★ ★ ★ ★ ★ ★ ★

Is MS Excel Really a Database?

★ ★ ★ ★ ★ ★ ★

2010

PCA Business Notes Series

The short answer: Yes, to a point.   Excel can be used as a "database" to store and manage business data, provided your needs do not exceed Excel limitations in the areas data integrity, security and multi-user support.

MS Excel allows you to manage data in two dimensions (Rows and Columns), and provides database-like functions like hlookup and vlookup, sorting and filtering (and the like), that enable some level of reference and re-use of data elements.   And with enough custom VB code, you can get your Excel spreadsheet to look and function like a real business application!

Managing Vs. Storing Data

One of biggest drawbacks of using Excel as a database: the data model that is necessary to optimize how your data is STORED – ideally, a single instance of every field element with no duplication – is often at odds with the data model that is ideal to MANAGE your data. It is very difficult for example to reference existing data elements inside Excel to insure a perfectly normalized set of data, in the same way that a relational database is designed to support.

Because of the difference between a data model that is optimized for storage and the data model that is suitable for managing edits, adds, and deletes, Excel developers often end up creating two (or more) instances of the same data element.

Whenever data is managed in a "denormalized" state (duplication of the same data element), it is only a matter of time before the data becomes fragmented and very cumbersome to maintain, and time and effort required to maintain an acceptable level of data integrity can become quite expensive! Attempting to produce reliable reports from a denormalized Excel spreadsheet is usually when this effort becomes most apparent.

Multi-User Limitations

Another major limitation of using Excel as a database is the inability to support multiple, concurrent End Users. Popular work-arounds to multi-user limitations include: putting the spreadsheet on a shared network drive, and for Internet access, exposing this shared drive to the Internet with a remote access program like Citrix or GoToMeeting. What neither these work-arounds will address however is locking out fields when more that one User is attempting to update information, and the resulting train wreck that occurs. Common consequences of using MS Excel in a multi-user environment include corrupt data, lost data, Excel crashing, or all of the above.

Excel is usually adequate to support 2 or 3 people involved with managing the data. This is most often accomplished by emailing around the spreadsheet, and keeping track of "who's got the ball." Any more than 2 or 3 Users however, and you're going to be spending a lot of time cutting-and-pasting data from different sources, and always wondering if you have the latest, most up-to-date and accurate set of data.

Security Limitations

Excel supports password protection for the entire workbook, for individual worksheets (tabs), and even protecting individual fields from being changed by Users. The security methods that Excel provides are usually adequate for small groups E.g. 2-5 Users.

If you need to support larger numbers of Users, you're much better off using a real database application like MS Access, or even SQL Server for more robust security management controls like role-based security, and/or integrating the database with network-wide security mechanism like Active Directory.

PCA Recommendation

No matter how much time and expertise you invest, MS Excel is a spreadsheet that is designed to support a small handful of End Users, and will never yield the levels of data integrity, scalability, ease-of-use or security that are the hallmarks of true database platforms like Oracle and SQL Server.

Excel is a great tool to move large amounts of data into a database from another application. As an alternative to building a direct interface via ODBC or API, Excel's CSV format is a very common method to share data among different applications.

Because Excel is so flexible and easy-to-use, with powerful built-in charting functions, Excel provides a great tool to display different views of business data, but falls short of providing many capabilities that tend to become important over time, especially within distributed multi-user environments.