Term Paper Report

CE432A: Geographical Information System **2023-24/II**

Name: Anukul Dwivedi Roll Number: 231030402

Satyam Agnihotri 231030406

Building Web GIS for Visualizing India's Demographic Distribution at Pincode Level in India

---

Abstract

This term paper presents the development of a Web Geographic Information System (GIS) which aims to visualizing India's demographic distribution at the pin code level using census data from 2011. By using Geographic Information System (GIS) technology and demographic data sourced from governmental websites, this study focuses on representing population density across India's geographic regions and the application of connecting pin code with the population details. The methodology involves the integration of PIN (Postal Index Number) data with demographic statistics, in our study we assumed an equal distribution of total population of district to each Pincode area of respective district. By using QGIS a open source software, we try to make a web map which can show the population detail at a particular pin code to the people in a user-friendly manner. The Web GIS platform provides an intuitive interface for users to explore and analyse demographic trends at a granular level, facilitating informed decision-making in various domains such as to know the migration pattern, urban planning, resource allocation, and public policy formulation. The findings demonstrate the potential of GIS technology in harnessing census data to generate valuable insights into population dynamics, thereby contributing to the advancement of spatial analysis and decision support system.

Introduction

Change is the law of nature, but here in case of locating human address we are using pincode from centuries which is unique 6-digit number for locating addresses and available in everyone’s AADHAR card so let make perfect use of it by joining them with the population data to unique pin code. The political way of acquiring these data ward wise or village wise is limited to make voter ID’s for the elections.

The next generation analytics and visualisation is the theme of this term paper, addresses the pressing need to develop a method for visualizing and analysing India's demographic distribution at the pin code level, using census data from 2011. The problem statement is the crucial importance of understanding demographics in shaping decisions across different fields like city planning, distributing resources, and creating public policies. Given that India is one of the world's most populous countries, having precise and easy-to-access demographic information is vital for making effective plans and policies.

The usual methods of examining information, such as using spreadsheets or simple maps, often lack the level of detail needed for effective decision-making. While governments share data, understanding and visualizing it can be challenging. Therefore, we aimed to fill this gap by

creating user-friendly tools that enable efficient utilization of this data. These tools are crucial for various decision-making processes, including locating hospitals, banks, managing road networks, and overseeing sanitation and water resources.

We are in search of something better to understand exactly where people are in India and how the demography is changing so our work of building this webGIS act as a reference point. Also, the private banks use Pincode to estimate the average income of the customer residing in the locality and accordingly sets the average credit limit for that locality, so by finding the lower performing Pincode in terms of income can be target easily and certain specific decisions can be taken to uplift them. This had become a one of the key points to build this WebGIS.

Methodology

WebGIS democratizes access to geographic information by making spatial data more accessible, interactive, and easy to use for a wide range of users, to make this user friendly webGIS which gives the demography of India based on the pincode we have performed following tasks:

1. Data Collection:
    - Collected demographic data pertaining to India.
    - Focused on data relevant to pincode areas.
2. Data Standardization and Filtering:
    - Standardized collected data for consistency.
    - Filtered data to ensure accuracy and relevance.
3. Data Integration:
    - Integrated different sets of demographic data.
    - Ensured seamless compatibility and coherence among datasets.
4. Data Visualization:
    - Utilized visualization techniques to represent demographic data effectively.
    - Created visual representations such as charts, graphs, and maps.
5. Web Map Development:
    - Developed a user-friendly web interface.
    - Integrated the visualized demographic data into an interactive web map.
    - Implemented features for easy navigation and exploration of demographic information based on pincode areas.

Data Collection

Data Sources and Data Tables:

1. Data Table 1:
    - Source: India Post.
    - Content: This table contains a pin code and their respective district, they lie.
2. Data Table 2:
    - Source: ArcGIS online.
    - Content: This table provides a relationship between pin codes, district names, and their corresponding states.
3. Data Table 3:
    - Source: Census data from Government of India.

- Content: This table contains population details associated with district names.

4. Shape Files:
    - Shape File 1:
    - Content: District boundaries of India
    - Usage: Used for spatial analysis and visualization.
5. Shape File 2:
    - Content: State boundaries of India
    - Usage: Also utilized for spatial analysis and visualization.
6. Shape File 3: Pin code as the point feature on the map with attached attribute table to it.

By using these different sources and datasets, we aimed to establish a common identifier that connects pin codes with population details. The first data table provided the pincode with name of the distict and the state. The second data table offered insights into the relationship between pin codes and latitude, longitude for pin codes facilitating spatial analysis and visualization. Additionally, the third data table enriched the analysis by providing population details associated with district names. Furthermore, the utilization of shape files containing district and state boundaries enabled comprehensive spatial analysis within the WebGIS platform. Through the integration and analysis of these datasets and shape files, our aim was to create a robust and informative platform for understanding demographic trends across India.

Data Standardization and Filtration

To standardize our data, we used MATLAB code to remove the unwanted points and data to make our dataset refined.

Data Integration

Joining Data Tables in QGIS:

1. Data Tables:
    - Three separate data tables containing distinct sets of information.
2. Import into QGIS:
    - Each data table was imported into QGIS for further processing and analysis.
3. Join Function:
    - Utilized the join function within QGIS to combine the data tables.
    - Defined primary and foreign keys to establish the relationship between the tables.
4. Joining Process:
    - Selected appropriate primary and foreign keys from each data table to establish connections.
    - Executed the join function to merge the tables based on the defined keys.
5. Output:
    - Combined dataset created, incorporating information from all three data tables.
    - Enabled comprehensive analysis and visualization within QGIS.

By employing the join function in QGIS and defining primary and foreign keys, we successfully merged the three separate data tables into a unified dataset. This integration

facilitated a more holistic understanding of the demographic information and spatial relationships within the dataset, enabling effective analysis and visualization.

Table 1

Pin Code	Latitude	Longitude	----
208016	26.511552	80.231247	
581109	14.550118	75.444374	
301702	28.015759	76.719094	
791104	28.191607	95.407424	

Table 2

Pin Code	District	---------
208016	Kanpur Nagar	
581109	Haveri	
301702	Alwar	
791104	East Siang	

Table 3

District	Total Household	Total Population	Male Population	Female Population	------
Kanpur Nagar	22137	117468	63072	54396	
Haveri	11013.8	53256	27305	25951	
Alwar	13328	76545	40396	36149	
East Siang	4972	25283	12728	12555	

Firstly, we joined ‘Table 1’ with ‘Table 2’ by using primary key as ‘Pin Code’ and then join ‘Table 2’ with ‘Table 3’ by using ‘District’ as a foreign key. Now we connected all the table with each other to get the relationship between Pin code and the population details.

This results a big data set of single attribute table with the available Latitude and Longitude in Geographic coordinate system which needs to be marked on map, for that vector layer of point feature is plotted by using the coordinates.

For further process and ease to locate coordinates, we transformed our Geographic Coordinate WGS-84 to Projected Coordinate System i.e., WGS-84 UTM projection in zone 44N.

In our work we have taken one assumption that we divided the population of the district equally to every pin code of that district, although data is available at the ward/sub-district level which need to organise effectively. Here we had not implemented this thing because of the huge data file brought high latency rate.

Data Visualisation

Now we have a shape file of the data which is obtained after joining small-small datasets. We also imported the district and state boundary shapefiles and plotted them on a map.
Visualised the pin code that it is correctly distributed in their respective district and the state. Also, for the confirmation of pin code distribution we combined different datasets using MATLAB and plotted that and compared with the plot obtained from GIS software.

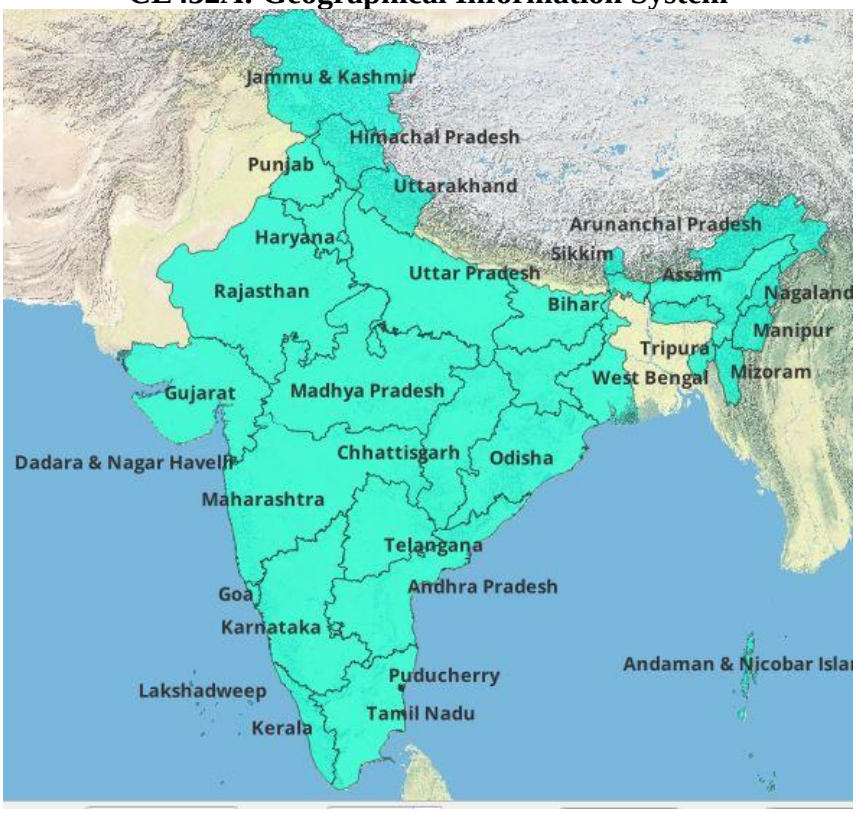

Figure 2: State boundary layer (Source: Open Source)

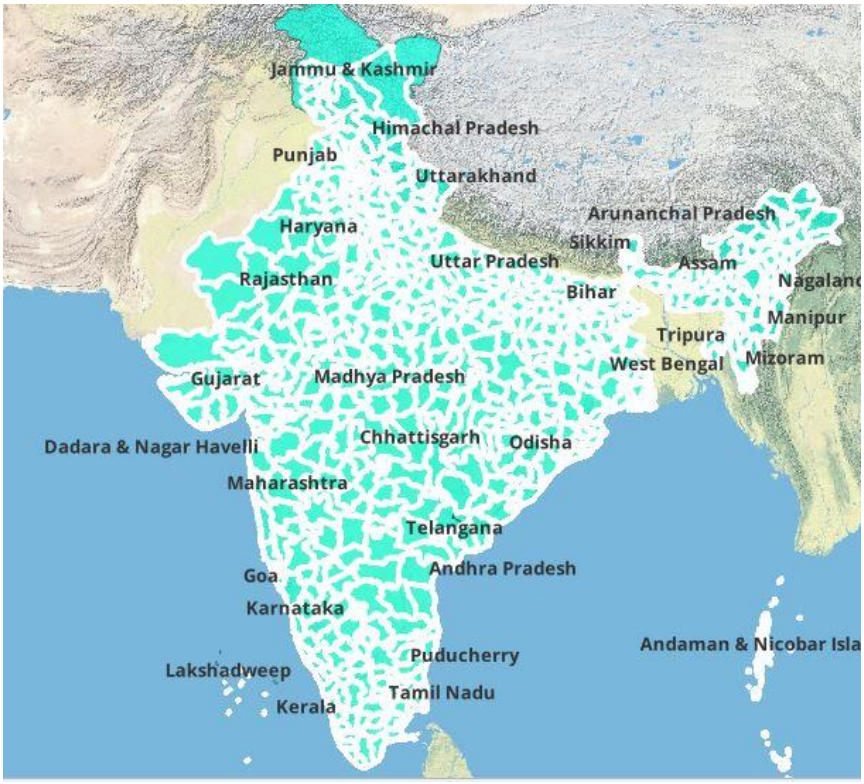

Figure 3: District boundary layer with state boundary (Source: Open Source)

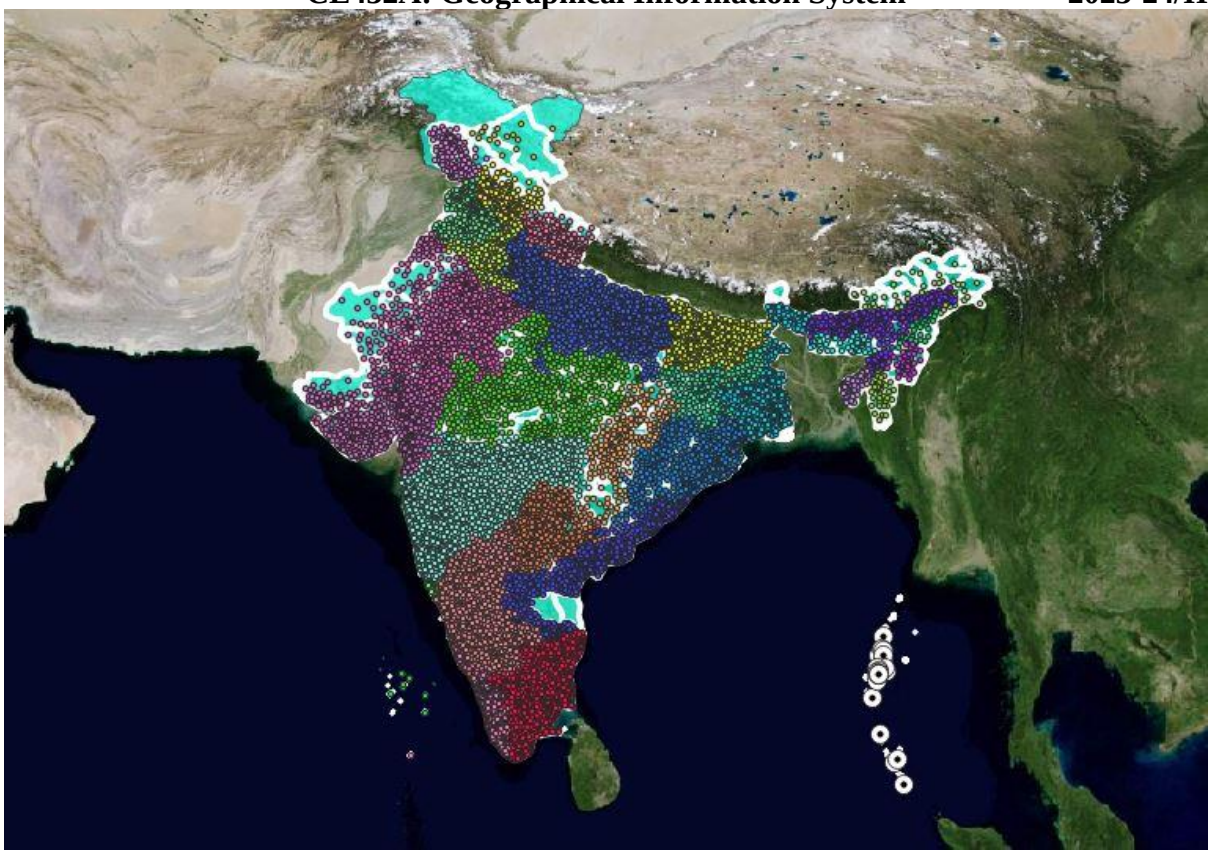

Figure 4: Pin code in India as point feature.

Figure 4 represents all the pin codes available in India except the pin code starting with 9 which is used by the Defence department of the India and some pin codes of the Andhra Pradesh due to problem in district name interpretation.

Till now we have integrated the data and also visualised it, now we are ready to create our web map.

Created Web map

The final step involves building the webGIS application, where users can interact with the data. we can use libraries like Leaflet.js or Mapbox to create interactive maps and integrate them into a web application for easy access.

Figure 1: Libraries (Google Images).

Step1: Imported all our layers to the QGIS and installed the QGIS plugin OpenLayers (This plugin used to integrate various web mapping libraries into QGIS) and qgis2web (This plugin is used to convert styled QGIS project into a web map format) to build web map.

Step2: Added google map as our base map and done some changes in symbols, labels and colours of the layers.

Step3: Now, we converted these layers as a web map by using qgis2web plugin.

Web->qgis2web->Create web map.

Figure 5: QGIS (location where we find Create web map)

Step4: By clicking on create web map a pop menu of export to web Map appeared in which we chose what layers, groups we want to show and what not to show and also have done settings for appearance of our web map (Such as: 'Add layer list', 'Geolocate user', 'Highlight o hover', etc). By clicking on update preview we can see the real time changes in our web map by changing the settings. Now, click on 'Export' to get export a web map, we have used 'leaflet' to export our web map but there are other options such as 'Open layer', 'Mapbox GL JS' which can be used.

Step6: We obtained a web map file comprising JavaScript, CSS (Cascading Style Sheet), Geopackage file and HTML formats. The HTML file allows us to use and display our web map.

Alternative way of web GIS: We can export the (.KML) file of our datasets and we can upload that on Google Earth pro to create an interactive web GIS. Google Earth Pro allows us to import KML (Keyhole Markup Language) files, which contain geographic data, and visualize them in a 3D environment. While Google Earth Pro provides a user-friendly interface for exploring geographic data, it is primarily a desktop application and does not offer the same level of customization and interactivity as web GIS platforms like Leaflet or Mapbox. However, if we are looking for a quick and simple way to share geographic data with others or visualize it in a 3D environment, uploading KML files to Google Earth Pro can be an effective alternative. Users will need to have Google Earth Pro installed on their computers to view the data, and the level of interactivity may be limited compared to web GIS platforms

Results and Discussion

After exporting layers to a web map using openlayer or leaflet, we get files in various formats to facilitate web map integration.

- **HTML (Hypertext Markup Language)**

This HTML file is the main document which contains the structure of the web map. This includes references to the necessary JavaScrip

- **JavaScript (JS) Files**
  These files contain the logic and functionality of the web map. It defines how the layers are displayed, interactions with the map, and any additional features or behaviours.
- **CSS (Cascading Style Sheet) Files**
  CSS files control the presentation and styling of the web map it defines the visual aspects such as colours, fonts and layout.
- **GeoPackage or GeoJSON Files (for layers)**
  Geographical data is exported to this file format, and it stores the spatial information such as geometry, attributes, and coordinate system details.

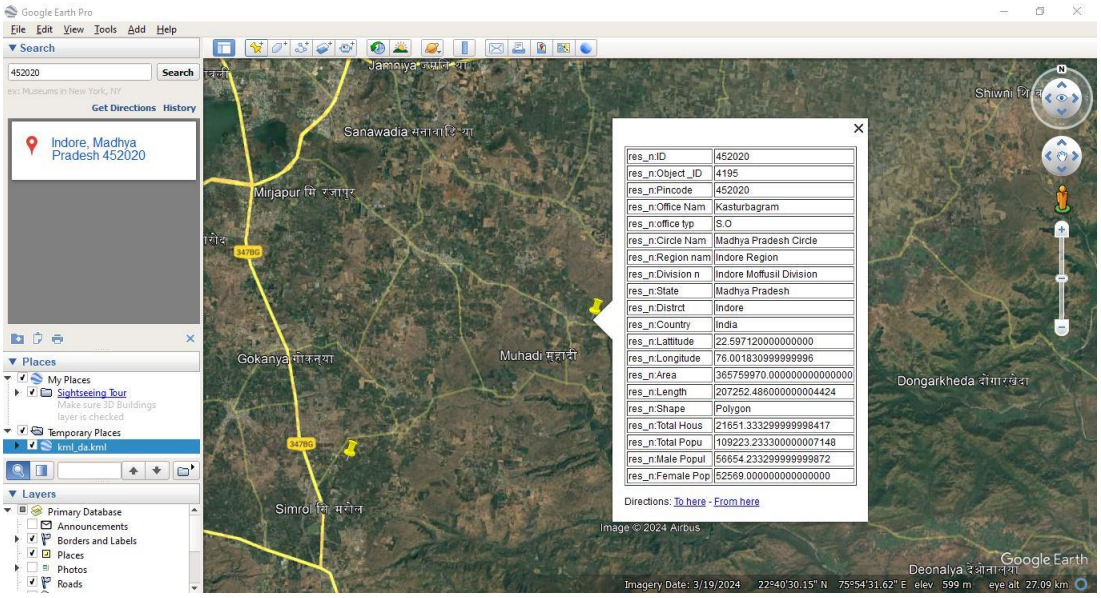

Figure 5: Stored data attached to pin code.

In our web map one can search the location based on the pin code and when the person click on the location it will gives the detail of total population, male population, female population and the corresponding district and the state name. This web map simplifies locating and accessing information for the general public. By utilizing this tool, we can accurately assess population distribution across the country. This knowledge enables us to allocate resources and govern more effectively.

This model will not be limited with the data of population further we can add all other data by recording with the primary key as the pin code and all other geographic related domains.

The new era of E-Governance can be enhanced by proper utilisation of this concept.

Further, by integrating village-level data, readily available, will provide a deeper understanding of population dynamics and gives us a clear picture of various scenario's related to population.

Following steps are followed to reach at the searched pin code:

We searched for the pin code '208016' which is a pin code of a IIT Kanpur region so to arrive at this location, we typically navigate through the geographic hierarchy.

State -- > District -- > Pin code.

Figure 2: State boundary.

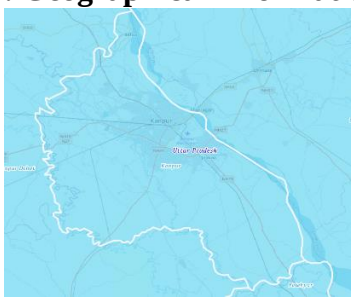

Figure 3: District boundary.

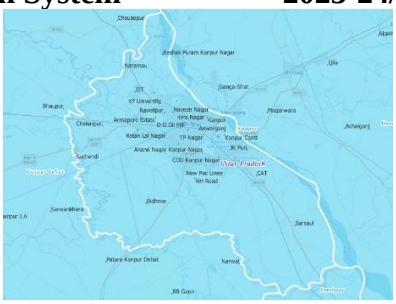

Figure 4: Pin Codes.

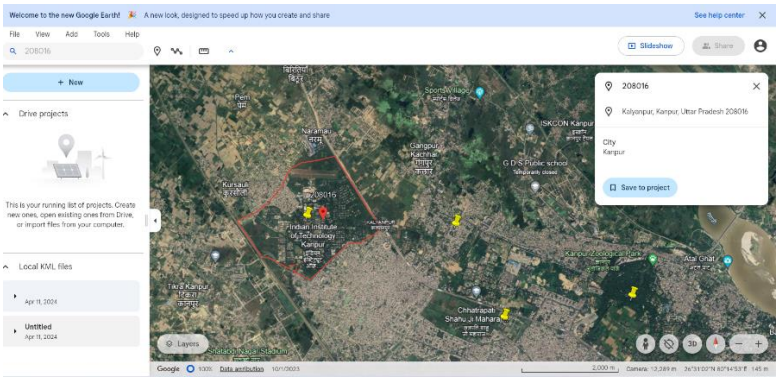

Figure 5: Searched pin code (Yellow).

There is a slight variation between the location of our point and the point given by the google map.

Conclusions and Future Work

With wide range of application of web GIS here we can conclude with the above layers of representing pin codes with the attached huge data base of different government agencies and to redesign the data with the unique Pincode and make a one stop destination for analysing the populations on different parameters which will enhance the governance in very effective way. This can be achieved by merging the different data base in future such as PAN card is linked with AADHAR card and these two are connected with pin codes, gives the information about the total tax collected from that particular area. The future work includes the integration of village and taluka level data to merge with Pincode which is available so that our assumptions of equally distributed population of district will be solved.

Today we all give location access to almost every app in day-to-day life where we are sharing location information by making such application, we can have real time data of traveling of personnel from one city to other which can be compared with our reference data at different epochs to know the demographic change and can predict the future requirements of resources such as water usage, sanitation etc.

The permanent shifting of person from one place to another then they have to update their address in AADHAR card with help of which we can know the demographic change within the city, city to city or state to state which will direct administration to plan a policy according to these changes, by linking total Aadhar issued at that pin code. Also, we can use this in joining

Gas cylinder connection data with the database will be effective way to channelise the available limited petroleum product.

References

1. Kuria, E., Kimani, S. and Mindila, A., 2019. A framework for web GIS development: a review. *International Journal of Computer Applications*, *178*(16), pp.6-10.
2. *"Census of India." (n.d.). Census Tables. Retrieved from* https://censusindia.gov.in/census.website/data/census-tables"
3. *"Esri. (n.d.). [Pincode Boundary of India 2022]. Retrieved from* https://www.arcgis.com/home/item.html?id=e1e9c110611b4e3383a3169901576544"
4. *Esri. (n.d.). Web GIS: Principles and Applications for GIS Web. Retrieved from* https://www.esri.com/content/dam/esrisites/en-us/esri-press/pdfs/gtk-web-gis-5e-sample.pdf