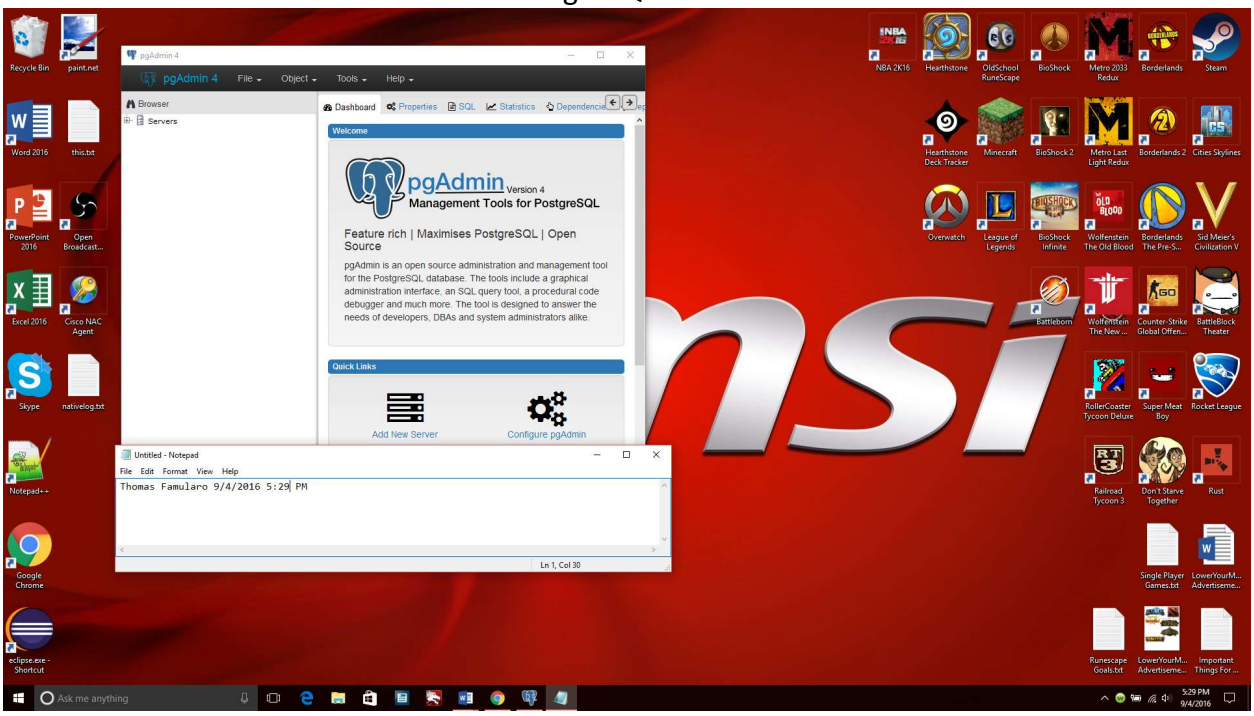
Thomas Famularo 9/4/2016

Lab 1 – PostgreSQL

1. 

2. Data vs. Information

Data and Information in everyday conversation are used interchangeably, but in the context of Computer Science and Data Base management they are different. The main difference is that information is data with context. It gives meaning to the numbers and facts that data contains. Data is just a number for example “5”, it is just a number that has no context. It could mean anything form how many Tacos I had for lunch to the score I would give the new IPhone 7. But once you give it context like “5 people” it starts to mean something. But just adding units is not enough, there are many ways “5 people” could be interred. But once you put more context for example, “5 People got of the seven train at Flushing Main Street on 9/5/2016 at 4:44 AM” you know exactly what that piece of data means, it turns in to information. If you were to record how many people got on and off that train all day over a month you could figure out patterns and see if all of the times the train stop at that station is necessary. Maybe on Thursday at 2:43 0 people got on or off the bus, so you might not want to stop that train at that station any more. Turning that data into information gave us the ability for us to make a more educated decision on the train service schedule.

A data base that I use every day is the relational database I use at work in the applications department of IT. It has data like:

4	13	Thomas	Famularo	200-64-092	8/7/2015 8:56 PM	0	0	1

Some of that data you might know what it means if you knew who I was and my CWID number but the rest you have no idea, but once I label the columns like this:

ID	RID	First Name	Last Name	CWID	ACID	ADID	DF	EF
4	13	Thomas	Famularo	200-64-092	8/7/2015 8:56 PM	0	0	1

It starts to make sense. This is information, it comes together to describe one specific user of the database. It gives you a complete picture of the User to help make decisions on all sorts of other things, like what roles should this user have, what permissions and much more. There are thousands of rows of users with all of this information in the table. Even with over one thousand rows of users the information is still extremely easy to read and compare to others information.

3. Data Models

There are many different types of data models that many smart people have come up with throughout the years of computing. From flat files at the beginning of computing to the Relational and Graphical Models that we use today, data models have improved a tremendous amount. Starting with the first actual system of data management the Hierarchical Model. The Hierarchical Model made many advances compared to the flat file system. It was set up like a family tree with one root at the top and split down the tree with the higher layers containing the lower layers. For example, one data base that could work in a Hierarchical Model are self-hygiene products. With Self-Hygiene products being the root and Shampoo and Body Wash and Tooth Brush being its children and under those there would be the brands of those products. Although it may seem that this Hierarchical system may work for this example it does have flaws. One being that a child cannot have two parents, which creates duplicate data. Another problem is that data might not fit under any of the parents so it should not have any relation to a parent but the system requires all of its data to have a parent except the root because of it being a Hierarchical model. The Network model fixes one of these problems so it is an improvement, it fixes the problem so that data can have more than one parent. This allows for most data to be stored only once. But this model still has problems, data still has to have a parent, there is no built in reporting system, and you still need to know how the entire model works to get any data out of it. When compared to a relational data base these two models fall short. They lack the flexibility that a relational model uses which makes it much harder to include all of the data may be required. The other major problem is data reporting, there is no system in Network or Hierarchical models. SQL or Structured Query Language allows for specific information to be called from the database and returned without the code writer having to know how the entire data base is set up. The Relational model also allows for information to be included without having a parent like in the other models. It can be included in a table but just not have any connections.

XML is another type of system that is used in data management, it is set up very similarly to a 2D array. XML is an array of objects, those objects have certain fields and information associated with them. The different objects are the rows in the 2d array and the fields are the columns and the cells are the information. This allows for it to be quickly set up and put into use if small amounts of data are being input. If there is a lot of information, data entry in this set up will take forever. XML does have a major shortcoming with not having functions like SQL. It does not have the ability to make complex queries without the user having to write loops and value checking them self's. This prevents XML from being that useful in any complex situation when compared to the relational model.