Statistical Approaches in Railway Transport

Daniel Dăneci-Pătrău
"Spiru Haret" University, Romania
danusidenima2@yahoo.com

Abstract

The analysis of the human resources performance in the railway transport, which is the main goal of this project, aims to identify the performance indicators of the wage earner on one hand, and on the other hand, the presentation of the relations and correlations between the wage level, seniority, labor conditions and requirements of the workplaces for the employee at the Constanta CFR railway station. The global economic reality over the last years allows the stressing of some indubitable defining features: the diversification and the renewal of the goods offer, the progress in the technology area, the globalization process, the increase in exigencies of clients and society. In this context, the goal of the development project is to offer managers quantitative and qualitative patterns, methods and work instruments for the estimate of human resources management systems performance in the railway transport.

Key words: human resources, railway transport, frequency curve
J.E.L. classification: M12, L92

1. Introduction

It can be stated that human resources management is the "backbone" of a modern company, directly related to productivity, performance, and success. That is why it must be viewed as a direct consequence of a double tendency, both related to obtaining superior performance and, in modern organizations, to the resizing, of the perspective on the individual as a distinct entity, (in terms of a social and economic philosophy focused on the human being) stressed by the human resources related issues, which are becoming increasingly complex and diversified.

The evaluation of the company's performance stands at the basis of excellence models allowing organizations to identify strengths, opportunities for improvement, and systematic monitoring. A proposed research direction is the development of its own system of quantitative indicators capable to finely register the company's performance upon each change dictated by the different stages of the HR management systems implementation, respectively their specific practices. This system of quantitative indicators will be built based on the enterprise's quality costs and economic and financial results. Since the indicators relevant for researching the performance of human resources management systems differ from one branch to another, and in such research, it is impossible to address all sectors of activity in an economy, the author aims to address the branch of rail transport.

2. Literature review

An important aspect of human resource management is the evaluation of performance in the organization, because through evaluation we can better understand the dynamic nature of professional development, as noted by some authors (Guskey, 2000). The evaluation helps us see professional development as a continuous process, and not as a common "event" produced in the employee's life. In addition to this characteristic of continuity, however, the evaluation process also has the attribute of complexity, so we must imagine that anything can be evaluated - even evaluation can be evaluated. From the perspective of human resources management, customers and shareholders are considered the dominant stakeholders (Daneci-Patrau, 2019).

The application of evaluation methods must be supported by the entire management body in order to obtain the highest possible efficiency. The lack of this support does nothing but distort the objective results without sufficient involvement in the project, diminishing or delaying the emergence of enhancing elements following the process. Yet, not all research, which investigated the effects of implementing HR management systems, speaks about the company's performance increase.

On the contrary, some researchers argue that the overall effects of this process are costly and time-consuming (Mallak et al., 2017) or lead to increased bureaucracy (Seddon, 2012).) In this context, one of the proposed research directions is the quantitative and qualitative view of the relationship between the different practices (methods, techniques and tools) of the quality management systems and company performance.

The aim of this research project was to provide managers with quantitative and qualitative work models, methods and tools for evaluating the HR management system performance. The research problem is timely and current (Hutu, 2013; Stancu, 2017) because during its transition to a market economy, Romania's efforts to join the structures of the European Union, its organizations must become more efficient to cope with the new competitive environment. Or, considering the implementation of different quality management systems as a way to streamline the enterprise, can be an additional chance for organizations in Romania to cope with the new conditions.

3. Research methodology

The research problem is related to a practical problem as it will provide managers with quantitative models, methods, and work tools for evaluating the performance of quality-related programs and making the most appropriate decisions. The research problem covers a field little researched both in the Romanian specialty literature and internationally. To date, very little on the effects of different practices of HR quality management systems has been covered in specialty literature. Most studies conducted in this regard consist of qualitative investigations of managers' perceptions on quality improvement.

The analysis of human resources performance in railway transport, which is the general objective of this project, aims to identify the performance indicators of employees on the one hand and on the other hand, to present the links and correlations between wage level, seniority, labor conditions and the requirements of the positions for the employees at the CFR Constanţa station. In addition, it is desired to determine the elements that would lead to the increase of the HR performance level in the railway transport activity. The creation of the data file is based on the information provided by the Human Resources department within the C.N.CFR Regional Infrastructure Constanta Company, applied to 45 people, employees of Constanta station. Initially, 14 variables were introduced (3 nominal and 11 numerical), namely: name, surname, gender, nationality, wage, age, background, level of education, position, age categories, salary categories, locality, marital status and seniority at work. Subsequently, 3 more numerical variables were introduced, namely: seniority increase, gross salary and tax.

In the *BD7sal* file, the data was entered into the cells of the *Data View* sheet in the *Data Editor* window opened by the *New Data* command from the *File* menu. The result of this approach is presented below:

Figure no. 1 Database of employees of CFR Constanta station

bdsal7.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	nume	String	15	0	numele salaria	None	None	8	Left	Nominal
2	prenume	String	12	0	prenumele sal	None	None	8	Left	Nominal
3	varsta	Numeric	8	0	varsta salariatu	None	None	8	Right	Scale
4	functie	Numeric	8	0	fuctia de incad	{1, sef statie}..	None	8	Right	Scale
5	salariu	Numeric	8	2	salariul de inca	None	None	8	Right	Scale
6	vechime	Numeric	8	1	vechimea total	None	None	8	Right	Scale
7	studii	Numeric	8	0	tipul de invata	{1, superioare}..	None	8	Right	Scale
8	stare_civila	Numeric	8	0		{1, casatorit}...	None	8	Right	Scale
9	sex	Numeric	8	0		{1, masculin}...	None	8	Right	Scale
10	nationalitat	Numeric	8	0		{1, romana}...	None	8	Right	Scale
11	mediu	Numeric	7	0		{1, urban}...	None	8	Right	Scale
12	localitate	String	15	0	localitate de d	None	None	8	Left	Nominal
13	categ_an	Numeric	8	0	categorii de var	{1, <25 ani}...	None	8	Right	Scale
14	categ_sal	Numeric	8	0	categorii de sa	{1, <800 lei}...	None	8	Right	Scale

bdsal7.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : nume balinisteanu

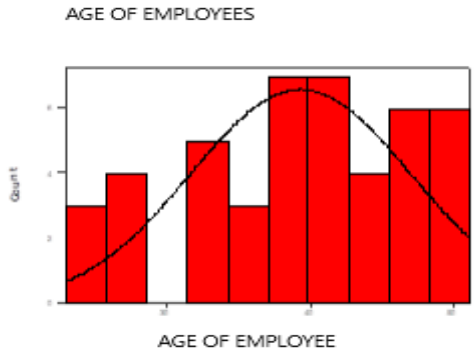
	nume	prenume	varsta	functie	salariu	vechime	studii	stare_civila	sex	nationalitate	mediu	localitate	categ_an	categ_sal
1	balinisteanu	marilena	39	2	1062,00	21,0	2	1	2	1	1	constanta	4	3
2	costea	gabriela	35	1	1093,00	12,0	1	1	2	1	1	constanta	4	3
3	pana	marian	39	2	1062,00	20,0	2	4	1	3	2	techirghiol	4	3
4	bledea	daniela	25	3	1000,00	2,0	2	2	2	3	1	mangalia	1	2
5	barleanu	maria	50	6	733,00	28,0	3	3	2	1	1	constanta	6	1
6	nastuta	florina	39	2	1062,00	16,5	1	2	2	1	2	cumpana	4	1
7	stan	smaranda	44	3	1000,00	24,0	2	3	2	1	1	constanta	5	2
8	mergeani	iuliana	40	6	733,00	18,5	3	1	2	3	1	constanta	4	1
9	gheorghiu	cristina	42	5	842,00	21,0	2	1	2	1	1	constanta	5	2
10	ghilinta	maricara	46	7	561,00	23,0	3	1	2	1	1	fetesti	6	1
11	mosteanu	ion	33	4	1000,00	14,0	1	1	1	1	1	efore	3	2
12	panait	daniela	37	3	1000,00	17,0	2	1	2	1	1	constanta	4	2
13	stanciu	costica	50	7	561,00	33,0	3	2	1	1	2	cobadin	6	1
14	pop	alexandru	43	2	1062,00	23,0	2	3	1	1	1	constanta	5	3
15	calin	maria	51	6	733,00	33,0	2	4	2	1	2	ciocarlia	6	1

Source: The author

The graphs synthetically show, in visual form, a statistical distribution. With the help of graphs, there is an overview of the data in one image. The choice of the graph for the representation of a distribution is made according to the purpose pursued and depends essentially on the number of variables considered, as well as their type.

The *frequency curve* is obtained by adjusting the histogram and is used to verify the normality of a distribution. The frequency curve can overlap the histogram, representing the corresponding theoretical distribution, with the same mean and variant.

Figure no. 2 Frequency Histogram and Curve for the age variable



Source: The author

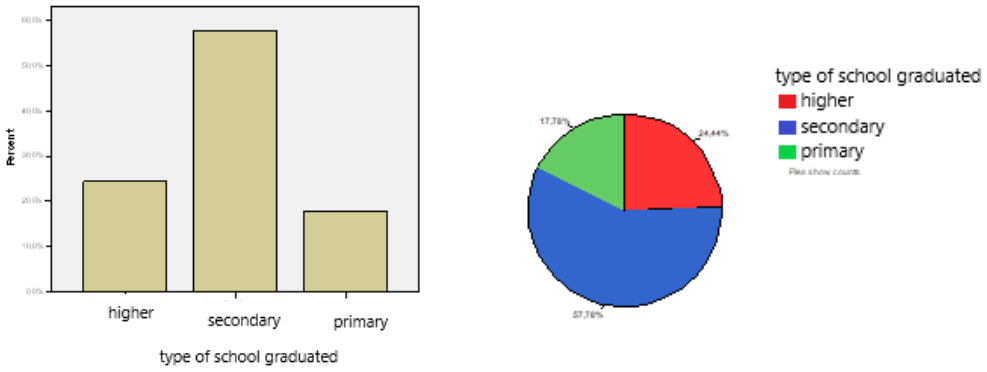
Following the construction algorithm, we obtained the histogram for the distribution of employees by age, which shows that, overall, the sample has an asymmetric distribution by age on the left, with a predominance in adulthood. Ages ranging from 38-42 and 46-50 have the highest frequencies, totaling about half of the sample volume.

As observed in Figure 3 in terms of type of school graduated, employees with high school education (57.78%) dominate over those with higher education (24.44%).

The Pie Chart and Bar chart are a method to synthesize a set of nominal (categorical) data. Each circle sector of the Pie chart represents a category, its area being proportional to the number of cases in this category of the nominal variable.

The Bar chart is often used to illustrate the categories of a distribution in a convenient way. The bars have the same base, equal to the unit, and the height proportional to the frequency of the category, so that the areas of each bar represent the number of cases of the considered category.

Figure no. 3 BAR and PIE charts for the variable (EDUCATION) using the GRAPH menu



Source: The author

The calculation of the indicators for the central tendency, dispersion and form of a univariate statistical distribution with the help of SPSS can be achieved in several ways. The usual is: *Analyze* menu - *Descriptive Statistics* command *Descriptives* option that opens the dialog box with the same title where we will select the variable/variables for which we want to calculate the parameters of a distribution, as well as the indicators to be calculated.

Table no.1 Age of employees and basic salary distribution parameters

Descriptives

		Statistic	Std. Erro
age of employee	Mean	39,24	1,136
	95% Confiden Lower Bound	36,96	
	Interval for M Upper Bound	41,53	
	5% Trimmed Mean	39,49	
	Median	40,00	
	Variance	58,053	
	Std. Deviation	7,619	
	Minimum	23	
	Maximum	51	
	Range	28	
	Interquartile Range	13	
	Skewness	-,491	,354
	Kurtosis	-,569	,695

Descriptives

		Statistic	Std. Error
basic salary	Mean	26,1333	8,72054
	95% Confidenc Lower Bound	88,2509	
	Interval for Me Upper Bound	84,0158	
	5% Trimmed Mean	29,7654	
	Median	00,0000	
	Variance	119,118	
	Std. Deviation	2,66322	
	Minimum	561,00	
	Maximum	1392,00	
	Range	831,00	
	Interquartile Range	329,00	
	Skewness	-,521	,354
	Kurtosis	-,252	,695

Source: The author

4. Results interpretation

1. The asymmetry coefficient (skewness) in the case of the age variable ranges between -1 and 0. This indicates the presence of a negative asymmetric distribution, with deviation to the left. This also emerged from the histogram for the same variable. The same interpretation applies for the distribution according to the basic salary.

2. The coefficient of vaulting (kurtosis) is a measure of the spread of each observation around a central value. In our case, for both variables it is negative, so it indicates a weaker grouping around the central value, the frequency curve is flatter and the distribution is called platykurtic distribution.

3. For the age variable, the mean is 39.24 years, very close to the median (40 years), and the standard deviation (sigma) is 7.69 years with a low value while the variant can be called close to mean value. In this case the interpretation is as follows: because the differences from the mean are not significant, it results that the mean is representative for the sample chosen as the basis of study.

4. For the *basic salary* variable the mean is 2926.1333 RON, the confidence interval for the mean (the interval in which most people are included) is 2868.25 - 2984.01 RON. The standard deviation (192.66 RON) has a small value compared to the mean, so the coefficient of variation has a value below 35%, which means that the population is approximately homogeneous. There is a weak concentration of values around the mean, so the mean is approximately representative. On the

other hand, the negative asymmetry coefficient indicates the presence of a negative asymmetric distribution, with a deviation to the left, while the negative kurtosis suggests that the frequencies have a more flattened curve and the distribution is platykurtic, as well as the distribution by age.

The frequency distribution "wage categories*age categories" expresses the distribution of the sample of people observed simultaneously after two variables were considered, showing how many people in a certain age category have a certain level of wage. The example presented in figure 5 shows that no person aged between 31-35 has a salary lower than 2800 RON, while the group over 45 years is comprised of the most (8) with salaries lower than 2800 RON.

*Table no.2. Frequency distribution "wage categories * age categories"*

	Age categories						
Wage categories	< 25 yrs	26-30 yrs	31-35 yrs	36-40 yrs	41-45 yrs	>45 yrs	Total
< 2800 ron	1	1	0	4	2	8	16
2801-3000 ron	1	0	2	3	4	3	13
3001-3200 ron	1	3	3	5	2	1	15
>3201 ron	0	0	0	1	0	0	1
Total	3	4	5	13	8	12	45

Source: The author

5. Testing the equality of means for two independent samples (Indep-Sample T test)

In SPSS program, the testing of two means can target either two independent populations or 2 dependent populations (groups). *Independent-Sample T Test* is a procedure that is applied to independent samples. This procedure tests whether the means of 2 groups are equal.Two tests will be performed: one regarding the basic salary by employees' origin, in which per the null hypothesis we will assume that the wage mean of the 2 groups (urban, rural) do not differ significantly and a second test regarding the basic salary by employees' gender, in which we will assume by the null hypothesis that there are no large pay gaps between males and females.

After applying the *Independent-Sample T Test* procedure for the first case, we obtained the *t test* equal to 1.139, with 43 degrees of freedom and a Sig probability of 0.228 (higher than 0.05), which shows us that for the means of the two groups (urban - 2946.90 RON and rural - 2875 RON), it cannot be concluded that it differs significantly, the null hypothesis being accepted. We reach the same finding by observing the confidence interval for the difference between the 2 values. The interval contains the value zero, therefore it cannot be concluded that the difference between the mean values of the two groups differs significantly (table 3).

Table no.3 Output from the Independent-Sample T Test for "basic salary - mean"

origin	N	Mean	Standard deviation	Std. Error Mean
Basic wage urban	32	29,469,063	186.89869	33.03933
rural	13	28,750,000	204.64970	56.75961

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
basic wage	Equal variances assumed	1,498	,228	1,139	43	,261	71,90625	63,15397	-55,45586	199,26836
	Equal variances not assumed			1,095	20,594	,286	71,90625	65,67535	-64,83709	208,64959

Source: The author

After applying the *Independent-Sample T Test* procedure for case 2, we obtained the *t test* equal to 2.321, with 43 degrees of freedom and a Sig. probability of 0.007 (lower than 0.05), which shows that for the means of the two groups - male - 3007.70 RON and female - 2876.60 RON) it can be concluded that it differs significantly, the null hypothesis, which assumed equality of means, being rejected. We reach the same finding by observing the confidence interval for the difference between the 2 values. The interval does not contain the value zero. As a result, it can be concluded that the difference between the mean values of the two groups differs significantly (table no. 4).

Table no.4 Output in the Independent-Sample T Test for "basic salary - gender"

sex	N	Mean	Standard deviation	Std. Error Mean
Basic salary male	17	3,007.70590	142.20222	34.4891
female	28	2,876.60710	204.64970	38.6255

Independent Samples Test

		Levene's Test for Equality of Variance		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
basic salary	Equal variances assumed	7,883	,007	2,321	43	,025	131,09874	56,48959	17,17662	245,02086
	Equal variances not assumed			2,532	42,079	,015	131,09874	51,78250	26,60322	235,59426

Source: The author

Testing the significance of the correlation coefficient *value* starts from **t**he hypothesis that there is no correlation between variables. Verification of hypothesis H0 is done using the t test for the simple correlation coefficient. The calculated value of t is compared with the theoretical value obtained in table t (Student), for n-2 degrees of freedom and the established level of significance. If *t calc. > t table*, then Ho is rejected, and it is concluded, with a considered risk (usually 5%), that the value of the correlation coefficient does not equal 0; respectively, that there is a significant connection between the investigated variables, ergo the correlation coefficient is statistically significant. In order to perform the correlation analysis, we will consider two distinct cases, namely: we will study the correlation between tax and employee's age, on the one hand and on the other hand the correlation between tax and gross salary (which we know are highly dependent, because the tax results from the gross salary to which the 16% quota is applied. The approach for each case is applied in SPSS as follows: *Analyze* menu - *Correlate* command - *Bivariate* option.

For the analysis of the correlation between the **age** and **tax** variables, by the Ho hypothesis, we assume that there is no connection between the tax paid by each employee and their age. Following the previous approach, the output in table 5 is obtained:

Table no.5 SPSS output for correlation analysis (case 1)

Correlations

		age of employee	tax
age of employee	Pearson Correlation	1	-,365*
	Sig. (2-tailed)		,014
	N	45	45
tax	Pearson Correlation	-,365*	1
	Sig. (2-tailed)	,014	
	N	45	45

*. Correlation is significant at the 0.05 level (2-tailed).

Source: The author

Interpretation. For the considered example, a Pearson correlation coefficient equal to -0.365 was obtained, which suggests that there is an inverse, weak correlation between the variables, the value of the coefficient being negative and closer to zero than to 1. The correlation coefficient is tested by use of test t. The corresponding Sig. value equal to 0.014, is lower than 0.05, the null hypothesis is accepted, that is, there is no significant correlation between the two variables (tax and age).

To highlight such correlation, I chose a second case to analyze the correlation between tax and gross salary. By the Ho hypothesis, I assumed that there is no link between the tax paid by each employee and the related gross salary. Following the approach as in the previous case, the output in figure 9 is obtained:

Table no.6 SPSS output for correlation analysis (case 2)

Correlations

		IMPOZIT	SAL_BRUT
tax	Pearson Correlation	1	1,000**
	Sig. (2-tailed)		,000
	N	45	45
gross wage	Pearson Correlation	1,000**	1
	Sig. (2-tailed)	,000	
	N	45	45

**. Correlation is significant at the 0.01 level (2-tailed).

Source: The author

Interpretation. For the considered example, a Pearson correlation coefficient equal to 1 was obtained, which suggests that there is a direct, perfect connection between the variables, the value of the coefficient being positive and exactly 1.

6. Conclusions

The aim of this research project was to provide managers with quantitative and qualitative work models, methods and tools for evaluating the quality management system performance.

Among the main contributions of the research to the development of knowledge in the field, we can list:

- The development of a coherent research conceptual framework and methodology, which would allow the quantitative and qualitative study of the quality management systems performance.
- Application of the research methodology within complex field research carried out in Romanian enterprises.
- Clarification and systematization of the basic terminology for the quality management systems performance assessment.
- Broadening knowledge and developing the existing volume of information in the field of quality management systems performance assessment.
- Defining and systematizing quality costs.
- Discussing some models and methods to assess the existing performance in the specialized literature.
- Developing an own model for the assessment of quality management systems performance.
- Building relevant performance indicators for evaluating the quality management systems performance.
- Identifying the effects of implementing quality management systems (ISO 9000: 2015).

As a future direction of research, the author aims to develop and implement software that will allow permanent monitoring of the HR management systems performance with specific quantitative indicators. Basically, this IT system will be able to gather, classify, summarize, and report that information that will underpin the decision-making process and quality-related activities planning in an organization.

The creation of an information system will eliminate the manual processing of the collected data, and the information user will be able to "shift" from analytical to synthetic, and vice versa, an essential condition for the good perception of the real situation in a company.

7. References

- Daneci-Patrau, D., 2019. *Comportament organizational* [Organizational conduct]. Craiova: Sitech Publishing House
- Guskey, T.R., 2000. *Evaluating Professional Development*. Thousand Oaks: Cirwin Press, Inc.
- Hutu, C.A., 2013. *Cultura, schimbare competitie* [Culture, Change, Competition]. Bucharest: Economica Publishing House
- Mallak, L.S., Larry A. Bringelson, D.M. Lynth, 2017. A cultural study of ISO 9000 certification. *International Journal of Quality & Reliability Management.*
- Seddon, H.R., DeLone, W.H. McLean, E.R., 2012. *Information systems success revisited*. Kogod Coll. of Bus., American Univ., Washington, DC, USA
- Stancu, I., 2017. *Finance: financial market theory, enterprise finance, analysis and financial management*. Bucharest: Economica Publishing House.