: What is relational algebra?

by Daniel J. Power

Editor, DSSResources.COM

Relational algebra is a formal system, i.e., a well-defined system of abstract thought, for manipulating relations among data elements. Relational algebra helps explain how to join, combine, and find data. Relational algebra is a formal mathematical systems. Relational algebra is abstract and for many of us the concepts are challenging and difficult to understand. The general formalization helps one understand how data can be manipulated and queried.

The Wikipedia page on relational algebra begins "In computer science, relational algebra is an offshoot of first-order logic and of algebra of sets concerned with operations over finitary relations ..." Wikipedia's explanation is difficult to understand and it discourages most of us from continuing to read the explanation. Despite the annoying mathematical terminology and Greek language symbols associated with relational algebra, those of us interested in decision support and especially data-driven applications need to make the effort to learn the basics so we can better understand relational databases and query languages.

Supposedly relational database management systems (RDBMS) use relational algebra (RA) to represent queries internally to facilitate query optimization and execution. Some of my colleagues in Computer Science also assert that knowledge of relational algebra will help in understanding structured query language (SQL) and relational database systems in general. I accept both of these reasons for learning more about relational algebra as true. Given those reasons, we can not accept or dismiss relational algebra as irrelevant and esoteric. Even if it is esoteric knowledge, applied decision support researchers should join the initiated and acquire and exhibit knowledge that has been restricted to a small group of computer scientists and mathematicians.

So how can we provide a simple, adequate explanation of the esoteric topic of relational algebra that is not simplistic? Most of us have forgotten the Greek alphabet so a quick review will help. Then define the key concepts with concrete examples. Finally, translate and compare to SQL queries. Let's begin.

Relational algebra uses a symbol notation to indicate actions and sets of elements. The Greek letters used for the basic select and project operators are: 1) the symbol σ for select, and 2) the symbol ∏ for project. The symbol σ is the minuscule or lower case Greek letter sigma. The ∏ is the majuscule or upper case Greek letter pi, pronounced "pie". The lower case π is also used as a symbol for the verb project, pronounced pruh-**jekt**, meaning to throw or to estimate.

: What is relational algebra?

Author: Daniel Power
Last update: 2014-08-07 04:42