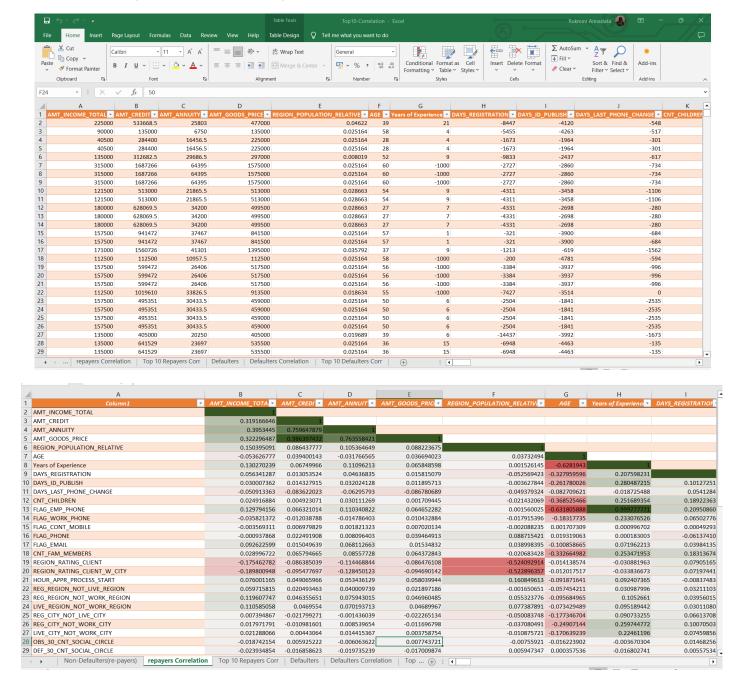*-Rukmini Annadata*

TOP 10 Correlation

To accommodate GitHub's file size limit of 25MB, I will be including screenshots of my Excel file for the bank loan case study project instead of uploading the actual 252MB file.

Identify Top Correlations for Different Scenarios: Understanding the correlation between variables and the target variable can provide insights into strong indicators of loan default.
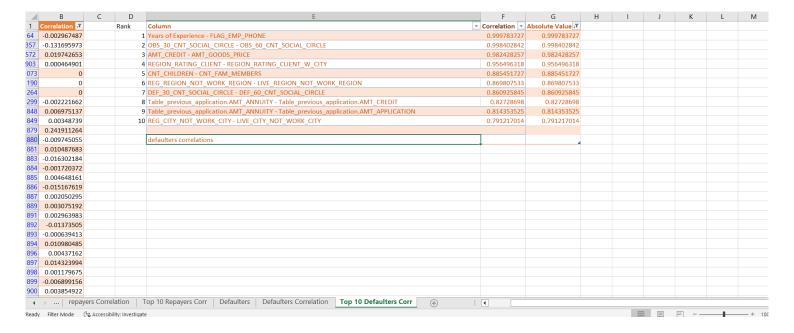
Task: Segment the dataset based on different scenarios (e.g., clients with payment difficulties and all other cases) and identify the top correlations for each segmented data using Excel functions.

AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	AGE	Years of Experience	DAYS_REGISTRATION	DAYS_ID_PUBLISH	DAYS_LAST_PHONE_CHANGE	CNT_CHILDREN
225000	533636.5	25803	477000	0.04622	39	21	-8447	-4120	-548	
90000	135000	6750	135000	0.025164	58	4	-5455	-4263	-517	
40500	284400	16456.5	225000	0.025164	28	4	-1673	-1964	-301	
40500	284400	16456.5	225000	0.025164	28	4	-1673	-1964	-301	
135000	312682.5	29686.5	297000	0.008019	52	9	-9833	-2437	-617	
315000	1687266	64395	1575000	0.025164	60	-1000	-2727	-2860	-734	
315000	1687266	64395	1575000	0.025164	60	-1000	-2727	-2860	-734	
315000	1687266	64395	1575000	0.025164	60	-1000	-2727	-2860	-734	
121500	513000	21865.5	513000	0.028663	54	9	-4311	-3458	-1106	
121500	513000	21865.5	513000	0.028663	54	9	-4311	-3458	-1106	
180000	628069.5	34200	499500	0.028663	27	7	-4331	-2698	-280	
180000	628069.5	34200	499500	0.028663	27	7	-4331	-2698	-280	
180000	628069.5	34200	499500	0.028663	27	7	-4331	-2698	-280	
157500	941472	37467	841500	0.025164	57	1	-321	-3900	-684	
157500	941472	37467	841500	0.025164	57	1	-321	-3900	-684	
171000	1560726	41301	1395000	0.035792	37	9	-1213	-619	-1562	
112500	112500	10957.5	112500	0.025164	58	-1000	-200	-4781	-594	
157500	599472	26406	517500	0.025164	56	-1000	-3384	-3937	-996	
157500	599472	26406	517500	0.025164	56	-1000	-3384	-3937	-996	
157500	599472	26406	517500	0.025164	56	-1000	-3384	-3937	-996	
112500	1019610	33826.5	913500	0.018634	55	-1000	-7427	-3514	0	
157500	495351	30433.5	459000	0.025164	50	6	-2504	-1841	-2535	
157500	495351	30433.5	459000	0.025164	50	6	-2504	-1841	-2535	
157500	495351	30433.5	459000	0.025164	50	6	-2504	-1841	-2535	
157500	495351	30433.5	459000	0.025164	50	6	-2504	-1841	-2535	
135000	405000	20250	405000	0.019689	39	6	-14437	-3992	-1673	
135000	641529	23697	535500	0.025164	36	15	-6948	-4463	-135	
135000	641529	23697	535500	0.025164	36	15	-6948	-4463	-135	

Sheet tabs: repayers Correlation | Top 10 Repayers Corr | Defaulters | Defaulters Correlation | Top 10 Defaulters Corr

Column1	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	AGE	Years of Experience	DAYS_REGISTRATION
AMT_INCOME_TOTAL								
AMT_CREDIT	0.319166646	1						
AMT_ANNUITY	0.3953445	0.759647879	1					
AMT_GOODS_PRICE	0.322296487	0.986397432	0.763558421	1				
REGION_POPULATION_RELATIVE	0.150395091	0.086437777	0.105364649	0.088223675	1			
AGE	-0.053626777	0.039400143	-0.031766565	0.036694023	0.03732494	1		
Years of Experience	0.130270239	0.06749966	0.11096213	0.065848598	0.001526145	-0.6281943	1	
DAYS_REGISTRATION	0.056341287	0.013053524	0.04636835	0.015815079	-0.052569423	-0.327959596	0.207598231	1
DAYS_ID_PUBLISH	0.030007362	0.014327915	0.032024128	0.011895713	-0.003627844	-0.261780026	0.280487215	0.10127251
DAYS_LAST_PHONE_CHANGE	-0.050913363	-0.083622023	-0.06295793	-0.086780689	-0.049379324	-0.082709621	-0.018725488	0.0541284
CNT_CHILDREN	0.024916884	0.004923071	0.030111269	0.001709445	-0.021432069	-0.368525466	0.251689354	0.18922363
FLAG_EMP_PHONE	0.129794156	0.066321014	0.110340822	0.064652282	0.001560025	-0.631805888	0.999777771	0.20950860
FLAG_WORK_PHONE	-0.035821372	-0.012038788	-0.014786403	0.010432884	-0.017915396	-0.18317735	0.233076526	0.06502776
FLAG_CONT_MOBILE	-0.003569311	0.006979829	0.001821323	0.007020134	-0.002088235	0.001707309	0.000996702	0.00049293
FLAG_PHONE	-0.000937868	0.022491908	0.008096403	0.039464913	0.088715421	0.019319063	0.000183003	-0.06137410
FLAG_EMAIL	0.092622599	0.015049639	0.068112663	0.01534832	0.038998395	-0.100858465	0.071962213	0.03984135
CNT_FAM_MEMBERS	0.028996722	0.065794665	0.08557728	0.064372843	-0.020683428	-0.332664982	0.253471953	0.18313674
REGION_RATING_CLIENT	-0.175462782	-0.086385039	-0.114468844	-0.086476108	-0.524092914	-0.014138574	-0.030881963	0.07905165
REGION_RATING_CLIENT_W_CITY	-0.189800948	-0.095477697	-0.128450123	-0.094690142	-0.522896357	-0.012017517	-0.033836673	0.07197441
HOUR_APPR_PROCESS_START	0.076001165	0.049065966	0.053436129	0.058039944	0.160849613	-0.091871641	0.092407365	-0.00837483
REG_REGION_NOT_LIVE_REGION	0.059715815	0.020493463	0.040009739	0.021897186	-0.001650651	-0.057454211	0.030987996	0.03211103
REG_REGION_NOT_WORK_REGION	0.119607747	0.046355651	0.075943015	0.046960485	0.055323776	-0.095684965	0.1052661	0.03956015
LIVE_REGION_NOT_WORK_REGION	0.110585058	0.0469554	0.070193713	0.04689967	0.077387891	-0.073429489	0.095189442	0.03011080
REG_CITY_NOT_LIVE_CITY	0.007394867	-0.021799271	-0.001436039	-0.022265134	-0.050083748	-0.177346704	0.090733255	0.06613708
REG_CITY_NOT_WORK_CITY	0.017971791	-0.010981601	0.008539654	-0.011696798	-0.037080491	-0.24907144	0.259744772	0.10070503
LIVE_CITY_NOT_WORK_CITY	0.021288066	0.00443064	0.014415367	0.003758754	-0.010875721	-0.170639239	0.22461196	0.07459856
OBS_30_CNT_SOCIAL_CIRCLE	-0.018742154	0.005925222	-0.006063622	0.007743721	-0.00755921	-0.016223902	-0.003670304	0.01468256
DEF_30_CNT_SOCIAL_CIRCLE	-0.023934854	-0.016858623	-0.019735239	-0.017009874	0.005947347	0.000357536	-0.016802741	0.00557534

Sheet tabs: Non-Defaulters(re-payers) | **repayers Correlation** | Top 10 Repayers Corr | Defaulters | Defaulters Correlation | Top ...

Top 10 Repayers Corr

	Column2			Column	Correlation	Absolute Value	Column
67	-0.014063653			Years of Experience - FLAG_EMP_PHONE	0.999777771	0.999777771	1
68	0			OBS_30_CNT_SOCIAL_CIRCLE - OBS_60_CNT_SOCIAL_CIRCLE	0.998553814	0.998553814	2
131	0.003409617			AMT_CREDIT - AMT_GOODS_PRICE	0.986397432	0.986397432	3
381	-0.023469268			REGION_RATING_CLIENT - REGION_RATING_CLIENT_W_CITY	0.94412424	0.94412424	4
612	-0.01287044			CNT_CHILDREN - CNT_FAM_MEMBERS	0.878421446	0.878421446	5
971	0.005953422			REG_REGION_NOT_WORK_REGION - LIVE_REGION_NOT_WORK_REGION	0.876379493	0.876379493	6
1157	-0.010650865			DEF_30_CNT_SOCIAL_CIRCLE - DEF_60_CNT_SOCIAL_CIRCLE	0.86231905	0.86231905	7
1286	-0.001762723			REG_CITY_NOT_WORK_CITY - LIVE_CITY_NOT_WORK_CITY	0.835634138	0.835634138	8
1368	0.009597189			AMT_ANNUITY - AMT_GOODS_PRICE	0.763558421	0.763558421	9
1407	0			AMT_CREDIT - AMT_ANNUITY	0.759647879	0.759647879	10
2147	-0.001032519						
2148	-0.002790565						
2149	-0.000126684						
2150	-0.00249655						
2151	-0.000497227						
2152	-6.86987E-05						
2153	-0.000366158						
2154	-9.71517E-06						
2155	-0.00037762						
2156	-0.000338755						
2157	-0.000233449						
2158	-0.000694758						
2159	-0.00011205						
2160	-0.000572853						
2161	-0.000181533						
2162	-0.000165473						
2163	-0.000132156						
2164	-0.000620322						

Sheet tabs: Non-Defaulters(re-payers) | repayers Correlation | **Top 10 Repayers Corr** | Defaulters | Defaulters Correlation | Top ...

Defaulters

AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	AGE	Years of Experience	DAYS_REGISTRATION	DAYS_ID_PUBLISH	DAYS_LAST_PHONE_CHANGE	CNT_CHILDF
135000	675000	53460	675000	0.035792	39	2	-482	-1721	-1800	
135000	675000	53460	675000	0.035792	39	2	-482	-1721	-1800	
202500	1193580	35028	855000	0.025164	47	4	-1182	-1029	-1075	
202500	1193580	35028	855000	0.025164	47	4	-1182	-1029	-1075	
135000	288873	16258.5	238500	0.007305	36	10	-45	-4409	-1480	
81000	112068	7488	99000	0.025164	26	2	-4717	-1721	-773	
81000	112068	7488	99000	0.025164	26	2	-4717	-1721	-773	
360000	539100	29245.5	450000	0.072508	35	10	-7009	-339	-229	
135000	601677	29394	423000	0.035792	47	5	-9969	-1031	-1568	
135000	601677	29394	423000	0.035792	47	5	-9969	-1031	-1568	
135000	601677	29394	423000	0.035792	47	5	-9969	-1031	-1568	
135000	601677	29394	423000	0.035792	47	5	-9969	-1031	-1568	
450000	1049094	30802.5	751500	0.072508	37	10	-7806	-4956	-1883	
157500	723996	30802.5	585000	0.007274	28	1	-387	-3190	-56	
292500	675000	36747	675000	0.035792	45	1	-5239	-114	-106	
157500	245619	12667.5	166500	0.022625	48	22	-774	-1082	-216	
90000	458460	36351	405000	0.018634	31	5	-9593	-3967	-630	
1575000	553806	28273.5	495000	0.032561	51	3	-5412	-2291	-1728	
1575000	553806	28273.5	495000	0.032561	51	3	-5412	-2291	-1728	
1575000	553806	28273.5	495000	0.032561	51	3	-5412	-2291	-1728	
135000	495216	26995.5	427500	0.008019	35	1	-6764	-126	-309	
157500	497520	33376.5	450000	0.028663	32	5	-5561	-4452	-252	
157500	497520	33376.5	450000	0.028663	32	5	-5561	-4452	-252	
112500	314055	19107	238500	0.016612	43	3	-7079	-4104	-336	
180000	170640	13612.5	135000	0.016612	41	3	0	-453	0	
180000	170640	13612.5	135000	0.016612	41	3	0	-453	0	
86400	1078200	31653	900000	0.00712	36	18	-4506	-5147	-365	

Sheet tabs: Non-Defaulters(re-payers) | repayers Correlation | Top 10 Repayers Corr | **Defaulters** | Defaulters Correlation | Top ...

Defaulters Correlation

Column1	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	AGE	Years of Experience	DAYS_REGISTRATION
AMT_INCOME_TOTAL	1							
AMT_CREDIT	0.034732656	1						
AMT_ANNUITY	0.044584554	0.745685336	1					
AMT_GOODS_PRICE	0.033961552	0.982428257	0.744917968	1				
REGION_POPULATION_RELATIVE	0.009349675	0.065861374	0.069290317	0.072988464	1			
AGE	-0.001149399	0.125337734	-0.003340074	0.127965017	0.051700791	1		
Years of Experience	0.013138335	-0.012214684	0.079597617	-0.019513733	-0.02442861	-0.583552214	1	
DAYS_REGISTRATION	-0.001298611	-0.020154859	0.046917968	-0.021160917	-0.0453924	-0.28063518	0.180775817	1
DAYS_ID_PUBLISH	-0.003162477	0.00824408	0.00824408	-0.034193643	-0.015145174	-0.236034607	0.2302206	0.09941500
DAYS_LAST_PHONE_CHANGE	-0.001398229	-0.10713107	-0.057222072	-0.113337724	-0.04332884	-0.136030307	0.007486295	0.06718219
CNT_CHILDREN	0.002880004	-0.011097018	0.029126834	-0.018316704	-0.02187898	-0.294013003	0.201110469	0.1498364
FLAG_EMP_PHONE	0.01314263	-0.013792129	0.078961068	-0.021170363	-0.024751122	-0.586799503	0.999783727	0.18253470
FLAG_WORK_PHONE	-0.011745266	-0.04765372	-0.029708138	-0.023837863	-0.022723555	-0.16913937	0.20664583	0.05713984
FLAG_CONT_MOBILE	0.000758684	0.01063325	0.003400453	0.009991556	-0.009263497	-0.005661463	0.005881698	0.008547
FLAG_PHONE	-0.00415243	0.014311336	-0.010392566	0.035063231	0.062153584	0.008570127	-0.019581673	-0.068918
FLAG_EMAIL	0.010283004	0.011135349	0.078242938	0.015060793	0.025462339	-0.067696612	0.045431684	0.02879694
CNT_FAM_MEMBERS	0.004160602	0.04288948	0.072256192	0.038113294	-0.023073341	-0.24757942	0.202130645	0.14932910
REGION_RATING_CLIENT	-0.021877767	-0.056500942	-0.069594439	-0.063384631	-0.430546204	-0.045964576	0.004792481	0.0909522
REGION_RATING_CLIENT_W_CITY	-0.023583865	-0.065936983	-0.083746657	-0.071066933	-0.433938097	-0.045465757	0.005088822	0.08497086
HOUR_APPR_PROCESS_START	0.012993933	0.02412127	0.028260231	0.036670721	0.140426971	-0.063327469	0.056783894	-0.03162676
REG_REGION_NOT_LIVE_REGION	0.006926149	0.014117129	0.037142185	0.015563183	-0.028814269	-0.042923143	0.024300901	0.01764255
REG_REGION_NOT_WORK_REGION	0.0136988	0.026698087	0.06617805	0.029217216	0.031651243	-0.065635292	0.087228178	0.02087155
LIVE_REGION_NOT_WORK_REGION	0.012461436	0.027782931	0.059250931	0.029645082	0.053489093	-0.048510423	0.077269731	0.01703788
REG_CITY_NOT_LIVE_CITY	-0.002335938	-0.039179118	-0.008647492	-0.040572755	-0.063364029	-0.155181952	0.084004822	0.04525080
REG_CITY_NOT_WORK_CITY	-0.003083915	-0.041387943	0.003249217	-0.042063095	-0.045812316	-0.218264942	0.247918235	0.08642893
LIVE_CITY_NOT_WORK_CITY	-0.001593168	-0.018989404	0.010496635	-0.018426005	-0.012159908	-0.131695973	0.203506899	0.06761170
OBS_30_CNT_SOCIAL_CIRCLE	-0.004138615	0.022480873	0.008233559	0.025899933	0.001721724	0.001310276	0.010101544	0.00961655
DEF_30_CNT_SOCIAL_CIRCLE	-0.00436615	-0.024142128	-0.022307207	-0.019616738	0.017064357	0.00280272	-0.018151509	0.00556253

	B (Correlation)	D (Rank)	E (Column)	F (Correlation)	G (Absolute Value)
64	-0.002967487	1	Years of Experience - FLAG_EMP_PHONE	0.999783727	0.999783727
357	-0.131695973	2	OBS_30_CNT_SOCIAL_CIRCLE - OBS_60_CNT_SOCIAL_CIRCLE	0.998402842	0.998402842
572	0.019742653	3	AMT_CREDIT - AMT_GOODS_PRICE	0.982428257	0.982428257
903	0.000464901	4	REGION_RATING_CLIENT - REGION_RATING_CLIENT_W_CITY	0.956496318	0.956496318
073	0	5	CNT_CHILDREN - CNT_FAM_MEMBERS	0.885451727	0.885451727
190	0	6	REG_REGION_NOT_WORK_REGION - LIVE_REGION_NOT_WORK_REGION	0.869807533	0.869807533
264	0	7	DEF_30_CNT_SOCIAL_CIRCLE - DEF_60_CNT_SOCIAL_CIRCLE	0.860925845	0.860925845
299	-0.002221662	8	Table_previous_application.AMT_ANNUITY - Table_previous_application.AMT_CREDIT	0.82728698	0.82728698
848	0.006975137	9	Table_previous_application.AMT_ANNUITY - Table_previous_application.AMT_APPLICATION	0.814353525	0.814353525
849	0.00348739	10	REG_CITY_NOT_WORK_CITY - LIVE_CITY_NOT_WORK_CITY	0.791217014	0.791217014
879	0.241911264				
880	-0.009745055		defaulters correlations		
881	0.010487683				
883	-0.016302184				
884	-0.001720372				
885	0.004648161				
886	-0.015167619				
887	0.002050295				
889	0.003075192				
891	0.002963983				
892	-0.01373505				
893	-0.000639413				
894	0.010980485				
896	0.00437162				
897	0.014323994				
898	0.001179675				
899	-0.006899156				
900	0.003854922				

Sheet tabs: ... | repayers Correlation | Top 10 Repayers Corr | Defaulters | Defaulters Correlation | **Top 10 Defaulters Corr** | ⊕

Ready   Filter Mode   Accessibility: Investigate

Key Factors / Conclusion:

• As the Age and Years of Experience Increase, the chances of defaulting decrease. So the bank should prioritize Older and Experienced Clients.

• Educated Clients tend to default lesser

• Male clients tend to default more than female clients(i.e., Female clients take more)

• As the Age increases, the amount taken by the clients is considerably higher and since the default percentage with higher age is lower these should be least risky and highly profitable clients for the bank