# Australian Fatal Crash Report (1989 to 2019)

## Introduction

Thousands of Australian die in road crashes each year. Although the number of deaths has been decreasing since it was first recorded in 1925, the Australian Government is still unable to keep the death toll below a thousand. By analysing the dataset provided from the Bureau of Infrastructure, Transport and Regional Economics (BITRE), common causes for road crashes can be identified in order to develop prevention measures for the improvement of Australian Road Safety and Regulations.

The aim of this report is to produce accurate and useful information from raw data to graphs and charts which could be visually seen and easily understood. To do that, it is essential to pre-process the dataset in order to extract important variables to reveal hidden factors for the maximisation of insightful information.

In this report, the methodology used for data pre-processing will be presented in Section 2. There are several data types in the dataset including numeric, nominal and spatial just to name a few, which will be described in Section 3. The generalisation of the results will be uploaded to the Kaggle notebook and presented in Section 4. Each graph and chart will have a clear explanation of the results and findings. Some recommendations for the future work will be discussed in Section 5 after presenting our research insight. Finally, Section 6 concludes our research for Australian Fatal Crashes between 1989 and 2019.

## Data Description

The study focuses on the fatal crashes on the states and territories in and around Australia. Based on the data updated by the Bureau of Infrastructure, Transport and Regional Economics (BITRE) on 14th every month where the data is preliminary and subject to revision. The study includes State (includes Australian jurisdiction), Time (Time of crash), Month(Month of crash), Year (Year of crash), Day/Week (Day of the week of crash), Crash Type (A crash is coded as a Pedestrian crash if a pedestrian was killed; otherwise, the number of vehicles involved), Number of Fatalities (fatal deaths in each crash),Speed Limit(Speed Limit on the road at the crash site), Road User Fatality(Driver, Motorcycle Rider, Passenger, Cyclist, Pedestrian and others), Gender(Male and Female), Age Group(Standard age group in 4 categories), Time of Day(plotted into 4 time frame and a histogram representation for each hour). Original data is available to the public at the Australian government [website](https://data.gov.au/dataset/ds-dga-5b530fb8-526e-4fbf-b0f6-aa24e84e4277).

This information is collected from the Australian Road Deaths Database and published by BITRE in two datasets as below:

* **Fatalities**: Each record in this dataset represents a person killed in the crash
* **Fatal crashes**: Each record in this dataset represents a crash

## Methodology

Data pre-processing is the method of removing ambiguous or noise data that may tamper the result. As a part of the data pre-processing, some data is altered on the two datasets, 'ardd\_fatal\_crashes.csv' and ‘ardd\_fatalities.csv’. For example, in 'ardd\_fatal\_crashes.csv', there is data where the column for speed limit exhibits a value which is '<40'. This value does not exhibit a constant value that can be used for data report generation. We replaced the value to 39 which satisfies the original data conditions. Similarly, the speed limits which are specified as 0 are also replaced to 'unspecified'. In addition to this, some rows that sound like noise data are removed from the dataset in order to avoid ambiguous reports. Generating a report is a complex procedure if the data given is not accurate and contains more noise data. In the Kaggle, we pre-processed the data by removing rows that have no value in it.

## Results and Findings

### Fatality per year per 100,000 People

As we can see in Figure 1, the number of fatalities all over Australia is represented for the population. In this instance, the fatality per 100,000 people is plotted and the rolling average of 3 years is plotted adjacent to the red line. It is so obvious that the fatality over the years has decreased to a level from more than 16% in 1989 to less than 6% by the end of 2019. It might be due to the strict laws and traffic regulations altered or changed over the years.

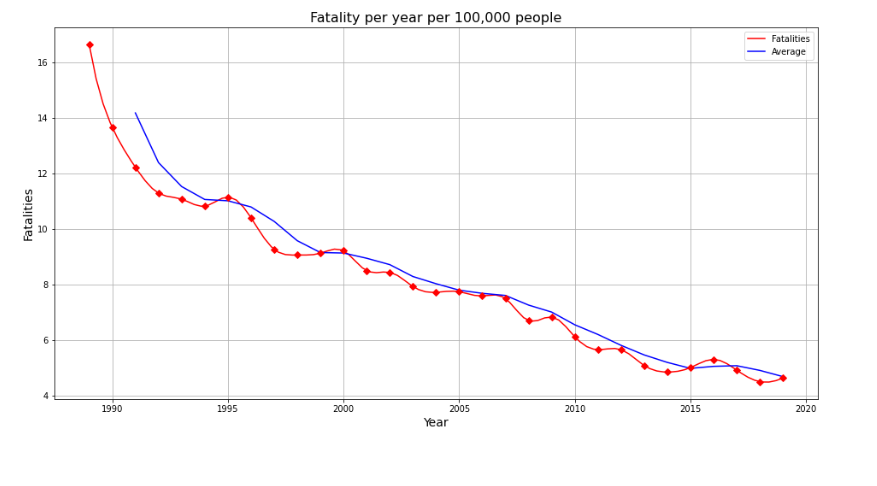


Figure 1. Fatality per year per 100,00 People

### Fatality by state per year per 100,000 people

All state-wise fatality per 100,000 people is represented in Figure 2. It can be seen from the graph that the fatality is 0 in ACT from 1995-mid 2000s. The data was not available to represent the fatalities in that period. However, it is surprisingly evident that Northern Territory (NT) has around 15 fatalities by 2019. All other states are demonstrating a downward trend to the fatalities.

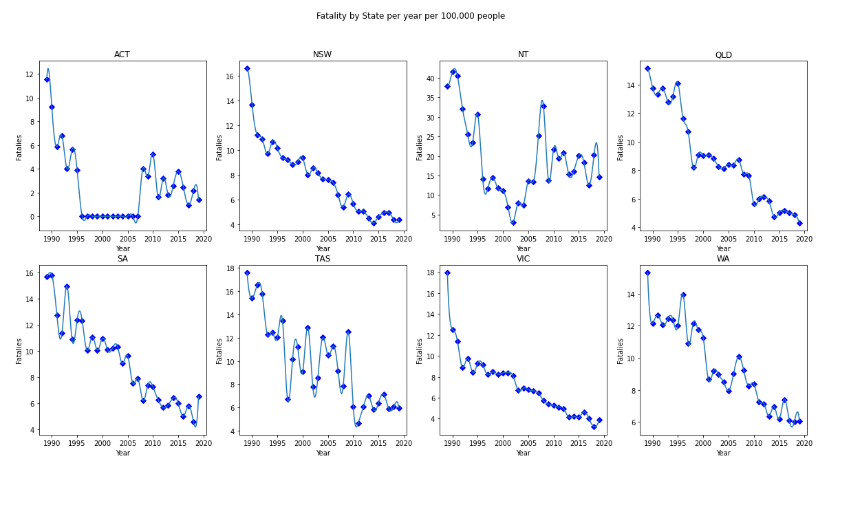


Figure 2. Fatality by state per year per 100,000 People

### Fatality rate by crash type

From the year 1989 to 2019, there are three categories in Figure 3 representing the fatality rate over the years. They are Multiple, Pedestrian and Single in each year. From multiple perspectives, it fluctuates between 40.02% and 43.81% with a margin of less than 4% over the years. The lowest rate was recorded in 2019 and highest in 2017. On the Pedestrian category, the percentage is twice lower than the Multiple, it looks quite steady between 13.25% and 15.62% which has less than 3% difference over the years. The lowest was recorded in 2017 and 2019 while the highest in 2018. The single fatality rate, on the other hand, has the highest percentage recorded amongst the three categories. The graph shows that between 42.9% and 46.73% which is almost three times greater than the Pedestrian. The fluctuation rate is just below 4% throughout the years, the lowest was recorded in 2017 and highest in 2019.

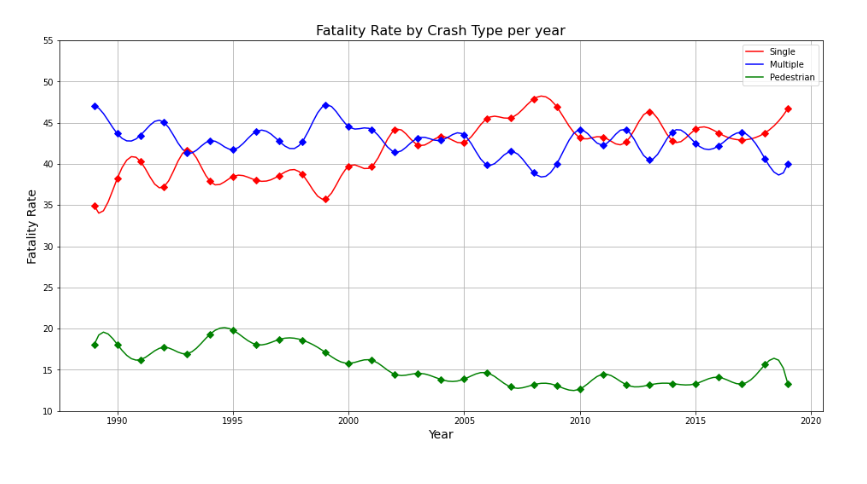


Figure 3. Australia Fatality Rate by Crash Type 1989 to 2019

### Road User Fatality Rate (1989 to 2019)

Knowing types of road users enable the government to make the road safer according to the user type. Figure 4 representing the fatality rate by crash type. There are six types of road users in our dataset, each type has its significance to the fatality rate. Begin from the highest to the lowest, Driver, represents 47%, the biggest contributor to fatality rate. The second is the Passenger category, which has 18%. The third is the Motorcycle Rider which is only 1% lower than the Passenger. The Pedestrian is the fourth with 14% in total. The fifth is Pedal Cyclist which has only 3% among all other categories. Lastly, there is 1% unspecified according to the dataset. From the graphs, it is evidently clear that the pedestrian fatality rate is more than 10% all around the course of time. This area must be taken into consideration while implementing the countermeasures.

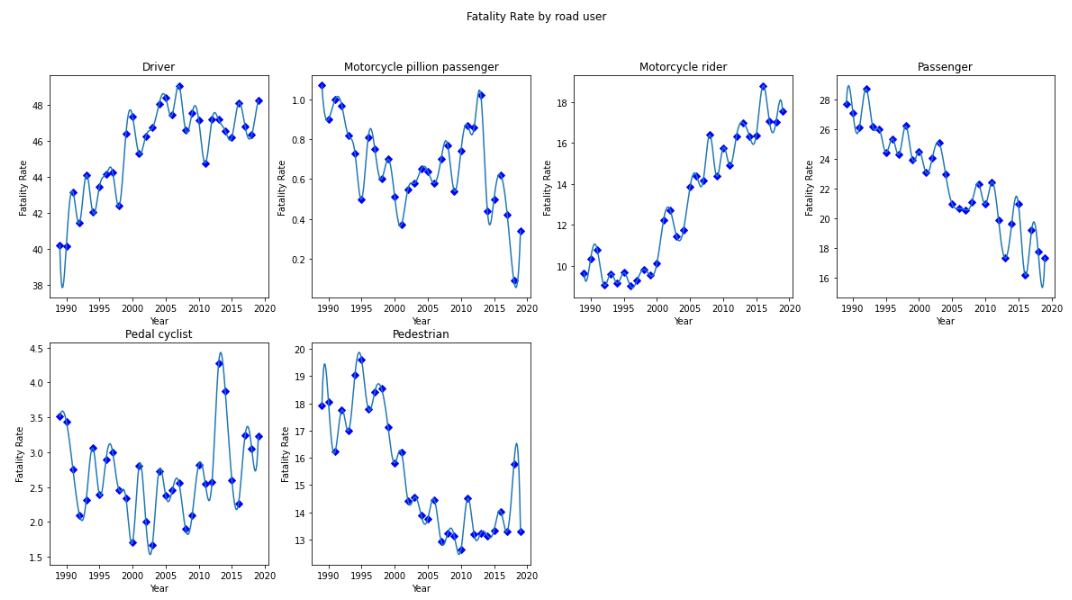


Figure 4. Road User Fatality Rate from 1989 to 2019

### Fatality by Speed Limit (1989 to 2019)

The speed limit is a substantial factor contributing to road crashes. Figure 5 shows that fatal crashes are more likely to happen above 75 km/h speed range and below 75 km/h. The fatality rate is more than 50% from Above 75-speed limit and this speed limit area might be a motorway that people attempt to drive faster. On the other hand, the local roads that have the speed limit of less than 75 were around 48%, in the beginning, and it is decreasing now and was 35% in 2019. This downtrend is due to the advanced speed cameras installed in different areas but limited to motorways.

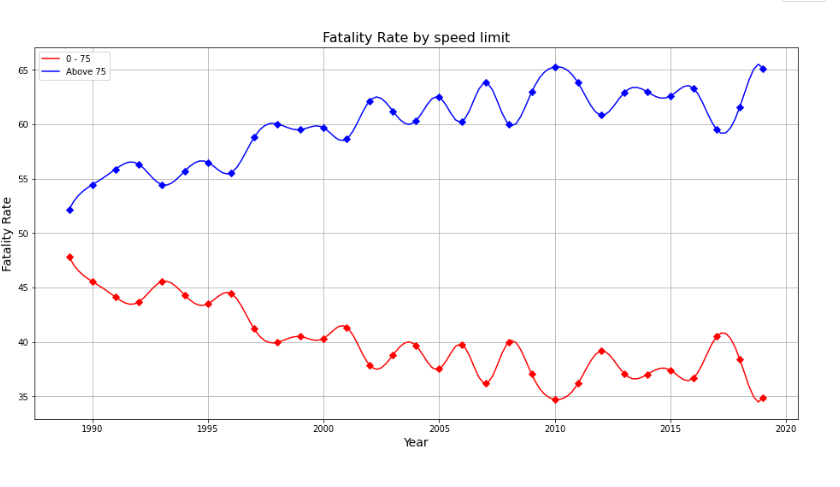


Figure 5. Road User Fatality Rate from 1989 to 2019

### Fatality Rate by time range

It is important to know the time frame in which the accidents are mostly happening. In order to achieve that, we have divided 24 hours into 4-time frames. From the graph of Figure 6, it is evident that the fatality rate is high during 12 PM-6 PM contributing more than 365 in 2019. In second graph, the fatality rate is high during 6 AM-12 PM which was initially less than 205 in 1989. Both time frames above mentioned are considered as the peak business hours in Australia. To further analyse and verify the above results, the rate of accidents every hour is plotted in a histogram form (00:00 to 23:00) every day from 1989 to 2019.

In Figure 7, fatalities are represented for every hour in order to verify the results. It is thus evident from the histogram that the fatality rate is increasing from 6 AM and reaches its maximum at 15:00 (3:00 PM) and starts to reduce from that point. However, the minimum fatality is approximately 1,000 at 04:00 AM and maximum is around 3,500 at 15:00 (3:00 PM).

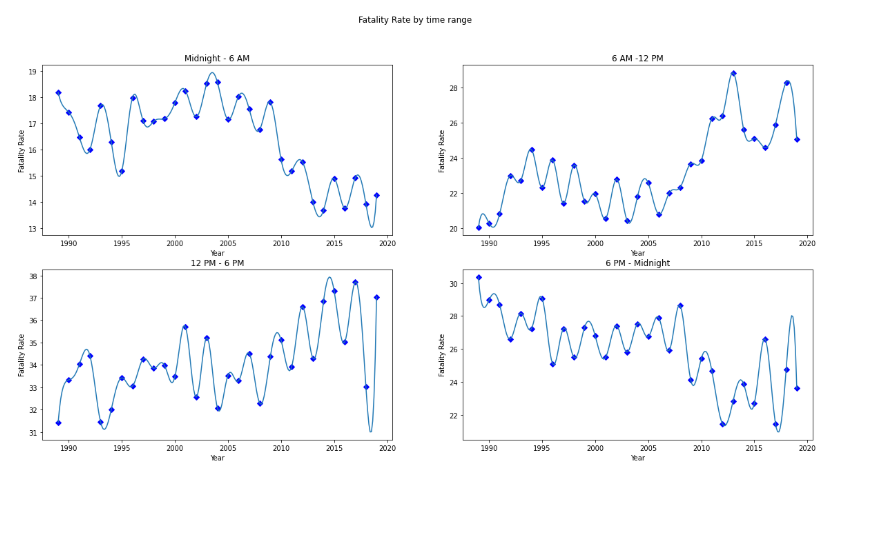


Figure 6. Fatal Crash by time range

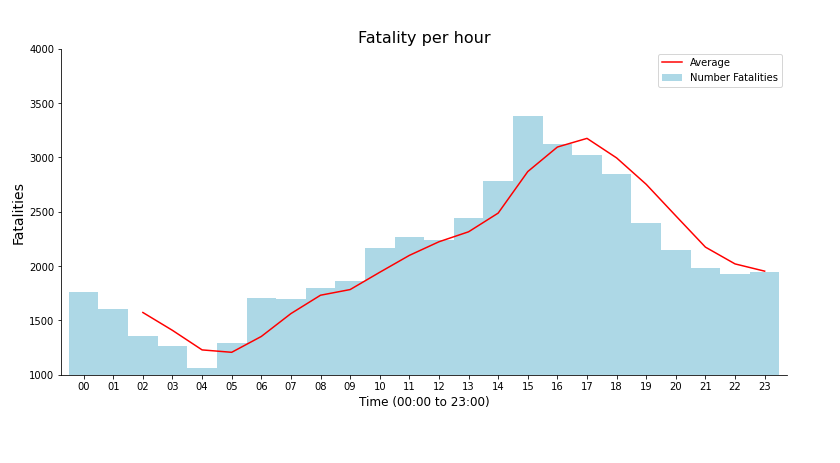


Figure 7. Fatality per hour

### Fatality Rate by Day of the Week

When it comes to understanding which day of the week fatal crashes is most likely to happen, the above bar chart illustrates evidently. Figure 8 representing fatality by day of the week. It reveals that from Friday to Sunday where the highest fatality rate was recorded from 15.44%, 17.53% and 15.57% respectively. The rest of the days during the week were below 13.1% in general, within them, however, Saturday has the highest fatality rate of 17.53%.

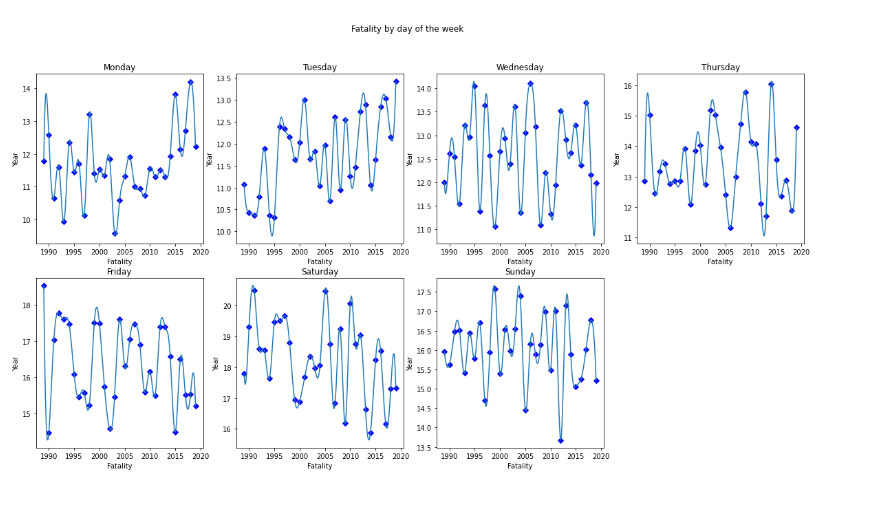


Figure 8. Fatality Rate by Day of the Week

### Fatality Rate by Age Group (1989 to 2019)

Driver’s age always plays a vital role when it comes to road crashes. Figure 9 shows the different age groups fatality rate over the years from 1989 to 2019. It can be easily observed that the age group between 51 and older has the highest fatality rate of over 37% by 2019 with an initial value of less than 25% in 1989. It also shows that the age group between 0 and 18 reduced from more than 18%in 1989 to 8% in2019. One interesting result should be noted that in the age group between 31 to 50 has increased dramatically and even peaked to more than 30% in 2016 and is now showing a downward trend. The age group between 19 and 30 was in its maximum of more than 34% and reduced with a sudden spike to more than 305 in 2005 and showed a decrease in the graph until now.

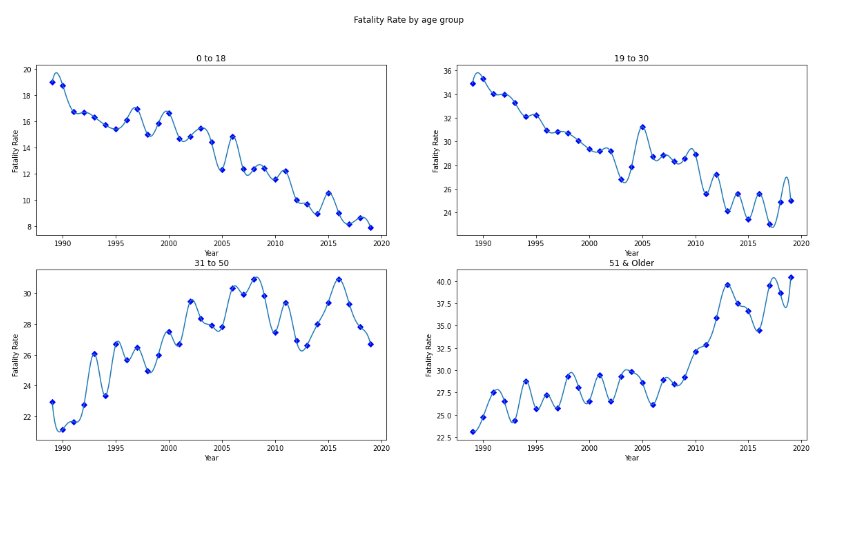


Figure 9. Fatality Rate by Age Group

### Fatality by Gender

In this category, the fatality between male and female during road crashes can be clearly seen in Figure 10 and 11. In general, female has a lower fatality rate than male. However, the gap between male and female has a significant difference. The fatality rate of male is almost three times larger than female in each year nationwide. The lowest fatality rate in the female category was recorded in 2019 with 23.43% while the highest in 2017 with 26.54% respectively. On the other hand, in the male category, the lowest fatality rate was recorded at 73.46% in 2017 while the highest in 2019 with 76.57% respectively. The margin over the years for both Male and Female categories is just over 3% which is considered steady. A rolling average on each graph is also represented to get a familiar view in the course of 3 years.

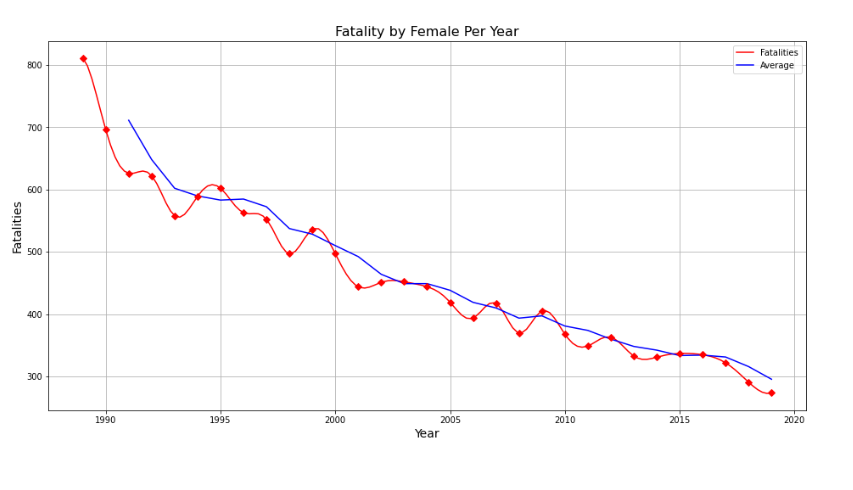


Figure 10. Fatality by Female per year

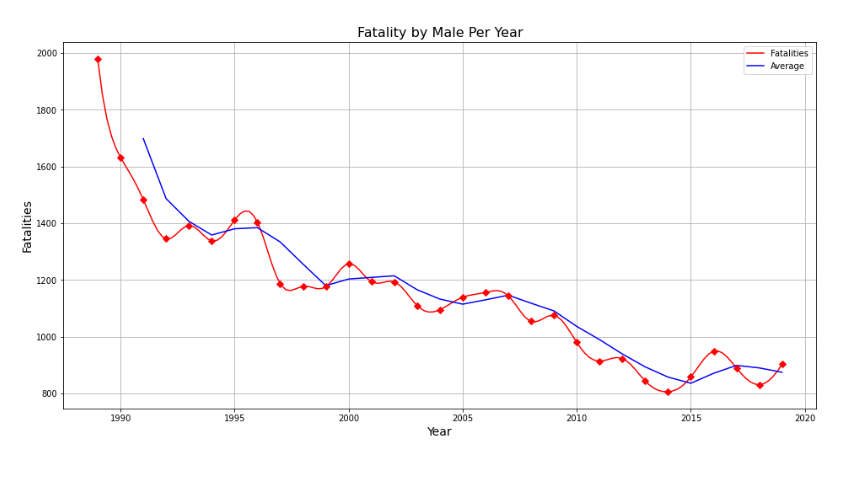


Figure 11. Fatality by Male per year

## Recommendation of Future Work

This report has identified several areas for the improvement of related work in the future. Firstly, within the datasets collected by ARDD, there are a considerable amount of missing values in several attributes which affects the integrity and accuracy of the result. Therefore, it is recommended that the recording of the dataset should be more thorough to improve the analysis of road crashes. Secondly, a new area of research which is called Machine Learning has been gradually introduced to this type of work. The clustering and association rules are two important machine learning techniques in finding patterns and categorising the similarities and dissimilarities which would be highly beneficial for the prevention of road accidents and the improvement of road safety. Thirdly, based on the results and findings, the speed limit was one of the significant factors for fatal crashes. It is recommended that more speed cameras should be installed in areas where accidents have been increased over the years. The recommendation also suggests that traffic rules should be regularly reviewed according to the improvement of public infrastructure and means of transportation.

## Conclusion

The land of Australia is vast, but the population is low. The demographic is therefore highly densified in one area and scattered in another. This report provides us with many insights in fatalities related to road crashes. It tells us that New South Wales has the highest fatality rate, pedestrians have the lowest fatality rate in terms of road crashes. In relation to vehicle involvement, Light vehicles contributed the majority of road crashes which is four times larger than the rest of other vehicles combined.

One interesting finding associated with the speed limit was that the fatality rate between speed limit above 75km/h and below 75km/h were almost identical. There was only a 2% difference between the two and the fatality rate was gradually dropping on the speed limit below 75km/h due to modern speed cameras. In terms of when the road crashes are more likely to happen, the result indicates that during daytime it was 1.5 times higher than night. In addition, the fatality rate was greater during weekends. For example, from Friday to Sunday with Saturday having the maximum fatality.

Age group and gender are important indicators when it comes to an insurance premium. According to the analysis, the most vulnerable age group is between 51 and older because they contributed over 30% of fatal crashes in average over the years. Female drivers are almost three times likely to be involved in fatal crashes compared to their counterparts.

## References

1. Bureau of Infrastructure, Transport and Regional Economics. (2020). Australian Road Death Database (ARDD). Retrieved from <https://data.gov.au/dataset/ds-dga-5b530fb8-526e-4fbf-b0f6-aa24e84e4277>
2. Budget Direct. (2019). Car Accident Statistics 2019. Retrieved from <https://www.budgetdirect.com.au/car-insurance/research/car-accident-statistics.html>
3. Harish S., Kevin J., Vikash S., Willis L. (2020). Australian Fatal Crashes Report from 1989 to 2019. Retrieved from <https://www.kaggle.com/kevinjames1993/australian-fatal-crashes-report-from-2015-to-2020>