Case Study: Registered Cats and Dogs in Greater Dandenong

Link: https://github.com/gabywu/Case-Study-Pet-Greater-Dandenong

Section

1. Overview
2. Objective
3. Data sources
4. Data Cleaning
5. Standardisation
6. Exploratory Data Analysis (EDA)
7. Key Insights
8. Summary
9. Recommendations

1. Overview

This project analyses pet registration data in Greater Dandenong to uncover trends in breed, gender, and geographic distribution. The aim is to inform council decision-making on future funding related to animal management services.

2. Objective

Stakeholders posed the question:

“What are the most common dog and cat breeds in Greater Dandenong?”

This case study focuses on breed and locality trends, identifying underrepresented breeds across suburbs and highlighting the most commonly registered breeds.

The database is intended to support council decisions regarding funding allocations for:

- Animal management services (e.g. off-leash parks)
- Pet adoption initiatives
- Veterinary and microchipping programs

Additional goals include:

- Estimating the population of registered cats and dogs in each suburb to assess microchipping compliance and shelter capacity
- Analysing dog breed sizes to inform off-leash park infrastructure and suitability for pet-friendly venues (e.g. cafés)

This analysis was conducted using Microsoft Excel, including the creation of a dashboard using Excel's graphing tools.

3. Data sources:

- Registered Cats Dataset
- Registered Dogs Dataset

Source: City of Greater Dandenong pet registration data (data.gov.au) – licensed under CC BY 3.0 AU

Field	Description
Locality	Suburb within the City of Greater Dandenong
Postcode	Postcode of the suburb
Animal_Type	Type of animal (Cat or Dog)
Breed_Description	Breed (abbreviated)
Colour_Description	Colour (abbreviated)
Gender	Sex of the pet

As both CSV files had the same structure, I combined them into a single dataset for analysis.

4. Data Cleaning

I checked for duplicates using COUNTIFS, but none were found.

I identified 13 missing records, accounting for only 0.00118% of the total dataset. As there were no reliable external sources to fill these gaps, and the impact was negligible, I left them as-is.

5. Standardisation

I changed the text case of values in the Locality field from UPPERCASE to Title Case and filled in full names for the Gender field.

There were 23 entries with unknown gender, accounting for 0.00208% of the data. These were also left unchanged due to the lack of reliable sources.

Breed abbreviations for cats and dogs lacked consistency. Attempts to resolve them using online sources and AI tools (e.g. ChatGPT and CoPilot) produced inconsistent results, and manual interpretation could introduce error. Therefore, I retained the original data unless duplicate abbreviations were found that could confuse dashboard outputs.

6. Exploratory Data Analysis (EDA)

I focused on the following variables:

- **Locality**
- **Animal_Type**
- **Breed_Description**

Key steps:

- Created new columns:
  - `Concat(type-breed)` to identify breed frequency
  - `Concat(type-loc-breed)` to analyse the most common breed in each suburb
- Used PivotTables to count:
  - Overall frequency of each breed
  - Distribution of breeds by locality
- Filtered for the most frequent breed per locality, separating cats and dogs into their own summaries.

In the **Working Sheet**, I used PivotTables to count:

- Unique breeds per animal type
- Gender counts
- Total counts of each animal type (to explore possible correlations)

In the **EDA Sheet**, I used PivotTables to:

- Count `Concat(type-breed)`
- Sort by `Animal_Type`, `Locality`, and `Breed_Description`
- Extract and evaluate the most frequent breed per suburb using an `Is_Max` formula to identify ties or dominant breeds

7. Key Insights

- **Cats:** Domestic Short Hair represents **52%** of all registered cats — by far the most common.
- **Dogs:** The majority of registered dog breeds are small to medium-sized.
- **Dogs show greater breed diversity** than cats.
- **Maltese (dogs)** and **Domestic Short Hair (cats)** have significantly higher frequencies compared to other breeds, both overall and by suburb.
- A majority of **unique breeds** appear only **once** in the dataset.

8. Summary

- The most common **dog breed** is the **Maltese**, making up **10%** of dog registrations.
- The most common **cat breed** is the **Domestic Short Hair**, accounting for **52%** of cat registrations.
- The data suggests that residents of Greater Dandenong tend to prefer **smaller pets**, likely due to housing type, lifestyle, or breed preferences.
- These breeds may also be more prevalent in **veterinary clinics** or **pet adoption centres**.

It is important to note a **potential undercoverage bias** — unregistered animals, including strays or those housed in shelters, are not captured in this data.

*This analysis is based on pet registration data dated **09/08/2023**, the most recent dataset available at the time. While pet ownership trends are relatively stable year to year, insights should be interpreted with consideration of any demographic or policy changes since 2023.*

9. Recommendations

- Improve **microchipping compliance** to ensure pet safety and improve the chances of reuniting lost pets with their owners.
- Enhance **fencing** in off-leash parks, especially for small breeds that may escape or be at risk from wildlife.
- Increase public awareness of **lesser-known breeds**, potentially through a specialist workforce focused on breed-specific care, adoption support, and community outreach.
- Enhance **designated spaces** and provide dog-friendly amenities at pet-friendly cafés.