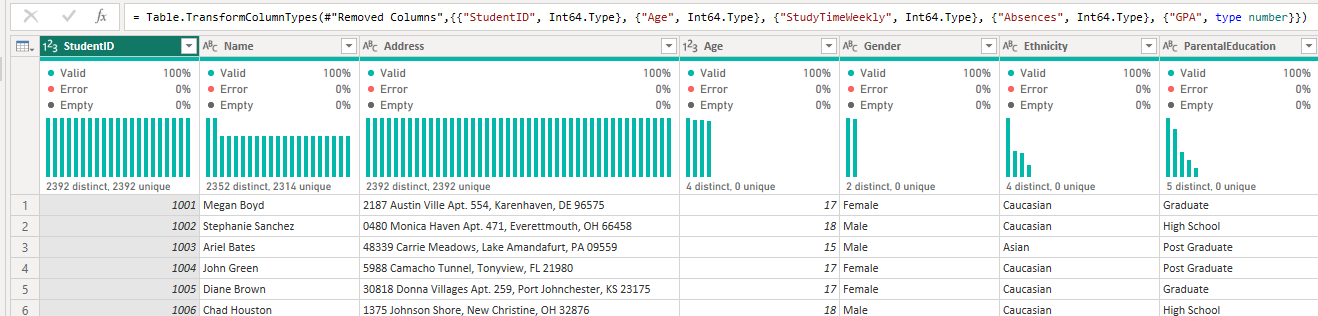
What is Normalization?

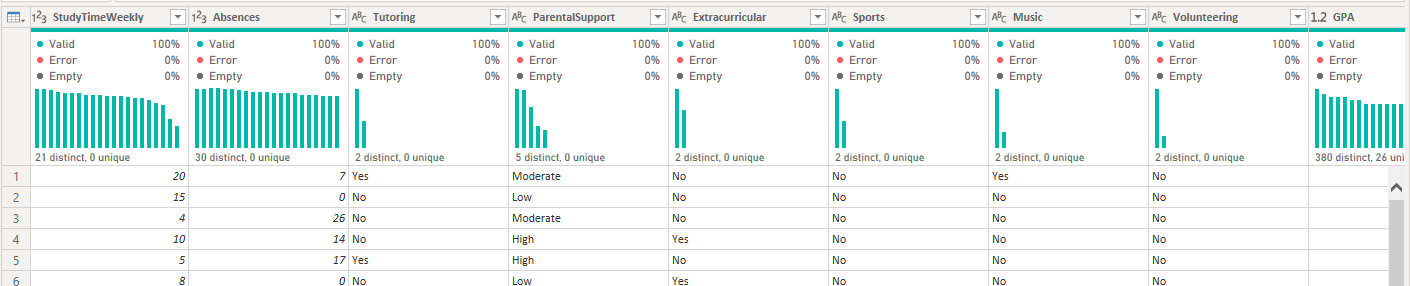
* Normalization is the process of organizing the data in the database.
* Normalization is used to minimize the redundancy from a relation or set of relations.
* Normalization divides the larger table into smaller and links them using relationships.
* The normal form is used to reduce redundancy from the database table.

Normalization Types

1NF - Ensure each table has a primary key and that each column contains atomic (indivisible) values.

All\_students\_Table:





* student\_id - A unique identifier assigned to each student (1001 to 3392) i.e. PRIMARY\_KEY
* student\_name - Name of the Student
* Address - Address where Students Resides
* Age - The age of the students ranges from 15 to 18 years
* Gender - Gender of the students
* Values - Male or Female
* Ethnicity - The ethnicity of the students
* Values - Caucasian, African American, Asian or Other
* Parental\_education - The education level of the parents
* Values - None, High School, Graduate, Post Graduate, PHD
* Study\_time - Weekly study time in hours, ranging from 0 to 20.
* Absences - Number of absences during the school year, ranging from 0 to 30
* Tutoring - Tutoring status (Yes or No)
* parental\_support - The level of parental support
* Values – None, Low, Moderate, High, Very High
* Extracurricular - Participation in extracurricular activities (Yes/No)
* Sports - Participation in sports (Yes/No)
* Music - Participation in music activities (Yes/No)
* Volunteering - Participation in volunteering (Yes/No)
* Gpa - Grade Point Average on a scale from 2.0 to 4.0, influenced by study habits, parental involvement, and extracurricular activities.
* grade\_class - Classification of students' grades based on GPA:
* 'A' (GPA >= 3.5)
* 'B' (3.0 <= GPA < 3.5)
* 'C' (2.5 <= GPA < 3.0)
* 'D' (2.0 <= GPA < 2.5)
* 'F' (GPA < 2.0)

2. Second Normal Form (2NF)

A relation will be in 2NF if it is in 1NF and all non-key attributes are fully functional dependent on the primary key (i.e.student\_id)

Split the data into multiple tables to remove partial dependencies:

* Students Table:

student\_id, student\_name, address, age, gender, ethnicity

* Parental Information Table:

student\_id, parental\_education, parental\_support

* Study Information Table:

student\_id, study\_time, absences, tutoring

* Extracurricular Activities Table:

student\_id, extracurricular, sports, music, volunteering

* Academic Performance Table:

student\_id, gpa, grade\_class

3. Third Normal Form (3NF)

Ensure the table is in 2NF and that all columns are non-transitively dependent on the primary key.

In this case, the tables are already in 3NF since there are no transitive dependencies

*Note: What is Transitive Dependency in DBMS?*

*A. Whenever some indirect relationship happens to cause functional dependency (FC), it is known as Transitive Dependency. Thus, if A -> B and B -> C are true, then A -> C happens to be a transitive dependency. Thus, to achieve*

*3NF, one must eliminate the Transitive Dependency.*