Exploring Pedestrian Activity and Park Utilisation in Casey Suburbs

Yehezkiel Efraim Darmadi
Ydar0001@student.monash.edu
34078215
DEP2 assignment

Table of Contents

Introduction

Urban areas worldwide are undergoing rapid growth, driven by both natural population increases and migration. Suburban municipalities, such as the City of Casey, Melbourne, are at the forefront of this expansion. For instance, Cranbourne East, located within the City of Casey, experienced one of the highest population growths in Victoria, driven by significant natural increases and internal migration (Australian Bureau of Statistics, 2020).

This rapid urbanisation brings challenges in managing public spaces like pedestrian pathways and parks. As cities expand, the demand for accessible, safe, and well-maintained public spaces grows. High pedestrian traffic may indicate areas requiring infrastructure enhancements, such as better lighting or more frequent crossings, while low usage might point to issues like underinvestment or barriers to access (Sharmin & Kamruzzaman, 2018). Additionally, the equitable distribution of parks and recreational facilities is crucial for community well-being, as these spaces are integral to promoting physical health, mental well-being, and social interaction in urban environments (Koohsari et al., 2014).

Motivation

The motivation for this project stems from the critical need to address the challenges posed by rapid urbanisation in suburban municipalities like the City of Casey. As these areas continue to grow, it is essential to understand and respond to the changing patterns of pedestrian activity and park usage. By analysing these patterns, city planners can make informed decisions that enhance public safety, community well-being, and sustainable development.

Understanding pedestrian movement is vital for creating safer urban environments. High-traffic areas may require infrastructure improvements such as additional crosswalks, better lighting, or expanded pathways to accommodate increasing numbers of pedestrians. Conversely, identifying areas with low pedestrian traffic can help uncover issues related to accessibility or safety, guiding targeted interventions (Koohsari et al., 2014).

In addition to pedestrian activity, the equitable distribution and usage of parks and green spaces play a crucial role in promoting physical and mental health. Parks are not merely recreational spaces; they are essential for fostering social interaction, reducing stress, and encouraging active lifestyles. Ensuring that these spaces are accessible and well-utilised can significantly contribute to the overall well-being of the community (Van Dyck et al., 2011).

Ultimately, this project aims to provide actionable insights that can guide urban planning and policy decisions in rapidly growing municipalities like the City of Casey. By leveraging data on pedestrian activity and park usage, planners can ensure that public spaces are designed and managed to meet the evolving needs of their residents, supporting the development of safer, more liveable, and sustainable urban environments.

Questions

1. What are the peak times for pedestrian activity in various parts of the City of Casey suburbs?
2. How does pedestrian activity differ between weekdays and weekends?
3. How do pedestrians utilise parks and reserves in the City of Casey, and are there significant trends in park usage across different seasons?

Data Wrangling

In this section, I will discuss the data description, outline the steps taken for data wrangling, and highlight the tools used to resolve any errors.

Data Description

Below are the data description for the project:

1. Pedestrian Count (Hourly): This dataset contains hourly pedestrian counts collected automatically by IoT sensors at various locations within the City of Casey. The Pedestrian Count dataset comprises two distinct files:

- The first dataset contains 360,401 rows and 27 columns, including spatial attributes, timestamps, and hourly pedestrian counts. It covers the period from August 12, 2021, to October 5, 2022. This dataset lacks a suburb column and has several columns that need to be dropped due to containing only a single value or being insignificant.
- The second dataset has 162,440 rows and 25 columns. This dataset covers the period of 19 September 2023 to 29 August 2024.

Source:
https://data.casey.vic.gov.au/explore/dataset/pedestrian/export/?disjunctive.location_suburb&disjunctive.dev_site&disjunctive.location_desc&disjunctive.sub_location&disjunctive.devicename&sort=date

2. Parks and Reserves Locations:

This dataset details the locations and classifications of parks and reserves within the City of Casey. The data is organised in a table with 16 columns and 3,237 rows, containing spatial attributes. The data covers park and reserves from 30 suburbs across the City of Casey.
Source: https://discover.data.vic.gov.au/dataset/parks-and-reserves-locations

Data Wrangling

Steps	Tool Used	Dataset
Selecting column and merging datasets	R	Pedestrian Count (Hourly)
Checking null and duplicate values	R	Pedestrian Count (Hourly), Parks and Reserves Locations
Handling extreme values and outliers	R	Pedestrian Count (Hourly)
Filtering out the unusable rows	R	Pedestrian Count (Hourly)
Dealing with missing data	R	Pedestrian Count (Hourly)
Creating Primary Keys	R	Pedestrian Count (Hourly), Parks and Reserves Locations
Selecting relevant park and reserves location	R	Parks and Reserves Locations
Merging datasets	R, Tableau	Pedestrian Count (Hourly), Parks and Reserves Locations

Table 1. Data Wrangling Steps

All of the data cleaning and data wrangling processes were performed using R, while some of the data visualisation was carried out using Tableau. The R libraries that are used for the assignment are: ggplot2, ggplot2, lubridate, naniar, tidyverse, readxl, geosphere, forecast, and ggridges.

Data Checking

The process begins by merging two separate datasets for the Pedestrian Count (Hourly). The reason for the approximately one-year gap between the datasets is unclear, as the site does not specify, nor is there any related news. Initially, I selected 8 fields relevant to the analysis from both datasets, including suburb, dev_site, location, date, and val_calibrated. Once the fields were standardised, I combined the datasets.

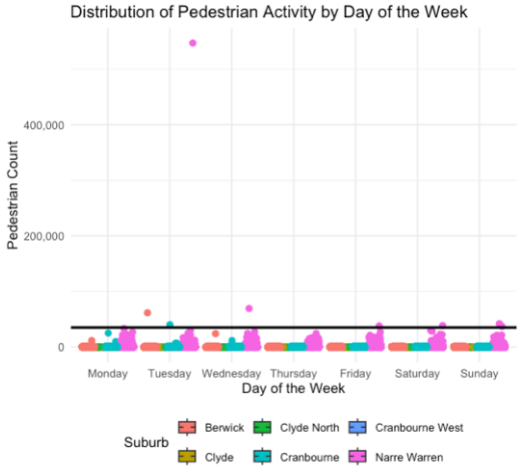

Figure 2. Extreme Values

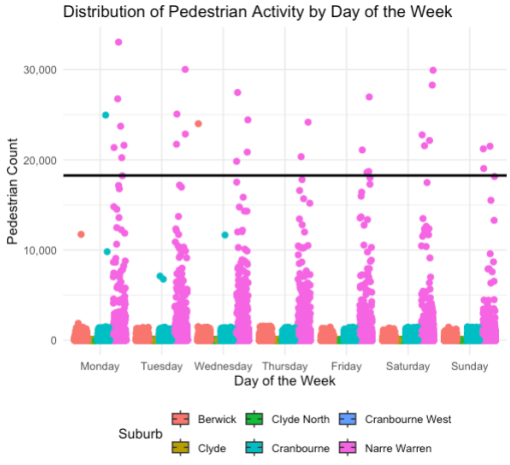

Figure 1. Outliers

The merged dataset showed no null values. Next, I identified outliers and extreme values. I needed to remove the extreme values first to properly assess the outliers, which are illustrated in figure 1. After that, I

removed the outliers, which are defined as 1,087 per hour (de Silva et al., 2023), though the number varies depending on the city road and day. Consequently, I decided to reduce the average by 30% and multiply it by 24 (to estimate the daily count), resulting in approximately 18,261.6, as shown in figure 2.

Additionally, I removed rows that were not significant to the analysis, such as data from sensors labelled "Under Maintenance" and "Offline," as well as those with a value type of "Time (Minutes)." These rows were excluded because they would introduce noise into the dataset, and my focus was on capturing the pedestrian count rather than the duration of time pedestrians spent under the sensors.

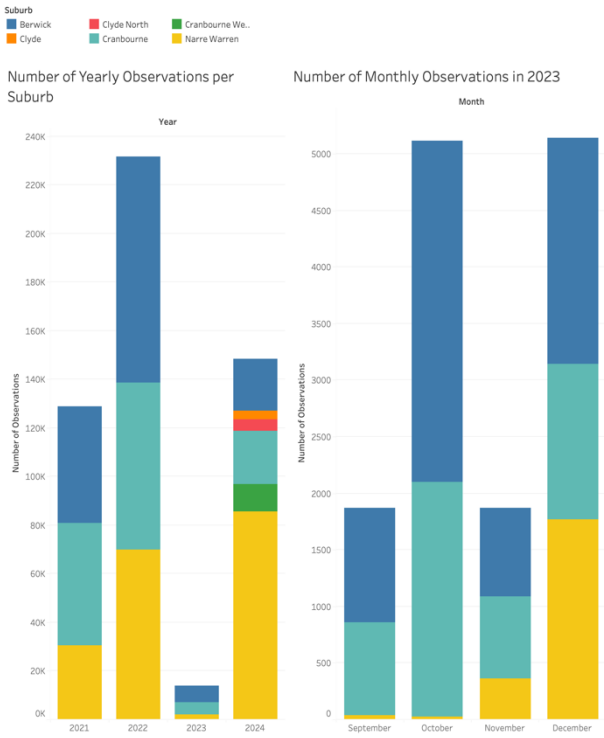

Figure 4. Number of Observations Yearly and Monthly per Suburb

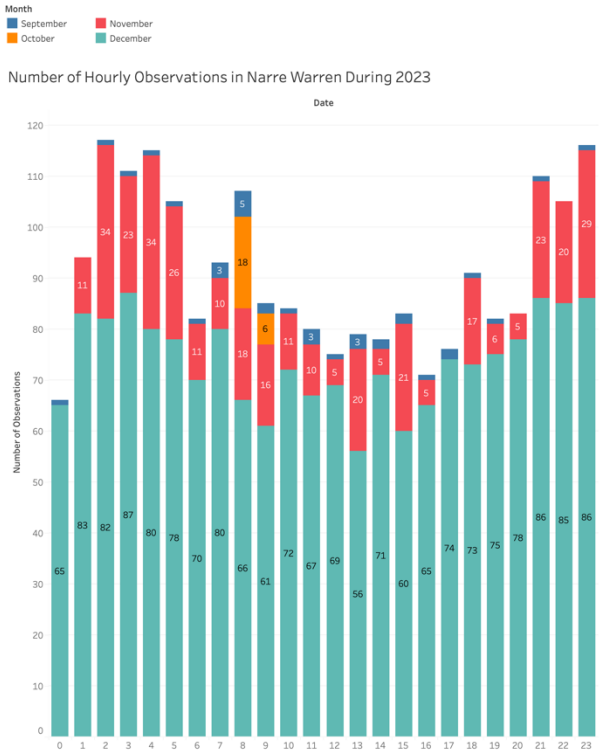

Figure 3. Number of Hourly Observation in Narre Warren in 2023

The yearly plot in figure 3 reveals that some suburbs do not have data available throughout the entire timeframe, leading us to focus only on the top three suburbs for further analysis. Additionally, there appears to be missing data for the year 2023 as shown in figure 4, which is evident from the inconsistencies observed in the yearly plot.

In the monthly plot for 2023, the data indicates that the number of observations in Narre Warren during September and October is a lot lower compared to other suburbs. This disparity suggests a potential gap or issue in data collection for those months.

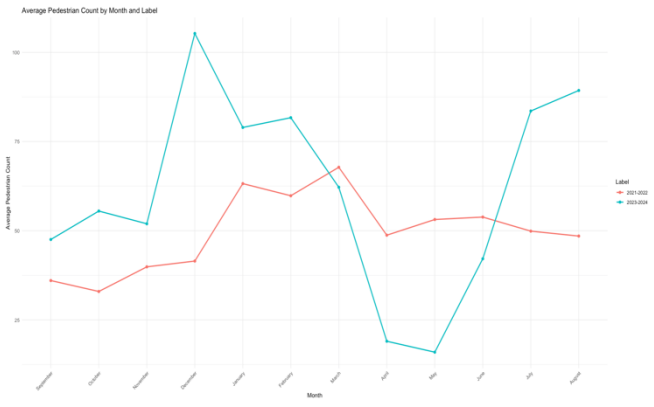

Figure 5. Monthly Average Pedestrian Trend

Furthermore, the hourly observations for Narre Warren during 2023 show a significant amount of missing data. The gaps in data collection are particularly noticeable and highlight the need for caution when interpreting the results for Narre Warren in this period.

The average pedestrian count by month and label reveals a quite similar pattern between the years 2021-2022 and 2023-2024 as shows by figure 5. Both periods exhibit comparable trends, with noticeable rises and falls in pedestrian counts across the months, with the only difference being in the extreme values.

Given this similarity in patterns, it may be feasible to address the missing data for Narre Warren in those years by utilising the entire dataset. One approach to achieve this is through the use of the linear interpolation method, which can help populate the missing data based on the observed trends and this method will be done in R.

For the park and reserve dataset, it is evident that there are null values present. The postcode variable has a particularly high percentage of null values at 79.2%, and since the postcode is not significant to the analysis, I decided to drop this variable from the dataset. Additionally, the area variable only contains two data points with null values, so I chose to remove those data points to maintain the dataset's integrity without significantly affecting the overall analysis.

Since there are numerous park and reserve locations, we will focus on selecting those closest to the pedestrian sensor locations, assuming that pedestrians are likely to wander around these areas. This approach allows the analysis to provide insights into pedestrian activity in relation to nearby parks and reserves. To achieve this, we will determine the minimum radius which is 0.75 km for each sensor location. The parks and reserves within this radius will be selected.

Data Exploration

Question 1

Suburb	Development_Site	count	Average	Median	SD	Minimum	Maximum	Q1	Q3
<chr>	<chr>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
Berwick	Berwick Village	133943	33.8	3	88.5	0	11737	0	27
Cranbourne	Cranbourne High Street	110346	41.1	3	155.	0	11661	0	14
Narre Warren	Bunjil Place	35920	120.	3	792.	0	18228	0	36
Narre Warren	Narre Warren Village	127888	81.1	3	198.	0	5526	0	45

Figure 6. Count Pedestrian Central Statistic

Figure 6 shows the pedestrian activity data in the City of Casey shows generally low foot traffic, with median counts around 3 across all development sites. However, the higher average counts indicate a positive skew, meaning occasional spikes in activity elevate the mean. Most of the time, pedestrian traffic remains low, as reflected by the first quartile (Q1) of 0 across sites.

There might be high variabilities exists in some locations, especially at Bunjil Place, where standard deviation and maximum counts are significantly higher. In contrast, sites like Narre Warren Village show more consistent pedestrian traffic, with lower variability and fewer spikes in activity.

The presence of low Q1 values (zero) across all sites implies that outside of peak hours, pedestrian traffic is minimal, especially in quieter areas. This suggests that peaks in pedestrian activity occur only during specific

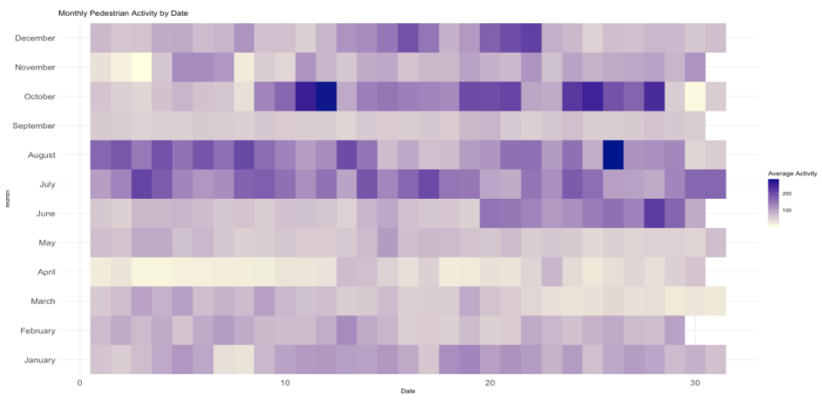

Figure 7. Distribution Average Count Pedestrian Monthly by Dates

hours, such as rush hours or event timings, and could vary based on the nature of the development site (e.g., commercial or residential).

Figure 7 shows the heatmap of monthly pedestrian activity with noticeable peaks in June, July, and August, likely due to the Winter Arts Festival held during these months, as suggested by the higher density of darker shades in this period (City of Casey, n.d.). This indicates increased foot traffic, especially around the middle of these months, correlating with local events. Conversely, April shows significantly lighter shades, representing lower pedestrian activity, possibly due to fewer events or colder weather at the start of autumn. Activity seems to fluctuate across the year, with a consistent rise in the middle of most months, possibly due to regular events or increased public activity.

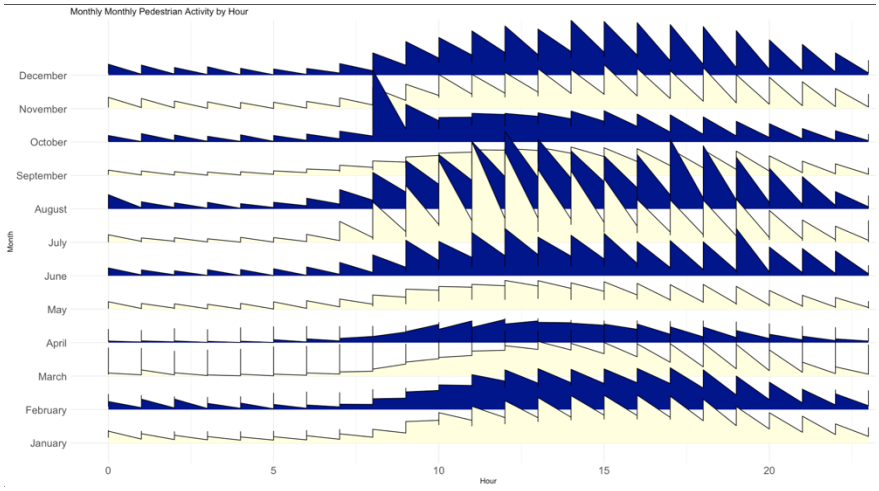

Figure 8. Hourly Distribution Count Pedestrian

Figure 8 shows clear trends in pedestrian activity across different months of the year. July and August stand out as the busiest months, with significantly higher foot traffic, likely due to events and festivals during the winter. Peak hours tend to fall between midday and early evening likely driven by lunch breaks and people commuting. In contrast, months like March, April, and February show much lower activity throughout the day, indicating fewer events or quieter periods. Early mornings and late evenings consistently have less activity across all months, emphasising the concentration of foot traffic during afternoon hours.

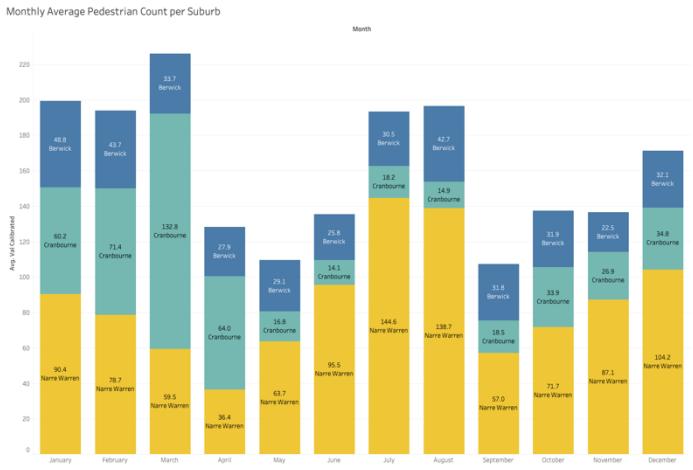

Figure 10. Monthly Average Pedestrian Count per Suburb

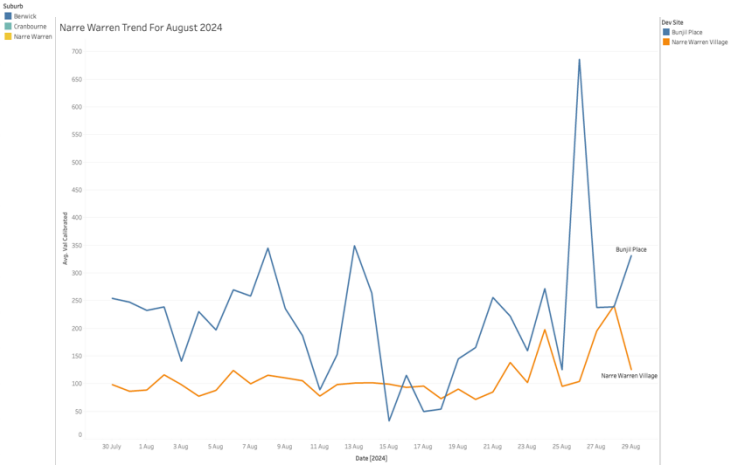

Figure 9. Pedestrian Count Trend for Narre Warren in 2024

Figure 10 shows that Narre Warren consistently experiences the highest pedestrian activity across multiple months, likely due to its strategic location as a transit point for people traveling from the CBD to other suburbs. Other suburbs like Berwick and Cranbourne also see significant traffic but fluctuate more. Figure 9 shows that within Narre Warren, Bunjil Place consistently draws higher foot traffic compared to Narre Warren

Village, with more frequent spikes, likely due to its role as a major community hub and event centre, offering arts, entertainment, and cultural attractions. This suggests that Bunjil Place is a key destination for locals, further reinforcing Narre Warren's importance as a central hub for both commuters and eventgoers in the City of Casey.

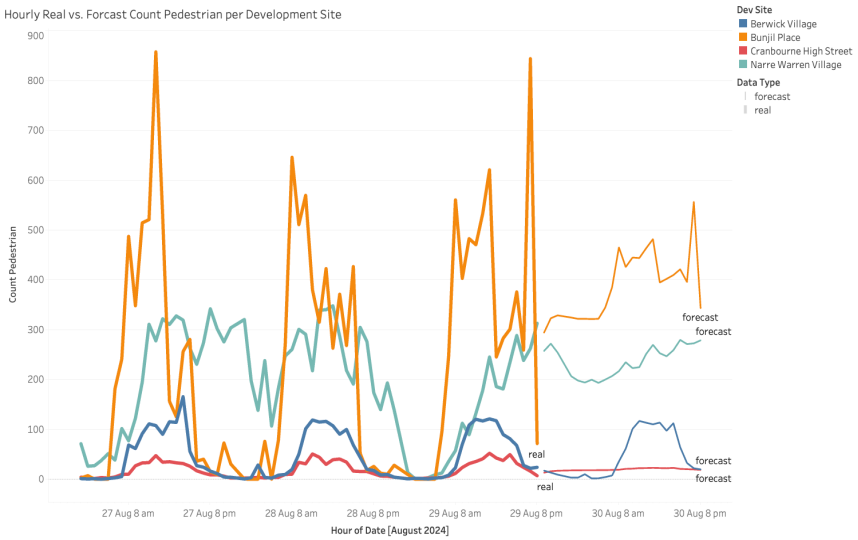

Figure 11. Hourly Forecast for Count Pedestrians per Development Site

Figure 11 shows the forecast data indicates that Bunjil Place consistently has the highest pedestrian activity across multiple days using ARIMA model (using the most current dataset which is 2024), with peak times occurring around midday and early evening. The forecast also predicts continued spikes, particularly during the afternoon on August 29 and 30. In contrast, Narre Warren Village shows steady but moderate activity, with forecasted peaks around mid-afternoon, indicating predictable foot traffic. Berwick Village and Cranbourne High Street exhibit lower forecasted pedestrian counts, with small increases during the same midday and evening periods. Overall, Bunjil Place is expected to continue seeing the highest foot traffic, making it the busiest development site in the City of Casey, especially during peak hours.

Pedestrian activity in the City of Casey peaks around midday and early evening, particularly in Narre Warren, which consistently shows the highest foot traffic. This is largely due to Narre Warren's strategic location as a transit point for people traveling from the CBD to other suburbs, as well as the presence of Bunjil Place, a key community hub that draws crowds for arts, entertainment, and events. Forecasted data also shows Bunjil Place experiencing peak activity in the afternoon, reinforcing its role as a primary destination. Other suburbs, such as Berwick and Cranbourne, have moderate activity levels, with peaks occurring at similar times but at lower volumes. July and August are the busiest months, driven by events like the Winter Arts Festival, while quieter months like March and April show less foot traffic. Across all suburbs, pedestrian movement is lowest in the early mornings and late evenings, with the busiest times concentrated in the afternoons, particularly around midday and early evening, likely driven by work commutes, lunch breaks, and events.

The City of Casey should leverage midday and early evening peaks by organising events and activities during these times, particularly in high-traffic areas like Narre Warren and Bunjil Place. Enhancing public transport links and amenities like food stalls and seating could further support and increase foot traffic during these peak hours.

Question 2

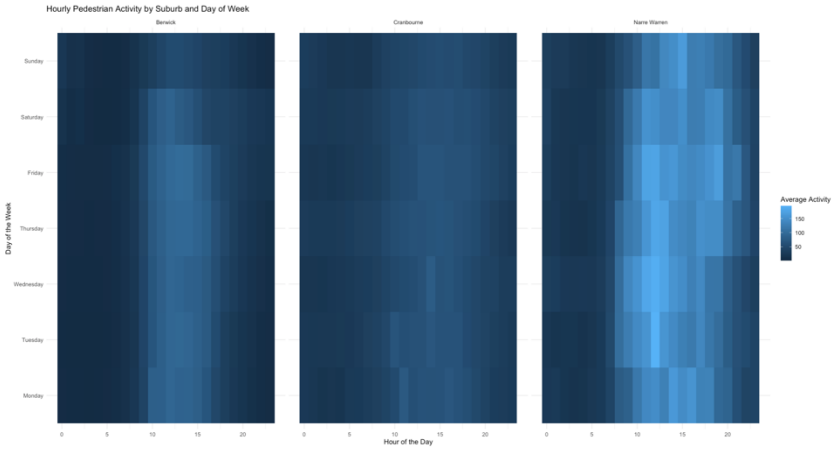

Figure 12. Hourly Pedestrian Activity by Suburb and Day of Week

Figure 12 shows pedestrian activity data shows clear trends across different days of the week. Sunday generally has lower foot traffic compared to other days, indicating it is a quieter day in most areas. Peak hours occur consistently from midday to early evening, likely driven by lunch breaks and people commuting home from work, particularly evident across all days. Friday stands out with notably higher activity later in the day, especially at Narre Warren, which also shows overall higher pedestrian traffic compared to other suburbs, making it a busier area. Morning hours are generally quiet, indicating that most activity is concentrated in the afternoon and evening.

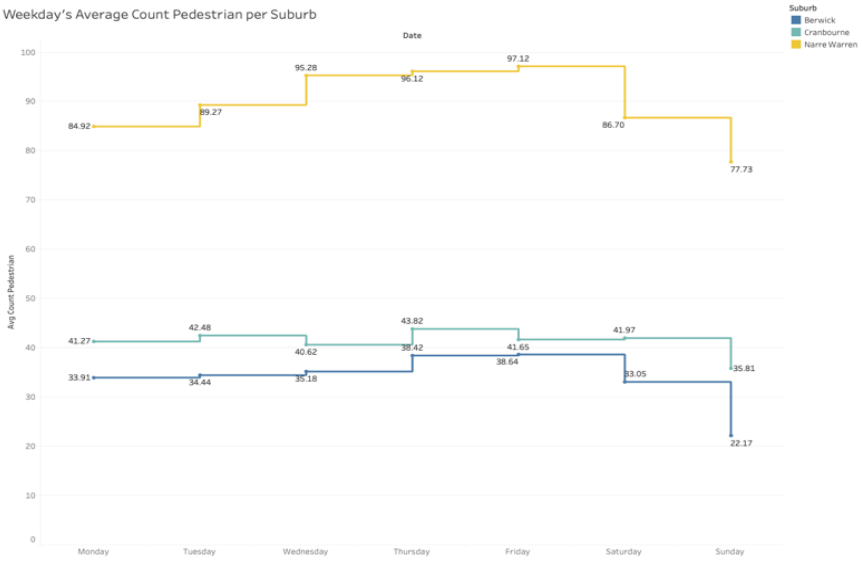

Figure 13. Weekday's Average Count Pedestrian per Subrub

Figure 13 shows pedestrian activity trends across different days of the week for the suburbs of Narre Warren, Cranbourne, and Berwick. Narre Warren consistently has the highest pedestrian counts, with activity peaking on Friday and gradually declining over the weekend. Cranbourne shows relatively stable traffic throughout the week, with only slight variations between weekdays and a small drop over the weekend. Berwick has the lowest pedestrian traffic, maintaining a steady trend on weekdays but experiencing a significant decrease on Sunday. Overall, pedestrian activity tends to be higher during the week, particularly on Fridays, and lower during weekends across all suburbs.

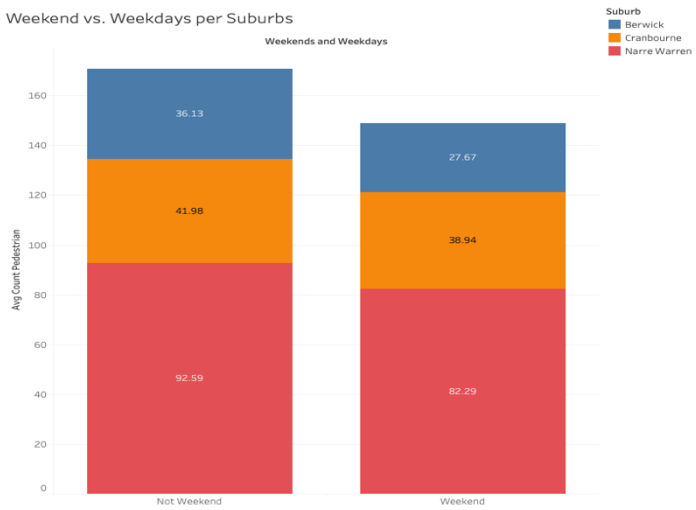

Figure 15. Weekend vs. Weekdays Average Count Pedestrian

Figure 15 shows clear differences in pedestrian activity between weekdays and weekends across the suburbs. Overall, pedestrian traffic is higher during weekdays compared to weekends in all suburbs. Narre Warren has the highest activity on both weekdays and weekends, but the weekday count is slightly higher. Cranbourne sees a moderate difference, with slightly more activity during weekdays than weekends. Berwick shows the lowest pedestrian traffic, with fewer people on both weekdays and weekends, and a noticeable drop-in activity on weekends. This suggests that weekdays are generally busier across all suburbs.

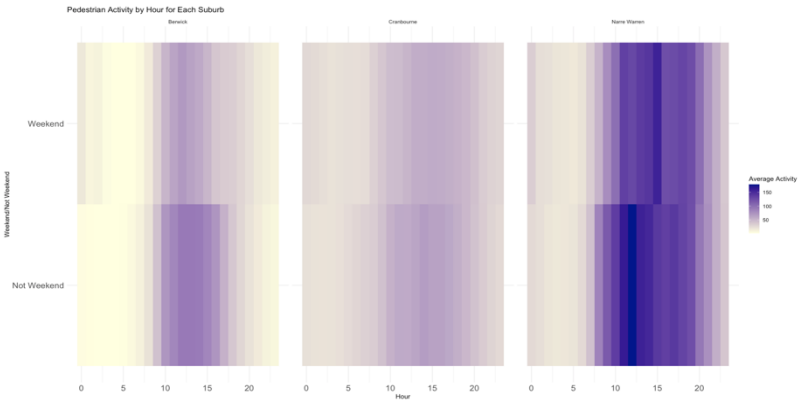

Figure 14. Pedestrian Activity by Hour for Each Suburb

Figure 14 illustrates pedestrian activity by hour for three suburbs—Berwick, Cranbourne, and Narre Warren—on both weekends and weekdays. Narre Warren consistently shows the highest activity, particularly during weekday afternoons between 10 AM and 6 PM, with notable peaks around midday. This pattern reflects Narre Warren's role as a central hub with high foot traffic. Cranbourne has moderate activity, with peak hours also around midday on both weekdays and weekends, although the intensity is lower than in Narre Warren. Berwick shows the least activity, with pedestrian counts concentrated in the afternoon but noticeably lower on weekends. Overall, weekday activity dominates in all three suburbs, with peak pedestrian movement occurring between mid-morning and late afternoon.

Pedestrian activity in the City of Casey differs between weekdays and weekends, with higher foot traffic during weekdays, especially in Narre Warren, which peaks on Fridays and gradually declines over the weekend. Cranbourne shows steady traffic with a slight drop on weekends, while Berwick sees the lowest activity, with a significant decline on Sundays. However, the hourly trend remains similar across both weekdays and weekends, with peak activity occurring between midday and early evening (10 AM to 6 PM). This suggests that while total foot traffic may decrease on weekends, the busy hours remain consistent across all days. Overall, weekdays are busier, especially Fridays, while weekends experience lower pedestrian traffic across all suburbs.

The City of Casey should focus on increasing weekend pedestrian traffic by organising events, festivals, or community activities, especially on Saturdays and Sundays. Enhancing accessibility, promoting recreational

facilities, and improving amenities like seating and lighting can also encourage more foot traffic during the quieter weekend periods across all suburbs.

Question 3

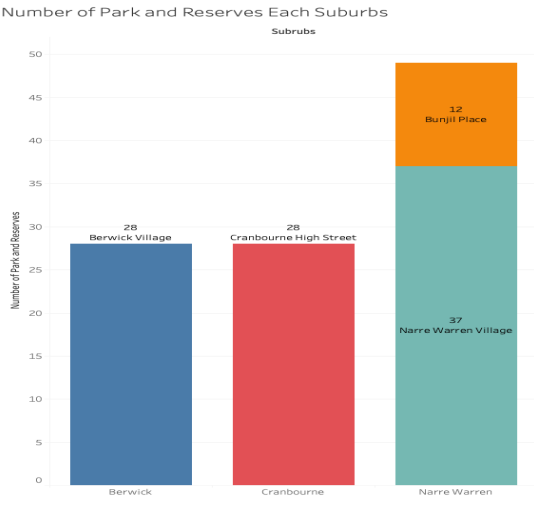

Figure 17. Number of Park and Reserves Each Suburb

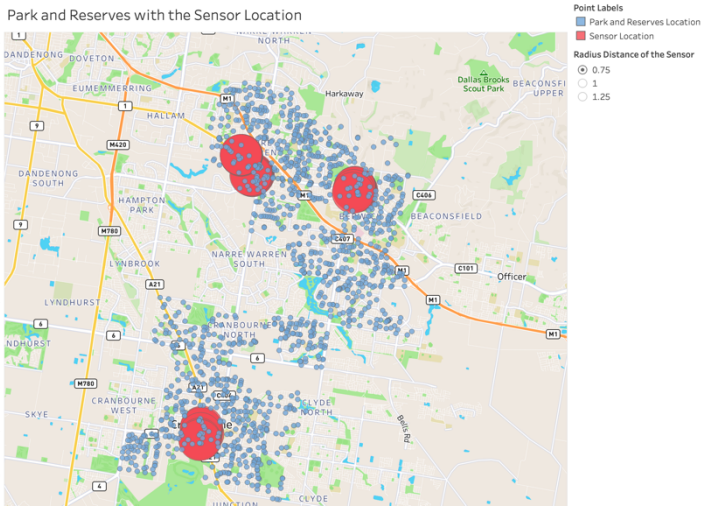

Figure 16. Park and Reserves with Sensor Location

Figure 17 shows that Narre Warren has the highest number of parks and reserves (49), including Bunjil Place, which may attract more visitors due to its role as a community hub. Berwick and Cranbourne both have 28 parks, indicating similar outdoor space availability.

Figure 16 shows the distribution of parks and reserves in the City of Casey with sensor locations marked for tracking pedestrian activity. The red circles represent areas with a higher density of sensors, particularly around Narre Warren, Cranbourne, and Berwick. These regions likely experience more foot traffic due to their proximity to key community hubs, like Bunjil Place in Narre Warren. The widespread distribution of parks suggests that residents across various suburbs have access to outdoor spaces, but the clustering of sensors in key areas indicates that pedestrian monitoring is focused on high-traffic zones.

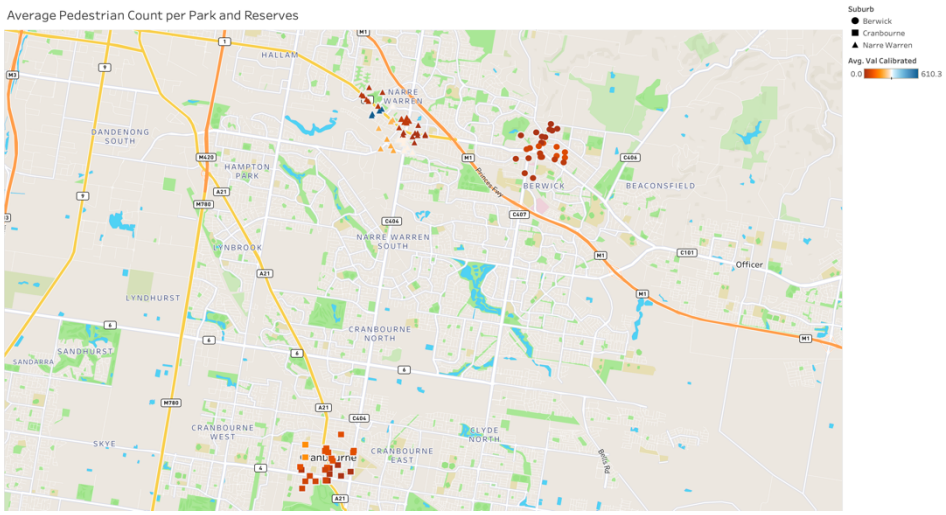

Figure 18. Average Pedestrian Count per Park and Reserves

Figure 18 highlights areas with higher pedestrian activity, particularly concentrated around Narre Warren and Cranbourne. The Narre Warren area, near Bunjil Place, shows the highest calibrated pedestrian counts, indicating it is a key hub for foot traffic. To enhance the pedestrian experience and safety in these high-traffic zones, the City of Casey should consider expanding amenities such as more shaded walkways, seating areas, and improved lighting. Additionally, introducing more public events and better connectivity between parks and reserves could further increase foot traffic and encourage greater community engagement across these areas.

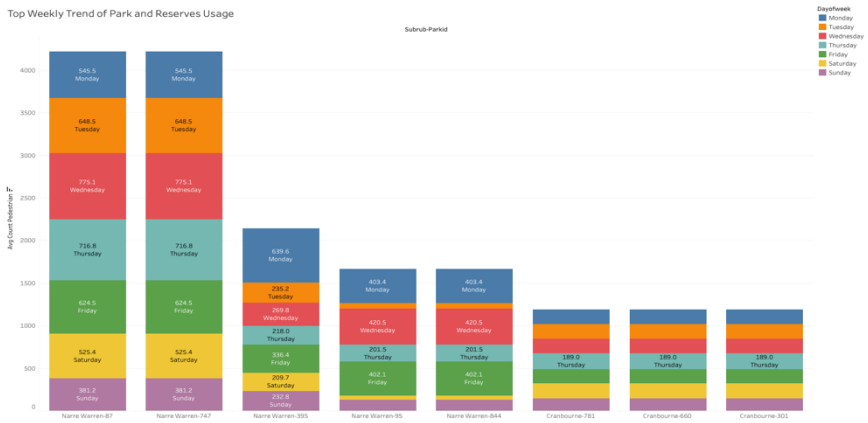

Figure 19. Top Weekly Trend of Park and Reserves Usage

Figure 19 highlights the weekly trends in park and reserve usage across different locations in Narre Warren and Cranbourne. Wednesday and Thursday consistently show the highest pedestrian counts, particularly in Narre Warren, with peaks over 700 pedestrians. Weekends experience lower usage, particularly on Sundays, with minimal activity compared to weekdays. To better balance park usage, the City of Casey could organise weekend events or community activities to encourage higher foot traffic on Saturdays and Sundays. Additionally, improving accessibility and promoting recreational opportunities could help distribute usage more evenly throughout the week.

Narre Warren has the highest number of parks and reserves (49), including Bunjil Place, which likely draws more visitors due to its role as a community hub. Cranbourne and Berwick have similar park availability, each with 28 parks. Pedestrian activity is concentrated in Narre Warren and Cranbourne, particularly around Bunjil Place, where foot traffic is highest. To improve pedestrian experience, the City of Casey could expand amenities like shaded walkways and seating, and organise events to boost weekend usage, especially on Saturdays and Sundays, when activity is lower. Wednesday and Thursday show the highest weekday pedestrian counts, suggesting regular activity during midweek. Increasing weekend events and recreational activities could help balance park usage throughout the week.

Conclusion

The data exploration highlights peak pedestrian activity occurring around midday and early evening, particularly in Narre Warren and Bunjil Place. Visualisations, such as time series plots, show consistent traffic during these times, driven by work commutes and events. These insights suggest opportunities to boost engagement during peak hours by organising events or improving public amenities in high-traffic areas.
Differences between weekdays and weekends are also evident, with higher foot traffic observed on weekdays, especially Fridays. The visualisations indicate that weekends, particularly Sundays, see lower activity across all suburbs. This suggests a need for weekend events or recreational activities to balance usage across the week.

Pedestrian activity within parks and reserves is concentrated near community hubs like Bunjil Place, as shown in spatial maps. July and August exhibit higher foot traffic, likely driven by seasonal events like the Winter Arts Festival. These trends suggest that the City of Casey could enhance pedestrian experience by improving park amenities and organising more events during quieter months to drive consistent foot traffic year-round.

Reflection

The exploration revealed valuable insights into pedestrian activity patterns in the City of Casey, particularly highlighting peak times, differences between weekdays and weekends, and trends in park utilization. However, there were challenges related to data gaps, particularly for Narre Warren, which required interpolation methods to address missing data. Another limitation was the lack of continuous data across the entire year for some suburbs, which affected the ability to draw consistent conclusions about trends over time.

To improve the analysis, more robust and continuous data collection, particularly in critical locations like Narre Warren, would help minimise gaps and provide a clearer picture. Additionally, expanding the analysis to cover more suburbs and longer time periods would yield a more comprehensive understanding of pedestrian behavior. Implementing real-time data monitoring and adjusting sensor placements to cover underrepresented areas could enhance the quality of insights derived from the data.

Reference

Australian Bureau of Statistics. (2020). Regional population growth, Australia, 2019-20. Retrieved from https://www.abs.gov.au/statistics/people/population/regional-population/2019-20#victoria.

de Silva, A. J., Yanotti, M., Sinclair, S., Angelopoulos, S., & Navon, Y. (2023, April 26). Mapping local economic recovery paths using pedestrian counts: A City of Melbourne case study. School of Economics, Finance and Marketing, RMIT University; Tasmanian School of Business and Economics, University of Tasmania.

City of Casey. (n.d.). Winter arts festival. https://www.casey.vic.gov.au/winter-arts-festival (Accessed September 9, 2024)

Koohsari, M. J., Kaczynski, A. T., Mcormack, G. R., & Sugiyama, T. (2014). Using Space Syntax to Assess the Built Environment for Physical Activity: Applications to Research on Parks and Public Open Spaces. Leisure Sciences, 36(2), 206–216. https://doi.org/10.1080/01490400.2013.856722

Sharmin, S., & Kamruzzaman, M. (2018). Meta-analysis of the relationships between space syntax measures and pedestrian movement. Transport Reviews, 38(4), 524–550. https://doi.org/10.1080/01441647.2017.1365101

Vargas-Hernández, J. G., & Zdunek-Wielgołaska, J. (2021). Urban green infrastructure as a tool for controlling the resilience of urban sprawl. Environment, Development and Sustainability, 23(2), 1335–1354. https://doi.org/10.1007/s10668-020-00623-2