Project Report: Montgomery County Office of Animal Services Data Analytics Dashboard

Merveille Lore Kuendzong
DATA 205
CRN: 22017

Introduction

Context of the Project:

The Montgomery County Office of Animal Services (OAS) handles the impounding and fostering of animals throughout the county. With a large amount of data being generated, there is a need to develop an efficient and actionable analytics dashboard to monitor animal impounds, foster placements, and their outcomes. This project leverages historical data from OAS to create a comprehensive dashboard that provides insights into the operational trends and key metrics related to animal services.

Data:

The dataset for this project includes four main datasets:

The ***OAS-Kennel*** *dataset* contains detailed records of animals housed in kennels, including information such as animal ID, color, intake date, species, breed, age, intake conditions, outcomes (e.g., adopted, transferred), and owner details.

The ***OAS-Foster*** *dataset* tracks animals placed in foster care, capturing data on foster start date, duration, reasons for fostering, and the condition of the animal.

The ***OAS-Impound*** *dataset* includes information on animals impounded by the service, such as impound location, reason for impoundment, and animal condition.

The ***OAS-Treatment*** *dataset* contains details of procedures and of treatments given to animals.

Goal:

The primary goal of this project is to develop an interactive dashboard that enables OAS staff to visualize and monitor key metrics related to animal care and services. These metrics include:

- Most Common Animal Types
- Intake and Adoption Trends
- Foster Care Placements and Outcomes
- Monthly populations
- Intake and Outcome Distributions
- Impound pathways
- Live release rates
- Treatments

This dashboard is crucial as it empowers OAS staff to make data-driven decisions, optimize operations, and allocate resources more effectively. Additionally, the project aims to tell the public a story of animal care by visualizing trends in animal intake, fostering, and adoption at OAS. By providing valuable insights to both the public and OAS staff, the dashboard will help inform the community about the processes and outcomes at OAS, reducing unnecessary calls for

animal-related inquiries. Furthermore, the analysis may prompt OAS to adjust certain procedures or improve its operational strategies based on the insights derived from the data.

Tools:

For this project, I used the following tools and methods:

- **R programming**: For data cleaning, processing, and analysis.

  Packages:

  - **lubridate**: For handling date-time data, such as calculating durations and extracting date components.
  - **tidyverse**: A collection of R packages (including ggplot2 and dplyr) that facilitated data manipulation, visualization, and analysis.
  - **plotly**: For creating interactive, web-based visualizations, allowing users to explore the data more dynamically.
  - **vcd**: For visualizing and analyzing categorical data.
  - **DescTools**: For specific statistical analysis.
- **Tableau**: To create interactive visualizations and the final dashboard.

Summary of Data Pre-processing

I focused primarily on the kennel dataset for statistical analysis, as it serves as the core dataset containing information on foster and impound activities. While a treatment dataset was later provided, time constraints prevented statistical analysis on that dataset.

Key steps in data pre-processing included:

- ***Filtering of data:***
  - Initially filtered out records with outcome types *'lost exp'* and *'find exp'* and kennel types *'lost'* and *'found'*, as these represent data specific to animals that are not in the facility.
  - Removed records with intake types labeled as 'disposal' and intake conditions labeled as 'dead,' as OAS aims to focus on animals that arrived alive.
  - Removed records with outcome types labeled as 'disposal', as 'disposal' is typically used for administrative purposes rather than tracking live animal outcomes.
  - Excluded entries with animal type *'wildlife'* before calculating and displaying adoption rates, as wildlife is not eligible for adoption.
- ***Creating New Variables:***
  - adopted: A binary variable to indicate whether an animal was adopted.
  - time_in_shelter: Calculated as the duration between intake and outcome dates.
  - Intake_season: Derived from the intake date to classify entries into seasons (e.g., fall, spring).

- ***Handling Missing Data:***
  After calculating time_in_shelter, some records with missing shelter times were identified. These observations were filtered out to ensure consistency and reliability in subsequent analyses.
- ***Date Formatting:***
  Converted intake and outcome dates from character format to date format using the lubridate package, ensuring accurate and consistent handling of time-related data across the dataset.

Basic Descriptive Statistics

Number of Records: *The* OAS-Kennel dataset contains **56,881 records** (data from January 1st, 2018, to September 11th, 2024, before cleaning and filtering of the data.

Key Variables after initial filtering:

- **Animal Type:** Bird (2,473), Cat (17,199), Dog (13,464), Livestock (138), Other (5,941), Wildlife (3,127).
- **Intake Types:** Boarding (792), Confiscate (1,561), Euthanasia Requested (1,501), Foster (6,150), Owner Surrender (9005), Return (718), Stray (15,009), Transfer (128), Veterinary (128), Wildlife (7122).
- **Outcome Types:** Adoption (14,075), Died (997), Euthanasia (7,568), Foster (6,268), Missing (8), Null (191), Relocate (165), Return to Owner (6,093), Transfer (6,977).
- **Adopted Status:** Not Adopted (28,267), Adopted (14,075).
- **Other Key Variables:** Intake Date, Outcome Date, intake season (A calculated variable with values *Fall*, *Winter*, *Spring*, and *Summer*, based on the intake date), time in the shelter (A calculated variable representing the duration (in days) an animal stayed in the shelter).

I initially considered working with the **Breed** and **Color** variables but chose not to because **Breed** has more than 1,300 unique values, and **Color** has over 500 unique values, making these variables challenging to analyze effectively within the project's scope. Similarly, the variables related to **pet age** exhibit inconsistent values, rendering them unreliable for analysis.

Description of final data product:

The final product of this project is a dashboard with eight tabs, each dedicated to a specific aspect of the data.

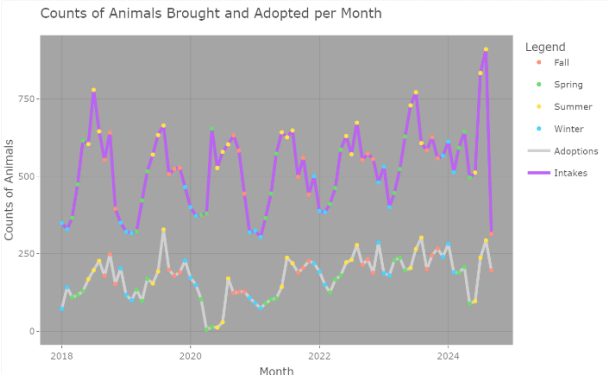
- The ***Home*** tab serves as the landing page and provides navigation buttons to access the other sections.
- The ***Shelter Pathways*** tab illustrates the various pathways animals take through the shelter system.
- The ***Impound Pathway*** tab shows the flow from animal intake through impound and onward to potential outcomes.

- The ***Animal Intake*** tab presents the different intake types of animals and explains the intake process.
- The ***Community Partners*** tab details partnerships involving foster homes and rescue organizations for each year.
- The ***Live Release Rates*** tab displays metrics related to live and non-alive outcomes.
- The ***Animal Outcomes*** tab provides insights into all possible outcomes for animals, including adoption, euthanasia, and transfer.
- The ***Animal Wellness*** tab focuses on metrics related to the health, care, and well-being of animals in the shelter system.
- The ***Glossary of Terms*** button leads to a dedicated glossary page, explaining key terms used throughout the dashboard.

The dashboard is accessible using the following Link to the dashboard on Tableau Public.

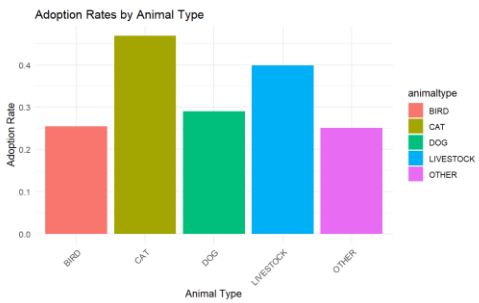
Some analysis:

1- Counts of Animals Brought and Adopted per Month



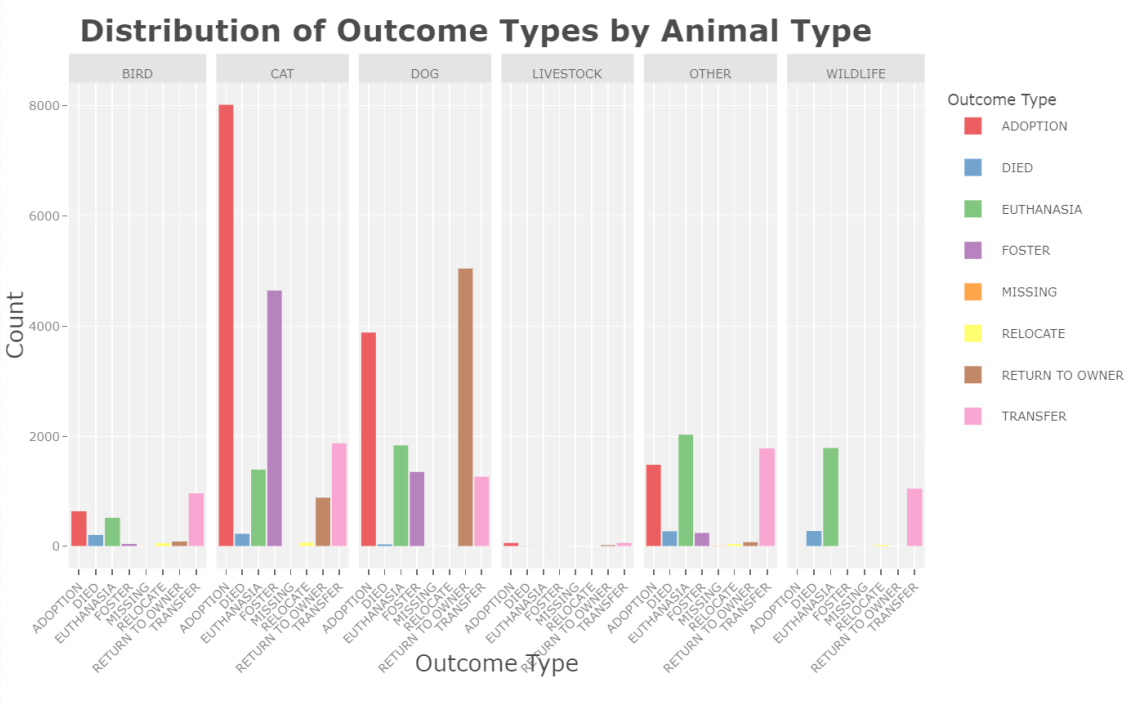
Intake numbers tend to be higher in summer and lower in winter. Adoptions generally decrease in winter, start to increase in spring and are higher around summer. Notably, there was a significant drop in adoptions around April 2020, likely due to the COVID-19 lockdowns.

2- Adoption rates by animal types:



Cats have the highest adoption rate, followed by livestock, dogs, birds, and other animals.

3- Outcome types by animal type



This graph confirms that cats have a very high number of adoptions compared to other outcomes, although a significant portion also remains unadopted. The most common outcome for dogs is returning to their owner, surpassing the number adopted, which makes adoption rates lower for dogs. A large proportion of birds were transferred to rescue organizations. Livestock has relatively few outcomes overall due to its low numbers. Among "other" animals, euthanasia is the most frequent outcome.

Wildlife is not adopted; most are either euthanized or die in care. Some are transferred to rescue organizations.

4- Linear regression model to predict time spent in the shelter based on animal type, intake type, and season

```
Call:
lm(formula = time_in_shelter ~ animaltype + intype + intake_season,
    data = fdata)

Residuals:
   Min     1Q Median     3Q    Max
-52.09 -13.00  -3.67   0.98 700.65

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        27.9914     1.2073  23.185  < 2e-16 ***
animaltypeCAT       3.0679     0.6836   4.488 7.21e-06 ***
animaltypeDOG      -0.1335     0.6888  -0.194 0.846316
animaltypeLIVESTOCK 16.8331     2.4749   6.802 1.05e-11 ***
animaltypeOTHER     5.0428     0.6709   7.517 5.73e-14 ***
animaltypeWILDLIFE  3.3994     0.8152   4.170 3.05e-05 ***
intypeCONFISCATE    7.4939     1.2315   6.085 1.17e-09 ***
intypeEUTH REQ    -24.6963     1.2352 -19.994  < 2e-16 ***
intypeFOSTER      -23.3931     1.0714 -21.833  < 2e-16 ***
intypeOWNER SUR    -3.2125     1.0466  -3.069 0.002146 **
intypeRETURN      -10.0183     1.4452  -6.932 4.21e-12 ***
intypeSTRAY       -13.0387     1.0281 -12.682  < 2e-16 ***
intypeTRANSFER     -7.1033     2.0906  -3.398 0.000680 ***
intypeVETERINARY  -26.9522     2.6729 -10.084  < 2e-16 ***
intypeWILDLIFE    -29.7486     1.1669 -25.494  < 2e-16 ***
intake_seasonSpring -1.4256     0.3958  -3.601 0.000317 ***
intake_seasonSummer -2.4131     0.3698  -6.526 6.85e-11 ***
intake_seasonWinter -0.2245     0.4181  -0.537 0.591282
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 27.92 on 42113 degrees of freedom
Multiple R-squared:  0.119,     Adjusted R-squared:  0.1186
F-statistic: 334.5 on 17 and 42113 DF,  p-value: < 2.2e-16
```

The coefficients show that livestock tend to stay longer in the shelter. Animals with intake types such as "Euth Req," "Foster," or "Veterinary" tend to spend less time in the shelter, while confiscated animals tend to stay longer. Animals admitted in Spring and Summer also spend less time in the shelter compared to other seasons. One possible explanation for the shorter shelter time in Summer and spring could be the increased adoption rates during these seasons, potentially due to more active adoption events or favorable weather encouraging outdoor activities and animal interaction.

The model explains only about 11.9% of the variance in time_in_shelter (R-squared = 0.119), but it is statistically significant (p-value< 2.2e-16). The residual standard error is 27.92, indicating moderate prediction variability.

5- **ANOVA Test Results: Differences in Shelter Time by Animal Type**

The ANOVA test (p-value: < 2.2e-16) shows that animal types have significant differences in shelter time, supporting the alternative hypothesis that at least one animal type's shelter time differs from the others. Possible explanation for this could be the differing adoption rates and care requirements for each animal type.

6- **T-test to compare Time Spent in the Shelter Between Adopted and Non-Adopted Animals**

The Two Sample t-test results show a t-statistic of -45.585, with a p-value < 2.2e-16. The negative t-statistic indicates that the mean time spent in shelter for non-adopted animals (mean = 9.42 days) is significantly less than that for adopted animals (mean = 24.26 days). The 95% confidence interval for the difference in means is between -15.48 and -14.2, further confirming this difference. One possible explanation for this is that non-adopted animals may include those that were returned to their owners or transferred to other facilities, which could explain their shorter shelter durations. In contrast, adopted animals tend to stay longer as they undergo more thorough adoption processes.

7- **Chi-squared and Cramér's V tests to determine the association between categorical variables**

There is strong statistical evidence of associations between **intake type and outcome type**, **animal type and outcome type**, and **animal type and intake type**. However, the strength of these relationships is **moderate**, as indicated by measures like Cramér's V coefficients. This suggests that while these variables are interrelated, other factors or random variations might also contribute to the observed patterns in the data.

Please refer to the detailed analysis here for further insights.

References:

How to Calculate Cramér's V in R." *GeeksforGeeks*, GeeksforGeeks, n.d., https://www.geeksforgeeks.org/how-to-calculate-cramers-v-in-r/.

Introduction to Tableau for Beginners." *YouTube*, uploaded by Data School, 15 May 2017, https://www.youtube.com/watch?v=oAIubTqg-Kw&list=PLO9LeSU_vHCU_DHaLzEvsLxFdmB3Qcao_.

R Tutorial." *GeeksforGeeks*, GeeksforGeeks, n.d., https://www.geeksforgeeks.org/r-tutorial/?ref=outind.

Tableau. "Getting Started with Tableau: Drag-and-Drop Interface." *Tableau Help*, Tableau Software, n.d., https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-drag.htm.

Acknowledgements:

I would like to express my sincere gratitude to my mentors, Mrs. **Maria Anselmo**, Mrs. **Victoria Liu**, Mrs. **Kathy Luh**, Mrs. **Caroline Hairtfield**, and Mrs. **Victoria Lewis**, for their time, availability, responsiveness, understanding, and constructive feedback. Their guidance and support have been instrumental to my understanding of the animal care field and my success in the project.

I would also like to express my deepest gratitude to my professor**, Dr. Jane Valentine**, for her guidance and support throughout this semester. Her insights, feedback, and encouragement have been invaluable to my learning and development, and I truly appreciate her dedication to helping me succeed.

I am also profoundly thankful to Professor **Rachel Saidi** for assisting me in finding my internship and offering valuable guidance with learning Tableau.