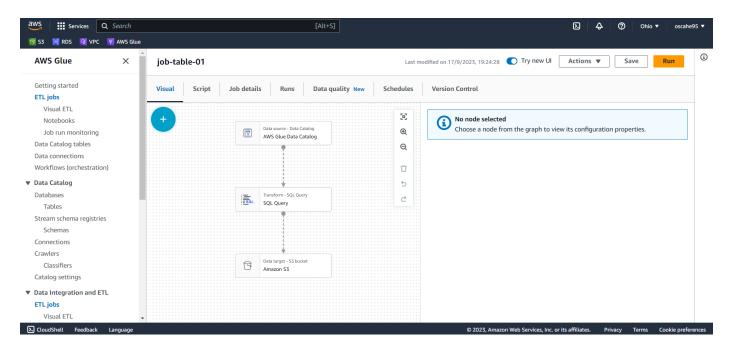
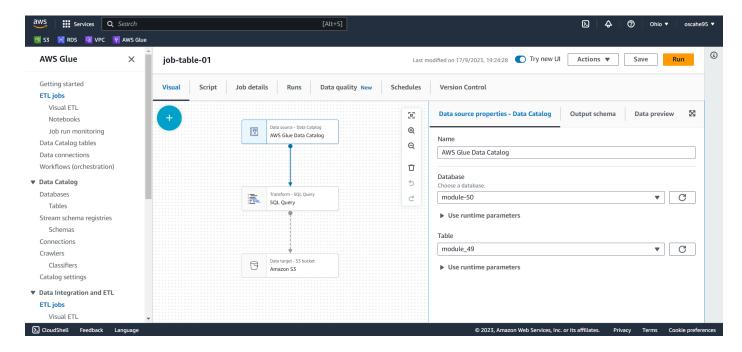
Ejercicio 50 - Proyecto Cloud Analytics II

- Construcción de un ETL Avanzado que genere dos tablas adicionales a la de housing:
- 1) promedio de precio, tamaño, número de habitaciones por estado de amobleamento:

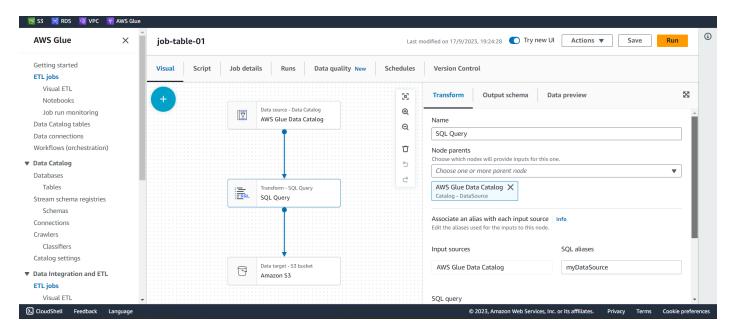
ETL JOB-01



Data Source



Transform



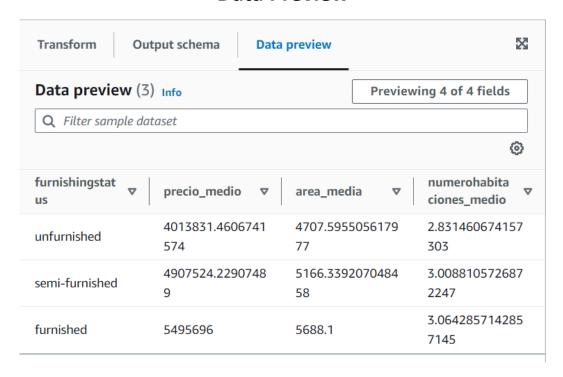
En vez de usar la opción de "Aggregate" realicé una query en SQL:

SQL query

Enter a SQL statement to add to your job.

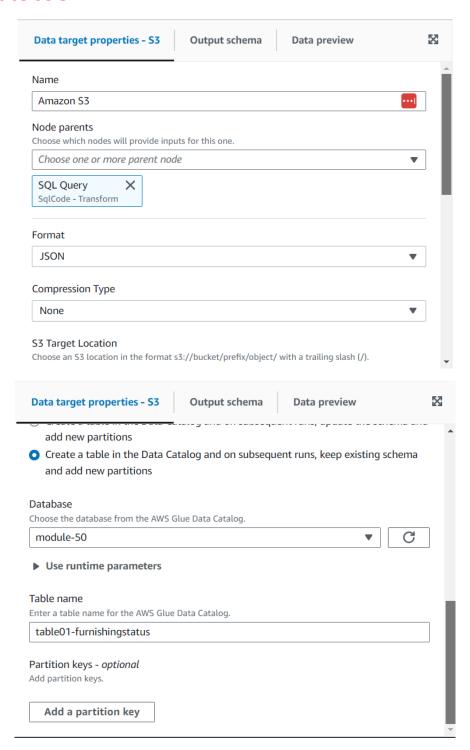
```
SELECT
2
       furnishingstatus,
3
       avg(price) as precio_medio,
       avg(area) as area_media,
4
5
       avg(bedrooms) as numerohabitaciones medio
6
       FROM myDataSource
7
       group by 1
8
       order by 2;
9
```

Data Preview

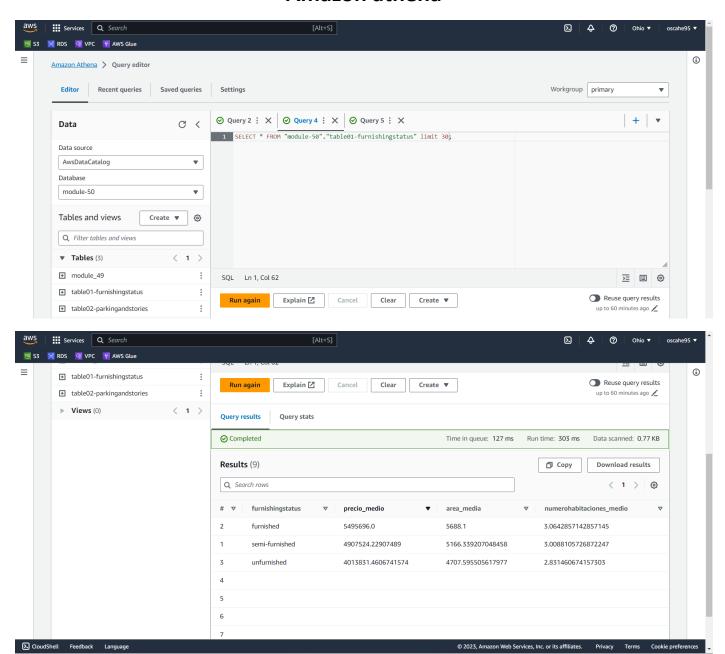


Data Target

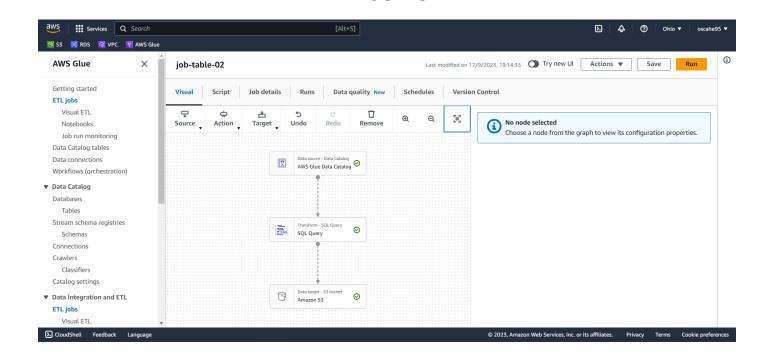
No se por cual motivo pero cuando utilice el formato csv la tabla final no era leida correctamente. En cambio, al hacerlo con JSON obtuve exactamente el mismo resultado del "data preview". Por este motivo usé el formato JSON.



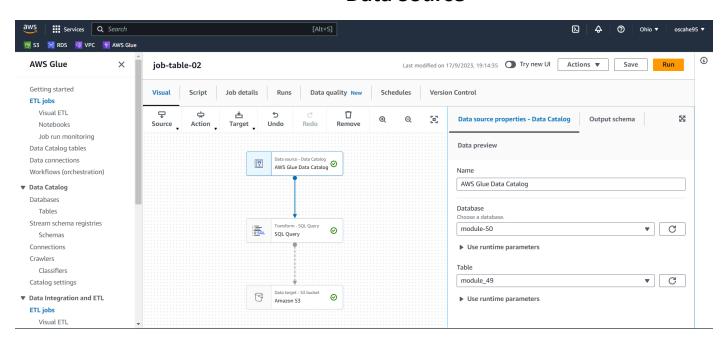
Amazon athena



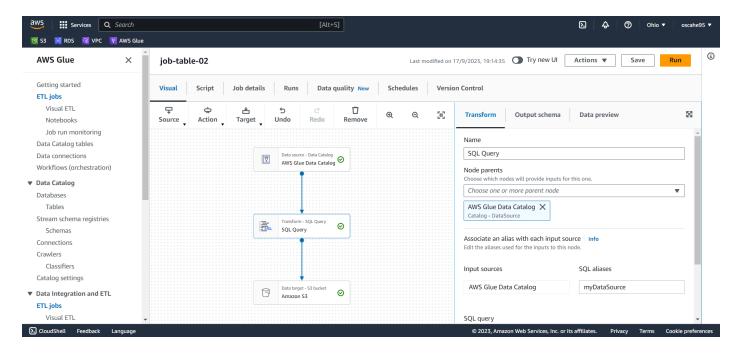
2) precio promedio de casa por stories y parking ETL JOB-02



Data Source



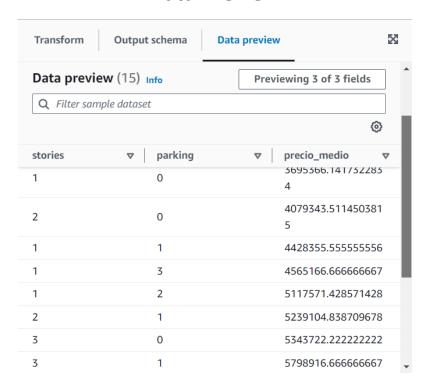
Transform



En vez de usar la opción de "Aggregate" realicé una query en SQL:

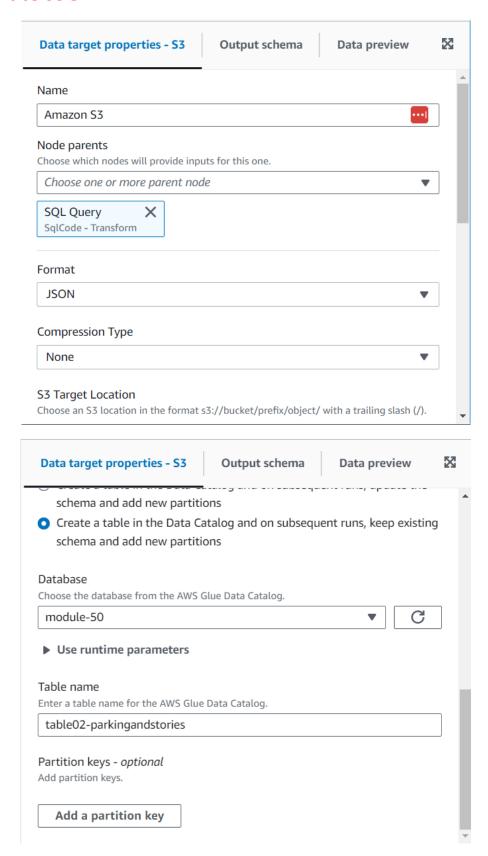


Data Preview

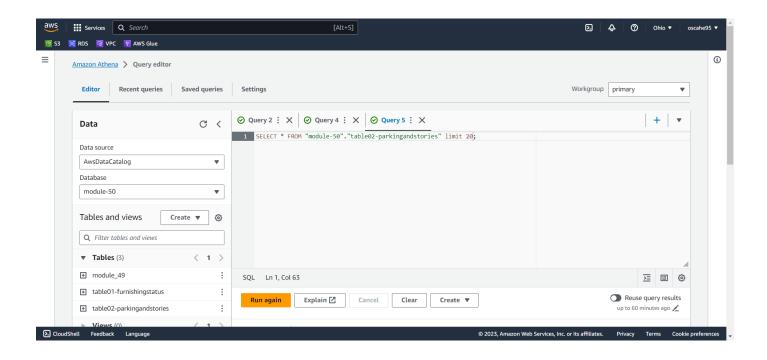


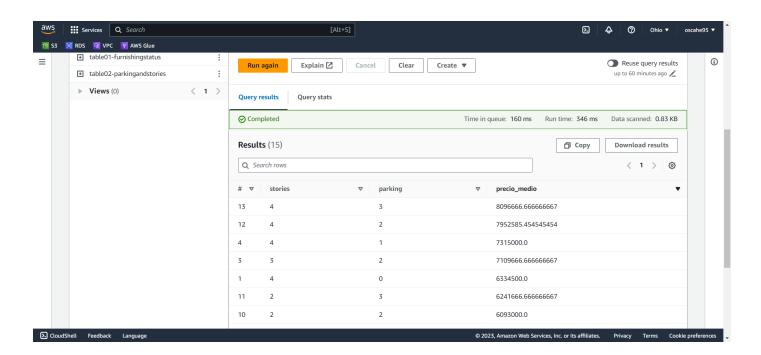
Data Target

No se por cual motivo pero cuando utilice el formato csv la tabla final no era leida correctamente. En cambio, al hacerlo con JSON obtuve exactamente el mismo resultado del "data preview". Por este motivo usé el formato JSON.



Amazon athena





Soportes adicionales

