

Statistical analysis of responses concerning the importance of human and production or services issues in various companies

Krzysztof Knop¹

¹ Institute of Production Engineering, Faculty of Management, Czestochowa University of Technology, Armii Krajowej 19B, 42-201 Czestochowa, Poland, e-mail: kknop@poczta.fm

Abstract. In the article the results of statistical analysis is presented with the use of basic statistical parameters and box and whisker plot conducted for the dataset, containing answers to the question E9 from BOST survey. Measures of central tendency and variation were employed for analysis of the data connected with the assessment of human problems and production-related issues in three analysed companies: a steel company, a plastics processing company and a retail chain.

Key words – statistical parameters, BOST research, 9th Toyota management principle, human and production/service issues

1. Introduction. Statistical parameters - measures of central tendency and variation. Box plot diagram

In addition to graphs and tables of numbers there are often used statistical parameters to describe sets of numbers. There are two major categories of these parameters. One group of parameters measures how a set of numbers is centered around a particular point on a line scale or, in other words, where (around what value) the numbers bunch together. This category of parameters is called measures of central tendency. The most famous statistical parameter from this category is *the mean or average*, next *the median*, *the mode*.

The second major category of statistical parameters is measures of variation. Measures of variation tell how far the numbers are scattered about the center value of the set. They are also called measures of dis-

persion. There are three common parameters of variation: *the range*, *standard deviation*, and *variance*. While measures of central tendency are indispensable in statistics, measures of variation provide another important yet different picture of a distribution of numbers. Measures of central tendency are useful in understanding how scores cluster about a center value, and measures of variation are useful in understanding how far, wide, or deep the high scores are scattered about the center value ([HTTP://WWW.SAGEPUB.COM UPMDATA/47775_CH_3.PDF](http://www.sagepub.com/upmdata/47775_CH_3.PDF)).

A box and whisker plot (called a boxplot), is a graph that presents information from a five-number summary (at a glance). It is especially useful for indicating whether a distribution is skewed and whether there are potential unusual observations (outliers) in the data set. Box and whisker plots are ideal for comparing distributions because the centre, spread and

overall range are immediately apparent. A box and whisker plot is a way of summarizing a set of data measured on an interval scale. It is often used in explanatory data analysis. This type of graph is used to show the shape of the distribution, its central value, and its variability (HTTP://WWW.STATCAN.GC.CA/EDU/POWER-POUVOIR/CH12/5214889-ENG.HTM).

A box and whisker plot makes medians visible, quartiles (Q_1 and Q_3) and the range (highest and lowest value) of dependent variable for each category of grouping variable (OSTASIEWICZ S., RUSNAK Z., SIEDLECKA U. 1999). The box-and-whisker plot provides a lot of information about empirical distribution. Location of the box in relation to the axis determines location of the distribution; a point inside the box determines central tendency. The length of the box representing the difference between lower and upper quartile shows differentiation of the data for 50% of central units. The wider the box, the bigger the differentiation. The length of the whole chart indicates dispersion of this data throughout the set (PUŁASKA-TURYNA B. 2008). An unsymmetrical position of the median and completely different length prove high differentiation and strong asymmetry of empirical distribution (LUSZNIEWICZ A., SŁABY T. 2008).

2. Statistical analysis of responses concerning E9 question

2.1. Analysis in the steel company

The attempt to transform 14 principles of Toyota management (LIKER J. K. 2005) into questions was reflected in BOST survey (BORKOWSKI S. 2012). Evaluation of 9th Toyota management principle was presented in BOST survey in E9 question: Assess using scale 1 to 9, the importance, in your company, of:

- human issues, - production/service issues.
where: 1 – lack of interest, 9 – high interest.

Responses to the E9 question can be analysed and presented in a different way, with the use, for example, of the managerial grid (BORKOWSKI S., PIESZCZOCH D. 2009) or with the use of graphical presentation of data (graphs, charts) in accordance to the principle that every picture tells more than a hundred words (data).

On the basis of the set of $n = 60$ answers to the question E9 asked among steelworks' employees, statistical analysis of the collected data was made.

The results for calculations of basic statistical parameters (i.e. mean \bar{x} , median Me , mode Mo , range R , standard deviation s , lower quartile Q_1 , upper quartile Q_3 , coefficient of variation V_s , skewness SK , kurtosis KU) in the analysed dataset are presented in Table 1.

Table 1. Statistical parameters of collected data concerning E9 question. Refers to the steel company

Class sign	\bar{x}	Me	Mo	$\sum Mo$	Σ	\bar{x}	Min	Max	Q_1	Q_3	R	S^2	S	V_s	SK	KU
SLI	5.5	5	5	11	333	2	9	9	4	7	7	3.95	1.99	35.8	0.02	-1.03
ZPI	7.2	8	8	23	436	3	9	9	7	8.5	6	3.22	1.79	24.7	-1.13	0.14

Source: own study

An evaluation of descriptive statistics for $n=60$ answers was obtained, i.e.:

- arithmetic mean for a group of assessments of human problems amounts to $\bar{x} = 5.5$ and $\bar{x} = 7.2$ for production issues,
- the lowest value of assessment of the degree of interest of the enterprise in human issues, made by an employee equals 2, highest being 9, while the range, calculated as a difference between the highest and lowest assessment in the dataset, amounts to $R = 7$; in the case of production issues the lowest observed assessment value is 3 (highest 8), results range amounts to $R = 6$
- assessment 5 is the most frequent value in SL set, in the case of production issues this value amounts to 8
- the calculated median at the level of $Me = 5$ for SL set means that half of all the answers had value at least equal to 5, the same number had value not lower than the determined one. In the case of production issues, the middle value amounts to 8
- a low average level of differentiation of answers, being $S = 1.99$ for SL set and $S = 1.79$ for ZP set, was observed
- the typical area of changeability defined by the equation of $\bar{x} - s < x_{typ} < \bar{x} + s$ for human problems amounts to $3.51 < x_{typ} < 7.49$, whereas for production issues $5.41 < x_{typ} < 8.99$. The values of about 2/3 of all the persons in the investigated population can be found in this area,
- the value of skewness amounts to 0.02 for SL set, which means a low, positive asymmetry of distribution of 60 answers and noticeable, negative asym-

metry which amounts to $SK = -1.13$ observed for assessment of production issues.

Box plot (Fig. 1) was used for graphical representation of relationships between the selected statistical parameters.

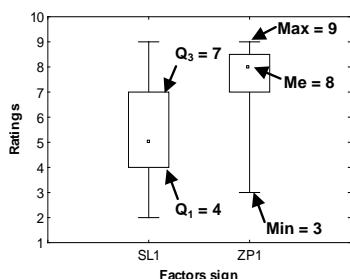


Fig. 1. Box plot. Refers to the steel company.

Source: own study

As results from Fig. 1 demonstrate, differentiation of the results for SL set is higher than for ZP set. Whiskers for SL set are of equal length, which proves symmetrical distribution. Longer left whisker in relation to the right one shows left skewness (asymmetry) for ZP set. Shift of the median closer to lower quartile Q_1 proves right skewness in 50% of SL population. Location of the median closer to the upper quartile Q_3 informs about left skewness for 50% of ZP population.

2.2. Analysis in the plastics processing company

The statistical evaluation of the set of data concerning the importance of human problems and production issues was undertaken in the plastics processing enterprise. The analysis was conducted on the basis of the results from answers of 84 employees. The results are presented in Table 2.

Table 2. Statistical parameters of collected data concerning E9 question. Refers to the plastics processing company

Class sign	\bar{x}	M_e	M_o	$\sum Mo$	Σ	Min	Max	Q_1	Q_3	R	S^2	S	V_1	SK	KU
SL2	6.2	7	7	30	524	1	9	6	7	8	3.29	1.81	29.1	-0.96	0.88
ZP2	6.8	7	9	28	572	1	9	5.5	9	8	5.24	2.29	33.6	-0.95	-0.01

Source: own study

Half of the data concerning human problems is maintained within the range of 6 to 7 whereas in the case of production issues, from 6 to 9. Mean arithmetic level of assessments concerning human problems

amounts to $\bar{x} = 6.2$ while average differentiation of ± 1.81 , for production issues is $\bar{x} = 6.8 \pm 2.29$. A typical area of variability defined by the equation of $\bar{x} - s < x_{typ} < \bar{x} + s$ for SL set amounts to $4.39 < x_{typ} < 8.01$, for ZP set being $4.51 < x_{typ} < 9$. Value 2 is the lowest value of assessment of human issues, 9 being the highest value while the range of results amounts to $R=7$. In the case of production issues, value 3 means the lowest assessment while 8 is the highest one with the range of results $R=6$. The value of assessment equal 7 is the most frequent among human problems. For assessment of production issues, the most frequent value is 9. The value of skewness of -0.96 and -0.95 denotes low and negative asymmetry of SL and ZP set. Kurtosis (KU) equal 0.88 means that the analysed distribution of SL is more pointed than in the case of the normal curve. The value of the coefficient of variation means 29.1 and 33.6 – percentage average differentiation of the observed variable for SL and ZP respectively. This corresponds to the assessment of mean differentiation of variants of the observed measurable variable.

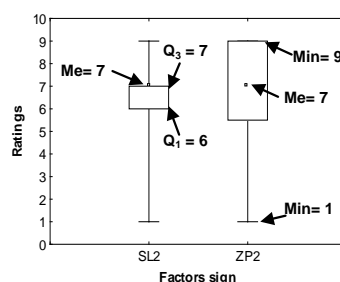


Fig. 2. Box plot. Refers to the plastics processing company.

Source: own study

Visible dispersion of assessments for the whole SL set can be observed. Identical dispersion can be observed for SL set. Differentiation of assessments for 50% of central units is low for the SL set. In order to assess skewness of the distribution visually on the basis of the described chart, one should bear in mind its two features: (1) length of whiskers located on the left and on the right to the box and (2) location of the median inside the box. As results from the chart show, left asymmetry of the distribution of SL data can be observed and in 50% of its central units, which can be proved by a longer left whisker in relation to right whisker and a shift of median to the right in relation to

the central line of the box. In the case of ZP set, right skewness is visible for 50% of units.

2.3. Analysis in the retail chain

In order to analyse basic statistical parameters concerning datasets for assessment of the level of interest in human problems and services-related issues the retail chain, BOST opinion poll was conducted among the group of 65 employees in the investigated unit. The values for descriptive statistics are presented in Table 3.

Table 3. Statistical parameters of collected data concerning E9 question. Refers to the retail chain

Class sign	\bar{x}	Me	Mo	ΣMo	Σ	Min	Max	Q_1	Q_3	R	S^2	S	V_s	SK	KU
SL3	5.4	6	7	14	349	1	9	4	7	8	5.42	2.33	43.4	-0.32	-1.15
ZP3	7.5	8	8	26	487	4	9	7	8	5	1.25	1.12	14.9	-1.08	1.76

Source: own study

Mean assessment of employees for importance of human problems in the company was formed at the level of 5.4 with average differentiation of the results $s=2.33$; for the production problems mean level of assessment was 7.5 ± 1.12 . Typical variability area for SL set amounts to $3.07 < x_{typ} < 7.73$, while for ZP set this value is $6.38 < x_{typ} < 8.62$. The most frequent assessment was 7 for human problems and 8 for production issues. Mean value (median) amounts to 6 for SL set, which means that half of assessments had values of 6 at the most, the same part of assessment had the value not lower than the determined. Range for the obtained results for SL amounts to $R = 8$ ($x_{max} - x_{min} = 9 - 1 = 8$), and $R = 5$ for ZP set ($x_{max} - x_{min} = 9 - 4 = 5$). Differentiation of the results measured by the coefficient of variation V_s is higher for SL set (43.4%) than in ZP set (14.9%).

Box plot (Fig. 3) was used for graphical verification of asymmetry, location and differentiation.

In both of the analysed datasets, left asymmetry of the whole distribution and 50% of its central units can be observed. Dispersion of the whole set of assessment is higher for SL than ZP; this also concerns differentiation for 50% of central units (box is wider for SL set).

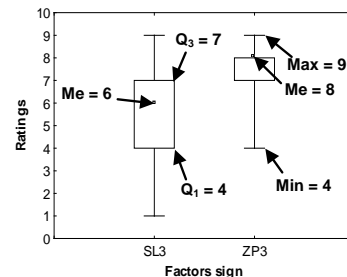


Fig. 3. Box plot. Refers to the retail chain.

Source: own study

3. Summary

In the article the statistical parameter usage was presented to analyse the answers concerning the E9 question contained in the BOST survey. A group of eleven statistical parameters in range of central tendency, variation and asymmetry allowed the analysis of the structure of datasets concerning assessments of interests in human problems and production/service issues in three of the analysed companies. Analysis of basic statistical parameters, supported by the box plots, provided a lot of valuable information about the responses of employees, which can be further used (by the division manager) for taking corrective actions.

References

1. BORKOWSKI S. 2012. *Zasady zarządzania Toyoty w pytaniach. Wyniki Badań BOST*. PTM. Warszawa.
2. BORKOWSKI S., PIESZCZOCH D. 2009. *Concepts of Leadership Maps and Managerial Grid*. In: Toyotarity. Styles of Management. Borkowski S., Shevtsova O. J. (ed.). Yurii V. Makovetsky. Dnipropetrovsk.
3. LIKER J. K. 2005. *Droga Toyoty. 14 zasad zarządzania wiodącej firmy produkcyjnej świata*. MT Biznes. Warszawa.
4. LUSZNIWICZ A., SŁABY T. 2008. *Statystyka z pakietem komputerowym Statistica PL. Teoria i zastosowania*. C.H. Beck. Warszawa.
5. OSTASIEWICZ S., RUSNAK Z., SIEDLECKA U. 1999. *Statystyka. Elementy teorii i zadania*. Publisher Akademia Ekonomiczna im. Oskara Langego. Wrocław.
6. PUŁASKA-TURYNA B. 2008. *Statystyka dla ekonomistów*. Difin. Warszawa.
7. http://www.sagepub.com/upm-data/47775_ch_3.pdf.
8. <http://www.statcan.gc.ca/edu/power-pouvoir/ch12/5214889-eng.htm>.