Mitchell Xanders
CMPT-308L
January 26, 2016

Data vs. Information

The Elias Sports Bureau is a trusted source of official statistics recorded for almost all sports at the professional level worldwide. An example of data they could have would include the number set [4,388; 34; 11; 70.6; 109.6]. This data has no context, which leads to ambiguity and lowers the value of the raw data. Without any added context, these numbers are seemingly random. Given context of who, what, and when the data becomes information. Added context allows more less ambiguity and more precision, making the data more valuable.

Add the fact that these are the statistics recorded by NFL quarterback Drew Brees in 2009 and the data becomes information. This information is still somewhat ambiguous, however. These numbers happen to be Brees' total yards passing, touchdowns thrown, interceptions thrown, completion percentage, and passer rating, respectively. Given all this context, the data is now valuable information. It becomes relevant in discussing the statistics of other NFL quarterbacks and can be used in comparisons with other quarterbacks of any year. Additional context such as Brees winning the Super Bowl in 2009 can help analysts compare Brees' numbers with other Super Bowl-winning quarterbacks. By doing so, some analysts may try to find a trend among these Super Bowl winners and make more accurate predictions for future Super Bowl winners. This is not possible if the data is left without context. The raw data could lead to inaccurate comparisons due to ambiguity if not added context.

Data Models

The first hierarchical model was the Information Management System, which developed by IBM in 1968 when Rockwell sought their help in developing a system for NASA. This system follows the logic that each segment has one parent but may have multiple children. It has its shortcoming when a record or records exist in the system but cannot be traced through the nodes of the model structure to the initial parent node. It can also create duplicate records if more than one entity contain the same attribute.

The network model was adapted from the Information Management System, in that it is also a parent-child tree structure. The network model, however, allowed for records to have multiple parents as well as multiple children. This helped solve the issue of duplicate records when not necessary, but the problem still remained for when certain records could not be traced through the same number of parent nodes from the same root row.

The XML data model is an example of a hierarchical model that takes some inspiration from relational models. As a hierarchical model, it is easier to query an XML database than a relational database, and the nature of the XML language allows for easier understanding by both the human operator and the machine storing the data. XML data models are a great compromise of the hierarchical and relational database models.

Screenshot of pgAdmin Tool