[**DATA WAREHOUSE OF PET FINDER** 2](#_Toc110030550)

[1. Thiết kế ERD: 2](#_Toc110030551)

[2. Câu lệnh SQL xây dựng Database: 3](#_Toc110030552)

[3. Xác định các truy vấn nghiệp vụ: 4](#_Toc110030553)

[4. Xây dựng ETL cho Dimension / Fact Table 4](#_Toc110030554)

[5. Câu lệnh SQL cho các truy vấn nghiệp vụ: 8](#_Toc110030555)

**DATA WAREHOUSE OF PET FINDER**

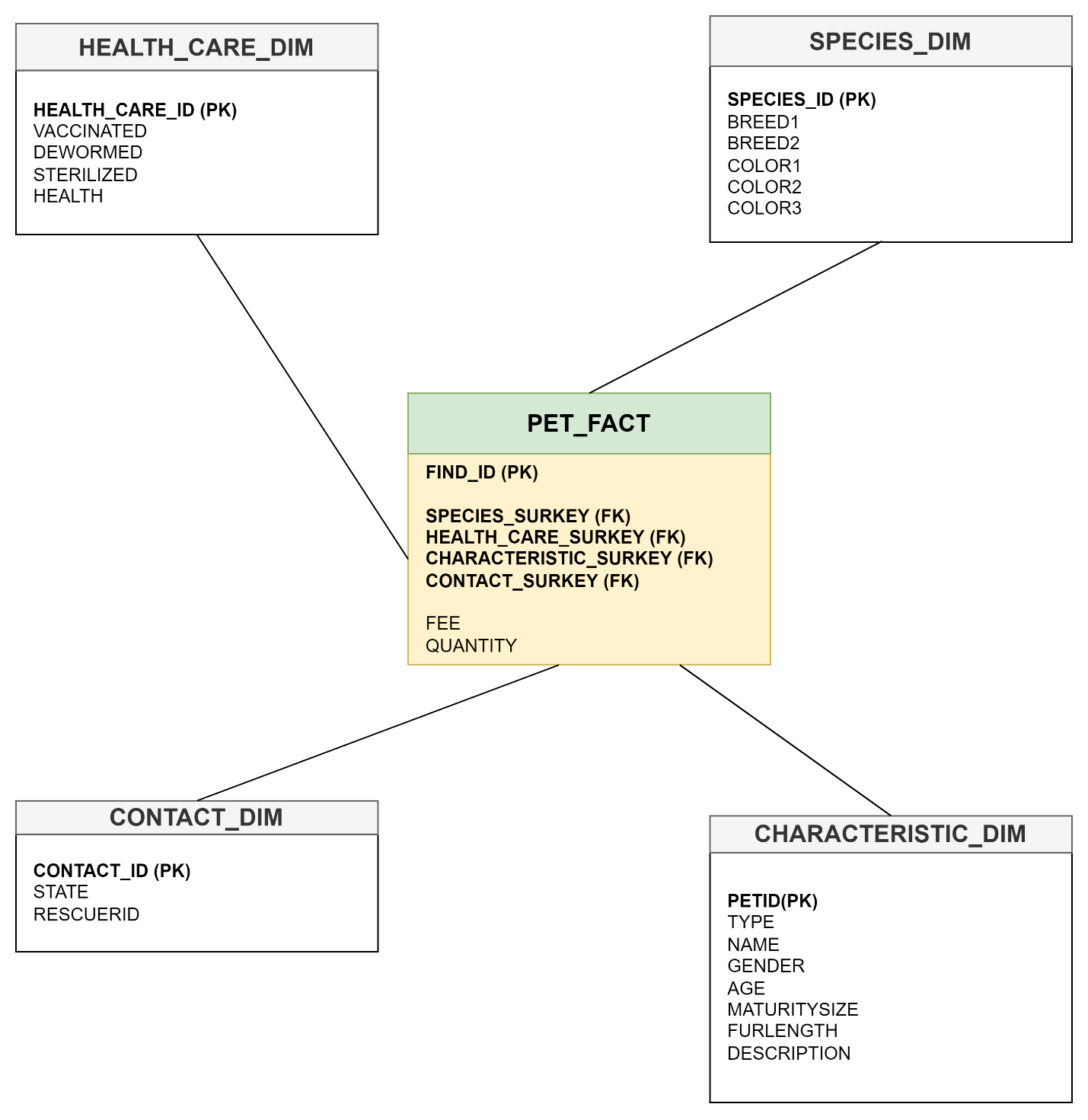
1. **Thiết kế ERD:**

Xác định phụ thuộc hàm:

* PetID → ALL

Để tối ưu các hệ thống lưu trữ data trong tương lai của bộ dữ liệu thú cưng, chia làm 4 bảng DIM và 1 bảng FACT:

* DIM:
  + Chế độ sức khỏe (Health\_care\_dim): ghi nhận các chế độ tiêm phòng
  + Chủng loài (Species\_dim): ghi nhận giống và màu sắc, phục vụ việc tìm kiếm
  + Liên hệ (Contact\_dim): ghi nhận thông tin người giải cứu – có thể thêm các thông tin chi tiết của người giải cứu
  + Đặc tính (Characteristic\_dim): ghi nhận chi tiết nhóm đặc tính của 1 con vật nuôi
* FACT:
  + Fee: chi phí để nhận nuôi, có thể tính toán tổng hay các mục đích thống kê khác
  + Quantity: số lượng để nhận nuôi, có thể tính toán tổng hay các mục đích thống kê khác

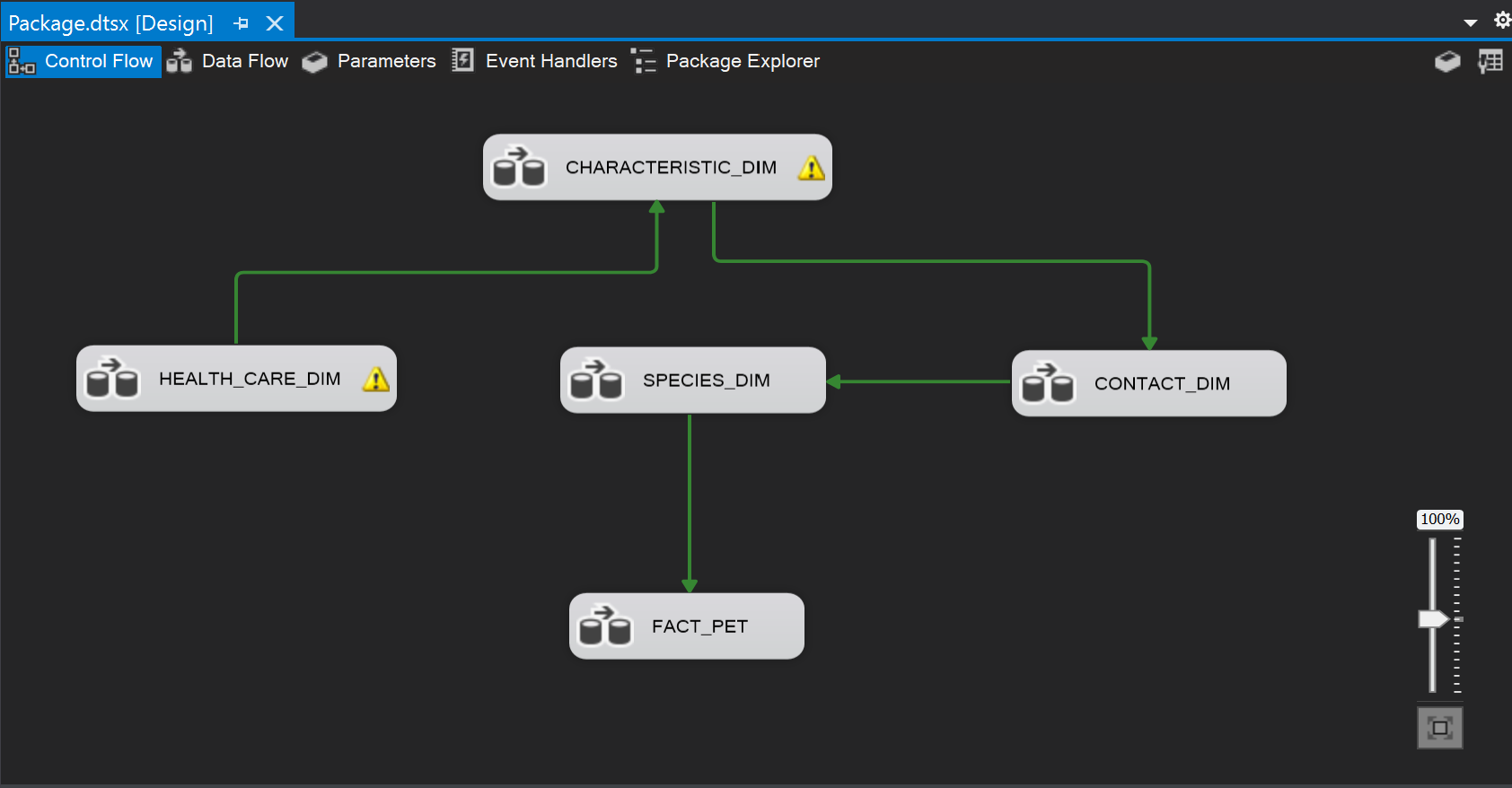


1. **Câu lệnh SQL xây dựng Database:**
2. -- TẠO BẢNG DIM VÀ FACT
3. CREATE TABLE HEALTH\_CARE\_DIM (
4. HEALTH\_CARE\_ID INT PRIMARY KEY,
5. VACCINATED VARCHAR(8),
6. DEWORMED VARCHAR(8),
7. STERILIZED VARCHAR(8),
8. HEALTH VARCHAR(14)
9. );
10. CREATE TABLE SPECIES\_DIM (
11. SPECIES\_ID INT PRIMARY KEY,
12. BREED1 VARCHAR(255),
13. BREED2 VARCHAR(255),
14. COLOR1 VARCHAR(255),
15. COLOR2 VARCHAR(255),
16. COLOR3 VARCHAR(255)
17. );
18. CREATE TABLE CONTACT\_DIM (
19. CONTACT\_ID INT PRIMARY KEY,
20. STATE VARCHAR(255),
21. RESCUERID INT
22. );
23. CREATE TABLE CHARACTERISTIC\_DIM (
24. PETID INT PRIMARY KEY,
25. TYPE VARCHAR(3),
26. NAME NVARCHAR(255),
27. AGE INT,
28. MATURITYSIZE VARCHAR(13),
29. FURLENGTH VARCHAR(13),
30. DESCRIPTION NVARCHAR(2000)
31. );
32. CREATE TABLE PET\_FACT (
33. SPECIES\_SURKEY INT FOREIGN KEY REFERENCES SPECIES\_DIM(SPECIES\_ID),
34. HEALTH\_CARE\_SURKEY INT FOREIGN KEY REFERENCES HEALTH\_CARE\_DIM(HEALTH\_CARE\_ID),
35. CHARACTERISTIC\_SURKEY INT FOREIGN KEY REFERENCES CHARACTERISTIC\_DIM(PETID),
36. CONTACT\_SURKEY INT FOREIGN KEY REFERENCES CONTACT\_DIM(CONTACT\_ID),
37. FEE INT,
38. QUANTITY INT,
39. );
40. **Xác định các truy vấn nghiệp vụ:**

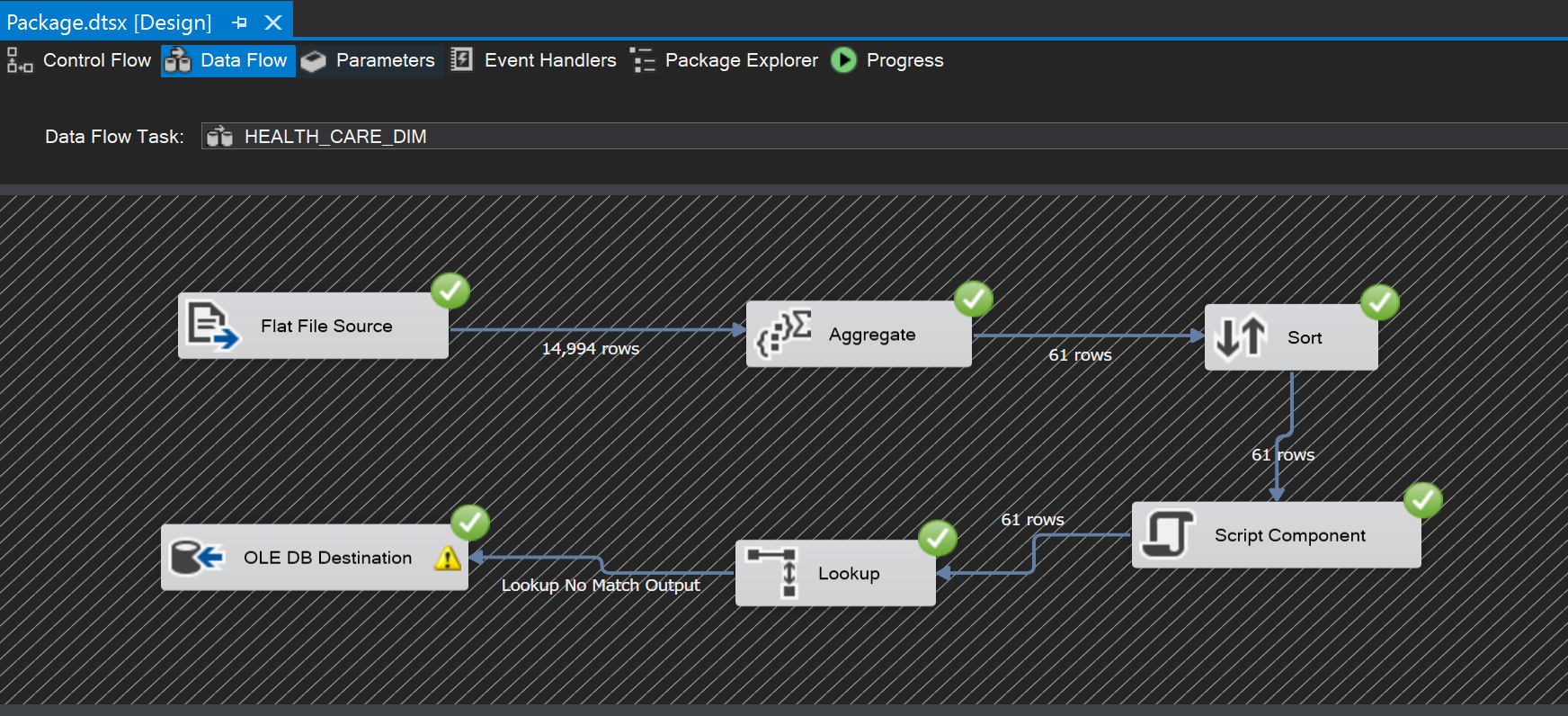
* **Có bao nhiêu con chó ở Kuala Lumpur có lông ngắn**
* **Tổng số mèo theo từng bang**
* **Số lượng mèo và chó dưới hoặc bằng 1 tuổi**

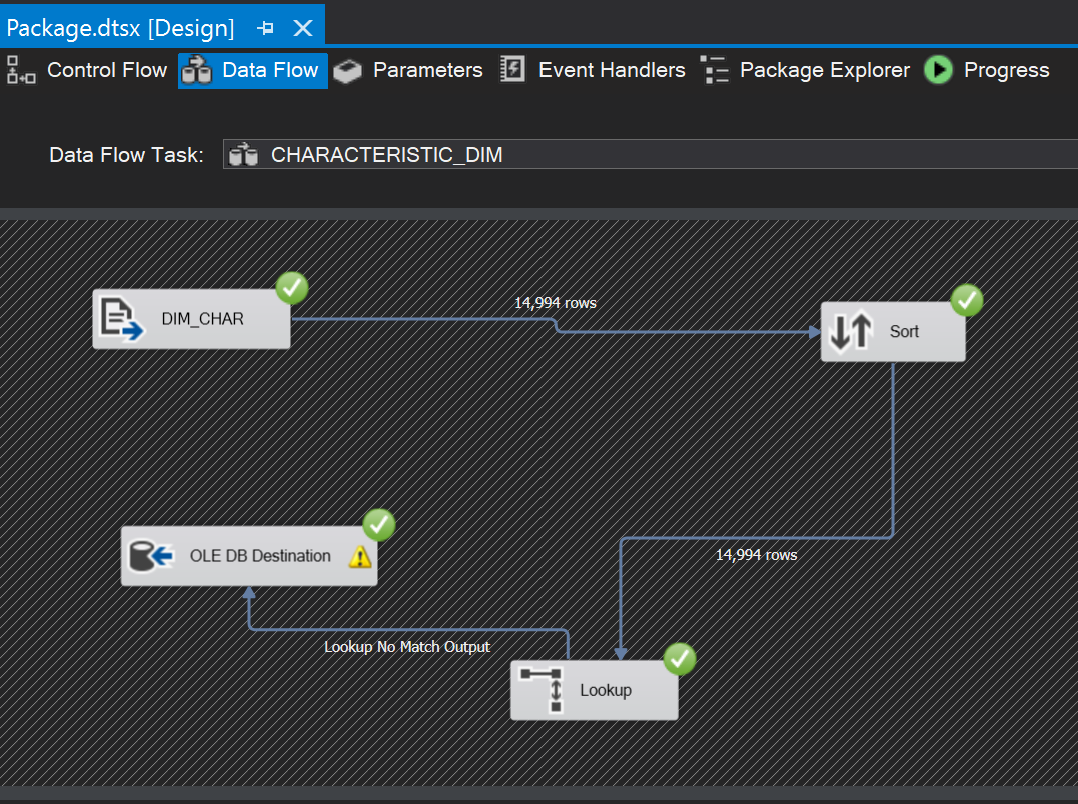
1. **Xây dựng ETL cho Dimension / Fact Table**

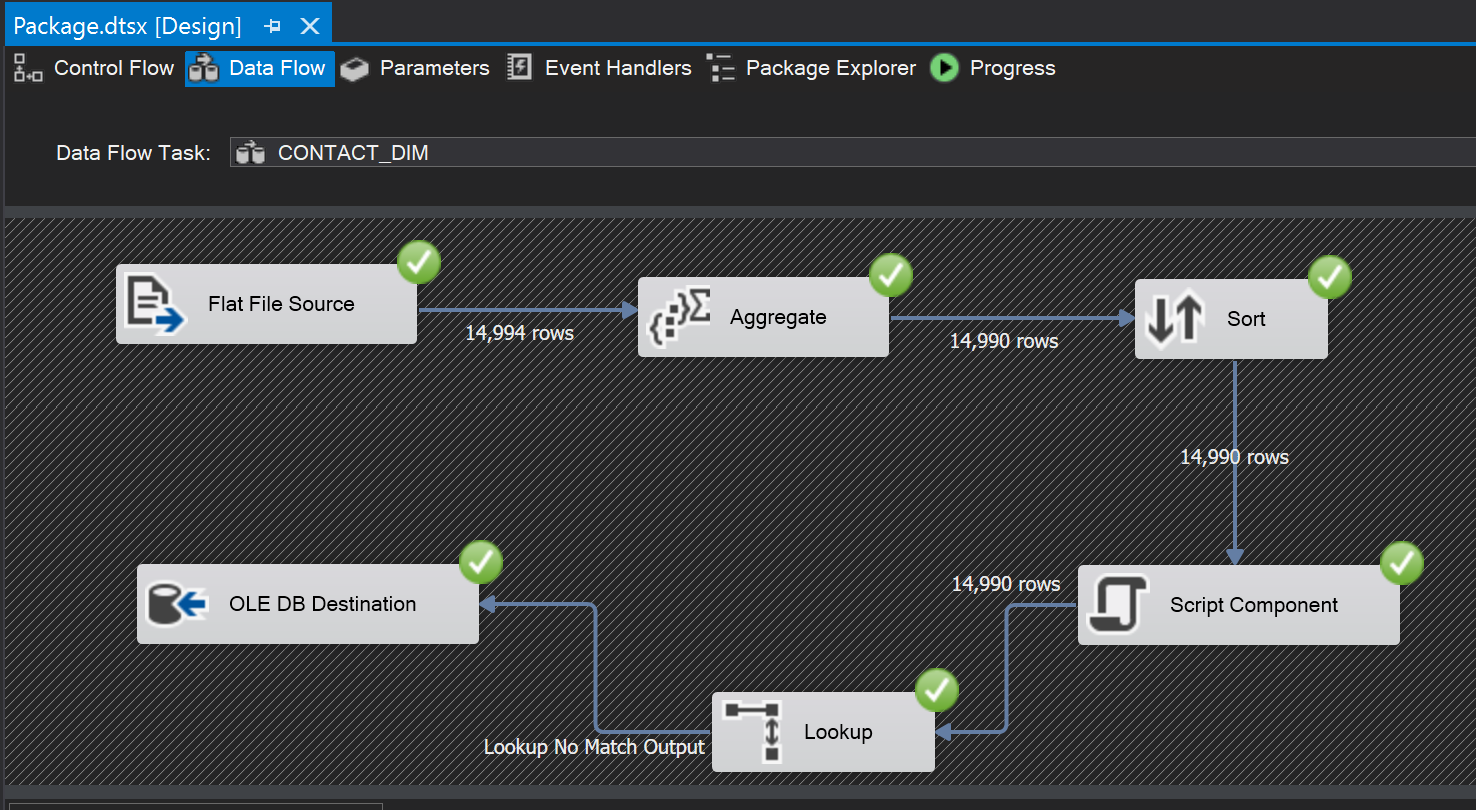
* 4 DIM VÀ 1 FACT

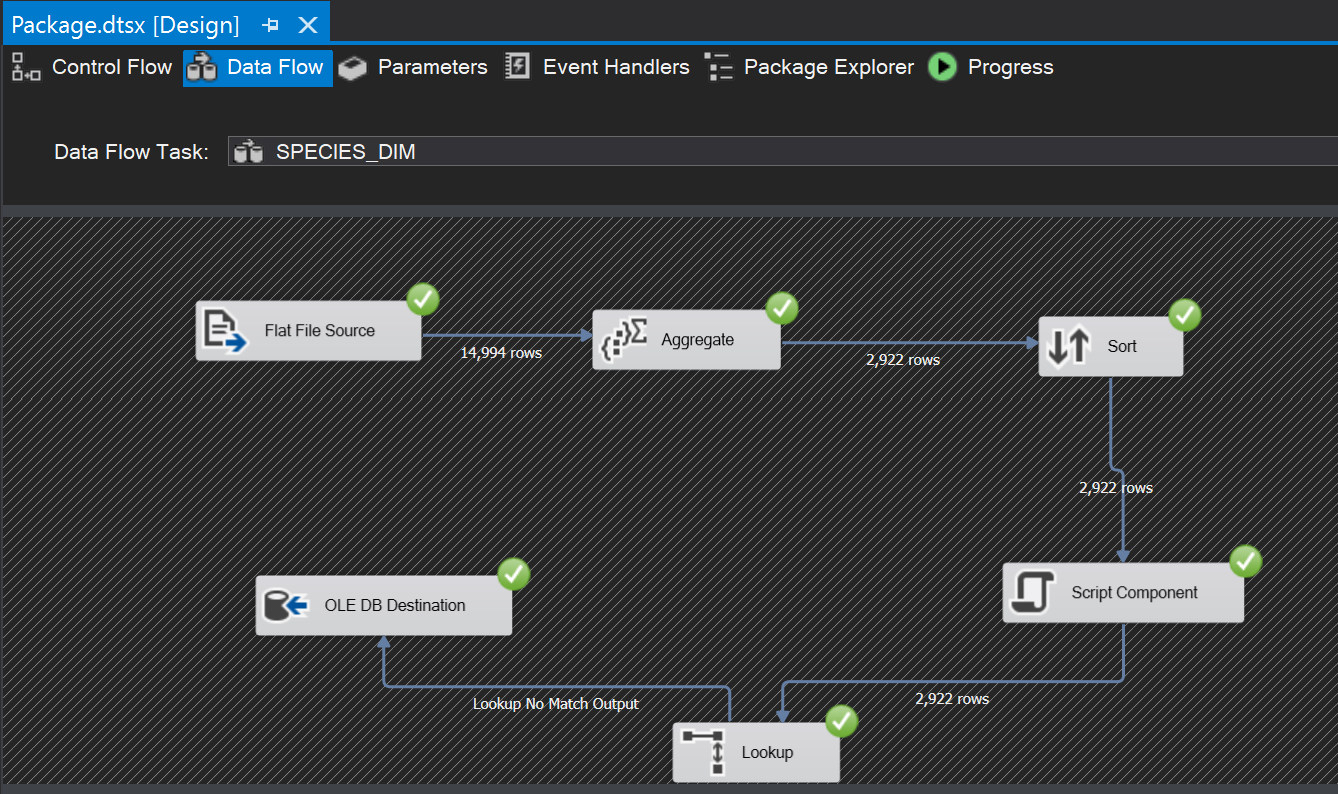


* NHẬP DATA TỪ FILE CSV NGUỒN. GROUP CÁC CỘT THUỘC TÍNH, SẮP XẾP LOẠI GIÁ TRỊ TRÙNG, TẠO KHÓA CHÍNH CHẠY TỪ 1. LOOKUP VỚI DỮ LIỆU TRONG BẢNG TƯƠNG ỨNG Ở SQL, NẾU KHÔNG TRÙNG SẼ IMPORT TIẾP VÀO BẢNG

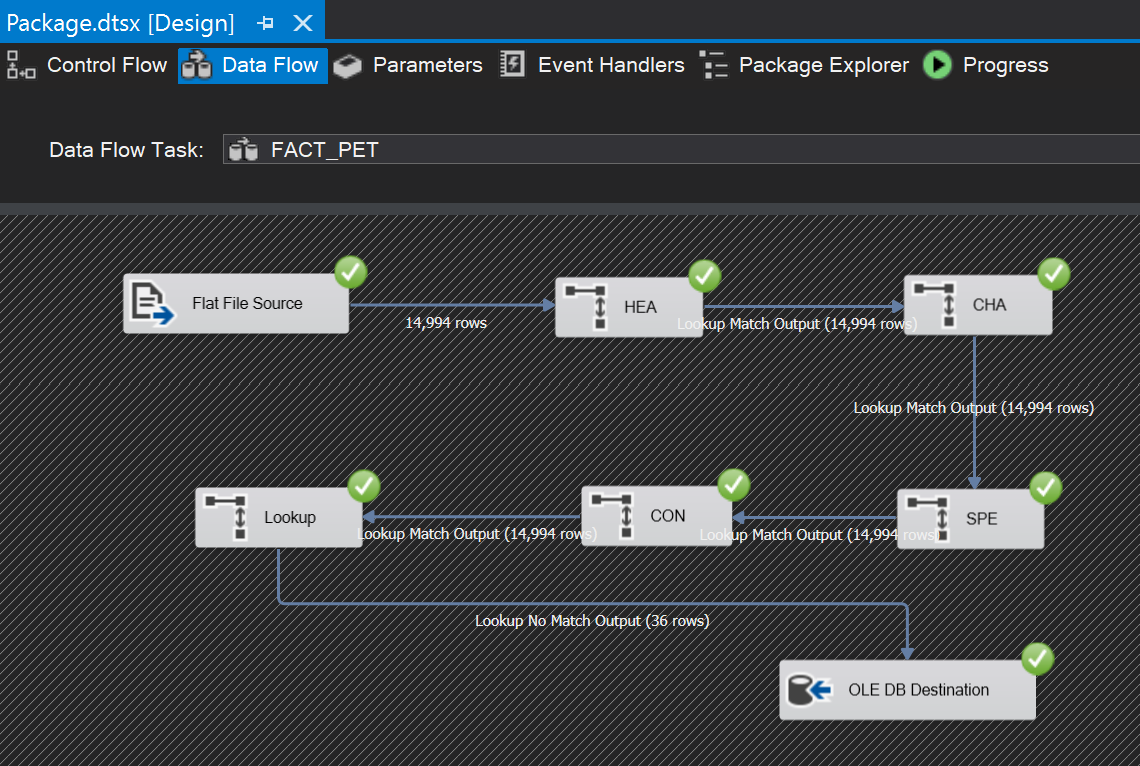








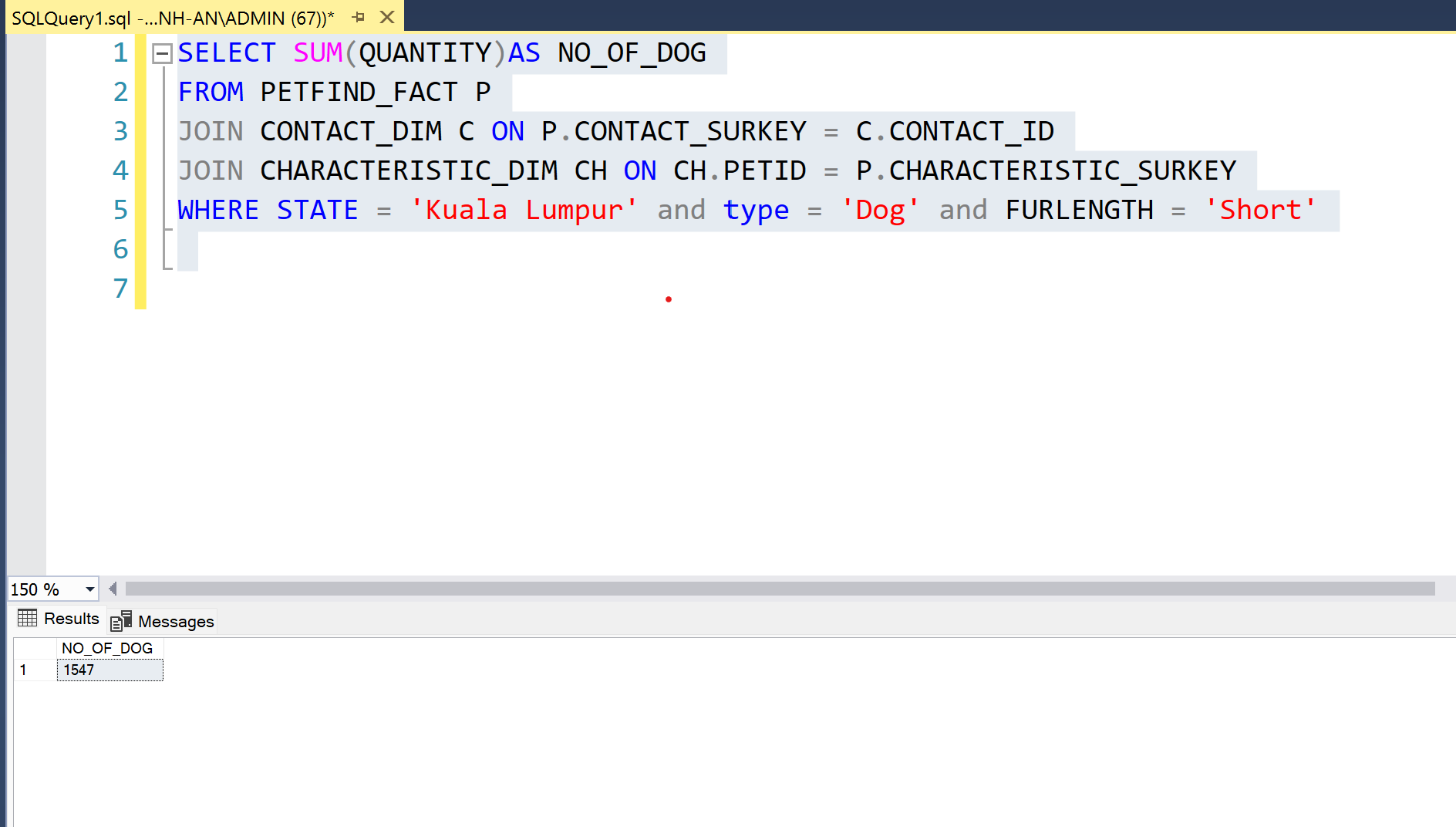
* THỰC HIỆN LOOKUP 4 LẦN TƯƠNG ỨNG VỚI 4 BẢNG DIM, SAU ĐÓ LOOKUP VỚI BẢNG FACT TRONG SQL ĐỂ CHỈ LẤY NHỮNG GIÁ TRỊ MỚI VÀ IMPORT VÀO FACT TABLE TRONG SQL



## **Câu lệnh SQL cho các truy vấn nghiệp vụ:**

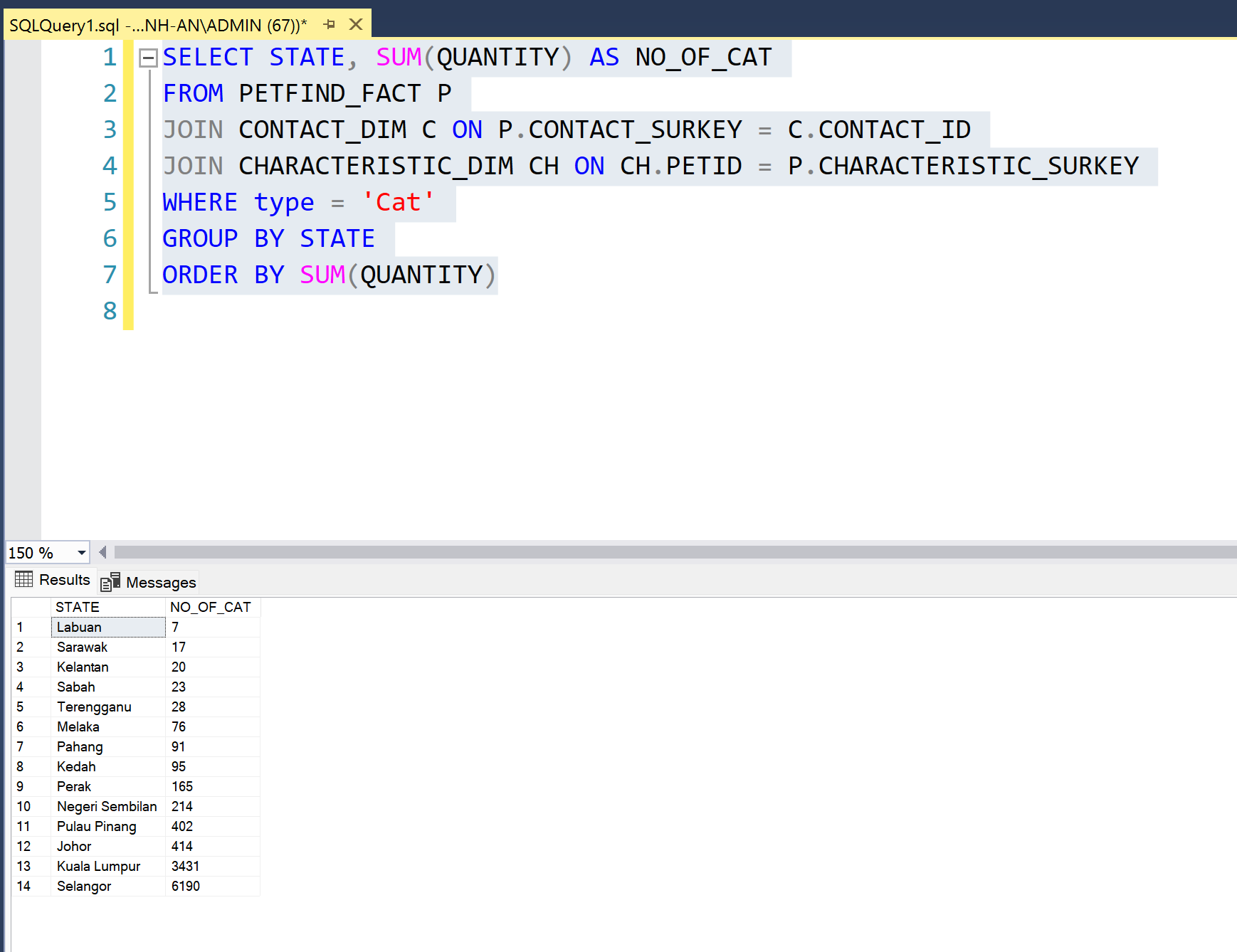
* **Có bao nhiêu con chó ở Kuala Lumpur có lông ngắn**

1. SELECT SUM(QUANTITY)AS NO\_OF\_DOG
2. FROM PETFIND\_FACT P
3. JOIN CONTACT\_DIM C ON P.CONTACT\_SURKEY = C.CONTACT\_ID
4. JOIN CHARACTERISTIC\_DIM CH ON CH.PETID = P.CHARACTERISTIC\_SURKEY
5. WHERE STATE = 'Kuala Lumpur' and type = 'Dog' and FURLENGTH = 'Short'



* **Tổng số mèo theo từng bang theo thứ tự tăng dần:**

1. SELECT STATE, SUM(QUANTITY) AS NO\_OF\_CAT
2. FROM PETFIND\_FACT P
3. JOIN CONTACT\_DIM C ON P.CONTACT\_SURKEY = C.CONTACT\_ID
4. JOIN CHARACTERISTIC\_DIM CH ON CH.PETID = P.CHARACTERISTIC\_SURKEY
5. WHERE type = 'Cat'
6. GROUP BY STATE
7. ORDER BY SUM(QUANTITY)

****

* **Số lượng mèo và chó dưới hoặc bằng 1 tuổi :**

1. SELECT TYPE, SUM(QUANTITY) AS NUMBER
2. FROM PETFIND\_FACT P
3. JOIN CONTACT\_DIM C ON P.CONTACT\_SURKEY = C.CONTACT\_ID
4. JOIN CHARACTERISTIC\_DIM CH ON CH.PETID = P.CHARACTERISTIC\_SURKEY
5. WHERE AGE = 1
6. GROUP BY TYPE
7. ORDER BY SUM(QUANTITY)

