

# IDS702 Final Project: HomeCredit Default Risk Dataset

William Huang

November 28, 2018

## Project Objective

Home Credit Group, founded in Czech Republic and headquartered in Netherlands, is a multinational non-bank financial company that provides consumer financial products, such as personal lending and credit card businesses. The company focuses primarily on people with no or little credit history. The project objective is to interpret how the information on the loan applications could affect the default risks of loan applicants.

## Variable Descriptions

- **TARGET:** Target variable (1 - client with payment difficulties: he/she had late payment more than X days on at least one of the first Y installments of the loan in our sample, 0 - all other cases)
- **CODE\_GENDER:** Gender of the client
- **NAME\_CONTRACT\_TYPE:** Identification if loan is cash or revolving
- **AMT\_INCOME\_TOTAL:** Income of the client
- **FLAG\_OWN\_CAR:** Flag if the client owns a car
- **FLAG\_OWN\_REALTY:** Flag if client owns a house or flat
- **CNT\_CHILDREN:** Flag if client owns a house or flat
- **AMT\_CREDIT:** Credit amount of the loan
- **AMT\_ANNUITY:** Loan annuity
- **AMT\_GOODS\_PRICE:** For consumer loans it is the price of the goods for which the loan is given
- **NAME\_TYPE\_SUITE:** Who was accompanying client when he was applying for the loan
- **NAME\_INCOME\_TYPE:** Clients income type (businessman, working, maternity leave,...)
- **NAME\_EDUCATION\_TYPE:** Level of highest education the client achieved
- **NAME\_FAMILY\_STATUS:** Family status of the client
- **NAME\_HOUSING\_TYPE:** What is the housing situation of the client (renting, living with parents, ...)
- **REGION\_POPULATION\_RELATIVE:** Normalized population of region where client lives (higher number means the client lives in more populated region)
- **DAYS\_BIRTH:** Client's age in days at the time of application
- **DAYS\_EMPLOYED:** How many days before the application the person started current employment
- **OWN\_CAR\_AGE:** Age of client's car
- **OCCUPATION\_TYPE:** What kind of occupation does the client have
- **CNT\_FAM\_MEMBERS:** How many family members does client have
- **ORGANIZATION\_TYPE:** Type of organization where client works
- **AMT\_REQ\_CREDIT\_BUREAU\_HOUR:** Number of enquiries to Credit Bureau about the client one hour before application
- **AMT\_REQ\_CREDIT\_BUREAU\_DAY:** Number of enquiries to Credit Bureau about the client one day before application (excluding one hour before application)
- **AMT\_REQ\_CREDIT\_BUREAU\_WEEK:** Number of enquiries to Credit Bureau about the client one week before application (excluding one day before application)
- **AMT\_REQ\_CREDIT\_BUREAU\_MON:** Number of enquiries to Credit Bureau about the client one month before application (excluding one week before application)
- **AMT\_REQ\_CREDIT\_BUREAU\_QRT:** Number of enquiries to Credit Bureau about the client 3 month before application (excluding one month before application)
- **AMT\_REQ\_CREDIT\_BUREAU\_YEAR:** Number of enquiries to Credit Bureau about the client one day year (excluding last 3 months before application)

## Load Data and Data Pre-processing

As the original dataset contains over 200 independent variables, I picked 23 variables for further analysis and interpretations based on my interests and previous experience in this industry. Besides, I also combined the following 6 variables together to improve the interpretability of my model: `AMT_REQ_CREDIT_BUREAU_HOUR`, `AMT_REQ_CREDIT_BUREAU_DAY`, `AMT_REQ_CREDIT_BUREAU_WEEK`, `AMT_REQ_CREDIT_BUREAU_MON`, `AMT_REQ_CREDIT_BUREAU_QRT`, `AMT_REQ_CREDIT_BUREAU_YEAR`. These variables represent the number of enquiries to Credit Bureau about the client one hour, one day(excluding one hour before), one week(excluding one day before), one month(excluding one week before), one quarter(excluding one month before) and one year(excluding one quarter before) before application. As

my goal is to construct an interpretation model, instead of a predictive model, the integration of these variables could enhance the interpretability of the model substantially. On the other hand, if I intended to build a predictive model, I should have kept them as 6 different variables.

```
library(mice)
```

```
## Loading required package: lattice
```

```
##  
## Attaching package: 'mice'
```

```
## The following objects are masked from 'package:base':  
##  
##      cbind, rbind
```

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':  
##  
##      filter, lag
```

```
## The following objects are masked from 'package:base':  
##  
##      intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## — Attaching packages ————— tidyverse 1.2.1 —
```

```
## ✓ ggplot2 3.1.0      ✓ readr   1.1.1  
## ✓ tibble  1.4.2      ✓ purrr   0.2.5  
## ✓ tidyr   0.8.2      ✓ stringr 1.3.1  
## ✓ ggplot2 3.1.0      ✓ forcats 0.3.0
```

```
## — Conflicts ————— tidyverse_conflicts() —  
## ✖ tidyr::complete() masks mice::complete()  
## ✖ dplyr::filter()   masks stats::filter()  
## ✖ dplyr::lag()      masks stats::lag()
```

```
library(pROC)
```

```
## Type 'citation("pROC")' for a citation.
```

```
##  
## Attaching package: 'pROC'
```

```
## The following objects are masked from 'package:stats':  
##  
##      cov, smooth, var
```

```
library(arm)
```

```
## Loading required package: MASS
```

```
##  
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':  
##  
##      select
```

```
## Loading required package: Matrix
```

```
##  
## Attaching package: 'Matrix'
```

```
## The following object is masked from 'package:tidyr':  
##  
##      expand
```

```
## Loading required package: lme4
```

```
##  
## arm (Version 1.10-1, built: 2018-4-12)
```

```
## Working directory is /Users/Macintosh/IDS702HomeCredit
```

```
#setwd("/Users/wh132/Desktop")  
setwd("/Users/Macintosh/Desktop")  
application_data <- read.csv("application_train.csv")
```

```
#select interested variables  
application_data_selected <- subset(application_data, select = c("TARGET", "CODE_GENDER", "NAME_CONTRACT_TYPE", "AM  
T_INCOME_TOTAL", "FLAG_OWN_CAR", "FLAG_OWN_REALTY", "CNT_CHILDREN", "AMT_CREDIT", "AMT_ANNUITY", "AMT_GOODS_PRICE",  
"NAME_TYPE_SUITE", "NAME_INCOME_TYPE", "NAME_EDUCATION_TYPE", "NAME_FAMILY_STATUS", "NAME_HOUSING_TYPE", "REGION_POPU  
LATION_RELATIVE", "DAYS_BIRTH", "DAYS_EMPLOYED", "OWN_CAR_AGE", "OCCUPATION_TYPE", "CNT_FAM_MEMBERS", "ORGANIZATION_TY  
PE", "AMT_REQ_CREDIT_BUREAU_HOUR", "AMT_REQ_CREDIT_BUREAU_DAY", "AMT_REQ_CREDIT_BUREAU_WEEK", "AMT_REQ_CREDIT_BUREAU  
_MON", "AMT_REQ_CREDIT_BUREAU_QRT", "AMT_REQ_CREDIT_BUREAU_YEAR"))  
  
#combine all the AMT_REQ_CREDIT_BUREAU variables together to compute the aggregated number of enquiries to Credi  
t Bureau about the client one year before the application  
application_data_selected$AMR_REQ_CREDIT_BUREAU_SUM <- application_data_selected$AMT_REQ_CREDIT_BUREAU_HOUR+appl  
ication_data_selected$AMT_REQ_CREDIT_BUREAU_DAY+application_data_selected$AMT_REQ_CREDIT_BUREAU_MON+application  
data_selected$AMT_REQ_CREDIT_BUREAU_QRT+application_data_selected$AMT_REQ_CREDIT_BUREAU_WEEK+application_data_se  
lected$AMT_REQ_CREDIT_BUREAU_YEAR  
  
#Drop AMT_REQ_CREDIT_BUREAU_HOUR, DAY, WEEK, MON, QRT, YEAR  
application_data_selected <- application_data_selected %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_HOUR)) %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_DAY)) %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_WEEK)) %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_MON)) %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_QRT)) %>%  
  dplyr::select(-(AMT_REQ_CREDIT_BUREAU_YEAR))
```

Based on the correlation table above, two pairs of independent variables have extremely high correlations:

- 1. (AMT\_CREDIT, AMT\_GOODS\_PRICE), corr = 0.987 AMT\_GOODS\_PRICE represents the goods price of good that client asked for on the previous application. AMT\_CREDIT represents the final credit amount on the previous application. Through the definitions of these two variables, it's clear that we could drop one of the variables.
- 2. (CNT\_FAM\_MEMBERS, CNT\_CHILDREN), corr = 0.914 CNT\_FAM\_MEMBERS represents how many family members clients have CNT\_CHILDREN represents how many children clients have Through the definitions of these two variables, it's clear that we should

drop one of the variables.

```
#check collinearity
num <- unlist(lapply(application_data_selected, is.numeric))
cor(cc(application_data_selected[,num]), use = "pair")
```

##	TARGET	1.000000000	-0.022374536	0.005679723
##	AMT_INCOME_TOTAL	-0.022374536	1.000000000	0.006490665
##	CNT_CHILDREN	0.005679723	0.006490665	1.000000000
##	AMT_CREDIT	-0.035478230	0.324854819	-0.021083017
##	AMT_ANNUITY	-0.015778897	0.411376003	-0.001362760
##	AMT_GOODS_PRICE	-0.044389541	0.332695445	-0.024611028
##	REGION_POPULATION_RELATIVE	-0.036110371	0.167362370	-0.032117581
##	DAYS_BIRTH	0.054896182	0.015586880	0.279672331
##	DAYS_EMPLOYED	-0.028323415	-0.103847372	-0.183114593
##	OWN_CAR_AGE	0.037232346	-0.132171880	0.008950126
##	CNT_FAM_MEMBERS	-0.001250596	-0.001726251	0.914413945
##	AMR_REQ_CREDIT_BUREAU_SUM	0.019143211	0.035277693	-0.029676739
##	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	
##	TARGET	-0.03547823	-0.01577890	-0.04438954
##	AMT_INCOME_TOTAL	0.32485482	0.41137600	0.33269544
##	CNT_CHILDREN	-0.02108302	-0.00136276	-0.02461103
##	AMT_CREDIT	1.000000000	0.74634289	0.98718062
##	AMT_ANNUITY	0.74634289	1.000000000	0.75161981
##	AMT_GOODS_PRICE	0.98718062	0.75161981	1.000000000
##	REGION_POPULATION_RELATIVE	0.09204365	0.11265566	0.09609678
##	DAYS_BIRTH	-0.11694218	-0.05319892	-0.11380798
##	DAYS_EMPLOYED	-0.01937100	-0.05245937	-0.01794351
##	OWN_CAR_AGE	-0.09280021	-0.09619776	-0.10185298
##	CNT_FAM_MEMBERS	0.02020290	0.03046515	0.01779781
##	AMR_REQ_CREDIT_BUREAU_SUM	-0.01731215	0.01144510	-0.01839987
##	REGION_POPULATION_RELATIVE	DAYS_BIRTH		
##	TARGET	-0.036110371	0.054896182	
##	AMT_INCOME_TOTAL	0.167362370	0.015586880	
##	CNT_CHILDREN	-0.032117581	0.279672331	
##	AMT_CREDIT	0.092043647	-0.116942177	
##	AMT_ANNUITY	0.112655663	-0.053198919	
##	AMT_GOODS_PRICE	0.096096778	-0.113807976	
##	REGION_POPULATION_RELATIVE	1.000000000	-0.038847969	
##	DAYS_BIRTH	-0.038847969	1.000000000	
##	DAYS_EMPLOYED	0.002191339	-0.508824142	
##	OWN_CAR_AGE	-0.081436107	0.001110481	
##	CNT_FAM_MEMBERS	-0.034209956	0.197954108	
##	AMR_REQ_CREDIT_BUREAU_SUM	0.025584835	-0.032038595	
##	DAYS_EMPLOYED	OWN_CAR_AGE	CNT_FAM_MEMBERS	
##	TARGET	-0.028323415	0.037232346	-0.001250596
##	AMT_INCOME_TOTAL	-0.103847372	-0.132171880	-0.001726251
##	CNT_CHILDREN	-0.183114593	0.008950126	0.914413945
##	AMT_CREDIT	-0.019371002	-0.092800207	0.020202897
##	AMT_ANNUITY	-0.052459368	-0.096197763	0.030465151
##	AMT_GOODS_PRICE	-0.017943510	-0.101852976	0.017797807
##	REGION_POPULATION_RELATIVE	0.002191339	-0.081436107	-0.034209956
##	DAYS_BIRTH	-0.508824142	0.001110481	0.197954108
##	DAYS_EMPLOYED	1.000000000	0.031961957	-0.152888423
##	OWN_CAR_AGE	0.031961957	1.000000000	-0.014499077
##	CNT_FAM_MEMBERS	-0.152888423	-0.014499077	1.000000000
##	AMR_REQ_CREDIT_BUREAU_SUM	0.007042087	-0.026658773	-0.018385506
##	AMR_REQ_CREDIT_BUREAU_SUM			
##	TARGET	0.019143211		
##	AMT_INCOME_TOTAL	0.035277693		
##	CNT_CHILDREN	-0.029676739		
##	AMT_CREDIT	-0.017312145		
##	AMT_ANNUITY	0.011445101		
##	AMT_GOODS_PRICE	-0.018399870		
##	REGION_POPULATION_RELATIVE	0.025584835		
##	DAYS_BIRTH	-0.032038595		
##	DAYS_EMPLOYED	0.007042087		
##	OWN_CAR_AGE	-0.026658773		
##	CNT_FAM_MEMBERS	-0.018385506		
##	AMR_REQ_CREDIT_BUREAU_SUM	1.000000000		

```
application_data_selected <- application_data_selected %>%  
  dplyr::select(-(AMT_GOODS_PRICE)) %>%  
  dplyr::select(-(CNT_CHILDREN))
```

Here we should factorize `CNT_FAM_MEMBERS` , representing how many family members the loan applicant has, and `AMT_REQ_CREDIT_BUREAU_SUM` , representing number of enquiries to Credit Bureau about the client one year before submitting his loan application, as both variables are more like categorical rather than continuous based on their definitions.

```
#factorize CNT_FAM_MEMBERS, AMR_REQ_CREDIT_BUREAU_SUM  
application_data_selected$CNT_FAM_MEMBERS <- as.factor(application_data_selected$CNT_FAM_MEMBERS)  
application_data_selected$AMR_REQ_CREDIT_BUREAU_SUM <- as.factor(application_data_selected$AMR_REQ_CREDIT_BUREAU_SUM)
```

As the dataset is extremely large and my computer's computing power doesn't support this size of computations. Therefore, I decided to randomly select 10,000 observations to run further analysis. Due to the random selection process, the random selection process should be able to represent the population. The summary tables for original dataset and the dataset after my random selection process in the Data Description section below also support this claim.

```
set.seed(100)  
application_data_sam <- application_data_selected[sample(nrow(application_data_selected), 10000), ]  
summary(application_data_sam)
```

```

##      TARGET      CODE_GENDER      NAME_CONTRACT_TYPE AMT_INCOME_TOTAL
## Min.      :0.0000    F   :6556    Cash loans      :9025    Min.      : 25650
## 1st Qu.:0.0000    M   :3444    Revolving loans: 975    1st Qu.: 112500
## Median :0.0000    XNA:   0                                Median : 148500
## Mean   :0.0829                                Mean   : 167765
## 3rd Qu.:0.0000                                3rd Qu.: 202500
## Max.    :1.0000                                Max.    :2250000
##
## FLAG_OWN_CAR FLAG_OWN_REALTY  AMT_CREDIT      AMT_ANNUITY
## N:6540      N:3070      Min.      : 45000  Min.      : 2844
## Y:3460      Y:6930      1st Qu.: 270000  1st Qu.: 16574
##                                         Median : 521280  Median : 25047
##                                         Mean   : 599379  Mean   : 27049
##                                         3rd Qu.: 808650  3rd Qu.: 34533
##                                         Max.    :2961000  Max.    :133848
##
##      NAME_TYPE_SUITE      NAME_INCOME_TYPE
## Unaccompanied :8115    Working      :5123
## Family        :1278    Commercial associate:2335
## Spouse, partner: 374    Pensioner      :1804
## Children      : 118    State servant   : 735
## Other_B       : 50     Businessman     : 1
##              : 39     Student           : 1
## (Other)       : 26     (Other)         : 1
##
##      NAME_EDUCATION_TYPE      NAME_FAMILY_STATUS
## Academic degree      : 5    Civil marriage      :1036
## Higher education     :2414   Married              :6320
## Incomplete higher    : 305   Separated           : 663
## Lower secondary      : 114   Single / not married:1469
## Secondary / secondary special:7162  Unknown              : 0
##                                         Widow              : 512
##
##      NAME_HOUSING_TYPE REGION_POPULATION_RELATIVE  DAYS_BIRTH
## Co-op apartment      : 45    Min.      :0.000938    Min.      : -25186
## House / apartment    :8863   1st Qu.:0.010006    1st Qu.: -19688
## Municipal apartment  : 356   Median :0.018850    Median : -15820
## Office apartment     : 80    Mean   :0.020710    Mean   : -16047
## Rented apartment     : 179   3rd Qu.:0.028663    3rd Qu.: -12469
## With parents         : 477   Max.    :0.072508    Max.    : -7721
##
## DAYS_EMPLOYED      OWN_CAR_AGE      OCCUPATION_TYPE CNT_FAM_MEMBERS
## Min.      : -17546.0  Min.      : 0.00      :3162    2      :5197
## 1st Qu.: -2729.0    1st Qu.: 5.00    Laborers   :1793    1      :2186
## Median : -1181.5    Median : 9.00    Sales staff:1041    3      :1686
## Mean   : 63952.6    Mean :11.97    Core staff : 878    4      : 814
## 3rd Qu.: -283.5    3rd Qu.:15.00    Managers   : 667    5      : 101
## Max.    :365243.0    Max.    :65.00    Drivers    : 623    6      : 11
##
##      NA's :6540    (Other) :1836    (Other): 5
##
##      ORGANIZATION_TYPE AMR_REQ_CREDIT_BUREAU_SUM
## Business Entity Type 3:2229    1      :1719
## XNA                  :1804    2      :1696
## Self-employed        :1263    0      :1672
## Other                : 529    3      :1304
## Business Entity Type 2: 354    4      : 906
## Government           : 350    (Other):1376
## (Other)              :3471    NA's    :1327

```

## Data Description

Through the two tables below, we can see that the distributions of each variable in the original and new datasets are fairly similar. Besides, we also have some missing values and should consider missing value imputations.

```
summary(application_data_selected) #original datasets with selected variables
```

```

##      TARGET      CODE_GENDER      NAME_CONTRACT_TYPE
## Min.      :0.00000    F :202448    Cash loans      :278232
## 1st Qu.:0.00000    M :105059    Revolving loans: 29279
## Median :0.00000    XNA:      4
## Mean      :0.08073
## 3rd Qu.:0.00000
## Max.      :1.00000
##
## AMT_INCOME_TOTAL    FLAG_OWN_CAR FLAG_OWN_REALTY    AMT_CREDIT
## Min.      :    25650    N:202924    N: 94199    Min.      : 45000
## 1st Qu.:    112500    Y:104587    Y:213312    1st Qu.: 270000
## Median :    147150
## Mean      :    168798
## 3rd Qu.:    202500
## Max.      :117000000
##
## AMT_ANNUITY          NAME_TYPE_SUITE          NAME_INCOME_TYPE
## Min.      : 1616    Unaccompanied :248526    Working      :158774
## 1st Qu.: 16524    Family        : 40149    Commercial associate: 71617
## Median : 24903    Spouse, partner: 11370    Pensioner    : 55362
## Mean      : 27109    Children      : 3267    State servant : 21703
## 3rd Qu.: 34596    Other_B       : 1770    Unemployed    : 22
## Max.      :258026
## NA's      :12      (Other)       : 1292    Student       : 18
##
## NAME_EDUCATION_TYPE    NAME_FAMILY_STATUS
## Academic degree        : 164    Civil marriage : 29775
## Higher education       : 74863    Married        :196432
## Incomplete higher      : 10277    Separated      : 19770
## Lower secondary        : 3816    Single / not married: 45444
## Secondary / secondary special:218391    Unknown        : 2
## Widow                 : 16088
##
## NAME_HOUSING_TYPE    REGION_POPULATION_RELATIVE    DAYS_BIRTH
## Co-op apartment      : 1122    Min.      :0.00029    Min.      :-25229
## House / apartment    :272868    1st Qu.:0.01001    1st Qu.: -19682
## Municipal apartment: 11183    Median :0.01885    Median : -15750
## Office apartment     : 2617    Mean      :0.02087    Mean      :-16037
## Rented apartment     : 4881    3rd Qu.:0.02866    3rd Qu.: -12413
## With parents         : 14840    Max.      :0.07251    Max.      :-7489
##
## DAYS_EMPLOYED        OWN_CAR_AGE        OCCUPATION_TYPE    CNT_FAM_MEMBERS
## Min.      : -17912    Min.      : 0.00      :96391    2      :158357
## 1st Qu.: -2760    1st Qu.: 5.00    Laborers :55186    1      : 67847
## Median : -1213    Median : 9.00    Sales staff:32102    3      : 52601
## Mean      : 63815    Mean :12.06    Core staff :27570    4      : 24697
## 3rd Qu.: -289    3rd Qu.:15.00    Managers :21371    5      : 3478
## Max.      :365243    Max.      :91.00    Drivers   :18603    (Other): 529
## NA's      :202929    (Other)    :56288    NA's      : 2
## ORGANIZATION_TYPE    AMR_REQ_CREDIT_BUREAU_SUM
## Business Entity Type 3: 67992    1      :53914
## XNA                    : 55374    2      :51559
## Self-employed         : 38412    0      :50911
## Other                 : 16683    3      :39380
## Medicine              : 11193    4      :27241
## Business Entity Type 2: 10553    (Other):42987
## (Other)               :107304    NA's     :41519

```

```
summary(application_data_sam) #the dataset after the random selection process
```



```

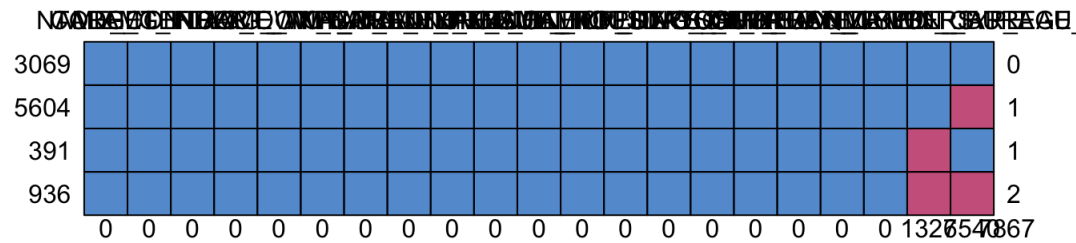
##      TARGET      CODE_GENDER      NAME_CONTRACT_TYPE AMT_INCOME_TOTAL
## Min.      :0.0000    F   :6556    Cash loans      :9025    Min.      : 25650
## 1st Qu.:0.0000    M   :3444    Revolving loans: 975    1st Qu.: 112500
## Median :0.0000    XNA:    0                                Median : 148500
## Mean   :0.0829                                Mean   : 167765
## 3rd Qu.:0.0000                                3rd Qu.: 202500
## Max.    :1.0000                                Max.    :2250000
##
## FLAG_OWN_CAR FLAG_OWN_REALTY  AMT_CREDIT      AMT_ANNUITY
## N:6540      N:3070      Min.      : 45000  Min.      : 2844
## Y:3460      Y:6930      1st Qu.: 270000  1st Qu.: 16574
##                                         Median : 521280  Median : 25047
##                                         Mean   : 599379  Mean   : 27049
##                                         3rd Qu.: 808650  3rd Qu.: 34533
##                                         Max.    :2961000  Max.    :133848
##
##      NAME_TYPE_SUITE      NAME_INCOME_TYPE
## Unaccompanied :8115    Working      :5123
## Family        :1278    Commercial associate:2335
## Spouse, partner: 374    Pensioner      :1804
## Children      : 118    State servant   : 735
## Other_B       : 50     Businessman     : 1
##              : 39     Student          : 1
## (Other)       : 26     (Other)         : 1
##
##      NAME_EDUCATION_TYPE      NAME_FAMILY_STATUS
## Academic degree      : 5    Civil marriage      :1036
## Higher education     :2414   Married              :6320
## Incomplete higher    : 305   Separated            : 663
## Lower secondary      : 114   Single / not married:1469
## Secondary / secondary special:7162  Unknown              : 0
##                                         Widow                : 512
##
##      NAME_HOUSING_TYPE REGION_POPULATION_RELATIVE  DAYS_BIRTH
## Co-op apartment      : 45    Min.      :0.000938    Min.      : -25186
## House / apartment    :8863   1st Qu.:0.010006    1st Qu.: -19688
## Municipal apartment: 356    Median :0.018850    Median : -15820
## Office apartment     : 80    Mean   :0.020710    Mean   : -16047
## Rented apartment     : 179   3rd Qu.:0.028663    3rd Qu.: -12469
## With parents         : 477   Max.    :0.072508    Max.    : -7721
##
## DAYS_EMPLOYED      OWN_CAR_AGE      OCCUPATION_TYPE CNT_FAM_MEMBERS
## Min.      : -17546.0  Min.      : 0.00      :3162    2      :5197
## 1st Qu.: -2729.0    1st Qu.: 5.00    Laborers   :1793    1      :2186
## Median : -1181.5    Median : 9.00    Sales staff:1041    3      :1686
## Mean   : 63952.6    Mean   :11.97    Core staff : 878    4      : 814
## 3rd Qu.: -283.5    3rd Qu.:15.00    Managers   : 667    5      : 101
## Max.    :365243.0    Max.    :65.00    Drivers    : 623    6      : 11
##
##      NA's :6540    (Other) :1836    (Other): 5
##
##      ORGANIZATION_TYPE AMR_REQ_CREDIT_BUREAU_SUM
## Business Entity Type 3:2229    1      :1719
## XNA                  :1804    2      :1696
## Self-employed        :1263    0      :1672
## Other                : 529    3      :1304
## Business Entity Type 2: 354    4      : 906
## Government           : 350    (Other):1376
## (Other)              :3471    NA's    :1327

```

## Missing values imputation

OWN\_CAR\_AGE and AMR\_REQ\_CREDIT\_BUREAU\_SUM are the two columns that contain missing values in the dataset. I used the mice package to conduct missing value imputation to generate complete datasets. For the imputation method, I chose cart, instead of the default method. I have tried to use the default method to impute missing values; however, it returned the following error "system is computationally singular". The cause of the problem here could probably be the large number of unbalanced factor variables, such as NAME\_CONTRACT\_TYPE, NAME\_TYPE\_SUITE and NAME\_EDUCATION\_TYPE, in the dataset. When these variables are turned into dummy variables, there's a high probability that one column is a linear combination of another. As the default imputation methods are parametric, which involve linear

```
#check pattern
md.pattern(application_data_sam)
```



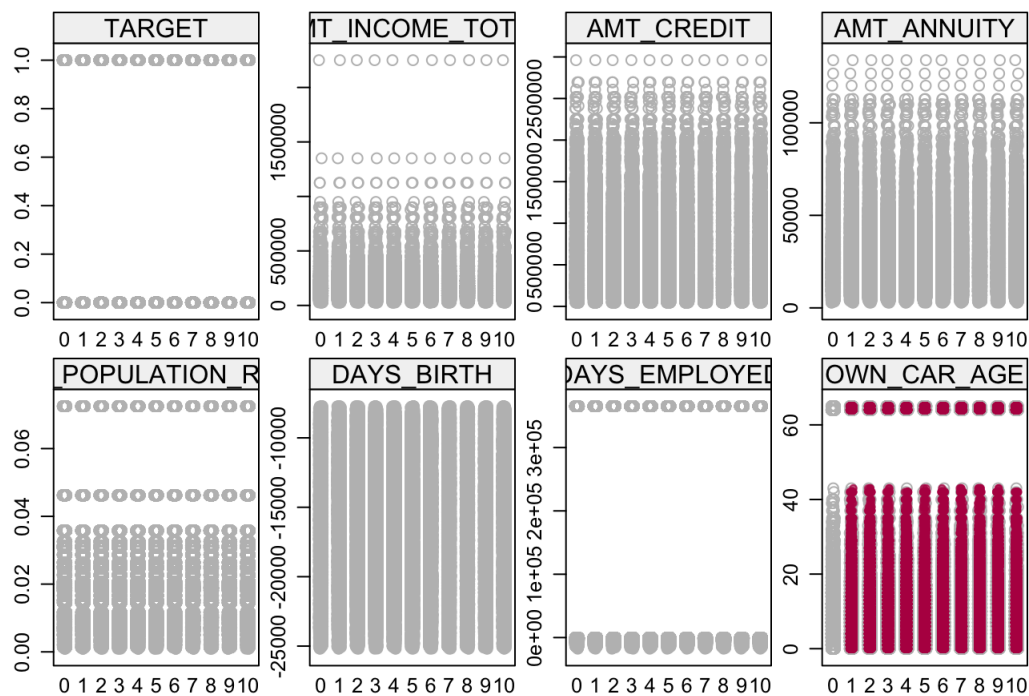
```
##      TARGET CODE_GENDER NAME_CONTRACT_TYPE AMT_INCOME_TOTAL FLAG_OWN_CAR
## 3069      1           1           1           1           1
## 5604      1           1           1           1           1
## 391       1           1           1           1           1
## 936       1           1           1           1           1
##          0           0           0           0           0
##      FLAG_OWN_REALTY AMT_CREDIT AMT_ANNUITY NAME_TYPE_SUITE
## 3069          1           1           1           1
## 5604          1           1           1           1
## 391          1           1           1           1
## 936          1           1           1           1
##          0           0           0           0
##      NAME_INCOME_TYPE NAME_EDUCATION_TYPE NAME_FAMILY_STATUS
## 3069          1           1           1
## 5604          1           1           1
## 391          1           1           1
## 936          1           1           1
##          0           0           0
##      NAME_HOUSING_TYPE REGION_POPULATION_RELATIVE DAYS_BIRTH DAYS_EMPLOYED
## 3069          1           1           1           1
## 5604          1           1           1           1
## 391          1           1           1           1
## 936          1           1           1           1
##          0           0           0           0
##      OCCUPATION_TYPE CNT_FAM_MEMBERS ORGANIZATION_TYPE
## 3069          1           1           1
## 5604          1           1           1
## 391          1           1           1
## 936          1           1           1
##          0           0           0
##      AMR_REQ_CREDIT_BUREAU_SUM OWN_CAR_AGE
## 3069          1           1   0
## 5604          1           0   1
## 391          0           1   1
## 936          0           0   2
##          1327          6540 7867
```

```
application_MI <- mice(application_data_sam, m = 10, method = "cart", seed = 8)
```

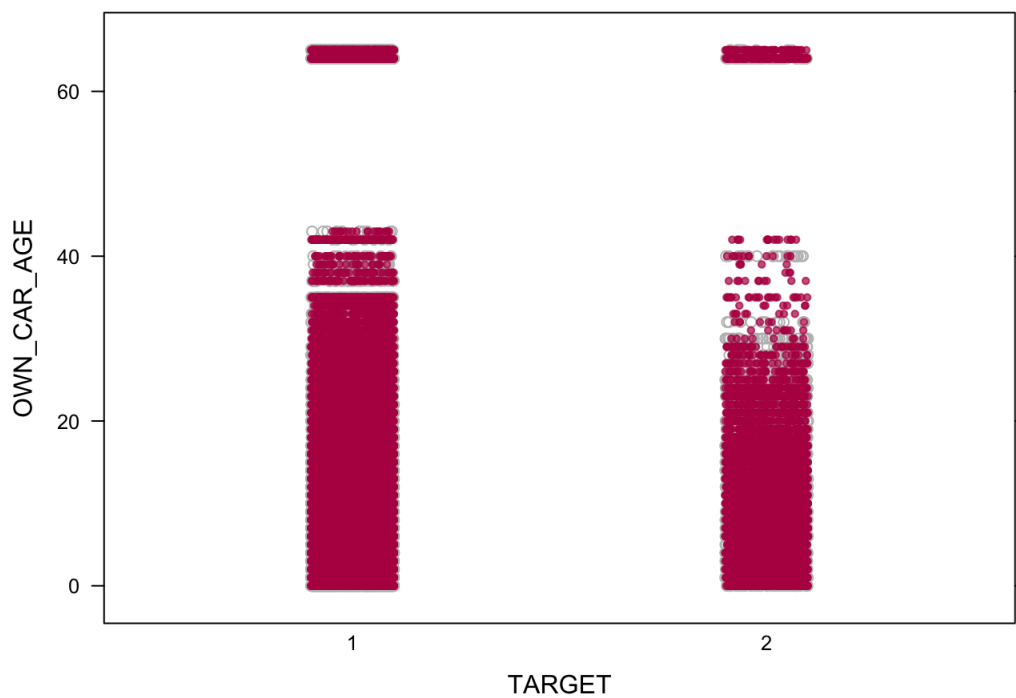
## Imputation Model Diagnostics

Based on the charts below, no problematic pattern with the imputations.

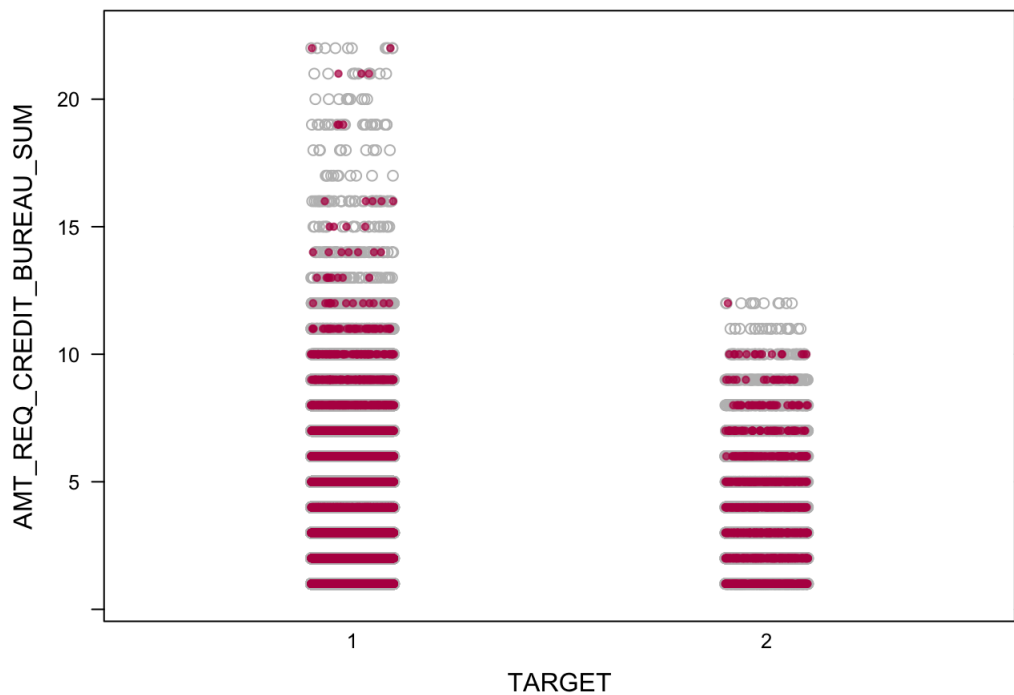
```
stripplot(application_MI, col=c("grey",mdc(2)),pch=c(1,20))
```



```
stripplot(application_MI, OWN_CAR_AGE~TARGET, col=c("grey",mdc(2)),pch=c(1,20), xlab = 'TARGET', ylab = "OWN_CAR_AGE")
```



```
stripplot(application_MI, AMR_REQ_CREDIT_BUREAU_SUM~TARGET, col=c("grey",mdc(2)),pch=c(1,20), xlab = 'TARGET', ylab = "AMT_REQ_CREDIT_BUREAU_SUM")
```



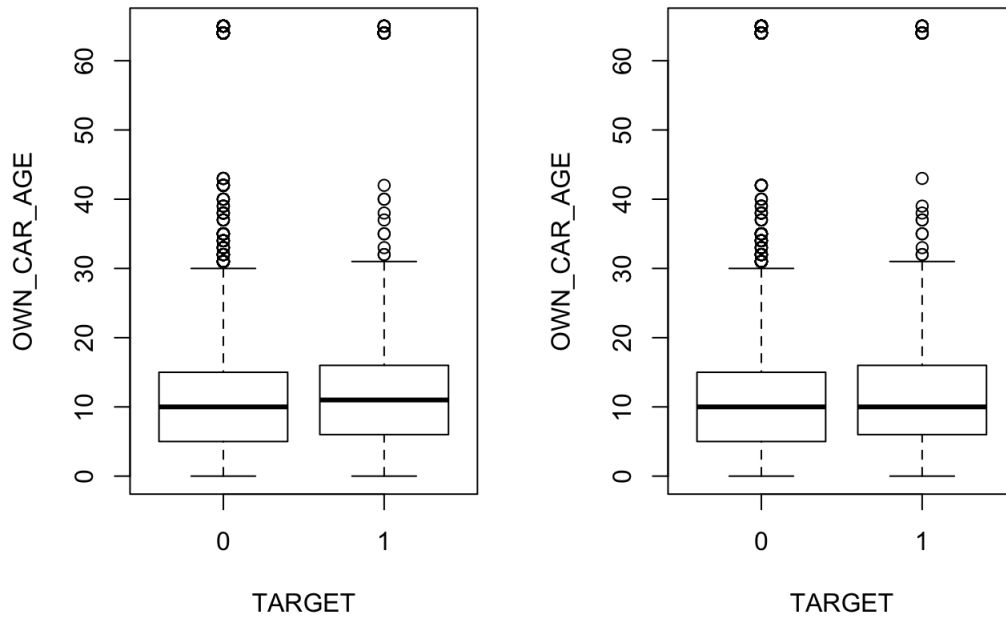
## Posterior Predictive Check on Two Complete Datasets

Both the histogram and boxplots look similar for replica and complete datasets; therefore, there is no evidence suggests that imputation models are poorly specified for what I want to do.

```
application_ppcheck <- rbind(application_data_sam, application_data_sam)
application_ppcheck[10001:20000, apply(is.na(application_data_sam), any, MARGIN = 2)] <- NA
application_ppcheck_MI <- mice(application_ppcheck, m = 10, method = "cart", seed = 12)
d1ppcheck <- mice::complete(application_ppcheck_MI, 1)
d2ppcheck <- mice::complete(application_ppcheck_MI, 2)
```

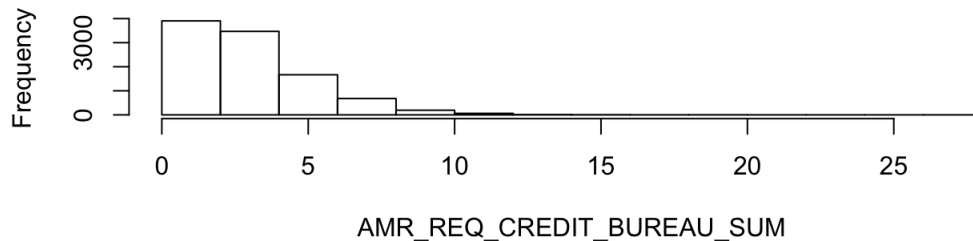
```
#dataset1
par(mfrow = c(1,2))
boxplot(d1ppcheck$OWN_CAR_AGE[1:10000]~d1ppcheck$TARGET[1:10000], ylab="OWN_CAR_AGE", xlab="TARGET", main = "OWN_CAR_AGE vs TARGET completed data")
boxplot(d1ppcheck$OWN_CAR_AGE[10001:20000]~d1ppcheck$TARGET[10001:20000], ylab="OWN_CAR_AGE", xlab="TARGET", main = "OWN_CAR_AGE vs TARGET completed data")
```

## WN\_CAR\_AGE vs TARGET complete

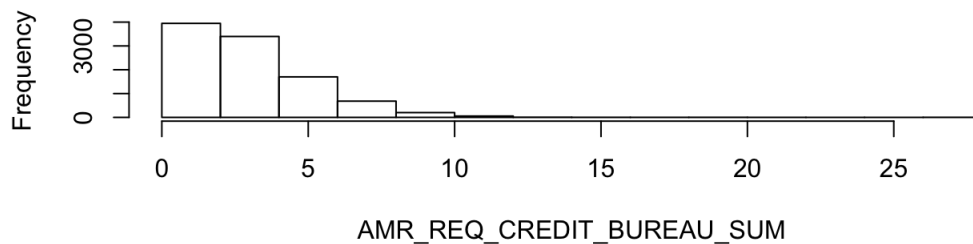


```
par(mfrow = c(2,1))
hist(as.numeric(dlpptest$AMR_REQ_CREDIT_BUREAU_SUM[1:10000]), xlab="AMR_REQ_CREDIT_BUREAU_SUM", main = "AMR_REQ_CREDIT_BUREAU_SUM complete data")
hist(as.numeric(dlpptest$AMR_REQ_CREDIT_BUREAU_SUM[10001:20000]), xlab="AMR_REQ_CREDIT_BUREAU_SUM", main = "AMR_REQ_CREDIT_BUREAU_SUM replicated data")
```

### AMR\_REQ\_CREDIT\_BUREAU\_SUM complete data



### AMR\_REQ\_CREDIT\_BUREAU\_SUM replicated data



## Exploratory Data Analysis

For categorical variables:

```
#Dataset 1
cd1 <- mice::complete(application_MI, 1)
tapply(cd1$TARGET, cd1$CODE_GENDER, mean)
```

```
##           F           M           XNA
## 0.06696156 0.11324042           NA
```

```
tapply(cd1$TARGET, cd1$NAME_CONTRACT_TYPE, mean)
```

```
##      Cash loans Revolving loans
##      0.08698061      0.04512821
```

```
tapply(cd1$TARGET, cd1$FLAG_OWN_CAR, mean)
```

```
##           N           Y
## 0.08837920 0.07254335
```

```
tapply(cd1$TARGET, cd1$FLAG_OWN_REALTY, mean)
```

```
##           N           Y
## 0.08534202 0.08181818
```

```
tapply(cd1$TARGET, cd1$NAME_TYPE_SUITE, mean)
```

```
##           Children           Family Group of people
##      0.07692308      0.05084746      0.08607199      0.00000000
##      Other_A      Other_B Spouse, partner      Unaccompanied
##      0.04761905      0.06000000      0.09090909      0.08280961
```

```
tapply(cd1$TARGET, cd1$NAME_INCOME_TYPE, mean)
```

```
##      Businessman Commercial associate      Maternity leave
##      0.00000000      0.07366167      NA
##      Pensioner      State servant      Student
##      0.05266075      0.04353741      0.00000000
##      Unemployed      Working
##      0.00000000      0.10345501
```

```
tapply(cd1$TARGET, cd1$NAME_EDUCATION_TYPE, mean)
```

```
##      Academic degree      Higher education
##      0.00000000      0.05675228
##      Incomplete higher      Lower secondary
##      0.10163934      0.12280702
## Secondary / secondary special
##      0.09033789
```

```
tapply(cd1$TARGET, cd1$NAME_FAMILY_STATUS, mean)
```

```
##      Civil marriage      Married      Separated
##      0.09266409      0.07721519      0.07843137
## Single / not married      Unknown      Widow
##      0.10619469      NA      0.07226562
```

```
tapply(cd1$TARGET, cd1$NAME_HOUSING_TYPE, mean)
```

##	Co-op apartment	House / apartment	Municipal apartment
##	0.08888889	0.08055963	0.07022472
##	Office apartment	Rented apartment	With parents
##	0.01250000	0.13407821	0.12788260

```
tapply(cdl$TARGET, cdl$OCCUPATION_TYPE, mean)
```

##		Accountants	Cleaning staff
##	0.06672992	0.04966887	0.06993007
##	Cooking staff	Core staff	Drivers
##	0.10194175	0.06036446	0.13162119
##	High skill tech staff	HR staff	IT staff
##	0.06024096	0.07142857	0.04761905
##	Laborers	Low-skill Laborers	Managers
##	0.10540993	0.16923077	0.05847076
##	Medicine staff	Private service staff	Realty agents
##	0.06800000	0.07920792	0.16666667
##	Sales staff	Secretaries	Security staff
##	0.09510086	0.11428571	0.14222222
##	Waiters/barmen staff		
##	0.20000000		

```
tapply(cdl$TARGET, cdl$CNT_FAM_MEMBERS, mean)
```

##	1	2	3	4	5	6
##	0.09103385	0.07677506	0.08718861	0.08353808	0.11881188	0.27272727
##	7	8	9	10	11	12
##	0.20000000	NA	NA	NA	NA	NA
##	13	14	15	16	20	
##	NA	NA	NA	NA	NA	

```
tapply(cdl$TARGET, cdl$ORGANIZATION_TYPE, mean)
```



##	Advertising	Agriculture	Bank
##	0.05555556	0.09523810	0.06451613
##	Business Entity Type 1	Business Entity Type 2	Business Entity Type 3
##	0.08080808	0.11299435	0.09600718
##	Cleaning	Construction	Culture
##	0.11111111	0.11607143	0.00000000
##	Electricity	Emergency	Government
##	0.06250000	0.04761905	0.06285714
##	Hotel	Housing	Industry: type 1
##	0.00000000	0.07766990	0.05263158
##	Industry: type 10	Industry: type 11	Industry: type 12
##	NA	0.09411765	0.16666667
##	Industry: type 13	Industry: type 2	Industry: type 3
##	0.00000000	0.00000000	0.13114754
##	Industry: type 4	Industry: type 5	Industry: type 6
##	0.10526316	0.10000000	0.00000000
##	Industry: type 7	Industry: type 8	Industry: type 9
##	0.11627907	NA	0.11403509
##	Insurance	Kindergarten	Legal Services
##	0.00000000	0.05957447	0.07142857
##	Medicine	Military	Mobile
##	0.08682635	0.02298851	0.00000000
##	Other	Police	Postal
##	0.07372401	0.01369863	0.05633803
##	Realtor	Religion	Restaurant
##	0.27272727	0.00000000	0.08196721
##	School	Security	Security Ministries
##	0.05147059	0.10101010	0.06153846
##	Self-employed	Services	Telecom
##	0.10530483	0.06122449	0.05263158
##	Trade: type 1	Trade: type 2	Trade: type 3
##	0.09090909	0.08928571	0.06611570
##	Trade: type 4	Trade: type 5	Trade: type 6
##	0.00000000	NA	0.04347826
##	Trade: type 7	Transport: type 1	Transport: type 2
##	0.11600000	0.00000000	0.14864865
##	Transport: type 3	Transport: type 4	University
##	0.26086957	0.04895105	0.05714286
##	XNA		
##	0.05266075		

For continuous variables:

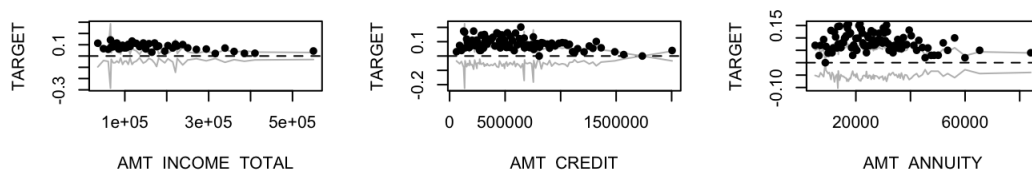
```

par(mfrow = c(3,3))
binnedplot(x=cd1$AMT_INCOME_TOTAL, y = cd1$TARGET, xlab = "AMT_INCOME_TOTAL", ylab = "TARGET",
  main = "Binned TARGET versus AMT_INCOME_TOTAL")
binnedplot(x=cd1$AMT_CREDIT, y = cd1$TARGET, xlab = "AMT_CREDIT", ylab = "TARGET",
  main = "Binned TARGET versus AMT_CREDIT")
binnedplot(x=cd1$AMT_ANNUITY, y = cd1$TARGET, xlab = "AMT_ANNUITY", ylab = "TARGET",
  main = "Binned TARGET versus AMT_ANNUITY")
binnedplot(x=cd1$REGION_POPULATION_RELATIVE, y = cd1$TARGET, xlab = "REGION_POPULATION_RELATIVE", ylab = "TARGET",
  main = "Binned TARGET versus REGION_POPULATION_RELATIVE")
binnedplot(x=cd1$DAYS_BIRTH, y = cd1$TARGET, xlab = "DAYS_BIRTH", ylab = "TARGET",
  main = "Binned TARGET versus DAYS_BIRTH")
binnedplot(x=cd1$DAYS_EMPLOYED, y = cd1$TARGET, xlab = "DAYS_EMPLOYED", ylab = "TARGET",
  main = "Binned TARGET versus DAYS_EMPLOYED")
binnedplot(x=cd1$OWN_CAR_AGE, y = cd1$TARGET, xlab = "OWN_CAR_AGE", ylab = "TARGET",
  main = "Binned TARGET versus OWN_CAR_AGE")
binnedplot(x=as.numeric(cd1$AMR_REQ_CREDIT_BUREAU_SUM), y = cd1$TARGET, xlab = "AMR_REQ_CREDIT_BUREAU_SUM", ylab = "TARGET",
  main = "Binned TARGET versus AMR_REQ_CREDIT_BUREAU_SUM")

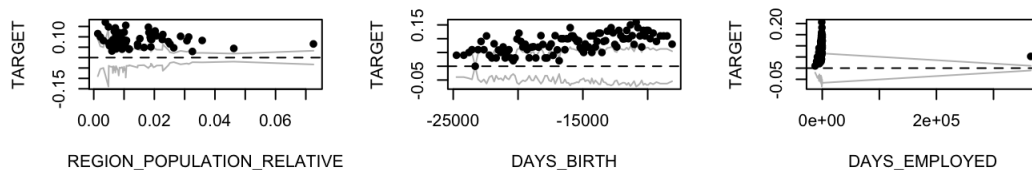
temp <- cd1[cd1$DAYS_EMPLOYED != 365243,]
binnedplot(x=temp$DAYS_EMPLOYED, y = cd1$TARGET, xlab = "DAYS_EMPLOYED", ylab = "TARGET",
  main = "Binned TARGET versus DAYS_EMPLOYED")

```

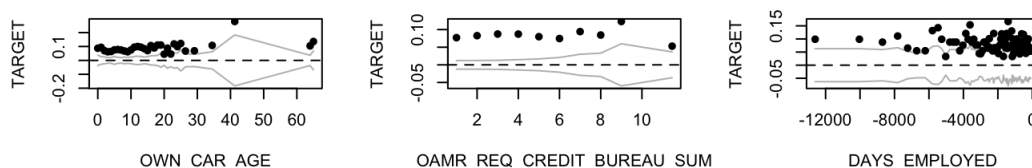
uned TARGET versus AMT\_INCOME\_ Binned TARGET versus AMT\_CRED Binned TARGET versus AMT\_ANNU



ARGET versus REGION\_POPULATIO Binned TARGET versus DAYS\_BIR3inned TARGET versus DAYS\_EMPLC



Binned TARGET versus OWN\_CAR\_ARGET versus AMR\_REQ\_CREDIT\_Binned TARGET versus DAYS\_EMPLC



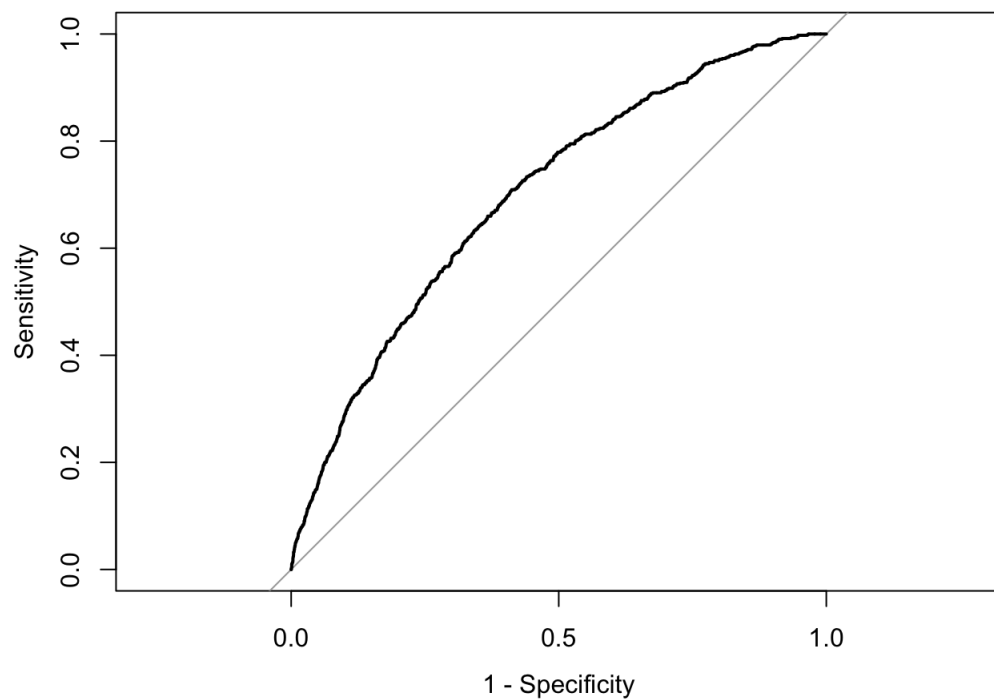
## Base Regression Model

Here I constructed a base logistic regression model without any transformation and interaction term. I also selected three complete datasets to see AUCs for the fitted models. The AUCs are fairly good at 0.6981, 0.6993 and 0.6986.

```
reg <- with(data = application_MI, glm(TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM), family = binomial)
```

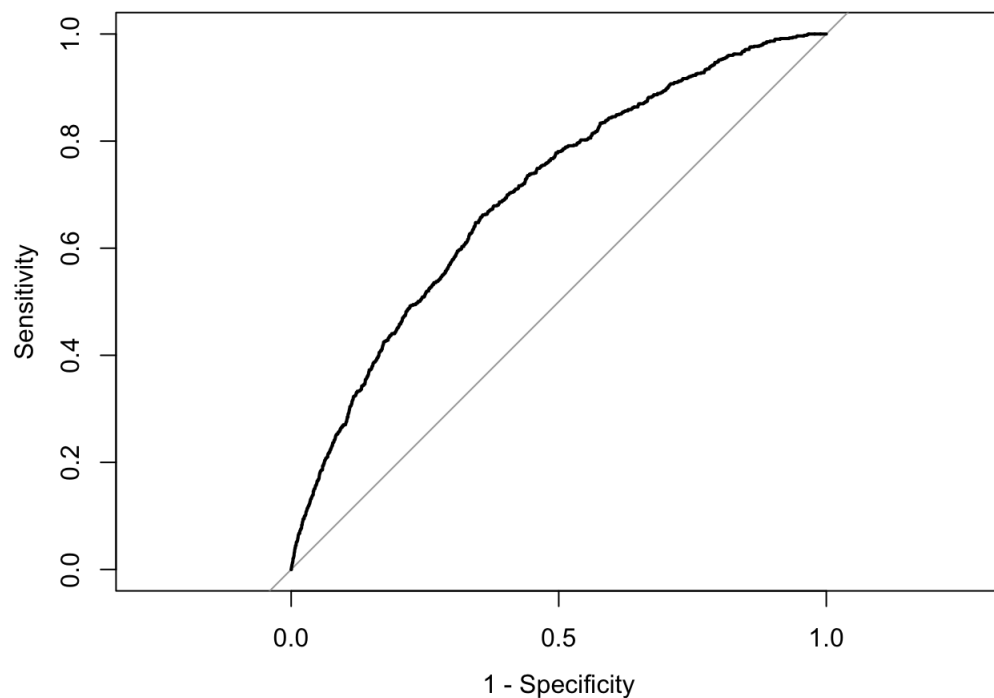
Use Three Completed Datasets to See AUCs for the Fitted Models

```
#Dataset 1
par(mfrow=c(1,1))
cd1 <- mice::complete(application_MI, 1)
reg_cd1 <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM, family=binomial)
roc(cd1$TARGET, fitted(reg_cd1), plot=T, legacy.axes=T)
```



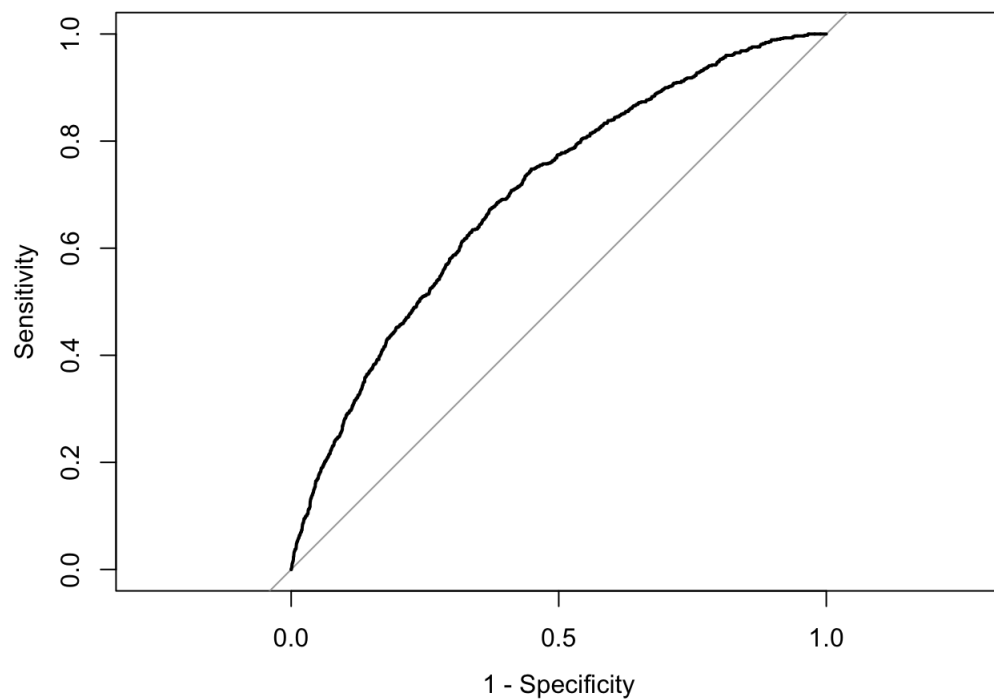
```
##
## Call:
## roc.default(response = cd1$TARGET, predictor = fitted(reg_cd1),      plot = T, legacy.axes = T)
##
## Data: fitted(reg_cd1) in 9171 controls (cd1$TARGET 0) < 829 cases (cd1$TARGET 1).
## Area under the curve: 0.6981
```

```
#Dataset 2
cd2 <- mice::complete(application_MI, 2)
reg_cd2 <- glm(data=cd2, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_RE
ALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS
+ NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM, family=binomial)
roc(cd2$TARGET, fitted(reg_cd2), plot=T, legacy.axes=T)
```



```
##
## Call:
## roc.default(response = cd2$TARGET, predictor = fitted(reg_cd2),      plot = T, legacy.axes = T)
##
## Data: fitted(reg_cd2) in 9171 controls (cd2$TARGET 0) < 829 cases (cd2$TARGET 1).
## Area under the curve: 0.6993
```

```
#Dataset 3
cd3 <- mice::complete(application_MI, 3)
reg_cd3 <- glm(data=cd3, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_RE
ALTY + AMT_CREDIT + AMT_ANNUITY
              + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TY
PE
              + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_F
AM_MEMBERS
              + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM,
              family=binomial)
roc(cd3$TARGET, fitted(reg_cd3), plot=T, legacy.axes=T)
```



```
##  
## Call:  
## roc.default(response = cd3$TARGET, predictor = fitted(reg_cd3),      plot = T, legacy.axes = T)  
##  
## Data: fitted(reg_cd3) in 9171 controls (cd3$TARGET 0) < 829 cases (cd3$TARGET 1).  
## Area under the curve: 0.6986
```

## Model Diagnostics

For continuous variables:

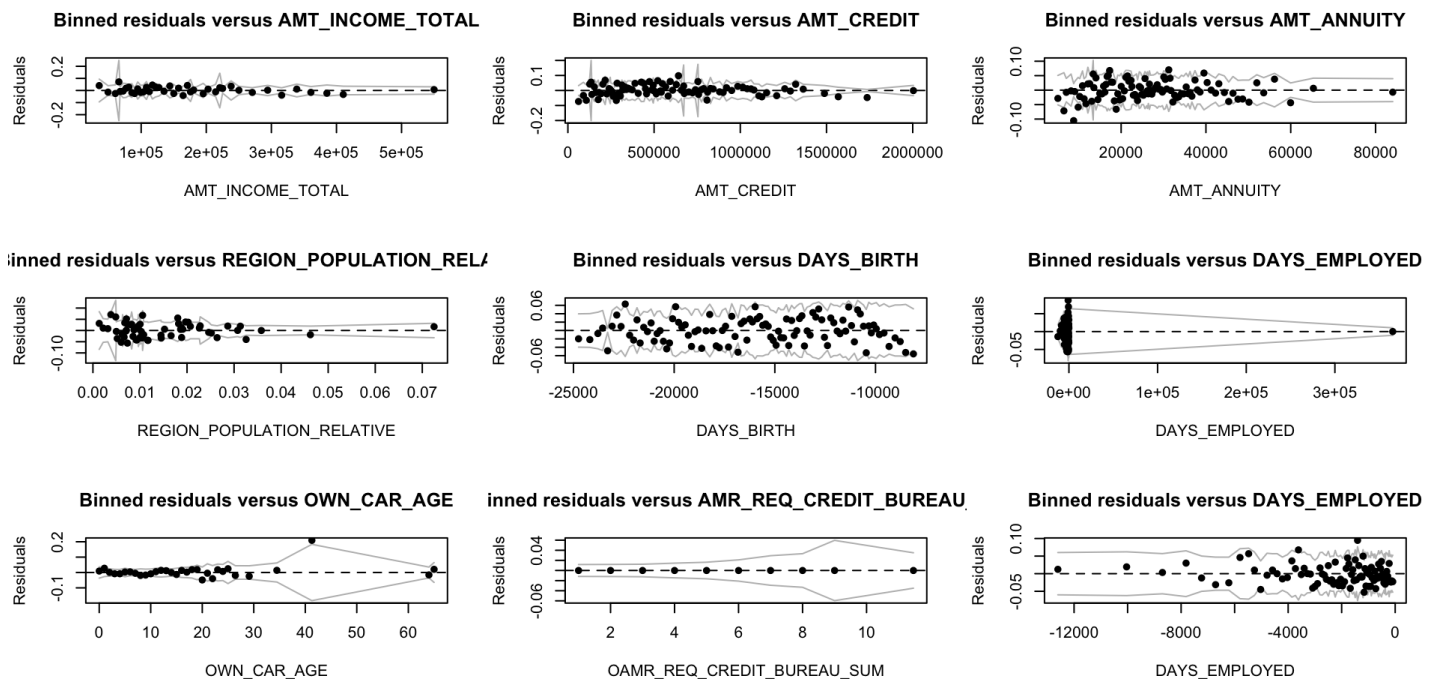
I used binned plots to check residuals to check validity of model assumptions for continuous variables. Based on plots below, there's no major violation of model assumptions, except for Binned residuals versus DAYS\_EMPLOYED. The plot is largely skewed by some points. Therefore, I plotted another residual plot to investigate if the other points form any pattern that might violate model assumptions.

```

par(mfrow = c(3,3))
#Dataset 1
cd1 <- mice::complete(application_MI, 1)
#Binned residual plots
rawresid1 = cd1$TARGET - fitted(reg_cd1)
#continuous variables
binnedplot(x=cd1$AMT_INCOME_TOTAL, y = rawresid1, xlab = "AMT_INCOME_TOTAL", ylab = "Residuals",
  main = "Binned residuals versus AMT_INCOME_TOTAL")
binnedplot(x=cd1$AMT_CREDIT, y = rawresid1, xlab = "AMT_CREDIT", ylab = "Residuals",
  main = "Binned residuals versus AMT_CREDIT")
binnedplot(x=cd1$AMT_ANNUITY, y = rawresid1, xlab = "AMT_ANNUITY", ylab = "Residuals",
  main = "Binned residuals versus AMT_ANNUITY")
binnedplot(x=cd1$REGION_POPULATION_RELATIVE, y = rawresid1, xlab = "REGION_POPULATION_RELATIVE", ylab = "Residuals",
  main = "Binned residuals versus REGION_POPULATION_RELATIVE")
binnedplot(x=cd1$DAYS_BIRTH, y = rawresid1, xlab = "DAYS_BIRTH", ylab = "Residuals",
  main = "Binned residuals versus DAYS_BIRTH")
binnedplot(x=cd1$DAYS_EMPLOYED, y = rawresid1, xlab = "DAYS_EMPLOYED", ylab = "Residuals",
  main = "Binned residuals versus DAYS_EMPLOYED")
binnedplot(x=cd1$OWN_CAR_AGE, y = rawresid1, xlab = "OWN_CAR_AGE", ylab = "Residuals",
  main = "Binned residuals versus OWN_CAR_AGE")
binnedplot(x=as.numeric(cd1$AMR_REQ_CREDIT_BUREAU_SUM), y = rawresid1, xlab = "OAMR_REQ_CREDIT_BUREAU_SUM", ylab = "Residuals",
  main = "Binned residuals versus AMR_REQ_CREDIT_BUREAU_SUM")

#DAYS_EMPLOYED
temp <- cd1[cd1$DAYS_EMPLOYED != 365243,]
binnedplot(x=temp$DAYS_EMPLOYED, y = rawresid1, xlab = "DAYS_EMPLOYED", ylab = "Residuals",
  main = "Binned residuals versus DAYS_EMPLOYED")

```



After I excluded the large value in the `DAYS_EMPLOYED` columns( `application_data_sam$DAYS_EMPLOYED == 365,243` ), the residual plot suggests no major violation of model assumptions. For those large value outliers, as they are also valid input data, I have no scientific reason to remove them.

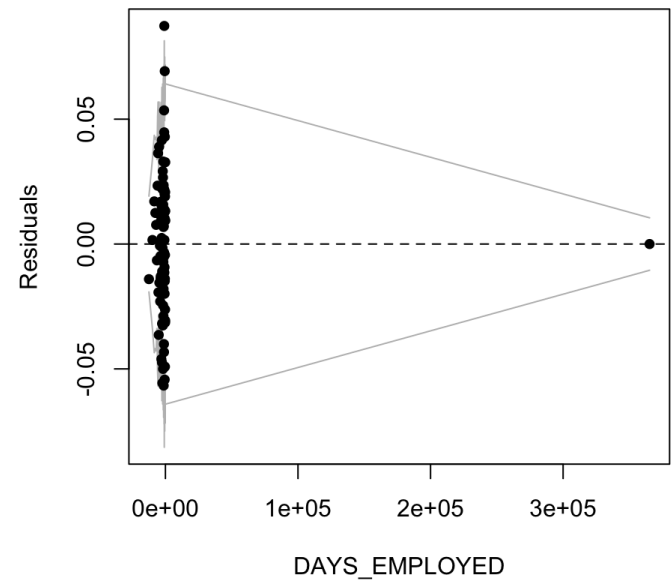
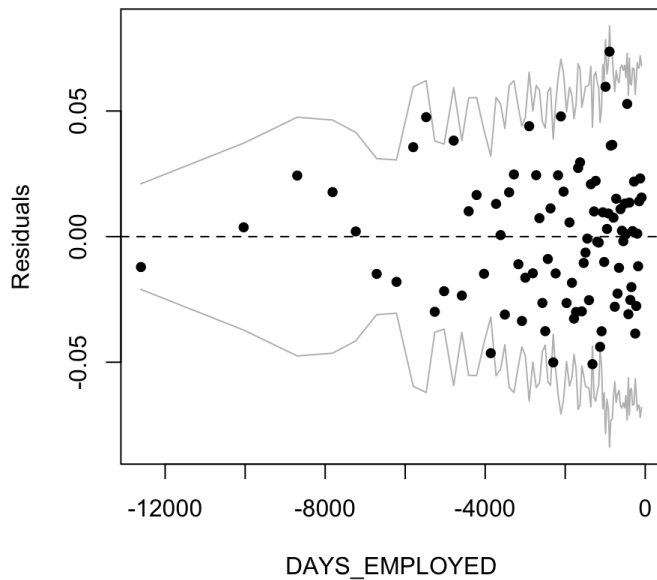
```

par(mfrow = c(1,2))
binnedplot(x=cd1$DAYS_EMPLOYED[cd1$DAYS_EMPLOYED != max(application_data_sam$DAYS_EMPLOYED)], y = rawresid1[cd1$DAYS_EMPLOYED != max(application_data_sam$DAYS_EMPLOYED)], xlab = "DAYS_EMPLOYED", ylab = "Residuals",
  main = "Binned residuals versus DAYS_EMPLOYED (without outliers)")
binnedplot(x=cd1$DAYS_EMPLOYED, y = rawresid1, xlab = "DAYS_EMPLOYED", ylab = "Residuals",
  main = "Binned residuals versus DAYS_EMPLOYED")

```

nned residuals versus DAYS\_EMPLOYED (without oi

Binned residuals versus DAYS\_EMPLOYED

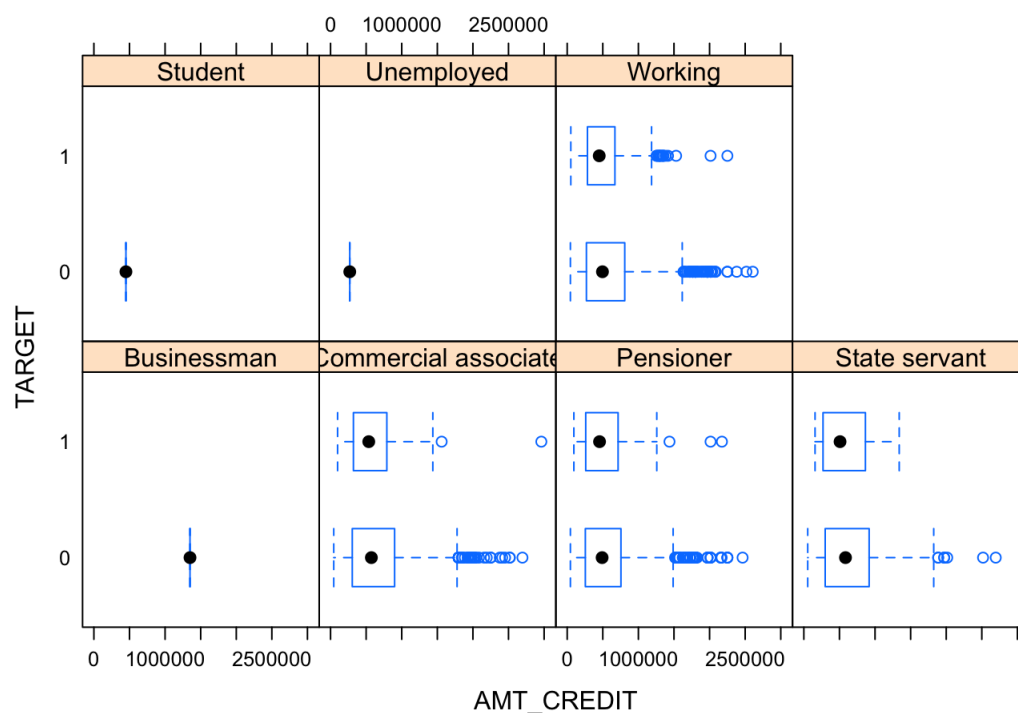


## Interaction Effect

There are several pairs of variables we would like to check for possible interaction effects based on the scientific reasons underneath.

- **NAME\_INCOME\_TYPE** versus **AMT\_CREDIT** The amount of credit a client has is typically based on his income type in the financial practice. For example, state servants' incomes are considered more stable and banks typically give them higher amounts of credits. Therefore, I believe there could be an interaction effect between these two variables. However, after taking an F-test, we could conclude that the interaction effect is not significant and I should not incorporate it into my model.

```
bwplot(as.factor(TARGET)~AMT_CREDIT|as.factor(NAME_INCOME_TYPE), data = cd1, ylab = "TARGET")
```

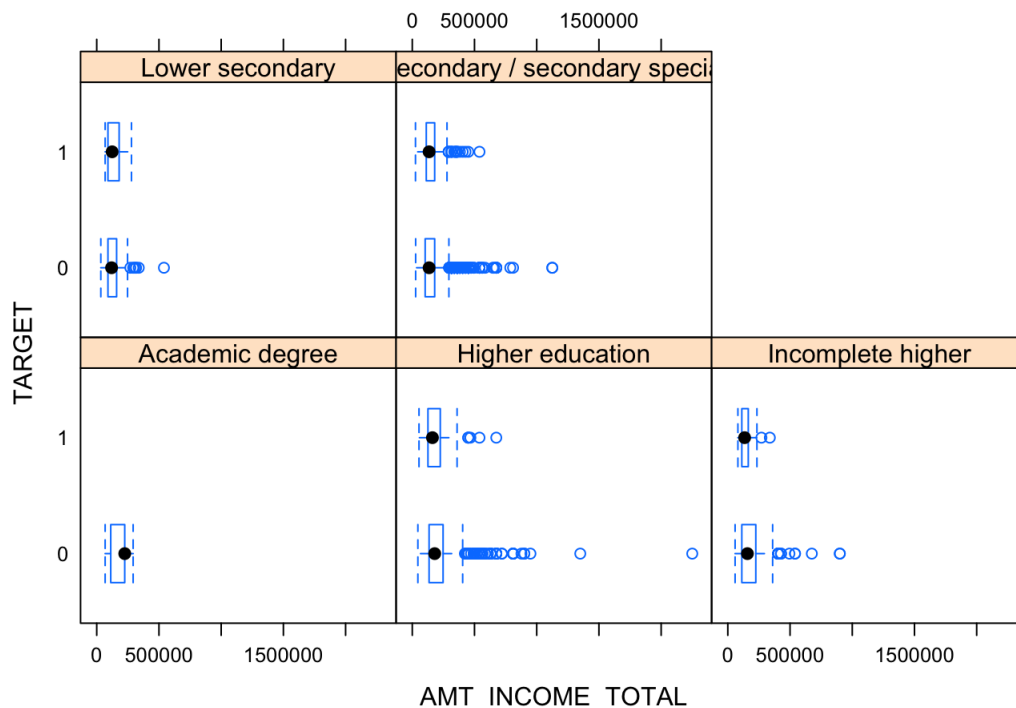


```
reg_cdl <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM, family=binomial)
reg_int <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM + AMT_CREDIT*NAME_INCOME_TYPE, family=binomial)
anova(reg_cdl, reg_int, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##   FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##   NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##   NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##   DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##   CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM
## Model 2: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##   FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##   NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##   NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##   DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##   CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM +
##   AMT_CREDIT * NAME_INCOME_TYPE
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      9863      5318.2
## 2      9860      5316.5  3    1.6452    0.6492
```

- NAME\_EDUCATION\_TYPE versus AMT\_INCOME\_TOTAL The amount of income is typically associated with the highest education he or her received. Therefore, I try to see if there's any interaction effect between these two variables. The plots below suggest that there might be an interaction effect, though not very clear. But the F-test indicates that I should not add an interaction term between these two variables in the regression model.

```
bwplot(as.factor(TARGET)~AMT_INCOME_TOTAL|as.factor(NAME_EDUCATION_TYPE), data = cd1, ylab = "TARGET")
```



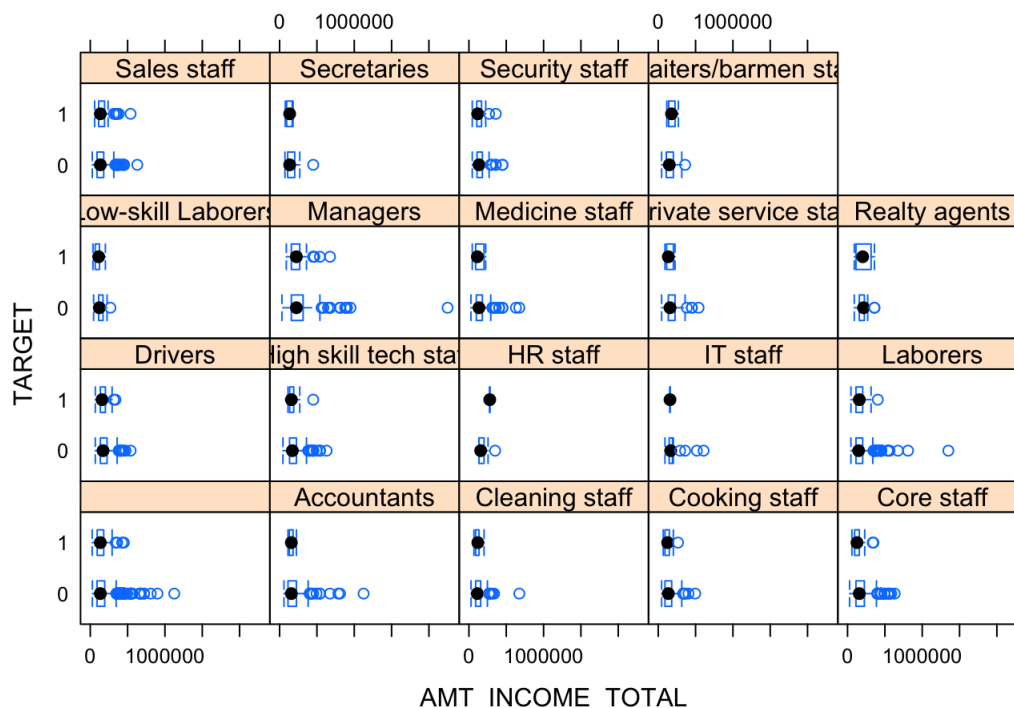


```
reg_cdl <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM, family=binomial)
reg_int <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM + AMT_INCOME_TOTAL*NAME_EDUCATION_TYPE, family=binomial)
anova(reg_cdl, reg_int, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##   FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##   NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##   NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##   DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##   CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM
## Model 2: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##   FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##   NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##   NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##   DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##   CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM +
##   AMT_INCOME_TOTAL * NAME_EDUCATION_TYPE
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      9863      5318.2
## 2      9859      5314.8  4    3.4027  0.4928
```

- **OCCUPATION\_TYPE** versus **AMT\_INCOME\_TOTAL** Different kinds of occupation generally have different levels of incomes. For example, high skill tech staffs typically have higher incomes than cleaning staffs. Hence, I think there might be an interaction effect. However, the F-test suggests that the interaction effect is not significant enough and I should not add it into the model.

```
bwplot(as.factor(TARGET)~AMT_INCOME_TOTAL|as.factor(OCCUPATION_TYPE), data = cd1, ylab = "TARGET")
```





```
reg_cdl <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM, family=binomial)
reg_int <- glm(data=cd1, TARGET~CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL + FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY + NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE + NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE + DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE + CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM + AMT_CREDIT*ORGANIZATION_TYPE, family=binomial)
anova(reg_cdl, reg_int, test = "Chisq")
```

```
## Analysis of Deviance Table
##
## Model 1: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##     FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##     NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##     NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##     DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##     CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM
## Model 2: TARGET ~ CODE_GENDER + NAME_CONTRACT_TYPE + AMT_INCOME_TOTAL +
##     FLAG_OWN_CAR + FLAG_OWN_REALTY + AMT_CREDIT + AMT_ANNUITY +
##     NAME_TYPE_SUITE + NAME_INCOME_TYPE + NAME_EDUCATION_TYPE +
##     NAME_FAMILY_STATUS + NAME_HOUSING_TYPE + REGION_POPULATION_RELATIVE +
##     DAYS_BIRTH + DAYS_EMPLOYED + OWN_CAR_AGE + OCCUPATION_TYPE +
##     CNT_FAM_MEMBERS + ORGANIZATION_TYPE + AMR_REQ_CREDIT_BUREAU_SUM +
##     AMT_CREDIT * ORGANIZATION_TYPE
##   Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      9863      5318.2
## 2      9809      5276.7 54   41.466   0.8942
```

## Model Interpretation

After examining all the possible pairs of variables that could have interaction effects; furthermore, there's no evidence suggests a logarithm or quadratic transformation on variables, therefore, the base model should be our ultimate model.

Significant predictors: - CODE\_GENDER  
M - NAME\_CONTRACT\_TYPE  
Revolving loans - FLAG\_OWN\_CAR  
Y - DAYS\_BIRTH - DAYS\_EMPLOYED - OCCUPATION\_TYPE  
Waiters/barmen staff - CNT\_FAM\_MEMBERS  
6 - ORGANIZATION\_TYPE  
Realtor - ORGANIZATION\_TYPE  
EXNA - AMR\_REQ\_CREDIT\_BUREAU\_SUM  
8

```
summary(pool(reg))
```

##	estimate
## (Intercept)	2.063259e-01
## CODE_GENDERM	4.352771e-02
## NAME_CONTRACT_TYPRevolving loans	-5.055284e-02
## AMT_INCOME_TOTAL	-5.315301e-08
## FLAG_OWN_CARY	-3.081047e-02
## FLAG_OWN_REALTYT	5.206398e-03
## AMT_CREDIT	-1.320492e-08
## AMT_ANNUITY	-1.769943e-07
## NAME_TYPE_SUITEChildren	-2.756176e-02
## NAME_TYPE_SUITEFamily	-9.752861e-03
## NAME_TYPE_SUITEGroup of people	-1.140243e-01
## NAME_TYPE_SUITEOther_A	-6.362593e-02
## NAME_TYPE_SUITEOther_B	-6.131117e-02
## NAME_TYPE_SUITESpouse, partner	-9.428628e-03
## NAME_TYPE_SUITEUnaccompanied	-1.514317e-02
## NAME_INCOME_TYPECommercial associate	-1.027763e-01
## NAME_INCOME_TYPEPensioner	-1.580174e-01
## NAME_INCOME_TYPEState servant	-1.063998e-01
## NAME_INCOME_TYPEStudent	-9.719079e-02
## NAME_INCOME_TYPEUnemployed	-2.097608e-01
## NAME_INCOME_TYPEWorking	-8.310605e-02
## NAME_EDUCATION_TYPEHigher education	3.705344e-02
## NAME_EDUCATION_TYPEIncomplete higher	5.441029e-02
## NAME_EDUCATION_TYPERLower secondary	7.767959e-02
## NAME_EDUCATION_TYPERSecondary / secondary special	5.559630e-02
## NAME_FAMILY_STATUSMarried	-6.035944e-03
## NAME_FAMILY_STATUSSeparated	-1.670707e-03
## NAME_FAMILY_STATUSSingle / not married	7.999912e-03
## NAME_FAMILY_STATUSWidow	1.727822e-02
## NAME_HOUSING_TYPEHouse / apartment	5.891942e-03
## NAME_HOUSING_TYPERMunicipal apartment	-2.051035e-03
## NAME_HOUSING_TYPEROffice apartment	-4.512994e-02
## NAME_HOUSING_TYPERRented apartment	3.482144e-02
## NAME_HOUSING_TYPERWith parents	2.526804e-02
## REGION_POPULATION_RELATIVE	-9.805655e-02
## DAYS_BIRTH	3.216099e-06
## DAYS_EMPLOYED	6.241027e-06
## OWN_CAR_AGE	1.037233e-04
## OCCUPATION_TYPEAccountants	-1.463772e-02
## OCCUPATION_TYPERCleaning staff	-1.694824e-02
## OCCUPATION_TYPERCooking staff	1.418259e-02
## OCCUPATION_TYPERCore staff	-5.402408e-03
## OCCUPATION_TYPERDrivers	1.028013e-02
## OCCUPATION_TYPERHigh skill tech staff	-1.774165e-02
## OCCUPATION_TYPERHR staff	-7.604784e-03
## OCCUPATION_TYPERIT staff	-2.806577e-02
## OCCUPATION_TYPERLaborers	-4.816502e-03
## OCCUPATION_TYPERLow-skill Laborers	2.681807e-02
## OCCUPATION_TYPERManagers	-5.676823e-03
## OCCUPATION_TYPERMedicine staff	-2.491503e-02
## OCCUPATION_TYPERPrivate service staff	-6.379935e-03
## OCCUPATION_TYPERRealty agents	7.190980e-02
## OCCUPATION_TYPERSales staff	-4.065929e-03
## OCCUPATION_TYPERSecretaries	4.072782e-02
## OCCUPATION_TYPERSecurity staff	4.084883e-02
## OCCUPATION_TYPERWaiters/barmen staff	1.002301e-01
## CNT_FAM_MEMBERS2	1.132116e-03
## CNT_FAM_MEMBERS3	-2.320473e-03
## CNT_FAM_MEMBERS4	-5.599006e-03
## CNT_FAM_MEMBERS5	3.401739e-02
## CNT_FAM_MEMBERS6	1.794479e-01
## CNT_FAM_MEMBERS7	8.906367e-02
## ORGANIZATION_TYPERAgriculture	4.382414e-03
## ORGANIZATION_TYPERBank	8.570364e-03
## ORGANIZATION_TYPERBusiness Entity Type 1	-1.214637e-03
## ORGANIZATION_TYPERBusiness Entity Type 2	3.669097e-02

## ORGANIZATION_TYPEBusiness Entity Type 3	8.953085e-03
## ORGANIZATION_TYPECleaning	3.671069e-02
## ORGANIZATION_TYPEConstruction	2.132647e-02
## ORGANIZATION_TYPECulture	-5.861336e-02
## ORGANIZATION_TYPEElectricity	-1.564123e-02
## ORGANIZATION_TYPEEmergency	-2.077316e-02
## ORGANIZATION_TYPEGovernment	1.162231e-03
## ORGANIZATION_TYPEHotel	-6.344843e-02
## ORGANIZATION_TYPEHousing	5.454882e-03
## ORGANIZATION_TYPEIndustry: type 1	-3.756790e-02
## ORGANIZATION_TYPEIndustry: type 11	1.558636e-02
## ORGANIZATION_TYPEIndustry: type 12	9.957355e-02
## ORGANIZATION_TYPEIndustry: type 13	-1.205705e-01
## ORGANIZATION_TYPEIndustry: type 2	-7.879773e-02
## ORGANIZATION_TYPEIndustry: type 3	4.168363e-02
## ORGANIZATION_TYPEIndustry: type 4	1.304441e-02
## ORGANIZATION_TYPEIndustry: type 5	1.303146e-02
## ORGANIZATION_TYPEIndustry: type 6	-7.803666e-02
## ORGANIZATION_TYPEIndustry: type 7	2.541053e-02
## ORGANIZATION_TYPEIndustry: type 9	3.122153e-02
## ORGANIZATION_TYPEInsurance	-7.338810e-02
## ORGANIZATION_TYPEKindergarten	-3.629599e-03
## ORGANIZATION_TYPELegal Services	1.982034e-02
## ORGANIZATION_TYPMedicine	3.709787e-02
## ORGANIZATION_TYPMilitary	-3.655942e-02
## ORGANIZATION_TYPMobile	-9.156294e-02
## ORGANIZATION_TYPEOther	1.135235e-03
## ORGANIZATION_TYPEPolice	-3.770840e-02
## ORGANIZATION_TYPEPostal	-6.280312e-03
## ORGANIZATION_TYPERealtor	2.169619e-01
## ORGANIZATION_TYPEReligion	-3.296117e-02
## ORGANIZATION_TYPERestaurant	-1.767501e-02
## ORGANIZATION_TYPSchool	8.833369e-04
## ORGANIZATION_TYPSecurity	-3.087008e-02
## ORGANIZATION_TYPSecurity Ministries	-3.000705e-03
## ORGANIZATION_TYPSelf-employed	1.969723e-02
## ORGANIZATION_TYPServices	-1.091431e-02
## ORGANIZATION_TYPTelecom	-2.065610e-02
## ORGANIZATION_TYPTTrade: type 1	1.810593e-02
## ORGANIZATION_TYPTTrade: type 2	-1.085085e-03
## ORGANIZATION_TYPTTrade: type 3	-1.529725e-02
## ORGANIZATION_TYPTTrade: type 4	-1.012080e-01
## ORGANIZATION_TYPTTrade: type 6	-9.123168e-03
## ORGANIZATION_TYPTTrade: type 7	3.726542e-02
## ORGANIZATION_TYPTTransport: type 1	-7.505025e-02
## ORGANIZATION_TYPTTransport: type 2	7.819173e-02
## ORGANIZATION_TYPTTransport: type 3	1.384465e-01
## ORGANIZATION_TYPTTransport: type 4	-3.459733e-02
## ORGANIZATION_TYPEUniversity	-3.775120e-04
## ORGANIZATION_TYPEXNA	-2.240731e+00
## AMR_REQ_CREDIT_BUREAU_SUM1	4.190379e-03
## AMR_REQ_CREDIT_BUREAU_SUM2	1.092473e-02
## AMR_REQ_CREDIT_BUREAU_SUM3	1.288864e-02
## AMR_REQ_CREDIT_BUREAU_SUM4	3.119740e-03
## AMR_REQ_CREDIT_BUREAU_SUM5	-7.012164e-03
## AMR_REQ_CREDIT_BUREAU_SUM6	2.503563e-02
## AMR_REQ_CREDIT_BUREAU_SUM7	1.588082e-02
## AMR_REQ_CREDIT_BUREAU_SUM8	5.843887e-02
## AMR_REQ_CREDIT_BUREAU_SUM9	1.691654e-02
## AMR_REQ_CREDIT_BUREAU_SUM10	-1.849285e-02
## AMR_REQ_CREDIT_BUREAU_SUM11	3.053713e-03
## AMR_REQ_CREDIT_BUREAU_SUM12	-4.140642e-02
## AMR_REQ_CREDIT_BUREAU_SUM13	-7.344946e-02
## AMR_REQ_CREDIT_BUREAU_SUM14	-1.746657e-02
## AMR_REQ_CREDIT_BUREAU_SUM15	-8.302970e-02
## AMR_REQ_CREDIT_BUREAU_SUM16	-2.768300e-02
## AMR_REQ_CREDIT_BUREAU_SUM17	-9.937823e-02
## AMR_REQ_CREDIT_BUREAU_SUM18	-7.341512e-02

## AMR_REQ_CREDIT_BUREAU_SUM19	-1.680945e-02	
## AMR_REQ_CREDIT_BUREAU_SUM20	-5.056997e-02	
## AMR_REQ_CREDIT_BUREAU_SUM28	-1.117664e-01	
##	std.error	statistic
## (Intercept)	3.238824e-01	0.637039695
## CODE_GENDERM	7.349921e-03	5.922200671
## NAME_CONTRACT_TYPERevolving loans	9.804538e-03	-5.156065232
## AMT_INCOME_TOTAL	3.806233e-08	-1.396472747
## FLAG_OWN_CARY	6.335356e-03	-4.863258282
## FLAG_OWN_REALTY	6.208820e-03	0.838548695
## AMT_CREDIT	1.107826e-08	-1.191966886
## AMT_ANNUITY	3.253000e-07	-0.544095497
## NAME_TYPE_SUITEChildren	5.078004e-02	-0.542767609
## NAME_TYPE_SUITEFamily	4.468996e-02	-0.218233848
## NAME_TYPE_SUITEGroup of people	1.298177e-01	-0.878341280
## NAME_TYPE_SUITEOther_A	7.422377e-02	-0.857217725
## NAME_TYPE_SUITEOther_B	5.868265e-02	-1.044792141
## NAME_TYPE_SUITESpouse, partner	4.628181e-02	-0.203722120
## NAME_TYPE_SUITEUnaccompanied	4.412372e-02	-0.343197977
## NAME_INCOME_TYPECommercial associate	2.844942e-01	-0.361259798
## NAME_INCOME_TYPEPensioner	3.946596e-01	-0.400389083
## NAME_INCOME_TYPEState servant	2.848237e-01	-0.373563697
## NAME_INCOME_TYPEStudent	3.901416e-01	-0.249116717
## NAME_INCOME_TYPEUnemployed	4.799008e-01	-0.437092064
## NAME_INCOME_TYPEWorking	2.846590e-01	-0.291949486
## NAME_EDUCATION_TYPEHigher education	1.222726e-01	0.303039512
## NAME_EDUCATION_TYPEIncomplete higher	1.231872e-01	0.441687961
## NAME_EDUCATION_TYPELower secondary	1.248090e-01	0.622387764
## NAME_EDUCATION_Typesecondary / secondary special	1.221992e-01	0.454964697
## NAME_FAMILY_STATUSMarried	9.324405e-03	-0.647327578
## NAME_FAMILY_STATUSSeparated	1.701766e-02	-0.098174855
## NAME_FAMILY_STATUSSingle / not married	1.680871e-02	0.475938598
## NAME_FAMILY_STATUSWidow	2.034232e-02	0.849372761
## NAME_HOUSING_TYPEHouse / apartment	4.101925e-02	0.143638448
## NAME_HOUSING_TYEMunicipal apartment	4.344237e-02	-0.047212774
## NAME_HOUSING_TYPEOffice apartment	5.133026e-02	-0.879207432
## NAME_HOUSING_TYERented apartment	4.573693e-02	0.761341852
## NAME_HOUSING_TYEWith parents	4.277699e-02	0.590692396
## REGION_POPULATION_RELATIVE	2.097866e-01	-0.467410822
## DAYS_BIRTH	9.715712e-07	3.310204273
## DAYS_EMPLOYED	1.476965e-06	4.225575963
## OWN_CAR_AGE	2.801423e-04	0.370252048
## OCCUPATION_TYPEAccountants	1.778898e-02	-0.822853488
## OCCUPATION_TYPECleaning staff	2.441296e-02	-0.694231084
## OCCUPATION_TYPECooking staff	2.101915e-02	0.674746192
## OCCUPATION_TYPECore staff	1.324128e-02	-0.407997419
## OCCUPATION_TYEDrivers	1.444968e-02	0.711443845
## OCCUPATION_TYPEHigh skill tech staff	1.559747e-02	-1.137469734
## OCCUPATION_TYEHHR staff	7.360465e-02	-0.103319349
## OCCUPATION_TYPEIT staff	6.037217e-02	-0.464879335
## OCCUPATION_TYELaborers	1.040000e-02	-0.463125263
## OCCUPATION_TYELow-skill Laborers	3.512250e-02	0.763558124
## OCCUPATION_TYEManagers	1.342852e-02	-0.422743750
## OCCUPATION_TYEMedicine staff	2.288624e-02	-1.088646858
## OCCUPATION_TYEPprivate service staff	3.046778e-02	-0.209399389
## OCCUPATION_TYERealty agents	5.679346e-02	1.266163399
## OCCUPATION_TYESales staff	1.227742e-02	-0.331171413
## OCCUPATION_TYESecretaries	4.763194e-02	0.855052630
## OCCUPATION_TYESecurity staff	2.237480e-02	1.825662323
## OCCUPATION_TYEWaiters/barmen staff	4.716424e-02	2.125129651
## CNT_FAM_MEMBERS2	1.464121e-02	0.077323907
## CNT_FAM_MEMBERS3	1.639724e-02	-0.141516061
## CNT_FAM_MEMBERS4	1.817850e-02	-0.308001501
## CNT_FAM_MEMBERS5	3.137761e-02	1.084129310
## CNT_FAM_MEMBERS6	8.416156e-02	2.132183844
## CNT_FAM_MEMBERS7	1.237411e-01	0.719758220
## ORGANIZATION_TYPEAgriculture	7.110094e-02	0.061636519
## ORGANIZATION_TYEBank	7.345444e-02	0.116675911

## ORGANIZATION_TYPEBusiness Entity Type 1	6.733533e-02	-0.018038620
## ORGANIZATION_TYPEBusiness Entity Type 2	6.608540e-02	0.555205375
## ORGANIZATION_TYPEBusiness Entity Type 3	6.467652e-02	0.138428667
## ORGANIZATION_TYPECleaning	1.114015e-01	0.329534866
## ORGANIZATION_TYPEConstruction	6.707111e-02	0.317968028
## ORGANIZATION_TYPECulture	9.999669e-02	-0.586153024
## ORGANIZATION_TYPEElectricity	8.055834e-02	-0.194160226
## ORGANIZATION_TYPEEmergency	8.792009e-02	-0.236273141
## ORGANIZATION_TYPEGovernment	6.604658e-02	0.017597149
## ORGANIZATION_TYPEHotel	8.257608e-02	-0.768363320
## ORGANIZATION_TYPEHousing	6.982835e-02	0.078118440
## ORGANIZATION_TYPEIndustry: type 1	7.841714e-02	-0.479077683
## ORGANIZATION_TYPEIndustry: type 11	7.098229e-02	0.219580928
## ORGANIZATION_TYPEIndustry: type 12	1.020855e-01	0.975393521
## ORGANIZATION_TYPEIndustry: type 13	2.036842e-01	-0.591948495
## ORGANIZATION_TYPEIndustry: type 2	9.731529e-02	-0.809715791
## ORGANIZATION_TYPEIndustry: type 3	6.914345e-02	0.602857174
## ORGANIZATION_TYPEIndustry: type 4	7.817763e-02	0.166856096
## ORGANIZATION_TYPEIndustry: type 5	8.878250e-02	0.146779649
## ORGANIZATION_TYPEIndustry: type 6	1.380145e-01	-0.565423688
## ORGANIZATION_TYPEIndustry: type 7	7.682659e-02	0.330751780
## ORGANIZATION_TYPEIndustry: type 9	6.945843e-02	0.449499499
## ORGANIZATION_TYPEInsurance	9.119419e-02	-0.804745374
## ORGANIZATION_TYPEKindergarten	6.709661e-02	-0.054095110
## ORGANIZATION_TYPELegal Services	9.761973e-02	0.203036166
## ORGANIZATION_TYPEMedicine	6.714712e-02	0.552486377
## ORGANIZATION_TYPEMilitary	7.111523e-02	-0.514086997
## ORGANIZATION_TYPEMobile	1.046480e-01	-0.874961544
## ORGANIZATION_TYPEOther	6.550997e-02	0.017329199
## ORGANIZATION_TYPEPolice	7.237334e-02	-0.521026129
## ORGANIZATION_TYPEPostal	7.224699e-02	-0.086928360
## ORGANIZATION_TYPERealtor	1.045732e-01	2.074736910
## ORGANIZATION_TYPEReligion	2.033270e-01	-0.162109178
## ORGANIZATION_TYPERestaurant	7.372632e-02	-0.239738171
## ORGANIZATION_TYPSchool	6.659657e-02	0.013263999
## ORGANIZATION_TYPESecurity	7.165784e-02	-0.430798349
## ORGANIZATION_TYPESecurity Ministries	7.303849e-02	-0.041083890
## ORGANIZATION_TYPESelf-employed	6.488590e-02	0.303567190
## ORGANIZATION_TYPEServices	7.689387e-02	-0.141939911
## ORGANIZATION_TYPETelecom	8.996517e-02	-0.229601124
## ORGANIZATION_TYPETrade: type 1	1.045225e-01	0.173225100
## ORGANIZATION_TYPETrade: type 2	7.420166e-02	-0.014623457
## ORGANIZATION_TYPETrade: type 3	6.905253e-02	-0.221530686
## ORGANIZATION_TYPETrade: type 4	2.036165e-01	-0.497052183
## ORGANIZATION_TYPETrade: type 6	8.610919e-02	-0.105948836
## ORGANIZATION_TYPETrade: type 7	6.674541e-02	0.558321819
## ORGANIZATION_TYPETransport: type 1	1.510872e-01	-0.496734785
## ORGANIZATION_TYPETransport: type 2	7.187849e-02	1.087832087
## ORGANIZATION_TYPETransport: type 3	7.668451e-02	1.805403139
## ORGANIZATION_TYPETransport: type 4	6.837164e-02	-0.506018664
## ORGANIZATION_TYPEUniversity	7.928623e-02	-0.004761382
## ORGANIZATION_TYPEXNA	6.089517e-01	-3.679653516
## AMR_REQ_CREDIT_BUREAU_SUM1	9.127409e-03	0.459098434
## AMR_REQ_CREDIT_BUREAU_SUM2	9.816924e-03	1.112846530
## AMR_REQ_CREDIT_BUREAU_SUM3	1.053364e-02	1.223569960
## AMR_REQ_CREDIT_BUREAU_SUM4	1.161135e-02	0.268680246
## AMR_REQ_CREDIT_BUREAU_SUM5	1.360600e-02	-0.515373121
## AMR_REQ_CREDIT_BUREAU_SUM6	1.725497e-02	1.450922280
## AMR_REQ_CREDIT_BUREAU_SUM7	1.898283e-02	0.836588629
## AMR_REQ_CREDIT_BUREAU_SUM8	2.739782e-02	2.132974918
## AMR_REQ_CREDIT_BUREAU_SUM9	3.608539e-02	0.468792048
## AMR_REQ_CREDIT_BUREAU_SUM10	4.468145e-02	-0.413882025
## AMR_REQ_CREDIT_BUREAU_SUM11	7.004005e-02	0.043599527
## AMR_REQ_CREDIT_BUREAU_SUM12	1.057255e-01	-0.391640785
## AMR_REQ_CREDIT_BUREAU_SUM13	1.157125e-01	-0.634758237
## AMR_REQ_CREDIT_BUREAU_SUM14	1.799169e-01	-0.097081323
## AMR_REQ_CREDIT_BUREAU_SUM15	1.321943e-01	-0.628088210
## AMR_REQ_CREDIT_BUREAU_SUM16	2.742893e-01	-0.100926295

## AMR_REQ_CREDIT_BUREAU_SUM17	2.726974e-01	-0.364426748
## AMR_REQ_CREDIT_BUREAU_SUM18	1.840078e-01	-0.398978291
## AMR_REQ_CREDIT_BUREAU_SUM19	2.746293e-01	-0.061207775
## AMR_REQ_CREDIT_BUREAU_SUM20	2.562267e-01	-0.197364150
## AMR_REQ_CREDIT_BUREAU_SUM28	2.599306e-01	-0.429985680
##	df	p.value
## (Intercept)	9842.13067	5.241138e-01
## CODE_GENDERM	9847.07017	3.282767e-09
## NAME_CONTRACT_TYPRevolving loans	9667.51517	2.570858e-07
## AMT_INCOME_TOTAL	9030.00715	1.626036e-01
## FLAG_OWN_CARY	9859.54073	1.172539e-06
## FLAG_OWN_REALTYT	9757.98866	4.017429e-01
## AMT_CREDIT	9807.46426	2.333029e-01
## AMT_ANNUITY	9841.48892	5.863881e-01
## NAME_TYPE_SUITEChildren	9835.74541	5.873021e-01
## NAME_TYPE_SUITEFamily	9851.82268	8.272514e-01
## NAME_TYPE_SUITEGroup of people	9854.89304	3.797799e-01
## NAME_TYPE_SUITEOther_A	9856.74596	3.913454e-01
## NAME_TYPE_SUITEOther_B	9853.70456	2.961447e-01
## NAME_TYPE_SUITESpouse, partner	9854.28935	8.385749e-01
## NAME_TYPE_SUITEUnaccompanied	9846.19867	7.314568e-01
## NAME_INCOME_TYPECommercial associate	9845.81218	7.179130e-01
## NAME_INCOME_TYPEPensioner	9850.37841	6.888786e-01
## NAME_INCOME_TYPEState servant	9846.11085	7.087370e-01
## NAME_INCOME_TYPEStudent	9858.09436	8.032757e-01
## NAME_INCOME_TYPEUnemployed	9852.56370	6.620542e-01
## NAME_INCOME_TYPEReWorking	9845.49705	7.703314e-01
## NAME_EDUCATION_TYPEHigher education	9857.85728	7.618661e-01
## NAME_EDUCATION_TYPEIncomplete higher	9857.92865	6.587247e-01
## NAME_EDUCATION_TYPEReLower secondary	9856.82413	5.337013e-01
## NAME_EDUCATION_TYPEReSecondary / secondary special	9858.02816	6.491447e-01
## NAME_FAMILY_STATUSSMarried	9851.19907	5.174350e-01
## NAME_FAMILY_STATUSSSeparated	9852.71466	9.217954e-01
## NAME_FAMILY_STATUSSSingle / not married	9855.64716	6.341287e-01
## NAME_FAMILY_STATUSSWidow	9801.82300	3.956945e-01
## NAME_HOUSING_TYPEReHouse / apartment	9850.15576	8.857889e-01
## NAME_HOUSING_TYPEReMunicipal apartment	9852.04325	9.623446e-01
## NAME_HOUSING_TYPEReOffice apartment	9855.89066	3.793102e-01
## NAME_HOUSING_TYPEReRented apartment	9857.25008	4.464711e-01
## NAME_HOUSING_TYPERewith parents	9853.18499	5.547401e-01
## REGION_POPULATION_RELATIVE	9810.09843	6.402163e-01
## DAYS_BIRTH	9852.36184	9.356256e-04
## DAYS_EMPLOYED	9851.66253	2.404757e-05
## OWN_CAR_AGE	62.44924	7.112027e-01
## OCCUPATION_TYPEReAccountants	9795.14360	4.106112e-01
## OCCUPATION_TYPEReCleaning staff	9849.07232	4.875537e-01
## OCCUPATION_TYPEReCooking staff	9859.54353	4.998529e-01
## OCCUPATION_TYPEReCore staff	9857.26052	6.832844e-01
## OCCUPATION_TYPEReDrivers	9849.62182	4.768261e-01
## OCCUPATION_TYPEReHigh skill tech staff	9857.56269	2.553696e-01
## OCCUPATION_TYPEReHR staff	9845.41388	9.177116e-01
## OCCUPATION_TYPEReIT staff	9852.56666	6.420281e-01
## OCCUPATION_TYPEReLaborers	9838.74883	6.432848e-01
## OCCUPATION_TYPEReLow-skill Laborers	9839.07310	4.451489e-01
## OCCUPATION_TYPEReManagers	9854.13464	6.724915e-01
## OCCUPATION_TYPEReMedicine staff	9845.39368	2.763363e-01
## OCCUPATION_TYPERePrivate service staff	9857.24427	8.341408e-01
## OCCUPATION_TYPEReRealty agents	9858.84639	2.054845e-01
## OCCUPATION_TYPEReSales staff	9859.23310	7.405221e-01
## OCCUPATION_TYPEReSecretaries	9855.14485	3.925428e-01
## OCCUPATION_TYPEReSecurity staff	9850.20513	6.793137e-02
## OCCUPATION_TYPEReWaiters/barmen staff	9857.58809	3.360059e-02
## CNT_FAM_MEMBERS2	9857.49802	9.383674e-01
## CNT_FAM_MEMBERS3	9858.30430	8.874652e-01
## CNT_FAM_MEMBERS4	9859.79692	7.580877e-01
## CNT_FAM_MEMBERS5	9859.60681	2.783340e-01
## CNT_FAM_MEMBERS6	9859.79036	3.301638e-02
## CNT_FAM_MEMBERS7	9859.90649	4.716909e-01



## ORGANIZATION_TYPEAgriculture	9851.96946	9.508535e-01
## ORGANIZATION_TYPEBank	9853.09656	9.071193e-01
## ORGANIZATION_TYPEBusiness Entity Type 1	9849.95852	9.856084e-01
## ORGANIZATION_TYPEBusiness Entity Type 2	9852.76155	5.787668e-01
## ORGANIZATION_TYPEBusiness Entity Type 3	9849.04393	8.899045e-01
## ORGANIZATION_TYPECleaning	9852.78481	7.417584e-01
## ORGANIZATION_TYPEConstruction	9846.97936	7.505159e-01
## ORGANIZATION_TYPECulture	9855.28352	5.577861e-01
## ORGANIZATION_TYPEElectricity	9855.34649	8.460544e-01
## ORGANIZATION_TYPEEmergency	9854.58245	8.132256e-01
## ORGANIZATION_TYPEGovernment	9849.30070	9.859606e-01
## ORGANIZATION_TYPEHotel	9856.82784	4.422898e-01
## ORGANIZATION_TYPEHousing	9853.55464	9.377354e-01
## ORGANIZATION_TYPEIndustry: type 1	9849.99755	6.318940e-01
## ORGANIZATION_TYPEIndustry: type 11	9848.57262	8.262021e-01
## ORGANIZATION_TYPEIndustry: type 12	9856.01441	3.293890e-01
## ORGANIZATION_TYPEIndustry: type 13	9852.40309	5.538987e-01
## ORGANIZATION_TYPEIndustry: type 2	9856.14096	4.181231e-01
## ORGANIZATION_TYPEIndustry: type 3	9849.73606	5.466176e-01
## ORGANIZATION_TYPEIndustry: type 4	9852.96031	8.674867e-01
## ORGANIZATION_TYPEIndustry: type 5	9855.15334	8.833089e-01
## ORGANIZATION_TYPEIndustry: type 6	9857.90329	5.717985e-01
## ORGANIZATION_TYPEIndustry: type 7	9848.23792	7.408390e-01
## ORGANIZATION_TYPEIndustry: type 9	9851.75824	6.530812e-01
## ORGANIZATION_TYPEInsurance	9853.96224	4.209860e-01
## ORGANIZATION_TYPEKindergarten	9850.80883	9.568605e-01
## ORGANIZATION_TYPELegal Services	9853.80085	8.391110e-01
## ORGANIZATION_TYEMedicine	9851.10286	5.806277e-01
## ORGANIZATION_TYEMilitary	9853.58736	6.072027e-01
## ORGANIZATION_TYEMobile	9857.00555	3.816161e-01
## ORGANIZATION_TYPEOther	9850.35832	9.861743e-01
## ORGANIZATION_TYPEPolice	9852.53848	6.023603e-01
## ORGANIZATION_TYPEPostal	9845.47055	9.307302e-01
## ORGANIZATION_TYERealtor	9857.75687	3.803680e-02
## ORGANIZATION_TYEReligion	9858.83227	8.712232e-01
## ORGANIZATION_TYERestaurant	9849.14060	8.105382e-01
## ORGANIZATION_TYESchool	9850.11849	9.894174e-01
## ORGANIZATION_TYESecurity	9851.16020	6.666244e-01
## ORGANIZATION_TYESecurity Ministries	9849.88328	9.672298e-01
## ORGANIZATION_TYESelf-employed	9850.71347	7.614641e-01
## ORGANIZATION_TYEServices	9855.21952	8.871304e-01
## ORGANIZATION_TYETelecom	9855.53402	8.184065e-01
## ORGANIZATION_TYETrade: type 1	9847.12069	8.624781e-01
## ORGANIZATION_TYETrade: type 2	9850.05884	9.883329e-01
## ORGANIZATION_TYETrade: type 3	9842.19071	8.246838e-01
## ORGANIZATION_TYETrade: type 4	9858.41253	6.191633e-01
## ORGANIZATION_TYETrade: type 6	9850.30181	9.156251e-01
## ORGANIZATION_TYETrade: type 7	9848.48774	5.766373e-01
## ORGANIZATION_TYETransport: type 1	9857.65978	6.193872e-01
## ORGANIZATION_TYETransport: type 2	9852.87199	2.766958e-01
## ORGANIZATION_TYETransport: type 3	9854.21526	7.104211e-02
## ORGANIZATION_TYETransport: type 4	9849.40373	6.128549e-01
## ORGANIZATION_TYPEUniversity	9851.63858	9.962011e-01
## ORGANIZATION_TYPEXNA	9855.66774	2.347956e-04
## AMR_REQ_CREDIT_BUREAU_SUM1	1417.16708	6.461736e-01
## AMR_REQ_CREDIT_BUREAU_SUM2	250.28318	2.658015e-01
## AMR_REQ_CREDIT_BUREAU_SUM3	260.84440	2.211437e-01
## AMR_REQ_CREDIT_BUREAU_SUM4	311.38428	7.881814e-01
## AMR_REQ_CREDIT_BUREAU_SUM5	499.95796	6.063039e-01
## AMR_REQ_CREDIT_BUREAU_SUM6	172.24968	1.468333e-01
## AMR_REQ_CREDIT_BUREAU_SUM7	521.47606	4.028441e-01
## AMR_REQ_CREDIT_BUREAU_SUM8	302.47679	3.295141e-02
## AMR_REQ_CREDIT_BUREAU_SUM9	276.96744	6.392286e-01
## AMR_REQ_CREDIT_BUREAU_SUM10	6861.17223	6.789695e-01
## AMR_REQ_CREDIT_BUREAU_SUM11	1787.91692	9.652245e-01
## AMR_REQ_CREDIT_BUREAU_SUM12	9659.66009	6.953321e-01
## AMR_REQ_CREDIT_BUREAU_SUM13	9710.15537	5.256008e-01
## AMR_REQ_CREDIT_BUREAU_SUM14	9819.14406	9.226638e-01

## AMR_REQ_CREDIT_BUREAU_SUM15	9501.72835	5.299607e-01
## AMR_REQ_CREDIT_BUREAU_SUM16	9859.90649	9.196110e-01
## AMR_REQ_CREDIT_BUREAU_SUM17	9859.35814	7.155472e-01
## AMR_REQ_CREDIT_BUREAU_SUM18	9491.65927	6.899178e-01
## AMR_REQ_CREDIT_BUREAU_SUM19	9859.90649	9.511950e-01
## AMR_REQ_CREDIT_BUREAU_SUM20	9858.65550	8.435467e-01
## AMR_REQ_CREDIT_BUREAU_SUM28	9833.27900	6.672155e-01

```
summary(pool(reg), conf.int=T)
```

##	estimate
## (Intercept)	2.063259e-01
## CODE_GENDERM	4.352771e-02
## NAME_CONTRACT_TYPRevolving loans	-5.055284e-02
## AMT_INCOME_TOTAL	-5.315301e-08
## FLAG_OWN_CARY	-3.081047e-02
## FLAG_OWN_REALTYT	5.206398e-03
## AMT_CREDIT	-1.320492e-08
## AMT_ANNUITY	-1.769943e-07
## NAME_TYPE_SUITEChildren	-2.756176e-02
## NAME_TYPE_SUITEFamily	-9.752861e-03
## NAME_TYPE_SUITEGroup of people	-1.140243e-01
## NAME_TYPE_SUITEOther_A	-6.362593e-02
## NAME_TYPE_SUITEOther_B	-6.131117e-02
## NAME_TYPE_SUITESpouse, partner	-9.428628e-03
## NAME_TYPE_SUITEUnaccompanied	-1.514317e-02
## NAME_INCOME_TYPECommercial associate	-1.027763e-01
## NAME_INCOME_TYPEPensioner	-1.580174e-01
## NAME_INCOME_TYPEState servant	-1.063998e-01
## NAME_INCOME_TYPEStudent	-9.719079e-02
## NAME_INCOME_TYPEUnemployed	-2.097608e-01
## NAME_INCOME_TYPEReading	-8.310605e-02
## NAME_EDUCATION_TYPEHigher education	3.705344e-02
## NAME_EDUCATION_TYPEIncomplete higher	5.441029e-02
## NAME_EDUCATION_TYPEReading secondary	7.767959e-02
## NAME_EDUCATION_TYPEReading secondary / secondary special	5.559630e-02
## NAME_FAMILY_STATUSSeparated	-6.035944e-03
## NAME_FAMILY_STATUSSeparated	-1.670707e-03
## NAME_FAMILY_STATUSSingle / not married	7.999912e-03
## NAME_FAMILY_STATUSSingle	1.727822e-02
## NAME_HOUSING_TYPEReading / apartment	5.891942e-03
## NAME_HOUSING_TYPEReading apartment	-2.051035e-03
## NAME_HOUSING_TYPEReading apartment	-4.512994e-02
## NAME_HOUSING_TYPEReading apartment	3.482144e-02
## NAME_HOUSING_TYPEReading parents	2.526804e-02
## REGION_POPULATION_RELATIVE	-9.805655e-02
## DAYS_BIRTH	3.216099e-06
## DAYS_EMPLOYED	6.241027e-06
## OWN_CAR_AGE	1.037233e-04
## OCCUPATION_TYPEReading staff	-1.463772e-02
## OCCUPATION_TYPEReading staff	-1.694824e-02
## OCCUPATION_TYPEReading staff	1.418259e-02
## OCCUPATION_TYPEReading staff	-5.402408e-03
## OCCUPATION_TYPEReading staff	1.028013e-02
## OCCUPATION_TYPEReading skill tech staff	-1.774165e-02
## OCCUPATION_TYPEReading staff	-7.604784e-03
## OCCUPATION_TYPEReading staff	-2.806577e-02
## OCCUPATION_TYPEReading staff	-4.816502e-03
## OCCUPATION_TYPEReading staff	2.681807e-02
## OCCUPATION_TYPEReading staff	-5.676823e-03
## OCCUPATION_TYPEReading staff	-2.491503e-02
## OCCUPATION_TYPEReading service staff	-6.379935e-03
## OCCUPATION_TYPEReading agents	7.190980e-02
## OCCUPATION_TYPEReading staff	-4.065929e-03
## OCCUPATION_TYPEReading staff	4.072782e-02
## OCCUPATION_TYPEReading staff	4.084883e-02
## OCCUPATION_TYPEReading staff	1.002301e-01
## CNT_FAM_MEMBERS2	1.132116e-03
## CNT_FAM_MEMBERS3	-2.320473e-03
## CNT_FAM_MEMBERS4	-5.599006e-03
## CNT_FAM_MEMBERS5	3.401739e-02
## CNT_FAM_MEMBERS6	1.794479e-01
## CNT_FAM_MEMBERS7	8.906367e-02
## ORGANIZATION_TYPEReading Agriculture	4.382414e-03
## ORGANIZATION_TYPEReading Bank	8.570364e-03
## ORGANIZATION_TYPEReading Business Entity Type 1	-1.214637e-03
## ORGANIZATION_TYPEReading Business Entity Type 2	3.669097e-02

## ORGANIZATION_TYPEBusiness Entity Type 3	8.953085e-03
## ORGANIZATION_TYPECleaning	3.671069e-02
## ORGANIZATION_TYPEConstruction	2.132647e-02
## ORGANIZATION_TYPECulture	-5.861336e-02
## ORGANIZATION_TYPEElectricity	-1.564123e-02
## ORGANIZATION_TYPEEmergency	-2.077316e-02
## ORGANIZATION_TYPEGovernment	1.162231e-03
## ORGANIZATION_TYPEHotel	-6.344843e-02
## ORGANIZATION_TYPEHousing	5.454882e-03
## ORGANIZATION_TYPEIndustry: type 1	-3.756790e-02
## ORGANIZATION_TYPEIndustry: type 11	1.558636e-02
## ORGANIZATION_TYPEIndustry: type 12	9.957355e-02
## ORGANIZATION_TYPEIndustry: type 13	-1.205705e-01
## ORGANIZATION_TYPEIndustry: type 2	-7.879773e-02
## ORGANIZATION_TYPEIndustry: type 3	4.168363e-02
## ORGANIZATION_TYPEIndustry: type 4	1.304441e-02
## ORGANIZATION_TYPEIndustry: type 5	1.303146e-02
## ORGANIZATION_TYPEIndustry: type 6	-7.803666e-02
## ORGANIZATION_TYPEIndustry: type 7	2.541053e-02
## ORGANIZATION_TYPEIndustry: type 9	3.122153e-02
## ORGANIZATION_TYPEInsurance	-7.338810e-02
## ORGANIZATION_TYPEKindergarten	-3.629599e-03
## ORGANIZATION_TYPELegal Services	1.982034e-02
## ORGANIZATION_TYPMedicine	3.709787e-02
## ORGANIZATION_TYPMilitary	-3.655942e-02
## ORGANIZATION_TYPMobile	-9.156294e-02
## ORGANIZATION_TYPEOther	1.135235e-03
## ORGANIZATION_TYPEPolice	-3.770840e-02
## ORGANIZATION_TYPEPostal	-6.280312e-03
## ORGANIZATION_TYPERealtor	2.169619e-01
## ORGANIZATION_TYPEReligion	-3.296117e-02
## ORGANIZATION_TYPERestaurant	-1.767501e-02
## ORGANIZATION_TYPSchool	8.833369e-04
## ORGANIZATION_TYPSecurity	-3.087008e-02
## ORGANIZATION_TYPSecurity Ministries	-3.000705e-03
## ORGANIZATION_TYPSelf-employed	1.969723e-02
## ORGANIZATION_TYPServices	-1.091431e-02
## ORGANIZATION_TYPTelecom	-2.065610e-02
## ORGANIZATION_TYPTTrade: type 1	1.810593e-02
## ORGANIZATION_TYPTTrade: type 2	-1.085085e-03
## ORGANIZATION_TYPTTrade: type 3	-1.529725e-02
## ORGANIZATION_TYPTTrade: type 4	-1.012080e-01
## ORGANIZATION_TYPTTrade: type 6	-9.123168e-03
## ORGANIZATION_TYPTTrade: type 7	3.726542e-02
## ORGANIZATION_TYPTTransport: type 1	-7.505025e-02
## ORGANIZATION_TYPTTransport: type 2	7.819173e-02
## ORGANIZATION_TYPTTransport: type 3	1.384465e-01
## ORGANIZATION_TYPTTransport: type 4	-3.459733e-02
## ORGANIZATION_TYPEUniversity	-3.775120e-04
## ORGANIZATION_TYPEXNA	-2.240731e+00
## AMR_REQ_CREDIT_BUREAU_SUM1	4.190379e-03
## AMR_REQ_CREDIT_BUREAU_SUM2	1.092473e-02
## AMR_REQ_CREDIT_BUREAU_SUM3	1.288864e-02
## AMR_REQ_CREDIT_BUREAU_SUM4	3.119740e-03
## AMR_REQ_CREDIT_BUREAU_SUM5	-7.012164e-03
## AMR_REQ_CREDIT_BUREAU_SUM6	2.503563e-02
## AMR_REQ_CREDIT_BUREAU_SUM7	1.588082e-02
## AMR_REQ_CREDIT_BUREAU_SUM8	5.843887e-02
## AMR_REQ_CREDIT_BUREAU_SUM9	1.691654e-02
## AMR_REQ_CREDIT_BUREAU_SUM10	-1.849285e-02
## AMR_REQ_CREDIT_BUREAU_SUM11	3.053713e-03
## AMR_REQ_CREDIT_BUREAU_SUM12	-4.140642e-02
## AMR_REQ_CREDIT_BUREAU_SUM13	-7.344946e-02
## AMR_REQ_CREDIT_BUREAU_SUM14	-1.746657e-02
## AMR_REQ_CREDIT_BUREAU_SUM15	-8.302970e-02
## AMR_REQ_CREDIT_BUREAU_SUM16	-2.768300e-02
## AMR_REQ_CREDIT_BUREAU_SUM17	-9.937823e-02
## AMR_REQ_CREDIT_BUREAU_SUM18	-7.341512e-02

## AMR_REQ_CREDIT_BUREAU_SUM19	-1.680945e-02	
## AMR_REQ_CREDIT_BUREAU_SUM20	-5.056997e-02	
## AMR_REQ_CREDIT_BUREAU_SUM28	-1.117664e-01	
##	std.error	statistic
## (Intercept)	3.238824e-01	0.637039695
## CODE_GENDERM	7.349921e-03	5.922200671
## NAME_CONTRACT_TYPEREvolving loans	9.804538e-03	-5.156065232
## AMT_INCOME_TOTAL	3.806233e-08	-1.396472747
## FLAG_OWN_CARY	6.335356e-03	-4.863258282
## FLAG_OWN_REALTY	6.208820e-03	0.838548695
## AMT_CREDIT	1.107826e-08	-1.191966886
## AMT_ANNUITY	3.253000e-07	-0.544095497
## NAME_TYPE_SUITEChildren	5.078004e-02	-0.542767609
## NAME_TYPE_SUITEFamily	4.468996e-02	-0.218233848
## NAME_TYPE_SUITEGroup of people	1.298177e-01	-0.878341280
## NAME_TYPE_SUITEOther_A	7.422377e-02	-0.857217725
## NAME_TYPE_SUITEOther_B	5.868265e-02	-1.044792141
## NAME_TYPE_SUITESpouse, partner	4.628181e-02	-0.203722120
## NAME_TYPE_SUITEUnaccompanied	4.412372e-02	-0.343197977
## NAME_INCOME_TYPECommercial associate	2.844942e-01	-0.361259798
## NAME_INCOME_TYPEPensioner	3.946596e-01	-0.400389083
## NAME_INCOME_TYPEState servant	2.848237e-01	-0.373563697
## NAME_INCOME_TYPEStudent	3.901416e-01	-0.249116717
## NAME_INCOME_TYPEUnemployed	4.799008e-01	-0.437092064
## NAME_INCOME_TYPEWorking	2.846590e-01	-0.291949486
## NAME_EDUCATION_TYPEHigher education	1.222726e-01	0.303039512
## NAME_EDUCATION_TYPEIncomplete higher	1.231872e-01	0.441687961
## NAME_EDUCATION_TYPERLower secondary	1.248090e-01	0.622387764
## NAME_EDUCATION_TYPERSecondary / secondary special	1.221992e-01	0.454964697
## NAME_FAMILY_STATUSSMarried	9.324405e-03	-0.647327578
## NAME_FAMILY_STATUSSSeparated	1.701766e-02	-0.098174855
## NAME_FAMILY_STATUSSSingle / not married	1.680871e-02	0.475938598
## NAME_FAMILY_STATUSSWidow	2.034232e-02	0.849372761
## NAME_HOUSING_TYPEHouse / apartment	4.101925e-02	0.143638448
## NAME_HOUSING_TYPERMunicipal apartment	4.344237e-02	-0.047212774
## NAME_HOUSING_TYPEROffice apartment	5.133026e-02	-0.879207432
## NAME_HOUSING_TYPERRented apartment	4.573693e-02	0.761341852
## NAME_HOUSING_TYPERWith parents	4.277699e-02	0.590692396
## REGION_POPULATION_RELATIVE	2.097866e-01	-0.467410822
## DAYS_BIRTH	9.715712e-07	3.310204273
## DAYS_EMPLOYED	1.476965e-06	4.225575963
## OWN_CAR_AGE	2.801423e-04	0.370252048
## OCCUPATION_TYPEAccountants	1.778898e-02	-0.822853488
## OCCUPATION_TYPERCleaning staff	2.441296e-02	-0.694231084
## OCCUPATION_TYPERCooking staff	2.101915e-02	0.674746192
## OCCUPATION_TYPERCore staff	1.324128e-02	-0.407997419
## OCCUPATION_TYPERDrivers	1.444968e-02	0.711443845
## OCCUPATION_TYPERHigh skill tech staff	1.559747e-02	-1.137469734
## OCCUPATION_TYPERHR staff	7.360465e-02	-0.103319349
## OCCUPATION_TYPERIT staff	6.037217e-02	-0.464879335
## OCCUPATION_TYPERLaborers	1.040000e-02	-0.463125263
## OCCUPATION_TYPERLow-skill Laborers	3.512250e-02	0.763558124
## OCCUPATION_TYPERManagers	1.342852e-02	-0.422743750
## OCCUPATION_TYPERMedicine staff	2.288624e-02	-1.088646858
## OCCUPATION_TYPERPrivate service staff	3.046778e-02	-0.209399389
## OCCUPATION_TYPERRealty agents	5.679346e-02	1.266163399
## OCCUPATION_TYPERSales staff	1.227742e-02	-0.331171413
## OCCUPATION_TYPERSecretaries	4.763194e-02	0.855052630
## OCCUPATION_TYPERSecurity staff	2.237480e-02	1.825662323
## OCCUPATION_TYPERWaiters/barmen staff	4.716424e-02	2.125129651
## CNT_FAM_MEMBERS2	1.464121e-02	0.077323907
## CNT_FAM_MEMBERS3	1.639724e-02	-0.141516061
## CNT_FAM_MEMBERS4	1.817850e-02	-0.308001501
## CNT_FAM_MEMBERS5	3.137761e-02	1.084129310
## CNT_FAM_MEMBERS6	8.416156e-02	2.132183844
## CNT_FAM_MEMBERS7	1.237411e-01	0.719758220
## ORGANIZATION_TYPEAgriculture	7.110094e-02	0.061636519
## ORGANIZATION_TYPERBank	7.345444e-02	0.116675911

## ORGANIZATION_TYPEBusiness Entity Type 1	6.733533e-02	-0.018038620
## ORGANIZATION_TYPEBusiness Entity Type 2	6.608540e-02	0.555205375
## ORGANIZATION_TYPEBusiness Entity Type 3	6.467652e-02	0.138428667
## ORGANIZATION_TYPECleaning	1.114015e-01	0.329534866
## ORGANIZATION_TYPEConstruction	6.707111e-02	0.317968028
## ORGANIZATION_TYPECulture	9.999669e-02	-0.586153024
## ORGANIZATION_TYPEElectricity	8.055834e-02	-0.194160226
## ORGANIZATION_TYPEEmergency	8.792009e-02	-0.236273141
## ORGANIZATION_TYPEGovernment	6.604658e-02	0.017597149
## ORGANIZATION_TYPEHotel	8.257608e-02	-0.768363320
## ORGANIZATION_TYPEHousing	6.982835e-02	0.078118440
## ORGANIZATION_TYPEIndustry: type 1	7.841714e-02	-0.479077683
## ORGANIZATION_TYPEIndustry: type 11	7.098229e-02	0.219580928
## ORGANIZATION_TYPEIndustry: type 12	1.020855e-01	0.975393521
## ORGANIZATION_TYPEIndustry: type 13	2.036842e-01	-0.591948495
## ORGANIZATION_TYPEIndustry: type 2	9.731529e-02	-0.809715791
## ORGANIZATION_TYPEIndustry: type 3	6.914345e-02	0.602857174
## ORGANIZATION_TYPEIndustry: type 4	7.817763e-02	0.166856096
## ORGANIZATION_TYPEIndustry: type 5	8.878250e-02	0.146779649
## ORGANIZATION_TYPEIndustry: type 6	1.380145e-01	-0.565423688
## ORGANIZATION_TYPEIndustry: type 7	7.682659e-02	0.330751780
## ORGANIZATION_TYPEIndustry: type 9	6.945843e-02	0.449499499
## ORGANIZATION_TYPEInsurance	9.119419e-02	-0.804745374
## ORGANIZATION_TYPEKindergarten	6.709661e-02	-0.054095110
## ORGANIZATION_TYPELegal Services	9.761973e-02	0.203036166
## ORGANIZATION_TYPEMedicine	6.714712e-02	0.552486377
## ORGANIZATION_TYPEMilitary	7.111523e-02	-0.514086997
## ORGANIZATION_TYPEMobile	1.046480e-01	-0.874961544
## ORGANIZATION_TYPEOther	6.550997e-02	0.017329199
## ORGANIZATION_TYPEPolice	7.237334e-02	-0.521026129
## ORGANIZATION_TYPEPostal	7.224699e-02	-0.086928360
## ORGANIZATION_TYPERealtor	1.045732e-01	2.074736910
## ORGANIZATION_TYPEReligion	2.033270e-01	-0.162109178
## ORGANIZATION_TYPERestaurant	7.372632e-02	-0.239738171
## ORGANIZATION_TYPSchool	6.659657e-02	0.013263999
## ORGANIZATION_TYPESecurity	7.165784e-02	-0.430798349
## ORGANIZATION_TYPESecurity Ministries	7.303849e-02	-0.041083890
## ORGANIZATION_TYPESelf-employed	6.488590e-02	0.303567190
## ORGANIZATION_TYPEServices	7.689387e-02	-0.141939911
## ORGANIZATION_TYPETelecom	8.996517e-02	-0.229601124
## ORGANIZATION_TYPETrade: type 1	1.045225e-01	0.173225100
## ORGANIZATION_TYPETrade: type 2	7.420166e-02	-0.014623457
## ORGANIZATION_TYPETrade: type 3	6.905253e-02	-0.221530686
## ORGANIZATION_TYPETrade: type 4	2.036165e-01	-0.497052183
## ORGANIZATION_TYPETrade: type 6	8.610919e-02	-0.105948836
## ORGANIZATION_TYPETrade: type 7	6.674541e-02	0.558321819
## ORGANIZATION_TYPETransport: type 1	1.510872e-01	-0.496734785
## ORGANIZATION_TYPETransport: type 2	7.187849e-02	1.087832087
## ORGANIZATION_TYPETransport: type 3	7.668451e-02	1.805403139
## ORGANIZATION_TYPETransport: type 4	6.837164e-02	-0.506018664
## ORGANIZATION_TYPEUniversity	7.928623e-02	-0.004761382
## ORGANIZATION_TYPEXNA	6.089517e-01	-3.679653516
## AMR_REQ_CREDIT_BUREAU_SUM1	9.127409e-03	0.459098434
## AMR_REQ_CREDIT_BUREAU_SUM2	9.816924e-03	1.112846530
## AMR_REQ_CREDIT_BUREAU_SUM3	1.053364e-02	1.223569960
## AMR_REQ_CREDIT_BUREAU_SUM4	1.161135e-02	0.268680246
## AMR_REQ_CREDIT_BUREAU_SUM5	1.360600e-02	-0.515373121
## AMR_REQ_CREDIT_BUREAU_SUM6	1.725497e-02	1.450922280
## AMR_REQ_CREDIT_BUREAU_SUM7	1.898283e-02	0.836588629
## AMR_REQ_CREDIT_BUREAU_SUM8	2.739782e-02	2.132974918
## AMR_REQ_CREDIT_BUREAU_SUM9	3.608539e-02	0.468792048
## AMR_REQ_CREDIT_BUREAU_SUM10	4.468145e-02	-0.413882025
## AMR_REQ_CREDIT_BUREAU_SUM11	7.004005e-02	0.043599527
## AMR_REQ_CREDIT_BUREAU_SUM12	1.057255e-01	-0.391640785
## AMR_REQ_CREDIT_BUREAU_SUM13	1.157125e-01	-0.634758237
## AMR_REQ_CREDIT_BUREAU_SUM14	1.799169e-01	-0.097081323
## AMR_REQ_CREDIT_BUREAU_SUM15	1.321943e-01	-0.628088210
## AMR_REQ_CREDIT_BUREAU_SUM16	2.742893e-01	-0.100926295

## AMR_REQ_CREDIT_BUREAU_SUM17	2.726974e-01	-0.364426748
## AMR_REQ_CREDIT_BUREAU_SUM18	1.840078e-01	-0.398978291
## AMR_REQ_CREDIT_BUREAU_SUM19	2.746293e-01	-0.061207775
## AMR_REQ_CREDIT_BUREAU_SUM20	2.562267e-01	-0.197364150
## AMR_REQ_CREDIT_BUREAU_SUM28	2.599306e-01	-0.429985680
##	df	p.value
## (Intercept)	9842.13067	5.241138e-01
## CODE_GENDERM	9847.07017	3.282767e-09
## NAME_CONTRACT_TYPRevolving loans	9667.51517	2.570858e-07
## AMT_INCOME_TOTAL	9030.00715	1.626036e-01
## FLAG_OWN_CARY	9859.54073	1.172539e-06
## FLAG_OWN_REALTYT	9757.98866	4.017429e-01
## AMT_CREDIT	9807.46426	2.333029e-01
## AMT_ANNUITY	9841.48892	5.863881e-01
## NAME_TYPE_SUITEChildren	9835.74541	5.873021e-01
## NAME_TYPE_SUITEFamily	9851.82268	8.272514e-01
## NAME_TYPE_SUITEGroup of people	9854.89304	3.797799e-01
## NAME_TYPE_SUITEOther_A	9856.74596	3.913454e-01
## NAME_TYPE_SUITEOther_B	9853.70456	2.961447e-01
## NAME_TYPE_SUITESpouse, partner	9854.28935	8.385749e-01
## NAME_TYPE_SUITEUnaccompanied	9846.19867	7.314568e-01
## NAME_INCOME_TYPECommercial associate	9845.81218	7.179130e-01
## NAME_INCOME_TYPEPensioner	9850.37841	6.888786e-01
## NAME_INCOME_TYPEState servant	9846.11085	7.087370e-01
## NAME_INCOME_TYPEStudent	9858.09436	8.032757e-01
## NAME_INCOME_TYPEUnemployed	9852.56370	6.620542e-01
## NAME_INCOME_TYPEReWorking	9845.49705	7.703314e-01
## NAME_EDUCATION_TYPEHigher education	9857.85728	7.618661e-01
## NAME_EDUCATION_TYPEIncomplete higher	9857.92865	6.587247e-01
## NAME_EDUCATION_TYPEReLower secondary	9856.82413	5.337013e-01
## NAME_EDUCATION_TYPEReSecondary / secondary special	9858.02816	6.491447e-01
## NAME_FAMILY_STATUSSMarried	9851.19907	5.174350e-01
## NAME_FAMILY_STATUSSSeparated	9852.71466	9.217954e-01
## NAME_FAMILY_STATUSSSingle / not married	9855.64716	6.341287e-01
## NAME_FAMILY_STATUSSWidow	9801.82300	3.956945e-01
## NAME_HOUSING_TYPERHouse / apartment	9850.15576	8.857889e-01
## NAME_HOUSING_TYPERMunicipal apartment	9852.04325	9.623446e-01
## NAME_HOUSING_TYPEROffice apartment	9855.89066	3.793102e-01
## NAME_HOUSING_TYPERRented apartment	9857.25008	4.464711e-01
## NAME_HOUSING_TYPERWith parents	9853.18499	5.547401e-01
## REGION_POPULATION_RELATIVE	9810.09843	6.402163e-01
## DAYS_BIRTH	9852.36184	9.356256e-04
## DAYS_EMPLOYED	9851.66253	2.404757e-05
## OWN_CAR_AGE	62.44924	7.112027e-01
## OCCUPATION_TYPERAccountants	9795.14360	4.106112e-01
## OCCUPATION_TYPERCleaning staff	9849.07232	4.875537e-01
## OCCUPATION_TYPERCooking staff	9859.54353	4.998529e-01
## OCCUPATION_TYPERCore staff	9857.26052	6.832844e-01
## OCCUPATION_TYPERDrivers	9849.62182	4.768261e-01
## OCCUPATION_TYPERHigh skill tech staff	9857.56269	2.553696e-01
## OCCUPATION_TYPERHR staff	9845.41388	9.177116e-01
## OCCUPATION_TYPERIT staff	9852.56666	6.420281e-01
## OCCUPATION_TYPERLaborers	9838.74883	6.432848e-01
## OCCUPATION_TYPERLow-skill Laborers	9839.07310	4.451489e-01
## OCCUPATION_TYPERManagers	9854.13464	6.724915e-01
## OCCUPATION_TYPERMedicine staff	9845.39368	2.763363e-01
## OCCUPATION_TYPERPrivate service staff	9857.24427	8.341408e-01
## OCCUPATION_TYPERRealty agents	9858.84639	2.054845e-01
## OCCUPATION_TYPERSales staff	9859.23310	7.405221e-01
## OCCUPATION_TYPERSecretaries	9855.14485	3.925428e-01
## OCCUPATION_TYPERSecurity staff	9850.20513	6.793137e-02
## OCCUPATION_TYPERWaiters/barmen staff	9857.58809	3.360059e-02
## CNT_FAM_MEMBERS2	9857.49802	9.383674e-01
## CNT_FAM_MEMBERS3	9858.30430	8.874652e-01
## CNT_FAM_MEMBERS4	9859.79692	7.580877e-01
## CNT_FAM_MEMBERS5	9859.60681	2.783340e-01
## CNT_FAM_MEMBERS6	9859.79036	3.301638e-02
## CNT_FAM_MEMBERS7	9859.90649	4.716909e-01

## ORGANIZATION_TYPEAgriculture	9851.96946	9.508535e-01
## ORGANIZATION_TYPEBank	9853.09656	9.071193e-01
## ORGANIZATION_TYPEBusiness Entity Type 1	9849.95852	9.856084e-01
## ORGANIZATION_TYPEBusiness Entity Type 2	9852.76155	5.787668e-01
## ORGANIZATION_TYPEBusiness Entity Type 3	9849.04393	8.899045e-01
## ORGANIZATION_TYPECleaning	9852.78481	7.417584e-01
## ORGANIZATION_TYPEConstruction	9846.97936	7.505159e-01
## ORGANIZATION_TYPECulture	9855.28352	5.577861e-01
## ORGANIZATION_TYPEElectricity	9855.34649	8.460544e-01
## ORGANIZATION_TYPEEmergency	9854.58245	8.132256e-01
## ORGANIZATION_TYPEGovernment	9849.30070	9.859606e-01
## ORGANIZATION_TYPEHotel	9856.82784	4.422898e-01
## ORGANIZATION_TYPEHousing	9853.55464	9.377354e-01
## ORGANIZATION_TYPEIndustry: type 1	9849.99755	6.318940e-01
## ORGANIZATION_TYPEIndustry: type 11	9848.57262	8.262021e-01
## ORGANIZATION_TYPEIndustry: type 12	9856.01441	3.293890e-01
## ORGANIZATION_TYPEIndustry: type 13	9852.40309	5.538987e-01
## ORGANIZATION_TYPEIndustry: type 2	9856.14096	4.181231e-01
## ORGANIZATION_TYPEIndustry: type 3	9849.73606	5.466176e-01
## ORGANIZATION_TYPEIndustry: type 4	9852.96031	8.674867e-01
## ORGANIZATION_TYPEIndustry: type 5	9855.15334	8.833089e-01
## ORGANIZATION_TYPEIndustry: type 6	9857.90329	5.717985e-01
## ORGANIZATION_TYPEIndustry: type 7	9848.23792	7.408390e-01
## ORGANIZATION_TYPEIndustry: type 9	9851.75824	6.530812e-01
## ORGANIZATION_TYPEInsurance	9853.96224	4.209860e-01
## ORGANIZATION_TYPEKindergarten	9850.80883	9.568605e-01
## ORGANIZATION_TYPELegal Services	9853.80085	8.391110e-01
## ORGANIZATION_TYEMedicine	9851.10286	5.806277e-01
## ORGANIZATION_TYEMilitary	9853.58736	6.072027e-01
## ORGANIZATION_TYEMobile	9857.00555	3.816161e-01
## ORGANIZATION_TYPEOther	9850.35832	9.861743e-01
## ORGANIZATION_TYPEPolice	9852.53848	6.023603e-01
## ORGANIZATION_TYPEPostal	9845.47055	9.307302e-01
## ORGANIZATION_TYERealtor	9857.75687	3.803680e-02
## ORGANIZATION_TYEReligion	9858.83227	8.712232e-01
## ORGANIZATION_TYERestaurant	9849.14060	8.105382e-01
## ORGANIZATION_TYESchool	9850.11849	9.894174e-01
## ORGANIZATION_TYESecurity	9851.16020	6.666244e-01
## ORGANIZATION_TYESecurity Ministries	9849.88328	9.672298e-01
## ORGANIZATION_TYESelf-employed	9850.71347	7.614641e-01
## ORGANIZATION_TYEServices	9855.21952	8.871304e-01
## ORGANIZATION_TYETelecom	9855.53402	8.184065e-01
## ORGANIZATION_TYETrade: type 1	9847.12069	8.624781e-01
## ORGANIZATION_TYETrade: type 2	9850.05884	9.883329e-01
## ORGANIZATION_TYETrade: type 3	9842.19071	8.246838e-01
## ORGANIZATION_TYETrade: type 4	9858.41253	6.191633e-01
## ORGANIZATION_TYETrade: type 6	9850.30181	9.156251e-01
## ORGANIZATION_TYETrade: type 7	9848.48774	5.766373e-01
## ORGANIZATION_TYETransport: type 1	9857.65978	6.193872e-01
## ORGANIZATION_TYETransport: type 2	9852.87199	2.766958e-01
## ORGANIZATION_TYETransport: type 3	9854.21526	7.104211e-02
## ORGANIZATION_TYETransport: type 4	9849.40373	6.128549e-01
## ORGANIZATION_TYPEUniversity	9851.63858	9.962011e-01
## ORGANIZATION_TYPEXNA	9855.66774	2.347956e-04
## AMR_REQ_CREDIT_BUREAU_SUM1	1417.16708	6.461736e-01
## AMR_REQ_CREDIT_BUREAU_SUM2	250.28318	2.658015e-01
## AMR_REQ_CREDIT_BUREAU_SUM3	260.84440	2.211437e-01
## AMR_REQ_CREDIT_BUREAU_SUM4	311.38428	7.881814e-01
## AMR_REQ_CREDIT_BUREAU_SUM5	499.95796	6.063039e-01
## AMR_REQ_CREDIT_BUREAU_SUM6	172.24968	1.468333e-01
## AMR_REQ_CREDIT_BUREAU_SUM7	521.47606	4.028441e-01
## AMR_REQ_CREDIT_BUREAU_SUM8	302.47679	3.295141e-02
## AMR_REQ_CREDIT_BUREAU_SUM9	276.96744	6.392286e-01
## AMR_REQ_CREDIT_BUREAU_SUM10	6861.17223	6.789695e-01
## AMR_REQ_CREDIT_BUREAU_SUM11	1787.91692	9.652245e-01
## AMR_REQ_CREDIT_BUREAU_SUM12	9659.66009	6.953321e-01
## AMR_REQ_CREDIT_BUREAU_SUM13	9710.15537	5.256008e-01
## AMR_REQ_CREDIT_BUREAU_SUM14	9819.14406	9.226638e-01



## AMR_REQ_CREDIT_BUREAU_SUM15	9501.72835 5.299607e-01
## AMR_REQ_CREDIT_BUREAU_SUM16	9859.90649 9.196110e-01
## AMR_REQ_CREDIT_BUREAU_SUM17	9859.35814 7.155472e-01
## AMR_REQ_CREDIT_BUREAU_SUM18	9491.65927 6.899178e-01
## AMR_REQ_CREDIT_BUREAU_SUM19	9859.90649 9.511950e-01
## AMR_REQ_CREDIT_BUREAU_SUM20	9858.65550 8.435467e-01
## AMR_REQ_CREDIT_BUREAU_SUM28	9833.27900 6.672155e-01
##	2.5 %
## (Intercept)	-4.285499e-01
## CODE_GENDERM	2.912035e-02
## NAME_CONTRACT_TYPERevolving loans	-6.977179e-02
## AMT_INCOME_TOTAL	-1.277638e-07
## FLAG_OWN_CARY	-4.322907e-02
## FLAG_OWN_REALTY	-6.964175e-03
## AMT_CREDIT	-3.492059e-08
## AMT_ANNUITY	-8.146490e-07
## NAME_TYPE_SUITEChildren	-1.271011e-01
## NAME_TYPE_SUITEFamily	-9.735433e-02
## NAME_TYPE_SUITEGroup of people	-3.684936e-01
## NAME_TYPE_SUITEOther_A	-2.091197e-01
## NAME_TYPE_SUITEOther_B	-1.763412e-01
## NAME_TYPE_SUITESpouse, partner	-1.001504e-01
## NAME_TYPE_SUITEUnaccompanied	-1.016347e-01
## NAME_INCOME_TYPECommercial associate	-6.604432e-01
## NAME_INCOME_TYPEPensioner	-9.316310e-01
## NAME_INCOME_TYPEState servant	-6.647126e-01
## NAME_INCOME_TYPEStudent	-8.619482e-01
## NAME_INCOME_TYPEUnemployed	-1.150465e+00
## NAME_INCOME_TYPEWorking	-6.410961e-01
## NAME_EDUCATION_TYPEHigher education	-2.026260e-01
## NAME_EDUCATION_TYPEIncomplete higher	-1.870618e-01
## NAME_EDUCATION_TypesLower secondary	-1.669716e-01
## NAME_EDUCATION_TypesSecondary / secondary special	-1.839391e-01
## NAME_FAMILY_STATUSSMarried	-2.431369e-02
## NAME_FAMILY_STATUSSSeparated	-3.502881e-02
## NAME_FAMILY_STATUSSSingle / not married	-2.494859e-02
## NAME_FAMILY_STATUSSWidow	-2.259693e-02
## NAME_HOUSING_TYPEHouse / apartment	-7.451419e-02
## NAME_HOUSING_TypesMunicipal apartment	-8.720697e-02
## NAME_HOUSING_TypesOffice apartment	-1.457478e-01
## NAME_HOUSING_TypesRented apartment	-5.483231e-02
## NAME_HOUSING_TypesWith parents	-5.858361e-02
## REGION_POPULATION_RELATIVE	-5.092816e-01
## DAYS_BIRTH	1.311621e-06
## DAYS_EMPLOYED	3.345873e-06
## OWN_CAR_AGE	-4.561931e-04
## OCCUPATION_TypesAccountants	-4.950778e-02
## OCCUPATION_TypesCleaning staff	-6.480265e-02
## OCCUPATION_TypesCooking staff	-2.701925e-02
## OCCUPATION_TypesCore staff	-3.135803e-02
## OCCUPATION_TypesDrivers	-1.804419e-02
## OCCUPATION_TypesHigh skill tech staff	-4.831588e-02
## OCCUPATION_TypesHR staff	-1.518850e-01
## OCCUPATION_TypesIT staff	-1.464076e-01
## OCCUPATION_TypesLaborers	-2.520263e-02
## OCCUPATION_TypesLow-skill Laborers	-4.202923e-02
## OCCUPATION_TypesManagers	-3.199947e-02
## OCCUPATION_TypesMedicine staff	-6.977675e-02
## OCCUPATION_TypesPrivate service staff	-6.610302e-02
## OCCUPATION_TypesRealty agents	-3.941700e-02
## OCCUPATION_TypesSales staff	-2.813218e-02
## OCCUPATION_TypesSecretaries	-5.264054e-02
## OCCUPATION_TypesSecurity staff	-3.010362e-03
## OCCUPATION_TypesWaiters/barmen staff	7.778561e-03
## CNT_FAM_MEMBERS2	-2.756765e-02
## CNT_FAM_MEMBERS3	-3.446242e-02
## CNT_FAM_MEMBERS4	-4.123259e-02
## CNT_FAM_MEMBERS5	-2.748915e-02

## CNT_FAM_MEMBERS6	1.447404e-02
## CNT_FAM_MEMBERS7	-1.534942e-01
## ORGANIZATION_TYPEAgriculture	-1.349900e-01
## ORGANIZATION_TYPEBank	-1.354154e-01
## ORGANIZATION_TYPEBusiness Entity Type 1	-1.332057e-01
## ORGANIZATION_TYPEBusiness Entity Type 2	-9.284995e-02
## ORGANIZATION_TYPEBusiness Entity Type 3	-1.178262e-01
## ORGANIZATION_TYPECleaning	-1.816591e-01
## ORGANIZATION_TYPEConstruction	-1.101466e-01
## ORGANIZATION_TYPECulture	-2.546274e-01
## ORGANIZATION_TYPEElectricity	-1.735521e-01
## ORGANIZATION_TYPEEmergency	-1.931145e-01
## ORGANIZATION_TYPEGovernment	-1.283026e-01
## ORGANIZATION_TYPEHotel	-2.253144e-01
## ORGANIZATION_TYPEHousing	-1.314230e-01
## ORGANIZATION_TYPEIndustry: type 1	-1.912816e-01
## ORGANIZATION_TYPEIndustry: type 11	-1.235535e-01
## ORGANIZATION_TYPEIndustry: type 12	-1.005350e-01
## ORGANIZATION_TYPEIndustry: type 13	-5.198332e-01
## ORGANIZATION_TYPEIndustry: type 2	-2.695556e-01
## ORGANIZATION_TYPEIndustry: type 3	-9.385171e-02
## ORGANIZATION_TYPEIndustry: type 4	-1.401997e-01
## ORGANIZATION_TYPEIndustry: type 5	-1.610004e-01
## ORGANIZATION_TYPEIndustry: type 6	-3.485733e-01
## ORGANIZATION_TYPEIndustry: type 7	-1.251853e-01
## ORGANIZATION_TYPEIndustry: type 9	-1.049312e-01
## ORGANIZATION_TYPEInsurance	-2.521474e-01
## ORGANIZATION_TYPEKindergarten	-1.351527e-01
## ORGANIZATION_TYPELegal Services	-1.715343e-01
## ORGANIZATION_TYEMedicine	-9.452424e-02
## ORGANIZATION_TYEMilitary	-1.759598e-01
## ORGANIZATION_TYEMobile	-2.966944e-01
## ORGANIZATION_TYPEOther	-1.272777e-01
## ORGANIZATION_TYPEPolice	-1.795750e-01
## ORGANIZATION_TYPEPostal	-1.478992e-01
## ORGANIZATION_TYERealtor	1.197700e-02
## ORGANIZATION_TYEReligion	-4.315237e-01
## ORGANIZATION_TYERestaurant	-1.621937e-01
## ORGANIZATION_TYESchool	-1.296596e-01
## ORGANIZATION_TYESecurity	-1.713341e-01
## ORGANIZATION_TYESecurity Ministries	-1.461711e-01
## ORGANIZATION_TYESelf-employed	-1.074924e-01
## ORGANIZATION_TYEServices	-1.616420e-01
## ORGANIZATION_TYETelecom	-1.970062e-01
## ORGANIZATION_TYETrade: type 1	-1.867797e-01
## ORGANIZATION_TYETrade: type 2	-1.465355e-01
## ORGANIZATION_TYETrade: type 3	-1.506544e-01
## ORGANIZATION_TYETrade: type 4	-5.003380e-01
## ORGANIZATION_TYETrade: type 6	-1.779148e-01
## ORGANIZATION_TYETrade: type 7	-9.356927e-02
## ORGANIZATION_TYETransport: type 1	-3.712120e-01
## ORGANIZATION_TYETransport: type 2	-6.270483e-02
## ORGANIZATION_TYETransport: type 3	-1.187089e-02
## ORGANIZATION_TYETransport: type 4	-1.686197e-01
## ORGANIZATION_TYPEUniversity	-1.557948e-01
## ORGANIZATION_TYPEXNA	-3.434401e+00
## AMR_REQ_CREDIT_BUREAU_SUM1	-1.371431e-02
## AMR_REQ_CREDIT_BUREAU_SUM2	-8.409579e-03
## AMR_REQ_CREDIT_BUREAU_SUM3	-7.853147e-03
## AMR_REQ_CREDIT_BUREAU_SUM4	-1.972689e-02
## AMR_REQ_CREDIT_BUREAU_SUM5	-3.374414e-02
## AMR_REQ_CREDIT_BUREAU_SUM6	-9.022792e-03
## AMR_REQ_CREDIT_BUREAU_SUM7	-2.141140e-02
## AMR_REQ_CREDIT_BUREAU_SUM8	4.524399e-03
## AMR_REQ_CREDIT_BUREAU_SUM9	-5.411993e-02
## AMR_REQ_CREDIT_BUREAU_SUM10	-1.060823e-01
## AMR_REQ_CREDIT_BUREAU_SUM11	-1.343153e-01
## AMR_REQ_CREDIT_BUREAU_SUM12	-2.486506e-01

## AMR_REQ_CREDIT_BUREAU_SUM13	-3.002701e-01
## AMR_REQ_CREDIT_BUREAU_SUM14	-3.701406e-01
## AMR_REQ_CREDIT_BUREAU_SUM15	-3.421588e-01
## AMR_REQ_CREDIT_BUREAU_SUM16	-5.653461e-01
## AMR_REQ_CREDIT_BUREAU_SUM17	-6.339209e-01
## AMR_REQ_CREDIT_BUREAU_SUM18	-4.341098e-01
## AMR_REQ_CREDIT_BUREAU_SUM19	-5.551390e-01
## AMR_REQ_CREDIT_BUREAU_SUM20	-5.528268e-01
## AMR_REQ_CREDIT_BUREAU_SUM28	-6.212837e-01
##	97.5 %
## (Intercept)	8.412018e-01
## CODE_GENDERM	5.793506e-02
## NAME_CONTRACT_TYPERevolving loans	-3.133389e-02
## AMT_INCOME_TOTAL	2.145779e-08
## FLAG_OWN_CARY	-1.839188e-02
## FLAG_OWN_REALTY	1.737697e-02
## AMT_CREDIT	8.510751e-09
## AMT_ANNUITY	4.606605e-07
## NAME_TYPE_SUITEChildren	7.197754e-02
## NAME_TYPE_SUITEFamily	7.784861e-02
## NAME_TYPE_SUITEGroup of people	1.404451e-01
## NAME_TYPE_SUITEOther_A	8.186785e-02
## NAME_TYPE_SUITEOther_B	5.371884e-02
## NAME_TYPE_SUITESpouse, partner	8.129319e-02
## NAME_TYPE_SUITEUnaccompanied	7.134837e-02
## NAME_INCOME_TYPECommercial associate	4.548906e-01
## NAME_INCOME_TYPEPensioner	6.155963e-01
## NAME_INCOME_TYPEState servant	4.519130e-01
## NAME_INCOME_TYPEStudent	6.675666e-01
## NAME_INCOME_TYPEUnemployed	7.309430e-01
## NAME_INCOME_TYPEWorking	4.748840e-01
## NAME_EDUCATION_TYPEHigher education	2.767328e-01
## NAME_EDUCATION_TYPEIncomplete higher	2.958824e-01
## NAME_EDUCATION_TYPELower secondary	3.223308e-01
## NAME_EDUCATION_TYPESecondary / secondary special	2.951317e-01
## NAME_FAMILY_STATUSMarried	1.224180e-02
## NAME_FAMILY_STATUSSeparated	3.168740e-02
## NAME_FAMILY_STATUSSingle / not married	4.094842e-02
## NAME_FAMILY_STATUSWidow	5.715336e-02
## NAME_HOUSING_TYPEHouse / apartment	8.629808e-02
## NAME_HOUSING_TYEMunicipal apartment	8.310490e-02
## NAME_HOUSING_TYPEOffice apartment	5.548787e-02
## NAME_HOUSING_TYERented apartment	1.244752e-01
## NAME_HOUSING_TYEWith parents	1.091197e-01
## REGION_POPULATION_RELATIVE	3.131685e-01
## DAYS_BIRTH	5.120578e-06
## DAYS_EMPLOYED	9.136180e-06
## OWN_CAR_AGE	6.636396e-04
## OCCUPATION_TYPEAccountants	2.023234e-02
## OCCUPATION_TYPECleaning staff	3.090617e-02
## OCCUPATION_TYPECooking staff	5.538443e-02
## OCCUPATION_TYPECore staff	2.055321e-02
## OCCUPATION_TYEDrivers	3.860446e-02
## OCCUPATION_TYPEHigh skill tech staff	1.283258e-02
## OCCUPATION_TYPEHR staff	1.366754e-01
## OCCUPATION_TYPEIT staff	9.027604e-02
## OCCUPATION_TYELaborers	1.556963e-02
## OCCUPATION_TYELow-skill Laborers	9.566536e-02
## OCCUPATION_TYEManagers	2.064583e-02
## OCCUPATION_TYEMedicine staff	1.994669e-02
## OCCUPATION_TYEPprivate service staff	5.334315e-02
## OCCUPATION_TYERealty agents	1.832366e-01
## OCCUPATION_TYESales staff	2.000032e-02
## OCCUPATION_TYESecretaries	1.340962e-01
## OCCUPATION_TYESecurity staff	8.470802e-02
## OCCUPATION_TYEWaiters/barmen staff	1.926817e-01
## CNT_FAM_MEMBERS2	2.983188e-02
## CNT_FAM_MEMBERS3	2.982147e-02

## CNT_FAM_MEMBERS4	3.003458e-02
## CNT_FAM_MEMBERS5	9.552393e-02
## CNT_FAM_MEMBERS6	3.444218e-01
## CNT_FAM_MEMBERS7	3.316215e-01
## ORGANIZATION_TYPEAgriculture	1.437548e-01
## ORGANIZATION_TYPEBank	1.525561e-01
## ORGANIZATION_TYPEBusiness Entity Type 1	1.307764e-01
## ORGANIZATION_TYPEBusiness Entity Type 2	1.662319e-01
## ORGANIZATION_TYPEBusiness Entity Type 3	1.357323e-01
## ORGANIZATION_TYPECleaning	2.550805e-01
## ORGANIZATION_TYPEConstruction	1.527996e-01
## ORGANIZATION_TYPECulture	1.374006e-01
## ORGANIZATION_TYPEElectricity	1.422696e-01
## ORGANIZATION_TYPEEmergency	1.515682e-01
## ORGANIZATION_TYPEGovernment	1.306271e-01
## ORGANIZATION_TYPEHotel	9.841759e-02
## ORGANIZATION_TYPEHousing	1.423328e-01
## ORGANIZATION_TYPEIndustry: type 1	1.161458e-01
## ORGANIZATION_TYPEIndustry: type 11	1.547262e-01
## ORGANIZATION_TYPEIndustry: type 12	2.996820e-01
## ORGANIZATION_TYPEIndustry: type 13	2.786922e-01
## ORGANIZATION_TYPEIndustry: type 2	1.119602e-01
## ORGANIZATION_TYPEIndustry: type 3	1.772190e-01
## ORGANIZATION_TYPEIndustry: type 4	1.662886e-01
## ORGANIZATION_TYPEIndustry: type 5	1.870633e-01
## ORGANIZATION_TYPEIndustry: type 6	1.925000e-01
## ORGANIZATION_TYPEIndustry: type 7	1.760064e-01
## ORGANIZATION_TYPEIndustry: type 9	1.673743e-01
## ORGANIZATION_TYPEInsurance	1.053712e-01
## ORGANIZATION_TYPEKindergarten	1.278935e-01
## ORGANIZATION_TYPELegal Services	2.111750e-01
## ORGANIZATION_TYEMedicine	1.687200e-01
## ORGANIZATION_TYEMilitary	1.028410e-01
## ORGANIZATION_TYEMobile	1.135685e-01
## ORGANIZATION_TYPEOther	1.295482e-01
## ORGANIZATION_TYPEPolice	1.041582e-01
## ORGANIZATION_TYPEPostal	1.353386e-01
## ORGANIZATION_TYERealtor	4.219468e-01
## ORGANIZATION_TYEReligion	3.656013e-01
## ORGANIZATION_TYERestaurant	1.268437e-01
## ORGANIZATION_TYESchool	1.314263e-01
## ORGANIZATION_TYESecurity	1.095940e-01
## ORGANIZATION_TYESecurity Ministries	1.401697e-01
## ORGANIZATION_TYESelf-employed	1.468869e-01
## ORGANIZATION_TYEServices	1.398134e-01
## ORGANIZATION_TYETelecom	1.556940e-01
## ORGANIZATION_TYETrade: type 1	2.229915e-01
## ORGANIZATION_TYETrade: type 2	1.443654e-01
## ORGANIZATION_TYETrade: type 3	1.200599e-01
## ORGANIZATION_TYETrade: type 4	2.979220e-01
## ORGANIZATION_TYETrade: type 6	1.596685e-01
## ORGANIZATION_TYETrade: type 7	1.681001e-01
## ORGANIZATION_TYETransport: type 1	2.211115e-01
## ORGANIZATION_TYETransport: type 2	2.190883e-01
## ORGANIZATION_TYETransport: type 3	2.887638e-01
## ORGANIZATION_TYETransport: type 4	9.942510e-02
## ORGANIZATION_TYPEUniversity	1.550397e-01
## ORGANIZATION_TYPEXNA	-1.047061e+00
## AMR_REQ_CREDIT_BUREAU_SUM1	2.209506e-02
## AMR_REQ_CREDIT_BUREAU_SUM2	3.025904e-02
## AMR_REQ_CREDIT_BUREAU_SUM3	3.363044e-02
## AMR_REQ_CREDIT_BUREAU_SUM4	2.596637e-02
## AMR_REQ_CREDIT_BUREAU_SUM5	1.971981e-02
## AMR_REQ_CREDIT_BUREAU_SUM6	5.909404e-02
## AMR_REQ_CREDIT_BUREAU_SUM7	5.317304e-02
## AMR_REQ_CREDIT_BUREAU_SUM8	1.123533e-01
## AMR_REQ_CREDIT_BUREAU_SUM9	8.795301e-02
## AMR_REQ_CREDIT_BUREAU_SUM10	6.909664e-02

## AMR_REQ_CREDIT_BUREAU_SUM11	1.404227e-01
## AMR_REQ_CREDIT_BUREAU_SUM12	1.658377e-01
## AMR_REQ_CREDIT_BUREAU_SUM13	1.533711e-01
## AMR_REQ_CREDIT_BUREAU_SUM14	3.352075e-01
## AMR_REQ_CREDIT_BUREAU_SUM15	1.760994e-01
## AMR_REQ_CREDIT_BUREAU_SUM16	5.099801e-01
## AMR_REQ_CREDIT_BUREAU_SUM17	4.351645e-01
## AMR_REQ_CREDIT_BUREAU_SUM18	2.872795e-01
## AMR_REQ_CREDIT_BUREAU_SUM19	5.215201e-01
## AMR_REQ_CREDIT_BUREAU_SUM20	4.516868e-01
## AMR_REQ_CREDIT_BUREAU_SUM28	3.977509e-01

## Model Limitation

The limitation of this logistic regressino model includes the following: - I cannot do extrapolation for prediction that is our of the range that I use to form the logistic regression model. - The predictability of this model is relatively week as most p-values are quite large. - The predictability is subject to the choice of threshold.