2810ICT Assignment A    
**Executive Summary**

**Victoria State Accident Data Analysis**

**s5254936 -** Bebin Roy   
**s5182075 -** Atticus Burgess   
**s5299401 –** Ansh

Table of Contents

[1.0 Introduction 3](#_Toc147695846)

[Objective 1.0 – Information for Accidents in a Selected Period 4](#_Toc147695847)

[Objective 2.0 – Chart Display for Accidents Average 5](#_Toc147695848)

[Objective 3.0 – Information for Accidents with Specific Keywords (Collision) 6](#_Toc147695849)

[Objective 4.0 – Information for Accidents with Alcohol 7](#_Toc147695850)

[Objective 5.0 – Information for Accidents with Hit & Runs 7](#_Toc147695851)

# Introduction

This report has been conducted to analyse and cover the details for the Victorian State Accident Dataset, in which data has been collected from the periods of 2015-2020. As the data encompasses 5 years’ worth of accidents, the task that was upon us was to address 5 objectives that the data analysis software must be able to achieve. For the purpose of our analysis, we have selected the initial days and months of the years 2017 and 2018 as our set time period. This decision was made to ensure that our analysis would cover a full 12 months' worth of data in a single program run, providing a comprehensive view of the data collected during this time frame.

This specific time period was strategically chosen as it consistently recorded results and is situated precisely in the middle of the data collection timeline, thereby offering a representative sample of the overall data set. Our analysis was meticulously conducted with the following objectives in mind:

1. Locate and identify data on all accidents that occurred within the selected time frame, providing a clear overview of the accident landscape during this period.
2. Produce and interpret a chart that displays the average number of accidents occurring each hour of the day, offering insights into the time distribution of accidents.
3. Retrieve and analyse all accidents involving specified accident types that contain certain keywords, allowing for targeted analysis of specific kinds of accidents.
4. Analyse and understand the effects of alcohol on accidents, shedding light on the role of alcohol consumption in the occurrence of accidents.
5. Filter out and exclude incidents that were recorded as "hit & run" incidents, ensuring that our analysis is not skewed by these specific types of accidents.

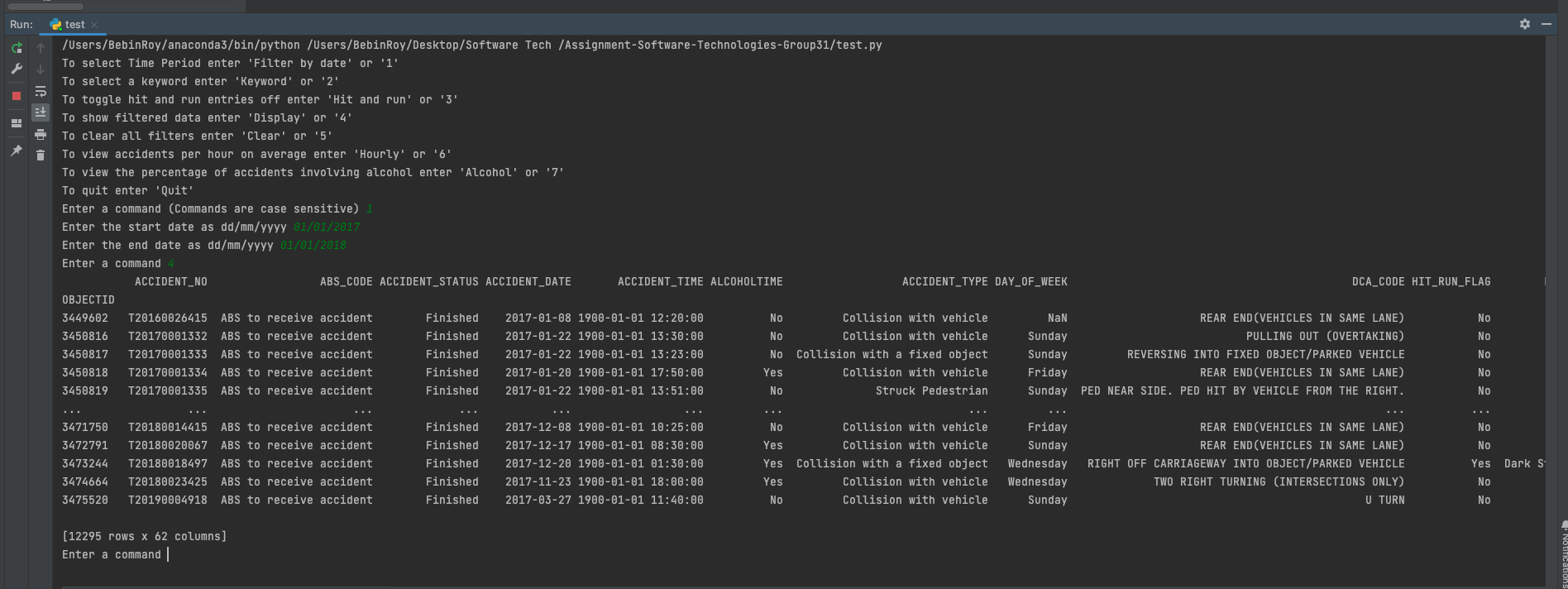
Through the careful execution of these objectives, our analysis aims to provide a thorough and insightful examination of the Victorian State Accident Dataset for the years 2015-2020, offering valuable insights and understanding of the accident trends and patterns during this period.

# Objective 1.0 – Information for Accidents in a Selected Period

Our data analysis program is efficient at identifying and allocating data within a user-specified time frame. For example, consider a period where we need to filter data between two distinct dates, spanning a twelve-month period. In this instance, the user would utilize the 'Filter by date' function, entering '01/01/2017' as the start date and '01/01/2018' as the end date.

Upon executing the 'Display' command post date selection, the system efficiently narrows down the data to specific rows that fall within the chosen time frame. For example, in one of the data rows, an accident numbered 'T20160026415' is recorded to have occurred on '2017-01-08' at '12:20:00'. This accident, categorized as 'Collision with vehicle', was finalized with the status 'Finished'. The accident type is further detailed as a 'Rear End' collision involving vehicles in the same lane, occurring during the day with police attendance, at a location not at an intersection, and resulting in other injuries. The speed zone of the accident location is recorded as '80 km/hr'.

The data set provides a comprehensive view of various accident types, including collisions with vehicles, collisions with fixed objects, vehicles overturning, and more, each with unique accident numbers and detailed information. For instance, another data row with accident number 'T20170001332' provides insights into an accident that occurred on '2017-01-22' at '13:30:00', involving a vehicle pulling out and overtaking, resulting in other injuries.



# Objective 2.0 – Chart Display for Accidents Average

A graph of blue bars

Description automatically generated with medium confidence

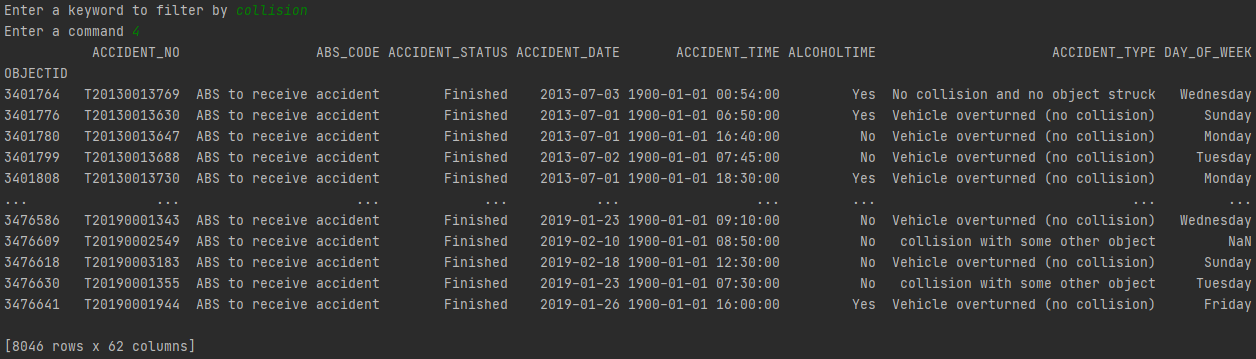
Our sophisticated data analysis program is adept at generating insightful visual data, particularly in The software tool had the ability to showcase the average number of accidents occurring each hour of the day within a user-defined period. This feature is invaluable for users seeking to understand the temporal distribution of accidents, providing a clear picture of the hours that are most prone to accidents.

For instance, in a scenario where a user is interested in analysing accident data over a specific timeframe, the first step involves selecting the desired period using the ‘Filter by date’ function. The user inputs the starting and ending dates, effectively narrowing down the data to the selected timeframe.

Once the period is set, the user can activate a function designed to calculate and visualize the average number of accidents occurring each hour of the day. The program meticulously processes the filtered data, calculating the average count of accidents for each hour and then presenting this information in a comprehensible chart.

# 

# Objective 3.0 – Information for Accidents with Specific Keywords (Collision)

  
  
Our data analysis tool is adept at identifying and visualizing the number of accidents occurring on with a specific keyword. For this, the user would can employ the ‘Keyword’ function. This function is designed to sift through the data and isolate rows that contain a specific keyword entered by the user. For the purpose of this example, the keyword used is ‘collision’.

Once the keyword ‘collision’ is entered and the ‘Display’ function is activated again, the software springs into action, meticulously searching and filtering the data. The result of this operation is a dataset reduced from the initial 8064 rows to a more focused 1685 rows of data. This reduction is significant, making the data more manageable and the analysis more precise.

Upon reviewing the displayed data, it is evident that each row is in some way associated with the keyword ‘collision’. This demonstrates the software’s capability to efficiently filter and present data that is closely aligned with the user’s specific requirements and input criteria.

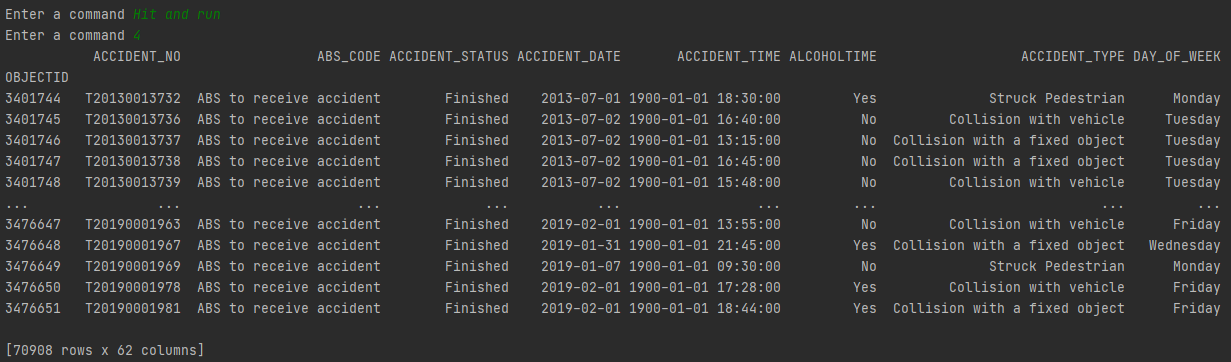
# A blue and orange pie chart Description automatically generatedObjective 4.0 – Information for Accidents with Alcohol



To further enhance our analysis, we proceed to apply the ‘Alcohol’ function. This function is designed to generate a pie chart based on the accumulated data specified between the starting and ending dates. The pie chart produced provides a visual representation with two distinct categories or keys: ‘Yes’ and ‘No’. The ‘Yes’ category refers to accidents where alcohol was involved, while the ‘No’ category represents accidents where alcohol was not a contributing factor.

This visual tool is invaluable as it offers immediate insight into the proportion of accidents related to alcohol consumption within the selected time frame, allowing for a more informed and detailed analysis of the data at hand.

# Objective 5.0 – Information for Accidents with Hit & Runs



In another insightful analysis conducted using our data analysis program, we focused on incidents where "hit and runs" were a significant factor. The objective of this analysis was to shed light on the prevalence and characteristics of accidents involving hit and run scenarios within a user-specified timeframe.

For this specific analysis, we engaged the ‘Hit and Run’ function embedded within our tool. This function is meticulously designed to filter and highlight accidents that were classified as hit and runs. Upon activation of this function, the program immediately sifts through the dataset, isolating and displaying only those incidents where hit and runs were identified as a contributing factor.