**PORTFOLIO OF FORMATIVE ASSESSMENTS OF DS-7002 MODULE**

**SPATIAL DATA ANALYSIS (SDA)**

**PREPARED BY:**

**AGBOOLA ISAAC OLUWATOMIWA**

**STUDENT NO: R2307D16915497**

**PROGRAM: MSC DATA SCIENCE**

**UNIVERSITY OF EAST LONDON**

**SUMMATIVE ASSESSMENT 1**

**21th JULY, 2024**

**TABLE OF CONTENTS**

[**LIST OF FIGURE** 5](#_Toc172371859)

[**ABSTACT** 7](#_Toc172371860)

[**INTRODUCTION** 8](#_Toc172371861)

[**WEEK 1 - FORMATIVE TASK 1** 9](#_Toc172371862)

[**QUESTIONS** 9](#_Toc172371863)

[**1.1 SPATIAL RESOLUTION** 10](#_Toc172371864)

[**1.2 RASTER DATA FILE** 12](#_Toc172371865)

[**1.3 TOPOLOGY** 14](#_Toc172371866)

[**1.4 RASTER AND VECTOR GIS MODELS** 15](#_Toc172371867)

[**1.5 GEOSPATIAL MODELS** 16](#_Toc172371868)

[**1.6 ADVANTAGE AND DISADVANTAGES OF RASTER AND VECTOR GIS MODELS** 17](#_Toc172371869)

[**COMPARE RASTER MODEL AND VECTOR MODEL** 17](#_Toc172371870)

[**1.7 DATA STRUCTURES** 18](#_Toc172371871)

[**CONCLUSION** 21](#_Toc172371872)

[**WEEK 2 - FORMATIVE TASK 2** 22](#_Toc172371873)

[**QUESTIONS** 22](#_Toc172371874)

[**2.1 SPATIAL RELATION AND SPATIAL QUERIES** 23](#_Toc172371875)

[**2.2 SPATIAL QUERY** 24](#_Toc172371876)

[**2.3 SPATIAL INFORMATION AND GIS** 25](#_Toc172371877)

[**2.4 SPACE-RELATED INFORMATION** 26](#_Toc172371878)

[**2.5 QUERY** 27](#_Toc172371879)

[**2.6 PARCELS AND DISTANCE OF BUILDING** 28](#_Toc172371880)

[**CONCLUSION** 28](#_Toc172371881)

[**WEEK 3 - FORMATIVE TASK 3** 29](#_Toc172371882)

[**QUESTIONS** 29](#_Toc172371883)

[**3.1 THE CENTROIDS MAP OF MUNICIPALITY OF THE STATE** 31](#_Toc172371884)

[**3.2 THE MEAN HUMAN DEVELOPMENT INDEX OF MUNICIPALITIES OF THE STATE IN BRAZIL** 35](#_Toc172371885)

[**3.3 THE POLYGON/SHAPEFILE MAPPING OF THE MUNICIPALITY** 37](#_Toc172371886)

[**3.4 THE BUFFER OF THE HOUSING PROJECT** 40](#_Toc172371887)

[**3.5 THE DISTRIBUTION OF THE TOTAL HOUSING UNITS BY MUNICIPALITY** 42](#_Toc172371888)

[**CONCLUSION** 44](#_Toc172371889)

[**WEEK 4 - FORMATIVE TASK 4** 45](#_Toc172371890)

[**QUESTIONS** 45](#_Toc172371891)

[**4.1 THEMATIC MAP TYPES** 46](#_Toc172371892)

[**4.2 POPULATION DENSITY IN A THEMATIC MAP** 48](#_Toc172371893)

[**4.3 THE DERIVED DATA OF THE CHOROPLETH MAP** 49](#_Toc172371894)

[**4.4 CLASSIFICATION OF THEMATIC MAP TYPES** 49](#_Toc172371895)

[**4.5 VISUAL VARIABLES AND SYMBOL DIMENSIONALITIES** 49](#_Toc172371896)

[**4.6 CHOROPLETH MAP FOR THE UNEMPLOYMENT RATE IN A COUNTRY** 52](#_Toc172371897)

[**4.8 THE LEGEND OF THEMATIC MAP TYPES** 54](#_Toc172371898)

[**CONCLUSION** 57](#_Toc172371899)

[**REFERENCES** 58](#_Toc172371900)

# **LIST OF FIGURE**

[Figure 1: The Diagram of a Digital Camera having one pixel of an Image 10](#_Toc172474185)

[Figure 2: The Diagram of a Digital Camera having nine pixels of images 11](#_Toc172474186)

[Figure 3: The Diagram shows the Run Length Encoding Method of Raster Representation 13](#_Toc172474187)

[Figure 4: The Diagram of Raster Data Model 19](#_Toc172474188)

[Figure 5: The Diagram of Vector Data Model 20](#_Toc172474189)

[Figure 6: Map of Brazilian Municipalities 31](#_Toc172474190)

[Figure 7: Features of the selected Municipality 32](#_Toc172474191)

[Figure 8: The Newly selected Layer from the Municipality 32](#_Toc172474192)

[Figure 9: The New Layer depicting the centroids of São Paulo 33](#_Toc172474193)

[Figure 10: The Sao Paulo showing the centroids 33](#_Toc172474194)

[Figure 11: The Sao Paulo representing the centroids and outer boundary of the state 34](#_Toc172474195)

[Figure 12: The New Layer of the Dissolved 35](#_Toc172474196)

[Figure 13: Displaying the states and their Human Development Index (MHDI) in Brazil 36](#_Toc172474197)

[Figure 14: The Indigenous Territory Diagram 37](#_Toc172474198)

[Figure 15: The Gaucha do Norte and Parque do Xingu new layer 38](#_Toc172474199)

[Figure 16: The area of the two regions Gaucha do Norte and Parque do Xingu 39](#_Toc172474200)

[Figure 17: Distance Matrix of the interaction between Gaucha do Norte and Parque do Xingu 40](#_Toc172474201)

[Figure 18: The new layer derived from housing and AC\_state 41](#_Toc172474202)

[Figure 19: 10km buffer around each housing project 41](#_Toc172474203)

[Figure 20: Attribute joining based on spatial location 42](#_Toc172474204)

[Figure 21: Map showing the distribution of total housing units 43](#_Toc172474205)

[Figure 22: The Unemployment rate of Brazil 52](#_Toc172474206)

[Figure 23: The Unemployment rate of Brazil using the graduated data 53](#_Toc172474207)

# **ABSTACT**

This paper presents weekly assessment for Module DS-7002 which is focused on Spatial Data Analysis. UNICAF oversees the University of East London's fourth module in its MSc Data Science program. The table of contents on the next page provides questions and answer to the formative assessment exercises that were conducting during week 1 to 4.

# **INTRODUCTION**

The module on formative assessment explores various concepts related to spatial data and Geographic Information System (GIS) models. Week 1 introduces topics like spatial resolution, raster data model, topology, vector model, and data structures. Week 2 focuses on spatial queries within a defined study area using GIS datasets, and the input-output processes of building GIS datasets. Week 3 involves illustrating and creating maps that display centroids, boundaries, mean Human Development Index, polygon mapping, creating kilometre-based buffers for housing project locations, and identifying the highest and lowest housing numbers across different states in Brazil using municipality datasets. Additionally, it includes showing distance matrices with zero factors represented diagonally. Finally, Week 4 covers thematic map types, such as population density and unemployment rate, using various symbols and diagrams to represent thematic map types. The explanation demonstrates how to solve complex problems and create maps using QGIS.

# **WEEK 1 - FORMATIVE TASK 1**

# **QUESTIONS**

1.1 Examine a digital photo you have taken recently. Can you estimate its spatial resolution?

1.2 If you were to create a raster data file showing the major land-use types in your county, which encoding method would you use? What method would you use if you were to encode a map of the major roads/rivers in your county and why?

1.3 What is topology? Explain topological and non-topological data structures.

1.4 What are raster and vector GIS models? Give an example of objects representing a point, line or polygon.

1.5 What are geospatial models? Explain raster and vector GIS models.

1.6 Explain the advantages and disadvantages of raster and vector GIS models. Compare both models.

1.7 Explain data structures in detail

## **1.1 SPATIAL RESOLUTION**

Determining the spatial resolution of a digital photo involves analyzing the image properties to establish its size, length and height. Spatial resolution refers to the quantity of pixels utilized by a sensor to generate a digital image (Lampros & Lambros, 2017). Images with higher spatial resolution contain more pixels than those with a lower resolution, enabling better information capture. Enlarging (Zooming), the image reveals small square elements forming the overall picture, with a coherent image resulting from the combination of pixels of different colours (Rao, 2017; Anh, et al., 2016).



Figure 1: The Diagram of a Digital Camera having one pixel of an Image



Figure 2: The Diagram of a Digital Camera having nine pixels of images

## **1.2 RASTER DATA FILE**

To create a raster data file displaying the major land use in my country, I prefer to use the Run-Length Encoding Data Structure method. This method provides a straightforward way to represent the habitat, and land cover of a particular area (Rao, 2017). The encoding process involves multiple iterations to ensure consistent values for each pixel in the image (Ron, 2008).

I would rather use vector lines to encode a map of my country's major roads/rivers. This choice is due to the easily indicated shape of linear features, which can effectively represent specific types of areas on the map. Vectors are quantities that have both magnitude and direction, mostly used in Geographic Information Systems (GIS) to represent the speed and force of movement and also the points, polygons, and lines (Christopher, 2014). In the case of roads and rivers, vectors cannot only indicate flow direction and strength but also represent various forces, such as currents or traffic patterns.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(0, 8)

(0, 2), (1, 4), (0, 2)

(0, 2), (1, 5), (0, 1)

(0, 3), (1, 4), (0, 1)

(0, 4), (1, 3), (0, 1)

(0, 4), (1, 3), (0, 1)

(0, 4), (1, 2), (0, 2)

(0, 8)

Figure 3: The Diagram shows the Run Length Encoding Method of Raster Representation

## **1.3 TOPOLOGY**

Topology is a set of rules that govern the space and time relationships between points, lines, and polygons. It encompasses the spatial connections like adjacency, containment, and connectivity in a Geographic Information System, studying the properties of geometric objects (Samuel, et al., 2016; Christopher, 2014). Topologies can be grouped into, Arc-node, polygon-arc, and right-left topology.

The topological data structure is called an intelligent data structure, spatial relationships among geographic features can be easily derived, conducting complex data analysis. Unlike in such a structure, lines cannot intersect without a node (Rao, 2017). Graphs and Networks exemplify topological data structures by representing interconnected nodes.

In contact, non-topological data structures store information on data element connection without relying on topological properties. “Arrays, lists, and tables are considered non-topological data structures” because they organize data in a linear or two-dimensional manner, making them effective for storing and arranging data without intricate hierarchical relationships (Rao, 2017).

## **1.4 RASTER AND VECTOR GIS MODELS**

In a Raster GIS model, spatial features are depicted as a grid of cells or pixels, where each cell contains a value representing the attribute, such as elevation or temperature, in a continuous phenomenon (Rao, 2017).

The vector GIS model represents spatial features as points, lines, and polygons, using points with their x, and y coordinates to construct spatial features in a discrete phenomenon (Rao, 2017).

Examples of objects representing points include cities, hospitals, and boreholes, while objects representing lines might include roads, rivers, railways, streams, and drainage gutters. Polygon objects can represent features like lakes, countries, land use, rock types, and buildings (Christopher, 2014).

## **1.5 GEOSPATIAL MODELS**

Geospatial models, also known as spatial models, are mathematical representations of geographic features that illustrate, predict patterns and understand spatial relationships within a given geographical or spatial component, showing locations on the Earth's surface (Reddy, 2018). This involves interpreting maps and specific geographic areas identified by latitude and longitude. These models are employed to illustrate data such as population density, land use, and the distribution of natural resources, frequently used in disciplines such as geography, environmental science, and urban planning.

Raster GIS models depict the “Earth's surface as a grid of cells, where each cell, or raster, holds a value such as elevation or land cover type, effectively representing continuous phenomena such as surface elevation or temperature”.

Vector GIS Models “use points, lines, and polygons to depict Earth’s surface, often portraying discrete phenomena, such as political boundaries or land use classifications” through vector data.

## **1.6 ADVANTAGE AND DISADVANTAGES OF RASTER AND VECTOR GIS MODELS**

**ADVANTAGES OF RASTER MODELS**

* Evaluation is straightforward with the appropriate software
* Easy display and visualization through maps.

**DISADVANTAGES OF RASTER MODELS**

* High storage consumption, particularly for fine-scale features or large areas.
* Poor representation of points, areas and lines.

**ADVANTAGES OF VECTOR MODELS**

* Easy editing and updating of vector data with strong network analysis capabilities.
* High spatial resolution in vector model.

**DISADVANTAGES OF VECTOR MODELS**

* More challenging to work with compared to raster data for complex analyses.
* Difficulty in representing continuous phenomena in vector data.

### **COMPARE RASTER MODEL AND VECTOR MODEL**

* Raster data represents continuous phenomena, while vector data represents discrete phenomena.
* In raster data, each grid cell contains a numerical value, while vector data consists of points, lines, and polygons.

## **1.7 DATA STRUCTURES**

The data structure in Geographic Information Systems refers to a collection of data of different data types represented as objects that can be used in programs.

Vector Data Structures illustrate geographical “features like points, lines, or polygons, and store their respective coordinates”, commonly used for discrete features like roads, buildings, or land parcels (Christopher, 2014).

Raster Data Structures illustrate geographical features through a grid of cells or a matrix, where each cell contains attributes illustrated by numbers in rows and columns, typically utilized for continuous surfaces, like elevation or land cover (Rao, 2017).

Topological Data Structures illustrate the spatial relationship between geographical features, containing information about how features are connected and the boundaries they share (Rao, 2017).

1. Point

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 |

1. Line

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 1 | 1 |
| 0 | 0 | 1 | 0 | 0 |
| 1 | 1 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 | 0 |

1. Polygon

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 |
| 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 |

Figure 4: The Diagram of Raster Data Model



Line



Point

Polygon

Figure 5: The Diagram of Vector Data Model

# 

# **CONCLUSION**

Lastly, our comprehension of spatial data and Geographic Information systems (GIS) is demonstrated by our capability to estimate spatial resolution, select suitable encoding techniques for various data types, elucidate topology and its variations, discuss data structures, and distinguish between raster and vector Geographic Information System (GIS). This is achieved through precise responses regarding the domain of Geographic Information Systems and spatial analysis.

# **WEEK 2 - FORMATIVE TASK 2**

# **QUESTIONS**

2.1 Which of the following is not a spatial relation that can be used in spatial queries? ◼ Distance-based relation ◼ Spatial autocorrelation ◼ Topological relation ◼ Direction relation 2.2 There is a GIS dataset of points of interest (POIs) in a region, and you would like to select only those located within a pre-defined study area. How will you translate it into a spatial query?

2.3 What does the spatial information in a GIS mainly consist of?

2.4 How can space-related information be divided?

2.5 Which approaches can be used to formulate a query?

2.6 Describe the inputs which should be used to answer the following questions and what would the outputs look like: "Find all buildings which are located on parcels with a minimal area of 1000 m2 and a distance of more than 250m to the highway"

## **2.1 SPATIAL RELATION AND SPATIAL QUERIES**

The spatial relation that cannot be used in spatial queries is spatial autocorrelation. It pertains to the statistical dependence of spatial attributes and the similarity between features is based on their spatial proximity. This can be evaluated using metrics that illustrate how closely observations are clustered together. Moran's I and Geary's C are the primary indices used in spatial autocorrelation analysis.

## **2.2 SPATIAL QUERY**

Spatial queries can be performed by utilising spatial operations like within or intersects to convert chosen points of interest (POIs) within a designated study area as a spatial geometry and employing spatial operators to extract POIs falling within that area. The SQL statement is given below:

SELECT pois.\*

FROM pois, study\_area

WHERE ST\_Within

(

Pois. Geometry,

study\_area.geometry

) ;

In this SQL statement, SELECT pois.\* denotes the selection of all columns from the POI table. FROM pois, study\_area specifies the tables from which the data will be retrieved. ST\_Within(pois. geometry, study\_area.geometry) is a spatial function in SQL that filters the data based on spatial relationships. ST\_Within function can be checked if the point falls within the specified points of interest (POIs). The geometry column stores spatial information of the points and the study area, while study\_area.geometry represents the pre-defined study area geometry.

## **2.3 SPATIAL INFORMATION AND GIS**

Spatial data in a GIS mainly comprises geographical location and features on the Earth's surface, such as vectors (points, lines, polygons), and raster that represent various attributes. They include:

* **Demographic features:** Representing population characteristics like age, migration, gender, and education level (Adam, 2024).
* **Physical features:** Natural Earth elements like mountains, forests, rivers, valleys, and deserts (Alex, 2024).
* **Human-made features:** Structures built by humans for various purposes including cities, buildings, bridges, dams, and transportation.
* **Natural resources:** Materials provided by nature without human intervention including Renewable energy, minerals, stone, water, and wildlife habitats (G., 2009).

## **2.4 SPACE-RELATED INFORMATION**

Spatial information can be divided into two main parts based on the nature of the objective of the analysis.

* **Features type:** Spatial information is segmented according to the feature type it represents, “such as points, lines, or polygons”. For example, a GIS dataset may include layers for buildings (polygons), and lakes (lines).

Top of Form

Bottom of Form

* **Aattribute type:** This division involves dividing spatial information based on the attribute type. For instance, a GIS dataset could include layers representing various land uses, such as residential, commercial, and industrial zones.
* **Geography:** This is when the spatial information of the geographical region in the Earth and universe can be divided. For example, a GIS dataset might have layers representing different countries, and states.

## **2.5 QUERY**

There are different ways to formulate queries in a geographic Information System (GIS) to assess Information accuracy (Rudolf, 2000).

* **Structured Query Language (SQL):** SQL is a powerful programming language for managing and manipulating data. In GIS, SQL empowers users to create intricate queries that define criteria for selecting, processing and sorting data (A.H.M, et al., 2004).
* **Query Builders:** Graphical query builders provide users with an intuitive point-and-click interface to construct queries. They simplify the process of defining criteria for data selection and extraction.
* **Map Selection tools:** These tools allow users to select features on a map directly. For instance, a user can delineate a specific area on a map containing all buildings and then use this selection to retrieve attribute data linked to those buildings.

## **2.6 PARCELS AND DISTANCE OF BUILDING**

**INPUT**

* The parcel boundaries and associated attributes, including detailed information on parcel areas such as parcel ID, owner details, and area measurement.
* The location and geometry of the highway such as street name and speed limit.
* The locations of buildings and their attributes including point features representing building locations with attributes such as building name, type, and year of construction.

**OUTPUT**

The query will generate either a list or a map displaying buildings that meet certain standard, such as being positioned on parcels with a minimum area of 1000m² and situated more than 250m away from the highway. The output may include a map highlighting the selected buildings or a table listing relevant building details such as establishment year, construction budget, and building ID. Additionally, the building layout will be determined using the QGIS software.

# **CONCLUSION**

Finally, we have demonstrated our comprehension of mapping building plans within Geographic Information Systems (GIS). This includes showcasing proficiency in spatial autocorrelation, handling datasets of points of interest (POIs) within specified study areas, executing spatial queries, and managing the input and output of building locations on parcels. This achievement was attained through accurate responses to inquiries concerning GIS.

Top of Form

Bottom of Form

# **WEEK 3 - FORMATIVE TASK 3**

# **QUESTIONS**

3.1 Produce a map showing the centroids of each municipality in just the state of Säo Paulo, and add the outer boundary of Säo Paulo state.

3.2 What is the mean Human Development Index of municipalities in each state of Brazil?

3.3 Produce a polygon/shapefile mapping the area of the municipality ‘Gaucha do Norte’ that is in the indigenous territory “Parque do Xingu”.

3.4 In the state of Acre (AC), which two social housing (MCMV) projects are closest to each other? Create a 10km buffer around each housing project.

3.5 Across Brazil, which municipalities have the lowest and highest number of MCMV housing units (UH) in its territory? Create a map of the distribution of total housing units by municipality.

## **3.1 THE CENTROIDS MAP OF MUNICIPALITY OF THE STATE**



Figure 6: Map of Brazilian Municipalities

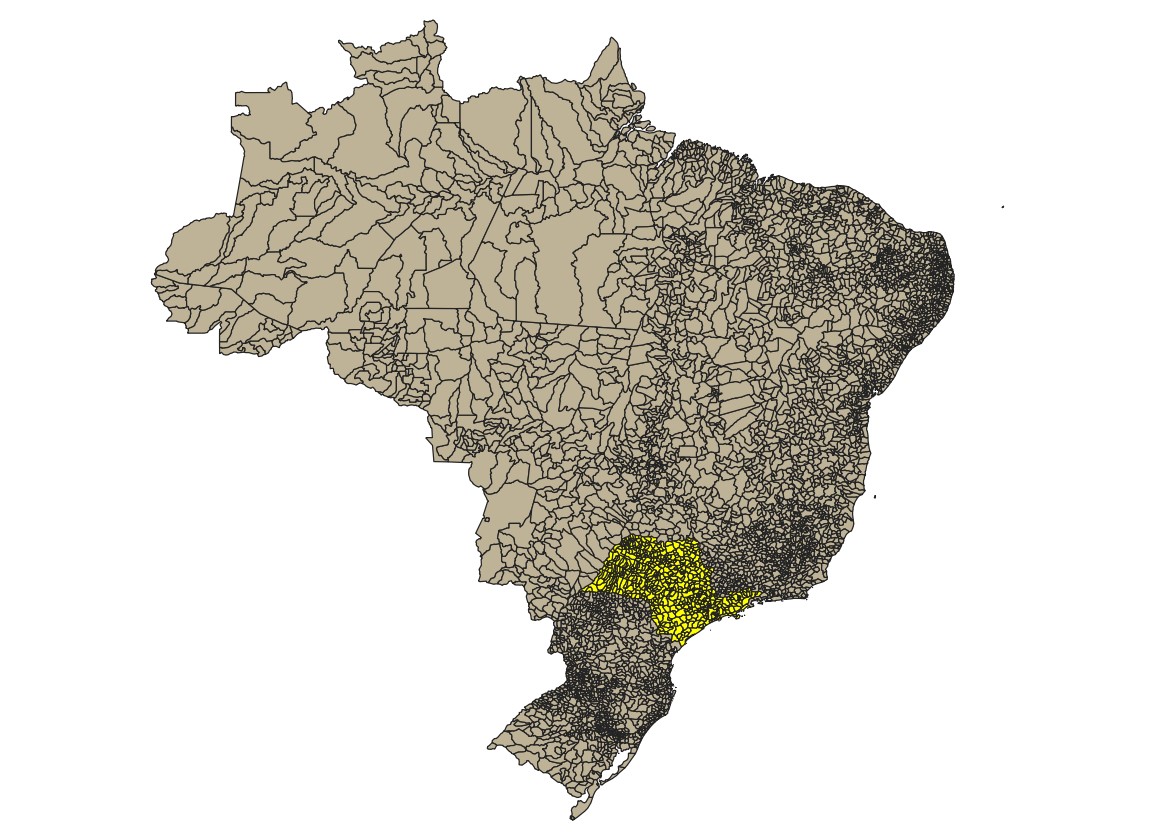


Figure 7: Features of the selected Municipality

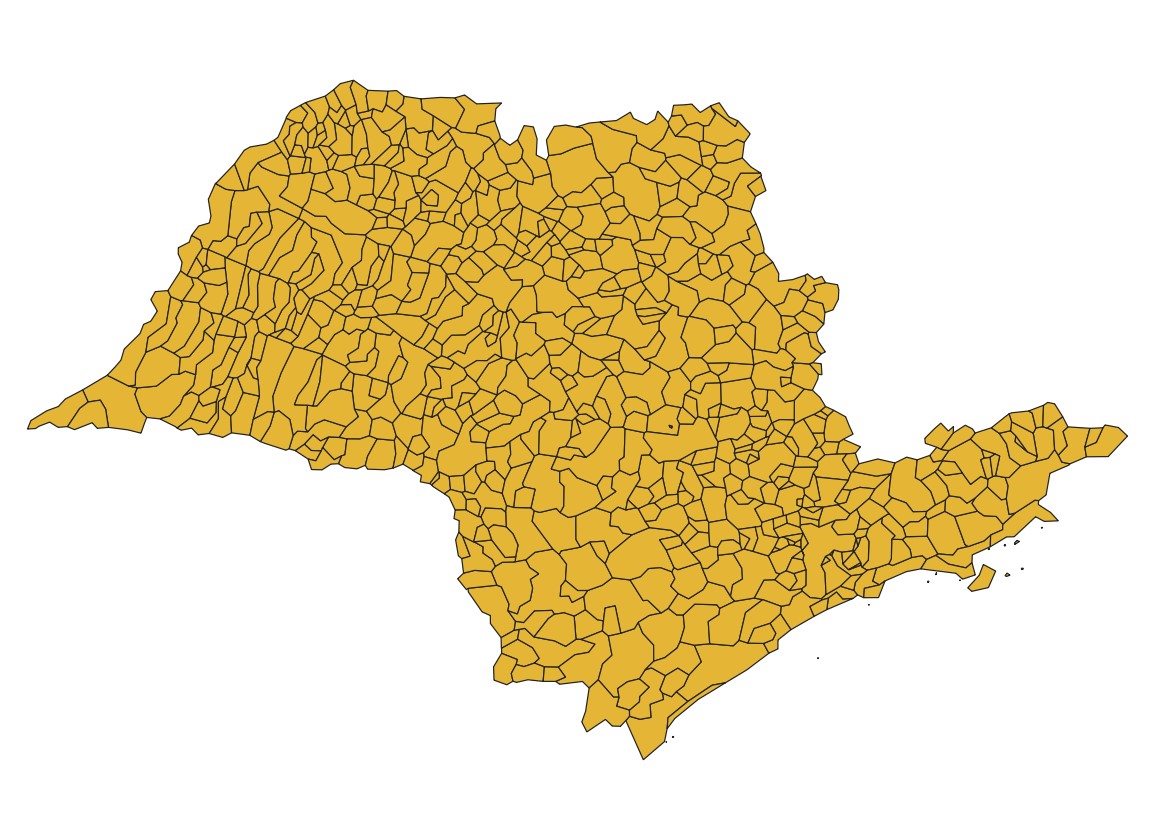


Figure 8: The Newly selected Layer from the Municipality

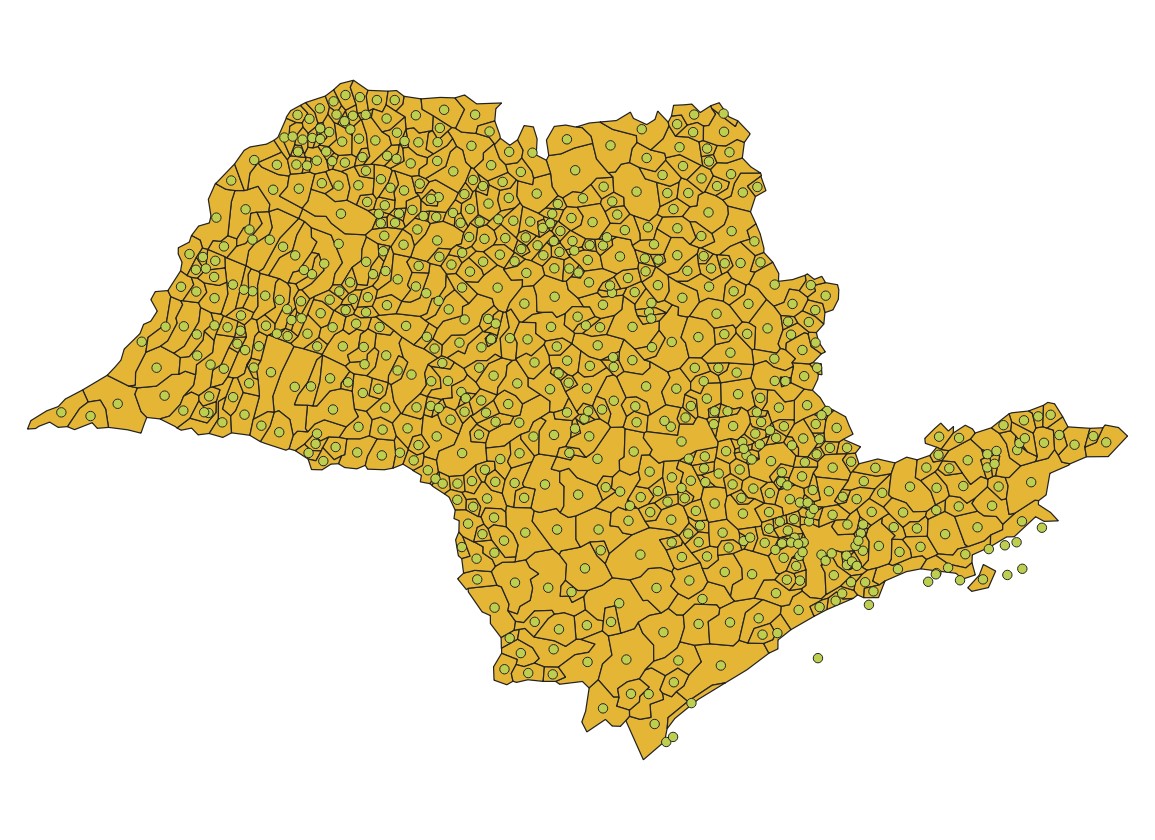


Figure 9: The New Layer depicting the centroids of São Paulo



Figure 10: The Sao Paulo showing the centroids

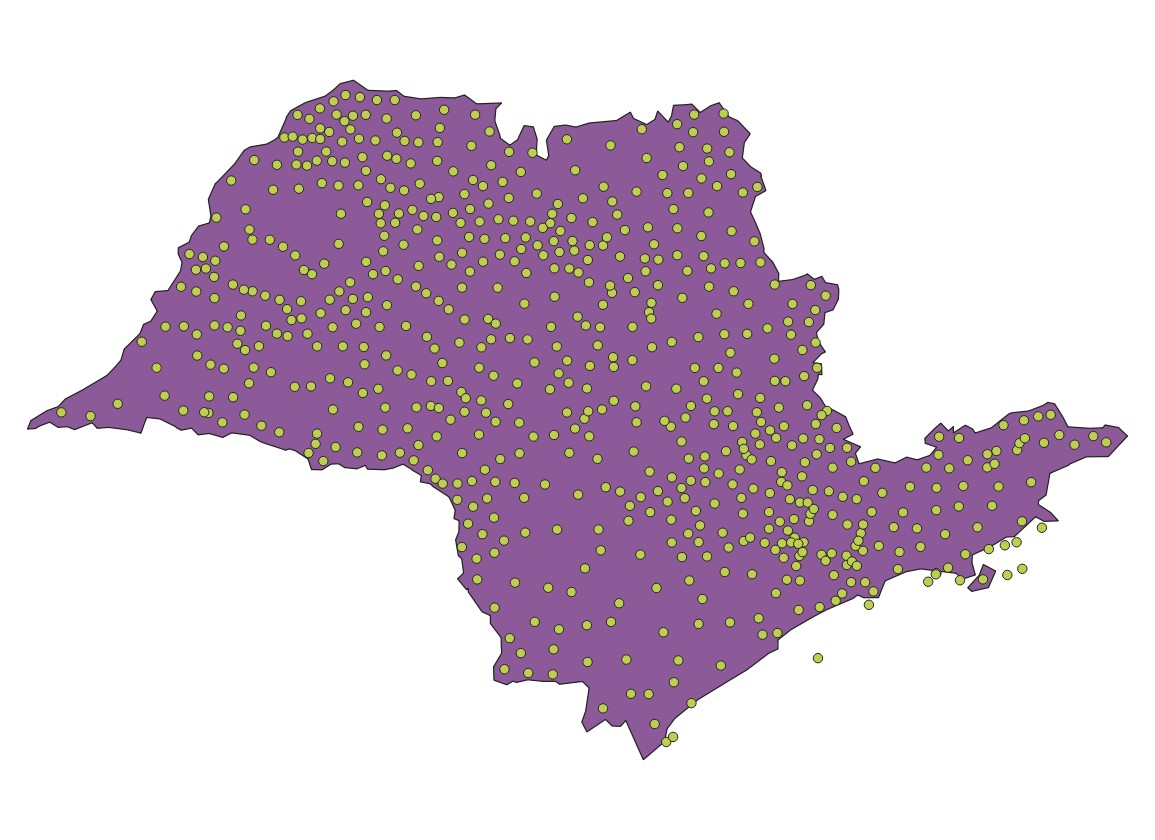


Figure 11: The Sao Paulo representing the centroids and outer boundary of the state

## **3.2 THE MEAN HUMAN DEVELOPMENT INDEX OF MUNICIPALITIES OF THE STATE IN BRAZIL**

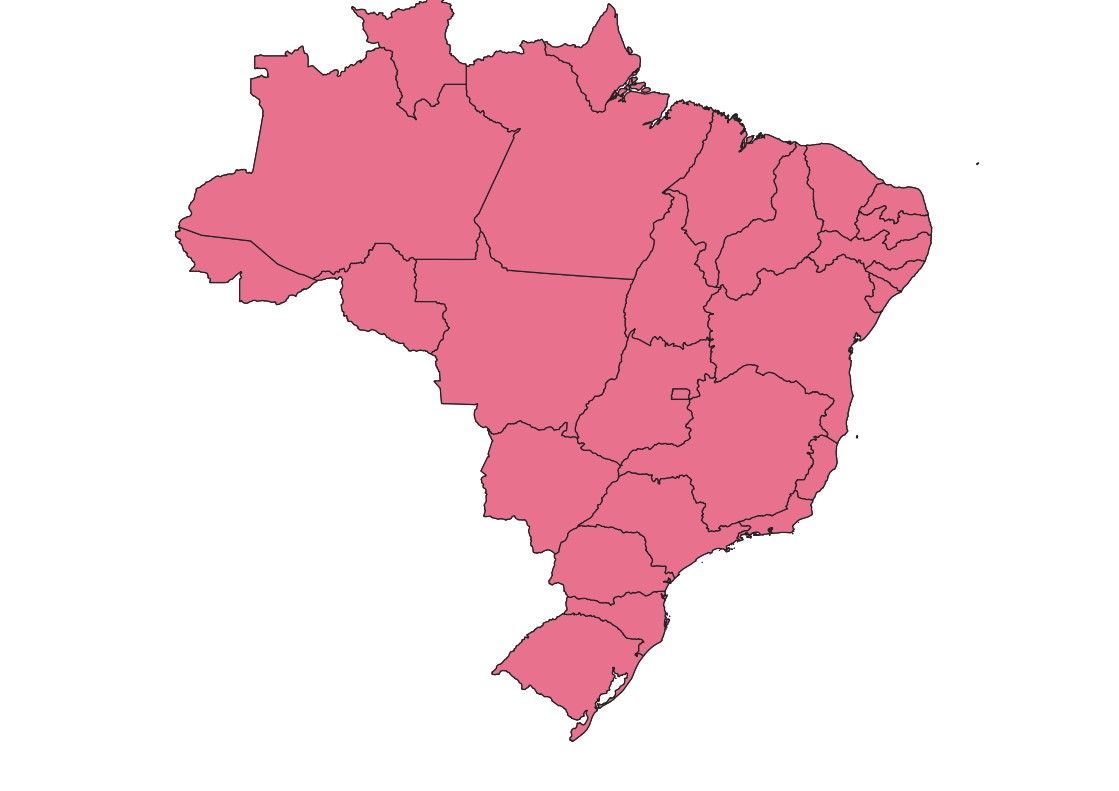


Figure 12: The New Layer of the Dissolved

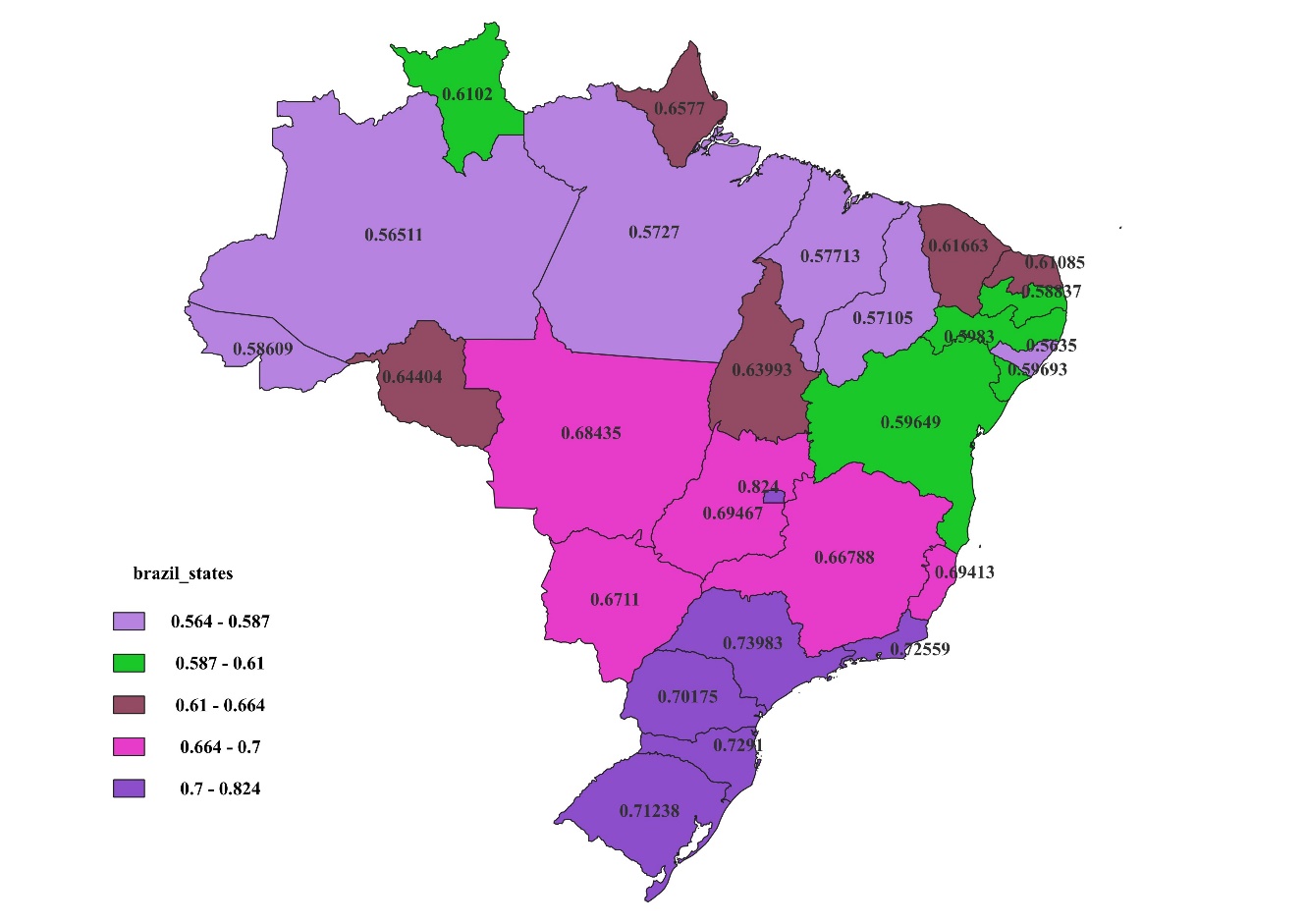


Figure 13: Displaying the states and their Human Development Index (MHDI) in Brazil

## **3.3 THE POLYGON/SHAPEFILE MAPPING OF THE MUNICIPALITY**



Figure 14: The Indigenous Territory Diagram

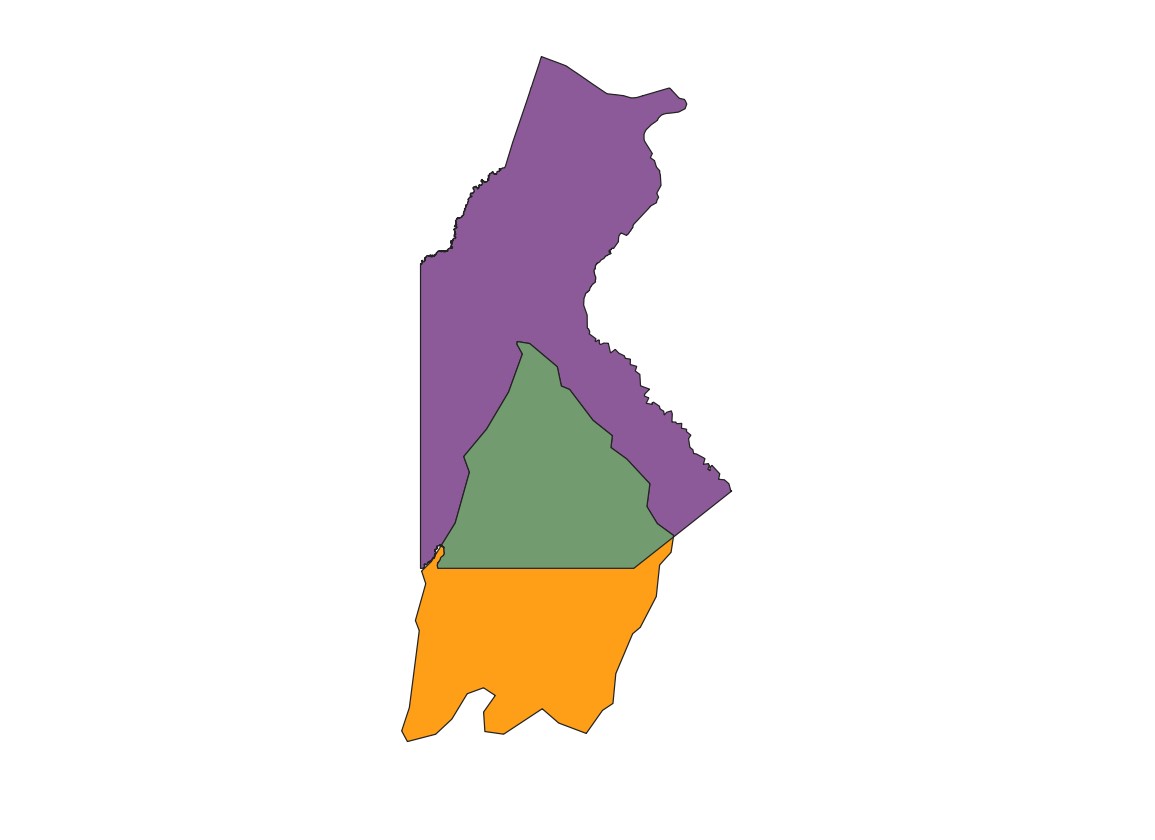


Figure 15: The Gaucha do Norte and Parque do Xingu new layer

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| id\_objeto | Nome | nomeabrev | geometriaa | perimetroo | Areaoficia | grupoetnic |
| 411 | Parque do Xingu | Parque do Xingu | NÃ£o | 33801 | 2642000 | Mentuktire, SuyÃ¡ |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| datasituac | situacaoju | Nometi | id\_produto | id\_element | codigofuna | COD\_MUN |
| 18/05/1987 | Declarada | Terra tradicional - Proc.concluÃ­do | 250002 | 29 |  | 5103858 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| UF | POP\_201 | IDHM\_10 | PIB\_PER | Area |
| MT | 6548 | 0.615 | 15926 | 8.08E+09 |

Figure 16: The area of the two regions Gaucha do Norte and Parque do Xingu

## 

## **3.4 THE BUFFER OF THE HOUSING PROJECT**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | 20 | 18 | 15 | 14 | 12 | 6 | 5 | 2 | 1 | 26 |
| 1 | 617681 | 595955 | 492471 | 592417 | 422161 | 287651 | 242104 | 24761 | 0 | 704221 |
| 2 | 594881 | 573004 | 468771 | 572810 | 398416 | 263272 | 217905 | 0 | 24761 | 680700 |
| 5 | 382700 | 359991 | 252046 | 387598 | 181720 | 46437.4 | 0 | 217905 | 242104 | 464337 |
| 6 | 343285 | 320150 | 209608 | 361577 | 139958 | 0 | 46437.4 | 263272 | 287651 | 421584 |
| 12 | 203928 | 180571 | 70390.8 | 248756 | 0 | 139958 | 181720 | 398416 | 422161 | 282639 |
| 14 | 153524 | 158106 | 216378 | 0 | 248756 | 361577 | 387598 | 572810 | 592417 | 240552 |
| 15 | 138203 | 114460 | 0 | 216378 | 70390.8 | 209608 | 252046 | 468771 | 492471 | 212292 |
| 18 | 23748.8 | 0 | 114460 | 158106 | 180571 | 320150 | 359991 | 573004 | 595955 | 116052 |
| 20 | 0 | 23748.8 | 138203 | 153524 | 203928 | 343285 | 382700 | 594881 | 617681 | 101458 |
| 26 | 101458 | 116052 | 212292 | 240552 | 282639 | 421584 | 464337 | 680700 | 704221 | 0 |

Figure 17: Distance Matrix of the interaction between Gaucha do Norte and Parque do Xingu

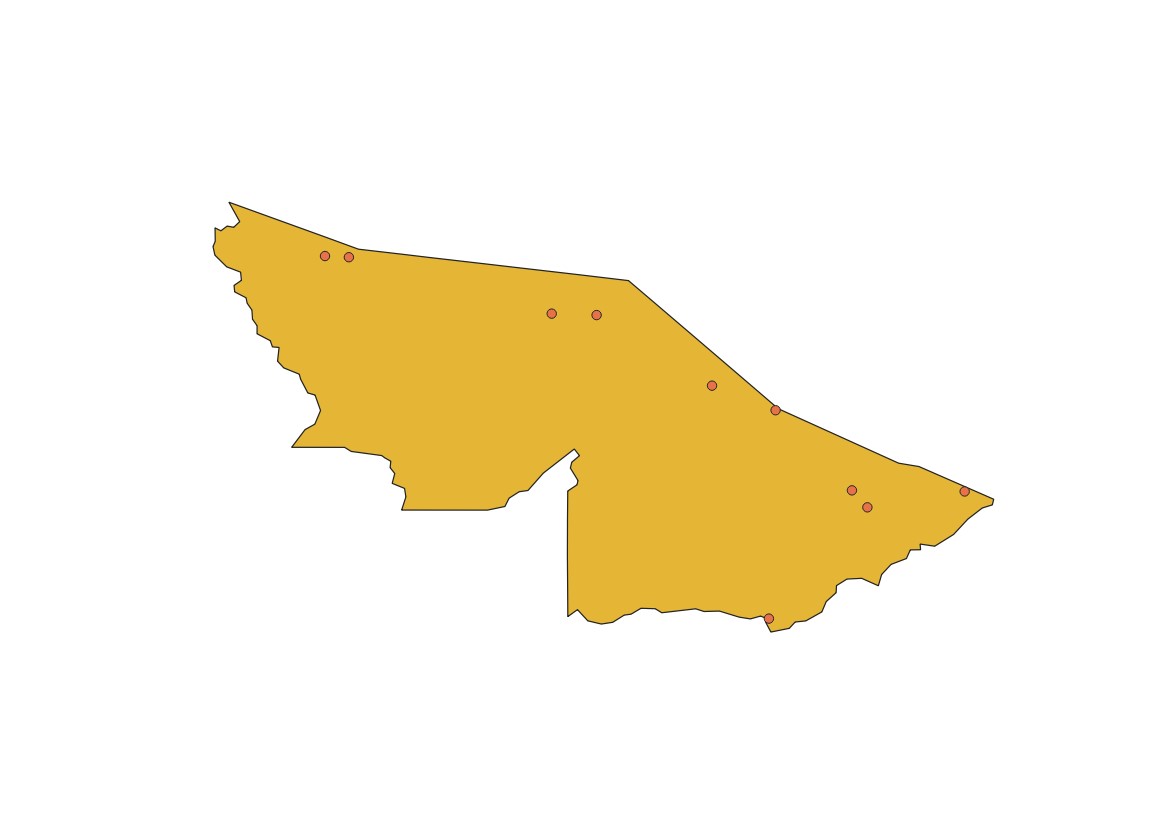


Figure 18: The new layer derived from housing and AC\_state

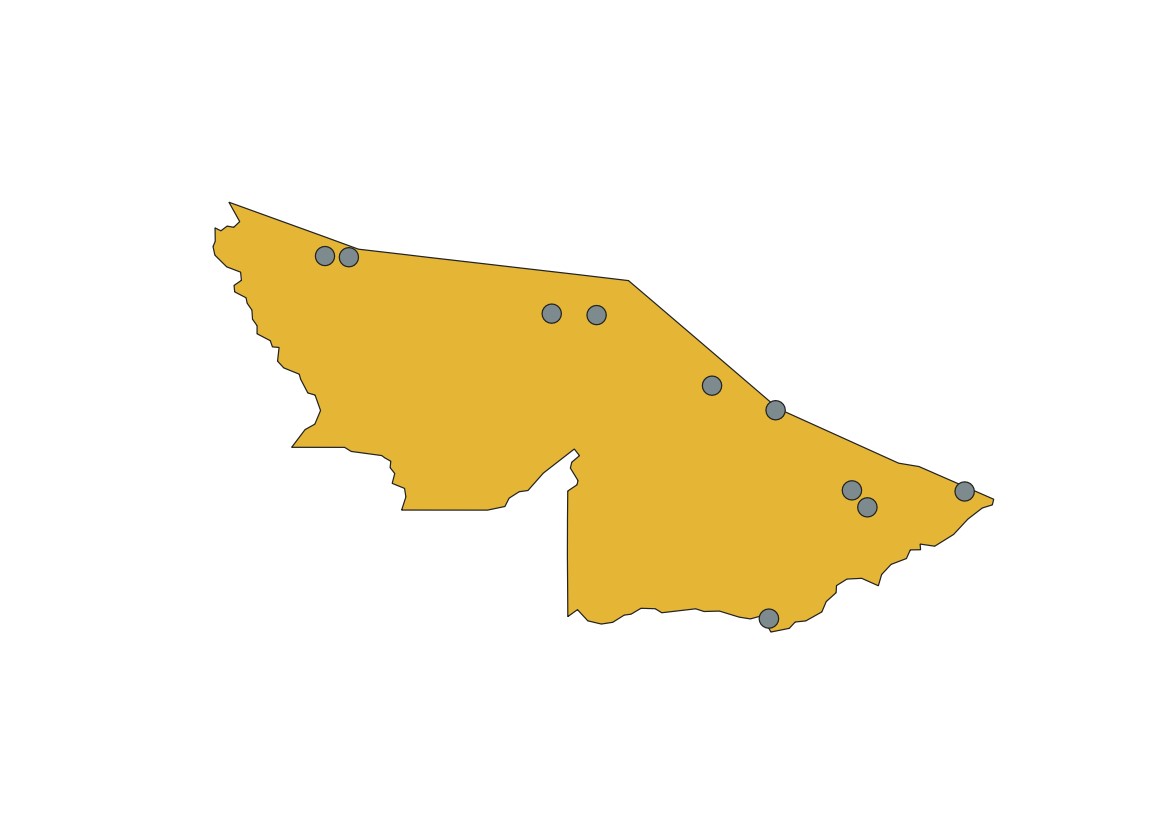


Figure 19: 10km buffer around each housing project

## 

## **3.5 THE DISTRIBUTION OF THE TOTAL HOUSING UNITS BY MUNICIPALITY**

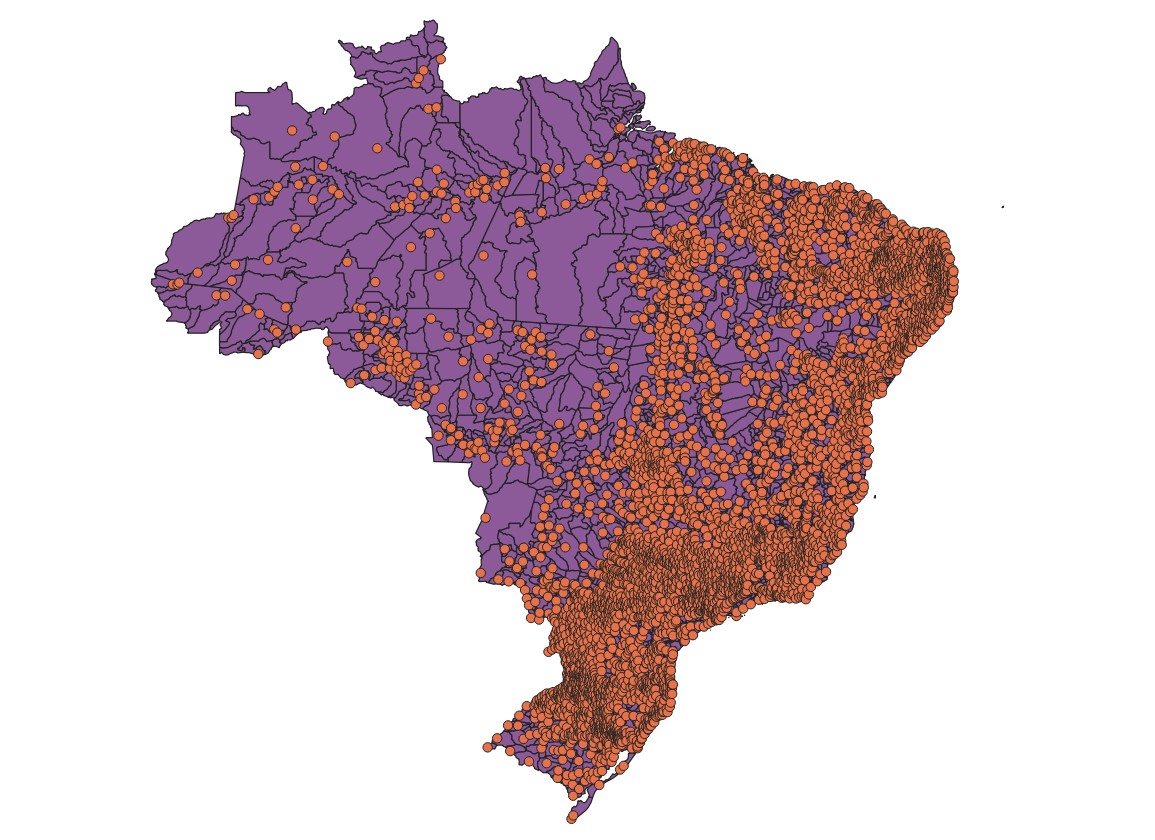


Figure 20: Attribute joining based on spatial location

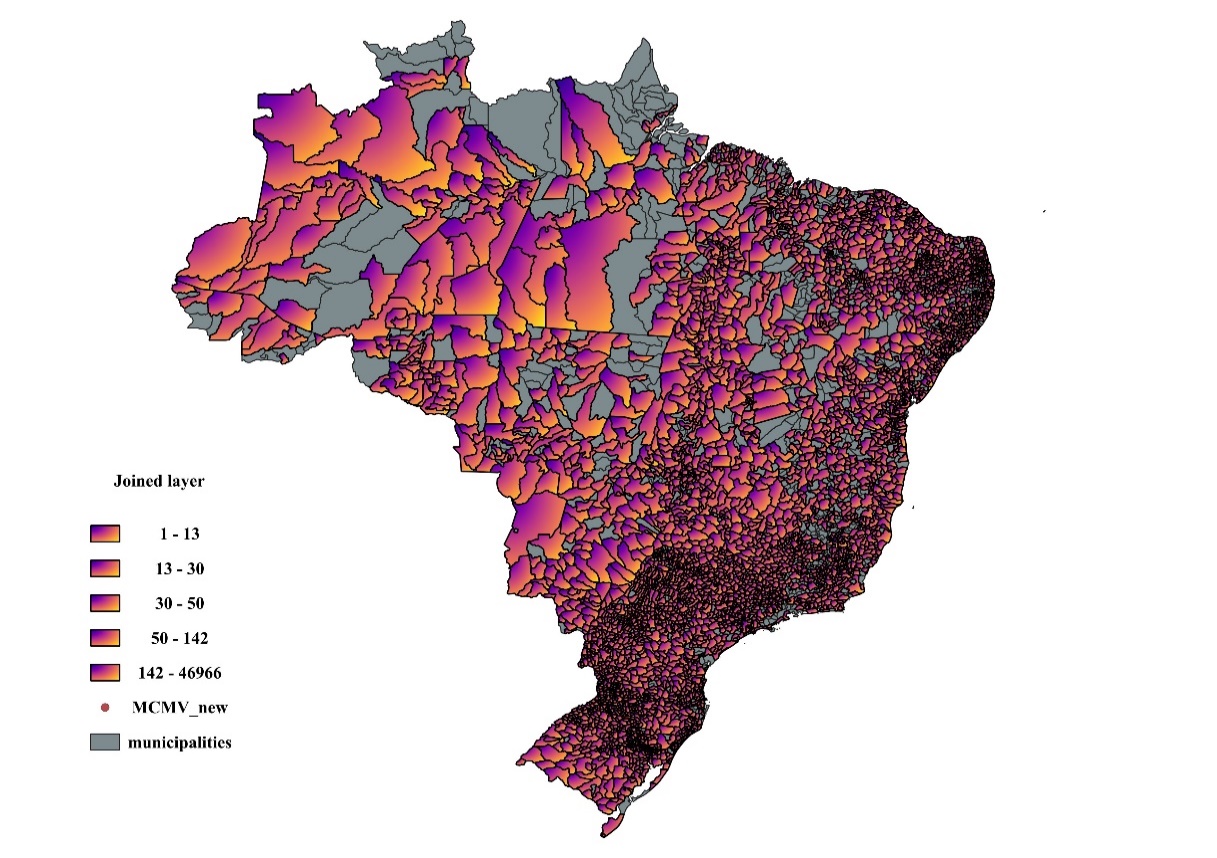


Figure 21: Map showing the distribution of total housing units

# **CONCLUSION**

The average “Human Development Index for municipalities in each state of Brazil was computed by” using a field calculator from the attribute table that was selected from the layer panel, the output field name was recognized as the mean\_HDI and the output field type was recognised as the Decimal number (real) which by selecting an expression as mean (“IDHM\_10”, group\_by:= “UF”) bring out the result of mean value under the field of mean\_HDI.

Indigenous and municipalities were filtered to get the Gaucha and Xingu layer by using an operator which is equal to represent the field and values to get the Gaucha do Norte and Parque do Xingu. The area of the two regions which are Gaucha do Norte and Parque do Xingu are 8.08E + 09 and also the buffer of 10km was created around each housing project and the matrix distance has a diagonal of zero factors which is an integer value (20 by 20, 18 by 18, 15 by 15, 14 by 14, 12 by 12, 6 by 6, 5 by 5, 2 by 2, 1 by 1, and 26 by 26), consisting of 10 Identification Number (ID) and 10 fields.

In the last question, we have indicated the highest and lowest number of housing by joining two layers which are MCMV and Municipalities by using the legend to indicate the territory in the map of a particular location.

# **WEEK 4 - FORMATIVE TASK 4**

# **QUESTIONS**

4.1 Compare and contrast the advantages and limitations of different thematic map types.

4.2 Identify and discuss the best methods to represent population density in a thematic map.

4.3 Why should a choropleth map (almost) always show derived data? Provide a mapping example when the exception to this rule applies.

4.4 Describe why you should classify some thematic map types but not others.

4.5 What visual variables and symbol dimensionalities are used for each thematic map type and how do these differences impact their design and use?

4.6 Create a choropleth map showing the unemployment rate within a selected country. Design an appropriate legend for the map.

4.7 Create proportional and graduated symbol maps for the same dataset and discuss the differences between the maps. Design an appropriate legend for both maps.

4.8 Sketch a legend for each thematic map type: choropleth, proportional symbol, graduated symbol, isoline, dot density, dasymetric, and flow.

## **4.1 THEMATIC MAP TYPES**

Thematic maps known as subject or specific maps, are created to illustrate particular topics or themes (Daniel, 2023). Rather than providing a general overview of a geographic area, their primary purpose is to convey detailed information on population density, political boundaries, and vegetation (Maurizio, 2004).

**CHOROPLETH MAPS**

These maps visually represent information through colour or pattern, typically at the scale of geographic units like census tracts or countries. Class intervals enhance data readability by grouping data into three or more categories using natural breaks or quantiles, thereby improving data presentation. (Daleska, 2022).

**ADVANTAGE OF CHOROPLETH MAPS**

• Choropleth maps visually group data into distinct colours, facilitating rapid comprehension of extensive datasets spanning multiple geographic regions.

• Creating a choropleth map is straightforward with tools like Maptive.

• They are highly effective for visualising variables into administrative units.

**DISADVANTAGES OF CHOROPLETH MAPS**

* Choropleth maps can potentially distort data and mislead viewers if not used properly since the map designer defines the data categories.
* Maps assume uniformity within enumeration areas despite potential fluctuations.
* Rapid changes in enumeration area boundaries can be misleading.

**DOT MAPS**

These maps use dots to represent the distribution of a commodity, with each dot corresponding to a specific value.

**ADVANTAGE OF DOT MAPS**

* They are a great option for those who prefer black-and-white maps.
* Dot density maps frequently offer more precise representations than other types, such as heat maps.

**DISADVANTAGE OF DOT MAPS**

Some software places dots randomly within enumeration areas, sometimes placing them far from the phenomena they represent. Moreover, choosing a single dot size that visually accommodates both high-density and low-density regions is challenging due to the diversity of data points. Contextual comprehension is crucial for interpreting dot density maps.

**PROPORTIONAL SYMBOL MAPS**

Proportional symbol maps use symbols whose sizes are proportional to the values they represent. These maps are adequate for analysing the “spatial distribution and magnitude” of variables. The advantage of “proportional symbol maps is their ability” to depict relative magnitudes among variables and handle large datasets efficiently. However, if data scaling is inadequate or if the scale lacks a zero value, the accuracy of proportional symbol maps can be compromised, leading to potential misinterpretations (Jonathan, 2024).

**ISOLINE MAP**

Isopleth maps are a type of isorhythmic mapping, where the “values reflect a standard rate per unit area”, directly derived from the data. In this method, lines represent areas of consistent values for a variable (Robert H., et al., n.d.). They are particularly effective for displaying the spatial distribution of continuous variables such as height, temperature, and precipitation.

## **4.2 POPULATION DENSITY IN A THEMATIC MAP**

**CHOROPLETH MAP**

These are a popular method for depicting population density on thematic maps. They use colour to represent visual variations in population across different geographic regions. This approach is intuitive and effective for highlighting patterns in population distribution (Daleska, 2022).

**DOT MAP**

Referred to as point density maps, dot maps depict “population distribution” by placing dots on a “grid” to represent individual residents. Dot maps handle large datasets effortlessly and pinpoint exact locations accurately. However, displaying too many points of interest can make point maps difficult to interpret (A. John, 2013).

**PROPORTIONAL MAP**

Maps can represent population density using proportional symbols, such as circles or squares, to display visually the density of people in an area. This method efficiently manages large amounts of data and is adequate for illustrating population densities across different regions. However, inaccuracies in scale or incorrect symbol sizes can lead to misleading representations on proportional symbol maps.

**ISOLINE MAP**

The Isoline map presents a three-dimensional perspective. This method uses lines to illustrate areas with consistent population densities, effectively portraying patterns of population density across geographical regions and handling extensive datasets efficiently.

## **4.3 THE DERIVED DATA OF THE CHOROPLETH MAP**

Choropleth maps typically display aggregated data at the level of geographic units like census tracts or counties, ensuring that derived data is represented consistently (Cromley, 1996). This approach enables accurate comparisons throughout the map, even when there are variations in the sizes of these units (Wangshu & Daoqin, 2020). Moreover, incorporating derived data on choropleth maps helps account for differences in the sizes of geographic units, facilitating more accurate data comparisons across the map.

## **4.4 CLASSIFICATION OF THEMATIC MAP TYPES**

Classification is essential for specific types of thematic maps, such as “choropleth and isorhythmic maps”, as they portray aggregated data at the level of geographic units, such as census tracts or counties (Cromley, 1996). Creating data into meaningful groups through classification simplifies analysis and comparison (Cromley, et al., 2015). It enables consistent comparisons throughout the map despite variations in the size of geographic units. For instance, when creating a choropleth map, data is classified by grouping census tracts or counties into distinct population density categories.

## **4.5 VISUAL VARIABLES AND SYMBOL DIMENSIONALITIES**

**CHOROPLETH MAPS**

Choropleth maps are a type of statistical decision where data is represented by colour or pattern, indicating aggregated findings within geographic areas like census tracts or countries. Typically, colour is used as the visual variable, with the geographic unit's size held constant. The variation in colour or pattern is used to convey data, influenced by the number of classes, colour scheme, and data classification method (Wangshu & Daoqin, 2020).

**PROPORTIONAL MAPS**

Symbols such as circles and squares, among others, are appropriate for showing the absolute value of a variable at specific points on a map with appropriate proportional symbols. Visual factors such as position, size, and colour are appropriate. To create a proportional symbol map, place a symbol at the variable's location and adjust its size accordingly. The capacity of each image is often chosen so that its area reflects the variable's magnitude, though alternative methods like classifying symbols by size ("small," "medium," "large") are also appropriate.

**DOT MAPS**

Dot maps, also known as dot density maps, depict the prevalence of a variable across a specific area using individual dots. Positional visual cues are created by placing dots where the variable is present. Appropriate position and quantity enable the visualisation of variable distribution throughout a region and are particularly effective for illustrating unusual events. However, excessive dot density on the map can hinder readability and interpretation.

**ISOPLETH MAPS**

Isopleth maps, which are a type of contour map, employ lines to indicate regions where a variable maintains consistent values. Since isopleths are one-dimensional, their location is the sole visual variable that can be adjusted. These maps create the spatial distribution of a variable by connecting locations that share the same value. However, if the lines on an isopleth map are too closely spaced, it can impede readability and interpretation.

**FLOW MAPS**

Flow maps, commonly called streamlined maps, illustrate the movement of people, goods, or information between different geographic areas. The visual variables used are position and orientation, and the symbols are flat. Creating a flow map involves drawing lines from origin to destination on a map and varying the line thickness to indicate the magnitude of the variable in question. While location and direction effectively represent the flow's direction and strength, maps with excessive or closely spaced lines can be challenging to interpret.

## 

## **4.6 CHOROPLETH MAP FOR THE UNEMPLOYMENT RATE IN A COUNTRY**

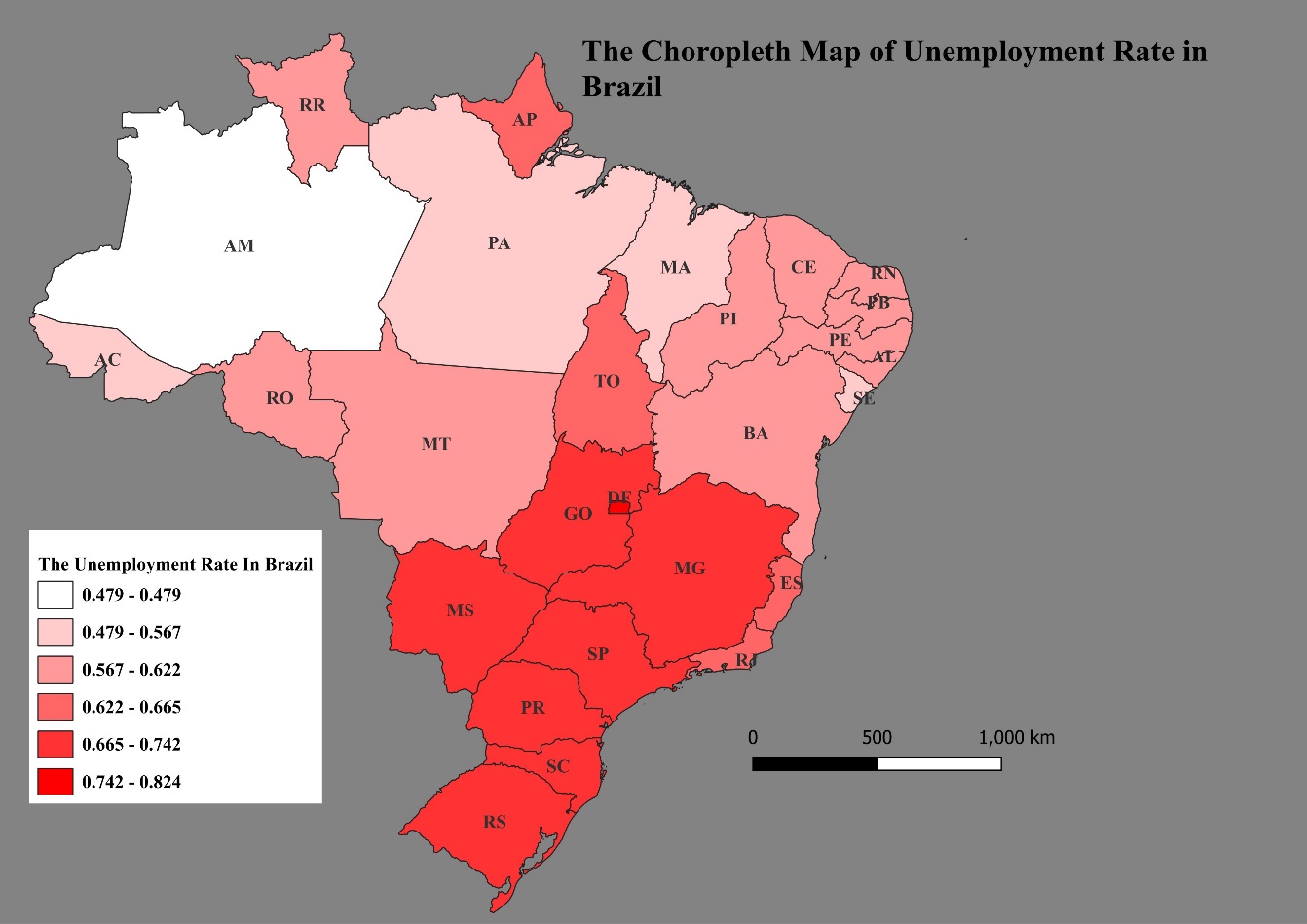


Figure 22: The Unemployment rate of Brazil

**4.7 PROPORTIONAL AND GRADUATED SYMBOL MAP**

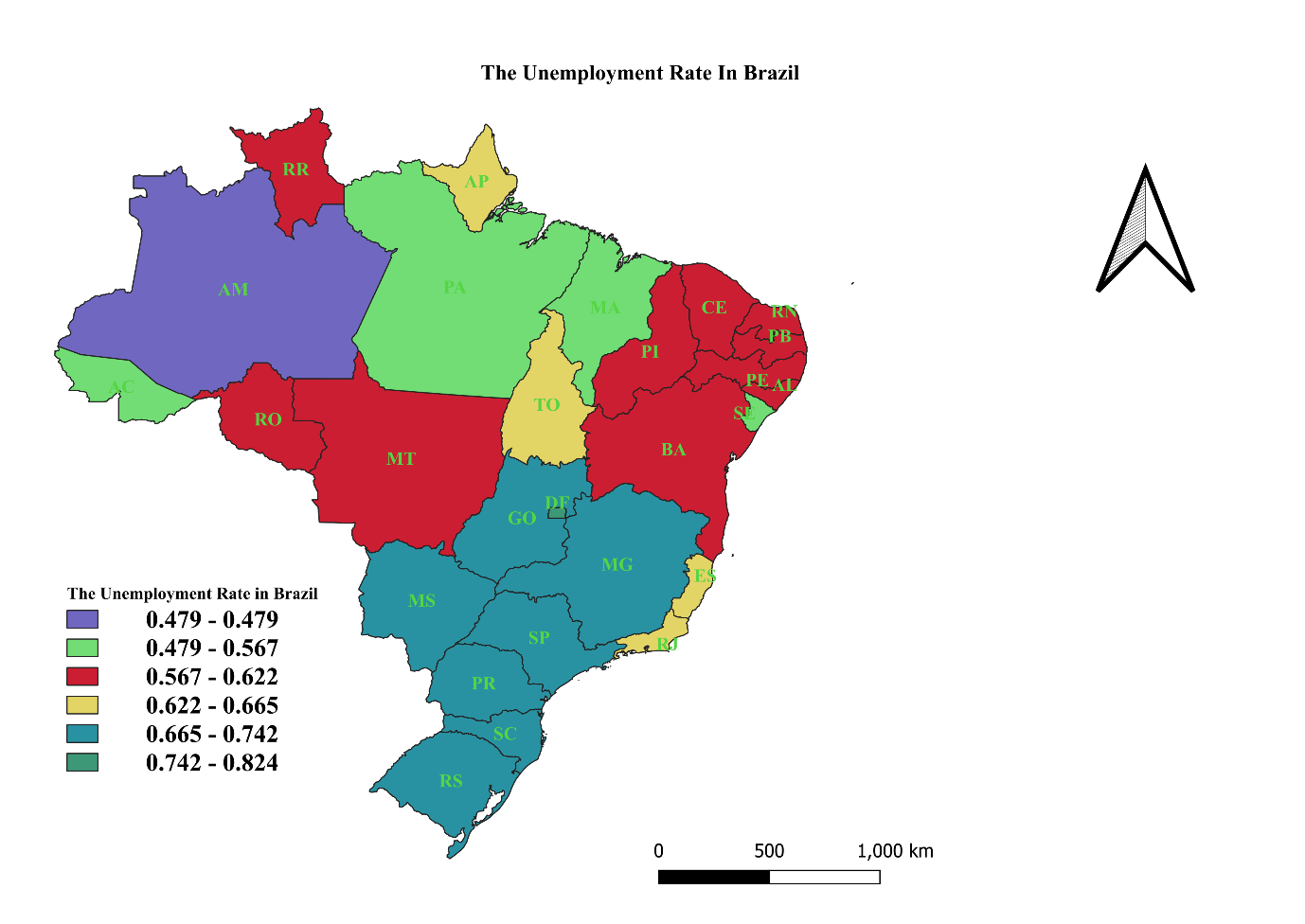


Figure 23: The Unemployment rate of Brazil using the graduated data

## **4.8 THE LEGEND OF THEMATIC MAP TYPES**

1. Choropleth Map

|  |
| --- |
|  |

10 – 20%

|  |
| --- |
|  |

20 – 30%

|  |
| --- |
|  |

30 – 40%

|  |
| --- |
|  |

40 – 50%

1. Graduated Map

|  |
| --- |
|  |

|  |
| --- |
|  |

10 – 20 %

20 – 30%

|  |
| --- |
|  |

30 – 40%

1. Proportional Symbol

|  |
| --- |
|  |

10%

|  |
| --- |
|  |

20%

|  |
| --- |
|  |

30%

|  |
| --- |
|  |

40 %

1. Dot Map

|  |
| --- |
| . |

10,000

|  |
| --- |
| ….. |

20,000

|  |
| --- |
| …….  …… |

30,0000

1. Flow

0 – 50 bottles

51 – 100 bottles

101 – 200 bottles

201 – 300 bottles

1. Dasymetric

1 - 50

51 - 150

151 – 250

# **CONCLUSION**

Lastly, we have understood the knowledge and understanding of spatial data and Geographic Information models (GIS) by showing the ability to answer and identify, explain and discuss thematic map type, advantages, limitations/disadvantages, uses, and unemployment rate of the country which is Brazil, the diagram and symbol to demonstrate the thematic map type in Geographic Information System models (GIS). This was done by answering the question correctly in Geographic Information Systems (GIS) and spatial analysis.

# **REFERENCES**

A. John, K., 2013. Dottiing the Map, Revisted. 14 March, 36(2), pp. 165 - 182.

A.H.M, t. H., H.A., P. & Th.P., V. d. W., 2004. Query Formulation as an Information Retrieval Problem. *The Computer Journal, 39(4):255-274, September 1996,* 23 June.

Adam, H., 2024. Demographics: How to collect, Analyze, and Use Demographic Data | Investopedia. 18 June.

Alex, W., 2024. What are the Physical Features in Geography? |Speeli. 31 January.

Anh, N., Jason, Y. & Jeff, C., 2016. Multifaceted Feature Visualization: Uncovering the Different Types of Features. arXiv preprint: 1602.03616. May.

Christopher, B. J., 2014. Geographical Information Systems and Computer Cartography. Routledge.

Cromley, R. G., 1996. A comparison of optimal classification strategies for choropleth displays of spatially aggregated data. *International Journal of Geographical Information Systems,* 10(4), pp. 405 - 424.

Cromley, R. G., Zhang, S. & Vorotyntseva, N., 2015. A concentration-based approach to data classification for choropleth mapping. *International Journal of Geographical Information Science,* 29(10), pp. 1845 - 1863.

Daleska, P., 2022. Choropleth Map: Definition, Uses & Creation Guide | Venngage. 17 May.

Daniel, O., 2023. What is a thematic Map. 6 March.

Debra, L., Michela, B., Sean, W. M. & Bianca Schon, 2008. Three-Dimensional Spatial Information Systems: State of the Art Review. *Recent Parents on Computer Science 2009, 2, 21 - 31,* 14 November.Volume 2.

G., B., 2009. Natural Resources | ScienceDirect. *International Encyclopedia of Human Geography.*

Jonathan, S., 2024. Bivariate Proportional Symbol Maps, art 1: An Introduction. *IPUMS Research Scientist, NHGIS Project Manager,* 7 February.

Lampros, K. M. & Lambros, S. A., 2017. Propagation of segmentation and imaging System Errors. *Atherosclerotic Plaque Characterization Methods Based on Coronary Imaging.*

Maurizio, P., 2004. Simple Thematic Mapping. *The Stata Journal: Promoting Communications on Statistics and Stata,* December, 4(4), pp. 361 - 378.

M, T., 2018. map: Thematic Maps in R. *Journal of Statistical Software,* 20 April.

Rao, K. N., 2017. Unit-5 GIS Data Models and Spatial Data Structure.

Reddy, O. G., 2018. Spatial Data Management, Analysis, and Modeling in GIS: Principles and Applications. *In Geospatial Technologies in Land |Resources Mapping, Monitoring and Management,* September. pp. 127 - 142.

Robert H., W., Russell H., C. & Tor H., N., n.d. Preparation and use of Isopleth maps of landslide deposits | USGS.

Ron, B., 2008. The Digital Representation of Visual Information. *The Art and Science of Digital Compositing (Second Edition).*

Rudolf, W. V. d. p., 2000. Knowledge-based Query Formulation in Information Retrieval. 14 September.

Samuel, T. T. et al., 2016. The topology of geology 1: Topological analysis. *Article in Journal of Structural Geology, 91,* August.

Wangshu, M. & Daoqin, T., 2020. Mapping uncertain geographical attributes: incorporating robustness into choropleth classification design. *International Journal of Geographical Information Science,* 03 Feb, 34(11), pp. 2204 - 2224.