



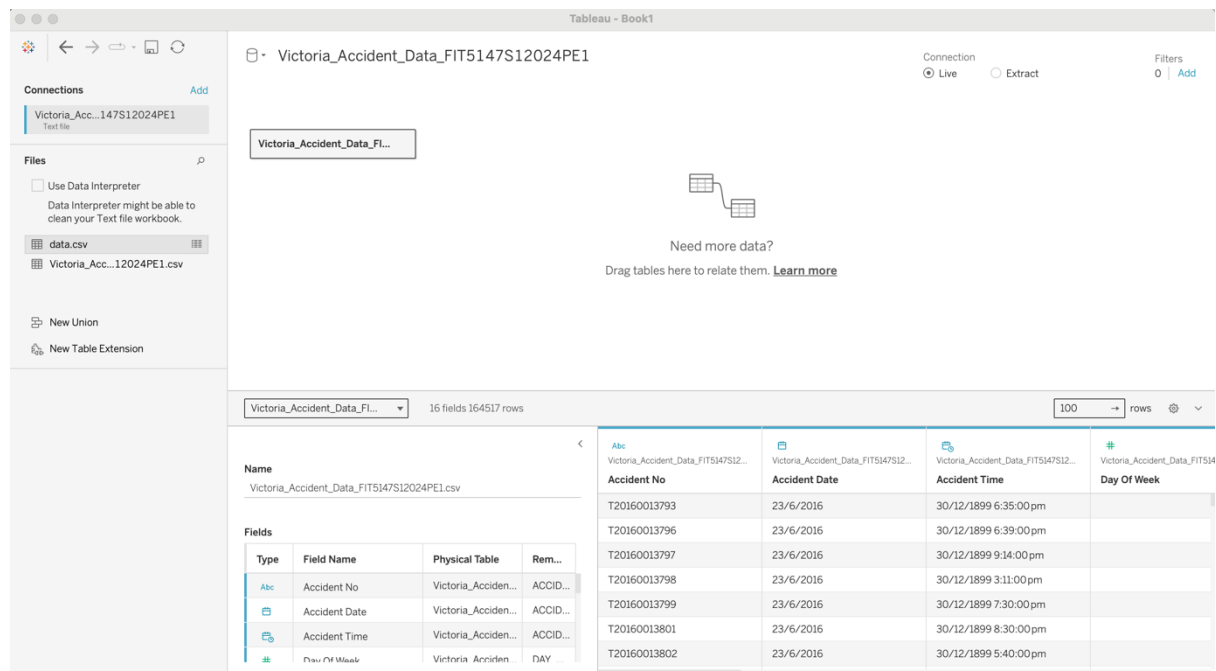
FIT5147 Data Exploration and Visualisation

Programming Exercise 1: Tableau

Bhavna Balakrishnan

33954437

Loading the Data Source-



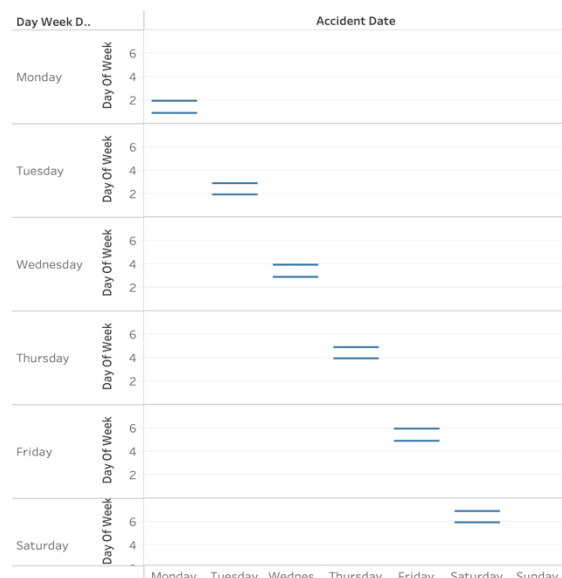
Data has been loaded into Tableau and the data type has been altered as per the need. Example, Accident Time was changed into Date-Time to visualise hours better. 37 Null values and 1 duplicate was removed from dataset prior.

ERRORS

Date with Days and Day Description Discrepancies

To ensure that all the days match the date. The below visualisation indicates that there are errors in both the date-day and the day number. It can be corrected in excel using the =TEXT(cell,"dddd") function and filtering.

Days of the week-



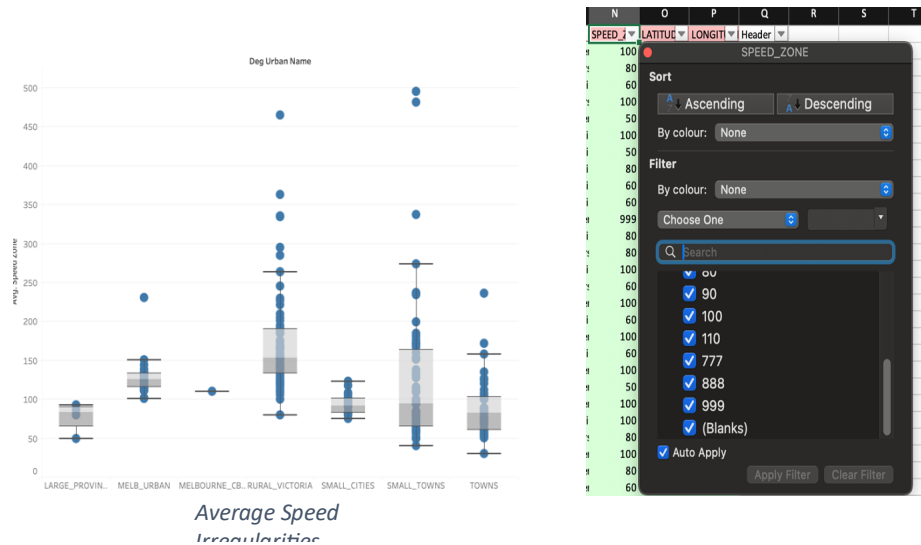
NUMBER	DAY
1	SUNDAY
2	MONDAY
3	TUESDAY
4	WEDNESDAY
5	THURSDAY
6	FRIDAY
7	SATURDAY

Figure 3: Corrected Data Descriptions

Here, we see the graph depicting that the days of the week have discrepancies in their description. For example, Monday is both day 1 and 2 which is incorrect. We will clean the data to correct this in excel by filtering out incorrect values and correcting them.

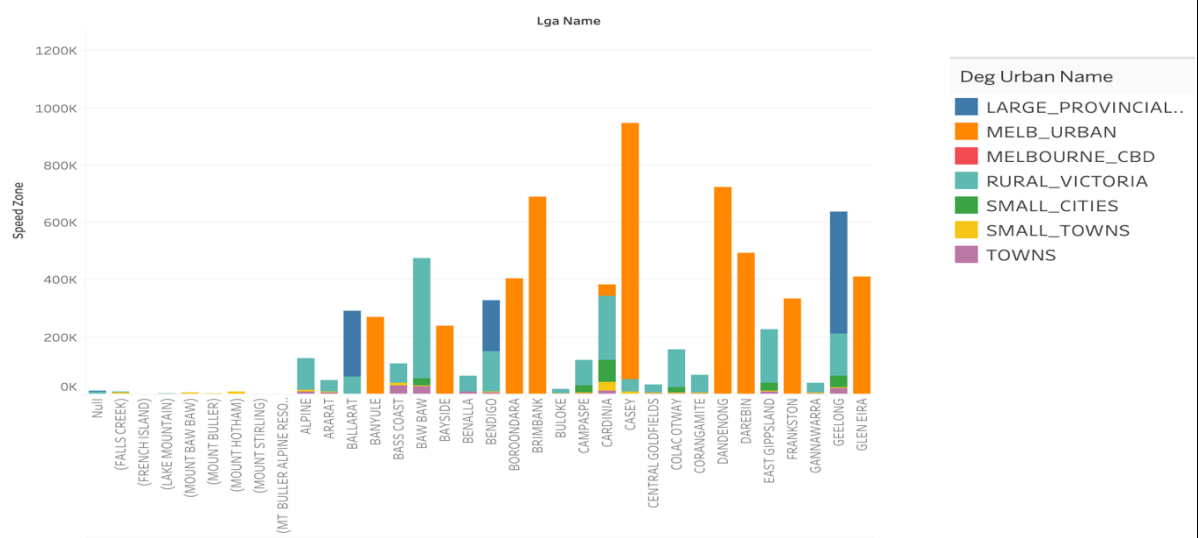
IRREGULARITIES: SPEED ZONE

Another irregularity that was found was in the speed zones which had values of 777,888 and 999. This is not possible since it exceeds the maximum speed at which a vehicle can travel at. Once again, we will filter out these values and correct them in excel by replacing them with the average speed from the other entries which is approximately 117. I am replacing instead of deleting because there are 529 irregular entries which is a significant number that should ideally be included in the dataset.



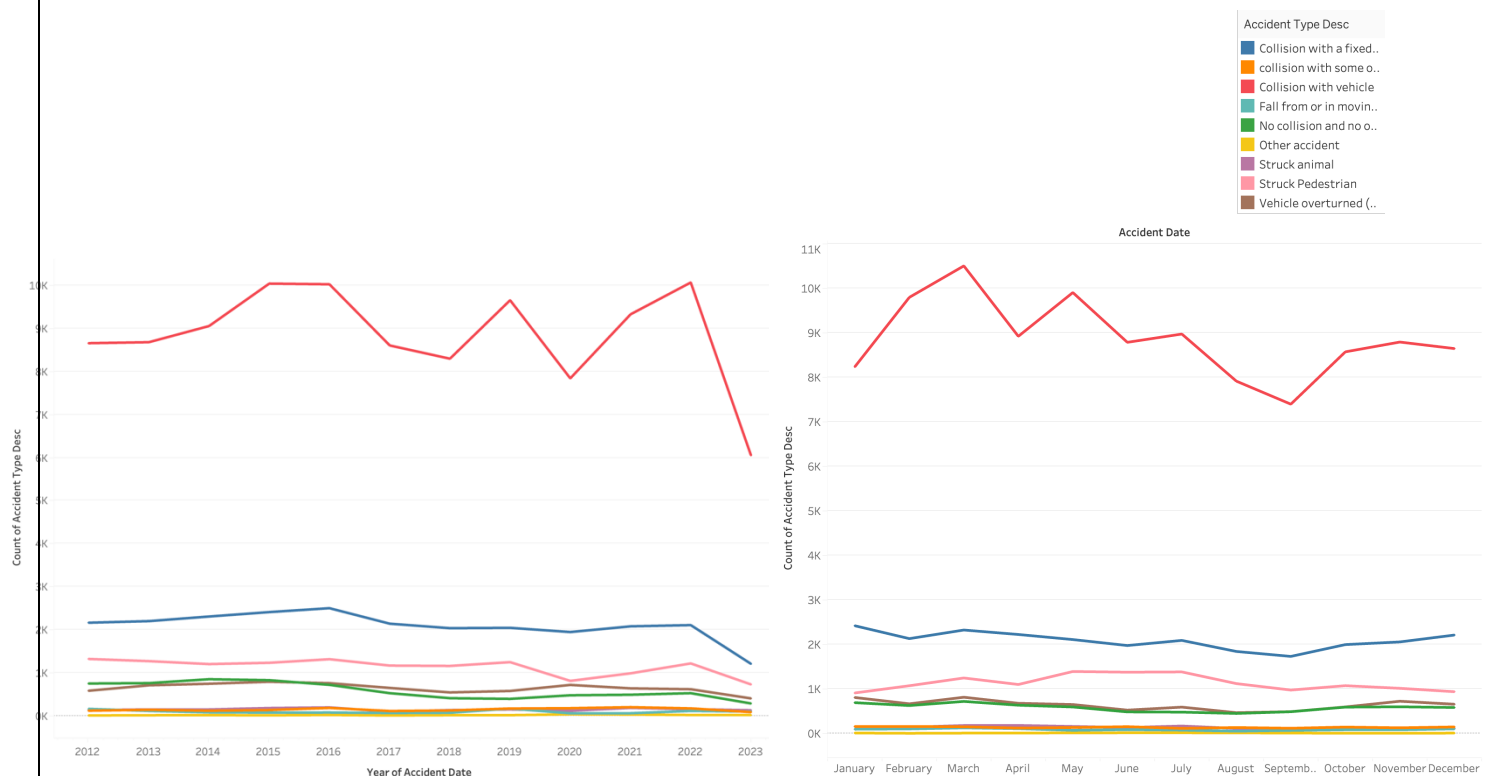
LGA Name Redundancy

Below we can see that there is a formatting difference so we should remove the brackets and null values as well as any duplicates in the entire dataset. Another thing is that there are redundancies like Mount Baw Baw and Baw Baw have to be merged. Similarly MT BULLER ALPINE RESORT to MOUNT BULLER.



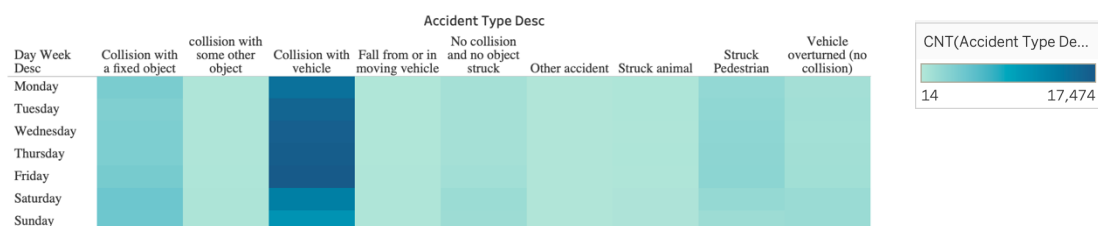
Q1. Accident Type Vs Time

Years and Months Vs Description



From above graphs we observe that the trends types of accidents remain same. The maximum type of accident happens to be a Collision with a vehicle and the least other. With maximum accidents happening in the years 2015, 2016 and 2022. Monthly data shows most of these accidents happen in the month of March.

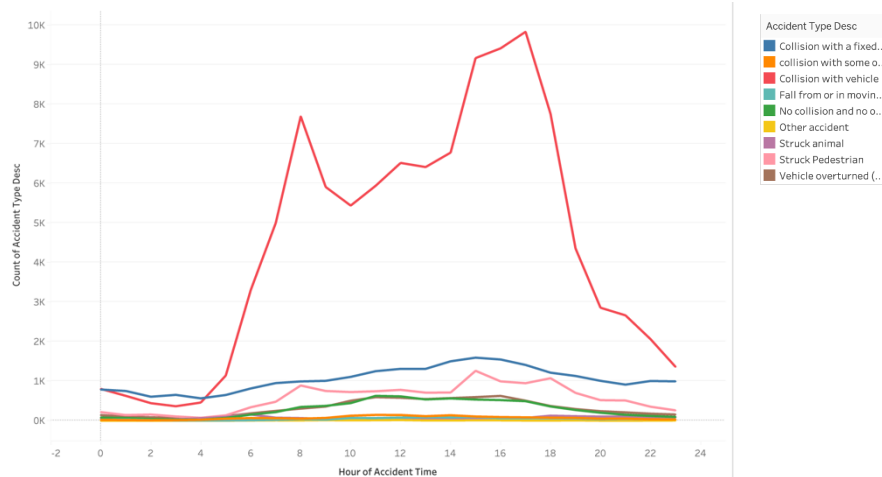
Days of the Week Vs Accident Description



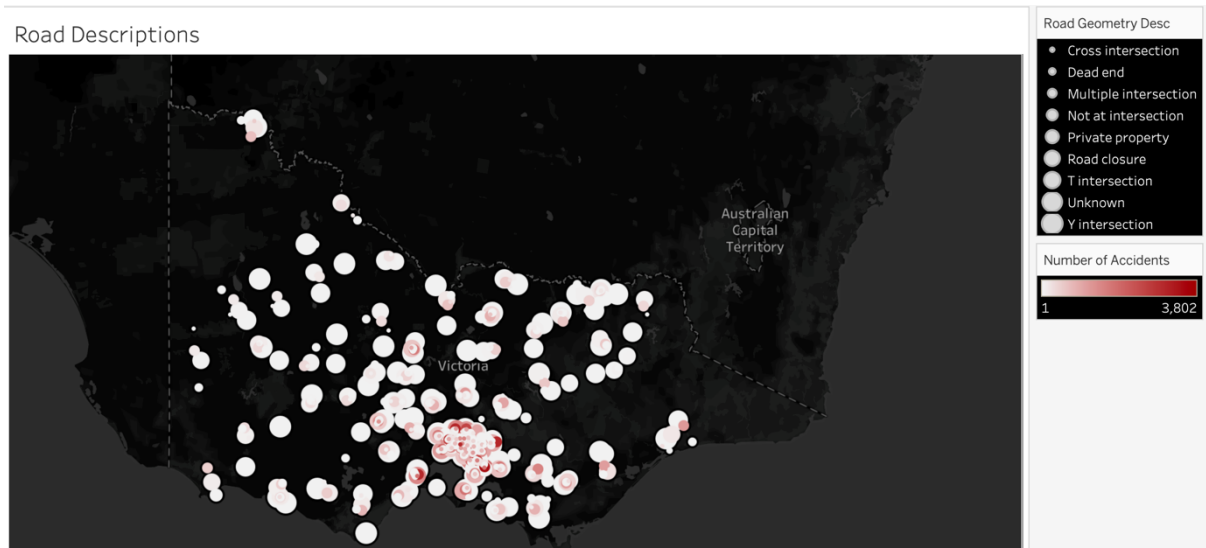
From above Choropleth, it can be concluded the maximum number of accidents typically happen on a Friday and that most of these accidents are collisions with other vehicles. It is highly possible that more accidents happen on Friday since it's the beginning of the weekend.

Hours of the Day Vs Accident Description

Below we can see that maximum accidents occur at 17 hours. We could possibly assume that this is caused by Office traffic of people on their way back home after a long day at work.

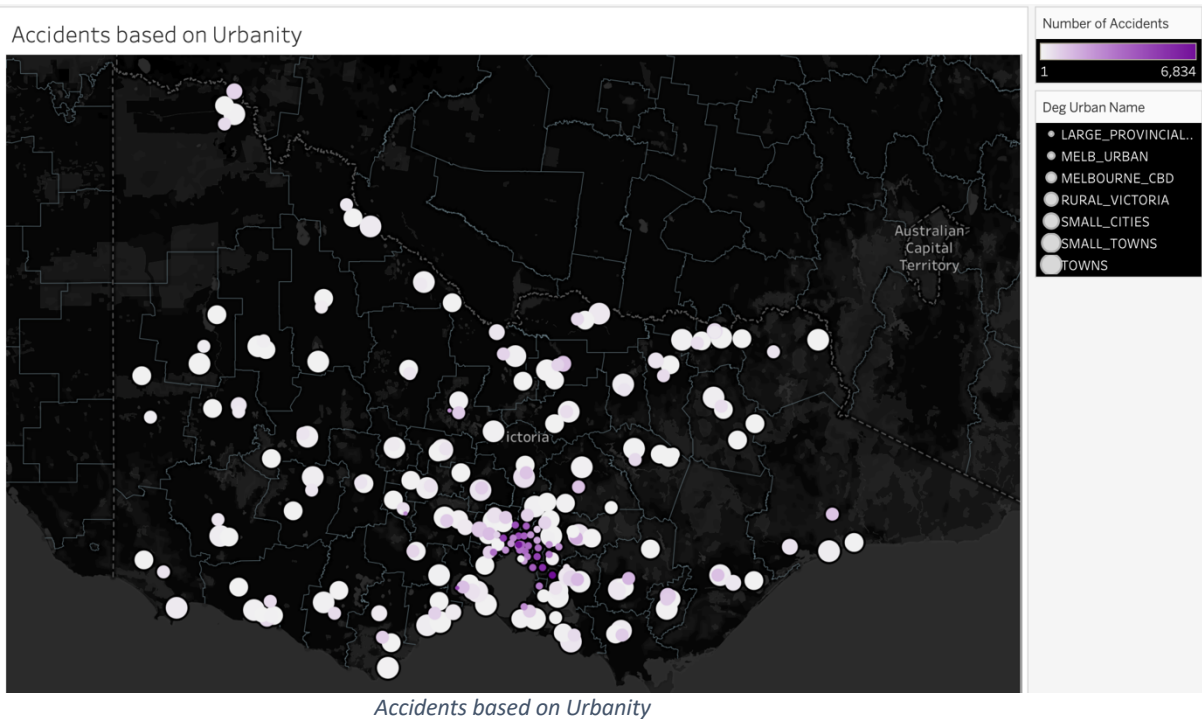


Q2: Geographic Mapping of Accidents

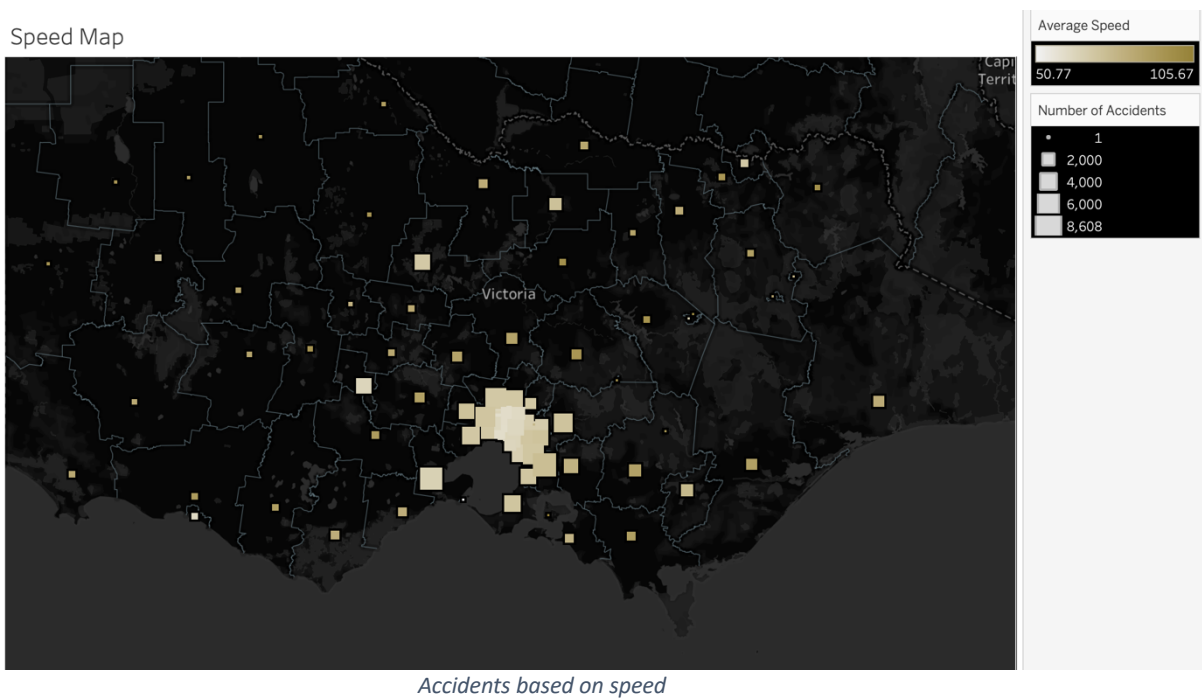


Accidents based on Road Geometry

From above, we observe that there is a mix of how many accidents are taking place based on the urban type of the place. The main pattern here is that most of the accidents are happening where there are no intersections.



From the above, we can observe that majority of the accidents are taking place more in Large-Provincial, urban places and Melbourne CBD. This makes sense since these places are more densely populated.



Puzzlingly, above shows that places with lesser average speeds have more accidents.