R exerices
SQL

Gilles Tredan

Abstract

SQL command line interaction

To get familiar with SQL we will first use the command line tool `mysql` to interact with the mysql database. To access the database, type `mysql -h mysql-drrp -u edsys -p edsys`. The password to access the database is `ed2015sys`.

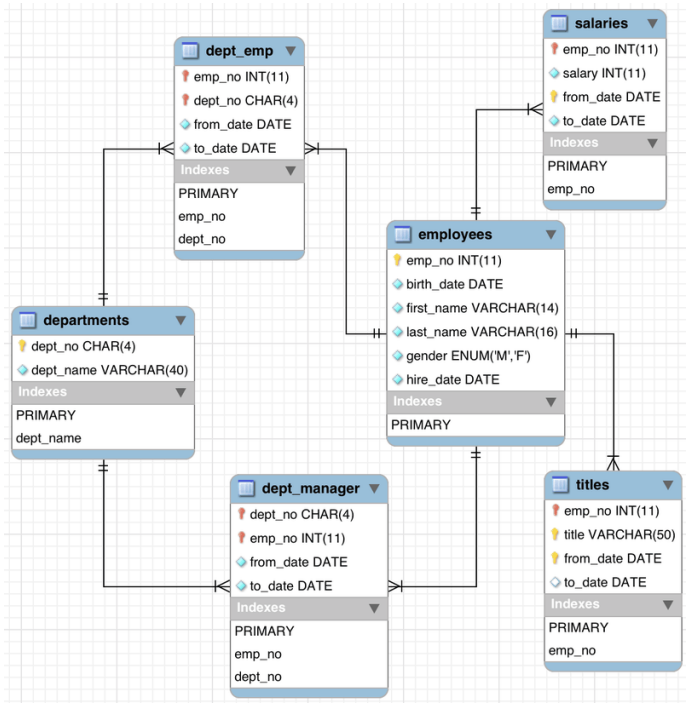

Figure 1: Sample Table Diagram

A diagram presenting the mysql structure is presented Figure 1. This table contains. The database contains about 300,000 employee records with 2.8 million salary entries.

- Connect to the database.
- List the tables and describe employees and salaries tables.
- Check for employees with your name

- Use `count` to count the number of male and female employees.
- When was the last employee hired ? When was the first employee hired ?
- How may employees share your birthday ? (tip= use `day(date)` (`month(date)`) to extract the day (month) of a date.
- How many employees have been employed between christmans and new eve ?
- Use the `group by` to extract the distribution of birthdays by day of the month.
- Use the `group by` to extract the distribution of hire dates by day and month, and order the result by month.
- what is the highest salary ? (`max`)
-
- Using the employee number `emp_no` of the person with the highest salary, what is the sum this person got as salary from the firm ?
- what is the name of the person with the highest salary
- using an inner join on the employee numbers, can you get the titles of the person with the highest salary in on request ?
- **bonus:** The title,name and salary of this person ?

SQL interaction within R

The package `RmySQL` allows you to flawlessly access mysql databases.

- Connect to the database within R and obtain the salary distribution
- Plot it.
- Let's say you want to study gender equality in this compagny. Prepare a set of plots presenting the phenomena.