

Uncovering Insights from UK Car Accidents

Abhi Zanzarukiya ^{*1} and Heli Mistry ^{†2}

¹University of Waterloo, Student Number: 20892635

²University of Waterloo, Student Number: 20985952

February 19, 2023

Road accidents are a major public health and safety issue, and understanding the factors that contribute to accidents is crucial for developing effective strategies to reduce their occurrence and severity. The **UK Car Accidents 2005-2015** dataset provides a wealth of information about road accidents in the UK, including detailed information about the location, time, and circumstances of each accident, as well as the characteristics of the vehicles and drivers involved.

For this project, we plan to design a clean and efficient database and apply data mining techniques to the dataset. By cleaning and restructuring the data, we plan on presenting the following findings:

1. Predicting Accident Severity: We aim to build a predictive model to determine the severity of a car accident based on factors such as weather, road conditions, location, time of day, and type of vehicle involved.
2. Identifying High-Risk Areas: We plan to identify areas where car accidents are more likely to occur, as well as the types of roads that are most and least accident-prone in each region.
3. Identifying Accident Types: We aim to identify different types of accidents, such as rear-end collisions, head-on, off-side, or near-side collisions in the accident hotspots.
4. Pedestrian Casualties: We will identify the locations where pedestrians were the casualties in accidents as a result of their failure to adhere to road crossing regulations, ranking them based on severity.
5. Identifying High-Risk Driver Profiles by Vehicle Type: We aim to investigate how the characteristics of drivers, such as age and gender, affect the likelihood and severity of accidents for different types of vehicles.

In addition, we will provide the client application in the form of a CLI interface which will facilitate CRUD operations of the database. We aim to provide CLI shortcuts for common types of queries that users may need to perform regularly, such as listing all accidents on a specific date or within a certain kilometer radius of a given lat-lon value.

We are still exploring the dataset further to see any other helpful information that can be drawn from it. Overall, this project has the potential to make a significant contribution to public health and safety by providing valuable insights into the causes and consequences of road accidents in the UK, and by identifying effective strategies for reducing their occurrence and impact.

^{*}azanzar@uwaterloo.ca

[†]h23mistr@uwaterloo.ca