

## Day 41 - 90 days of Analytics: Window functions

In today's video, we looked at Windows functions with SQL queries

The following were mentioned

-Window functions apply aggregate and ranking functions over a particular window (set of rows). The **OVER** clause is used with window functions to define that window. The **OVER** clause does two things:

- Partitions rows into form set of rows. (**PARTITION BY** clause is used)
- Orders rows within those partitions into a particular order. (**ORDER BY** clause is used)

-We should note that If partitions aren't done, then **ORDER BY** orders all rows of table.

-The syntax is as follows

```
SELECT coulmn_name1,  
window_function(cloumn_name2)  
OVER ([PARTITION BY column_name3] [ORDER BY column_name3]) AS new_column  
FROM table_name;
```

Where

- **window\_function**: any aggregate or ranking function
- **column\_name1**: column to be selected
- **coulmn\_name2**: column on which window function is to be applied
- **column\_name3**: column on whose basis partition of rows is to be done
- **new\_column**: Name of new column
- **table\_name**: Name of table

-Various aggregate functions such as **SUM()**, **COUNT()**, **AVERAGE()**, **MAX()**, **MIN()** applied over a particular window (set of rows) are called **aggregate window functions**.

-Examples with window functions

```
SELECT *, ROUND(AVG(Salary)  
OVER (PARTITION BY JobTitle),2) AS Average_Salary_Per_Title  
FROM staff_db.staffsalary;
```

```
SELECT *, ROUND(AVG(Salary)  
OVER (PARTITION BY JobTitle ORDER BY Salary),2) AS Average_Salary_Per_Title  
FROM staff_db.staffsalary;
```

Ranking Window Functions include **RANK()**, **DENSE\_RANK()**, **ROW\_NUMBER()**

-**RANK()**: the rank function assigns rank to all the rows within every partition. Rank is assigned such that rank 1 given to the first row and rows having same value are assigned same rank. For the next rank after two same rank values, one rank value will be skipped.

-**DENSE\_RANK()**: It assigns rank to each row within every partition. Just like rank function first row is assigned rank 1 and rows having same value have same rank. The difference between **RANK()** and **DENSE\_RANK()** is that in **DENSE\_RANK()**, for the next rank after two same rank, consecutive integer is used, no rank is skipped.

-**ROW\_NUMBER ()**: It assigns consecutive integers to all the rows within partition. Within a partition, no two rows can have same row number.

**Note** : **ORDER BY()** should be specified compulsorily while using rank window functions.

-Example with the **RANK()** function

```
SELECT *,RANK()  
OVER (PARTITION BY JobTitle ORDER BY Salary) AS Salary_Rank  
FROM staff_db.staffsalary;
```

Link to the YouTube Recording: <https://www.youtube.com/watch?v=4RD8A03JvGE>

[#90daysofanalytics](#) [#community](#) [#dataanalysis](#) [#dataanalyst](#) [#microsoft](#) [#msexcel](#) [#SQL](#)