

## Telecommunication Churn Customers

The project aims to exploit the problem of churn of customers from a telecommunications company. The data set contains observations from the portfolio of customers of a telecommunications company, with some detail on the usage during the previous period. There is also some demographic information.

The aim of the project is to check which variables are related to churn. Important: we do not care at this project for prediction.

Things that you need to consider are the following (but clearly not restricted to them)

- Which variables are important?
- Do we need to transform them?
- How good is the model?
- Are there any assumptions that need to be checked carefully?

You have to write a detailed enough report to explain to your boss about that, explaining the model you used together with sufficient technical details on what you have done. Use tables and plots that really contribute to your story. Explain what other information you may need. It is important that the report is self-contained and if somebody reads it could follow the arguments even with limited statistical knowledge.

You should not

- a. Copy and paste R output
- b. Discuss the ability to predict with that model, this is the next project. There will be negative grading if you do this now.

You need to upload 2 files and not one zip file

1. The word/Latex or pdf file with your report.
2. The R-code used as a separate txt file

Deadline Sunday 19/02/2023, 23.59. After that date you will have -2 (out of 10) points for every day of delay