Project Dissertation Report on

PREDICTING EMPLOYEE'S PERFORMANCE USING HR ANALYTICS

Submitted by:

Mayuri Verma

Roll No: 2K19/DMBA/051

Under the Guidance of:

Asst. Prof. Meha Joshi

Delhi School of Management



DELHI SCHOOL OF MANAGEMENT Delhi Technological University

Bawana Road, Delhi – 110042

CERTIFICATE

This is to certify that the work titled 'Predicting Employee's Performance using HR **Analytics'** as part of the final year Major Research Project submitted by Mayuri Verma in the 4th Semester of MBA, Delhi School of Management, Delhi Technological University during January-May 2021 is her original work and has not been submitted anywhere else for the award of any credits/ degree whatsoever.

The project is submitted to Delhi School of Management, Delhi Technological University in partial fulfillment of the requirement for the award of the degree of Master of Business Administration.

Prof. Archana Singh

Asst. Prof. Meha Joshi

Faculty Advisor

Head of Department (DSM, DTU)

DECLARATION

I hereby declare that the work titled 'Predicting Employee's Performance using HR Analytics' as part of the final year Major Research Project submitted by me in the 4th Semester of MBA, Delhi School of Management, Delhi Technological University, during January-May 2021 under the guidance of Asst. Prof. Meha Joshi is my original work and has not been submitted anywhere else.

The report has been written by me in my own words and not copied from elsewhere. Anything that appears in this report which is not my original work has been duly and appropriately referred/cited/acknowledged.

Mayuri Verma

(Roll No 2K19/DMBA/051)

ACKNOWLEDGMENT

It is a great pleasure for me to acknowledge the kind of help and guidance received during the research work. I would like to thank my faculty advisor Asst. Prof. Meha Joshi, who helped me to take up the topic 'Predicting Employee's Performance using HR Analytics' and guided me to complete this project properly. The project provided me with an excellent opportunity to explore the areas of Performance Management and HR Analytics.

I am highly indebted to Delhi School of Management, Delhi Technological University for giving me an opportunity to work on this project. Lastly, I would like to express my gratitude to all the honorable faculty members for sharing their experience and expertise on this project.

I have put all my efforts to ensure that the project is completed in the best possible manner and also ensured that the project is error-free.

Mayuri Verma

(Roll No. 2K19/DMBA/051)

iii

ABSTRACT

In light of unpredictable business sectors and uplifted financial vulnerability, firms have put expanded focus on individual, group and company performance. For sure, one could consider the HR analytical capability to foresee performance in future as the sacred goal of HR professionals. The capability to understand, foresee, distinguish and impact undeniable degrees of individual or group performance can be considered basic to improving performance of the company and driving employee's strategy.

Performance Management Systems in firms has advanced throughout the most recent time from an independent performance evaluation process to a more coordinated arrangement of measures and cycles, including client input, marketing projections, group efficiency and 360-degree criticism, to give some examples. The pattern has likewise been to move duty regarding performance management away from the HR work towards line supervisors. Though, this move has offered line supervisors the chance to assume greater liability for the performance and development of their group members, it has presented more prominent variety in the individual employee's experience, hence setting more noteworthy focus on the strength of the performance appraisal framework and the attitude and competence of line supervisors to view their obligations towards their team members in a serious way.

The drawback of this decentralized cycle is that it is frequently hard to get one single perspective on the performance, all things considered, and people across the firm and across the entirety of the pointers of performance evaluation. In reality, how do senior pioneers realize who is doing admirably across these, what they are doing well, and where would it be advisable for them to contribute their, regularly restricted, learning and improvement investments to guarantee the best return? How would we utilize this data to have the greatest effect to the main concern?

TABLE OF CONTENTS

Ce	ertificate	i
De	eclarationeclaration	ii
Ac	cknowledgement	iii
Al	bstract	iv
	able of Contents	
Li	st of Figures	vii
1.	Introduction	1
	1.1. Background	1
	1.2. Objectives of the Research	2
	1.3. Scope of the Research	2
	1.4. Structure of the Research	3
2.	Literature Review	4
	2.1. HR Analytics	4
	2.2. Application of HR Analytics	5
	2.3. Future of HR Analytics.	5
	2.4. Concluding Remark	6
3.	Research Methodology	7
	3.1. Data Coverage	7
	3.2. Tools & Techniques	7
	3.2.1 Multiple Linear Regression.	7
	3.2.2 Stepwise Multiple Regression	9
	3.3 Concluding Remark	10

4.	Empirical Analysis	11
	4.1 Exploratory Data Analysis	11
	4.1.1 Count of number of Employees by Performance Rating in 2014	13
	4.1.2 Count of number of Employees by Performance Rating in 2015	14
	4.1.3 Count of number of Employees by Number of Sick Days in 2014	15
	4.1.4 Count of ID by Gender	15
	4.1.5 Count of number of Employees by Job Tenure in years	16
	4.1.6 Count of ID by POS2014	17
	4.1.7 Count of number of Employees by Job Strain in 2014	18
	4.1.8 Count of number of Employees by Distributive Justice Perception in 2014	19
	4.1.9 Count of number of Employees by Procedural Justice Perception in 2014	20
	4.2 Predictive Modelling using SPSS Statistical Software	22
	4.2.1 Multiple Linear Regression based Prediction Model	24
	4.2.2 Stepwise Regression based Prediction Model	27
	4.3 Concluding Remark	34
5.	Recommendations & Conclusion	35
	5.1 Concluding Remark	36
Re	eferences	37

LIST OF FIGURES

Figure 4.1 Count of ID by Performance Rating2014	13
Figure 4.2 Count of ID by Performance Rating2015	14
Figure 4.3 Count of ID by SickDays2014	15
Figure 4.4 Count of ID by Gender	16
Figure 4.5 Count of ID by JobTenure2014	17
Figure 4.6 Count of ID by POS2014	18
Figure 4.7 Count of ID by JobStrain2014	19
Figure 4.8 Count of ID by DistJust2014	20
Figure 4.9 Count of ID by ProcJust2014	21
Figure 4.10 Data-View in SPSS Software	23
Figure 4.11 Variable-View in SPSS Software	23
Figure 4.12 Analyze, Regression and Linear Options in SPSS	24
Figure 4.13 Moving Independent and Dependent Variables to respective selection boxes	25
Figure 4.14 Output of Multiple Linear Regression Model in SPSS	25
Figure 4.15 Output of Multiple Linear Regression Model in SPSS	26
Figure 4.16 Analyze, Regression and Linear Options in SPSS	28
Figure 4.17 Moving Independent and Dependent Variables to respective selection boxes	29
Figure 4.18 Output of Stepwise Regression Model in SPSS	29
Figure 4.19 Output of Stepwise Regression Model in SPSS	30
Figure 4.20 Output of Stepwise Regression Model in SPSS	31
Figure 4.21 Output of Stepwise Regression Model in SPSS	31

Chapter 1

INTRODUCTION

HR analytics assists businesses in discovering any existing gaps in employee performance and works to help them address those gaps, utilizing the information that data-analytics offers. With better information comes better ability to plan and planning well can help set the founding stone for the most ideal performance outcomes. This chapter discusses what all factors can we measure to indicate performance along with defining the objectives, scope, and structure of the research.

1.1 Background

One of the vital difficulties with carrying out performance evaluation with HR information is choosing what kind of performance to attempt to foresee. One may accept that the yearly or half-yearly performance evaluation rating ought to be something we center around attempting to predict. In any case, while there may be valid justifications why this is something we could attempt to do, there are different kinds of performance related indices that can likewise give some very valuable data. What is significant here is to recall that we not only need to evaluate the performance, however to consider what factors we may have that might be an indicator of employee performance.

Moving away from assembling ventures to the service industries, consumer loyalty has become a significant measure that we could zero in on as an indicator of performance. In the service industry, consumer loyalty, client dependability and the probability that clients reinvest could be viewed as solid markers of performance, in light of the fact that these variables eventually sway the main concern.

Numerous companies regularly gather consumer loyalty information; notwithstanding, it is uncommon that these companies have the HR analytical ability to have the option to completely comprehend and investigate what indicators may foresee these key performance results.

While measures like staff turnover, absence due to sickness and commitment scores may not clearly be indicators of performance from a conventional perspective, these factors are all measures of performance of a sort.

High staff turnover, high sickness absence rates and low commitment will all be related with an unfavorable change in the company's primary concern and are everything that a fruitful HR capacity will need to stay away from. Along these lines, they would all be able to be viewed as indicators of performance that we should attempt to systematically anticipate.

In some cases, data that we would need to consider for our insightful model isn't accessible at the individual level. Singular commitment scores may not generally be accessible because of the confidentiality of most employee reviews. Additionally, with performance, it may not generally be proper to pass judgment on performance at an individual level, particularly when there might be groups engaged with the work interaction. As such we should zero in group level performance or the like.

The real kind of investigation and models that we can run will be reliant upon the accessibility of the information that we have, and firms will vary in a general sense in this. For instance, a firm that includes some type of selling could zero in on getting new clients and the measure of income produced straightforwardly from deals. A call community, for instance, would zero in available to come in to work rates, call loads and call term.

1.2 Objectives of the Research

The major objectives for carrying out the research work are:

- To find out factors which can be measured to indicate Performance
- Use Multiple Linear Regression to predict an Employee's Performance
- Use Stepwise Multiple Linear Regression to Model Performance

1.3 Scope of the Research

The research has been conducted to study various factors which can be measured to indicate performance of an employee. A comprehensive study of the application of HR Analytics in measuring Employee's performance has been carried out to draw results.

Multiple Linear Regression and Stepwise Multiple Linear Regression methods have been used to predict Employee's Performance.

1.4 St	tructure of the Research
	ter 1 discusses the introduction, followed by chapter 2 that brings out a review of the existing ture about the application of HR Analytics in Performance Management. Chapter 3 present
	esearch methodology discussing the tool and techniques used. Chapter 4 brings out the
	rical analysis followed by chapter 5 recommendations and conclusion.
cmpn	ical analysis followed by enapter 3 recommendations and conclusion.

Chapter 2

LITERATURE REVIEW

The goal behind writing the literature review is to give a more extensive perspective on scholarly writing accessible on subject of HR analytics, how analytics can be coordinated with functions being carried out in the companies, and further extent of research that studies published have provided to scholarly community and industry experts. For accomplishing the expressed exploration questions, the literature review will give a possible highlight for conversation about the investigated topic.

2.1 HR Analytics

Baron A (2011) the point of this paper was to recognize how HR experts can push ahead towards the estimation of human resources. There are a few territories recognized in which information is valuable to the executive's activity. How human insight is being pulled in, selected and held; sustained and used; remunerated and roused; how information and execution are together overseen. There are three components of estimating human resources: proficiency of HR work, viability of individuals cycles and estimating the effect or return of speculation.

Bassi (2011) recommended HR examination comprises from essential report of HR data or estimations to make forecasts identified with HR. According to **Saraswathy et al** (2016) HR Analytics is strategy for comprehension and surveying the causal connection between HR rehearses and authoritative execution results by applying measurable apparatuses and investigations which depended on required measurements of proficiency, adequacy and effect.

Venkatakrishnan C et al (2017) Human asset examination will profit associations by taking existing operational information and changing it over to dynamic way to deal with future HR issues. Test information source which are of insightful use incorporates Electronic, online media, portable sound data, versatile video appraisal catch, video talk with maintenance, virtual acceptance, social communication.

2.2 Application of HR Analytics

Harris et al (2011) Analytical techniques have capacity to help leaders to relate contributions to human resources to an organization's results in their monetary capital. HR pioneers can effectively define association's future through overseeing basic ability and organizing programs towards long haul prerequisites of business.

Kong et al (2013) analyzed the part of HR rehearses for creating skills for development in Indian IT area. For satisfying the examination destinations incompletely planned inside and out interviews with senior chiefs of ITSPs in India were directed. The outcomes featured that improvement of information and learning capacities structure a significant piece of ability the board design. Likewise, the association need to create HR methodologies to oversee enormous pool of specific IT experts and guarantee that their ability in industry stays applicable.

Ghosh et al (2013) Aim of this examination was to find the elements which help to decide the workers who expect to leave the association and the individuals who plan to remain with the association. The consequences of study showed that likelihood of representative with high emotional and regularizing obligation to the association would have lesser expectation to leave which is needed to be improved among workers. A Discriminant examination led can assist an association with being proactive and which can additionally help in setting up the maintenance technique.

2.3 Future of HR Analytics

Ulrich (2010) deduces in their investigation it is the propensity of HR experts to keep away from quantitative part of business however it will no more be workable for them to stay away from the information.

Johnson et al (2012) states that best improvement in HRIS understanding will be acquired through speculation of time, ability and assets accessible by HR experts.

2.4 Concluding Remark

In this chapter, the literature review was conducted to give a more extensive perspective on scholarly writing accessible on subject of HR analytics, how analytics can be coordinated with functions being carried out in the companies, and further extent of research that studies published have provided to scholarly community and industry experts. The next chapter deals with the research methodology used, highlighting the data coverage and the tools and techniques to be employed to draw results.

Chapter 3

RESEARCH METHODOLOGY

The research is a comprehensive explorative study with an attempt to employ mathematical models and perform statistical modeling to attain the defined objectives. This chapter deals with the research methodology and briefly discusses the tools and techniques adopted for the study.

3.1 Data Coverage

Secondary data has been gathered from the Kaggle.com, which is a web-based platform for consulting, data analytics competitions and predictive modelling, and utilized for the purpose of project. To discover factors which can be estimated to predict employees' performance and to build prediction models, employee-attitudes based survey-information from a manufacturing-based firm has been utilized for a time-frame of 2-years from 2014-2015.

The secondary data-set being used for the study incorporates: Employee-attitudes based survey-information gathered in November-2014, Performance evaluation information gathered in January-2014, Performance evaluation information gathered in January-2015 and Number of leaves due to illness, taken per worker in the year 2014.

3.2 Tools & Techniques

The collected data was classified and tabulated in SPSS for analysis and interpretation. SPSS Statistics is a software product used for statistical, batched and intelligent analysis. The following listed statistical techniques have been employed in the study:

3.2.1 Multiple Linear Regression

Multiple-Linear-Regression (MLR), likewise referred to just as multiple-regression, is a statistical-method that utilizes a few illustrative factors to anticipate the result of a dependent variable. The objective of Multiple-Linear-Regression (MLR) is to show the direct connection between the informative (autonomous) factors and reaction (dependent) factor.

Generally, multiple-regression is the expansion of conventional least-squares-regression (OLS) since it includes more than one informative variable.

Formula and calculation for Multiple-Linear-Regression:

$$yi=\beta 0+\beta 1xi1+\beta 2xi2+...+\beta pxip+\epsilon$$

Where, for $i=n$ observations:

 $yi=$ dependent variable

 $xi=$ explanatory variables

 $\beta 0=$ y-intercept (constant term)

 $\beta p=$ slope coefficients for each explanatory variable

 $\epsilon=$ the model's error term (also known as the residuals)

The multiple-regression-model is based on the following presumptions:

- There is a linear-relationship between the dependent-factors and the independent-factors
- The independent-factors are not very profoundly related with one another
- yi observations are selected randomly and autonomously from the population-dataset
- Residuals must be normally-distributed with variance σ and zero mean

The coefficient-of-determination (R squared) is a statistical-metric that is utilized to calculate the amount of the variance in result can be clarified by the amount of variance in independent-factors. R2 consistently increments as more indicators are added to the M-L-R model, despite the fact that the indicators may not be identified with the result variable.

R2 without anyone else can't hence be utilized to recognize which indicators ought to be used for a model and which ought to be avoided. R2 must be somewhere in the range of 0 and 1, where 0 shows that the result can't be anticipated by any of the independent-factors and 1 demonstrates that the result can be anticipated without blunder from the autonomous-factors.

When analyzing the outcomes of multiple-regression model, beta coefficients are legitimate while holding any remaining factors consistent ("all else equivalent"). The outcome from a multiple-regression model can be shown on a level plane as a condition, or in an upward direction in table structure.

3.2.2 Stepwise Multiple Linear Regression

Stepwise-Regression is a technique that iteratively looks at the factual meaning of every autonomous variable in a linear-regression model. The forward determination approach begins with nothing and adds each new factor steadily, testing for measurable importance. The backward-elimination strategy starts with a full model stacked with a few factors and afterward eliminates one variable to test its significance comparative with over-all results. Stepwise-Regression has its drawbacks, in any case, as it is a methodology that fits information into a model to accomplish the ideal outcome.

The basic objective of stepwise-regression is, through a progression of tests (e.g., F-tests, t-tests) to track down a bunch of autonomous factors that essentially impact the reliant variable. This is finished with PCs through iterative cycles, which is the way toward showing up at results or choices by going through continuous iteration or patterns of examination. Carrying out tests naturally with assistance from statistical-software-packages enjoys the benefit of saving time and restricting missteps.

Stepwise-Regression can be accomplished either by evaluating each autonomous factor in turn and remembering it for the regression-model in the event that it is genuinely huge or by remembering all expected autonomous factors for the model and disposing of those that are not measurably critical. Some utilize a blend of the two strategies and in this manner, there are three ways to deal with stepwise-regression:

- 1. Forward-selection starts without any factors in the model, tests every factor as it is added to the model, at that point keeps those that are considered most measurably huge—iterating the cycle until the outcomes are ideal.
- 2. Backward-elimination begins with a bunch of independent-factors, erasing one-at-a-time, at that point testing to check whether the eliminated variable is measurably critical.
- 3. Bidirectional-elimination is a blend of the initial two techniques that test which factors ought to be incorporated or barred.

33	Concluding Remark
	s chapter dealt with the tools and techniques discussing the research methodology for the study
	e next chapter discusses the analysis of the data based on various indices and statistical tests
	cussed in this chapter. Further, the next chapter highlights the findings of the analysis furnishing
	ropriate reasons behind it.

Chapter 4

EMPIRICAL ANALYSIS

An empirical analysis is based on the statistical techniques pre-decided as the research methodology. This chapter includes the analysis of the secondary dataset divided into two sections, namely Exploratory Data Analysis through Visualizations and Predictive Modelling and Analysis through SPSS. The findings of the empirical analysis have been discussed in detail with the appropriate reasoning behind it.

4.1 Exploratory Data Analysis

The research started through by exploring the secondary data set gathered from Kaggle.com, by making a graphical representation of the numeric data in PowerBI Visualization Software by Microsoft.

In this project, a secondary data-set, which is a collection of 626 representatives, containing their scores, absence due to illness records, performance-evaluation scores, from the yearly worker-attitude-survey from a manufacturing-based company, has been utilized. The information, taken over the course of 2-years, incorporates the accompanying information:

- Employee-attitude-survey information gathered in November-2014.
- Performance-appraisal rating gathered in January-2014.
- Performance-appraisal information gathered in January-2015.
- Number of leaves taken due to illness, per worker, in the year 2014.

Here, the information has been connected from two sources to build composite survey-estimations (normal measures got from different survey-questionnaires) utilizing the distinctive Employee-ID. Some new terms have been presented and a short depiction for each is given underneath:

 Perceived-Organizational-Support (POS): this is how much workers accept that their company gives value to their commitments, thinks often about their prosperity and satisfies emotional and social-needs.

- **Distributive-Justice-Perceptions:** this is encouraged where prizes and results are seen to be steady with verifiable standards for their allocation to employees, like value or fairness
- Procedural-Justice-Perceptions: this is cultivated through voice during a decision making
 interaction or impact over the result or by adherence to reasonable cycle models like
 consistency, absence of predisposition, correctability, portrayal, precision and ethicality
- Job-Satisfaction: this is characterized as 'a pleasurable or positive passionate state coming about because of the examination of one's work or professional adventures. It is said to have passionate, intellectual and social segments where the enthusiastic segment alludes to sentiments viewing the work like fatigue, nervousness or energy; the psychological segment of occupation fulfillment alludes to convictions seeing one's work, for example, feeling that one's work is intellectually requesting and testing; lastly the conduct segment remembers individuals' activities for connection to their work, which may incorporate being late, remaining late, or professing to be sick to keep away from work.
- Person-Organization-Fit: this alludes to the similarity between a worker and their association. It can prompt expanded degrees of trust and a common feeling of corporatecommunity.
- **Job-Strain:** this alludes to a mental reaction to a demanding-work environment and can be a result of having the view of having little power over one's work, while confronting high occupation-demands.

The data looks like this:

- 1. **ID** (Unique Employee-ID).
- 2. **Gender** (1=female,2=male).
- 3. Sick Days 2014 (total no. of leaves taken by employees due to sickness in 2014).
- 4. **Performance Rating 2015** (performance-appraisal-rating given in January-2015 on a scale of 1-5, with the highest rating being 5).
- 5. **Performance Rating 2014** (performance-appraisal-rating given in January-2014 on a scale of 1-5, with the highest rating being 5).
- 6. **Job Strain 2014** (composite of job-strain-questions on a scale of 1-5 as at November-2014).

- 7. **Value Fit 2014** (composite of person-organization-fit-questions on a scale of 1-5 as at November-2014).
- 8. **JOB SAT 2014** (composite of job-satisfaction-questions on a scale of 1-5 as at November-2014).
- 9. **Proc Just 2014** (composite of procedural-justice-questions on a scale of 1-5 as at November-2014).
- 10. **Dist Just 2014** (composite of distributive-justice-questions on a scale of 1-5 as at November-2014).
- 11. **POS 2014** (composite of perceived-organizational-support-questions on a scale of 1-5 as at November-2014).
- 12. **Job Tenure 2014** (years with the company as at November-2014).

4.1.1 Count of number of Employees by Performance Rating in 2014

Figure 4.1 illustrates Count of Number of Employees versus their Performance Rating in the year 2014.

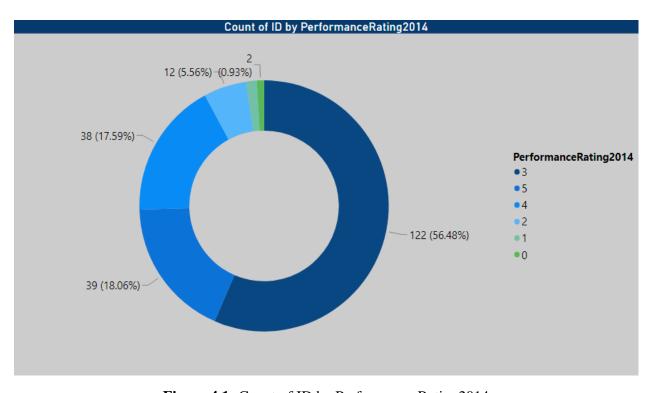


Figure 4.1: Count of ID by Performance Rating2014

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.1 illustrates that a large majority of employees i.e., 56.48% had an average performance rating i.e., 3 in the year 2014. Around 35.65% of employees had an above average performance rating in the year 2014. And, a negligible number of employees i.e., 6.49% had below average performance in the year 2014.

4.1.2 Count of number of Employees by Performance Rating in 2015

Figure 4.2 illustrates Count of Number of Employees versus their Performance Rating in the year 2015.

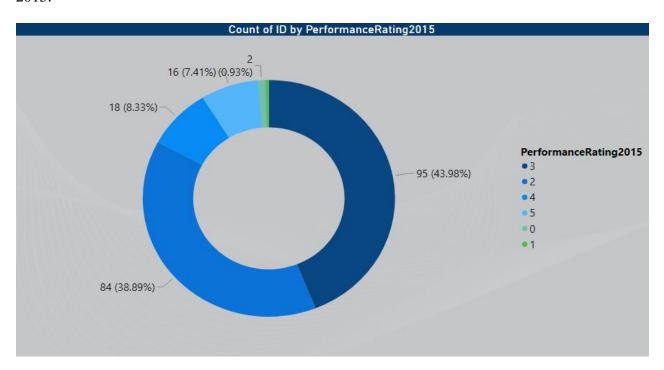


Figure 4.2: Count of ID by Performance Rating2015

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.2 illustrates that a large majority of employees i.e., 82.87% had an average performance rating i.e., 2 and 3 in the year 2015. Around 15.74% of employees had an above average performance rating in the year 2015.

This shows that the count of employees having above average performance declined in the year 2015 as compared to the year 2014. And, a negligible number of employees i.e., 0.93 % had below average performance in the year 2015.

4.1.3 Count of number of Employees by Number of Sick Days in 2014

Figure 4.3 illustrates Count of Number of Employees versus number of sick days taken by them in the year 2014.

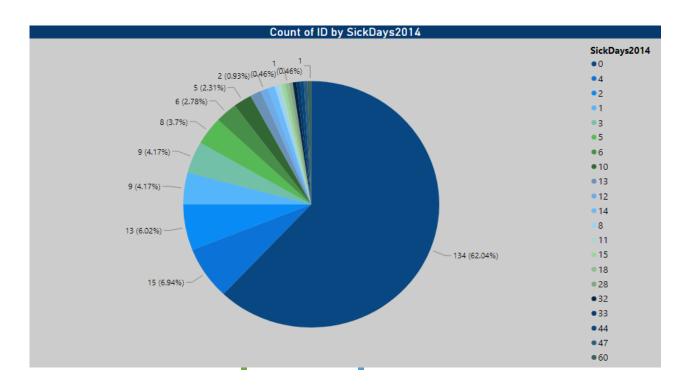


Figure 4.3: Count of ID by SickDays2014

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.3 illustrates that a large majority of employees i.e., 62.04% took 0 sick leaves in the year 2014. Around 3.09% of employees took sick leaves in the range of 1 to 10 in the year 2014. And, a negligible number of employees i.e., 1.85 % took sick leaves in the range of 12 to 60 in the year 2014.

4.1.4 Count of number of Employees by Gender

Figure 4.4 illustrates Count of Number of Employees by Gender [Male, Female] in the organization under study.

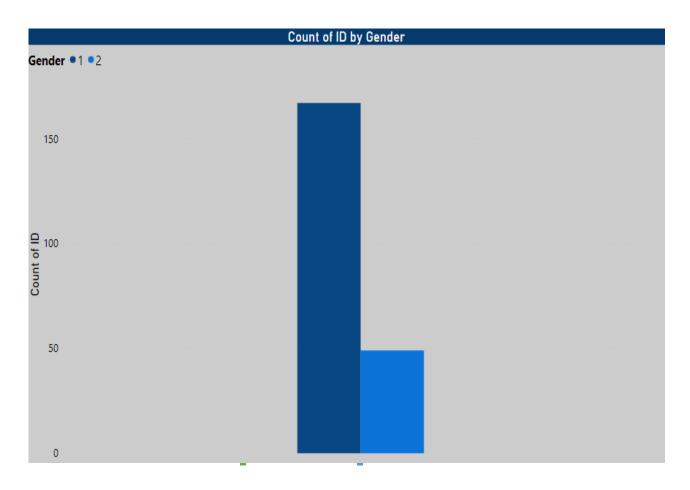


Figure 4.4: Count of ID by Gender

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.4 illustrates that the manufacturing organization under study had large number of male employees as compared to female employees.

4.1.5 Count of number of Employees by Job Tenure in years

Figure 4.5 illustrates Count of Number of Employees by their Job Tenure in years with the organization under study as at November 2014.

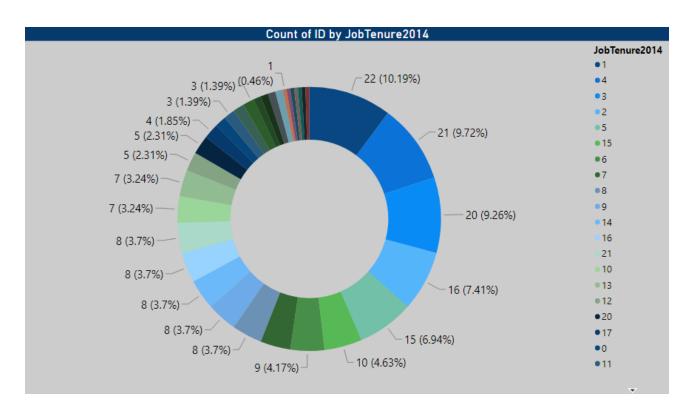


Figure 4.5: Count of ID by JobTenure2014

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.5 illustrates that only 3.7% of employees had Job Tenure of 21 years which is also the maximum job tenure. 10.19% of employees had Job tenure of only 1 year.

Around 43.52% of employees had Job Tenure between 1 to 5 years. Around 15.27% of employees had Job Tenure between 6 to 10 years. Around 41.21% of employees had Job Tenure in the range of 11 to 21 years.

4.1.6 Count of number of Employees by Perceived Organizational Support in 2014

Figure 4.6 illustrates Count of Number of Employees by their Perceived Organizational Support (POS) in the year 2014.

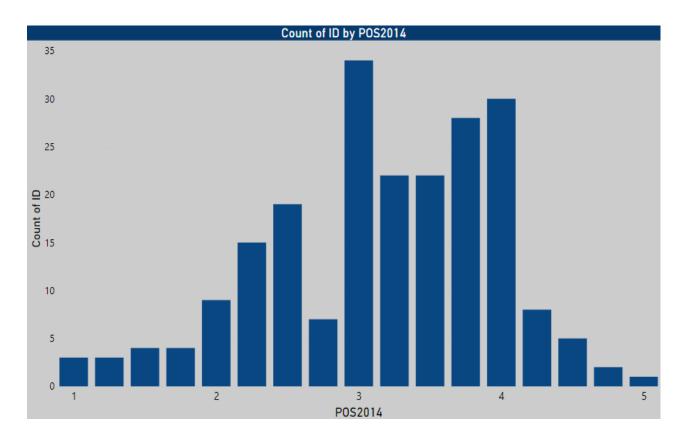


Figure 4.6: Count of ID by POS2014

Source: Own Analysis - Own Analysis - Built-in Microsoft PowerBI

Figure 4.6 illustrates that majority of employees had POS (Perceived Organizational Support) score in the range of 3 to 4. And very negligible number of employees had a low POS Score (1 to 2). Also, a smaller number of employees had a high POS Score of 5.

Here, POS Score is the degree to which employees believe that their organization values their contributions, cares about their well-being and fulfils socio-emotional needs.

4.1.7 Count of number of Employees by Job Strain in 2014

Figure 4.7 illustrates Count of Number of Employees by their Job Strain in the year 2014.

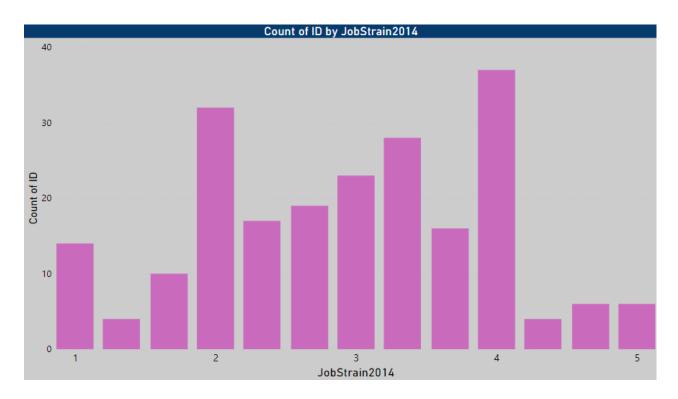


Figure 4.7: Count of ID by JobStrain2014

Source: Own Analysis - Own Analysis - Built-in Microsoft PowerBI

Figure 4.7 illustrates that majority of employees experienced Job Strain in the range of 2 to 4 at the organization under study. A smaller number of employees experienced a very high or a very low Job Strain.

Here, Job Strain refers to a psychological response to a demanding work context and can be a consequence of having the perception of having little control over one's work, while facing high job demands.

4.1.8 Count of number of Employees by Distributive Justice Perception in 2014

Figure 4.8 illustrates Count of Number of Employees by their Perception regarding Distributive Justice in the year 2014.

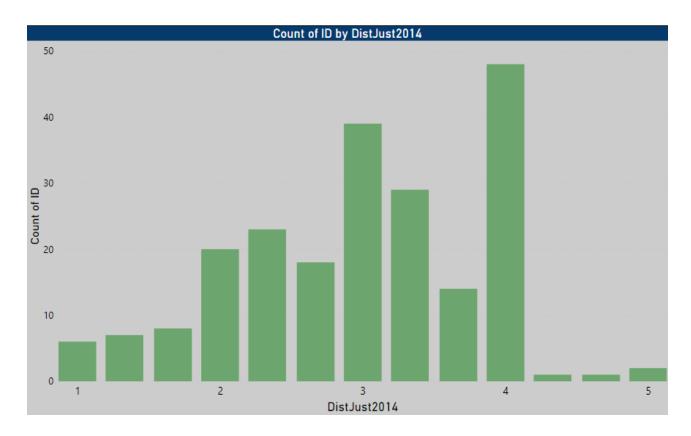


Figure 4.8: Count of ID by DistJust2014

Source: Own Analysis - Built-in Microsoft PowerBI

Figure 4.8 illustrates that majority of employees at the organization under study believed that the rewards were distributed according to the outcomes produced by the employees.

Distributive Justice Perception is fostered where rewards and outcomes are perceived to be consistent with implicit norms for allocation, such as equity or equality.

4.1.9 Count of number of Employees by Procedural Justice Perception in 2014

Figure 4.9 illustrates Count of Number of Employees by their Perception regarding Procedural Justice in the year 2014.

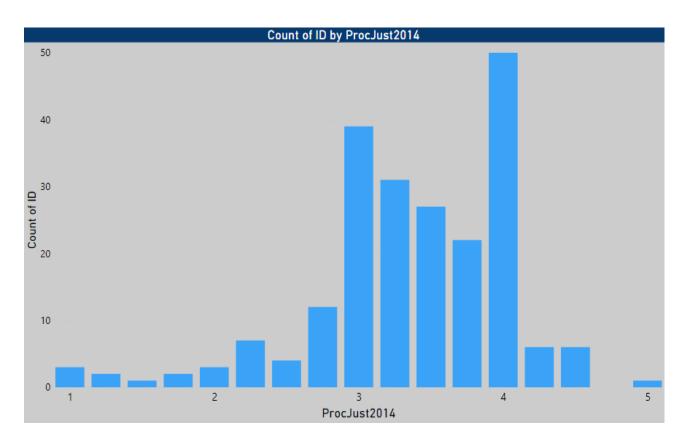


Figure 4.9: Count of ID by ProcJust2014

Source: Own Analysis - Own Analysis - Built-in Microsoft PowerBI

Figure 4.9 illustrates that majority of employees at the organization under study perceived high procedural justice. Also, a negligible number of employees perceived a very high or a very low procedural justice being met at the organization under study.

Procedural justice is fostered through voice during a decision-making process or influence over the outcome or by adherence to fair process criteria such as consistency, lack of bias, correctability, representation, accuracy and ethicality.

4.2 Predictive Modelling using SPSS Statistical Software

In choosing which model to test here, one can settle on numerous decisions (a circumstance that the HR professional will regularly end up in). The best option to make is which performance measure one should utilize.

In our dataset two measures are accessible January 2014 and January 2015 (giving the appraisals from the earlier year's performance).

Our linear regression-based prediction models, make a certain presumption that A causes B, or for this situation that our independent factors may cause variety in employee's performance. Thus, if the independent factors which are accessible are gathered in 2014, we likely need to ensure that the measure of employee's performance that we are to foresee with these variables is the 2015 measure of performance.

In the event that we attempt to anticipate the January 2014 measure of employee's performance with November 2014 staff perspectives then we are doing things in reverse. Along these lines, we ought to pick the 2015 measure of employee's performance as a dependent factor in the model.

The multiple regression algorithm-based prediction model was run in SPSS, taking a look at the connection between employee's performance, absence due to sickness and attitudes of employees. Hence, the model shown below has been tested in SPSS Software:

Performance-Rating-2015 (January) = a + b1 (POS 2014) + b2 (Dist Just 2014) + b3 (Proc Just 2014) + b4 (JOB SAT 2014) + b5 (Value Fit 2014) + b6 (Job Strain 2014) + b7 (sick days 2014)

Figure 4.10 illustrates the Data-View and Figure 4.11 illustrates the Variable-View after the dataset is loaded into SPSS Software as shown below:

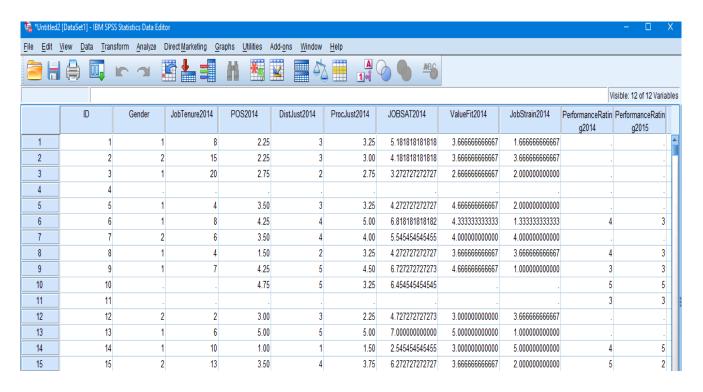


Figure 4.10: Data-View in SPSS Software

Source: Own Analysis - Built-in SPSS Statistical Software

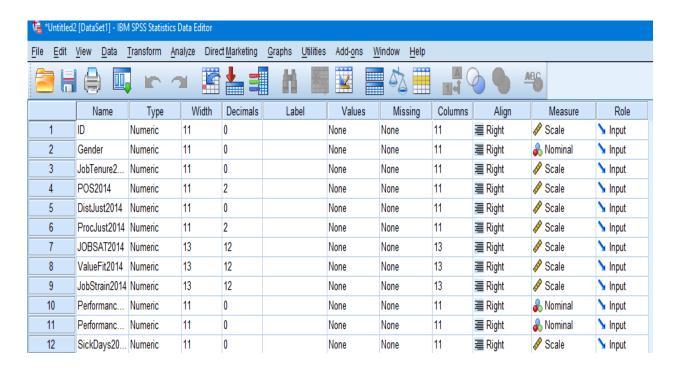


Figure 4.11: Variable-View in SPSS Software

Source: Own Analysis - Built-in SPSS Statistical Software

4.2.1 Multiple Linear Regression based Prediction Model

Utilizing the very same linear regression-based strategy, that has been used elsewhere, we can test this model by choosing 'Analyze-----→Regression-----→Linear' and moving the Performance Rating 2015 variable into the 'Dependent:' selection box and selecting the accompanying factors as 'Independent Variables' and moving them into the 'Independent(s):' selection box: POS-2014, DistJust-2014, ProcJust-2014, JOBSAT-2014, ValueFit-2014, JobStrain-2014 and sickdays-2014 as shown in the *Figure 4.12* and *Figure 4.13*.

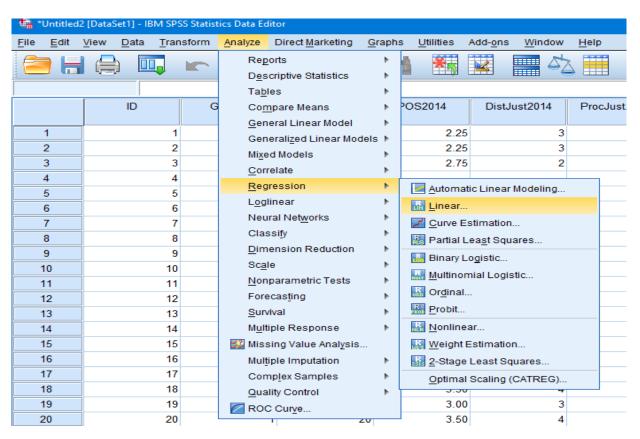


Figure 4.12: Analyze, Regression and Linear Options in SPSS Source: Own Analysis - Built-in SPSS Statistical Software

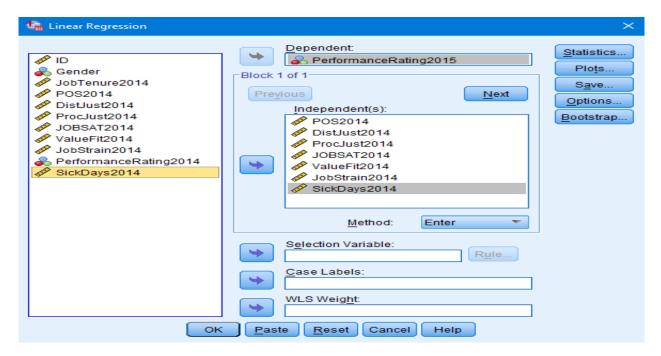


Figure 4.13: Moving Independent and Dependent Variables to respective selection boxes

Source: Own Analysis - Built-in SPSS Statistical Software

At the point, when we select 'OK' option, we get the output, as shown in the *Figure 4.14* and *Figure 4.15*:

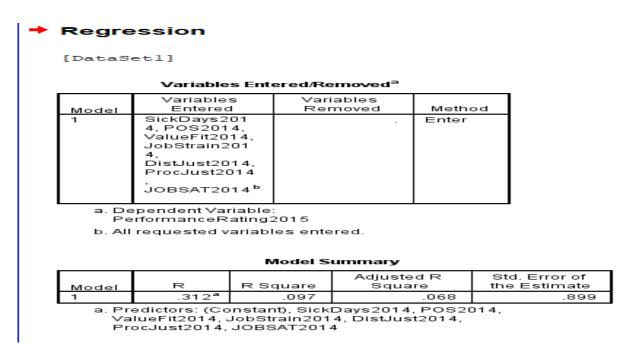


Figure 4.14: Output of Multiple Linear Regression Model in SPSS

Source: Own Analysis - Built-in SPSS Statistical Software

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.553	7	2.650	3.278	.002 ^b
1	Residual	172.207	213	.808		
	Total	190.760	220			

- a. Dependent Variable: PerformanceRating2015
- b. Predictors: (Constant), SickDays2014, POS2014, ValueFit2014, JobStrain2014, DistJust2014, ProcJust2014, JOBSAT2014

Coefficients^a

		Unstandardized Coefficients		Standardized Coefficients		
Model		В	B Std. Error		t	Sig.
1	(Constant)	.959	.584		1.641	.102
l	POS2014	.196	.130	.168	1.500	.135
l	DistJust2014	126	.102	116	-1.235	.218
l	ProcJust2014	.163	.140	.121	1.163	.246
l	JOBSAT2014	.033	.100	.037	.331	.741
l	ValueFit2014	.062	.114	.041	.544	.587
l	JobStrain2014	.247	.072	.267	3.415	.001
	SickDays2014	018	.008	144	-2.145	.033

a. Dependent Variable: PerformanceRating2015

Figure 4.15: Output of Multiple Linear Regression Model in SPSS

Source: Own Analysis - Built-in SPSS Statistical Software

In the above model, we can see that the R-square = 0.097, or that 9.7 percent (around one-10th) of the fluctuation in the employees' rating of performance given in January 2015 can be represented by the predictive factors being utilized here.

Taking a look at the investigation of change in *Figure 4.15*, i.e., the regression-based ANOVA table, which tests the level of how well the model predicts change in the 2015's rating of employees' performance, we can see that the regression-based ANOVA F value is 3.278, with a related all out degrees-of-freedom of 220 and the significance-level of 0.002 (p<0.01).

So, we could say it is under 0.005, which implies there is not exactly a 5 of every 1,000 possibility that we would arbitrarily discover this example of shared difference (between the 2015's rating of employees' performance and the independent factors in the model). Therefore, the model is significant.

Along these lines, while the model just records for one-10th of the result (2015's rating for employees' performance), it is significant and henceforth we can be certain about suggestions that can be made based on the critical outcomes created with this model (in that we may have discovered

variables that may affect the employees' performance). It is critical to set the assumptions so that the effect will be traditionalist (and consistently admonition our cases on the premise that our regression-based discoveries don't conclusively 'demonstrate' causality).

This is a regression-based technique where you put all components in as performance indicators and see what come out as critical indicators. This aides control for various factors and considers noise in data and connections between the components is estimated.

Taking a look at the Beta coefficients, one can see that the quantity of days off taken in 2014 has a negative relationship with the 2015's rating of employees' performance, and the measure of occupation strain has a critical and positive relationship with the 2015's rating of employees' performance (this finding could be an illustration of where one should investigate communications between the predictive variables under study).

The perceived degree of occupation strain and the quantity of days off taken in 2014 are the key likely drivers of performance-on-performance appraisal ratings given in 2015. Curiously, the more an individual feels they have work strain, the almost certain they are to get a better performance appraisal rating.

4.2.2 Stepwise Regression based Prediction Model

One could also utilize a somewhat extraordinary technique called stepwise-regression. This strategy, as a result, simply chooses the primary factors adding to the predictivity of the dependent variable under study. It takes a look at the independent factors, it tracks down the most grounded relationship, and afterward it enters this as a regression-based model.

At that point it hopes to check whether the variable with the following most grounded relationship adds to the regression-based model in the event that it is incorporated in the model. This iterative-process proceeds until the predictivity of the dependent factor isn't improved any more. Hence, one gets the most effective and efficient regression-based model. This cycle is called stepwise-regression.

In the event, if one enters the above independent factors into a stepwise-regression model anticipating employees' performance then we get a cleaner, more straightforward, more tightfisted model. The idea of stinginess here alludes to not having pointless components incorporated in the model with the contention 'basic is better'. One should remember, in any case, that when we do this, the model isn't 'controlling' for the conceivable impact of other (barred) factors, so the expert requirements to ponder what method can be used.

In the stepwise-regression based prediction model, the same dataset has been used as in the case of simple multiple linear regression-based prediction model. To run stepwise-regression in SPSS first of all, 'Regression' option is selected, then 'Linear' from the 'Analyze' menu as shown in the *Figure 4.16*:

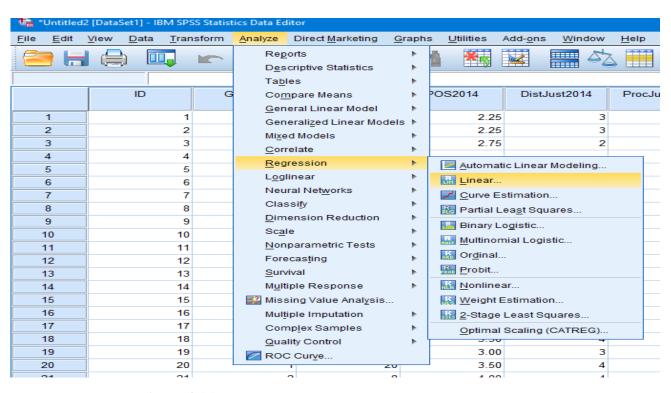


Figure 4.16: Analyze, Regression and Linear Options in SPSS

Source: Own Analysis - Built-in SPSS Statistical Software

On the next screen, the Independent and Dependent variables are moved to the selection boxes as done before in the previous model. The primary difference in stepwise-regression is that we select 'Stepwise' in the 'Method' option as shown in the *Figure 4.17*:

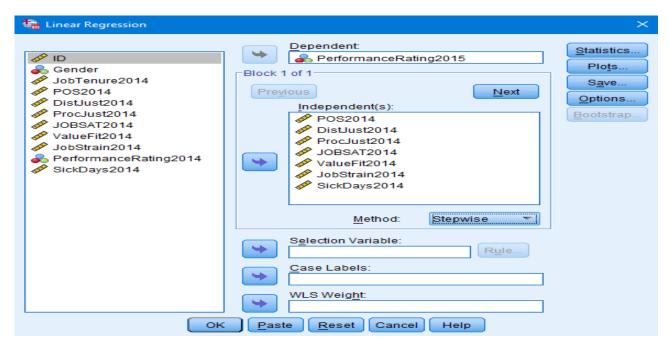


Figure 4.17: Moving Independent and Dependent Variables to respective selection boxes

Source: Own Analysis - Built-in SPSS Statistical Software

Then, 'OK' option is selected. This resulted in the output as shown in the *Figure 4.18*, *Figure 4.19*, *Figure 4.20* and *Figure 4.21*:

Regression

[DataSetl]

Variables Entered/Removed^a

1	Variables	Variables	
Model	Entered	Removed	Method
1	JobStrain201 4	-	Stepwise (Criteria: Probability-of- F-to-enter <= . 050, Probability-of- F-to-remove >= .100).
2	ProcJust2014	-	Stepwise (Criteria: Probability-of- F-to-enter <= . 050, Probability-of- F-to-remove >= .100).
3	SickDays201 4	-	Stepwise (Criteria: Probability-of- F-to-enter <= . 050, Probability-of- F-to-remove >= .100).

a. Dependent Variable: PerformanceRating2015

Figure 4.18: Output of Stepwise Regression Model in SPSS

Source: Own Analysis - Built-in SPSS Statistical Software

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.153ª	.024	.019	.922
2	.250 ^b	.063	.054	.906
3	.282°	.080	.067	.900

a. Predictors: (Constant), JobStrain2014

b. Predictors: (Constant), JobStrain2014, ProcJust2014

c. Predictors: (Constant), JobStrain2014, ProcJust2014, SickDays2014

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.486	1	4.486	5.274	.023 ^b
	Residual	186.274	219	.851		
	Total	190.760	220			
2	Regression	11.965	2	5.982	7.294	.001°
	Residual	178.796	218	.820		
	Total	190.760	220			
3	Regression	15.166	3	5.055	6.247	.000 ^d
	Residual	175.595	217	.809		
	Total	190.760	220			

a. Dependent Variable: PerformanceRating2015

b. Predictors: (Constant), JobStrain2014

c. Predictors: (Constant), JobStrain2014, ProcJust2014

d. Predictors: (Constant), JobStrain2014, ProcJust2014, SickDays2014

Figure 4.19: Output of Stepwise Regression Model in SPSS

Source: Built-in SPSS Statistical Software

Coefficients^a

		Unstandardized Coefficients		Standardized Coefficients		
Model		В	Std. Error	Beta	t	Sig.
1	(Constant)	2.404	.191		12.568	.000
	JobStrain2014	.142	.062	.153	2.297	.023
2	(Constant)	1.311	.408		3.216	.001
	JobStrain2014	.194	.063	.210	3.083	.002
	ProcJust2014	.278	.092	.206	3.020	.003
3	(Constant)	1.351	.405		3.333	.001
	JobStrain2014	.208	.063	.225	3.299	.001
	ProcJust2014	.269	.092	.199	2.936	.004
	SickDays2014	016	.008	131	-1.989	.048

a. Dependent Variable: PerformanceRating2015

Figure 4.20: Output of Stepwise Regression Model in SPSS

Source: Built-in SPSS Statistical Software

Excluded Variables^a

					Partial	Collinearity Statistics
Model		Beta In	t	Sig.	Correlation	Tolerance
1	POS2014	.215 ^b	2.886	.004	.192	.780
	DistJust2014	.071 ^b	.985	.326	.067	.848
	ProcJust2014	.206 ^b	3.020	.003	.200	.923
	JOBSAT2014	.185 ^b	2.491	.013	.166	.791
	ValueFit2014	.117 ^b	1.687	.093	.114	.916
	SickDays2014	140 ^b	-2.103	.037	141	.984
2	POS2014	.119°	1.212	.227	.082	.446
	DistJust2014	069°	805	.421	055	.588
	JOBSAT2014	.070°	.723	.471	.049	.459
	ValueFit2014	.042°	.557	.578	.038	.772
	SickDays2014	131°	-1.989	.048	134	.982
3	POS2014	.143 ^d	1.462	.145	.099	.440
	DistJust2014	063 ^d	742	.459	050	.587
	JOBSAT2014	.061 ^d	.632	.528	.043	.458
	ValueFit2014	.059 ^d	.786	.433	.053	.763

a. Dependent Variable: PerformanceRating2015

b. Predictors in the Model: (Constant), JobStrain2014

c. Predictors in the Model: (Constant), JobStrain2014, ProcJust2014

d. Predictors in the Model: (Constant), JobStrain2014, ProcJust2014, SickDays2014

Figure 4.21: Output of Stepwise Regression Model in SPSS

Source: Own Analysis - Built-in SPSS Statistical Software

The distinction here, in stepwise-regression, is that one presently, can contrast a couple of models and can develop a view in order to track down the model that best fits in the most effective manner. It is imperative to consider that while stepwise-regression enjoys a few benefits in that the models introduced toward the end just incorporate critical indicators, subsequently they are more basic and miserly models, there are some likely disadvantages to this methodology.

The principal drawback is that one is not 'controlling' for the impacts of other conceivable influencers of the dependent-variable. It may really be critical to do this and show that one has control of the potential impacts of different elements.

In the event that, for instance, somebody cross examining the discoveries can feel that some different variables (that the stepwise-regression-method neglected to incorporate as a model) could be significant, it might give the researcher or other people, more noteworthy assurance, on the off chance, if in case other variables are also included in the model, despite the fact that they don't come out as critical indicators.

Model-Summary can be looked at comprising of coefficients-table and ANOVA-table in *Figure 4.18*, *Figure 4.19*, *Figure 4.20* and *Figure 4.21*. Data-analysis has identified 3-best-models as described below:

Model-1

Model-1 shows that the vital driver in deciding the performance-appraisal-rating given in January-2015 is really the measure of occupation-strain informed, two months before in November-2014 (Beta=0.153, p<0.05).

The basic regression-model is in the form: Performance-Rating-2015 (January) = 2.404 + 0.142(Job-Strain-2014)

Model-2

Model-2 depicts the vital drivers of the performance-appraisal-rating given in January-2015 as a mix of occupation-strain (Beta=0.210, p<0.05) revealed in November-2014 and the procedural-justice (Beta=0.206, p<0.05) felt by the employee in November-2014.

The basic regression-model is in the form:

Performance-Rating-2015 (January) = 1.311 + 0.194(Job-Strain-2014) + 0.278(Proc-Just-2014)

• Model-3

Model-3 is more extensive and outlines the indicators of the performance-appraisal-rating given in January-2015 as a mix of occupation-strain (Beta=0.225, p<0.05) felt in November-2014, procedural-justice (Beta=0.199, p<0.05) felt in November-2014, and the all-out number of days fallen ill in 2014 (Beta=-0.131, p<0.05).

The basic regression-model is in the form:

Performance-Rating-2015 (January) = 1.351 + 0.208(Job-Strain-2014) + 0.269(Proc-Just-2014) - 0.016(sick-days-2014)

The rest of the factors (as demonstrated in the 'Excluded-Variables' table in *Figure 4.21*) don't assist us with foreseeing employees' performance and hence are excluded from these models.

What does the information show us? All things considered, curiously, apparently individuals who felt more under strain (who lost rest stressing over work) in November-2014 appear to have better performance-evaluation appraisal-ratings in January-2015. This is possibly dangerous from a wellbeing and prosperity viewpoint.

More strain and stress will in general prompt better? One would not have any desire to consider attempting to expand the strain and stress on people as a organizational-performance improvement technique. However, it very well may be an impression of the way that individuals who try sincerely and are feeling the squeeze, and in reality, who may lose rest stressing over their work, are in this manner compensated with a better appraising (whether or not they are really performing at a more elevated level).

The story doesn't stop there, be that as it may. The subsequent model demonstrates that individuals who feel that their manager acts in an equitable and reasonable manner additionally have better performance-appraisals. For what reason may this be? Indeed, one could decipher this as the

representatives who feel their manager isn't reasonable or just may have lower levels of inspiration or they may have a more displeased mentality than the individuals who experience (or see) more noteworthy degrees of procedural-justice; along these lines, this may work out in how their administrator at that point rates them.

It merits considering briefly that it is so critical to have the performance-appraisal-measure subsequent to the attitude-survey-rating in this model. In the event that, for instance, the model embedded the 2014-performance-measure here as a dependent-variable rather than the 2015-performance-measure, and a positive connection has been tracked down between November-2014's employees' perception of justice and January-2014 performance-appraisals, it is possible that a lower performance-rating delivered negative impression of procedural-justice, hence the causality would be switched. Thusly, it is fundamental that the performance-rating used in the model (as a dependent-variable) is given after the employee-attitude-measure. This issue has been addressed in the next model below.

In the event that we take a look at the third model that the model produces, we likewise see that number of days' leave due to illness, in 2014 is a huge indicator of 2015-performance-rating. This implies that the recurrence of days for which individuals were missing because of disease is by all accounts reflected in their performance-evaluation. Thus, in outline, the representatives who get the best appraisals are the individuals who have the most un-number of days off due to illness, who encounter and see that their manager acts in a procedurally way, and who feel a feeling of strain and stress over their work.

4.3 Concluding Remark

This chapter detailed the empirical analysis based on two Prediction Models modelled in SPSS i.e., Linear Regression and Stepwise Regression Model for predicting employees' performance. The next chapter discusses the recommendations and conclusions of the study.

Chapter 5

RECOMMENDATIONS & CONCLUSION

This chapter details the recommendations and conclusion of the study 'Predicting Employee's Performance using HR Analytics' based on the empirical analysis and findings.

Strategically, how is it possible that we would utilize this data to make suggestions to the company's employee strategy? To begin with, it is worth featuring to senior pioneers in the company that the information proposes we might be remunerating individuals for being under strain. This could be harming, on the grounds that it is the superior workers who are viewed as good examples in the association – and assuming they are losing rest and are feeling the stress, it very well might be a sign to others that in the event that they need to be remunerated additionally, they should stress themselves, which isn't good for them or for the company.

An undeniable degree of occupation strain is probably going to make the individual become sick, which would, in future, cause more days taken in nonattendance due to sickness and in the end (as per this model) lead to a lower rating due to poor performance. Thus, this methodology would not prompt consistent, long-term superior performance.

How about we take a look at the other key drivers: procedural justice and sickness absence. To start with, nonattendance sue to sickness. Numerous companies put resources into representative prosperity projects to assist workers with remaining and genuinely solid, and there are numerous experts in the field who can exhort on representative prosperity programs.

Indeed, even little, straightforward, minimal effort ideas like giving healthy fruits, empowering exercise through noon or after-work clubs, or getting gymnasium membership on discount for staff can go far to assist representatives with remaining sound and healthy. Different ideas incorporate teaching solid habits to workers through outside speakers, programmes to help representatives who are encountering pressure, offering nearby influenza inoculations or even an on-location medical specialist's office.

At last, procedural justice. The more workers see their manager to be reasonable and just, the higher they are probably going to perform. To improve this, it would in all likelihood require further examinations concerning which strategies are considered reasonable and just (and which

are not) so prescriptive move might be made. This should be possible through an audit of cycles joined with contribution from review respondents.

The apparent degree of occupation strain, the apparent degree of procedural equity felt and the quantity of days off taken in 2014 are the vital drivers of performance appraisal given in 2015. Strangely, the more an individual feels they have work strain, and the more they feel their manager is reasonable and just, the almost certain they are to get a better appraisal. The lower the quantity of days off taken, the higher the performance appraisal that they will in general be given.

5.1 Concluding Remark

This chapter discussed the recommendations and the conclusion.

REFERENCES

- Armstrong, M and Baron, A (2005) Managing Performance: Performance management in action, Chartered Institute of Personnel and Development, London
- Bach, S (2012) Performance management, in Managing Human Resources: HRM in transition, 5th edn, ed S Bach and M R Edwards, John Wiley and Sons, New York
- Bernstein, D A and Nash, P W (2008) [accessed November 2015] Essentials of Psychology, 4th edn, Cengage Learning, Boston [Online] http://books.google.com/books?id=4DobFrt9tUC.
- Boon, C and Den Hartog, D N (2011) Human resource management, person—environment fit,
 and trust, Trust and Human Resource Management, pp 109–21
- Brown, D and Hirsh, W (2011) Performance management: fine intentions, People Management, September, pp 34–37
- Colquitt, J A (2001) On the dimensionality of organizational justice: a construct validation of a measure, Journal of Applied Psychology, 86 (3), pp 386–400
- Cooper, C L, Dewe, P J and O'Driscoll, M P (2001) Organizational Stress: A review and critique of theory, research and applications, Sage, California
- Locke, E A (1976) The nature and causes of job satisfaction, in Handbook of Industrial and Organizational Psychology, pp 1297–349, ed M D Dunnette, Rand McNally, Chicago
- House, R J and Rizzo, J R (1972) Role conflict and ambiguity as critical variables in a model of organizational behavior, Organizational Behavior and Human Performance, 7, pp 467–505
- Eisenberger, R, Huntington, R, Hutchison, S and Sowa, D (1986) Perceived organizational support, Journal of Applied Psychology, 71, pp 500–07
- Daily Mail [accessed 25 November 2015] One Checkout Item Every Three Seconds, Daily Mail [Online] http://www.dailymail.co.uk/femail/article-2170377/Morrisons-imposes-new-second-targets-cut-till-queues.html