

Workers Compensation Claims Intelligence Dataset

WCCID v1.0 — Read Me First

Welcome to WCCID v1.0

Thank you for acquiring the Workers Compensation Claims Intelligence Dataset from IntelligentPlatforms.net. This document provides essential guidance for working with this dataset. Please read it in full before loading or processing the data.

Package Contents

Your package tier determines which files are included:

Dataset.csv — Primary dataset file in comma-separated values format

Dataset.parquet — Dataset in Apache Parquet format (Professional/Enterprise tiers)

Sample_Dataset.csv — 500-record evaluation sample (Sample tier only)

Data_Dictionary.xlsx — Complete field definitions, data types, and lineage tags

Schema.json — Machine-readable schema definition

Metadata.json — Dataset metadata and provenance information

Source_Lineage.xlsx — Full field-level source lineage documentation

Coverage_Report.pdf — Geographic and temporal coverage analysis

Quality_Assessment.pdf — Data quality scoring and validation results

Certification_Report.pdf — Platinum certification documentation

***_Registry.xlsx** — Domain-specific reference registries (Enterprise tier)

Field Lineage Tags

Every field in this dataset carries four mandatory lineage tags that document its provenance:

Source_System	The authoritative source system (e.g., NCCI, OSHA_300, DOL_BLS, STATE_WC)
Source_Type	Classification: PUBLIC_SOURCE, DERIVED_INTELLIGENCE, or SYNTHETIC_SIMULATION
Confidence_Score	Field-level confidence score (0.00–1.00) based on source authority
Transformation_Method	How the field was derived: SOURCE_ENRICHMENT, CALCULATION, MODELING, etc.

Technical Requirements

- Python 3.8+ with pandas, pyarrow for Parquet files
- Microsoft Excel 2016+ or compatible for .xlsx files
- Minimum 16 GB RAM recommended for full 1M-record dataset

- UTF-8 encoding — all text fields use standard UTF-8
- Date format: YYYY-MM-DD (ISO 8601) throughout
- Numeric fields use period (.) as decimal separator
- Boolean fields: True/False string values

Licensing

Use of this dataset is governed by the License Agreement included in your package. Redistribution, resale, or sublicensing is prohibited except under OEM licensing. For licensing inquiries, contact IntelligentPlatforms.net.