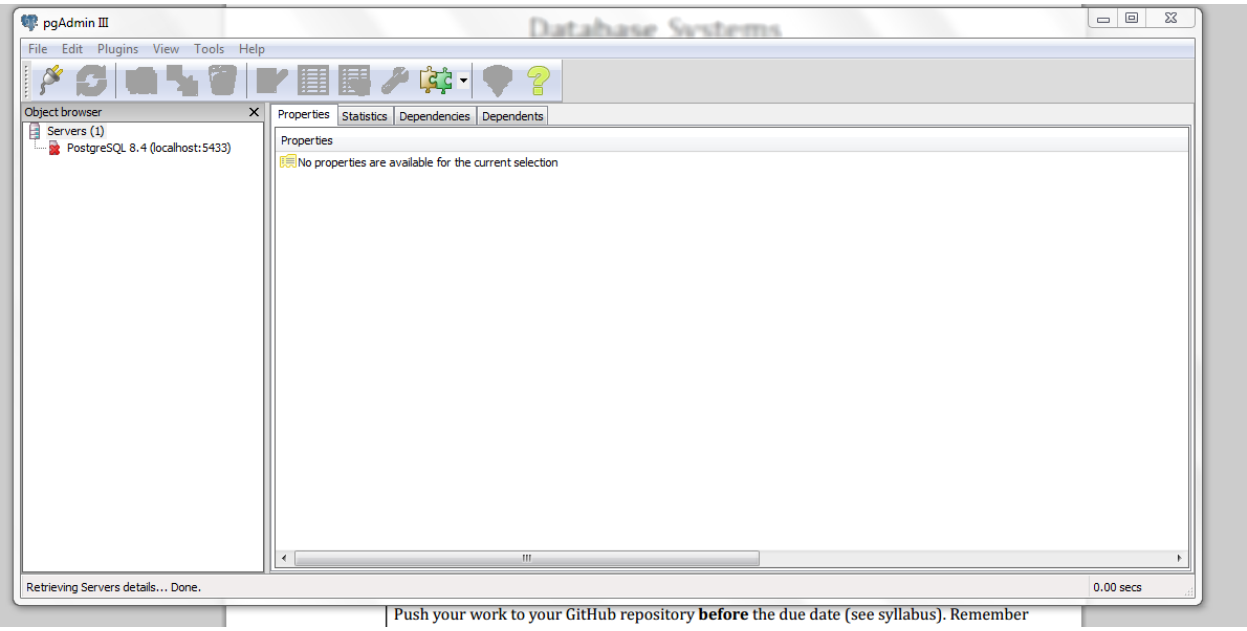
Ryan Fredericks

Data Management

Assignment 1



Short essay #1: Data vs. Information - Select a database in use today (real or imagined) and identify the elements of "data" stored therein and describe how the database organizes the "data" into "information". Give contrasting examples of "data" and "information" that illustrate the meaninglessness of "data" without context and organization. Talk about the value the "information" provides once the component data is given context.

Data is any value that is put into a database, though when meaning behind that data is given within the database, it becomes information. In a database, data fills the rows of each of the tables within the database, however the database finds a definition of what that data means by giving it identifiers in the form of the columns within the table, which can be used to create patterns and be used in a useful way. This distinction is important because without the definition and identification of data, it is essentially meaningless. When it is not given a unique identifier, the value that the data represents could mean anything, and without a way to define that data it cannot be compared and contrasted to anything else. This also means that the data, when organized in a way that it can be used to find a pattern or be used to impact other values, becomes information when it is used to do this. For example, when there a set of values that is defined as the mileage of a car would be data, but when it is ordered and it is decided that the mileage of the car correlates with the year it was created is information, which can be useful to whoever is using it. Another example would be if the critic reviews of a movie were stored as a number, which would be data. This data can become information if they compare a director's name in the database with this data, and find a correlation behind a good movie and the director behind it. The value that "information" has is that it gives meaning to any data that you may have accumulated, which

allows a designer to use this data in as many ways as possible, and makes the data worth something to whoever is using the database.

Short Essay #2: Data Models - Briefly describe the hierarchical and network prerelational data models. Explain their shortcomings in relation to the relational model. Considering this, what do you think of XML as a model for data storage?

a. The hierarchical database model has a tree-like structure and is based on a series of records which are connected by links. The limitation with this is that every record can only have one parent, although parent records can have multiple children. This pattern goes all the way up to the root node, which has no parent and has the children which create the model. This data model has limitations because it does not allow multiple parents to a child node, and it requires information to be stored in duplicates in multiple nodes. It also cannot handle many to many relationships. It can have slow queries because it requires the system to look top to bottom as well.
b. The network model, which while similar to the hierarchical database model has differences which makes it unique to that model. Network models allow children nodes to have multiple parents, although the model itself is similarly tree-like. There is an issue with this data model because it does not have structural independence. In addition to this, it is a very complex model and is very difficult to change.
c. I believe that XML as a network model has several advantages, the most importance of which is that is extremely flexible in how it can store data, and it is very strong when it is dealing with a schema that changes very often, which is important because it can changes based on the users' needs. One fault that I think will make the relational data model used for a significant amount of time is that it has a higher performance, and this is why I think that while the XML data model definitely can fit for some systems, the relational model will continue to be used, especially in large environments.