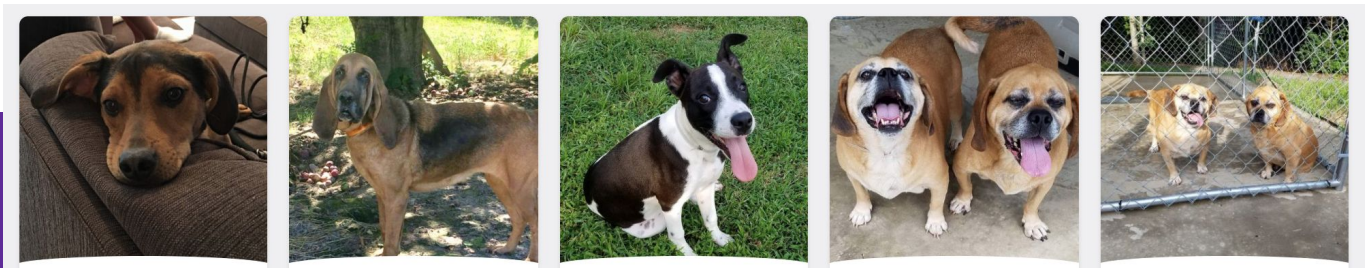


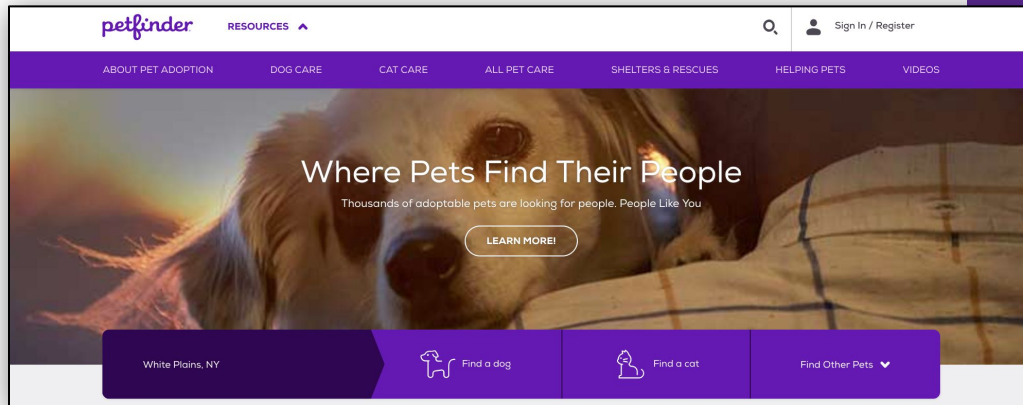
WebScraping Project:

petfinder



Erin Dugan • 08.08.2018 • NYC-Data Science Academy

Overview



petfinder.com

Largest online pet adoption website in North America

Almost 300,000 animals listed, including:

cats, dogs, fish, birds, horses, lizards, etc.

project objectives

Scrape subset of dog postings in NYC metro area

Explore distribution by age, breed, size, etc

Find adoption trends

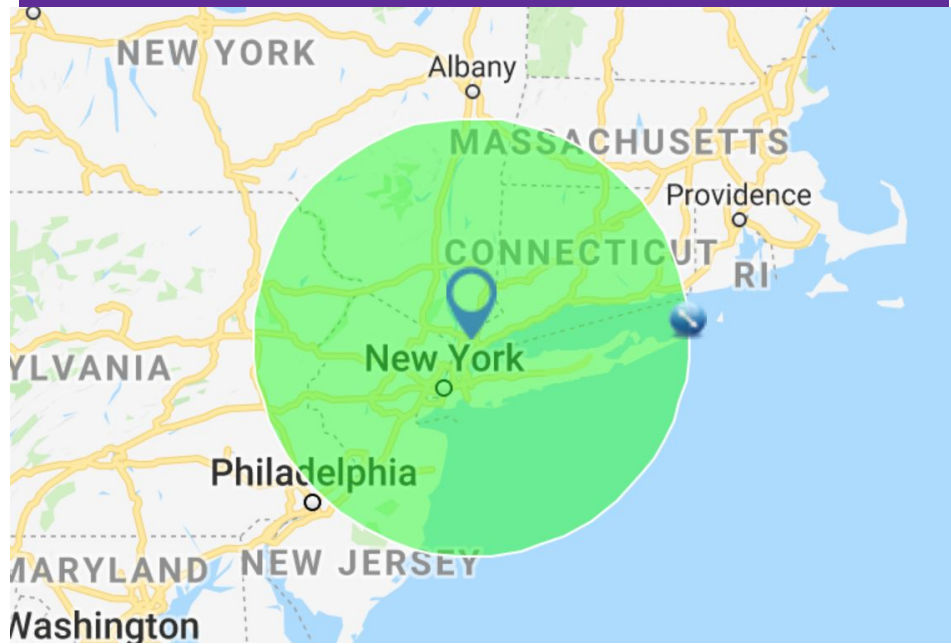
Examine characteristics of common dog breeds

Webscraping

8000+ dogs from
urban/suburban/rural areas
in 5 states

breed, age, size, health
posting date, location, contact info
brief description

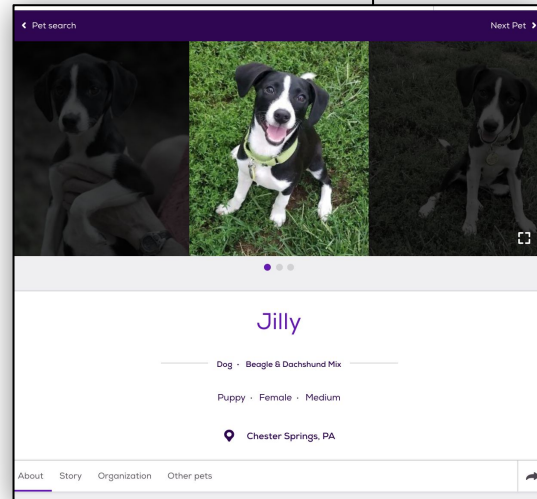
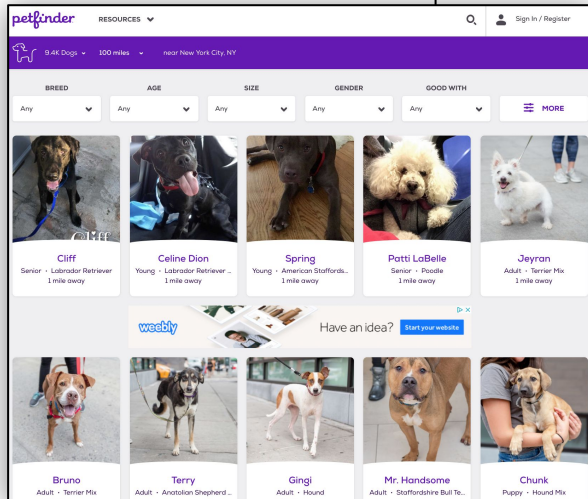
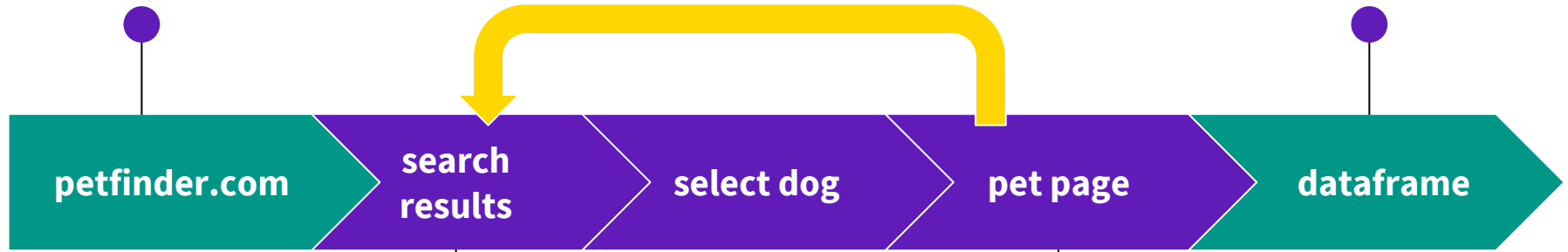
**All dogs available for adoption
within 100-mile radius of White
Plains NY**



Search: dogs within
100-mi of White Plains NY

Selenium: step through 40
dogs per page, 217 pages

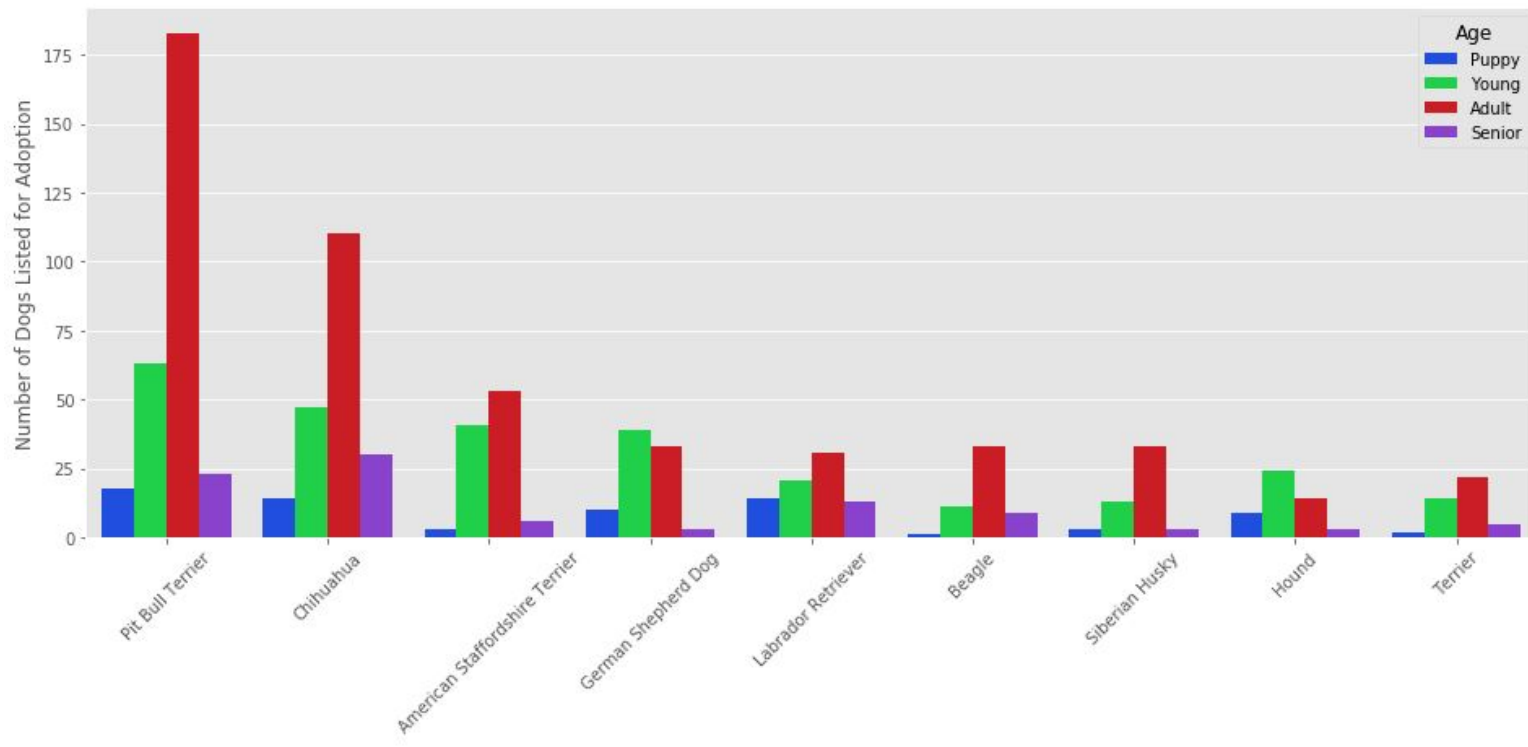
Python: clean & parse
data for analysis



Selenium & Xpath:
scrape visible &
hidden data within
HTML code

Demographics

Purebred dogs with most listings in NYC region



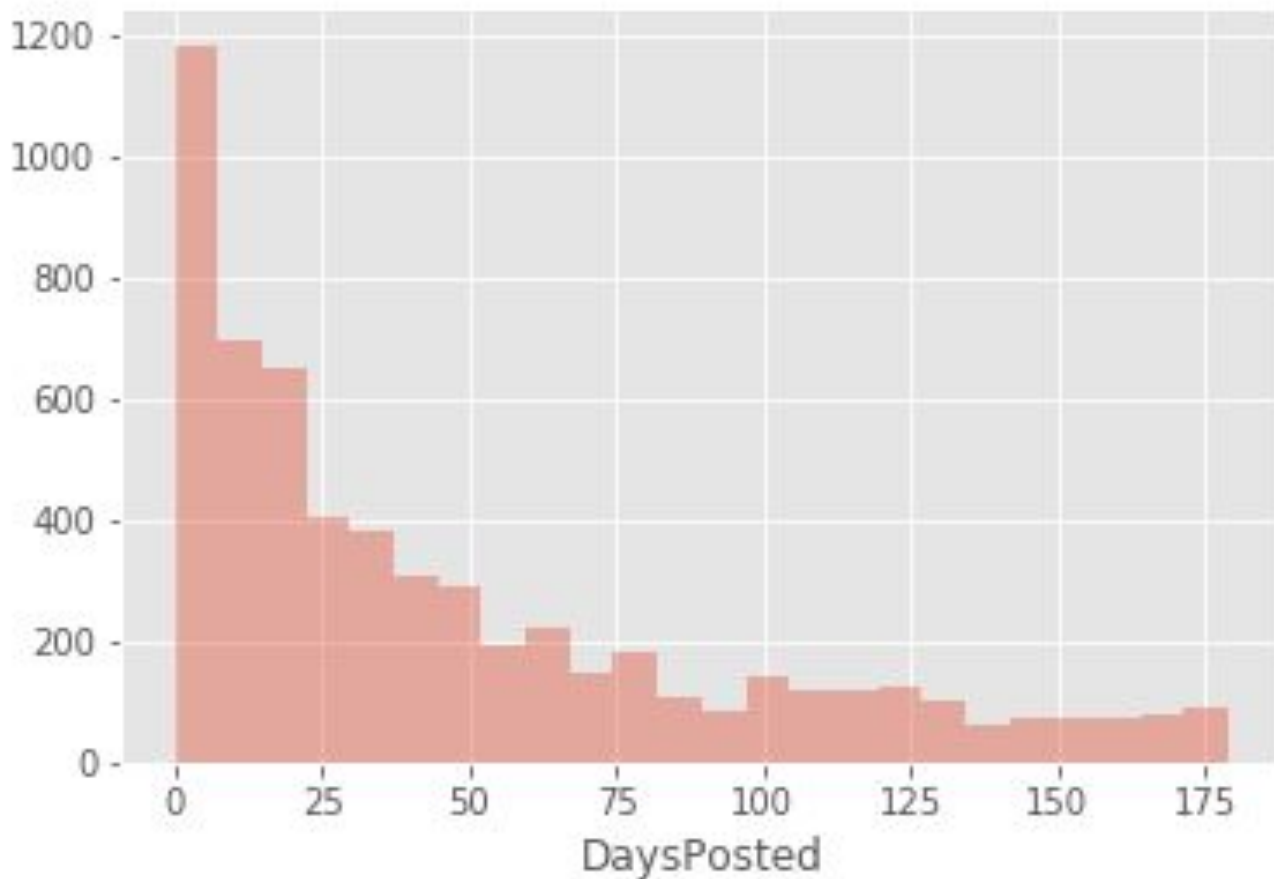
Adoption Trends

No adoption dates provided

Assume constant posting & adoption rate

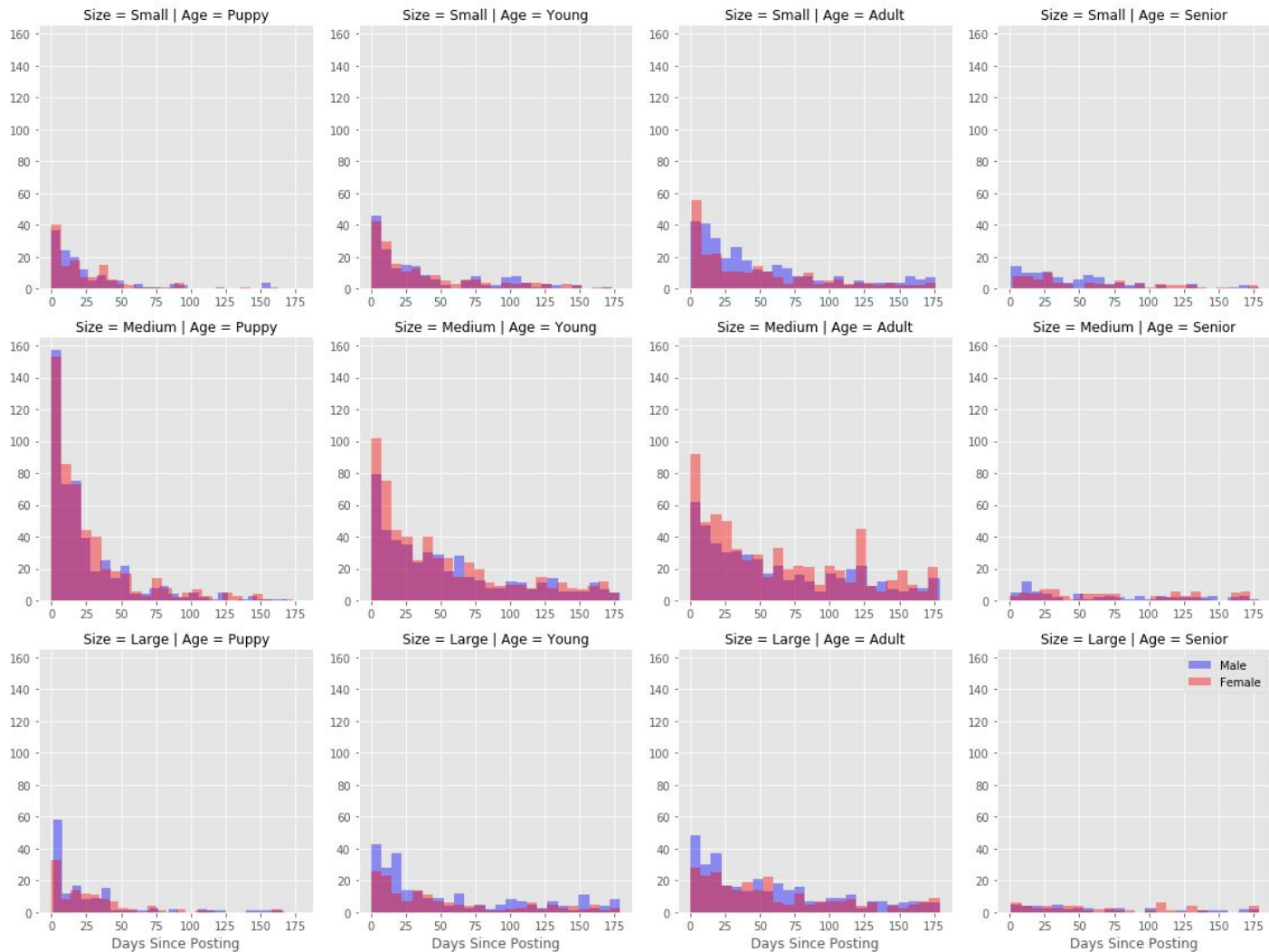
Number of days dogs have been listed on petfinder.com (before adoption)

6-month time period



Adoption Trends

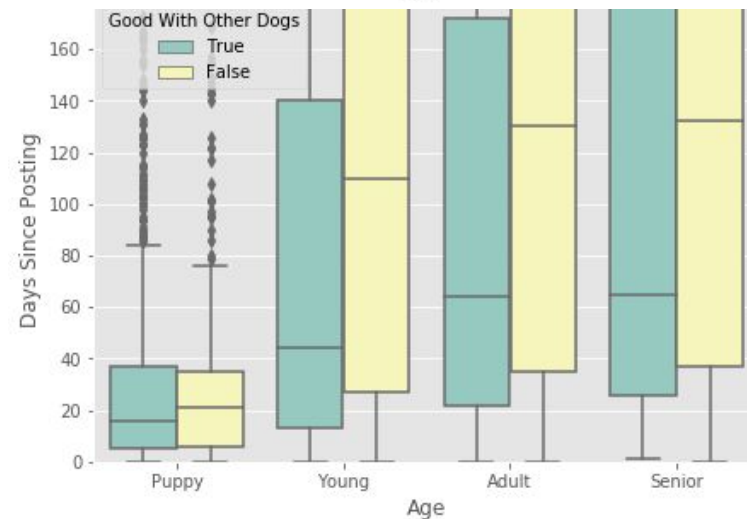
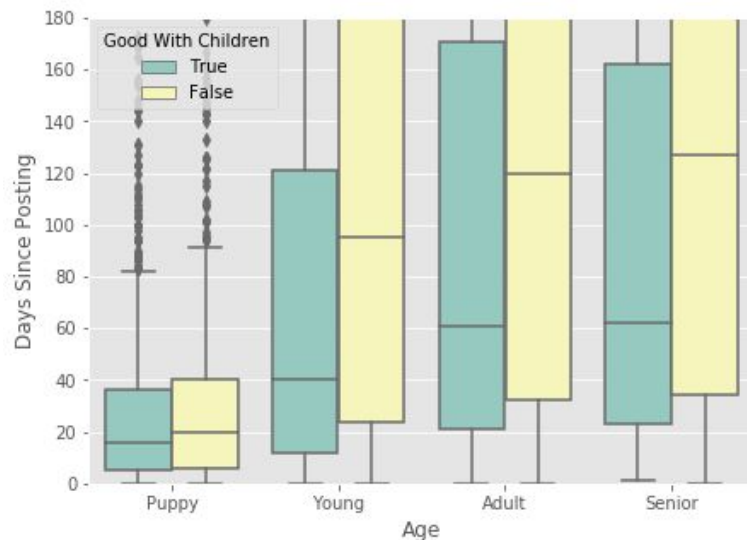
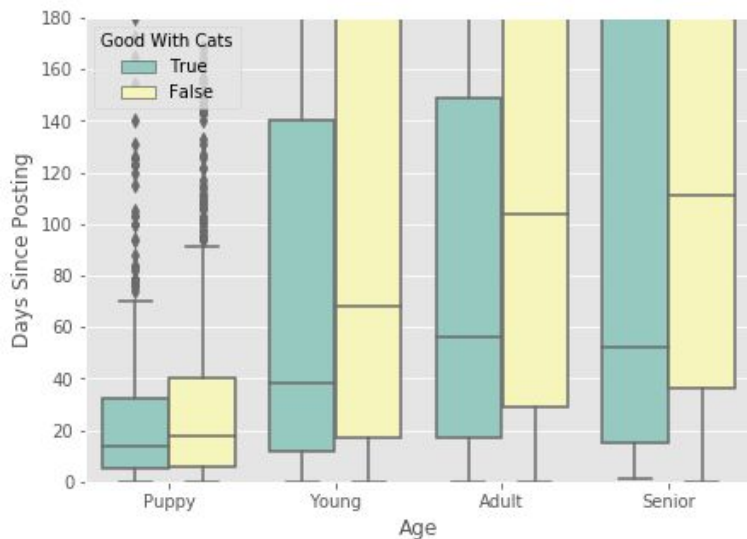
How quickly does each category get adopted?



Adoption Trends

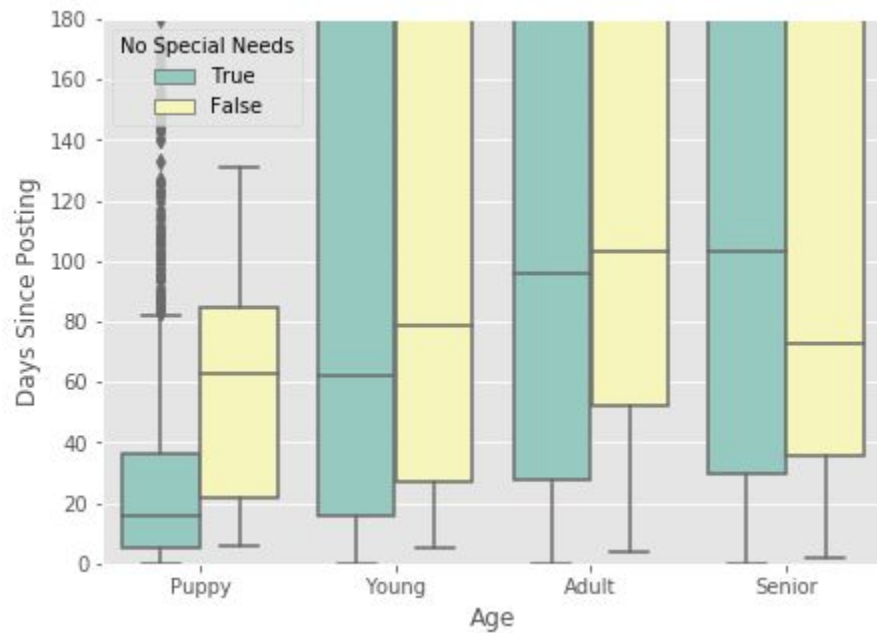
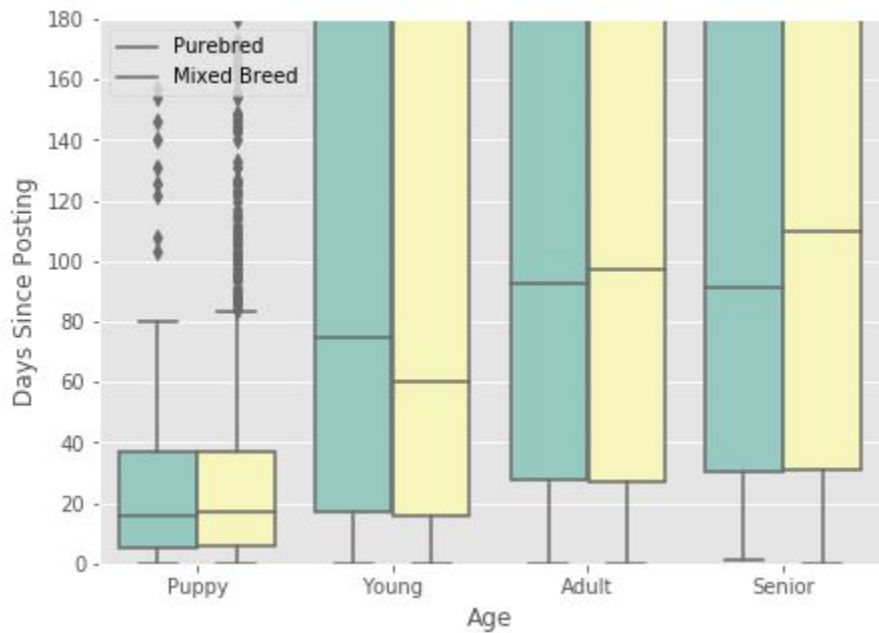
Median number of days posted, for all dogs:
55.5 days

What qualities are most desirable at each age?



Adoption Trends

What other factors increase adoption rate? Does it vary by age?



Wordclouds by Breed



Continued Work

Future Applications

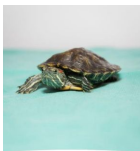
- Analyze other pet types & regional areas
- Evaluate most popular mixed dog breeds
- Help to find homes for pets evacuated from devastated areas due to fires, hurricanes, etc.
- Use similar analyses for other search sites: Dating apps, Employment search, etc.



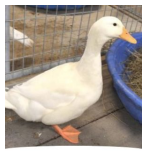
Yuri
Lionhead & Lop Eared Mix
44 miles away



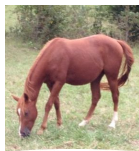
Waffles
Domestic Short Hair
26 miles away



Leah
Turtle
225 miles away



PUCK
Duck
932 miles away



Bingo
Quarterhorse
840 miles away



Atlas
Vietnamese Pot Bellied
440 miles away