**-Pewlett-Hackard-Analysis.-**

Retirement plan

**Overview**

Pewlett-Hackard is a company with thousands of employees, some of wich are getting close to retirement. The company needs to know:

* How many people are eligible for a retiring package.
* Wich and how many positions will have an opening.

And, in order to start off a mentorship program, aimed to train the future generation:

* How many employees are eligible for a mentorship program.

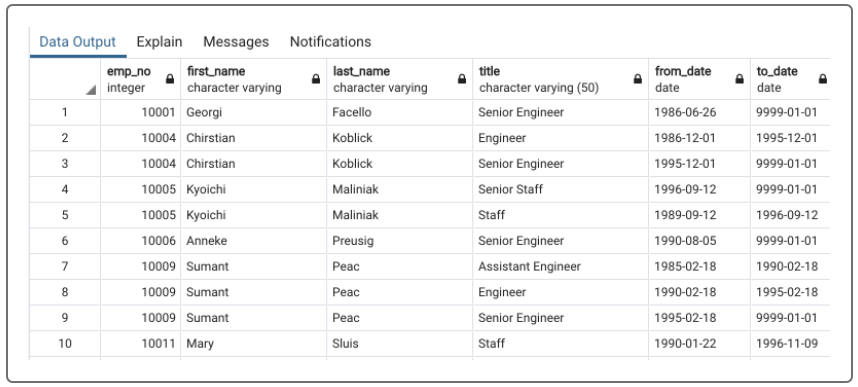
Based on six CSV files provided by the company, we'll create a set of lists that hold the employee's number, name and title (among other data) who meet certain criteria, wich would include them in one of the groups named above.

**Results**

**Deliverable 1**

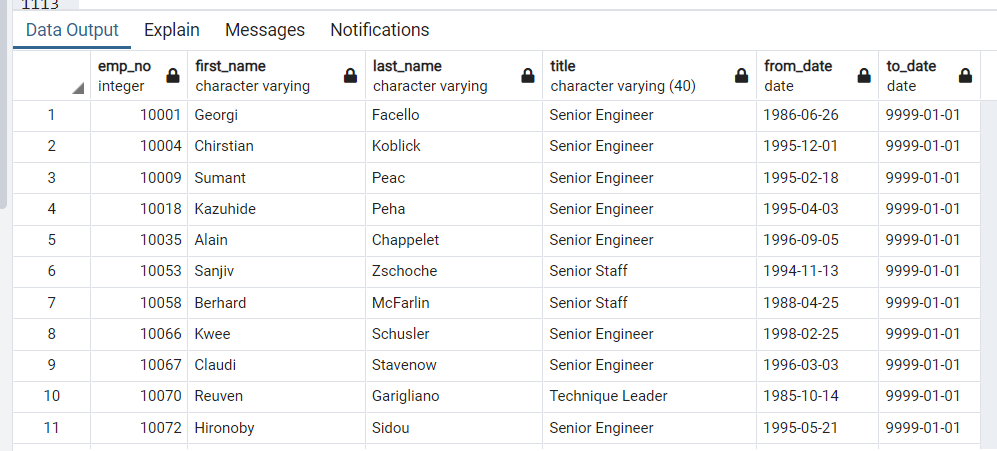
To get the number of retirees by title, we have two options.

* Option #1: We can follow the instructions in the challenge (retrieve data from employees table) to create the next table (retirement\_titles.csv). In this table there are duplicate entries for some employees because they have switched titles over the years.

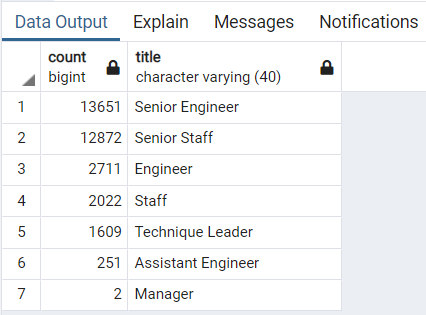
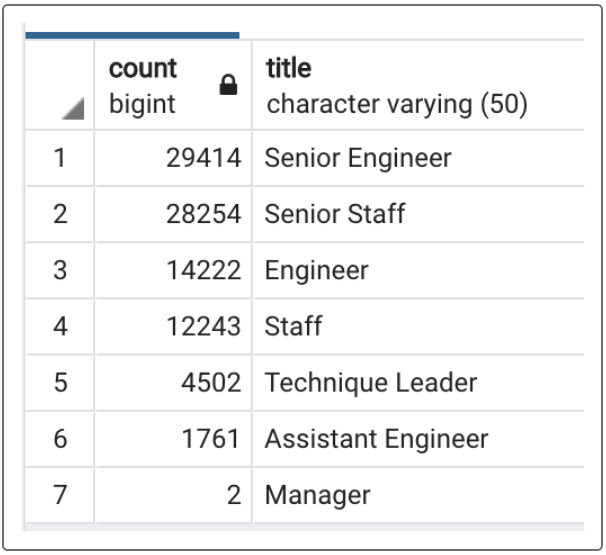
[](https://github.com/Francisco170594/-Pewlett-Hackard-Analysis.-/blob/main/Images/emp_retirement%20titles.png)

-> Because of the duplicate entries, we need to used the DISTINCT ON statement to retrieve the most recent position.

* Option #2: We can look for the same data only this time, we search in the current employees table instead, thus getting a "clean table" (most recent title). The DataFrame we obtain would be the same one as "unique\_titles.csv" wich is asked in deliverable 1 on the challenge.

[](https://github.com/Francisco170594/-Pewlett-Hackard-Analysis.-/blob/main/Images/ce_retirement%20titles.png)

Using the COUNT() function we will retrieve the number of employees by their most recent job title who are about to retire. The website shows us the next DataFrame as a reference of how our table should look like.

[](https://github.com/Francisco170594/-Pewlett-Hackard-Analysis.-/blob/main/Images/demo%20retiring%20titles.png)

However, the sum of these numbers (90,398) doesn't match the values obtained on the current employee table ( 33,118). A more accurate table would look like this.