PalArch's Journal of Archaeology of Egypt / Egyptology

HR- ANALYTICS FOR EFFECTIVE PERFORMANCE APPRAISAL

Priyansh Gupta[1], Manoj Hudnurkar[2], Suhas Ambekar[3]

Symbiosis Centre for Management and Human Resource Development,

Symbiosis International (Deemed University), Pune, India[1,2,3]

Email: [3]suhas_ambekar@scmhrd.edu

ABSTRACT

This paper aims to remove subjectivity from the performance appraisal process through data analytics. In doing so, the paper lists down all the biases that affects the ratings of an employee and develop methods that flag off those potential biases. This paper uses statistical parameters, statistical tools as well as parametric tests to devised methods to quantify the identified biases. We have considered weighted average ratings given by a manager to its subordinates over the period of years and utilized it to classify the consistent behaviour of the manager into several biases. The distribution pattern of ratings given to employees by most of the managers shows highly negative skewed normal distribution curve indicating the presence of leniency as well as central tendency bias. We have also found out that 52% of the employees exhibit subjective bias indicating the presence of biases in the performance appraisal system. This paper provides implications for all the people associated with the performance management area for reducing the subjectivity in the appraisal process using data analytics and for future researchers to test and analyse the suggested methods for different organizational settings. The paper offers insights about how employee performance data can be used to identify different existing biases in performance appraisal system by suitable analytical methods and thus contributing in making a more transparent, more objective and hence, a more effective Performance Appraisal System.

1. Introduction

Employees are the heart for any organisation and thus it becomes necessary to keep the employees motivated and engaged. Researchers termed employees as an asset to the organisation who helps in increasing the productivity and thus building and maintaining the goodwill of the company/organisation, he is associated with (Sastry, 2018). Performance Appraisal system is a critical Human Resource domain which is used as a measure for employees' career management, career development, talent mapping, high -flyers' identification, succession planning (promotions), team composition, and diversity analysis and hence covers almost all the strategic human resource practices. Performance appraisal is critical as it tends to affect the organisational performance (Ayers, 2015) and research also shows it impacts employees' participation in informal learning activities at work (Bednall, 2014). It also plays a very important role for achieving the goals and objectives of the organisation (Mishra, 2013).

Performance appraisal has a high correlation with the promotion (Jansirani, 2013) and also affects the motivation levels of an individual triggering innovative behaviour amongst the employees (Jansirani, 2013). The progress and growth of an individual is dependent on their performance and performance appraisal system has the capability to improve quality of both employee as well as the organisation he\she is associated with (Xavier, 2015).

Performance Appraisal System tends to affect each and every facet of a successful human resource management system of an organization and hence it becomes imperative to have an error free, bias free, an efficient as well as a transparent performance appraisal system. We have seen that there is a lot of literature around performance appraisal processes, their pros and cons, introduction of biases and impact of biases on the satisfaction and motivation level of the employee as well as on the organization as a whole. However, there is not much research that is been done for finding and quantifying these biases leveraging the potential of appraisal data. Hence through this paper we have tried to fill this gap and devised several methods using analytics, central tendency measures coupled with certain statistical tools such as box-plot and chi-square tests to quantify certain most important biases and through this paper organizations will be able to list down all the managers who tends to follow certain behaviors inducing biases in the overall evaluation system.

The rest of the paper is organized as follow: the following section i.e. Literature Review section presents an overview of performance appraisal system covered by prior literature, it also discusses certain methods used for performance evaluations with their pros and cons and finally introduces biases with its overall effects on the organisation. it is followed by a research question section that lists couple of questions that this paper aims to solve. The next section is Research methodology which briefly describes the mode and criteria behind data collection, description of tools used and finally describing certain terms used as the paper proceeds. It is been followed by results and analysis section which provides the step wise elaboration of the devised methods with

the expected outputs. The next section is Discussion section followed by Managerial Implication section covering the exact usage of this paper for the organisations. The results are discussed in detail and concluded in the next section which is Conclusion and finally the final section namely Limitations presents the limited scope of the paper, and hence concluding with mentioning certain future research directions that future researchers can consider to take in the Performance Appraisal/evaluation or even Human Resource Management domain.

2. Literature Review

Performance appraisal system can be defined as evaluating behaviour of the employees of the organisation. It covers both subjective as well as quantifiable aspects of performance indicating the potential of an employee in fulfilling the tasks and demands expected for the job. It is always measured in terms of the result or outcome of an employee's actions and efforts. It is an important analysis which is utilised by organisations to identify employees worthy of continued employment, of promotion, pay hike, bonuses or any other incentives and acknowledgement (Bagul, 2014). According to Oladimeji (1999) performance appraisal as a method of evaluating the success, which an employee or organization has achieved in performing assigned tasks or meeting set goals over a period of time. It is hence a very important method for evaluating how the individual is performing with respect to the assigned work targets and goals (Oladimeji, 1999). (Mukesh, 2017) cites a book by (Atiomo, 2000) and defines performance appraisal system as a system which not only just measures an individual performance but it also results in identifying potential areas of improvement for the optimum utilisation of Human Resource. He also emphasises the importance of communicating the strengths, weaknesses and the overall performance evaluation results to the employee in study for the potential improvement of his performance in the coming years, which according to him is the purpose of the performance appraisal process. Performance appraisal is the process of evaluation by merit factors of the behaviour and ability of employee in their work and a process of rewarding a percentage above his salary based on the assessed merit value. (Sastry D. R., 2018). Performance management is the process of evaluating the performance and qualifications of the employees as against the defined expectations of the job he is recruited for, serving different purposes of administration such as placement, providing financial incentives, checking eligibility for promotion, and similar actions which demands differential treatment among the peers or other team members, it as a method of evaluating the value excellence, qualities or status of some object, person or thing contributing to recruitment, selection, staffing processes, placement and indoctrination (Xavier, 2015). The entire appraisal or performance management process is the funnel through which an employee is rewarded or penalised. It gives opportunities to companies and organisations to strengthen and diversify their talent pools and ensure success by promoting the best employees. On the other hand, this

process is critical to employees as this is the period where their value to the company is validated and their future careers is identified (Traub, 2013).

2.1. Performance Appraisal Methods

Different methods used by the organisations for evaluating their employees are very properly described in the literature (Jansirani, 2013). A brief description of all the methods has been included in the paper. Primarily all the methods have been divided into two parts -Traditional Methods and Modern Methods with each part further classified into a list of relevant methods.

2.1.1 Graphic Rating Scales

In this method performance of employees are being compared to an absolute standard. The ratters are provided with forms against each employee and contains number of trait-based qualities, objectives and behaviours to be rated such as amount of work done, quality of work, attitude, knowledge, skills, initiative taking capabilities and many more. Points are assigned by the rater against all the pre-defined parameters and employees is finally ranked according to the total points assigned to each individual. The standardise nature of the scales does not consider aspects of trait or behaviour relevance. Also, relevance of traits and skills varies with the type of job an employee does and hence workplace should also be considered in evaluation (Armstrong, 2009).

2.1.2 Ranking Method

In this method, employees are rated from best to worst against some well-defined parameters. The appraiser starts with highest performance and with lowest performance in that job category and rates them best and poorest respectively. Then the appraiser moves to next highest and lowest performers till he rates all the employees.

2.1.3 Paired Comparison Method

It is the appraisal method which is mainly used in the organisations or teams with a smaller number of employees to rate. In this method, the rater ranks the employees with all other employees in the group, one at a time. The biggest advantage of paired comparison is that it does not force distribution of ratings amongst the employees in each department.

2.1.4 Forced Distribution Method

In the earlier methods, there is no restriction on the rater on the distribution of the ratings but this method is developed to prevent the rater from rating too high or too low. Under this, after evaluating the employees against pre-decided parameters, the rater has to distribute the ratings according to the ideal pattern agreed in the organisation, which is mostly normal frequency distribution curves. This method can reduce the or completely eliminates, central tendency as well as leniency biases, if the distribution of the ratings matches with the

ideal curve prescribed by the organisation. However, in this method employees are placed in certain ranks within the groups or teams.

2.1.5 Checklist method

In this method, rater is given a list of statements or words against which he has to check statements which represents the behaviour and performance of the employee at best.

2.1.6 Critical Incident Method

Unlike all the above methods we discussed, it is a continuous evaluation process in which the rater records all the critical incidents of the employee's performance and behaviour related to all the parameters specified under performance appraisal. It is not necessary to have a separate rating system but it can act as a documentation of the reason for an individual's rating.

2.1.7 Essay or Free Form Appraisal

In this method, rater needs to write a short essay emphasising on the employee's overall performance based on the strength/weaknesses. This method minimises supervisory bias and halo effect as the supervisor/rater is asked to support his arguments with specific examples.

2.1.8 Group Appraisal

In this technique, an employee is been appraised by a group of raters. This group can consist of his immediate manager, managers who have some contact due to work, head of department and consultants. This group can use any of the above appraisal method to rate the employee and in the entire process, immediate manager can act as a co-ordinator for all the group activities.

2.1.9 Behaviourally Anchored Rating Scales (BARS)

This method is a combination of conventional rating scales and critical incident methods. Using BARS, effective as well as ineffective behaviours are described more objectively. The method asks employees who are familiar with the job to identify its important and major components, then the behaviour of the appraisee is then validated and ranked against those pre-decided components.

2.1.10   Assessment Centre

In 1930, this method was first used by German Army and after that the application of the method was followed by businesses and organisations. It is a system in itself, where evaluation of an individual is done by several experts by using various techniques which includes case-study, role playing, stimulation insight, transactional analysis etc.

2.1.11   Human Resource Accounting

This method deals with the contribution of and cost of an employee to the organisation. Employee contribution is the value in money of an individual's service which can be measured by labour productivity or value added by the employee. Cost of Human Resource includes cost of manpower planning, placement, selection, training and development, wages and all other incentives, perks associated to an individual.

2.1.12   Management by Objectives

Management by objective is a method where the manager and other sub-ordinate managers who are directly or indirectly involved in the work jointly decides on the goals and objectives that an employee must accomplish during the appraisal period. Managers define individuals' areas of responsibility in terms of the final outcomes expected from him and then these guidelines are used to evaluate the employees. The goals are measurable, tangible and verifiable and can also be altered in the middle of the appraisal period to evolve according to the business environment.

2.1.13   Behaviours Observation Scales (BOS)

In this method, the manager or the supervisor plays the role of an observer rather than the evaluator and he measures how frequently each of the behaviour has been observed and provides continuous feedback to employees to strive to meet the expected outcomes. Advantages includes increase in objectivity, identification of a need for training better than all other methods, easily able to differentiate good performers and bad performers, continuous feedback mechanism.

2.1.14   Result Method

This method of evaluation focuses more on the outcomes and results that an employee has achieved during the appraisal period rather than focusing on behavioural traits and factors. Results can be measured in terms of number of customers served, number of complaints handled, number of sales achieved, number of units manufactured and other similar metrics defined depending on the kind of work an employee is involved in.

2.1.15   Productivity Measures

Under this method of appraisal, employees are evaluated based on the ration of the output they were able to achieve to the input provided to them. Some of the metrics may be number of customers served in a day, total sales by an employee to employee's salary and many more. One of the limitations of this method is that not all the measures can be quantified thus making it difficult to evaluate employees on a common ground.

2.1.16   The Balance Scorecard

This method of appraisal method strives to align company's mission and vision with the goals and objectives to be accomplished by the employees. In this the focus is on linking processes, learning and overall profitability of the organisations. It believes that learning and development contributes to better and enhanced processes which are very critical for increasing customer satisfaction as well as loyalty. This satisfaction then leads to overall customer value creation hence contributes to the overall business and profitability of the organisation.

2.1.17   360 Degree/Multi-Rater Performance Appraisal

360-degree method is a process that involves views of different group of reviewers who are associated with the employee in examination. This reviewer may be mangers, supervisors, customers, clients, peers or any other stakeholders who is directly or indirectly involved or affected by the employee's action or work. It also includes self-appraisal by the employee by which the employee expresses his/her opinion, feedback about the work undertaken by him (Idowu, 2017). The 4 assessments which are must to implement 360 Degree method are – self-assessment, immediate supervisor assessment, peer assessment and sub-ordinate assessment (Horng, 2011). Although numerous advantages have been listed down in the literature, there are certain drawbacks which are identified by some researchers. The use of single kind of expression numerical or linguistic limits the ability to gather the quality of data that is been gathered by all the reviewers (Espinilla, 2013). Also, the accurate interpretation of all the reviews is hard to make as it is difficult for quantitative data to exactly represent qualitative data accurately (Hsu, 2005)).

2.1.18   Management by Objectives (MBO)

MBO is a method in which the goals of the performance appraisal system are defined by the mutual consensus of all the key stakeholders which includes peers, sub-ordinates as well as supervisors. There are several steps that needs to be followed in the method starting with establishing a set of well-defined objectives that an employee needs to achieve followed by developing a detailed action plan indicating the way by which the objective is to be achieved, then the employee then moves forward with the decided action plans (Idowu, 2017). Shortcoming of these methods include the fact that it does not allow the reviewer to see how the employee is dealing with the events and issues hence their approach or way may not represent most effective use of the resources. Also, there is a very little scope for comparative study as no benchmark is present for changing workplace environment (Bipp, 2011).

In the discussion regarding several methods used for appraisal in organisations, one common thing that we observed is in almost all the methods is that the rating of an employee completely depends on the perception of the employee in the eyes of the supervisor/manager and also a lot of methods still rely on

parameters which are perceived differently by different managers thus introducing subjectivity in the process and hence it may introduce biases and errors in the overall performance appraisal process. Researches have showed that managers who are rating their employees are influenced by some factors which are beyond the employee's control, resulting in subjectivity in the process and hence introducing biases and errors in the performance appraisal process (Cardy, 1998). Even managerial community in Indian organisations are not very satisfied with the prevailing performance appraisal system, dissatisfaction is more in public sector companies compared to private sector companies (Singh, 1981). Evaluation errors during judgement affects the accuracy and relevance of Performance Management. Also, inaccurate and less effective performance appraisal system results in increase in dissatisfaction levels of employees (Mahmoud Javidmehr, 2015). Presence of subjectivity in the performance appraisal process makes it difficult to differentiate between employees and hence result in problems in decision making and followed by difficulty in selections for promotion and other incentives (Moers, 2005).The techniques used for evaluating performance of employees in organisations are very old and traditional and needs to be updated and modernized for the wellbeing of the employee and organisation as a whole (Sastry, 2018).

Immense amount of research has been carried in the field of performance management system and its limitations and hence, several biases have been established in the literature.

2.2. Performance Appraisal Biases
2.2.1 Leniency and Strictness Bias

Leniency bias is the tendency of a manager or supervisor to give relatively high rating to almost all the employees depending upon the standards set by them and their physical and mental perception at the time of the appraisal discussion (Mahmoud Javidmehr, 2015). Empirical studies have also shown that leniency bias has affected evaluation related decisions such as promotions and bonus roll-outs. strictness bias as when managers believe in exact evaluations, focusing more on the weaknesses of an employee and thus making the evaluation excessively strict (Mahmoud Javidmehr, 2015). Leniency and Strictness bias are easily visible in graphical methods of rating and thus methods where employees are been ranked in comparison with performance of other employees is a better alternative and can reduce the mentioned biases.

Leniency bias is associated to the behaviour of the evaluator rather than the work and efforts put by the employee (Vance, 1983). Leniency bias is similar to heuristics in the investment's decision. Due to Leniency bias most of the employees are rated higher than what they actually deserved (Raveendra, 2018).

2.2.2   Central Tendency Bias –

Central tendency bias is the attitude of the manager to refrain from extreme judgement of employees and thus results in rating all the employees as average

irrespective of their individual performances (Mahmoud Javidmehr, 2015). One of the recommendations to tackle this bias is the use of ranking method where all employees cannot be given a common rank. Central tendency bias can be compared to herding in investment decisions as lack of proper knowledge about the employees and tendency to play safe are the main reasons for this type of behaviour by the manager (Raveendra, 2018). Central tendency bias totally defeats the sole purpose of performance appraisal system of differentiating between good performers and poor performers (Singh, 1981).

2.2.3   Contrast Bias and order effect:

Contrast bias occurs when a manager has to rate or evaluate a large number of employees under him, hence in the process of evaluation the manager tends to compare performance of employees with one another forgetting about the ideal standards set by the organisation as a result some of the employees may end up at the lower end of the rating even his performance was well above organisation standards (Mahmoud Javidmehr, 2015) (Raveendra, 2018) . Order effect says that employees rated first are mostly rated higher compared to employees who are rated at last, the effect increases with increase in the time interval between evaluations.

2.2.4   Discrimination between insider and outsider Employees/ Contract Bias:

There is a bias that exists between insider and outsider groups (may be employees on contract). Managers tend to trust insiders more than outsider groups thus providing them challenging work and hence receive higher ratings however outsider group employees tend to receive trivial and less important tasks (Mahmoud Javidmehr, 2015). Such biases occur due to mental blocks developed in the mind of the evaluator for outsiders and insiders (Raveendra, 2018).

2.2.5   Subjective Bias/ Personal Bias –

Managers tend to execute subjective biases when they allow their personal biases and prejudices to impact and influence the evaluation of the performance of the employees under them (Singh, 1981). The perception of the manager about all the employees working under him, whether the manager likes or dislikes them has a significant impact on rating their performances (Malik, 2018).

2.2.6   Gender Bias –

Gender Bias refers to the tendency of a manager to rate employees of a particular genders higher or lower as compared to other gender. However, it is very difficult to conclude anything about gender bias amounting to inconclusive results amounting to contrasting studies (Millmore, 2007). Good performing female employees does not always receive deserved ratings as their performance is evaluated on parameters other than ability (Nieva, 1980). A

contrary conclusion to this is performance ratings tend to favour female employees compared to their male counterparts (Jacobson, 1974).

Leniency and strictness biases introduces skewness in the performance appraisal process resulting in a less effective evaluation system. Improvement in employee performance directly results in improvement in organisational performance (Mahmoud Javidmehr, 2015). Organisational success is directly impacted by employees' performance (Gardner, 2011). Customer retention is possible through employee retention which is possible through employee satisfaction and one of the critical components of employee satisfaction is effective performance appraisal system without biases (Raveendra, 2018). Also, the existence of biases is inversely related to the quality of the outcome from appraisal system, more the biases, less the quality of the results.

3. Research Questions

The purpose of the present paper is to explore two research questions:

RQ1. How can we quantify different performance appraisal biases and list down all the managers exhibiting those biases using analytics on appraisal data?

RQ2. How to identify subjective bias in a manager using the accuracy derived after considering the ratings before and after review?

To explore the above-stated research questions, the paper in the following sections, lists down step wise methods that needs to be followed to arrive at lists of managers exhibiting the respective biases.

4. Research Methodology

In the research paper, we have worked on primary data collected from a questionnaire. Which was filled by employees with more than 2 years of experience in any organisation which followed a 6-point rating scale. The form meant to collect the information regarding their manger, past ratings as well as certain demographic information.

First, we established a method for quantifying a bias for a year, and then tried to extend it to other years, if we find consistency in our findings for all the years, we conclude that the selected method can be used to detect and quantify respective biases. Tools which are used in the process are Chi-Square Analysis, Box-Plot and Outlier detection and Pearson Coefficient of Skewness. In the research paper we have used chi-square to test for a significant relation between high performance rating and gender of an employee, as well as relation between ratings and the type of employee (permanent or contractual).

4.1. Chi-square test

Chi-square test is a non-parametric test used for comparing a sample variance with a theoretical variance. It is implemented on categorical variables (e.g. Gender – Male or Female) (Rana, 2015).

4.2. Box-Plot

Also known as box and whiskers plot is statistical technique mainly used as an exploratory data analysis technique for identifying hidden patterns from the data. (Williamson, 1989). It is normally a representation of quartiles and the visualisation contains a maximum value, upper quartile, median, lower quartile and minimum value (Thirumalai, 2017).

4.3. Outlier

An outlier is generally defined as the data point that is far outside the norm for a given distribution of datapoints of a variable or a population (Osborne, 2004).

4.4. Pearson Coefficient of Skewness

it is a statistical method used to determine the amount of skewness of a data compared to the ideal normal distribution curve (Salma Shorna, 2019).

4.5. Description of Key terms used

4.5.1   Pre-TR rating

This is the rating given to employee by the manager

4.5.2   Post-TR rating

This is the final rating that the employee gets after the Pre-TR ratings have been reviewed by a Talent Review (TR) bench

4.5.3   Team Size

It is the number of employees under the supervision of the manager

4.5.4   Weighted Average

It is a kind of mean that takes varying importance of the numbers, known as weights into account while calculating the results

4.5.5   Odds Ratio

In our case odds ratio is defined as the odds/chances of a particular gender of getting high ratings

4.5.6   Performance Ratings

These are the scores given by a manager or supervisor to the employees under him as an outcome of their performance evaluation. We have considered a 6-point rating scale in the analysis. The rating scale consists of 6 ratings – (E, A+, A, B+, B, C), where E stands for the best rating (more than expected performance) and C stands for the worst rating (performance need improvement).

5.   Results and Analysis

5.1. Leniency, Central Tendency and Strictness Bias

To calculate Leniency, Central Tendency and Strictness bias we will consider the 3 years ratings of employees, Pre-TR ratings of the employees and Managers having team size more than or equal to 5.

Step 1- Calculate the average ratings given by each manager for all the 3 years.

Table 1: A sample of the output after step1 showing number of employees (Team Size) and their average rating under each manager for a period of 3 years

| Manager _ID	TeamSize _2019	AvgRating_ 2019	TeamSize _2018	AvgRating_ 2018	TeamSize _2017	AvgRating_ 2017
939946	10.00	4.90	4.00	5.00	1.00	4.00
925891	5.00	4.80	5.00	4.20	4.00	4.25
926160	7.00	4.14	7.00	4.29	3.00	4.33
927634	4.00	4.25	4.00	4.25	4.00	4.25
927949	4.00	4.50	6.00	4.00	7.00	4.29

Step 2 – Considering the team size, calculate the weightage average rating against each manager.

Table 2: A sample of the output after step 2 showing weighted average calculated of 3 years for each manager

Manager ID	Weighted Average
1432036	3.89
1432004	4.33
1431942	4.40
1431940	4.15
1431926	4.80
1431838	4.54
1431818	5.00
1431789	4.50
1431764	4.70
1431753	4.00

Step 3 – Plot a box-plot of all the weighted average ratings.

Figure 1. Box-plot distribution of the weighted average ratings obtained in step 2
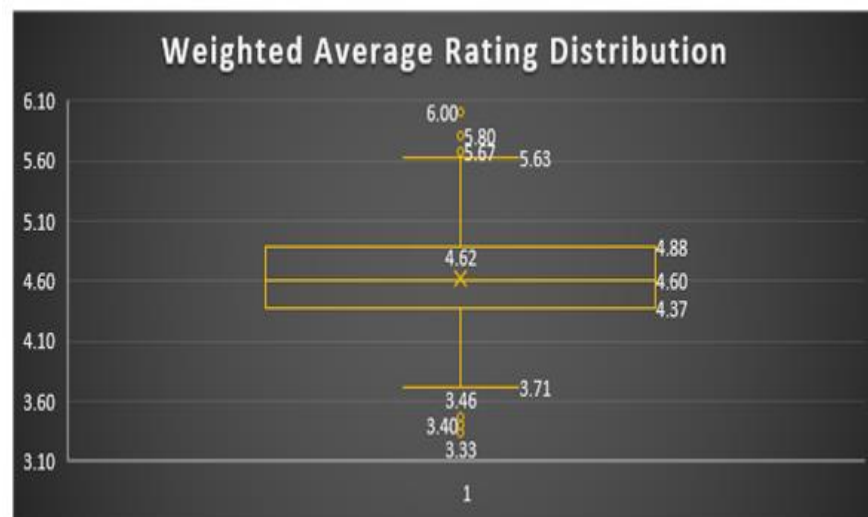
As can be seen from the box plot, rating distribution follows an ideal normal distribution curve with mean equals to median (4.62 and 4.60 respectively).

An important conclusion that we can form looking the box-plot is that most of the managers are lenient if we consider the overall rating distribution. The ideal average of the ratings comes out to be 3.5 (average of ratings scores 1,2,3,4,5,6) but our distribution shows an average of 4.62 which substantiates the leniency behaviour of the managers.

To identify the managers who possess either Leniency bias or Strictness bias, we can take the help of outliers from the box-plot and formulise a rule that if the weighted average rating of a manager is above 5.63 (maximum) then the manager is definitely showing Leniency bias and if the weighted average for the manager is below 3.71 (minimum) then the manager is surely showing strictness bias.

An alternate approach to find out the mentioned biases is to go by definition. Leniency bias arises when manager always give good ratings (A+ and E) hence all the managers whose weighted average lies between 5 (A+) and 6 (E) can be termed as managers showing potential behaviour of leniency.

For central tendency bias we take all the managers whose weighted average rating is distributed around overall weighted average rating (4.62) considering the entire dataset. Hence, we say that managers having weighted average rating in the range of (4.58,4.66) i.e. 1% above and below the overall weighted average rating, are the managers showing central tendency behaviour and needs to be flagged-off.

Organisations can also change these thresholds to suit their companies' culture and requirements and to suit the objectives they define for performance appraisal.

5.2. Contrast Bias

To identify contrast bias, we will try to evaluate the percentage difference between the Pearson coefficient of Skewness for all the employees and for the employees who are in teams with size more than 5. Below tables outline the results.

Table 3,4,5 shows results highlighting percentage difference in Pearson coefficient of skewness for all the employees compared to those with team size more than 5 for year 2019,2018 and 2017 respectively.

Table 3: Year-2019

Year-2019	Pearson Coefficient of Skewness	% Difference
All the Employees	-1.21	
Employees team size more than 5	-1.29	6.54%

Table 4: Year-2018

Year-2018	Pearson Coefficient of Skewness	% Difference
All the Employees	-1.26	
Employees team size more than 5	-1.33	5.08%

Table 5: Year-2017

Year-2017	Pearson Coefficient of Skewness	% Difference
All the Employees	-1.47	
Employees team size more than 5	-1.51	2.34%

We observe that for all the 3 years, the percentage difference is in the range of 6% to 7% and hence can be concluded that for the data under observation, contrast bias does not exist

5.3. Subjective Bias/ Personal Bias

For identifying the list of managers practising subjective or personal bias, we will introduce a new concept called Accuracy and define it as the percentage of employees for whom the manager was able to defend the ratings in front of the review panel out of the total employees under his supervision.

Table 6: Manager wise accuracies calculated using team size and number of ratings defended by the manager in the review

Manager ID	Team Size	Number of Defended Ratings	Accuracy
1431838	5	2	40.00%
1428296	6	3	50.00%
1428253	5	3	60.00%
1427906	5	0	0.00%
1427882	5	2	40.00%
1427819	6	3	50.00%
1412063	7	6	85.71%
1412056	5	4	80.00%
1412054	5	4	80.00%

Once we have a dataset (as Table 6) consisting of accuracies for all the managers we can find out very easily, a list of managers having accuracy below some threshold as decided by the company or the organisation. In our case we have considered 60% as the threshold.

5.4. Gender Bias

First, we will try to check and establish whether the organisation as a whole possess any gender bias and if yes, what is the odds ratio. After checking for gender bias for the entire dataset we will try to establish a method through which we can generate a list of managers showing gender bias.

We will start with using Chi-Square method to test if there is any significant relation between high performance ratings and the gender.

Hence, we will apply chi-square test on two categorical variables – High performance (if rating more than or equal to 5 i.e. A+ or E then 1 otherwise 0), Gender (If Male then 1 otherwise 0).
We found that for all the 3 years under examination, p-value is well below 0.05 (95% Confidence Interval) and hence there is a significant relation that exists between high ratings and the gender.

```
        Pearson's Chi-squared test

data:  performance_2018$Gender and performance_2018$high_performer
X-squared = 17.164, df = 1, p-value = 3.429e-05
```

Figure 2. A sample output for the Year-2019 showing the Chi-Square results

```
Point estimates and 95% CIs:
--------------------------------------------------------------
Inc risk ratio                              1.21 (1.10, 1.34)
Odds ratio                                  1.38 (1.18, 1.60)
Attrib risk *                               7.61 (4.11, 11.11)
Attrib risk in population *                 6.74 (3.26, 10.21)
Attrib fraction in exposed (%)              17.65 (9.29, 25.24)
Attrib fraction in population (%)           15.95 (8.28, 22.99)
--------------------------------------------------------------
 Test that OR = 1: chi2(1) = 17.164 Pr>chi2 = <0.001
 Wald confidence limits
 CI: confidence interval
 * Outcomes per 100 population units
```

Figure 3. Odds Ratio output for the Year-2019

Table 7: Summary of Chi-square test and odds ratio for all 3 years

Year	P-Value	X-Squared	Odds Ratio	DF
2019	3.43E-05	17.164	1.38	1
2018	0.0005975	11.784	1.33	1
2017	0.002372	9.237	1.32	1

Considering the odds-ratio for the year-2019, we say that 1.38 indicates that the chances of a male getting a high rating is 38% more than the female considering all other parameters to be constant.

5.5. Manager Level – Gender Bias

Since now we know that gender bias exists in the dataset, we will try to find out the managers who are showing gender bias. Analysis is done to find the average ratings of male and female employees under each manager.

Table 8: Percentage difference in average ratings of males and females given by each manager

Manager ID	Male	Male Average Rating	Female	Female Average Rating	% Difference
893866	2	5.00	2	4.00	20.00%
896601	59	4.66	17	4.00	14.18%
897196	3	5.33	3	4.33	18.75%
897234	5	4.60	3	4.33	5.80%
897331	14	4.00	4	4.00	0.00%
897332	8	4.75	4	4.00	15.79%
897879	2	6.00	4	5.25	12.50%
900998	7	4.29	2	4.00	6.67%
901432	12	4.58	2	4.00	12.73%

% Difference column indicates the difference in the average rating of male and female evaluated by the same manager.

Once we have the dataset as Table 8, we will keep only those managers who have more than or equal to 2 males and 2 females so that the results are not skewed. Finally, we will flag-off all the managers having a % difference in rating of more than 20%.

6. Discussions

This research paper aims to quantify different biases existing in the performance appraisal system in an organisation and hence make an effort to reduce the subjectivity already present in the performance appraisal system. This research paper tries to combine two different fields of study – Analytics and Performance Appraisal system. Analytics predominantly belongs to Information Technology Domain (Sharma, 2017) (Pemmaraju, 2007) and performance appraisal system belongs to Human Resource Management domain. This research has used exploratory data analytics coupled with some existing statistical tools to establish different methods and approaches to measure and quantify different kinds of performance biases present in the system. It also tries to address the future scope of a lot of researches to deal with subjectivity in the performance appraisal system by suggesting to leverage the potential of data and HR analytics as the way forward for bringing objectivity in the process. (Sharma, 2017) (Moers., 2005). Now is the time for the organisations to shift from their intuition and memory-based   performance evaluation to fact and data-based evaluation and decision making.

7. Managerial Implications

The implications of this paper are numerous. The paper highlights the limitations of performance appraisal and in the process develops some background regarding biases that exists in the system and how to find out who are exhibiting those biases and who are mostly getting affected by those biases. This paper can be used by all the HR professionals in the organisation to gauge the effectiveness of their performance appraisal system. It can be used by Head HR, Head PMS, CHRO, HR-BP's and everyone who is involved in strategy making with regards to employees' performance, employees' career management, talent mapping, promotions, team composition, diversity analysis staffing, talent recruitment analysis, learning and development planning as well as internal job posting decisions. The results after implementing the described methods can also be used by HR departments to identify managers who are not objectively evaluating employees under their supervision, who are considering factors other than the performance for rating and evaluating an individual and hence devising a strategy around training them and work towards a more transparent, more objective and hence, a more effective Performance Appraisal System.

8. Conclusion

In the research, we found ways through which we can quantify various biases existing in the literature as well as are found in most of the organisations. In our analysis we found out that number of managers showing Leniency Bias is highest at 20% in line with the study (Moers., 2005) then Central Tendency Bias at 12% followed by Strictness Bias at 3 %. However, our data failed to establish the presence of contrast bias in the data, it may not be the case with all the organisations and hence organisation can use the approach used in the paper on their respective appraisal data before any conclusions. Through the analysis of Subjective Bias, we found that 52% of the managers have an accuracy of less than 60% which is an alarming finding as it indicated the amount of subjectivity and hence the bias present in the performance appraisal system. Through the detailed study for gender bias we also found that 20% of the managers shows gender bias or in the other words 20% of all the managers eligible for this study (Number of Male and Female employees under them each is greater or equal to 2), had deviation of Male and Female average ratings of more or equal to 20%.

In the recent times, competition between organisations have increased tremendously and hence the goals and objectives of organisations have become more complex and dynamic. In order to fulfil these, organisations need to rely on their employees who needs to be relevant and updated amidst the continuously and ever evolving working climates. This requires a lot of commitment, hard work and efforts from employees and hence employees expect a similar or better commitment from their employer. To have a proper check on the performance of employees and to know if the organisation is performing as per the goals and milestones set at the start of the year, performance appraisal is the best tool that an employer can use. An effective and bias free performance appraisal system ensures high motivation levels and engagement levels of employees throughout the year and hence it is the key and critical factor in deciding the overall growth and development of the organisations. The trend has already started as organisations are now motivated enough for exploring tremendous amount of opportunities and recruiting the right talent for bringing sophistication in HR functions (Fink, 2017).

9.   Limitations

Regardless of this study's efforts and findings about the various kind of appraisal techniques and system present in the literature and thereby answering the question how to quantify those biases, a number of biases must be noted. It is important to know that HR analytics as a potential solution depends entirely on the quality of data which is a limitation of these approach. (Sharma, 2017). Also, the research was confined to only some specific biases, although numerous different biases are already established in the literature such as Halo-effect, Recency bias, Spill-over effect, proximity bias, Recency bias (Mahmoud Javidmehr, 2015) (Raveendra, 2018) and hence future researchers can think of devising a method using the statistical tools available, to quantify the remaining biases so that the outcomes from the overall study can be used by the organisations to bring in objectivity in the performance appraisal system.

References

Fink, A. A., & Sturman, M. C. (2017). HR Metrics and Talent Analytics. The Oxford Handbook of Talent Management, 375-390.

Anshu Sharma, T. S. (2017). HR analytics and performance appraisal system. Management Research Review.

Armstrong, M. (2009). Armstrong's Handbook of Performance Management: An Evidence-Based Guide to Delivering High Performance.

Atiomo, A. (2000). Human Resource Management. Lagos: Malthouse Management Science Books.

Ayers, R. S. (2015). Aligning individual and organizational performance: Goal alignment in federal government agency performance appraisal programs. Public Personnel Management, 44(2), 169-191.

Bagul, D. B. (2014). A research paper on "study of employee's performance appraisal system". Scholarly research journal for humanity science and English language, 1, 287-292.

Bednall, T. C., Sanders, K., & Runhaar, P. (2014). Stimulating informal learning activities through perceptions of performance appraisal quality and human resource management system strength: A two-wave study. Academy of Management Learning & Education, 13(1), 45-61.

Bipp, T., & Kleingeld, A. (2011). Goal- setting in practice. Personnel Review.

Thirumalai, C., Vignesh, M., & Balaji, R. (2017, April). Data analysis using Box and Whisker plot for Lung Cancer. In 2017 Innovations in Power and Advanced Computing Technologies (i-PACT) (pp. 1-6). IEEE.

Williamson, D. F., Parker, R. A., & Kendrick, J. S. (1989). The box plot: a simple visual method to interpret data. Annals of internal medicine, 110(11), 916-921.

Malik, R. (2018). Study of Performance Appraisal Systems in Service Industry in India.

Raveendra, P. V., & Singh, P. PERFORMANCE APPRAISAL BIASES AND BEHAVIORAL BIASES IN DECISION MAKING: AN EMPIRICAL STUDY. International Journal of Mechanical Engineering and Technology.

Jansirani, S., Hatrikrishnan, R., Jaya, D., & Saisathya, A. (2013). A Study on Performance Appraisal System at Wipro Infrastructure Engineering Pvt Ltd. IOSR Journal of Business and Management, 9(3), 8-23.

Espinilla, M., de Andrés, R., Martínez, F. J., & Martínez, L. (2013). A 360-degree performance appraisal model dealing with heterogeneous information and dependent criteria. Information Sciences, 222, 459-471.

Gardner, T. M., Wright, P. M., & Moynihan, L. M. (2011). The impact of motivation, empowerment, and skill- enhancing practices on aggregate voluntary turnover: The mediating effect of collective affective commitment. Personnel psychology, 64(2), 315-350.

Jeou-Shyan, H., Hsuan, H., Chih-Hsing, L., Lin, L., & Chang-Yen, T. (2011). Competency analysis of top managers in the Taiwanese hotel industry. International Journal of Hospitality Management, 30(4), 1044-1054.

Salma Shorna, U., & Hossain, F. (2019). A New Approach to Determine the Coefficient of Skewness and An Alternative Form of Boxplot. arXiv, arXiv-1908.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. Science, 310(5754), 1680-1683.

Idowu, A. (2017). Effectiveness of performance appraisal system and its effect on employee motivation. Nile Journal of Business and Economics, 3(5), 15-39.

Jacobson, M. B., & Effertz, J. (1974). Sex roles and leadership: Perceptions of the leaders and the led. Organizational behavior and human performance, 12(3), 383-396.

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). Practical Assessment, Research, and Evaluation, 9(1), 6.

Javidmehr, M., & Ebrahimpour, M. (2015). Performance appraisal bias and errors: The influences and consequences. International Journal of Organizational Leadership, 4, 286-302.

Millmore, M., Biggs, D., & Morse, L. (2007). Gender differences within 360-degree managerial performance appraisals. Women in Management Review.

Mishra, L. (2013). A Research Study on Employee Appraisal System-Case of Hong Kong and Shanghai Banking Corporation (HSBC Bank). International Journal of Business and Management Invention, 2(2), 60-67.

Moers, F. (2005). Discretion and bias in performance evaluation: the impact of diversity and subjectivity. Accounting, Organizations and Society, 30(1), 67-80.

Kumar, M., Shirley, D.N., & Singh, D.G. (2017). A Research Paper on "Employee's Performance Appraisal System and its Implication for Individual and Organizational Growth".

Nieva, V. F., & Gutek, B. A. (1980). Sex effects on evaluation. Academy of management Review, 5(2), 267-276.

Oladimeji, A. (1999). Human resource management in Nigeria. Lagos: Business and Institutional Support Associates Limited.

Singh, P., Maggu, A., & Warrier, S. K. (1981). Performance Appraisal Systems: A Critical Analysis. Indian Journal of Industrial Relations, 315-343.

Pemmaraju, S. (2007). Converting HR data to business intelligence. Employment Relations Today, 34(3), 13-16.

Rana, R., & Singhal, R. (2015). Chi-square test and its application in hypothesis testing. Journal of the Practice of Cardiovascular Sciences, 1(1), 69.

Cardy, R. L., Sutton, C. L., Carson, K. P., & Dobbins, G. H. (1998). Person and system effects in performance appraisal: Ratings as a function of the degree of performance responsibility and errorfulness. Journal of Quality Management, 3(1), 79-99.

Sastry, D. R. (2018). A Study of Performance Appraisal and Employee‟ s Performance by Measuring the Latest Role of Motivation in an Organization. International Journal of Trend in Research and Development Volume 5(6), ISSN: 2394-9333.

Traub, L. (2013). Bias in Performance Management Review Process.

Vance, R. J., Winne, P. S., & Wright, E. S. (1983). A longitudinal examination of rater and ratee effects in performance ratings. Personnel Psychology, 36(3), 609-620.

Xavier, J. V. (2015). A Study on the Effectiveness of Performance Appraisal System and its Influence with the Socio-Demographic Factors of the Employees of a Manufacturing Industry in Tamil Nadu. International Journal of Research in Management & Business Studies, 2(1).