CREDIT CARD TRANSACTION FRAUD DETECTION

Ge Zeng, Xiaoying (Shawn) Zou, Xinyue Yu, Zihan (Johnson) Ling, Zihang (Charlie) Li

DSO 560 | Fraud Analysis | Project 3

Content

1. Executive Summary

With the development of internet technology, transaction frauds have become a serious problem. According to KPMG international study on banking risk, the transaction fraud is one of the top 5 challenges which banks are confronted with nowadays. It leads to billions of fraud losses each year and there is a growing need for detection and prevention.

This project focuses on helping banks identify credit card transaction frauds in active accounts. After analyzing each transaction, we designed several supervised machine learning models and selected the best-performing model to help the bank identify the risk, detect the frauds, and mitigate fraudulent activities.

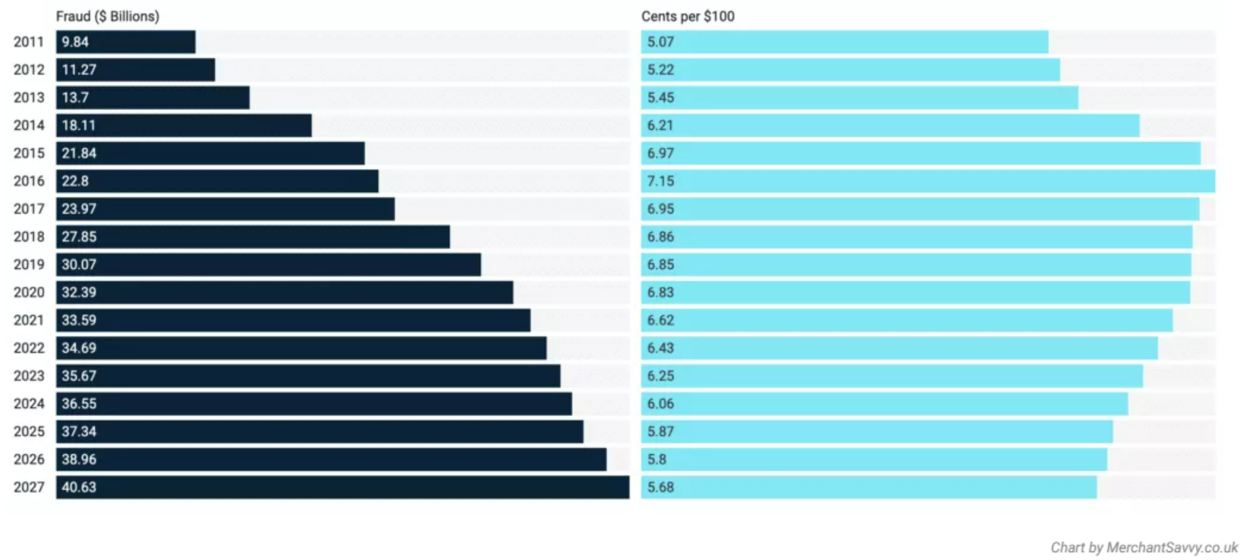
The original dataset contains credit card transaction information to identify fraudulent actions in 2006. There are 96,753 records and 10 columns in this dataset. We started by exploring the data and writing the data quality report to provide a basic description of the data. With a general idea of our dataset, we created 1,108 new variables through feature engineering, and selected 20 important features as our final features through the feature selection process. By reserving the last two months of data as the out of time sample and randomly splitting the training and testing data in the proportion of 8:2, we constructed logistic regression, single decision tree, random forest, boosted tree, neural network, and adaptive boosting models using these 20 features, experimented with different hyperparameters, and found out the best performer, a random forest model. Finally, we evaluated its performance on the Training, Testing, and out-of-time dataset. Our random forest model detected 90.91% of fraud in the Training data, 86.36% of fraud in the Testing data, and 60.34% of fraud in the out-of-time data by looking at the top 3% of the corresponding dataset.

2. Background

The transaction fraud is one of the most serious issues of online security. The perpetrator takes away funds, personal property, interest or sensitive information from victims by illegally obtaining credit card information. This is not only harmful to the interests of consumers, but also problematic for merchants.

According to the Nilson report, payment fraud losses have increased each year. Fraud losses have tripled since 2011 and are expected to exceed $40 Billion by 2027.

Figure 1. *Payment Fraud Losses Each Year*



According to the report from Aite-Novarica, the number of the card-not-present frauds conducted online or over the phone continues to grow.

Figure 2. *Card-Not-Present Fraud Losses*

Although the percentage of attempted fraudulent transactions in the U.S. is substantially lower than in Mexico, the value of these transactions is much higher in the U.S. except for Q4 2020 ($149 in the U.S. vs $155 in Mexico). This shows that although frauds may be not as prevalent in the U.S. as in Mexico, merchants should still establish and enforce an anti-fraud system to protect U.S. transactions given the higher value of each fraudulent transaction.

Figure 3. *Fraudulent Transactions in U.S. and Mexico*

Based on the statistical data above, we realize the importance of building models to help banks identify credit card transaction frauds, including credit card transactions, money transfers from compromised accounts, insurance claims, etc.

3. Description of Data

This dataset contains credit card transaction information from a U.S. governmental organization and synthetic fraud labels. The dataset includes 96,753 records of transactions and 10 fields. It covers records from 1/1/2006 to 12/31/2006. This dataset includes 1,059 transaction frauds, accounting for 1.09% of all records.

3.1 Summary statistics table

According to the numeric field summary table, the maximum transaction amount is over 3 million dollars, which is suspicious since the mean was only 427 dollars.

Table 1. *Numeric field summary*

Field Name	% Populated	Min	Max	Mean	Stdev	*% Zero
Date	100	2006-01-01	2006-12-31	-	-	0
Amount	100	0.01	3,102,045.53	427.89	10,006.14	0

*% Zero: only including record whose value is 0

Based on the categorical fields summary table below, Merchnum, Merch state and Merch zip have missing values. We will pay attention to these values in the data cleaning section.

Table 2. *Categorical field summary*

Field Name	% Populated	*Unique Values	Most Common Value
Recnum	100	96,753	-
Cardnum	100	1,645	5142148452
Merchnum	96.51	13,091	930090121224
Merch description	100	13,126	GSA-FSS-ADV
Merch state	98.76	227	TN
Merch zip*	95.19	4,567	38118

Field Name	% Populated	*Unique Values	Most Common Value
Transtype	100	4	P
Fraud	100	2	0

*Unique Values: does not include Nan

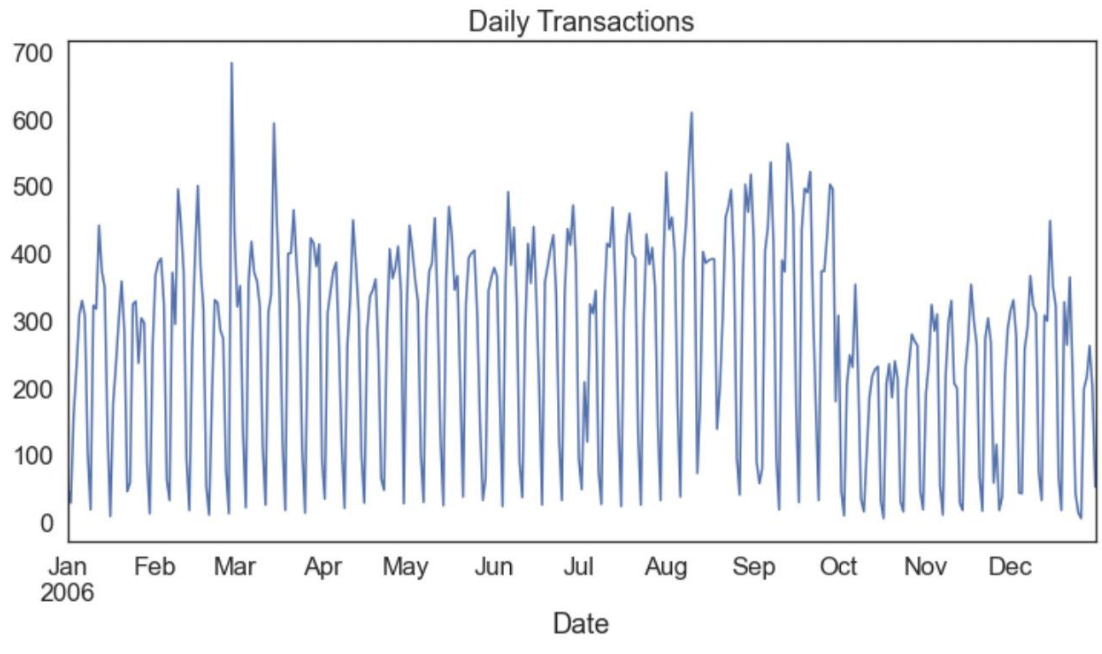
*Transformed Merch zip from float as int

3.2 Field Distribution

3.2.1 date

This field is the date of all transaction records. The plot below shows the number of transactions by day. The graph is spiky with weekly patterns, indicating that transaction shrinks on weekends. The daily transaction graph shows that the average number of transactions from October to December is significantly smaller than in previous months. One explanation is that the government fiscal year starts in October and people would save their spending and allocate budget for the following year. Therefore, it may affect the accuracy of our prediction when we use November and December data as our Out of Time validation data.

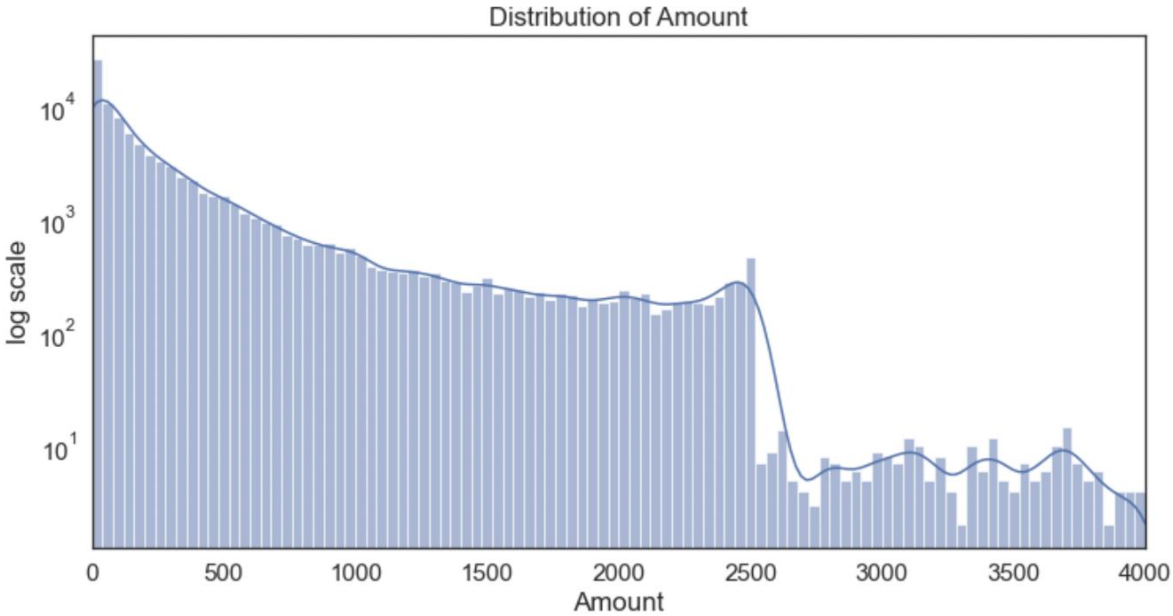
Figure 4. *Daily Transactions*



3.2.2 Amount

This field is a numerical variable which specifies the transaction amount. The graph below shows the distribution of amount with log scale, and it covers 99.50% of all values. Most transactions are small amounts transactions.

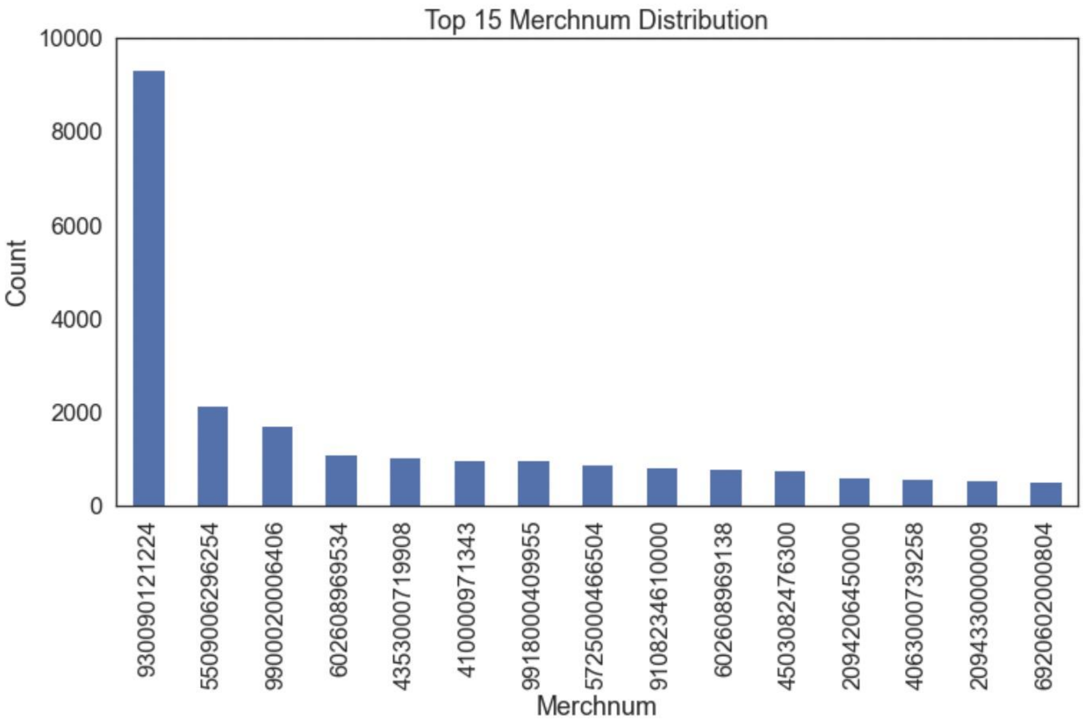
Figure 5. *Distribution of Amount*



3.2.3 Merchnum

This field is the merchant number of all transaction records. The plot below shows the top 15 merchants with the most transactions. The Merchnum 930090121224 has over 8,000 transactions, and we find out that this merchant is FedEx.

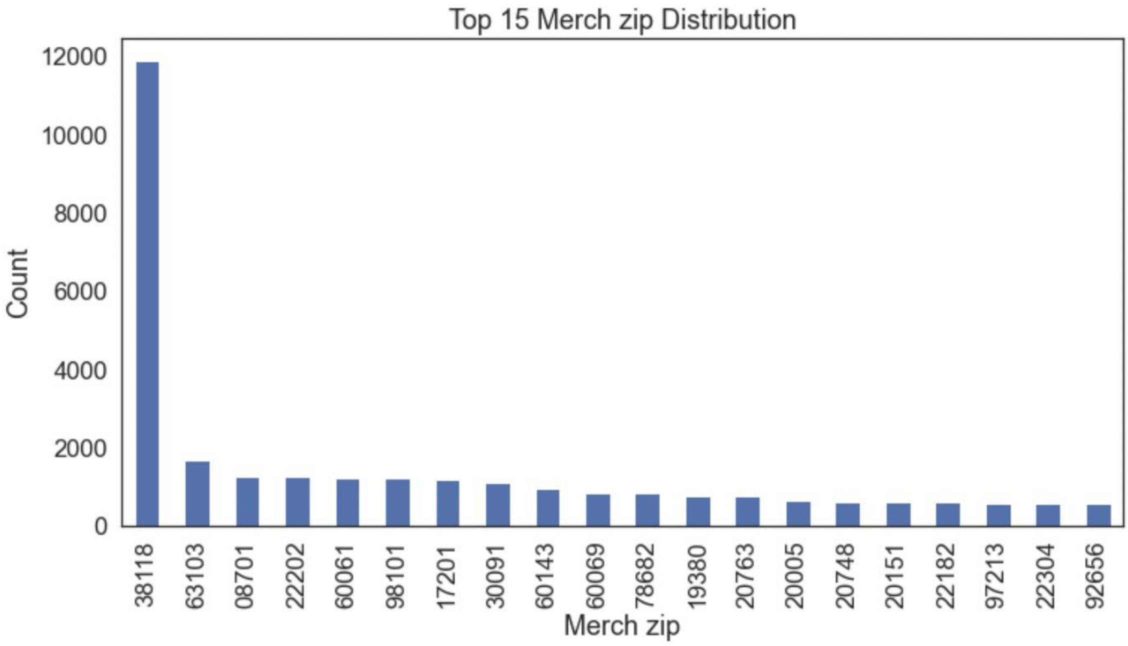
Figure 6. *Distribution of Merchnum*



3.2.4 Merch zip

This field is the Merch zip code of all transaction records. The plot below shows the top 15 Merch zip code with the most transactions. We can see that zip code 38118 has much more transactions than other areas, and we found out that the zip code is in Memphis, Tennessee, where FedEx's headquarter is located. That explains why we get a large number of transactions from this zip code.

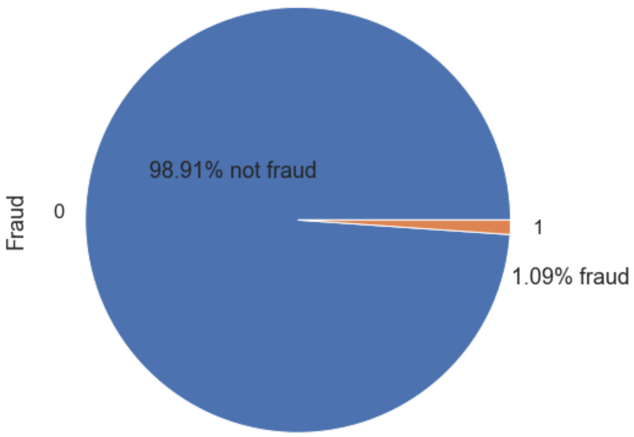
Figure 7. *Distribution of Merch zip.*



3.2.5 Fraud

This field is a fraud label indicating whether the record is fraud or not. 0 represents not fraud, and 1 represents fraud. As shown in the pie chart below, there is only 1.09% of fraud.

Figure 8. *Distribution of Fraud*



4. Data Cleaning

4.1 Fix Data Type

We changed the type of Date variable into datetime.

4.2 Remove Irrelevant Transaction Type

We only kept the transaction with type "P" to conduct our analysis since we mainly focus on transaction fraud on purchase activity.

4.3 Remove Outlier

We removed the outlier of Amount because the number is suspicious and will affect our result.

4.4 Fill in Merchnum Field

There are 3.49% of records missing Merchnum field. Although the Merchnum is missing, we could use some related information to infer the Merchnum. Therefore, we found the most common (mode) Merchnum over all records that has the particular Merch description. Then we matched and filled with the mode Merchnum with a particular Merch description. Finally, we assigned unknown for the remaining transactions.

4.5 Fill in Merch State

There are 1.24% of records missing Merch state. Since state has a strong relationship with zip code, we could use Merch zip to infer the missing Merch state. Firstly, we built dictionary to match zip code and state. Then we filled in missing values by mapping with zip code. After filling in the missing value by utilizing zip code, we built dictionaries to match the Merch state with Merchnum and Merch description, and fill in missing values by mapping with Merchnum and Merch description. Finally, we assigned unknown for the remaining transactions.

4.6 Fill in Merch Zip

There are 4.81% of records missing Merch zip. Firstly, we built dictionaries to match Merchnum and Merch description with zip code. Then we filled in missing values by mapping with Merchnum and Merch description. Finally, we assigned unknown for adjustment transactions and the remaining transactions.

5. Candidate Variables

To build expert variables for our credit card transaction project, we investigated when and how transaction fraud occurs. According to Paygilant, credit card transaction fraud happens when someone uses a stolen credit card to conduct unauthorized transaction from a merchant. There are two scenarios that can be derived from such behavior.

On the one hand, a person can utilize a stolen credit card at various merchant locations and conduct multiple transactions. For instance, a credit card happens to have several numbers of transactions in three different states with large amounts. There are burst of activities, abnormally large purchases and transactions in different merchants and geographical regions associated with this card number. On the other hand, suspicious transactions can be associated with a high risk or fictitious merchant. For example, if suddenly on one day 50 new card numbers appear with large transaction amounts from a single merchant, we might postulate that somebody at this merchant is using stolen information to conduct unauthorized transactions. Based on these two scenarios, we built new entities besides the 5 original entities (Cardnum, Merchnum, Merch description, Merch state, Merch zip) to ensure that our algorithm could catch such abnormalities as many as possible.

Table 3. *New Entities and Compositions*

New Entity	Composition
Fulladdress	Merch state + Merch zip
Merchnum_Merch description	Merchnum + Merch description
Merchnum_fulladdress	Merchnum + Fulladdress
Cardnum_fulladdress	Cardnum + Fulladdress
Cardnum_Merchnum	Cardnum + Merchnum
Cardnum_Merch description	Cardnum + Merch description
Cardnum_Merch state	Cardnum + Merch state
Cardnum_Merch zip	Cardnum + Merch zip

Final entity list: ['Cardnum', 'Merchnum', 'Merch description', 'Merch state', 'Merch zip', 'fulladdress', 'Merchnum_Merch description', 'Merchnum_fulladdress', 'Cardnum_fulladdress', 'Cardnum_Merchnum', 'Cardnum_Merch description', 'Cardnum_Merch state', 'Cardnum_Merch zip']

We built 6 different types of variables based on these entities and in total, we created 1,108 new variables. (For a full list of variables created, please refer to appendix 10.2) While most variables track unusual activities in the past (0,1,3,7,14,30) days, we extended the number of days in these variables from 30 days to 60 and 90 days to track long term fraudulent activities. For each type of the variables, we provided the formula, logic, example, number of variables created as follows:

1) Frequency variables

Formula:

$$\text{Frequency} = \frac{\text{count of entity e over past n days}}{n} \qquad \text{for n} \in \{0,1,3,7,14,30,60,90\}$$

Logic: The number of times that we see the same entity appear for a given time period (from 0 to 90 days). It measures the frequency of certain entity being used. So a transaction might be considered fraudulent if a single entity appears too often in past records.

Example: Cardnum_count_30

Number of created variables: 104

2) Day since variables

Formula:

$$d_{since} = d_{new} - d_{old}$$

Logic: Difference in the number of days since we last saw this entity. This type of variable is useful because we can find out the interval between the appearance of the same entity and easily identify abnormal transactions if an entity shows up too often.

Example: Cardnum_day_since

Number of created variables: 13

3) Amount variables

Formula:

$$\text{Amount variables} = \frac{f(Amount) \quad \text{at entity e over past n days}}{n} \qquad \text{for n} \in \{0,1,3,7,14,30,60,90\},$$

$$\text{for } f(Amount) \in \{\text{avg(Amount), max(Amount), median(Amount), sum(Amount)}\} \ \&$$
$$\{\text{Actual/avg(Amount), Actual/max(Amount), Actual/median(Amount), Actual/sum(Amount)}\}$$

Logic: Different numeric calculation of Amount at given entity for a given time period. It describes the amount of transactions this entity involves in different measurement criteria. A transaction would be considered fraudulent if it is related to unusual amount of transactions in a given time period.

Example: Cardnum_avg_30

Number of created variables: 832

4) Velocity change variables

Formula:

$$\text{Velocity change variables}_{(x,y)} = \frac{\text{count of entity e over past x days}}{\text{average count of entity e over past y days}} \qquad \text{for x} \in \{0,1\}, \quad \text{for y} \in \{3,7,14,30,60,90\}$$

Logic: The number of times that we see this entity appears in the last (0, 1) day divided by the average number of times we see this entity appears in the last (3,7,14,30,60,90) days. This type of variables measures whether we see a surge in transactions with the same entity in a relatively short period of time.

Example: Cardnum_count_0_by_30

Number of created variables: 156

5) Target Encoded Variables

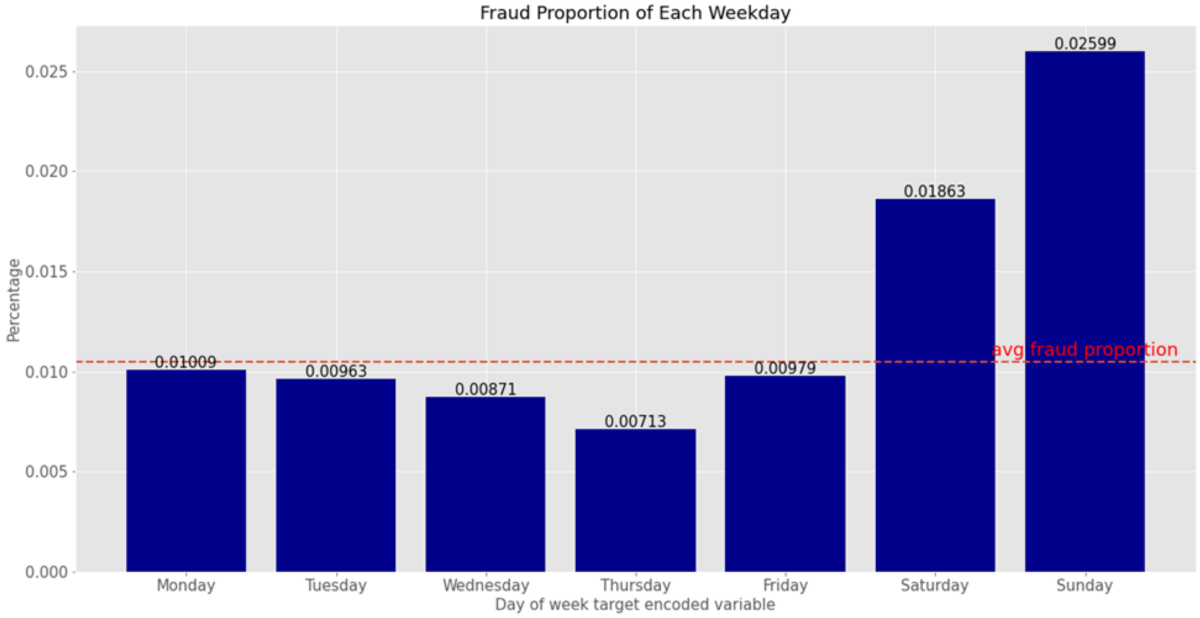
Formula:

$$Y_{dow} = \frac{\text{\# of frauds}}{\text{\# of records for each day of week}} \qquad \text{(prior to 11/01/2006)}$$

Statistical smoothing: $\text{Value} = Y_{\text{low}} + \dfrac{Y_{\text{high}} - Y_{\text{low}}}{1 + e^{-(n-n_{mid})/c}}$ where c = 4 and nmid = 20

Logic: For each day of the week, we can use the target encoding to convert the categorical field into a numeric field and find the relevant risk of fraud transactions on that day.

Figure 9. *Fraud Proportion of Each Weekday*



Number of created variables: 1

6) Benford's law variables: removing all transactions from FedEx and group transactions by cardnum and merchnum, calculating the ratio of 1s and 2s for first digit of amount versus 3s to 9s.

Formula:

$$R = \frac{1.096 \times n_{\text{low}}}{n_{\text{high}}} \qquad U = \max(R, 1/R)$$

where low represents number of transactions starting with digit 1 or 2 for amount and high represents number of transactions starting with 3 to 9 for amount.

Statistical Smoothing: $U^* = 1 + \left(\dfrac{U-1}{1+\exp^{-t}}\right) \qquad t = (n - n_{mid})/c$ where nmid = 15 and c = 3

Logic: The distribution of first digit of amount field for each cardholder and merchant should be around a ratio of 1.096 for digits (3 through 9) over (1 and 2). If the cardnum or merchnum violates too much of this ratio then this number becomes suspicious.

Table 4. *Cardnum and Merchnum Examples According to Benford's Law*

Cardnum	initial1-2	initial3-9	n	R	1/R	U	t	U*
5142253356	61	5	66	13.37	0.07	13.37	17	13.37
5142299705	25	3	28	9.13	0.11	9.13	4.33	9.03
Merchnum	**initial1-2**	**initial3-9**	**n**	**R**	**1/R**	**U**	**t**	**U***
991808369338	1	181	181	0.01	165.15	165.15	55.33	165.15
8078200641472	59	1	60	64.66	0.02	64.66	15	64.66

Number of created variables: 2

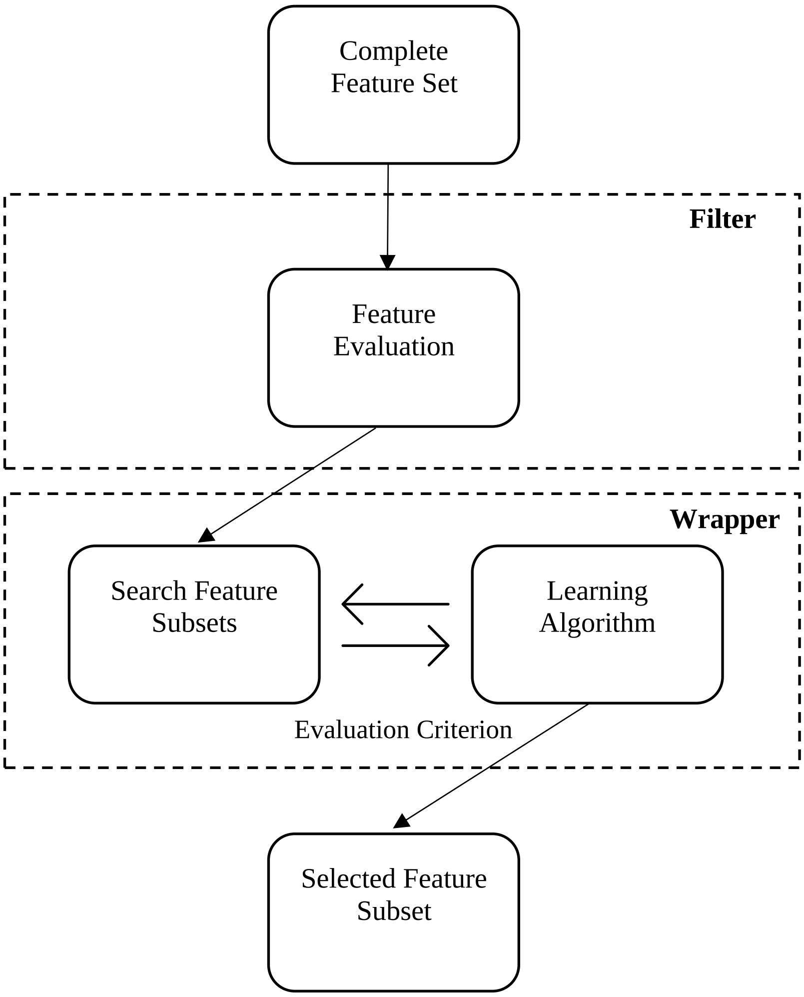
6. Feature Selection Process

6.1 Feature Selection Workflow

Through feature engineering, we created 1,108 candidates variables, generating a high dimensional space. According to the curse of dimensionality, data becomes sparse and all points become outliers in the high dimensional space, so we need much more data for model fitting. In addition, high dimensions make non-linear models run much slower to find the optimal solution. Therefore, we need to go through feature selection process before training the models. In this project, we applied filter and wrapper methods in the feature selection process.

As the feature selection workflow (Figure 10) shown below, we first applied a 2-sample KS as a filter to select the top 80 variables in univariate significance. We then used Random Forest Classifier as our wrapper for sequential forward selection to select the top 20 variables.

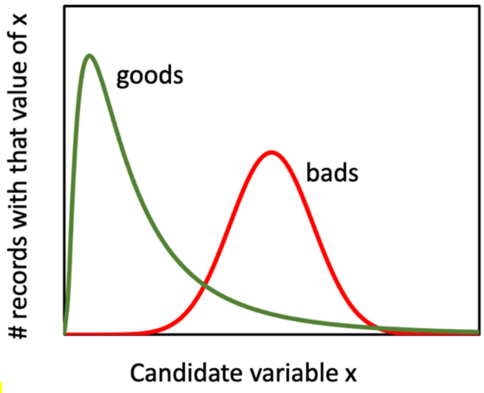
Figure 10. *Feature Selection Workflow*



6.2 Filter Methods

The filter method examines the univariate relationship between the independent variable and the dependent variable. In this project, we used the two-sample Kolmogorov-Smirnov (KS) test to filter variables. As Figure 11 shown below, the two-sample KS test quantifies the distance between two distributions. After applying the filter, we picked the top 80 variables with high KS scores from the 1,108 candidate variables.

Figure 11. *Distribution Plot of Kolmogorov-Smirnov Test*

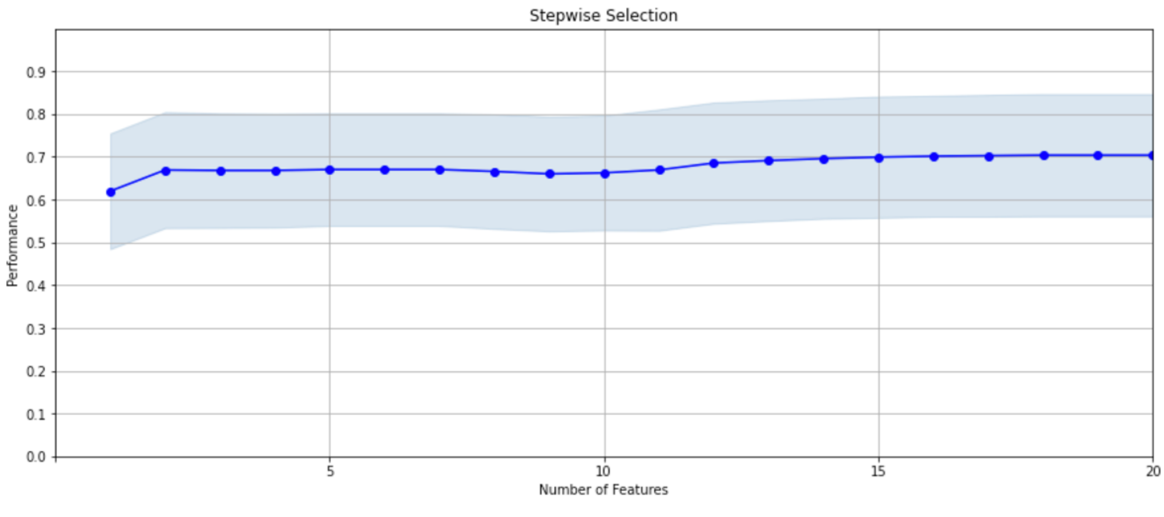


6.3 Wrapper Methods

The wrapper method uses a subset of features and train a model to choose the features that have the the strongest multivariate relationship with the dependent variable.

We used Random Forest Classifier as our wrapper for forward selection. Forward selection is an iterative method in which we start with having no features in the model. In each iteration, we keep adding the feature which best improves our model until adding a new variable does not significantly improve the model performance. We used the fraud detection rate as an evaluation criterion of the wrapper in the 10-fold Cross Validation. After applying the wrapper, we selected the 20 final features from 80 variables.

Figure 12. *Forward Selection Performance of Top 20 Variables*



6.4 Final Variables

The table below is the top 20 variables sorted by the wrapper order. We also provide the description and the KS score of each variable.

Table 5. *Top 20 Variables and Variable Descriptions*

Order	Variable Name	Description	KS Score
1	Cardnum_Merchnum_total_7	total transactions in that card number and merchandise number over the past 7 days	0.681
2	Cardnum_Merch zip_max_90	the maximum transaction in that card number and merchandise zip code over the past 90 days	0.621
3	Cardnum_Merchnum_total_30	total transactions in that card number and merchandise number over the past 30 days	0.659
4	Cardnum_Merch zip_total_30	total transactions in that card number and merchandise zip code over the past 30 days	0.656
5	Cardnum_Merchnum_total_60	total transactions in that card number and merchandise zip code over the past 60 days	0.643
6	Cardnum_fulladdress_total_60	total transactions in that card number and full address over the past 60 days	0.647
7	Cardnum_Merch zip_total_60	total transactions in that card number and merchandise zip code over the past 60 days	0.646
8	Cardnum_Merch description_total_60	total transactions in that card number and merchandise description over the past 60 days	0.646
9	Cardnum_Merch description_total_3	total transactions in that card number and merchandise description over the past 3 days	0.661
10	Cardnum_fulladdress_max_30	the maximum transaction in that card number and full address over the past 30 days	0.652
11	Merch state_actual/med_90	total transactions in that merchandise state of that day divided by the median transactions over the last 90 days	0.615
12	Cardnum_Merch state_max_14	the maximum transaction in that card number and merchandise state over the past 14 days	0.631
13	Merch description_max_0	the maximum transaction in that merchandise description in the day of transaction happened	0.609
14	Cardnum_Merch description_max_90	the maximum transaction in that card number and merchandise description over the past 90 days	0.631
15	Cardnum_Merch description_total_90	total transactions in that card number and merchandise description over the past 90 days	0.640
16	Cardnum_Merchnum_max_90	the maximum transaction in that card number and merchandise number over the past 90 days	0.629
17	Cardnum_fulladdress_max_7	the maximum transaction in that card number and full address over the past 7 days	0.658

Order	Variable Name	Description	KS Score
18	Cardnum_Merch description_max_7	the maximum transaction in that card number and merchandise description over the past 7 days	0.657
19	Cardnum_Merch zip_max_14	the maximum transaction in that card number and merchandise zip code over the past 14 days	0.658
20	Cardnum_Merch zip_max_7	the maximum transaction in that card number and merchandise zip code over the past 7 days	0.657

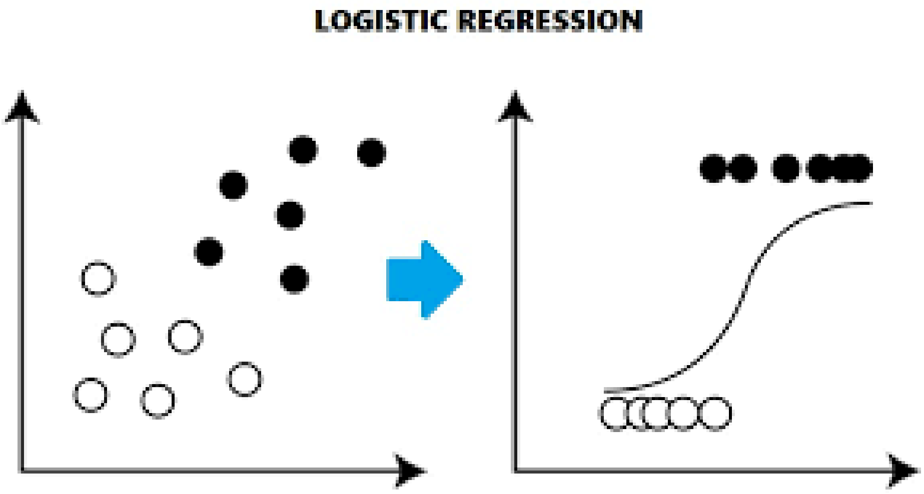
7. Model Algorithms

To select the most effective model, we compared results among 7 different models including logistic regression, single decision tree, random forest, boosted tree, neural network, and adaptive boosting models. We also adjusted critical hyperparameters in different models to achieve the best performance of each model. We used records before 11/01/2006 as our training and testing dataset and randomly split the training and testing data in the proportion of 4:1. Then we used records of the last two months (11/01/2006-12/31/2006) as out-of-time (OOT) data. To measure the results, we used 3% of the population in the training dataset, testing dataset, and OOT dataset respectively.

7.1 Logistic Regression

The logistic regression model is used to predict the class of individuals based on one or multiple predictor variables. It is a supervised algorithm that learns a linear relationship from the given dataset and then introduces non-linearity through the Sigmoid function. In our case, the model is used to model the binary outcome to predict whether a record is a fraud or not.

Figure 13. *Illustration of Logistic Regression*



Hyperparameters:

- penalty: It specifies the norm of the penalty. It has three values, ‘l1’, ‘l2’, ‘elasticnet’ and ‘none’, and the default value of penalty is ‘l2’.
- solver: It indicates the algorithm to use in the optimization problem. It has five values, ‘newton-cg’, ‘lbfgs’, ‘liblinear’, ‘sag’ and ‘saga’. The default value is ‘lbfgs’.

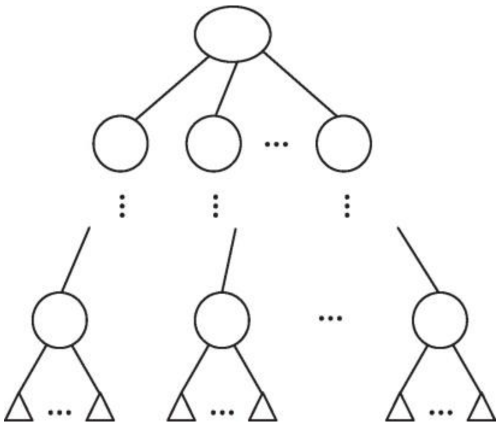
Table 6. *Hyperparameter Tuning of Logistic Regression*

Model		Parameters			Avg FDR at 3%		
	Iteration	Variables	penalty	solver	trn	tst	oot
Logistic Regression	1	10	none	lbfgs	0.657	0.681	0.360
	2	10	l2	lbfgs	0.659	0.673	0.362
	3	10	none	saga	0.652	0.677	0.358
	4	10	l2	saga	0.658	0.669	0.361
	5	15	none	lbfgs	0.678	0.674	0.366
	6	15	l2	lbfgs	0.688	0.667	0.373
	7	15	none	saga	0.680	0.697	0.359
	8	15	l2	saga	0.678	0.708	0.380
	9	20	none	saga	0.684	0.666	0.362
	10	20	l2	saga	0.677	0.699	0.377
	11	20	none	lbfgs	0.679	0.679	0.363
	12	20	l2	lbfgs	0.684	0.690	0.380

7.2 Single Decision Tree

The decision tree model is a supervised machine learning algorithm that can be used for classification and regression. A decision tree is a flowchart resembling a tree structure where each internal node denotes a sub-classifier on an attribute, each branch represents an outcome of the classification, and each leaf node (terminal node) holds a class label. In our case, we use the decision tree to classify a record as fraud or not a fraud

Figure 14. *Illustration of Single Decision Tree*



Hyperparameters:

- max_depth: It represents the maximum depth of the tree. If its value is None, nodes are expanded until all leaves are pure, or until all leaves contain less than min_samples_split samples. The max_depth input should be integer and the default value of max_depth is None.
- splitter: It denotes the strategy to choose the split at each node. It has two values, 'best' and 'random', which means choosing the best split and choosing the best random split respectively. The default value of the splitter is 'best'.

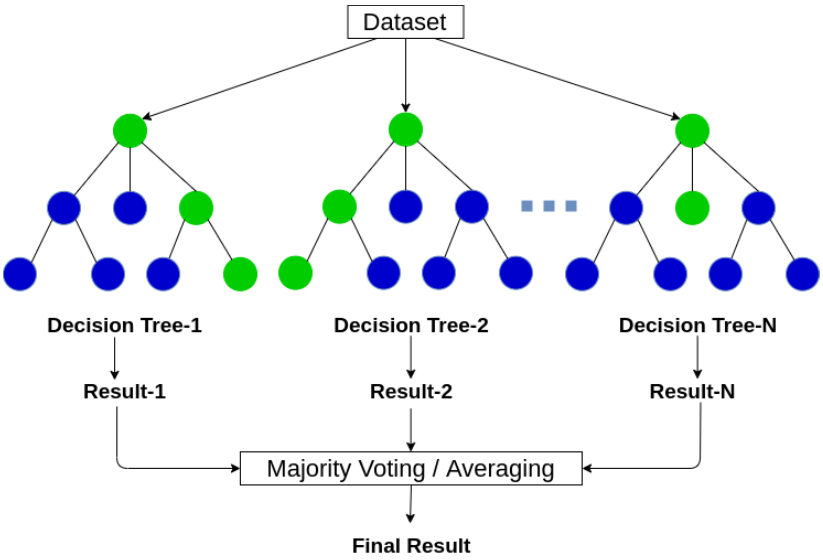
Table 7. *Hyperparameter Tuning of Single Decision Tree*

Model		Parameters			Avg FDR at 3%		
	Iteration	Variables	max_depth	splitter	trn	tst	oot
Single Decision Tree	1	10	None	random	1.000	0.594	0.263
	2	10	5	best	0.707	0.711	0.380
	3	15	10	random	0.725	0.650	0.340
	4	15	10	best	0.839	0.682	0.302
	5	20	7	random	0.666	0.644	0.361
	6	20	7	best	0.751	0.740	0.396

7.3 Random Forest

The random forest is a classification algorithm consisting of many relatively strong and deep decision trees. It combines ensemble techniques and training randomness when building each individual tree to create an uncorrelated forest of trees which predicts by averaging or majority voting and outperforms predictions of any individual tree in the forest. In our case, we use it to classify a record as fraud or not a fraud.

Figure 15. *Illustration of Single Decision Tree*



Hyperparameters:

- n_estimators: It denotes the number of trees in the forest. The input should be an integer and the default value of n_estimator is 100.
- max_depth: It represents the maximum depth of the tree. If the value is None, nodes are expanded until all leaves are pure or until all leaves contain less than min_samples_split samples. The input of max_depth should be an integer and the default value of max_depth is 100.
- min_samples_leaf: It represents the minimum number of samples required at a leaf node. The default value of max_depth is 1.

- min_samples_split: It represents the minimum number of samples required to split an internal node. The default value of max_depth is 2.

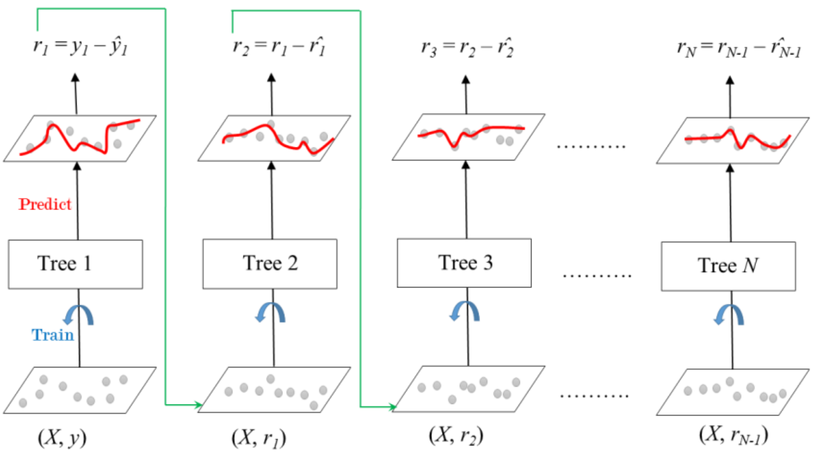
Table 8. *Hyperparameter Tuning of Random Forest*

Model	Parameters						Avg FDR at 3%		
	Iteration	Variables	n_estimators	max_depth	min_samples_leaf	min_samples_split	trn	tst	oot
Random Forest	1	10	10	10	1	2	0.880	0.760	0.513
	2	10	20	10	10	10	0.845	0.810	0.543
	3	15	20	20	20	20	0.929	0.844	0.575
	4	15	50	30	20	20	0.931	0.828	0.591
	5	20	50	50	10	10	0.997	0.854	0.607
	6	20	50	50	20	20	0.937	0.860	0.591
	7	20	50	50	30	50	0.899	0.851	0.598
	8	20	50	50	15	20	0.997	0.865	0.606
	9	20	100	50	15	20	0.941	0.866	0.606
	10	20	200	50	15	30	0.922	0.882	0.610

7.4 Boosted Tree

The boosted tree model is a supervised learning classification and regression model consisting of many relatively weak and shallow trees. These trees are built sequentially to train on the residual errors of the current sum, each adding more correction. Boosting is a way to train a series of weak models to form a strong model. A gradient boosted tree model is built in a stage-wise fashion and allows optimization of an arbitrary differentiable loss function.

Figure 16. *Illustration of Boosted Tree*



Hyperparameters:

- n_estimators: It denotes the number of trees in the forest to fit. The default value is 100.

- max_depth: It denotes the maximum depth of the tree. If the value is None, nodes are expanded until all leaves are pure or until all leaves contain less than min_samples_split samples. When max_depth is less than or equal to 0, it means that the depth has no limit. The default value is -1.
- learning_rate: It represents boosting learning rate. To prevent overfitting the dataset, we can use a smaller learning rate to prevent overfitting and improve model performance. A learning rate in the range of 0.1 to 0.3 usually gives good results. The default learning rate is 0.1.

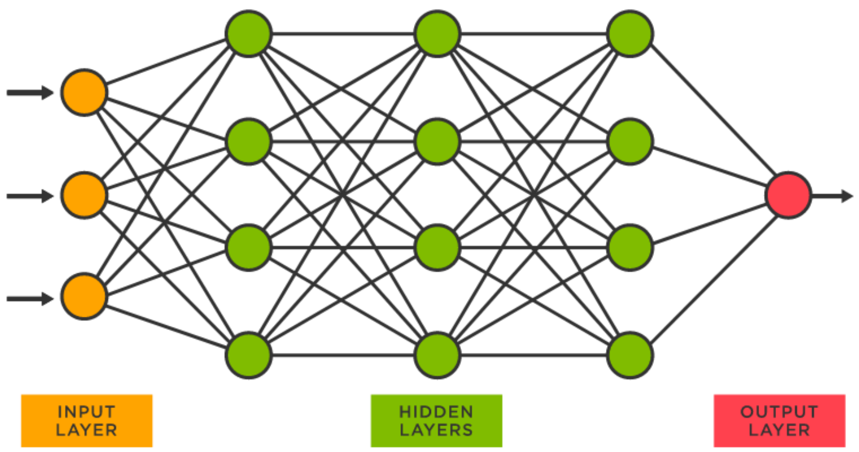
Table 9. *Hyperparameter Tuning of Boosted Tree*

Model		Parameters				Avg FDR at 3%		
	Iteration	Variables	n_ estimators	max_ depth	learning_ rate	trn	tst	oot
Boosted Tree	1	10	10	3	3	0.764	0.755	0.499
	2	10	10	5	5	0.867	0.821	0.466
	3	15	15	5	5	0.969	0.842	0.505
	4	15	15	5	5	0.857	0.807	0.463
	5	20	20	5	5	1.000	0.837	0.418
	6	20	20	5	5	0.908	0.848	0.504

7.5 Neural Network

Neural network algorithm is inspired by the biological neural networks in brains. It consists of an input layer, some hidden layers, and an output layer with nodes resembling the neurons in the brain. Each node transmits signals to nodes in the next layer and the next nodes process signals and decides whether to release signals depending on whether the aggregate level reaches the threshold. Each node receives weighted signals from nodes in previous layers and performs a transformation based on the linear combination of signals received. The algorithm adjusts weights through backpropagation and the records are passed through many times until weights reach local optimum.

Figure 17. *Illustration of Neural Network*



Hyperparameters:

- hidden_layer_sizes: It denotes the ith element represents the number of neurons in the ith hidden layer. The default value is (100,), which means that the model has 1 hidden layer with 100 hidden neurons.
- n_layers_: It represents the number of layers in the neural network model.
- learning_rate: It schedules weight updates with three values, 'constant', 'invscaling' and 'adaptive'. The default value is 'constant'.
- activation: It indicates the activation function for the hidden layer. It has four values, 'identity', 'logistic', 'tanh' and 'relu'. The default value is 'relu'.

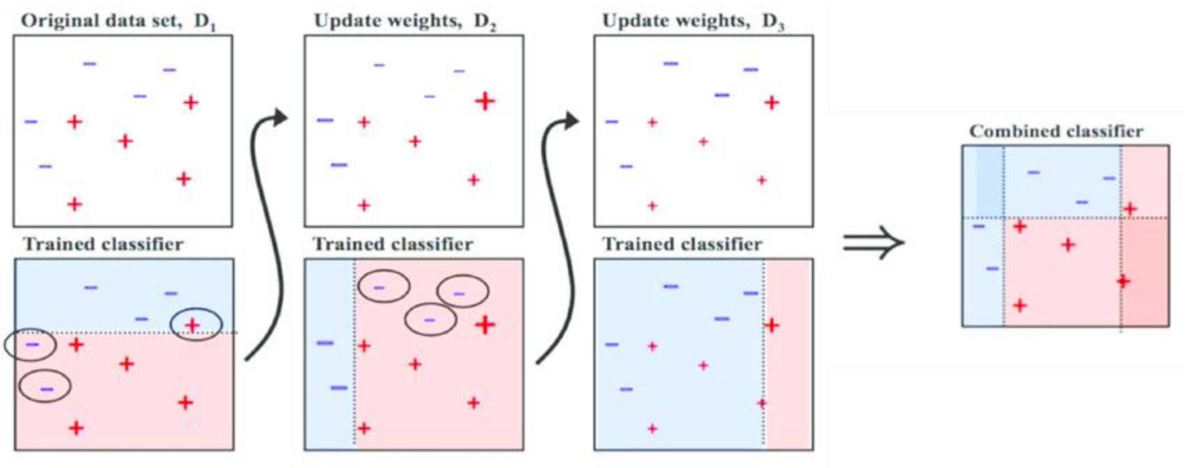
Table 10. *Hyperparameter Tuning of Neural Network*

Model	Parameters						Avg FDR at 3%		
	Iteration	Variables	hidden_ layer_ sizes	n_ layers_	learning_ rate	activation	trn	tst	oot
Neural Network	1	10	5	1	constant	relu	0.714	0.709	0.476
	2	10	5	1	adaptive	logistic	0.717	0.682	0.518
	3	15	10	1	constant	relu	0.754	0.753	0.532
	4	15	10	2	adaptive	logistic	0.701	0.705	0.491
	5	20	20	2	adaptive	relu	0.708	0.720	0.478
	6	20	20	2	constant	logistic	0.715	0.680	0.516

7.6 Adaboost

Adaptive boosting is a boosted algorithm that increases the weight on misclassified records so that the training would focus more on these misclassified records in the next iteration. It is an iterative approach to learn from past mistakes of weak learners and gradually converge to a stronger learner. In each stage, adaptive boosting learns the relative 'hardness' of classifying each sample so that later trees pay more attention to harder examples.

Figure 18. *Illustration of Neural Network*



Hyperparameters:

- n_estimators: It denotes the maximum number of estimators at which boosting is terminated. The default value is 50.
- learning_rate: It represents the weight applied to each classifier at each boosting iteration. A higher learning rate increases the contribution of each classifier. There is a trade-off between n_estimators and learning rate. The higher the learning rate, the smaller the estimators when the algorithm stops. The default value is 1.
- algorithm: It has 2 values, 'SAMME' and 'SAMME.R'. 'SAMME' uses real boosting algorithm while 'SAMME.R' uses discrete boosting algorithm. 'SAMME.R' usually converges faster with fewer iterations.

Table 11. *Hyperparameter Tuning of Adaboost*

Model		Parameters				Avg FDR at 3%		
	Iteration	**Variables**	**n_ estimators**	**learning_ rate**	**algorithm**	**trn**	**tst**	**oot**
Adaboost	1	10	50	1	SAMME.R	0.746	0.757	0.384
	2	10	50	1	SAMME	0.719	0.708	0.407
	3	15	50	0.1	SAMME	0.658	0.664	0.264
	4	15	100	1	SAMME	0.713	0.708	0.374
	5	20	200	1	SAMME	0.781	0.752	0.537
	6	20	300	1	SAMME	0.790	0.767	0.507

Figure 19. *Hyperparameter Tuning of All Models*

Model		Parameters					Avg FDR at 3%		
	Iteration	Variables	penalty		solver		trn	tst	oot
Logistic Regression	1	10	none		lbfgs		0.657	0.681	0.360
	2	10	l2		lbfgs		0.659	0.673	0.362
	3	10	none		saga		0.652	0.677	0.358
	4	10	l2		saga		0.658	0.669	0.361
	5	15	none		lbfgs		0.678	0.674	0.366
	6	15	l2		lbfgs		0.688	0.667	0.373
	7	15	none		saga		0.680	0.697	0.359
	8	15	l2		saga		0.678	0.708	0.380
	9	20	none		saga		0.684	0.666	0.362
	10	20	l2		saga		0.677	0.699	0.377
	11	20	none		lbfgs		0.679	0.679	0.363
	12	20	l2		lbfgs		0.684	0.690	0.380
	Iteration	Variables	max_depth		splitter		Avg FDR at 3%		
Single Decision Tree	1	10	None		random		1.000	0.594	0.263
	2	10	5		best		0.707	0.711	0.380
	3	15	10		random		0.725	0.650	0.340
	4	15	10		best		0.839	0.682	0.302
	5	20	7		random		0.666	0.644	0.361
	6	20	7		best		0.751	0.740	0.396
	Iteration	Variables	n_estimators	max_depth	min_samples_leaf	min_samples_split	Avg FDR at 3%		
Random Forest	1	10	10	10	1	2	0.880	0.760	0.513
	2	10	20	10	10	10	0.845	0.810	0.543
	3	15	20	20	20	20	0.929	0.844	0.575
	4	15	50	30	20	20	0.931	0.828	0.591
	5	20	50	50	10	10	0.997	0.854	0.607
	6	20	50	50	20	20	0.937	0.860	0.591
	7	20	50	50	30	50	0.899	0.851	0.598
	8	20	50	50	15	20	0.997	0.865	0.606
	9	20	100	50	15	30	0.941	0.866	0.606
	10	20	200	50	15	30	0.922	0.882	0.610
	Iteration	Variables	n_estimators	max_depth	learning_rate		Avg FDR at 3%		
Boosted Tree	1	10	20	3	0.1		0.764	0.755	0.499
	2	10	100	5	0.01		0.867	0.821	0.466
	3	15	200	5	0.1		0.969	0.842	0.505
	4	15	200	5	0.01		0.857	0.807	0.463
	5	20	500	5	0.1		1.000	0.837	0.418
	6	20	300	5	0.01		0.908	0.848	0.504
	Iteration	Variables	hidden_layer_sizes	n_layers	learning_rate	activation	Avg FDR at 3%		
Neural Network	1	10	5	1	constant	relu	0.714	0.709	0.476
	2	10	5	1	adaptive	logistic	0.717	0.682	0.518
	3	15	10	1	constant	relu	0.754	0.753	0.532
	4	15	10	2	adaptive	logistic	0.701	0.705	0.491
	5	20	20	2	adaptive	relu	0.708	0.720	0.478
	6	20	20	2	constant	logistic	0.715	0.680	0.516
	Iteration	Variables	n_estimators		learning_rate	algorithm	Avg FDR at 3%		
Adaboost	1	10	50		1	SAMME.R	0.746	0.757	0.384
	2	10	50		1	SAMME	0.719	0.708	0.407
	3	15	50		0.1	SAMME	0.658	0.664	0.264
	4	15	100		1	SAMME	0.713	0.708	0.374
	5	20	200		1	SAMME	0.781	0.752	0.537
	6	20	300		1	SAMME	0.790	0.767	0.507

8. Results

After building 7 different machine learning models, implementing hyperparameter tuning with each model, and comparing model performance on the training dataset, testing dataset, and out-of-time dataset, we found out that the random forest model with 20 variables, 50 estimators, maximum depth is 50, minimum samples of leaf is 30 and minimum samples of split is 50 is our best performer. The model achieves the fraud detection rate of 90.91% on training data, 86.36% on testing data, and 60.34% on out-of-time data in the top 3% population. And based on out of time performance, we plotted fraud savings calculation curves and suggested to set score cut off as 3%. We now take a closer look at the model performance on top 20% records of training data, testing data, and out-of-time data and the plot to suggest score cutoff.

8.1 Training Data

Figure 20. *Boosted Tree Model Performance on Top 20% Records of Training Data*

Training												
# Records			# Goods			# Bads			Fraud Rate			
67440			66736			704			0.010438909			
	Bin Statistics					Cumulative Statistics						
Population Bins %	# Records	# Goods	# Bads	% Goods	% Bads	Total # Records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	674	238	436	35.35%	64.65%	674	238	436	0.36%	61.93%	61.57%	0.55
2	674	526	148	78.05%	21.95%	1349	765	584	1.15%	82.95%	81.81%	1.31
3	674	618	56	91.70%	8.30%	2023	1383	640	2.07%	90.91%	88.84%	2.16
4	674	650	24	96.44%	3.56%	2698	2034	664	3.05%	94.32%	91.27%	3.06
5	674	656	18	97.33%	2.67%	3372	2690	682	4.03%	96.88%	92.84%	3.94
6	674	661	13	98.07%	1.93%	4046	3351	695	5.02%	98.72%	93.70%	4.82
7	674	667	7	98.96%	1.04%	4721	4019	702	6.02%	99.72%	93.69%	5.72
8	674	672	2	99.70%	0.30%	5395	4691	704	7.03%	100.00%	92.97%	6.66
9	674	674	0	100.00%	0.00%	6070	5366	704	8.04%	100.00%	91.96%	7.62
10	674	674	0	100.00%	0.00%	6744	6040	704	9.05%	100.00%	90.95%	8.58
11	674	674	0	100.00%	0.00%	7418	6714	704	10.06%	100.00%	89.94%	9.54
12	674	674	0	100.00%	0.00%	8093	7389	704	11.07%	100.00%	88.93%	10.50
13	674	674	0	100.00%	0.00%	8767	8063	704	12.08%	100.00%	87.92%	11.45
14	674	674	0	100.00%	0.00%	9442	8738	704	13.09%	100.00%	86.91%	12.41
15	674	674	0	100.00%	0.00%	10116	9412	704	14.10%	100.00%	85.90%	13.37
16	674	674	0	100.00%	0.00%	10790	10086	704	15.11%	100.00%	84.89%	14.33
17	674	674	0	100.00%	0.00%	11465	10761	704	16.12%	100.00%	83.88%	15.29
18	674	674	0	100.00%	0.00%	12139	11435	704	17.13%	100.00%	82.87%	16.24
19	674	674	0	100.00%	0.00%	12814	12110	704	18.15%	100.00%	81.85%	17.20
20	674	674	0	100.00%	0.00%	13488	12784	704	19.16%	100.00%	80.84%	18.16

8.2 Testing Data

Figure 21. *Boosted Tree Model Performance on Top 20% Records of Testing Data*

Testing												
# Records			# Goods			# Bads			Fraud Rate			
16860			16684			176			0.010438909			
	Bin Statistics					Cumulative Statistics						
Population Bins %	# Records	# Goods	# Bads	% Goods	% Bads	Total # Records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR
1	169	60	109	35.35%	64.65%	169	60	109	0.36%	61.93%	61.57%	0.55
2	169	138	31	81.61%	18.39%	337	197	140	1.18%	79.55%	78.36%	1.41
3	169	157	12	92.88%	7.12%	506	354	152	2.12%	86.36%	84.24%	2.33
4	169	166	3	98.22%	1.78%	674	519	155	3.11%	88.07%	84.96%	3.35
5	169	168	1	99.41%	0.59%	843	687	156	4.12%	88.64%	84.52%	4.40
6	169	167	2	98.81%	1.19%	1012	854	158	5.12%	89.77%	84.66%	5.40
7	169	168	1	99.41%	0.59%	1180	1021	159	6.12%	90.34%	84.22%	6.42
8	169	168	1	99.41%	0.59%	1349	1189	160	7.13%	90.91%	83.78%	7.43
9	169	169	0	100.00%	0.00%	1517	1357	160	8.14%	90.91%	82.77%	8.48
10	169	168	1	99.41%	0.59%	1686	1525	161	9.14%	91.48%	82.34%	9.47
11	169	168	1	99.41%	0.59%	1855	1693	162	10.15%	92.05%	81.90%	10.45
12	169	169	0	100.00%	0.00%	2023	1861	162	11.16%	92.05%	80.89%	11.49
13	169	169	0	100.00%	0.00%	2192	2030	162	12.17%	92.05%	79.88%	12.53
14	169	169	0	100.00%	0.00%	2360	2198	162	13.18%	92.05%	78.87%	13.57
15	169	168	1	99.41%	0.59%	2529	2366	163	14.18%	92.61%	78.43%	14.52
16	169	167	2	98.81%	1.19%	2698	2533	165	15.18%	93.75%	78.57%	15.35
17	169	168	1	99.41%	0.59%	2866	2700	166	16.18%	94.32%	78.13%	16.27
18	169	169	0	100.00%	0.00%	3035	2869	166	17.19%	94.32%	77.12%	17.28
19	169	169	0	100.00%	0.00%	3203	3037	166	18.21%	94.32%	76.11%	18.30
20	169	169	0	100.00%	0.00%	3372	3206	166	19.22%	94.32%	75.10%	19.31

8.3 Out of Time Data

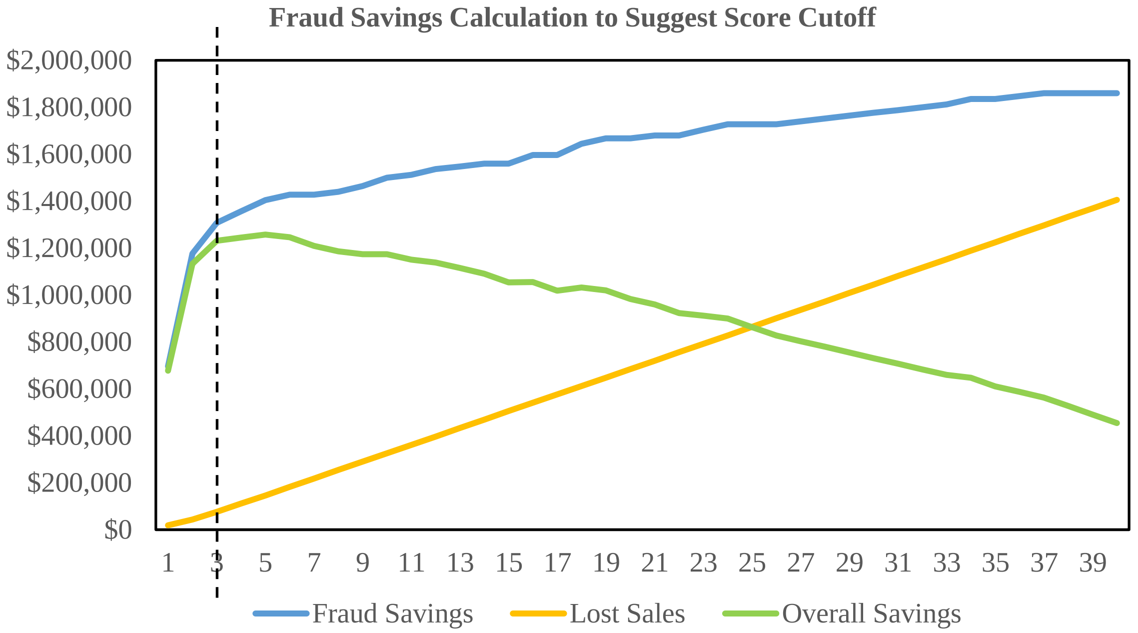
Figure 22. *Boosted Tree Model Performance on Top 20% Records of Out of Time Data*

OOT														
# Records			# Goods			# Bads			Fraud Rate					
12097			11918			179			0.014797057					
	Bin Statistics					Cumulative Statistics								
Population Bins %	# Records	# Goods	# Bads	% Goods	% Bads	Total # Records	Cumulative Goods	Cumulative Bads	% Goods	% Bads (FDR)	KS	FPR		
1	121	60	61	49.57%	50.43%	121	60	61	0.50%	34.08%	33.58%	0.98		
2	121	83	38	68.59%	31.41%	242	143	99	1.20%	55.31%	54.11%	1.44		
3	121	112	9	92.56%	7.44%	363	255	108	2.14%	60.34%	58.20%	2.36		
4	121	113	8	93.39%	6.61%	484	368	116	3.09%	64.80%	61.72%	3.17		
5	121	118	3	97.52%	2.48%	605	486	119	4.08%	66.48%	62.40%	4.08		
6	121	120	1	99.17%	0.83%	726	606	120	5.08%	67.04%	61.96%	5.05		
7	121	120	1	99.17%	0.83%	847	726	121	6.09%	67.60%	61.51%	6.00		
8	121	121	0	100.00%	0.00%	968	847	121	7.10%	67.60%	60.49%	7.00		
9	121	121	0	100.00%	0.00%	1089	968	121	8.12%	67.60%	59.48%	8.00		
10	121	119	2	98.35%	1.65%	1210	1087	123	9.12%	68.72%	59.60%	8.83		
11	121	118	3	97.52%	2.48%	1331	1205	126	10.11%	70.39%	60.28%	9.56		
12	121	121	0	100.00%	0.00%	1452	1326	126	11.12%	70.39%	59.27%	10.52		
13	121	119	2	98.35%	1.65%	1573	1445	128	12.12%	71.51%	59.39%	11.29		
14	121	120	1	99.17%	0.83%	1694	1565	129	13.13%	72.07%	58.94%	12.13		
15	121	119	2	98.35%	1.65%	1815	1684	131	14.13%	73.18%	59.06%	12.85		
16	121	121	0	100.00%	0.00%	1936	1805	131	15.14%	73.18%	58.04%	13.77		
17	121	118	3	97.52%	2.48%	2056	1922	134	16.13%	74.86%	58.73%	14.35		
18	121	121	0	100.00%	0.00%	2177	2043	134	17.15%	74.86%	57.71%	15.25		
19	121	121	0	100.00%	0.00%	2298	2164	134	18.16%	74.86%	56.70%	16.15		
20	121	120	1	99.17%	0.83%	2419	2284	135	19.17%	75.42%	56.25%	16.92		

8.4 Suggestion for Cutoff

We calculated the Fraud Savings (blue), Lost Sales (orange), and Overall Savings (green) by assuming $2,000 gain for every fraud caught and $50 loss for every false positive case, which means a good that is flagged as a bad. Additionally, we recommended that the client should set 3% as a score cutoff.

Figure 23. *Fraud Savings Calculation to Suggest Score Cutoff as 3%*

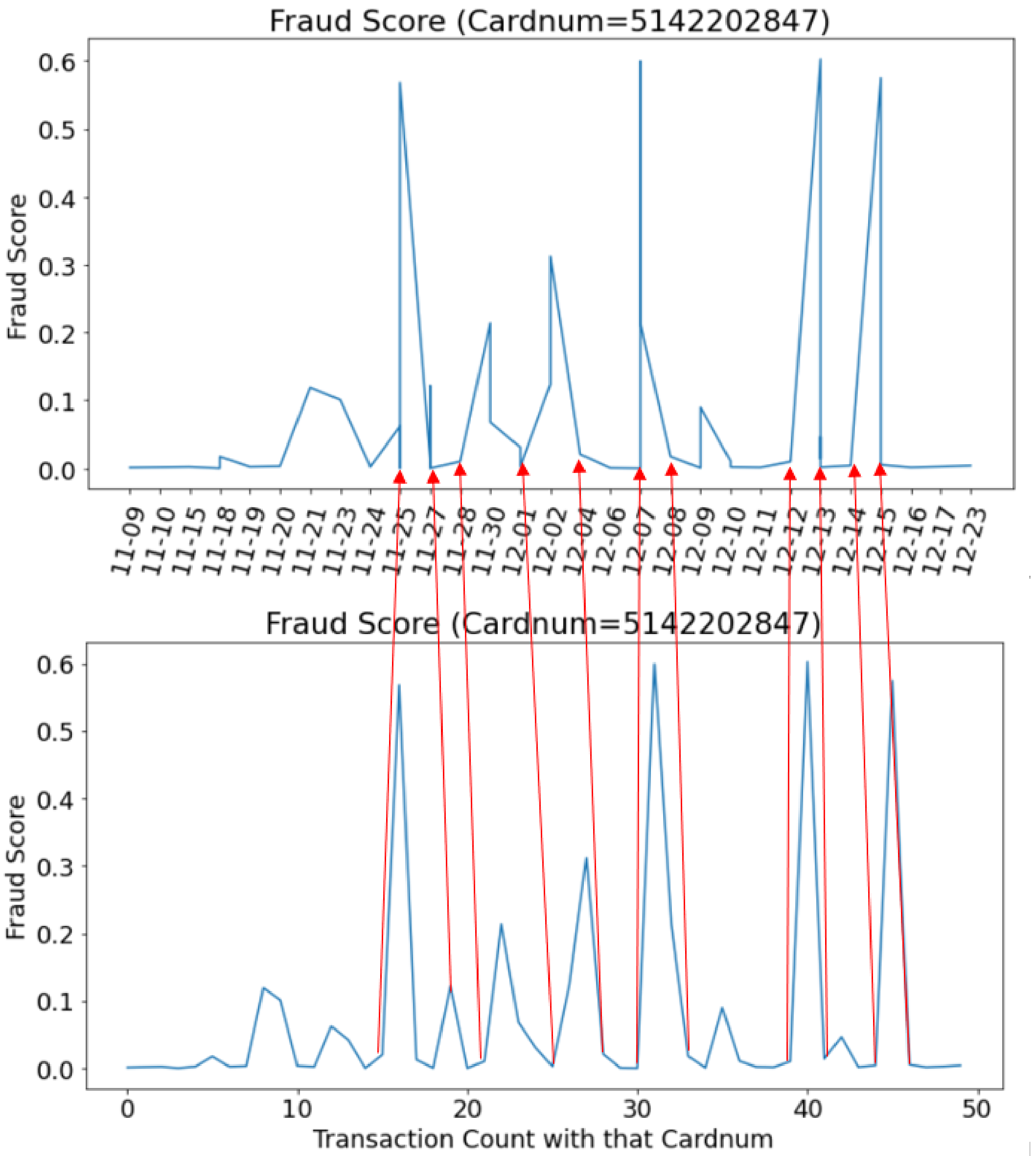


8.5 Time Dependency of Fraud Scores

We looked into a specific card (Cardnum=5142202847) and a merchant (Merchnum=4353000719908) to see their changes in fraud score over time.

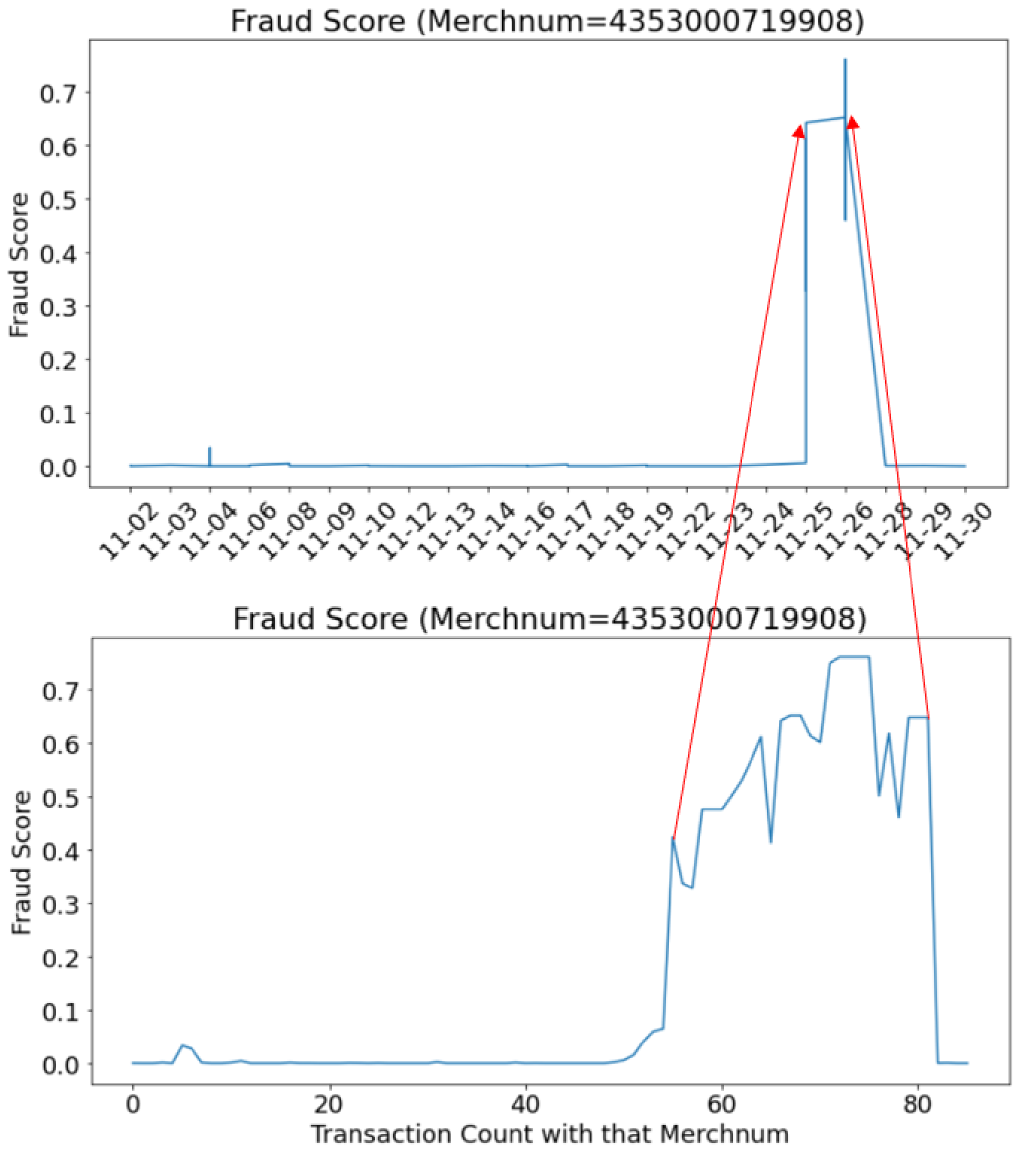
For the card (Cardnum=5142202847), 50 transactions happened in November and December and they spread across 2 months. 5 transactions happened on 11/25, 4 transactions happened on 11/27, 3 transactions happened on 12/7, 4 transactions happened on 12/13 and 2 transactions happened on 12/15. These transactions caused 4 rapid rises in fraud score. Small bunch of transactions also continuously happened, and they caused a slight rise in fraud score between these 4 humps.

Figure 24. *Fraud Scores for Cardnum = 5142202847*



For the Merchant (Merchnum=4353000719908), 86 transactions happened in November. 17 transactions happened on 11/25 and 15 transactions happened on 11/26. Most transactions happened in these two days, so the fraud score grows rapidly when transactions occurred in these two days.

Figure 25. *Fraud Scores for Merchnum = 4353000719908*



9. Conclusion

Credit card transaction fraud is a serious threat for the banking system. According to Nilson Report, the amount of payment fraud losses has grown over three times since 2011. Transaction frauds not only harm the revenue of banks, but also lead to the inefficiency of social economics. Therefore, it is necessary to develop a robust statistical model to detect fraud and protect the banks.

In this project, we investigated the transaction dataset in 2006 which contains 96,753 transaction records and 10 fields. Among them, 1,059 records of transactions are labeled as fraud. After fixing the missing values and removing the irrelevant transactions and outliers, we created 1,108 variables based on our domain knowledge. We applied filter and wrapper methods to select the top 20 most relevant variables. In the model constructing process, we kept the data in the last two months of 2006 as out-of-time (OOT) data and randomly split 80% of remaining data as the training data and 20% as the testing data. We used 6 different models including logistic regression, decision tree, random forest, boosted tree, neural network, and adaboost and tuned the models with different hyperparameters to find the best performer.

By constructing a supervised model based on a synthetic dataset of credit card transactions, we successfully completed several important stages of fraud analytics including data preparation, feature engineering, feature selection, model building, and business interpretation. For each step, there is still room for improvements for us to better build our model for out-of-time data prediction. In the data generating process, we may have a greater scope to detect frauds if we can have more demographic information, such as gender and age, and geographical information, such as transaction locations. It will also be helpful if we can gather information about detailed transaction time. We also realized that the transaction data was outdated because it was collected in 2006. To better detect future transaction frauds, we need to collect more recent data to understand changing fraud patterns. In the data preparation process, since only 1% of the data are frauds, we can use oversampling techniques such as SMOTE to increase fraud data to see if it works better. We also noticed that the number of daily transactions of the last two months is lower than in the previous ten months, which may affect our evaluation accuracy. In feature engineering stage, we can try to create more new entities and variables associated with card number since most of our effective variables are built upon card number in this dataset. Finally, although we have used FDR at top 3% as our main evaluation metric for this project, we can explore more metrics to adapt to different business objectives.

10. Appendix

10.1 Data Quality Report on Transaction Data

Part 1. File Description

This dataset contains credit card transaction information from a U.S. governmental organization and synthetic fraud labels. The dataset includes 96,753 records of transactions and 10 fields. It covers records from 1/1/2006 to 12/31/2006. This dataset includes 1,059 transaction frauds, accounting for 1.09% of all records.

Part 2. Summary statistics table

Table 12. *Numeric fields summary (same as **Table 1**)*

Field Name	% Populated	Min	Max	Mean	Stdev	*% Zero
Date	100	2006-01-01	2006-12-31	-	-	0
Amount	100	0.01	3,102,045.53	427.89	10,006.14	0

*% Zero: only including record whose value is 0

Table 13. *Categorical fields summary (same as **Table 2**)*

Field Name	% Populated	*Unique Values	Most Common Value
Recnum	100	96,753	-
Cardnum	100	1,645	5142148452
Merchnum	96.51	13,091	930090121224
Merch description	100	13,126	GSA-FSS-ADV
Merch state	98.76	227	TN
Merch zip*	95.19	4,567	38118
Transtype	100	4	P
Fraud	100	2	0

*Unique Values: does not include Nan
*Transformed Merch zip from float as int

Part 3. Fields Distribution

1. Date

Figure 26. *Daily Transactions (same as* ***Figure 4****)*

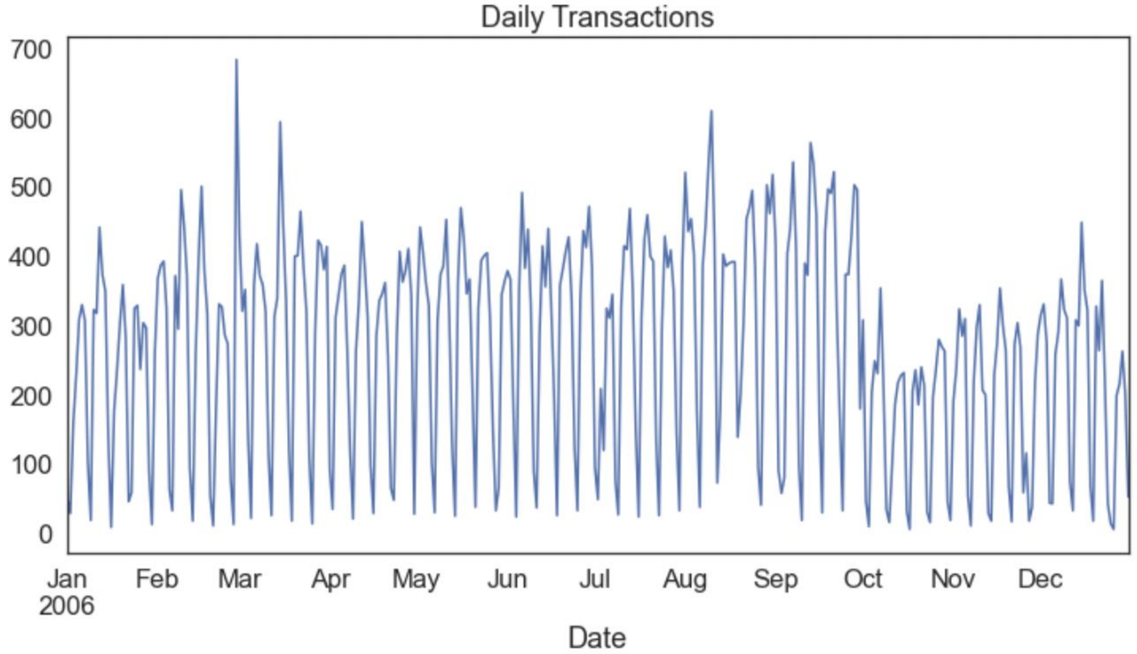
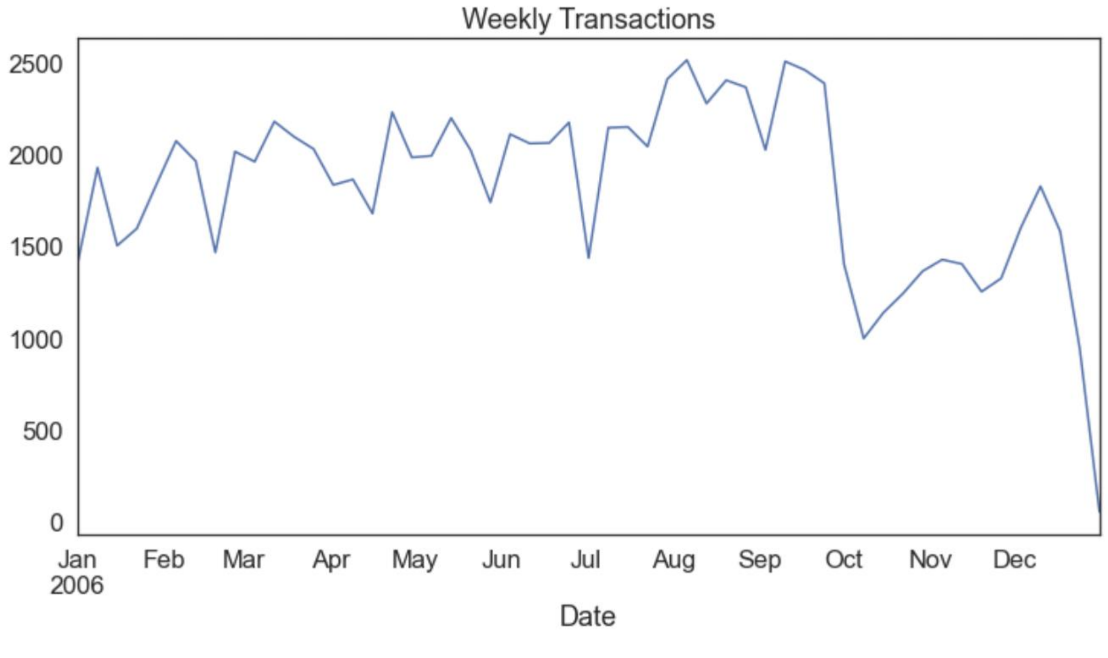
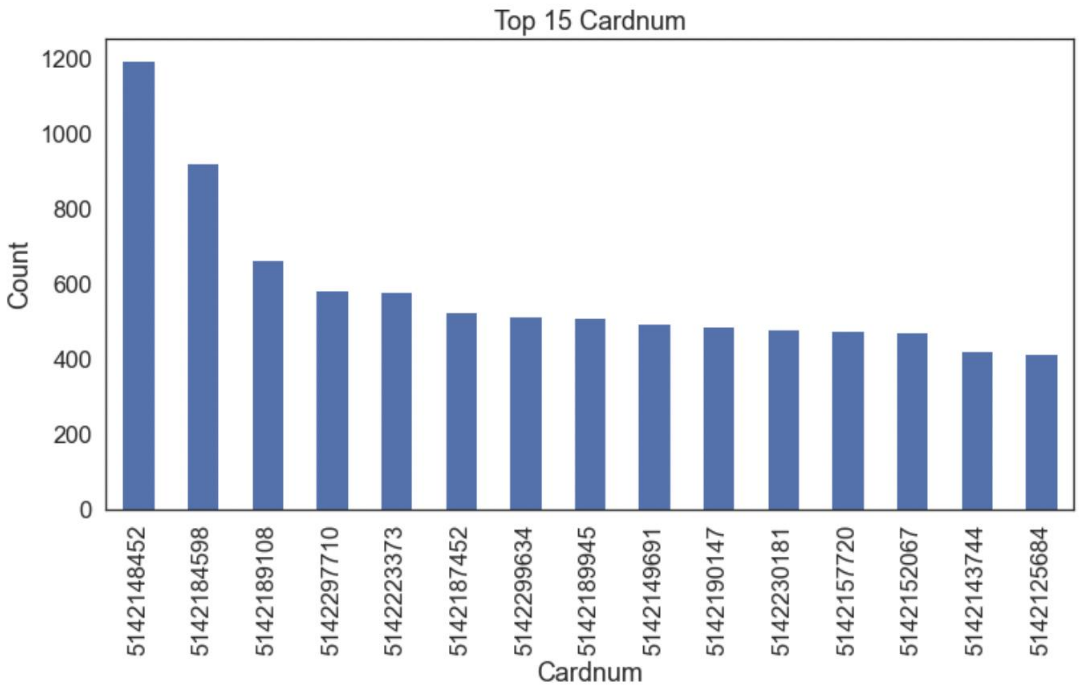


Figure 27: *Weekly Transactions*



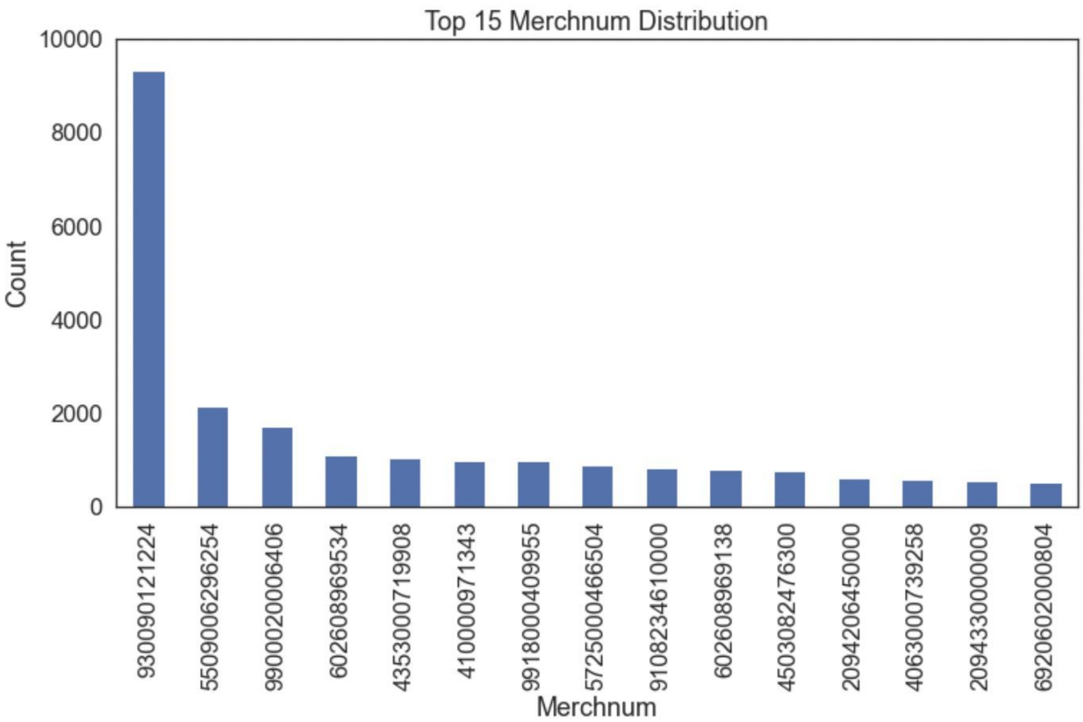
2. Cardnum

Figure 28. *Distribution of Cardnum*



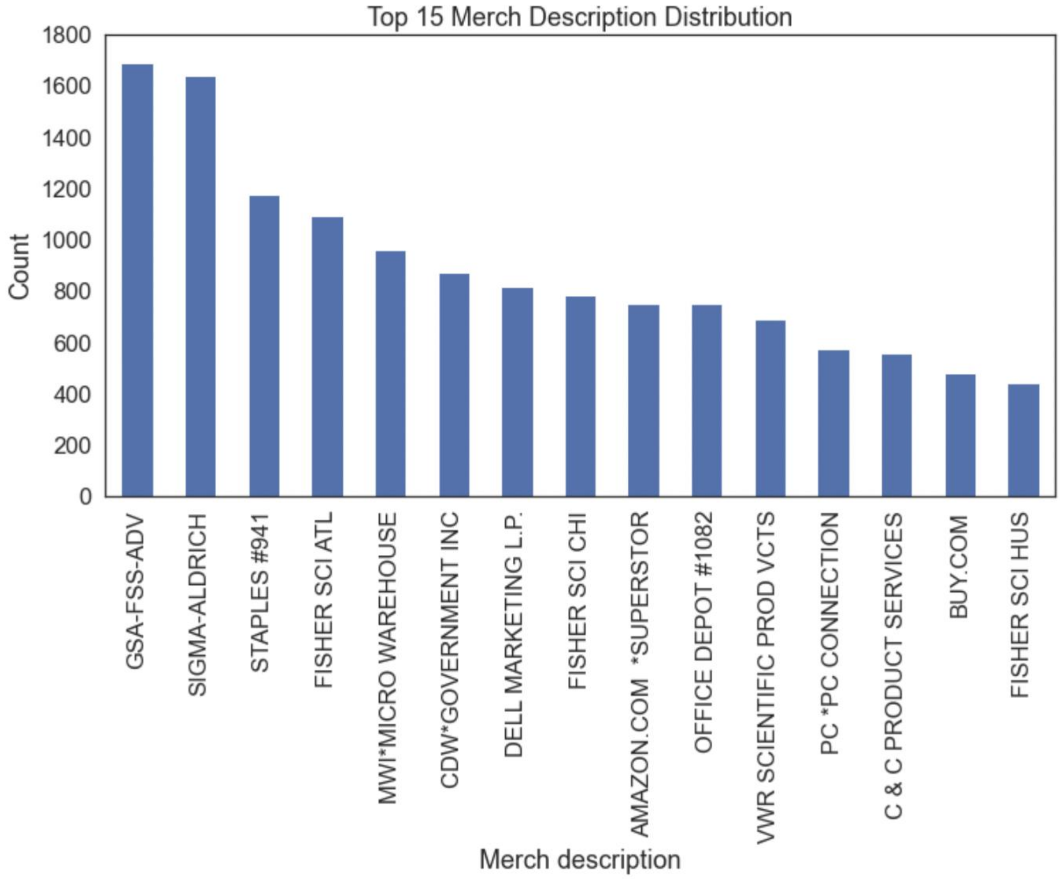
3. Merchnum

Figure 29. *Distribution of Merchnum. It doesn't include null value (nan). (same as* ***Figure 6****)*



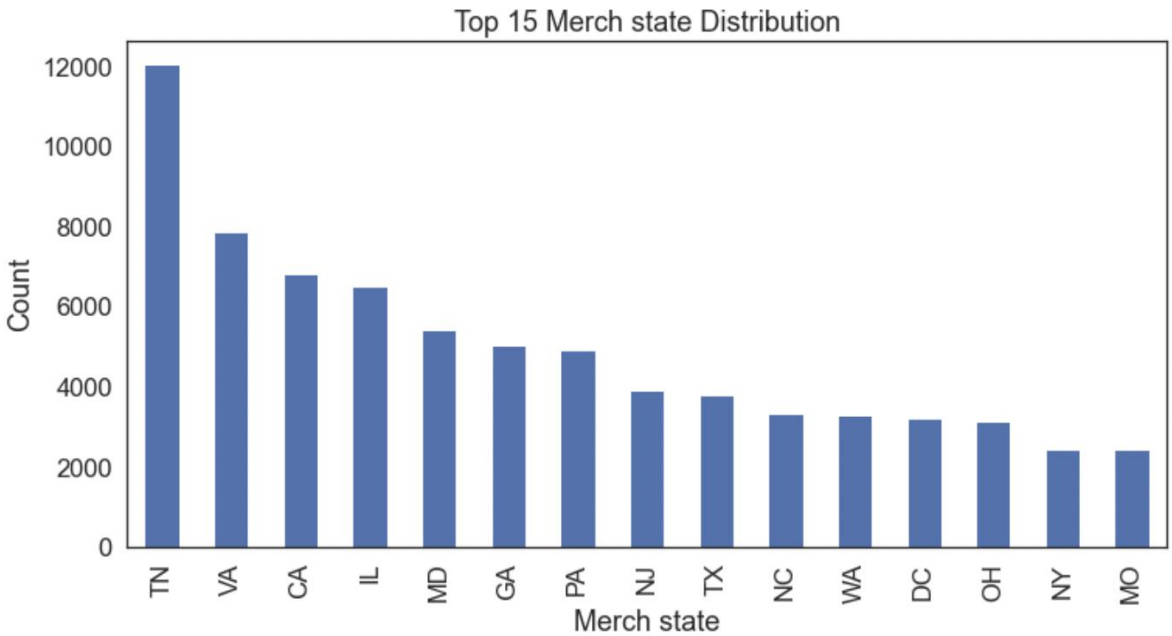
4. Merch description

Figure 30. *Distribution of Merch description*



5. Merch state

Figure 31. *Distribution of Merch state*



6. Merch zip

Figure 32. *Distribution of Merch zip. The value of y-axis is Count. (same as* ***Figure 7****)*

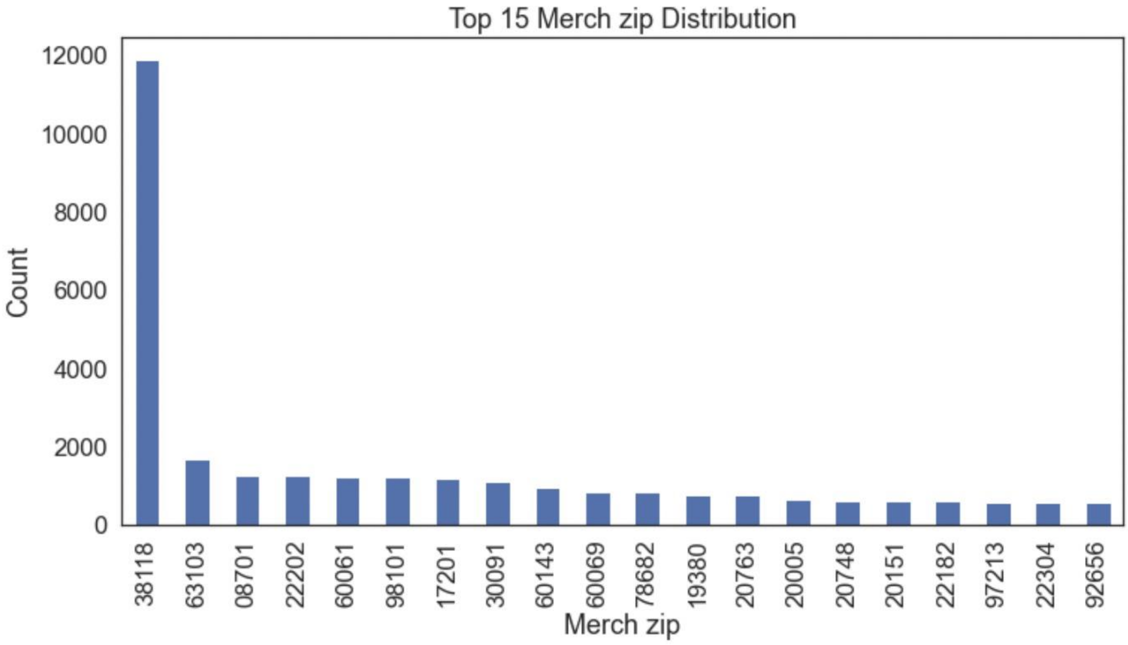
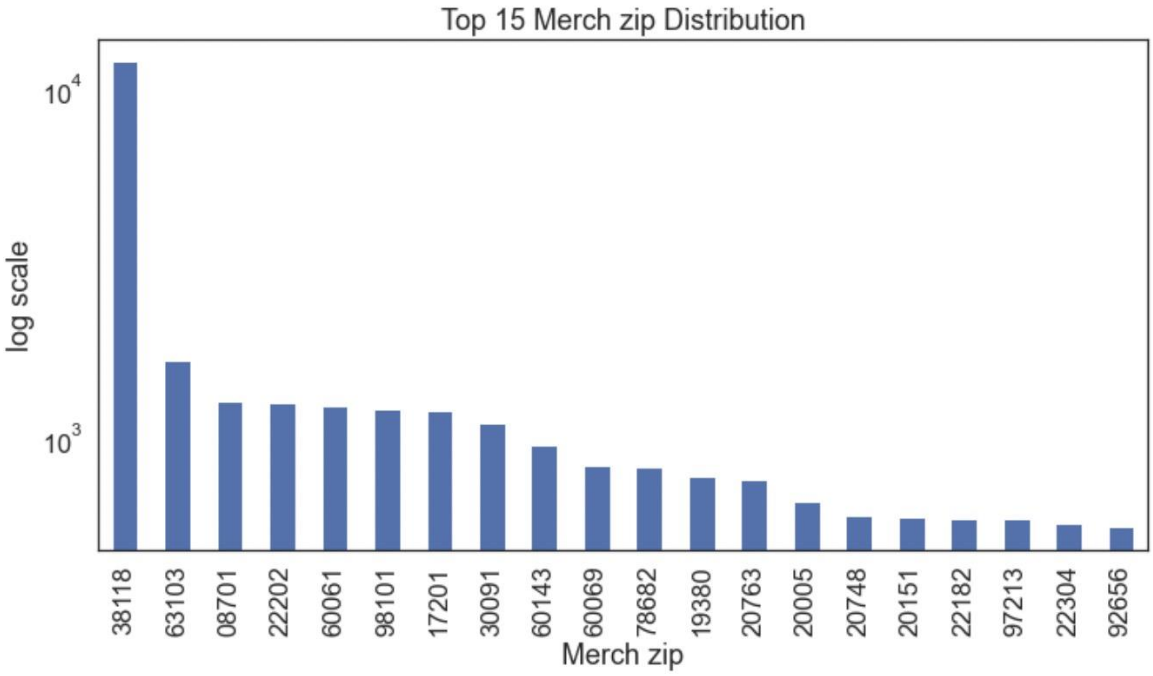
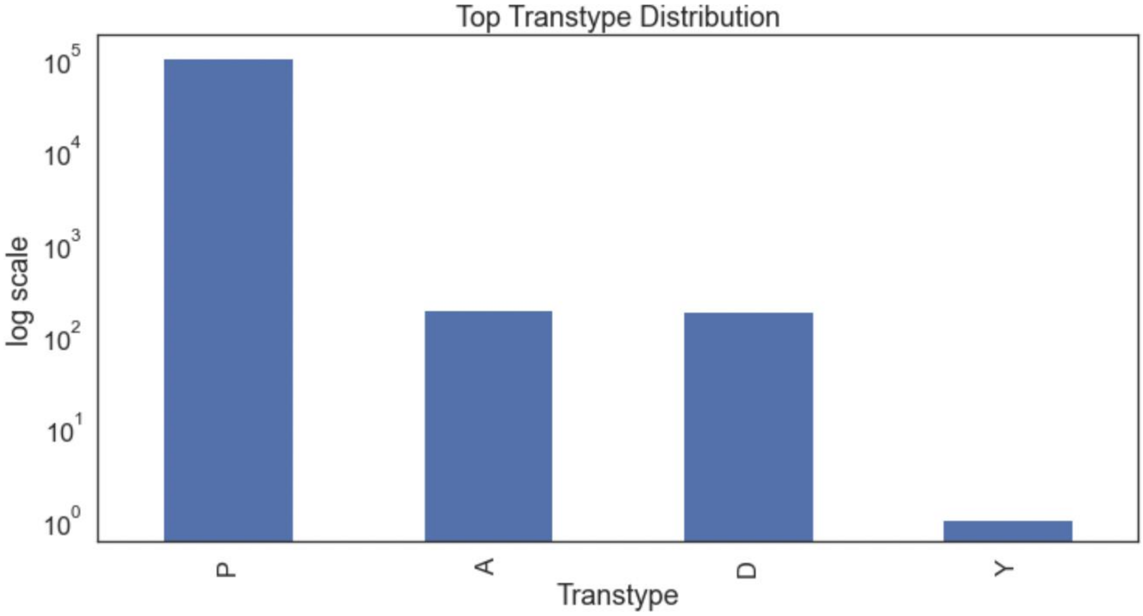


Figure 33. *Distribution of Merch zip. The value of y-axis is log(count).*



7. Transtype

Figure 34. *Distribution of Transtype. The value of y-axis is log(count).*



8. Amount

Figure 35. *Distribution of Amount. (Count, it covers 99.50% of all values)*

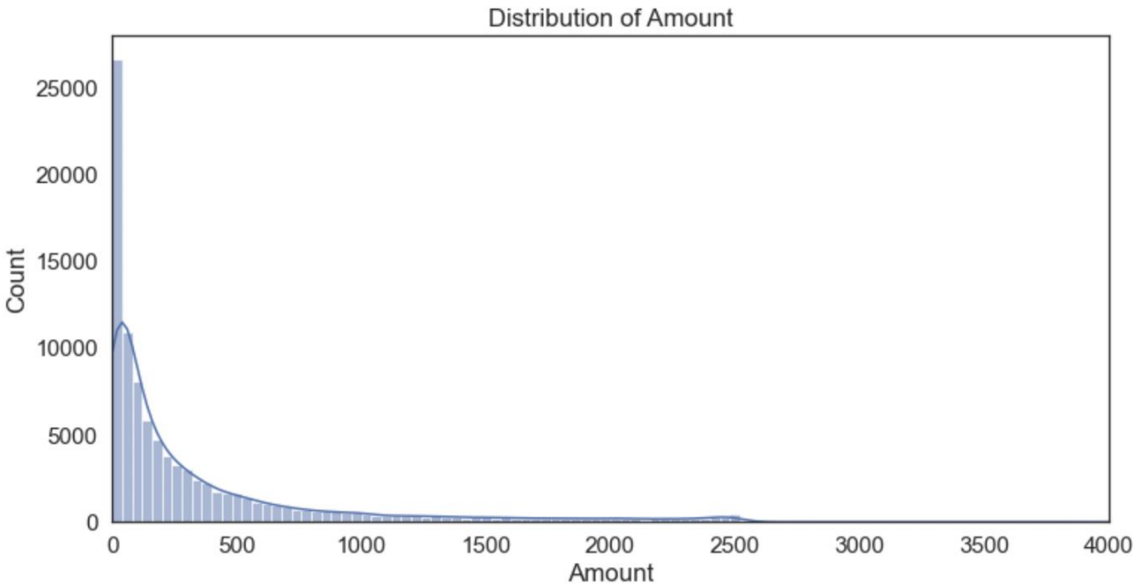
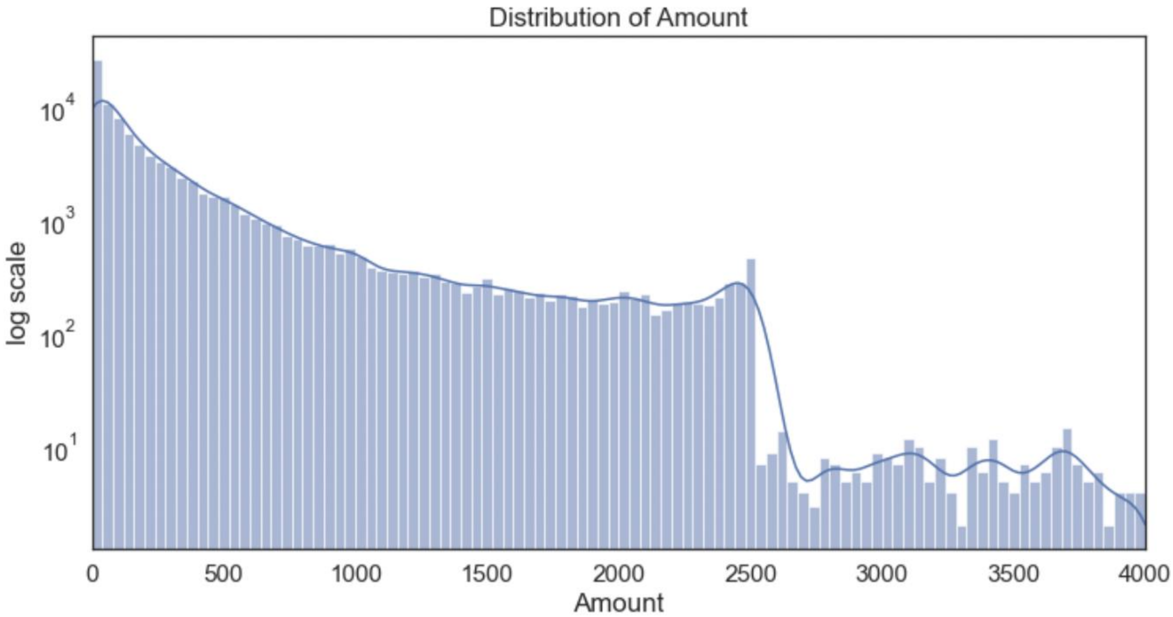


Figure 36. *Distribution of Amount. (Log(count), it covers 99.50% of all values) (same as* ***Figure 5****)*



9. Fraud

Figure 37. *Distribution of Fraud. (Fraud_0 : Fraud_1 = 95,694 : 1,059)*

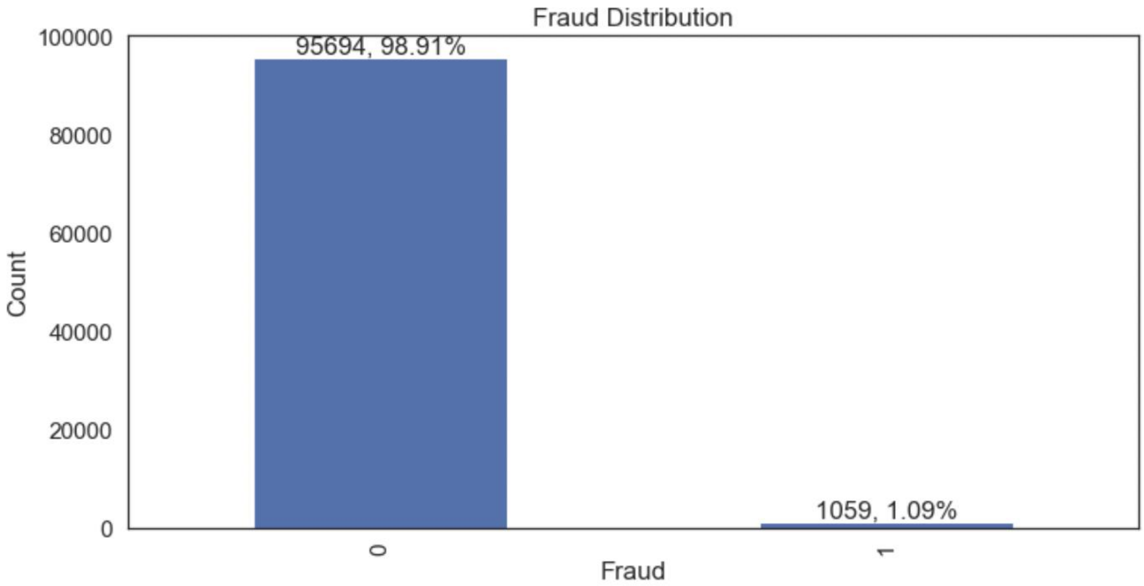


Figure 38. *Bad Transactions Weekly Proportion Distribution*

Bad (red): Fraud=1

Proportion: (weekly count of bad transactions) / (weekly count of total transactions)

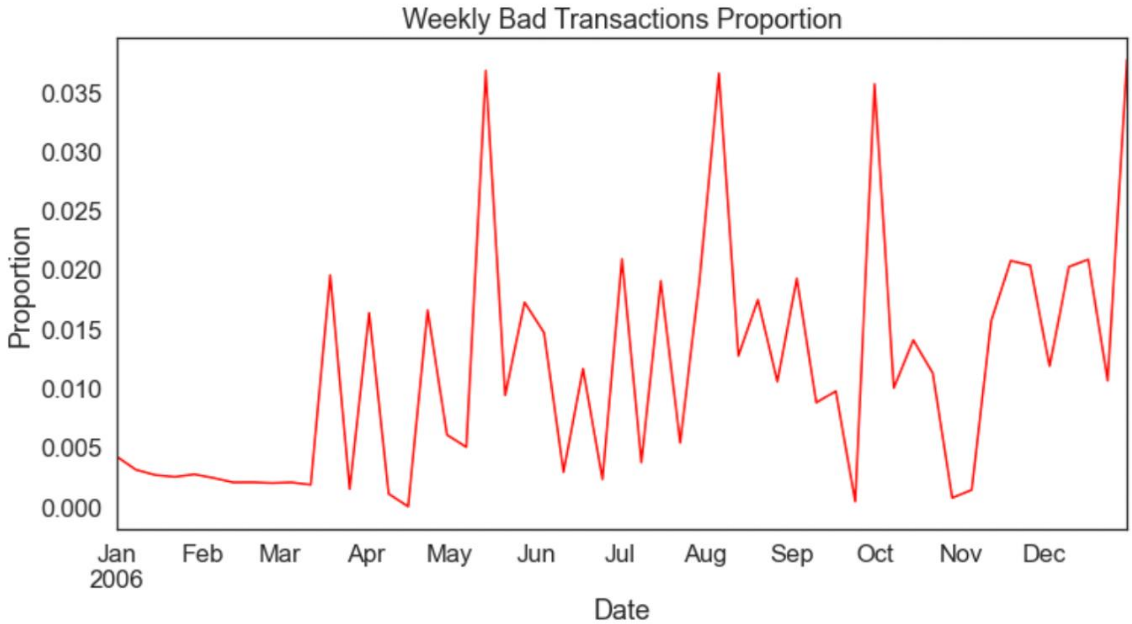
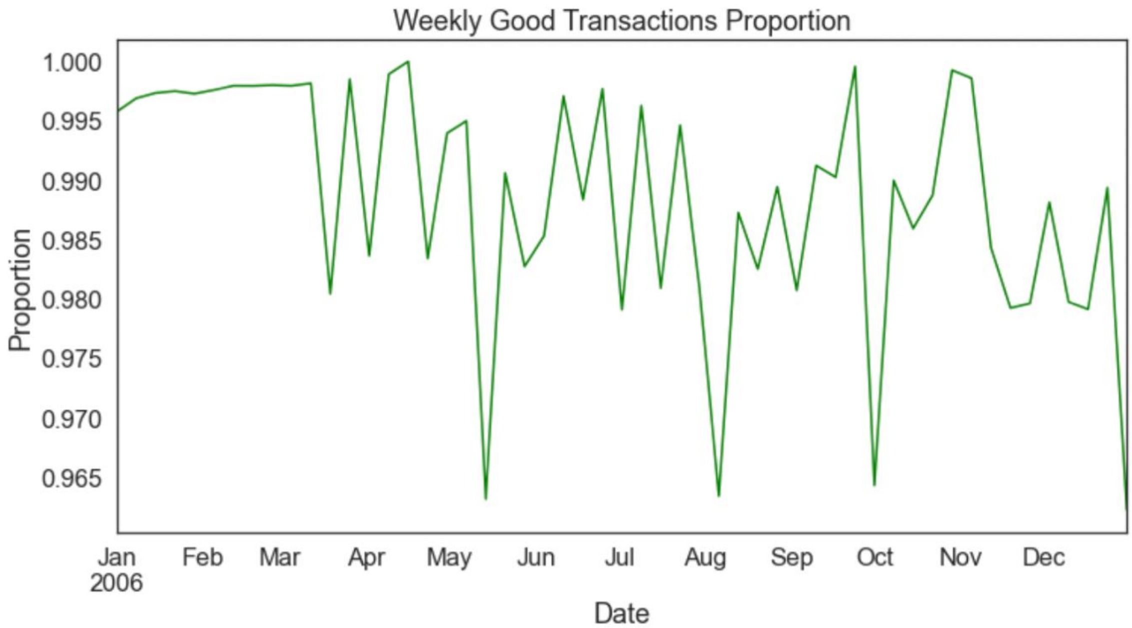


Figure 39. *Good Transactions Weekly Proportion Distribution*

Good (green): Fraud=0

Proportion: (weekly count of good transactions) / (weekly count of total transactions)



10.2 Variable List

Table 14. *List of variables created*

Cardnum_day_since	Merch state_med_60	Merchnum_fulladdress _actual/max_14	Cardnum_Merch state_count_7
Cardnum_count_0	Merch state_total_60	Merchnum_fulladdress _actual/med_14	Cardnum_Merch state_avg_7
Cardnum_avg_0	Merch state_actual/avg_60	Merchnum_fulladdress _actual/toal_14	Cardnum_Merch state_max_7
Cardnum_max_0	Merch state_actual/max_60	Merchnum_fulladdress _count_30	Cardnum_Merch state_med_7
Cardnum_med_0	Merch state_actual/med_60	Merchnum_fulladdress _avg_30	Cardnum_Merch state_total_7
Cardnum_total_0	Merch state_actual/toal_60	Merchnum_fulladdress _max_30	Cardnum_Merch state_actual/avg_7
Cardnum_actual/avg_0	Merch state_count_90	Merchnum_fulladdress _med_30	Cardnum_Merch state_actual/max_7
Cardnum_actual/max_0	Merch state_avg_90	Merchnum_fulladdress _total_30	Cardnum_Merch state_actual/med_7
Cardnum_actual/med_0	Merch state_max_90	Merchnum_fulladdress _actual/avg_30	Cardnum_Merch state_actual/toal_7
Cardnum_actual/toal_0	Merch state_med_90	Merchnum_fulladdress _actual/max_30	Cardnum_Merch state_count_14
Cardnum_count_1	Merch state_total_90	Merchnum_fulladdress _actual/med_30	Cardnum_Merch state_avg_14
Cardnum_avg_1	Merch state_actual/avg_90	Merchnum_fulladdress _actual/toal_30	Cardnum_Merch state_max_14
Cardnum_max_1	Merch state_actual/max_90	Merchnum_fulladdress _count_60	Cardnum_Merch state_med_14
Cardnum_med_1	Merch state_actual/med_90	Merchnum_fulladdress _avg_60	Cardnum_Merch state_total_14
Cardnum_total_1	Merch state_actual/toal_90	Merchnum_fulladdress _max_60	Cardnum_Merch state_actual/avg_14
Cardnum_actual/avg_1	Merch zip_day_since	Merchnum_fulladdress _med_60	Cardnum_Merch state_actual/max_14
Cardnum_actual/max_1	Merch zip_count_0	Merchnum_fulladdress _total_60	Cardnum_Merch state_actual/med_14
Cardnum_actual/med_1	Merch zip_avg_0	Merchnum_fulladdress _actual/avg_60	Cardnum_Merch state_actual/toal_14

Cardnum_actual/toal_1	Merch zip_max_0	Merchnum_fulladdress_actual/max_60	Cardnum_Merch state_count_30
Cardnum_count_3	Merch zip_med_0	Merchnum_fulladdress_actual/med_60	Cardnum_Merch state_avg_30
Cardnum_avg_3	Merch zip_total_0	Merchnum_fulladdress_actual/toal_60	Cardnum_Merch state_max_30
Cardnum_max_3	Merch zip_actual/avg_0	Merchnum_fulladdress_count_90	Cardnum_Merch state_med_30
Cardnum_med_3	Merch zip_actual/max_0	Merchnum_fulladdress_avg_90	Cardnum_Merch state_total_30
Cardnum_total_3	Merch zip_actual/med_0	Merchnum_fulladdress_max_90	Cardnum_Merch state_actual/avg_30
Cardnum_actual/avg_3	Merch zip_actual/toal_0	Merchnum_fulladdress_med_90	Cardnum_Merch state_actual/max_30
Cardnum_actual/max_3	Merch zip_count_1	Merchnum_fulladdress_total_90	Cardnum_Merch state_actual/med_30
Cardnum_actual/med_3	Merch zip_avg_1	Merchnum_fulladdress_actual/avg_90	Cardnum_Merch state_actual/toal_30
Cardnum_actual/toal_3	Merch zip_max_1	Merchnum_fulladdress_actual/max_90	Cardnum_Merch state_count_60
Cardnum_count_7	Merch zip_med_1	Merchnum_fulladdress_actual/med_90	Cardnum_Merch state_avg_60
Cardnum_avg_7	Merch zip_total_1	Merchnum_fulladdress_actual/toal_90	Cardnum_Merch state_max_60
Cardnum_max_7	Merch zip_actual/avg_1	Cardnum_fulladdress_day_since	Cardnum_Merch state_med_60
Cardnum_med_7	Merch zip_actual/max_1	Cardnum_fulladdress_count_0	Cardnum_Merch state_total_60
Cardnum_total_7	Merch zip_actual/med_1	Cardnum_fulladdress_avg_0	Cardnum_Merch state_actual/avg_60
Cardnum_actual/avg_7	Merch zip_actual/toal_1	Cardnum_fulladdress_max_0	Cardnum_Merch state_actual/max_60
Cardnum_actual/max_7	Merch zip_count_3	Cardnum_fulladdress_med_0	Cardnum_Merch state_actual/med_60
Cardnum_actual/med_7	Merch zip_avg_3	Cardnum_fulladdress_total_0	Cardnum_Merch state_actual/toal_60
Cardnum_actual/toal_7	Merch zip_max_3	Cardnum_fulladdress_actual/avg_0	Cardnum_Merch state_count_90

Cardnum_count_14	Merch zip_med_3	Cardnum_fulladdress_actual/max_0	Cardnum_Merch state_avg_90
Cardnum_avg_14	Merch zip_total_3	Cardnum_fulladdress_actual/med_0	Cardnum_Merch state_max_90
Cardnum_max_14	Merch zip_actual/avg_3	Cardnum_fulladdress_actual/toal_0	Cardnum_Merch state_med_90
Cardnum_med_14	Merch zip_actual/max_3	Cardnum_fulladdress_count_1	Cardnum_Merch state_total_90
Cardnum_total_14	Merch zip_actual/med_3	Cardnum_fulladdress_avg_1	Cardnum_Merch state_actual/avg_90
Cardnum_actual/avg_14	Merch zip_actual/toal_3	Cardnum_fulladdress_max_1	Cardnum_Merch state_actual/max_90
Cardnum_actual/max_14	Merch zip_count_7	Cardnum_fulladdress_med_1	Cardnum_Merch state_actual/med_90
Cardnum_actual/med_14	Merch zip_avg_7	Cardnum_fulladdress_total_1	Cardnum_Merch state_actual/toal_90
Cardnum_actual/toal_14	Merch zip_max_7	Cardnum_fulladdress_actual/avg_1	Cardnum_Merch zip_day_since
Cardnum_count_30	Merch zip_med_7	Cardnum_fulladdress_actual/max_1	Cardnum_Merch zip_count_0
Cardnum_avg_30	Merch zip_total_7	Cardnum_fulladdress_actual/med_1	Cardnum_Merch zip_avg_0
Cardnum_max_30	Merch zip_actual/avg_7	Cardnum_fulladdress_actual/toal_1	Cardnum_Merch zip_max_0
Cardnum_med_30	Merch zip_actual/max_7	Cardnum_fulladdress_count_3	Cardnum_Merch zip_med_0
Cardnum_total_30	Merch zip_actual/med_7	Cardnum_fulladdress_avg_3	Cardnum_Merch zip_total_0
Cardnum_actual/avg_30	Merch zip_actual/toal_7	Cardnum_fulladdress_max_3	Cardnum_Merch zip_actual/avg_0
Cardnum_actual/max_30	Merch zip_count_14	Cardnum_fulladdress_med_3	Cardnum_Merch zip_actual/max_0
Cardnum_actual/med_30	Merch zip_avg_14	Cardnum_fulladdress_total_3	Cardnum_Merch zip_actual/med_0
Cardnum_actual/toal_30	Merch zip_max_14	Cardnum_fulladdress_actual/avg_3	Cardnum_Merch zip_actual/toal_0
Cardnum_count_60	Merch zip_med_14	Cardnum_fulladdress_actual/max_3	Cardnum_Merch zip_count_1

Cardnum_avg_60	Merch zip_total_14	Cardnum_fulladdress_actual/med_3	Cardnum_Merch zip_avg_1
Cardnum_max_60	Merch zip_actual/avg_14	Cardnum_fulladdress_actual/toal_3	Cardnum_Merch zip_max_1
Cardnum_med_60	Merch zip_actual/max_14	Cardnum_fulladdress_count_7	Cardnum_Merch zip_med_1
Cardnum_total_60	Merch zip_actual/med_14	Cardnum_fulladdress_avg_7	Cardnum_Merch zip_total_1
Cardnum_actual/avg_60	Merch zip_actual/toal_14	Cardnum_fulladdress_max_7	Cardnum_Merch zip_actual/avg_1
Cardnum_actual/max_60	Merch zip_count_30	Cardnum_fulladdress_med_7	Cardnum_Merch zip_actual/max_1
Cardnum_actual/med_60	Merch zip_avg_30	Cardnum_fulladdress_total_7	Cardnum_Merch zip_actual/med_1
Cardnum_actual/toal_60	Merch zip_max_30	Cardnum_fulladdress_actual/avg_7	Cardnum_Merch zip_actual/toal_1
Cardnum_count_90	Merch zip_med_30	Cardnum_fulladdress_actual/max_7	Cardnum_Merch zip_count_3
Cardnum_avg_90	Merch zip_total_30	Cardnum_fulladdress_actual/med_7	Cardnum_Merch zip_avg_3
Cardnum_max_90	Merch zip_actual/avg_30	Cardnum_fulladdress_actual/toal_7	Cardnum_Merch zip_max_3
Cardnum_med_90	Merch zip_actual/max_30	Cardnum_fulladdress_count_14	Cardnum_Merch zip_med_3
Cardnum_total_90	Merch zip_actual/med_30	Cardnum_fulladdress_avg_14	Cardnum_Merch zip_total_3
Cardnum_actual/avg_90	Merch zip_actual/toal_30	Cardnum_fulladdress_max_14	Cardnum_Merch zip_actual/avg_3
Cardnum_actual/max_90	Merch zip_count_60	Cardnum_fulladdress_med_14	Cardnum_Merch zip_actual/max_3
Cardnum_actual/med_90	Merch zip_avg_60	Cardnum_fulladdress_total_14	Cardnum_Merch zip_actual/med_3
Cardnum_actual/toal_90	Merch zip_max_60	Cardnum_fulladdress_actual/avg_14	Cardnum_Merch zip_actual/toal_3
Merchnum_day_since	Merch zip_med_60	Cardnum_fulladdress_actual/max_14	Cardnum_Merch zip_count_7
Merchnum_count_0	Merch zip_total_60	Cardnum_fulladdress_actual/med_14	Cardnum_Merch zip_avg_7

Merchnum_avg_0	Merch zip_actual/avg_60	Cardnum_fulladdress_actual/toal_14	Cardnum_Merch zip_max_7
Merchnum_max_0	Merch zip_actual/max_60	Cardnum_fulladdress_count_30	Cardnum_Merch zip_med_7
Merchnum_med_0	Merch zip_actual/med_60	Cardnum_fulladdress_avg_30	Cardnum_Merch zip_total_7
Merchnum_total_0	Merch zip_actual/toal_60	Cardnum_fulladdress_max_30	Cardnum_Merch zip_actual/avg_7
Merchnum_actual/avg_0	Merch zip_count_90	Cardnum_fulladdress_med_30	Cardnum_Merch zip_actual/max_7
Merchnum_actual/max_0	Merch zip_avg_90	Cardnum_fulladdress_total_30	Cardnum_Merch zip_actual/med_7
Merchnum_actual/med_0	Merch zip_max_90	Cardnum_fulladdress_actual/avg_30	Cardnum_Merch zip_actual/toal_7
Merchnum_actual/toal_0	Merch zip_med_90	Cardnum_fulladdress_actual/max_30	Cardnum_Merch zip_count_14
Merchnum_count_1	Merch zip_total_90	Cardnum_fulladdress_actual/med_30	Cardnum_Merch zip_avg_14
Merchnum_avg_1	Merch zip_actual/avg_90	Cardnum_fulladdress_actual/toal_30	Cardnum_Merch zip_max_14
Merchnum_max_1	Merch zip_actual/max_90	Cardnum_fulladdress_count_60	Cardnum_Merch zip_med_14
Merchnum_med_1	Merch zip_actual/med_90	Cardnum_fulladdress_avg_60	Cardnum_Merch zip_total_14
Merchnum_total_1	Merch zip_actual/toal_90	Cardnum_fulladdress_max_60	Cardnum_Merch zip_actual/avg_14
Merchnum_actual/avg_1	fulladdress_day_since	Cardnum_fulladdress_med_60	Cardnum_Merch zip_actual/max_14
Merchnum_actual/max_1	fulladdress_count_0	Cardnum_fulladdress_total_60	Cardnum_Merch zip_actual/med_14
Merchnum_actual/med_1	fulladdress_avg_0	Cardnum_fulladdress_actual/avg_60	Cardnum_Merch zip_actual/toal_14
Merchnum_actual/toal_1	fulladdress_max_0	Cardnum_fulladdress_actual/max_60	Cardnum_Merch zip_count_30
Merchnum_count_3	fulladdress_med_0	Cardnum_fulladdress_actual/med_60	Cardnum_Merch zip_avg_30
Merchnum_avg_3	fulladdress_total_0	Cardnum_fulladdress_actual/toal_60	Cardnum_Merch zip_max_30

Merchnum_max_3	fulladdress_actual/avg_0	Cardnum_fulladdress_count_90	Cardnum_Merchzip_med_30
Merchnum_med_3	fulladdress_actual/max_0	Cardnum_fulladdress_avg_90	Cardnum_Merchzip_total_30
Merchnum_total_3	fulladdress_actual/med_0	Cardnum_fulladdress_max_90	Cardnum_Merchzip_actual/avg_30
Merchnum_actual/avg_3	fulladdress_actual/toal_0	Cardnum_fulladdress_med_90	Cardnum_Merchzip_actual/max_30
Merchnum_actual/max_3	fulladdress_count_1	Cardnum_fulladdress_total_90	Cardnum_Merchzip_actual/med_30
Merchnum_actual/med_3	fulladdress_avg_1	Cardnum_fulladdress_actual/avg_90	Cardnum_Merchzip_actual/toal_30
Merchnum_actual/toal_3	fulladdress_max_1	Cardnum_fulladdress_actual/max_90	Cardnum_Merchzip_count_60
Merchnum_count_7	fulladdress_med_1	Cardnum_fulladdress_actual/med_90	Cardnum_Merchzip_avg_60
Merchnum_avg_7	fulladdress_total_1	Cardnum_fulladdress_actual/toal_90	Cardnum_Merchzip_max_60
Merchnum_max_7	fulladdress_actual/avg_1	Cardnum_Merchnum_day_since	Cardnum_Merchzip_med_60
Merchnum_med_7	fulladdress_actual/max_1	Cardnum_Merchnum_count_0	Cardnum_Merchzip_total_60
Merchnum_total_7	fulladdress_actual/med_1	Cardnum_Merchnum_avg_0	Cardnum_Merchzip_actual/avg_60
Merchnum_actual/avg_7	fulladdress_actual/toal_1	Cardnum_Merchnum_max_0	Cardnum_Merchzip_actual/max_60
Merchnum_actual/max_7	fulladdress_count_3	Cardnum_Merchnum_med_0	Cardnum_Merchzip_actual/med_60
Merchnum_actual/med_7	fulladdress_avg_3	Cardnum_Merchnum_total_0	Cardnum_Merchzip_actual/toal_60
Merchnum_actual/toal_7	fulladdress_max_3	Cardnum_Merchnum_actual/avg_0	Cardnum_Merchzip_count_90
Merchnum_count_14	fulladdress_med_3	Cardnum_Merchnum_actual/max_0	Cardnum_Merchzip_avg_90
Merchnum_avg_14	fulladdress_total_3	Cardnum_Merchnum_actual/med_0	Cardnum_Merchzip_max_90
Merchnum_max_14	fulladdress_actual/avg_3	Cardnum_Merchnum_actual/toal_0	Cardnum_Merchzip_med_90

Merchnum_med_14	fulladdress_actual/max_3	Cardnum_Merchnum_count_1	Cardnum_Merchzip_total_90
Merchnum_total_14	fulladdress_actual/med_3	Cardnum_Merchnum_avg_1	Cardnum_Merchzip_actual/avg_90
Merchnum_actual/avg_14	fulladdress_actual/toal_3	Cardnum_Merchnum_max_1	Cardnum_Merchzip_actual/max_90
Merchnum_actual/max_14	fulladdress_count_7	Cardnum_Merchnum_med_1	Cardnum_Merchzip_actual/med_90
Merchnum_actual/med_14	fulladdress_avg_7	Cardnum_Merchnum_total_1	Cardnum_Merchzip_actual/toal_90
Merchnum_actual/toal_14	fulladdress_max_7	Cardnum_Merchnum_actual/avg_1	Cardnum_count_0_by_3
Merchnum_count_30	fulladdress_med_7	Cardnum_Merchnum_actual/max_1	Cardnum_count_0_by_7
Merchnum_avg_30	fulladdress_total_7	Cardnum_Merchnum_actual/med_1	Cardnum_count_0_by_14
Merchnum_max_30	fulladdress_actual/avg_7	Cardnum_Merchnum_actual/toal_1	Cardnum_count_0_by_30
Merchnum_med_30	fulladdress_actual/max_7	Cardnum_Merchnum_count_3	Cardnum_count_0_by_60
Merchnum_total_30	fulladdress_actual/med_7	Cardnum_Merchnum_avg_3	Cardnum_count_0_by_90
Merchnum_actual/avg_30	fulladdress_actual/toal_7	Cardnum_Merchnum_max_3	Cardnum_count_1_by_3
Merchnum_actual/max_30	fulladdress_count_14	Cardnum_Merchnum_med_3	Cardnum_count_1_by_7
Merchnum_actual/med_30	fulladdress_avg_14	Cardnum_Merchnum_total_3	Cardnum_count_1_by_14
Merchnum_actual/toal_30	fulladdress_max_14	Cardnum_Merchnum_actual/avg_3	Cardnum_count_1_by_30
Merchnum_count_60	fulladdress_med_14	Cardnum_Merchnum_actual/max_3	Cardnum_count_1_by_60
Merchnum_avg_60	fulladdress_total_14	Cardnum_Merchnum_actual/med_3	Cardnum_count_1_by_90
Merchnum_max_60	fulladdress_actual/avg_14	Cardnum_Merchnum_actual/toal_3	Merchnum_count_0_by_3
Merchnum_med_60	fulladdress_actual/max_14	Cardnum_Merchnum_count_7	Merchnum_count_0_by_7

Merchnum_total_60	fulladdress_actual/med_14	Cardnum_Merchnum_avg_7	Merchnum_count_0_by_14
Merchnum_actual/avg_60	fulladdress_actual/toal_14	Cardnum_Merchnum_max_7	Merchnum_count_0_by_30
Merchnum_actual/max_60	fulladdress_count_30	Cardnum_Merchnum_med_7	Merchnum_count_0_by_60
Merchnum_actual/med_60	fulladdress_avg_30	Cardnum_Merchnum_total_7	Merchnum_count_0_by_90
Merchnum_actual/toal_60	fulladdress_max_30	Cardnum_Merchnum_actual/avg_7	Merchnum_count_1_by_3
Merchnum_count_90	fulladdress_med_30	Cardnum_Merchnum_actual/max_7	Merchnum_count_1_by_7
Merchnum_avg_90	fulladdress_total_30	Cardnum_Merchnum_actual/med_7	Merchnum_count_1_by_14
Merchnum_max_90	fulladdress_actual/avg_30	Cardnum_Merchnum_actual/toal_7	Merchnum_count_1_by_30
Merchnum_med_90	fulladdress_actual/max_30	Cardnum_Merchnum_count_14	Merchnum_count_1_by_60
Merchnum_total_90	fulladdress_actual/med_30	Cardnum_Merchnum_avg_14	Merchnum_count_1_by_90
Merchnum_actual/avg_90	fulladdress_actual/toal_30	Cardnum_Merchnum_max_14	Merch description_count_0_by_3
Merchnum_actual/max_90	fulladdress_count_60	Cardnum_Merchnum_med_14	Merch description_count_0_by_7
Merchnum_actual/med_90	fulladdress_avg_60	Cardnum_Merchnum_total_14	Merch description_count_0_by_14
Merchnum_actual/toal_90	fulladdress_max_60	Cardnum_Merchnum_actual/avg_14	Merch description_count_0_by_30
Merch description_day_since	fulladdress_med_60	Cardnum_Merchnum_actual/max_14	Merch description_count_0_by_60
Merch description_count_0	fulladdress_total_60	Cardnum_Merchnum_actual/med_14	Merch description_count_0_by_90

Merch description_avg_0	fulladdress_actual/avg_60	Cardnum_Merchnum_actual/toal_14	Merch description_count_1_by_3
Merch description_max_0	fulladdress_actual/max_60	Cardnum_Merchnum_count_30	Merch description_count_1_by_7
Merch description_med_0	fulladdress_actual/med_60	Cardnum_Merchnum_avg_30	Merch description_count_1_by_14
Merch description_total_0	fulladdress_actual/toal_60	Cardnum_Merchnum_max_30	Merch description_count_1_by_30
Merch description_actual/avg_0	fulladdress_count_90	Cardnum_Merchnum_med_30	Merch description_count_1_by_60
Merch description_actual/max_0	fulladdress_avg_90	Cardnum_Merchnum_total_30	Merch description_count_1_by_90
Merch description_actual/med_0	fulladdress_max_90	Cardnum_Merchnum_actual/avg_30	Merch state_count_0_by_3
Merch description_actual/toal_0	fulladdress_med_90	Cardnum_Merchnum_actual/max_30	Merch state_count_0_by_7
Merch description_count_1	fulladdress_total_90	Cardnum_Merchnum_actual/med_30	Merch state_count_0_by_14
Merch description_avg_1	fulladdress_actual/avg_90	Cardnum_Merchnum_actual/toal_30	Merch state_count_0_by_30
Merch description_max_1	fulladdress_actual/max_90	Cardnum_Merchnum_count_60	Merch state_count_0_by_60
Merch description_med_1	fulladdress_actual/med_90	Cardnum_Merchnum_avg_60	Merch state_count_0_by_90
Merch description_total_1	fulladdress_actual/toal_90	Cardnum_Merchnum_max_60	Merch state_count_1_by_3
Merch description_actual/avg_1	Merchnum_Merch description_day_since	Cardnum_Merchnum_med_60	Merch state_count_1_by_7

Merch description_actual/max_1	Merchnum_Merch description_count_0	Cardnum_Merchnum_total_60	Merch state_count_1_by_14
Merch description_actual/med_1	Merchnum_Merch description_avg_0	Cardnum_Merchnum_actual/avg_60	Merch state_count_1_by_30
Merch description_actual/toal_1	Merchnum_Merch description_max_0	Cardnum_Merchnum_actual/max_60	Merch state_count_1_by_60
Merch description_count_3	Merchnum_Merch description_med_0	Cardnum_Merchnum_actual/med_60	Merch state_count_1_by_90
Merch description_avg_3	Merchnum_Merch description_total_0	Cardnum_Merchnum_actual/toal_60	Merch zip_count_0_by_3
Merch description_max_3	Merchnum_Merch description_actual/avg_0	Cardnum_Merchnum_count_90	Merch zip_count_0_by_7
Merch description_med_3	Merchnum_Merch description_actual/max_0	Cardnum_Merchnum_avg_90	Merch zip_count_0_by_14
Merch description_total_3	Merchnum_Merch description_actual/med_0	Cardnum_Merchnum_max_90	Merch zip_count_0_by_30
Merch description_actual/avg_3	Merchnum_Merch description_actual/toal_0	Cardnum_Merchnum_med_90	Merch zip_count_0_by_60
Merch description_actual/max_3	Merchnum_Merch description_count_1	Cardnum_Merchnum_total_90	Merch zip_count_0_by_90
Merch description_actual/med_3	Merchnum_Merch description_avg_1	Cardnum_Merchnum_actual/avg_90	Merch zip_count_1_by_3
Merch description_actual/toal_3	Merchnum_Merch description_max_1	Cardnum_Merchnum_actual/max_90	Merch zip_count_1_by_7
Merch description_count_7	Merchnum_Merch description_med_1	Cardnum_Merchnum_actual/med_90	Merch zip_count_1_by_14
Merch description_avg_7	Merchnum_Merch description_total_1	Cardnum_Merchnum_actual/toal_90	Merch zip_count_1_by_30

Merch description_max_7	Merchnum_Merch description_actual/avg_1	Cardnum_Merch description_day_since	Merch zip_count_1_by_60
Merch description_med_7	Merchnum_Merch description_actual/max_1	Cardnum_Merch description_count_0	Merch zip_count_1_by_90
Merch description_total_7	Merchnum_Merch description_actual/med_1	Cardnum_Merch description_avg_0	fulladdress_count_0_by_3
Merch description_actual/avg_7	Merchnum_Merch description_actual/toal_1	Cardnum_Merch description_max_0	fulladdress_count_0_by_7
Merch description_actual/max_7	Merchnum_Merch description_count_3	Cardnum_Merch description_med_0	fulladdress_count_0_by_14
Merch description_actual/med_7	Merchnum_Merch description_avg_3	Cardnum_Merch description_total_0	fulladdress_count_0_by_30
Merch description_actual/toal_7	Merchnum_Merch description_max_3	Cardnum_Merch description_actual/avg_0	fulladdress_count_0_by_60
Merch description_count_14	Merchnum_Merch description_med_3	Cardnum_Merch description_actual/max_0	fulladdress_count_0_by_90
Merch description_avg_14	Merchnum_Merch description_total_3	Cardnum_Merch description_actual/med_0	fulladdress_count_1_by_3
Merch description_max_14	Merchnum_Merch description_actual/avg_3	Cardnum_Merch description_actual/toal_0	fulladdress_count_1_by_7
Merch description_med_14	Merchnum_Merch description_actual/max_3	Cardnum_Merch description_count_1	fulladdress_count_1_by_14
Merch description_total_14	Merchnum_Merch description_actual/med_3	Cardnum_Merch description_avg_1	fulladdress_count_1_by_30

Merch description_actual/avg_14	Merchnum_Merch description_actual/toal_3	Cardnum_Merch description_max_1	fulladdress_count_1_by_60
Merch description_actual/max_14	Merchnum_Merch description_count_7	Cardnum_Merch description_med_1	fulladdress_count_1_by_90
Merch description_actual/med_14	Merchnum_Merch description_avg_7	Cardnum_Merch description_total_1	Merchnum_Merch description_count_0_by_3
Merch description_actual/toal_14	Merchnum_Merch description_max_7	Cardnum_Merch description_actual/avg_1	Merchnum_Merch description_count_0_by_7
Merch description_count_30	Merchnum_Merch description_med_7	Cardnum_Merch description_actual/max_1	Merchnum_Merch description_count_0_by_14
Merch description_avg_30	Merchnum_Merch description_total_7	Cardnum_Merch description_actual/med_1	Merchnum_Merch description_count_0_by_30
Merch description_max_30	Merchnum_Merch description_actual/avg_7	Cardnum_Merch description_actual/toal_1	Merchnum_Merch description_count_0_by_60
Merch description_med_30	Merchnum_Merch description_actual/max_7	Cardnum_Merch description_count_3	Merchnum_Merch description_count_0_by_90
Merch description_total_30	Merchnum_Merch description_actual/med_7	Cardnum_Merch description_avg_3	Merchnum_Merch description_count_1_by_3
Merch description_actual/avg_30	Merchnum_Merch description_actual/toal_7	Cardnum_Merch description_max_3	Merchnum_Merch description_count_1_by_7
Merch description_actual/max_30	Merchnum_Merch description_count_14	Cardnum_Merch description_med_3	Merchnum_Merch description_count_1_by_14
Merch description_actual/med_30	Merchnum_Merch description_avg_14	Cardnum_Merch description_total_3	Merchnum_Merch description_count_1_by_30

Merch description_actual/toal_30	Merchnum_Merch description_max_14	Cardnum_Merch description_actual/avg_3	Merchnum_Merch description_count_1_by_60
Merch description_count_60	Merchnum_Merch description_med_14	Cardnum_Merch description_actual/max_3	Merchnum_Merch description_count_1_by_90
Merch description_avg_60	Merchnum_Merch description_total_14	Cardnum_Merch description_actual/med_3	Merchnum_fulladdress_count_0_by_3
Merch description_max_60	Merchnum_Merch description_actual/avg_14	Cardnum_Merch description_actual/toal_3	Merchnum_fulladdress_count_0_by_7
Merch description_med_60	Merchnum_Merch description_actual/max_14	Cardnum_Merch description_count_7	Merchnum_fulladdress_count_0_by_14
Merch description_total_60	Merchnum_Merch description_actual/med_14	Cardnum_Merch description_avg_7	Merchnum_fulladdress_count_0_by_30
Merch description_actual/avg_60	Merchnum_Merch description_actual/toal_14	Cardnum_Merch description_max_7	Merchnum_fulladdress_count_0_by_60
Merch description_actual/max_60	Merchnum_Merch description_count_30	Cardnum_Merch description_med_7	Merchnum_fulladdress_count_0_by_90
Merch description_actual/med_60	Merchnum_Merch description_avg_30	Cardnum_Merch description_total_7	Merchnum_fulladdress_count_1_by_3
Merch description_actual/toal_60	Merchnum_Merch description_max_30	Cardnum_Merch description_actual/avg_7	Merchnum_fulladdress_count_1_by_7
Merch description_count_90	Merchnum_Merch description_med_30	Cardnum_Merch description_actual/max_7	Merchnum_fulladdress_count_1_by_14
Merch description_avg_90	Merchnum_Merch description_total_30	Cardnum_Merch description_actual/med_7	Merchnum_fulladdress_count_1_by_30

Merch description_max_90	Merchnum_Merch description_actual/avg_30	Cardnum_Merch description_actual/toal_7	Merchnum_fulladdress_count_1_by_60
Merch description_med_90	Merchnum_Merch description_actual/max_30	Cardnum_Merch description_count_14	Merchnum_fulladdress_count_1_by_90
Merch description_total_90	Merchnum_Merch description_actual/med_30	Cardnum_Merch description_avg_14	Cardnum_fulladdress_count_0_by_3
Merch description_actual/avg_90	Merchnum_Merch description_actual/toal_30	Cardnum_Merch description_max_14	Cardnum_fulladdress_count_0_by_7
Merch description_actual/max_90	Merchnum_Merch description_count_60	Cardnum_Merch description_med_14	Cardnum_fulladdress_count_0_by_14
Merch description_actual/med_90	Merchnum_Merch description_avg_60	Cardnum_Merch description_total_14	Cardnum_fulladdress_count_0_by_30
Merch description_actual/toal_90	Merchnum_Merch description_max_60	Cardnum_Merch description_actual/avg_14	Cardnum_fulladdress_count_0_by_60
Merch state_day_since	Merchnum_Merch description_med_60	Cardnum_Merch description_actual/max_14	Cardnum_fulladdress_count_0_by_90
Merch state_count_0	Merchnum_Merch description_total_60	Cardnum_Merch description_actual/med_14	Cardnum_fulladdress_count_1_by_3
Merch state_avg_0	Merchnum_Merch description_actual/avg_60	Cardnum_Merch description_actual/toal_14	Cardnum_fulladdress_count_1_by_7
Merch state_max_0	Merchnum_Merch description_actual/max_60	Cardnum_Merch description_count_30	Cardnum_fulladdress_count_1_by_14
Merch state_med_0	Merchnum_Merch description_actual/med_60	Cardnum_Merch description_avg_30	Cardnum_fulladdress_count_1_by_30

Merch state_total_0	Merchnum_Merch description_actual/toal_60	Cardnum_Merch description_max_30	Cardnum_fulladdress_count_1_by_60
Merch state_actual/avg_0	Merchnum_Merch description_count_90	Cardnum_Merch description_med_30	Cardnum_fulladdress_count_1_by_90
Merch state_actual/max_0	Merchnum_Merch description_avg_90	Cardnum_Merch description_total_30	Cardnum_Merchnum_count_0_by_3
Merch state_actual/med_0	Merchnum_Merch description_max_90	Cardnum_Merch description_actual/avg_30	Cardnum_Merchnum_count_0_by_7
Merch state_actual/toal_0	Merchnum_Merch description_med_90	Cardnum_Merch description_actual/max_30	Cardnum_Merchnum_count_0_by_14
Merch state_count_1	Merchnum_Merch description_total_90	Cardnum_Merch description_actual/med_30	Cardnum_Merchnum_count_0_by_30
Merch state_avg_1	Merchnum_Merch description_actual/avg_90	Cardnum_Merch description_actual/toal_30	Cardnum_Merchnum_count_0_by_60
Merch state_max_1	Merchnum_Merch description_actual/max_90	Cardnum_Merch description_count_60	Cardnum_Merchnum_count_0_by_90
Merch state_med_1	Merchnum_Merch description_actual/med_90	Cardnum_Merch description_avg_60	Cardnum_Merchnum_count_1_by_3
Merch state_total_1	Merchnum_Merch description_actual/toal_90	Cardnum_Merch description_max_60	Cardnum_Merchnum_count_1_by_7
Merch state_actual/avg_1	Merchnum_fulladdress_day_since	Cardnum_Merch description_med_60	Cardnum_Merchnum_count_1_by_14
Merch state_actual/max_1	Merchnum_fulladdress_count_0	Cardnum_Merch description_total_60	Cardnum_Merchnum_count_1_by_30
Merch state_actual/med_1	Merchnum_fulladdress_avg_0	Cardnum_Merch description_actual/avg_60	Cardnum_Merchnum_count_1_by_60
Merch state_actual/toal_1	Merchnum_fulladdress_max_0	Cardnum_Merch description_actual/max_60	Cardnum_Merchnum_count_1_by_90

Merch state_count_3	Merchnum_fulladdress_med_0	Cardnum_Merch description_actual/med_60	Cardnum_Merch description_count_0_by_3
Merch state_avg_3	Merchnum_fulladdress_total_0	Cardnum_Merch description_actual/toal_60	Cardnum_Merch description_count_0_by_7
Merch state_max_3	Merchnum_fulladdress_actual/avg_0	Cardnum_Merch description_count_90	Cardnum_Merch description_count_0_by_14
Merch state_med_3	Merchnum_fulladdress_actual/max_0	Cardnum_Merch description_avg_90	Cardnum_Merch description_count_0_by_30
Merch state_total_3	Merchnum_fulladdress_actual/med_0	Cardnum_Merch description_max_90	Cardnum_Merch description_count_0_by_60
Merch state_actual/avg_3	Merchnum_fulladdress_actual/toal_0	Cardnum_Merch description_med_90	Cardnum_Merch description_count_0_by_90
Merch state_actual/max_3	Merchnum_fulladdress_count_1	Cardnum_Merch description_total_90	Cardnum_Merch description_count_1_by_3
Merch state_actual/med_3	Merchnum_fulladdress_avg_1	Cardnum_Merch description_actual/avg_90	Cardnum_Merch description_count_1_by_7
Merch state_actual/toal_3	Merchnum_fulladdress_max_1	Cardnum_Merch description_actual/max_90	Cardnum_Merch description_count_1_by_14
Merch state_count_7	Merchnum_fulladdress_med_1	Cardnum_Merch description_actual/med_90	Cardnum_Merch description_count_1_by_30
Merch state_avg_7	Merchnum_fulladdress_total_1	Cardnum_Merch description_actual/toal_90	Cardnum_Merch description_count_1_by_60
Merch state_max_7	Merchnum_fulladdress_actual/avg_1	Cardnum_Merch state_day_since	Cardnum_Merch description_count_1_by_90
Merch state_med_7	Merchnum_fulladdress_actual/max_1	Cardnum_Merch state_count_0	Cardnum_Merch state_count_0_by_3

Merch state_total_7	Merchnum_fulladdress _actual/med_1	Cardnum_Merch state_avg_0	Cardnum_Merch state_count_0_by_7
Merch state_actual/avg_7	Merchnum_fulladdress _actual/toal_1	Cardnum_Merch state_max_0	Cardnum_Merch state_count_0_by_14
Merch state_actual/max_7	Merchnum_fulladdress _count_3	Cardnum_Merch state_med_0	Cardnum_Merch state_count_0_by_30
Merch state_actual/med_7	Merchnum_fulladdress _avg_3	Cardnum_Merch state_total_0	Cardnum_Merch state_count_0_by_60
Merch state_actual/toal_7	Merchnum_fulladdress _max_3	Cardnum_Merch state_actual/avg_0	Cardnum_Merch state_count_0_by_90
Merch state_count_14	Merchnum_fulladdress _med_3	Cardnum_Merch state_actual/max_0	Cardnum_Merch state_count_1_by_3
Merch state_avg_14	Merchnum_fulladdress _total_3	Cardnum_Merch state_actual/med_0	Cardnum_Merch state_count_1_by_7
Merch state_max_14	Merchnum_fulladdress _actual/avg_3	Cardnum_Merch state_actual/toal_0	Cardnum_Merch state_count_1_by_14
Merch state_med_14	Merchnum_fulladdress _actual/max_3	Cardnum_Merch state_count_1	Cardnum_Merch state_count_1_by_30
Merch state_total_14	Merchnum_fulladdress _actual/med_3	Cardnum_Merch state_avg_1	Cardnum_Merch state_count_1_by_60
Merch state_actual/avg_14	Merchnum_fulladdress _actual/toal_3	Cardnum_Merch state_max_1	Cardnum_Merch state_count_1_by_90
Merch state_actual/max_14	Merchnum_fulladdress _count_7	Cardnum_Merch state_med_1	Cardnum_Merch zip_count_0_by_3
Merch state_actual/med_14	Merchnum_fulladdress _avg_7	Cardnum_Merch state_total_1	Cardnum_Merch zip_count_0_by_7
Merch state_actual/toal_14	Merchnum_fulladdress _max_7	Cardnum_Merch state_actual/avg_1	Cardnum_Merch zip_count_0_by_14
Merch state_count_30	Merchnum_fulladdress _med_7	Cardnum_Merch state_actual/max_1	Cardnum_Merch zip_count_0_by_30
Merch state_avg_30	Merchnum_fulladdress _total_7	Cardnum_Merch state_actual/med_1	Cardnum_Merch zip_count_0_by_60
Merch state_max_30	Merchnum_fulladdress _actual/avg_7	Cardnum_Merch state_actual/toal_1	Cardnum_Merch zip_count_0_by_90
Merch state_med_30	Merchnum_fulladdress _actual/max_7	Cardnum_Merch state_count_3	Cardnum_Merch zip_count_1_by_3
Merch state_total_30	Merchnum_fulladdress _actual/med_7	Cardnum_Merch state_avg_3	Cardnum_Merch zip_count_1_by_7

Merch state_actual/avg_30	Merchnum_fulladdress _actual/toal_7	Cardnum_Merch state_max_3	Cardnum_Merch zip_count_1_by_14
Merch state_actual/max_30	Merchnum_fulladdress _count_14	Cardnum_Merch state_med_3	Cardnum_Merch zip_count_1_by_30
Merch state_actual/med_30	Merchnum_fulladdress _avg_14	Cardnum_Merch state_total_3	Cardnum_Merch zip_count_1_by_60
Merch state_actual/toal_30	Merchnum_fulladdress _max_14	Cardnum_Merch state_actual/avg_3	Cardnum_Merch zip_count_1_by_90
Merch state_count_60	Merchnum_fulladdress _med_14	Cardnum_Merch state_actual/max_3	dow_risk
Merch state_avg_60	Merchnum_fulladdress _total_14	Cardnum_Merch state_actual/med_3	CardnumU*
Merch state_max_60	Merchnum_fulladdress _actual/avg_14	Cardnum_Merch state_actual/toal_3	MerchnumU*