**INTRODUCTION TO DATA MANAGEMENT**

**PROJECT REPORT**

(Project Semester January-April2025)

***For the partial fulfilment of the BTech Computer Science Engineering***

***(Road Accident Dataset Analysis)***

Submitted by

**NAME: Abhay Chaudhary**

**Registration No:12315216**

Programme and Section BTech CSE

Course Code: INT217

Under the Guidance of

**(Mrs. ASHU**

**Assistant Professor**

**SCSE (Data Science)**UID-23631

**Discipline of CSE/IT**

**Lovely School of Computer Science and Engineering**

**Lovely Professional University, Phagwara**

# Declaration

I, ABHAY CHAUDHARY, student of under B. TECH CSE Discipline at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 9/4/2025

Signature: ABHAY

Registration No:12315216

Name: ABHAY CHAUDHARY

# Certificate

This is to certify that Abhay Chaudhary bearing Registration No. 12315216, has completed the INT217 project titled “ROAD ACCIDENT ANALYSIS DASHBOARD” under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort, and study.

Signature: ABHAY

Name of the Supervisor: MRS. ASHU

School of Computer Applications

Lovely Professional University, Phagwara, Punjab

# Acknowledgement

I would like to express my heartfelt gratitude to my project guide, Mrs. Aashu, for her continuous support, encouragement, and valuable insights throughout this project. Her expert guidance has been crucial in helping me shape the direction of my analysis and improve the quality of my work.

I am thankful to the School of Computer Science and Engineering at Lovely Professional University for providing the academic infrastructure and resources that made this project possible. The availability of labs and tools like Microsoft Excel played a vital role in performing the analysis effectively.

I sincerely appreciate the creators of the Road dataset available on Kaggle, which served as the foundation for this project. Their contribution enabled me to gain practical experience in data preprocessing, analysis, and visualization.

I also extend my thanks to the faculty members whose lectures and coursework helped me build a strong foundation in data handling and reporting.

A special note of gratitude to my parents and family members for their consistent encouragement, emotional support, and patience throughout the duration of this work. Their faith in me has been my biggest strength.

Finally, I acknowledge the self-discipline, dedication, and perseverance it required to complete this as an individual project. This report is a reflection of my personal growth, academic learning, and independent effort.

Table of Contents

1. **Introduction**
   * 1.1 Background
   * 1.2 Objective
2. **Dataset Description**
   * 2.1 Data Source
   * 2.2 Key Columns Used and Their Relevance
3. **Dashboard Layout & Insights**
   * 3.1 Key Performance Indicators (KPIs)
   * 3.2 Monthly Casualties Trend
   * 3.3 Casualties by Vehicle Type
   * 3.4 Casualties vs Road Type
   * 3.5 Casualties by Road Surface Condition
   * 3.6 Casualties by Urban vs Rural Area
   * 3.7 Casualties by Light Condition
4. **Dashboard Features**
   * 4.1 Slicers
   * 4.2 Cards and Visual Indicators
   * 4.3 Linked Visuals
5. **Conclusion**
   * 5.1 Key Takeaways
   * 5.2 Recommendations
6. **Future Scope**
7. **References**
8. **Author Information**

**Road Accident Dashboard Analysis Report (2021–2022)**

# 1. Introduction

### I. Background and Motivation:

### Road traffic accidents are one of the leading causes of death and injury worldwide. According to the World Health Organization (WHO), approximately 1.3 million people die each year due to road traffic crashes. Beyond fatalities, millions more suffer non-fatal injuries, often resulting in long-term disabilities. The impact spans across economic loss, social burdens, and emotional trauma. Understanding patterns in road accidents through data analysis is crucial for formulating evidence-based policies, improving infrastructure, and raising public awareness.

**II. Objective**

**The primary goal of this project is to create a visual dashboard using Excel that captures key metrics related to road accidents in 2021 and 2022. The dashboard helps:**

* **Identify high-risk road types, weather conditions, and time periods.**
* **Examine the role of vehicle type and environmental factors in accident severity.**
* **Empower policymakers and road safety authorities with actionable insights.**
* **Offer a comparative analysis across years to detect any increasing or decreasing trends.**

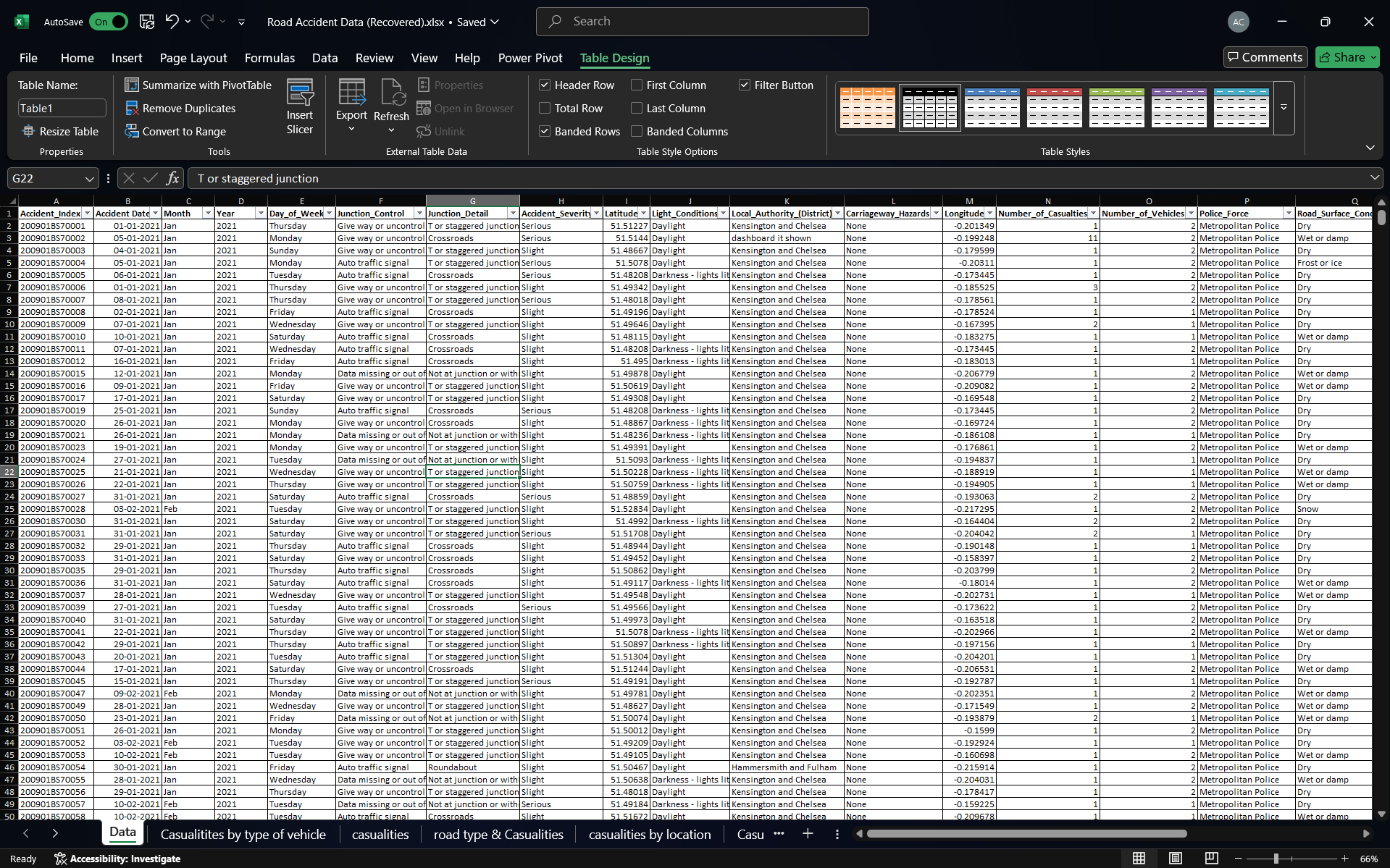
### 

**2. Dataset Description**

**I. Data Source**

The data has been compiled from **UK Department for Transport** datasets and includes accidents reported by the police. The dataset is considered highly reliable as it contains information on:

* Exact date and time of occurrence
* Environmental and road conditions
* Type and number of vehicles involved
* Severity of accidents
* Urban or rural location classification
* <https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-annual-report-2021>



These variables are essential in correlating **external conditions with accident occurrences**.

**II. Key Columns Used and Their Relevance**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **Column Name** | **Description & Purpose** | | Accident\_Index | Unique identifier, helpful for filtering and tracking individual incidents. | | Accident\_Severity | Critical for categorizing accidents into Fatal, Serious, or Slight based on outcome. | | Date, Time | Helps understand temporal patterns and peak accident times. | | Light\_Conditions | Identifies if visibility (e.g., darkness) is a contributing factor. | | Weather\_Conditions | Establishes impact of fog, rain, frost on accident probability. | | Road\_Surface\_Conditions | Links slipperiness or dryness of the surface to accident causes. | | Road\_Type | Assesses how road design/layout affects safety. | | Vehicle\_Type | Analyzes which types of vehicles are more accident-prone. | | Number\_of\_Casualties | Helps quantify the human impact of accidents. | | Urban\_or\_Rural\_Area | Provides a contextual difference between high-density and low-density regions. | |

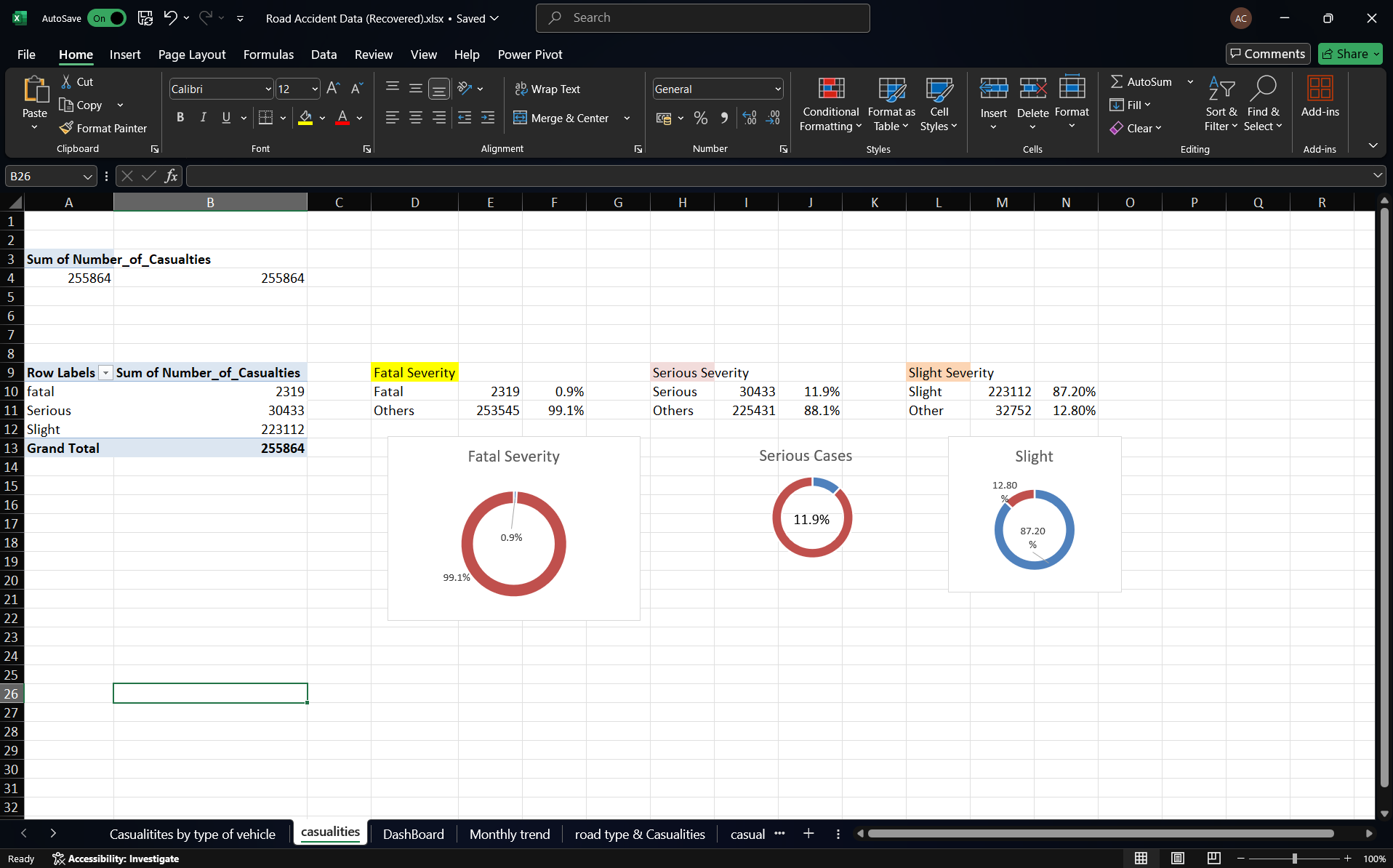
**3. Dashboard Layout & Insights**

**I. 📊 Key Performance Indicators (KPIs)**

KPIs are important **summary metrics** used to evaluate the overall performance and impact of accidents. These are shown as **interactive cards** at the top of the dashboard for quick insights.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | KPI | Value | Interpretation | | Total Casualties | 255,864 | Total number of people affected in the accidents (includes all severity levels). | | Fatal Severity | 2,319 | Represents deaths—although a smaller percentage, these have the highest cost. | | Serious Cases | 3,043 | Injuries requiring hospitalization or major medical attention. | | Slight Casualties | 223,112 | Injuries that are minor—yet, contribute to a large portion of the burden. | | Car Casualties | 203,357 | The highest contributor to total casualties—likely due to high car density. | |
|  |

**Theoretical Note:**  
Using KPIs aligns with **descriptive analytics**, which is the first stage of the data analytics lifecycle, helping summarize historical data in a meaningful way.



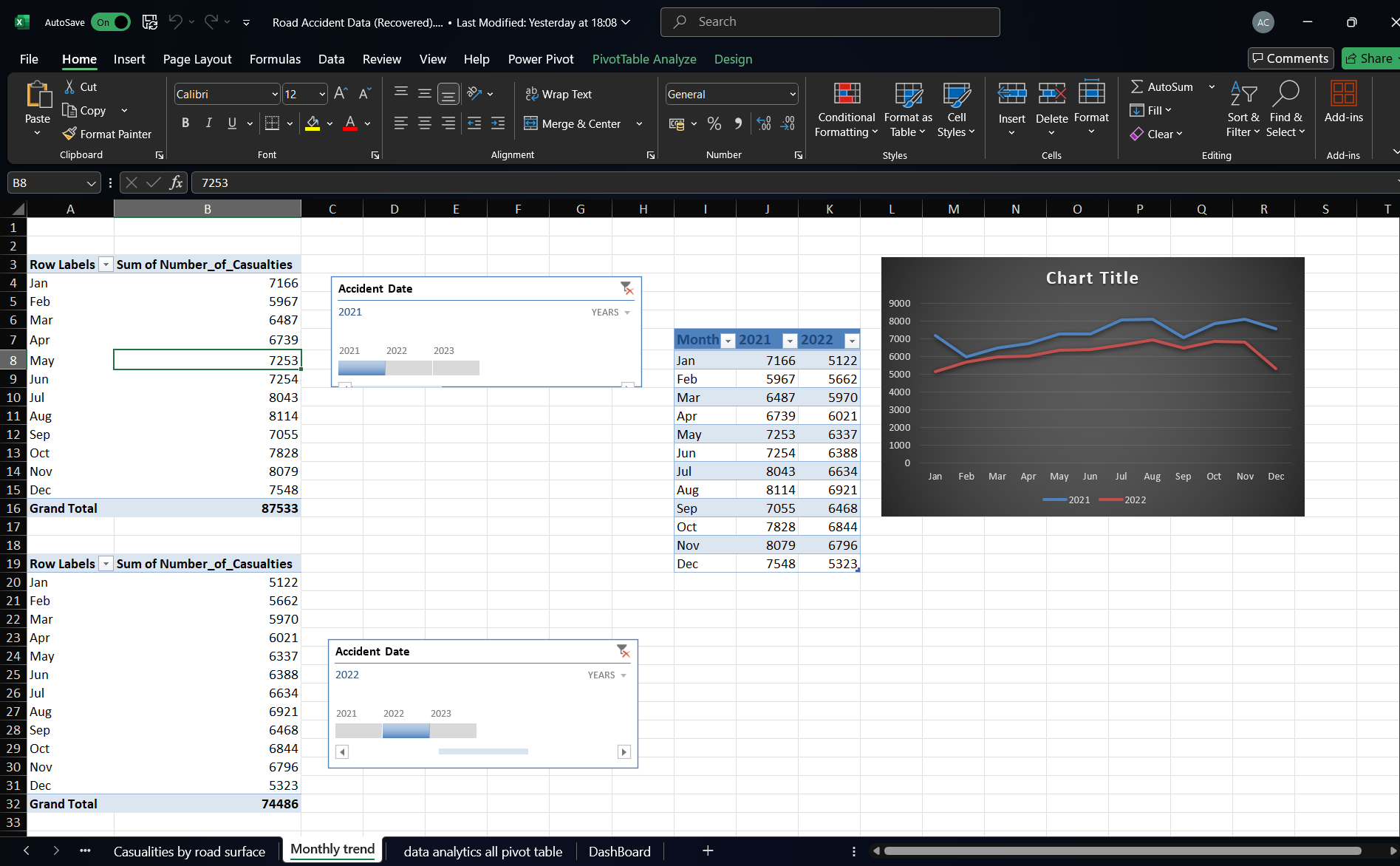
**II. 📈 Monthly Casualties Trend**

A **dual-line chart** presents a comparison of monthly casualties for 2021 and 2022. This helps spot seasonal fluctuations and anomaly periods.

* **Observation:** Casualties peak during **October and November**, which may relate to:
  + Shorter daylight hours.
  + Poorer weather (fog/rain).
  + Festivals or holiday season traffic congestion.

**Theoretical Note:**

This time-series visualization supports **trend analysis** and forecasting. It helps in proactive planning by identifying high-risk months and resource deployment windows.



**III. 🚙 Casualties by Vehicle Type**

This bar chart visualizes the **total number of casualties based on the type of vehicle** involved.

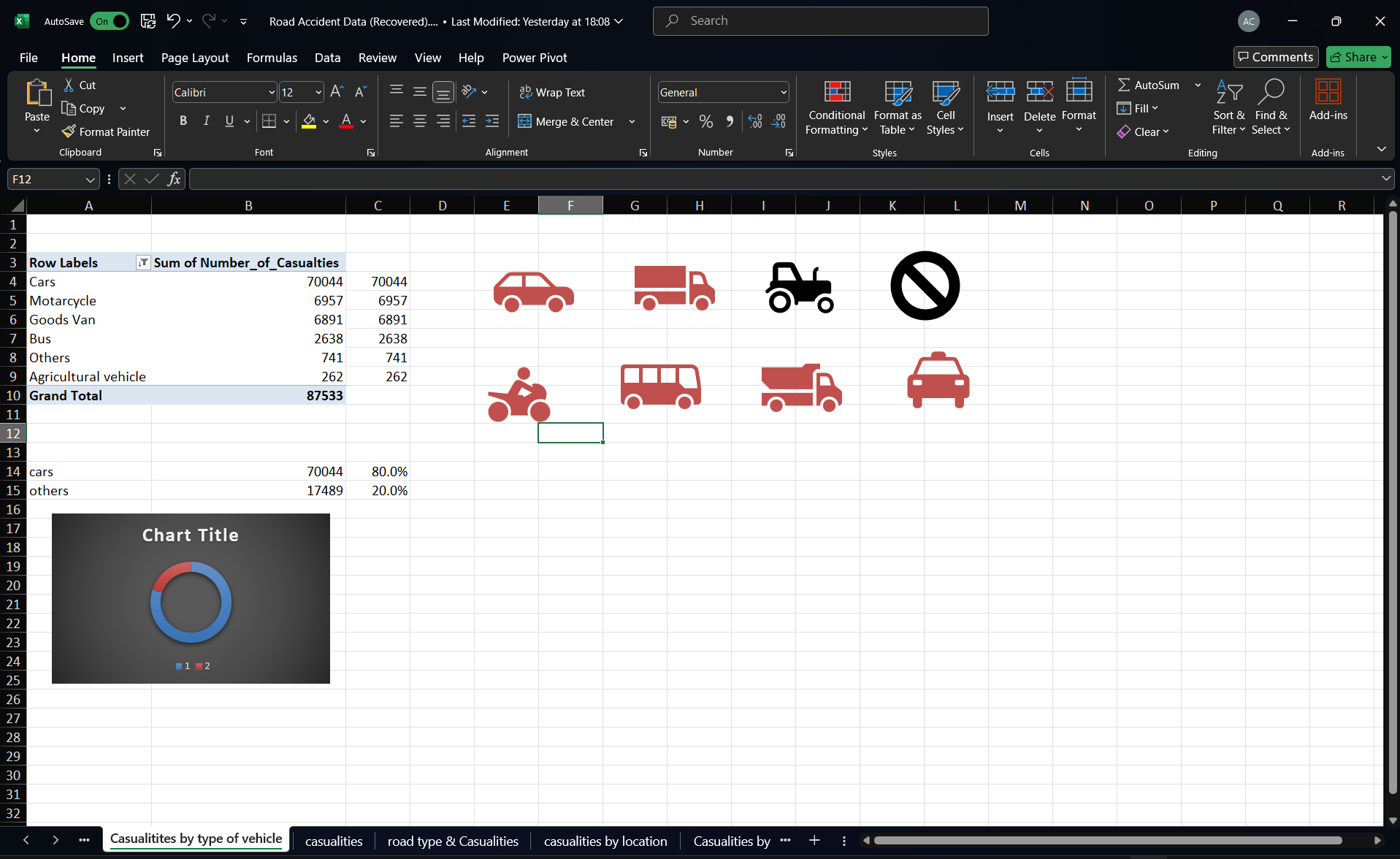
|  |  |
| --- | --- |
| Vehicle Type | Casualty Count |
| Cars | 203,357 |
| Vans | 20,416 |
| Buses | 8,135 |
| Motorcycles | 21,218 |
| Pedal Cycles | 598 |
| Other | 2,140 |

**Insights:**

* **Cars** contribute the most, which is consistent with their dominance on the roads.
* **Motorcycles**, though fewer in number, show relatively high casualty counts, indicating higher vulnerability.

**Theoretical Note:**

This represents **categorical data analysis**, helping understand which entities contribute the most to the observed phenomenon.



**IV. 🛣️ Casualties vs Road Type**

Road type is a major determinant in accident severity due to visibility, lane width, and traffic management.

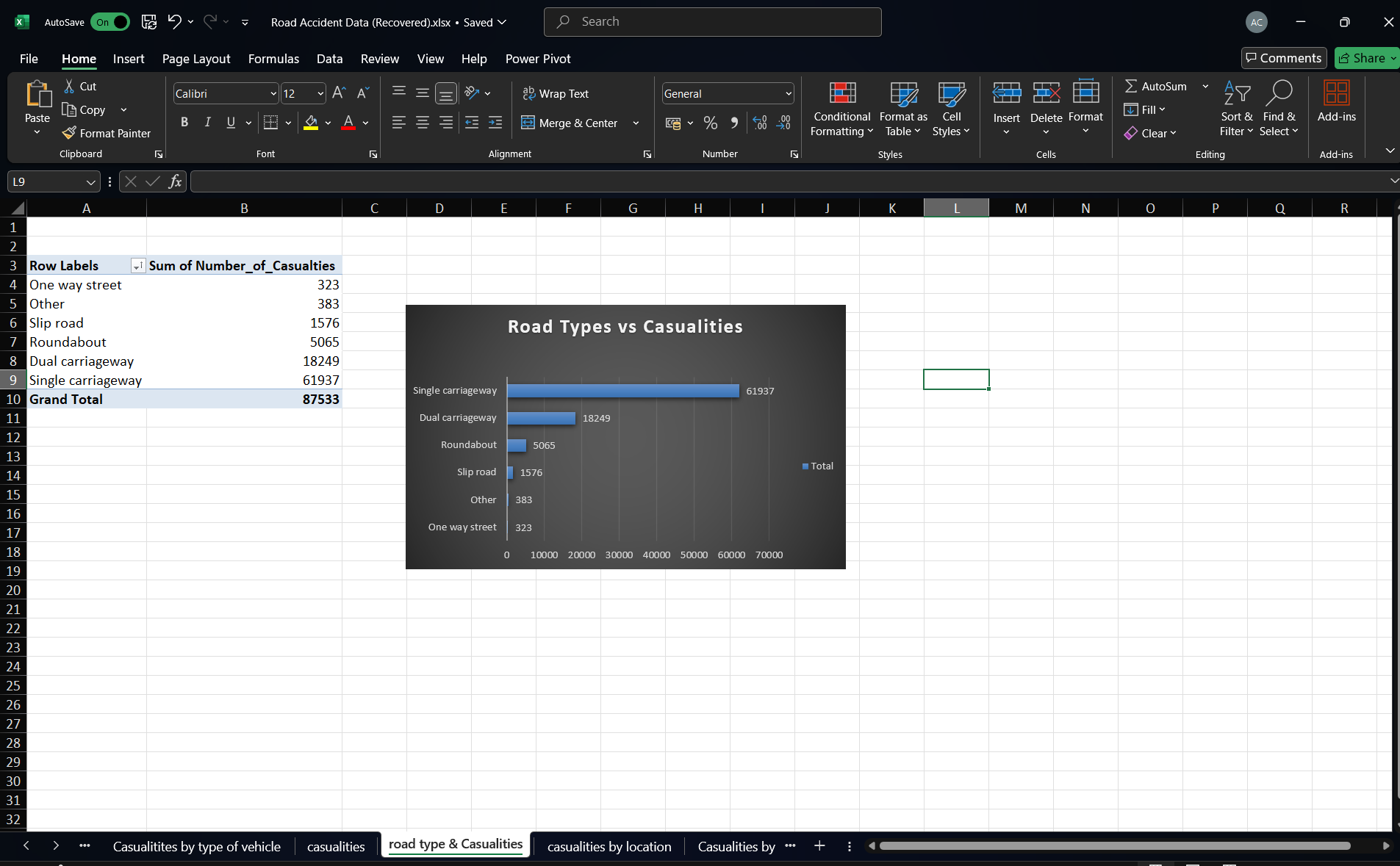
|  |  |
| --- | --- |
| Road Type | Casualties |
| Single Carriageway | 196,600 |
| Dual Carriageway | 32,400 |
| Roundabout | 17,200 |
| One-Way Street | 6,800 |
| Slip Road | 1,700 |
| Other/Unknown | 1,300 |

**Insight:**

* **Single carriageways** pose the most risk due to the absence of a central divider, leading to head-on collisions.
* **Roundabouts**, while designed for safety, still have significant accident numbers.

**Theoretical Note:**

Analyzing the role of road geometry is part of **spatial risk modeling**, often used in urban planning and civil engineering.



**V. 🌧️ Casualties by Road Surface Condition**

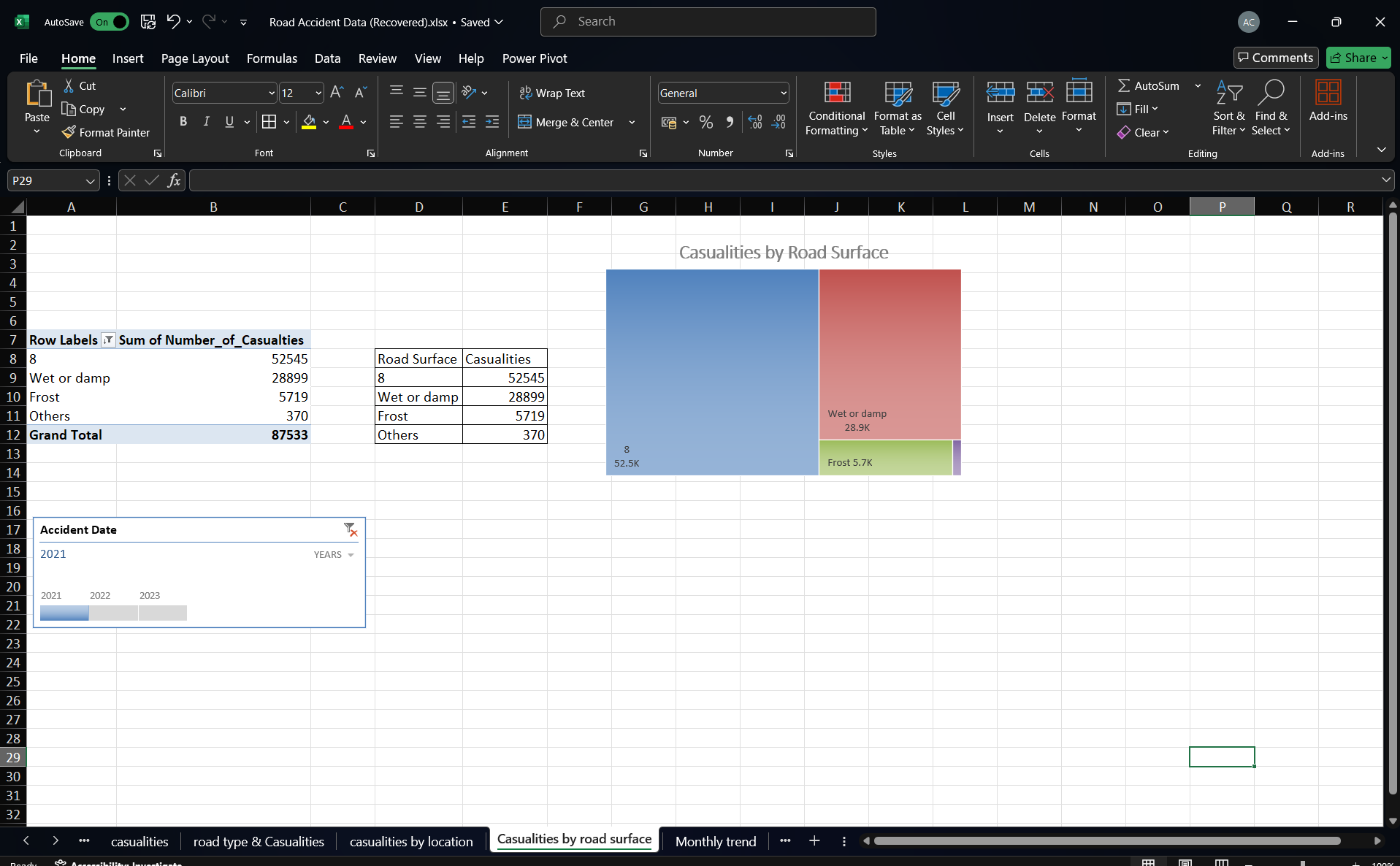
Road surface condition contributes directly to vehicle control and stopping distance.

|  |  |
| --- | --- |
| Surface Condition | Casualties |
| Dry | 181,300 |
| Wet/Damp | 64,400 |
| Frost/Ice | 9,800 |

**Insight:**  
While wet and icy roads are more dangerous, **more accidents occur on dry roads**, likely due to overconfidence and higher speeds.

**Theoretical Note:**

This supports the idea that human behaviour (e.g., risk compensation) affects accident likelihood more than environmental factors alone.



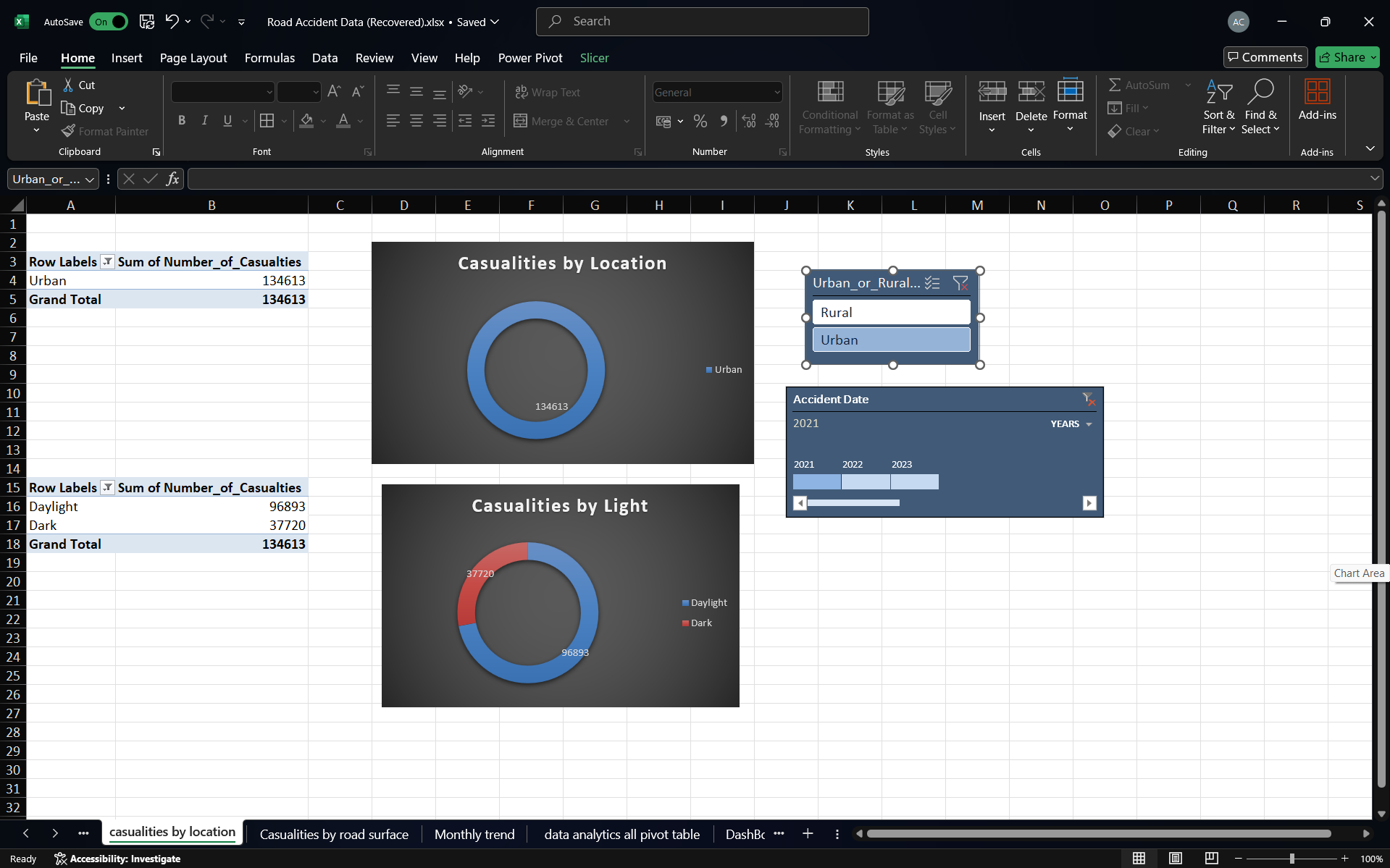
**VI. 🏙️ Casualties by Urban vs Rural**

|  |  |  |
| --- | --- | --- |
| Area Type | Casualties | Percentage |
| Urban | 186,782 | ~73% |
| Rural | 69,082 | ~27% |

**Insight:**  
Urban areas, due to higher population and vehicle density, naturally lead to more accidents. However, rural accidents tend to be more **severe** due to higher speeds and delayed medical attention.

**Theoretical Note:**

This reflects **geographical stratification** in data science, which classifies records based on location-specific attributes.



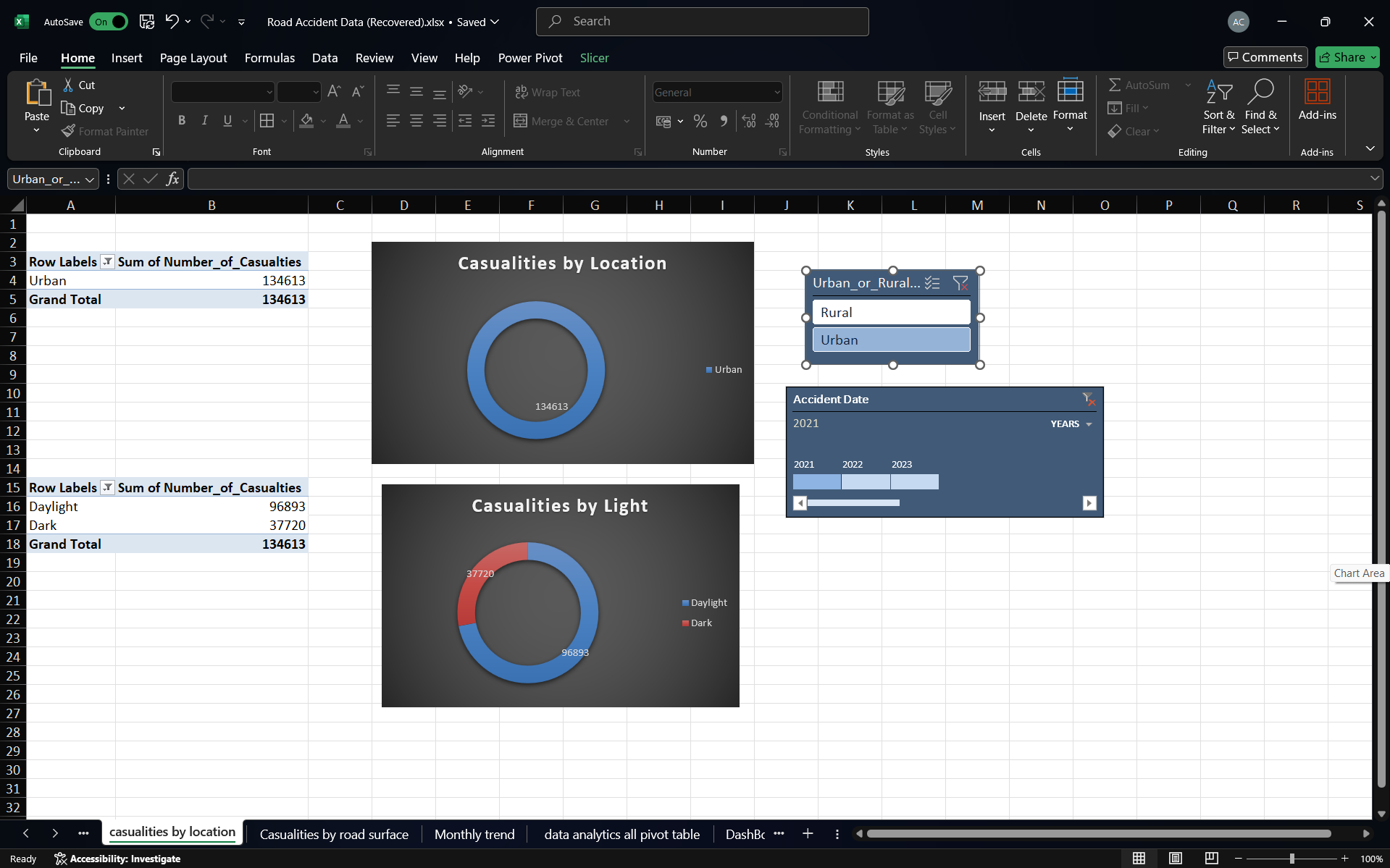
**VII. 💡 Casualties by Light Condition**

|  |  |
| --- | --- |
| Light Condition | Casualties |
| Daylight | 186,200 |
| Darkness | 69,700 |

**Insight:**  
More accidents occur during daylight because that’s when traffic volume is higher. However, accidents in darkness often result in **greater severity** due to reduced visibility and delayed response times.

**Theoretical Note:**

This represents a **confounding variable**—light conditions may not directly cause accidents but influence them through visibility and driver perception.



**4. Dashboard Features**

* **Slicers:**  
  Allow filtering by year, region, or road type. Makes the dashboard **interactive and user-friendly**.
* **Cards and Visual Indicators:**

Conditional formatting (e.g., red for fatalities) makes **at-a-glance reading** effective for decision-makers.

* **Linked Visuals:**

All visuals are interconnected—changing a slicer updates every chart dynamically.

**Theoretical Note:**

The dashboard follows principles of **visual analytics**, combining **data mining, human perception, and design** to extract insights.



**5. Conclusion**

**I. Key Takeaways**

* **Cars and single carriageways** account for the majority of casualties.
* **Urban areas** experience more accidents due to traffic congestion, while rural accidents are more fatal.
* **Dry weather and daylight** are not safe guarantees—more accidents happen during these conditions.
* **Monthly trends** show a spike in October–November, suggesting event-based traffic surges.

**II. Recommendations**

* Install better signage, speed bumps, and lighting on single carriageways.
* Conduct **awareness programs** in urban areas targeting car drivers.
* Encourage use of **advanced driver-assistance systems (ADAS)** in vehicles.
* Initiate **seasonal safety drives** before October and November.

**6. Future Scope**

* Integrating **real-time data** from traffic sensors and IoT devices.
* Using **machine learning algorithms** to predict accident-prone zones.
* **Mapping accidents using GIS** tools to visualize spatial clusters.
* Developing a mobile dashboard for **on-the-go decision-making** by city planners or traffic police.

**7. References**

1. **World Health Organization (WHO).**  
   *Global status report on road safety 2018.*  
   Retrieved from: <https://www.who.int/publications/i/item/9789241565684>
2. **UK Department for Transport.**  
   *Reported road casualties in Great Britain: annual reports and data tables.*  
   Retrieved from: https://www.gov.uk/government/collections/road-accidents-and-safety-statistics
3. **National Highway Traffic Safety Administration (NHTSA).**  
   *Traffic Safety Facts Annual Report Tables.*  
   Retrieved from: https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars
4. **Road Safety Authority (RSA) – Ireland.**  
   *Road collision statistics.*  
   Retrieved from: https://www.rsa.ie/road-safety/statistics
5. **Peden, M. et al. (2004).**  
   *World report on road traffic injury prevention.*  
   World Health Organization. ISBN: 9241562609.
6. **Elvik, R. (2008).**  
   *Road safety management by objectives: effectiveness of Speed Cameras.*  
   *Accident Analysis & Prevention*, 40(1), 349–357. doi:10.1016/j.aap.2007.06.006

### ****8. Author Information****

**Prepared by:** Abhay Chaudhary  
**LinkedIn:** <https://www.linkedin.com/posts/abhay-chaudhary042_exceldashboard-dataanalysis-roadsafety-activity-7311863276483555328-KdV1?utm_source=share&utm_medium=member_desktop&rcm=ACoAAEMNkP4BO7UjkJpurcWDPvOD-WlycvU-Oow>

**GitHub:** <https://github.com/abhaychaudhary18/Interactive-Road-Accident-Analysis-Dashboard-UK-2021-2022->