Brunel University London

Student ID – 2267302

Module Code – CS5803

Module Title – Data Visualisation

Academic Year – 2022/23

Assignment Title – Visualisation Design Task

<u>Topic: Implementation of the Australian Road Deaths Dashboard in Tableau</u>

1. Introduction

The Australian Road Deaths Database (ARDD) is a comprehensive resource containing detailed information on road fatalities in Australia, collected by the police and reported to the State and Territory road safety authorities. This dataset, maintained by the Bureau of Infrastructure, Transport and Regional Economics, offers both preliminary and revised data, with information compiled specifically for individuals killed in road crashes.

The ARDD captures crucial demographic and crash-related information, with each record representing either a fatal crash or a fatality. These records are linked through the unique Crash ID field. The database includes valuable data on factors such as the time, location, and type of crash, as well as the involvement of different types of vehicles, speed limits, and the road user type of the deceased. It also provides information on the demographics of the individuals involved, such as their age and gender.

Additionally, the dataset incorporates geographic classification data, including National Remoteness Areas, Statistical Areas Level 4, and National Local Government Areas. Other fields capture details about the road type, whether the crash occurred during specific holiday periods, and standardized age groupings. With this extensive range of data, the ARDD offers valuable insights into road safety in Australia, helping to inform policy and prevention efforts aimed at reducing the number of road fatalities in the country.

1.1Data Dictionary

Field	Description	Format	Values
Crash ID	National crash identifying number	Text	8-characters, unique to each fatal crash
State	Australian jurisdiction	Text	Abbreviation for each state/territory NSW Vic Qld SA WA Tas NT ACT
Month	Month of crash	Integer	01 to 12
Year	Year of crash	Integer	1989 – Current year
Dayweek	Day of week of crash	Text	Monday Tuesday Wednesday Thursday Friday Saturday Sunday
Time	Time of crash	Time	hh:mm

Crash Type	A crash is coded as a Pedestrian crash if a pedestrian was killed; otherwise, the number of vehicles involved.	Text	Single Multiple
Number Fatalities	Number of killed persons (fatalities) in the crash	Integer	Yes No
Bus Involvement	Indicates involvement of a bus in the crash	Text	Yes No
Heavy Rigid Truck Involvement	Indicates involvement of a heavy rigid truck in the crash	Text	Yes No
Articulated Truck Involvement	Indicates involvement of an articulated truck in the crash	Text	Yes No
Speed Limit	Posted speed limit at location of crash	Text	10 15 20 25 30 <40 40 50 60 70 75 80 90 100 110 130
Road User	Road user type of killed person	Text	Driver Passenger Pedestrian Motorcycle rider Motorcycle pillion passenger Pedal cyclist (Note: includes pillion passenger)
Gender	Sex of killed person	Text	Male Female
Age	Age of killed person (years)	Integer	

National Remoteness Areas 2016		Text	Major Cities of Australia Inner Regional Australia Outer Regional Australia Remote Australia Very Remote Australia
Statistical Areas Level 4 (SA4) Name 2016		Text	
National Local Government Areas (LGAs) Name 2017		Text	
National Road Type	Geoscape Australia, Transport and Topography 2020	Text	National or State Highway Arterial Road Sub-arterial Road Collector Road Local Road Access Road Pedestrian Thoroughfare
Christmas Period	Indicates if crash occurred during the 12 days commencing on December 23rd	Text	Yes No
Easter Period	Indicates if crash occurred during the 5 days commencing on the Thursday before Good Friday	Text	Yes No
Age Group	Standard age groupings used in the Road Deaths Australia monthly bulletin	Text	0_to_16 17_to_25 26_to_39 40_to_64 65_to_74 75_or_older
Day of week	Indicates if crash occurred during the weekday or weekend. (Note: 'Weekday' refers to 6am Monday through to 5:59pm Friday)	Text	Weekday Weekend
Time of day	Indicates if crash occurred during the day or night (Note: 'Day' refers to 6am through to 5:59 pm)	Text	Day Night

Table: 1 Data description

Source of the dataset: https://www.bitre.gov.au/statistics/safety/fatal_road_crash_database

1.2 User Persona:

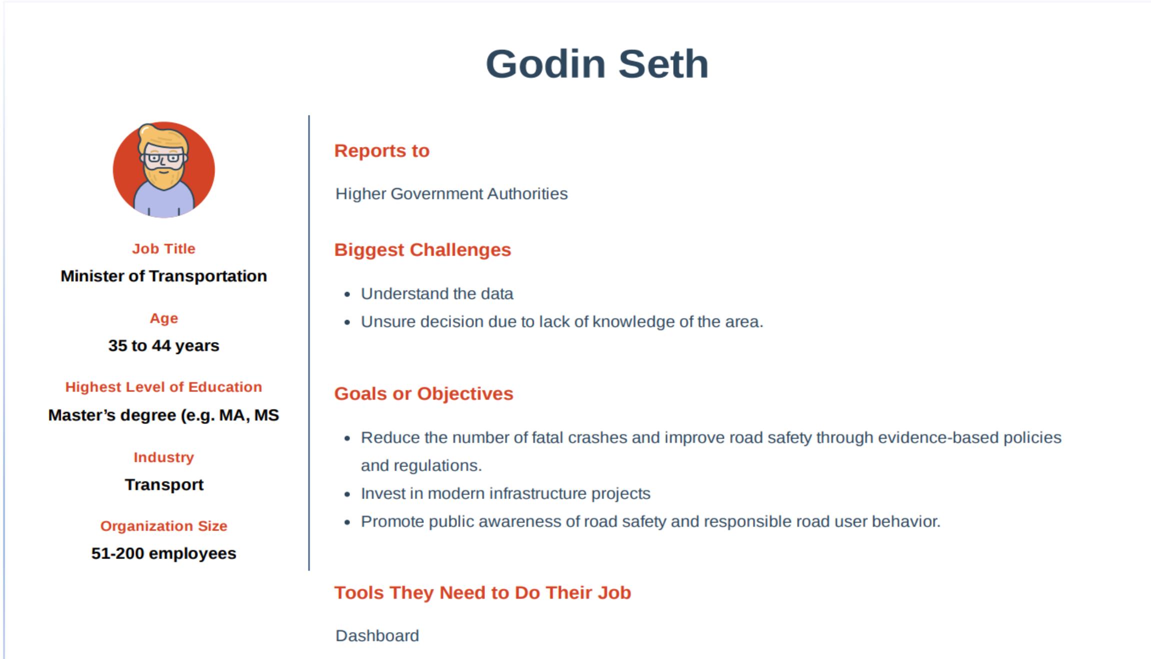

Fig 1: Persona Details

1.3 Persona Questions and Requirements

Q1. What is the annual trend of the fatal crashes in Australia? Which part was most affected in the country by fatal crashes and on which weekday?

R1: To answer Q1, the user needs to visualize the yearly number of fatal crashes, enabling them to observe trends over the years. A good view would be a line chart with years on the x-axis and the number of fatal crashes on the y-axis.

R2: To identify the most affected part of the country, the user needs to see the geographic distribution of fatal crashes. A suitable view would be a map, color-coded by the density of fatal crashes per region/state.

R3: To analyse the frequency of fatal crashes by weekday, the user needs to see a comparison between weekdays. A treemap with weekdays on the x-axis and the number of fatal crashes on the y-axis would be suitable. A treemap, on x-axis the number

Q2. How various factors like speed limit, road types and crash types have an adverse impact on fatal crashes? Do the proportion changes base on the weekdays or weekends?

R4: To answer Q2, the user needs to visualize the relationship between speed limits and the number of fatal crashes among different states. A good view would be a Box-Whisker plot with speed limit on the y-axis and the number of fatal crashes on the y-axis and states on x-axis.

R5: To analyse the impact of road types on fatal crashes, the user needs to visualize a comparison between different road types. A suitable view would be a stack chart with road types on the x-axis and the number of fatal crashes on the y-axis.

Non-functional requirements:
NF1: The dashboard is made in consideration of user-friendly and intuitive, catering to users with varying experience in interpreting statistical graphics.

NF2: The visualizations would be clear and easy to read, with proper colour schemes and labelling to ensure useability for users with varying visual abilities.

NF3: The dashboard should be compatible with different devices and screen resolutions, ensuring optimal display and user experience across various platforms.

NF4: The mode of interaction with the dashboard would be easy-to-use, allowing users to easily explore the data and apply filters or selections as needed.

2. Design

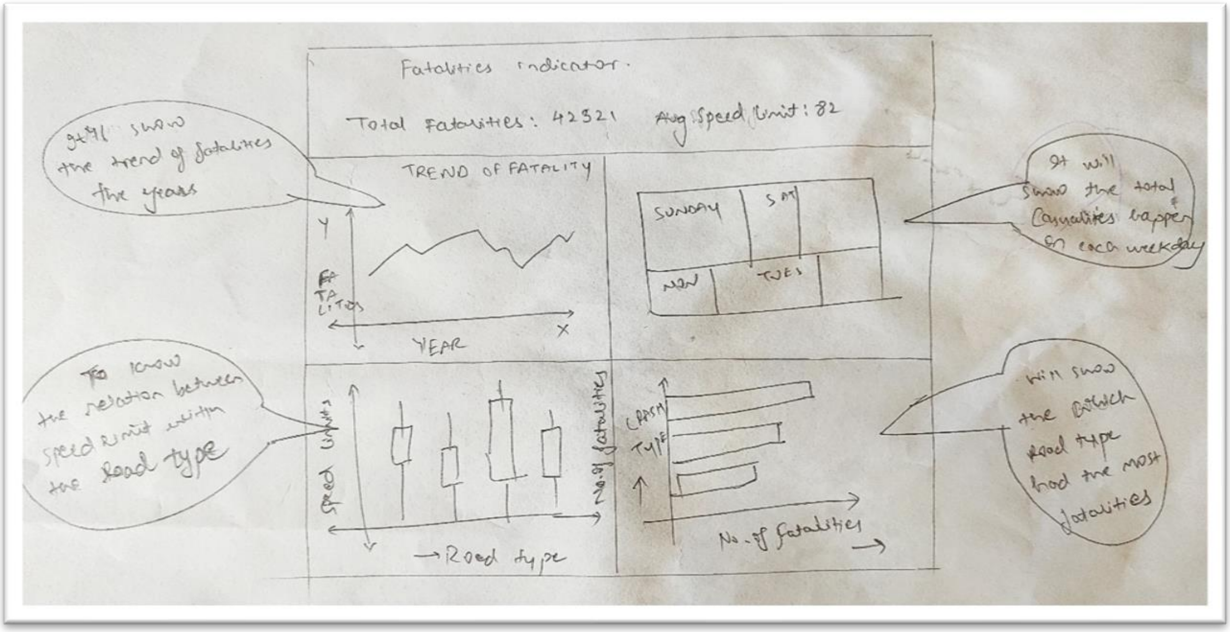

Fig 2: Prototype Design

During the proposal feedback, I presented the initial rough figure prior to implementing it in Tableau. The feedback highlighted that the research questions were loosely connected to the graphs I proposed. As a result, it became necessary to refine either the visualizations or the research questions. In response, I opted to work on both aspects, and the improvements can be observed in the subsequent work.

Furthermore, due to the advanced features and limitations inherent in the Tableau software, the final implementation resulted in visualizations and interactions that differ significantly from the original prototype. These differences stem from adapting to the software's capabilities, optimizing the design to enhance user experience, and making adjustments to better convey the information and insights derived from the data. In the process, the refined visualizations and interactions not only adhere to the software constraints but also more effectively address the research questions and present a more coherent narrative to the end-users.

The final view of the dashboard that our persona would use:

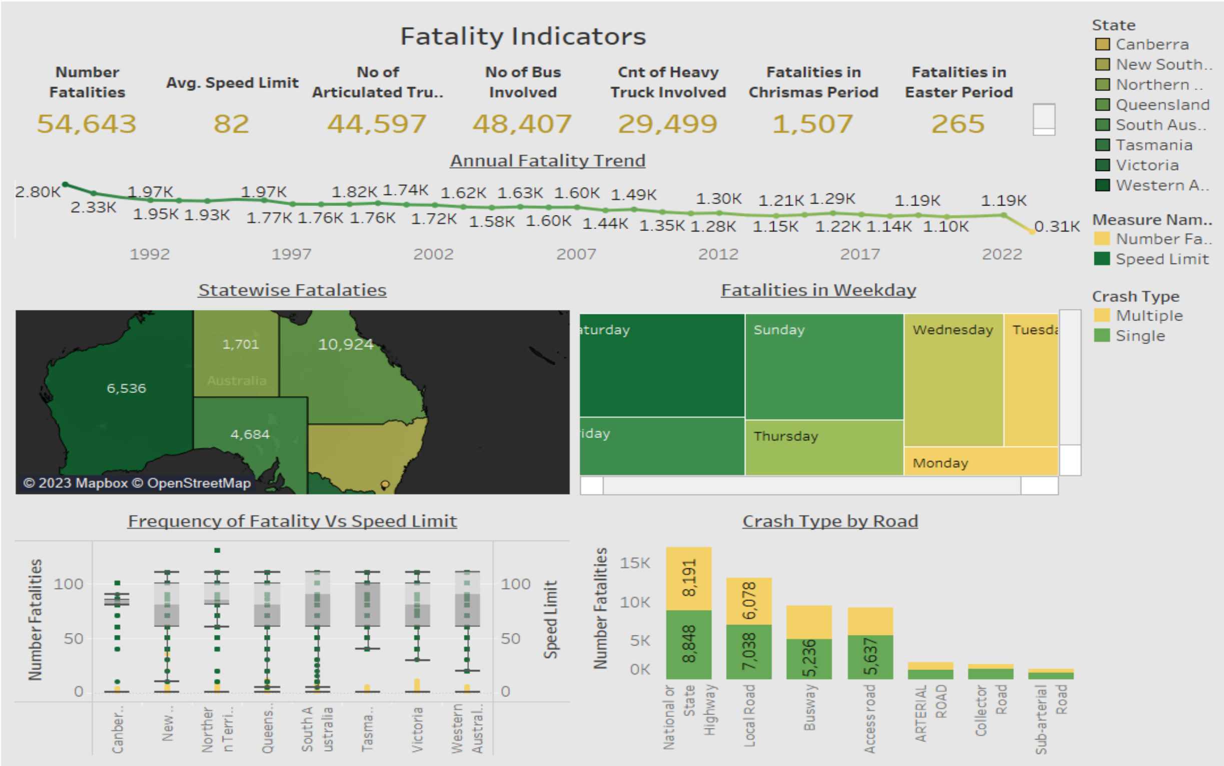

Fig 3: Final Dashboard

As it can be clearly observed that there are significant changes in the final design compared to the prototype.

In the initial prototype of the Australian Road Deaths Dashboard, I had included two primary indicators that highlighted total fatalities and the average speed limit. However, as I progressed from the prototype to the final implementation, it demanded a more comprehensive view of the data. As a result, I expanded the indicators to include seven essential metrics, such as the number of articulated buses and trucks involved, as well as fatalities during the Christmas and Easter periods which will give important information at a single glance.

Additionally, while the prototype featured four different types of charts – a line graph, treemap, boxplot, and bar chart – I decided to further enhance the final implementation with the addition of geographical maps. These maps provide users with a clearer visual representation of the regions most affected by fatal crashes, helping them quickly identify areas that may require more attention and policy intervention.

Moreover, in the final implementation, I replaced the bar chart with a stacked bar chart to better address the research questions. This change allows users to see not only the total number of crashes per road type but also the proportion of crashes for each speed limit within each road type. This richer visualization offers users a more in-depth understanding of how speed limits and road types influence the occurrence of fatal crashes.

Throughout the development process, I made these adjustments to evolve the dashboard from the initial prototype to the final implementation, ensuring that the visualizations and interactive features effectively answer the research questions and provide a valuable tool for decision-making in road safety policy.

3. Implementation

3.1 Tableau

The implementation process for the Australian Road Deaths Dashboard in Tableau began with the recreation of the initial prototype as accurately as possible. This involved using various chart types, visual encodings, and interactivity features to address the research questions (RQ1 and RQ2) effectively. Some adjustments were made to the original design based on the limitations and functionalities of Tableau. This section outlines the key steps in the implementation process and highlights the significant issues encountered and how they were addressed. The theme of the tableau would be the colour of Australia's flag. In the first stage, the dataset had some missing values and duplicated columns so therefore it was removed by the process Data Cleaning and Preparation.

3.1.1 Fatalities Across Australia Statistics

3.1.1.1 Fatalities Indicator Implementation

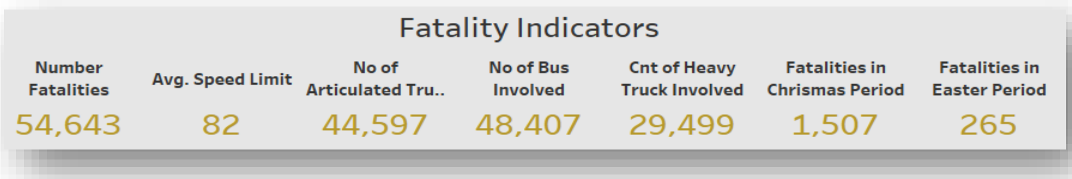

Fig 4: Fatality Indicator (Top part of Dashboard)

The figure shows the dashboard's upper horizontal part. Users can rapidly determine the overall number of fatalities across the country by using this section. Furthermore, Tableau's features enable users to acquire statistics on total fatalities for specific regions. Similarly, users can investigate the average speed limit in relation to different road types or geographical places. The dashboard also shows the number of articulated trucks, buses, and trucks involved in accidents, emphasizing the importance of heavy vehicles in these incidents. Users can also learn about fatalities that occur around holiday seasons such as Christmas and Easter. These statistics help to widen understanding of transportation concerns and promote more effective transportation management.

3.1.1.2 Calculative field created to make Fatality Indicator

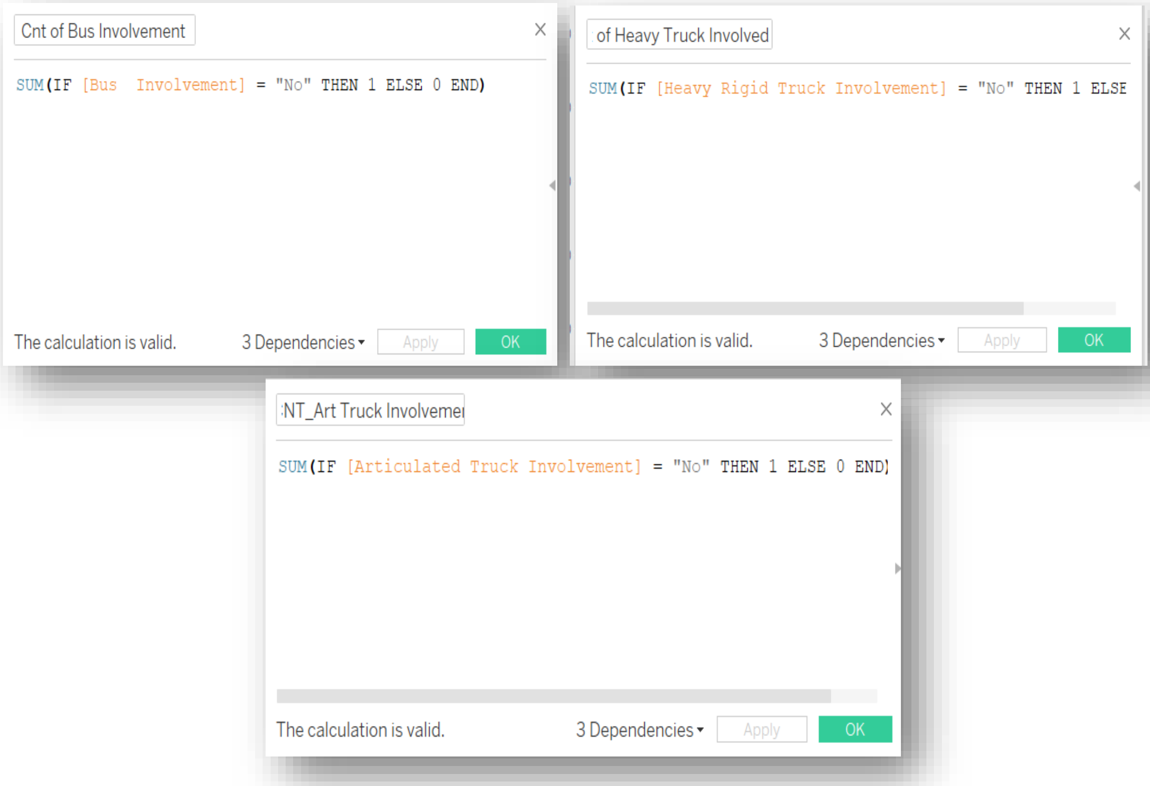

Fig 5: Calculative fields

The dashboard employs a straightforward logic utilizing the formula SUM(IF "Variable Name" = "No" then 1 else 0 end). This code is employed to calculate the number of bus involvement, heavy rigid truck involvement, and articulated truck involvement. In other words, the count for each type of heavy vehicle involved in accidents is derived using this method. This information could potentially provide users with insights for future infrastructure developments, such as separating travel roads for various vehicle types or implementing other safety enhancements to reduce accident risks involving heavy vehicles.

3.1.2 Understanding Fatalities Trend by geography and weekdays.

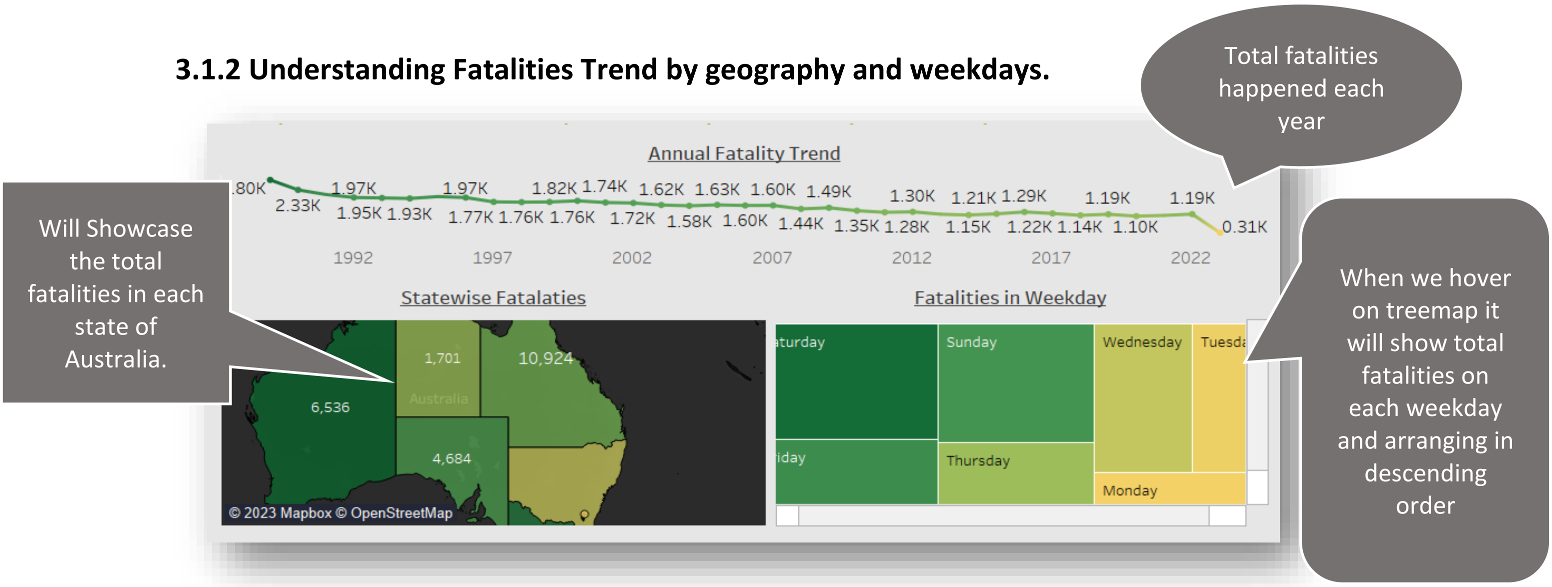

Fig 6: Dashboard 2nd section

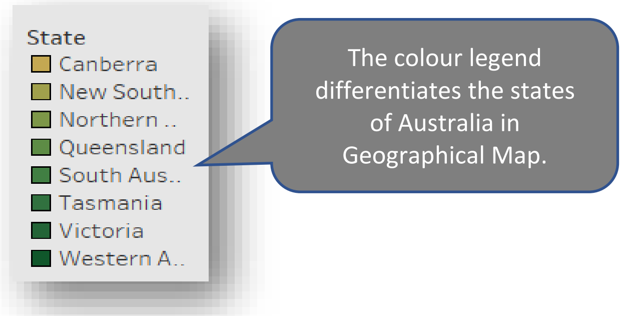

Fig 7: Color Legend for states

One section of the dashboard given in the above figure shows that distribution of fatalities across states in Australia along with the tree map which depicts the number of fatalities happened on weekday or weekend. In geographical map, the color used here depicts the name of the state in the graph. So that it would be easy for the user to see the insights on fatalities over the given regions of Australia. Moreover, the line chart shows the trend of crashes happening since the 1990s. which depicts that there is that the trend decreasing slowly by each year. At the same time, the "Fatalities in Weekday" graph shows the total number of fatalities with the highest weekday will have bigger size of box.

3.1.3 Impact on fatalities by speed limit and the type of roads

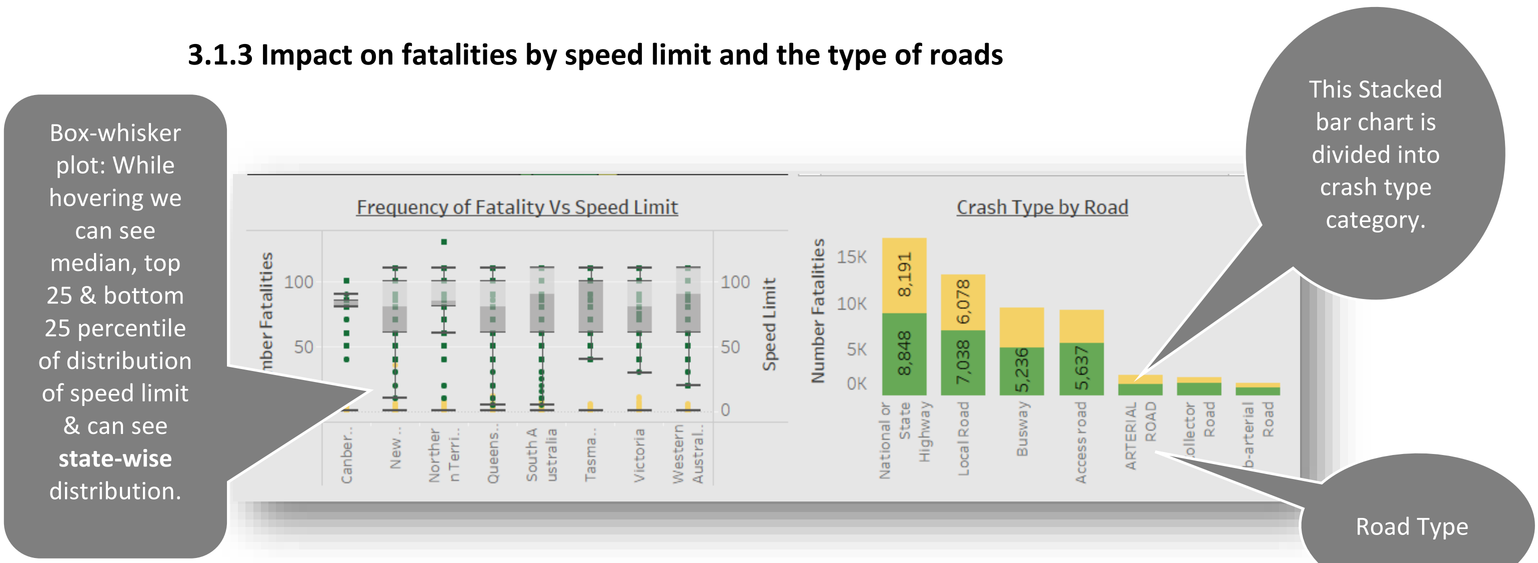

Fig 8: 3rd section of dashboard

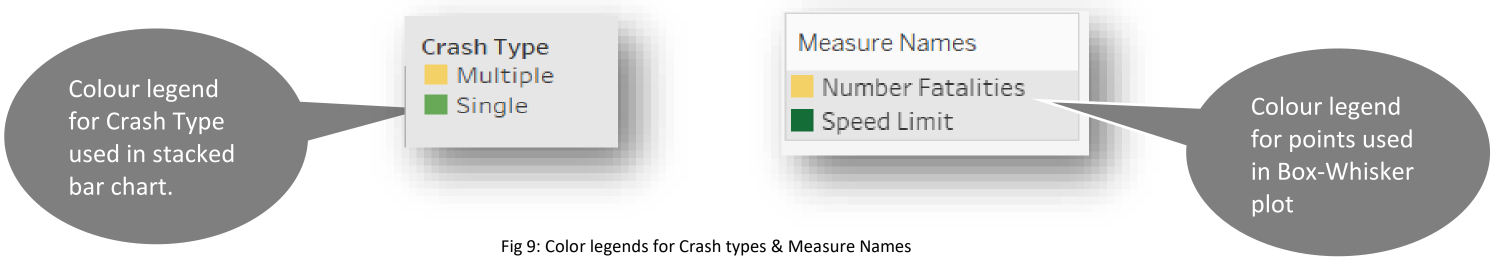

Fig 9: Color legends for Crash types & Measure Names

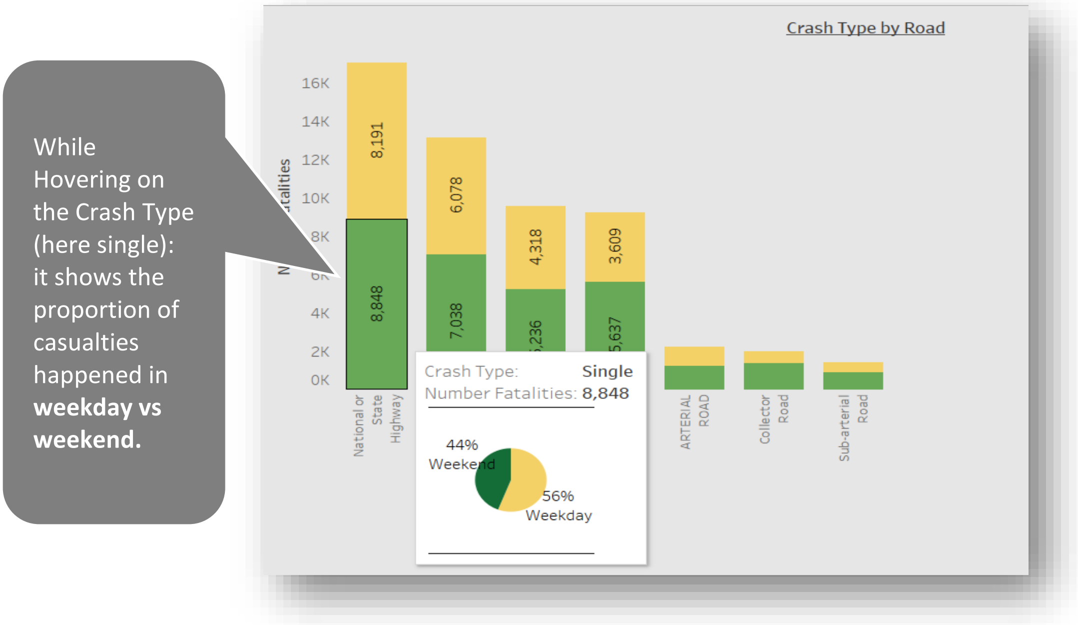

Fig 10: Stacked bar chart

This section highlights the second part of our dashboard, delving into the relationship between casualties and factors such as speed limits and road types. The box-whisker plot effectively illustrates the correlation between fatalities and speed limits across different Australian states, revealing that higher speed limits often result in a higher number of casualties.

The stacked bar chart, on the other hand, examines the relationship between crash type, road type, and state-wise proportions of fatalities during weekdays and weekends. The figure above demonstrates that by hovering over specific sections, users can access further insights into the data. For example, one can determine which crash type has the highest fatalities on national or state highways and the proportion of fatalities occurring on weekdays. In this case, the chart indicates that 56% of the 8848 single crashes happened during weekdays.
The colour legends featured in the annotated screenshot aid users in quickly comprehending the visualizations.

4. Walkthrough

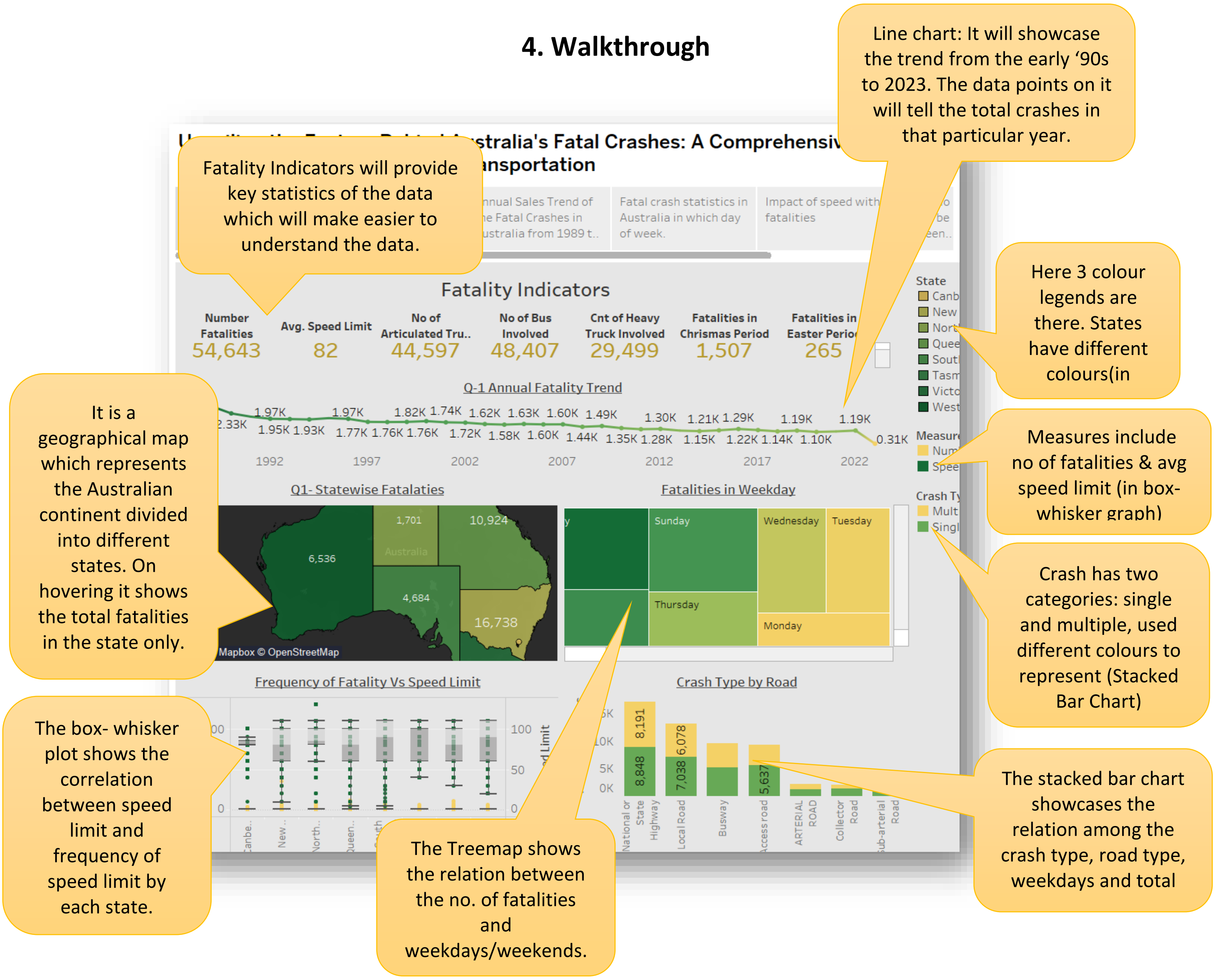

Fig 11: Walkthrough of Dashboard

This section will provide a walkthrough to the dashboard to address the research question that user would be satisfied with. The responsibility to reduce the number of fatal crashes and improve road safety through evidence-based policies and regulations.

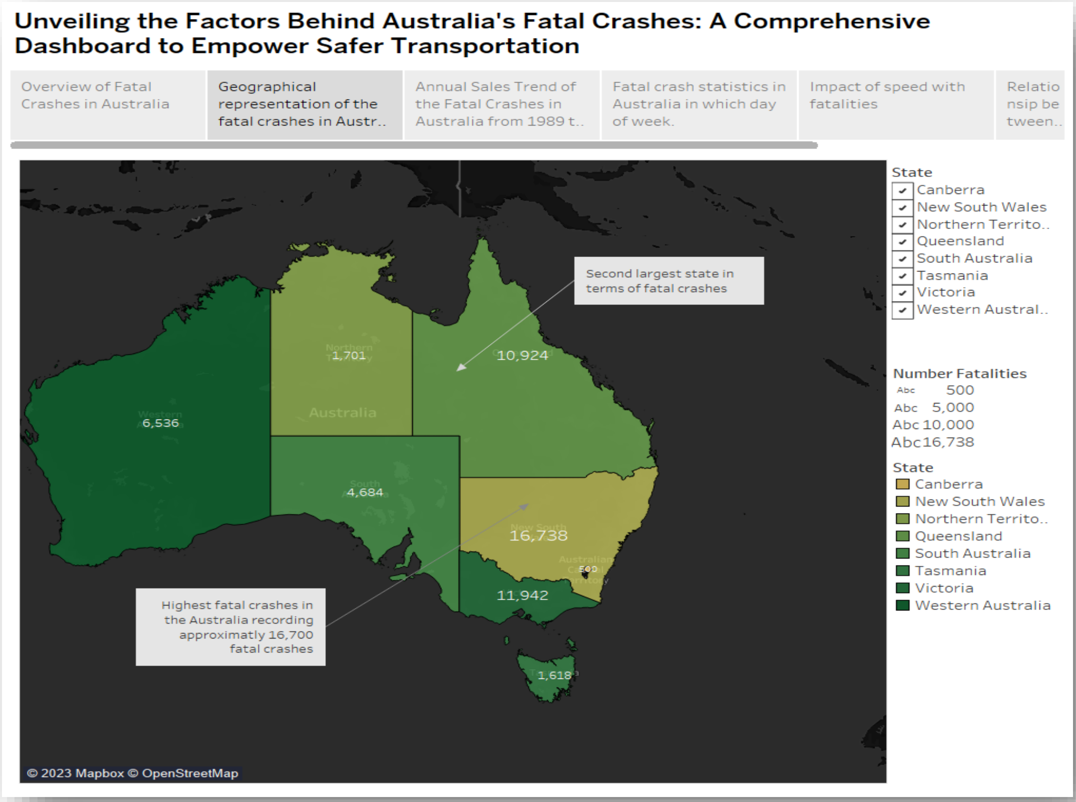

Fig 12: Story-1 view

The graph correctly distinguishing the state by the shades of green color while maintaining the theme of the dashboard. The above story satisfy the R2 to partly answer the research question (Q1). The geograhical map shows the states of Australia and depicts the total fatalities region wise. By observing the graph, it can be clearly said that Canberra is the state with the highest cases followed by Quensland. Which answers the part of our Q1

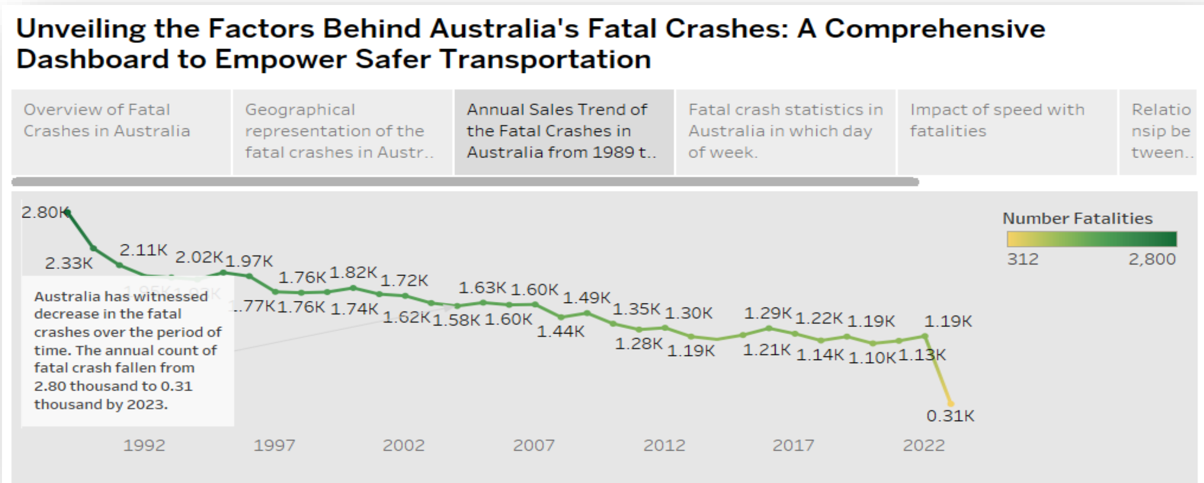

Fig 13: Story-2 view

From the line chart, the visual combinations of color and declining line in the line chart gives a perfect visual to understand the trend of the fatalities over the year. Which fulfill the R2 for Q1 and depicts that Australia has witnessed decrease in the fatal crashes over the period. The annual count of fatal crashes has fallen from 2.8 thousand to 0.31 thousand by 2023. It seems like a huge dip in the accidents. Moreover, because of the interactive nature of the software, the user can select any time frame to understand in-depth pattern.

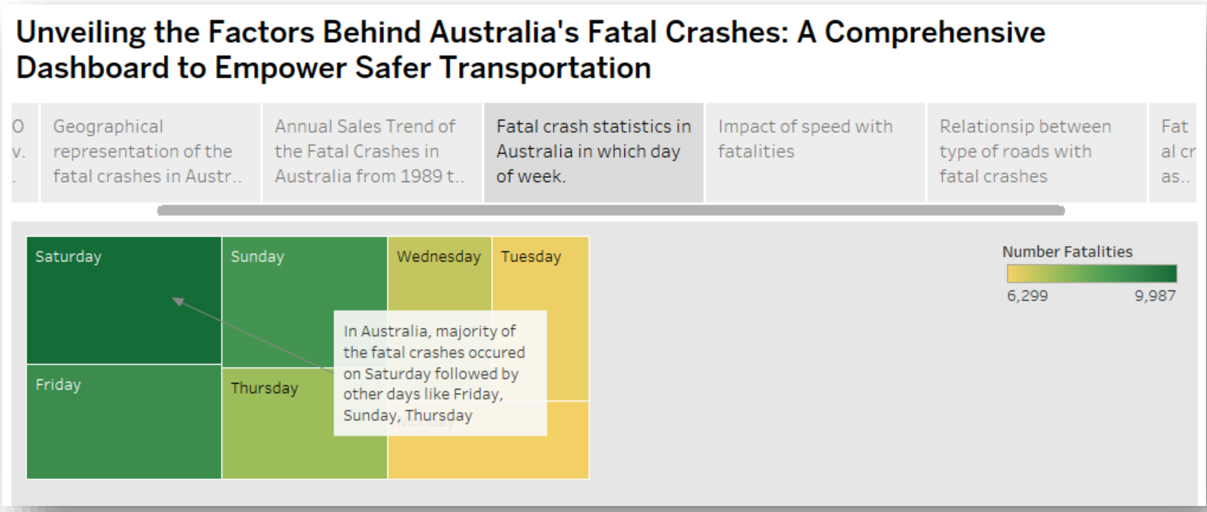

Fig 14: Story-3 view

From the treemap, it satisfies the R3, and it qualifies to answer the last part of the Q1 of the user. In Australia, majority of the fatal crashes occurred on Saturday followed by other days like Friday, Sunday, and Thursday. To gain in-depth analysis, the dashboard is equipped with color scale and size, to arrange the weekdays in descending order, and user on once glance can identify the highest fatalities is recorded on Saturday, and lowest is on Tuesday.

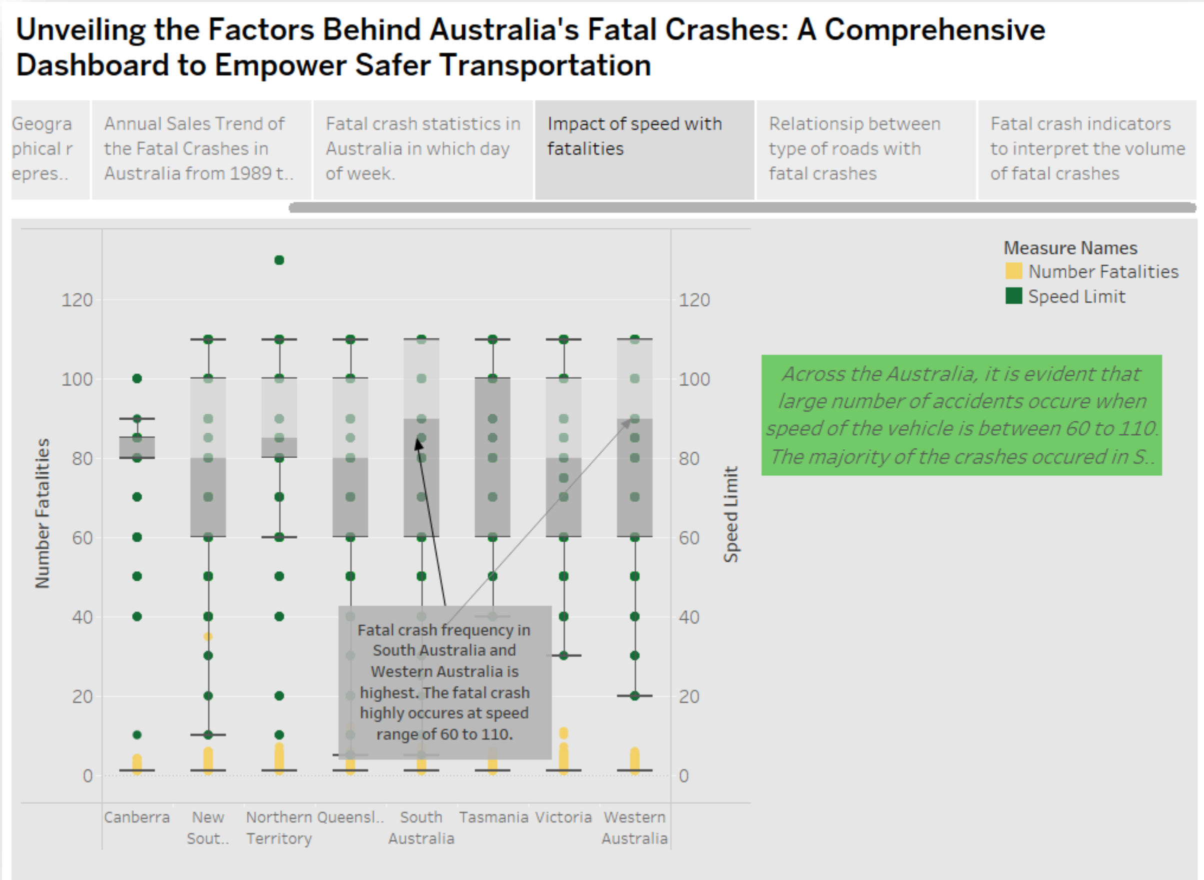

Fig 15: Story-4 view

From Box-Whisker plot, the user requirement (R4) is fulfilled by the plot. The graph is able to answer the part of Q2. The fatal crash frequency in South Australia and Western Australia is highest. The fatal crash highly occurs at speed range of 60 to 100. The insight would be helpful to know which states are more prone to crashes but would be incomplete to answer the Q2 without considering other influencing factors.

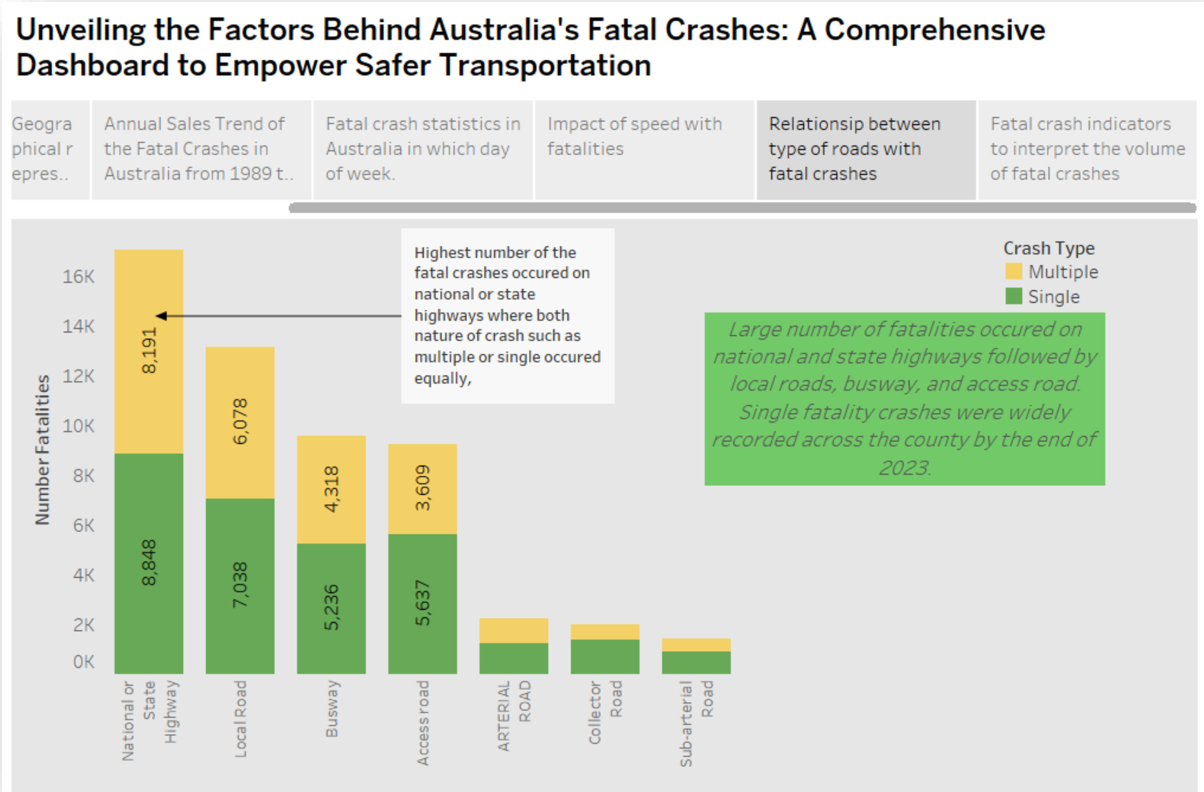

Fig 16: Story- 5 view

Finally, the stacked bar chart satisfies the R5 requirements that will be helpful to answer the remaining part. The highest number of fatal crashes occurred on national or state highways where both nature of crash such as multiple or single occurred equally. However, the user didn't get the proportion of accidents in weekdays versus weekends. Therefore, Tableau's one of feature to add additional chart to tooltip made it easier for user to get additional insights.

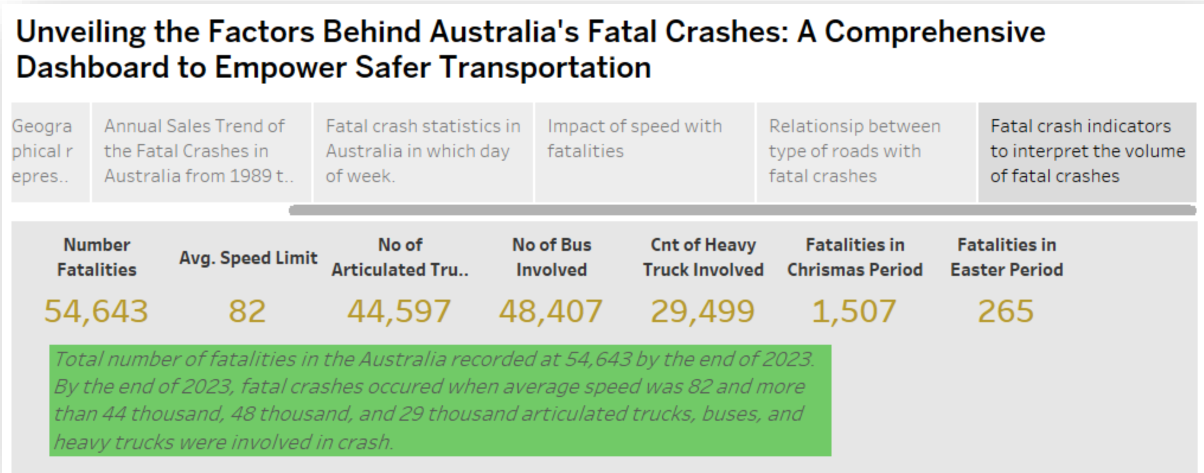

Fig 17: Story-6 view

From these indicators, the statistics can be directly extracted by the implementation of any interaction by the user.

5. Reflective Discussion

Through the implementation of this project, got to learn to think in terms of the user-centric design. It is very important to have well-defined research questions. I got to know that there is almost always possibly huge gap between the idea and implementation. Therefore, the project demanded a high amount of feedback and constant improvement of the design. This iterative approach ultimately resulted in a more efficient and informative dashboard that can assist in shaping policy decisions.

Got to know about the importance of the using suitable visual encodings, chart types, interaction methods and colour encoding. I learned that the developer should focus on how we can answer the complex questions easily through simple graphs and proper interaction. Which was the base of my whole implementation I had done in this project.

However, there were some limitations I faced during implementation because of uncleaned data and Tableau's own limitations to manage uncleaned dataset which forced me to clean dataset first. However, overall, the experience is moderate using Tableau.

6. Conclusion

In conclusion, this project has not only been an insightful experience but has also provided a comprehensive understanding of the importance and application of data visualization in addressing real-world problems. Throughout the project, various stages such as data collection, research question formulation, prototype design, and implementation using Tableau have been meticulously completed. The final dashboard effectively answers the research questions and presents valuable insights that can aid the Minister of Transportation in making informed decisions to enhance road safety and infrastructure.

Despite the challenges encountered during the implementation process, this project has greatly enriched my skills in crafting persuasive visualizations and adapting to the intricacies of Tableau. Furthermore, the reflective discussion has allowed me to critically evaluate the project and my personal learning journey. Overall, this project has demonstrated the power of data visualization in transforming complex data into actionable insights and has laid a solid foundation for future endeavors in the field of data analysis and visualization.