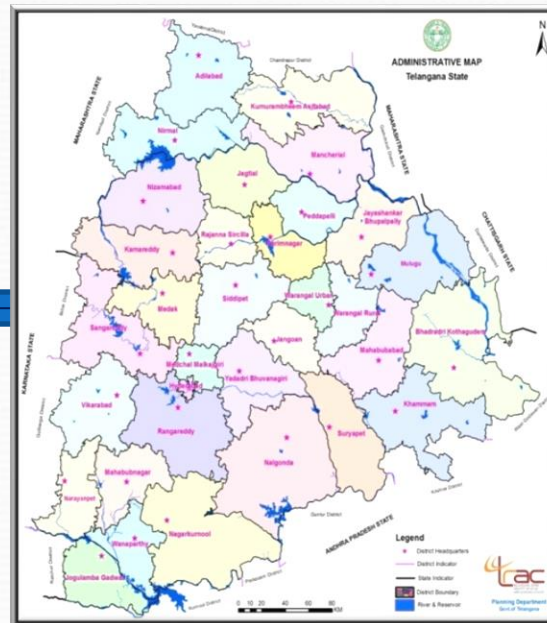




# 'Samagra Kutumba Survey'

A case Study of Big Data Analytics from Telangana State



March, 2019







# Big Data

- extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions.
- *Big data* is an evolving term that describes a large volume of structured, semi-structured and unstructured data that has the potential to be mined for information and used in machine learning projects and other advanced analytics applications.

# Relevance of 6Vs of big data with

## The six Vs of big data

Big data is a collection of data from various sources, often characterized by what's become known as the 3Vs: *volume, variety and velocity*. Over time, other Vs have been added to descriptions of big data:

VOLUME	VARIETY	VELOCITY	VERACITY	VALUE	VARIABILITY
The amount of data from myriad sources.	The types of data: structured, semi-structured, unstructured.	The speed at which big data is generated.	The degree to which big data can be trusted.	The business value of the data collected.	The ways in which the big data can be used and formatted.
					

1 crore households  
3.68 crore persons

Structured format  
Separate qn for R/U  
8 section  
94 parameters

ONE SINGLE DAY

Matched with Growth Rate.  
Data collected from each household.  
Voluntary collection so people are expected to provide correct data.

Most recent data when compared to Census, so all departments are dependant on SKS

It is being maintained by NIC  
A web portal has developed with dashboards  
A server is allocated  
A mobile App is developed and being used by officials





**Welcome**  
**to**  
**Intensive Household Survey (HIS)**  
***Samagra Kutumba Survey***

**Planning Department**  
**Telangana State**







# Intensive Household Survey (IHS) – 2014 (*Samagra Kutumba Survey-SKS*)

## An Overview



ఆగస్టు 19 2014

ఇంటింటి సర్వే

సమగ్ర సామాజిక,  
ఆర్థిక సర్వేకు సహకరిద్దాం!  
మనమంతా కలలు కంటున్న  
బంగారు తెలంగాణా  
సాకారం చేసుకుందాం!!

# Concept

- Enumeration of Socio Economic status of all the Households in Telangana State formed on 02.06.2014.
- Survey conducted on one single day (19.08.2014) to ensure objectivity and prevent duplication.
- Entire Government machinery at the field level was deployed.
- Simple but effective format for survey.
- Compared to other surveys taken up so far, unique and most comprehensive exercise.

# Objectives

- To create a reliable database on each household of Telangana State.
- To facilitate usage by different Government Depts /agencies to implement programmes.
- Effective targeting of the Welfare Programmes.
- To ensure that the intended benefits reach the deserving poor.
- To plug the leakages in the implementation of programmes.

*Every paisa must reach the deserving poor*



# Process

- Nearly 4 lakh government employees including police utilized to collect details of the households
- Information collected on voluntary basis to avoid litigation (WPs pending on privacy issues)
- Not notified under the Collection of Statistics Act, 2008 (mandatory disclosure)
- Collected the actual information on 8 broad areas covering 94 items
- Verification of documents issued by the State / Central governments such as:
  - Aadhaar card/Ration card
  - Bank Account Details
  - LPG connection booklet
  - Physically Handicapped certificate (if applicable)
  - Electricity Bill
  - Pattedar Passbook
  - Caste certificate
  - Vehicle Registration details



# Process...

- Natives of Telangana, residing in other states and abroad who were physically not present included in the list, based on relevant documents produced.
- Training imparted to all the enumerators and supervisors for the effective conduct of Intensive Household Survey.
- Wide publicity given for participating of the people in the survey.

ఆంధ్ర 19 2014  
ఇంటింటి సర్వే

సమగ్ర సామాజిక, ఆర్థిక సర్వేకు సహకరిద్దాం!  
మనమంతా కలలు కంటున్న బంగారు తెలంగాణా సాకారం చేసుకుందాం!!

**KEEP THESE DETAILS READY**

LPG CONNECTION	CASTE CERTIFICATE	INCOME TAX PAYEE OR NOT	BANK BRANCH AND ACCOUNT NUMBER
PENSION DETAILS	AADHAR CARD	IF PHYSICALLY CHALLENGED, CERTIFICATE ID NO.	
LONG TERM ILLNESSES, IF ANY	PROPERTY OWNERSHIP AND STATUS OF RESIDENCE	ELECTRICITY CONNECTION AND METER NUMBER	
OWNERSHIP OF AGRICULTURAL LANDS	OWNERSHIP OF VEHICLES CAR/ MOTORCYCLE/JEEP	IF HAILING FROM OTHER STATES, LANGUAGE SPOKEN AND YEAR OF COMING HERE	
WHETHER AIR-CONDITIONER IS INSTALLED			

# Process...

- Enumeration conducted from 8 AM to 8 PM in rural areas, 7 AM to 7 PM in urban areas, the time was also extended to cover all the households.
- Holiday declared under Negotiable Instrument Act.
- Took a minimum of 15 minutes to canvass schedule for each household.





# Census, NSSO Vs IHS 2014

Census	NSSO Survey	IHS 2014
Information collected on every citizen	Survey collects information from a sample households on selected subject for the round	Information collected on all the members of the Households.
Conducted every 10 years	Surveys conducted theme wise on a 5-year cycle	As decided by the Government
Statutory status under the Census Act, 1948	Results are being estimated on sample data collected	Voluntary disclosure of data
Preparatory process for house listing over 1 year - enumeration over 20 days	Process covers 6 months – one year	Though conducted in census mode, IHS completed in ONE DAY
Only consolidated data of households is available on public domain	Only sample survey	Individual data available for use by Government Depts., (Not on public domain due to litigation)



# Census, NSSO Vs IHS 2014...

Census	NSSO Survey	IHS 2014
Provides demographic details on households, amenities, workers, economic status and residence, literacy, schooling, married population, age at marriage etc., at village, district, state wise.	Focuses on socio-economic, demographic, agricultural and industrial subjects for collecting data from house holds and from enterprises located in villages and in the towns	More comprehensive as it covers 94 items under 8 different parameters
Socioeconomic Caste Census (SECC) as part of Census 2011 provides abstract caste wise details only for SC, STs.	The NSSO's mandate is to conduct surveys on socioeconomic issues (more subjective)	Provides sub-caste wise details of SC, ST, BCs as well as minorities

# Census 2011 Vs IHS 2014 on Key Parameters

Parameter	Census 2011	IHS 2014
Households	83,03,612	1,03,95,629
Total Population	3,50,03,674*	3,68,76,544
Male	1,76,11,633	1,84,11,741
Female	1,73,92,041	1,84,05,885
Rural	2,13,95,009	2,42,42,966
Urban	1,36,08,665	1,26,33,578
SC	54,08,800	64,44,584
ST	31,77,940	36,44,453

\* Annual growth of 1.35% as per Decadal Growth Rate

# Achievements of IHS

- Survey was conducted on a single day i.e., on 19.08.2014 throughout State.
- Entire Government machinery i.e., 3,85,892 employees were involved.
- Simple but effective format was used
- All the enumerators were effectively trained before the survey
- Compared to other surveys taken up so far, unique and most comprehensive exercise.
- About 1.05 crore survey formats were printed centrally



ఆంధ్ర 19 2014  
ఇంటింటి సర్వే



సమగ్ర కుటుంబ సర్వే - 2014

26 ఆగస్టు

(మీరు నింపవలసినది)

GHMC

సమగ్ర కుటుంబ సర్వే



సమగ్ర కుటుంబ సర్వే  
పూర్తి అయింది

తెక్కతేలింది  
1,05,76,922 కుటుంబాల్లో 100% పూర్తి



19 2014  
ఇంటింటి సర్వే

సమగ్ర లింగాంకనం, జాతీయ సర్వేను సహజంగా ముగింపుగా కడయ కంటున్న వంశారు తెలంగాణ లాభారం చేసుకుందాం!!



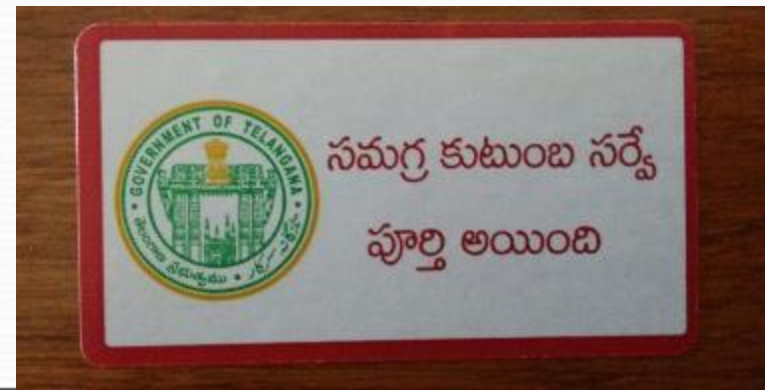


# Coverage

- Total Number of Districts covered: 10 now 33
- Total Number of Mandals Covered: 464 now 589
- Total Number of Municipalities Covered: 57 now 136
- Total Number of Municipal Corporations Covered : 6
- Total Households surveyed: 1.03 crore
- Total Enumerators deployed: 385,892
- Total number of computers and Data Entry Operators deployed: 25,000

# Survey Format

- The format contained information on 8 broad areas covering 98 items
- Modification for urban areas as per local need
- Each booklet consisted 30 HH survey formats.
- Total HH in GP divided by 30 to arrive at total booklets needed.
- ‘Door Stickers’ affixed after filling the format, to ensure comprehensive coverage.



# 8 Broad Areas & (parameters)

- Identification particulars (13)
- Family particulars (21)
- Housing details (16)
- Family member details (13)
- Persons with disability details (6)
- Chronic diseases (3)
- Movable assets of the household (5)
- Household land details (15)
- Live stock and pets details (6)













# Samagra Kutumba Survey, 2014

## Part-H: Household land details (including of all lands in all places)

1. Does the household hold agricultural land? (1- Yes 2- Yes but land is on the name of mother/father/ grandfather /brother, 3- Land less )
2. If family holds land, give details

S. No	Name of the person holding land	Patta /Account No. (Khatha No.)	Extent						How land is acquired*	If land is in different place then Extent						If Dry/Fallow land then reason *		
			Wet Land			Dry Land				Wet Land			Dry Land					
			Acres	Gu.	Type of irrigation *	Cultivable		Fallow land*		Acres	Gu.	Type of irrigation*	Cultivable		Fallow land*		Acres	Gu.
						Acres	Gu.						Acres	Gu.				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15				

## Part I: Livestock and Pets details: 1. Do the household have livestock/pets? 1-Y, 2-N) 2. If yes then give details:

1) S. No	2) Type of live stock/pets	3) Number	S. No	Type of live stock/pets	Number
1	Buffalo/Cow/Ox		4	Poultry birds (including the birds in poultry farm )	
2	Sheep/goat		5	Pet Dogs	
3	Pigs		6	Others (Specify)	

### Declaration

I certify that the above information is true. If the above information is found to be wrong I may be declared ineligible for Government benefits/schemes and I will be liable for any legal action. I solemnly declare the above in the name of God/Self.

<b>Head of the family/members signature/left thumb impression*</b>  Name: <input style="width: 100%;" type="text"/>	<b>Enumerator Signature</b> Name: <input style="width: 100%;" type="text"/> Designation: <input style="width: 100%;" type="text"/> Mobile No: <input style="width: 100%;" type="text"/>	<b>Supervisor signature</b> Name: <input style="width: 100%;" type="text"/> Designation: <input style="width: 100%;" type="text"/> Mobile No: <input style="width: 100%;" type="text"/>
---	--	--

# Software Management

- National Informatics Centre :
  - Developed the software for survey
  - Database development
  - Processing the data
  - Provision of server space
  - Maintenance of the data
  - Maintaining helpline for software problems
  - Developed Mobile App which is being used by Government officials

# Monitoring of Survey

- Empowered District Collectors to:
  - Monitor the overall implementation of household survey.
  - Designated Mandal Special Officers for monitoring
  - Designated four Zonal officers (Mandal divided into four Zones) to monitor the zone allotted
  - Designated GP Special Officers to monitor the survey in GPs
  - To ensure high quality data collection and data entry



# Expenditure incurred

- Rs. 2.0 crores sanctioned to each of the 10 erstwhile districts for logistical arrangements initially
- No honoraria paid to enumerators
- Printing of formats through SERP – Rs.6.0 crores
- Data entry at district level through local vendors selected through tender system
- Infrastructure available in Govt. / Pvt. Colleges used
- Total expenditure – Rs.33.94 crores
- Annual Maintenance by NIC for managing database - Rs.14.0 lakhs
-

# Immediate benefits of the Survey

- Use to identify eligible beneficiaries for welfare programmes, such as:
  - Food Security Cards
  - Rythu Bheema
  - Aasara Pensions
  - Financial assistance to Beedi workers
  - Scholarships
  - Selection of beneficiaries for 2BHK Housing
  - Individual Sanitary Latrines
  - Database used by ST and BC Commissions of State
  - Integrated People Information Hub IPIH - Citizen 360
  - Database for district reorganization exercise
  - Assess impact of demonetization by using bank/PO account details
  - Providing Community based schemes such as sheep distribution / fisheries / artisans / single women pensions etc.



# Intensive Household Survey - 2014

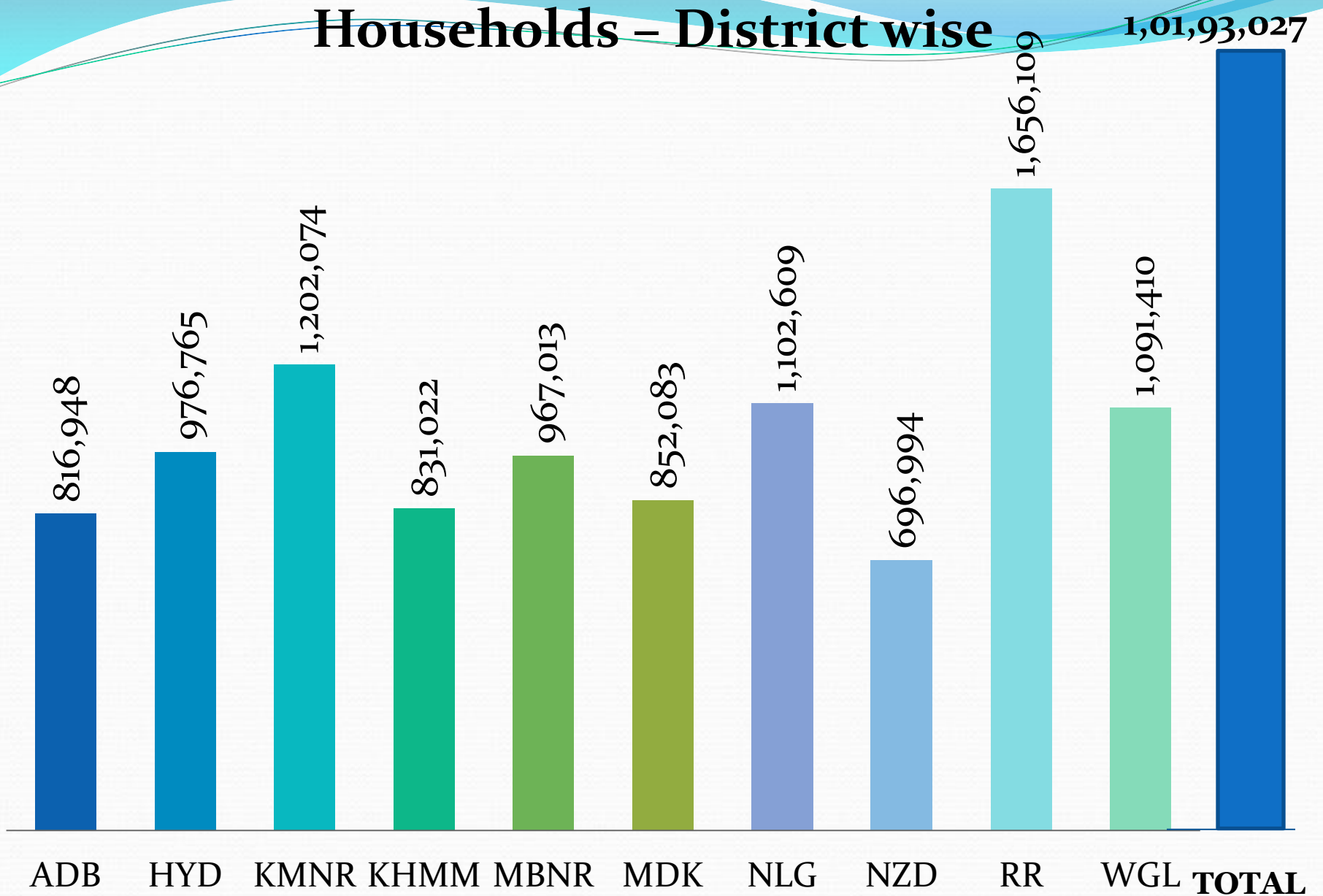
## Key Findings



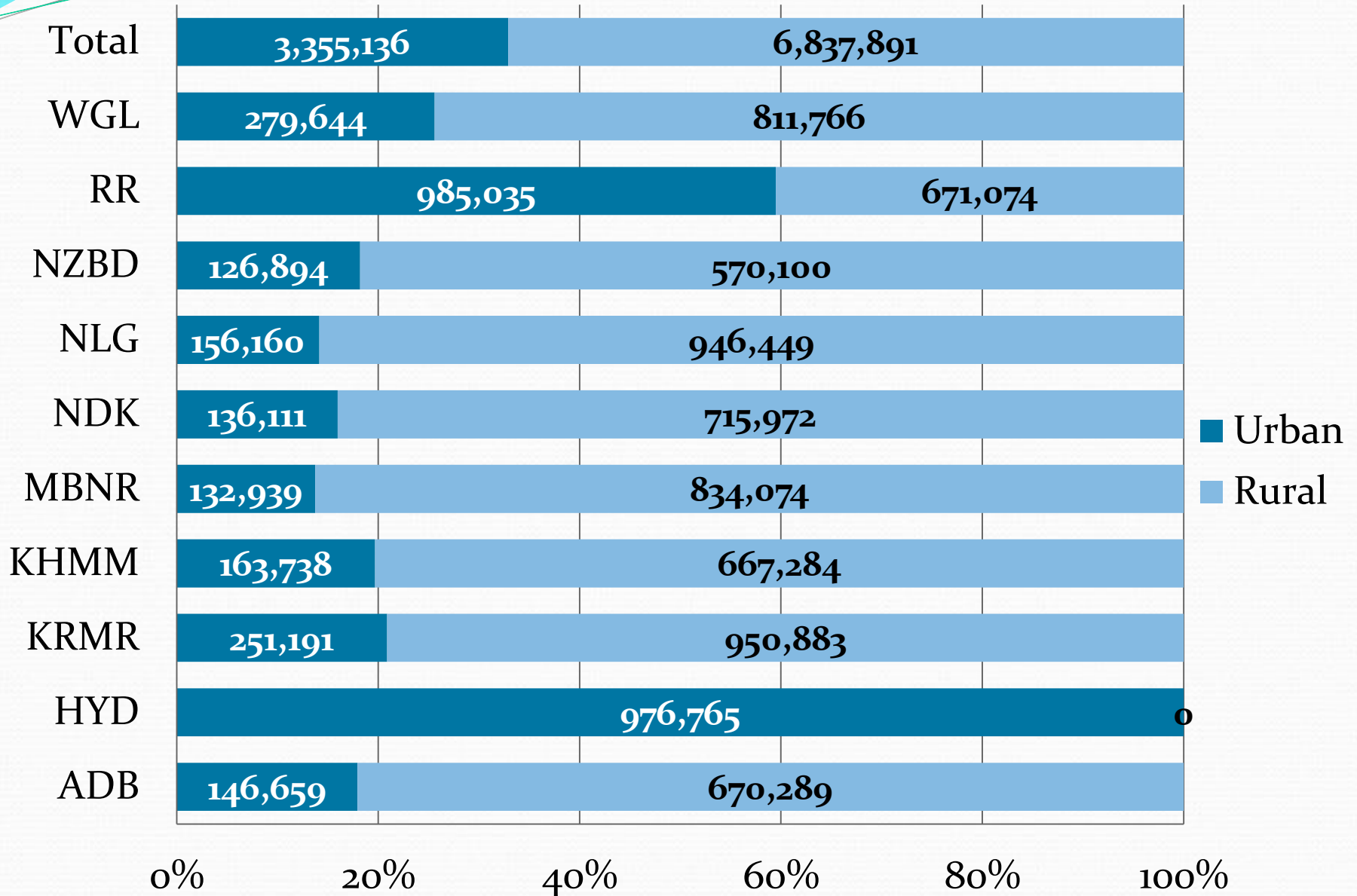
# IHS – District wise Details

<b>District Name</b>	<b>Total Households</b>	<b>Total Population</b>	<b>Total Eligible Households for Government scheme</b>	<b>Total Eligible Population for Government Schemes</b>
ADB	8,16,948	28,24,953	4,90,169	16,94,972
HYD	9,76,765	37,94,218	5,86,059	22,76,531
KRMR	12,02,074	38,38,323	7,21,244	23,02,994
KHMM	8,31,022	26,23,072	4,98,613	15,73,843
MBNR	9,67,013	42,84,024	5,80,208	25,70,414
MDK	8,52,083	30,92,584	5,11,250	18,55,550
NLG	11,02,609	35,95,203	6,61,565	21,57,122
NZBD	6,96,994	24,67,312	4,18,196	14,80,387
R R	16,56,109	61,36,368	9,93,665	36,81,821
WGL	10,91,410	36,46,955	6,54,846	21,88,173

# Households - District wise



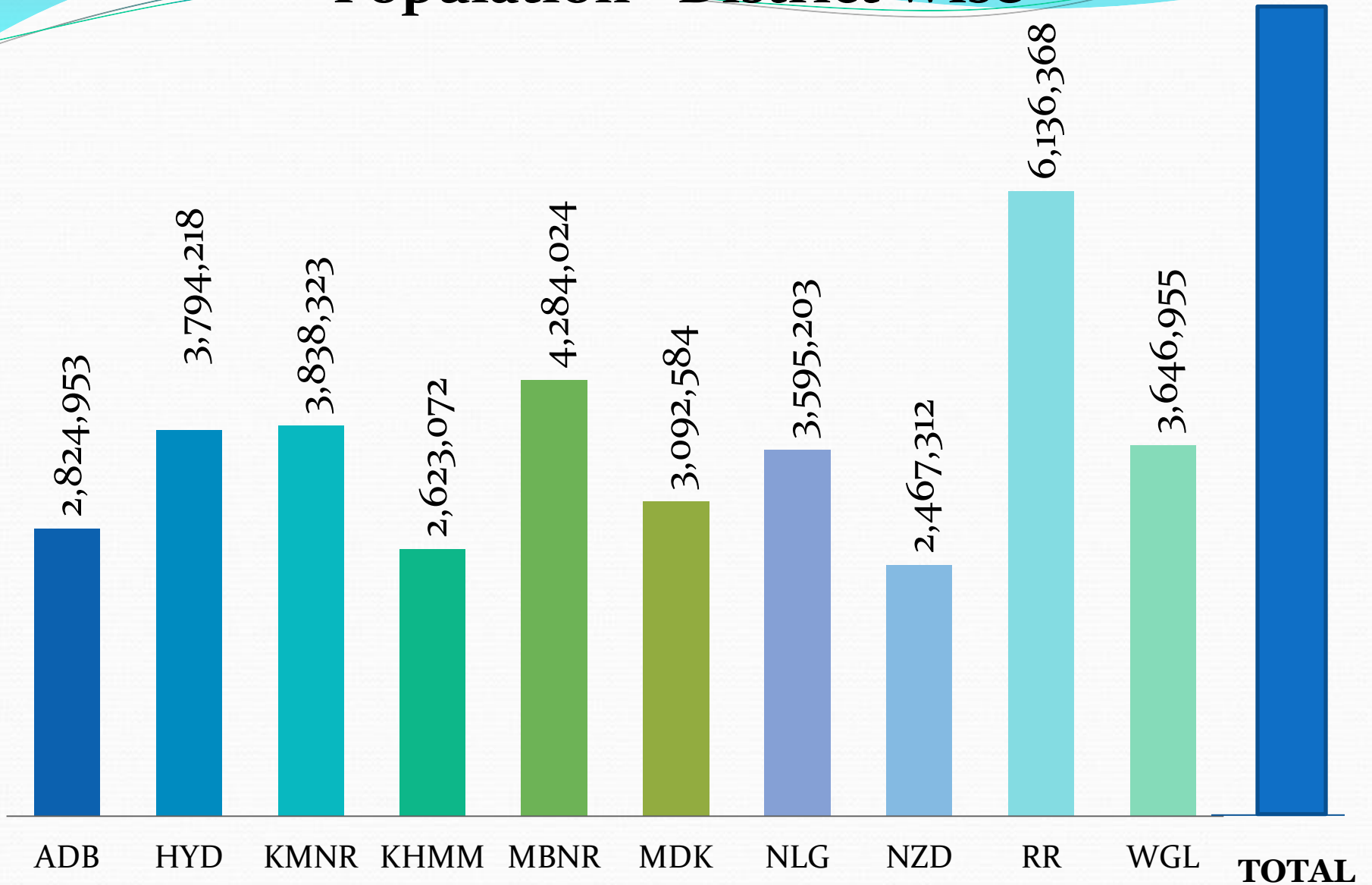
# Household - Rural & Urban



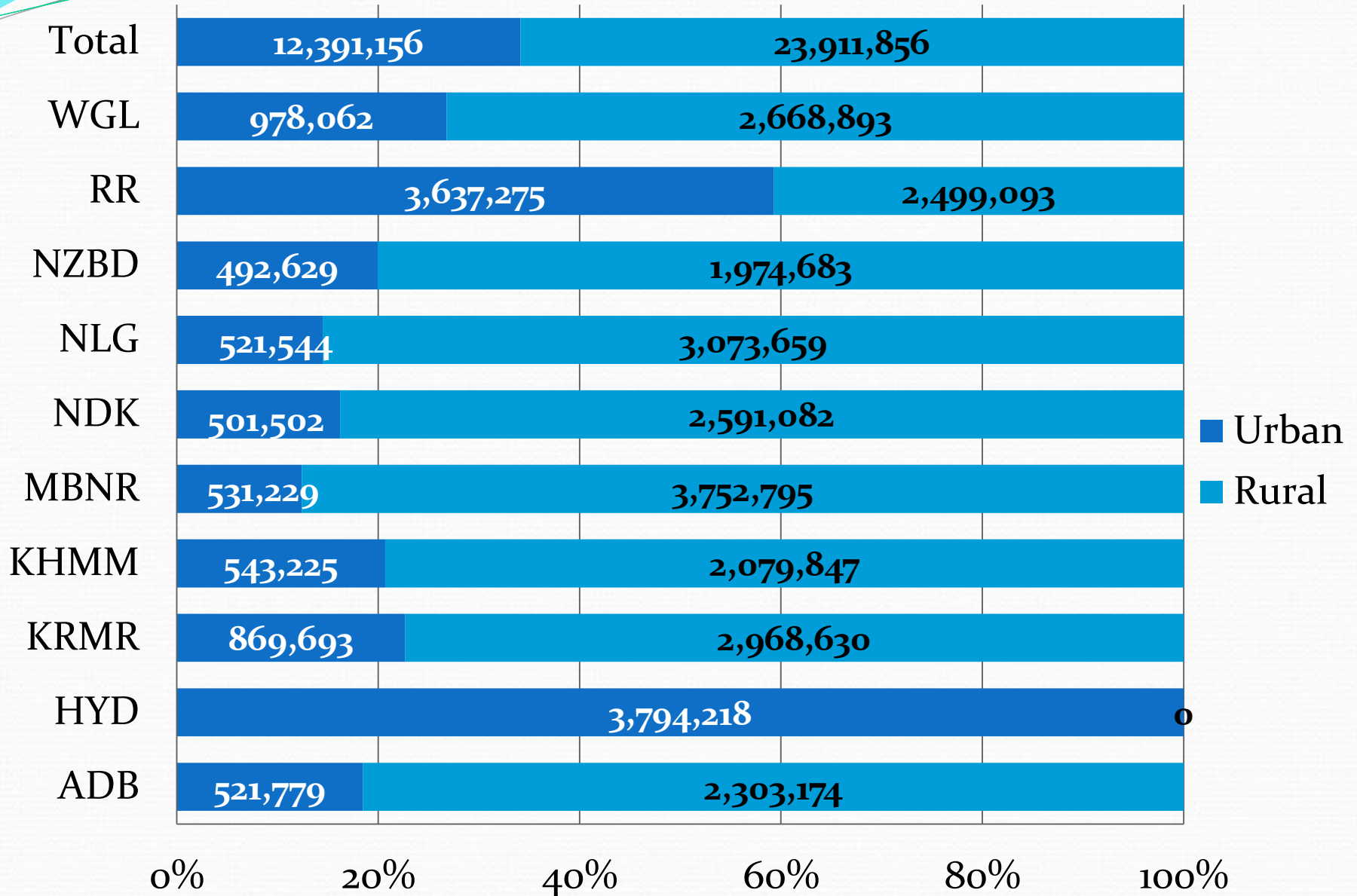


# Population - District Wise

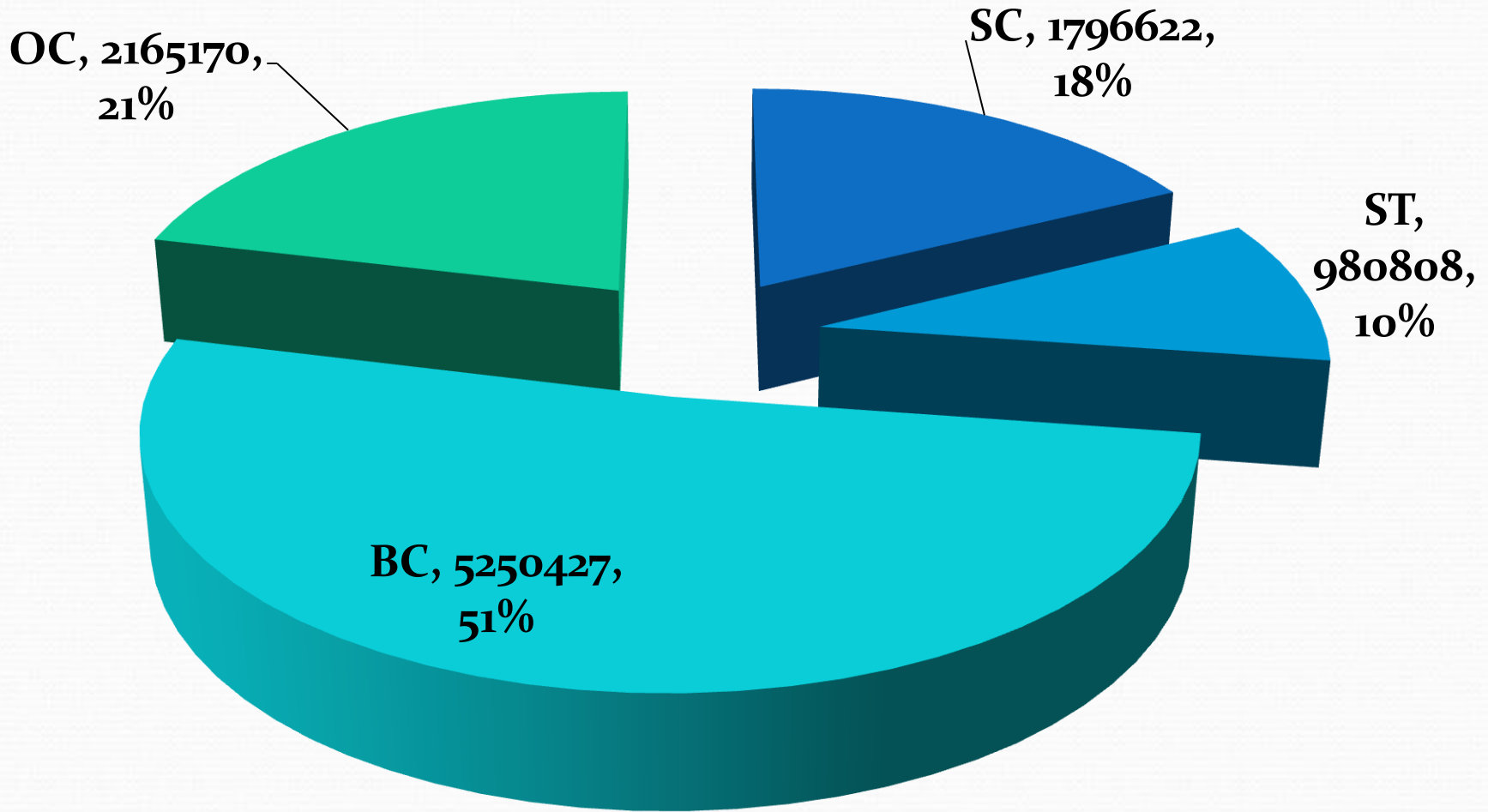
3,63,03,012



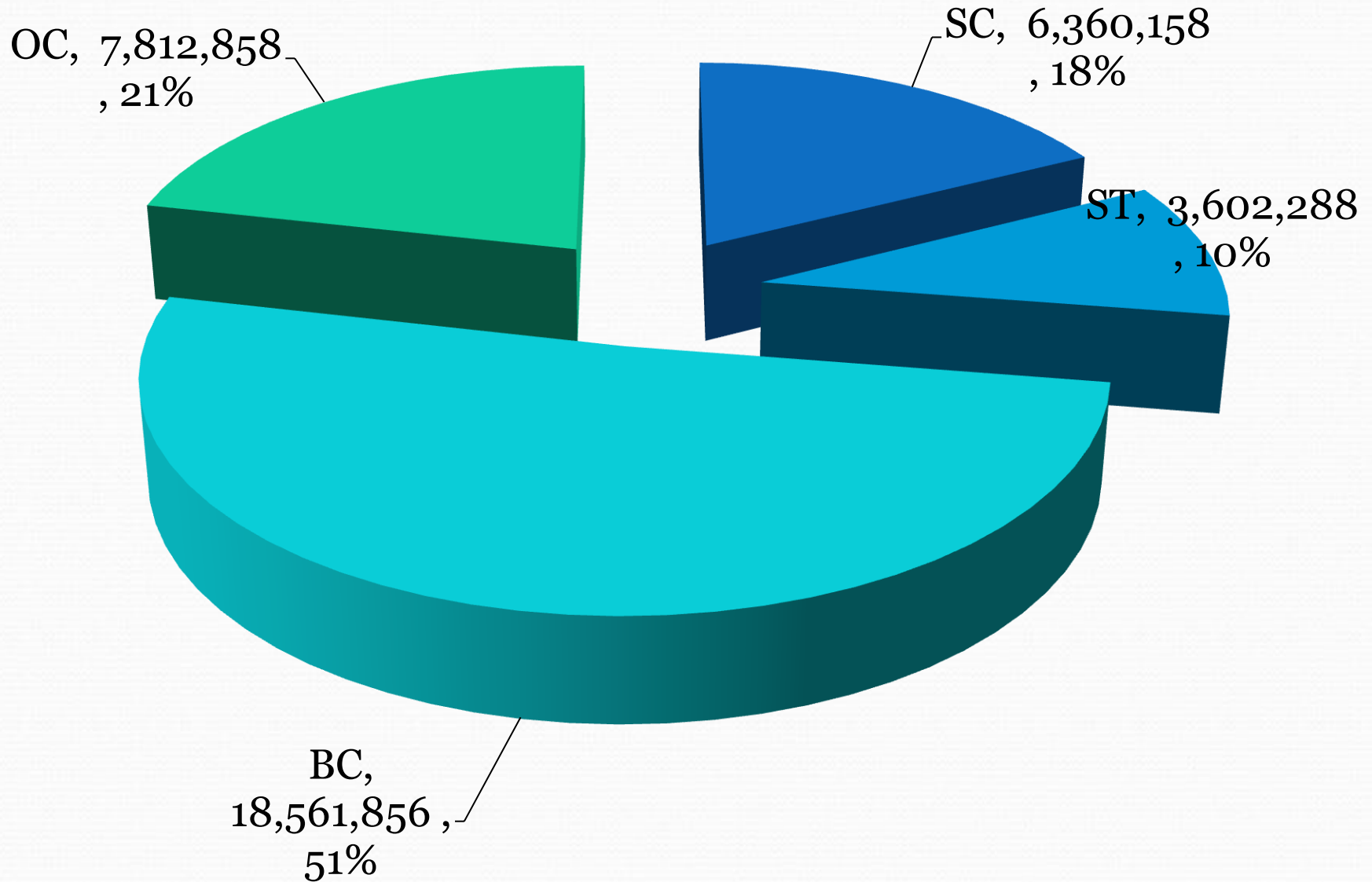
# Population - Rural & Urban



# Households – Caste Wise

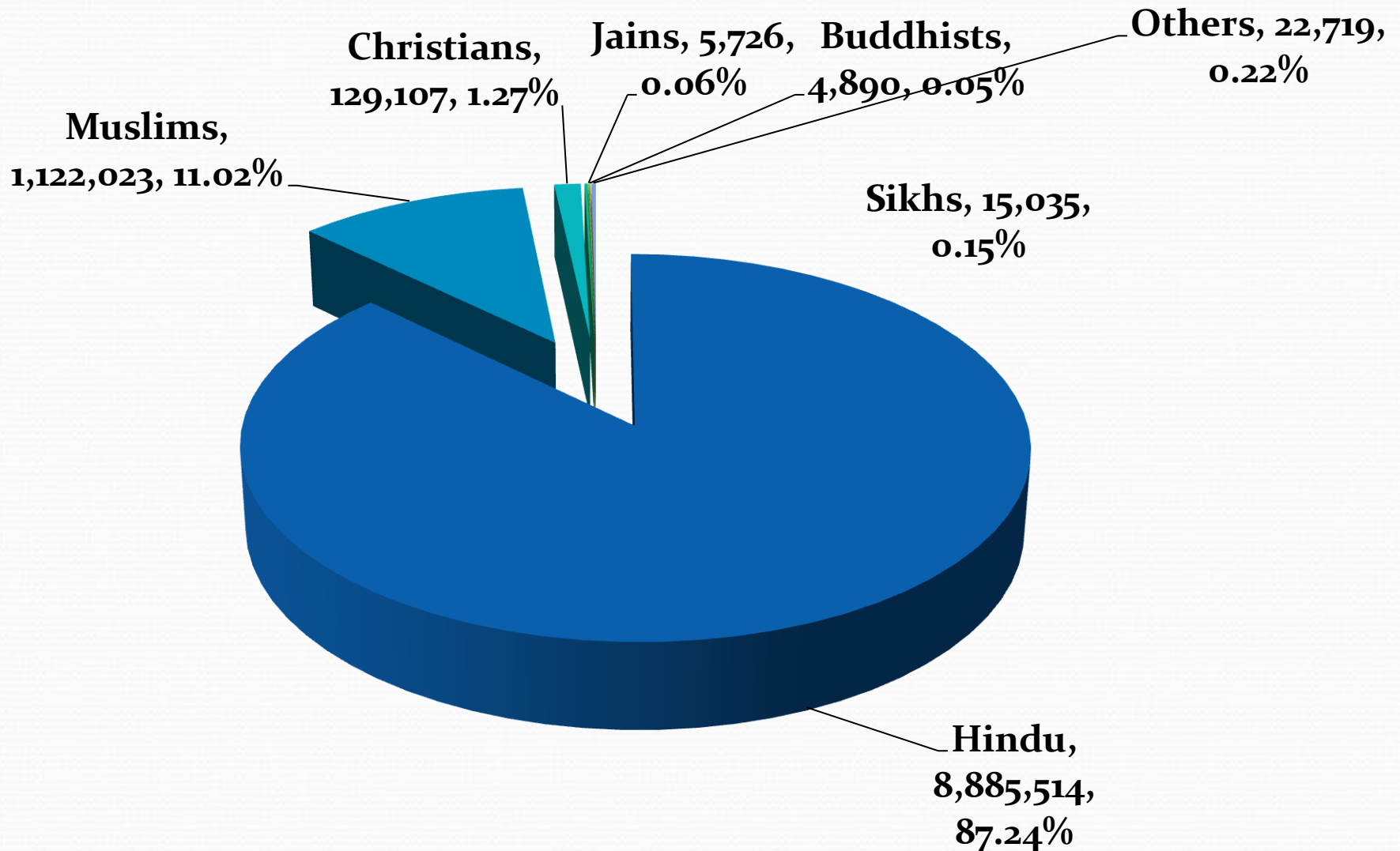


# Caste Wise - Population

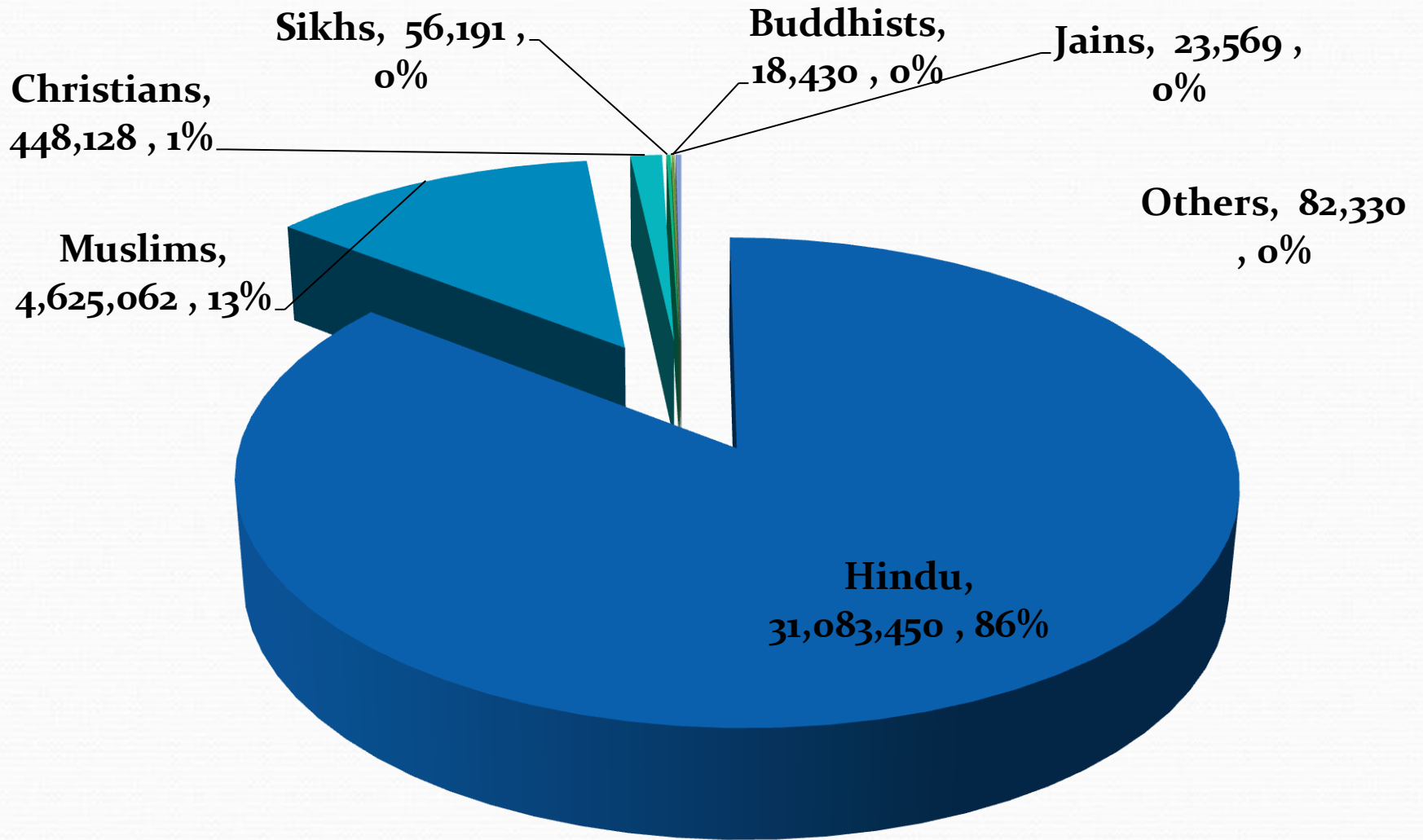




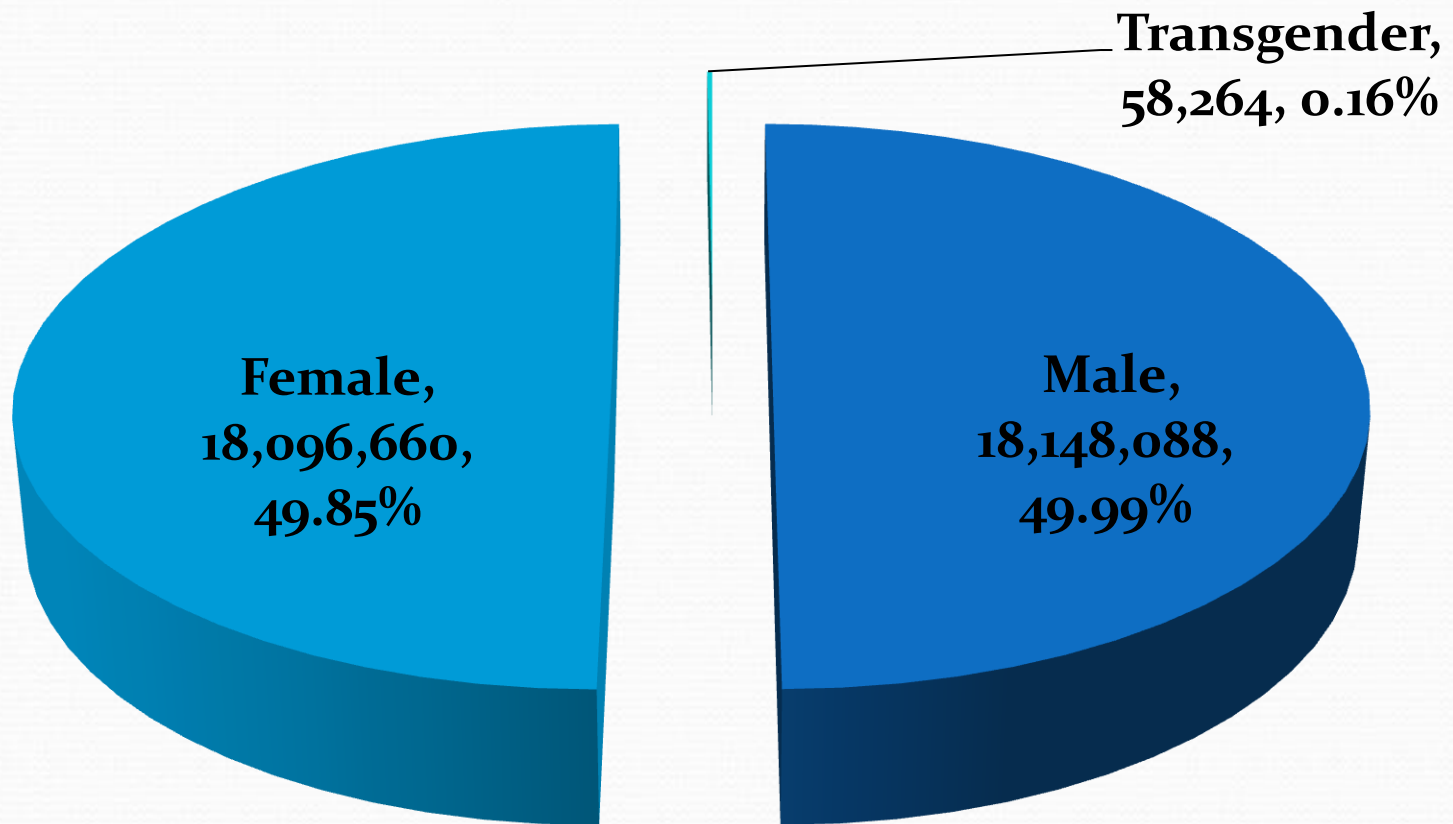
# Household - Religion wise



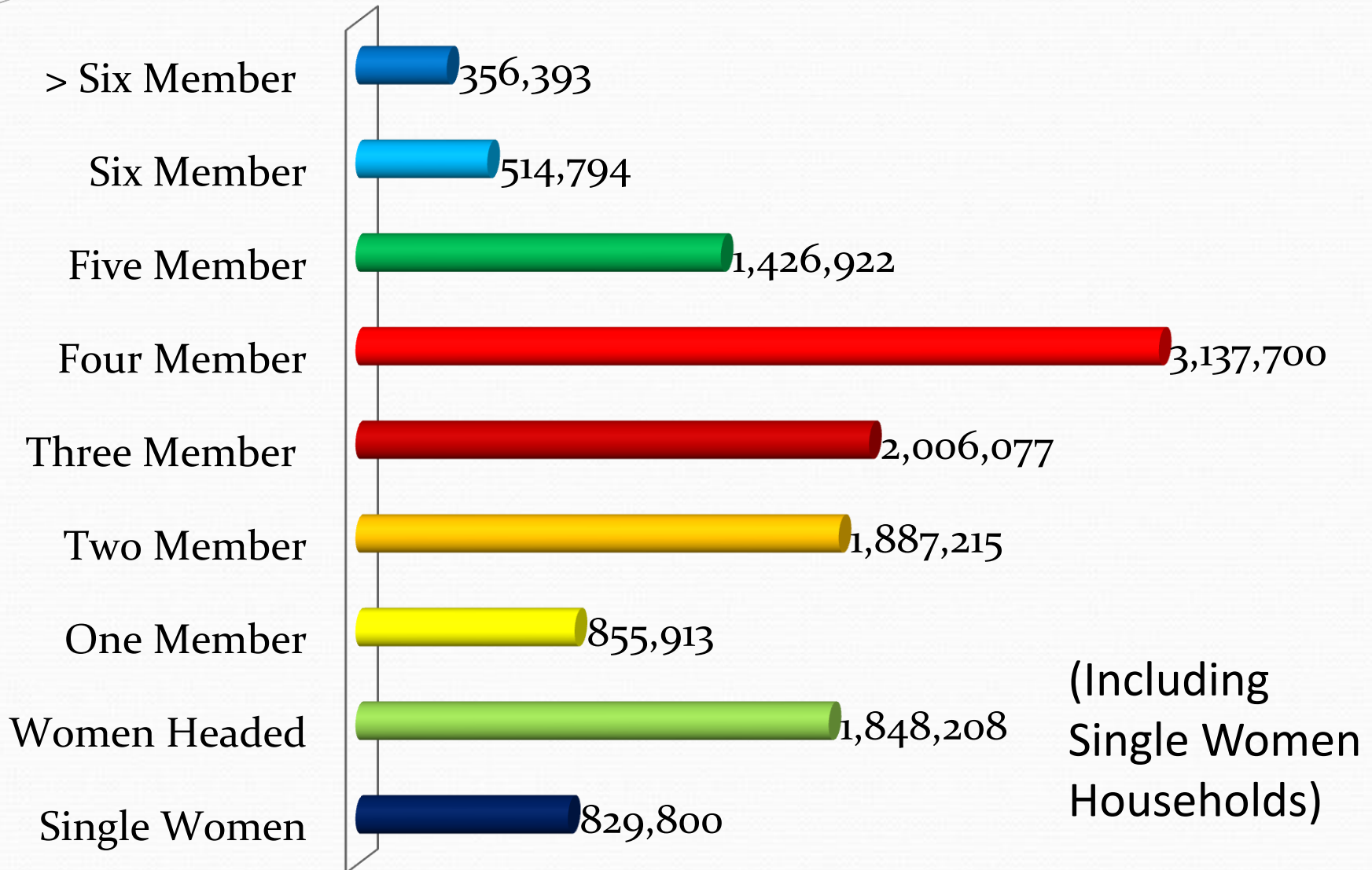
# Religion - Population



# Gender wise Population

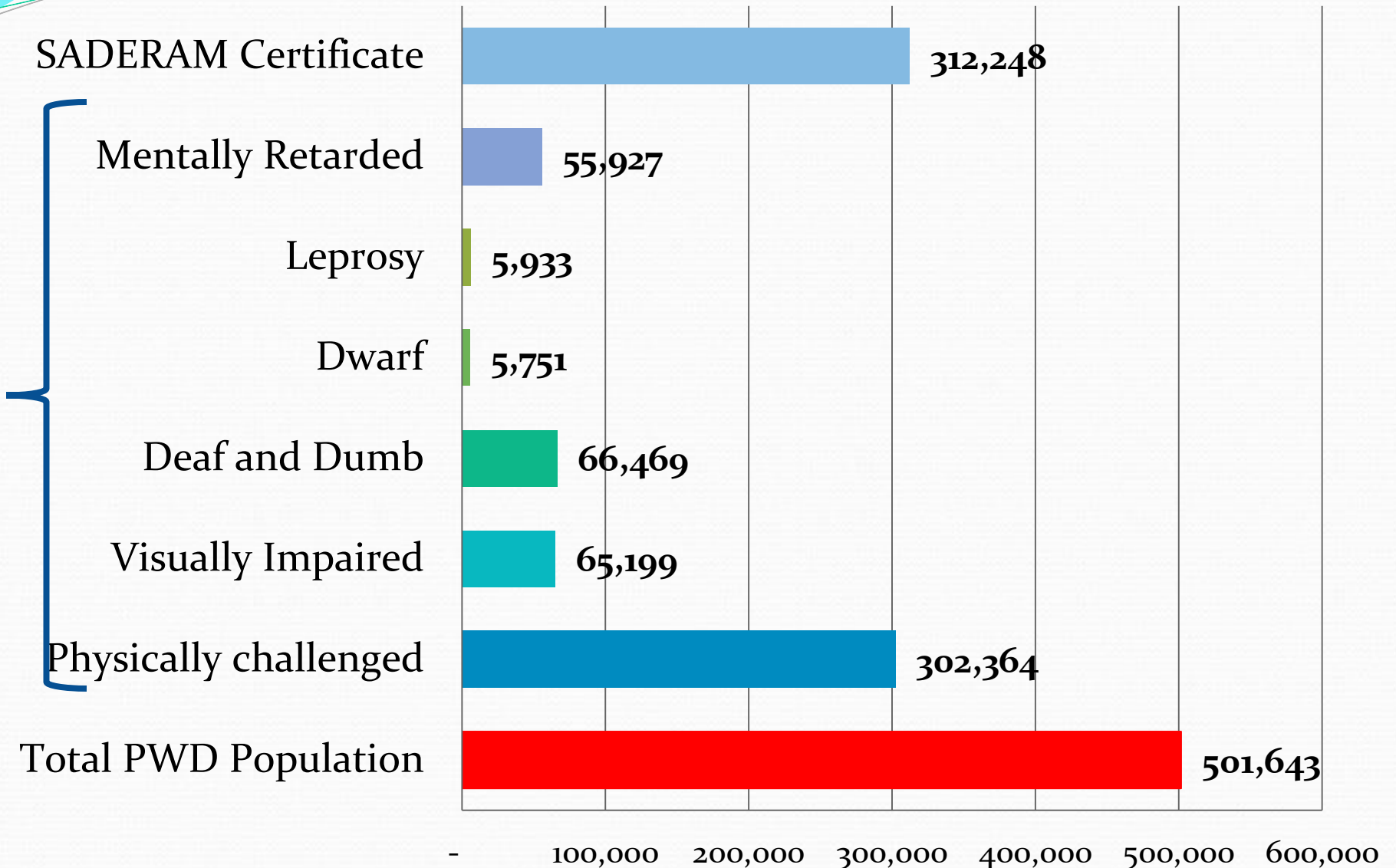


# Member wise - Household Count

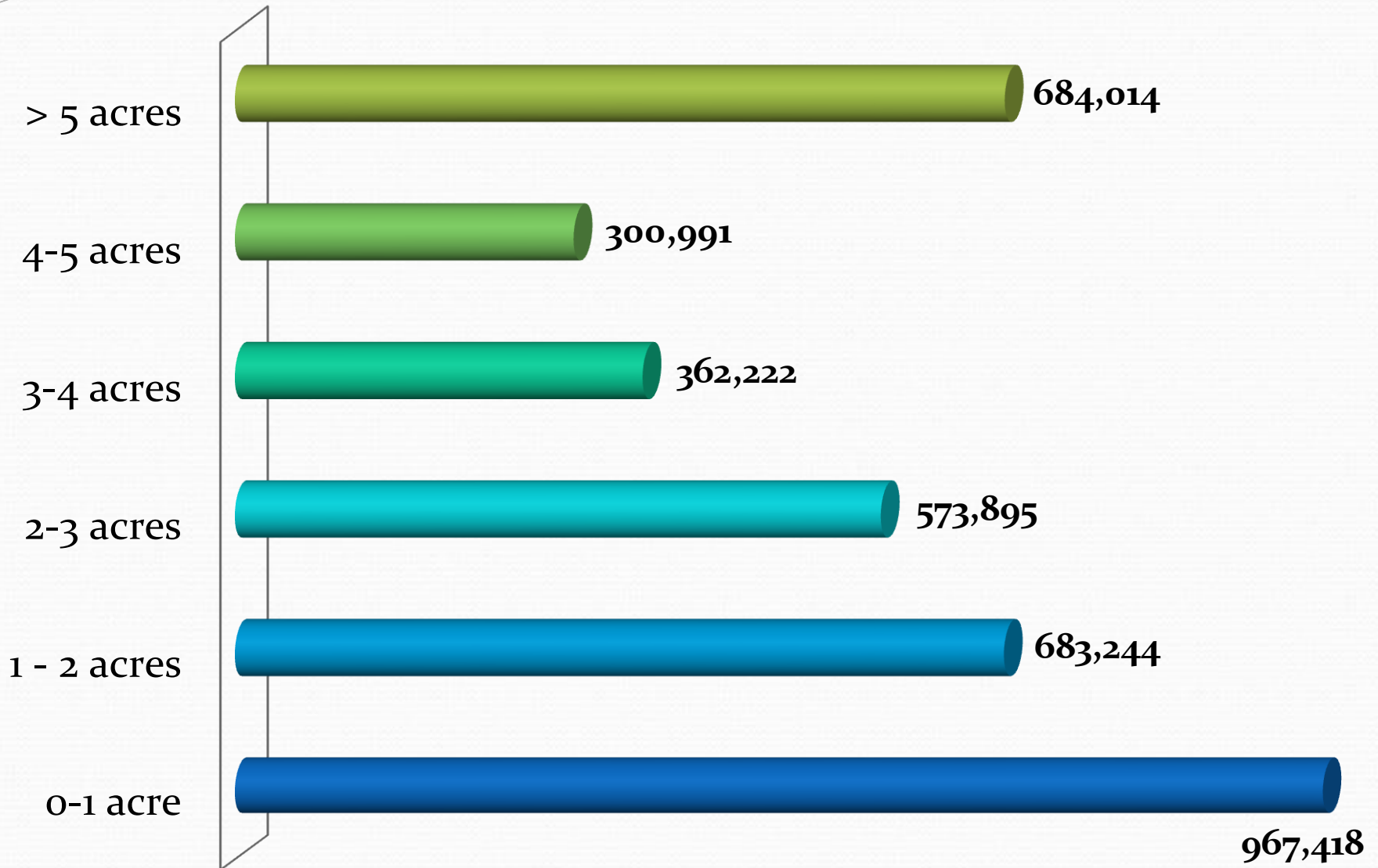




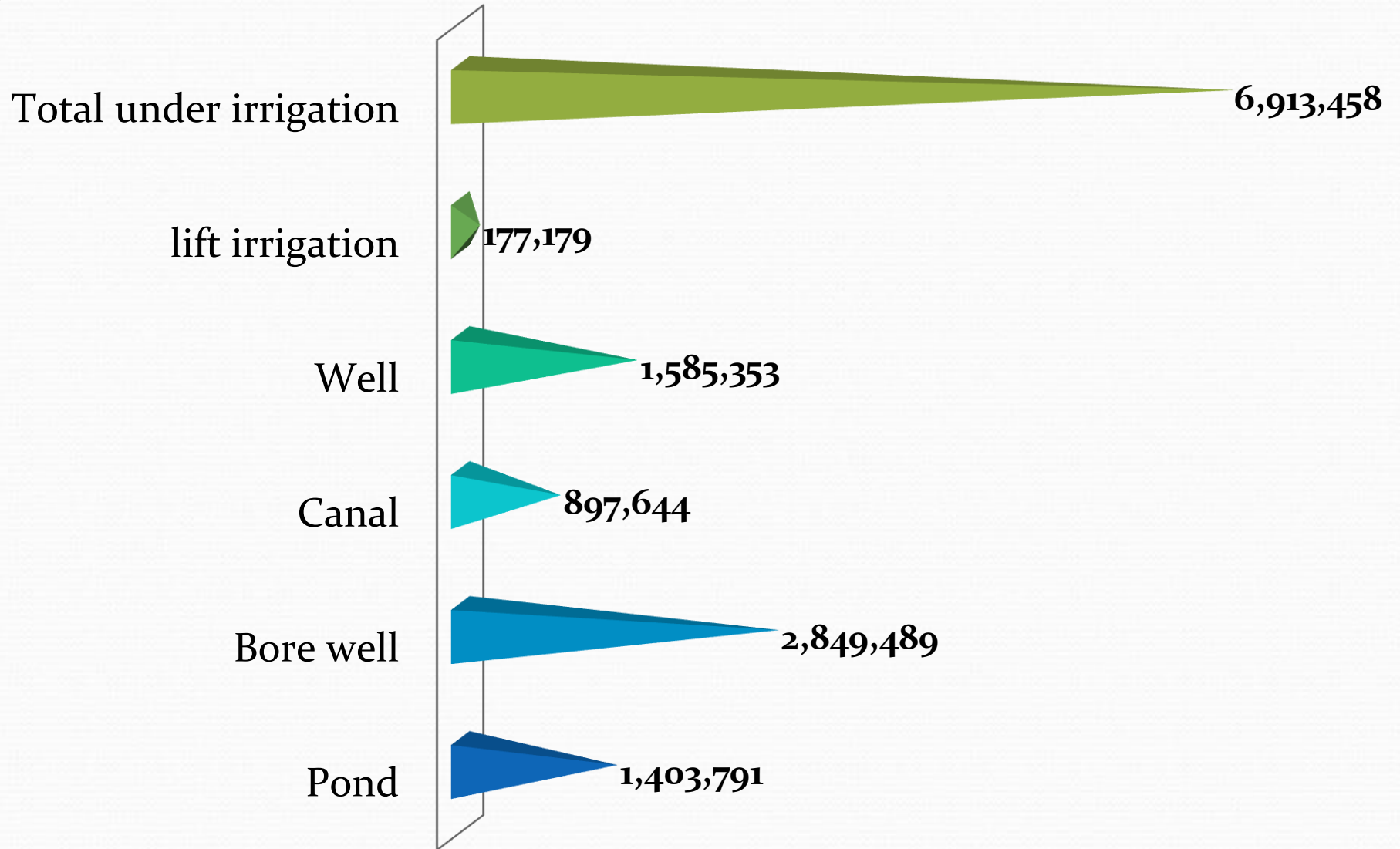
# Disability Population



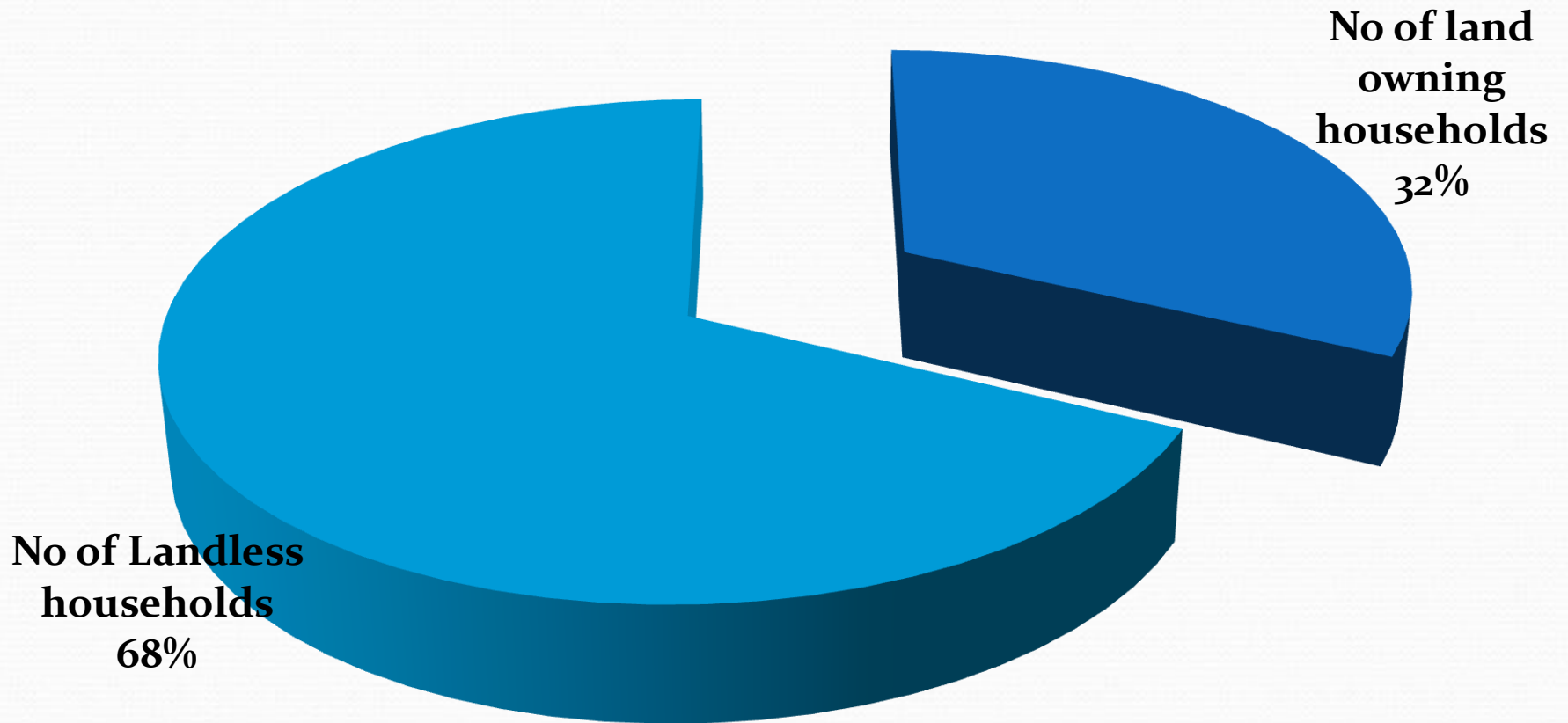
# Households - Land Possession



# Land under Irrigation - Acres

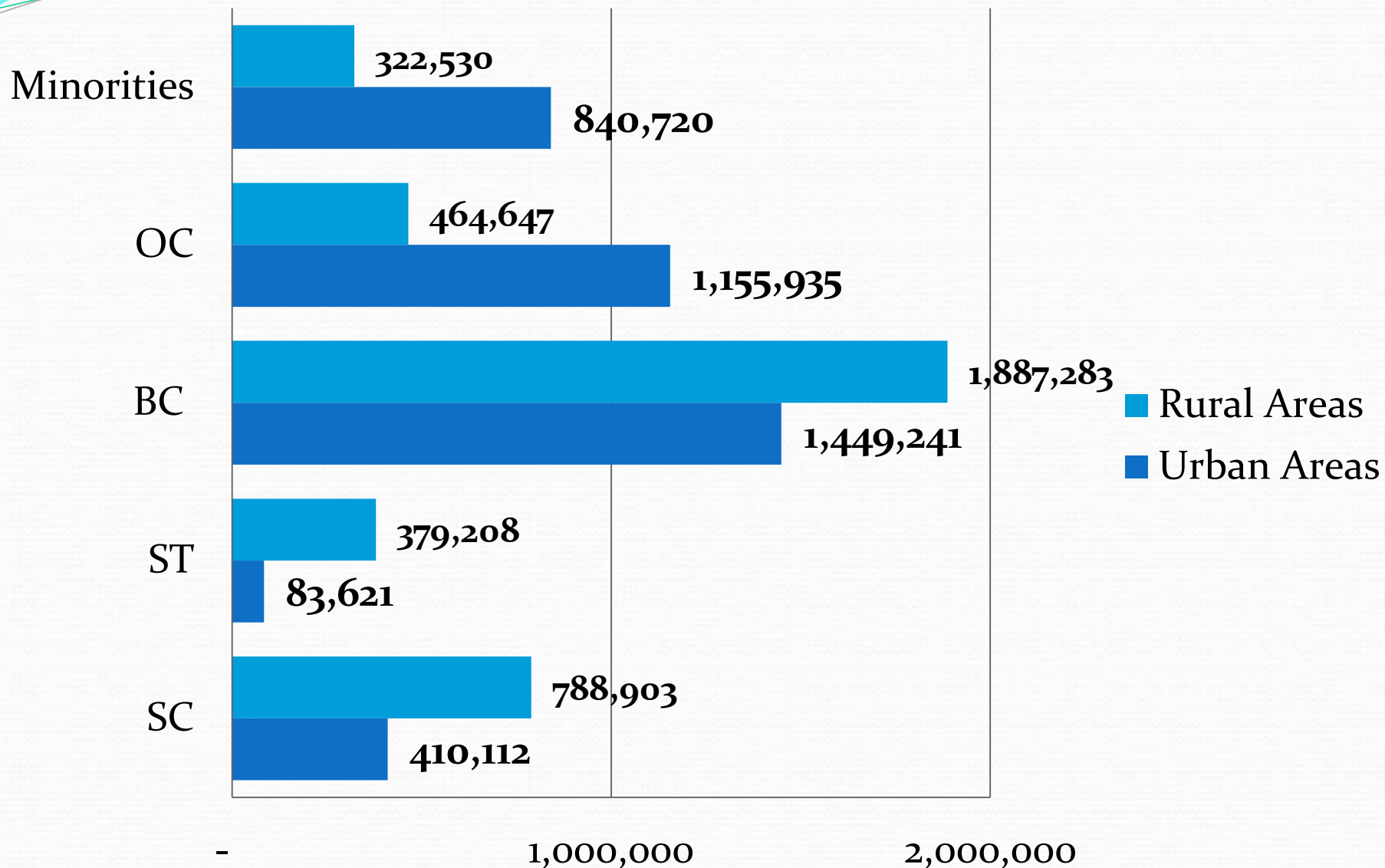


# Land Owners Vs Landless

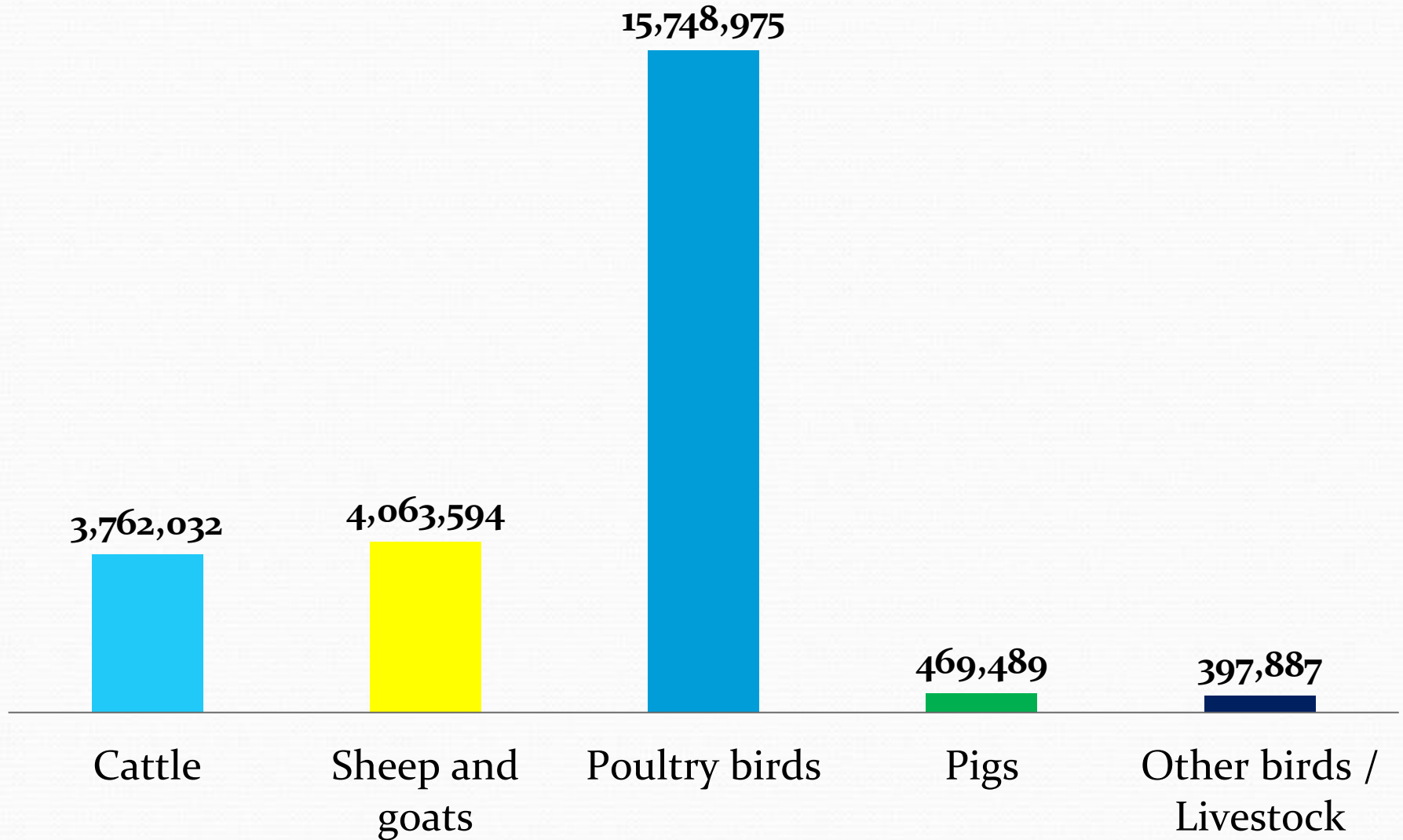




# Landless Households - Caste wise



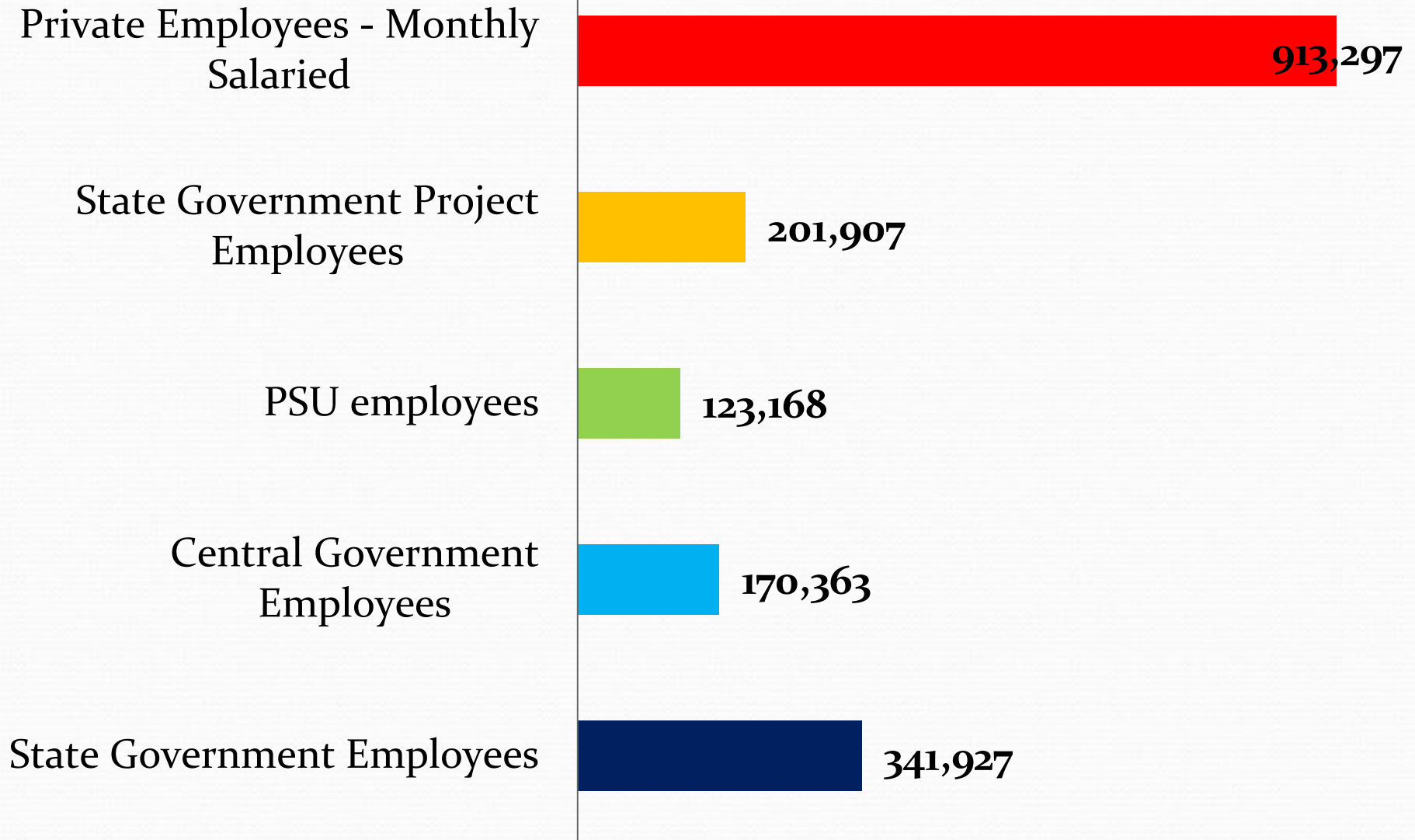
# Livestock - Status



# Employees Status

Sl. No	Employee Type	GHMC	Rest of GHMC (Districts including Municipalities and Municipal corporations)	Total
1	Total State Government Employees	64,821	2,77,106	3,41,927
2	Total Central Government Employees	56,989	1,13,374	1,70,363
3	Total PSU employees	20,465	1,02,703	1,23,168
4	Total State Government Project Employees	16,393	1,85,514	2,01,907
5	Total Private Employees - Monthly Salaried	4,42,349	4,70,948	9,13,297
6	Total Employees	<b>6,01,017</b>	<b>11,49,645</b>	<b>17,50,662</b>

# Employee Status

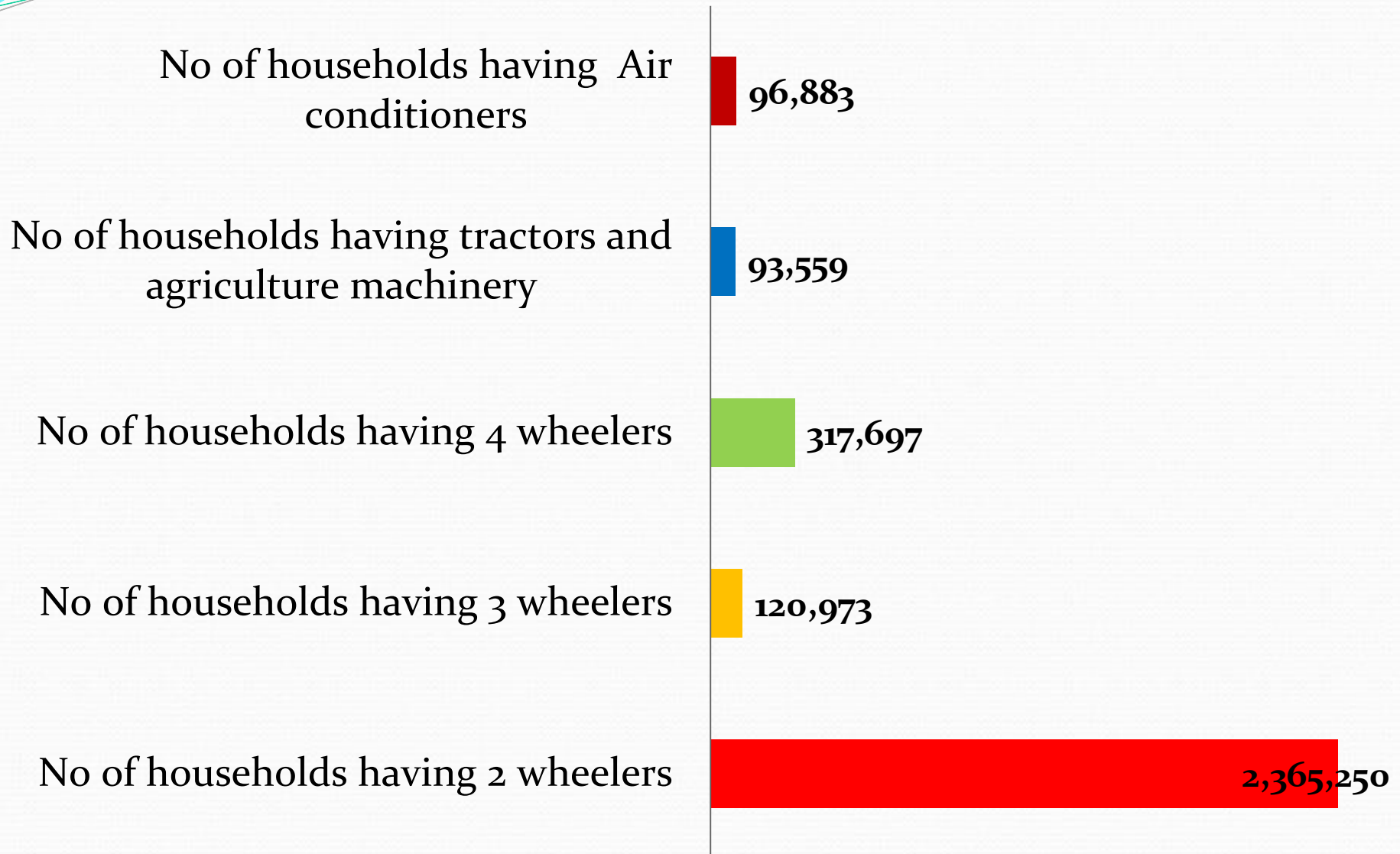




# Mobility - Statistics

<b>Sl. No</b>	<b>Vehicle Type</b>	<b>GHMC</b>	<b>Rest of GHMC (Districts including Municipalities and Municipal corporations)</b>	<b>Total</b>
1	Two Wheelers	6,44,603	17,20,647	23,65,250
2	Three Wheelers	19,898	1,01,075	1,20,973
3	Four Wheelers	1,57,381	1,60,316	3,17,697
4	Agricultural Equipment	3,154	90,405	93,559
5	Air Conditioner	64,487	32,396	96,883
6	Total	<b>8,89,523</b>	<b>21,04,839</b>	<b>29,94,362</b>

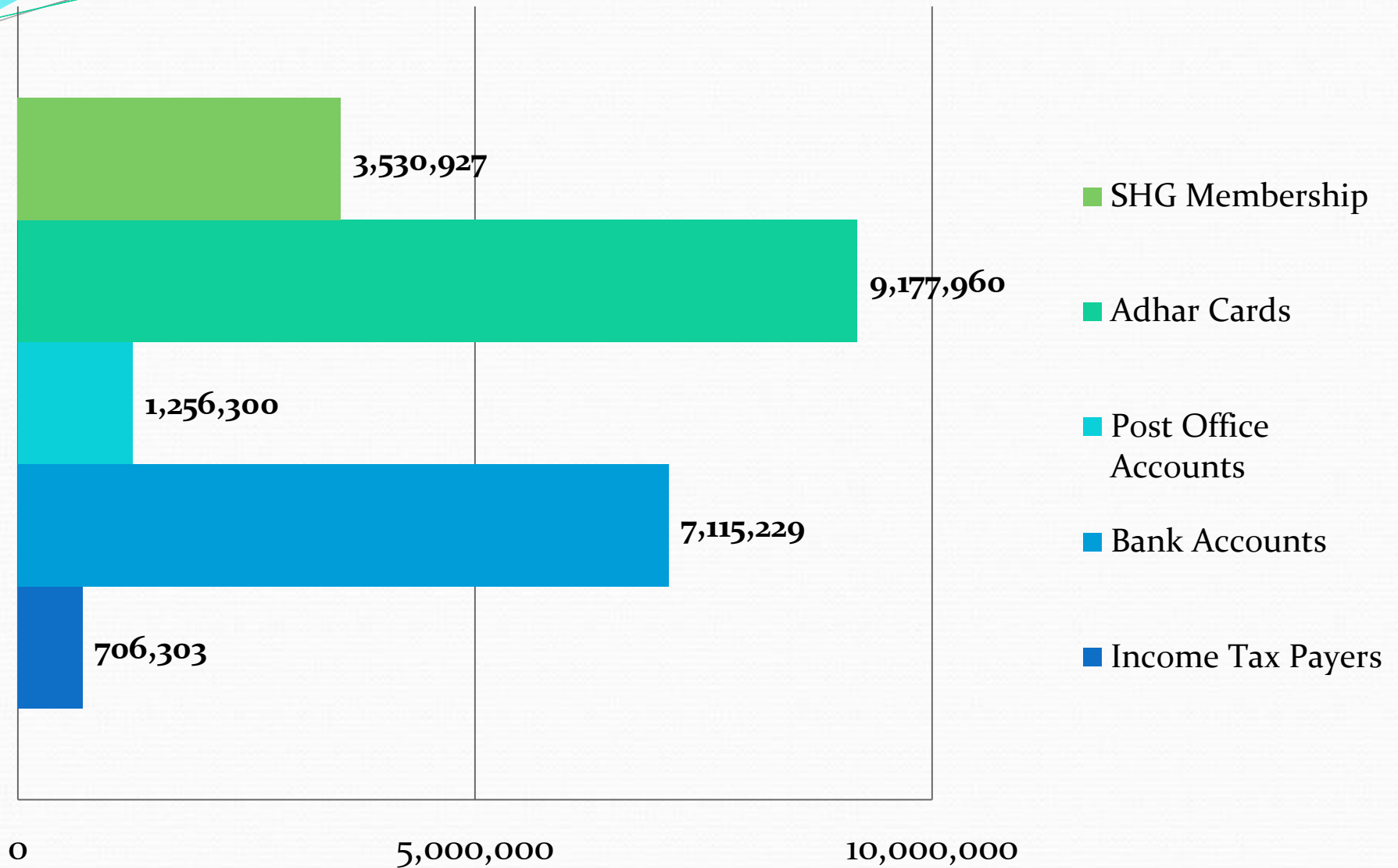
# Household – Movable Assets



# Household - Other Economic Indicators

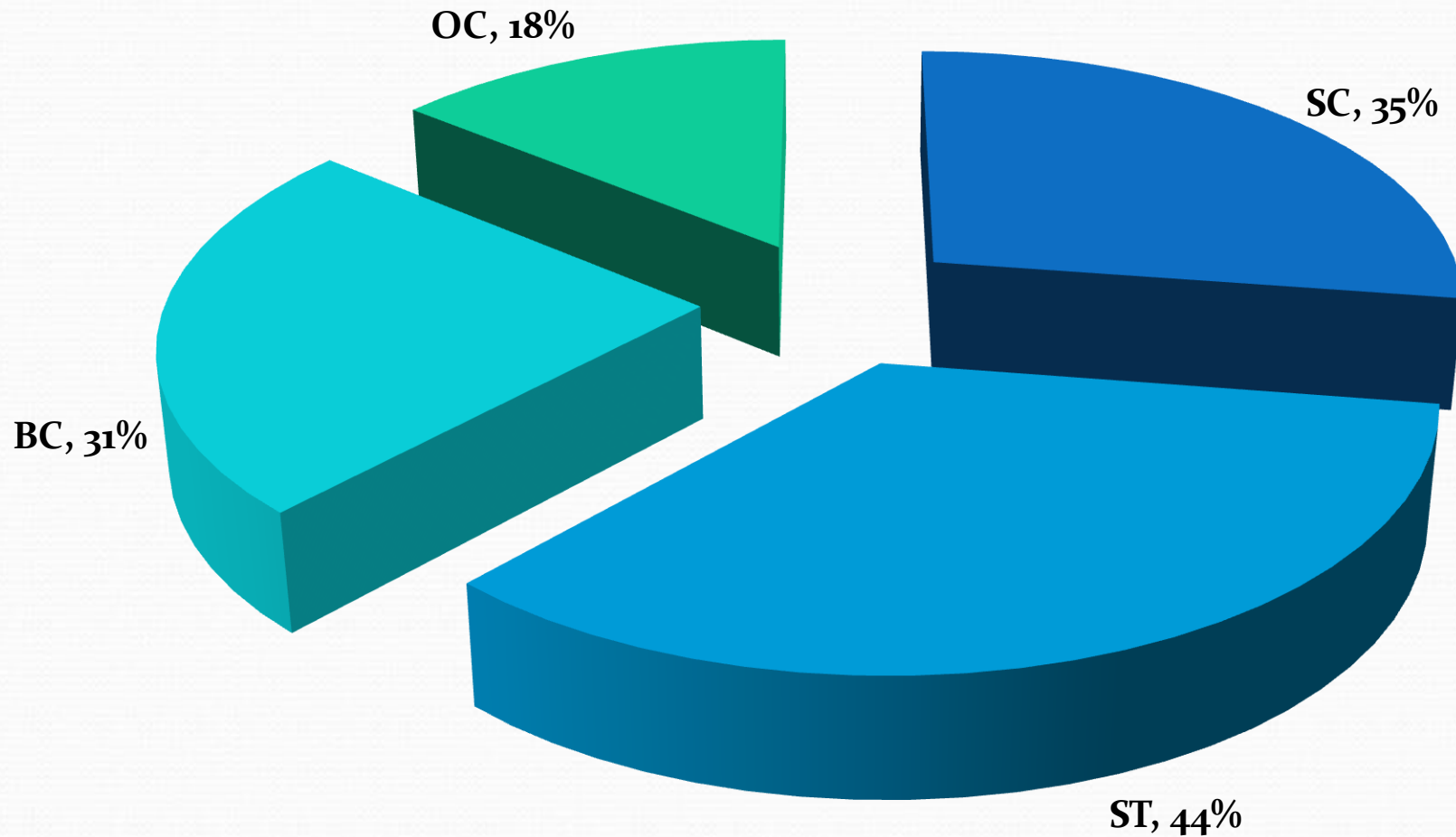
Sl. No	Economic Indicators	GHMC	Rest of GHMC	Total
1	IT Tax Payers	2,92,885	4,13,418	7,06,303
2	Large Business	7,461	19,741	27,202
3	Household having more than 10 cattle , 100 Small Ruminants, 3000 poultry birds	2,246	24,842	27,088
4	Bank Accounts	8,52,850	62,62,379	71,15,229
5	Post Office Accounts	7,809	12,48,491	12,56,300
6	SHG Membership	33,934	34,96,993	35,30,927

# Other Economic Indicators

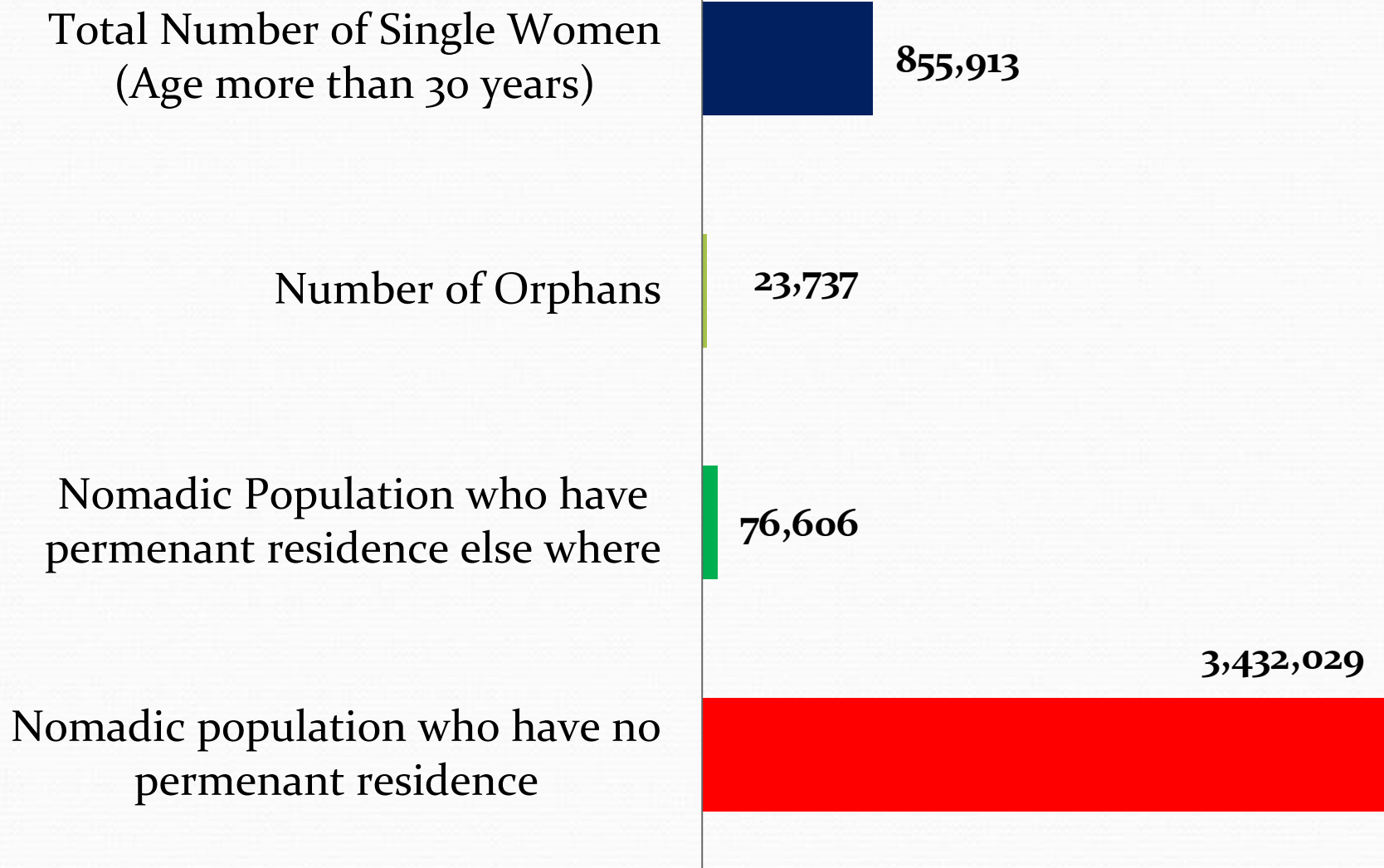




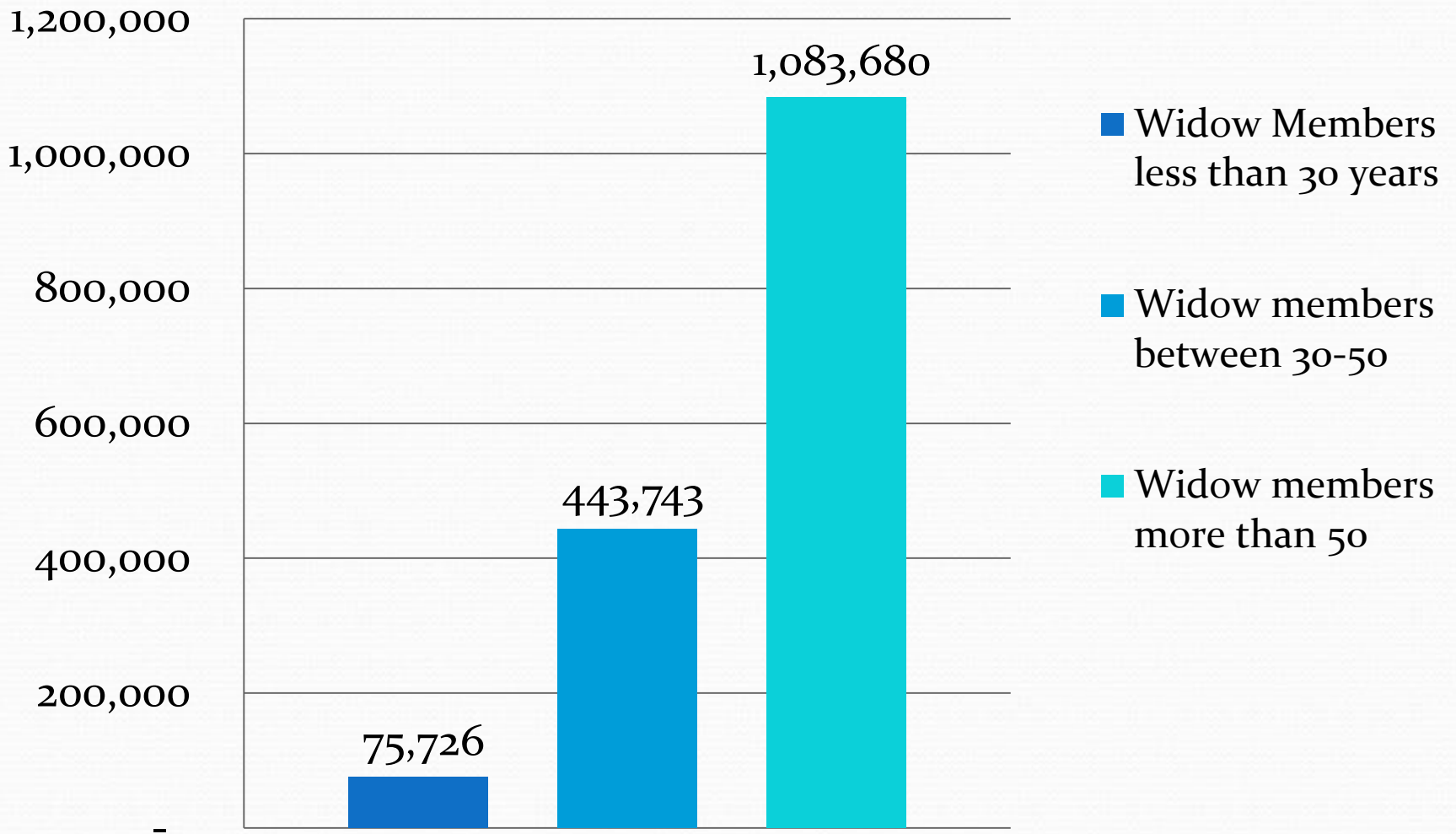
# Illiterates – Caste wise



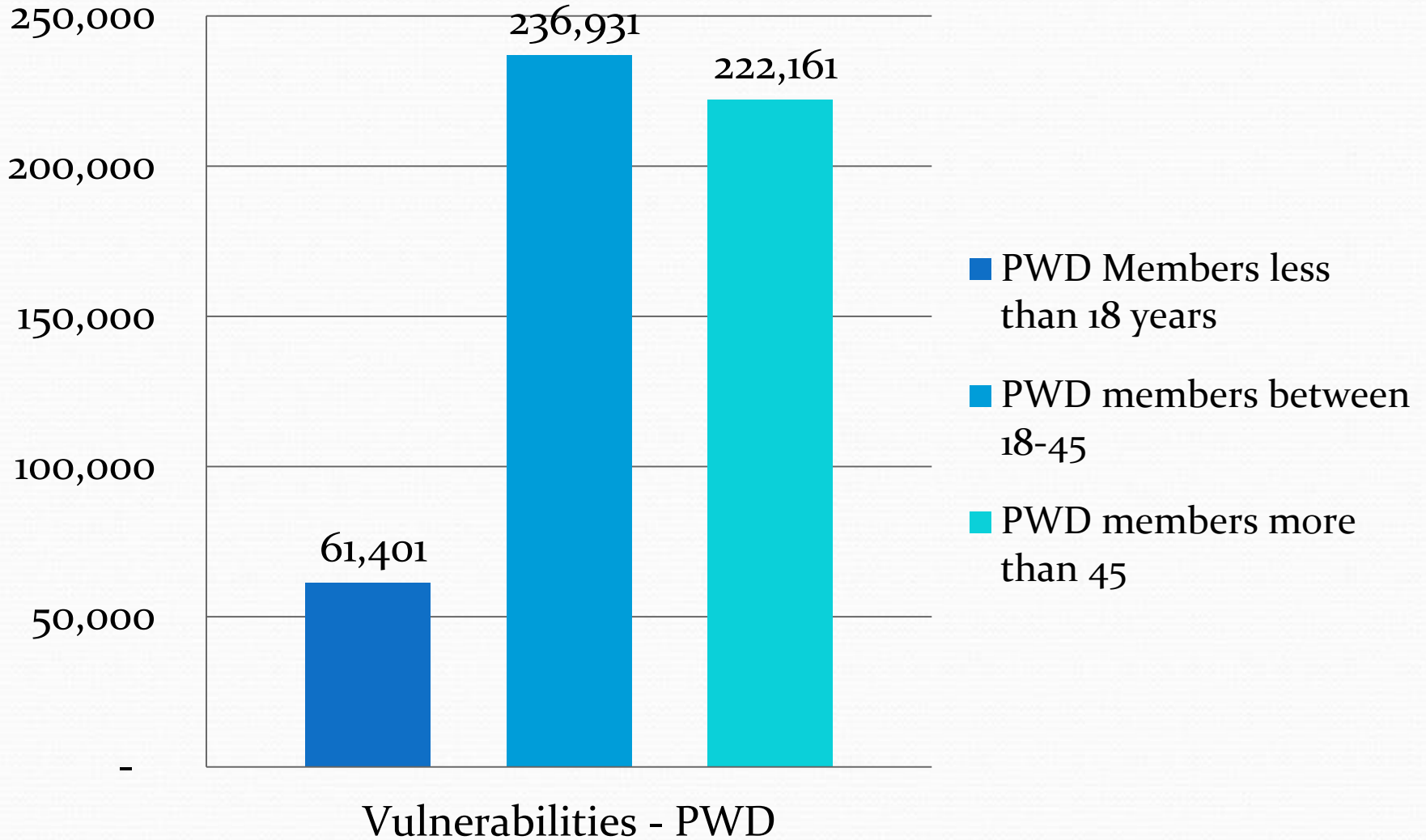
# Vulnerability



# Vulnerability - Widow

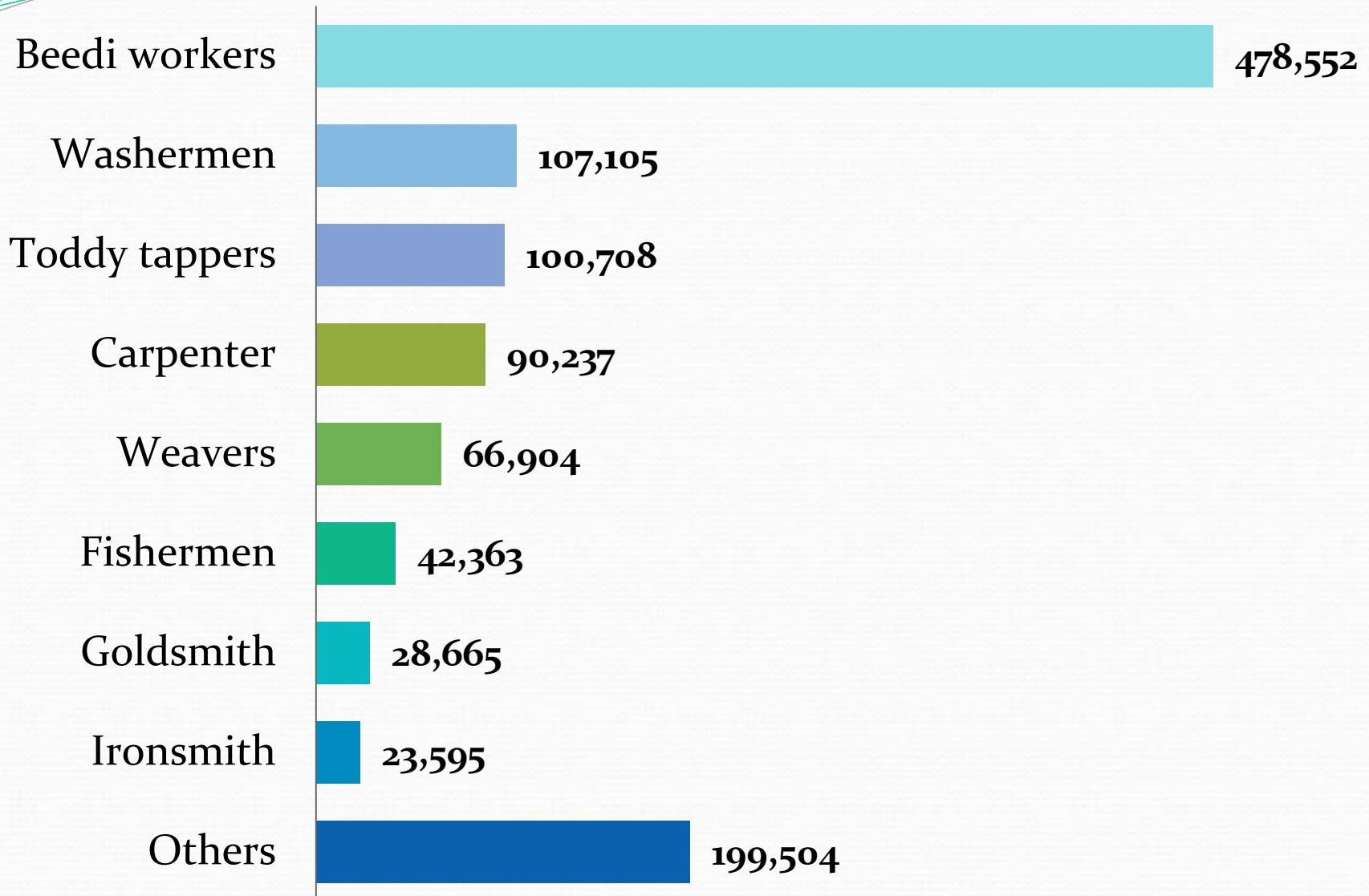


# Vulnerabilities - PWD





# Artisans – Population



# Households - Housing Status

Housing with Govt assistance



Having no electricity facility



Having no toilets facility



Living in rented house



Living in Own house



# House Facilities

Four rooms and above



Three rooms



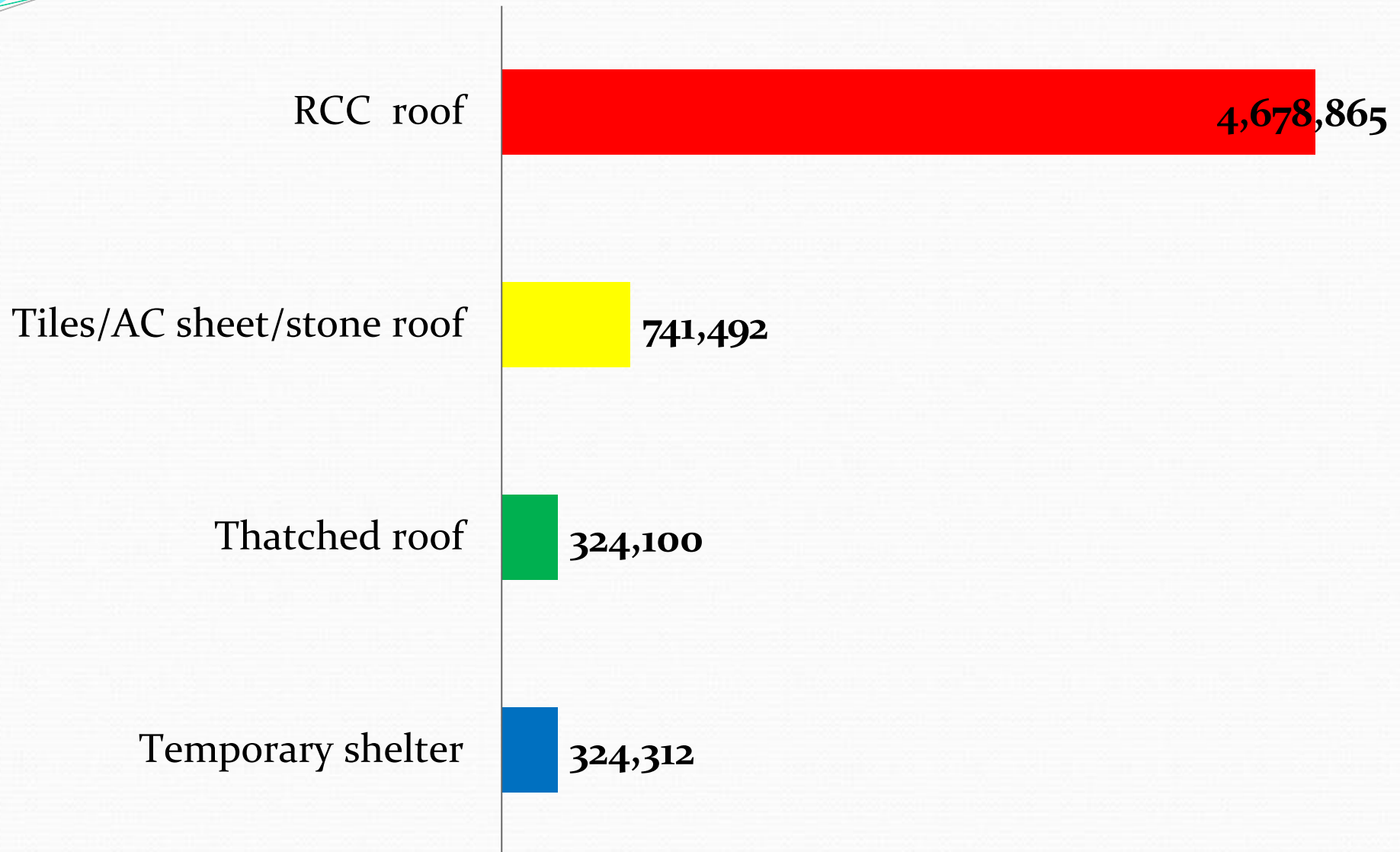
Two rooms



One room

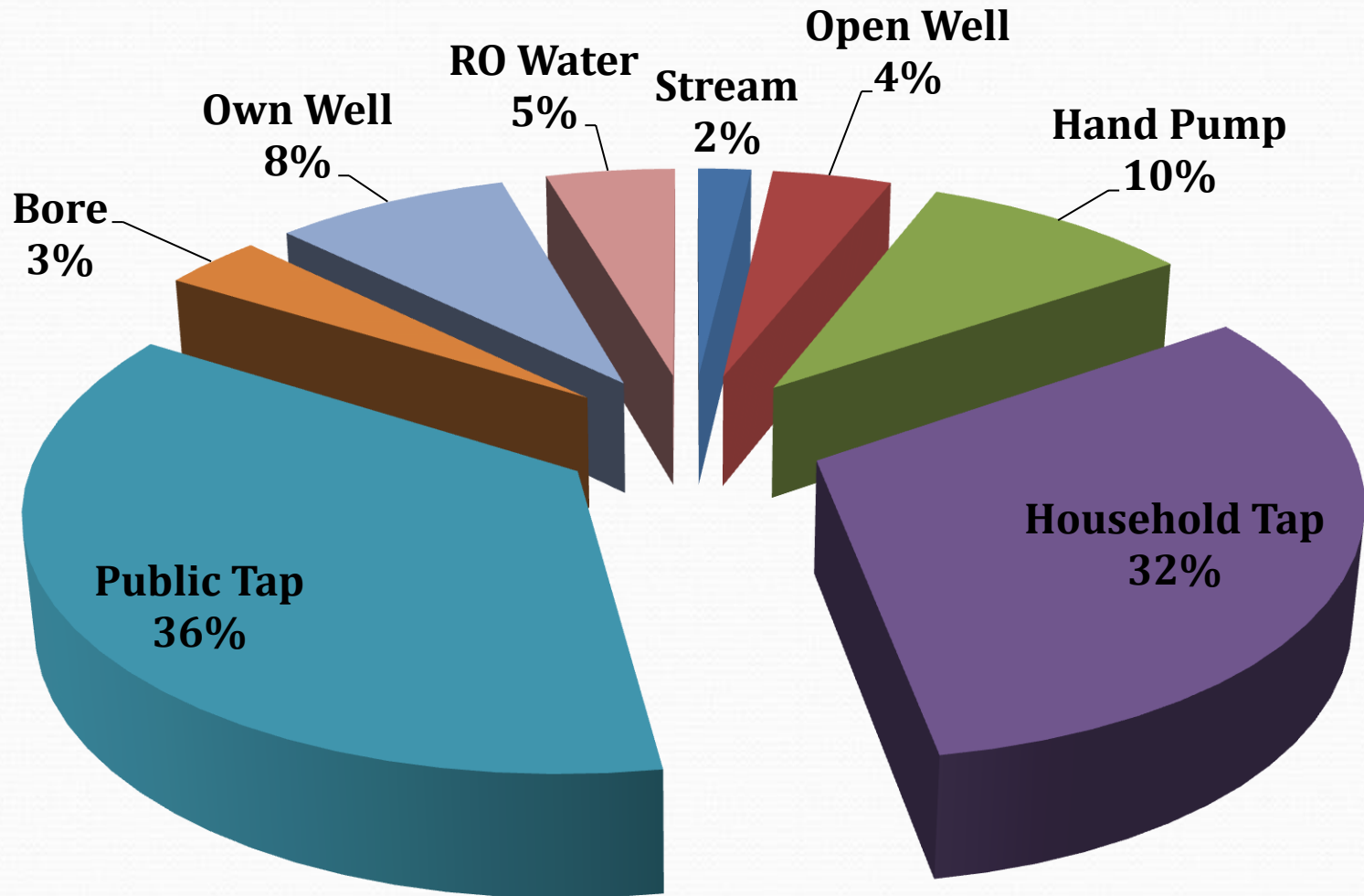


# Housing - Roof Pattern

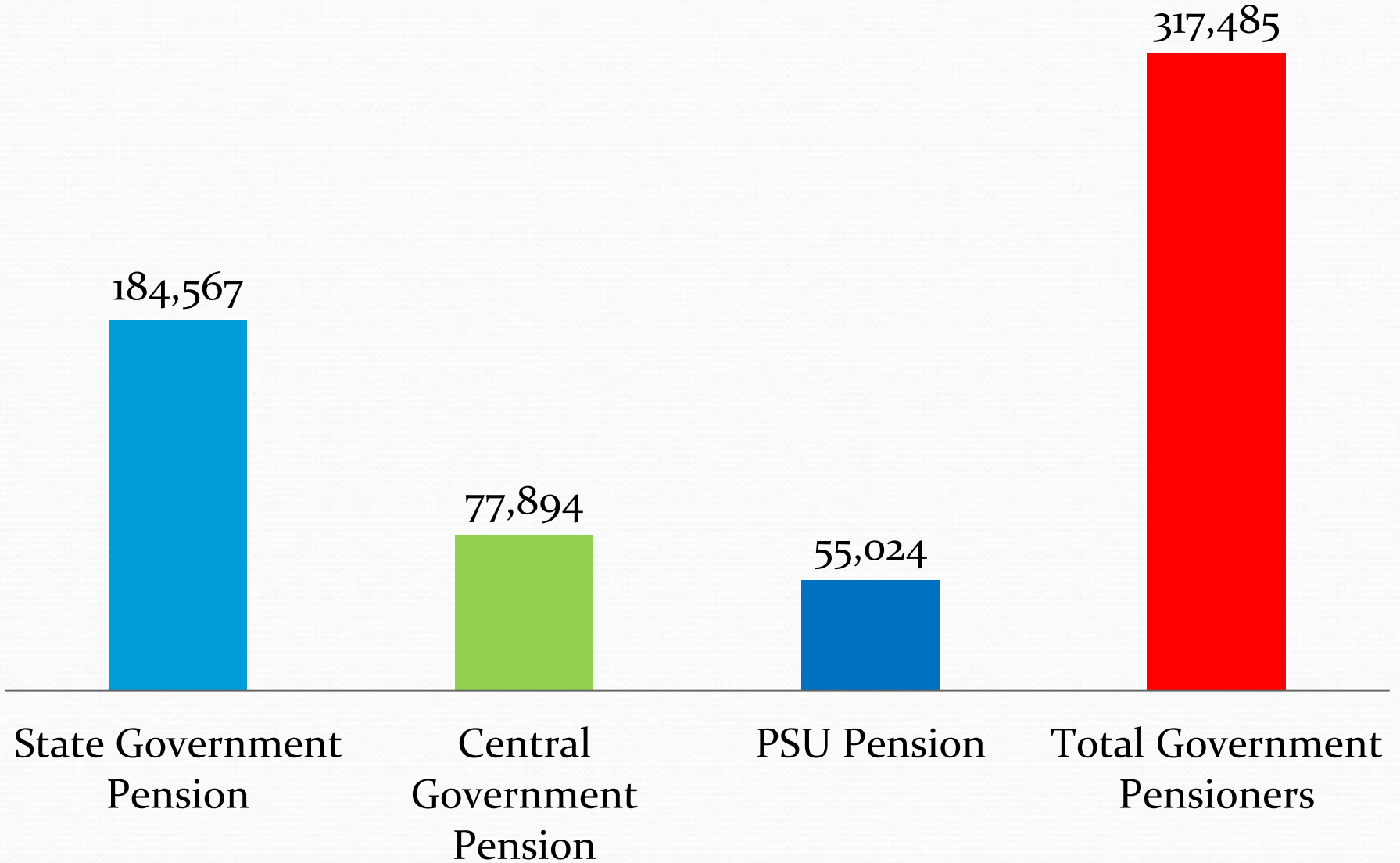




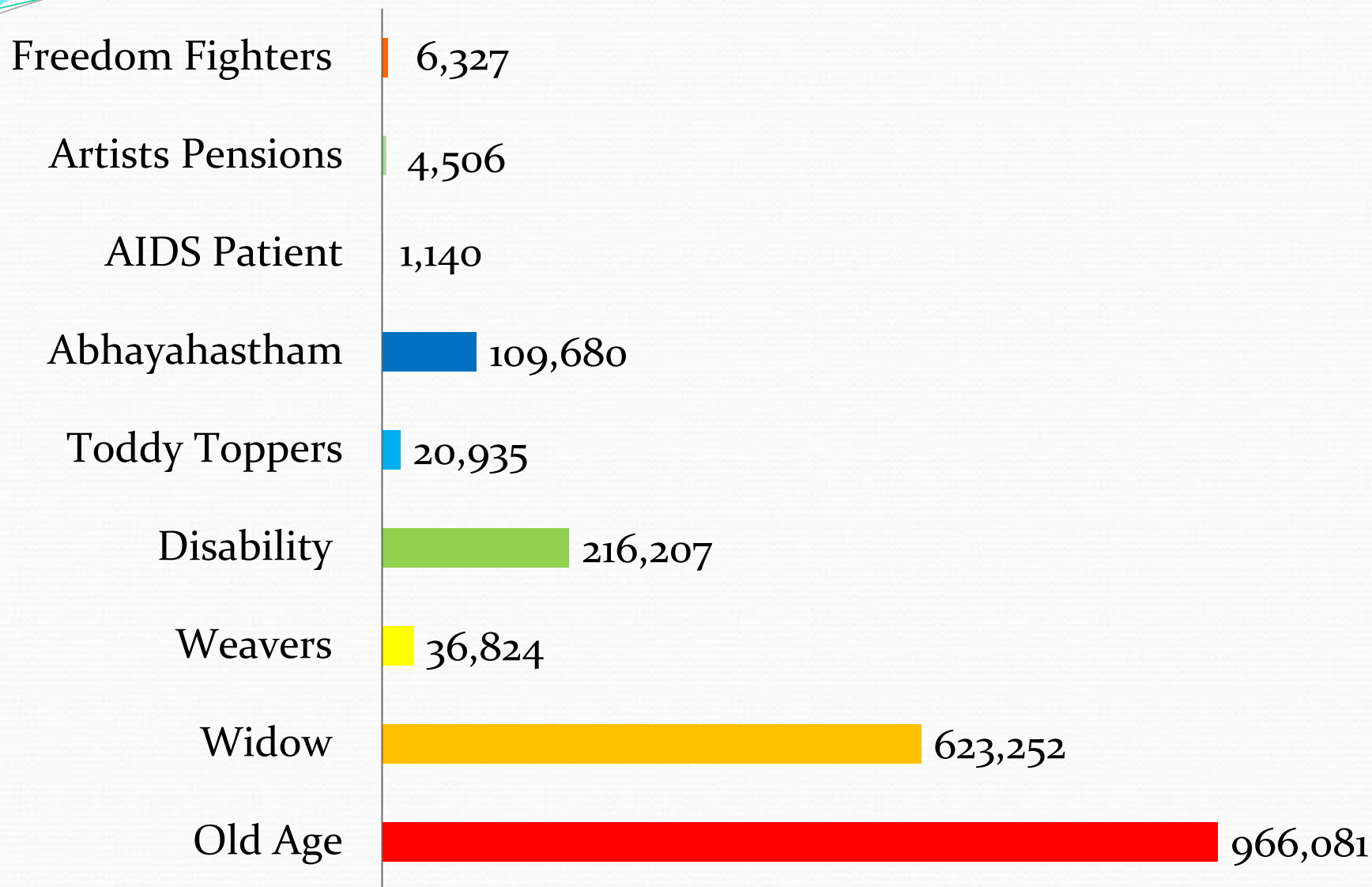
# Drinking Water - Status



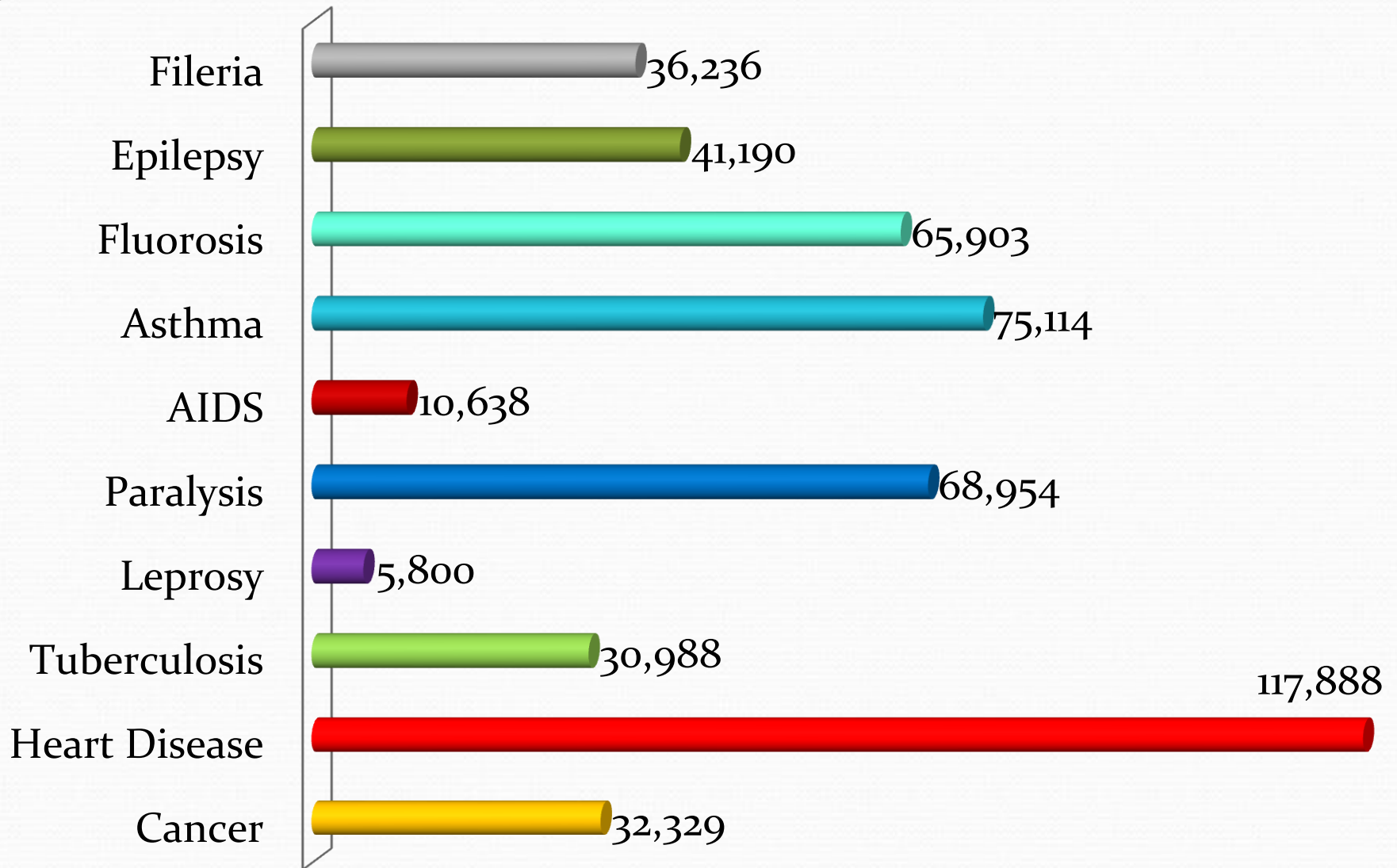
# Government Pensioners



# Social Security Pensioners

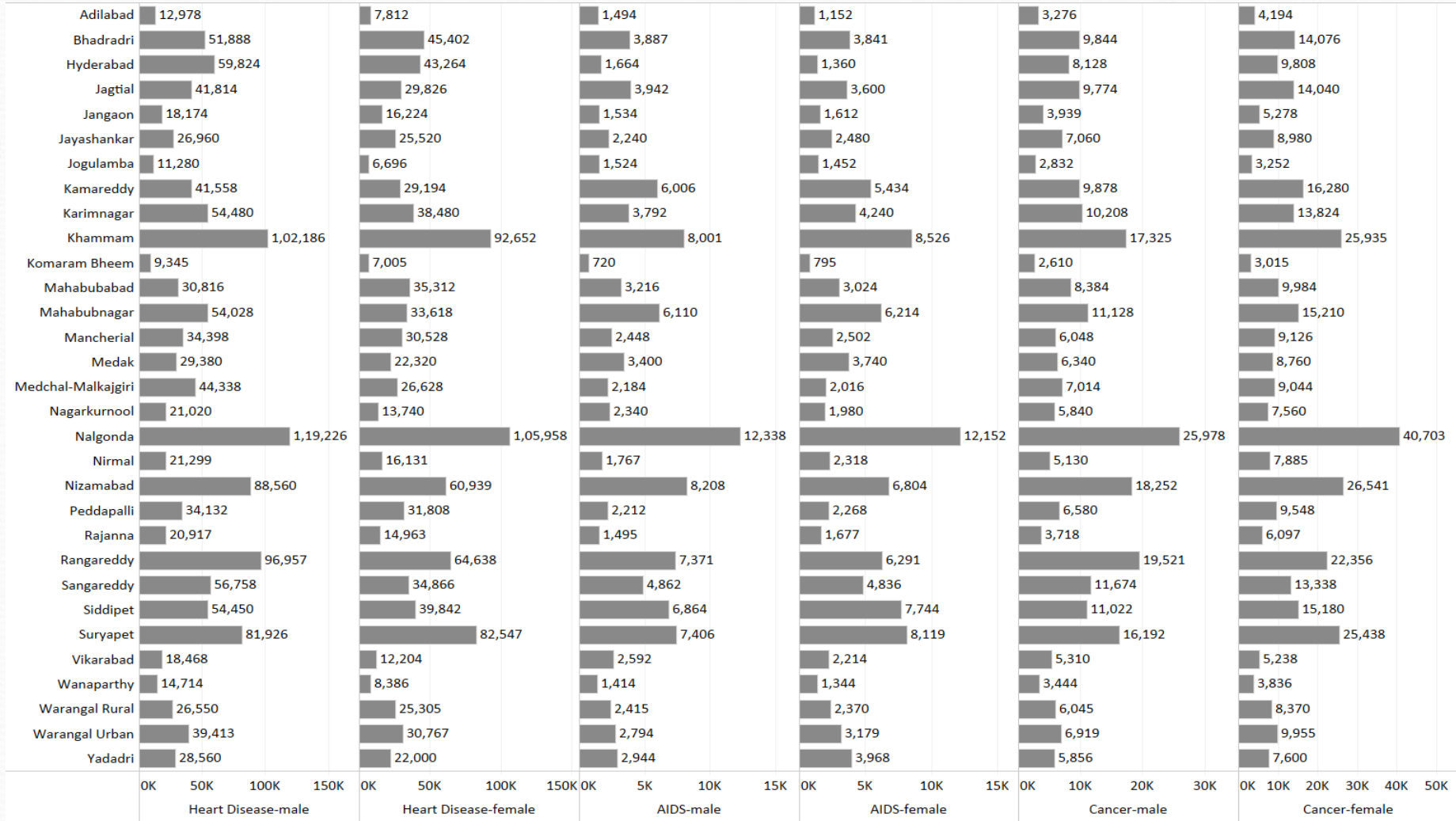


# Chronic Diseases

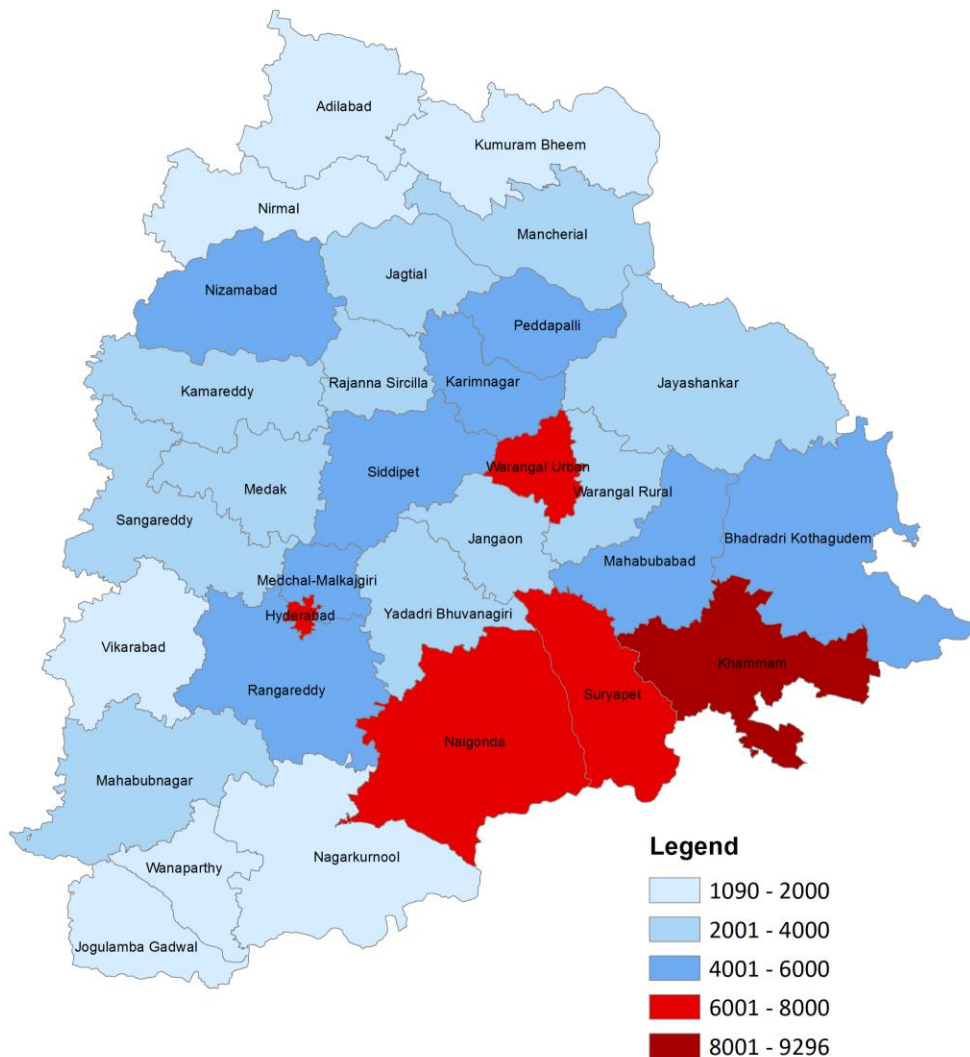




# Gender wise Diseases (Heart, AIDS, Cancer)

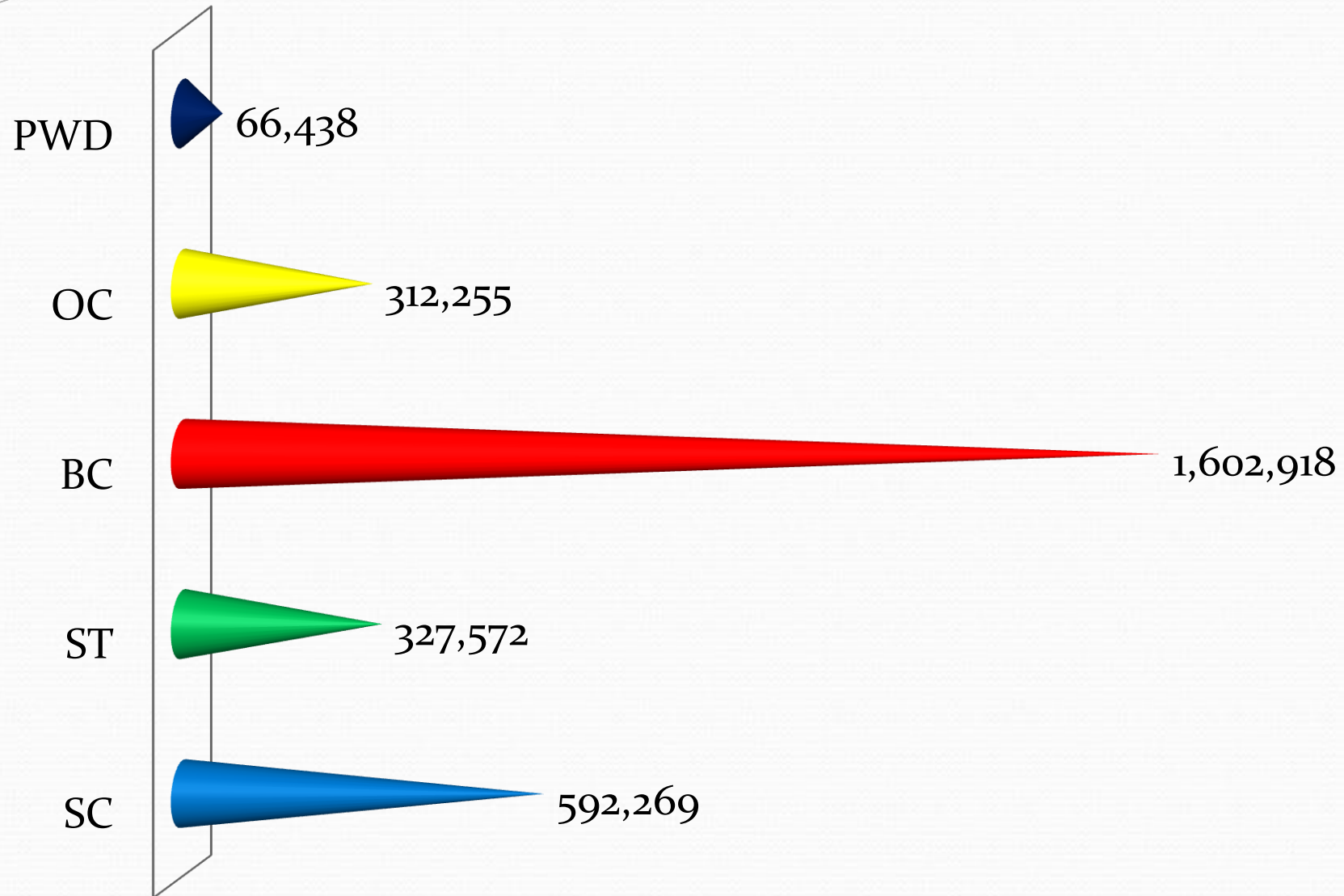


# District wise disease burden – Heart Diseases

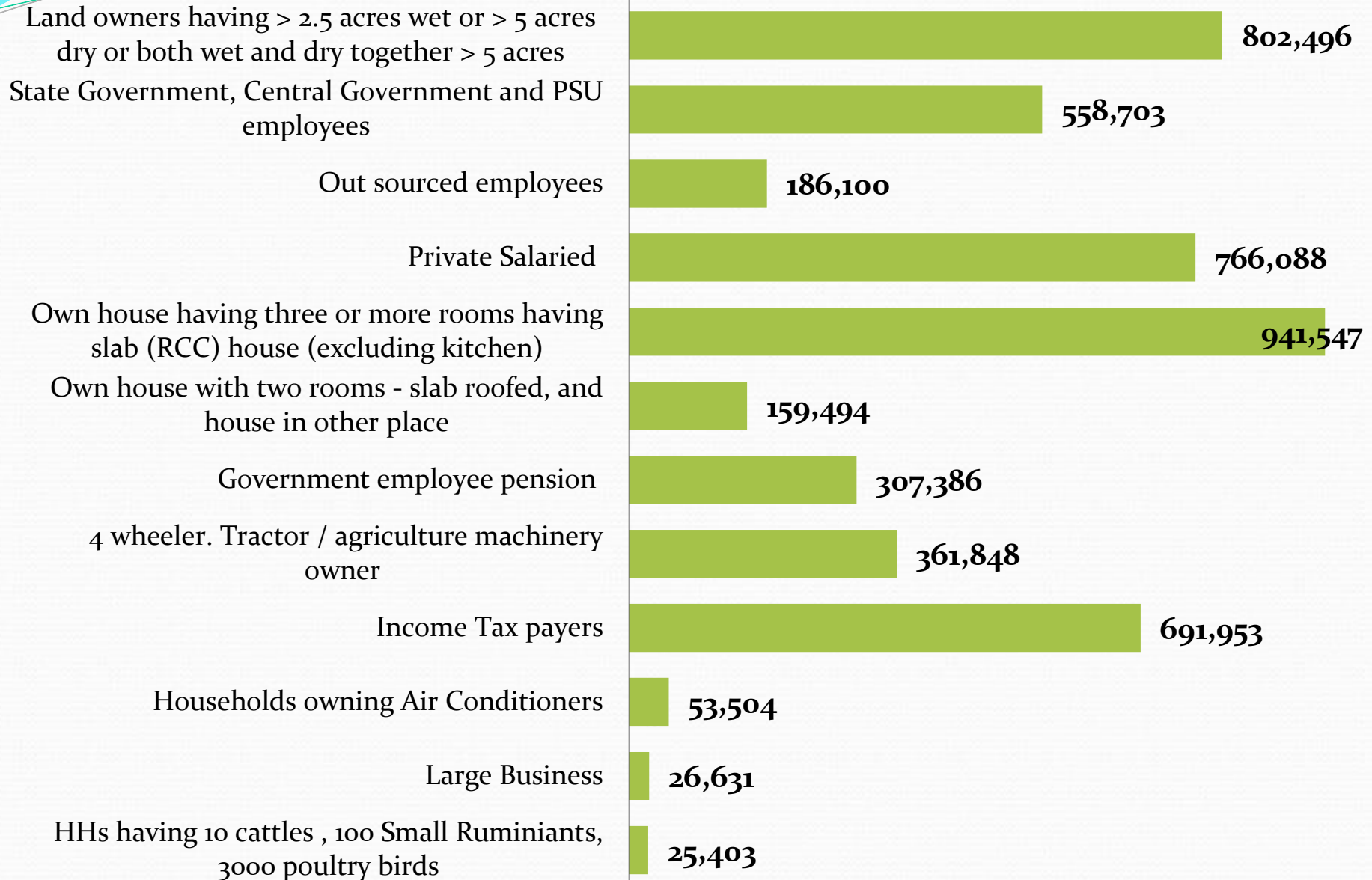


S.No.	District	No. of Heart Diseases
1	Kumuram Bheem	1,090
2	Adilabad	1,156
3	Jogulamba Gadwal	1,499
4	Wanaparthy	1,653
5	Vikarabad	1,708
6	Nagarkurnool	1,743
7	Nirmal	1,972
8	Medak	2,588
9	Jayashankar	2,627
10	Jangaon	2,649
11	Rajanna	2,765
12	Yadadri	3,168
13	Kamareddy	3,224
14	Mahabubnagar	3,374
15	Warangal Rural	3,464
16	Sangareddy	3,532
17	Mancherial	3,614
18	Jagtial	3,986
19	Mahabubabad	4,141
20	Bhadradi	4,236
21	Siddipet	4,295
22	Peddapalli	4,717
23	Medchal-Malkajiri	5,075
24	Nizamabad	5,546
25	Karimnagar	5,815
26	Rangareddy	5,995
27	Warangal Urban	6,389
28	Hyderabad	6,449
29	Suryapet	7,157
30	Nalgonda	7,269
31	Khammam	9,296
<b>Total</b>		<b>1,22,192</b>

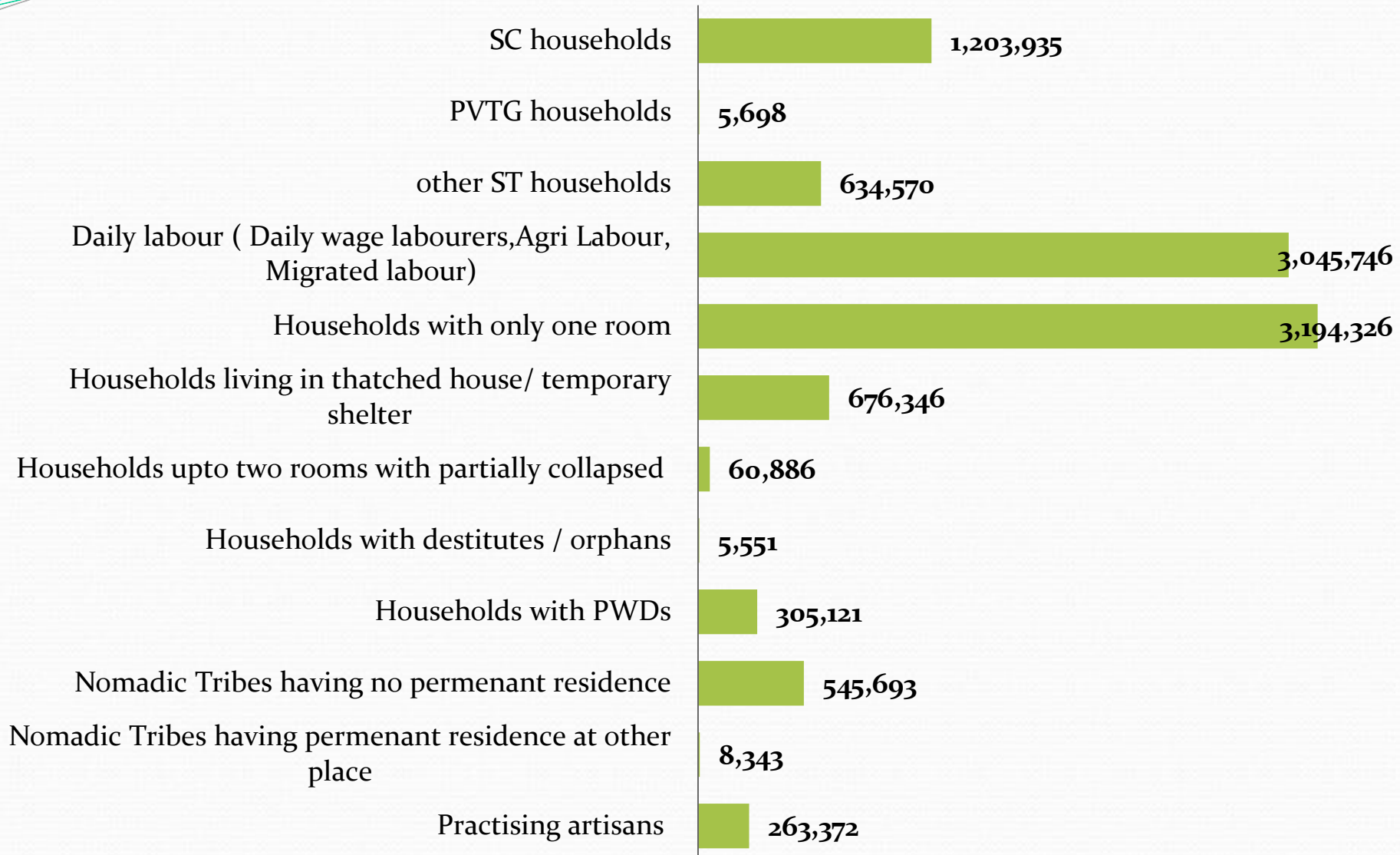
# Household - SHG Membership



# Exclusion Criteria – wise Households

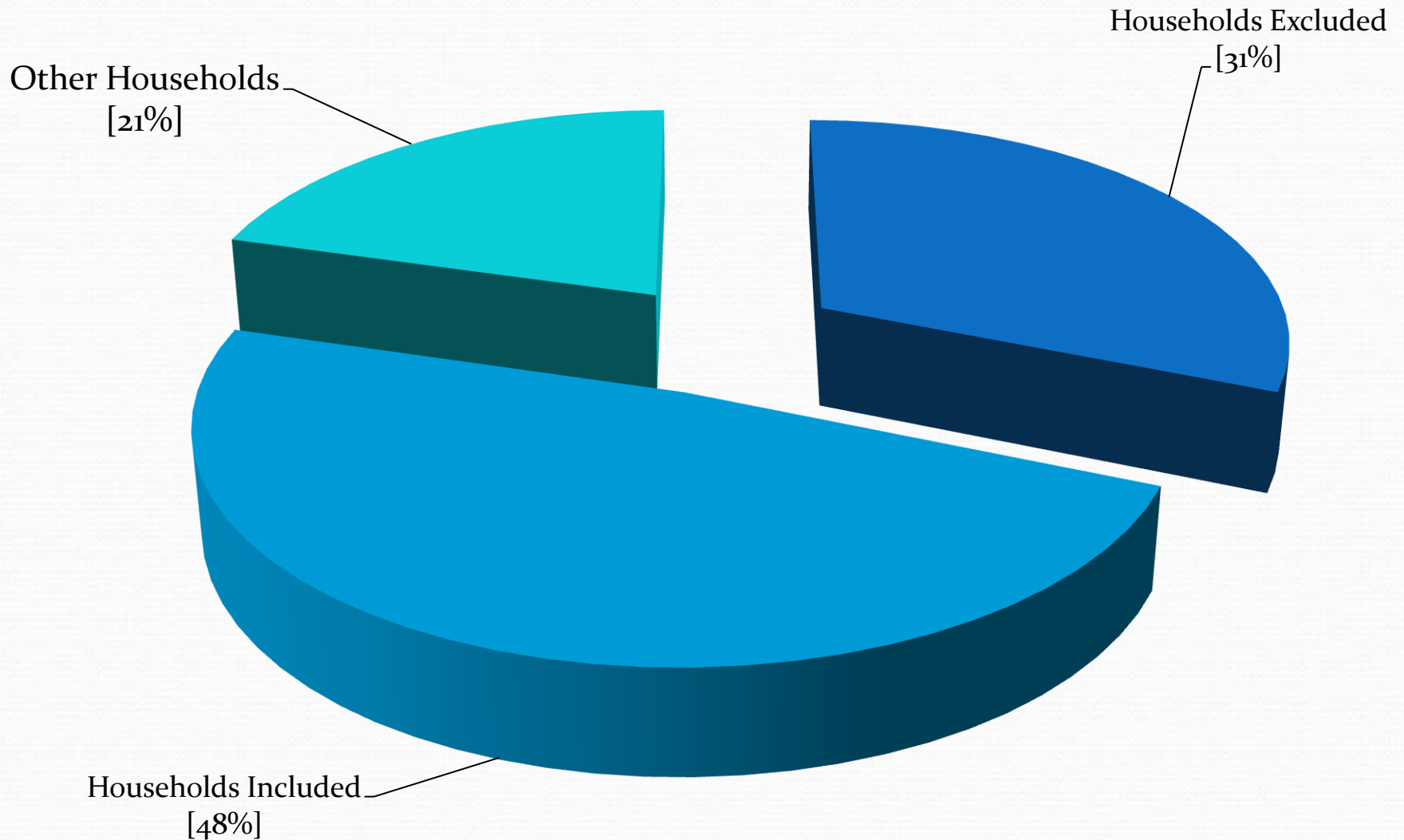


# Inclusion Parameter – wise households



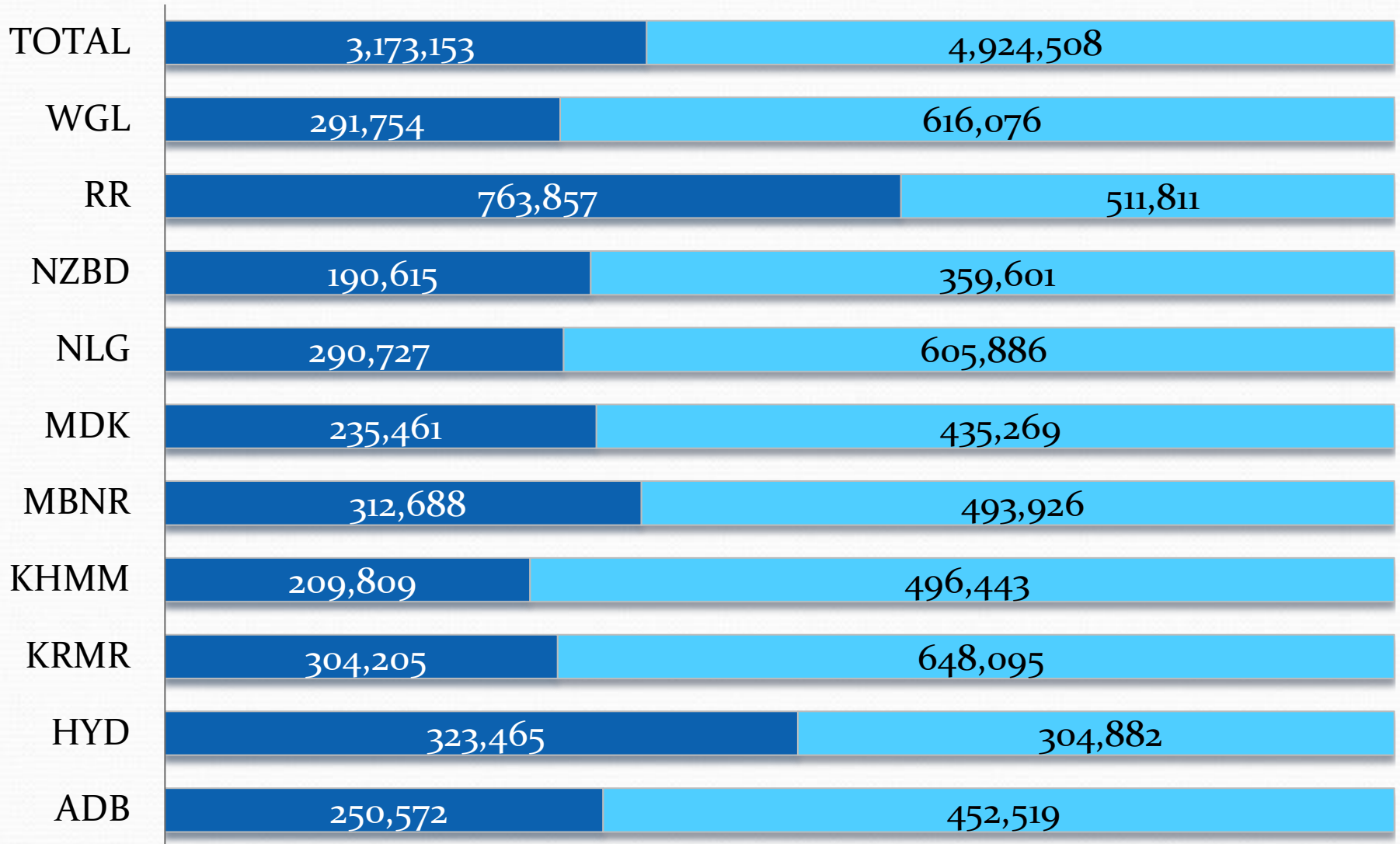


# Exclusion & Inclusion



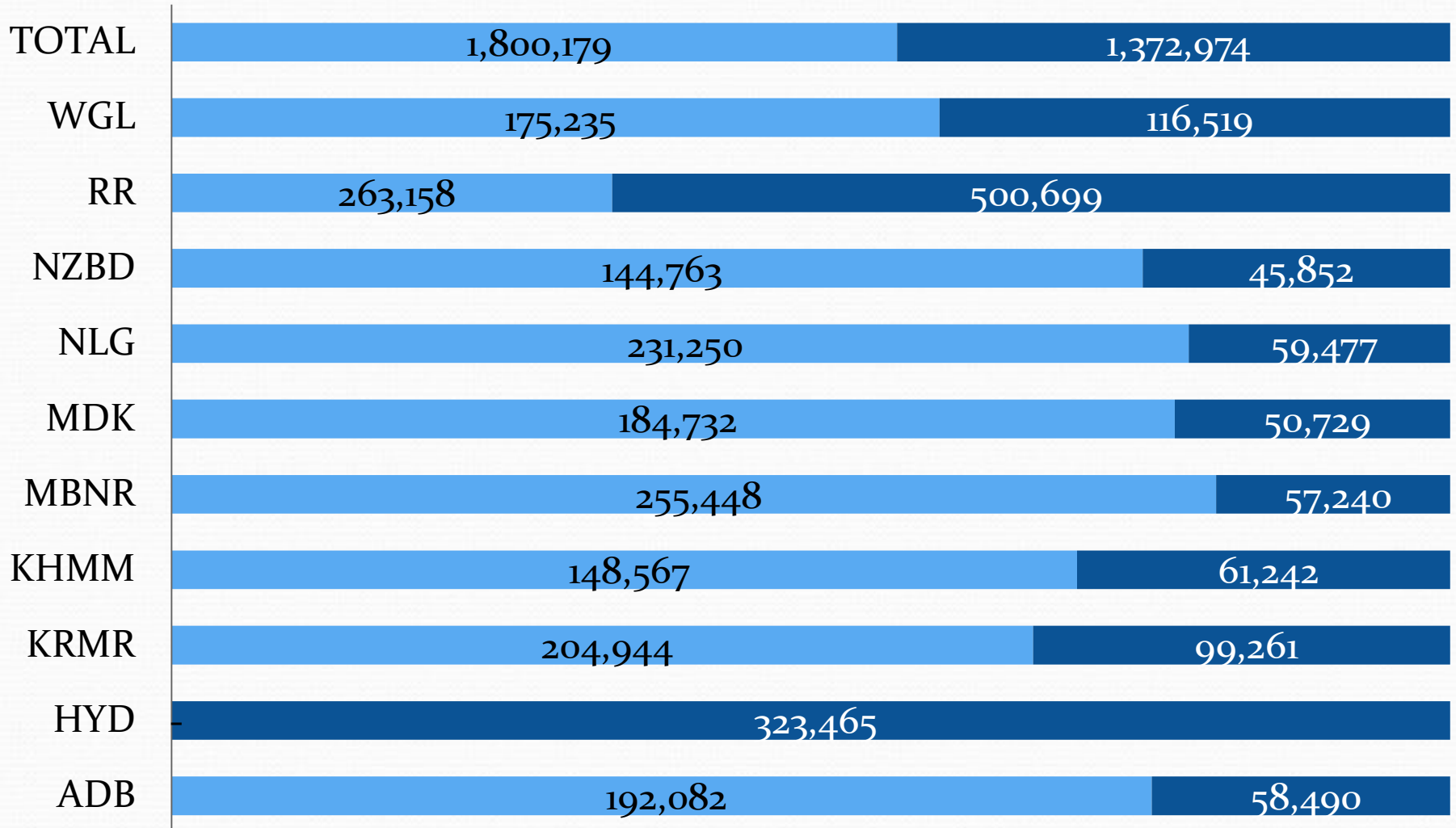
# Exclusion & Inclusion – District Wise

■ Excluded Households    ■ Included Households



# Exclusion Households

■ Rural ■ Urban

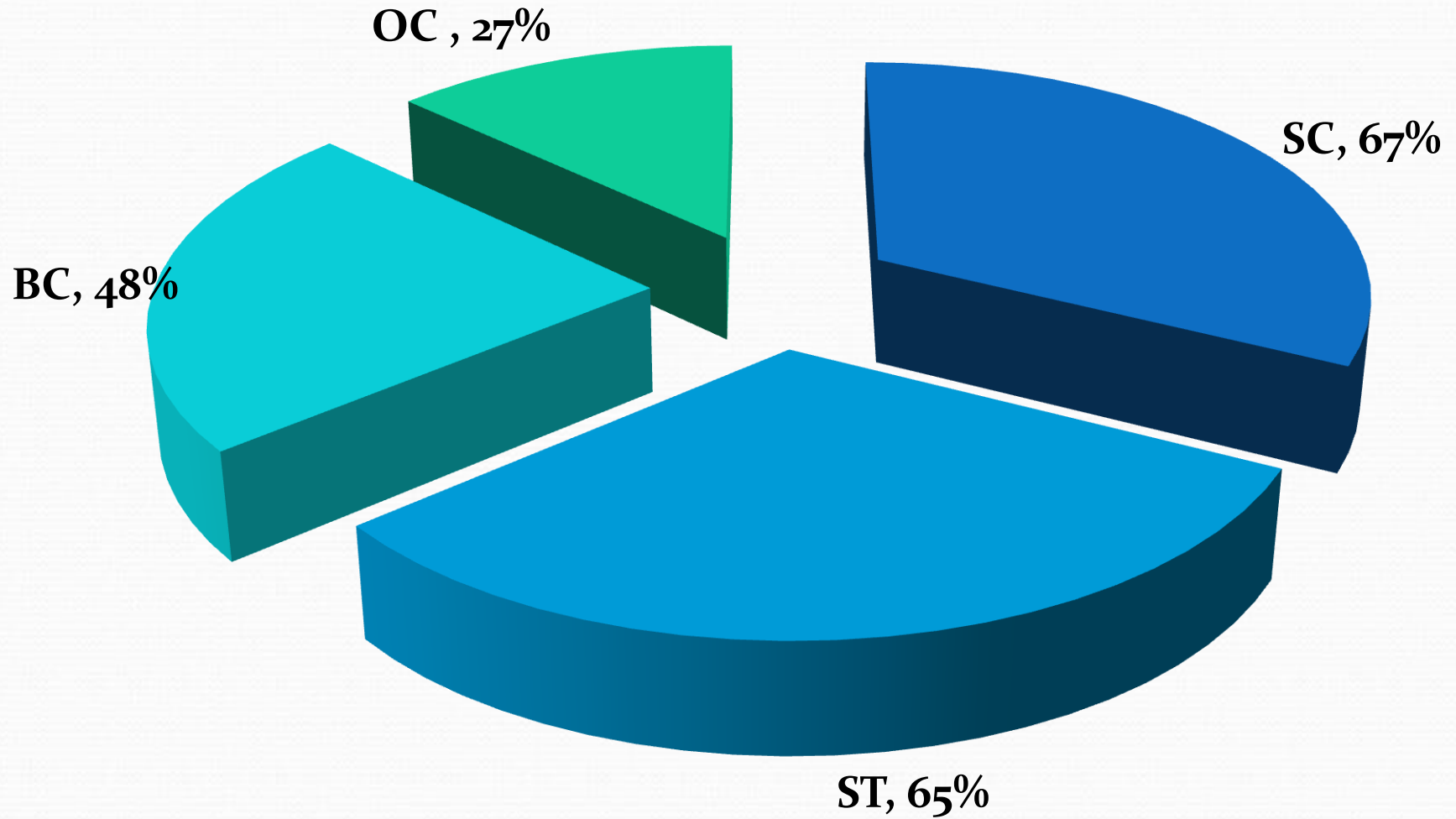


# Inclusion Households

■ Rural ■ Urban

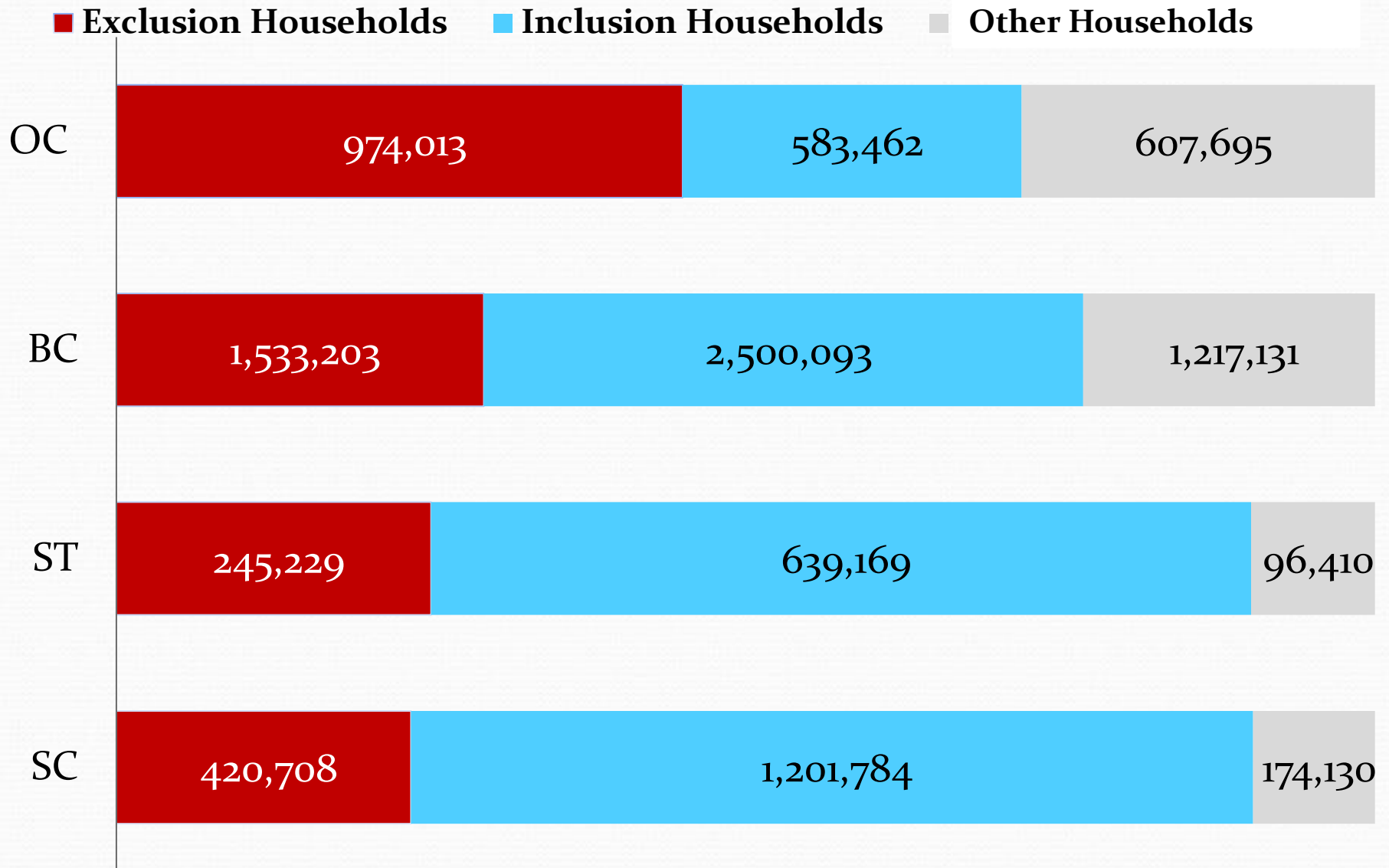


# Inclusion Households – Caste Wise





# Households – Caste wise



# Limca

## Book of Records

### National Record

The Government of Telangana conducted the most intensive household survey by a state, the state-wide **Intensive Household Survey-2014**, to ascertain the social and economic status of the citizens living in all the 10 districts of the State on August 19, 2014. The exercise had nearly 4,00,000 government employees covering over 1.09 crore households in a single day. The survey began at 7.00 am and continued till 8.00 pm. The enumerators collected data such as names, age, educational qualification, profession, Aadhaar card numbers, bank account number, property and other details of the family members.

*Vijaya Ghose*

**Vijaya Ghose**  
Editor, Limca Book of Records



"LIMCA BOOK OF RECORDS" IS THE COPYRIGHT OF THE COCA-COLA COMPANY. "LIMCA" IS THE REGISTERED TRADEMARK OF THE COCA-COLA COMPANY.  
THIS CERTIFICATE DOES NOT NECESSARILY DENOTE AN ENTRY INTO LIMCA BOOK OF RECORDS.



# Thank you