Introduction

Analyzing road accident datasets can provide valuable insights into the factors contributing to accidents, thereby enabling policymakers, traffic engineers and law enforcement agencies to implement targeted interventions for prevention and mitigation. I came across the dataset on Kaggle and decided to apply my skills and expertise in data cleaning, analysis and visualization on it, with the hope of generating valuable insights that can be helpful for the stakeholders.

Microsoft Excel Concepts Applied:

* Pivot table Analysis
* Pivot chart
* Dynamic chart & buttons
* Visualization filtering(Slicer & Timeline)
* Creation and modification of new and existing fields/items respectively

Problem Statement

This analysis aims to understand how factors like road type, road surface, location/area, time of the day and vehicle type determine the number of survivors of road accidents in the UK between 2021 and 2022. We will therefore create a Report Dashboard so that we can have insight into the following requirements and KPIs:

* **Primary KPI** – Total Casualties recorded for each year and the entire period under consideration
* **Primary KPIs** – Number and percentage for each accident severity type to total casualties.
* **Primary KPI** – Number and percentage of maximum casualties by type of vehicle.
* **Secondary KPIs** – Total Casualties for each vehicle type.
* **Secondary KPIs** – Monthly trend showing the comparison of casualties for both years under examination.
* Breakdown of total casualties by Road Type.
* Distribution of total casualties by Road Surface.
* Distribution of total casualties by urban-rural classification and time of the day.

Dataset

Description: The dataset of road accidents in 2021 and 2022. The dataset contains 1 sheet/table of 307973 unique rows with 21 fields/columns. The Accident\_Index column is the primary key for the dataset. The file extension is .xlsx

Data source: https://www.kaggle.com/datasets/syedibrahim03/road-accidents-dataset

Data Cleaning

* Each column was inspected and error was corrected in the severity column. Typographical error of fetal was modified to fatal with “find and replace” tool.
* All fields checked to ensure there are no null values in the fields that SHOULD NOT have null values

Data Processing

* New field Month was inserted and TEXT function was used to populate that field from the Accident Date field
* New field Year was inserted and TEXT function was used to populate that field from the Accident Date field