

# CIS64A Spring 2020

## Homework 1

Joshua Saunders

4/20/2020

**Question:**

In business operation, companies can store data using Microsoft Excel, that is a “spreadsheet” software, not a DBMS. Why do many businesses need a real database management system instead of using Excel? Give the details of the reasons.

**Answer:**

Microsoft Excel certainly has its usefulness in business, from creating charts and graphs to performing financial analyses and more. However, something that Microsoft Excel, and spreadsheets in general, do not excel at is acting as a store for business data. That is not to say that spreadsheets can’t be used as a database management system (DBMS), it’s that spreadsheets shouldn’t be used as a DBMS.

An important aspect of business information is ensuring that every employee that needs that information has access to it. This can be a problem when using Excel for a number of reasons [2]. First, the same file can be used by different versions of Microsoft Excel (such as Office 365, the web, Android, or Excel 2019), but the problem is that not every version of Excel has the capability to share access to a file with other users (Microsoft calls this “co-authoring”). When a user uses a version of Excel that’s incompatible with co-authoring on a shared file the file becomes locked, preventing other users from using the spreadsheet. If an Excel spreadsheet is used as the database for a business and it becomes locked making it not possible for others to get the information necessary to conduct business. By using an actual DBMS this problem is avoided because a DBMS allows for concurrent access to the data meaning that multiple people have access to the data at the same time [6].

Having an Excel spreadsheet as a DBMS instead of an actual DBMS prevents customers from having access to a business’s data. An example of this is with an e-commerce website. Currently, Excel limits the number of concurrent users to 256 if the file is set to allow changes by more than one user [3]. This is important because a large e-commerce website may have millions or billions of visits per month, such as Walmart.com and Amazon.com, which had approximately 388M and 2.2B visits in the month of March 2020, respectively, which means that a spreadsheet would be unsuitable as a DBMS for an e-commerce site [7, 8]. An actual DBMS such as MySQL can handle 100,000 simultaneous connections out of the box [5].

There is also a potential for different people or departments within a company to create different spreadsheets duplicating the data which can then diverge as time goes on [6]. If the data is stored in a central DBMS then this problem is avoided.

Another problem with using an Excel spreadsheet as a DBMS is that of security. Although Excel does offer some security features such as encrypting the file and read-only access, there exist tools that allow encrypted Excel files to be easily cracked and for someone with access to the spreadsheet to copy sensitive data formulae [4, 1]. An actual DBMS provides authorization features that can be applied to restrict users from certain actions such as reading certain tables or from deleting records [6].

Lastly, in an Excel spreadsheet formulae are embedded in the spreadsheet itself and are viewable to anyone with access to the spreadsheet. This can be a problem because this means that the formulae can be viewed or changed without the proper protections enable on the spreadsheet if the formulae are proprietary, thus exposing valuable intellectual property. With an actual DBMS, the application logic can be hidden behind a program layer and not be exposed to users [6].

# Bibliography

- [1] Software Security Ltd. Limitations of excel's 'protection'. <https://www.spreadsheetsentry.com/limitation-of-excel-protection.html>, 2019. [Online; accessed 19-April-2020].
- [2] Microsoft. Excel file is locked for editing by another user. <https://support.office.com/en-us/article/excel-file-is-locked-for-editing-by-another-user-6fa93887-2c2c-45f0-abcc-31b04aed68b3>, 2020. [Online; accessed 19-April-2020].
- [3] Microsoft. Excel specifications and limits. <https://support.office.com/en-us/article/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3>, 2020. [Online; accessed 19-April-2020].
- [4] Microsoft. Protection and security in excel. <https://support.office.com/en-us/article/Protection-and-security-in-Excel-be0b34db-8cb6-44dd-a673-0b3e3475ac2d>, 2020. [Online; accessed 19-April-2020].
- [5] Oracle. Mysql reference manual. [https://dev.mysql.com/doc/refman/8.0/en/server-system-variables.html#sysvar\\_max\\_connections](https://dev.mysql.com/doc/refman/8.0/en/server-system-variables.html#sysvar_max_connections), 2020. [Online; accessed 19-April-2020].
- [6] Abraham Silberschatz, Henry F. Korth, and S. Sudarshan. *Database system concepts*. McGraw-Hill, 2020.
- [7] SimilarWeb. Amazon.com march 2020 overview. <https://www.similarweb.com/website/amazon.com>, 2020. [Online; accessed 19-April-2020].
- [8] SimilarWeb. Walmart.com march 2020 overview. <https://www.similarweb.com/website/walmart.com>, 2020. [Online; accessed 19-April-2020].