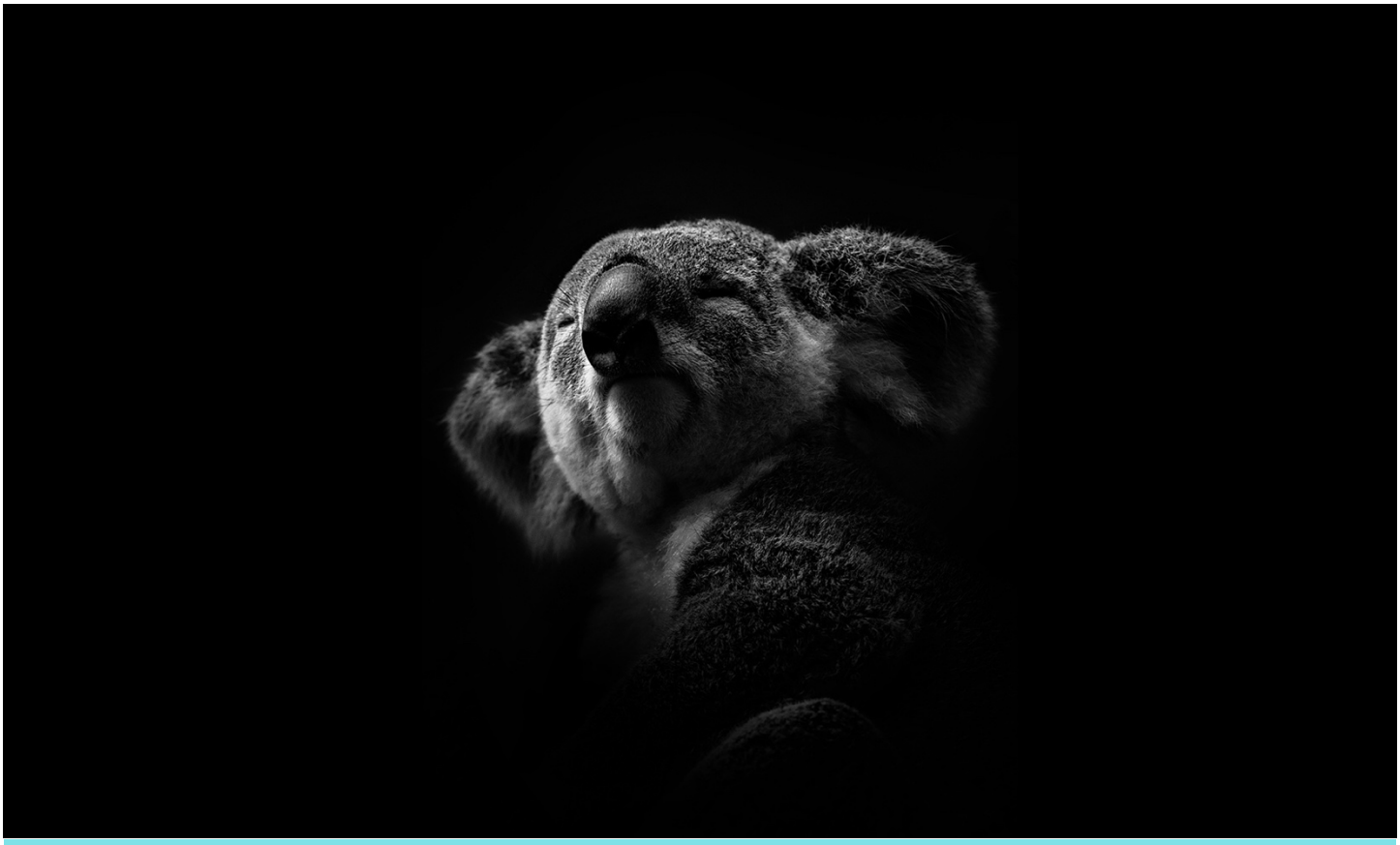


It's surely our responsibility to do everything within our power to create a planet that provides a home not just for us, but for all life on Earth.

Sir David Attenborough

Title: *An Analysis of Koala Hospital Dataset Using Tableau and Python*
Author: *Yasaman Mohammadi*

Autumn 2024

1. Introduction

Koalas, one of Australia's most iconic and beloved marsupials, face a precarious future. In 2012, the Australian government officially declared koalas as vulnerable and listed them as a threatened species, highlighting the urgent need for concerted conservation efforts (Quigley & Timms, 2020). Against this backdrop, the importance of exploratory data analysis (EDA) on koala populations cannot be overstated, mainly when focused on identifying the reasons for their death and injury. Such an analysis is invaluable for organizations dedicated to wildlife preservation, as it provides critical insights that can shape effective conservation strategies and inform policy decisions. By dissecting the factors contributing to koala mortality and injuries, the research aids in developing targeted interventions that contribute to the broader scientific understanding of koala ecology and support the mobilization of resources and public awareness towards their protection.

RESEARCH QUESTIONS

1. *What are the primary causes of mortality and injury in koala populations?*
2. *What is the role of disease in the decline of the koala population?*
3. *How do human-induced factors contribute to koalas' mortality and injury rates compared to natural causes?*

2. Dataset Description

This dataset originates from the Department of Environment and Science's Koala Base, focusing on koala records from 2016 to 2023. This specific dataset was chosen for several reasons:

1. **Relevance to Conservation**: Koalas are iconic Australian marsupials and are considered vulnerable due to various threats. This dataset provides crucial insights into the factors affecting koala populations and their interactions with human environments.
2. **Comprehensiveness**: With over 10310 records, the dataset offers a substantial volume of data to analyse trends, identify patterns, and make informed decisions for koala conservation.
3. **Richness in Features**: The dataset has 37 features allows for a multifaceted analysis.

Table 1- Dataset Description

	Category	Features	Explanation
1	***Identification and Tracking Features***	Record No, Koala Name, Call Date Time	These features are fundamental for organizing the dataset, tracking individual koalas, and analysing temporal trends in koala encounters or incidents.
2	***Geolocation Features***	LAT, LNG, Koala Found Address, Koala LGA	These spatial coordinates play a vital role in geographical analysis, pinpointing areas where koalas face a heightened risk of injury or fatality. They facilitate spatial assessments that inform efforts in habitat preservation and urban development planning.
3	***Demographic and Health Features***	Adult Size, Gender, Sick, Injured, Conjunctivitis, Cystitis, Wasted, Sick Other	These features provide insights into the health and demographic profile of the koala population. Health-related attributes are pivotal for assessing disease prevalence, injury causes, and overall health status, guiding healthcare and intervention strategies for koalas.
4	***Incident and Outcome Features***	Adult Situation, Adult Situation Other, Young Present, Adult Fate, Young Fate Other, Dead, Orphaned	These features detail the circumstances of each koala encounter and their outcomes, offering a comprehensive view of the challenges koalas face. Analysing these can help identify common threats, such as predation, accidents, or environmental stressors, and the effectiveness of intervention measures.
5	***Threat Identification Features***	Caused By Dog, Under Threat, Vehicle Hit, Road Speed Limit	These indicators highlight specific threats to koalas, such as traffic collisions and dog attacks. Understanding these threats is crucial for developing targeted mitigation strategies, like wildlife crossings, speed limit enforcement, or public awareness campaigns to reduce dog attacks.

6	***Post-Rescue Features***	Release Date, Release Location, Release Suburb, Release Post Code, Release LAT, Release LNG	Information on the post-rescue release of koalas is vital for assessing the success of rehabilitation efforts and the viability of release sites. Monitoring post-release survival and habitat suitability helps refine rescue and rehabilitation protocols and conservation strategies. 

3. Tools Used

2.1 Python

Python was chosen for its adaptability and the depth of analysis it allows, particularly with its extensive data science libraries. Libraries such as Pandas for data manipulation and Matplotlib for visualization were extensively utilized. The aim of using Python for this project was twofold: firstly, to efficiently perform exploratory data analysis on the koala dataset using existing skills; and secondly, to delve deeper into more advanced visualization techniques, such as interactive plots with Plotly, which can provide more nuanced insights into complex datasets.

3.2 Tableau

Tableau was selected for its user-friendly interface and visual analytics capabilities. Despite limited prior experience, the adoption of Tableau was motivated by its speed and ease in creating data narratives. The tool's functionalities, particularly in rapidly generating dashboards and geographical visualizations, were leveraged for their suitability to the spatial aspects of the koala dataset. Tableau's intuitive interface enabled quicker data exploration, reducing the need for extensive coding, and enhancing immediate data comprehension.

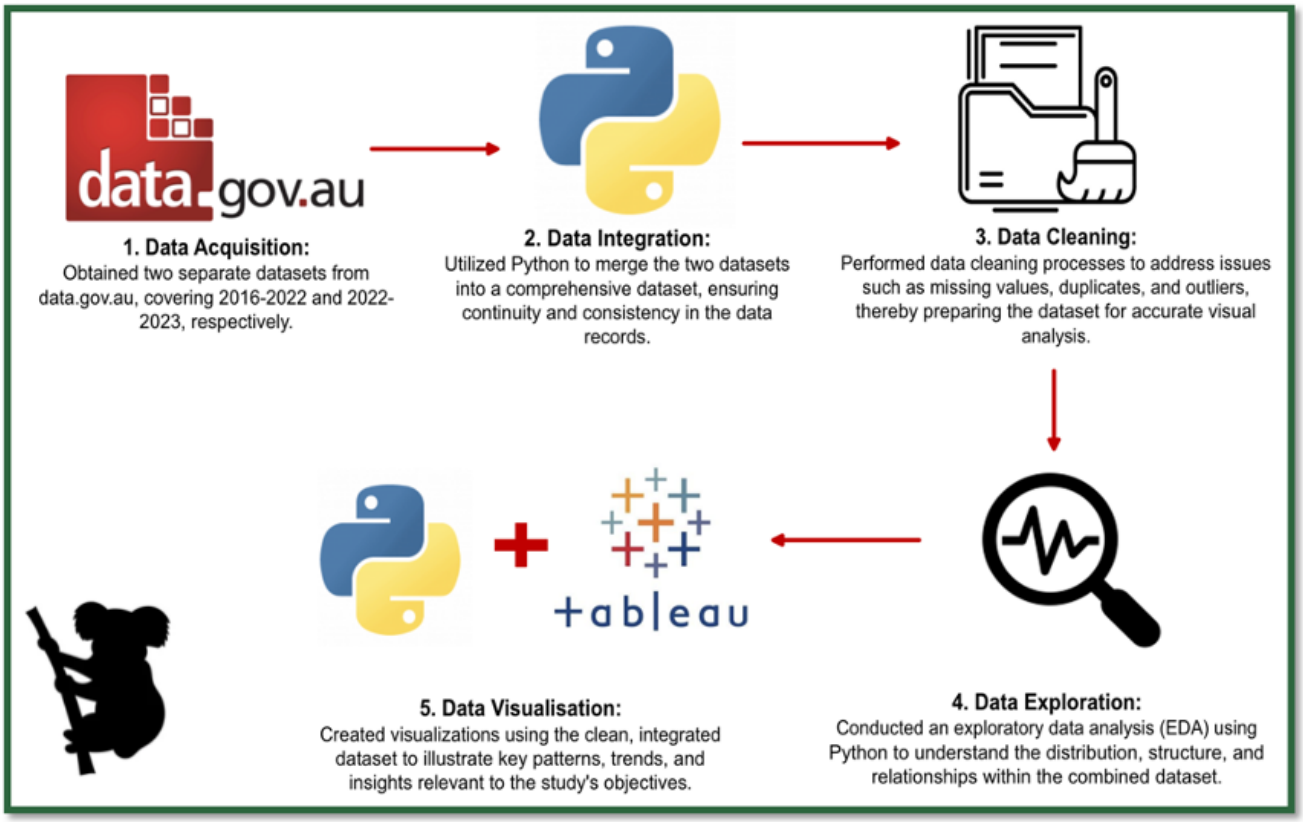


Figure 1- Workflow

4. Visualization and Analysis

4.1 Visualization with Python

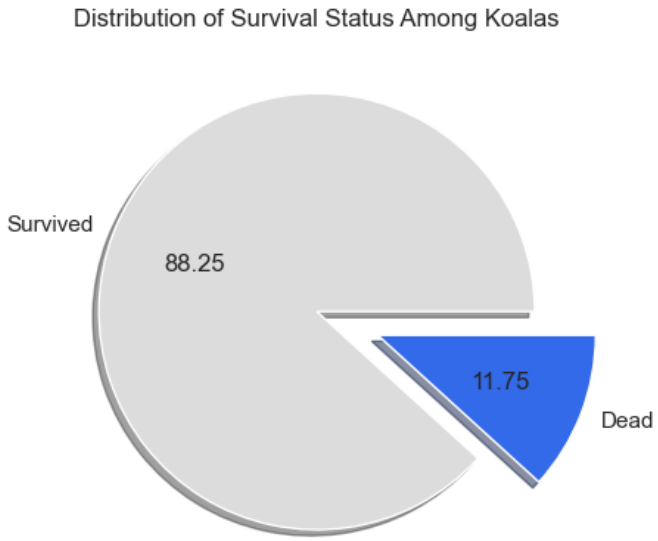
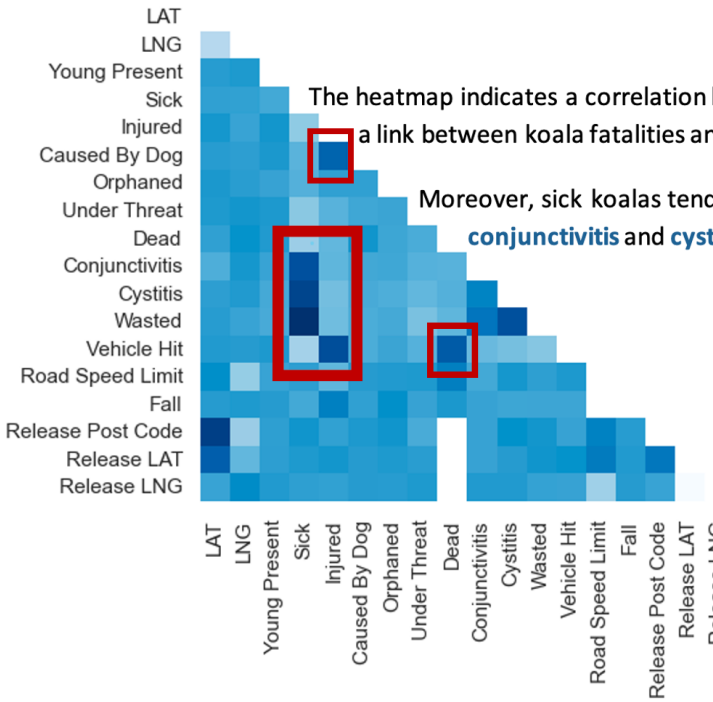


Figure 2| Distribution of Survival Status Among Koalas

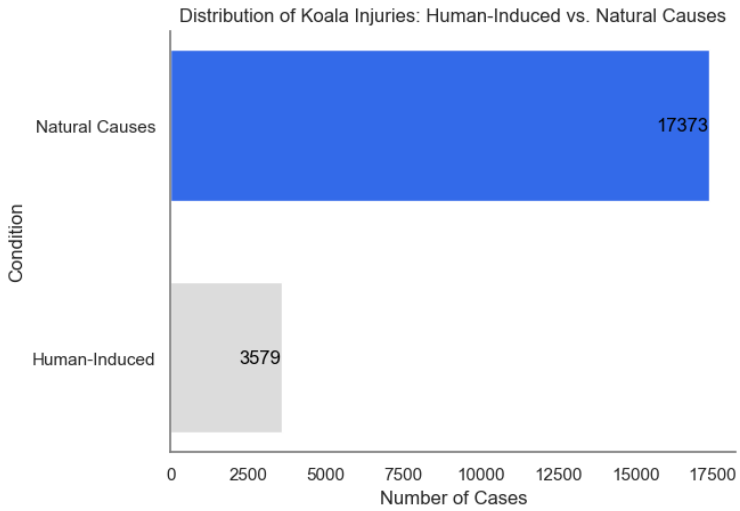
The dataset predominantly comprises data about **survived koalas**.



The heatmap indicates a correlation between koalas being injured and **dog attacks**, as well as a link between koala fatalities and **vehicle collision**.

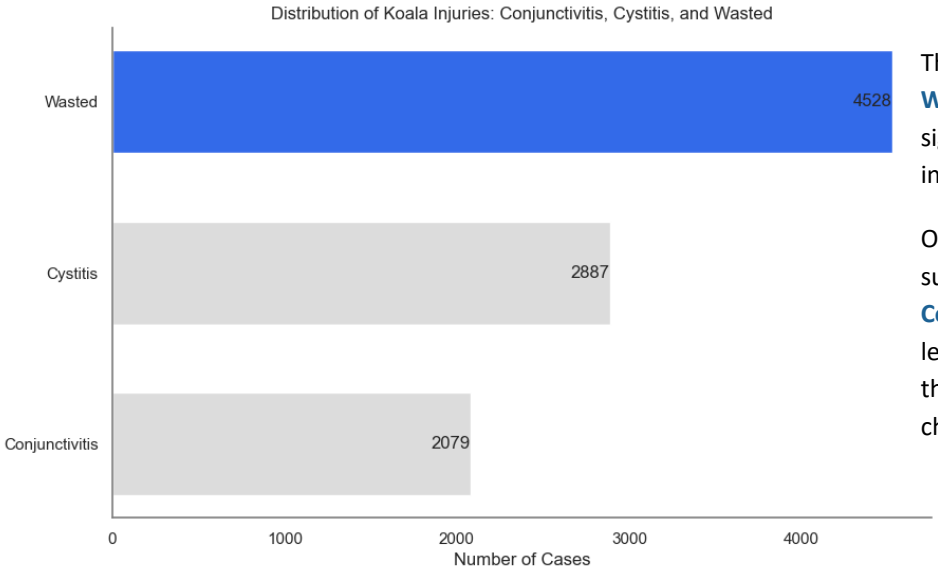
Moreover, sick koalas tend to be in a **wasted** condition, with specific ailments like **conjunctivitis** and **cystitis** being common factors.

Figure 3|A correlation heatmap between all features of the dataset.



The bar chart illustrates that koala injuries and mortality are primarily attributed to **Natural Causes**, such as disease and falls, which significantly outweigh **Human-Induced** factors like vehicle collisions, dog attacks, and environmental hazards.

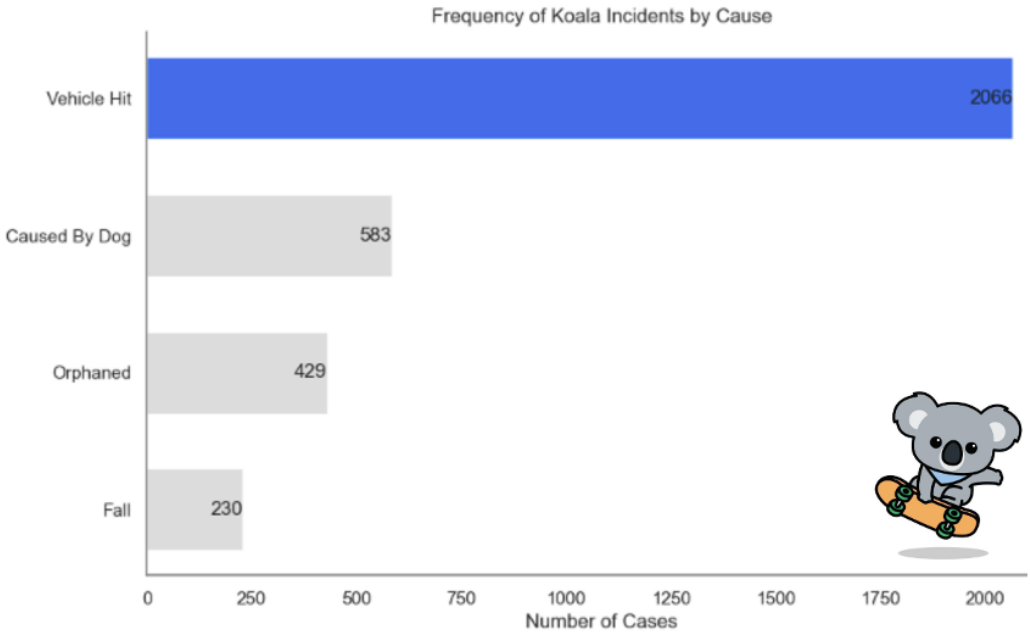
Figure 4|Distribution of Koala Injuries: Human-Induced vs. Natural Causes



The data highlights that **Wasted** represents the most significant proportion of injuries observed.

On the other hand, ailments such as **Cystitis** and **Conjunctivitis** are relatively less prevalent, indicated by their lower frequency in the chart.

Figure 5|Distribution of Koala Injuries: Conjunctivitis, Cystitis, and Wasted



Data indicates that **vehicle hits** are the predominant cause of koala incidents.

Figure 6|Frequency of Koala Incidents by Cause

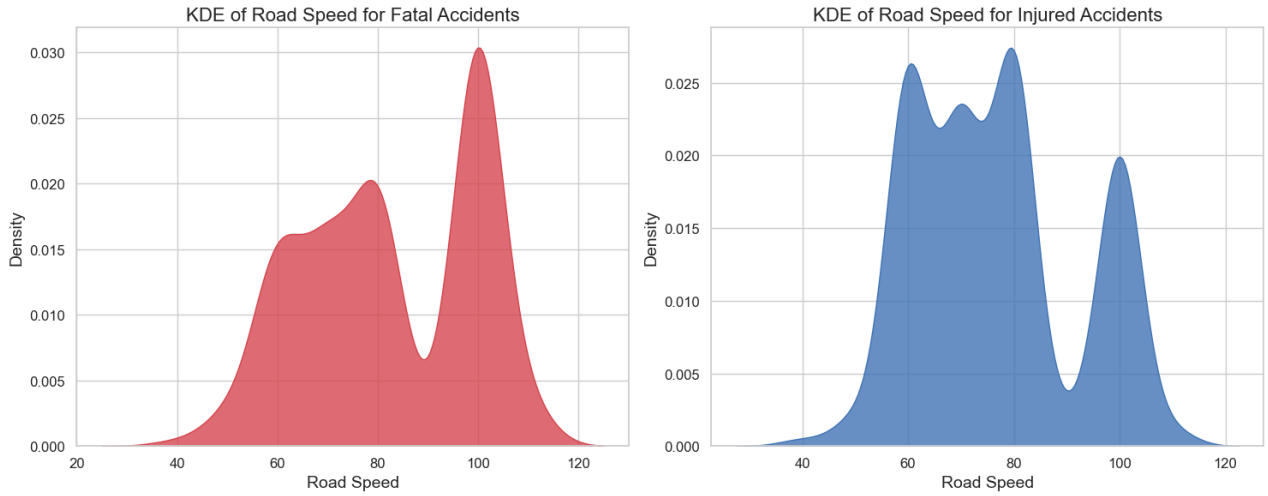


Figure 7| KDE plots of road speed

Fatal accidents are most likely to happen at two specific road speed ranges. The first peak occurs between 40-50 km/h, while the second peak, more pronounced, is around 100 km/h.
Similarly, most **injuries** occur at two distinct speed ranges on the road. Peaks are observed around 60 km/h and again near 100 km/h.
This suggests that reducing vehicle speeds in these zones could be critical for koala conservation, as lower speeds might decrease the severity of accidents.

4.2 Visualization with Tableau

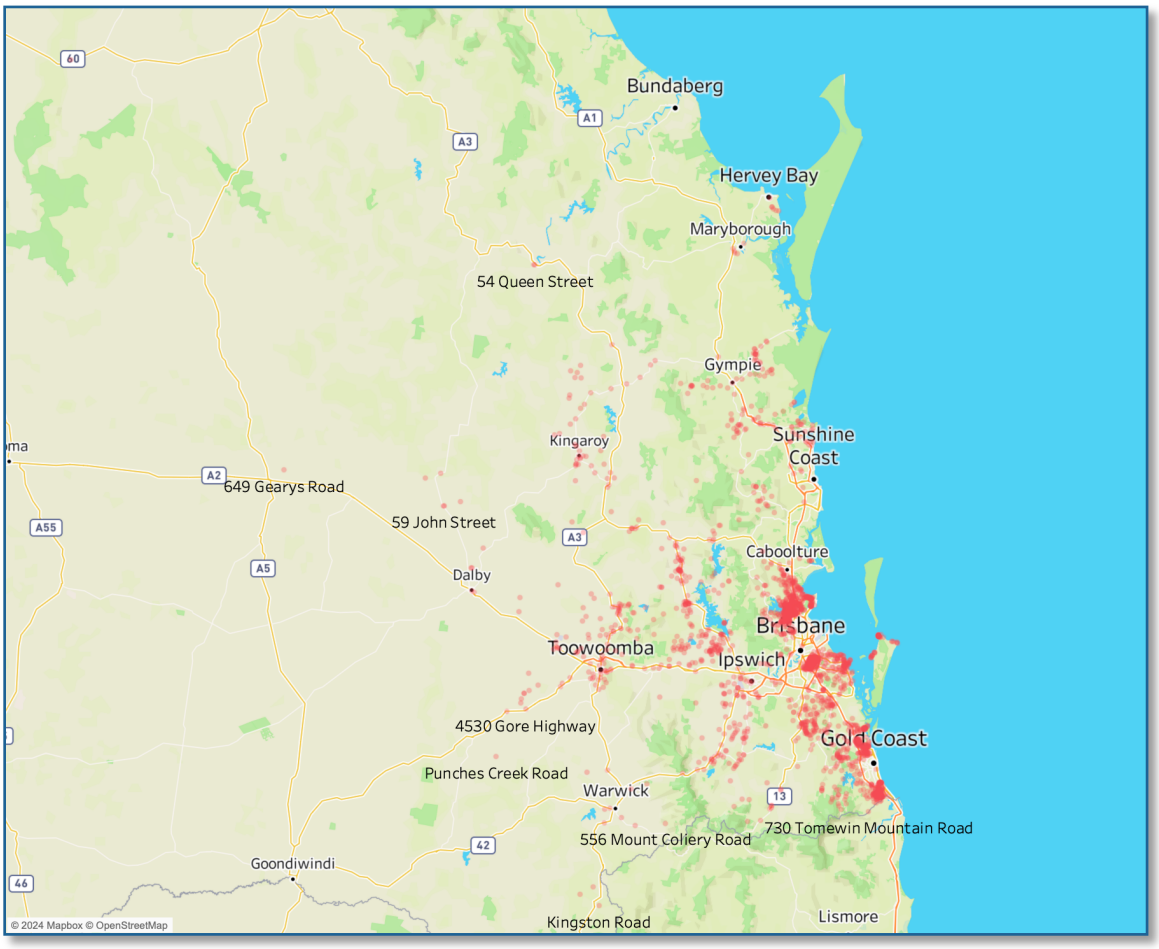


Figure 8 |Map of Injured koalas' Locations,

Based on the preceding map, it is evident that most koala injuries occur in **Gold Coast** and **Brisbane**.

Figure 9 | Word cloud of top 10 addresses where the koala was found

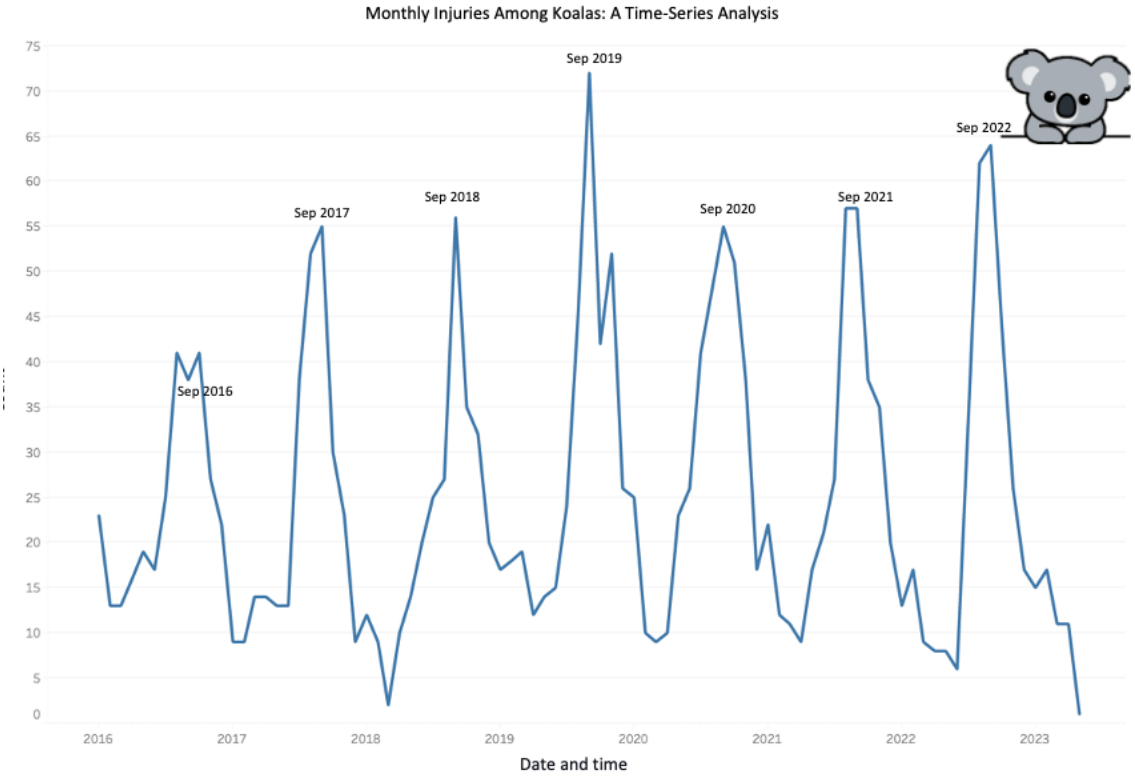


Figure 10 |Monthly Injuries Among Koalas: A Time-Series Analysis

The chart clearly shows that injury incidents among koalas peak predominantly in September, suggesting a seasonal surge in these months.

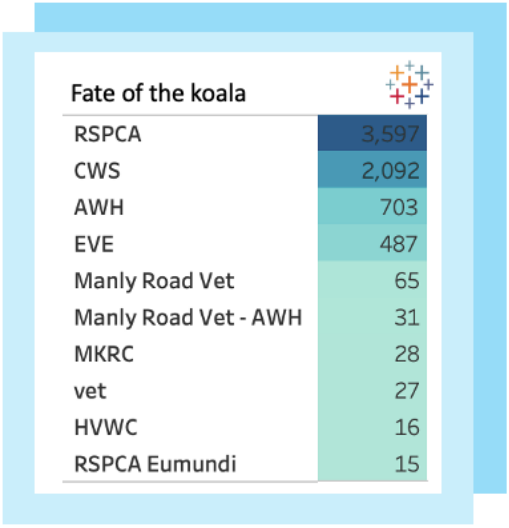


Figure 11| Top 10 Outcomes for Adult Koalas: Details of Their Fate

The chart distinctly indicates that most injured koalas are treated at leading facilities such as the **RSPCA** (Royal Society for the Prevention of Cruelty to Animals), **CWS** (Currumbin Wildlife Sanctuary), and **AWH** (Australian Wildlife Hospital).

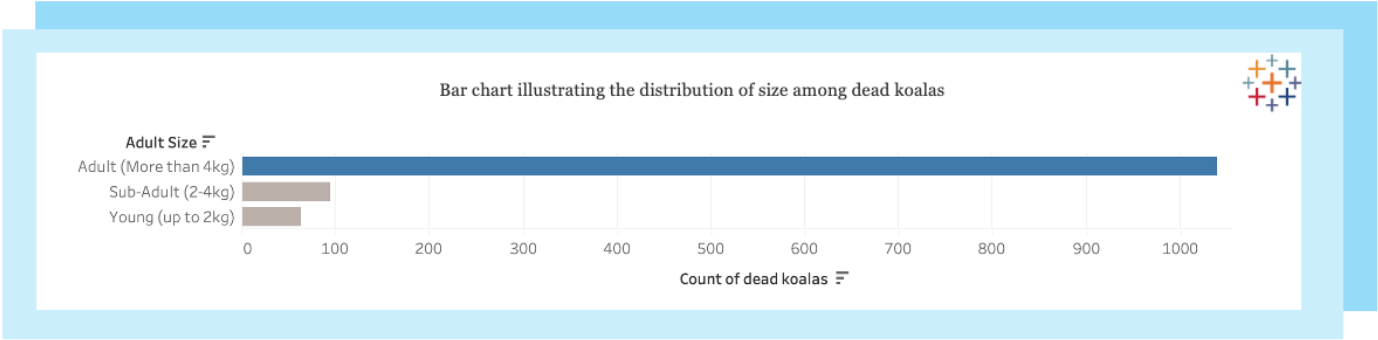


Figure 12 |Bar chart illustrating the distribution of size among dead koalas

Adult koalas, **weighing more than 4 kg**, exhibit the highest mortality rate, whereas their smaller counterparts experience lower death rate.

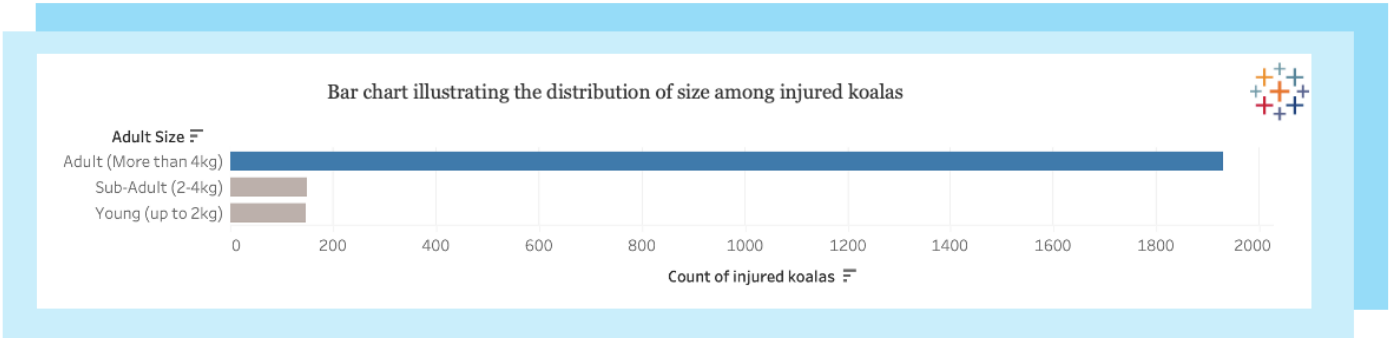


Figure 13 |bar chart illustrating the distribution of size among injured koalas

Adult koalas, **weighing more than 4 kg**, exhibit the highest injury rate, while their smaller counterparts experience a lower rate of injury.

5. Conclusion

Koalas, an endangered species, represent an iconic aspect of Australia's wildlife. Despite data showing that only 11.75% of koalas were found dead **(Figure 2)**, their conservation remains a critical issue. Various factors contribute to the mortality and injuries of koalas, including both natural causes like disease and falls, as well as human-induced causes such as vehicle collisions and dog attacks **(Figure 3)**. Remarkably, natural causes account for 82% of these incidents **(Figure 4)**, with "wasted" being the predominant disease affecting koalas (Figure 5). However, vehicle collisions stand out as a significant human-induced threat **(Figure 6)**, with speeds of 100 km/h being most lethal and 60 km/h frequently causing injuries **(Figure 7)**.

Geographical analysis indicates that the Gold Coast and Brisbane are hotspots for koala injuries **(Figure 8)**, with roads like Brisbane Valley Highway, M1, Boundary Road, and Youngs Crossing being common locations for these incidents **(Figure 9)**. There's a noticeable increase in injuries in September **(Figure 10)**, aligning with the beginning of the koala breeding season, which extends from September to July (Watters, F., Ramsey, D., Molsher, R., & Cassey, P. (2021)). This period also marks the onset of spring, a peak time for tourism, potentially exacerbating the risk to koalas. Facilities like the RSPCA, CWS, and AWH play a pivotal role in the care of injured koalas **(Figure 11)**. Furthermore, data suggests that adult koalas weighing over 4 kg are at a higher risk of injury or death **(Figures 12 and 13)**, underlining the importance of targeted conservation efforts to safeguard these vulnerable marsupials.

In conclusion, using Python and Tableau to analyse the koala hospital dataset yielded valuable insights. Python's comprehensive data manipulation capabilities enabled a deep dive into the factors affecting koala welfare. Tableau's visualisation strength was beneficial in the spatial aspects of the data.

The analysis revealed critical areas and times for koala incidents, offering actionable information for conservation efforts. These tools provided a clearer picture of the current state of koala welfare and equipped us with the means to communicate these findings effectively.

The learnings from this project have far-reaching implications, demonstrating the power of data analysis in guiding strategic conservation initiatives. This approach could serve as a model for data-driven decision-making in wildlife preservation, emphasising the need for specialised focus in vulnerable regions and times.

6. References

1. Data.gov.au. (n.d.). Koala.
   Retrieved from https://data.gov.au/dataset/ds-qld-9e3147cc-df40-4359-b603-845673401568/details?q=koala

2. Quigley, B. L., & Timms, P. (2020). Helping koalas battle disease – Recent advances in Chlamydia and koala retrovirus (KoRV) disease understanding and treatment in koalas. FEMS Microbiology Reviews, 44(5), 583–605. https://doi.org/10.1093/femsre/fuaa024

3. Watters, F., Ramsey, D., Molsher, R., & Cassey, P. (2021). Breeding dynamics of overabundant koala. Wildlife Research (East Melbourne), 48(7), 663–672. https://doi.org/10.1071/WR20162

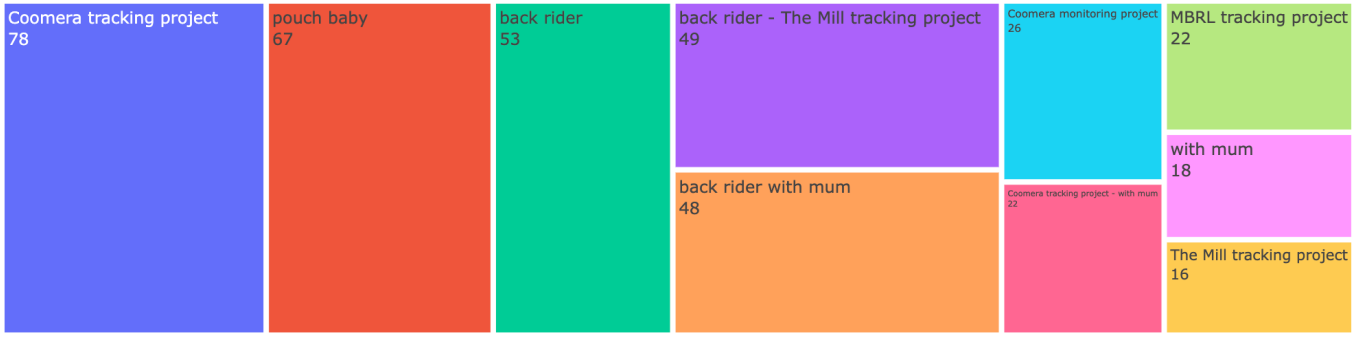
7. Appendices

Data dictionary:

1. 'Record No': Unique identifier for each record.
2. 'Koala Name': Name or identifier of the koala.
3. 'Call Date Time': Date and time when the call/report was made.
4. 'LAT': Latitude coordinate of the incident location.
5. 'LNG': Longitude coordinate of the incident location.
6. 'Adult Fate Other': Additional details about the fate of the Adult koalas.
7. 'Adult Size': Size classification of the adult koala.
8. 'Adult Gender': Gender of the adult koala.
9. 'Adult Situation': Situation of the adult koala.
10. 'Adult Situation Other': Other details about the situation of the adult koala.
11. 'Young Present': Indicator of whether young koalas were present.
12. 'Adult Fate': Fate of the adult koala (e.g., released, deceased).
13. 'Young Fate Other': Additional details about the fate of the young koalas.
14. 'Koala Found Address': Address where the koala was found.
15. 'Koala LGA': Local Government Area where the koala was found.
16. 'Sick': Indicator of whether the koala was sick.
17. 'Injured': Indicator of whether the koala was injured.
18. 'Caused By Dog': Indicator of whether the incident was caused by a dog.
19. 'Orphaned': Indicator of whether the koala was orphaned.
20. 'Under Threat': Indicator of whether the koala was under threat.
21. 'Dead': Indicator of whether the koala was deceased.
22. 'Conjunctivitis': Indicator of whether the koala had conjunctivitis.
23. 'Cystitis': Indicator of whether the koala had cystitis.
24. 'Wasted': Indicator of whether the koala appeared wasted.
25. 'Sick Other': Additional details about the sickness of the koala.
26. 'Vehicle Hit': Indicator of whether the koala was hit by a vehicle.
27. 'Road Speed Limit': Speed limit of the road where the incident occurred.
28. 'Fall': Indicator of whether the koala fell.
29. 'Injury Other': Additional details about the koala's injury.
30. 'Field Comments': Additional comments or notes about the incident.
31. 'Status Other': Other status details about the koala.
32. 'Release Date': Date when the koala was released.
33. 'Release Location': Location where the koala was released.
34. 'Release Suburb': Suburb where the koala was released.

35. 'Release Post Code': Postal code of the release location.
36. 'Release LAT': Latitude coordinate of the release location.
37. 'Release LNG': Longitude coordinate of the release location.

Appendix Figure 1| Word cloud depicting the most common streets where koalas were hit



Appendix Figure 2| Top 10 Situations Leading to Adult Injuries: A Treemap Visualisation