



# UNIVERSITY OF LEEDS

MATH5741M Statistical Theory and Methods

## Dogs Rehoming Times: A Statistical Comparison between Three Breeds

**Report Prepared by:**

Nilavan Sritharan (shvj0061)

201997034

Data Science & Analytics

University of Leeds

**Submitted to:**

Prof. Elham Mirfarah

School of Mathematics

University of Leeds

**Submission Date:** 12 December 2025

# 1 Introduction

The dogs that are unwanted, neglected, or stray are often sent to a rehoming centre in the hope of finding them a new home. The main outcome for each dog is the time it takes, in weeks, to be rehomed. Previous research has suggested that the average rehoming time is 27 weeks. However, it is unclear whether this trend applies to all breeds within the local shelter context.

This report investigates rehoming times for three specific breeds: Border Collies, Mixed Breeds, and West Highland White Terriers. The aim is to determine whether the average time to rehome these breeds differs from the benchmark of 27 weeks using exploratory summaries and confidence intervals.

## 2 Data & Methods

### 2.1 Data Description and Cleaning

The dataset comprised 825 observations from an animal shelter, recording rehoming time (weeks), breed, health score, age, reason for surrender, and return status. Data cleaning removed 16 observations (1.9%): 6 with missing breed information and 10 with invalid rehoming times (coded as 99999 or negative). The final dataset contained 809 observations: Mixed Breed (n=709, 87.6%), Border Collie (n=78, 9.6%), and West Highland White Terrier (n=22, 2.7%).

### 2.2 Exploratory Analysis and Normality Assessment

Descriptive statistics, boxplots, histograms, and Q-Q plots were produced for each breed. Normality was assessed using Q-Q plots (comparing sample quantiles to theoretical normal quantiles) and the Shapiro-Wilk test, which provides a W statistic (values near 1.0 indicate a good fit) and a p-value testing the null hypothesis of normality.

### 2.3 Statistical Methods

**Distribution Modelling:** The Normal distribution  $N(\mu, \sigma^2)$  was proposed with parameters estimated via Maximum Likelihood Estimation:  $\hat{\mu}$  = sample mean,  $\hat{\sigma}$  = sample standard deviation.

**Inference:** 95% confidence intervals were calculated using the t-distribution:  $CI = \bar{x} \pm t^* \times (s/\sqrt{n})$ , where  $t^*$  is the critical value with (n-1) degrees of freedom. The t-based confidence interval was chosen because the population standard deviation is unknown. For Mixed Breed (n=709), the Central Limit Theorem (CLT) ensures approximate normality of the sampling distribution. For Border Collie (n=78), CLT applies with a moderate sample size. For the West Highland White Terrier (n=22), the t-test is reasonably robust despite departure from normality.

**Comparison:** Pairwise comparisons used Welch's t-test (unequal variances) due to differing sample sizes and variances. Confidence intervals excluding zero indicate significant differences at  $\alpha=0.05$ .

## 3 Results

### 3.1 Descriptive Statistics

Table 1: Summary Statistics for Rehoming Time

Breed	n	Mean (weeks)	Median	SD	Variance
Mixed Breed	709	18.97	17.00	9.49	90.08

Border Collie	78	20.47	18.00	11.87	140.98
West Highland White Terrier	22	20.23	18.50	10.11	102.28

All breeds (Table 1) showed means well below 27 weeks. Mixed Breed dominates the sample (709 of 809 dogs). Border Collie exhibited the highest variance, indicating greater heterogeneity. Medians slightly below the means suggest mild right skewness.

**Boxplots** (Figure 1) reveal that all three breeds have median rehoming times around 17-18 weeks, substantially below the 27-week (red dashed line), with Border Collie showing the greatest variability and Mixed Breed displaying several high-value outliers.

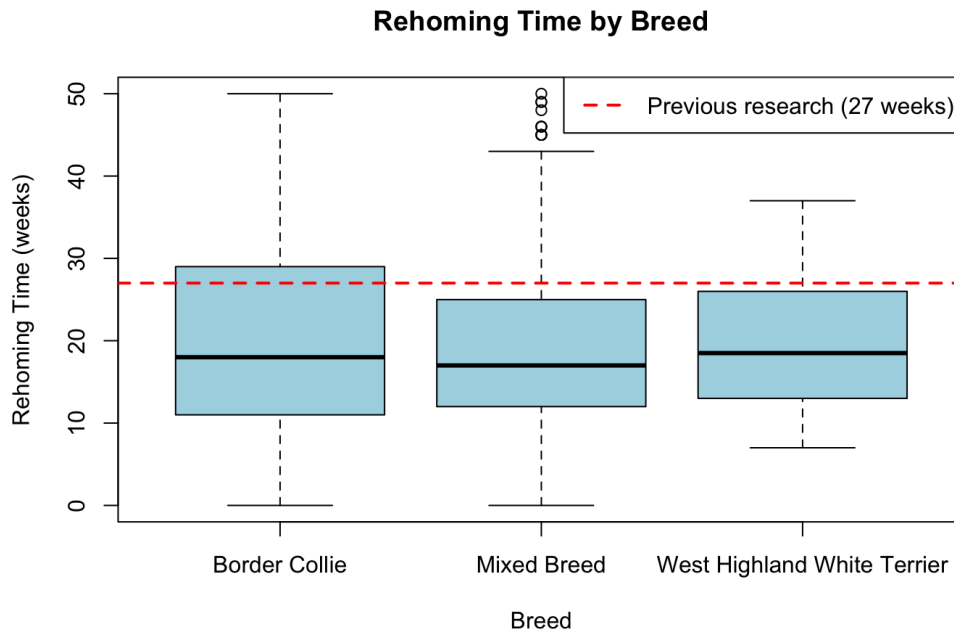


Figure 1 – Boxplot of Rehoming Time by breed

### 3.2 Distribution Characteristics

**Histograms** (Figure 2) revealed Mixed Breed has an unimodal, right-skewed distribution; Border Collie showed a less smooth distribution, with hints of multiple clusters; West Highland White Terrier displayed irregular patterns due to a small sample size.

**Q-Q plots** (Figure 3) showed Mixed Breed followed the theoretical line closely (approximate linearity in the centre with upper-tail deviations), Border Collie showed a reasonable fit with slight deviation, and West Highland White Terrier showed a substantial departure.

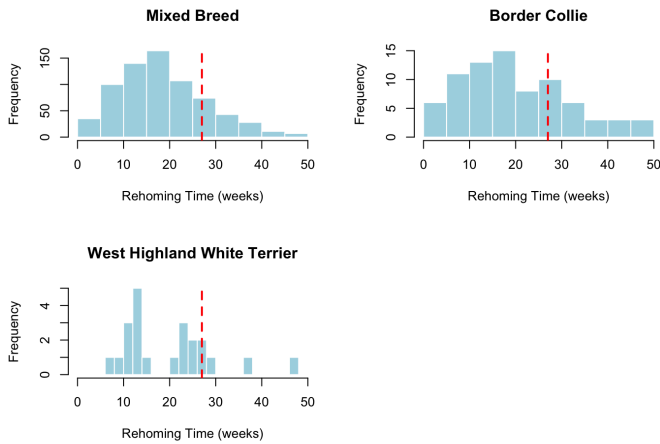


Figure 2 – Histograms by Breed

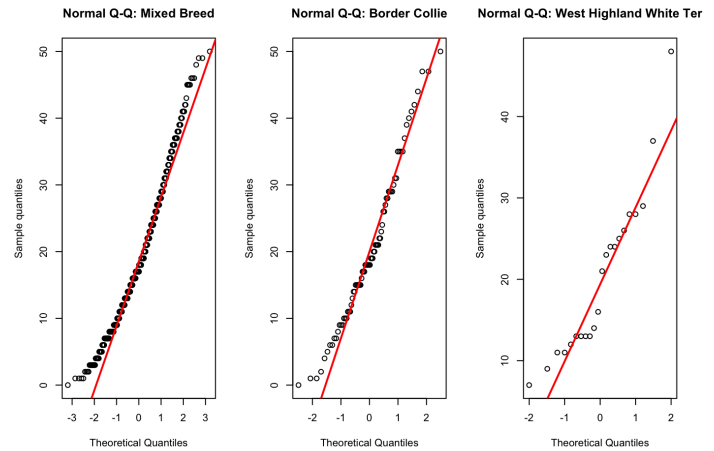


Figure 3 – Q-Q plots by Breed

Table 2: Normality Test Results

Breed	$\hat{\mu}$	$\hat{\sigma}$	W	p-value	Assessment
Mixed Breed	18.97	9.49	0.9711	<0.001	Excellent fit
Border Collie	20.47	11.87	0.9650	0.0306	Good fit
West Highland White Terrier	20.23	10.11	0.9025	0.0333	Poor fit

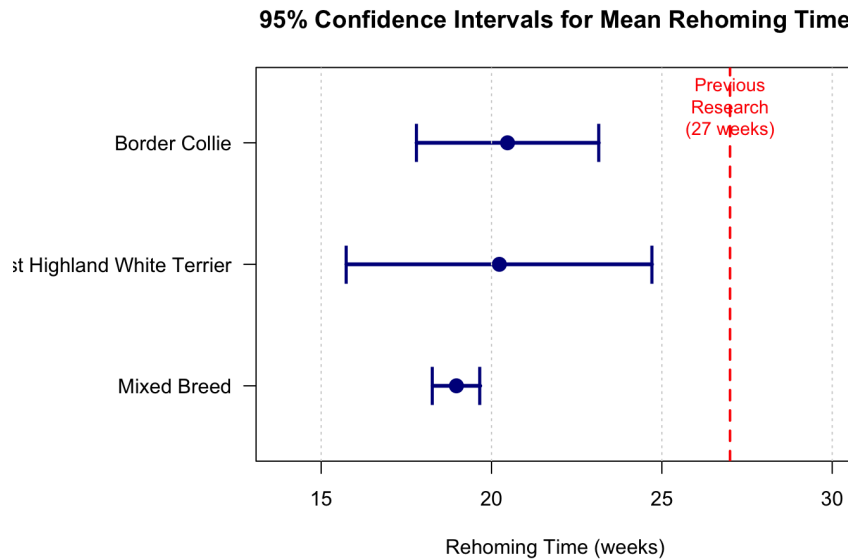
Formally, Shapiro–Wilk (Table 2) rejects normality for all three breeds ( $p < 0.05$ ). However, the W statistic (closeness to 1) provides a better indicator of practical fit. Mixed Breed's  $W=0.9711$  (very close to 1.0) indicates an excellent fit despite the significant p-value, which reflects the large sample size ( $n=709$ ), making the test extremely sensitive to minor deviations. Border Collie's  $W=0.9650$  indicates a good fit, while West Highland White Terrier's  $W=0.9025$  confirms a meaningful departure from normality.

### 3.3 Confidence Intervals

Table 3: 95% Confidence Intervals for Mean Rehoming Time

Breed	Mean	SE	95% CI	Contains 27?
Mixed Breed	18.97	0.356	(18.27, 19.67)	No
Border Collie	20.47	1.344	(17.79, 23.16)	No
West Highland White Terrier	20.23	2.156	(15.75, 24.71)	No

All confidence intervals (Table 3) exclude 27 weeks, indicating that each breed was rehomed significantly faster ( $p < 0.05$ ), with a 24-30% reduction (6.5-8 weeks faster).



*Figure 4 – 95% confidence intervals for Mean Rehoming Time*

Figure 4 visualises these confidence intervals with a reference line at 27 weeks. All three intervals lie entirely below the 27-week line, demonstrating that:

1. All three breeds in this sample rehome **significantly faster** than the previously reported average of 27 weeks.
2. The magnitude of the difference is substantial (6.5 to 8 weeks faster), representing a 24-30% reduction.
3. Precision varies considerably by sample size, with Mixed Breed having a very narrow interval and West Highland White Terrier having a much wider interval.

### 3.4 Pairwise Comparisons

*Table 4: Pairwise Comparisons*

Comparison	Difference (weeks)	95% CI	Significant?
Border Collie - Mixed Breed	1.50	(-1.20, 4.19)	No
Mixed Breed - West Highland White Terrier	-1.26	(-6.26, 3.74)	No
Border Collie - West Highland White Terrier	0.24	(-5.47, 5.96)	No

All confidence intervals contain zero, indicating **no significant differences** between breeds ( $p \geq 0.05$ ).

## 4 Discussion & Conclusion

### 4.1 Main Findings

This analysis found that all three breeds rehome significantly faster than the 27-week average, with mean times of 18.97-20.47 weeks (24-30% reduction). However, no significant differences were detected between the breeds themselves. The three studied breeds differ from the previously reported overall average.

## 4.2 Practical Significance

The 6-8 week reduction in rehoming time has substantial practical implications for animal shelters: reduced housing costs, increased capacity to help more animals, improved animal welfare through less kennel time, and higher adoption rates. The similarity between breeds does not provide strong evidence that breed-specific programmes are needed for these three breeds in this shelter.

## 4.3 Limitations

**Sample size imbalance:** Mixed Breed dominates the sample (87.6%) while West Highland White Terrier comprises only 2.7% (n=22), resulting in wide confidence intervals and low power for this breed.

**Normality violations:** West Highland White Terrier departed substantially from normality. While the t-test is reasonably robust, this remains a concern with small samples.

**Data source:** Results are based on shelter data from a specific context. Generalizability to other shelters, regions, or time periods is uncertain without further multi-centre studies.

**Outcome only:** you focused on Rehomed and Breed; other factors (health, age, reason for surrender) might also affect rehoming time, but weren't modelled.

**Multiple testing:** Six hypothesis tests were conducted without adjustment (e.g., Bonferroni correction), which increases the risk of Type I error, though results were consistently significant or non-significant across all tests.

## 4.4 Conclusion

This analysis provides evidence that Mixed Breed, Border Collie, and West Highland White Terrier dogs are rehomed significantly faster than the previously reported 27-week average (18.97-20.47 weeks), although the three breeds do not differ significantly from each other. While sample size limitations and normality violations (particularly for West Highland White Terrier) require cautious interpretation, the findings suggest dogs from these breeds can be expected to find homes relatively quickly. The practical implications, including reduced costs, increased capacity, and improved animal welfare, are substantial. Further research with larger, more balanced samples and investigation of additional factors would strengthen the understanding of what drives rehoming success.

## 5 References

- MATH5741M Lecture Notes (2025/26). Statistical Theory and Methods. University of Leeds.
- MATH5741M Practical Worksheet (2025/26). Statistical Theory and Methods. University of Leeds.