# New York City Dogs

Python for Data Analytics

Capstone Project

# Introduction

The goal of this project is to analyze New York City dog licensing data to identify trends in dog breeds, gender distribution, popular names, and other characteristics across boroughs. Understanding these patterns is simply interesting for creating a portrait of an average New-Yorker-Dog. As well as it can help researchers, pet owners, and policymakers assess population trends, breed popularity, and licensing behavior.

# Data Sources

The data provider is **New York City Department of Health and Mental Hygiene.** The dataset was sourced from NYC Open Data web page https://opendata.cityofnewyork.us/ (last update February 2025)

**Total Rows: Over 722,000 records**

**Columns Include:** Dog's name, Gender M/F, Birth year, Breed type, ZIP code of registration, Date the license was issued, Expiration date of the license, Extract date

# Methodologies

## I used to analyze the data :

**Pandas**:

**- Data cleaning, filtering, aggregations**

**Matplotlib & Seaborn**:

**- Visualizations**

**NumPy**:

**- Numerical transformations (age calculations)**

**Data Cleaning:**

**- Removed missing values**

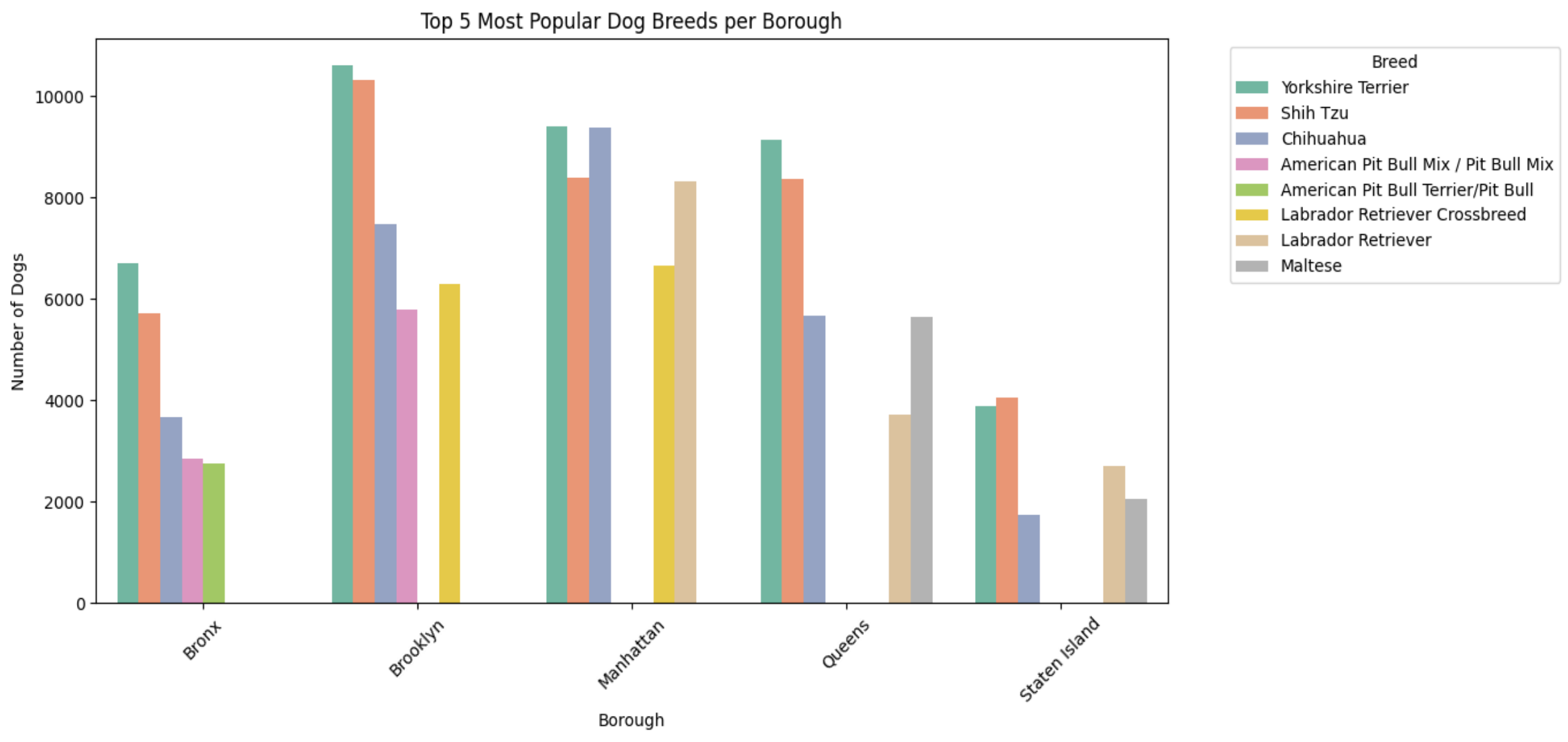
**- Filtered unrealistic birth years**

**Data Transformations:**

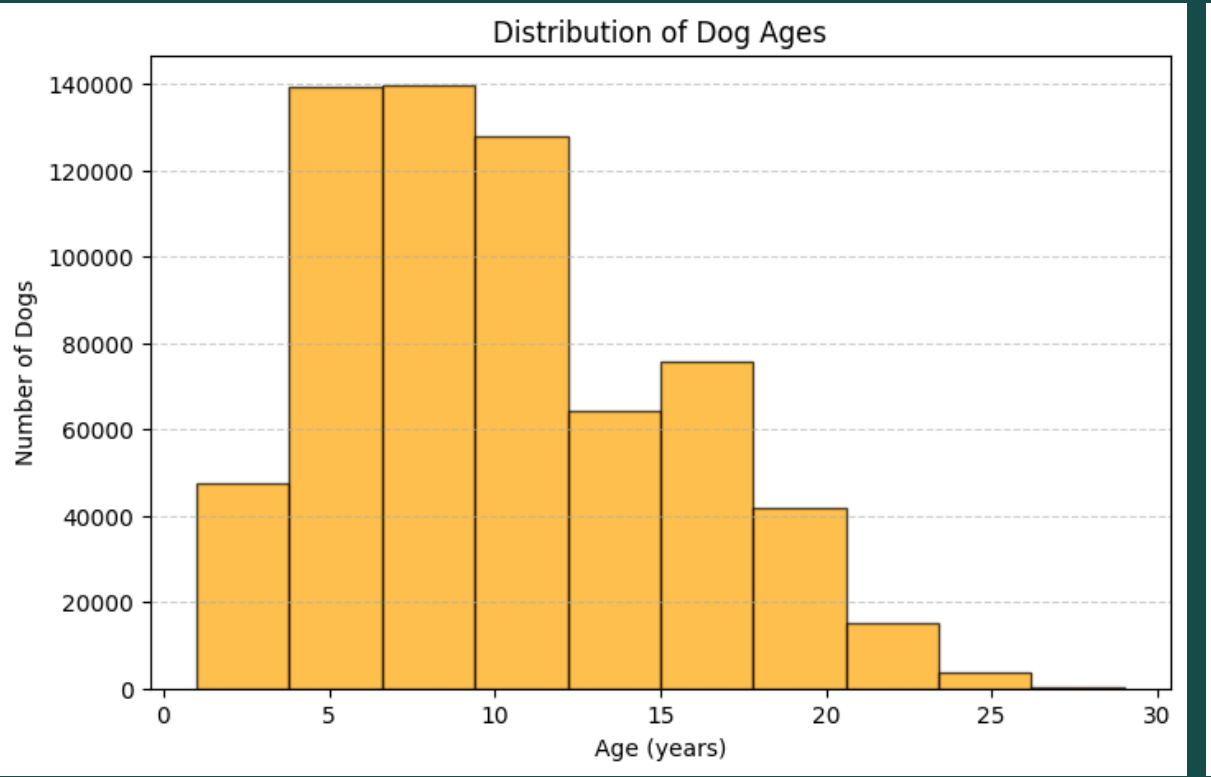
**- Grouped data into categories (breed)**

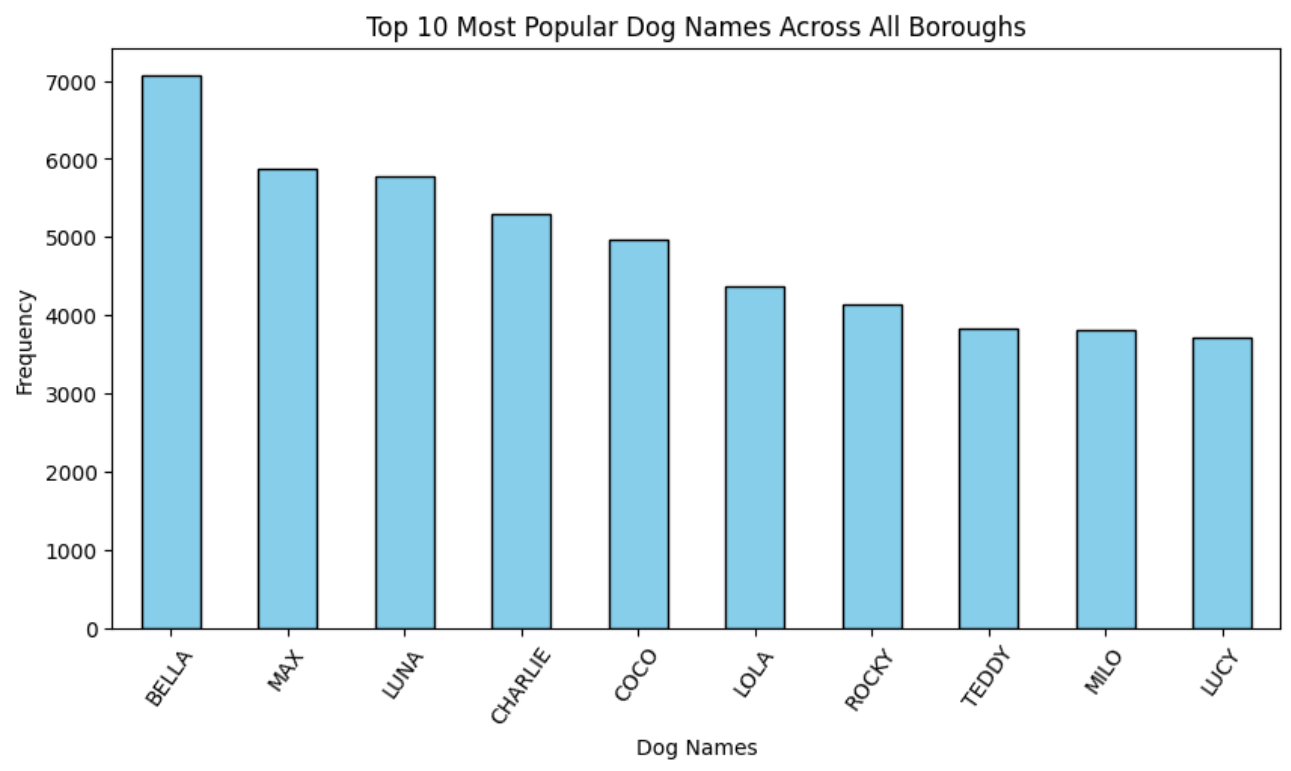
**- Computed proportion**

# The Most Popular Breeds per Borough



# Age Distribution & Common Names



# Conclusions

- The most popular breed varies by borough, but small-to-medium breeds dominate.
- Male dogs outnumber female dogs slightly across all boroughs (~55% male).
- The average dog age is around 10 years, aligning with typical pet aging trends.
- Unexpected insights: Certain boroughs show distinct preferences for specific breeds, influenced by factors like apartment size and community trends.

# Thank you

Photo Credit:

Instagram: Dogs of New York (@dogs_ny)