Adoptions in Austin Animal Shelter – Review of Correlation Between Breed Attributes, Surrender Conditions

Team Ten – Chris Flint, Robert Perron, Cindy Gorbas

**April 14, 2019**

Approximately 68 % of American households have pets. Unfortunately, sometimes there are circumstances when pet owners are not longer able to keep their pets. We have decided to look at possible factors in successful dog adoptions. Using data from Dogtime.com and the Austin Animal Shelter, we sought to discover which factors could lead to successful adoption.

**Methodolgy**

**Extraction**

We decided we would limit our efforts to dogs…because we love dogs. From Kaggle.com we found two data sources that we believed would be helpful.

* From dogtime.com we retrieved a json file (rating.json). This json contained data such as breed, personality, health and activity. The data was ultimately read into a dataframe (dogs\_df). This data was read into a dataframe (shelter\_df)
* From the Austin Animal Center we used a csv file (aac\_intakes\_outcomes.csv). This file was extensive, and provided data such as reason for intake (admittance to the shelter), how long animal spent in the shelter, etc.

**Transform**

Once both sources of data were in dataframes, the data was ‘cleaned up by

renaming columns to eliminate wanted spaces in the column names  
eliminating animal types (cat, other, etc) to ensure we were only looking at dogs

parsing columns and removing unwanted characters

dropping unwanted columns.

The dataframes from different sources were merged into one dataframe

New columns, the means in shelter days, were calculated and added to the combined dataframe.

**Load**

The team loaded the data using Mongo. Dataframes were converted into json files then loaded into mongo ‘database’.

* ETL\_project\_dogs – mongo database
* ‘dog\_data’ – mongo collection created by inserting/loading the final\_dog\_json (*final\_dog\_data\_df*) data
* ‘shelter\_data’ - mongo collection created by inserting/loading the final\_shelter\_json (*shelter\_details\_df*) data
* ‘total\_data’ - mongo collection created by inserting/loading the final\_data\_json (*final\_dog\_data\_df*) data

**Resources**

**Data Sources:**

<https://www.kaggle.com/aaronschlegel/austin-animal-center-shelter-intakes-and-outcomes/version/1>

<https://www.kaggle.com/hocop1/cat-and-dog-breeds-parameters>

**Other Sources:**

<https://www.statista.com/statistics/198086/us-household-penetration-rates-for-pet-owning-since-2007/>