

Problem n.1

The file `nutrients.txt` reports the nutrients (energy, proteins, fats, carbohydrates, sugars and fibers) of 295 different breakfast cereals. Each statistical unit corresponds to a different brand and kind of cereals.

- a) Decide whether it is appropriate to use the original variables or the standardized ones, and perform a Principal Component Analysis of the dataset. Report the loadings of the first principal component.
- b) Report a plot of the loadings of the first 3 principal components and interpret them.
- c) Report the biplot of the data along the first two principal components and describe what characterizes the cereals with positive scores for both the first and second components.
- d) Report the screeplot. Propose (with motivations) a dimensionality reduction of the dataset. Report the variance explained along the principal components retained.
- e) Project a new cereal with 400kcal, 9g of proteins, 5g of fats, 100g of carbohydrates, 30g of sugar and 4g of fiber, on the reduced space identified at point (d).

Upload your results here:

<https://forms.office.com/Pages/ResponsePage.aspx?id=K3EXCvNtXUKAjjCd8ope6-9AS0GWf2lHjvGX24HiqFVUNUhJQ0tDWkgyNOFTNFVHNEhRUDVaR0xCUS4u>