Lettuce Spectral Analysis

**Data description:** 232 reflectance spectral signatures collected with the CropSign hardware (AS7341 spectral sensor and a white LED as light source). The sensor includes eight VIS channels (415, 445, 480, 515, 555, 590, 630, 680 nm) and a NIR channel (910 nm). The data was collected at seven moments (from 1 to 7) during the growing cycle of two varieties of lettuce (TA and TB) subjected to different nitrogen fertilisation. Four levels were defined: zero, where the plants grew in the peat substrate with no fertiliser added, two, four and eight where 200, 400 and 800 mg of nitrogen (urea fertiliser) were added to the pot, respectively.

**Exploratory data analysis:** By analysing the average of the signatures for each level (shown in the image below), it can be said that, in general, the TA variety has higher reflectance intensities. This can be explained by the green colour of its leaves, which contrasts with the purple colour of the leaves of the TB variety, which may cause higher levels of light absorption.

Uma imagem com diagrama, texto, file, Gráfico

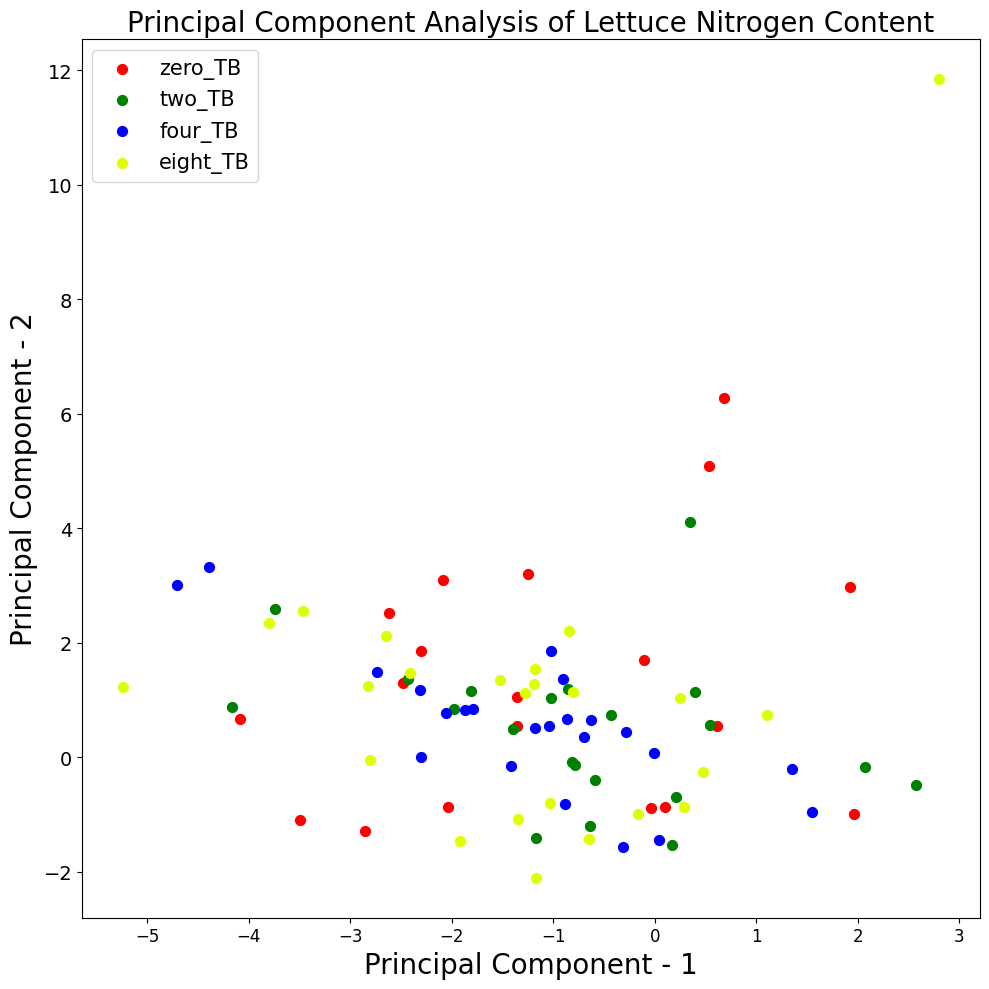
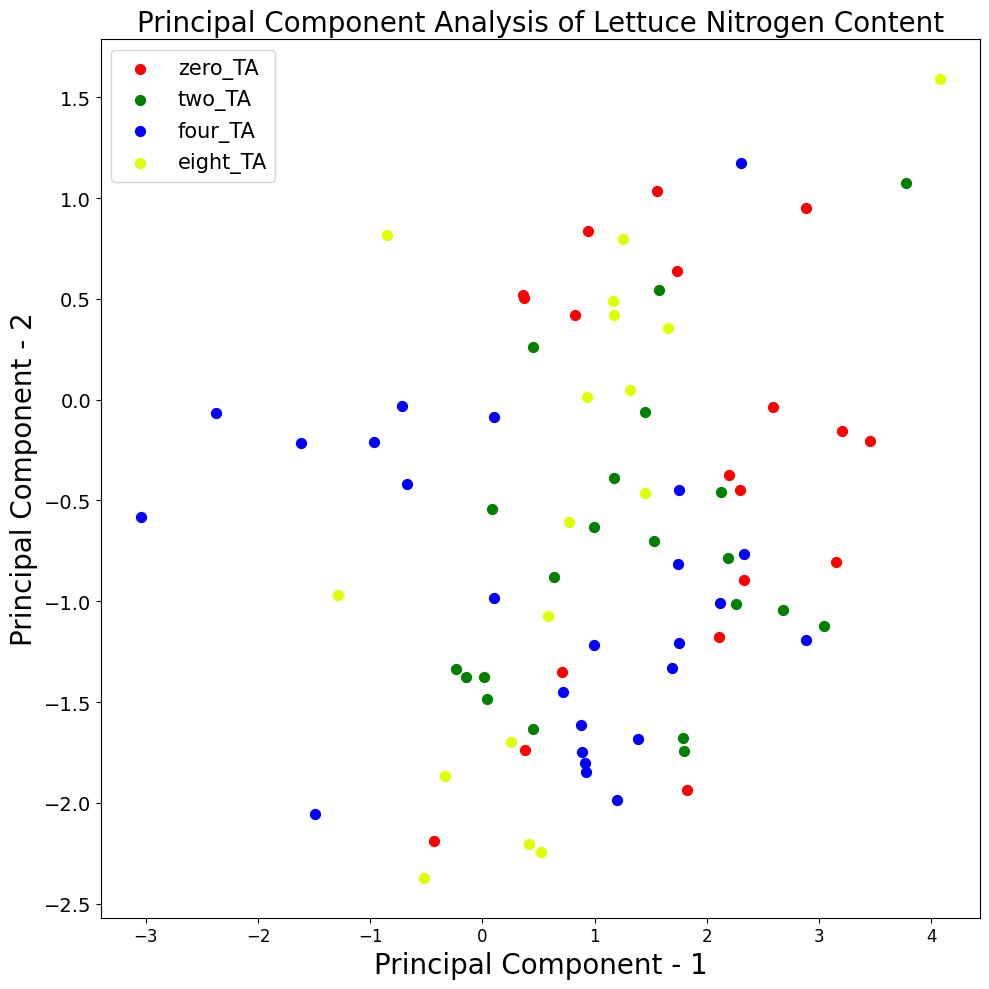
Descrição gerada automaticamente

**Dimensionality reduction:** without any pre-processing the complete dataset was given as input to a PCA (Principal Component Analysis). The result is shown in the figure below.

Uma imagem com texto, captura de ecrã, Saturação de cores, diagrama

Descrição gerada automaticamente

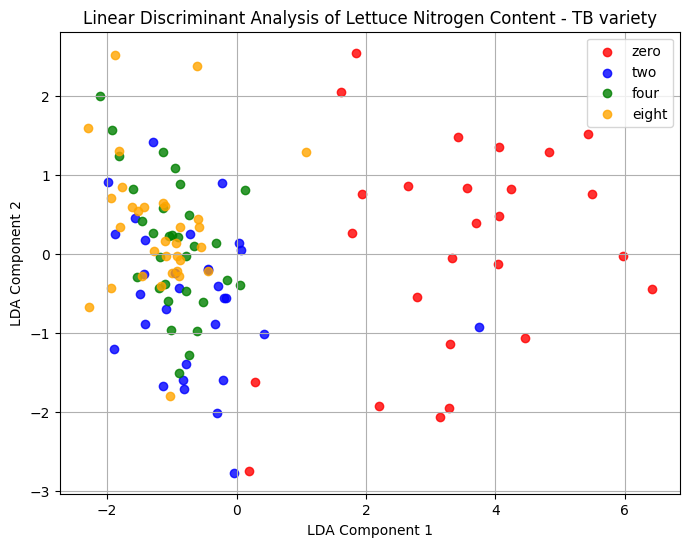
Next, the data of each variety was analysed individually, as shown in the following image.



In both cases, it is not possible to correctly separate the levels under study. Therefore, the performance of another technique, LDA (Linear Discriminant Analysis), was tested. LDA is a supervised technique and aims to maximize the separation between different classes in the data. So, can be more powerful in our context of classification of lettuce nutritional status.

Uma imagem com texto, captura de ecrã, diagrama, Gráfico

Descrição gerada automaticamente



In this case, and particularly for the TB variety, the LDA was able to group the “zero” elements. Bearing in mind that this analysis considers the values for different dates, this may be compromising the analysis, as the level of nitrogen absorption is variable throughout the crop cycle. The following figures show the result of the LDA applied only to the data collected at the first moment, one week after the lettuce plants were transplanted to the corresponding pots. To improve the results and apply the data to predictive models, such as SVM, it may be necessary to apply normalization filters (e.g. MSC - Multiplicative signal correction or Savitzky–Golay filter) to the data and enhance the dataset.

