

Application Dataset:

ID: client's identification

CODE_GENDER: male/female

FLAG_OWN_CAR: if client owns a car (Y/N)

FLAG_OWN_REALTY: if client owns property (Y/N)

CNT_CHILDREN: # of children

AMT_INCOME_TOTAL: yearly income

NAME_INCOME_TYPE: income category (working, commercial associate, pensioner, state servant, student)

NAME_EDUCATION_TYPE: level of education attained (secondary/special secondary, higher education, incomplete higher, lower secondary, academic degree)

NAME_FAMILY_STATUS: marital status (married, single/not married, civil marriage, separated, widow)

NAME_HOUSING_TYPE: type of housing (house/apartment, with parents, municipal apartment, rented apartment, office apartment, co-op apartment)

DAYS_BIRTH: days client has been alive (counting backwards)

DAYS_EMPLOYED: days client has been employed (counting backwards, positive value means unemployed)

FLAG_MOBIL: if client has a mobile phone (1/0)

FLAG_WORK_PHONE: if client has a work phone (1/0)

FLAG_PHONE: if client has a phone (1/0)

FLAG_EMAIL: if client has an email (1/0)

OCCUPATION_TYPE: type of job (not specified, laborers, core staff, sales staff, managers, drivers, high skill tech staff, accountants, medical staff, cooking staff, security staff, cleaning staff, private service staff, low-skill laborers, waiters/barmen staff, secretaries, HR staff, realty agents, IT staff)

CNT_FAM_MEMBERS: family size

Credit Dataset:

ID: client's identification

MONTHS_BALANCE: months on book (0 is current month, counting backwards)

STATUS: payment status (0: 1-29 days late, 1: 30-59 days late, 2: 60-89 days late, 3: 90-119 days late, 4: 120-149 days late, 5: writeoffs/150+ days late, C: on time, X: no loan for the month)

Final Modeling Dataset:

MONTHS_WITH_BALANCE: number of months client has used the card (standardized)

AMT_INCOME_TOTAL: yearly income (standardized)

CNT_FAM_MEMBERS: family size (standardized)

AGE: in years (standardized)

YEARS_EMPLOYED: years employed (standardized)

CODE_GENDER: dummified (excluded factor: female)

FLAG_OWN_CAR: dummified (excluded factor: no)

FLAG_OWN_REALTY: dummified (excluded factor: no)

NAME_EDUCATION_TYPE: dummified (excluded factor: academic degree)

NAME_INCOME_TYPE: dummified (excluded factor: commercial associate)

NAME_FAMILY_STATUS: dummified (excluded factor: civil marriage)

NAME_HOUSING_TYPE: dummified (excluded factor: co-op apartment)

FLAG_WORK_PHONE: dummified (excluded factor: 0)

FLAG_PHONE: dummified (excluded factor: 0)

FLAG_EMAIL: dummified (excluded factor: 0)

OCCUPATION_TYPE: dummified (excluded factor: accountants)

GOOD_BAD: whether or not a client is a good or bad credit card candidate (based on criteria in code. 0 for good, 1 for bad)