Animal Shelter Trends

By: Taylor Woodington

1. Business Problem

Many shelters experience an inflow of animals where some get adopted very quickly, and others take a long time or never get adopted. This project aims to help these businesses identify what important factors are likely to affect the time of adoption for an animal. Will the breed of the animal or age affect the adoption? Do the animals' coat color or gender alter their time of adoption? From there, if I can identify the main factors, I can target marketing for these animals better to aide in their adoption, which betters the wellbeing and life of the animal, but also decreases the strain on the shelter resources thus costing less.

1. Background/History

To aide in this business problem, I will be using an Austin Animal Center Shelter dataset that is found on Kaggle. The Austin Animal Center is the largest no kill animal shelter in the United States, so while this dataset does not contain data from shelters across the country, I believe it can raise areas of focus or questions for other shelters to investigate. This dataset includes information such as the individual animal identification number, their age when they were adopted, what type of animal they are, the breed, their coat color, the outcome of their adoption, and gender.

1. Data Preparation

In order to prepare for the data some cleaning and readjusting needed to be done. The dataset included different types of animals such as cats, dogs, and an “other” category. Due to “other” being an unknown category where it could mean any kind of animal, I removed those rows from the dataset for less confusion. Moving forward I converted the ages of all the animals to be uniform with one another on the unit of years, rather than ages being days, months, or years. Following this, I dropped unnecessary columns which will not be used such as the “name”, “date of birth”, “outcome subtype” and previous “age upon outcome” column. In addition to this cleaning, I extracted the month and year from the month and year column for future data analysis. Finally, I checked for missing values amongst the left-over columns and found only a few, I removed those missing values.

1. Methods

Once the data was prepared, I did some exploratory analysis and visualized the distribution of age across the years, amount in each outcome type, and the top 10 most common breeds in the dataset. After the exploratory analysis I dove deeper into the data to look for possible trends or features which affect the adoption of animals. I decided to make a linear regression model showing the adoption likelihood versus the age in years of the animal to see if there was an effect on adoption depending on how old the animal became. In addition, I chose to graph the adoptions across months to see if the holiday seasons or times of the year had an impact on animal adoptions. Following this basic graph, I made a more complex map that shows the number of adoptions made in different breeds throughout the months of the year. Lastly, I did a larger logistic regression model to find the top influential features of the dataset when put up against adoptions.

1. Analysis

Upon exploratory data analysis I was able to see that majority of animals were under five years old, the highest counted outcome of the different types of outcomes was adoption, and the most common breed amongst all the animals was the “Domestic Shorthair Mix”. Once the exploratory data analysis was done, I decided to focus a little more on the likelihood of adoption at different ages, what months and breeds where most popular during the year, and which features in the logistic regression showed a relationship with the outcome of adoption. When looking at the linear regression for adoption likelihood versus age in years as shown below, we see that as animals get older, the likelihood of them getting adopted does decrease. (See Appendix A.1)

When looking at what times of the year will be most popular, it seems to be that the second half of the year is more popular for adopting animals, and the most common adopted animal overall is the most common breed in the dataset, the Domestic Shorthair Mix cat for every single month. (See Appendix.2)

For the last model, the graph below visualizes the different features which were affecting adoption of the animal the most. This model performed with a 74% accuracy rate which is not the highest rating but is still a positive percentage which can give us areas for focus in shelters as a total. In the visual we can see that it is important to know the gender and have a spayed/neutered animal when it comes to adoption. There is a positive relationship with those animals as opposed to the animals with unknown genders. (See Appendix A.3)

1. Conclusion and Assumptions

Overall, from the data analysis I have concluded that the animals' age is important for adoption as well as knowing the gender of the animal and if it was spayed/neutered. These results do not surprise me as I had previously assumed that the older animals were the less likely they were to be adopted. Many individuals want a young animal for multiple reasons. These reasons can include they are more appealing with the idea of getting older with the owner and possibly less medical expenses. An article from the National Library of Medicine stated that gaps in shelters programs for older cats and dogs lead to that population being the most at risk for euthanasia. (*Hawes, 2018*)

1. Limitations and Challenges

With the Austin Animal Center being a no kill shelter, it is difficult to see the relationship discussed in the National Library of Medicine article. This article states how the likeliness of euthanasia increases as the age of an animal increases because of their decreased appeal in adoption. While we can see there is a decrease in adoptions as animals get older, it is difficult to understand what the relationship would be like without that regulation in place at another shelter.

1. Future Uses and Recommendations

While this analysis and shelter shows the expected relationship for older animals and adoption, other shelters may not. In the future, comparing shelters that does not have the no kill regulation to shelters that does could be beneficial in really understanding the relationship between animals age and adoption. I would recommend all shelters to analyze their data when it comes to older cats and dogs in their shelters and how often they are adopted in comparison to younger animals.

1. Implementation Plan

Once the understanding of relationships between animals' ages and adoption rates are supported by other shelters, promotion plans can be put in place for animals like those of older age. For example, spay/neuter programs which will decrease the number of young animals in the shelter could be beneficial in promoting the adoption of older animals. As well as programs which partner health care benefits for the older animals if adopted from the shelter, and promotional signage at shelters which educate people on the care for animals of different ages.

1. Ethical Assessment

Ethically it is important to consider that there may be bias in the dataset that does not reflect the relationship in other regions of the U.S. shelters. This data analysis is not to reflect on all shelters but to be a part of a step for other shelters and what they should investigate further. In addition, to acknowledge that the intention is not to deprioritize the other animals that have a higher chance of adoption, there needs to be balance in the approach. While promoting those less likely to be adopted, all adoptions and the factors of those adoptions should be recorded and kept in mind while removing the personal records of those documents such as the personal information of individuals adopting those animals.

1. Audience Questions and Answers
   1. Why did you remove missing values? Were they values that could have been filled in at all?
      1. Missing values were removed because it makes the data inconsistent and unusable from that row. There was categorized data missing this information, putting in the average category through the dataset would be misleading and cause bias.
   2. Would a random forest model been useful to this analysis?
      1. I considered random forest and tried this for the model. However, some random forest models overfit the training data and these models can be more complex and lack transparency in the data that is necessary to move forward.
   3. Why didn’t you separate the data analysis to focus on dogs and cats rather than them together?
      1. This may be useful in the future, but for preliminary analysis, it is useful to see overall what is affecting the shelters. If I were to compare cats and dogs equally, I could create a bias of there being an equal number of them going into the shelters (when there could be more cats than dogs, or vice versa). This alters the campaigning for specific breeds that need more promotions for adoption.
   4. If the animals are not getting adopted at a shelter which is a no kill shelter, what were other common outcomes for the animal?
      1. Other common outcomes for the animals that do not get adopted include transfer to another facility (which could practice euthanasia) and being returned to their owner.
   5. How many animals go to other shelters?
      1. About 22,937 animals from this sampled data went to other shelters with the outcome type as “Transfer”.
   6. What are the reasons animals may go to other shelters?
      1. Animals may go to other shelters due to lack of room at the current shelter or specific medical needs that the shelter cannot provide whether that is financial or otherwise.
   7. How would we monitor the ethical concerns of individuals personal information being included in the data that is analyzed?
      1. Personal information such as who has adopted an animal is often a part of animal records at the shelters. However, in the analysis of the shelters trend, this should not be included in the overall dataset being analyzed. To ensure this, the confidential data will be kept in a separate location from the data that is being analyzed for trends.
   8. Can you explain the implementation of the health care benefits with older aged animal adoption?
      1. To promote adoption of older aged animals, shelters can partner with local veterinary hospitals to create a deal which decreases the cost of veterinary visits or medical casts of the animal. This will bring business into that specific veterinary hospital but also promote the adoption of these animals by decreasing the financial burden.
   9. What are the other most common breeds of animals getting adopted?
      1. Some other common breeds of animals getting adopted are Pit Bull Mix, Labrador Retriever Mix, and the Chihuahua Shorthair Mix.
   10. What is the average age of the animals in the shelter?
       1. The average age of animals in the shelter is 2 years old.
2. References

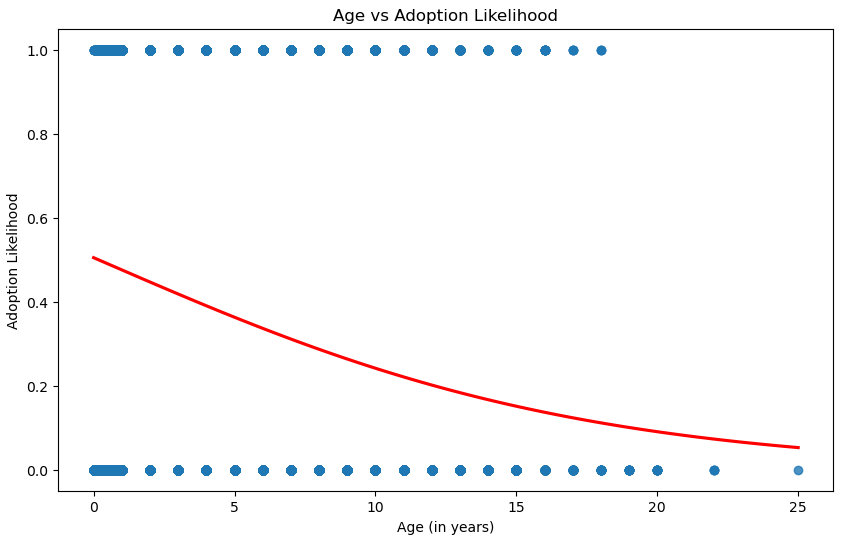
AaronSchlegel. (2018, February 16). *Austin Animal Center Shelter Outcomes*. Kaggle. <https://www.kaggle.com/datasets/aaronschlegel/austin-animal-center-shelter-outcomes-and/data>

Hawes, S., Kerrigan, J., & Morris, K. (2018, March 7). *Factors informing outcomes for older cats and dogs in animal shelters*. Animals : an open access journal from MDPI. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5867524/>

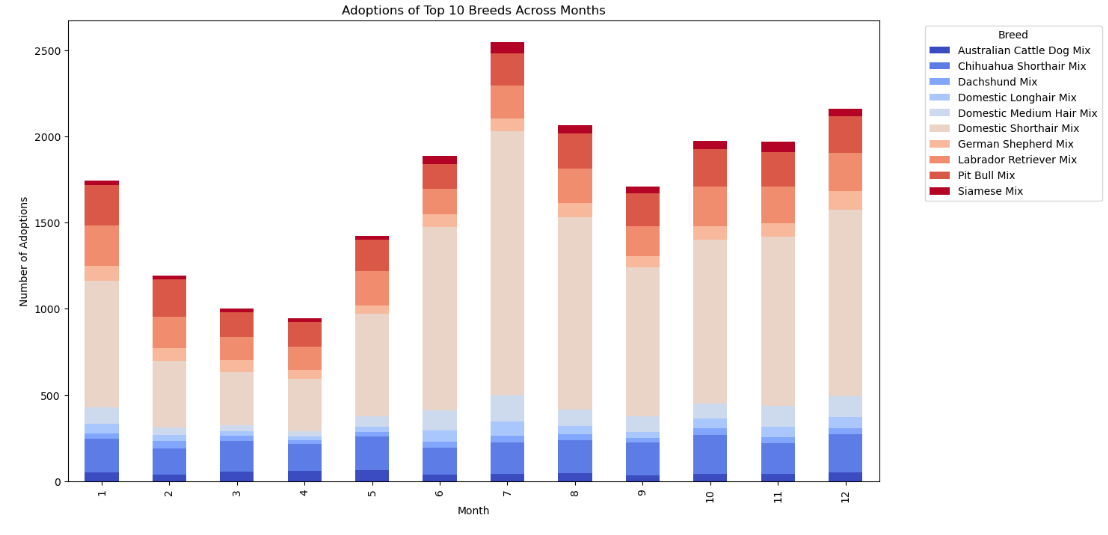
Publisher data.austintexas.gov. (2024, November 25). *City of austin - austin animal center outcomes*. Catalog. <https://catalog.data.gov/dataset/austin-animal-center-outcomes>

1. Appendix A

A.1: Age of the animal versus the adoption likelihood for them.



A.2: Adoptions of the top 10 breeds across all months of the year.



A.3: The top features affecting adoption through a logistic regression model.

