Preserving Harmony: Analyzing Traffic Patterns and Wildlife Interactions in Boonsong Lekagul Nature Preserve

COMP4449 Capstone

Yixuan Huang

# Introduction

## Purpose

This project endeavors to investigate and analyze traffic patterns within the Boonsong Lekagul Nature Preserve, with a primary focus on understanding potential correlations with the observed decline in nesting pairs of the Rose-Crested Blue Pipit. The goal is to employ data analytics methodologies to gain valuable insights into the movements and behaviors of vehicles within the preserve. By identifying patterns, we aim to shed light on factors that may impact the bird population. This analysis serves the overarching mission of contributing to evidence-based conservation strategies for the preservation of the Rose-Crested Blue Pipit and its habitat. The paper encompasses comprehensive dataset analysis, exploratory visual analytics, and the application of data-driven insights to inform conservation efforts.

## Significant

The significance of this project extends beyond the realm of data analysis; it is a call to action for preserving the delicate balance between human activities and the thriving biodiversity within the nature preserve. Through visualizations and data-driven insights, we aspire to empower conservationists and policymakers with the tools needed to make informed decisions and safeguard the natural harmony of Boonsong Lekagul.

## Research Question

My research questions for this analysis are: Which areas within the preserve experience the highest concentration of traffic? Which areas within the preserve experience the highest concentration of traffic?

## Dataset

The dataset for this project comprises a comprehensive collection of 171,477 instances recorded from sensors strategically placed around Boonsong Lekagul Nature Preserve. The data spans from May 2015 to May 2016, capturing a year-long snapshot of vehicular activity within the preserve. Including columns: Timestamp (Indicates the date and time of the sensor reading), Car-id (Represents the assigned car ID from the entry gate), Car-type (Enumerates the vehicle type, including 2 axle car (or motorcycle), 2 axle truck, 3 axle truck, 4 axle (and above) truck, 2 axle bus, and 3 axle bus), Gate-name: Specifies the name of the sensor taking the reading, categorized as Entrances, General-gates, Gates, Ranger-stops, and Camping.

A black background with white lines

Description automatically generated

In addition to the sensor data, a map of the park has been provided, featuring distinct colors for each sensor type: Entrances (green), General-gates (blue), Gates (red), Ranger-stops (yellow), and Camping (orange). This map, oriented with north at the top, will play a crucial role in visualizing and understanding the paths taken by vehicles within the preserve.

# Data Preprocessing

Data preparation involves data type conversion, time series features creation, including day of the week, month, and hour, were derived to facilitate in-depth temporal exploration. Gate type and number extraction, and path data frame creation.

## Exploratory data analysis and visualization techniques

A graph of travel breakdown

Description automatically generated

This plot, depicting travel popularity by months (excluding Ranger vehicles), indicates a peak in visitor activity from May, reaching a peak during July, gradually declining since August, and remaining low until the following May. This pattern suggests a seasonal trend, with visitors predominantly exploring Boonsong Lekagul Nature Preserve in the warmer months.

A graph with blue rectangular bars

Description automatically generated with medium confidence

Delving into the intricate pathways within Boonsong Lekagul Nature Preserve, a spotlight was cast on the top 6 paths, each unveiling unique facets of vehicular movement.

A screenshot of a computer screen

Description automatically generated A screen shot of a graph

Description automatically generated

The first path emerged as a dedicated route for ranger vehicles, following consistent daily patterns of approximately 50 or 30 minutes. A reasonable assumption is that these vehicles may traverse these routes for shift switches or essential tasks at ranger stops. Notably, a temporal gap in January suggests a potential influence from holiday activities around the end of the year.

![A map of a city

Description automatically generated]() A graph with numbers and dots

Description automatically generated

The second high-traffic path, intriguingly, witnesses visitors entering and exiting without making any stops. Despite an entrance fee, this behavior hints at the possibility of using it as an express toll or a shortcut to another location or just checking out the view of the park, with an average completion time of around 20 minutes.

A map of a city

Description automatically generatedA graph of different colored bars

Description automatically generated

The third path sheds light on camper routes, portraying a right-skewed plot indicative of shorter stays rather than extended ones. These insights, extracted from the top paths, provide a nuanced understanding of the varied motivations driving vehicular activity within the Preserve, contributing to a more comprehensive view of visitor behaviors and potential correlations with the natural environment. Please check the code for more route examples.

A graph of different colored lines

Description automatically generatedA graph showing different colored lines

Description automatically generated

A screenshot of a computer

Description automatically generated A screenshot of a graph

Description automatically generatedA screen shot of a graph

Description automatically generated

A glance at the "Monthly Traffic Volumes at Different Gates" plot unveils several notable patterns. As anticipated, general gates (depicted in green) unsurprisingly register the highest vehicle traffic in comparison to entrances and camping areas throughout the months. What stands out is the intriguingly consistent popularity trend month over month, forming a stable distribution in vehicles passing through these gates, raising questions about the nature of visitor behaviors. While it hints at a steady flow, the plot also suggests no significant changes in specific gates throughout the year, underscoring the robustness and stability of vehicular activities within Boonsong Lekagul Nature Preserve.

A screenshot of a graph

Description automatically generated

The plot illustrating the "Weekly Distribution of Car Types at Gates" provides a straightforward yet insightful view of vehicular activities within Boonsong Lekagul Nature Preserve. With the x-axis representing days of the week and the y-axis indicating car counts, the plot showcases a uniform distribution across all gates for each car type.

![A computer screen shot of a tree

Description automatically generated]()

The amalgamation of insights gleaned from various plots has spotlighted a compelling concentration of vehicular activity in the northwest of Boonsong Lekagul Nature Preserve. Pinpointing the first two highest volume occurrences on the map reveals that car type 1 (2-axle car or motorcycle) and car type 2 (2-axle truck) dominate this region. While discerning whether 2-axle cars generate more noise than their 3 or 4-axle counterparts remains uncertain, as it hinges on individual car engines or systems, but motorcycles, in general, are recognized for producing louder noises. This potential auditory disturbance introduces an intriguing element that might contribute to the observed decline in the Rose-Crested Blue Pipit population.

However, arriving at a definitive conclusion necessitates more data on the specific locations of bird habitats within the Preserve. Furthermore, we recognize the importance of obtaining information about the timeframe when the bird population started declining. If feasible, having vehicle data recorded during the same timeframe becomes crucial to unravel the real factors impacting the bird population. This multi-faceted investigation underscores the complexity of understanding the interplay between vehicular activities and the well-being of the avian inhabitants, demanding a holistic approach for comprehensive insights.

## Data Splitting

Data was resampled monthly and split into training (May 01, 2015 – May 15, 2016) and testing sets (May 15, 2016 – May 31, 2016).

# Model Building and Evaluation

A table with numbers and letters

Description automatically generated

In the pursuit of forecasting future vehicle volumes, ARIMA model candidates were considered, and the one with the lowest AIC emerged as the optimal choice. The dependent variable was identified as car volume, and statistical tests indicated that the errors exhibit characteristics of white noise, as reflected in a Ljung-Box p-value above 0.05. Moreover, the p-value for heteroscedasticity surpassed 0.05, suggesting consistent variance in the error residuals. The evaluation metrics for the ARIMA model showcased a MAE of 7.54 and a MSE of 94.72. Comparatively, the Prophet model, another contender in the forecasting endeavor, demonstrated a higher MAE of 11.79 but a lower MSE of 11.79.

# Conclusion

While ARIMA outperformed in MAE, Prophet exhibited superior accuracy in minimizing squared errors. This dual-metric evaluation underscores the nuanced trade-offs in model performance and assists in selecting the most fitting model for predicting future vehicle volumes within Boonsong Lekagul Nature Preserve.

## Lessons Learned

The project underscores the critical role of comprehensive data in drawing accurate conclusions about the factors contributing to the decline in the Rose-Crested Blue Pipit population within Boonsong Lekagul Nature Preserve. The absence of specific information about vehicle characteristics and bird habitat locations highlights the challenges in pinpointing the root cause of the decline. This emphasizes the necessity for a more detailed and targeted data collection process aligned with the project's goals. Incorporating coordinates would significantly enhance the spatial analysis, allowing for more efficient mapping of concentrated traffic locations. The project's exploratory nature could further benefit from additional data, enabling the application of diverse analytical techniques for a deeper understanding of the patterns observed. With the provision of more information, the project has the potential to offer more accurate insights into the factors influencing the decline in the bird population, underscoring the importance of robust data collection strategies in ecological studies.