Lydia Rogers
STATS 406
March 21, 2021
Final Project

Stigmas Surrounding Shelter Animals: A Case Study from Austin Animal Control

Introduction

Animal shelters across the United States face the daunting challenge of intaking and caring for millions of animals each year, while also working towards finding these animals forever homes. Shelters often face constraints due to lack of space, medical supplies, staff, and other resources necessary to care for a variety of animals. Due to these constraints, shelters, especially city-funded animal controls, often face the dilemma of either transferring or euthanizing animals in order to free up resources to intake new animals rather than risk waiting for these animals to be adopted.

In addition to the challenge of finding suitable homes willing to adopt these animals, certain superstitions or preconceptions about animal breeds or characteristics may lead to them being adopted less often, or spending more time at the shelter before being adopted. Two prominent examples of this phenomenon are pit bulls, which have a tendency to be seen as a more "aggressive" dog breed, and black cats, which are rumored in many cultures to be sources of bad luck. In this paper, I dive further into these two groups to determine whether or not they are less likely to be adopted than other animals of their species.

Although this analysis only uses data from a single shelter, Austin Animal Control located in Austin, Texas, similar analysis can be performed with other shelters in order to determine if these trends hold when observed on the state or national level. The results presented in this paper apply to the shelter population of Austin, Texas, where our data is derived from.

For this project, I aim to answer the following research questions:

1. Do black cats taken into AAC take longer to be adopted, or are adopted at lower rates than other cats?
2. Do pitbulls taken into AAC take longer to be adopted, or are adopted at lower rates than other dogs?
3. What variables have the greatest effect on the length of stay of an animal adopted from AAC?

A similar study to what is discussed in this paper is presented in Kogan (2013). In this analysis, the authors performed group-based analysis of black cats, cats with some black in their coats, and other cats to determine differences in the amount of time spent in the shelter. Overall, through analysis of cats and kittens at two different animal shelters, the authors estimated the means and standard deviations for the number of days spent at the shelter for each color of cat. The authors found that fully black cats did take a longer amount of time to adopt than partially-black cats and other colored cats, on average. Cats with coats that were not fully black did take less time to adopt than black cats at both of the shelters studied. Seeing as this trend is apparent, I choose to also see how trends surrounding superstitions apply to dog breeds, specifically pit bulls, as this trend appears to have not been studied.

This paper begins with a brief overview of the data and its contents. Then, in order to identify significant differences in adoption rates across animals based on breed and color to answer questions (1) and (2), permutation testing is performed with regards to the length of stay of pit bulls and black cats compared to all other dogs and cats, as well as the proportion of pit

bulls and black cats adopted as opposed to another outcome of interest. Finally, least squares regression analysis is performed to identify animal characteristics which may affect how quickly a particular animal is adopted in order to answer (3).

Data

The data used are provided by the City of Austin (TX). The two datasets respectively represent intakes and outcomes of animals at the Austin Animal Center from October 1st, 2013 to March 2021. This gives us data for 104,833 animals. Variables available for both datasets include identifying information of the animal (Animal ID and name) and characteristics about the animal (date of birth, sex and age upon intake and outcome, breed, and color). Additionally, the intake data provide details about the intake (location, time, type of intake, animal's condition during intake), while other outcome information is available in the outcome data (date of outcome, type and subtype of outcome). The outcome type describes what happened to an animal leaving the shelter (euthanasia, adopted, transferred to another rescue, returned to owner, etc.).

For the purposes of my analysis, several different data cleaning methods needed to be performed. First, I create new indicator variables for whether an animal is a black cat or a pit bull. I used the difference between the date of intake and the date of outcome for a particular animal to calculate the number of days the animal spent at AAC. This resulted in a few negative values, which indicate that an animal was adopted, but then returned to the shelter again. Since only the first intake for each animal is recorded, I use the difference in the date of first intake and the date of the final outcome to determine the length of stay. One difficulty with this dataset is that age at the time of intake and outcome is a character variable, represented as "X weeks", "Y months", "Z years", and so on. For this reason, it is not easy to convert this into an actual age, so I instead choose to create an indicator variable for whether or not a particular animal is over a year old. I am interested in the effect of age on length of stay and adoption, as I suspect potential adopters may show preference for a younger animal rather than an adult animal.

In order to define my adoption rates, I look at the possible outcomes for different animals. If the animal does not have an associated outcome (a value of NA), I may assume it to still be under the care of AAC. Outcomes I am not concerned with are those which are returned to owners or missing animals. This leaves animals that either died under the care of the shelter, were successfully adopted through AAC, or were transferred to another shelter. An overview of these desired outcomes for dogs and cats is shown in the table below.

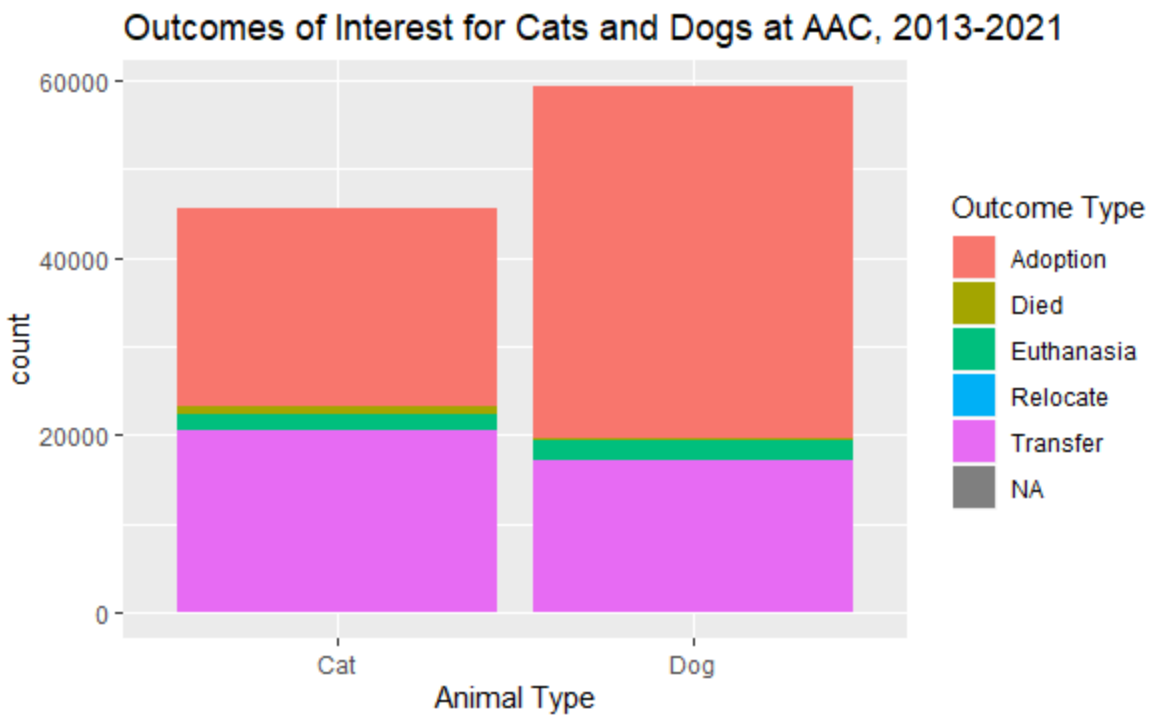

Figure 1: Outcomes of interest for cats and dogs taken into Austin Animal Center between 2013 and 2021. Data Source: City of Austin, TX

By using both of these datasets, I am able to track when an animal leaves and enters the AAC, and how long they may have stayed there without being adopted. By dividing the data further by species of animal, I am able to estimate the average amount of time a particular breed or species is spending at the shelter. I choose to use the median for these estimates, since some animals remain at the shelter for a very long period of time (up to seven years in some instances), making the mean an inaccurate measure for average length of stay. A portion of the distribution of length of stay for cats and dogs at AAC is shown in Figure 2 below.

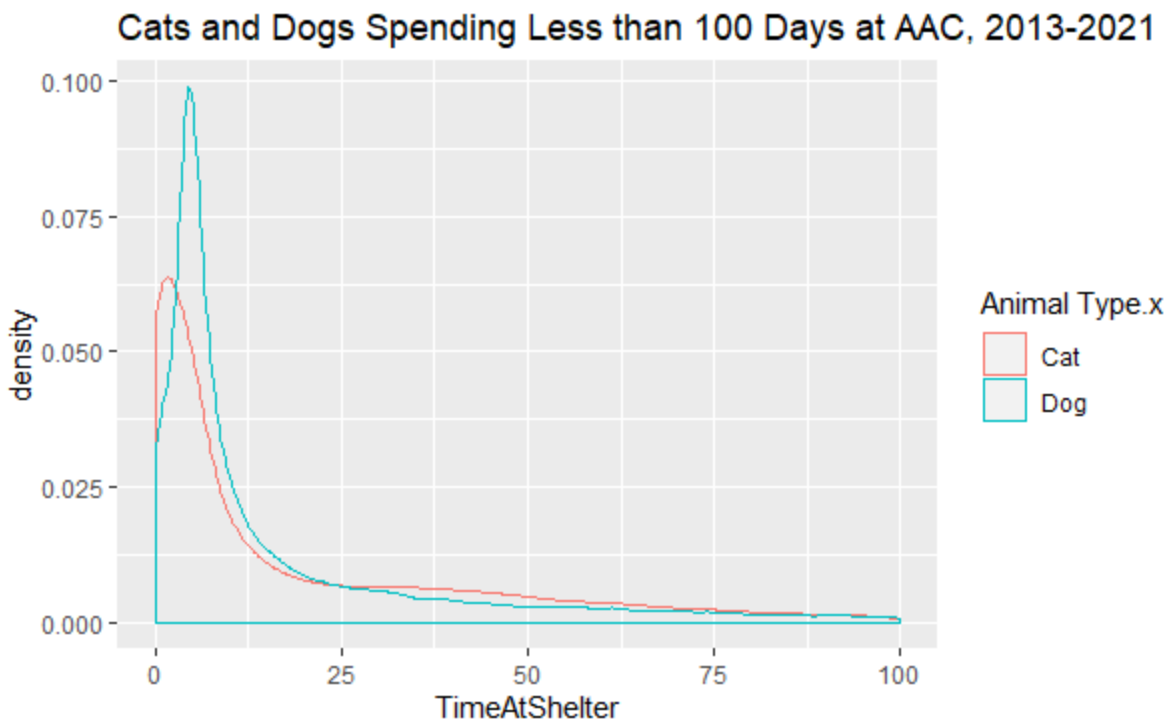

Figure 2: Distribution of length of stay for animals spending less than 100 days at AAC, 2013-2021. This excludes 10,098 cats and dogs that spent more than 100 days at the shelter. Data Source: City of Austin, TX

Looking at these data, it is apparent that a majority of the animals at this particular shelter spend a very short amount of time there. This is likely due to the shelter constraints mentioned in the introduction. The city of Austin, TX prides itself on being a no-kill city, which means that animals under the care of shelters there are not euthanized unless they are deeply sick or suffering. Therefore, I see many more animals being transferred to other shelters to make room for more intakes. I suspect that transferring or another outcome may occur more frequently for the less-desirable animals of black cats and pit bulls, and since very few animals intaken by this shelter are euthanized, I see more motivation to find these animals homes. I wish to see whether or not these efforts are successful for black cats and pit bulls.

Methods

The animals I am interested in for the purposes of this project are those that were surrendered, found, or otherwise rescued by the Austin Animal Center. From the possible outcomes for these animals, I exclude those who were returned to their owners or missing. This allows us to see direct adoption rates from the shelter itself as opposed to transferred animals and euthanized animals.

The statistics I am particularly interested in are the difference in adoption rates between pitbulls and other dogs, as well as black cats compared to all other cats at the AAC. I also am interested in the length of time spent at the shelter by each of these groups. If the difference in proportions is negative, this shows that these particular characteristics cause an animal to be adopted less frequently than others. If the difference in average number of days at the shelter is negative, this shows that particular characteristics lead to longer stays for certain animals.

To examine the difference in adoption rates between pit bulls and all dogs, as well as between black cats and all cats, I employ Fisher's exact test. This allows us to compare the two sample proportions in a non-parametric manner, as well as also estimate the odds that one group is more likely to be adopted than the other. This test has relaxed assumptions compared to other tests, such as a standard t-test. One of the most useful assumptions of this test is that I assume that the number of animals constituting each category (adopted, type of animal) are fixed rather than random. This is in fact the case for these data, as I assume that one animal being adopted over another animal is not random. Also, with animals at the shelter at the same time, it becomes very apparent that these counts are not random, as an adopter can only adopt animals that are at AAC at a particular time, rather than having an equal choice between all animals at the shelter

To estimate the difference in time spent at the shelter between these two groups, I perform a simple permutation test for difference in medians. This is done by sampling with replacement to determine whether or not a cat is a black cat or a dog is a pit bull, then comparing the difference in sample medians between black cats and other cats and pit bulls and other dogs. The median was chosen due to the fact that the distribution of length of time spent at the shelter is very right-skewed, meaning that some animals spend a disproportionately long time at AAC. Therefore, the median will give us a more accurate estimate of how long most of these animals are spending at the shelter. The advantages of this test are very similar to the advantages of Fisher's exact test, specifically that I do not need to assume that the data are random. Also, I need not assume that the two groups (in this case, black cats vs. other cats and pit bulls vs. other dogs) are independent. As discussed previously, these groups are likely not independent due to the nature of the adoption process.

In order to estimate the individual effects of several different variables on how long an animal spends in the shelter, I perform ordinary least squares regression analysis. This form of analysis is very flexible, allowing us to use both indicator variables and continuous variables to estimate the average influence of each variable while also accounting for interactions between variables. The variables I use for this case are an indicator for whether an animal is a black cat, an indicator for whether an animal is a pit bull, the type of intake (abandoned, euthanasia request, owner surrender, public assist, stray), an indicator for whether an animal is less than one year old, the animal's sex upon their outcome, and whether the animal is a cat or a dog. These variables are chosen for their significance in the dataset, as well as hypotheses as to which of the variables in the data are most likely to have an effect on the time it takes to get an animal adopted.

Simulation

In order to conduct simulations to get a sense of the procedures I employ, I treat the assignment of black cats and pit bulls as random binomial variables with a range of probabilities (0.01, 0.1, 0.25, 0.5, 0.75) of a cat being black or a dog being a pit bull. Notice that by doing this, I am treating the assignment of breed of dog and color of cat as random, which violates the original assumptions I am making about the data being non-random. In particular, by altering the sizes of the two groups in each scenario, I am testing the power of Fisher's exact test and my permutation test, as well as the prediction power of my proposed ordinary least squares model.

The result of Fisher's exact test is an odds ratio describing the odds of a positive outcome (in this case, a black cat being adopted and a non-black cat not being adopted) versus a negative outcome (a black cat not being adopted and a black cat being adopted). This, in more simple

terms, will give us the odds that a black cat is adopted over another cat. This same logic applies for the odds ratio of pit bulls compared to other dogs.

When performing Fisher's exact test on the new simulated data, I do see differences in power based on the sizes of each group. I particularly see that the power is largest when working with groups of similar sizes, i.e. the probability of being in a particular group is equal to 0.5.

To study the usefulness of the linear regression model, I consider the sum of squared residuals for the model fit later in the Analysis section. I also consider the p-value of the F-statistic of this model, which describes the model fit more accurately. The sum of squared residuals for this model is quite large at over 15 million, but this is to be expected due to the large number of rows in the data. However, the large value of this does raise some questions about the accuracy of model fit with this particular model. However, this model does have an associated p-value close to zero, as well as a number of statistically significant predictors. Although this model may not be the most accurate model for predicting new observations of this data (in this case, the number of days it takes for an animal to be adopted), this model does fit the data rather well, meaning that its coefficients are useful to describe the average differences in length of stays at AAC imposed by different variables.

Analysis

I perform Fisher's exact test on the populations of cats and dogs to determine the odds ratio for a black cat being adopted over another cat, and a pit bull being adopted over another dog. This is done by comparing the number of adoptions for each group to the number of all other outcomes for each group.

For black cats vs. other cats, I find an odds ratio of 0.8926, with an associated 95% confidence interval for this estimate of (0.8446, 0.9432). This analysis also gives an extremely low p-value. Based on these results, I can conclude that black cats are less likely to end up with an adoption outcome than all other cats, supporting my initial hypothesis that black cats are less likely to be adopted than other cats.

For pit bulls vs. other dogs, I perform the same test, and find an odds ratio of 0.8014, with an associated 95% confidence interval of (0.7653, 0.8392). Once again, I also find an extremely low p-value. This supports the hypothesis that pit bulls are less likely to be adopted than other dogs, as once again, I see an odds ratio less than one, which would indicate an equal chance of pit bulls being adopted as other dogs.

I then perform permutation tests for the difference in median length of stay for black cats compared to other cats, as well as for pit bulls vs. other dogs. For black cats, the difference in median length of stay is 7.66 days compared to other cats, meaning that the average length of stay is 7.66 days greater for black cats than other cats. In order to test the probability of this particular statistic, I resample the data 10,000 times to redistribute the assignment of black cats versus other cats with replacement. The resulting proportion of black cats versus other cats should thus be similar to the sample data. Then, using these new samples, I calculate the difference in medians between black cats and other cats. Using these new median differences, I compare these values to the observed median difference of 7.66 days to estimate the probability that I see the observed median difference, which is used as the p-value for the alternative hypothesis that the difference in medians is positive, meaning that black cats spend more time in shelters. In this case, the estimated p-value is zero, concluding that the observed difference in means is not due to chance, and supports my conclusion. The resulting distribution of differences in medians, as well as the actual observed value, is shown in Figure 3 below.

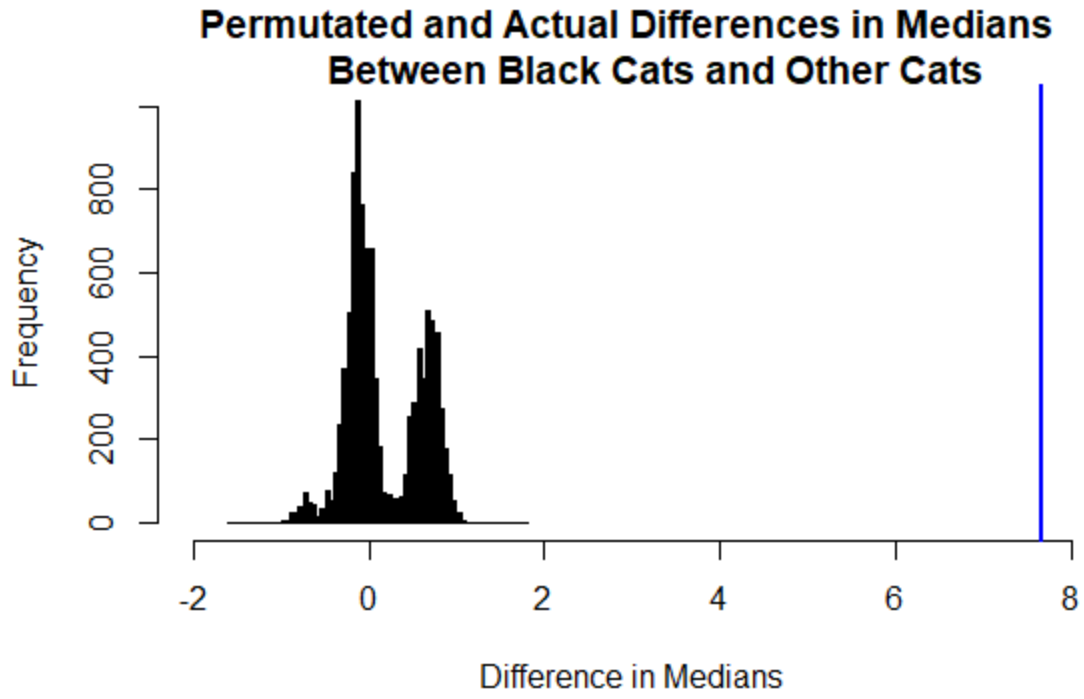

Figure 3: Permuted differences in medians generated through random assignment of black cats (black) versus actual observed difference in medians between black cats and other cats (blue). Data Source: City of Austin, TX

Next, this same analysis is performed, only this time will pit bulls versus all other dogs. Here, I see a much larger difference in the median length of stay, with a difference of 28.8, meaning that pit bulls spend almost 29 more days in the shelter compared to other dogs, on average. Unsurprisingly, this gives similar results to the same experiment performed on black cats, again giving a p-value of zero. These results support the alternative hypothesis that pit bulls take longer to be adopted than other dogs. A plot of permuted medians is given in Figure 4 below.

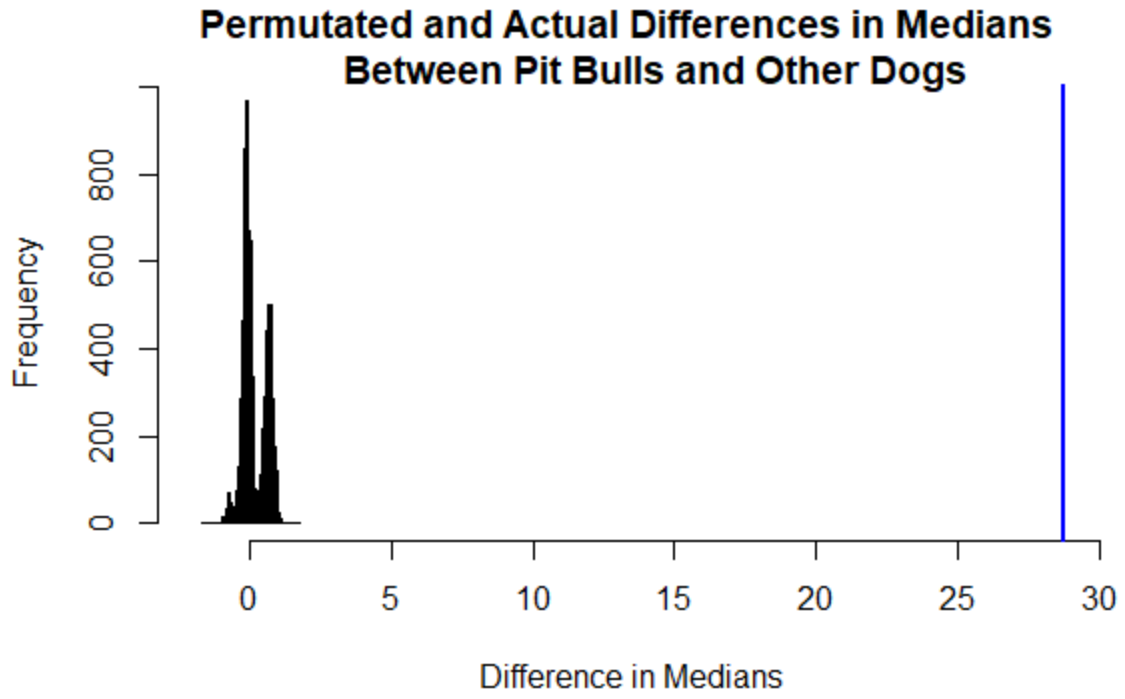

Figure 4: Permutated (black) and actual (blue) differences in medians for pit bulls versus other dogs. Data source: City of Austin, TX

	Coefficient Estimate	P-value
(Intercept)	59.396149	0.0000214
BlackCatTRUE	7.001565	0.0298024
PitbullTRUE	51.916079	0.0000000
`Intake Type`Euthanasia Request	32.179846	0.4282675
`Intake Type`Owner Surrender	12.292431	0.3493375
`Intake Type`Public Assist	110.615176	0.0000000
`Intake Type`Stray	26.697045	0.0415827
youngTRUE	-67.591264	0.0000000
`Sex upon Outcome`Intact Male	-1.278468	0.8572558
`Sex upon Outcome`Neutered Male	21.879873	0.0000065
`Sex upon Outcome`Spayed Female	19.953235	0.0000392
`Sex upon Outcome`Unknown	-21.190814	0.8171811
`Animal Type.y`Dog	-19.993332	0.0000000

Table 1: Ordinary least squares regression results predicting number of days spent at AAC, with associated p-values. Here, the reference variables used for different coefficients are “Abandoned” for Intake Type, “Intact Female” for Sex upon Outcome, and “Cat” for Animal Type. Data source: City of Austin, TX

Finally, I perform ordinary least squares regression analysis using the variables described in Methods on only the set of adopted animals. The coefficients describe the average difference in length of stay of adopted animals that fall into a particular group when all other variables are held constant. A positive coefficient means that an animal with the characteristic is likely to stay longer than an animal without, while a negative coefficient indicates that an animal with this coefficient is likely to spend less time in the shelter.

Interestingly, according to the results above, with all other variables held constant, adopted dogs spend around 20 days less in the shelter compared to cats, on average. The associated small p-value suggests that this relationship is strong. This suggests support for a hypothesis that wasn't of interest for this paper, but is one worth considering: the idea that dogs are more popular than cats and therefore adopted more quickly. This idea would require further analysis, of course.

I also see several other interesting findings based on the regression analysis. As previously hypothesized, this analysis does appear to support the idea that younger animals are adopted more quickly than older animals. With a regression coefficient of -67.6 for young animals, this suggests that with all other variables held constant, younger animals are adopted in 67.6 fewer days than other animals of their same characteristics. I also see that pit bulls spend more time in the shelter than all other adopted animals, as their corresponding coefficient is 51.9. This is almost two months of additional time that it takes for pit bulls to be adopted on average. Interestingly, the coefficient isn't nearly as large for black cats. Although black cats are also stigmatized, they only take about a week longer to be adopted than other animals on average.

Discussion

As a whole, the analysis and tests I have performed show a statistically significant difference in the amount of time it takes for black cats and pit bulls to be adopted compared to other members of their species. I have also shown that an adoption outcome is less likely for these animals than other members of their species.

Looking further at the results, it appears that pit bulls receive more discrimination, and therefore have proportionately fewer adoption outcomes and longer shelter stays, than black cats do. Although my research was originally driven by the stigma surrounding black cats, these results show that getting pit bulls adopted is a far greater problem that needs addressing. While the median difference in stay for black cats is only around a week greater than other cats, the median difference for pit bulls is estimated to be almost a month compared to other dogs. This shows that further publicity for pit bulls is needed, and more of an effort needs to be made to find these loving animals homes. Further research also needs to be done to determine why pit bulls are less preferred than other breeds, and if this is actually caused by the stereotype of pit bulls being violent or some other characteristic of their breed. For example, pit bulls are shorter-haired dogs, which may be less appealing to potential adopters than dogs of more visually-appealing breeds.

To extrapolate this data to other shelters, I would need to assume that all of the animals from a particular shelter are drawn randomly from a population of animals, which may not necessarily be the case. I can, however, use this data to draw conclusions about the Austin Animal Center specifically, and the animal preferences of the adopters from this shelter.

During the course of the analysis performed by this project, I find several other variables that have an effect on the length of stay and outcome of certain animals that may be worth

exploring. The bulk of these come from the regression analysis performed. First, I find that younger animals spend significantly less time at the shelter before being adopted. This is almost to be expected, as puppies and kittens are adorable and give adopters the opportunity to raise an animal from a young age and have an influence on its personality. However, the difference in the number of days spent in the shelter is surprising, with younger animals spending almost two fewer months than older animals on average. This indicates a need to promote the adoption of older animals, who also are in dire need of homes, and are also lower maintenance than younger animals. Additionally, I see that dogs spend less time in shelters than cats on average. It has long been standard that dogs are viewed as more popular than cats, but the extent to which this occurs is an interesting area for more in-depth research.

Regardless of these differences, more resources are needed in shelters on a constant basis. Many of the animals who enter AAC are transferred to other shelters, usually due to a lack of space, food, or staff needed to care for these animals. We are still seeing a good number of these animals being adopted, but further research is needed to determine whether the no-kill status of Austin, TX has any effect on this. In other shelters, it is very well possible that these less desirable animals would be euthanized more frequently than other animals of the same species. There are millions of animals in shelters in need of homes, including those animals with less desirable characteristics, and it is important that we find them loving homes.

Bibliography

1. Austin Animal Center. “Austin Animal Center Outcomes | Open Data.” City of Austin Texas, 21 Mar. 2021, data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238.
2. Austin Animal Center. “Austin Animal Center Intakes | Open Data.” City of Austin Texas, 24 Mar. 2021, data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Intakes/wter-evkm.
3. Kogan, Lori R., Regina Schoenfeld-Tacher, and Peter W. Hellyer. "Cats in animal shelters: Exploring the common perception that black cats take longer to adopt." The Open Veterinary Science Journal 7.1 (2013).