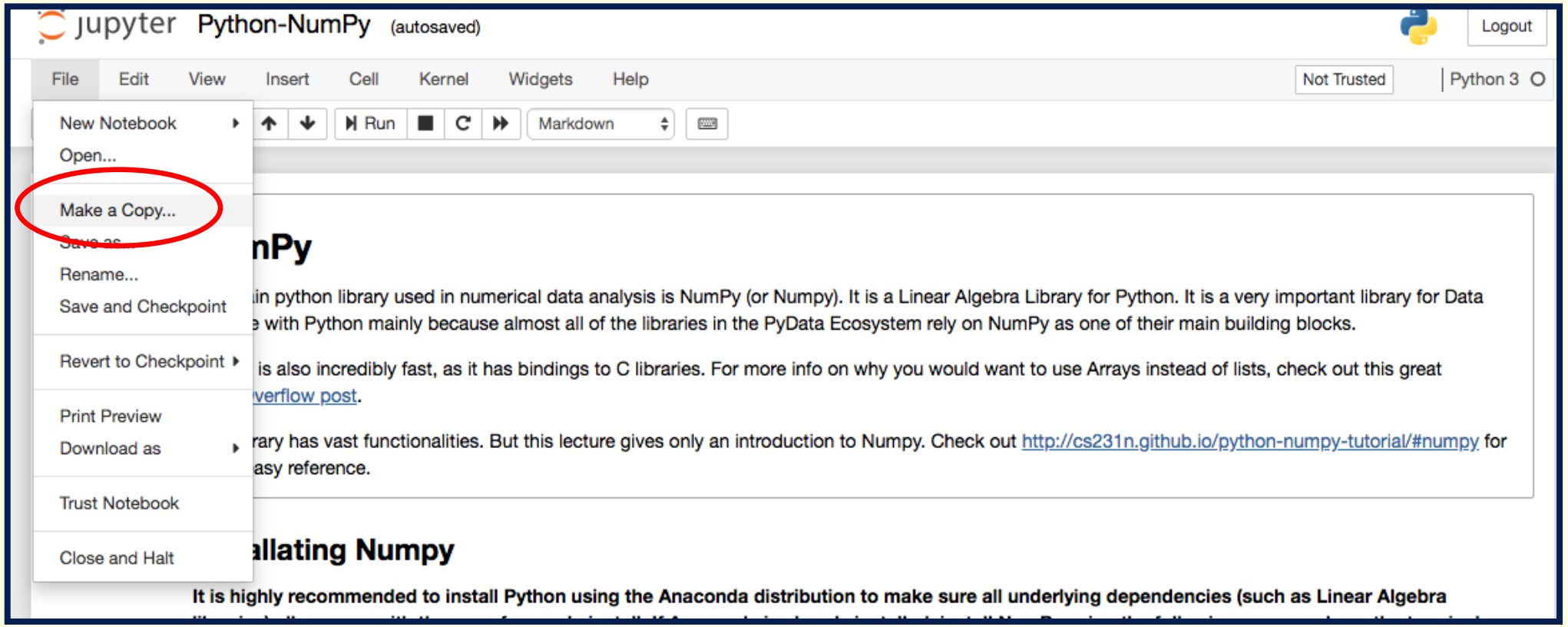
Command Mode (press `Esc` to enable)

`F`: find and replace	`Shift-J`: extend selected cells below
`Ctrl-Shift-P`: open the command palette	`A`: insert cell above
`Enter`: enter edit mode	`B`: insert cell below
`Shift-Enter`: run cell, select below	`X`: cut selected cells
`Ctrl-Enter`: run selected cells	`C`: copy selected cells
`Alt-Enter`: run cell, insert below	`Shift-V`: paste cells above

Keyboard shortcuts can be viewed from Help → Keyboard Shortcuts

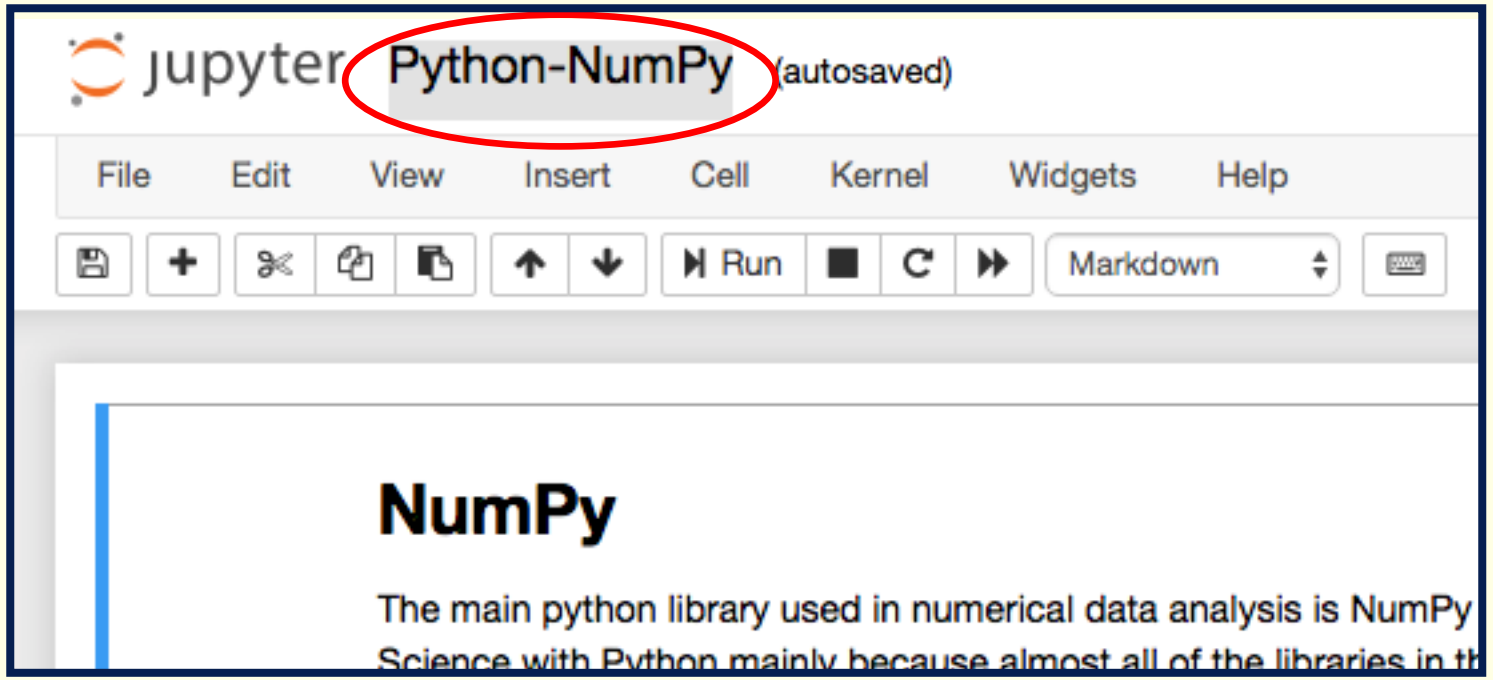
Introduction to Jupyter Notebook

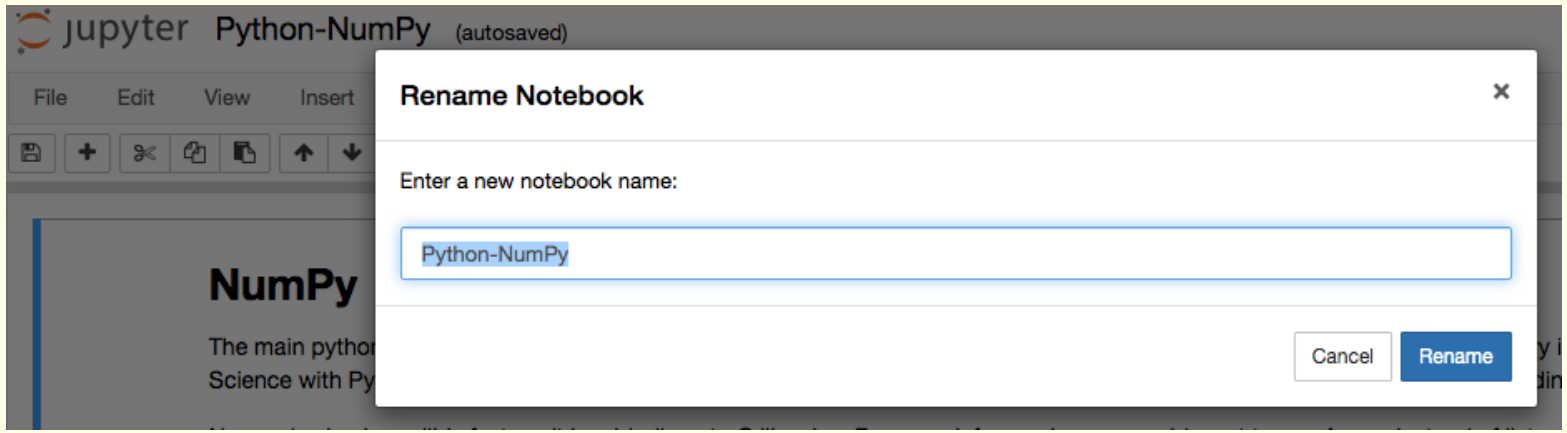
How to save it ?



Introduction to Jupyter Notebook

How to save it ?





Practicals

- Go to Jupyter Notebook
- Open ‘EDA-notebook.ipynb’
- Type in an empty cell all the code from a picture in the middle such as:

2.png

```
#Use the list to initialise a numpy array.
my_array = np.array(my_list)
my_array
```

The sell for Typing your code :

```
In [16]: #Use the list to initialise a numpy array.

In [ ]:

In [3]:

Out[3]: array([100, 300, 500, 700, 900])
```

The answer for comparison. Do not type the code here, please!