MANAGEMENT STATISTICS USING DATABASES

Raul ŞERBAN*

Abstract. *In our days, open and low-cost access to appropriate, timely and reliable information and statistics on emerging economies and international financial flows is of heightened importance to policy makers, investors and civil society. Therefore this paper will present the possibility to use a management system with interconnected databases.*

Keywords: *system, database, statistic database, information fluxes.*

1. Introduction

The concept of the statistics has evolved over a period of time comprising many different facets of the organizational function. Statistics is a necessity of all the organizations. The initial concept of statistics was to process data from the organization and to present it in the form of reports at regular intervals.

In other words statistic means things like designing new products, streamlining a production process or evaluating current vs. prospective customers, today's business managers face greater complexities than ever before. Running a shop on instinct no longer suffices. Statistics provide managers with more confidence in dealing with uncertainty in spite of the flood of available data, enabling managers to more quickly make smarter decisions and provide more stable leadership to staff relying on them.

Management based on statistics in an organisation comprises the following processes:

a) Initiation of actions;
b) Ratification of initiatives;
c) Implementation of actions;
d) Monitoring of implementations and actions.

These viewpoints in Management statistics using databases are:

a) The viewpoint of quality management;

* Hyperion University, Bucharest, raulserban@gmail.com

b) The viewpoint of economics;
c) The empirical viewpoint;
d) The historical viewpoint;
e) The viewpoint of statistical programmes.

In order to speak about modern statistics we must refer to a system largely capable of handling the data from collection to processing

Not many years ago, processing information and statistic of a business was performing manual and transmission was done on paper. The invention of electronic computers has led to the idea that a modern information system can only be fully mechanized and automated. This idea was quickly abandoned because not all the management and execution can be standardized, thus remains a series of activities that cannot be driven solely by computer. In this component of the system people are operating information according to their meaning. Such processed information appears on the entry and exit, where people are identified them.

The primary objective of an information system for economic analysis is provided with maximum information necessary to the management board. We can appreciate that, the computerization of information systems is in fact a process of increasing their efficiency and therefore the notice of decision.

The level of computerization of an economic analysis system depends strictly on the level of complexity for decision aggregation and potential financing in commercial enterprise. Therefore, in order to achieve a permanent system, to be compatible with all platforms on the market but with a reduced share of the level of funding – Web 2.0 technology has allowed the appearance of such systems that run in web browsers on secure systems, allowing a total compatibility between existing platforms at the moment.

This paper will try to present the model of an information system that can make an economic analysis used for generating statistics.

2. Statistical database (SDB)

A statistical database is one that provides data of a statistical nature. The term statistical database is used in two contexts:

- Pure statistical database: This type of database only stores statistical data. An example is a census database. Typically, access control for a pure SDB is straight forward: Certain users are authorized to access the entire database.

- Ordinary database with statistical access: This type of database contains individual entries; this is the type of database discussed so far in this chapter. The database supports a population of users who are allowed access to selected portions of the database using DAC, RBAC, or MAC. In addition, the database supports a set of statistical users who are only permitted statistical queries. For these latter users, aggregate statistics based on the underlying raw data are generated in response to a user query, or may be precalculated and stored as part of the database.

In order to design an SDB we can first design the structures of triggers, query, tables, forms etc.

A form trigger in SDB is a set of SQL statements which affect record insertions, updates and deletions during form completion. Before creating triggers, it is important to fully understand the use of SQL in the database system.

The database system provides for several types of triggers in the construction of forms. First, it includes built-in triggers of its own. In addition, it enables you to create your own triggers:

– Column (or Field) triggers;
– Row triggers;
– Form triggers.

It should be noted that the term form trigger is used in two senses:

- generically, to refer to all triggers activated by either exiting a form column, exiting the row or entering/exiting the form;
- and to refer to a specific type of trigger – Pre-Form and Post-Form – that takes effect when the form is entered or exited, respectively.

One example of using triggers is when we manage warehouse balances using statistics.

Warehouse balances are affected whenever there is a receipt of goods from a supplier, a transfer of items from one warehouse to another or a shipment of goods to a customer. These three transactions are itemized in the *RECITEMS*, *TRANSITEMS* and *SHIPITEMS* forms, respectively. All three forms share the same base table: DOCITEMS.

Because of the similarities between these forms, a trigger in one form can easily be included in the other two. The following example makes efficient use of included triggers and included buffers.

The starting point was the *TRANSITEMS* form, as it records transactions for *both* receiving and sending warehouses (*RECITEMS* only refers to receiving warehouses, whereas *SHIPITEMS* only refers to sending

warehouses). We created two buffers: one to check warehouse balances and the other to update.

The Pre-Insert trigger for *TRANSITEMS* checks the balance of the sending warehouse, designated in the upper-level form (:$$.WARHS), since goods cannot be sent out of the warehouse if their quantity exceeds the warehouse balance. It is not necessary to check the balance in the receiving warehouse (the quantity received is in no way related to the existing balance). Thus, the trigger sets the values for the variables in the *WareBal* buffer as follows and then activates that buffer:

```
SELECT :$$.WARHS, :$.PROD, :$.QUANT
INTO :WARHS, :PROD, :QUANT
FROM DUMMY ;
#INCLUDE TRANSITEMS/WareBal.
```

The Pre-Delete trigger performs a similar task, checking balances in the *receiving* warehouse before allowing for record deletion. This is to ensure that the balance is large enough to cover the cancellation of the recorded receipt. The trigger is nearly identical to the Pre-Insert trigger, the only difference being the reference to: $$.TOWARHS rather than to: $$.WARHS.

The Pre-Update trigger checks balances for *both* receiving and sending warehouses.

3. Using Oracle managing optimizer statistics

In Oracle Database, optimizer statistics collection is the gathering of optimizer statistics for database objects, including fixed objects.

Because the cost-based approach relies on statistics, we must generate statistics for all tables and clusters and all indexes accessed by our SQL statements before using the cost-based approach. If the size and data distribution of the tables change frequently, then regenerate these statistics regularly to ensure the statistics accurately represent the data in the tables.

Oracle generates statistics using the following techniques:

- Estimation based on random data sampling;
- Exact computation;
- User-defined statistics collection methods.

To perform an exact computation, Oracle databases requires enough space to perform a scan and sort of the table. If there is not enough space in memory, then temporary space might be required. For estimations, Oracle databases require enough space to perform a scan and sort of only the rows

in the requested sample of the table. For indexes, computation does not take up as much time or space.

Some statistics are computed exactly, such as the number of data blocks currently containing data in a table or the depth of an index from its root block to its leaf blocks.

Oracle databases recommend setting the ESTIMATE_ PERCENT parameter of the DBMS_STATS gathering procedures to DBMS_ STATS.AUTO _SAMPLE_SIZE to maximize performance gains while achieving necessary statistical accuracy. AUTO_SAMPLE_SIZE lets Oracle databases determine the best sample size for good statistics.

For example, to collect table and column statistics for all tables in the OE schema with auto-sampling:

EXECUTE
DBMS_STATS.GATHER_SCHEMA_STATS('OE',DBMS_STATS.AUTO_ SAMPLE_SIZE);

To estimate statistics, Oracle selects a random sample of data. You can specify the sampling percentage and whether sampling should be based on rows or blocks. Also the documentation of Oracle databases recommends using DBMS_STATS.AUTO_SAMPLE_SIZE for the sampling percentage. When in doubt, choose row sampling.

- **Row sampling** reads rows without regard to their physical placement on disk. This provides the most random data for estimates, but it can result in reading more data than necessary. For example, in the worst case a row sample might select one row from each block, requiring a full scan of the table or index.
- **Block sampling** reads a random sample of blocks and uses all of the rows in those blocks for estimates. This reduces the amount of I/O activity for a given sample size, but it can reduce the randomness of the sample if rows are not randomly distributed on disk. Block sampling is not available for index statistics.

When you generate statistics for a table, column, or index, if the data dictionary already contains statistics for the object, then database updates the existing statistics. Database also invalidates any currently parsed SQL statements that access the object.

The next time such a statement executes, the optimizer automatically chooses a new execution plan based on the new statistics. Distributed statements issued on remote databases that access the analyzed objects use the new statistics the next time Oracle database parses them.

When you associate a statistics type with a column or domain index, Oracle database calls the statistics collection method in the statistics type, if you analyze the column or domain index.

4. Conclusions

Database economic statistics systems in many companies are largely in their infancy. However, with rapidly growing usage of mainframe, minicomputers and microcomputers we can expect to see considerable developments in this area in the next few years. There is a good deal to be done by many companies to improve their analytic systems. To achieve this firms need help and assistance in choosing a system which best meets their own particular need. Obtaining a system is not by itself the solution to the problem, however. Employees need to be trained both how to use the system and how to operate it.

The primary objective of this paper work has been highlighting the importance of the databases in designing an economic statistical system that can be used in any enterprises. As we mentioned in this paper, information, specifically the power of information, is the main competitive advantage of companies on the market.

The idea of this statistic system summarized in this article, not only manages to combine into a universal and easy way the two concepts: that the dynamic web language and databases interconnection, but it manages to bring a new concept – the software can be adapted to any economic activity and can use any multiple database.

REFERENCES

[1] Hideki Takayasu, Misako Takayasu, Mitsuhiro P. Okazaki, Kouhei Marumo, Tokiko Shimizu, *Fractal Properties In Economics.*

[2] Hong Zeng, *Fixed Income Database Design & Architecture*, Worcester Polytechnic Institute, 2005.

[3] http://library.thinkquest.org/26242/full/ap/ap7.html.

[4] http://mattfleming.com/node/192

[5] B. B. Mandelbrot, *The variation of certain speculative prices*, J. of Business (Chicago) 36 (1963), pp. 394-419.

[6] J. Kertesz and I. Kondor (Eds.), *Econophysics: an emerging science,* Kluwer Academic Publisher, Dordrecht.

[7] R. Mantegna and E. H. Stanley, *Introduction to Econophysics,* Cambridge Univ. Press, London, 1995.

[8] H. Takayasu, A.-H. Sato and M. Takayasu, *Power law behaviours of dynamic numerical models of stock market prices,* in Econophysics, 1998.

[9] H. Takayasu and M. Takayasu, *Critical fluctuations of demand and supply*, Physica **A 269** (1999), pp. 24-29.
[10] D. Sornette, D. Stauffer and H. Takayasu, *Market fluctuations II: multiplicative and percolation models, size effects and predictions*, (Proceedings book of conference in Rauischholzhausen 1999, to appear (Eds. A. Bunde) (http://xxx.lanl.gov/abs/condmat /9909439)).
[11] H. Takayasu, H. Miura, T. Hirabayashi and K. Hamada, *Statistical properties of deterministic threshold elements – the case of market price*, Physica **A 184** (1992), pp. 127-134.
[12] R. N. Mantegna and H. E. Stanley, *Scaling behaviour in the dynamics of an economic index*, Nature, 376 (1995), pp. 46-49.
[13] Oracle 9i *Database Administration*, Published by Oracle Corp.
[14] Len Boss, From Carnegie Mellon, *Attribute Driven Design Method*.
[15] Baicu F., *Security and audit of information systems,* Victor Publishing House, Bucharest, 2006.
[16] Bob A., Visean M., Fulea M., Saseanv A., *Information systems in Commerce*, ASE Publishing House, Bucharest, 2003.
[17] Gheorghiu A., Bichis C*, Databases*, Victor Publishing House, Bucharest, 2004.
[18] Mitru D., *Integrated management of the company,* ASE Publishing House, Bucharest, 2004.
[19] Moga T., Radulescu C., *Management Fundamentals,* ASE Publishing House, Bucharest, 2004.
[20] Nicolescuo, Verboncv I., *Organization Management Fundamentals*, Economics Tribune Publishing House, Bucharest, 2001.
[21] Paun M., Hurtulari C., *Analysis, Diagnosis and Evaluation of Systems in the Economy,* ASE Publishing House, Bucharest, 2004.
[22] Pop A. D., *General Management*, ASE Publishing House, Bucharest, 2005.
[23] Scarlat E., Chirita N., *Economic Systems' Cybernetics*, ASE Publishing House, Bucharest, 2003.
[24] Stanciu V., Gavrila, Mangiuc D., Sehleang, *Financial-monetary information systems*, Economics Tribune Publishing House, Bucharest, 1999.
[25] *Management and Organization Dictionary*, Politics Publishing House, Bucharest, 1985.
[26] Raul Serban, Ana-Maria Ramona, *Use Of PHP Language and MYSQL Databases in Information Systems*, Titu Maiorescu University, Bucharest, 2008.
[27] Raul Serban, *Using Fractal Theory in Databases Design,* Victor Publishing House, Bucharest, 2011.
[28] http://www.dbtalks.com/uploadfile/anjudidi/open-system-interconnection-model-in-oracle/
[29] http://docs.oracle.com/cd/E16655_01/server.121/e15858/tgsql_s tats.htm# TGSQ L390.