

## Assessment Students' Performance through Principal Component Analysis and Factor Analysis: A Case Study

Debasmita Dutta, Rabi Sankar Pandey<sup>1</sup>, Sumanta Ray<sup>2</sup>, Sanat Kumar Mahato<sup>3</sup>

<sup>1</sup> Research Scholar, Department of Mathematics, SKBU, email: [rabi.pandey@gmail.com](mailto:rabi.pandey@gmail.com)

<sup>2</sup> Assistant Professor (Contractual), Department of Computer Science, SKBU, email: [raysumanta1@gmail.com](mailto:raysumanta1@gmail.com)

<sup>3</sup> Professor, Department of Mathematics, SKBU, email: [sanatkmahato@gmail.com](mailto:sanatkmahato@gmail.com)

Received: 05.02.2024; revised: 26.04.2024; accepted: 27.04.2024; published online: April, 2024

### Abstract

This research paper is anticipated to focus on the assessment of class 10 level students using Principal Component Analysis and Factor Analysis. Society is always expecting a high level quality of services from all the Educational Institutions. In West Bengal, there are so many Government schools, Government sponsored schools, Self-financed schools. All the schools has most dynamic curriculum. The study has been conducted to improve the academic performance of class 10 level students. Initially, we have selected one Madhyamik level school from Burdwan district in West Bengal to collect the data. Nine parameters associated with class 10 level students has been chosen. Assessment has been done using Principal component Analysis and Factor Analysis. Principal Component Analysis and Factor analysis, both are the examples of unsupervised learning method. They are also most widely used data analysis technique. In this paper, Principal Component Analysis and Factor Analysis both are used to analyse the students' parameters. A new assessment technique has also been introduced in this research paper.

**Keywords:** PCA, FA, Academic Performance, Students' Assessment, Parameters

### Introduction

In this present scenario, class 10 level students mainly evaluated based on their performance in the examination. (pay main attention in the curriculum and pay less attention in co-curriculum and extra-curriculum activity) Here, in this work one new method has been introduced to assess class 10 level students in a different way. Therefore, this paper introduced a more comprehensive study for the assessment of class 10 level students. Principal Component Analysis and Factor Analysis, both allows us to find out the hidden pattern behind the data set. By applying Principal Component

Analysis and Factor Analysis to the data set, we can increase a deeper perception into the various features that impact class 10 level students' academic achievements. By applying PCA and FA methods, the factors and more important parameters associated with class 10 level students has been identified.

### Data Collection

Data has been collected from Belar-Bhurkunda High School, a Madhyamik school in Burdwan district of West Bengal. Details of 25 students in 2011 Madhyamik batch and 25 students of 2012 Madhyamik batch has been selected randomly. Nine parameters has been chosen, they are (x1) Attendance in School, (x2) Discipline in school, (x3) Result in different class test, (x4) Result in different surprise test, (x5) Financial background, (x6) Communication skill, (x7) Comprehensive skill, (x8) Professional skill, (x9) Dress code and the Total Marks.

### Methodology:

Principal component analysis (PCA) is one example of Unsupervised Learning Method. Although it is oldest technique, but most widely used multivariate data analysis. It was first introduced by Pearson in 1901, and developed by Hotelling in 1933. It is applied in economic and statistical research. It is a method for dimension reduction and feature extraction. By applying Principal Component Analysis, the weights of Principal Components are automatically generated. Therefore, Principal Component Analysis is a more powerful tool for feature extraction.

Factor Analysis is also most widely used statistical tool for multivariate data analysis. By applying Factor Analysis, we will be able to find the factors from a huge number of variables. It has some unique features. Factor Analysis is performed on continuous variables and categorical variables. Generally, Factor Analysis has two types, EFA (Exploratory Factor Analysis) and CFA (Confirmatory Factor Analysis).

## Assessment Students' Performance through Principal Component Analysis and Factor Analysis: A Case Study

Principal Component Analysis and Factor Analysis has implemented by using R programming language. R is open source programming language used for data analysis. It is also the leading tool for Machine Learning, Statistics etc.

**Table 1.**

Index	X1	X2	X3	X4	X5	X6	X7	X8	X9	Total Marks in class 10 Examination (800)
1	9	8	9	6	5	6	7	8	7	555
2	9	9	9	7	4	9	9	9	8	480
3	8	8	8	6	8	6	8	7	7	434
4	9	9	8	7	9	7	7	7	7	480
5	7	9	8	7	7	7	7	7	9	480
6	8	8	8	6	9	7	7	8	7	320
7	7	8	8	7	4	7	8	7	8	375
8	8	7	7	6	7	6	7	8	7	359
9	8	8	7	7	5	8	7	8	7	301
10	7	7	8	6	5	7	8	7	8	327
11	8	8	7	7	4	7	8	7	7	311
12	8	7	8	6	4	8	7	8	7	274
13	8	8	7	7	4	7	7	8	8	324
14	7	7	8	6	7	6	8	7	8	270
15	8	8	7	7	4	7	7	8	8	273
16	9	9	9	9	8	9	9	9	9	715
17	9	9	9	9	7	8	8	9	9	633
18	9	8	8	8	8	8	8	8	9	630
19	8	8	8	8	7	7	8	8	7	491
20	8	8	7	7	6	8	8	7	8	482
21	8	7	7	8	6	7	8	8	7	500
22	7	7	7	6	7	7	7	7	7	338
23	7	7	7	7	4	8	8	7	7	392
24	8	8	8	7	8	8	8	7	8	387
25	7	7	7	8	7	7	7	8	7	305
26	8	7	8	8	4	8	8	9	9	282
27	8	9	9	9	8	8	9	9	9	493
28	8	8	8	8	7	7	8	8	8	348
29	9	9	9	9	4	9	9	9	9	420
30	9	9	8	9	4	8	8	8	8	420
31	8	9	9	9	9	9	8	7	8	522
32	8	8	8	7	4	7	7	7	7	261
33	8	8	8	8	7	8	8	8	8	377
34	8	8	8	8	5	8	8	8	8	376
35	8	7	7	8	7	8	8	7	8	236
36	9	9	9	9	4	8	8	8	8	233
37	8	8	7	7	4	7	7	8	7	324
38	8	8	8	8	4	8	8	7	7	211
39	9	9	9	9	7	8	8	8	8	379
40	7	7	8	7	4	8	8	7	7	233
41	9	9	9	9	9	9	9	9	9	422
42	9	9	8	8	8	8	8	8	8	229
43	8	8	7	7	4	8	8	8	8	276
44	8	8	8	8	8	8	8	8	8	377
45	8	7	7	7	7	7	8	8	8	270
46	9	9	9	9	6	8	9	9	9	383
47	8	8	8	8	4	8	8	7	7	235
48	9	9	9	9	4	8	8	9	9	360
49	8	8	8	8	7	8	8	7	7	238
50	7	7	7	7	4	7	7	7	7	208

Step 2: Correlation matrix of Table-1 has been calculated and stored in Table-2.

**Table-2:**

INDEX	C1	C2	C3	C4	C5	C6	C7	C8	C9
1	1.0000	0.6156731	0.5452753	0.3726780	0.21305422	0.2740243	0.23473428	0.61885257	0.1976934
2	0.6156731	1.0000	0.6008521	0.4402603	0.23497170	0.2946789	0.26289626	0.32152065	0.5247148
3	0.5452753	0.6008521	1.0000	0.1936492	0.19038649	0.1545288	0.38809117	0.43773277	0.4824939
4	0.3726780	0.4402603	0.1936492	1.0000	0.16385860	0.4274910	0.4294875	0.52164053	0.4822428
5	0.2130542	0.2349717	0.1903865	0.1638586	1.0000	-0.1662474	-0.09695267	-0.01709506	0.1488203
6	0.2740243	0.2946789	0.1545288	0.4274910	-0.1662474	1.0000	0.51648711	0.47052287	0.2803701
7	0.2347342	0.2628962	0.3880912	0.4294875	-0.09695267	0.5164871	1.0000	0.25762056	0.4107455
8	0.6188525	0.3215206	0.4377328	0.5216405	-0.01709506	0.4705229	0.25762056	1.0000	0.4107455
9	0.1976934	0.5247148	0.4824939	0.4822428	0.1488203	0.2803701	0.4107455	0.21696826	1.0000

Step 2: The Eigenvalues, Variance Contribution rate, % of variance, cumulative % of variance of Table-1 has been fitted out in Table-3.

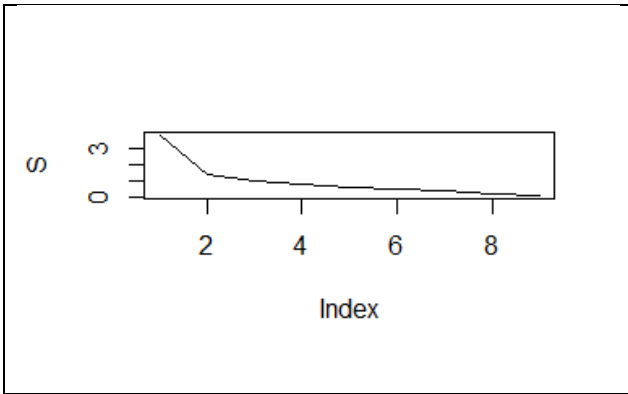
**Table-3:**

Eigenvalues	Variation Contribution Rate
3.7831200	0.4203466667
1.4389563	0.1598840333
1.0522145	0.1169127222
0.8611653	0.0956850333
0.5842439	0.0649159889
0.5089683	0.0565520333
0.4177129	0.0464125444
0.2314381	0.0257153444
0.1221808	0.0135756444

Step-3: Eigenvalues has been plotted in a graph, and observed more important parameters related to class 10 level students.

### Principal Component Analysis

Step 1: Data has been stored in Table-1.



Step-4: Eigenvectors of Table-2 has also been furnished in Table-4.

Table-4

INDICATORS	C1	C2	C3	C4	C5	C6	C7	C8	C9
X1	-0.37438782	-0.22307603	0.47786487	0.07278616	0.08257105	0.16589158	0.3040177	0.61816341	-0.25986044
X2	-0.39352403	0.274039	0.071017	-0.1393970	0.31660947	0.5594798	0.2114965	-0.33342311	0.41973923
X3	-0.36518111	0.27196947	0.0891717	-0.5471159	0.15763384	0.2793363	0.1956435	0.40293914	0.3533271
X4	-0.36347310	0.141375	0.103224	0.5928844	0.245344	-0.1653652	0.3987373	-0.2871551	0.38139649
X5	-0.09519851	0.61238317	0.5247916	0.0999679	0.5014977	0.4054505	0.2611432	0.005660311	0.0585744
X6	-0.301441	0.49098	-0.0548700	0.13831542	-0.16389	0.52567061	-0.52360	-0.03036546	0.2436944

	31	5	0		524		90	35	6
X7	-0.31134450	0.37022985	0.38462	-0.16817353	-0.60725302	-0.11147770	0.3827417	0.07178626	0.0768549
X8	-0.35970523	0.1332755	-0.2056824	0.121644	0.08925797	-0.45342596	-0.290964	0.076694293	0.5037496
X9	-0.33652606	0.04808062	0.56683	-0.093816	0.08174493	0.101898	0.1963477	0.30481740	0.499489

**Factor Analysis:**

- Step 1: Evaluation of Data. (Correlation Matrix, KMO)
- Step 2: Factor Extraction. (PCA)
- Step 3: Eigenvalues, Variance Explained.
- Step 4: New Constructs.

**Conclusion**

Principal Component Analysis and Factor Analysis both are used for dimensionality reduction, i.e. higher dimension to lower dimension. In this research work, Principal Component Analysis has been implemented to find more and less important features related to class 10 level students. Factor Analysis has been implemented to find the factors related to class 10 level students. In case of PCA more important attributes are equal to the number of original variables, and in case of Factor Analysis, the factors are equal to the original variables. The steps of PCA and FA, are almost same. We only can differentiate them in terms of their applications. The future researchers of this domain can apply both PCA and FA for education systems assessment. In this research work, it has been shown that first two components are most correlated with the indicators “Result in Surprise Test” and “Financial Background” respectively. So we can conclude by applying Principal Component analysis these two are more important parameters to assess class 10 level students. It has also been shown that by applying Factor Analysis we will find two

factors among nine components for class 10 level students assessment.

Our main aim was to identify the class 10 level students, those has the ability to improve their academic performance, but they won't be able to implement due to several reasons. The idea of this paper is also be helpful for the future researchers of this domain.

Based on the primary data, this research paper attains a complete evaluation methods of class 10 level students quality. By calculating the scores of Principal Components and find Factors, we can observe the strengths and weaknesses of class 10 level students.

### **References**

- [1] I.T.Jolliffe, "Principal Component Analysis", Springer,
- [2] T.Zhang,J.Guo,N.Zhang, "Application of PCA in Quality Assessment of College Students", International Conference on Information Technology, Computer Engineering and Management Science, IEEE Computer Society, 2011, pp.134-137.
- [3] S.N.Mandal, J.Pal Choudhury, .R.Bhadra Choudhury, "Neuro-PCA-Factor Analysis in Prediction of Time Series Data," American Journal of Intelligent Systems, 2012, pp.45-52
- [4] G.Hong-bo, L.Xuan, Research on Optimizing Quality of Higher Vocational Education Based on Principal Component Analysis of Data Mining", International Conference on Intelligence Science and Information Engineering, IEEE Computer Society, 2011, pp.534-537
- [5] Z.Sumei,"The Comprehensive Evaluation of Teaching Quality Based on Principal Component Analysis", International Conference on Advanced Computer Theory and Engineering(ICACTE), IEEE, 2010, pp.v3-314-v3-317.
- [6] Jamilah, M.; Zakaria, A., Md. Shakaff, A.Y., Idayu, N.; Hamid, H., Subari, N., Mohamad, J(2012). Principal Component Analysis-A Realization of Classification Success in Multi Sensor Data Fusion. In Principal Component Analysis-Engineering Applications; InTech: Philadelphia, PA, USA, 2012; pp. 1–25.
- [7] Paul, L.C., Suman, A.A., Sultan, N(2013). Methodological Analysis of Principal Component Analysis (PCA) Method. Int. J. Comput. Eng. Manag, no.16, pp.32–38.
- [8] Yu M Z 2020 Analysis and reflection on the effect of the implementation of the second class transcript system for university students: an example from Huabei Normal University for example Science & Technology Vision 2020(34):46-47.
- [9] Bartlett, M. S., A note on the multiplying factors for various chi square approximations, Journal of the Royal Statistical Society, 16, series B, 296 – 298, 1954.
- [10] Comrey, A. L., A First Course in Factor Analysis, Academic Press, New York 1973.