

Assignment 2 Starting with R

1. Load the `biofam` data set that comes with the TraMineR library.
2. Print the variable names.
3. Create an `age` variable by subtracting the birth year from the year of the survey and add it to the `biofam` data frame.
4. What is the minimum, maximum, median and mean age in the sample?
5. What is the minimum, maximum, median and mean age of the women?
6. Add a `cohort` factor to the `biofam` data frame grouping the birth years into the following categories: 1900-1929, 1930-1939, 1940-1949, 1950-1959.
7. Generate an histogram of the distribution of `birthyear` using the above birth year classes. (Look at the help of the `hist` function for how to do that.)
8. Produce a frequency table of the `cohort` factor.
9. Cross tabulate the cohort with the state at 25 years old.
10. Fit a logistic regression for the probability to be married with a child and having left home at 25 years old in terms of the language of the questionnaire and the sex. Comment the results.
11. Fit the same logistic regression, but for the youngest cohort only.