

# The Logistic Regression as an Instrument of the Personnel Policy in an Organization

David B. Katz\*<sup>1</sup>, Diana I. Makhmutova<sup>2</sup>, Nadezhda A. Opokina<sup>3</sup>

<sup>1</sup>Kazan Federal University, Institute of Mathematics and Mechanics named after N.I. Lobachevsky, Candidate of Physical and Mathematical Sciences, <sup>2</sup>Kazan Federal University, Institute of Mathematics and Mechanics named after N.I. Lobachevsky, <sup>3</sup>Kazan Federal University, Institute of Mathematics and Mechanics named after N.I. Lobachevsky, Candidate of Physical and Mathematical Sciences,

E-mail: [katzdavid89@gmail.com](mailto:katzdavid89@gmail.com), [dimakhmutova@gmail.com](mailto:dimakhmutova@gmail.com), [opnadin@mail.ru](mailto:opnadin@mail.ru), Contact: 89172570455

Received: 21st October 2017 Accepted: 16th November 2017, Published: 31st December 2017

## Abstract

Satisfaction with work is one of the most important factors in ensuring high productivity and stability of an enterprise. The purpose of this paper is to identify and analyze the factors influencing the satisfaction, and to develop a model that allows us to determine the possible problem of high personnel turnover and the mechanisms for its resolution. In the survey there were used data from the questionnaire provided by the Russian Monitoring of the Economic Situation and People Health of the National Research University - Higher School of Economics (RLMS-HSE) and the binary logistic regression method, in which the dependent variable is labor satisfaction, and the independent variables are groups of social, economic and institutional factors. The model for the construction activity sphere developed with the help of the statistical software package SPSS, makes it possible to calculate the probability that an employee under study is satisfied with his/her work, and showed that the average duration of a working day at the primary work location, the level of education, the satisfaction with one's material situation, opportunities to improve housing conditions are the most valuable among the factors under consideration. The authors obtained an effective tool that can be used in personnel management both to determine the threat of staff turnover and to select the necessary measure of influence on sensitive indicators in order to increase job satisfaction of an employee under study.

**Keywords:** Job Satisfaction, Binary Logistic Regression, Character Selection, Employee Classification, Classification Table, ROC Analysis

## Introduction

In this paper, the attention of the authors is drawn to the method of determining the satisfaction of employees in the construction sector with their work. The relevance of the research lies in the fact that advanced approaches in personnel management that take into account the modern challenges of a difficult period of the economy development are one of the key factors in increasing the competitiveness and profitability of production, since the human factor is one of the determining factors in production growth [1].

Methods of personnel management, where sociological studies are taken into account in addition to economic

analysis, create an effective management mechanism for working with personnel in various areas of human professional activity and are indicators of the enterprise development level [2].

The idea of the nature and trends of changes in job satisfaction provides important information about the stability within the organization that is necessary to improve the effectiveness of management strategies, productivity and quality of productive activities, promptly stimulating staff and reducing the number of violations of labor discipline and the level of staff turnover [3, 4].

Analysis of the literature helped to find out that "job satisfaction" is the object of consideration of various sciences and researchers from various countries since the 20s of the last century [5]. The authors of the paper focused on the job satisfaction position as a personal assessment of their work situation. According to E. Mayo, the factors influencing the satisfaction with work include the form of ownership of the enterprise and their share in the ownership as a form of external control and the possibility of participation in management. According to F. Herzberg, satisfaction / dissatisfaction with work is due to the influence of a set of motivating factors responsible for the process of "satisfaction - lack of satisfaction" and hygiene factors responsible for the process of "dissatisfaction - the lack of dissatisfaction" [6]. These factors included: the content of labor (the diversity of skills, the identity and importance of the task, autonomy and feedback), personal and collective achievements such as qualifications, education, length of service, prestige of the position, responsibility, career opportunities, labor conditions, relationships in the team, pay, methods used to control employees, some aspects of the employee's personal life, and security [7,8].

Satisfaction with work is dynamic. The activity type [9,10] and the social group of the employee under survey can lead to a change in the factors affecting satisfaction [11], therefore, we focused on one area of activity when developing the satisfaction model.

Factors motivating labor activity also have influence [12]. Different researchers include the following factors to their list: salary, social insurance, the possibility of career growth, the administrative regime in the enterprise, working microclimate, maintaining a positive assessment of achievements.

The purpose of our study is to identify and analyze the degree of influence caused by various groups of factors on the level of job satisfaction, which is necessary to determine the practical means and instruments that affect the increase of satisfaction, and also to build a mathematical model of job satisfaction by the example of employees in the construction sector as a convenient tool for monitoring and management.

Logistic regression as a part of Data Mining methods is increasingly used, including in personnel management, as a method of studying dependencies and finding regularities in analyzing data of an arbitrary nature, both quantitative and qualitative [13]. Logistic regression is a convenient tool both in forecasting the state of a binary dependent variable, in our case - satisfaction with the work of personnel, and when selecting the attributes that affect this variable [14]. The independent variables are the socio-demographic and economic data on employees considered. In the computational experiment, the results of calculating the probability from the model developed on the basis of the data obtained in the construction industry are shown. The selection of characteristics is made on the basis of the received evaluations of logistic regression coefficients. Personnel decisions can be made on the basis of the calculated probabilities.

The evaluation of the validity of the developed model is based on the classification table and ROC analysis [15].

**Methods**

In this paper, the authors consider job satisfaction only in terms of "yes-no" indicators, so such a statistical method is used to construct the model as binary logistic regression based on the logistic law of probability distribution. In this method, the dependent variable takes only two values, and the independent variable can have an arbitrary data type, namely, they can be quantitative, logistic, and categorical. The constructed logit model calculates for each employee of an enterprise the probability of occurrence of such events as "satisfied with work" and "dissatisfied with work", which depends on the values of the factor attributes chosen on the basis of the earlier literature data

analysis. To calculate the probability of these events, the formula is used:

$$p = \frac{1}{1 + e^{-z}}, (1)$$

Where

$$z = \sum_{i=1}^n b_i X_i + a,$$

$i$  - the quantity of factor attributes affecting the resulting characteristic;  $X_i$ -values of factor attributes ( $i = \overline{1, n}$ );  $b_i$  and  $a$  are the model coefficients to be found ( $i = \overline{1, n}$ ).

In a case if the probability calculated by the formula (1) will be less than 0.5, it can be assumed that an event will not occur, i.e., the corresponding employee is dissatisfied and in this case it is necessary to make personnel decisions; otherwise the employee is satisfied with his/her work [16].

To assess the adequacy of the model, a classification table and an index for the area under the ROC curve, AUC, are used. ROC curve shows the dependence of proportion of correctly classified positive outcomes (sensitivity) on the proportion of incorrectly classified negative outcomes (unit minus specificity). The area under the AUC curve changes from 0.5 (no separation) to 1 (ideal separation) [17].

**Results**

We have constructed a model of satisfaction with the work for the subjects from the Russian Federation. The information base for this study was data provided by the Russian Monitoring of the Economic Situation and Population Health of the National Research University - Higher School of Economics (RLMS-HSE) [18].

The studies were carried out with a breakdown of 10.2016-01.2017, which corresponds to 25th Wave [18].

We proposed the questionnaire for adults and the following indicators were selected for this study (Table 1).

**Table 1. Indicators Influencing Satisfaction with the Work of a Respondent**

Types of Indicators	Indicators
Resultant Indicator	1. Satisfaction with the work as a whole
Social Indicators	2. Respondent's place of residence
	3. Sex of the respondent
	4. Respondent's age
Economic Indicators	5. The level of wages at the primary work location
	6. The average length of a working day at the primary work location
	7. The average length of the work week at the primary work location
	8. Work experience in today's workplace
	9. The form of ownership of the enterprise where the respondent works
	10. Ownership of the respondent in the enterprise
Institutional Indicators	11. The level of education
	12. Method of employment
	13. Satisfaction with your material situation

	14. Having the opportunity to save money for buying expensive things
	15. Possibility to spend holidays abroad
	16. Possibility to improve housing conditions
	17. Having the opportunity to pay for the education of children in a university

To construct the model, we used logistic regression choosing "Satisfaction with work as a whole" as the resultant indicator, and we chose the other indicators listed in Table 2 in the capacity of independent variables. Respondents who answered the question of satisfaction with their work as a whole could choose the following options: 1 - "Completely satisfied", 2 - "More satisfied", 3- "Half-and-half ", 4 - "Not very satisfied" 5 -"Rather dissatisfied". Interviewees who choose answers 3, 4 or 5 are among those who work with insufficient performance or who can change their place of work. With the purpose to reveal the factors influencing the resultant indicator "Satisfaction with

work as a whole", it was re-coded from the categorical variable into a logistic one. The variable was assigned the value "1" if the respondent answered "1" or "2"; "0" if the respondent gave the answer either "3", or "4", or "5". This model was developed using the statistical software package SPSS. The sample was the respondents who indicated the type of their activity "Construction industry". The number of such respondents, i.e. the sample size was 310 people.

For the sample under study, a valid percentage was calculated for satisfaction with the work of the respondents (Fig. 1).

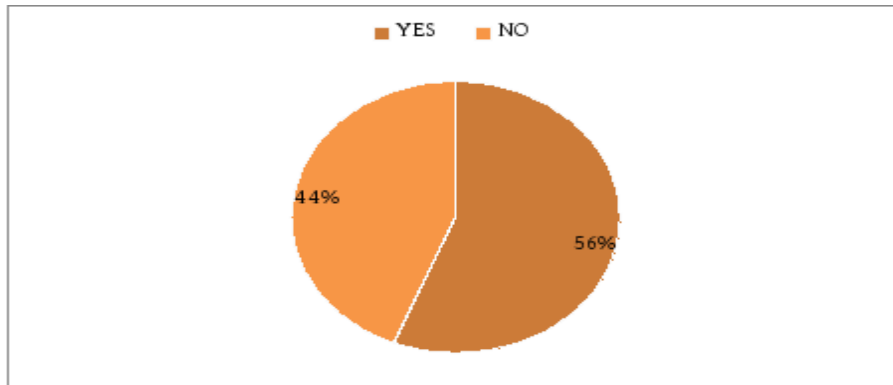


Fig. 1 - Valid Percentage for Job Satisfaction of Respondents

It can be seen from the diagram presented above, that the number of people unsatisfied with their work in this industry is quite high. Construction is one of the priority sectors in the development of the country and a high percentage of people who are not satisfied with

their work give grounds for investigating the causes and indicators affecting satisfaction.

When using logistic regression, we obtained a classification table (Table 2) from which it can be seen that the number of correctly recognized cases was 72.6%.

Table 2 Classification Table

Satisfaction with Work	Yes	No	Percentage of Correct Results
Yes	110	28	79.7
No	41	73	64.0
Total Percentage			72.6

When developing this model, we have received that the following indicators affect the satisfaction of respondents: average hours of work at the primary work location, level of education, satisfaction with their material situation, and the availability of opportunities to improve housing conditions. The

significance of these variables is 0.004; 0.003; 0.001 and 0.009, respectively.

In our study, the categorical variables of the average duration of the working day at the primary work location, level of education, and satisfaction with their material situation were re-coded into logistic ones. The coding of variables is presented in Table 3.

Table 3. Coding of Categorical Variables

	Value si	Parameter Coding										
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
How many hours on average	4	1	0	0	0	0	0	0	0	0	0	0
	7	0	1	0	0	0	0	0	0	0	0	0

	8	0	0	1	0	0	0	0	0	0	0	0
	9	0	0	0	1	0	0	0	0	0	0	0
	10	0	0	0	0	1	0	0	0	0	0	0
	11	0	0	0	0	0	1	0	0	0	0	0
	12	0	0	0	0	0	0	1	0	0	0	0
	14	0	0	0	0	0	0	0	1	0	0	0
	15	0	0	0	0	0	0	0	0	1	0	0
	24	0	0	0	0	0	0	0	0	0	1	0
	Irregular working hours	0	0	0	0	0	0	0	0	0	0	1
	Difficult to answer	0	0	0	0	0	0	0	0	0	0	0
	What is your highest level of education, for which you have received a certificate, diploma? (Y <sub>i</sub> )	Primary or modern secondary school	1	0	0	0	0	0				
Higher grade school		0	1	0	0	0	0					
Professional courses of drivers, tractor drivers, accountants, typists, etc.		0	0	1	0	0	0					
Technical school, factory apprenticeship school, factory apprenticeship training without secondary education		0	0	0	1	0	0					
Vocational school with secondary education, technical school		0	0	0	0	1	0					
Technical school, medical, musical, art, or pedagogical school		0	0	0	0	0	1					
Post graduate study, residency		0	0	0	0	0	0	1				
Institute, university, academy - specialist diploma		0	0	0	0	0	0	0				
How satisfied are you with your current financial situation? (U <sub>i</sub> )	Completely satisfied	1	0	0	0							
	Rather satisfied	0	1	0	0							
	Half and half	0	0	1	0							
	Not really satisfied	0	0	0	1							
	Completely dissatisfied	0	0	0	0							

We also introduce the variable  $W$  as follows:  $W = 1$  if the family has the opportunity to improve housing conditions, otherwise  $W = 2$ .

Next, the coefficients were calculated; they are presented in Table 4.

**Table 4. Coefficients for Independent Variables**

Variables	Coefficients	Variables	Coefficients
X <sub>1</sub>	-21.301	U <sub>1</sub>	19.937
X <sub>2</sub>	-21.796	U <sub>2</sub>	-1.620
X <sub>3</sub>	-21.564	U <sub>3</sub>	-0.332
X <sub>4</sub>	-22.225	U <sub>4</sub>	-1.127
X <sub>5</sub>	-22.642	W	1.585
X <sub>6</sub>	-42.116	Y <sub>1</sub>	2.562
X <sub>7</sub>	-20.806	Y <sub>2</sub>	0.262
X <sub>8</sub>	-0.778	Y <sub>3</sub>	0.760
X <sub>9</sub>	-42.906	Y <sub>4</sub>	0.071
X <sub>10</sub>	-21.544	Y <sub>5</sub>	1.277
X <sub>11</sub>	-42.854	Y <sub>6</sub>	0.624
Constant	18.382	Y <sub>7</sub>	24.420

A positive coefficient indicates that this factor increases the probability of the outcome being analyzed and vice versa. Using these coefficients, we can calculate the probability  $p$  for the respondent under

study would be satisfied with his/her work. For example, some employee graduated from the technical school ( Y<sub>6</sub> ), has an eight-hour day ( X<sub>3</sub> ); he/she has no opportunity to improve his/her living conditions

(W) and is not very satisfied with his/her financial situation (U4). Then  $Z = 18.382 + 0.624 - 21.564 + 1.585 * 2 - 1.127 = -3.685$ , and the probability that he/she will be satisfied with the current work is

$$p = \frac{1}{1 + e^{3.685}} = 0,024. \text{ Since the result is less than}$$

50%, then in this case we can say about the negative attitude of this person to his/her work.

The authors have constructed the ROC-curve and calculated the AUC value which was equal to 0.738.

### Summary

Based on the results of the study, a model was constructed for calculating the probability of work satisfaction. Assumptions about the influence on the satisfaction of various groups of factors were confirmed: they are social, economic, and institutional. The use of data from the annual national survey of the Russian Monitoring of the Economic Situation and Population Health of the National Research University - Higher School of Economics hindered the possibility of studying groups of psychological and motivational factors. In the developed model, the share of correctly predicted results is 72.6%, and  $0.7 < \text{AUC} < 0.8$ , what corresponds to the good quality of the model. The use of questionnaires that include more factors and the restriction of polls by employees of the same enterprise will lead to an increase in the accuracy of the model. Also, according to the results of the research, it was concluded that each of the following factors: the average length of the working day at the primary work location, the level of education, the satisfaction with one's material situation, the availability of an opportunity to improve housing conditions, is particularly sensitive to the model. Knowledge of this fact allows management to build a strategy for personnel management and shows the levers of effective influence to improve work satisfaction (arrange a shorter day, raise wages or give out a bonus) and prevent the loss of valuable personnel. The model constructed by the authors makes it possible to calculate how the probability will be changed upon this or that change of factors, and choose the best option.

### Conclusion

In the work, the indicators influencing the satisfaction of the respondent's work were considered. In the course of the study, factors influencing the respondent's satisfaction with his/her work were selected from them, i.e. a list of variables for modeling was generated. Using the statistical software package SPSS, a model of respondent satisfaction with their work based on binary logistic regression was constructed. Special attention is paid to the study of the quality and prognostic properties of the constructed model.

### Acknowledgments

This work was carried out in accordance with the Russian Government Program of Competitive Growth of Kazan Federal University

### References

- [1] Kabasheva, I. A., Rudaleva, I. A., Bulnina, I. S., Askhatova, L. I. Organizational factors affecting employee innovative behavior // *Mediterranean Journal of Social Sciences*. - 2015, - Vol. 6, Issue 1S3. - P. 435-439.
- Personnel Management in an Organization: Textbook // Edited by A.Ya. Kibanova. - 4th edition, revised and enlarged - M.: INFRA-M. - 2010. - 695 p.
- [2] Blau, G. J., Boal, K. P. Conceptualizing how job involvement and organizational commitment affect turnover and absenteeism // *Academy of Management Review*. - 1987. - 12. - P. 288-300.
- Iaffaldano, M. T., Muchinsky, P. M. Job Satisfaction and Job Performance: A meta-analysis // *Psychological Bulletin*. - 1985. - 97(2). - P.P. 251-273.
- [3] Rudaleva, I. A., Kabasheva, I. A., Kovaleva, E. R. Factor analysis of labor satisfaction of the managerial staff working in an organization // *Life Science Journal*. - 2014. - V. 11, №12. - P. 157.
- Herzberg, F., Mausner, B. and Snyderman, B. B. *Motivation to Work* // John Wiley and Sons Inc., New York, NY. - 1959.
- [4] Ivanova, T. Yu. Socio-demographic and organizational factors of job satisfaction / T. Yu. Ivanova, E.I.Rasskazova // *Psychological journal*. - 2013. - V. 34. - No. 6. - P. 40-52.
- [5] Spector, P.E. *Job Satisfaction: Application, Assessment, Causes and Consequences* // Thousand Oaks, CA: Sage. - 1997.
- [6] Blegen, M. *Nurses Job Satisfaction: A meta-analysis of related variables* // *Nursing Research*. - 1993. - 42 (1). - P.P. 36-41.
- [7] Malyakov, P.L. *Labor of a Teacher: Analysis of Approaches to the Concept of Satisfaction with Work* // *Actual Directions of Scientific Research: From Theory to Practice*. - 2015. - No. 1 (3). - P. 195-197.
- [8] Patrushev, V.D. *Satisfaction with work: socio-economic aspects* / V.D.Patrushev, N.A.Kalmakan. - Moscow: Nauka. - 1993. - 112 p.
- [9] Maslow, A. *Motivation and Personality* / A. Maslow. Harper and Row, New York. 1954.
- [10] Hastie, T., Tibshirani, R., Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* // Springer, 2009. Corr. 10th printing 2013. - 745 p.
- [11] Hosmer, D. W. *Applied logistic regression*. [Text] / D. W. Hosmer, S. Lemeshow, R. X. Sturdivant. - New York: John Wiley and Sons. - 2013. - 528 p. - 3rd ed.
- [12] Pallant, J. *SPSS Survival Manual. A step by step guide to data analysis using the SPSS program* / J. Pallant. - (4th ed.). Sydney: Allen&Unwin. - 2011.

[13] Kleinbaum, D. G. Logistic Regression /D. G. Kleinbaum, M. Klein. – New-York: Springer-Verlag. – 2002. – 513 p.

[14] Sorokin, A.S. On the validation of the logistic regression model in credit scoring. Internet-magazine "NAUKOVODENIE" Issue 2, March - April 2014 - p.1-14. (<http://naukovedenie.ru>)

[15] Russian Monitoring of the Economic Situation and People Health by the NRU-HSE (RLMS-HSE), conducted by the National Research University "Higher School of Economics" and "Demoscope" with the participation of the Population Center of the University of North Carolina in Chapel Hill and the Institute of Sociology RAS. (Websites of RLMS-HSE surveys: <http://www.cpc.unc.edu/projects/rlms> and <http://www.hse.ru/rlms>)".