Connect ▸ **Support** ▸ **Advance**

The Institute of
Internal Auditors
Australia

White paper

Data Analytics using SQL

UPDATED 2020

Data Analytics using SQL

Contents

Background

Purpose

Data Analytics tools broaden the investigative weaponry that an auditor has at their disposal. Using joins, compares, logical operators, sums and counts gives an auditor a fast and efficient population view as opposed to a sample view.

Internal auditors can use analytics to:

- Identify possible frauds.
- Test entire populations rather than sampling.
- Save time and effort.
- Test for regulatory compliance.
- Identify trends or patterns.
- Test data quality.
- Identify business improvement opportunities.

Background

This white paper is intended to give the reader a starting point for initial implementation of SQL analytics in an organisation. As the paper is targeted for beginners, the tool adopted for education throughout the paper is Microsoft Access which is a relatively inexpensive and easily obtained interrogative tool for an organisation.

The white paper will explore the issues surrounding limitations of current tools, the history and evolution of analytic tools, and in the discussion examples of code and associated logic will be given to demonstrate uses of the technology.

Discussion

Issue

Patterns or anomalies in information are identified once data is aggregated. Generally, data does not mean much in isolation. In order to find relationships amongst data sets, the use of SQL enables disparate data sets to be matched using a common identifier, such as a member or customer number.

History

Tools that auditors initially have used for auditing can be very limited in their capability. For example, the use of Microsoft Excel is limited by the way it can match (or 'join') various data sets. It can also be limited by the number of records being manipulated. In today's data-rich environments, more and more data is being stored, and hence more manipulation is required.

Other audit tools (names withheld) require the user to input program parameters prior to run-time, and generally an entire routine is required to be run before results can be obtained.

Modern tools such as Microsoft Access, SAS, and Teradata allow the user to enter the number of rows to be returned, to get a better feel for the output prior to execution of a larger routine.

Discussion

Essentially, the concept of SQL will be shown below, together with some practical examples, some statements and their use.

SQL = Structured Query Language, a programming language which converts user entered statements into machine language. It is dependent upon data being stored in tables, otherwise known as a relational database management system (RDBMS), with the information in those tables related to one another by a 'key'. A key is a common piece of data with exactly the same data replicated in one table to another,

Data Analytics using SQL

for example the customer number mentioned above.

Below is the data structure shown for a customer's transaction records at a Bank.

Customer_Master_table

Name	Customer_ number	Address_ Line_1	Suburb_City	Post-code
Fred Smith	X12345	1 Happy Street	Happy Valley	5999
Jane Johns	C9988	4 Somewhere Street	Johnstown	2100

Customer_Transaction_table

Customer_number	Tran_date	Amount	Type
X12345	01/01/2016	20	W
X12345	02/01/2016	30	D
C9988	03/02/2016	85	D

One of the key concepts of a RDBMS is relationships. As you can see, one table (Customer Master) records all customer related information in one table, and another table (Customer Transaction) records customer transactions. Using the Customer Number, a 'join' is performed using SQL and all transactions can be retrieved for a customer using customer number, as the relational piece of joining data.

Based on other conditions such as date, amount or type, a join can enable a user to select customer attributes in one table, and attributes from another table. In this case, the arrow represents this join, as Customer Number X12345 has highlighted related records in the Customer Transaction table to that of the Customer Master table. More than one join can occur, and various pieces of data / information can be sourced from numerous tables.

Example of SQL

Below, a useful feature of SQL is explained. Imagine you wanted to obtain a list of transactions for Deposits (Type = "D")

Coding wise, this would be coded as:

```
SELECT Customer_Transaction_Table.Customer_Number,
Customer_Transaction_Table.Tran_Date, Customer_
Transaction_Table.Amount, Customer_Transaction_Table.
Type
FROM Customer_Transaction_Table
WHERE (((Customer_Transaction_Table.Type)="D"));
```

The output would then be:

Query1

Customer_number	Tran_Date	Amount	Type
X12345	01/01/2016	$30.00	D
C9988	03/02/2016	$85.00	D

If you wanted to obtain more information about the customer, the power of SQL becomes apparent. For example, you may want to know transaction patterns for all people who live in different areas. Assume you wanted to know specifically about all transactions made by customers who live in Johnstown. The code then becomes:

```
SELECT Customer_Master_Table.Customer_Name,
Customer_Master_Table.Customer_Number, Customer_
Master_Table.Suburb_City, Customer_Transaction_Table.
Tran_Date, Customer_Transaction_Table.Amount,
Customer_Transaction_Table.Type
FROM Customer_Master_Table INNER JOIN Customer_
Transaction_Table ON Customer_Master_Table.Customer_
Number = Customer_Transaction_Table.Customer_Number
WHERE (((Customer_Master_Table.Suburb_
City)="Johnstown"));
```

You can see from above, an additional SQL keyword (INNER JOIN) has been introduced into this query. This key word / statement tells the program to perform functions which allow for further processing to be undertaken and output to be displayed. In this example, the output would be as follows:

Query1

Customer_name	Customer_ number	Suburb_ City	Tran_date	Amount	Type
Jane Johns	C9988	Johnstown	03/02/2016	$85.00	D

You can imagine how powerful this would be when you have thousands of transactions for a customer in Johnstown. Suburbs or cities could then be grouped and summed and sorted to determine patterns, trends etc. Taking raw data and converting this into insights and an easy to understand summary view is the real power of SQL programming.

Some SQL keywords and their uses

The following keywords perform the relevant functions

SELECT – User specifies what database fields are being selected in the query (eg name, address etc).

Data Analytics using SQL

FROM — The table(s) being queried.

JOIN — Specifies a way of combining data from two tables whenever there are matching values in a common field, and forms the tables into a single result.

WHERE — Conditional expressions. For example, "WHERE name = 'Smith'". Can be used with less than or greater than expressions (< >) and like operators, e.g. "like (%complaint%)".

This is very useful when a user wants to search for keywords in a field of information, such as "like %complain%" (will return complaint, complained, complains etc), "like %angry%" etc. These expressions will return all the text in a field where these words are present, so text mining is very useful in obtaining information from within larger sets of information.

ORDER BY — Outputs can be sorted in ascending or descending order.

There are a number of useful advanced SQL functions, such as Common Table Expressions (CTE). CTE can be used to return information in a hierarchy, such as a list of staff and their managers starting with the highest-ranking officer, screenshot below.

CTE returning employment hierarchy

Staff Name	Manager	Manager Title	Staff ID	Manager ID
Fred	Jim	CEO	123	678
John	Fred	HR Manager	400	123
Amy	Fred	HR Manager	300	123
Rachel	Fred	HR Manager	200	123

Conclusion

Summary

SQL programming is an efficient way to perform analysis on databases. Enormous amounts of information (millions of rows) can be condensed (grouped) into easy to understand output in a fraction of the time other tools take.

Conclusion

Any audit function should embrace and use analytics wherever possible. They can often be undertaken with standard office software. The time and cost savings will be measurable, and so will the insights gleaned from the results.

Bibliography

Institute of Internal Auditors, Inc. (2011, Aug). GTAG 16: Data Analysis Technologies. Retrieved from The Institute of Internal Auditors: Global:
https://global.theiia.org/standards-guidance/recommended-guidance/practice-guides/Pages/GTAG16.aspx

Microsoft Corporation. (n.d.). Accessing and Changing Database Data. Retrieved from TechNet:
https://technet.microsoft.com/en-us/library/ms181080(v=sql.105).aspx

Data Analytics using SQL

Purpose of White Papers

A White Paper is an authoritative report or guide that informs readers concisely about a complex issue and presents the issuing body's philosophy on the matter. It is meant to help readers understand an issue, solve a problem, or make a decision.

Author's Biography

Written by: Malcolm Webster
BComm(Mgmt), GradDip(CIT), DipFP, MIIA

Mal Webster is Chief Risk Officer of Australian Mutual Bank Limited (AMBL), a secure and modern financial institution offering access to a wide range of products and services through enhanced delivery channels combined with a discounted fee structure.

During his career, Mal worked in a variety of risk and audit roles. At the Commonwealth Bank of Australia, he worked on retail network audits, analytics and ICT audits. At Police Bank, he was the head of internal audit and introduced analytics efficiencies to the wider internal audit team. Prior to that, Mal worked for ASX Top 50 companies ANZ, TabCorp and Westfield in a variety of roles.

Mal is a firm believer that analytics can quantify business problems by scanning and concluding on the entire population, therefore resulting in better business outcomes.

Edited by: Michael J A Parkinson
BSc(Hons), GradDipComputing, PFIIA, CIA, CISA, CRMA, CRISC

About the Institute of Internal Auditors–Australia

The Institute of Internal Auditors (IIA) is the global professional association for Internal Auditors, with global headquarters in the USA and affiliated Institutes and Chapters throughout the world including Australia.

As the chief advocate of the Internal Audit profession, the IIA serves as the profession's international standard-setter, sole provider of globally accepted internal auditing certifications, and principal researcher and educator.

The IIA sets the bar for Internal Audit integrity and professionalism around the world with its 'International Professional Practices Framework' (IPPF), a collection of guidance that includes the 'International Standards for the Professional Practice of Internal Auditing' and the 'Code of Ethics'.

IIA–Australia ensures its members and the profession as a whole are well-represented with decision-makers and influencers, and is extensively represented on a number of global committees and prominent working groups in Australia and internationally.

The IIA was established in 1941 and now has more than 200,000 members from 190 countries with hundreds of local area Chapters. Generally, members work in internal auditing, risk management, governance, internal control, information technology audit, education, and security.

Copyright

Disclaimer