Traffic Accident Analysis Project

Data Analyst: Agustin A. Zavagnini

Purpose:

To analyze historical traffic accident data to identify the **primary conditions, temporal factors, and causal elements** contributing to crash occurrence and injury severity. The ultimate goal is to generate **actionable insights** for transportation safety officials, enabling evidence-based decision-making to mitigate future traffic risks and fatalities. A baseline predictive model will be used to quantify the relative weight of identified risk factors.

Scope / Major Project Activities:

Activity	Description
1. Data Understanding	Examine dataset structure, variable types, summary statistics, missing values, and distributions across temporal, environmental, and causal features. Prepare a clean data dictionary and define initial hypotheses.
2. Data Cleaning & Preparation	Standardize categories (e.g., control devices, weather conditions), handle missing values, convert data types (especially date and time fields to derive temporal features like Hour of Day), and ensure dataset consistency for analysis.
3. Exploratory Data Analysis (EDA)	Conduct univariate, bivariate, and multivariate analysis to explore relationships between key factors: **Temporal trends** (peak hours/days), **Environmental conditions** (lighting, weather, road surface), and **Infrastructure** (traffic control, roadway type).
4. Risk & Accident Insights	Identify characteristics strongly associated with **high severity outcomes** (fatal or incapacitating injuries). This includes correlating primary causes, road defects, type of crash (e.g., head-on vs. rear-end), and intersection involvement with injury metrics.

5. Recommendations & Conclusions	Provide data-driven conclusions to improve road safety policy, traffic control device deployment, and public safety campaigns. Highlight the most **high-impact variables** (e.g., most dangerous hours, specific critical causes) based on the analysis and the model.
6. Visualization Assets	Create data visualizations including temporal trend lines, causal factor breakdowns, severity segmentation, and heatmaps. Develop an interactive Tableau or PowerBI dashboard summarizing key insights for executive review.

This project does not include:

- **Advanced Model Optimization:** Techniques such as GridSearch, complex feature engineering, or fine-tuning of hyperparameters are out of scope.
- **Real-Time Data Integration:** The analysis is strictly historical; integration with live weather APIs or real-time traffic flow data is excluded.
- **Geospatial Mapping:** Precise geospatial mapping (lat/long) is excluded unless specific coordinates are derived or provided; analysis will focus on roadway types.

Deliverables:

Deliverable	Description/Details
1. Jupyter Notebook	A fully documented notebook containing the data cleaning process, visual statistical analysis, and conclusions
2. Executive Summary Report	A document (PDF/Markdown) outlining "Key Insights": primary accident factors, hazardous conditions, and safety recommendations.
3. Analytical Dashboard	A visual dashboard (Tableau) allowing for the interactive exploration of accidents by cause, weather, and severity.