SQL Practice exercises

Goal: Put into practice what you have learned so far in the SQL lectures (basic queries and more advanced queries)

There are 15 questions to solve with queries that are roughly of increasing difficulty:

* Level 1: Stretching
* Level 2: Pouring some spices
* Level 3: Why so serious?

**The document to submit should be either a doc file or a pdf file with the answers to each question:**

1. **The query you ran**
2. **The result you got**

# Setup instructions

1. [Download the repository](https://github.com/datacharmer/test_db/archive/refs/heads/master.zip). Unzip the file.
2. Launch your terminal/command shell and go to the repository where you downloaded the file. Enter the folder (that you have unzipped).
3. Run the following command: *mysql -u root -p < employees.sql*

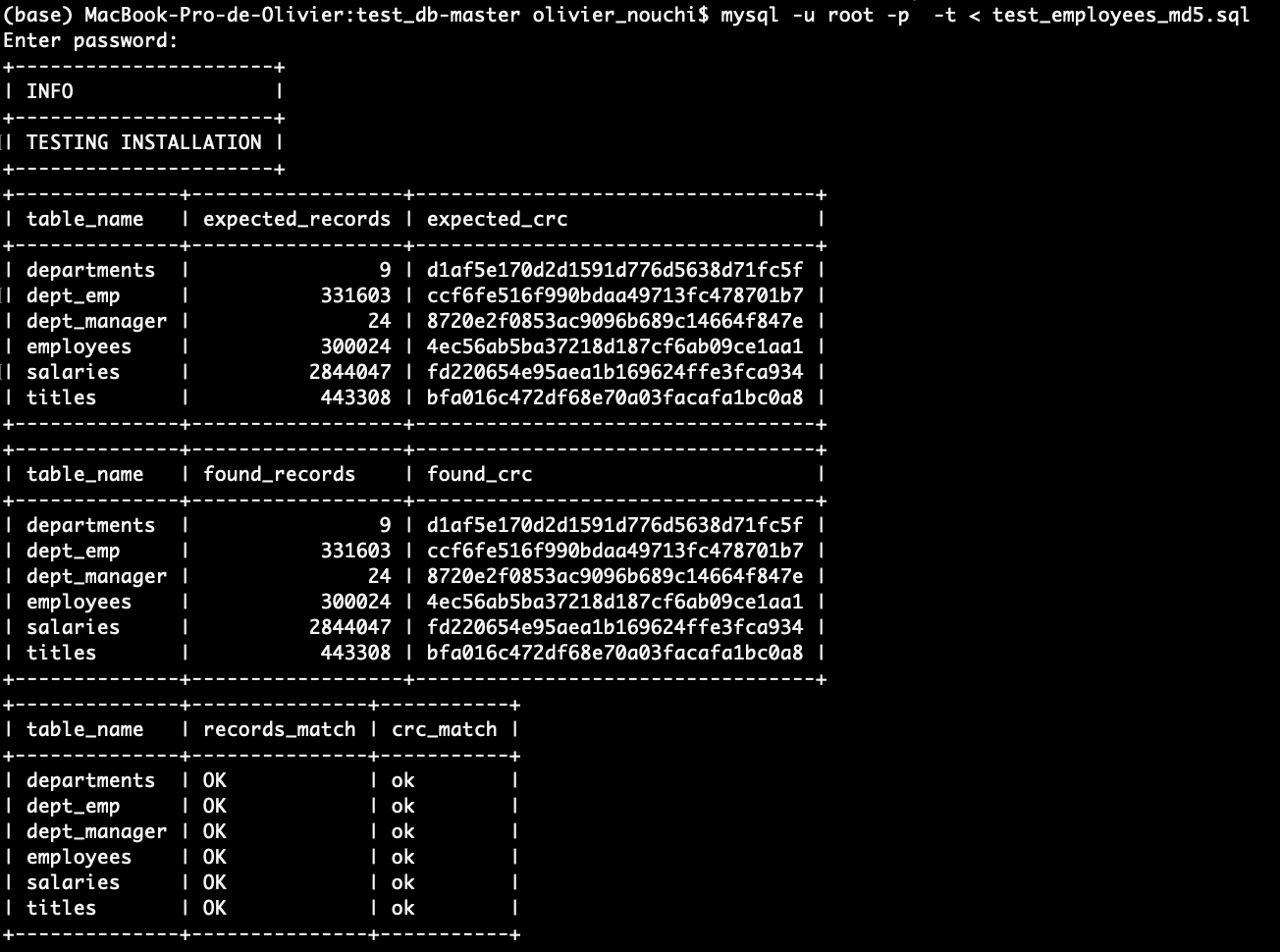
You should see the tables loading:



You have created the employees’ table with the tables cited above.

1. Then run the following command line: *mysql -u root -p -t < test\_employees\_md5.sql*

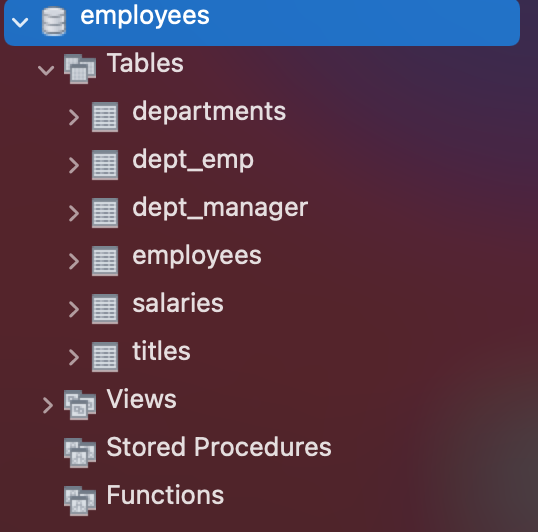
It will make sure that you have loaded everything correctly. You should be able to see the following (records matched).



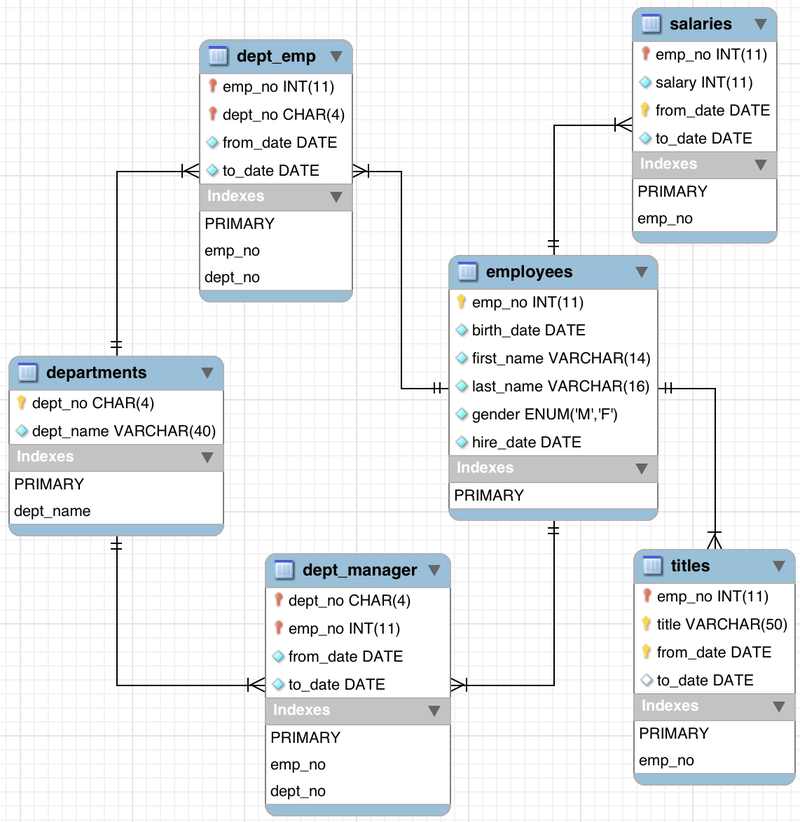
You are all set!

You can now work in MySQL Workbench.

Don’t forget to “Refresh All” if you don’t see it appear in MySQL Workbench.



# Database schema



# 

# 

# Stretching

Q1 - How many unique employees are there?

Q2 - How many males and females employees are there? Order the counts in descending order.

Q3 - Display the year and total hires for the year with the most hires

Q4 - What is the name of the department with the most employees

Q5 - How many employees were born on November 12? What's the percentage out of all the employees?

Q6 - What are the 3 most common employee titles (display the employee titles and the number of times they occur)

Q7 - Find the avg salary for each department (department name). Round to the nearest integer and order by avg salary from the highest to the lowest.

Q1: select distinct count(emp\_no) from employees.employees

**300024**

Q2: select gender , count(gender) as quantity\_per\_gender from employees.employees group by gender order by gender desc

|  |  |
| --- | --- |
| **F** | **120051** |
| **M** | **179973** |

Q3: SELECT DATE\_FORMAT(hire\_date, "%Y") as year\_hired, count(hire\_date) as total\_per\_year from employees.employees

group by year\_hired

order by count(hire\_date) desc

|  |  |
| --- | --- |
| **1986** | **36150** |

Q4: SELECT employees.dept\_emp.dept\_no, count(emp\_no) from employees.dept\_emp group by dept\_no order by emp\_no asc

|  |  |
| --- | --- |
| d005 | 85707 |

development

Q5: SELECT DATE\_FORMAT(birth\_date, "%M %d") as born\_in\_this\_date, count(birth\_date) from employees.employees group by born\_in\_this\_date

having born\_in\_this\_date = "November 12"

|  |  |
| --- | --- |
| **November 12**  Select count(emp\_no) as count employees,  (count(emp\_no)/(select count(emp\_no) from employees))\*100 as perc  From employees  Where month(birth\_date)=”11” and day(birth\_date)=”12”  **0.2666** | **800** |
| **800** |  |

Q6: SELECT title, count(title) from employees.titles group by title order by count(title) desc limit 3

|  |  |
| --- | --- |
| **Engineer** | **115003** |
| **Staff** | **107391** |
| **Senior Engineer** | **97750** |

Q7: Select ds.dept\_name as dept\_name, round(avg(s.salary)) as avg\_salary

From salaries as s

join employees as e

on e.emp\_no = s.emp\_no

Join dept\_emp as de

On de.emp\_no = e.emp\_no

Join departments as ds

On ds.dept\_no = de.dept\_no

Group by ds.dept\_name

Order by avg\_salary desc

|  |  |
| --- | --- |
| **Sales** | **80668** |
| **Marketing** | **71913** |
| **Finance** | **70489** |
| **Research** | **59665** |
| **Production** | **59605** |
| **Development** | **59479** |
| **Customer Service** | **58770** |
| **Quality Management** | **57251** |
| **Human Resources** | **55575** |

# Pouring some spices

Q8 - Find the average salary by employee title. Round to 2 decimals and order by descending order

Select t.title, round(avg(s.salary),2) as avg\_title

From salaries as s

Join titles as t

On s.emp\_no=t.emp\_no

Group by t.title

Order by avg\_title desc

|  |  |
| --- | --- |
| **Senior Staff** | **70470.84** |
| **Staff** | **69309.10** |
| **Manager** | **66924.27** |
| **Senior Engineer** | **60543.22** |
| **Engineer** | **59508.04** |
| **Assistant Engineer** | **59304.99** |
| **Technique Leader** | **59294.37** |

Q9 - Find the number of employees who have worked in at least 2 departments

Select count(emp\_no)

From (

Select emp\_no

From dept\_emp

Group by emp\_no

Having count(dept\_no)>=2

) as t

**31579**

Q10 - Get the distribution of the year of the hire dates. (hint: you should end up with a number of employees per year of hiring date)

Do you notice any pattern? Assuming there is no missing data, is the company hiring more or less as time goes by?

Select year(hire\_date) year, count(emp\_no) as count\_employees

From employees

Group by year

Order by year asc

|  |  |
| --- | --- |
| **1985** | **35316** |
| **1986** | **36150** |
| **1987** | **33501** |
| **1988** | **31436** |
| **1989** | **28394** |
| **1990** | **25610** |
| **1991** | **22568** |
| **1992** | **20402** |
| **1993** | **17772** |
| **1994** | **14835** |
| **1995** | **12115** |
| **1996** | **9574** |
| **1997** | **6669** |
| **1998** | **4155** |
| **1999** | **1514** |
| **2000** | **13** |

Q11 - Display the first name, last name, and salary of the highest paid employee

Select e.first\_name, e.last\_name, s.salary

From employees as e

Join salaries as s

On s.emp\_no = e.emp\_no

Order by s.salary desc

Limit 1

|  |  |  |
| --- | --- | --- |
| **Tokuyasu** | **Pesch** | **158220** |

Q12 - Display the first name, last name, and salary of the THIRD highest paid employee

Select e.first\_name, e.last\_name, s.salary

From employees as e

Join salaries as s

On s.emp\_no = e.emp\_no

Order by s.salary desc

Limit 3, 1

|  |  |  |
| --- | --- | --- |
| **Xiahua** | **Whitcomb** | **155709** |

# Why so serious?

Q13 - Display each department name and the age of the youngest employee at hire date

SELECT ds.dept\_name,

floor(datediff(current\_date(),max(e.hire\_date)) / 365.25) as age\_hired

FROM employees as e

JOIN dept\_emp as de

ON de.emp\_no = e.emp\_no

JOIN departments as ds

ON ds.dept\_no = de.dept\_no

GROUP BY dept\_name

|  |  |
| --- | --- |
| **21** | **Customer Service** |
| **21** | **Development** |
| **21** | **Finance** |
| **21** | **Human Resources** |
| **21** | **Marketing** |
| **21** | **Production** |
| **21** | **Quality Management** |
| **21** | **Research** |
| **21** | **Sales** |

Q14 - What's the range of age the employees would be today (calculate their age whole years)

SELECT min(age), max(age) from

(select (year(current\_date()) - year(birth\_date)) as age

from employees) as e

|  |  |
| --- | --- |
| 56 | 69 |

Q15 - How many employees were born on the same date (day-month-year) in 1955.

NB: Change the according to preferences parameters if your query fails because of a timeout

select sum(quantity) from

(SELECT count(birth\_date) as quantity

from employees

where year(birth\_date) = '1955'

group by birth\_date) as b

23104