Factors Influencing the Duration of Stay for Animals in Shelters

Group17

University of Glasgow

CONTENTS

1 Introduction

2 Method And Model

3 Conclusion

University of Glasgow

1.1 Problem statement

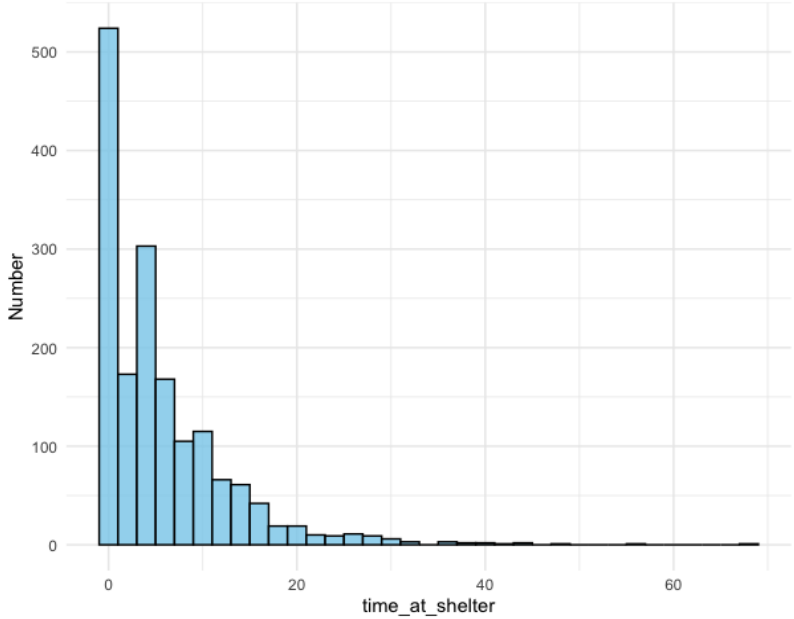
- **Problem Statement:**
- Which factors influence the number of days an animal spends in the shelter before its final outcome is decided?

1.2 Data Overview

Key variables:

- Animal_type – The type of animal admitted to the shelter
- Month – Month the animal was admitted, recorded numerically with January=1
- Year. – Year the animal was admitted to the shelter.
- Intake_type – Reason for the animal being admitted to the shelter
- Outcome_type – Final outcome for the admitted animal
- Chip_Status – Did the animal have a microchip with owner information?
- Time_at_Shelter – Days spent at the shelter between being admitted and the final outcome.

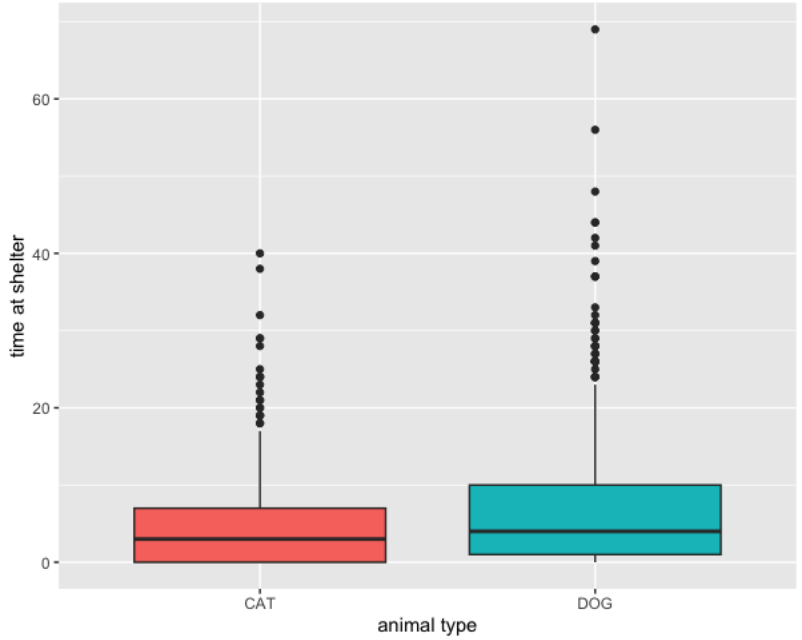
1. **Histogram** of the number of days an animal spends in the shelter



The chart shows a right-skewed distribution, with most animals staying 0-10 days and few staying beyond 30 days.

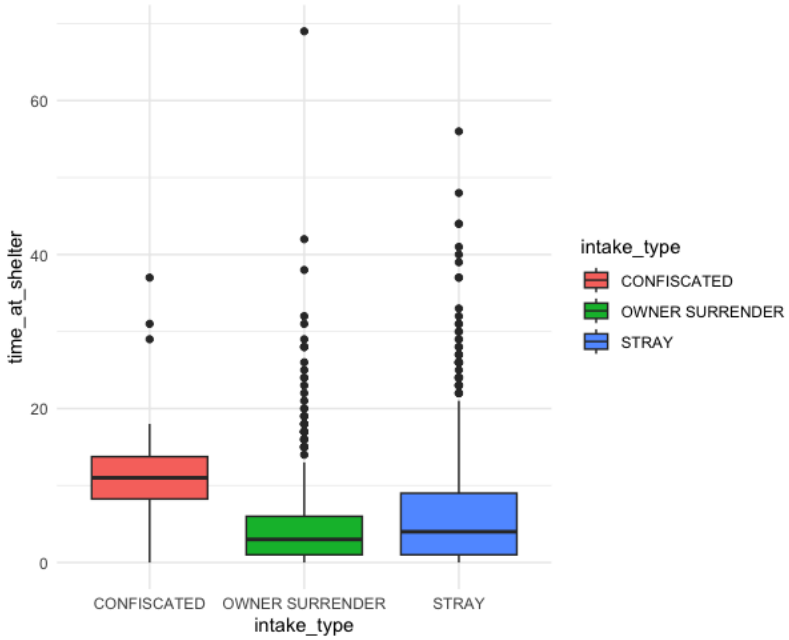
1.2 Data Overview

2. **Boxplot** of time_at_shelter and animal_type



Cats have a lower median stay than dogs, but dogs show a wider distribution with longer stays.

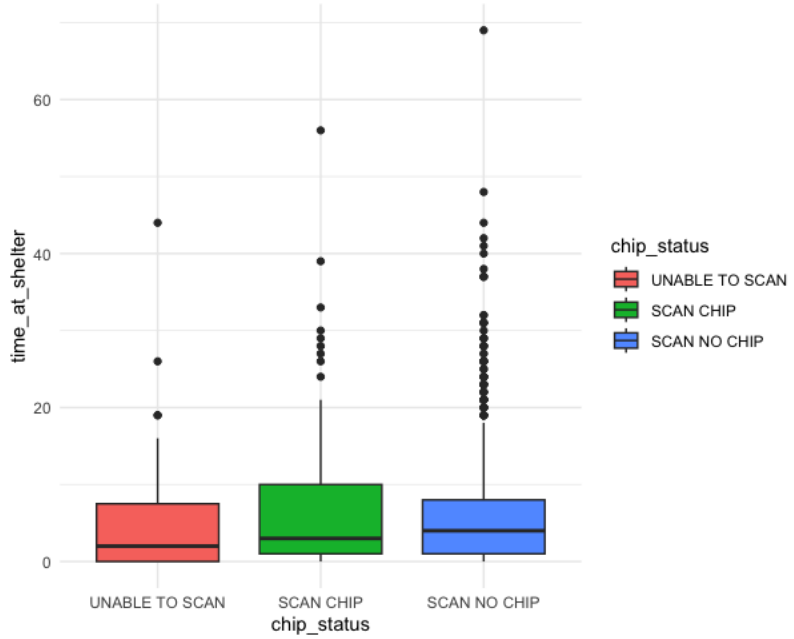
3. **Boxplot** of time_at_shelter and intake_type



The median of confiscated animals was significantly higher than that of the other two groups, indicating that confiscated animals stayed concentrated and longer.

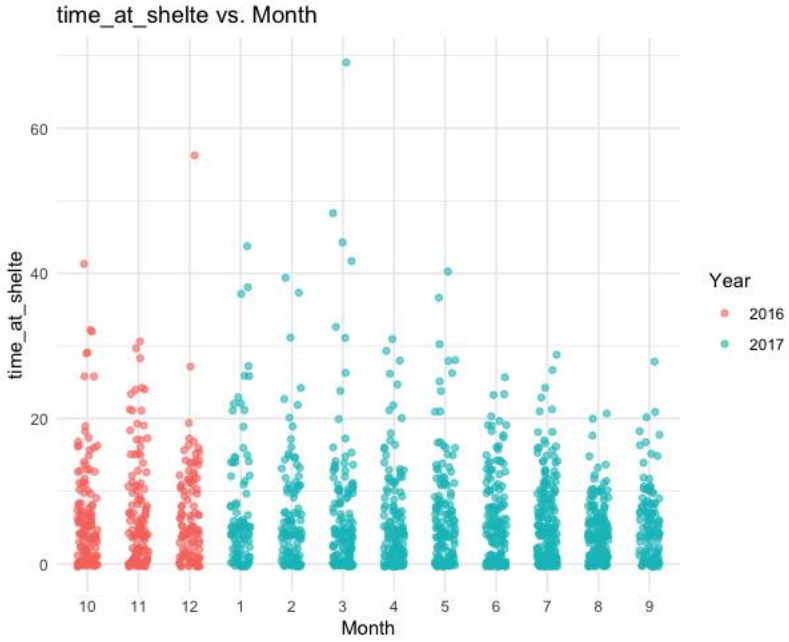
1.2 Data Overview

4. Boxplot of time_at_shelter and chip_status



The medians were similar across chip statuses, suggesting comparable shelter stays. However, the IQR was wider for scanned animals, indicating greater data dispersion.

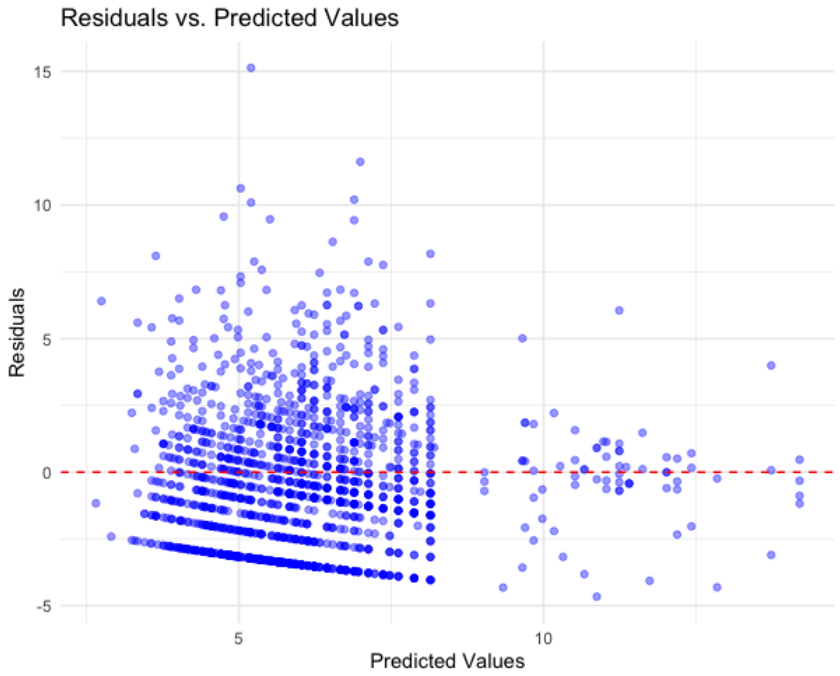
5. Scatter plot of time_at_shelter and intake_type



The dataset spans from October 2016 to September 2017 with complete monthly data. Animal shelter stays were mostly short (under 20 days) and showed no significant seasonal variation.

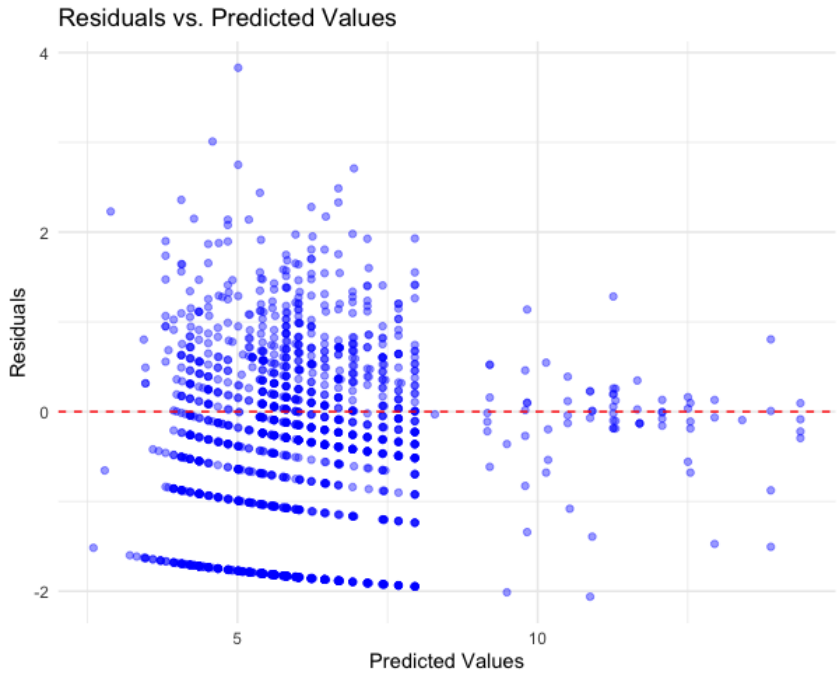
2. Residual Plot

Model1 :
Possion



Residual variability is higher at lower predicted values, indicating excessive dispersion, consistent with the previously calculated dispersion parameter.

Model2 :
nb_model_2

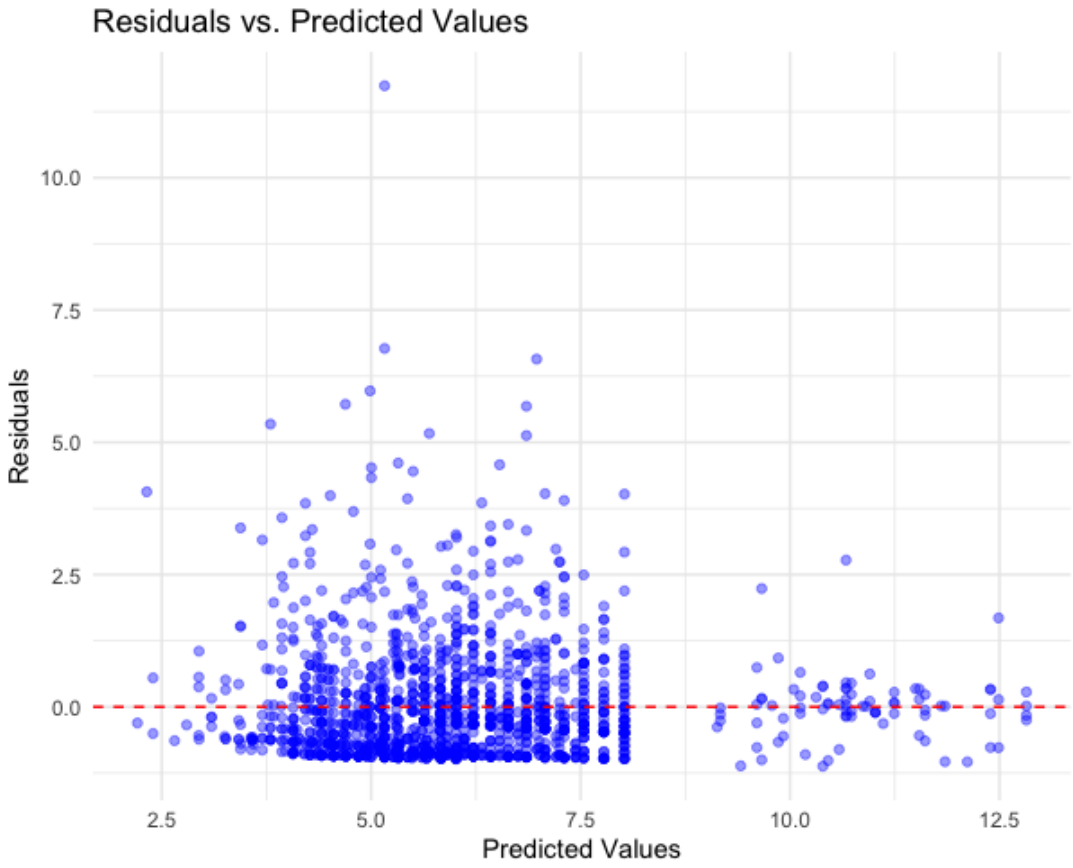


The variance of the residuals is generally in the range of -2 to +2. The negative binomial regression model basically solves the overdispersion problem.

2. Residual Plot

Model3 :
Zero-flated_model

The residual plot highlights useful insights, confirming the presence of dispersion and validating the need for a more flexible model.



University of Glasgow

3.1 Key Findings Summary

Distribution

This report examines animal shelter stays, showing a right-skewed distribution with most animals staying under 10 days.

Predictors

Using Poisson and Negative Binomial models with predictors like intake_type, animal_type, chip_status, and month, we found that owner-surrendered or stray animals stay shorter, dogs stay slightly longer, and chip status and intake month have weaker effects.

Modeling

Given many zero-day stays, a Zero-Inflated Negative Binomial model provided the best fit, distinguishing structural zeros from regular counts for deeper insight.

3.2 Limitations & Future Work

Problem statment

- The dataset contains a large number of zero values.
- Multiple explanatory variables are present.
- GLM does not fit well.

Proposed solution

- Explore ANN as an alternative predictive model.
- Investigate methods for transforming non-numerical variables into numerical ones.
- Consider whether the frequency of observer occurrences can serve as a conversion criterion.

University of Glasgow
Thank you!
WORLD CHANGING GLASGOW
A WORLD TOP 100 UNIVERSITY