

HORIZONTAL AGGREGATION- GOLD IN COAL MINE

SHILPA KADAM¹, VIJAY BHOSALE² & UMESH KULKARNI³

¹Department of Computer, MGM College of Engineering and Technology, Mumbai, Maharashtra, India

²Department of Computer, MGM College of Engineering and Technology, Mumbai, Maharashtra, India

³Department of Computer, Vidyalankar Institute of Technology, Mumbai, Maharashtra, India

ABSTRACT

Datasets are widely used data mining, with columns in horizontal tabular layout to efficiently analyze the data. To extract trends and pattern from the historical data to prepare the future strategies, Data mining is the widely used domain. And this is a tedious task as it requires many complex queries, aggregation columns and joining tables. For statistical analysis, information gathered and represented in a summary form in data aggregation process. With one variable per column, surplus records plus horizontal layouts with data sets are the requirement of many algorithms. Data mining tools will create a SQL code representing horizontal aggregation as a template is the first.

Advantage of horizontal aggregation, which cuts shorts the human work in data preparation mode of data mining. End user written SQL code is inefficient then the code generated automatically is the second advantage plus this takes less time. Next advantage is that the data sets can be created in the DBMS.

KEYWORDS: Aggregation, SQL (Structured Query Language), SPJ, PIVOT, CASE, Data Set