Mapping and Analysing Urban Growth Patterns in Indian Cities Using Geospatial and Socioeconomic Data

Project Report

Presented as part of the completion of coursework requirements

Team Leader: Vishal Singh (CS23I1039)

Team Members:
Sharad Kumar Dubey (CS23I1052)
Gyan Chandra (CS23I1053)
Suvam Sahu (CS23I1038)

Department of Computer Science and Engineering
Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram

November 8, 2025

Contents

1 Introduction and Methodology

1.1 Introduction

This project focuses on analysing the spatial and demographic growth of major Indian cities using open geospatial and census datasets. The study integrates population statistics from the Census of India with land-use and land-cover (LULC) data from ISRO's Bhuvan portal to quantify and visualize patterns of urbanization between 1991 and 2015.

1.2 Aim and Objectives

- To analyse urban expansion, population density, and land-use change across major Indian cities.
- To correlate demographic and land-use data for identifying spatial growth patterns.
- To develop visual and data-driven insights supporting sustainable urban planning.

1.3 Data Sources

- Census of India (1991, 2001, 2011)
- ISRO Bhuvan – Land Use and Land Cover Maps (2011, 2015)
- Derived metrics: Built-up ratio, population density, and growth rates

1.4 Methodology

The methodology followed in this project integrates geospatial processing, statistical analysis, and data-driven modeling to study the dynamics of urban expansion. A combination of census and remote sensing data was used to ensure both demographic and spatial perspectives were covered.

1. **Data Acquisition:** Population and demographic data were collected from the Census of India (1991, 2001, 2011), while Land Use and Land Cover (LULC) data were obtained from ISRO Bhuvan for 2011 and 2015.

2. **Preprocessing and Cleaning:** Census data were formatted into structured Excel sheets, handling missing or inconsistent entries. LULC maps were processed through image-based and OCR-based extraction pipelines to identify land-use categories and their corresponding areas.

3. **Feature Normalization:** All datasets were standardized for consistent temporal and spatial scales. Derived metrics such as population density, built-up ratio, and growth rates were computed to ensure comparability between cities and across decades.

4. **Data Integration:** Cleaned demographic and spatial datasets were merged into a harmonized table for each city, covering population, built-up area, land-use proportions, and density for the years 1991–2015.

5. **Statistical and Predictive Analysis:** Correlation matrices and linear regression models were applied to identify relationships between population and built-up area growth. Trend estimation for 2031 was performed using extrapolated regression lines.

6. **Visualization:** Graphical representations such as heatmaps, growth trajectories, and scatter plots were generated to illustrate temporal changes and inter-variable dependencies. These were complemented by geospatial visualizations of land-use transitions derived from satellite maps. .

2 Results and Visualizations

2.1 Correlation and Growth Patterns

Population density was found to correlate strongly with urban built-up expansion. Cities such as Delhi, Mumbai, and Bengaluru displayed rapid increases in both total population and built-up share between 1991–2015.

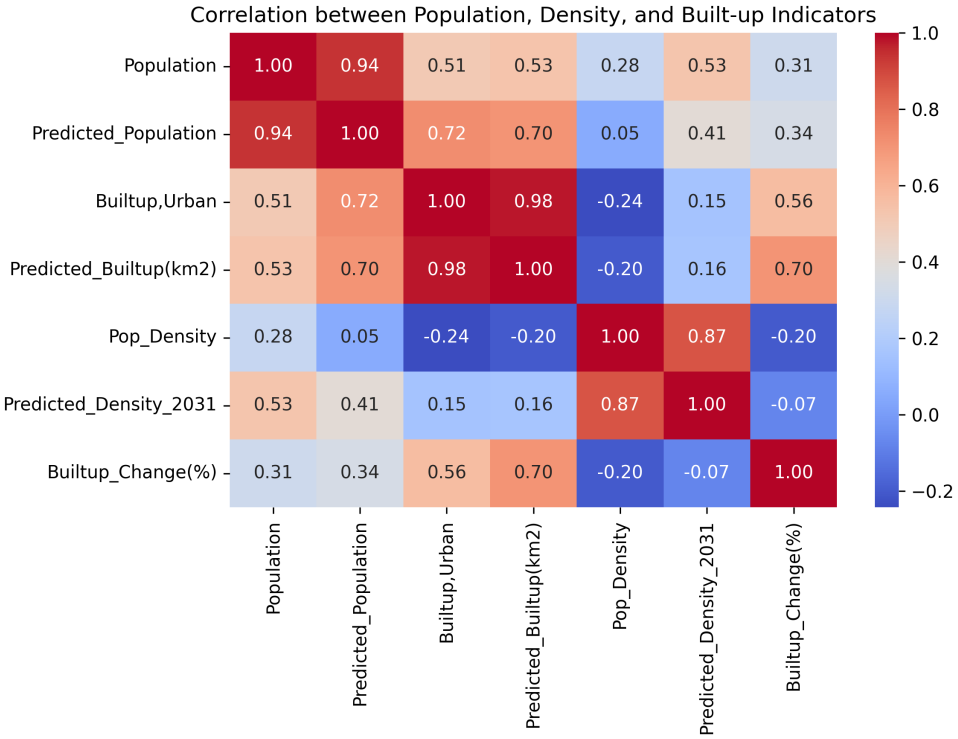

Figure 2.1: Correlation heatmap between demographic and land-use metrics

2.2 Temporal Growth Visualization

The population and built-up growth trends were visualized using line plots and scatter comparisons to demonstrate the changing spatial dynamics.

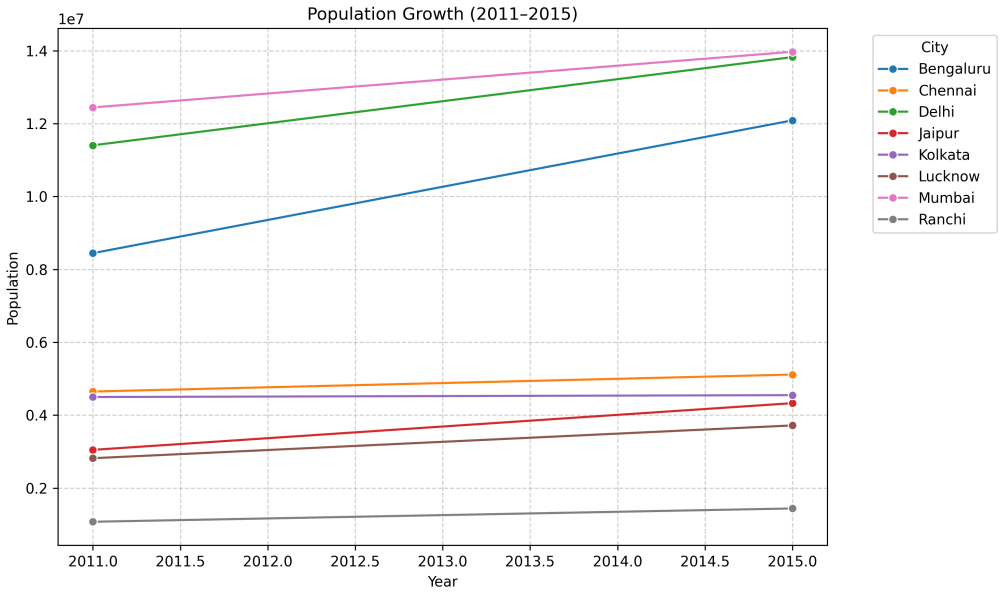

Figure 2.2: City-wise population growth (2011–2015)

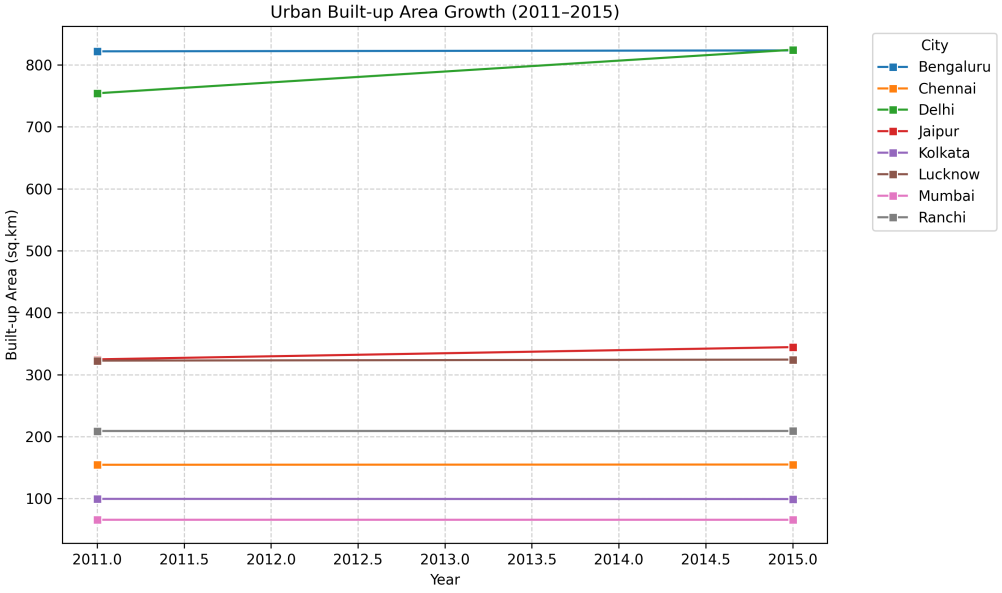

Figure 2.3: Urban built-up expansion trends (2011–2015)

2.3 Regression Analysis

Linear regression was applied to estimate the projected growth of urban population and built-up area till 2031, revealing that cities with higher economic activity showed stronger linear trends of densification.

3 Conclusion and References

3.1 Conclusion

The integration of census and geospatial data enabled an accurate mapping of urban growth patterns. Findings indicate that population and built-up area expansion are closely linked, with green and agricultural land being the most affected. The results can guide policymakers in planning infrastructure expansion, optimizing land use, and ensuring sustainability in high-growth regions.

3.2 References

- Towns and urban agglomerations classified by population size class in 2011 with variation between 1901 and 2011 - Class I (population of 100,000 and above).
- Primary Census Abstract - Urban Agglomeration India - 2011.
- ISRO Bhuvan (2011–2015). Land Use and Land Cover Data Portal.
- Project source code and datasets: https://drive.google.com/drive/folders/1YMfBNlqDWtO4-f30IDYSgEsBvmkbbETz?usp=sharing

4 README: Code Execution Guide

4.1 Environment Setup

- Platform: Google Colab / Jupyter Notebook
- Dependencies: `pandas`, `numpy`, `matplotlib`, `seaborn`, `opencv-python`, `tabula-py`, `pytesseract`
- Mount your Google Drive to store datasets:

```
from google.colab import drive
drive.mount('/content/drive')
```

4.2 Project Folder Structure

```
Urban_Growth_Project/
|-- data/
|   |-- excel/
|   |-- pdfs/
|   |-- maps/
|   |-- legends/
|-- output/
|   |-- Urban_Land_and_Population.xlsx
|   |-- Urban_LandUse_Data_OCR.xlsx
|   |-- Urban_Harmonized_2011_2015.xlsx
|   |-- Urban_Correlation_Heatmap.png
|   |-- Population_Trends.png
|   |-- Builtup_Trends.png
|-- Urban_Growth_Analysis.ipynb
```

4.3 Execution Steps

1. Open `Urban_Growth_Analysis.ipynb` in Google Colab.

2. Mount Google Drive.

3. Run all cells sequentially.

4. Outputs will be stored in the `output/` directory.

4.4 Outputs

- Harmonized dataset (Excel)
- Visualizations (Heatmaps, Growth Curves, Scatter Plots)
- Linear Regression Forecasts (2031)