“UK Traffic Accidents Analysis using Tableau.”

Team Members:
Apurva Padwal (Apadwal2)
Soham Narvekar (Sn24)

Under the guidance of **Prof. Michael Wonderlich**

INTRODUCTION

The number of traffic accidents worldwide increases by an average rate of 15% annually. The United Kingdoms is no stranger to this trend and experience a similar growth in its road accident rate. 137,013 reported casualties of all severities. In pursuit of lowering accidents and fatalities, improving infrastructure, and promoting safer driving practices, societies and governments may benefit from the analysis of road accidents. A greater quality of life for residents, lower healthcare expenditures, and more economic growth are potential findings of such a study.

OBJECTIVE

Our goal is to create an intuitive and user-friendly platform that facilitates comprehensive data interpretation, identifies contributing factors, and enhances public awareness by leveraging the dynamic and descriptive dashboard capabilities of Tableau. Our objective is to position ourselves as a valuable resource for businesses across multiple sectors, including marketing, finance, healthcare, and retail.

DATASET DESCRIPTION

We have opted to utilize a dataset comprising of details on prior traffic incidents. This dataset encompasses information regarding the accident sites' location, denoted by their respective longitude and latitude coordinates, alongside their corresponding district regions. Furthermore, it incorporates various accident-related particulars, such as the severity of the incident, date of occurrence, number of casualties, and the involved vehicle's types. Additionally, it encompasses various environmental factors that may have impacted the probability of an accident or injury, including weather conditions, location (urban or rural), road type, and surface conditions.

Dataset: UK Traffic Accidents
Data Source: UK Accidents
Columns: 14
Rows: 642,810

DATA CLEANING APPROACH

To extract detailed insights, we wanted to undertake a comprehensive examination of the dataset, consisting of 660,000 road accidents, utilizing the parameters. Missing values were taken care of by removing them. The reason for this decision was as most values were categorical and thus it would be better to remove them rather than skewing the dataset. Date format was changed to ensure better accessibility to reading data.

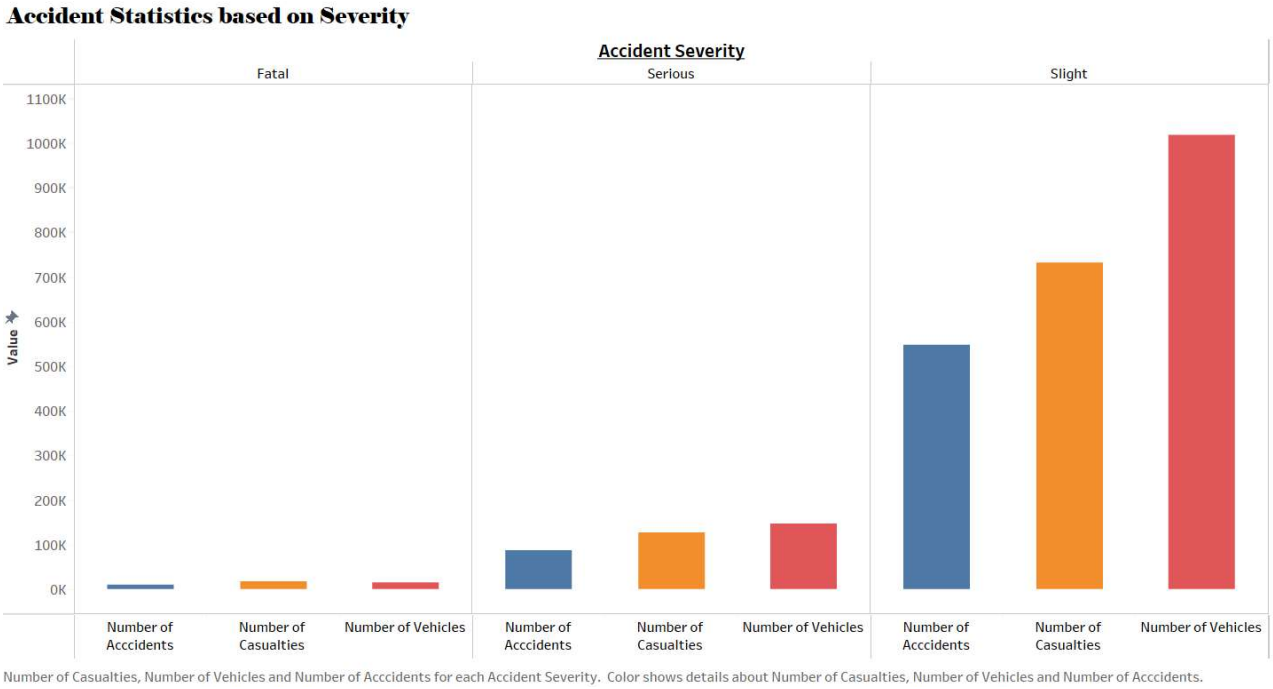
Next, we used Tableau itself to clean some data for Vehicle Type and Weather conditions. These two parameters had multiple values in them as categories and we did not need all of them to be separate. E.g., There were 4 different types of motorcycles, 2 types of buses, Vans, etc. And for Weather, it was again distributed in types based on wind and season. To make the data simplified, we decided to group these similar values and use them for analysis. This ensured complexities and confusion was removed for viewers to grasp better understanding of the actual goal.

CHALLENGES

1. The most challenging part of any project relying on data is the processing of the raw data to fit your need. Procuring the right dataset is also a crucial step. We found a dataset on Kaggle which had all the right parameters required for us to create and insightful visualization. The dataset had missing values in various columns which were cleaned to present meaningful insights.
2. The next challenging part was finding correlations between the datasets which can give us information about traffic accidents. For example, the severity of the accidents paired with the type of vehicle involved would help authorities understand the trend between how diverse types of vehicles impact the severity accidents.
3. One more challenge was figuring out how to use calculated parameters within the dashboard to filter the data and visualize it based on its type on the same screen for all small visuals/tables. We watched multiple YouTube videos and tutorials to find the workaround and move ahead with achieving what we needed.
4. One final challenge we faced was to design a good layout for the Dashboard. Choosing appropriate colors, making the visuals interactive, overall design, and further making the Dashboard easy for navigation while being visually appealing were all the tiny obstacles that came our way.

EXPLORATORY ANALYSIS

- To better understand the situation, we first conduct some initial exploratory data analysis.
- In Fig 1, we try to find statistics about crashes based on the severity of accidents. We find that most accidents that take place are not severe, and this can be observed from the grouped bar chart below.
- There is a stark difference in the number of accidents with fatal and slight severity. Even so when looking at the ratio of casualties to accidents, those with slight severity have a higher ratio as compared to the accidents labelled fatal.



(Fig 1. Accident Statistics based on Severity)

- In Fig 2, We try to find out statistics about accidents based on the types of vehicles involved in the accidents. We see that most of the accidents in the spectrum fall be involved of high CC vehicles.
- This trend is particularly evident for Motorcycles as they are categorized in the dataset based on their CC. While vehicles are based on transport and how many goods they can transport.

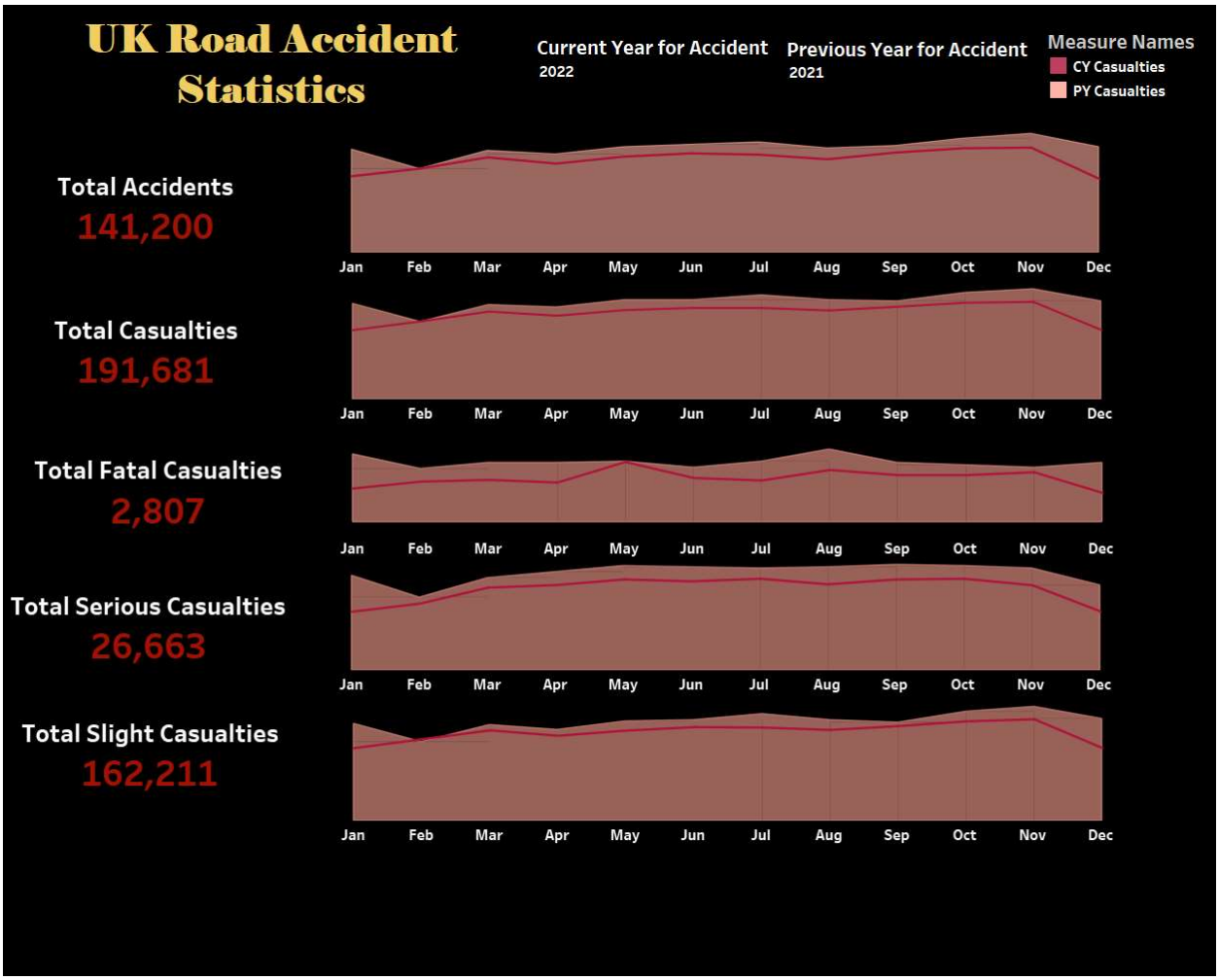
Accident Statistics based on Vehicle Type

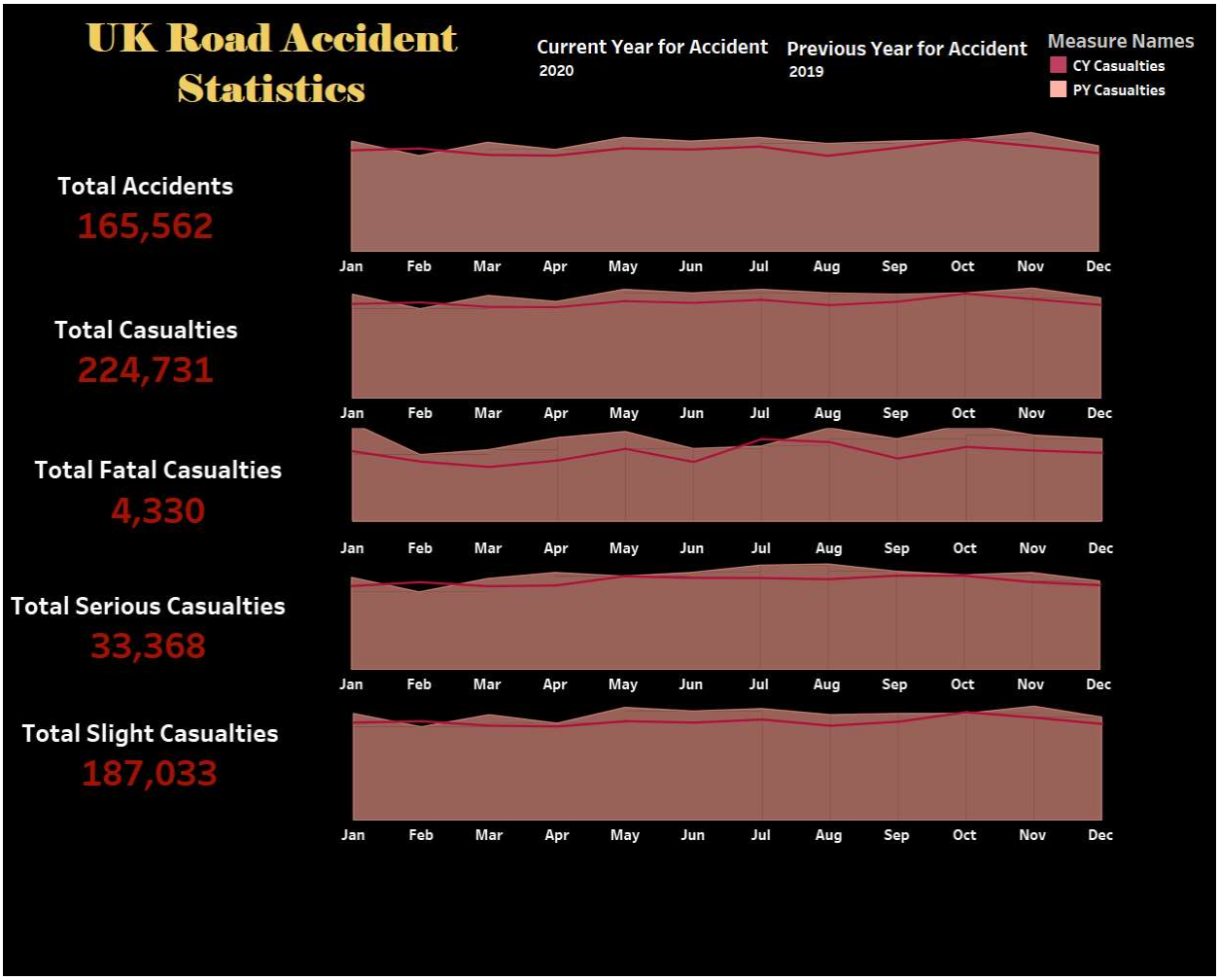
Vehicle Type	Number of Vehicles	Number of Casualties
Car	888,390	659,998
Van / Goods 3.5 tonnes mgw or under	60,710	45,142
Bus or coach (17 or more pass seats)	46,095	34,009
Motorcycle over 500cc	45,881	34,043
Goods 7.5 tonnes mgw and over	30,904	22,841
Motorcycle 125cc and under	26,985	19,824
Taxi/Private hire car	23,793	17,767
Motorcycle over 125cc and up to 500cc	13,641	10,052
Motorcycle 50cc and under	13,678	9,956
Goods over 3.5t. and under 7.5t	10,894	8,083
Other vehicle	9,965	7,313
Minibus (8 - 16 passenger seats)	3,487	2,599
Agricultural vehicle	3,417	2,548
Pedal cycle	347	258
Data missing or out of range	9	8
Ridden horse	6	5

Number of Vehicles and Number of Casualties broken down by Vehicle Type. Color shows Number of Casualties. The marks are labeled by Number of Vehicles and Number of Casualties.

(Fig 2. Accident Statistics based on Vehicle Type)

RESULTS AND INSIGHTS





(Fig 3 & 4. Annual Accident Statistics Comparison)

- The above dashboard compares the statistics between 2 years by month to find if there had been a spike in consecutive years or months which will help us create boundaries for the period to determine potential reasons for the rise/fall in accidents.
- The reasons for the spikes/drops can be associated with many phenomena such as new traffic regulations, weather conditions or can be pinpointed to faults in specific vehicles.
- From the visualization, we see that as summer approaches, the number of accidents rises in March and follows the same trend to the end of November after which Holiday season which is responsible for the drop in accidents.
- Looking Specifically at the statistics from 2021, we see that this was a period of lockdown for the United Kingdom, yet it observed more accidents compared to that in 2022.
- In February 2022, the lockdown restrictions were lifted in the UK and after which we can see there was a rise in the number of accidents as more people started going out for work or obtaining daily necessities.

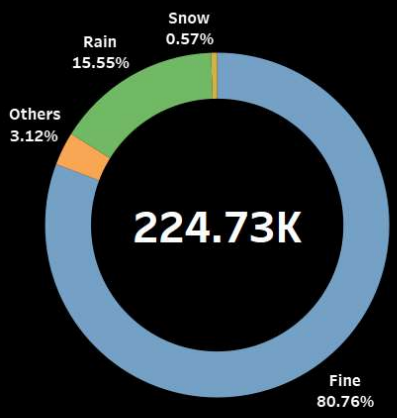
UK Road Accident Statistics

Select Accident Severity
All

All Casualties by the Vehicle Type

704	9,693	173,144	19,952	19,117	2,121

All Casualties by the Weather Conditions



All Casualties by Location/ District Area

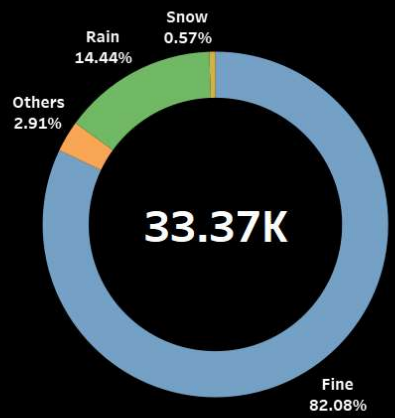
UK Road Accident Statistics

Select Accident Severity
Serious

Serious Casualties by the Vehicle Type

88	1,323	25,934	2,979	2,730	314

Serious Casualties by the Weather Conditions



Serious Casualties by Location/ District Area

In the above dashboard we show statistics about accidents in the UK with Geo positional insights.

- The visualization shows statistics about accidents based on their severity.
- The very top shows the number of specific vehicles involved in accidents, for example, the number of cars, buses, motorcycles etc.
- Below we have a donut chart which shows the weather condition at the time of accidents.
- Next to the donut chart we have a map of the United Kingdoms with marking on the maps showing statistics about the accidents taking place segregated by districts. Hovering over the points shows us the number of accidents, total fatalities, and casualties.

CONCLUSION

- The purpose behind this project were to analyze the trend in accidents happening in the United Kingdom to analyze and find if there is any pattern to the events. So that we could improve the existing system using these insights.
- This project helps us understand there are multiple factors associated with the reasons why accidents take place. The dataset provides us with various external factors such as road conditions, vehicle types and location. We understand that the severity of accidents is not solely dependent on heavy vehicles. Many accidents with serious severity have noticeable involvement of normal vehicles.
- Similarly, weather conditions have little correlation to the number of accidents happening.
- Unfortunately, the dataset used fails to include driver conditions which are very crucial for analysis of the subject at hand. This data might help us understand what part human error or poor judgement has to the severity of accidents.
- The observations and references from this project can help us draft better regulations and driving guidelines to improve road safety.

REFERENCES

[1] Ambani Verma. (2018, November 22). UK traffic accidents. Kaggle. Retrieved March 10, 2023, from https://www.kaggle.com/code/ambaniverma/uk-traffic-accidents/notebook

[2] Transport, D. for. (2022, November 24). Reported road casualties in Great Britain, Provisional Estimates: Year ending June 2022. GOV.UK. https://www.gov.uk/government/statistics/reported-road-casualties-in-great-britain-provisional-estimates-year-ending-june-2022