CS 425 – Database Organization

Fall 2023

Homework 1.3

Group Members:

Shriya Prasanna (A20521733)

Girish Rajani-Bathija (A20503736)

Due Date: 9/22/23

Contributions:
Both members completed all questions together, compared/discussed answers, and then took 1 file as a submission.

Database Schema Implementation:

```
CREATE DATABASE assignment3;
USE assignment3;

CREATE TABLE IF NOT EXISTS employee (
row_num INT PRIMARY KEY,
first_name VARCHAR(50),
last_name VARCHAR(50),
salary double);

INSERT INTO employee (row_num, first_name, last_name, salary) VALUES
(1, 'Karen', 'Colmenares', 2500),
(2, 'Guy', 'Himuro', 2600),
(3, 'Irene', 'Mikkilineni', 2700),
(4, 'Sigal', 'Tobias', 2800),
(5, 'Shelli', 'Baida', 2900),
(6, 'Alexander', 'Khoo', 3100),
(7, 'Britney', 'Everett', 3900),
(8, 'Sarah', 'Bell', 4000),
(9, 'Diana', 'Lorentz', 4200),
(10, 'Jennifer', 'Whalen', 4400),
(11, 'David', 'Austin', 4800),
(12, 'Valli', 'Pataballa', 4800),
(13, 'Bruce', 'Ernst', 6000),
(14, 'Pat', 'Fay', 6000),
(15, 'Charles', 'Johnson', 6200);
```

PART I: advanced window functions

1. Write a query to compute for the **FIRST_VALUE()** given the above dataset and return the value along with the entire row.

2. Write a query to compute for the **LAST_VALUE()** and return the value along with t,he entire row.
3. Write a query to compute for **LEAD(2)** for Guy and return the value along with the Guy's row.
4. Write a query to compute for **LAG(4)** for Pat and return value along with Pat's row.
5. Write a query to compute the **RANK()** and **DENSE_RANK()** and return the entire dataset, including the rank and dense rank for each employee.
6. Write a query to compute the **RANK()** and **DENSE_RANK()** but only return Valli's and Bruce's rank and dense rank.
7. Write a query to compute the **ROW_NUMBER()** for Irene and Sarah and only return the rows corresponding to them.
8. Write a query to compute the **PERCENT_RANK()** and return the entire dataset, including the percent rank for each employee. Format your PERCENT_RANK() values to 100%.
9. Write a query to compute the **CUME_DIST()** and return the entire dataset, including the percentage rank for each employee. Format your CUME_DIST() values to 2 decimal places.
10. Write a query to compute the **NTILE(4)** and return the entire dataset showing approximately equal groups/buckets.

Windowing activity