

Road Traffic Accident Analysis

**COURSE: DATA 200 –
Applied Statistical Analysis**

**Group Mememebers: Prasanna,
Ankita, Sambriddhi**



Introduction

- Road traffic accidents are a major cause of injuries, deaths, and economic loss worldwide.
- Accidents occur due to multiple factors such as time of day, weather conditions, road conditions, and vehicle type.
- Large amounts of road accident data are available, but they need proper analysis to extract useful insights.



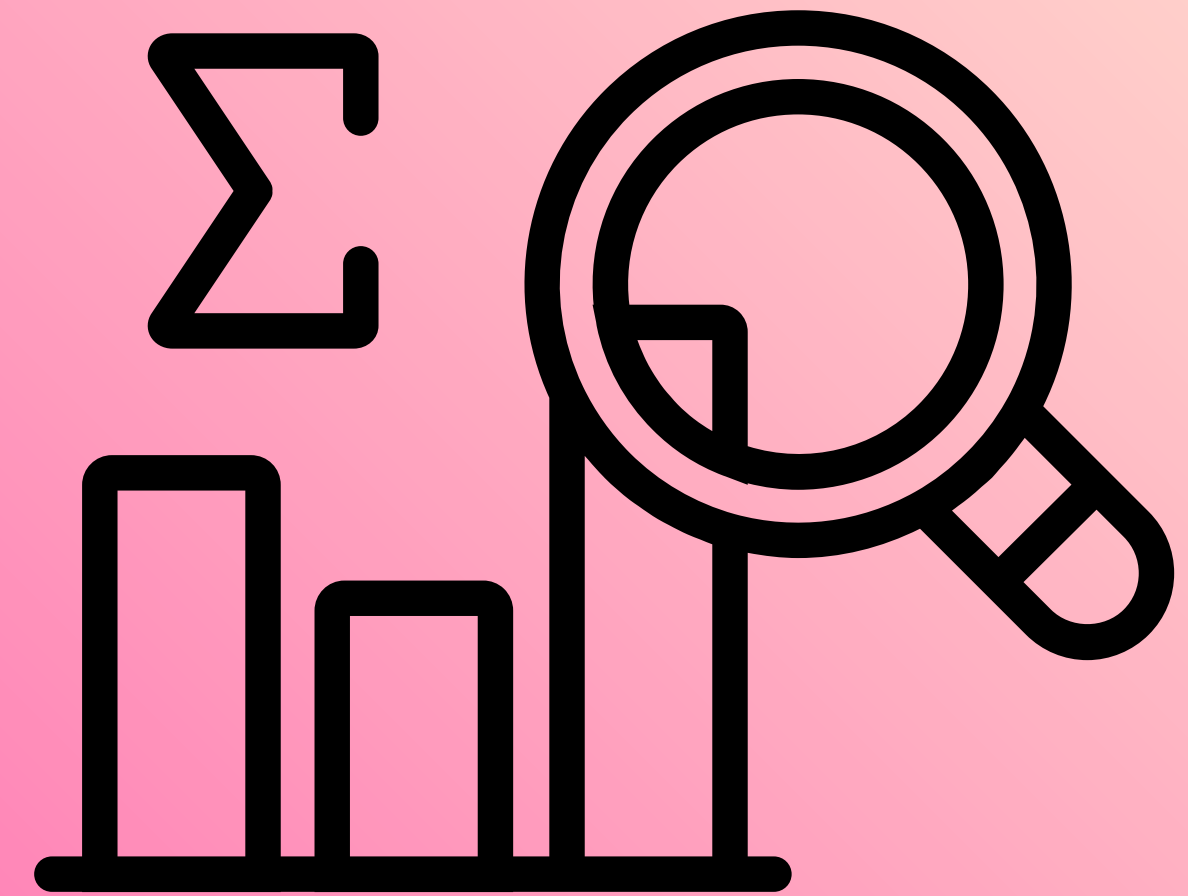
Problem Statement

- Road accidents are increasing, but the key contributing factors are not clearly identified.
- Accident data is often underutilized and not properly analyzed.
- There is limited understanding of how factors like time, weather, and road conditions affect accident severity.
- Without data-driven insights, it is difficult to design effective road safety measures.
- There is a need to analyze road accident data to identify high-risk patterns and relationships.



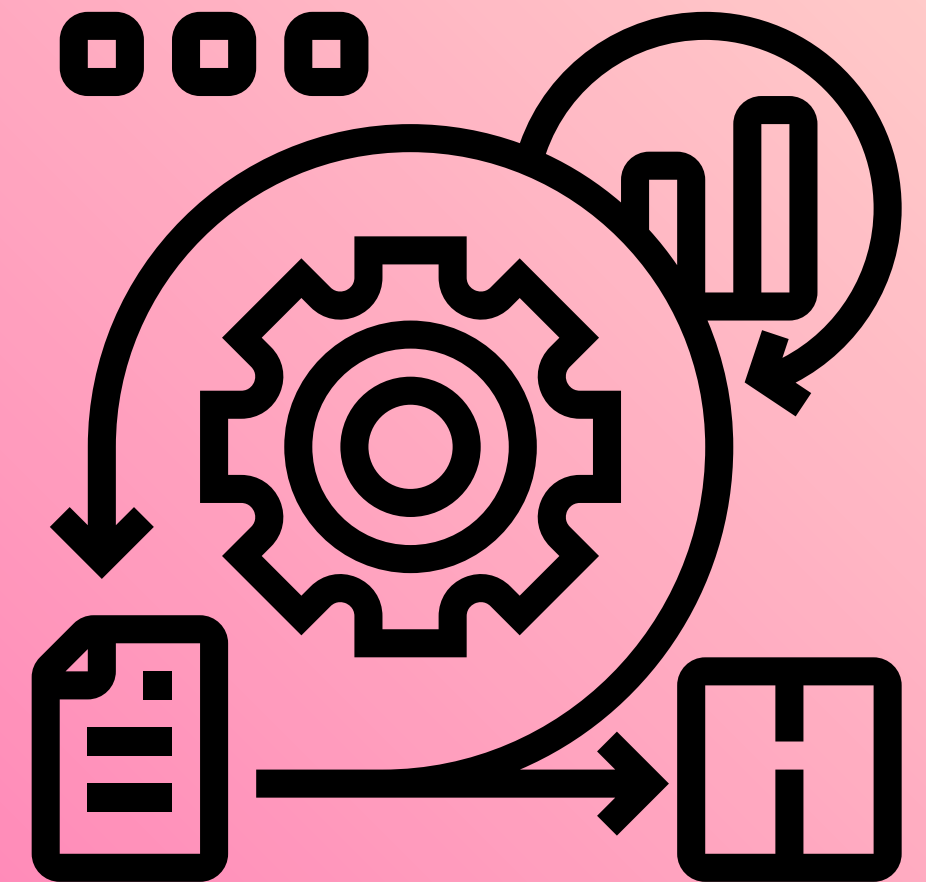
Selected Statistical Models

- **Logistic Regression:** Used to analyze accident severity (severe vs. non-severe).
- **Linear Regression:** Used to study the relationship between number of vehicles and casualties.
- **ANOVA (Analysis of Variance):** Used to compare accident severity across different groups (weather, road condition, time).



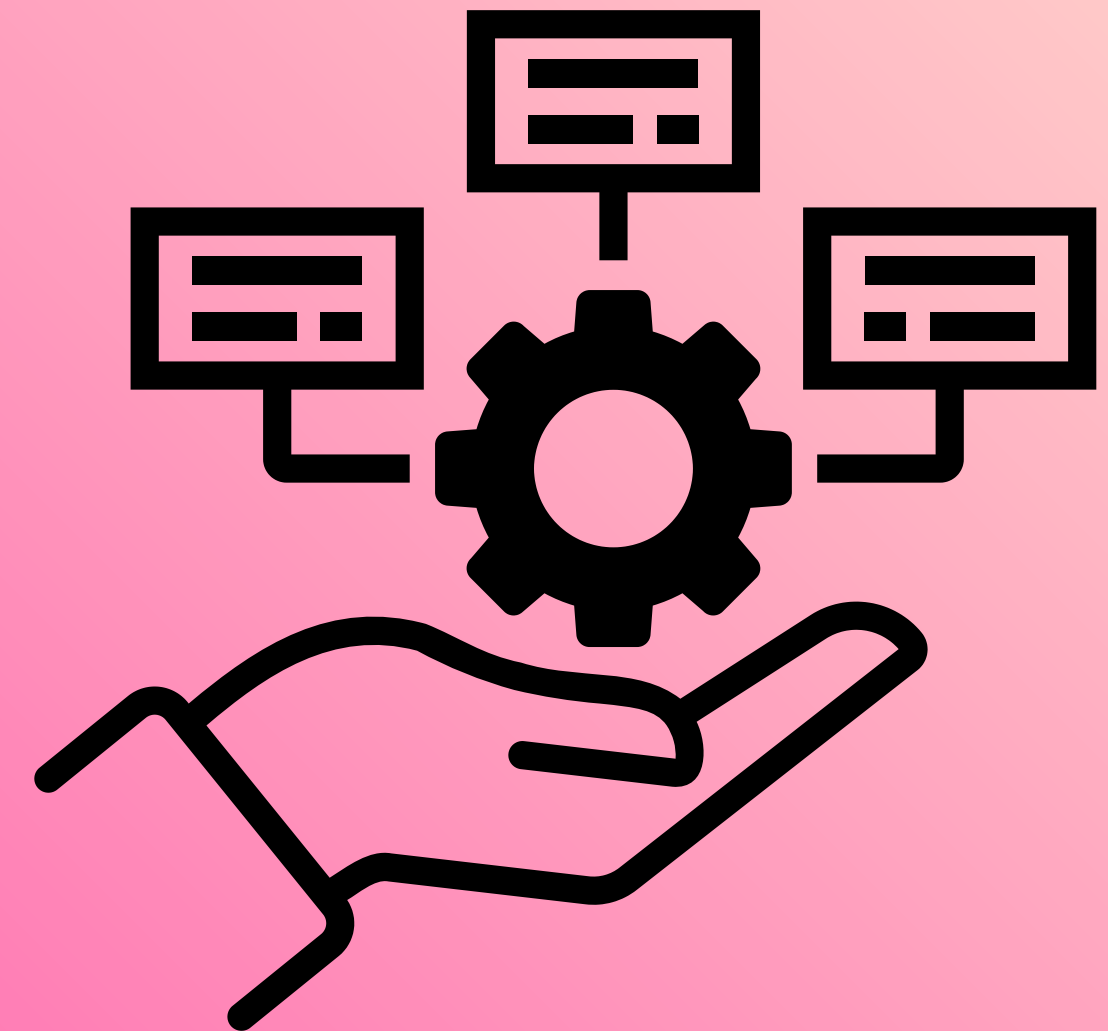
Justification for Model Choice

- Accident severity is a categorical variable, making Logistic Regression appropriate.
- Casualty count is a numerical variable, suitable for Linear Regression.
- Road and weather conditions are categorical groups, best analyzed using ANOVA.
- These models are simple, interpretable, and widely used in accident analysis studies.



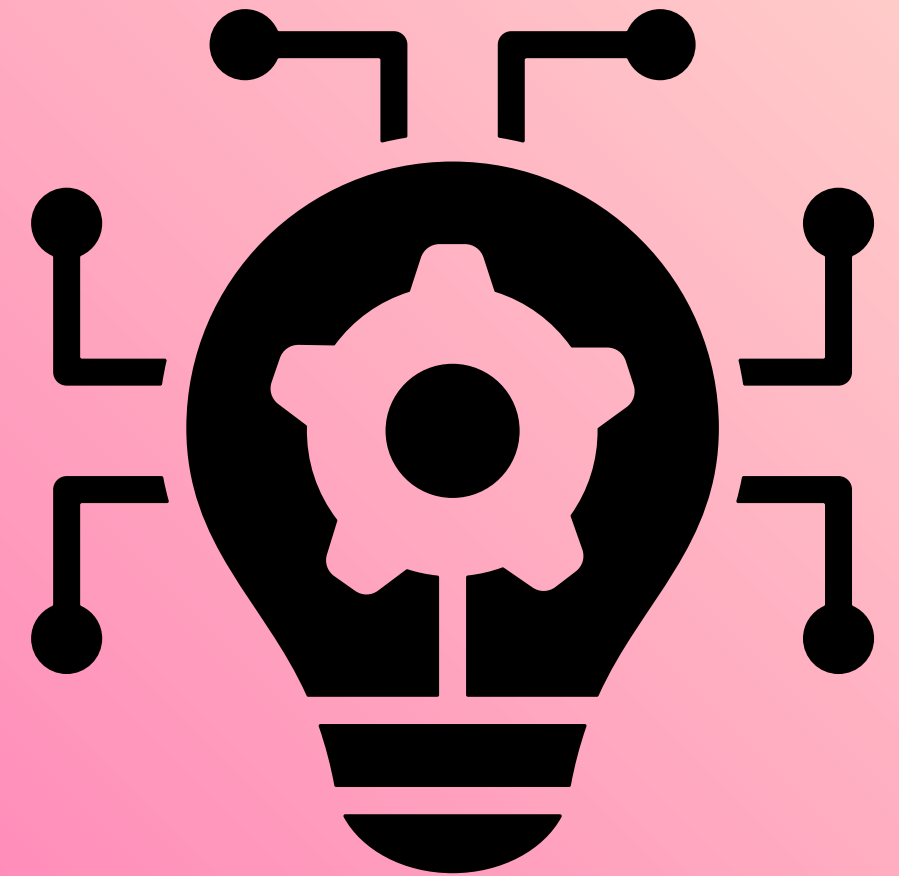
Feature Selection Rationale

- **Features were selected based on:**
 - Exploratory Data Analysis (EDA) results.
 - Literature review findings.
 - Data availability and low missing values.
- **Selected Features:**
 - Time of day.
 - Weather condition.
 - Road condition.
 - Number of vehicles involved.
- **Target Variables:**
 - Accident severity.
 - Number of casualties.



Develop Hypotheses

- H_0 : Time of day has no significant effect on accident severity.
- H_1 : Time of day has a significant effect on accident severity.
- H_0 : Weather condition does not affect accident severity.
- H_1 : Weather condition affects accident severity.
- H_0 : Number of vehicles involved has no relationship with casualties.
- H_1 : Number of vehicles involved is positively related to casualties.



*Thank
you!*