Road Casualty Data Analysis 2022

Introduction

- Road accidents are traumatic experience for everyone and almost every government try to find its causes in order to reduce the casualties.
- In 2022, the UK reported over 61k casualties caused by road accidents. Of those, nearly 800 people were killed and over 11k people had serious injuries.
- The aim of this project is to understand the nature of these casualties, analyze its data to find probable causes and try to suggest ways to reduce them.
- The project applies statistical methods, visualizing methods and machine learning to extract useful insights and predict serious casualties in advance.

Casualty terms

- The dataset contains information about age, and sex of casualty, whether they are pedestrian, driver or passenger, severity of injury, etc.
- Some important features:
- casualty_class: Driver/Pedestrian/Passenger
- sex_of_casualty: Male/Female
- age_of_casualty: integer number
- age_band_of_casualty: 0-5/6-10/11-15/16-20/21-25/6-35/36-45/46-55/56-65/66-75/75+
- casualty_severity: Fatal/Serious/Slght
- casualty_type: 31 situation where the casualty accured
- casualty_home_area_type: Urban area/Small town/Rural
- casualty_imd_decile: 10 classes from the Most deprived 10% to the Least deprived 10%

Data Pre-processing

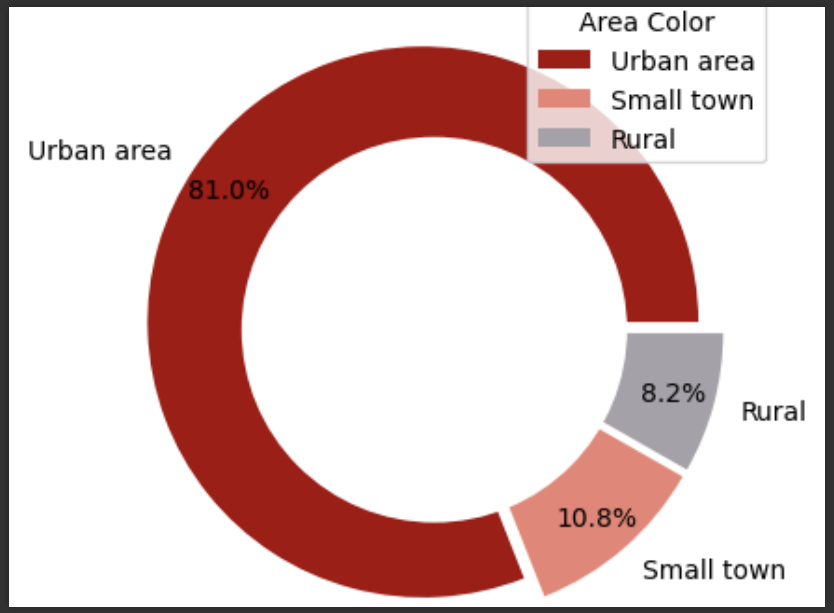
- Dropping unnecessary columns
- Missing or duplicated data
  - Checking for duplicated entries(the dataset did not have one)
  - Checking for NaN or null (the dataset did not have one)
  - Missing data encoded as -1: removing these entries results in better analysis and insights
- Re-encoding the values
  - All data were categorical but encoded in integer numbers which is not suitable for visualizing. So the values were re-encoded into meaningful strings.

Methodology

- As the data was almost clean, no special data cleaning is required
- Presenting key variables in the dataset with the help of visualizing
- Exploratory analysis of road casualty data

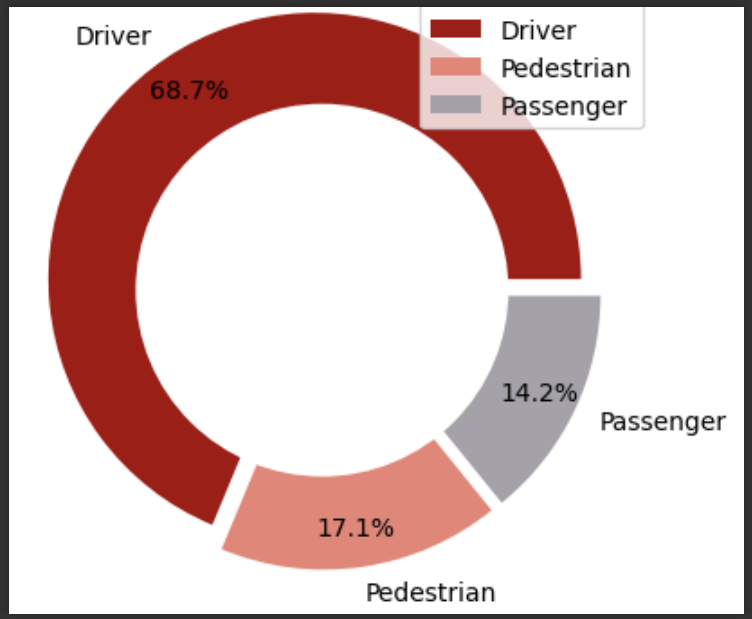
Results at a glance

Total Casualties
61352



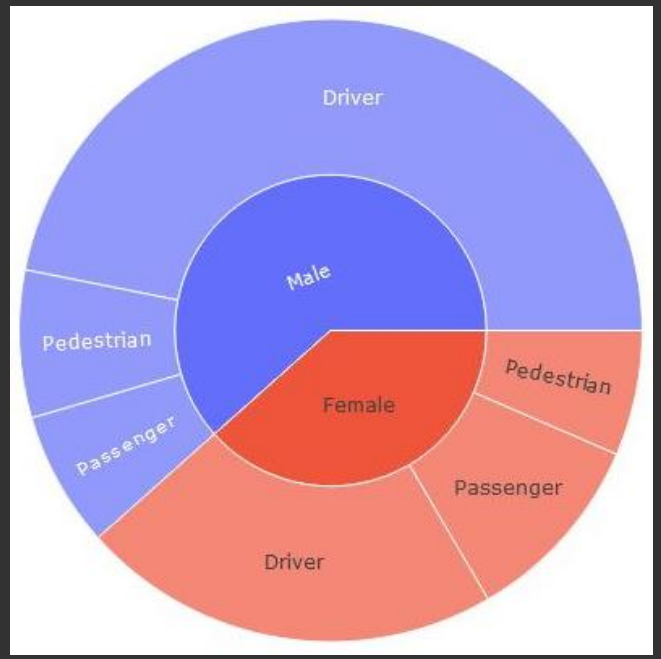
Casualties by Area
81% Urban Area

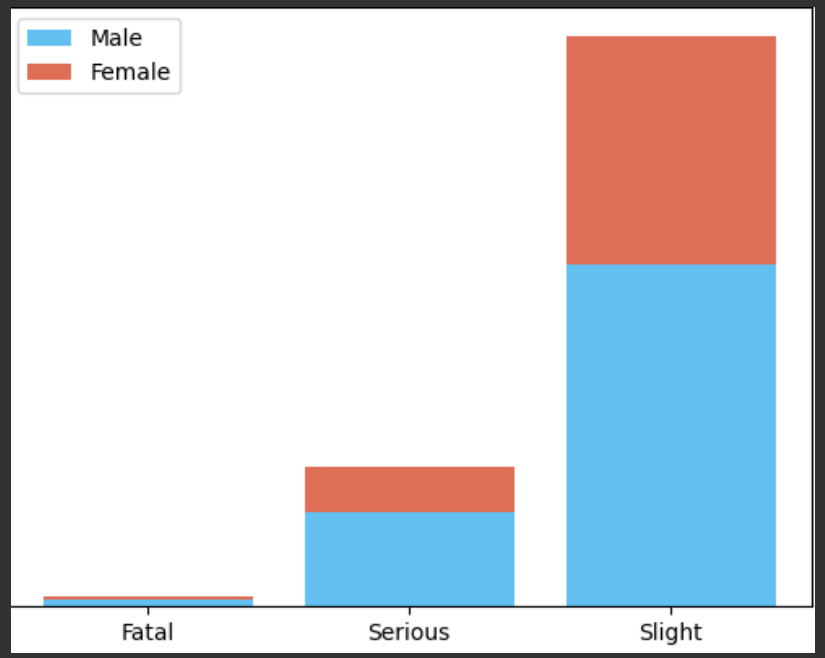
Casualties by Class
68.7% Driver



Results at a glance

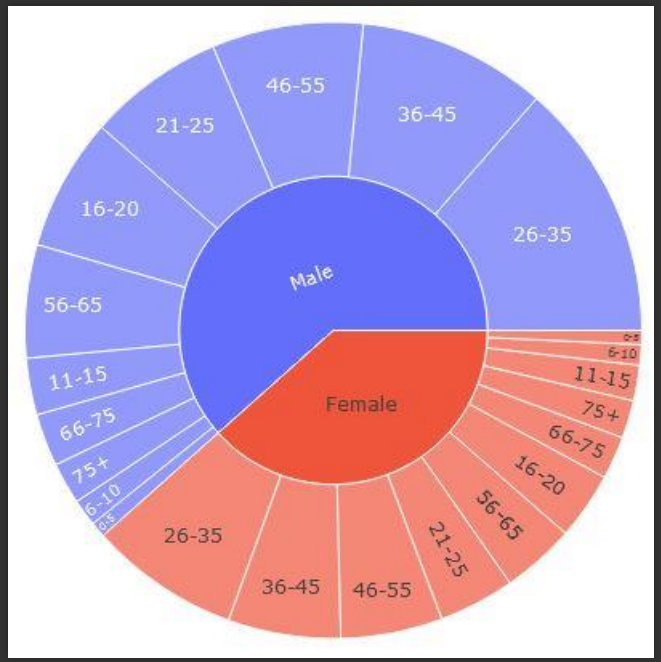
Casualty Sex and Class
Driver Men & Passenger Women





Casualty Sex & Severity

Casualty Sex and Age band



Results

- The most endangered citizens are men drivers with ages between 26-35
- More men were injured as drivers, while more women were injured as passengers
- Beside car occupants, most injured citizens are cyclists and pedestrians
- Pedestrians are more vulnerable and a great deal of Fatal casualties happen to them.

Conclusion

- The government should focus on educating men drivers with ages between 26-35 as they are the most endangered citizens and probably the main cause of road accidents!
- The government should provide more safety for pedestrians as they are more vulnerable to fatal casualties
- The most casualties happen in urban areas, which is reasonable.