MOC 55232 A: Writing Analytical Queries for Business Intelligence

Course Summary

Description

This three-day instructor led course is about writing TSQL queries for the purpose of database reporting, analysis, and business intelligence. Specifically, this course presents TSQL within the context of data analysis – in other words, making meaning from the data rather than transaction-oriented data-tier application development.

The course starts with a brief discussion of levels of measurement and quantitative research methodology, and integrates these concepts into each TSQL topic presented. The goal is to provide a consistent, direct, and purposeful learning path for RDBMS data retrieval for use in analytical tools such as SQL Server Reporting Services, PowerBI, R, and Excel.

Objectives

After taking this course, students will be able to:

- Identify independent and dependent variables and measurement levels in their own analytical work scenarios.
- Identify variables of interest in relational database tables.
- Choose a data aggregation level and data set design appropriate for the intended analysis and tool.
- Use TSQL SELECT queries to produce ready-to-use data sets for analysis in tools such as PowerBI, SQL Server Reporting Services, Excel, R, SAS, SPSS, and others.
- Create stored procedures, views, and functions to modularize data retrieval code.

Topics

- Introduction to TSQL for Business Intelligence
- Turning Table Columns into Variables for Analysis: SELECT List Expressions, WHERE, and ORDER BY
- Combining Columns from Multiple Tables into a Single Dataset: The JOIN Operators
- Creating an Appropriate Aggregation Level Using GROUP BY
- Subqueries, Derived Tables and Common Table Expressions
- Encapsulating Data Retrieval Logic
- Getting Your Dataset to the Client

Audience

This course is intended for information workers and data science professionals who seek to use database reporting and analysis tools such as Microsoft SQL Server Reporting Services, Excel, Power BI, R, SAS and other business intelligence tools, and wish to use TSQL queries to efficiently retrieve data sets from Microsoft SQL Server relational databases for use with these tools.

Prerequisites

Before attending this course, students must have context knowledge of data analysis and business intelligence scenarios. For example, an understanding of a work-related business intelligence project or need. Students should have basic knowledge of the Windows operating system and its core functionality, including file system navigation. Basic understanding of the purpose of relational database management systems such as SQL Server is also required.

Duration

Three days

ProTech *"Charting the Course ...*

protechtraining.com

... to Your Success!"

MOC 55232 A: Writing Analytical Queries for Business Intelligence

Course Outline

I. Introduction to TSQL for Business Intelligence

This module discusses writing analytical queries vs. transactional DML queries, and describes the typical architecture of a business intelligence environment. It discusses the role of SELECT queries in retrieving data for analysis from relational databases. It introduces the sample database to be used in the course, and begins a presentation of the SELECT query.

A. Two Approaches to SQL Programming
B. TSQL Data Retrieval in an Analytics Environment
C. The Database Engine
D. SQL Server Management Studio and the CarDeal Sample Database
E. Identifying Variables in Tables
F. SQL is a Declarative Language
G. Introduction to the SELECT Query

Lab:

- Create a database diagram
- Create and execute basic SELECT queries

II. Turning Table Columns into Variables for Analysis: SELECT List Expressions, WHERE, and ORDER BY

This module covers the identification of and relationship between levels of measurement and column data types. It continues a discussion of the SELECT query and adds the WHERE and ORDER BY clauses.

A. Turning Columns into Variables for Analysis
B. Column Expressions, Data Types, and Built-in Functions
C. Column aliases
D. Data type conversions
E. Built-in Scalar Functions
F. Table Aliases
G. The WHERE clause
H. ORDER BY

Lab: Write queries using:

- Column and table aliases
- DISTINCT
- WHERE
- ORDER BY
- Built-in functions
- Explicit and implicit data type conversion

III. Combining Columns from Multiple Tables into a Single Dataset: The JOIN Operators

Module 3 discusses creating single datasets for analysis by combining results from multiple database tables using JOIN.

A. Primary Keys, Foreign Keys, and Joins
B. Understanding Joins, Part 1: CROSS JOIN and the Full Cartesian Product
C. Understanding Joins, Part 2: The INNER JOIN
D. Understanding Joins, Part 3: The OUTER JOINS
E. Understanding Joins, Part 4: Joining more than two tables
F. Understanding Joins, Part 5: Combining INNER and OUTER JOINs
G. Combining JOIN Operations with WHERE and ORDER BY

Lab: Write SELECT queries using:

- Inner join
- Left, right, and full join
- Joins of more than two tables
- Join operators, in addition to WHERE and ORDER BY

IV. Creating an Appropriate Aggregation Level Using GROUP BY

This module covers the aggregation of quantitative column values across grouping factors for the purpose of group-wise comparisons and/or changing the granularity of a dataset.

A. Identifying required aggregation level and granularity
B. Aggregate Functions
C. GROUP BY
D. HAVING
E. Order of operations in SELECT queries

Lab: Write queries using:

- Aggregate functions
- Aggregate function with HAVING
- Aggregate function with GROUP BY and HAVING
- Aggregate function with GROUP BY, HAVING, WHERE, and ORDER BY

MOC 55232 A: Writing Analytical Queries for Business Intelligence

Course Outline (cont'd)

V. Subqueries, Derived Tables and Common Table Expressions

This module covers the use of subqueries, derived tables, and common table expressions in SELECT queries as techniques for creating intermediate result sets.

A. Non-correlated and correlated subqueries
B. Derived tables
C. Common table expressions

Lab: Write queries using:

- Non-correlated subqueries
- Correlated subqueries
- Derived tables
- Common table expressions
- Subqueries, derived tables, and common table expressions and other topics

VI. Encapsulating Data Retrieval Logic

This module discusses the encapsulation of data retrieval logic in views, table-valued functions, and stored procedures. It also describes scenarios in which these techniques are useful for producing datasets for analysis. Finally, it describes the database security issues involved, and techniques for creating and using these database objects while maintaining current permission sets on source data.

A. Views
B. Table-valued functions
C. Stored procedures
D. Creating objects for read-access users
E. Creating database accounts for analytical client tools

Lab:

- Create a SQL login
- Create a database user and assign required permissions
- Create a database schema for views, functions, and stored procs
- Create a view
- Create a table-values function
- Create a stored procedure
- Allow a user with read-only access to use views, table-valued functions, and store procedures

VII. Getting Your Dataset to the Client

This module covers common techniques for making datasets produced by SELECT queries available to analytical client tools such as SQL Server Reporting Services, PowerBI, Excel, and R. It discusses running queries directly from the client tool, in addition to exporting datasets to text files which can then be accessed by the client tool.

A. Connecting to SQL Server and Submitting Queries from Client Tools
B. Connecting and running SELECT queries from:
C. Excel
D. PowerBI
E. RStudio
F. Exporting datasets to files using
G. Results pane from SSMS
H. The bcp utility
I. The Import/Export Wizard

Lab:

- Retrieving the results of a view in Excel
- Running an ad-hoc SELECT query from Excel
- Running an ad-hoc query from PowerBI
- Running an ad-hoc query from RStudio
- Using the Import/Export wizard to write the results of a query to a text file