

-Rukmini Annadata

Task A: Identify the missing data and use appropriate method to deal with it. (Remove columns/or replace it with an appropriate value)

Task B: Identify if there are outliers in the dataset

Task C: Analyze data imbalance. Find the ratio of data imbalance

Task D (EDA): Univariate Analysis/ Segmented Uni/Bivariate Analysis

A	B	C	D	E	F	G	H	I	J	
SK_ID_PREV	SK_ID_CURR	NAME_CONTRACT_TYPE	AMT_ANNUITY	AMT_APPLICATION	AMT_CREDIT	AMT_GOODS_PRICE	NAME_CASH_LOAN_PURPOSE	NAME_CONTRACT_STATUS	DAYS_DECISION	NAME_PAYM
2030495	271877	Consumer loans	1730.43	17145	17145	17145	XAP	Approved		-73 Cash through
2802425	108129	Cash loans	25188.615	607500	679671	607500	XNA	Approved		-164 XNA
2523466	122040	Cash loans	15060.735	112500	136444.5	112500	XNA	Approved		-301 Cash through
2819243	176158	Cash loans	47041.335	450000	470790	450000	XNA	Approved		-512 Cash through
1784265	202054	Cash loans	31924.395	337500	404055	337500	Repairs	Refused		-781 Cash through
1383531	199383	Cash loans	23703.93	315000	340573.5	315000	Everyday expenses	Approved		-684 Cash through
2315218	175704	Cash loans	11250	0	0	111511.58	XNA	Canceled		-14 XNA
1656711	296299	Cash loans	11250	0	0	111511.58	XNA	Canceled		-21 XNA
2367563	342292	Cash loans	11250	0	0	111511.58	XNA	Canceled		-386 XNA
2579447	334349	Cash loans	11250	0	0	111511.58	XNA	Canceled		-57 XNA
1715995	447712	Cash loans	11368.62	270000	335754	270000	XNA	Approved		-735 Cash through
2257824	161140	Cash loans	13832.775	211500	246397.5	211500	XNA	Approved		-815 Cash through
2330894	258628	Cash loans	12165.21	148500	174361.5	148500	XNA	Approved		-860 Cash through
1397919	321676	Consumer loans	7654.86	53779.5	57564	53779.5	XAP	Approved		-408 Cash through
2273188	270658	Consumer loans	9644.22	26550	27252	26550	XAP	Approved		-726 Cash through
1232483	151612	Consumer loans	21307.455	126490.5	119853	126490.5	XAP	Approved		-699 Cash through
2163253	154602	Consumer loans	4187.34	26955	27297	26955	XAP	Approved		-1473 Cash through
1285768	142748	Revolving loans	9000	180000	180000	180000	XAP	Approved		-336 XNA
2933109	396305	Cash loans	10181.7	180000	180000	180000	XNA	Approved		-700 Cash through
1173070	199178	Cash loans	4666.5	45000	49455	45000	Everyday expenses	Refused		-584 XNA
1506815	166490	Cash loans	25454.025	450000	491580	450000	XNA	Refused		-401 Cash through
1182516	267782	Cash loans	20361.6	405000	451777.5	405000	XNA	Approved		-686 Cash through
1172842	302212	Cash loans	11250	0	0	111511.58	XNA	Refused		-239 XNA
1172937	302212	Cash loans	39475.305	1129500	1277104.5	1129500	XNA	Refused		-594 Cash through
1555330	199353	Cash loans	11250	0	0	111511.58	XNA	Canceled		-202 XNA
1543131	275707	Cash loans	22619.52	229500	241920	229500	XNA	Approved		-370 Cash through
2536650	338725	Cash loans	16708.32	369000	369000	369000	XNA	Approved		-1487 XNA
1676258	433469	Cash loans	22242.825	247500	268083	247500	XNA	Approved		-1883 XNA
2075578	418383	Consumer loans	7656.705	74610	65610	74610	XAP	Approved		-2702 Cash through
1583704	315664	Cash loans	11250	0	0	111511.58	XNA	Refused		-430 XNA
2138466	170289	Consumer loans	3286.485	32562	32562	32562	XAP	Approved		-307 Cash through

1	SK_ID_CURR	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	NAME_TYPE_SUITE	NAME_INCOME_TYPE	NAME_EDUCATION_TYPE	NAME_FAMILY_STATUS	NAME_HOUSING_TYPE
2	100002	Cash loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
3	100003	Cash loans	F	N	N	Family	State servant	Higher education	Married	House / apartment
4	100004	Revolving loans	M	Y	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
5	100006	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Civil marriage	House / apartment
6	100007	Cash loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
7	100008	Cash loans	M	N	Y	Spouse, partner	State servant	Secondary / secondary special	Married	House / apartment
8	100009	Cash loans	F	Y	Y	Unaccompanied	Commercial associate	Higher education	Married	House / apartment
9	100010	Cash loans	M	Y	Y	Unaccompanied	State servant	Higher education	Married	House / apartment
10	100011	Cash loans	F	N	Y	Children	Pensioner	Secondary / secondary special	Married	House / apartment
11	100012	Revolving loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
12	100014	Cash loans	F	N	Y	Unaccompanied	Working	Higher education	Married	House / apartment
13	100015	Cash loans	F	N	Y	Children	Pensioner	Secondary / secondary special	Married	House / apartment
14	100016	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
15	100017	Cash loans	M	Y	N	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
16	100018	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
17	100019	Cash loans	M	Y	Y	Family	Working	Secondary / secondary special	Single / not married	Rented apartment
18	100020	Cash loans	M	N	N	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
19	100021	Revolving loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
20	100022	Revolving loans	F	N	Y	Other_A	Working	Secondary / secondary special	Widow	House / apartment
21	100023	Cash loans	F	N	Y	Unaccompanied	State servant	Higher education	Single / not married	House / apartment
22	100024	Revolving loans	M	Y	Y	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
23	100025	Cash loans	F	Y	Y	Unaccompanied	Commercial associate	Secondary / secondary special	Married	House / apartment
24	100026	Cash loans	F	N	N	Unaccompanied	Working	Secondary / secondary special	Married	Rented apartment
25	100027	Cash loans	F	N	Y	Unaccompanied	Pensioner	Secondary / secondary special	Married	House / apartment
26	100029	Cash loans	M	Y	N	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
27	100030	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Married	House / apartment
28	100031	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Widow	House / apartment
29	100032	Cash loans	M	N	Y	Family	Working	Secondary / secondary special	Married	House / apartment
30	100033	Cash loans	M	Y	Y	Unaccompanied	State servant	Higher education	Single / not married	House / apartment
31	100034	Revolving loans	M	N	Y	Unaccompanied	Working	Higher education	Single / not married	With parents
32	100035	Cash loans	F	N	Y	Unaccompanied	Commercial associate	Secondary / secondary special	Civil marriage	House / apartment

	A	B	C	D	E	F	G	H	I	J
1	Column1	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	AGE	Years of Experience	DAYS_REGISTRATION	
2	AMT_INCOME_TOTAL									
3	AMT_CREDIT	0.138771991	1							
4	AMT_ANNUITY	0.172323385	0.758630818	1						
5	AMT_GOODS_PRICE	0.139851962	0.986091461	0.762111661	1					
6	REGION_POPULATION_RELATIVE	0.063753773	0.085766155	0.103129834	0.088211435					
7	AGE	-0.022068053	0.047325658	-0.029075319	0.045591016	0.040791045	1			
8	Years of Experience	0.056397807	0.061069512	0.108456229	0.05868915	-0.001756113	-0.62576	1		
9	DAYS_REGISTRATION	0.022615292	-0.009651473	0.046069972	0.011784029	-0.053437007	-0.32622	0.207225792		
10	DAYS_ID_PUBLISH	0.011308948	0.009622066	0.029777004	0.006694873	-0.006274449	-0.2624	0.278197383		0.103081977
11	DAYS_LAST_PHONE_CHANGE	-0.021033188	-0.086513064	-0.062828558	-0.090376009	-0.050906304	-0.09098	-0.014022299		0.057550976
12	CNT_CHILDREN	0.010859378	0.003198692	0.029864456	-0.000480755	-0.022069485	-0.36229	0.247940972		0.186385809
13	FLAG_EMP_PHONE	0.056200683	0.059834852	0.10782093	0.057424598	-0.00177992	-0.62938	0.999778134		0.209151151
14	FLAG_WORK_PHONE	-0.017743596	-0.015600653	-0.01617311	0.006711513	-0.019230713	-0.18337	0.231468641		0.065416932
15	FLAG_CONT_MOBILE	-0.001256044	0.007234092	0.00193326	0.007212967	-0.006266044	0.001069	0.001372442		0.001151174
16	FLAG_PHONE	-0.001455611	0.022352027	0.006871606	0.03977638	0.087418157	0.019991	-0.002220679		-0.062803557
17	FLAG_EMAIL	0.040258362	0.014676665	0.068830805	0.01523224	0.037828141	-0.09796	0.069920513		0.038983263
18	CNT_FAM_MEMBERS	0.012860451	0.063576791	0.084337728	0.061768071	-0.021304144	-0.32539	0.249539461		0.180633533
19	REGION_RATING_CLIENT	-0.076774034	-0.085153839	-0.111148215	-0.086270142	-0.517422506	-0.02095	-0.025473472		0.082226848
20	REGION_RATING_CLIENT_W_CITY	-0.083010094	-0.094290468	-0.125095857	-0.094486589	-0.516549676	-0.01919	-0.028029941		0.075367931
21	HOUS_APPR_PROCESS_START	0.034242475	0.0475839	0.051590644	0.057030173	0.159801545	-0.0875	0.088460254		-0.011245182
22	REG_REGION_NOT_LIVE_REGION	0.02601143	0.019845912	0.039691726	0.021211135	-0.003949193	-0.0563	0.030564113		0.03098896
23	REG_REGION_NOT_WORK_REGION	0.052060897	0.044538701	0.075012301	0.04523454	0.053100004	-0.09329	0.103890952		0.038200345
24	LIVE_REGION_NOT_WORK_REGION	0.048095575	0.045281565	0.069228736	0.045342171	0.075254616	-0.07137	0.093766094		0.029142509
25	REG_CITY_NOT_LIVE_CITY	0.002202687	-0.024248146	-0.002376997	-0.025142729	-0.052408341	-0.17719	0.091439709		0.065582575
26	REG_CITY_NOT_WORK_CITY	0.006326544	-0.014655908	0.007726444	-0.01577892	-0.039380545	-0.24883	0.25985092		0.101256619
27	LIVE_CITY_NOT_WORK_CITY	0.008100379	0.001695341	0.013826989	0.000850445	-0.012101186	-0.16895	0.223564895		0.075229268
28	OBS_30_CNT_SOCIAL_CIRCLE	-0.008694889	0.006898958	-0.005042434	0.008695138	-0.007317979	-0.01573	-0.00198408		0.014834914
29	DEF_30_CNT_SOCIAL_CIRCLE	-0.010868533	-0.018144009	-0.02009904	-0.018148606	0.005802836	-0.00173	-0.015377301		0.006877567
30	OBS_60_CNT_SOCIAL_CIRCLE	-0.008610784	0.006894975	-0.004676607	0.008634119	-0.006789724	-0.01521	-0.002156972		0.014662222
31	DEF_60_CNT_SOCIAL_CIRCLE	-0.010840789	-0.020849267	-0.021405585	-0.020825513	0.003555206	-0.00373	-0.013216572		0.006138788
32	FLAG_DOCUMENT_2	-0.000618717	0.013825415	0.006675743	0.016469147	-0.003794963	0.001309	0.004134894		-0.006911186

	A	B	C	D	E	F	G	H	I	J
1	SK_ID_CURR	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	NAME_TYPE_SUITE	NAME_INCOME_TYPE	NAME_EDUCATION_TYPE	NAME_FAMILY_STATUS	NAME_HOUSING_TYPE
2	271877	Cash loans	F	N	Y	Unaccompanied	State servant	Incomplete higher	Married	House / apartment
3	108129	Revolving loans	F	N	Y	Family	Commercial associate	Secondary / secondary special	Civil marriage	House / apartment
4	122040	Cash loans	F	N	Y	Spouse, partner	Working	Secondary / secondary special	Married	House / apartment
5	122040	Cash loans	F	N	Y	Spouse, partner	Working	Secondary / secondary special	Married	House / apartment
6	100006	Cash loans	F	N	Y	Unaccompanied	Working	Secondary / secondary special	Civil marriage	House / apartment
7	176158	Cash loans	M	Y	Y	Family	Pensioner	Secondary / secondary special	Married	House / apartment
8	176158	Cash loans	M	Y	Y	Family	Pensioner	Secondary / secondary special	Married	House / apartment
9	176158	Cash loans	M	Y	Y	Family	Pensioner	Secondary / secondary special	Married	House / apartment
10	100007	Cash loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
11	100007	Cash loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
12	202054	Cash loans	F	N	N	Unaccompanied	State servant	Secondary / secondary special	Married	House / apartment
13	202054	Cash loans	F	N	N	Unaccompanied	State servant	Secondary / secondary special	Married	House / apartment
14	202054	Cash loans	F	N	N	Unaccompanied	State servant	Secondary / secondary special	Married	House / apartment
15	199383	Cash loans	M	Y	Y	Family	Working	Secondary / secondary special	Civil marriage	House / apartment
16	199383	Cash loans	M	Y	Y	Family	Working	Secondary / secondary special	Civil marriage	House / apartment
17	100009	Cash loans	F	Y	Y	Unaccompanied	Commercial associate	Higher education	Married	House / apartment
18	175704	Cash loans	F	N	Y	Family	Pensioner	Secondary / secondary special	Married	House / apartment
19	296299	Cash loans	F	N	Y	Unaccompanied	Pensioner	Secondary / secondary special	Married	House / apartment
20	296299	Cash loans	F	N	Y	Unaccompanied	Pensioner	Secondary / secondary special	Married	House / apartment
21	296299	Cash loans	F	N	Y	Unaccompanied	Pensioner	Secondary / secondary special	Married	House / apartment
22	100011	Cash loans	F	N	Y	Children	Pensioner	Secondary / secondary special	Married	House / apartment
23	342292	Cash loans	F	N	N	Unaccompanied	Commercial associate	Secondary / secondary special	Married	House / apartment
24	342292	Cash loans	F	N	N	Unaccompanied	Commercial associate	Secondary / secondary special	Married	House / apartment
25	342292	Cash loans	F	N	N	Unaccompanied	Commercial associate	Secondary / secondary special	Married	House / apartment
26	342292	Cash loans	F	N	N	Unaccompanied	Commercial associate	Secondary / secondary special	Married	House / apartment
27	100012	Revolving loans	M	N	Y	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
28	334349	Cash loans	F	N	N	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
29	334349	Cash loans	F	N	N	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
30	334349	Cash loans	F	N	N	Unaccompanied	Working	Secondary / secondary special	Single / not married	House / apartment
31	447712	Cash loans	F	N	Y	Unaccompanied	Pensioner	Secondary / secondary special	Widow	House / apartment
32	161140	Revolving loans	F	N	N	Unaccompanied	Pensioner	Secondary / secondary special	Separated	House / apartment

	A	B	C
298			
299	AMT_INCOME_TOTAL	Count of	Count of TARGET_1
300	65650-75650	2695	30117
301	75650-85650	2543	25523
302	85650-95650	5426	58509
303	95650-105650	1929	18133
304	105650-115650	8191	83954
305	115650-125650	1821	16823
306	125650-135650	10337	109663
307	135650-145650	1289	11100
308	145650-155650	866	7237
309	155650-165650	8042	78671
310	165650-175650	1652	16269
311	175650-185650	6992	70300
312	185650-195650	690	8318
313	195650-205650	4801	49917
314	215650-225650	6031	63908
315	245650-255650	1683	18482
316	265650-275650	2896	32151
317	285650-295650	916	10968
318	305650-315650	1325	18554
319	355650-365650	851	11895
320			

	A	B	C
1			
2	Table_previous_application.NAME_CASH_LOAN_PURPOSE	TARGET	Count of TARGET
3	Building a house or an annex		216
4	Business development		28
5	Buying a garage		3
6	Buying a holiday home / land		35
7	Buying a home		49
8	Buying a new car		49
9	Buying a used car		201
10	Car repairs		83
11	Education		86
12	Everyday expenses		124
13	Furniture		51
14	Gasification / water supply		27
15	Hobby		7
16	Journey		70
17	Medicine		152
18	Money for a third person		2
19	Other		1043
20	Payments on other loans		165
21	Purchase of electronic equipment		64
22	Refusal to name the goal		1
23	Repairs		1598
24	Urgent needs		701
25	Wedding / gift / holiday		52
26	XAP		41290
27	XNA		30657
28			
29			

