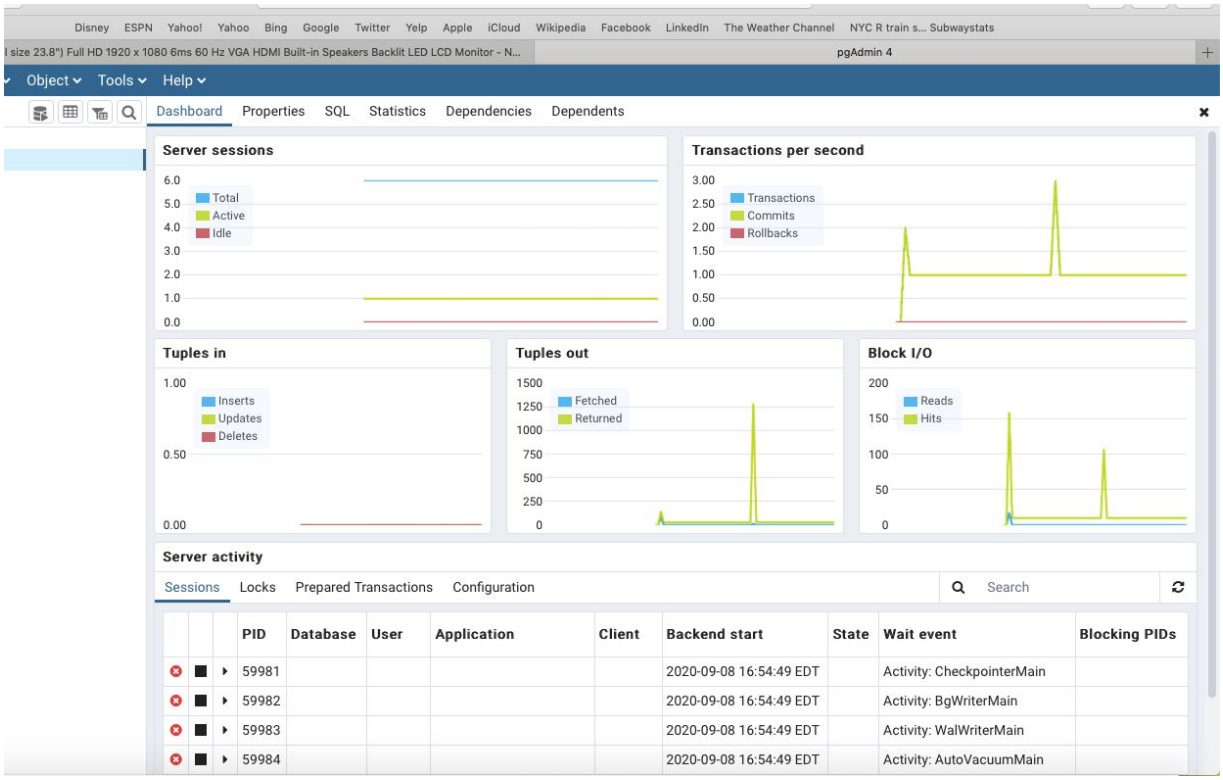
Madeline Atwood

Lab 1

Database Management

Professor Labouseur

August 31, 2020



Essay #1:

Data vs. Information

Data would be meaningless without context. One huge database that controls a lot of my personal everyday life is social media databases. There is lots of data that help me use instagram in a smooth manner. To create the algorithms in my feed there must be data to support what I am seeing. There may be a collection of my contacts, my followers, who I follow, previous

messages, comments, blocked users, previous searches etc. All of that information in their database connects together to help me see what I want where I want it. If the data of who I followed did not have context to it, and just listed their User IDs somewhere, then they may end up showing up under my recently searched and then everything I search could end up showing up in my daily feed. Whats worse is I could have everybody who blocked me or that I've blocked end up showing up as my friends if there is no context to that specific list of users. The data is the user ID that connects with what category it should go under (such as friend, follower, mutual, blocked).

Another reason that data is meaningless without context is that it could end up being a liability. If there is an extremely weak database and there is somehow a police investigation and the police need to access previous messages but there are unable to see who sent what message in a group chat/ direct message then they would be unable to find who is the victim in the situation. With the idea of "social media influencers" on the rise, it is important that their follower number is dictated properly to anybody who wants see it. When you have a certain number of followers there are certain access points that you can reach on the app that other people cannot, such as live streaming. When celebrities use social media, they become verified so the stakeholders of instagram have more control over their account to prevent hackings and improv security of instagram promotions they may bring. If somebody such as Donald Trump or Obama was not a verified user in instagram's database system, then that could be an enormous problem to many Americans and people who follow those accounts to receive information. Cross referencing data through the systems is the only way to keep what we see when using the app

accurate. It is most important to process the relationship between all the inputs of data to verify that everybody has a safe and productive experience when using social media.

Essay #2:

Data Models

In the past, there have been multiple different approaches developed for structuring access to information in a database. IBM's IMS (Information Management System) was one of the two earliest data models created for structuring information, the other one being Cullinet Software's IDMS. IMS created the hierarchical data model which is where different kinds of records relate to one another in hierarchical form where certain data that holds higher value (ex cooperate headquarters location) is at the top and then data that holds lower value (location of smaller branches of the company) would be placed at the bottom. The IDMS product was created as a result of the 1971 CODASYL report of an industry database task group and made the network data model. The network data model is similar to the hierarchical data model but the network data model is more of a generalization of the hierarchical model where a set of records in one layer may be on different hierarchies on the next layer.

The hierarchical model is structured as a tree with the root at the top and the branches at the bottom. The network model is structured as a directed graph without circuits. The main problem with both of these models of data was that they are difficult to pose and execute and often required an expert programmer who actually understood what could be a complex navigational structure of data. IMS and IDMS are still running today but are running on legacy systems which is harder to switch out of although it may be running perfectly fine. Most companies will now choose DBMS over either of these two now.

The most commonly used model for database systems is the relational model because it is easier for non-expert programmers to use. Even more recently, another newer data model called the object-relational model is been replacing the relational model which is an extension of the relational model.

An XML document kind of looks like an upside down tree to me that spreads it's leafs into little nets. I think XML makes it easier to specify data due to the organizational properties the tree creates, always allowing room for more branches but easily checkable to see what overall branch it is under. XML databases are a category of NoSQL databases.