Road Accidents Analysis

Name & ID - Muskan Muskan (C70303902)

Name & ID - Vatsal Rameshbhai Gohel (C70303897)

Name & ID - Sakina Hassani (C70306262)

Course - Business Intelligence (MSIT3820-02-F23)

Due Date - 12/10/2023

Croup 8: Final Project

Executive Summary:

In this project, the business analysts at the National Road Safety Intelligence Bureau of the UK Department for Transportation analyzed the road safety data from 2019 to 2022 (Kaggle.com). The goal of this project was to address the challenges the Department for Transportation faces: a nationwide increase in road accidents and a tragic loss of human lives. The Bureau analysts analyzed the dataset to find the root cause of road accidents by investigating the patterns and trends in the data. The analysis findings revealed that road conditions, weather conditions, road and vehicle types, and other external factors are not the reason for the increase in accidents. The drivers' lack of information or uninterest in adhering to the traffic rules are the causes of the accidents.

Thus, this project's recommended plan of action is to leverage advanced Business intelligence (BI) tools in the long run, starting by developing a mobile application for data collection, communication, and education purposes. The BI tool will help the bureau start with a user-friendly mobile application for collecting real-time data, utilizing geospatial analysis to find the exact location of the accidents, communicating with the users, and providing training and awareness programs. This will later be improved by advanced machine learning models and analytical methods to enable immediate and accurate decision-making consequently. Moreover, integrating the BI tools into road safety systems and approaches will create collaborative work and be useful for different entities, authorities, and stakeholders.

Statement of the Problem:

The Department for Transport of the UK is currently grappling with a surge in nationwide road accident complaints, which are not only causing chaos but also resulting in tragic loss of human lives. In 2016, recognizing the escalating surge in road accidents and the alarming yearly statistics, the Department for Transportation embarked on a collaborative venture by establishing the National Road Safety Intelligence Bureau. The bureau's clear mission was to delve into the extensive data surrounding major road accidents and decipher patterns and trends to devise effective solutions.

The National Road Safety Intelligence Bureau is dedicated to addressing recurring road accidents and tackling road safety challenges and is aimed to leverage available data efficiently. Our strategy at the National Road Safety Intelligence Bureau involved not only understanding the intricacies of accident data but also integrating this knowledge with cutting-edge technology. Through this holistic approach, we sought to develop innovative solutions that, when implemented, would significantly reduce casualties on the roads.

Causes of the Problem

Data Analysis:

To investigate the root cause of the surge in road accidents and tragic loss of human lives, the Bureau conducted a comprehensive data analysis on the data collected by the Department for Transportation (Hypothetically). This dataset equipped us with valuable insights that can significantly benefit stakeholders such as the Department for Transport, Police Force, Emergency ServicesDepartment, Road Safety Corps, Transport Operators, Traffic Management Agencies, as well as the general public and media. Notably, the dataset included critical geospatial data like longitude and latitude, enabling us to pinpoint the locations of accidents precisely. This geographical information holds immense potential for enhancing our ability to strategize and implement targeted measures aimed at reducing the overall number of casualties. Our research was poised to provide actionable insights that can contribute meaningfully to the efforts of various stakeholders in the realm of road safety.

At present, we do not have a Business Intelligence system in place, but we organized our data based on the dataset we have. As a result of this structured schema, we were able to efficiently analyze and interpret insights from the data to address our road safety concerns. The logical components of our Star Schema include

a fact table and dimension tables. The fact table contains quantitative data, including number of accidents, the number of vehicles, the number and severity of casualties and the primary key of the dimensions. The dimension tables include attributes such as date, location, weather conditions, road conditions, vehicle types, junction, etc.

Data Cleaning:

To enhance data quality, we executed a comprehensive data cleaning and transformation process. In alignment with our primary analytical objectives and the specific challenges at hand, we judiciously omitted certain columns from our initial analysis. These columns, namely "Day_of_Week," "Junction_Control," "Junction_Detail," "Local_Authority_(District)," "Carriageway_Hazards," "Police_Force," "Speed_limit," and "Time," were identified as temporarily dispensable for the current phase of investigation.

However, it is imperative to note that we envision incorporating these excluded columns in subsequent stages of our analysis to facilitate a more nuanced and granular examination of the data. These columns harbor potential insights that may prove instrumental in unraveling deeper patterns and correlations.

Furthermore, we identified and rectified several spelling errors within the dataset. Notably, we corrected the spelling discrepancies in category names to ensure uniformity and coherence. For instance, in the case of "Accident_Severity (Fatal & Fetal)," we rectified the spelling to establish it as a sub-category, thereby fostering clarity and consistency in our dataset. This attention to detail reinforces the integrity of our data, paving the way for robust and reliable analyses in our ongoing and future endeavors.

Our commitment to refining data interpretation led us to consolidate categories sharing similar meanings. For instance, we streamlined classifications such as "Darkness - lighting unknown," "Darkness - lights unlit," "Darkness - no lighting," and "Darkness - lights lit" into a unified "Darkness" category. Similarly, we harmonized other categories in columns like Light Conditions, Road Surface Conditions, Vehicle Type, and Weather Conditions to facilitate a more coherent analysis.

To enhance the user experience and offer a dynamic perspective, we implemented an "Accident Severity Filter" on our dashboard. This feature allows users to seamlessly toggle between different accident severity levels, enabling a focused exploration of statistics for each severity category.

In addition, we introduced calculated fields to distill key metrics, such as CY Accidents, PY Accidents, CY Causalities, PY Causalities, CY Fatal Causalities, PY Fatal Causalities, CY Serious Causalities, PY Serious

Causalities, CY Slight Causalities, and PY Slight Causalities. These fields provide a nuanced understanding of various accident-related aspects, aiding in a comprehensive analysis.

To offer a longitudinal perspective, we incorporated a calculated field named "YoY Accidents," leveraging the accident date's year component. This addition empowers users to assess year-over-year trends, contributing valuable insights into the evolving nature of accidents over time.

Further enhancing interactivity, we introduced parameters such as "Current Year," "Previous Year," and "Select Accident Severity." These parameters enhance the adaptability of our dynamic dashboard, enabling users to tailor their analyses based on specific criteria and refine their exploration of the data according to their unique requirements.

Findings:

After the data was cleaned and ready for analysis, we used Tableau to visualize the data and find the patterns and trends demonstrating the root cause of the problem. The data analysis first confirmed that the number of accidents has increased since 2019. The total accidents have a 20.70% year-over-year increase (See more in appendix A). Second, 64.43% of the accidents happen on roads with dry surfaces and 80.71% in fine weather conditions compared to other types of road surface conditions and weather conditions (See more in Appendix B). 144,539 accidents happened during daylight, and only 51,198 were in darkness regarding the light conditions. Looking at the vehicle types, we found that the three major groups of vehicles causing accidents are cars (155,804), vans (15,905), and motorcycles (15,579). The analysis of road type revealed that a vast number of road accidents happened on single-carriageways (144,653), followed by dual-carriageways (31,912), and roundabouts (12,683) (See Appendix C). Finally, accidents happen all over the UK. However, the majority of them, including the severe accidents, happened in big cities, an example of which is Bermingham (See the map in Appendix D).

To conclude, our analysis revealed that most accidents happen in normal situations, indicating that the road, weather, or light conditions are not the reasons for the accidents. The reason for the accidents is the driver’s carelessness, which might be due to a lack of information on the traffic and safety rules or disinterest in obeying the rules. Considering the findings of the initial analysis by the National Road Safety Intelligence Bureau, the

recommended solutions for decreasing accidents, increasing road safety, and saving human lives will be considered next.

Decision Criteria and Alternative Solutions:

The bureau is dedicated to a primary mission: raising awareness about road safety with the ultimate goal of reducing the number of accidents. We aim to achieve this by harnessing the wealth of data derived from past accidents. Through comprehensive data analysis, we intend to identify recurring trends and patterns, subsequently using these insights to disseminate critical information about road safety. In addition to our core objective, our bureau is committed to addressing road safety concerns in collaboration with transportation authorities, law enforcement agencies, and other keystakeholders. Our collective aim is to enhance safety measures and mitigate the tragic consequences of road accidents. The recommended solution needs to be effective in the long term and is preferred to be achieved in a short time. Thus, it is expected to provide a solution that could impact the stakeholders with rapid access to data and enable immediate decision-making.

In order to further support the bureau's goals, we propose several initiatives. First, a few public awareness campaigns, including social media, are proposed to educate the public about safe driving practices, adherence to traffic rules, and distracted driving dangers. The second option is a collaboration with smart city initiatives in an effort to incorporate road safety data into broader urban planning efforts, which will result in safer and more efficient transportation. The third alternative solution is to conduct predictive analytics, which would allow us to identify hotspots and high-risk areas based on historical data, allowing us to allocate resources and take preventative measures in advance. Finally, the fourth alternative solution is to leverage advanced Business intelligence (BI) tools in the long run, starting by developing a mobile application for data collection, communication, and education purposes.

The first three alternative solutions will help the Department for Transportation and its partners decrease road accidents and improve road safety. However, none is a long-term solution that could provide real-time data and help with immediate decision-making. The fourth solution, however, enables the department to collect and analyze real-time data, which will be used for immediate decision-making and, consequently, improve road safety. Additionally, integrating the data into a user-friendly mobile application would enable paramedics and towing services to be dispatched efficiently according to precise accident location coordinates. Moreover, the

mobile app will enable us to launch public awareness and educational programs.

One of the critical challenges with integrating BI is the occasional difficulty in obtaining data from all regions, leading to a reduction in the accuracy of our predictive models. This lack of accuracy poses a significant hurdle in our mission to prevent road accidents and increases the burden on the National Road Safety Intelligence Bureau. Additionally, connectivity issues with transportation authorities, stemming from inadequate infrastructure, hinder our ability to collaborate effectively with medical and law enforcement agencies. The task of influencing legislative changes to enhance road safety can be particularly daunting, given that these matters involve state and federal laws. Implementing changes in these domains requires extensive processing and time, making it a complex and time-consuming endeavour.

Recommended Solutions, Implementation and Justification:

Despite all these challenges, our focus has shifted toward Business Intelligence (BI) as the recommended solution to enhance road safety since it is the most comprehensive approach. We aim to integrate the data we have collected into a mobile application. The ultimate aim is to provide valuable support to various departments and agencies involved in road safety, addressing this urgent issue effectively.

It is important to note that the development of this app will take some time, as it entails a meticulous analysis of the collected data to identify patterns. Subsequently, the app will be built based on this information.
Our data management and analysis approach is well-structured at the National Road Safety Intelligence Bureau. We harness a combination of powerful tools to address our mission of enhancing road safety effectively.

Comprehensive data analysis is key to the bureau's primary mission of promoting road safety and reducing accidents. We work with transportation authorities, law enforcement agencies, and key stakeholders to improve safety measures and minimize the consequences of road accidents. We initiate our data collection process using Excel to gather historical information related to road accidents and safety incidents. To maintain this valuable data, we employ SQL databases, ensuring that we retain only the information essential for our specific use case, which is road safety data. This app will also utilize the longitude and latitude coordinates of accident locations to dispatch paramedic services and towing assistance promptly. The goal is to ensure the swift and efficient management of accident scenes without causing extended road closures. Data analysis is a critical

step in our workflow, and we turn to Python for this purpose. Python allows us to conduct descriptive and predictive analyses, providing us with actionable insights into road safety trends and potential areas of concern.

To present this data in a comprehensible and visually engaging manner, we utilize Tableau. This tool empowers us to transform our tabular data into insightful visualizations and interactive dashboards. These visuals are designed with the aim of making complex information easily understandable for our business stakeholders and partners in various agencies, thereby fostering effective communication and collaboration in pursuit of our organization's shared goals.

In the future, we are poised to expand our capabilities by incorporating APIs and application development. These elements will play a pivotal role in the realization of our earlier-mentioned app, which is a key component of our ongoing efforts to improve road safety.

Road safety analysis provides Business Intelligence (BI) insights to a wide array of stakeholders interested in understanding and improving road safety. Government entities such as the Department for Transport, which shapes transportation policies and regulations, are the primary consumers of BI insights in this context. These insights become crucial in improving road safety across the nation. Also, agencies like the Road Transport Department use BI to implement targeted interventions and enhance overall road infrastructure based on data-driven insights to manage and regulate road transport effectively.

The Bureau will take further steps to improve the Business Intelligence (BI) systems for road safety analysis. As a result of integrating advanced analytics and machine learning algorithms into predictive modelling, insights can be significantly more accurate, aiding in developing proactive accident prevention measures. Real-time data integration is crucial to ensuring the BI system reflects the most recent information, especially when responding to road safety emergencies as soon as possible. To enhance the accuracy and reliabilityof BI insights, robust data quality processes and governance standards, such as regular data cleaning and validation, are essential.

By incorporating social media, traffic camera feeds, and other IoT devices into data sources, roadsafety factors can be seen from a broader perspective. The planned mobile application must be continuously enhanced for effective emergency response, incorporating user feedbackand ensuring user-friendliness. Other road safety agencies can collaborate by integrating APIs with external data sources and systems. The BI system is effective

when users receive awareness, assistance, documentation, and training sessions. The ability to notify stakeholders of critical insights or emerging trends through automated reporting and alert systems ensures timely awareness and response to changing road safety conditions. In order to accommodate future increases in data volume and user demand, scalabilityplanning is essential. Furthermore, if data visualization techniques are aligned with best practices,insights will be presented intuitively and actionable. Collaborative BI platforms promote real-time collaboration and insight sharing between different stakeholders.

For sustained effectiveness, monitoring and optimising the BI system for efficiency, including addressing bottlenecks and optimizing query performance is essential. Continuous improvement is driven by creating a feedback mechanism for users to provide input on usability and functionality.Security measures, such as encryption, access controls, and regular security audits, help protectsensitive road safety information. By maintaining a culture of continuous improvement within the BI team, the bureau can remain on top of its BI game. With the implementation of these recommendations, the bureau can optimize its business intelligence systems, ensuring they are effective, scalable, and aligned with stakeholders' evolving needs to improve road safety.

BI insights significantly impact law enforcement agencies, as represented by the Police Force. These insights can reduce road accidents by identifying accident patterns, optimizing resource allocation, and refining enforcement strategies. In order to improve preparation and coordination, emergency service departments, such as fire departments and paramedics, benefit from BI for prompt and efficient responses to accidents. In addition to tailored campaigns, educational programs, and enforcement initiatives, road safety organizations can use BI insights to focus on targeted campaigns, education, and enforcement. Public transportation and logistics companies use BI insights to optimize routes, improve vehicle safety, and reduce accident risks. By leveraging BI, traffic management agencies can improve traffic flow, signage, and road safety.

In summary, we firmly believe that harnessing business intelligence to scrutinize the data will unveil critical information about accident-prone scenarios. This approach enables us to proactively address and rectify the conditions and circumstances contributing to accidents, ultimately making our roadways safer. Taking these initiatives will not only meet the bureau’s mission but will also enhance organizational efficiency, decision-making capacity, and overall effectiveness of the shared-mission stakeholders.

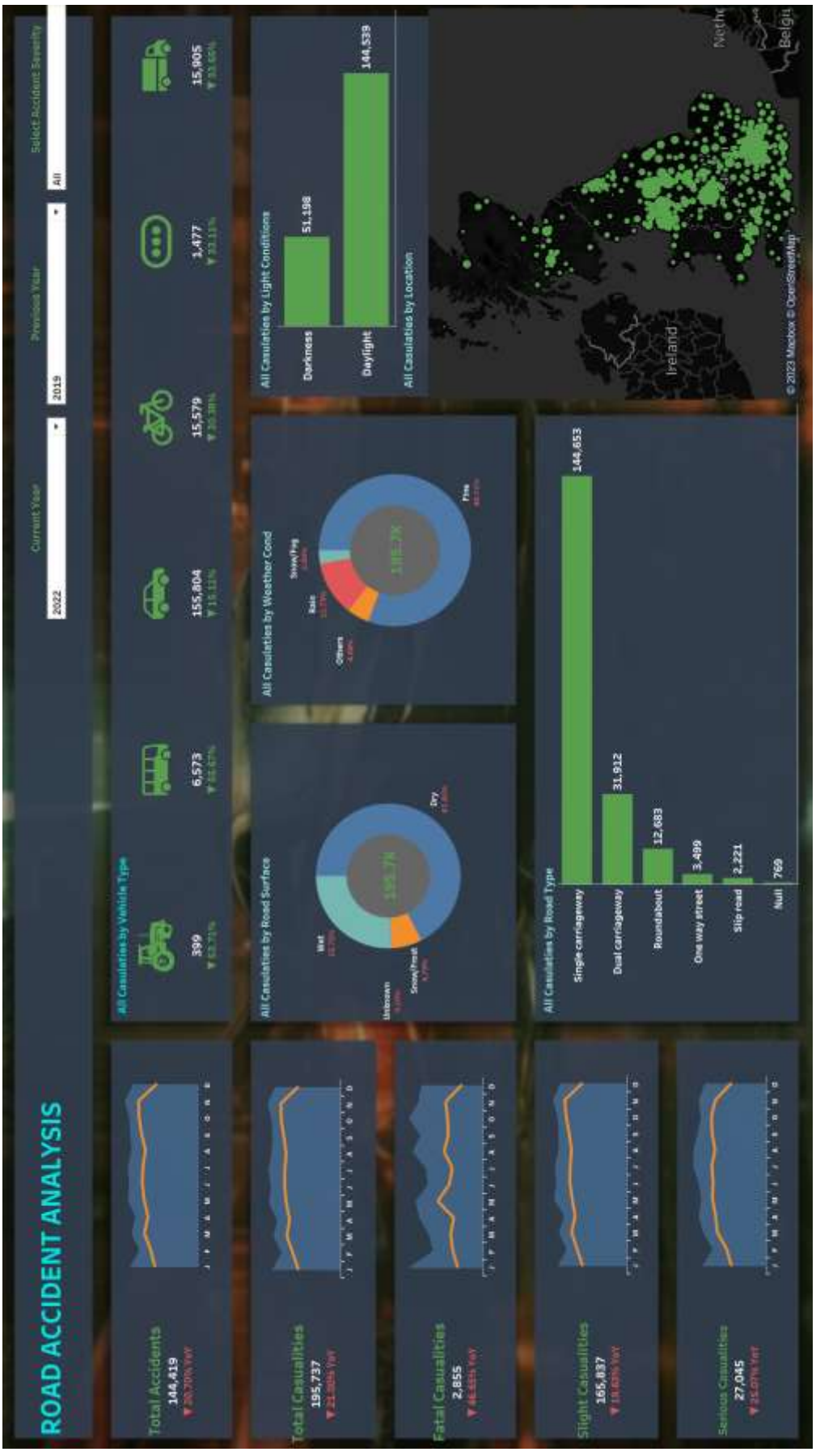
Reference:

Agrawal, R. (n.d.). *Road Accident*. Www.kaggle.com. Retrieved December 8, 2023, from https://www.kaggle.com/datasets/richeeee/accident-data?resource=download

Department for Transport. (2019, January 14). *Department for Transport*. GOV.UK. https://www.gov.uk/government/organisations/department-for-transport

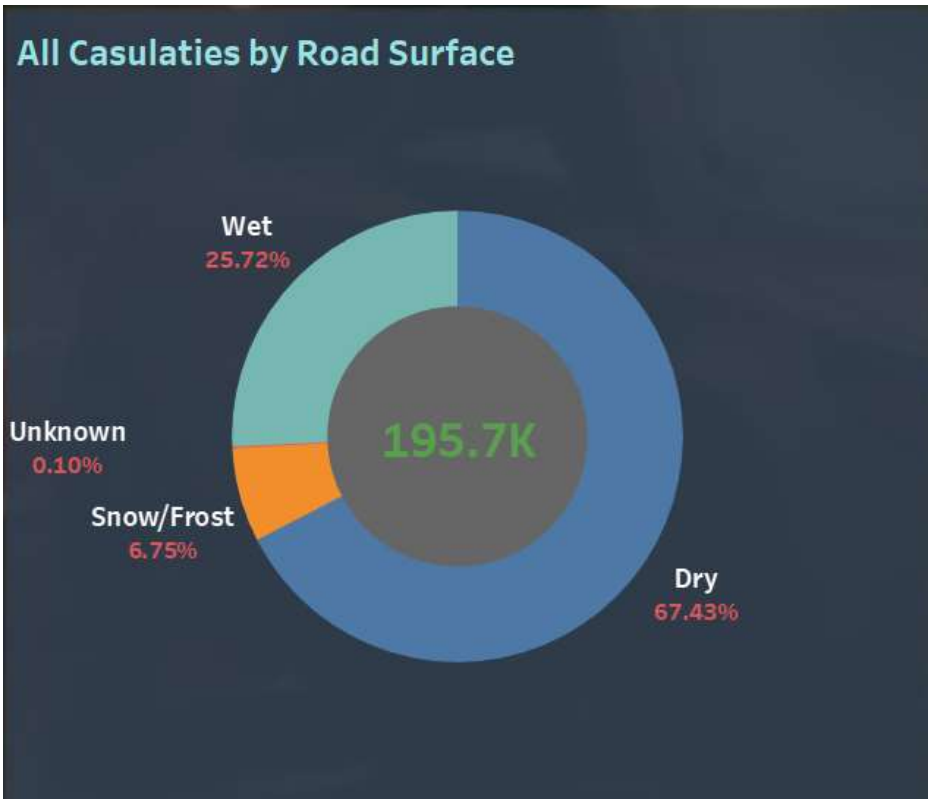
Appendix A

Road Accident Analysis Dashboard

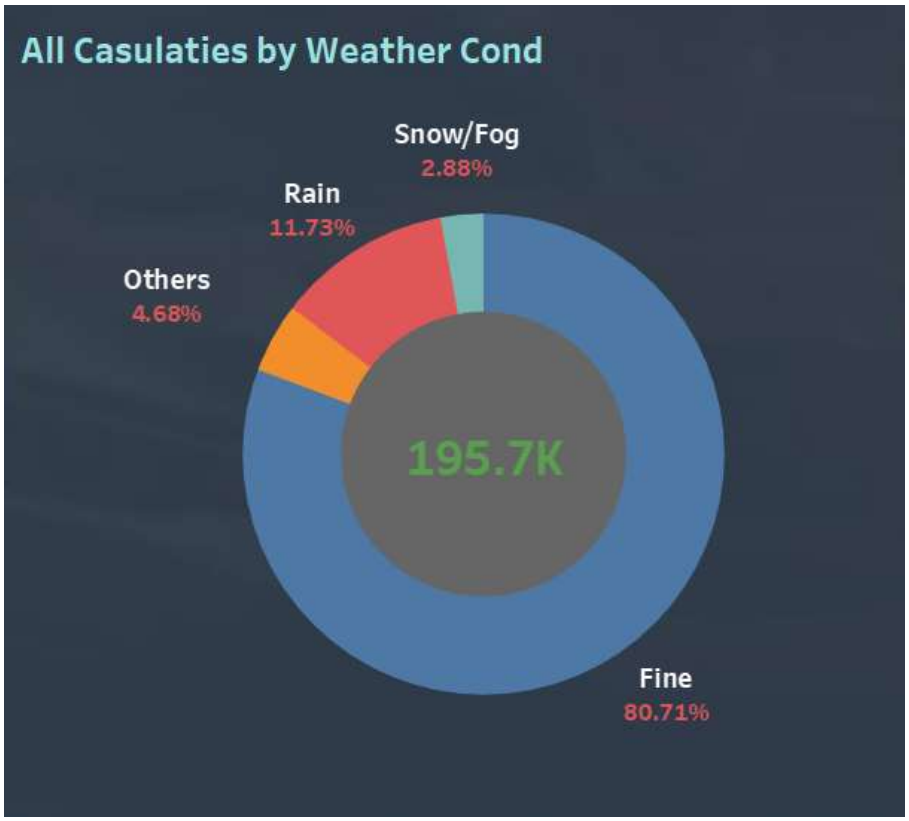


Appendix B

B.1. Accidents by Road Condition

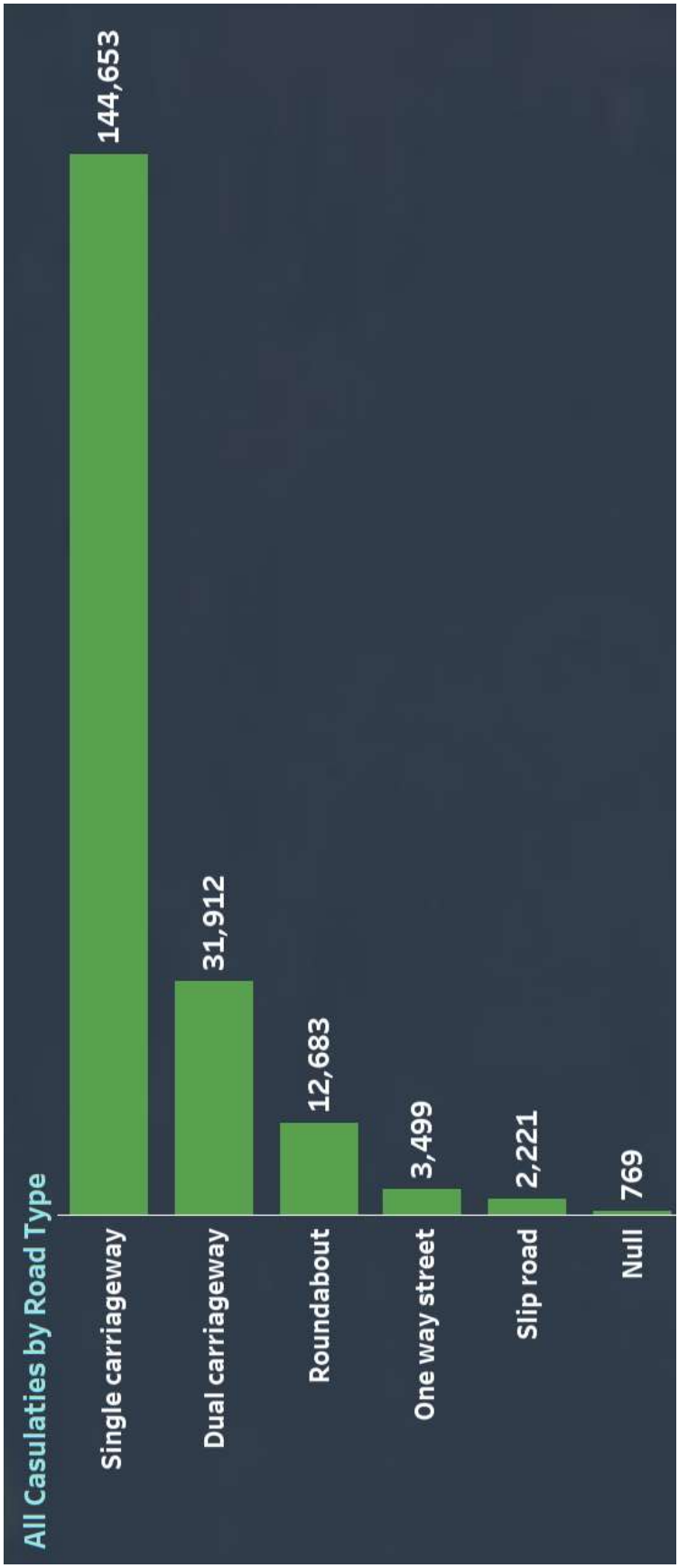


B.2. Accidents by Weather Condition



Appendix C

Accidents by road types



Appendix D: Accidents by Location