

# Age at Intake and Length of Stay Among Adopted Dogs at San Jose Animal Shelter (2018–2025)\*

Not Just About Age: Rethinking Adoption Timelines

Anna Wadlow

September 24, 2025

San Jose Animal Care and Services (ACS) has faced sustained overpopulation, making it important to understand drivers of animals' length of stay (LOS). Using ACS open data (July 2018–present), we examined whether a dog's age at intake predicts LOS among dogs whose outcome was adoption ( $n = 8,846$ ). We fit a simple linear regression of LOS (days) on age (years). The estimated slope is small (0.85 days per additional year; intercept 18.38 days), indicating only a weak, positive association. Model diagnostics show violations of linear model assumptions and a clear shift in residual patterns beginning around 2022. These patterns suggest unmodeled factors, such as intake condition, behavior, breed/size, or time-varying shelter operations, play a larger role than age alone. We conclude that age at intake, by itself, is a poor predictor of LOS at ACS and recommend incorporating additional variable into the model to better inform operational decision-making and welfare outcomes.

## 1 Introduction

San Jose Animal Care and Services (ACS) is a full-service shelter that has served San Jose and surrounding cities since 2004. In recent years, ACS has frequently operated over capacity, triggering an official city audit in the fall of 2024 (City Auditor 2024). Shelter population directly affects animal welfare and operational efficiency. Longer lengths of stay (LOS) increase the risk of disease — especially contagious respiratory infections — and constrain staff resources,

---

\*Project repository available at: <https://github.com/anna-wadlow/MATH261A-Project-1>.

which can in turn lower adoption rates and contribute to the overpopulation (Maddie’s Shelter Medicine Program 2018). At the extreme, prolonged LOS can intersect with population management tools, such as euthanasia.

What this paper aims to understand is if a dog’s age at intake is associated with its LOS, among dogs ultimately adopted from ACS.

Prior work in shelter contexts generally treats age as one of several potential correlates of LOS, alongside factors such as intake type/condition, breed/size, color/appearance, behavior notes, and risk of euthanasia (Garrison and Weiss 2014). However, the effects of age can vary by facility and period, motivating an ACS-specific analysis that uses the shelter’s own records to quantify the association.

We use the City of San Jose’s open ACS data (updated daily) and restrict the analysis to dogs whose outcome was adoption between July 2018 and the present. A simple linear regression is used to relate LOS to age. We found a modest, positive estimated slope of about 0.85 additional days of LOS per year of age increase. But, diagnostic plots indicate the linear model performs poorly (heteroskedasticity, non-normal residuals, and time-related shifts) suggesting the need for a more robust model.

Section 2 describes the ACS data and key variables used. Section 3 outlines the linear modeling approach and Section 4 presents estimates and diagnostics. Section 5 interprets findings, reviews limitations, and proposes methodological and operational next steps.

## 2 Data

Data for this paper was obtained from the City of San Jose Open Data Portal. ACS provides intake and outcome records for animals housed at their facility with daily updates covering July 1, 2018 through present. This paper focuses on dogs with an adoption outcome, yielding 8,846 records after data filtering.

The key variables, organized in columns, are animal type, date of birth (DOB), intake date, outcome date, and outcome type. Additional descriptors include the animals’ primary breed, sex/sterilization status, primary/secondary colors, intake/outcome condition, intake type/sub-type, and intake reason.

The proposed predictor variable, dog’s age in years at intake is calculated from intake date and DOB variables. The model’s dependent variable, length of stay, is the difference in days between the outcome date and intake date. The median age is 1.56 years and the median length of stay is 10 days.

Figure 1 shows quite a bit of variability, with no clear pattern, between the dogs’ age at intake and their length of stay at the shelter.

Many dogs lacked data for date of birth and were excluded from the study. It is also unclear in the raw data if there are records for animals returned to the shelter after adoption, creating a “new” record. The 2024 city audit (City Auditor 2024) also flagged opportunities for ACS to improve their data management and reporting, confirming the need for careful consideration when interpreting results.

The standard linear regression assumptions (independent errors, constant variance, normality of residuals, and linear relationship) are evaluated in Section 4.

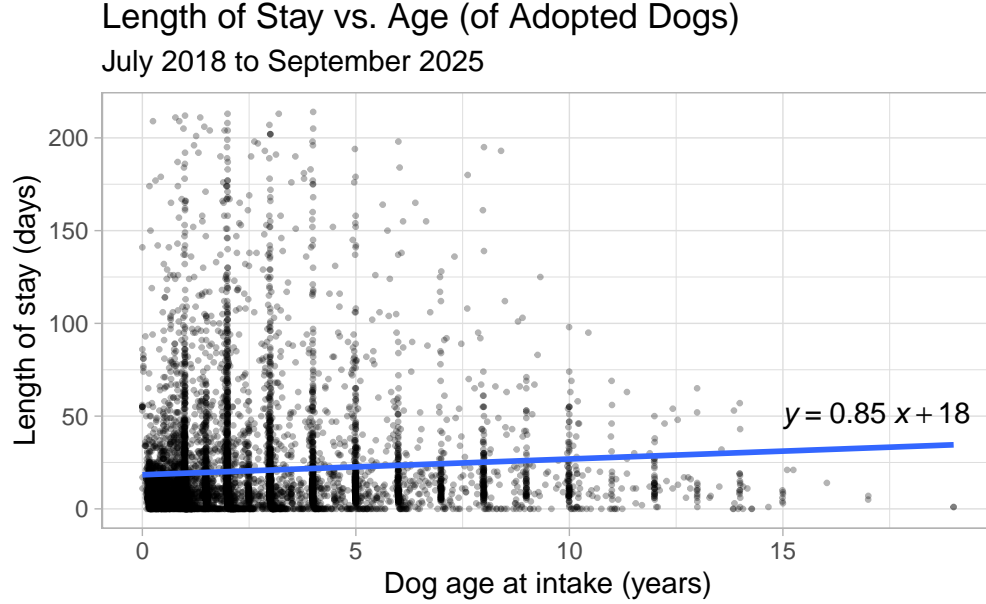


Figure 1

### 3 Methods

We use a simple linear regression model, comparing length of stay (LOS) to age at intake.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

The predictor variable, age of dogs at intake, is denoted by  $X_i$  and the response variable, length of stay, is denoted by  $Y_i$ .

The slope,  $\beta_1$ , represents the estimated change in length of stay per additional year of age. The intercept,  $\beta_0$ , represents the estimated length of stay for a zero-year-old dog. The residuals,  $\varepsilon_i$ , are assumed to be approximately normally distributed.

Using a simple linear regression model for this analysis may be overly rudimentary, ignoring other variables that can influence length of stay, but provides a baseline investigation.

We use the base R's `lm()` function to implement the linear model (R Core Team 2025) alongside functions from the `dplyr` (Wickham et al. 2023) and `tidyverse` (Wickham et al. 2019) packages for data cleaning and visualization.

Model fit was evaluated using diagnostic plots discussed in Section 4.

## 4 Results

The estimated slope parameter is  $b_1 = 0.85$ , which can be interpreted to mean that a one-year increase in dog's age, on average, corresponds to  $\sim 0.85$  additional days in LOS. The estimated intercept is  $b_0 = 18.38$ , the expected LOS for a 0-year-old dog.

Residual plots Figure 2 show important assumption violations. Mean residuals appear positive at younger ages and trend negative at older ages, suggesting under-prediction for young dogs and over-prediction for older dogs. The variance of residuals is much larger among younger dogs, violating constant variance (heteroskedasticity). Residuals plotted against intake date show a marked structural change beginning around 2022, consistent with time-varying shelter conditions and, possibly, data management shifts not captured by the model.

Finally, the QQ plot Figure 3 shows pronounced right-skew and heavy tails, indicating residual non-normality.

## 5 Discussion

Age at intake has, at best, a weak positive association with LOS among adopted dogs at ACS. Visualizations in Section 4 reinforce that LOS varies widely at all ages and that a straight-line relationship is an oversimplification.

Consistent with previous research in this area, age at intake is a weak and unreliable predictor of LOS among adopted dogs at ACS. For operations and welfare decisions, relying solely on age as the predictor of length of stay would likely misdirect resources. A multi-factor profile should be explored to better understand the true relationship.

Several data limitations were identified and pose accuracy concerns of the model. Many dogs lack a recorded or reliable DOB and as such, these cases were removed, potentially introducing a selection bias. It is unclear whether each record reflects a unique dog or whether returns/re-adoptions appear as separate rows. More generally, the 2024 audit flagged opportunities to improve ACS data management and reporting.

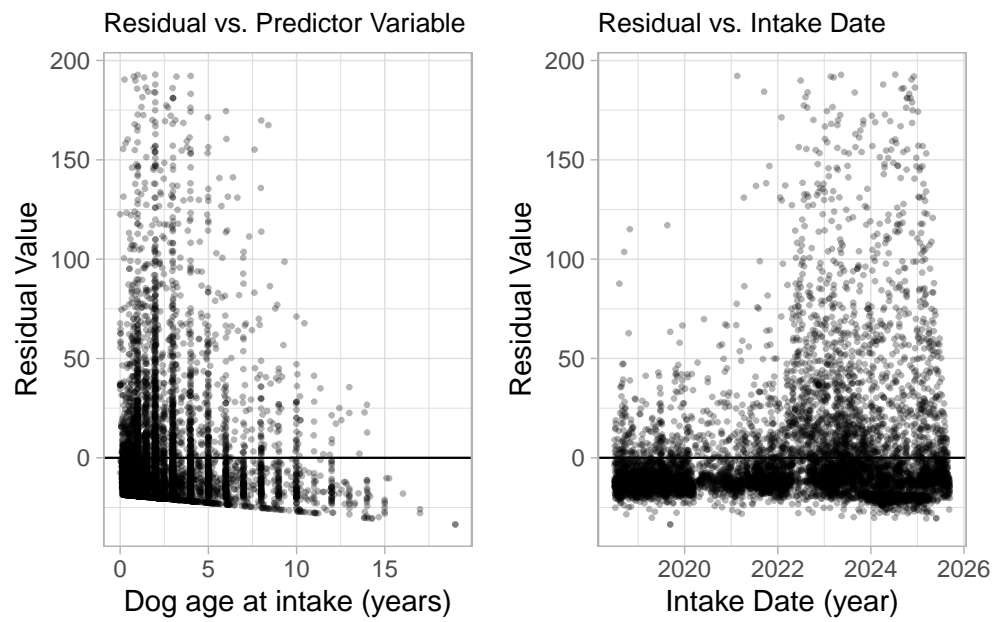


Figure 2

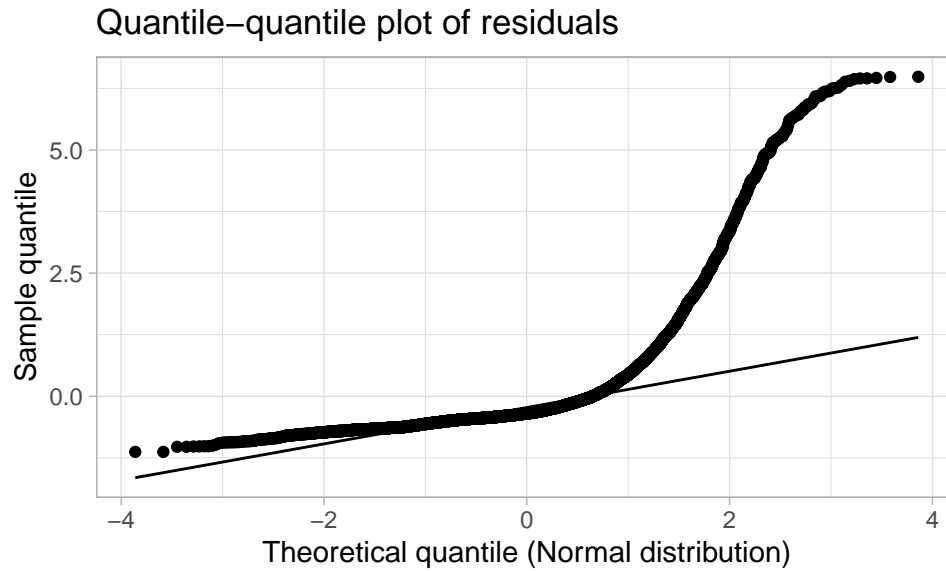


Figure 3

As a modeling extension, additional variables should be considered to develop a better fitting model. Suggestions for these include the available descriptive variables: intake type/condition, sex/sterilization, breed/size, color/appearance, behavior notes, outcome subtype, and calendar time (e.g., year, month, pandemic/post-pandemic indicators). And, shifting the model’s goal to look at the probability of exceeding operational thresholds (e.g., LOS > 30 days) would better support capacity planning at ACS.

## References

- City Auditor, Office of the. 2024. “24.06 Animal Care and Services: Addressing Capacity Constraints and Updating Shelter Protocols Will Help Improve Animal Outcomes.” <https://www.sanjoseca.gov/home/showdocument?id=116106>.
- Garrison, Laurie, and Emily Weiss. 2014. “What Do People Want? Factors People Consider When Acquiring Dogs, the Complexity of the Choices They Make, and Implications for Nonhuman Animal Relocation Programs.” *Journal of Applied Animal Welfare Science : JAAWS* 18 (August): 1–17. <https://doi.org/10.1080/10888705.2014.943836>.
- Maddie’s Shelter Medicine Program, College of Veterinary Medicine. 2018. “Canine Respiratory Infections in Animal Shelters.” *University of Florida*. <https://sheltermedicine.vetmed.ufl.edu/files/2017/01/Canine-respiratory-infections-in-shelters.2018.pdf>.
- R Core Team. 2025. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://doi.org/10.32614/CRAN.package.dplyr>.