A PRELIMINARY ANALYSIS:
INCREASING ADOPTION RATES AT THE AUSTIN ANIMAL CENTER

August 2023

Prepared by:
Sumaiya Ahmed
Andrea Messmer
Shamim Munshi
Charlotte Schnoebelen
Yana Zykova

1.0 INTRODUCTION

With animal shelters in the United States collectively taking in millions of companion animals each year and being stretched to capacity, there is interest in maximising the number of animals with positive outcomes - that is, adoption into loving forever homes as quickly as possible. Using data from the Austin Animal Center (AAC), this project set out to investigate how the shelter can increase the number of animals who are adopted and the speed at which they are rehomed.

This report outlines the development team's approach to project delivery, provides insight into our analysis design and implementation strategy, and identifies some high-level findings. It starts with an overview of the problem, followed by our project objectives. It goes on to outline the project specifications and design, then our implementation process and challenges. We provide an overview of our data sources before concluding with high-level findings and recommendations for next steps.

2.0 BACKGROUND

Across the United States, approximately 6.3 million companion animals are taken into care at animal shelters each year. While roughly 4.1 million shelter animals are adopted annually, almost one million (920,000) are euthanised.[1] More animals are killed in Texas each year than in any other state; in fact, its euthanisations per annum represent nearly 20% of the total kills for the entire country.[2]

The Austin Animal Center (AAC), a publicly-funded sanctuary serving the City of Austin and Travis County, is one of Texas's 371 shelters. It is one of only a third (36%) of the state's shelters that has achieved a "no-kill" designation.[2] That is, "euthanasia is reserved for animals that are irremediably suffering or cannot be safely placed into the community and no animals are killed for space only."[3] While the shelter does not euthanize to free up capacity, shelter space is limited and its services are in high demand. In fact, as of August 2023, its intake capacity for cats and medium-to-large sized dogs was classed at "emergencies only" and its intake capacity for small dogs was "limited."

3.0 PROJECT AIMS AND OBJECTIVES

3.1 Research Questions

In order for the AAC to ensure that its services reach as many animals in need as possible in its catchment area, while also sustaining its no-kill commitment, this project set out to conduct data analysis to answer the question: *How can the AAC increase the number of animals who are adopted and the speed at which they are rehomed?*

To address this leading question, our analysis was guided by the following sub-questions:

- *What are the characteristics of animals who are adopted? How are their characteristics different to animals who have other outcomes?*
- *For animals who are adopted, how long do they stay in the shelter before being rehomed?*
- *Are there particular times during the year that most adoptions happen?*
- *How have adoption trends changed over time?*

3.2 Tiered Objectives

We set out to achieve a set of primary objectives, which would constitute a minimal viable product, with the intention of tackling secondary and tertiary objectives if time allowed. These are shown in Table 1.

---

[1] https://www.aspca.org/helping-people-pets/shelter-intake-and-surrender/pet-statistics
[2] https://www.veterinarians.org/animal-shelter-statistics/
[3] https://www.austintexas.gov/department/aac/faq

Table 1: Primary, secondary, and tertiary objectives

Primary objectives (MVP)	Secondary objectives	Tertiary objectives
• Generate a clean and analysable dataset (involving sourcing, merging, and cleaning) • Conduct exploratory descriptive analysis (with visualisations) to address research questions • Conduct preliminary time series analysis (with visualisations) to address research questions	• More in depth analysis guided by findings of exploratory/preliminary analysis • Generate a dataset that can be used for predictive modelling (ie using one-hot encoding, etc) • Use machine learning to develop a model that predicts animal outcomes at the point of intake to the shelter	• Develop a SQL database in third normal form for AAC data (and for ultimate storage of data from other shelters)

4.0 SPECIFICATIONS AND DESIGN

4.1 Requirements

To be successful, our minimal viable product (MVP) needed to meet the technical and non-technical requirements listed in Table 2. These requirements are a combination of those set by our team - in interest of completing a quality analysis that is accessible to our target audience - as well as those set out by Code First Girls. It would have also been important to consult our primary stakeholders, AAC management, on these, if we had had direct access to them.

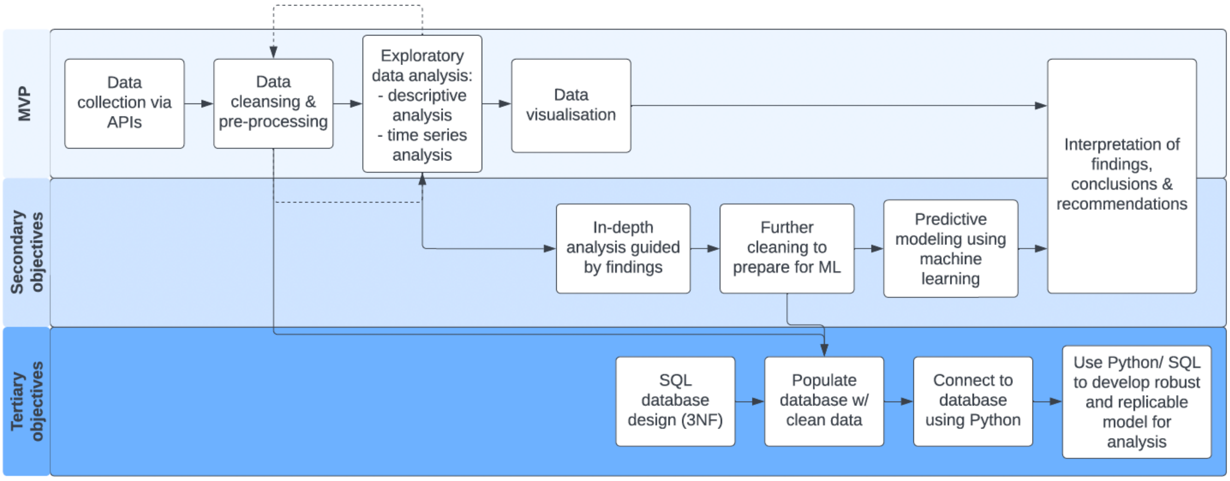
Table 2: Project Requirements

MVP Technical Requirements	MVP Non-Technical Requirements
• Written in Python in Jupyter Notebook with clearly defined sections • Uses key analytical libraries (Pandas, Matplotlib, NumPy) • Code successfully retrieves data from API and generates .csv file • Code merges two datasets without duplicates or other negative impacts to data integrity • Cleaned dataset is free from duplicated, missing, and irrelevant data and all columns have data that are analysable (ie no leading/trailing whitespaces, data in usable formats, one unique feature per column, etc) • Code runs without error and yields accurate data analysis and visualisations	• Project has clear objective and use case • ReadMe.md file explains how to run the code • Findings in the Jupyter Notebook are explained using visualisations and common language so they are accessible to our non-technical stakeholders • Slides for 3-5 min presentation to communicate findings • Project is complete by 27 Aug deadline
Nice to Have Technical Components *(for secondary/tertiary objectives)* • A dataset that is appropriately cleaned/encoded for machine learning • Use of scikit library and machine learning • SQL database	

4.2 Project design and architecture

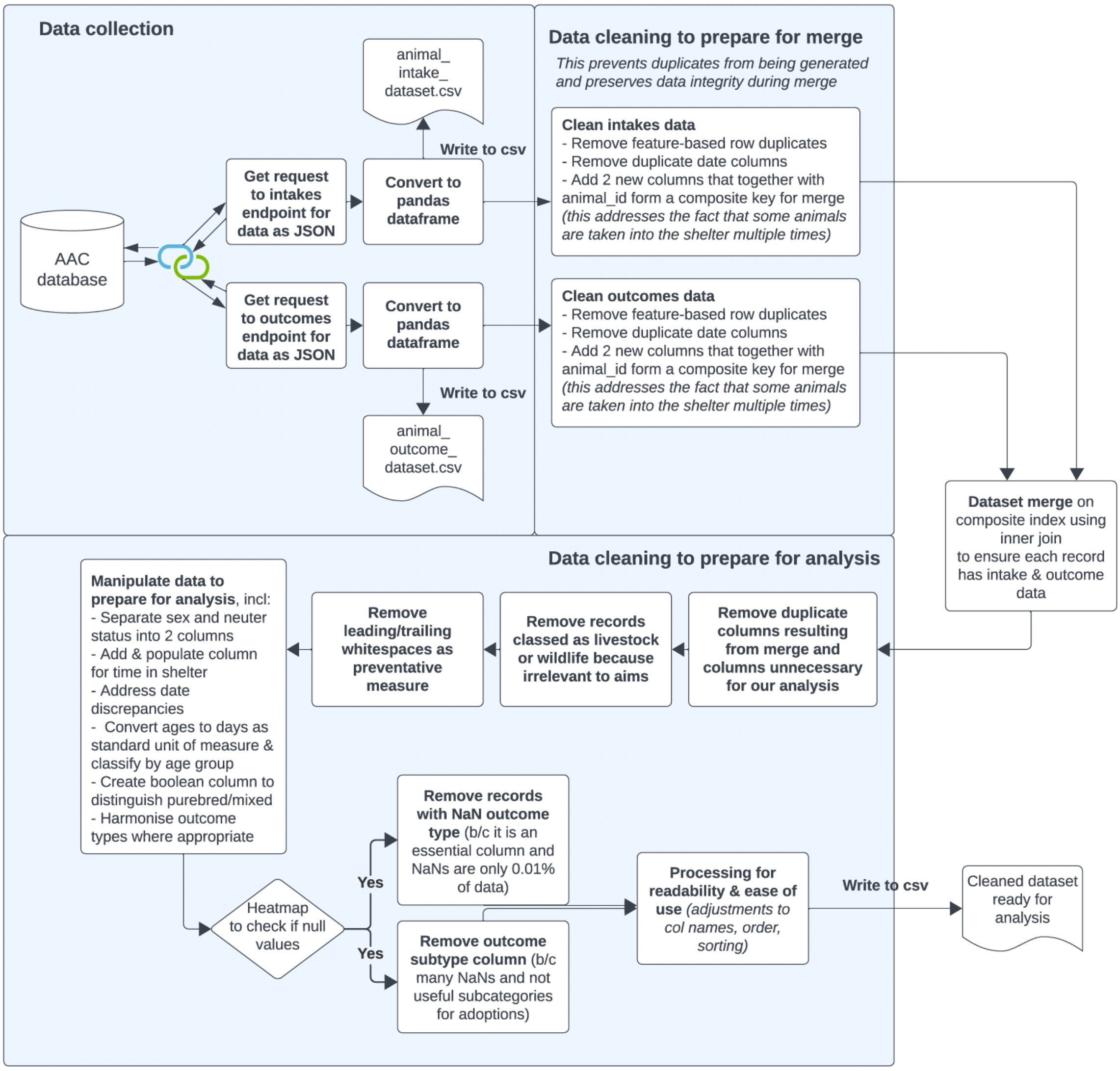
Figure 1 provides a high-level overview of our project design, distinguishing the various stages by tiered objective. While each stage is represented here by a distinct box, these stages are intended to be fluid and carried out using an iterative, agile method (see section 5.1).

Figure 1: Project Design by Tiered Objective



Our research sub-questions guided our identification of tasks to carry out in the descriptive and time series analysis. An essential phase that happened prior, and in tandem with, the analysis was data pre-processing. Figure 2 details our pipeline from data collection through data pre-processing, including high-level justifications for the steps we took.

Figure 2: Data Collection & Pre-Processing



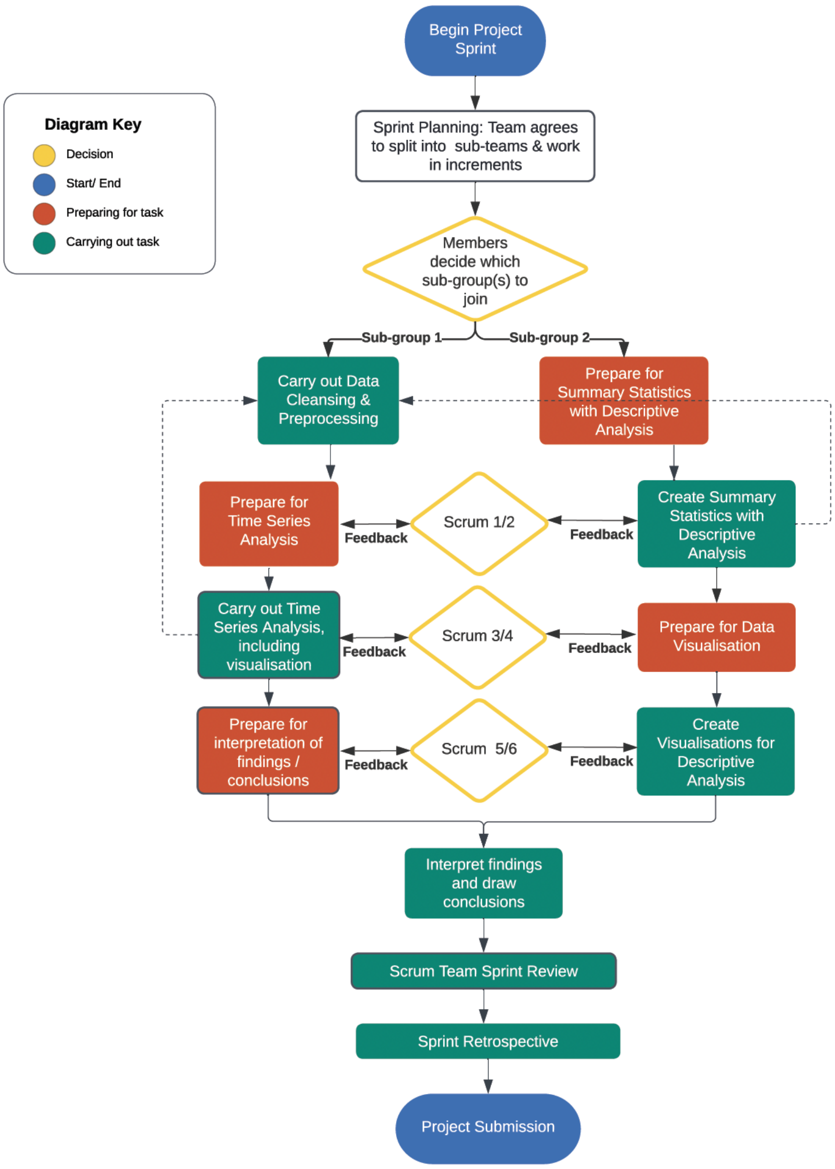
5.0 IMPLEMENTATION AND EXECUTION

5.1 Agile development approach

Our team split into sub-groups and followed a parallel Agile working method using Scrum. Per Figure 3, this approach saw one group start on one stage of the project while the other group gathered information, resources, and tools to prepare for subsequent stages.

We held frequent scrum meetings to update each other, discuss challenges, and identify where iterative work needed to happen (represented by dotted lines in Figure 1). Separately, our sub-groups met for pair programming and code reviews - especially for troubleshooting - as needed.

Figure 3: Parallel Agile Approach to Development



In addition to scrum meetings, we held other ceremonies including:

- Sprint Planning meeting to agree initial tasks and deadlines
- Sprint Review to assess what we accomplished, what remains outstanding (section 5.4), and our initial recommendations to the AAC (section 8.1)
- Sprint Retrospective to capture what went well, what could be improved, and areas of focus for a hypothetical future sprint (see section 8.2)

5.2 Team members' roles

Table 3 shows how project responsibilities were divided across our team. This was largely decided by each member's availability over the course of the project period which we identified early on via a team calendar. Members' interests and expertise (discussed during an initial SWOT analysis) were also taken into consideration.

Table 3: Division of Responsibilities Across Team

Team Member	**Responsibilities**												
	Project definition & dataset selection	Scrum master	Kanban board manager	Compilation of group homework	Data mining	Data merging & cleaning	Descriptive analysis & visualisations	Time series analysis & visualisations	Compilation of final Jupyter notebook	Compilation of Readme.md file	Compilation of Project Doc	Compilation of presentation	Contributions to sprint review & retrospective
Sumaiya Ahmed	x				x	x		x	x	x			x
Andrea Messmer	x					x					x		x
Shamim Munshi	x	x					x					x	x
Charlotte Schnoebelen	x							x					x
Yana Zykova	x		x	x			x						x

5.3 Tools and libraries

As identified in Table 4, we used a variety of tools and libraries to carry out this project. While some were essential to the technical aspects of our work, others enabled task management and team collaboration.

Table 4: Tools & Libraries Employed

Tool / Library Used	**Description of Use**
Technical tools	
Requests library	For making get request to the API for data collection
Pandas and NumPy libraries	For storing data in dataframe, then cleaning, manipulating, performing calculations and analysing it
Matplotlib and Seaborn libraries	For creating visualisations during data cleaning and data analysis
Power BI	For creating dashboards for project presentation
Task management tools	
Trello	For transparent progress tracking (see Annex B for snapshot of our board)
Collaboration and communication tools	
GitHub	For sharing code
Slack / Zoom	For communication
Google Docs, Sheets	For collaborating in real-time, especially for project documentation
Lucid	For creating flow charts for communication

5.4 Implementation: achievements, pivot points, and challenges

Our team met our MVP objectives (Table 1), satisfying the relevant project requirements (Table 2). We were able to draw some preliminary recommendations for AAC, as well as recommendations for future sprints (section 8.0), based on the findings from our preliminary analysis. Implementation highlights included great teamwork and attitude, effective communication through slack and stand-ups, and time efficiency resulting from division of roles via sub-groups.

We maintained a flexible approach throughout the project, which allowed us to nimbly make two important pivots:

- We drastically scaled back our research questions after starting our analysis, which made our aims more achievable given the time we had available.
- We re-worked how we merged our two datasets in order to ensure improved data integrity after spotting inconsistencies in our analysis. This involved creatively generating a composite key as a basis for the merge.

Our greatest challenges were time constraints and competing demands on team members' time. This constrained our iterative work, as we would have liked to refine our analysis even at this preliminary stage. To compensate, we note what further preliminary analysis we would do, if time allowed, in the accompanying Jupyter notebook. Some additional challenges are outlined alongside suggested improvements in Table 5.

Table 5: Project Implementation Challenges

Challenge	How Addressed	Suggestions for Improvement
Time constraints and competing demands on team's time	Focused on MVP objectives, reserving more advanced objectives for future sprint. Limited number of iterations we could do on our analysis.	Objectives and sprint goal could have been narrowed earlier on to make more efficient use of time
Version control issues with GitHub and Jupyter Notebook	Rectified by team members' keeping work separate, only merging at end	Explicitly agree procedures for version control at outset
Course structure meant some useful topics to project (ie exploratory analysis) were taught close to deadline, leaving us unable to fully incorporate them	Relied on online research to fill in gaps where needed; acknowledged that we did the best we could with the knowledge we had at the time	More refined and advanced analysis can be undertaken in future sprints with the knowledge we have now gained

6.0 DATA COLLECTION

6.1 Criteria for Data Selection

For our analysis, we needed data from at least one animal shelter that would fulfil the criteria defined in Table 6. We chose to work with two datasets from the AAC - one for animal intakes and one for animal outcomes.

Table 6: An assessment of available data against our defined criteria

Criterion	Available Data		
	AAC Intakes Dataset	AAC Outcomes Dataset	Meets criteria ?
For each animal, contains: - Outcome (ie adoption, euthanasia, transfer) to serve as target for any predictive modelling - Characteristics (ie animal type, breed, colour, sex, neuter status, condition, means of intake, etc) as independent variables - Dates for intake and outcome for time-based analysis	- Characteristics: animal type, breed, colour, sex at intake, age at intake, condition at intake, intake type - Intake date	- Outcomes identified - Characteristics: animal type, breed, colour, sex at outcome, age at outcome - Outcome date	Y
Sufficient number of records (2,000+)	150,000+ records	150,000+ records	Y

Allows for merging of datasets	Two separate datasets, each with an Animal_id column, which enabled merging (though more complex than we anticipated - see section 5.4)	Y
At least one data source accessed via API	Accessible via API	Y
Publicly available	Free and accessible without token	Y

6.2 Data Sources and Collection

Both AAC datasets have been updated monthly since October 2013. The intakes dataset contains information on the status of each animal as it arrives at the shelter, including how it came to the shelter, intake date, and various characteristics of the animal. The outcomes dataset contains information on the status of each animal as they leave the shelter, including outcome type, outcome date, and characteristics of the animal.

We accessed both datasets via the Austin API, which is managed by Tyler Technologies. We used the Python requests library to make a get request to each of the specified endpoints, and received the responses in JSON format. It was important for us to set upper limits of 160,000 on the requests in order to receive all 150,000+ records in the responses.

7.0 SUMMARY OF FINDINGS

In this section, we present an overview of our key findings by sub-question (see section 3.1). Visualisations evidencing these findings are in the accompanying Jupyter Notebook (file 'group5_Project.ipynb').

7.1 Characteristics of animals who are adopted (versus those with other outcomes)

- Just under half (49%) of shelter intakes are adopted, while 30% are transferred to a partner shelter; under a fifth (17%) are returned to their owner (plot 10.1.2 in Jupyter Notebook).
- Dogs and cats comprise the majority of animals at the shelter (plot 10.2.1, subplot 'Distribution of Animals'), and each have adoption rates around 49% (though differ on rates of other outcomes) (plot 10.5, subplot 'Outcome Types for Cats' and 'Outcome Types for Dogs').
- Adopted animals tend to be younger than animals with other outcomes; dogs who were 1-3 years old at the time of their outcome and cats up to one year old were adopted in greatest numbers (plot 10.6.3).
- Neutered and spayed cats and dogs are more likely to be adopted, while intact animals make up the largest groups transferred to partner organisations (plots 10.8.8. and 10.8.9).
- Most adopted animals are in good health (or "normal" condition) at the time of their intake (plots 10.8.14 and 10.8.15).
- Animals who come to the shelter as strays or by owner surrender comprise the greatest number of adoptions (plot 10.4).
- Sex, breed, and colour do not appear to influence adoption rates for cats or dogs (plots 10.3.1, 10.3.2, 10.8.4, 10.8.5, 10.8.6, 10.8.7, 10.8.10, 10.8.11), though further analysis could help verify this.

7.2 Duration of shelter stay for animals who are adopted

- On average, animals who are adopted are rehomed within 4 months. However, the older an animal, the longer they tend to stay in the shelter before adoption. The exceptions are kittens and puppies up to 12 weeks old, who stay in the shelter for longer periods, presumably because they need the shelters' care while at such a young age (plots 10.6.4, 10.7.1, 10.7.2, 10.7.3, 11.2.2).

7.3 Seasonality of adoptions

- Animal adoption rates are highest in summer, but this is largely driven by cat adoptions which peak at this time of year (perhaps because of their spring breeding season). Dogs, on the other hand, tend to have more consistent adoption rates over the course of the year. (plots 11.3.2, 11.3.2.2)

7.4 Changes to adoption trends over time

- Adoption numbers mirror intake numbers. The start of the Covid-19 pandemic (2019) saw a slight increase in adoptions, followed by a sharp decline in 2020. Thereafter, adoption rates have plateaued at a lower level than pre-pandemic. (plots 11.3.4, 11.4.2)

8.0 CONCLUSIONS AND RECOMMENDATIONS

8.1 Recommendations to AAC

This preliminary analysis forms the basis of some initial recommendations to AAC for increasing the adoption rate of intakes and minimising the time that animals spend in the shelter, including:

- A social media / awareness campaign is launched to improve cat adoption rates, particularly in spring and winter, and the adoption of older cats and dogs. The latter, who tend to have the longest stays in the AAC, could help free up much-needed capacity.
- Enhance and expedite adoption procedures for the most adoptable groups, which from this preliminary analysis could be dogs 1-3 years old and kittens 0-3 months old, though further analysis could verify this.
- A fundraising effort is made to support the neutering of more animals to improve adoption rates.

8.2 Expanding on this work: potential focus areas for future sprints

To expand on this work and provide further insight to AAC, next steps may include:

- Using predictive modelling to predict outcomes for an animal upon intake. Based on our initial analysis, we suggest paying particular attention to age, neuter status, and intake type and condition as independent variables that may influence adoption.
- Investigating repeat intakes: In our pre-processing stage, we learned that thousands of animals were taken into the shelter multiple times over the last 10 years, placing strain on AAC's capacity. Analysing trends/patterns related to repeat intakes may be useful to inform ways to reduce them.
- Investigating transfers: There are also significant numbers of animals who are transferred from AAC to a partner organisation. For a more complete analysis of outcomes, sourcing data from partners on ultimate outcomes for these animals could be prioritised.

This analysis has led to some interesting insights that may help inform AAC's operations and help ensure positive outcomes for more animals. Internet research suggests that other US shelters - including those that have not attained a "no-kill" designation - collect similar data on intakes and outcomes, even though it may not be publicly available. Future work could focus on adapting and refining the code that we have developed so that it can be used to analyse data from any shelter, thereby empowering their management with the information that they need to secure more happy outcomes for more animals.