



BABD

INTERNATIONAL MASTER IN BUSINESS ANALYTICS AND BIG DATA

Programming Tools



POLITECNICO DI MILANO
GRADUATE SCHOOL
OF BUSINESS



Executive Education
Ranking 2019



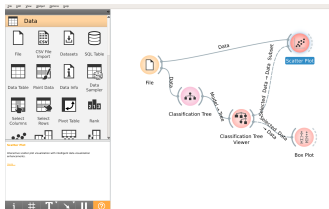
European Business Schools
Ranking 2018



Programming Tools

1. Orange

<https://orange.biolab.si/>



- ▶ Intuitive interface
- ▶ Fast development

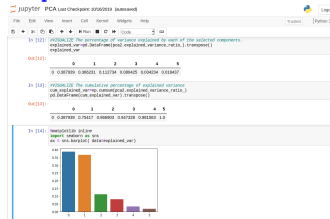


A library featuring various ML algorithms designed to inter-operate with the Python numerical and scientific libraries e.g. NumPy, Pandas.

<https://scikit-learn.org/stable/>

2. Jupyter-Notebook (Anaconda)

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



- ▶ Advanced functions
- ▶ Customization

Zero-based indexed



Three programmers walk into a bar...

BABD

List indexes

Forward Indexing

[1:4]

0	1	2	3	4
a	b	c	d	e
-5	-4	-3	-2	-1

Backward Indexing

[-5:-3]

included [start:end] excluded

Exercise 1. Functions

Create a function that given two integers m, n computes the greatest common divisor between m and n .

Exercise 2. Numpy

1. Create a random array of length 100.
Hint: `np.random.rand()`
2. Sort your array.
3. Compute the mean, median and sample variance.

Exercise 3. Pandas

1. Import the dataset *iris* as a DataFrame
2. Add the columns names (sepal length, sepal width, petal length, petal width).
3. Create a new column that contains the ratio between the sepal and petal length.
4. Add a new column named `target` with value 1 if the type is *setosa* and 0 otherwise.