Road Accident Analysis Report

Problem Statement:-

Introduction

In a world where safety on the roads is of paramount concern, the government has taken significant steps to ensure the well-being of its citizens. Over the years 2019 to 2022, the government has introduced a series of policies aimed at enhancing road safety. These policies span a spectrum of measures, including stricter speed limits, heightened law enforcement, public awareness campaigns, and substantial infrastructure improvements. The overarching goal of these initiatives is clear: to reduce road accidents and casualties, making our streets safer for all.

At the heart of this endeavor lies the need for informed and data-driven decision-making. The government recognizes that to understand the true impact of these policies, it must turn to data analysis and insights. In response to this call, a dedicated team of analysts has been entrusted with the task of delving deep into the wealth of road accident data spanning the aforementioned years.

As data analysts, our mission is twofold: to unearth trends, patterns, and insights from the data, and to quantify the reduction in road accidents achieved through the diligent implementation of these policies. We recognize the profound significance of this analysis. The implications extend far beyond numbers and statistics; they resonate in the lives and safety of our citizens.

Our manager has conveyed the urgency of this task, underscoring the government's commitment to data-driven decision-making. It is our collective responsibility to harness the power of data, extract meaningful insights, and provide actionable recommendations that will continue to shape the future of road safety.

Data Source:-

The data source for this analysis project consists of a comprehensive dataset spanning the years 2019 to 2022. The dataset has been provided by the government, and it contains a wealth of information regarding road accidents and associated factors. The data has been meticulously collected from various government agencies, law enforcement reports, and hospital records, ensuring its accuracy and reliability.

Key Data Points:

- Accident date and time.
- Location and road conditions.
- Vehicle types and specifications.
- Casualty details, including fatalities, slight injuries, and serious injuries.
- Weather conditions at the time of accidents.

Data Integrity:

Ensuring the integrity of the data is paramount. The government has taken measures to verify and validate the data to minimize errors and discrepancies.

Data Privacy and Security:

The dataset complies with all relevant data privacy and security regulations. Personal information has been anonymized and encrypted to protect individuals' identities and privacy.

Data Accessibility:

The dataset is accessible through our secured servers, ensuring that only authorized personnel can access and analyze the data.

Data Limitations:
While every effort has been made to collect accurate and comprehensive data, it's essential to acknowledge potential limitations. These limitations may include data gaps, reporting discrepancies, or variations in data collection methods over the four-year period.

Data Link:- accident data.csv

Project Objective:

The primary objective of this project is to assess the impact of government policies on road safety by quantifying the reduction in road accidents, casualties, and fatalities from 2019 to 2022.

After having a glance of data we come to conclusion that we need to address following questions:-

Questions to Address:

- What are the year-to-year changes in the total number of accidents?
- How did the implementation of stricter speed limits affect the number of accidents?
- Has enhanced law enforcement led to a decrease in accident rates?
- What is the trend in casualties over the years?

Dataset Overview:-

The dataset under examination provides a comprehensive snapshot of road safety incidents within the city of London. These records encompass a range of essential data fields that offer insights into various facets of road accidents specific to London. Below is a detailed summary of each column within the dataset:

Index: This column is likely utilized as a unique identifier for each entry in the dataset, simplifying data organization and retrieval.

Accident_Severity: A categorical variable that classifies the severity of each accident into one of three categories: fatal, serious, or slight. It plays a pivotal role in gauging the seriousness of each incident.

Accident Date: The date on which each accident took place. This variable is vital for uncovering temporal trends, including monthly patterns and long-term shifts in accident frequencies, particularly within the unique traffic dynamics of London.

Latitude and Longitude: These geographical coordinates precisely pinpoint the location of each accident within London. They serve as the foundation for spatial

analysis, enabling the identification of accident-prone zones, the mapping of accident hotspots, and the assessment of geographic factors contributing to accidents.

Light_Conditions:A categorical variable describing the lighting conditions at the time of each accident. It is indispensable for understanding how visibility, which can be influenced by factors like street lighting and weather, affects accident rates.

District Area: Provides information about the specific district or borough within London where each accident occurred. It can be pivotal in identifying regions with higher accident frequencies, which may require targeted safety interventions.

Number_of_Casualties: A numerical variable representing the total number of casualties (injuries and fatalities) resulting from each accident. This metric quantifies the impact of each incident on public safety within the London context.

Number_of_Vehicles: The count of vehicles involved in each accident. This variable is valuable for assessing the complexity of accidents, including multi-vehicle incidents, which are common on London's busy roads.

Road_Surface_Conditions: Describes the condition of the road surface at the accident location. It plays a crucial role in addressing factors such as rain, snow, or ice, which can affect road safety, especially in a city with variable weather conditions like London.

Road_Type: Categorizes the type of road where each accident occurred. Whether it's a major highway, urban street, or rural road, this variable helps assess the characteristics of different road types contributing to accidents within London's urban landscape.

Urban_or_Rural_Area: Distinguishes between accidents that occurred in urban and rural areas within London. This classification assists in tailoring safety measures and policies to specific urban and suburban environments.

Weather_Conditions: Records the prevailing weather conditions at the time of each accident. This field is vital for understanding how adverse weather conditions, common in London, contribute to accidents.

Vehicle_Type: Categorizes the type of vehicle involved in each accident. Analyzing accidents by vehicle type can reveal distinct patterns related to specific vehicle categories prevalent in London, such as cars, buses, taxis, bicycles, and more.

Given the unique characteristics of London's road network, traffic density, and diverse population, this dataset represents a valuable resource for evaluating road safety in the city. It offers the opportunity to uncover specific trends, patterns, and factors contributing to accidents within the context of London, thereby informing data-driven strategies and policies aimed at enhancing road safety and reducing casualties in this urban environment. Subsequent sections of this report will delve into an in-depth analysis of this dataset to derive meaningful insights and recommendations tailored to London's road safety challenges.

Data Cleaning & Preprocessing:- Data is provided to us through government agencies , so it not contain much discrepancies, but still we do some data cleaning using excel tool. Now we have our cleaned data.

BrainStorming/Roadmap:-

We have following fields in Dataset provided:-

Fields

Type	Field Name	Physical Table	Remote Field Name
Abc	Accident Severity	accident data.csv	Accident_Severity
	Accident Date	accident data.csv	Accident Date
	Latitude	accident data.csv	Latitude
Abc	Light Conditions	accident data.csv	Light_Conditions
Abc	District Area	accident data.csv	District Area
	Longitude	accident data.csv	Longitude
#	Number of Casualties	accident data.csv	Number_of_Casualties
#	Number of Vehicles	accident data.csv	Number_of_Vehicles
Abc	Road Surface Conditions	accident data.csv	Road_Surface_Conditions
Abc	Road Type	accident data.csv	Road_Type
Abc	Urban or Rural Area	accident data.csv	Urban_or_Rural_Area
Abc	Weather Conditions	accident data.csv	Weather_Conditions
Abc	Vehicle Type	accident data.csv	Vehicle_Type

Field Name

We make following inferences after doing some brainstorming on data fields.

Columns	Uses	Conversion/Application/Brainstorming
Accident Severity	Can be used as Parameter for different KPI results and charts.	Used as Parameter to give overall view on values according to severity
Accident Date	Can be used to calculate monthwise and year wise calculations	Converted in Parameters ,**Current Year** and **Previous Year** to provide better user experience.
Latitude/Longitude	Can be used to make maps	Make geographical map using suitable filters.
Number of Casualities	Can be used to calculate total number of Casualtities	Make KPI as it is number and indicator directly.
Number of Vehicles	Can be used to calculate no. of vehicles accordance with vehicle type	Make KPI as direct indicator, and we have different vehicle type also
Road Surface Condition	Can be used to calculate accident based on road surface condition	As here not direct involvement in imapct of Policies so we can make in chart form,Pie chart will be suitable for it.
Road Type	Used to calculate accident on road type condition	As here direct involvement but not to give exact figures to disclosing so we can make bar chart due to different categories.
Weather Condition	Used to calculate accident due to different weather condition	As here not direct involvement to disclosing figures so we can make it in chart form, pie chart will be useful
Vehicle Type	Can be used to calculate accident according to vehicle type	Involve in making KPI as we have number of vehicle involved,and have direct relation to disclosing figures.

Addition of Calculation field according to requirements:-

Fields

Type	Field Name	Physical Table	Remote Field Name
#=	PY Accidents	Calculation	CY Accidents (copy)_1412722...
#=	CY Fatal Casualities	Calculation	CY Accidents (copy)_1412722...
#=	PY Casualities	Calculation	CY Casualities (copy)_141272...
#=	CY Serious Casualities	Calculation	CY Fatal Casualities (copy)_14...
#=	CY Slight Casualities	Calculation	CY Serious Casualities (copy)...
#=	Year of Accident Date	Calculation	Calculation_14127229669496...
#=	CY Accidents	Calculation	Calculation_14127229669672...
#=	YoY Accidents	Calculation	Calculation_14127229669694...
#=	CY Casualities	Calculation	Calculation_14127229669739...
=T|F	Accident Severity filter	Calculation	Calculation_141272296710105...
#=	PY Fatal Casualities	Calculation	PY Accidents (copy)_1412722...
#=	PY Serious Casualities	Calculation	PY Fatal Casualities (copy)_14...
#=	PY Slight Casualities	Calculation	PY Serious Casualities (copy)...

#	YoY Casualities	Calculation	YoY Accidents (copy)_1412722...
#	YoY SeriousCasualities	Calculation	YoY Accidents (copy)_1412722...
#	YoY Slight Casualities	Calculation	YoY SeriousCasualities (copy)...
📎	Road Surface Conditions (gr...	Group	Road Surface Conditions (gro...
📎	Vehicle Type (group)	Group	Vehicle Type (group)
📎	Weather Conditions (group)	Group	Weather Conditions (group)

Addition/Grouping of Fields as per requirement

Data Comprehension:-

1. Key Performance Indicators (KPIs)

We've taken a good look at the data and figured out what we really need to understand road safety. So, we've come up with some Key Performance Indicators (KPIs) that are like our cheat sheet to get the big picture.

Total Accidents:-

It tells us about total no. of accidents in a current year.

As no direct column so we need to make calculation field.

Formula Used:-

Current Year *(CY Accidents)*:-

count(IF Year([Accident Date]) = [Current Year] then [Index] END)

Previous Year*(PY Accidents)*:-

count(IF Year([Accident Date]) = [Previous Year] then [Index] END)

Year on Year decrement/increment *(YOY Accidents)*:-

([CY Accidents]-[PY Accidents])/[PY Accidents])

Total Accident KPI

Total Casualties:-

Total Casualties: This one's about adding up all the people who got hurt or worse, in all the accidents. It helps us see how bad things really are.

As no direct column there so we need to make calculation field.

Formula Used:-

Current Year*(CY Casualities):-*

sum(if year([Accident Date]) = [Current Year] then [Number of Casualties] END)

Previous Year*(PY Casualities):-*

sum(if year([Accident Date]) = [Previous Year] then [Number of Casualties] END)

Year on Year*(YoY Casualities):-*

([CY Casualities]-[PY Casualities])/[PY Casualities])

Total Casualities KPI

Fatal Casualties:-

Fatal Casualties: We're focusing on the accidents where people sadly lost their lives. It's a super important part of the safety puzzle.

As no direct column there so we need to make calculation field.

Formula Used:-

Current Year*(CY Fatal Causalities)*:-

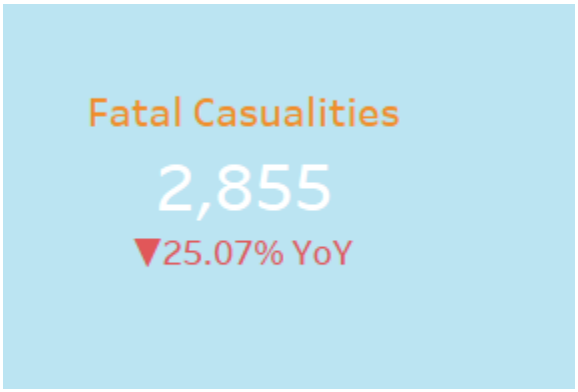
sum (if [Accident Severity] = 'Fatal' and Year([Accident Date]) = [Current Year] then [Number of Casualties] end)

Previous Year*(PY Fatal Casualities)*:-

sum (if [Accident Severity] = 'Fatal' and Year([Accident Date]) = [Previous Year] then [Number of Casualties] end)

Year on Year*(YoY Casualities)*:-

([CY Fatal Casualities]-[PY Fatal Casualities])/[PY Fatal Casualities]



Fatal Casualiites KPI

Slight Casualties

Slight Casualties:- Not all accidents are super serious. This KPI helps us spot accidents where folks got minor injuries.

As again there is no direct column so we need to make calculation field.

Current Year(*CY Slight Casualities)*:-

sum (if [Accident Severity] = 'Slight' and Year([Accident Date]) = [Current Year] then [Number of Casualties] end)

Previous Year*(PY Slight Casualities)*:-

sum (if [Accident Severity] = 'Slight' and Year([Accident Date]) = [PreviousYear] then [Number of Casualties] end)

Year on Year*(YoY Slight Casualities)*:-

([CY Slight Casualities]-[PY Slight Casualities])/[PY Slight Casualities])

Slight Casualities KPI

Serious Casualties

Serious Casualties: Now, we're looking at the accidents where people got really hurt. This helps us zoom in on the serious stuff.

As again no direct column so we need to make calculation field.

Current Year(*CY Serious Casualities)*:-

sum (if [Accident Severity] = 'Serious' and Year([Accident Date]) = [Current Year] then [Number of Casualties] end)

Previous Year*(PY Serious Casualities)*:-

sum (if [Accident Severity] = 'Serious' and Year([Accident Date]) = [PreviousYear] then [Number of Casualties] end)

Year on Year*(YoY Serious Casualities)*:-

([CY Serious Casualities]-[PY Serious Casualities])/[PY Serious Casualities])

Serious Casualities KPI

Monthly Trends/SparkLine:-

Monthly Trends:- We're making a graph to show how accidents change over each month. It's like looking at a pattern to see if accidents are going up, down, or staying about the same. We use it to comparison with previous year and current year accidents.

Color Decode

Total Accident KPI with Sparkline

Total Casualities with Sparkline

Fatal Casualities with Sparkline

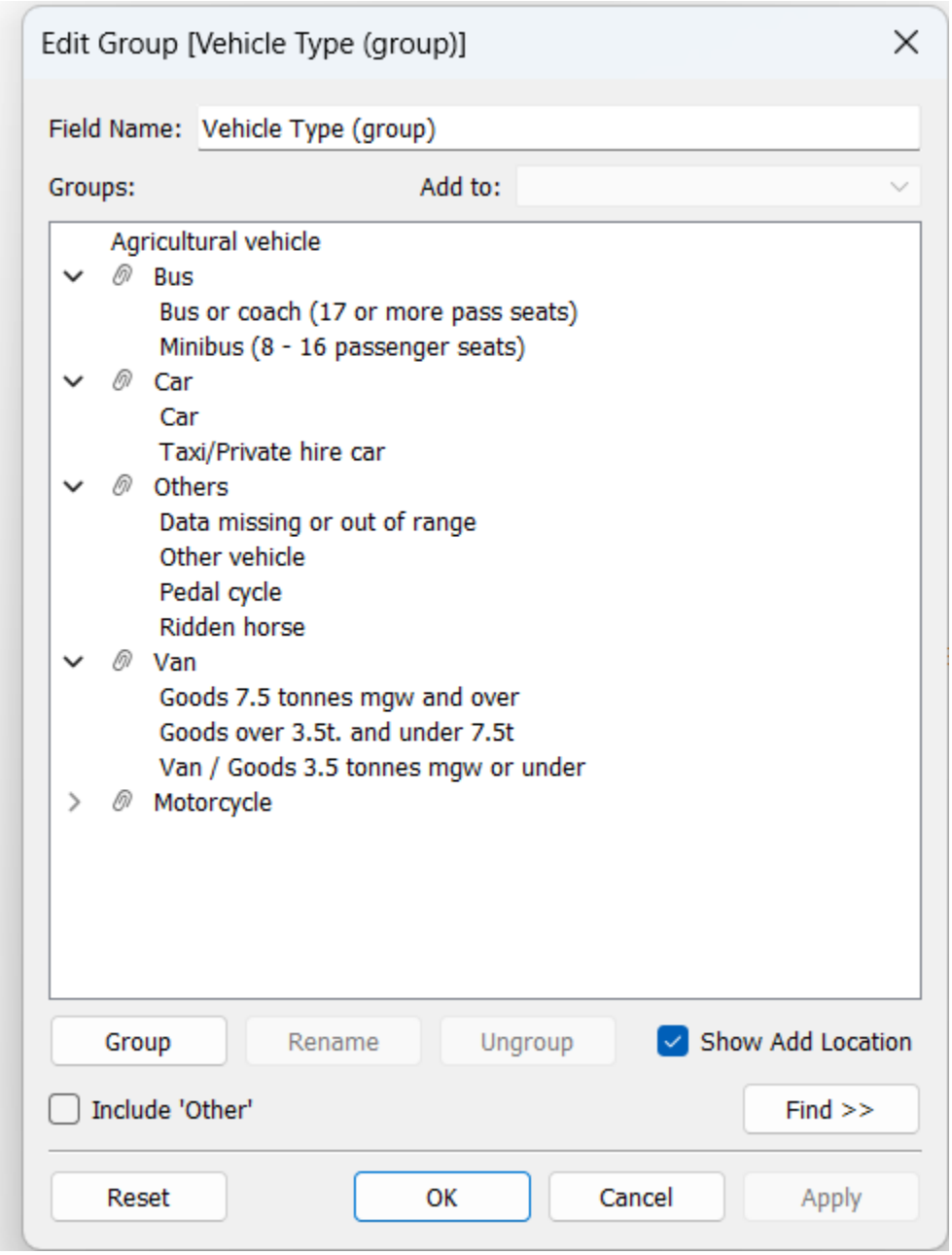
Serious Casualities with SparkLine

Slight Casualities with SparkLine

2. Vehicle Type Analysis

Vehicle Type Analysis: Instead of lumping all accidents together, we're breaking them down by the types of vehicles involved. This helps us understand the unique challenges of each type.

We get not grouped column but on glancing vehicle type column there are entities which may be fall in same category so we grouped them.



Grouping Vehicle Type

Agriculture Vehicle:- We're checking accidents with agricultural vehicles. They have their own set of issues on the road.

Bus:- Accidents with buses matter, especially for public transportation.

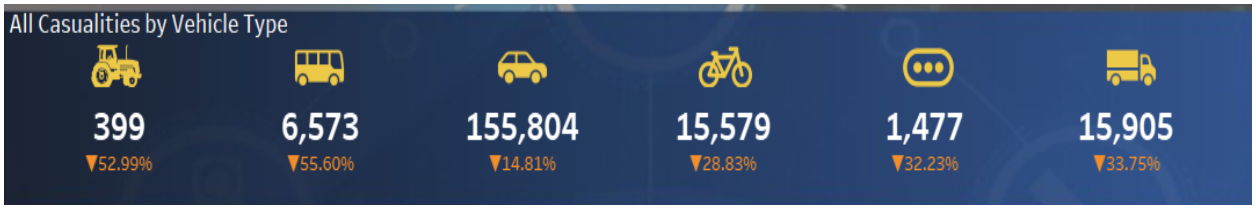
Car: Since most people drive cars, we're focusing on accidents involving them.

MotorCycle: We're looking at accidents with bicycles because they have their own safety concerns.

Others: Some accidents involve different types of vehicles not mentioned above. We're keeping an eye on those too.

Van:- Accidents with big trucks, especially the heavy ones, are a special concern.

These KPIs are like our roadmap to understand road safety better. By using them, we can make smarter decisions, figure out what needs fixing, and work on making our roads safer for everyone.



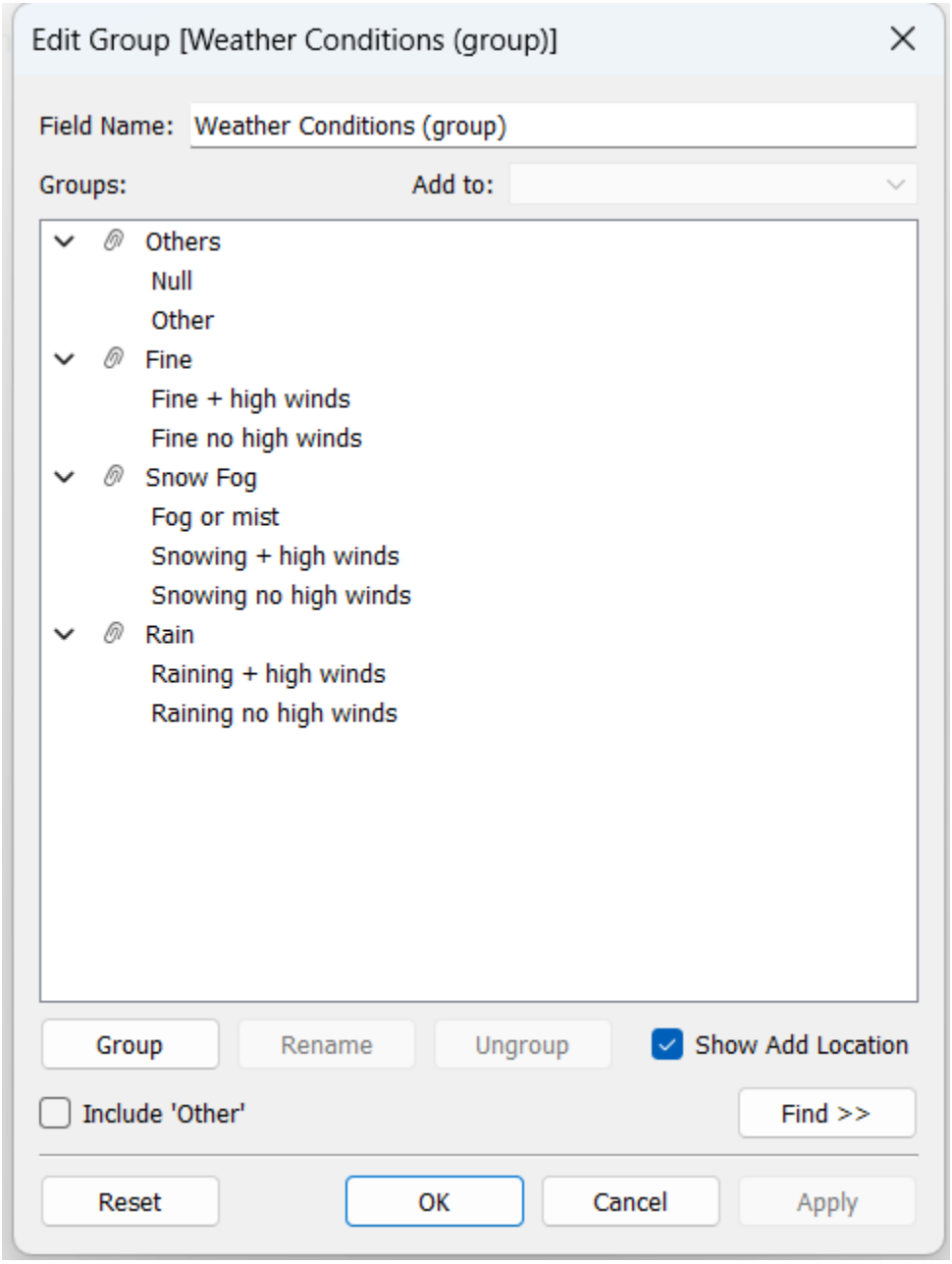
All Casualities by Vehicle Type

Visualizations(Charts/Maps)

Weather induced Casualities Analysis

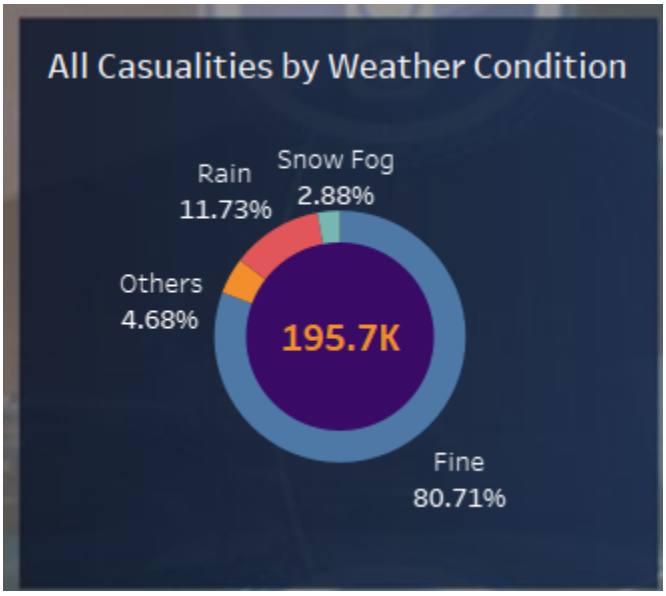
This is important because different weather conditions like snow, fog, rain, or fine weather can impact accidents. We're going to analyze how weather affects accidents and casualties to make our roads safer, rain or shine.

As we get column of different similar categories so we grouped them according to main weather condition.



Grouped Weather Conditions

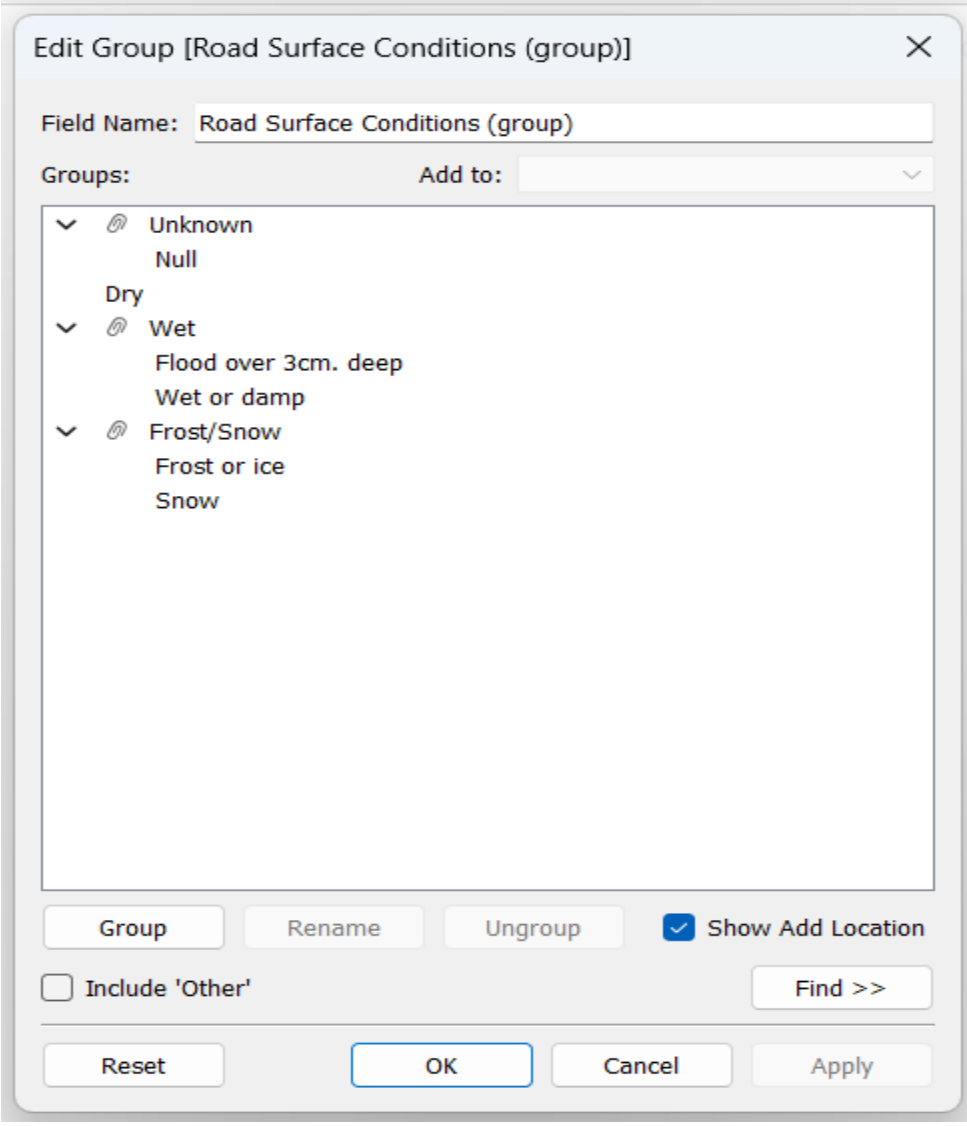
We make pie chart for this grouped weather condition. As pie chart is not directly available in Tableau so we use two supporting axis of average(0) and using dual axis concept we make chart.



Casualties due to Weather Condition

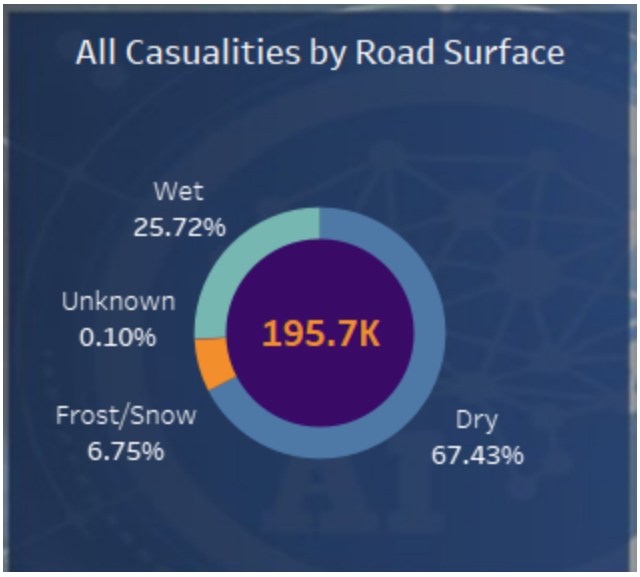
Road Surface Casualities Analysis

The road surface matters too. We're going to look at accidents on unknown surfaces, dry roads, wet roads, and frost/snow-covered roads. This helps us understand how road conditions can lead to accidents and casualties, ensuring our roads are safer in all conditions.



Grouped Road Surfaces

The road surface matters too. We're going to look at accidents on unknown surfaces, dry roads, wet roads, and frost/snow-covered roads. This helps us understand how road conditions can lead to accidents and casualties, ensuring our roads are safer in all conditions.



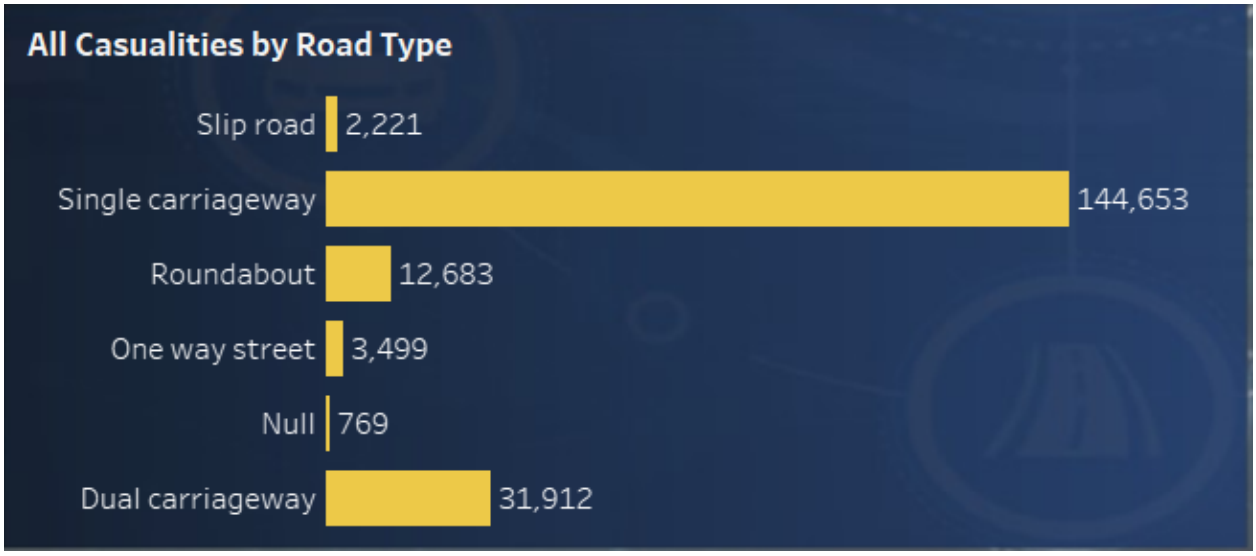
Casualities by Road Surface

Road Type Casualities Analysis

Roads come in different types like sliproads, single carriage ways, roundabouts, one-way streets, dual carriageways, and sometimes, we don't know what type they are. We're going to analyze accidents on each type to see how road design can impact accidents and casualties, making our roads safer no matter where we drive.

As for Road type casualties analysis we have already column of Road type so no need to group anymore.

For Road type Casualities Analysis we make **Horizontal chart.**



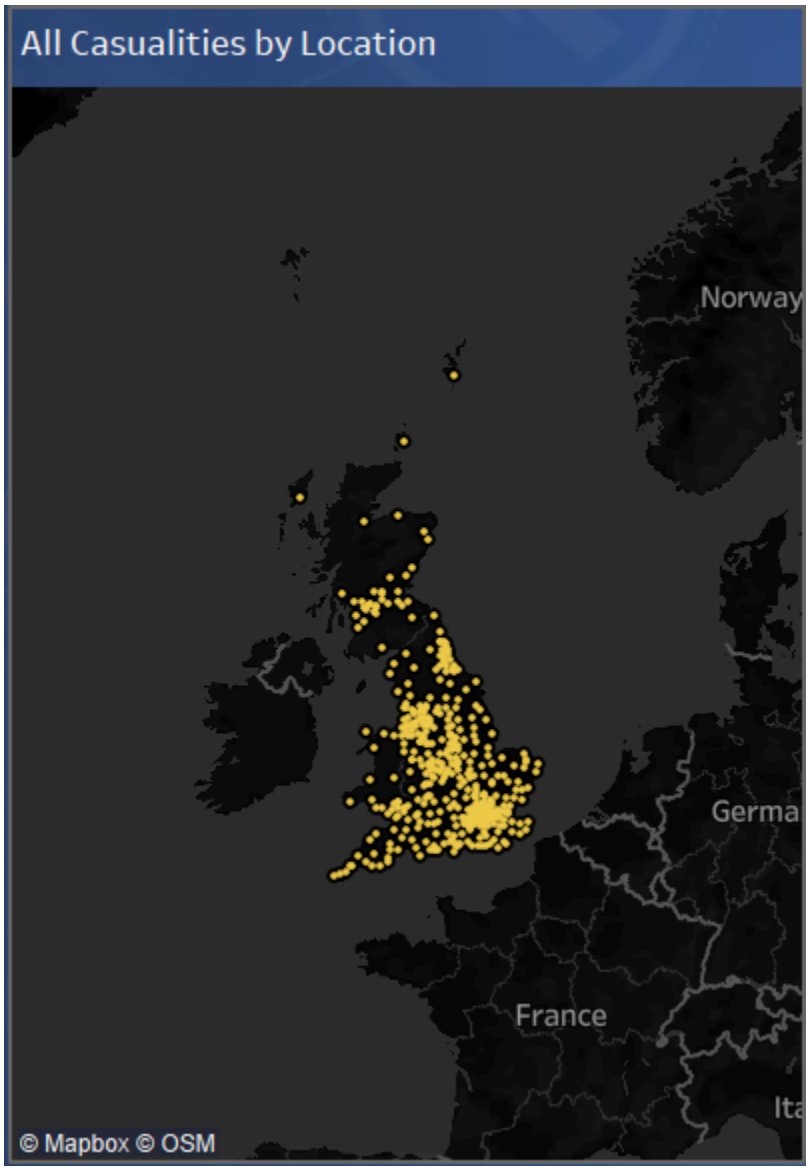
Casualties by Road Type

Maps

Casualities By location

We're also looking at where accidents happen. By understanding the specific locations where casualties occur, we can better focus our safety efforts and resources to make those areas safer for everyone.

As we have columns **Latitude and Longitude** this enables us to make geographic map.



Casualties By Location

Dashboard Overview:-

Dashboard

Link:-https://public.tableau.com/views/RoadAccidentAnalysis_16941874327320/Dashboard12?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

KPI Section:- This section directly contain KPI's which directly deal with figures, having current year accidents, casualities , fatal casualities, serious casualities,slight casualties.Along with that having **Year on Year increment/decrement** as per requirement and main variable to determine.

Along with KPIs directly relating to vehicle type and casulaities is also there.

Charts Section:- This section contain charts such as two **Pie charts** dealing with casualities due to weather conditions and casualties due to Road Surface.

Along with these charts there is also **horizontal bar chart** dealing with casualities by road type.

Maps Section:- This section contain **maps** directly deals with location .

Conclusion

As government frame different policies on road safety and to avoid road accidents.These policies span a spectrum of measures, including stricter speed limits, heightened law enforcement, public awareness campaigns, and substantial infrastructure improvements. The overarching goal of these initiatives is clear: to reduce road accidents and casualties, making our streets safer for all.

Implementing these policies government has achieved significant reduction in accidents.

Inferences:-

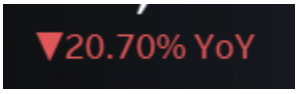
- What are the year-to-year changes in the total number of accidents
  Ans:- 20.70%

- How did the implementation of stricter speed limits affect the number of accidents?
  Ans:- It decreases the number of accidents due to various heavy loaded vehicles

- Has enhanced law enforcement led to a decrease in accident rates?
  Ans:- Yes

- What is the trend in casualties over the years?
  Ans:- We got **decreasing** trend over the years on casualties.

Result:- Achieved Reduction of **20.70%** from year 2019 to 2022.



****