User-Defined Functions

Riley Major
@RileyMajor

This is a repeat after me song.

Happy Llama, Sad Llama,
Totally Rad Llama

Super Llama, Drama Llama,
Big Fat Mama Llama

Baby Llama, Crazy Llama,
Don’t Forget Barack O’Llama

Moose! Fish! More Fish!

Unicorn! Peacock! More Fish!

See, repetition is fun, right?

But not in your code.

Fundamentally, the purpose of a user-defined function is to prevent code repetition.

- Scalar
- Table-Valued
  - Multi-Statement
  - Inline

Scalar

- Accepts parameters.
- Returns a value.
- Used in SQL as if it were a variable or literal value.
- Contains a bunch of SQL statements.
- Opaque to Query Optimizer.
- *Might* be slow.

Multi-Statement Table-Valued

- Accepts parameters.
- Returns rows with pre-defined columns.
- Used in SQL as if it were a table.
- Contains a bunch of SQL statements.
- SQL statements must INSERT into output table.
- Opaque to Query Optimizer.
- *Might* be slow.

Inline Table-Valued

- Accepts parameters.
- Returns rows with derived columns.
- Used in SQL as if it were a table.
- Contains a single SELECT statement.
- Transparent to Query Optimizer.
- No slower than equivalent SQL.


•APPLY
•JOIN
Yes
No
Working Weeken
seven years prior to the application
felony less than
automatic disqualification
EMPLOYMENT

Restrictions

- Read Only.
  - No external data modification.
- No “side-effecting operator”.
  - No NEWID().
  - No RAISERROR.
- Must pass all parameters.
  - Default values require DEFAULT keyword (unless using EXEC).

https://msdn.microsoft.com/en-us/library/ms186755.aspx

Why bother?

- Don’t Repeat Yourself (DRY)
  - Less code to write.
  - Less code to maintain.
  - Consistent behavior across application.
- Smaller Code Chunks
  - Less likely to conflict with another developer.
- Eaiser to Test
- Easier to Understand
- Encourages Set-Based Logic (Discourages Cursors)

But…

- Can hide complexity.
- Can cause performance problems if misused.

Suggested Architecture

- Separate Business Logic from Data Logic – all in SQL.
- Outer “Get” function gathers relevant information.
- Inner “Calc” function computes result or logical controls.
  - Deterministic.
  - Easy to add Unit Testing.
- No logic in Stored Procedure SELECTs– they just call functions.
- Stored Procedures only used for data modification or by external apps.
- Die-hards can use XML to pass “arrays” to Calc functions.

Determinism

- A given input will always have the same output.
- Pure calculation.
- Easy to test.
- Can’t use any tables (data could change).
- Can’t use system data (e.g. current date, as that changes).
- Can’t use random numbers.
- Can only call deterministic system functions.
- WITH SCHEMABINDING allows indexing of computed columns.

Future?

- Inline Scalar
  - There’s a connect item for that.
  - Go vote!
- Named Parameters
  - There’s a connect item for that, too.
  - Closed as Won’t Fix. Vote anyway.
- Arrays?
  - Table Valued Parameters are quite limited.
  - Can use XML as a workaround. Good for code, bad for performance.

https://connect.microsoft.com/SQLServer/feedback/details/273443/the-scalar-expression-function-would-speed-performance-while-keeping-the-benefits-of-functions

https://connect.microsoft.com/SQLServer/feedback/details/361449/table-valued-udf-scalar-udf-functions-named-parameter-input

Image Credits

- Llama
  - 6582898025_d17c76e44c_o.jpg
  - Gabriel Gonz·lez
  - https://www.flickr.com/photos/gaby1/6582898025

- dry
  - 114348973_2e83cf28d2_o.jpg
  - Jared Tarbell Follow
  - https://www.flickr.com/photos/generated/114348973

- 3
  - 26565220686_3eff3de883_o.jpg
  - Andy Maguire Follow
  - https://www.flickr.com/photos/andymag/26565220686

- The grindstone
  - 5621810815_848d0578cc_o.jpg
  - Kathryn Decker Follow
  - https://www.flickr.com/photos/waponigirl/5621810815