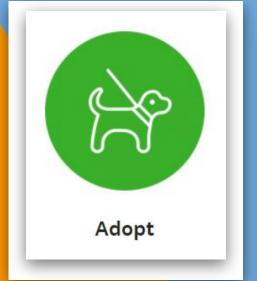
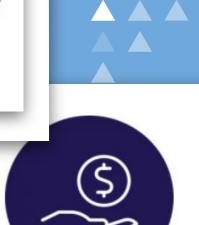
AUSTIN ANIML RESCUE











Donate



This project studies data from the [Austin Animal Center](http://www.austintexas.gov/content/austin-animal-center) in Austin, TX.

We chose this topic because we would like to help the Austin Animal Center better understand information on their animals so that as many animals as possible have a positive outcome.



Data Source Description

There are two datasets Animal Center as csv

Intakes (information on animals brought to the shelter) and outcomes (information on animals when they leave the shelter).

The information dates from 2013 to present and is updated daily on the Austin Animal Center data website. Variables in the dataset include an Animal ID, Animal Name, Animal Type, Breed, Color, Dates of intake and outcome, Sex, etc.

Links to the raw original datasets:

- [Austin Animal Center Intakes](https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Intakes/wter-evkm

- [Auslin Animal Center Outcomes](https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238)

We have two main questions that we are interested in answering, along with related subquestions.

Question 1 is a categorical prediction
Question 2 is a regression prediction.

- Question 1: Can we predict the outcome for an animal based on other characteristics?
 - What are the possible outcomes that we should consider?
 - What factors most influence the determination of the outcome?

- Question 2: Can we predict the length of stay at the shelter for animals?
 - What factors most influence the length of stay?
- Are the factors that influence the length of stay different for different animals?



Preliminary Machine Learning Mo<mark>del</mark>

predicting For the animal outcome, our current sketch of our pipeline is to move from the SQL data into Python for cleaning, splitting into the training and testing sets, and then utilize a Random Forest Classifier.

Load in data output from SQL server Clean, convert, and scale data Split data into training and testing sets

Train model with Random Forest Classifier

Test model and evaluate performance

The roles will change during the final project so that each member has the opportunity to learn about each piece of the project and practice the related skills.

Communication Protocol

- **Role 1:** Repository management. This team member leads efforts to maintain the GitHub repository, including resolving merge conflicts, and help keep the main branch as the source of our most recent working code. This person also updates the Readme.md file on the main branch as changes are made to the GitHub repo.

- Role 2: Project management. This team member leads the efforts for knowing what deliverables are required for the UT Bootcamp at each stage and assuring that the work the team is doing leads to successful fulfillment of the deliverables. This includes comparing the work to the rubric requirements as posted for class and helping to decide which technologies are used at each step of the project.

Communication Protocol

- **Role 3:** ML modeling. This team member works on the coding aspect of the ML model as well as data cleaning and exploratory data analysis.
- **Role 4:** Database & Dashboard management. This team member maintains and updates the database (PgAdmin 4) as needed and leads the efforts for creating and maintaining our final dashboard.
- **Role 5:** Presentation management. This team member writes the presentation files and helps other team members as needed.



