Studio: Setup and Basic SQL

Introduction

In this studio you will:

- Set up an instance of MySQL
- Import some sample data
- Explore the general structure of a database
- Execute some simple queries
- Analyze query execution

You are encouraged to work in groups of up to 4 people. Please do not just let one person do all of the work while everyone else watches. It is important for everyone to follow these steps and participate in the studio.

Installation

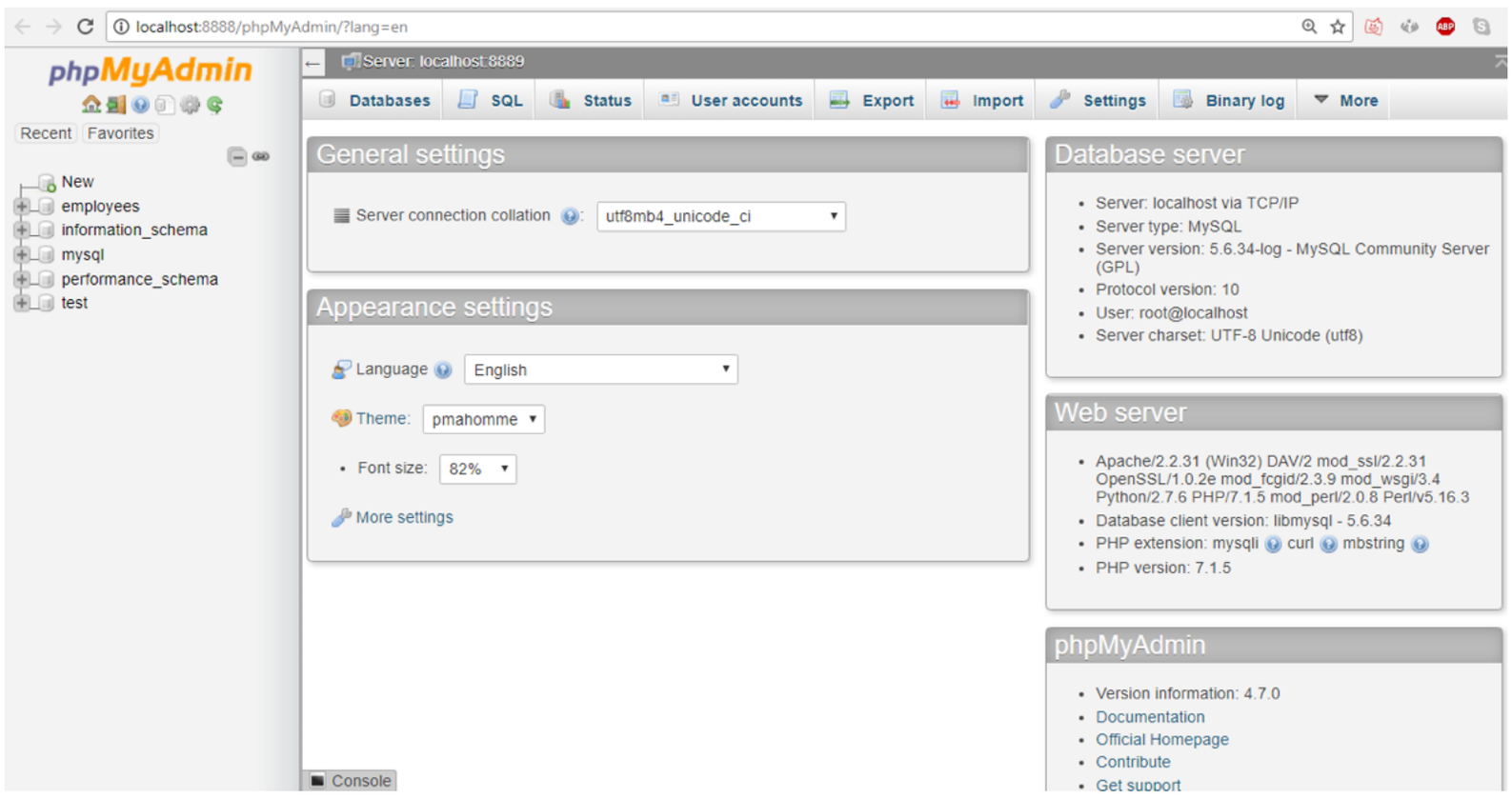
For Windows and Mac: The easiest way to install MySQL is to install a program called MAMP. MAMP includes some extra tools (like phpMyAdmin) that will be useful for us. Go to the MAMP website and follow the download and installation instructions for your platform. Make sure you do not select the MAMP Pro version during installation.

Once installation is complete, start the MAMP program. There should be a button marked "preferences." Open the preferences, select the "Ports" tab, then click the button that says "Set MAMP ports to default." You should then be able to start MAMP by clicking the "Start Servers" button on the main page. If you have trouble getting past this part, please ask for help!

For Linux: MAMP is not available for Linux. For now, all you need to install is two things: MySQL (often found as mysql-server) and phpMyAdmin. The instructions for doing this will vary based on the distribution of Linux that you are using. Make liberal use of the internet and please feel free to ask for help if you need it. You can also consider working with a partner who can install MAMP if you get stuck on this part.

Initial Investigations

Once everything is set up, open a browser and navigate to http:/localhost:8888/MAMP. You can also get there by clicking the "Open Start Page" button in the MAMP application. Next, select the "Tools" menu at the top of the screen and then click "phpMyAdmin." You should see a screen similar to the one below:



On the left side of this page, you will notice the names of a few databases (your list may look slightly different than the list shown in the picture). Your next task is to answer the following question:

What information is contained in these databases?

Use phpMyAdmin to explore these databases. Get a feel for how navigation works with this setup. Dig into each one of these databases and look at the information contained within. Is there anything that surprises you? Anything that you don't understand? Make some notes of your findings so that we can discuss them at the end of class.

Loading Some Data

The databases from the previous section are useful, but to explore how to run queries, we will need to load up some data of our own. MySQL provides a sample database that we can use to explore database functionality. The actual database information can be found here. Go to that page and grab a copy of the data, either by using git to create a clone of the repository or downloading the repository from GitHub (both options can be accomplished by clicking the "Clone or Download" button).

Once you have downloaded and unzipped the data, we need to load it onto the database server. The easiest way to do this is to use the command line. Open a terminal and navigate to the directory where you downloaded the database. Next, you will need to connect to the MySQL server. To do this you need to run the `mysql` command. The location of this command will be different depending on what kind of computer you are using:

Windows : `C:\MAMP\bin\mysql\bin\mysql.exe`

Mac : `/Applications/MAMP/Library/bin/mysql`

Linux: `mysql`

Please note that these locations may be slightly different depending on how you installed MySQL, but they should work on most machines. You also need to include the port number and login information (which should be the same default values for everyone, user=root, password=root). So for example, if you are on a Windows machine, you would run a command similar to the following:

```
C:\MAMP\bin\mysql\bin\mysql.exe --port=8889 --user=root --password=root
```

If you are successful, you should see a welcome mesage and a mysql prompt. At the prompt, run the following command:

```
source employees.sql
```

This command tells the server to run the employees.sql script (which itself is linked to several other scripts). You should see a series of success messages displayed on your screen, if everything worked properly.

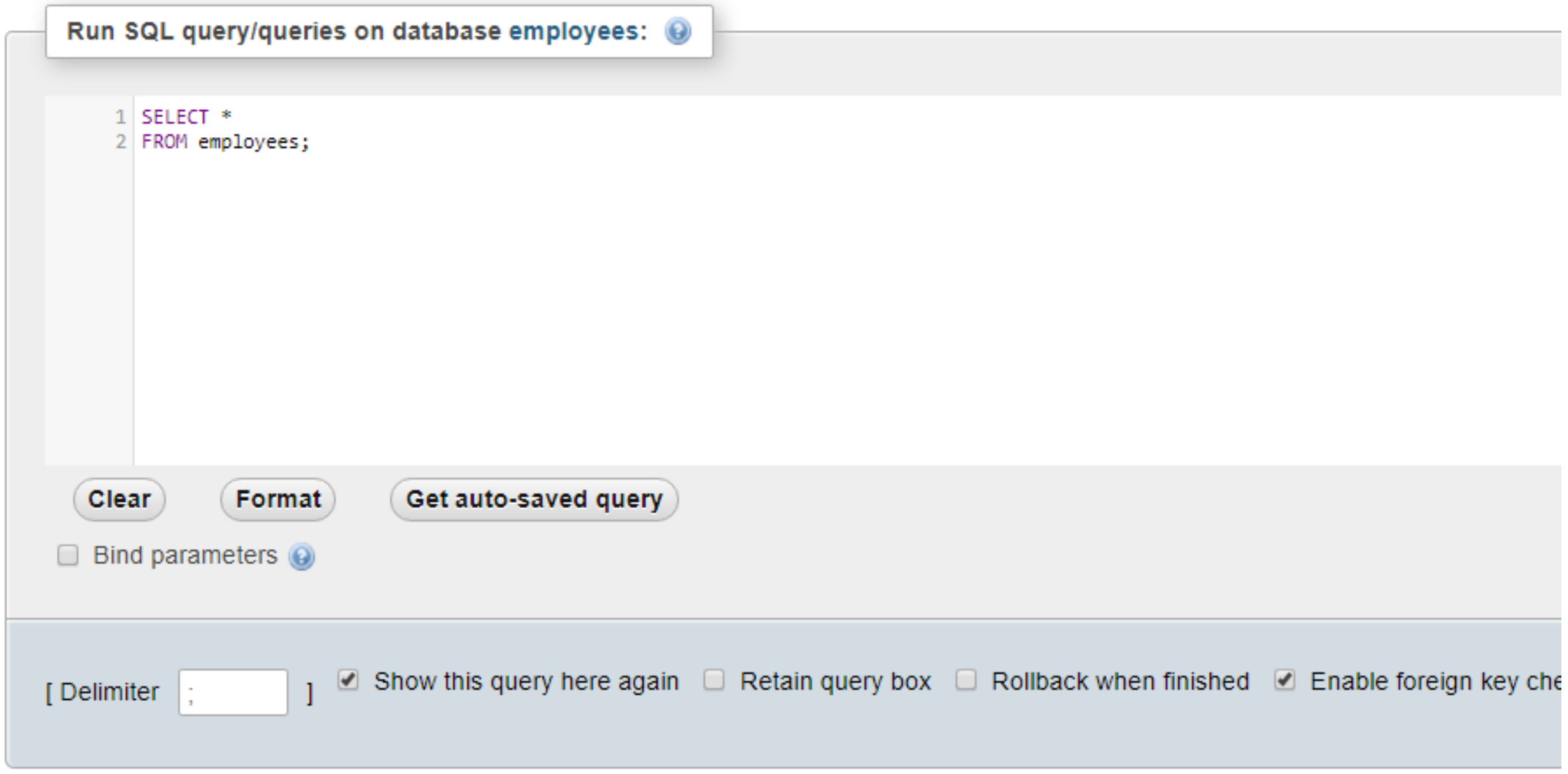
Exploring the Data

If you return to phpMyAdmin and refresh, you should now see an "employees" database on the left side of the screen. Take a few moments to explore the contents of this database. Try to become familiar with what each table contains, and how it relates to the other tables.

You can pull up an entity-relationship diagram (like the one we saw previously in class) by clicking the "More" tab at the top, then selecting "Design." It may be helpful to have this diagram up as you explore the rest of the studio.

Writing Queries

Next, put yourself in the shoes of someone who needs to use this database on a regular basis. Click the "SQL" tab at the top to open a SQL editor. Try running a simple query like the one seen below:



Running this query should return all columns and all rows from the employees table. Note the formatting here. Each clause (SELECT, FROM) is capitalized and on a separate line.

Your next task is to use SQL to answer the following questions. Please use the documentation and feel free to look things up as necessary. Each of these questions should be able to be answered by using some or all of the following clauses:

```
SELECT
FROM
WHERE
GROUP BY
ORDER BY
```

You may have to use some other modifiers in certain cases, but you should not need any clauses other than these. Also, you will not need to use any `JOIN`s to solve these problems.

Write SQL queries to answer the following questions:

What are the names and birthdays of all of the employees?

Generate a list of all employees that were hired after January 1, 1990.

Generate a list of all employees born in the year 1961. Hint: refer to the information on comparison operators.

How many employees are there? Hint: it may help to refer to the information on aggregates.

What is the highest salary? The lowest salary?

What is the current average salary. Be careful! Look at the salary table and be sure you understand its structure.

Which department has the most employees? The least?

Generate a query showing how many male and female employees work for the company.

Come up with at least one meaningful question of your own to answer.

Once you are finished, review the queries that you wrote and try to answer the following question:

How are queries processed by the server? Is there a particular order that clauses are executed in? If so, what is that order and why is it set up that way?