

■ Data Dictionary – Healthcare RCM Project

This document defines the core datasets and fields used in the Healthcare Revenue Cycle & Utilization Analytics project. The goal is to ensure clarity, consistency, and business alignment.

Eligibility (stg_eligibility / dim_member_scd2)

Column	Description
member_id	Unique identifier for member/patient
first_name, last_name	Member name
dob	Date of birth
gender	Gender (M/F/O)
payer_id	Payer identifier
plan_id	Insurance plan identifier
plan_name	Plan descriptive name
product_type	Type of product (HMO, PPO, EPO)
coverage_start	Coverage effective start date
coverage_end	Coverage end date (null = active)
metal_tier	ACA tier (Bronze, Silver, Gold, Platinum)
member_group_id	Employer or group identifier
subscriber_id	Subscriber identifier (policy holder)
address_zip	Member ZIP code
state	Member state
_ingested_at	Timestamp when ingested

Claims Header (stg_claims_header)

Column	Description
claim_id	Unique identifier for the claim
member_id	Links to eligibility
payer_id	Insurance payer
provider_id	Rendering/billing provider ID
billing_npi	NPI of billing provider
rendering_npi	NPI of rendering provider
taxonomy_code	Provider taxonomy

tin	Provider tax ID
claim_type	Professional, Institutional
service_from, service_to	Claim service period
place_of_service	Place of service code
facility_type_code	Facility type
claim_frequency_code	Frequency (original/corrected)
claim_submit_date	Date submitted
claim_status	Paid, Denied, Submitted
total_charge	Total billed charge
drg_code	DRG (if institutional)
icn	Internal control number
tcn	Transaction control number
_ingested_at	Timestamp when ingested

Claims Line (stg_claims_line → fact_claims_line)

Column	Description
claim_id	Claim identifier
line_num	Line sequence number
cpt_code	CPT/HCPCS procedure code
icd_code	ICD diagnosis code
modifier1..modifier4	CPT modifiers
revenue_code	Hospital revenue code
pos_code	Place of service code
rendering_npi	Rendering provider NPI
claim_type	Professional, Institutional
claim_status	Line adjudication status
service_from, service_to	Line service dates
units	Units billed
line_charge_amount	Provider billed charge
allowed_amount	Allowed by payer (claim file)
paid_amount	Paid by payer (claim file)
adjustment_amount	Adjustment amount
latest_denial_code	Denial code if denied
last_remit_dt	Last remit applied

_ingested_at	Ingestion timestamp
--------------	---------------------

Remits (stg_remits → int_remits_latest)

Column	Description
remit_id	Unique remit identifier
claim_id	Claim reference
line_num	Line reference
paid_amount	Amount actually paid
allowed_amount	Allowed amount
adj_amount	Adjustment amount
denial_code	Denial code (if any)
remit_dt	Date of remit
group_code	Adjustment group code
remark_code	Remark code
reason_description	Explanation of denial/adjustment
payment_id	Payment identifier
payment_method	ACH, Check
payment_dt	Payment date
payer_claim_control_number	Payer reference
trace_number	Trace number
_ingested_at	Ingestion timestamp

Providers (Seed: providers.csv)

Column	Description
provider_id	Provider identifier
npi	National Provider Identifier
tin	Taxpayer Identification Number
taxonomy	Provider taxonomy code
specialty	Specialty description
location_id	Practice location reference
active_flag	Is provider active?