

POSTGRESQL

TUTORIAL!!!!

CONTENTS

DOWNLOADS.....	3
SELECT	5
SELECT DISTINCT.....	6
COUNT	7
SELECT WHERE.....	12
ORDER BY	19
LIMIT.....	22
BETWEEN.....	23
IN	25
LIKE AND ILIKE	26
AGGREGATE FUNCTIONS	32
GROUP BY.....	34
HAVING	36
AS.....	39

DOWNLOADS

Time to get set-up for the course!

We will install

PostgreSQL (also known as Postgres) is an open-source relational database management system (RDBMS) that is known for its robustness, reliability, and advanced features. It provides a powerful and scalable platform for managing large volumes of structured data. PostgreSQL supports a wide range of data types, indexing options, and query optimization techniques, making it suitable for various applications. It offers ACID (Atomicity, Consistency, Isolation, Durability) compliance and supports transactions, which ensures data integrity and reliability. With its extensibility and support for various programming languages, PostgreSQL is widely used in enterprise environments and by developers for building high-performance, data-driven applications

Time to get set-up for the course!

We will install

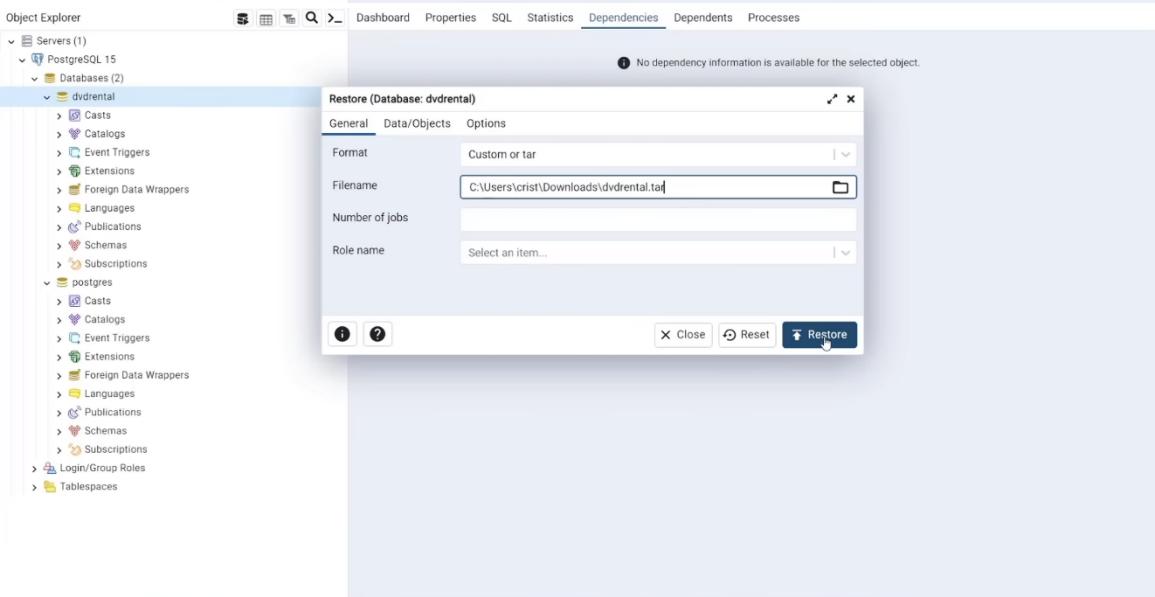
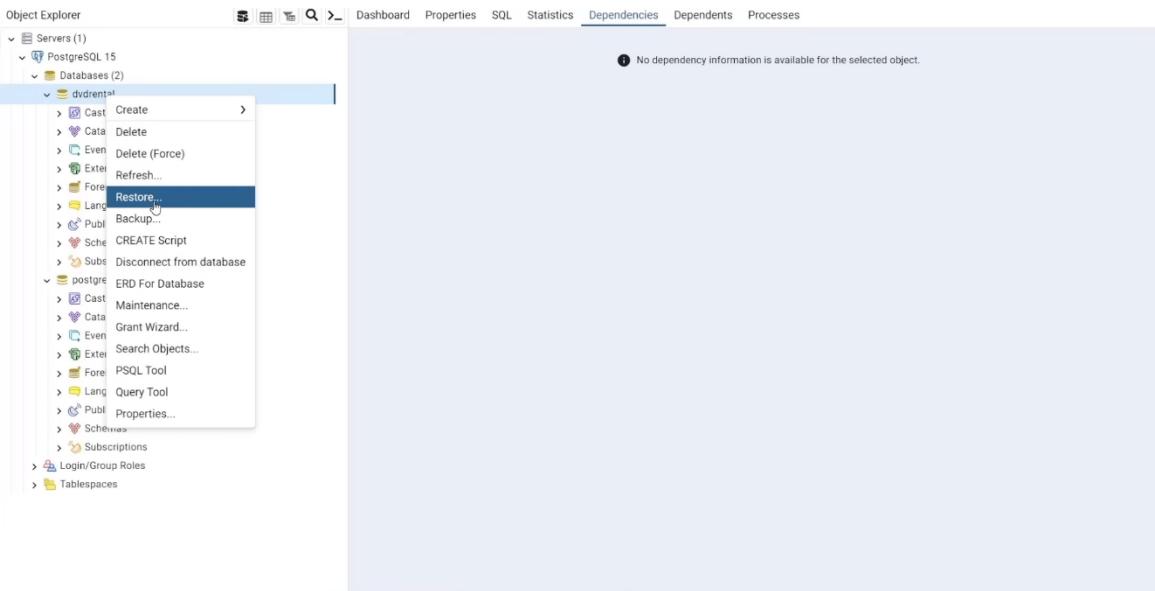
pgAdmin is a free and open-source administration and development platform for managing PostgreSQL databases. It provides a graphical user interface (GUI) that allows users to interact with the database and perform various administrative tasks. pgAdmin offers a range of features, including database object management (such as creating tables, views, and indexes), querying and editing data, monitoring database activity, and managing server settings. It provides a user-friendly environment for database administrators and developers to efficiently work with PostgreSQL databases, allowing them to visually design database schemas, write and execute SQL queries, and monitor database performance. pgAdmin is available for multiple operating systems and is widely used as a primary tool for PostgreSQL database administration.

Time to get set-up for the course!

We will install

PostgreSQL - SQL Engine that stores data and read queries and returns information.

PgAdmin - Graphical User Interface for connecting with PostgreSQL



SELECT

SELECT statement is used to retrieve data from one or more database tables. It allows you to specify the columns you want to retrieve, the table(s) from which you want to retrieve the data, and any conditions that must be met for the data to be included in the result set.

Later we will learn how to combine **SELECT** statement with other SQL statements to perform more complex queries.

SELECT c1, c2 FROM table_3

Database

Table 1

c1	c2	c3
g	54	a
t	67	b
m	89	d

Table 2

c1	c2	c3
1	p	r
2	y	w
4	k	v

Table 3

c1	c2	c3
r	8	78
b	9	99
q	1	85

SELECT * FROM table_1

Database

Table 1

c1	c2	c3
g	54	a
t	67	b
m	89	d

Table 2

c1	c2	c3
1	p	r
2	y	w
4	k	v

Table 3

c1	c2	c3
r	8	78
b	9	99
q	1	85

SELECT DISTINCT

In SQL, the **DISTINCT** statement is used to filter the result set of a query and retrieve only unique rows from the specified column or combination of columns. It is particularly useful when you want to eliminate duplicate records from the query results.

Example:

Consider a table called "employees" with the following data:

employee_id	first_name	last_name	department
1	John	Smith	HR
2	Jane	Doe	IT
3	John	Smith	HR
4	Alice	Johnson	HR

If you run the following query:

```
SELECT DISTINCT first_name, last_name  
FROM employees;
```

The result will be:

first_name	last_name
John	Smith
Jane	Doe
Alice	Johnson

COUNT

COUNT Function in SQL:

The **COUNT** function in SQL is an aggregate function that allows you to count the number of rows that match a specified condition within a table or a result set. It returns a single value representing the count of rows that meet the criteria.

How to Use **COUNT** with the **DISTINCT** Statement:

The **DISTINCT** keyword is used to eliminate duplicate rows from the result set. When combined with the **COUNT** function, it allows you to count the number of unique values in a specific column. Here's how you can use it:

Let's consider a sample table "employees":

emp_id	emp_name	departament
1	John	HR
2	Alice	IT
3	Bob	Finance
4	Emily	HR
5	Michael	IT

Example 1: Count all employees in the "employees" table.

SELECT COUNT(*) FROM employees;

Count
5

Example 2: Count the number of unique departments in the "employees" table using **DISTINCT**.

SELECT COUNT(DISTINCT department)
FROM employees;

emp_id	emp_name	departament	Count
1	John	HR	
2	Alice	IT	
3	Bob	Finance	
4	Emily	HR	
5	Michael	IT	2

Example 2: Count the number of unique departments in the "employees" table using **DISTINCT**.

SELECT COUNT(DISTINCT department)
FROM employees;

emp_id	emp_name	departament	Count
1	John	HR	
2	Alice	IT	
3	Bob	Finance	
4	Emily	HR	
5	Michael	IT	3

In this example, we used the **COUNT** function with **DISTINCT** to count the unique departments in the "employees" table. It returned the count as 3 because there are three unique departments: HR, IT, and Finance.

I hope these examples help illustrate the usage and usefulness of the **COUNT** function in SQL, particularly when combined with the **DISTINCT** statement. It allows you to perform various data analysis tasks and retrieve valuable information from your database.

In SQL, the usage of the **COUNT** function with and without asterix depends on what you want to count. Let's explore the differences and when to use each form:

1. **COUNT with Parentheses and asterix:**

When you use **COUNT(*)** with parentheses and asterix, it counts all the rows in the specified table, regardless of the values in the columns. Here's the syntax:

```
SELECT COUNT(*) FROM table_name;
```

Usage:

Use **COUNT(*)** when you want to count the total number of rows in a table, irrespective of whether there are NULL values or not. It counts all rows, including those with NULL values in any column.

COUNT(*):

The **COUNT(*)** form of the **COUNT** function counts all rows in the specified table or result set, regardless of whether there are NULL values in the columns. It essentially counts the total number of rows, including those with NULL values in any column.

Here's the syntax:

Example:

Consider a table "students":

student_id	student_name	age
1	Alice	20
2	Bob	21
3	Chris	NULL
4	David	19
5	NULL	22

COUNT(column_name):

The **COUNT(column_name)** form of the **COUNT** function counts the number of non-NULL values in the specified column. It ignores NULL values in that particular column and only counts the rows where the specified column has a non-NULL value.

Here's the syntax:

```
SELECT COUNT(student_name) FROM students;
```

Usage:

We use **COUNT(column_name)** when you want to count the occurrence of NON-NULL values in a specific column. It ignore NULL values in the column and counts only NON-NULL entries.

**SELECT COUNT(student_name)
FROM table_name;**

student_id	student_name	age
1	Alice	20
2	Bob	21
3	Chris	NULL
4	David	19
5	NULL	22

Count
4

In summary:

COUNT(*): Counts all rows in the table, including rows with NULL values in any column.

COUNT(column_name): Counts the number of non-NULL values in the specified column, ignoring rows with NULL values in that column.

Choose the appropriate form of the **COUNT** function based on your specific requirement. If you want to count all rows in the table, regardless of NULL values, use **COUNT(*)**. If you want to count the number of non-NULL values in specific column, use **COUNT(column_name)**.

Object Explorer

- Servers (1)
 - PostgreSQL 15
 - Databases (2)
 - dvrental
 - Casts
 - Catalogs
 - Event Triggers
 - Extensions
 - Foreign Data Wrappers
 - Languages
 - Publications
 - Schemas (1)
 - public
 - Aggregates
 - Collations
 - Domains
 - FTS Configurations
 - FTS Dictionaries
 - Aa FTS Parsers
 - FTS Templates
 - Foreign Tables
 - Functions
 - Materialized Views
 - Operators
 - Procedures
 - Sequences
 - Tables (15)
 - actor
 - address
 - category
 - item

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT COUNT(DISTINCT(amount)) FROM payment;
```

Data Output Messages Notifications

count	bigint
1	19

Total rows: 1 of 1 Query complete 00:00:00.194 Ln 1, Cx

Object Explorer

 - Servers (1)
 - PostgreSQL 15
 - Databases (2)
 - dvrental
 - Casts
 - Catalogs
 - Event Triggers
 - Extensions
 - Foreign Data Wrappers
 - Languages
 - Publications
 - Schemas (1)
 - public
 - Aggregates
 - Collations
 - Domains
 - FTS Configurations
 - FTS Dictionaries
 - Aa FTS Parsers
 - FTS Templates
 - Foreign Tables
 - Functions
 - Materialized Views
 - Operators
 - Procedures
 - Sequences
 - Tables (15)
 - actor
 - address
 - category
 - item

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT COUNT(DISTINCT(amount))
```

Execute/Refresh F5

Data Output Messages Notifications

count	bigint
1	19

Total rows: 1 of 1 Query complete 00:00:00.194 Ln 1, Cx

Object Explorer

 - Servers (1)
 - PostgreSQL 15
 - Databases (2)
 - dvrental
 - Casts
 - Catalogs
 - Event Triggers
 - Extensions
 - Foreign Data Wrappers
 - Languages
 - Publications
 - Schemas (1)
 - public
 - Aggregates
 - Collations
 - Domains
 - FTS Configurations
 - FTS Dictionaries
 - Aa FTS Parsers
 - FTS Templates
 - Foreign Tables
 - Functions
 - Materialized Views
 - Operators
 - Procedures
 - Sequences
 - Tables (15)
 - actor
 - address
 - category
 - item

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT COUNT(DISTINCT(amount)) FROM payment;
```

Data Output Messages Notifications

count	bigint
1	19

Total rows: 1 of 1 Query complete 00:00:00.155 Ln 1, Cx

SELECT WHERE

The SQL **WHERE** statement is a crucial part of querying and filtering data in a relational database. It allows you to retrieve specific rows from a table that match certain conditions. In this detailed step-by-step explanation, I'll cover everything you need to know about the SQL **WHERE** statement:

Basic Syntax:

The **WHERE** clause is typically used in combination with the **SELECT** statement to filter rows based on specific conditions. The basic syntax is as follows:

Basic Syntax:

The **WHERE** clause is typically used in combination with the **SELECT** statement to filter rows based on specific conditions. The basic syntax is as follows:

```
SELECT column1, column2, ...
      FROM table_name
      WHERE condition;
```

- **SELECT**: Specifies the columns you want to retrieve from the table.
- **FROM**: Indicates the name of the table you are querying.
- **WHERE**: Introduces the condition that filters the rows. The condition is evaluated for each row in the table, and only the rows that meet the condition will be included in the result set

A condition is a logical expression that evaluates to true or false for each row in the table. The **WHERE** clause uses these conditions to determine which rows to include in the result set. Common operators used in conditions include:

- **=**: Equal to
- **<> or !=**: Not equal to
- **>**: Greater than
- **<**: Less than
- **>=**: Greater than or equal to
- **<=**: Less than or equal to
- **BETWEEN**: Between a range of values (inclusive)
- **LIKE**: Pattern matching using wildcards (often used with % and _)
- **IN**: Matches any value in a list
- **IS NULL**: Checks for NULL values
- **AND, OR, NOT**: Logical operators for combining multiple conditions

Equal To Operator (=):

The equal to operator (=) checks whether two values are exactly equal. It is used to filter rows where the specified column is equal to a given value.

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department = 'Sales';**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department = 'Sales';**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department <> 'Finance';**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE age > 30;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE salary < 50000;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE experience >= 5;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE rating <= 4.5;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE age BETWEEN 25 AND 40;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE first_name LIKE 'J%';**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department IN ('HR', 'Finance', 'Admin');**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE phone_number IS NULL;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department = 'HR' AND age >= 30;**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

**SELECT * FROM Employees
WHERE department = 'Finance' OR department =
'Marketing';**

employee_id	first_name	last_name	job_title	age	salary	department	experience	rating	phone_number
1	John	Doe	Manager	35	80000	Sales	7	4.5	555-123-4567
2	Jane	Smith	Supervisor	28	60000	HR	5	4.2	555-987-6543
3	Robert	Johnson	Analyst	42	75000	Finance	10	4.7	NULL
4	Jennifer	Williams	Manager	38	90000	Marketing	12	4.9	555-444-3333
5	Michael	Brown	Supervisor	32	650000	Sales	8	4.3	555-777-8888
6	Laura	Davis	Coordinator	27	550000	Admin	3	4.0	555-222-1111
7	James	Miller	Analyst	29	70000	HR	6	4.6	555-999-0000
8	Emily	Anderson	Supervisor	31	68000	Marketing	7	4.4	555-999-0000

The screenshot shows a PostgreSQL database interface with the following details:

- Project Explorer:** Shows various database objects like FTS Parsers, FTS Templates, Foreign Tables, Functions, Materialized Views, Operators, Procedures, Sequences, and Tables (15).
- Tables (15):** Includes actor, address, category, city, country, customer, and film.
- Columns (13) for film:** Includes film_id, title, description, release_year, language_id, rental_duration, rental_rate, length, replacement_cost, rating, last_update, special_features, and fulltext.
- Query Editor:** Displays the SQL query:


```
1 SELECT COUNT(*) FROM film WHERE rental_rate > 4 AND replacement_cost >= 19.99
2 AND rating = 'R';
```
- Data Output:** Shows the result of the query:

count	bigint
1	34
- Messages:** Total rows: 1 of 1 Query complete 00:00:00.278
- Notifications:** Ln 2, C1

ORDER BY

The **ORDER BY** clause in PostgreSQL is used to sort the result set of a query in a specified order. It is used when you want the query's output to be presented in a particular sequence based on one or more columns' values. This can be crucial for making query results more meaningful and easier to interpret.

Usage of **ORDER BY**:

The syntax of the **ORDER BY** clause is as follows:

```
SELECT column1, column2, ...
FROM table_name
WHERE conditions
ORDER BY column1 [ASC | DESC], column2 [ASC | DESC], ...;
```

Here's what each component does:

- **SELECT:** Specifies the columns you want to retrieve from the table.
- **FROM:** Specifies the table from which you're retrieving data.
- **WHERE:** Optional condition to filter rows before sorting.
- **ORDER BY:** Specifies the columns by which the result set should be sorted. You can sort by one or more columns, and for each column, you can specify ascending (ASC, which is the default) or descending (DESC) order.

The screenshot shows the pgAdmin 4 interface. On the left is the Object Browser tree, which includes categories like Functions, Materialized Views, Operators, Procedures, Sequences, Tables (15), Triggers, and various system catalogs. The Tables node is expanded, showing tables such as actor, address, category, city, country, customer, film, film_actor, film_category, inventory, language, payment, rental, staff, and store. The right side of the screen has a toolbar at the top with various icons for file operations, followed by a Query History tab and a Scratch Pad tab. Below the toolbar is a large text area containing a single-line SQL query: "1 SELECT * FROM customer ORDER BY first_name ASC;". At the bottom of this area, it says "Total rows: 599 of 599" and "Query complete 00:00:00.289". Below the query editor is the Data Output tab, which displays the results of the query as a table. The table has columns: customer_id [PK] integer, store_id smallint, first_name character varying (45), last_name character varying (45), email character varying (50), address_id smallint, activebool boolean, and create_date. The data consists of 599 rows, with the first few rows being: (1, 375, Aaron, Selby, aaron.selby@sakilacustomer.org, 380, true, 20), (2, 367, Adam, Gooch, adam.gooch@sakilacustomer.org, 372, true, 20), (3, 525, Adrian, Clary, adrian.clary@sakilacustomer.org, 531, true, 20), (4, 217, Agnes, Bishop, agnes.bishop@sakilacustomer.org, 221, true, 20), (5, 389, Alan, Kahn, alan.kahn@sakilacustomer.org, 394, true, 20), (6, 352, Albert, Crouse, albert.crouse@sakilacustomer.org, 357, true, 20), (7, 568, Alberto, Henning, alberto.henning@sakilacustomer.org, 574, true, 20), and (8, 454, Alex, Gresham, alex.gresham@sakilacustomer.org, 450, true, 20). A green success message at the bottom right of the Data Output tab says "Successfully run. Total query runtime: 289 msec. 599 rows affected".

Object Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvdrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT * FROM customer ORDER BY first_name DESC;
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
5	359	Willie	Markham	willie.markham@sakilacustomer.org	364	true	2006-02-01
6	219	Willie	Howell	willie.howell@sakilacustomer.org	223	true	2006-02-01
7	303	William	Satterfield	william.satterfield@sakilacustomer.org	308	true	2006-02-01
8	578	Willard	Lumpkin	willard.lumpkin@sakilacustomer.org	584	true	2006-02-01
9	469	Wesley	Bull	wesley.bull@sakilacustomer.org	474	true	2006-02-01
10	115	Wendy	Harrison	wendy.harrison@sakilacustomer.org	119	true	2006-02-01
11	370	Wayne	Truong	wayne.truong@sakilacustomer.org	375	true	2006-02-01
12	462	Warren	Sherrod	warren.sherrod@sakilacustomer.org	463	true	2006-02-01

Total rows: 599 of 599 Query complete 00:00:00.302 ✓ Successfully run. Total query runtime: 302 msec. 599 rows affected

Object Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvdrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT * FROM customer ORDER BY store_id
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
4	3	Linda	Williams	linda.williams@sakilacustomer.org	7	true	2006-02-01
5	101	Peggy	Myers	peggy.myers@sakilacustomer.org	105	true	2006-02-01
6	102	Crystal	Ford	crystal.ford@sakilacustomer.org	106	true	2006-02-01
7	103	Gladys	Hamilton	gladys.hamilton@sakilacustomer.org	107	true	2006-02-01
8	104	Rita	Graham	rita.graham@sakilacustomer.org	108	true	2006-02-01
9	105	Dawn	Sullivan	dawn.sullivan@sakilacustomer.org	109	true	2006-02-01
10	106	Connie	Wallace	connie.wallace@sakilacustomer.org	110	true	2006-02-01
11	107	Florence	Woods	florence.woods@sakilacustomer.org	111	true	2006-02-01

Total rows: 599 of 599 Query complete 00:00:00.313 ✓ Successfully run. Total query runtime: 313 msec. 599 rows affected

Project Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvrrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT * FROM customer ORDER BY store_id, first_name;
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	cre
1	367	Adam	Gooch	adam.gooch@sakilacustomer.org	372	true	201
2	389	Alan	Kahn	alan.kahn@sakilacustomer.org	394	true	201
3	352	Albert	Crouse	albert.crouse@sakilacustomer.org	357	true	201
4	51	Alice	Stewart	alice.stewart@sakilacustomer.org	55	true	201
5	152	Alicia	Mills	alicia.mills@sakilacustomer.org	156	true	201
6	548	Allan	Cornish	allan.cornish@sakilacustomer.org	554	true	201
7	196	Alma	Austin	alma.austin@sakilacustomer.org	200	true	201
8	139	Amber	Dixon	amber.dixon@sakilacustomer.org	140	true	201

Total rows: 599 of 599 Query complete 00:00:00.348 ✓ Successfully run. Total query runtime: 348 msec. 599 rows affected

Project Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvrrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT store_id, first_name, last_name FROM customer
2 ORDER BY store_id ASC, first_name DESC;
```

Data Output Messages Notifications

store_id	first_name	last_name
1	Zachary	Hite
2	Wendy	Harrison
3	Wanda	Patterson
4	Wallace	Sloane
5	Wade	Delvalle
6	Vivian	Ruiz
7	Virgil	Wofford
8	Violet	Rodriquez
9	Vincent	Ralston

Total rows: 599 of 599 Query complete 00:00:00.214 ✓ Successfully run. Total query runtime: 214 msec. 599 rows affected

LIMIT

The **LIMIT** command in PostgreSQL is used to restrict the number of rows returned by a query. It allows you to specify a maximum number of rows to be included in the result set. This can be particularly useful when you have a large dataset and only need to retrieve a subset of the data for analysis or display purposes.

Here's how the **LIMIT** command works:

```
SELECT column1, column2, ...
FROM table_name
LIMIT number_of_rows;
```

In this SQL query:

- **column1, column2, etc.:** The columns you want to retrieve data from.
- **table_name:** The name of the table you want to query.
- **number_of_rows:** The maximum number of rows you want to retrieve.

For example, if you have a table named employees and you want to retrieve the first 10 employees, you would use the following query:

```
SELECT first_name, last_name
FROM employees
LIMIT 10;
```

BETWEEN

- The **BETWEEN** command in PostgreSQL is used to filter data within a specified range of values. It is typically used in conjunction with the **WHERE** clause to retrieve rows that fall within a specified range of values for a particular column.

Here's a breakdown of its usage and significance:

1. Syntax: The basic syntax of the **BETWEEN** command is as follows:

```
SELECT column_name(s)  
FROM table_name  
WHERE column_name BETWEEN  
value1 AND value2;
```

Performance: When using the BETWEEN command, make sure the column you're applying it to has an appropriate index. Indexing the column can significantly improve query performance, as PostgreSQL can quickly narrow down the rows within the specified range.

Examples:

- Retrieving orders made within a specific date range:

```
SELECT order_id, order_date  
FROM orders  
WHERE order_date BETWEEN '2023-01-01' AND '2023-06-30'
```

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT * FROM payment
2 WHERE amount NOT BETWEEN 8 AND 9;
```

Data Output Messages Notifications

payment_id	customer_id	staff_id	rental_id	amount	payment_date
1	17503	341	2	1520	7.99 2007-02-15 22:25:46.996577
2	17504	341	1	1778	1.99 2007-02-16 17:23:14.996577
3	17505	341	1	1849	7.99 2007-02-16 22:41:45.996577
4	17506	341	2	2829	2.99 2007-02-19 19:39:56.996577
5	17507	341	2	3130	7.99 2007-02-20 17:31:48.996577
6	17508	341	1	3382	5.99 2007-02-21 12:33:49.996577
7	17509	342	2	2190	5.99 2007-02-17 23:58:17.996577
8	17510	342	1	2914	5.99 2007-02-20 02:11:44.996577
9	17511	342	1	3081	2.99 2007-02-20 13:57:39.996577

Total rows: 1000 of 14157 Query complete 00:00:00.368

✓ Server connected Ln 2

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT COUNT(*) FROM payment
2 WHERE amount NOT BETWEEN 8 AND 9; I
```

Data Output Messages Notifications

count	bigint
1	14157

Total rows: 1 of 1 Query complete 00:00:00.152

✓ Server connected Ln 2

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT * FROM payment
2 WHERE amount NOT BETWEEN 8 AND 9;
```

Data Output Messages Notifications

count	bigint
1	14157

Total rows: 1 of 1 Query complete 00:00:00.152

✓ Server connected Ln 1

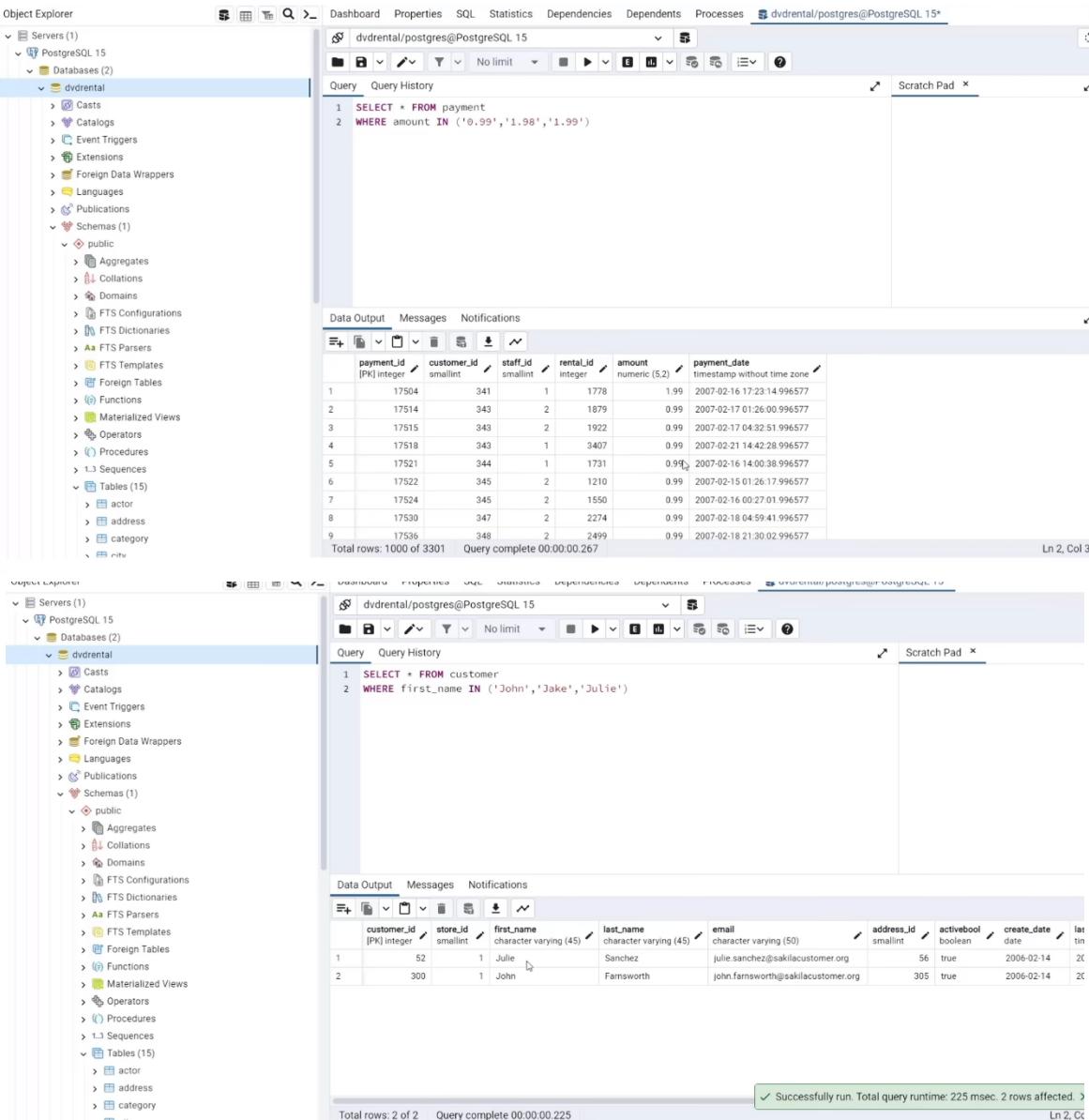
IN

The **IN** command in PostgreSQL is used to filter data based on a specified list of values for a particular column. It allows you to retrieve rows where a specified column's value matches any of the values in the given list.

Here's a detailed explanation of the **IN** command and its significance:

1. Syntax: The basic syntax of the IN command is as follows:

```
SELECT column_name(s)  
FROM table_name  
WHERE column_name IN (value1, value2, ...);
```



The screenshot shows two separate sessions in pgAdmin 4. Both sessions are connected to the 'dvdrental' database on PostgreSQL 15.

Session 1 (Top): The query is:`1 SELECT * FROM payment
2 WHERE amount IN ('0.99','1.98','1.99')`

The results are:

payment_id	customer_id	staff_id	rental_id	amount	payment_date
1	17504	341	1	1.99	2007-02-16 17:23:14.996577
2	17514	343	2	1.98	2007-02-17 01:26:00.996577
3	17515	343	2	1.99	2007-02-17 04:32:51.996577
4	17518	343	1	1.99	2007-02-21 14:42:28.996577
5	17521	344	1	1.99	2007-02-16 14:00:38.996577
6	17522	345	2	1.99	2007-02-15 01:26:17.996577
7	17524	345	2	1.99	2007-02-16 00:27:01.996577
8	17530	347	2	1.99	2007-02-18 04:59:41.996577
9	17536	348	2	1.99	2007-02-18 21:30:02.996577

Total rows: 1000 of 3301 Query complete 00:00:00.267 Ln 2, Col 3!

Session 2 (Bottom): The query is:`1 SELECT * FROM customer
2 WHERE first_name IN ('John','Jake','Julie')`

The results are:

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date	last_update
1	52	John	Farmsworth	john.farmsworth@sakilacustomer.org	305	true	2006-02-14 20:20:45.0	2006-02-14 20:20:45.0
2	300	Jake	Sanchez	julie.sanchez@sakilacustomer.org	56	true	2006-02-14 20:20:45.0	2006-02-14 20:20:45.0

Total rows: 2 of 2 Query complete 00:00:00.225 ✓ Successfully run. Total query runtime: 225 msec. 2 rows affected. Ln 2, Col 3!

LIKE AND ILIKE

The **LIKE** and **ILIKE** commands in PostgreSQL are used for pattern matching within text columns. They allow you to filter rows based on patterns defined using wildcard characters. The main difference between the two commands is that **LIKE** performs case-sensitive matching, while **ILIKE** performs case-insensitive matching.

Here's a detailed explanation of the **LIKE** and **ILIKE** commands and their significance:

1. Syntax:

- The basic syntax of the **LIKE** command is:

```
SELECT column_name(s)  
FROM table_name  
WHERE column_name LIKE pattern;
```

- The basic syntax of the **ILIKE** command is:

```
SELECT column_name(s)  
FROM table_name  
WHERE column_name ILIKE pattern;
```

2. Wildcard Characters:

Wildcards are special characters used to represent one or more characters in a pattern.

- **'%'**: Represents any sequence of characters (including zero characters).
- **'_'**: Represents a single character.
- **'[]'**: Represents a character class, allowing you to match any single character within the specified set.

4. Case Sensitivity:

LIKE: Performs case-sensitive pattern matching. For example, if you use LIKE 'apple%', it will match rows with values like "apple", "applesauce", but not "Apple".

ILIKE: Performs case-insensitive pattern matching. It will match rows regardless of the case. For example, ILIKE 'apple%' would match "apple", "Applesauce", and "APPLEJUICE"

Examples:

- Retrieving all customers with names starting with "J":

```
SELECT customer_name, contact_email  
FROM customers  
WHERE customer_name LIKE 'J%';
```

- Finding products with names containing "blue":

```
SELECT product_name, category  
FROM products  
WHERE product_name LIKE '%blue%'
```

- Searching for employees with last names ending in "son" (case-insensitive):

```
SELECT first_name, last_name  
FROM employees  
WHERE last_name ILIKE '%son';
```

Wildcard characters are special symbols used in pattern matching within SQL queries. They are extremely useful for performing flexible searches and filtering based on specific patterns within text columns. Here's an explanation of the common wildcard characters, their benefits, and when to use them:

% (Percent Sign):

Represents any sequence of characters (including zero characters).

Use cases:

- Searching for values that start with or end with a specific sequence of characters.

- Finding records that contain a particular substring within a larger string.

Example:

```
SELECT product_name  
FROM products  
WHERE product_name LIKE '%apple%';
```

This query retrieves all products with names containing the word "apple".

_ (Underscore):

- Represents a single character.

- Use cases:

- Searching for values with a specific character at a certain position.

- Finding records with a particular pattern where you know the character at a specific position.

Example:

```
SELECT customer_name  
FROM customers  
WHERE customer_name LIKE 'J_n%';
```

This query retrieves customers with names that start with "J" and have any second character followed by "n".

[] (Square Brackets):

- Represents a character class, allowing you to match any single character within the specified set.
 - Use cases:
 - Searching for values that can have variations in a specific character position (e.g., for alternate spellings or formats).
- Example:

```
SELECT product_name  
FROM products  
WHERE product_name LIKE 'colo[u]r%';
```

Example:

```
SELECT product_name  
FROM products  
WHERE product_name LIKE 'colo[u]r%';
```

This query matches products with names starting with "color" and accounts for variations in spelling, such as "colour" in British English.

Combining the percent sign (%), underscore (_), and square brackets ([])) allows you to create even more complex and specific pattern-based searches within SQL queries. This combination is particularly useful when you need to search for data with a specific structure or pattern while accounting for variations. Here's an explanation of how to use this combination, along with a use case:

Syntax for Combining Wildcard Characters:

'%': Represents any sequence of characters (including zero characters).

_: Represents a single character.

[]: Represents a character class, allowing you to match any single character within the specified set.

Use Case - Searching for Email Domains:

Suppose you have a database table containing email addresses and you want to retrieve all email addresses from a specific domain, while allowing for variations in the domain name, such as different subdomains or top-level domains (TLDs). Let's say you're interested in retrieving email addresses from both "example.com" and "example.net".

Query:

```
SELECT email_address  
FROM users  
WHERE email_address LIKE '%example[_].[cno]m';
```

Explanation:

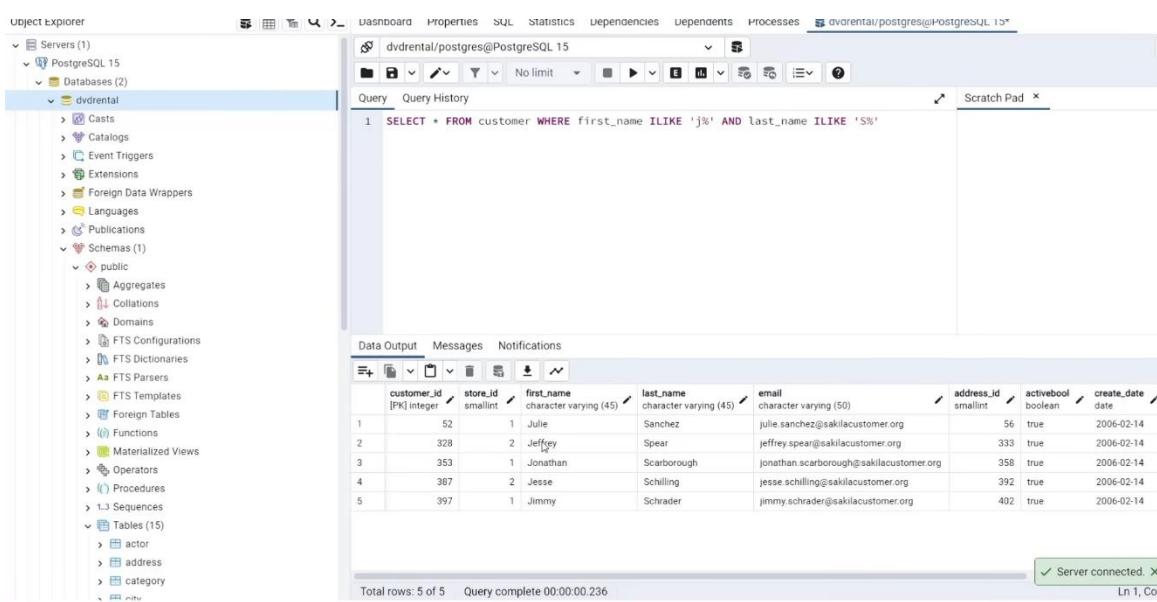
'%**example**' matches any email address that ends with "example".

'**[_]**' matches a single character, which can be any character (including a dot) in the email address.

'**.[cno]**' matches a single character from the character set [cno], which covers variations of the TLD: "c", "n", or "o" (for .com, .net, or .org).

Results:

This query would retrieve email addresses like "user@example.com", "admin@example.net", and any other variations that match the specified pattern.



The screenshot shows the pgAdmin 4 interface. The left pane, 'Object Explorer', displays the database structure under 'PostgreSQL 15' and 'dvrental'. The 'Tables' section lists 'customer' and 'address'. The right pane contains a query editor window with the following content:

```
dvrental/postgres@PostgreSQL 15
SELECT * FROM customer WHERE first_name ILIKE 'j%' AND last_name ILIKE 'S%'
```

The results of the query are displayed in a table:

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
1	52	Julie	Sanchez	julie.sanchez@sakilacustomer.org	56	true	2006-02-14
2	328	Jeffrey	Spear	jeffrey.spear@sakilacustomer.org	333	true	2006-02-14
3	353	Jonathan	Scarborough	jonathan.scarborough@sakilacustomer.org	358	true	2006-02-14
4	387	Jesse	Schilling	jesse.schilling@sakilacustomer.org	392	true	2006-02-14
5	397	Jimmy	Schrader	jimmy.schrader@sakilacustomer.org	402	true	2006-02-14

Total rows: 5 of 5 Query complete 00:00:00.236

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT * FROM customer
2 WHERE first_name LIKE '%er%'
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
1	6	Jennifer	Davis	jennifer.davis@sakilacustomer.org	10	true	2006-02-14
2	24	Kimberly	Lee	kimberly.lee@sakilacustomer.org	28	true	2006-02-14
3	46	Catherine	Campbell	catherine.campbell@sakilacustomer.org	50	true	2006-02-14
4	53	Heather	Morris	heather.morris@sakilacustomer.org	57	true	2006-02-14
5	54	Teresa	Rogers	teresa.rogers@sakilacustomer.org	58	true	2006-02-14
6	59	Cheryl	Murphy	cheryl.murphy@sakilacustomer.org	63	true	2006-02-14
7	61	Katherine	Rivera	katherine.rivera@sakilacustomer.org	65	true	2006-02-14
8	72	Theresa	Watson	theresa.watson@sakilacustomer.org	76	true	2006-02-14

Total rows: 58 of 58 Query complete 00:00:00.314

✓ Server connected. Ln 2, Col 2

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT * FROM customer
2 WHERE first_name LIKE '_er%'
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
1	54	Teresa	Rogers	teresa.rogers@sakilacustomer.org	58	true	2006-02-14
2	156	Bertha	Ferguson	bertha.ferguson@sakilacustomer.org	160	true	2006-02-14
3	158	Veronica	Stone	veronica.stone@sakilacustomer.org	162	true	2006-02-14
4	161	Geraldine	Perkins	geraldine.perkins@sakilacustomer.org	165	true	2006-02-14
5	172	Bernice	Willis	bernice.willis@sakilacustomer.org	176	true	2006-02-14
6	206	Terri	Vasquez	terri.vasquez@sakilacustomer.org	210	true	2006-02-14
7	207	Gertrude	Castillo	gertrude.castillo@sakilacustomer.org	211	true	2006-02-14
8	218	Vera	Mccoy	vera.mccoy@sakilacustomer.org	222	true	2006-02-14

Total rows: 26 of 26 Query complete 00:00:00.233

✓ Server connected. Ln 2, Col 2

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes [dvrental/postgres@PostgreSQL 15*](#)

Query Query History

```
1 SELECT * FROM customer
2 WHERE first_name LIKE '_er_'
```

Data Output Messages Notifications

customer_id	store_id	first_name	last_name	email	address_id	activebool	create_date
1	218	Vera	Mccoy	vera.mccoy@sakilacustomer.org	222	true	2006-02-14

Total rows: 1 of 1 Query complete 00:00:00.187

✓ Server connected. Ln 2

AGGREGATE FUNCTIONS

In PostgreSQL, **AGGREGATE** functions are special functions that perform calculations on a set of values and return a single value as a result. These functions are commonly used to summarize or derive meaningful insights from large datasets in a database. Instead of operating on individual rows, aggregate functions work on a group of rows and return a single value for each group. They are crucial for tasks such as calculating averages, sums, counts, minimum and maximum values, and more.

Here are some key aspects of aggregate functions in PostgreSQL:

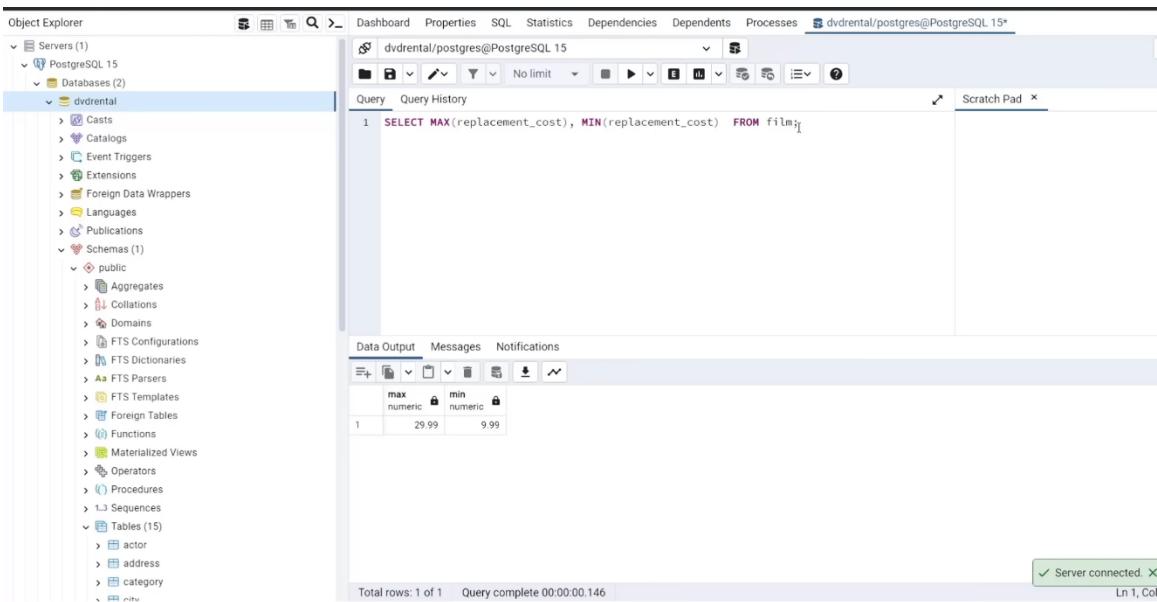
1. Grouping Rows:

Aggregate functions are often used in conjunction with the **GROUP BY** clause. The **GROUP BY** clause divides the result set into groups based on one or more columns. The aggregate functions then operate on each group separately, producing a single result for each group.

2. Common Aggregate Functions:

PostgreSQL provides a wide range of aggregate functions that cater to various analytical needs. Some of the most common aggregate functions include:

- **SUM:** Calculates the sum of a numeric column within a group.
- **AVG:** Computes the average of a numeric column within a group.
- **COUNT:** Counts the number of rows within a group.
- **MIN:** Returns the minimum value from a column within a group.
- **MAX:** Returns the maximum value from a column within a group.

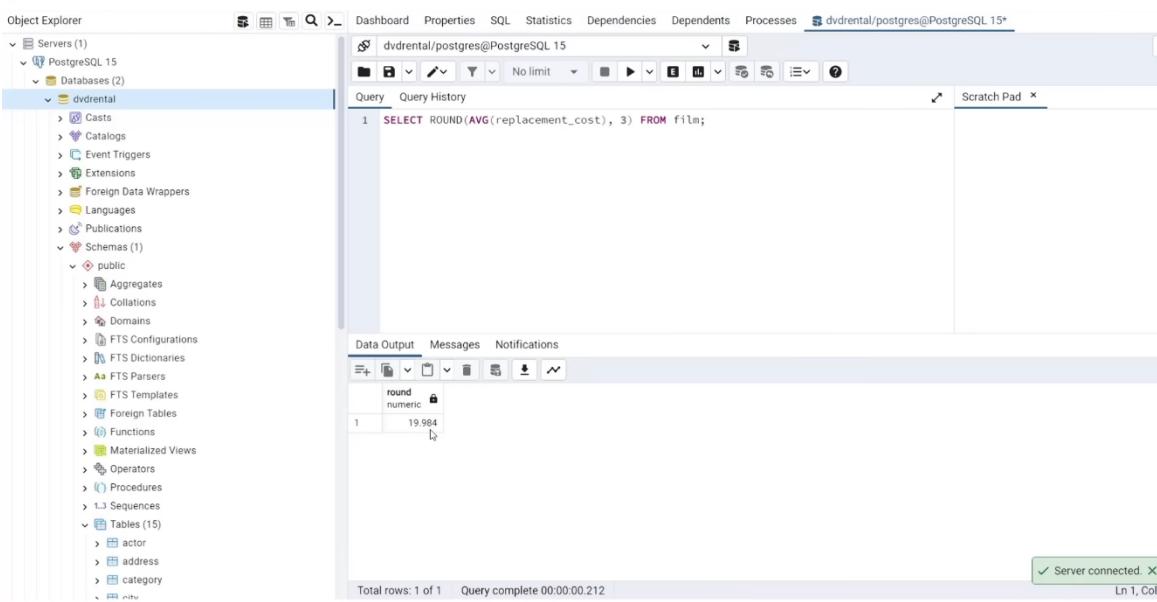


The screenshot shows the Object Explorer on the left with the 'dvrental' database selected. The 'Tables' node under 'dvrental' has 15 tables listed: actor, address, category, city, customer, employee, film, film_actor, film_category, inventory, language, nation, payment, rental, staff, and store. In the main pane, a query window displays the following SQL and its results:

```
1 SELECT MAX(replacement_cost), MIN(replacement_cost) FROM film;
```

	max	min
1	29.99	9.99

Total rows: 1 of 1 Query complete 00:00:00.146



The screenshot shows the Object Explorer on the left with the 'dvrental' database selected. The 'Tables' node under 'dvrental' has 15 tables listed: actor, address, category, city, customer, employee, film, film_actor, film_category, inventory, language, nation, payment, rental, staff, and store. In the main pane, a query window displays the following SQL and its results:

```
1 SELECT ROUND(AVG(replacement_cost), 3) FROM film;
```

	round
1	19.984

Total rows: 1 of 1 Query complete 00:00:00.212

GROUP BY

What is **GROUP BY** and Why is it Necessary?

In SQL, the **GROUP BY** clause is used to group rows from a table based on one or more columns. This grouping enables the application of aggregate functions, such as **SUM**, **COUNT**, **AVG**, **MIN**, and **MAX**, on subsets of data sharing common values in the specified columns. Essentially, it transforms individual rows into summarized information, aiding in the analysis of large datasets and revealing trends and insights that might not be immediately apparent.

The **GROUP BY** operation follows a **Split, Apply, Combine** process. Let's consider a hypothetical sales table with columns "Region," "Product," and "SalesAmount." To analyze total sales for each region, we apply the following steps:

Split: The dataset is divided into groups based on unique values in the "Region" column.

Apply: Aggregate functions, such as **SUM(SalesAmount)**, are applied to the "SalesAmount" column within each region group.

Combine: The aggregated results are combined, creating a summary table that displays the total sales for each region.

The screenshot shows the pgAdmin 4 interface. On the left is the Object Explorer, displaying a tree structure for a PostgreSQL 15 database named 'dvrental'. Under 'Tables (15)', there are entries for 'actor', 'address', and 'category'. The main pane shows a query editor with the following SQL code:

```
1 SELECT customer_id, SUM(amount) FROM payment
2 GROUP BY customer_id
```

Below the query editor is a Data Output tab showing the results of the query:

customer_id	sum
1	184
2	87
3	477
4	273
5	550
6	51
7	394
8	272
9	70

Total rows: 599 of 599 Query complete 00:00:00.242

Object Explorer

Servers (1)

- PostgreSQL 15
- Databases (2)
 - dvrental
 - public

Query Query History

```

1 SELECT customer_id, SUM(amount) FROM payment
2 GROUP BY customer_id
3 ORDER BY SUM(amount) DESC

```

Data Output Messages Notifications

customer_id	sum
148	211.55
526	208.58
178	194.61
137	191.62
144	189.60
459	183.63
181	167.67
410	167.62
236	166.61

Successfully run. Total query runtime: 235 msec. 599 rows affected.

PostgreSQL 15/dvrental - Database connected.

Object Explorer

Servers (1)

- PostgreSQL 15
- Databases (2)
 - dvrental
 - public

Query Query History

```

1 SELECT customer_id, staff_id, SUM(amount) FROM payment
2 GROUP BY customer_id, staff_id
3 ORDER BY SUM(amount)

```

Data Output Messages Notifications

customer_id	staff_id	sum
320	1	5.96
281	1	9.96
318	2	9.97
248	1	10.94
395	2	11.95
344	1	12.95
272	1	14.96
22	2	14.96
319	9	16.46

PostgreSQL 15/dvrental - Database connected.

Object Explorer

Servers (1)

- PostgreSQL 15
- Databases (2)
 - dvrental
 - public

Query Query History

```

1 SELECT DATE(payment_date), SUM(amount) FROM payment
2 GROUP BY DATE(payment_date)
3 ORDER BY SUM(amount)

```

Data Output Messages Notifications

date	sum
2007-02-14	116.73
2007-04-05	273.36
2007-03-16	299.28
2007-04-26	347.21
2007-05-14	514.18
2007-02-21	917.87
2007-02-16	1154.18
2007-02-17	1188.17
2007-02-15	1188.92

Object Explorer

Servers (1)

- PostgreSQL 15
- Databases (2)
 - dvrental
 - public

Query Query History

```

1 SELECT customer_id, COUNT(amount) FROM payment
2 GROUP BY customer_id
3 ORDER BY COUNT(amount)

```

Data Output Messages Notifications

customer_id	count
318	7
281	10
110	12
248	13
272	13
61	13
310	13
48	14
464	14

HAVING

The **HAVING** command in PostgreSQL is used in conjunction with the **GROUP BY** clause to filter the results of a query based on the aggregated values of a column or columns. It allows you to apply conditions to the result set after grouping has been performed. In other words, while the **WHERE** clause filters individual rows before they are grouped, the **HAVING** clause filters groups of rows after they have been grouped.

Here's a breakdown of its key aspects:

1. Purpose: The **HAVING** clause is used to filter rows in the result set after grouping and aggregation, specifically when you want to filter groups based on aggregated values.
2. Usage: It follows the **GROUP BY** clause and comes after it in a SQL query. The basic syntax is:

```
SELECT column1, aggregate_function(column2)  
FROM table  
GROUP BY column1  
HAVING condition;
```

3. Need for **HAVING**: While the **WHERE** clause filters individual rows before grouping, the **HAVING** clause filters groups of rows based on the aggregated results. This is especially useful when you want to filter out groups that meet certain criteria.

4. Scenario to use: You would use **HAVING** when you need to filter out groups based on aggregate functions like **COUNT**, **SUM**, **AVG**, etc. For example, finding the average salary of departments with an average salary greater than a certain value.

5. Comparison with **WHERE**: Use **WHERE** to filter individual rows before grouping and **HAVING** to filter groups after aggregation.

6. Avoid unnecessary complexity: While **HAVING** is powerful, avoid overusing it or making it too complex. Sometimes, it might be more efficient to use subqueries or other methods to achieve the desired result.

Performance Consideration: When using **HAVING**, consider the performance impact, especially when dealing with large datasets. Improper use of **HAVING** can lead to slower query execution times.

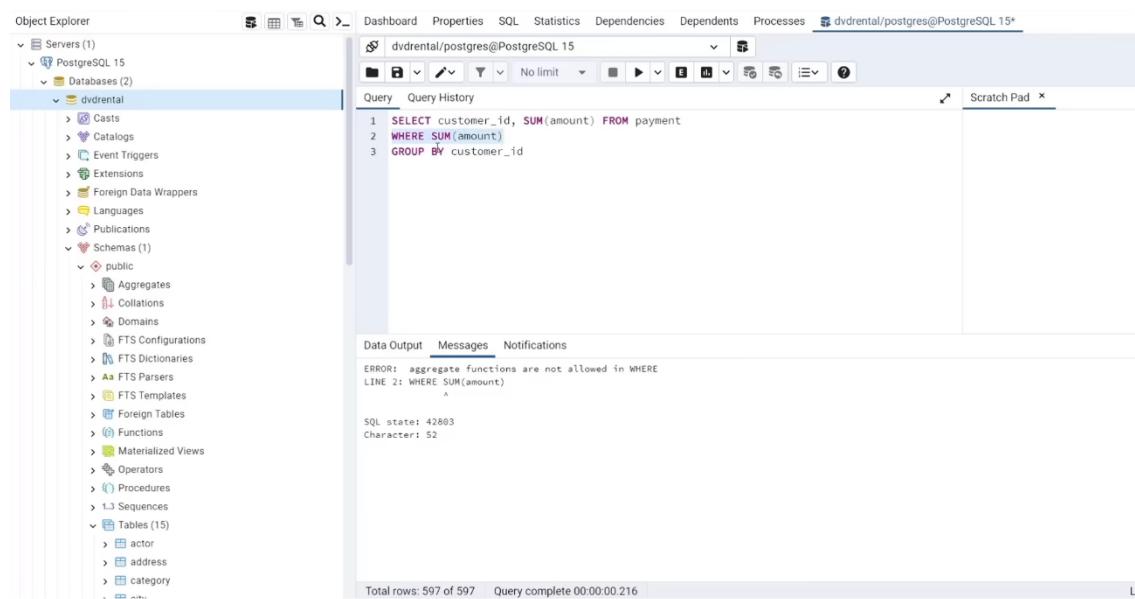
Examples:

To find departments with more than 10 employees:

```
SELECT department, COUNT(*) FROM employees  
GROUP BY department  
HAVING COUNT(*) > 10;
```

- To find the average salary of departments with an average salary above \$50000:

```
SELECT department, AVG(salary) FROM employees  
GROUP BY department  
HAVING AVG(salary) > 50000;
```



The screenshot shows the pgAdmin 4 interface with the Object Explorer on the left and a query editor on the right. The query editor contains the following code:

```
1 SELECT customer_id, SUM(amount) FROM payment
2 WHERE SUM(amount)
3 GROUP BY customer_id
```

The output pane shows an error message:

ERROR: aggregate functions are not allowed in WHERE
LINE 2: WHERE SUM(amount)
^

SQL state: 42803
Character: 52

Total rows: 597 of 597 Query complete 00:00:00.216

Object Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvdrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT customer_id, SUM(amount) FROM payment
2 GROUP BY customer_id
3 HAVING SUM(amount) > 100
```

Data Output Messages Notifications

customer_id	sum
1	87
2	477
3	273
4	550
5	51
6	190
7	424
8	406
9	176

Total rows: 296 of 296 Query complete 00:00:00.246 ✓ Successfully run. Total query runtime: 246 msec. 296 rows affected. Ln 3, Cc

This screenshot shows the pgAdmin 4 interface. The left pane displays the Object Explorer with the 'dvdrental' database selected. The right pane shows a query editor with the following SQL code:

```
1 SELECT customer_id, SUM(amount) FROM payment
2 GROUP BY customer_id
3 HAVING SUM(amount) > 100
```

The results are displayed in a data grid:

customer_id	sum
1	87
2	477
3	273
4	550
5	51
6	190
7	424
8	406
9	176

A message at the bottom indicates the query was successfully run with a runtime of 246 msec and 296 rows affected.

Object Explorer

Dashboard Properties SQL Statistics Dependencies Dependents Processes dvdrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT store_id, COUNT(*) FROM customer
2 GROUP BY store_id
3 HAVING COUNT(*) > 300
4
```

Data Output Messages Notifications

store_id	count
1	326

Total rows: 1 of 1 Query complete 00:00:00.201 ✓ Successfully run. Total query runtime: 201 msec. 1 rows affected. Ln 3, Cc

This screenshot shows the pgAdmin 4 interface. The left pane displays the Object Explorer with the 'dvdrental' database selected. The right pane shows a query editor with the following SQL code:

```
1 SELECT store_id, COUNT(*) FROM customer
2 GROUP BY store_id
3 HAVING COUNT(*) > 300
4
```

The results are displayed in a data grid:

store_id	count
1	326

A message at the bottom indicates the query was successfully run with a runtime of 201 msec and 1 row affected.

AS

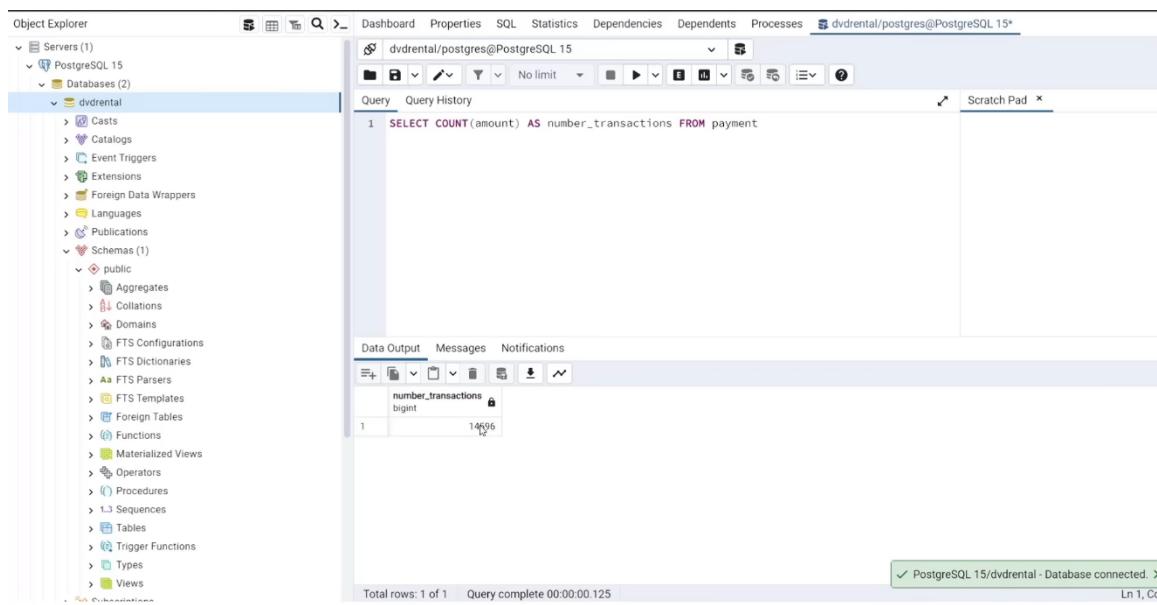
In PostgreSQL, think of the **AS** statement as a tool that lets you give nicknames to tables, columns, or calculations within your queries. Just like how you might give your friends nicknames, you can give these elements easier-to-remember names to make your queries less confusing.

You'd want to use the "**AS**" statement when:

1. You want to give friendlier names to columns in your query results.
2. You're dealing with math or complicated stuff, and you want to simplify things by using a simpler name.
3. You're working with a table and its own copy (kind of like a twin), and you need to tell them apart

Here's how you do it:

```
SELECT original_name AS new_name  
FROM table_name;
```



The screenshot shows the pgAdmin 4 interface. The left pane is the Object Explorer, displaying the database structure under the 'dvrental' schema. The right pane is the Query Editor, showing a single query: 'SELECT COUNT(amount) AS number_transactions FROM payment'. The results pane below shows the output of this query:

	number_transactions
1	14996

Total rows: 1 of 1 Query complete 00:00:00.125

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes dvlrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT customer_id, SUM(amount) AS total_spent FROM payment
2 GROUP BY customer_id
```

Data Output Messages Notifications

customer_id	total_spent
184	80.80
87	137.72
477	106.79
273	130.72
550	151.69
51	123.70
394	77.80
272	65.87
70	75.83

Total rows: 599 of 599 Query complete 00:00:00.237

✓ PostgreSQL 15/dvlrental - Database connected

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes dvlrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT customer_id, SUM(amount) AS total_spent FROM payment
2 GROUP BY customer_id
3 HAVING SUM(amount) > 100
```

Data Output Messages Notifications

customer_id	total_spent
538	109.73
595	110.71
410	167.62
38	127.66
193	105.77
453	111.77
186	111.71
145	107.73
27	123.70

Total rows: 296 of 296 Query complete 00:00:00.217

✓ Successfully run. Total query runtime: 217 msec. 296 rows affected.

✓ PostgreSQL 15/dvlrental - Database connected.

Object Explorer Dashboard Properties SQL Statistics Dependencies Dependents Processes dvlrental/postgres@PostgreSQL 15*

Query Query History

```
1 SELECT customer_id, SUM(amount) AS total_spent FROM payment
2 GROUP BY customer_id
3 HAVING total_spent > 100
```

Data Output Messages Notifications

ERROR: column "total_spent" does not exist
LINE 3: HAVING total_spent > 100
^

SQL state: 42703
Character: 89

Total rows: 296 of 296 Query complete 00:00:00.418