

Correlation-regression analysis of the relationship between the volume of sales of goods and the use of rotary photo tables for their presentation in online stores

ABSTRACT

Aims: The aim of this article was to determine the existence of a relationship between the volume of sales of goods by an online store and the use of rotary photo tables for the presentation of goods, using mathematical methods and models.

Study design: In order to determine the relationship between the volume of sales of goods at the online store and the use of rotary photo tables for the presentation of goods, economic and mathematical modeling was carried out on the base of the data of online store "Ukrasa".

Place and Duration of Research: Russian Federation

Methodology: To achieve the goals, the research used such methods as analysis and synthesis, induction and deduction, comparison and grouping, methods of mathematical modeling (correlation-regression analysis).

Results: On the basis of mathematical methods, it was established that there is a relationship between the volume of sales of goods by the online store and the number of goods that are presented with the help of photos and videos made with the use of turntables for photography. As a result of the research, it was established that 78,85% of the total change in the volume of sales of goods is explained by the change in the presentation of goods on the site with the help of rotary photo tables.

Conclusion: The conducted research showed the importance of using rotary photo tables for presenting goods in online store. The economic interpretation of the parameters of the model shows that the increase in the number of goods presented with the help of photos and videos made with the use of rotary tables for photography will increase by one hundred units, then the volume of sales of goods will increase by 1661,9 thousand hryvnias.

Keywords: *Internet trade, rotary photo tables, product photography, scope of implementation, correlation, regression, influence.*

1. INTRODUCTION

Product photography in modern conditions is gaining more and more popularity. Many stores go online and need high-quality photos that will help to show the products from all sides and in the most attractive light. It is advisable to use rotary photo tables for product photographs for this purpose. Under such conditions, the question of determining the relationship between the volume of sales of goods in the online store and the use of rotary photo tables for the presentation of goods becomes particularly relevant. It is expedient to carry out such research with the help of a mathematical model, which is considered to be the most effective tool for the research of complex economic problems and allows modeling the main properties of real processes and phenomena.

The purpose of this research was to determine the existence of a relationship between the volume of sales of goods by an online store and the use of rotary photo tables for the

presentation of goods, using mathematical methods and models, namely correlation-regression analysis for the assessment of economic processes and trends, as well as the development of a dependence model between the volume of sales of goods by the online store and the use of rotary photo tables for the presentation of goods.

Mathematical modeling actively penetrates into various spheres of human activity and enriches many sciences. In the field of economic processes, mathematical modeling plays a very important role. Since, mathematical models allow to describe essential correlation between economic processes and phenomena, to forecast economic indicators, to develop management strategies of economic objects.

The use of mathematical methods significantly expands modeling capabilities, allows formulating new tasks, and improves the quality of management decisions. Since economic-mathematical models simulate the main properties of real processes and phenomena with the help of various mathematical apparatuses, they are one of the most effective tools for researching complex economic problems.

As a result of the review of literary sources no researches were found about determination of the relationship between the volume of sales of goods by online stores and the use of rotary photo tables to present goods to consumers. The aim of this article was to determine such a relationship.

2. MATERIAL AND METHODS

To achieve the goals, the research used such methods as analysis and synthesis, induction and deduction, comparison and grouping, as well as methods of mathematical modeling and mathematical statistics. The information base of the research is made up of the works of recognized specialists, as well as data from the online store "Ukrasa". To estimate the level of relationship between the volume of sales of goods by an online store and the use of rotary photo tables for presenting goods the following steps were applied:

1. Data collection and preliminary analysis of the collected information.

When assessing the volume of sales of goods by an online store using rotary photo tables for presenting goods, it was necessary to have all the information about the company's activities for a certain period. As part of the research, information was obtained from the open sources.

2. Selection of one or more impact indicators.

We had to apply comparative approach for such indicators as the total number of goods in the store, the number of goods presented with a description and photo, the number of goods presented with photos and videos made using photo turntables, and some others from which we had to select the base of multipliers. The choice of specific factors depended, as usually, on the availability of information, the significance of the selected factors for a particular company, as well as on the preferences of the valuer. As soon as the necessary data prepared, processing started.

3. Calculation of the values of the matrix of pair correlation coefficients, identification of the dependence of the resulting indicator on the selected factors (factor).

In the course of correlation analysis, the strength of the connection is determined as no connection, weak connection or strong connection (correlation). The nature of the relationship of factors (direct or inverse). The information obtained during the correlation analysis about the nature and strength of the relationship (correlation coefficient lies in the range from -1 to +1) needs to be used to plan the sequence of calculating the parameters of regression equations.

The dependence is considered sufficiently strong if the correlation coefficient exceeds 0.8 in absolute value, and weak if it does not exceed 0.5. In the case if this coefficient is equal to zero, the connection (correlation) is completely absent. The correlation coefficient give an objective assessment of the degree of dependence (correlation) only with a linear regression model.

4. Selecting the type of model and calculating the parameters of this model. Practice shows that one- and two-factor models are the most reliable. If it is found that two factors have a strong or complete relationship with each other, then it will be sufficient to include one of them in the regression equation. For example as a model for assessing the market value of a company, it is also possible to use a linear function but other functions described by parabolic and hyperbolic dependencies. In addition, power, exponential, logarithmic, exponential, and mixed models are also used.

5. Study of the model for significance and adequacy. To analyze the quality of the model, various techniques can be used, among which are residual analysis, outlier analysis, and coefficient of determination analysis. However, the most adequate way to check the quality of the resulting model is to calculate the coefficient of determination, which is calculated as the squared correlation coefficient (R^2). It shows which proportion of the variation of the estimated variable is taken into account in the model and is due to the influence of the selected factors. The closer the R^2 value is to 1, the more significant the model is. The boundary value is the coefficient of determination equal to 0.7. If the resulting value is less than 0.7, then the model cannot be used for estimation. The calculation of this coefficient, as well as the plotting of residuals, is performed using EXCEL tables automatically during the simulation of the regression equation.

When constructing economic and mathematical models, the adequacy of the model is revealed, that is, its compliance with the modeled object or process, which is considered essential for the study. Fisher's F-test is usually used in such cases to determine the significance of a regression.

The F-criterion value calculated from the observational data (F_p) is compared with the corresponding critical value (F_k). In the case when the value of F_p is less than the critical F_k , the equation cannot be considered significant.

6. Calculation of the value of the resulting indicator using the constructed regression model. The resulting model involves the use of information on the volume of sales of goods by an online store and the use of rotary photo tables for their presentation to consumers.

7. The stage of interpretation of the obtained results is also important for the possibility of subsequent application of the simulation results. It allows to translate simulation data from the language of statistics and mathematics into the language of economics.

3. RESULTS AND DISCUSSION

In recent years, there has been a steady trend towards the reorientation of the trade sphere into an online format.

The global pandemic has changed the attitude of users towards online trading, which has significantly affected the development of the e-commerce market. Mall closures and quarantine restrictions in 2020 forced shoppers to adopt new habits [4].

According to the [6], at the beginning of the pandemic, namely from February to April 2020, there was a decline in purchasing activity. This was due to the fact that many consumers were not yet accustomed to the new conditions of self-isolation and shopping in online stores. However, since April 2020, the situation has leveled off - the share of online purchases started to increase, and by the fourth quarter of 2020, the number of visits to online stores increased by almost 13% compared to the same period in 2019 [2].

It is noteworthy that in 2020 the share of buyers of online stores increased, but not the turnover of trade. That is, users switched to online under the influence of the pandemic, but did not buy more than before. But in 2021, the growth of both traffic and purchasing activity is observed. In 2021, the share of visits to online stores increased by 7%, and the turnover of electronic commerce increased by 34% compared to 2020 [13].

Quarantine restrictions have influenced consumer habits: people continue to buy not only food and beverages, but also electronics, jewelry, clothing, furniture, and more. Online trade has rapidly entered everyday life, thanks to which online stores have significantly improved their positions, as well as gained new customers - people of the older generation and people living in small settlements [19].

In 2021, consumer engagement metrics increased compared to last year — buyers began to spend more time on sites and visit more pages on one resource. At the same time, audience loyalty decreased: in 2020, 73% of users visited more than six sites of the same topic per day, and in 2021, this indicator increased to 76% [10].

It can be said that buyers began to browse the site even more carefully when choosing a product, because they want to study various offers before making a purchase [20].

Under such conditions, competition among online stores is steadily growing, so it is especially important to determine the factors that force the consumer to buy goods in the particular online store. One of these factors can be the creation of a high-quality image of the product on the website using the rotary tables for product photography.

The product photography is becoming more and more popular every year, and the main reason for this phenomenon is the active development of the Internet business sphere, namely: internet stores. For trading in the "network", it is desirable to have high-quality photos of what needs to be sold. A rotary table for product photography greatly facilitates the process of creating photos and videos, and also improves the quality of images [8]. A rotary table or as it is also called "3D-table" consists of a disk that rotates and a fixed base on which it is fixed. The use of this device allows you to shoot products and to create GIF-animation, as well as obtain three-dimensional images of products. The advantages of such pictures are obvious, as a result, a potential buyer can look at the product he is interested in from all sides, which has a positive effect on the volume of sales.

To build a correlation-regression model, we chose the following factors:

x_1 - Number of product items presented in the store, in total;

x_2 - The number of items of the product presented in the store, with a description and photo;

x_3 - Number of items of the product presented in the store, presented with a description and photo using a rotary table, pcs.

The data for correlation regression analysis are presented in Table 1.

Table 1. Data for carrying out a correlation-regression analysis of the influence of the relationship between the volume of goods sold and the use of rotary photo tables for their presentation in the online store "Ukrasa"

Period	The volume of sales of goods, thousand UAH.	The number of items of the product presented in the store		
		Total, pcs.	presented with description and photo, pcs.	presented with description and photo using a rotary table, pcs.
	Y	X1	X2	X3
1 quarter 2015	35587,35	3975	3211	2886

2 quarter 2015	32951,25	3978	3213	2888
3 quarter 2015	30315,15	3985	3219	2894
4 quarter 2015	32951,25	3982	3217	2891
1 quarter 2016	23559,9	2940	2761	2053
2 quarter 2016	21747,6	2955	2775	2063
3 quarter 2016	22653,75	2960	2779	2067
4 quarter 2016	22653,75	2945	2765	2057
1 quarter 2017	21452,1	3000	2787	2164
2 quarter 2017	23317,5	3155	2931	2276
3 quarter 2017	24250,2	3045	2829	2193
4 quarter 2017	24250,2	3200	2973	2308
1 quarter 2018	28038,5	4080	3200	2560
2 quarter 2018	30281,58	4100	3216	2572
3 quarter 2018	25795,42	4060	3184	2547
4 quarter 2018	28038,5	4080	3200	2560
1 quarter 2019	35324,64	4412	3133	2677
2 quarter 2019	32708	4379	3110	2657
3 quarter 2019	30091,36	4400	3125	2670
4 quarter 2019	32708	4411	3133	2677
1 quarter 2020	35369,63	4065	2915	2690
2 quarter 2020	38445,25	4235	3037	2803
3 quarter 2020	39983,06	4207	3017	2784
4 quarter 2020	39983,06	4173	2992	2762
1 quarter 2021	35727,51	4012	3312	3187
2 quarter 2021	38834,25	3788	3127	3009
3 quarter 2021	40387,62	3900	3220	3098
4 quarter 2021	40387,62	3900	3220	3098

Source: created by the author

We will analyze the relationship between the volume of sales of goods and the number of goods presented in the online store with only a description, with a description and a photo, as well as with a photo taken using a rotary photo table from 2015 to 2021 using the correlation coefficient. This coefficient characterizes the linear correlation (that is, the relationship that is given by a certain value and direction) of two or more variables.

To calculate the correlation, the Pearson formula is used:

$$r_{xy} = \left(\frac{1}{n-1}\right) \sum \left(\frac{x-\mu_x}{\sigma_x}\right) * \left(\frac{y-\mu_y}{\sigma_y}\right) \quad (1)$$

where:

r_{xy} - correlation coefficient;

n – number of periods;

x_n, y_n – values of variables x and y ;

μ_x, μ_y – average arithmetic value of x and y ;

σ_x, σ_y – standard deviations x and y .

The value of the correlation coefficient is between -1 and +1, that is, the correlation can be both positive and negative. When the value of the correlation coefficient is -1, there is a perfect negative correlation; when the value of the correlation coefficient is +1, there is a perfect positive correlation. In other cases, there is a positive correlation, a negative correlation, or no correlation between the two variables [12].

To determine the correlation between the selected indicators in the period from 2015 to 2021, we will calculate the value of the correlation indicators. The level of correlation of the factors selected for building the model with the resulting indicator is presented in Table 2.

Table 2. Correlation level of the factors selected for building the model of correlation regression analysis with the volume of sales of goods by the online store "Ukrasa"

<i>Indicator</i>	<i>Y</i>	<i>X₁</i>	<i>X₂</i>	<i>X₃</i>
Volume of sales of goods	Y	1		
The number of items of the product presented in the store	X ₁	0,7326204	1	
The number of items of the product presented in the store with a description and photo	X ₂	0,6049287	0,781582806	1
The number of product positions presented in the store with photos using rotary photo tables	X ₃	0,8879974	0,757276793	0,750261417

Source: created by the author

The value of the correlation coefficient is always in the range from -1 to +1, the closer the value is to 0, the less dependence of the indicators on each other, and accordingly, the closer the value is to 1, the higher the level of dependence [15]. Note that a correlation coefficient with a positive value means a direct relationship, and with a negative value, an inverse relationship [16]. A high degree of dependence is considered to be the value of the correlation coefficient (modulo) greater than 0.80.

The assessment of the closeness of the connection between the studied features was carried out using the Chaddock scale, the characteristics of the obtained ties are presented in the figure 1.

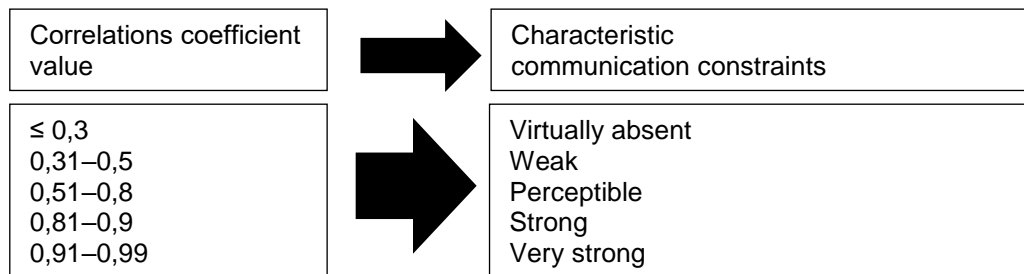


Fig. 1. Chaddock scale

Source: [17]

Based on this, it can be argued that factor X_3 which is the number of products presented with photos and videos made using rotary tables for photography (table 3), has a significant influence on the volume of sales of goods in the online store, since the correlation coefficient for this factor is larger than 0.8. Factors X_1 and X_2 do not have a significant impact on the resulting indicator, since the value of the correlation coefficients for these factors is less than 0.8. So, in order to build a correlation-regression model, it is advisable to use the indicator of the number of items of the product presented in the store with photos using rotary photo tables.

Table 3. Characteristics of the closeness of the relationship between the performance indicator and factor characteristics

Indicator	Characteristic of correlation strength
The number of items of the product presented in the store	Notable
The number of items of the product presented in the store with a description and photo	Notable
The volume of foreign economic operations with the countries of the European Union	Strong

Source: created by the author

The significance check of the regression model is carried out using Fisher's F-Criterion, the calculated value of which is defined as the ratio of the variance of the original series of observations of the studied indicator and the unbiased estimate of the variance of the residual sequence for this model.

If the estimated value with $k_1=(m)$ and $k_2=(n-m-1)$ (m is the number of factors in the model) "degrees of freedom" is larger than the tabular value at the specified level of significance, then the model is considered significant. To assess the statistical significance of the paired linear regression, the null hypothesis is put forward that the equation as a whole is statistically insignificant: $R^2=0$ at the level of significance α [18].

Next, the actual value of the F- Criterion is determined by the formula:

$$F = \frac{R^2}{1-R^2} * \frac{n-m-1}{m} (2)$$

where:

$m=1$ for paired regression.

The tabular value (F tab.) is determined based on the Fisher distribution table for a given level of significance, taking into account that the number of degrees of freedom for the total sum of squares (larger variance) is 1, and the number of degrees of freedom of the residual sum of squares (smaller variance) in linear regression is $n-2$.

F table. is the maximum possible value of the criterion under the influence of random factors with the given degrees of freedom and significance level α . The significance level α represents the probability of rejecting the correct hypothesis, given that it is true. Usually, α is taken equal to 0.05.

If the actual value of the F-Criterion is less than the table value, then there is no reason to reject the null hypothesis [16]. And if not, then the null hypothesis is rejected and the alternative hypothesis about the statistical significance of the equation as a whole is accepted with probability $(1-\alpha)$.

In our case, the tabular value of the criterion with degrees of freedom $k_1=1$ and $k_2=5$, tabular $F = 6.61$, and calculated for the studied model $F = 96.95$.

Since the actual value of $F > F$ tabular, the coefficient of determination is statistically significant (the regression equation estimate found is statistically reliable), and the proposed model is significant.

To confirm the significance of the correlation coefficient of the constructed model, we use the Student's criterion. The formula which is used to calculate it:

$$t = \frac{|r|(n-2)^{1/2}}{(1-r^2)^{1/2}} (3)$$

The calculated value of the Student's t criterion is 9.8465, the table value is 2.571. Since the actual value of the Student's coefficient is larger than the tabular one, the linear correlation coefficient is considered significant, and the relationship between the volume of sales of goods in the online store and the number of goods that are presented with the help of photos and videos made with the use of rotary tables for photography is real.

Based on the data in Table 1, using formula 1, we write down the system of equations, as a result of solving which, we get the regression coefficients: $b = 16,6190$, $a = -12389,58$. By substituting them into the equation of linear dependence, we get a regression equation which looks as following:

$$y = 16,6190x_1 - 12389,58 (4)$$

where y is the volume of sales of goods;

x_1 – the number of products that are presented with the help of photos and videos made using rotary tables for photography.

The value of the coefficient of determination R^2 for the constructed model reaches the level of 0.7885, which means that 78.85% of the variations in the sales of goods are linearly related to the factor included in the model.

So, if the number of goods presented with the help of photos and videos made using turntables for photography is increased by one hundred hryvnias, then the volume of sales of goods will increase by 1,661.9 thousand hryvnias. It should be clarified that in this research, the change in sales volumes was considered exclusively in conjunction with the use of rotary photo tables for presenting goods to consumers. The other factors that may have an impact on changes in the resulting indicator in this were not taken into account but this may become the direction of the further research.

4. CONCLUSION

The conducted research showed the importance of using rotary photo tables for presenting goods in an online store. In the course of the research, it was proved that there is a close relationship between the volume of sales of goods by the online store and the number of goods that were presented with the help of photos and videos made with the use of rotary tables for photography. It was also established that the correlation between the volume of sales of goods by the online store and the number of goods presented without the use of rotary photo tables is weak. This trend is quite clear and is explained by the fact that consumers of goods from online stores make a choice in favor of goods that are most widely and comprehensively presented on the website of the online store. It is the use of rotary tables for product photography that makes it possible to present the product from all possible angles and demonstrate all the features and advantages of the offered product to the potential consumer. Paired linear regression was chosen in the research. Its parameters were estimated by the method of least squares. The statistical significance of the equation was checked using the coefficient of determination and Fisher's criterion. It was established that 78,85% of the total variability of Y in the model is explained by the change in x. It was also established that the parameters of the model are statistically significant. The economic interpretation of the parameters of the model shows that the increase in the number of goods presented with the help of photos and videos created with the use of rotary tables for photography will increase by one hundred units, then the volume of sales of goods will increase by 1661,9 thousand hryvnias.

REFERENCES

1. Akoglu H. User's guide to correlation coefficients. Turkish journal of emergency medicine. 2018; 18(3): 91–93.
2. Amankwah-Amoah J., Khan Z., Wood G. COVID-19 and business failures: The paradoxes of experience, scale, and scope for theory and practice. European Management Journal. 2021; 39(2): 179–184.
3. Bassey E., Akra U. Canonical Correlation of Multivariate Regression Analysis on Economic Factors in Nigeria. Asian Journal of Probability and Statistics. 2021; 4: 1-7.
4. Cumming D., Johan S., Khan Z., Meyer M. E-Commerce Policy and International Business. Manag Int Rev. 2022; 2: 1-23.
5. Cong H., Hu X., Du Y. Correlation Analysis and Regression Fitting of Multiple Sulfide Aging Characteristic Parameters. IEEE Transactions on Dielectrics and Electrical Insulation. 2022; 2: 1-17.
6. Dos Santos E., Gomes J. Interactions between e-commerce users during the COVID-19 pandemic period: What came and what remained. Work. 2022; 73(s1): 177-187.
7. Duan L. Correlation Analysis between Atmospheric Environment and Public Sentiment Based on Multiple Regression Model. Wireless Communications and Mobile Computing. 2022; 3: 1-8.
8. Grechyshkin Y. Mathematical modeling of the development of the potential of integrated business structures. Management. 2021; 33: 141-148.
9. Jobsta L., Heck D, Moshagena M. A comparison of correlation and regression approaches for multinomial processing tree models. Journal of Mathematical Psychology. 2020; 98:123-131.

10. Homapour E., Su L., Caraffini F., Chiclana, F. Regression Analysis of Macroeconomic Conditions and Capital Structures of Publicly Listed British Firms. *Mathematics*. 2022; 10: 11-19.
11. Hwang S., Kim J. Correlation Between Tractor Variables and Loan Support Limit in South Korea Through Regression Analysis. *Journal of Biosystems Engineering*. 2022; 47: 101-107.
12. Khatti J., Grover K. A Study of Relationship among Correlation Coefficient, Performance, and Overfitting using Regression Analysis. *International Journal of Scientific and Engineering Research*. 2022; 13: 1074-1085.
13. Mellos E., Paparrigopoulos T. Substance use during the COVID-19 pandemic: What is really happening? *Psychiatriki*. 2022; 33(1): 17-20.
14. Mohammed M., Feroskhan M., Sreekanth K., Karunamurthy R., Shivakumar S., Nazaruddin T., Yunus K. Regression-Analysis-Based Empirical Correlations to Design Regenerative Flow Machines. *Energies*. 2022; 15: 1-23.
15. Yanqing Y. Model-free tests for series correlation in multivariate linear regression. *Journal of Statistical Planning and Inference*. 2020; 206: 179-195.
16. Papageorgiou S. On correlation coefficients and their interpretation. *J Orthod*. 2022; 49(3): 359-361.
17. Schober P., Boer C., Schwarte LA. Correlation Coefficients: Appropriate Use and Interpretation. *Anesth Analg*. 2018; 126(5): 1763-1768.
18. Tabatabai M., Bailey S., Bursac Z., Tabatabai H., Wilus D. An introduction to new robust linear and monotonic correlation coefficients. *BMC Bioinformatics*. 2021; 22(1): 170-186.
19. Tzeremes, P. Productivity, efficiency and firm's market value: Microeconomic evidence from multinational corporations. *Bull. Appl. Econ*. 2020; 7: 95-101.
20. Vetter T., Mascha E. Unadjusted Bivariate Two-Group Comparisons: When Simpler is Better. *Anesth Analg*. 2018; 126(1): 338-342.