**What is Autovacuum?**

In PostgreSQL (and Azure Cosmos DB for PostgreSQL), **Autovacuum** is a background process that:

* Reclaims storage by removing **dead tuples**
* Updates statistics for the query planner
* Prevents table and index bloat

In Cosmos DB for PostgreSQL, **Autovacuum runs automatically**, but you can **manually trigger** VACUUM or ANALYZE commands as well.

**🧪 DEMO SCENARIO OVERVIEW**

We'll:

1. Create a sample table
2. Insert and delete data
3. Observe autovacuum activity
4. Manually run VACUUM and ANALYZE
5. Monitor results

**✅ Prerequisites**

* Azure Cosmos DB for PostgreSQL cluster
* psql client or Azure Data Studio
* Superuser access (or an admin account for vacuum analysis)

**🧱 Step 1: Create a Table**

CREATE TABLE employees (

id SERIAL PRIMARY KEY,

name TEXT,

department TEXT,

salary NUMERIC

);

**➕ Step 2: Insert Sample Data**

INSERT INTO employees (name, department, salary)

SELECT

'Emp\_' || i,

'Dept\_' || (i % 5),

50000 + (random() \* 10000)::int

FROM generate\_series(1, 10000) AS s(i);

**❌ Step 3: Delete Some Rows**

DELETE FROM employees

WHERE id <= 5000;

Now the table has **"dead tuples"**, which autovacuum will eventually clean.

**🔍 Step 4: Check Autovacuum Settings**

SHOW autovacuum;

SHOW autovacuum\_vacuum\_threshold;

SHOW autovacuum\_vacuum\_scale\_factor;

These control **when** autovacuum triggers.

**🧼 Step 5: Manually Trigger Vacuum**

**🔸 Option 1: Basic VACUUM**

VACUUM employees;

**🔸 Option 2: Full Vacuum (Locks Table!)**

VACUUM FULL employees;

**🔸 Option 3: Analyze Statistics**

ANALYZE employees;

You can combine:

VACUUM ANALYZE employees;

**🧠 Step 6: Monitor Table Statistics**

To check table bloat and vacuum stats:

SELECT relname AS table\_name,

n\_live\_tup AS live\_rows,

n\_dead\_tup AS dead\_rows,

last\_vacuum,

last\_autovacuum,

last\_analyze,

last\_autoanalyze

FROM pg\_stat\_user\_tables

WHERE relname = 'employees';

**⚙️ Step 7: Enable Logging (Optional – requires admin rights)**

Enable logging to see autovacuum actions:

SET log\_autovacuum\_min\_duration = 0; -- Logs all autovacuums

Then perform some deletes and observe logs.

**🧾 Best Practices**

| **Area** | **Recommendation** |
| --- | --- |
| Autovacuum thresholds | Tune autovacuum\_vacuum\_threshold and autovacuum\_vacuum\_scale\_factor for high-churn tables |
| Regular cleanup | Use VACUUM ANALYZE during low-usage periods if autovacuum is not fast enough |
| Avoid full vacuum | Only use VACUUM FULL when table bloat is significant and you can afford the downtime |

**Step-by-Step: Check if Table Needs Vacuum**

**🔍 1. Use pg\_stat\_user\_tables to Inspect Table Health**

Run this SQL query:

SELECT

relname AS table\_name,

n\_live\_tup AS live\_rows,

n\_dead\_tup AS dead\_rows,

last\_vacuum,

last\_autovacuum,

vacuum\_count,

autovacuum\_count

FROM

pg\_stat\_user\_tables

WHERE

relname = 'your\_table\_name'; -- Replace with actual table name

**What to Look For:**

* n\_dead\_tup: High number of dead tuples → **table likely needs vacuum**
* last\_autovacuum: If it’s NULL or a long time ago → vacuum may be overdue
* autovacuum\_count: Low value over time = autovacuum may not be running enough

**📏 2. Rule of Thumb: Dead Tuples Threshold**

Autovacuum runs when:

n\_dead\_tup > autovacuum\_vacuum\_threshold + autovacuum\_vacuum\_scale\_factor \* n\_live\_tup

Default values (unless customized):

| **Setting** | **Default Value** |
| --- | --- |
| autovacuum\_vacuum\_threshold | 50 |
| autovacuum\_vacuum\_scale\_factor | 0.2 (20%) |

For example, if your table has 1000 live rows:  
50 + 0.2 \* 1000 = 250 → If dead tuples exceed **250**, vacuum will trigger.

**🧠 3. Estimate Bloat (Optional but Useful)**

You can use this query to estimate table bloat:

SELECT

schemaname,

relname AS table\_name,

n\_live\_tup,

n\_dead\_tup,

ROUND(100.0 \* n\_dead\_tup / (n\_live\_tup + n\_dead\_tup), 2) AS dead\_pct

FROM

pg\_stat\_user\_tables

WHERE

n\_dead\_tup > 0

ORDER BY

dead\_pct DESC;

If dead\_pct is high (e.g., >20–30%), it's a sign to **manually vacuum**.

**🛠 4. Run Manual Vacuum (if needed)**

VACUUM ANALYZE your\_table\_name;