

Classification of the States of India Based on Higher Education Development Indicators

Tushar Kanti Ghara¹

¹*Joint Director of Public Instruction & State Nodal Officer, AISHE Government of West Bengal Bikash Bhavan, Salt Lake City, Kolkata 700091*

Abstract: *There are development indicators for elementary education as developed by NEUPA. There is no such for higher education in relation to India. All India Survey on Higher Education (AISHE) started in 2010-2011 on annual basis. Based on AISHE database, a number of factors has been taken as indicators for higher education development. Using cluster analysis, an attempt has been made to classify the states on India into different homogeneous clusters. Thus, a ranking of the states of India has also been made.*

Keywords: *Development Indicators for Higher Education, Classification, Cluster Analysis*

I. Introduction

The educational planners, researchers and administrators may frequently need to evaluate various indicators of education using the data. The Education for all Development Index (EDI) (framed by UNESCO) used 4 goals – Universal education, Adult Literacy, Quality of education and Gender. The proxies used to quantify EDI are – (1) adjusted net enrolment ratio (ANER), (2) adult literacy rate, (3) quality of education (may be inters of retention rate, drop-out rate, success rate, etc.), (4) gender equality (gender parity index – access, drop-out, success)¹. The main aim is to choose indicators for quantifying development in education. In literature, the indicators used are – (i) percentage of population in the age group to the total population, (ii) apparent intake rate, (iii) net intake rate, (iv) gross enrolment ratio, (v) net enrolment ratio, (vi) age-specific enrolment ratio, (vii) percentage of private enrolment, (viii) enrolment by gender and social group/s, (ix) percentage of girls/female enrolment, (x) percentage of teachers by gender and social group, (xi) student class-room ratio, (xii) percentage of institutions with/without toilet, (xiii) percentage of institutions with furniture, (xiv) percentage of institutions with medium as mother tongue, (xv) percentage of institutions with/without library facilities, (xvi) percentage of trained teachers, (xvii) pupil teacher ratio, (xviii) expenditure, (xix) transition rate, (xx) percentage of repeaters, (xxi) repetition rate, (xxii) survival rate by grade, (xxiii) coefficient of efficiency, (xxiv) percentage of under-aged and over-aged students, (xxv) percentage of teachers in private institutions, etc.. To portray the status of higher education in the country, Ministry of Human Resource Development has endeavoured to conduct an annual web-based effort All India Survey on Higher Education (AISHE) since 2010-11. The survey covers all the institutions in the country engaged in imparting of higher education. Data is being collected on several parameters such as teachers, student enrolment, programmes, examination results, education finance and infrastructure. Indicators of educational development such as Institution Density, Gross Enrolment Ratio, Pupil-teacher ratio, Gender Parity Index, Per Student Expenditure will also be calculated from the data collected through AISHE. These are useful in making informed policy decisions and research for development of education sector. The AISHE is now an annual event. Based on AISHE database, in this paper, it is being tried to quantify the development in higher education by framing an indicator and attempt is being made to rank the States of India. Cluster Analysis has been used to choose a smaller set of variables objectively. Based on the variables selected, a classification of the states on India has been made (using Statistica Software ver 20 trial). Thus ranking of the States has been made.

II. Data

AISHE has been taken as main source of data². The following variables are being used for indexing the higher education growth in 2014-15. The AISHE survey for 2014-15 was closed on 31.03.2016. The reports were considered from the portal www.aishe.gov.in. The author being APEX user of the portal (as approved by MHRD), the state-wise reports downloaded and put into the common database. 179 sheets are considered and the following variables(X##) are being considered – (i) proportion of females (female enrolment / male enrolment)(X₁); (ii) Pupil Teacher Ratio(X₂); (iii) percentage of filled up teachers(X₃); (iv) percentage of filled up non-teaching(X₄); (v) Staff per student(X₅); (vi) percentage of students getting education loan (X₆); (vii) percentage of students getting scholarship (X₇); (viii) percentage of private institutions (X₈); (ix) percentage of institutions (colleges) having NAAC accreditation (active)(X₉); (x) percentage of institutions (universities) having NAAC accreditation (active)(X₁₀); (xi) number of books per student (X₁₁); (xii) percentage of institutions with running cafeteria (X₁₂); (xiii) percentage of enrolment in Government & Government-aided institutions

(X₁₃); (xiv) population in the age group 18-23 (1bd) years (X₁₄); (xv) percentage of enrolment in distance mode (X₁₅) and (xvi) average expenditure per student (X₁₆). Except the population figures, all other variables are calculated based on the reports from AISHE portal. The population figures are taken from a report of MHRD.

Analysis

X_{ij} is the value of X_i (ith variable) corresponding to jth state; i=1(1)16(p) and j=1(1)36.

It is to note that the variables are in different scales and units. To make variables comparable, a standardised transformation is used as follows –

For fixed j, $Z_{ij} = (X_{ij} - \text{minimum value}) / (\text{Maximum value} - \text{minimum value})$ for each of the variavles, i=1(1)16 and j=1(1)36.

PCA is used on Z_{ij}'s to determine weights for each variables and each state. The factor weights(F1 and F2) are being used as weights for variavles.

The total score(Total Raw Score) may be sum of Z-values for selected the variables. i.e. for fixed j,

$\sum Z_{ij}$, the index for jth state.

The jth State index SI1_j = $\sum_{i=0}^p \{W1_{ij}Z_{ij}\}$

Where Z_{ij} = the transformed score for ith variable and jth State

$W1_{ij}$ = the weight for ith variable as contribution corresponding to (F1) (obtained using PCA) and jth State

i = 1,2, ..., 16(p)

j = j th State – Andaman & N Island,, West Bengal

The jth State index SI2_j = $\sum_{i=0}^p \{W2_{ij}Z_{ij}\}$

Where Z_{ij} = the transformed score for ith variable and jth State

$W2_{ij}$ = the weight for ith variable as contribution corresponding to (F2)(obtained using PCA) and jth State

i = 1,2, ..., 16(p)

j = j th State – Andaman & N Island,, West Bengal

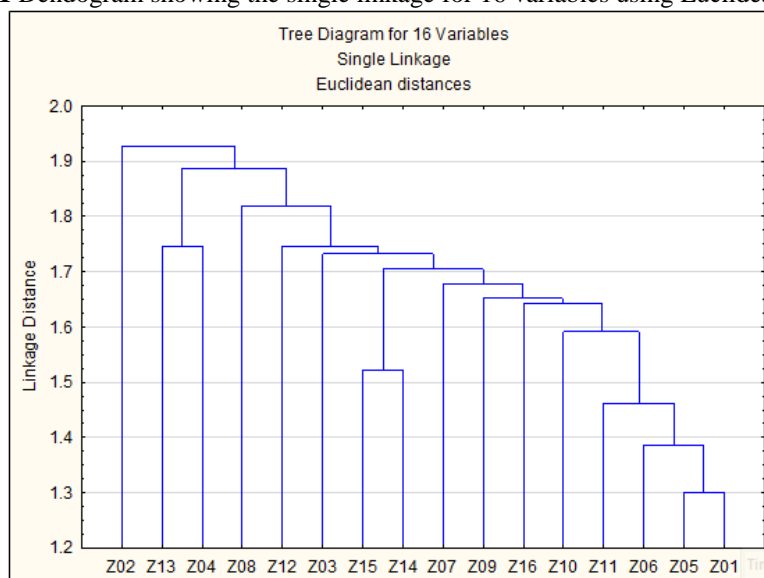
Table – 1.1 showing Z-values for 1st 9 variables for all states

STATE(j)	Z ₀₁	Z ₀₂	Z ₀₃	Z ₀₄	Z ₀₅	Z ₀₆	Z ₀₇	Z ₀₈	Z ₀₉
Andaman & Nicobar Island	0.0197	0.5665	0.0857	0.6701	0.0406	0.0000	0.3092	0.0000	0.0000
Andhra Pradesh	0.0697	0.1278	0.6913	0.1838	0.2700	0.0996	0.9190	0.9776	0.1543
Arunachal Pradesh	0.1046	0.7355	0.2916	0.7318	0.0588	0.0000	0.1989	0.3701	0.7808
Assam	0.1632	0.3564	0.3161	0.7485	0.1148	0.0400	0.1824	0.1305	0.6026
Bihar	0.0302	1.0000	0.3211	0.6273	0.0000	0.0138	0.0819	0.1664	0.1841
Chandigarh	0.2013	0.4855	0.0777	0.6222	0.2498	0.0017	0.0447	0.1814	0.7540
Chhatisgarh	0.1604	0.2947	0.3081	0.4291	0.1563	0.1655	0.7559	0.5464	0.1522
Dadra & N H	0.1280	0.4028	0.7012	0.7774	0.1138	0.0000	0.0000	0.6853	0.0000
Daman & Diu	0.0000	0.1500	0.0600	0.6224	0.3299	0.0000	0.1813	0.5483	0.1813
Delhi	0.0768	0.9834	0.5299	0.4695	0.0156	0.0101	0.0339	0.4771	0.3066
Goa	0.2012	0.1857	0.1806	0.7793	0.4666	0.0902	0.2287	0.2847	0.6591
Gujarat	0.0298	0.3871	0.0956	0.1102	0.1039	0.0468	0.4570	0.7127	0.3138
Haryana	0.1105	0.1979	0.3416	0.5998	0.2609	0.0550	0.3042	0.8224	0.5397
HP	0.2310	0.2776	0.2872	0.6612	0.2475	0.0879	0.1681	0.5745	0.2267
Jharkhand	0.1276	0.9465	0.5376	0.6187	0.0236	0.0586	0.1778	0.4911	0.1540
JK	0.1641	0.4562	0.2593	0.8020	0.1535	0.0265	0.0952	0.4905	0.3427
Karnataka	0.1611	0.0684	0.9395	0.4355	0.3586	0.5386	0.5214	0.8244	0.2735
Kerala	0.3718	0.1343	0.5552	0.7173	0.2873	1.0000	0.2595	0.7445	0.2453
Lakshadweep	1.0000	0.0600	0.1587	0.8448	0.3440	0.0000	0.8877	0.0000	0.0000
Madhya Pradesh	0.0470	0.3139	0.4528	0.2025	0.1617	0.0524	0.6279	0.7318	0.1277
Maharashtra	0.0675	0.3511	0.3565	0.2904	0.1545	0.1561	0.6869	0.7530	0.4356
Manipur	0.1772	0.3071	0.3233	0.7515	0.1021	0.0170	0.2400	0.2578	0.4143
Meghalaya	0.1854	0.3469	1.0000	0.7213	0.1208	0.0427	0.6788	0.3190	0.3867
Mizoram	0.0779	0.2086	0.1094	0.7696	0.2104	0.0032	0.1435	0.1050	1.0000
Nagaland	0.1678	0.1588	0.5603	0.8448	0.3264	0.0000	0.4700	0.2372	0.2762
Odisha	0.0901	0.2551	0.4090	0.3204	0.1936	0.1979	0.3522	0.4009	0.3212
Puducherry	0.1540	0.0000	0.6597	0.2230	1.0001	0.5217	0.2807	0.6294	0.3006
Punjab	0.1511	0.1350	0.0000	0.4803	0.3141	0.1559	0.5148	0.7585	0.3644
Rajasthan	0.0752	0.3861	0.5965	0.3256	0.1169	0.0368	0.2345	0.8887	0.1342
Sikkim	0.0488	0.2849	0.8269	0.7604	0.1463	0.0408	0.0561	0.3290	0.1318
Tamil Nadu	0.1501	0.1278	0.6904	0.3254	0.3045	0.8194	0.5849	0.9350	0.4081
Telangana	0.1136	0.1114	0.6954	0.0000	0.2676	0.0628	1.0000	1.0000	0.1134
Tripura	0.0252	0.4631	0.3315	1.0000	0.1889	0.2159	0.4825	0.1645	0.2610
Uttar Pradesh	0.1516	0.5840	0.6969	0.4743	0.0455	0.0195	0.3709	0.9135	0.1339
Uttrakhand	0.1308	0.2744	0.3230	0.2241	0.2429	0.0230	0.1668	0.4945	0.2190
West Bengal	0.1074	0.5908	0.3191	0.6659	0.0278	0.1169	0.2681	0.5100	0.3949

Table – 1.2 showing Z-values for last 7 variables for all states

STATE(j)	Z ₁₀	Z ₁₁	Z ₁₂	Z ₁₃	Z ₁₄	Z ₁₅	Z ₁₆
Andaman & Nicobar Island	0.0000	0.5282	0.6250	0.6115	0.0016	0.0074	0.00108
Andhra Pradesh	0.5893	0.5216	0.5219	0.0695	0.2324	0.4129	0.00000
Arunachal Pradesh	0.5000	0.0585	0.2917	0.9604	0.0064	0.0204	0.00014
Assam	0.3947	0.4528	0.3871	0.9455	0.1512	0.1237	0.00001
Bihar	0.4091	0.1281	0.2620	0.9489	0.4525	0.1707	0.00000
Chandigarh	0.0000	0.2473	0.7188	0.2520	0.0066	0.0457	0.00042
Chhatisgarh	0.2727	0.4335	0.3226	0.5736	0.1257	0.0936	0.00001
Dadra & N H	0.0000	0.2419	0.7500	0.5640	0.0021	0.0000	0.00078
Daman & Diu	0.0000	0.0000	0.3214	1.0000	0.0019	0.0000	1.00000
Delhi	0.7500	0.1512	0.7841	0.2327	0.0906	1.0000	0.00003
Goa	0.7500	0.7114	0.7212	0.3830	0.0066	0.0114	0.00032
Gujarat	0.2895	0.3173	0.4642	0.2310	0.2947	0.1247	0.00020
Haryana	0.3529	0.4673	0.6550	0.3805	0.1308	0.0495	0.00001
HP	0.1200	0.2132	0.5107	0.6357	0.0307	0.0524	0.00003
Jharkhand	0.2143	0.1875	0.4458	0.8833	0.1525	0.0454	0.00001
JK	0.4500	0.2524	0.5018	0.7563	0.0558	0.1150	0.00001
Karnataka	0.6300	1.0001	0.4459	0.3049	0.2957	0.1114	0.00001
Kerala	0.5250	0.5746	0.7206	0.1243	0.1265	0.2792	0.00001
Lakshadweep	0.0000	0.1889	0.0000	0.0000	0.0000	0.0000	0.00060
Madhya Pradesh	0.3571	0.4576	0.3957	0.4637	0.3589	0.2896	0.00000
Maharashtra	0.9667	0.4579	0.3909	0.2045	0.5501	0.9865	0.00000
Manipur	0.5000	0.3474	0.3384	0.9274	0.0118	0.0095	0.00009
Meghalaya	0.2500	0.2479	0.6689	0.2856	0.0140	0.0138	0.00131
Mizoram	0.5000	0.4890	0.4655	0.8453	0.0051	0.0101	0.00045
Nagaland	0.3750	0.4027	0.5287	0.3247	0.0100	0.0095	0.00010
Odisha	0.5714	0.4192	0.2448	0.3538	0.1919	0.0620	0.00001
Puducherry	0.7500	0.5707	0.6576	0.3756	0.0057	0.0111	0.00020
Punjab	0.3600	0.4314	0.8578	0.2613	0.1351	0.0591	0.00001
Rajasthan	0.3214	0.3439	0.6433	0.4227	0.3534	0.1527	0.00000
Sikkim	0.3000	0.2674	1.0000	0.3013	0.0030	0.0192	0.00039
Tamil Nadu	1.0000	0.4798	0.5590	0.1550	0.3049	0.7404	0.00000
Telangana	0.9167	0.3002	0.5394	0.1136	0.1683	0.1048	0.00000
Tripura	0.5000	0.3053	0.3913	0.9375	0.0179	0.0166	0.00022
Uttar Pradesh	0.6048	0.1830	0.4437	0.1460	1.0000	0.1124	0.00000
Uttarakhand	0.3214	0.2960	0.5900	0.4331	0.0502	0.0462	0.00003
West Bengal	0.3529	0.2706	0.5054	0.6949	0.4493	0.2949	0.00000

Figure 1.1 Dendrogram showing the single linkage for 16 variables using Euclidean distances



It is observed that Z02 dominates all others. We may take any section based on clusters. Z02, Z13, Z04 & Z08 are upper cluster as compared to other variables. We are considering these 4 variables as most favoured indicators as development indicators for higher education in India. By considering the 4 indicators, the Principal Component Analysis (PCA) eliminates 100% variability of data.

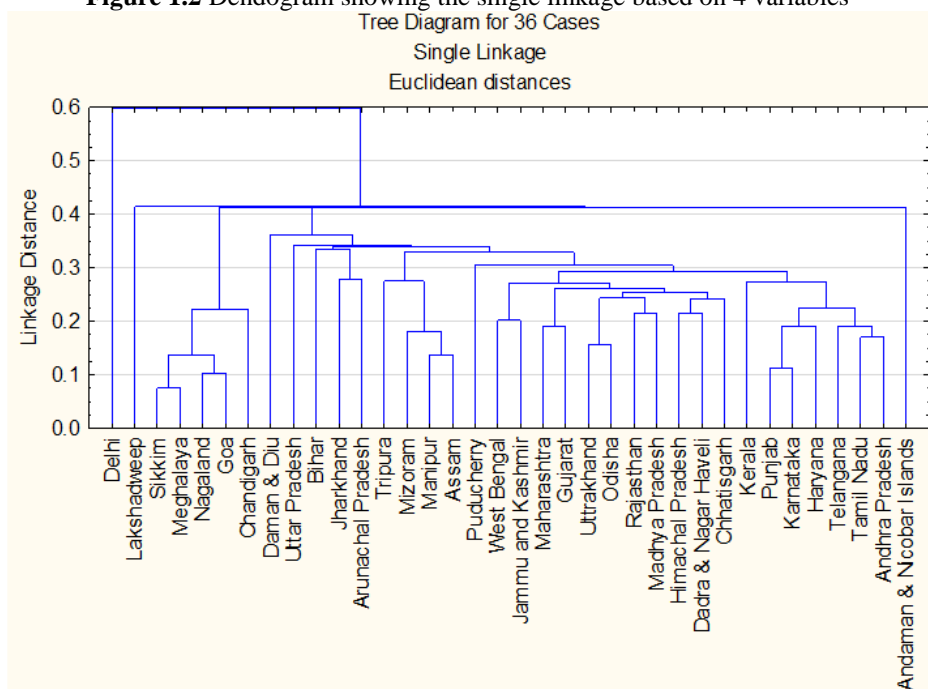
Table 1.2 showing the Eigen Values of correlation matrix (PCA)

Eigen Value	% Total Variance	Cumulative Eigen Value	Cumulative %
2.279279	56.98198	2.279279	56.9820
0.916403	22.91007	3.195682	79.8920
0.500620	12.51550	3.696302	92.4075
0.303698	7.59246	4.000000	100.0000

Table 1.3 showing the variable contributions based on correlations (PCA)

Variable	Factor 1	Factor 2	Factor 3	Factor 4
Z02	0.132329	0.642697	0.208669	0.016305
Z13	0.267983	0.060148	0.660298	0.011571
Z04	0.293951	0.183875	0.014807	0.507366
Z08	0.305737	0.113280	0.116227	0.464757

Figure 1.2 Dendrogram showing the single linkage based on 4 variables



It is observed that Delhi and all other states are two clusters. On further observations, it is true that there are groups/clusters among states almost on same level but differently. The clusters are – Sikkim-Meghalaya-Nagaland-Goa, Jharkhand-Arunachal Pradesh-Tripura-Mizoram-Manipur-Assam, West Bengal-Jammu & Kashmir-Maharashtra-Gujarat-Uttarakhand-Odisha-Rajasthan-Madhya Pradesh-Chhatisgarh, Kerala-Punjab-Haryana-Telangana-Tamil Nadu-Andhra Pradesh.

Table 1.4 showing Total Scores and ranks for the States

STATE	Z02	Z13	Z04	Z08	Tot_Raw	Rank	CPCA1	RPCA1	CPCA2	RPCA2
Andaman & Nicobar Island	0.5665	0.6115	0.6701	0.0000	1.8481	17	0.4358	25	0.5241	9
Andhra Pradesh	0.1278	0.0695	0.1838	0.9776	1.3587	32	0.3885	29	0.2308	33
Ar Pradesh	0.7355	0.9604	0.7318	0.3701	2.7978	2	0.6830	2	0.7069	4
Assam	0.3564	0.9455	0.7485	0.1305	2.1809	10	0.5605	11	0.4384	12
Bihar	1.0000	0.9489	0.6273	0.1664	2.7426	3	0.6219	7	0.8340	1
Chandigarh	0.4855	0.2520	0.6222	0.1814	1.5411	29	0.3701	30	0.4621	11
Chhatisgarh	0.2947	0.5736	0.4291	0.5464	1.8438	18	0.4859	17	0.3647	19
Dadra & N Haveli	0.4028	0.5640	0.7774	0.6853	2.4296	7	0.6425	5	0.5134	10
Daman & Diu	0.1500	1.0000	0.6224	0.5483	2.3208	8	0.6384	6	0.3331	24
Delhi	0.9834	0.2327	0.4695	0.4771	2.1626	11	0.4764	18	0.7864	3
Goa	0.1857	0.3830	0.7793	0.2847	1.6327	25	0.4433	22	0.3179	25
Gujarat	0.3871	0.2310	0.1102	0.7127	1.4410	30	0.3634	32	0.3637	20
Haryana	0.1979	0.3805	0.5998	0.8224	2.0007	15	0.5559	12	0.3536	21
HP	0.2776	0.6357	0.6612	0.5745	2.1489	12	0.5771	10	0.4033	16
Jharkhand	0.9465	0.8833	0.6187	0.4911	2.9396	1	0.6940	1	0.8308	2
JK	0.4562	0.7563	0.8020	0.4905	2.5051	5	0.6488	4	0.5417	8

Classification of the States of India Based on Higher Education Development Indicators

Karnataka	0.0684	0.3049	0.4355	0.8244	1.6332	24	0.4708	19	0.2358	32
Kerala	0.1343	0.1243	0.7173	0.7445	1.7204	19	0.4895	16	0.3100	26
Lakshadweep	0.0600	0.0000	0.8448	0.0000	0.9048	36	0.2563	36	0.1939	34
Madhya Pradesh	0.3139	0.4637	0.2025	0.7318	1.7119	20	0.4491	21	0.3498	22
Maharashtra	0.3511	0.2045	0.2904	0.7530	1.5991	26	0.4169	28	0.3767	18
Manipur	0.3071	0.9274	0.7515	0.2578	2.2437	9	0.5889	9	0.4205	14
Meghalaya	0.3469	0.2856	0.7213	0.3190	1.6728	22	0.4320	26	0.4089	15
Mizoram	0.2086	0.8453	0.7696	0.1050	1.9286	16	0.5125	15	0.3383	23
Nagaland	0.1588	0.3247	0.8448	0.2372	1.5655	27	0.4289	27	0.3038	27
Odisha	0.2551	0.3538	0.3204	0.4009	1.3302	33	0.3453	35	0.2896	29
Puducherry	0.0000	0.3756	0.2230	0.6294	1.2280	34	0.3586	33	0.1349	36
Punjab	0.1350	0.2613	0.4803	0.7585	1.6351	23	0.4610	20	0.2767	30
Rajasthan	0.3861	0.4227	0.3256	0.8887	2.0231	14	0.5318	14	0.4341	13
Sikkim	0.2849	0.3013	0.7604	0.3290	1.6756	21	0.4425	23	0.3783	17
Tamil Nadu	0.1278	0.1550	0.3254	0.9350	1.5431	28	0.4400	24	0.2572	31
Telangana	0.1114	0.1136	0.0000	1.0000	1.2250	35	0.3509	34	0.1917	35
Tripura	0.4631	0.9375	1.0000	0.1645	2.5650	4	0.6567	3	0.5565	7
Uttar Pradesh	0.5840	0.1460	0.4743	0.9135	2.1178	13	0.5351	13	0.5748	6
Uttarakhand	0.2744	0.4331	0.2241	0.4945	1.4261	31	0.3694	31	0.2996	28
West Bengal	0.5908	0.6950	0.6659	0.5100	2.4618	6	0.6161	8	0.6017	5

Rank, RPCA1 & RPCA2 are the ranks of the states based on Total Raw Score, SI1 & SI2 respectively. It is interesting to note that the ranks for the states do not differ markedly from Total Raw Score to SI1 (correlation is about 0.96).

Remarks

Using PCA and Cluster Analysis, the education development indicators have been chosen. The weighted average of selected variable's standard scores has been calculated. Based on the education development indicators for higher education, the states have been classified and ranked. West Bengal has the rank 8. It is an attempt to classify the states of India based on higher education development indicators. Each year rating may be calculated based on the data for the corresponding year.

Reference

- [1] 2015 Report Education Development Index : http://en.unesco.org/gem-report/sites/.../2015Report_EDI2012.pdf
- [2] All India Survey on Higher Education, MHRD, Govt. on India: www.aishe.gov.in
- [3] Bhandari, P(2012), Refining State Level Comparisons in India, Working Paper Series, Planning Commission, India
- [4] Educational Statistics at a glance(2014), MHRD, Government of India
- [5] Everitt, Brian (2011). Cluster analysis. Chichester, West Sussex, U.K: Wiley
- [6] Global Monitoring Report (2006), Planning Commission of India, Govt. of India
- [7] Global Education Monitoring Report(2015), The Education for All Development Index
- [8] OECD Report(2012), How is the global talent pool changing ?
- [9] Mehta A C, Indicators of Educational Development with focus on elementary education : Concept and Definitions
- [10] Rencher A C(2013), Methods of Multivariate Analysis, 2nd Edition, Wiley
- [11] Sarkar, D and Jhingran, D (2012), Educational Development Index, Working Paper Series, MHRD, Govt. of India