CensusScope: Exploring India's Census through Visualization

CS661 Course Project Proposal

Group 14

Indian Institute of Technology, Kanpur

Team Members

- Tanish Bansal (231071) – btanish23@iitk.ac.in – MTH, BS
- Aditya Anand (230065) – adityaa23@iitk.ac.in – MTH, BS
- Aurav Pratap Singh (230243) – auravps23@iitk.ac.in – MTH, BS
- Abdul Ahad Zareef (230030) – abdulahad23@iitk.ac.in – MTH, BS
- Ananya Pandey (230135) – ananyap23@iitk.ac.in – ME, BTech
- Vibha Narayan (231141) – vibhan23@iitk.ac.in – PHY, BS
- Ruchika Raj (230879) – ruchikar23@iitk.ac.in – EE, BTech
- Kartikey Singh Tomar (240529) – kartikey24@iitk.ac.in – ECO, BS

1. Introduction and Motivation

India's Census data is a vast source of demographic and socio-economic information: from population, literacy, sex ratio, and others, right down to state and district level. However, this vast amount of data is cumbersome to handle and interpret without the acumen or the tools. That's where our project comes in.

We desire to build an online web dashboard that is intuitive and interactive, transforming the 2011 Census data into meaningful visualizations. Users will be able to explore district-wise data across various metrics, compare performance across regions, and uncover patterns, trends, and anomalies with ease. The dashboard will also support drill-down functionality—allowing users to start at the state level and seamlessly dive deeper into the top- and bottom-performing districts based on the selected metric.

This tool is useful for anyone interested in exploring India's demographic landscape through an interactive and visually engaging platform.

2. Datasets and Sources

We will use:

- India Census 2011 - District-level data (uploaded dataset)
- India District GeoJSON – for mapping district and state boundaries (from Datameet or GADM)
- Supplementary demographic ratios (computed: sex ratio, literacy %, etc.)

3. Key Tasks and Objectives

- **Interactive Statewise Map:** An interactive choropleth map of India will display state-level metrics using color gradients. Clicking on a state will zoom in and load district-specific data for detailed analysis. .
- **Drill-Down to Districts:** Upon state selection, the dashboard displays the top-5 and bottom-5 districts competing on the chosen metric, exhibited through bar charts and ranking plots for comparison.
- **Metric Comparison and Filtering:** The platform will allow users to compare different socio-demographic indicators through dropdown menus. Additionally, filter controls such as sliders will let users define value ranges—enabling focused analysis of districts or states that meet specific criteria
- **Multidimensional Visualization (VTK):** 3D scatter plots built with VTK will show multivariate relationships (e.g., literacy vs. population vs. sex ratio) across districts..
- **Dynamic Dashboard Layout:** The dashboard will have a filter sidebar, central interactive map, and a interactive panel with responsive charts and insights.
- **Insight Cards and Narratives:** The dashboard will generate brief, readable insights—such as identifying districts with extreme values—to help users grasp key trends at a glance.

4. Planned Solution

Framework: Dash will serve as the main development framework, with Plotly being used to render the interactive 2D maps and charts. VTK will be integrated for advanced 3D visualizations to represent multidimensional data relationships.

Data Workflow:

- Clean and normalize Census and GeoJSON data to ensure consistency and accuracy.
- Merge and pre-process metrics (e.g., compute sex ratio) and also computer derived metrics.
- Standardize and link state and district names across all sources for smooth integration.

Web Interface:

- Statewise interactive choropleth to visualize metrics geographically.
- District-level filtering, ranking, and comparison for deeper exploration.
- Metric selection via dropdown and real-time updates to all relevant plots.

5. Team Responsibilities

Task	Members
Map & GeoJSON Integration & Dash App Backend Logic	Aurav, Tanish, Aditya
Dash App Backend Logic (Core Logic + State Filtering)	Aurav, Tanish, Aditya
UI Design	Abdul, Kartikey
Visualization (Charts, Interactivity, Layout)	Ruchika, Vibha, Ananya
Data Cleaning & Preprocessing	All Members
Documentation & Final Report	All Members
Testing & Deployment	All Members

6. References

- India Census 2011: `https://censusindia.gov.in/`
- Kaggle: `https://www.kaggle.com/datasets/danofer/india-census`
- Dash by Plotly: `https://dash.plotly.com/`
- VTK Python: `https://vtk.org/`
- GeoJSON Data: `https://github.com/datameet/maps`