SQL FOR DATA ANALYSIS

a) Using SELECT, WHERE, ORDER BY, GROUP BY

1. Basic SELECT

```
SELECT ID, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping;
```

This query retrieves only the columns ID, Mode of Shipment, and Cost of the Product from the dataset.

2. Filter Records with WHERE

```
SELECT *FROM ecommerce_shipping
WHERE Reached_on_Time_YN = 0;
```

This query selects only those shipments that were not delivered on time (value = 0).

3. Sort Data with ORDER BY

```
SELECT ID, Cost_of_the_Product, Discount_offered
FROM ecommerce_shipping
ORDER BY Cost_of_the_Product DESC;
```

Here, shipments are ordered by product cost in descending order to find the most expensive items first.

4. Group and Aggregate

```
SELECT Warehouse_block, AVG(Customer_rating) AS avg_rating
FROM ecommerce_shipping
GROUP BY Warehouse_block;
```

This groups shipments by warehouse and calculates the average customer rating for each warehouse.

b) Using JOINS (INNER, LEFT, RIGHT)

We created another table warehouse_details to demonstrate joins.

1. INNER JOIN

```
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
INNER JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
```

Fetches only the orders where warehouse information exists in both tables.

2. LEFT JOIN

```
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
LEFT JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
```

Fetches all orders, even if there is no matching warehouse in the reference table.

3. RIGHT JOIN

```
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
RIGHT JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
```

Fetches all warehouses, even if no orders were shipped from them.

C) Using Subqueries

```
SELECT ID, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping
WHERE Cost_of_the_Product = (
SELECT MAX(Cost_of_the_Product)
FROM ecommerce_shipping);
```

The subquery finds the maximum product cost, and the outer query fetches shipment(s) with that cost.

d) Using Aggregate Functions (SUM, AVG, COUNT)

1. Total Revenue

```
SELECT SUM(Cost_of_the_Product) AS total_revenue
FROM ecommerce_shipping;
```

Calculates the sum of product costs to estimate total revenue.

2. Average Discount by Shipment Mode

```
SELECT Mode_of_Shipment, AVG(Discount_offered) AS avg_discount
FROM ecommerce_shipping
GROUP BY Mode_of_Shipment;
```

Groups shipments by mode and calculates the average discount offered.

3. Count Orders by Gender

```
SELECT Gender, COUNT(*) AS total_orders
FROM ecommerce_shipping
GROUP BY Gender;
```

Counts how many orders were placed by each gender.

e) Creating Views for Analysis

1. Late Deliveries View

```
CREATE VIEW late_deliveries AS
SELECT ID, Warehouse_block, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping
WHERE Reached_on_Time_YN = 0;
```

Creates a reusable virtual table showing only late deliveries.

2. Average Cost by Shipment Mode

```
CREATE VIEW shipment_avg_cost AS
SELECT Mode_of_Shipment, AVG(Cost_of_the_Product) AS avg_cost
FROM ecommerce_shipping
GROUP BY Mode_of_Shipment;
```

Creates a view that stores the average cost per shipment mode.

f) Query Optimization with Indexes

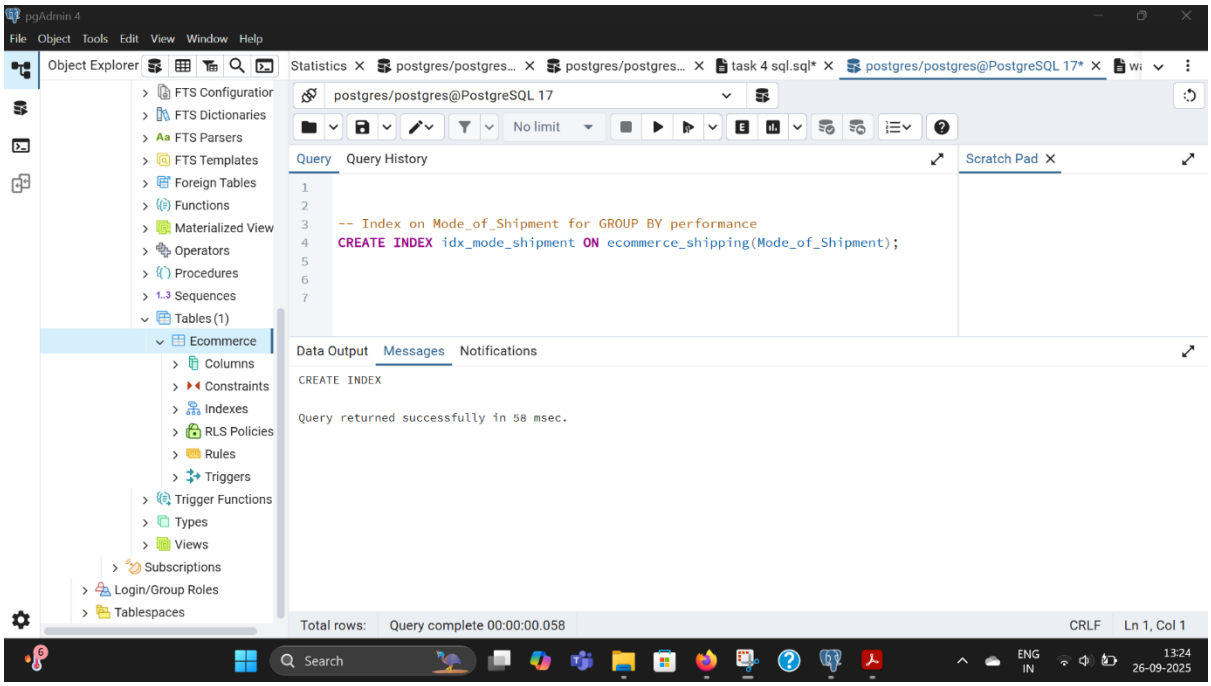
```
CREATE INDEX idx_reached_time ON ecommerce_shipping(Reached_on_Time_YN);
CREATE INDEX idx_mode_shipment ON ecommerce_shipping(Mode_of_Shipment);
```

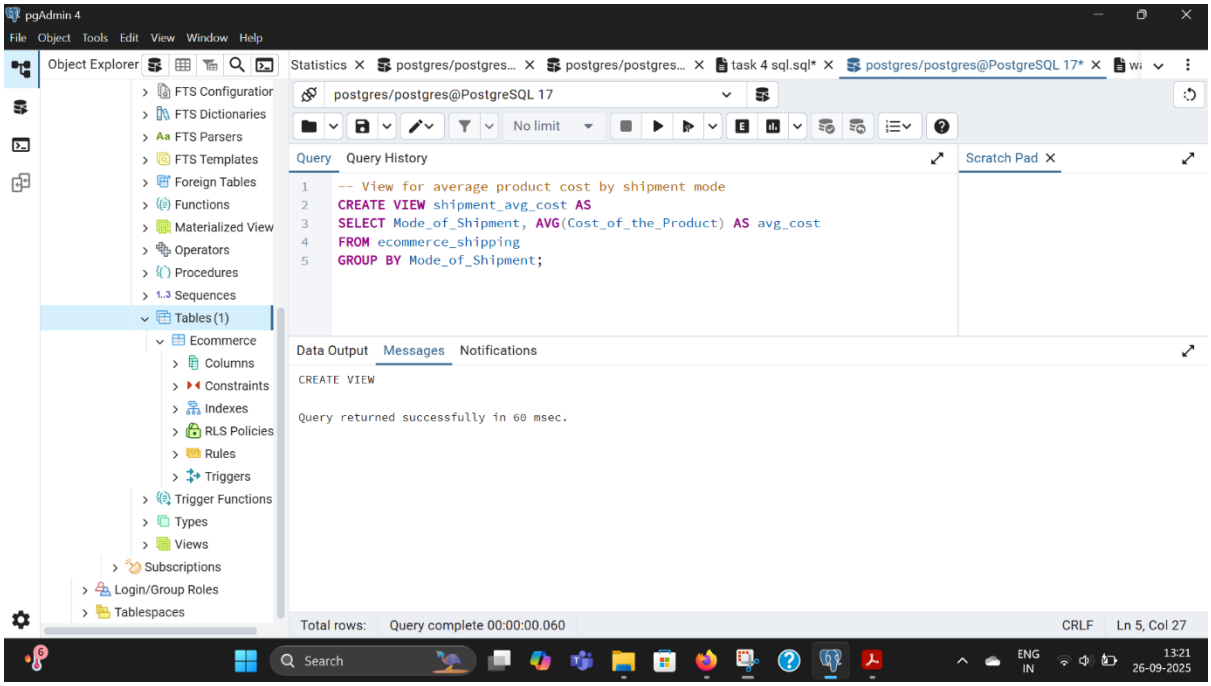
Indexes speed up queries that filter or group by these columns.

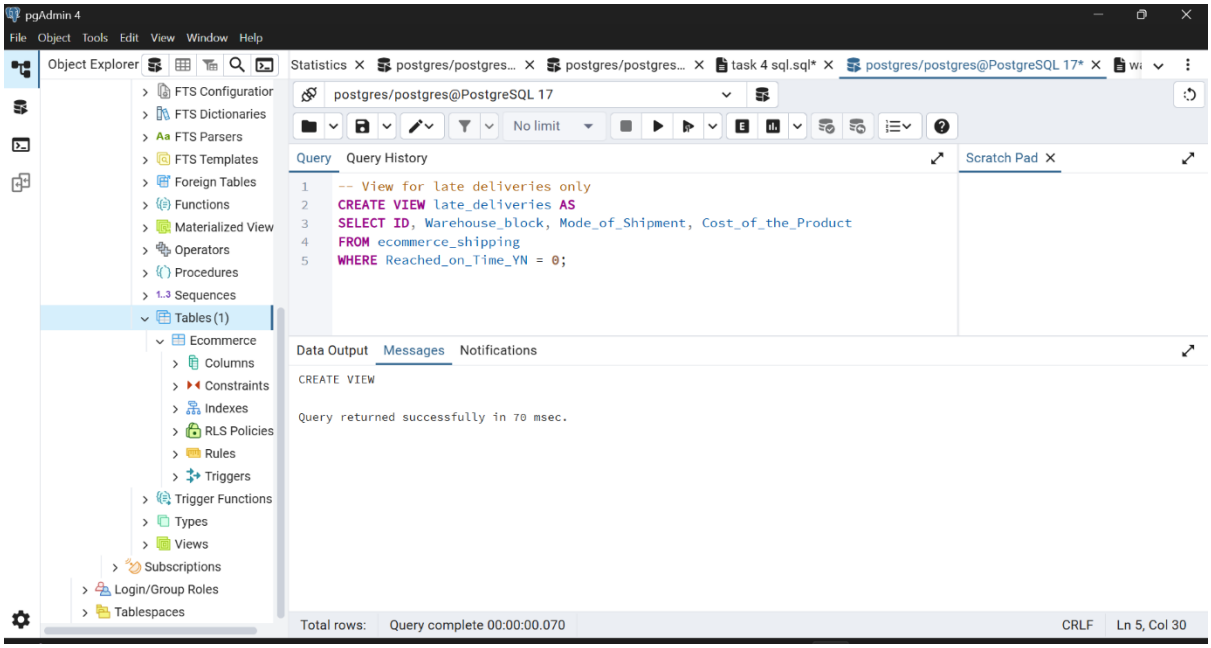
Summary of Work Done So Far

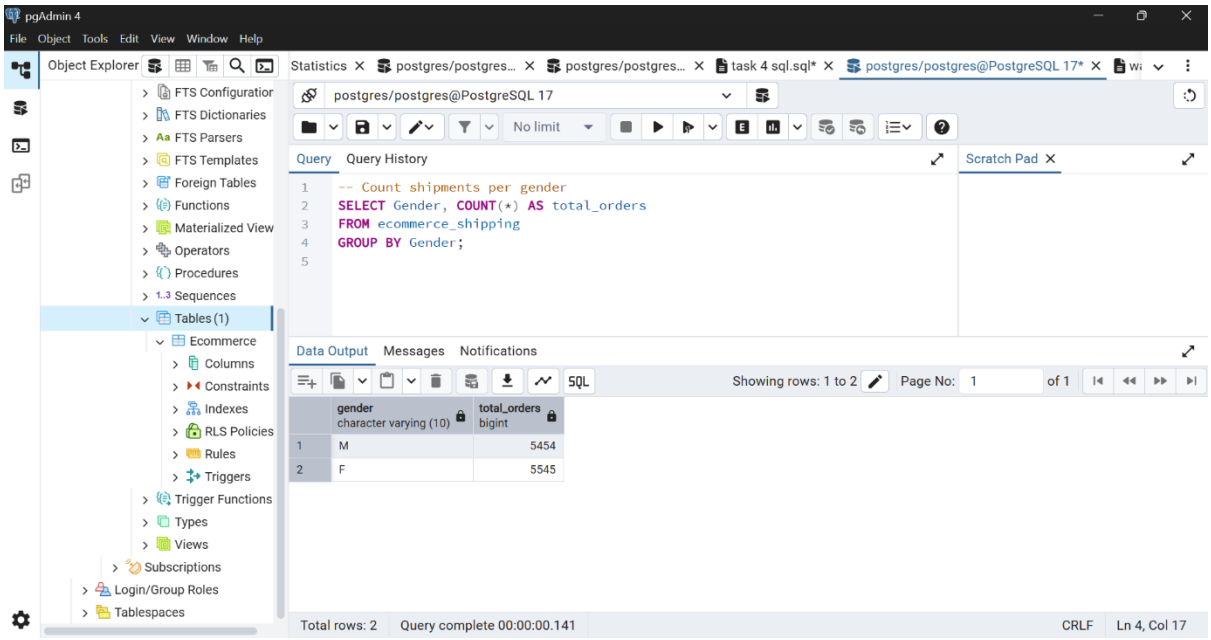
1. Created a table for the dataset.
2. Wrote queries using SELECT, WHERE, ORDER BY, and GROUP BY.
3. Used JOINS to combine shipment data with warehouse details.
4. Applied subqueries to find maximum values.
5. Implemented aggregate functions (SUM, AVG, COUNT).
6. Built views for reusable analysis.
7. Added indexes to improve query performance.

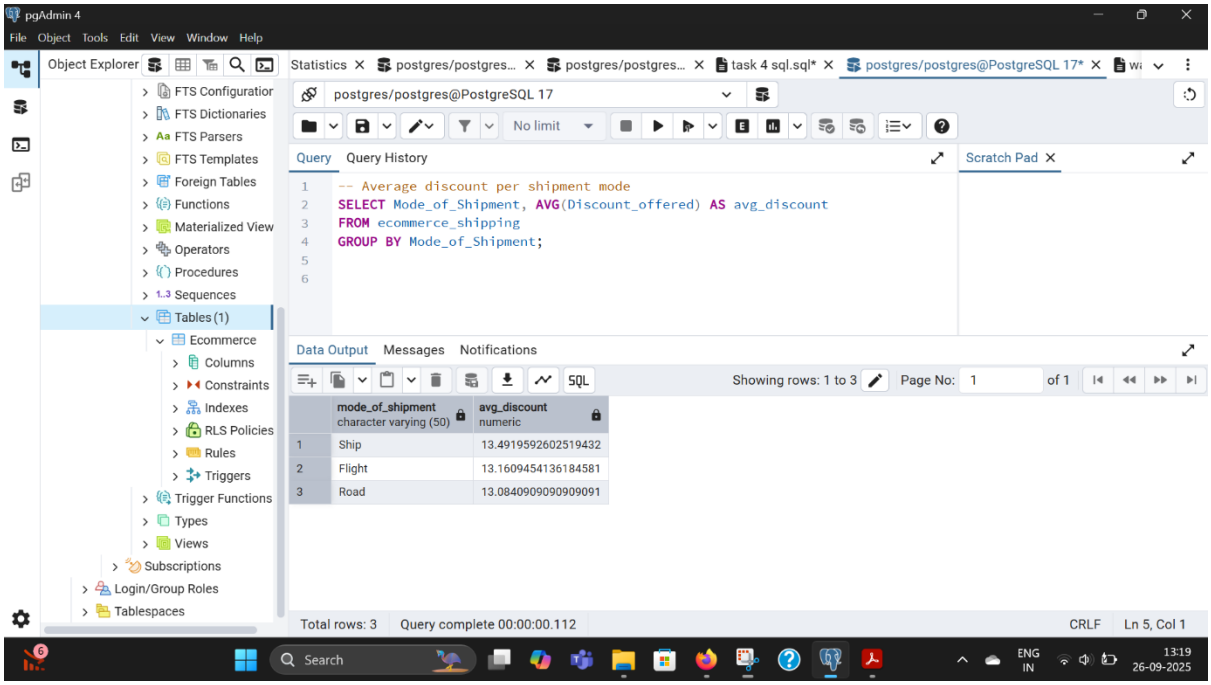
OUTPUT:

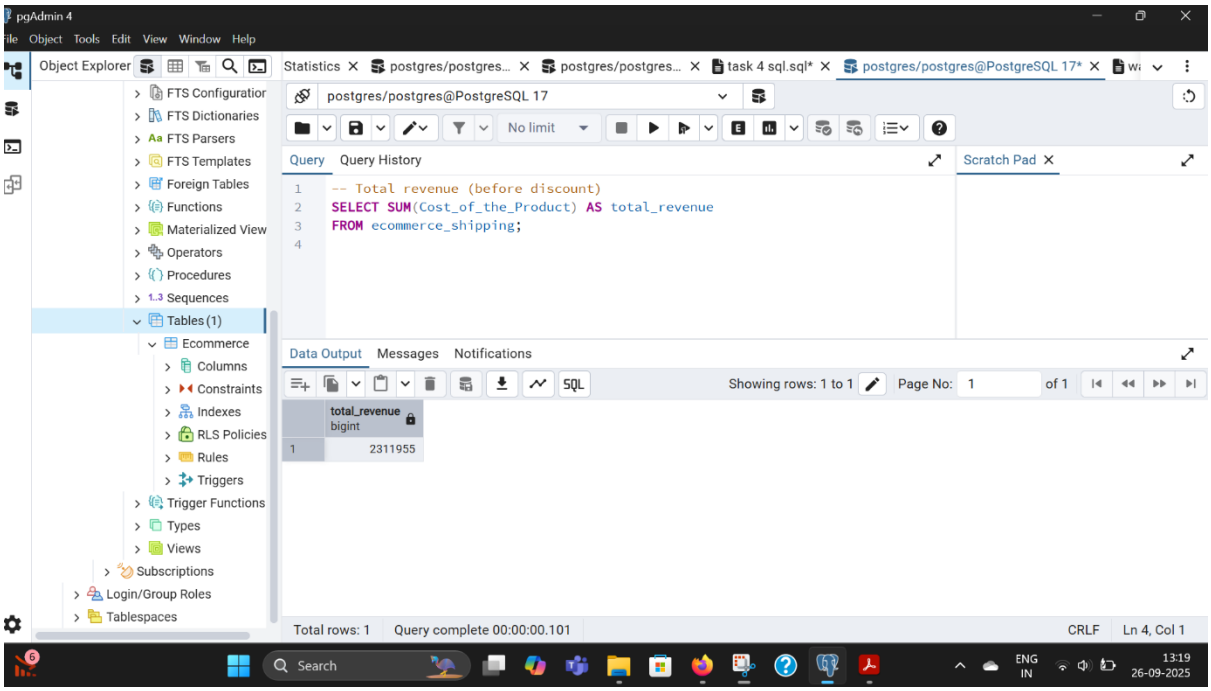

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Index on Mode_of_Shipment for GROUP BY performance
CREATE INDEX idx_mode_shipment ON ecommerce_shipping(Mode_of_Shipment);
Data Output
Messages
Notifications
CREATE INDEX
Query returned successfully in 58 msec.
Total rows:
Query complete 00:00:00.058
CRLF
Ln 1, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:24
26-09-2025

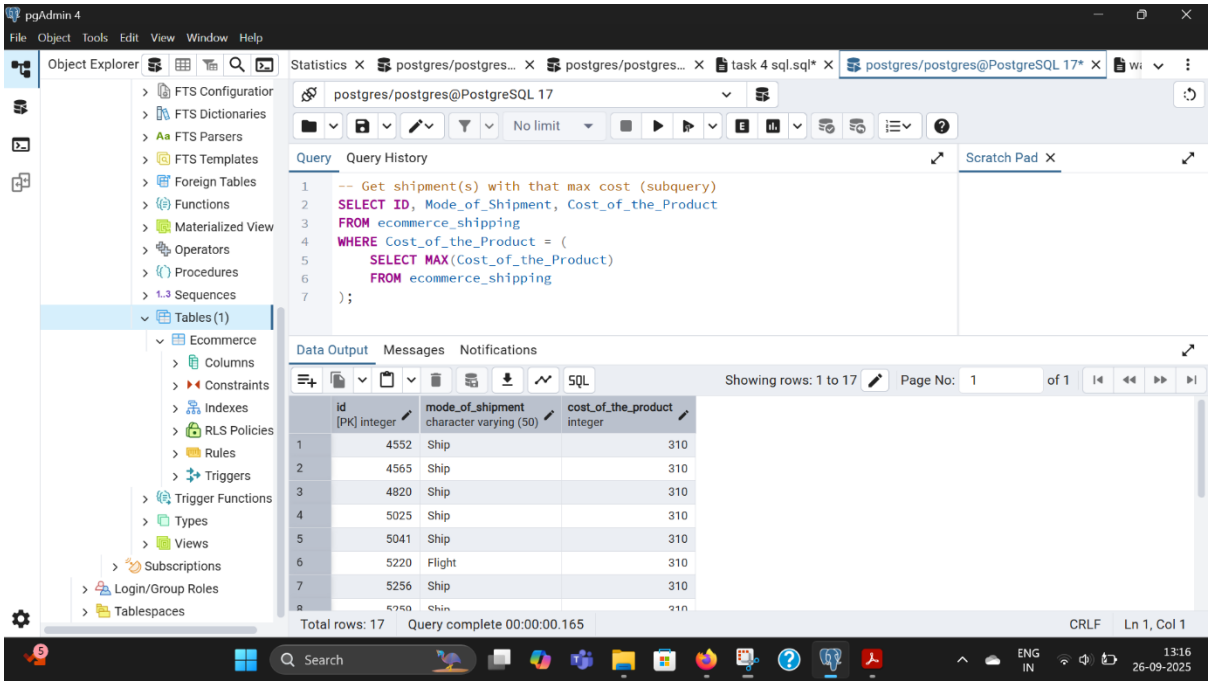

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- View for average product cost by shipment mode
CREATE VIEW shipment_avg_cost AS
SELECT Mode_of_Shipment, AVG(Cost_of_the_Product) AS avg_cost
FROM ecommerce_shipping
GROUP BY Mode_of_Shipment;
Data Output
Messages
Notifications
CREATE VIEW
Query returned successfully in 60 msec.
Total rows:
Query complete 00:00:00.060
CRLF
Ln 5, Col 27
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:21
26-09-2025

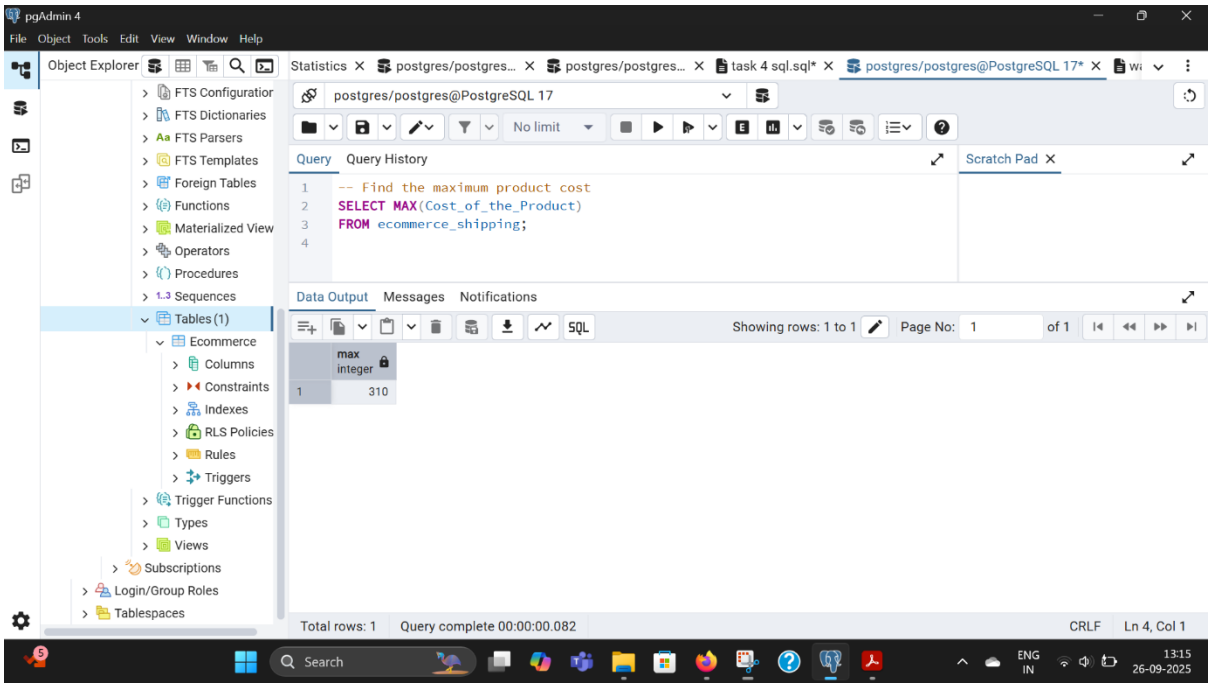

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- View for late deliveries only
CREATE VIEW late_deliveries AS
SELECT ID, Warehouse_block, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping
WHERE Reached_on_Time_YN = 0;
Data Output
Messages
Notifications
CREATE VIEW
Query returned successfully in 70 msec.
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Total rows:
Query complete 00:00:00.070
CRLF
Ln 5, Col 30

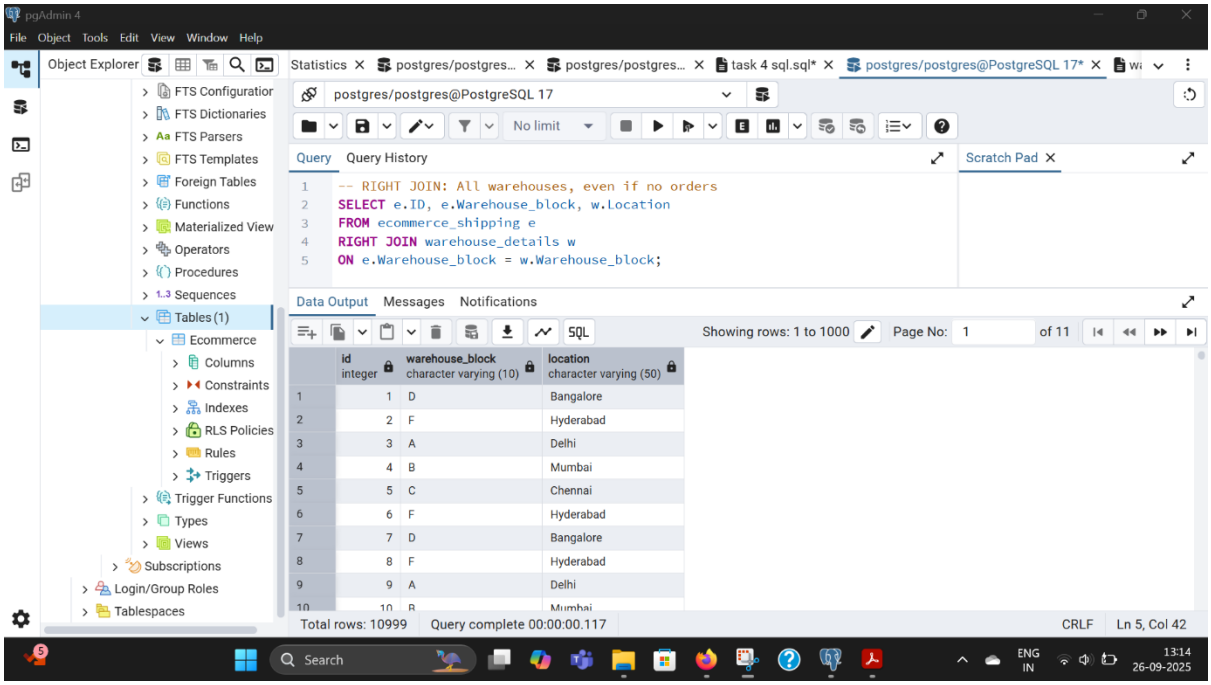

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Count shipments per gender
SELECT Gender, COUNT(*) AS total_orders
FROM ecommerce_shipping
GROUP BY Gender;
Data Output
Messages
Notifications
SQL
Showing rows: 1 to 2
Page No: 1
of 1
gender
character varying (10)
total_orders
bigint
1 M 5454
2 F 5545
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Total rows: 2
Query complete 00:00:00.141
CRLF
Ln 4, Col 17

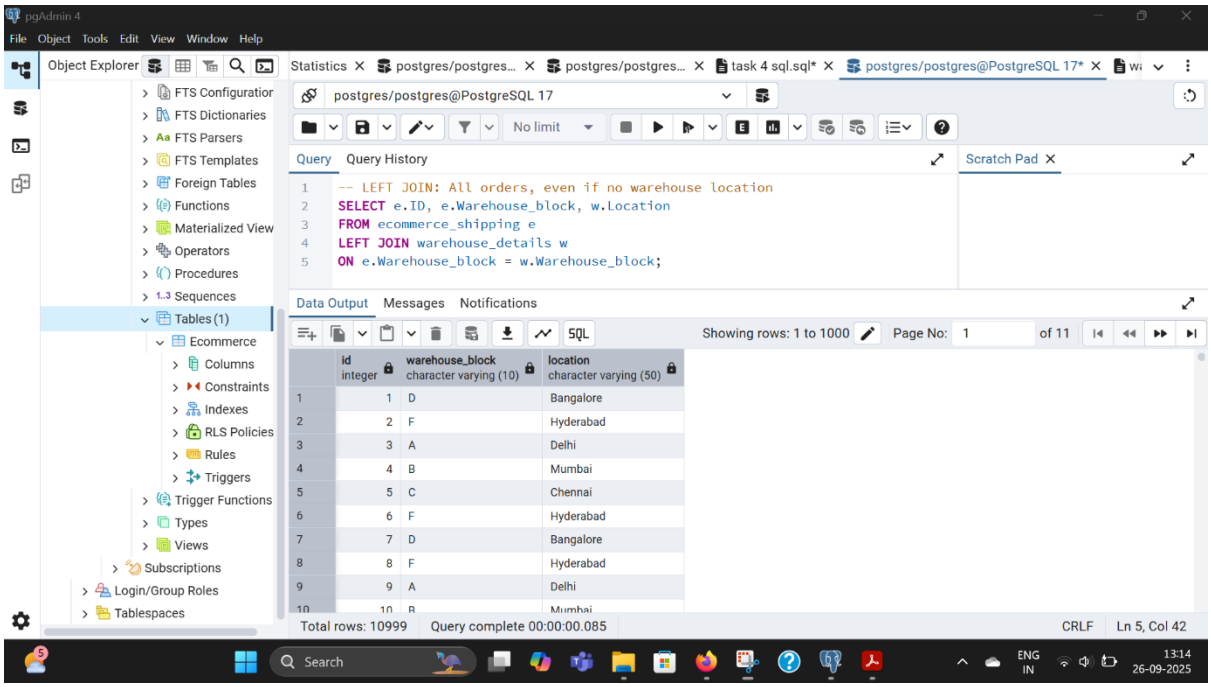

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Average discount per shipment mode
SELECT Mode_of_Shipment, AVG(Discount_offered) AS avg_discount
FROM ecommerce_shipping
GROUP BY Mode_of_Shipment;
Data Output
Messages
Notifications
SQL
Showing rows: 1 to 3
Page No: 1
of 1
mode_of_shipment character varying (50)
avg_discount numeric
1 Ship 13.4919592602519432
2 Flight 13.1609454136184581
3 Road 13.0840909090909091
Total rows: 3
Query complete 00:00:00.112
CRLF
Ln 5, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:19
26-09-2025

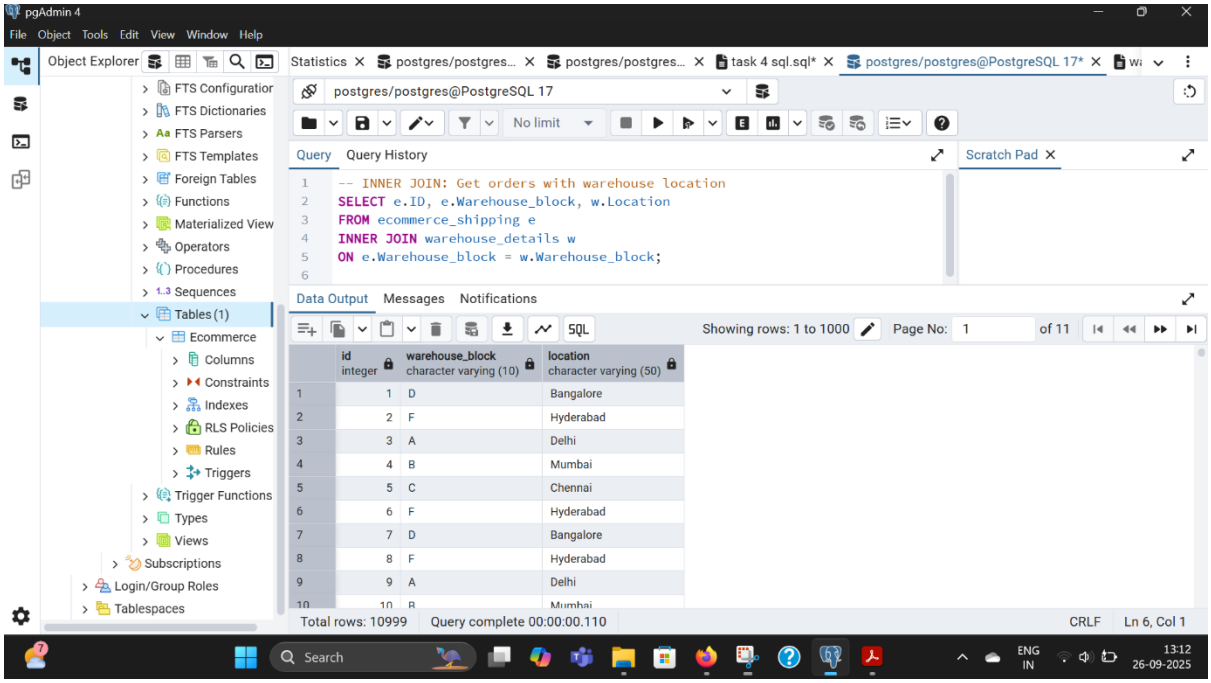

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Total revenue (before discount)
SELECT SUM(Cost_of_the_Product) AS total_revenue
FROM ecommerce_shipping;
Data Output
Messages
Notifications
SQL
Showing rows: 1 to 1
Page No: 1
of 1
total_revenue bigint
1 2311955
Total rows: 1
Query complete 00:00:00.101
CRLF
Ln 4, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:19
26-09-2025

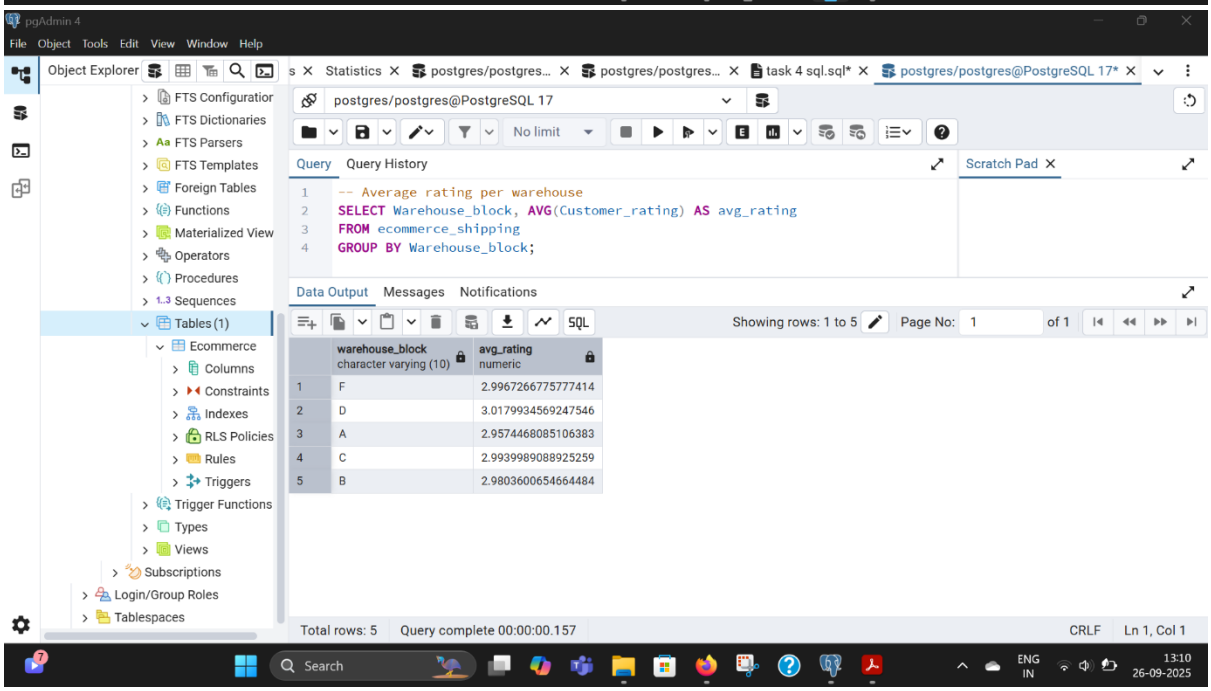

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Get shipment(s) with that max cost (subquery)
SELECT ID, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping
WHERE Cost_of_the_Product = (
SELECT MAX(Cost_of_the_Product)
FROM ecommerce_shipping
);
Data Output
Messages
Notifications
Showing rows: 1 to 17
Page No: 1 of 1
id [PK] integer
mode_of_shipment character varying (50)
cost_of_the_product integer
1 4552 Ship 310
2 4565 Ship 310
3 4820 Ship 310
4 5025 Ship 310
5 5041 Ship 310
6 5220 Flight 310
7 5256 Ship 310
Total rows: 17
Query complete 00:00:00.165
CRLF
Ln 1, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:16
26-09-2025

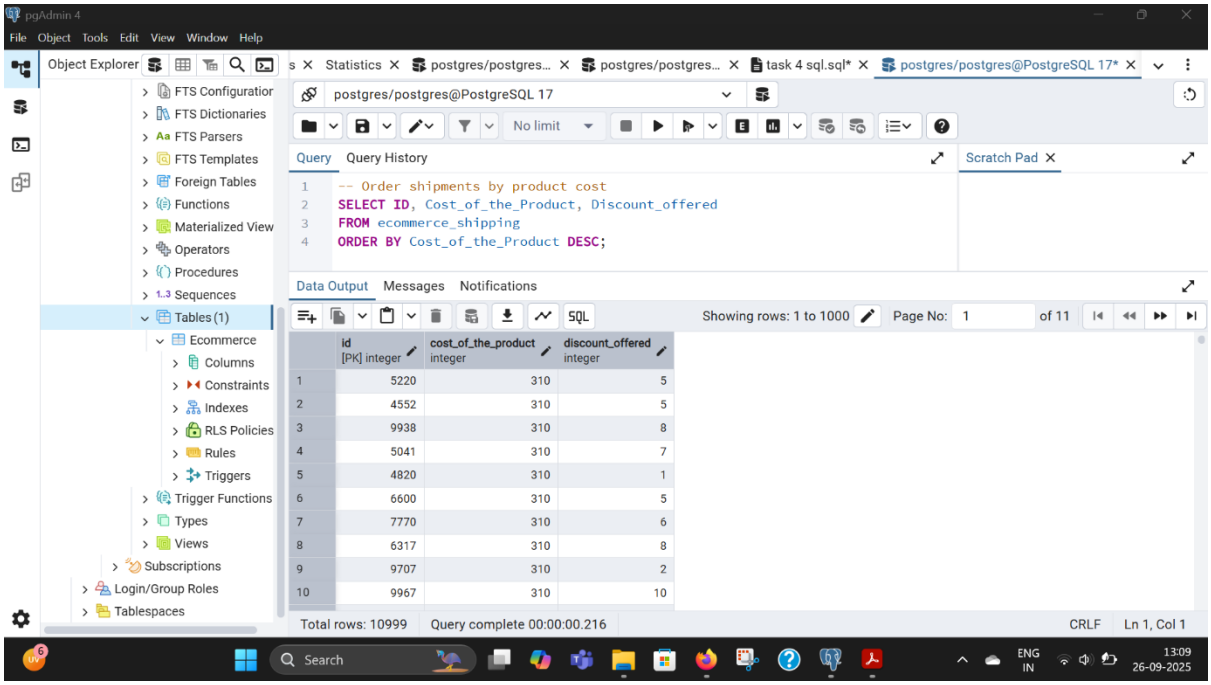

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Find the maximum product cost
SELECT MAX(Cost_of_the_Product)
FROM ecommerce_shipping;
Data Output
Messages
Notifications
Showing rows: 1 to 1
Page No: 1 of 1
max integer
1 310
Total rows: 1
Query complete 00:00:00.082
CRLF
Ln 4, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:15
26-09-2025

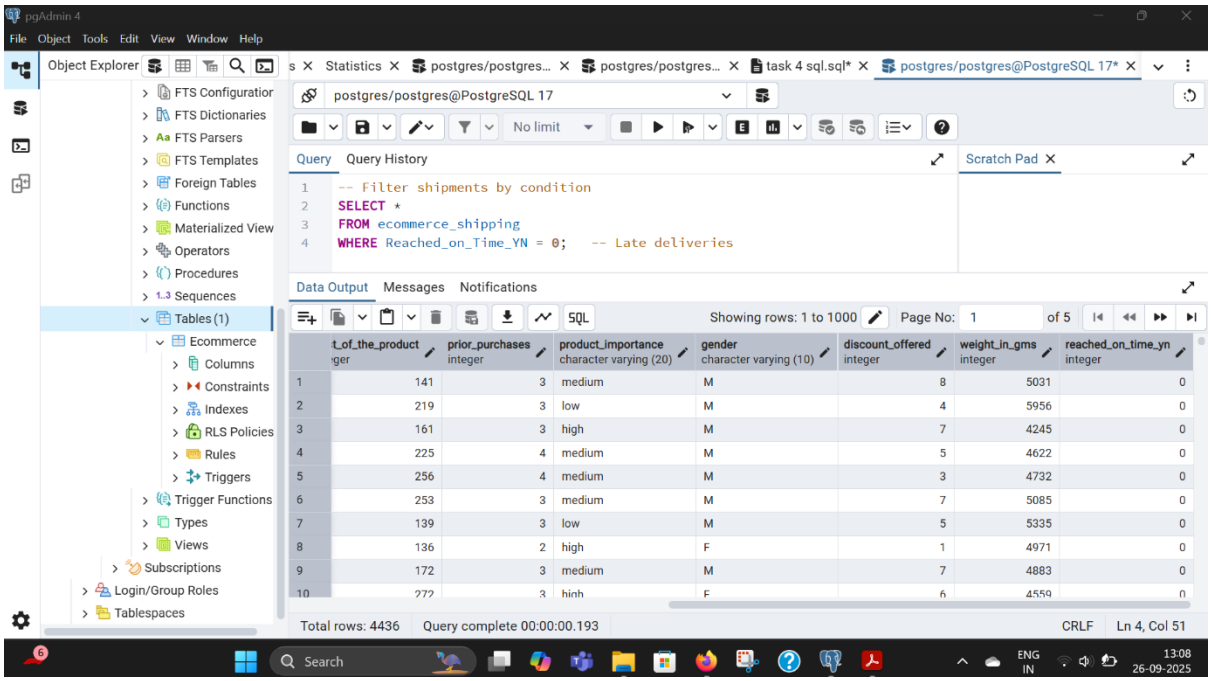

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- RIGHT JOIN: All warehouses, even if no orders
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
RIGHT JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
Data Output
Messages
Notifications
Showing rows: 1 to 1000
Page No: 1 of 11
id integer
warehouse_block character varying (10)
location character varying (50)
1 1 D Bangalore
2 2 F Hyderabad
3 3 A Delhi
4 4 B Mumbai
5 5 C Chennai
6 6 F Hyderabad
7 7 D Bangalore
8 8 F Hyderabad
9 9 A Delhi
10 10 B Mumbai
Total rows: 10999
Query complete 00:00:00.117
CRLF
Ln 5, Col 42
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:14
26-09-2025

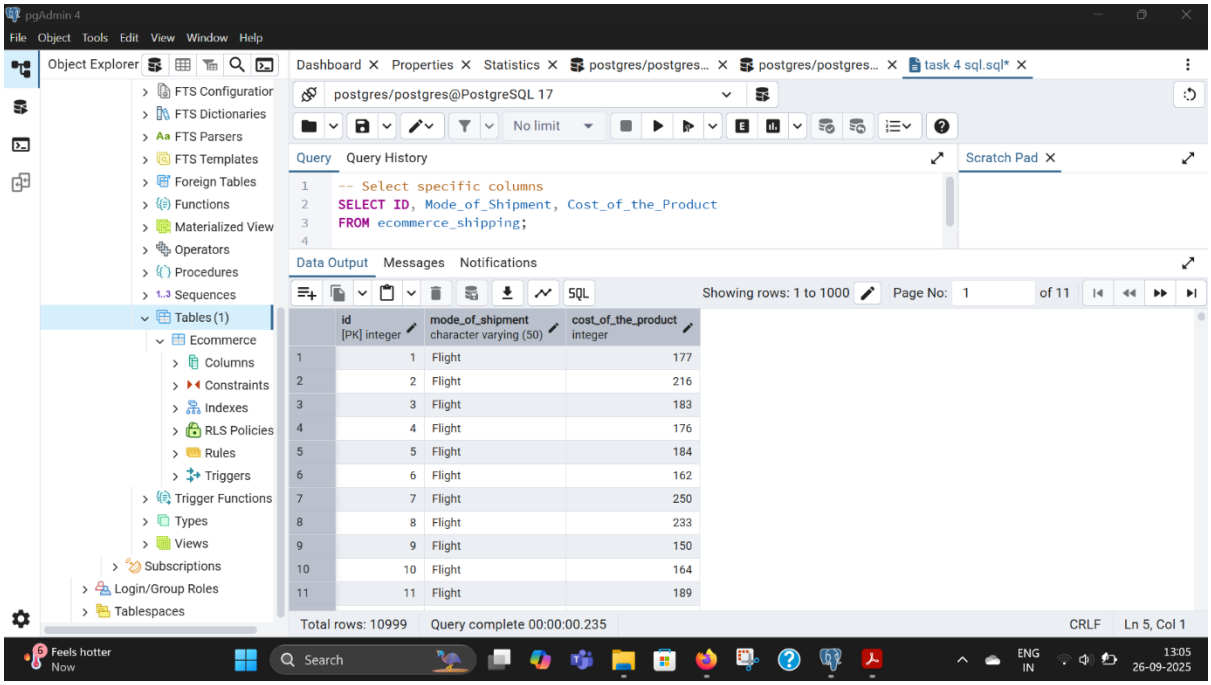

pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- LEFT JOIN: All orders, even if no warehouse location
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
LEFT JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
Data Output
Messages
Notifications
Showing rows: 1 to 1000
Page No: 1 of 11
id integer
warehouse_block character varying (10)
location character varying (50)
1 1 D Bangalore
2 2 F Hyderabad
3 3 A Delhi
4 4 B Mumbai
5 5 C Chennai
6 6 F Hyderabad
7 7 D Bangalore
8 8 F Hyderabad
9 9 A Delhi
10 10 B Mumbai
Total rows: 10999
Query complete 00:00:00.085
CRLF
Ln 5, Col 42
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:14
26-09-2025


pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- INNER JOIN: Get orders with warehouse location
SELECT e.ID, e.Warehouse_block, w.Location
FROM ecommerce_shipping e
INNER JOIN warehouse_details w
ON e.Warehouse_block = w.Warehouse_block;
Data Output
Messages
Notifications
Showing rows: 1 to 1000
Page No: 1 of 11
id integer
warehouse_block character varying (10)
location character varying (50)
1 1 D Bangalore
2 2 F Hyderabad
3 3 A Delhi
4 4 B Mumbai
5 5 C Chennai
6 6 F Hyderabad
7 7 D Bangalore
8 8 F Hyderabad
9 9 A Delhi
10 10 B Mumbai
Total rows: 10999
Query complete 00:00:00.110
CRLF
Ln 6, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:12
26-09-2025


pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
Statistics
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Average rating per warehouse
SELECT Warehouse_block, AVG(Customer_rating) AS avg_rating
FROM ecommerce_shipping
GROUP BY Warehouse_block;
Data Output
Messages
Notifications
Showing rows: 1 to 5
Page No: 1 of 1
warehouse_block character varying (10)
avg_rating numeric
1 F 2.9967266775777414
2 D 3.0179934569247546
3 A 2.9574468085106383
4 C 2.9939989088925259
5 B 2.9803600654664484
Total rows: 5
Query complete 00:00:00.157
CRLF
Ln 1, Col 1
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Search
ENG
IN
13:10
26-09-2025


pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
postgres/postgres@PostgreSQL 17
No limit
Query Query History
Scratch Pad
-- Order shipments by product cost
SELECT ID, Cost_of_the_Product, Discount_offered
FROM ecommerce_shipping
ORDER BY Cost_of_the_Product DESC;
Data Output Messages Notifications
Showing rows: 1 to 1000 Page No: 1 of 11
id [PK] integer
cost_of_the_product integer
discount_offered integer
1 5220 310 5
2 4552 310 5
3 9938 310 8
4 5041 310 7
5 4820 310 1
6 6600 310 5
7 7770 310 6
8 6317 310 8
9 9707 310 2
10 9967 310 10
Total rows: 10999
Query complete 00:00:00.216
CRLF
Ln 1, Col 1
Search
ENG IN
13:09 26-09-2025


pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
postgres/postgres@PostgreSQL 17
No limit
Query Query History
Scratch Pad
-- Filter shipments by condition
SELECT *
FROM ecommerce_shipping
WHERE Reached_on_Time_YN = 0; -- Late deliveries
Data Output Messages Notifications
Showing rows: 1 to 1000 Page No: 1 of 5
prior_purchases integer
product_importance character varying (20)
gender character varying (10)
discount_offered integer
weight_in_gms integer
reached_on_time_yn integer
1 141 3 medium M 8 5031 0
2 219 3 low M 4 5956 0
3 161 3 high M 7 4245 0
4 225 4 medium M 5 4622 0
5 256 4 medium M 3 4732 0
6 253 3 medium M 7 5085 0
7 139 3 low M 5 5335 0
8 136 2 high F 1 4971 0
9 172 3 medium M 7 4883 0
Total rows: 4436
Query complete 00:00:00.193
CRLF
Ln 4, Col 51
Search
ENG IN
13:08 26-09-2025


pgAdmin 4
File Object Tools Edit View Window Help
Object Explorer
FTS Configuration
FTS Dictionaries
FTS Parsers
FTS Templates
Foreign Tables
Functions
Materialized View
Operators
Procedures
Sequences
Tables (1)
Ecommerce
Columns
Constraints
Indexes
RLS Policies
Rules
Triggers
Trigger Functions
Types
Views
Subscriptions
Login/Group Roles
Tablespaces
Dashboard
Properties
Statistics
postgres/postgres...
postgres/postgres...
task 4 sql.sql*
postgres/postgres@PostgreSQL 17
No limit
Query
Query History
Scratch Pad
-- Select specific columns
SELECT ID, Mode_of_Shipment, Cost_of_the_Product
FROM ecommerce_shipping;
Data Output
Messages
Notifications
SQL
Showing rows: 1 to 1000
Page No: 1
of 11
id [PK] integer
mode_of_shipment character varying (50)
cost_of_the_product integer
1 1 Flight 177
2 2 Flight 216
3 3 Flight 183
4 4 Flight 176
5 5 Flight 184
6 6 Flight 162
7 7 Flight 250
8 8 Flight 233
9 9 Flight 150
10 10 Flight 164
11 11 Flight 189
Total rows: 10999
Query complete 00:00:00.235
CRLF
Ln 5, Col 1
Feels hotter Now
Search
ENG IN
13:05 26-09-2025