

Mladá veda

Young Science



Mladá veda

Young Science

MEDZINÁRODNÝ VEDECKÝ ČASOPIS MLADÁ VEDA / YOUNG SCIENCE

Číslo 2, ročník 12., vydané v júni 2024

ISSN 1339-3189, EV 167/23/EPP

Kontakt: info@mladaveda.sk, tel.: +421 908 546 716, www.mladaveda.sk

Fotografia na obálke: Èze, Francúzsko. © Branislav A. Švorc, foto.branisko.at

REDAKČNÁ RADA

prof. Ing. Peter Adamišin, PhD. (Katedra environmentálneho manažmentu, Prešovská univerzita, Prešov)

doc. Dr. Pavel Chromý, PhD. (Katedra sociálnej geografie a regionálneho rozvoje, Univerzita Karlova, Praha)

Mgr. Jakub Köry, PhD. (School of Mathematics & Statistics, University of Glasgow, Glasgow)

prof. Dr. Paul Robert Magocsi (Chair of Ukrainian Studies, University of Toronto; Royal Society of Canada)

Ing. Lucia Mikušová, PhD. (Ústav biochémie, výživy a ochrany zdravia, Slovenská technická univerzita, Bratislava)

doc. Ing. Peter Skok, CSc. (Ekomos s. r. o., Prešov)

Mgr. Monika Šavelová, PhD. (Katedra translitológie, Univerzita Konštantína Filozofa, Nitra)

prof. Ing. Róbert Štefko, Ph.D. (Katedra marketingu a medzinárodného obchodu, Prešovská univerzita, Prešov)

prof. PhDr. Peter Švorc, CSc., predseda (Inštitút histórie, Prešovská univerzita, Prešov)

doc. Ing. Petr Tománek, CSc. (Katedra verejnej ekonomiky, Vysoká škola báňská - Technická univerzita, Ostrava)

Mgr. Michal Garaj, PhD. (Katedra politických vied, Univerzita sv. Cyrila a Metoda, Trnava)

REDAKCIA

Mgr. Branislav A. Švorc, PhD., šéfredaktor (Vydavateľstvo UNIVERSUM, Prešov)

Mgr. Martin Hajduk, PhD. (Banícke múzeum, Rožňava)

PhDr. Magdaléna Keresztesová, PhD. (Fakulta stredoeurópskych štúdií UKF, Nitra)

RNDr. Richard Nikischer, Ph.D. (Ministerstvo pro místní rozvoj ČR, Praha)

PhDr. Veronika Trstianska, PhD. (Ústav stredoeurópskych jazykov a kultúr FSŠ UKF, Nitra)

Mgr. Veronika Zuskáčová (Geografický ústav, Masarykova univerzita, Brno)

VYDAVATEĽ

Vydavateľstvo UNIVERSUM, spol. s r. o.

www.universum-eu.sk

Javorinská 26, 080 01 Prešov

Slovenská republika

© Mladá veda / Young Science. Akékoľvek šírenie a rozmnožovanie textu, fotografií, údajov a iných informácií je možné len s písomným povolením redakcie.

ANALYSIS OF THE QUALITY LEVEL OF TRANSPORT SERVICES USING SPECIALISED SOFTWARE

ANALÝZA ÚROVNE KVALITY DOPRAVNÝCH SLUŽIEB POMOCOU
ŠPECIALIZOVANÉHO SOFTVÉRU

Dominika Jonasíková¹

Dominika Jonasíková pôsobí ako interná doktorandka na Katedre cestnej a mestskej dopravy na Fakulte prevádzky a ekonomiky dopravy a spojov Žilinskej univerzity v Žiline. Vo svojej dizertačnej práci sa venuje výskumu vplyvu kvality dopravných služieb v cestnej nákladnej doprave na výkonnosť a spoľahlivosť cestnej nákladnej dopravy. Cieľom výskumu je objektívnym spôsobom hodnotiť vplyv kvality a prínosy vyplývajúce zo zvyšovania kvality dopravných služieb.

Dominika Jonasíková works as an internal doctoral student at the Department of Road and Urban Transport at the Faculty of Operation and Economics of Transport and Communications of the University of Žilina. In her dissertation thesis, she is researching the impact of the quality of transport services in road freight transport on the performance and reliability of road freight transport. The aim of her research is to objectively assess the impact of quality and the benefits resulting from improving the quality of transport services.

Abstract

Quality assessment enables organisations to identify the strengths and weaknesses of their services, which is essential for continuous improvement and adaptation to customer needs. IBM SPSS Statistics is one of the most widely used statistical analysis tools that enables efficient processing and evaluation of survey data. With its wide range of features and intuitive interface, SPSS is the ideal tool for data analysis in a variety of fields, including transportation. The aim of this paper is to present methods for data analysis in transport services, focusing on the use of IBM SPSS Statistics. Using chi-square test, the relationships between different quality criteria will be investigated, which will provide valuable information for improving the services provided and increasing customer satisfaction. The results of this analysis will highlight the importance of a systematic approach to quality management and the benefits it brings to organisations operating in a dynamic and competitive environment.

Keywords: chi-square, IBM SPSS statistics, analysis, methods

¹ Work address: Ing. Dominika Jonasíková, Katedra cestnej a mestskej dopravy, Fakulta prevádzky a ekonomiky dopravy a spojov, Žilinská univerzita v Žiline, Univerzitná 8215/1, 010 26 Žilina
E-mail: jonasikova1@stud.uniza.sk

Abstrakt

Hodnotenie kvality umožňuje organizáciám identifikovať silné a slabé stránky ich služieb, čo je nevyhnutné pre neustále zlepšovanie a prispôsobovanie sa potrebám zákazníkov. IBM SPSS Statistics je jedným z najpoužívanejších nástrojov na štatistickú analýzu, ktorý umožňuje efektívne spracovanie a vyhodnocovanie dát z dotazníkových prieskumov. Vďaka širokej škále funkcií a intuitívnemu rozhraniu je SPSS ideálnym nástrojom pre analýzu údajov v rôznych oblastiach, vrátane dopravy. Cieľom článku je predstaviť metódy analýzy údajov v dopravných službách, so zameraním na využitie IBM SPSS Statistics. Pomocou chí-kvadrát testu budú skúmané vzťahy medzi rôznymi kritériami kvality, čo poskytne cenné informácie pre zlepšenie poskytovaných služieb a zvýšenie spokojnosti zákazníkov. Výsledky tejto analýzy poukážu na dôležitosť systematického prístupu k riadeniu kvality a prínosy, ktoré z neho vyplývajú pre organizácie pôsobiace v dynamickom a konkurenčnom prostredí.

Kľúčové slová: chí-kvadrát, IBM SPSS Statistics, analýza, metódy

Introduction

A questionnaire is a structured set of questions that are necessary to collect specific information. A questionnaire can equally be described as a focused conversation that is on paper, online or in digital format. This 'conversation' consists of a set of well thought out questions designed to elicit specific information from people (respondents). The questions in a questionnaire can cover a wide range of topics and thus help researchers, organisations, or educators to understand opinions preferences, facts, or behaviours [1], [2], [3].

A survey is considered a broader research method that involves collecting and analysing data from multiple respondents. Surveys can include a variety of instruments such as just questionnaires, interviews, focus groups, or observations [3].

In general, data analysis is key to uncovering trends in development or relationships between certain criteria. Through data analysis, it is possible to identify opportunities, solve problems or predict future events with greater accuracy. For example, in business or organizational management, data analysis can help identify certain customer preferences or streamline operational processes. In quality management, feedback analysis can help identify areas for improvement, while in product or service quality data analysis can reveal production bottlenecks that need to be addressed.

Data evaluation is a process that involves validating and verifying results, assessing performance, or visualising results. The aim is to ensure that the results of the data analysis are accurate, reliable, and usable for decision-making. Different types of methods and tools are used in practice to analyse and evaluate data effectively.

The process of processing the questionnaire survey data

Data analysis is a complex method of inspecting, cleaning, transforming and modelling data to discover useful information, draw conclusions and support decision-making. It is a multifaceted process involving different techniques and methodologies to interpret data from different sources in different formats, both structured and unstructured [4].

The data analysis process is a structured sequence of steps that lead from raw data to actionable insights. The data analysis process includes the following steps:

1. *Defining objectives and questions* - The first step in the data analysis process is to define the objectives and formulate specific questions that the analysis is intended to answer. This is a key step in the data analysis process because it sets the direction for the entire process. It also involves understanding the situation or problem at hand and identifying the data needed to address it and defining metrics or indicators to evaluate the results.
2. *Data collection* - gathering relevant data from pre-selected sources (e.g., a questionnaire survey) to ensure data quality and integrity. The data collected may be quantitative (numerical) or qualitative (verbal), depending on the nature of the problem and the questions being asked.
3. *Data cleaning* - involves identifying or correcting erroneous information, missing values and inconsistencies in the data set. Clean data is critical for accurate analysis.
4. *Data analysis* - involves applying statistical or mathematical techniques to data to discover patterns, relationships, or trends. Various tools and software are available for this purpose, such as MS Excel, Python, and specialized software such as SPSS, SAS.
5. *Data interpretation and visualization* - after data analysis, the next step is to interpret the results and visualize them. This may involve creating tables, graphs or other visual representations of the data. Data visualisation helps to make complex data more understandable and provides a clear picture of the findings.
6. *Deployment* - the findings are implemented into real solutions or strategies to ensure that the recommendations are implemented based on the data [4], [5], [6], [7].

Methods of analysis of the questionnaire survey data

A variety of methods and tools are used in data analysis, each with its own unique purpose and application. Several types of methods are used to analyse data, some of which will be discussed in the following sections.

1. *Descriptive analysis* - Answers the question: "What happened?". It involves sorting, data processing, and descriptive characterization. It is a set of techniques that summarizes the attributes of the data in a user-friendly form [5]. Descriptive analytics allows us to characterize data based on its properties [7].
2. *Exploratory analysis* - Answering the question: "How to explore data relationships?". Exploratory data analysis (EDA) focuses on exploring and understanding data without pre-understood hypotheses. EDA involves data visualizations that allow researchers to identify and define patterns and characteristics in a dataset that they would not otherwise be able to look for. It helps to generate hypotheses for further analysis [7], [8], [9].
3. *Diagnostic analysis* - Answers the question: "Why did this happen?". Diagnostic analysis is a branch of data analysis that focuses on examining historical data and identifying the root causes of various outcomes, events, or trends. The goal of diagnostic analysis is to understand cause and effect relationships in data [7], [10].
4. *Predictive analysis* - Answers the question: "What will happen?". Predictive analysis is a category of data analysis aimed at making predictions about future outcomes based on historical data and analytical techniques such as statistical modelling and

machine learning. The science of predictive analytics can generate future insights with a significant degree of accuracy [7], [11].

5. *Prescriptive analysis* - Answers the question: "How does this happen?". Prescriptive analysis is the process of using data to determine the optimal course of action. After considering all relevant factors, this type of analysis provides recommendations for next steps. For this reason, prescriptive analysis is a valuable tool for data-driven decision making [7], [12].

IBM SPSS Statistics

Today, data analytics is one of the fastest growing areas as organizations around the world look for ways to gain insights from their data. In practice, there are a number of data analysis tools available, such as SPSS - Statistical Package for the Social Sciences, which is used for statistical analysis.

SPSS is a set of software programmes for analysing scientific data that was first used in the social sciences. Today, SPSS is used in fields ranging from economics to transportation. SPSS offers an environment for rapid visual modelling that ranges from the smallest to the most complex models. The data obtained from SPSS are used for surveys, data evaluation, market research, etc [13].

IBM SPSS Statistics software is used for the elaboration of the paper, in which the data from the questionnaire survey concerning the collection and dispensing points of shipments are evaluated. The main features offered by IBM SPSS Statistics software include:

1. *Data collection and import* - the software can import data from various sources such as Excel spreadsheets, databases, text files, etc. It is also possible to create new data files directly in the software and manually enter data.
2. *Data cleaning* - it is possible to remove errors, add missing values or transform variables in the software.
3. *Statistical analysis* - the software offers various forms of statistics, which include:
 - *Descriptive statistics* - calculating basic statistical indicators such as mean, median, variance, standard deviation, and percentiles.
 - *Inferential statistics* - hypothesis testing, t-tests, ANOVA, correlations, regression and time series analysis.
 - *Advanced statistical methods* - factor analysis, cluster analysis, discriminant analysis, logistic regression and more.
4. *Data visualization* - the software can create graphs, charts, histograms, box plots, scatter plots, etc.
5. *Reporting* - in the software it is possible to create reports with the results of analyses, which can be exported to different formats (PDF, MS Word, MS Excel).
6. *Integration* - the software can be integrated with other software tools and programming languages such as R and Python for enhanced analytical capabilities [13], [14], [15].

Creation of the questionnaire survey data set

In the case where data are not recorded on a computer during the questionnaire survey, it is necessary to convert the responses obtained into electronic form when evaluating them. The

rows of the table are defined for the answers of individual respondents, the columns contain the answers to individual questions or parts of questions. The rows are most referred to as cases in statistical software systems, while the term record is used in database programs. The aim of a question is to find out the value of a statistical characteristic. In statistical software systems, the equivalent term variable is used. The detected values of a variable are recorded in a certain column of a table, which is referred to as a field (item) in database systems (Řezanková, 2017). Table 1 shows the most common terminology used in statistical software systems.

	1. variable	2. variable	3. variable	...
1. case				
2. case				
3. case				
...				

Table 1 – Structure of the basic data matrix

Resource: Own elaboration by the author according to [14]

Types of variables in the dataset

Answers to closed questions are assigned word or number codes. Answers to open questions are recorded in their original form. The term category is used for coded and integer values. Variables whose values are categories are called categorical. Examples are as follows:

- nationality (Slovak, Czech, ...),
- level of education (primary, secondary, university, ...),
- number of children (0, 1, 2, ...) [14].

At the same time, the examples define different levels of relationships between categories - in the first case the categories cannot be arranged, in the second case the categories can be arranged, and in the third case we can additionally calculate the difference. In this we can speak of scales (scales) of measurement. The basic breakdown of the scales is as follows:

- *nominal* - for these values we can only determine that they are different, we cannot determine their order,
- *ordinal* - for these values we can determine their order, but we cannot determine how much greater one is than the other;
- *interval* - for these values, we can determine how much greater one value is than the other (these are numeric values);
- *proportional* - with these values we can determine how much and how many times one value is greater than the other (the scale must contain only positive values) [14], [16].

Each response from the respondent must have recorded in a separate variable. The type of scale is the basis for dividing the variables into:

- *nominal* - e.g. type of employment, type of school attended, type of product or service,
- *ordinal* - for example, level of education attained, degree of importance of a factor or criterion, degree of satisfaction,
- *quantitative* - which are generally divided into:
 - interval - temperature, degree of humidity,
 - proportional - number of members in the household,

- discrete - number of computers in the household,
- continuous - age of the respondent, area of the dwelling, monthly household expenditure on food [14], [16].

Evaluation of the questionnaire survey in IBM SPSS Statistics software

Principles of detecting the dependence of two variables from a questionnaire

The dependence of two variables can be either symmetric (reciprocal) or asymmetric (unilateral). The basic test used for detecting the dependence of two categorical traits (without considering the direction of dependence) is the chi-square test of independence [14].

The dependence of two nominal variables is called contingency. The chi-square test of independence is the basis for detecting this dependence. Denote the relative abundances in the base set as π_{ij} (their point estimates are the abundances of p_{ij}), then write the null hypothesis of independence in the form $H_0: \pi_{ij} = p_i + p_{+j}$, where $p_i + p_{+j}$ is the relative abundance expected under independence. We test this null hypothesis against the hypothesis $H_1: \pi_{ij} \neq p_i + p_{+j}$, for at least one pair i, j ($i \neq j$). The Pearson chi-square statistic can be used as a test criterion [14]. The Pearson chi-square statistic is calculated according to the formula:

$$\chi_P^2 = \sum_{i=1}^R \sum_{j=1}^S \frac{(n_{ij} - m_{ij})^2}{m_{ij}} \quad (1)$$

where: R is the number of rows in the contingency table,

S is the number of columns of the contingency table,

$(n_{ij} - m_{ij})^2$ is the difference between the observed counts and the expected counts in the i -th row and j -th column,

m_{ij} is the expected abundance in the i -th row and j -th column.

In addition to chi-square statistics, likelihood ratio is used to test the independence of two nominal variables [14]. The likelihood ratio is calculated according to the relation:

$$G^2 = 2 \sum_{i=1}^R \sum_{j=1}^S n_{ij} \ln \frac{n_{ij}}{m_{ij}} \quad (2)$$

where: n_{ij} is the observed frequency in the i -th row and j -th column,

\ln is the natural logarithm.

Dependence measures for two nominal variables can also be used for other types of variables - for example, for one nominal and one ordinal variable [14].

Case study of detecting the dependence of two variables from a questionnaire

A questionnaire survey was used to develop the case study. The questionnaire was distributed to respondents who use the services of the dispensaries. A total of 65 responses were received from respondents, however 65 respondents completed the identification questions, where the identification section asked whether the respondent used the services of the collection and dispensing points. 7 respondents did not use the services of the collection and dispensing points and therefore did not continue to complete the questionnaire. Two respondents, for an undetermined reason, did not respond further after the question "How often do you use the dispensing points designated for parcels?". This is why the part of the questionnaire dealing with the importance of selected criteria for this form of service was completed by 56 respondents.

Hypothesis:

- Null Hypothesis - H_0 : There is no statistically significant relationship between gender (I1) and services used by customers (I3).
- Alternative Hypothesis - H_1 : There is a statistically significant relationship between gender (I1) and services used by customers (I3).

Let us take an example of the test of independence of the variables I1 - Gender and I3 - Services used by customers. Both these variables are nominal. The contingency table of absolute counts, shown in Table 2, shows the observed absolute counts (Count) and, in the case of independence, the expected absolute counts (Expected Count).

Gender * Services used by customers Crosstabulation

Count

		Services used by customers			Total
		dispatch shipments	dispatch and submint shipments	I don't use the dispensing points	
Gender	Female	40	0	0	40
	Male	10	8	7	25
Total		50	8	7	65

Table 2 – Contingency table of observed and expected absolute abundances
Resource: Author's own processing using IBM SPSS Statistics [17]

Without the use of statistical software, we would calculate the Pearson chi-square statistic according to the following formula (1):

$$\chi_P^2 = \frac{(40 - 30,8)^2}{30,8} + \frac{(0 - 4,9)^2}{4,9} + \dots + \frac{(8 - 3,1)^2}{3,1} + \frac{(7 - 2,7)^2}{2,7} = 31,200$$

We would calculate the belief ratio according to formula (2):

$$G^2 = 2 * (40 * \ln(40/30,8) + \dots + 7 * \ln(7/2,7)) = 36,576$$

Testing at the 5% significance level, we compare the calculated value of the statistic χ^2 , or G^2 , with the quantile $\chi_{0,95}^2[(2-1)(3-1)] = \chi_{0,95}^2[2] = 5,99$. In both cases, we obtained a value higher than the stated critical value. That is why in both cases we reject the null hypothesis of independence of the variables I1 - Gender and I3 - Services used by customers.

Using IBM SPSS Statistics statistical software, make the Analyse → Descriptive Statistics → Crosstabs option, then select Statistics and choose Chi-square. We obtain the result shown in Table 3. The first line of Pearson Chi-Square refers to the Pearson chi-square statistic. The second line of Likelihood Ratio refers to the likelihood ratio. The third row gives the total range of the set. For both statistics, the first column Value gives the calculated value of the statistic. The second column, df, lists the number of degrees of freedom. The third column lists Asymptotic Significance (2-sided), the minimum significance level from which we reject the H_0 hypothesis of independence of the observed variables. Below the table is a note that in four cells of the contingency table the expected frequencies are less than 5, which may affect the reliability of the Chi-Square test results. The minimum expected frequency in these cells is 2.69.

The high values of the test statistics (Pearson Chi-Square and Likelihood Ratio) together with the very low Asymptotic Significance values (<0.001) indicate that the differences between observed and expected values are statistically significant. This means that there is a relationship between the variables under study.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	31,200 ^a	2	<,001
Likelihood Ratio	36,576	2	<,001
N of Valid Cases	65		

a. 4 cells (66,7%) have expected count less than 5. The minimum expected count is 2,69.

Table 3 – Independence tests based on chi-square statistics
Resource: Author's own processing using IBM SPSS Statistics [17]

The obtained results which are calculated without and with statistical software agree. The minimum level of significance is less than three thousandths, therefore at 1% level of significance we reject the H_0 hypothesis of independence.

Conclusion

The evaluation of the questionnaire survey in IBM SPSS Statistics software provided important insights into the relationship between the gender of customers and the services they use. Based on the Chi-Square tests performed, we found that there is a statistically significant relationship between these variables, suggesting that gender may influence customers' preferences in their choice of services.

The results of the analysis revealed that the differences between the observed and expected values are significant, confirming that gender plays a significant role in customers'

decision making. These findings can help organisations to better tailor their services to the needs of different customer groups and thus improve overall customer satisfaction and loyalty.

The use of IBM SPSS Statistics software enabled accurate and efficient processing and interpretation of the questionnaire survey data. The results of this analysis provide valuable information that can be used for strategic decision making and service optimization within the organization. In this way, a higher degree of customer satisfaction can be achieved, ultimately contributing to the growth and profitability of the organisation.

*This article was recommended for publication in the scientific journal Young Science by:
doc. Ing. Vladimír Konečný, PhD.*

Resources

1. Oppenheim, A. N. (2001). *Questionnaire design, interviewing and attitude measurement (New edition)*. CONTINUUM London and New York.
2. Sreejesh, S., Mohapatra, S., & Anusree, M. R. (2014). *Questionnaire Design*. SpringerBooks. Business Research Methods. p. 143-159.
3. SurveySparrow. (2024). *What Are Questionnaires? Benefits, Importance and Examples*. Online. Available at: <https://surveysparrow.com/blog/questionnaires/>.
4. DataCamp. (2023). *What is Data Analysis? An Expert Guide With Examples*. Online. Available at: <https://www.datacamp.com/blog/what-is-data-analysis-expert-guide>.
5. Švábová, L., Ďurana, P. & Ďurica, M. (2022). *Deskriptívna a indukčívna štatistika*. Žilinská univerzita v Žiline: EDIS – vydavateľské centrum ŽU, 2022. ISBN 978-80-554-1839-1.
6. SimpliLearn. (2024). *What Is Data Analysis: A Comprehensive Guide*. Online. Available at: <https://www.simplilearn.com/data-analysis-methods-process-types-article>.
7. Datapine. (2023). *Your Modern Business Guide To Data Analysis Methods And Techniques*. Online. Available at: <https://www.datapine.com/blog/data-analysis-methods-and-techniques/>.
8. Stack, J. (1995). *The analysis of survey data*. *European journal of operational research*, 81(1). p. 1-16.
9. Alchemer. (no date). *What is Exploratory Data Analysis?*. Online. Available at: <https://www.alchemer.com/resources/blog/what-is-exploratory-data-analysis/>.
10. Wolniak, R., & Grebski, W. (2023). *The concept of diagnostic analytics*. Silesian University of Technology Scientific Papers. Organization and Management Series, 175. p. 650-669.
11. Kumar, V., & Garg, M. L. (2018). *Predictive analytics: a review of trends and techniques*. *International Journal of Computer Applications*, 182(1). p. 31-37.
12. HBS. (2021). *What is prescriptive analytics? 6 examples*. Online. Available at: <https://online.hbs.edu/blog/post/prescriptive-analytics>.
13. Stehlik-Barry, K., & Babinec, A. J. (2017). *Data analysis with IBM SPSS statistics*. Packt Publishing Ltd.
14. Řezanková, H. (2017). *Analýza dat z dotazníkových šetření*. Professional Publishing, Praha. Čtvrté vydání, 2017. ISBN 978-80-906594-8-3.
15. Aljandali, A. (2016). *Quantitative analysis and IBM® SPSS® statistics*. Cham, Switzerland: Springer International Publishing.
16. Almquist, Y. B., Kwart, S., & Brännström, L. (2020). *A practical guide to quantitative methods with SPSS*.
17. IBM. (no date). IBM SPSS Statistics. Available at: <https://www.ibm.com/account/reg/us-en/signup?formid=urx-19774>.

Mladá veda

Young Science

ISSN 1339-3189