Road Collision Analysis of Leeds (UK)

Biking towards carbon-neutrality and
safe urban mobility in 2030

Abstract Promoting the use of cycling, walking, and public transportation as a means to navigate the city is part of the city of Leeds' plan to achieve being carbon neutral by 2030[1]. The potential benefits of embracing, and advocating for, the widespread use of these means of transportation also extends towards increased road safety. With Leeds being the 9th most congested city in the whole of UK[2], the local authorities need to take action to develop a future-proof urban mobility plan. This report sets the statistical foundation for data-driven policy to make the city greener and safer.

Aidan Stocks, *aist@itu.dk*
Christian M. Hansen, *chmh@itu.dk*
Jonas-Mika, *jsen@itu.dk*
Malthe Pabst, *mrla@itu.dk*
Rasmus Bondo Hansen, *rabh@itu.dk*

1. LEEDS.GOV.UK, *"Climate change"*, Accessed: 25/02-2021, Retrieved from: https://www.leeds.gov.uk/your-council/plans-and-strategies/climate-change
2. Wikipedia, *"Leeds"*, Accessed: 25/02-2021, Retrieved from: https://en.wikipedia.org/wiki/Leeds

Background and Motivation

Cities that prioritise cycling tend to experience fewer traffic related casualties and experience better overall environmental conditions, especially when supported by the appropriate infrastructure, such as separated and protected bike lanes[3]. This was recently made evident in Oslo, which reported only one traffic related fatality in 2019[4] as a result of their extensive work to make the city more bike-friendly and improve overall traffic safety. While Leeds has experienced a reduction in road casualties in recent years[5], the annual number of people killed or seriously injured has remained largely unchanged.

This report is an independent study setting out to investigate the overall road safety in Leeds, while specifically analysing which factors contribute to road accidents that involve cyclists, and how these might be addressed to effectively follow in Oslo's bike tracks.

Data

Description of Datasets

The investigation was carried out with the use of three data sets that were obtained from official statistics provided by the UK Department of Transport[6]. The *Road Safety Data* recorded for the whole of the UK for 2019 was split into three related data sets.

Main Dataset. The primary data set - referred to as *Accidents* for the remainder of the report - provided detailed information on the general circumstances of the accidents in question. The accident attributes within the data set were separated into categorical (ie. *Accident Severity*), geographic (ie. *Latitude*, *Longitude*) and relational attributes (ie. *Accident Index*).

Sub Datasets. Two other sub-data sets provided additional information about the vehicles and casualties involved for each accident reported in *Accidents*. These two data sets are respectively referred to as *Vehicles* and *Casualties* from now on. The *Casualties* data set contained 15 attributes relevant to the individuals that were casualties of each accident; such as the casualty severity (slight, serious, fatal), casualty type (driver, passenger, or pedestrian), sex, and age. The *Vehicles* data set consisted of 22 attributes relating to the characteristics and operation of the vehicle involved in the accident. This included the type of vehicle, junction location, vehicle manoeuvre, and age of vehicle.

Data Processing

The quality of the data was first assessed by checking *Accidents* for duplicate unique Accident IDs. These unique IDs were then cross-referenced with the IDs in the *Casualties* and *Vehicles* data sets, to ensure that the entries from the latter two data sets corresponded to an entry in the former. While there were no duplicate entries, the analysis revealed that there were roughly 20,000 entries in each sub-data set that did not refer to the main data set *Accidents*. With this in mind, the final step was filtering for accidents that specifically took place in the city of Leeds, using Leeds' Local Authority District (LAD) identifier. The filtered data set, now specific to road accidents in Leeds for 2019, was again checked for missing Accident IDs and found to be complete.

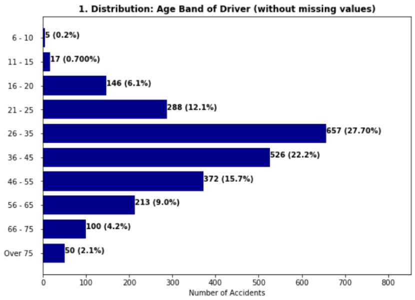
Results and Discussion

Single Variable Analysis

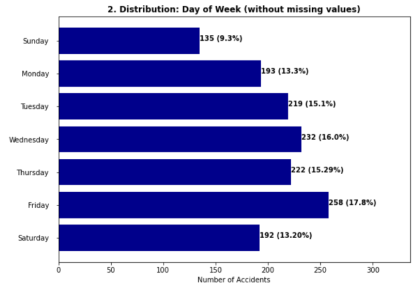
The initial investigation into the accident data was focused on the frequency of road collisions in relation to different accident, vehicle, and casualty attributes. The initial attributes investigated were the age bands of those involved, the time at which the accidents took place, and the sex of the driver (*Figure 1*).

Subplot 1 illustrates the age band distribution of the individuals involved in the accidents, indicating that people between 26-35 are most involved in accidents. This is a somewhat expected result, as it matches the age distribution of the population of Leeds[7].

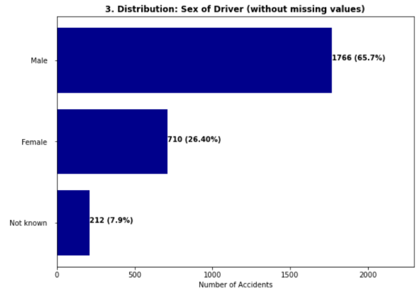
3. Science Daily, *"Cycling lanes reduce fatalities for all road users, study shows"*, Accessed: 25/02-2021, Retrieved from: https://www.sciencedaily.com/releases/2019/05/190529113036.htm
4. Anders Hartmann and Sarah Abel (2020), *"How Oslo Achieved Zero"*, Retrieved from: http://staging.nxtbook.com/ygsreprints/ITE/ITE_May2020/stage.php#/p/32
5. LEEDS.GOV.UK,*"Road traffic collision statistics"*, Accessed: 25/02-2021, Retrieved from: https://www.leeds.gov.uk/parking-roads-and-travel/connecting-leeds-and-transforming-travel/road-safety/road-traffic-collision-statistics
6. Data.gov.uk, *"Road Safety Data"*, Accessed: 05/02-2021, Retrieved from: https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data
7. Plumplot, *"Leeds population statistics"*, Accessed: 25/02-2021, Retireved from: http://www.plumplot.co.uk/Leeds-population.html



Accident Frequency for different Age Band of Driver



Accident Frequency for different Days of the Week



Accident Frequency for different Sexes

FIGURE 1. *Plots for Single Variable Analysis*

Subplot 2 revealed that the majority of accidents took place on weekdays. These findings correspond to increased traffic during the week due to work commutes.

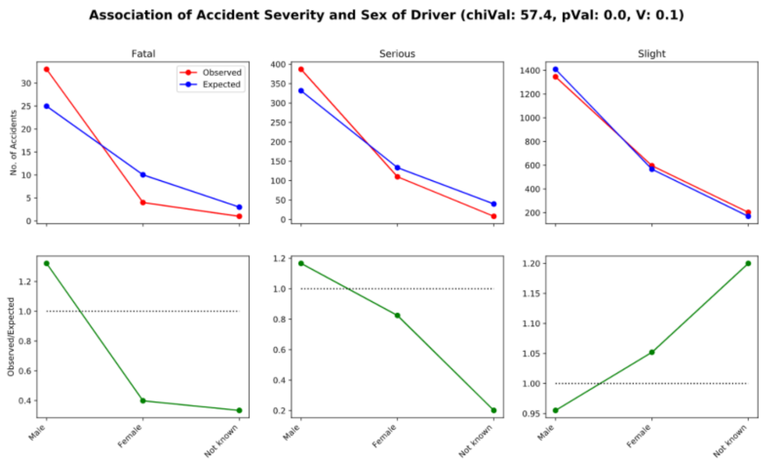
Subplot 3 investigates the significance of the sex of the driver in relation to the frequency of accidents. The figure reports males being involved in more than twice as many accidents than females.

Associations

In this section, our goal was to investigate, which accident, vehicle, or casualty property resulted in the most severe accidents. To investigate these associations, we used the chi-squared test, which involves computing the p-value and the Cramér's V-value for each relationship between a categorical attribute and the accident severity.

A relevant association was found between the accident severity and the sex of driver. *Figure 2* shows this association, clearly illustrating that men are more likely to be involved in severe accidents than women. In turn, women are more likely to be involved in slight accidents. Through the low p-value and the Cramer's V of 0.1 we can conclude that the sex of the driver is a factor related to the severity of the accidents.

This insight might be used for urban transport planning. To reduce the amount of severe accidents, it is reasonable to counteract against the more reckless driving behavior of men. This could ie. be achieved through periodic seminars, specifically designed to raise awareness for conscious driving behavior. Another -more drastic- method might be the obligation to retake the drivers test in a specific time frame.



Notes: Statistical Association by Pearson Chi Squared. Reported p-Value: 0, and Cramer's V: 0.1

FIGURE 2. *Association of Accident Severity and Sex of Driver*

Spatial Visualisation

Data is abstract as long as we cannot see it - the same goes for our Road Collision Data. 1,451 accidents appear as 1,451 rows of abstract numbers. This part of our data analysis therefore focused on making the data interpretable. To do so, we plotted the reported road collisions on a scalable interactive map (*Figure 3*). The map can be opened in any browser and is intuitive to navigate. Different layers can be turned on and off to provide different ways of inspecting the data. Each individual accident is color coded by its severity (where 'black' implies a 'fatal', 'red' a 'serious' and 'green' a 'slight' level of accident severity). For more detailed information, each accidents' circumstances can be inspected through a popup-text appearing on hover.

The map shows an increasingly high density of road collisions towards the center of the city. Besides the increased frequency of accidents in urban areas, the map also reveals especially dangerous roads.

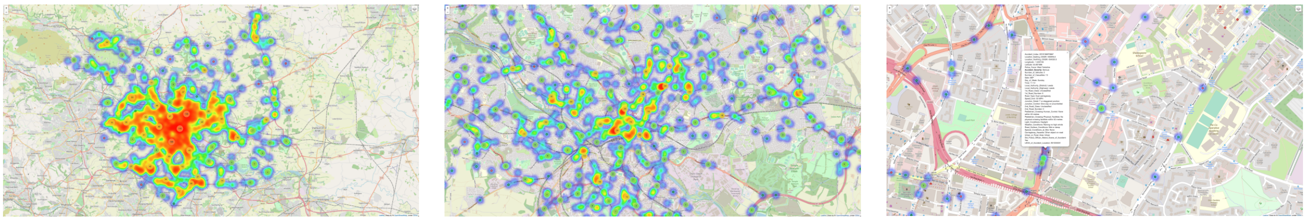


FIGURE 3. *Spatial Visualisation of Leeds Accidents in 2019*

Bike Safety in Leeds

Our analysis of traffic flow in Leeds, as well as external research, has proven that Leeds is currently lacking adequate alternative means of travelling the city. To this day, Leeds does not have a metro system and the number of cars significantly outnumbers the cyclists in the city (*Figure 4*).

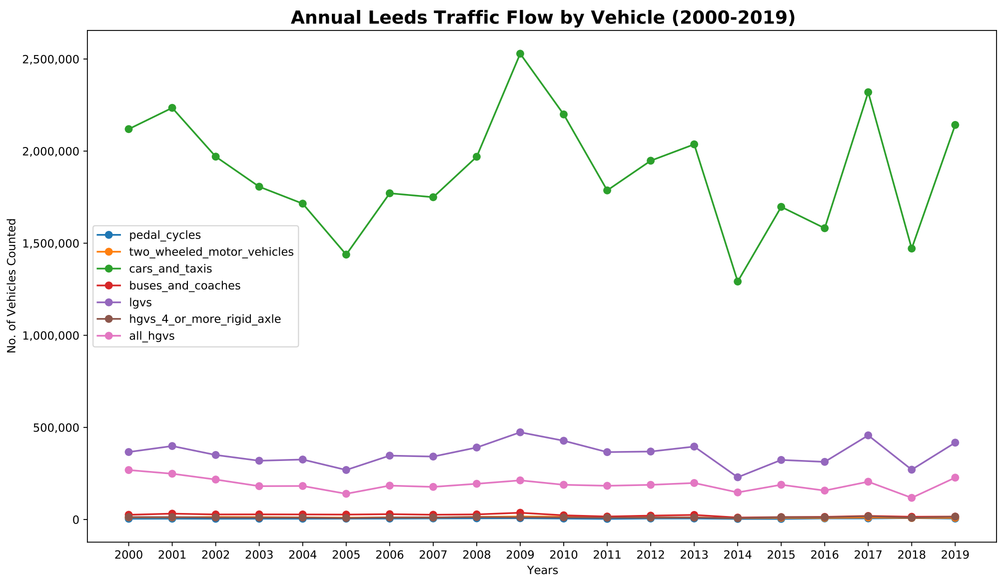


FIGURE 4. *Annual Traffic Flow by Vehicle Type in Leeds (2000-2019)*

Although there are hardly any bikes in the streets of Leeds compared to the number of cars out of 1,451 total reported accidents in Leeds, 238 involved bikes. *Figure 4* gives motivation to increase bike safety in Leeds and explore what measures could be taken to reduce the number of bike accidents.

Most Common Bike Accident

94.2% of the bicycle accidents occur with cars or taxis, and in almost all of those cases the bicyclist gets injured. In contrast, only four drivers of motorised vehicles had been injured in all recorded bike accidents. By looking at where these accidents occur, it can be seen that in most cases the accidents are at junctions, where the cyclist is going straight and the car is turning, and the collision is front to front. This suggests that in most bicycle accident scenarios the car driver turns into the bicyclist on accident. This recurring pattern of accident participants, manoeuvres, and location of accidents can, and should, be used to actively prevent these patterns of accident scenarios from happening. This could ie. be achieved through bike-specific traffic lights, longer delay periods after red lights, or similar measures.

On the Way to Work

The data gives various arguments for the fact that most bike accidents happen during the rush hour times from 6-9 and 15-19. Firstly, most of the participants involved in bike accidents stated to be on the way to or from their workplace. This can also be seen in the distribution of accidents in different times of the day. The road collision numbers peak during the rush hours. The distribution of bike accidents during the week and the year further strengthen the above argument. On days of the weekends there occur noticeably less accidents when compared to regular week days. In the distribution of bike accidents per month, it is visible that the number of bike collisions drop in the month of August, which is the month of national summer holiday in the UK. Therefore less people commute to or from work, and thus the number of bike accidents decreases.

Hazardous, Big Roads

Most of the accidents happen on streets classified as class A, which in the UK refers to main carriageways with high speed and little bends. The three most hazardous roads are the roads with road number 660 (Woodhouse Lane or Otley Road), 65 (Kirkstall Road) and 61 (Regent Street). Together over 15% of all accidents involving bikes occur on these streets. Naturally, these roads experience the majority of accidents since they are used more frequently throughout the year than the less central roads. However, the high percentage of bike accidents happening at the three mentioned roads, gives reason to believe that these roads are poorly secured for safe bike-travel, and should therefore be a focus of political measurements trying to reduce the number of bike accidents in Leeds.

Limitations

A difficult limitation was to avoid drawing biased conclusions. During a group discussion we talked about many associations we could draw from the data, such as how there are more men than women involved in bicycle accidents. However, since the data does not show the amount of men and women that cycle in general, we cannot say for sure whether men are more prone to having bicycle accidents because they are men, or because there are more of them cycling. We are lacking more general data that would give us perspective into such associations.

Concluding Remarks and Future Work

While there exists plans to redevelop parts of the city to better accommodate cycling and public transportation, there are simple changes that could be implemented to improve traffic safety. Most noticeably, as a significant amount of cycling related accidents take place at junctions, the city could implement more toucan crossings or protected bike paths to improve cyclist safety in these perilous areas.

To further and improve our analysis, if more general information such as traffic data not specifically tied to accidents was provided, there would be less limitations in the associations we could draw. Additionally, location data on all currently existing bike paths within Leeds, and the frequency of their use, would be extremely valuable for further analysis. Not only would this reveal if the existing paths are actually being used in lieu of riskier but more convenient routes, but would allow us to pinpoint how significantly bike paths aid in the reduction of cycling related traffic collisions.

We hope for this study encourages the local authorities in Leeds to rethink the current road safety conditions, and future urban mobility plans to pave the way towards a green and safe urban mobility environment.

Disclosure Statement

The majority of the Jupyter Notebook was coded out by Jonas-Mika Senghaas.