

### Keys

The distinctions between the primary key, candidate key and super key are that a super key is a field or set of fields that ensure every row will be unique as well as all of the columns. It is a super set of what is needed and will uniquely identify each row and column. A candidate key utilizes the fewest number of columns so that all of the fields are not needed. It is basically a concise super key or minimal super key. It is atomic and will throw everything away until it is able to stay unique. A primary key is a candidate key that I, as the designer, have chosen to be the primary key. It is unique and identifies all rows. It is a minimal super key that the designer has chosen from candidate key to primary key.

### Data Types

I might create a table for current students at Marist College to keep a record of their first and last names, hometowns, majors, and year. The name of the table would be Marist Student Records. This way the school could keep record of all of the current students at Marist and filter by first and last name, hometown, majors, and their year in school. None of these would be nullable unless a student did not have a last name or did not have a hometown. Otherwise a student has a first name, and a major (even if it is undecided) and a year in school. The data types used in the fields that Marist stores in First Name, Last Name, Hometown, Major, and Year would all be text. While the data kind used in all of these fields that Marist stores would be string of letters.

First Name	Last Name	Hometown	Major	Year
Pearl	Amin	Bridgewater, NJ	Finance	Senior
Girl	Lain	San Francisco, CA	Art	Freshman

### Relational Rule

The “first normal form” rule is important because it states that at every intersection of cells of rows and columns, the table must be atomic (indivisible). Therefore, it cannot be broken down more and must be unstructured (no lists). For example, a violation of this rule would be if in the FIRST NAME row, there was more than one name. In the table below, there should only be one name in the first name row, not many in a list because it then it is no longer atomic, breaking rule “first normal form” rule.

SID	FIRST NAME	LAST NAME
1	PEARL, ALAN	AMIN
1	PEARL	AMIN

The “access rows by consent only” rule, or the “what? Not where?!” rule is important because it talks about how we ask and interact with the data. The contents of the table, or what is there,

is more important than where it is because there is not guarantee that the rows will be in the same order when you rerun it. Because the order can change, the set has no original data. For example, it would be more important to remember that FIRST NAME is PEARL as opposed to remembering the location cell (2,2) because the data, when rerun, could change and the new information in cell (2,2) could end up being FIRST NAME is ALAN and that would not be the information that you are looking for, therefore, it is more important to remember the contents of the table as opposed to where the information is located.

The “all rows must be unique” rule states that in theory, the same name cannot be in the same row twice with the same details. Duplicate rows will give more than one and that is unnecessary and unneeded, so at least one thing must be unique. For example, the table below would violate rule number three because nothing is unique in these rows and the duplication is unnecessary.

SID	FIRST NAME	LAST NAME
1	PEARL	AMIN
1	PEARL	AMIN