Analyzing UK Road Safety

Insights from Traffic Accidents and Vehicle Data

*by*

*SAHIL PATEL & NAREN BABU*



CONTENTS

* Introduction
* Problem Statement
* Motivation of the topic
* Data Cleaning and Prep
* Descriptive Analysis
* Conclusion

**INTRODUCTION**

This report, we delve into the critical issue of road safety in the United Kingdom, leveraging a comprehensive dataset sourced from 2019-2022 made publicly available via Kaggle. Our primary objective is to uncover patterns and insights related to traffic accidents and vehicle involvements, with a focus on enhancing road safety strategies. This study not only aims to identify the key factors contributing to road accidents but also seeks to propose data-driven recommendations for improving traffic management and safety measures. The urgency of this investigation stems from the ongoing challenges in road safety, highlighting the need for continuous assessment and improvement in traffic regulations and infrastructure.

**PROBLEM STATEMENT**

Despite advancements in vehicle technology and road safety measures, traffic accidents remain a significant concern in the UK, posing risks to life and property. This report addresses the pressing need to analyze accident data to identify underlying causes and trends. By examining factors such as accident locations, times, vehicle types, and driver behaviors, we aim to provide a comprehensive understanding of the current state of road safety. The ultimate goal is to inform policy-making and public awareness initiatives, contributing to the reduction of accident rates and enhancing overall road safety.

**MOTIVATION FOR THIS TOPIC**

We chose to focus on UK road safety and traffic accidents for a deeply personal reason. Having witnessed the devastating impact of a severe road accident in our community, we became acutely aware of the fragility of life and the critical importance of road safety. This incident left a lasting impression, driving our curiosity to understand the broader patterns and causes behind such tragedies. By analyzing accident data, we aim to contribute to making roads safer, hoping that our research can play a part in preventing similar incidents in the future. This project is more than an academic endeavour for me; it's a pursuit born out of a desire to make a tangible difference in the realm of road safety.

**DATA CLEANING & PREP**

We acquired a comprehensive dataset from Kaggle, detailing 660,680 instances of traffic accidents across the UK, structured into 14 informative columns. This dataset encompasses a broad spectrum of variables, each providing insights into different aspects of traffic accidents:

Accident\_Severity: Categorized into three types—'Fatal,' 'Serious,' and 'Slight,' indicating the severity of each accident.

Accident Date: Documenting the date on which each accident occurred.

Latitude and Longitude: Geographical coordinates pinpointing the accident locations.

Light\_Conditions: Describing the visibility conditions at the time of the accident.

District Area: With 420 unique districts represented, this variable specifies the area where each accident happened

Number\_of\_Casualties: The count of individuals injured or killed in the accident.

Number\_of\_Vehicles: The number of vehicles involved in each incident.

Road\_Surface\_Conditions: The state of the road surface during the accident, such as wet or dry.

Road\_Type: The type of road on which the accident took place.

Urban\_or\_Rural\_Area: Classification of the accident site as either urban or rural.

Weather\_Conditions: Grouped into 'Fine,' 'Rain,' 'Snow or Fog,' and 'Others' to describe the weather during the accident.

Vehicle\_Type: Grouped in Tableau for analysis into categories including 'Agricultural Vehicle,' 'Bus,' 'Car,' 'Goods Vehicle,' 'Van,' 'Motor Cycle,' and 'Other Modes of Transport.'

To ensure the dataset's integrity, we undertook a rigorous data cleaning process. This included the elimination of duplicate entries and careful management of missing values. We standardized the data formats, such as converting the 'Accident Date' to a uniform date format suitable for time-series analysis. We ensured consistency in the labeling of categorical variables like 'Light\_Conditions,' 'Road\_Surface\_Conditions,' and 'Weather\_Conditions,' as inconsistencies could lead to flawed analytical outcomes.

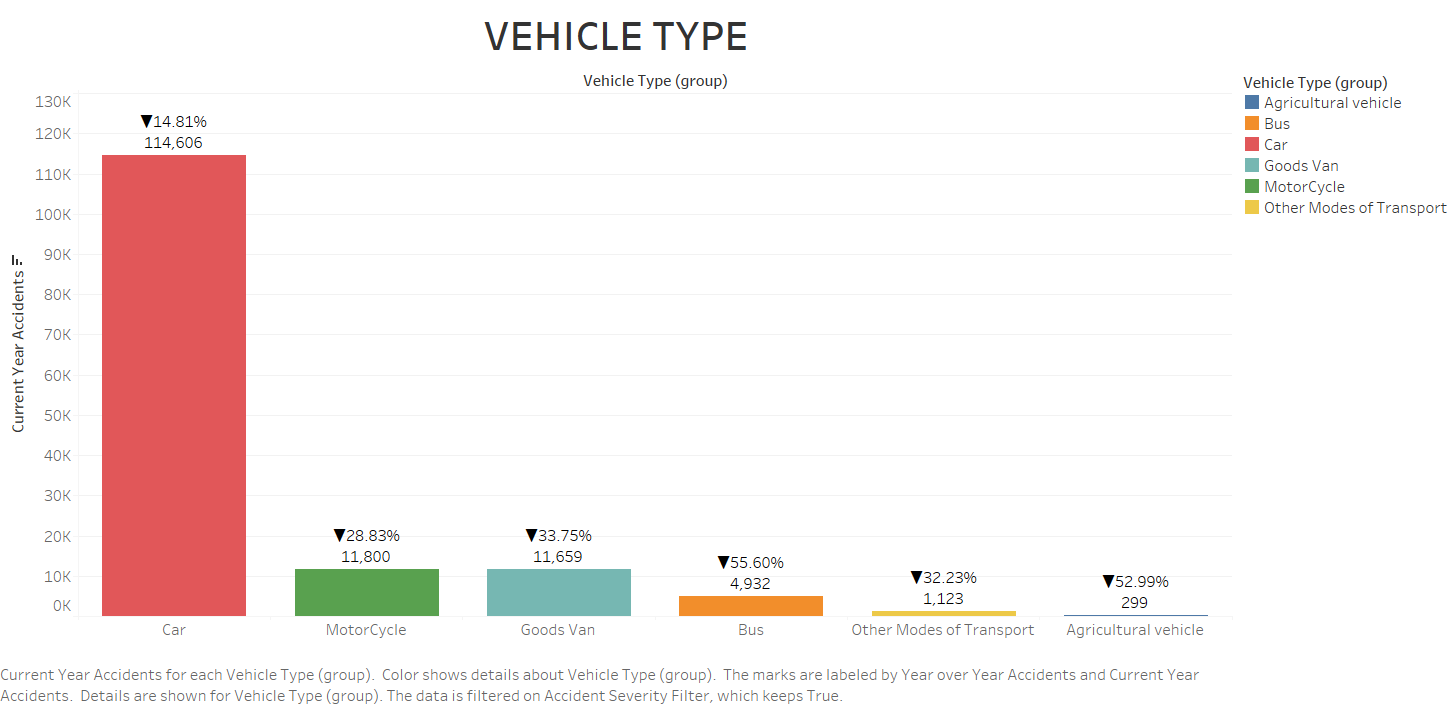
Furthermore, we created derived variables to enhance our analysis. This included extracting the day of the week, month, and year from the 'Accident Date' to explore trends over time and categorizing 'Vehicle\_Type' into broader groups for a macro-level understanding.

These meticulous steps in preparing our dataset were essential for establishing a solid foundation for our analysis, aiming to draw meaningful conclusions and recommendations for improving road safety.

**DESCRIPTIVE ANALYSIS**

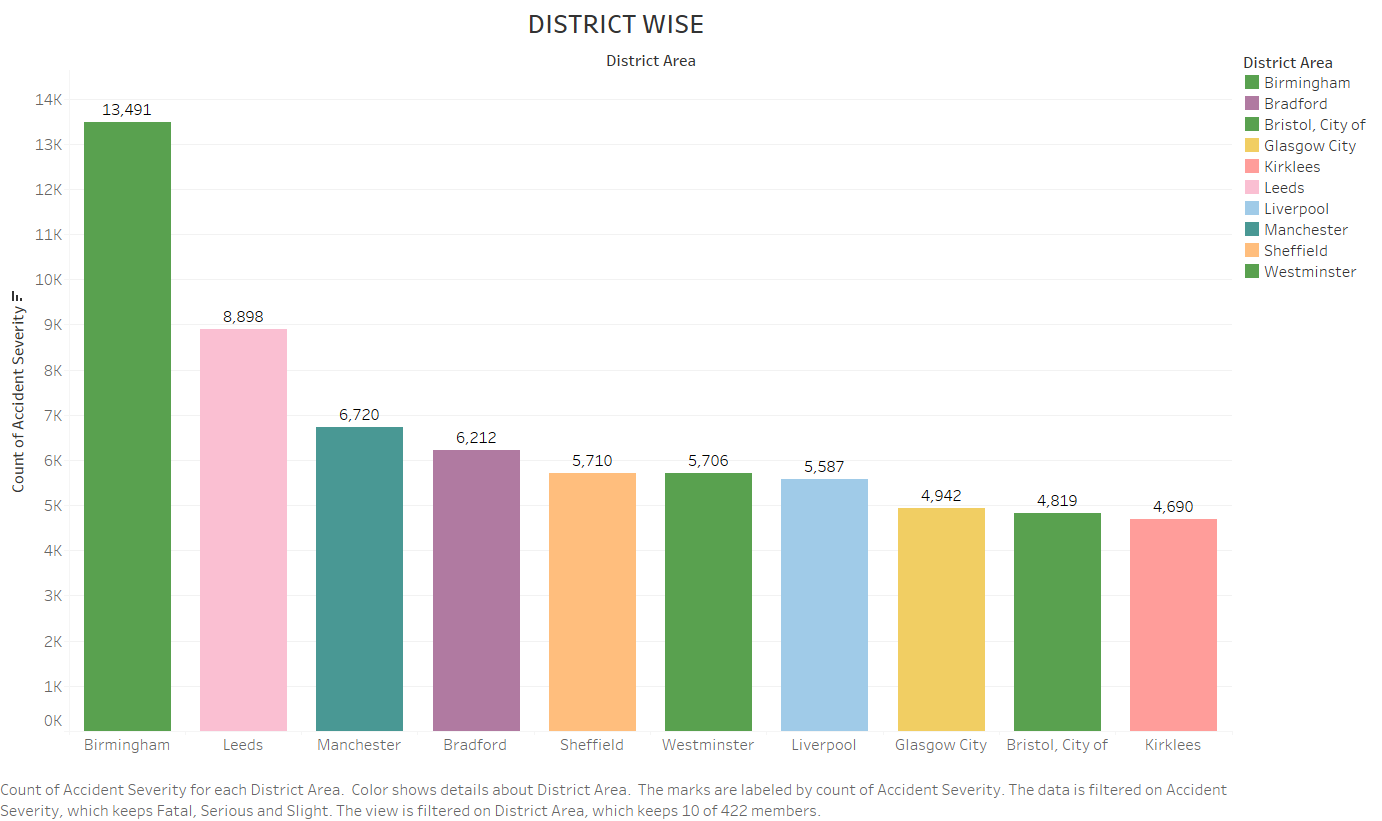
In our Tableau analysis, we meticulously crafted each visualization to uncover and communicate specific insights from the traffic accident data:

Bar Chart for Vehicle Types: We chose a bar chart to display the number of accidents by vehicle type because of its clear comparative abilities. The bar chart showed that cars are involved in the highest number of accidents, which is critical information for stakeholders aiming to reduce traffic incidents. This visualization makes it evident that safety initiatives should prioritize car-related incidents.



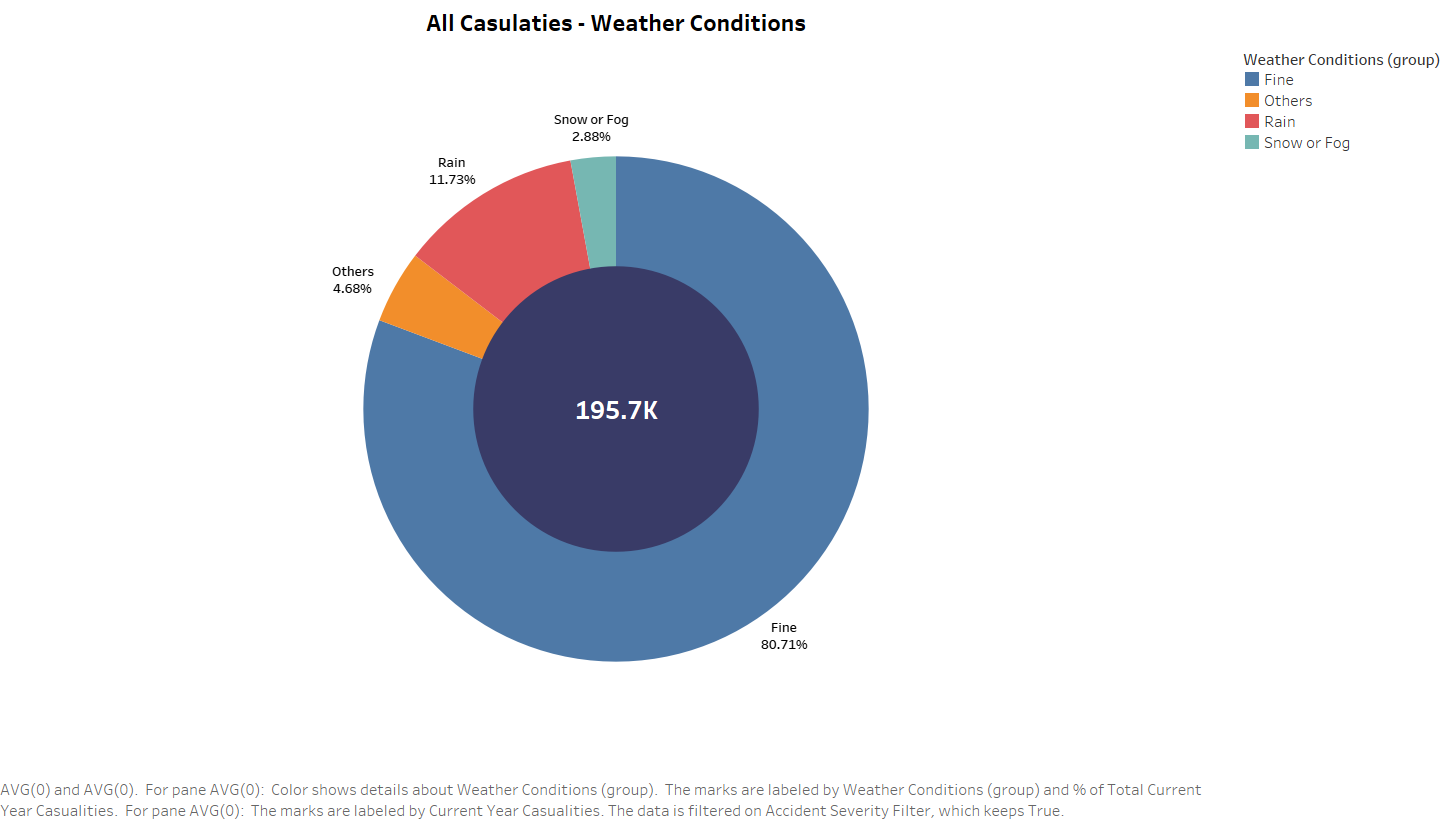
Year-on-Year Percent Change: We used a calculated field to depict the year-on-year percentage change in accidents. This provided a clear visual cue on the trend of accidents over the years, enabling viewers to quickly ascertain whether the situation is improving or worsening.

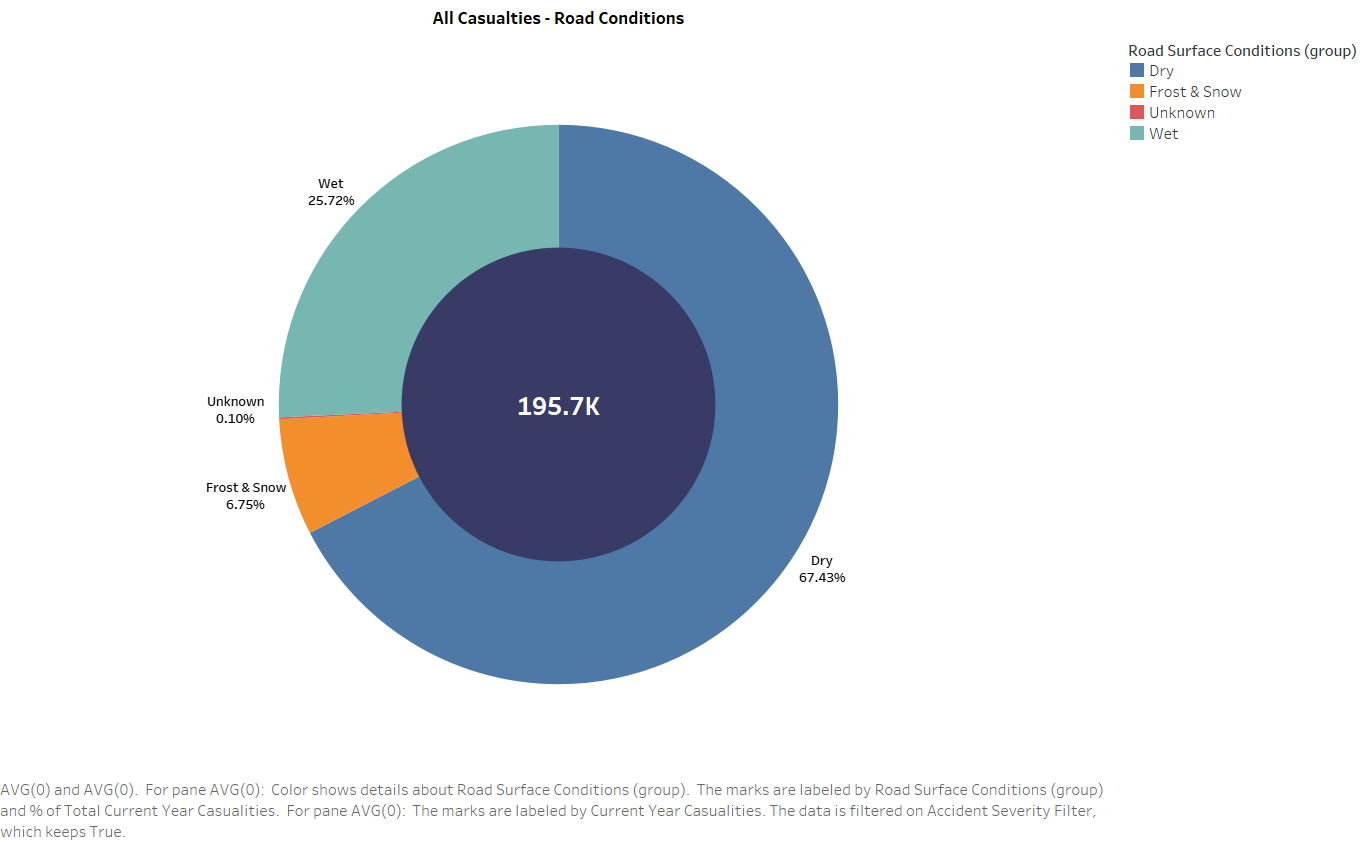
Bar Chart for District Accident Frequencies: To identify which districts have higher accident occurrences, we used a bar chart that ranks the top 10 districts. Birmingham, Leeds, and Manchester topped the list, signalling the need for targeted safety measures in these areas. Birmingham's prominence in this chart with the highest number of weather-affected accidents necessitates weather-specific road safety measures.



Bar Chart for Weather-Impacted Districts: A bar chart was also used to illustrate which districts' accident rates were most affected by weather conditions. Birmingham again stood out, suggesting that weather-proofing measures could be particularly effective there. Numbers here depicted a stark hierarchy of accident frequencies, with Birmingham at 13,491 accidents, substantially more than Leeds at 8,898, painting a clear picture of regional disparities.

Pie Charts for Casualties by Weather and Road Conditions: We employed pie charts to represent the proportion of casualties under different weather and road conditions. These charts effectively communicated that the majority of casualties occurred during fine weather on dry roads, potentially indicating overconfidence or complacency under seemingly 'safe' conditions.The pie charts quantified with 80.71% occurring in fine weather and 67.43% on dry roads, challenging assumptions about 'safe' conditions.

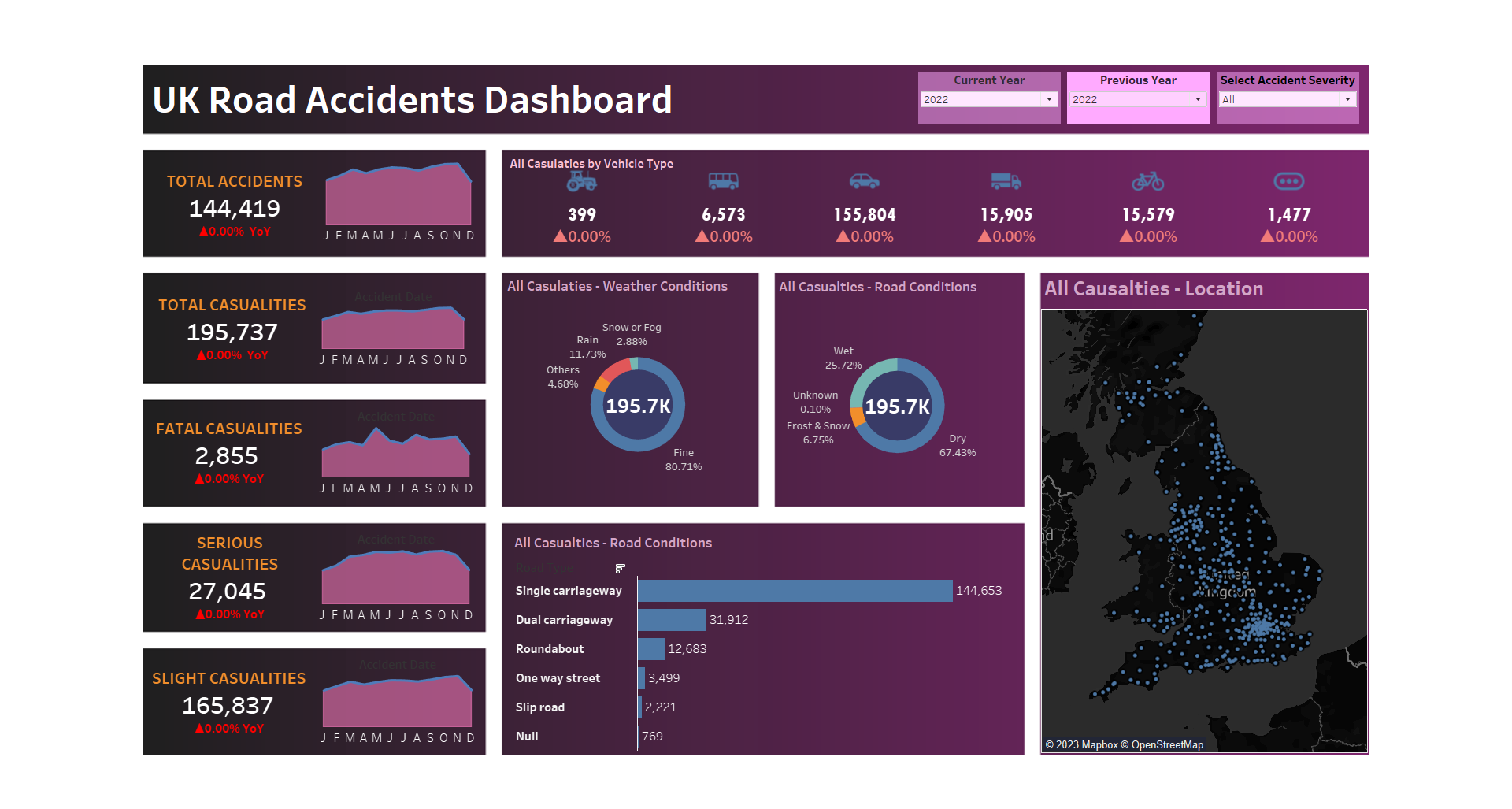




Geographic Heat Map: Lastly, the use of a geographic heat map was pivotal in visually identifying accident hotspots. This type of visualization is particularly effective in spatial analysis, allowing for the easy pinpointing of areas that require immediate attention regarding road safety. The density of incidents in certain areas was highlighted, with brighter intensities corresponding to higher numbers of accidents, signalling urgent need for interventions in those hotspots.

**CONCLUSION**

Our analysis of four years' worth of UK road traffic accident data offers crucial insights that directly address our initial problem statement. We observed that cars are predominantly involved in accidents, and a significant number of casualties occur under fine weather conditions and on dry roads, which may indicate a sense of complacency among drivers during good weather.



*DASHBOARD*

The district-level data revealed that urban areas, particularly Birmingham, have higher accident rates, suggesting a need for urban-specific traffic management strategies. Our year-on-year comparison highlighted a need for continuous improvement in road safety measures, as the fluctuations in accident rates point to varying effectiveness.

In conclusion, based on our findings, we recommend:

* Focused educational campaigns on safe driving, especially during conditions perceived as safe.
* Urban infrastructure improvements in high-frequency accident areas.
* Further research into the causality behind weather-related accidents.
* Continuous monitoring and analysis of road accident data to evaluate the impact of implemented safety measures.

These recommendations aim to support policymakers in formulating targeted interventions to enhance road safety and ultimately reduce traffic-related casualties.

**REFERENCES**

*UK Vehicle Accident Database 2019-2022*. (2023, July 19). Kaggle. https://www.kaggle.com/datasets/charliescott556/uk-vehicle-accident-database-2019-2022/data