6/26/2023

Capstone Project
(Data Engineering and Intelligence Unit)

Ajeyomi Adedoyin Samuel
Adedoyinsamuel25@gmail.com

Table of Contents

Task Guideline 2

ASSESSMENT OF HEALTH FACILITY LOCATIONS AND DATA VALIDATION IN BORNO STATE 3

Problem with existing data 4

Data Used and sources 4

Methodology: 4

Results and Findings 5

DASHBOARD 12

Findings 13

Recommendations 14

Task Guideline

Objective: To test the level of expertise across

- Data acquisition
- Data analysis
- Data Visualization
- Data interpretation (insight generation)

Data Category: Health, Energy, Religion, Agriculture, Hospitality, Finance

Capstone Rule:

- All datasets must be from Nigerian geography.
- Try to make your deliverable as native to the Nigerian ecosystem as possible.
- Be original in your deliverable.

Task One: Select one or more data categories, search the internet, and acquire enough geospatial and statistical dataset

Deliverables:

- Acquired dataset and their referenced source

Task Two: Clean and analyze the acquired dataset, generation at least 10 analytical outputs you believe are useful

(*hint: try to focus on what can help create solution, impact or business intelligence*).

Deliverable:

- Cleaned dataset
- Analytical report and method used

Task Three: Create a visualization based on the analysis done, visuals should be in both statistical charts and maps.

(*hint: your output here can be either static or interactive or both)*

Deliverables:

- Statistical Chart(s)
- Embellished map(s)

Task Four: Interpret maps and charts in a summarized format (no more than 500 words) providing a relatable context on how the visualization relates to reality in terms of solution, impact, or profitability

Deliverables:

- A document having all visualization with their appending interpretation

CATEGORY 1: HEALTH

ASSESSMENT OF HEALTH FACILITY LOCATIONS AND DATA VALIDATION IN BORNO STATE

Introduction

Access to quality healthcare services is vital for the well-being and development of any location, especially in regions affected by conflicts and emergencies. Borno State in Nigeria has faced significant challenges in delivering adequate healthcare due to various conflict that has occur in the state in the past. The displacement of thousands of people and the destruction of critical infrastructure, including health facilities, have exacerbated the healthcare crisis. To address these challenges and improve healthcare delivery, it is imperative to assess the existing health facility locations and validate the available data to guide resource allocation and enhance access to healthcare services.

The accuracy and reliability of data are paramount when making informed decisions, particularly in the healthcare sector. In the context of Borno State's healthcare challenges, there is a need to question the integrity of health facility location data collected through surveys. This is crucial to ensure that the information used for planning health solutions and allocating resources is accurate and trustworthy.

In our analysis, we aim to validate the collected health facility data to establish its reliability. One critical aspect of data validation involves examining the population residing in proximity to the reported health facility locations. By overlaying population data onto satellite imagery, we can cross-reference the existence of these facilities with the actual distribution of the population. This validation process enables us to identify any discrepancies or inaccuracies in the reported locations of health facilities.

By leveraging satellite imagery and employing spatial analysis techniques, we aim to validate the existence and accessibility of health facilities in Borno State. This rigorous approach allows us to determine the precise locations of these facilities, ensuring that the information used for healthcare planning and resource allocation is based on reliable and up-to-date data.

Our analysis seeks to provide valuable insights into the accuracy and integrity of health facility data in Borno State. By validating and, if necessary, invalidating certain data points, we contribute to enhancing the reliability of healthcare planning efforts. Our findings will guide stakeholders, including policymakers, healthcare providers, and organizations involved in post-conflict recovery efforts, in making informed decisions about resource allocation and developing targeted healthcare interventions that address the needs of the affected population.

Ultimately, through rigorous data validation, we aim to ensure that healthcare solutions in Borno State are based on accurate information. This approach will contribute to improving access to quality healthcare services, supporting efficient resource allocation, and fostering the overall well-being and development of the region.

Problem with existing data

- Inaccurate Reporting: One of the key problems with existing healthcare facility data is the presence of inaccurate reporting. This means that the reported locations of health facilities may not align with their actual geographic coordinates, leading to a distorted representation of the healthcare landscape.
- Outdated Information: The data may suffer from outdated information, failing to capture the changes that have occurred over time. This is particularly relevant in regions affected by conflicts or emergencies, where displacement and infrastructure destruction can significantly alter the healthcare landscape.
- Limited Validation Processes: The validation processes for healthcare facility data are often insufficient, resulting in a lack of accuracy and reliability. Without robust mechanisms to verify the reported data, its integrity cannot be assured, hindering effective decision-making and resource allocation.

Data Used and sources

- E-health health facilities dataset - https://data.humdata.org/dataset/fea18f4e-0463-4194-a21c-602e48e098e1/resource/ce73440b-3fa2-46c0-8613-b674c8c42d45/download/nigeria.geojson

- Borno Administrative Boundary dataset – https://data.humdata.org/dataset/nigeria-borno-operational-boundaries-admin-1-3-state-lga-ward-archived

- Grid3 Population Estimate data – https://data.grid3.org/maps/GRID3::nigeria-gridded-population-estimates-version-02

Methodology:

- Data Collection: The existing healthcare facility data for Nigeria was obtained from E-health Africa dataset. The dataset included information on the location of health centers in the country.

- Buffer Creation: A buffer of 1km radius was created around each health center location using geospatial analysis techniques. This buffer represented the area within which the population was assessed for healthcare access.

- Population Extraction: Within each 1km buffer, the population data was extracted from relevant sources, such as census data or population databases. This provided an estimate of the population residing within the vicinity of each health center.

- Validation Threshold: A threshold was set to determine the validity of each location. Any health center location with a population of less than 100 people within the 1km buffer was considered invalid. This threshold was established to ensure if a health center truly exist in that said location due to the integrity of the data.

- Validation with Grid3 Population Dataset: The Grid3 raster population dataset was utilized as a validation dataset to cross-reference the total population within each area. By comparing the population extracted from the health center dataset with the Grid3 dataset, the accuracy and reliability of the health facility data were evaluated.

- LGA Population Comparison: The population data extracted from the Grid3 dataset was also used to estimate the population of each Local Government Area (LGA). This information was then compared with the total number of health facilities within each LGA to assess the average population that should have access to a health facility based on the population data.

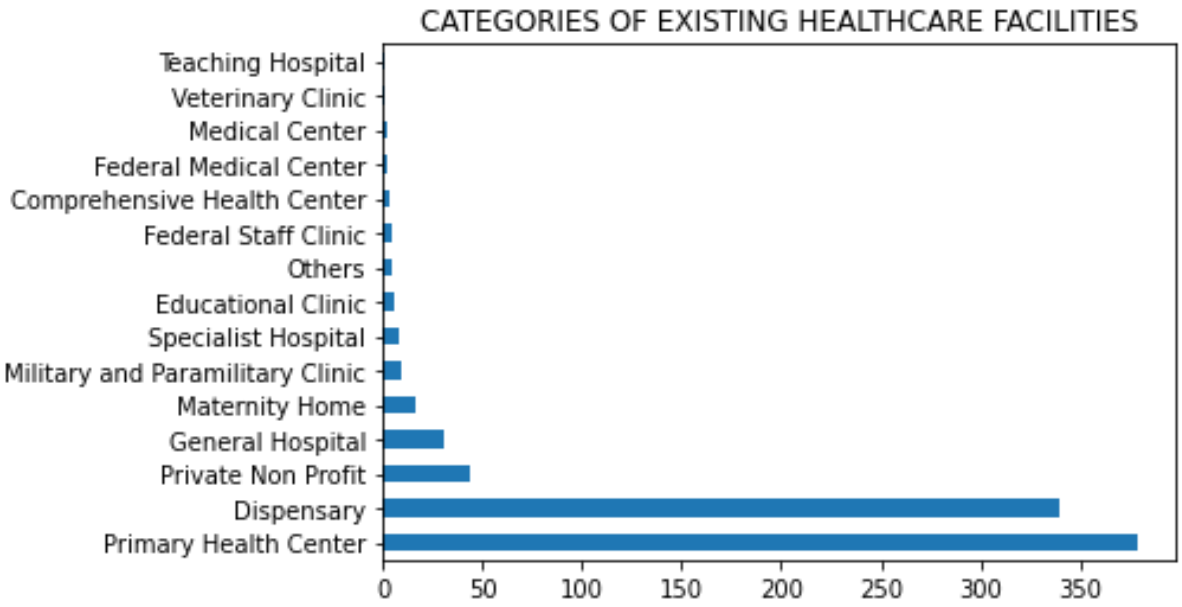
- Invalidated Data Analysis: The percentage of invalidated data, i.e., health center locations that does exist in real time but was collected by enumerators as existing health facilitates location. This provided insights into the overall data integrity issues within the healthcare facility dataset.

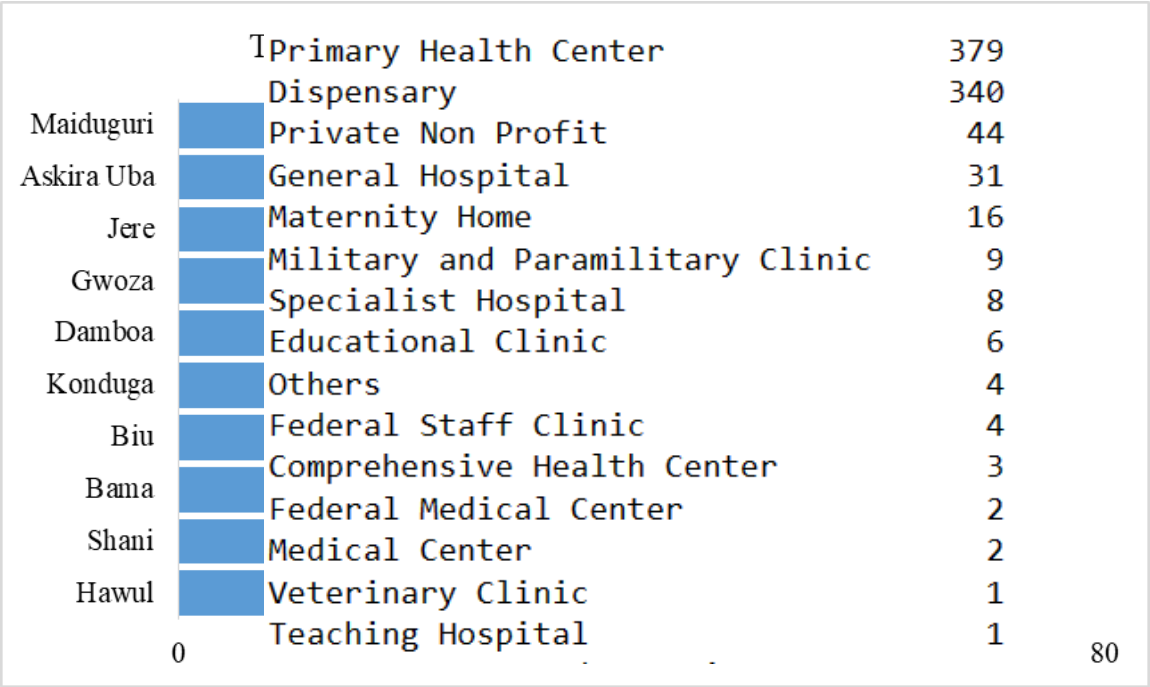
Results and Findings

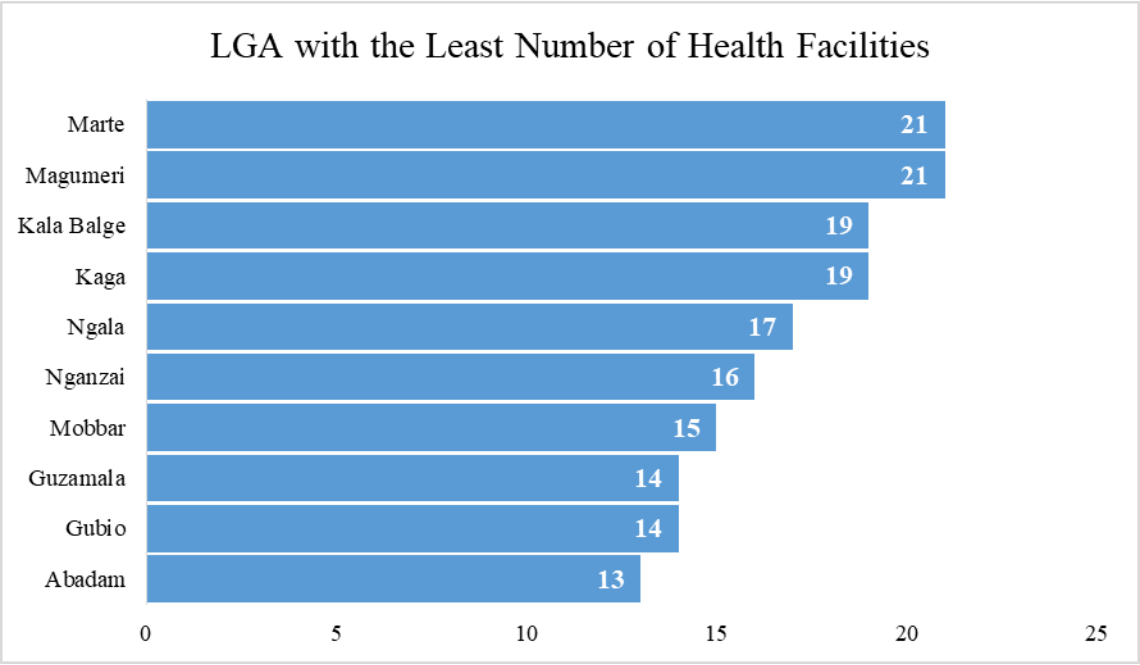
LGA and Total Number of Facilities locations

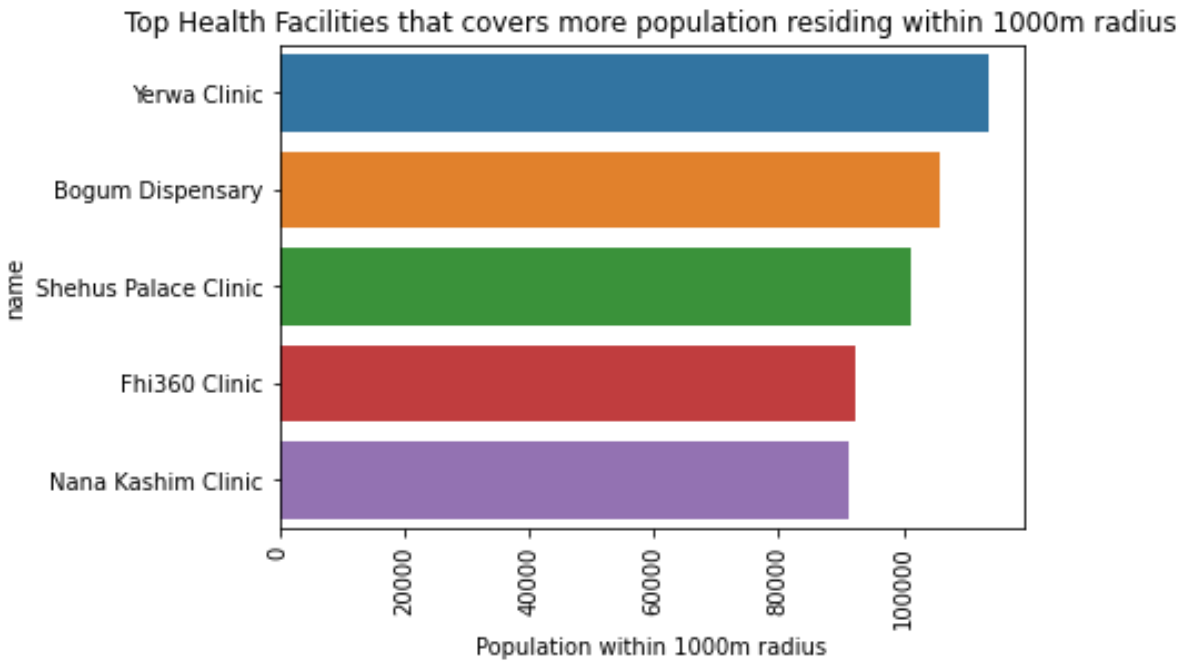
LGA	No Health Facilities
Abadam	13
Askira Uba	68
Bama	43
Bayo	34
Biu	43
Chibok	32
Damboa	48
Dikwa	27
Gubio	14
Guzamala	14
Gwoza	50
Hawul	40
Jere	58
Kaga	19
Kala Balge	19
Konduga	46

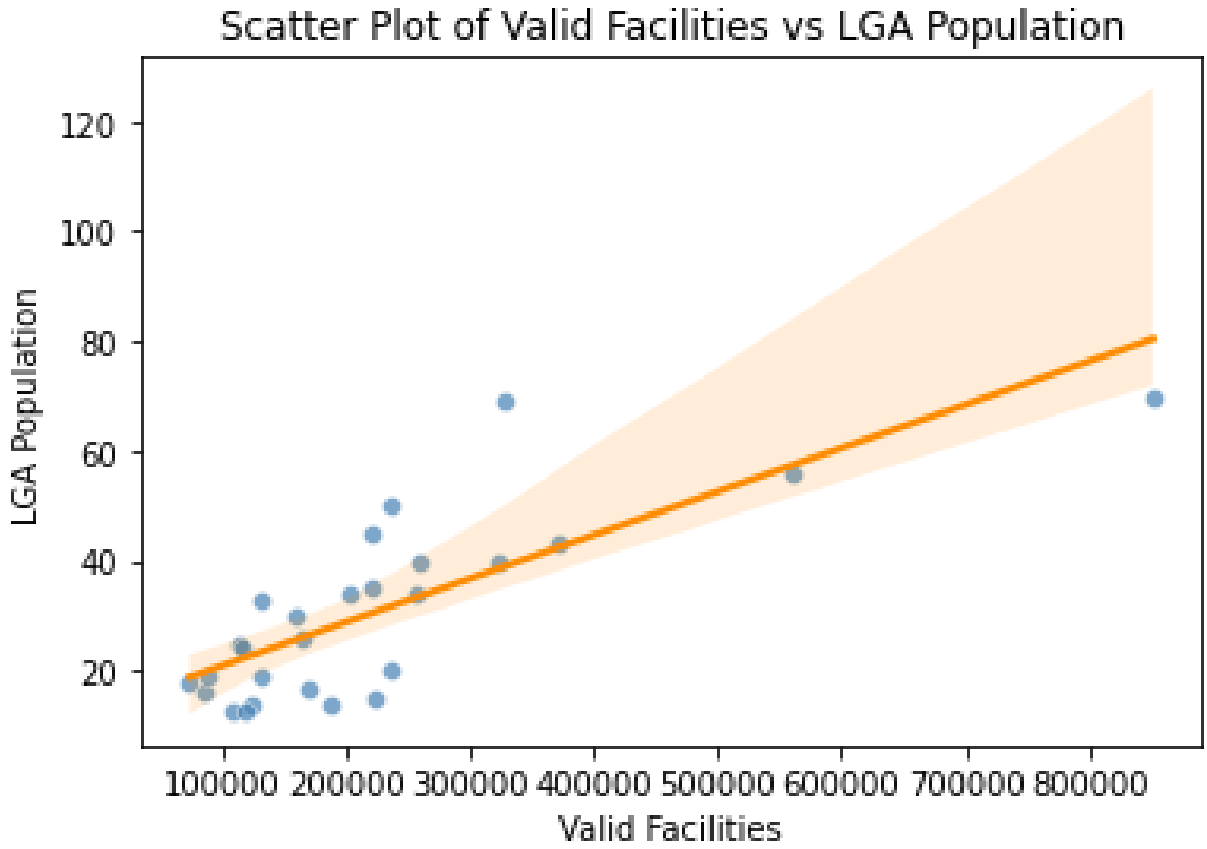
Kukawa	26
Kwaya Kusar	30
Mafa	31
Magumeri	21
Maiduguri	69
Marte	21
Mobbar	15
Monguno	30
Ngala	17
Nganzai	16
Shani	40
Total	**884**

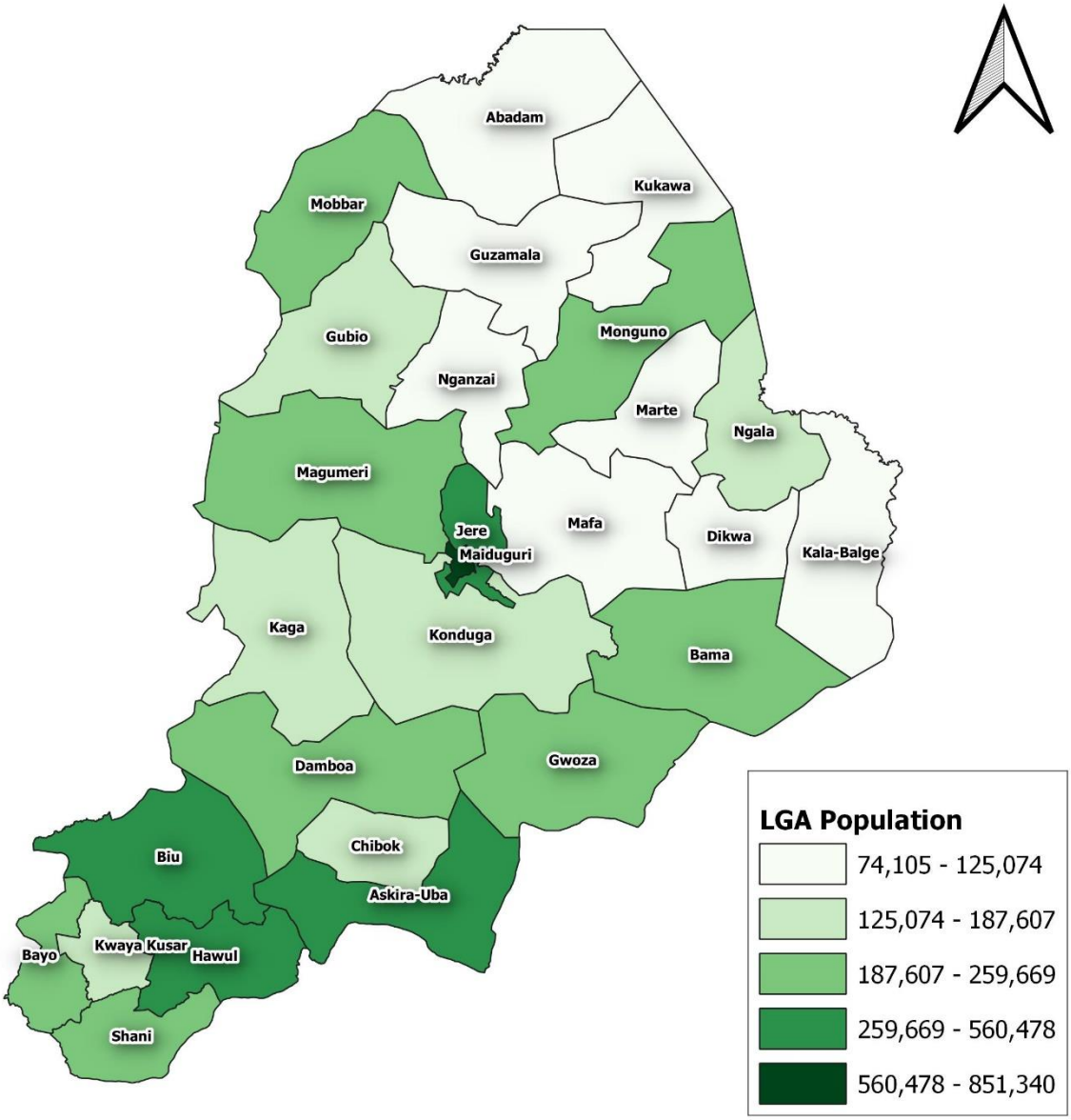



Primary Health Center 379
Dispensary 340
Private Non Profit 44
General Hospital 31
Maternity Home 16
Military and Paramilitary Clinic 9
Specialist Hospital 8
Educational Clinic 6
Others 4
Federal Staff Clinic 4
Comprehensive Health Center 3
Federal Medical Center 2
Medical Center 2
Veterinary Clinic 1
Teaching Hospital 1
Maiduguri
Askira Uba
Jere
Gwoza
Damboa
Konduga
Biu
Bama
Shani
Hawul
0
80

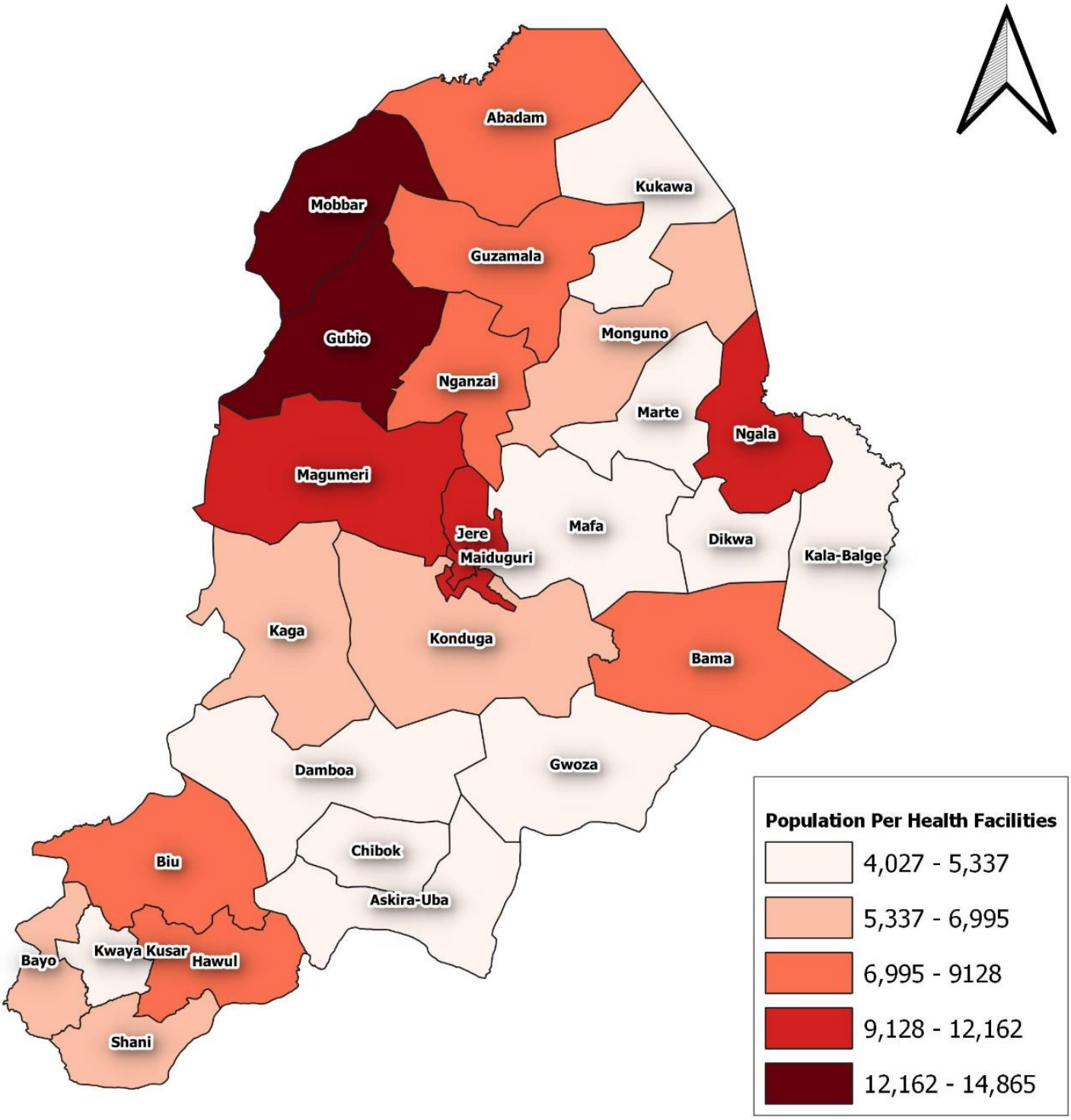

LGA with the Least Number of Health Facilities
Marte 21
Magumeri 21
Kala Balge 19
Kaga 19
Ngala 17
Nganzai 16
Mobbar 15
Guzamala 14
Gubio 14
Abadam 13
0
5
10
15
20
25


Top Health Facilities that covers more population residing within 1000m radius
Yerwa Clinic
Bogum Dispensary
Shehus Palace Clinic
Fhi360 Clinic
Nana Kashim Clinic
name
0
20000
40000
60000
80000
100000
Population within 1000m radius

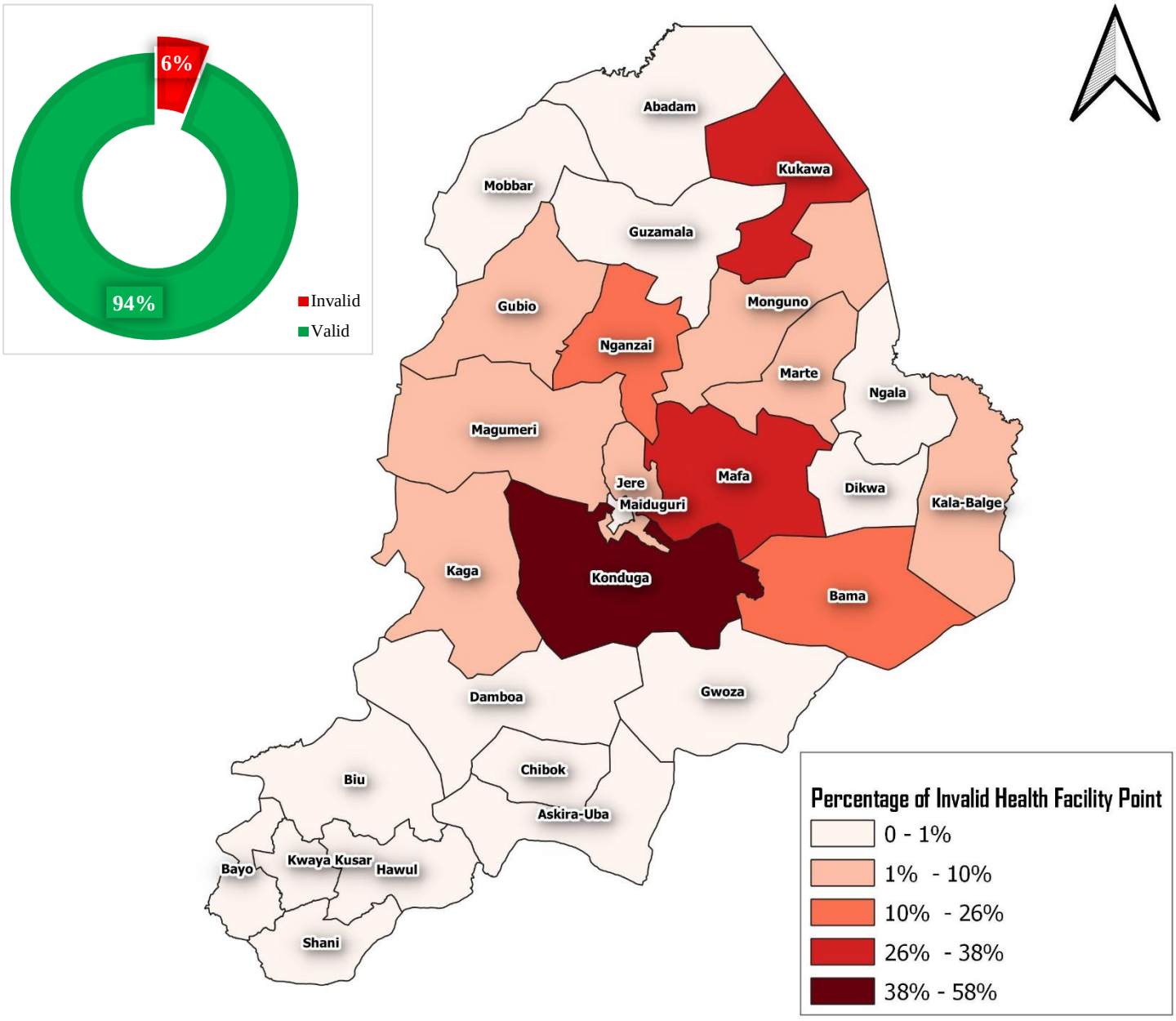

Scatter Plot of Valid Facilities vs LGA Population
LGA Population
120
100
80
60
40
20
100000 200000 300000 400000 500000 600000 700000 800000
Valid Facilities



Map showing LGA Population of Borno State.



Map showing Population Per Health Facilities

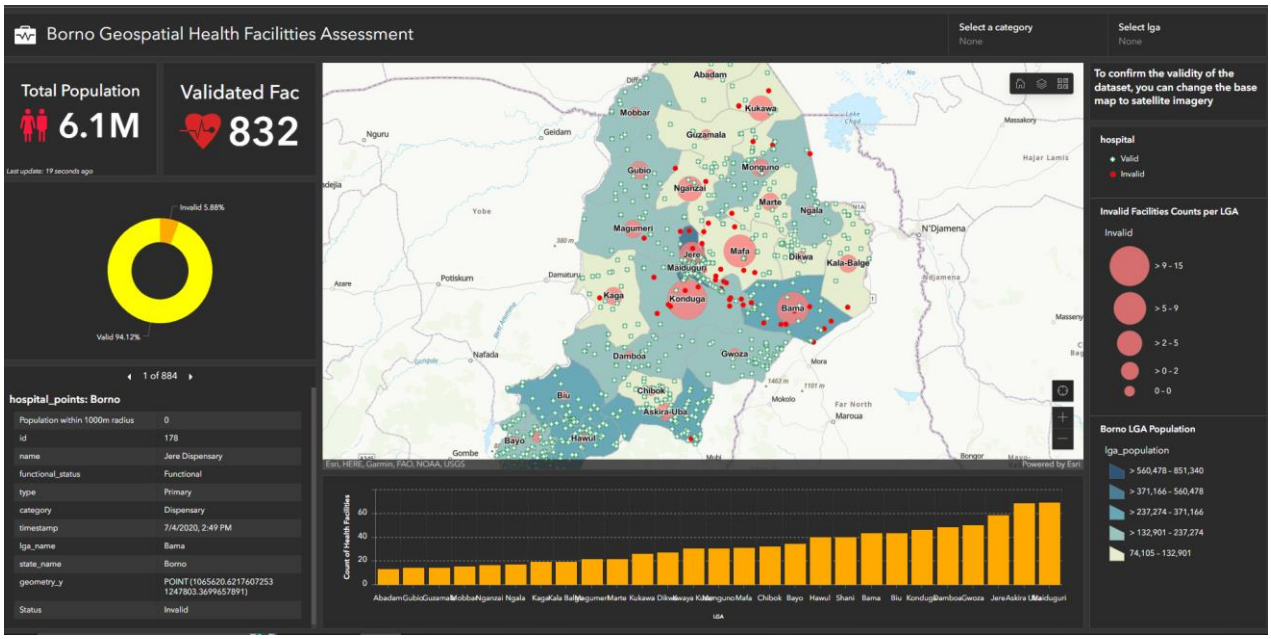


Map showing the total number of invalid health facility point per LGA

LGAName	Invalid	Valid	Percentage of invalid	lga_population
Konduga	15	26	58	164643
Mafa	9	24	27	115344
Kukawa	5	16	24	85359
Bama	9	34	21	258106

Nganzai	3	13	19	108600
Magumeri	2	20	9	235283
Gubio	1	14	7	187607
Kala-Balge	1	18	5	74105
Jere	3	56	5	560478
Kaga	1	19	5	132899
Marte	1	19	5	87701
Monguno	1	35	3	220503
Askira-Uba	1	69	1	329181
Bayo	0	34	0	202314
Damboa	0	45	0	220173
Chibok	0	33	0	132901
Hawul	0	40	0	323674
Dikwa	0	25	0	114713
Gwoza	0	50	0	237274
Guzamala	0	14	0	125074
Mobbar	0	15	0	222978
Maiduguri	0	70	0	851340
Kwaya Kusar	0	30	0	160104
Abadam	0	13	0	118661
Ngala	0	17	0	170496
Shani	0	40	0	259669
Biu	0	43	0	371166

DASHBOARD



Link to Dashboard:
https://www.arcgis.com/apps/dashboards/95342543f17e4cf3b0b22a7e9d6e94f2

Findings

During the validation process, it has come to light that the e-health Africa dataset raises significant concerns about its integrity. A substantial number of reported health facility points are located within forested or bush areas, casting doubt on the reliability of the dataset for effective health planning and resource allocation. It emphasizes the importance of conducting thorough validations to ensure that the information used for decision-making is accurate, reliable, and aligned with the actual needs of the population.

Based on the collected data, Maiduguri emerges as the top local government area (LGA) with the highest number of health facilities, totaling 69. Following closely are Askira Uba with 68 facilities, Jere with 58, Gwoza with 50, Damboa with 48, and Konduga and Biu with 46 each. Hawul and Shani both have 40 health facilities.

Conversely, the LGAs with the fewest health facilities are Magumeri and Marte, each having 21 facilities. Kala Balge and Kaga have 19 facilities each, Ngala has 17, Nganzai has 16, and Mobbar has 15. Guzamala and Gubio both have 14 facilities, while Abadam has the lowest number with only 13.

It is important to note that the provided data does not include information about the population served by each health facility. However, according to the grid3 gridded population estimates, the total estimated population of Borno State is 6,070,348.

To conduct a more comprehensive analysis and determine if certain areas are underserved, further investigation is necessary. This analysis could consider geospatial factors such as population density in each LGA, the proximity between health facilities, and other variables that influence access to healthcare services. Additionally, incorporating location data and utilizing geospatial modeling techniques can provide a more accurate assessment of the number of hospitals required in each location to adequately serve the population.

Regarding the validation of the health facility data, it is worth noting that 93% (832) of the collected data points are considered valid, while 7% (52 points) are deemed invalid based on the validation threshold. Further analysis reveals that Konduga has the highest number of invalid health facility points registered in the database. Consequently, the validation of data in this particular location should be scrutinized. Utilizing an interactive dashboard can offer users an advantage in exploring and visualizing the data more effectively. Users can even switch the dashboard basemap to satellite imagery to verify real-world conditions.

The population map indicates that Maiduguri has the highest population among the LGAs. To identify underserved areas, the population per health facility can be calculated by dividing the total population by the total number of hospitals. From the map, it is evident that locations such as Gubio, Mobbar, Magumeri, Jere, and even

Maiduguri itself—despite having the highest number of health facilities—are underserved. Enhancing the validation process and improving the findings will contribute to a more comprehensive and accurate assessment.

Recommendations

To address the challenges of data integrity, the key components of this recommended are as follows:

- Data Validation Services: Comprehensive data validation by employing a team of experts proficient in data validation techniques. Through desk-based verification, on-site visits, and field surveys, the organization would thoroughly validate the reported health facility locations, cross-referencing them with reliable sources of information. This validation process would enhance the integrity of the dataset, enabling stakeholders to rely on accurate and up-to-date healthcare facility data.

- Collaborative Partnerships: Establishing collaborative partnerships with key stakeholders involved in healthcare planning and delivery in Borno State. This would include government agencies, healthcare providers, NGOs, and local authorities. By working closely with these partners, the organization would gain access to relevant data sources and foster a collaborative approach to data integrity. Through these partnerships, stakeholders can collectively address the challenges of data accuracy and build a shared understanding of the healthcare landscape in the region.

- Technological Solutions: Leveraging advanced technologies, such as remote sensing, satellite imagery, and geospatial analysis, to enhance the accuracy and efficiency of data validation processes. By harnessing these technologies, one can identify discrepancies between reported data and the physical environment, ensuring greater precision and reliability in validating healthcare facility locations.

- Standardized Protocols and Capacity Building: Developing standardized protocols and provide capacity building programs to enhance data collection and validation processes. This would involve defining clear guidelines, implementing quality control measures, and offering training programs for data collectors and stakeholders involved in data management.
- Continuous Monitoring and Maintenance: Establishing a system for continuous monitoring and maintenance of healthcare facility data. Regular updates, periodic audits, and feedback mechanisms would be implemented to promptly identify and rectify any discrepancies or changes. This monitoring and maintenance system would ensure that the dataset remains accurate, reliable, and up-to-date, enabling stakeholders to make well-informed decisions based on the most current information available.

- Data Analysis and Reporting: In addition to data validation services, the organization would provide data analysis and reporting capabilities. By analyzing the validated data, the organization would generate meaningful insights and reports that help stakeholders understand healthcare trends, gaps, and priorities. These insights would guide informed decision-making processes, leading to improved healthcare planning, resource allocation, and service delivery in Borno State.

.