Window Function in SQL

Ami Rai E.

Introduction

- An aggregate function returns a single result for the values from one or multiple rows
- Although a window function does the same functionality of an aggregate function, it is different from an aggregate function
- A Window function returns value for each row for the values from one or multiple rows
- MySQL support window functions

Relation Schema – Employees

- Consider the Employees schema used in the following examples

```
CREATE TABLE Employees (
    Emp_no      INT             PRIMARY KEY,
    First_name  VARCHAR(20)     NOT NULL,
    Last_name   VARCHAR(20)     NOT NULL,
    Gender      ENUM ('M','F')  NOT NULL,
    Dept_no     CHAR(4)         NOT NULL,
    Salary      INT             NOT NULL
);
```

Relation State

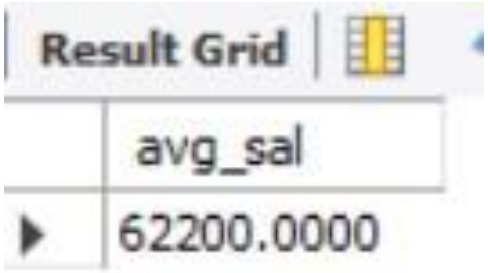
- There are 50 rows in the Employees table
- Consider the part of the populated table

Emp_no	First_name	Last_name	Gender	Dept_no	Salary
10001	Georgi	Facello	M	d009	62000
10002	Bezalel	Simmel	F	d007	65000
10003	Parto	Bamford	M	d004	43000
10004	Chirstian	Koblick	M	d004	42000
10005	Kyoichi	Maliniak	M	d003	54000
10006	Anneke	Preusig	F	d001	83000
10007	Tzvetan	Zielinski	F	d008	47000
10008	Saniya	Kalloufi	M	d002	63000
10009	Sumant	Peac	F	d006	71000
10010	Duangkaew	Piveteau	F	d004	84000
10011	Mary	Sluis	F	d006	69000
10012	Patricio	Bridgland	M	d009	85000
10013	Eberhardt	Terkki	M	d009	79000
10014	Berni	Genin	M	d003	44000
10015	Guoxiang	Nooteboom	M	d002	54000

Aggregate Function in SQL

Average salary of all employees

```
SELECT
    AVG(Salary) AS avg_sal
FROM
    Employees;
```



- AVG() function is used
- A single result is obtained for all the rows

Average salary of employees in each department

```
SELECT
    Dept_no, AVG(Salary) AS avg_sal
FROM
    Employees
GROUP BY Dept_no;
```

Result Grid | Filter

Dept_no	avg_sal
d009	71400.0000
d007	59333.3333
d004	60428.5714
d003	59500.0000
d001	66200.0000
d008	62000.0000
d002	61000.0000
d006	64250.0000
d005	58714.2857

- GROUP BY clause is used to group the employees according to the department number
- A single result for all rows in each group is obtained
- Order of department number is not considered

Average salary of employees in each department (in an order)

```
SELECT
    Dept_no, AVG(Salary) AS avg_sal
FROM
    Employees
GROUP BY Dept_no
ORDER BY Dept_no;
```

Result Grid | Filter Rc

Dept_no	avg_sal
d001	66200.0000
d002	61000.0000
d003	59500.0000
d004	60428.5714
d005	58714.2857
d006	64250.0000
d007	59333.3333
d008	62000.0000
d009	71400.0000

- ORDER BY clause is used to get the result in the order of department numbers
- A single result for all rows in each group is obtained

Number of employees in each department

```
SELECT
    Dept_no, COUNT(Emp_no) AS Count
FROM
    Employees
GROUP BY
    Dept_no
ORDER BY Dept_no;
```

Result Grid

Dept_no	Count
d001	5
d002	7
d003	4
d004	7
d005	7
d006	4
d007	6
d008	5
d009	5

- A single result for all rows in each group is obtained
- COUNT() is used to get the number of rows matching the condition

Number of employees in each department based on gender

```
SELECT
    Dept_no, Gender, COUNT(Emp_no) AS Count
FROM
    Employees
GROUP BY
    Dept_no, Gender
ORDER BY Dept_no;
```

Result Grid | Filter Rows

Dept_no	Gender	Count
d001	M	3
d001	F	2
d002	M	6
d002	F	1
d003	M	3
d003	F	1
d004	M	5
d004	F	2
d005	M	5
d005	F	2
d006	M	1
d006	F	3
d007	M	3
d007	F	3
d008	M	3

- A part of the complete result is shown
- A single result for all rows in each group is obtained
- Grouping is done based on department number and gender

Window Function in SQL

Average salary of all employees

```
SELECT
    Dept_no, AVG(Salary) OVER() AS avg_sal
FROM
    Employees
ORDER BY Dept_no;
```

Result Grid | Filter

Dept_no	avg_sal
d001	62200.0000
d001	62200.0000
d001	62200.0000
d001	62200.0000
d001	62200.0000
d002	62200.0000
d002	62200.0000
d002	62200.0000
d002	62200.0000
d002	62200.0000
d002	62200.0000
d002	62200.0000
d003	62200.0000
d003	62200.0000
d003	62200.0000

- Result for each row is obtained
- OVER clause is empty
- OVER clause specifies how to partition query rows into groups

Average salary of employees in each department

```
SELECT
    Dept_no, AVG(Salary) OVER(PARTITION BY Dept_no) AS avg_sal
FROM
    Employees
ORDER BY dept_no;
```

Result Grid | Filter

Dept_no	avg_sal
d001	66200.0000
d001	66200.0000
d001	66200.0000
d001	66200.0000
d001	66200.0000
d002	61000.0000
d002	61000.0000
d002	61000.0000
d002	61000.0000
d002	61000.0000
d002	61000.0000
d002	61000.0000
d003	59500.0000
d003	59500.0000
d003	59500.0000

- A part of the complete result is shown
- OVER clause partitions rows by department number
- Result for each partition row is obtained

Average salary of employees in each department (remove duplicates)

```
SELECT
    DISTINCT Dept_no,
    AVG(salary) over(partition by Dept_no) AS avg_sal
FROM
    Employees
ORDER BY Dept_no;
```

Result Grid | Filter

dept_no	avg_sal
d001	66200.0000
d002	61000.0000
d003	59500.0000
d004	60428.5714
d005	58714.2857
d006	64250.0000
d007	59333.3333
d008	62000.0000
d009	71400.0000

- The duplicated rows are removed by using DISTINCT keyword

Number of employees in each department

```
SELECT
    Dept_no,
    COUNT(Emp_no) over(partition by Dept_no) AS Count
FROM
    Employees
ORDER BY Dept_no;
```

Result Grid

Dept_no	Count
d001	5
d001	5
d001	5
d001	5
d001	5
d002	7
d002	7
d002	7
d002	7
d002	7
d002	7
d002	7
d002	7
d003	4
d003	4

- A part of the complete result is shown
- Partition is based on department number

Number of employees in each department based on gender

```
SELECT
    Dept_no,
    Gender,
    COUNT(Emp_no) over(partition by Dept_no, Gender) AS Count
FROM
    Employees
ORDER BY Dept_no;
```

Result Grid | Filter Rows

Dept_no	Gender	Count
d001	M	3
d001	M	3
d001	M	3
d001	F	2
d001	F	2
d002	M	6
d002	M	6
d002	M	6
d002	M	6
d002	M	6
d002	M	6
d002	F	1
d003	M	3
d003	M	3

- A part of the complete result is shown
- Partition is based on department number and gender

Number of employees in each department based on gender (remove duplicates)

```
SELECT
    DISTINCT Dept_no,
    Gender,
    COUNT(Emp_no) over(partition by Dept_no, Gender) AS Count
FROM
    Employees
ORDER BY Dept_no;
```

Result Grid | Filter Rows:

Dept_no	Gender	Count
d001	M	3
d001	F	2
d002	M	6
d002	F	1
d003	M	3
d003	F	1
d004	M	5
d004	F	2
d005	M	5
d005	F	2
d006	M	1
d006	F	3
d007	M	3
d007	F	3
d008	M	3

- A part of the complete result is shown
- Partition is based on department number and gender
- DISTINCT keyword is used to remove the duplicates

Conclusion

- Both window function and aggregate functions serve the same purpose
- Difference – The aggregate function returns a single result for all the rows and the window function returns result for each row

Thank You!!