Using Animal Control Incident Data to Improve Animal Welfare by Community

Rebecca V. Lewis

Bellevue University

DSC 680 – Fadi Alsaleem

April 4, 2021

Table of Contents

Abstract

According to research estimates from the American Society for the Prevention of Cruelty to Animals (ASPCA), over 6.5 million animals enter shelters each year [7]. Of those, approximately 1.5 million animals are euthanized [7]. Animals of all species and breeds enter shelters in multiple ways. They can be surrendered by their former owners, rescued by individuals or authorities as strays, or confiscated from a cruelty case. [9] Most municipalities have an animal control or humane law entity that is responsible for responding to animal complaints and emergencies. Sometimes these agencies are publicly funded, and sometimes they are part of a private nonprofit animal welfare organization. [10] Animal control officers sometimes have a bad reputation, but often their desire is to maintain the best quality of life for animals in the community. Common problems that they work towards improving include pet overpopulation, cruelty and dogfighting, and feral cat community control. The objective of this project is to analyze animal control incident data collected by an animal control agency and use it to provide vital information to address animal welfare problems.

Introduction/Background

Animal control entities and nonprofit animal rescues often work together to improve the lives of animals in their local communities. Common problems that face the animal community can benefit from a closer look at animal control statistics. A few of these problems include pet overpopulation, cruelty and dogfighting, and feral cat community control.

Pet overpopulation is caused by factors like allowing animals to breed without the likelihood of finding homes for the offspring. [5] This should not be confused with reputable breeders who specialize in a specific breed of dog or cat because they often have an audience of people searching for those animals. Unintentional breeding happens when owners do not spay and neuter their pets due to a lack of education or resources. Communities with the most stray animal incidents or incidents involving young/pregnant animals can benefit from spay/neuter education and grants.

Dog fighting is a brutal underground criminal activity that has led to a bad reputation for pit bull type dogs. In Louisiana specifically, our shelters are overrun with “bully” type breeds and are among the hardest to place in homes. Dogs who win fights are often overbred and their puppies are sold for high dollar values. The weaker dogs are often used as bait dogs to train the fighters or are disposed of in an inhumane manner. [2] Animal control incident records indicating the conditions of dogs with wounds indicative of fighting can bring attention to a potential ring in the area.

Feral cats are free roaming cats that are considered wild and cannot easily be domesticated. [11] They can offer benefits to their respective communities but are often picked up by animal control because they are considered a nuisance by neighbors. These benefits include offering rodent and other pest control as well as giving the community members who may not be able to care for a pet of their own a sense of purpose.[6] Feral cats are difficult to place once in a shelter because they are not accustomed to human contact and are very likely to be euthanized. Locations that have a high population of cat incidents may benefit from education about the benefits of feral cat colonies and may be eligible for grants to start new feral at programs.

Far more problems plague the animal welfare community but the scenarios above provide a starting point for discovering opportunities that may lie in the animal control incident data.

Data

The dataset used for this analysis comes from the Baton Rouge Animal Control and Rescue Center (ACRC) who describes their mission as “protecting animals from people and people from animals” in the community. [3] The ACRC is responsible for investigating animal related complaints, resolving animal related issues, and enforcing the local laws surrounding animals. They work with the Companion Animal Alliance (CAA) which is the nonprofit agency that takes care of animal sheltering, adoptions, rescue, fostering and euthanasia. [3] The raw data was downloaded from Baton Rouge's Open data website.[4] A previously cleaned dataset with geocoded locations was found on Kaggle using Census.gov and Google geocoding services in addition to custom Python scripts. [1] The original raw dataset contained 95,329 observations of animal control incidents from 2012 through 2020 with 38 data points. The clean dataset with geocoded locations contained 52,585 observations with 16 variables between 2012 and 2019. The raw data was used as the primary dataset and was filtered down to include the incidents with geocoded locations only. After data preparation steps were completed, the final dataset contained 38,288 observations and 26 variables for exploratory data analysis. A dictionary of both the raw dataset and final dataset used for analysis can be found in Appendix A.

Methods

Data Preparation

To clean and prepare the data for analysis, Python 3 was used in a Jupyter notebook to view general information and summary statistics about each of the variables in the raw dataset. The file number, latitude, and longitude from the Kaggle dataset were extracted, deduped, and joined with the raw data using an inner join that reduced the number of observations to 55,889. Columns were renamed to remove spaces and correct spelling errors.

Variables from the original datasets were removed for various reasons. The complainant street was removed to protect the identity of the person reporting the incident. Variables that were irrelevant to the analysis and removed include call taker, dispatcher, dispatched other, officer, equipment number, collar color, disposition date, and disposition officer. The file number and impound number were also removed after they were used for dataset merging.

Missing values were handled in different ways depending on the column. The dispatched time, arrival time, available time all had very few missing values; however, when one was missing, the others were also missing. Due to the low impact on the number of observations and to improve analysis performed on time deltas, observations with these values missing were removed. If there were too many missing values to provide useful information, they were removed as well: vaccination number, vaccination date, pet name, request type, and age. Missing values for categorical variables were grouped into an unknown category if one was already present in the data.

Categorical variables were evaluated for frequency and duplicates. Categories with similar names and common meanings were combined to reduce cardinality in service code, breed, and sex. The species variable was primarily cats and dogs. Initially, the non-cat and non-dog values were going to be eliminated, but 30% of the data would be removed. Instead, all wildlife, exotic and livestock animals were grouped into the other category.

Inconsistent data was observed in the species and breed combinations. For example, DSH stands for domestic shorthair and is a general breed of cat. The dog species contained some breeds of DSH as well as parrot, rabbit, and other non-dog categories. Any observations with a combination of species and breed that did not make sense were removed from the dataset to avoid any confusing results in analysis.

New variables were created from the data provided in the initial raw dataset. The incident, dispatched, arrival, available, and impound time values were listed separately from the date. To make calculations on the dates easier, the incident, dispatched, arrival and available times were combined with the incident date to create new variables. The impound time was combined with the impound date. Then the difference was taken between some of these times to create some variables that could be used in analysis including the difference between the incident call time to the dispatch, arrival, completion and impound time. The difference was also taken between the dispatch to arrival time and the arrival time to the completion time. These variables can be explored to see if they have an impact on the outcome of the incident.

Exploratory Data Analysis

The pandas_profiling library was used in a Jupyter notebook using Python to explore the distribution and preliminary visualizations of the variables. Phik's coefficient was used to review correlations between the variables because it can handle both categorical and numerical data. The incident call to dispatch time was highly correlated with incident call to completion time. This makes sense because longer times between the time the call occurred and the time the officer was dispatched would increase the time taken to complete the call. The high correlation between the dispatch to arrival time and the arrival to completion time is interesting since they are not necessarily dependent on each other. The longer it takes for an officer to arrive at an incident positively influences the amount of time it takes to complete the incident. The species and service code are also highly correlated meaning the type assigned to the incident by is influenced by the species. Species is also highly correlated to the sex of the animal. Sometimes it is hard to determine the sex of some wildlife due to anatomy or their temperament so the unknown sex can be prevalent in the other species.

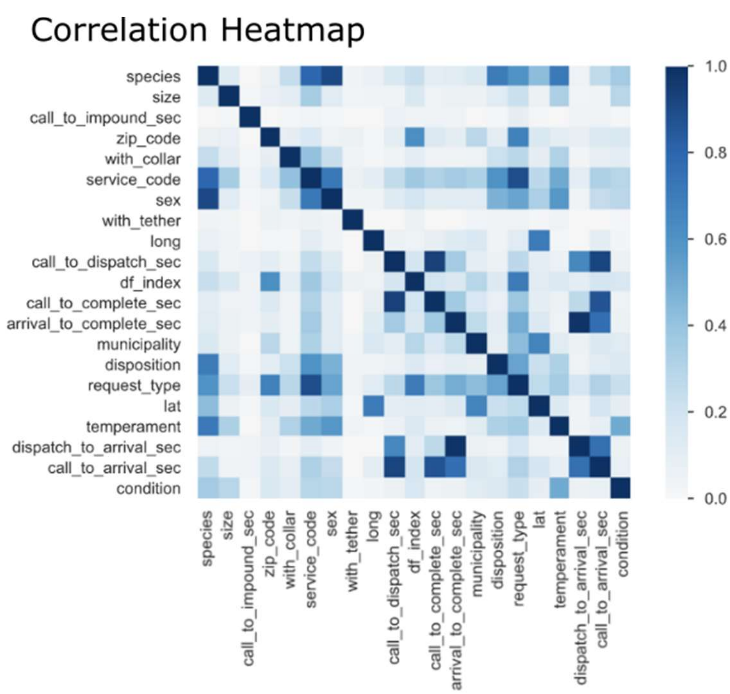

Figure 1 Correlation Heatmap

The impound date and time values contain a low percentage of missing values. The zip code column contains a high number of zero values and can be removed since we already have geocoded locations. The calculated time deltas contain some zeros as well as negative values. The zero values could be valid however negative values indicate that something is wrong. Because they make up a small percentage of the total data, they will be removed before analysis.

Data Visualization

Preliminary visualizations were performed using Tableau to view the geographic disbursement of categorical variables to answer our primary questions. The graphics presented in this report were enhanced using Inkscape. When plotting the incidents on a map, outliers were noted both outside of East Baton Rouge parish and outside of Louisiana. These could either be anomalies due to geocoding errors or isolated cases that required the officers to cross parish and state boundaries, noted in red in Figure 2. Regardless, before training any models, the data was filtered to only include spatial coordinates that fell with in the max coordinates for East Baton Rouge Parish.

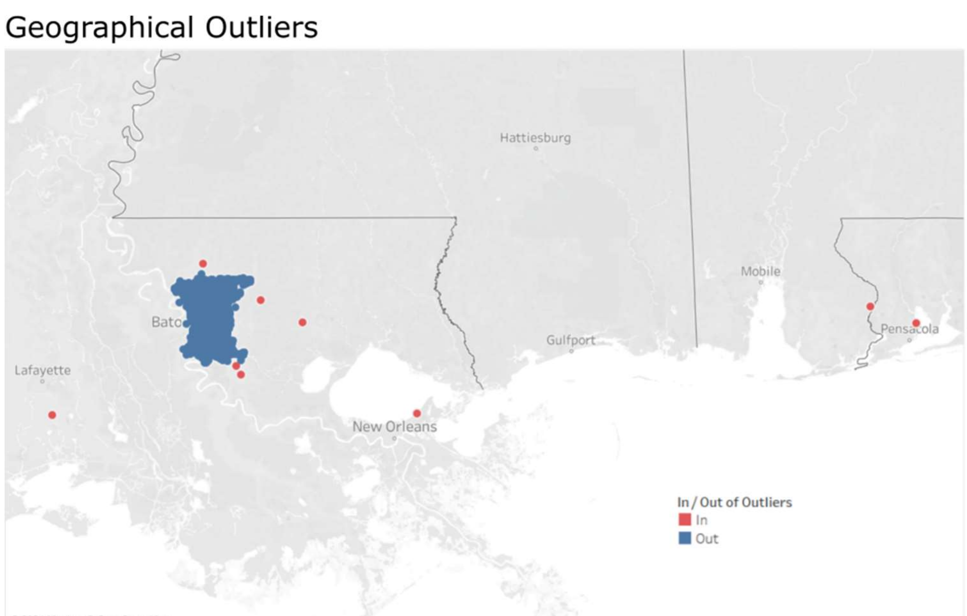

Figure 2 Geographical Outliers (noted in red)

Target Variables

Condition and temperament describe the health and demeanor of the animal. It would be useful to predict this information based on the information available to the dispatcher at the time of the call so that officers can be prepared for an emergency or dangerous situation. The condition category is moderately imbalanced with 63% of the observations in the Fair category. Temperament is also imbalanced with 48% of observations in the Normal category. For both models, only the features available to the dispatcher at the time the call is received will be used as features: incident call time, location, species, and request type.

Feature Engineering

The objective is to predict the condition and temperament of the animal before officers get to the scene. To avoid feature leakage, any variables that are most likely assigned after the officer arrives to the scene were excluded from the features. Temporal features were extracted form the incident date and time for the month, day of the week, and hour of the incident. A K-means clustering algorithm was used to group the latitude and longitude into clusters for the location feature. The final dataset used for modeling contained 38,288 observations and 12 features.

Modeling

Modeling was performed in Python through a Jupyter notebook using Pycaret's classification library to evaluate the best model for the dataset for the temperament and condition separately. Because the target variables were highly imbalanced, the AUC (area under the curve) and the F1 score were used to grade the model's performance. Ten models were tested for both the condition and the temperament target variables using a stratified k-fold cross validation and SMOTE to account for the imbalance. The training set contained 26,801 observations and the test set contained 11,487 observations.

Analysis

The dataset contains a variable for municipalities within East Baton Rouge Parish, the area served by the BRACRC. Four are specific cities: Baker, Baton Rouge, Zachary and Central (see Figure 3). The remaining observations are either attributed to an Unknown location or the Parish as a whole. A total of 85% of incidents are recorded in Baton Rouge; however, many of the latitude and longitude values were outside of the immediate city. Because of the unreliability of this variable, the latitude and longitude variables were used to create visualizations to understand the locations of incidents.

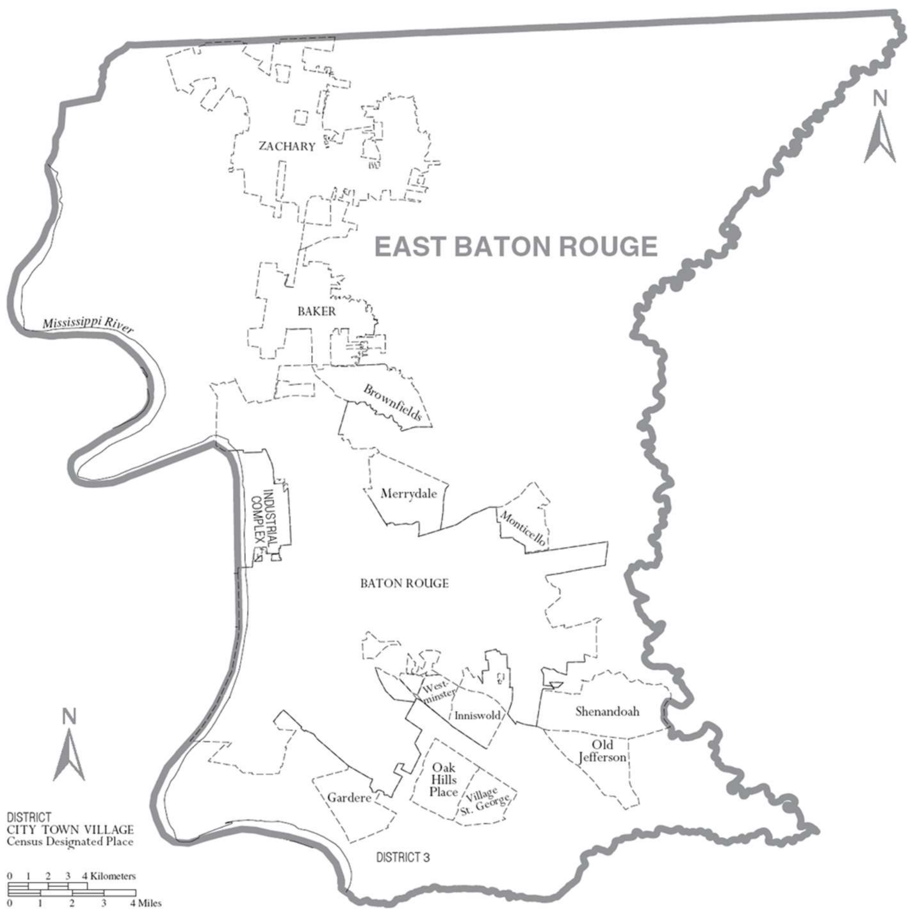

Figure 3 East Baton Rouge Parish Municipalities [8]

Pet Overpopulation

The preferred method to assess overpopulation was to use the age to determine if young animals were found more often in certain areas. However, the age variable was missing or zero in 99% of the incidents in the dataset. To get a view of the general overpopulation across the parish, the data was filtered to include service codes of Abandoned Animal, Animal Rescue, Stray, and Trapped Animal. These service codes are relevant because they describe situations where litters of animals can possibly be picked up. In addition to filtering service codes, species other than cat or dog were excluded. The largest grey dot representing stray animals is located just south of the Industrial Complex in Baton Rouge along the river, which contains plants, refineries and manufacturing facilities (Figure 4). Unfortunately, the Mississippi River is a common dumping ground for unwanted animals. A large group of incidents involving trapped animals can be spotted in the rural area of Zachary. There seems to be other clusters for stray and trapped animals scattered throughout south Baton Rouge.

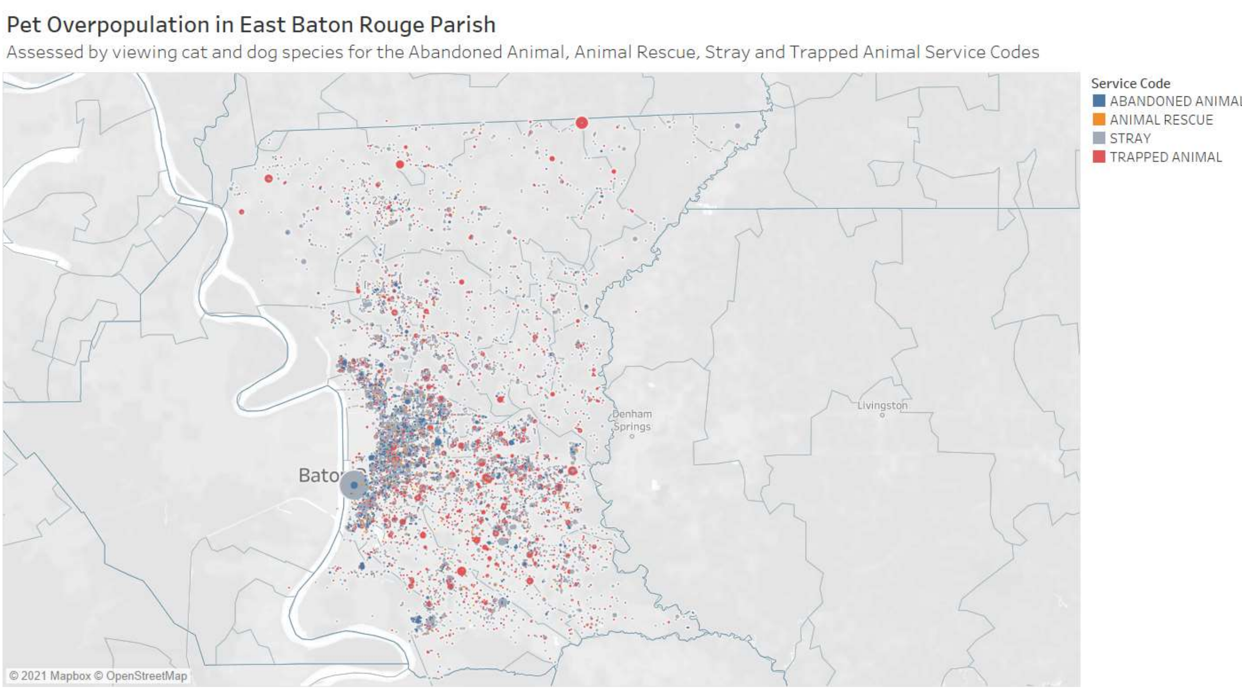

Figure 4 Pet Overpopulation Incidents

Dog Fighting

To assess dog fighting in the area, I first looked at the incidents that were coded as dog fighting calls along with the temperament of the animals that were involved. Most dog fighting incidents seem to involve nervous dogs and some clusters can be observed again near the industrial complex and in Zachary (Figure 5).

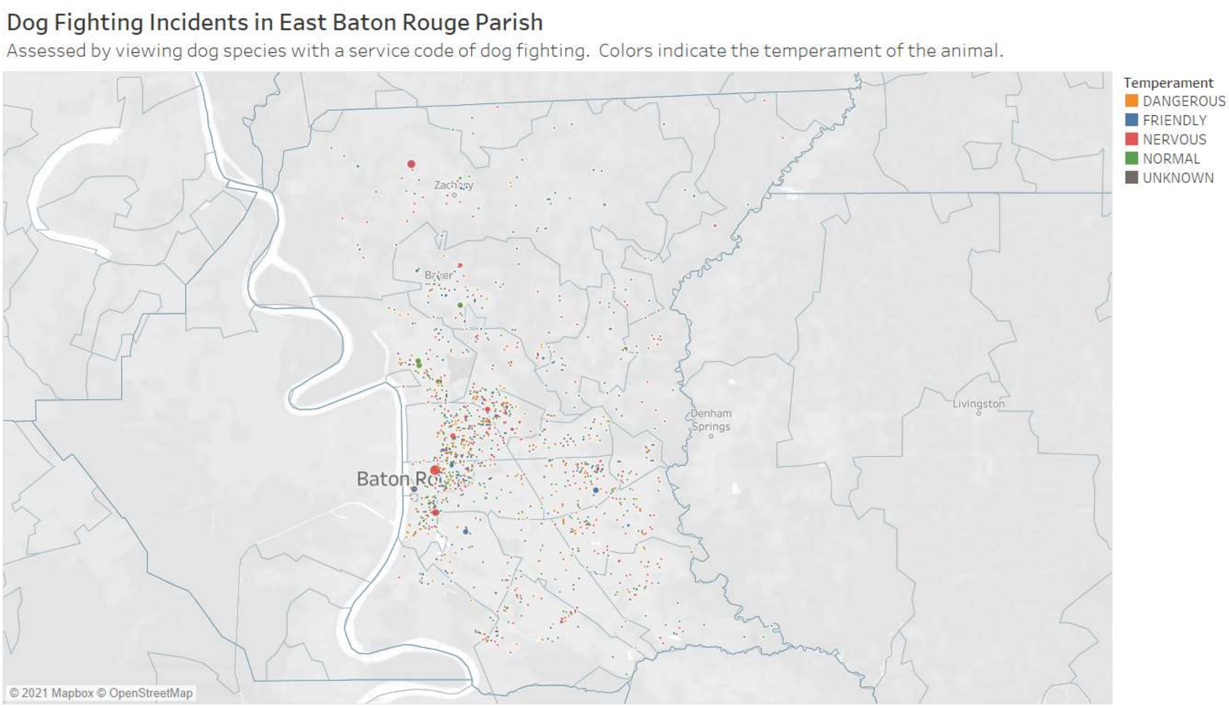

Figure 5 Dog Fighting Incidents

Another approach to identifying areas that may contain unknown dog fighting rings is to look at pit bull incidents and the conditions in which the dogs were found. Unfortunately, the data was not specific as to whether the animals had wounds; however, the most relevant conditions would be fair, emaciated, and poor. The emaciated clusters of incidents are located mostly in the city limits of Baton Rouge. The largest concentration is again located south of the industrial complex (Figure 6).

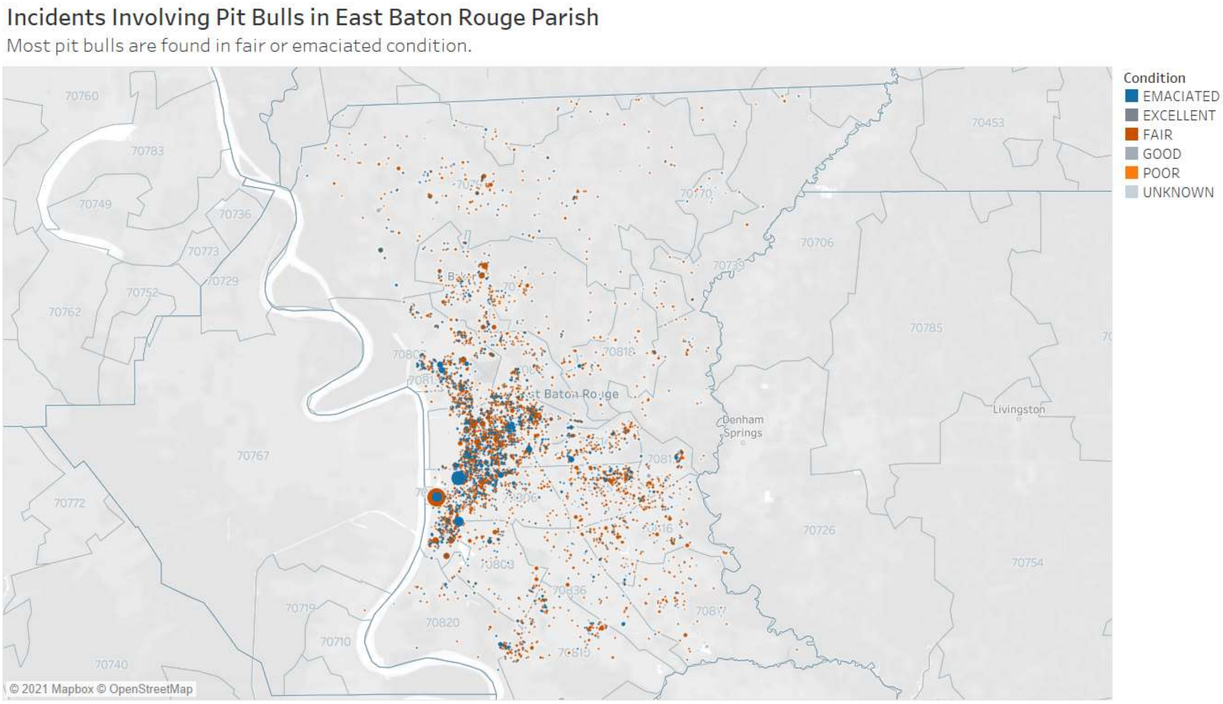

Figure 6 Pit Bull Incidents

It is important to note that Pit Bulls picked up in incidents are most often in Fair condition at 57% followed by emaciated at 22% (Figure 7). Despite their likelihood of being in poor health, 94% are transferred to the affiliated rescue for routing through the adoption process. More good news for pit bulls is that 70% have a friendly or normal temperament. Overall, dog fighting only represents .41% of all service codes reported during incidents in the time frame.

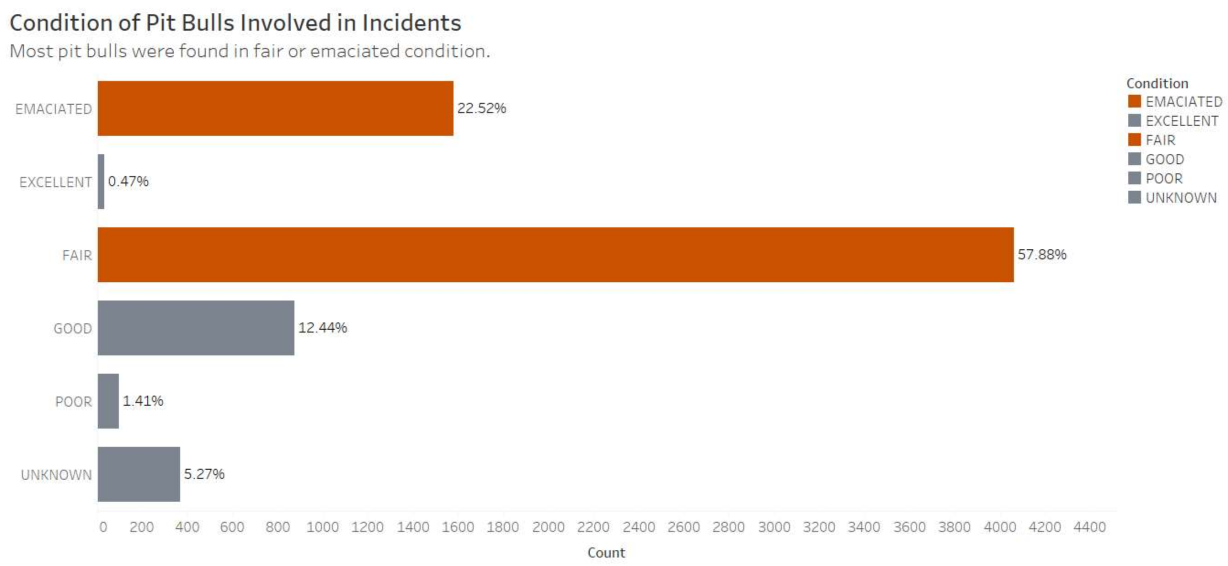

Figure 7 Pit Bull Conditions

Feral Cats

The service code assigned to most of cat incidents is trapped animal at 58% followed by stray at 22%. Feral cats must be trapped because they will not tolerate human interaction. Looking closer at the trapped animals, 31% of them were labeled with a nervous temperament and 3% were labeled as dangerous (Figure 8).

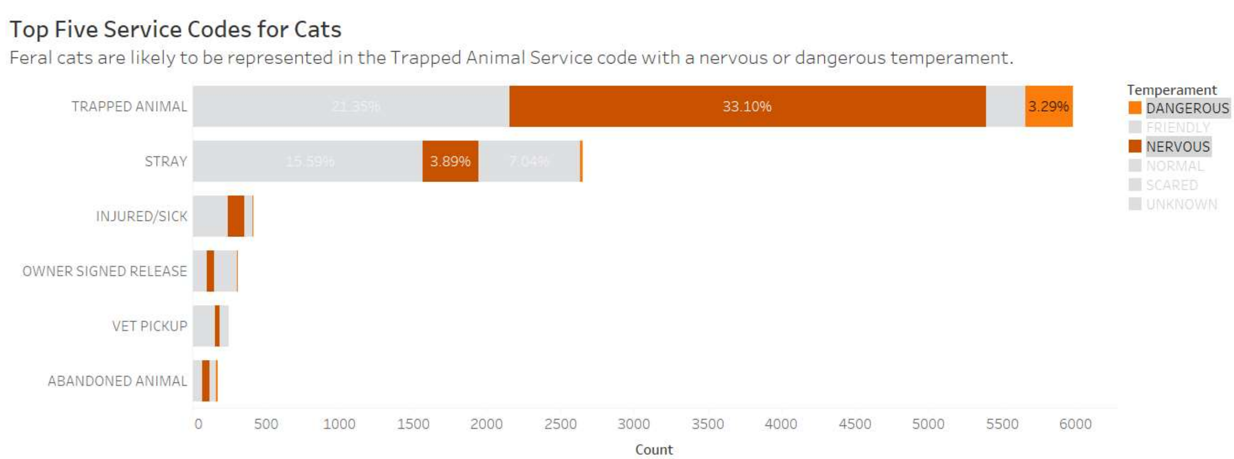

Figure 8 Feral Cat Service Code Distribution

To assess potential feral cat communities, incidents with the stray and trapped animal categories were plotted and filtered to only include cats that had nervous, scared or dangerous temperaments. Most of these incidents are in the city of Baton Rouge, with a large number below the industrial complex and a small cluster in Zachary (Figure 9). Ninety-seven percent of cats in this group were transferred to the affiliated rescue.

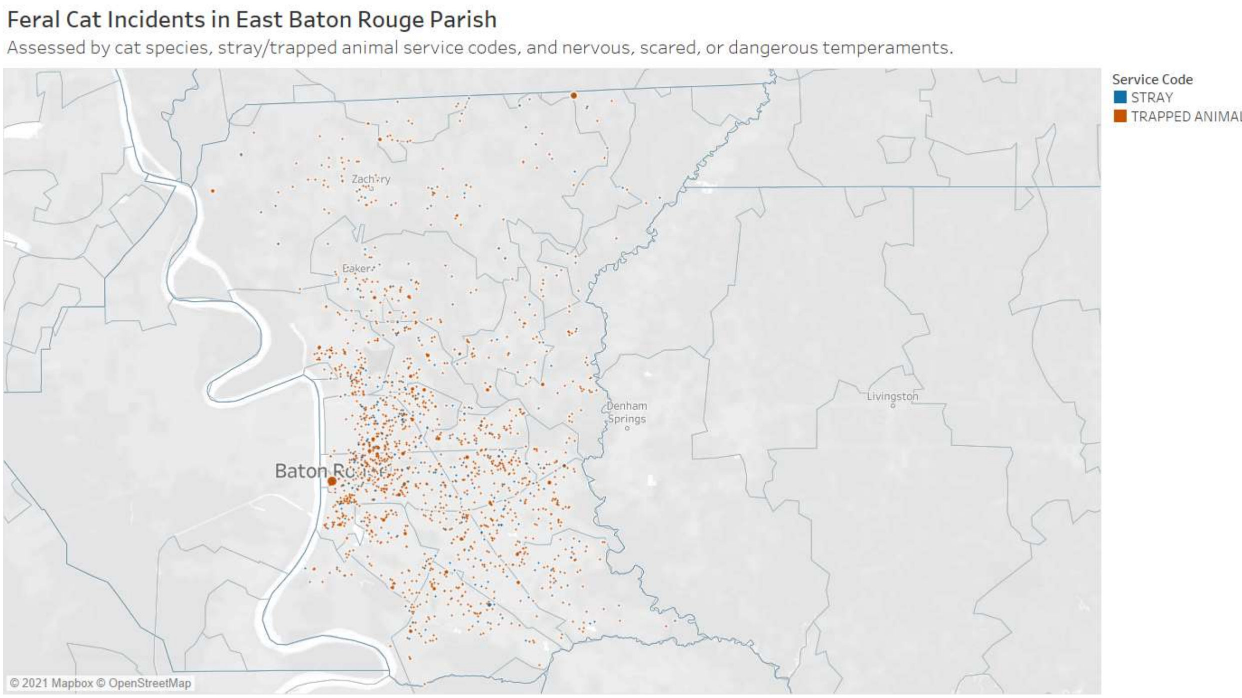

Figure 9 Potential Feral Cat Incidents

Results

Condition

The Random Forest classifier was the best performing AUC true positive rate of 70.65% and F1 score of 62.67%. Tuning the hyperparameters did not result in improved results so the original model was used for training and testing. The model was most accurate at predicting animals that were found in an excellent condition with a true positive prediction rate of .84; however, the false positive rate at that score was .65. We do not want to send unprepared officers out to a call and risk it being a dangerous situation. A more acceptable cutoff for the excellent condition's AUC would be at .70 which provides a false positive rate of 10%. It was least effective at prediction animals that were found in fair condition at with an AUC of 70%; however, that is still considered to be a good true positive rate. The false positive rate is around 50% at that cutoff. The other classes were also fairly accurate: emaciated 80%, good 72%, and poor 72%. The amount of time taken to dispatch the officer was the most important feature that influenced the prediction with a size of extra-large in a close second. When predicting the result on the test data, the AUC score increased to 71.54% and the F1 score increased to 63.34%.

Temperament

The Random Forest classifier had the best AUC score of 81.47% and F1 score of 61.10%. Tuning the hyperparameters did not result in improved results so the original model was used for training and testing. It was most accurate at predicting friendly animals with a true positive rate of 86%; the false positive rate was acceptable at 20%. It was least effective at prediction animals that were found to be scared at with an AUC of 49%. Animals that were considered normal had a true positive rate of 81%; however, the false positive rate was almost 50%. This would not be acceptable if the animal was actually dangerous. A more appropriate cutoff was chosen with a true positive rate of 70% and false positive of 20%. The amount of time taken to dispatch the officer was the most important feature again followed by the other species. When predicting the result on the test data, both the AUC and F1 score increased slightly to 81.55% and 61.69% respectively.

Discussion

An initial look at both models would indicate that they are both good at predicting the condition and temperament of animals involved in incidents. When predicting the condition, the AUC for all the classes were at least 70% accurate or above. The most important classes to predict would be animals in poor and emaciated conditions so that officers with veterinary training or local veterinarians could be prepared. The emaciated animals were predicted true accurately 80% of the time and the poor animals were predicted accurately 72% of the time. For temperament, the AUC scores were a little more dispersed. The most useful temperaments to predict would be dangerous, nervous, or scared to make sure the more seasoned officers could respond and bring backup. The model resulted in true positives for dangerous animals at 76% and nervous animals at 79%. Scared animals only resulted in a 49% true positive rate so the model would be ineffective at predicting that class correctly.

Conclusion

The objective of this project was to analyze animal control incident data to determine if it can provide vital information that could improve the welfare of animals in each community. Overall, we can say that yes, this data can be helpful in directing locations for additional animal welfare support programs. In each of the problems we assessed, pet overpopulation, dog fighting, and feral cat control, most incidents were centered around the city of Baton Rouge with a large number just south of the Industrial Complex. Additional patrols can be set up in this area to determine why these problems are so prevalent there. Before introducing new programs into the area, further analysis should be completed on the specific neighborhoods within the city of Baton Rouge to isolate any clusters. The predictive models were effective at predicting conditions and temperaments that would be most beneficial to know before an animal control officer arrives on the scene. This model could be used as part of an alert system within the animal control dispatch software to notify them that initial support is needed. The model should be tested further on unseen data to further evaluate its effectiveness.

References

[1] *(2012-2019) Baton Rouge, LA Animal Control Calls*. (n.d.). Retrieved March 20, 2021, from https://kaggle.com/dustinec/2019-baton-rouge-la-animal-control-calls

[2] *A Closer Look at Dogfighting*. (n.d.). ASPCA. Retrieved March 20, 2021, from https://www.aspca.org/animal-cruelty/dogfighting/closer-look-dogfighting

[3] *Animal Control & Rescue Center | Baton Rouge, LA*. (n.d.). Retrieved March 20, 2021, from https://www.brla.gov/148/Animal-Control-Rescue-Center

[4] *Animal Control Incidents | Open Data BR*. (n.d.). Retrieved March 20, 2021, from https://data.brla.gov/Public-Safety/Animal-Control-Incidents/qmns-hw3s

[5] *Animal Population Control*. (n.d.). American Humane. Retrieved March 20, 2021, from https://www.americanhumane.org/position-statement/animal-population-control/

[6] Collaborator, B. (500, 58:00). *8 Benefits of Community & Feral Cats (infographic)*. K&H Pet Products. https://khpet.com/blogs/cats/8-benefits-of-community-feral-cats

[7] *Pet Statistics | Shelter Intake and Surrender | ASPCA*. (n.d.). Retrieved March 20, 2021, from https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics

[8] Ruhrfisch, U. C. (2007). *Map of East_Baton_Rouge Parish, Louisiana, United States with municipal boundaries*. taken from US Census website [1] and modified by User:Ruhrfisch. https://commons.wikimedia.org/wiki/File:Map_of_East_Baton_Rouge_Parish_Louisiana_With_Municipal_Labels.PNG

[9] *Shelter Intake and Surrender*. (n.d.). ASPCA. Retrieved March 20, 2021, from https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender

[10] *Understanding Your Local Government & Animal Control*. (n.d.). Alley Cat Allies. Retrieved March 20, 2021, from https://www.alleycat.org/resources/guide-to-local-government-animal-control/

[11] *What is a Feral Cat?* (n.d.). Feral Cat Spay/Neuter Project. Retrieved March 20, 2021, from

http://www.feralcatproject.org/what-is-a-feral-cat

Appendix A: Data Dictionary

Raw Dataset from Baton Rouge Open Data Site [3]

Column Name	Description	Type
FILE NUMBER	Internal file number of incident	Number
INCIDENT DATE	Date incident was called in	Date & Time
INCIDENT TIME	Time incident was called in	Plain Text
DISPATCHED TIME	Time an officer was dispatched to incident	Plain Text
CALL TAKER	ID of employee who took the call	Plain Text
DISPATCHER	ID of employee who dispatched the officer to incident	Plain Text
DISPATCHED OTHER	Not dispatched by cad but there is dispatch info in record	Plain Text
DISPATCHED SITUATION	Indicates if officer was dispatched (D) or the call had to be postponed (K) and addressed the following day.	Plain Text
ARRIVAL TIME	Time officer arrived on scene	Plain Text
AVAILABLE TIME	Time officer was finished with incident and was once again available for another call	Plain Text
COMPLAINANT STREET	Street of complaintant	Plain Text
REQUEST TYPE	Type of incident being reported	Plain Text
OFFICER	ID of officer who was dispatched to incident	Plain Text
EQUIPMENT NUMBER	ID of vehicle used to respond to incident	Plain Text
REMARKS	Remarks provided during initial call, if any	Plain Text
IMPOUND NUMBER	Number used to track animals being impounded into shelter	Plain Text
IMPOUND DATE	Date animal was brought into shelter	Date & Time
IMPOUND TIME	Time animal was brought into shelter	Plain Text
LOCATION	Location of incident	Plain Text
MUNICIPALITY	Municipality of incident	Plain Text
ZIP CODE	ZIP code of incident	Plain Text
SERVICE CODE	Code is used at the responding officer's discretion; may be the same as CPREMARKS but may be different if issue is not what complaint reported	Plain Text
SPECIES	Species of animal	Plain Text
BREED	Breed of animal	Plain Text
SEX	Sex of animal	Plain Text
SIZE	Size of animal	Plain Text
COLOR	Color of animal	Plain Text
AGE	Age of animal	Plain Text
CONDITION	Condition of animal	Plain Text
TEMPERMENT	Temperment of animal	Plain Text
COLLAR	Type of collar, if any	Plain Text

COLLAR COLOR	Collar color	Plain Text
VACCINATION NUMBER	Vaccination id number of animal, provided by owner	Plain Text
VACCINATION DATE	Date of vaccination of animal, provided by owner	Date & Time
PET NAME	Pet has a name. Used when turning in animal to be adopted.	Plain Text
DISPOSITION	How animal was processed once at ACC (e.g. transferred to the Companion Animal Alliance, an organization with which the City-Parish is contracted to perform all animal sheltering, adoptions, and animal reclaimations)	Plain Text
DISPOSITION DATE	Date of disposition	Date & Time
DISPOSITION OFFICER	Name of officer responsible for disposition of animal	Plain Text

Final Dataset used for Analysis

Column	Description	Type
request_type	Type of incident being reported	string
location	Location of incident	string
municipality	Municipality of incident	string
zip_code	ZIP code of incident	int64
service_code	Code is used at the responding officer's discretion.	string
species	Species of animal	string
breed	Breed of animal	string
sex	Sex of animal	string
size	Size of animal	string
condition	Condition of animal	string
temperament	Temperament of animal	string
disposition	How animal was processed once at ACC	string
incident_datetime	The combined date and time of when incident was called in	datetime64[ns]
dispatched_datetime	The combined date and time of when the officer was dispatched.	datetime64[ns]
arrival_datetime	The combined date and time of when the officer arrived at the incident	datetime64[ns]
available_datetime	The combined date and time of when the officer completed the incident and was available for another call.	datetime64[ns]
impound_datetime	The combined date and time of when the animal was brought into the shelter	datetime64[ns]
lat	The geocoded latitude of the incident location. Retrieved from previously cleaned dataset on Kaggle.[1]	float64
long	The geocoded longitude of the incident location. Retrieved from previously cleaned dataset on Kaggle.[1]	float64

call_to_dispatch_sec	The time in seconds between the incident call time and dispatch time	float64
call_to_arrival_sec	The time in seconds between the incident call time and time the officer arrived	float64
dispatch_to_arrival_sec	The time in seconds between the dispatch time and the time the officer arrived	float64
arrival_to_complete_sec	The time in seconds between when the officer arrived at the incident and when he/she finished	float64
call_to_complete_sec	The time in seconds between the incident call time and the completion time of the incident	float64
call_to_impound_sec	The time in seconds between the incident call and when the animal was impounded	float64
with_collar	Indicates whether the animal was found wearing a traditional collar or non-controversial item (shirt, bandana, sweater). This can indicate the animal had previously been in a home.	bool
with_tether	Indicates whether the animal was found wearing a chain, choke collar, rope, belt or other non-traditional collar or leash type. This can indicate the animal came from an environment that may need more education on responsible pet ownership or a cruelty environment.	bool