Credit risk game

- Team assignment.
- Upload your decision(s) before 11:55PM (EDT) on December 19th, 2022.

Context:

In order to minimize loss and maximize profits from lending, banks require careful assessment of their borrowers. This has led to credit risk computation becoming a commonly used application of statistics and data science. You are working with a large bank to help them optimize their profits from financing individuals who apply for a loan. The bank has provided you with the records and results of lending money to some customers. You will be using this information and R to decide how many and which individuals to lend money to. Your objective is to maximize the banks' profits obtained from these loans.

Data:

A zip file containing the data is available on ZoneCours. The file *CreditGame_TRAIN.csv* contains records of previous loans provided by the bank. They include features about the customers and information about their default status as well as the magnitude of profit or loss incurred for each loan. The file *CreditGame_Applications.csv* contains information about current loan applications. You must decide who on this list get their loan approved.

Method:

Your objective is based on a business outcome: profit. The process of prediction will involve cleaning, analyzing, modeling, and getting results. We do not give further instructions on the methods used; you are on your own for that. There is no single good answer and multiple strategies that can support the business problem reasonably well. We expect each student to come up with their own approach.

Disclaimer:

There are additional nuances for credit risk assessment in a real-life setting. Banks need to abide by the Basel accords and must comply with some rules to assess their credit risk. Credit risk typically implies a need to interpret the results of a model, and some standards in methodology apply. Although this business simulation is very realistic, both in terms of context and data, it does not depict those field-specific particularities.

Variables:

Variable name	Description
ID_TRAIN	Unique borrower ID
TYP_FIN	Type of funding requested (Car, Mortgage, or Credit)
NB_EMPT	Number of borrowers (borrower variables are only for the principal borrower)
R_ATD	Total Debt Amortization (TDA) Ratio, i.e., monthly financial commitments over monthly income
PRT_VAL	The requested loan amount over the value of the goods
DUREE	Requested loan duration
AGE_D	Age of the borrower
REV_BT	Gross Income
REV_NET	Net Income
TYP_RES	Residence Type – P: Owner, L: Tenant, A: Others
ST_EMPL	Employment Status – R: Regular, P: Part-Time, T: Self Employed
MNT_EPAR	Savings Value
NB_ER_6MS	Number of transactions refused due to insufficient funds in the last 6 months
NB_ER_12MS	Number of transactions refused due to insufficient funds in the last 12 months
NB_DEC_12MS	Number of overdrafts in the last 12 months
NB_OPER	Total number of transactions in record
NB_SATI	Total number of satisfactory transactions in record (No payment delay)
NB_COUR	Number of current transactions
NB_INTR_1M	Number of inquiries in the last month
NB_INTR_12M	Number of inquiries in the last 12 months
PIR_DEL	Worst current delinquency
NB_DEL_30	Number of 30–59 day delinquency in the last 12 months
NB_DEL_60	Number of 60–89 day delinquency in the last 12 months
NB_DEL_90	Number of 90+ day delinquency in the last 12 months
MNT_PASS	Value of financial Liabilities
MNT_ACT	Value of financial Assets
MNT_AUT_REN	Total authorized amount of revolving credit
MNT_UTIL_REN	Total used amount of revolving credit
MNT_DEMANDE	Loan amount requested
DEFAULT	Default is considered when payment is 90 days or more late within 24 months 1: Default, 0: Did not default
PROFIT_LOSS	Profit made or loss incurred with this loan after two years