**The Business Case for Corporate Use of Stored Procedures (SPs)**

Purpose: This document provides the business case for organizations to create and leverage a library of stored procedures to automate ETL and dashboard refreshes.

In this organizational view of IT usage, a stored procedure is a saved query that is stored in your database server. SPs are used by database administrators and business analysts to accomplish (and thankfully automate) many data manipulation and data management tasks. This document discusses SP usage in report automation and ETL processing.

Once you develop, validate and perfect a query, you never have to write it again, hooray! Rather you can save the query as a stored procedure. Many developers create a library of stored procedures which is stored functionality awaiting usage. When a report is needed there is no time to write a query, so DBA/analysts are required to listen to reporting needs and produce the reports. Rather than produce reports when management requests it, if the report is often needed, the reporting process can be automated with SPs. One DBA/analyst can produce hundreds of reports, dashboards and datasets which supply an organization with the data they need to make decisions. The stored procedure then is the IT that is used to add organization and automation to the query/reporting process.

Business managers indirectly (an unknowingly) use SPs when they display a report or refreshed dashboard, because when the report is run, the SP fetches and compiles the data behind the scenes (as compared to hours of Excel ‘work’). Automation of the data refresh process is important because the value added activity is viewing the report and making a decision. An employee fiddling with Excel data is not a value added process. There is never a shortage of information requests, new metrics requested, and changes to reports. DBA’s could never serve management’s new report requests if they were manually producing queries and reports. Query/reporting needs to be automated for many reasons, if no other that the work day can be more orderly and shorter. A manager saying “I need this report before you go home” becomes a less common event.

Overall SPs are an IT asset that DBA’s create to organize data and provide access to compiled data. SPs are used to build datasets that reports and dashboards can be built upon. Firms that use SPs to refresh reports and dashboards are more efficient in their data management. It is very common that employees store data in Excel spreadsheets, and then spend hours each week to pull different spreadsheets together to produce the analysis and reports needed. This Excel copy/paste and column creation process continues to be viewed as a value added clerical activity, however businesspeople don’t really want to create the report in Excel (even if they feel empowered), they want to use the report to analyze business performance and make decisions that enable continuous improvement. SPs automate the report/dashboard refresh process, eliminating the need for manual calculations and formulas which (because we are human) are at known to be wrong 10% of the time.

SPs are also used to automate the data merging, integration, manipulation, and storage that together form ETL processes. There are many data manipulation and data management processes needed (such as backing up data), SPs are the way to automate them. SPs can be scheduled to run using SQL triggers.

The purpose of this document then is to make the business case for SPs, to describe their benefits, and discuss the managerial best practices for using them.

Why use SPs? - Stored procedures are used for many reasons

1. SPs turn DBA effort into managed corporate IT assets – SPs are the vehicle to save the work of DBA and data engineers. By building a library of saved queries, the SPs can be cataloged and leveraged. If an employee creates amazing queries but does not save and share them, what if the employee leaves the organization for other opportunities? By instilling a SP library and governance structure, useful SQL queries can be understood, vetted, and documented giving stability to their use. If a DBA leaves the company, their queries can continue to be used with confidence. Using SPs then enable a governance structure to be built, which can then document the different ways that data is being manipulated.
2. SPs as a time saver - SPs are saved effort and have a value as measured in labor hours. SPs can pull data together from different tables and data sources and represent hours or days of effort. If you write a query that produces the intended resultset, or produces the needed data management, then by saving the query as a SP *you never have to write the query again.* Rather than perform data manipulation in Excel (or similar) in a manual process, the SP provides automation of the data transformation process. Could you imagine if corporate ETL processes were manual? Oh wait that describes many small companies.

When a new dataset (or report) is needed the DBA can review their library of SPs (and reports) and hopefully repurpose an existing IT asset thereby saving time, effort and aggravation. Analysts and managers that need a dataset can first check the SP library for a solution or partial solution. Similarly they can look thru a catalog of reports that use SPs for their datasource and calculations.
3. SPs enable reliable, repeatable data processing - SPs can be used to build automated processes. SPs do need to be tested for fragility based on dirty data (data that can’t be correctly parsed can crash a process or just provide a subset of the calculations needed. So SPs need to be carefully tested and made to not ‘assume’ incoming data in correctly formed and available. As compared to manual data processing, SPs are much less likely to fail and inject errors into the data. Due to removing human data manipulation labor, the data management process can be repeatable, reliable and error-free. SPs just work. Again and again. They do need to be updated however.
4. SPs as a way to rapidly disseminate new metrics – if a new metric is created that for example helps to identify important retail trends, if the metric is inside a SP stored on a database server, then the metric can be pushed out to the entire organization in an instant. If reports are based on SPs, then as the SPs evolve the reports refresh immediately, nationwide. Compare this to the pockets of brilliance which are common in analytics and reporting. The spreadsheet or Tableau report that one analyst creates in one region will more slowly be adopted by other analysts in other regions because ‘it wasn’t built here’.

If SPs are managed, cataloged, documented and therefore trusted then they can be re-used, and repurposed. Larger companies that have many DBAs and analysts, the advances made by one DBA can be shared quickly to other DBA’s if the SPs are viewed as vetted, managed assets.
5. As a mechanism to improve and create consistent decision making processes – Managers make decisions based on the information at hand, personal experience, intuition, staff etc. Related to the above point, a company can have too many reports across regions. The use of reports can be haphazard with little consistency in how the reports are used to make decisions. Each manager has the ability or request special purpose reports, but the reporting process for a firm can be unorganized. Reports that use approved SPs can improve the consistency of the reporting process as managers know which reports are used to analyze problems.

Innovation will always occur and is to be encouraged, however firms can use expanding sets of updated SPs to keep the managers making decisions on validated data.

In summary this author believes data belongs in databases, and when it needs to be analyzed, if the analysis is repetitive such as weekly or monthly, then the dataset and/or report should be produced in an automated “hands off” process using SPs. Human labor can introduce errors and manually produced reports are rarely checked for accuracy of the numbers. Data management processes can be analyzed, vetted, and treated as IT assets when saved as stored procedures.