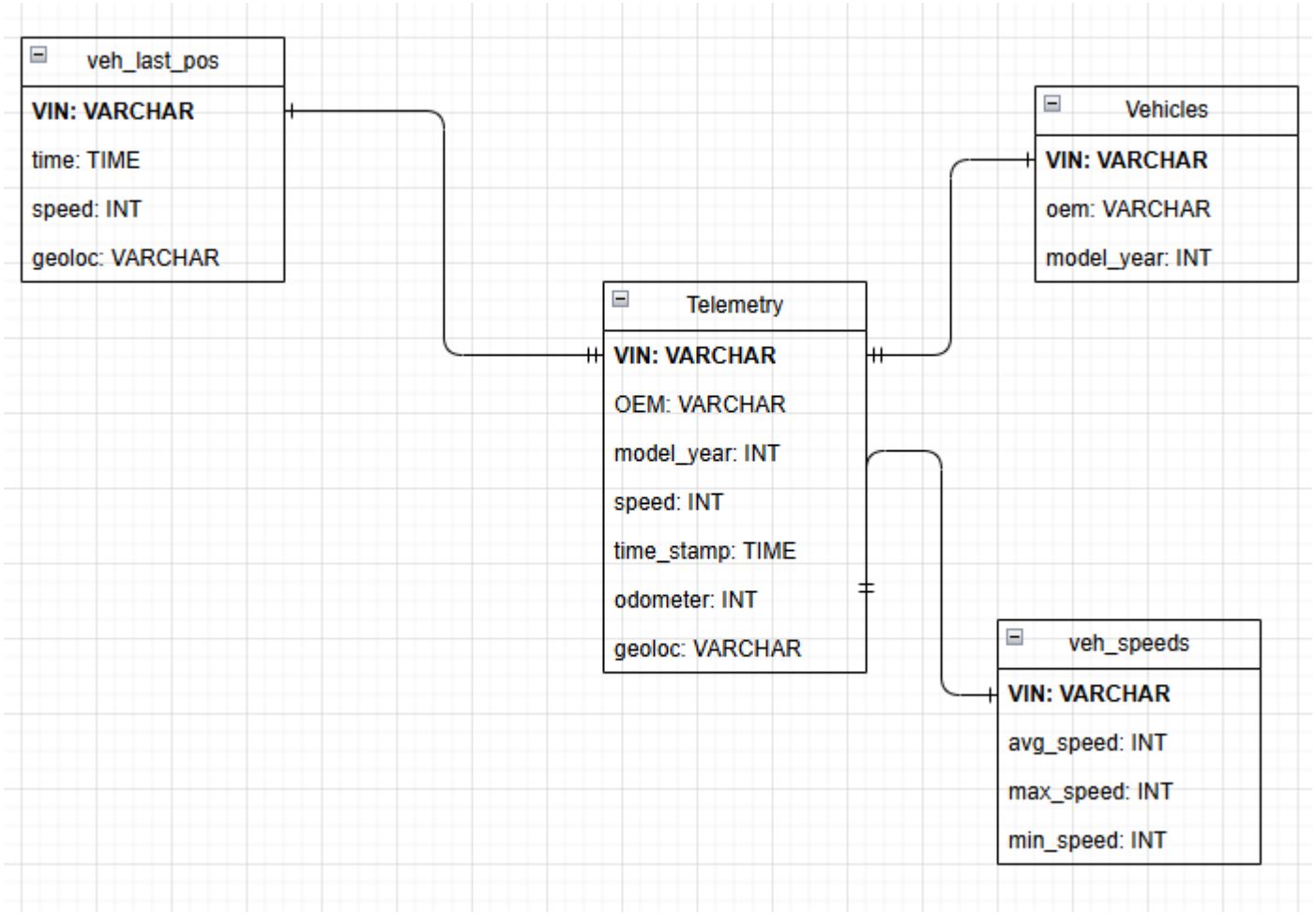
Proyecto Integrador

Enrique Rodríguez Toscano

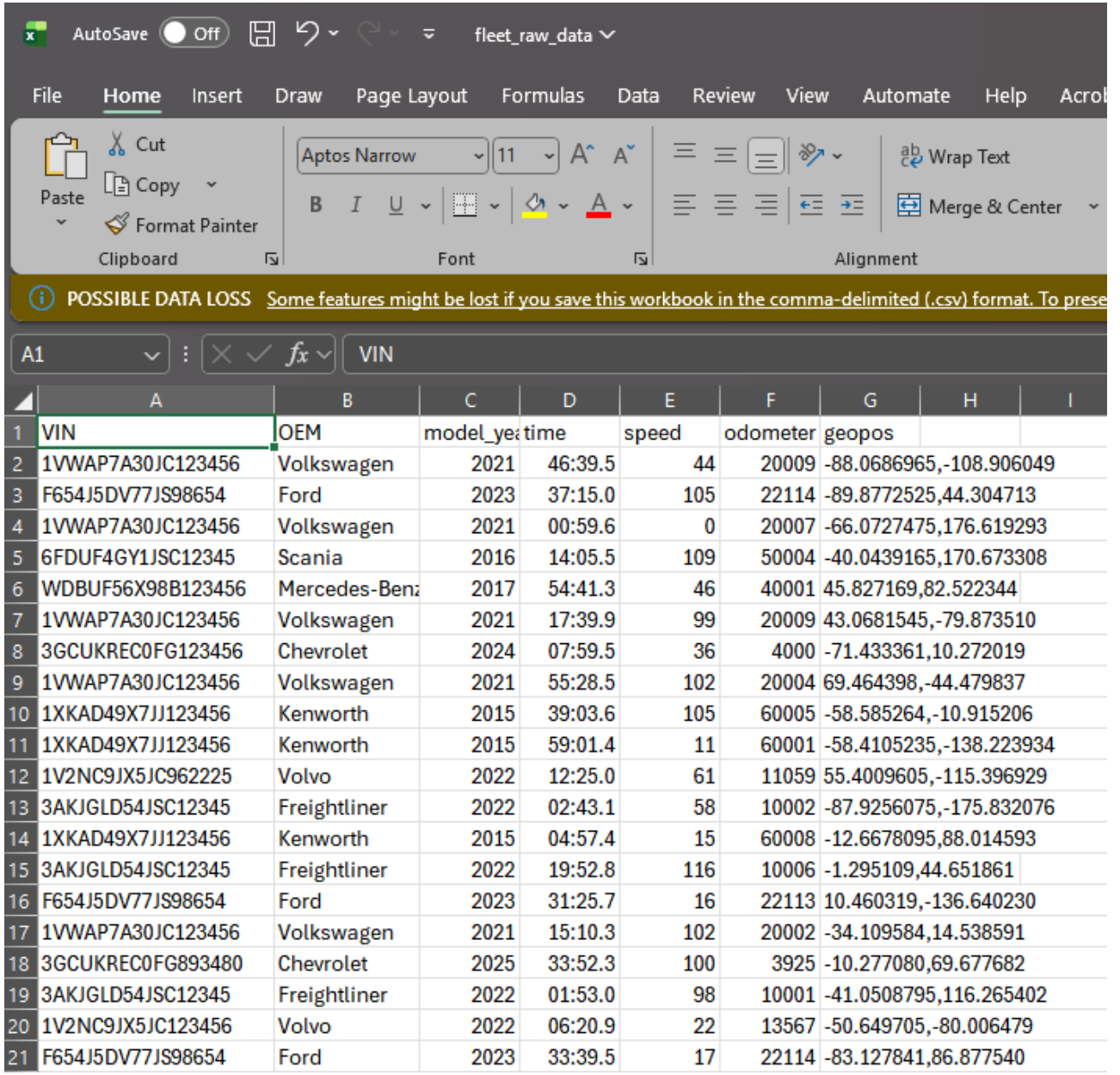
Vehicle fleet Simple Worflow (12 Vehicle fleet)

DBC Design

\ProyectoFinal_BigData\data\dbc_diagram.dio



Synthetic Raw Data generated

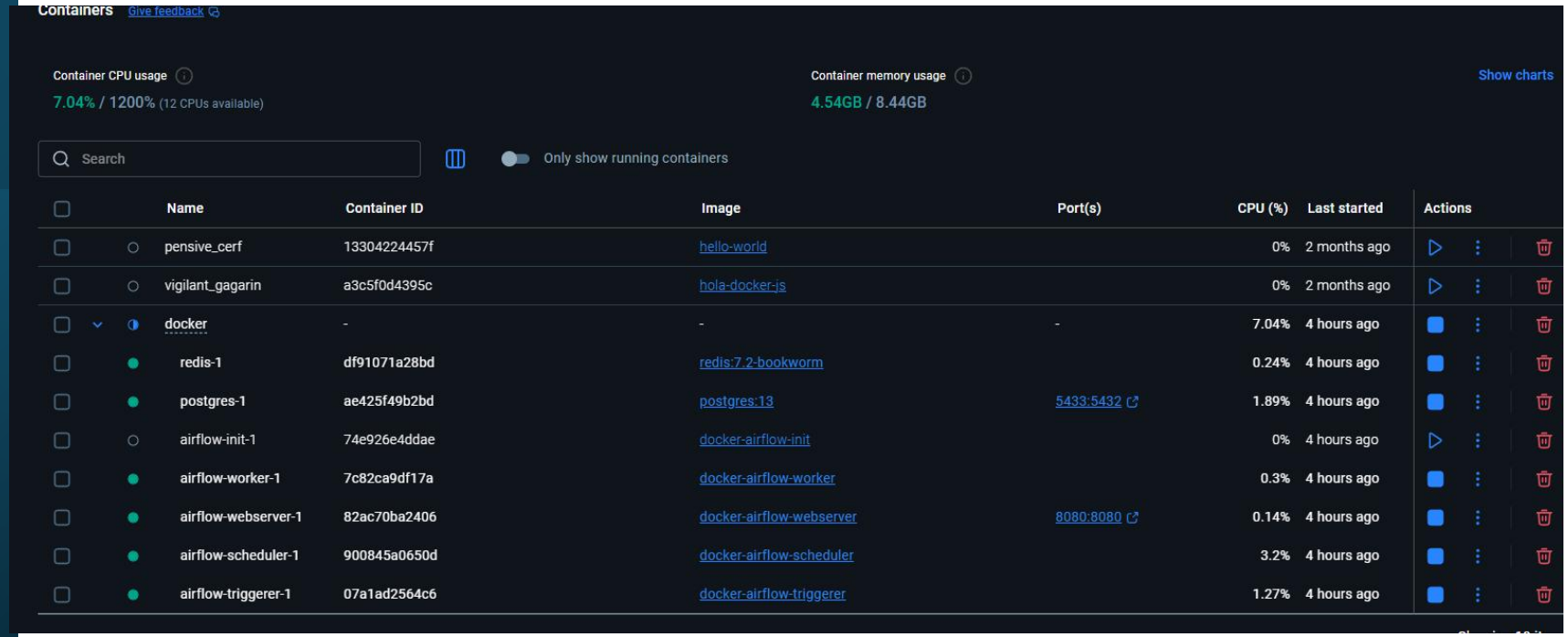


	A	B	C	D	E	F	G	H	I
1	VIN	OEM	model_ye	time	speed	odometer	geopos		
2	1VWAP7A30JC123456	Volkswagen	2021	46:39.5	44	20009	-88.0686965,-108.906049		
3	F654J5DV77JS98654	Ford	2023	37:15.0	105	22114	-89.8772525,44.304713		
4	1VWAP7A30JC123456	Volkswagen	2021	00:59.6	0	20007	-66.0727475,176.619293		
5	6FDUF4GY1JSC12345	Scania	2016	14:05.5	109	50004	-40.0439165,170.673308		
6	WDBUF56X98B123456	Mercedes-Benz	2017	54:41.3	46	40001	45.827169,82.522344		
7	1VWAP7A30JC123456	Volkswagen	2021	17:39.9	99	20009	43.0681545,-79.873510		
8	3GCUKREC0FG123456	Chevrolet	2024	07:59.5	36	4000	-71.433361,10.272019		
9	1VWAP7A30JC123456	Volkswagen	2021	55:28.5	102	20004	69.464398,-44.479837		
10	1XKAD49X7JJ123456	Kenworth	2015	39:03.6	105	60005	-58.585264,-10.915206		
11	1XKAD49X7JJ123456	Kenworth	2015	59:01.4	11	60001	-58.4105235,-138.223934		
12	1V2NC9JX5JC962225	Volvo	2022	12:25.0	61	11059	55.4009605,-115.396929		
13	3AKJGLD54JSC12345	Freightliner	2022	02:43.1	58	10002	-87.9256075,-175.832076		
14	1XKAD49X7JJ123456	Kenworth	2015	04:57.4	15	60008	-12.6678095,88.014593		
15	3AKJGLD54JSC12345	Freightliner	2022	19:52.8	116	10006	-1.295109,44.651861		
16	F654J5DV77JS98654	Ford	2023	31:25.7	16	22113	10.460319,-136.640230		
17	1VWAP7A30JC123456	Volkswagen	2021	15:10.3	102	20002	-34.109584,14.538591		
18	3GCUKREC0FG893480	Chevrolet	2025	33:52.3	100	3925	-10.277080,69.677682		
19	3AKJGLD54JSC12345	Freightliner	2022	01:53.0	98	10001	-41.0508795,116.265402		
20	1V2NC9JX5JC123456	Volvo	2022	06:20.9	22	13567	-50.649705,-80.006479		
21	F654J5DV77JS98654	Ford	2023	33:39.5	17	22114	-83.127841,86.877540		

Docker up and runing

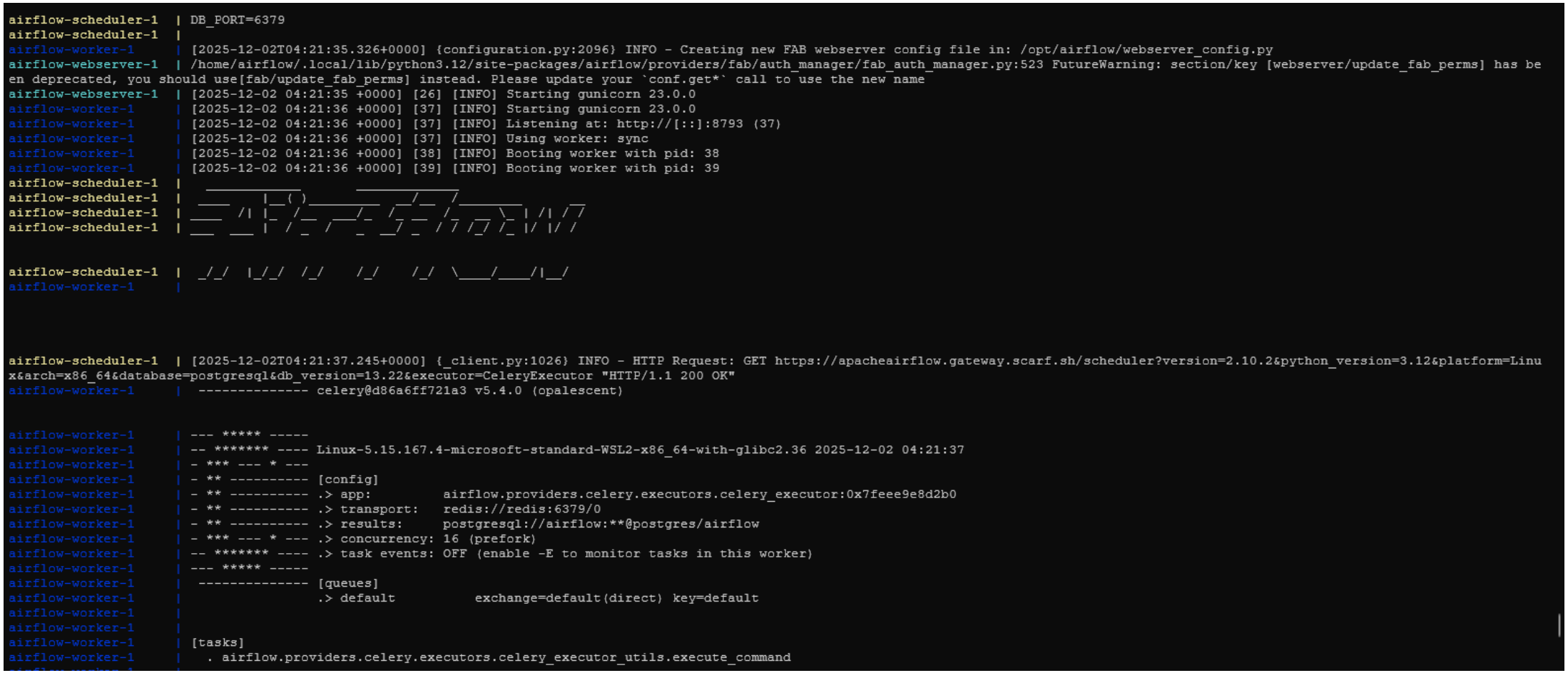
ProyectoFinal_BigData\docker\Dockerfile

To start Docker go to folder “ProyectoFinal_BigData” and run:
docker compose -f docker/docker-compose.yaml up --build



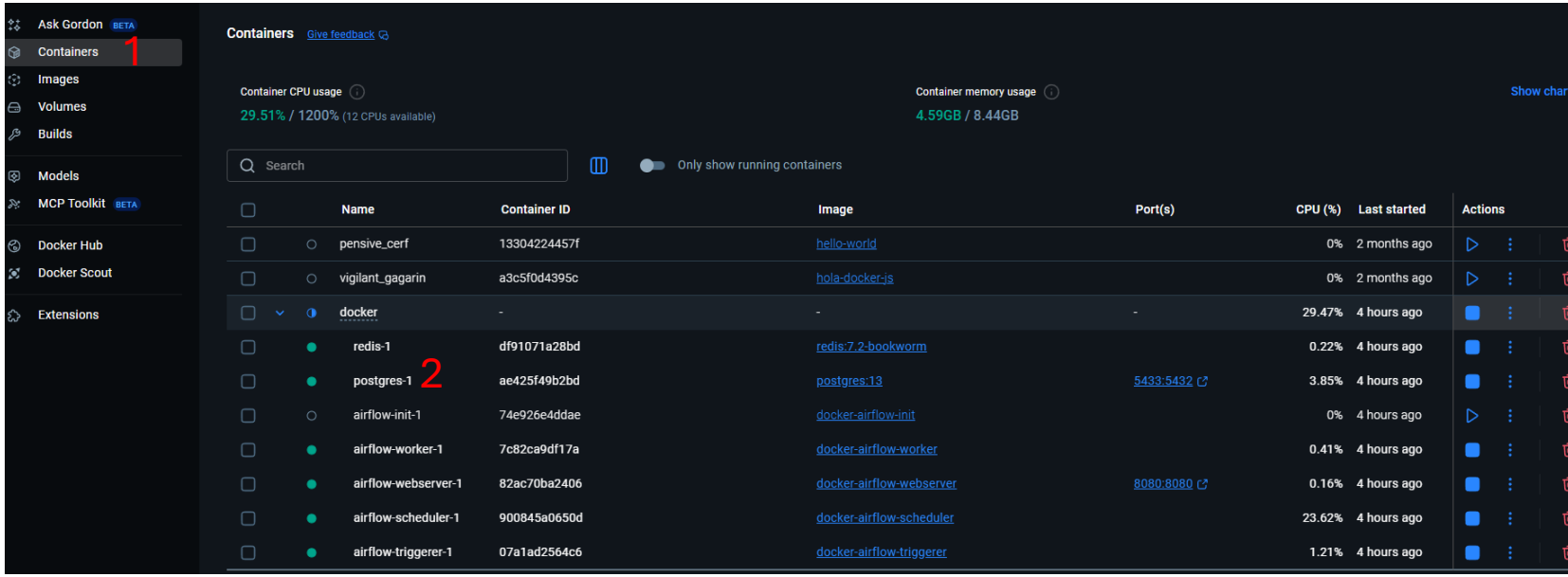
Docker up and running

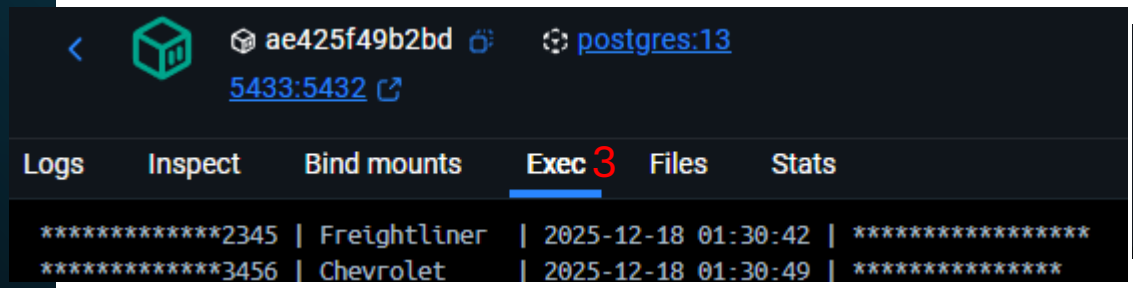
To start Docker go to folder “ProyectoFinal_BigData” and run:
docker compose -f docker/docker-compose.yaml up --build

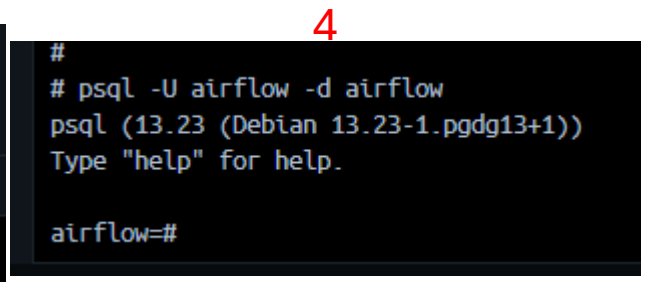


Raw Schema creation and raw data import

To see schemas and table inside Docker open Docker desktop
1 Click on Containers
2 Double click on postgres-1 container
3 Clck in Exec
4. run: psql –U airflow –d airflow







Raw Schema creation and raw data import

Tables

Once inside execute \dn to see al schemas

```
airflow=# \dn
     List of schemas
      Name       |  Owner
-----------------+---------
 driven_raw      | airflow
 driven_staging  | airflow
 driven_trusted  | airflow
 public          | airflow
(4 rows)
```

```
airflow=# dt\
invalid command \
Try \? for help.
airflow-# \dt
airflow-# \dn
     List of schemas
      Name       |  Owner
-----------------+---------
 driven_raw      | airflow
 driven_staging  | airflow
 driven_trusted  | airflow
 public          | airflow
(4 rows)
```

To see all tables inside a schema execute \dn *schema_name*.* to see schema content

```
airflow=# DROP TABLE driven_raw.raw_batch_data ;
DROP TABLE
airflow=# \dt driven_raw.*
              List of relations
   Schema   |      Name      | Type  |  Owner
------------+----------------+-------+---------
 driven_raw | raw_batch_data | table | airflow
(1 row)
```

To see a table content runs psql query, and ends with “;”

\dn *schema_name*.* to see schema content

```
airflow=# SELECT * FROM driven_raw.raw_batch_data limit 20;
        vin        |      oem      | model_year |          time          | speed | odometer |        geopos
-------------------+---------------+------------+------------------------+-------+----------+-----------------------
 1VWAP7A30JC123456 | Volkswagen    |       2021 | 2025-11-28 17:46:39.508461 |    44 |    20009 | -88.0686965,-108.906049
 F654J5DV77JS98654 | Ford          |       2023 | 2025-11-28 20:37:15.007511 |   105 |    22114 | -89.8772525,44.304713
 1VWAP7A30JC123456 | Volkswagen    |       2021 | 2025-11-28 21:00:59.620985 |     0 |    20007 | -66.0727475,176.619293
 6FDUF4GY1JSC12345 | Scania        |       2016 | 2025-11-29 03:14:05.500707 |   109 |    50004 | -40.0439165,170.673308
 WDBUF56X98B123456 | Mercedes-Benz |       2017 | 2025-11-28 22:54:41.271485 |    46 |    40001 | 45.827169,82.522344
 1VWAP7A30JC123456 | Volkswagen    |       2021 | 2025-11-28 14:17:39.858323 |    99 |    20009 | 43.0681545,-79.873510
 3GCUKREC0FG123456 | Chevrolet     |       2024 | 2025-11-28 12:07:59.541961 |    36 |     4000 | -71.433361,10.272019
 1VWAP7A30JC123456 | Volkswagen    |       2021 | 2025-11-28 18:55:28.499485 |   102 |    20004 | 69.464398,-44.479837
 1XKAD49X7JJ123456 | Kenworth      |       2015 | 2025-11-28 11:39:03.588197 |   105 |    60005 | -58.585264,-10.915206
 1XKAD49X7JJ123456 | Kenworth      |       2015 | 2025-11-28 07:59:01.391861 |    11 |    60001 | -58.4105235,-138.223934
 1V2NC9JX5JC962225 | Volvo         |       2022 | 2025-11-28 13:12:24.993437 |    61 |    11059 | 55.4009605,-115.396929
 3AKJGLD54JSC12345 | Freightliner  |       2022 | 2025-11-28 19:02:43.074903 |    58 |    10002 | -87.9256075,-175.832076
 1XKAD49X7JJ123456 | Kenworth      |       2015 | 2025-11-28 08:04:57.37322  |    15 |    60008 | -12.6678095,88.014593
 3AKJGLD54JSC12345 | Freightliner  |       2022 | 2025-11-28 22:19:52.821633 |   116 |    10006 | -1.295109,44.651861
 F654J5DV77JS98654 | Ford          |       2023 | 2025-11-28 22:31:25.735616 |    16 |    22113 | 10.460319,-136.640230
 1VWAP7A30JC123456 | Volkswagen    |       2021 | 2025-11-28 08:15:10.300023 |   102 |    20002 | -34.109584,14.538591
 3GCUKREC0FG893480 | Chevrolet     |       2025 | 2025-11-28 21:33:52.336908 |   100 |     3925 | -10.277080,69.677682
 3AKJGLD54JSC12345 | Freightliner  |       2022 | 2025-11-29 01:01:53.048155 |    98 |    10001 | -41.0508795,116.265402
 1V2NC9JX5JC123456 | Volvo         |       2022 | 2025-11-28 06:06:20.928829 |    22 |    13567 | -50.649705,-80.006479
 F654J5DV77JS98654 | Ford          |       2023 | 2025-11-28 09:33:39.464821 |    17 |    22114 | -83.127841,86.877540
(20 rows)
```

Staging Schema creation and table creation

```
airflow=# \dt driven_staging.*
                List of relations
     Schema     |     Name     | Type  |  Owner
----------------+--------------+-------+---------
 driven_staging | dim_last_pos | table | airflow
 driven_staging | telemetry    | table | airflow
 driven_staging | veh_speeds   | table | airflow
 driven_staging | vehicles     | table | airflow
(4 rows)
```

```
airflow=# SELECT * FROM driven_staging.veh_speeds;
        vin        |     avg_speed      | max_speed | min_speed
-------------------+--------------------+-----------+-----------
 1V2NC9JX5JC123456 | 60.0423508821264210 |       120 |         0
 1V2NC9JX5JC962225 | 60.1924282827770986 |       120 |         0
 1VWAP7A30JC123456 | 60.1354669548688792 |       120 |         0
 1XKAD49X7JJ123456 | 60.0737079755013145 |       120 |         0
 3AKJGLD54JSC12345 | 59.7743354539003060 |       120 |         0
 3GCUKREC0FG123456 | 60.0788512846368774 |       120 |         0
 3GCUKREC0FG893480 | 59.8236931178722479 |       120 |         0
 6FDUF4GY1JSC12345 | 59.9128900738672324 |       120 |         0
 F654J5DV77JS46566 | 60.0938576815380168 |       120 |         0
 F654J5DV77JS98654 | 60.0127507843244310 |       120 |         0
 LZB12345678901234 | 60.0771827116136532 |       120 |         0
 WDBUF56X98B123456 | 59.9946310537823213 |       120 |         0
(12 rows)
```

```
airflow=# SELECT * FROM driven_staging.dim_last_pos LIMIT 20;
        vin        |      oem      |            time            |        geopos
-------------------+---------------+----------------------------+----------------------
 1V2NC9JX5JC123456 | Volvo         | 2025-12-03 03:52:55.282961 | 47.850364,47.445878
 1V2NC9JX5JC962225 | Volvo         | 2025-12-03 04:21:12.847493 | 38.8466925,-53.461134
 1VWAP7A30JC123456 | Volkswagen    | 2025-12-03 04:27:42.483603 | -3.9792025,131.287153
 1XKAD49X7JJ123456 | Kenworth      | 2025-12-03 04:22:18.382928 | -35.650918,-58.772821
 3AKJGLD54JSC12345 | Freightliner  | 2025-12-03 04:13:42.907018 | 56.4041145,-61.153963
 3GCUKREC0FG123456 | Chevrolet     | 2025-12-03 04:09:41.051617 | 87.0324565,118.591120
 3GCUKREC0FG893480 | Chevrolet     | 2025-12-03 04:20:08.740631 | 34.471792,-165.425157
 6FDUF4GY1JSC12345 | Scania        | 2025-12-03 04:24:43.276448 | -1.864793,59.415064
 F654J5DV77JS46566 | Ford          | 2025-12-03 03:14:08.523712 | 8.677037,119.675660
 F654J5DV77JS98654 | Ford          | 2025-12-03 04:28:21.59087  | 47.4118195,-61.047921
 LZB12345678901234 | FAW           | 2025-12-03 04:23:26.574768 | -87.3454665,23.925814
 WDBUF56X98B123456 | Mercedes-Benz | 2025-12-03 04:11:25.587765 | 3.266145,55.275334
(12 rows)
```

Driven Trusted + no PII

Tables

```
airflow=# \dt driven_trusted.*
              List of relations
     Schema     |     Name      | Type  |  Owner
----------------+---------------+-------+---------
 driven_trusted | veh_last_pos  | table | airflow
 driven_trusted | veh_speeds    | table | airflow
 driven_trusted | veh_telemetry | table | airflow
(3 rows)
```

```
airflow=# SELECT * FROM driven_trusted.veh_telemetry limit 10;
    masked_vin     |     oem      | model_year | speed | odometer |        time         |   masked_geopos
-------------------+--------------+------------+-------+----------+---------------------+-------------------
 **********1234 | FAW          |       2023 |    97 |     8000 | 2025-12-17 02:43:00 | ****************
 **********3456 | Volkswagen   |       2021 |    67 |    20004 | 2025-12-17 12:29:43 | ****************
 **********3456 | Volvo        |       2022 |     3 |    13560 | 2025-12-17 12:09:57 | ****************
 **********3480 | Chevrolet    |       2025 |   115 |     3926 | 2025-12-17 18:40:56 | ****************
 **********8654 | Ford         |       2023 |    65 |    22113 | 2025-12-17 04:14:37 | ****************
 **********2345 | Freightliner |       2022 |   103 |    10007 | 2025-12-17 21:46:05 | ****************
 **********3456 | Chevrolet    |       2024 |   107 |     4008 | 2025-12-17 16:14:17 | ****************
 **********3480 | Chevrolet    |       2025 |   119 |     3930 | 2025-12-16 23:37:46 | ****************
 **********3456 | Volvo        |       2022 |   117 |    13567 | 2025-12-17 21:41:36 | ****************
 **********1234 | FAW          |       2023 |     2 |     8007 | 2025-12-17 16:25:04 | ****************
(10 rows)
```

```
airflow=# SELECT * FROM driven_trusted.veh_speeds limit 10;
        vin        |    masked_vin     |      avg_speed      | max_speed | min_speed
-------------------+-------------------+---------------------+-----------+-----------
 1V2NC9JX5JC123456 | **********3456 | 60.0423508821264210 |       120 |         0
 1V2NC9JX5JC962225 | **********2225 | 60.1924282827770986 |       120 |         0
 1VWAP7A30JC123456 | **********3456 | 60.1354669548688792 |       120 |         0
 1XKAD49X7JJ123456 | **********3456 | 60.0737079755013145 |       120 |         0
 3AKJGLD54JSC12345 | **********2345 | 59.7743354539003060 |       120 |         0
 3GCUKREC0FG123456 | **********3456 | 60.0788512846368774 |       120 |         0
 3GCUKREC0FG893480 | **********3480 | 59.8236931178722479 |       120 |         0
 6FDUF4GY1JSC12345 | **********2345 | 59.9128900738672324 |       120 |         0
 F654J5DV77JS46566 | **********6566 | 60.0938576815380168 |       120 |         0
 F654J5DV77JS98654 | **********8654 | 60.0127507843244310 |       120 |         0
(10 rows)
```

Airflow DAG execution