

## Assignment: Data Normalize/Standardize

Original Data

	0	1	2	3	4
LOAN	1100.000000	1300.000000	1500.000000	1500.0	1700.000000
MORTDUE	25860.000000	70053.000000	13500.000000	NaN	97800.000000
VALUE	39025.000000	68400.000000	16700.000000	NaN	112000.000000
YOJ	10.500000	7.000000	4.000000	NaN	3.000000
DEROG	0.000000	0.000000	0.000000	NaN	0.000000
DELINQ	0.000000	2.000000	0.000000	NaN	0.000000
CLAGE	94.366667	121.833333	149.466667	NaN	93.333333
NINQ	1.000000	0.000000	1.000000	NaN	0.000000
CLNO	9.000000	14.000000	10.000000	NaN	14.000000
DEBTINC	NaN	NaN	NaN	NaN	NaN

	LOAN	MORTDUE	VALUE	YOJ	DEROG	\
count	5960.000000	5442.000000	5848.000000	5445.000000	5252.000000	
mean	18607.969799	73760.817200	101776.048741	8.922268	0.254570	
std	11207.480417	44457.609458	57385.775334	7.573982	0.846047	
min	1100.000000	2063.000000	8000.000000	0.000000	0.000000	
25%	11100.000000	46276.000000	66075.500000	3.000000	0.000000	
50%	16300.000000	65019.000000	89235.500000	7.000000	0.000000	
75%	23300.000000	91488.000000	119824.250000	13.000000	0.000000	
max	89900.000000	399550.000000	855909.000000	41.000000	10.000000	

	DELINQ	CLAGE	NINQ	CLNO	DEBTINC	\
count	5380.000000	5652.000000	5450.000000	5738.000000	4693.000000	
mean	0.449442	179.766275	1.186055	21.296096	33.779915	
std	1.127266	85.810092	1.728675	10.138933	8.601746	
min	0.000000	0.000000	0.000000	0.000000	0.524499	
25%	0.000000	115.116702	0.000000	15.000000	29.140031	
50%	0.000000	173.466667	1.000000	20.000000	34.818262	
75%	0.000000	231.562278	2.000000	26.000000	39.003141	
max	15.000000	1168.233561	17.000000	71.000000	203.312149	

Normalize the Data

	nor_LOAN	nor_MORTDUE	nor_VALUE	nor_YOJ	nor_DEROG	nor_DELINQ	\
0	0.000000	0.059869	0.036590	0.256098	0.0	0.000000	
1	0.002252	0.171050	0.071234	0.170732	0.0	0.133333	
2	0.004505	0.028773	0.010261	0.097561	0.0	0.000000	
3	0.004505	NaN	NaN	NaN	NaN	NaN	
4	0.006757	0.240856	0.122655	0.073171	0.0	0.000000	

	nor_CLAGE	nor_NINQ	nor_CLNO	nor_DEBTINC	\
0	0.080777	0.058824	0.126761	NaN	
1	0.104289	0.000000	0.197183	NaN	
2	0.127942	0.058824	0.140845	NaN	
3	NaN	NaN	NaN	NaN	
4	0.079893	0.000000	0.197183	NaN	

	nor_LOAN	nor_MORTDUE	nor_VALUE	nor_YOJ	nor_DEROG	\
count	5960.000000	5442.000000	5848.000000	5445.000000	5252.000000	
mean	0.197162	0.180378	0.110597	0.217616	0.025457	
std	0.126210	0.111847	0.067679	0.184731	0.084605	
min	0.000000	0.000000	0.000000	0.000000	0.000000	
25%	0.112613	0.111231	0.068493	0.073171	0.000000	
50%	0.171171	0.158385	0.095807	0.170732	0.000000	
75%	0.250000	0.224976	0.131882	0.317073	0.000000	
max	1.000000	1.000000	1.000000	1.000000	1.000000	

	nor_DELINQ	nor_CLAGE	nor_NINQ	nor_CLNO	nor_DEBTINC	\
count	5380.000000	5652.000000	5450.000000	5738.000000	4693.000000	
mean	0.029963	0.153879	0.069768	0.299945	0.163991	
std	0.075151	0.073453	0.101687	0.142802	0.042418	
min	0.000000	0.000000	0.000000	0.000000	0.000000	
25%	0.000000	0.098539	0.000000	0.211268	0.141111	
50%	0.000000	0.148486	0.058824	0.281690	0.169112	
75%	0.000000	0.198216	0.117647	0.366197	0.189748	
max	1.000000	1.000000	1.000000	1.000000	1.000000	

The data is normalized as we can see from the nor\_ variable descriptive statistics. The minimum of each normalized variable is 0 and the maximum of each is 1 while the original variables had minimum and maximums not within the range of 0 to 1.

I selected to use CLNO to normalize by hand. The following are the results:

```

CLNO  nor_CLNO
0    9.0  0.126761
1   14.0  0.197183
2   10.0  0.140845
3    NaN     NaN
4   14.0  0.197183

count    5738.000000
mean      21.296096
std       10.138933
min        0.000000
25%       15.000000
50%       20.000000
75%       26.000000
max       71.000000
Name: CLNO, dtype: float64

TEMP = ( X_TEST["CLNO"] - 0 ) / ( 71 - 0 )
X_TEST = X_TEST.assign( calc_CLNO = TEMP.values )
print( X_TEST.head() )

CLNO  nor_CLNO  calc_CLNO
0    9.0  0.126761  0.126761
1   14.0  0.197183  0.197183
2   10.0  0.140845  0.140845
3    NaN     NaN     NaN
4   14.0  0.197183  0.197183

```

The calculated normalization of CLNO is represented by calc\_CLNO. The output for the first calculated 5 rows are the same values as that of nor\_CLNO, which was built from using MinMaxScaler.

## Complete Normalized Dataframe

	LOAN	MORTDUE	VALUE	YOJ	DEROG	DELINQ	CLAGE	NINQ	CLNO	\
0	1100	25860.0	39025.0	10.5	0.0	0.0	94.366667	1.0	9.0	
1	1300	70053.0	68400.0	7.0	0.0	2.0	121.833333	0.0	14.0	
2	1500	13500.0	16700.0	4.0	0.0	0.0	149.466667	1.0	10.0	
3	1500	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	1700	97800.0	112000.0	3.0	0.0	0.0	93.333333	0.0	14.0	

  

	DEBTINC	nor_LOAN	nor_MORTDUE	nor_VALUE	nor_YOJ	nor_DEROG	nor_DELINQ	\
0	NaN	0.000000	0.059869	0.036590	0.256098	0.0	0.000000	
1	NaN	0.002252	0.171050	0.071234	0.170732	0.0	0.133333	
2	NaN	0.004505	0.028773	0.010261	0.097561	0.0	0.000000	
3	NaN	0.004505	NaN	NaN	NaN	NaN	NaN	
4	NaN	0.006757	0.240856	0.122655	0.073171	0.0	0.000000	

  

	nor_CLAGE	nor_NINQ	nor_CLNO	nor_DEBTINC	TARGET_BAD_FLAG	\
0	0.080777	0.058824	0.126761	NaN	1	
1	0.104289	0.000000	0.197183	NaN	1	
2	0.127942	0.058824	0.140845	NaN	1	
3	NaN	NaN	NaN	NaN	1	
4	0.079893	0.000000	0.197183	NaN	0	

  

	TARGET_LOSS_AMT	REASON	JOB
0	641.0	HomeImp	Other
1	1109.0	HomeImp	Other
2	767.0	HomeImp	Other
3	1425.0	NaN	NaN
4	NaN	HomeImp	Office

## Standardize the Data

```

std_LOAN std_MORTDUE std_VALUE std_YOJ std_DEROG std_DELINQ \
0 -1.562299 -1.077548 -1.093588 0.208329 -0.300922 -0.398738
1 -1.544453 -0.083409 -0.581658 -0.253822 -0.300922 1.375631
2 -1.526606 -1.355591 -1.482655 -0.649951 -0.300922 -0.398738
3 -1.526606 NaN NaN NaN NaN NaN
4 -1.508759 0.540771 0.178177 -0.781994 -0.300922 -0.398738

```

```

std_CLAGE std_NINQ std_CLNO std_DEBTINC
0 -0.995304 -0.107639 -1.212866 NaN
1 -0.675189 -0.686169 -0.719675 NaN
2 -0.353132 -0.107639 -1.114228 NaN
3 NaN NaN NaN NaN
4 -1.007348 -0.686169 -0.719675 NaN

```

```

count std_LOAN std_MORTDUE std_VALUE std_YOJ std_DEROG \
mean -1.522868e-15 -1.790189e-16 -4.884089e-16 -3.400402e-16 -4.777426e-17
std 1.000084e+00 1.000092e+00 1.000086e+00 1.000092e+00 1.000095e+00
min -1.562299e+00 -1.612871e+00 -1.634274e+00 -1.178124e+00 -3.009218e-01
25% -6.699632e-01 -6.182821e-01 -6.221681e-01 -7.819944e-01 -3.009218e-01
50% -2.059485e-01 -1.966507e-01 -2.185493e-01 -2.538222e-01 -3.009218e-01
75% 4.186868e-01 3.987801e-01 3.145334e-01 5.384362e-01 -3.009218e-01
max 6.361645e+00 7.328759e+00 1.314259e+01 4.235642e+00 1.151988e+01

```

```

count std_DELINQ std_CLAGE std_NINQ std_CLNO std_DEBTINC
mean 2.319334e-15 -2.628137e-16 -4.635232e-16 -5.634314e-17 -3.706579e-16
std 1.000093e+00 1.000088e+00 1.000092e+00 1.000087e+00 1.000107e+00
min -3.987384e-01 -2.095117e+00 -6.861694e-01 -2.100611e+00 -3.866536e+00
25% -3.987384e-01 -7.534696e-01 -6.861694e-01 -6.210362e-01 -5.394693e-01
50% -3.987384e-01 -7.341987e-02 -1.076386e-01 -1.278447e-01 1.207263e-01
75% -3.987384e-01 6.036654e-01 4.708922e-01 4.639851e-01 6.072932e-01
max 1.290903e+01 1.152026e+01 9.148855e+00 4.902709e+00 1.971115e+01

```

The data is standardized as we can see from the std\_ variable descriptive statistics. The range of values between the standardized values are -2 and 19 while the original variables had a much larger range of minimum and maximum values from 0 to the ten thousands respectively. Most standardized data range from -3 to 3, but the outliers were not removed from this data prior to standardization, so the range is a lot larger.

I selected to use CLNO to normalize by hand. The following are the results:

```

CLNO std_CLNO
0 9.0 -1.212866
1 14.0 -0.719675
2 10.0 -1.114228
3 NaN NaN
4 14.0 -0.719675

```

```

count 5738.000000
mean 21.296096
std 10.138933
min 0.000000
25% 15.000000
50% 20.000000
75% 26.000000
max 71.000000
Name: CLNO, dtype: float64

```

```

TEMP = ( Y_TEST["CLNO"] - 21.296096 ) / 10.138933
Y_TEST = Y_TEST.assign( calc_CLNO = TEMP.values )
print( Y_TEST.head() )

```

```

CLNO std_CLNO calc_CLNO
0 9.0 -1.212866 -1.212760
1 14.0 -0.719675 -0.719612
2 10.0 -1.114228 -1.114131
3 NaN NaN NaN
4 14.0 -0.719675 -0.719612

```

The calculated standardization of CLNO is represented by calc\_CLNO. The calculated output for the first 5 rows are very similar values as that of std\_CLNO, which was built from using StandardScaler.

### Complete Standardized Dataframe

	LOAN	MORTDUE	VALUE	YOJ	DEROG	DELINQ	CLAGE	NINQ	CLNO	\
0	1100	25860.0	39025.0	10.5	0.0	0.0	94.366667	1.0	9.0	
1	1300	70053.0	68400.0	7.0	0.0	2.0	121.833333	0.0	14.0	
2	1500	13500.0	16700.0	4.0	0.0	0.0	149.466667	1.0	10.0	
3	1500	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
4	1700	97800.0	112000.0	3.0	0.0	0.0	93.333333	0.0	14.0	

  

	DEBTINC	std_LOAN	std_MORTDUE	std_VALUE	std_YOJ	std_DEROG	std_DELINQ	\
0	NaN	-1.562299	-1.077548	-1.093588	0.208329	-0.300922	-0.398738	
1	NaN	-1.544453	-0.083409	-0.581658	-0.253822	-0.300922	1.375631	
2	NaN	-1.526606	-1.355591	-1.482655	-0.649951	-0.300922	-0.398738	
3	NaN	-1.526606	NaN	NaN	NaN	NaN	NaN	
4	NaN	-1.508759	0.540771	0.178177	-0.781994	-0.300922	-0.398738	

  

	std_CLAGE	std_NINQ	std_CLNO	std_DEBTINC	TARGET_BAD_FLAG	\
0	-0.995304	-0.107639	-1.212866	NaN	1	
1	-0.675189	-0.686169	-0.719675	NaN	1	
2	-0.353132	-0.107639	-1.114228	NaN	1	
3	NaN	NaN	NaN	NaN	1	
4	-1.007348	-0.686169	-0.719675	NaN	0	

  

	TARGET_LOSS_AMT	REASON	JOB
0	641.0	HomeImp	Other
1	1109.0	HomeImp	Other
2	767.0	HomeImp	Other
3	1425.0	NaN	NaN
4	NaN	HomeImp	Office

## BINGO BONUS – Remove outliers before data transformation

### Descriptive Statistics of Truncated Dataframe

	count	mean	std	min	25%	50%	75%	max
TARGET_BAD_FLAG	5960.0	0.199497	0.399656	0.000000	0.000000	0.000000	0.000000	1.0
TARGET_LOSS_AMT	1189.0	13414.576955	10839.455965	224.000000	5639.000000	11003.000000	17634.000000	78987.0
O_LOAN	5960.0	0.015940	0.125252	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_LOAN	5960.0	18362.256711	10148.976515	1100.000000	11100.000000	16300.000000	23300.000000	52230.0
O_MORTDUE	5960.0	0.014765	0.120621	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_MORTDUE	5442.0	73002.219441	41139.438729	2063.000000	46276.000000	65019.000000	91488.000000	207134.0
O_VALUE	5960.0	0.014933	0.121295	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_VALUE	5848.0	100704.418953	50961.444513	8000.000000	66075.500000	89235.500000	119824.250000	273933.0
O_YOJ	5960.0	0.002852	0.053336	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_YOJ	5445.0	8.909780	7.531208	0.000000	3.000000	7.000000	13.000000	32.0
O_DEROG	5960.0	0.012081	0.109255	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_DEROG	5252.0	0.218012	0.616441	0.000000	0.000000	0.000000	0.000000	3.0
O_DELINQ	5960.0	0.015101	0.121964	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_DELINQ	5380.0	0.411338	0.924025	0.000000	0.000000	0.000000	0.000000	4.0
O_CLAGE	5960.0	0.005034	0.070775	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_CLAGE	5652.0	178.832933	81.085167	0.000000	115.116702	173.466667	231.562278	437.0
O_NINQ	5960.0	0.020302	0.141043	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_NINQ	5450.0	1.125872	1.472925	0.000000	0.000000	1.000000	2.000000	6.0
O_CLNO	5960.0	0.006208	0.078553	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_CLNO	5738.0	21.254444	9.990744	0.000000	15.000000	20.000000	26.000000	52.0
O_DEBTINC	5960.0	0.004530	0.067160	0.000000	0.000000	0.000000	0.000000	1.0
TRUNC_DEBTINC	4693.0	33.641411	7.534484	0.524499	29.140031	34.818262	39.003141	60.0

## Normalized Truncated Descriptive Statistics

	nor_TRUNC_LOAN	nor_TRUNC_MORTDUE	nor_TRUNC_VALUE	nor_TRUNC_YOJ	\
count	5960.000000	5442.000000	5848.000000	5445.000000	
mean	0.337615	0.345925	0.348601	0.278431	
std	0.198494	0.200611	0.191633	0.235350	
min	0.000000	0.000000	0.000000	0.000000	
25%	0.195580	0.215599	0.218384	0.093750	
50%	0.297281	0.306996	0.305474	0.218750	
75%	0.434187	0.436068	0.420498	0.406250	
max	1.000000	1.000000	1.000000	1.000000	

  

	nor_TRUNC_DEROG	nor_TRUNC_DELIQ	nor_TRUNC_CLAGE	nor_TRUNC_NINQ	\
count	5252.000000	5380.000000	5652.000000	5450.000000	
mean	0.072671	0.102835	0.409229	0.187645	
std	0.205480	0.231006	0.185550	0.245488	
min	0.000000	0.000000	0.000000	0.000000	
25%	0.000000	0.000000	0.263425	0.000000	
50%	0.000000	0.000000	0.396949	0.166667	
75%	0.000000	0.000000	0.529891	0.333333	
max	1.000000	1.000000	1.000000	1.000000	

  

	nor_TRUNC_CLNO	nor_TRUNC_DEBTINC
count	5738.000000	4693.000000
mean	0.408739	0.556816
std	0.192130	0.126682
min	0.000000	0.000000
25%	0.288462	0.481131
50%	0.384615	0.576603
75%	0.500000	0.646966
max	1.000000	1.000000

The normalized truncated variables are represented by the prefix `nor_TRUNC_`. Similar to the normalized dataframe without the removed outliers, the minimum and maximum are 0 and 1 respectively. The mean for each of the normalized truncated variables are larger than the corresponding mean of the normalized variables. This makes sense since the maximum value each value is divided by during normalization is smaller, so the resulting normalized value is bigger.

## Complete Normalized Truncated Dataframe

	TRUNC_LOAN	TRUNC_MORTDUE	TRUNC_VALUE	TRUNC_YOJ	TRUNC_DEROG	\
0	1100	25860.0	39025.0	10.5	0.0	
1	1300	70053.0	68400.0	7.0	0.0	
2	1500	13500.0	16700.0	4.0	0.0	
3	1500	NaN	NaN	NaN	NaN	
4	1700	97800.0	112000.0	3.0	0.0	

  

	TRUNC_DELIQ	TRUNC_CLAGE	TRUNC_NINQ	TRUNC_CLNO	TRUNC_DEBTINC	\
0	0.0	94.366667	1.0	9.0	NaN	
1	2.0	121.833333	0.0	14.0	NaN	
2	0.0	149.466667	1.0	10.0	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	0.0	93.333333	0.0	14.0	NaN	

  

	nor_TRUNC_LOAN	nor_TRUNC_MORTDUE	nor_TRUNC_VALUE	nor_TRUNC_YOJ	\
0	0.000000	0.116043	0.116665	0.328125	
1	0.003912	0.331544	0.227125	0.218750	
2	0.007823	0.055771	0.032715	0.125000	
3	0.007823	NaN	NaN	NaN	
4	0.011735	0.466848	0.391076	0.093750	

  

	nor_TRUNC_DEROG	nor_TRUNC_DELIQ	nor_TRUNC_CLAGE	nor_TRUNC_NINQ	\
0	0.0	0.0	0.215942	0.166667	
1	0.0	0.5	0.278795	0.000000	
2	0.0	0.0	0.342029	0.166667	
3	NaN	NaN	NaN	NaN	
4	0.0	0.0	0.213577	0.000000	

  

	nor_TRUNC_CLNO	nor_TRUNC_DEBTINC	TARGET_BAD_FLAG	TARGET_LOSS_AMT	\
0	0.173077	NaN	1	641.0	
1	0.269231	NaN	1	1109.0	
2	0.192308	NaN	1	767.0	
3	NaN	NaN	1	1425.0	
4	0.269231	NaN	0	NaN	

  

	REASON	JOB
0	HomeImp	Other
1	HomeImp	Other
2	HomeImp	Other
3	NaN	NaN
4	HomeImp	Office



## Standardized Truncated Descriptive Statistics

	std_TRUNC_LOAN	std_TRUNC_MORTDUE	std_TRUNC_VALUE	std_TRUNC_YOJ	\
count	5.960000e+03	5.442000e+03	5.848000e+03	5.445000e+03	
mean	5.472133e-16	-1.916471e-16	-1.173442e-16	-8.286403e-17	
std	1.000084e+00	1.000092e+00	1.000086e+00	1.000092e+00	
min	-1.701029e+00	-1.724519e+00	-1.819264e+00	-1.183156e+00	
25%	-7.156255e-01	-6.497093e-01	-6.795702e-01	-7.847774e-01	
50%	-2.032155e-01	-1.940705e-01	-2.250701e-01	-2.536054e-01	
75%	4.865671e-01	4.493858e-01	3.752144e-01	5.431527e-01	
max	3.337340e+00	3.260718e+00	3.399499e+00	3.066220e+00	

  

	std_TRUNC_DEROG	std_TRUNC_DELINQ	std_TRUNC_CLAGE	std_TRUNC_NINQ	\
count	5.252000e+03	5.380000e+03	5.652000e+03	5.450000e+03	
mean	4.261633e-17	1.468528e-15	5.788796e-17	7.645261e-16	
std	1.000095e+00	1.000093e+00	1.000088e+00	1.000092e+00	
min	-3.536965e-01	-4.452007e-01	-2.205690e+00	-7.644481e-01	
25%	-3.536965e-01	-4.452007e-01	-7.858634e-01	-7.644481e-01	
50%	-3.536965e-01	-4.452007e-01	-6.618647e-02	-8.546470e-02	
75%	-3.536965e-01	-4.452007e-01	6.503534e-01	5.935187e-01	
max	4.513414e+00	3.884089e+00	3.184182e+00	3.309452e+00	

  

	std_TRUNC_CLNO	std_TRUNC_DEBTINC
count	5.738000e+03	4.693000e+03
mean	-1.502032e-15	-1.330144e-16
std	1.000087e+00	1.000107e+00
min	-2.127599e+00	-4.395847e+00
25%	-6.260784e-01	-5.975007e-01
50%	-1.255716e-01	1.562119e-01
75%	4.750366e-01	7.117011e-01
max	3.077672e+00	3.498766e+00

The standardized truncated variables are represented with the prefix std\_TRUNC\_. The minimum and maximum values of the variables range from -4 to 5 respectively. This range is a lot smaller than the range of the non-truncated minimum and maximum values of -2 to 19. Without the large outliers, the mean is smaller because it is not as positively skewed and the standard deviation is also smaller because there is less variance, resulting in a smaller range when the values are standardized.

## Complete Normalized Truncated Dataframe

	TRUNC_LOAN	TRUNC_MORTDUE	TRUNC_VALUE	TRUNC_YOJ	TRUNC_DEROG	\
0	1100	25860.0	39025.0	10.5	0.0	
1	1300	70053.0	68400.0	7.0	0.0	
2	1500	13500.0	16700.0	4.0	0.0	
3	1500	NaN	NaN	NaN	NaN	
4	1700	97800.0	112000.0	3.0	0.0	

  

	TRUNC_DELINQ	TRUNC_CLAGE	TRUNC_NINQ	TRUNC_CLNO	TRUNC_DEBTINC	\
0	0.0	94.366667	1.0	9.0	NaN	
1	2.0	121.833333	0.0	14.0	NaN	
2	0.0	149.466667	1.0	10.0	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	0.0	93.333333	0.0	14.0	NaN	

  

	std_TRUNC_LOAN	std_TRUNC_MORTDUE	std_TRUNC_VALUE	std_TRUNC_YOJ	\
0	-1.701029	-1.146018	-1.210419	0.211170	
1	-1.681321	-0.071695	-0.633953	-0.253605	
2	-1.661613	-1.446488	-1.648533	-0.651984	
3	-1.661613	NaN	NaN	NaN	
4	-1.641905	0.602829	0.221668	-0.784777	

  

	std_TRUNC_DEROG	std_TRUNC_DELINQ	std_TRUNC_CLAGE	std_TRUNC_NINQ	\
0	-0.353696	-0.445201	-1.041790	-0.085465	
1	-0.353696	1.719444	-0.703022	-0.764448	
2	-0.353696	-0.445201	-0.362198	-0.085465	
3	NaN	NaN	NaN	NaN	
4	-0.353696	-0.445201	-1.054535	-0.764448	

  

	std_TRUNC_CLNO	std_TRUNC_DEBTINC	TARGET_BAD_FLAG	TARGET_LOSS_AMT	\
0	-1.226687	NaN	1	641.0	
1	-0.726180	NaN	1	1109.0	
2	-1.126585	NaN	1	767.0	
3	NaN	NaN	1	1425.0	
4	-0.726180	NaN	0	NaN	

  

	REASON	JOB
0	HomeImp	Other
1	HomeImp	Other
2	HomeImp	Other
3	NaN	NaN
4	HomeImp	Office