

Predicting Shelter Animal Outcomes

Lisa Ryan
May 31st, 2016

Overview

- Kaggle competition – Austin Animal Center
- Insights from competition could help shelters focus their energy on specific animals who need a little extra help finding a home.
- 7.6 million pets end up in US shelters every year.
- Competition judged using logarithmic loss. Lower numbers are better.

Project Goals

- Predict which of 5 outcomes a shelter animal will have
- Understand trends in animal outcomes
- Identify types of animals that may need additional resources for a good outcome

Target Variable

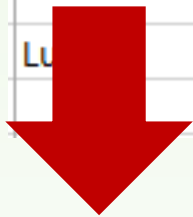
- Five possible outcomes for each animal:
 - Adoption (40.3%)
 - Transferred (35.3%)
 - Returned to owner (17.9%)
 - Euthanasia (5.8%)
 - Died (0.7%)
- 26,729 total records in dataset

Predictor Variables

- Many strings of text that need to be separated (e.g. breed, color, sex)
- Age variable should be numeric
- Several indicator variables can be created (e.g. name indicator, mix indicator)

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Pit Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lu	Dog	Spayed Female	5 months	American Pit Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White



Name indicator:

- Has name
- No name

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Do	Spayed Female	5 months	American Pit Bull Terrier Mix	Red/White
	Do	Spayed Female	1 year	Cairn Terrier	White

Animal Type:

- Dog
- Cat

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Pit Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White

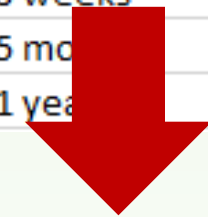
Sex Type:

- Altered
- Intact

Sex:

- Male
- Female

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Pit Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White



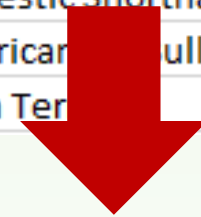
Age:
 $1/12 = 0.083$

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White

Breed Groups:

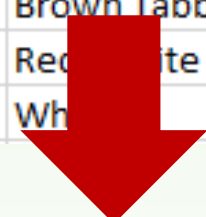
- Herding
- Hound
- Non-sporting
- Sporting
- Terrier
- Toy
- Working
- Pit Bull

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Bull Terrier Mix	Red/White
	Dog	Spayed Female	1 year	Cairn Terrier	White



- Mix Indicator
- Purebred Indicator

Name	AnimalType	SexuponOutcome	AgeuponOutcome	Breed	Color
Hambone	Dog	Neutered Male	1 year	Shetland Sheepdog Mix	Brown/White
Emily	Cat	Spayed Female	1 year	Domestic Shorthair Mix	Cream Tabby
Pearce	Dog	Neutered Male	2 years	Pit Bull Mix	Blue/White
	Cat	Intact Male	3 weeks	Domestic Shorthair Mix	Blue Cream
	Dog	Neutered Male	2 years	Lhasa Apso/Miniature Poodle	Tan
Elsa	Dog	Intact Female	1 month	Cairn Terrier/Chihuahua Shorthair	Black/Tan
Oreo	Dog	Intact Female	1 month	Irish Setter/Pit Bull	Black/White
	Cat	Unknown	3 weeks	Domestic Shorthair Mix	Brown Tabby
Lucy	Dog	Spayed Female	5 months	American Pit Bull Terrier Mix	Red White
	Dog	Spayed Female	1 year	Cairn Terrier	White



Color:

- Black
- White

Color Type:

e.g., Tabby

Random Forest Model

- Good for predicting categories
- Better predictions than other models (e.g. decision trees)
- Easy to find most important variables
- Top variables:
 - Age
 - Altered/Intact/Unkown
 - Name flag
 - Dog/Cat/Unknown
 - Female/Male/Unknown
 - Pit bull

Random Forest Model

- Null accuracy = 0.402
- Accuracy = 0.586
- Log loss = 0.967
- Mean 5-fold cross-validation accuracy = 0.58
- Sensitivity: how often is a class predicted correctly? (True positive rate)
- Specificity: when not in a class, how often do we predict that an animal is not in that class? (True negative rate)

Outcome	Sensitivity	Specificity
Adoption	0.77	0.61
Transfer	0.53	0.87
Return to Owner	0.48	0.88
Euthanasia	0.02	1.00
Died	0.00	1.00

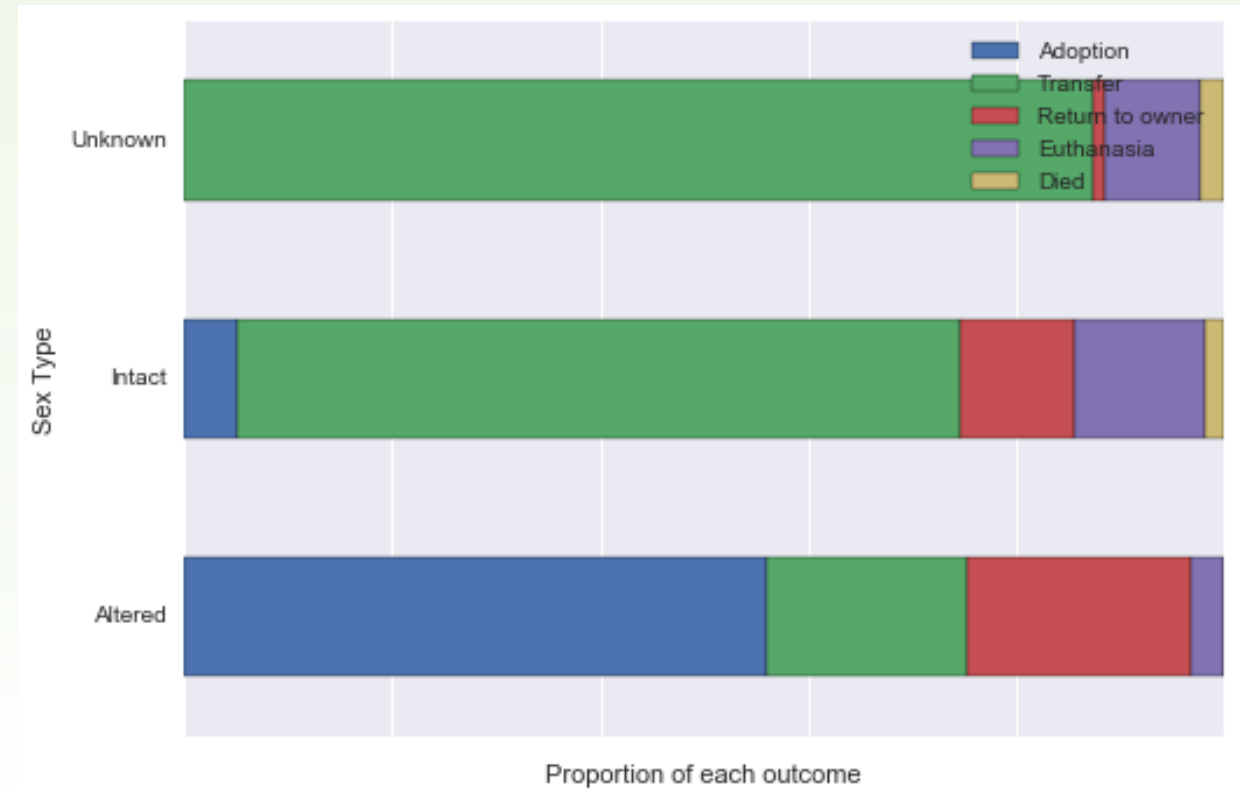
Decision Tree

- Good for predicting categories
- Easily interpretable model
- Null accuracy = 0.402
- Accuracy = 0.553
- Log loss = 1.016
- Mean 5-fold cross-validation accuracy = 0.544
- Sensitivity: how often is a class predicted correctly? (True positive rate)
- Specificity: when not in a class, how often do we predict that an animal is not in that class? (True negative rate)

Outcome	Sensitivity	Specificity
Adoption	0.60	0.71
Transfer	0.55	0.83
Return to Owner	0.65	0.80
Euthanasia	0.03	1.00
Died	0.00	1.00

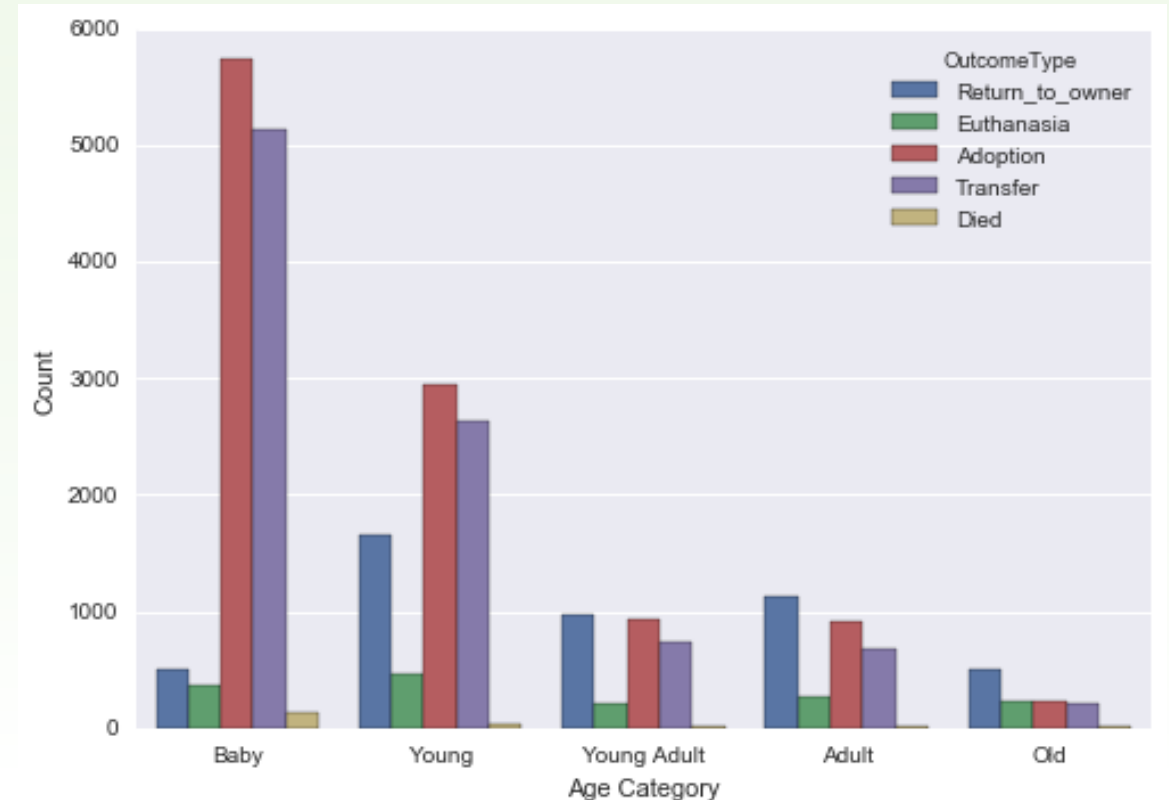
Altered/Intact Flag may not be a good predictor

- Shelters most likely spay/neuter animals prior to adopting them out.
- They probably don't invest in that effort if they're transferring them.
- Flag was derived from *sex upon outcome*.
- This variable was excluded from both models.



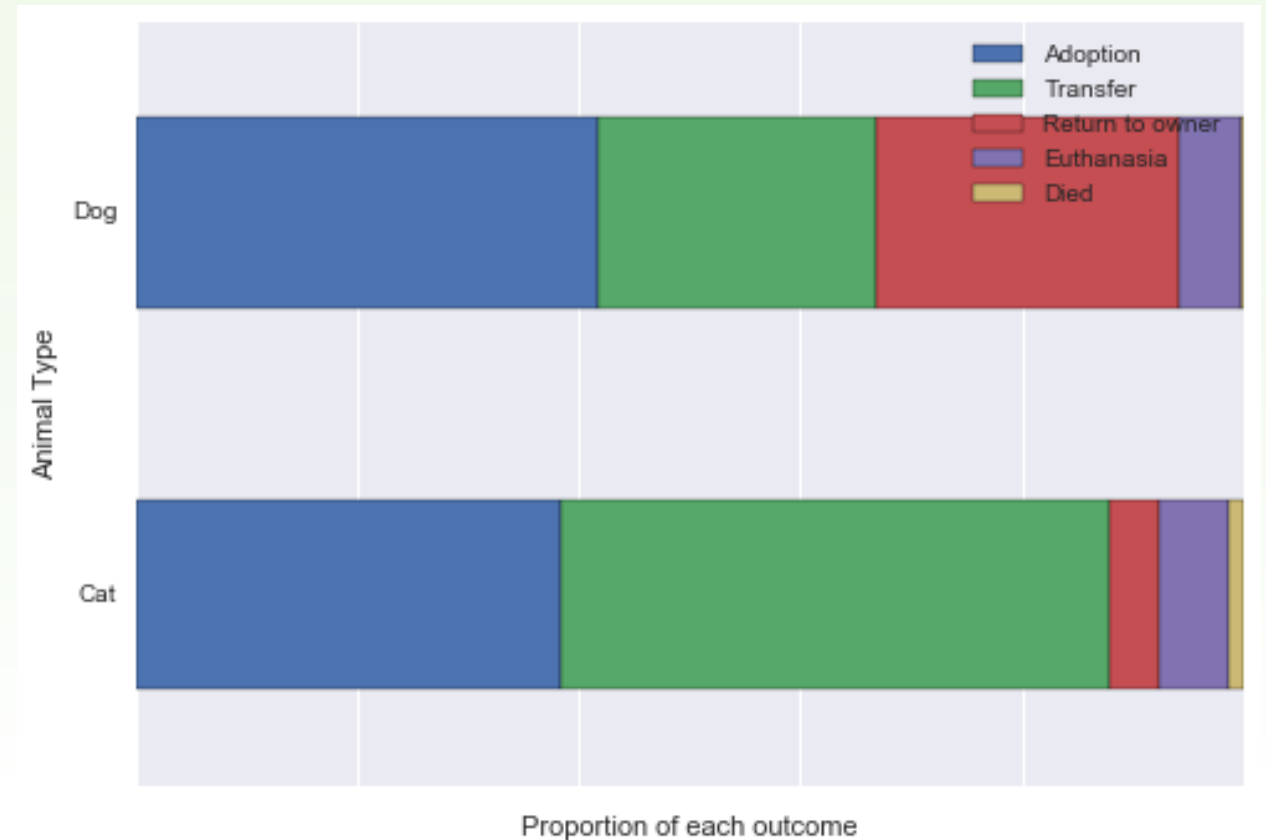
Age is one of the best predictors of outcome

- Categorized age into:
 - Baby (<1 year)
 - Young (<3 years)
 - Young Adult (<5 years)
 - Adult (<10 years)
 - Old (≥ 10 years)
- Younger animals have much higher Adoption and transfer rates.
- Older animals need more help getting adopted.



Cats are much less likely than dogs to be returned to owner

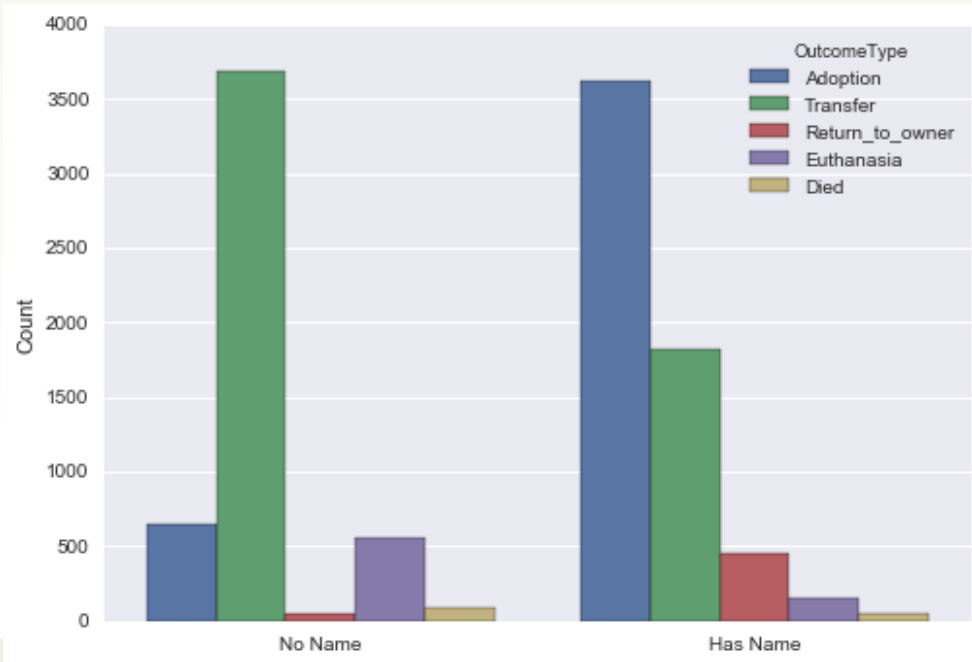
- Possibly because more cats are found feral.
- The shelter could start gathering information on whether or not each animal was an owner surrender.
- The shelter should invest more time and resources in achieving good outcomes for cats.



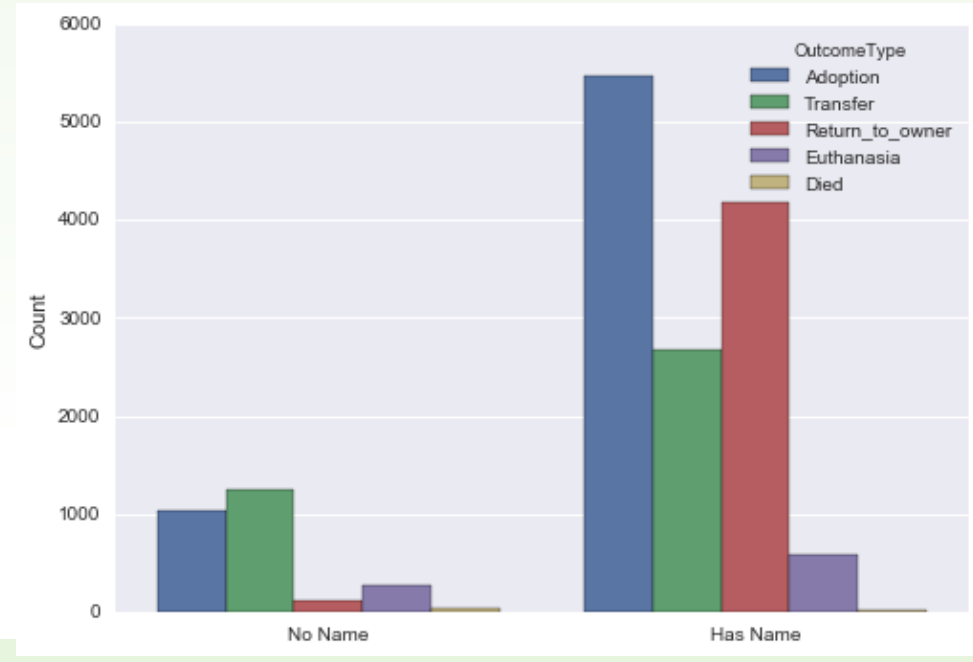
Naming animals may help them get adopted

- Animals with names have higher adoption rates.
- Cats without names have high transfer rates, but that could be because cats aren't given names prior to transfer.

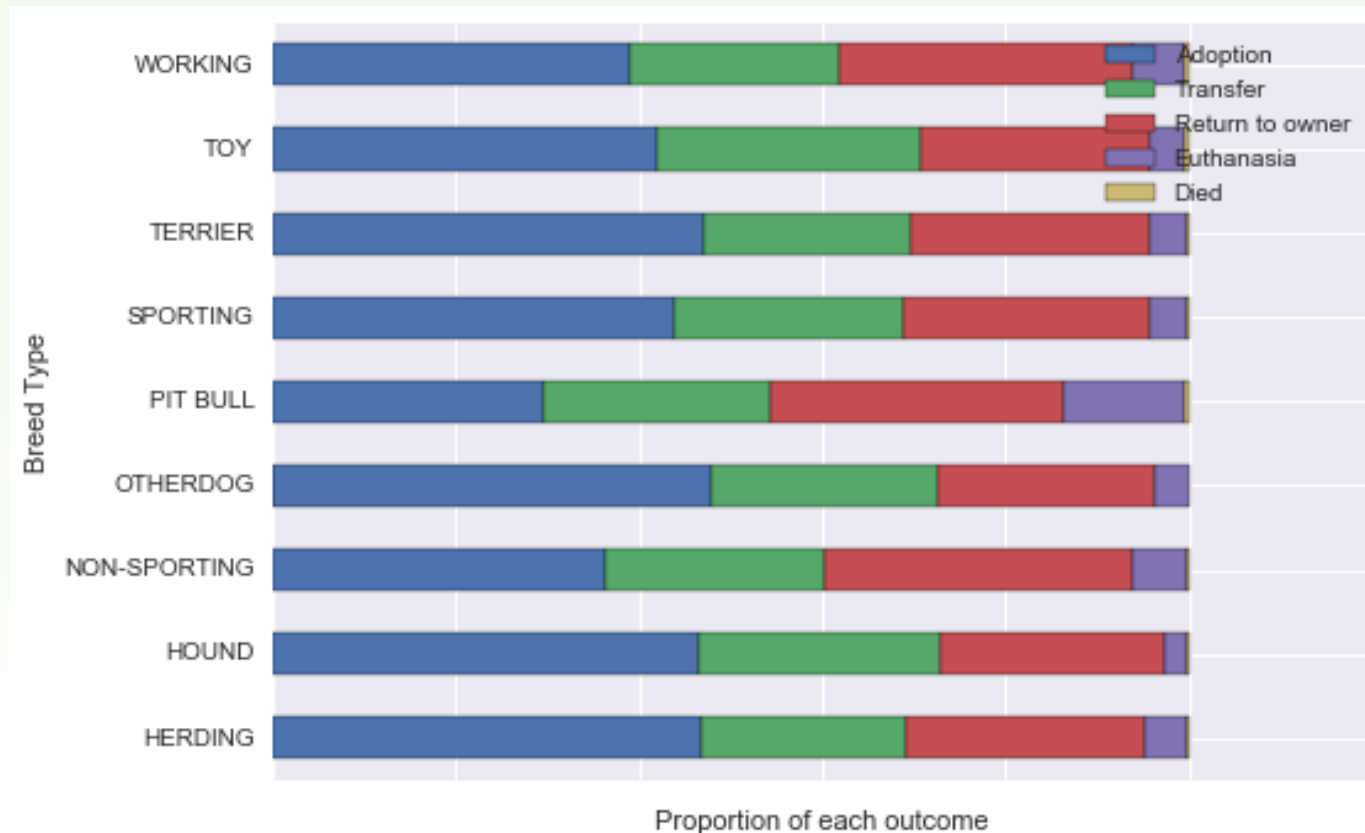
Cats



Dogs



Pit bulls have low adoption rates and high euthanasia rates



Summary of Recommendations

- Older animals, cats, and pit bulls may need more help getting adopted.
- Animals should be given names prior to putting them up for adoption.

Further Research

- The shelter should consider tracking animal behavior, which may be a good predictor.
- The shelter could also start tracking which animals are an owner surrender.
- They could consider turning this into a binary outcome problem, for example good outcome vs. bad outcome.
- The dataset doesn't have much information to predict euthanasia and death. The shelter may want to consider tracking animal health at the time of intake.
- I can also try multinomial logistic regression, knn, support vector machines, and neural networks.

Appendix

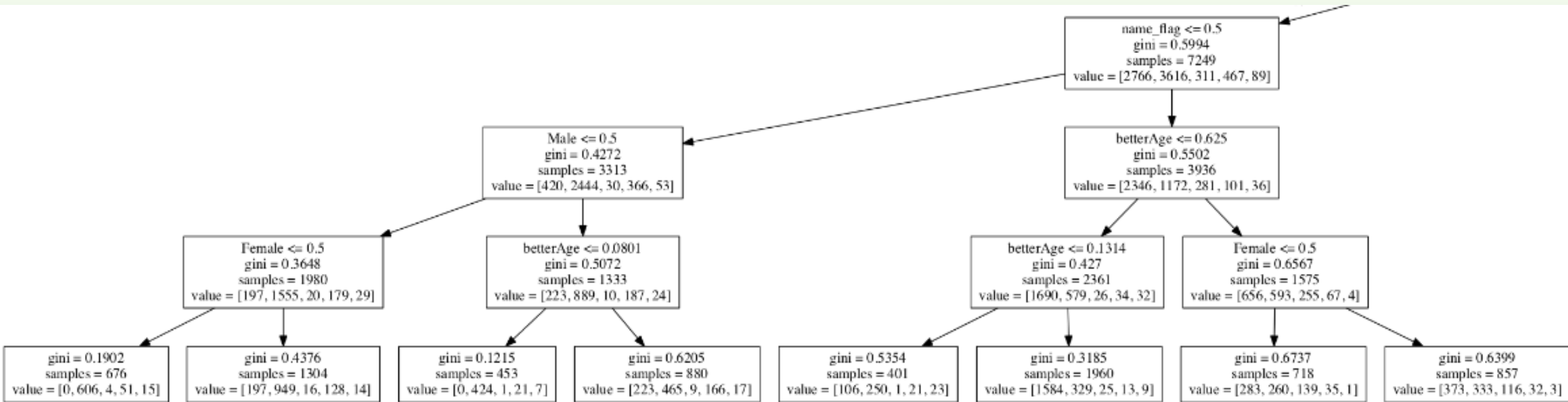
Logarithmic Loss – “Log loss”

- The logarithm of the likelihood function for a Bernoulli random distribution.

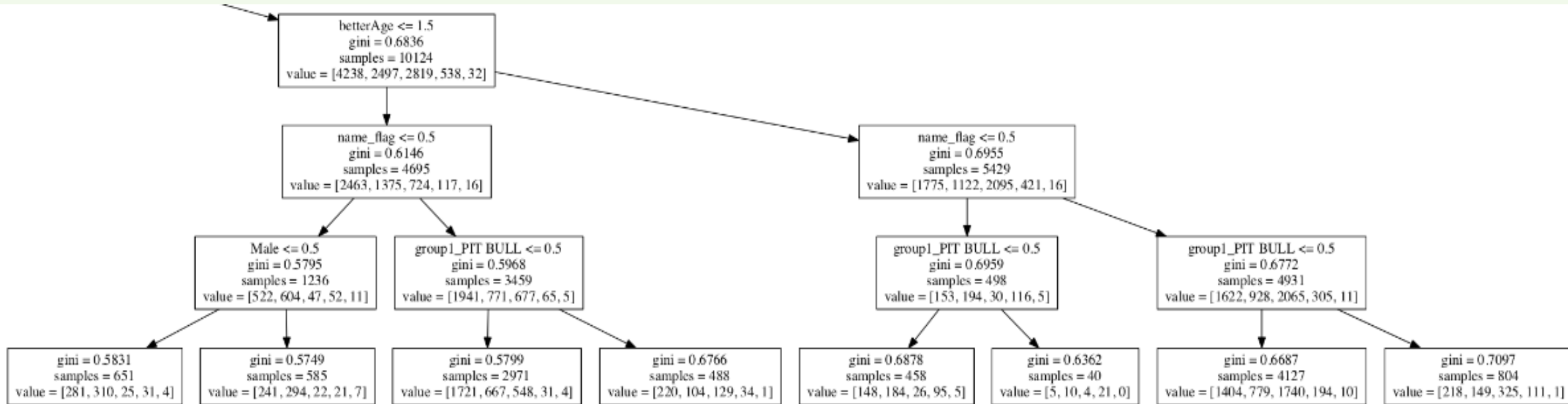
$$-\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} \log(p_{ij})$$

- N is the number of samples in the dataset
- M is the number of possible labels
- y_{ij} is a binary indicator of whether or not label j is the correct classification for instance i
- p_{ij} is the model probability of assigning label j to instance i.
- A perfect classifier has a log loss of 0.

Cat Decision Tree

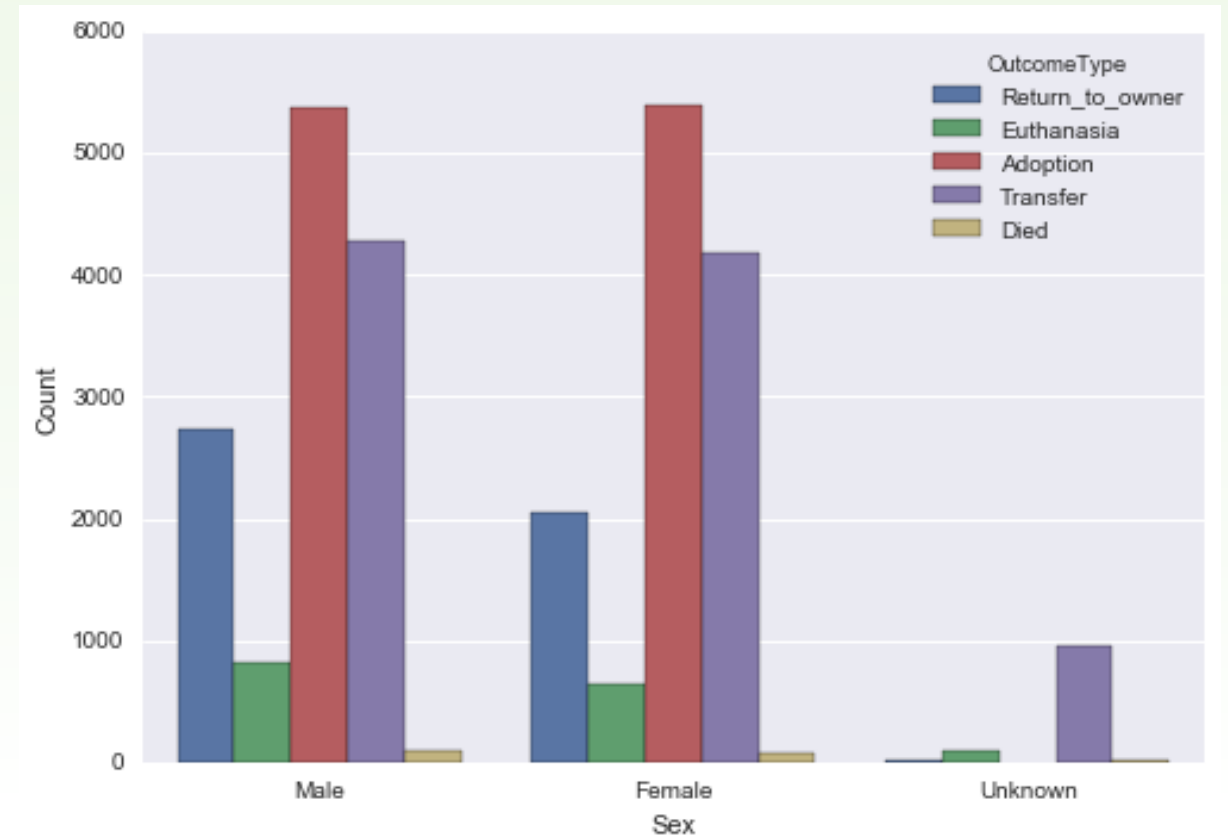


Dog Decision Tree

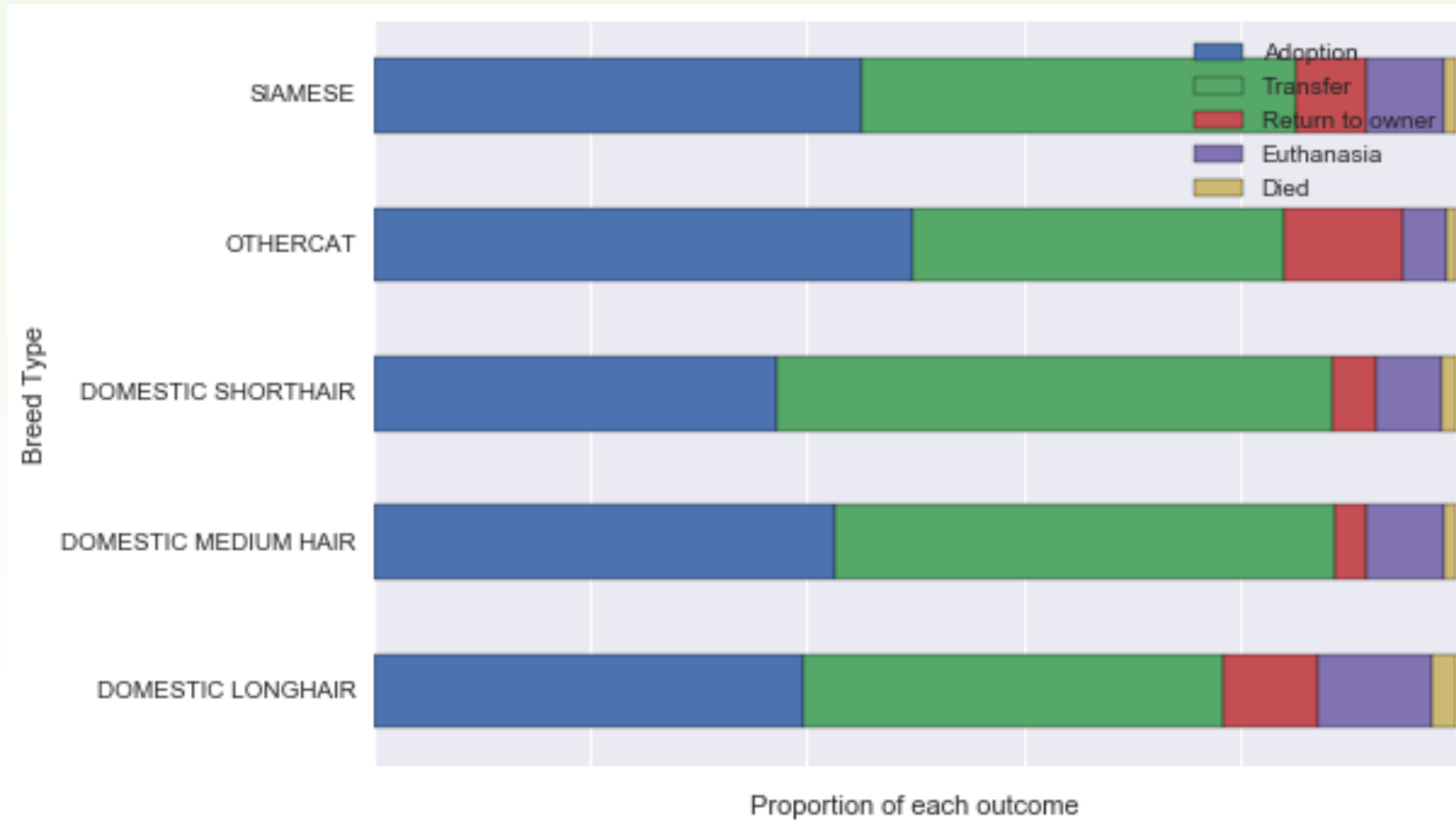


Female animals are less likely to be returned to owners

- Interesting finding that the shelter may want to explore.
- Why is this happening?
- Maybe female animals have puppies and the owners keep a puppy?



Cats with long hair have the highest euthanasia rates



Saturday and Sunday have the highest adoption rates

