CS440 Spring 2024 - Homework 3 (SQL part)

Due Date: 2/25/2024 at 11:59 PM

Database import instructions

For this assignment, use employee.db SQLite database included with this assignment. For nice layout in SQLite shell, run the commands

.mode column

.headers on

The SQLite3 will automatically truncate your output to 10 characters. This auto-truncating won't affect your score.

Database Schema

```
CREATE TABLE employee (
    employee_id INT,
    first_name VARCHAR(50),
    last_name VARCHAR(50),
    age INT,
    gender VARCHAR(50),
    company_name VARCHAR(50),
    department VARCHAR(50),
    country VARCHAR(50),
    salary INT,
    extra JSON
);
CREATE TABLE company_info(
    country VARCHAR(50),
    security_clearance INT
);
```

Column Description

The employee_id is a system-generated identification number that is used to uniquely identify each employee.

first_name and last_name refer to a user's actual name.

age refers to a user's self-identified age.

gender refers to a user's self-identified gender.

department refers to the employee's department.

country refers to the employee's current place of residence.

salary represents the current amount of money that each employee is paid ($).

extra contains information about the employee and it is JSON format

security_clearance is the security clearance level of current country

1. (5 points) Please order employees in Germany by salary in descending order. For each employee, list his/her first name, last name, salary and the salary difference with last one's salary.

   Expected Result:

first_name	last_name	salary	diff
Ronald	Wharby	258606	
Shara	Pigeon	249557	9049
Georgie	Kubek	247364	2193
Windham	Mulcock	220334	27030
Aldric	Faires	216381	3953
Shurwood	Whiffen	198983	17398
Marcille	Saggs	184837	14146
Genia	Cluely	140918	43919
Lelah	Jeschner	136254	4664
Vinson	Danilewicz	116862	19392
Maggi	Strasse	113888	2974
Leo	Pimblotte	88728	25160

2. (5 points) For employees in "United Kingdom" and "Germany", order employees by their salary within their own country in descending order. For each employee, list his/her first name, last name, salary, country, rank in his country and which quartile his salary fails into in his country.

   Expected Result:

first_name	last_name	salary	country	rank	quartile
Ronald	Wharby	258606	Germany	1	1
Shara	Pigeon	249557	Germany	2	1
Georgie	Kubek	247364	Germany	3	1
Windham	Mulcock	220334	Germany	4	2
Aldric	Faires	216381	Germany	5	2
Shurwood	Whiffen	198983	Germany	6	2
Marcille	Saggs	184837	Germany	7	3
Genia	Cluely	140918	Germany	8	3
Lelah	Jeschner	136254	Germany	9	3
Vinson	Danilewicz	116862	Germany	10	4
Maggi	Strasse	113888	Germany	11	4
Leo	Pimblotte	88728	Germany	12	4
Papagena	MacCosto	143930	United Kingdom	1	1
Fonz	Colvine	120224	United Kingdom	2	1
Rutger	Klimushev	104536	United Kingdom	3	2
Karel	Huskisson	103318	United Kingdom	4	3
Joell	Matt	73975	United Kingdom	5	4

3. (5 points) For each gender, list a running count of the gender employees categorized by age. Include only the ages 24 to 30 in your result. The count at age X includes all employees whose age ¡= X. For example, the expected result indicates that there are 7 female employees aged 24 or younger, and 14 female employees aged 25 or younger. It is assumed that all employees are aged 24 or above, and for each age between 24 and 30, there is at least one female and one male employee.

   Expected Result:

age	gender	employee_num
24	Female	7
25	Female	14
26	Female	26
27	Female	37
28	Female	50
29	Female	59
30	Female	79
24	Male	13
25	Male	26
26	Male	37
27	Male	45
28	Male	52
29	Male	60
30	Male	73

4. (5 points) Calculate the salary rank among employees of the same gender (marked as 'rank1') and the salary rank among both female and male employees (marked as 'rank2') for employees aged 24. For example, Evie's salary ranks first among all female employees aged 24 and ranks second among both female and male employees.

   Expected Result:

first_name	last_name	age	gender	salary	rank1	rank2
Evie	McNamara	24	Female	279021	1	2
Dione	Carlow	24	Female	275333	2	4
Gustie	Feria	24	Female	247568	3	6
Terese	Heeley	24	Female	145341	4	13
Leila	Morena	24	Female	139996	5	14
Lusa	Shermore	24	Female	138795	6	15
Laetitia	Rate	24	Female	109811	7	18
Tracie	Gutherson	24	Male	284044	1	1
Herc	Hearon	24	Male	277560	2	3
Darnall	Trevallion	24	Male	266780	3	5
Alick	Scrowson	24	Male	219293	4	7
Ruy	Fritschmann	24	Male	206802	5	8
Hughie	Flemmich	24	Male	158433	6	9
Thedrick	Covely	24	Male	154725	7	10
Uriel	Mordin	24	Male	153293	8	11
Biron	Tubbles	24	Male	152936	9	12
Hubey	Chevins	24	Male	134068	10	16
Farr	Godthaab	24	Male	116553	11	17
Niel	Blamphin	24	Male	109003	12	19
Skye	Kinman	24	Male	71591	13	20

5. (5 points) Retrieve the department names and the average salary for employees in each department, categorizing them into "High," "Medium," or "Low" salary brackets based on the average salary. Departments with an average salary of $200,000 or more are considered "High" salary, those with an average salary greater than or equal $180,000 and less than $200,000 are considered "Medium" salary, and those with an average salary below $180,000 are considered "Low" salary.

   Expected Result:

department	average_salary	salary_bracket
Accounting	181819.397727273	Medium
Business Development	190324.164705882	Medium
Engineering	183213.817073171	Medium
Human Resources	201889.197183099	High
Legal	179371.084337349	Low
Marketing	183964.863636364	Medium
Product Management	186738.936842105	Medium
Research and Development	167820.6125	Low
Sales	189182.320987654	Medium
Services	186227.038961039	Medium
Support	178205.153846154	Low
Training	182908.032608696	Medium

6. (5 points) The security clearance of each country ranges from 1 to 5, which correspond to low clearance to top secret clearance. Retrieve the number of countries in each security clearance level.

   Expected Result:

low_clearance	moderate_clearance	high_clearance	very_high_clearance	top_secret_clearance
2	1	3	2	1

7. (5 points) Categorize employees based on their age groups and gender and retrieve the number of employees of each category. Employees with age smaller than 30 are considered "Young", those with age no less than 30 and smaller than 50 are considered "Middle-aged", and those with age no less than 50 are considered "Senior".

   Expected Result:

Age Group	Gender	Num Employees
Middle-aged	Female	249
Middle-aged	Male	222
Senior	Female	196
Senior	Male	214
Young	Female	59
Young	Male	60

8. (5 points) Retrieve employees whose emails are under federal trade commission (i.e., @ftc.gov).

   Expected Result:

first_name	last_name	email
Alano	Screase	ascrease1@ftc.gov
Maryanna	Klejna	mklejna71@ftc.gov
Stefan	Moakes	smoakesay@ftc.gov

9. (5 points) Retrieve the average number of tags per employee.

   Expected Result:

avg_tags_per_employee
5.51

10. (5 points) Count the number of employees who have both a cell phone and home phone number listed.

    Expected Result:

num_employees
643