## **Summary:**

**Relational databases**

A **relational database** is a database that contains a series of tables that can be connected to show relationships. Basically, they allow data analysts to organize and link data based on what the data has in common.

In a non-relational table, you will find all of the possible variables you might be interested in analyzing all grouped together. This can make it really hard to sort through. This is one reason why relational databases are so common in data analysis: they simplify a lot of analysis processes and make data easier to find and use across an entire database.

### **Database Normalization**

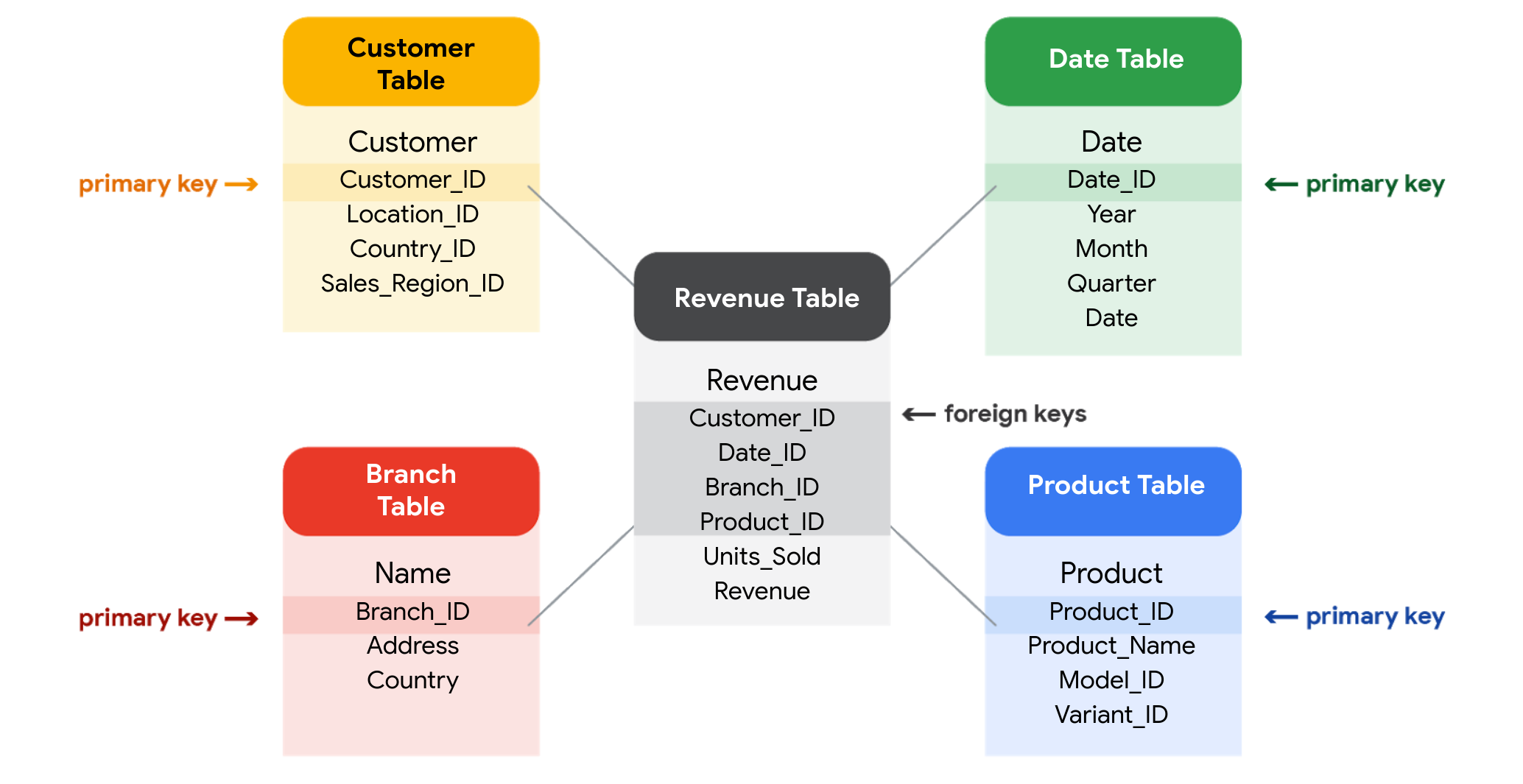
**Normalization** is a process of organizing data in a relational database. For example, creating tables and establishing relationships between those tables. It is applied to eliminate data redundancy, increase data integrity, and reduce complexity in a database.

### **The key to relational databases**

Tables in a relational database are connected by the fields they have in common. You might remember learning about primary and foreign keys before. As a quick refresher, a **primary key** is an identifier that references a column in which each value is unique. In other words, it's a column of a table that is used to uniquely identify each record within that table. The value assigned to the primary key in a particular row must be unique within the entire table. For example, if customer\_id is the primary key for the customer table, no two customers will ever have the same customer\_id.

By contrast, a **foreign key** is a field within a table that is a primary key in another table. A table can have only one primary key, but it can have multiple foreign keys. These keys are what create the relationships between tables in a relational database, which helps organize and connect data across multiple tables in the database.

Some tables don't require a primary key. For example, a revenue table can have multiple foreign keys and not have a primary key. A primary key may also be constructed using multiple columns of a table. This type of primary key is called a **composite key**. For example, if customer\_id and location\_id are two columns of a composite key for a customer table, the values assigned to those fields in any given row must be unique within the entire table.



**Metadata:**

Metadata is data that describes other data. It can be used to provide information about a file, such as its size, content, and when it was created. There are different types of metadata, including:

1. Descriptive metadata: This type of metadata provides information about the content of a file, such as its title, author, and subject.
2. Structural metadata: This type of metadata describes the organization of a file, such as how it is divided into chapters or sections.
3. Administrative metadata: This type of metadata provides information about how a file is managed, such as who has access to it and how it is stored.