**« How to properly analyze quantitative data and document this work »**

**Let’s the doctoral students talk:** *« I don’t know how to analyze my data »*

**When:** This document must accompany you when you start processing your data.

**Why:** It documents how you will validate, pre-process and process your data.

**Inputs for the thesis writing:** It allows you to write the data processing and results part. It also allows you to write the results part of a research paper.

*Platforms such as Jupyter can be useful to provide this documentation.*

**Objectives of treatment**

* Can you recall the objectives of your experimentation?

* What questions/hypotheses does your data processing have to answer?

**Literature method**

* What methods of analysis are used in the literature? *(e.g. visualization, ANOVA, modeling, thematic analysis, etc.)*

* What are the tools, software used in the literature? *(e.g. R, SPSS, python, CAQDAS, etc.)*

* Why are you going to use these methods of analysis?

* If not, why don’t you use these methods of analysis?

* Finally, what methods and software will you use?

**Data pre-processing, some tips**

The **pre-processing of data** makes it possible to check the quality of the data before processing. For this, it is recommended to study the distribution of the different variables:

* Calculation of frequencies, averages, medians, standard deviations, min, max, quartiles.
* Make histograms and box plots for numerical variables in order to detect outliers.
* Make bar charts for nominal and ordinal variables to detect outliers.

For work on **data quality**, use this book: Di Ruocco, & al. “la qualité des données concepts de base et techniques d’amélioration” in L.Berti Equille (dri.) la qualité et la gouvernance des données, Hermès-Lavoisier, Paris 2012

* Describe the pre-treatment you will be doing.

* Write what you observe about these pre-treatments. (There is always a lot of information in these pre-treatments)

* If you observed any outliers, what did you do with them? Why?
* Did these pre-treatments enrich your research questions?

* Did these pre-treatments allow you to create new variables to anwser your research questions *(e.g. combining variables, recoding, etc.)*? Which ones?

**Data processing**

* Describe the treatment your are going to do? *(e.g. methods, software, language, etc.)*

* What are the bibliographical references on these methods and software?

* What are the graphical representations that will make your data talk?

**Analyses and results**

* Identify the data processing that you will retain to answer your research question?

* Write down what you observe about these processing operations.

* How do these processing operations do or do not answer your research questions?

**Limits of the data processing**

* Indicate the limits of this data processing.

* Indicate how it could have been improved.