1. Introduction

I used San Jose's animal shelter data to forecast daily and weekly animal intakes for the next six months. This is critical for the shelter to ensure the effective resource allocation and operational efficiency. In order to help decision-makers better manage shelter capacity, staffing, and purchasing, this project aims to create a dependable forecast model that offers practical insights. Shelter managers can enhance planning and allocate resources more efficiently by examining past trends and forecasting future ones.

1. Data Overview

The Animal shelter data contains the date of intake, unique animal identifiers, types of intakes and other variables from 2022-2024. To prepare this data, the intake records were sorted by date, and daily counts of animal intakes were calculated through aggregation. I also excluded entries with the intake condition "Dead" to focus on active shelter operations. These daily counts are the basis for creating the time series data. The time series plot of daily intakes shows fluctuations and has potential seasonality. The pie chart was created to show the distribution of intake types, and it shows that the Stray animal contributed significantly to the overall intake numbers.

1. Time series analysis

I created two time series with different frequency. The daily counts converted into time series with frequency of 365 capture temporal patterns and the weekly aggregations has frequency of 52 and it capture the seasonal trends. I decide to use frequency of 52 since I want to understand the seasonal trends of the intake amount. Then the time series data was decomposed using STL decomposition.

1. Forecast Model

I use naïve forecast, mean forecast. But to account for seasonality, I use seasonal naïve forecast model to predict future values based on the same period in pervious season. I also using exponential smoothing and ARIMA methods to capture both trend and seasonal patterns.

1. Accuracy Metrics and Residual Analysis

The main accuracy metric I used to evaluate the models was MAPE, because it gives a clear understanding of forecast accuracy in comparison to actual values. In the forecast models, the ETS model and ARIMA model performed the best in terms of accuracy with lower MAPE values compared to simpler models like naïve and mean forecasts. The seasonal naïve model captures the seasonal repetition but was less reliable for long-term projections. For the ETS and ARIMA models, residual analysis showed consistent trends with minimal autocorrelation.

1. Tableau Visualizations

I created Tableau dashboards to provide additional details and visualizations to the forecasting models. Trends in monthly shows seasonal peaks and expected rises in early 2025. The type of Intake shows the stray animals accounted for most intakes, which highlights a steady demand for resources for shelter. Overall adoption rates were found to be high, however there were some declines in recent months. Seasonality heatmap highlighted the significance of seasonal forecasting by displaying daily and monthly patterns.