Road Car Accident Data Analysis and Reporting

Dataset Overview

This dataset contains detailed records of road accidents, providing insights into accident severity, contributing factors, and environmental conditions. It includes information on accident location, time, road characteristics, weather conditions, vehicle involvement, and casualties. The dataset includes 21 columns.

Objective of the Dataset

The dataset is used for analyzing and understanding road accident trends, such as:

- Identifying accident-prone areas based on location data.
- Examining the impact of environmental conditions (weather, lighting) on accident severity.
- Studying vehicle involvement and accidents statistics.
- Providing insights to improve traffic management and reduce accident rates.

Description of the Dataset

Each row represents a single road accident incident, with the following key features:

Column Name	Description
Accident_Index	Unique identifier for each accident
Accident Date	Date when the accident occurred
Day_of_Week	Day of the week (e.g., Monday, Tuesday)
Time	Time of the accident
Accident_Severity	Severity level: Slight, Serious, or Fatal

Weather_Conditions	Weather during the accident
Light_Conditions	Light condition during the accident
Road_Surface_Conditions	Condition of the road surface
Speed_limit	Road speed limit (numeric)
Vehicle_Type	Type of vehicle involved
Number_of_Casualties	Number of people injured or killed
Number_of_Vehicles	Total vehicles involved
Road_Type	Type of road: One-way, Dual, etc.
Junction_Control	Type of traffic control at the junction
Junction_Detail	Specific junction layout
Carriageway_Hazards	Hazards present (if any)
Latitude / Longitude	Coordinates of the accident location
Local_Authority_(District)	Geographic district of accident
Urban_or_Rural_Area	Urban or Rural classification
Police_Force	Police division reporting the accident

Data Cleaning Process

1. Data Preparation

- **Handle Missing Values:** Fill missing values with defaults like "Unknown".
- **Remove Duplicates:** Ensure unique entries for accidents for Dim weather, Dim Vehicles, Dim Carriageway hazard, Dim light conditions, Data table and Road surface conditions
- **Standardize Formats:**
    - Convert dates to UK time (YYYY-MM-DD) format.
    - Ensure columns data types are properly formatted.

- **Validate and Normalize Data:**
    - Merge latitude and longitude into a location key.
    - Merge road attributes into a road key.
    - Correct specific typos:
        - "Auto traffic Sigl" → "Auto traffic signal".
        - "Data missing of out of range" → "Out of range".
        - "Give way or uncontrolled" → "Give way".
        - "Fetal" → "Fatal".
        - T or staggered junction → T junction
        - blank in table of Carriageway_Hazards replaced by None
        - Wet or damp → Wet
        - Frost or ice → ice
        - Motorcycle over 500cc /Motorcycle 125cc and under /Motorcycle 50cc and under/Motorcycle over 125cc and up to 500cc → Motorcycle
        - Goods 7.5 tonnes mgw and over/Van / Goods 3.5 tonnes mgw or under /Goods over 3.5t. and under 7.5t → Goods
        - Fog or mist→ Fog

2. Data Normalization and Modeling

Fact Table: FactAccidents

Column	Description
AccidentID	Primary Key
DateID	FK to Dim Date
Time	Time of day
Severity	Accident severity
LocationID	FK to DimLocation
RoadID	FK to DimRoad
WeatherID	FK to DimWeather
LightConditionID	FK to Dim Light Condition
SurfaceConditionID	FK to Dim Surface Condition
HazardID	FK to Dim Hazard
Vehicle Type ID	FK to Dim Vehicle Type
Number_of_Casualties	Numeric measure
Number_of_Vehicles	Numeric measure

Dimension Tables:

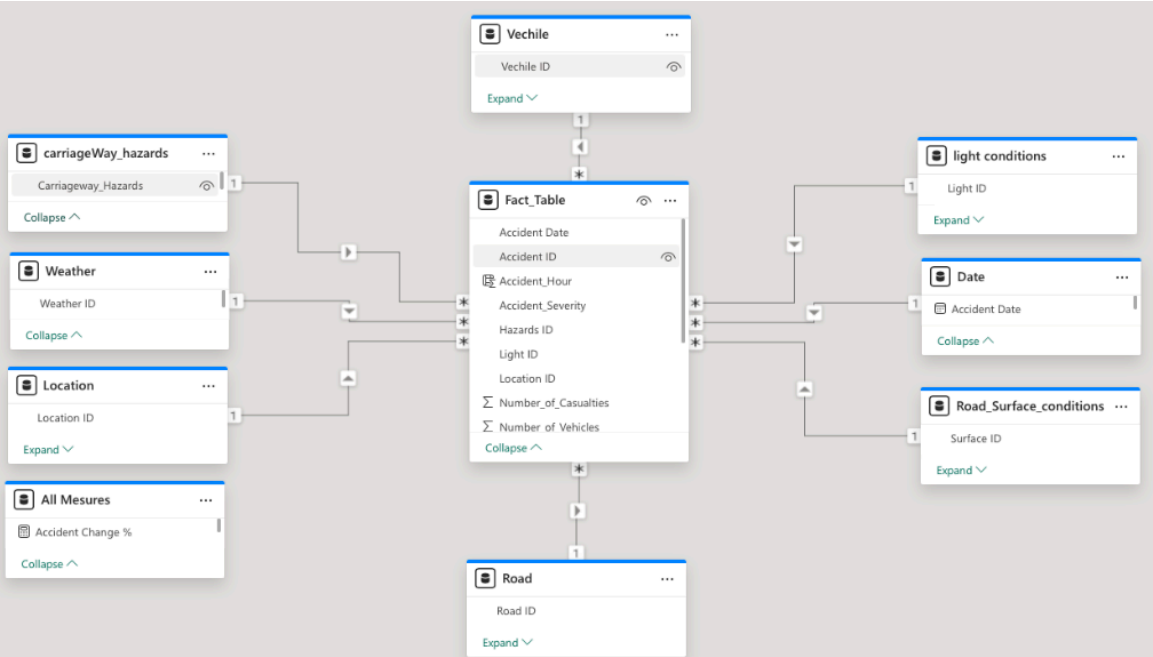
- **Date (DimDate):**
    - Date ID, Date, Day Of Week, Year, Month, Week, Quarter.
- **Location (Dim Location):**
    - Location ID, Latitude, Longitude, Local Authority District, Urban or Rural Area.
- **Road (DimRoad):**

    - RoadID, RoadType, SpeedLimit, JunctionControl, JunctionDetail.

- **Weather (DimWeather):**
    - WeatherID, WeatherCondition.
- **LightCondition (DimLightCondition):**
    - LightConditionID, LightCondition.
- **SurfaceCondition (DimSurfaceCondition):**
    - SurfaceConditionID, SurfaceCondition.
- **Hazard (DimHazard):**
    - HazardID, HazardType.
- **Vehicle (DimVehicleType):**
    - VehicleTypeID, VehicleType.

3. Relationships (Star Schema)

- Each dimension table has a one-to-many relationship with FactAccidents.



4. Creation of IDs and Custom Columns

- **Location key in fact table and Dim location :** Merge Latitude and Longitude.
- **Road ID:** Merge Road Type, Junction Control, Junction Details, and Speed Limit.
- **Vehicle, Hazard, Surface, Light, Weather IDs:** Generate using index columns in Power BI.

5. Additional Enhancements

- Add columns for Accident Hour and Rush Hour based on the Time column in fact table.
- Using power query we added season column

Steps to Implement Visualizations in Power BI

- Load normalized tables.
- Establish relationships.
- **Measures:**

1. NumOfAccidents = COUNTROWS('FactAccidents')

2. NumOfCasualties = SUM('FactAccidents'[Number_of_Casualties])
3. Accident Change % = [VAR PreviousYearAccidents = CALCULATE([NumOfAccidents], PREVIOUSYEAR('DimDate'[Date])) RETURN IF(NOT(ISBLANK(PreviousYearAccidents)), ([NumOfAccidents] - PreviousYearAccidents) / PreviousYearAccidents, BLANK())]
4. Trend Analysis = IF([Accident Change %] > 0, "Worsening", "Improving")
5. Accident Change % = VarPreviousYearAccidents=CALCULATE([NumOfAccidents],PREVIOUSYEAR('Date'[Accident Date].[Date])) return if(NOT(ISBLANK(PreviousYearAccidents)),([NumOfAccidents]-PreviousYearAccidents)/PreviousYearAccidents,BLANK()
6. Fatal_accidents = CALCULATE(COUNT('Fact_Table'[Accident ID]),'Fact_Table'[Accident_Severity]="Fatal")

Overview Questions

1. What is the total number of road accidents?

- **Answer:** 308K road accidents

2. What is the total number of casualties resulting from accidents?

- **Answer:** 417.9K casualties.

3. Which season has the highest number of accidents?

- **Answer:** Autumn is the most affected season

4. Is the trend of accidents improving or worsening?

- Answer: Accidents are improving.

5. Are casualties more frequent in urban or rural areas?

- Urban areases

6. Which road type experiences the most accidents?

- Single carriageways have the highest number of accidents compared to other road types

7. How are accidents distributed by severity (Slight, Serious, Fatal)?

- Most accidents are slight (85.49%), followed by serious (13.23%), and fatal accidents (1.28%).

8. How do accident and casualty numbers vary across the months of the year?

- Accidents and casualties are highest in November and gradually decrease throughout the year, with the lowest numbers around February.

Time Questions

1. **Which day of the week has the highest number of accidents?**
- Answer: Saturday has the highest number of accidents
2. **How are accidents distributed by rush hour?**
- Answer: Morning Rush Hour: 50.24% (155K accidents), Evening Rush Hour: 28.83% (89K accidents) and Off-Peak: 20.93% (64K accidents)
3. **What season has the highest number of accidents?**
- Answer: Autumn has the highest number of accidents while Season with the fewest accidents Winter
4. **At what hour of the day do most accidents occur?**
- Answer: Most accidents occur between 8 and 17 , with a noticeable peak around 15 PM.

Trends

1. **Which districts have the highest number of accidents?**

- Birmingham, Manchester, London area, and two northern districts (based on circle size and color).

2. **How has the total number of accidents changed compared to previous years?**

- It's currently high (307,973), but based on the forecast, it is expected to decrease in future years.

3. **How significant is accident severity in impacting fatalities?**

- Very significant — accident severity is listed first, suggesting it's a major contributor to fatalities.

4. **How do carriageway hazards correlate with accident severity?**

- Carriageway hazards (like road surface defects, flooding) likely increase accident severity and fatality rates.

Location and vehicles questions

1. On which type of carriageway do accidents involving multiple vehicles occur most frequently?

   Answer: Dual carriageway has the highest number of accidents involving multiple vehicles (98 accidents).

2. Which vehicle type is involved in the highest number of accidents?

   Answer: Cars are involved in the highest number of accidents

3. Are more accidents reported in urban areas or rural areas?

   Answer: Urban areas report more accidents (**199K**) than rural areas (**109K**).

4. What percentage of the total accidents occur in urban areas?

   Answer: Urban accidents account for around **65%** of total accidents.

5. Which district has the highest number of reported accidents?

   Answer: Birmingham with 6.2K accidents.

6. Are there districts with similar accident numbers among the top 5 listed?

   Answer: Yes, **Manchester** (3.1K) and **Bradford** (3.0K) have very close accident numbers.

7. How do light conditions impact the number of accidents in urban versus rural areas?

- Most accidents happen during daylight (especially in urban areas), but accidents during darkness with lights lit are also significant. Rural areas have fewer accidents overall compared to urban areas across all light conditions.

8. Which districts report the highest number of fatal accidents?

- **London** reports the highest number of fatal accidents, followed by **Birmingham**. They are the most critical regions needing road safety interventions.

Road Conditions questions

1. **What is the most dangerous road condition?**

- **Answer:** Dry surfaces lead to the highest number of casualties (0.28M), likely due to higher speeds and more traffic.

2. **Are icy or snowy roads causing many casualties?**

- **Answer:** No, casualties on icy (0.02M) and snowy (0.01M) roads are much fewer, possibly because drivers are more cautious or avoid driving.

3. **At which speed limit do severe accidents peak?**

- **Answer:** Around 30 and 60 mph show noticeable peaks in severe accidents.

4. **Do higher speed limits always mean more severe accidents?**

- **Answer:** Not necessarily. The graph shows fluctuations rather than a steady increase with speed.

5. **Which type of junctions have the highest number of accidents?**

- T-junction: 97K accidents

6. Do dark conditions without street lights cause more accidents than those with lighting?

- **Answer:** Accidents in darkness **without lights** are significantly fewer than those **with lights**, suggesting visibility isn't the only factor—traffic density might be more influential.

Weather & Insights questions

1. **Which weather condition has the highest number of slight accidents?**
   Answer: *Fine no high winds*

2. **Is there a relationship between accident severity and weather condition?**
   Answer: Yes. Poor weather (like rain or fog) is associated with a higher proportion of serious and fatal accidents, while fine weather has more slight accidents.

3. **Which weather condition results in the highest number of casualties?**
   Answer :*Fine no high winds*

4. **Does rainy weather cause a significant number of casualties despite fewer accidents?**
   Answer: Yes. Rainy weather results in a relatively high number of casualties, showing that those accidents are often more severe
5. **Which road type experiences the highest number of serious or fatal accidents?**
   Answer: Single carriageways

6. **Are fatal accidents more common on single or dual carriageways?**
   Answer: Fatal accidents are more common on single carriageways than on dual carriageways.

7. Which carriageway hazard causes the most accidents?
   Answer: None

Conclusion questions

1. What factors most influence accident severity to be fatal?

Answer:

- The most influential factor is **"Light Conditions = Darkness – no lighting"**, increasing fatal severity likelihood by **3.24x**.

- Other influential factors include:

    - Road surface = Flood over 3 cm deep (1.97x)

    - Road type = Dual carriageway (1.59x)

    - Weather = Fine + high winds (1.55x)

Key Insights and Recommendations

- **Visibility and Lighting:**
    - Improving street lighting, especially in rural or dark areas, could reduce accident severity.
- **Road Conditions:**
    - Proactive maintenance of roads during adverse weather (floods, ice) can significantly reduce fatal accident risk.
    - Single carriageway more crowded so lead to more accident based on this we suggest build new bridges and road expansion
- **Speed Management:**
    - Enforcing appropriate speed limits, especially around 30 and 60 mph zones, could curb severe accidents.

- **Traffic Management at Junctions:**
    - Enhanced signage, better traffic control at T-junctions and crossroads, and visibility improvements could lower accident counts.
- **Seasonal Awareness Campaigns:**
    - Targeted safety campaigns in autumn and winter focusing on weather-related driving tips could be effective.
- **Vehicle Type-Specific Strategies:**
    - Special safety programs for motorcyclists could address the higher severity of their accidents.

Conclusion

This analysis provides a detailed understanding of road accident patterns based on environmental, road, and vehicle-related factors. By focusing efforts on improving lighting, managing speeds, maintaining road conditions, and targeted safety campaigns for vulnerable vehicle types, accident rates and severities can be effectively reduced.