



# CAR ACCIDENTS IN THE UNITED KINGDOM

  
Natibundit Suntraranusorn 



## Table of Contents

<b>Introduction .....</b>	<b>3</b>
<b>Project Summary .....</b>	<b>3</b>
<b>Dataset Selection .....</b>	<b>3</b>
<b>Analysis Questions .....</b>	<b>4</b>
<b>Working with Tableau .....</b>	<b>5</b>
<b>Connecting to Data .....</b>	<b>5</b>
<b>Preparing Visualization .....</b>	<b>6</b>
Pie Chart .....	6
Line Graph .....	7
Bar Chart .....	8
Stacked Bar Chart .....	9
Area Chart .....	10
Tree Map .....	11
Forecasting Data .....	12
<b>Preparing Dashboard .....</b>	<b>13</b>
<b>Create Story .....</b>	<b>16</b>
<b>Conclusion .....</b>	<b>18</b>
<b>References .....</b>	<b>20</b>

# Introduction

## Project Summary

Every year, millions of people are affected by road accidents which lead to a significant impact on public health systems, economics, and families. Understanding the causes and patterns of road accidents could decrease and protect the accidents which could lead to a sustainable plan for decreasing accidents. Road accidents can be caused by various elements which include but are not limited to road conditions, vehicle types, and weather conditions.

This study works on an in-depth analysis of road accident statistics in the United Kingdom, covering a critical two-year period from 2021 to 2022. The emphasis is on analyzing the different types of traffic accidents and discovering the behaviours, patterns, and underlying reasons that contribute to them. The project uses broad accident data to provide an understanding of the different aspects at play, which range from temporal variables (time of day, day of the week) to environmental circumstances (light, weather), as well as vehicle types involved and road features. The ultimate goal is to inform and guide evidence-based measures for enhancing road safety, lowering the frequency and severity of accidents, and also protecting the well-being of the United Kingdom citizens and tourists.

## Dataset Selection

The dataset covers details of road accidents in the United Kingdom from 2021 to 2022 which includes information such as accident date, accident time, vehicle types, and accident severity, accident location. The dataset contains the following fields:

- **Accident\_Index:** A unique number of accident records.
- **Accident Date:** The date of the accident occurred (format: DD/MM/YYYY).
- **Day\_of\_Week:** The certain day of the week of the accident.
- **Junction\_Control:** Indicates the sort of junction control at the accident location.
- **Junction\_Detail:** Additional information regarding the accident location.
- **Accident\_Severity:** Specifies the severity of the accident.
- **Latitude:** The latitude of the accident location.
- **Light\_Conditions:** Indicates the lighting conditions at the accident time.
- **Local\_Authority\_(District):** The district where the accident happened.
- **Carriageway\_Hazards:** Indicates any dangers on the carriageway at accident time.
- **Longitude:** The longitude of the accident location.

- **Number\_of\_Casualties:** Total number of casualties in the accident.
- **Number\_of\_Vehicles:** Total number of vehicles involved in the accident.
- **Police\_Force:** The police force that responded to the accident.
- **Road\_Surface\_Conditions:** Identifies the road surface conditions.
- **Road\_Type:** Specifies the sort of road where the accident happened.
- **Speed\_limit:** The speed limit where the accident occurred.
- **Time:** The time of the accident (format: HH: MM).
- **Urban\_or\_Rural\_Area:** Specifies the area of the accident happened.
- **Weather\_Conditions:** Describes the weather conditions at the time of the accident.
- **Vehicle\_Type:** Identifies the type of vehicle in the accident.

### Analysis Questions

According to the dataset, we identify the following analysis questions which could seek meaningful insights and guide effective policy and safety measures.

1. What are the peak and trough times for road accidents during the day?
2. What does lighting condition play in the occurrence of accidents?
3. How does the severity of accidents vary by vehicle type?
4. What is the impact of weather conditions on road accidents?
5. How does the time of year affect the number of road accidents?
6. What are the common types of collisions and their frequencies by day of the week?
7. How does the number of vehicles on the road correlate with the number of accidents?

## Working with Tableau

This project uses Tableau, the visual analytics platform, to analyze the dataset, employing visuals to further analyze the deep reason inside the pattern of data. In addition, Tableau could be used as an end-to-end analytics platform which provides the connection to the dataset, analyzing data as well as performing the dashboard and story for presenting the analyzed data. By using Tableau, we could understand and interpret the data for efficiency analysis via the visualizations.

### Connecting to Data

In this project, we use only one data source which comes as a .csv file and we use Tableau to connect the dataset. Tableau could perform the visualization with the live dataset which means when there is a change in the data, it will reflect with the Tableau file in real-time. However, in this project, we will not change any data after importing to Tableau, the connection as an Extract file was chosen.

The screenshot shows the Tableau interface for connecting to a data source. The sidebar on the left contains 'Connections' and 'Files' sections. The 'Connections' section shows 'Road Accident Data' as a 'Text file'. The 'Files' section shows 'Road Accident Data.csv' as a file. The main area displays a preview of the 'Road Accident Data.csv' file, showing 24 fields and 307973 rows. A table of data is visible, including columns for Accident Index, Accident Date, Day of Week, and Junction Control.

Accident Index	Accident Date	Day of Week	Junction Control
200901BS70001	1/1/2021 AD	Thursday	Give way or uncontrolled
200901BS70002	5/1/2021 AD	Monday	Give way or uncontrolled
200901BS70003	4/1/2021 AD	Sunday	Give way or uncontrolled
200901BS70004	5/1/2021 AD	Monday	Auto traffic signal
200901BS70005	6/1/2021 AD	Tuesday	Auto traffic signal
200901BS70006	1/1/2021 AD	Thursday	Give way or uncontrolled
200901BS70007	8/1/2021 AD	Thursday	Give way or uncontrolled
200901BS70008	2/1/2021 AD	Friday	Auto traffic signal
200901BS70009	7/1/2021 AD	Wednesday	Give way or uncontrolled

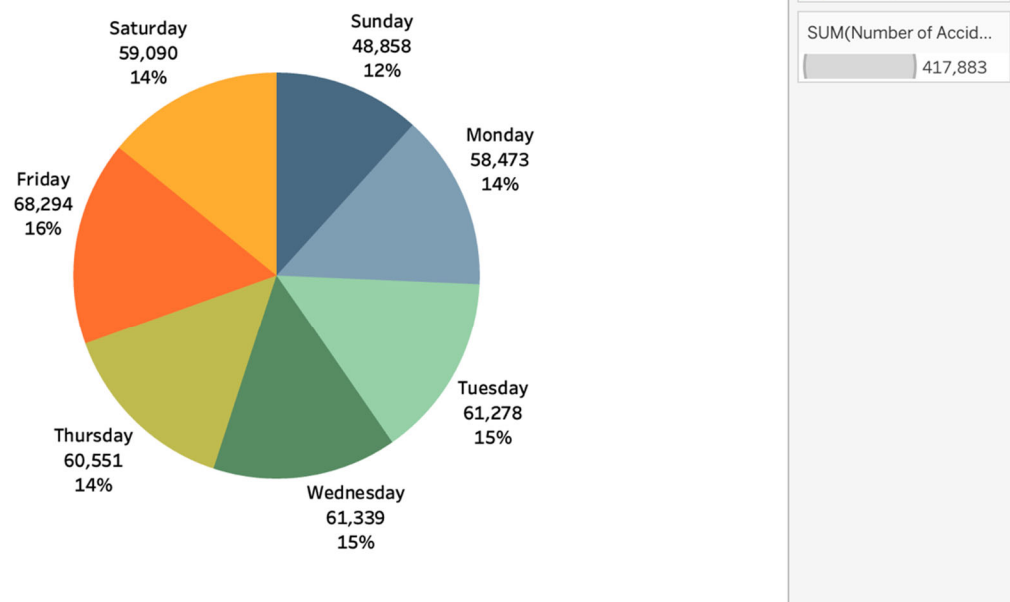
## Preparing Visualization

With the purpose of answering the analysis questions, the visualizations were created in Tableau. The following is some of the visualization that was created in Tableau.

### Pie Chart

The pie chart is mostly used to illustrate the portion or percentage of analyzed data which can easily distinguish the pattern of the data.

1. Percentage of Accidents by day - display the proportion of accidents by day of the week

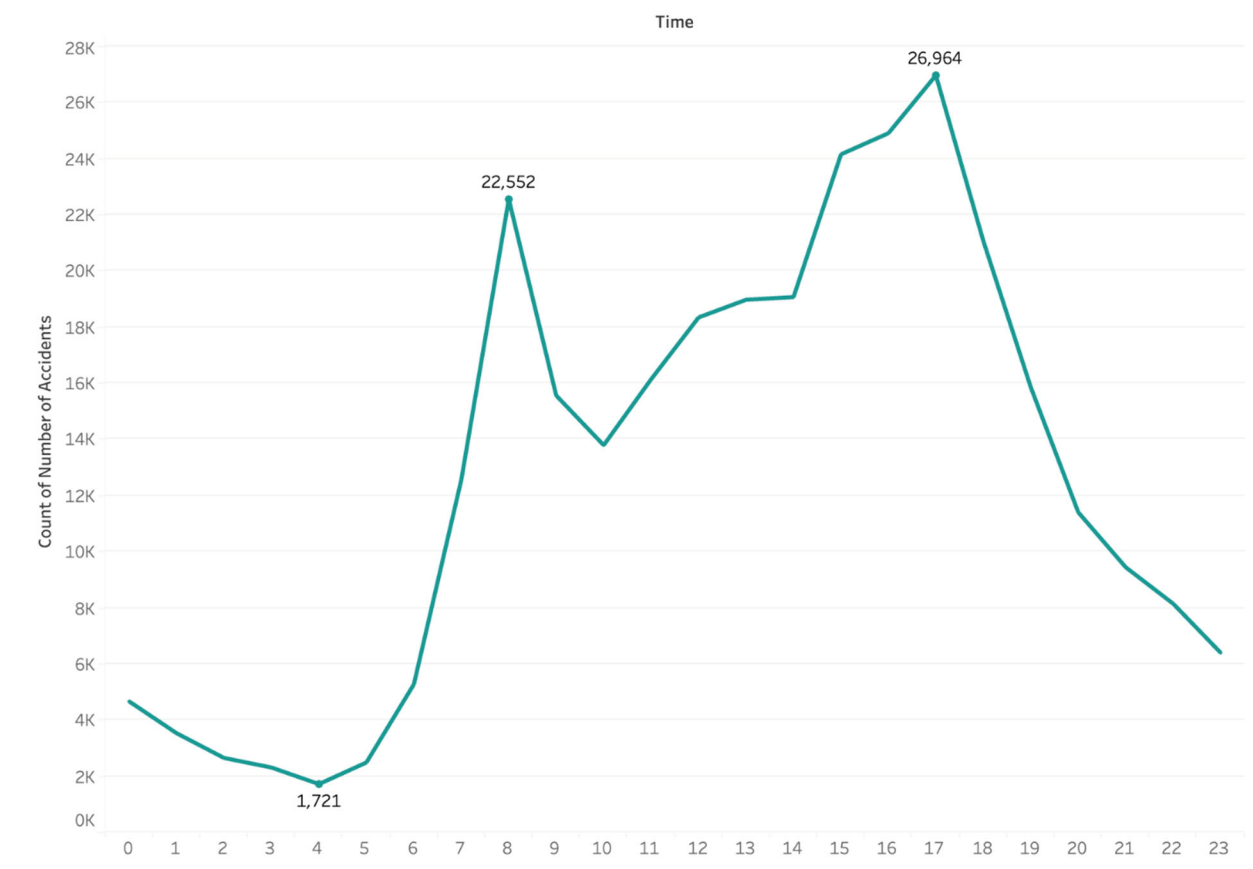


The pie chart illustrates the percentage of accidents by day of the week with a total of 417,883 accidents in seven days of week. The chart shows that the portion of the accident in each day of the week is similar to other days which is around 14 to 16 percent. The day of the week that has the lowest number of accidents is Sunday which is 12 percent or 48,858 car accidents. The day of the week that has the highest number of accidents is Friday which is 16 percent or 68,294 car accidents.

## Line Graph

The line graph can be used to analyze the trend and the cycle of data. In addition, it can be used to compare changes over a period which can compare more than one group of data.

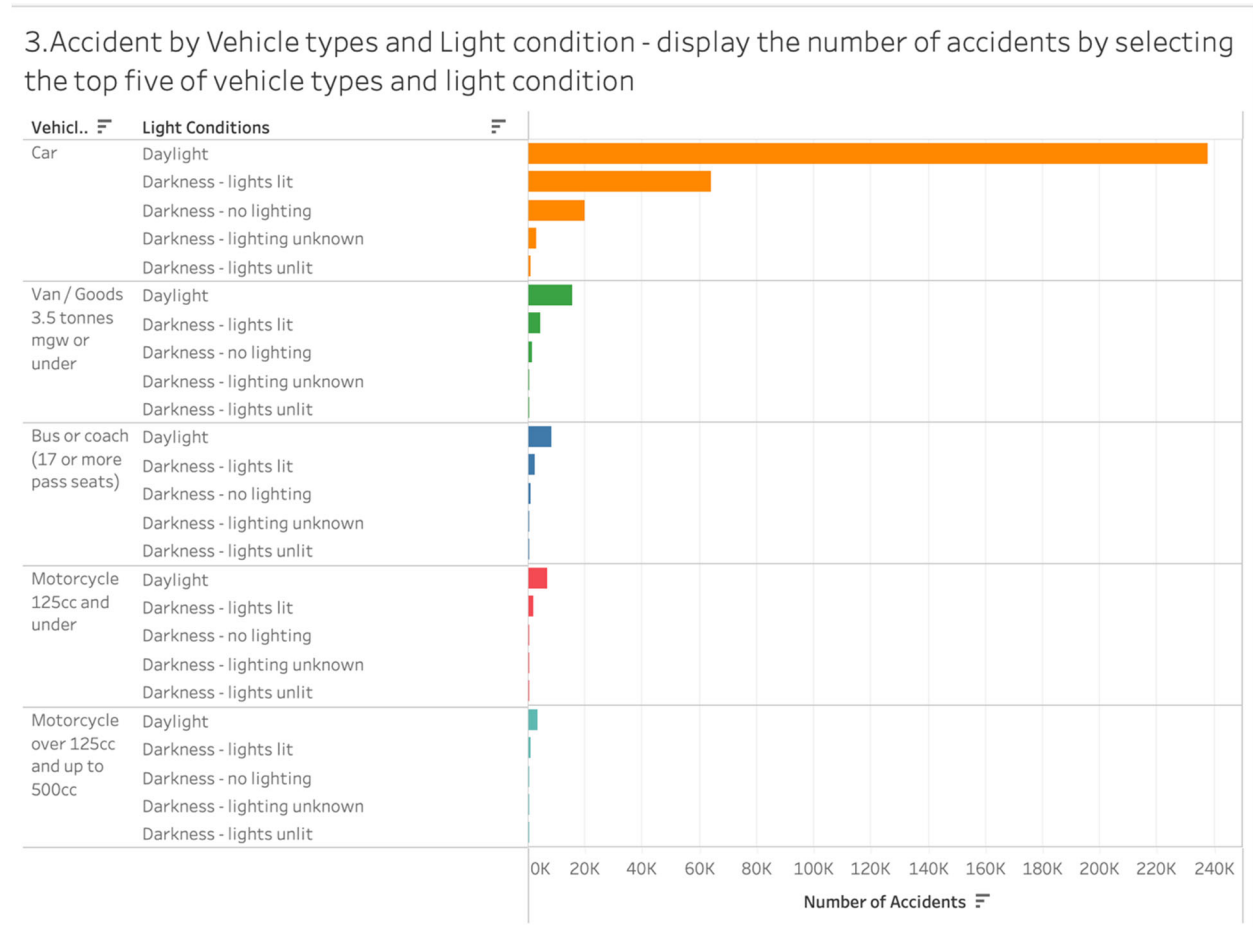
2.Number of Accidents per each hour



The picture displays the number of accidents per hour which ranges of time from 0.00 to 23.00. Two peaks of time have a higher number of accidents than other periods which are 8.00 and 17.00(22,552 car accidents and 26,964 car accidents respectively). On the other hand, the lowest accident is at 4.00 which is 1,721. As we can see, the graph shows, after 4.00 the number of accidents is increasing until it reaches its peak at 8.00 and 17.00. After passing the peek at 17.00, the number of accidents continuously decreased until the end of the day.

## Bar Chart

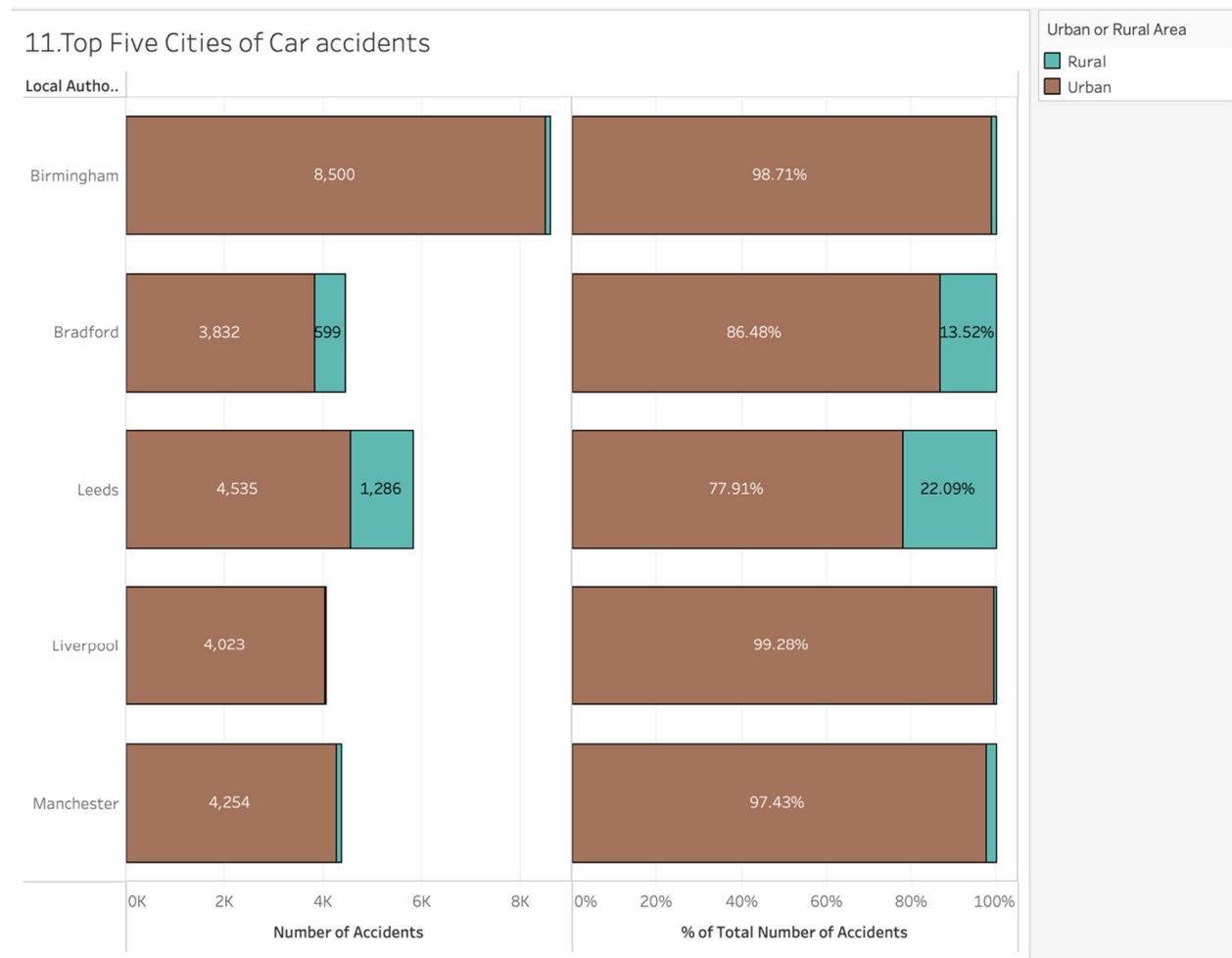
The bar chart usually be used when showing segments of information. The vertical bar charts also have an advantage in comparing different categorical variables.



The picture of the bar chart displays accidents by the top 5 vehicle types and light conditions. The type of vehicle that has the highest number of accidents is a car followed by a van, bus, and motorcycle with the light condition is daylight.



## Stacked Bar Chart



The two stacked bars display the number of accidents between urban and rural areas and the amount of car accidents in urban areas is significantly high compared with rural areas.

## Area Chart

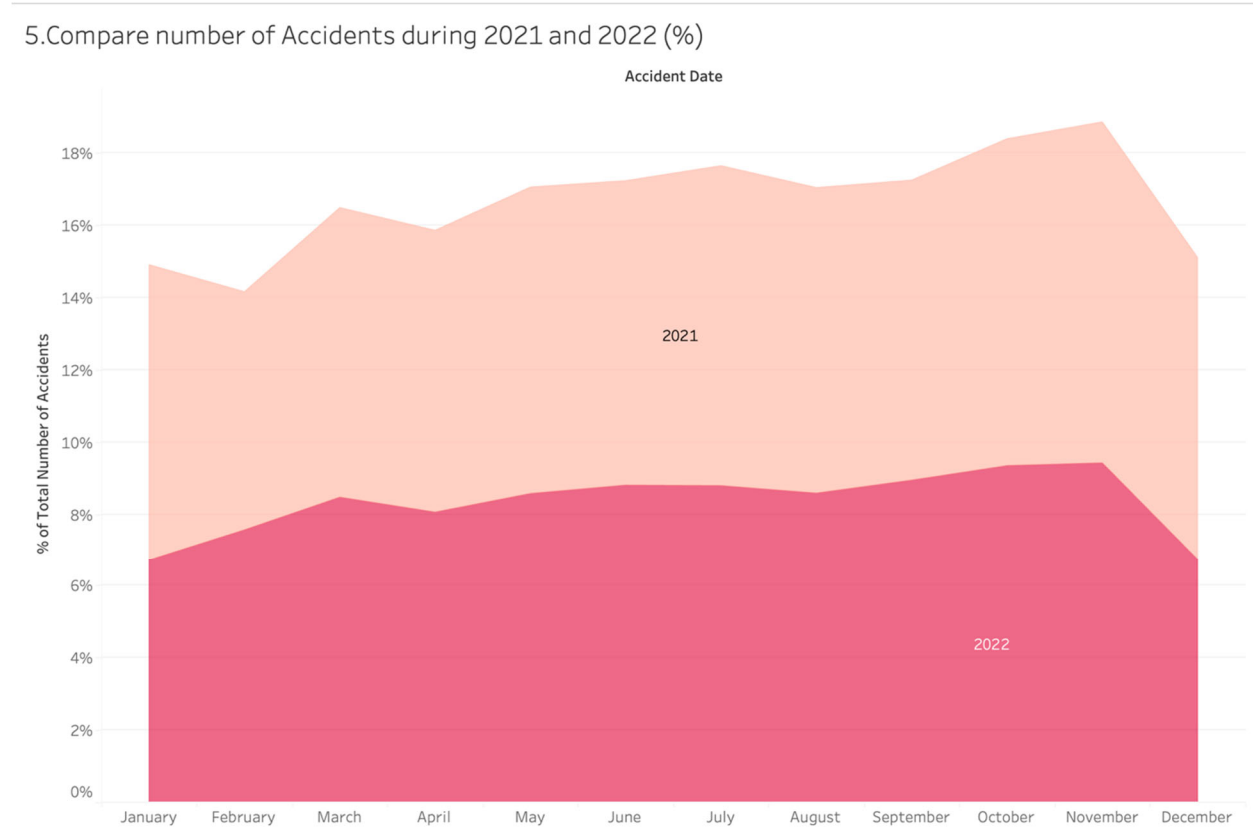
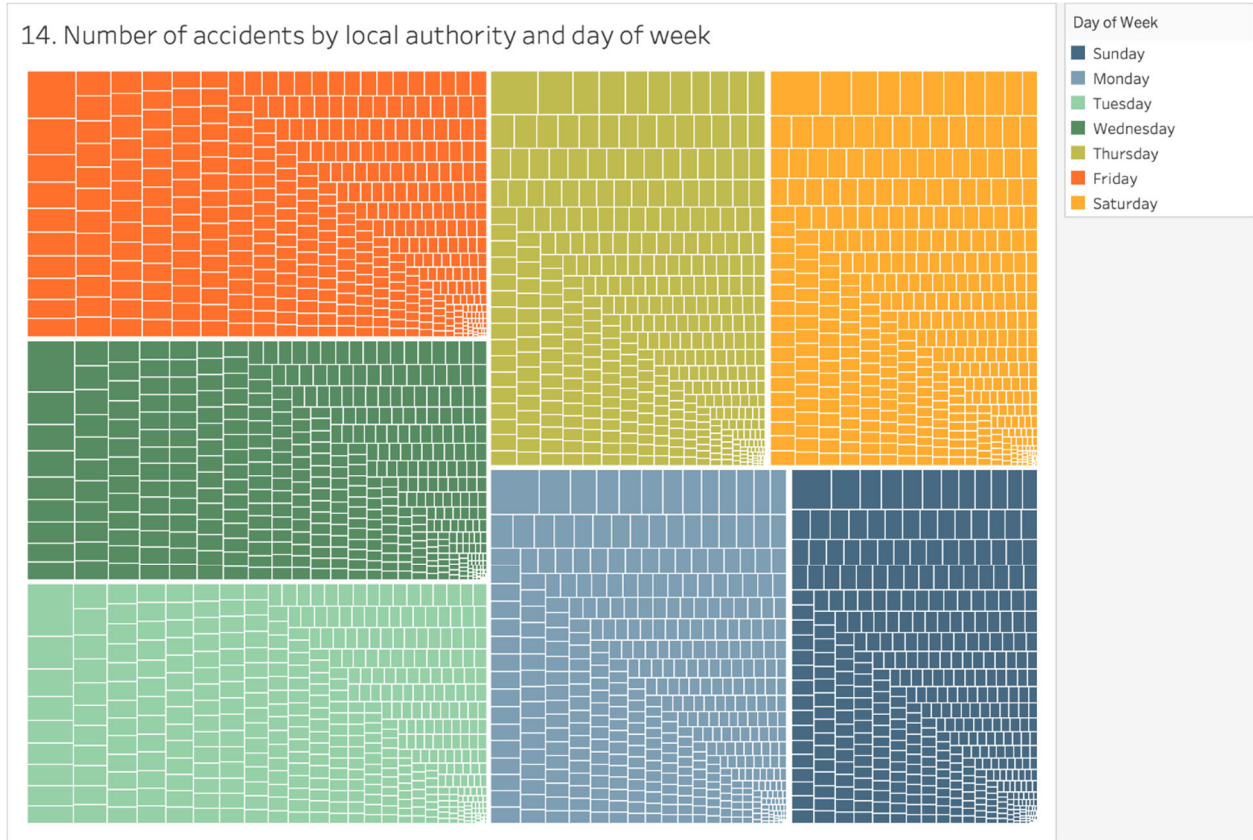


Tableau could also be used to do a table calculation. For example, the graph above illustrates the percentage of accidents with a comparison between 2021 and 2022. We use Tableau to calculate the percent of the total amount of accidents compared in 2 years. The area chart compares the percentage of accidents between 2021 and 2022; both years have a similar percentage pattern. The chart illustrates the highest percentage is in November in both 2 years whereas the lowest percentage is different in each year.

## Tree Map



The tree map displays the number of accidents by day of the week in a different color and the size of each square illustrates the local authority of the accident location. The bigger the square more number of accidents that occurred in that location.

## Forecasting Data

### 4. Number of Accidents over period of time (Low-High month) with forecasting

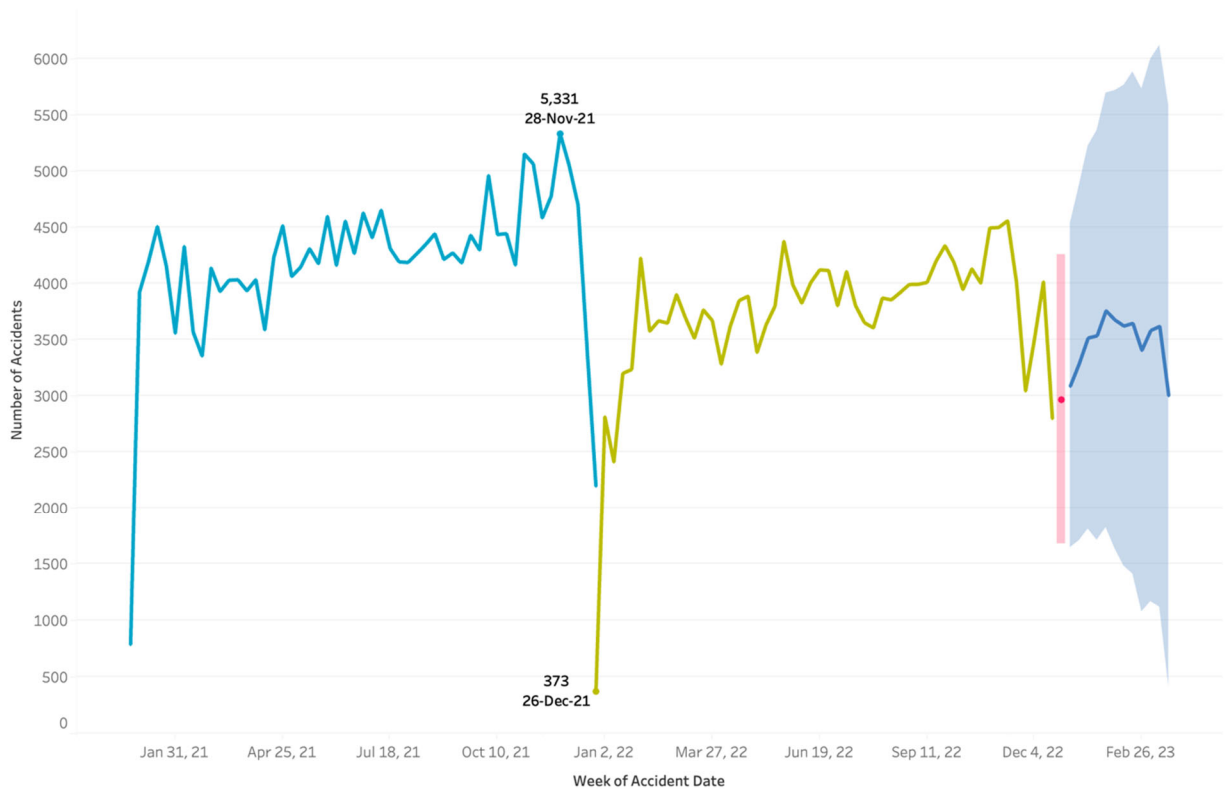
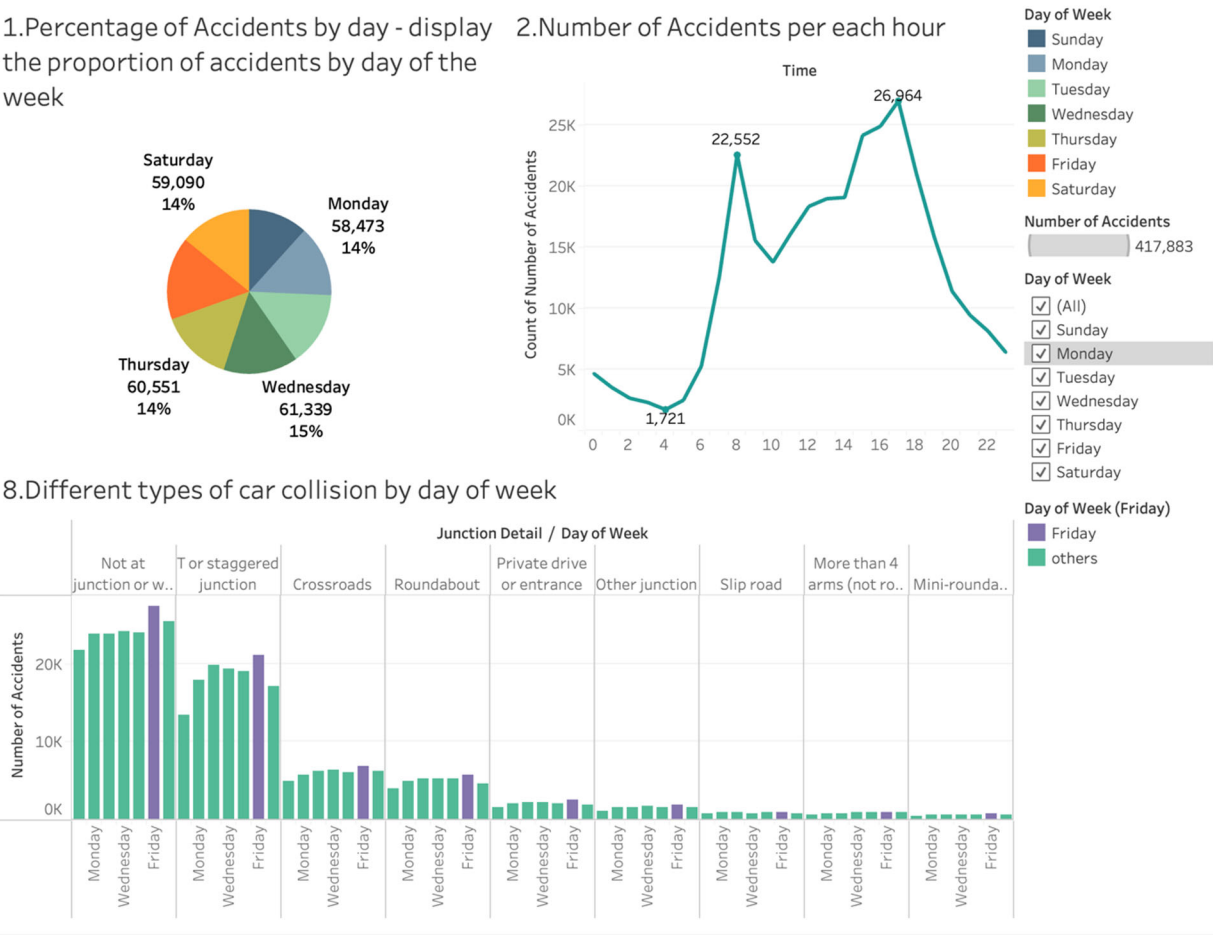


Tableau could be used to forecast the data in the future as the following line chart. The line chart shows the forecasting of the number of accidents in 2023 which the number of accidents might decrease from the previous year.

Preparing Dashboard

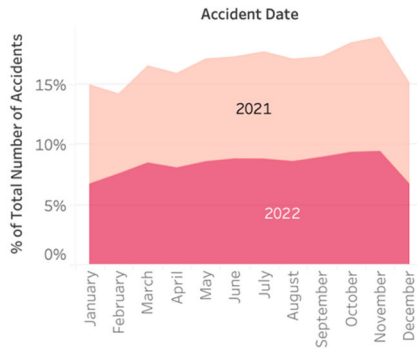
Number of Road Accident by Time and Day of Week

- 1.Percentage of Accidents by day - display the proportion of accidents by day of the week
- 2.Number of Accidents per each hour

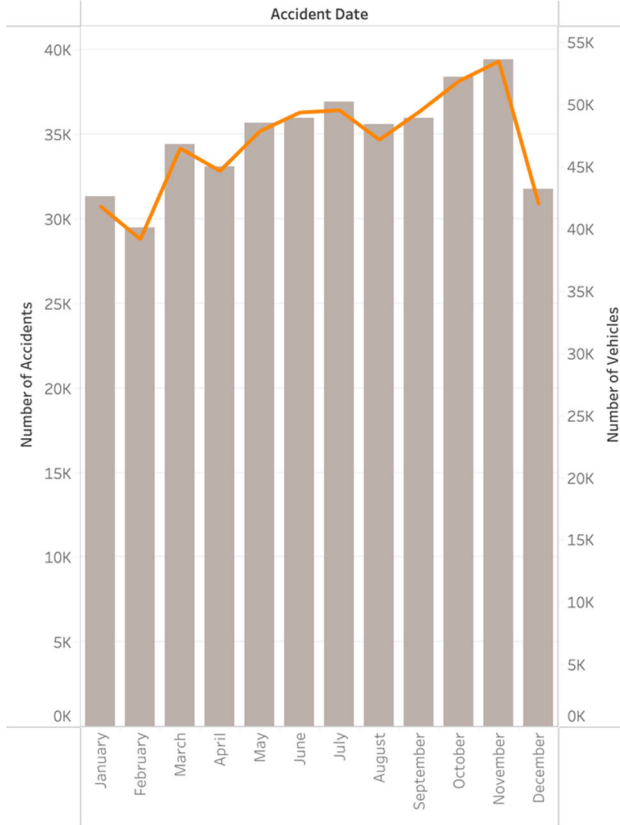


## Number of Road Accident Over the Period of Time

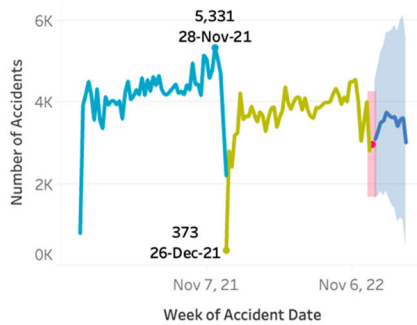
5. Compare number of Accidents during 2021 and 2022 (%)



4. Number of Accidents vs Number of vehicles over period of time (2)

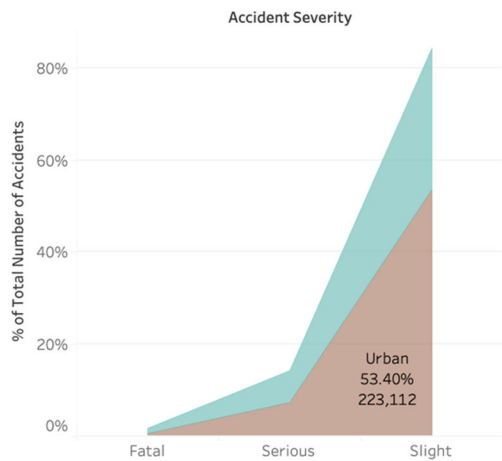


4. Number of Accidents over period of time (Low-High month) with forecasting

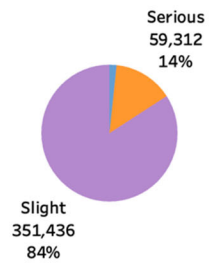


## Number of Accident by Weather Condition and Types of Accidents

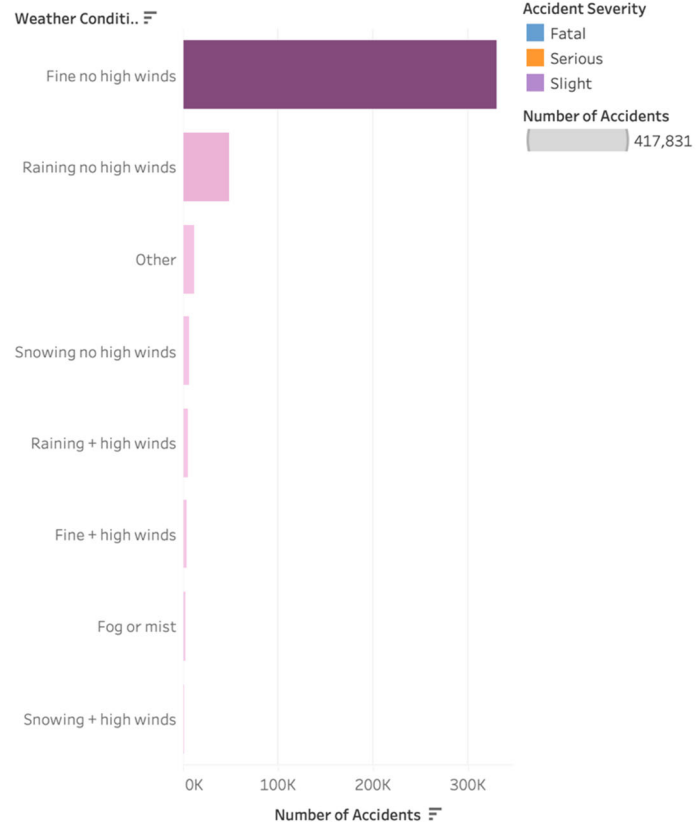
13. Number of accidents by area and accident severity



6. Number of accidents by Type of Accidents

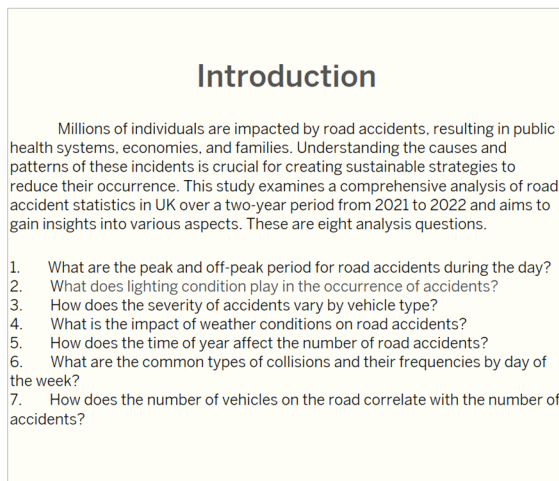
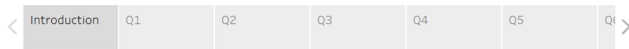


9. Number of accidents by weather condition

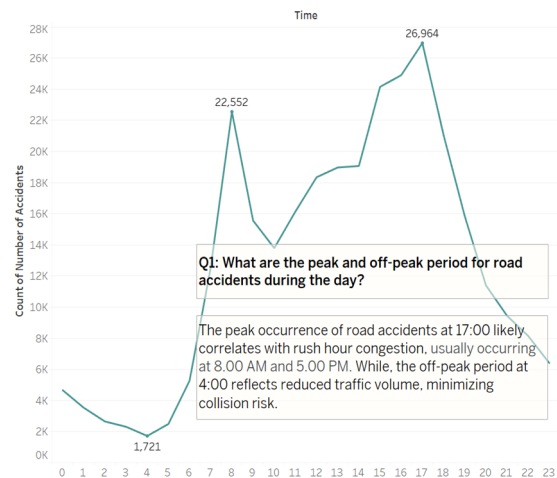


## Create Story

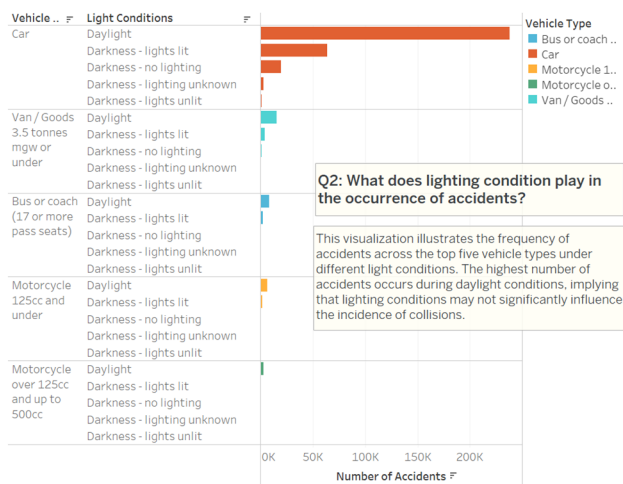
### Car Accident in UK Story



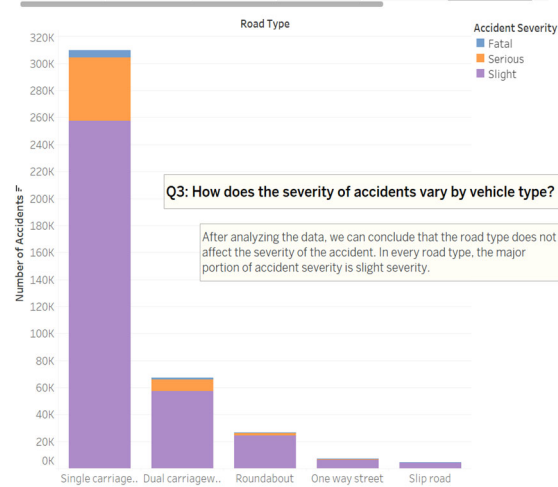
### Car Accident in UK Story



### Car Accident in UK Story

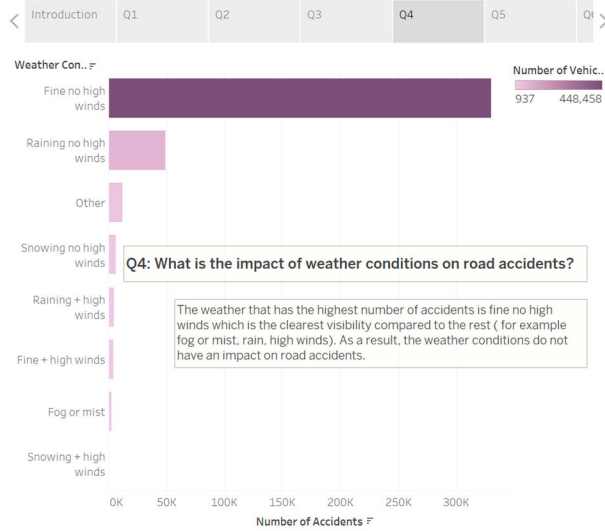


### Car Accident in UK Story

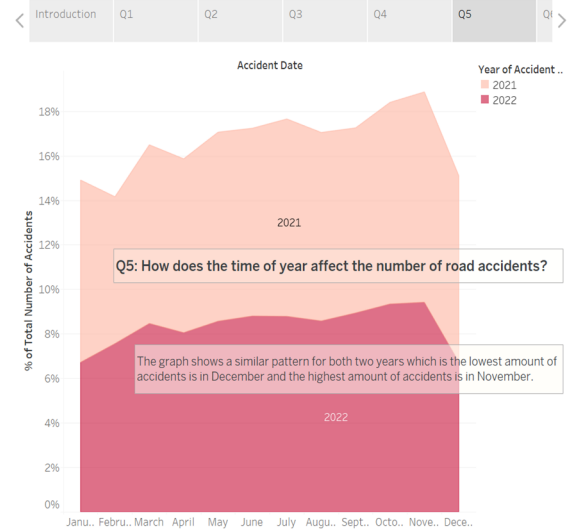




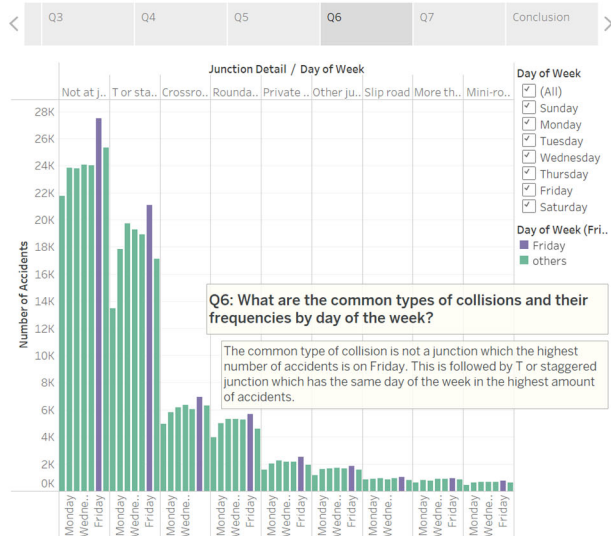
## Car Accident in UK Story



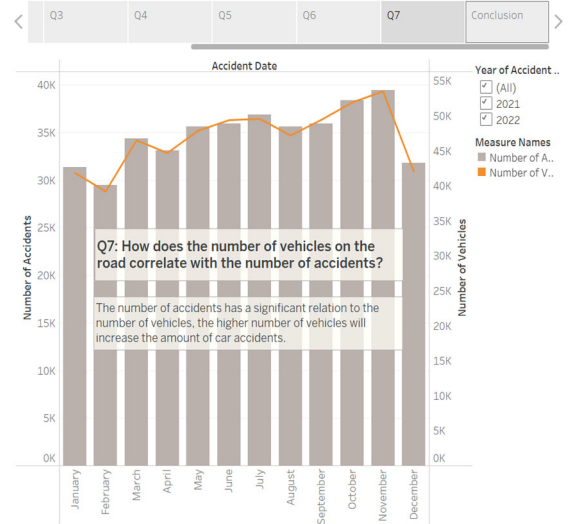
## Car Accident in UK Story



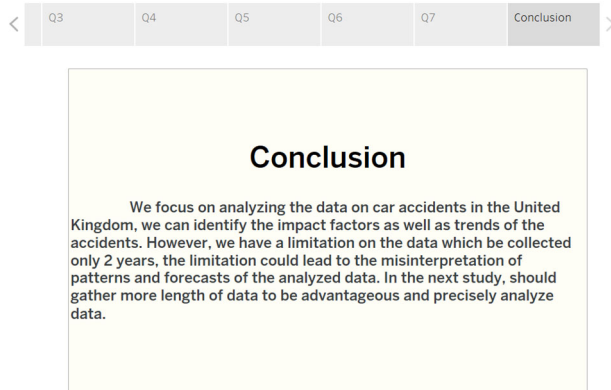
## Car Accident in UK Story



## Car Accident in UK Story



## Car Accident in UK Story



## Conclusion

After analyzing the data by using Tableau, we could answer the analysis questions in the following paragraph.

1. What are the peak and trough times for road accidents during the day?
  - The peak occurrence of road accidents at 17:00 likely correlates with rush hour congestion, usually occurring at 8.00 AM and 5.00 PM. While, the off-peak period at 4:00 reflects reduced traffic volume, minimizing collision risk.
2. What does lighting condition play in the occurrence of accidents?
  - This visualization illustrates the frequency of accidents across the top five vehicle types under different light conditions. The highest number of accidents occurs during daylight conditions, implying that lighting conditions may not significantly influence the incidence of collisions.
3. How does the severity of accidents vary by road type?
  - After analyzing the data, we can conclude that the road type does not affect the severity of the accident. In every road type, the major portion of accident severity is slight severity.
4. What is the impact of weather conditions on road accidents?
  - The weather that has the highest number of accidents is fine no high winds which is the clearest visibility compared to the rest ( for example fog or mist, rain, high winds). As a result, the weather conditions do not have an impact on road accidents.
5. How does the time of year affect the number of road accidents?
  - The graph shows a similar pattern for both two years which is the lowest amount of accidents is in December and the highest amount of accidents is in November.
6. What are the common types of collisions and their frequencies by day of the week?
  - The common type of collision is not a junction which the highest number of accidents is on Friday. This is followed by T or staggered junction which has the same day of the week in the highest amount of accidents.

7. How does the number of vehicles on the road correlate with the number of accidents?
- The number of accidents has a significant relation to the number of vehicles, the higher number of vehicles will increase the amount of car accidents.

In conclusion, as we focus on analyzing the data on car accidents in the United Kingdom, we can identify the impact factors as well as trends of the accidents. However, we have a limitation on the data which be collected only 2 years, the limitation could lead to the misinterpretation of patterns and forecasts of the analyzed data. In the next study, should gather more length of data to be advantageous and precisely analyze data.

## References

- <https://www.kaggle.com/datasets/nextmillionaire/car-accident-dataset>