Databases - They're all around us
By Phil Sorrentino, Contributing Writer, The Computer Club, Florida
The Journal of The Computer Club, Inc.
http://scccomputerclub.org / Philsorr.wordpress.com
philsorr (at) yahoo.com

Think about it; during the course of the day we might use a telephone directory, a dictionary, an encyclopedia, an airline flight guide, a bibliography. All of which are databases. We use databases that store all kinds of knowledge that we use on a regular basis. At home or in your office, the paper filing system you have set up is really a database. But as the computer has become more a part of the fabric of our homes and offices, we use more databases that are resident on our home computers. Databases are one of the reasons that make home computers so useful. Databases show up in many places and are at the heart of many computer software programs or Apps. Many of our databases are "on-line", available only through the use of a browser on your home computer. Some of these databases include your Healthcare Providers Formulary, the Library of Congress Online Catalog, and Facebook.

Your contact list is a database, as are your collection of MP3 music, and your collection of friends (in a social media environment, such as Facebook, or LinkedIn, or Twitter). We all use a multitude of databases, most of the time without knowing anything about the database details or operations. Software applications allow us to use the database and extract the information that we are looking for. For example, take Google. Google has created an enormous database of locations on the internet and through its application, at www.Google.com, allows a user to search the database and come up with websites that might provide answers to their questions. Many years ago, I authored a Nutrition Analysis program that provided the totals of 21 nutrients for a person's one day food intake. The basis of this App was a Food Nutrition database that held the amounts of each of 21 nutrients for each food. At that time, the database was only available in handbook form so the 21 values for each of around 2500 foods had to be entered into a computer database so that the App, "Nutriplan", could extract the values when the user selected the particular food. The App provided the answers and reports, but the database held all of the possible nutrition information for each individual food.

Just as background, a database is an organized collection of data. Data, in the home computer sense, are words, numbers, images, (and maybe even moving images, a.k.a. videos), represented in a structure, possibly tabular or graphical. Spreadsheets are an easily understood representation of databases. A typical spreadsheet represents something of interest in the real world, on the computer as rows and columns. Usually, for an item database, the rows are the items and the columns are things that describe the item. So, to expand on my example of the Food Nutrition database, let's say the rows are the food names, and there are 21 columns for the nutrients. The rows might be "Apple", "Pear", "Peach", "Banana", etc. Each column would be another nutrition component; for example, column 1 might be protein, column 2 might be vitamin A, column 3 vitamin C…, I think you get the idea, all the way up to column 21, maybe fiber. You could even add a column that had a picture of the food, or a recommendation for

how the food could be prepared. Once the database is defined, a clever programmer could develop all sorts of Apps that could produce useful nutritional evaluations, (once the user had input the day's food intake).

This type of row and column database is typically referred to as a "flat file", maybe because it seems to have two dimensions, row and column. A more powerful database organization is called a Relational database. A Relational database is a collection of data items organized as a set of tables. The tables are connected by "keys" or key variables. The "keys" allow data to be accessed quickly and efficiently. Software that uses or manipulates a Relational database is called a Relational Database Management System (RDBMS). Access and Oracle are examples of an RDBMS. These complex database systems are typically manipulated by using SQL (Structured Query Language), a database language, which has become the standard (circa 1987) for defining, manipulating and managing Relational Databases. (For those of you who have been around computers since the beginning (circa 1982), you might remember dBase II. dBase II was the first widely used database management system for microcomputers. dBase II included a database engine and a database language all in one package. dBase II was around before SQL, and coincidentally lost favor as the industry switched to SQL as it became the standard for Relational Database manipulation. By the way, there never was a dBase I, the developer thought the II in the name of the software package would make it more acceptable (read that salable). The II was just a marketing trick. "Ah, marketing, where the rubber meets the sky.")

Access and Oracle are both examples of an RDBMS, but are at the extreme ends of size and complexity, although each uses SQL as the managing software. Oracle is a professional system designed for very large, critical use, databases where rapid, multiple accesses are anticipated. Access, although a very capable RDBMS, is designed for smaller database projects (possibly those projects that can be easily handled by a small group of clever programmers). Access is on many business systems, and is ubiquitous because it is part of the more expensive version of Microsoft Office.

So, you see, databases are at the core of some very simple Apps, like Nutriplan, and some very comprehensive Apps, like Google.com. Undoubtedly, due to the size and scope of the data involved, Google probably employs the Relational database model. Nutriplan, on the other hand, because of its relative simplicity, worked very well with the flat file model.