Screenshots of SQL queries in Databricks Notebook—

query-8-running-total.png

query-9-top-3-per-category.png

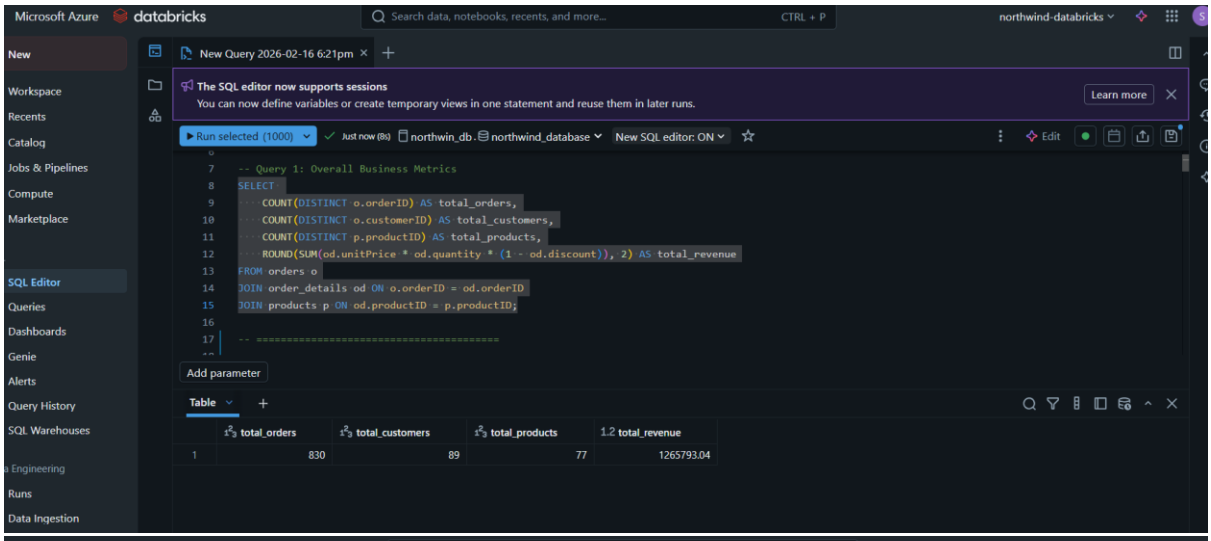
query-10-customer-ltv.png

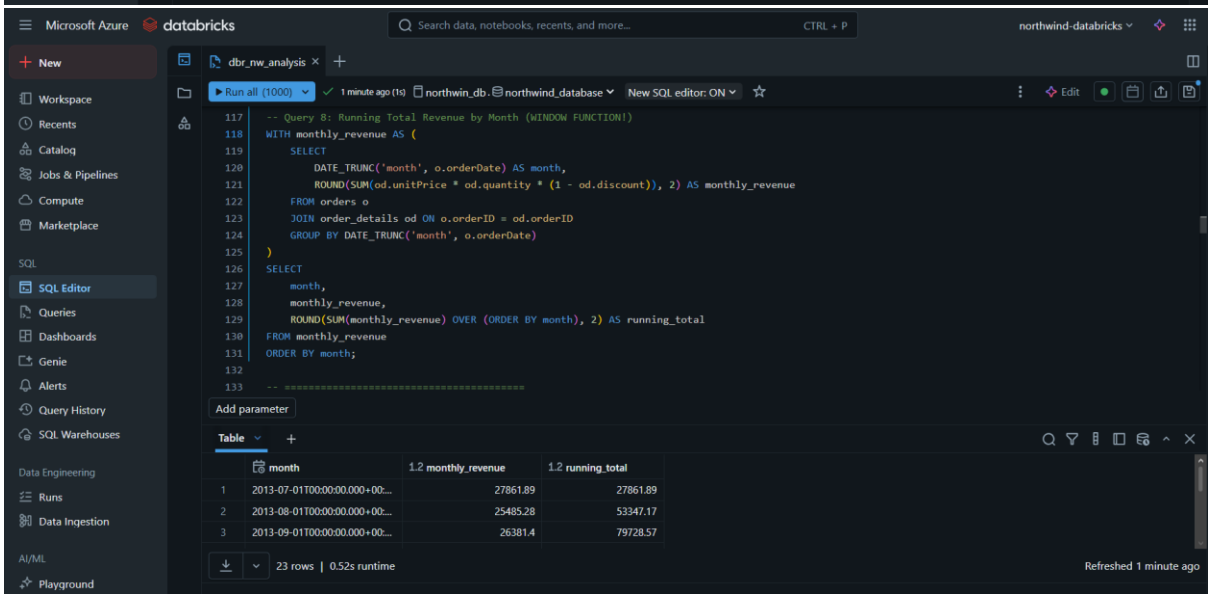
query-11-yoy-growth.png

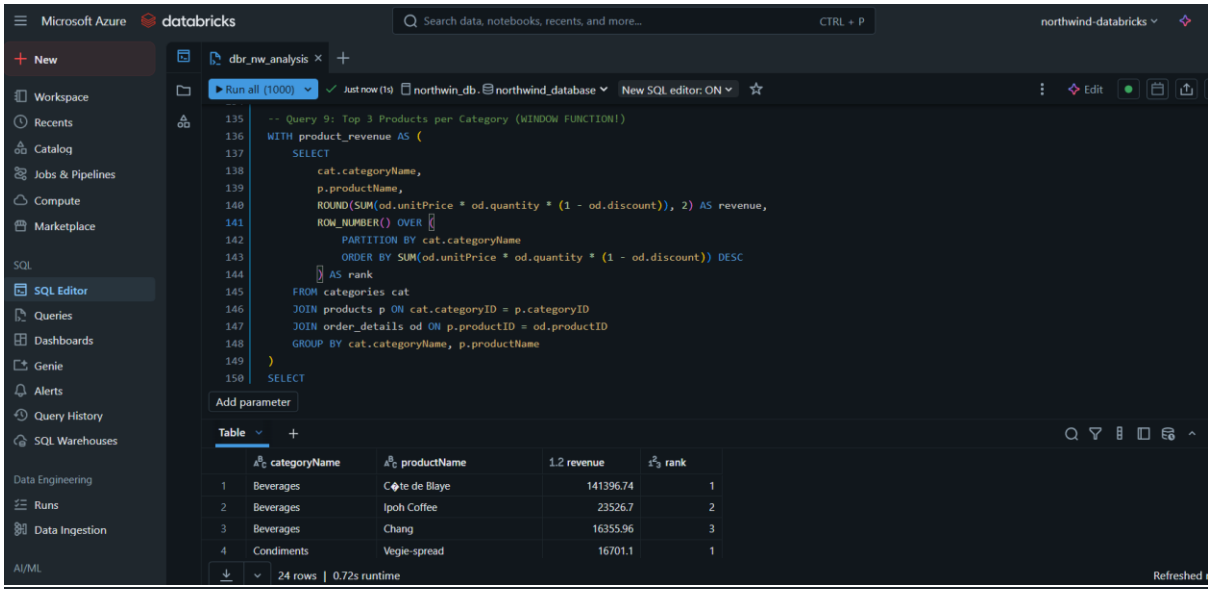
query-12-cohort.png

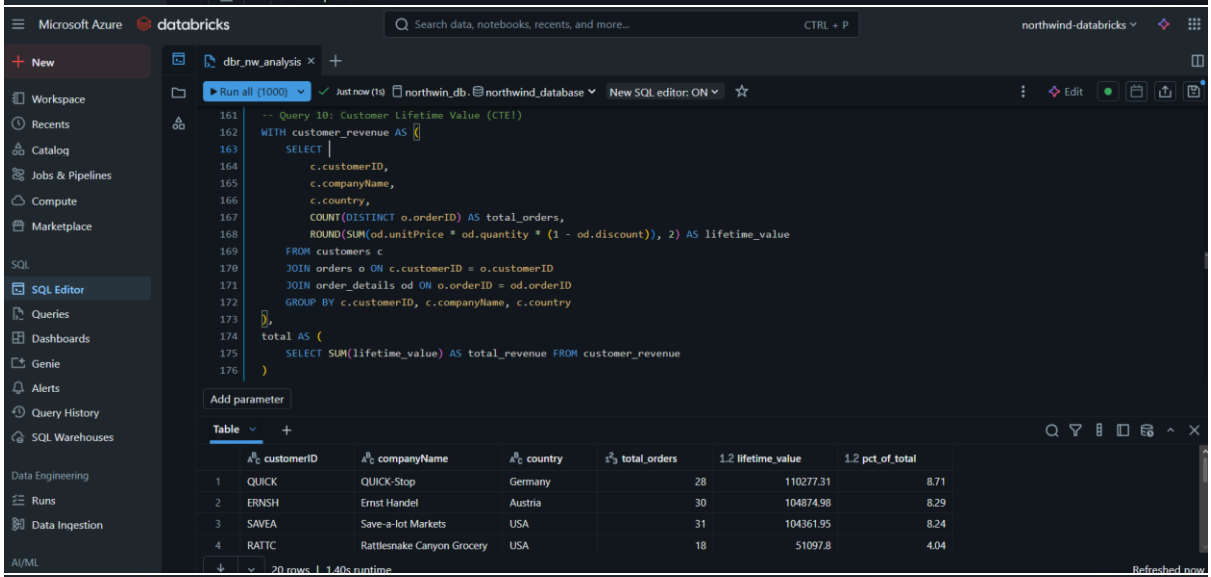
query-13-shipping.png

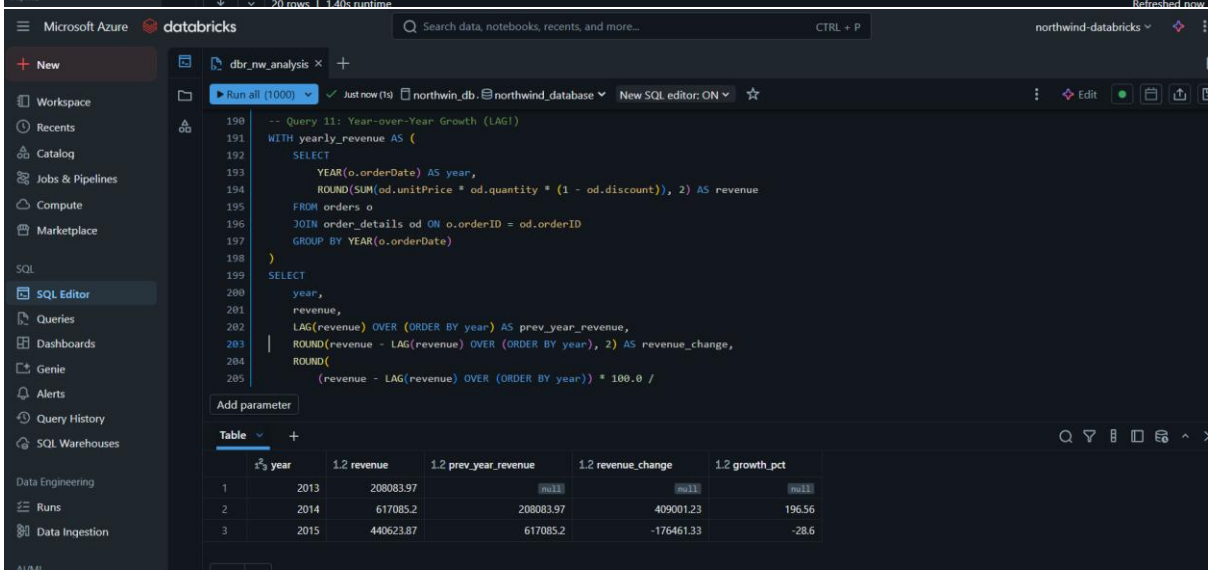
query-14-market.png

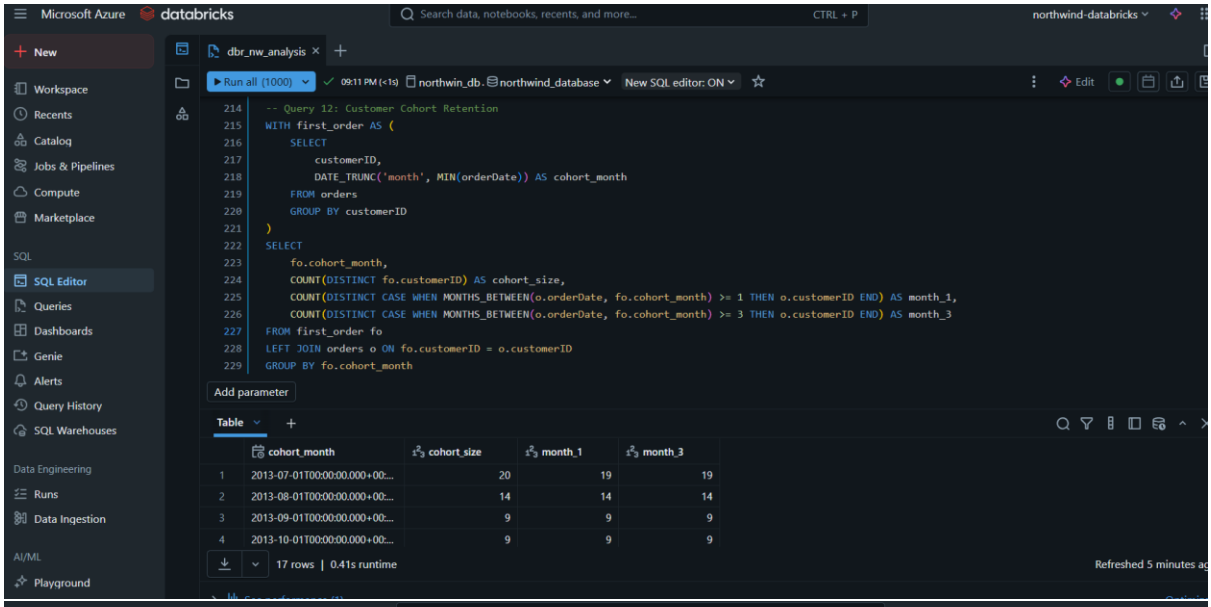


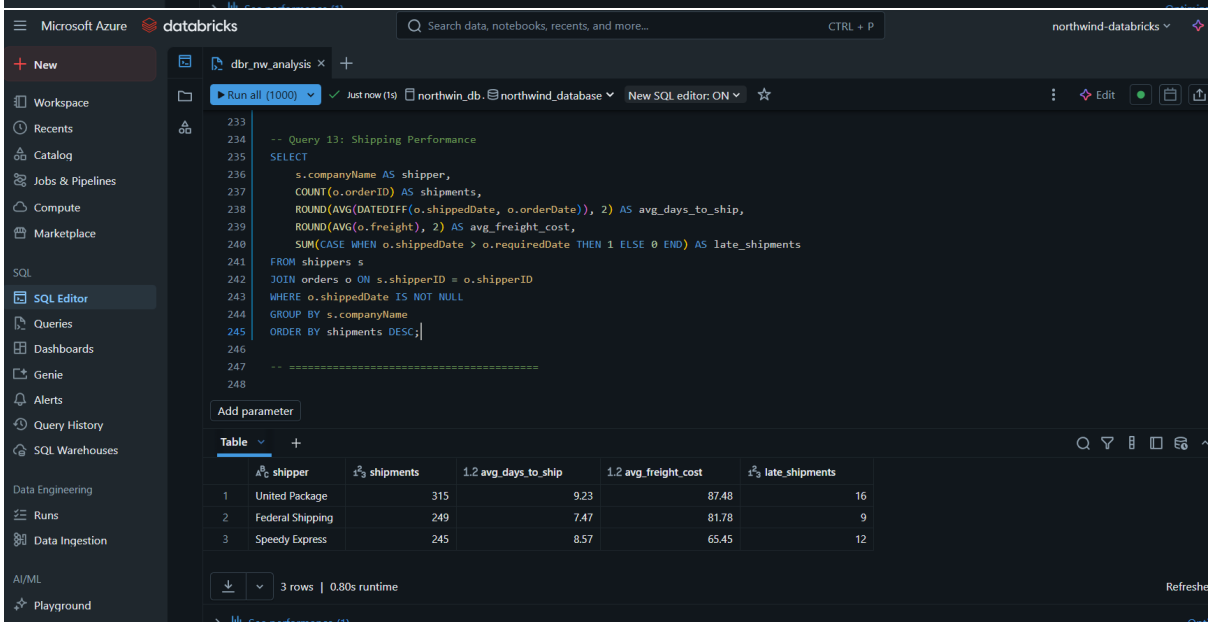


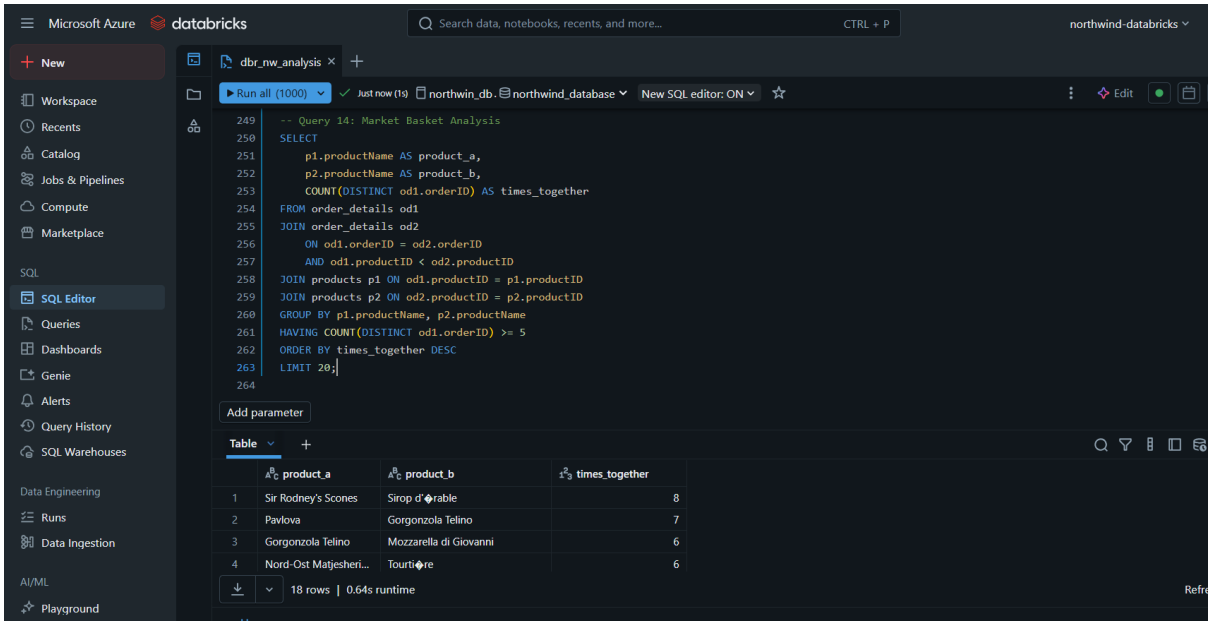

Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
dbr_nw_analysis
Run all (1000)
Just now (1s)
northwin_db
northwind_database
New SQL editor: ON
Edit
-- Query 9: Top 3 Products per Category (WINDOW FUNCTION!)
WITH product_revenue AS (
SELECT
cat.categoryName,
p.productName,
ROUND(SUM(od.unitPrice * od.quantity * (1 - od.discount)), 2) AS revenue,
ROW_NUMBER() OVER (
PARTITION BY cat.categoryName
ORDER BY SUM(od.unitPrice * od.quantity * (1 - od.discount)) DESC
) AS rank
FROM categories cat
JOIN products p ON cat.categoryID = p.categoryID
JOIN order_details od ON p.productID = od.productID
GROUP BY cat.categoryName, p.productName
)
SELECT
Add parameter
Table
categoryName
productName
revenue
rank
1 Beverages Côte de Blaye 141396.74 1
2 Beverages Ipoh Coffee 23526.7 2
3 Beverages Chang 16355.96 3
4 Condiments Vegie-spread 16701.1 1
24 rows | 0.72s runtime
Refreshed

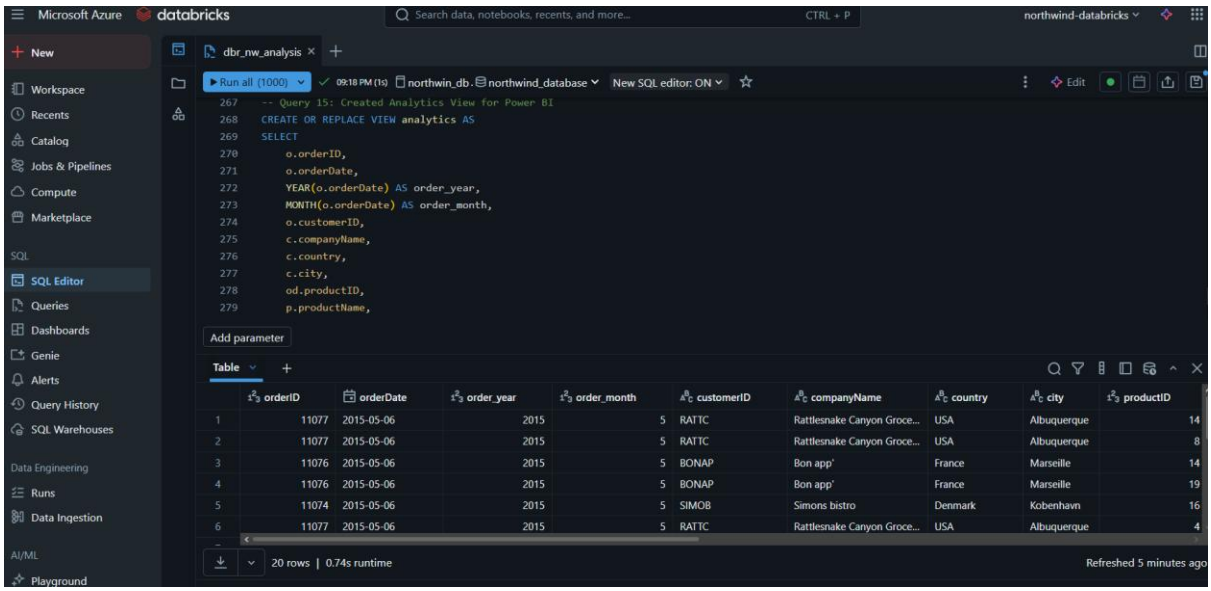

Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
dbr_nw_analysis
Run all (1000)
Just now (1s)
northwin_db
northwind_database
New SQL editor: ON
Edit
-- Query 10: Customer Lifetime Value (CTE!)
WITH customer_revenue AS (
SELECT
c.customerID,
c.companyName,
c.country,
COUNT(DISTINCT o.orderID) AS total_orders,
ROUND(SUM(od.unitPrice * od.quantity * (1 - od.discount)), 2) AS lifetime_value
FROM customers c
JOIN orders o ON c.customerID = o.customerID
JOIN order_details od ON o.orderID = od.orderID
GROUP BY c.customerID, c.companyName, c.country
),
total AS (
SELECT SUM(lifetime_value) AS total_revenue FROM customer_revenue
)
Add parameter
Table
customerID
companyName
country
total_orders
lifetime_value
pct_of_total
1 QUICK QUICK-Stop Germany 28 110277.31 8.71
2 ERNSH Ernst Handel Austria 30 104874.98 8.29
3 SAVEA Save-a-lot Markets USA 31 104361.95 8.24
4 RATTC Rattlesnake Canyon Grocery USA 18 51097.8 4.04
20 rows | 1.40s runtime
Refreshed now


Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
dbr_nw_analysis
Run all (1000)
Just now (1s)
northwin_db
northwind_database
New SQL editor: ON
Edit
-- Query 11: Year-over-Year Growth (LAG!)
WITH yearly_revenue AS (
SELECT
YEAR(o.orderDate) AS year,
ROUND(SUM(od.unitPrice * od.quantity * (1 - od.discount)), 2) AS revenue
FROM orders o
JOIN order_details od ON o.orderID = od.orderID
GROUP BY YEAR(o.orderDate)
)
SELECT
year,
revenue,
LAG(revenue) OVER (ORDER BY year) AS prev_year_revenue,
ROUND(revenue - LAG(revenue) OVER (ORDER BY year), 2) AS revenue_change,
ROUND(
(revenue - LAG(revenue) OVER (ORDER BY year)) * 100.0 /
Add parameter
Table
year
revenue
prev_year_revenue
revenue_change
growth_pct
1 2013 208083.97 null null null
2 2014 617085.2 208083.97 409001.23 196.56
3 2015 440623.87 617085.2 -176461.33 -28.6


Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
Playground
dbr_nw_analysis
Run all (1000)
09:11 PM (<1s)
northwin_db.northwind_database
New SQL editor: ON
Edit
214 -- Query 12: Customer Cohort Retention
215 WITH first_order AS (
216 SELECT
217 customerID,
218 DATE_TRUNC('month', MIN(orderDate)) AS cohort_month
219 FROM orders
220 GROUP BY customerID
221)
222 SELECT
223 fo.cohort_month,
224 COUNT(DISTINCT fo.customerID) AS cohort_size,
225 COUNT(DISTINCT CASE WHEN MONTHS_BETWEEN(o.orderDate, fo.cohort_month) >= 1 THEN o.customerID END) AS month_1,
226 COUNT(DISTINCT CASE WHEN MONTHS_BETWEEN(o.orderDate, fo.cohort_month) >= 3 THEN o.customerID END) AS month_3
227 FROM first_order fo
228 LEFT JOIN orders o ON fo.customerID = o.customerID
229 GROUP BY fo.cohort_month
Add parameter
Table
cohort_month
cohort_size
month_1
month_3
1 2013-07-01T00:00:00.000+00:... 20 19 19
2 2013-08-01T00:00:00.000+00:... 14 14 14
3 2013-09-01T00:00:00.000+00:... 9 9 9
4 2013-10-01T00:00:00.000+00:... 9 9 9
17 rows | 0.41s runtime
Refreshed 5 minutes ag


Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
Playground
dbr_nw_analysis
Run all (1000)
Just now (1s)
northwin_db.northwind_database
New SQL editor: ON
Edit
233
234 -- Query 13: Shipping Performance
235 SELECT
236 s.companyName AS shipper,
237 COUNT(o.orderID) AS shipments,
238 ROUND(AVG(DATEDIFF(o.shippedDate, o.orderDate)), 2) AS avg_days_to_ship,
239 ROUND(AVG(o.freight), 2) AS avg_freight_cost,
240 SUM(CASE WHEN o.shippedDate > o.requiredDate THEN 1 ELSE 0 END) AS late_shipments
241 FROM shippers s
242 JOIN orders o ON s.shipperID = o.shipperID
243 WHERE o.shippedDate IS NOT NULL
244 GROUP BY s.companyName
245 ORDER BY shipments DESC;
246
247 -- ==
248
Add parameter
Table
shipper
shipments
avg_days_to_ship
avg_freight_cost
late_shipments
1 United Package 315 9.23 87.48 16
2 Federal Shipping 249 7.47 81.78 9
3 Speedy Express 245 8.57 65.45 12
3 rows | 0.80s runtime
Refreshe


Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
dbr_nw_analysis
Run all (1000)
Just now (1s)
northwin_db
northwind_database
New SQL editor: ON
Edit
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
Playground
249 -- Query 14: Market Basket Analysis
250 SELECT
251 p1.productName AS product_a,
252 p2.productName AS product_b,
253 COUNT(DISTINCT od1.orderID) AS times_together
254 FROM order_details od1
255 JOIN order_details od2
256 ON od1.orderID = od2.orderID
257 AND od1.productID < od2.productID
258 JOIN products p1 ON od1.productID = p1.productID
259 JOIN products p2 ON od2.productID = p2.productID
260 GROUP BY p1.productName, p2.productName
261 HAVING COUNT(DISTINCT od1.orderID) >= 5
262 ORDER BY times_together DESC
263 LIMIT 20;
264
Add parameter
Table
product_a
product_b
times_together
1 Sir Rodney's Scones Sirop d'�rable 8
2 Pavlova Gorgonzola Telino 7
3 Gorgonzola Telino Mozzarella di Giovanni 6
4 Nord-Ost Matjesheri... Tourti�re 6
18 rows | 0.64s runtime


Microsoft Azure
databricks
Search data, notebooks, recents, and more...
CTRL + P
northwind-databricks
New
dbr_nw_analysis
Run all (1000)
09:18 PM (1s)
northwin_db
northwind_database
New SQL editor: ON
Edit
Workspace
Recents
Catalog
Jobs & Pipelines
Compute
Marketplace
SQL
SQL Editor
Queries
Dashboards
Genie
Alerts
Query History
SQL Warehouses
Data Engineering
Runs
Data Ingestion
AI/ML
Playground
267 -- Query 15: Created Analytics View for Power BI
268 CREATE OR REPLACE VIEW analytics AS
269 SELECT
270 o.orderID,
271 o.orderDate,
272 YEAR(o.orderDate) AS order_year,
273 MONTH(o.orderDate) AS order_month,
274 o.customerID,
275 c.companyName,
276 c.country,
277 c.city,
278 od.productID,
279 p.productName,
Add parameter
Table
orderID orderDate order_year order_month customerID companyName country city productID
1 11077 2015-05-06 2015 5 RATTC Rattlesnake Canyon Groce... USA Albuquerque 14
2 11077 2015-05-06 2015 5 RATTC Rattlesnake Canyon Groce... USA Albuquerque 8
3 11076 2015-05-06 2015 5 BONAP Bon app' France Marseille 14
4 11076 2015-05-06 2015 5 BONAP Bon app' France Marseille 19
5 11074 2015-05-06 2015 5 SIMOB Simons bistro Denmark Kobenhavn 16
6 11077 2015-05-06 2015 5 RATTC Rattlesnake Canyon Groce... USA Albuquerque 4
20 rows | 0.74s runtime
Refreshed 5 minutes ago