The University of Manchester

A step-change in quantitative social science skills

Funded by the Nuffield Foundation, ESRC and HEFCE

Transport for London: Evaluating the Impact of Interventions on Safety Performance

Dionysios Stamatelakis | BA (Hons) Social Sciences (Politics & Criminology)

Overview of the Data Fellowship

The project aimed to set out and trail a template methodology which will enable TfL to successfully quantify the impact of different safety measures on number and type of incidents on different surface Modes. In the long-term, this will enable more data-driven target setting and intervention planning to maximise the benefits for the customers and the businesses.

Data Analysis

The data I analysed are injury and incident data from buses. The first step was to identify which columns I needed to extract from the IRIS database. To simplify the process, the objective was to scale it down as much as possible at this early stage in order to be able to pick the ones with the least variables that changed over the relevant period of time. Then, I decided on how many years of data I wanted to include. For my research project to be reliable, I chose 5 as older data would be unrepresentative. After that, I did an extract that included all the columns and years that I needed for my analysis. This enabled the analysis to be quickly run on updated datasets in the future as needed. Finally, I made an extract from SQL to Excel, which allowed me to do the further univariate and bivariate analises. Then i proceeded with cleaning and gathering as well as interpreted the data and generated tables.

Findings

Due to the confidential nature of some parts of the project, i am unable to share any of my findings that specifically name bus operators or safety intervantions. Unfortunately, because of data limitations in the extractions and raw data, the project was not completed. However, the discriptive and afterwards exploratory analyses allowed to produce a very constructive piece that not only impressed the stakeholders,

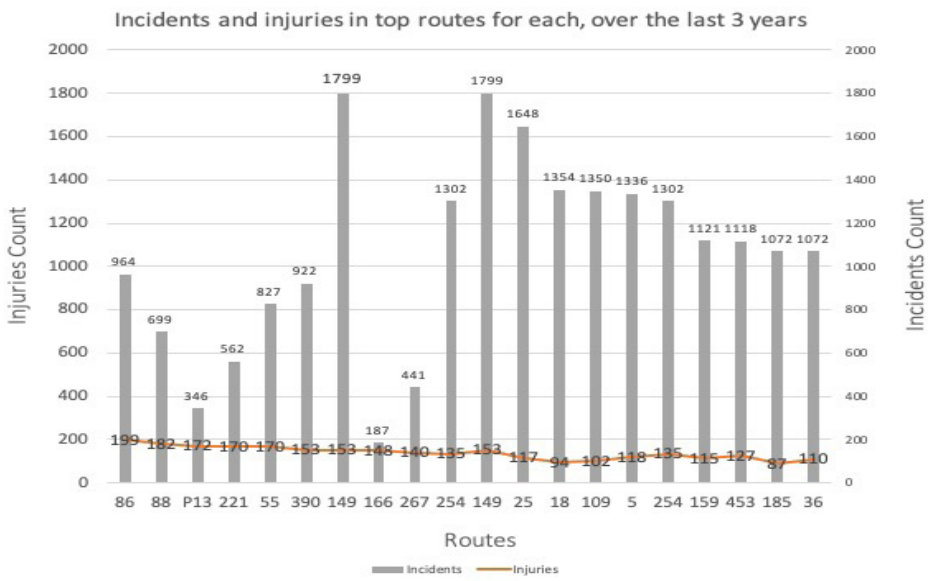

Figure 1 showcases one of the overview findings. Here we can see how the most dangerous routes with the most injuries dont necessarily have the most incidents. This allwoed to further scale down our analysis.

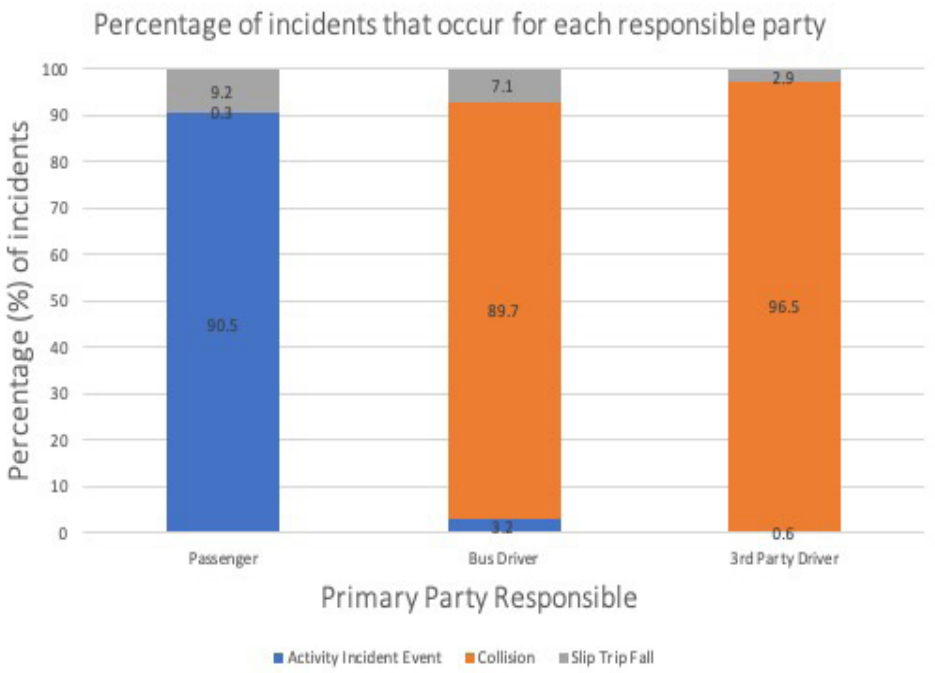

Figure 2 highlights what percentage of the accidents as well as what accidents each primary party responsible accounts for.

giving them important insight in bus incidents over the last 5 years, but also shed light on the raw data imperfections of TfL. This can help improve both the way incidents are recorded to make sure every detail is logged on IRIS but also allow TfL to be more critical about the data extracts used for research.

Key Skills Learnt

The analytical and research skills I learnt enabled me to do the research including extracting data from the IRIS database using SQL and R which then I exported to Excel for cleaning, filtering and further analysis. The skill of extracting datasets from SQL has taught me many new things about functions, commands, raw data analysis and programming. Communication and networking skills were developed through arranging and leading team meetings. Presenting our draft analysis to key stakeholders, the Senior SHE Business Partners for Bus Operations, gave me the chance to practice my presentation skills in a real life scenario and broaden my networking with co-workers from various sectors of the company. Collaboration, team working, adaptability and time management were advanced through the various steps of the project I led. For example, during the extraction of data and draft analysis, we encountered numerous issues in the data such as missing data and wrong dates that had to be dealt with all in good time to present.

Contact: dionysios.stamatelakis@manchester.ac.uk