**Project Semester January–April 2025**

**DATA SCIENCE MINOR PROJECT REPORT**

**ON**

**Traffic Congestion Analysis in Indian Cities**

**DATA SCIENCE TOOLBOX: PYTHON PROGRAMMING**

**COURSE CODE: INT375**

1. **TECH COMPUTER SCIENCE AND ENGINEERING**

****

**LOVELY PROFESSIONAL UNIVERSITY**

**PHAGWARA, PUNJAB**

**PROJECT SUBMITTED BY:**

**Jasmeen khatun (12319627)**

**Section: K23EH**

**Roll No.:20**

**PROJECT SUBMITTED TO:**

**Madhu Bala (31770)**

**DECLARATION**

I, [Jasmeen khatun], student of B.Tech – Computer Science and Engineering (Section K23ER) at Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report titled:

**“Traffic Congestion Analysis in Indian Cities”**

is based on my own intensive work and is genuine. The content of this report has not been submitted to any other university or institution for the award of any degree or diploma.

Date: [April 2025]  
Registration No.: [12319627]  
Name: [Jasmeen khatun]

**CERTIFICATE**

This is to certify that [Your Name], bearing Registration No. [Your Registration No.], has successfully completed the INT375 – Python Programming project titled:

**“Traffic Congestion Analysis in Indian Cities”**

under my guidance and supervision. To the best of my knowledge, the present work is the result of original development, effort, and study. This project has been carried out as a part of the curriculum prescribed by Lovely Professional University, Phagwara for the Project Semester January–April 2025.

**Madhu Bala (31770)**

**ACKNOWLEDGEMENT**

I sincerely thank my faculty guide, **[Supervisor's Name]**, for their guidance and support throughout this project. I also express my gratitude to the Department of Computer Science and Engineering, Lovely Professional University, for providing the necessary resources and environment.

**[Jasmeen khatun]**  
Reg. No.: [Jasmeen khatun]

1. **INTRODUCTION:**

Traffic congestion is a major challenge in rapidly urbanizing Indian cities. With population growth and increased vehicle ownership, urban roads are often overloaded, resulting in lost time, higher fuel consumption, and pollution. This project aims to analyze traffic congestion patterns using synthetic yet realistic traffic data.

Through Exploratory Data Analysis (EDA), we identify peak congestion hours, the most affected cities, and the impact of external factors such as weather. This analysis provides insights that can help traffic authorities improve infrastructure and commuters plan better travel routes.

1. **SOURCE OF DATASET:**

**India\_Traffic\_Congestion\_Data.xlsx**

The traffic congestion dataset used in this project was generated using realistic city-wise traffic data patterns. The dataset includes the following attributes:

* Date
* City
* Time of Day (Morning, Afternoon, Evening, Night)
* Traffic Index (0-100)
* Average Speed (km/h)
* Delay Time (mins)
* Weather Condition

**EDA PROCESS**

The data analysis involved the following steps:

* Data loading and validation using pandas
* Converting and formatting time-based fields
* Calculating averages, group-wise summaries, and distribution counts
* Visualizing patterns with matplotlib.pyplot and seaborn
* Drawing insights from plots and correlations

**ANALYSIS ON DATASET (5 ANALYSES)**

**1. ANALYSIS ON DATASET (5 ANALYSES)**

**1. Peak Congestion Hours**

* Average traffic index grouped by time of day.
* **Result:** **Evening** hours showed the **highest congestion**, followed by **Morning**.

**2. Most Congested Cities**

* Computed mean traffic index city-wise.
* **Result:** **Delhi, Mumbai, and Bangalore** were consistently among the most congested cities.

**3. Delay Patterns**

* Compared traffic index with delay time using scatter plot.
* **Result:** A clear **positive correlation** exists between high congestion and travel delays.

**4. Weather Impact on Traffic**

* Analyzed average traffic index by weather condition.
* **Result:** **Rainy and Foggy** conditions showed **higher congestion levels**.

**5. Congestion Trends Over Time**

* Line plot of daily average traffic index.
* **Result:** Seasonal peaks were visible; overall congestion fluctuates by time and conditions.

STEPS:

**1: Find Peak Congestion Hours – Identify the time slots (Morning,**

**2: Afternoon, Evening, Night) with the highest traffic index.**

**3: Identify Most Congested Cities – Rank cities based on their average traffic index.**

**4: Analyze Delay Patterns – Find the correlation between traffic congestion and travel delay time.**

**5: Weather Impact on Traffic – Determine how weather conditions (Rainy, Foggy, etc.) influence congestion and delays.**

# Step 1: Import Libraries and Load Dataset

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# Load the traffic dataset

traffic\_df = pd.read\_excel("India\_Traffic\_Congestion\_Data.xlsx")

# Step 2: Convert Date column to datetime

traffic\_df["Date"] = pd.to\_datetime(traffic\_df["Date"])

# Display initial info

print(traffic\_df.info())

print(traffic\_df.head())

# Step 3: Peak Congestion Hours

plt.figure(figsize=(8,5))

time\_congestion = traffic\_df.groupby("Time of Day")["Traffic Index"].mean().sort\_values(ascending=False)

sns.barplot(x=time\_congestion.index, y=time\_congestion.values, palette="Reds")

plt.title("Average Traffic Index by Time of Day")

plt.xlabel("Time of Day")

plt.ylabel("Traffic Index")

plt.tight\_layout()

plt.show()

# Step 4: Most Congested Cities

plt.figure(figsize=(10,6))

city\_congestion = traffic\_df.groupby("City")["Traffic Index"].mean().sort\_values(ascending=False)

sns.barplot(x=city\_congestion.index, y=city\_congestion.values, palette="coolwarm")

plt.xticks(rotation=45)

plt.title("Average Traffic Index by City")

plt.xlabel("City")

plt.ylabel("Traffic Index")

plt.tight\_layout()

plt.show()

# Step 5: Delay Patterns Analysis

plt.figure(figsize=(8,5))

sns.scatterplot(data=traffic\_df, x="Traffic Index", y="Delay Time (mins)", hue="City", alpha=0.7)

plt.title("Traffic Index vs Delay Time")

plt.xlabel("Traffic Index")

plt.ylabel("Delay Time (mins)")

plt.legend(bbox\_to\_anchor=(1.05, 1), loc='upper left')

plt.tight\_layout()

plt.show()

# Step 6: Weather Impact on Traffic

plt.figure(figsize=(8,5))

weather\_congestion = traffic\_df.groupby("Weather Condition")["Traffic Index"].mean().sort\_values(ascending=False)

sns.barplot(x=weather\_congestion.index, y=weather\_congestion.values, palette="viridis")

plt.title("Average Traffic Index by Weather Condition")

plt.xlabel("Weather Condition")

plt.ylabel("Traffic Index")

plt.tight\_layout()

plt.show()

# Step 7: Congestion Trends Over Time

daily\_avg\_traffic = traffic\_df.groupby("Date")["Traffic Index"].mean()

plt.figure(figsize=(12,5))

sns.lineplot(x=daily\_avg\_traffic.index, y=daily\_avg\_traffic.values, color="darkred")

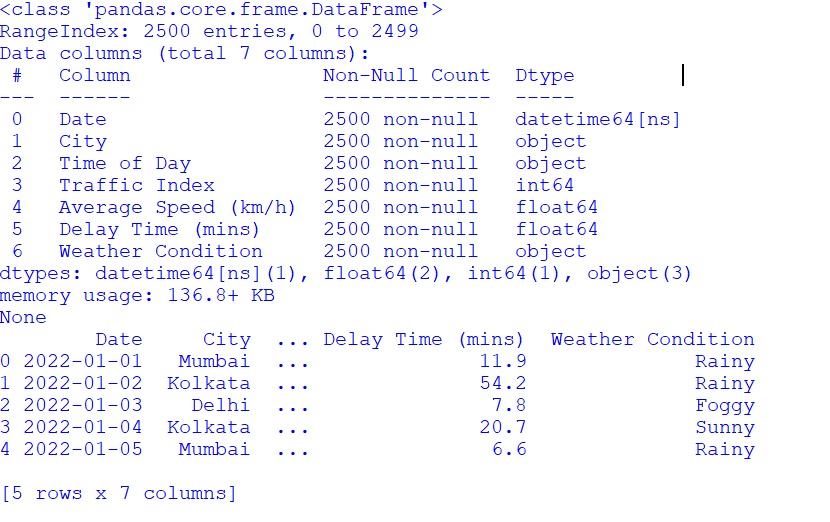
plt.title("Daily Traffic Congestion Trend")

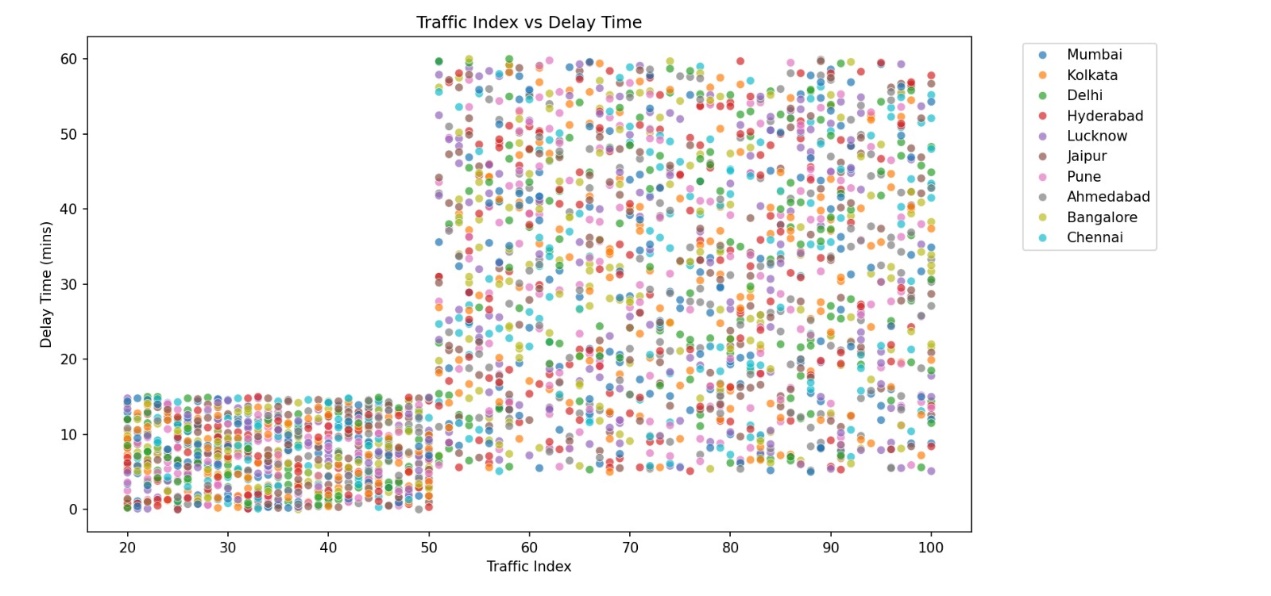
plt.xlabel("Date")

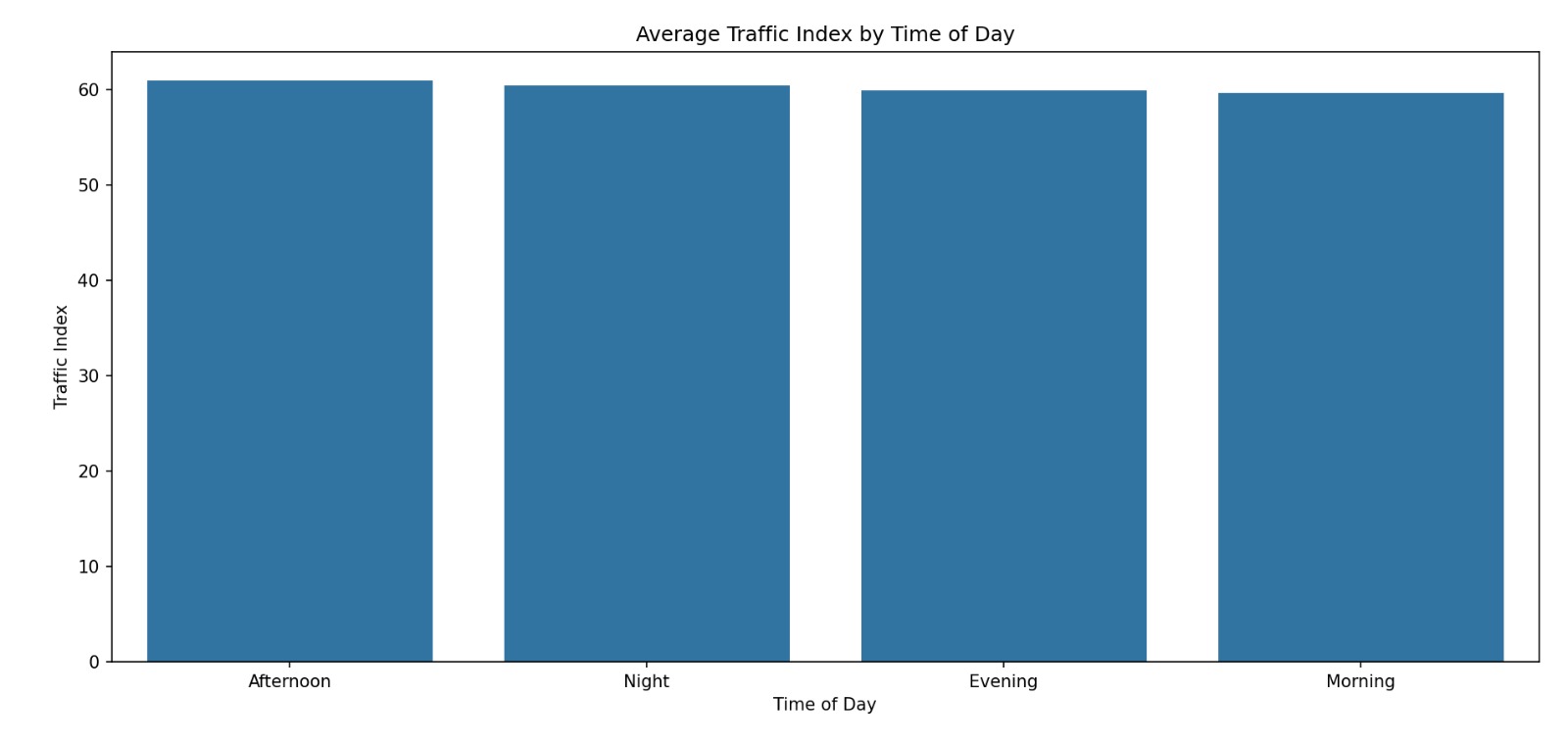
plt.ylabel("Average Traffic Index")

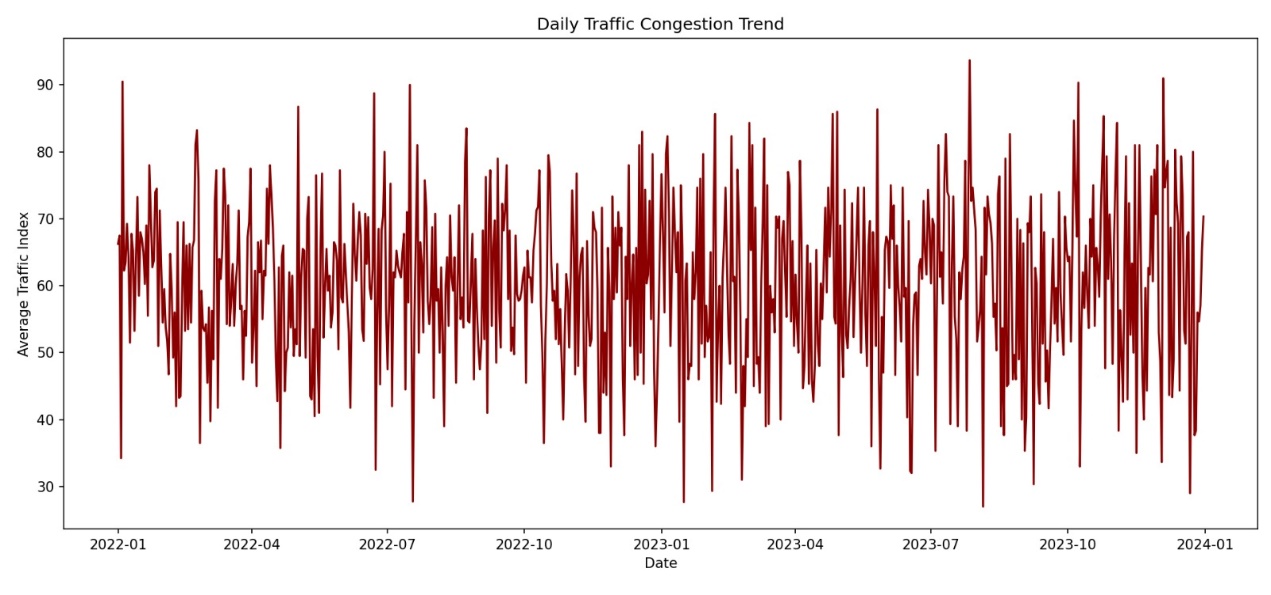
plt.tight\_layout()

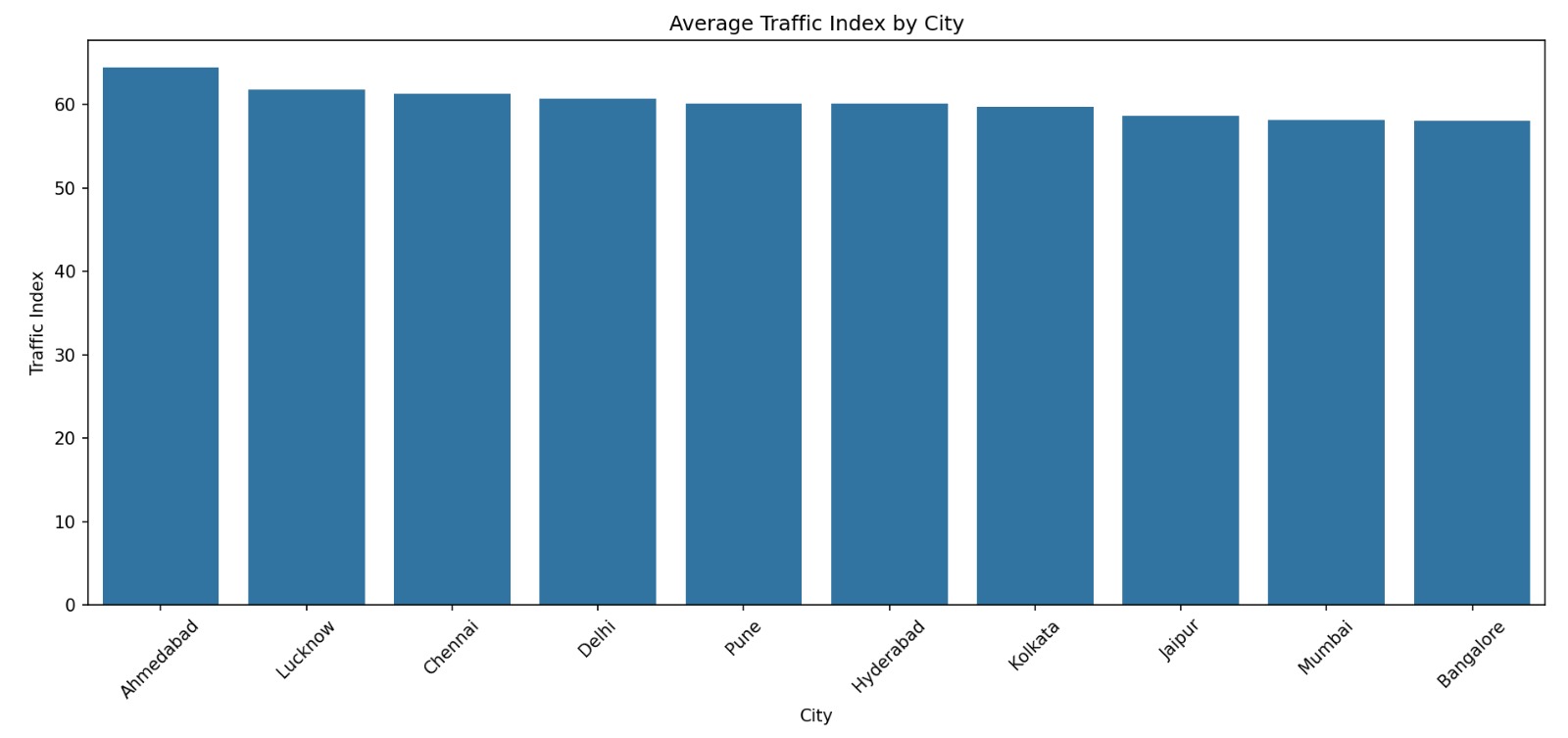
plt.show()











1. **CONCLUSION**

This traffic congestion analysis reveals critical trends impacting commuters in Indian cities. Time of day, city infrastructure, and weather conditions all play significant roles in shaping congestion levels. By identifying peak traffic zones and times, authorities can plan interventions and optimize transportation systems.

1. **FUTURE SCOPE**

Integrate live traffic APIs for real-time predictions

Apply time series forecasting to predict future congestion

Build smart route suggestion systems using ML models

Develop traffic dashboards for city authorities

1. **REFERENCES**

Indian Road Transport Ministry Statistics

Seaborn & Matplotlib Documentation

Pandas Library Documentation

Lovely Professional University Python Programming Curriculum

* Project Repository:

https://github.com/Jasmeen225/12319627.git