



Comparative Analysis of Urban Delineation Approaches for Pakistan

By Lydia Teinfalt
George Washington University
Fall 2022
lteinfalt@gwu.edu

What is Urban vs Rural?

PROBLEM SETTING

- Pakistan is fifth most populous country in the world
- Pakistan's 2017 Population Study reported **36.4%** was urban. The European Commission Global Human Settlement Layer published a significantly higher figure **73.1%**
- Each country has differing definition for urban areas, including Pakistan

LITERATURE REVIEW

- Literature highlights two main approaches to defining urban areas:
 - Degree of Urbanization (Dijkstra, 2021)
 - Dartboard (de Bellefon, 2021)

CAPSTONE PROJECT

- Conduct geospatial analysis of Pakistan by incorporating gridded population data using Degree of Urbanization and Dartboard approaches to classify urban/rural areas
- Compare results to Pakistan's 2017 Population Study published by Pakistan Bureau of Statistics (PBS)

Data

Geospatial Information
Population Data

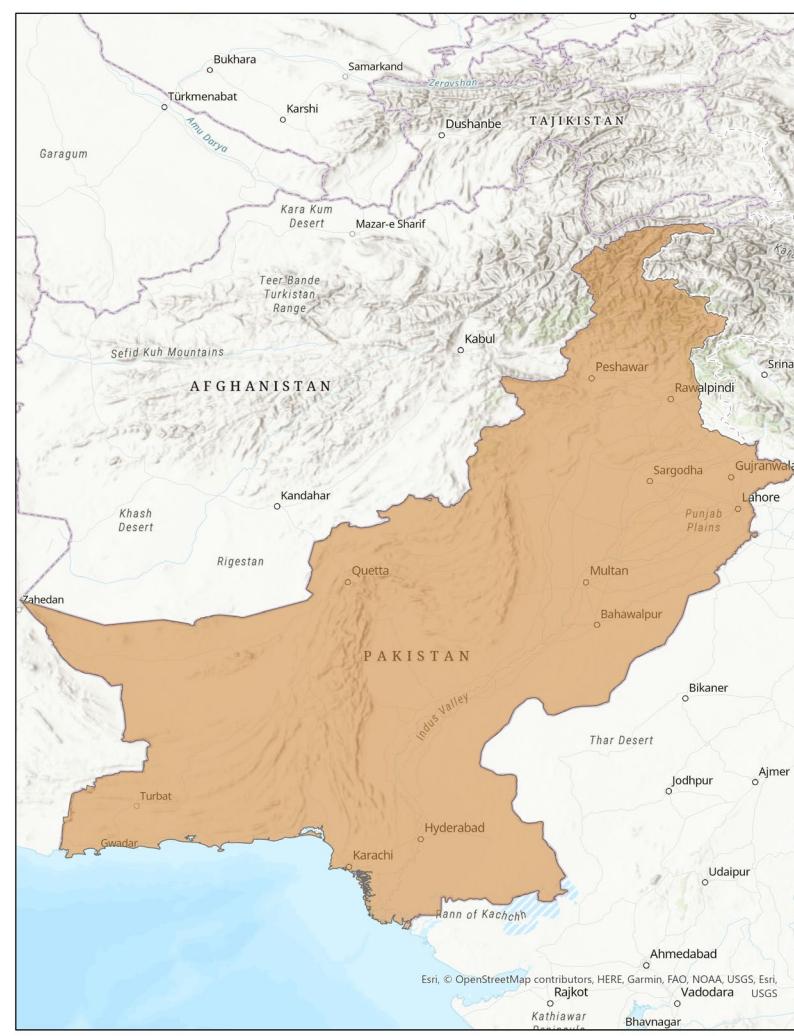
PART.02



Geospatial Data

VECTOR & RASTER DATA

- Two types of geospatial data: vector and raster. Vector data uses geometry to represent the geographical shapes
 - Pakistan's shape files represent its administrative boundaries using polygons
 - Raster data is a matrix of pixels of uniform size. Raster cell sizes for this project are **1 km x 1 km and 250 m x 250 m**
 - Gridded population is expressed as population density per cell and in raster format

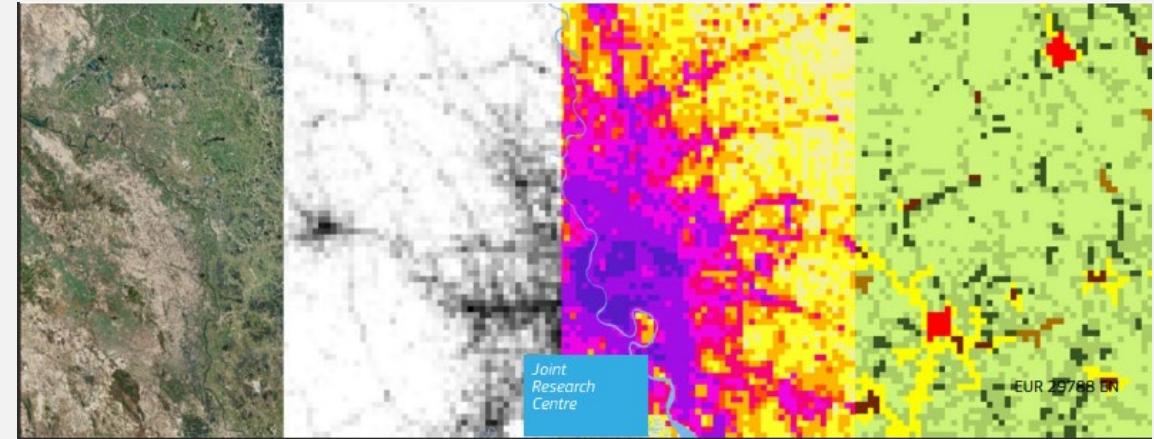


(UN OCHA, OCT 2022)
[HTTPS://DATA.HUMDATA.ORG/DATASET/COD-AB-PAK?](https://data.humdata.org/dataset/COD-AB-PAK?)

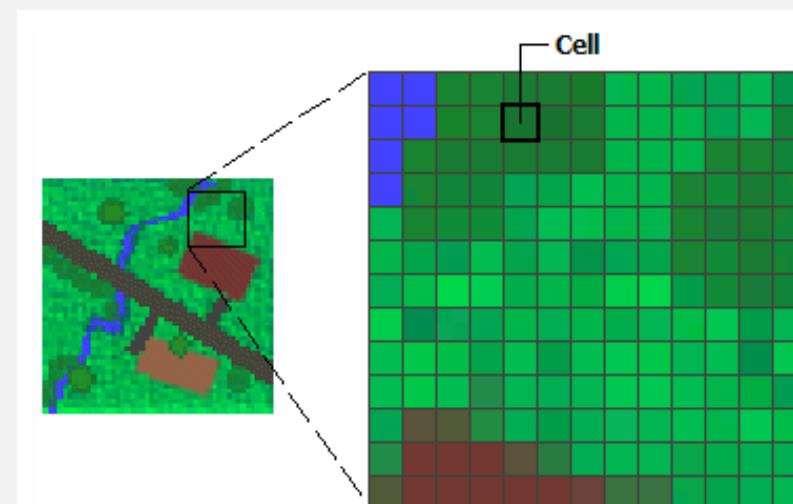
Gridded Population Data

GHS-POP & WORLDPOP

- **GHS-POP:** Combines satellite-derived built-up areas with census data. Grid cells with more built-up land -- higher population¹. (GHS-POP, 2015)
- **WorldPop:** Machine learning algorithms are used to produce population grid. Excludes a portion of 'rural' pop from census if satellite-derived image does not show built-up areas². (WorldPop, 2015)



From Satellite Image to Raster Data of Built-up Areas
(GHSL, n.d.)



What is Raster Data? (ESRI, n.d.)

¹ https://ghsl.jrc.ec.europa.eu/ghs_pop2019.php

² <https://hub.worldpop.org/project/categories?id=18>

Methodology

Degree of Urbanization (DOU)

Spatial Analysis

Dartboard (DB)

PART.02



Degree of Urbanization (DOU)

PROVIDES A GLOBAL EXTERNAL PERSPECTIVE

- **Urban:** minimum 300 people/km² and a minimum of total population of 5000 people
- **High Density Urban:** minimum of 1500 people/km² and a minimum of 50,000 people total population
- Developed by Dijkstra & Poelman (2014)

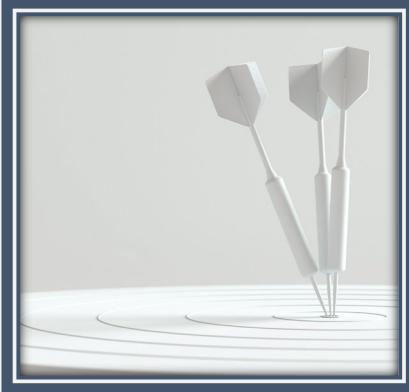


GEOSPATIAL ANALYSIS

- Create intersection of gridded population raster with boundary shapefile
- Find neighboring cells based on Queen's contiguity
- Create cluster based minimum population thresholds

DARTBOARD (DB)

PROVIDES STATISTICALLY SIGNIFICANT RELATIVE PERSPECTIVE



- Compute building density per pixels, smooth density across pixels using a kernel, randomly reshuffles pixels 2,000 times over all habitable pixels
- **Counterfactual:** calculates non-buildable areas due to water, elevation, slope and desert. Applies statistical methods to find density greater than 95% of counterfactual
 - **Urban Areas:** Contiguous pixels where pop density is greater than 95th percentile of counterfactual
 - **Cores:** Urban areas that have a core
 - **Cities:** Contiguous pixels within urban areas that are above the 95th percentile of the counterfactual within the urban areas
- **Cores vs Cities:** Cores are relative to all urban pixels in the country. Cities are defined relative to their own urban areas.

Research Questions

01

Methodology

H_0 : There is no relationship between urban estimate & methodology (DOU / DB)

H_a : A relationship exists between urban estimate and methodology (DOU / DB)

02

Gridded Population Data Source

H_0 : There is no relationship between urban estimate & data source (GHS-POP / WorldPop)

H_a : A relationship exists urban estimate and data source (GHS-POP / WorldPop)

03

Size of Population Data Grid

H_0 : There is no relationship between urban estimate & grid resolution (1 km x 1 km / 250 m x 250 m)

H_a : A relationship exists between urban estimate & grid resolution

Results

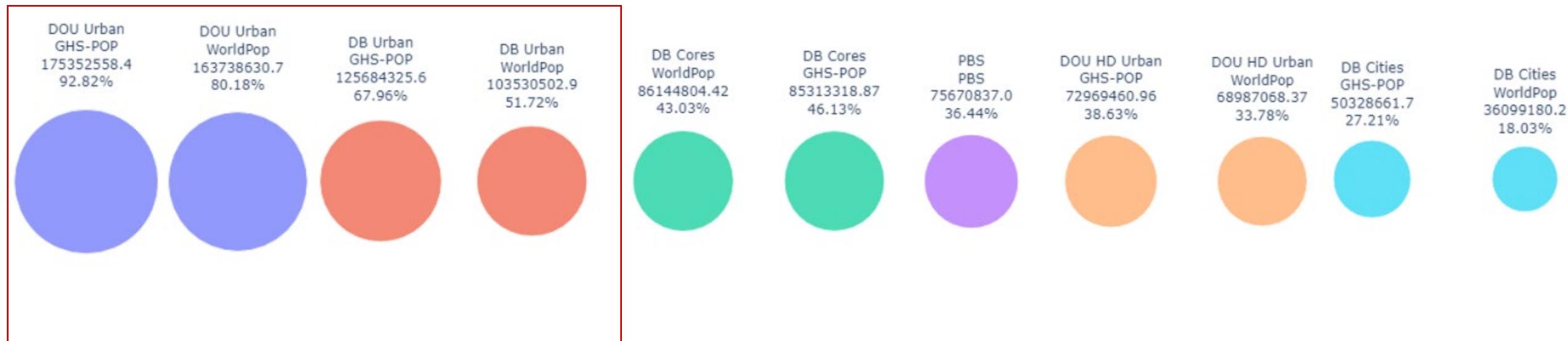
PART.03



Urban Estimates order by Population

- Results similar for 250 m x 250 m grid size
- Definition matters here

Urban Percent (1 km²)

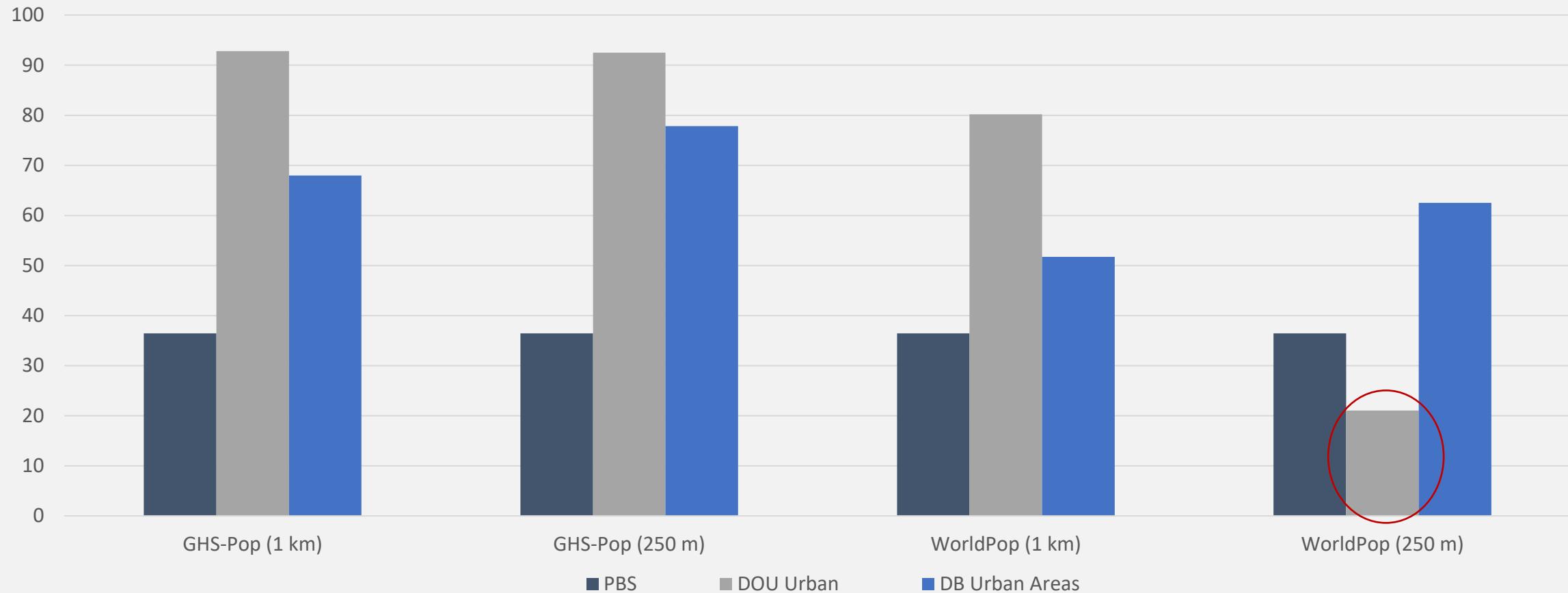


DOU = Degree of Urbanization

DB = Dartboard

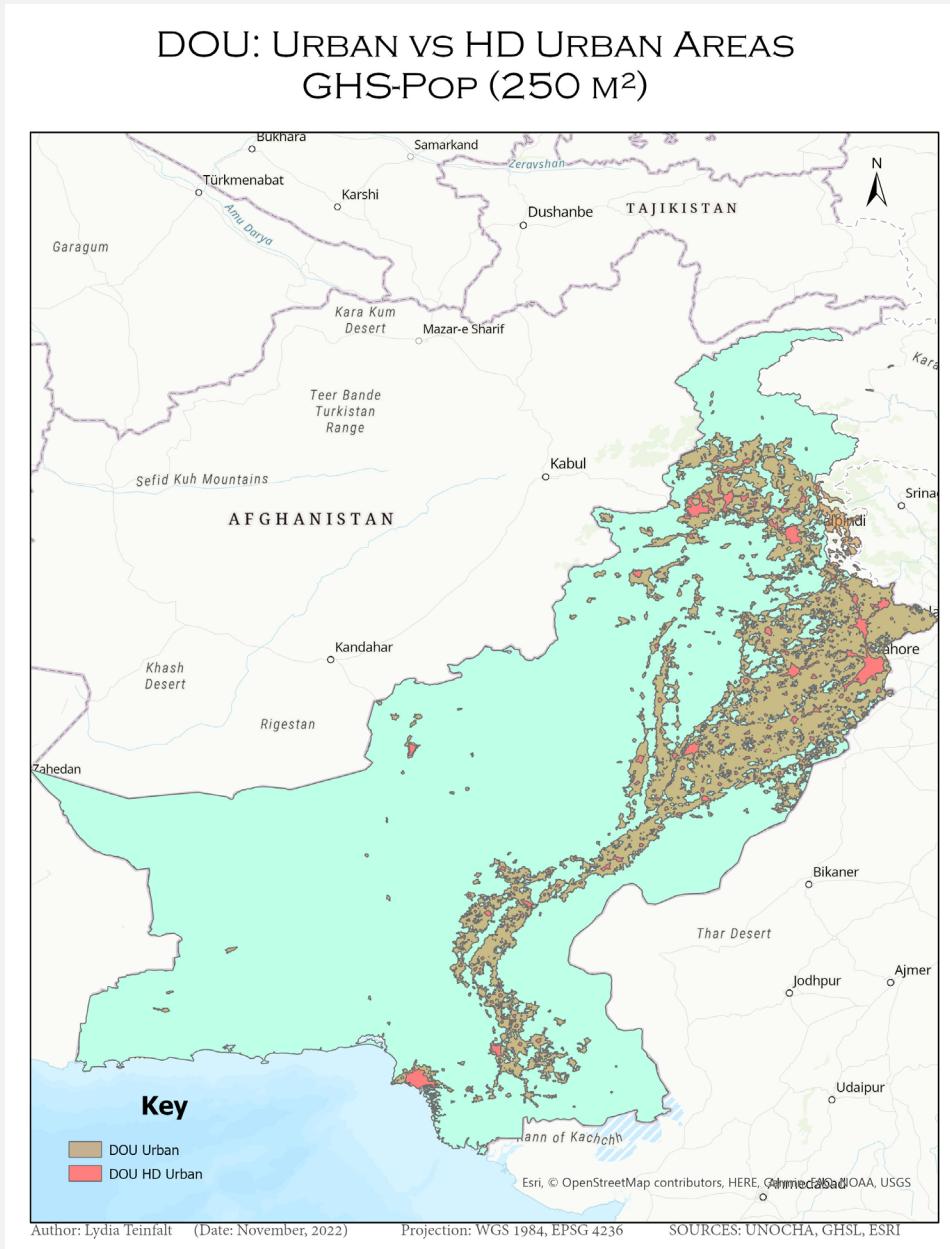
PBS = Pakistan Bureau of Statistics

Pakistan's Percent Urban Areas - National Level



PBS: Pakistan Bureau of Statistics

DOU: GHS-Pop and WorldPop (250m x 250 m)

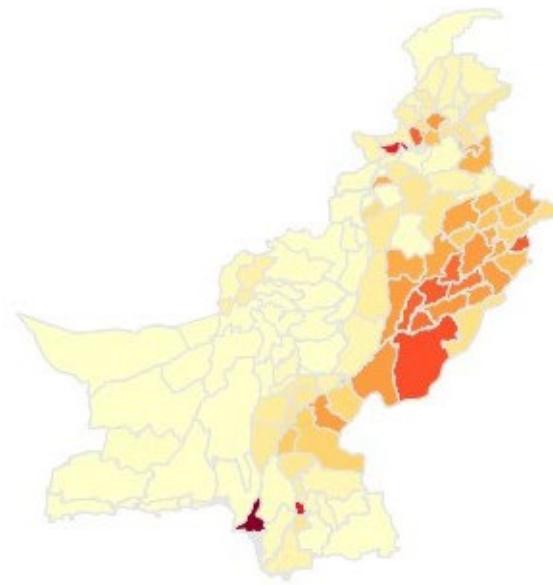


DOU: GHS-Pop (1 km) Administrative Levels 1-3

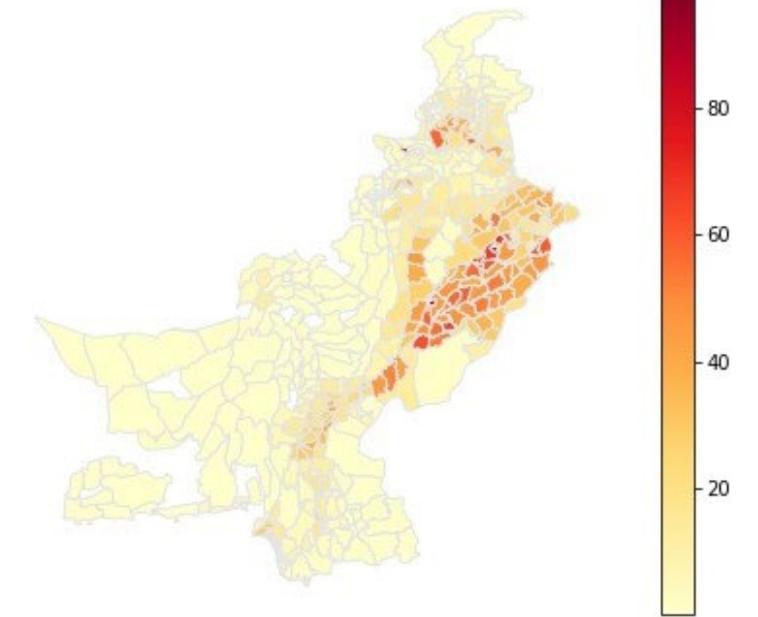
PROVINCE



DISTRICT



TEHSIL



Research Findings

01

Methodology

H_0 : There is no relationship between urban estimate & methodology (DOU / DB)

H_a : A relationship exists between urban estimate and methodology (DOU / DB)

02

Gridded Population Data Source

H_0 : There is no relationship between urban estimate & data source (GHS-POP / WorldPop)

H_a : A relationship exists urban estimate and data source (DOU GHS-POP / WorldPop)

03

Size of Population Data Grid

H_0 : There is no relationship between urban estimate & grid size (1 km x 1 km / 250 m x 250 m)

H_a : A relationship exists between urban estimate and grid size (for WorldPop)

Summary

PART.04



Findings/Discussion

URBAN DELINEATION APPROACHES:

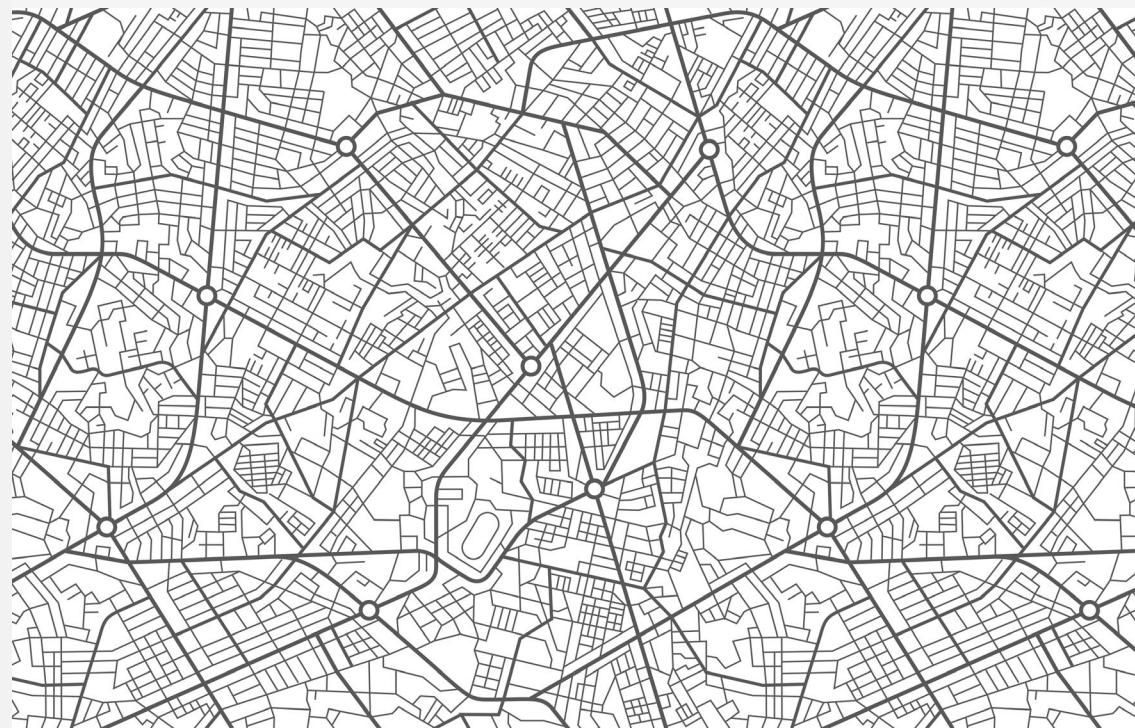
- DOU finds 80-93% of Pakistan's population living in urban areas.
- In comparison to the 36.4% figure from PBS, DB shows that statistically significant urban areas to be 52-68%.

POP DATA SOURCES:

- GHS-POP & DOU: urban rate is 93%
- WorldPop & DOU: urban rate is 80%

GRID RESOLUTION:

- DOU & WorldPop (250 m x 250 m): 21% urban
- DOU & WorldPop (1 km x 1 km): 80% urban





Summary

- Gridded population is not a replacement for census data.
- Both DOU and DB serve as benchmarks and highlight the need to move beyond binary classification of urban/rural.
- Results based on geospatial analysis conducted at national level but may vary at province, district, and tehsil levels.

Thank you!

References

1. One UN Pakistan : Annual Report 2021. (n.d.). Pakistan. <https://pakistan.un.org/en/197948-one-un-pakistan-annual-report-2021>
2. Bellefon, M. P., Combes, P. P., Duranton, G., Gobillon, L., & Gorin, C. (2021, September). Delineating urban areas using building density. *Journal of Urban Economics*, 125, 103226. <https://doi.org/10.1016/j.jue.2019.103226>
3. Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., & Schiavina, M. (2021, September). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*, 125, 103312. <https://doi.org/10.1016/j.jue.2020.103312>
4. What is raster data?—ArcMap | Documentation. (n.d.). Retrieved October 28, 2022, from <https://desktop.arcgis.com/en/arcmap/latest/manage-data/raster-and-images/what-is-raster-data.htm>
5. Pakistan's Country Fact Sheet. (n.d.). European Commission Global Human Settlement Layer. Retrieved October 24, 2022, from <https://ghsl.jrc.ec.europa.eu/CFS.php>
6. Final Results (Census-2017). (n.d.). Pakistan Bureau of Statistics. Retrieved November 1, 2022, from <https://www.pbs.gov.pk/content/final-results-census-2017>