

Provider Network Optimization Dataset

PNOD v1.0 — Read Me First

Welcome to PNOD v1.0

Thank you for acquiring the Provider Network Optimization Dataset from IntelligentPlatforms.net. This document provides essential guidance for working with this dataset.

Key Field Groups

Provider_ID	Unique provider identifier (PNOD-XXXXXXXX format)
NPI	10-digit National Provider Identifier (CMS NPPES)
Taxonomy_Code	NUCC provider taxonomy code (10-character alphanumeric)
Network_Tier	Tier 1 Preferred / Tier 2 In-Network / Tier 3 Out-of-Network
Network_Adequacy_Score	Composite adequacy score (0–100) based on distance, ratio, and access standards
Quality_Score	Composite quality score (0–100) based on HEDIS, Star, CAHPS measures
Referral_Leakage_Rate	Percentage of referrals going out-of-network (lower is better)
Retention_Probability	Predicted probability (0.00–1.00) provider remains in network
Source_System	Primary source: NPPES, PECOS, CMS_UTILIZATION, NCQA_HEDIS, IP_MODEL

Technical Requirements

- Python 3.8+ with pandas, pyarrow for Parquet files
- Microsoft Excel 2016+ or compatible for .xlsx registry files
- Minimum 16 GB RAM recommended for full 1M-record dataset
- UTF-8 encoding throughout; dates in YYYY-MM-DD (ISO 8601) format
- Boolean fields: True/False string values; numeric decimals use period (.)