DOI 10.31489/2022Ec4/112-124
JELH26
УДК336.228.34

T.K. Shuren[*1], T.P. Pritvorova[1], E.A. Vechkinzova[2], A. Kizimbayeva[3]

[1]Karaganda University of the name of academician E.A. Buketov, Kazakhstan;
[2]Institute of Management Problems of the Russian Academy of Sciences named after V.A. Trapeznikov, Moscow, Russia
[3]Caspian University of Technology and Engineering named after Sh.Yessenov;

[1]itoktar@gmail.com, [2]pritvorova_@mail.ru,
[3]ea_vechkinzova@guu.ru, [4]kizimbaeva@mail.ru

Digital transformation and the relationship with economic growth

Abstract

Object: The main purpose of this article is to identify and assess the impact of digital transformation indicators on economic growth in Kazakhstan.

Methods: For this study, we used methods of statistical multiple correlation and regression analysis based on the software package "Data Analysis" offered by MS Excel. We used data from the official website of the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK at stat.gov.kz.

Findings: The primary selection of statistical indicators was carried out and a group of factors (and corresponding indicators) was determined hypothetically influencing economic growth for the period 2007-2020. There are 14 factors that have a significant impact on gross value added. Based on the selection of the most significant factors, a regression equation is constructed that demonstrates the degree of influence on the resulting GVA. The obtained regression model was evaluated. The found regression equation is significant according to the Fisher criterion, all its parameters, including the free term, are significant according to the Student's criterion with a maximum error of 0.07. The multiple correlation coefficient is 0.99. The obtained results can be useful in planning GDP and GVA, both at the regional and national level.

Conclusions: In the system of gross value added indicators, an important place is occupied by the indicators of the number of organizations using the Internet, the unemployment rate and computer literacy of the population. The analysis demonstrates a strong relationship between these indicators. As a result, we saw that the relationship between these indicators can be explained by a linear equation with an average accuracy of 97%. At the same time, for a more adequate analysis of the situation, it is also necessary to take into account the inverse relationship between changes in unemployment rates in the Republic of Kazakhstan and added gross value. The negative correlation between these indicators confirms the vulnerability and instability of the economy from changes in the unemployment rate.

Keywords: digital economy, economic growth, the impact of digitalization, multiple regression, correlation.

Introduction

The article discusses the indicators of the digital transformation of the economy and business in Kazakhstan and its potential impact on the economic growth of the country.

The digital revolution is in full swing and gaining momentum. The established indicators and assessment tools cannot keep up with the rapid pace of digital transformation (OECD, 2019a). OECD (2019b), which reveals many gaps in the existing system of measuring digital transformation and reports from international organizations suggest new indicators and recommend improving the international comparability of currently used ones.

The article examines such statistical data to assess the development of the digital economy, starting with the number of large and medium-sized enterprises using digital technologies, the number of Internet and computer users, the total costs of information and communication technologies, the main goals of using the Internet by household members, indicators of the use of information and communication technologies in organizations, indicators, characterizing the development of E-commerce in the Republic of Kazakhstan, digital literacy of the population, export and import of goods related to information and communication technologies.

The research questions that will be answered in this article are as follows:

(1) What indicators can be used to calculate the aggregated indicator of Kazakhstan's digital transformation?

[*] Corresponding author. E-mail address: itoktar@gmail.com

(2) Can these values be used to predict the GVA indicator for Kazakhstan?

To answer these research questions, we use data from the reports of the Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan for 2007-2020.

Literature Review

Digitalization of the economy opens up huge opportunities in such areas as economics, innovation, education, healthcare, management and lifestyle (Mühleisen, 2018). Less than 1 percent of technologically processed information worldwide was in digital format in the late 1980s, and more than 99 percent by 2012 (Hilbert, 2020). Moreover, every 2.5-3 years, humanity can accumulate more knowledge than before the birth of civilization (United Nations Conference on Trade and Development (UNCTAD), 2019).

A lot of work has been written about the impact of digital transformation on the economy, but, as far as we know, no attempt has yet been verified to measure the relationship of the above indicators with the real growth of the economy in Kazakhstan.

We want to fill this gap by trying to assess the digital transformation in Kazakhstan by several criteria and exploring its relationship with economic growth. The article discusses the question of which indicators to use to assess the level of digitalization of the economy and business. We also use economic data analysis models to check whether changes in these indicators affect economic growth in Kazakhstan, in particular Gross Value Added (GVA).

The novelty of the article is the assessment of the impact of digitalization on economic growth and the construction of a mathematical equation of this connection. The model we propose can become one of the tools for forecasting the GVA for future years. We point out that digitalization contains the potential for economic growth, and innovation supports the sustainability of the economy.

In the economy, the digital revolution began on a large scale at the end of the 20th century, when the Internet was introduced into economic use. The positive effects of the digital economy can be seen on countless fronts. Molinari & Torres write that, first of all, digitalization supports economic growth, but the power of influence depends on the research methodology used in the study and the geographical configuration (Molinari, Torres, 2018; Solomon, van Klyton, 2020). Other researchers have found that this also greatly changes the structure of the labor market, reducing the demand for routine work and low-skilled workers (Peetz, 2019). In addition, digitization is transforming the way businesses work and interact with their customers and suppliers. This has a significant impact on improving the efficiency of business operations (Ritter, Pedersen, 2020). In addition, it is strongly recommended to better adapt existing statistical systems to the rapid changes caused by digital

In 2018, in order to study the impact of digitalization on the economy, the ECB conducted a special survey of leading companies in the eurozone (Elding, C., Morris, R., 2018), the main purpose of which was also to measure how digital transformation affects macroeconomic aggregates. This study examined 74 leading non-financial companies. In our article, we studied not individual firms, but the economy as a whole. As a measure of economic growth, we chose added gross value, since it shows the economic well-being of the population, including all primary incomes.

According to the survey results, the overwhelming majority of respondents felt that digitalization has a positive impact on their company's sales. More than half expect that the introduction of digital technologies will lead to a “slight increase” in sales over the next three years, while about a third expect a “significant increase”. To some extent, this positive opinion may reflect the relative size.

Methods

The research methods are multiple correlation analysis and regression analysis.

As independent variables of the model, statistical data used to assess the development of the digital economy by international organizations are taken, starting with the number of large and medium-sized enterprises using digital technologies, the number of Internet and computer users, the total costs of information and communication technologies, the main goals of using the Internet by household members, indicators of the use of information and communication technologies in organizations, indicators characterizing the development of E-commerce in the Republic of Kazakhstan, digital literacy of the population, export and import of goods related to information and communication technologies, investments in fixed assets and the unemployment rate.

Results

The research work consisted of the following stages:

- To determine the indicators that are indicators of the digital transformation of the economy of Kazakhstan and search for statistical data on the above indicators for the period from 2007 to 2020, also select indicators of GVA, exports and imports for these periods.
- To process statistical data first and then filter them out.
- To identify the relationships between variables and evaluate the strength of this relationship. Understand how they affect each other and determine how strong.
- To select the variables x that we will use for the equation of the GVA calculation model. It's not necessary that all variables will remain. It is possible that 2-3 or even 1 indicator will remain in the calculation.
- To get a regression model and evaluate how it corresponds to the data that we have.

There are many software products for analyzing statistical data. For calculations, the authors used the built-in MS Excel tools, as well as an additional add-in “Data Analysis”, where there are many different statistical tools.

To begin with, in Table 1, we collected statistical data on the GVA and 14 other indicators that could hypothetically be related to the digitalization of the economy and business in one way or another, and determined which of the indicators are independent (arguments) and which dependent (function).

Independent variables:

X1 - The number of organizations using the Internet units network;

X2 - Labor productivity Index;

X3 - Computer literacy at the age of 6 years and older: Novice user (was changed to a quantitative indicator taking into account the population for the corresponding periods), units;

X4 - Computer literacy at the age of 6 years and older: An ordinary user (was changed to a quantitative indicator taking into account the population for the corresponding periods), units;

X5 - Computer literacy at the age of 6 years and older: Experienced user (was changed to a quantitative indicator taking into account the population for the corresponding periods), units;

X6 - The level of innovative activity of enterprises in all types of innovations;

X7 - Innovative products and services produced in 1 year, units;

X8 - The volume of manufactured industrial products (goods, services) in the field of information and communication technologies (in current prices of enterprises), million tenge;

X9 - Indicators of the global competitiveness index by the factor “Level of technological development”;

X10 - Exports, million US dollars;

X11 - Imports, million US dollars;

X12 - Investments in fixed assets, million tenge;

X13 - Number of employed people;

X14 - Unemployment rate, as a percentage.

Dependent is the value measured in connection with changes in independent values.

In this case, the GVA (million tenge, signed as Y) is considered depending on changes in other indicators.

Table 1. Statistical data on GVA and 14 indicators from 2007 to 2020

Year	GVA million tenge	Number of organizations using the Internet	Labor Productivity Index	Computer literacy at the age of 6 years and older: Novice user	Computer literacy at the age of 6 years and older: Ordinary user	Computer literacy at the age of 6 years and older: Experienced user	The level of innovation activity of enterprises and organizations on technological innovations	Innovative products and services produced in 1 year	The volume of manufactured industrial products (goods, services) in the field of information and communication technologies (in current prices of enterprises) million tenge.	Indicators of the global competitiveness index by the factor "Level of technological development"	Exports, million US dollars	Import million US dollars	Investments in fixed assets million tenge	The number of active people, units.	Unemployment rate
	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14
2007	12772498,2	6803	107,5	498456000	828180000	201240000	4,8	92145227	12461,596	81,3	3424,8326	11868,0825	3 392 122	14349960	7,3
2008	13056532,9	35089	100	356628000	943644000	250902000	4	11076514,2	10883,038	103	4292,41577	11218,929	4 210 878	14738520	6,6
2009	14506780,8	44046	100,2	296056000	1076421000	186644000	4	8333747,03	9733,085	106	4103,67878	10081,7276	4 585 298	15028060	6,6
2010	21115891,00	45354	103,7	269280000	1062432000	172992000	4,3	14180086,1	11428,677	119	4118,96403	11368,5429	4 653 528	15373440	5,80
2011	25741874,80	48064	105	264960000	1107864000	145728000	5,7	23724164,3	16099,222	152	4337,73992	10972,9484	5 010 231	15665760	5,40
2012	28528090,1	49853	102,5	292146000	1114856000	196443000	5,7	37838527,7	22851,831	180	5430,9083	14344,5498	5 473 161	15900130	5,30
2013	32896601,00	58456	105,1	298200000	1226880000	175512000	8	57958430,4	29638,6	178	5970,58406	14083,5205	6 072 687	16153920	5,20
2014	36651572,20	52630	104,6	299117000	1241422000	174629000	8,1	57926716	30168,6	165	7002,48434	13845,9469	6 591 482	16425500	5,00
2015	38783900,40	65186	100,6	456040000	1173426000	124534000	8,1	37613402,9	17493	150	6177,43215	10897,7382	7 024 709	16645460	5,10
2016	44337585,50	75779	100,2	373590000	983787000	97845000	9,3	44622592,5	22805	143	6084,52982	9846,94536	7 762 303	16900500	5,00
2017	51195859,30	79658	104,3	416724000	1003024000	90200000	9,6	84287229,6	21245	148	6504,8801	10082,6496	8 770 572	17156040	4,90
2018	57706553,30	100702	103,1	429580000	1043788000	124304000	10,6	106329603	22509,3	143	7319,91358	11981,3684	11 179 036	17384280	4,90
2019	64681604,80	105531	103,7	436836000	1093941000	131421000	11,3	111252202	23265	148	7745,29788	11462,1922	12 576 793	17621520	4,80
2020	66828235,10	110246	97,5	421875000	1164375000	148125000	11,5	171677151	22879	152	5032,03767	8096,35806	12 270 144	17831250	4,90

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.

The first thing we investigated was how homogeneous our data was for each variable. That is, possible outliers were excluded from the model, the values of which were very different from the rest. They could have arisen either due to human error (typo), or it was a unique case (crisis, pandemic, lockdowns, sanctions). Outliers were determined visually using the MS Excel tool - color scales (Table 2). If there are no outliers, that data is distributed more or less evenly. Otherwise, the outliers are very different in color with neighboring cells. The years 2007, 2015 and 2020 were chosen as such emissions, and it was decided to exclude them from the model. As a result, we received the primary filtered data.

Table 2. Visual definition of emissions

year	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14
2007	12772498,2	6803	107,5	4984560	8281800	2012400	4,8	92145227	12461,596	81,3	3424,8326	11868,0825	3 392 122	14349960	7,3
2008	13056532,9	35089	100	3566280	9436440	2509000	4	11076514,2	10883,038	103	4292,41577	11218,929	4 210 878	14738520	6,6
2009	14506780,8	44046	100,2	2960560	10764210	1866440	4	8333747,03	9733,085	106	4103,67878	10081,7276	4 585 298	15028060	6,6
2010	21115891,00	45354	103,7	2692800	10624300	1729920	4,3	14180086,1	11428,677	119	4118,96403	11368,5429	4 653 528	15373440	5,80
2011	25741874,80	48064	105	2649600	11078640	1457280	5,7	23724164,3	16099,222	152	4337,73992	10972,9484	5 010 231	15665760	5,40
2012	28528090,1	49853	102,5	2921460	11148560	1964430	5,7	37838527,7	22851,831	180	5430,9083	14344,5498	5 473 161	15900130	5,30
2013	32896601,00	58456	105,1	2982000	12268800	1755120	8	57958430,4	29638,6	178	5970,58406	14083,5205	6 072 687	16153920	5,20
2014	36651572,20	52630	104,6	2991170	12414220	1746290	8,1	57926716	30168,6	165	7002,48434	13845,9469	6 591 482	16425500	5,00
2015	38783900,40	65186	100,6	4560400	11734260	1245340	8,1	37613402,9	17493	150	6177,43215	10897,7382	7 024 709	16645460	5,10
2016	44337585,50	75779	100,2	3735900	9837870	978450	9,3	44622592,5	22805	143	6084,52982	9846,94536	7 762 303	16900500	5,00
2017	51195859,30	79658	104,3	4167240	10030240	902000	9,6	84287229,6	21245	148	6504,8801	10082,6496	8 770 572	17156040	4,90
2018	57706553,30	100702	103,1	4295800	10437880	1243040	10,6	106329603	22509,3	143	7319,91358	11981,3684	11 179 036	17384280	4,90
2019	64681604,80	105531	103,7	4368360	10939410	1314210	11,3	111252202	23265	148	7745,29788	11462,1922	12 576 793	17621520	4,80
2020	66828235,10	110246	97,5	4218750	11643750	1481250	11,5	171677151	22879	152	5032,03767	8096,35806	12 270 144	17831250	4,90

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.

Up to this point, the indicators were studied separately from each other. Next, we will look at their connections with each other. For a preliminary analysis of the relationship, the authors calculated the correlation coefficient, and based on it, a correlation matrix was constructed (Table 3). Correlation, as a relationship between phenomena, can be more or less close, i.e. the dependence of one quantity on another is more or less clearly expressed. The main task of the correlation method is to establish the closeness (strength) of the connection between phenomena. The closer the connection, the greater the influence of the studied factor on the result and the less influence of extraneous factors for this case. In order to most fully identify the dependence of the factorial and effective indicator in the dynamics of the studied indicators, a larger number of periods should be taken. After the initial filtering of statistical data, there are 11 periods left to calculate the model. As described above, this is done in order to exclude periods with unique cases or errors. At the intersection of a row and a column, you can observe the correlation coefficient between variables. The matrix is symmetric with respect to the diagonal.

Table 3. Correlation matrix

	y	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14
y	1														
x1	0,9687867	1													
x2	0,3388247	0,2027064	1												
x3	0,7973520	0,846712	-0,134927	1											
x4	0,0138902	-0,105786	0,6557071	-0,469200	1										
x5	-0,767327	-0,743650	-0,293616	-0,518378	0,1171118	1									
x6	0,9778858	0,9324928	0,3374694	0,7690235	0,0850493	-0,754696	1								
x7	0,9620863	0,9319283	0,4137646	0,7872893	0,1191196	-0,622042	0,9469313	1							
x8	0,6051244	0,4498662	0,5267970	0,2085913	0,6181037	-0,359209	0,7054611	0,6398856	1						
x9	0,4094084	0,2475183	0,6388532	-0,074922	0,6637337	-0,273528	0,4637613	0,4376812	0,8665849	1					
x10	0,9233334	0,8486155	0,3470033	0,7067139	0,2389969	-0,558677	0,9457355	0,9434351	0,7875204	0,5213488	1				
x11	-0,028240	-0,146958	0,4465264	-0,331572	0,7487323	0,4027213	0,0174491	0,1260495	0,6130990	0,7020885	0,2452275	1			
x12	0,9698435	0,9860683	0,2206046	0,8588197	-0,063201	-0,666336	0,9295630	0,9532796	0,4685435	0,2446952	0,8879662	-0,088442	1		
x13	0,9903573	0,9412034	0,3582176	0,7432813	0,0646419	-0,809512	0,9827305	0,9397425	0,6680515	0,4788197	0,9261819	-0,000878	0,9322748	1	
x14	-0,862110	-0,739250	-0,599037	-0,434775	-0,308134	0,7467576	-0,872433	-0,806768	-0,814568	-0,749289	-0,834430	-0,252295	-0,725887	-0,904857	1

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.

We see that there is a connection between some variables X. In particular, there is a very strong direct correlation between X1 and X6, between X1 and X7. On Figure 1, you can see this relationship. All points lie approximately on the same straight line. There is also a significant relationship between the other variables. According to the data presented in Table 3, it can be seen that the variables X6, X7, X12 and X13 correlate with the rest of the indicators. If there is a strong correlation between variables, this is called multicollinearity. At the same time, one of the variables should be excluded from the calculation. If this is not done, it can lead to the following problems:

- Small changes in the source data will lead to large changes in the coefficients.
- Instability of the solution.
- There is a high probability of a model error.

We are also looking at the Y column. These values show how strongly the variables X affect Y. In correlation analysis, it is established when the correlation coefficient:

r > 0,7 - the relation is considered close;
0.5 < r < 0.7 - the relation is average;
r < 0.5 - the relation is weak.

With a weak connection between the function Y and the argument X, the influence of this factor, taken as X, is insignificant and can be neglected. And the change in the performance indicator is mainly due to other factors.

On the table we see that X2, X4, X9, X11 have little effect on Y. Only X1, X3, X5, X6, X7, X10, X12 have a significant effect on the function. The indicator X8 has an average relationship. For a better analysis, values with a close correlation were selected, that is, where r > 0.7. The remaining indicators were excluded from the calculation.

Relationship between X1 and X6 Relationship between X1 and X7

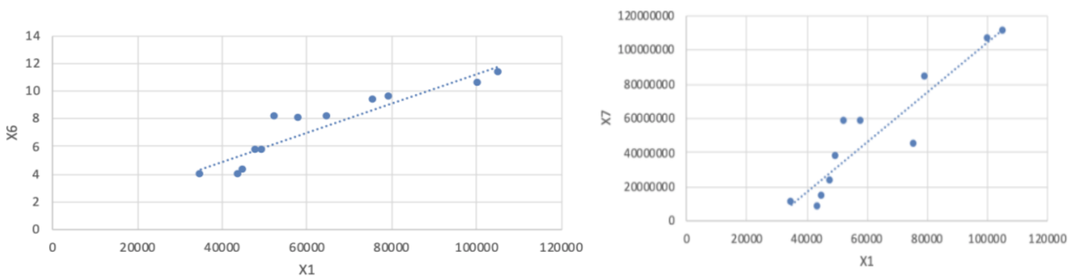

Figure 1. Multicollinearity between X1 and 6, X1 and X7

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.

As a result, the correlation matrix was reduced to 5 variables (Table 4). We can observe that the most significant impact on the GVA is X1, that is, the number of organizations using the Internet is slightly less than 0.97. The next variable in influence was X10, exports from the Republic of Kazakhstan with a value of 0.92. Indicator X3, Computer literacy at the age of 6 years and older: a novice user, slightly below 0.80. Unlike the first three variables, which have a direct correlation with the GVA, the remaining two indicators under consideration are X14 (Unemployment rate) and X5 (Computer literacy at the age of 6 years and older: Experienced user) has a negative effect on the VDS: approximately -0.86 and -0.77, respectively. The fact that the reduction in unemployment will have a positive impact on the GVA was expected, however, it is surprising that despite the decrease in the number of experienced PC users, gross value added is steadily growing.

Table 4. Correlation table, with variables where r > 0.7 with respect to Y

	y	x1	x3	x5	x10	x14
y	1					
x1	0,96878674	1				
x3	0,79735205	0,846712	1			
x5	-0,7673273	-0,7436505	-0,5183783	1		
x10	0,92333349	0,84861553	0,70671391	-0,5586777	1	
x14	-0,8621109	-0,7392504	-0,4347756	0,74675768	-0,8344306	1

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz

Compilation of a regression model.

The correlation coefficient indicates only the degree (closeness) of the relationship in the variation of the two variables. But it does not give an idea of how one quantity changes quantitatively as the other changes.

With the help of regression, the task is to establish how the effective indicator Y changes quantitatively when the factorial indicators X change by one. Thus, a model is formed that makes it possible to predict the change in the result Y with a given change in factors X.

At this stage, a formula has been drawn up by which we could, knowing the variables X, or rather the number of organizations using the Internet, the number of experienced PC users and the number of novice PC users, calculate the VDS. A straight-line regression model will be applied here:

$$Y = a0 + a1x1 + \ldots + anxn,$$

where Y is the GVA;

a0, a1, an - the regression coefficients;

x1, xn - variables.

This equation reflects a uniform change in the performance indicator with a change in factor indicators. The projected calculation of GDP is made by substituting the values of the corresponding factors into the planned equation.

The values of the parameters of the regression coefficients (a_0, a_1, etc.) can be found in various ways. The most common is the least squares method. With this method, the line that aligns the empirical data should pass so that the sum of the squares of deviations from this line is the smallest.

Using the Data Analysis tool from MS Excel, we can get the following data:

Table 5. Regression analysis with variables X1, X3, X5, X10, X14.

Output of results

Regression statistics	
Multiple R	0,99789005
R-square	0,99578456
Normalized R-square	0,99156912
Standard error	1583038,68
Observations	11

Analysis of variance

	df	*SS*	*MS*	*F*	*Significance of F*
Regression	5	2,96E+15	5,9198E+14	236,22308	6,24E-06
Remains	5	1,25E+13	2,506E+12		
Total	10	2,97E+15			

	Coefficients	Standard error	t-statistics	P-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Y-intersection	32629014,2	13579462	2,40282084	0,06140631	-2278103,99	67536132,5	-2278104	67536132,5
x1	343,808752	69,6979927	4,93283577	0,00434892	164,644358	522,973146	164,644358	522,973146
x3	0,03521539	0,01879043	1,87411327	0,11977892	-0,01308694	0,08351773	-0,0130869	0,08351773
x5	-0,0146264	0,0232805	-0,628269	0,55740518	-0,07447085	0,04521802	-0,0744708	0,04521802
x10	2124,85111	1214,79854	1,74913867	0,14067841	-997,887967	5247,59018	-997,88797	5247,59018
x14	-7514230,9	2414396,35	-3,1122607	0,02648297	-13720634,3	-1307827,5	-13720634	-1307827,5

Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.

The most important indicators in this table were highlighted in bold. The first is the R-square. This value shows how much the change in Y can be explained by changes in variables X and shows the adequacy of the equation. In this case, the indicator 0.99 is very high. The next thing to pay attention to is the significance of F. The significance of F allows you to check the significance of the regression equation, i.e. to determine whether the mathematical model expressing the dependence between variables corresponds to experimental data and whether the explanatory variables included in the equation (one or several) are sufficient to describe the dependent variable. For the significance of the model, it should not exceed 0.05. In our case, it is equal to 6.24E-06, therefore the overall significance is confirmed.

In addition, P-values were determined. The P-value is the lowest value of the significance level (i.e., the probability of rejection of a fair hypothesis) for which the calculated verification statistics leads to rejection of the null hypothesis. Usually, the p-value is compared with the generally accepted standard significance levels of 0.005 or 0.01. For example, if the value of the test statistics calculated from the sample corresponds to p = 0.005, this indicates a probability of validity of the hypothesis of 0.5%. Thus, the smaller the p-value, the better, since this increases the “strength” of the rejection of the null hypothesis.

However, in the table we can observe that the P-values for X5 and X10 critically exceed the permissible levels (0.55, and 0.14, respectively). Therefore, further construction of the model based on these indicators is not statistically significant. Thus, it can be concluded that the resulting multiple regression equation is significant, but its adequacy is rather low, therefore, the recommendation is to remove statistically insignificant factors in order to ensure the accuracy and quality of the model.

Based on this, it was decided to exclude the indicators X5 and X10 and leave in the model only the GVA, the number of organizations using the Internet, Computer literacy at the age of 6 years and older; the novice user and the unemployment rate. And as a result, we simplify the model to three variables and build Table 6 on its basis.:

Table 6. Regression analysis with variables X1, X3, X14.

Output of results

Regression statistics	
Multiple R	0,99642293
R-square	0,99285866
Normalized R-square	0,98979809
Standard error	1741388,78
Observations	11

Analysis of variance

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	2,95E+15	9,84E+14	324,40287	7,15E-08
Remains	7	2,12E+13	3,03E+12		
Total	10	2,97E+15			

	Coefficients	Standard error	t-statistics	P-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Y-intersection	52476960,2	8689588,66	6,03906148	0,00052159	31929348,1	73024572,3	31929348,1	73024572,3
x1	370,257029	68,8772607	5,3756062	0,00103525	207,388188	533,12587	207,388188	533,12587
x3	5,011375	0,01866986	2,68420572	0,03134769	0,00596654	0,09426095	0,00596654	0,09426095
x14	-10610029	1479690,18	-7,1704394	0,00018208	-14108940	-7111117,4	-14108940	-7111117,4
Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.								

Table 6 shows that all 4 coefficients, Y-intersection, X1, X3, 14 of them are statistically significant. The P-values of 5.2E-04, 0.001, 0.031 and 1.8E-04, respectively, are very scanty, which means that the randomness of the correct result in the first coefficient is 0.05%, in the second 0.1%, in the third 3%, and in the fourth is almost zero. The significance of F also does not exceed 0.05. The R-square of the model is greater than 99%, which indicates a very high approximation accuracy (the model describes the phenomenon well).

The regression equation will be calculated using the following formula:

$$Y1 = a_0 + a_1x_1 + a_2x_3 + a_3x_{14},$$

where Y1 is the estimated GVA;

a_0,a_1,a_2,a_3 - regression coefficients;

x_1 - number of organizations using the Internet units network;

x_3 - computer literacy at the age of 6 years and older: Novice user (was changed to a quantitative indicator taking into account the population for the corresponding periods), units;

x_{14} - unemployment rate, as a percentage.

The coefficients $a_0, a_{1,}, a_{2,}, and\ a_3$ can be found by the least squares method, or you can look in Table 6 in the coefficients column. The equation in its final form, taking into account the found parametersa$a_0, a_{1,}, a_{2,}, and\ a_3$ will take the form:

$$Y1 = 52476960{,}2 + 370{,}257029x_1 + 5{,}011375x_3 - 10610029x_{14},$$

To check the calculated values of a0, a1, a2, and a3, the values are substituted into both initial equations of the system. Performing equalities in the original equations with the calculated data a0 and a1 will indicate the correctness of the calculation.

If we substitute various values of the number of organizations using the Internet, the number of novice PC users and the unemployment rate into the regression equation, we get the theoretical values of the GVA (Y1) corresponding to these indicators (Table 7).

Table 7. Theoretical values of the GVA (Y1) and the percentage of error of the model.

year	X1	X3	X14	Y	Y1	Error
2008	35089	3566280	6,6	13056532,9	13314684,41	2%
2009	44046	2960560	6,6	14506780,8	13595586,85	6%
2010	45354	2692800	5,8	21115891	21226060,36	1%
2011	48064	2649600	5,4	25741874,8	26256977,01	2%
2012	49853	2921460	5,3	28528090,1	29342761,98	3%
2013	58456	2982000	5,2	32896601	33892474,69	3%
2014	52630	2991170	5	36651572,2	33903317,28	7%
2016	75779	3735900	5	44337585,5	46206518,19	4%
2017	79658	4167240	4,9	51195859,3	50865354,36	1%
2018	100702	4295800	4,9	57706553,3	59301305,59	3%
2019	105531	4368360	4,8	64681604,8	62513904,99	3%
					Average error	3%
					Maximum error	7%
Note – compiled by the authors on the basis of Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the RK, https://stat.gov.kz.						

The parameters of the regression equation $a_{1,}$, a, $and\ a_3$ are called regression coefficients, which are the main indicators in the equation. Regression coefficients $a_{1,}, a_{2,}, and\ a_3$ show by how many units on average the effective indicator Y changes with a change in factor indicators X by one.

Based on the results of the calculations, it is advisable to draw the following conclusions:

An increase in the number of organizations using the Internet by 10,000 units over the period corresponded to an increase in the GVA by about 3,702,000 million tenge. In other words, the equation shows that an increase in the number of companies using the Internet by 9.47% will entail an increase in GVA by more than 5.7%.

An increase in the number of people with computer knowledge at the Novice user level by 10% will lead to an increase in the GVA by 3.38%.

And, finally, a decrease in the unemployment rate by 1% corresponds to an increase in the GVA by 16.4%.

The free term of the equation a_0 is a certain basis that must be taken into account when using the regression coefficient.

Using Y1, we can estimate the accuracy of the model. In column 7 (errors) of Table 7, the percentage of model error was calculated. The average error rate is 3%. Considering this and other above-mentioned checks, in particular R-squared (greater than 99%), P-values (5.2E-04, 0.001, 0.031 and 1.8E-04 for Y and X1, X3 and X14, respectively) and the significance of F (does not exceed 0.05), we can conclude that the model gives quite an acceptable good result, and with its help it is possible to make a forecast of the GVA at the specified (planned) values of the number of organizations using the Internet, the number of people with a computer proficiency level at the level of the initial user and the unemployment rate.

Discussions

The results of the regression analysis (Table 6) suggest that the GVA can be explained, among other things, by an increase in the number of organizations using the Internet and the level of computer literacy of the population, as well as a decrease in unemployment. The estimated coefficients are statistically significant, which means that digitalization is a significant indicator of economic growth. Thus, the results obtained confirm the study of Solomon and van Clayton (2020) on the positive impact of digitalization on the economy. In addition, by the method of correlation and regression analysis, a fairly correct mathematical model was built in our article that determines the degree of this influence.

In addition, our study, which establishes a link between the indicators of digital transformation and economic growth, supports the idea of focusing investments of the Republic of Kazakhstan on digital convergence within the framework of the Digital Kazakhstan program plan, which aims to accelerate the pace of development of the economy of the Republic of Kazakhstan and improve the quality of life of the population through the use of digital technologies in the medium term, as well as the creation of conditions for the transition of Kazakhstan's economy to a fundamentally new development trajectory, ensuring the creation of the digital economy of the future in the long term. So, we recommend Kazakh investments to focus on the process of digital transformation and its acceleration for further growth.

Conclusions

Regression analysis showed that the influence of each indicator on the GVA is unbalanced and strongly depends on the correlation between variables. As expected, the problem of strong correlation led to double counting and inaccuracies in calculations. You can get a very accurate model by reducing the set of diagnostic variables.

The results of the analysis positively confirmed the hypothesis that economic growth measured by gross value added can be reliably explained by indicators of digital transformation. At this stage of Kazakhstan's development, the digital transformation of the country's enterprises has a positive impact on its economic growth. The authors concluded that the quantitative growth of companies using the Internet in their activities significantly affects the country's GVA indicator. The level of computer literacy affects this indicator a little less. The unemployment rate was the leader among the indicators. It is assumed that, as technology develops, the digitization process will be carried out faster and cheaper, which will entail a greater positive impact on economic growth.

Our results are of great importance to government authorities in terms of measuring, supporting and deepening digital transformation. If government agencies in the Republic of Kazakhstan want to support and even stimulate economic growth, it is recommended to legislatively encourage digital transformation, paying special attention to companies using the Internet in their activities, improving computer literacy of the population and reducing unemployment, since these indicators gave the highest connection and accuracy with economic growth in our model. It is advisable, in our opinion, to introduce and calculate indicators of digitalization of workplaces.

Our research has some limitations. When using the model, after filtering and excluding variables, there are only 3 indicators that can be used with high accuracy for planning the GVA. When trying to use additional variables, the accuracy of the model is significantly reduced. We plan to conduct a series of analyses and studies to identify other indicators to improve our model.

References

Abdel Azim, R., Salman, O., & El Henawy, I. (2020). The Role of e-Government as a stimulus for economic growth. *The International Journal of Business Management and technology, 4,* 69–79. http://www.theijbmt.com/archive/0935/1632455370.pdf

Aker, J. (2017). Using digital technology for public service provision in developing countries potential and pitfalls (in S. Gupta, M. Keen, A. Shah, & G. Verdier eds.). *Digital revolutions in public finance,* 201–224. International Monetary Fund.

Arvin, B., & Pradhan, P. (2014). Broadband penetration and economic growth nexus: Evidence from cross-country panel data. *Applied Economics, 46,* 4360–4369.

Birlea, S., & Capatina, A. (2017). The impact of internet and e-commerce on economic growth. *Journal of Danubian Studies and Research, 7,* 48–57.

Elding, C., & Morris, R. (2018). Digitalization and its impact on the economy: Insights from a survey of large companies. *ECB Economic Bulletin, 7,* 1–9. https://www.ecb.europa.eu/pub/economic-bulletin/focus/2018/html/ecb.ebbox201807_04.en.html

Hilbert, M. (2020). Digital technology and social change: The digital transformation of society from a historical perspective. *Dialogues in Clinical Neuroscience, 22,* 189–194. https://doi.org/10.31887/dcns.2020.22.2/mhilbert

Karavasilis I., Zafiropoulos K., & Vrana V. (2010). Extending TAM to Understand E-Governance Adoption by Teachers in Greece. *Organizational, Business, and Technological Aspects of the Knowledge Society, 112*, 57-68. https://doi.org/10.1007/978-3-642-16324-1_7

Kim, P. S., & Thanh H. T. (2016). The Impact of E-Government on Competitiveness-based Economic Growth in Vietnam. *Korea Open Access Journals.* https://DOI.org/10.21485/hufsea.2016.26.2.003

Kuttybaeva, N. B., Raikhanova, G. A., Zhaparova, R. Ye., Kozhabatchina, G. M., & Akmolda, M. N. (2022). Technological audit in the system of optimization of innovative activity. *Bulletin of the Karaganda university. Economy series, 106,* 71-80. https://doi.org/10.31489/2022Ec2

Macdougald, J. J. (2011). Internet use and economic development: Evidence and policy implications. http://scholarcommons.usf.edu/etd/3225

Molinari, B., & Torres, J. L. (2018). Technological sources of economic growth in Europe and the U.S. *Technological and Economic Development of Economy, 24,* 1178–1199. https://doi.org/10.3846/20294913.2017.1280557

Mühleisen, M. (2018). The long and short of the digital revolution. *Finance and Development, 55,* 4–8.

Nelder, J. A., & Mead, R. (1965). Simplex method for function minimization. *Computer Journal*, 7 (4), 308–313. https://doi.org/10.1093/comjnl/7.4.308

Niebel, T. (2018). ICT and economic growth – Comparing developing, emerging and developed countries. *World Development, 104,* 197–211.https://doi.org/10.1016/j.worlddev.2017.11.024

OECD. (2019). OECD skills Outlook 2019. *Thriving in a Digital World. Paris.* https://doi.org/10.1787/df80bc12-en

OECD. (2019b). Going digital: Shaping policies, improving lives. *Going digital: Shaping policies, improving lives.* https://doi.org/10.1787/9789264312012-en

Panda, P., Sahu, G. P. & Gupta B. (2019). E-government procurement implementation in India: evolving decision parameters for project success. *International Journal of Business Information Systems, 31*, 414-454.

Peetz, D. (2019). The realities and futures of work. *ANU Press.* https://doi.org/10.22459/RFW.2019

Ritter, T., & Pedersen, C. (2020). Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future. *Industrial Marketing Management, 86,* 180–190. https://doi.org/10.1016/j.indmarman.2019.11.019

Roszko-Wójtowicz, E., & Grzelak, M. M. (2020). Macroeconomic stability and the level of competitiveness in EU member states: A comparative dynamic approach. *Oeconomia Copernicana, 11,* 657–688. https://doi.org/10.24136/oc.2020.027

Sabbagh, K., Friedrich, R., El-Darwiche, B., Singh, M., & Koster, A. (2013). Digitization for economic growth and job creation: Regional and industry perspectives. *In The global information technology report,* 35–42.

Shakeyev, S. S., Nevmatulina, K. A., & Vladimirov Z. (2021). Theoretical foundations and main stages of the transformation of the digitalization of the economy. *Bulletin of the Karaganda university. Economy series, 104,* 85-93. https://doi.org/10.31489/2021Ec4

Solomon, E. M., & van Klyton, A. (2020). The Impact of Digital Technology Usage on Economic Growth in Africa.

Utilities Policy, 67. https://doi.org/10.1016/j.jup.2020.101104

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70,* 65–94. https://doi.org/10.2307/1884513

Spirakis, G., Spiraki, C., & Nikolopoulos, K. (2010). The impact of electronic government on democracy: E-democracy through e-participation. *Electronic Government, 7,* 75–88. https://doi.org/10.1504/EG.2010.029892

Sredojević D., Cvetanović S., & Bošković G. (2016). Technological changes in economic growth theory: Neoclassical, endogenous, and evolutionary-institutional approach. *Economic Themes, 54,* 177-194. https://doi.org/10.1515/ethemes-2016-0009

Statistics of information and communication technologies (2007-2020). *Agency for Strategic planning and reforms of the Republic of Kazakhstan Bureau of National statistics.* Retrived from https://stat.gov.kz/official/industry/29/statistic/7

Thomassen, L., Derrida, J., & Habermas, J. (2006). The Derrida-Habermas Reader.

T. T. Hoa, S. K. Pan. (2016). The Impact of E-Government on Competitiveness-based Economic Growth in Vietnam. https://doi.org/10.21485/hufsea.2016.26.2.003

United Nations Conference on Trade and Development (2019). *Digital economy report 2019: Value creation and capture – implications for developing countries. UNCTAD.* Retrived from https://unctad.org/system/files/official-document/der2019_en.pdf

Vyshnevskyi, O., Stashkevych, I., Shubna, O., & Barkova, S. (2020). Economic growth in the Conditions of digitalization in the EU countries. *Studies of Applied Economics, 38,* 2–17. https://doi.org/10.25115/eea.v38i4.4041

Zhao, F., Wallis, J., & Singh, M. (2015). E-government development and the digital economy: A reciprocal relationship. *Internet Research, 25,* 734–766. https://doi.org/10.1108/IntR-02-2014-0055

T.K. Shuren, T.P. Pritvorova, E.A. Vechkinzova, A. Kizimbayeva

Цифрлық трансформация және экономикалық өсумен өзара байланыс

Аңдатпа

Мақсаты: Мақаланың негізгі мақсаты цифрлық трансформация көрсеткіштерінің Қазақстандағы экономикалық өсуге әсерін анықтау және бағалау.

Әдісі: Зерттеу барысында статистикалық көп корреляциялық-регрессиялық талдау әдістері MS Excel бағдарламасының «Мәліметтерді талдау» қолданбалы пакеті негізінде қолданылды. Зерттеудің ақпараттық базасы Қазақстан Республикасы Стратегиялық жоспарлау және реформалар агенттігінің Ұлттық статистика бюросының statgov.kz ресми сайтының мәліметтері.

Қорытынды: Статистикалық көрсеткіштерді бастапқы іріктеу жүргізілді және 2007-2020 жылдар кезеңінде экономиканың өсуіне гипотетикалық әсер ететін факторлар (және тиісті индикаторлар) тобы айқындалды. Ең маңызды факторларды таңдау негізінде регрессия теңдеуі құрылды, ол алынған ЖҚҚ-ға әсер ету дәрежесін көрсетеді. Алынған регрессиялық модель бағаланды. Табылған регрессия теңдеуі Фишер критерийі бойынша маңызды, оның барлық параметрлері, оның ішінде бос термин, Стьюдент критерийі бойынша максималды қатесі 0,07-ге тең. Корреляцияның бірнеше коэффициенті-0,99. Алынған нәтижелер аймақтық және ұлттық деңгейде ЖІӨ мен ЖҚҚ жоспарлау кезінде пайдалы болуы мүмкін.

Тұжырымдама: Жалпы қосылған құн индикаторлары жүйесінде Интернет желісін пайдаланатын ұйымдар саны, жұмыссыздық деңгейі және халықтың компьютерлік сауаттылығы маңызды орын алады. Талдау осы көрсеткіштер арасындағы күшті байланысты көрсетеді. Нәтижесінде осы көрсеткіштер арасындағы қатынасты 97% орташа дәлдікпен сызықтық теңдеумен түсіндіруге болатындығы көрінді. Сонымен қатар, жағдайды неғұрлым барабар талдау үшін ҚР-дағы жұмыссыздық көрсеткіштерінің өзгеруі мен жалпы қосылған құн арасындағы кері өзара байланысты да ескеру қажет. Осы көрсеткіштер арасындағы теріс корреляция жұмыссыздық деңгейінің өзгеруінен экономиканың осалдығы мен тұрақсыздығын растайды.

Кілт сөздер: цифрлық экономика, экономикалық өсу, цифрландырудың әсері, бірнеше регрессия, корреляция.

Т.К. Шурен, Т.П. Притворова, Е.А. Вечкинзова, А. Кизимбаева

Цифровая трансформация и взаимосвязь с экономическим ростом

Аннотация:

Цель: Основной целью данной статьи является выявление и оценка влияния показателей цифровой трансформации на экономический рост в Казахстане.

Методы: При проведении исследования были использованы методы статистического множественного корреляционно-регрессионного анализа на основе прикладного пакета «Анализ данных» программы MS Excel.

Информационной базой исследования послужили данные официального сайта Бюро национальной статистики Агентства по стратегическому планированию и реформам РК: statgov.kz.

Результаты: Проведен первичный подбор статистических показателей и определена группа факторов (и соответствующих индикаторов), гипотетически влияющих на рост экономики за период 2007–2020 гг. Выделены 14 факторов, имеющих значимое влияние на валовую добавленную стоимость. На основе отбора наиболее значимых факторов построено уравнение регрессии, демонстрирующее степень влияния на результирующий ВДС. Выполнена оценка полученной регрессионной модели. Найденное уравнение регрессии значимо по критерию Фишера, все его параметры, в том числе и свободный член, значимы по критерию Стьюдента с максимальной ошибкой 0,07. Множественный коэффициент корреляции равен 0,99. Полученные результаты могут быть полезными при планировании ВВП и ВДС как на региональном, так и национальном уровне.

Выводы: В системе индикаторов валовой добавленной стоимости важное место занимают показатели количество организации, использующих сеть Интернет, уровень безработицы и компьютерная грамотность населения. Проведенный анализ демонстрирует сильную взаимосвязь между этими показателями. В итоге, мы увидели, что взаимосвязь между этими показателями можно объяснить линейным уравнением со средней точностью в 97 %. В то же время для более адекватного анализа ситуации следует также учитывать обратную взаимосвязь между изменением показателей безработицы в РК и валовой добавленной стоимостью. Отрицательная корреляция между этими показателями подтверждает уязвимость и неустойчивость экономики от изменений уровня безработицы.

Ключевые слова: цифровая экономика, экономический рост, влияние цифровизации, множественная регрессия, корреляция.