

Tony Tran Project Proposal

Problem statement:

How can Austin Animal Center increase the overall adoption rate of their dogs by adjusting the adoption page by increasing the type of animal filters, adjusting how adoptees are displayed on the adoption page, and by promoting recommendations based on adoption trends?

Context:

Austin Animal Center is an open-intake animal rescue where lost, wild, and surrendered animals in need of shelter are accepted regardless of age, health, or status. Currently, intake of animals is by appointment only and appointments are only available Monday through Friday. Only sick or injured animals will be accepted without an appointment. Their adoption page, not available on their front page, only allows filtering via species (cat/dog), sex, age, and size. According to their 2019 fiscal reports, 6092 adult dogs were admitted (9291 total, including puppies). For outcomes, a total of 10,856 outcomes were reported, 5337 dogs being adopted, 2887 returned to owners, and 2457 transferred.

Criteria for Success:

Provide insight into adoption trends and recommendations for ways to improve their adoption page to increase the overall adoption rate of dogs.

Scope of solution space:

- Multiple animal types accepted at Austin Animal Center
- Multiple different outcomes for animals, including a 2nd column for outcome type, which may be empty

Constraints within solution space:

- Only for the Austin Animal Center, not representative of other cities
- No reason given for owner surrenders/returned animals

Stakeholders:

Austin Animal Center

Key Data Sources:

-Outcomes

<https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238>

-Intake

<https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Intakes/wter-evkm>

The problem I am aiming to uncover insight into or solve is a way to increase the adoption rates of dogs at the Austin Animal Center by analyzing and interpreting trends in the intake and outcome data. I will be using the Austin Animal Center provided datasets, one containing the information of animal intake and one containing the information of the animal outcome. Both datasets are updated daily and both uniquely identify each animal by a 'Animal ID'.

This information could benefit Austin Animal Center by more easily promoting or displaying specific animals or filter options on their adoption page that would allow more personalized dogs to be visible to those looking to adopt. Austin Animal Center would benefit from any insight into their provided data because it would increase their revenue created through adoptions, decrease the amount of boarded animals, and benefit the local community by decreasing the amount of stray animals on the streets.

My plan to approach this project would be to first, import and clean the data. The two data sets can be merged on the Animal ID. There may be repetitions in the Animal ID column due to animals returning and being adopted on multiple occasions. Next, I would filter the merged dataset for 'Dog' entries in the 'Animal' column. I would clean the dataset by removing rows with no listed outcomes, that seemingly were data entry errors. I would check each column for unique entries and duplicate values to make sure there were no other data entry errors or misspellings. Next would be to create some informational charts to view distributions in the data, possibly charting on a horizontal bar chart the average number of intake and outcomes by dates, sorting by month and type. I could also sort the data by breed, size, age, color and chart the distribution of time spent at the Animal Center between intake/outcome on a box plot.

After exploring the data, I would use Exploratory Data Analysis to perform summary statistics, review data relationships, and create more in depth visualizations for the Austin Animal Center data on dogs. I would then fit and train my data to begin the process of selecting a model, training that model, and deploy the model to get predictive insight into ways to increase adoption rates. Lastly I'd display my research through visualizations and ultimately deliver a GitHub repo containing my work, a slide deck, and a completed project report.