Clarence V. Kimbrell Jr.
Assignment #6

When discussing databases there are a few things that need to be mentioned first. The units of measurement that are associated with a database. There is almost an upside-down pyramid of measurements starting with bits and then going up a layer to bytes. Then with many bytes, it becomes a field then with many fields becomes a record, and then a file. At the top of the pyramid, we have databases which is a huge collection of files of data. An entity is a part of a database that categorizes generalized things, this can include a supplier, a place, and so on. Within an entity, some attributes are characteristics of the entity mentioned above. These are like descriptions of the entity, and these things are like specific addresses and item descriptions. A key field is almost like a title for each record that uniquely identifies it from others. A relational database is the most commonly used today. These are the types of databases that most think and are aware of. These types of databases are separated into two different tables. Each table contains information about entities and their attributes. They are normally split into the supplier table and the customer table to allow for ease of access for the user. There are a lot of parts within databases and one of the final things that help construct one is entity-relationship diagrams and normalization design. These are diagrams that clarify the relationship between tables in relational databases such as one-to-one or many relationships. Normalization of design as stated before helps the user quickly and efficiently find information by simplifying data.

Databases require a lot of effort to create and maintain which is why many big businesses like google and amazon created programs to help future business owners create their own within their services. Database management systems commonly referred to as DBMS is a type of software that helps create databases. And allows its user to store, organize, and access said information from the database. Most of these programs support relational databases since they are the most common. It benefits the organization by alleviating the problem of understanding the ins and outs of the foundational support that provides the databases with their functionality. This saves a lot of time for businesses which could be used in other aspects. There are three main capabilities of DBMS and those are data definition, data dictionary, and querying and reporting. The data definition can be confused with the data dictionary because they sound similar and their functionality is also similar. Data definition is the specific structure of the content of databases. The data dictionary is almost like a file sorted. It stores definitions of data and its contents. Then we have querying and reporting which also include things such as data manipulation. Which is used to add, delete, or edit information in databases.

The very cleverly named big data is exactly what it states. This is a huge amount of data that programs such as DBMS are not capable of computing. Big data are normally used in the field of engineering, science, and social media because calculations in the trillions are not necessary for a small business inventory. Analyzing big data comes much quicker than traditional databases because with their far more sophisticated technology they can compute things way faster. There are many components of business intelligence infrastructure and those are data warehouses, data marts, Hadoop, in-memory computing, and analytical platforms. The data warehouse can be described as a database that stores current and historical data of potential interest to decision-makers throughout out the company. If all things are documented correctly historical data can be used in the future decision-making process to alter ideas for the better. Data marts are smaller sections of a data warehouse because the focus group of information is separated from the entire

content for a specific set of users. Hadoop is an open-source program that organizes unstructured data or semi-structured data in large quantities. In-memory computing is done by utilizing the computer's ram to compute calculations on big data analysis. Normally this is done with are far better computer than what you would find in the common household. Analytical platforms were developed by businesses to analyze large data sets. This pre-configured software is designed for query processing and analytics. Online analytical processing supports multidimensional data analysis. This allows users to see data from different angles without distributing said data.

Data mining is something I'm aware of by being a gamer. This is essential in digging for information that is not exactly readily available to everyone yet. In the gaming world if there is an update for a game there are thousands of data miners ready to see what was added before the developers release the content to everyone. This can relate to databases too because they can be used to see patterns and sequences and use them to predict future behavior. The types of information that can be found but are not limited to are associations, sequences, classifications, clusters, and forecasts. This is different from OLAP because this type of data discovery is not obtainable within these programs. Text mining is normally used for things such as surveys, patent descriptions, and emails. Businesses use text mining to analyze data from these sources to extract key elements from unstructured data. Web mining is the discovery and analysis to find patterns and information from web sources. These can be used to study customer behavior and quantify the success of a marketing campaign.

In my pc repair business, the use of databases will be great. I've stated that I wanted to have a database that has customer's pc parts that will allow for more understanding behind each computer. With the knowledge of which types of databases, I will use the relational one for a customer's name and the following components in their computer. This information can be used on a website that is called pc part picker which will allow me to see what upgrades or fixes will best suit the customer's price range. With the use of this type of database, it will speed up repairs a lot instead of manually seeing if things are compatible or not. Then I can have more information about each customer like visit frequency and what they purchase. This will need to provide patterns that I can then use to improve my business.