

A Study on Identifying Rehoming Patterns in Different Dog Breeds

ABSTRACT

Objective

To identify the rehoming trends for three different breeds of dogs – Rottweiler, Shih Tzu and West Highland White Terrier and deduce whether the average time taken for each breed to find a new home deviates from the hypothesized value of 27 weeks (for all breeds).

Data Design

A sample of 159 dogs in total were collected for the three breeds of dogs, which comprised of 97 Rottweilers, 25 Shih Tzus and 26 West Highland White Terriers. These dogs were sent to animal shelters mostly due to them being stray or being abandoned. The dataset consists of 7 variables:

- Rehomed – Number of weeks from the dog's arrival at shelter till being adopted.
- Visited - Number of weeks from the dog's arrival at the shelter until the first visit from a potential new owner.
- Health – A Measure of the dog's physical health upon arriving at the shelter, on a scale of 0 (worst possible health) to 100 (perfect health).
- Breed - Breed of the dog.
- Age - Categorised as either "puppy" or "fully grown".
- Reason – States the reason the dog was taken to the shelter(stray, neglected, unwanted health condition)
- Returned - States whether the dog has previously been rehomed but was returned to the shelter for any reason.

Main Outcomes and Overview of Results

The primary objective is to find the average rehoming times for each breed. The secondary objective involves trying to find the breed with shorter rehoming time by performing comparison tests to identify confidence intervals. Additional objectives aim to deduce characteristics that could affect the rehoming patterns.

The average rehoming time was found to be 11 weeks for Rottweilers [95% CI : 9.2,11.7], 19.5 weeks for Shih Tzus [95% CI : 15.7,23.2] and 20 weeks for White Terriers [95% CI : 15.7,24.7]. It was also observed that independent samples of a pair of breeds did not exhibit similar average rehoming times.

Conclusion

It was observed that none of the breeds had an average rehoming time of 27 weeks, with the number of weeks being shorter than the hypothesized value. Moreover, it was noted that Rottweilers were adopted more frequently than the other breeds, indicating a possible relation between the breed and rehoming times.

A Study on Identifying Rehoming Patterns in Different Dog Breeds

Introduction

This report examines the rehoming patterns of three different dog breeds—Rottweiler, Shih Tzu, and West Highland White Terrier. The primary objective is to investigate whether the average time taken for a dog to be rehomed or adopted is 27 weeks for all breeds of dogs, or whether this trend varies depending on the breed of dog. Moreover, the report also tries to identify any anomalies in the data, explore relationships between the variables and to derive meaningful insights that would help paint a better picture on the numerous factors affecting the rehoming times of dogs. These insights are drawn by making use of the information from the variables in the dataset— First visit, Rehoming status, health condition, breed, age, reason for rehoming, and whether the dog was returned to the shelter.

Results

Data Analysis and Exploration

The initial phase of analysis involved cleaning and summarization of the different numerical variables. Of the 154 total observations, 9.74% of the data were found to contain null or irrelevant values in the variables 'Breed' and 'Rehomed', thereby removing them to ensure data clarity and precision. Adoption of Rottweilers was more frequent, suggesting a preference for this breed. [Figure 1](#) represents this as a boxplot indicating the reduced rehoming times for Rottweilers' when compared to Shih Tzus' and White Terriers'.

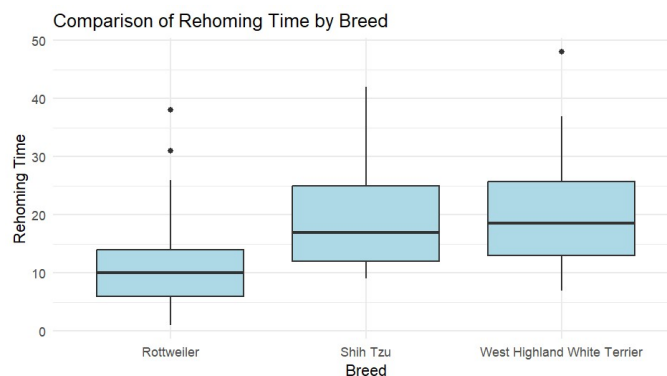


Figure 1: A boxplot of Rehoming Time for each of the three dog breeds (Rottweiler, Shih Tzu, and West Highland White Terrier)

Summary statistics ([Table 1](#)) was obtained to determine the variability and spread of the data. Rottweilers were observed to have a relatively shorter time for their first visit and for being rehomed, while also having a higher health score. The high variance in the health scores could stem from the fact that about 89.9% of the dogs were neglected or stray dogs from the streets, with little to no resources for food or shelter.

In terms of age distribution, Rottweilers and West Highland White Terriers mostly had Fully Grown Dogs (96.7% and 90.1% respectively), while Shih Tzus' showed a more balanced distribution, with about 58.3% of them being Fully Grown Dogs and the rest were Puppies.

Breed	Time till First Visit (Weeks)			Time until Rehomed (Weeks)			Health Scores		
	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
Rottweiler	-1.00	26.00	8.07	1.00	38.00	10.51	21.00	86.00	66.27
Shih Tzu	5.00	42.00	16.67	9.00	42.00	19.50	9.00	67.00	50.58
West Highland White Terrier	7.00	37.00	17.55	7.00	48.00	20.23	1.00	78.00	44.18

Table 1: The Minimum, Maximum and Mean of variables (Visited, Rehomed and Health) for the three dog breeds (Rottweiler, Shih Tzu, and West Highland White Terrier)

Data Modelling and Estimation

The distributions of Rehoming times for each breed were calculated by making use of histogram plots to show the distribution curve. It is evident from [Figure 2 \(a\)](#) that the none of the density curves showed behaviors of a normal distribution, suggesting that these could be multi-modal distributions. Upon applying a log transform to the data, initial observations hinted at potential normality for Rottweilers and White Terriers, while the Shih Tzu data appeared to follow either a normal or an exponential distribution, which can be seen clearly in [Figure 2 \(b\)](#). The plots appear to be more symmetrical, therefore, the model is designed based on the assumption that the data could be either normally or exponentially distributed.

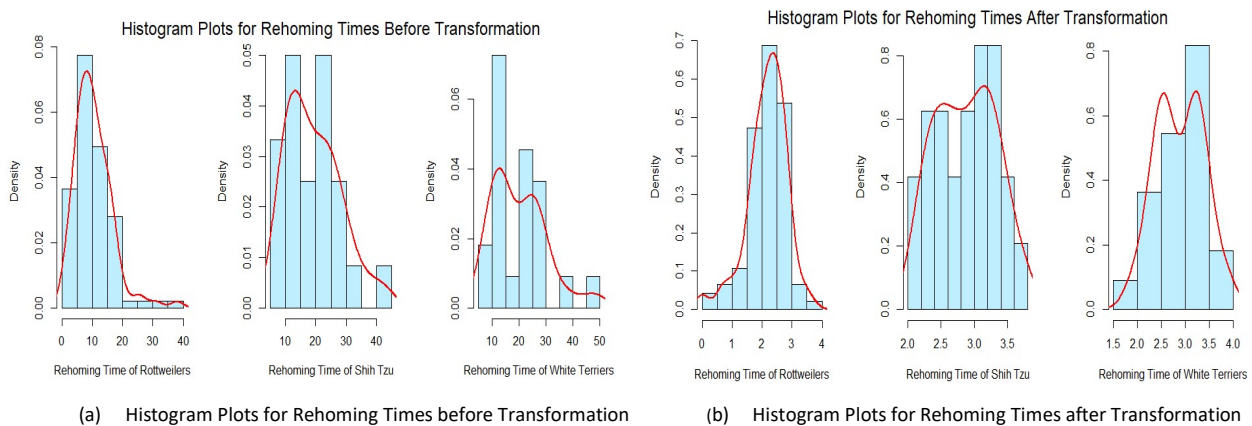


Figure 2: Histograms of Rehoming Times of three dog breeds (Rottweiler, Shih Tzu, and West Highland White Terrier) showing a comparison of the density curve lines (in red) for model estimation of each breed.

The parameters of each of our assumed models are estimated by making use of Maximum Likelihood Estimator approach, and the estimates have been tabulated in [Table 2](#). These estimates are used to identify a suitable model for modelling the rehoming time for each breed.

Estimates (for Distributions)		Breeds		
		Rottweiler	Shih Tzu	West Highland White Terrier
Normal	μ	10.505	19.500	20.227
	σ^2	38.535	77.913	102.279
Exponential (λ)		0.095	0.051	0.049

Table 2: Estimates of Parameters for the Normal and Exponential Distribution Models for the three dog breeds three dog breeds (Rottweiler, Shih Tzu, and West Highland White Terrier) by using Maximum Likelihood Estimator.

The assumptions made on the distribution model are verified by using Kolmogorov-Smirnov test and Q-Q Plots. It is clear from [Figure 3](#) that none of the Q-Q plots for the exponential distribution follow the reference line, thereby negating the assumption of the model following an exponential distribution. On the other hand, the Q-Q plots for normal distribution almost closely follow the reference lines. However, the Shih Tzu's shows some deviation from normality. Despite this, the overall pattern still suggests a tendency towards normality more than exponentiality.

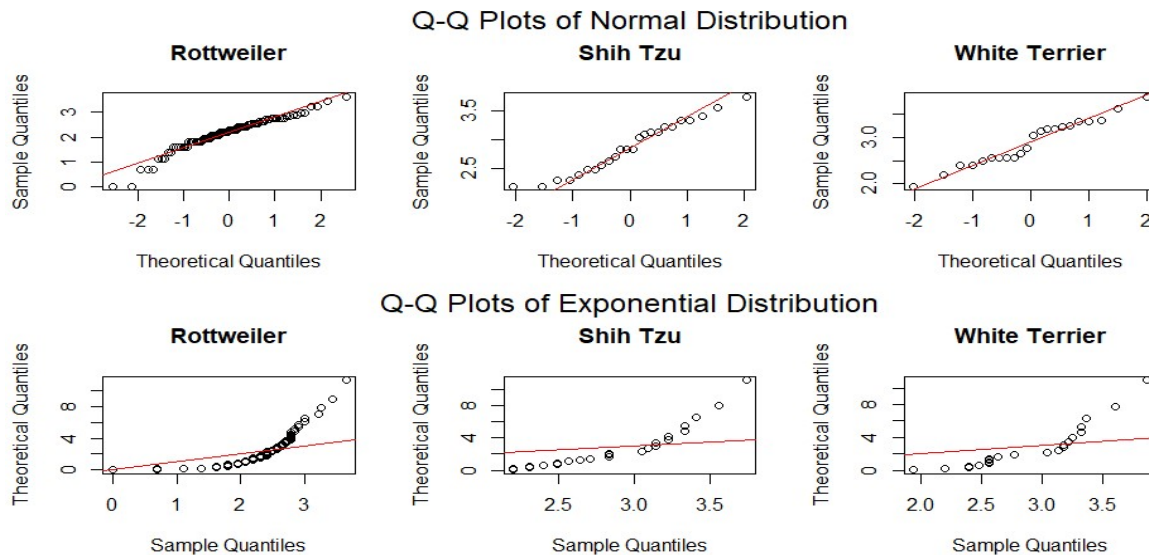


Figure 3: QQ Plots of sample and theoretical rehoming times of each dog breeds (Rottweiler, Shih Tzu, and West Highland White Terrier), the first row of plots displaying QQ plots of Normal Distribution while the second row of plots display those of Exponential Distribution.

A statistical analysis performed by using the **Kolmogorov-Smirnov test** gives us a clearer understanding of the type of distribution best suited for the rehoming time data model. The p-values for both distributions are very low (close to $<2.2e-16$), thereby strongly rejecting the hypotheses that the distribution follows normality or exponentiality. Despite the rejection of normality by the Kolmogorov-Smirnov test, the visual alignment of the data points with the theoretical line in the Q-Q plots for the normal distribution, particularly for Rottweilers and West Highland White Terriers, suggests that the data is more closely approximated by a normal distribution than an exponential one.

Confidence Interval Estimation

Moving forward under the assumption that the rehoming times for all breeds follows a normal distribution, the calculated confidence intervals give insights on the assumption of 27 weeks being the average time taken for the dogs to be rehomed. For Rottweilers, the 95% confidence interval (9.2 to 11.7 weeks) is significantly lesser than the speculated 27 weeks, thereby providing enough evidence to reject this assumption. This also suggests Rottweilers have much shorter average times in finding new homes. On the other hand, the confidence intervals of Shih Tzu (15.7 to 23.2 weeks) and West Highland White Terrier (15.7 to 24.7 weeks) is not far off from the hypothesized average rehoming time, indicating that there is insufficient evidence to reject this assumption.

The confidence intervals were calculated using T-tests Shih Tzus' and West Highland White Terriers, despite having the normality assumption and variance values, since the sample size of these breeds are small, thereby not being able to give an accurate confidence interval when calculated using Z-Test. The larger sample size of Rottweilers was a reason Z – score was used to calculate the confidence interval since it is in accordance with Central Limit Theorem.

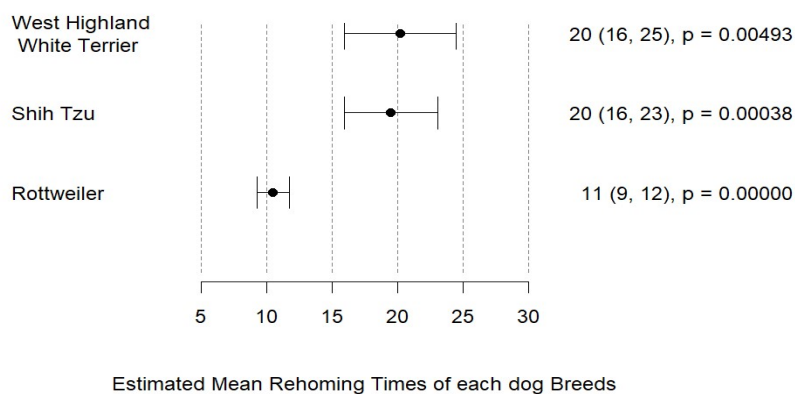


Figure 4: A forest plot showing 95%-confidence intervals for the three breeds of dogs that shows the range of mean rehoming times. The plot also displays the rehoming time estimate, the confidence intervals, and the p-values.

This is visually represented by plotting a forest plot ([Figure 4](#)) which clearly shows that the mean rehoming times of each dog breed is not equal to 27 weeks, thereby giving strong evidence to reject the Null Hypothesis (H_0 : mean rehoming time = 27). The p-values of all breeds are very low as well, further proving that the estimated mean rehoming times are statistically significant, and not due to random probability.

Moreover, a 95% confidence interval was found for each breed-pair using a two sample T-test to check two breeds had similar rehoming times. The comparison between West Highland White Terrier and Shih Tzu shows no substantial difference in mean rehoming times, as indicated by a p-value >0.05 in [Table 3](#) and a confidence interval spanning zero, while the same could not be said for the other two pairs.

These results were obtained based on the assumptions that the distribution is normal, and the variances of the two independent breeds as equal (since the standard deviation of the breed-pair lies in the range $1/4$ to 4). These findings show that there are breed specific dynamics that play a significant role in determining the rehoming times of each dog breed.

Breed Pair	Difference of Mean (Weeks)	Confidence Interval (95% Confidence)	p-values
(Shih Tzu, Rottweiler)	8.994624	(5.082219, 12.907028)	5.764e-05
(West Highland White Terrier, Rottweiler)	9.721896	(5.086241, 14.357552)	0.000219
(West Highland White Terrier, Shih Tzu)	0.727272	(-4.943660, 6.398205)	0.797000

Table 3: Results of two-sample T-tests performed over pairs of dog breeds to identify whether two breeds could have similar mean rehoming times, making use of 95% Confidence Interval and p-values.

Discussion and Conclusion

The analysis of rehoming times for the dog breeds - Rottweiler, Shih Tzu, and West Highland White Terrier show significant variations in the adoption patterns, and proved that the mean rehoming time for all breeds is not around 27 weeks. The results suggested animal shelters could allocate resources more efficiently for the adoption of Rottweilers, while urging them to use better strategies to improve the adoption rates of Shih Tzus and White Terriers.

The analyses presented in this study have certain limitations as well. Firstly, the confidence intervals and the tests were computed based on the assumption that the rehoming times follows normal distribution. Secondly, the data was assumed to be independent while having equal variances which may not hold true. Finally, there was not enough data to work with some breeds, leading to discrepancies.

Despite these limitations, several strengths marked the analysed results. The ability to use statistical and graphical analysis to analyse the rehoming times provided a strong basis for identifying the patterns and differences.

In conclusion, the comprehensive analysis gave meaningful insights into the rehoming patterns of different dog breeds, helping us understand the various factors affecting the adoption of dogs, thereby helping us improve the condition of stray and neglected animals.