

The Importance of Data Normalization

In today's business, data is one of the most important assets, and data analysis – one of the most important procedures.

And as with every important asset, we have to make sure that we take good care of our **data**. We need to actively seek out and remove every issue in our databases to be able to use the data to its fullest potential. Data normalization is a process created specifically for such a clean-up of databases. Considering the central role of data in business, it is a procedure that one cannot afford to overlook.

Picture 1 of The Importance of Data Normalization

Getting your data in shape

Data **normalization** can be compared to the idea of getting fit by taking care of your body. The main reason to do both is health. As we take care of our bodies to stay in good health, so should we take care of our datasets to keep them clean and in great condition.

And normalization of the database is also a kind of getting in shape. In this case, the shape refers to standard forms which we aim to give to our datasets, which would reduce the likelihood of errors to a minimum. There are six database arrangement ways, currently recognized as normal forms, following the naming introduced by Edgar F. Codd, the pioneer of data normalization.

These forms have a character of levels, meaning that every form from first to sixth is considered to be more normalized than the one before it. However, a database is thought to be sufficiently normalized when the third normal form is reached.

Specifically, normalization is a process by which redundancies and inaccuracies are methodically removed from the database. This is done by following a set of rearrangement rules that define the aforementioned normal forms. The information stored in the company's databases might be structured in various ways. Often, the structure is not clear and ordered enough to properly evaluate the data stored. Additionally, different databases might have conflicting information due to structural issues.

Putting data units into the columns and tables as required by such rules allows to see such issues as empty fields, redundancies, and inconsistencies of the data and fix them. The higher the level of normalization, that is the latter normal form, is satisfied, the better we are enabled to do this.

Important Benefits of Data Normalization

Thus, data normalization is a process that brings out issues with the data to the forefront by ordering the database. Once everything is in order, it is easy to make sure that new problems do not occur as data is updated

further.

The general benefits of such a procedure are easy to understand. Let us consider a few more specific advantages provided by data normalization, that make this process extremely important.

1) **Protection of data.** Unorganized data is at constant risk to be lost when handled due to such issues as deletion anomaly. Such processes when data is lost when the database is being updated is a costly and easily avoidable problem. After normalization data can be easily protected from being mishandled and misplaced among various databases. And as with money, saved data is earned data, thus normalization adds value to the assets of the business.

2) **Correction of data.** When data is presented in normal forms, conflicting or simply incorrect information is easy to notice. This allows to correct the errors and fill in the blanks, making data more consistent and complete. Incorrect data is actually worse than no data at all, as it leads to misconceptions about the company's assets which may turn into misinformed decisions. Thus, normalizing data may protect the company from huge making huge mistakes.

3) **Cleaning the database.** Redundant data gives data managers the wrong impression of how much data is actually owned by the company. For business intelligence even more important than knowing about other companies is knowing your own thoroughly. And as the importance of data in business is always growing, being mistaken about the scope of data owned is a big flaw in intelligence. One of the main purposes of data normalization is cleaning the database of such redundancies and making it clear how much and what kind of data is stored there.

Two paths to choose from

When a company owns a lot of data and especially when this data is stored in multiple databases, a process like data normalization is extremely necessary. When one arrives at this conclusion there are two possible paths to choose from.

The first is decomposition, which is an attempt to improve the existing design through rearrangement. Another is the creation of the new design by the process of synthesis. The right way to choose depends on the current state of the database. One should arrive at this decision after a thorough inspection of this state.

Whichever way is chosen, one can be sure that after normalization this state will be drastically improved.

You finished reading the article "**The Importance of Data Normalization**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.