Fundamentals of Data Science
Homework/Laboratory session #8

This week you will create simple regression models. Please, create Python scripts solving Problems 1-3. You can use any environment, Spyder, Colab, Jupyter Notebook etc (although, Spyder is preferred).

Problem 1

You are given a data file, `electrons.csv`. This two-column file can be inspected using any basic text editing software, such as Gedit or Notepad.

The data show observations of electrons accelerated in a uniform electric field. The left column shows the observed path length of a particle in the electric field, the right column shows the energy it has after leaving the electric field.

The energy of an electron accelerated in a uniform electric field can be calculated as
$U_a = U_0 + |q|\, E\, S$
where U_a and U_0 are the electron energy before and after acceleration, respectively, $q = 1.6 \cdot 10^{-19}$ C is the electron charge, E is electric field strength, and S is the particle path in the electric field.

Use linear regression to fit the data and evaluate the electric field strength and average electron energy before acceleration.

Problem 2

You are given a data file, `gauss2d.csv`, which is a two-column file containing X and Y coordinates of around 400 data points. Find the principal components.