**WINTER SEMESTER 2024 - 2025**

**CSE3040-EXPLORATORY DATA ANALYSIS**

**PROJECT REVIEW-1**

**A COMPREHENSIVE ANALYSIS**

**ON**

**TRAFFIC ACCIDENTS**

****

**TEAM MEMBERS**

**LAKSHARAA A S (23MIA1053)**

**AMRITHA A (23MIA1067)**

**NITISH RAM J (23MIA1095)**

**GUIDED BY**

**ASNATH VICTY PHAMILA Y**

**SUBMITTED ON**

**25/01/2025**

**SDG:** Good Health and Well being

**SDG NO:** 3

### This goal emphasizes ensuring safety and well-being in urban environments by addressing traffic collisions and their impact on public health. It highlights the need to reduce road traffic injuries and fatalities, identify contributing factors to accidents, and promote safer and more sustainable transport systems. By focusing on data-driven insights and preventive measures, this goal aims to enhance road safety, minimize casualties, and improve overall quality of life in urban communities by 2030.

### **Description:**

SDG 3 aims to ensure healthy lives and promote well-being for people of all ages. This goal emphasizes reducing road traffic injuries and fatalities, improving healthcare systems, and addressing public health challenges to enhance global health outcomes. By focusing on preventive measures, accessible healthcare, and public awareness, it seeks to build safer and healthier communities.

This project uses Python for data analytics to enhance road safety and minimize traffic-related casualties. By leveraging libraries like Pandas, NumPy, and Matplotlib, it analyzes collision data to identify trends, contributing factors, and high-risk areas. The insights gained can be used to implement targeted interventions, optimize urban planning, and develop evidence-based policies for safer transport systems.

Furthermore, SDG 3 encourages the adoption of innovative technologies, data-driven strategies, and collaborative efforts among governments, organizations, and communities to reduce health risks and ensure safety. It emphasizes the importance of creating resilient healthcare and safety frameworks to achieve better well-being for all.

**Cross-linkages with Other SDGs:**

Traffic accidents also relate to other SDGs:

1. **SDG 11: Sustainable Cities and Communities**
   * **Target 11.2**: Ensure access to safe, affordable, and sustainable transport systems for all, improving road safety and prioritizing vulnerable users like pedestrians, cyclists, and public transport users.
2. **SDG 9: Industry, Innovation, and Infrastructure**
   * Focus on building resilient infrastructure and promoting sustainable transportation systems.
3. **SDG 12: Responsible Consumption and Production**
   * Encourages the development of safer and more sustainable vehicles and transportation systems.
4. **SDG 17: Partnerships for the Goals**
   * Promotes global cooperation for implementing road safety measures and sharing best practices.

### **ABSTRACT:**

This review focuses on exploratory data analysis (EDA) of a dataset containing information about traffic collisions in New York City. The dataset includes key attributes such as the date, time, and location of collisions, as well as contributing factors, vehicle types involved, and the number of injuries and fatalities among different groups (pedestrians, cyclists, and motorists). By leveraging Python's data analysis libraries, such as Pandas, NumPy, and Matplotlib, the analysis aims to uncover patterns, trends, and high-risk areas that contribute to accidents.

The primary objectives of this analysis are to:

1. Identify the temporal and spatial distribution of collisions.
2. Analyse the most common contributing factors to accidents.
3. Assess the impact on different road users, including pedestrians, cyclists, and motorists.
4. Provide actionable insights for improving road safety and minimizing injuries and fatalities.

### **DATASET**

**Link:** [Traffic Accidents](https://maven-datasets.s3.amazonaws.com/NYC+Traffic+Accidents/NYC_Collisions.zip)

**Description:** The dataset provides critical insights into traffic collisions in New York City, serving as a resource for analysing progress toward SDG 3. It offers detailed information on accident locations, times, contributing factors, and the impact on various road users, enabling data-driven approaches to enhance road safety and reduce fatalities.

**Key Features:**

1. **Collision Details**:
   * Information on the date, time, and location of collisions, including borough, street name, and cross streets.
   * Geographic coordinates (latitude and longitude) for precise mapping and spatial analysis.
2. **Contributing Factors**:
   * Primary causes of accidents such as driver behaviour, environmental conditions, and mechanical failures.
3. **Impact Metrics**:
   * Number of persons injured and killed, categorized by road user type (pedestrians, cyclists, motorists).
   * Data on the severity of accidents, providing insights into high-impact collisions.
4. **Vehicle Involvement**:
   * Types of vehicles involved in collisions, such as passenger cars, trucks, and buses.
5. **Spatial and Temporal Data**:
   * Data grouped by boroughs and time periods, allowing for regional and temporal trend analysis.

This dataset is a valuable resource for understanding traffic safety dynamics, identifying high-risk areas, and developing data-driven interventions to reduce accidents and fatalities. It supports efforts aligned with SDG 3, focusing on improving public health and safety.

### **References:**

1. **Dataset:**
   * Traffic Accidents Dataset. [Traffic Accidents](https://maven-datasets.s3.amazonaws.com/NYC+Traffic+Accidents/NYC_Collisions.zip)
2. **United Nations Sustainable Development Goal:**
   * United Nations. "Sustainable Development Goal 3: Good Health and Well-being."  
     <https://sdgs.un.org/goals/goal3>
3. **Python Libraries:**
   * Matplotlib Documentation.  
     <https://matplotlib.org/stable/contents.html>
   * Pytorch Documentation.

<https://pytorch.org/docs/stable/index.html>

* + Plotly Documentation.

<https://plotly.com/python/getting-started/>

* + Sci-Kit Learn Documentation.

<https://scikit-learn.org/0.21/documentation.html>

* + Numpy Documentation.

<https://numpy.org/devdocs/>

* + Pandas Documentation.

<https://pandas.pydata.org/docs/>