

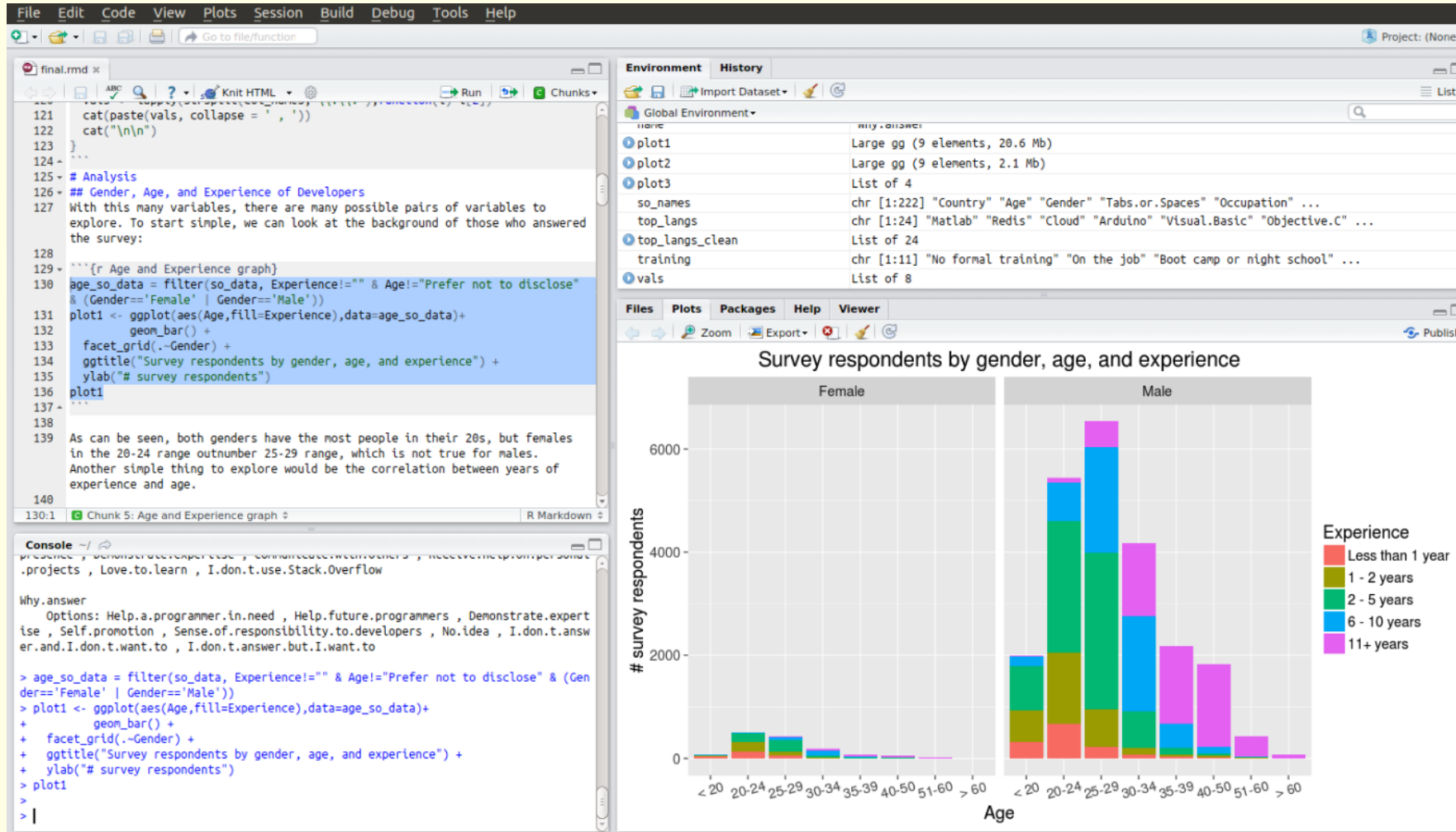


Anglia Ruskin University

School of Computing and Information Science

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

Introduction to Anaconda, Jupyter
Notebook



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

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.

The screenshot shows the Anaconda Distribution website. At the top is a navigation bar with links: Products, Why Anaconda?, Solutions, Resources, Company, and a green Download button. The main header area has a green background with the text 'Anaconda Distribution' and 'The World's Most Popular Python/R Data Science Platform' followed by another Download button. Below this, a paragraph describes Anaconda as the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X, mentioning over 15 million users. A bulleted list highlights key features: downloading 1,500+ packages, managing libraries with Conda, developing models with scikit-learn, TensorFlow, and Theano, analyzing data with Dask, NumPy, pandas, and Numba, and visualizing results with Matplotlib, Bokeh, Datashader, and Holoviews. To the right of the text is a grid of 15 logos for various data science libraries and tools. At the bottom, there are icons for Windows, macOS, and Linux.

ANACONDA

Products Why Anaconda? Solutions Resources Company **Download**

Anaconda Distribution

The World's Most Popular Python/R Data Science Platform **Download**

The open-source **Anaconda Distribution** is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 15 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

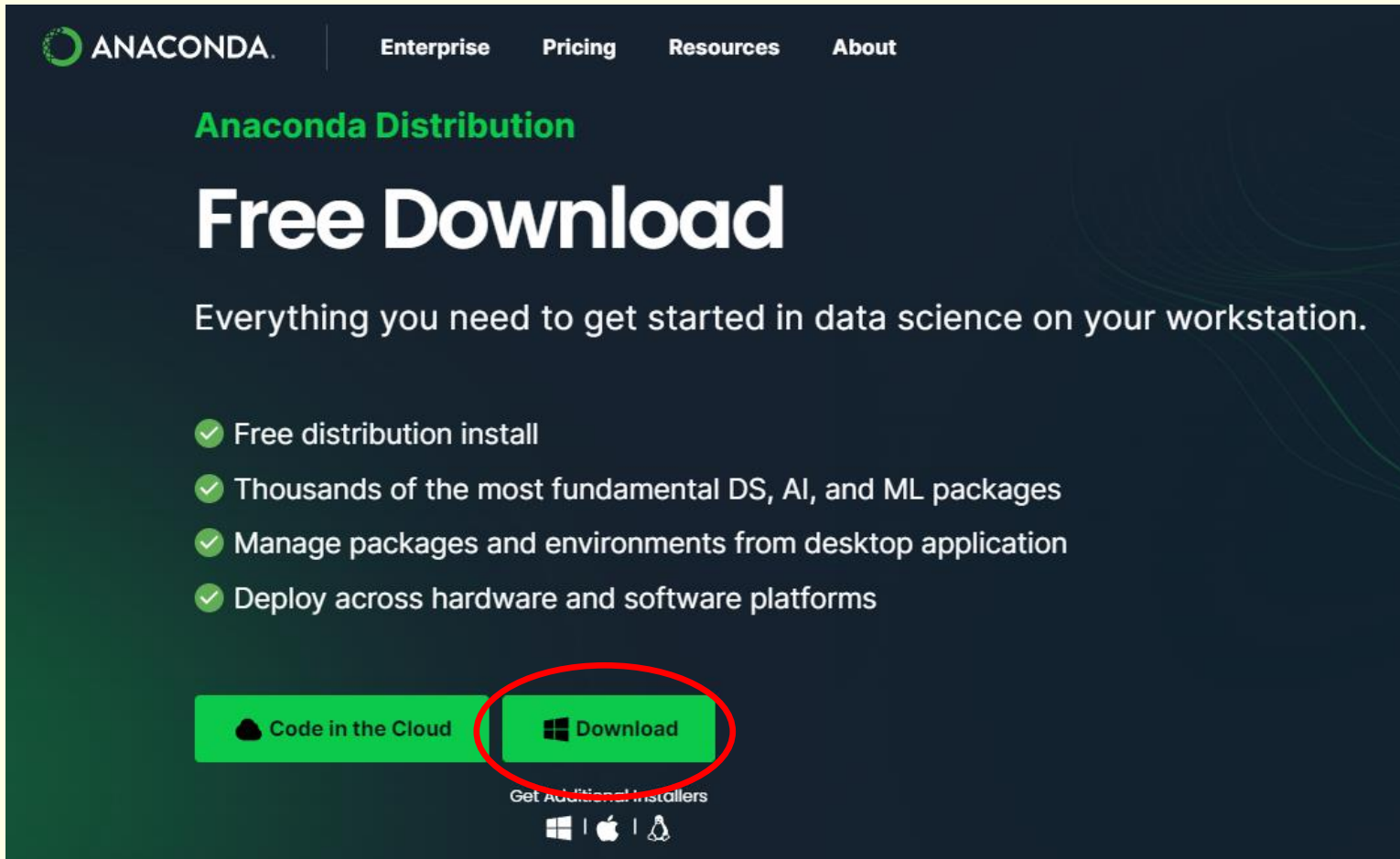
- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with **Conda**
- Develop and train machine learning and deep learning models with **scikit-learn**, **TensorFlow**, and **Theano**
- Analyze data with scalability and performance with **Dask**, **NumPy**, **pandas**, and **Numba**
- Visualize results with **Matplotlib**, **Bokeh**, **Datashader**, and **Holoviews**

Logos displayed: jupyter, spyder, NumPy, SciPy, Numba, pandas, DASK, Bokeh, HoloViews, Datashader, matplotlib, learn, H₂O.ai, TensorFlow, CONDA

Windows | macOS | Linux

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

Anaconda and Python-3 installation





ANACONDA. Enterprise Pricing Resources About

Anaconda Distribution




Free Download

Everything you need to get started in data science on your workstation.

- ✓ Free distribution install
- ✓ Thousands of the most fundamental DS, AI, and ML packages
- ✓ Manage packages and environments from desktop application
- ✓ Deploy across hardware and software platforms

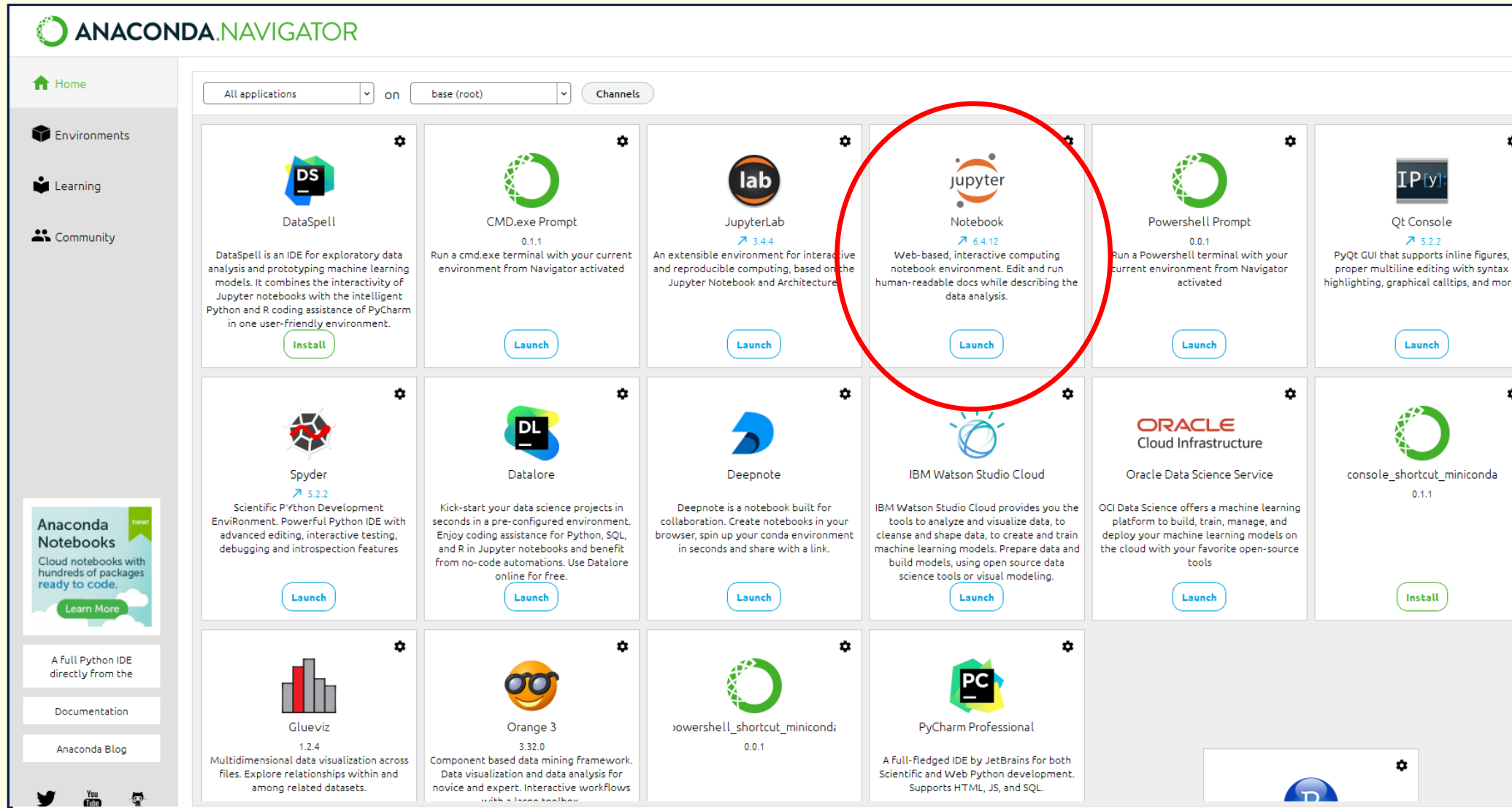
 Code in the Cloud  Download

Get Additional Installers

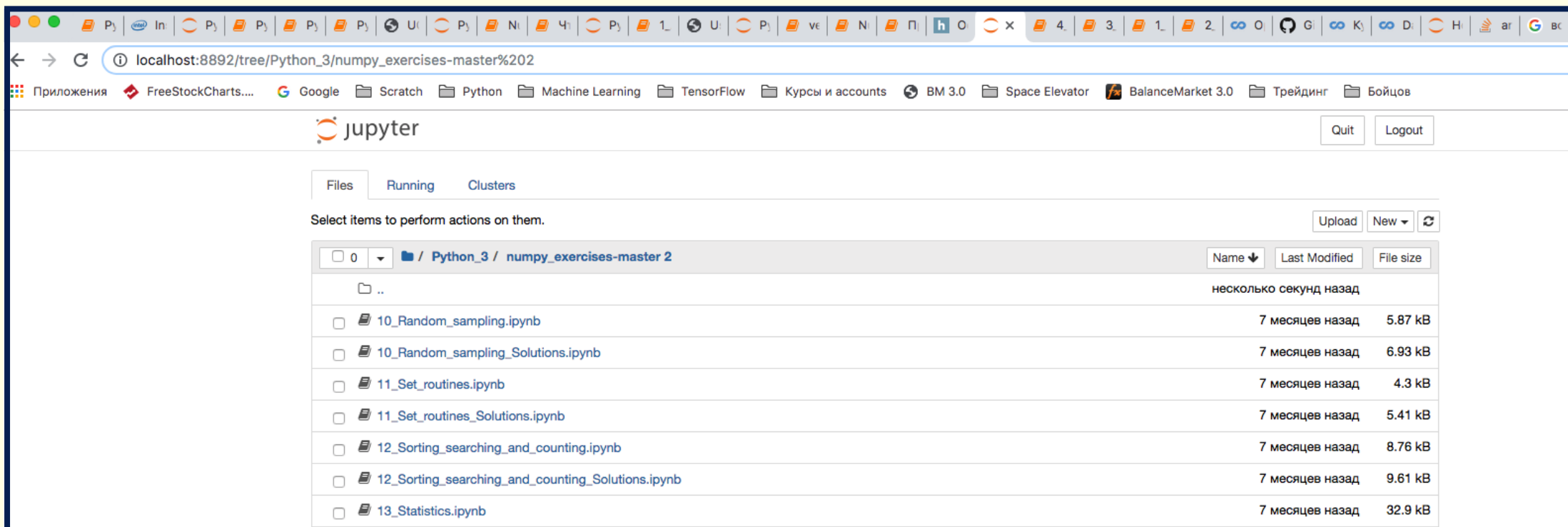
  

The open-source **Anaconda** is the frame to perform Python/R data science and machine learning tools on Linux, Windows, and Mac OS X

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



Introduction to Jupyter Notebook

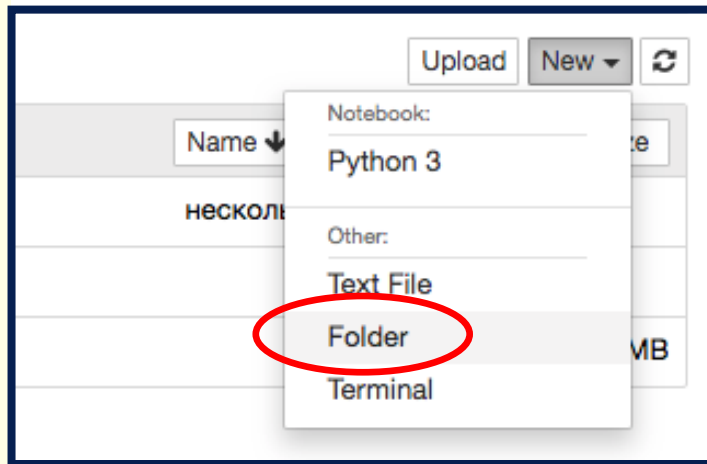


	Name	Last Modified	File size
<input type="checkbox"/>	..	несколько секунд назад	
<input type="checkbox"/>	10_Random_sampling.ipynb	7 месяцев назад	5.87 kB
<input type="checkbox"/>	10_Random_sampling_Solutions.ipynb	7 месяцев назад	6.93 kB
<input type="checkbox"/>	11_Set_routines.ipynb	7 месяцев назад	4.3 kB
<input type="checkbox"/>	11_Set_routines_Solutions.ipynb	7 месяцев назад	5.41 kB
<input type="checkbox"/>	12_Sorting_searching_and_counting.ipynb	7 месяцев назад	8.76 kB
<input type="checkbox"/>	12_Sorting_searching_and_counting_Solutions.ipynb	7 месяцев назад	9.61 kB
<input type="checkbox"/>	13_Statistics.ipynb	7 месяцев назад	32.9 kB

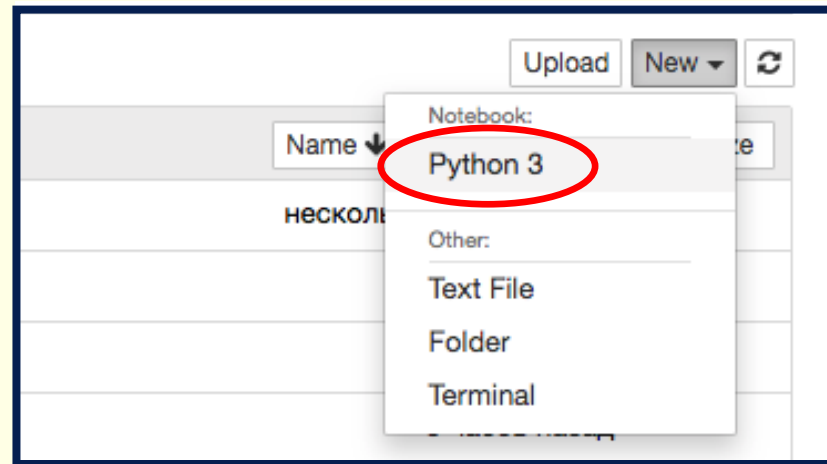
As a result, Jupyter Notebook will be launched in your browser

Introduction to Jupyter Notebook

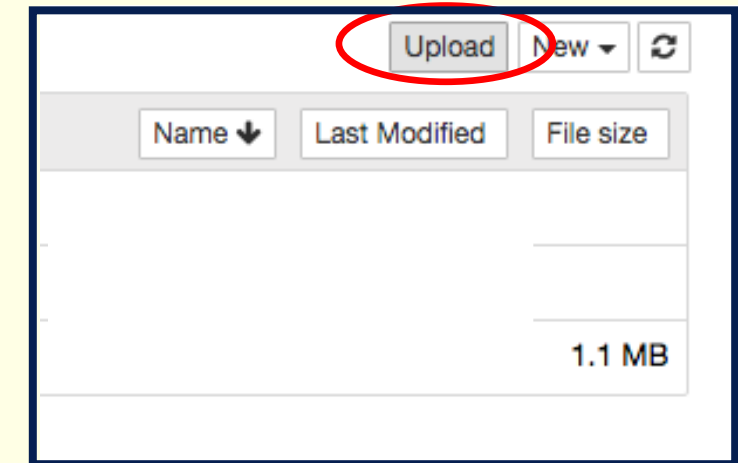
Create a new Folder :



Create a new Jupyter Notebook :



or upload a new Notebook :



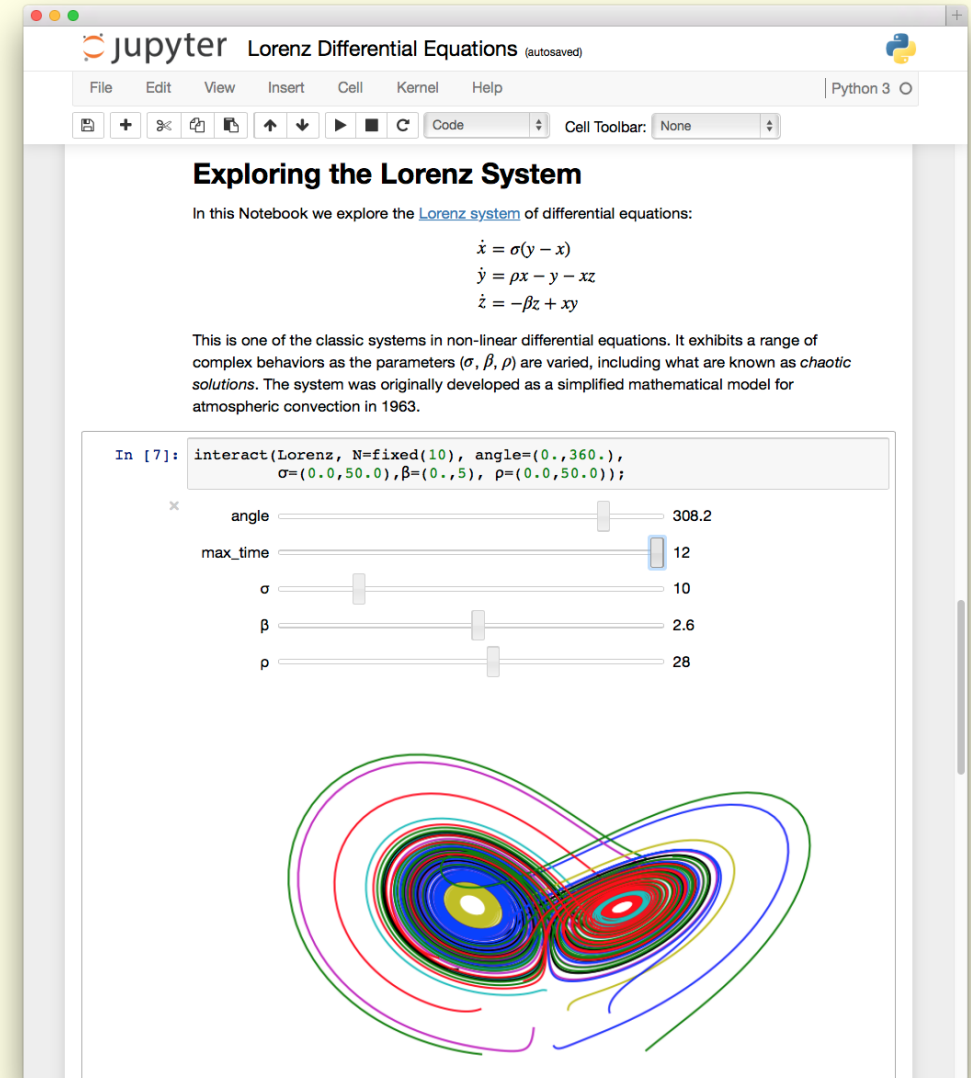
Introduction to Jupyter Notebook

“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

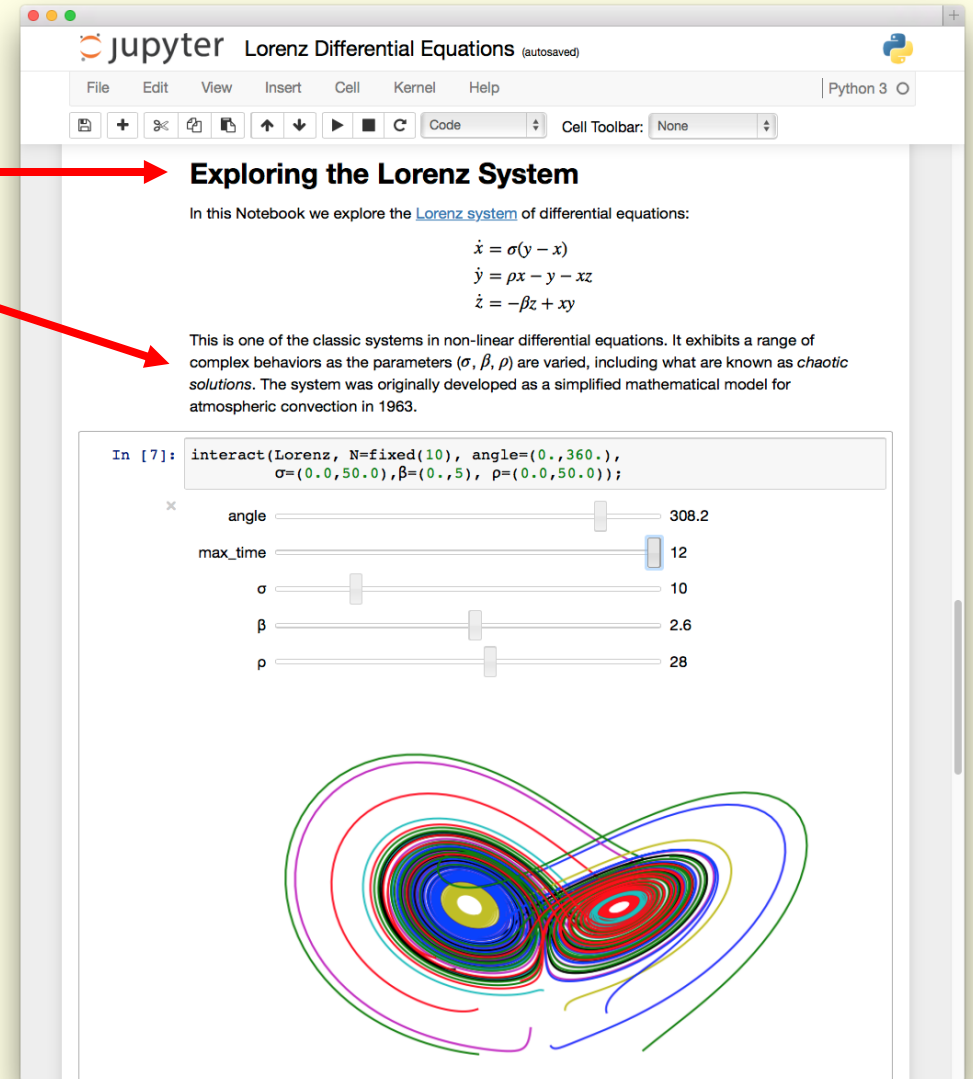


Source: <http://jupyter.org/>

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/>

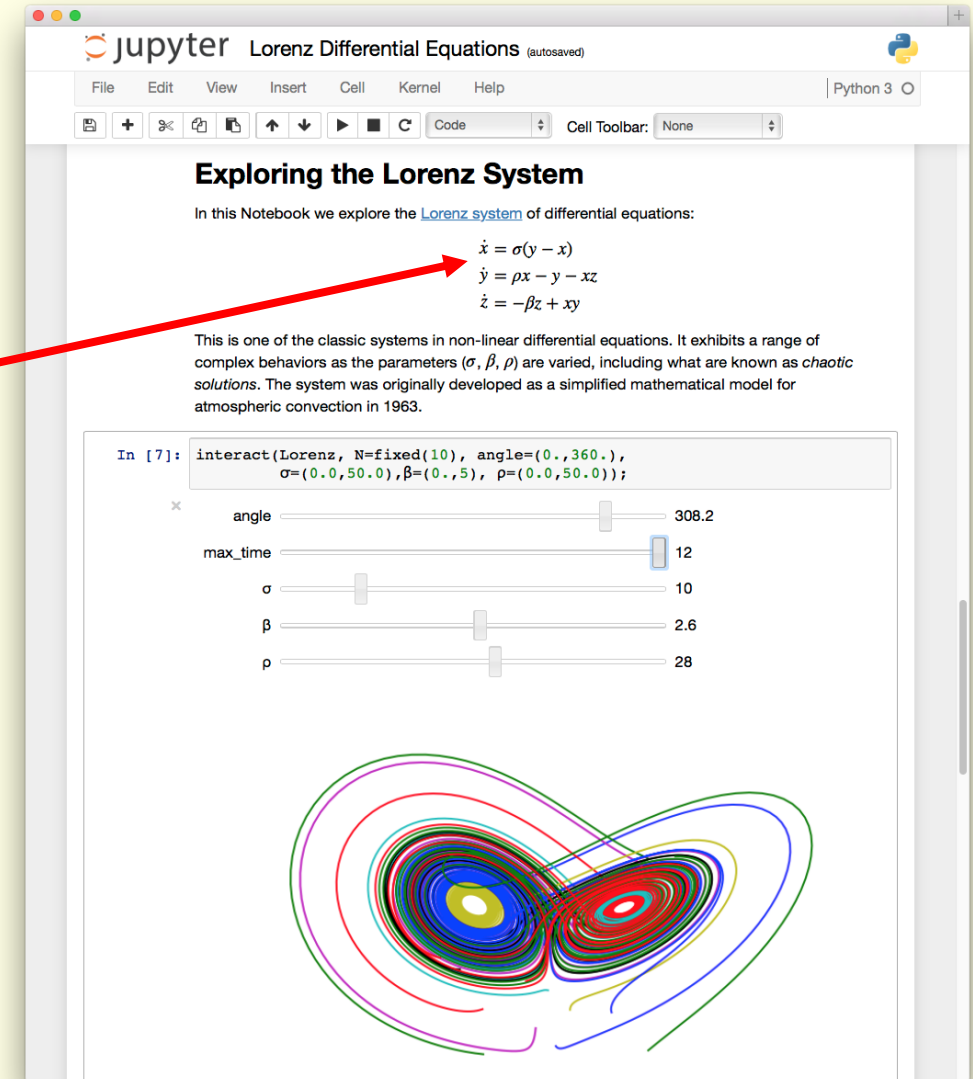
Introduction to Jupyter Notebook

- **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**

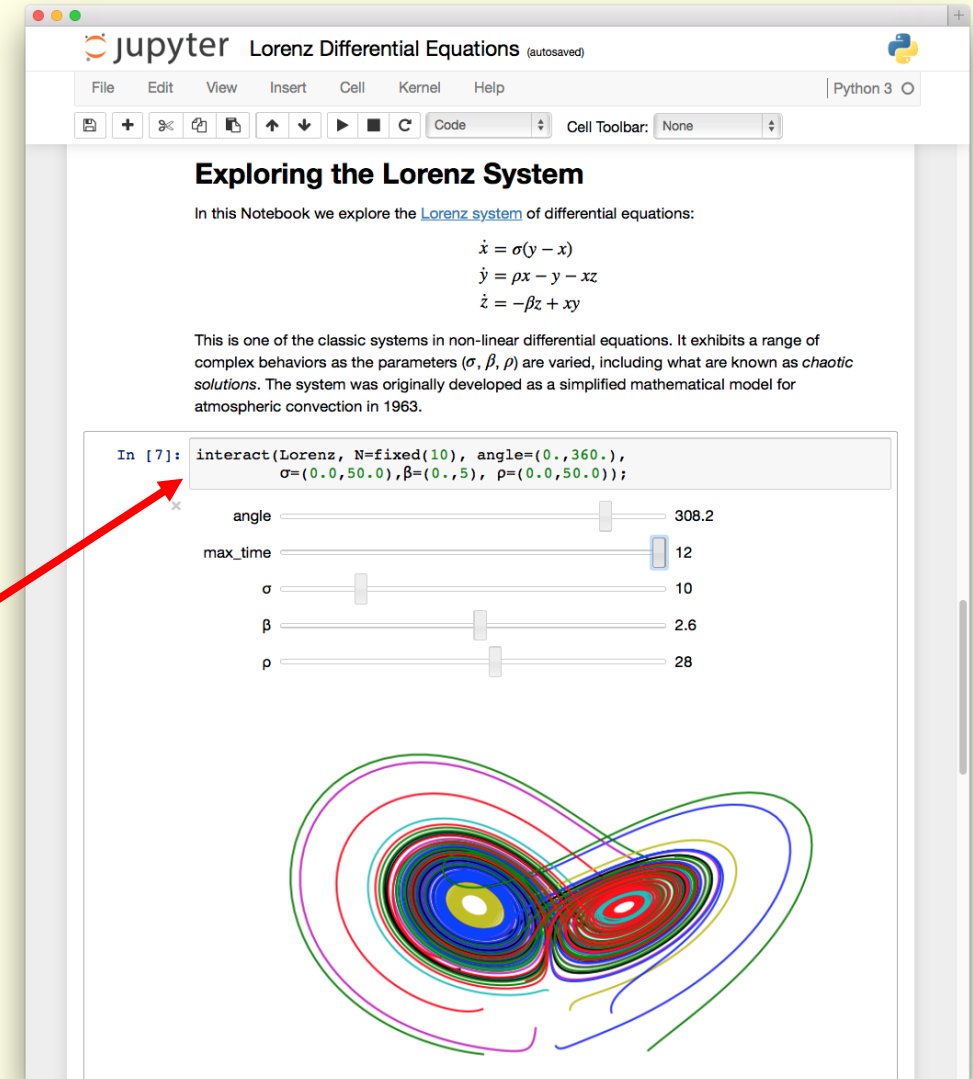
It has a header and some text explaining (commenting) what we are about to do.

- **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.



Introduction to Jupyter Notebook

- **HTML & Markdown**

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

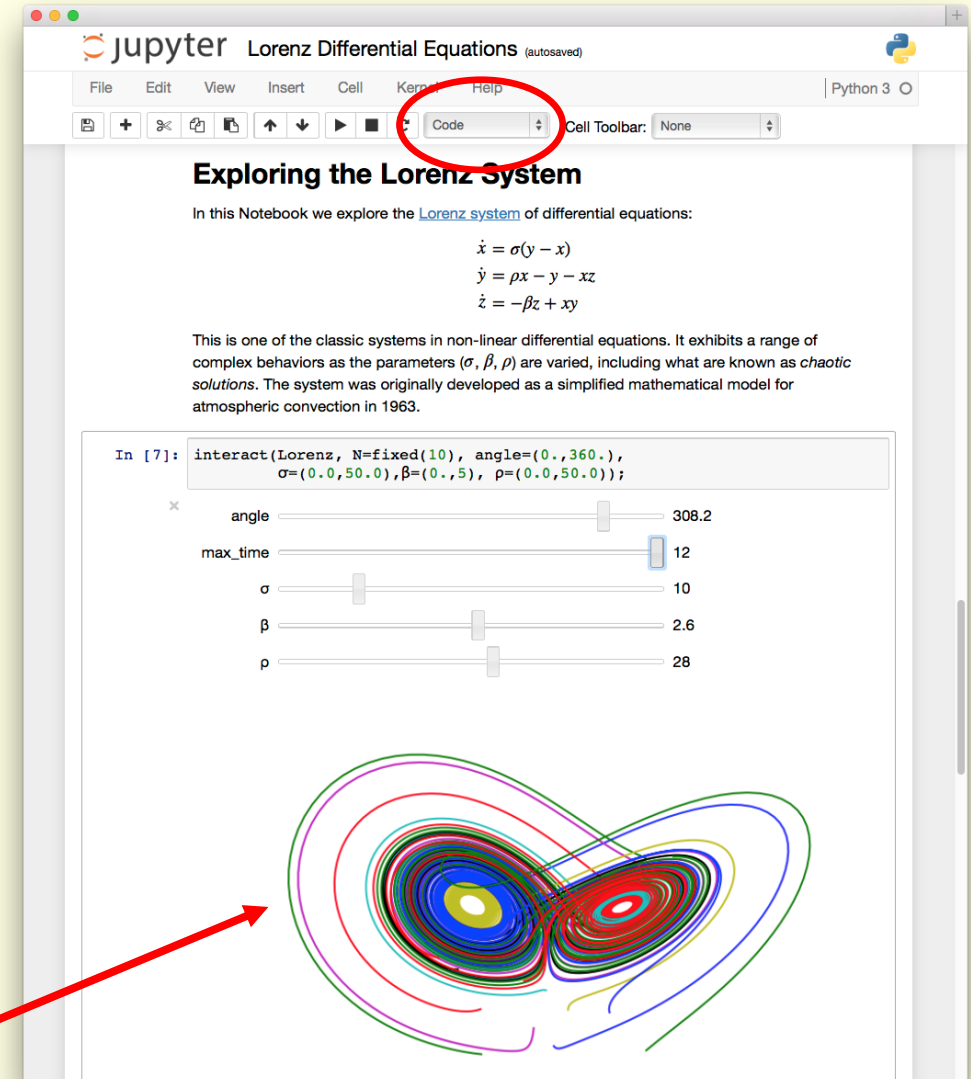
- **LaTeX (equations)**

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()

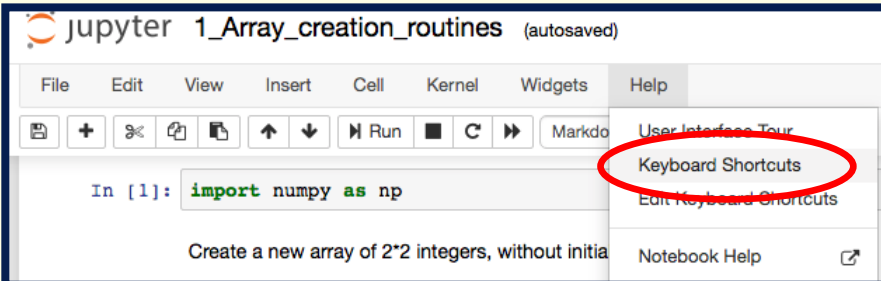
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3738 entries, 0 to 3737
Data columns (total 2 columns):
ID      3738 non-null object
R/L id   3738 non-null int32
dtypes: int32(1), object(1)
memory usage: 43.9+ KB
```

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

IPython Notebook: An Overview

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.

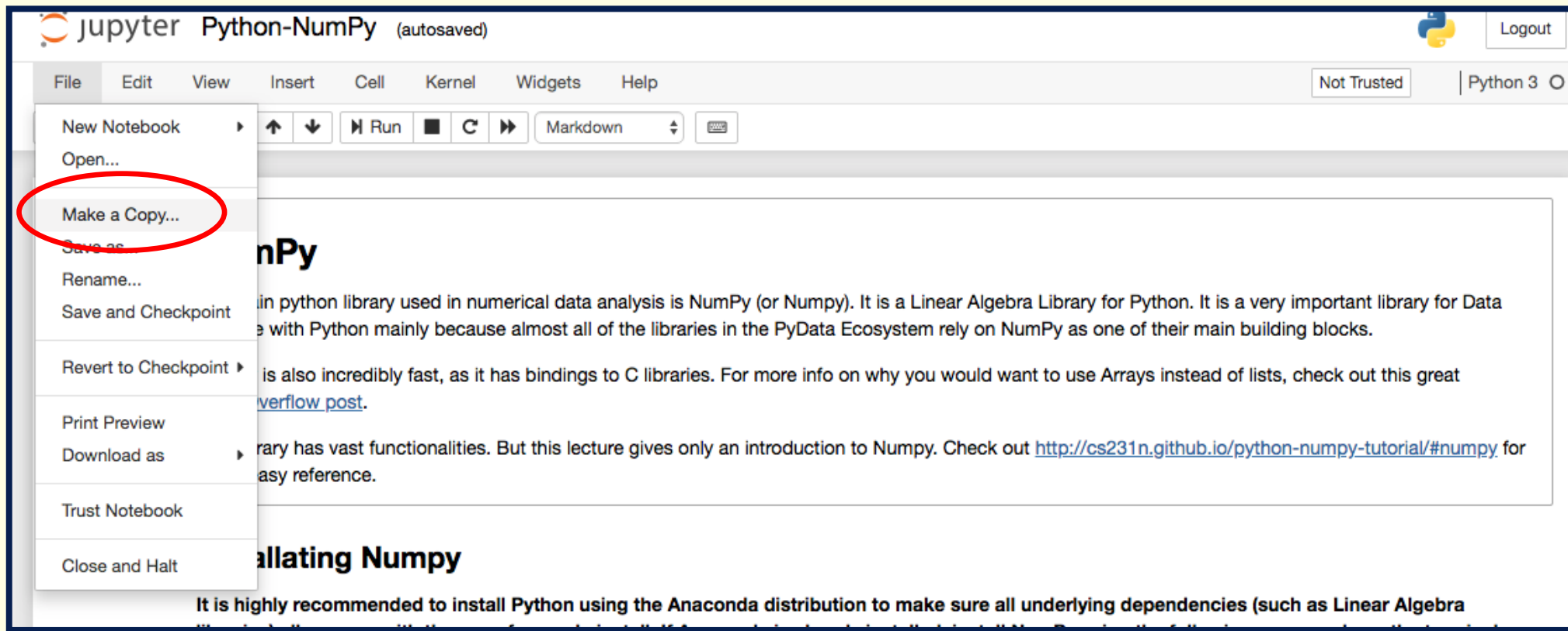
Command Mode (press `ESC` to enable)

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

Keyboard shortcuts can be viewed from Help → Keyboard Shortcuts

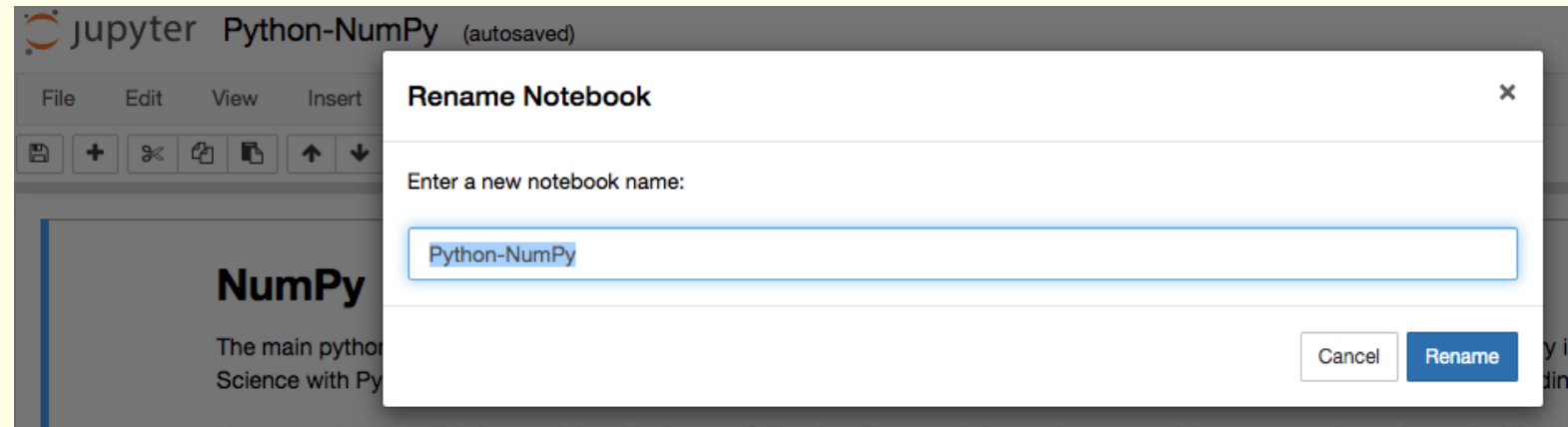
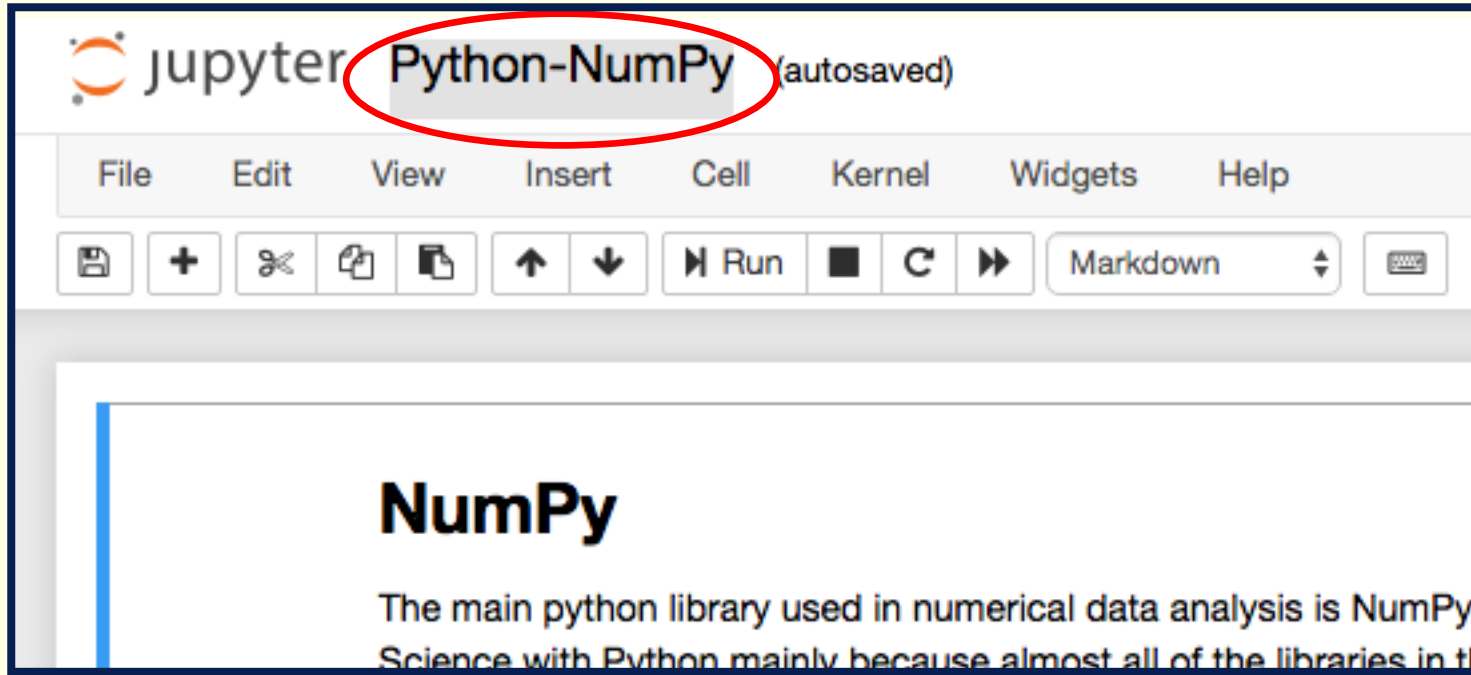
Introduction to Jupyter Notebook

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:

The sell for Typing your code :

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

2.png

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

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

In []:

In [3]:

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