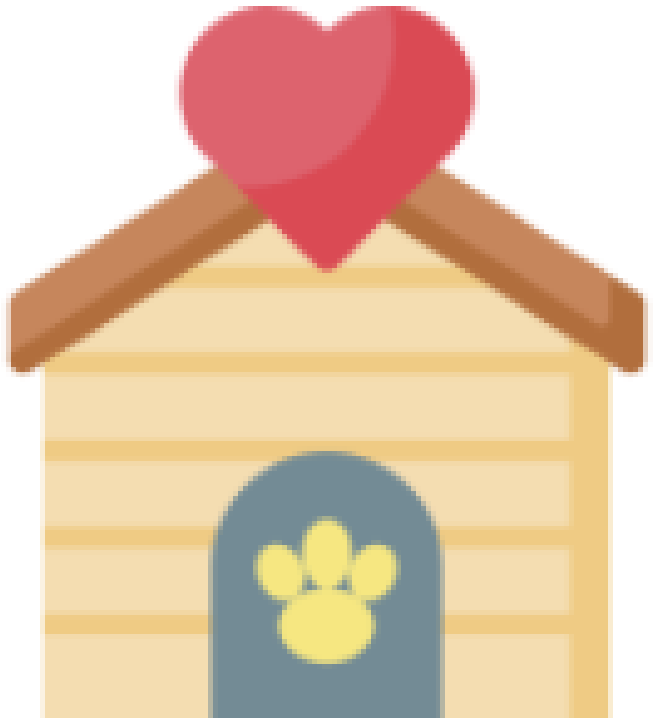
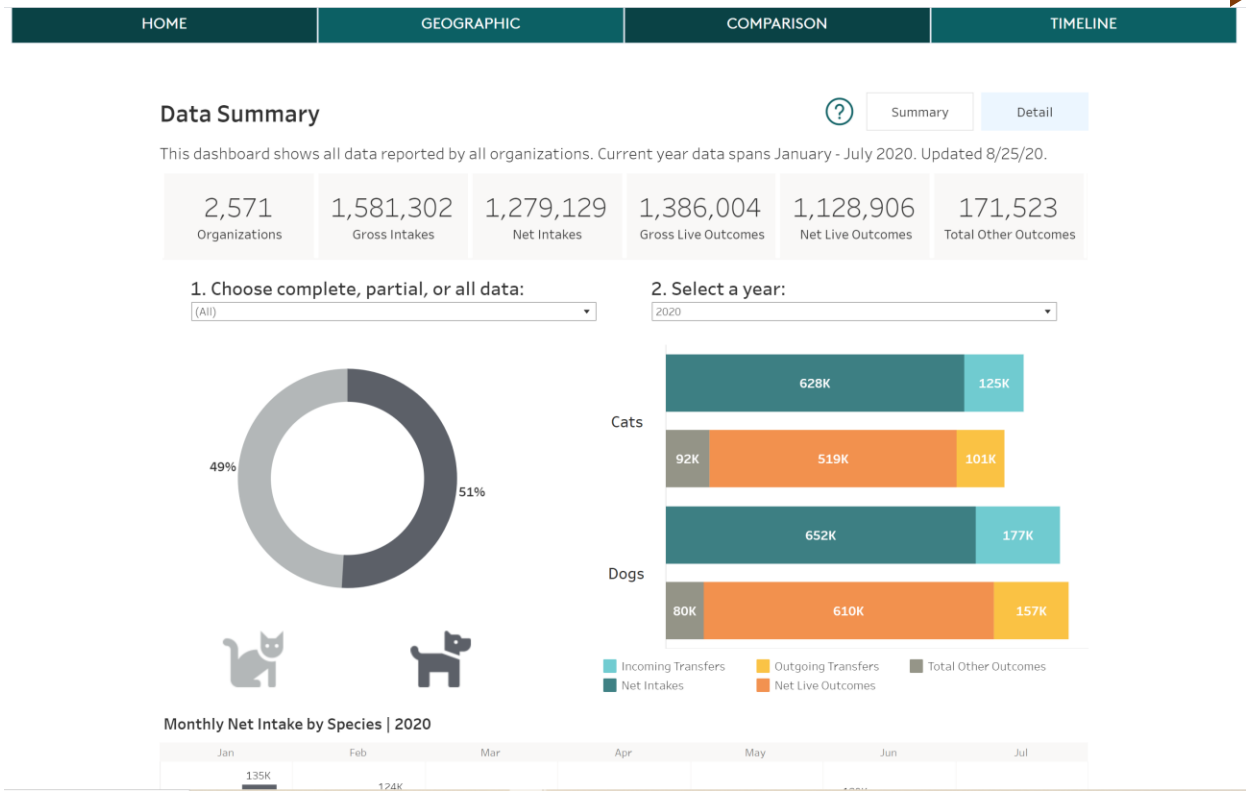


Gimme Shelter

Ana and the Chrises

Picking a Topic

- The topic of this project is the Dallas Animal Shelter data. The data for this project was pulled from dallasopendata.com catalog. The data was used to analyze the surrendered and stray animals, adoptions, and transferring of animals to rescue groups
- The focus was on developing machine learning algorithms to predict the adoptability of cats and dogs
- The inspiration was the following dashboard, which provides the latest animal shelter counts



Animal Shelter Statistics

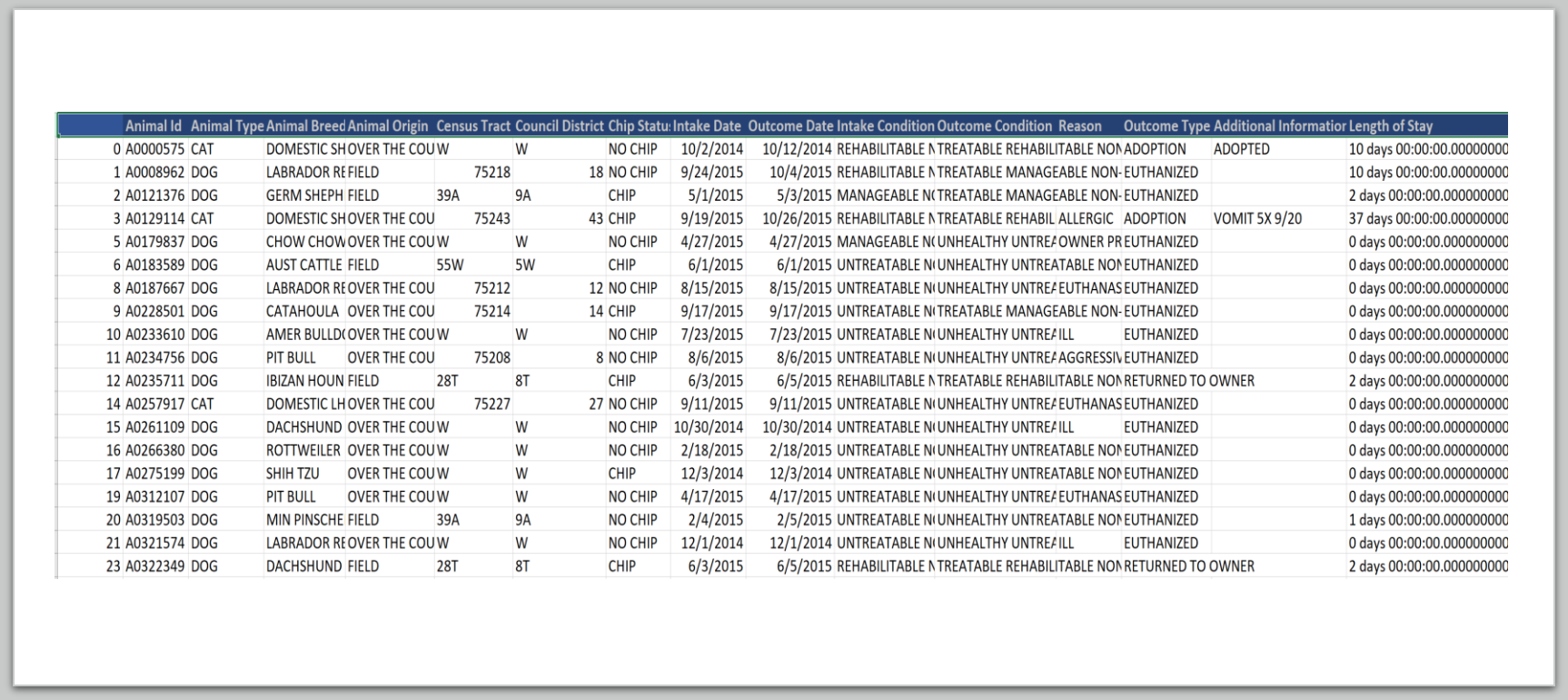
- Approximately 6.5 million companion animals enter U.S. animal shelters nationwide every year. Of those, approximately 3.3 million are dogs and 3.2 million are cats
- Each year, approximately 1.5 million shelter animals are euthanized (670,000 dogs and 860,000 cats)
- Approximately 3.2 million shelter animals are adopted each year (1.6 million dogs and 1.6 million cats)
- About 710,000 animals who enter shelters as strays are returned to their owners. Of those, 620,000 are dogs and only 90,000 are cats

https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics

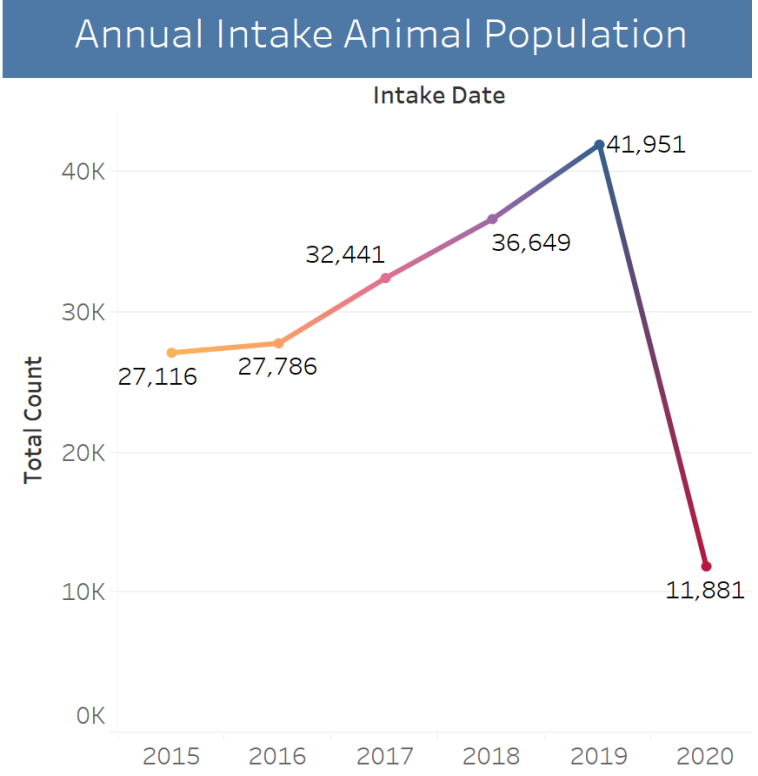
Dallas Animal Services / Facebook

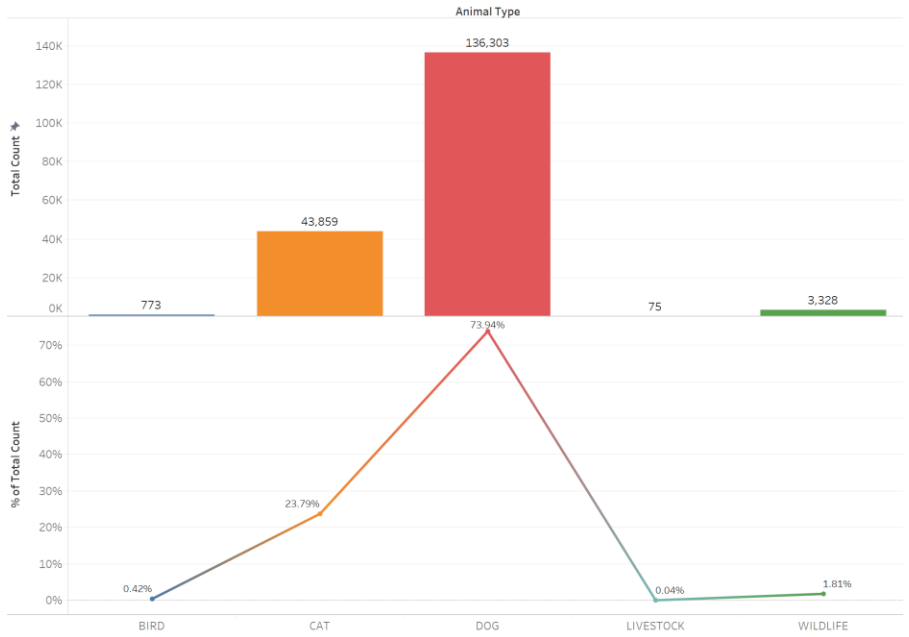
What is in this Dataset?

- 30.9K rows and 34 columns
- From January 2015 through July 2020 the original Dataset consisted of 211,216 animal records
- Data analyzed includes a period of 60 days with a total of 184,338 animals



	Animal Id	Animal Type	Animal Breed	Animal Origin	Census Tract	Council District	Chip Statu	Intake Date	Outcome Date	Intake Condition	Outcome Condition	Reason	Outcome Type	Additional Informatior	Length of Stay
0	A0000575	CAT	DOMESTIC SH	OVER THE COU	W	W	NO CHIP	10/2/2014	10/12/2014	REHABILITABLE N	TREATABLE REHABILITABLE NON		ADOPTION	ADOPTED	10 days 00:00:00.00000000
1	A0008962	DOG	LABRADOR RE	FIELD	75218	18	NO CHIP	9/24/2015	10/4/2015	REHABILITABLE N	TREATABLE MANAGEABLE NON-		EUTHANIZED		10 days 00:00:00.00000000
2	A0121376	DOG	GERM SHEPH	FIELD	39A	9A	CHIP	5/1/2015	5/3/2015	MANAGEABLE N(	TREATABLE MANAGEABLE NON-		EUTHANIZED		2 days 00:00:00.00000000
3	A0129114	CAT	DOMESTIC SH	OVER THE COU	75243	43	CHIP	9/19/2015	10/26/2015	REHABILITABLE N	TREATABLE REHABIL	ALLERGIC	ADOPTION	VOMIT 5X 9/20	37 days 00:00:00.00000000
5	A0179837	DOG	CHOW CHOW	OVER THE COU	W	W	NO CHIP	4/27/2015	4/27/2015	MANAGEABLE N(	UNHEALTHY UNTREA	OWNER PR	EUTHANIZED		0 days 00:00:00.00000000
6	A0183589	DOG	AUST CATTLE	FIELD	55W	5W	CHIP	6/1/2015	6/1/2015	UNTREATABLE N(	UNHEALTHY UNTREATABLE NON		EUTHANIZED		0 days 00:00:00.00000000
8	A0187667	DOG	LABRADOR RE	OVER THE COU	75212	12	NO CHIP	8/15/2015	8/15/2015	UNTREATABLE N(	UNHEALTHY UNTREA	EUTHANAS	EUTHANIZED		0 days 00:00:00.00000000
9	A0228501	DOG	CATAHOULA	OVER THE COU	75214	14	CHIP	9/17/2015	9/17/2015	UNTREATABLE N(	TREATABLE MANAGEABLE NON-		EUTHANIZED		0 days 00:00:00.00000000
10	A0233610	DOG	AMER BULLD(	OVER THE COU	W	W	NO CHIP	7/23/2015	7/23/2015	UNTREATABLE N(	UNHEALTHY UNTREA	ILL	EUTHANIZED		0 days 00:00:00.00000000
11	A0234756	DOG	PIT BULL	OVER THE COU	75208	8	NO CHIP	8/6/2015	8/6/2015	UNTREATABLE N(	UNHEALTHY UNTREA	AGGRESSIV	EUTHANIZED		0 days 00:00:00.00000000
12	A0235711	DOG	IBIZAN HOUN	FIELD	28T	8T	CHIP	6/3/2015	6/5/2015	REHABILITABLE N	TREATABLE REHABILITABLE NON		RETURNED TO OWNER		2 days 00:00:00.00000000
14	A0257917	CAT	DOMESTIC LH	OVER THE COU	75227	27	NO CHIP	9/11/2015	9/11/2015	UNTREATABLE N(	UNHEALTHY UNTREA	EUTHANAS	EUTHANIZED		0 days 00:00:00.00000000
15	A0261109	DOG	DACHSHUND	OVER THE COU	W	W	NO CHIP	10/30/2014	10/30/2014	UNTREATABLE N(	UNHEALTHY UNTREA	ILL	EUTHANIZED		0 days 00:00:00.00000000
16	A0266380	DOG	ROTTWEILER	OVER THE COU	W	W	NO CHIP	2/18/2015	2/18/2015	UNTREATABLE N(	UNHEALTHY UNTREATABLE NON		EUTHANIZED		0 days 00:00:00.00000000
17	A0275199	DOG	SHIH TZU	OVER THE COU	W	W	CHIP	12/3/2014	12/3/2014	UNTREATABLE N(	UNHEALTHY UNTREATABLE NON		EUTHANIZED		0 days 00:00:00.00000000
19	A0312107	DOG	PIT BULL	OVER THE COU	W	W	NO CHIP	4/17/2015	4/17/2015	UNTREATABLE N(	UNHEALTHY UNTREA	EUTHANAS	EUTHANIZED		0 days 00:00:00.00000000
20	A0319503	DOG	MIN PINSCHE	FIELD	39A	9A	NO CHIP	2/4/2015	2/5/2015	UNTREATABLE N(	UNHEALTHY UNTREATABLE NON		EUTHANIZED		1 days 00:00:00.00000000
21	A0321574	DOG	LABRADOR RE	OVER THE COU	W	W	NO CHIP	12/1/2014	12/1/2014	UNTREATABLE N(	UNHEALTHY UNTREA	ILL	EUTHANIZED		0 days 00:00:00.00000000
23	A0322349	DOG	DACHSHUND	FIELD	28T	8T	CHIP	6/3/2015	6/5/2015	REHABILITABLE N	TREATABLE REHABILITABLE NON		RETURNED TO OWNER		2 days 00:00:00.00000000



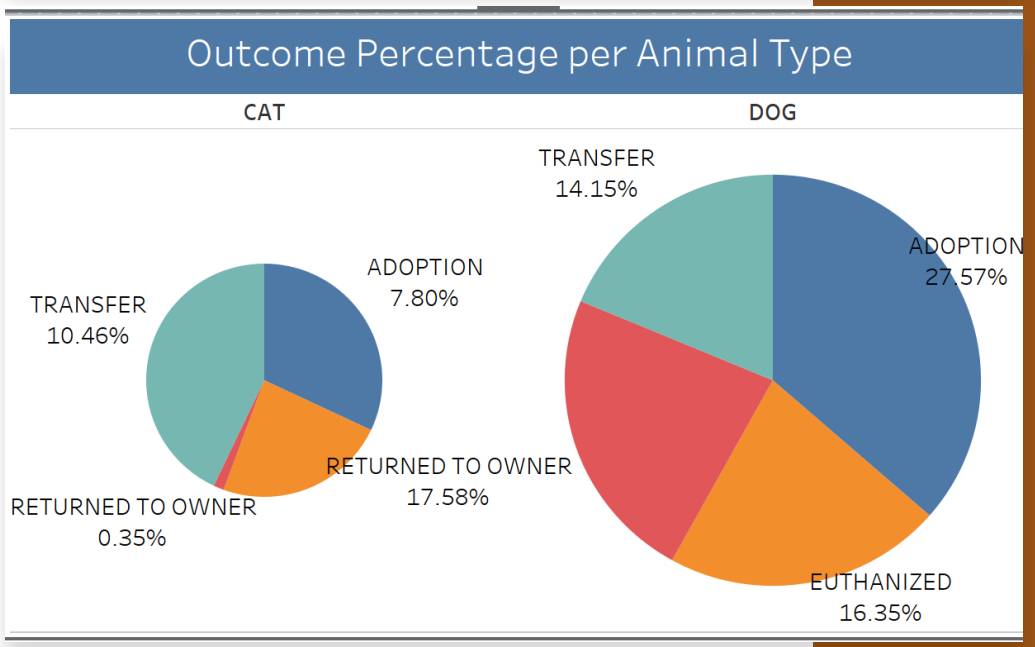


Data Overview

- The Annual Intake Animal Population chart shows steady yearly increases from 2016 with a sharp decline during 2020
- Dog intakes were 136,303 (73.94%) of the shelter population, followed by 43, 859 cats (23.79%)

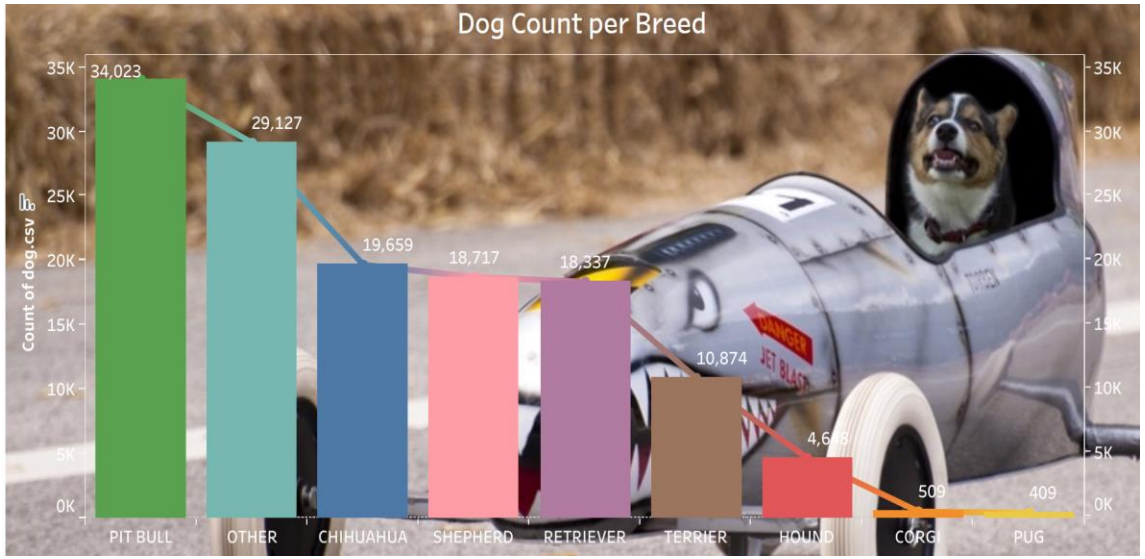
Exploratory Data Analysis

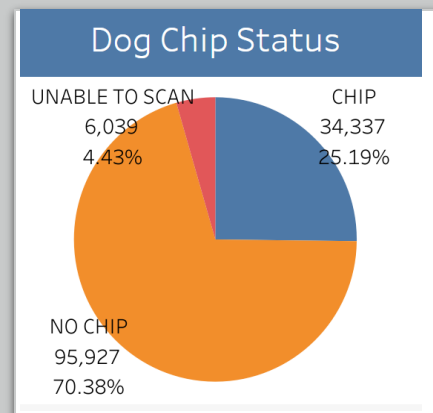
- Top 5 reasons given why animals are in shelter include: too many, owner problem, move, no time and Landlord.
- Out of the dogs and cats that were impounded, 28% of dogs and 8% of cats were adopted
- 16% of dogs and 6% of cats were euthanized.

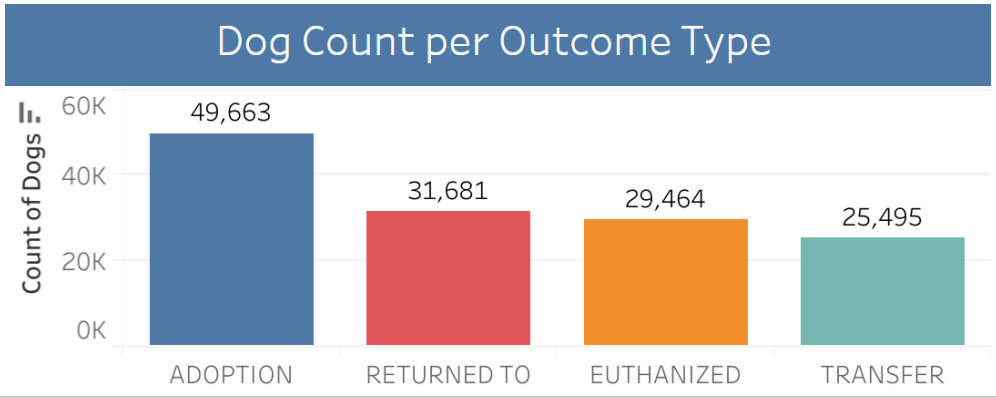


Exploratory Data Analysis

- Pitbull breed make up the population of dogs most impounded, while the least most impounded dog population is Corgis and Pugs.
- 25.19% of dogs impounded have a tracking microchip and 24% were returned to owner
- Of 136,303 dogs, 49,663 (36.44%) were adopted and 29,464 (21.62%) were euthanized

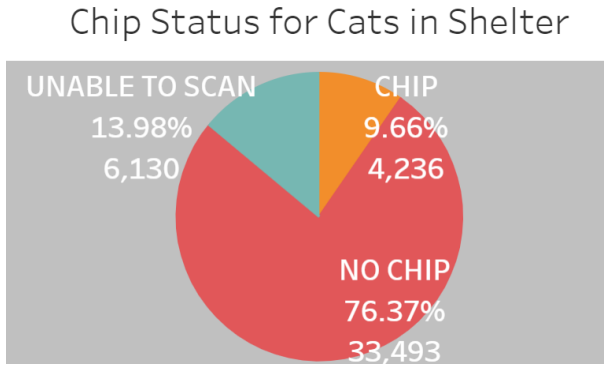


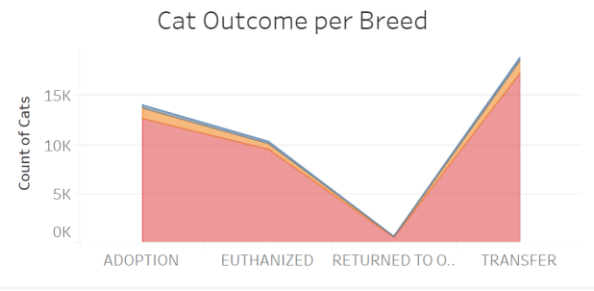


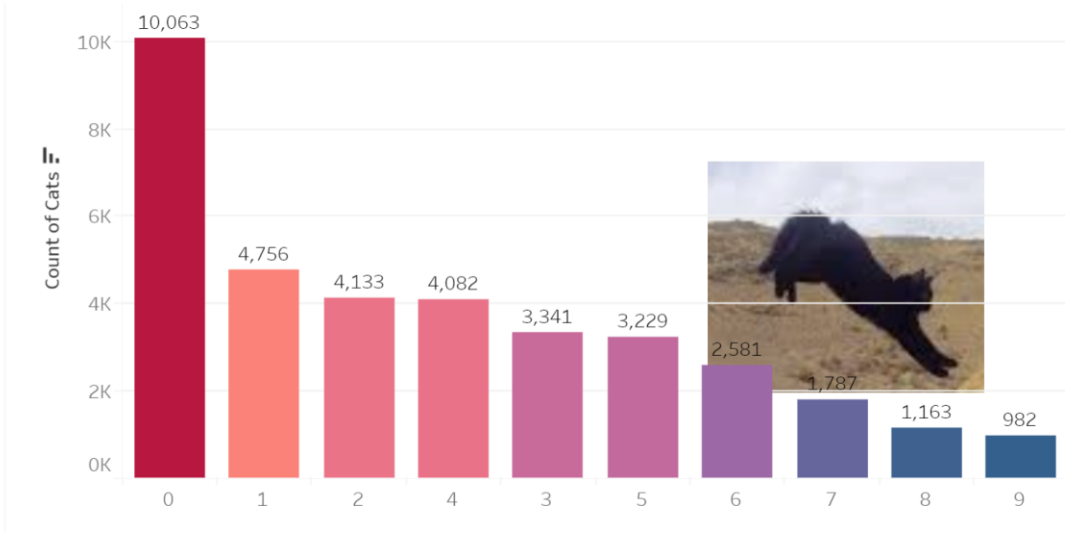


Exploratory Data Analysis

- 9.66% of cats impounded have a tracking microchip
- Of the cats that were impounded less than 2% were returned to owner
- Of 43,859 cats, 10,063 on the first day of impoundment were either returned to owner, adopted or euthanized







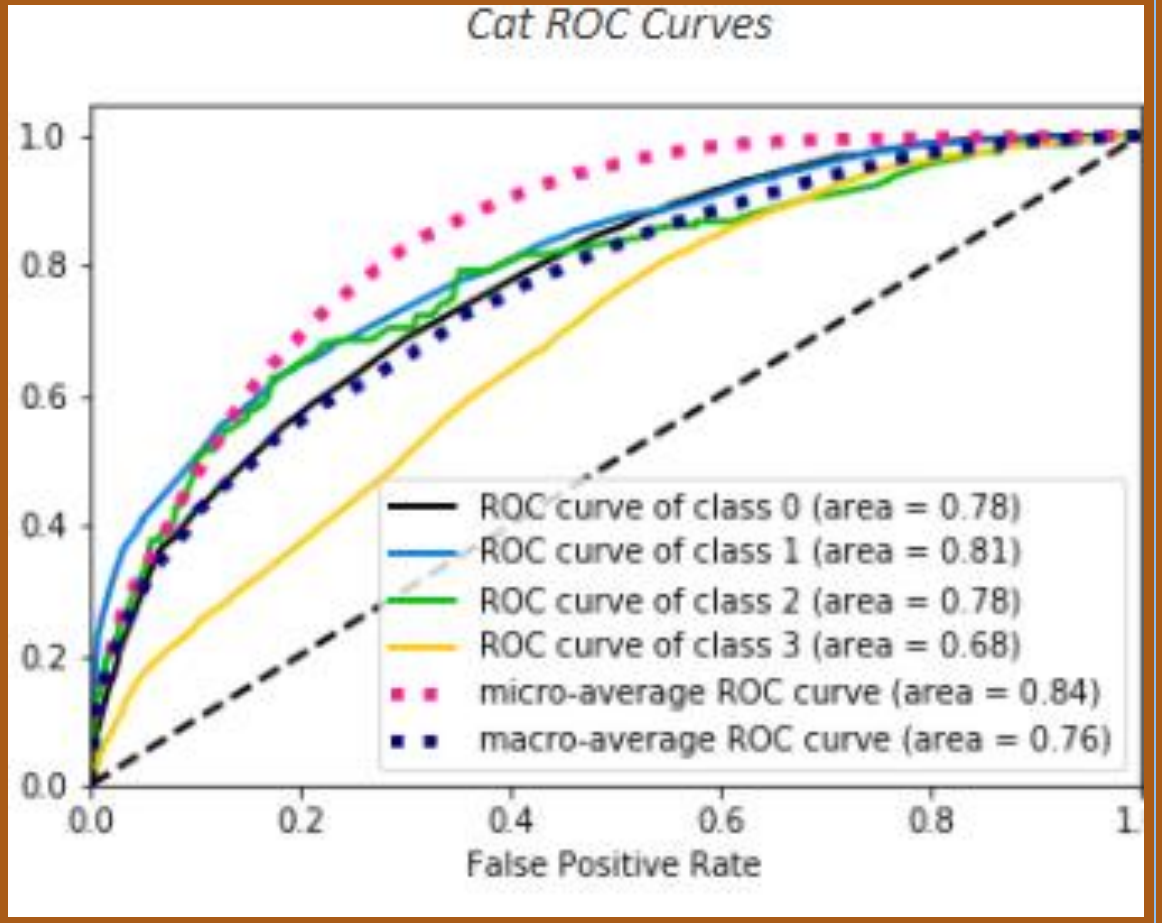
Data Modeling

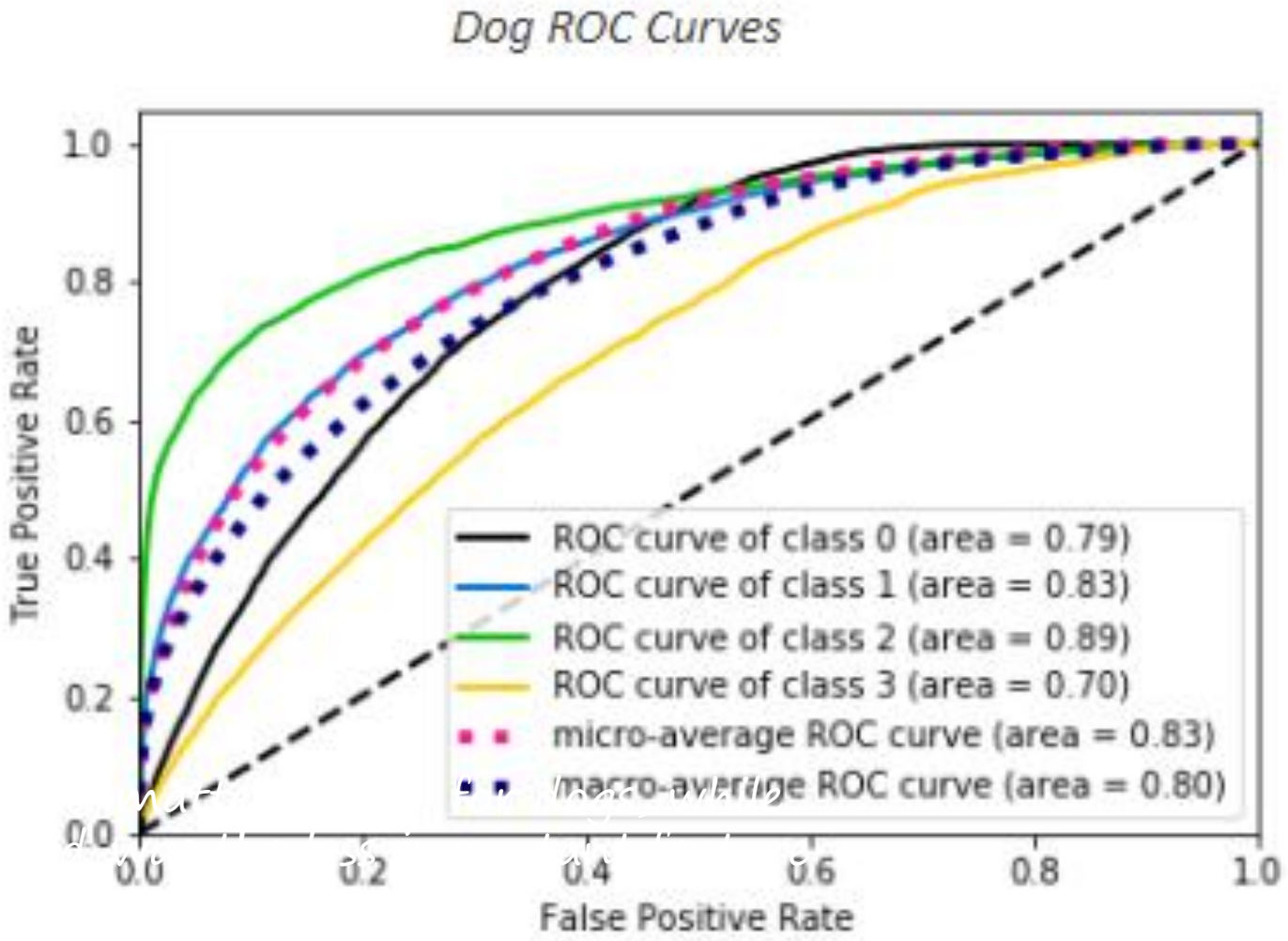
- While each model has its strengths and weaknesses, we believe that, for both cats and dogs, the Gradient Boost Classifier works the best. It had similar accuracy scores (0.58 for cats, 0.59 for dogs) with the other tree ensemble models but had better F1 scores than the rest of the models. It also falsely predicted euthanization at lower rates than the other models. While it did predict more euthanized animals would be adopted than other models, those models had lower F1 scores for predicting euthanization.

Data Modeling

Cat Gradient Boost Classification Analysis Report

While length of stay was the most important feature for both (0.26 for cats, 0.40 for dogs), the origin of a dog seems to matter more than its health, and the opposite for cats.



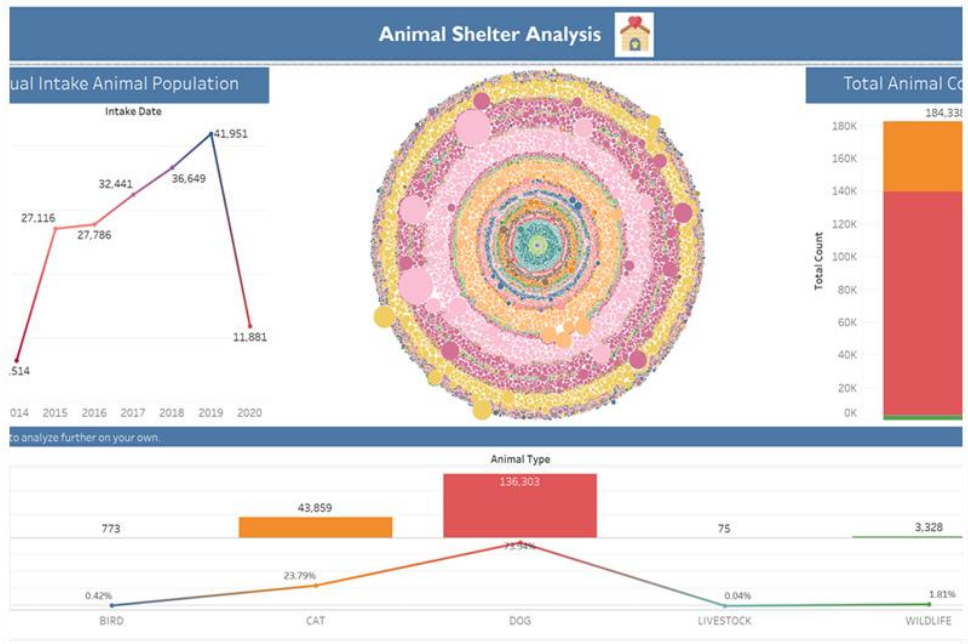


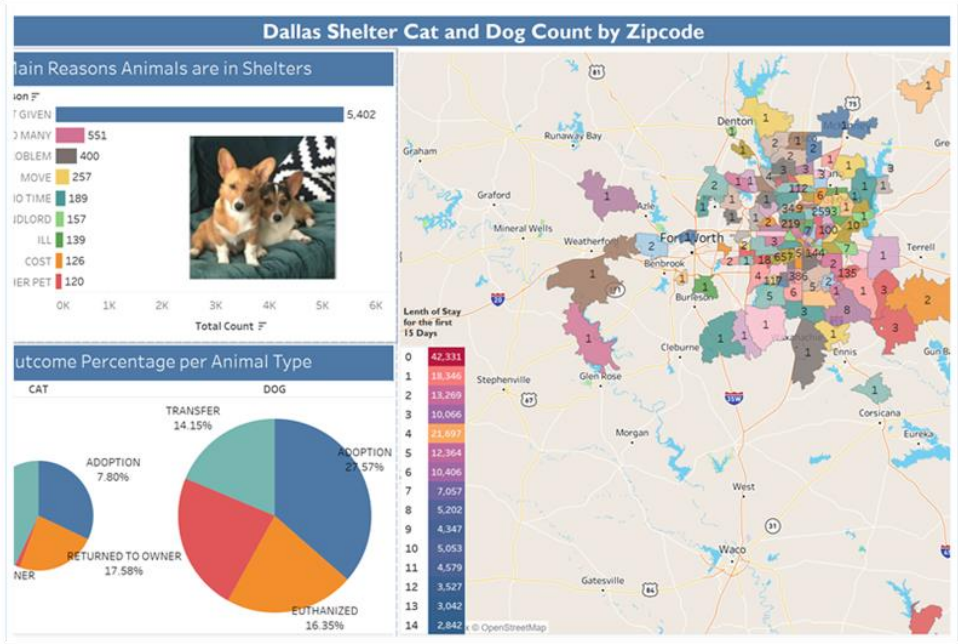
Data Modeling

Dog Gradient Boost Classification Analysis Report

Dashboard & Mobile Applications

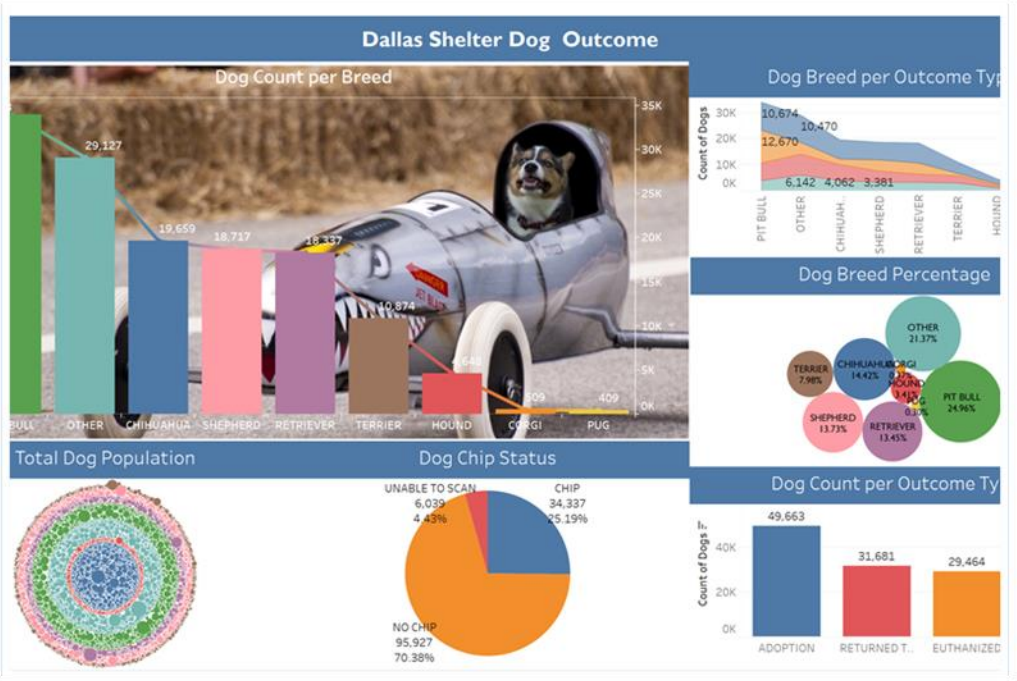
- The following dashboards were developed to provide options to uncover additional insight for this analysis.

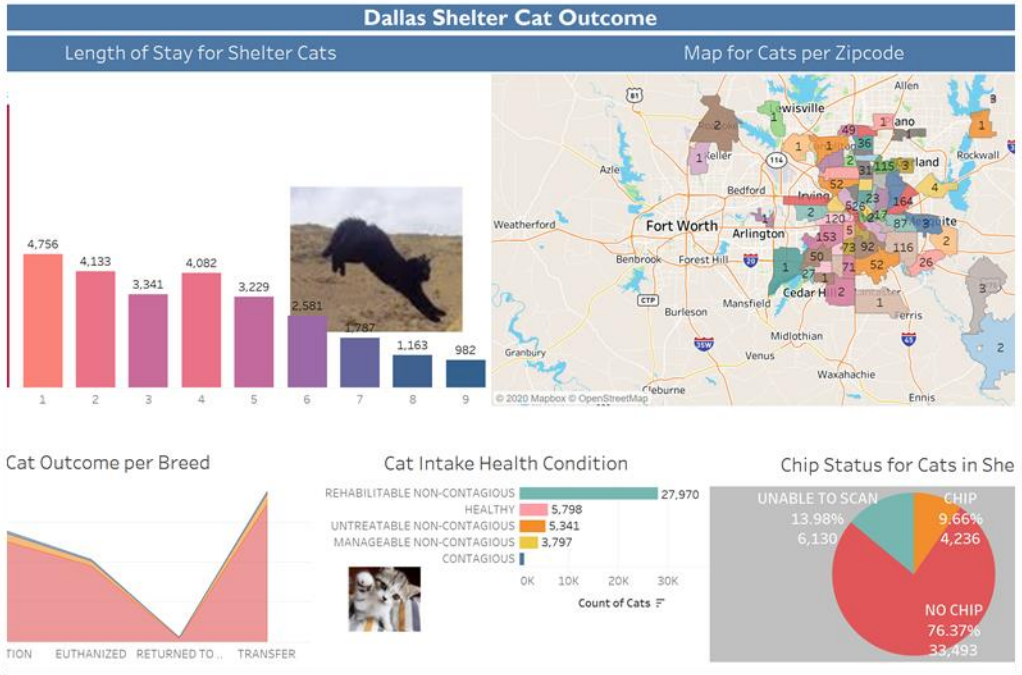




Dashboard & Mobile Applications

- The following dashboards were developed to provide options to uncover additional insight for this analysis.





Modeling Applications

- The following dashboards were developed to provide options to uncover additional insight for this analysis.

Cat Predictions

Cat Breed	Animal Origin	Chip Status	Intake Condition	Length of Stay(Days)
Medium Hair	Field	No Chip	Healthy	3

Submit

Cat Breed:	Medium Hair
Cat Origin:	Field
Chip Status:	No Chip
Intake Condition:	Healthy
Length of Stay:	3

Your cat is likely to be adopted!

Result	Probability
Adopted:	47.44%
Euthanized:	2.76%
Returned to Owner:	3.01%
Transfer/Foster:	46.8%

Dog Predictions

Dog Breed	Origin	Chip Status	Intake Condition	Length of Stay(Days)
Chihuahua	Field	Chip	Healthy	

Submit

Dog Breed:	Pug
Dog Origin:	Field
Chip Status:	No Chip
Intake Condition:	Healthy
Length of Stay:	1

Your dog is likely to be adopted!

Result	Probability
Adopted:	37.82%
Euthanized:	2.5%
Returned to Owner:	33.13%
Transfer/Foster:b	26.55%

Limitations / Future Work

- Model accuracy is limited (59% for Dogs, 58% for Cats)
- Geodata was difficult to clean
  - Mapsco - ##A – likely not possible
  - Zip Code - 752### - Able to map in Tableau
  - Census Tract - ###.## - Needs further cleaning/Tableau manipulation
- Large number of dog breeds

- Future Work:
- - Get Census Tract and associated data involved with the Maps on Tableau
- - Web Scrape on a monthly basis as the CSV files are published on Dallas OpenData
- - Parse the animals into Breeds (Pitbull / Pug) rather than just by Type (Cat / Dog)
- - Continuously test other Machine Learning Models to ensure we are using the best one
  - Gradient Boost Classifier won't necessarily always be the best

Conclusions

- Total shelter population is down for 2020
- Dogs make up 74% and Cats 24%
- Dogs are more likely to be adopted
- Accuracy limits the use of models

Q&A