New Variables for DSO562 Project 2 (Credit Cards and Cell Phone Application Data)

Note: Variables with Red Color are new-created variables.

<u>Note: variables with Red Color are new-created variables.</u>					
Description of Variables	# Varia bles Creat ed	Approximate CPU Time (Minute)	Names of New Columns		
Day of Week of Each Application (including Monday to Sunday, 7 unique values)	1	<0.1 min	dow		
Target Encoding for dow column (transferring categorical variables to numerical variables)	1	<0.1min	Dow_risk		
Age of the applicant when they deliver application (treat missing values as 0)	1	<0.1m	age		
Age level of the age filed (7 categories, treat missing values as "MISSING" type)	1	<0.1m	Age_level		
Target Encoding for age column (transferring categorical variables to numerical variables)	1	<0.1m	Age_level_risk		
Full name of applicant (including first and last name)	1	<0.1m	name		
Full address of applicant (including address and 5-digit zip code)	1	<0.1m	Fulladdress		
Combination of name and date of birth, address and home phone number of the applicant	3	<0.1m	Name_dob Name_fulladdress Name_homephone		
Combination of full address and dob, home phone number of the applicant	2	<0.1m	Fulladdress_dob Fulladdress_homephone		
Combination of home phone number and date of	2	<0.1m	Dob_homephone Homephone_name_dob		

birth, name & date of birth			
of the applicant			
Combination of age level			
and address, home phone	2	<0.1m	Avg_level_address
number of the applicant		VO.1 111	Avg_level_homephone
Combination of social			
security number and {first			
name, last name, address,			Ssn_ {firstname, lastname, address,
5-digit zip code, date of			zip5, dob, homephone, age_level,
birth, home phone	12	<0.1m	name, fulladdress, name _dob,
number, age level, name,			age_level_address,
full address, name & date			age_level_homephone}
of birth, age level &			0 = = ' '
address, age level & home			
phone number}			
Number of days since an			
application with that entity	ļ		{ssn, address, dob, homephone, name, fulladdress, name_dob, name_fulladdress, name_homephone, fulladdress dob,
has been seen. Entities are			
{social security number,			
address, date of birth,			
home phone number,			
name, full address, name	ĺ		
& date of birth/ full			
address/ home phone			fulladdress_homephone,
number, full address &			dob homephone,
date of birth/ home phone			homephone name dob,
number, date of birth &			age_level_address,
home phone number/	27	About 5 mins	age_level_address, age_level_homephone, ssn_firstname,
home phone number with			
name, age level & address/			ssn_lastname, ssn_address, ssn_zip5,
home phone number,			ssn_dob, ssn_homephone,
social security number &			ssn_age_level, ssn_name,
first name/ last name/			ssn_fulladdress, ssn_name_dob,
address/ 5-digit zip code/			ssn_age_level_address,ssn_age_level_
date of birth/ home phone			homephone}_day_since
number/ age level/ name/			
full address/ name & date			
of birth/ age level &			
address/ age level & home			
phone number}			
Number of applications at			{ssn, address, dob, homephone, name,
that entity over the past n	162	About 47 min	fulladdress, name dob,
days. Entities are {social			name_fulladdress, name_homephone,
security number, address,			fulladdress dob,

date of birth, home phone			fulladdress_homephone,
number, name, full			dob_homephone,
address, name & date of			homephone_name_dob,
birth/ full address/ home			age_level_address,
phone number, full			age_level_homephone, ssn_firstname,
address & date of birth/			ssn_lastname, ssn_address, ssn_zip5,
home phone number, date			ssn_dob, ssn_homephone,
of birth & home phone			ssn_age_level, ssn_name,
number/ home phone			ssn_fulladdress, ssn_name_dob,
number with name, age			ssn_age_level_address,ssn_age_level_
level & address/ home			homephone}_count_{0,1,3,7,14,30}
phone number, social			
security number & first			
name/ last name/ address/			
5-digit zip code/ date of			
birth/ home phone			
number/ age level/ name/			
full address/ name & date			
of birth/ age level &			
address/ age level & home			
phone number}, n is			
{0,1,3,7,14,30}			
Number of applications at			
that entity seen in a short			
time window (today or			
past day} compared to			{ssn, address, dob, homephone, name,
number of applications at			fulladdress, name_dob,
that entity over a longer			name_fulladdress, name_homephone,
time window {past			fulladdress_dob,
3,7,14,30 days}. Entities			fulladdress homephone,
are {social security			dob homephone,
number, address, date of			homephone name dob,
birth, home phone	0.4.6		age level address,
number, name, full	216	About 22.58 s	age_level_homephone, ssn_firstname,
address, name & date of			ssn_lastname, ssn_address, ssn_zip5,
birth/ full address/ home			ssn dob, ssn homephone,
phone number, full			ssn age level, ssn name,
address & date of birth/			ssn_fulladdress, ssn_name_dob,
home phone number, date			ssn_age_level_address,ssn_age_level_
of birth & home phone			homephone}_count_{0,1}_by_{3,7,14,
number/ home phone			30}
number with name, age			
level & address/ home			
phone number, social			
			<u>I</u>

security number & first name/ last name/ address/ 5-digit zip code/ date of birth/ home phone number/ age level/ name/ full address/ name & date of birth/ age level & address/ age level & home phone number}			
Number of unique identy in the specific field during the past {1,3,7,14,30,60} days. Both entity and field are {social security number, full address, name & date of birth/ full address, full address & date of birth, date of birth & home phone number, age level & address/ home phone number, social security number & last name/ 5-digit zip code/ age level/ name/ full address/ name & date of birth/ age level & address/ age level & home phone number}, but with different values.	1440	About 88 mins	{ssn, fulladdress, name_dob, name_fulladdress, fulladdress_dob, dob_homephone, age_level_address, age_level_homephone, ssn_lastname, ssn_zip5, ssn_age_level, ssn_name, ssn_fulladdress, ssn_name_dob, ssn_age_level_address, ssn_age_level_homephone}_unique_count_in_{ ssn, fulladdress, name_dob, name_fulladdress, fulladdress_dob, dob_homephone, age_level_address, age_level_homephone, ssn_lastname, ssn_zip5, ssn_age_level, ssn_name, ssn_fulladdress, ssn_name_dob, ssn_age_level_address, ssn_age_level_homephone}_for_{1,3,7,14,30,60}
Total	1873	About 140 mins	