

Power BI Road Accident Data Project

STEPS IN THE PROJECT

- Requirement Gathering
- Stakeholders in a project
- Raw data overview
- Connecting data with Power BI
- Data Cleaning
- Data Processing
- Data Modeling
- Data Visualization
- Dashboard building
- Insights

REQUIREMENTS

Creating Road Accidents dashboard for a year 2021 and 2022

- Primary KPI – Total Casualties and Total Accident values for Current Year and Year and Year growth (*This is about tracking the **total number of casualties (injuries + deaths)** and the **total number of accidents** in the current year (e.g., 2022). Then, compare it to the previous year (e.g., 2021) to calculate the **Year-over-Year (YoY) growth.***)
- Primary KPI's – Total Casualties by Accident Severity for Current Year and Year and Year growth (*This KPI tells you **how many people were injured or killed in accidents this year, broken down into fatal, serious, and slight injuries, and how those numbers have changed compared to last year.***)

- Secondary KPI's – Total Casualties with respect to vehicle type for Current Year (*Break down casualties based on what type of vehicle was involved in the accident. Cars, motorcycles, bicycles, trucks, buses, etc.*)
- Monthly trend showing comparison of casualties for Current Year and Previous Year (Show casualties' month by month (*Jan–Dec for the current year and compare them with the same months in the previous year.*))
- Casualties by Road Type for Current year (*Show where the accidents happened, depending on road type. Highways, urban roads, rural roads, intersections, etc.*)
- Current Year Casualties by Area/ Location & by Day/ Night (*Compare casualties based on geography (area/location) and whether the accident happened in the daytime or nighttime.*)
- Total Casualties and Total Accidents by Location (*Show both the number of casualties and the number of accidents for each location (city, region, province, etc.).*)

STAKEHOLDERS ANALYSIS

- **Ministry of Transport** → Policymaking, national road safety strategy.
- **Road Transport Department** → Operational oversight of roads and accident reporting.
- **Police Force** → Law enforcement, accident investigations, traffic control.
- **Emergency Services Department** → Ambulance, fire, medical teams responding to accidents.
- **Road Safety Corps** → Organizations focused on reducing accidents through campaigns, enforcement, and education.
- **Transport Operators** → Bus, taxi, trucking, logistics companies who want safer operations.
- **Traffic Management Agencies** → Manage signals, congestion, and traffic flow to prevent accidents.
- **Public** → Drivers, pedestrians, passengers who need safer roads.
- **Media** → Share accident statistics with the public, raise awareness.

Data Cleaning

How to fix this in Power BI (Power Query Editor):

Go to Power Query Editor (you're already there in your screenshot).

Select the column Accident Severity.

On the toolbar, click Transform > Replace Values.

In the pop-up:

*Value to Find: **Fetal***

*Replace With: **Fatal***

Click OK.

Review to confirm that only the incorrect spelling was changed.

Apply changes (Close & Apply) to load the cleaned data back into Power BI.

Data Modeling

Why create a Calendar table?

- Accident datasets (or any transactional data) usually just have a **date column** (e.g., accident_date = 2022-04-15).
- To do **time-based analysis** (like YoY growth, monthly trends, comparing years, etc.), you need a **Calendar/Date table** that Power BI can use for calculations and filtering.
- Without it, Power BI won't know how to roll up data by **month, quarter, or year** properly.

What goes in a Calendar Table?

At minimum, you'll want these fields:

- **Date** (continuous list of all days in your dataset range, e.g., 2021–2022)
- **Year** (e.g., 2021, 2022)
- **Month Number** (1–12, useful for sorting)
- **Month Name** (Jan, Feb, Mar ...)
- **Quarter** (Q1, Q2, etc.)
- **Day Name** (Monday, Tuesday, etc.)
- **Day/Month/Year columns**

Optional (but useful for your accident project):

- **Day/Night Flag** (if your dataset can map accident time to day vs night).
- **Year-Month column** (e.g., 2022-Jan, for trend charts).

How it works in modeling

- You create the **Calendar table** (either with DAX or Power Query).
- Then, in the **Model view**, you **link** your Calendar[Date] to your Accident Data[Accident_Date].
- This relationship lets you slice and analyze casualties/accidents across different time levels (year, month, quarter).

Power BI Desktop interface showing a table named 'Calendar' with columns: Date, Year, Month. The table contains data for January 2021, with dates ranging from 01/01/2021 to 31/01/2021. The 'Month' column is highlighted. The 'Data' pane on the right shows a list of tables and columns, including 'Calendar', 'Date', 'Month', 'Year', 'Data', 'Accident Date', 'Accident Index', 'Accident Severity', 'Carriageway Hazards', 'Day of Week', 'Junction Control', 'Junction Detail', 'Latitude', 'Light Conditions', 'Local Authority (District)', 'Longitude', 'Number of Casualties', 'Number of Vehicles', 'Police Force', 'Road Surface Conditions', 'Road Type', 'Speed Limit', 'Time', 'Urban or Rural Area', 'Vehicle Type', and 'Weather Conditions'.

Table: Calendar (30 rows) Columns: Month (12 distinct values)

Power BI Desktop interface showing the 'New relationship' dialog box. The dialog is used to create a relationship between the 'Calendar' table and the 'Data' table. The 'From table' is 'Data' and the 'To table' is 'Calendar'. The relationship is defined by the 'Date' column in 'Data' and the 'Date' column in 'Calendar'. The cardinality is set to 'Many to one (1:1)' and the cross-filter direction is 'Single'. The 'Make this relationship active' checkbox is checked. The 'Assume referential integrity' checkbox is unchecked. The 'Apply security filter in both directions' checkbox is unchecked. The 'There's already a relationship between these two columns' error message is displayed. The 'Properties' pane on the right shows the 'General' tab for the 'Calendar' table, including fields like 'Description', 'Synonyms', 'Display folder', 'Is hidden', 'Formatting', 'Data type', and 'Date time format'.

New relationship

Select tables and columns that are related.

From table: Data

To table: Calendar

Cardinality: Many to one (1:1)

Cross-filter direction: Single

Make this relationship active: ☒

Assume referential integrity: ☐

Apply security filter in both directions: ☐

There's already a relationship between these two columns.

Properties

General

Description: Enter a description

Synonyms: date

Display folder: Enter the display folder

Is hidden: ☐ No

Formatting

Data type: Date/Time

Date time format: *14/03/2001 13:30:55 (General Date)

1. Primary KPI – Total Casualties and Total Accidents for Current Year and YoY Growth

For Total Casualties

- **What it means:**
 - You need the total number of casualties and accidents **for the selected/current year**.
 - YoY growth is the percentage change compared to the previous year.

Code: Current Year Casualties = TOTALYTD(SUM(Data [Number_of_Casualties]), Calender[Date])

Code: PreviousYearCasualties = CALCULATE(SUM(Data[Number_of_Casualties]),
SAMEPERIODLASTYEAR(Calender[Date]))

=

CasualtiesYoYGrowth =

DIVIDE([CurrentYearCasualties] - [PreviousYearCasualties], [PreviousYearCasualties], 0)

For Total Accidents

Code: Current year Accidents = TOTALYTD(COUNT(Data[Accident_Index]), Calender[Date])

Code: Previous year Accidents =
CALCULATE(COUNT(Data[Accident_Index]),SAMEPERIODLASTYEAR(Calender[Date]))

Code: Year on Year Accidents = DIVIDE(
[Current year Accidents] - [Previous year Accidents],
[Previous year Accidents],
0)

2. Primary KPI – Total Casualties by Accident Severity for Current Year and YoY Growth

- **What it means:**
 - Break down casualties by **severity levels**, e.g., Minor, Serious, Fatal.
 - Show current year totals and compare with previous year.

3. Secondary KPI – Total Casualties by Vehicle Type for Current Year

- **What it means:**
 - Show how many casualties were caused by different types of vehicles (Car, Truck, Bike, etc.) **for the current year.**

4. Monthly Trend – Current Year vs Previous Year

- **What it means:**
 - Compare casualties month by month for current year and previous year.

5. Casualties by Road Type for Current Year

- **What it means:**
 - Show casualties based on the type of road (Highway, City Road, Rural Road, etc.) for the current year.
- **Visualization:** Column chart or stacked bar chart

6. Current Year Casualties by Area/Location & Day/Night

- **What it means:**
 - Show how casualties vary by **location** and by **time of day**.
- **Visualization:** Matrix or stacked bar chart

7. Total Casualties and Total Accidents by Location

- **What it means:**
 - Aggregate casualties and accidents **per location**.
- **Visualization:** Map or bar chart