**Table of Content**

[**Introduction** 2](#_Toc199534235)

[**Preprocessing Data** 2](#_Toc199534236)

[ **Connect to the data:** 3](#_Toc199534237)

[ **Initial data type review and correction** 3](#_Toc199534238)

[ **Create a usable ‘DateTime’ field** 4](#_Toc199534239)

[ **Clean categorical and numerical fields** 5](#_Toc199534240)

[**Visualization and Reason** 6](#_Toc199534241)

[ **Visualization 1: Crash Trends Over Time** 6](#_Toc199534243)

[**Findings analysis:** 6](#_Toc199534244)

[ **Visualization 2: Weekday Total Crashes** 8](#_Toc199534245)

[**Findings analysis:** 8](#_Toc199534246)

[ **Visualization 3: Crash Timing Heatmap** 9](#_Toc199534247)

[**Findings analysis:** 9](#_Toc199534248)

[ **Visualisation 4: Severity by Condition** 10](#_Toc199534249)

[**Findings analysis:** 10](#_Toc199534250)

[ **Visualization 5: Crash Severity by Day of Week** 12](#_Toc199534251)

[**Findings analysis:** 12](#_Toc199534252)

[**Pattern discussion and interpretation** 13](#_Toc199534253)

[ **Temporal Patterns** 13](#_Toc199534254)

[ **Geographical Pattern** 13](#_Toc199534255)

[ **Insights on severity** 13](#_Toc199534256)

[**Data Visualization Ethics** 13](#_Toc199534257)

[**Conclusion:** 14](#_Toc199534258)

# **Introduction**

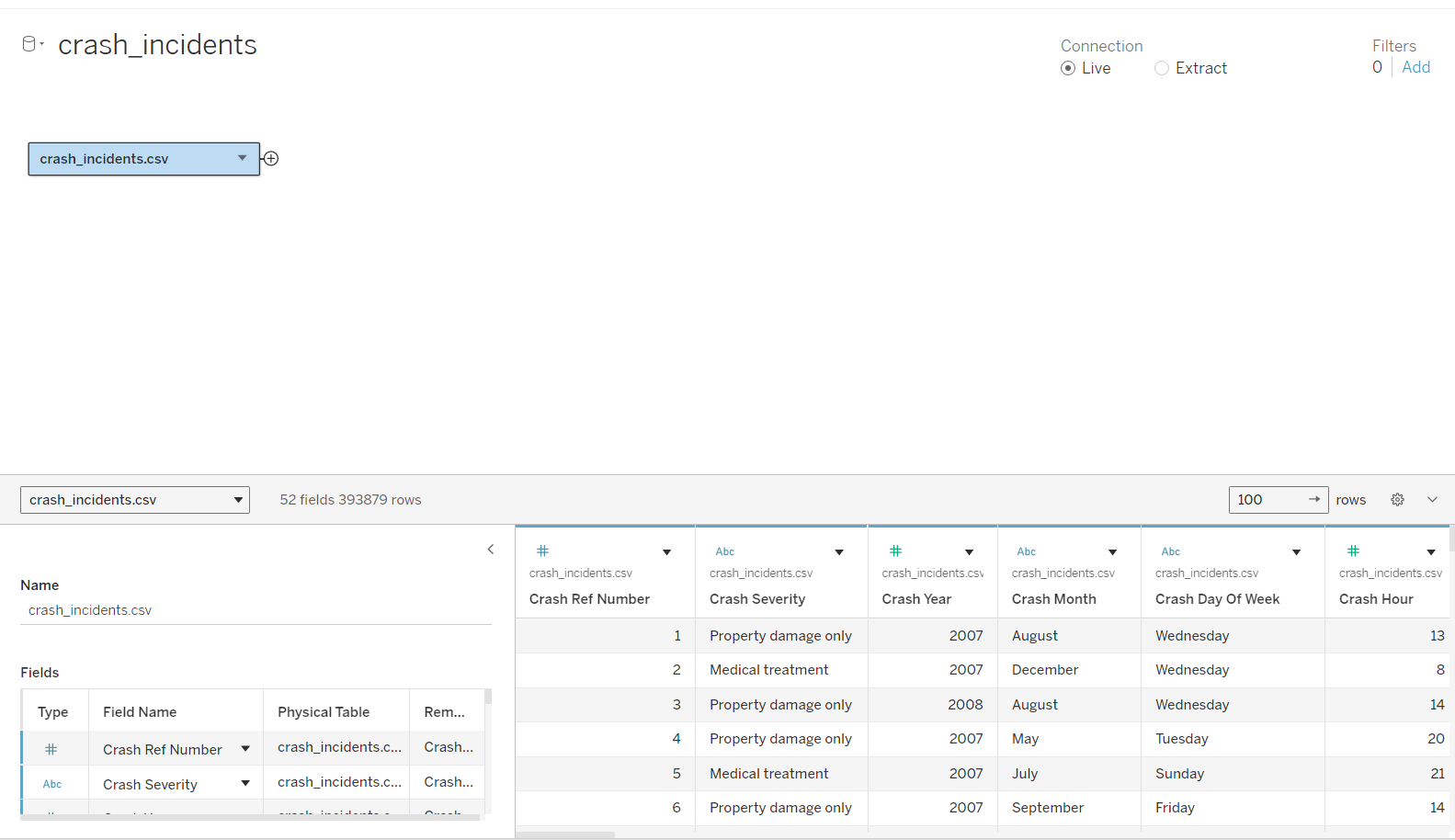
Safety when using vehicles for transportation is vital around the world, particularly in Queensland. As a data analysis we need to understand road accident patterns and trends in order to implement effective interventions and public safety measures.   
Exploring this report will provide you with a general understanding of Queenlands accidents and road safety. Using Tableau to pre-processing data and exporting 5 visualizations will provide insight into crash patterns, severity, geographical distribution, and contributing variables.  
After plotting various plots to acquire a general understanding, we will make some assumptions about the findings and address some ethical concerns concerning data analytics' role.

# **Preprocessing Data**

To get some insight with the data, at first we need to conduct a cleaning and preprocessing in Tableau following these steps:

* Connect to the data
* Initial data type review and correction
* Create a usable ‘DateTime’ field
* Handle geographic data issues (problematic coordinates)
* Clean categorical (text) fields
* Clean and transform ‘Crash\_Speed\_Limit’
* Review ‘Crash\_DCA\_Code, Crash\_DCA\_Description, Crash\_DCA\_Group\_Description’

## **Connect to the data:**



**Figure 1: Connecting the data to dashboard**

## **Initial data type review and correction**

Tableau automatically assigns data types to each column. Carefully review these in the data grid.

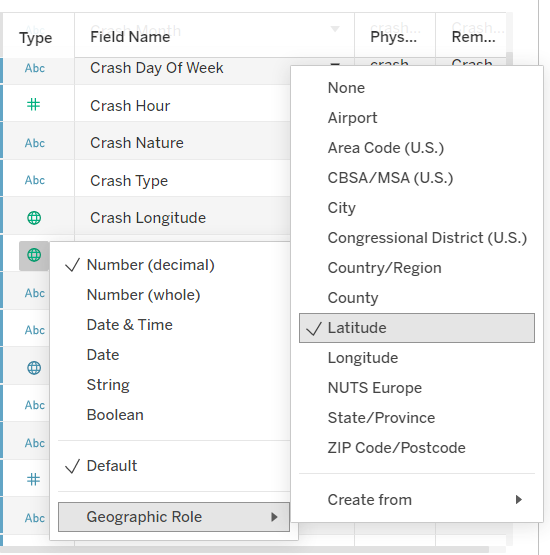
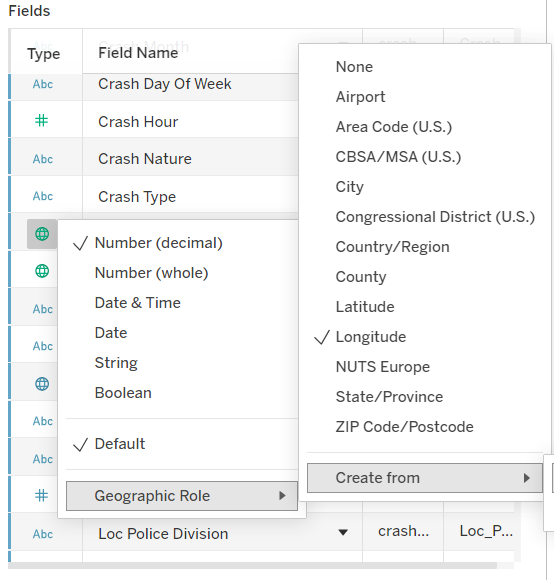
Ensuring Crash\_Ref\_Number, Count\_Casualty\_Hospitalised, Count\_Casualty\_MedicallyTreated, Count\_Unit\_Car…. Are whole number

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 2: Crash\_Ref\_Number Configuration**

Checking for Crash\_Longitude, Crash\_Latitude: Ensure these are number (Decimal => This is important for the step creating map, if you don’t change it into decimal, tableau going to recognize it as a **null value**) and assign them into correct geographic roles

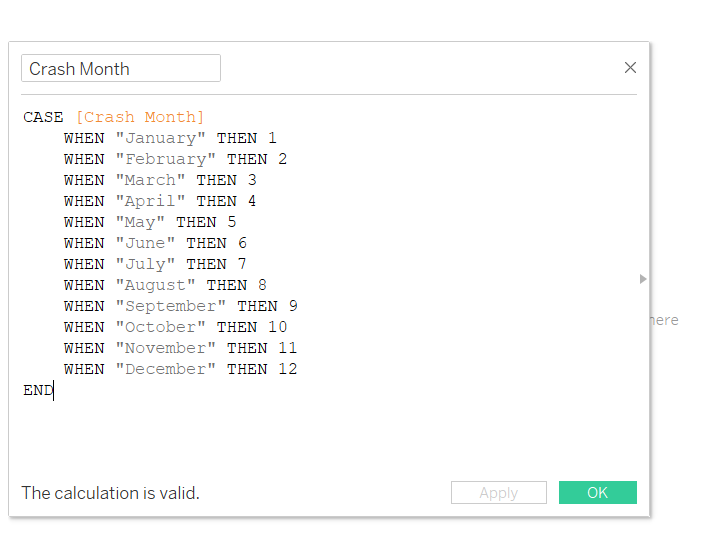


**Figure 3: Crash\_Longitude Configuration**

**Figure 4: Crash\_Latitude Configuration**

## **Create a usable ‘DateTime’ field**

Convert month to number.



**Figure 5: Converting month to number**

Create Crash Date, due to we a time series line chart that can filter each year, and every month of that year to have a in depth observation from the dataset.

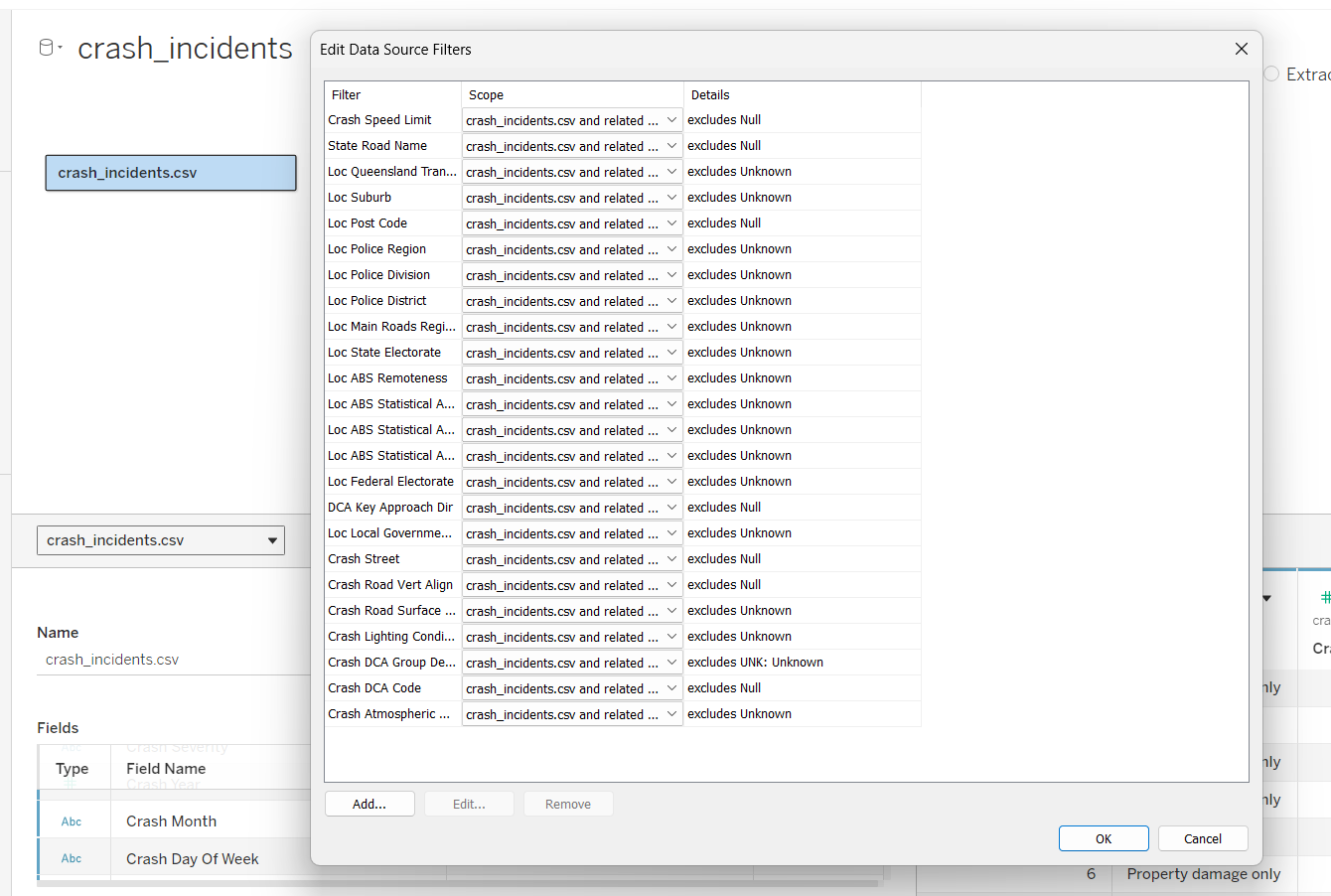
A screenshot of a computer

AI-generated content may be incorrect.

**Figure 6: Creating Crash Date**

## **Clean categorical and numerical fields**

Remove unknown value rows by filtering all the null or unknown value to gain best pratice of exploring the data.



**Figure 7: Cleanning null and unknown value**

# **Visualization and Reason**



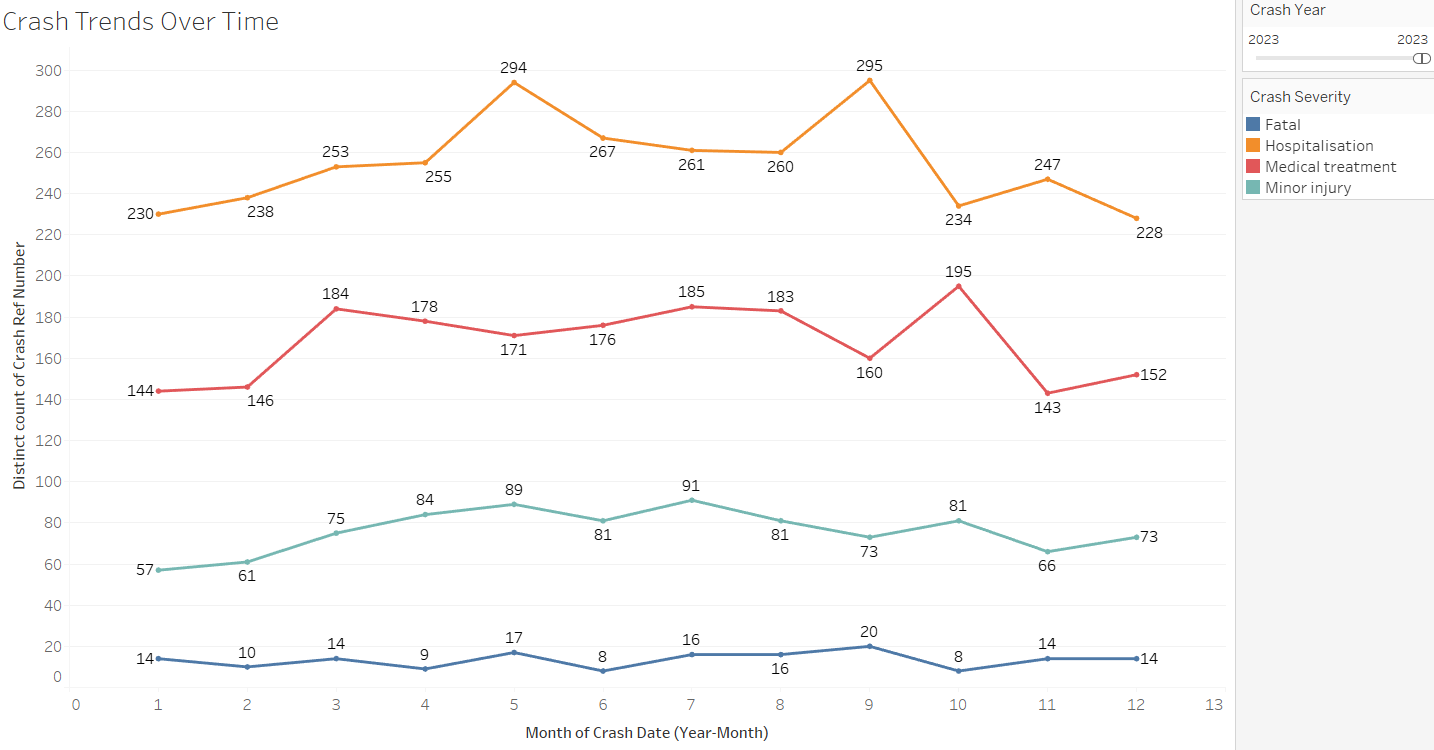
## **Visualization 1: Crash Trends Over Time**

A graph of different colored lines

AI-generated content may be incorrect.

**Figure 8: Crash Trends Over Time (sort by year)**

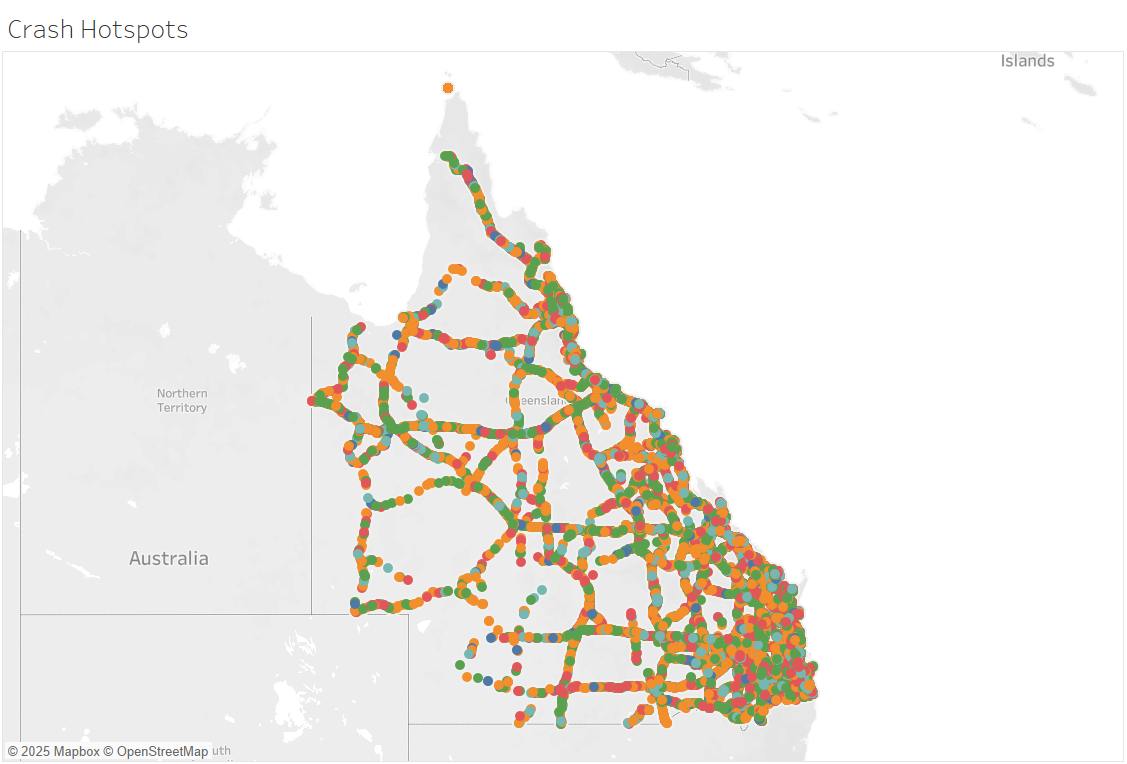
This plot indicates that "Minor injury" crashes (mint line) were consistently the most frequent type of incident reported between 2001 and 2021. Following similar overall trends, "Medical treatment" (red) and "Hospitalisation" (orange) crashes occurred less frequently but also peaked around 2007-2008 before generally declining.  
A major finding is a considerable decline in all collision in severity levels, including "Fatal" (blue) crashes, in recent years, significantly around 2020-2021. While fatal crashes are the least widespread, they appear to be following a current decreasing pattern. It suggests a significant overall decline in reported crash events near the end of the time represented.

****

**Figure 9: Crash Trends Over Time (sort by year)**

We can explore deeply though adding a filter of year, to gain more depth insight of this time series data. This can help us to gain more information from the dataset by checking the latest collision severity trend. By navigating and filtering in the year of 2023 for further exploration.

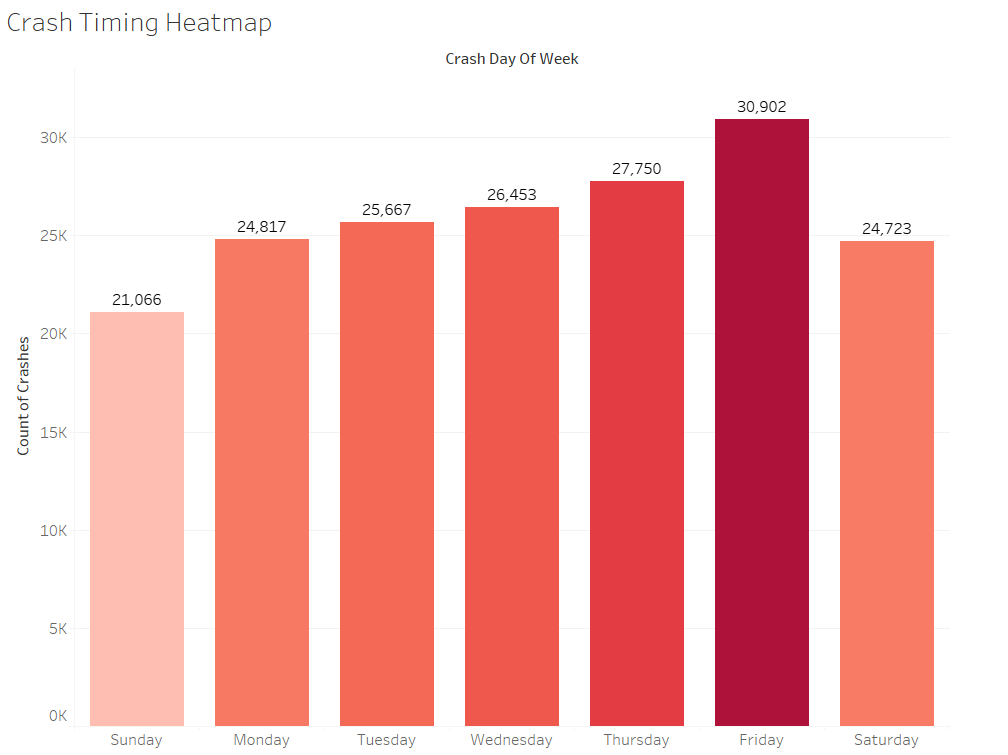
## **Visualization 2: Crash Hotspots**



**Figure 10: Crash Hotspots**

The "Crash Hotspots" map, giving us a clearly observation that crash incidents are not evenly distributed, with several distinct geographical concentrations. A major hotspot, featuring numerous crashes including fatal (blue) and hospitalization (orange) incidents, is prominent in the densely populated area detailed in the figure supplied.  
  
Other clusters appear along what seem to be arterial routes and within other populated zones. While minor injuries (mint) are visually the most widespread, severe crashes like medical treatment (red), hospitalisations, and fatalities are also present within these hotspot areas. The varying circle sizes effectively highlight specific locations experiencing a higher frequency of distinct crash events, pinpointing problematic zones.

## **Visualization 3: Crash Timing Heatmap**

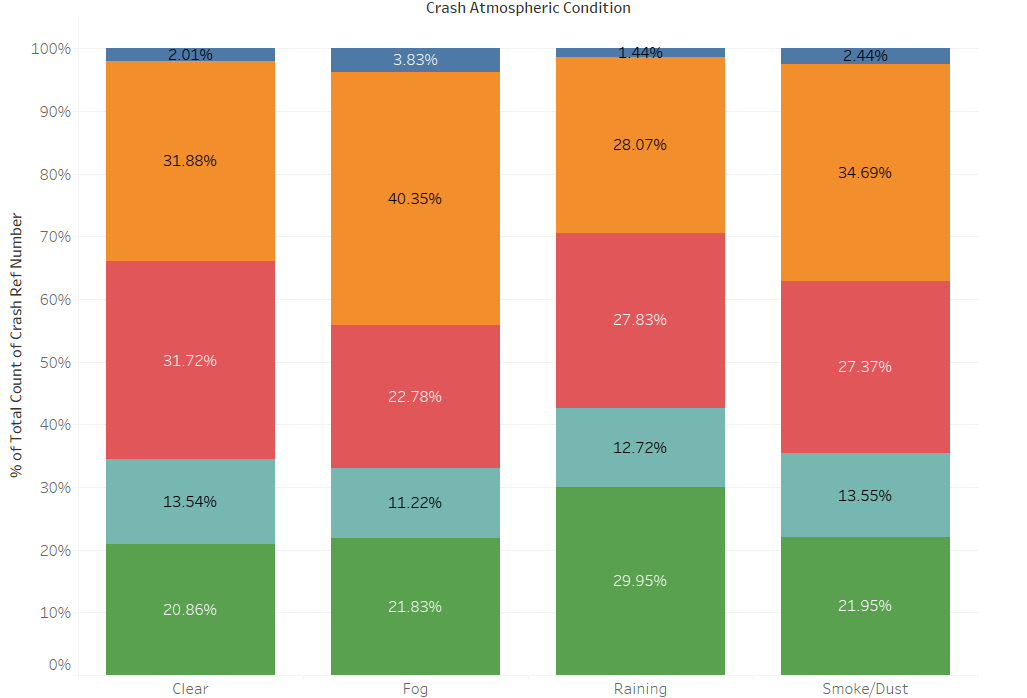


**Figure 11: Crash Timing Heatmap**

The "Crash Timing Heatmap" clearly illustrates that crash occurrences vary significantly by time of day and day of the week. Weekdays, from Monday to Friday, exhibit a pronounced concentration of incidents during the afternoon and evening commute hours, roughly between 2 PM and 6 PM. Friday evenings often show a sustained period of this heightened crash activity.

A secondary peak on weekdays is also visible during morning commute times. In contrast, weekend patterns differ, with Saturdays, in particular, showing increased crash frequencies extending later into the evening and night. Consistently, the early morning hours (approximately 2 AM to 5 AM) across all days experience the lowest number of reported crash incidents.

## **Visualisation 4: Severity by Condition**



**Figure 12: Severity by Condition (Atmospheric Specific)**

This chart reveals that the overwhelming majority of crash incidents, encompassing all severity levels, occur when the road surface condition is 'Dry'. 'Wet' road surfaces account for the second most significant number of crashes, although these are substantially fewer than those on dry roads.

Across the visible conditions, 'Minor injury' (mint) crashes consistently represent the most common outcome. This is followed in frequency by 'Medical treatment' (red), then 'Hospitalisation' (orange), and finally 'Fatal' (blue) incidents. While the absolute number of severe crashes (fatal and hospitalisation) is highest on 'Dry' roads due to the sheer volume of total incidents, these severe outcomes are also notably present when conditions are 'Wet'.

A screenshot of a computer

AI-generated content may be incorrect.Furthermore, if you want to explore many data insight you can change the data want to see, drag it into the column

**Figure 13: Data Pane**

A graph with numbers and a number of objects

AI-generated content may be incorrect.

A graph with different colored squares

AI-generated content may be incorrect.

**Figure 14: Severity by Condition (Road Surface Specific)**

**Figure 15: Severity by Condition (Lighting Specific)**

## **Visualization 5: Vehicle Involvement in Severe Crashes**

**Figure 16: Crash Severity Distribution by Day of Week**

The "Vehicle Involvement in Severe Crashes" bar chart clearly demonstrates that cars ('Count Unit Car') are the most prevalent unit type involved in major incidents. Their involvement far outnumbers that of all other groups, underscoring the significant role passenger vehicles play in serious traffic incidents.  
  
Pedestrians ('Count Unit Pedestrian') are the second most common unit involved in serious crashes, indicating a substantial risk to this vulnerable road user group. Motorcycles and mopeds ('Count Unit Motorcycle Moped') had the third highest number of involvements, closely followed by bicycles ('Count Unit Bicycle'), underscoring the dangers to two-wheeled transport users.   
  
Trucks ('Count Unit Truck'), while less common than the preceding groups, still account for a significant proportion of serious crash incidents. In contrast, busses ('Count Unit Bus') and 'Other' unit types have the lowest figures in this dataset. This chart clearly quantifies the involvement of various road users and vehicle types in high-impact events, highlighting critical areas for safety measures aimed at minimizing major crashes.

# **Interpretation of Patterns, Trends, Anomalies, and Potential Contributing Factors**

**Visualization 1:**

* Minor injuries were most common (2001–2021). Most severities peaked about 2007-2008 and then fell, with a notable drop in all severities in 2020-2021. Fatal crashes have also decreased slightly, particularly recently. With the monthly views (Figure 9) indicate seasonality.
* The significant 2020-2021 decline across all severity categories.
* The 2020-2021 fall was most likely caused by less traffic due to the covid-19 lockdowns. Long-term trends may apply to safety programs.

**Visualization 2:**

* Crashes are more concentrated in densely inhabited areas near arterial routes. Minor injuries are normal, but catastrophic incidents may occur in specific regions.
* Increased traffic volume and specific road network designs in heavily populated or critical transit regions may be contributory factors.

**Visualization 3:**

* Crashes peak on Friday, fall on Sunday, and increase overall from Monday to Friday.
* Peak hours are weekday afternoon/evening commutes (2-6 PM) and mornings. Friday evenings are bustling. Saturdays had higher crash rates late in the evening and night. Early mornings (2-5 a.m.) have generally been the lowest.
* Weekday commuting patterns and increased social/leisure activities on Friday and Saturday evenings influence these periods.  
    
  **Visualization 4:**
* Patterns/Trends: The majority of crashes occur on 'dry' road surfaces, with 'wet' coming in second. Fatal crashes are higher in 'Fog' and 'Smoke/Dust' (Figure 12) and on 'Unsealed' roads (Figure 14) than in other situations.
* 'Minor injury' is the most prevalent consequence in all circumstances.
* Reduced vision in fog, smoke, or dust, as well as impaired vehicle control on wet or untreated terrain, can all contribute to increased severity proportions.

**Visualization 5:**

* Cars are the most commonly involved unit type in serious crashes. Pedestrians rank second, followed by motorcycles/mopeds and bicycles.
* The huge number of cars on the road (exposure) and the inherent vulnerability of pedestrians and two-wheeler users all contribute to these results.

# **Data Visualization Ethics**

**Truth and accuracy (objectivity)**

* Principle: is to provide data honestly, without misleading distortion.
* Application: to the data pretreatment methods in your report (for example, type correction and cleaning unknowns) are designed to improve data quality and enable accurate visualizations. Using common chart types is also beneficial.
* Relevance: is also Critical for establishing trust and facilitating solid, data-driven decision-making.

**Clarity and understandability (accessibility).**

* Principle: Create visuals that are simple for the intended audience to understand.
* Application: Including titles and "Findings analysis" for each visualization improves comprehension.
* Relevance: ensures that insights are accessible to people other than data professionals, allowing for more effective communication and action.

**Context and impartiality (avoiding bias)**

* Principle: Presenting data with sufficient context while reducing biases in display or perception.
* Application: The "Pattern discussion and interpretation" section and data preparation explanations provide some background. Recognizing data constraints and interpreting findings cautiously are critical.
* Relevance: Encourages fair readings and appropriate data use, avoiding distorted perspectives and unjust conclusions.

# **Conclusion:**

To conclude, in order to increase road safety, through this study has identified some important trends and patterns in Queensland road crash statistics. Using Tableau preprocess and show data to expose temporal crash distributions, including peak times on Friday and during commuter hours, geographical hotspots, and how conditions and vehicle type impact event severity.  
  
These results support focused interventions comprising highway feature safety audits and Friday road safety events. This data-driven research demonstrates how visual analytics can transform challenging information into insightful analysis for Queensland road safety.