Final Project – Paper

# Spatial Analysis of Ethiopia's Health center locations using R:

*Topic of Focus: Availability of Primary health care (PHC) access throughout Ethiopia using open source data*

By: Kidus Y Berhe

UNL: GEOG 891 Special Topics - Spatial Analysis in R

# Introduction:

Any nations that wish to ensure a stable environment for their population must take steps to provide basic health care to its population in some form or manor, however the differences in health care access around the world is startling. While some nations are able to provide universal health care to a certain degree such as Austria, others struggle to provide similar services and so must focus on populated regions, such as Ethiopia.

Health care, or more specifically Primary Health Care (PHC), is defined by the World Health Organization (WHO) as addressing the majority of an individual's health needs throughout their lives. PHC is distinct from secondary health care (SHC) that focuses on medical care that is provided by a specialist that requires more advanced procedures or checks (e.g.: Computed tomography (CT) scans). Access to PHC is considered a human right under the "Universal Declaration of Human Rights" and must be provided without any prejudices to race, sex and other personal defining characteristics. These criteria however do not normally apply to income or the ability to pay for proper PHC. As healthcare requires continuous financial backing, areas that are generally unable to pay for healthcare and so cannot keep the facilities running tend to be neglected, this is primarily seen in the rural regions of developing nations (Freeman et al, 2020).

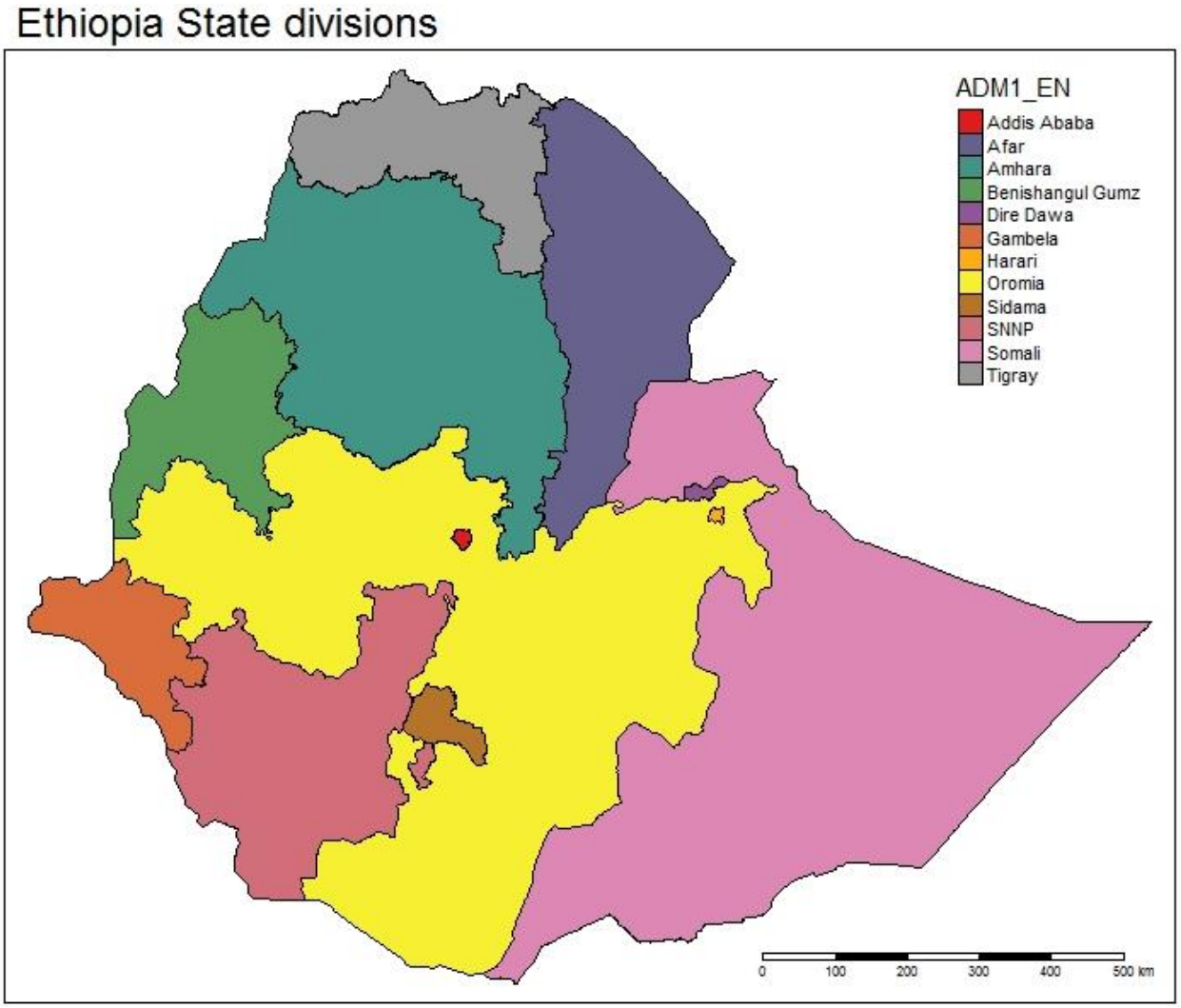
Health care, or lack of it, in Ethiopia has been a major issue that causes problems for the nation during times of stability at best, while suffering greatly during times of stress placed on the population's health. These stresses can be in the form of famines, local/endemic diseases, global pandemics like Covid-19, or natural disasters such as floods. This being said, healthcare in general within Ethiopia has improved over the years, bringing life expectancy from 38 years in the 1960s, to approximately 66 years as of 2018 (Freeman et al, 2020). The life expectancy increase can be attributed to the improvement of the quality of health care provided, as well as in part to the more available access to certain PHC such as HIV/AIDS detection and medication (Freeman et al, 2020). However, it is important to note that while approximately 80% of Ethiopia's population live in rural areas, 90% of the hospitals that can grant adequate healthcare are located in urban regions and while strides have been made to improve the distribution of PHC healthcare throughout the country, mismanagement and a population focus have resulted in only urban or close to urban regions attaining sufficient PHC. For many health issues access is still the main challenge in rural areas. A study on the socio-demographic factors affecting cervical cancer screening in Ethiopia (Woldetsadick et al, 2020) found that rural residence was significantly associated with the amount of screening in the study, in that women were less likely to receive a screening the further they lived into rural areas.

The hypothesis of this study/project will follow similar trends as those mentioned above where it is expected that rural areas have fewer and less evenly distributed health care, as well as have a less dense/lower percentage population.

# Background into Topic:

Ethiopia is an ancient land with deep rooted cultural and social similarities and differences found between the nation's ethnicities. This is so deep rooted that Ethiopia's states

are classified based of ethnic dispersion, resulting in states that have irregular shapes, as well as irregular sub-divisions. These divisions can be seen in the following map.




The issue of health care has been covered many times in an ongoing effort to produce information that may benefit the advancement of PHC and SHC as well as their access to as many people as possible. Most of the studies found are conducted using interviews and surveys on a specific diseases or issues that are affecting the populace of the study region, and this is no exception for the studies carried out in Ethiopia.

The first example of a study done in Ethiopia on the health care distribution is an assessment of the Health management Information System (HMIS) in public health centers of Addis Ababa, Ethiopia (Adane et al, 2017). This study was focused on determining how effective the HMIS is and how much it has progressed using daily health center data. The resulting conclusions were that the HMIS has was initiated but had not advanced much with the lack of technical and department level support being the major contributors (Adane et al, 2017). This is significant as even within the capital city of Addis Ababa, implemented programs on health care go miss-administrated, making it even harder for rural areas to receive similar

services. This study used primarily statistics generated on accumulated data on visits and data usage, without any mention on the spatial consequences and implications.

Another study focusing on the Physical, Behavioral and sociodemographic factors affecting hypertensions in adults for Nekemte town of western Ethiopia (Geleta et al, 2019). This study was carried out on a community based level, indicating the spatial extent of Nekemte town. The study found hypertension to be prevalent in adults 18 years or older with further sociodemographic and physiological factors influencing the results (Geleta et al, 2019). Though the findings were informative, they did not place a spatial application aside from screening accessibility as a factor. The town was assumed to be spatially constant with no information on the distribution of the population within the town and how this affects screening access.

A study conducted in Ethiopia on the patient volume and quality of PHC aimed to understand their relationship to provide insights for service delivery in Ethiopia (Arsenault et al, 2021). This study found that understanding this relationship would improve the distribution of PHC, stating that policy makers can use this to determine how to more evenly distribute services (Arsenault et al, 2021). This study focused more on the availability of healthcare staff to patients than the ability to access the health center in distant regions. They provided statistics to determine the relationship between rural and urban PHC and SHC however provided no spatial visualizations to interpret the data.

Much like Ethiopia, Nigeria has its own healthcare distribution issues, and the following study tried to analysis and represent the spatial distribution of health centers (Nwakeze and Kandala, 2011). This study carried out a great deal of analysis on the distribution of the health centers within Nigeria as well as used different criteria to understand how the distributions affect health care, from public or private ownership, to the quantity of health care centers in each state. These factors however were not shown in comparison to the nation's population dynamics that would also show relationships such as population densities of the different states (Nwakeze and Kandala, 2011). They do however use patients per medical practitioners though this is an unstable variable that can be influenced easily, unlike the amount of influence required to alter population dynamics.

Finally a study conducted in Germany to determine the spatial distribution of health care and the relationship with government structure show another way of approaching the issue spatially (Bennema-Broos, M, 2001). Though this study is not as applicable to Ethiopia due to the difference in government types and the roles they play with German, the paper offers insight into other factors aside from socio-demographic or solely spatial influences that would affect health care distribution. The study's conclusion found that certain government types tend to provide differing rates of health care access, with more "left-wing" regions enjoying a more even distribution (Bennema-Broos, M, 2001). The study using interesting and useful data however does not represent any of it spatially only using statistical comparisons to come to a result.

As can be seen from the examples above, health care distribution and its effectiveness/access can be and has been studied for many years, using many different perspectives. From the studies indicated above, a noticeable trend in health care distribution can be noticed, that health care requires a multi viewed analysis, from as many fields that affect it to

be able to understand how to provide PHC and SHC to more people that need it, especially in developing nations where factors may over lap and so result in confusion.

## Methods:

Data used:

- Ethiopia administrative boundaries
  - (OCHA admin, 2021)
- Ethiopia Health Facility
  - (OCHA health, 2021)
- Ethiopia Health sites
  - (OCHA sites, 2021)
- Ethiopia District capitals
  - (OCHA cities, 2021)
- Ethiopia population data, 2021 estimates
  - (OCHA pop, 2021)

Software used:

- Rstudio (main)
  - Packages:
    - Tidyverse
    - Dyplyr
    - Sf
    - Tmap
    - Raster
    - spdep
    - spData
    - RcolourBrewer
- QGIS
  - Preliminary data viewing and spatially joining health sites and health facilities for point data as R studio could not run the required code
- Excel
  - Viewing population data and selecting required sheets

Projection throughout the project:

- Africa Equidistant Conic ESRI:102023
  - Datum: WGS84

Methodology:

The project started with acquiring the data that was needed for the process to follow. These datasets had to be open source so that the idea for quick and easily accessible code that would not cause copyright issues with use. The overarching goal of the project code was to be able to provide an r based code that could be easily adapted for different statistics due to the data all located together and being accessible. The data collected is noted above and contained shape and csv files that were used in future analysis.

QGIS was used to merge the health site data as r studio found difficulties performing the same spatial task for points. This and initial visualizations of the shapefiles where the only use of

QGIS. The created file is located in the same repository and so should not cause issues when used with the code.

The selected data was then read into r studio using the projection stated above (Africa Equidistant Conic) to preserve distance as it would be used later. The datas were then validated and checked. Following this the shapefiles and the population files were joined using the administrative level 3 names for the districts (also known as Woredas) and the administrative level 2 names for the Zones.

The joined data was then used to calculate the percentage of population to total that each district and zone had. Alongside this, the population density of the districts and zones, similar values but of different administrative boundary levels, were calculated using the sf package's "st_area" to spatially calculate the area each polygon made up. These values were significantly different from the provided values in $km^2$ from the shapefile. After a quick comparison, the calculated values were seen to be more precise to recorded areas for districts and zones.

The calculated data was then quickly plotted to see the distribution. The population density measurements were excessively skewed due to Addis Ababa's (capital city) large population to area ratio. Because of this, Addis Ababa was removed from the population density maps to better represent to rest of the nation.

The health center data was then used to estimate the number of health care sites to districts for a base comparison.

A 20km buffer was created around the health center points. This buffer was decided upon as the time an average person would take to complete 20km back and forth (40km max in total) within the span of the day, as night time treks would be hazardous. In total, the journey would take 8 – 12 hours, assuming variations in geography and road networks.

Following this the spatial analysis of the data began. First the "NA" values were removed from the required data, as most of the analysis required full data. There was only one row that contained "NA" and this was filtered out. This data corresponded to a national nature reserve and so would not influence the population data. From here the neighbours of the district polygons were found, a spatial weights matrix using style "W" was row standardized and the counts of neighbours plotted on a histogram to show the distribution of neighbours.

Moran's I was then calculated using two different functions. The first was moran.test, which provided a strange p-value as it was far too small for use. The second was moran.mc which computed the required values analytically to provide a more believable and expected p-value. The morans'I was plotted showing the distribution of district neighbours and their associated MI value.

The final step was to find the Local Indicator of Spatial Autocorrelation (LISA) of the population percentage of the total population and plot it to show the clustering of populations throughout the nation and the areas that share spatial similarities across the calculated data.

# Results:

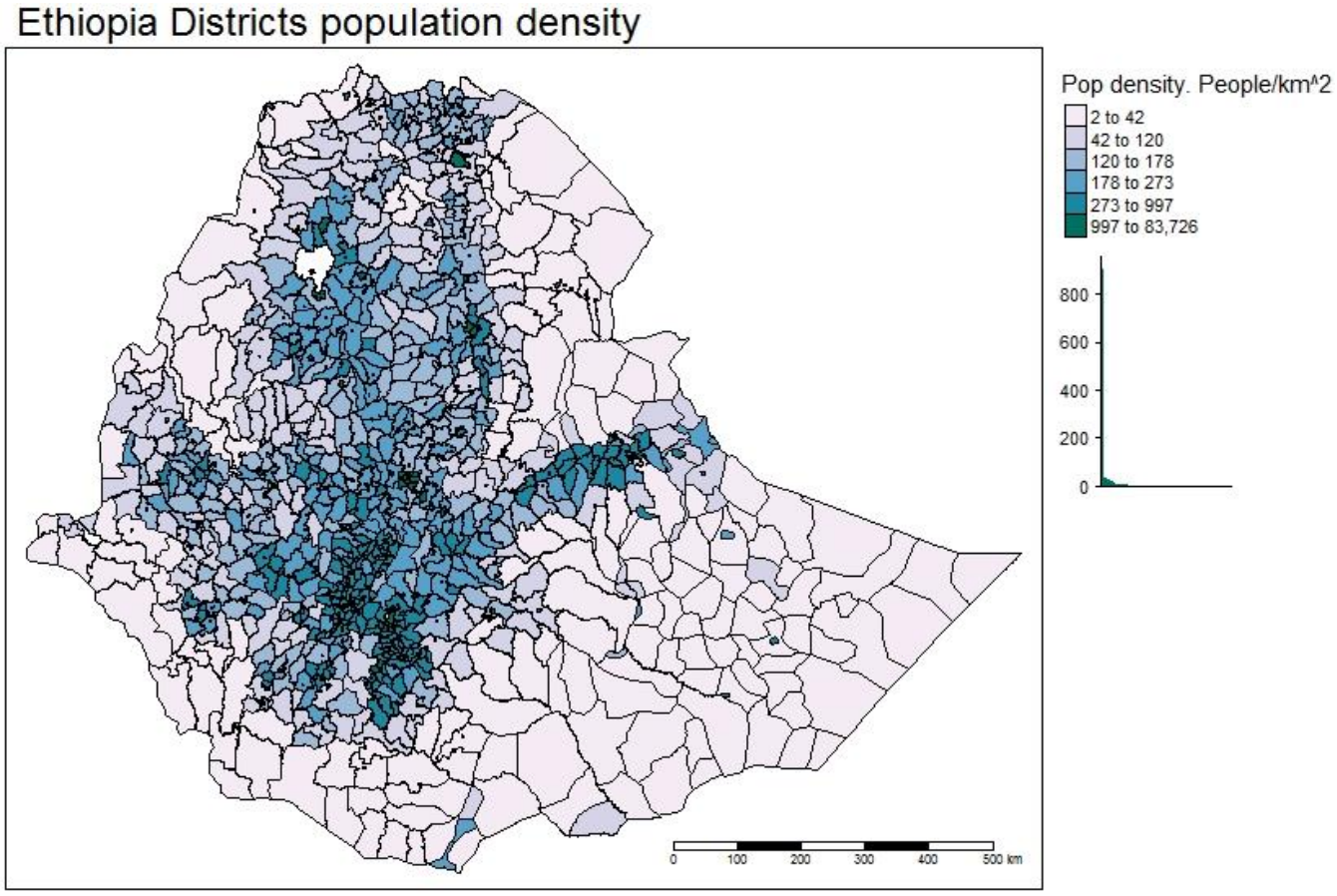
Ethiopia Population Density:

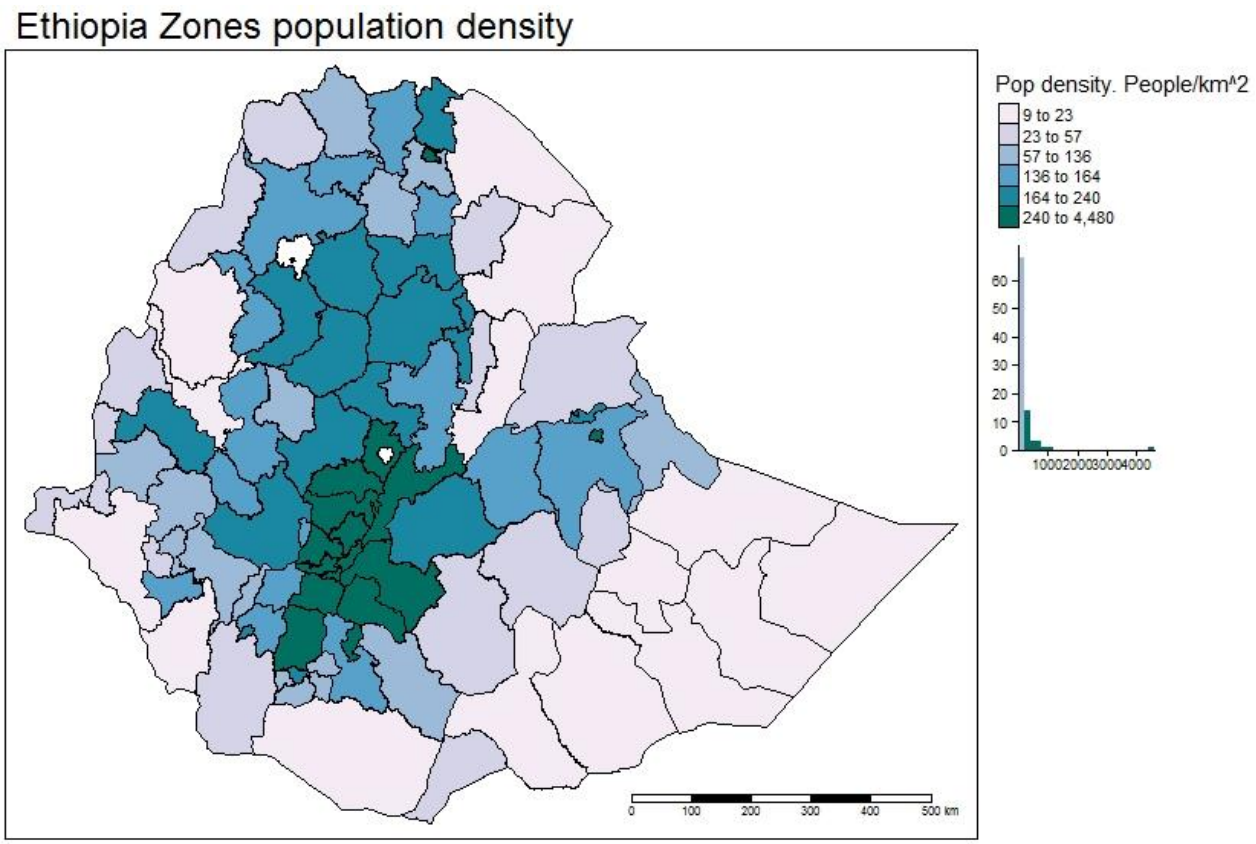
**Districts**

The map produced for population density for districts of Ethiopia shows there is a large population focus along the center and north of Ethiopia. The large diversity in the ranges demonstrates the how much of a difference there is in the highly populated urban regions of the center verses the rural outskirts surrounding the central highlands. The histogram provided with the map shows that a massive portion of the districts fell in the first class, with the remaining unable to be displayed due to the large range.

**Zones**

The map showing Ethiopia's zones population density (minus Addis Ababa) shows a similar trend to the districts, however the districts shows a more in-depth view of the population dispersal. The histogram provided shows that even without Addis Ababa (which was >8000 people per square kilometer) the differences in the classes are caused by a few points of data, however to represent the whole nation the natural jenks of the data were used to create classes.






Ethiopia Percent of total Population:

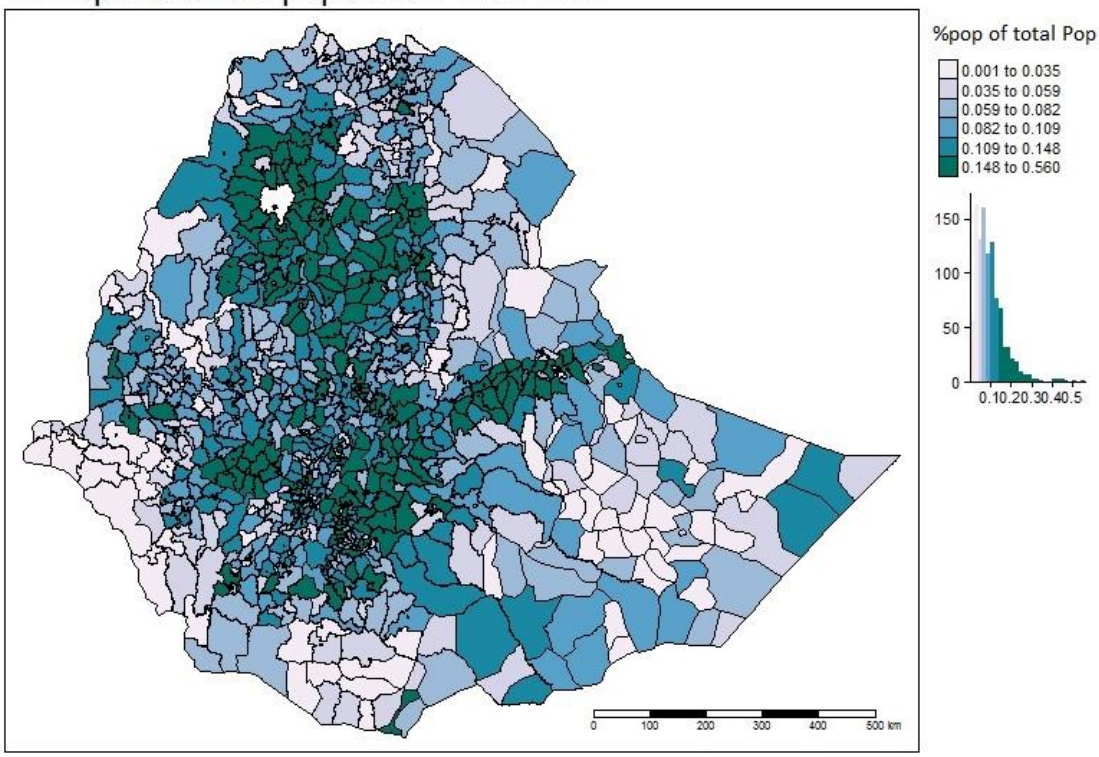
**Districts**

The percentage of total maps are able to better represent the population contribution of each region, however is unable to provide the populations relation with the area they encompass. This being said, the map still shows a similar trend of population located around the more urbanized center of the country.

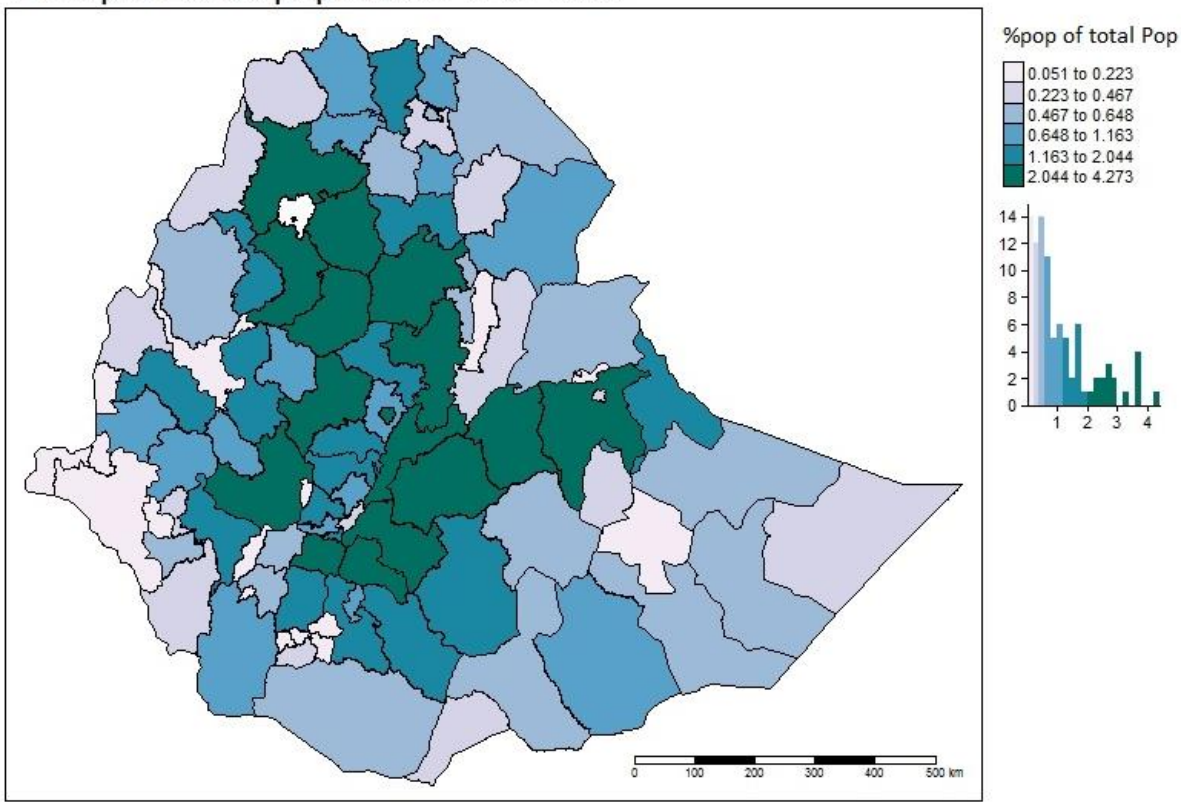
**Zones**

The zones percent of total population map is better at providing a quick and simpler visual representation to understand. The least populated regions are still seen to be in the west and east of Ethiopia.

Ethiopia Districts population % of Total



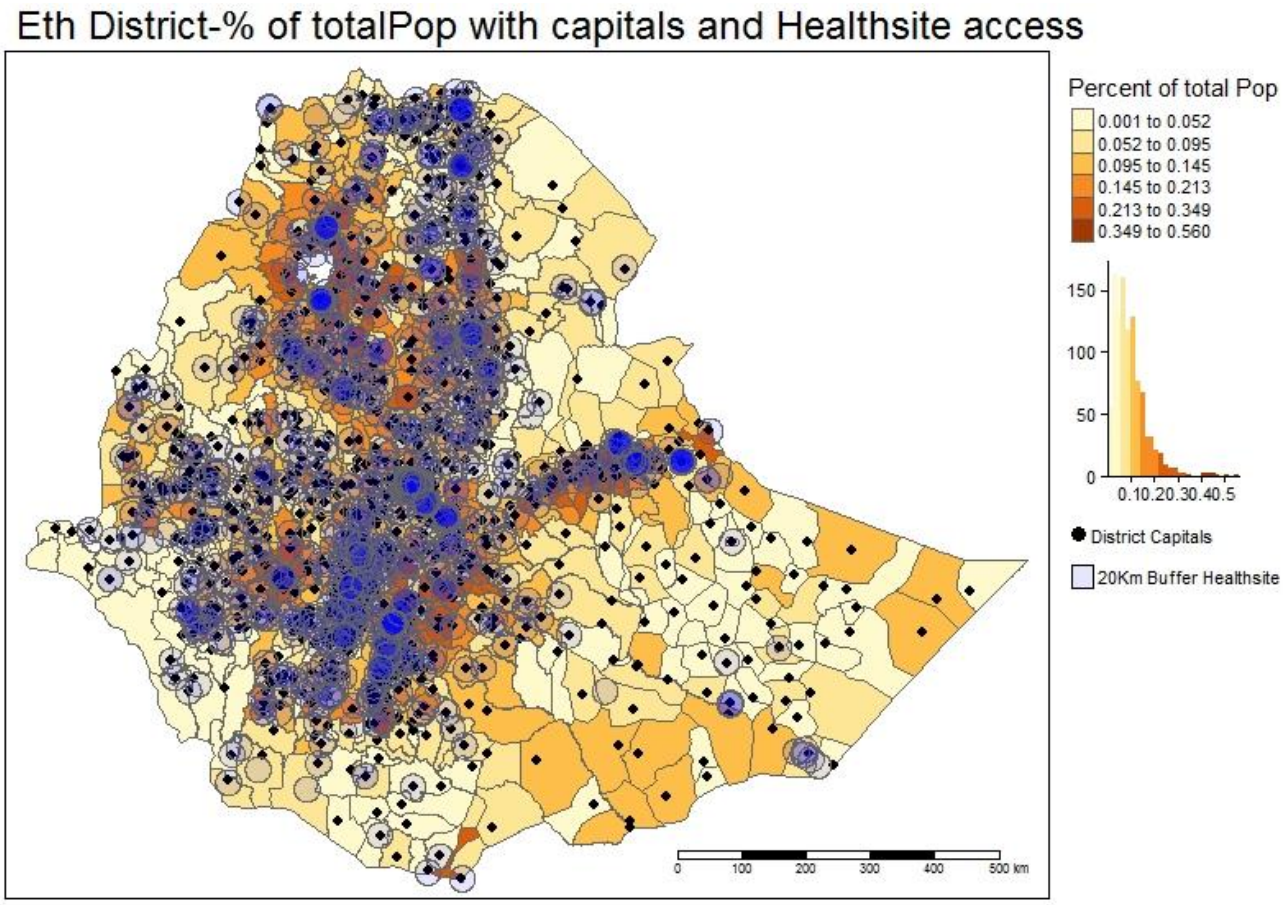

Ethiopia Zones population % of Total

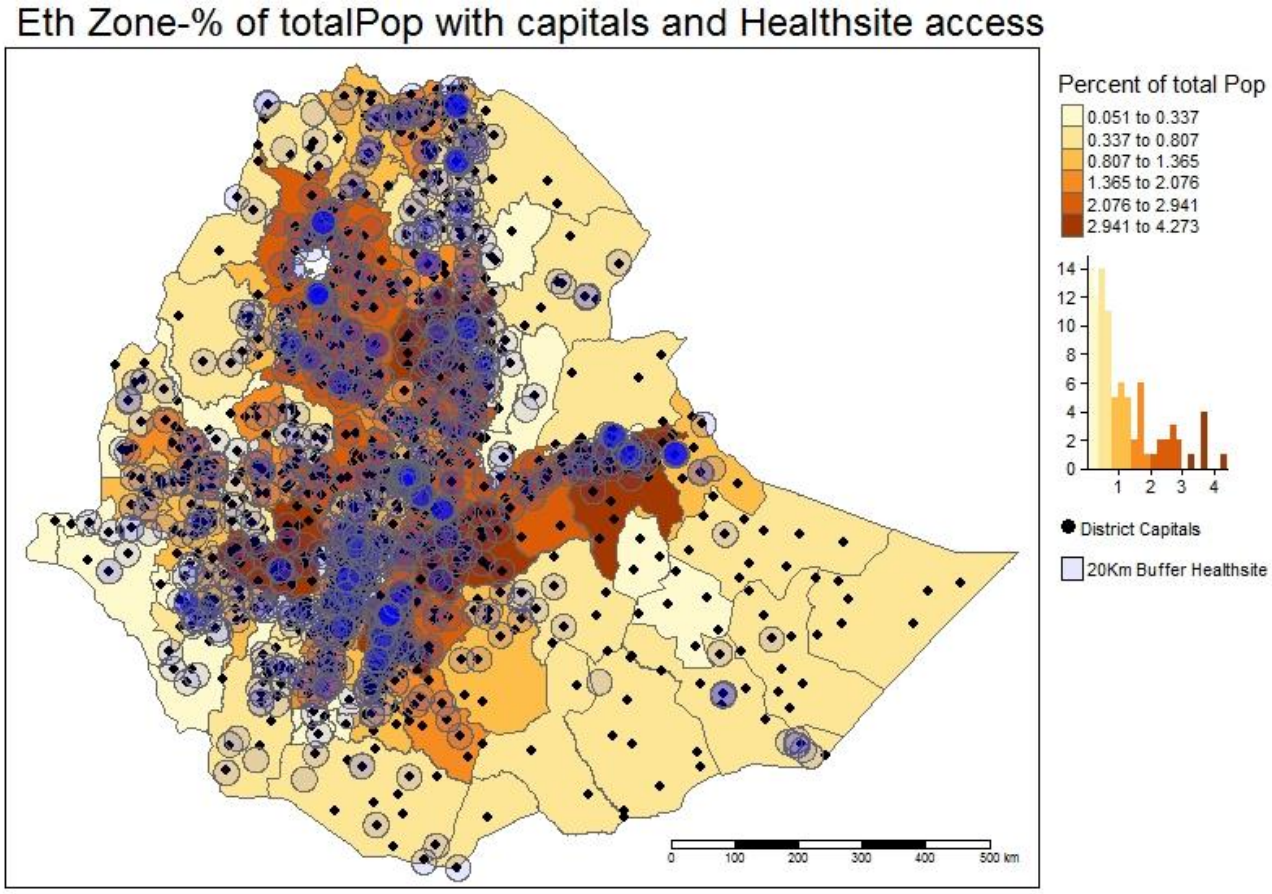


Ethiopia Health site and cities locations:

**Districts and Zones**

These maps represent the dispersal of health sites and district capitals with a choropleth map beneath showing the percent of population that these regions have. Along with reported information stated earlier about the percentages, this map shows health centers, these ranging from hospitals to pharmacies, with a 20 km buffer. Transparency has been applied to the buffers, as such, the darker blue they are, the more that region have health care centers. While district capitals are dispersed throughout the country with a focus in central Ethiopia due to the many districts found there, health care centers do not share the same dispersal. These show a similar range that the heavily populated regions follow, primarily central and northern Ethiopia.



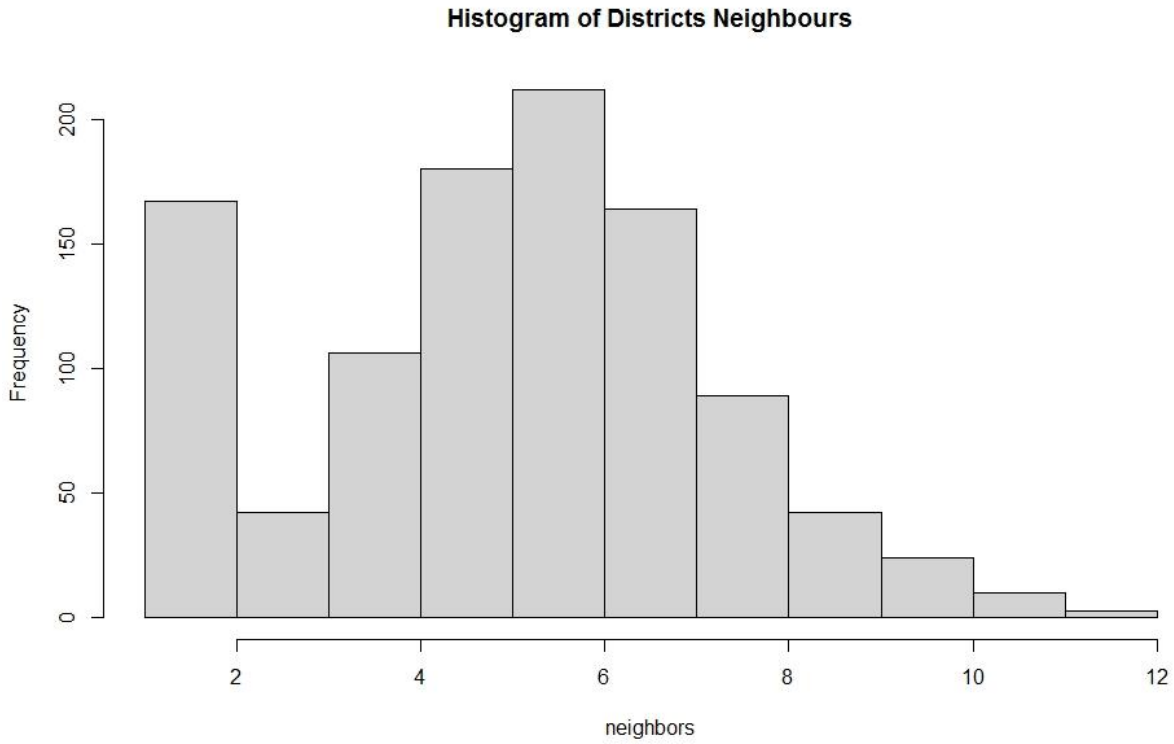


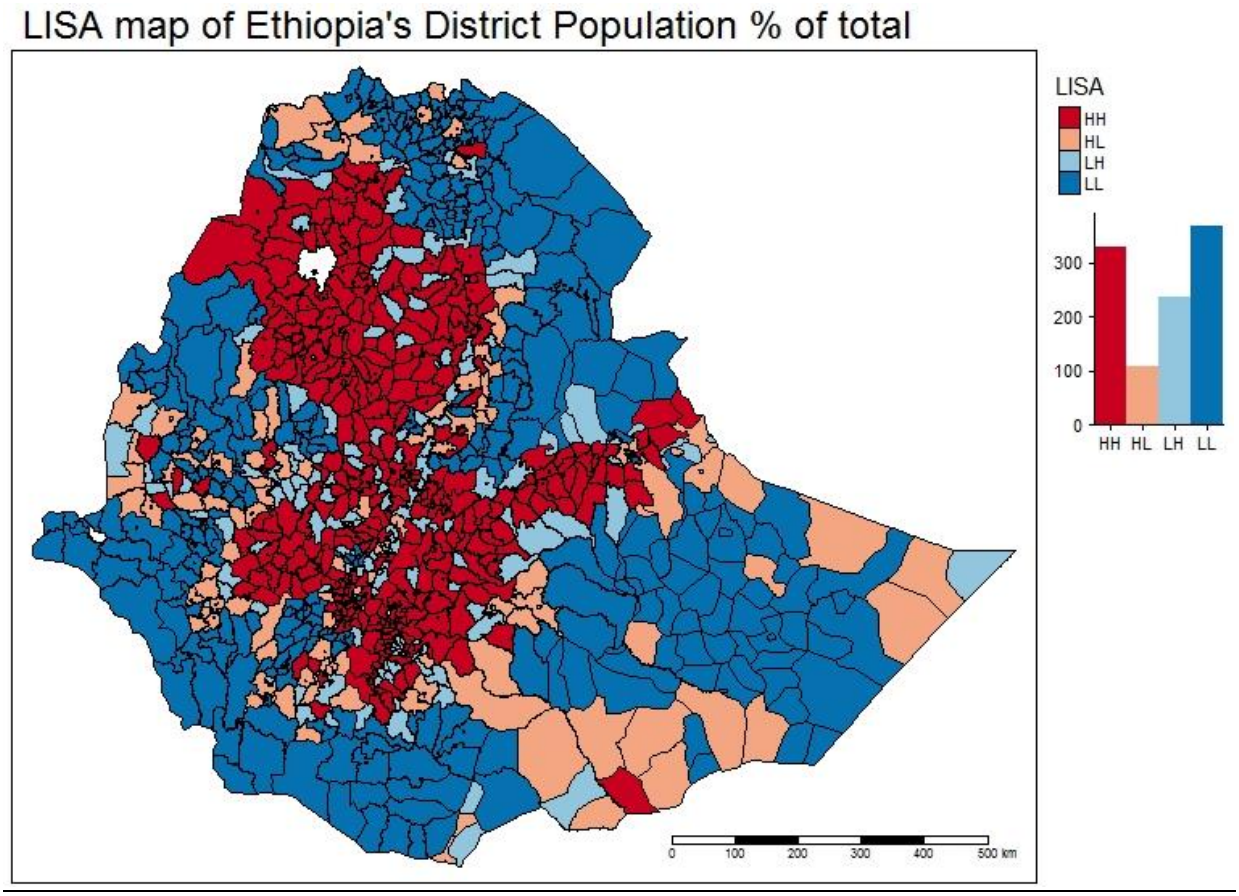

Ethiopia Spatial Autocorrelation:

**Districts**

The neighbours created from the code were plotted onto a histogram to show dispersion of the district neighbours and their counts. This graph shows that there is a large quantity of districts with only 1 neighbour while the peak have 5 nieghbours. The larger classes of neighbours can be attributed to the irregular subdivisions found in central and northern Ethiopia.

The LISA map created from the indicators of spatial autocorrelation shows the clustering regions follow the general dispersion of the population with areas of high values surrounded by high values (HH) in primarily the center and north of Ethiopia. The histogram provided shows that there are more LL areas indicating greater clustering of low values around low values.


Histogram of Districts Neighbours
Frequency
0
50
100
150
200
2
4
6
8
10
12
neighbors


LISA map of Ethiopia's District Population % of total
LISA
HH
HL
LH
LL
300
200
100
0
HH HL LH LL
0 100 200 300 400 500 km

Goals and success:

Success for this project was defined as being able to visually represent the spatial relationships between health centers and population clusters. Another goal was to find evidence showing that health centers are primarily found within greater populated areas with more cities not under a health care buffer located in rural and exterior districts such as in the sparsely populated areas.

# Discussion:

Ethiopia Population Density:

Ethiopia's population density maps show that the population found within the interior of the country tend to have more people per square kilometer. As stated in the results, the range shows that a large portion of the population lives in the smaller districts, as the surrounding rural areas along the border of Ethiopia have very low people per $km^2$. Urbanization is concentrated around the center of Ethiopia, however as most of Ethiopia's population is located in rural regions. Many of the central plain and highland regions (central Ethiopia) consist of rural areas, however as these regions are also highly populated within small boundaries. Ethiopia's largest ethnic group is the Oromo peoples and this can be seen with the distribution of the population density when compared to the state map. These comparisons can also been seen with the percentage of population maps.

The take away from the population density maps for districts and zones, is that the exterior, rural regions tend to have more land per person, this means that as the district capitals are more sparsely located, regions with health centers in these areas would not be able to accommodate neighbouring regions, whereas due to the smaller size but larger population density of the central areas of Ethiopia, regions are likely to have health centers as well as be able to accommodate for deficiencies that they may have due to their proximity to other health centers that are not the closets option.

Ethiopia Percent of total Population

These maps show a similar trend to the population densities, however further enforces that the Oromia region is the most populated state when compared to the Ethiopia state map. The districts map shows more information, allowing the map to visualize the differences between the populations of the external rural regions. These areas are not represented well on the zones map, showing that it is important to consider the spatial extent at which policy makers must look at the data. Generalizing with zones for example may leave a large portion of areas under represented. The histograms beside the maps show that the distribution of the calculated percentage data. Due to the districts having smaller regions of divisions, they provide a much better graph for even dispersion than the zones map which has aggregated the districts, reducing the precision.

The importance of this is to note that the spatial extent that is used matters to the data that is being viewed. For the purpose of this project, the percentage of population is a valuable metric to use when wanting to visualize the spatial dispersal of the population, in this case more so than density as while the density map is useful for understanding areas associated with individuals, the final decisions will be based of the actual populations living within the regions and the volume of patients to health centers.

Ethiopia Health site and cities locations:

This map shows that the areas with a low percentage of the total population are not as nearly provided for than the areas with higher population percentages. PHC or even the base levels of service such as pharmacies are not present within 20km of many of the exterior district capitals. When calculated, there is approximately 1.4 health centers for every district, however this is calculating using raw values. The dispersion shown in the maps indicate that this figure is not followed. Again, like the population, the health sites are primarily located within the center in north of Ethiopia. The darker the blue circle buffer is, the more health care centers are located in the area, where the 20km buffers overlap.

Eastern desert regions are the worst affected by the lack of health centers, with many of the district capitals not within a day's travel from a health center.

Ethiopia Spatial Autocorrelation:

When the LISA map is compared to the previous maps, it can be implied that the regions of low values surrounded by low values represent the regions that would likely require the more even dispersal of healthcare systems. The accumulation of health care centers in the high clustered regions indicated that the greater volume of patients require greater number of health centers, though as stated in the earlier background information, Ethiopia has a high volume of patients with a low ability to provide quality care (Arsenault et al, 2021). This information in association with the LISA map can show a trend of piling more centers in populated regions instead of ensuring proper services are distributed evenly.

The calculated p-value of the data set came out to 0.001, suggesting that the null hypothesis that population is evenly disperse has strong evidence against it. This clustering of population then affects the dispersal of health care centers.

# Conclusion:

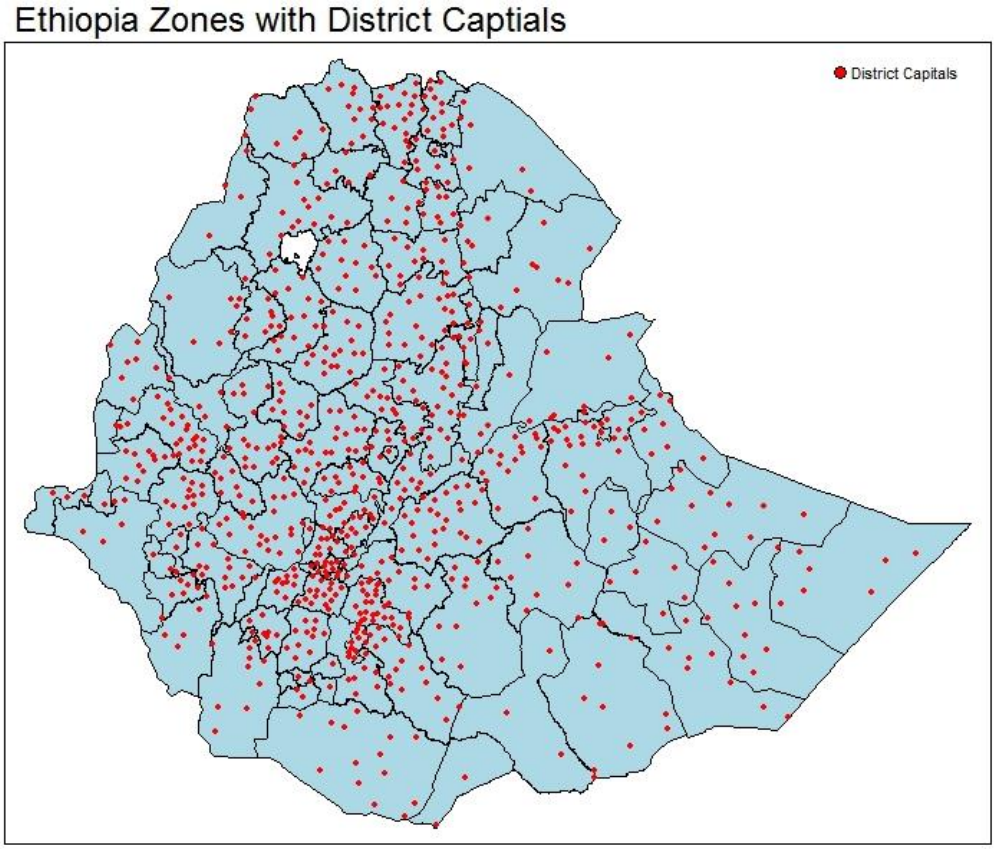
The background information and introduction showed how Ethiopia is experiencing a health care issue in terms of its distribution. These maps show that rural regions, namely in the east are the most neglected areas of the country, with the central regions have fair access to

health care but not to the quality required. This is where large movements are required for SHC for example involving going to the main federal hospital in the Capital city.

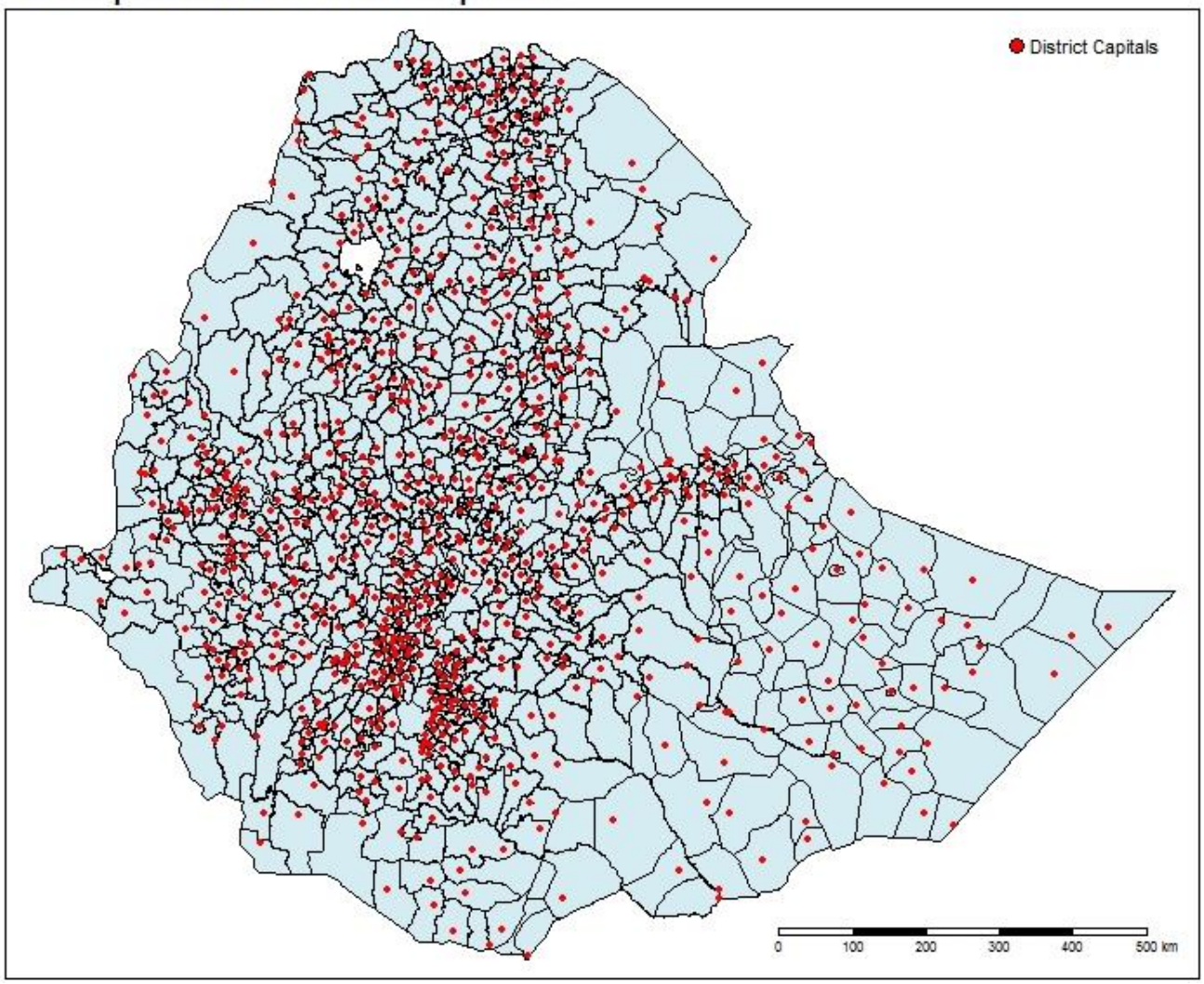
In additions to the assumptions made in this project, as Ethiopia is currently undergoing a civil conflict, many of the region within the north will have had their data changed, from health site locations to population densities due to refugee movements and civilian casualties. As the conflict is fairly unpredictable, these maps can be used to visualize areas that require health care until the fighting has ceased and aid is allowed to return to the afflicted areas, allowing for new data collection from international organizations.

In conclusion, health care is a massive undertaking that require many fields of study to provide interpretations of information so as to better know how and where to provide health care. PHC (primary health care) in Ethiopia is lacking in rural regions, primarily in the Eastern districts, however the maps help to show this information visually and so can help pin point areas that require assistance. The accumulated data also acts as a source for population data that is tied into the administrative data of Ethiopia, a data set that was not possible to find in open sources. Rstudio allowed for the quick calculation and visualization of the datasets, however was not ideal for visually representing the data.

## Additional map info:



# Ethiopia District with Captials



References:


- Freeman, T., Gesesew, H. A., Bambra, C., Giugliani, E. R. J., Popay, J., Sanders, D., Macinko, J., Musolino, C., & Baum, F. (2020). Why do some countries do better or worse in life expectancy relative to income? An analysis of Brazil, Ethiopia, and the United States of America. *International Journal for Equity in Health*, *19*(1). https://doi.org/10.1186/s12939-020-01315-z
    - (Freeman et al, 2020)
- Woldetsadik, A. B., Amhare, A. F., Bitew, S. T., Pei, L., Lei, J., & Han, J. (2020). Socio-demographic characteristics and associated factors influencing cervical cancer screening among women attending in St. Paul's Teaching and Referral Hospital, Ethiopia. *BMC Women's Health*, *20*(1). https://doi.org/10.1186/s12905-020-00927-5
    - (Woldetsadick et al, 2020)
- Adane, T. (2017). Assessment on Utilization of Health Management Information System at Public Health Centers Addis Ababa City Administrative, Ethiopia. *Internet of Things and Cloud Computing*, *5*(1), 7. https://doi.org/10.11648/j.iotcc.20170501.12
    - (Adane et al, 2017)
- Geleta, G. T., Cheme, M. C., & Roro, E. M. (2019). Physical, behavioral and sociodemographic determinants of hypertension among the adult population in Nekemte town, western Ethiopia: community based study. *BMC Research Notes*, *12*(1). https://doi.org/10.1186/s13104-019-4804-0
    - (Geleta et al, 2019)
- Arsenault, C., Yakob, B., Tilahun, T., Nigatu, T. G., Dinsa, G., Woldie, M., Kassa, M., Berman, P., & Kruk, M. E. (2021). Patient volume and quality of primary care in Ethiopia: findings from the routine health information system and the 2014 Service Provision Assessment survey. *BMC Health Services Research*, *21*(1). https://doi.org/10.1186/s12913-021-06524-y
    - (Arsenault et al, 2021)
- Nwakeze, N. M., & Kandala, N. B. (2011). The spatial distribution of health establishmentsin Nigeria. *African Population Studies*, *25*(2). https://doi.org/10.11564/25-2-251
    - (Nwakeze and Kandala, 2011)
- Bennema-Broos, M. (2001). Social democratic government and spatial distribution of health care facilities: The case of hospital beds in Germany. *The European Journal of Public Health*, *11*(2), 160–165. https://doi.org/10.1093/eurpub/11.2.160
    - (Bennema-Broos, M, 2001)
- *Ethiopia - Subnational Administrative Boundaries - Humanitarian Data Exchange*. (2021, November 23). [Dataset]. OCHA. https://data.humdata.org/dataset/ethiopia-cod-ab
    - (OCHA admin, 2021)
- *Ethiopia - Location of Health Facilities - Humanitarian Data Exchange*. (2021, November 11). [Dataset]. OCHA. https://data.humdata.org/dataset/ethiopia-health
    - (OCHA health, 2021)
- *Ethiopia-healthsites - Humanitarian Data Exchange*. (2021, November 8). [Dataset]. OCHA. https://data.humdata.org/dataset/ethiopia-healthsites
    - (OCHA sites, 2021)
- *Ethiopia - Cities, towns and villages - Humanitarian Data Exchange*. (2021, July 22). [Dataset]. OCHA. https://data.humdata.org/dataset/ethiopia-settlements
    - (OCHA cities, 2021)
- *Ethiopia - Subnational Population Statistics - Humanitarian Data Exchange*. (2021). [Dataset]. OCHA. https://data.humdata.org/dataset/ethiopia-population-data-_-admin-level-0-3
    - (OCHA pop, 2021)