



A MINI PROJECT REPORT ON
**“ Analysis of Road Accident Trends in India
(2021-2022)”**

FOR

Mini Project

***Bachelors of Computer Application in Artificial Intelligence
& Machine Learning(BCA - AIML)***

Year 2024-2025

Ajeenkya DY Patil University, Pune

-Submitted By-

Mr. Moaiz Suzain

Under the guidance of

Prof. Vivek More



Ajeenkya DY Patil University

D Y Patil Knowledge City,
Charholi Bk. Via Lohegaon,
Pune - 412105
Maharashtra (India)

Date: / / 2025

CERTIFICATE

This is to certified that Mr. Moaiz Suzain
A student of **(BCA - AIML)**Sem-IV URN No: 2023-B-02092005A has Successfully Completed the Dashboard Report On

“Analysis of Road Accident Trends in India (2021-2022)” “

As per the requirement of
Ajeenkya DY Patil University, Pune was carried out under my
supervision.

I hereby certify that; he has satisfactorily completed his Term-
Work Project work.

Place: - Pune

Examiner

INDEX		
		Page No.
ABSTRACT		4
CHAPTER – 1	INTRODUCTION	4
CHAPTER – 2	METHODOLOGY	4
CHAPTER – 3	ANALYSIS AND INTERPRETATION	4
CHAPTER – 4	RESULTS AND VISUALIZATIONS	5
CHAPTER – 5	CONCLUSIONS, SUMMARY AND RECOMMENDATIONS	8
BIBLIOGRAPHY		8
APPENDIX AND CODE REVIEW		8

Abstract

This project presents a data-driven analysis of road accident trends across Indian states for the years 2021 and 2022. Using publicly available data and visualization tools, we highlight the states with the highest and lowest accident counts, observe shifts over the two years, and draw insights into accident-prone regions.

Chapter 1: Introduction

Road safety remains a critical public health challenge in India. Each year, thousands perish or are severely injured due to traffic collisions. This project aims to analyse state-wise accident data from 2021 and 2022 to detect patterns and changes, particularly in high-incident states like Tamil Nadu, Madhya Pradesh, and Uttar Pradesh.

Chapter 2: Methodology

Data was sourced from official government records and cleaned using Python. Libraries such as Pandas, Matplotlib, and Seaborn were used for data analysis and visualization. Key steps included:

- Handling missing data
- Standardizing formats
- Visualizing trends using bar and line charts to compare annual figures

Chapter 3: Analysis and Interpretation

The visual analysis revealed the following:

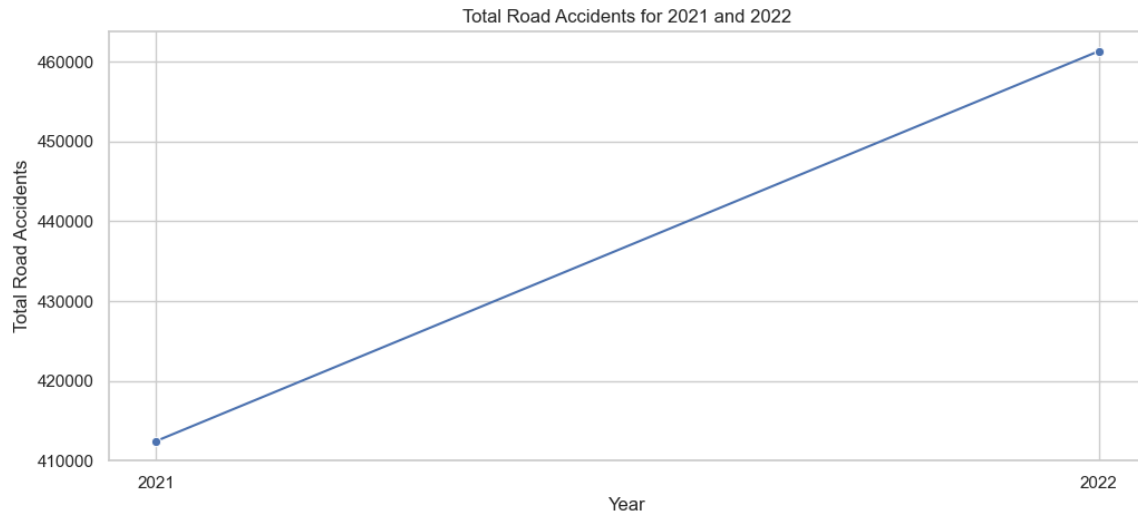
- **Tamil Nadu** recorded the highest number of accidents in both years, with a slight increase in 2022.
- **Madhya Pradesh** and **Uttar Pradesh** followed closely in accident counts.
- **Kerala** and **Karnataka** also featured in the top five states, showing consistent patterns year over year.
- States like **Mizoram** and **Nagaland** had significantly fewer incidents, suggesting better road safety or lower vehicle density.
- A few states, such as **Punjab**, saw a noticeable decrease in accidents in 2022.

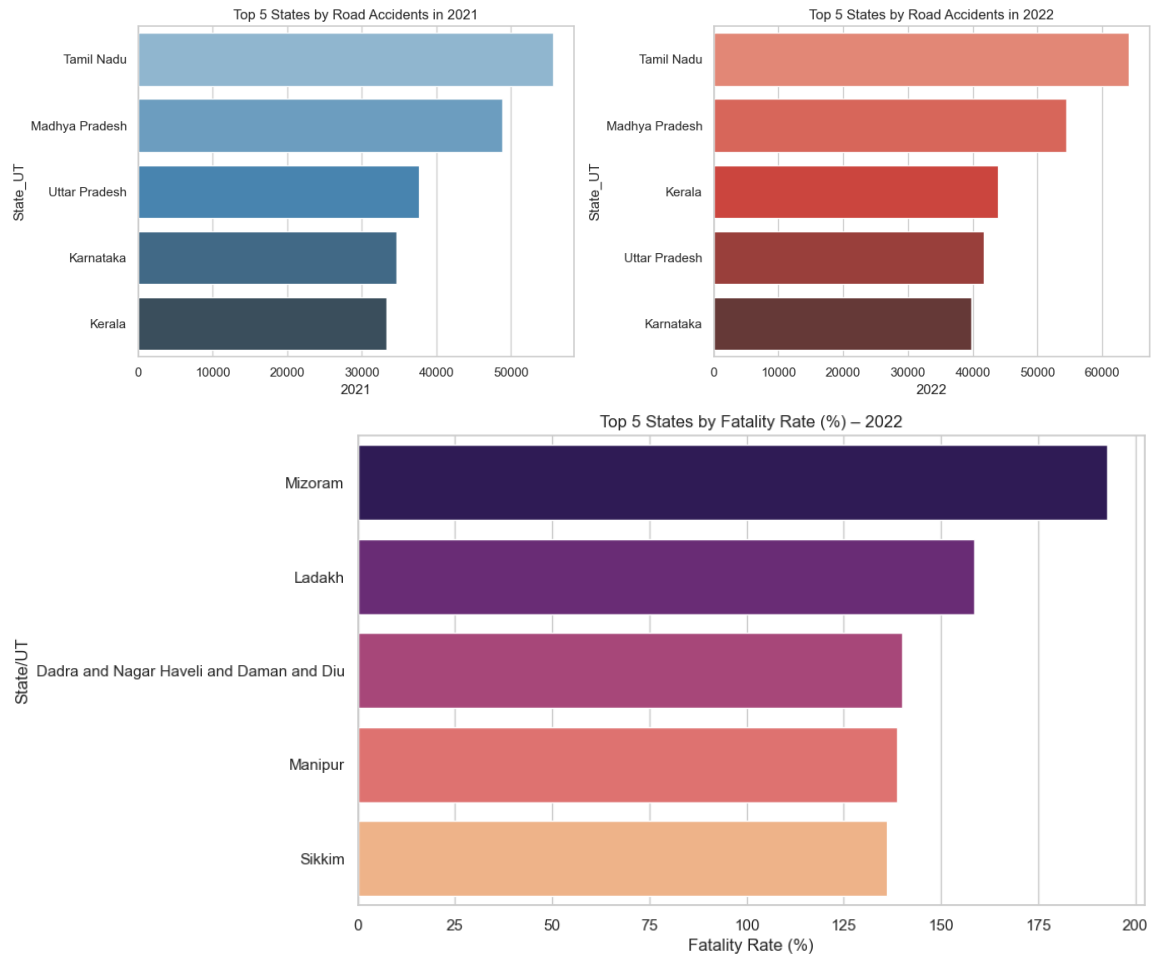
Chapter 4: Results and Visualizations

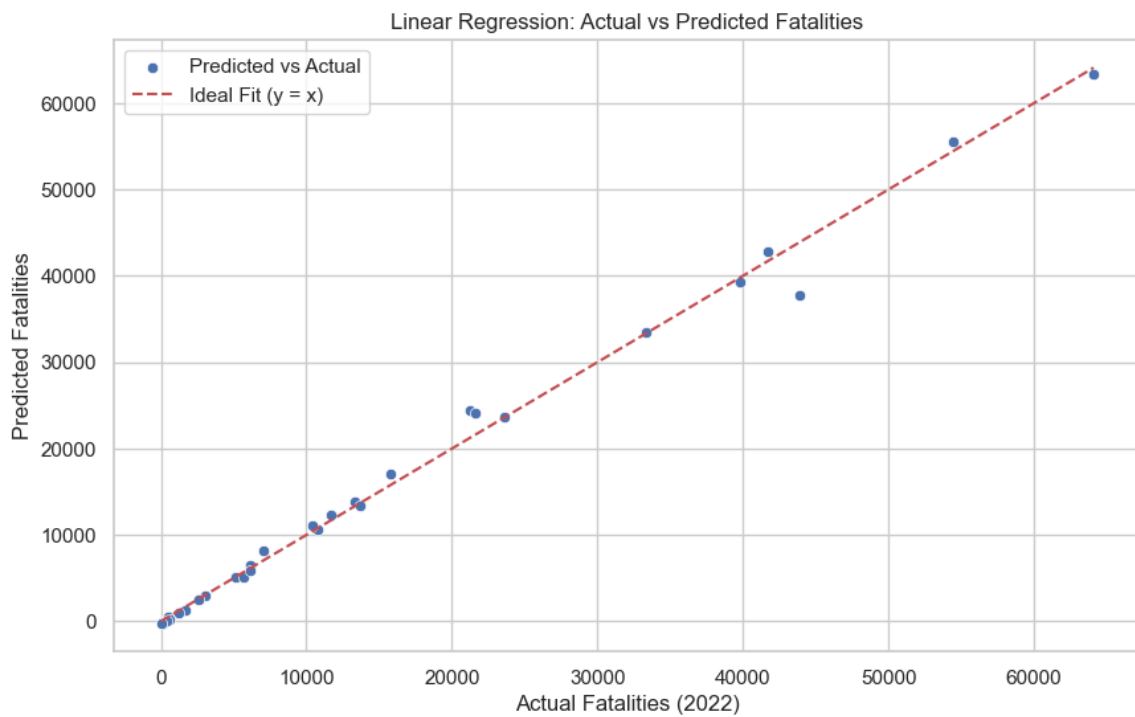
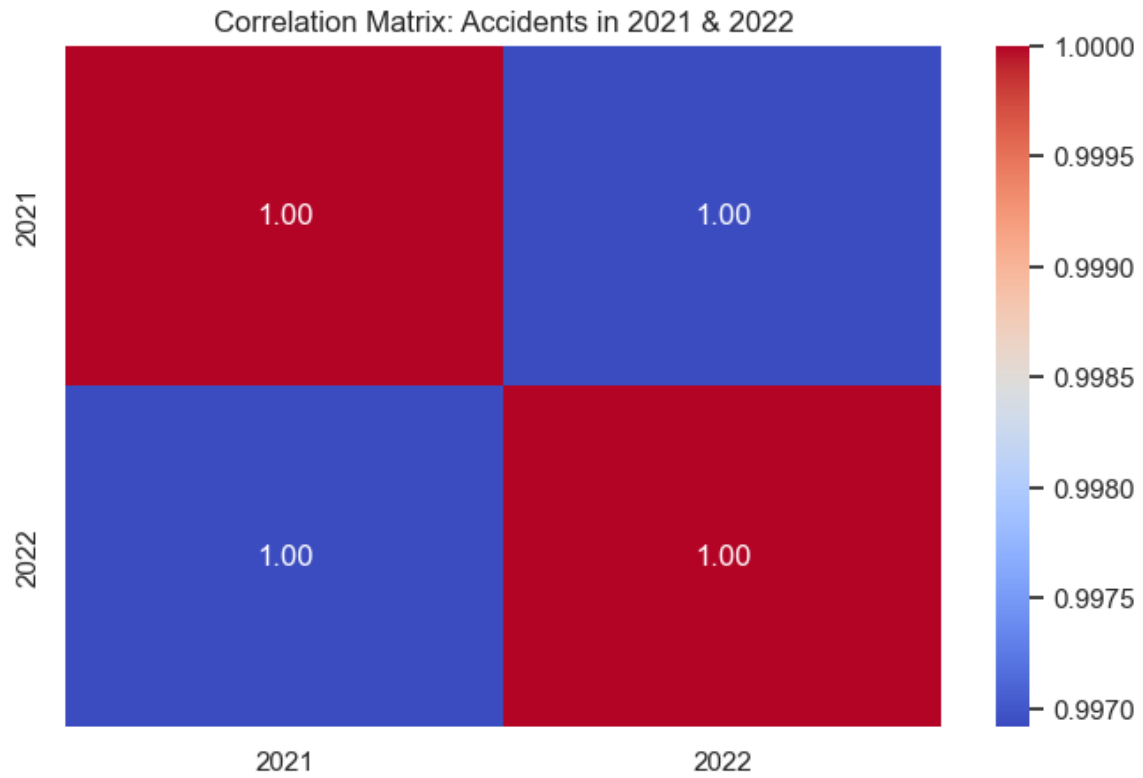
Key Insights from the Charts:

- The **bar chart** comparing 2021 and 2022 indicates a **slight overall increase** in accidents.
- The **line plot** highlights a **stable or rising trend** in the top accident-prone states.
- States with low accident numbers remained largely unchanged across both years.

These findings suggest that while overall accident rates are stable, specific states require focused safety interventions.







Chapter 5: Conclusion and Future Scope

This study underscores the need for stronger traffic management and public safety campaigns in states with persistently high accident rates. Future work could include:

- Analysis by time of day, type of vehicle, or road condition
- Integration of machine learning for accident prediction
- Use of GIS tools for spatial visualization

Bibliography

- Ministry of Road Transport & Highways (morth.nic.in)
- Python Libraries: Pandas, Matplotlib, Seaborn
- Jupyter Notebook environment

Appendix - Code Overview

Python scripts used in this project are structured for clarity and modular analysis:

- Data ingestion and cleaning
- Year-wise and state-wise visualizations
- Annotated charts for ease of interpretation

Code Snippets:

Importing Libraries

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
from sklearn.preprocessing import StandardScaler
```

Loading Datasets

```
df = pd.read_csv("RS_Session_266_AU_2647_1.csv")

df.columns = df.columns.str.strip().str.replace(" ", "_")

df.head()
```

🔗 Open 'df' in Data Wrangler

Total Accidents in 2021 and 2022

```
yearly_stats = df[["2021", "2022"]].sum().reset_index()

yearly_stats.columns = ['Year', 'Total_Road_Accidents']

plt.figure(figsize=(12, 5))
sns.lineplot(data=yearly_stats, x="Year", y="Total_Road_Accidents", marker="o")
plt.title("Total Road Accidents for 2021 and 2022")
plt.ylabel("Total Road Accidents")
plt.xlabel("Year")
plt.show()
```

Top 5 States by Accidents in 2021 and 2022

```
top_accidents_2021 = df[["State_UT", "2021"]].sort_values(by="2021", ascending=False).head(5)
top_accidents_2022 = df[["State_UT", "2022"]].sort_values(by="2022", ascending=False).head(5)

plt.figure(figsize=(14, 5))

plt.subplot(1, 2, 1)
sns.barplot(data=top_accidents_2021, y="State_UT", x="2021", palette="Blues_d")
plt.title("Top 5 States by Road Accidents in 2021")

plt.subplot(1, 2, 2)
sns.barplot(data=top_accidents_2022, y="State_UT", x="2022", palette="Reds_d")
plt.title("Top 5 States by Road Accidents in 2022")

plt.tight_layout()
plt.show()
```

Actual vs Predicted Fatalities

```
X = df[["2021"]]
y = df[["2022"]]

model = LinearRegression()
model.fit(X, y)

y_pred = model.predict(X)

df["Predicted_Fatalities"] = y_pred

plt.figure(figsize=(10, 6))
sns.scatterplot(x=y, y=y_pred, label="Predicted vs Actual")
plt.plot([y.min(), y.max()], [y.min(), y.max()], '--r', label="Ideal Fit (y = x)")
plt.xlabel("Actual Fatalities (2022)")
plt.ylabel("Predicted Fatalities")
plt.title("Linear Regression: Actual vs Predicted Fatalities")
plt.legend(loc="upper left")
plt.grid(True)
plt.show()

r2 = r2_score(y, y_pred)
print(f"R² Score: {r2:.2f}")
```