Module 7 Challenge

Start Assignment

Due Sunday by 11:59pm **Points** 100 **Submitting** a text entry box or a website url

Background

Now that Bobby has proven his SQL chops, his manager has given both of you two more assignments: determine the number of retiring employees per title, and identify employees who are eligible to participate in a mentorship program. Then, you'll write a report that summarizes your analysis and helps prepare Bobby's manager for the "silver tsunami" as many current employees reach retirement age.

What You're Creating

This new assignment consists of two technical analysis deliverables and a written report. You will submit the following:

- Deliverable 1: The Number of Retiring Employees by Title
- Deliverable 2: The Employees Eligible for the Mentorship Program
- Deliverable 3: A written report on the employee database analysis (README.md)

Files

Use the following link to download the Challenge starter code.

Download challenge starter code (https://2u-data-curriculum-

team.s3.amazonaws.com/dataviz-online/module_7/Employee_Challenge_starter_code.sql)

Deliverable 1: The Number of Retiring Employees by Title (50 points)

Deliverable 1 Instructions

Using the ERD you created in this module as a reference and your knowledge of SQL queries, create a Retirement Titles table that holds all the titles of employees who were born between January 1, 1952 and December 31, 1955. Because some employees may have multiple titles in the database—for example, due to promotions—you'll need to use the `DISTINCT ON` statement to create a table that contains the most recent title of each employee. Then, use the `COUNT()` function to create a table that has the number of retirement-age employees by most recent job title. Finally, because we want to include only current employees in our analysis, be sure to exclude those employees who have already left the company.

REWIND

For this deliverable, you've already done the following in this module:

- **Lesson 7.3.1:** Create new tables with the `INTO` statement
- **Lesson 7.3.1:** Export a table as a CSV file
- **Lesson 7.3.1:** Filter queries with the `WHERE` clause
- **Lesson 7.3.3:** Use the `INNER JOIN` clause to join two tables on a primary key

- **Lesson 7.3.3:** Use the `ON ()` clause
- **Lesson 7.3.3:** Use an alias instead of a full table name
- **Lesson 7.3.4:** Use the `ORDER BY` clause
- **Lesson 7.3.4:** Use the `COUNT()` function to retrieve the total number of rows that matches a specified criteria

Create a SQL file in the Queries folder of your Pewlett-Hackard-Analysis GitHub folder, and name it `Employee_Database_challenge.sql`.

Follow the instructions below to complete Deliverable 1.

1. Retrieve the `emp_no`, `first_name`, and `last_name` columns from the Employees table.
2. Retrieve the `title`, `from_date`, and `to_date` columns from the Titles table.
3. Create a new table using the `INTO` clause.
4. Join both tables on the primary key.
5. Filter the data on the `birth_date` column to retrieve the employees who were born between 1952 and 1955. Then, order by the employee number.

6. Export the Retirement Titles table from the previous step as `retirement_titles.csv` and save it to your Data folder in the Pewlett-Hackard-Analysis folder.
7. Before you export your table, confirm that it looks like this image:

Data Output | Explain | Messages | Notifications

	emp_no integer	first_name character varying	last_name character varying	title character varying (50)	from_date date	to_date date
1	10001	Georgi	Facello	Senior Engineer	1986-06-26	9999-01-01
2	10004	Chirstian	Koblick	Engineer	1986-12-01	1995-12-01
3	10004	Chirstian	Koblick	Senior Engineer	1995-12-01	9999-01-01
4	10005	Kyoichi	Maliniak	Senior Staff	1996-09-12	9999-01-01
5	10005	Kyoichi	Maliniak	Staff	1989-09-12	1996-09-12
6	10006	Anneke	Preusig	Senior Engineer	1990-08-05	9999-01-01
7	10009	Sumant	Peac	Assistant Engineer	1985-02-18	1990-02-18
8	10009	Sumant	Peac	Engineer	1990-02-18	1995-02-18
9	10009	Sumant	Peac	Senior Engineer	1995-02-18	9999-01-01
10	10011	Mary	Sluis	Staff	1990-01-22	1996-11-09

Note: There are duplicate entries for some employees because they have switched titles over the years. Use the following instructions to remove these duplicates and keep only the most recent title of each employee.

8. Copy the query from the `Employee_Challenge_starter_code.sql` and add it to your `Employee_Database_challenge.sql file`.
9. Retrieve the employee number, first and last name, and title columns from the Retirement Titles table.
    - These columns will be in the new table that will hold the most recent title of each employee.
10. Use the `DISTINCT ON` statement to retrieve the first occurrence of the employee number for each set of rows defined by the `ON ()` clause.

If you'd like a hint on using the `DISTINCT ON` statement, that's totally okay. If not, that's great too. You can always revisit this later if you change your mind.

SHOW HINT

SHOW HINT

11. Exclude those employees that have already left the company by filtering on `to_date` to keep only those dates that are equal to `'9999-01-01'`.
12. Create a Unique Titles table using the `INTO` clause.
13. Sort the Unique Titles table in ascending order by the employee number and descending order by the last date (i.e., `to_date`) of the most recent title.
14. Export the Unique Titles table as `unique_titles.csv` and save it to your Data folder in the Pewlett-Hackard-Analysis folder.
15. Before you export your table, confirm that it looks like this image:

Data Output | Explain | Messages | Notifications

	emp_no integer	first_name character varying	last_name character varying	title character varying (50)
1	10001	Georgi	Facello	Senior Engineer
2	10004	Chirstian	Koblick	Senior Engineer
3	10005	Kyoichi	Maliniak	Senior Staff
4	10006	Anneke	Preusig	Senior Engineer

5	10009	Sumant	Peac	Senior Engineer
6	10018	Kazuhide	Peha	Senior Engineer
7	10019	Lillian	Haddadi	Staff
8	10020	Mayuko	Warwick	Engineer
9	10022	Shahaf	Famili	Engineer
10	10023	Bojan	Montemayor	Engineer

16. Write another query in the `Employee_Database_challenge.sql` file to retrieve the number of employees by their most recent job title who are about to retire.
17. First, retrieve the number of titles from the Unique Titles table.
18. Then, create a Retiring Titles table to hold the required information.
19. Group the table by title, then sort the count column in descending order.
20. Export the Retiring Titles table as `retiring_titles.csv` and save it to your Data folder in the Pewlett-Hackard-Analysis folder.
21. Before you export your table, confirm that it looks like this image:

	count bigint	title character varying (50)
1	25916	Senior Engineer
2	24926	Senior Staff

2	24926	Senior Staff
3	9285	Engineer
4	7636	Staff
5	3603	Technique Leader
6	1090	Assistant Engineer
7	2	Manager

22. Save your `Employee_Database_challenge.sql` file in your Queries folder in the Pewlett-Hackard folder.

Deliverable 1 Requirements

You will earn a perfect score for Deliverable 1 by completing all requirements below:

- A query is written and executed to create a Retirement Titles table for employees who are born between January 1, 1952 and December 31, 1955. **(10 pt)**
- The Retirement Titles table is exported as `retirement_titles.csv`. **(5 pt)**

- A query is written and executed to create a Unique Titles table that contains the employee number, first and last name, and most recent title. **(15 pt)**
- The Unique Titles table is exported as `unique_titles.csv`. **(5 pt)**
- A query is written and executed to create a Retiring Titles table that

contains the number of titles filled by employees who are retiring. **(10 pt)**

- The Retiring Titles table is exported as `retiring_titles.csv`. **(5 pt)**

Deliverable 2: The Employees Eligible for the Mentorship Program (30 points)

Deliverable 2 Instructions

Using the ERD you created in this module as a reference and your knowledge of SQL queries, create a mentorship-eligibility table that holds the current employees who were born between January 1, 1965 and December 31, 1965.

REWIND

For this deliverable, you've already done the following in this module:

- **Lesson 7.3.1:** Create new tables with the `INTO` statement
- **Lesson 7.3.1:** Export a table as a CSV file
- **Lesson 7.3.1:** Filter queries with the `WHERE` clause
- **Lesson 7.3.3:** Use the `INNER JOIN` clause to join two tables on a similar column
- **Lesson 7.3.3:** Use the `ON ()` clause
- **Lesson 7.3.3:** Use an alias instead of a full table name
- **Lesson 7.3.4:** Use the `ORDER BY` clause

In the `Employee_Database_challenge.sql file`, write a query to create a Mentorship Eligibility table that holds the employees who are eligible to participate in a mentorship program.

1. Retrieve the `emp_no`, `first_name`, `last_name`, and `birth_date` columns from the Employees table.
2. Retrieve the `from_date` and `to_date` columns from the Department Employee table.
3. Retrieve the `title` column from the Titles table.
4. Use a `DISTINCT ON` statement to retrieve the first occurrence of the employee number for each set of rows defined by the `ON ()` clause.
5. Create a new table using the `INTO` clause.
6. Join the Employees and the Department Employee tables on the primary key.
7. Join the Employees and the Titles tables on the primary key.
8. Filter the data on the `to_date` column to all the current employees, then filter the data on the `birth_date` columns to get all the employees whose birth dates are between January 1, 1965 and December 31, 1965.
9. Order the table by the employee number.
10. Export the Mentorship Eligibility table as `mentorship_eligibilty.csv` and save it to your Data folder in the Pewlett-Hackard-Analysis folder.
11. Before you export your table, confirm that it looks like this image:

	emp_no integer	first_name character varying	last_name character varying	birth_date date	from_date date	to_date date	title character varying (50)
1	10095	Hilari	Morton	1965-01-03	1994-03-10	9999-01-01	Senior Staff
2	10122	Ohad	Esposito	1965-01-19	1998-08-06	9999-01-01	Technique Leader
3	10291	Dipayan	Seghrouchni	1965-01-23	1987-03-30	9999-01-01	Senior Staff
4	10476	Kokou	Iisaka	1965-01-01	1987-09-20	9999-01-01	Senior Staff
5	10663	Teunis	Noriega	1965-01-09	1999-02-12	9999-01-01	Technique Leader
6	10762	Lech	Himler	1965-01-19	1992-01-21	9999-01-01	Staff
7	10933	Juyoung	Seghrouchni	1965-01-24	1993-08-02	9999-01-01	Senior Engineer
8	12155	Keiichiro	Glinert	1965-01-21	1993-09-16	9999-01-01	Engineer
9	12408	Rasiah	Sudkamp	1965-01-10	1995-04-18	9999-01-01	Senior Engineer
10	12643	Morrie	Schurmann	1965-01-30	1998-12-31	9999-01-01	Staff

Deliverable 2 Requirements

You will earn a perfect score for Deliverable 2 by completing all requirements below:

- A query is written and executed to create a Mentorship Eligibility table for current employees who were born between January 1, 1965 and December 31, 1965. **(25 pt)**
- The Mentorship Eligibility table is exported and saved as `mentorship_eligibilty.csv`. **(5 pt)**

Deliverable 3: A written report on the employee database analysis (20 points)

Deliverable 3 Instructions

For this part of the Challenge, you'll write a report to help the manager prepare for the upcoming "silver tsunami."

The analysis should contain the following:

1. **Overview of the analysis:** Explain the purpose of this analysis.

2. **Results:** Provide a bulleted list with four major points from the two analysis deliverables. Use images as support where needed.
3. **Summary:** Provide high-level responses to the following questions, then provide two additional queries or tables that may provide more insight into the upcoming "silver tsunami."
    - How many roles will need to be filled as the "silver tsunami" begins to make an impact?
    - Are there enough qualified, retirement-ready employees in the departments to mentor the next generation of Pewlett Hackard employees?

Deliverable 3 Requirements

Structure, Organization, and Formatting (6 points)

The written analysis has the following structure, organization, and formatting:

- There is a title, and there are multiple sections. **(2 pt)**
- Each section has a heading and subheading. **(2 pt)**
- Links to images are working and displayed correctly. **(2 pt)**

Analysis (14 points)

The written analysis has the following:

1. Overview of the analysis:
    - The purpose of the new analysis is well defined. **(3 pt)**
2. Results:
    - There is a bulleted list with four major points from the two

- There is a bulleted list with four major points from the two analysis deliverables. **(6 pt)**

3. Summary:

    - The summary addresses the two questions and contains two additional queries or tables that may provide more insight. **(5 pt)**

Submission

Once you're ready to submit, make sure to check your work against the rubric to ensure you are meeting the requirements for this Challenge one final time. It's easy to overlook items when you're in the zone!

As a reminder, the deliverables for this Challenge are as follows:

- Deliverable 1: The Number of Retiring Employees by Title
- Deliverable 2: The Employees Eligible for the Mentorship Program
- Deliverable 3: A written report on the employee database analysis (README.md)

Upload the following to your Pewlett-Hackard-Analysis GitHub repository:

1. The Queries folder with the `Employee_Database_challenge.sql` file
2. The Data folder with the `retirement_titles.csv`, `unique_titles.csv`, `retiring_titles.csv`, and `mentorship_eligibilty.csv` files
3. An updated README.md that has your written analysis

To submit your challenge assignment in Canvas, click Submit, then provide the URL of your Pewlett-Hackard-Analysis GitHub repository for grading.

Comments are disabled for graded submissions in BootCampSpot. If you have questions about your feedback, please notify your instructional staff or the Student Success Manager. If you would like to resubmit your work for an improved grade, you can use the **Re-Submit Assignment** button to upload new links. You may resubmit up to 3 times for a total of 4 submissions.

IMPORTANT

> Once you receive feedback on your Challenge, make any suggested updates or adjustments to your work. Then, add this week's Challenge to your professional portfolio.

NOTE

> You are allowed to miss up to two Challenge assignments and still earn your certificate. If you complete all Challenge assignments, your lowest two grades will be dropped. If you wish to skip this assignment, click Next, and move on to the next Module.

Module-7 Rubric

Criteria	Ratings					Pts
Deliverable 1: The number of retiring employees by title	**50 to >46.0 pts Demonstrating Proficiency** ✓A query is written and executed to create a retirement titles table, which is exported. ✓A query is written and executed to create the unique titles table, which is exported. ✓A query is written and executed to create a retiring titles table, which is exported.	**46 to >42.0 pts Approaching Proficiency** ✓A query is written and executed to create a retirement titles table, which is exported. ✓A query is written to retrieve the unique titles table, which is exported but not sorted correctly. ✓A query is written to create a retiring titles table, which is exported but not sorted correctly.	**42 to >39.0 pts Developing Proficiency** ✓A query is written and executed to create a retirement titles table, which is exported ✓A query is written to retrieve the unique titles table, which is exported but does not have the correct columns OR it is not sorted correctly. ✓A query is written to create a retiring titles table, which is exported, but the number of titles is incorrect OR it is not sorted correctly.	**39 to >0.0 pts Emerging** ✓A query is written and executed to create a retirement titles table, which is exported. ✓A query is written to retrieve the unique titles table, which is exported but does not have the correct columns, duplicates are not removed, and it is not sorted correctly. ✓A query is written to create a retiring titles table, which is exported, but the number of titles OR columns is incorrect.	**0 pts Incomplete**	50 pts

Criteria	Ratings					Pts
Deliverable 2: The employees eligible for the mentorship program	**30 to >26.0 pts Demonstrating Proficiency** ✓A query is written and executed to create a mentorship eligibility table for current employees born in 1965. ✓The mentorship eligibility table is exported, and the columns are correct.	**26 to >23.0 pts Approaching Proficiency** ✓A query is written to retrieve the data, both inner join clauses are correct, only ONE of TWO filtering conditions are met, and the table may be ordered correctly. ✓The mentorship eligibility table is exported and saved, but the columns are not correct.	**23 to >19.0 pts Developing Proficiency** ✓A query is written to retrieve data, but only ONE inner join clause is used, only ONE of TWO filtering conditions are met, and the table is not ordered correctly. ✓The mentorship eligibility table is exported and saved, but the columns are not correct.	**19 to >0.0 pts Emerging** ✓A query is written to retrieve the data, but only ONE inner join clause is used, the where clauses do not filter correctly, and the table is not ordered correctly. ✓The mentorship eligibility table is exported and saved, but the columns are not correct	**0 pts Incomplete**	30 pts
Deliverable 3: Structure, Organization, and Formatting	**6 to >5.0 pts Demonstrating Proficiency** ✓The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. ✓Each section has a heading and subheading. ✓There are images, which are formatted and displayed correctly.	**5 to >4.0 pts Approaching Proficiency** ✓The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. ✓Each section has a heading and subheading. ✓There are images, which are formatted and displayed correctly with one or two minor errors.	**4 to >3.0 pts Developing Proficiency** ✓The written analysis has ALL of the following: ✓There is a title, and there are multiple sections. AND ONE of the following: ✓Each section may have a heading and subheading. ✓There are images, which are formatted and displayed correctly with one or two minor errors.	**3 to >0.0 pts Emerging** ✓The written analysis has ALL of the following: ✓There is a title. ✓There may be a subheading for a section. ✓There are no headings for each section, but there are three sections.	**0 pts Incomplete**	6 pts

Criteria	Ratings					Pts
Deliverable 3: Analysis	**14 to >12.0 pts Demonstrating Proficiency** ✓The purpose is well defined. ✓There is a bulleted list that addresses FOUR major results. ✓The summary addresses the TWO questions and contains TWO additional queries or tables.	**12 to >10.0 pts Approaching Proficiency** ✓The purpose is well defined. ✓There is a bulleted list that addresses THREE of the FOUR major results. ✓The summary addresses the TWO questions and contains ONE of TWO additional queries or tables.	**10 to >8.0 pts Developing Proficiency** ✓The purpose is well defined ✓There is a bulleted list that addresses TWO of the FOUR major results. ✓The summary addresses ONE or TWO questions and contains ONE or TWO additional queries or tables.	**8 to >0.0 pts Emerging** ✓The purpose is well defined. ✓There is a bulleted list that addresses ONE of FOUR major results. ✓The summary addresses ONE of the questions and contains ONE additional query or table.	**0 pts Incomplete**	14 pts
						Total Points: 100