

## Homework 1

### Python Essentials

In the Python Essentials module are attached three files, each corresponding to a NumPy array. The file “breast\_data.npy” contains a data matrix where each row corresponds to various data points of a breast cancer tumor. The file “column\_names.npy” contains an array of characters with a length equal to the number of columns in the previous matrix; these are the names of the columns (tumor features). The file “breast\_labels.npy” is a binary array that specifies whether a row in the matrix (i.e., a tumor) is benign (0) or malignant (1). The task is as follows:

1. **(40% of the grade)** Load these data into a Jupyter notebook and convert them into a DataFrame so that the columns have the corresponding names. Add the column from the “breast\_labels.npy” data to the DataFrame, but instead of 0 or 1, it should display “Benign” and “Malignant,” respectively.
2. **(20% of the grade)** Calculate the maximum, minimum, and average value of each tumor feature, but only for the Benign tumors and separately for the Malignant tumors.
3. **(40% of the grade)** Choose two features of the tumors and create a graph with one feature on the X-axis and the other on the Y-axis, plotting the data points where those corresponding to malignant tumors are one color, and those for benign tumors are another. Create another graph, different from the first, that you find interesting to represent the data. Remember that the graphs should have at least a title, axis labels, and legends identifying the points.

#### Conditions:

- All codes should be submitted, and each part of each item must be working. If a code does not work, you will receive a score of 0 for that item.
- All codes will be tested in Google Colab.
- Do an organized notebook, with comments in each part of it.