Group 1 Project 2 Project Proposal

Team members: Hibaa Aldubai, Farjana Rowther, Astha Nayak, Vivian Nnadozie, Abraham Ofolu

UK Road Traffic Collision Dataset

Data Extraction

2 files

1 – accident\_data.csv

2 – vehicle\_data.csv

Citing Data Source

Both sets of data to come from the following site:

[UK Road Traffic Collision Dataset | Kaggle](https://www.kaggle.com/datasets/salmankhaliq22/road-traffic-collision-dataset)

Data Transforming (cleaning, joining, filtering and aggregating)

* Removing unwanted columns

(keeping from vehicle-data – Age\_Band\_of\_Driver, Age\_of\_Vehicle, make, model, sex\_of\_driver)

(keeping from accident\_data – accident\_severity, day\_of\_week, light\_conditions)

* Removing empty cells
* Joining on “Accident\_Index”

Loading Data into Database

* Loading data on Postgresql (relational database)

Findings

From the CSV files, we would be able to extract data about accidents in the UK to include vehicle information such as the make and model as well as the gender and age of the driver. From this we would be able to analyse the data to see if there are any patterns or correlations. We will also be able to look at the type of weather during the time of accidents and day of the week.