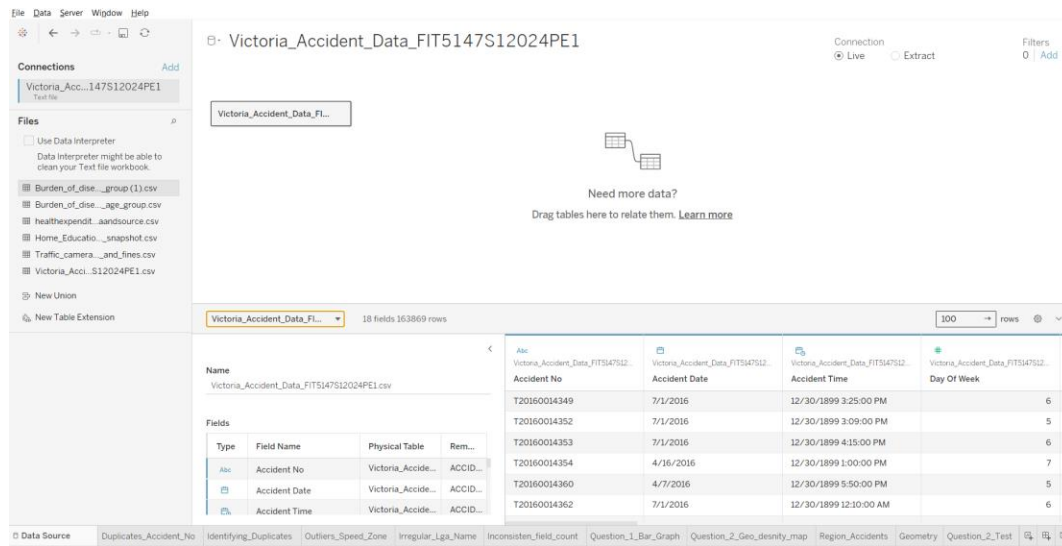


FIT5147
DATA EXPLORATION & VISUALISATION
Exploring Data on Accidents in Victoria

Table of Contents

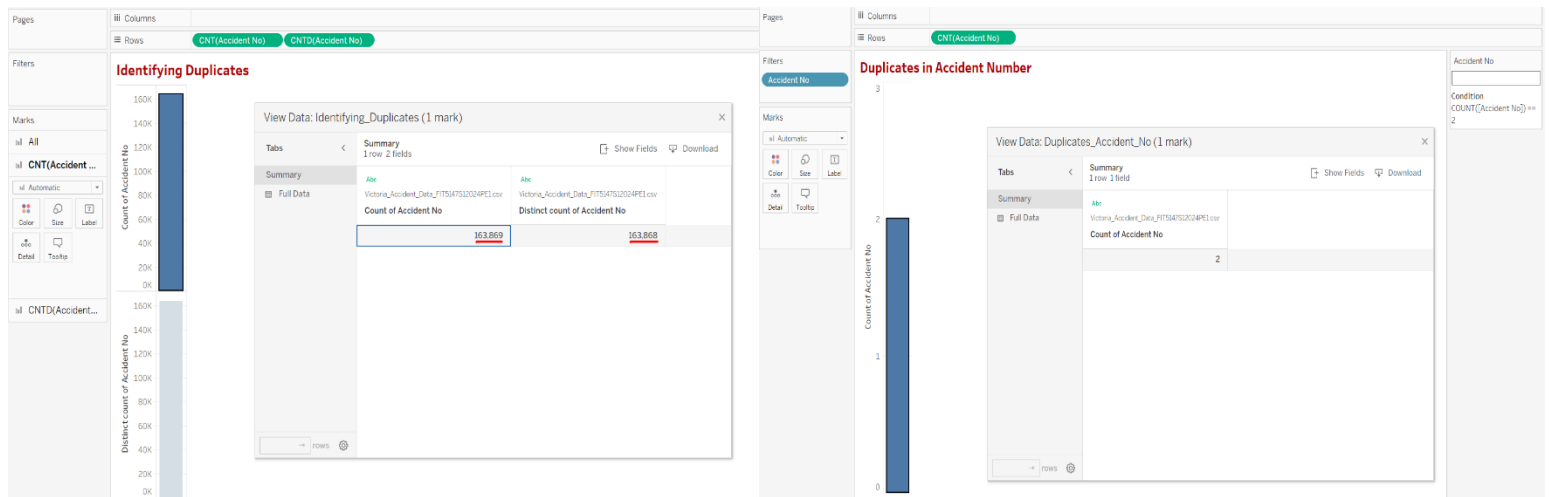
1. Data After Being Loaded to Tableau	2
2. Data Checking and Cleaning	2
1. Duplicates in Accident Number.....	2
2. Outliers in Speed Zone Attribute.....	3
3. Inconsistencies in a Single Attribute's Values.....	3
3.Data Exploration & Visualization.....	4
1. Contrast on when and what type of accidents occur over time (2012-2024).	4
2. Understanding Accidents Geographically with Spatial mapping of the data.	4
3. Where Did the Majority of Accidents Happen in the State of Victoria?	5
4. Accidents with respect to Geometry	5
Conclusion	6
References.....	6

1. Data After Being Loaded to Tableau:



2. Data Checking and Cleaning:

2.1 Duplicates in Accident Number:



Normal Count and Distinct Count of Accident Number With total number of duplicates.

With the Accident number being Unique it is evident that these number should not have any duplicates, by using filter and count features of Tableau I was able to differentiate and find the duplicate values in the accident number. The duplicate T20170014004 was searched and deleted in Excel.

2.2 Outliers in Speed Zone Attribute



Detecting Outliers by conducting Z-Score Analysis

The highest speed zone in VIC is 110 KMPH, anything more than this is clearly an outlier. Z-score analysis has been performed to visualise and depict the outliers. Significant data were deleted in the data sheet since the outliers count is negligible when compared to the whole data set count. Another remedy which has been used is that the outliers has been replaced with the median value of the speed zone using excel in order to avoid data loss. We can also use filter to limit the speed zone value from 30kmph to 110kmph instead of deleting the outliers.

Filter [Speed Zone]

Range of values At least At most Special

Range of values

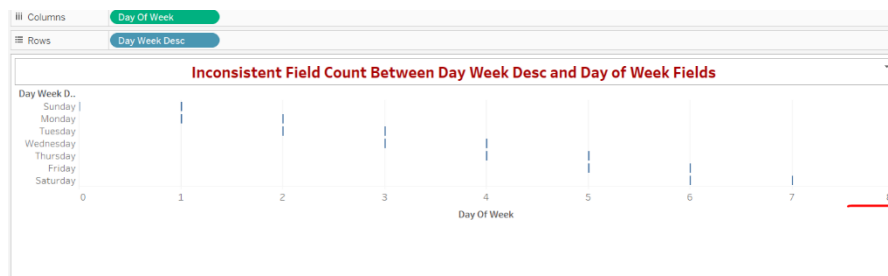
30 110

30 999

Show: Only Relevant Values ☐ Include Null Values

Reset OK Cancel Apply

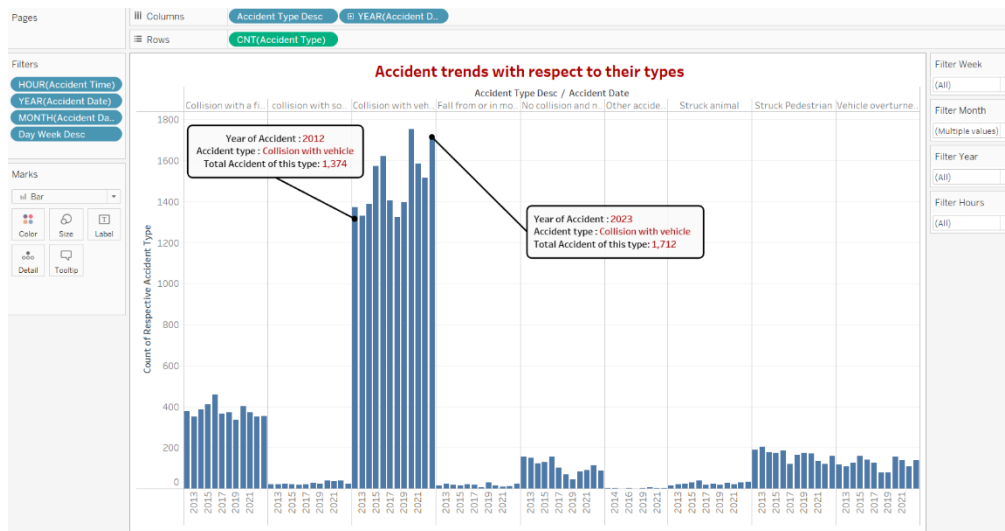
2.3 Inconsistencies in a Single Attribute's Values



A week has 7 days with Sunday being the start of the week. But the Data shows 8 days in a week instead. The extra day has been filtered and deleted due to negligible count using excel.

Data Exploration and Visualization:

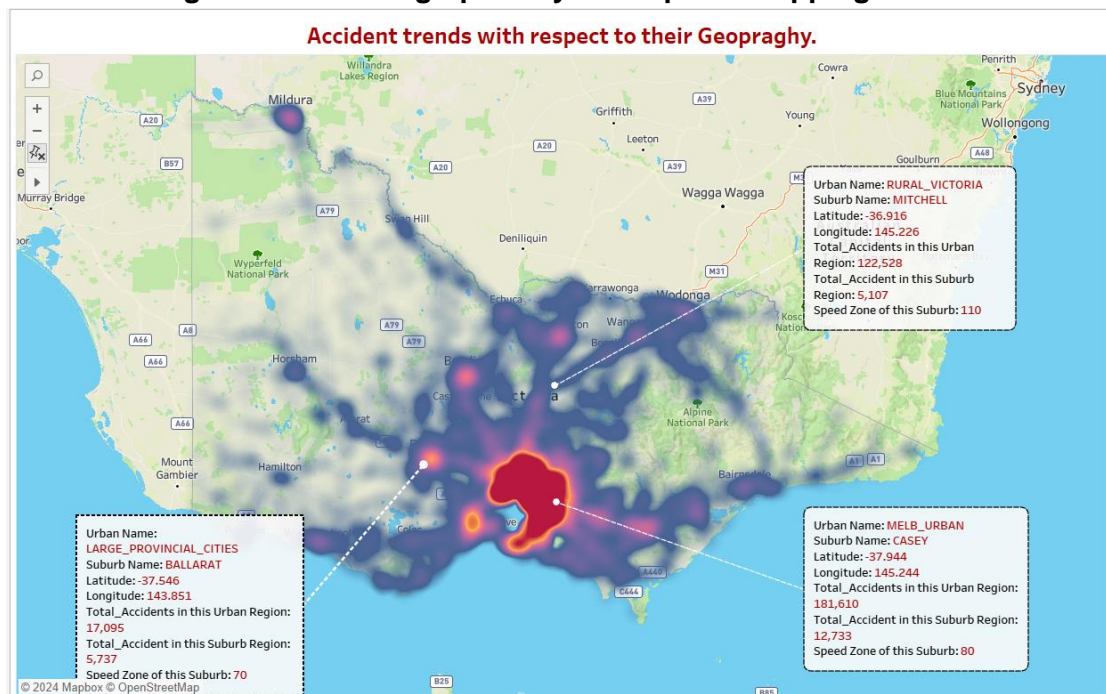
3.1 Contrast on when and what type of accidents occur over time (2012-2024).



Visualizing accident trends with respect to their type with usage of bar charts and filters

In order to understand what type accidents are high with respect to time periods, bar charts can be very helpful as they are pretty much straight forward while conveying the numbers. In the above graph it is very clear and evident that collision with vehicles resulted in more accidents across all years from 2012 till 2023 when compared to any other type of accidents in VIC. By the usage of the filters on the right we can further dive deep into months, weeks and hours.

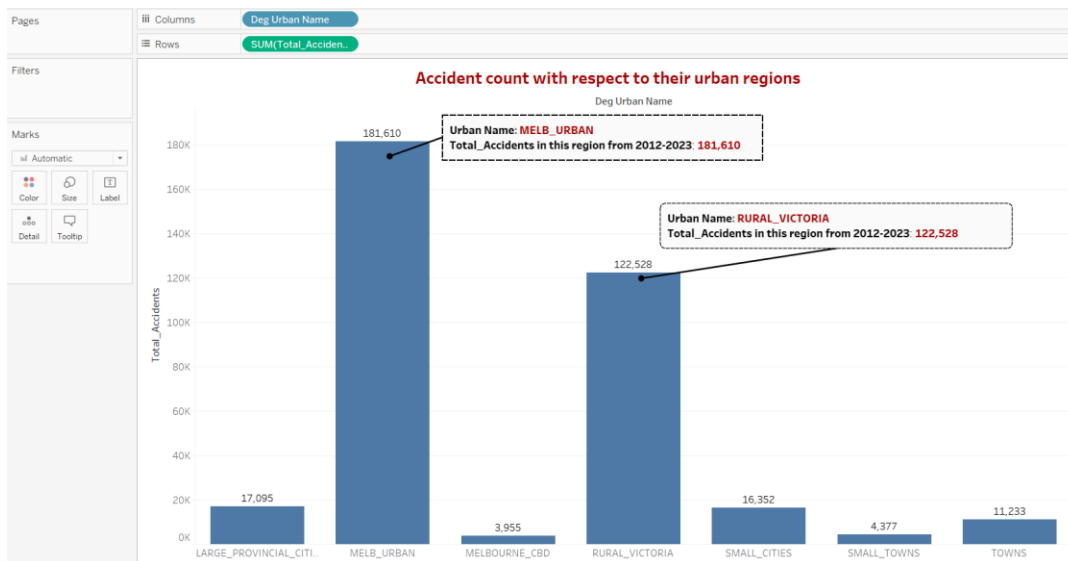
3.2 Understanding Accidents Geographically with Spatial mapping of the data.



Visualization of accidents by its region with usage of Density feature in tableau

Spatial visualization can be very easy and impactful to understand which region had more accidents and which regions had less. With the usage of filter, we can further squeeze down our conclusions while understanding the type of accidents that happened in those specific regions over time. From the above visualization we can clearly see how accidents have lined up towards the urban Melbourne region.

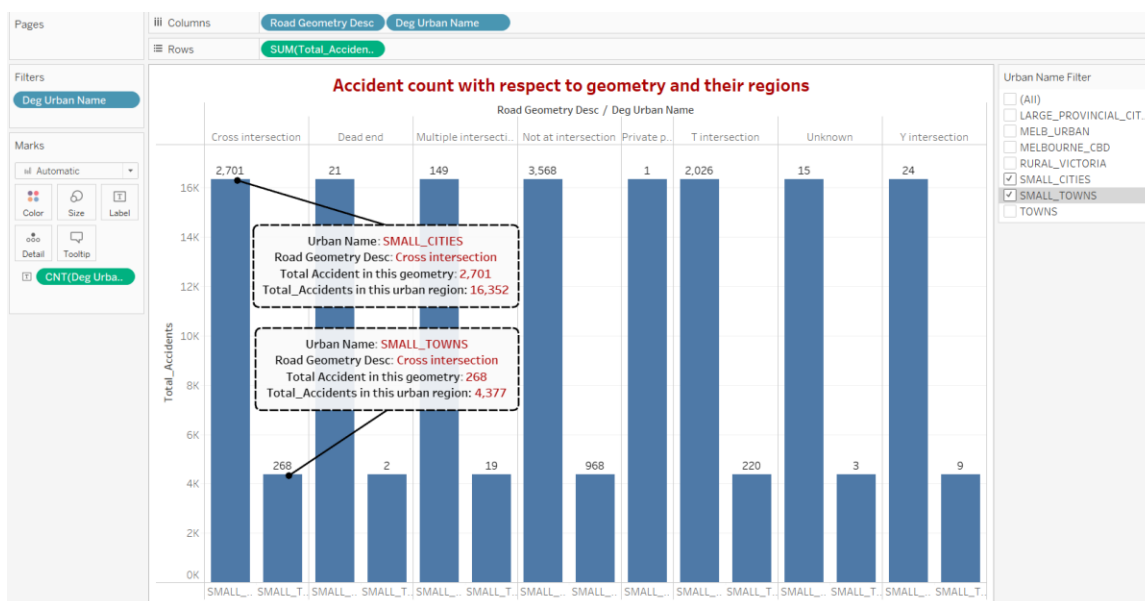
3.2 Where Did the Majority of Accidents Happen in the State of Victoria?



Visualization of total accidents with respect to Urban Locations

Once again, Bar graph make's things easier to understand which urban area dominated in accidents. An interesting observation is that though rural Victoria is spread wide across the state, majority of accidents have happened in the Melbourne Urban region.

3.2 Accidents with respect to Geometry.



Visualization of total accidents with respect to its Geometry

By the usage of filter, we can also understand what type of geometry has majority of the accidents with respect to its region. Thus, SOP's can be planned accordingly in order to avoid the same accidents that occur frequently.

Conclusion:

By exploring and analyzing the data that depicted VIC accidents it is very evident that the urban region faces more accidents when compared to the other regions. Though rural regions are spread wide across the state, accidents can be witnessed more as we move closer to the Urban regions. With the combination of filters, we can further understand the injuries and fatality rate that are caused as a byproduct which helps us to create strategies to evade accidents.

References:

- [Implementing density features using tableau.](#)
- [Understanding Geographic rules using tableau](#)
- [Data Filtering using tableau](#)