

Population based bus stop analysis of Lahore

Muhammad Murtaza Nadeem (MSDS24040)

Session: Spring 2025

Supervisor: Dr. Adnan Siddique

Course: Spatial Data Science

1. Introduction

- Efficient public transport relies on equitable bus stop distribution.
- Lahore's rapid population growth strains current infrastructure.
- This project analyzes how well bus stops serve population.
- Combining spatial data helps identify underserved areas for improvement.

2. Problem Statement

Are Lahore's bus stops aligned with population distribution?
Which areas are underserved or overcrowded?

3. Data

Population Raster: Gridded population data (2020) from WorldPop.

Bus Stops: OSM-based bus stop locations in Lahore.

Lahore Boundary: Used to clip relevant spatial data.

Preprocessing:

- Removed no-data values from raster.
- Reprojected all datasets to UTM (EPSG:32643) for accurate spatial analysis.
- Extracted population near each bus stop using spatial sampling.

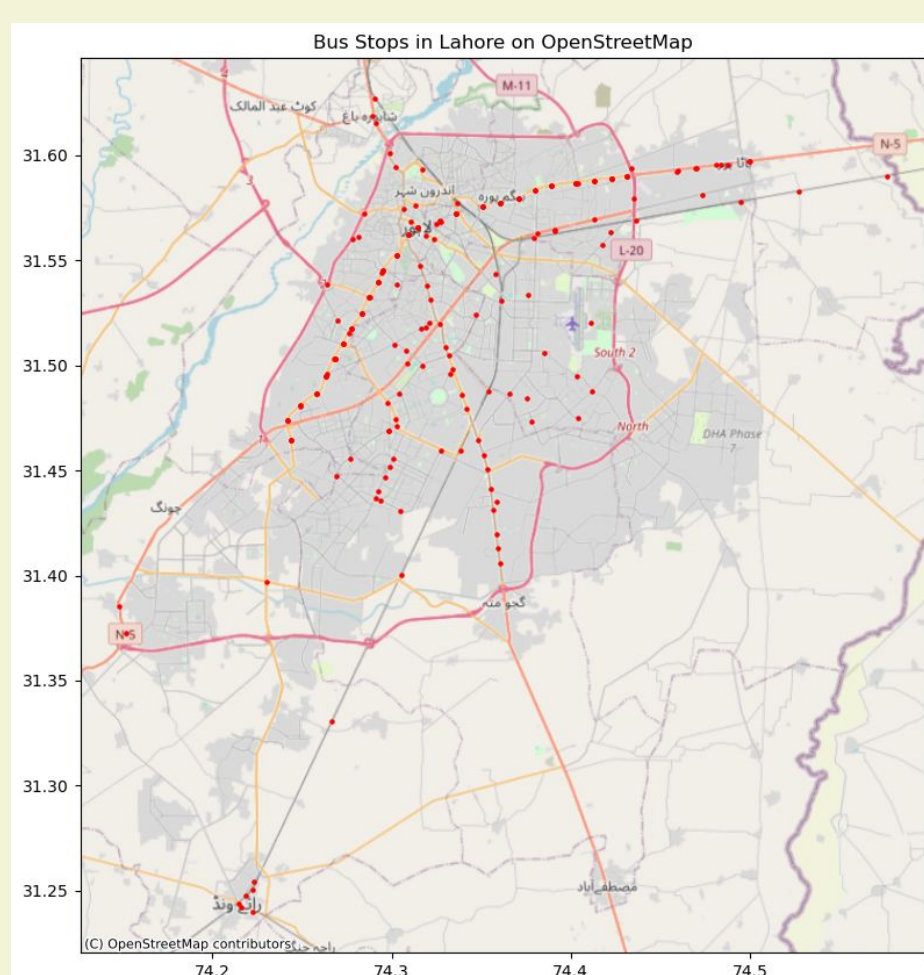


Fig. 1: Public bus stop locations extracted from OpenStreetMap for the Lahore region.

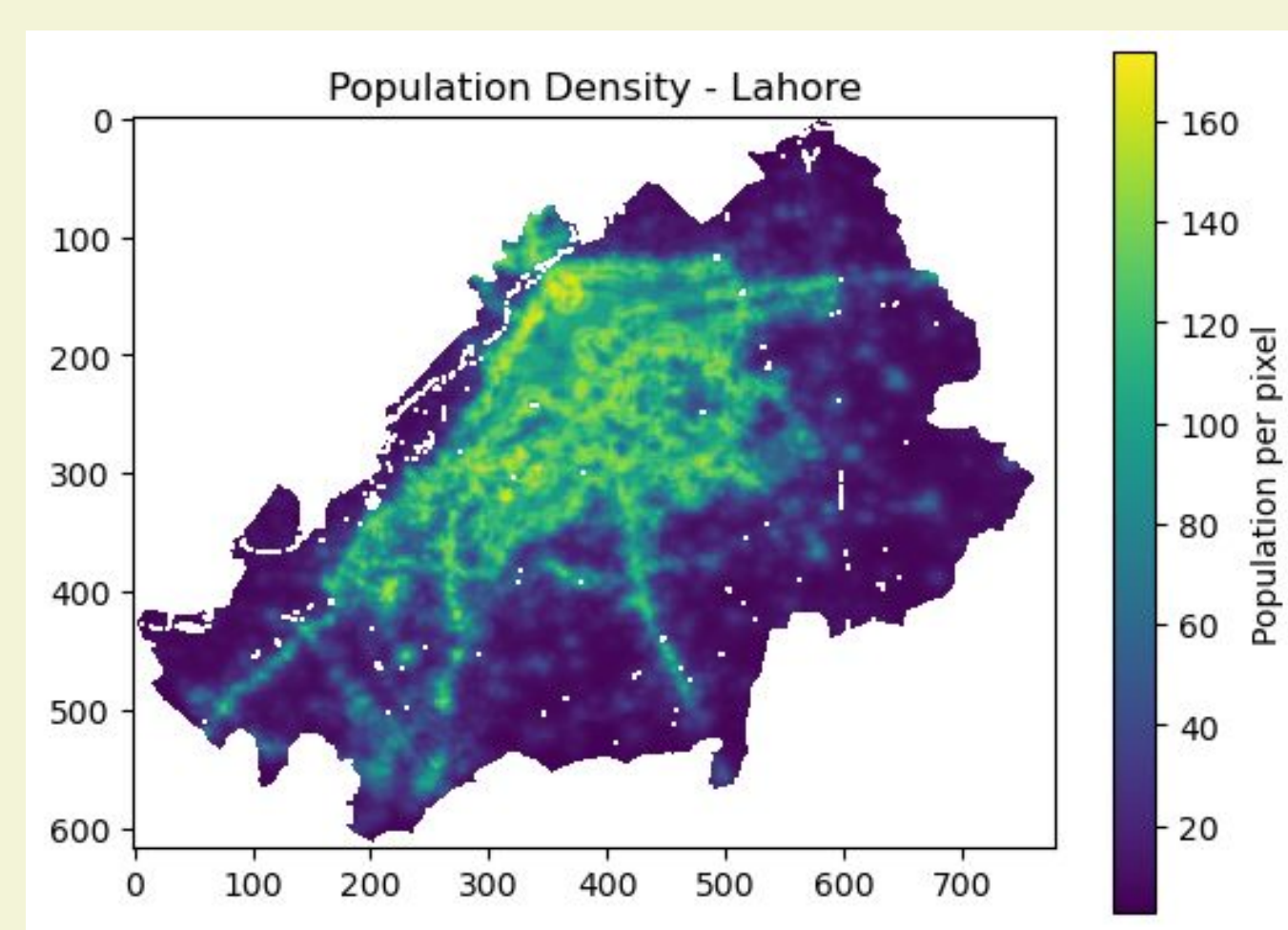


Fig. 2: 2020 gridded population estimates showing densely and sparsely populated areas of Lahore.

4. Methodology

- Population Raster (WorldPop 2020) clipped to Lahore boundary and cleaned by masking no-data values.
- Bus stop locations obtained from OpenStreetMap; filtered to retain only those within Lahore.
- Both datasets reprojected to UTM (EPSG:32643) for accurate spatial analysis in meters.
- Population values extracted at each bus stop using raster sampling.
- Nearest Neighbor Distance (NND) calculated to study bus stop clustering.
- Spatial Autocorrelation (Moran's I) used to detect clustering of population values around stops.
- Equity of service measured using Gini Coefficient and Lorenz Curve.
- Underserved areas identified based on population density and distance from nearest stop.

5. Results & Discussion

- High population areas (e.g., central Lahore) show higher concentration of bus stops, while peripheral regions (e.g., Raiwind) have sparse coverage.
- Mean Nearest Neighbor Distance: 543 m; lower than expected under CSR (1514 m) → indicates clustering of stops.
- Moran's I = 0.49 ($p = 0.001$) → significant spatial autocorrelation in population served per stop.
- Gini Coefficient = 0.094 → relatively equitable distribution of population coverage among bus stops.
- Underserved areas identified through raster highlighting regions with high population but no nearby stops.

Solution: Use population coverage and spatial analysis to guide future stop placement — particularly in southern and eastern Lahore.

Future Work: Extend analysis by incorporating locations of schools, colleges, and office zones to assess adequacy of public transport access for daily commuters.

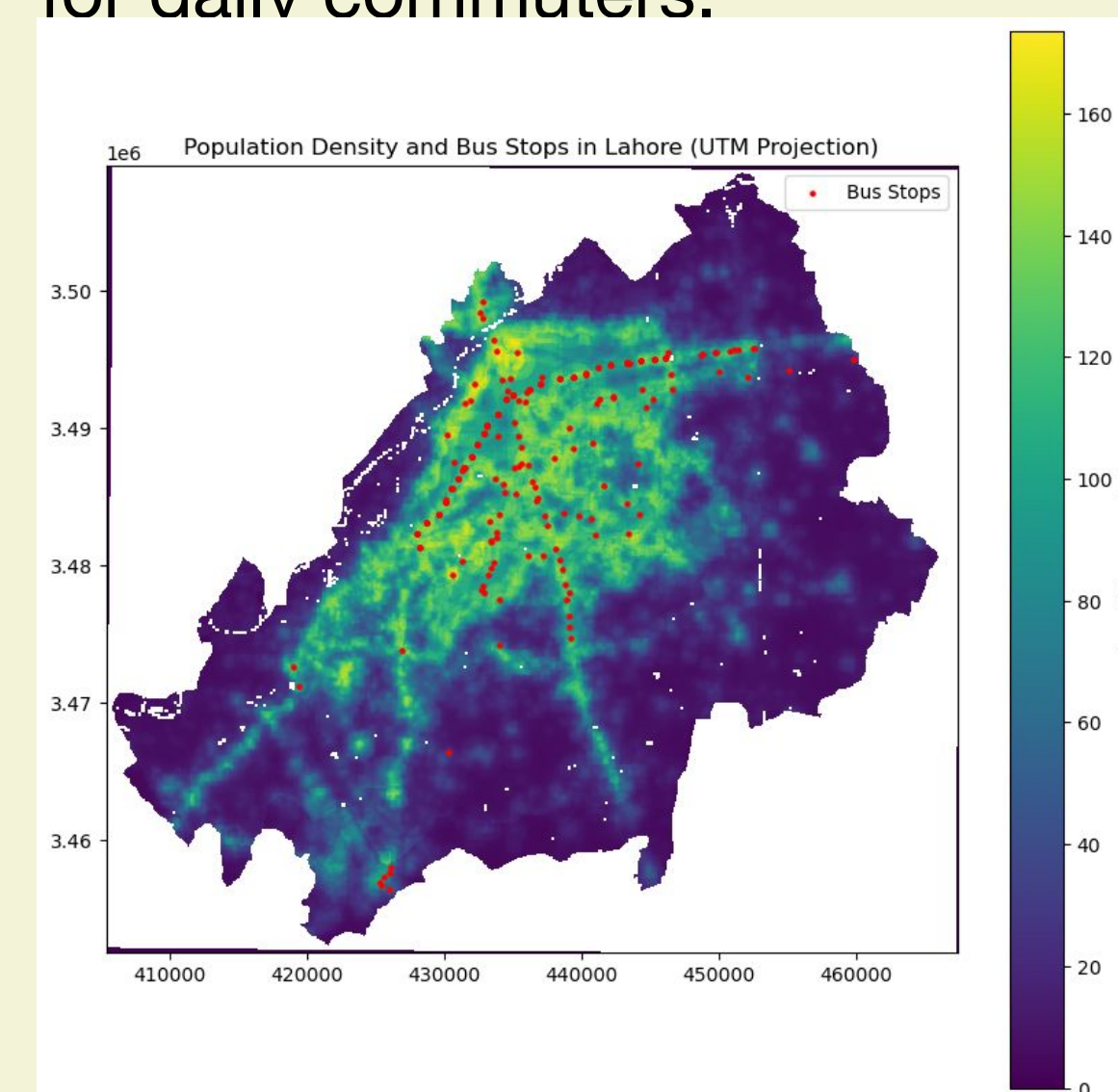


Fig. 3,4: Population & bus stop overlay

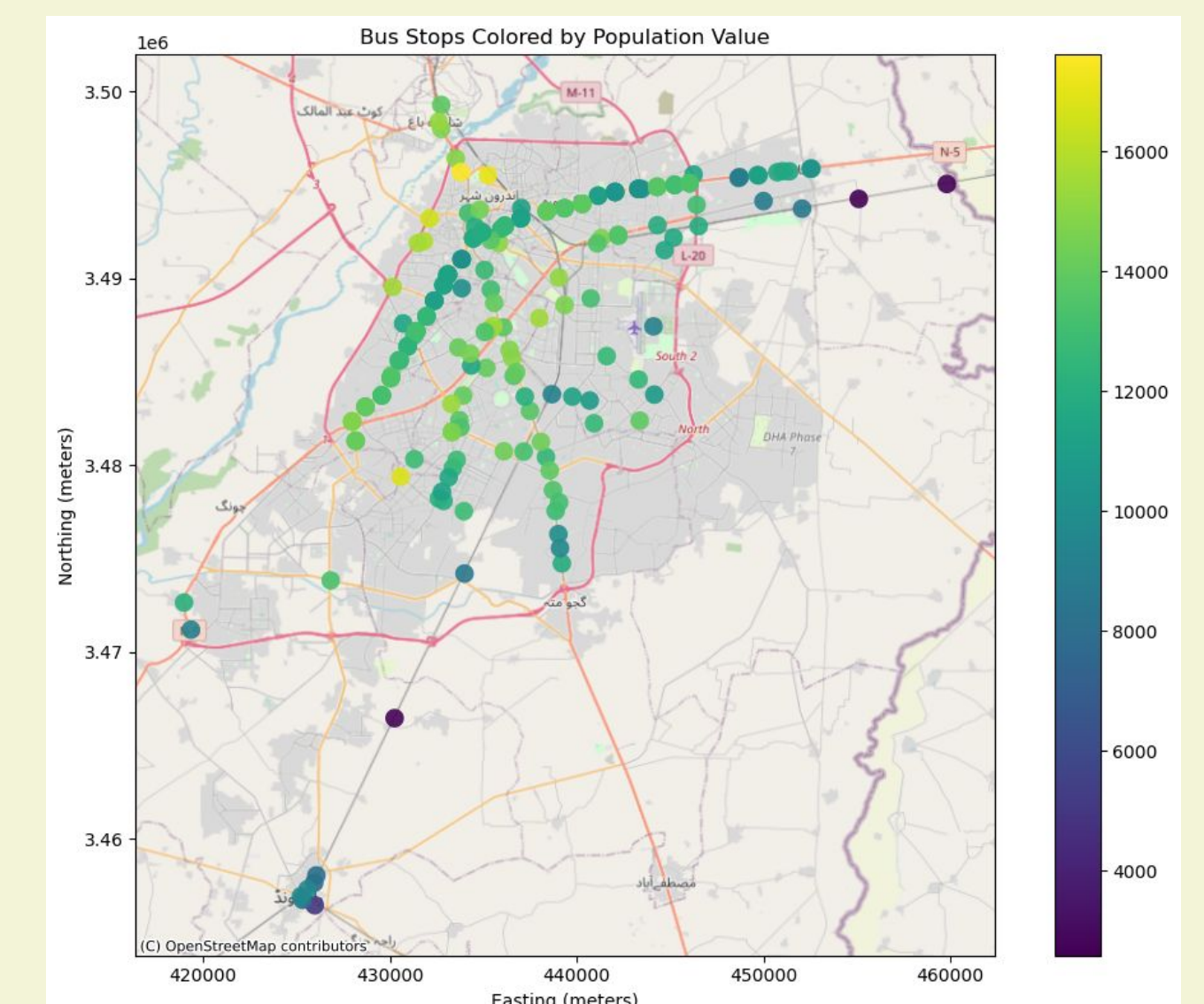


Fig. 4: Population served by each bus stop

Acknowledgments: WorldPop <https://dx.doi.org/10.5258/SOTON/WP00645>