**Pewlett-Hackard Analysis**

1. In your first paragraph, introduce the problem that you were using data to solve.

In this challenge, we were required to find the total number of employees who will be retiring according to their title. From the list we have created another table that included employees who were eligible to participate in this program. Then, in order to be part of the mentorship program, the employee had to be born between date of birth that falls between January 1, 1965 and December 31, 1965. The difficulties that involved creating these table are that some names were duplicated. Another problem is for the second table, we needed to count the number of employees.

1. In your second paragraph, summarize the steps that you took to solve the problem, as well as the challenges that you encountered along the way. This is an excellent spot to provide examples and descriptions of the code that you used.

In order to complete the challenge part one, I had to join tables together to find the total number of employees who will be retiring according to their title. In order to make this new table, called retirement\_ready, I joined current\_emp, dept\_manager, salaries and titles together. I used an inner join to join the four tables. After joining the tables, there were names that were duplicated because an employee can have more than one title throughout the years of being in the company. In order to fix this, I partitioned the same set of tables to create a new table called ret\_ready\_partition. After partitioning I used the COUNT method to find how many times a name appears. With this new information, I created a new table called retirement\_ready\_count which consists of the employee name, number, to\_date and the count of how many titles they have received.

After completing the retirement\_ready\_count table, I created the part 2 challenge table. For part 2 in the challenge, I had to find employees who will able eligible to participate in the mentorship program. The employees needed to have a date of birth that falls between January 1, 1965 and December 31, 1965. In order to create this table, I had to do two sets of merges. On the first merge (which became the mentorship table), I merged the employee and the dept\_emp tables suing an INNER merge. On this merge, I also included the WHERE (birth\_date BETWEEN '1965-01-01' AND '1965-12-31'), to indicate what their birthdate needed to fall between January 1, 1965 and December 31, 1965, and an AND (de.to\_date = '9999-01-01'); to indicate that they still needed to be working in the company. After creating this table, I then joined this table (mentorship table) with the salaries table and the titles table. I once again used INNER JOIN to join the three tables. The problem that occurred after making this table if that the names would once again there were names that were duplicated. To fix this problem, I used PARTITION the names to delete the duplicates. The problem that occurred after making this table is that there were names that were duplicated. To fix this problem, I used PARTITION the names to delete the duplicates which created the mentor\_partition table.

1. In your final paragraph, share the results of your analysis and discuss the data that you’ve generated. Have you identified any limitations to the analysis? What next steps would you recommend?

After creating these sets of tables, I have found that there are 81 people who are ready to retire and there 1,549 employees who are eligible to be mentors. The implications with the tables are that it does show which employees have agreed to be part of the mentorship program. Another implication is the birthdate. The birthdate for the mentorship table only looked at people who were born between January 1, 1965 to December 31, 1965. This means every year one has to make a new table of eligibility table every year. An additional implication about the date is that, the employee can be born between January 1, 1965 to December 31, 1965, but the table does not show how long they have been in the company. If it is their first year working in the company, would they still eligible to be part of the mentorship program?In order to fix the implications, I would suggest these following steps. I would first lengthen the employee birthday to five years so that there are more people eligible to become a mentor. I would also include the amount of years they have been working for the company. I would only include in the list employees that have been in the company for at least five years. Lastly, I would include the employee’s contact information, such as an email address and/or telephone number, so that we can ask each employee is they would like to be part of the mentorship program.

A screenshot of a cell phone

Description automatically generatedThe retirement table also needs readjustments. One major implication that the retirement table has, is that it does not include each employee’s age. If a person who is looking at the table, he/she will not know which employee is about to retire or considering retirement due to age. Therefore, for the retirement table, I would add the employee’s birth date and age. I would also add a column that states the amount of years working in the company. There are people who wants to stay in a company for a certain amount of years because of benefits.