The order of writing will be the following:

Starting with the big picture, a dendrogram analysis was performed from the whole dataframe obtaining the following plots:

Clustering optimal selection

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Dendrogram of the samples

Gráfico, Gráfico de proyección solar

Descripción generada automáticamente

Afterwards a PCA analysis was performed followed by a clustering techniques (k-means) applied on top of the PCA and will be explained here in the with the following images

Scree plot to track dimension to work on

Gráfico, Histograma

Descripción generada automáticamente

Individuals PCA representation

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Variables PCA representation

Gráfico, Gráfico radial

Descripción generada automáticamente

Biplot representation

Diagrama

Descripción generada automáticamente

After observation of individual and variables the analysis is completed with a k-means in order to select the optimal number of clusters based on their dissimilarities

Patrón de fondo

Descripción generada automáticamente

Silhouette optimal clusters

Gráfico, Gráfico de líneas

Descripción generada automáticamente

Clusterization of the samples

Gráfico

Descripción generada automáticamente con confianza media

Basically cluster 1= Fungi, Cluster 2= bacteria in bean and Cluster 3 bacteria in bean + rice

Visualization of the relevant variables per cluster

Gráfico, Gráfico de cajas y bigotes

Descripción generada automáticamente

At this point after understanding the best performance of Fungi, a zoom in is performed in the variables Rafinose and Estaquiose, due to their negative importance and wide spread, therefore, single variable is graphed and analyzed statistically through ANOVA and pos-hoc TukeyHSD

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Gráfico de dispersión

Descripción generada automáticamente con confianza media

All the methodology and discussion will be described here. All the code is attached and will be available on the repository along the data frame and metadata file.