Predicting Time to Adoption in Animal Shelters Using Survival Analysis

M.Srinija

Hyderbad, India

12 May, 2025

I. Abstract

The US animal shelters house stray or unwanted animals and are vital to animal welfare. Survival analysis methods are used here on Austin Animal Center (2013–2025) intake and outcome data to find factors influencing animal adoption time. The aim is to estimate how different characteristics (e.g., age, sex, type of intake) impact the time animals spend in shelters prior to adoption. Kaplan-Meier survival curves and log-rank tests are applied to compare group adoption survival (e.g., species, sex). A multivariate model is created to test the impact of multiple features on adoption velocity. The outcomes provide actionable insights that can inform shelters on how to improve adoption approaches and shorten animal wait times.

Keywords: animal shelter, adoption, survival analysis, Cox model, Kaplan-Meier, Austin Animal Center

II. Dataset Description

Dataset used in this project is publicly available from the Austin Open Data Portal. Animal Center Intakes and Animal Center Outcomes from Oct, 1st 2013 to date. Intakes are the condition of animals as they come into the Animal Center. All animals are assigned an individual Animal ID upon intake. More than 90% of animals entering the facility each year, are adopted, sent to rescue or reclaimed by their owners. Outcomes are the disposition of animals when they depart the Animal Center. The Outcomes data set indicates Austin, TX. is the nation's largest "No Kill" city. The data set consists of two files:

a. Intake data: Has fields such as Animal ID, Intake DateTime, Intake Type, Intake Condition, Sex upon Intake, Age upon Intake, Breed, and Color.
b. Outcome data: Has fields such as Animal ID, Outcome DateTime, Outcome Type (such as Adoption, Transfer), Outcome Subtype, Sex upon Outcome, and Age upon Outcome.

Following the merging of both datasets based on Animal ID, we compute a new column Duration to capture the duration (in days) from intake to adoption. The isAdopted column is used as the event indicator (1 = adopted, 0 = not adopted or other).

III. Method

All analysis was performed using Python, with key libraries including pandas, lifelines, matplotlib, and seaborn. The Kaplan-Meier estimator is a statistical estimation procedure applied during survival analysis that estimates the likelihood of an event not happening within a specific point in time. It produces a stepwise survivor curve that describes the probability of subjects "surviving" as time proceeds. One of its greatest attributes is the way it can address the event that fails to happen through the observation phase. Here, it was applied to estimate the duration for which animals remain in the shelter until adoption, and to contrast survival times between groups. It produces an interpretable curve that indicates differences in adoption patterns by time, which can inform shelters about the kind of animals get adopted early.

IV. Data Cleaning and Pre-processing

In order to get the dataset ready for analysis, we have initially filtered the data so that only the adoption-related outcomes were kept. The outcome and intake dates were changed into datetime format, and a new column Duration was calculated to express the number of days from intake to adoption. Negative or zero duration records were dropped and a binary column isAdopted was added to signify the occurrence of an adoption event. Categorical variables like sex, type of intake, and intake condition were cleaned by normalizing categories. Age, initially recorded in different formats (weeks, months, years), was transformed uniformly into days. In addition, new fields such as age group, breed type, and intake month/year were designed to facilitate trend analysis and model building.

V. Results & Discussion

Distribution of adoption duration is right-skewed (long tail), such that the majority of animals get adopted within 30 days and a lesser proportion remain for a longer time(Fig 1).
This could be possibly due to medical or behavioral issues or low-demand characteristics.

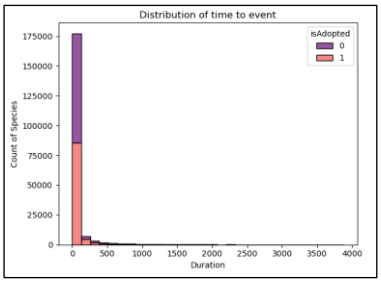

Fig 1: Distribution of Adoption Duration

Dogs and cats are the most adopted animals, but there can be a little preference or quicker turnover for cats in the statistics (Fig 2). Species by itself might be a poor predictor, but still helpful when paired with age or condition of intake. Sterilized animals will have lower median length (Fig 3). Owner-surrendered animals will be adopted quicker than strays. Public opinion of health and behavior may drive quicker adoption.

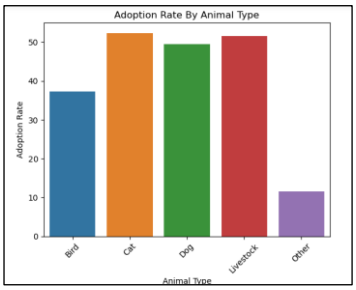

Fig 2: Adoption Rate Based on Animal Type

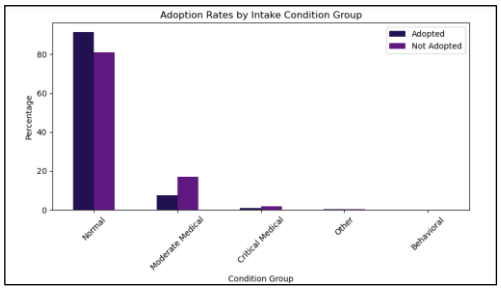

Fig 3: Adoption Rate Based on the Intake Condition

Adoption by Month or Year based on Seasonal patterns were apparent (Fig 4(a)). Adoptions peak in spring and early summer months (Fig 4(b)). Arranging adoption drives or campaigns during peak-adoption months may be more efficient. There is a steady rise in the number of adoptions by the years.

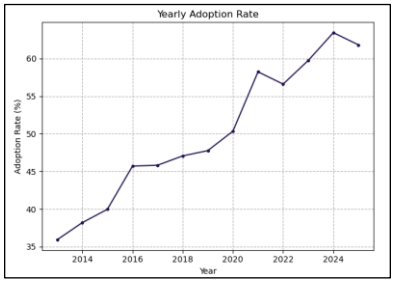

Fig 4(a): Variation of Adoption rate throughout years

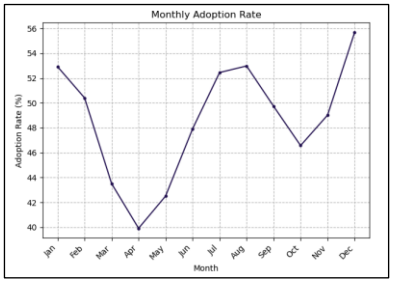

Fig 4(b): Variation of Adoption rate throughout months

VI. Survival Analysis

Kaplan-Meier survival plots were paired with log-rank tests to statistically validate whether the observed differences were significant. The survival curve for cats declines more steeply (Fig 5), meaning they are adopted sooner than dogs. Species affects adoption speed; shelter efforts might focus more on long-stay dogs.

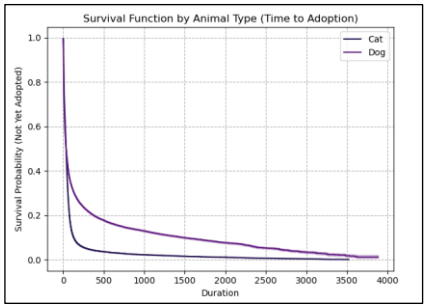

Fig 5: Kaplan-Meier Curve – Cats vs Dogs

The survival curves revealed that neutered males and spayed females had steeper declines (Fig 6) in survival probability compared to their intact counterparts. This pattern suggests that sterilized animals were adopted faster. Sterilization may reassure adopters and speed up decisions.

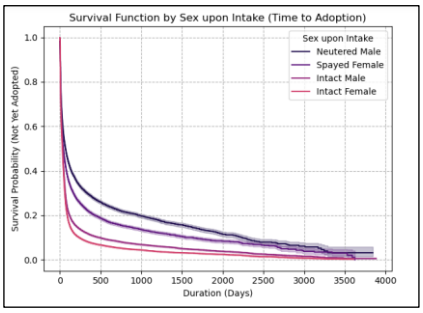

Fig 6: Kaplan-Meier Curve by Sex upon Intake

Furthermore, Kaplan-Meier plots by age group showed that younger animals were adopted more rapidly (Fig 7) than seniors or adults. Senior animals represented a flatter curve, showing extended shelter stays. This comports with general trends when adopters tend to prefer younger pets, seeing them as more trainable and living longer.

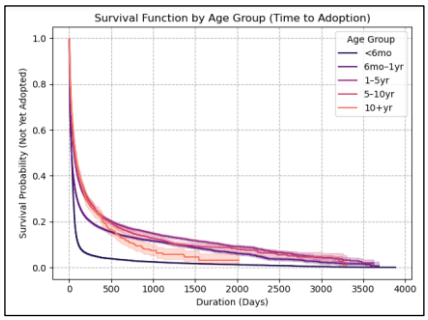

Figure 7: Kaplan-Meier Curve by Age Group

VII. Conclusion

Survival analysis established that features including species, age, sex, and intake condition significantly impact how rapidly animals get adopted from the shelter. The Kaplan-Meier curve as well as Cox model established that younger animals, particularly kittens and puppies, exhibit much shorter lengths of stay within shelters when compared to adults and seniors. Animals spayed or neutered upon intake were adopted more quickly, showing that sterilization status is an important factor in adopter choice. Animals categorized as "Normal" or healthy upon intake were more likely to be adopted than animals tagged as sick, injured, or with behavioral issues, implying the utility of medical care and rehabilitation. Methods such as Kaplan-Meier estimation, log-rank testing, and Cox Proportional Hazards modeling provide the shelters with an effective tool to learn and optimize adoption strategies through identifying important variables that affect animal stay length.

VIII. References

[1] E. L. Kaplan and P. Meier, “Nonparametric estimation from incomplete observations,” *J. Am. Stat. Assoc.*, vol. 53, no. 282, pp. 457–481, 1958.

[2] C. Davidson-Pilon, *lifelines: Survival Analysis in Python*, 2019.

[3] Austin Animal Center, “Animal Intakes and Outcomes Dataset,” *Austin Open Data Portal*.

[4] H. Ural and B. Tas, “Predicting Animal Shelter Pet Adoption Times and Feature Importance Analysis using CatBoost,” *ResearchGate*, Sep. 2023.

[5] A. Diesel, J. Smith, and S. Pfeiffer, “Factors affecting time to adoption of dogs re-homed by a charity in the UK,” *Animal Welfare*, vol. 16, no. 3, pp. 353–360, 2007.