SIMATS ENGINEERING
Saveetha Institute of Medical and Technical Sciences
Chennai-602105.

DSA0613- Data Handling and Visualization for Data Analytics

Geospatial Visualization of Population Density for Regional Pattern and Trend Analysis

Monika R
1923234281

Anu R
192324278

Supervisor
Dr. Kumaragurubaran . T, Professor
Dr. Senthilvadivu. S, Professor
Department of Computer Science and Engineering
SIMATS Engineering

Abstract

This project focuses on the geospatial visualization of population density to identify regional patterns and analyze trends over time. Population data is collected from reliable sources and pre-processed to remove inconsistencies and missing values. Population density is calculated using standard metrics and mapped onto geographic regions using geospatial visualization techniques such as choropleth maps and color scales. The visual representations help in clearly identifying high-density, medium-density, and low-density regions. Trend analysis is performed to observe changes in population distribution across different time periods, enabling the study of urbanization and regional growth patterns. The proposed system provides an intuitive and analytical approach to understanding spatial population data, making complex datasets easier to interpret. The results of this project can support applications such as urban planning, resource allocation, and policy formulation by offering meaningful visual insights into population distribution and trends.

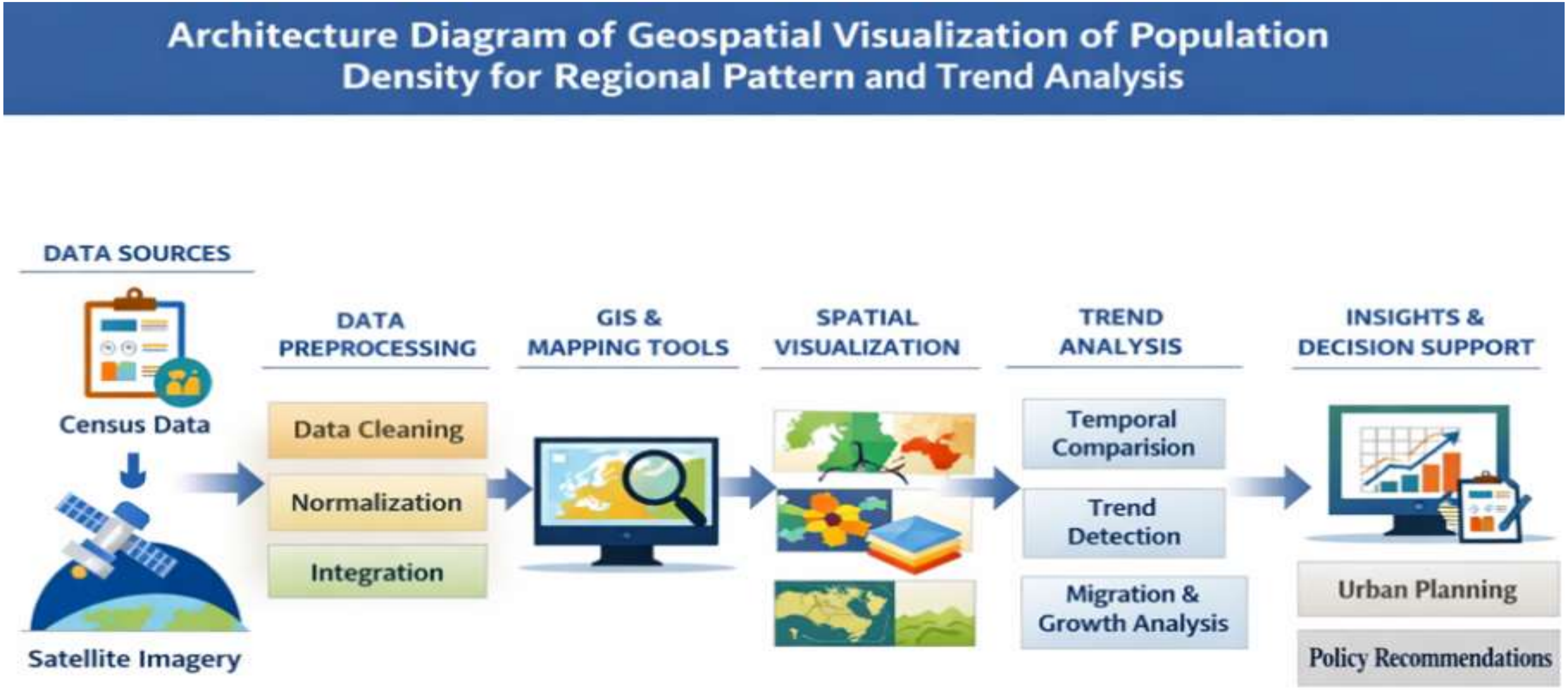
Problem statement

- Population density varies significantly across regions, creating planning and management challenges
- Traditional data formats do not effectively represent spatial population patterns
- Lack of clear visualization makes it difficult to identify high-density and low-density regions
- Regional population growth and decline trends are hard to analyze without geospatial tools
- Decision-makers need spatial insights for urban planning and resource allocation
- Integrating population data with geographic maps is essential for meaningful analysis
- Geospatial visualization helps in understanding regional patterns and temporal trends

Objectives

- To collect and preprocess population density data from reliable sources
- To integrate population data with geospatial maps for spatial analysis
- To visualize population density distribution across different regions
- To identify high-density and low-density population areas
- To analyze regional population patterns and variations
- To study population growth and decline trends over multiple years
- To compare population distribution between regions

System Architecture diagram



MODULES

Module 1: Population Density Pattern Detection

Module 2: Geospatial Population Density Trend Prediction

Module 1: Population Density Pattern Detection

What We Do?

- Collect and preprocess population density data
- Integrate population data with geospatial boundary maps
- Analyze spatial population distribution across regions
- Visualize population density patterns using maps and charts
- Compare population density variations between regions
- Detect high-density and low-density population areas
- Identify regional population clustering and dispersion patterns

Why We Need

- To identify spatial population patterns and regional disparities
- To clearly understand population distribution across regions
- To detect population concentration and sparse areas
- To support urban planning and infrastructure development

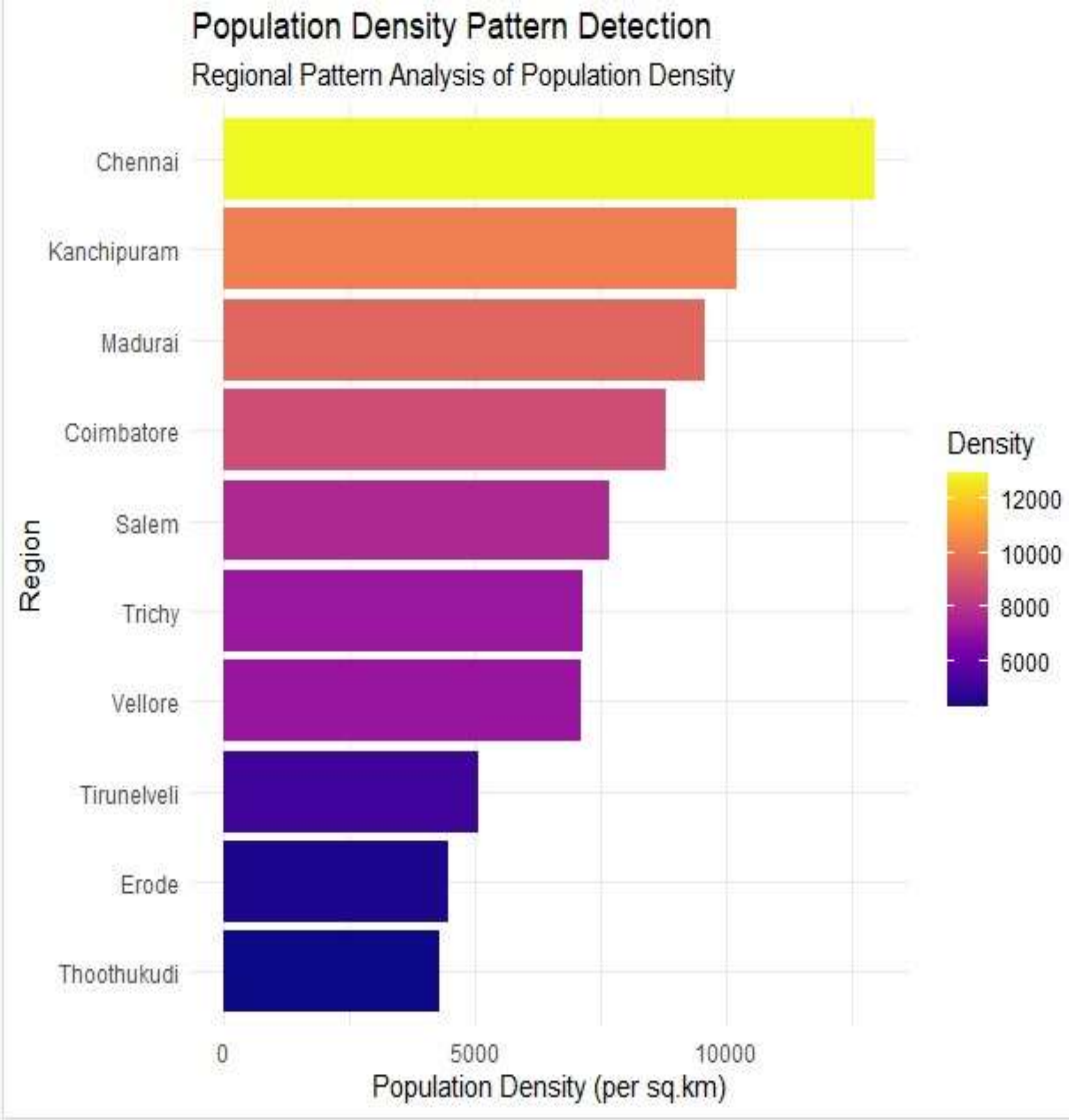
Visualization Techniques Used

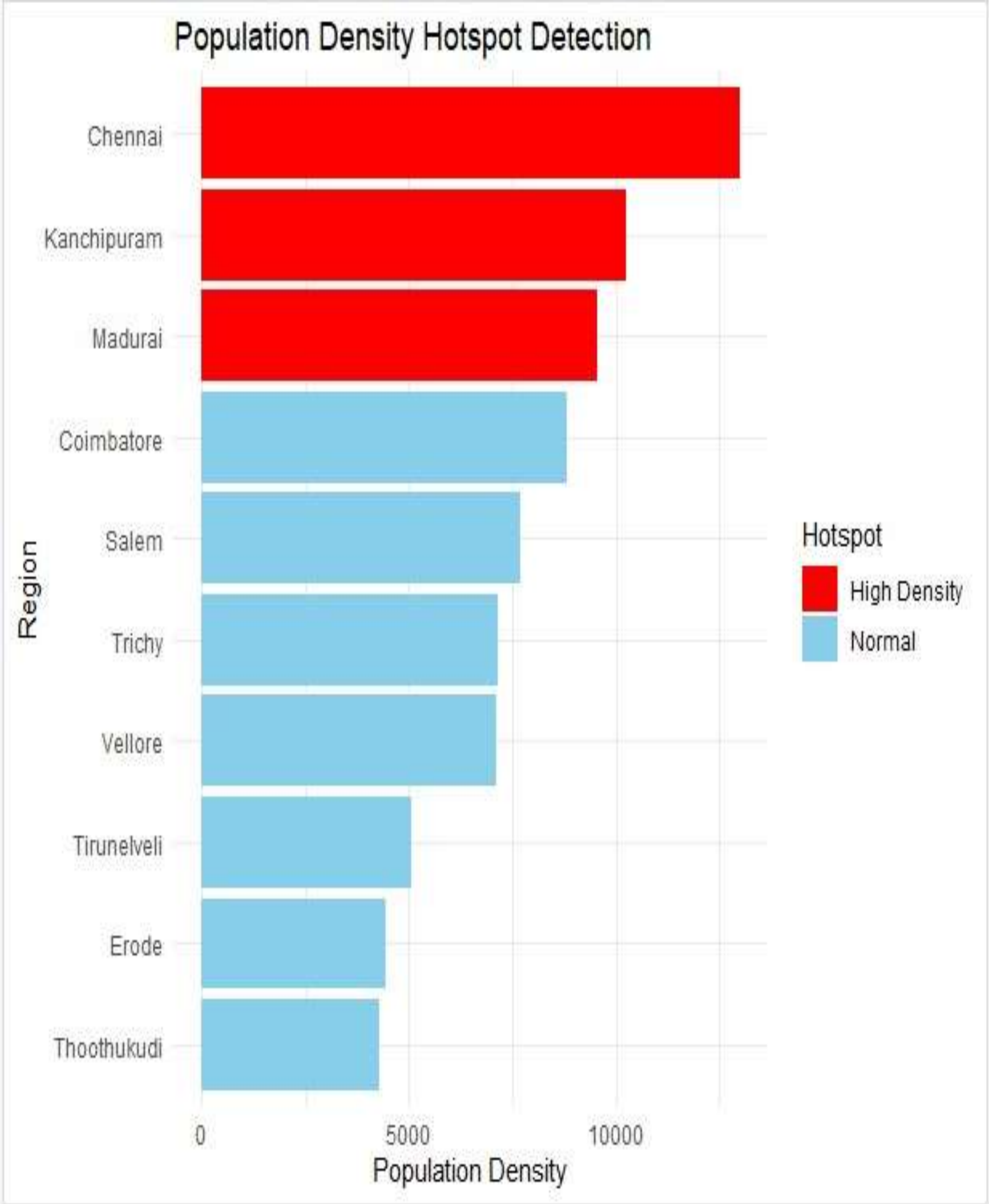
- **Point/Bubble maps** – city-level population representation
- **Thematic maps** – spatial comparison of population density
- **Bar charts** – regional population density comparison.

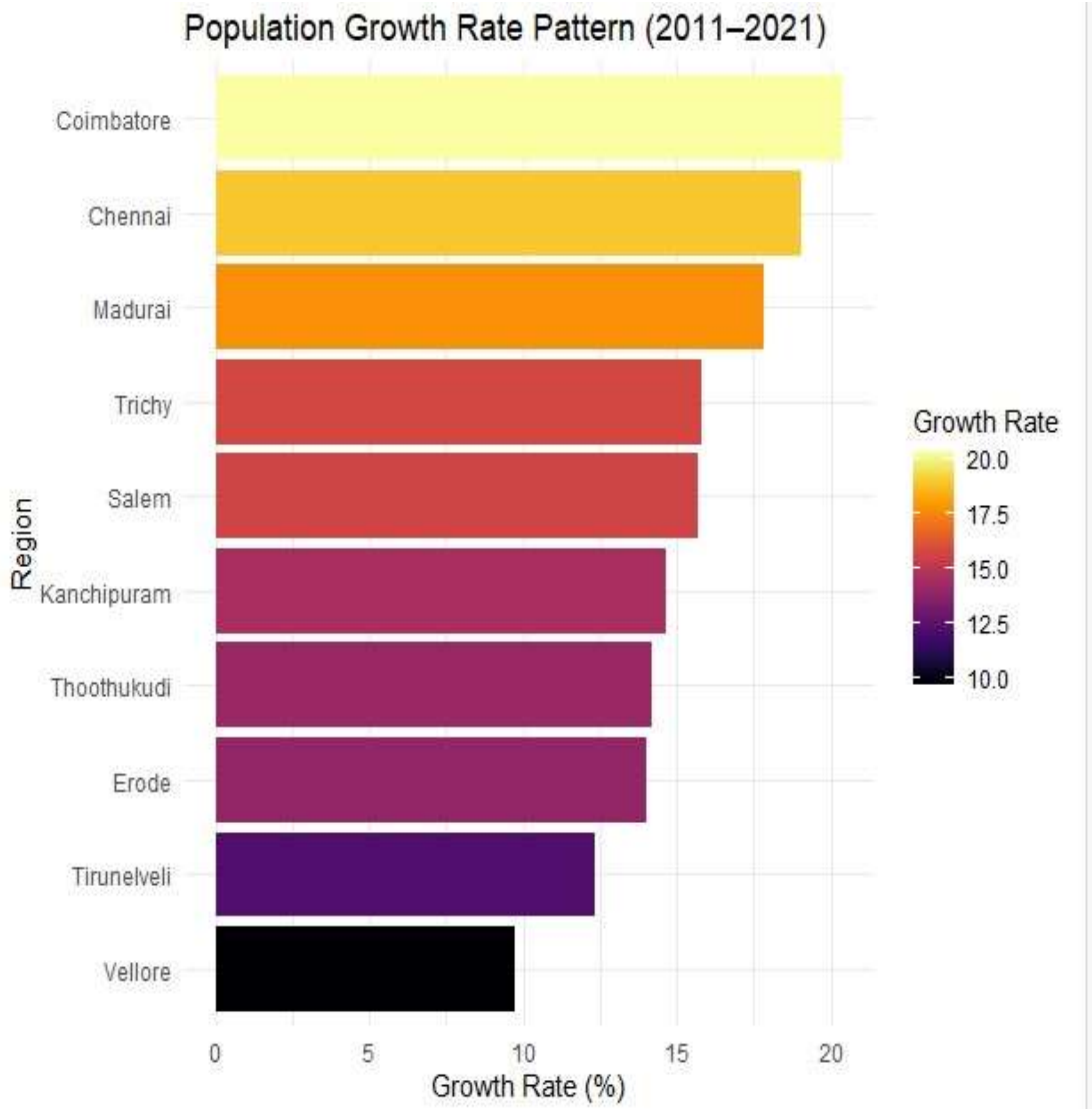
Module 1: Implementation

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("viridis")
library(ggplot2)
library(dplyr)
library(viridis)
setwd("C:/Users/tamil/Documents/population_density.csv")
population <- read.csv("population_density.csv")
head(population)
population <- population %>%
  mutate(Population_Density = Population_2021 / Area_km2)
ggplot(population, aes(x = reorder(Region,
Population_Density),
                       y = Population_Density,
                       fill = Population_Density)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_viridis(option = "plasma") +
  theme_minimal() +
  labs(
    title = "Population Density by Region",
    subtitle = "Geospatial Visualization of Population Density",
    x = "Region",
    y = "Population Density (per sq.km)"
  )
population <- population %>%
  mutate(Growth_Rate = ((Population_2021 -
Population_2011) /
                        Population_2011) * 100)
ggplot(population, aes(x = reorder(Region, Growth_Rate),
                       y = Growth_Rate,
                       fill = Growth_Rate)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_viridis(option = "inferno") +
  theme_minimal() +
  labs(
    title = "Population Growth Rate (2011–2021)",
    x = "Region",
    y = "Growth Rate (%)"
  )
threshold <- quantile(population$Population_Density, 0.75)
population <- population %>%
  mutate(Hotspot = ifelse(Population_Density >= threshold,
                          "High Density", "Normal"))

ggplot(population, aes(x = Region,
                       y = Population_Density,
                       fill = Hotspot)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_manual(values = c("High Density" = "red",
                               "Normal" = "skyblue")) +
  theme_minimal() +
  labs(
    title = "Population Density Hotspot Regions",
    x = "Region",
    y = "Population Density"
  )
threshold <- quantile(population$Population_Density, 0.75)

population <- population %>%
  mutate(Hotspot = ifelse(Population_Density >= threshold,
                          "High Density", "Normal"))
ggplot(population, aes(x = Region,
                       y = Population_Density,
                       fill = Hotspot)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_manual(values = c("High Density" = "red",
                               "Normal" = "skyblue")) +
  theme_minimal() +
  labs(
    title = "Population Density Hotspot Regions",
    x = "Region",
    y = "Population Density"
  )
```

Module 1: Outputs







Module 2: Geospatial Population Density Trend Prediction

What We Do?

• Collect and preprocess historical population density data

• Integrate multi-year population datasets with geospatial boundary maps

• Analyze spatial and temporal population patterns

• Visualize population density trends across regions and time

• Compare population growth and decline between regions

• Apply trend analysis to predict future population density

• Identify emerging high-density and low-density zones

Why We Need

• To understand how population distribution changes over time

• To identify long-term spatial population trends

• To predict future population growth or decline patterns

• To support urban planning, infrastructure development, and policy decisions

Visualization Techniques Used

- Choropleth maps – region-wise population density trends
- Heatmaps – population concentration and predicted hotspots
- Point/Bubble maps – city-level population growth indicators
- Time-based maps – visualization of population change over year
- Line and bar charts – comparison of historical and predicted trends

Module 2: Implementation

```r
install.packages(c("ggplot2", "dplyr", "viridis"))
library(ggplot2)
library(dplyr)
library(viridis)
```

2. Load the Dataset

```r
setwd("D:/population_project")
population <- read.csv("population_density.csv")

head(population)
```

3. Calculate Population Density

```r
population <- population %>%
  mutate(Pop_Density_2011 = Population_2011 / Area_km2,
         Pop_Density_2021 = Population_2021 / Area_km2)
```

4. Calculate Annual Growth Rate

Assuming linear growth over 10 years (2011 → 2021):

```r
population <- population %>%
  mutate(Annual_Growth_Rate = (Pop_Density_2021 - Pop_Density_2011) / 10)
head(population)
```

5. Predict Future Population Density

Predict for 2025 & 2030 using linear trend:

```r
population <- population %>%
  mutate(
    Pred_2025 = Pop_Density_2021 + Annual_Growth_Rate * 4,
    Pred_2030 = Pop_Density_2021 + Annual_Growth_Rate * 9
  )
head(population)
```

6. Combine Data for Plotting

```r
population_long <- population %>%
  select(Region, Pop_Density_2011, Pop_Density_2021, Pred_2025, Pred_2030) %>%
  pivot_longer(cols = c(Pop_Density_2011, Pop_Density_2021, Pred_2025, Pred_2030),
               names_to = "Year",
  values_to = "Density")

population_long$Year <- gsub("Pop_Density_", "", population_long$Year)
```

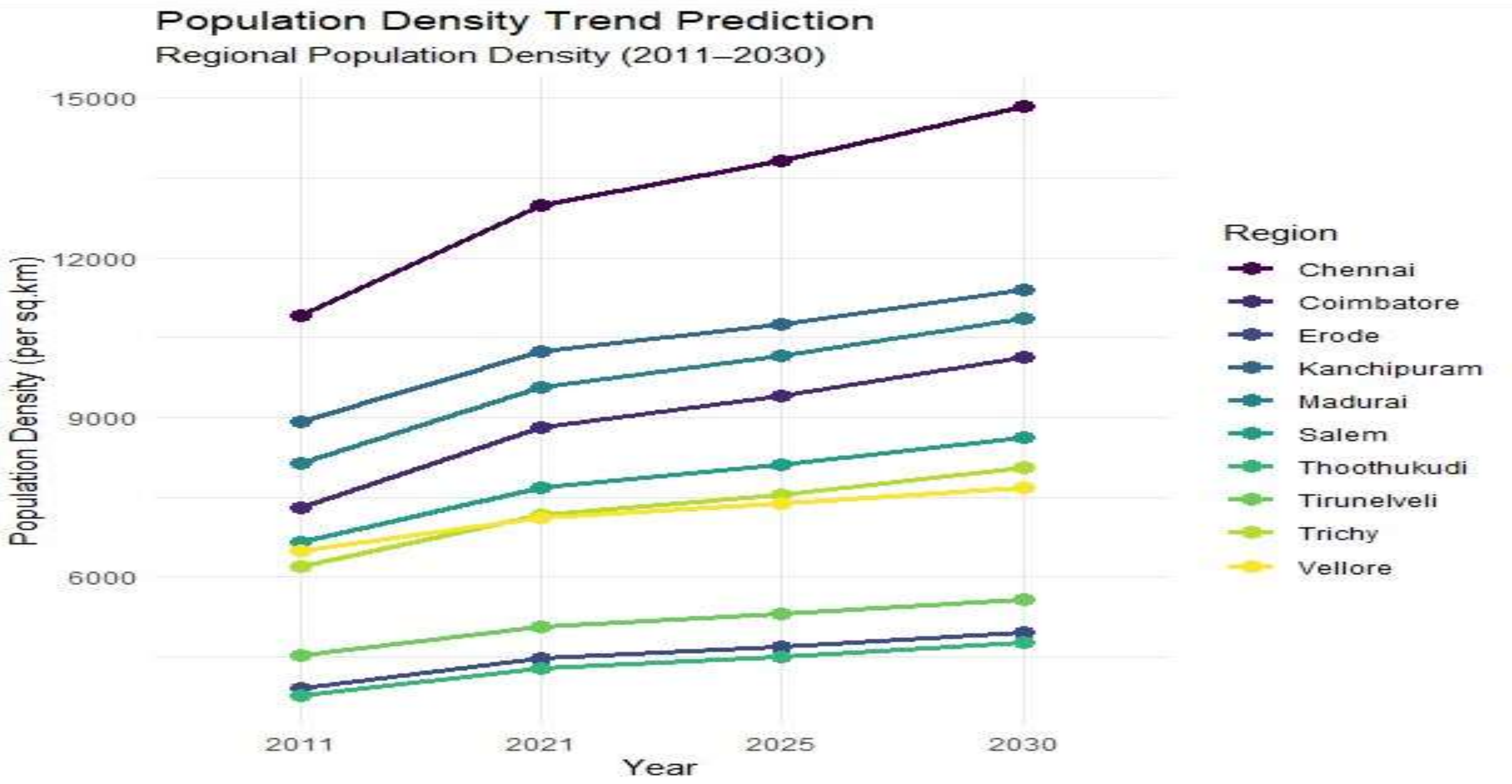
7. Plot Population Density Trend

```r
ggplot(population_long, aes(x = Year, y = Density, group = Region, color = Region)) +
  geom_line(size = 1.2) +
  geom_point(size = 3) +
  scale_color_viridis(discrete = TRUE) +
  theme_minimal() +
  labs(
    title = "Geospatial Population Density Trend Prediction",
    subtitle = "Regional Population Density (2011–2030)",
    x = "Year",
    y = "Population Density (per sq.km)",
    color = "Region"
  )
```

8. Optional: Bar Plot for Predicted 2030 Density

```r
ggplot(population, aes(x = reorder(Region, Pred_2030), y = Pred_2030, fill = Pred_2030)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  scale_fill_viridis(option = "plasma") +
  theme_minimal() +
  labs(
    title = "Predicted Population Density for 2030",
    x = "Region",
    y = "Population Density (per sq.km)"
  )
```

Module 2: Output



Results

- Regional population distribution was clearly visualized using thematic maps
- High-density and low-density population regions were effectively identified
- Spatial population patterns and regional disparities were clearly observed
- Temporal visualizations revealed population growth and decline trends across regions
- Comparative analysis enabled clear understanding of regional population changes
- Geospatial visualizations simplified complex spatio-temporal population data
- The results provide valuable insights to support planning and policy decision-making

Conclusion

- Geospatial visualization techniques effectively mapped population density across regions
- Choropleth maps clearly identified high-density, medium-density, and low-density areas
- Spatial population patterns and regional variations were easily interpreted
- Trend analysis revealed changes in population distribution over different time periods
- Visual representations simplified complex spatial and temporal datasets
- The system provides meaningful insights to support urban planning, resource allocation, and policy formulation

Future scope

- Inclusion of advanced machine learning models to improve population trend prediction accuracy
- Expansion of the system to support finer spatial scales such as ward or street-level analysis
- Development of interactive web-based dashboards for real-time visualization and user exploration
- Incorporation of socio-economic and environmental factors to enhance pattern detection

Reference

1. **Huang, W., & Xia, J. (2021).**
Geovisualization of urban population density: Methods and applications. ISPRS International Journal of Geo-Information, 10(3), 145.

2. **Andrienko, N., Andrienko, G., Fuchs, G., & Wood, J. (2020).**
Visual analytics for spatio-temporal data. Springer Handbook of Geographic Information, 1257–1281.

3. **Batty, M. (2021).**
Urban analytics defined. Environment and Planning B: Urban Analytics and City Science, 48(1), 3–6.Top of Form

4. **Mota, R., et al. (2022).**
A comparison of spatiotemporal visualizations for 3D urban analytics. IEEE Transactions on Visualization and Computer Graphics, 29(1), 1277–128.

5. **Kovacova, M., et al. (2022).**
3D simulation and geospatial data mining in smart city governance. Geopolitics, History, and International Relations, 14(1), 43–58.