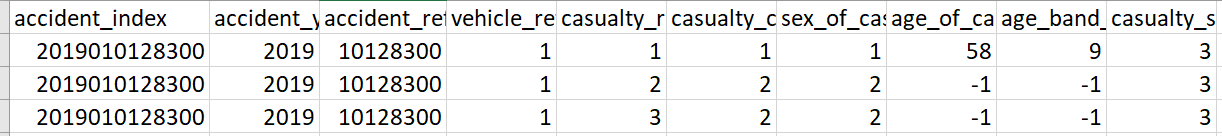
**Analysis Case Study On Road Safety Data For England**

1. **Introduction**

Safety is a key issue for transport policy globally. The following is an analysis case study trying to investigate the occurrence of road traffic accidents in England at a national scale. STATS19 data for road accidents, vehicles involved in road accidents and casualties occurring during 2019 were analysed and visualized using various statistical techniques. The main aims of this study were to investigate the patterns of road accidents, casualties, and vehicles and give a suggestion for road safety management.

1. **Data quality**
   1. **Misleading**





Some records have the same accident index, which may refer to different persons injured in one particular accident.

* 1. **Data cleaning**

As the raw data is for the whole of the UK, the data was filtered as it was only needed to analyse data for England. Figure 1 shows the distribution of accidents based on the raw data and Figure 2 shows the distribution of accidents in England.

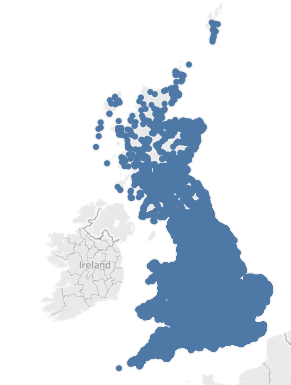
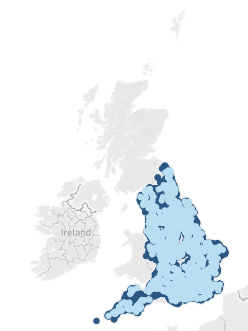
 

Fig.1.Distribution of accidents across the UK Fig.2.Distribution of accidents across the UK

1. **Data characterization**

**3.1 Accident Data**

In road safety, an accident is an incident occurring on a public highway which involves at least one vehicle. The “dft-road-casualty-statistics-accident-2019.scv” contains information on accidents. It contains all kinds of various attributes about the whole accident.

The accident\_index is a unique value for each accident. The accident\_index combines the accident\_year and accident\_reference to form a unique ID. It can be used to join Casualty tables.

The severity of an Accident is defined as the highest severity of the injury suffered by any resultant casualty. Accident Severity can be shown as a separate table based on Vehicles or Casualties. It should not be confused with Casualty Severity, which is shown only in reports based on Casualties.

All the Accident attributes have lookup tables related to them which refer to the meaning of each field. Detailed description of each variable is shown in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **DESCRIPTION** | **RANGE** |
| accident\_index | TEXT | Unique value for each accident | None |
| accident\_year | INT | Years on which accident took place | 2019 |
| accident\_reference | TEXT | Unique value for each accident | None |
| location\_easting\_osgr | FLOAT | Location easting on Ordnance Survey National Grid | [64084.0,655244.0] |
| location\_northing\_osgr | FLOAT | Location northing on Ordnance Survey National Grid | [ 10814.0  ,1167366.0] |
| longitude | FLOAT | Longitude of location at which accident took place | [ -7.52528,  1.757476] |
| latitude | FLOAT | Latitude of location at which accident took place | [ 49.917760,  60.388637] |
| police\_force | INT | Looks up the police force at which accident took place | [1,99] |
| accident\_severity | INT | Refers to the severity class of accident | [1,3] |
| number\_of\_vehicles | INT | Number of vehicles involved in accident | [1,17] |
| number\_of\_casualties | INT | Number of casualties involved in accident | [1,52] |
| date | TIMESTAMP\_NTZ | Date on which accident took place | 01/01/2019­-31/12/2019 |
| day\_of\_week | INT | Day of week on which accident took place | [1,7] |
| time | TIMESTAMP\_NTZ | Date at which accident took place | 00:00-23:59 |
| local\_authority\_district | INT | Looks up the local authority district at which accident took place | [1,941] |
| local\_authority\_ons\_district | TEXT | ONS( Office for National Statistics) location of local authority district at which accident took place | None |
| local\_authority\_highway | TEXT | Looks up the local authority highway at which accident took place | None |
| first\_road\_class | INT | Refers to the first road class on which accident took place | [1,6] |
| first\_road\_number | INT | Refers to the first road number on which accident took place | [-1,9999] |
| road\_type | INT | Refers to the road type on which accident took place | [-1,12] |
| speed\_limit | INT | Speed limit of road on which accident took place | [-1,99] |
| junction\_detail | INT | Describes the junction at which accident occurred | [-1,99] |
| junction\_control | INT | Refers to the junction control at which accident took place | [-1,9] |
| second\_road\_class | INT | Refers to the second road class on which accident took place | [0,6] |
| second\_road\_number | INT | Refers to the second road number on which accident took place | [-1,9999] |
| pedestrian\_crossing\_human\_control | INT | Category of pedestrian crossing human control | [-1,9] |
| pedestrian\_crossing\_physical\_facilities | INT | Describes the pedestrian crossing physical facilities at which accident occured | [-1,9] |
| light\_conditions | INT | Light conditions of the scene while accident took place | [-1,7] |
| weather\_conditions | INT | Weather conditions while accident took place | [-1,9] |
| road\_surface\_conditions | INT | Road surface conditions of scene while accident took place | [-1,9] |
| special\_conditions\_at\_site | INT | Describes the special conditions if any at which accident occurred | [-1,9] |
| carriageway\_hazards | INT | Details on the carriageway hazards caused by accident | [-1,9] |
| urban\_or\_rural\_area | INT | Refers to the type of area where the accident took place | [-1,3] |
| did\_police\_officer\_attend\_scene\_of\_accident | INT | Checks if the police officer attend scene of accident | [-1,3] |
| trunk\_road\_flag | INT | Checks if accident occurred is on trunk road | [-1,2] |
| lsoa\_of\_accident\_location | TEXT | ONS( Office for National Statistics) location of England and Wales only | None |

Table 1. Description of Accident Data

**3.2 Casualty Data**

The “dft-road-casualty-statistics-casualty-2019.csv” contains information on the casualty. It contains all kinds of various attributes about the casualty. One accident might have more than one casualty, and this could be identified uniquely by looking into the casuality\_reference under the same accident\_index.

All the Casualty attributes have lookup tables related to them which refer to the meaning of each field. Detailed description of each variable is shown in the table below.

|  |  |  |  |
| --- | --- | --- | --- |
| **COLUMN\_NAME** | **DATA\_TYPE** | **DESCRIPTION** | **RANGE** |
| accident\_index | TEXT | Unique value for each accident. | None |
| accident\_year | INT | Years on which accident took place | 2019 |
| accident\_reference | TEXT | Unique value for each accident | None |
| vehicle\_reference | INT | Unique value for each vehicle | [1,20] |
| casualty\_reference | INT | Unique value for each casualty, refers to the casualty under a accident | [ 1,991] |
| casualty\_class | INT | Refers to the class of casualty | [ 1,2,3] |
| sex\_of\_casualty | INT | Refers to the sex of casualty | {-1,1,2,9} |
| age\_of\_casualty | INT | Refers to the age of casualty | [-1,102] |
| age\_band\_of\_casualty | INT | Refers to the age band of casualty | [-1,11] |
| casualty\_severity | INT | Refers to the severity of casualty | [1,3] |
| pedestrian\_location | INT | Refers to the pedestrian location while the accident took place | [-1,10] |
| pedestrian\_movement | INT | Refers to the pedestrian location while the accident took place | [-1,9] |
| car\_passenger | INT | Refers to the situation of a car passenger when the accident took place | {-1,0,1,2,9} |
| bus\_or\_coach\_passenger | INT | Refers to the situation of a bus or coach passenger when the accident took place | {-1,0,1,2,3,4,9} |
| pedestrian\_road\_maintenance\_worker | INT | Refers if the casualty belongs to the category pedestrian and he or she is a road maintenance worker | [-1,3] |
| casualty\_type | INT | Refers to which casualty type category the casualty is associated | [0,113] |
| casualty\_home\_area\_type | INT | Refers to the type of home areas from which the casualty is from | [-1,3] |
| casualty\_imd\_decile | INT | Deprivation scores of individual areas into one of the groups of equal frequency | [-1,10] |
| Isoa\_of\_casualty | TEXT | ONS( Office for National Statistics) location of England and Wales only | None |

Table 2. Description of Casualty Data

1. **Detailed analysis**
   1. **What patterns are there in the demographics of casualties?**
      1. **Change in the number of casualties by month**

Figure 3 shows the number of car accident cases by month in 2019, averaging around 6,911 cases. The Lowest number of events in February and the highest number of events in November during the year. The change in caseload was relatively flat in the first half of the year, with an increase in the second half over the first.

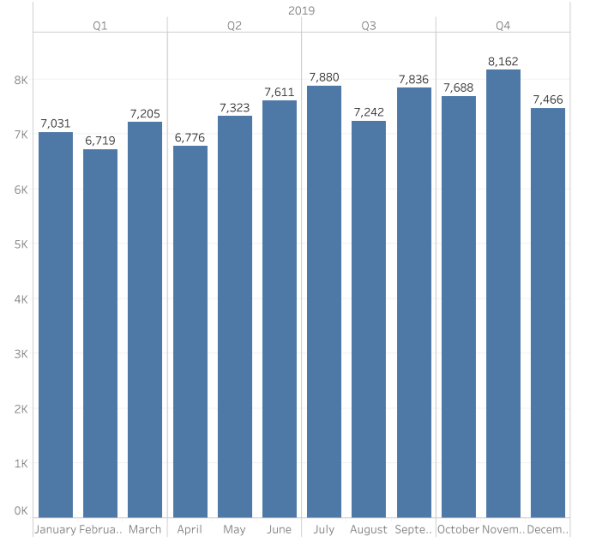


Fig.3.Number of casualties

* + 1. **Comparison of different levels of injury by month**

Figure 4 and Figure 5 also show a slow rise in the number of injuries starting in January, peaking in November, and declining slightly in December. The percentage chart shows that the largest number of people suffered slight injuries, with the highest percentage around 84%.

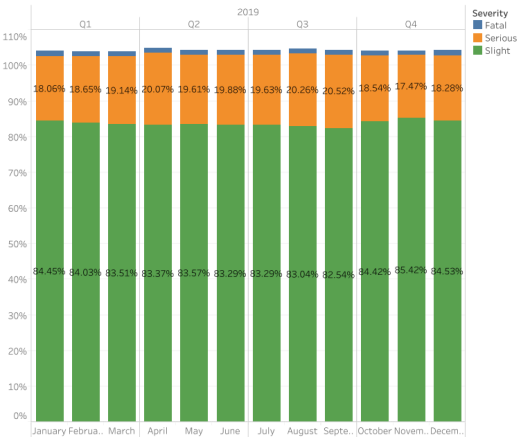
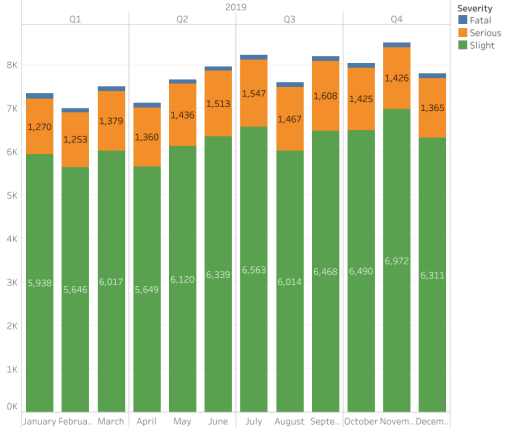
 

Fig.4. Percentage of injury types by month Fig.5. Number of injury types by month

* + 1. **The number of crashes at different times of the day**

The sub-tables in Figures 6 and 7 show the number of serious injuries and deaths in a 24-hour period (in decreasing order of magnitude). For 2019 as a whole, the highest number of serious injuries reached 1,403 at 17:00, and the highest number of fatalities reached 75 at 14:00, with the number of serious injuries and fatalities reaching their lowest values at 3:00 and 2:00 eperately. The number of crashes is highest at 5 pm, which coincides with people's off-duty hours, and the increase in the number of crashes is likely to be caused by the off-duty rush hour.

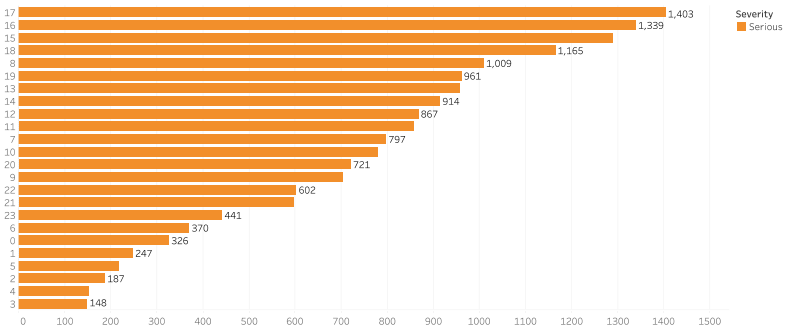


Fig.6. Serious 24 hours

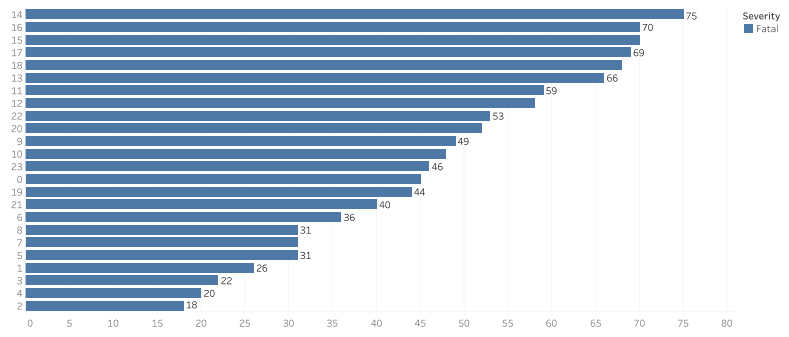


Fig.7. Fatal in 24 hours

* 1. **Looking at accidents which included a killed or seriously injured (KSI) casualty, what patterns are there between the local authorities?**
     1. **Relationship between the number of KSI casualties and local authority district**

In the analysis of local authorities, we used numbers because there are too many districts in the original data, and there are 941 districts in total. As can be seen in Figure 8, Region 300 (i.e. Birmingham) has the highest number of cases, with 420 cases.

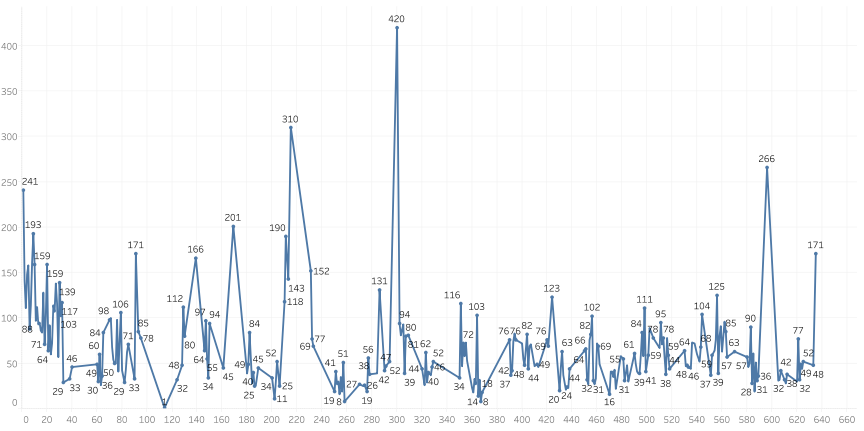


Fig.8. Number of KSI casualties by Local Authority District

* + 1. **Relationship between the number of KSI casualties and police force**

Similarly, when analysing the number of different police stations handling incidents, it was encountered that the number of police stations was too large to show specific names directly in the graph, so numbers were used instead. By looking at Figure 9, there is no doubt that Station 1 (i.e. Metropolitan Police) handled the highest number of KSI incidents, with 3,720 for the year. Comparatively, Police Station 6 (Greater Manchester) handled the fewest, with only one incident.

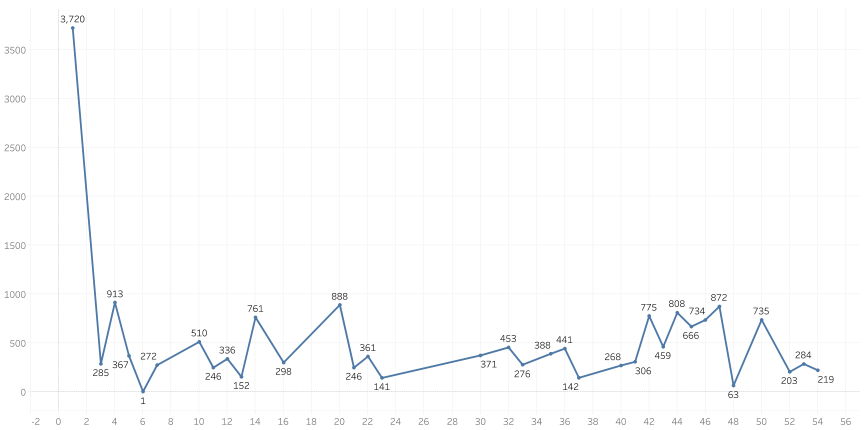


Fig.9. Number of KSI casualties by Police Force

* 1. **What patterns are there in pedestrians who were KSI casualties?**
     1. **Comparison of the number of different types of KSI casualties**

Table 3 gives the number of injured people in different categories. There are three categories: Driver or rider, Passenger, and Pedestrian. Of these, pedestrians had the lowest number of injuries, while the number of drivers or riders which has 4.6 times the number of pedestrians. It can be seen that people who have a certain driving speed are more likely to be injured in a car accident.

|  |  |
| --- | --- |
| **Casualties Class** | **Number** |
| Driver or rider | 75,627 |
| Passenger | 23,632 |
| Pedestrian | 16,199 |

Table 3. Number of KSI casualties by Casualties Class

* + 1. **KSI pedestrians of all ages**

Figure 10 gives a clear picture of the age groups of KSI pedestrians, including the number and the percentage. It can be concluded that the 26-35 age group has the highest number of people, with 2,292 people accounting for 14.48%. The next highest number of young people, at 2,087, were aged 11-15, accounting for only about 1% less than those aged 26-35. It is common sense that middle-aged and older people over the age of 66 should be involved in many pedestrian accidents due to their health and reaction time, but the number is actually not very high, totalling 2,143.



Fig.10. Age, Number, Percentage of KSI Pedestrian

* + 1. **Analysis of the circumstances of the accidents**

In addition to age, the location and movement of the pedestrian at the time of the crash are also relevant. As shown in Figure 11, the highest number of crashes and the most dangerous occur when pedestrians cross from driver’s nearside. Instead, if the pedestrian is in carriageway, stationary - not crossing(standing or playing) - masked by parked or stationary vehicle, he/she has the least chance of being involved in a car accident. When looking at the relationship between pedestrian movement and the number of crashes, it was found that people crossing in zig-zag exit lines had the fewest crashes, while those crossing elsewhere in the carriageway take the largest number of all, whether seriously injured or killed.

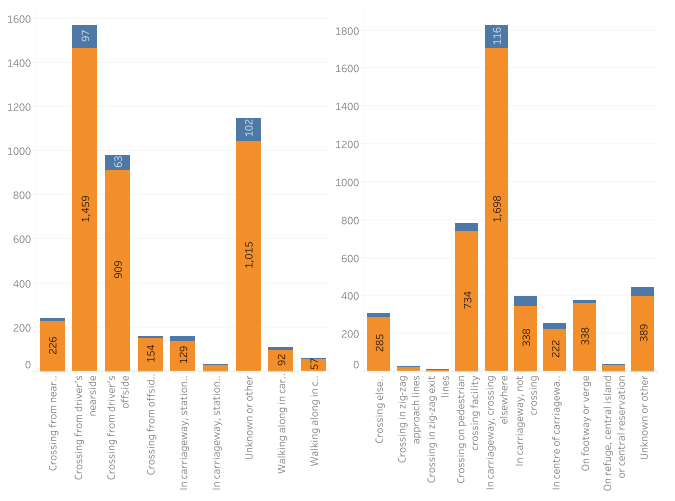


Fig.11. Number of KSI Pedestrian by Different Location and Movement

1. **Conclusion**

According to this report and data, the number of casualties is highest at the end of the year, but there are no major fluctuations. Most casualties are slight injuries, and only a small percentage are fatal. The number of casualties at 4-5pm accounts for the largest proportion of the day with the high volume of traffic and people at the end of the day being the main cause of crashes. The number of local police calls in Birmingham was high, and the government should pay more attention to solving such problem in this district.

Among the three ways, including Driver or Rider, Passenger and Pedestrian, pedestrians were least likely to be injured compared to others, suggesting that walking is the safest way to travel. Based on research data and graphs, young people aged 11-15 are most at risk of crashes, and parents need to improve education in this area to reduce the chance of teenage injuries in car accidents.