

Data Normalization Documentation

From Raw Data to 1NF, 2NF, and 3NF

This document outlines the data normalization process, transforming unstructured data into an organized and efficient structure across First Normal Form (1NF), Second Normal Form (2NF), and Third Normal Form (3NF). Each stage of normalization minimizes redundancy and dependencies, ensuring a streamlined and well-structured dataset.

Overview of Normalization Stages

- **1NF (First Normal Form):** Ensures that each attribute contains atomic (indivisible) values and eliminates duplicate rows.
 - **2NF (Second Normal Form):** Addresses partial dependencies by moving attributes to new tables if they rely on part of a composite key, ensuring that all non-key attributes are fully dependent on the primary key.
 - **3NF (Third Normal Form):** Removes transitive dependencies, ensuring that every attribute depends only on the primary key, further reducing redundancy and enhancing data integrity.
-

Table Summaries

1NF Structure

- **Person Table:** Contains personal details like First Name, Last Name, Race, Sex, Age, Marital Status, and Work. All fields are atomic, and duplicate entries are eliminated.
- **Family Table:** Manages family details such as Family ID, Family Name, and Dwelling, alongside HOH (Head of Household) details.

2NF Structure

- **Person Table:** Retains personal details.
- **Personal Attributes Table:** Created to store attributes like Race, Sex, Age, and Marital Status, separating them based on full primary key dependencies.
- **Family Table:** Further refined to include only family-specific data, such as Family ID, Family Name, and HOH details.

3NF Structure

- **Person Table:** Stores essential individual identifiers such as First Name and Last Name.
- **Personal Attributes Table:** Holds unique attributes tied to individuals, like Race, Age, and Sex.
- **Relationships Table:** Manages relational data, such as family relationships and household roles.
- **Marital Level Table:** Separates marital status information to avoid transitive dependencies.
- **Occupations Table:** Holds work information to ensure that occupational data remains independent.
- **Property Ownership Table:** Stores ownership and rental information.

Attribute	1NF	2NF	3NF	Description
Person ID (PK)	Person Table	Person Table	Person Table	Unique identifier for each person.
First Name	Person Table	Person Table	Person Table	First name of the individual.
Last Name	Person Table	Person Table	Person Table	Last name of the individual.
Attribute ID (PK)	–	Personal Attributes Table	Personal Attributes Table	Unique identifier for each personal attribute entry.
Race	Person Table	Personal Attributes Table	Personal Attributes Table	Ethnic or cultural background of the individual.

Sex	Person Table	Personal Attributes Table	Personal Attributes Table	Gender identity of the individual.
Age	Person Table	Personal Attributes Table	Personal Attributes Table	Age of the individual.
Marital Status	Person Table	Personal Attributes Table	Marital Level Table	Marital status, such as single or married.
Work	Person Table	Personal Attributes Table	Occupations Table	Current occupation or job title.
Owned/Rented	Person Table	Personal Attributes Table	Property Ownership Table	Indicates if the individual rents or owns their residence.
Relation to HOH	Person Table	Personal Attributes Table	Relationships Table	Defines relationship to the head of household (e.g., spouse, child).
Family ID (PK)	Family Table	Family Table	Family Table	Unique identifier for each family.
Family	Family Table	Family Table	Family Table	Family name or identifier.
Dwelling	Family Table	Family Table	Family Table	Description of the family's dwelling type.

HOH First Name	Family Table	Family Table	Family Table	First name of the head of household.
HOH Last Name	Family Table	Family Table	Family Table	Last name of the head of household.
Relationship ID (PK)	-	-	Relationships Table	Unique identifier for each family relationship.
Marital Level ID (PK)	-	-	Marital Level Table	Unique identifier for marital level entries.
Occupation ID (PK)	-	-	Occupation Table	Unique identifier for each occupation entry.
Ownership ID (PK)	-	-	Property Ownership Table	Unique identifier for each property ownership entry.