Data Dictionary

heat_resilience_scores.csv

National Transit Station Atlas — Tree Canopy & Heat Resilience Module

TrAMS Camp & Consulting | Schneider Analytics LLC

Overview

This file contains tree canopy cover, impervious surface coverage, and heat exposure classifications for 5,115 transit stations across the United States. Each row represents one station. The data is organized into 12 columns covering station identifiers, geographic attributes, two NLCD-derived environmental metrics, a composite heat score, and two heat exposure classification fields.

All environmental metrics are measured within a ½-mile circular walkshed centred on the station coordinate. Values represent the mean pixel value of the relevant NLCD 2021 raster layer within that circle, computed from a 64×64 pixel GeoTIFF exported from the MRLC GeoServer. This walkshed-mean approach reflects the land surface environment that riders experience traveling to and from the station, not just the station platform itself.

The primary key station_id can be used to join this file with other National Transit Station Atlas datasets, including stormwater_risk_scores.csv and stations_flood_zones.csv.

National Distribution Summary

The table below summarises key statistics from the 5,115-station national dataset.

Metric	Statistic	Interpretation
tcc_pct (Tree Canopy)	Min 0% | Max 73% | Mean 11% | Median 6%	The typical transit station walkshed has very low tree canopy. Only 20 of 5,115 stations reach the Dense Canopy threshold (≥ 60%).
imp_pct (Impervious Surface)	Min 0% | Max 95% | Mean 62% | Median 66%	Impervious surface dominates the typical station environment. Most stations fall in the Moderate or High Imperviousness tiers.
Composite Score	Min −71 | Max 95 | Mean 51 | Median 57	The national median is strongly positive, confirming impervious surface substantially outweighs tree canopy at most transit stations.
heat_quadrant distribution	Urban Heat Island: 83% | Bare/Open Land: 8% | Cool Green Corridor: 6% | Mixed/Partial: 1% | No data: 1%	More than 4 in 5 U.S. transit stations qualify as Urban Heat Islands under the current thresholds.

Column Reference

The tables below define all 12 columns in the file, organized by thematic section.

Section 1 — Station Identifiers

Column Name	Data Type	Unit / Values	Definition
station_id	Integer	Unique ID	Sequential integer uniquely identifying each transit station record. Primary key for joining with other Atlas datasets (stormwater_risk_scores.csv, stations_flood_zones.csv).

Column Name	Data Type	Unit / Values	Definition
station_name	Text	—	Full name of the transit station as reported to the National Transit Database (NTD). Names are standardized to all-caps and may include line or direction suffixes for multi-platform stations.
agency	Text	—	Name of the transit agency that operates the station, as reported to the NTD.

Section 2 — Geographic Attributes

Column Name	Data Type	Unit / Values	Definition
latitude	Float	Decimal degrees	Geographic latitude of the station centroid in WGS 84 decimal degrees.
longitude	Float	Decimal degrees	Geographic longitude of the station centroid in WGS 84 decimal degrees. Negative values indicate locations west of the Prime Meridian (i.e., the continental United States).

Section 3 — Tree Canopy Cover (NLCD 2021)

Column Name	Data Type	Unit / Values	Definition
tcc_pct	Integer	0–100 (%)	Walkshed-mean tree canopy cover percentage within a ½-mile circular area centred on the station. Derived from the USDA Forest Service NLCD 2021 Tree Canopy Cover raster (nlcd_tcc_conus_2021_v2021-4) at 30-metre resolution. Computed by requesting a 64×64 GeoTIFF from the MRLC GeoServer, applying a circular mask, and averaging valid pixels (0–100). Blank if outside NLCD CONUS coverage or if the GeoServer returned no valid pixels.
tcc_category	Text	Tier label	Human-readable classification of tcc_pct into five tiers. See Tree Canopy Category Reference table for threshold definitions.

Section 4 — Impervious Surface Coverage (NLCD 2021)

Column Name	Data Type	Unit / Values	Definition
imp_pct	Integer	0–100 (%)	Walkshed-mean impervious surface percentage within a ½-mile circular area centred on the station. Derived from the NLCD 2021 Impervious Surface raster (NLCD_2021_Impervious_L48) at 30-metre resolution using the same pipeline as tcc_pct. Each pixel carries a value of 0–100 representing the fraction of that 30m × 30m cell covered by roads, rooftops, parking lots, and other impermeable surfaces. Blank if outside NLCD CONUS coverage.
imp_category	Text	Tier label	Human-readable classification of imp_pct into five tiers. See Impervious Surface Category Reference table for threshold definitions.

Section 5 — Composite Heat Score

Column Name	Data Type	Unit / Values	Definition
Composite Score	Integer	−100 to +100	Net heat exposure score computed as: Composite Score = imp_pct − tcc_pct. Positive values indicate

Column Name	Data Type	Unit / Values	Definition
			impervious surface dominates (higher heat risk); negative values indicate tree canopy dominates (lower heat risk). Observed range across 5,115 stations: −71 to +95 (mean = 51, median = 57). Blank if either tcc_pct or imp_pct is blank.

Section 6 — Heat Exposure Matrix Classification

Column Name	Data Type	Unit / Values	Definition
heat_quadrant	Text	Key string	Machine-readable key identifying which of four heat exposure quadrants the station falls into, based on binary thresholds applied to tcc_pct (≥ 35% = high canopy) and imp_pct (≥ 40% = high impervious). Values: cool_green | mixed | bare_open | heat_island | unknown. See Heat Exposure Matrix Reference for full definitions.
heat_label	Text	Display label	Human-readable label corresponding to heat_quadrant, suitable for display in charts and reports. Values: Cool Green Corridor | Mixed / Partial Cover | Bare / Open Land | Urban Heat Island | Insufficient data.

Tree Canopy Category Reference

The tcc_category field is derived from tcc_pct using the thresholds below. Thresholds were calibrated to the NLCD distribution for developed urban land in CONUS; a score of 35% or above is meaningfully above average for a U.S. transit station walkshed.

tcc_category Value	tcc_pct Range	Heat Implication
Dense Canopy	≥ 60%	Low heat exposure. Excellent cooling from shade and evapotranspiration. Rare at U.S. transit stations (20 of 5,115 stations).
Moderate Canopy	35–59%	Moderate cooling. Above-average canopy for urban transit environments. Trees provide meaningful but not complete heat relief.
Sparse Canopy	15–34%	Elevated heat exposure. Some tree cover but insufficient to substantially cool the walkshed. Riders are largely unshaded during peak heat.
Minimal Canopy	1–14%	High heat exposure. Very limited tree cover. The most common tier for U.S. transit stations (3,407 of 5,115).
No Canopy	0%	Very high heat exposure. No measurable tree canopy within the ½-mile walkshed.
No data	—	Station is outside NLCD CONUS coverage (Alaska, Hawaii, PR/USVI) or the MRLC GeoServer returned no valid pixels.

Impervious Surface Category Reference

The imp_category field is derived from imp_pct using the thresholds below. The 40% threshold used in the heat matrix corresponds to approximately the national median for developed urban land.

imp_category Value	imp_pct Range	Heat Implication
High Imperviousness	≥ 70%	Dominant paved/built surface. Strong urban heat island effect; surfaces absorb and re-radiate solar heat throughout the day and into the evening.

imp_category Value	imp_pct Range	Heat Implication
Moderate Imperviousness	40–69%	Significant impervious coverage. Heat island effect is measurable; limited permeable surfaces for evaporation.
Low Imperviousness	15–39%	Mixed surfaces. Some permeable land moderates heat; typical of low-density suburban environments.
Minimal Imperviousness	1–14%	Predominantly natural or undeveloped land. Minimal contribution to urban heat island effect.
No Impervious Surface	0%	No measurable impervious surface within the walkshed. Rare for transit station environments.
No data	—	Outside NLCD CONUS coverage or MRLC GeoServer returned no valid pixels.

Composite Score Reference

The Composite Score is computed as imp_pct − tcc_pct. Positive values indicate impervious surface dominates the walkshed; negative values indicate tree canopy dominates. The national observed range is −71 to +95.

Score Range	Interpretation
≥ 50	Heat island conditions dominate. High impervious, low canopy. Riders face elevated heat stress. Most U.S. transit stations fall here (median = 57).
10 – 49	Moderate heat exposure. Impervious surface outweighs canopy but conditions are not extreme.
−10 – 9	Roughly balanced. Neither canopy nor impervious surface dominates the walkshed character.
< −10	Canopy-dominated walkshed. Tree cover substantially exceeds impervious surface; cooling conditions prevail.
Minimum −71	A high negative score indicates a densely wooded station area with minimal paving — the national minimum observed value.

Heat Exposure Matrix Reference

The heat_quadrant field places each station in one of four quadrants based on binary thresholds applied simultaneously to tcc_pct and imp_pct. Both thresholds must be met to qualify for a given quadrant.

heat_quadrant	heat_label	Qualifying Conditions	Interpretation
cool_green	Cool Green Corridor	TCC ≥ 35% AND Impervious < 40%	Best conditions. Shade and permeable surfaces moderate heat. Riders experience relatively cool walking environments.
mixed	Mixed / Partial Cover	TCC ≥ 35% AND Impervious ≥ 40%	Partial cooling. Trees provide relief but significant impervious surfaces still absorb and re-radiate heat.
bare_open	Bare / Open Land	TCC < 35% AND Impervious < 40%	Moderate risk. Limited canopy but also limited paving. Typical of parks, fields, or low-density fringe areas.
heat_island	Urban Heat Island	TCC < 35% AND Impervious ≥ 40%	Highest risk. Minimal shade combined with extensive hard surfaces. Greatest rider heat stress, especially during heat waves. 83% of U.S. transit stations fall here.
unknown	Insufficient data	tcc_pct or imp_pct is	One or both NLCD layers returned no valid

heat_quadrant	heat_label	Qualifying Conditions	Interpretation
		null	pixels. Classification cannot be assigned.

Notes on Data Sources & Methodology

NLCD Tree Canopy Cover (TCC)

Tree canopy cover values are derived from the USDA Forest Service NLCD 2021 Tree Canopy Cover product (layer identifier: nlcd_tcc_conus_2021_v2021-4), produced in partnership with the Multi-Resolution Land Characteristics (MRLC) Consortium using Landsat and Sentinel-2 satellite imagery at 30-metre resolution. Pixel values represent percent canopy cover (0–100) estimated by a regression tree model trained on field-verified reference data. Accuracy is approximately ±10 percentage points at the pixel level, with better performance in dense urban areas.

NLCD Impervious Surface

Impervious surface values are derived from the NLCD 2021 Impervious Surface layer (layer identifier: NLCD_2021_Impervious_L48), produced by the same MRLC Consortium pipeline. Each 30-metre pixel carries a value of 0–100 representing the fraction of that cell covered by roads, rooftops, parking lots, and other hard surfaces that shed water. This sub-pixel precision makes the layer more accurate than a binary classification.

Walkshed Mean Computation

For each station, a 64×64 pixel GeoTIFF tile was requested from the MRLC GeoServer (https://www.mrlc.gov/geoserver) covering the ½-mile bounding box centred on the station coordinate. A circular boolean mask (radius = 32 pixels) was applied. All pixels with values ≤ 100 were retained; pixels with values > 100 (NLCD no-data sentinels) were excluded. The mean of valid pixels was rounded to the nearest integer. At 64 pixels across a ½-mile bounding box, each pixel covers approximately 40 metres, close to NLCD's native 30-metre resolution. The circle contains approximately 3,200 valid pixels for a typical CONUS station.

Relationship to Stormwater Risk Scores

The NLCD-derived impervious surface metric (imp_pct) in this file differs from the impervious surface component of stormwater_risk_scores.csv. The stormwater score uses an OpenStreetMap polygon-based metric counting explicitly mapped parking lots, road polygons, and building footprints — a land-use-specific measure with a national median of approximately 8%. The NLCD metric counts every hard surface pixel detected by satellite imagery — a total-imperviousness measure with a national median of approximately 66%. Both figures can be correct simultaneously; they answer different questions and should not be compared or substituted directly.

Coverage Limitations

The NLCD CONUS layers cover the contiguous United States only. Transit stations in Alaska, Hawaii, Puerto Rico, and the U.S. Virgin Islands will have blank tcc_pct and imp_pct values, tcc_category and imp_category set to 'No data', and heat_quadrant set to 'unknown'. Separate OCONUS NLCD layers exist but are not included in this version of the dataset.