DSO 562 Fraud Analytics

Project 2

March 24, 2021

Identity Fraud Detection

Team Member

GU, Xinxue (Sue)	**SUN, Yang**	**WANG, Bufan**
XIE, Binghong	**YANG, Siqin**	**ZHAO, Zihao**

Table of Content

1. Executive Summary

This project is aimed at developing machine learning algorithms to identify credit card application fraud. We started our project by examining data quality and conducting data cleaning. Specifically, we inspected the statistics and distribution of each field, and we detected frivolous data values and inconsistent data formats, both of which are handled in the data cleaning process.

After data cleaning, we built 680 candidate variables and selected 30 variables that have the highest predictive power. We then built a broad scope of machine learning models to identify application frauds, including logistic regressions, boosted trees, random forests and neural networks.

Rigorously tuning hyperparameters, we eventually achieved 54% fraud detection rate on training data, 54% on testing data and 51.9% on out of time data with an Extreme Gradient Boosting model.

2. Description of Data

The data has 1,000,000 product application records, with 9 fields and 1 fraud label. Synthetically generated from real data distribution, the dataset is built by an identity fraud prevention company to reproduce the important univariate and multivariate field distributions of real data.

The dataset consists of 8 categorical variables and 2 datetime variables.

Table 2.1 Summary Table for Categorical Fields

Column Name	# of Records	% populated	Unique Values	Most Common Field Value
ssn	1,000,000	100	835819	999999999
firstname	1,000,000	100	78136	EAMSTRMT
lastname	1,000,000	100	177001	ERJSAXA
address	1,000,000	100	828774	123 MAIN ST
zip5	1,000,000	100	26370	68138
dob	1,000,000	100	42673	19070626
homephone	1,000,000	100	28244	9999999999
fraud_label	1,000,000	100	2	0

Table 2.2 Summary Table for Datetime Fields

Column Name	# of Records	% populated	Minimum	Maximum	Most Common Field Value
date	1,000,000	100	2016-01-01	2016-12-31	2016-08-16
dob	1,000,000	100	1900-01-01	2016-10-31	1907-06-26

We examined each field of the dataset and found some interesting facts below.

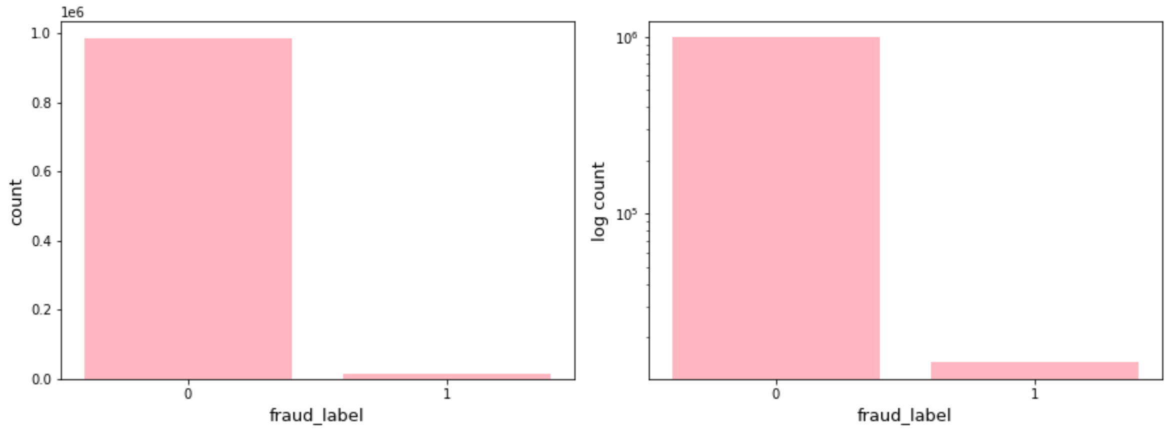


Figure 2.1 Histogram of Fraud Label

In the fraud label field, 14,393 records are labeled as fraudulent, with a proportion of 1.439%.

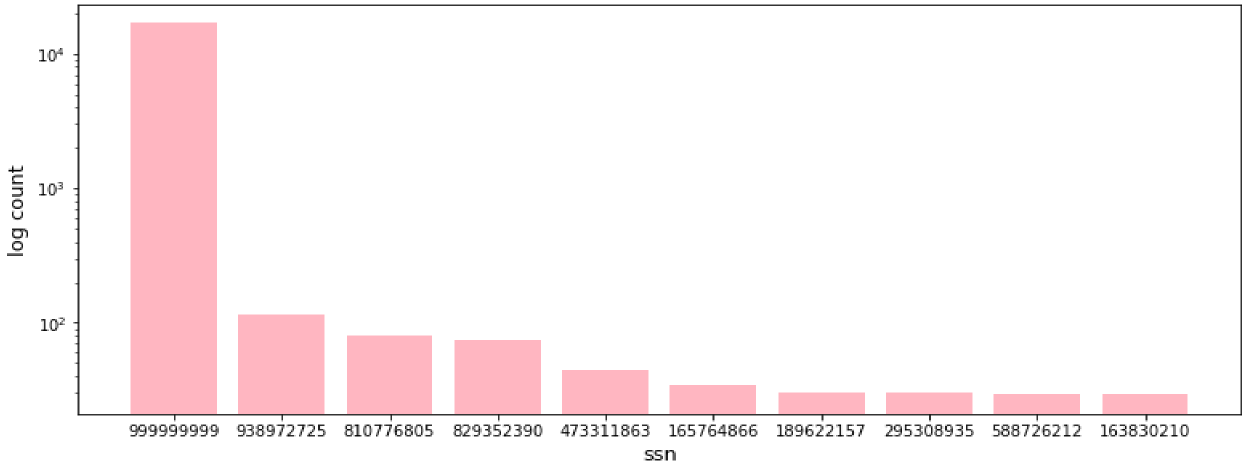


Figure 2.2 Histogram of Social Security Number

In the social security number field, 1.693% of records have the value of 999999999, which is most likely a frivolous value.

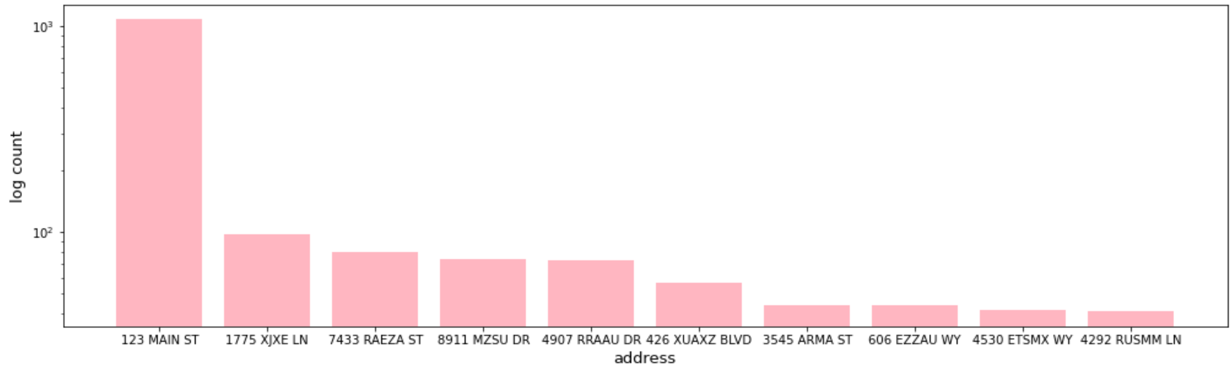


Figure 2.3 Histogram of Address

In the address field, 0.108% of records have the value of 123 MAIN ST, which is most likely a frivolous value.

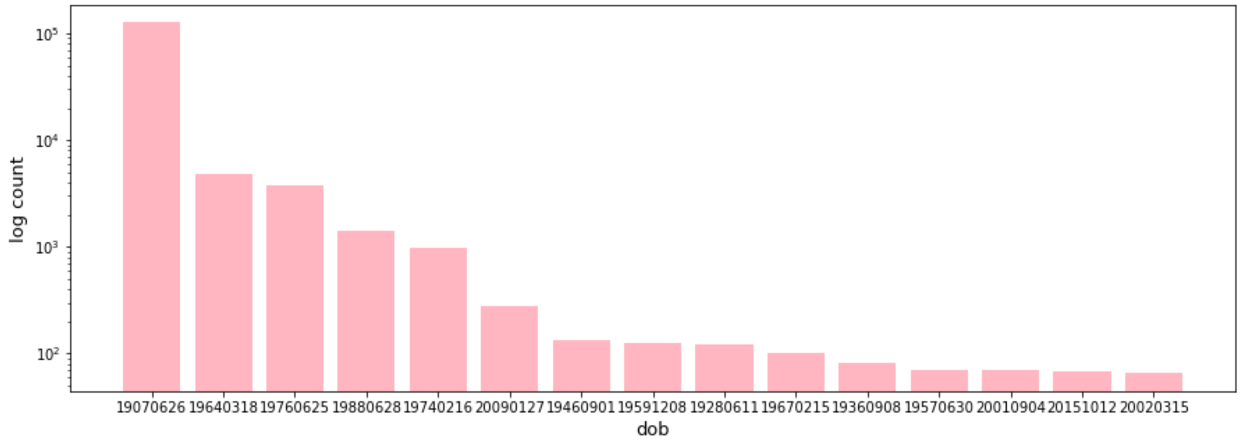


Figure 2.4 Histogram of Date of Birth

In the date of birth field, 12.657% of records have the value of 06/26/1907, which is most likely a frivolous value.

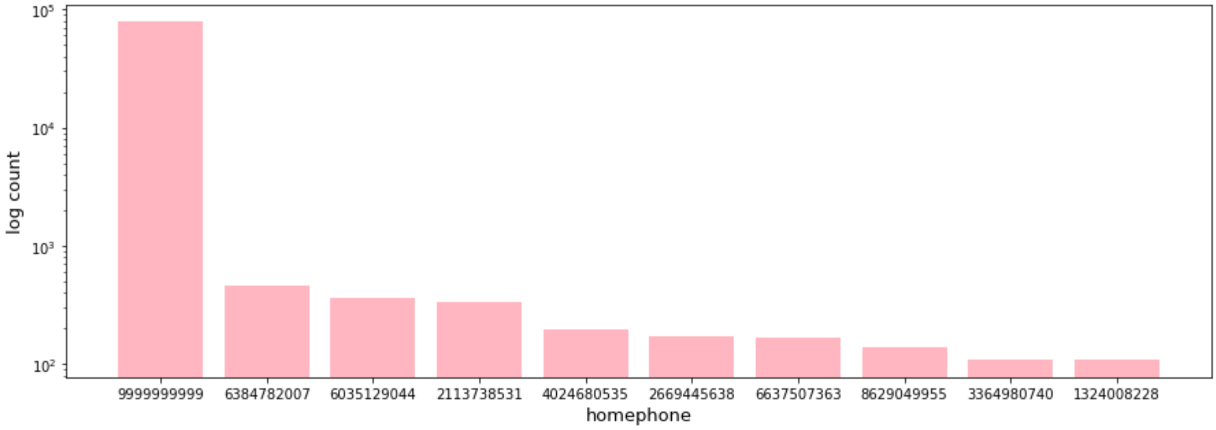


Figure 2.5 Histogram of Home Phone Number

In the home phone number field, 7.851% of records have the value of 9999999999, which is most likely a frivolous value.

More detailed data description can be found in Appendix 9.1.

3. Data Cleaning

After checking data quality, we found that there are no missing values for any fields in this data. However, the data has two major problems.

First, there exists format inconsistency in some fields. For example, some zip5 codes are 4-digit numbers such as 2765. We also found similar inconsistencies in social security number field, date of birth field and home phone number field. To fix format inconsistencies, we picked the format used by most records as the standard and applied it to all records. In the example above, we changed the value 2765 into 02765 to meet the 5-digit standard.

Second, we found certain frivolous values in some fields. For example, in social security number field, the value of 999999999 accounted for 1.693% of total records, which is also an unusual value for social security number. Following this pattern, we reckoned the values below as frivolous values:

- 999999999 in social security number field
- 123 MAIN ST in address field
- 19070626 in date of birth field
- 9999999999 in home phone number field

To fix frivolous values, we replaced them with negative record numbers that would not link to any other records.

Based on time, we further split the data into two parts: modeling data and out of time (OOT) data.

- Modeling data: all records before Nov 1, 2016
- Out of time data: all records on and after Nov 1, 2016

We built variables for all data records. However, when performing feature selection, we used only modeling data. On the other hand, we used OOT data to represent incoming data from different time periods, which provides insights into the model performance in actual practice.

4. Candidate Variables

4.1 Feature Engineering

We used 7 fields to perform feature engineering, including **ssn, firstname, lastname, address, zip5, dob** and **homephone.** We wanted to find the frequency of each record based on those categorical variables or the combination of each categorical variable. For instance, if a particular combination of a social security number and a home phone number appears 30 times a day, or that particular social security number has been paired with more than 10 different names and date of birth, then it is reasonable to consider that records associated with this social security number are more suspicious. Having that logic in mind, we started our variable creation in this direction.

4.2 Variable Creation Process

The entire process of variable creation can be broken down into four steps.

(1) Two Identifiers

We created 2 new fields by concatenating firstname and lastname, address and zip5 to use as identifiers.

1. name = firstname + lastname
2. fulladdress = address + zip

As a result, we got 5 entities:

1. zip5
2. dob
3. homephone
4. name
5. Fulladdress

(2) Twenty-One Combination Groups

We further linked existing fields to 1 or 2 of entities to form 21 combination groups, listed in the table below.

Table 4.1 Combination Groups

No.	Combination Group	Formula
1	name_dob	name + dob
2	name_homephone	name + homephone
3	name_fulladdress	name + address + zip
4	fulladdress_dob	address + zip5 + dob
5	fulladdress_homephone	address + zip5 + homephone
6	dob_homephone	dob + homephone
7	homephone_name_dob	homephone + name + dob
8	ssn_firstname	ssn + firstname
9	ssn_lastname	ssn + lastname
10	ssn_address	ssn + address
11	ssn_zip5	ssn + zip5
12	ssn_dob	ssn + dob
13	ssn_homephone	ssn + homephone
14	ssn_fulladdress	ssn + address + zip5

15	ssn_name_dob	ssn + name + dob
16	ssn_name_homephone	ssn + name + homephone
17	ssn_name_fulladdress	ssn + name + address + zip5
18	ssn_dob_homephone	ssn + dob + homephone
19	ssn_fulladdress_homephone	ssn + address + zip5 + homephone
20	ssn_fulladdress_dob	ssn + address + zip5 + dob
21	ssn_homephone_name_dob	ssn + homephone + name + dob

Eventually, we got 26 attributes (5 entities and 21 combination groups).

(3) Four Series of Variables

Day-since variables are a series of variables defined by how many days since the last time the entity or combination group was seen. 26 variables were created in total using this formula. The alias for variables in this series have the following structure: [attribute name + _day_since].

Velocity variables are a series of variables defined by the number of records with same attributes over the last {0, 1, 3, 7, 14, 30} days. 156 variables were created in total **(**26 attributes x 6 different day periods) using this formula. The alias for variables in this series have the following structure: [attribute name + _count_ + day].

Relative velocity variables are a series of variables defined by the number of applications with attributes seen in the past {0, 1} day divided by the number of applications with the same attributes seen in past {3, 7, 14 ,30} days. 208 variables were created in total (26 attributes x 2 days x 4 days) with this formula. The alias for variables in this series have the following structure: [attribute name + _count_ + day (0 or 1) + _by_ + day (3, 7, 14, or 30)].

Uniqueness variables are a series of variables defined by the number of unique attributes for a particular (different) attribute over the past {3,7,14,30} days (Here 9 useful attributes were chosen out of the 26 attributes). 288 variables were created in total (72 permutations x 4 different past days) using this formula. The alias for variables in this series have the following structure: [unique_num_ + attribute name + _for_each_ + attribute name + _day].

(4) Risk Variable

Day of week (DOW) risk value is defined by the likelihood of fraud for a particular day of the week. This is computed by grouping days of the week and assign corresponding average value of y to each of days. 1 variable was created in total using this formula which is named dow_risk. The risk values were calculated based on training and testing data, not from out of time data.

4.3 Summary of Candidate Variables

We created 680 features in this stage, which can be divided in 6 categories: 26 day-since variables, 156 velocity variables, 208 relative velocity variables, 288 uniqueness variables, 1 DOW risk variable, and 1 random variable that is designed to examine the effectiveness of all variables created. With all these carefully created variables, we were ready to go to the next step — feature selection.

For more details and a full list of 680 variables, please refer to Appendix 9.2.

5. Feature Selection

5.1 Methodology for Feature Selection

Feature selection is the process of selecting a subset of relevant features that contribute the most to model performance. To maximize model accuracy, we wanted to exclude any redundant or irrelevant features that can incur loss of information. In this section, we demonstrated how we used two robust methods: filter and wrapper, to selected 30 top performance variables out of 680 features.

The data we used for feature selection are applications hapspened between 01/15/2016 and 11/01/2016. We removed first two weeks' records because they are not well built (fraud label usually comes a few months later after the applications are submitted).

5.2 Filter Method

The filter method evaluates the predictive power of each feature using various kinds of univariate measures, such as Pearson Correlation, Fisher Score, Kolmogorov-Smirnov, and mutual information. In this project, we chose two measures: Kolmogorov–Smirnov test (KS) and Fraud Detection Rate (FDR). KS is a statistical measure of how well two distributions are separated. In our case, it calculates the maximum difference of cumulative distributions of "Goods" (fraud_label = 0) and "Bads" (fraud_label = 1). FDR is a common and robust measure for fraud model. It calculates the percentage of all the fraud being caught at a score cutoff. For instance, 50% FDR at 3% cutoff means that in the top 3% applications ranked by predicted fraud score, 50% of all the frauds can be caught (% Frauds caught = # frauds caught / total # frauds 20% in that data set (training/testing/OOT)).

At the end of filter stage, we got a list of candidate variables sorted by the average ranks of KS and FDR. The higher the ranking, the more important variables are to predict our target field.

First, we split the data into "Goods" and "Bads" according to the fraud label. Next, we used the *scipy.stats.ks_2samp* function to calculate the KS score for each candidate variable, with Goods and Bads record of each column as the input. Then, we calculated the FDR at 3% for all our candidate variables. To do this, we made a DataFrame for every variable together with the fraud label, sorted the DataFrame by the variable value, selected the top 3% record, and count how many "Bads" are in it. We then divided this number by the total number of "Bads" in the data to get our FDR.

After sorting, ranking, and averaging two scores, we selected 80 variables out of 706 candidates for the next stage – Wrapper.

5.3 Wrapper Method

The wrapper method gives us information on the combined importance of variables. In other words, variables that are not adding value to the combination will be thrown away. There are two types of wrapper: Forward Selection and Backward Selection. Forward Selection starts from one variable and adds one at a time. On the contrary, Backward Selection starts from modelling with all the variables and removes one at a time. Whether to add or remove a variable depends on how much influence on the model goodness measure if variables were added or removed. In this project, we use backward selection method, and we use Receiver Operating Characteristic (ROC) Curve, which is a plot of true positive against the false positive, for goodness measure.

For the wrapper step, the desired output is a list of variables sorted by their importance. The feature selection algorithm we used is the RFE function from *sklearn.feature_selection* package, which selects features by recursively considering smaller and smaller sets of features. However, running models recursively is time-consuming. Thus, we chose to use a simple linear model – logistic regression, for better efficiency. We also rescaled the data using *sklearn.preprocessing.StandardScaler* to avoid exceeding the iteration limit. After preparing the data and model, we put all 80 variables selected from the filter step into the wrapper, set the number of features to select (n_features_to_select) to 30. As a result, the RFE function gave us a list of 30 variables. These would be our final variables to put in the fraud score prediction model.

5.4 Selected Variables

Table 5.1 Top 30 Variables with Highest Predictive Power

No.	Variable Name	Description
1	fulladdress_count_0_by_7	the number of fulladdress appearances per day within the last 7 days
2	fulladdress_homephone_day_since	the number of days since the same fulladdress and homephone combination occurred in this field
3	name_dob_count_14	the number of appearances of the same name and dob combination within 14 days
4	name_dob_count_30	the number of appearances of the same name and dob combination within 30 days
5	name_dob_day_since	the number of days since the same name and dob combination occurred in this field
6	ssn_dob_count_30	the number of appearances of the same ssn and dob combination within 30 days
7	ssn_dob_day_since	the number of days since the same ssn and dob combination occurred in this field
8	ssn_firstname_count_30	the number of appearances of the same ssn and firstname combination within 30 days
9	ssn_lastname_count_30	the number of appearances of the same ssn and lastname combination within 30 days
10	ssn_name_dob_count_30	the number of appearances of the same ssn, name and dob combination within 30 days
11	ssn_name_dob_day_since	the number of days since the same ssn, name and dob combination occurred in this field
12	unique_num_dob_for_each_fulladdress_30	the number of unique dob for a particular fulladdress over the past 30 days
13	unique_num_dob_for_each_ssn_name_dob_30	the number of unique dob for a particular ssn over the past 30 days
14	unique_num_homephone_for_each_ssn_30	the number of unique homephone for a particular ssn over the past 30 days
15	unique_num_name_for_each_ssn_30	the number of unique names for a particular ssn over the past 30 days
16	unique_num_name_for_each_ssn_dob_30	the number of unique names for a particular ssn and dob combination over the past 30 days
17	unique_num_name_fulladdress_for_each_ssn_30	the number of unique names and fulladdress combination for a particular ssn over the past 30 days
18	unique_num_name_fulladdress_for_each_ssn_dob_30	the number of unique names and fulladdress combination for a ssn and dob combination over the past 30 days

19	unique_num_name_homephone_for_each_ssn_30	the number of unique names and homephone combination for a ssn over the past 30 days
20	unique_num_ssn_for_each_fulladdress_30	the number of unique ssn for a fulladdress over the past 30 days
21	unique_num_ssn_for_each_ssn_name_dob_30	the number of unique ssn for a ssn, name and dob combination over the past 30 days
22	unique_num_ssn_fulladdress_for_each_fulladdress_30	the number of unique ssn and fulladdress combination for a fulladdress over the past 30 days
23	unique_num_ssn_fulladdress_for_each_ssn_30	the number of unique ssn and fulladdress combination for a ssn over the past 30 days
24	unique_num_ssn_fulladdress_for_each_ssn_name_dob_30	the number of unique ssn and fulladdress combination for a ssn and name combination over the past 30 days
25	unique_num_ssn_name_dob_for_each_ssn_30	the number of unique ssn and name combination for a ssn over the past 30 days
26	unique_num_ssn_name_dob_for_each_ssn_dob_30	the number of unique ssn and name combination for a ssn and dob combination over the past 30 days
27	unique_num_ssn_name_fulladdress_for_each_fulladdress_30	the number of unique ssn, name and fulladdress combination for each fulladdress over the past 30 days
28	unique_num_ssn_name_fulladdress_for_each_ssn_30	the number of unique ssn, name and fulladdress combination for each ssn over the past 30 days
29	unique_num_ssn_name_fulladdress_for_each_ssn_name_dob_30	the number of unique ssn, name fulladdress combination for each ssn, name and dob over the past 30 days
30	unique_num_zip5_for_each_fulladdress_7	the number of unique zip5 for each fulladdress over the past 30 days

6. Model Algorithms

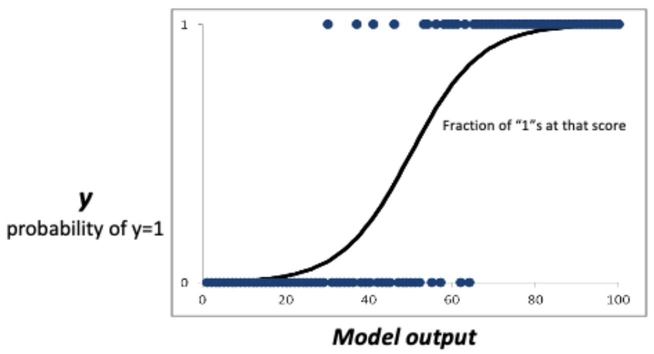
6.1 Algorithm Description

After feature engineering and feature selection, we got down to 30 high performance features, and we were ready to build a binary classification model using a wide range of machine learning algorithms. In this section, we will include a brief description for each algorithm used, treatment for imbalanced dataset, a high-level summary of performance for each model (FDR for training, testing, and OOT), and our final model choice.

We built the following machine learning algorithms.

(1) Logistic Regression

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable (i.e., fraud label). Logistic regression is estimating the parameters of logistic model given all the X parameters for binary classification problem with logit function.



$$y = \frac{1}{1 + e^{-\sum b_i x_i}}$$

Figure 6.1 Logistic Regression

We used logistic regression as our base model, which meant all other models should perform better than logistic regression. We created two versions of logistic regression. For the first version, we used Ridge Regularization to minimize the mean square error, and for the second version, we did not apply any regularization. For other model parameters, we kept the same for both model with default values.

(2) Gradient Boosted Trees

Gradient Boosting is a machine learning technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. Decision tree divides the independent variable (x's) space into boxes and places a step above each box at the height of the average of the dependent variable y in that box. A two-dimensional decision tree is shown below:

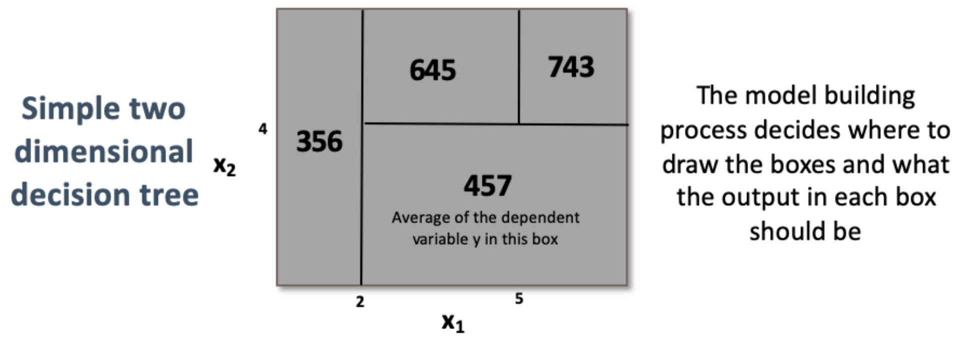


Figure 6.2 Decision Tree

When a decision tree is the weak learner, the resulting algorithm is called the gradient boosted tree. A boosted tree starts with a very simply model, then keep adding in weak models to get closer to the right answer. Each weak model is trained to predict the residual error of the current sum, and each added model makes the overall model slightly better.

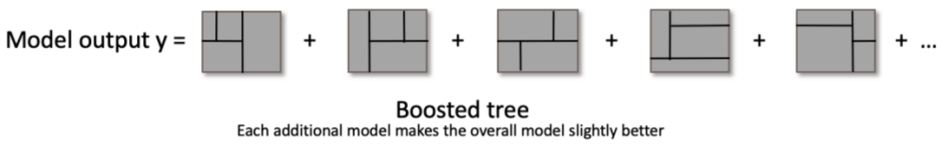


Figure 6.3 Boosted Tree

We tried three versions of boosted trees.

Gradient Boosting Classifier **(GB)** builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions, and it is capable of multiclass classification. For each class, a tree is fit on the negative gradient of the binomial or multinomial deviance loss function. Binary classification is a special case where only a single regression tree is induced.

Extreme gradient boosting (**XGBoost)** is another ensemble model, which is constructed from decision tree. We wanted to use this method to compare with random forest because XGboost uses the method called gradient boosting as the loss gradient. Specifically, models are fit using any arbitrary differentiable loss function and gradient descent optimization algorithm. We focused on couple parameters, such as learning rate, number of estimators to see the model performance by adjusting different parameters.

Light Gradient Boosting Machine **(LightGBM)** is an open source distributed gradient boosting framework for machine learning originally developed by Microsoft. It is based on decision tree algorithms and used for ranking, classification and other machine learning tasks. Instead of growing a tree level-wise, LightGBM grows a tree leaf-wise. It chooses the leaf it believes will yield the largest decrease in loss. Besides, LightGBM does not use the widely used sorted-based decision tree learning algorithm, which searches the best split point on sorted features values. Instead, LightGBM implements a highly optimized histogram-based decision tree learning algorithm, which yields great advantages on both efficiency and memory consumption.

(3) Random Forest

Random forest is an ensemble learning method for classification that operate by constructing multitude of decision trees. We used random forest to detect possible non-linear boundary in the feature space. One advantage of random forest is the ensembles of the model. It created many trees and then output the mode of classes for different tree algorithms.

We created 8 random forest models with different parameters. We specifically selected and modified number of estimators, the depth of each estimator and the smallest sample needed to split a node. Generally, we wanted a deep tree for the random forest because we can avoid overfitting by creating N different trees.

(4) Neural Network

A neural network is a machine learning algorithm construct that maps an input vector to an output scalar, or typically a vector of axes into a single dependent variable. It is inspired by the biological neural networks that constitute human brains. A typical neural network consists of an input layer, some hidden layers and an output layer.

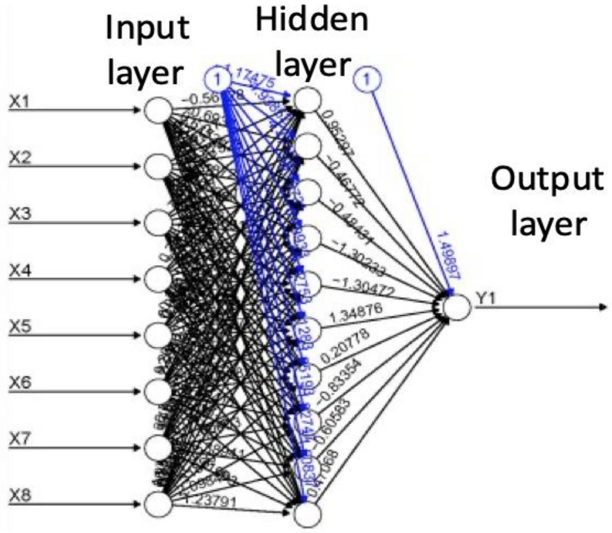


Figure 6.4 Neural Network

The input layer is formed by all independent variables(x's), so it has as many nodes as the input vector dimension. The output layer, which is the last layer, has as many nodes as the output dependent variable dimension, which is typically one. We also have many hidden layers in between the input layer and the output layer. Each node in the hidden layer receives weighted signals from all nodes in the previous layer and then go through some kinds of activation/transformation function, such as logistic function and sigmoid function, to output a single dependent variable (y).

The node weight is trained by backpropagating the error. For each training record shown to the neural net, an error is calculated and propagates backward to each node. Then, neural net would calculate the gradient of that error, with respect to the weight in that node. The gradient, associated with a learning rate, is used to slightly adjust each node weight. Each record is passed through many times as the weights settle into a local optimum.

6.2 Treatment for Imbalanced Dataset

For training datasets, we implemented Synthetic Minority Oversampling Technique (SMOTE) as an oversampling method to create artificial "bad" record.

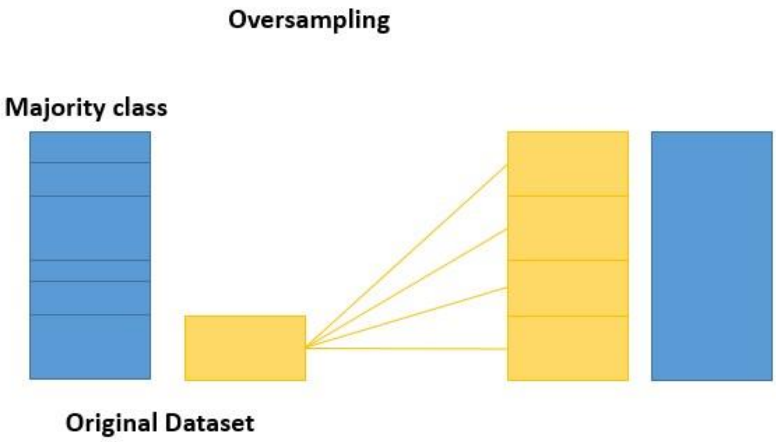


Figure 6.5 Oversampling

Specifically, SMOTE takes a "bad" record from the dataset and consider its k nearest neighbors (in the feature space). Then, it invents new "bad" records that lie somewhere between neighboring know "bad" records. Thus, for each training dataset, we were able to get equal numbers of fraud and non-fraud records.

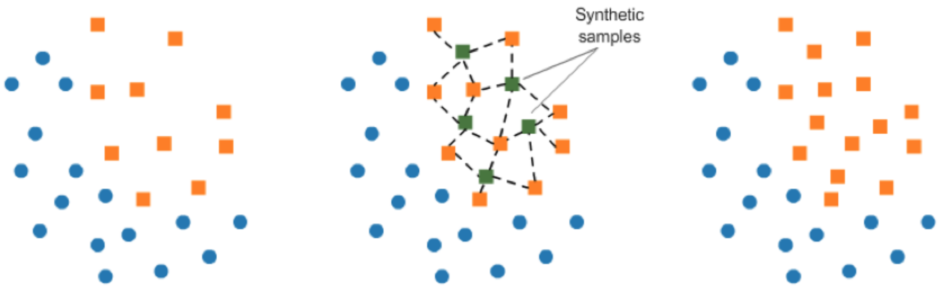


Figure 6.6 Synthetic Minority Oversampling Technique

6.3 Summary of Model Exploration

For each set of hyperparameters of each model, we conducted 10 times cross-validation and calculated the average FDR at 3% of population for training, testing, and OOT dataset. For each validation, we randomly divided our dataset (excluded OOT and first two-weeks of data) into training and testing. We implemented SMOTE on training dataset and then fit our model. Then, we predicted on training, testing, and OOT datasets and calculated FDR for all of them. After 10 times cross-validation, we calculated our average FDR by taking the mean of 10 FDRs.

Table 6.1 Model Summary

Model	Parameters						Average FDR at 3%		
Logistic Regression	max_iter		Penalty				Train	Test	OOT
1	800		l2				52.6%	52.8%	50.9%
2	800		none				52.5%	52.8%	50.9%
lightgbm	# of trees	min_split_gain	max_depth	num_leaves	learning_rate	min_leaf_size	Train	Test	OOT
1	500	0	2	4	0.01	30	51.2%	51.2%	49.6%
2	500	0	4	10	0.1	30	54.1%	53.4%	48.4%
3	800	0	4	10	0.1	30	54.9%	53.5%	49.6%
4	800	0.3	6	25	0.1	40	54.1%	53.8%	49.9%
5	1000	0	4	12	0.1	40	54.4%	53.5%	50.1%
6	1000	0	6	25	0.05	40	54.2%	53.3%	51.1%
7	1500	0	4	10	0.05	40	54.0%	53.4%	50.8%
8	1500	0.3	6	50	0.1	50	53.7%	53.2%	51.2%
9	2000	0	2	4	0.1	50	53.6%	53.4%	52.2%
10	2000	0	4	8	0.1	50	54.2%	52.2%	51.7%
Random Forest	n_estimators		max_depth		min_samples_split		Train	Test	OOT
1	100		Default		Default (1)		55.1%	52.8%	49.6%
2	200		Default		Default (1)		55.1%	52.9%	48.8%
3	300		Default		Default (1)		55.1%	52.9%	49.0%
4	300		7		20		53.7%	53.7%	51.2%
5	300		10		200		54.1%	53.7%	51.9%
6	300		13		500		54.2%	54.0%	51.9%
7	300		15		20		54.4%	53.9%	51.6%
8	300		20		5		54.8%	53.6%	50.1%
Neural Net	hidden_layer_sizes	activation	learning_rate		learning_rate_init	alpha	Train	Test	OOT
1	(10,) 1 layer and 10 nodes	logistic	constant		0.001	0.001	52.9%	52.7%	52.3%
2	(10,)	tanh	constant		0.001	0.001	52.6%	52.7%	51.7%
3	(5,5)	logistic	invscaling		0.01	0.001	52.7%	52.1%	51.6%
4	(10,5)	logistic	constant		0.001	0.005	52.1%	52.2%	51.4%
5	(20,)	logistic	constant		0.001	0.001	52.5%	52.7%	51.9%
6	(10,10)	tanh	constant		0.001	0.001	52.3%	51.4%	51.3%
7	(20,5)	logistic	invscaling		0.005	0.0008	50.8%	50.4%	50.3%
8	(20,10)	logistic	constant		0.001	0.0001	52.6%	53.0%	51.6%
9	(20,20)	logistic	constant		0.001	0.0001	52.8%	53.4%	51.7%
GradientBoostingClassifier	# of trees	max_depth	min_samples_split		min_samples_leaf		Train	Test	OOT
1	500	3	1500		30		54.0%	53.5%	52.5%
2	800	**3**	1500		30		53.1%	52.9%	53.1%
3	800	5	1000		25		53.8%	53.4%	53.7%
4	1000	3	1500		25		53.7%	52.8%	53.1%
5	1000	5	1000		20		53.7%	53.3%	52.8%
6	1300	3	1200		20		53.9%	52.6%	53.3%
7	1300	5	1000		20		54.2%	52.2%	52.9%
Extreme Gradient Boosting	Learning Rate	min_child_weight	n_estimator		max_depth		Train	Test	OOT
1	0.3	1	100		default (6)		54.6%	53.9%	52.0%
2	0.1	20	250		default (6)		54.2%	53.9%	51.8%
3	0.1	20	150		10		54.2%	53.8%	51.9%
4	**0.05**	**100**	**150**		**15**		**54.0%**	**54.0%**	**51.9%**
5	0.01	500	150		10		52.2%	52.1%	51.6%
6	0.3	200	100		10		53.9%	53.9%	51.9%

The table above is our high-level summary of hyperparameter tuning and model performance. We decided to use **Extreme Gradient Boosting (XGBoost),** which has the highest average testing FDR, as our final model.

Table 6.2 Hyperparameters of Final Model

Final Model	Learning Rate	Min_child_weight	n_estimator	Max_depth
Xgboost	0.05	100	150	15

7. Final Model Results

Our model results are displayed in the three tables below. The first two tables are the training and testing results, which only covers the period from 2016-01-15 to 2016-11-01. The first two weeks' records are excluded because they are not reliable. The third table is OOT population result for the validation purpose.

For each population table, we assigned a fraud score to each record predicted using our best model. Then we sorted the table based on this fraud score in descending order. After sorting the table, we separated the population equally into 100 bins, each covering 1% of the table population. For each bin, bins statistics and cumulated statistics are calculated. Lastly, for details of the statistics calculated, please refer to the tables below, which show the top 20 bins' results. For the full result tables, kindly refer to Appendix 9.3.

Table 7.1 Training Population Results (top 20%)

Training	# Records		# Goods		# Bads		Fraud Rate					
	638172		628951		9221		0.014449083					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	6382	1645	4737	25.78%	74.22%	6382	1645	4737	0.26%	51.37%	51.11	0.35
2	6382	6213	169	97.35%	2.65%	12764	7858	4906	1.25%	53.20%	51.96	1.60
3	6382	6315	67	98.95%	1.05%	19146	14173	4973	2.25%	53.93%	51.68	2.85
4	6382	6332	50	99.22%	0.78%	25528	20505	5023	3.26%	54.47%	51.21	4.08
5	6382	6337	45	99.29%	0.71%	31910	26842	5068	4.27%	54.96%	50.69	5.30
6	6382	6327	55	99.14%	0.86%	38292	33169	5123	5.27%	55.56%	50.28	6.47
7	6382	6330	52	99.19%	0.81%	44674	39499	5175	6.28%	56.12%	49.84	7.63
8	6382	6341	41	99.36%	0.64%	51056	45840	5216	7.29%	56.57%	49.28	8.79
9	6382	6330	52	99.19%	0.81%	57438	52170	5268	8.29%	57.13%	48.84	9.90
10	6382	6335	47	99.26%	0.74%	63820	58505	5315	9.30%	57.64%	48.34	11.01
11	6382	6335	47	99.26%	0.74%	70202	64840	5362	10.31%	58.15%	47.84	12.09
12	6382	6331	51	99.20%	0.80%	76584	71171	5413	11.32%	58.70%	47.39	13.15
13	6382	6337	45	99.29%	0.71%	82966	77508	5458	12.32%	59.19%	46.87	14.20
14	6382	6338	44	99.31%	0.69%	89348	83846	5502	13.33%	59.67%	46.34	15.24
15	6382	6334	48	99.25%	0.75%	95730	90180	5550	14.34%	60.19%	45.85	16.25
16	6382	6347	35	99.45%	0.55%	102112	96527	5585	15.35%	60.57%	45.22	17.28
17	6382	6336	46	99.28%	0.72%	108494	102863	5631	16.35%	61.07%	44.71	18.27
18	6382	6345	37	99.42%	0.58%	114876	109208	5668	17.36%	61.47%	44.10	19.27
19	6382	6332	50	99.22%	0.78%	121258	115540	5718	18.37%	62.01%	43.64	20.21
20	6382	6331	51	99.20%	0.80%	127640	121871	5769	19.38%	62.56%	43.19	21.13

Table 7.2 Testing Population Results (top 20%)

Testing	# Records		# Goods		# Bads		Fraud Rate				
	159543		157239		2304		0.014441248				
	Bins Statistics					Cumulative Statistics					
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	1596	404	1192	25.31%	74.69%	1596	404	1192	0.26%	51.74%	51.48	0.34
2	1596	1549	47	97.06%	2.94%	3192	1953	1239	1.24%	53.78%	52.53	1.58
3	1596	1588	8	99.50%	0.50%	4788	3541	1247	2.25%	54.12%	51.87	2.84
4	1596	1586	10	99.37%	0.63%	6384	5127	1257	3.26%	54.56%	51.30	4.08
5	1596	1587	9	99.44%	0.56%	7980	6714	1266	4.27%	54.95%	50.68	5.30
6	1596	1588	8	99.50%	0.50%	9576	8302	1274	5.28%	55.30%	50.02	6.52
7	1596	1585	11	99.31%	0.69%	11172	9887	1285	6.29%	55.77%	49.48	7.69
8	1596	1589	7	99.56%	0.44%	12768	11476	1292	7.30%	56.08%	48.78	8.88
9	1596	1580	16	99.00%	1.00%	14364	13056	1308	8.30%	56.77%	48.47	9.98
10	1596	1582	14	99.12%	0.88%	15960	14638	1322	9.31%	57.38%	48.07	11.07
11	1596	1591	5	99.69%	0.31%	17556	16229	1327	10.32%	57.60%	47.27	12.23
12	1596	1578	18	98.87%	1.13%	19152	17807	1345	11.32%	58.38%	47.05	13.24
13	1596	1581	15	99.06%	0.94%	20748	19388	1360	12.33%	59.03%	46.70	14.26
14	1596	1589	7	99.56%	0.44%	22344	20977	1367	13.34%	59.33%	45.99	15.35
15	1596	1584	12	99.25%	0.75%	23940	22561	1379	14.35%	59.85%	45.50	16.36
16	1596	1589	7	99.56%	0.44%	25536	24150	1386	15.36%	60.16%	44.80	17.42
17	1596	1585	11	99.31%	0.69%	27132	25735	1397	16.37%	60.63%	44.27	18.42
18	1596	1587	9	99.44%	0.56%	28728	27322	1406	17.38%	61.02%	43.65	19.43
19	1596	1581	15	99.06%	0.94%	30324	28903	1421	18.38%	61.68%	43.29	20.34
20	1596	1584	12	99.25%	0.75%	31920	30487	1433	19.39%	62.20%	42.81	21.27

Table 7.3 Out Of Time Population Results (top 20%)

OOT	# Records		# Goods		# Bads		Fraud Rate					
	163771		161427		2344		0.014312668					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	1638	512	1126	31.26%	68.74%	1638	512	1126	0.32%	48.04%	47.72	0.45
2	1638	1563	75	95.42%	4.58%	3276	2075	1201	1.29%	51.24%	49.95	1.73
3	1638	1629	9	99.45%	0.55%	4914	3704	1210	2.29%	51.62%	49.33	3.06
4	1638	1622	16	99.02%	0.98%	6552	5326	1226	3.30%	52.30%	49.00	4.34
5	1638	1619	19	98.84%	1.16%	8190	6945	1245	4.30%	53.11%	48.81	5.58
6	1638	1628	10	99.39%	0.61%	9828	8573	1255	5.31%	53.54%	48.23	6.83
7	1638	1623	15	99.08%	0.92%	11466	10196	1270	6.32%	54.18%	47.86	8.03
8	1638	1625	13	99.21%	0.79%	13104	11821	1283	7.32%	54.74%	47.41	9.21
9	1638	1628	10	99.39%	0.61%	14742	13449	1293	8.33%	55.16%	46.83	10.40
10	1638	1618	20	98.78%	1.22%	16380	15067	1313	9.33%	56.02%	46.68	11.48
11	1638	1629	9	99.45%	0.55%	18018	16696	1322	10.34%	56.40%	46.06	12.63
12	1638	1629	9	99.45%	0.55%	19656	18325	1331	11.35%	56.78%	45.43	13.77
13	1638	1626	12	99.27%	0.73%	21294	19951	1343	12.36%	57.30%	44.94	14.86
14	1638	1624	14	99.15%	0.85%	22932	21575	1357	13.37%	57.89%	44.53	15.90
15	1638	1626	12	99.27%	0.73%	24570	23201	1369	14.37%	58.40%	44.03	16.95
16	1638	1626	12	99.27%	0.73%	26208	24827	1381	15.38%	58.92%	43.54	17.98
17	1638	1621	17	98.96%	1.04%	27846	26448	1398	16.38%	59.64%	43.26	18.92
18	1638	1629	9	99.45%	0.55%	29484	28077	1407	17.39%	60.03%	42.63	19.96
19	1638	1623	15	99.08%	0.92%	31122	29700	1422	18.40%	60.67%	42.27	20.89
20	1638	1630	8	99.51%	0.49%	32760	31330	1430	19.41%	61.01%	41.60	21.91

8. Conclusions

Throughout the project, we have built a real time fraud algorithm to identify application fraud. Given the 1,000,000 records of application data, we delicately checked data quality. We then carefully dealt with frivolous field values. With cleaned data ready, we created 680 candidate variables, which took time flow characteristics into consideration. Then, we selected the top 30 variables in terms of predictive power for model building. During the model building phase, we began with building baseline logistic regression models with different combination of parameters. Overall, these models performed reasonably good. We also tried a wide range of non-linear models (Gradient Boosted Trees, Random Forest and Neural Network) and tuned the hyperparameters to achieve the best model performance. With the performance results for all these models, we finally selected an extreme gradient boosting model with all parameters set out in section 6 and 7. Eventually, we were able to achieve a fraud detection rate of 51.62% at 3% population on OOT data.

We could make further improvements to our model in several aspects if given more time. Firstly, we can create more candidate variables. Currently we have built 680 variables. If our computer capacity and time allow, we can build as many as thousands of variables. Also, if time permits, we can consult with several domain experts for advice in building expert variables. These might change our best 30 variables used for model building and consequently the results. In addition, we can also include external data, which associates with each application, to give us more insights about each record.

9. Appendix

9.1 Data Quality Report for Application Fraud

(1) Description

Dataset Name: Product Applications Data
Dataset Purpose: Real-time product application data to identify fraud.
Data Source: Synthetic data generated from real data distribution, over a billion real U.S. applications over about 10 years. Built by an identity fraud prevention company to reproduce the important univariate and multivariate field distributions of real data.
Time Period: 01/01/2016 – 12/31/2016
Number of Fields: 10
Number of Records: 1,000,000

(2) Summary Tables

Table 9.1 Summary Table for Categorical Fields

Column Name	# of Records	% populated	Unique Values	Most Common Field Value
ssn	1,000,000	100	835819	999999999
firstname	1,000,000	100	78136	EAMSTRMT
lastname	1,000,000	100	177001	ERJSAXA
address	1,000,000	100	828774	123 MAIN ST
zip5	1,000,000	100	26370	68138
dob	1,000,000	100	42673	19070626
homephone	1,000,000	100	28244	9999999999
fraud_label	1,000,000	100	2	0

Table 9.2 Summary Table for Datetime Fields

Column Name	# of Records	% populated	Minimum	Maximum	Most Common Field Value
date	1,000,000	100	2016-01-01	2016-12-31	2016-08-16
dob	1,000,000	100	1900-01-01	2016-10-31	1907-06-26

(3) Data Field Exploration

Field 1:
Name: record
Description: Unique identifier of each application in the data.

Field 2:
Name: date
Description: Date of each application with the format of year-month-day.
All happened in 2016. Frequency of each month/day/day of week is almost same.

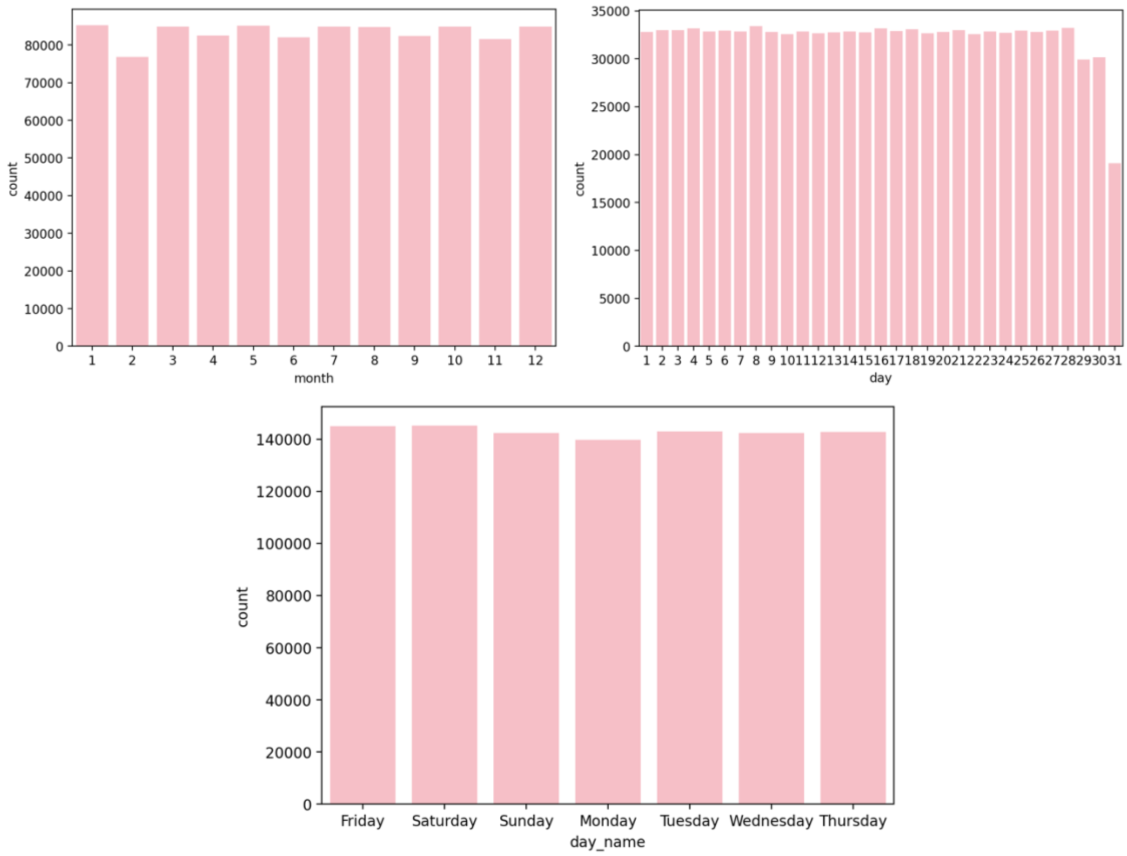


Figure 9.1 Histograms of Date

Field 3:
Name: ssn
Description: Social security number of each applicant.
% of 999999999: 1.693%

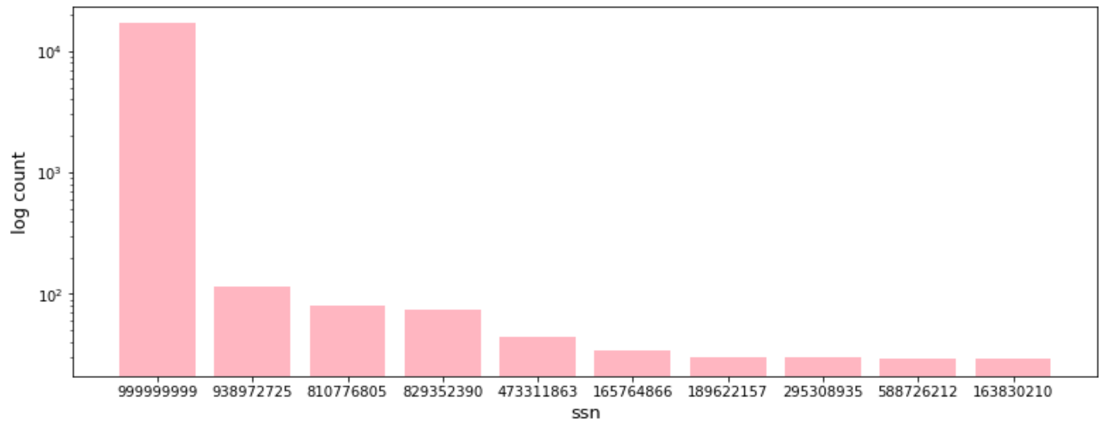


Figure 9.2 Histogram of Social Security Number

Field 4:
Name: firstname
Description: First name of each applicant.

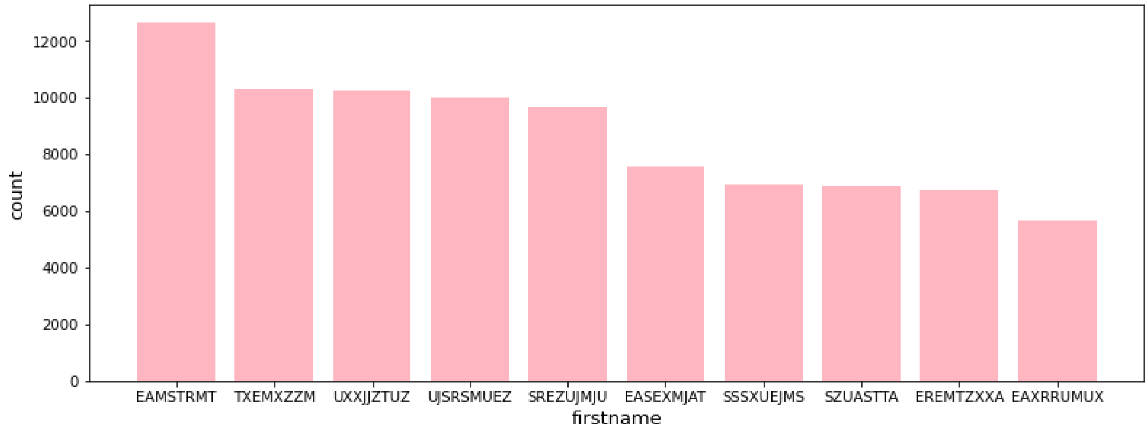


Figure 9.3 Histogram of First Name

Field 5:
Name: lastname
Description: Last name of each applicant.

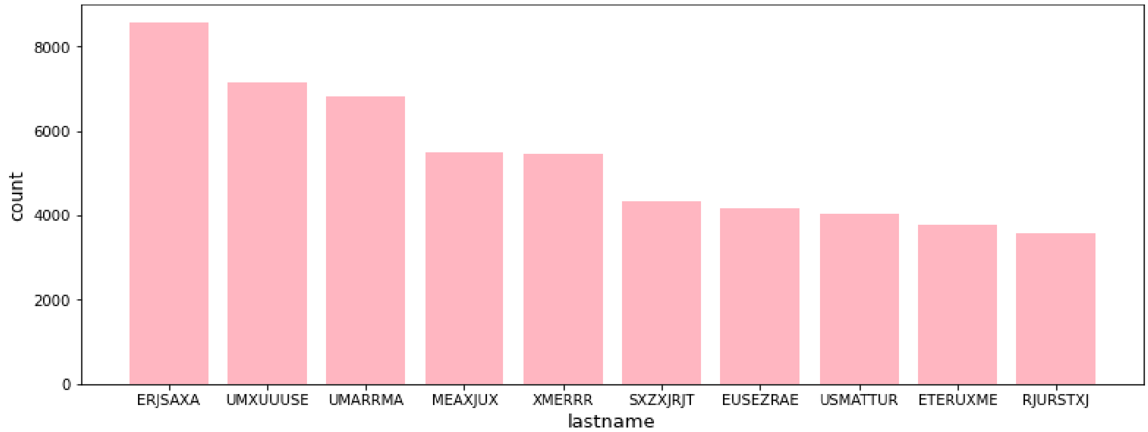


Figure 9.4 Histogram of Last Name

Field 6:
Name: address
Description: Address of each applicant.
% of 123 MAIN ST: 0.108%

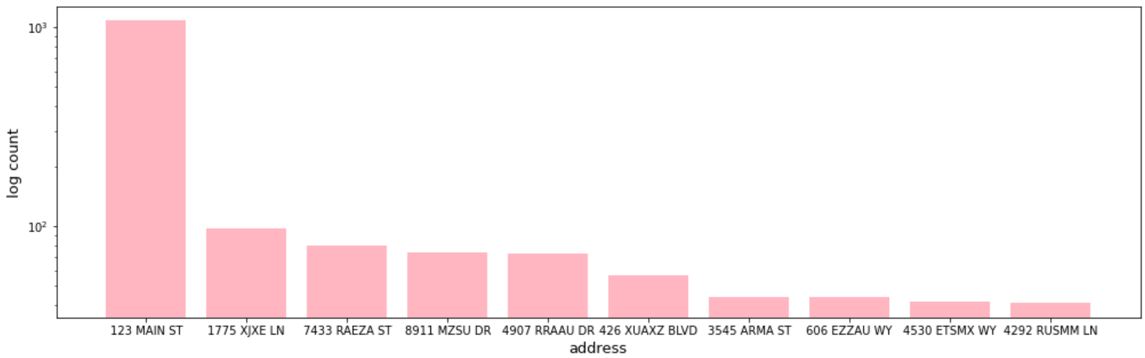


Figure 9.5 Histogram of Address

Field 7:
Name: zip5
Description: Zip5 code of each applicant.

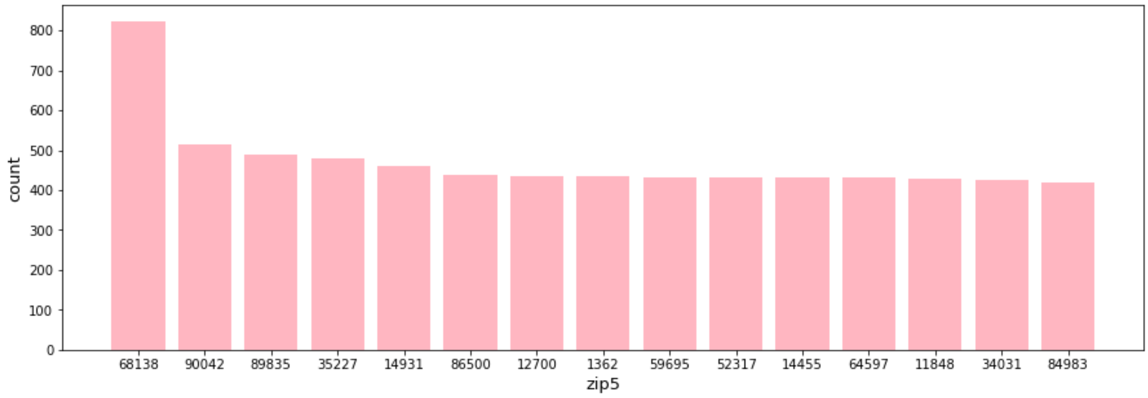


Figure 9.6 Histogram of Zip5 Code

Field 8:
Name: dob
Description: Date of birth of each applicant.
% of 06/26/1907: 12.657%

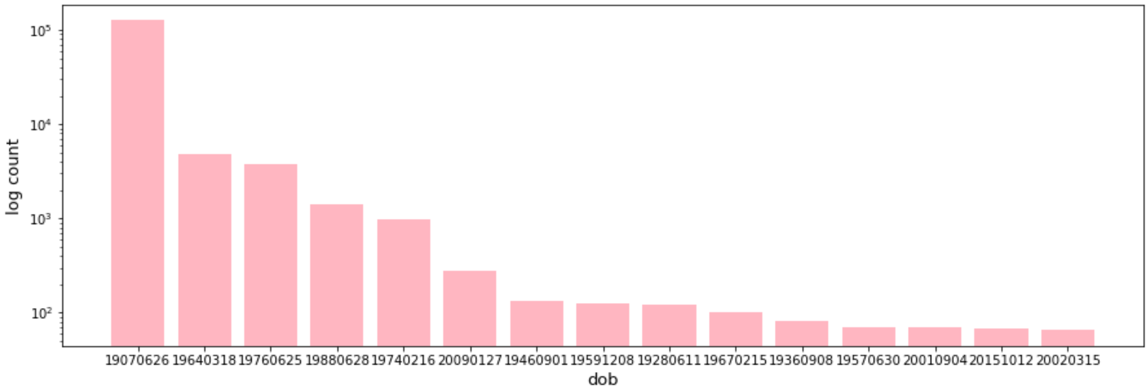


Figure 9.7 Histogram of Date Of Birth

Field 9:
Name: homephone
Description: Home phone number of each applicant.
% of 9999999999: 7.851%

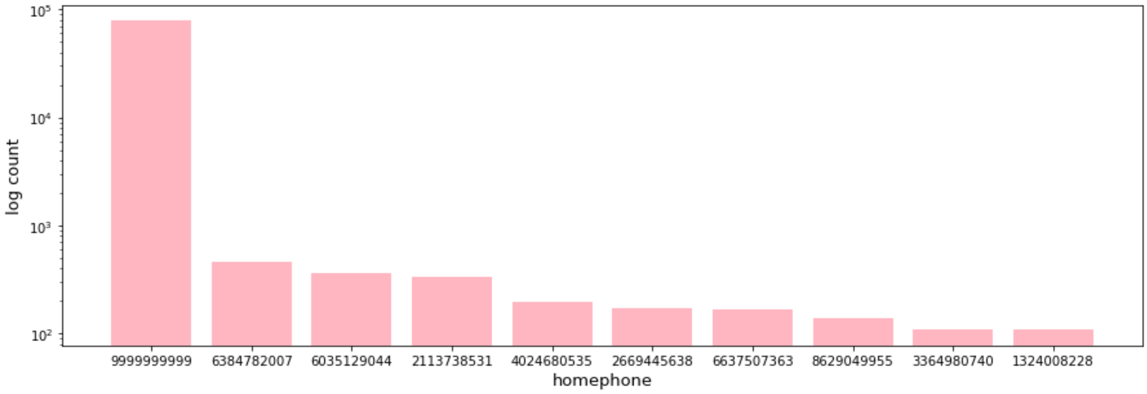


Figure 9.8 Histogram of Home Phone Number

Field 10:
Name: fraud_label
Description: Fraud label for each applicant.
1.439% applicants are labeled as fraudulent.

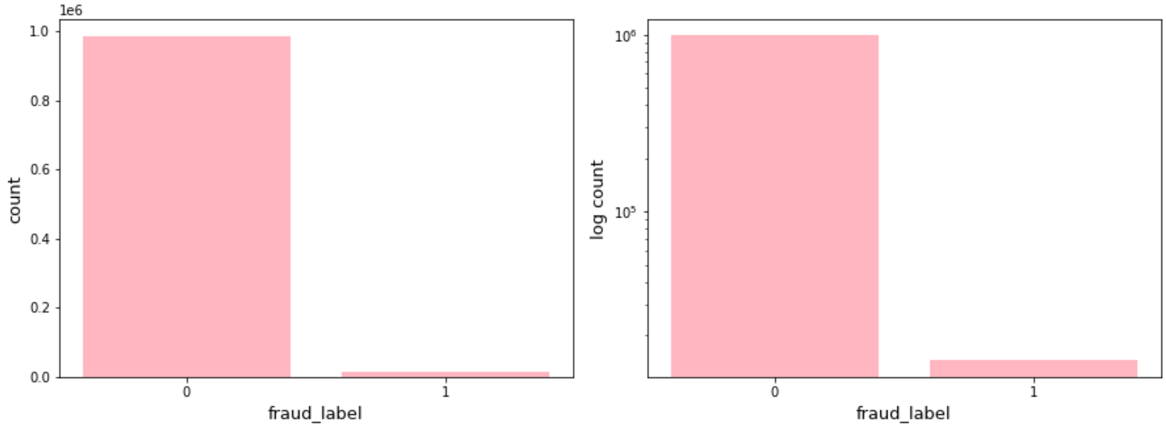


Figure 9.9 Histograms of Fraud Label

9.2 Feature Engineering Result

Table 9.3 Statistics of Candidate Variables

No.	Variable name	mean	std	min	50%	max
1	age	57.80	31.40	-1.00	57.00	117.00
2	dob_count_0	1.09	0.73	1.00	1.00	23.00
3	dob_count_0_by_14	10.24	4.14	0.06	14.00	14.00
4	dob_count_0_by_3	2.76	0.58	0.05	3.00	3.00
5	dob_count_0_by_30	16.94	9.37	0.07	15.00	30.00
6	dob_count_0_by_7	5.90	1.78	0.06	7.00	7.00
7	dob_count_1	1.25	1.79	1.00	1.00	40.00
8	dob_count_14	3.27	15.90	1.00	1.00	236.00
9	dob_count_1_by_14	10.52	4.01	0.09	14.00	14.00
10	dob_count_1_by_3	2.84	0.46	0.08	3.00	3.00
11	dob_count_1_by_30	17.42	9.29	0.08	15.00	30.00
12	dob_count_1_by_7	6.06	1.64	0.09	7.00	7.00
13	dob_count_3	1.57	3.97	1.00	1.00	72.00
14	dob_count_30	5.63	32.79	1.00	2.00	446.00
15	dob_count_7	2.19	8.33	1.00	1.00	129.00
16	dob_day_since	75.83	131.49	0.00	15.00	365.00
17	dob_homephone_count_0	1.00	0.02	1.00	1.00	3.00
18	dob_homephone_count_0_by_14	13.92	0.76	1.75	14.00	14.00
19	dob_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
20	dob_homephone_count_0_by_30	29.64	2.33	2.73	30.00	30.00
21	dob_homephone_count_0_by_7	6.98	0.27	1.17	7.00	7.00
22	dob_homephone_count_1	1.00	0.04	1.00	1.00	3.00
23	dob_homephone_count_14	1.01	0.12	1.00	1.00	8.00
24	dob_homephone_count_1_by_14	13.92	0.73	1.75	14.00	14.00
25	dob_homephone_count_1_by_3	3.00	0.06	0.75	3.00	3.00
26	dob_homephone_count_1_by_30	29.65	2.29	2.73	30.00	30.00
27	dob_homephone_count_1_by_7	6.98	0.25	1.17	7.00	7.00
28	dob_homephone_count_3	1.00	0.06	1.00	1.00	5.00
29	dob_homephone_count_30	1.03	0.17	1.00	1.00	11.00
30	dob_homephone_count_7	1.01	0.08	1.00	1.00	6.00
31	dob_homephone_day_since	332.80	89.73	0.00	365.00	365.00
32	dow_risk	0.01	0.00	0.01	0.01	0.02
33	fulladdress_count_0	1.01	0.28	1.00	1.00	24.00
34	fulladdress_count_0_by_14	13.87	1.02	0.64	14.00	14.00
35	fulladdress_count_0_by_3	2.99	0.14	0.18	3.00	3.00
36	fulladdress_count_0_by_30	29.49	2.83	1.30	30.00	30.00
37	fulladdress_count_0_by_7	6.96	0.41	0.32	7.00	7.00
38	fulladdress_count_1	1.02	0.48	1.00	1.00	30.00
39	fulladdress_count_14	1.05	0.61	1.00	1.00	30.00
40	fulladdress_count_1_by_14	13.89	0.89	0.93	14.00	14.00
41	fulladdress_count_1_by_3	2.99	0.09	0.43	3.00	3.00
42	fulladdress_count_1_by_30	29.54	2.64	1.30	30.00	30.00
43	fulladdress_count_1_by_7	6.97	0.33	0.58	7.00	7.00
44	fulladdress_count_3	1.03	0.54	1.00	1.00	30.00
45	fulladdress_count_30	1.06	0.63	1.00	1.00	30.00
46	fulladdress_count_7	1.04	0.58	1.00	1.00	30.00
47	fulladdress_day_since	325.10	99.04	0.00	365.00	365.00
48	fulladdress_dob_count_0	1.00	0.02	1.00	1.00	3.00
49	fulladdress_dob_count_0_by_14	13.92	0.77	1.75	14.00	14.00
50	fulladdress_dob_count_0_by_3	3.00	0.08	0.75	3.00	3.00
51	fulladdress_dob_count_0_by_30	29.63	2.34	2.73	30.00	30.00
52	fulladdress_dob_count_0_by_7	6.98	0.27	1.17	7.00	7.00
53	fulladdress_dob_count_1	1.00	0.04	1.00	1.00	3.00
54	fulladdress_dob_count_14	1.01	0.12	1.00	1.00	8.00
55	fulladdress_dob_count_1_by_14	13.92	0.74	1.75	14.00	14.00
56	fulladdress_dob_count_1_by_3	3.00	0.06	0.75	3.00	3.00
57	fulladdress_dob_count_1_by_30	29.65	2.30	2.73	30.00	30.00
58	fulladdress_dob_count_1_by_7	6.98	0.25	1.17	7.00	7.00
59	fulladdress_dob_count_3	1.00	0.06	1.00	1.00	5.00

60	fulladdress_dob_count_30	1.03	0.18	1.00	1.00	11.00
61	fulladdress_dob_count_7	1.01	0.08	1.00	1.00	6.00
62	fulladdress_dob_day_since	332.44	90.13	0.00	365.00	365.00
63	fulladdress_homephone_count_0	1.01	0.22	1.00	1.00	21.00
64	fulladdress_homephone_count_0_by_14	13.89	0.91	0.64	14.00	14.00
65	fulladdress_homephone_count_0_by_3	2.99	0.12	0.18	3.00	3.00
66	fulladdress_homephone_count_0_by_30	29.57	2.59	1.30	30.00	30.00
67	fulladdress_homephone_count_0_by_7	6.97	0.36	0.32	7.00	7.00
68	fulladdress_homephone_count_1	1.02	0.38	1.00	1.00	30.00
69	fulladdress_homephone_count_14	1.03	0.48	1.00	1.00	30.00
70	fulladdress_homephone_count_1_by_14	13.91	0.82	0.93	14.00	14.00
71	fulladdress_homephone_count_1_by_3	3.00	0.08	0.50	3.00	3.00
72	fulladdress_homephone_count_1_by_30	29.60	2.45	1.30	30.00	30.00
73	fulladdress_homephone_count_1_by_7	6.98	0.30	0.58	7.00	7.00
74	fulladdress_homephone_count_3	1.02	0.43	1.00	1.00	30.00
75	fulladdress_homephone_count_30	1.05	0.51	1.00	1.00	30.00
76	fulladdress_homephone_count_7	1.03	0.46	1.00	1.00	30.00
77	fulladdress_homephone_day_since	329.79	93.68	0.00	365.00	365.00
78	homephone_count_0	1.07	0.36	1.00	1.00	22.00
79	homephone_count_0_by_14	7.50	4.28	0.39	7.00	14.00
80	homephone_count_0_by_3	2.54	0.73	0.09	3.00	3.00
81	homephone_count_0_by_30	10.97	8.78	0.63	7.50	30.00
82	homephone_count_0_by_7	4.89	2.08	0.21	3.50	7.00
83	homephone_count_1	1.20	0.63	1.00	1.00	31.00
84	homephone_count_14	2.76	1.74	1.00	2.00	36.00
85	homephone_count_1_by_14	7.98	4.22	0.41	7.00	14.00
86	homephone_count_1_by_3	2.69	0.61	0.12	3.00	3.00
87	homephone_count_1_by_30	11.66	8.76	0.65	7.50	30.00
88	homephone_count_1_by_7	5.20	1.97	0.21	7.00	7.00
89	homephone_count_3	1.45	0.86	1.00	1.00	32.00
90	homephone_count_30	4.60	2.86	1.00	4.00	50.00
91	homephone_count_7	1.93	1.21	1.00	2.00	33.00
92	homephone_day_since	47.38	110.70	0.00	6.00	365.00
93	homephone_name_dob_count_0	1.00	0.02	1.00	1.00	3.00
94	homephone_name_dob_count_0_by_14	13.92	0.76	1.75	14.00	14.00
95	homephone_name_dob_count_0_by_3	3.00	0.08	0.75	3.00	3.00
96	homephone_name_dob_count_0_by_30	29.64	2.31	2.73	30.00	30.00
97	homephone_name_dob_count_0_by_7	6.98	0.27	1.17	7.00	7.00
98	homephone_name_dob_count_1	1.00	0.04	1.00	1.00	3.00
99	homephone_name_dob_count_14	1.01	0.12	1.00	1.00	8.00
100	homephone_name_dob_count_1_by_14	13.92	0.73	1.75	14.00	14.00
101	homephone_name_dob_count_1_by_3	3.00	0.06	0.75	3.00	3.00
102	homephone_name_dob_count_1_by_30	29.65	2.27	2.73	30.00	30.00
103	homephone_name_dob_count_1_by_7	6.98	0.25	1.17	7.00	7.00
104	homephone_name_dob_count_3	1.00	0.06	1.00	1.00	5.00
105	homephone_name_dob_count_30	1.03	0.17	1.00	1.00	11.00
106	homephone_name_dob_count_7	1.01	0.08	1.00	1.00	6.00
107	homephone_name_dob_day_since	333.26	89.16	0.00	365.00	365.00
108	name_count_0	1.01	0.23	1.00	1.00	21.00
109	name_count_0_by_14	13.57	1.79	0.64	14.00	14.00
110	name_count_0_by_3	2.97	0.20	0.14	3.00	3.00
111	name_count_0_by_30	28.39	5.09	1.36	30.00	30.00
112	name_count_0_by_7	6.88	0.68	0.32	7.00	7.00
113	name_count_1	1.02	0.39	1.00	1.00	34.00
114	name_count_14	1.10	0.60	1.00	1.00	34.00
115	name_count_1_by_14	13.61	1.68	0.82	14.00	14.00
116	name_count_1_by_3	2.98	0.16	0.43	3.00	3.00
117	name_count_1_by_30	28.46	4.94	1.50	30.00	30.00
118	name_count_1_by_7	6.90	0.60	0.47	7.00	7.00
119	name_count_3	1.04	0.45	1.00	1.00	34.00
120	name_count_30	1.18	0.84	1.00	1.00	34.00
121	name_count_7	1.06	0.51	1.00	1.00	34.00
122	name_day_since	284.73	134.20	0.00	365.00	365.00

123	name_dob_count_0	1.01	0.22	1.00	1.00	21.00
124	name_dob_count_0_by_14	13.89	0.89	0.64	14.00	14.00
125	name_dob_count_0_by_3	2.99	0.12	0.14	3.00	3.00
126	name_dob_count_0_by_30	29.58	2.55	1.36	30.00	30.00
127	name_dob_count_0_by_7	6.97	0.35	0.32	7.00	7.00
128	name_dob_count_1	1.01	0.38	1.00	1.00	34.00
129	name_dob_count_14	1.03	0.47	1.00	1.00	34.00
130	name_dob_count_1_by_14	13.91	0.81	0.88	14.00	14.00
131	name_dob_count_1_by_3	3.00	0.08	0.43	3.00	3.00
132	name_dob_count_1_by_30	29.62	2.41	1.76	30.00	30.00
133	name_dob_count_1_by_7	6.98	0.29	0.50	7.00	7.00
134	name_dob_count_3	1.02	0.42	1.00	1.00	34.00
135	name_dob_count_30	1.05	0.50	1.00	1.00	34.00
136	name_dob_count_7	1.03	0.45	1.00	1.00	34.00
137	name_dob_day_since	331.11	92.16	0.00	365.00	365.00
138	name_fulladdress_count_0	1.00	0.02	1.00	1.00	3.00
139	name_fulladdress_count_0_by_14	13.91	0.80	1.75	14.00	14.00
140	name_fulladdress_count_0_by_3	3.00	0.08	0.75	3.00	3.00
141	name_fulladdress_count_0_by_30	29.59	2.46	2.73	30.00	30.00
142	name_fulladdress_count_0_by_7	6.98	0.29	1.17	7.00	7.00
143	name_fulladdress_count_1	1.00	0.04	1.00	1.00	3.00
144	name_fulladdress_count_14	1.01	0.12	1.00	1.00	8.00
145	name_fulladdress_count_1_by_14	13.91	0.77	1.75	14.00	14.00
146	name_fulladdress_count_1_by_3	3.00	0.07	0.75	3.00	3.00
147	name_fulladdress_count_1_by_30	29.61	2.41	2.73	30.00	30.00
148	name_fulladdress_count_1_by_7	6.98	0.27	1.17	7.00	7.00
149	name_fulladdress_count_3	1.00	0.06	1.00	1.00	5.00
150	name_fulladdress_count_30	1.03	0.18	1.00	1.00	11.00
151	name_fulladdress_count_7	1.01	0.09	1.00	1.00	6.00
152	name_fulladdress_day_since	328.85	94.23	0.00	365.00	365.00
153	name_homephone_count_0	1.00	0.02	1.00	1.00	3.00
154	name_homephone_count_0_by_14	13.91	0.78	1.75	14.00	14.00
155	name_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
156	name_homephone_count_0_by_30	29.61	2.40	2.73	30.00	30.00
157	name_homephone_count_0_by_7	6.98	0.28	1.17	7.00	7.00
158	name_homephone_count_1	1.00	0.04	1.00	1.00	3.00
159	name_homephone_count_14	1.01	0.12	1.00	1.00	8.00
160	name_homephone_count_1_by_14	13.92	0.76	1.75	14.00	14.00
161	name_homephone_count_1_by_3	3.00	0.07	0.75	3.00	3.00
162	name_homephone_count_1_by_30	29.63	2.36	2.73	30.00	30.00
163	name_homephone_count_1_by_7	6.98	0.26	1.17	7.00	7.00
164	name_homephone_count_3	1.00	0.06	1.00	1.00	5.00
165	name_homephone_count_30	1.03	0.18	1.00	1.00	11.00
166	name_homephone_count_7	1.01	0.09	1.00	1.00	6.00
167	name_homephone_day_since	330.66	92.21	0.00	365.00	365.00
168	ssn_address_count_0	1.00	0.02	1.00	1.00	3.00
169	ssn_address_count_0_by_14	13.91	0.80	1.75	14.00	14.00
170	ssn_address_count_0_by_3	3.00	0.08	0.75	3.00	3.00
171	ssn_address_count_0_by_30	29.60	2.45	2.73	30.00	30.00
172	ssn_address_count_0_by_7	6.98	0.29	1.17	7.00	7.00
173	ssn_address_count_1	1.00	0.04	1.00	1.00	3.00
174	ssn_address_count_14	1.01	0.12	1.00	1.00	8.00
175	ssn_address_count_1_by_14	13.91	0.77	1.75	14.00	14.00
176	ssn_address_count_1_by_3	3.00	0.07	0.75	3.00	3.00
177	ssn_address_count_1_by_30	29.61	2.40	2.73	30.00	30.00
178	ssn_address_count_1_by_7	6.98	0.26	1.17	7.00	7.00
179	ssn_address_count_3	1.00	0.06	1.00	1.00	5.00
180	ssn_address_count_30	1.03	0.18	1.00	1.00	11.00
181	ssn_address_count_7	1.01	0.09	1.00	1.00	6.00
182	ssn_address_day_since	329.16	93.91	0.00	365.00	365.00
183	ssn_dob_count_0	1.01	0.22	1.00	1.00	21.00
184	ssn_dob_count_0_by_14	13.89	0.89	0.64	14.00	14.00
185	ssn_dob_count_0_by_3	2.99	0.12	0.14	3.00	3.00

186	ssn_dob_count_0_by_30	29.59	2.54	1.36	30.00	30.00
187	ssn_dob_count_0_by_7	6.97	0.35	0.32	7.00	7.00
188	ssn_dob_count_1	1.01	0.38	1.00	1.00	34.00
189	ssn_dob_count_14	1.03	0.47	1.00	1.00	34.00
190	ssn_dob_count_1_by_14	13.91	0.80	0.88	14.00	14.00
191	ssn_dob_count_1_by_3	3.00	0.08	0.43	3.00	3.00
192	ssn_dob_count_1_by_30	29.62	2.40	1.76	30.00	30.00
193	ssn_dob_count_1_by_7	6.98	0.29	0.50	7.00	7.00
194	ssn_dob_count_3	1.02	0.42	1.00	1.00	34.00
195	ssn_dob_count_30	1.05	0.50	1.00	1.00	34.00
196	ssn_dob_count_7	1.03	0.45	1.00	1.00	34.00
197	ssn_dob_day_since	331.28	91.97	0.00	365.00	365.00
198	ssn_dob_homephone_count_0	1.00	0.02	1.00	1.00	3.00
199	ssn_dob_homephone_count_0_by_14	13.92	0.76	1.75	14.00	14.00
200	ssn_dob_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
201	ssn_dob_homephone_count_0_by_30	29.64	2.31	2.73	30.00	30.00
202	ssn_dob_homephone_count_0_by_7	6.98	0.27	1.17	7.00	7.00
203	ssn_dob_homephone_count_1	1.00	0.04	1.00	1.00	3.00
204	ssn_dob_homephone_count_14	1.01	0.12	1.00	1.00	8.00
205	ssn_dob_homephone_count_1_by_14	13.92	0.73	1.75	14.00	14.00
206	ssn_dob_homephone_count_1_by_3	3.00	0.06	0.75	3.00	3.00
207	ssn_dob_homephone_count_1_by_30	29.66	2.27	2.73	30.00	30.00
208	ssn_dob_homephone_count_1_by_7	6.98	0.25	1.17	7.00	7.00
209	ssn_dob_homephone_count_3	1.00	0.06	1.00	1.00	5.00
210	ssn_dob_homephone_count_30	1.03	0.17	1.00	1.00	11.00
211	ssn_dob_homephone_count_7	1.01	0.08	1.00	1.00	6.00
212	ssn_dob_homephone_day_since	333.36	89.05	0.00	365.00	365.00
213	ssn_firstname_count_0	1.01	0.22	1.00	1.00	21.00
214	ssn_firstname_count_0_by_14	13.89	0.93	0.64	14.00	14.00
215	ssn_firstname_count_0_by_3	2.99	0.12	0.14	3.00	3.00
216	ssn_firstname_count_0_by_30	29.55	2.65	1.36	30.00	30.00
217	ssn_firstname_count_0_by_7	6.97	0.36	0.32	7.00	7.00
218	ssn_firstname_count_1	1.01	0.38	1.00	1.00	34.00
219	ssn_firstname_count_14	1.03	0.47	1.00	1.00	34.00
220	ssn_firstname_count_1_by_14	13.90	0.84	0.88	14.00	14.00
221	ssn_firstname_count_1_by_3	3.00	0.08	0.43	3.00	3.00
222	ssn_firstname_count_1_by_30	29.58	2.52	1.76	30.00	30.00
223	ssn_firstname_count_1_by_7	6.97	0.30	0.50	7.00	7.00
224	ssn_firstname_count_3	1.02	0.42	1.00	1.00	34.00
225	ssn_firstname_count_30	1.05	0.50	1.00	1.00	34.00
226	ssn_firstname_count_7	1.03	0.45	1.00	1.00	34.00
227	ssn_firstname_day_since	327.52	96.15	0.00	365.00	365.00
228	ssn_fulladdress_count_0	1.00	0.02	1.00	1.00	3.00
229	ssn_fulladdress_count_0_by_14	13.91	0.80	1.75	14.00	14.00
230	ssn_fulladdress_count_0_by_3	3.00	0.08	0.75	3.00	3.00
231	ssn_fulladdress_count_0_by_30	29.60	2.44	2.73	30.00	30.00
232	ssn_fulladdress_count_0_by_7	6.98	0.29	1.17	7.00	7.00
233	ssn_fulladdress_count_1	1.00	0.04	1.00	1.00	3.00
234	ssn_fulladdress_count_14	1.01	0.12	1.00	1.00	8.00
235	ssn_fulladdress_count_1_by_14	13.92	0.77	1.75	14.00	14.00
236	ssn_fulladdress_count_1_by_3	3.00	0.07	0.75	3.00	3.00
237	ssn_fulladdress_count_1_by_30	29.61	2.40	2.73	30.00	30.00
238	ssn_fulladdress_count_1_by_7	6.98	0.26	1.17	7.00	7.00
239	ssn_fulladdress_count_3	1.00	0.06	1.00	1.00	5.00
240	ssn_fulladdress_count_30	1.03	0.18	1.00	1.00	11.00
241	ssn_fulladdress_count_7	1.01	0.09	1.00	1.00	6.00
242	ssn_fulladdress_day_since	329.25	93.80	0.00	365.00	365.00
243	ssn_fulladdress_dob_count_0	1.00	0.02	1.00	1.00	3.00
244	ssn_fulladdress_dob_count_0_by_14	13.92	0.76	1.75	14.00	14.00
245	ssn_fulladdress_dob_count_0_by_3	3.00	0.08	0.75	3.00	3.00
246	ssn_fulladdress_dob_count_0_by_30	29.64	2.33	2.73	30.00	30.00
247	ssn_fulladdress_dob_count_0_by_7	6.98	0.27	1.17	7.00	7.00
248	ssn_fulladdress_dob_count_1	1.00	0.04	1.00	1.00	3.00

249	ssn_fulladdress_dob_count_14	1.01	0.12	1.00	1.00	8.00
250	ssn_fulladdress_dob_count_1_by_14	13.92	0.73	1.75	14.00	14.00
251	ssn_fulladdress_dob_count_1_by_3	3.00	0.06	0.75	3.00	3.00
252	ssn_fulladdress_dob_count_1_by_30	29.65	2.28	2.73	30.00	30.00
253	ssn_fulladdress_dob_count_1_by_7	6.98	0.25	1.17	7.00	7.00
254	ssn_fulladdress_dob_count_3	1.00	0.06	1.00	1.00	5.00
255	ssn_fulladdress_dob_count_30	1.03	0.17	1.00	1.00	11.00
256	ssn_fulladdress_dob_count_7	1.01	0.08	1.00	1.00	6.00
257	ssn_fulladdress_dob_day_since	332.84	89.65	0.00	365.00	365.00
258	ssn_fulladdress_homephone_count_0	1.00	0.02	1.00	1.00	3.00
259	ssn_fulladdress_homephone_count_0_by_14	13.91	0.78	1.75	14.00	14.00
260	ssn_fulladdress_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
261	ssn_fulladdress_homephone_count_0_by_30	29.62	2.37	2.73	30.00	30.00
262	ssn_fulladdress_homephone_count_0_by_7	6.98	0.28	1.17	7.00	7.00
263	ssn_fulladdress_homephone_count_1	1.00	0.04	1.00	1.00	3.00
264	ssn_fulladdress_homephone_count_14	1.01	0.12	1.00	1.00	8.00
265	ssn_fulladdress_homephone_count_1_by_14	13.92	0.75	1.75	14.00	14.00
266	ssn_fulladdress_homephone_count_1_by_3	3.00	0.06	0.75	3.00	3.00
267	ssn_fulladdress_homephone_count_1_by_30	29.64	2.33	2.73	30.00	30.00
268	ssn_fulladdress_homephone_count_1_by_7	6.98	0.26	1.17	7.00	7.00
269	ssn_fulladdress_homephone_count_3	1.00	0.06	1.00	1.00	5.00
270	ssn_fulladdress_homephone_count_30	1.03	0.18	1.00	1.00	11.00
271	ssn_fulladdress_homephone_count_7	1.01	0.08	1.00	1.00	6.00
272	ssn_fulladdress_homephone_day_since	331.39	91.36	0.00	365.00	365.00
273	ssn_homephone_count_0	1.00	0.02	1.00	1.00	3.00
274	ssn_homephone_count_0_by_14	13.91	0.78	1.75	14.00	14.00
275	ssn_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
276	ssn_homephone_count_0_by_30	29.62	2.39	2.73	30.00	30.00
277	ssn_homephone_count_0_by_7	6.98	0.28	1.17	7.00	7.00
278	ssn_homephone_count_1	1.00	0.04	1.00	1.00	3.00
279	ssn_homephone_count_14	1.01	0.12	1.00	1.00	8.00
280	ssn_homephone_count_1_by_14	13.92	0.75	1.75	14.00	14.00
281	ssn_homephone_count_1_by_3	3.00	0.07	0.75	3.00	3.00
282	ssn_homephone_count_1_by_30	29.63	2.35	2.73	30.00	30.00
283	ssn_homephone_count_1_by_7	6.98	0.26	1.17	7.00	7.00
284	ssn_homephone_count_3	1.00	0.06	1.00	1.00	5.00
285	ssn_homephone_count_30	1.03	0.18	1.00	1.00	11.00
286	ssn_homephone_count_7	1.01	0.09	1.00	1.00	6.00
287	ssn_homephone_day_since	330.85	92.02	0.00	365.00	365.00
288	ssn_homephone_name_dob_count_0	1.00	0.02	1.00	1.00	3.00
289	ssn_homephone_name_dob_count_0_by_14	13.92	0.75	1.75	14.00	14.00
290	ssn_homephone_name_dob_count_0_by_3	3.00	0.08	0.75	3.00	3.00
291	ssn_homephone_name_dob_count_0_by_30	29.64	2.31	2.73	30.00	30.00
292	ssn_homephone_name_dob_count_0_by_7	6.98	0.27	1.17	7.00	7.00
293	ssn_homephone_name_dob_count_1	1.00	0.04	1.00	1.00	3.00
294	ssn_homephone_name_dob_count_14	1.01	0.12	1.00	1.00	8.00
295	ssn_homephone_name_dob_count_1_by_14	13.92	0.73	1.75	14.00	14.00
296	ssn_homephone_name_dob_count_1_by_3	3.00	0.06	0.75	3.00	3.00
297	ssn_homephone_name_dob_count_1_by_30	29.66	2.26	2.73	30.00	30.00

298	ssn_homephone_name_dob_count_1_by_7	6.98	0.25	1.17	7.00	7.00
299	ssn_homephone_name_dob_count_3	1.00	0.06	1.00	1.00	5.00
300	ssn_homephone_name_dob_count_30	1.03	0.17	1.00	1.00	11.00
301	ssn_homephone_name_dob_count_7	1.01	0.08	1.00	1.00	6.00
302	ssn_homephone_name_dob_day_since	333.48	88.89	0.00	365.00	365.00
303	ssn_lastname_count_0	1.01	0.22	1.00	1.00	21.00
304	ssn_lastname_count_0_by_14	13.89	0.92	0.64	14.00	14.00
305	ssn_lastname_count_0_by_3	2.99	0.12	0.14	3.00	3.00
306	ssn_lastname_count_0_by_30	29.55	2.65	1.36	30.00	30.00
307	ssn_lastname_count_0_by_7	6.97	0.36	0.32	7.00	7.00
308	ssn_lastname_count_1	1.01	0.38	1.00	1.00	34.00
309	ssn_lastname_count_14	1.03	0.47	1.00	1.00	34.00
310	ssn_lastname_count_1_by_14	13.90	0.84	0.88	14.00	14.00
311	ssn_lastname_count_1_by_3	3.00	0.08	0.43	3.00	3.00
312	ssn_lastname_count_1_by_30	29.58	2.52	1.76	30.00	30.00
313	ssn_lastname_count_1_by_7	6.97	0.30	0.50	7.00	7.00
314	ssn_lastname_count_3	1.02	0.42	1.00	1.00	34.00
315	ssn_lastname_count_30	1.05	0.50	1.00	1.00	34.00
316	ssn_lastname_count_7	1.03	0.45	1.00	1.00	34.00
317	ssn_lastname_day_since	327.53	96.14	0.00	365.00	365.00
318	ssn_name_dob_count_0	1.01	0.22	1.00	1.00	21.00
319	ssn_name_dob_count_0_by_14	13.90	0.89	0.64	14.00	14.00
320	ssn_name_dob_count_0_by_3	2.99	0.12	0.14	3.00	3.00
321	ssn_name_dob_count_0_by_30	29.59	2.53	1.36	30.00	30.00
322	ssn_name_dob_count_0_by_7	6.97	0.35	0.32	7.00	7.00
323	ssn_name_dob_count_1	1.01	0.38	1.00	1.00	34.00
324	ssn_name_dob_count_14	1.03	0.47	1.00	1.00	34.00
325	ssn_name_dob_count_1_by_14	13.91	0.80	0.88	14.00	14.00
326	ssn_name_dob_count_1_by_3	3.00	0.08	0.43	3.00	3.00
327	ssn_name_dob_count_1_by_30	29.62	2.40	1.76	30.00	30.00
328	ssn_name_dob_count_1_by_7	6.98	0.29	0.50	7.00	7.00
329	ssn_name_dob_count_3	1.02	0.42	1.00	1.00	34.00
330	ssn_name_dob_count_30	1.05	0.50	1.00	1.00	34.00
331	ssn_name_dob_count_7	1.03	0.45	1.00	1.00	34.00
332	ssn_name_dob_day_since	331.43	91.79	0.00	365.00	365.00
333	ssn_name_fulladdress_count_0	1.00	0.02	1.00	1.00	3.00
334	ssn_name_fulladdress_count_0_by_14	13.91	0.80	1.75	14.00	14.00
335	ssn_name_fulladdress_count_0_by_3	3.00	0.08	0.75	3.00	3.00
336	ssn_name_fulladdress_count_0_by_30	29.60	2.44	2.73	30.00	30.00
337	ssn_name_fulladdress_count_0_by_7	6.98	0.29	1.17	7.00	7.00
338	ssn_name_fulladdress_count_1	1.00	0.04	1.00	1.00	3.00
339	ssn_name_fulladdress_count_14	1.01	0.12	1.00	1.00	8.00
340	ssn_name_fulladdress_count_1_by_14	13.92	0.77	1.75	14.00	14.00
341	ssn_name_fulladdress_count_1_by_3	3.00	0.07	0.75	3.00	3.00
342	ssn_name_fulladdress_count_1_by_30	29.61	2.40	2.73	30.00	30.00
343	ssn_name_fulladdress_count_1_by_7	6.98	0.26	1.17	7.00	7.00
344	ssn_name_fulladdress_count_3	1.00	0.06	1.00	1.00	5.00
345	ssn_name_fulladdress_count_30	1.03	0.18	1.00	1.00	11.00
346	ssn_name_fulladdress_count_7	1.01	0.09	1.00	1.00	6.00
347	ssn_name_fulladdress_day_since	329.32	93.72	0.00	365.00	365.00
348	ssn_name_homephone_count_0	1.00	0.02	1.00	1.00	3.00
349	ssn_name_homephone_count_0_by_14	13.91	0.78	1.75	14.00	14.00
350	ssn_name_homephone_count_0_by_3	3.00	0.08	0.75	3.00	3.00
351	ssn_name_homephone_count_0_by_30	29.62	2.38	2.73	30.00	30.00
352	ssn_name_homephone_count_0_by_7	6.98	0.28	1.17	7.00	7.00
353	ssn_name_homephone_count_1	1.00	0.04	1.00	1.00	3.00
354	ssn_name_homephone_count_14	1.01	0.12	1.00	1.00	8.00
355	ssn_name_homephone_count_1_by_14	13.92	0.75	1.75	14.00	14.00
356	ssn_name_homephone_count_1_by_3	3.00	0.07	0.75	3.00	3.00
357	ssn_name_homephone_count_1_by_30	29.63	2.34	2.73	30.00	30.00
358	ssn_name_homephone_count_1_by_7	6.98	0.26	1.17	7.00	7.00
359	ssn_name_homephone_count_3	1.00	0.06	1.00	1.00	5.00
360	ssn_name_homephone_count_30	1.03	0.18	1.00	1.00	11.00

361	ssn_name_homephone_count_7	1.01	0.09	1.00	1.00	6.00
362	ssn_name_homephone_day_since	331.06	91.75	0.00	365.00	365.00
363	ssn_zip5_count_0	1.00	0.02	1.00	1.00	3.00
364	ssn_zip5_count_0_by_14	13.91	0.80	1.75	14.00	14.00
365	ssn_zip5_count_0_by_3	3.00	0.08	0.75	3.00	3.00
366	ssn_zip5_count_0_by_30	29.60	2.45	2.73	30.00	30.00
367	ssn_zip5_count_0_by_7	6.98	0.29	1.17	7.00	7.00
368	ssn_zip5_count_1	1.00	0.04	1.00	1.00	3.00
369	ssn_zip5_count_14	1.01	0.12	1.00	1.00	8.00
370	ssn_zip5_count_1_by_14	13.91	0.77	1.75	14.00	14.00
371	ssn_zip5_count_1_by_3	3.00	0.07	0.75	3.00	3.00
372	ssn_zip5_count_1_by_30	29.61	2.41	2.73	30.00	30.00
373	ssn_zip5_count_1_by_7	6.98	0.27	1.17	7.00	7.00
374	ssn_zip5_count_3	1.00	0.06	1.00	1.00	5.00
375	ssn_zip5_count_30	1.03	0.18	1.00	1.00	11.00
376	ssn_zip5_count_7	1.01	0.09	1.00	1.00	6.00
377	ssn_zip5_day_since	329.05	94.02	0.00	365.00	365.00
378	unique_num_homephone_name_dob_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	3.00
379	unique_num_homephone_name_dob_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
380	unique_num_homephone_name_dob_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	4.00
381	unique_num_homephone_name_dob_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
382	unique_num_homephone_name_dob_for_each_name_homephone_14	0.01	0.11	0.00	0.00	3.00
383	unique_num_homephone_name_dob_for_each_name_homephone_3	0.00	0.06	0.00	0.00	2.00
384	unique_num_homephone_name_dob_for_each_name_homephone_30	0.03	0.16	0.00	0.00	4.00
385	unique_num_homephone_name_dob_for_each_name_homephone_7	0.01	0.08	0.00	0.00	2.00
386	unique_num_homephone_name_dob_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
387	unique_num_homephone_name_dob_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
388	unique_num_homephone_name_dob_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
389	unique_num_homephone_name_dob_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
390	unique_num_homephone_name_dob_for_each_ssn_address_14	0.01	0.12	0.00	0.00	3.00
391	unique_num_homephone_name_dob_for_each_ssn_address_3	0.00	0.06	0.00	0.00	2.00
392	unique_num_homephone_name_dob_for_each_ssn_address_30	0.03	0.16	0.00	0.00	4.00
393	unique_num_homephone_name_dob_for_each_ssn_address_7	0.01	0.08	0.00	0.00	2.00
394	unique_num_homephone_name_dob_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
395	unique_num_homephone_name_dob_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
396	unique_num_homephone_name_dob_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
397	unique_num_homephone_name_dob_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
398	unique_num_homephone_name_dob_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	3.00
399	unique_num_homephone_name_dob_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	2.00
400	unique_num_homephone_name_dob_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	4.00
401	unique_num_homephone_name_dob_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	2.00

402	unique_num_homephone_name_dob_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	3.00
403	unique_num_homephone_name_dob_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	2.00
404	unique_num_homephone_name_dob_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	4.00
405	unique_num_homephone_name_dob_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	2.00
406	unique_num_homephone_name_dob_for_each_zip5_14	4.61	4.06	0.00	4.00	47.00
407	unique_num_homephone_name_dob_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
408	unique_num_homephone_name_dob_for_each_zip5_30	9.37	7.69	0.00	8.00	86.00
409	unique_num_homephone_name_dob_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
410	unique_num_name_fulladdress_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
411	unique_num_name_fulladdress_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
412	unique_num_name_fulladdress_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	2.00
413	unique_num_name_fulladdress_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
414	unique_num_name_fulladdress_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
415	unique_num_name_fulladdress_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
416	unique_num_name_fulladdress_for_each_name_homephone_30	0.03	0.16	0.00	0.00	2.00
417	unique_num_name_fulladdress_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
418	unique_num_name_fulladdress_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
419	unique_num_name_fulladdress_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
420	unique_num_name_fulladdress_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
421	unique_num_name_fulladdress_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
422	unique_num_name_fulladdress_for_each_ssn_address_14	0.01	0.12	0.00	0.00	2.00
423	unique_num_name_fulladdress_for_each_ssn_address_3	0.00	0.06	0.00	0.00	2.00
424	unique_num_name_fulladdress_for_each_ssn_address_30	0.03	0.16	0.00	0.00	3.00
425	unique_num_name_fulladdress_for_each_ssn_address_7	0.01	0.08	0.00	0.00	2.00
426	unique_num_name_fulladdress_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
427	unique_num_name_fulladdress_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
428	unique_num_name_fulladdress_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
429	unique_num_name_fulladdress_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
430	unique_num_name_fulladdress_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	2.00
431	unique_num_name_fulladdress_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	2.00
432	unique_num_name_fulladdress_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	3.00
433	unique_num_name_fulladdress_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	2.00

434	unique_num_name_fulladdress_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	2.00
435	unique_num_name_fulladdress_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	2.00
436	unique_num_name_fulladdress_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	3.00
437	unique_num_name_fulladdress_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	2.00
438	unique_num_name_fulladdress_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
439	unique_num_name_fulladdress_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
440	unique_num_name_fulladdress_for_each_zip5_30	9.35	7.68	0.00	8.00	86.00
441	unique_num_name_fulladdress_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
442	unique_num_name_homephone_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	1.00
443	unique_num_name_homephone_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
444	unique_num_name_homephone_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	1.00
445	unique_num_name_homephone_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
446	unique_num_name_homephone_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	2.00
447	unique_num_name_homephone_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
448	unique_num_name_homephone_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	3.00
449	unique_num_name_homephone_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
450	unique_num_name_homephone_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
451	unique_num_name_homephone_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
452	unique_num_name_homephone_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
453	unique_num_name_homephone_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
454	unique_num_name_homephone_for_each_ssn_address_14	0.01	0.12	0.00	0.00	2.00
455	unique_num_name_homephone_for_each_ssn_address_3	0.00	0.06	0.00	0.00	2.00
456	unique_num_name_homephone_for_each_ssn_address_30	0.03	0.16	0.00	0.00	3.00
457	unique_num_name_homephone_for_each_ssn_address_7	0.01	0.08	0.00	0.00	2.00
458	unique_num_name_homephone_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
459	unique_num_name_homephone_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
460	unique_num_name_homephone_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
461	unique_num_name_homephone_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
462	unique_num_name_homephone_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	2.00
463	unique_num_name_homephone_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	2.00
464	unique_num_name_homephone_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	3.00
465	unique_num_name_homephone_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	2.00

466	unique_num_name_homephone_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	2.00
467	unique_num_name_homephone_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	2.00
468	unique_num_name_homephone_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	3.00
469	unique_num_name_homephone_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	2.00
470	unique_num_name_homephone_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
471	unique_num_name_homephone_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
472	unique_num_name_homephone_for_each_zip5_30	9.36	7.68	0.00	8.00	86.00
473	unique_num_name_homephone_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
474	unique_num_ssn_address_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
475	unique_num_ssn_address_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
476	unique_num_ssn_address_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	3.00
477	unique_num_ssn_address_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
478	unique_num_ssn_address_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	2.00
479	unique_num_ssn_address_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
480	unique_num_ssn_address_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	3.00
481	unique_num_ssn_address_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
482	unique_num_ssn_address_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
483	unique_num_ssn_address_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
484	unique_num_ssn_address_for_each_name_homephone_30	0.03	0.16	0.00	0.00	3.00
485	unique_num_ssn_address_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
486	unique_num_ssn_address_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
487	unique_num_ssn_address_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
488	unique_num_ssn_address_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
489	unique_num_ssn_address_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
490	unique_num_ssn_address_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
491	unique_num_ssn_address_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
492	unique_num_ssn_address_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
493	unique_num_ssn_address_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
494	unique_num_ssn_address_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	1.00
495	unique_num_ssn_address_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	1.00
496	unique_num_ssn_address_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	1.00
497	unique_num_ssn_address_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	1.00

498	unique_num_ssn_address_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	1.00
499	unique_num_ssn_address_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	1.00
500	unique_num_ssn_address_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	2.00
501	unique_num_ssn_address_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	1.00
502	unique_num_ssn_address_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
503	unique_num_ssn_address_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
504	unique_num_ssn_address_for_each_zip5_30	9.36	7.68	0.00	8.00	86.00
505	unique_num_ssn_address_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
506	unique_num_ssn_dob_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
507	unique_num_ssn_dob_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
508	unique_num_ssn_dob_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	3.00
509	unique_num_ssn_dob_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
510	unique_num_ssn_dob_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	3.00
511	unique_num_ssn_dob_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
512	unique_num_ssn_dob_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	4.00
513	unique_num_ssn_dob_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
514	unique_num_ssn_dob_for_each_name_homephone_14	0.01	0.11	0.00	0.00	3.00
515	unique_num_ssn_dob_for_each_name_homephone_3	0.00	0.06	0.00	0.00	2.00
516	unique_num_ssn_dob_for_each_name_homephone_30	0.03	0.16	0.00	0.00	4.00
517	unique_num_ssn_dob_for_each_name_homephone_7	0.01	0.08	0.00	0.00	2.00
518	unique_num_ssn_dob_for_each_ssn_14	0.02	0.14	0.00	0.00	10.00
519	unique_num_ssn_dob_for_each_ssn_3	0.01	0.08	0.00	0.00	4.00
520	unique_num_ssn_dob_for_each_ssn_30	0.03	0.20	0.00	0.00	14.00
521	unique_num_ssn_dob_for_each_ssn_7	0.01	0.11	0.00	0.00	7.00
522	unique_num_ssn_dob_for_each_ssn_address_14	0.01	0.12	0.00	0.00	3.00
523	unique_num_ssn_dob_for_each_ssn_address_3	0.00	0.06	0.00	0.00	2.00
524	unique_num_ssn_dob_for_each_ssn_address_30	0.03	0.16	0.00	0.00	4.00
525	unique_num_ssn_dob_for_each_ssn_address_7	0.01	0.08	0.00	0.00	2.00
526	unique_num_ssn_dob_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	3.00
527	unique_num_ssn_dob_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	2.00
528	unique_num_ssn_dob_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	4.00
529	unique_num_ssn_dob_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	2.00
530	unique_num_ssn_dob_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	3.00
531	unique_num_ssn_dob_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	2.00

532	unique_num_ssn_dob_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	4.00
533	unique_num_ssn_dob_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	2.00
534	unique_num_ssn_dob_for_each_zip5_14	4.61	4.06	0.00	4.00	47.00
535	unique_num_ssn_dob_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
536	unique_num_ssn_dob_for_each_zip5_30	9.37	7.69	0.00	8.00	86.00
537	unique_num_ssn_dob_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
538	unique_num_ssn_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
539	unique_num_ssn_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
540	unique_num_ssn_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	3.00
541	unique_num_ssn_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
542	unique_num_ssn_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	2.00
543	unique_num_ssn_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
544	unique_num_ssn_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	3.00
545	unique_num_ssn_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
546	unique_num_ssn_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
547	unique_num_ssn_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
548	unique_num_ssn_for_each_name_homephone_30	0.03	0.16	0.00	0.00	3.00
549	unique_num_ssn_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
550	unique_num_ssn_for_each_ssn_address_14	0.01	0.12	0.00	0.00	1.00
551	unique_num_ssn_for_each_ssn_address_3	0.00	0.06	0.00	0.00	1.00
552	unique_num_ssn_for_each_ssn_address_30	0.03	0.16	0.00	0.00	1.00
553	unique_num_ssn_for_each_ssn_address_7	0.01	0.08	0.00	0.00	1.00
554	unique_num_ssn_for_each_ssn_dob_14	0.02	0.12	0.00	0.00	1.00
555	unique_num_ssn_for_each_ssn_dob_3	0.01	0.08	0.00	0.00	1.00
556	unique_num_ssn_for_each_ssn_dob_30	0.03	0.16	0.00	0.00	1.00
557	unique_num_ssn_for_each_ssn_dob_7	0.01	0.10	0.00	0.00	1.00
558	unique_num_ssn_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	1.00
559	unique_num_ssn_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	1.00
560	unique_num_ssn_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	1.00
561	unique_num_ssn_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	1.00
562	unique_num_ssn_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	1.00
563	unique_num_ssn_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	1.00
564	unique_num_ssn_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	1.00
565	unique_num_ssn_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	1.00
566	unique_num_ssn_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
567	unique_num_ssn_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
568	unique_num_ssn_for_each_zip5_30	9.35	7.68	0.00	8.00	86.00
569	unique_num_ssn_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
570	unique_num_ssn_fulladdress_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
571	unique_num_ssn_fulladdress_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00

572	unique_num_ssn_fulladdress_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	3.00
573	unique_num_ssn_fulladdress_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
574	unique_num_ssn_fulladdress_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	2.00
575	unique_num_ssn_fulladdress_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
576	unique_num_ssn_fulladdress_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	3.00
577	unique_num_ssn_fulladdress_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
578	unique_num_ssn_fulladdress_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
579	unique_num_ssn_fulladdress_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
580	unique_num_ssn_fulladdress_for_each_name_homephone_30	0.03	0.16	0.00	0.00	3.00
581	unique_num_ssn_fulladdress_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
582	unique_num_ssn_fulladdress_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
583	unique_num_ssn_fulladdress_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
584	unique_num_ssn_fulladdress_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
585	unique_num_ssn_fulladdress_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
586	unique_num_ssn_fulladdress_for_each_ssn_address_14	0.01	0.12	0.00	0.00	2.00
587	unique_num_ssn_fulladdress_for_each_ssn_address_3	0.00	0.06	0.00	0.00	1.00
588	unique_num_ssn_fulladdress_for_each_ssn_address_30	0.03	0.16	0.00	0.00	2.00
589	unique_num_ssn_fulladdress_for_each_ssn_address_7	0.01	0.08	0.00	0.00	1.00
590	unique_num_ssn_fulladdress_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
591	unique_num_ssn_fulladdress_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
592	unique_num_ssn_fulladdress_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
593	unique_num_ssn_fulladdress_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
594	unique_num_ssn_fulladdress_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	1.00
595	unique_num_ssn_fulladdress_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	1.00
596	unique_num_ssn_fulladdress_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	2.00
597	unique_num_ssn_fulladdress_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	1.00
598	unique_num_ssn_fulladdress_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
599	unique_num_ssn_fulladdress_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
600	unique_num_ssn_fulladdress_for_each_zip5_30	9.36	7.68	0.00	8.00	86.00
601	unique_num_ssn_fulladdress_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
602	unique_num_ssn_zip5_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
603	unique_num_ssn_zip5_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00

604	unique_num_ssn_zip5_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	3.00
605	unique_num_ssn_zip5_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
606	unique_num_ssn_zip5_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	2.00
607	unique_num_ssn_zip5_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	2.00
608	unique_num_ssn_zip5_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	3.00
609	unique_num_ssn_zip5_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	2.00
610	unique_num_ssn_zip5_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
611	unique_num_ssn_zip5_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
612	unique_num_ssn_zip5_for_each_name_homephone_30	0.03	0.16	0.00	0.00	3.00
613	unique_num_ssn_zip5_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
614	unique_num_ssn_zip5_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
615	unique_num_ssn_zip5_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
616	unique_num_ssn_zip5_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
617	unique_num_ssn_zip5_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
618	unique_num_ssn_zip5_for_each_ssn_address_14	0.01	0.12	0.00	0.00	2.00
619	unique_num_ssn_zip5_for_each_ssn_address_3	0.00	0.06	0.00	0.00	1.00
620	unique_num_ssn_zip5_for_each_ssn_address_30	0.03	0.16	0.00	0.00	2.00
621	unique_num_ssn_zip5_for_each_ssn_address_7	0.01	0.08	0.00	0.00	1.00
622	unique_num_ssn_zip5_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
623	unique_num_ssn_zip5_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
624	unique_num_ssn_zip5_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
625	unique_num_ssn_zip5_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
626	unique_num_ssn_zip5_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	1.00
627	unique_num_ssn_zip5_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	1.00
628	unique_num_ssn_zip5_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	1.00
629	unique_num_ssn_zip5_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	1.00
630	unique_num_ssn_zip5_for_each_zip5_14	4.60	4.06	0.00	4.00	47.00
631	unique_num_ssn_zip5_for_each_zip5_3	1.16	1.49	0.00	1.00	33.00
632	unique_num_ssn_zip5_for_each_zip5_30	9.35	7.68	0.00	8.00	86.00
633	unique_num_ssn_zip5_for_each_zip5_7	2.43	2.44	0.00	2.00	40.00
634	unique_num_zip5_for_each_homephone_name_dob_14	0.01	0.11	0.00	0.00	2.00
635	unique_num_zip5_for_each_homephone_name_dob_3	0.00	0.06	0.00	0.00	1.00
636	unique_num_zip5_for_each_homephone_name_dob_30	0.02	0.15	0.00	0.00	2.00
637	unique_num_zip5_for_each_homephone_name_dob_7	0.01	0.08	0.00	0.00	1.00
638	unique_num_zip5_for_each_name_fulladdress_14	0.01	0.12	0.00	0.00	1.00
639	unique_num_zip5_for_each_name_fulladdress_3	0.00	0.06	0.00	0.00	1.00

640	unique_num_zip5_for_each_name_fulladdress_30	0.03	0.16	0.00	0.00	1.00
641	unique_num_zip5_for_each_name_fulladdress_7	0.01	0.08	0.00	0.00	1.00
642	unique_num_zip5_for_each_name_homephone_14	0.01	0.11	0.00	0.00	2.00
643	unique_num_zip5_for_each_name_homephone_3	0.00	0.06	0.00	0.00	1.00
644	unique_num_zip5_for_each_name_homephone_30	0.03	0.16	0.00	0.00	2.00
645	unique_num_zip5_for_each_name_homephone_7	0.01	0.08	0.00	0.00	1.00
646	unique_num_zip5_for_each_ssn_14	0.03	0.48	0.00	0.00	33.00
647	unique_num_zip5_for_each_ssn_3	0.02	0.42	0.00	0.00	33.00
648	unique_num_zip5_for_each_ssn_30	0.05	0.51	0.00	0.00	33.00
649	unique_num_zip5_for_each_ssn_7	0.03	0.45	0.00	0.00	33.00
650	unique_num_zip5_for_each_ssn_address_14	0.01	0.12	0.00	0.00	2.00
651	unique_num_zip5_for_each_ssn_address_3	0.00	0.06	0.00	0.00	1.00
652	unique_num_zip5_for_each_ssn_address_30	0.03	0.16	0.00	0.00	2.00
653	unique_num_zip5_for_each_ssn_address_7	0.01	0.08	0.00	0.00	1.00
654	unique_num_zip5_for_each_ssn_dob_14	0.03	0.47	0.00	0.00	33.00
655	unique_num_zip5_for_each_ssn_dob_3	0.02	0.42	0.00	0.00	33.00
656	unique_num_zip5_for_each_ssn_dob_30	0.04	0.49	0.00	0.00	33.00
657	unique_num_zip5_for_each_ssn_dob_7	0.02	0.45	0.00	0.00	33.00
658	unique_num_zip5_for_each_ssn_fulladdress_14	0.01	0.12	0.00	0.00	1.00
659	unique_num_zip5_for_each_ssn_fulladdress_3	0.00	0.06	0.00	0.00	1.00
660	unique_num_zip5_for_each_ssn_fulladdress_30	0.03	0.16	0.00	0.00	1.00
661	unique_num_zip5_for_each_ssn_fulladdress_7	0.01	0.08	0.00	0.00	1.00
662	unique_num_zip5_for_each_ssn_zip5_14	0.01	0.12	0.00	0.00	1.00
663	unique_num_zip5_for_each_ssn_zip5_3	0.00	0.06	0.00	0.00	1.00
664	unique_num_zip5_for_each_ssn_zip5_30	0.03	0.16	0.00	0.00	1.00
665	unique_num_zip5_for_each_ssn_zip5_7	0.01	0.08	0.00	0.00	1.00
666	zip5_count_0	1.18	0.55	1.00	1.00	27.00
667	zip5_count_0_by_14	4.85	3.95	0.29	3.50	14.00
668	zip5_count_0_by_3	2.08	0.90	0.09	2.00	3.00
669	zip5_count_0_by_30	6.59	7.27	0.35	3.75	30.00
670	zip5_count_0_by_7	3.50	2.16	0.18	2.80	7.00
671	zip5_count_1	1.51	0.97	1.00	1.00	33.00
672	zip5_count_14	5.64	4.09	1.00	5.00	48.00
673	zip5_count_1_by_14	5.54	3.92	0.31	4.67	14.00
674	zip5_count_1_by_3	2.39	0.76	0.09	3.00	3.00
675	zip5_count_1_by_30	7.48	7.35	0.38	5.00	30.00
676	zip5_count_1_by_7	4.01	2.07	0.19	3.50	7.00
677	zip5_count_3	2.16	1.49	1.00	2.00	34.00
678	zip5_count_30	10.49	7.79	1.00	9.00	87.00
679	zip5_count_7	3.44	2.45	1.00	3.00	41.00
680	zip5_day_since	16.27	59.62	0.00	3.00	365.00

9.3 Full Final Results of Model Selected

Table 9.4 Training Population Results

Training	# Records		# Goods		# Bads		Fraud Rate					
	638172		628951		9221		0.014449083					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	6382	1645	4737	25.78%	74.22%	6382	1645	4737	0.26%	51.37%	51.11	0.35
2	6382	6213	169	97.35%	2.65%	12764	7858	4906	1.25%	53.20%	51.96	1.60
3	6382	6315	67	98.95%	1.05%	19146	14173	4973	2.25%	53.93%	51.68	2.85
4	6382	6332	50	99.22%	0.78%	25528	20505	5023	3.26%	54.47%	51.21	4.08
5	6382	6337	45	99.29%	0.71%	31910	26842	5068	4.27%	54.96%	50.69	5.30
6	6382	6327	55	99.14%	0.86%	38292	33169	5123	5.27%	55.56%	50.28	6.47
7	6382	6330	52	99.19%	0.81%	44674	39499	5175	6.28%	56.12%	49.84	7.63
8	6382	6341	41	99.36%	0.64%	51056	45840	5216	7.29%	56.57%	49.28	8.79
9	6382	6330	52	99.19%	0.81%	57438	52170	5268	8.29%	57.13%	48.84	9.90
10	6382	6335	47	99.26%	0.74%	63820	58505	5315	9.30%	57.64%	48.34	11.01
11	6382	6335	47	99.26%	0.74%	70202	64840	5362	10.31%	58.15%	47.84	12.09
12	6382	6331	51	99.20%	0.80%	76584	71171	5413	11.32%	58.70%	47.39	13.15
13	6382	6337	45	99.29%	0.71%	82966	77508	5458	12.32%	59.19%	46.87	14.20
14	6382	6338	44	99.31%	0.69%	89348	83846	5502	13.33%	59.67%	46.34	15.24
15	6382	6334	48	99.25%	0.75%	95730	90180	5550	14.34%	60.19%	45.85	16.25
16	6382	6347	35	99.45%	0.55%	102112	96527	5585	15.35%	60.57%	45.22	17.28
17	6382	6336	46	99.28%	0.72%	108494	102863	5631	16.35%	61.07%	44.71	18.27
18	6382	6345	37	99.42%	0.58%	114876	109208	5668	17.36%	61.47%	44.10	19.27
19	6382	6332	50	99.22%	0.78%	121258	115540	5718	18.37%	62.01%	43.64	20.21
20	6382	6331	51	99.20%	0.80%	127640	121871	5769	19.38%	62.56%	43.19	21.13
21	6382	6342	40	99.37%	0.63%	134022	128213	5809	20.39%	63.00%	42.61	22.07
22	6382	6337	45	99.29%	0.71%	140404	134550	5854	21.39%	63.49%	42.09	22.98
23	6382	6333	49	99.23%	0.77%	146786	140883	5903	22.40%	64.02%	41.62	23.87
24	6382	6346	36	99.44%	0.56%	153168	147229	5939	23.41%	64.41%	41.00	24.79
25	6382	6338	44	99.31%	0.69%	159550	153567	5983	24.42%	64.88%	40.47	25.67
26	6382	6333	49	99.23%	0.77%	165932	159900	6032	25.42%	65.42%	39.99	26.51
27	6382	6336	46	99.28%	0.72%	172314	166236	6078	26.43%	65.91%	39.48	27.35
28	6382	6342	40	99.37%	0.63%	178696	172578	6118	27.44%	66.35%	38.91	28.21
29	6382	6340	42	99.34%	0.66%	185078	178918	6160	28.45%	66.80%	38.36	29.05
30	6382	6324	58	99.09%	0.91%	191460	185242	6218	29.45%	67.43%	37.98	29.79
31	6382	6347	35	99.45%	0.55%	197842	191589	6253	30.46%	67.81%	37.35	30.64
32	6382	6340	42	99.34%	0.66%	204224	197929	6295	31.47%	68.27%	36.80	31.44
33	6382	6334	48	99.25%	0.75%	210606	204263	6343	32.48%	68.79%	36.31	32.20
34	6382	6325	57	99.11%	0.89%	216988	210588	6400	33.48%	69.41%	35.92	32.90
35	6382	6346	36	99.44%	0.56%	223370	216934	6436	34.49%	69.80%	35.31	33.71
36	6382	6329	53	99.17%	0.83%	229752	223263	6489	35.50%	70.37%	34.87	34.41
37	6382	6347	35	99.45%	0.55%	236134	229610	6524	36.51%	70.75%	34.24	35.19
38	6382	6326	56	99.12%	0.88%	242516	235936	6580	37.51%	71.36%	33.85	35.86
39	6382	6335	47	99.26%	0.74%	248898	242271	6627	38.52%	71.87%	33.35	36.56
40	6382	6340	42	99.34%	0.66%	255280	248611	6669	39.53%	72.32%	32.80	37.28
41	6382	6336	46	99.28%	0.72%	261662	254947	6715	40.54%	72.82%	32.29	37.97
42	6382	6333	49	99.23%	0.77%	268044	261280	6764	41.54%	73.35%	31.81	38.63
43	6382	6332	50	99.22%	0.78%	274426	267612	6814	42.55%	73.90%	31.35	39.27
44	6382	6346	36	99.44%	0.56%	280808	273958	6850	43.56%	74.29%	30.73	39.99
45	6382	6325	57	99.11%	0.89%	287190	280283	6907	44.56%	74.91%	30.34	40.58
46	6382	6340	42	99.34%	0.66%	293572	286623	6949	45.57%	75.36%	29.79	41.25
47	6382	6342	40	99.37%	0.63%	299954	292965	6989	46.58%	75.79%	29.21	41.92
48	6382	6343	39	99.39%	0.61%	306336	299308	7028	47.59%	76.22%	28.63	42.59
49	6382	6345	37	99.42%	0.58%	312718	305653	7065	48.60%	76.62%	28.02	43.26
50	6382	6338	44	99.31%	0.69%	319100	311991	7109	49.60%	77.10%	27.49	43.89
51	6382	6338	44	99.31%	0.69%	325482	318329	7153	50.61%	77.57%	26.96	44.50
52	6382	6330	52	99.19%	0.81%	331864	324659	7205	51.62%	78.14%	26.52	45.06
53	6382	6338	44	99.31%	0.69%	338246	330997	7249	52.63%	78.61%	25.99	45.66
54	6382	6339	43	99.33%	0.67%	344628	337336	7292	53.63%	79.08%	25.45	46.26
55	6382	6332	50	99.22%	0.78%	351010	343668	7342	54.64%	79.62%	24.98	46.81
56	6382	6345	37	99.42%	0.58%	357392	350013	7379	55.65%	80.02%	24.37	47.43
57	6382	6351	31	99.51%	0.49%	363774	356364	7410	56.66%	80.36%	23.70	48.09
58	6382	6340	42	99.34%	0.66%	370156	362704	7452	57.67%	80.82%	23.15	48.67
59	6382	6328	54	99.15%	0.85%	376538	369032	7506	58.67%	81.40%	22.73	49.16
60	6382	6331	51	99.20%	0.80%	382920	375363	7557	59.68%	81.95%	22.27	49.67

Training	# Records		# Goods		# Bads		Fraud Rate						
	638172		628951		9221		0.014449083						
	Bins Statistics					Cumulative Statistics							
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	6382	6346	36	99.44%	0.56%	389302	381709	7593	60.69%	82.34%	21.65	50.27
62	6382	6341	41	99.36%	0.64%	395684	388050	7634	61.70%	82.79%	21.09	50.83
63	6382	6329	53	99.17%	0.83%	402066	394379	7687	62.70%	83.36%	20.66	51.30
64	6382	6339	43	99.33%	0.67%	408448	400718	7730	63.71%	83.83%	20.12	51.84
65	6382	6339	43	99.33%	0.67%	414830	407057	7773	64.72%	84.30%	19.58	52.37
66	6382	6341	41	99.36%	0.64%	421212	413398	7814	65.73%	84.74%	19.01	52.90
67	6382	6339	43	99.33%	0.67%	427594	419737	7857	66.74%	85.21%	18.47	53.42
68	6382	6333	49	99.23%	0.77%	433976	426070	7906	67.74%	85.74%	18.00	53.89
69	6382	6347	35	99.45%	0.55%	440358	432417	7941	68.75%	86.12%	17.37	54.45
70	6382	6331	51	99.20%	0.80%	446740	438748	7992	69.76%	86.67%	16.91	54.90
71	6382	6332	50	99.22%	0.78%	453122	445080	8042	70.77%	87.21%	16.45	55.34
72	6382	6331	51	99.20%	0.80%	459504	451411	8093	71.77%	87.77%	16.00	55.78
73	6382	6335	47	99.26%	0.74%	465886	457746	8140	72.78%	88.28%	15.50	56.23
74	6382	6335	47	99.26%	0.74%	472268	464081	8187	73.79%	88.79%	15.00	56.69
75	6382	6336	46	99.28%	0.72%	478650	470417	8233	74.79%	89.29%	14.49	57.14
76	6382	6323	59	99.08%	0.92%	485032	476740	8292	75.80%	89.93%	14.13	57.49
77	6382	6325	57	99.11%	0.89%	491414	483065	8349	76.80%	90.54%	13.74	57.86
78	6382	6317	65	98.98%	1.02%	497796	489382	8414	77.81%	91.25%	13.44	58.16
79	6382	6332	50	99.22%	0.78%	504178	495714	8464	78.82%	91.79%	12.97	58.57
80	6382	6344	38	99.40%	0.60%	510560	502058	8502	79.82%	92.20%	12.38	59.05
81	6382	6345	37	99.42%	0.58%	516942	508403	8539	80.83%	92.60%	11.77	59.54
82	6382	6333	49	99.23%	0.77%	523324	514736	8588	81.84%	93.14%	11.29	59.94
83	6382	6337	45	99.29%	0.71%	529706	521073	8633	82.85%	93.62%	10.78	60.36
84	6382	6331	51	99.20%	0.80%	536088	527404	8684	83.85%	94.18%	10.32	60.73
85	6382	6333	49	99.23%	0.77%	542470	533737	8733	84.86%	94.71%	9.85	61.12
86	6382	6347	35	99.45%	0.55%	548852	540084	8768	85.87%	95.09%	9.22	61.60
87	6382	6338	44	99.31%	0.69%	555234	546422	8812	86.88%	95.56%	8.69	62.01
88	6382	6337	45	99.29%	0.71%	561616	552759	8857	87.89%	96.05%	8.17	62.41
89	6382	6326	56	99.12%	0.88%	567998	559085	8913	88.89%	96.66%	7.77	62.73
90	6382	6344	38	99.40%	0.60%	574380	565429	8951	89.90%	97.07%	7.17	63.17
91	6382	6342	40	99.37%	0.63%	580762	571771	8991	90.91%	97.51%	6.60	63.59
92	6382	6341	41	99.36%	0.64%	587144	578112	9032	91.92%	97.95%	6.03	64.01
93	6382	6350	32	99.50%	0.50%	593526	584462	9064	92.93%	98.30%	5.37	64.48
94	6382	6356	26	99.59%	0.41%	599908	590818	9090	93.94%	98.58%	4.64	65.00
95	6382	6350	32	99.50%	0.50%	606290	597168	9122	94.95%	98.93%	3.98	65.46
96	6382	6365	17	99.73%	0.27%	612672	603533	9139	95.96%	99.11%	3.15	66.04
97	6382	6358	24	99.62%	0.38%	619054	609891	9163	96.97%	99.37%	2.40	66.56
98	6382	6361	21	99.67%	0.33%	625436	616252	9184	97.98%	99.60%	1.62	67.10
99	6382	6364	18	99.72%	0.28%	631818	622616	9202	98.99%	99.79%	0.80	67.66
100	6354	6335	19	99.70%	0.30%	638172	628951	9221	100.00%	100.00%	-	68.21

Table 9.5 Testing Population Results

Testing	# Records		# Goods		# Bads		Fraud Rate					
	159543		157239		2304		0.014441248					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	1596	404	1192	25.31%	74.69%	1596	404	1192	0.26%	51.74%	51.48	0.34
2	1596	1549	47	97.06%	2.94%	3192	1953	1239	1.24%	53.78%	52.53	1.58
3	1596	1588	8	99.50%	0.50%	4788	3541	1247	2.25%	54.12%	51.87	2.84
4	1596	1586	10	99.37%	0.63%	6384	5127	1257	3.26%	54.56%	51.30	4.08
5	1596	1587	9	99.44%	0.56%	7980	6714	1266	4.27%	54.95%	50.68	5.30
6	1596	1588	8	99.50%	0.50%	9576	8302	1274	5.28%	55.30%	50.02	6.52
7	1596	1585	11	99.31%	0.69%	11172	9887	1285	6.29%	55.77%	49.48	7.69
8	1596	1589	7	99.56%	0.44%	12768	11476	1292	7.30%	56.08%	48.78	8.88
9	1596	1580	16	99.00%	1.00%	14364	13056	1308	8.30%	56.77%	48.47	9.98
10	1596	1582	14	99.12%	0.88%	15960	14638	1322	9.31%	57.38%	48.07	11.07
11	1596	1591	5	99.69%	0.31%	17556	16229	1327	10.32%	57.60%	47.27	12.23
12	1596	1578	18	98.87%	1.13%	19152	17807	1345	11.32%	58.38%	47.05	13.24
13	1596	1581	15	99.06%	0.94%	20748	19388	1360	12.33%	59.03%	46.70	14.26
14	1596	1589	7	99.56%	0.44%	22344	20977	1367	13.34%	59.33%	45.99	15.35
15	1596	1584	12	99.25%	0.75%	23940	22561	1379	14.35%	59.85%	45.50	16.36
16	1596	1589	7	99.56%	0.44%	25536	24150	1386	15.36%	60.16%	44.80	17.42
17	1596	1585	11	99.31%	0.69%	27132	25735	1397	16.37%	60.63%	44.27	18.42
18	1596	1587	9	99.44%	0.56%	28728	27322	1406	17.38%	61.02%	43.65	19.43
19	1596	1581	15	99.06%	0.94%	30324	28903	1421	18.38%	61.68%	43.29	20.34
20	1596	1584	12	99.25%	0.75%	31920	30487	1433	19.39%	62.20%	42.81	21.27
21	1596	1587	9	99.44%	0.56%	33516	32074	1442	20.40%	62.59%	42.19	22.24
22	1596	1585	11	99.31%	0.69%	35112	33659	1453	21.41%	63.06%	41.66	23.17
23	1596	1580	16	99.00%	1.00%	36708	35239	1469	22.41%	63.76%	41.35	23.99
24	1596	1578	18	98.87%	1.13%	38304	36817	1487	23.41%	64.54%	41.13	24.76
25	1596	1588	8	99.50%	0.50%	39900	38405	1495	24.42%	64.89%	40.46	25.69
26	1596	1583	13	99.19%	0.81%	41496	39988	1508	25.43%	65.45%	40.02	26.52
27	1596	1583	13	99.19%	0.81%	43092	41571	1521	26.44%	66.02%	39.58	27.33
28	1596	1586	10	99.37%	0.63%	44688	43157	1531	27.45%	66.45%	39.00	28.19
29	1596	1585	11	99.31%	0.69%	46284	44742	1542	28.45%	66.93%	38.47	29.02
30	1596	1581	15	99.06%	0.94%	47880	46323	1557	29.46%	67.58%	38.12	29.75
31	1596	1582	14	99.12%	0.88%	49476	47905	1571	30.47%	68.19%	37.72	30.49
32	1596	1585	11	99.31%	0.69%	51072	49490	1582	31.47%	68.66%	37.19	31.28
33	1596	1584	12	99.25%	0.75%	52668	51074	1594	32.48%	69.18%	36.70	32.04
34	1596	1584	12	99.25%	0.75%	54264	52658	1606	33.49%	69.70%	36.22	32.79
35	1596	1582	14	99.12%	0.88%	55860	54240	1620	34.50%	70.31%	35.82	33.48
36	1596	1584	12	99.25%	0.75%	57456	55824	1632	35.50%	70.83%	35.33	34.21
37	1596	1585	11	99.31%	0.69%	59052	57409	1643	36.51%	71.31%	34.80	34.94
38	1596	1587	9	99.44%	0.56%	60648	58996	1652	37.52%	71.70%	34.18	35.71
39	1596	1586	10	99.37%	0.63%	62244	60582	1662	38.53%	72.14%	33.61	36.45
40	1596	1589	7	99.56%	0.44%	63840	62171	1669	39.54%	72.44%	32.90	37.25
41	1596	1583	13	99.19%	0.81%	65436	63754	1682	40.55%	73.00%	32.46	37.90
42	1596	1584	12	99.25%	0.75%	67032	65338	1694	41.55%	73.52%	31.97	38.57
43	1596	1583	13	99.19%	0.81%	68628	66921	1707	42.56%	74.09%	31.53	39.20
44	1596	1584	12	99.25%	0.75%	70224	68505	1719	43.57%	74.61%	31.04	39.85
45	1596	1582	14	99.12%	0.88%	71820	70087	1733	44.57%	75.22%	30.64	40.44
46	1596	1588	8	99.50%	0.50%	73416	71675	1741	45.58%	75.56%	29.98	41.17
47	1596	1590	6	99.62%	0.38%	75012	73265	1747	46.59%	75.82%	29.23	41.94
48	1596	1587	9	99.44%	0.56%	76608	74852	1756	47.60%	76.22%	28.61	42.63
49	1596	1582	14	99.12%	0.88%	78204	76434	1770	48.61%	76.82%	28.21	43.18
50	1596	1584	12	99.25%	0.75%	79800	78018	1782	49.62%	77.34%	27.73	43.78
51	1596	1583	13	99.19%	0.81%	81396	79601	1795	50.62%	77.91%	27.28	44.35
52	1596	1584	12	99.25%	0.75%	82992	81185	1807	51.63%	78.43%	26.80	44.93
53	1596	1577	19	98.81%	1.19%	84588	82762	1826	52.63%	79.25%	26.62	45.32
54	1596	1582	14	99.12%	0.88%	86184	84344	1840	53.64%	79.86%	26.22	45.84
55	1596	1582	14	99.12%	0.88%	87780	85926	1854	54.65%	80.47%	25.82	46.35
56	1596	1586	10	99.37%	0.63%	89376	87512	1864	55.66%	80.90%	25.25	46.95
57	1596	1583	13	99.19%	0.81%	90972	89095	1877	56.66%	81.47%	24.80	47.47
58	1596	1579	17	98.93%	1.07%	92568	90674	1894	57.67%	82.20%	24.54	47.87
59	1596	1583	13	99.19%	0.81%	94164	92257	1907	58.67%	82.77%	24.10	48.38
60	1596	1589	7	99.56%	0.44%	95760	93846	1914	59.68%	83.07%	23.39	49.03

Testing	# Records		# Goods		# Bads		Fraud Rate					
	159543		157239		2304		0.014441248					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	1596	1586	10	99.37%	0.63%	97356	95432	1924	60.69%	83.51%	22.81	49.60
62	1596	1590	6	99.62%	0.38%	98952	97022	1930	61.70%	83.77%	22.06	50.27
63	1596	1587	9	99.44%	0.56%	100548	98609	1939	62.71%	84.16%	21.45	50.86
64	1596	1584	12	99.25%	0.75%	102144	100193	1951	63.72%	84.68%	20.96	51.35
65	1596	1579	17	98.93%	1.07%	103740	101772	1968	64.72%	85.42%	20.69	51.71
66	1596	1588	8	99.50%	0.50%	105336	103360	1976	65.73%	85.76%	20.03	52.31
67	1596	1584	12	99.25%	0.75%	106932	104944	1988	66.74%	86.28%	19.54	52.79
68	1596	1583	13	99.19%	0.81%	108528	106527	2001	67.75%	86.85%	19.10	53.24
69	1596	1585	11	99.31%	0.69%	110124	108112	2012	68.76%	87.33%	18.57	53.73
70	1596	1581	15	99.06%	0.94%	111720	109693	2027	69.76%	87.98%	18.22	54.12
71	1596	1585	11	99.31%	0.69%	113316	111278	2038	70.77%	88.45%	17.68	54.60
72	1596	1582	14	99.12%	0.88%	114912	112860	2052	71.78%	89.06%	17.29	55.00
73	1596	1581	15	99.06%	0.94%	116508	114441	2067	72.78%	89.71%	16.93	55.37
74	1596	1585	11	99.31%	0.69%	118104	116026	2078	73.79%	90.19%	16.40	55.84
75	1596	1592	4	99.75%	0.25%	119700	117618	2082	74.80%	90.36%	15.56	56.49
76	1596	1580	16	99.00%	1.00%	121296	119198	2098	75.81%	91.06%	15.25	56.82
77	1596	1588	8	99.50%	0.50%	122892	120786	2106	76.82%	91.41%	14.59	57.35
78	1596	1588	8	99.50%	0.50%	124488	122374	2114	77.83%	91.75%	13.93	57.89
79	1596	1589	7	99.56%	0.44%	126084	123963	2121	78.84%	92.06%	13.22	58.45
80	1596	1591	5	99.69%	0.31%	127680	125554	2126	79.85%	92.27%	12.43	59.06
81	1596	1586	10	99.37%	0.63%	129276	127140	2136	80.86%	92.71%	11.85	59.52
82	1596	1586	10	99.37%	0.63%	130872	128726	2146	81.87%	93.14%	11.28	59.98
83	1596	1579	17	98.93%	1.07%	132468	130305	2163	82.87%	93.88%	11.01	60.24
84	1596	1588	8	99.50%	0.50%	134064	131893	2171	83.88%	94.23%	10.35	60.75
85	1596	1585	11	99.31%	0.69%	135660	133478	2182	84.89%	94.70%	9.82	61.17
86	1596	1583	13	99.19%	0.81%	137256	135061	2195	85.90%	95.27%	9.37	61.53
87	1596	1587	9	99.44%	0.56%	138852	136648	2204	86.90%	95.66%	8.76	62.00
88	1596	1582	14	99.12%	0.88%	140448	138230	2218	87.91%	96.27%	8.36	62.32
89	1596	1588	8	99.50%	0.50%	142044	139818	2226	88.92%	96.61%	7.69	62.81
90	1596	1581	15	99.06%	0.94%	143640	141399	2241	89.93%	97.27%	7.34	63.10
91	1596	1589	7	99.56%	0.44%	145236	142988	2248	90.94%	97.57%	6.63	63.61
92	1596	1586	10	99.37%	0.63%	146832	144574	2258	91.95%	98.00%	6.06	64.03
93	1596	1587	9	99.44%	0.56%	148428	146161	2267	92.95%	98.39%	5.44	64.47
94	1596	1590	6	99.62%	0.38%	150024	147751	2273	93.97%	98.65%	4.69	65.00
95	1596	1592	4	99.75%	0.25%	151620	149343	2277	94.98%	98.83%	3.85	65.59
96	1596	1590	6	99.62%	0.38%	153216	150933	2283	95.99%	99.09%	3.10	66.11
97	1596	1593	3	99.81%	0.19%	154812	152526	2286	97.00%	99.22%	2.22	66.72
98	1596	1588	8	99.50%	0.50%	156408	154114	2294	98.01%	99.57%	1.55	67.18
99	1596	1589	7	99.56%	0.44%	158004	155703	2301	99.02%	99.87%	0.85	67.67
100	1539	1536	3	99.81%	0.19%	159543	157239	2304	100.00%	100.00%	-	68.25

Table 9.6 Out Of Time Population Results

OOT	# Records	# Goods	# Bads	Fraud Rate
	163771	161427	2344	0.014312668

	Bins Statistics					Cumulative Statistics						
Population bin %	**# Record**	**# Goods**	**# Bads**	**% Goods**	**% Bads**	**Total # of records**	**Cumulative Goods**	**Cumulative Bads**	**% Goods**	**% Bads (FDR)**	**KS**	**FPR**
1	1638	512	1126	31.26%	68.74%	1638	512	1126	0.32%	48.04%	47.72	0.45
2	1638	1563	75	95.42%	4.58%	3276	2075	1201	1.29%	51.24%	49.95	1.73
3	1638	1629	9	99.45%	0.55%	4914	3704	1210	2.29%	51.62%	49.33	3.06
4	1638	1622	16	99.02%	0.98%	6552	5326	1226	3.30%	52.30%	49.00	4.34
5	1638	1619	19	98.84%	1.16%	8190	6945	1245	4.30%	53.11%	48.81	5.58
6	1638	1628	10	99.39%	0.61%	9828	8573	1255	5.31%	53.54%	48.23	6.83
7	1638	1623	15	99.08%	0.92%	11466	10196	1270	6.32%	54.18%	47.86	8.03
8	1638	1625	13	99.21%	0.79%	13104	11821	1283	7.32%	54.74%	47.41	9.21
9	1638	1628	10	99.39%	0.61%	14742	13449	1293	8.33%	55.16%	46.83	10.40
10	1638	1618	20	98.78%	1.22%	16380	15067	1313	9.33%	56.02%	46.68	11.48
11	1638	1629	9	99.45%	0.55%	18018	16696	1322	10.34%	56.40%	46.06	12.63
12	1638	1629	9	99.45%	0.55%	19656	18325	1331	11.35%	56.78%	45.43	13.77
13	1638	1626	12	99.27%	0.73%	21294	19951	1343	12.36%	57.30%	44.94	14.86
14	1638	1624	14	99.15%	0.85%	22932	21575	1357	13.37%	57.89%	44.53	15.90
15	1638	1626	12	99.27%	0.73%	24570	23201	1369	14.37%	58.40%	44.03	16.95
16	1638	1626	12	99.27%	0.73%	26208	24827	1381	15.38%	58.92%	43.54	17.98
17	1638	1621	17	98.96%	1.04%	27846	26448	1398	16.38%	59.64%	43.26	18.92
18	1638	1629	9	99.45%	0.55%	29484	28077	1407	17.39%	60.03%	42.63	19.96
19	1638	1623	15	99.08%	0.92%	31122	29700	1422	18.40%	60.67%	42.27	20.89
20	1638	1630	8	99.51%	0.49%	32760	31330	1430	19.41%	61.01%	41.60	21.91
21	1638	1628	10	99.39%	0.61%	34398	32958	1440	20.42%	61.43%	41.02	22.89
22	1638	1627	11	99.33%	0.67%	36036	34585	1451	21.42%	61.90%	40.48	23.84
23	1638	1632	6	99.63%	0.37%	37674	36217	1457	22.44%	62.16%	39.72	24.86
24	1638	1627	11	99.33%	0.67%	39312	37844	1468	23.44%	62.63%	39.18	25.78
25	1638	1627	11	99.33%	0.67%	40950	39471	1479	24.45%	63.10%	38.65	26.69
26	1638	1625	13	99.21%	0.79%	42588	41096	1492	25.46%	63.65%	38.19	27.54
27	1638	1625	13	99.21%	0.79%	44226	42721	1505	26.46%	64.21%	37.74	28.39
28	1638	1624	14	99.15%	0.85%	45864	44345	1519	27.47%	64.80%	37.33	29.19
29	1638	1631	7	99.57%	0.43%	47502	45976	1526	28.48%	65.10%	36.62	30.13
30	1638	1633	5	99.69%	0.31%	49140	47609	1531	29.49%	65.32%	35.82	31.10
31	1638	1630	8	99.51%	0.49%	50778	49239	1539	30.50%	65.66%	35.15	31.99
32	1638	1629	9	99.45%	0.55%	52416	50868	1548	31.51%	66.04%	34.53	32.86
33	1638	1632	6	99.63%	0.37%	54054	52500	1554	32.52%	66.30%	33.77	33.78
34	1638	1621	17	98.96%	1.04%	55692	54121	1571	33.53%	67.02%	33.50	34.45
35	1638	1621	17	98.96%	1.04%	57330	55742	1588	34.53%	67.75%	33.22	35.10
36	1638	1624	14	99.15%	0.85%	58968	57366	1602	35.54%	68.34%	32.81	35.81
37	1638	1627	11	99.33%	0.67%	60606	58993	1613	36.54%	68.81%	32.27	36.57
38	1638	1628	10	99.39%	0.61%	62244	60621	1623	37.55%	69.24%	31.69	37.35
39	1638	1630	8	99.51%	0.49%	63882	62251	1631	38.56%	69.58%	31.02	38.17
40	1638	1624	14	99.15%	0.85%	65520	63875	1645	39.57%	70.18%	30.61	38.83
41	1638	1629	9	99.45%	0.55%	67158	65504	1654	40.58%	70.56%	29.99	39.60
42	1638	1624	14	99.15%	0.85%	68796	67128	1668	41.58%	71.16%	29.58	40.24
43	1638	1631	7	99.57%	0.43%	70434	68759	1675	42.59%	71.46%	28.86	41.05
44	1638	1631	7	99.57%	0.43%	72072	70390	1682	43.60%	71.76%	28.15	41.85
45	1638	1621	17	98.96%	1.04%	73710	72011	1699	44.61%	72.48%	27.87	42.38
46	1638	1623	15	99.08%	0.92%	75348	73634	1714	45.61%	73.12%	27.51	42.96
47	1638	1627	11	99.33%	0.67%	76986	75261	1725	46.62%	73.59%	26.97	43.63
48	1638	1625	13	99.21%	0.79%	78624	76886	1738	47.63%	74.15%	26.52	44.24
49	1638	1633	5	99.69%	0.31%	80262	78519	1743	48.64%	74.36%	25.72	45.05
50	1638	1621	17	98.96%	1.04%	81900	80140	1760	49.64%	75.09%	25.44	45.53
51	1638	1627	11	99.33%	0.67%	83538	81767	1771	50.65%	75.55%	24.90	46.17
52	1638	1627	11	99.33%	0.67%	85176	83394	1782	51.66%	76.02%	24.36	46.80
53	1638	1627	11	99.33%	0.67%	86814	85021	1793	52.67%	76.49%	23.82	47.42
54	1638	1624	14	99.15%	0.85%	88452	86645	1807	53.67%	77.09%	23.42	47.95
55	1638	1628	10	99.39%	0.61%	90090	88273	1817	54.68%	77.52%	22.83	48.58
56	1638	1628	10	99.39%	0.61%	91728	89901	1827	55.69%	77.94%	22.25	49.21
57	1638	1625	13	99.21%	0.79%	93366	91526	1840	56.70%	78.50%	21.80	49.74
58	1638	1628	10	99.39%	0.61%	95004	93154	1850	57.71%	78.92%	21.22	50.35
59	1638	1624	14	99.15%	0.85%	96642	94778	1864	58.71%	79.52%	20.81	50.85
60	1638	1621	17	98.96%	1.04%	98280	96399	1881	59.72%	80.25%	20.53	51.25

OOT	# Records		# Goods		# Bads		Fraud Rate					
	163771		161427		2344		0.014312668					
	Bins Statistics					Cumulative Statistics						
Population bin %	# Record	# Goods	# Bads	% Goods	% Bads	Total # of records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
61	1638	1628	10	99.39%	0.61%	99918	98027	1891	60.73%	80.67%	19.95	51.84
62	1638	1619	19	98.84%	1.16%	101556	99646	1910	61.73%	81.48%	19.76	52.17
63	1638	1628	10	99.39%	0.61%	103194	101274	1920	62.74%	81.91%	19.17	52.75
64	1638	1625	13	99.21%	0.79%	104832	102899	1933	63.74%	82.47%	18.72	53.23
65	1638	1623	15	99.08%	0.92%	106470	104522	1948	64.75%	83.11%	18.36	53.66
66	1638	1633	5	99.69%	0.31%	108108	106155	1953	65.76%	83.32%	17.56	54.35
67	1638	1627	11	99.33%	0.67%	109746	107782	1964	66.77%	83.79%	17.02	54.88
68	1638	1625	13	99.21%	0.79%	111384	109407	1977	67.77%	84.34%	16.57	55.34
69	1638	1624	14	99.15%	0.85%	113022	111031	1991	68.78%	84.94%	16.16	55.77
70	1638	1630	8	99.51%	0.49%	114660	112661	1999	69.79%	85.28%	15.49	56.36
71	1638	1621	17	98.96%	1.04%	116298	114282	2016	70.79%	86.01%	15.21	56.69
72	1638	1627	11	99.33%	0.67%	117936	115909	2027	71.80%	86.48%	14.67	57.18
73	1638	1625	13	99.21%	0.79%	119574	117534	2040	72.81%	87.03%	14.22	57.61
74	1638	1624	14	99.15%	0.85%	121212	119158	2054	73.82%	87.63%	13.81	58.01
75	1638	1631	7	99.57%	0.43%	122850	120789	2061	74.83%	87.93%	13.10	58.61
76	1638	1622	16	99.02%	0.98%	124488	122411	2077	75.83%	88.61%	12.78	58.94
77	1638	1621	17	98.96%	1.04%	126126	124032	2094	76.83%	89.33%	12.50	59.23
78	1638	1623	15	99.08%	0.92%	127764	125655	2109	77.84%	89.97%	12.13	59.58
79	1638	1628	10	99.39%	0.61%	129402	127283	2119	78.85%	90.40%	11.55	60.07
80	1638	1626	12	99.27%	0.73%	131040	128909	2131	79.86%	90.91%	11.06	60.49
81	1638	1620	18	98.90%	1.10%	132678	130529	2149	80.86%	91.68%	10.82	60.74
82	1638	1625	13	99.21%	0.79%	134316	132154	2162	81.87%	92.24%	10.37	61.13
83	1638	1622	16	99.02%	0.98%	135954	133776	2178	82.87%	92.92%	10.05	61.42
84	1638	1628	10	99.39%	0.61%	137592	135404	2188	83.88%	93.34%	9.47	61.88
85	1638	1630	8	99.51%	0.49%	139230	137034	2196	84.89%	93.69%	8.80	62.40
86	1638	1630	8	99.51%	0.49%	140868	138664	2204	85.90%	94.03%	8.13	62.91
87	1638	1629	9	99.45%	0.55%	142506	140293	2213	86.91%	94.41%	7.50	63.39
88	1638	1632	6	99.63%	0.37%	144144	141925	2219	87.92%	94.67%	6.75	63.96
89	1638	1625	13	99.21%	0.79%	145782	143550	2232	88.93%	95.22%	6.30	64.31
90	1638	1624	14	99.15%	0.85%	147420	145174	2246	89.93%	95.82%	5.89	64.64
91	1638	1625	13	99.21%	0.79%	149058	146799	2259	90.94%	96.37%	5.44	64.98
92	1638	1633	5	99.69%	0.31%	150696	148432	2264	91.95%	96.59%	4.64	65.56
93	1638	1632	6	99.63%	0.37%	152334	150064	2270	92.96%	96.84%	3.88	66.11
94	1638	1624	14	99.15%	0.85%	153972	151688	2284	93.97%	97.44%	3.47	66.41
95	1638	1628	10	99.39%	0.61%	155610	153316	2294	94.98%	97.87%	2.89	66.83
96	1638	1626	12	99.27%	0.73%	157248	154942	2306	95.98%	98.38%	2.40	67.19
97	1638	1626	12	99.27%	0.73%	158886	156568	2318	96.99%	98.89%	1.90	67.54
98	1638	1631	7	99.57%	0.43%	160524	158199	2325	98.00%	99.19%	1.19	68.04
99	1638	1625	13	99.21%	0.79%	162162	159824	2338	99.01%	99.74%	0.74	68.36
100	1609	1603	6	99.63%	0.37%	163771	161427	2344	100.00%	100.00%	-	68.87