

# 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

< 1 >

<input type="checkbox"/>	Nombre	Tipo	Última modificación	Tamaño	Clase de almacenamiento
<input type="checkbox"/>	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

Aurora and RDS

Dashboard

Databases

Performance insights

Snapshots

Exports in Amazon S3

Automated backups

Reserved instances

Proxies

Subnet groups

Parameter groups

Option groups

housing-postgres

Modify

Actions

Summary

DB identifier

housing-postgres

Status

Available

Role

Instance

Engine

PostgreSQL

Recommendations

CPU

4.03%

Class

db.t4g.micro

Current activity

0.00 sessions

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

test-housing

AWS Glue

Getting started

ETL jobs

Visual ETL

Notebooks

Job run monitoring

Data Catalog tables

Data connections

Workflows (orchestration)

Zero-ETL integrations

Data Catalog

Databases

Tables

Stream schema registries

Schemas

Connections

Crawlers

Classifiers

Catalog settings

Announcing new optimization features for Apache Iceberg tables

Optimize storage for Apache Iceberg tables with automatic snapshot retention and orphan file deletion. [Learn more](#)

July 24, 2025 at 04:37:55

Edit

Delete

test-housing

Database properties

Name

test-housing

Description

-

Location

-

Created on (UTC)

July 24, 2025 at 04:35:25

Tables (1)

View and manage all available tables.

Filter tables

< 1 >

<input type="checkbox"/>	Name	Database	Location	Classification	Deprecated	View data	Data quality	Column statis...
<input type="checkbox"/>	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 a PostgreSQL interface with the following components:

- Object Explorer:** A tree view on the left showing the database structure. The 'housing' table is selected under the 'public' schema.
- Query Editor:** The main area where the SQL query is written:
 

```
SELECT area, AVG(price::numeric) AS promedio_costo
FROM housing
GROUP BY area;
```
- Data Output:** A table showing the results of the query. It has two columns: 'area' (integer) and 'promedio\_costo' (numeric). The results are as follows:
 

area	promedio_costo
1	2660000.000000000000
2	4631666.666666666667
3	3640000.000000000000
4	5880000.000000000000
5	3395000.000000000000
6	4410000.000000000000
7	4165000.000000000000
8	3136000.000000000000
9	5530000.000000000000
10	2835000.000000000000
11	2380000.000000000000
12	4550000.000000000000
13	4070000.000000000000
- Summary:** At the bottom, it indicates 'Total rows: 284' and '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 interface for a bucket named 'housing-csv1'. The 'Objects' tab is selected, showing a list of objects:

Name	Type	Last modified	Size	Storage class
Housing (1).csv	csv	July 23, 2025, 20:24:13 (UTC-06:00)	29.3 KB	Standard
run-1753334543111-part-r-00000	-	July 23, 2025, 23:22:32 (UTC-06:00)	28.7 KB	Standard