sciendo

Financial Internet Quarterly 2021, vol. 17 / no. 1, p. 62-69

DOI: 10.2478/fiqf-2021-0007

Financial Internet Quarterly

COMPREHENSIVE ANALYSIS OF A COMPANY'S ACTIVITY BY MEANS OF STATISTICAL MODELING AS SUPPORT FOR ITS DECISION-MAKING SYSTEM

Janusz Soboń[1], Natalia Burkina[2], Kostiantyn Sapun[3], Ruslana Seleznova[4]

Abstract

An important role in ensuring effective forms of management and increasing competitiveness is played by the process of forecasting the activity of the enterprise. This work analyzed the performance of a food industry enterprise, for which a wide range of statistical methods were applied such as methods of cluster, correlational and regression analysis, statistical tests of Fisher, Student, Farrar-Glauber, Durbin-Watson, Goldfeld-Quandt, μ-criterion, multifactor regression, trend, auto-regression models, and models of seasonal fluctuations, which provided a view of the economic properties of the enterprise profit process, in particular the auto-regression component of revenue dependence on its value last year, seasonal quarterly dependence on sales and marketing costs, product price, etc. The detected patterns will allow us to take into account these features for forecasting future revenues and for adjusting the enterprise's decision-making system taking into account seasonal features and results of the previous year.

JEL classification: A1, C1, C5, C6, C8

Keywords: multifactor regression, forecasting revenue, correlational and regression analysis

Received: 14.01.2021 Accepted: 20.03.2021

[1] AJP w Gorzowie Wielkopolskim, Poland, ORCID: 0000-0001-9161-6967.
[2] Donetsk National University after V. Stus, Ukraine, ORCID: 0000-0003-2993-5128.
[3] Varna Free University, Bulgaria, ORCID: 0000-0002-5966-5125.
[4] Taras Shevchenko Kiev National University, Ukraine, ORCID: https://orcid.org/0000-0003-3080-773X.

Introduction

The confectionery market of Ukraine is a dynamic market, where there are constant fluctuations in the volume of production, revenues and profits of enterprises, as well as every year there is increased competition (today in this area there are 5,459 companies: 6 large, 274 medium and 5,179 small firms). The events of 2014 significantly weakened the positions of large confectionery companies located in the temporarily occupied territories, in particular PJSC AVK and JSC PU "Konti", so these companies were forced to return their leadership positions. An important role in ensuring effective forms of management and increasing competitiveness is played by the process of forecasting the activity of the enterprise, which is responsible for the selection, evaluation and interpretation of financial, economic and other data affecting the process of making important decisions.

Literature review

The main issues of forecasting the activity of enterprises are highlighted in scientific works by such scientists as Drobysheva[5], Zakabluk[6], Sidorova and Burkina[7], Sapun, Selezneva and Soboń[8], Rogozińska-Mitrut and Soboń[9] and others. However, as practice shows, in rapidly changing market situations not enough attention is paid to the forecasting process on the part of managers of enterprises, which negatively affects the final results of their activities.

[5] Drobysheva, O.O., Kernychyshyn,Yu.V. *Forecasting the Development of Industrial Enterprise in Modern Conditions [Prognozuvannya rozvytku promyslovogo pidpryyemstva v suchasnyh umovah].* Retrieved from: http://www.zgia.zp.ua-/ga-zeta/evzdia_5_158.pdf.

[6] Zakabluk, G.O. (2018). Prognostication of Income and Expenses of the Machine-Building Enterprise on the Basis of the Holt-Winters Method. ЕКОНОМІЧНА НАУКА, No. 6, 51-55. Retrieved from:http://www.economy-.in.-ua/pdf/6_2018/13.pdf.

[7] Sidorova, A.V., Burkina, N.V. (2019). Analysis of the Impact of Socio-economic Factors on the Financial Results of Small and Medium-sized Enterprises. Collection of Scientific Works of Professor and Teaching Staff of VasylStusDonNU,[S.l.], 27-29, July 2019. Retrieved from: http://jpvs.donnu.edu.ua/article/view/7089.

[8] Soboń, J, Sapun, K., Bezpalchuk, O., Seleznova, R. (2019). Perspectives of Management Development of International Tourist Business in Ukraine. *Proceedings of Fourth International Scientific Conference Telecommunications, Informatics, Energy and Management*; Greece; TIEM 2019; Kavala.

[9] Soboń J., Rogozińska-Mitrut, J. (2005). Znaczenie bezpośrednich inwestycji zagranicznych w rozwoju regionalnym. Zeszyty naukowe z zagadnień jakości w nowoczesnej gospodarce, No. 15, 87-105.

company's activities and forecast the indicators of its activity to support the enterprise decision-making system.

Roshen Confectionery Corporation is one of the world's largest confectionery manufacturers. ROSHEN produces more than 320 high-quality confectionery products. The corporation's range includes chocolates and jelly candies, caramels, chocolate bars and cookies, waffles, biscuit rolls and cakes. Some of them have no analogues in the Ukrainian market. The total output reaches about 300 thousand tons per year. Roshen confectionery products are made using the most modern technologies. The corporation consists of Ukrainian factories (Kyiv, Kremenchug, Boryspilska and two production sites in Vinnitsa), Klaipeda Confectionery Factory (Lithuania), Bonbonetti Choco Kft (Budapest, Hungary), as well as Vinnytsia dairy plant, which provides factories with natural high-quality raw dairy materials.

Thus, this company is a bright representative of the confectionery industry, which led to the research on its example. For the analysis, the data of PrJSC "VMZ "Roshen" for 2014-2018 were used on the main types of goods sold by the enterprise:

1) milk fat dehydrated (t 32°),

2) butter “Extra” 82.5% fat,

3) the lactosecent product condensed with vegetable fat and sugars "Moles” 8,5% fat,

4) skimmed milk powder,

5) whole milk powder 26% fat,

6) other.

The first step of the study was to identify uniformity on the set of indicators of the enterprise's activity by the main types of products.

Indicators chosen for the analysis are:

Y - revenue from sales of products, thousand UAH,

X_1 - sales costs, thousand UAH,

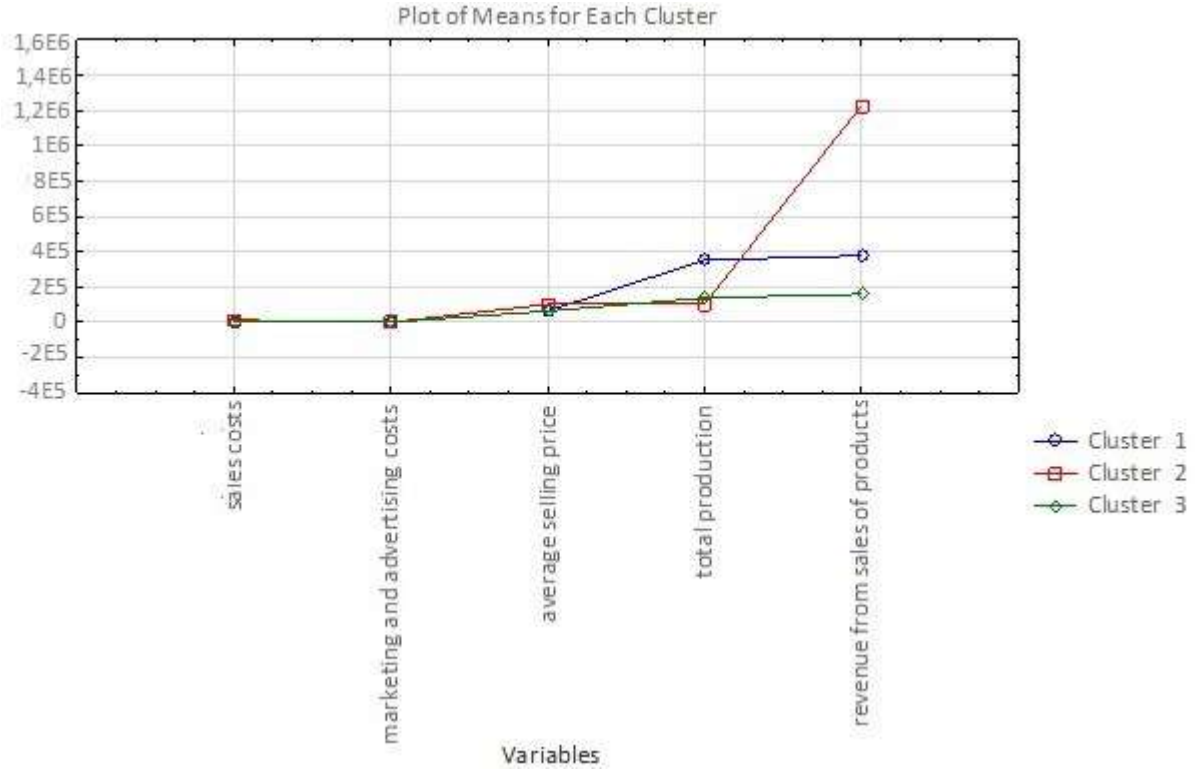
X_2 - marketing and advertising costs, thousand UAH,

X_3 - average selling price, UAH/ton,

X_4 - total production, UAH.

The first stage of the study with the help of the Statistica Enterprise 10.0 program was implemented clustering on the selected features by the k-means method of 3 clusters, which is reflected in Figure 1.

Figure 1: Clustering of the main types of goods of the enterprise, left to right: X_1; X_2; X_3; X_4; Y



Source: Author's elaboration

The graph shows that in clustering on the result past revenue, clusters differ significantly, so the purpose of clustering is realized. Clusters do not differ significantly in factors for sales costs, marketing and price, but significantly differ in the volume of total production and revenue.

The second cluster draws the attention with the greatest value of revenue at the lowest price and low outputs unlike other clusters. This is single cluster (C_3) Milk fat for 2016. This product differs from the other largest value of revenue at slightly higher sales costs and average marketing costs, negligible price and small volume of production. That is, it was in 2016 that dairy fat was produced and sold eficiently, which led to increased demand and as a consequence with savings on its marketing, but taking into account the low price and low volume, it is unlikely that the company was able to get an inflated profit on this product. That is why the data on this position raise doubts, and were attributed to "emissions" and therefore excluded from further analysis. For all other goods there is a greater homogeneity, which proves a relatively unchanging strategy for the production and sale of these products.

The first cluster includes 15 observations. It has average values in terms of revenues and the largest in terms of manufactured products. The largest amount of money is spent on the promotion of goods in this cluster. The third cluster consists of 14 observations. It has the lowest revenue value, average marketing costs, and sales. Thus, cluster analysis allowed us to polish the inputs, to detect their relative homogeneity for further analysis.

The study of the impact of all indicators on revenue was carried out by correlational and regression analysis. The correlation matrix, given in Table 1, shows the impact of all selected factors on the resulting revenue from the sale of products.

The construction of a multifactory regression model took place by step-by-step exclusion. Four factor linear model $\hat{y} = \hat{a}_0 + \sum_{j=1}^{4} \hat{a}_j x_j$ is built on all factors selected:

$\hat{y}$ = 78885.91 + 14.26 x_1 + 295.74 x_2 + 5.43 x_3 - 0.28 x_4
(R^2 = 0.984 R^2_n = 0.979).

Table 1: Correlation matrix of factors of the enterprise

	Sales costs, thousand UAH	Marketing and advertising costs, thousand UAH	Average selling price, UAH/ton	Total production, thousand UAH	Revenue from sales of products, thousand UAH
Sales costs, thousand UAH	1				
Marketing and advertising costs, thousand UAH	0.001949359	1			
Averagesales price, UAH/ton	0.654027305	0.217208649	1		
Volume of manufactured products, thousand UAH	0.398120650	0.907875708	0.488470552	1	
Revenue from sales of products, thousand UAH	0.636067789	0.496587375	0.941233385	0.723728068	1

Source: Author's elaboration

When checking the significance of the model, it is found that:

F_{fact} = 17,8, F_{table} (α = 0.05, v_1 = 4, v_2 = 15) = 3.0556.

Since $F_{fact}>F_{table}$ (17.8>3.0556), $P(F<F_{fact}) \leq \alpha$: $2.8610\times^{-13}$ <0.05,

the resulting model is statistically significant with a reliability of at least 95%. However, there are no significant coefficients for variables $X_1$and X_4, the error of which is greater than the specified valueα = 0.05. In the process of step-by-step selection of factors, a two-factor linear model $\hat{y} = \hat{a}_0 + \hat{a}_2x_2 + \hat{a}_3x_3$ was obtained, which is significant in all criteria with a relative approximation error of 5.09% in the form of:

$\hat{y}$ = 40875.58 + 161.1 x_2 + 5.85 x_3 (R^2 = 0.975 R^2_n= 0.973).

The resulting model allowed us to find that while increasing the cost of marketing and advertising for each thousand UAH (subject to the immutability of other factors) revenue from sales of products increases by an average of 161.1 thousand UAH. With the increase in the average selling price of 1 UAH/t (subject to the immutability of other factors), revenue from the sale of products will increase by an average of 5.85 thousand UAH.

To verify the results of the model, we will conduct a study of the prerequisites used by the method of the least squares, which is the basis for obtaining the coefficients of the model.

Verification of multicollinearity between the two factor arguments was carried out according to the Farrar-Glauber test, which allowed us to draw the following conclusions:

since $\chi^2_{fact}<\chi^2_{table}$ (0.846<3.841), then there is no multicollinearity in the X_2, X_3 variable array.

Since condition $F_{fact}<F_{table}$ (F_{table} = 3.592) is performed for statistics F_2 and F_3, then the conclusion (with a probability of 95%) about the statistical non-significance of the multiple correlation coefficients of X_2, X_3 and the absence of close linear dependence of each factor with the other is confirmed.

Since statistics t_{23} are met by the condition $t_{fact}<t_{table}$ (t_{table} = 2.1), there is no close linear dependency between X_2 and X_3.

Thus, the Farrar-Glauber test proved the absence of multicollinearity between the selected factors.

To check the fulfillment of the conditions of homoskedasticity for the two-factor model $\hat{y}$ = *40875.58 + 161.1* x_2 *+ 5.85* x_3 was applied the μ criterion that proved that the constructed model has heteroscedastic remnants, because $\mu>\chi^2_{table}$ (8.59> 7,81).

To determine which of the variables is the cause of heteroskedasticity, a check was made for each variable using the Goldfeld-Quandt test, which concluded (with 95% reliability) that variable X_2 is not the cause of heteroskedasticity of the model residues, as $F^{(2)}_{fact} < F_{table}$ (0.047 < 5.05) and $F^{(3)}_{fact} < F_{table}$ (0.007 < 5.05). So, variable X_3 is not the cause of heteroscedacticity of the remnants of the model. Thus, none of the indicators causes heteroskedasticity alone.

Due to the fact that the current trend is influenced by the latest values of the indicator, it was decided to consider the reduced sample. After removing the four first observations y_1, y_2, y_3, y_4 to the μ criterion the hypothesis of heteroscedasticity was rejected: $\mu < \chi^2_{table}$ (3.33 < 7.81).

$$\tilde{y} = 36527.50 + 161.65\, x_2 + 5.91\, x_3$$

High determination ratio R^2 = 0.95, Fisher's criterion F = 121.3406, and low average relative error of approximation ξ = 4.94%, the presence of homoskedasticity and lack of multicollinearity characterize the sufficiently high quality of the model.

Another important aspect that often arises in economic research and impairs the quality of regression models is the presence of autocorrelation. To test the hypothesis of the presence of autocorrelation of the constructed model, the Durbin-Watson test was applied.The actual value of the criterion is close to zero, so it was concluded that there is autocorrelation. To improve the quality of the model, minding the detected autocorrelation, another factor y_{t-1} was introduced into the model, in result the model took the form:

$$y = 23233.66 + 90.33\, x_2 + 4.18\, x_3 + 0.32\, y_{t-1}$$

The resulting model has even higher quality indicators. In particular, the coefficient of determination R2 = 0.974 became higher and proves that the constructed model explains the variance of the effective characteristic Y by 97.4%, and only 2.6% is due to factors not accounted in the model. At the significance level α = 0.05 all coefficients of the model are significant, and the whole model is also significant according to Fisher's criterion: $F_{fact} > F_{table}$ (150.84 > 3.49). The approximation error is only ξ = 3.27%, which demonstrates an even higher quality of the model.

The results of the models constructed are shown in Table 2.

Table 2: Comparison of econometric models

#	Model	n	R^2	R^2_{norm}	ξ
1	$\hat{y} = 40875.58 + 161.1\, x_2 + 5.85\, x_3$	20	0.975	0.973	5.09%.
2	$\tilde{y} = 36527.50 + 161.65\, x_2 + 5.91\, x_3$	16	0.949	0.941	4.94%
3	$y = 23233.66 + 90.33\, x_2 + 4.18 X_3 + 0.32\, y_{t-1}$	16	0.974	0.968	3.27

Source: Author's elaboration

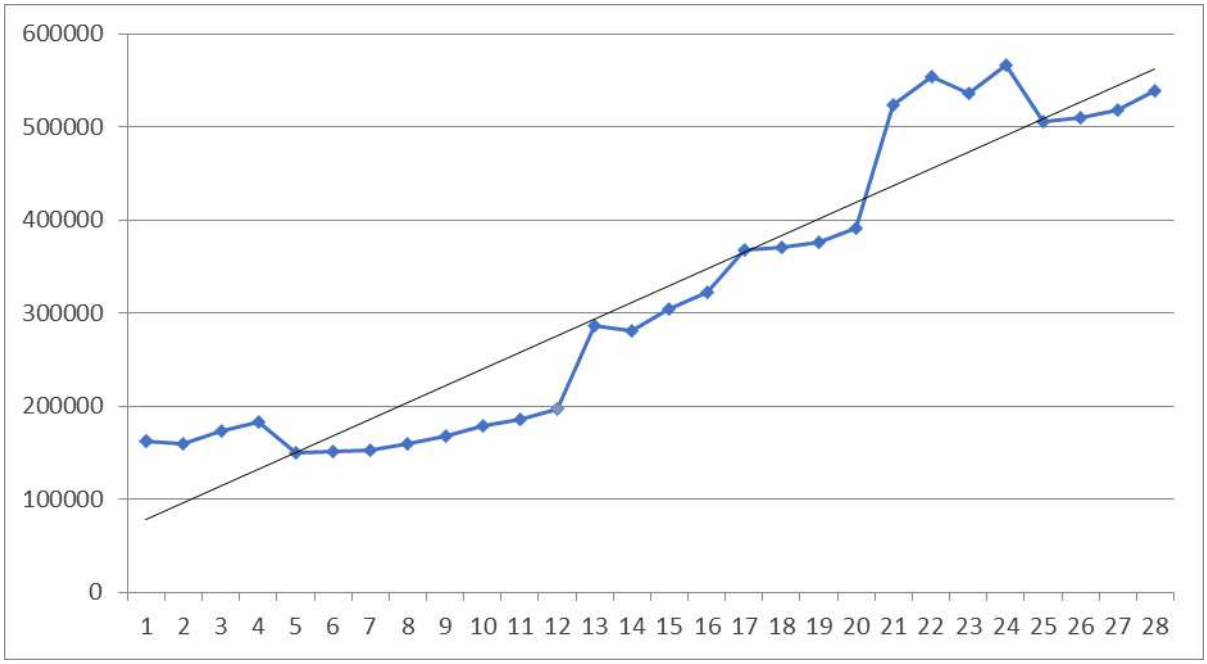
The latest model is the best, so it was used for forecasting. As a result, the forecast values of the volume of revenues from sales

y_{17}= 492369.2779 thousand UAH, and
y_{18}= 478112.7284 thousand UAH

were obtained with an average fluctuation of about 6%.

The next step of the study was to find out the seasonal properties of the sale of products, thousand UAH. To do this, the time interval was extended and the data was chosen quarterly. 28 quarters for 2012-2018 were analyzed, according to which the graphic depicted in Fig. 2 was constructed and demonstrates a clear trend toward revenue growth.

Figure 2: Quarterly revenues from sales of products of PJSC "VMZ "Roshen" for 2012-2018



Source: Author's elaboration

The graph clearly demonstrates the presence of long-term cyclical fluctuations in about 12-13 quarters (3 years). The test of the hypothesis was carried out using the autocorrelation coefficient. To calculate the coefficient of autocorrelation deviations from the trend, a trend model was built in the form of a polynomial of a third order. Based on the value of the coefficient of determination R^2 = 0.964, it is concluded that the constructed model explains the variance of the resultant trait *Y* by 96.4%, and only 3.6% is the share of factors not accounted in the model.

Since $F_{fact}>F_{table}$ (214.98>3.009), the resulting equation is statistically significant. The average error of the approximation ξ = 7.71%, which shows sufficient accuracy, that is, of course, less than in the previous model, but much better, given that the sample was expanded and reviewed quarterly. Therefore, the constructed model can be used for economic analysis.

Since there are few local extremums in a number of deviations, the logical thing is the presence of cyclical fluctuations.The coefficient of autocorrelation of deviations with a lag in one year is 0.37, which is more than + 0.3, so we can assume that in the general fluctuations of the levels of a number of dynamics there is a significant cyclic component.

To find the cycle length, the coefficients of autocorrelation of deviations with a lag of 1, 2, 3, 4, etc. were successively calculated. The largest in absolute value autocorrelation coefficient determines the length of the cycle in 12 quarters.

The calculated indicators of the absolute oscillation force provided the following results. The amplitude of fluctuations between the largest and smallest in absolute magnitude deviation from the trend is 64,406.18711 thousand UAH. The average linear deviation is 23,033.82, that is, the specific values of revenue from the sale of time series products deviate from the trend by an average of 23,033.82 thousand UAH. The average standard deviation shows that the levels of the time series are in a wide range of values.

Seasonality indices were calculated according to two methods – by average without the trend and on trend indicators that characterize the existing trend. According to the first method, the peak falls in the fourth quarter, and the second method gives the first quarter, and the smallest value in the first method is for the first quarter, the second method gives the third quarter, which once again confirms the importance of taking into account a sufficiently stable trend toward growth.

Taking into account the fact that the existing seasonal fluctuations do not have a clear size shape, seasonality indices were calculated using the algorithm of adjustment of parameters for asymmetric seasonal fluctuations taking into account the trend.

The trend equation for a conditional time looks like: $\hat{y}_i = 320643.96 + 18106.6 * t_i$.

When calculating the parameters, the peaks are sharply allocated for the fourth quarter, for t_i on average it will be zero weight. The minimum levels of the first quarters fall on the values of t_i, which on average give a negative weight. The values of the levels of the 2nd quarter are lower than the average for the year, their weight t_i is on average zero. The positive weight of "emissions" falls on the 3rd quarter. To adjust the calculation, it is necessary to "remove" from the numerator of the parameter *b* the specified inequality, i.e. the excess of positive deviations from the trend in the third quarters over the negative values of deviations in other quarters. The adjusted equation of the trend obtained has the form:

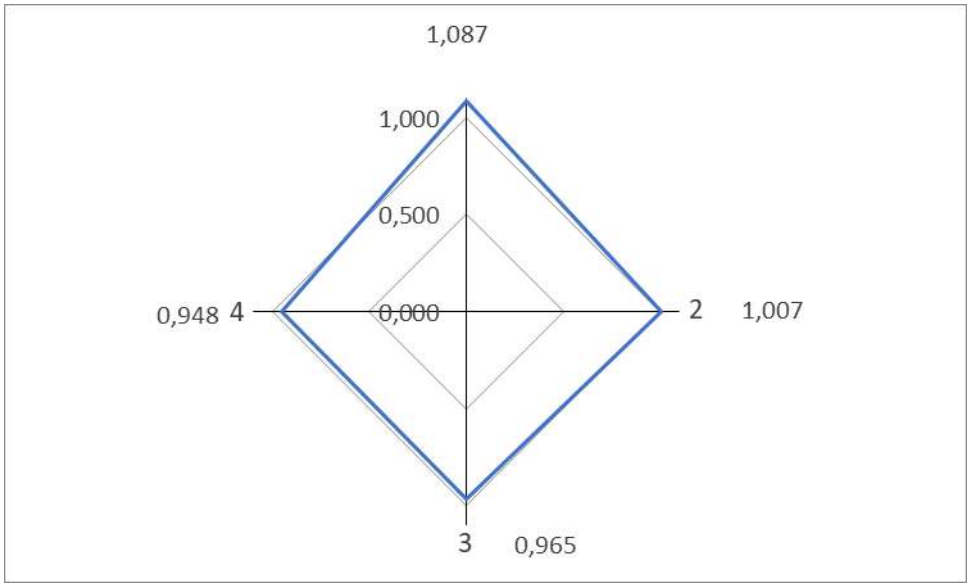
$$\hat{y} = 320643.96 + 17951.84 * t_i$$

Thus, the exaggeration of the average increase in levels of the series per quarter due to the asymmetric distribution of seasonal peaks was 0.86%.

Seasonality indices are not much different. With the adjusted trend, the value of seasonal fluctuations decreased by less than 1%, which once again emphasizes the significance of the existing fluctuations, which should be taken into account in the design of the company's further activities, although they are not strongly manifested, but present.

Graphical indices of seasonality are presented in Figure 3.

Figure 3: Schedule of seasonal fluctuations in sales revenues



Source: Author's elaboration

According to the results of clustering, a series of models was also built, but clustering on the one hand increases the uniformity of input data, and on the other hand reduces the volume of sampling. In this case, the sample reduction was critical and the resulting models on individual clusters had poorer quality than the module obtained throughout the database, which once again emphasizes the sufficient uniformity of input data and a sufficiently stable decision-making system implemented at the enterprise, according to which the production and sales strategy does not depend heavily on the time and type of products.

Conclusion

Thus, the work analyzed the performance of the food industry enterprise, for which a wide range of statistical methods were applied such as methods of

cluster, correlational and regression analysis, statistical tests of Fisher, Student, Farrar-Glauber, Durbin-Watson, Goldfeld-Quandt, μ-criterion, multifactor regression, trend, auto-regression models, models of seasonal fluctuations, which allowed us to reveal the economic properties of the enterprise profit process, in particular the auto-regression component of revenue dependence on its value last year, seasonal quarterly dependence on sales and marketing costs, product price, etc. The detected patterns will allow us to take into account these features for forecasting revenues for the coming years and adjust the system of acceptance of decisions at the enterprise taking into account seasonal features and results of the previous year.

References

Drobysheva, O.O., Kernychyshyn,Yu.V. *Forecasting the Development of Industrial Enterprise in Modern Conditions [Prognozuvannya rozvytku promyslovogo pidpryyemstva v suchasnyh umovah].* Retrieved from: http://www.zgia.zp.ua-/ga-zeta/evzdia_5_158.pdf.

Soboń, J, Sapun, K., Bezpalchuk, O., Seleznova, R. (2019). Perspectives of Management Development of International Tourist Business in Ukraine. *Proceedings of Fourth International Scientific Conference Telecommunications, Informatics, Energy and Management*; Greece; TIEM 2019; Kavala.

Soboń J., Rogozińska-Mitrut, J. (2005). Znaczenie bezpośrednich inwestycji zagranicznych w rozwoju regionalnym. *Zeszyty naukowe z zagadnień jakości w nowoczesnej gospodarce*, No. 15, 87-105.

Zakabluk, G.O. (2018). Prognostication of Income and Expenses of the Machine-Building Enterprise on the Basis of the Holt-Winters Method. ЕКОНОМІЧНА НАУКА, No. 6, 51-55. Retrieved from:http://www.economy-.in.-ua/pdf/6_2018/13.pdf.

Sidorova, A.V., Burkina, N.V. (2019). Analysis of the Impact of Socio-economic Factors on the Financial Results of Small and Medium-sized Enterprises. Collection of Scientific Works of Professor and Teaching Staff of VasylStusDonNU,[S.l.], 27-29, July 2019. Retrieved from: http://jpvs.donnu.edu.ua/article/view/7089.