NATIONAL COLLISION DATABASE ANALYSIS

Capstone Project
Group 1

WHY THIS ANALYSIS ?

High Rate of Motor Vehicle Collisions

Economic loss, human loss, trauma

Need for Proactive and Reactive Measures

Prevention and efficient emergency response

Implementation of Data-Driven Strategies

Use of NCDB for predictive modeling

Improved Road Safety and Emergency Preparedness

Reduced accidents, better emergency response, Improved Proactive Measures Improved Reactive Measure

Problem

Action

Solution

Output

WHY THIS ANALYSIS ?

Low Profitability, Low Revenue Growth

Over Utilisation of Assets

Implement profitabiility analysis

Excute Cost Optimisation

Process Reengineering

Cross-training Employees

Increase in profitability and revenue

Better cost optimisation, More sustainable asset utilisation

Problem

Action

Solution

Output

PHASES OF IMPLEMENTATION

1

Evaluate current procurement processes and machinery.

Audit for inefficiencies and redundant tasks.

Assess benefits of bulk purchasing and quality machinery investment.

2

Strategise for bulk purchasing, including vendor negotiations.

Design a plan for process reengineering.

Acquire high quality machinery, considering budget and ROI.

3

Implement bulk purchasing agreements and revise ordering procedures.

Execute Process reengineering and staff Cross-training.

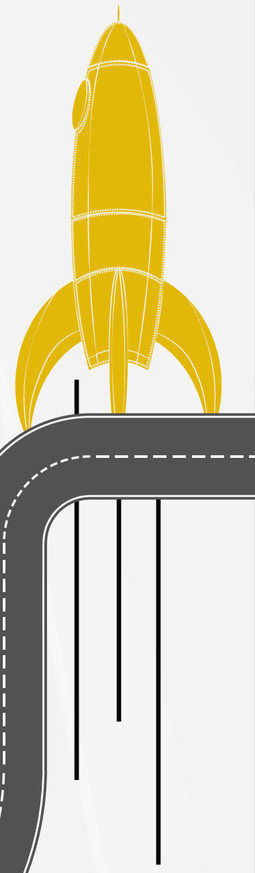
Install new machinery while minimising operational disruption.

4

Review impact of changes on costs and efficiency.

Strategise based on feedback & performance data.

Regularly reassess processes for continuous improvement and future investments.



Research Questions

Identify patterns and risk factors in road traffic collisions, thereby enabling the implementation of targeted preventive measures.

To enhance the efficiency and effectiveness of first responders, including paramedics, firefighters, and police, through predictive modeling that anticipates high-risk scenarios and optimizes resource deployment.

Manage healthcare resources, ensuring immediate and adequate medical attention for accident victims by prioritizing emergency care and optimizing the use of medical facilities during peak accident times.

Analysing the road surface, we could give awarness to the road authorities about the repair.

DATASET & METHODOLOGY

Objective & Data Types:

Analyze collision severity using a dataset of 6.9 million records.

Quantitative data for robust numerical analysis.

Data Source & Coverage:

Sourced from Canada Open Data Portal (Open.Canada.ca).

Spans over a significant period for comprehensive trend analysis.

Data Attributes & Integrity:

Key variables: vehicle type, safety precautions, time, environmental conditions, involved parties.

High-quality data with minimal missing/inconsistent entries.

Data Preparation:

Cleaning and preprocessing using Python & Excel.

Exploratory Data Analysis:

Conducted with Python to analyze and summarize.

Data Visualization:

Insights presented through Tableau & Python.

Interpretation & Validation:

Pattern identification with graphs.

Ensuring analysis aligns with proposed objectives.

Final Results:

Tableau Dashboard, PowerPoint Presentation.

EVERY LIFE MATTERS

Between 2000 to 2019

6,913,204

Accidents Reported

Fatality 46,735

Injured 3,617,076

Analyzing the Key Trend

Yearly Trend: A **decreasing trend in accidents** from 2000 to 2019 may reflect improved safety measures or reporting changes.

Monthly Peaks: Higher accident rates in **January, July, and October** may be due to seasonal driving conditions and travel habits.

Weekly Highs: More accidents on **Thursdays and Fridays** could be linked to increased travel and changes in driving behavior near weekends.

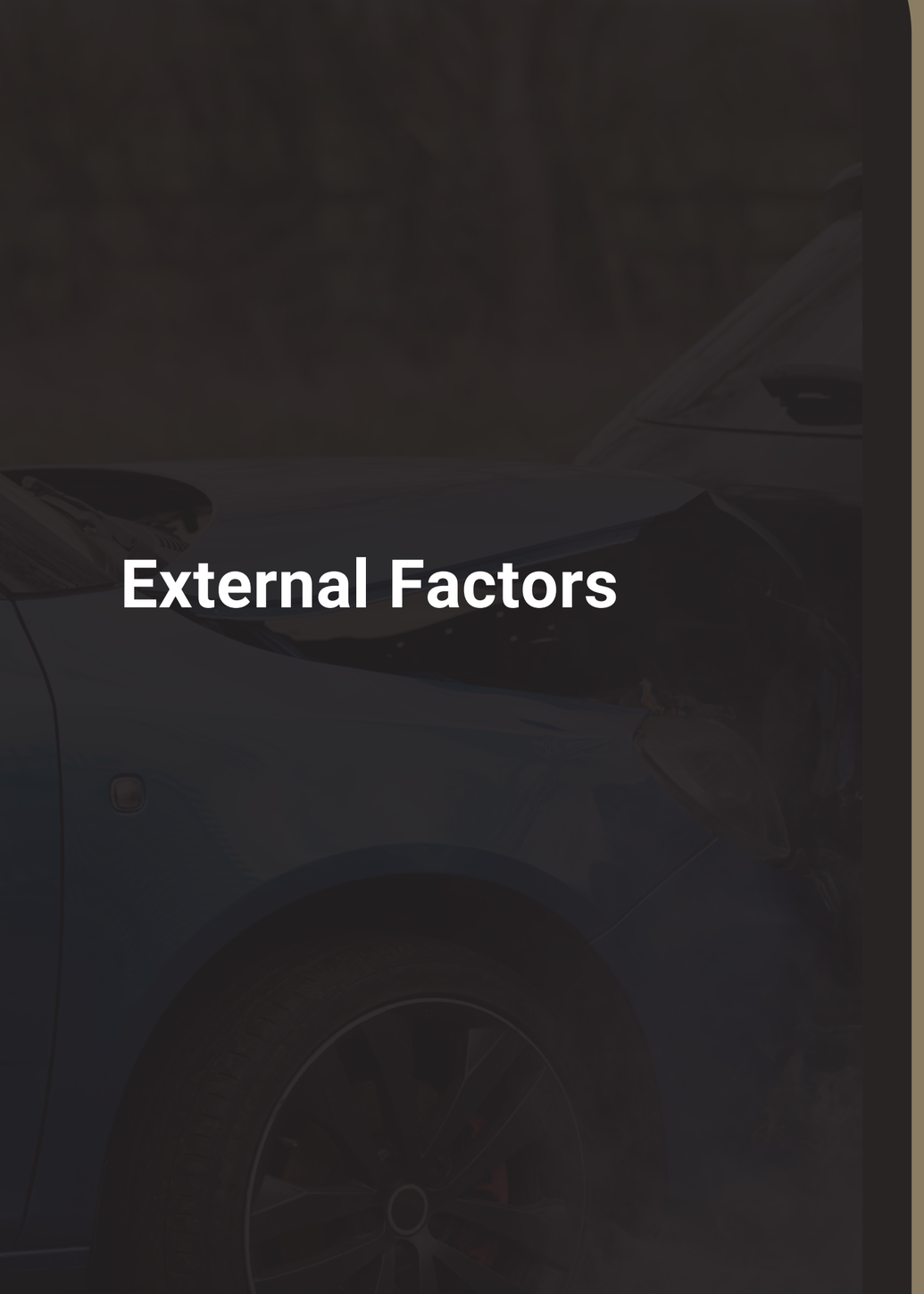
Hourly Spike: A rise in accidents from **3 PM to 8 PM** coincides with the evening rush hour and diminishing daylight.

January, July, and October
Thursdays and Fridays

3 PM to 8 PM

21 Yr to 30 Yrs - 22%

Male 56.07%



External Factors

The majority of accidents occur under clear and sunny conditions on dry, straight, and level roads at intersections.

Accidents predominantly happen at intersections on straight roads without traffic controls.

Areas without traffic controls, especially non-intersections, experience the highest number of accidents, with a significant spike at intersections featuring parking and on curved roads.

A lack of traffic control is commonly associated with a higher incidence of accidents.

The risk of accidents escalates markedly during freezing rain when traffic controls are not in place.

Internal Factors

The driver's seat is the most common position for both non-fatal injuries and fatalities in accidents.

The unsafe positions, with the highest numbers of fatalities, are the driver’s, front row and second-row right position and passengers in motorcycle.

Most of the accidents are due to the Light Duty Vehicles and the involvement of these vehicles in accidents are decreasing over the years,

The vehicle model years 2000-2003 had the highest number of accidents compared to other years in the past decade.

Recommendations

Traffic Control Setup:

Implement traffic controls at key risk spots like non-intersections, curved roads,intersections with parking.

For Healthcare Strategy:

Run campaigns for risk periods and advise healthcare to strategize for peak times.

For Police: Seasonal Hazard Education:

Boost education on driving risks during high-accident months.

Road Authorities:

Work with authorities to fix road issues and enhance traffic management.

For Users - seat Safety Instruction:

Educate on optimal seat positioning and seat belt use, focusing on drivers and front-row passengers.

Future Work

Predictive Modeling Development: K Means Clustering for Identifying High Risk Scenarios

ML models can predict high-risk scenarios, allowing for proactive measures to prevent accidents and better prepare emergency services. K Means Clustering

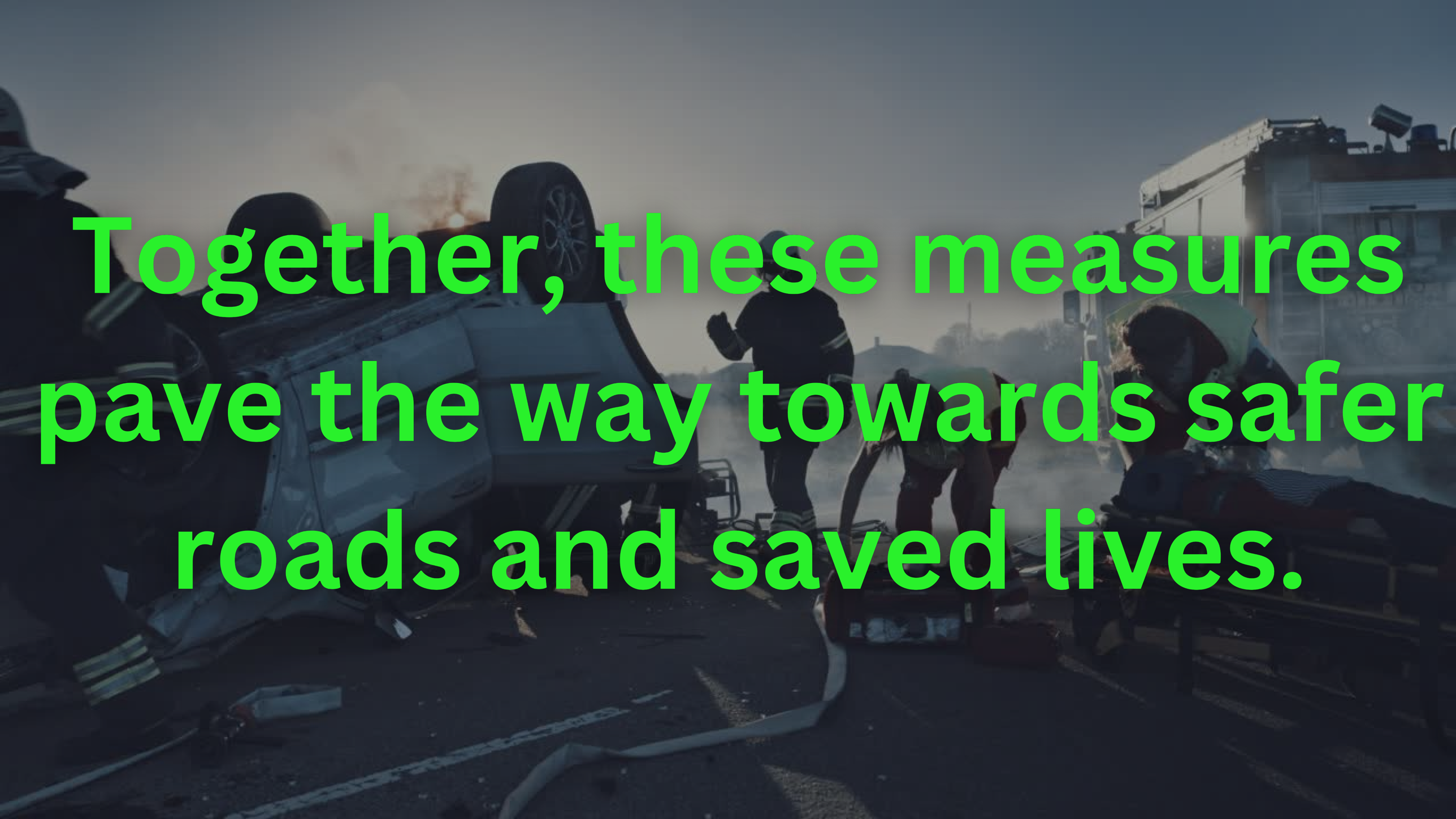
Time Series Analysis for Trend Prediction and Prevention Strategies

K-Means Clustering for Identifying High-Risk Scenarios

Random Forests for Insightful Risk Factor Analysis

Continual Data Analysis: Update the model every 6 months, as data collection and analysis of accident reports to refine predictive models, ensuring they remain accurate and reflective of current trends.

Collaborative Research Initiatives:
Foster partnerships with academic and research institutions to explore advanced machine learning techniques for better understanding and mitigating road traffic collision risks.
Innovative Traffic Safety Technologies:


Together, these measures
pave the way towards safer
roads and saved lives.