HR Metrics & Analytics
MSB21T2001

Index- Session 36

Session 36 - Objective

Predicting Employee Performance using SPSS

Recap

- HR measurements can be done with metrics, analytics and predictive analytics by using big data. Understanding the process of using these requires specialized skills and knowledge on various statistical tools for developing suitable algorithm.
- To develop predictive models for HR decisions, right at the beginning, we critically study the data volume.
- Predictive analytics can be used as a critical HR decision-making tool duly diagnosing the underlying problems. To develop a predictive model, it is important to select the appropriate predictor variables. This requires HR managers to understand which predictor variables can have potential effect on specific HR issues.
- The success of HR analytics and predictive modelling of HR decisions largely depends on the use of statistical tools combining HR and business data. Relating HR with business data can churn out information on customer satisfaction, market share trend and nature of products sold, e.g., premium products, mass products and so on.
- Predictive analytics extract information from raw data and make use of the same to predict the future pattern of behaviour or trends for getting better insights into HR decision-making. Broadly, predictive analytics tools are divided into regression techniques and machine learning.

Concept discussion -Predicting Employee Performance using SPSS

- Organizations are more performance driven to sustain in competition. Individual, team and organizational performances are now constantly measured and used as critical inputs for major decisions pertaining not only to employees but also business and strategy framing.
- Using HR analytics, organizations predict the future performance; identify who are the high-performing employees, who can have critical role in improving team performance; driving organizational performance and so on.
- Also, such predictive measurement of employees' performances helps in making investment decisions for learning and development on those employees who are chronic underperformers. Using HR analytics, we can predict performance and take decisions that are more accurate and business driven.
- For an HR manager, the first challenge here is to decide what sort of performance he/she will be predicting.
- Although we naturally assume that the annual performance rating data are adequate for us to predict the future performance, but this can only help us in getting a partial view.
- We have many other performance indicators that are capable of providing useful information.

Concept discussion

- For example, for predicting future team performance, we need to measure team function, competence of a team leader, attitude of individual team members, team engagement, customer feedback, customer loyalty, repeat business from customers, employee turnover rate in a team and so on.
- Similarly, for predicting the future individual performance, in addition to individual performance rating, we need to study behaviour rating, sales performance figures, individual customer feedback, peer feedback, gender, age group, sickness absence, job satisfaction, person–organization fit, compensation and benefits, perceptions on organizational equity and justice and so on.
- But operationally, many organizations may not have all these data sets. Obviously, in such cases, we have to restrict our predictive performance analysis based on the available data set in organizations.
- As performance measures widely vary across organizations, we cannot have any universal approach to collect performance data.
- In many organizations, customer satisfaction data may not be considered as the performance measure, so also the sickness absences, attrition data and employee engagement score.
- But all these have direct bearing on organizational performance.
- Hence, once we are capable to churn these data, we can encapsulate these in our predictive analysis.

Process of Analysis

- For performance analysis, we need to look into the relationships and predictors. Some of the points for our line of enquiry are as follows:
  - Characteristics of top individual and team performers
  - Investment on the training and development of these top performers
  - Key employee characteristics, capabilities and attributes of top performers
- To answer all the aforementioned points, we need to do multiple linear regressions.
- It can help us to analyze the number of independent variables, such as gender, age, team leader, country location and so on, and predict a dependent variable (e.g., performance level) by developing a 'best-fit' model.
- The multiple linear regressions will tell us the proportion of variance in the dependent variable that is accounted for by all the independent variables collectively and also can indicate which independent variable has a significant impact on the dependent variable, and the relative emphasis of each dependent variable.
- As the process of doing multiple linear regressions has been understood from a NPTEL course Predictive or any statistics book, we are not discussing the process here, rather focusing on data collection and analysis for predicting the future performance level of employees.

Based on our imaginary data sets (see the Annexure: Employee Data Set), we can perform the following multiple linear regressions:

- Predicting the individual performance level combining performance ratings data, employees' record of sickness
- Predicting the individual performance level with performance ratings data, sickness records and customer's loyalty
- Predicting the individual performance level with performance ratings data, sickness records and attitude of employee
- Predicting the individual performance level with performance ratings data, sickness records and employee's profile

This is a tentative description of multiple linear regressions that can be done in a hypothetical situation. Depending on the availability of the data sets and understanding how several factors can influence employees' performance, we can have even more multiple linear regressions to build our predictive model. As our purpose is to explain how predictive modelling for decision-making can be done by using SPSS, we will explain with a specific case of predicting employees' performance level and retention (dependent variables) in relation to some independent variables.

We have to analyze based on the imaginary data sets. Ideally, however, in a real-life situation, for this type of analysis, we collect some inputs from employees' through survey and some inputs from the existing data sets available in the organization. Our purpose here is to understand what factors can be attributed as important for employees' performance and retention. For this example, we have created imaginary data sets of 100 employees working in a hypothetical company. For each employee, we have the following seven survey questions: four of these questions are pertaining to employees' performance level issues and the rest three to understand how loyal they are to their organization.

The first phase of the four questions which are intended to measure the performance level in terms of employees' competence and skills are as follows:

- Understanding job (Sat 1)
- Functional autonomy (Sat 2)
- Frequency of interaction with boss (Sat 3)
- Use of technology in job (Sat 4)

Using a scale of 1–5, employees were asked to indicate their responses against the aforementioned questions, where 5 indicates highly satisfied and 1 indicates unsatisfied.

Similarly, the next three questions were asked to measure the likelihood of their retention, again using a scale of 1–5, except in case of gender, as follows: Data sets, represent imaginary responses from 100 employees

- Engagement with the organization (Eng)
- Competitive compensation and benefits (Ccb)
- Gender (1 for female and 2 for male)

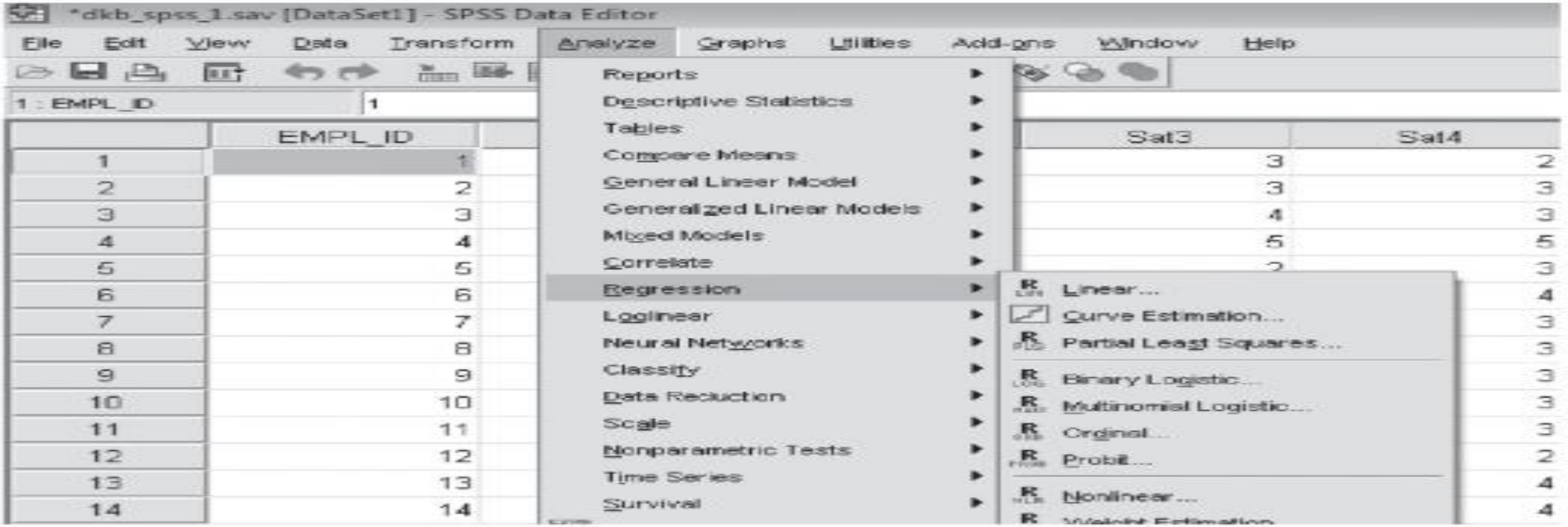
Looking at the problem, we can see that we have two dependent or outcome variables and seven independent or predictor variables. Our predictive analysis will help us in understanding whether the organization, by bringing change in predictor variables, can improve employees' performance levels and retention. Our regression model will look as follows:

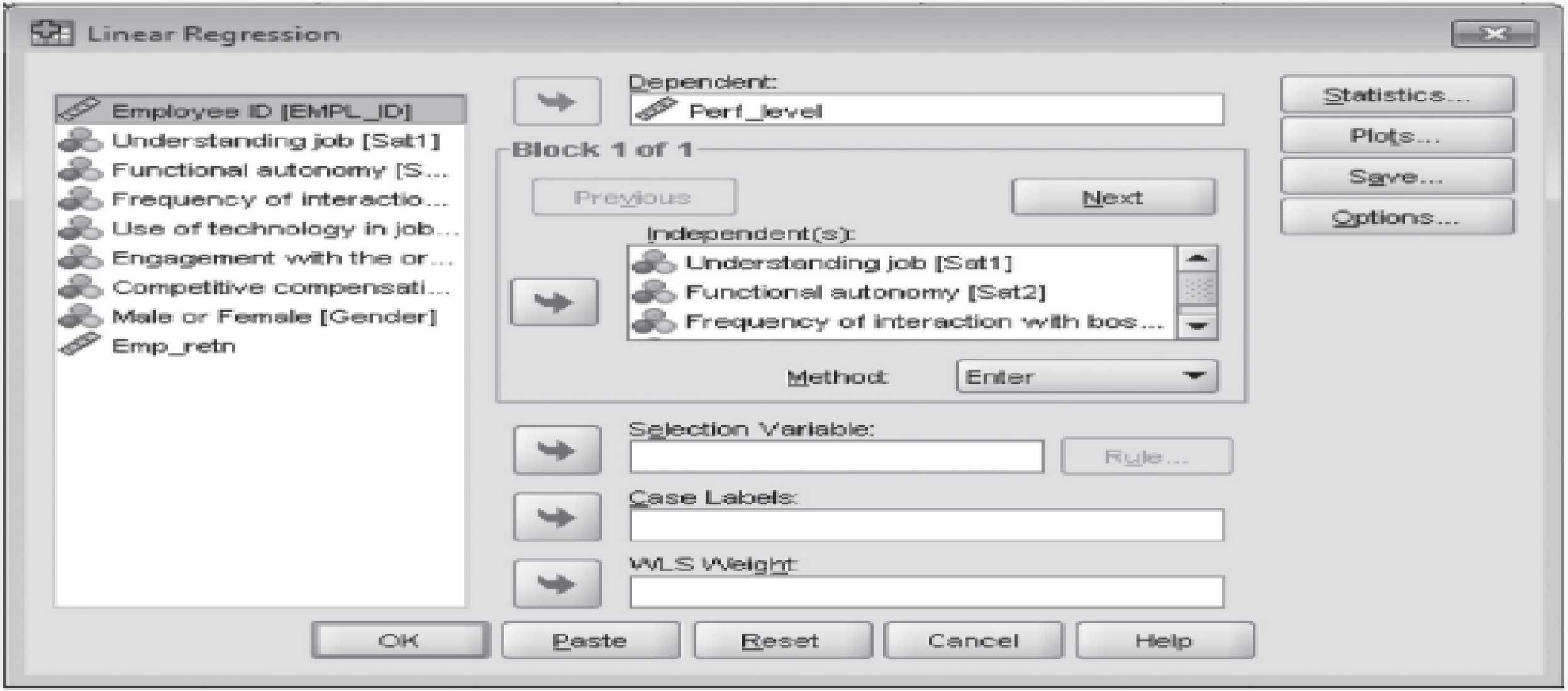
The performance level = a + b1 (understanding job) + b2 (functional autonomy) + b3 (frequency of interaction with boss) + b4 (use of technology in job).

Similarly, for employee retention, our model will be as follows:

Employee retention = a + b1 (engagement with the organization) + b2 (competitive compensation and benefits) + b3 (gender, i.e., 1 or 2).

Now we need to fit these data sets in SPSS, selecting ‘Analyze’, ‘Regression’ and ‘Linear’. Then, we set the ‘performance level’ as a dependent variable and four predictor or independent variables, mentioned earlier. The same is repeated for predicting employee retention. This is explained using the following screenshots of SPSS




Linear Regression
Employee ID [EMPL_ID]
Understanding job [Sat1]
Functional autonomy [S...
Frequency of interactio...
Use of technology in job...
Engagement with the or...
Competitive compensati...
Male or Female [Gender]
Emp_retn
Dependent:
Perf_level
Block 1 of 1
Previous
Next
Independent(s):
Understanding job [Sat1]
Functional autonomy [Sat2]
Frequency of interaction with bos...
Method:
Enter
Selection Variable:
Rule...
Case Labels:
WLS Weight:
Statistics...
Plots...
Save...
Options...
OK
Paste
Reset
Cancel
Help

After arranging data in SPSS (importing from the appended excel sheet), we then click 'OK' to get the output. In our case, output is

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Use of technology in job, Understanding job, Frequency of interaction with boss, Functional autonomy[a]		Enter

a. All requested variables entered.
b. Dependent Variable: Perf_level

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.235[a]	.055	.016	1.41788

a. Predictors: (Constant), Use of technology in job, Understanding job, Frequency of interaction with boss, Functional autonomy

The summary of the model -R-square is 0.055. R-square is the square of the multiple regression coefficients between the dependent and predictor variables together. So, the greater the R-square, the more the predictor variables are jointly predictive about the dependent variable. Multiply it by 100, it tells the % of variance in dependent variable,(accounted for by other variables). More technically, it is percentage of variance accounted for in dependent variable when taking into account its shared linear relationship with independent variables (interrelationships between the independent variables). So, in this case, the R-square is 0.055, and thus we can say that 5.5 percent of variance in expressions of the performance level is accounted for by the particular combination of predictor variables.

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	11.205	4	2.801	1.393	.242[a]
	Residual	190.985	95	2.010		
	Total	202.190	99			

a. Predictors: (Constant), Use of technology in job, Understanding job, Frequency of interaction with boss, Functional autonomy

b. Dependent Variable: Perf_level

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.060	1.057		4.787	.000
	Understanding job	-.408	.286	-.191	-1.428	.157
	Functional autonomy	.074	.295	.037	.251	.802
	Frequency of interaction with boss	.168	.230	.090	.729	.468
	Use of technology in job	-.307	.211	-.171	-1.452	.150

Looking at the ANOVA table, we can understand how well this model predicts variation in the performance level. Here, the regression ANOVA F-value is 1.393 with total degrees of freedom of 99. Although such statistical significance can be interpreted from a statistical table commonly available in any book on statistics, here we could get it from SPSS directly. Here, SPSS has calculated the statistical significance for us to be 0.242 (at 5 percent level of significance). So, we can say that the significance level reached has a p-value of 0.242, which means there is less than about 24 percent chance that we would find this pattern of shared variance (between the performance level and the other survey questions) by chance alone. Therefore, our model is not significant, as the data set used is imaginary.

Had the model being significant from the ANOVA outcome, we could understand

- Which of our independent variables have a significant impact on the performance level,
- To what extent and also in
- What direction (i.e., some may potentially improve the performance while others may decrease it).

Using our knowledge of the levels of significance and looking at the right-hand column of the table (th coefficients table), it appears that four of the predictor variables could not significantly predict the performanc level.

The data sets used for this analysis are imaginary and used only to explain analytical process.

But from this example and with the actual data, it would be reasonable to make the recommendation to invest i helping employees to better understand their jobs, similarly giving more functional autonomy and enhancing th frequency of interaction with boss.Likewise, we can also predict the employee retention in relation t engagement with the organization, competitive compensation and benefits, and gender.

REFERENCES

- Bhattacharya ,Deepak (2019). HR Metrics and Analytics, Pearson
- Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. Academy of Management Journal, 45(3), 587–597.
- Becker, B. E., & Gerhart, B. (1996). Human resources and organizational performance: Progress and prospects. Academy of Management Journal [Special Issue: Human Resources and Organizational Performance], 39(4): 779–801.
- Becker, Gary, S. (1962). Investment in human capital: A theoretical analysis. Journal of Political Economy, 70(2), 437–448.
- ———. (1964). Human capital. New York: National Bureau for Economic Research.
- Beer, M., Spector, R., Lawrence, P., Quinn Mills, D., & Walton, R. (1984). Human resource management: A general managers perspective. Glencoe, IL: Free Press.
- Begin, J.P. (1993). Identifying patterns in HRM systems: Lessons learned from organizational theory. In J. Shaw, P. Kirkbride and K. Rowlands (Eds.), Research in Personnel and Human Resource Management, 3 (pp. 3–20), Greenwich, CT: JAI Press.
- Bhattacharyya, D. K. (2012). Human resource management (3rd ed.). New Delhi: Excel Books

In case, you find any difficulty in understanding the concepts of lecture, please feel free to contact.

Dr Mamta Gaur

9891121396

mamtagaur@galgotiasuniversity.edu.in

Thanks