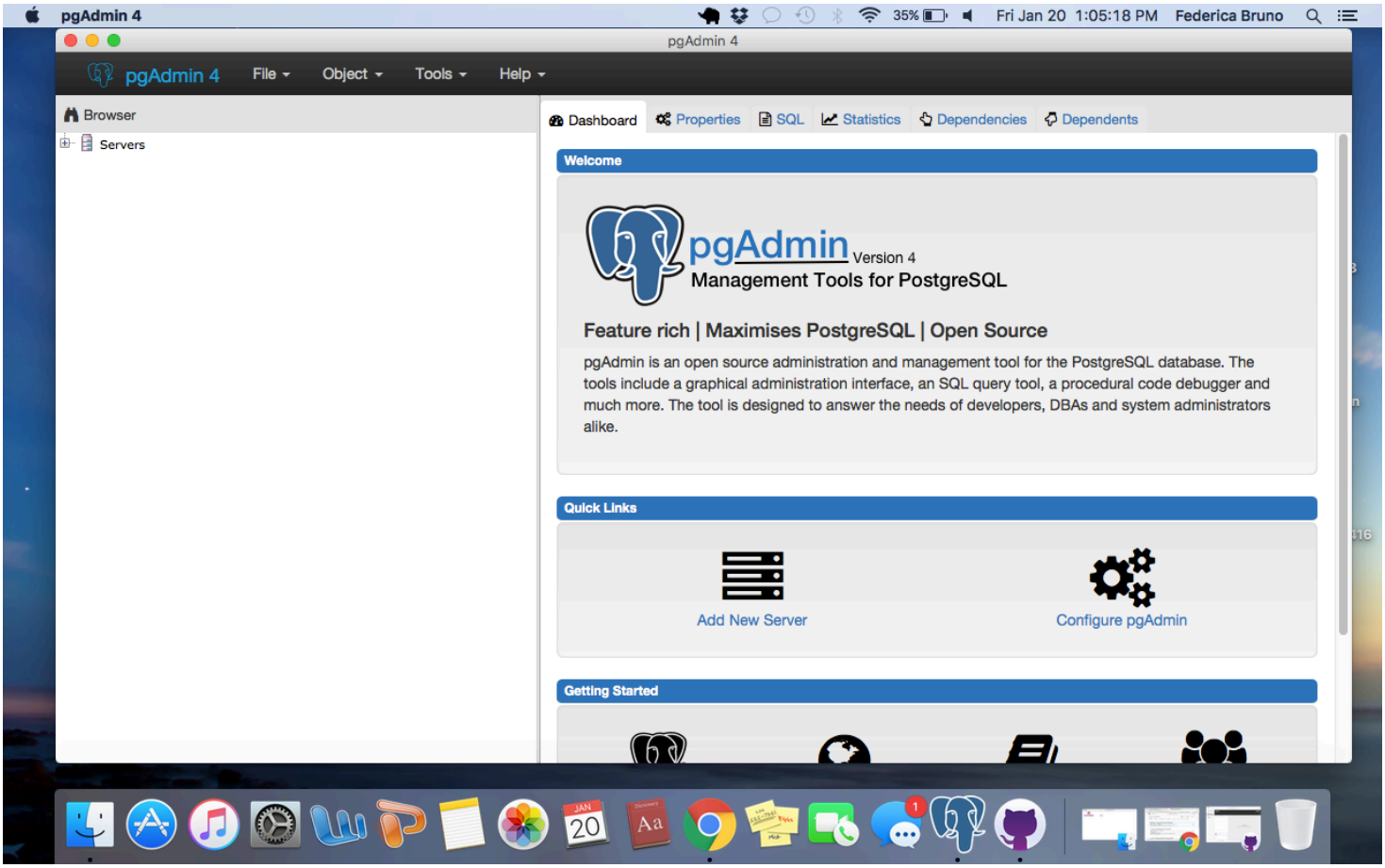
Federica Bruno
Database Management 308
Alan Labouseur
Due Date: January 24, 2017

Screenshot of pgAdmin tool



Short Answers:

1. Netflix, an online video-streaming site, is an example of a database that is popularly used today. I learned in System Design that the best user interface is one that is unnoticed; in other words, the interface is the system to the user and the user should be and usually is unaware of the backend system such as the database, so that the site naturally seems easy to use and easy to learn. Until Database Management, I never realized how much data Netflix manipulates to create its popular user interface and only then realized how I was guilty of the "unaware user".

Netflix is able to conclude user preferences to display recommended viewing based on the user's previously viewed material through its database. Netflix uses both a key-value and column orientated database management system called *Apache Cassandra*, which is a free and open-source DMS designed to handle large amount of incoming data. An example of how Netflix achieves personalization and recommendation using a database is capturing raw data and processing it with context to create information. Data is raw and unorganized numbers or words. Context is bringing meaning to the raw data. Data and context combined produce an organized interpretation of facts known as information. For instance, Netflix wants to know whether to cancel a show or not. The database gathers 70% and 30%. Those are just numbers, until you apply context to them. 70% of users finished watching all three seasons of a show, and 30% only watched the first season. This allows Netflix to analyze the data generated and infer that it should continue airing the show.

Netflix analyzes different types of data such as what day of the week you watch videos, the time you watch videos, where you watch videos, when you pause and leave content, ratings, searches, etc. Once figures are applied with context and are able to become information, this becomes an essential component to their business. Concluding information about user preferences reduces member cancelation and attracts new members. Tracking user interaction not only helps them to better understand the user, but also avoids displaying unpopular content that is taking up space, improving their service.

2. The Hierarchical model is structured much like a tree. It presents a parents-child relationship where the parent is the root of the tree and it sub-branches into individual data elements like nodes. This structure only supports one-to-one or one-to-many relationships. Another shortcoming to the hierarchical model is that it duplicates data being stored and searching for data is slow because it requires the DBMS to run from top to bottom of the structure. The Networking model is similar to the Hierarchical model except it allows many-to-many relationships by linking to multiple parent records. A shortcoming to this structure is that it involves many link adjustments to make insertion, deletion, and updating changes.

The XML is good for storing small file sizes, but can become complicated when a project grows; a database at this point would probably work out more than just this simple file storage alternative. You can break down the large xml

file size to avoid this complication; however, this requires a lot of manual labor and effort and having a database also avoids this extra step. Unlike a database, storing an unlimited amount of xml files in the XML directory is unadvised because code must iterate through the whole collection of xml files, which slows down the whole system.

Links used:

https://blog.kissmetrics.com/how-netflix-uses-analytics/
http://unixspace.com/context/databases.html
https://readmystuff.wordpress.com/2009/07/24/a-pitfall-of-using-xml-for-data-storage/