# Factorial discrimant analysis with Python and scikit learn

## Exercise 1 : Factorial discrimant analysis of synthetic data

1. Generate two dataset of 3D 500 vectors, following a standard Gaussian distribution and transform these two datasets :
   ```
   >>> import numpy as np
   >>> s1 = np.array([[3,0,0],[0,1,0],[0,0,0.01]])
   >>> r1 = np.array([[0.36,0.48,-0.8],[-0.8,0.6,0],[0.48,0.64,0.6]])
   >>> rndn3d1 = np.random.randn(500,3)
   >>> rndef1 = rndn3d1.dot(s1).dot(r1)
   >>> rndn3d2 = np.random.randn(500,3)
   >>> rndef2 = rndn3d2.dot(s1).dot(r1) + [0, 0, 1]
   >>> rndef = np.concatenate((rndef1, rndef2))
   ```

2. Generate the corresponding labels : 1 for the `rndef1` data, 2 for those of `rndef2` and visualize the data

3. Apply factorial discriminant analysis to these data and visualize the data

4. (a) Reduce the dimension of data with PCA with two components

   (b) Apply the factorial discrimant analysis to the reduced data of the previous question and visualize. Compare the two approaches

## Exercise 2 : Factorial discrimant analysis of the dataset `leaf` with scikit-learn

We apply factorial discriminant analysis on the data : `https://archive.ics.uci.edu/ml/datasets/Leaf`

1. Import the data.

2. Apply factorial discriminant analysis

3. Compare the results with PCA.