



---

### EXERCISES SESSION 3

The purpose of this lab is to become familiar with the different datasets that will be used during this course and to learn how to deal with missing values.

#### Datasets

The datasets that will be used during the course are the following:

- Transportation Mode Choice in the Netherlands
- Residential Telephone Services
- Swissmetro
- Airline Itinerary Choice (Boeing)
- Mode choice in Switzerland (Optima)

For each of them perform the following steps:

1. Download the dataset and the dataset description that can be found in the website.
2. Read carefully the dataset descriptions.
3. Open the dataset with a text editor and copy its contents in Excel.
4. Go through all the columns to make sure that you understand what they represent.
5. Choose randomly a few observations (each row corresponds to one observation) and study them in detail. Does the choice of this respondent correspond to your expectations? Are there any missing values? How are they coded?

#### Missing values

From the previous exercise you have seen that there are often missing values in the observations. It is important to model this accordingly. Go over the presentation and apply what is proposed in the model files from Exercise session 2.