

CY 2024 HMIS Demographic Dashboard Data Methods

AUTHOR

NJ - 504 Newark/Essex - Continuum of Care Planning

AFFILIATION

Essex County - Department of Citizen Services

PUBLISHED

Jun 11, 2025

ABSTRACT

This report provides an synopsis of the data preparation and cleaning methodology used to create the HMIS demographics dashboard instance.

Public Demographics Dashboard Initiative | All Systems & All Features Report

Why this Matters

Documenting data cleaning methodology enhances transparency, improves reproducibility, addresses biases and reporting errors, boosts trustworthiness, and educates the audience. Moreover, data cleaning provides an explanation for handling messy data, prevents misinterpretation, provides ethical rationalization for analysis, and set standards for future work.

About our Data

Data for this dashboard was extracted from the HMIS Report Builder, Demographic Report Builder, and Supportive Services Report.

- **Data Date Range:** *CY 2024*
- **Sample Size:** $n = 18,916$
- Data merged on Client ID
- All individuals counted once per project admission within given date range
- May contain duplicate individuals across projects
- NJ - 211 Admissions/Calls removed from dashboard view
 - NJ - 211 Admissions/Calls skew admission data as all calls are entered into HMIS

Selected Variables/Features for Analysis

- Project Admission Date
 - Date & Time variable
- Households with Children
 - Nominal variable
- Destination Category (exit data)
 - Nominal variable
- HMIS Project Type
 - Cross tabulation count of consumers per project
 - Cross tabulation percent of total consumers per project
- Client ID
 - Cross tabulation count of consumers
 - Unique identifier per project admission
- Personal ID
 - Nominal variable
 - NJ HMFA HMIS global unique identifier
- Domestic Violence Victim Survivor
 - Nominal variable
- Veteran Status
 - Nominal variable
- Chronic Homelessness at Admission
 - Nominal variable
- Race
 - Nominal variable
- Gender
 - Nominal variable
- Age Category
 - Persons aged 51 years and older
 - Boolean value
- Unsheltered
 - Cross tabulation count of consumers per Street Outreach project
- NJ SNAP
 - Boolean value
- Individual Family Type (Household Composition)
 - Cross tabulation count of consumers per household composition value (reported at project admission)

Data Cleaning, Modeling, & Feature Engineering

To date, as the data currently stands in the NJ HMFA HMIS database, some values require additional amendments in order to present a clean and uniform visual for business intelligence tools and dashboards.

The following methods were followed during the data preprocessing step.

Missing values

The following list includes data elements whose missing data was replaced with *"Data Not Collected"*:

- Individual Family Type (Household Composition)
- Gender
- Race
- Ethnicity
- Domestic Violence Victim Survivor
- Veteran Status
- Chronically Homeless at Admission

Domestic Violence & Veteran Status

Data elements with the following string/text objects were re-labeled into two new categories (*Data not collected & Client refused*):

- Client doesn't know
- Client prefers not to answer
- Null

Age

In some instances, specifically *Street Outreach Projects*, partial admission data is reported during project intake. The rationale is for the outreach worker to engage with the consumer over multiple observations, eventually collecting as much Personal Identifiable Information (PII) as possible. During this process, Date of Birth data is not collected at first observation. This leads to incorrect age values for consumers in these project types.

- *Age values above ninety-five (95) and less than zero (0) were replaced with the dataset median age value.*

Key Findings

Data visualization and dynamic filtering revealed gaps in sub-populations accessing homeless services in Essex County, New Jersey. Specifically, Emergency Shelter (night-by-night) programs service predominantly male consumers (over 81 % CY 2024).

In addition, Coordinated Entry consumers accessing services in the HMIS system were mostly women. Individuals over fifty-one (51) years of age access Emergency Shelter (entry-exit) more often than prevention services. Similar to Coordinated Entry data, single-parent families with Female head-of-households are accessing Emergency Shelter (entry-exit) services more than Male head-of-households.

Key performance indicators regarding Veterans, Seniors, Chronically Homeless, Domestic Violence Survivors, and Households with Children will be continuously measured for CY 2025.

In addition, further key programmatic data regarding program access will be added to bolster Essex County's quantitative findings to ensure that the allocation of homeless service resources is strategically distributed.

Limitations & Summary

Action steps

To facilitate resource deployment and program access efficiency, many organizations and NGO providers need access to local homeless service data. This dashboard helps fill the gap for locally aggregated homeless service data that can be used for future grant opportunities & renewal cycles.

Data caveats

One of the obvious concerns with data smoothing is if the data undergoes a drastic change then the insight can be limiting.

In addition, sample bias is always a threat as data was not re-sampled after missing values were filled. Due to these preprocessing steps the data is not a perfect sample of all system observations which could limit meaningful inference.