Description of the Candies database

From Luciano & Næs (2009) and Martin & Govaerts (2019)

This article presents the Candies database originally published in Luciano & Næs (2009) and Liland *et al.* (2018). It has been used in this thesis in Chapter 5 and in Martin & Govaerts (2019).

Statistical experimental design

Three factors of variability were considered in the experimental design:

1. The Assessors: a panel of 11 judges are rating the candies.
2. The Candies: 5 different candies are rated with respect to a series of 9 attributes: transparency, acidity, sweet taste, raspberry flavour, sugar coated texture tested with a spoon, biting strength in the mouth, hardness, elasticity in the mouth and stick to teeth in the mouth.
3. The replicates: Each candy was evaluated in triplicate by each assessor.

The final Candies database

This dataset is available in the *Candies.RData* data file where the *design* data.frame corresponds to the experimental design matrix with variables *Judges* and *Candies*, and the *outcomes* matrix corresponds to the sensory attributes evaluated by the judges.

The resulting balanced design is presented here in an unfolded data table with $165 = 11 \times 5 \times 3$ observations and 9 variables (attributes). The observations are identified by 4 digits as: judge-candy-replicate, *e.g.* 0712 corresponds to judge 7, candy 1 and replicate 2.

In Figure 1, the raw values of the 9 different attributes are displayed and colour-coded according to the candies. Clearly, distinct patterns are appearing: candies {5, 1} and {2, 3, 4} already seem to share similar attribute rating profiles.

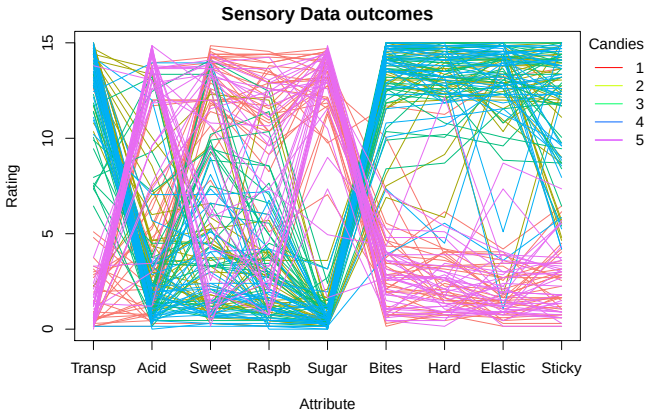

Figure 1: *Plot of the sensory attributes matrix.*

References

Liland, K. H., Smilde, A., Marini, F., & Næs, T. 2018. Confidence ellipsoids for ASCA models based on multivariate regression theory. *Journal of Chemometrics*, **32**(5), e2990.

Luciano, G., & Næs, T. 2009. Interpreting sensory data by combining principal component analysis and analysis of variance. *Food Quality and Preference*, **20**(3), 167 – 175.

Martin, M., & Govaerts, B. 2019. *LiMM-PCA : combining ASCA$^+$ and linear mixed models to analyse high dimensional designed data.* Discussion Paper DP 2019/21. Institute of Statistics, Biostatistics and Actuarial Sciences, UCLouvain, Belgium.