ROAD ACCIDENT ANALYSIS – UK (2021)

An Analytical Study Using Microsoft Power BI

ABSTRACT

Road accidents pose a serious threat to public safety and remain a major challenge for transport authorities worldwide. This project aims to conduct a comprehensive analysis of road accident data from the United Kingdom for the year 2021 using Microsoft Power BI. With over 307,000 records, the dataset includes critical attributes such as accident severity, weather and lighting conditions, road types, junction control, speed limits, vehicle categories and more.

The primary objective of the analysis is to identify key contributing factors behind accidents and uncover trends based on time, location and environmental conditions. Using Power Query and DAX, the data was cleaned, transformed and modeled to support insightful, interactive dashboard visualizations.

Key findings from the analysis indicate that most accidents occurred on single carriageways, particularly in areas with uncontrolled junctions and inadequate lighting — especially in rural zones. Afternoon hours and Fridays showed the highest accident frequencies, with cars being the most involved vehicle type. Interestingly, most accidents happened in dry and fine weather, suggesting that poor weather alone is not the main contributor. a significant number of accidents occurred in 30–40 mph speed zones, indicating that moderate-speed roads, often found in urban settings, require enhanced safety measures such as better signage, crossings and traffic calming.

The insights derived from this project support targeted recommendations, including the installation of traffic control mechanisms at high-risk junctions, improvements in rural lighting infrastructure and driver awareness programs focused on specific road and time conditions. The interactive dashboard built as part of this project allows stakeholders to explore accident patterns and take data-driven decisions to enhance road safety.

INTRODUCTION

Road accidents are a critical issue affecting public safety and mobility, often resulting in significant human and economic loss. With the rapid growth of transportation networks, understanding the root causes behind road accidents has become more essential than ever for ensuring safer infrastructure and informed policy-making.

This project presents an in-depth analysis of road accident data from the United Kingdom for the year **2021**, leveraging the power of **Microsoft Power BI** for visualization and insight generation. The dataset comprises over **307,000 accident records**, covering various parameters such as accident severity, weather and lighting conditions, road type, speed limits, vehicle involvement and geographical factors.

The primary objective was to explore, clean and transform the data, then model and visualize it to uncover patterns and derive actionable conclusions. Using tools like Power Query and DAX, the data was structured to enable advanced filtering and analysis across different dimensions—time, location, environment and behavior.

1. PROJECT OBJECTIVE

The primary objective of this project is to conduct a thorough analysis of UK road accident data from the year 2021 using Power BI. The aim is to:

- Identify and quantify major contributors to road accidents (e.g., road type, vehicle type, lighting conditions).
- Highlight patterns based on time, location and weather conditions.
- Provide visual, evidence-based insights that can aid policymakers and safety officials.
- Recommend potential interventions to improve road safety and reduce accident severity.

This analysis ultimately helps in driving informed decisions by utilizing visual analytics and interactive dashboards.

2. TOOLS USED

Power BI Desktop – for visualization and data modeling

Microsoft Excel – for preliminary data checks and structure

DAX (Data Analysis Expressions) – for creating calculated measures

Power Query Editor – for cleaning and transforming data

3. DATASET OVERVIEW

The dataset includes over 307,000 records with various attributes like:

Accident Index, Time, Accident Severity

Weather Conditions, Light Conditions

Road Surface, Speed Limit, Junction Type

Vehicle Type, Urban/Rural classification

Data was restricted to the year 2021 to maintain temporal consistency and clarity in trend analysis.

4. DATA UNDERSTANDING & PREPARATION

This phase combined **Exploratory Data Analysis (EDA)** and **Data Cleaning** to assess data quality, identify key patterns and prepare the dataset for accurate modeling and visualization.

Data Overview

- **Records**: ~307,000 road accident cases (UK, 2021)
- **Key Identifier**: Accident Index (unique for each entry)

Completeness & Missing Data Handling

Column	Action Taken
Carriageway Hazards	Dropped – over 99% were "None"; minimal analytical value
Junction Control	Missing values labeled as "Insufficient Data"
Light Conditions	Sparse missing values handled similarly
Other Critical Fields	Rows with null Date, Time, or Severity were removed

Redundant Columns Excluded

Police Force was excluded since its role was overlapping with Local District Authority and didn't add extra analytical insight.

Categorical Distribution Highlights

Feature	Dominant Values
Accident Severity	Slight (≈70%), Serious (≈25%), Fatal (≈5%)
Road Type	Single Carriageway (≈75%)
Light Conditions	Daylight (≈67%), Darkness w/ lighting
Weather Conditions	Fine (≈80%), Rain (~10%)
Vehicle Type	Cars (≈60%), LGVs, Motorcycles

Standardization & Label Cleaning

Replaced inconsistent or misspelled entries:

- *"Auto traffic sigl"* → *"Traffic Signal"*
- *"Give way or uncontrolled"* → *"Uncontrolled"*
- *"Frost or ice"* → *"Ice/Frost"*
- Lighting and junction terms were cleaned for better grouping

Reformatted column names (e.g., Road Type → Road Type) for readability

Data Transformations (Power Query)

- Removed duplicates using Accident Index
- Filtered invalid or irrelevant rows
- Applied formatting for Date, Time
- Created new fields:
  - Month Name (from Date)
  - Hour (from Time)
  - Time Band (Morning, Afternoon, Evening, Night)

Outlier Detection

During the exploratory analysis, several **low-frequency outliers** were observed:

- **Non-Motorized Vehicles** were involved in only **69 accidents**, a sharp contrast to over **239,000** for cars. This reflects either **limited road exposure** or **underreporting**.

- **Manual Control** junctions and **Stop Signs** showed **very few accidents** relative to other junction types, suggesting either **limited usage** or **effective control mechanisms**.
- **Snow-related weather conditions** had very low accident counts, likely due to their **infrequent occurrence** in the dataset's geographic coverage.

Feature Engineering

New Column	Description
Month Name	Extracted from Accident Date
Hour	Extracted from Time
Time Band	Grouped Hour field (Morning, Afternoon, etc.)
Day of Week	Derived from Date
Severity Index	Weighted numeric scoring (e.g., Fatal = 3)

This combined EDA and cleaning process revealed:

- High-risk segments (Single Carriageways, Uncontrolled Junctions)
- Time hotspots (Afternoon, Fridays)
- Environmental patterns (dry weather, poor lighting in rural zones)
- Sparse or inconsistent data fields that were cleaned or grouped

These insights informed the dashboard design and helped ensure reliable, insightful analysis.

5. DASHBOARD DEVELOPMENT IN POWER BI

Page 1: Overview Dashboard

The first page of the Power BI dashboard presents a comprehensive summary of road accident metrics, enabling a quick understanding of key trends and distributions. The visuals included are:

Summary Metrics (Cards):
Displays KPIs such as total accidents, fatal accidents and vehicle involvement.

Bar Charts:

- **Accidents by Time Band** – Compares accident frequency during Morning, Afternoon, Evening and Night.
- **Accidents by Day of Week** – Highlights which weekdays saw the highest number of incidents.
- **Accidents by Month** – Shows monthly distribution and seasonality.
- **Accident Severity** – Categorizes the severity (Slight, Serious, Fatal) for comparative analysis.

Pie Chart:

Accidents by Weather Condition – Visualizes how different weather types (e.g., Fine, Rain, Fog) impact accident counts.

Donut Chart:

Urban vs Rural Accidents – Compares accident occurrences between urban and rural areas, providing insight into geographic risk.

This page offers a consolidated view for stakeholders to quickly assess the general patterns of road accidents across time, severity, location and environmental conditions.

Page 2: Cause Analysis Dashboard

The second page of the dashboard dives deeper into the contributing factors behind accidents, focusing on conditions, severity and vehicle involvement. It includes the following visuals:

Bar Chart:

Total Accidents by Vehicle Type and Severity – Shows which vehicle categories (e.g., Car, Motorcycle, Goods Vehicle) are most involved and the severity distribution within each type.

Stacked Column Chart:

Total Accidents by Urban/Rural and Light Conditions – Compares accident counts across urban and rural zones based on lighting (e.g., daylight, darkness with/without lighting).

Column Chart:

Severity of Accidents by Road Surface Conditions – Highlights how surface types (Dry, Wet, Ice/Frost) relate to accident severity.

Bar Chart:

Total Accidents by Speed Limit and Severity – Analyzes how different speed zones (e.g., 30, 40, 50 mph) affect accident severity.

Multi-row Card:

Displays key indicators including dominant weather condition during accidents, severity type most frequently reported and average number of casualties per accident.

This page aims to uncover root causes and patterns, supporting strategic planning for road safety interventions based on environmental and contextual factors.

Page 3: Insights & High-Risk Zones Dashboard

The third page presents deeper insights into the locations and structural conditions most associated with road accidents. It helps stakeholders identify where targeted interventions can be most effective. Key visual elements include:

Uncontrolled Junction Analysis:

- **Bar Chart and Pie Chart:** Show that the majority of accidents occurred at uncontrolled junctions, emphasizing the lack of traffic control measures.
- **KPI Card:** Displays the total number of accidents at uncontrolled junctions for quick reference.

Road Type Insight – Single Carriageways:

- **Bar Chart and Donut Chart:** Indicate that single carriageways account for the highest proportion of road accidents, likely due to two-way traffic with limited separation.
- **KPI Card:** Highlights the accident count specifically from single carriageways.

Junction Detail Severity Chart:

Bar Chart: Shows the distribution of accident severity (Slight, Serious, Fatal) across different junction types (e.g., Crossroads, T-junctions, Roundabouts).

Top 20 Districts by Road Accidents:

Bar Chart: Displays the districts with the highest accident counts to help identify regional hotspots. This focuses on the top 20 to ensure clarity, given the large number of district entries.

This page synthesizes the analytical findings from the dataset into actionable insights, emphasizing infrastructure elements (junctions and road types) that require attention.

6. VISUAL TYPES USED

Bar & Column Charts – For comparisons (e.g., accidents by road type, junction, speed)

Pie & Donut Charts – For proportional data (e.g., weather, urban vs. rural)

Multi-row Cards – To display multiple key metrics side-by-side

KPI Cards – For highlighting totals (e.g., total accidents, fatalities)

Slicers – For interactive filtering by fields like severity, time, weather

Filled Map – Visualizing top accident-prone districts

Matrix Table – For detailed cross-tab analysis

Ribbon Chart – To visualize changes in category rank over time or across groups

Line Charts – For analyzing trends over time (e.g., monthly accident patterns)

Treemap – For hierarchical data like accident distribution across districts or vehicle types

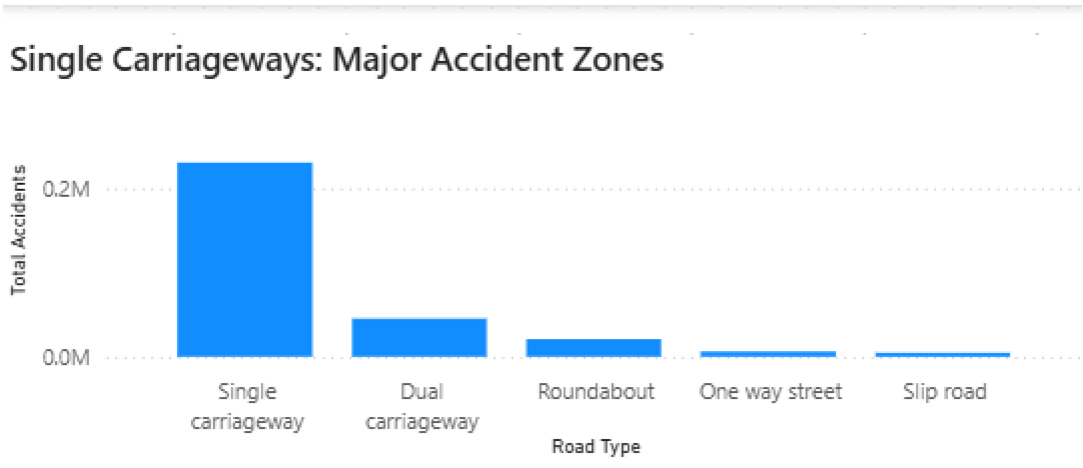
7. KEY INSIGHTS & INTERPRETATIONS

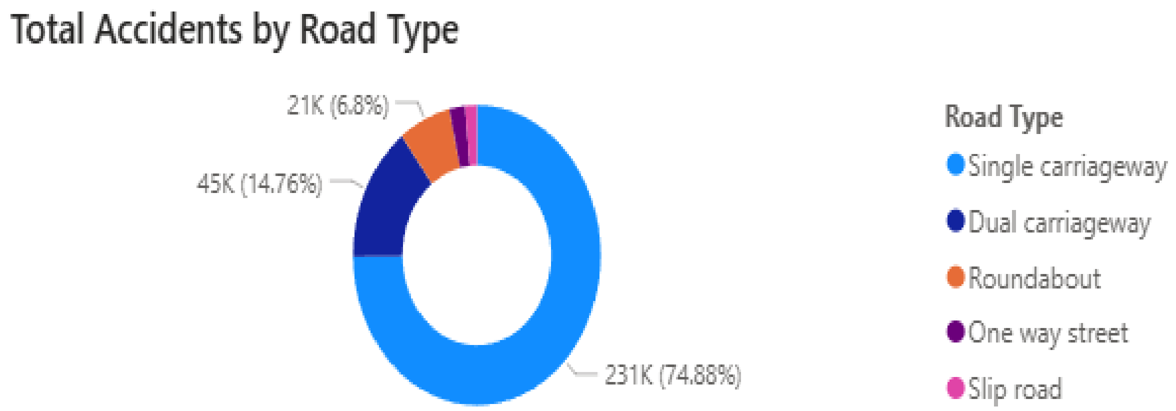
ROAD TYPE

Over **75% of total accidents** occurred on **Single Carriageways**, highlighting them as the most accident-prone road type.

Interpretation:

Single carriageways, which often allow two-way traffic without any physical separation, increase the risk of head-on collisions, overtaking errors and reduced reaction time — especially in high-traffic or poor visibility conditions.





Recommendation:

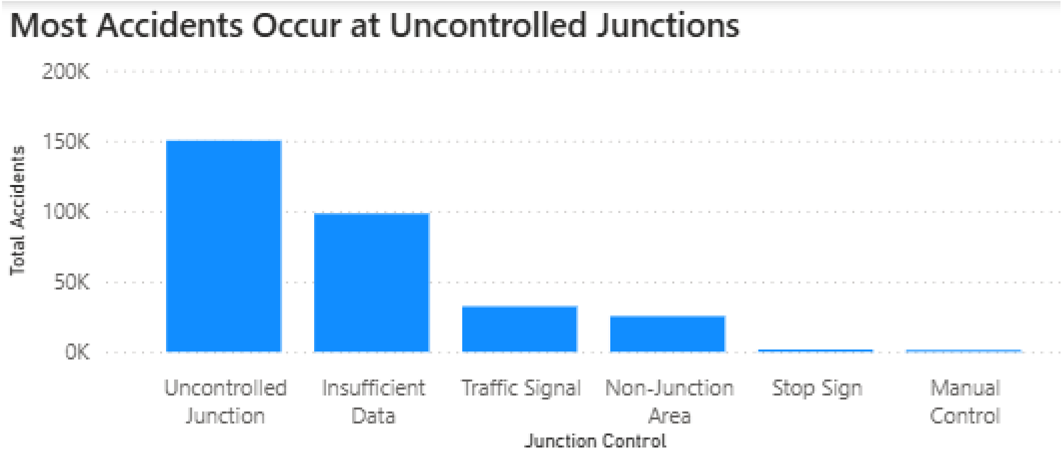
Road safety audits and possible **upgrades to dual carriageways or inclusion of medians** should be considered in high-risk zones.

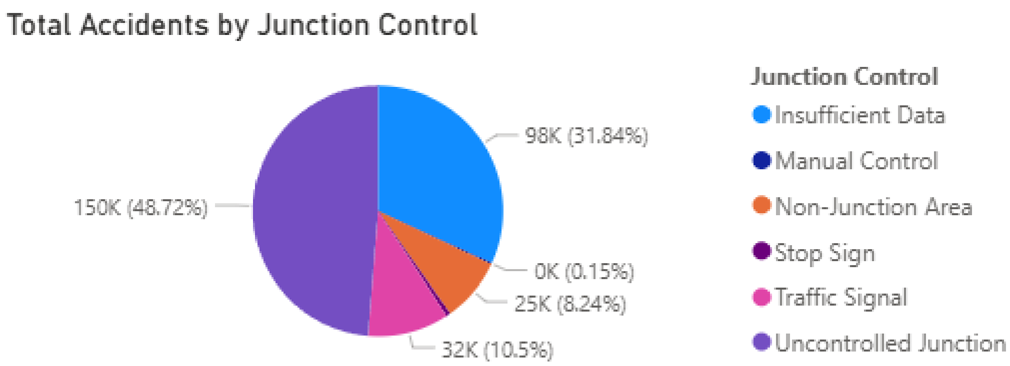
JUNCTION CONTROL

A significant proportion of accidents were observed at **uncontrolled junctions**, making them a critical contributor to road safety risks.

Interpretation:

Uncontrolled junctions lack regulatory mechanisms such as traffic signals, stop signs, or roundabouts, increasing the likelihood of driver confusion, misjudgement and right-of-way violations — especially during peak traffic hours or in low-visibility conditions. A significant portion of records under the Junction Control field were marked as *missing* or *unknown*. This suggests that proper documentation and reporting procedures are not being consistently followed, potentially due to lapses by field officers or inadequate data collection protocols.





Recommendation:

Authorities should prioritize the **installation of traffic control systems** (signals, signage, or roundabouts) at high-risk or high-traffic junctions to reduce collisions and enhance navigational clarity for road users.

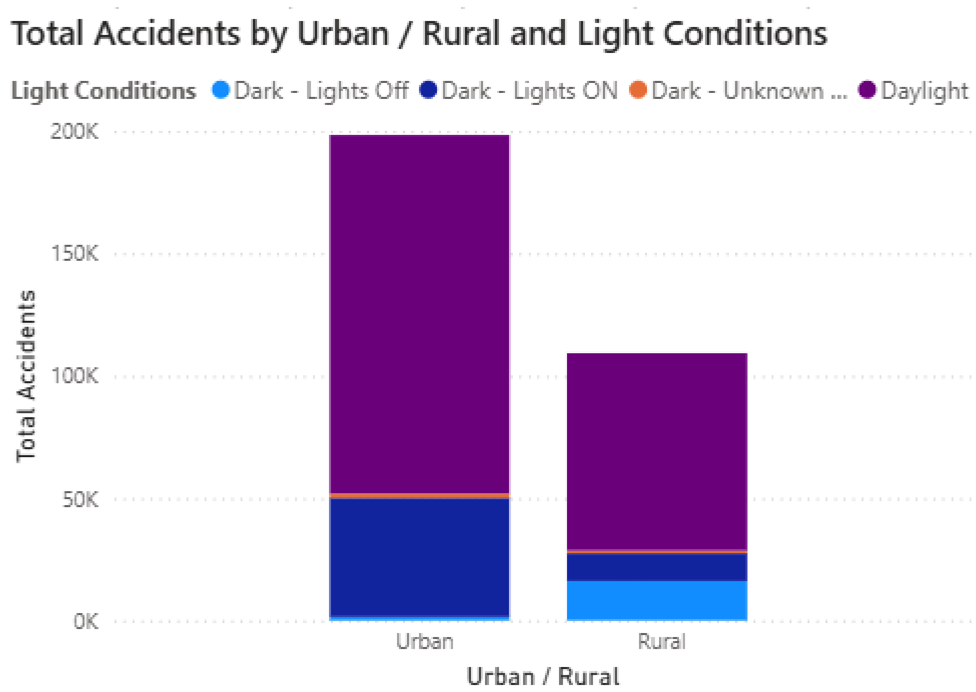
LIGHTING & LOCATION

Many **severe accidents** occurred on **dark roads without lighting**, especially in **rural areas**.

Interpretation:

Poor visibility at night, combined with limited or no street lighting, significantly reduces reaction time, increases the chances of missing road signs or obstacles and contributes to higher accident severity.

While urban areas recorded more total accidents, the majority were minor due to better infrastructure and lighting. In contrast, rural areas had fewer accidents, but a higher proportion were fatal or serious, likely due to inadequate lighting and emergency response limitations.



Recommendation:

Authorities should **upgrade lighting infrastructure in rural and poorly lit areas**, especially in high-risk junctions and stretches with recurring accidents. Strategic streetlight placement and visibility enhancements can reduce accident severity and save lives.

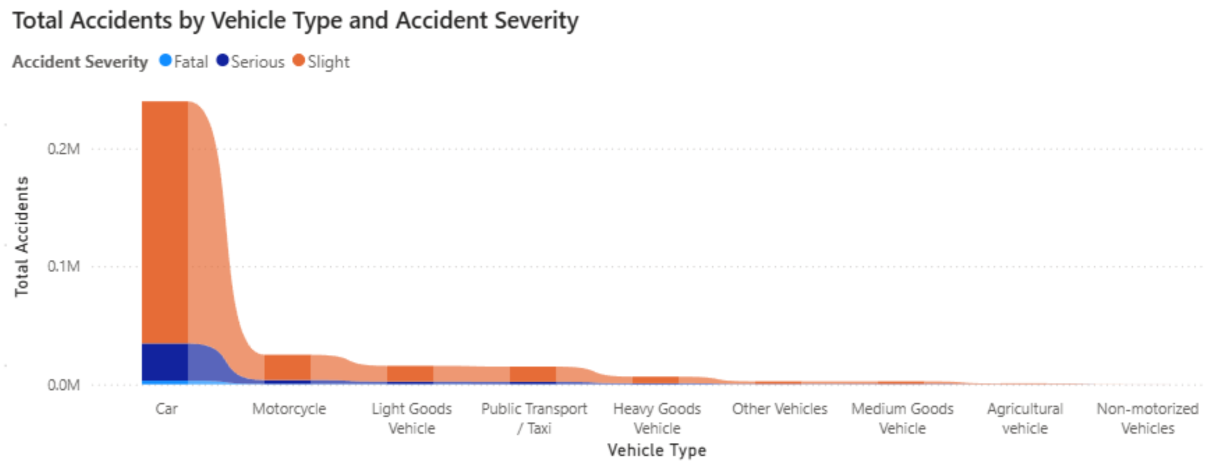
VEHICLE TYPE

Cars were the most involved vehicle type in road accidents across all regions.

Interpretation:

Cars accounted for the majority of total accidents, largely due to their high volume on roads. However, when analyzing accident severity, motorcycles and goods vehicles had a disproportionately higher percentage of serious and fatal outcomes.

This trend can be attributed to the vulnerability of motorcyclists, lack of protective structure and sometimes risky maneuvering in mixed traffic. Goods vehicles, due to their size and momentum, contribute to greater impact force during collisions.



Recommendation:

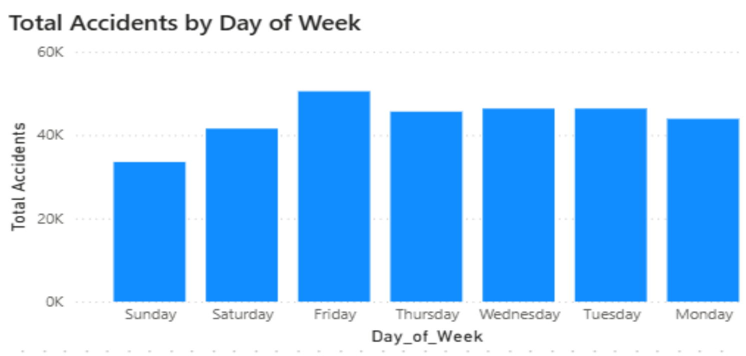
Awareness campaigns and enforcement of protective gear use should be intensified for motorcyclists, who show higher fatality rates. Additionally, driver training programs and speed monitoring systems for goods vehicle operators can help reduce severe outcomes. Urban planners should also consider implementing dedicated motorcycle lanes or vehicle-type-specific safety signage in high-risk areas to enhance overall road safety.

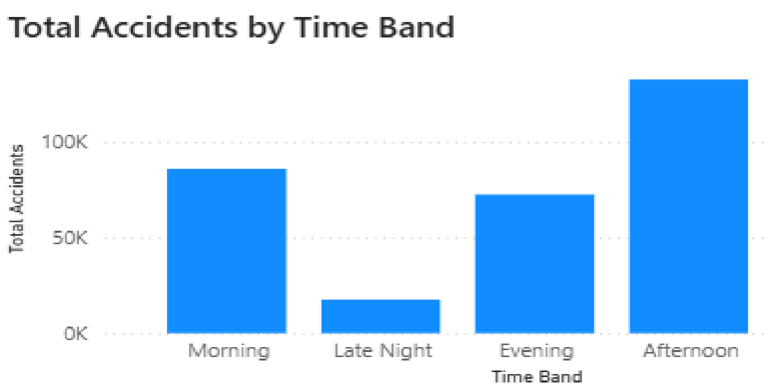
TIME-BASED RISK

Accident frequency peaked during **Afternoon hours**, with **Fridays** consistently showing the highest number of incidents.

Interpretation:

This trend likely reflects increased traffic congestion during post-lunch office hours and pre-weekend travel rush. Reduced alertness or distractions during this time may also contribute to accident occurrence.





Recommendation:

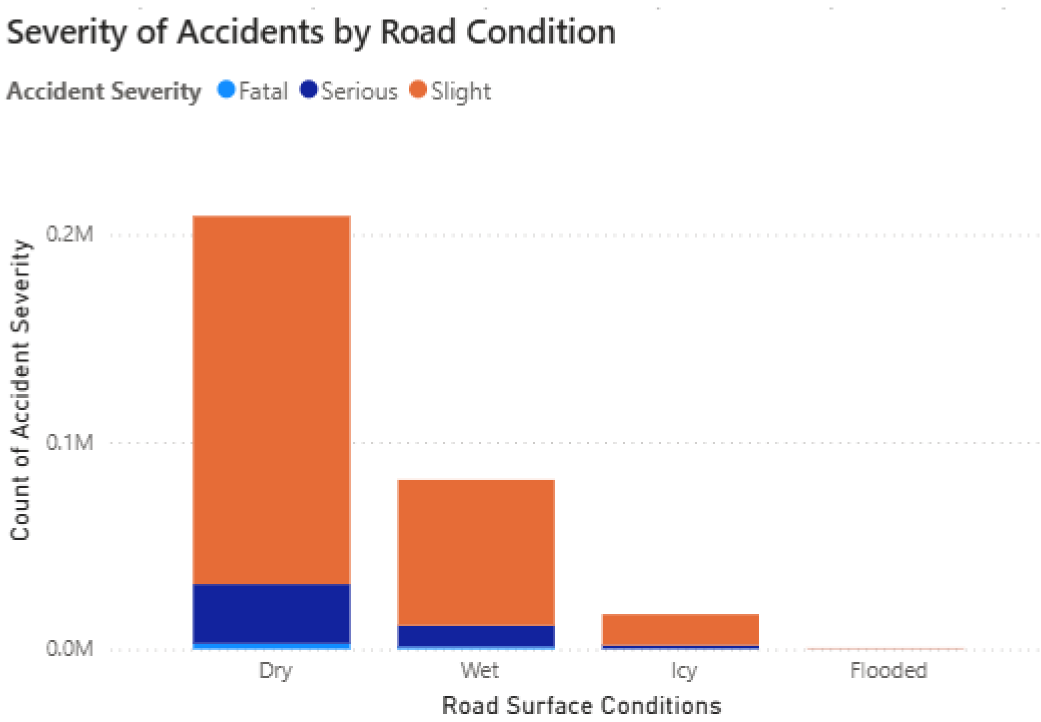
Implement **targeted enforcement** and **speed calming measures** during peak afternoon periods, especially on Fridays. Increase awareness campaigns focusing on alertness and responsible driving during busy time bands.

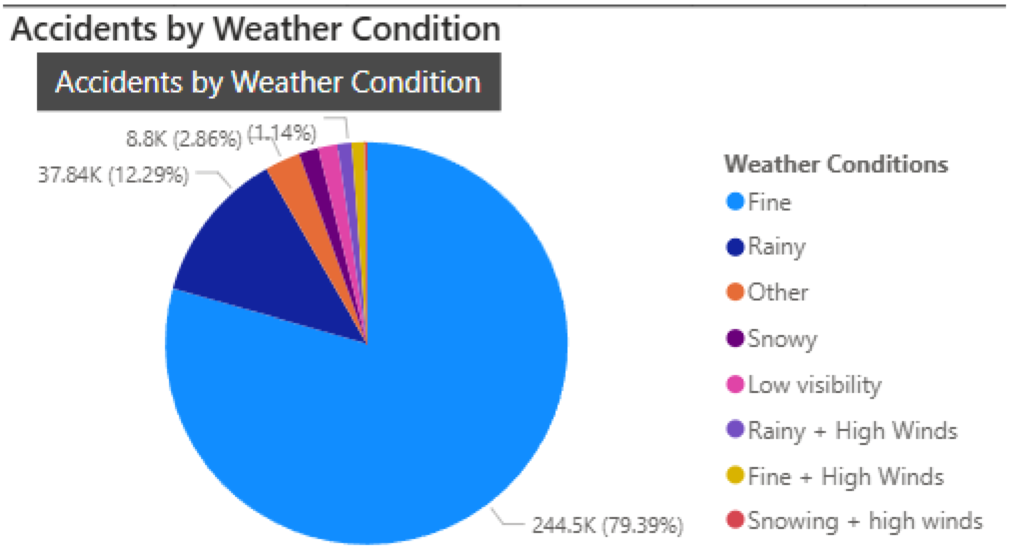
SURFACE & WEATHER

The majority of accidents occurred on **dry road surfaces** and during **clear/fine weather**.

Interpretation:

Contrary to expectation, adverse weather was not a major contributor. Instead, overconfidence and reduced caution during favourable driving conditions may lead to inattentive or aggressive driving behaviours.





Recommendation:

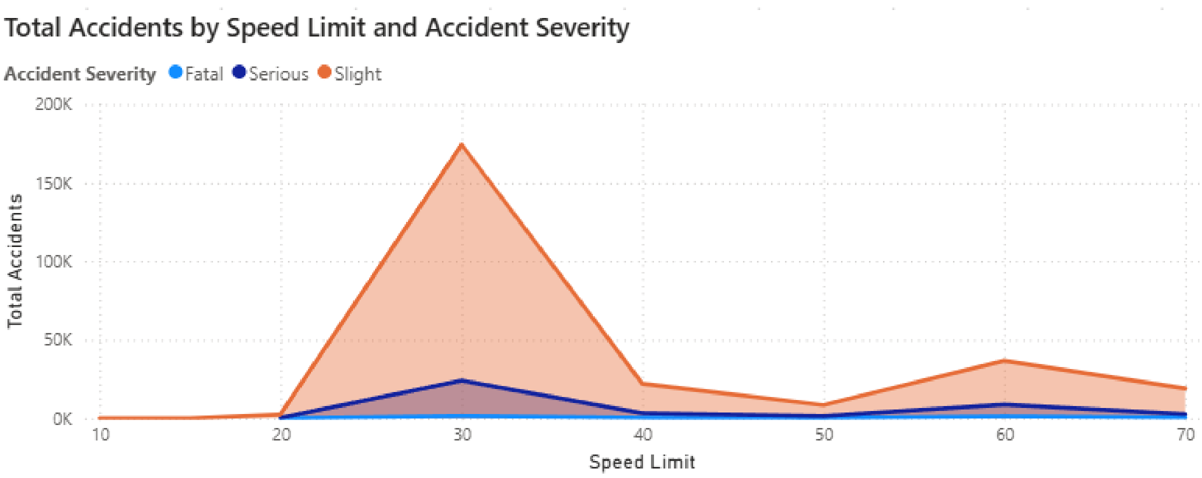
Continue **driver education** emphasizing that most accidents happen in normal weather. Always encourage defensive driving practices, regardless of apparent safety.

SPEED ZONES

Most accidents were recorded in the **30–40 mph speed limit zones**.

Interpretation:

These speed zones are typically found in urban and residential areas with high pedestrian activity. The issue is likely not just speed, but road design, lack of crossings, or insufficient signage that fail to protect vulnerable users.



Recommendation:

Authorities should audit road layouts in these zones and consider adding pedestrian crossings, traffic calming measures and better visual cues to reduce collision risk.

11. RISK MITIGATION STRATEGIES

Install traffic control measures at high-risk junctions

Many accidents occurred at uncontrolled junctions, indicating the absence of signals or roundabouts. Installing proper traffic control systems—such as traffic lights, roundabouts, or stop signs—can significantly reduce conflict points and improve flow management.

Improve street lighting in rural and poorly lit zones

Severe accidents were frequently recorded in dark, unlit rural areas. Upgrading lighting infrastructure in these zones will enhance visibility for drivers, reduce nighttime collision risk and improve pedestrian safety.

Redesign roads in 30–40 mph zones

The majority of accidents happened in moderate-speed zones, particularly in urban areas. Redesigning such roads with clearer signage, better pedestrian crossings and physical traffic calming measures (like speed bumps or narrowed lanes) will improve safety for all users.

Ensure road surface quality and drainage

Although most accidents happened on dry roads, weather-sensitive regions require proper road surface maintenance. Ensuring grip, drainage and hazard-free surfaces during rain or frost conditions is vital for reducing slippage and hydroplaning.

Policy & Data-Driven Enforcement

Target high-risk time slots with strategic enforcement Afternoon hours and Fridays emerged as peak accident times. Local authorities should increase patrolling, deploy mobile traffic units, or use digital signage during these periods to enforce speed limits and deter reckless driving.

Promote standardized accident reporting protocols

Incomplete data in fields like Junction Control suggests a lack of reporting discipline. Mandating standardized and complete data entry can help authorities collect better insights for long-term planning and intervention.

Deploy smart surveillance systems in accident-prone areas

Automated tools like traffic cameras and speed detectors can replace manual enforcement and ensure continuous monitoring, particularly in districts with high accident rates.

Conduct targeted safety campaigns

Motorcyclists and goods vehicle operators showed higher accident severity. Awareness programs should emphasize helmet usage, protective gear and cautious lane behaviour for motorcyclists and speed/load management for commercial vehicle drivers.

Launch seasonal and regional road safety drives

Slight increases in accidents during spring and summer suggest seasonal influence. Road safety campaigns during these months, especially in urban zones, can remind drivers of key precautions.

Reinforce driver training and licensing requirements

Improving driver education, especially for new and commercial drivers, ensures better understanding of road risks. Regular refresher courses or stricter licensing tests may help reduce careless or untrained driving behaviours.

This integrated safety action plan leverages insights from accident data to recommend infrastructure changes, improve reporting systems, enforce policy more effectively and influence driver behaviours through education. Implementing even a subset of these strategies can help authorities mitigate accident rates and build safer roads for the public.

CONCLUSION

This project successfully leveraged the analytical capabilities of Microsoft Power BI to conduct a comprehensive exploration of the 2021 UK road accident dataset. Through a systematic approach involving data cleaning, transformation, feature engineering and interactive dashboard development, key insights were uncovered about the conditions under which road accidents most commonly occur.

Findings from the analysis pointed to several high-risk patterns: the predominance of accidents on single carriageways, the critical role of uncontrolled junctions and the disproportionate severity of incidents occurring in rural zones with inadequate lighting. Time-based trends showed peaks during afternoon hours and on Fridays, while vehicle-type analysis highlighted that motorcyclists and goods vehicle operators face higher fatality risks. Surprisingly, most accidents occurred under fine weather and dry road conditions, underscoring the influence of driver behaviours rather than environmental extremes.

The dashboards created as part of this project offer a powerful tool for stakeholders—including policymakers, traffic authorities and urban planners—to interactively explore accident patterns across multiple dimensions. These visualizations facilitate evidence-based decision-making and allow for targeted interventions.

To address the safety challenges revealed by this analysis, a multi-faceted action plan was proposed. This includes infrastructure improvements like signal installations and lighting upgrades, data governance enhancements that ensure reporting completeness and educational initiatives aimed at high-risk groups.

This project demonstrates how data-driven visual analytics can uncover hidden patterns, drive impactful recommendations and ultimately contribute to safer, more responsive transport systems. With continued efforts in data collection, real-time monitoring and public engagement, such analytical models can play a crucial role in reducing accidents and saving lives on the road.