

## Geospatial & Econometric Analysis of Household Data of Haryana

### Data Importing & Cleaning:

For this project, I have taken the State Wise Household 2011 Survey Data from the survey of India website.<sup>1</sup>

The original data set “**DDW\_PCA0000\_2011\_Indiastatedist**” in the repository contains data for the whole of India in the workbook titled: “India States”

MARG_OT_3_6_F								
G	H	I	J	K	L	M	N	
Level	Name	TRU	No_HH	TOT_P	TOT_M	TOT_F	P_06	M
India	India	Total	249501663	1210854977	623270258	587584719	164515253	
India	India	Rural	168612897	833748852	427781058	405967794	121322865	
India	India	Urban	80888766	377106125	195489200	181616925	43192388	
STATE	JAMMU & KASHMIR	Total	2119718	12541302	6640662	5900640	2018905	
STATE	JAMMU & KASHMIR	Rural	1553433	9108060	4774477	4333583	1593008	
STATE	JAMMU & KASHMIR	Urban	566285	3433242	1866185	1567057	425897	
DISTRICT	Kupwara	Total	113929	870354	474190	396164	188798	
DISTRICT	Kupwara	Rural	101930	765625	412038	353587	169588	
DISTRICT	Kupwara	Urban	11999	104729	62152	42577	19210	
DISTRICT	Badgam	Total	103363	753745	398041	355704	155202	
DISTRICT	Badgam	Rural	89417	655833	343385	312448	140608	
DISTRICT	Badgam	Urban	13946	97912	54656	43256	14594	
DISTRICT	Leh(Ladakh)	Total	21909	133487	78971	54516	12016	
DISTRICT	Leh(Ladakh)	Rural	14905	87816	48411	39405	9053	
DISTRICT	Leh(Ladakh)	Urban	7004	45671	30560	15111	2963	
DISTRICT	Kargil	Total	18338	140802	77785	63017	19928	
DISTRICT	Kargil	Rural	16147	124464	67703	56761	18359	
DISTRICT	Kargil	Urban	2191	16338	10082	6256	1569	
DISTRICT	Punch	Total	90261	476835	251899	224936	84674	

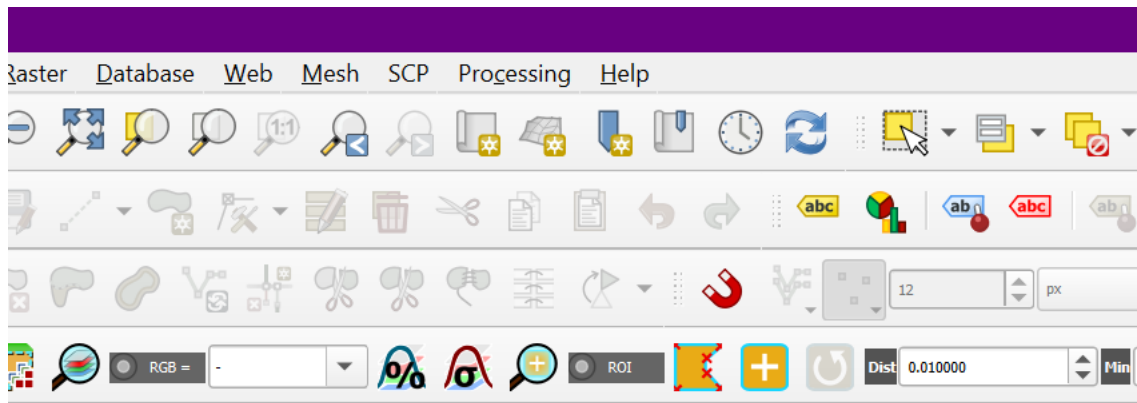
The data for our target state has been extracted, cleaned and transformed to make it ready for geo-spatial analysis. New indexes have been calculated which can be seen in the workbook titled “**computed\_indx**”.

### Digitization:

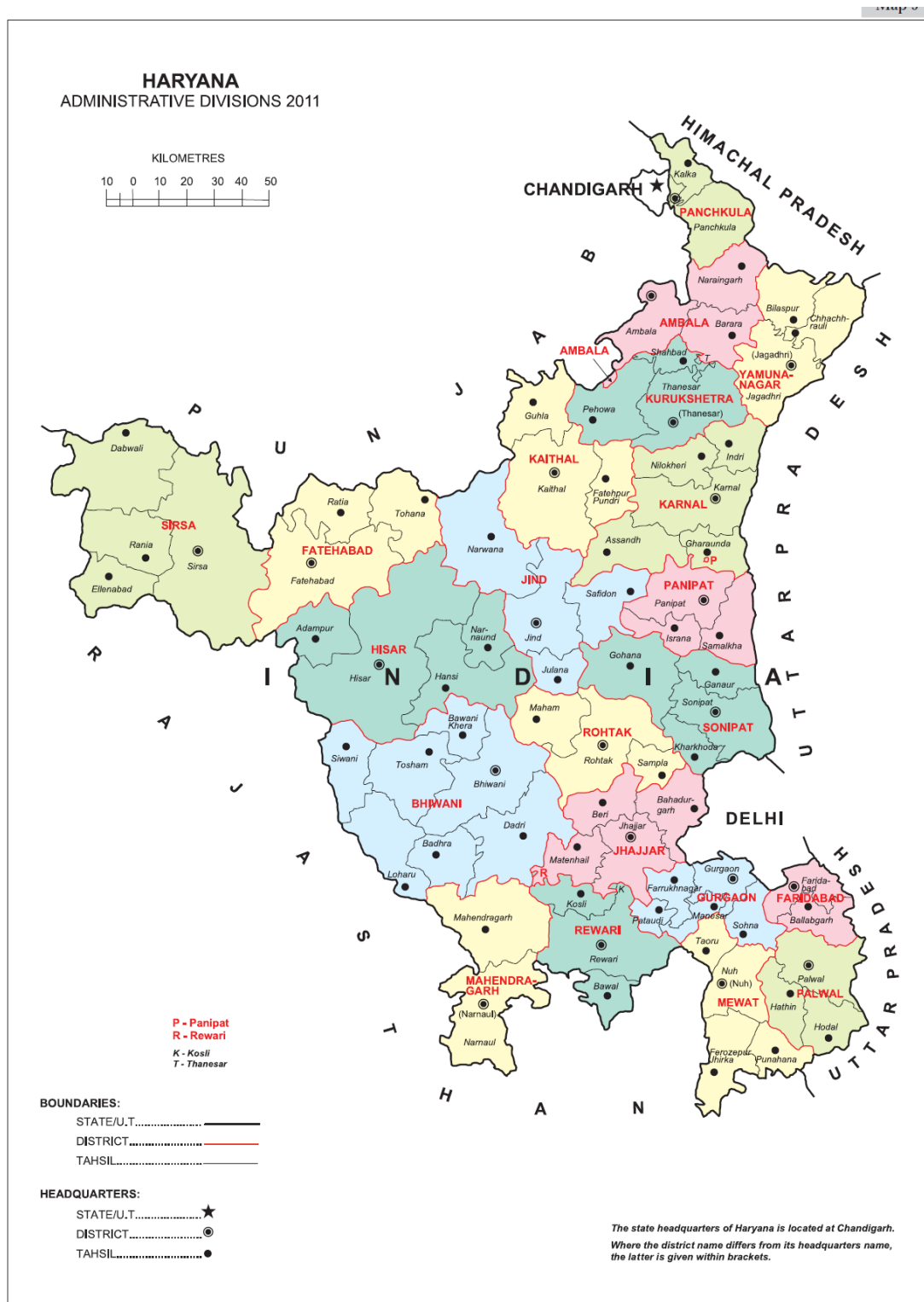
qGis which stands for Quantum Geographic Information System, is an open-source cross-platform desktop geographic information system (GIS) application. It allows users to create, edit, visualize, analyze, and publish geospatial information.

For digitisation, I have taken a shape file of India States “**India\_States.shp**” in the repository.

<sup>1</sup> <https://censusindia.gov.in/census.website/data/census-tables>



And overlaid it with a tiff file (“**Haryana.tiff**”) of Haryana that has district wise boundaries:



To digitise the polygons of the states have been manually traced to make a digital map.

### Data Transformation:

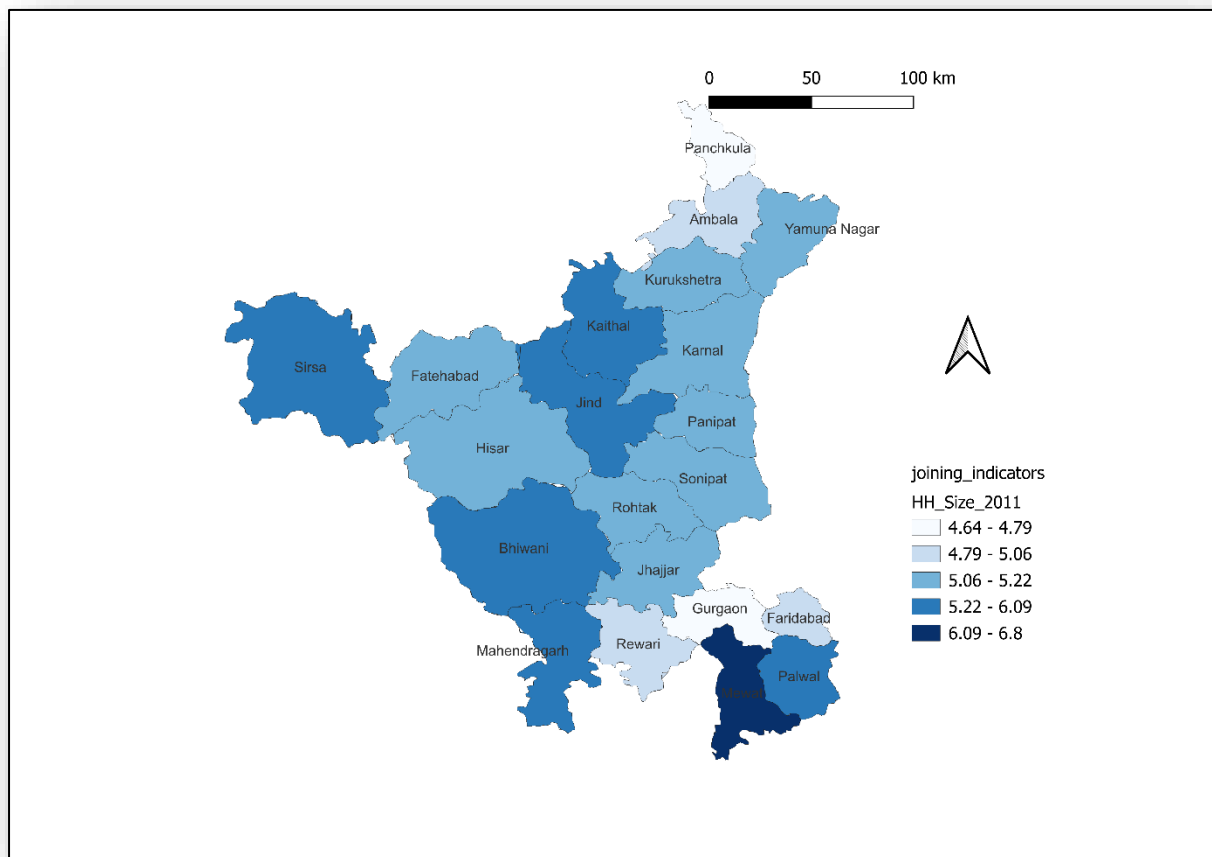
The computed indexes have been converted into a csv file and joined with the shape file to make our shape file ready for geospatial analysis.

### District Geospatial Analysis:

I have used qGis to make maps that show the district wise distribution of various household factors such as:

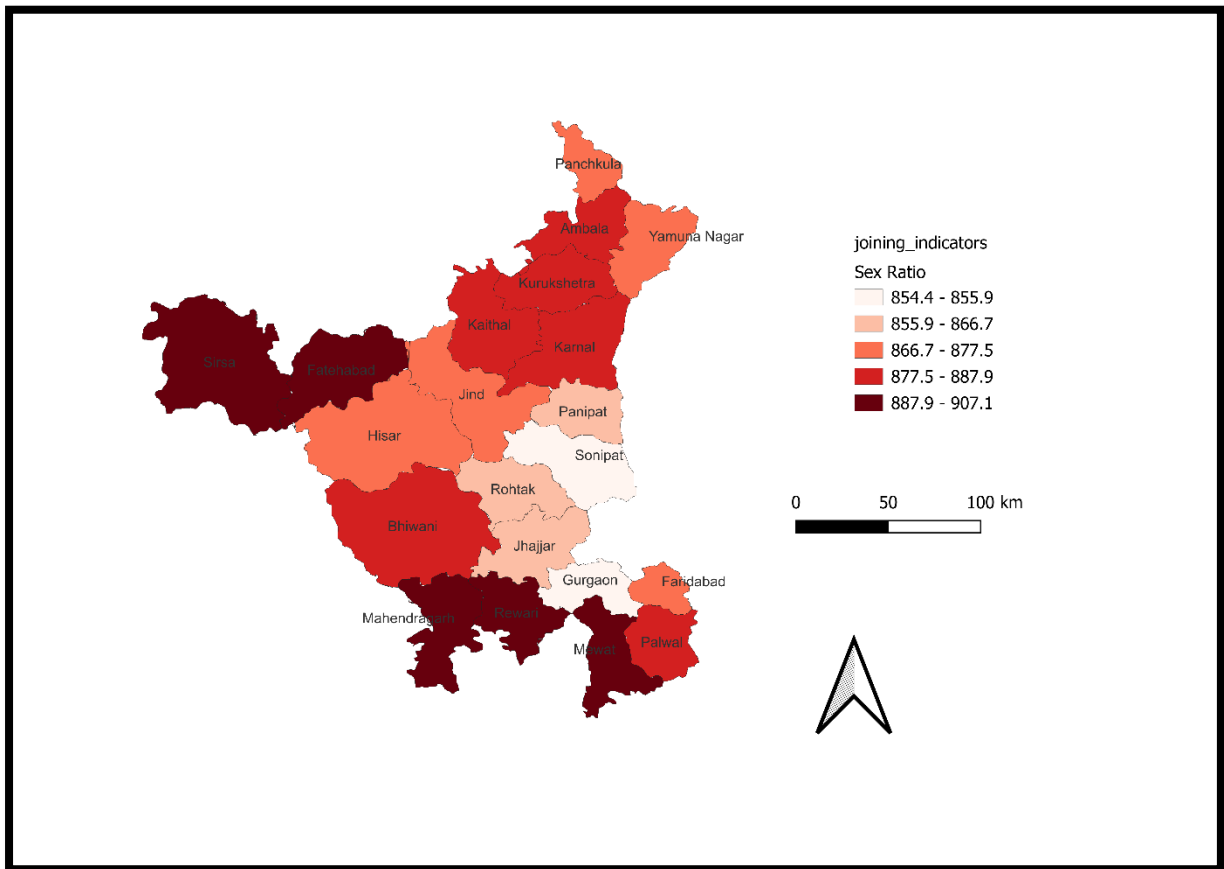
**Figure 1: Household Size:**

Larger household sizes are indicative of agrarian households, as expected more industrialized and affluent districts such as **Gurgaon, Faridabad, Ambala and Panchkula** have lower household sizes whereas districts such as **Sirsa, Bhiwani and Mewat** have relatively higher household sizes.



Source: Based on data from Census of India, 2011

Figure: 2: Sex Ratio

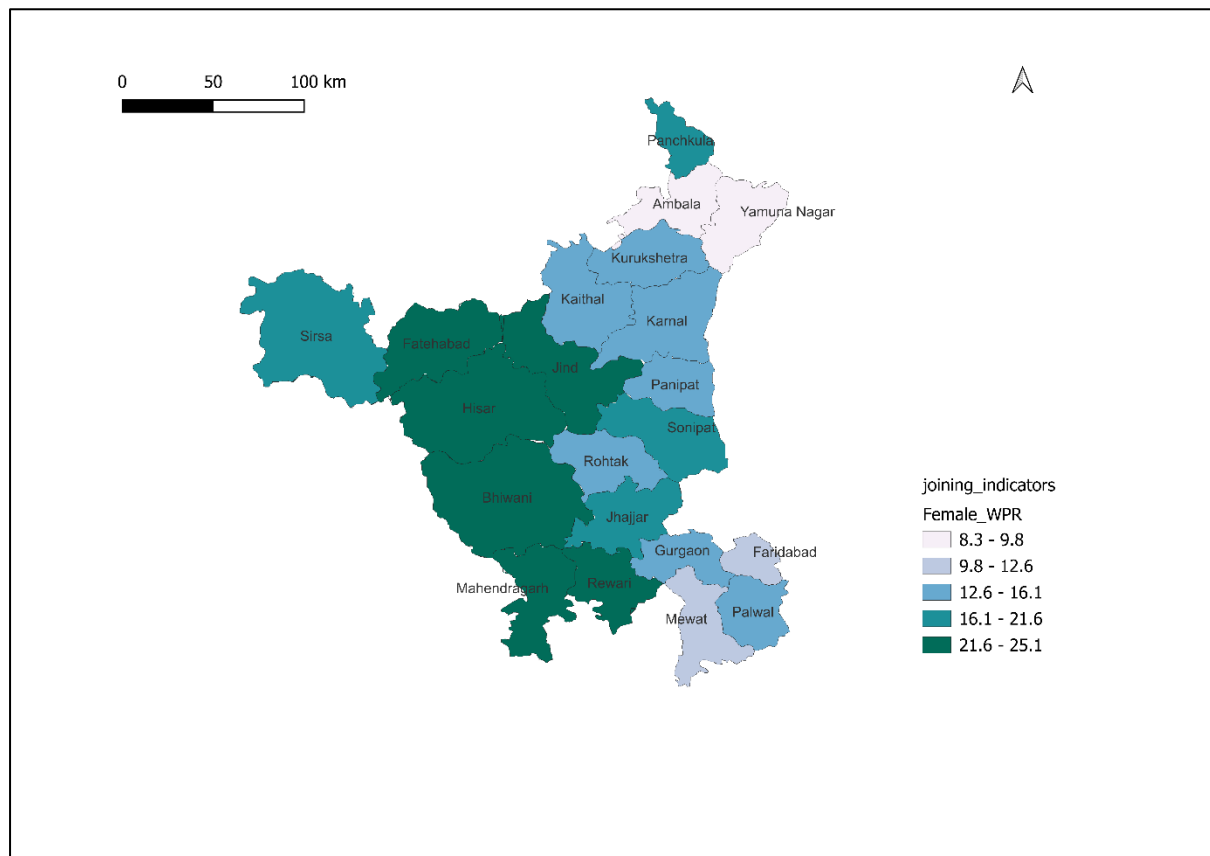


Source: Based on data from Census of India, 2011

As we can see from the map, districts of **Sonapat** and **Gurgaon** are the worst performing in the state, this may be because of a confluence of factors such as sex-selective abortions, female foeticide, societal preference for male offspring, often rooted in cultural or traditional beliefs, economic factors such as dependence on male heirs, and the persistence of a dowry system also play roles. Lack of education for women, inadequate healthcare access, gender discrimination, son preference, and traditional gender roles further contribute to the imbalance.

Thus, more efforts need to be poured into these districts it will Addressing a necessitate comprehensive efforts, including legal measures, educational initiatives, economic empowerment, and cultural awareness campaigns tailored to promote gender equality and discourage discriminatory practices against female children.

**Figure 3: Female Work Participation Rate:**



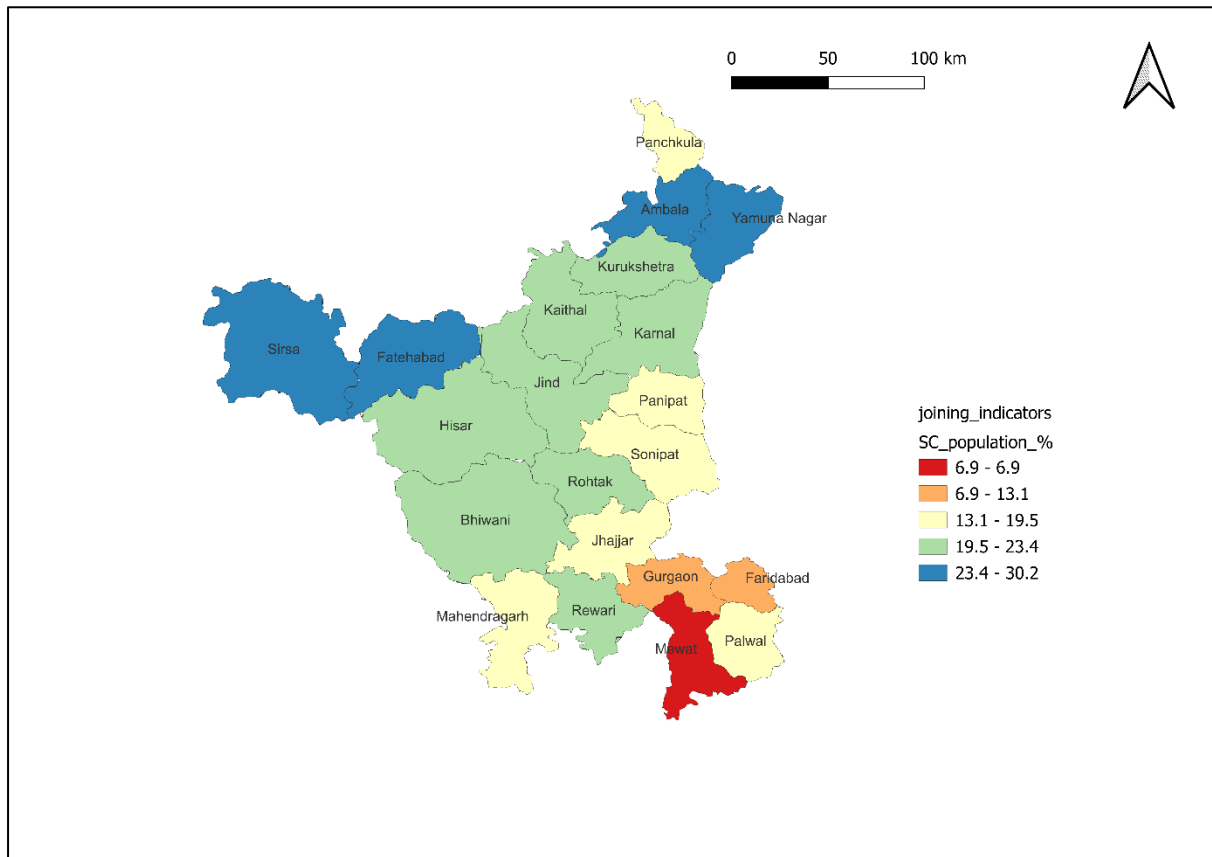
**Source: Based on data from Census of India, 2011**

Mewat, Yamuna Nagar, Faridabad and Ambala have the worst female work participation rate in the state. Most concerning is the last two as both have a significant amount of factories and textiles respectively which can be used as an engine of growth as an increase in work participation is associated with positive externalities with reproductive health as well<sup>2</sup>.

Low female work participation rates can be attributed to a range of factors, including societal norms and gender roles that limit women's career choices, lack of access to education and skill development opportunities, inadequate family support, and prevailing stereotypes about women's capabilities. Unequal distribution of domestic responsibilities, limited availability of flexible work arrangements, and safety concerns also contribute to the underrepresentation of women in the workforce. Addressing these challenges requires targeted interventions, such as promoting education, changing societal perceptions, implementing supportive workplace policies, and creating an enabling environment for women to pursue careers.

<sup>2</sup> Finlay, J. E. (2021). Women's reproductive health and economic activity: A narrative review. *World Development*, 139, 105313.

**Figure 4: Scheduled Caste Population:**



**Source: Based on data from Census of India,2011**

The scheduled caste population is mostly concentrated in 4 districts:

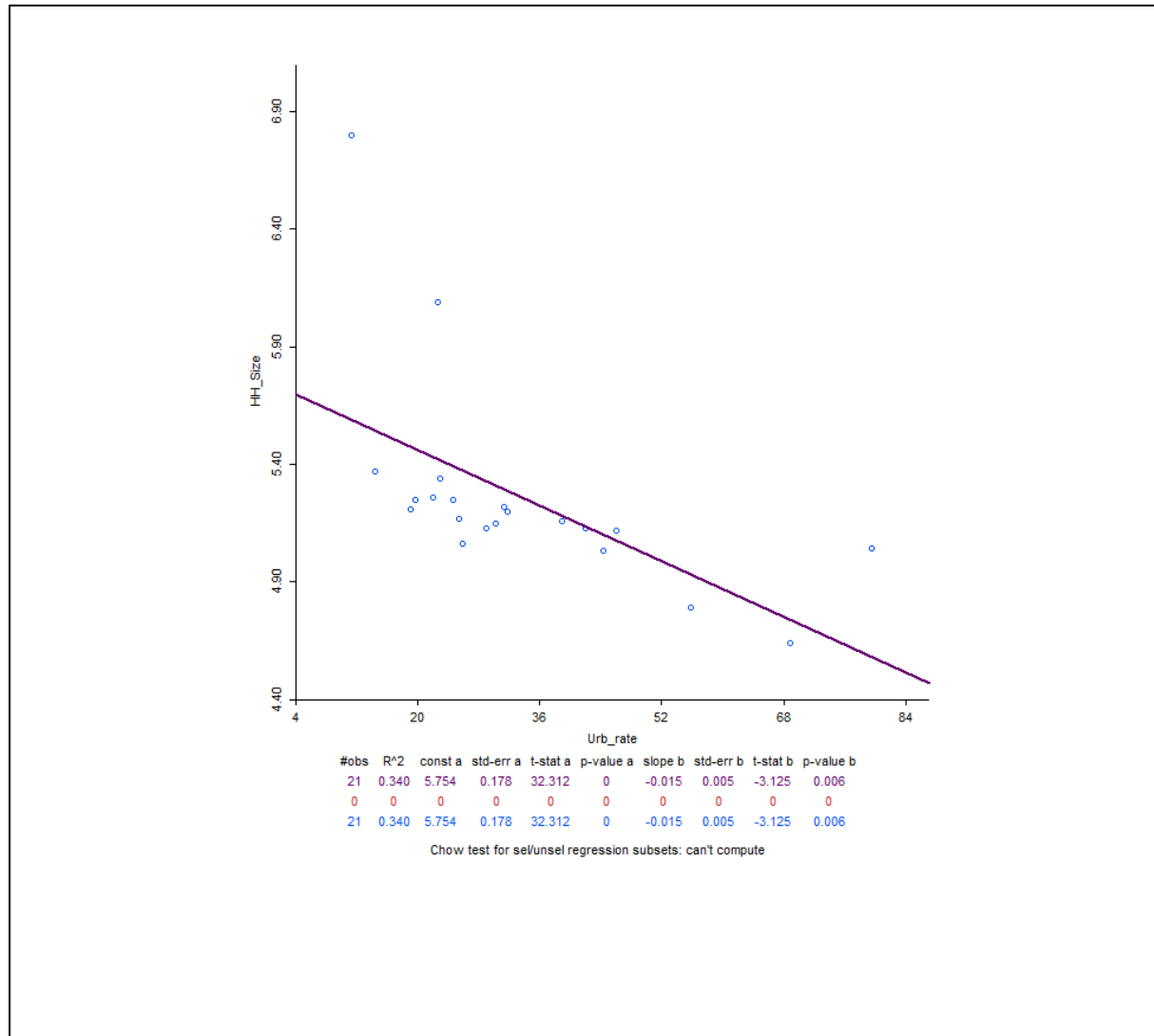
1. Sirsa
2. Fatehabad
3. Ambala
4. Yamuna Nagar

Efforts regarding upliftment of marginalized communities should be focussed on this area.

## Econometric Analysis using GeoDa:

Although qGIS is able to make high quality digitised vector layer maps, it's capabilities are limited when it comes to deploying econometric tests. For that purpose I have used GeoDa.

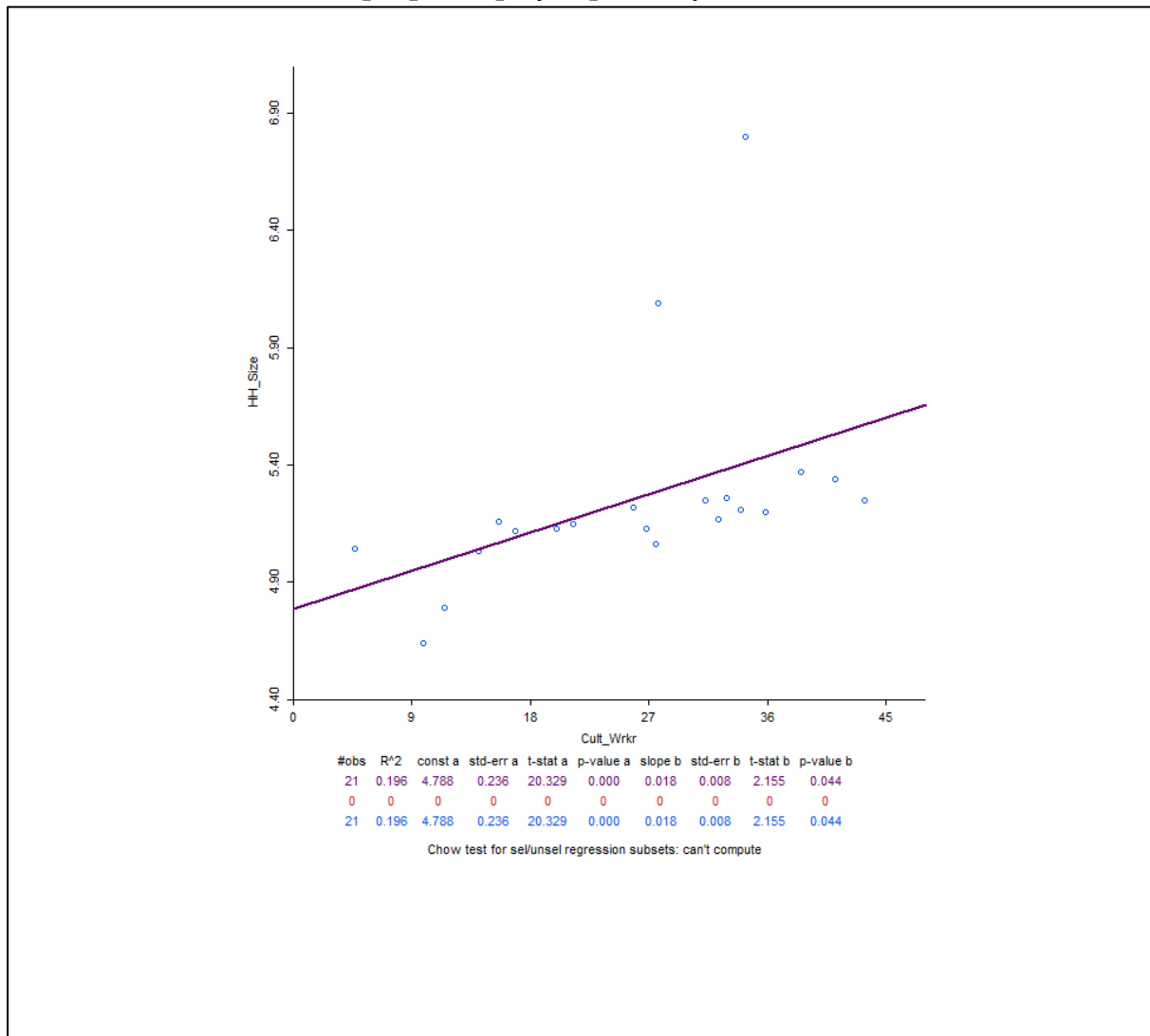
### 1: Scatter Plot of Urbanisation Rate and Household Size



The downward sloping line on the scatter plot indicates an inverse relationship between urbanization rate and household size. As the urbanization rate increases, household size tends to decrease.



## 2: Scatter Plot of number of people employed primarily as cultivators to household size:



As expected, areas with higher ratio of workers involved in cultivation activities are associated with a larger household size.