

DATA QUALITY LOG						
Table Name	Field / Column	Issue Found	Root Cause (If Known)	Fix Applied / Planned	Fix Location (Raw/View/Final)	Status (Open / Closed)
providers	All fields	None	N/A	N/A	N/A	Closed
members	chronic_condition	Categorical inconsistency + Missing clinical attributes .	Chronic condition not recorded for some members in the source dataset.	Do NOT replace NULL values. Maintain 'None' & 'NULL' as separate categories. NULL will be preserved to indicate missing clinical information, while 'None' indicates members with no chronic condition.	View	Closed
diagnosis_codes	All fields	None	N/A	N/A	N/A	Closed
claims	paid_amount	Missing paid_amount values (40 records).	Payment outcome not captured in source, paid_amount missing for 40 claims ( likely due to incomplete adjudication or extraction error).	Do NOT replace NULL with 0. NULL represents missing financial data, whereas 0 represents denied/no payment. In the cleaned_claims view create a derived field to flag these rows separately. (payment_status = 'Missing' when paid_amount IS NULL).	View	Closed
claims	paid_amount vs allowed_amount	40 rows where paid_amount > allowed_amount.	Synthetic dataset did not enforce adjudication rules creating invalid payment amounts	Add a flag column (paid_gt_allowed) BOOLEAN in cleaned_claims view to label rows where paid_amount > allowed_amount. Exclude flagged rows from financial KPIs while retaining them for data-quality monitoring.	View	Closed