File DATOSCLIN.XLS contains the data of a sample of patients collected in a North American Hospital in the decade of the 70s. For each patient, data were taken at the time of admission and at the exit, either by death or discharge of the patient. The state of survival was also observed of the patient. The variables are described in the following table. The variables whose name ends with "1" refer to the data obtained in the initial examination and the variables whose name ends in \ 2 "refer to the data obtained in the final exam.

IDNUM	Patient ID
EDAD	Age (years)
ALTURA	Height (cm)
SEXO	Sex (1=man, 2=woman)
SUPERVIVENCIA	Survival (1=lives, 2=dies)
TIPOSHOCKE	Shock type (2=No shock,, 3=Hypovolemic, 4=Cardiogenic, 5= bacterial, 6=Neurogenic, 7=Others)
SBP1	Systolic Pressure at start (mm Hg)
MAP1	Mean arterial pressure at start (mm Hg)
HEART1	Beats per minute at start
CARDIAC1	Cardiac Index at start
CTIEMPO1	Mean Urine circulation time at start (secs.)
URINA1	Urine at start (ml/hr)
HGB1	Hemoglobin at start (gm)
SBP2	Systolic Pressure at end (mm Hg)
MAP2	Mean arterial pressure at end (mm Hg)
HEART2	Beats per minute at end
CARDIAC2	Cardiac index at end
CTIEMPO2	Mean Urine circulation time at end (secs.)
ORINA2	Urine at end (ml/hr)
HGB2	Hemoglobin at end (gm)

We will start reading and preparing the database.
1. The basic data for the study can be found in the Excel file DATOSCLIN.XLS. You can read the data directly from Rcommander (in Windows).
2. The first thing we will do is to take a sample of the data, so that each of you works with a different dataset.
   a. To do this create a new database extracting 60 individuals randomly from the 113 contained in the database (this can be get a call to the sample instruction of R).
3. Carry out a brief descriptive study of the variables in the file.
4. Make a graphic and numeric description of every one of them. It does not have to be exhaustive but simply provide an overview of the composition of the dataset.
5. Recode the TIPOSHOCK variable in a new SHOCKBIN that is 0 if there has been no Shock and 1 otherwise.

Next we do a power analysis
1. Calculate the number of individuals that we should study to be able to estimate the percentage of patients with Shock with an error lower than 5%.
2. If, as in this case, we have decided to extract a sample of size 60, which is the precision with we are estimating the percentage of patients with Shock?

Last we will make some comparisons.
1. Perform a normality test to decide whether the variables "MAP1", HEART1 and CARDIAC1 can be considered to follow a normal distribution (start with some plots). Do it also with their counterparts at the final examination (values ended by "2").
2. Compare the mean values of the three previous variables between individuals with Shock and those who did not present this. Use a parametric test and one non-parametric test and compare the results.
3. Compare the values of the three previous variables, corresponding to the initial time with the respective value corresponding to the instant of each patient. Use a parametric test and a non-parametric test and compare the results obtained.
4. Study whether the TYPESHOCK and SURVIVAL variables are independent. Do the same for variables SHOCKBIN and SURVIVAL.