LAB SESSION 1

The objective of this first lab session is to get familiar with the datasets that will be used during the course. To do so, we recommend to perform the following tasks:

1. Go through the description of the case studies that will be used during the lab sessions. You will find the corresponding .pdf files for each case study under **Datasets descriptions** section of the Laboratories tab. Each file contains the following information:

   - description of the choice context,
   - methodology of the data collection,
   - description of the available data,
   - statistics on certain data variables.

2. Understand the structure of the data file (.dat) and compute descriptive statistics for the Netherlands dataset. *Note* that the Netherlands dataset contains records for both RP and SP choices.

   For this task, open the netherlands.dat file using excel and answer the following:

   (a) What does each row of this file correspond to?

   (b) What does each column of this file correspond to?

   (c) Identify which variables are related to (i) attributes of the alternatives and (ii) characteristics of the decision makers. How are the variables related to the characteristics of the decision makers coded?

   (d) Compute the mean, mode, standard deviation, min and max for each variable.

   (e) Visualize the qualitative variables (e.g. gender) by generating charts.

   (f) Visualize the attributes of the alternatives (e.g. travel time) using histograms.

   (g) According to your expectations about the variables that would be the most important in explaining the mode choice, compute the mode shares for specific groups (e.g. age, gender, etc.).

   (h) Investigate correlation between the independent variables.

   (i) Use scatter plots to investigate bias towards specific alternatives. One idea is to plot the travel time of the rail alternative against the travel time of the car alternative for those who chose car/ rail.