



Virtual Internship Experience

T-SQL Filter, Aggregate, and Duplicate

Data Engineer VIX

A. Filter, Aggregate, and Duplicate pada TSQL

Filter, Aggregate, dan Duplicate merupakan beberapa dari fungsi di TSQL yang sangat penting untuk di pelajari dan dipahami serta fungsi yang sering digunakan oleh role Data Engineer.

1. Filter

Filter digunakan untuk melakukan subset item data dari database. Beberapa filter yang sering digunakan :

a. *WHERE*

WHERE merupakan perintah yang di gunakan untuk mem-filter hasil *SELECT* dengan mengekstrak record yang memenuhi persyaratan tertentu.

```
SELECT kolom1, kolom2, ... FROM nama_tabel WHERE kondisi;
```

b. *AND, OR, NOT*

Operator AND, OR dan NOT merupakan perintah yang biasanya di kombinasikan dengan perintah *WHERE*. Ketiganya di gunakan untuk mem-filter record berdasarkan suatu kondisi, operator *AND* akan menampilkan record apabila semua kondisi bernilai *TRUE*, operator *OR* akan menampilkan record apabila salah satu kondisi bernilai *TRUE*, sedangkan operator *NOT* akan menampilkan record apabila semua kondisi bernilai *FALSE*.

```
SELECT kolom1, kolom2, ... FROM nama_tabel WHERE kondisi1 AND kondisi2 AND kondisi3;
```

```
SELECT kolom1, kolom2, ... FROM nama_tabel WHERE kondisi1 OR
kondisi2 OR kondisi3 ...;
```

C. BETWEEN

Operator BETWEEN memilih nilai dalam rentang tertentu. Nilai dapat berupa angka, teks, atau tanggal. Operator BETWEEN bersifat inklusif, nilai awal dan akhir disertakan.

```
SELECT column_name(s)
```

```
FROM table_name
```

```
WHERE column_name BETWEEN value1 AND value2;
```

2. Aggregate

Agregat dalam TSQL melakukan perhitungan pada sekelompok nilai dan kemudian mengembalikan satu nilai. Berikut ini adalah beberapa fungsi agregat yang paling umum digunakan:

a. SUM

Fungsi SUM digunakan untuk menghitung jumlah dari sekumpulan nilai

Syntax :

```
SELECT SUM(ColumnName)
```

```
FROM TableName;
```

b. AVG

Fungsi AVG digunakan untuk mengembalikan nilai rata-rata

Syntax :

```
SELECT AVG(ColumnName)
```

```
FROM TableName;
```

c. COUNT

Fungsi COUNT digunakan untuk menghitung jumlah baris yang dikembalikan dalam pernyataan SELECT.

Syntax :

```
SELECT COUNT(ColumnName)
```

```
FROM TableName;
```

d. MIN

Fungsi MIN digunakan untuk mengembalikan nilai minimum kolom.

Syntax :

```
SELECT MIN(ColumnName)
```

```
FROM TableName;
```

e. MAX

Fungsi MAX digunakan untuk mengembalikan nilai maksimum kolom.

Syntax :

```
SELECT MAX(ColumnName)
```

```
FROM TableName;
```

3. DUPLICATE

Duplicate dalam TSQL merupakan terdapat dua atau lebih data yang sama. Kita dapat menghapus duplicate dengan menggunakan beberapa perintah berikut :

a. SELECT DISTINCT

Pernyataan SELECT DISTINCT digunakan untuk mengembalikan hanya nilai yang berbeda (berbeda).

```
SELECT DISTINCT column1, column2, ...
```

```
FROM table_name;
```

b. COUNT

Fungsi COUNT mengembalikan jumlah baris yang cocok dengan kriteria yang ditentukan.

```
SELECT COUNT(column_name)
```

```
FROM table_name;
```

c. *GROUP BY*

Pernyataan GROUP BY mengelompokkan baris yang memiliki nilai yang sama ke dalam baris ringkasan, seperti "temukan jumlah pelanggan di setiap negara". Pernyataan GROUP BY sering digunakan dengan fungsi agregat (COUNT, MAX, MIN, SUM, AVG) untuk mengelompokkan kumpulan hasil menurut satu atau beberapa kolom.

```
SELECT column_name(s)
```

```
FROM table_name
```

```
WHERE condition
```

```
GROUP BY column_name(s)
```

```
ORDER BY column_name(s);
```

Referensi

<https://www.ibm.com/docs/en/tivoli-netcoolimpact/7.1?topic=filters-sql>

<https://towardsdatascience.com/data-science-lesson-3-filtering-data-using-sql-45d90986487f>

<https://idcloudhost.com/panduan-sql-fungsi-cara-kerja-serta-perintah-dasarnya/>

<https://www.w3schools.com/sql>

<https://www.edureka.co/blog/sql-functions>