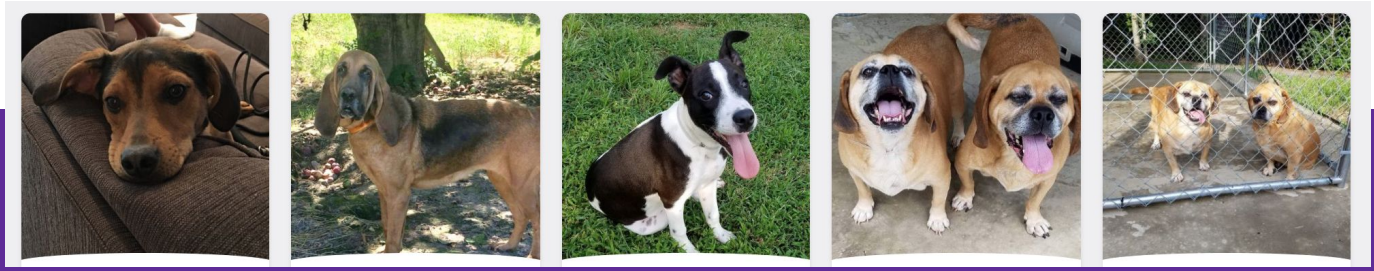
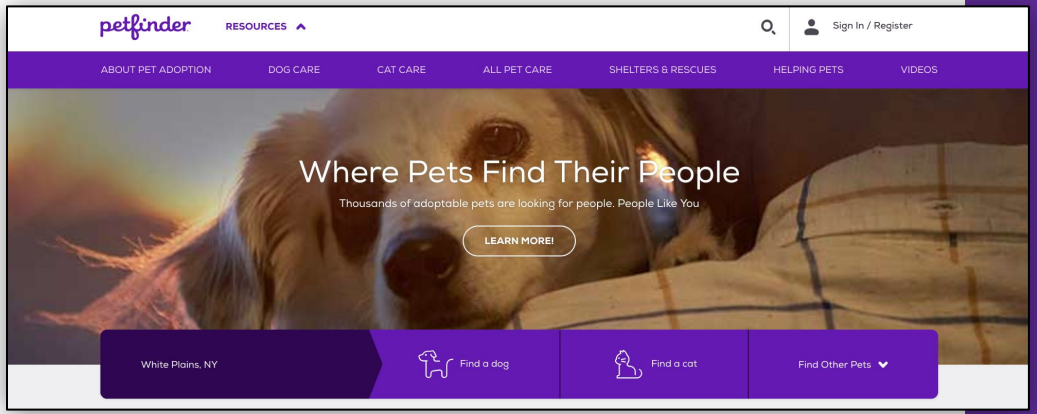
WebScraping Project:

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

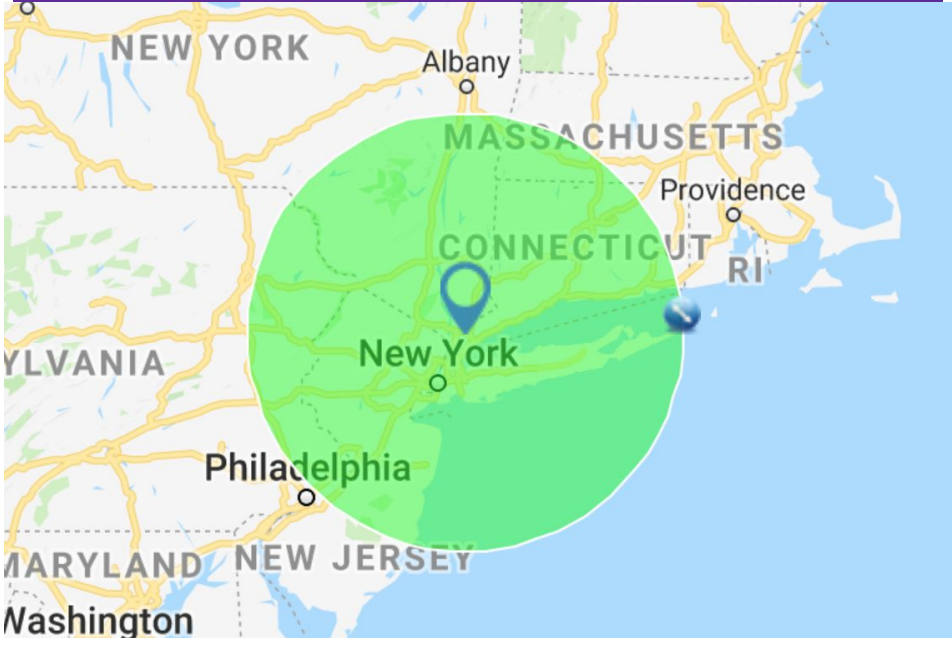
Examine characteristics of common dog breeds

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

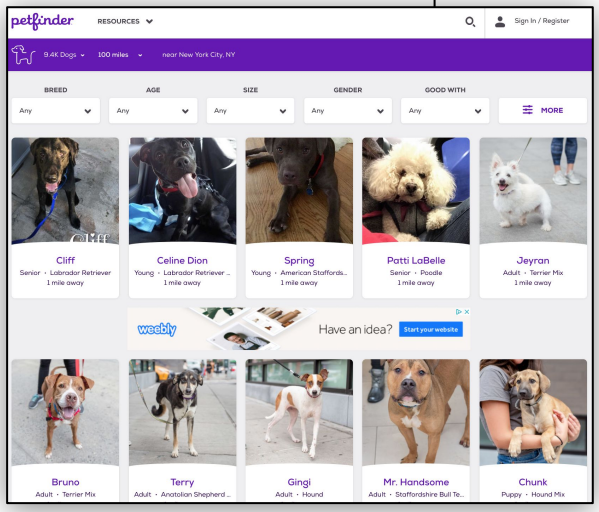


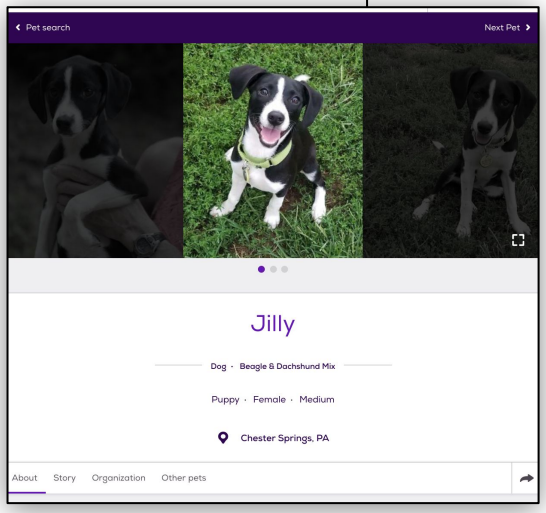
Search: dogs within 100-mi of White Plains NY

Selenium: step through 40 dogs per page, 217 pages

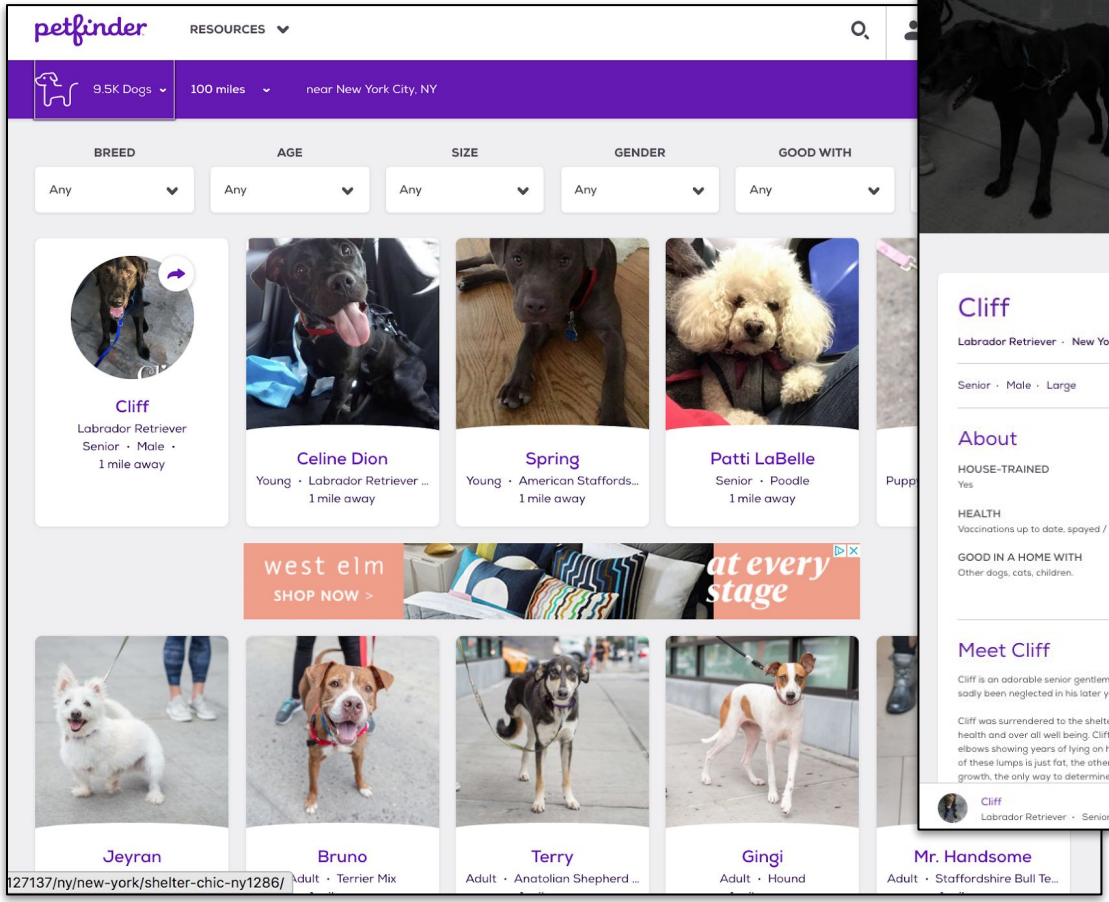
Python: clean & parse data for analysis

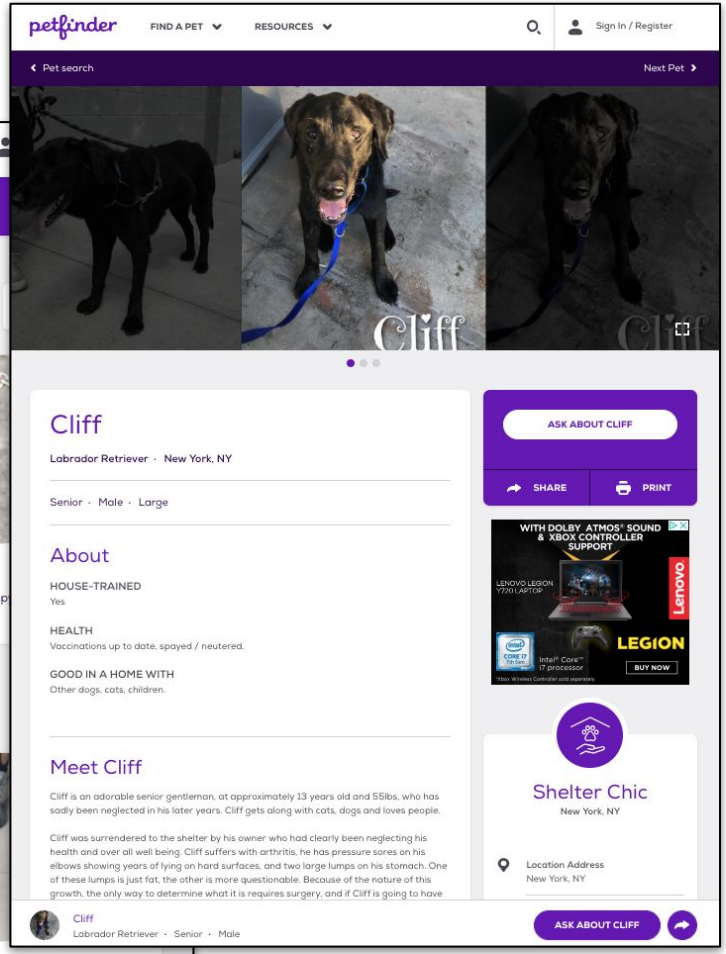
petfinder.com → search results → select dog → pet page → dataframe





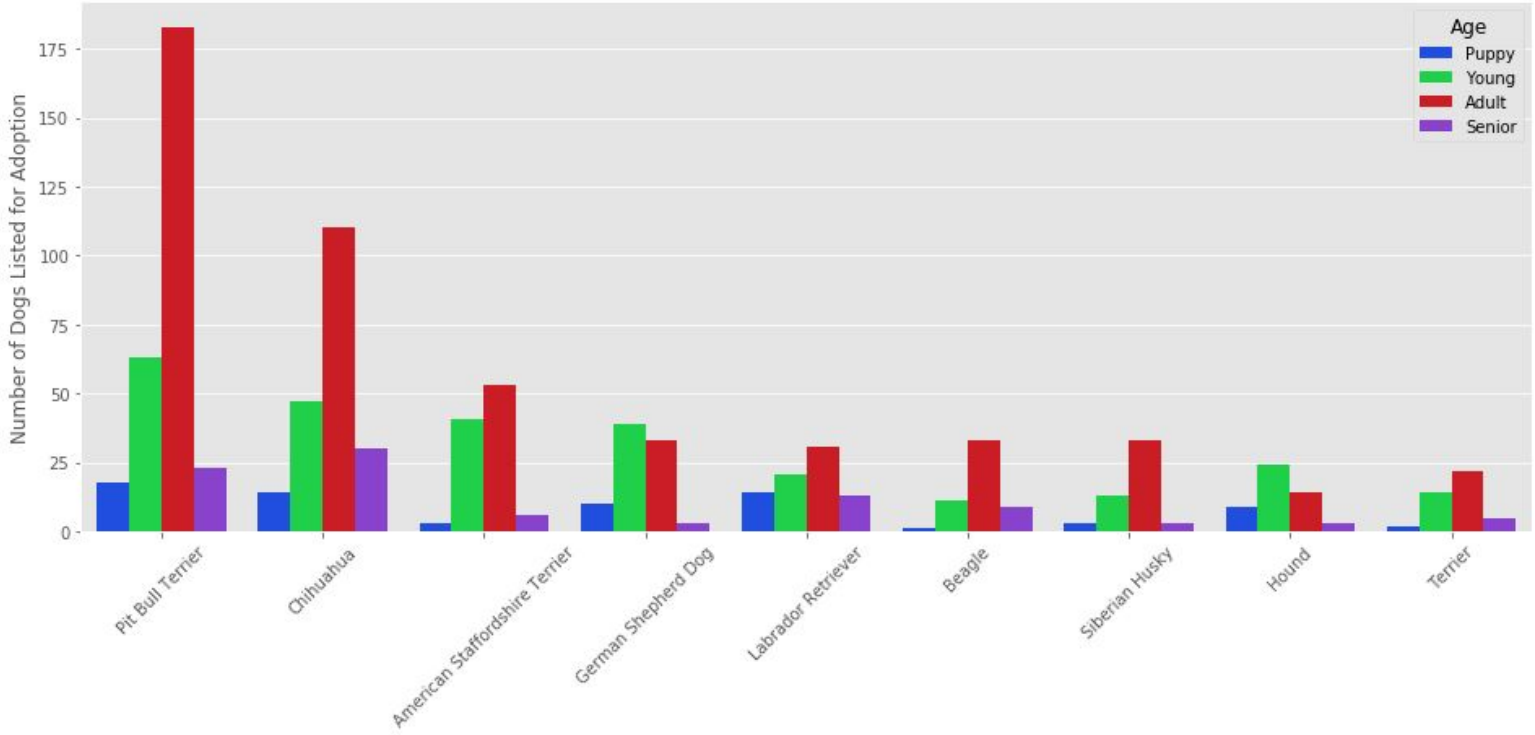
Selenium & Xpath: scrape visible & hidden data within HTML code


petfinder
RESOURCES
9.5K Dogs
100 miles
near New York City, NY
BREED
AGE
SIZE
GENDER
GOOD WITH
Any
Any
Any
Any
Any
Cliff
Labrador Retriever
Senior · Male ·
1 mile away
Celine Dion
Young · Labrador Retriever ...
1 mile away
Spring
Young · American Staffords...
1 mile away
Patti LaBelle
Senior · Poodle
1 mile away
west elm
SHOP NOW >
at every stage
Jeyran
Bruno
Adult · Terrier Mix
Terry
Adult · Anatolian Shepherd ...
Gingi
Adult · Hound
Mr. Handsome
Adult · Staffordshire Bull Te...
127137/ny/new-york/shelter-chic-ny1286/


petfinder
FIND A PET
RESOURCES
Sign In / Register
Pet search
Next Pet
Cliff
Cliff
Labrador Retriever · New York, NY
Senior · Male · Large
About
HOUSE-TRAINED
Yes
HEALTH
Vaccinations up to date, spayed / neutered.
GOOD IN A HOME WITH
Other dogs, cats, children.
Meet Cliff
Cliff is an adorable senior gentleman, at approximately 13 years old and 55lbs, who has sadly been neglected in his later years. Cliff gets along with cats, dogs and loves people.
Cliff was surrendered to the shelter by his owner who had clearly been neglecting his health and over all well being. Cliff suffers with arthritis, he has pressure sores on his elbows showing years of lying on hard surfaces, and two large lumps on his stomach. One of these lumps is just fat, the other is more questionable. Because of the nature of this growth, the only way to determine what it is requires surgery, and if Cliff is going to have
ASK ABOUT CLIFF
SHARE
PRINT
WITH DOLBY ATMOS® SOUND & XBOX CONTROLLER SUPPORT
LENOVO LEGION Y720 LAPTOP
LEGION
BUY NOW
Shelter Chic
New York, NY
Location Address
New York, NY
Cliff
Labrador Retriever · Senior · Male
ASK ABOUT CLIFF

Demographics

Purebred dogs with most listings in NYC region

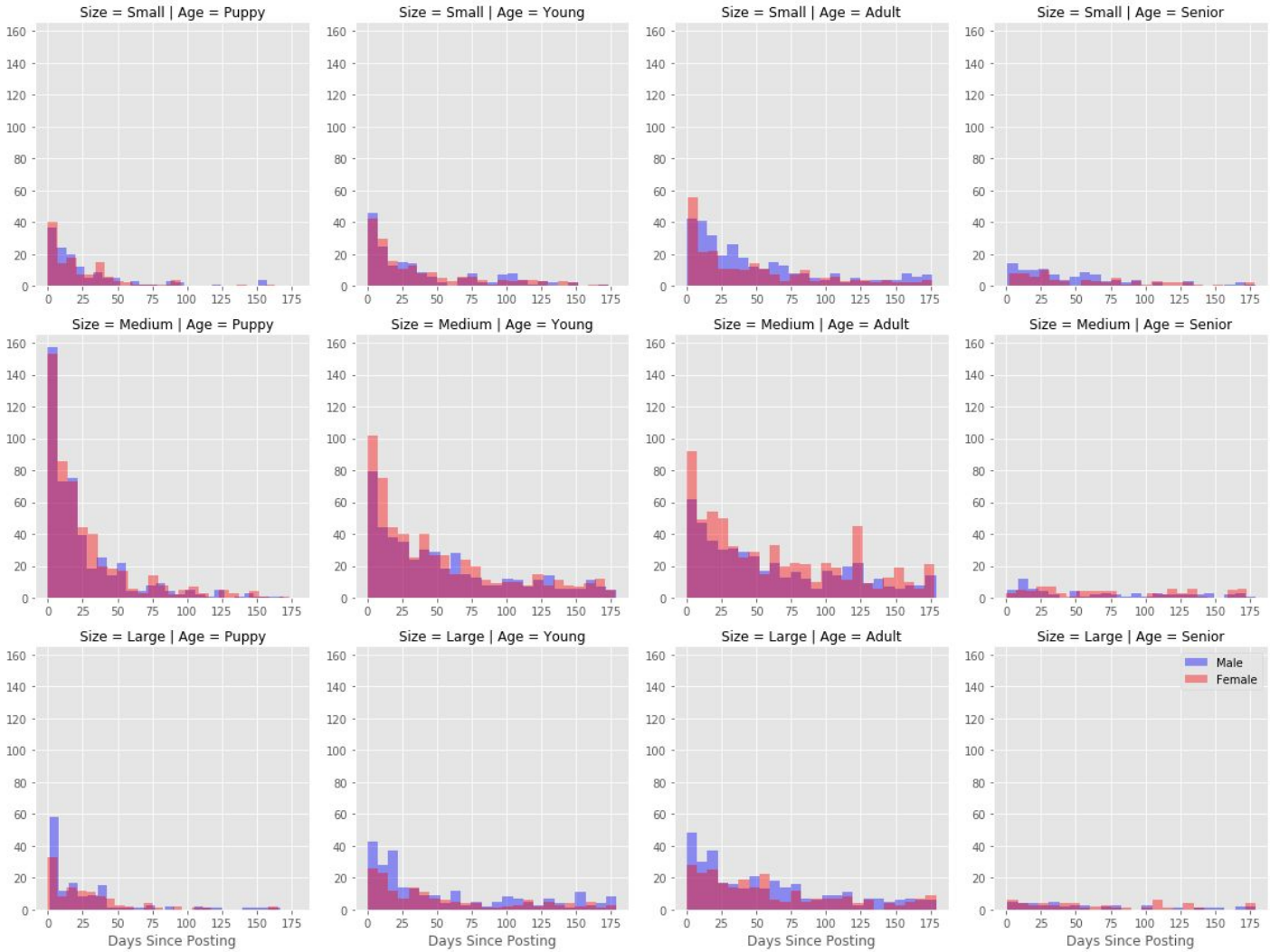


Adoption Trends

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

6-month time period

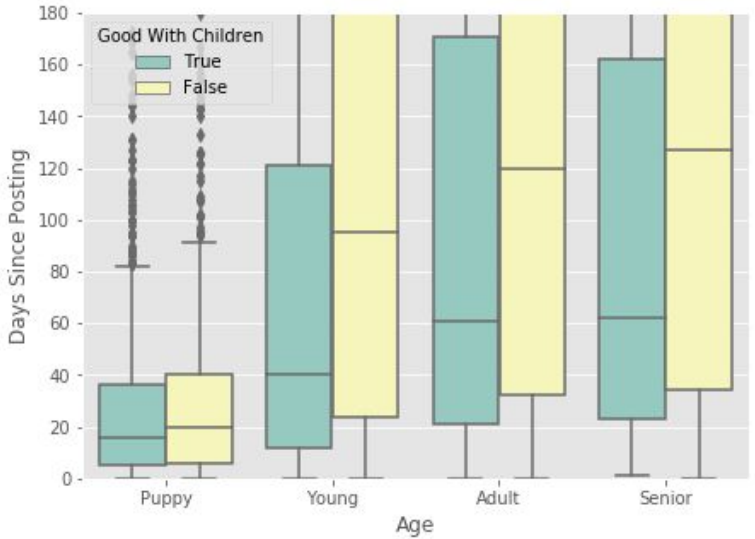
How quickly does each category get adopted?

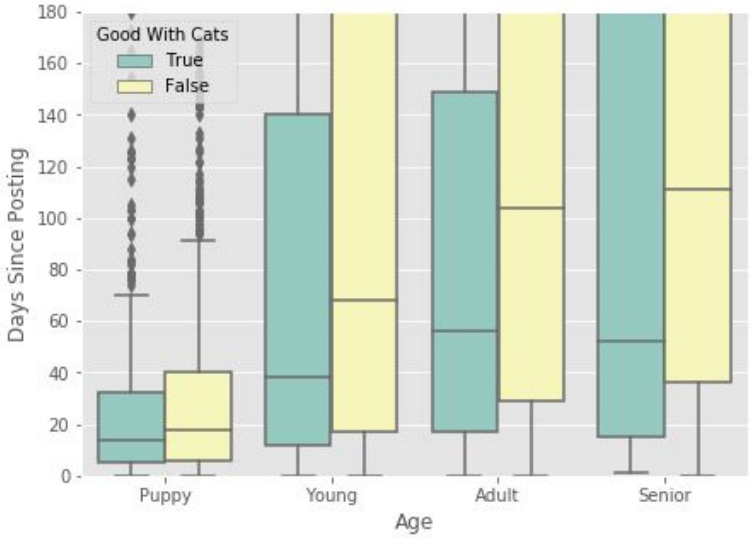


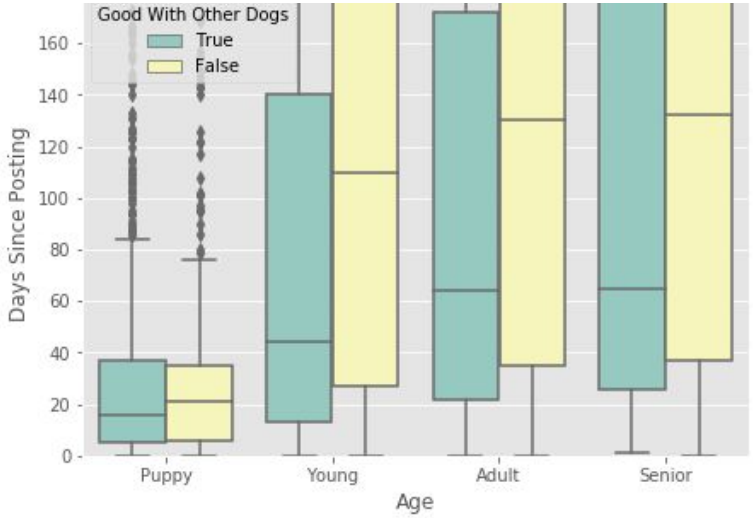
Adoption Trends

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

What qualities are most desirable at each age?

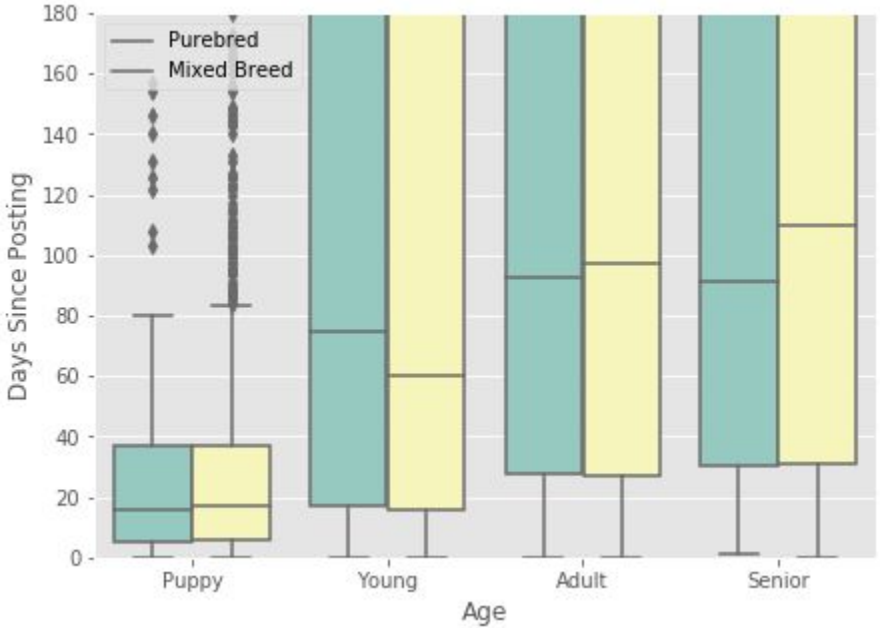


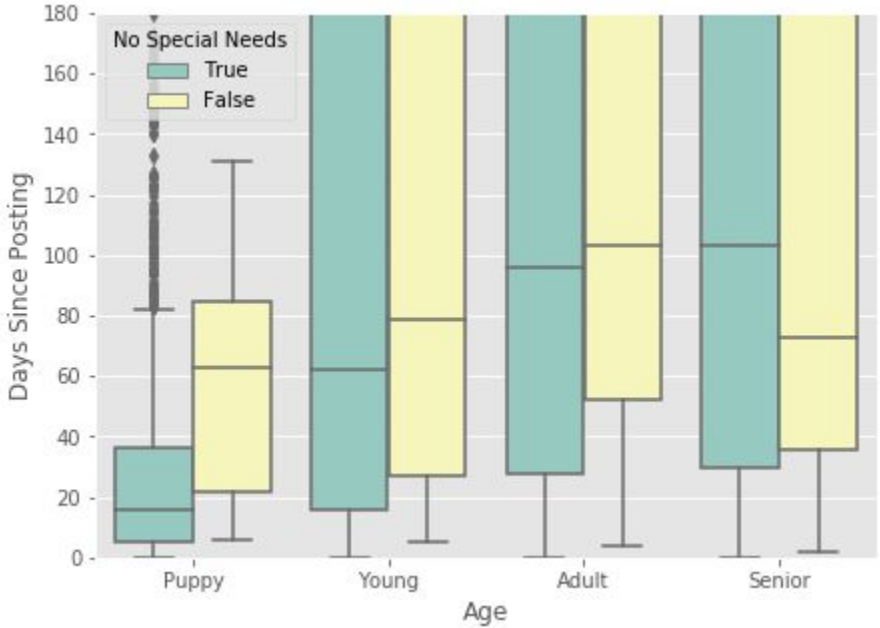




Adoption Trends

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

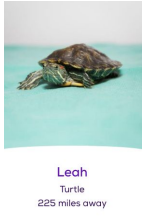


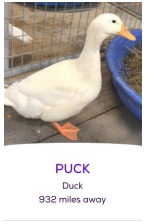


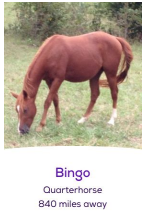
Wordclouds by Breed

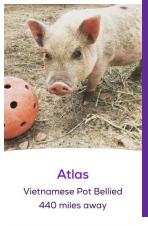
Continued Work

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.

GitHub