IJCTA, 10(01), 2017, pp. 19-29

Multivariate Clustering Evaluation of Principal Component Analysis and Singular Value Decomposition

Ann Baby* A. Kannammal and Sania Correya*****

Abstract : Quality of work life is an experience and Job satisfaction is the expression of employees about their job. The purpose of this research is to identify the relationship between the Quality of work Life and Job satisfaction. The findings of the study are done by implementing Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) algorithms. These algorithms aim to reduce the dimensionality ofvariables and to find the most appropriate factors. The results of this research point out a strong divergence in the perception of employees towards quality of work life and job satisfaction which is remarkable as it reflects the different requirements and priorities of employees in IT industry.

Keywords : Principal Component Analysis (PCA), Singular value decomposition (SVD), Fuzzy *c*-means (FCM).

1. INTRODUCTION

The quality of work life (QWL) can be defined as the quality of relationship between the employees and the work environment such that employees have a significant influence in shaping organizational environments in methods used to increase not only their own motivations and job satisfaction but also the productivity and profits of the company. QWL covers a number of areas like work environment, career growth, relationship behavior ,salary, security, bonus and incentives, organization culture, training, facilities, work life balance, time management, gender inclusive, organization support, role conflict, recognition and health. Satisfaction is the job induced motivation and interest in work. This study is to identify relation between QWL and job satisfaction leads to positive attitude and behavior. Job satisfaction leads to decreased turnover and less absenteeism. The antecedents that affects the job satisfaction are job autonomy, job security, employee commitment, job satisfaction, job designation, work exposure, rewarding system, interpersonal development, risk factors, stress, strain, support of top management, transportation, training, facilities, transportation, technical help, motivation, professional recognition, clear statement of projects.

A measure of job satisfaction and QWL (or any other subjective attitude) can be obtained by constructing multiple items indicating different aspects of the attitude. These multivariate data can be reduced into univariate scales in many ways. A well-known technique that allows computing such weights is the linear Principal Components Analysis (PCA). When all the correlations among variables are large, the simple sum is a reasonable choice for a univariate compression of variables. However, PCA is frequently used in data analysis to obtain not just one scale, but a set of subscales, taking into account the multidimensionality of the variables measuring the concept of interest. PCA is used to reduce a number of variables to a much

* Rajagiri College of Social Sciences, Kalamassery, Kerala. *ann@rajagiri.edu*

** Professor, Coimbatore Institute of Technology. Coimbatore

*** Rajagiri College of Social Sciences, Kalamassery, Kerala.

smaller number of Principal Components, linear combinations of the initial variables, such that they retain as much information as possible. Consequently, instead of the large number of original variables coming from a job satisfaction survey, subscales of the concept can be established and used in other models as dependent or independent variables. Usually, a subscale is the simple sum score of a subgroup of variables loading high on one particular Principal Component. (Bloom, 2007)

The most appropriate method for computing Principal Components is the Singular Value Decomposition algorithm (SVD) (CalisKan, 2010) .Since the data matrix can be decomposed directly, and with appropriate preliminary scaling yields the same solution as factor analyzing the matrix. Singular value decomposition (SVD) can be looked at from three mutually compatible points of view. On one hand, we can see it as a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items. At the same time, SVD is a method for identifying and ordering the dimensions along which data points exhibit the most variation. (Dawson, 1999)

JS and QWL variables are of high dimensions. The variables which effect more can be identified by the algorithms PCA and SVD which are used for dimensionality reductions.

2. LITERATURE REVIEW

This section provides an understanding of the relationship between two variables, *i.e*. Quality work life and Job satisfaction, through various studies and researches.

A. Quality of work life

In the present scenario, there is need of improvement in quality of work life programs for enhancing the productivity and satisfaction level of organizations' employees (Haque, 1992). Better quality of work life encourages human self-esteem and development, people compatibilities, collaboration for work and organizational goals. When employees get quality of work life environment, they feel satisfied, motivated, committed at work place (Dhawar, 2014)

B. Job Satisfaction

Job satisfaction is the most studied construct in business science and organizational behavior. Job satisfaction is the expression of employees about their job and the expectations from the job that is a desired outcome of employees for their involvement in the organization. (Armstrong, 2013). Job satisfaction is the degree to which people like their jobs. Some people enjoy work and find it to be a central part of life. Others hate to work and do so only because they must (paul, 2009). Job satisfaction is more of an attitude, an internal state. It is affected by a wide range of variables relating to individual, social, cultural, organizational and environmental factor. (Mullins, 2005)

C. Principal Component Analysis

PCA is based on some assumptions that are often not true in social sciences: all of the variables are assumed to be of numeric (interval or ratio) measurement level and the relationships between variables are assumed to be linear. As in social and behavioral sciences, many variables are nominal or ordinal and relationships between variables are frequently nonlinear, linear or standard PCA is often not the most appropriate analysis method, although it is commonly used. The standard PCA could be appropriately used in the presence of categorical variables only after verified the existence of linearity in the variables and in the relationships with variables. To avoid the limitations of standard PCA, nonlinear PCA has been introduced and developed during the last 40 years (Hagedorns, 2000).

D. Singular Value Decomposition

SVD is a factorization of a matrix (real or complex). It has many useful applications in signal processing and statistics. Singular Value Decomposition (SVD) is a method for identifying and ordering the dimensions along which data points exhibit the most variation. (Chambers, 1977)

3. OBJECTIVES OF THE STUDY

The objectives of this study are:

1. To implement algorithms such as principal Component Analysis (PCA) and Singular Value Decomposition (SVD).
2. To identify the factors which influence on quality of work life and job satisfaction of the IT employees in Kerala.
3. To determine the relationship between quality of work life and job satisfaction.
4. To perform comparison between the results of algorithms PCA and SVD.

4. METHODOLOGY

In the present study, attributes were generated from the previous studies related to quality of work life and job satisfaction. The generated attributes were consolidated, repetition in attributes, similar attributes and unclear attributes were deleted from the list of attributes. To ensure the inner consistency of the present instrument 'Cronbach's Alpha' reliability test was applied. Application of the Cronbach's alpha co-efficient with the purpose of testing the inner consistency of the instruments presented the following results.

Table 1

Relaibility Statistics

Variables	*Cronbach's Alpha*	*No: of items*	*Scale Statistics*	
			Mean	*S.D*
Job Satisfaction	.92	32	221.5	14.8
Quality of work life	.85	17	.65.6	57.9

Table 1 presents the reliability coefficient associated with quality of work life and job satisfaction scales, number of items in the scale, its mean and standard deviation. Higher values of Alpha indicate higher reliability. Reliability values are greater than 0.75 for both the scales. Hence, the questionnaire meets the reliability requirements of minimum value of 0.75 as recommended by Rodrigues. (Rodrigs, 2012)

5. PROPOSED SYSTEM

Primary data were collected from two hundred and six IT employees by administering the research instrument. Descriptive statistics, factor analysis (PCA, SVD) and correlation were used to analyze the data. The proposed system performs PCA algorithm, SVD algorithm, Comparison of both factor analysis algorithm and finds the relationship between job satisfaction and quality of work life. Analysis is done on the basis of the model developed. A behavioral model was initially developed by the help of literature review.

A. Principal Component Analysis

Reduction into small major dimensions was performed by factor analysis using Principal Component Analysis with Varimax Rotation is applied. In order to examine the suitability of the data for factor analysis, the following steps are considered.

1. The correlation matrices are computed and examined. It reveals that there are enough correlations to go ahead with factor analysis.
2. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for the individual variables is studied. The KMO calculated is found to be0.610. This score indicates that the sample is good enough for sampling.
3. The overall significance of correlation matrices is tested with Bartlett's Test of Sphericity which proved to be highly significant. It indicates valid inter correlations between the items and proved goodness of fit to the data.

B. Principal component analysis of quality of work life

There are 17 variables associated with job satisfaction which are work environment, career growth, relationshipbehaviour, salary, sequrity, bonus and incentives, organisation culture, training, facilities, workbalance, gender inclusive, time management, training , recognisation, health, and role conflict.

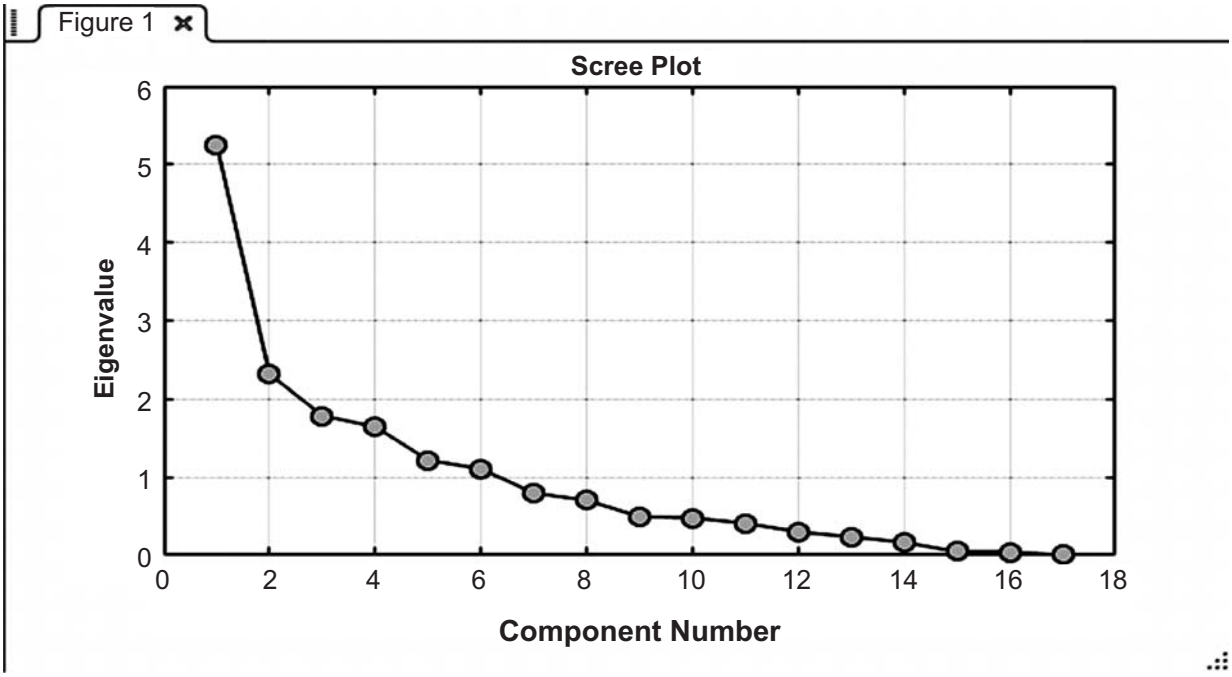

Figure 1: Quality ofwork life eigen values

The study presents an exploratory factor analysis model to explain the factors influencing the QWL of employees working the life IT companies. All the six factors are found to be highly influencing life of IT company employees.

Factor 1: This factor consists of employees' opinion about work environment, career growth and relationship behavior. The factors are positively loaded. Hence factor 1 is identified as **"Work life Balance"**

Factor 2: The second factor considers the employee opinion about salary, security and bonus. The factors are positively loaded. Hence the factor 2 is identified as **"Organization Support"**

Factor 3: The third factor consists of employees opinion regarding men /women equality, facilities, job designation and roles managed. The factors were positively loaded. Hence factor 3 is considered as**" Gender Inclusive"**

Factor 4: The fourth factor contains the employee opinion about job autonomy, interpersonal skills and recognitions. The factors are loaded positively. Hence the factor 4 is considered as **"Recognition"**

Factor 5: The fifth factor contains the employee opinion about organization support and opportunities for career enhancement. The factors are loaded positively. Hence the factors are considered as **"Career Advancement"**

Factor 6: The sixth factor contains the employee opinion about time management and health. The factors are loaded positively. Hence factor 6 is considered as **"Time Management"**

C. Results of quality of work life compared to other statistical tools

The results of factor analysis of quality of work life were compared to other statistical tools. The error plotted in bar graph is shown in fig-3 .A negligible error which is of the power 10^{-5} is found.

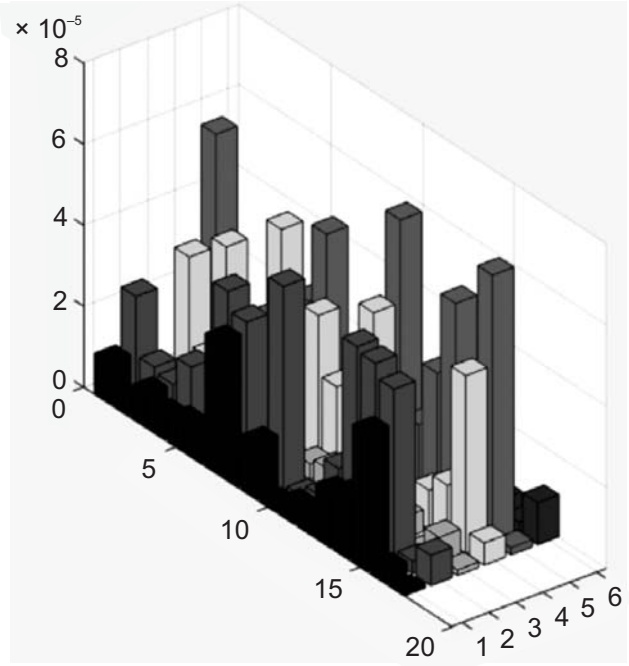

Figure 2: Results of quality of work life compared to other tools

D. PCA factor analysis of job satisfaction

There are 32 variables (table-5) which represent job satisfaction. Eigen values greater than one is considered significant. All the factors with latent root less than one are concluded to be insignificant and ignored. The 9 dimensions of quality of work life are identified and labeled aswere employee satisfaction and engagement, interpersonal relationships and recognition growth and support, internal environment, time management and resources, stress and strain, work place environment, facilities and development.

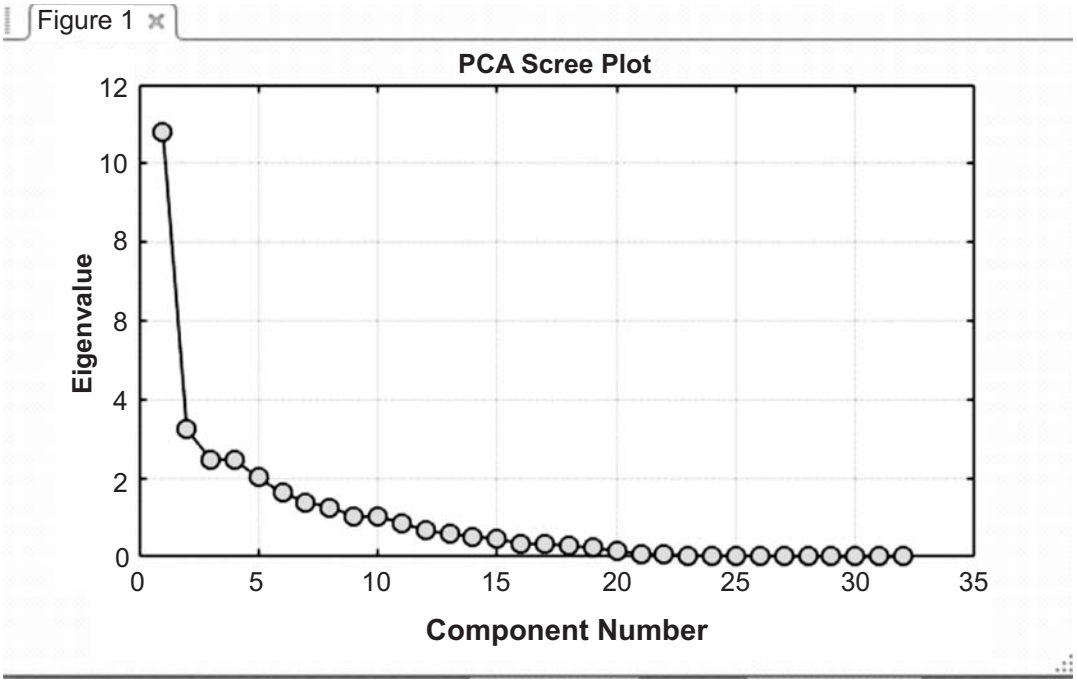

Figure 3: Eigen values of job satisfaction

Factor 1: This factor consist of employees' opinion about Quality for job, in service training, Work exposure, Lack of motivation, Direction by managers and Research policies. The factors are positively loaded. Hence factor 1 is identified as **"Employee satisfaction and engagement.**

Factor 2: The second factor considers the employee opinion political problems, feedbacks, communication. The factors are positively loaded. Hence the factor 2 is identified as **"Interpersonal relationship and recognition".**

Factor 3: The third factor consists of employees' opinion regarding professional growth, job autonomy, rewarding system, career growth and bonus. The factors were positively loaded. Hence factor 3 is considered as**" Career growth and support".**

Factor 4: The fourth factor contains the employee opinion about clear statement of projects, relation between staff and admin and continuity of programs .The factors are loaded positively. Hence the factor 4 is considered as "**Internal Environment".**

Factor 5: The fifth factor contains the employee opinion about training and tools, efficient support for family and transportation. The factors are loaded positively. Hence the factors are considered as **"Time Management and Resources".**

Factor 6: The sixth factor contains the employee opinion about medical assurance and insurance .The factors are loaded positively. Hence factor 6 is considered as **"Stress and Strain".**

Factor 7: The seventh factor contains the employee opinion about job specialization and motivation. The factors are loaded positively. Hence factor 7 is considered as**" Workplace environment".**

Factor 8: The eighth factor contains the employee opinion about technical help and opportunities for publishing. The factors are loaded positively. Hence factor 8 is considered as" **Facilities and Development".**

Factor 9: The ninth factor contains the employee opinion about managerial style. The factors are loaded positively. Hence factor 9 is considered as" **Top Management".**

The results of factor analysis of job satisfaction were compared to a statistical tool. The error plotted in bar graph is shown in fig-5.A negligible error which is of the power 10^{-5} is found out.

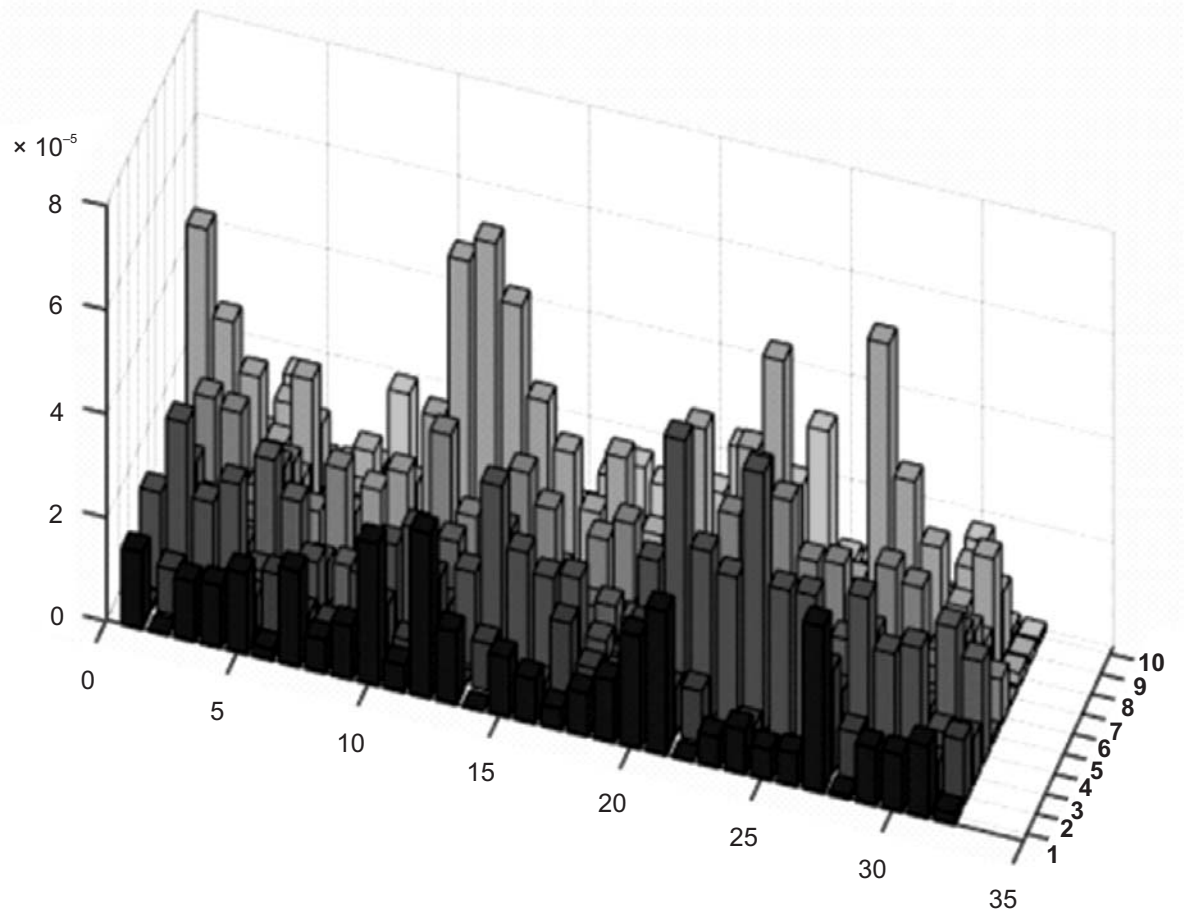

Figure 4: Results of job satisfaction compared to other statistical tool

E. Singular value decomposition

SVD is a factorization of a matrix (real or complex). It has many useful applications in signal processing and statistics. Singular value decomposition (SVD) is a method for identifying and ordering the dimensions along which data points exhibit the most variation.

The singular value decomposition of an mxn real or complex matrix A can be broken down into the product of 3 matrices

$$A_{mn} = u_{mn} s_{mn} v_{nn}^{T} \quad (1)$$

Where U = *an* orthogonal matrix S = *a* diagonal matrix with non-negative real numbers and all its entries are singular values of A. V* = transpose of an orthogonal matrix V (real or complex unitary matrix, if *v* is real) Also we have to put in notice that, the n columns of V are called the left-singular vectors and the m columns of U and and right-singular vectors of M, respectively. The eigen-vector for XX*(X multiplied to X transpose) and arrange them as a column vectors in in a matrix ordered by the size of the corresponding eigenvalue. Convert this matrix into an orthogonal matrix which we do by applying the Gram-Schmidt ortho-normalization process to the column vectors

Calculate similarly (step 1) for V* using X*X. (X transpose multiplied to X). For matrix S, take the square roots in the non-zero eigenvalues and also populate the diagonal with these, putting the major in *s*11, the next major in s22 and so forth until the smallest value results in smm.

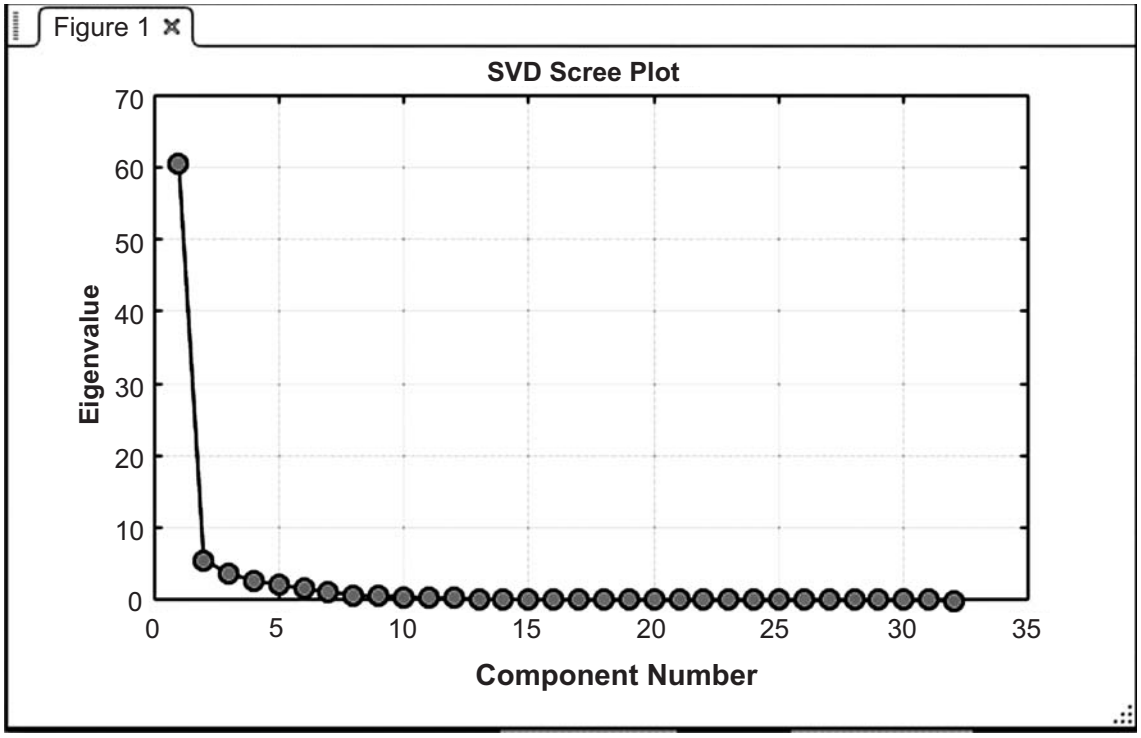

Figure 5: Eigen values of job satisfaction using SVD

The variables of job satisfaction has 32 variables. Eigen values greater than one is considered significant. All the factors with latent root less than 1 are concluded to be insignificant and ignored. The 10 dimensions of job satisfaction are identified and labeled same as that of PCA

The variables of quality of work life has 17 variables. Eigen values greater than one is considered significant. All the factors with latent root less than 1 are concluded to be insignificant and ignored. The 6 dimensions of quality of work life are identified and labeled same as that of PCA.

Eigen values greater than one are considered significant. All the factors with latent root less than one are concluded to be insignificant and ignored. The six dimensions of quality of work life are identified and labeled as work life balance, organization support, gender inclusive, recognition, career advancement and time management.

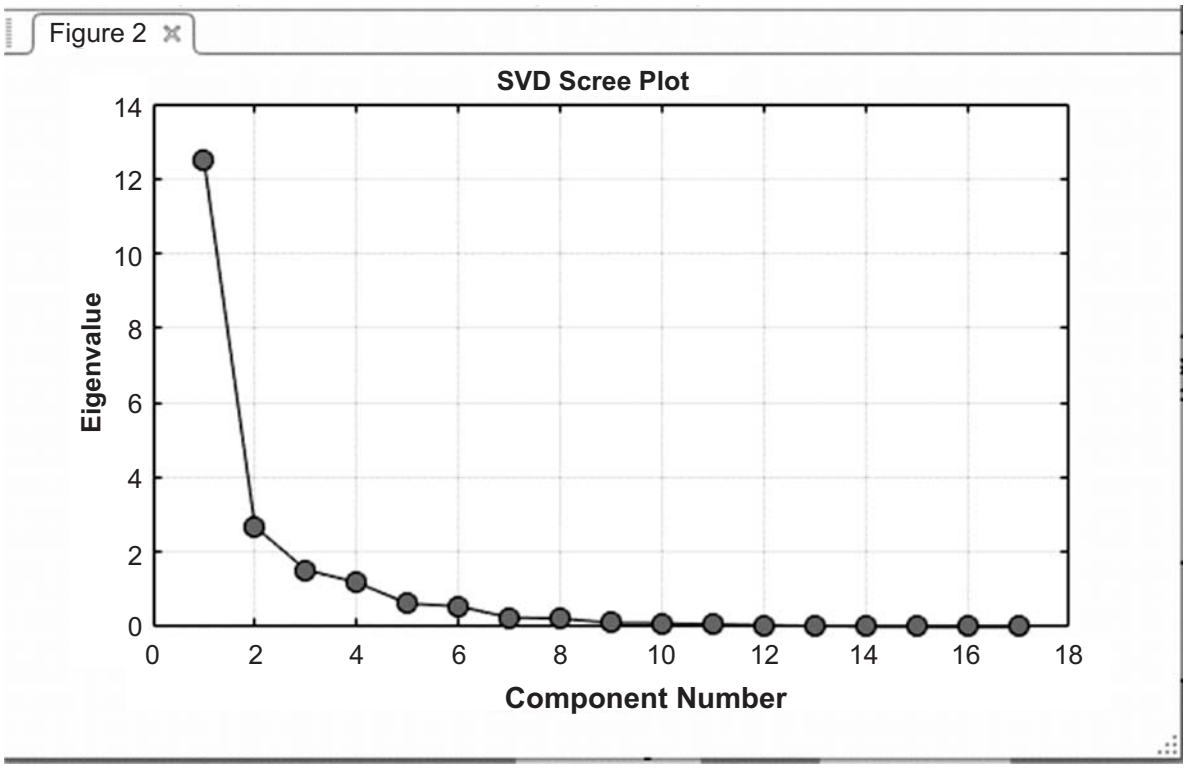

Figure 6: Eigen values of quality of work life using SVD

The factor loading with Varimax Rotation for quality of work life was done to investigate the underlying relationships of a large number of items and to determine whether they can be reduced to a smaller set of factors.

F. Comparison between SVD and PCA

There is a very direct mathematical relation between SVD and PCA.For this reason, the two algorithms essentially deliver the same result: a set of "new axes "constructed from linear combinations of the original feature space axes in which the dataset is plotted. These "new axes" are useful because they systematically break down the variance in the data points (how widely they are distributed) the direction in which there is great variation is referred as Principal Components.

There are many benefits by reducing dimensionality of the features per data point.

Firstly, a reduction computational cost can be achieved. In many data sets, most of the variance can be explained by relatively small number of principal components, when compared to the dimensionality of the entire feature set. In such cases the computational cost per data point may be reduced by many orders of magnitude with a procedure like SVD without losing much of the information stored by looking at the data in the unreduced feature space.

Secondly, a rotation in data can eliminate collinearity in the original feature space. By reducing the number of features with SVD, such collinearity may be eliminated by attending to a single axis.

Thirdly, it reduces the impact of a phenomenon called curse of dimensionality. The cause of this effects many machine learning algorithms to flatter if the dimensionality is too high. By performing SVD the effect of curse of dimensionality can be mitigated.

According to theory, the PCA viewpoint requires that one compute the eigenvalues and eigenvectors of the XX^T covariance matrix, which is the product where X is the data matrix. Since the covariance matrix is symmetric, the matrix is diagonalizable, and the eigenvectors can be normalized such that they are orthonormal:

$$XX^T = WDW^T \quad (2)$$

On the other hand, applying SVD to the data matrix X as follows:

$$X = U\sum V^T$$

$$X = U\sum V^T \quad (3)$$

Attempting to construct the covariance matrix from this decomposition gives

$$XX^T = (U\sum V^T)(U\sum V^T)$$

$$XX^T = (U\sum V^T)(U\sum V\text{^}T)^T \quad (4)$$

Since VV is an orthogonal matrix ($V^T V = I$),

$$XX^T = U\sum^T U^T \quad (5)$$

The correspondence is easily seen (the square roots of the eigenvalues of XX^T are the singular values of X, etc.)

In fact, using the SVD to perform PCA makes much better sense numerically than forming the covariance matrix to begin with, since the formation of XX^T can cause loss of precision.

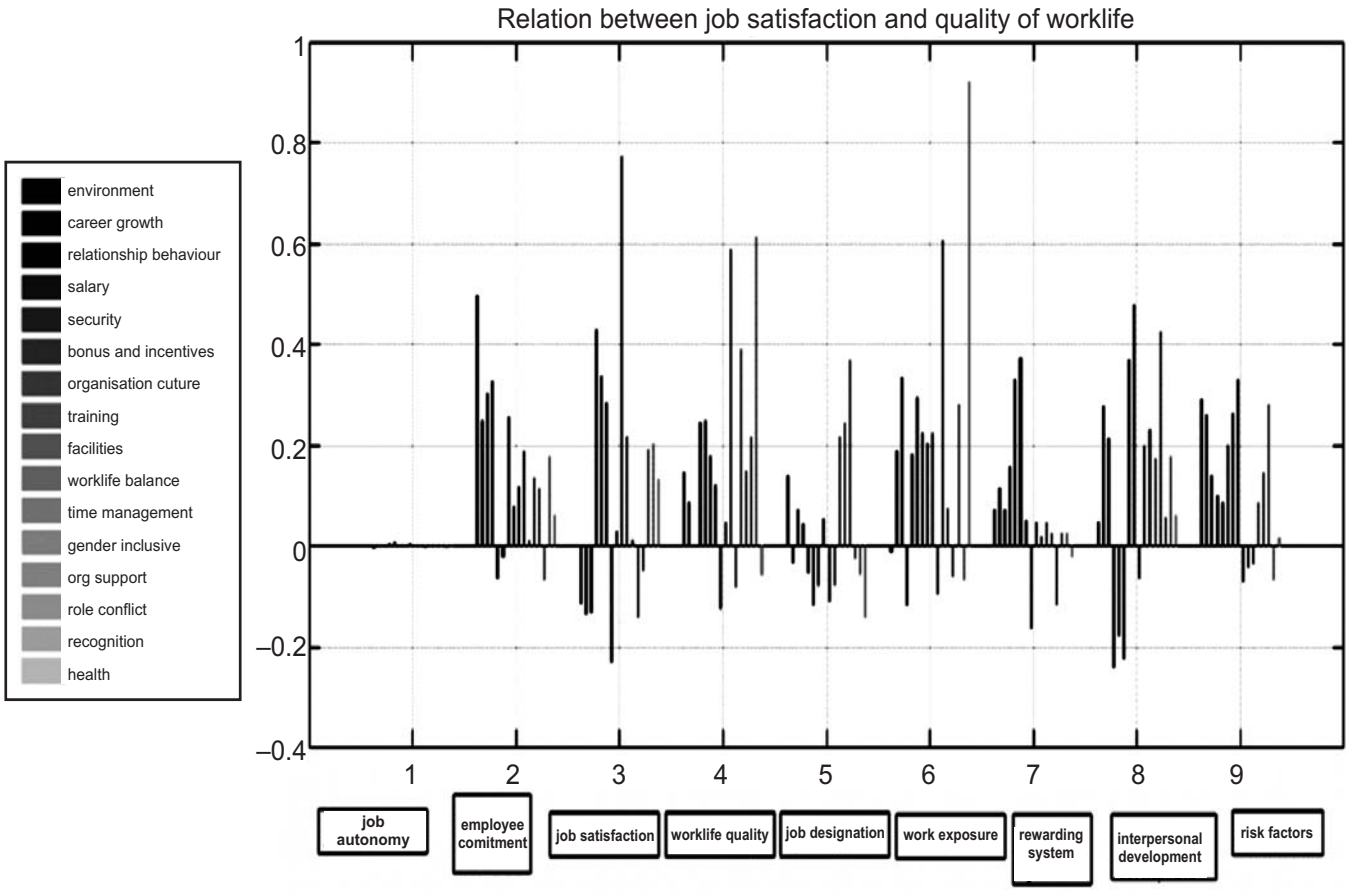

Figure 7: Relation between job satisfaction and quality of work life

The variables of quality of work life is compared to job satisfaction with the help of linear regression coefficients. Linear regression is an approach for modelling the relationship between a scalar dependent variable Y(quality of work life) and independent variables X(job satisfaction).A linear regression of the form, $Y = aX + b$, where X is the explanatory variable and Y is the dependent variable. The slope of the line is *b*, and is the intercept (the value of Y when X = 0). The bar graph shows the variation in the association of QWL and JS variables. The results shows that job autonomy has no much association with job satisfaction. All other variables have positive or negative association to each other. Health and work exposure has highest association. Training and job satisfaction has least association.

G. Analyzing relation between demographic variables and job satisfaction

The demographic variables such as gender, age, marital status and experience were collected. Using two clustering algorithms, Fuzzy c-means and subtractive clustering the patterns of job satisfaction were studied with respect to these demographic variables. Age of person plays a vital role in his job satisfaction. While performing clustering two cluster groups were found, one which contains highest age and another with lowest age. It is seen that a person having highest age (45-50) seems to be satisfied more than a person with less age. The also results shows that middle age (30-35) person seems to be less satisfied to their job and they prefer to try out different organization instead of sticking into one. A person in the age group of 20-25 seems to be highly satisfied in an organization of IT.

Weak relationships were observed between gender and salary. Females rated these constructs higher than males, indicating a higher level of satisfaction with personal learning and growth opportunities at work, job security, and compensation.

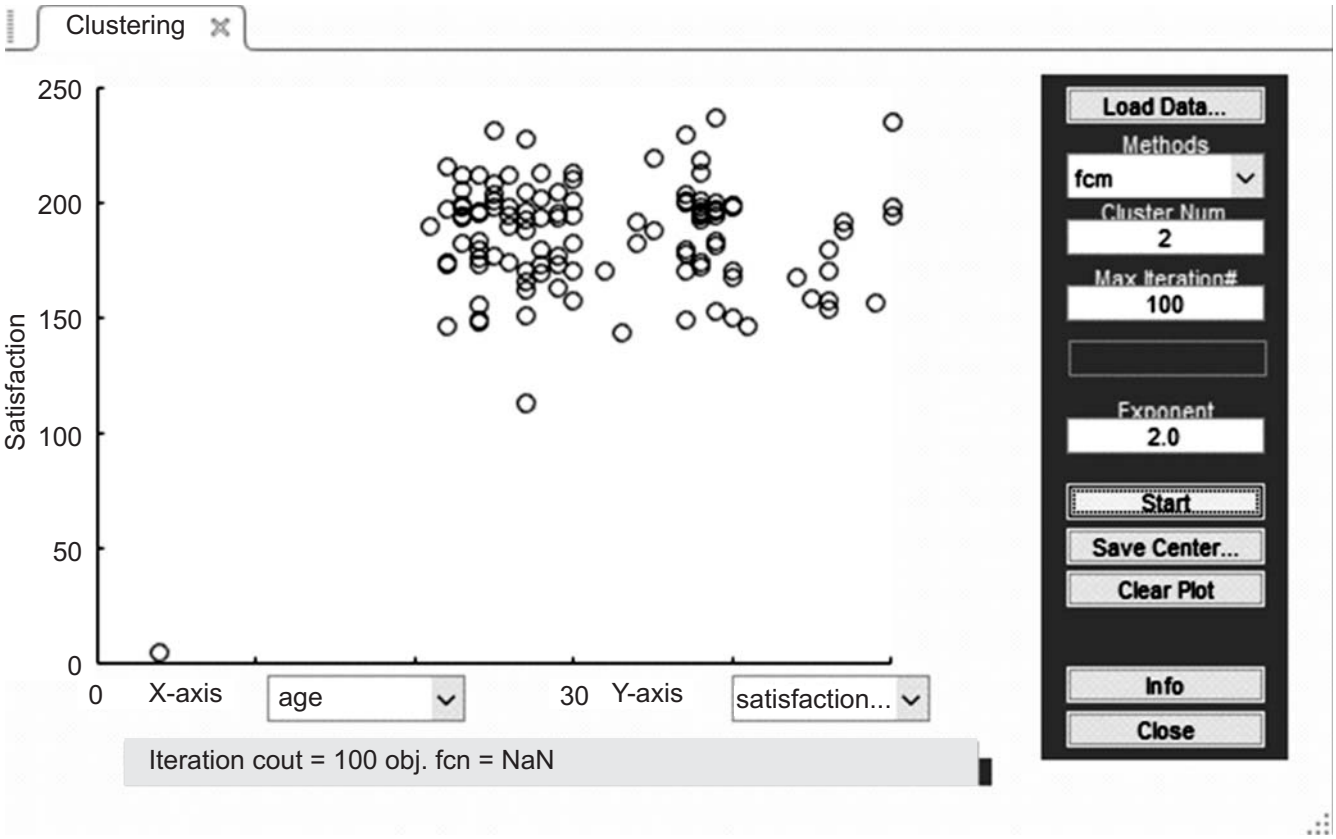

Figure 8: Age compared to job satisfaction using clustering

6. FUTURE SCOPE

The data generated is growing exponentially and the problem is large data sets are challenging to manipulate and handle while analyzing a problem. PCA and SVD can be used for retaining the information were high dimensional data has to be used. More number of variables can be added on to the system in further use. The absentees am and employee turnover can be reduced in an IT industry by satisfying the highly effected factors which reduce the loss of losing a well-trained employee

7. CONCLUSION

Job satisfaction level among IT company employees is positively correlated with the quality of work life factors. The study found that among the quality of work life variables work life balance, organization support, gender inclusive, recognition career advancement and time management of IT employees were

the factors that depend highly on QWL of an employee. Quality of work life cannot be isolated from modern human resource management practices in the corporate companies. Tough control of officers cannot provide appropriate result to the organization as well as officers. The reason for selecting IT sector for the study is that it has been playing a momentous role in Industry economy after 1991. As per research estimation in 2020, Indian IT industry will be on 3rd position in top level ranking IT companies worldwide (Drawer, 2014). But this sector is also facing a severe problem that is high attrition rate of employees in IT companies. This problem can only be solved if IT companies develop strategies for improving the quality of work life of employees.

The main factors that depend on the job satisfaction of an employee are employee satisfaction and engagement, interpersonal relationships and recognition growth and support, internal environment, time management and resources, stress and strain, work place environment, facilities and development. There is a direct relationship between job satisfaction and quality of work life. The current thinking in numerical analysis is that the most appropriate method for computing Principal Components is the singular value decomposition algorithm .This procedure does not require the computation of the matrix of phi coefficients.

8. REFERENCES

1. R. J. Staw, "QWL:New Working Relationship In Communication Industry," Labour Studies Journal, Pp. 261-274, 8 1984.
2. Bloom, "Measuring And Explaining Management Practice Across Firms And Country," Pp. 1351-1408, 2007.
3. Caliskan, "The Impact Of Stratergic Human Resources Management On Organisational Performance," Journal Of Naval Science And Engeineering, Pp. 100-116, 2010.
4. Dawson, "Computer Technology And The Job Of First Line Supervisor In Work Environment," Study Using Singular Value Decomposition , Pp. 47-59, 1999.
5. A. Haque, "Quality Of Work Life And Job Satisfaction Of Industrial Workers," Bangladesh Study Of Psycological Studies, P. 43, 1992.
6. S. Dhawar, "Dynamics Of Quality Of Work Life:A Study Of Life Insurance Sector," Asian Journal Of Management, Pp. 312-217, 2014.
7. Armstrong, A Hand Book Of Human Resources Management Practice, Kogan Page , 2013.
8. Paul, "Job Satisfaction,Application,Assement Causes And Consequences," Usa, Sage Publications, 2009.
9. Mullins, "Management Of Organisational Behaviour," United Kingdom, Person Education, 2005.
10. Hagedorns, "Conceptualizing Faculty Job Satisfaction :Components ,Theaory ,Outcomes.," New Directions For Research, Pp. 105,5-20, 2000.
11. Chambers, "Computational Methods For Data Analysis," New York, 1977.
12. L. Davis, "Quality Of Worklife:," In Problem,Prospects And The State-Of-The Art, New York, 1975.
13. F. A. Rodrigs, "Work Related Quality Of Work Life," Staistical Measure , P. 12, 2012.
14. Gani, "Quality Of Work Life In A State :The Indian Journal Of Labour Ecnomics," In Findings Of An Emperica Study, 1993.
15. Yousaf, "Evaluating Quality Of Work Life," In Management And Labour , United States, 2015.
16. Louis, "Comparison Of Singular Value Decomposition," Factor Analysis Techiniques And Study On Factors, 1999.
17. Lawler, "Motivation In Work Organization," In Asian Journal Of Management Research, 2012.
18. Likert, "A Techniques On Measurement Of Attributes," Archieves Of Psycology, Pp. 15-20, 2008.
19. Stewart, "The Application And Misapplication Of Relationship Analysis," In Seminar On Regression Analysis , Florida, 2010.
20. Jerome, "A Study On Regression Analysis," International Journal Of Advance Research In Computer Science And Management Studies , 2000.
21. Burke, "The Changing Work Of Nature And Organization," 2012.
22. Reddy, "Comparitive Analysis Of Quality Of Job Satisfaction," In Conference On Job Satisfaction, 2015.
23. P. Drawer, "Dynamics Of IT Industry," Asian Journal Of Management, Pp. 312-217, 2014.
24. Walton, "Inter National Labour Organisation:Recommended," In National Seminar On Quality Of Work Life, New York, 1982.
25. Scott, "Extention agents persception of fundamental job characteristics and their level of job satisfaction," journal of southern educational research, pp. 88-101, 2005.