

I. List of final variables along with respective univariate KS's (filter score)

Wrapper Order	Variable	Filter Score
1	max_count_by_address_30	0.355872
2	max_count_by_ssn_dob_7	0.225774
3	max_count_by_homephone_7	0.232724
4	zip5_unique_count_for_fulladdress_dob_1	0.219167
5	name_day_since	0.209358
6	max_count_by_fulladdress_30	0.356094
7	max_count_by_ssn_firstname_30	0.236333
8	max_count_by_homephone_30	0.226131
9	address_count_30	0.32996
10	max_count_by_address_7	0.343199
11	address_day_since	0.331458
12	max_count_by_address_3	0.328576
13	max_count_by_fulladdress_3	0.328381
14	address_count_14	0.320782
15	fulladdress_count_14	0.320202
16	max_count_by_address_1	0.314053
17	max_count_by_fulladdress_1	0.313893
18	address_count_7	0.301069
19	fulladdress_count_7	0.300964
20	address_count_0_by_30	0.288978
21	fulladdress_unique_count_for_ssn_homephone_60	0.287489
22	fulladdress_unique_count_for_name_homephone_60	0.287085
23	fulladdress_unique_count_for_homephone_name_do...	0.286135
24	fulladdress_unique_count_for_dob_homephone_60	0.286088
25	fulladdress_unique_count_for_ssn_name_60	0.284964
26	fulladdress_unique_count_for_ssn_lastname_60	0.284933
27	fulladdress_unique_count_for_ssn_firstname_60	0.284925
28	fulladdress_unique_count_for_ssn_60	0.284917
29	fulladdress_unique_count_for_ssn_name_dob_60	0.283268
30	fulladdress_unique_count_for_ssn_dob_60	0.283222
31	fulladdress_unique_count_for_ssn_homephone_30	0.282986
32	fulladdress_unique_count_for_name_60	0.282725
33	fulladdress_unique_count_for_name_homephone_30	0.282455
34	fulladdress_unique_count_for_homephone_name_do...	0.282018
35	fulladdress_unique_count_for_dob_homephone_30	0.281967

Wrapper Order	Variable	Filter Score
36	fulladdress_unique_count_for_name_dob_60	0.281498
37	fulladdress_unique_count_for_dob_60	0.28145
38	fulladdress_unique_count_for_ssn_name_30	0.281084
39	fulladdress_unique_count_for_ssn_firstname_30	0.281032
40	fulladdress_unique_count_for_ssn_lastname_30	0.281032
41	fulladdress_unique_count_for_ssn_30	0.280996
42	fulladdress_unique_count_for_ssn_name_dob_30	0.280299
43	fulladdress_unique_count_for_ssn_dob_30	0.280248
44	address_count_0_by_14	0.280014
45	fulladdress_count_0_by_14	0.278876
46	fulladdress_unique_count_for_name_30	0.278808
47	fulladdress_count_3	0.278379
48	address_count_3	0.278345
49	fulladdress_unique_count_for_name_dob_30	0.278146
50	fulladdress_unique_count_for_dob_30	0.278025

II. Model performance vs # variables

