1 Assignment

Instructions

Submit your answers to the four first exercises through Aula Global. **Remember to round longer decimals to the second place after the comma**. The remainder of the assignment concerns self-studies.

Descriptive statistics

Imagine you measured the (average) voice pitch of 15 Korean speakers (in Hz)[1]. The data looks as follows:

208, 158, 153.3, 239.7, 102.9, 83.1, 231.9, 285.1, 181.2, 266, 252.5, 185.5, 110.7, 236.8, 207.5

Codified as a vector, called `pitch`, in R:

```r
pitch <- c(208, 158, 153.3, 239.7, 102.9,
           83.1, 231.9, 285.1, 181.2, 266,
           252.5, 185.5, 110.7, 236.8, 207.5)
```

Report the following summaries of this sample

1. Mean
2. Median
3. Variance
4. Standard deviation

Self-study

1. Does the mean give a good summary of this sample? Why (not)?
2. What is suggested by the mean and median leading to different summaries?
3. Work through Chapter 1, Chapter 2 (until the end of subsection 2.2) and Chapter 3 of Introduction to Data Analysis (Franke 2021);
4. Make sure you have a working R environment: either locally –on your computer– or by having made sure you can execute code in a "colab" document
5. Look at the guidelines for the analysis report and read the illustrative report. Think of a general question about language that interests you and write it down.

References

Franke, Michael. 2021. *An Introduction to Data Analysis.*

Winter, Bodo, and Sven Grawunder. 2012. "The Phonetic Profile of Korean Formal and Informal Speech Registers." *Journal of Phonetics* 40 (6): 808–15. https://doi.org/10.1016/j.wocn.2012.08.006.

---

[1] This data was randomly sampled from the pre-processed and shortened version of the study of Winter and Grawunder (2012), as provided in Franke (2021). The data is accessible here.