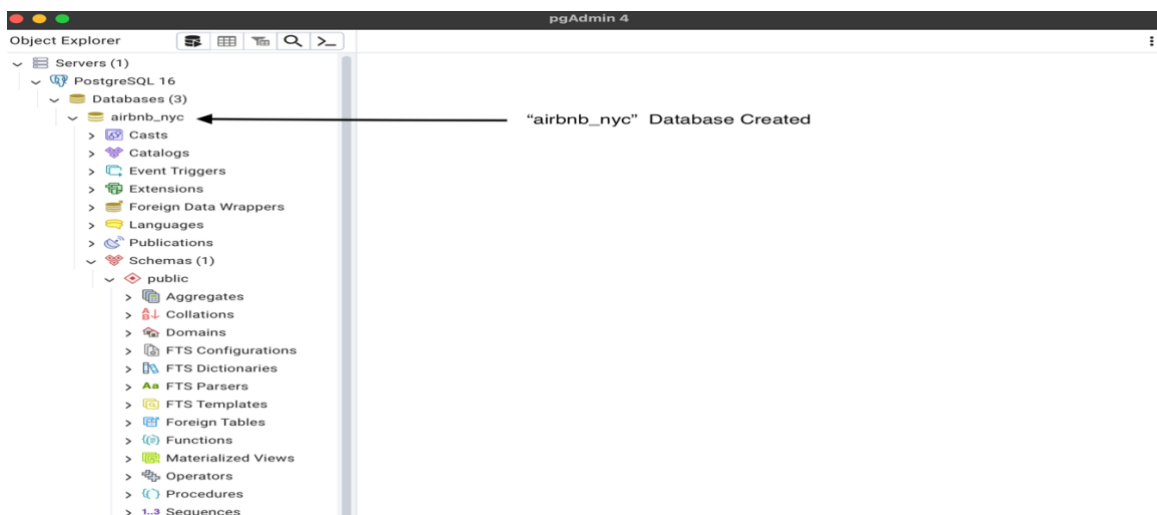


# Demonstration

## 1. Activating virtual environment and creating the database

```
AI_Planet_DE_Assignment --zsh -- 163x47
~/Desktop/AI_Planet_DE_Assignment --zsh

Last login: Sat Jul 13 16:34:28 on ttys000
(base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % source data_engineer_env/bin/activate
(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % psql -U postgres -c "CREATE DATABASE airbnb_nyc;"
Password for user postgres:
CREATE DATABASE
(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment %
```



## 2. Creating the “listings” table and loading it with raw data

The terminal window shows the following commands and output:

```
Last login: Sat Jul 13 17:39:52 on ttys000
(base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % source data_engineer_env/bin/activate
(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % python models.py
Table 'listings' created successfully!
(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % python -m src.load_data
Table 'listings' created successfully!
Data loading complete! Check the 'listings' table in your database.
(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment %
```

The pgAdmin 4 interface shows the 'public' schema in the 'airbnb\_nyc' database. The 'listings' table is highlighted in the Object Explorer. The Query Editor shows the following SQL query:

```
1 SELECT * FROM public.listings
2 LIMIT 100
3
```

The Data Output pane displays the first 10 rows of the 'listings' table:

id	name	host_id	host_name	neighbourhood
1	Clean & quiet apt home by the park	2787	John	Brooklyn
2	Skylit Midtown Castle	2845	Jennifer	Manhattan
3	THE VILLAGE OF HARLEM.... NEW YORK !	4632	Elisabeth	Manhattan
4	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn
5	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan
6	Large Cozy 1 BR Apartment In Midtown East	7322	Chris	Manhattan
7	BlissArtsSpace!	7356	Garon	Brooklyn
8	Large Furnished Room Near B'way	8967	Shunichi	Manhattan
9	Cozy Clean Guest Room - Family Apt	7490	MarvEllen	Manhattan

A message in the Messages pane states: "listings" table created and raw data was ingested.

3. Metaflow ETL pipeline was executed and “listings\_transformed” table was created.

```

AI_Planet_DE_Assignment -- -ssh - 163x47
~/Desktop/AI_Planet_DE_Assignment -- -ssh

(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % python airbnb_flow.py run
Metaflow 2.12.7 executing AirbnbETLFlow for user:saiteja
Validating your flow...
    The graph looks good!
Running pylint...
    Pylint is happy!

2024-07-13 17:32:33.190 Workflow starting (run-id 1720872153189611):
2024-07-13 17:32:33.200 [1720872153189611/start/1 (pid 55506)] Task is starting.
2024-07-13 17:32:33.796 [1720872153189611/start/1 (pid 55506)] 2024-07-13 17:32:33,796 - INFO - Starting Airbnb ETL flow...
2024-07-13 17:32:33.886 [1720872153189611/start/1 (pid 55506)] Task finished successfully.
2024-07-13 17:32:33.894 [1720872153189611/extract/2 (pid 55509)] Task is starting.
2024-07-13 17:32:34.426 [1720872153189611/extract/2 (pid 55509)] 2024-07-13 17:32:34,426 - INFO - Extracting data from the 'listings' table...
2024-07-13 17:32:34.934 [1720872153189611/extract/2 (pid 55509)] Task finished successfully.
2024-07-13 17:32:34.940 [1720872153189611/transform/3 (pid 55525)] Task is starting.
2024-07-13 17:32:35.443 [1720872153189611/transform/3 (pid 55525)] 2024-07-13 17:32:35,443 - INFO - Transforming data...
2024-07-13 17:32:35.683 [1720872153189611/transform/3 (pid 55525)] Task finished successfully.
2024-07-13 17:32:35.688 [1720872153189611/load/4 (pid 55528)] Task is starting.
2024-07-13 17:32:36.191 [1720872153189611/load/4 (pid 55528)] 2024-07-13 17:32:36,191 - INFO - Loading data into listings_transformed...
2024-07-13 17:32:38.894 [1720872153189611/load/4 (pid 55528)] 2024-07-13 17:32:38,894 - INFO - Data successfully loaded into listings_transformed
2024-07-13 17:32:39.034 [1720872153189611/load/4 (pid 55528)] Task finished successfully.
2024-07-13 17:32:39.039 [1720872153189611/end/5 (pid 55557)] Task is starting.
2024-07-13 17:32:39.516 [1720872153189611/end/5 (pid 55557)] 2024-07-13 17:32:39,516 - INFO - Airbnb ETL flow completed successfully!
2024-07-13 17:32:39.599 [1720872153189611/end/5 (pid 55557)] Task finished successfully.
2024-07-13 17:32:39.599 Done!

(data_engineer_env) (base) saiteja@Saitejas-MacBook-Air AI_Planet_DE_Assignment % █

```

The screenshot shows the pgAdmin 4 interface. In the left sidebar, the 'listings\_transformed' table is highlighted under the 'public' schema. The main window displays the table's structure and data. A green text annotation points to the table name in the sidebar.

**Table Structure:**

id	bigint	name	text	host_id	bigint	neighbourhood_group	text	neigh	text
1	2539	Clean & quiet apt home by the park		2787	Brooklyn	Kens			
2	2595	Skylit Midtown Castle		2845	Manhattan	Midt			
3	3647	THE VILLAGE OF HARLEM....NEW YORK !		4632	Manhattan	Harle			
4	3831	Cozy Entire Floor of Brownstone		4869	Brooklyn	Clint			
5	5022	Entire Apt: Spacious Studio/Loft by central park		7192	Manhattan	East			
6	5099	Large Cozy 1 BR Apartment In Midtown East		7322	Manhattan	Murr			
7	5121	BlissArtsSpace!		7356	Brooklyn	Bedf			
8	5178	Large Furnished Room Near B'way		8967	Manhattan	Hell'			
9	5203	Cozy Clean Guest Room - Family Ant		7490	Manhattan	Uttom			

**Query Results:**

Query: `SELECT * FROM public.listings_transformed`

Total rows: 1000 of 48879 | Query complete 00:00:00.248 | Ln 1, Col 1