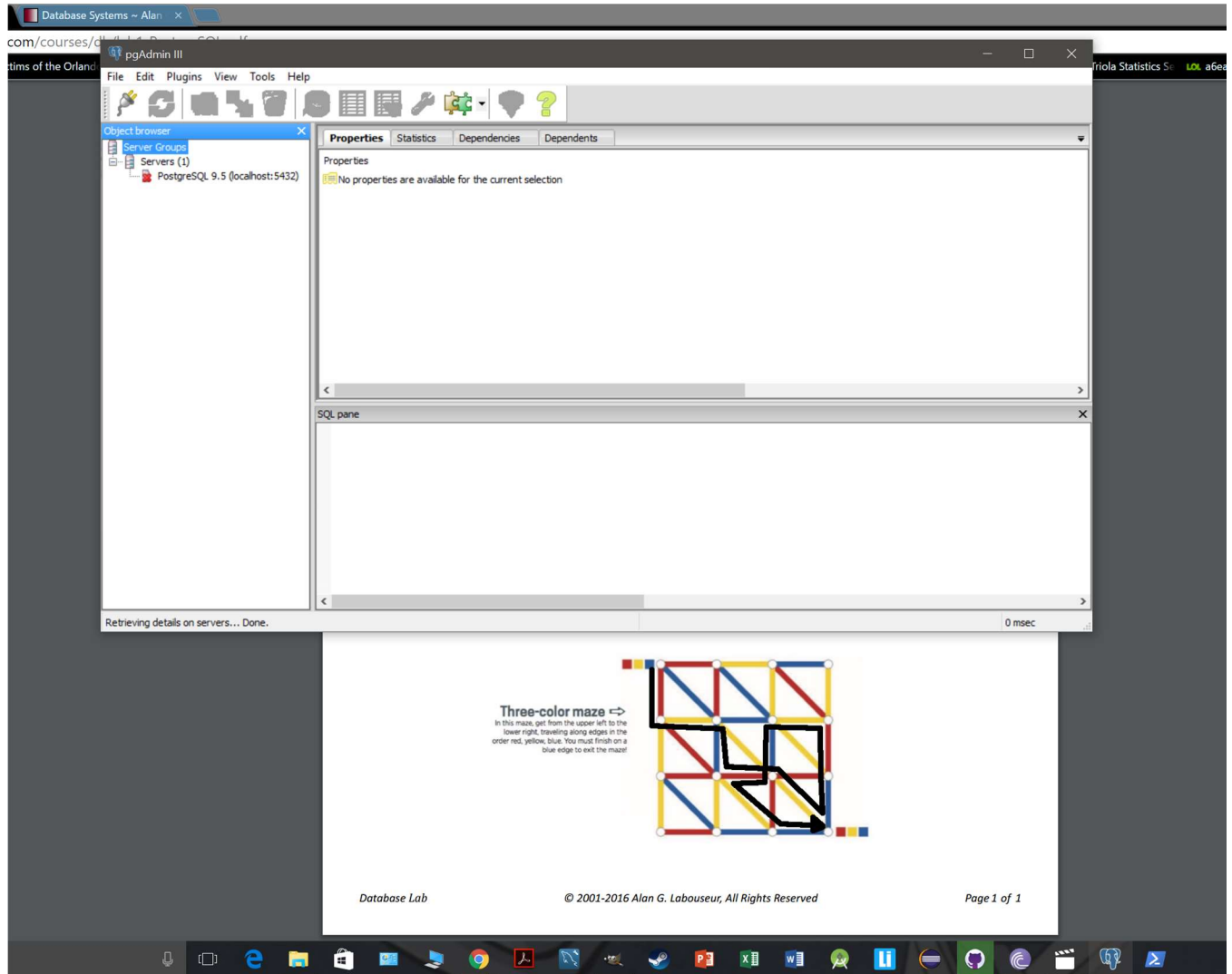


Michael Read

September 8, 2016

PgAdmin Screenshot



Michael Read

September 8, 2016

Short Essay – Data vs. Information

When you work with an airline to book your flight for your summer vacation, that airline is using many databases to keep track of the flights available, where those flights are going, who the crew of the flight will be, as well as so many other pieces of information about that flight. Without the context behind the database tables, looking at the data stored inside of them would be meaningless. If you pulled out a single number, 108 for example, you would have no idea whether that number was the price of some flight, or how many flights a certain pilot had flown, or the capacity of a certain aircraft. It is only when you give context to the data that you get the information saying aircraft 37 has a capacity of 108 passengers. The tables give these random bits of data context by organizing them into rows and columns, so that each entry in a specific column has the same meaning, and each entry in a row identifies a certain attribute of the row. In other words, rows represent an entity of what the table is, while the columns represent the attributes of the entities in the table. With this organization, the data becomes information that we can use to make more informed decisions.

Information is just data without context, and without that context there is no meaning. For example, looking at the string of numbers “1 23 4 668 25 4 7” without being given any context, it is very hard to come up with what that data means. It could be a phone number, or a series of grades a student got in a course, the possibilities without context are nearly endless. Once you are given the context of these numbers, for example these numbers represent the number of people who looked at a certain webpage over a one week period starting on Sunday, you can start making decisions about the data, and drawing some conclusions as well. From these numbers, we can see that on Wednesday the data seemed to spike, so we could figure out what happened on Wednesday to draw in the much higher view count, and see if that information could be used to help increase the view count for the rest of the days of a week. Being able to have data with context is a very powerful tool to help make informed decisions that can help to steer you in the right direction.

Michael Read

September 8, 2016

Short Essay – Data Models

The first model created for data storage was the Hierarchical Model, which used a tree model to describe relationships. This was good because it made it easy to transfer data from one machine to another, making it physical data independent. However, there was a problem with redundancy, as well as some other problems that came about because of the Hierarchical nature of the model. There were no loops or cycles allowed in the model, and you couldn't have any unrelated data, which could be a problem for some types of applications. One other problem with this model was the fact that you needed to know the structure of the model in order to correctly work with the data.

The next advance was the Network Model, which took the Hierarchical, and allowed you to add in cycles and loops, which solved the problem of duplication. This model still had the problems of unrelated data being difficult to incorporate, as well as the fact that you needed to know the structure of the model before working with it.

Finally, Dr. Edgar Codd came up with the idea for the Relational Model. This model solved all of the problems that the previous two systems had, by making the model into a collection of tables of rows and columns. This solved the problem of duplication, as all of the pieces of data are entered into the tables once, and then whenever they need to be related to some other entry, they can be referenced by pieces of data called foreign keys. With this system, whenever any piece of data needs to be changed, you can simply change it in once place, and every place that piece of data occurs knows the change. It also makes it easier to learn the structure of the model, as the id's of the columns are built in to the model.

Looking at the development of these models, I do not believe that XML is a good model for data storage, compared to the Relational Model. XML would work best with a small amount of data, but apart from being able to organize pieces of data by tags, there is no way to easily relate two or more pieces of data together, which is one of the great parts of the Relational Model. Also, because there is no easy way to relate two pieces of data together, duplication of data becomes an issue once again. While an XML file is easy to write quickly for some small, unorganized data, the Relational Model is definitely a superior method for data storage and management.