Final task

You have a dataset in two csv file about life insurance contracts. Try to analyze the data among the given tasks. The detailed list of variable is given in the task. You have 3 hours to finish the analysis and to conclude your main results in a presentation.

a) Merge the 2 csv files in one dataset!

b) Create a training and test set! Use the Event_NP variable as dependent variable. Try to predict which contracts are cancelled due to overdue fees! (Event_NP value is 1 in that cases) In the presentation, highlight the efficiency of the model! If you use different methods, compare them to each other! List explanatory variables that are not included in the dataset, but are considered important for the estimation!

c) Filter the dataset to only the contracts that are cancelled due to overdue fees. Try to build a model by which you can determine the factors of „survival time" of these contracts. (Dependent variable: Time) In the presentation, highlight the method/way of model selection and identify the significant variables!

d) Conclude the results in a presentation. We also need the way of calculation to check the validity of results!

Variables:

Time	period measured in months from the signing of the contract
Event_NP	1, if cancelled due to overdue fees; 0 otherwise

Explanatory:

TARTAM_EV	Duration of the contract in years
UzletiKedv	Rate of business discount (%), contracting cost reduced
BELEPESI_KOR	Contractor's entry age in years
KEZD_BOSSZEG	Sum insured at the time of contracting
MEGYE	County of residence of the contractor
REGIO	Region of residence of the contractor
ANNUAL_PREMIUM	Initial annual premium for the contract
NATIONALITY	Nationality of the contractor
KULFOLDI_E	Is the contractor a foreign national (1: yes, 0: no)
SZERZO_KORA	How many years has the agent been in business with the insurer?
ERT_CSAT	Through which sales channel did the insurer conclude the contract?
HEALTH_CATEGORY	Health category of the contractor (standard, priority)
HITELFEDZETI	Is the contract taken out as credit insurance (1: yes, 0: no)
CEGES_E	Was the contract concluded with a legal entity (1: yes, 0: no)
BALESETI_ROKK	Is the contract covered by supplementary insurance for accidental disability (1: yes, 0: no)
BALESETI_HALAL	Is the contract covered by supplementary accidental death insurance (1: yes, 0: no)
BALESETI_KORHAZ	Is supplementary insurance against accidental hospitalisation included in the contract (1: yes, 0: no)?
BALESETI_MUTET	Is supplementary insurance for accidental surgical compensation included in the contract (1: yes, 0: no)?
MUTET	Is additional insurance for surgical compensation included in the contract (1: yes, 0: no)
CRIT_ILLNESS	Is supplementary cover for critical illness included in the contract (1: yes, 0: no)
EUG_KIEG	Is other supplementary health insurance included in the contract (1: yes, 0: no)?
KOZLEKEDESI	Is supplementary insurance for traffic accidents included in the contract (1: yes, 0: no)
MUTET_DIJ	Annual premium for supplementary insurance for surgical compensation
CRIT_ILLNESS_DIJ	Annual premium for supplementary insurance for critical illness
EUG_KIEG_DIJ	Annual premium for other supplementary health insurance
KOZLEKEDESI_DIJ	Annual premium for supplementary insurance for traffic accidents

NOI_EUG_DIJ	Annual premium for supplementary health insurance for women
OTHER_LIFE	Does the policyholder have other life insurance with the insurer?
OTHER_NON_LIFE	Does the policyholder have other non-life insurance with the insurer?
TOROLT_LIFE	Does the policyholder have any other cancelled life insurance with the insurer?
TOROLT_NON_LIFE	Did the policyholder have other cancelled non-life insurance with the insurer?