Data Quality Plan

Strategies for inconsistencies

Feature ⬍	Data Quality Issue ⬍	Handling Strategy ⬍
ExternalRiskEstimate	negative value	imputation
PercentTradesWBalance	negative value	imputation
MSinceOldestTradeOpen	negative values	imputation
MSinceMostRecentInqexcl7days	negative values (>24%)	imputation
NetFractionRevolvingBurden	negative values	imputation
NetFractionInstallBurden	negative values	imputation
NumRevolvingTradesWBalance	negative values	imputation
NumInstallTradesWBalance	negative values	imputation
NumBank2NatlTradesWHighUtilization	negative values	imputation
MSinceMostRecentDelq	negative values (>50%)	imputation
NumTrades60Ever2DerogPubRec	many 0s	do nothing
NumTrades90Ever2DerogPubRec	many 0s	do nothing
NumTradesOpeninLast12M	many 0s	do nothing
NumInqLast6M	many 0s	do nothing
NumInqLast6Mexcl7days	many 0s	do nothing
MSinceMostRecentTradeOpen	outliers	do nothing
AverageMInFile	outliers	do nothing
NumSatisfactoryTrades	outliers	do nothing
NumTotalTrades	outliers	do nothing

Clean data results

	count	unique	top	freq
RiskPerformance	939	2	Bad	475
MaxDelq2PublicRecLast12M	939	9	7	404
MaxDelqEver	939	7	8	436

	count	mean	std	min	25%	50%	75%	max
AverageMInFile	939.0	78.570820	34.428568	5.0	56.0	75.0	96.0	245.0
PercentTradesNeverDelq	939.0	92.716720	11.654435	27.0	90.5	98.0	100.0	100.0
PercentInstallTrades	939.0	33.903088	17.344379	0.0	21.0	33.0	45.0	100.0
NetFractionRevolvingBurden	939.0	34.702875	28.941675	0.0	9.0	30.0	56.0	135.0
NetFractionInstallBurden	939.0	45.419595	37.734041	0.0	0.0	51.0	80.0	190.0
PercentTradesWBalance	939.0	66.379127	21.411369	0.0	50.0	67.0	82.5	100.0
ExternalRiskEstimate	939.0	72.108626	9.856157	40.0	65.0	71.0	80.0	93.0
MSinceOldestTradeOpen	939.0	196.111821	105.732309	0.0	128.5	183.0	259.0	565.0
MSinceMostRecentTradeOpen	939.0	9.367412	11.716637	0.0	3.0	6.0	12.0	152.0
NumSatisfactoryTrades	939.0	21.684771	11.766323	1.0	13.0	20.0	29.0	78.0
NumTrades60Ever2DerogPubRec	939.0	0.588924	1.224886	0.0	0.0	0.0	1.0	11.0
NumTrades90Ever2DerogPubRec	939.0	0.391906	0.954050	0.0	0.0	0.0	0.0	7.0
MSinceMostRecentDelq	939.0	11.566560	19.465365	0.0	0.0	0.0	15.0	83.0
NumTotalTrades	939.0	22.992545	13.272563	0.0	14.0	21.0	31.0	100.0
NumTradesOpeninLast12M	939.0	1.903088	1.938330	0.0	0.0	1.0	3.0	19.0
MSinceMostRecentInqexcl7days	939.0	1.894569	4.303834	0.0	0.0	0.0	1.0	23.0
NumInqLast6M	939.0	1.566560	2.072407	0.0	0.0	1.0	2.0	20.0
NumInqLast6Mexcl7days	939.0	1.510117	2.046541	0.0	0.0	1.0	2.0	20.0
NumRevolvingTradesWBalance	939.0	4.166134	3.049010	0.0	2.0	3.0	5.5	21.0
NumInstallTradesWBalance	939.0	2.342918	1.731922	0.0	1.0	2.0	3.0	13.0
NumBank2NatlTradesWHighUtilization	939.0	1.093717	1.612105	0.0	0.0	1.0	2.0	12.0