

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

Partho Pratim Bhadra¹, Tushar Kanti Ghara²

¹Deputy Director, Directorate of Census Operations, West Bengal, Govt. of India

IB-199, Sector –III, Salt Lake, Kolkata 700106

²Joint Director of Public Instruction, Department of Higher Education, Govt. of West Bengal

8th Floor, Bikash Bhavan, Salt Lake City, Kolkata 700091

Corresponding Author: **Tushar Kanti Ghara**

Abstract: Population Census in India is conducted once in every ten years under the census Act 1948 and Census rules 1990. The last decennial census was conducted in the year 2011. The three broad categories of Houselisting & Housing Census are Housing condition, Amenities and Assets. Here, an attempt has been made to use four different approach using Probability theory for finding Ranks based on these House listing and Housing Census Data for rural areas. Although, there may be different Index developed to find such ranking, ranks obtained here using these different approaches are then compared and statistically tested. Statistical techniques like Probability, Weighted Geometric mean and Weighted Arithmetic mean, Correlation coefficient, and other Descriptive Statistical theories have been applied on this.

Keywords: Probability, Weighted Geometric mean, Proportionate weight, ranking, correlation

INTRODUCTION

India is a welfare State. Since independence, various welfare schemes have been launched for the welfare of the common man. This information is provided by the Census. Census in India is conducted in two phases, (i) House listing & Housing Census is the 1st Phase. and (ii) Population Census is the 2nd phase. The Houselisting and Housing Census has immense utility as it provides comprehensive data on the conditions of human settlements, housing deficit and consequently the housing requirement to be taken care of in the formulation of housing policies. This also provide a wide range of data on amenities and assets available to the households, information much needed by various departments of the Union and State Governments and other non-Governmental agencies for development and planning at the local level as well as the State level. This also provide the base for Population Enumeration. For Census purposes, total geographical area is broadly classified into Rural & Urban. The basic Unit of rural area is revenue Village (CENSUS OF INDIA, 2011). The rural data set is considered here for the State of West Bengal. The ranking of the districts have been calculated based on the individual values of each observations (Bhadra & Ghara 2020). Different approach has been adopted earlier to this to rank the district based on the dataset. Household quality of living (HQL) refers to three broad aspects including housing condition, amenities and assets. The study is an attempt to investigate regional variation of the districts of West Bengal in terms of HQL based on published 2011 Census data (Das and Mistri 2013). There are about 69% of the population lived in rural areas. States are compared using different methods based on household data (Panda 2014).

Here, an attempt is made to rank the districts of the state of West Bengal based on Rural data considering the Housing infrastructure, Availability of Household Amenities and Assets. Using the probability approach and then adopting the Simple Arithmetic mean, Simple Geometric mean, Weighted Arithmetic Mean and Weighted Geometric Mean the analytical process have been done, which is further tested by Correlation and Rank Test. The data analysis has been done by using SPSS.

DATA

Census of India 2011 is used for the purpose of the study. The entire exercise is done on the Rural data set of House listing & Housing Census which was conducted in 2010. The main variables are Housing infrastructure, Availability of Household Amenities and Assets. We consider here 11 such main variables and 78 sub variables. These 78 variables are actually sub variables of 11 main

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

variables. The data considered for rural West Bengal covering all districts except Kolkata. The breakup of those selected 11 main and 78 sub variables are as below. The variables are – *Census house* (residence, residence-cum-other use, shop/office, school/college/etc, hotel/ lodge/ guest house/etc, hospital/ dispensary/etc, factory/ workshop/ work shed/etc., place of worship, other non-residential use); *Condition of census houses* (good, livable, dilapidated); *Material of roof* (grass/thatch/bamboo/wood/mud/etc, plastic/polythene, handmade tiles, machine made tiles, burnt brick, stone/slate, GI/metal/ asbestos sheets, concrete, any other material); *Material of wall* (grass/ thatch/ bamboo/ wood/ mud/etc, plastic/ polythene, Mud/ Unburnt brick, Wood, Stone not packed with mortar, Stone packed with mortar, GI/metal/asbestos sheets, Burnt brick,concrete, any other material); *Material of floor* (mud, wood/bamboo, burnt brick, stone, cement, mosaic/floor tiles, any other material); *Availability of assets* (radio/transistor, television, computer/laptop , bicycle, scooter/ motorcycle/moped, none of the assets,); Here the variable Computer/laptop reflects the total of Computer/laptop with Internet and Without Internet. In Census database, there is another asset termed as availability of Car/ Jeep/Van. This is also precluded in the analysis because of the uses of this asset varies significantly as someone is using it as luxury goods and someone is for his daily earnings and for this reason it won't be very justified to use this variable in this study. *Main source of lighting* (electricity, kerosene, solar energy, other oil, any other, no lighting); *Type of latrine facility within the premises* (piped sewer system, septic tank, other system of flush latrine, with slab/ventilated improved pit, without slab/open pit of Pit Latrine, night soil disposed into open drain, night soil serviced by animal) *No latrine within premises* (public latrine, open latrine); *Number of households having bathing facility* within the premises (bathroom, enclosure without roof) and no bathroom; *Waste water outlet connected to* (closed drainage, open drainage, no drainage); *Cooking facility* (has kitchen-fire-wood, crop residue, cow dung cake, coal/lignite/charcoal, kerosene, lpg/png, does not have kitchen -fire-wood, crop residue, cow dung cake, coal/lignite/charcoal, kerosene, lpg/png, and no cooking).

The sub variables under each main variable are weighted according to their importance in social structure.

Table 1.1: showing the weightage scheme (shown in the last column in each table)

Uses of census houses (X1..)								
Residence	Residence -cum- other use	Shop/ Office.	School/ College etc.	Hotel/ Lodge/ Guest house etc	Hospital/ Dispensary etc..	Factory/ Workshop/ Workshed etc.	Place of worship.	Other non-residential use
4	3	5	6	7	8	9	2	1

*Number of census houses used as Schools, Colleges, etc were more than 15 lakhs (From Census 2001)

Condition if Census Houses X2..		
Good	Livable	Dilapidated
3	2	1

*10 million households in the houses which were in dilapidated condition. (From Census 2001)

*50% of all the households in India categorized the condition of census houses occupied by them as 'Good' (From Census 2001)

*44.3% categorized as 'Livable'. (From Census 2001)

Material of Roof X3..								
Grass/ Thatch/ Bamboo/ Wood/ Mud, etc..	Plastic/ Polythene.	Handmade Tiles.	Machine made Tiles.	Burnt Brick.	Stone/Slate.	G.I./ Metal/ Asbestos sheets.	Concrete.	Any other material.
2	3	4	6	7	8	5	9	1

42 million or 22 percent household had roof made of grass, thatch, mud etc. (From Census 2001)

Material of Wall X4..									
Grass/ Thatch/ Bamboo etc..	Plastic/ Polythene.	Mud/ Unburnt brick.	Wood.	Stone not packed with mortar.	Stone packed with mortar.	G.I./Metal/ Asbestos sheets.	Burnt brick.	Concrete.	Any other material.
2	3	4	5	8	9	6	7	10	1

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

*61 million households had wall made of mud and unburnt bricks. (From Census 2001)

Material of Floor						
Mud	Wood/ Bamboo	Burnt Brick	Stone	Cement	Mosaic/ Floor tiles	Any other material
2	3	4	6	5	7	1

*110 million or 57 percent households had Mud floors in the country (From Census 2001)

Availability of assets					
Radio/ Transist or	Televisi on	Computer/Lapt op	Bicycle	Scooter/ Motorcycle/M oped	None of the assets specified in col. 10 to 19
2	4	6	3	5	1

*6.62 crore out of 19.17 crore households in India did not possess any of the assets. In the beginning of this millennium 67 million [35.1%] households had radio/ transistor. In the beginning of this millennium 61 million households had television (From Census 2001)

Main Source of lighting					
Electricity	Kerosene	Solar energy	Other oil	Any other	No lighting
6	4	5	3	2	1

*In the rural the mainsourceswere kerosene (55.6%) and electricity (43.5%) (From Census2001)

Type of latrine facility within the premises							No latrine within premises		
Flush/pour flush latrine connected to			Pit latrine		Night soil disposed into open drain	Service Latrine		Alternative source	
Piped sewer system	Septic tank	Other system	With slab/ ventilated improved pit	Without slab/ open pit		Night soil serviced by animal	Public latrine	Open	
10	9	8	7	6	4	3	5	1	

*63.6% households did not have any latrine within the house (From Census2001).

Number of households having bathing facility within the premises		
Yes		No
Bathroom	Enclosure without roof	
3	2	1

*Only 6.93 crore out of 19.17 crore households in India had bathroom facility.(From Census2001)

Waste water outlet connected to		
Closed drainage	Open drainage	No drainage
3	2	1

*More than half the households in the country did not have any drainage connectivity for waste water outlet.(From Census2001)

Has Kitchen						Does not have kitchen						No kitchen
Fire- wood	Crop residue	Cow dung cake	Coal, Lignite, Charcoal	Ker ose ne	LP G/P NG	Fire- wood	Crop residue	Cow dung cake	Coal, Lignite, Charcoal	Ker ose ne	LPG /PN G	
5	6	7	9	11	13	2	3	4	8	10	12	1

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

MATERIAL AND METHODS

11 main variables have been considered for analytical purpose in this study. These 11 main variables are consisting of several sub variables. Very few sub variables are excluded due to lack of importance, insignificant small data values and lesser impact in 2020 corresponding to 2010.

Dataset obtained by excluding a very few sub variables in the above process and finally it consists of 78 variables.

$$x_{ijk} = \text{Value of } j^{\text{th}} \text{ subvariable of } i^{\text{th}} \text{ main variable for the } k^{\text{th}} \text{ district. } i = 1(1)11, \quad j = 1(1)n(i); \quad k = 1(1)18$$

We define for fixed i and k,

$$p_{ijk} = x_{ijk} / \sum_{j=1}^{n(i)} x_{ijk}; i=1(1)11 \text{ and } k=1(1)18., \quad n(i) = \text{number of sub variables in } i^{\text{th}} \text{ main variable.}$$

1st Method: Considering $w_{ijk} = \text{weightage of } p_{ijk} = 1/N$, where $N = \text{sum of } n(i) = 78$

$$\text{Index (ik)}[\text{I1(ik)}] = \frac{\sum P_{ijk}^{w_{ijk}}}{\sum P_{ijk}^{w_{ijk}} + \sum (1 - P_{ijk})^{w_{ijk}}}$$

2nd Method): Considering $w_{ijk} = \text{weightage of } p_{ijk}$

Weightage has been considered as shown above in Table 1.1. Here the weightage of variable varies unlike the first method.

$$\text{Index(ik)}[\text{I2(ik)}] = \frac{\sum P_{ijk}^{w_{ijk}}}{\sum P_{ijk}^{w_{ijk}} + \sum (1 - P_{ijk})^{w_{ijk}}}$$

3rd Method: Considering $w_{ijk} = \text{weightage of } p_{ijk} = 1/N$, where $N = 78$

$$\text{Index (ik)} [\text{I3(ik)}] = \frac{\prod P_{ijk}^{w_{ijk}}}{\prod P_{ijk}^{w_{ijk}} + \prod (1 - P_{ijk})^{w_{ijk}}}$$

4th Method: Considering $w_{ijk} = \text{weightage of } p_{ijk}$

Weightage has been considered as shown above in Table 1.1. Here the weightage of variable varies unlike the first method.

$$\text{Index(ik)}[\text{I4(ik)}] = \frac{\prod P_{ijk}^{w_{ijk}}}{\prod P_{ijk}^{w_{ijk}} + \prod (1 - P_{ijk})^{w_{ijk}}}$$

Table – 2.1 showing the ranks (R1) of the districts based on 1st method

Rank(R1)	District code	District name	Index Value(I1)
1	338	Hugli	0.63324
2	341	Haora	0.63089
3	335	Barddhaman	0.62953
4	344	PaschimMedinipur	0.62829
5	339	Bankura	0.62820
6	334	Birbhum	0.62589
7	336	Nadia	0.62568
8	337	North Twenty Four Parganas	0.62559
9	327	Darjiling	0.62534
10	345	PurbaMedinipur	0.62519
11	340	Puruliya	0.62512
12	343	South Twenty Four Parganas	0.62474
13	328	Jalpaiguri	0.62432
14	331	DakshinDinajpur	0.62429
15	333	Murshidabad	0.62367
16	330	Uttar Dinajpur	0.62352
17	332	Maldah	0.62275
18	329	Koch Bihar	0.62166

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

Table – 2.2 showing the ranks(R2) of the districts based on 2nd method

Ranking(R2)	District code	District name	Index value(I2)
1	338	Hugli	0.17622
2	327	Darjiling	0.16936
3	341	Haora	0.16813
4	335	Barddhaman	0.16754
5	337	North Twenty Four Parganas	0.15868
6	336	Nadia	0.15862
7	345	PurbaMedinipur	0.15272
8	339	Bankura	0.15219
9	344	PaschimMedinipur	0.1506
10	333	Murshidabad	0.14947
11	334	Birbhum	0.14809
12	343	South Twenty Four Parganas	0.14807
13	331	DakshinDinajpur	0.14739
14	328	Jalpaiguri	0.14724
15	332	Maldah	0.14575
16	329	Koch Bihar	0.14505
17	330	Uttar Dinajpur	0.14409
18	340	Puruliya	0.14289

Table – 2.3 showing the ranks (R3) of the districts based on 3rdmethod

Rank(R3)	District code	District name	Index value(I3)
1	337	North Twenty Four Parganas	0.04513
2	338	Hugli	0.04453
3	335	Barddhaman	0.04254
4	327	Darjiling	0.04249
5	336	Nadia	0.04104
6	341	Haora	0.04092
7	343	South Twenty Four Parganas	0.03844
8	345	PurbaMedinipur	0.03829
9	333	Murshidabad	0.03652
10	332	Maldah	0.03405
11	328	Jalpaiguri	0.03328
12	344	PaschimMedinipur	0.03297
13	334	Birbhum	0.03149
14	331	DakshinDinajpur	0.03140
15	339	Bankura	0.03026
16	330	Uttar Dinajpur	0.02933
17	340	Puruliya	0.02662
18	329	Koch Bihar	0.02653

Table – 2.4 showing the ranks(R4) of the districts based on 4thmethod

Rank(R4)	District code	District name	Index value(I4)
1	338	Hugli	8.52389E-14
2	327	Darjiling	4.97226E-14

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

3	337	North Twenty Four Parganas	3.15052E-14
4	335	Barddhaman	2.95542E-14
5	341	Haora	2.01082E-14
6	336	Nadia	1.13277E-14
7	345	PurbaMedinipur	2.51883E-15
8	343	South Twenty Four Parganas	2.30413E-15
9	333	Murshidabad	1.76082E-15
10	332	Maldah	6.70960E-16
11	328	Jalpaiguri	5.98242E-16
12	344	PaschimMedinipur	5.51777E-16
13	334	Birbhum	3.78713E-16
14	339	Bankura	3.28377E-16
15	331	DakshinDinajpur	2.40617E-16
16	330	Uttar Dinajpur	1.53489E-16
17	329	Koch Bihar	5.43194E-17
18	340	Puruliya	2.88899E-17

Table 2.5 showing final table of ranking

Code	District name	R1	R2	R3	R4
339	Bankura	5	8	15	14
335	Barddhaman	3	4	3	4
334	Birbhum	6	11	13	13
331	DakshinDinajpur	14	13	14	15
327	Darjiling	9	2	4	2
341	Haora	2	3	6	5
338	Hugli	1	1	2	1
328	Jalpaiguri	13	14	11	11
329	Koch Bihar	18	16	18	17
332	Maldah	17	15	10	10
333	Murshidabad	15	10	9	9
336	Nadia	7	6	5	6
337	North Twenty Four Parganas	8	5	1	3
344	PaschimMedinipur	4	9	12	12
345	PurbaMedinipur	10	7	8	7
340	Puruliya	11	18	17	18
343	South Twenty Four Parganas	12	12	7	8
330	Uttar Dinajpur	16	17	16	16

Table 2.6 Correlation between ranks

Correlations					
		R1	R2	R3	R4
R1	Pearson Correlation	1	.981**	.523*	.833**
	Sig. (2-tailed)		0.000	0.026	0.000
R2	Pearson Correlation	.981**	1	.542*	.882**
	Sig. (2-tailed)	0.000		0.020	0.000
R3	Pearson Correlation	.523*	.542*	1	.779**

Ranking of Districts of West Bengal by an Index Developed Through Probabilistic Approach Based on Household Data for Rural Areas of Census 2011

	Sig. (2-tailed)	0.026	0.020		0.000
R4	Pearson Correlation	.833**	.882**	.779**	1
	Sig. (2-tailed)	0.000	0.000	0.000	

REMARKS

An attempt has been made to use four different approaches for finding ranks based on House listing and Housing Census Data for rural areas. The approaches are simple average, weighted average, simple geometric mean, weighted geometric mean. Correlation coefficient shows that the methods are not significantly different (Table 2.6). From the Table 2.5, it is observed that the optimal district is either Hugli or Koch Bihar. The next better district is Barddhaman, Haora, etc. The deviation of rankings is maximum for the District Bankura, Paschim Medinipur, Malda, Birbhum, etc.

ACKNOWLEDGEMENT

Authors want to acknowledge - Office of the Registrar General, India and Director of Census Operations, West Bengal.

Disclaimer: The findings of this paper is the view of the Authors. It has nothing to do to any Organization.

REFERENCE

- 1) Bhadra P P and Ghara T K (2020): Ranking Of Districts Of West Bengal By An Index Developed Through Household Data For Rural Areas Of Census 2011 (accepted IJRSR)
- 2) Bhadra P P and Ghara T K (2020): Ranking of Districts of West Bengal by an Index Developed Through Census 2011 Household Data (Urban & Total), International Journal of Humanities and Social Science Invention (IJHSSI), Vol-9 No-7 Pp31-36
- 3) Census 2011: Census of India 2011 Provisional Population Totals Urban Agglomerations and Cities : www.censusindia.gov.in/2011-prov-results/paper2/data_files/India2/1. Data Highlight.pdf
- 4) Das, B and Mistri, A (2013): Household Quality of Living in Indian States: Analysis of 2011, Census: <https://doi.org/10.1177/0975425313477759>
- 5) Directorate of Census Operators, West Bengal.
- 6) Ghara, T K(2016): Classification of the States of India Based on Higher Education Development Indicators, IOSR Journal of Research & Method in Education, Volume 6, Issue 6 Ver. I, PP 65-70
- 7) Jian L, Saha B and Deshpande A(2009) : A Unified Approach to Ranking in Probabilistic Databases
- 8) Leclerc G, Nelson A and Knapp E B(2000): The use of unit-level census data for research on poverty: a multiscale approach; International Center for Tropical Agriculture (CIAT), Colombia