Project NYC Dog names

Anna
Xinly
Katya

Project Overview

This project analyzes **dog name trends** in NYC using available **datasets** and external breed **API** data. The aim is to explore the factors influencing naming trends, the popularity of certain dog names over time, and potential connections between a dog's name, its breed, and behavior.

NYC CSV dataset:

	AnimalName	AnimalGender	AnimalBirthYear	BreedName	ZipCode	LicenseIssuedDate	LicenseExpiredDate	Extract Year
0	PAIGE	F	2014	American Pit Bull Mix / Pit Bull Mix	10035.0	09/12/2014	09/12/2017	2016
1	YOGI	M	2010	Boxer	10465.0	09/12/2014	10/02/2017	2016
2	ALI	M	2014	Basenji	10013.0	09/12/2014	09/12/2019	2016
3	QUEEN	F	2013	Akita Crossbreed	10013.0	09/12/2014	09/12/2017	2016

The dog API:

	breedname	lifespan	temperament	breedgroup	weight	height	bredfor	origin
0	affenpinscher	10 - 12 years	Stubborn, Curious, Playful, Adventurous, Activ...	Toy	3 - 6	23 - 29	Small rodent hunting, lapdog	Germany, France
1	afghan hound	10 - 13 years	Aloof, Clownish, Dignified, Independent, Happy	Hound	23 - 27	64 - 69	Coursing and hunting	Afghanistan, Iran, Pakistan
2	african hunting dog	11 years	Wild, Hardworking, Dutiful	None	20 - 30	76	A wild pack animal	
3	airedale terrier	10 - 13 years	Outgoing, Friendly, Alert, Confident, Intellig...	Terrier	18 - 29	53 - 58	Badger, otter hunting	United Kingdom, England

Hypotheses:

1. The **Popularity of Names** shifts over time and is shaped by **trends** in popular culture.
2. A **Pattern** Exists Between **Dog Personalities and Popular Names.**
3. The majority of **Dog Names are 4 characters** long.
4. Some **Names** are popular for both **male and female** dogs.
5. **Certain names** are more commonly associated with **specific breeds.**
6. There is a pattern between **Dog Size** and Popular **Names.**

Data Wrangling and Cleaning

NYC database:

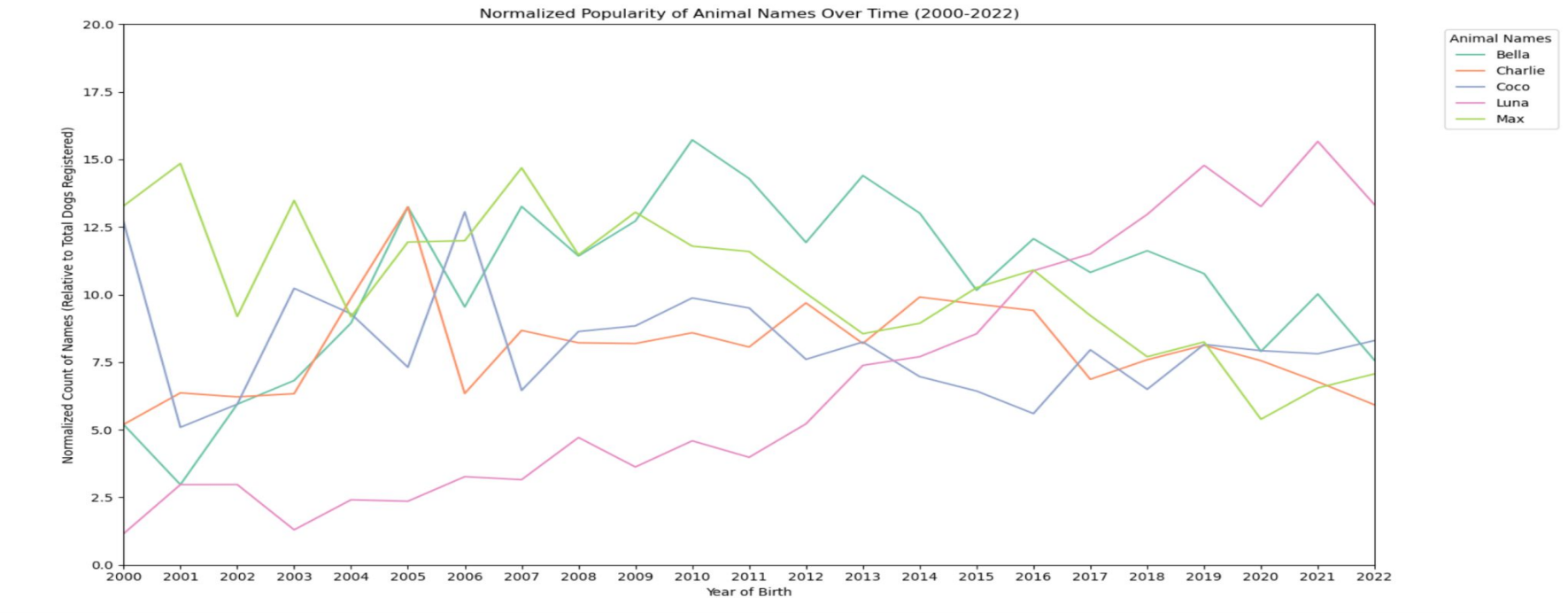
- Formatting issues.
- Dropped null values that were common for all columns.
- Dropped duplicate values.
- Dropped not relevant columns.
- Grouped breed names into relevant breed groups.

API:

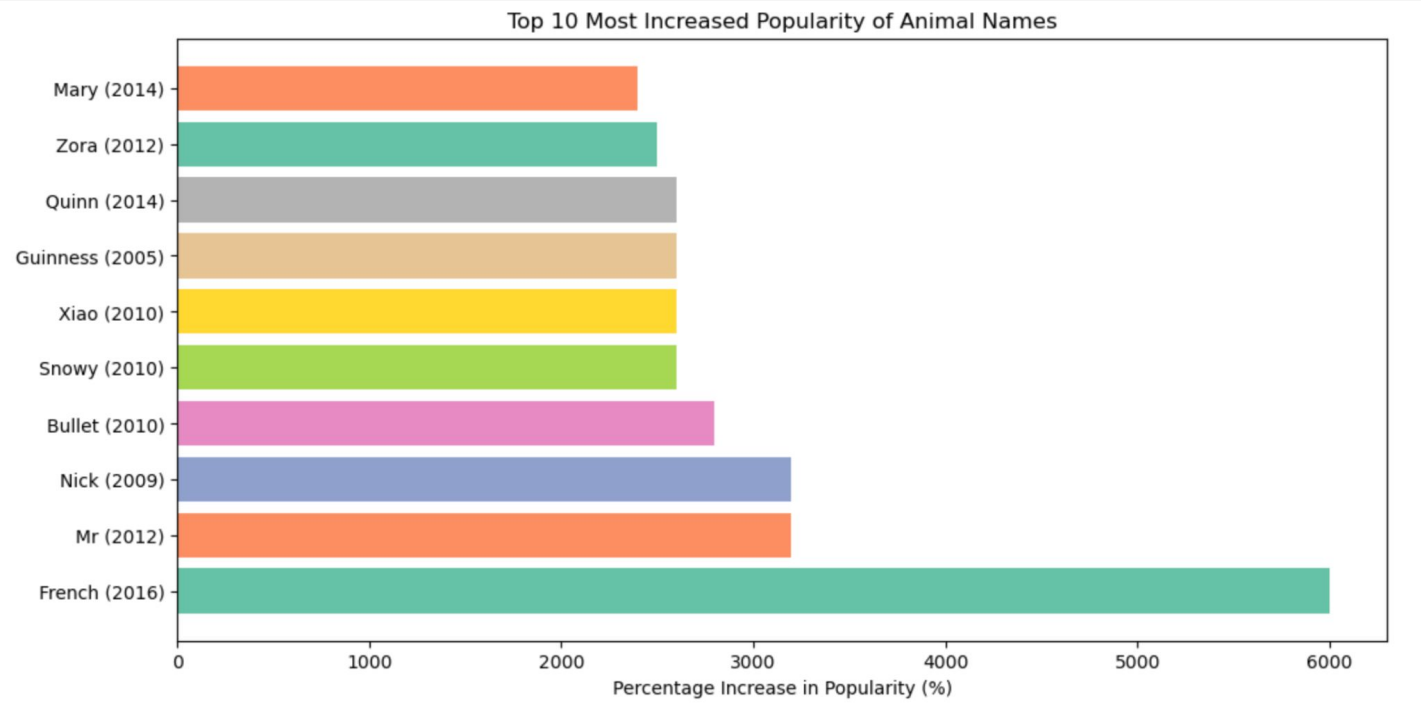
- Analysed the temperament values and aggregated them to finally ungroup them into new columns.
- Not every dog breed from the API was present or matched the ones on the database, therefore we had to group and redefine them before merging.

All data was **merged** taking the common column of **breed group.**

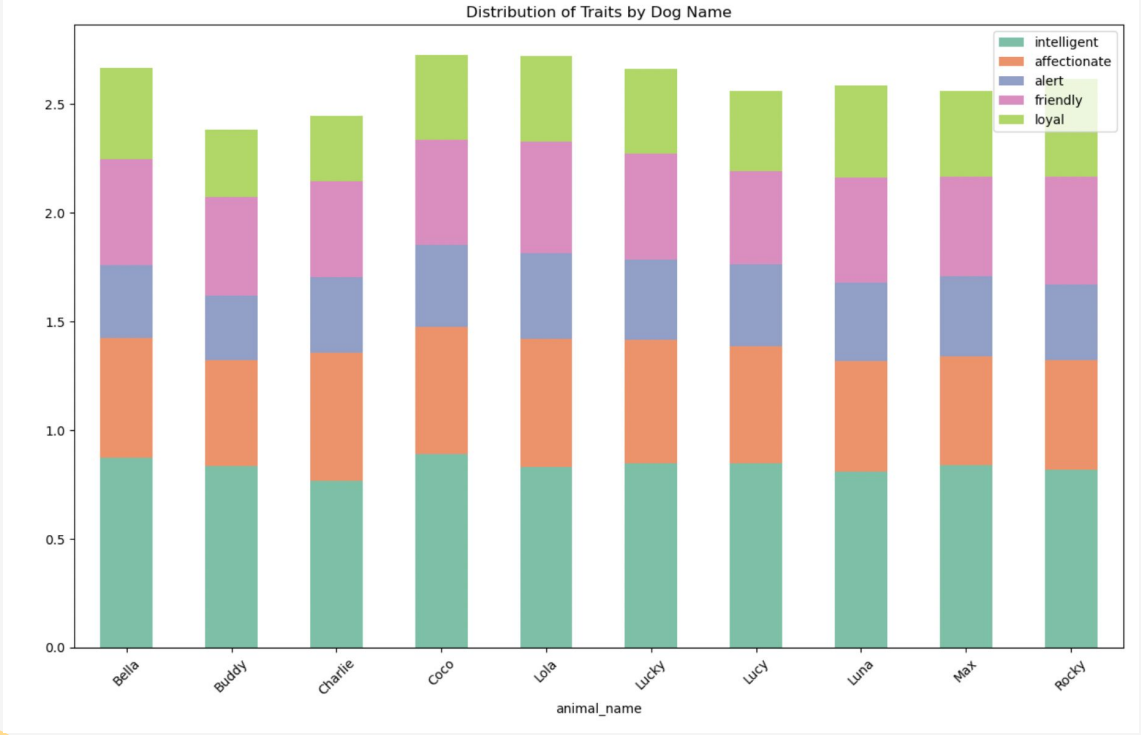
1. Hypothesis: The popularity of names shifts over time and is shaped by trends in popular culture

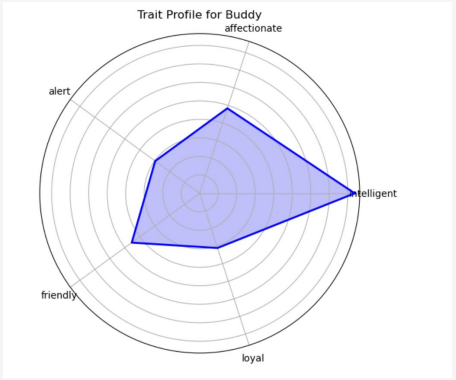


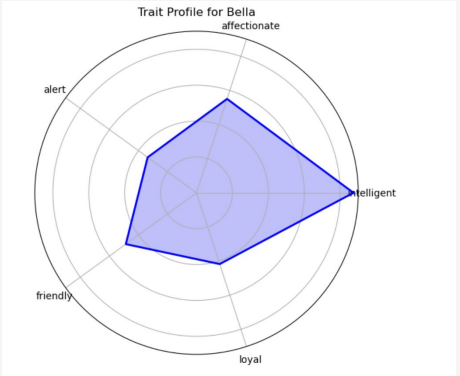
1. Hypothesis: The popularity of names shifts over time and is shaped by trends in popular culture



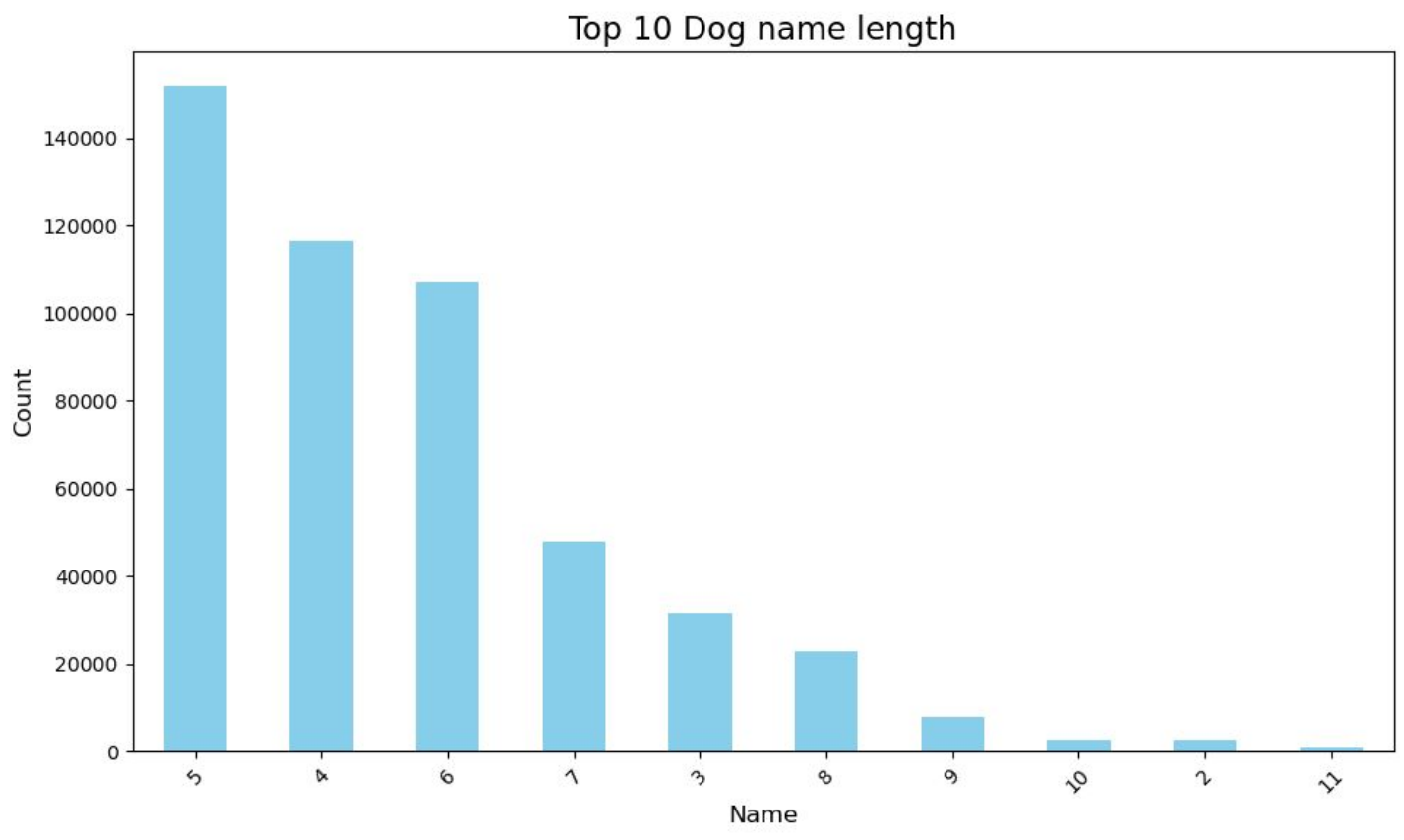
2. Hypothesis: A pattern exists between dog personalities and popular names



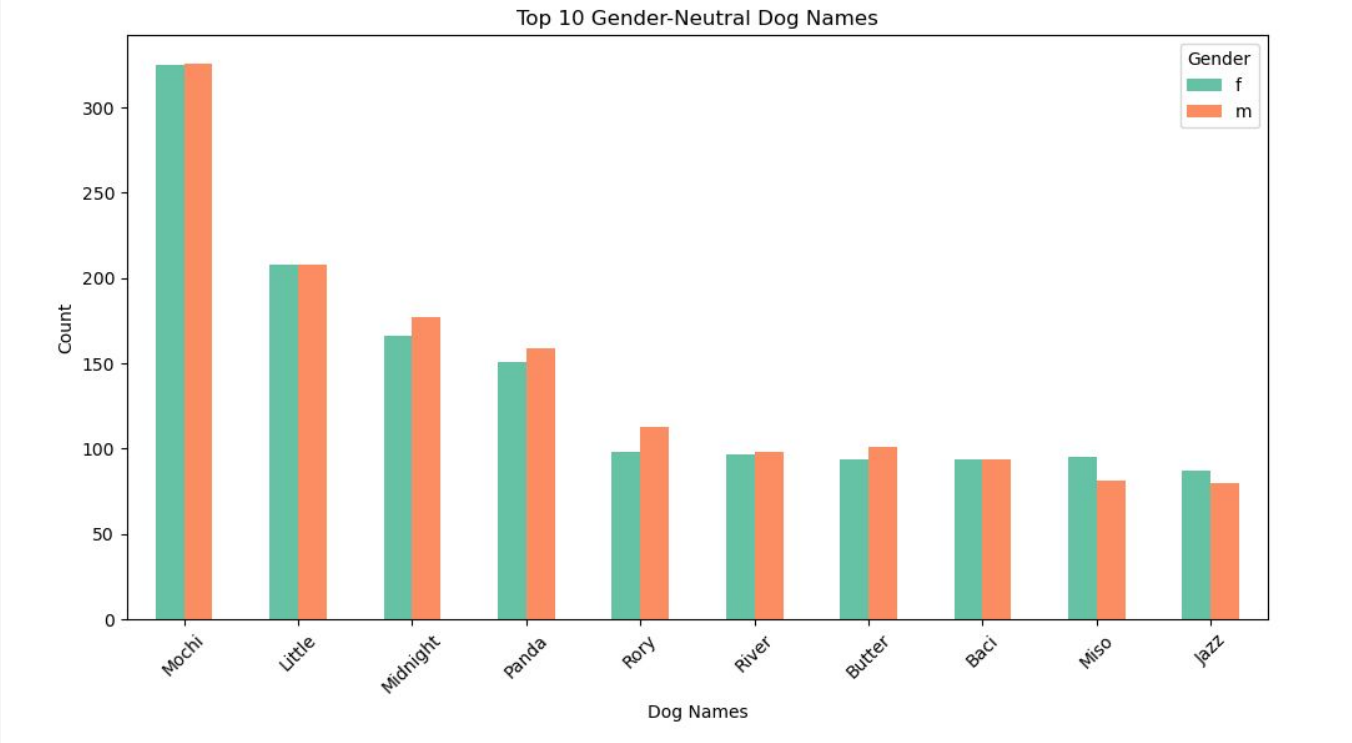




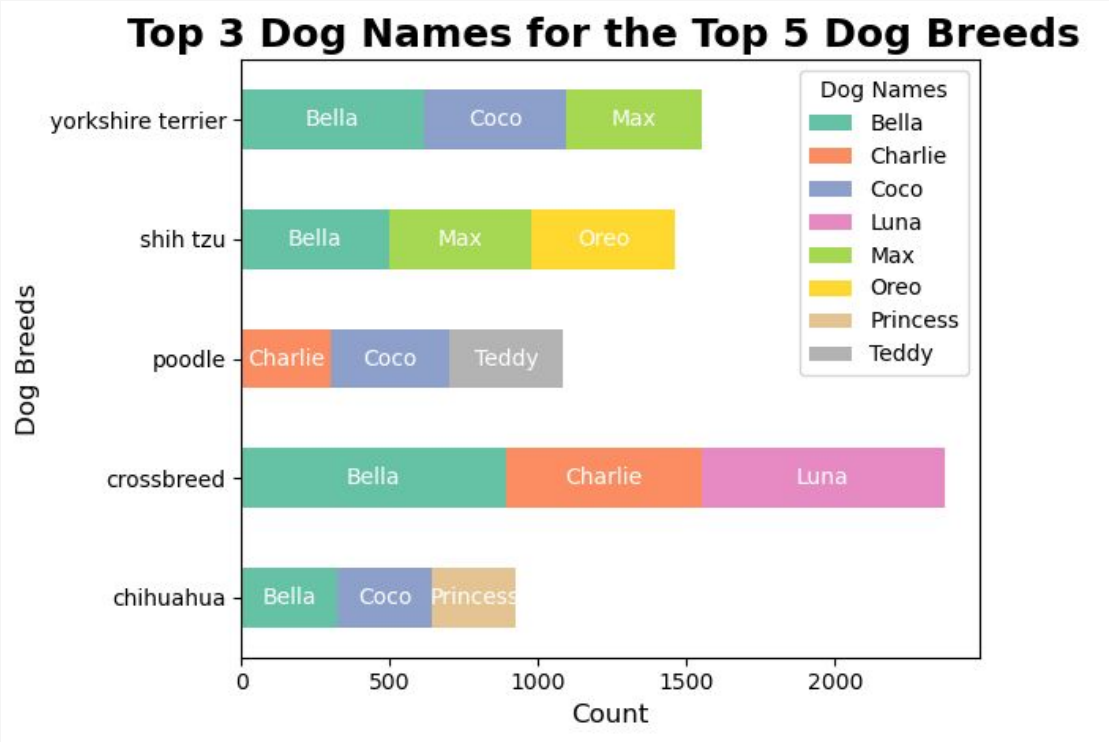
3. Hypothesis: The majority of dog names are 4 characters long



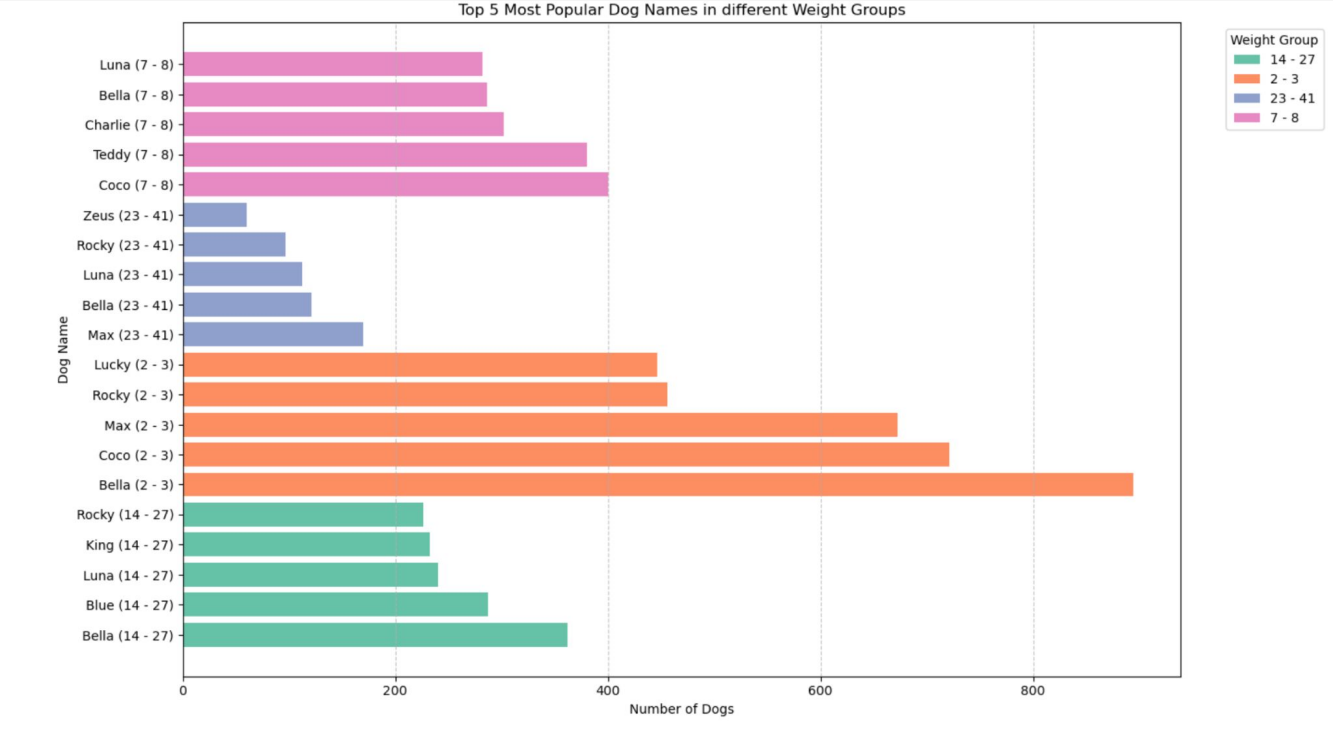
4. Hypothesis: Some popular names are gender neutral



5. Hypothesis: Certain names are more commonly associated with specific breeds



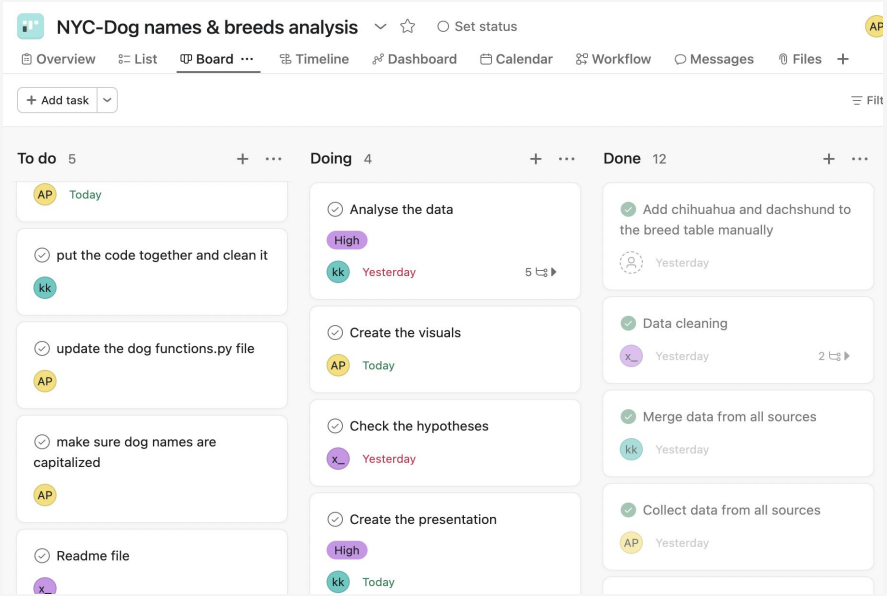
6. Hypothesis: There is a pattern between dog size and popular names



Teamwork & Project Management

We established a workflow in **Asana** from the start, making a few adjustments along the way, but it remained largely **consistent** throughout the project.

It was a major improvement from previous projects.



Major Obstacles

The **greatest challenge** we faced was cleaning all the datasets and organizing the different animal breeds. The data was **not consistently formatted**, and we had thousands of rows to categorize.

We also realised that the **variety of names** in the dataset made it challenging to find patterns, and we shifted to **focus on most popular names and breeds.**

Also working with **git** proved to be difficult at the end of the project when we were merging our code together.

Conclusion and Insights

Some names, like "**Bella," "Luna" and "Charlie**," were popular across multiple breeds, sizes and timeframes.

While we found no pattern between dogs' personalities and names, we observed **correlations between names and breeds**, as well as between **names and gender and size**.

When **naming dogs**, people seem influenced more by cultural trends than the dog's traits. They also tend to prefer **shorter names**.

These insights on popular names and breeds can be **valuable** for **targeted marketing and sales** in the pet industry.

THANK YOU!

Project NYC Dog Names

Anna
Xinly
Katya