Step	Code start	Code end	Start df name	End df name	Before shape	After shape	Study part
Data load	62	83	NA	data	0-0	1259-27	#2.1.3
>20% MV	84	88	data	data	1259-27	1259-24	#2.2.2
NMAR	89	92	data	data	1259-24	1241-24	#2.2.2
Dom know	93	96	data	data	1241-24	1241-21	#2.2.3
ValCap	97	115	data	data	1241-21	1233-21	#2.2.4
HighCard	116	188	data	data	1233-21	1233-21	#2.2.5
Cramér's	189	285	data	data	1233-21	1233-16	#2.3
1HE	286	320	data	data_encoded_float	1233-16	1233-41	#2.4.1
RobSca_1	321	333	data_encoded_float	scaled_data	1233-41	1233-41	#2.4.2
VT()	334	360	scaled_data	data_VT	1233-41	1233-38	#2.5.1
Spearman	361	426	data_VT	data_corr_final	1233-38	1233-27	#2.5.2
MDS	427	500	data_corr_final	MDS_clean_data	1233-27	1228-23	#2.6.1
LLE	501	584	MDS_clean_data	LLEed_data	1228-23	~1160-23	#2.6.2
DFS	585	694	LLEed_data	data_dfs_scaled	~1160-23	~1160-157	#2.6.3
Feat Imp	695	797	data_dfs_scaled	filtered_data	~1160-157	~1160-7	#2.6.4
Clustering	798	END	filtered_data	data_with_HDBSCAN_2	~1160-7	END	#2.7