

Tarea M51 – Bryan Alberto Coronado García

- Configuración de Plataforma ETL y generación de una nueva función.

1) Se generó en S3 el bucket para cargar el archivo “Housing”.

housing-csv1 [Información](#)

[Objetos](#) [Metadatos](#) [Propiedades](#) [Permisos](#) [Métricas](#) [Administración](#) [Puntos de acceso](#)

Objetos (1)

[Copiar URI de S3](#) [Copiar URL](#) [Descargar](#) [Abrir](#) [Eliminar](#) [Acciones](#) [Crear carpeta](#) [Cargar](#)

Los objetos son las entidades fundamentales que se almacenan en Amazon S3. Puede utilizar el [inventario de Amazon S3](#) para obtener una lista de todos los objetos de su bucket. Para que otras personas obtengan acceso a sus objetos, tendrá que concederles permisos de forma explícita. [Más información](#)

Buscar objetos por prefijo

<input type="checkbox"/> Nombre	<input type="checkbox"/> Tipo	<input type="checkbox"/> Última modificación	<input type="checkbox"/> Tamaño	<input type="checkbox"/> Clase de almacenamiento
Housing (1).csv	csv	23 Jul 2025 8:24:13 PM CST	29.3 KB	Estándar

2) Se creó una database “housing-postgres”

[Aurora and RDS](#) > [Databases](#) > [housing-postgres](#)

housing-postgres

[Modify](#) [Actions](#)

Summary

DB identifier housing-postgres	Status Available	Role Instance	Engine PostgreSQL	Recommendations
CPU <div style="width: 4.03%;">4.03%</div>	Class db.t4g.micro	Current activity <div style="width: 0.00%;">0.00 sessions</div>	Region & AZ us-east-2b	

[Connectivity & security](#) [Monitoring](#) [Logs & events](#) [Configuration](#) [Zero-ETL integrations](#) [Maintenance & backups](#) [Data](#)

Connectivity & security

[Endpoint & port](#) [Networking](#) [Security](#)

3) Se realizó el crawler en AWS Glue:

[AWS Glue](#)

Getting started
ETL jobs
 Visual ETL
 Notebooks
 Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations [New](#)

test-housing

[Edit](#) [Delete](#)

Database properties

Name test-housing	Description -	Location -	Created on (UTC) July 24, 2025 at 04:35:25
----------------------	------------------	---------------	---

Tables (1)

Last updated (UTC)
July 24, 2025 at 04:41:01

[Delete](#) [Add tables using crawler](#) [Add table](#)

Filter tables

<input type="checkbox"/> Name	<input type="checkbox"/> Database	<input type="checkbox"/> Location	<input type="checkbox"/> Classification	<input type="checkbox"/> Deprecated	<input type="checkbox"/> View data	<input type="checkbox"/> Data quality	<input type="checkbox"/> Column statis...
housing-ebac	test-housing	s3://housing-csv1/	CSV	-	Table data	View data quality	View statistics

- Esta función generará una tabla adicional en la base de datos con los precios promedios por año y por zipcode. NOTA: NO existe columna de años, se toma “área” como tabla para “zipcode”

The screenshot shows the pgAdmin interface. On the left, the Object Explorer pane displays the database schema, including publications, schemas (public), aggregates, collations, domains, FTS configurations, dictionaries, parsers, templates, foreign tables, functions, materialized views, operators, procedures, sequences, and tables (housing). The 'Columns (13)' under the housing table are listed: price, area, bedrooms, bathrooms, stories, mainroad, guestroom, basement, hotwaterheating, airconditioning, and parking.

In the central pane, a query window titled 'postgres/postgres@housing-postgres*' contains the following SQL code:

```

1 v SELECT area, AVG(price::numeric) AS promedio_costo
2 FROM housing
3 GROUP BY area;

```

The 'Data Output' tab shows the results of the query:

	area integer	promedio_costo numeric
1	3934	2660000.00000000000000
2	5400	4631666.666666666667
3	2850	3640000.000000000000
4	7160	5880000.000000000000
5	4785	3395000.000000000000
6	3968	4410000.000000000000
7	4046	4165000.000000000000
8	3850	3136000.000000000000
9	3760	5530000.000000000000
10	2640	2835000.000000000000
11	2475	2380000.000000000000
12	2550	4550000.000000000000
13	7606	4070000.000000000000

Total rows: 284 Query complete 00:00:00.406

- Se hará un dump de la información de la nueva table al bucket de archivos.

The screenshot shows the Amazon S3 console. The left sidebar lists general purpose buckets, directory buckets, table buckets, vector buckets, access grants, and various access points. The main pane shows the 'housing-csv1' bucket. The 'Objects' tab is selected, displaying two objects: 'Housing (1).csv' (CSV file, 29.3 KB, Standard storage) and 'run-1753334543111-part-r-00000' (CSV file, 28.7 KB, Standard storage).