Unsupervised Learning on Corporate Card Transaction Data

Jiaying Gu
Ruoyu Sun
Jiaxi Yu
Siyu Zhang
Xinyi Zhao

Executive Summary

Project 2, Unsupervised Learning on Corporate Card Transaction Data, is an exploration of 95,272 credit card transactions made during the year of 2010 from corporate card dataset. For each transaction record, there's detailed information of card number, transaction date, merchant number, merchant state and zip, etc. Among these 95,272 transactions, there are 206 transactions that are fraudulent. For example, Fraud Type 1 transactions are all made using the same card, with the same merchant, over 2 days. Fraud Type 2 transactions are made with the same merchant across many cards in 5 days, also with usual high dollar amounts. Through unsupervised learning method, our ultimate goal is to catch as many transactions with the fraud label as possible by building models to compute a fraud score for each record. The labeled frauds are not used in our model but are used to check the robustness of the model.

We started the project by conducting data quality report on the dataset, briefly explored the dataset and found interesting and unusual things about the data. After understanding what the data is about, we began to build variables and prepared the variables for model construction. When choosing entity levels, we decided that card number (CARDNUM) and merchant number (MERCHNUM) are proper because basically they help explain behavior variation in our data. Since we want to study the transaction behavior, two things we focus on are the number of transactions and the amount of transaction. The variables we built are number of transactions in past N (N = 1,2,3,7) days and amount of transactions in past N (N = 1,2,3,7) days based on CARDNUM and MERCHNUM entity levels, and standardize them by the dividing the corresponding activities in the past 90 days. This gives us a total of 16 calculated variables. Setting the first 90 days as baseline, we are only able to score the transactions starting from April 1. We experimented several scoring methods, the method of z-scaling the variables and taking the weighted average of the highest z-scores gives us the best result, which catches 67% of the fraud transactions in our top 5% scored records and 70% of the fraud transactions in our top 10% scored records.

In the end of the report, we will discuss the fraud transactions that have been caught through our model and the high fraud score transactions in detail.

Summary of Data

The data contains credit card transaction records, along with the card number, merchant information, and transaction date and type. There is a total of 95272 records with 8 fields. The timeframe was from 1/1/2010 to 12/31/2010 and the original format is Excel.

Looking at the data, we first separated the fields into numerical and categorical variables. The basic standard we follow is to put most fields with continuous numerical values as a numerical variable, and put fields with word descriptions as a categorical variable. However, for variables like CARDNUM, MERCHNUM and MERCHZIP which seem like numerical, we think it makes more sense to put those fields as categorical. We can conduct our analysis by selecting categorical variables to account for the difference between each level. Below is a summary of the field names and the percent populated in each field.

Numerical Fields	
Field Name	% Populated
AMOUNT	100%

Categorical Fields	
Field Name	% Populated
CARDNUM	100%
MERCHNUM	96%
MERCHDESCRIPTION	100%
MERCHSTATE	99%
TRANSTYPE	100%
MERCHZIP	95%
DATE	100%

From our basic exploration of the data, a few findings may help guide further analysis:

- The number of transactions associated with each card number varies greatly, with largest number over 1,000. The number of transactions for each merchant also varies greatly, with largest number over 9,000. It might be interesting to explore the high values within these two entities.
- The zip code with each merchant has the highest number of missing values. The zip codes listed also have different length and formats. Due to the unexplainable irregularity in this field, we might not choose it as an entity for our analysis.
-

CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCH	MERCHZIP	TRANSTYPE	AMOUNT	fraud?
5142189135	7/13/2010		INTERMEXICO			P	$3,102,045.53	

This is the largest amount of payment in the dataset, and has a significantly higher value than other records. There are many missing information in this row and the information of Merchant description which is “INTERMEXICO” is very suspicious associated with this payment amount. The existence of this record might have an influence on our scoring of other records.

Entities and Variables

Entities:

We considered dividing the data based on two entity levels: CARDNUM and MERCHNUM. Observing anomalies on these two entity levels may help account for the user differences among different card holders and different merchants.

Variables:

We added a total of 16 variables to model our data. Our intention is to find anomalies based on the number of transactions and the total transaction amount during a time frame. We calculated each variable on its entity level. Due to the usual patterns of credit card fraud, we selected the time frame to be in the past 1, 2, 3, or 7 days. Since we are assuming that we have no knowledge of records that happened after each existing record, we standardized the variable by setting the activity on each entity level in the past 90 days as normal.

As we define fraud as those records who have an unusually high score, we set all negative values in the variables to be 0. We also removed records that are within the first 90 days, from 1/1/2010 to 3/31/2010, from our ranking, since these records do not have a 90 day history that we are using for standardization. If the 90 day history of a record is 0, which cannot be divided, we set the variable for that record to be 1, the assumed normal value.

We excluded the record with the highest transaction amount, to prevent it from influencing our analysis.

Below are our variables:

- $card\_scale\_trans\_N = (90/N) \cdot \frac{Number\ of\ transactions\ in\ the\ past\ N\ days\ on\ this\ card}{Number\ of\ transactions\ in\ the\ past\ 90\ days\ on\ this\ card}$

  For N = 1, 2, 3, 7
- $card\_scale\_amount\_N = (90/N) \cdot \frac{Total\ transaction\ amount\ in\ the\ past\ N\ days\ on\ this\ card}{Total\ transaction\ amount\ in\ the\ past\ 90\ days\ on\ this\ card}$

  For N = 1, 2, 3, 7
- $merch\_scale\_trans\_N = (90/N) \cdot \frac{Number\ of\ transactions\ in\ the\ past\ N\ days\ from\ merchant}{Number\ of\ transactions\ in\ the\ past\ 90\ days\ from\ merchant}$

  For N = 1, 2, 3, 7
- $merch\_scale\_amount\_N = (90/N) \cdot \frac{Total\ trans\ amount\ in\ the\ past\ N\ days\ from\ merchant}{Total\ trans\ amount\ in\ the\ past\ 90\ days\ from\ merchant}$

  For N = 1, 2, 3, 7

Model Algorithm

After creating the 16 variables as mentioned above, we experimented with two types of scoring algorithms:

1. Conduct Principle Component Analysis and choose top components that together contribute roughly 70% to 80% of the variation. Then perform z-scaling on the top components. To calculate the fraud score, we tried different approaches:
   a) Score = sum of z-scores
   b) Score = average of z-scores

2. Perform z-scaling on those 16 variables. Consequently, we have 16 groups of z-scores. We explored three ways of calculating fraud score:
   a) Score = sum of all z-scores
   b) Score = average of all z-scores
   c) Score = weighted average of the top 4 z-scores (one for each N). For example, find the highest z-score among the groups of 4 variables for number of transactions in n days (N = 1, 2, 3, 7). Do the same for other three groups.

In both approaches, we make sure that we are only looking at the positive z-scores when calculating the final fraud score. The reason is that in this project, we are mainly interested in finding cases where the card is lost or stolen, or a bad merchant gets the card. In theses cases, records tend to have higher-than-usual amount or number of transactions in a given period. Therefore, we transformed all negative z-scores into zero to only focus on positive deviations.

In addition, we did not score the first 3 months of the records. The fraud algorithms need to “mature” before they can do a good job in finding frauds. We chose 90 days as a training period so that the algorithms can get a good sense of what is normal behavior for each card number or merchant. We experimented with scoring all records, and the result was not as good as eliminating the first 90 days of records.

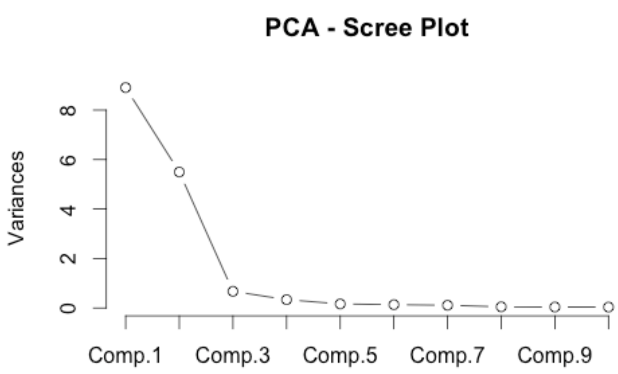
We found out that the second approach with Score = weighted average of the top 4 z-scores (one for each N) gives the best result. Detailed result will be presented later in the report.

Model Result:

As mentioned in Model Algorithm, we experimented with two different approaches as well as a number of ways to compute the final fraud score. We will briefly talk about the result of these algorithms and focus on the one that we decided to implement.

1. Conduct Principle Component Analysis then z-scale

We chose the first two components to do z-scaling on. From the Scree plot below, we can see that the top two components can together explain the majority of the variation in the data.



Then we did z-scaling on the scores of these two components. Then we tried Score = sum of z-scores and Score = average of z-scores, and they gave very similar results. We can only catch 2 fraud records in the top 10%.

2. Perform z-scaling on those 16 variables

This method does better than the previous one. We calculated the z-scores associated with the 16 variables. Here again, we experimented with different ways to score. We found that taking the weighted average of the top 4 z-scores generate the best result. The top 4 z-scores are determined by finding the highest z-score among the groups of 4 variables. Specifically, find the maximum values for number of transactions and amount of transaction in n days (N = 1, 2, 3, 7) on card number entity level and merchant number entity level. Then we took a weighted average of these four z-scores. The weights are:

card_scale_trans	merch_scale_trans	card_scale_amount	merch_scale_amount
0.2	0.05	0.7	0.05

The graph bellows show the top scores:

	A	AC	AD	AE
1	Record Number	Fraud Score	Rank	fraud?
2	23911	3.90	1	
3	23920	3.90	1	
4	24313	3.90	1	
5	24414	3.90	1	
6	24451	3.90	1	
7	24636	3.90	1	
8	24660	3.90	1	
9	25055	3.90	1	
10	25265	3.90	1	
11	25467	3.90	1	
12	25570	3.90	1	
13	25711	3.90	1	
14	25716	3.90	1	
15	25756	3.90	1	
16	25850	3.90	1	
17	25921	3.90	1	
18	26426	3.90	1	
19	27471	3.90	1	
20	27510	3.90	1	

Then we checked what percent of all fraud records can be caught by our model in the top 5% and top 10% records:

	Top 5%	Top 10%
Number of Fraud Caught	137	144
% of all fraud	**67%**	**70%**

Our model is quite robust as it captures 67% of all the fraud in the top 5% records. Later in the report, we will look into some of the top scores and explain why they are unusual.

High Score Analysis

Within fraud labeled 1~12, we caught 137 out of 206 in top 5% highest scores (about 4500 records), with most within the labels #1~#7.

We ranked the top 5% highest score records starting from 1 and did analysis on why they are fraudulent. Here are some records we think are typical of each kind of fraud that our algorithm find. Each table represent one fraud, including all original information, a score for each record and a corresponding rank.

Some records that we included in our report are not in the top 5% (the score and rank cells are empty), but we still put it there because it can help us better understand why a particular record was labeled fraud by our algorithm. Since we use the past 90 days as our definition of normal, these records may help us understand how far away our target record is from normal, and possibly also the reason why the target record was labeled as fraud.

In our analysis below, we only interpreted records that had the highest fraud score (ranked as 1), and we grouped them by the type of fraud possible. Under each type of fraud, we grouped the suspicious records by the entity level that conveys more information, and analyzed each target record in context of date and entity.

Fraud type: Large transaction amount in short amount of time

Entity: CARDNUM

Explanation: this cardholder made 10 transactions in one day. The merchant is the same and the amount of each transaction is huge.

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
32569	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	3,640		3.887340703	1
32587	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	3,225		3.887340703	1
32591	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	2,250		3.887340703	1
32625	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	700		3.887340703	1
32667	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	14,625		3.887340703	1
32730	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	2,100		3.887340703	1
32765	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	2,100		3.887340703	1
32785	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	2,475		3.887340703	1
32815	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	2,156		3.887340703	1
32888	5142288601	5/4/2010	460450006HRI6	SENTINEL, INC.	AL	35801	P	800		3.887340703	1

Entity: CARDNUM

Explanation: this cardholder made 4 transactions in a very short period of time. The merchant is the same and the amount of each transaction is huge and almost identical.

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
61425	5142308889	8/11/2010	6054006890063	BAR CODE DISCOUNT WAREHOU	OH	44133	P	2,245		3.887340703	1
62204	5142308889	8/14/2010	6054006890063	BAR CODE DISCOUNT WAREHOU	OH	44133	P	2,241		2.875472119	394
62681	5142308889	8/15/2010	6054006890063	BAR CODE DISCOUNT WAREHOU	OH	44133	P	2,241		2.875472119	394
62855	5142308889	8/16/2010	6054006890063	BAR CODE DISCOUNT WAREHOU	OH	44133	P	2,242		2.875472119	394

Entity: CARDNUM

Explanation: our algorithm captured 4 records of this cardholder so we just put it together. Transactions of large amount in a week.

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
32720	5142182016	5/4/2010	7129011009306	CBQ-NEWPORT #2	RI		P	$2,177.40		3.887340703	1
32842	5142182016	5/4/2010	8006000808492	CONWAY'S TOURS/GRAY LI	RI	02864	P	$575.00		3.887340703	1
34304	5142182016	5/10/2010	89200600057	ARAMARK HYNES CON	MA	02115	P	$6,116.01		2.894465643	384
41411	5142182016	6/5/2010	7129011006606	CBQ-NEWPORT #1	RI	02841	P	$262.40			
50001	5142182016	7/5/2010	8060633001300	BOSTON PARK PLAZA HOTEL	MA	02116	P	$7,947.90		1.406695437	2255

Entity: CARDNUM

Explanation: this cardholder made 4 transactions in a very short period of time. The amount is huge. Also it is worth further analysis how this cardholder was able to made transactions in different states in 1 day (see record #71835, #72089). Is it online transactions or not?

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
70305	5142310347	9/8/2010	4620006308197	A-Z SALES & SERVICE INC	CO	80524	P	$1,697.30		3.887340703	1
71249	5142310347	9/12/2010	607990940336	TOOL & ANCHOR SUPPLY #2	CO	80204	P	$2,180.00		2.894465643	384
71835	5142310347	9/13/2010	06-3666163370	SWINTEC CORPORATION	NJ	07074	P	$1,696.32		2.894465643	384
72089	5142310347	9/13/2010	604906862335	INTERMTN SAFETY SHOES SR	CO	80907	P	$496.75		2.759679847	469
72345	5142310347	9/14/2010	997536508333	FIVE R REPAIR INC	CO	80401	P	$509.97		2.786916818	445

<u>Fraud type: Large transaction amount compare to other transactions</u>

Entity: CARDNUM

Explanation: this type of fraud is common in our captured records. They easily got high scores because the amount is huge compared to transactions before. In order to make it clearer, we kept all the records before high score transactions regardless they have scores or not.

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
2661	5142186335	1/13/2010	465094667331	AGILENT SAP	GA	30319	P	299.78			
3995	5142186335	1/19/2010	465094667331	AGILENT SAP	GA	30319	P	163			
6968	5142186335	2/1/2010	900009091152	BESSENBERG BINDERY CORP	MI	48104	P	162			
7132	5142186335	2/2/2010	465094667331	AGILENT SAP	GA	30319	P	120.16			
33376	5142186335	5/8/2010	08-3508724258	TOOL CRIB OF THE NORTH	ND	58201	P	24.2		3.887340703	1
33642	5142186335	5/8/2010	08-3508724258	TOOL CRIB OF THE NORTH	ND	58201	P	1748.79		3.887340703	1
34278	5142186335	5/10/2010	955666251221	VALCU INSTRUMENTS	TX	77255	P	141.85		3.211706018	228

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
30805	5142125025	4/28/2010	8060639563700	RADISSON OLD TOWNE ALEXA	VA	22314	P	$6,068.00		3.887340703	1
33234	5142125025	5/6/2010	1988500018501	TROPICANA RESORT	NV	89109	P	$4,225.22		1.440039417	2179
35784	5142125025	5/16/2010	1988500018501	TROPICANA RESORT	NV	89109	P	5,448			
37043	5142125025	5/19/2010		RETAIL CREDIT ADJUSTMENT			P	6,068			
38847	5142125025	5/25/2010	1988500018501	TROPICANA RESORT	NV	89109	P	213			
39342	5142125025	5/27/2010	1988500010006	TROPICANA RESORT	NV	89109	P	213			
39345	5142125025	5/27/2010	1988500010006	TROPICANA RESORT	NV	89109	P	5,448			
39487	5142125025	5/29/2010	1988500010006	TROPICANA RESORT	NV	89109	P	539			
39500	5142125025	5/29/2010	1988500018501	TROPICANA RESORT	NV	89109	P	213			
39970	5142125025	5/31/2010	1988500018501	TROPICANA RESORT	NV	89109	P	539			
40033	5142125025	5/31/2010	1988500010006	TROPICANA RESORT	NV	89109	P	213			
40211	5142125025	5/31/2010	1988500010006	TROPICANA RESORT	NV	89109	P	326			
42682	5142125025	6/8/2010		RETAIL DEBIT ADJUSTMENT			P	6,068			
92769	5142125025	12/17/2010	2094890001832	S.R. COVEY LEADERSHIP CTR	UT	84606	P	4,815		3.782707154	28
92849	5142125025	12/17/2010	2094890001832	S.R. COVEY LEADERSHIP CTR	UT	84606	P	3,858		3.782707154	28

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
350	5142207606	1/4/2010	600000201284	AIR & WASTE MGMT ASSOC	PA	15222	P	$549.00			
8505	5142207606	2/7/2010	4800000698423	RADIO SHACK 00123489	NC	27707	P	$22.39			
40216	5142207606	5/31/2010	6880098991453	MOTIONEX	NC	28217	P	2,305		3.887340703	1
78533	5142207606	10/4/2010	605060869	WHATMAN INC	MA	02173	P	280		3.887340703	1
78643	5142207606	10/5/2010	9415000739457	INPRISE CORPORATION	CA	95066	P	500		3.59070494	128
78970	5142207606	10/6/2010	9415000739457	INPRISE CORPORATION	CA	95066	P	30		3.393119195	166
79787	5142207606	10/12/2010	970610004229	QUANTEX MICROSYSTEMS INC	NJ	08873	P	1,871		2.34359241	802

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
3459	5142164130	1/16/2010	83400040006	KEY COMMUNICATIONS GP	MD	20815	P	$143.20			
10901	5142164130	2/15/2010	9900020006406	GSA-FSS-ADV	VA		P	$25.48			
19641	5142164130	3/17/2010	2722000970806	AMAC KC PUBS	KS	12983	P	203			
58955	5142164130	8/4/2010	9100020008106	OPM-ATLANTA SVC CTR036	GA	30303	P	1,900		3.887340703	1

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
993	5142186909	1/6/2010	9013200007658	BIRCLAR ELECTRIC&ELECTRIC	MI	48174	P	$2,045.00			
22985	5142186909	3/30/2010	9233400065101	INTERNATIONAL LIBRARY	UT	84604	P	$68.60			
23462	5142186909	3/31/2010	991904849338	AMER NATL STDS INST INC	NY	10036	P	18			
23976	5142186909	4/3/2010	972610657332	IEEE BOOK ORDERS	NJ	08855	P	6			
23995	5142186909	4/3/2010	972610657332	IEEE BOOK ORDERS	NJ	08855	P	115			
53763	5142186909	7/19/2010	9013200007658	BIRCLAR ELECTRIC&ELECTRIC	MI	48174	P	2,045		3.887340703	1

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
3459	5142164130	1/16/2010	83400040006	KEY COMMUNICATIONS GP	MD	20815	P	$143.20			
10901	5142164130	2/15/2010	9900020006406	GSA-FSS-ADV	VA		P	$25.48			
19641	5142164130	3/17/2010	2722000970806	AMAC KC PUBS	KS	12983	P	203			
58955	5142164130	8/4/2010	9100020008106	OPM-ATLANTA SVC CTR036	GA	30303	P	1,900		3.887340703	1

<u>Fraud type: Suspicious merchant</u>

Entity: MERCHNUM

Explanation: This merchant has multiple merchant number, in the same zip code. Transactions made in a short period of time with high amount by a single cardholder, which is also suspicious.

Record #	CARDNUM	DATE	MERCHNUM	MERCHDESCRIPTION	MERCHSTATE	MERCHZIP	TRANSTYPE	AMOUNT	Fraud?	Score	Rank
34023	5142149691	5/10/2010	330400610031	OMNI INNER HARBOR	MD	21201	P	$336.00			
52696	5142295584	7/14/2010	330400610006	OMNI INNER HARBOR	MD	21201	P	$3,836.25		3.887340703	1
55015	5142295584	7/22/2010	330400610006	OMNI INNER HARBOR	MD	21201	P	$48.36			
55067	5142295584	7/22/2010	330400615555	OMNI INNER HARBOR	MD	21201	P	$495.00			
55636	5142295584	7/25/2010	330400615555	OMNI INNER HARBOR	MD	21201	P	$6,294.00		2.102439357	1043
57087	5142295584	7/29/2010	330400610033	OMNI INNER HARBOR	MD	21201	P	$123.75			

Appendix-Data quality report

Dataset Name	Payments
File format	Excel
# of records	95272
# of fields	1 Unique identifier
	8 Independent variables
	1 Dependent variable (fraud?)

Numerical Fields	Field Name	Description	Length	Non-missing	Missing	Missing Percent	Cumulative Distribution		Standard Deviation	Mean	Min	Max
1	AMOUNT	Payment amount	7	95272	0	0.00%	p1	3.57	10077.35	413.33	0.01	3102045.53
							p5	3.62				
							p10	4.37				
							p25	33.2				
							p50	136.31				
							p75	420.09				
							p90	1054.63				
							p95	1721.03				
							p99	2480				

Categorical Fields	Field Name	Description	Length	Non-missing	Missing	Missing Percent	Frequent Value	Counts	Unique
1	CARDNUM	Card number	10	95272	0	0.00%	5142148452	1192	1636
							5142184598	921	
							5142189108	663	
							5142297710	583	
							5142223373	577	
							5142187452	526	
							5142299634	515	
							5142189945	512	
							5142149691	497	
							5142190147	488	
2	MERCHNUM	Merchant number	13	91690	3582	3.76%	930090121224	9157	10388
							5509006296254	2131	
							9900020006406	1713	
							602608969534	1091	
							410000971343	981	
							9918000409955	953	
							4353000719908	940	
							5725000466504	868	
							9108234610000	784	
							602608969138	782	
3	MERCHDESCRIPTION	Merchant description	25	95272	0	0.00%	GSA-FSS-ADV	1687	13121
							SIGMA-ALDRICH	1632	
							STAPLES #941	1131	
							FISHER SCI ATL	1092	
							MWI*MICRO WAREHOUSE	955	
							CDW*GOVERNMENT INC	868	
							DELL MARKETING L.P.	804	
							FISHER SCI CHI	782	
							OFFICE DEPOT #1082	744	
							AMAZON.COM *SUPERSTOR	723	
4	MERCHSTATE	State of merchant's location	5	94081	1191	1.25%	TN	11834	227
							VA	7698	
							CA	6653	
							IL	6485	
							MD	5343	
							GA	4930	
							PA	4844	
							NJ	3904	
							TX	3748	
							WA	3215	
5	TRANSTYPE	Transaction Type	1	95272	0	0.00%	P	94917	4
							A	181	
							D	173	
							Y	1	
6	MERCHZIP	Zip code of merchant's location	5	90633	4639	5.12%	38118	11669	4567
							63103	1647	
							8701	1263	
							60061	1217	
							22202	1216	
							17201	1136	
							98101	1117	
							30091	1091	
							60143	940	
							60069	825	
7	DATE	Date of merchant	8	95272	0	0.00%	5/17/10	30	67
							7/4/10	19	
							6/24/10	18	
							9/16/10	13	
							5/28/10	11	
							5/27/10	9	
							9/15/10	9	
							7/5/10	8	
							5/29/10	7	
							9/17/10	5	