
Table of Contents

Preface	7
<i>(Marek Miłosz, Ualsher Tukeyev)</i>	
1. Complex Technology of Machine Translation Resources Extension for the Kazakh Language	9
<i>(Ualsher Tukeyev, Diana Rakhimova, Zhandos Zhumanov, Dina Amirova, Aidana Karibayeva)</i>	
2. The Structural Transformation of Sentences for the Kazakh-Russian and Kazakh-English Language Pairs in Machine Translation System	23
<i>(Ualsher Tukeyev, Aida Sundetova, Diana Rakhimova, Zhandos Zhumanov, Balzhan Abduali, Aliya Turganbaeva)</i>	
3. Independent Computer Presentation of Adverbs in Natural Languages	45
<i>(Batyigul Bayachorova, Pavel Pankov)</i>	
4. Recognition of Kazakh Speech in Smart University Strategy	51
<i>(Altynbek Sharipbay, Alibek Barlybayev, Bibigul Razakhova, Gulmira Bekmanova)</i>	
5. Development of Algorithms for the Magnetic Attitude Control of Small Spacecraft	71
<i>(Zaure Rakisheva, Nazgul Kaliyeva, Anna Sukhenko, Gulnara Mayemerova)</i>	
6. Algorithms Development of the Attitude Determination and Control of the Low-Orbit Small Satellites	87
<i>(Zaure Rakisheva, Anna Sukhenko, Nazgul Kaliyeva, Kuanysh Alipbayev)</i>	
7. The Earth's Remote Sensing for Exploration of Mineral Resources of the Kyrgyz Republic.....	107
<i>(Gulnaz Tultemirova)</i>	
8. Mathematical and Computer Modeling of Spillway Medeu	121
<i>(Belgibayev Baurzhan, Berik Magzhan, Zhakupbekova Assem, Kurbanalieva Ayiman)</i>	

9. Simulation Systems	145
<i>(Nella Israilova)</i>	
10. Introducing Informatics Topics to Students in ISC	159
<i>(Petra Sauer, Agathe Merceron, Heike Ripphausen-Lipa)</i>	
11. Optimal Stabilization of Continuous Linear Systems with Constrained Control on a Finite Time Interval	173
<i>(Zainel Murzabekov)</i>	
12. Development of Information Systems for the Simulation of Urbanization Processes in Almaty	189
<i>(Saya Sapakova, Umit Anarbek)</i>	
13. Correction of Incomplete and Incorrect Time Series for Traffic Prediction	217
<i>(Petra Sauer, Stephan Pieper, Felix Kunde, Salke Waterloo)</i>	
14. Aspects of EEG Acquisition Process	227
<i>(Safia Elidrissi Tuir, Nabil Benamar, Grzegorz Koziel, Marek Miłosz, Nada Sabil)</i>	
15. Composition Problem in Programming Language Java	241
<i>(Telman Boltaev)</i>	
16. Some Joint Applications of the Bootstrap and Kernel Density Estimation Methods	267
<i>(Mirbulat Sikhov, Kalamkas Kuanyshbekova, Raushan Serikzhan)</i>	
17. A Comparative Study of the Static Code Quality of Different PHP Frameworks	291
<i>(Marek Milosz, Mykyta Kochetkov)</i>	
Author Index	303

Development of Information Systems for the Simulation of Urbanization Processes in Almaty

In this chapter we consider the problem of modeling the process of urbanization in Almaty, collected and structured statistical data, on the basis of which defining the major factors influencing to this process. As a result of work was developed information system, which run a multiple regression and analysis of variance to assess, evaluate the quality of the regression equation with Fisher criterion and to assess the statistical significance of the regression coefficients and the correlations are calculated Student's t-test. In the process of creation an information system for the simulation of urbanization used web-technology and a text editor JetBrains PhpStorm. To retrieve information from another site feature is designed for robot-parser technology. The created information system allows us to make reliable forecasts of the urbanization process for the next years, which is important for companies that implement long-term projects in our country.

12.1. INTRODUCTION

Urbanization issues are becoming increasingly relevant in connection with the formulation of the task of Kazakhstan's entry into the list of 30 highly developed states of the world [4]. In the Address of the President of the Republic of Kazakhstan, the people are told about the need to ensure the dynamic development of the infrastructure triad - agglomerations, transport, energy. Agglomerations are regarded

as the framework of the knowledge-based economy of Kazakhstan. The first modern urban centers of Kazakhstan should be Astana, Almaty, Shymkent and Aktobe. Specific instructions have been given to the Government to develop, by the end of current year, draft strategies for the formation of agglomerations in the cities of Astana and Almaty for the period until 2030 [1, 4].

The first part of the thesis contains the meaning and importance of the process as well as a report for the achievement of the goal. In the second part, a model of the urbanization process was created, and a regression model in the tabular processor was developed. In the third part the practical implementation of this work, the features of the information system of the urbanization process model was considered.

12.2. THE PROCESS OF URBANIZATION

Urbanization tends to correlate positively with industrialization. With the promise of greater employment opportunities that come from industrialization, people from rural areas will go to cities in pursuit of greater economic rewards. Urbanization is relevant to a range of disciplines, including geography, sociology, economics, urban planning and public health. The phenomenon has been closely linked to modernization, industrialization and the sociological process of rationalization. So urbanization can be quantified either in terms of, the level of urban development relative to the overall population, or as the rate at which the urban proportion of the population is increasing. The process creates enormous social, economic and environmental changes [1].

Urbanization is not merely a modern phenomenon, but a rapid and historic transformation of human social roots on a global scale, whereby predominantly rural culture is being rapidly replaced by predominantly urban culture. Rural inhabitants come to the city to seek their fortune and alter their social position. Businesses, which provide jobs and exchange capital, are more concentrated in urban areas. Whether the source is trade or tourism, it is also through the ports or banking systems, commonly located in cities, that foreign money flows into a country.

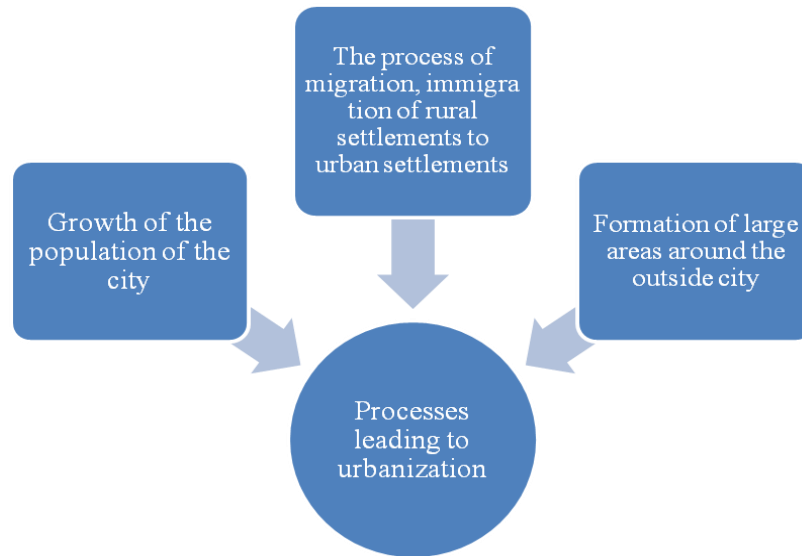


Fig. 12.1. Processes leading to urbanization

Source: [1, 2]

Demographic aspects of the process are determined by the population of the city. This figure differs in each country. For the integration of data, a certain number of people in the region will be given. For example, the United Nations projected that half of the world's population would live in urban areas at the end of 2008. It is predicted that by 2050 about 64% of the developing world and 86% of the developed world will be urbanized. That is equivalent to approximately 3 billion urbanites by 2050, much of which will occur in Africa and Asia. Notably, the United Nations has also recently projected that nearly all global population growth from 2017 to 2030 will be absorbed by cities, about 1.1 billion new urbanites over the next 13 years.

The main aspect of the current urbanization process is determined by the number of people in urban areas. Australia, Europe and North America are facing a high number of people in urban areas. Countries in Europe that provide a high level of urbanization: Great Britain - 91%, Belgium 72%, and 77% urban population of the United States of America, Canada, 76%.

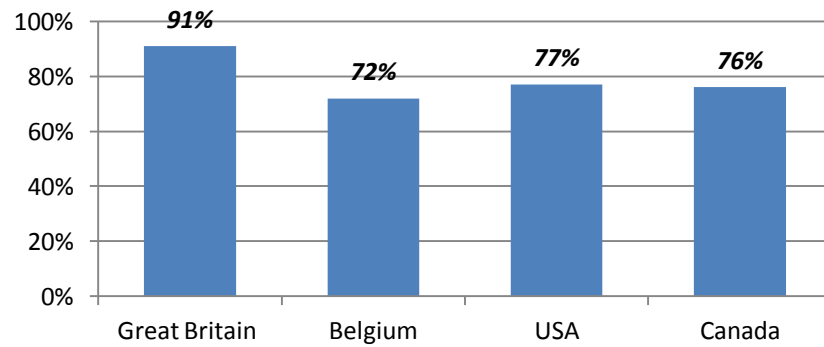


Fig. 12.2. Urban Level Chart

Source: [5]

The growth of the population, the structure and location of modern processes lead to numerous difficulties. In recent years, the whole world has been paying great attention to the development of the urbanization process. In this regard, the modeling of urbanization processes is underway in many developed countries [2].

12.3. FACTORS AFFECTING THE URBANIZATION PROCESS

The process of urbanization, the growth of cities, have an impact on the formation of national economic growth and development. Factors that lead to the growth of cities will be divided into three levels: people, culture and globalization. A person defines a modern description of the process of urbanization, or the level of quality of life in the city. The quality of life in the city, the natural and climatic conditions, the quality of life in terms of living standards, housing, health, transport, education, culture, the environment are factors in the development of the urbanization process [3].

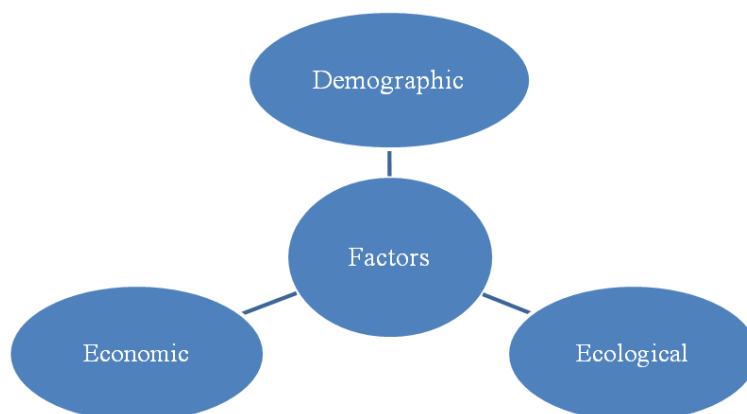


Fig. 12.3. Factors affecting the urbanization process

Source: [2]

The role of demographic factors, population, migration, urban settlements are determined by the results of calculations. The main factor for modeling urbanization processes is the demographic index.

Table 12.1. Demographic indicators in Almaty, thousands of people

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	1 247.9	1 278.3	1 324.7	1 361.9	1 390.7	1 413.5	1 449.8	1 475.4	1 507.5	1 642.3	1 703.5

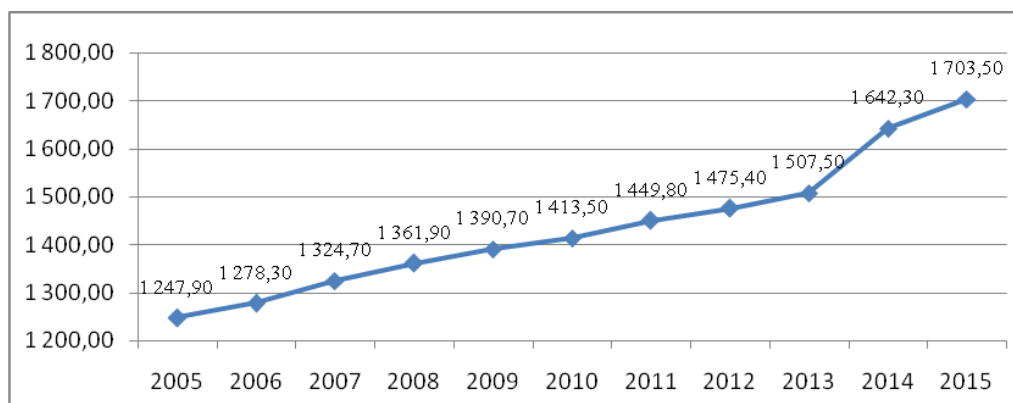


Fig. 12.4. Demographic indicators in Almaty

Source: [6]

Economic factors include changes in the structure of employment, concentration of production and cross-border economic development. Tab. 12.2 shows the standard of living in Almaty from 2005 to 2015, the average nominal monthly, and the unemployment rate in Tab. 12.3 and Tab. 12.4 shows the result of 10 years of price index in the city.

Table 12.2. The standard of living of the population in Almaty, tenge

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	39 614	49 201	59 240	78 021	90 239	95 139	106 597	121 674	134 378	142 310	155 242

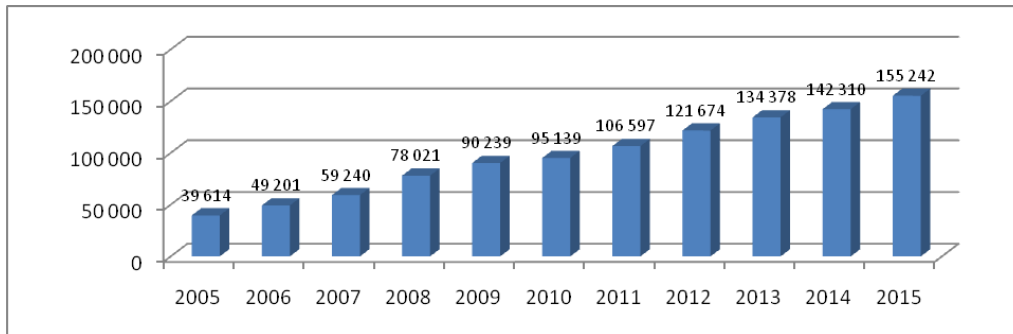


Fig. 12.5. Standard of living in Almaty

Source: [6]

Table 12.3. Unemployment rate, percentage

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	8.8	8.4	8.2	7.8	7.4	7.7	6.3	5.6	5.6	5.6	5.5

The unemployment rate is one of the factors affecting the urbanization process. Nevertheless, the demographic value of the index, and the unemployment rate are inversely proportional to the value of the index values. The unemployment rate has decreased, the number of urban population growth is increasing, which is an indicator of the development of the urbanization process [4].

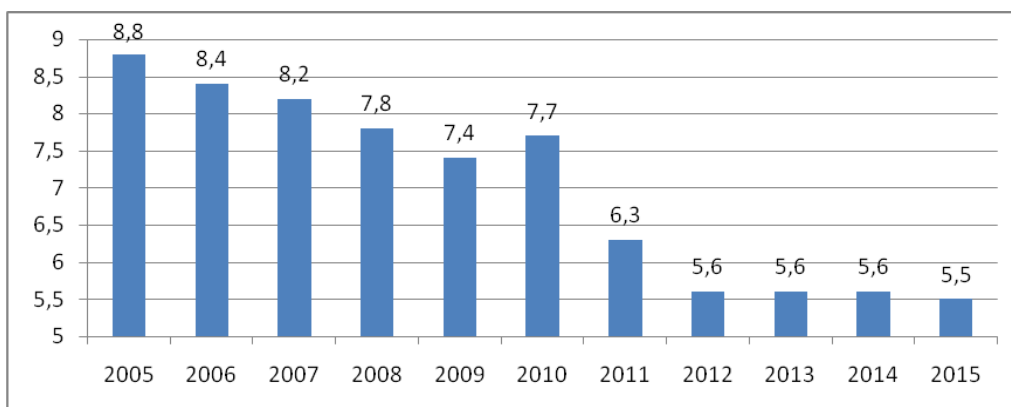


Fig. 12.6. Unemployment rate in Almaty

Source: [6]

Table 12.4. Consumer price index, percentage

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	109.6	112.0	127.1	110.7	107.5	108.0	106.8	106.1	104.4	107.0	114.8

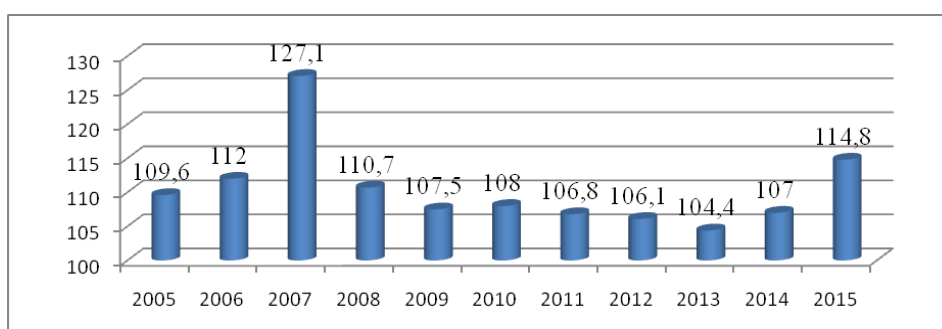


Fig. 12.7. The level of the price index in Almaty

Source: [6]

Economic factors, such as additional aspects of health care, transport and housing, can be attributed to factors [5]. The quality of life in the city, the natural and climatic conditions, the quality of life in terms of living standards, housing, health, transport, education, culture, the environment are factors in the development of the urbanization process.

Table 12.5. Health of Almaty city, registered for the diagnosis of diseases, 1000 per capita

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	695.9	624.	781.3	778.5	757.0	733.5	780.1	752.7	753.6	725.2	728.8

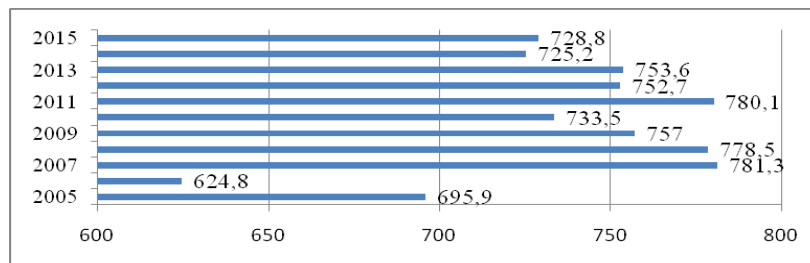


Fig. 12.8. Health of Almaty city

Source: [6]

Table 12.6. Provision of housing per capita

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	16.7	18.4	18.5	18.6	19.1	19.3	19.5	20.3	20.8	24.4	28.0

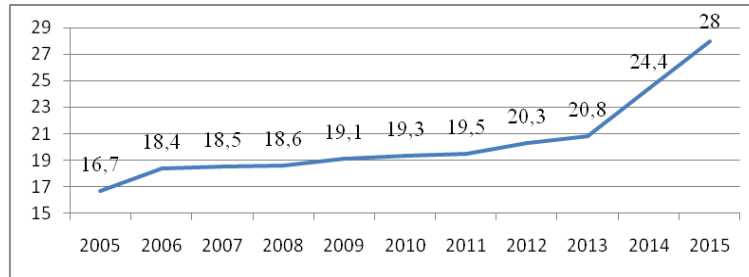


Fig. 12.9. The level of housing in Almaty

Source: [6]

Table 12.7. Freight turnover of transport, mln.ton

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	5 843.6	6 260.9	7 418.9	8 024.8	8 212.0	8 591.3	11 476.8	16 130.5	17 957.6	18 928.4	21 780.3

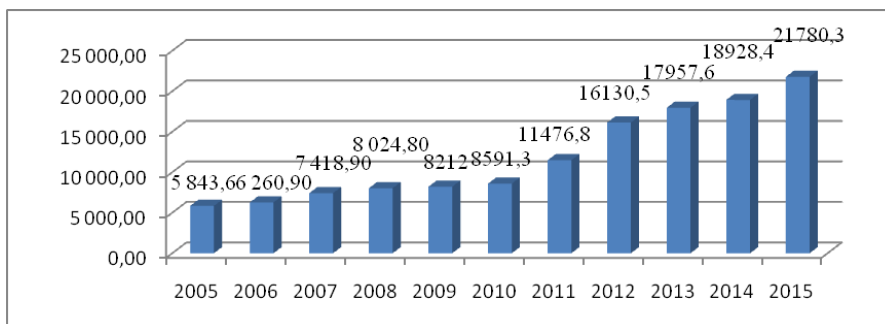


Fig. 12.10. Freight turnover of transport in Almaty

Source: [6]

Today, environmental factors of overcrowding in urban areas, the population of the city and suburban areas determine the level of industrial development. In particular,

the process of urbanization is properly carrying out environmental pollution, which leads to environmental problems.

Table 12.8. The volume of harmful substances in the atmosphere, thousands of tons

Years	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Almaty	6	8	8.7	10.1	10.6	18.2	11.0	11.7	12.1	12.4	43.5

Mobility in urban and rural residents is the main factor affecting the strengthening of communication. By combining economic, environmental, and social sustainability, cities will become equitable, resilient, and more appealing than urban sprawl that overuses land, promotes automobile use, and segregates the population economically [6].

12.4. CREATION OF A MATHEMATICAL MODEL OF THE URBANIZATION PROCESS

Requirements and stages of creating an urbanization process model

Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable). The variables we are using to predict the value of the dependent variable are called the independent variables.

Many techniques for carrying out regression analysis have been developed. Familiar methods such as linear regression and ordinary least squares regression are parametric, in that the regression function is defined in terms of a finite number of

unknown parameters that are estimated from the data. Nonparametric regression refers to techniques that allow the regression function to lie in a specified set of functions, which may be infinite-dimensional.

In general, the multiple regression equation of Y on X_1, X_2, \dots, X_k is given by:

$$y = \alpha' + \beta_1' x_1 + \beta_2' x_2 + \dots + \beta_p' x_p + \varepsilon, \quad (12.1)$$

here b_0 is the intercept and $b_1, b_2, b_3, \dots, b_k$ are analogous to the slope in linear regression equation and are also called regression coefficients. They can be interpreted the same way as slope. Thus if $b_i = 2.5$, it would indicate that Y will increase by 2.5 units if X_i increased by 1 unit. Determining the parameters of the regression equation is the most widely used method of least squares. The purpose of the regression equation by the γ method and the values of the dependent variable γ is to minimize the sum of the root-mean-square deviation. $\alpha', \beta_1', \beta_2', \dots, \beta_p'$ the parameters are random variables, the regression equation, that is, the process model is not immediately apparent [9]. Therefore, instead of the theoretical regression equation, the empirical regression equation is used:

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_p x_p + e, \quad (12.2)$$

$a, b_1, b_2, \dots, b_p - \alpha', \beta_1', \beta_2', \dots, \beta_p'$ - the empirical or theoretical value of the regression coefficients of the regression equation, e - is a measure of the fluctuations ε . The expression can be computed:

$$y = a + b_1 x_1 + b_2 x_2 + \dots + b_p x_p. \quad (12.3)$$

The parameters of the regression equation to determine the number of parameters must be greater than the number of observations and must describe the formula $n \geq p + 1$.

Independent variables $n \times (p+1)$ are transmitted in the form:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}. \quad (12.4)$$

Empirical equation:

$$Y = XB + e, \quad (12.5)$$

where: B –vector parameters, X –set of values of independent variables. e - random deviation vector:

$$e = Y - XB. \quad (12.6)$$

In accordance with the method of least squares $Q = \sum e_i^2$:

$$Q = e'e = (Y - XB)'(Y - XB). \quad (12.7)$$

In accordance with the method of least squares Q and B vectors is expressed as follows:

$$\frac{\partial Q}{\partial B} = -2X'Y + 2(X'X)B, \quad (12.8)$$

(12.8) gives the expression defines the parameters of the regression equation for the vector:

$$B = (X'X)^{-1} X'Y. \quad (12.9)$$

Multiple linear regression analysis of the empirical equation. It is determined by the importance of the quality of the parameters of the statistical regression equation.

In order to check the quality of the regression equation, the determination coefficient R^2 is used, the ratio is calculated as follows:

$$R^2 = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2}. \quad (12.10)$$

The value of the coefficient is from 0 to 1, each value is closer to 1 indicates the importance of the equation y.

The statistical significance of the coefficient of determination, that is, F is the statistic calculated by the following formula:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - p - 1}{p} \quad (12.11)$$

if $F = 0$, then the coefficient of determination $R_2 = 0$ equals the regression will be decisive. With $F_{table}(\alpha; p; n-p-1)$ the null hypothesis is verified. If the table has the value $F > F_{table}$, so the statistical significance of the coefficient R_2 is [7].

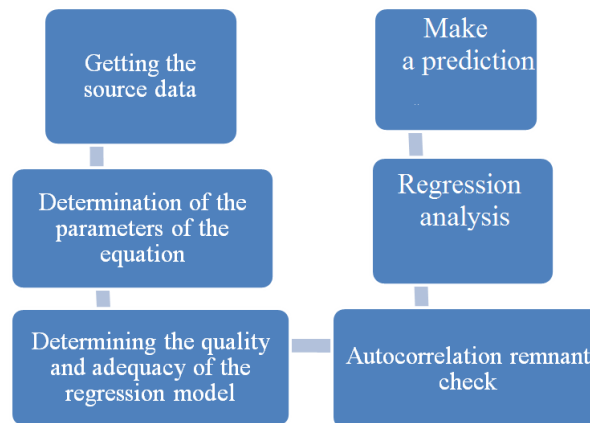


Fig. 12.11. Stages of multifactorial regression analysis

Source: own work

All major statistical software packages perform least squares regression analysis and inference. Simple linear regression and multiple regression using least squares can be done in some spreadsheet applications and on some calculators. While many statistical software packages can perform various types of nonparametric and robust regression, these methods are less standardized; different software packages implement different methods, and a method with a given name may be implemented differently in different packages. Specialized regression software has been developed for use in fields such as survey analysis and neuroimaging.

Mathematical model of the urbanization process

Econometric regression analysis is carried out in several stages:

During the **initial data preparation phase** (Tab. 12.9) the following symbols are used:

- n – count of experimental data between 2005 and 2015 years ;
- Y – demographic indicators in Almaty between 2005 and 2015 years;
- X₁ - the standard of living of the population in Almaty, tenge
- X₂ - unemployment rate, percentage;
- X₃ - consumer price index, percentage;
- X₄ -registered for the diagnosis of diseases, 1000 per capita;
- X₅ - provision of housing per capita;
- X₆ - freight turnover of transport, mln.ton;
- X₇ - the volume of harmful substances in the atmosphere, thousands of tons.

Table 12.9. Table of performance indicators and factors

N	Y (tho, people)	X ₁ (tg)	X ₂ (%)	X ₃ (%)	X ₄ (1000)	X ₅ (sq.m)	X ₆ (mln ton)	X ₇ (tho, ton)
2005	1247.9	39614	8.8	109.6	695.9	16.7	5843.6	6
2006	1278.3	49201	8.4	112	624.8	18.4	6260.9	8
2007	1324.7	59240	8.2	127.1	781.3	18.5	7418.9	8.7
2008	1361.9	78021	7.8	110.7	778.5	18.6	8024.8	10.1
2009	1390.7	90239	7.4	107.5	757	19.1	8212	10.6
2010	1413.5	95139	7.7	108	733.5	19.3	8591.3	18.2
2011	1449.8	106597	6.3	106.8	780.1	19.5	11476.8	11
2012	1475.4	121674	5.6	106.1	752.7	20.3	16130.5	11.7
2013	1507.5	134378	5.6	104.4	753.6	20.8	17957.6	12.1
2014	1642.3	142310	5.6	107	725.2	24.4	18928.4	12.4
2015	1703.5	155242	5.5	114.8	728.8	28	21780.3	43.5

Determining the parameters of the equation by using the method of least squares.The vector of the regression parameters is calculated by the formula

$$B = (X'X)^{-1}X'Y.$$

The regression equation is:

$$y = 496,43 + 0,0019x_1 + 12,02x_2 - 1,08x_3 + 0,23x_4 + 31,8x_5 - 0,00053x_6 - 1,89x_7 \quad (12.12)$$

The coefficients of the regression equation X_3 , X_5 and X_7 are negative (see (12.12)). As well as factor X_2 is greater than the value of the standard error rate results [8]. The value of the remaining three factors is derived from the regression analysis .

Table 12.10. Table of Source Data

N	Y(tho, people)	X1(tg)	X4(1000)	X5(sq. m.)
2005	1247.9	39614	695.9	16.7
2006	1278.3	49201	624.8	18.4
2007	1324.7	59240	781.3	18.5
2008	1361.9	78021	778.5	18.6
2009	1390.7	90239	757	19.1
2010	1413.5	95139	733.5	19.3
2011	1449.8	106597	780.1	19.5
2012	1475.4	121674	752.7	20.3
2013	1507.5	134378	753.6	20.8
2014	1642.3	142310	725.2	24.4
2015	1703.5	155242	728.8	28

The regression equation and the equation are determined by the results of an analysis based on the forecast.

Table 12.11. Table of performance indicators and factors

Years	Demographic indicator	Average monthly volume	Volume of housing provision
N	Y	X1	X5
	(tho, people)	(tg)	(sq, m)
2005	1247,9	39614	16,7
2006	1278,3	49201	18,4
2007	1324,7	59240	18,5
2008	1361,9	78021	18,6
2009	1390,7	90239	19,1
2010	1413,5	95139	19,3
2011	1449,8	106597	19,5
2012	1475,4	121674	20,3
2013	1507,5	134378	20,8
2014	1642,3	142310	24,4
2015	1703,5	155242	28

The results of the regression analysis show the factors had positive coefficients, Therefore, the factors included in the parameters of the regression equation, but the factor X_4 is greater than the value of the standard error rate. The values of X_1 and X_5 as a result of the analysis of factors included in the regression equation [6].

The period of estimating the parameters of the equation by means of the method of least squares. The vector of the regression parameters is calculated by the formula

$$B = (X'X)^{-1}X'Y.$$

$$X = \begin{pmatrix} 1 & 39614 & 16.7 \\ 1 & 49201 & 18.4 \\ 1 & 59240 & 18.5 \\ 1 & 78021 & 18.6 \\ 1 & 90239 & 19.1 \\ 1 & 95139 & 19.3 \\ 1 & 106597 & 19.5 \\ 1 & 121674 & 20.3 \\ 1 & 134378 & 20.8 \\ 1 & 142310 & 24.4 \\ 1 & 155242 & 28 \end{pmatrix}, \quad Y = \begin{pmatrix} 1247.9 \\ 1278.3 \\ 1324.7 \\ 1361.9 \\ 1390.7 \\ 1413.5 \\ 1449.8 \\ 1475.4 \\ 1507.5 \\ 1642.3 \\ 1703.5 \end{pmatrix}.$$

The matrix X consists of the value of each factor that affects to the process.

$$X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 39614 & 49201 & 59240 & 78021 & 90239 & 95139 & 106597 & 121674 & 134378 & 142310 & 155242 \\ 16.7 & 18.4 & 18.5 & 18.6 & 19.1 & 19.3 & 19.5 & 20.3 & 20.8 & 24.4 & 28 \end{pmatrix}$$

$X^T X$ matrix is determined by determining the average of the quadratic deviation

$$X^T X = \begin{pmatrix} 11 & 1071655 & 223,6 \\ 1071655 & 119358312113 & 22836556,5 \\ 223,6 & 22836556,5 & 4647,3 \end{pmatrix}$$

The value of the coefficient vector of the regression equation is calculated by:

$$Y(X) = (X^T X)^{-1} X^T Y = \begin{pmatrix} 812,001 \\ 0,000745 \\ 20,694 \end{pmatrix}.$$

The regression equation is given in (12.13):

$$y = 812,001 + 0,000745x_1 + 20,694x_5. \quad (12.13)$$

12.5. IMPLEMENTATION OF THE INFORMATION SYSTEM FOR MODELING URBANIZATION PROCESSES

A programming tool or software development tool is a computer program that software developers use to create, debug, maintain, or otherwise support other programs and applications.

Table data processing will be divided into three groups of processes in our work:

- Regulation, automatic filter, advanced filter to generate data processing;
- Solution of optimal problems;
- Statistical processing of data, analysis and forecasting.

In our work we used the following software tools:

- PHP is a server-side scripting language designed primarily for web development, but also used as a general-purpose programming language.
- MySQL is an open-source relational database management system. MySQL Fabric is an integrated system for managing a collection of MySQL servers, and a framework on top of which high availability and database sharding is built. MySQL Fabric is open-source, and supports procedure execution in the presence of failure, providing an execution model usually called resilient execution. MySQL client libraries are extended so they are hiding the complexities of handling failover in the event of a server failure, as well as correctly dispatching transactions to the shards.
- AJAX is a set of Web development techniques using many Web technologies on the client side to create asynchronous Web applications. With Ajax, Web applications can send data to and retrieve from a server asynchronously (in the background) without interfering with the display and behavior of the existing page.
- Robot Parser is the process of analysing a string of symbols, either in natural language or in computer languages, conforming to the rules of a formal grammar.

The traditional grammatical exercise of parsing, sometimes known as clause analysis, involves breaking down a text into its component parts of speech with an explanation of the form, function, and syntactic relationship of each part. This is determined in large part from study of the language's conjugations and declensions, which can be quite intricate for heavily inflected languages. To parse a phrase such as 'man bites dog' involves noting that the singular noun 'man' is the subject of the sentence, the verb 'bites' is the third person singular of the present tense of the verb 'to bite', and the singular noun 'dog' is the object of the sentence. Techniques such as sentence diagrams are sometimes used to indicate relation between elements in the sentence [7].

Structure of the information system

The information system operation scheme is shown in Fig. 12.12.

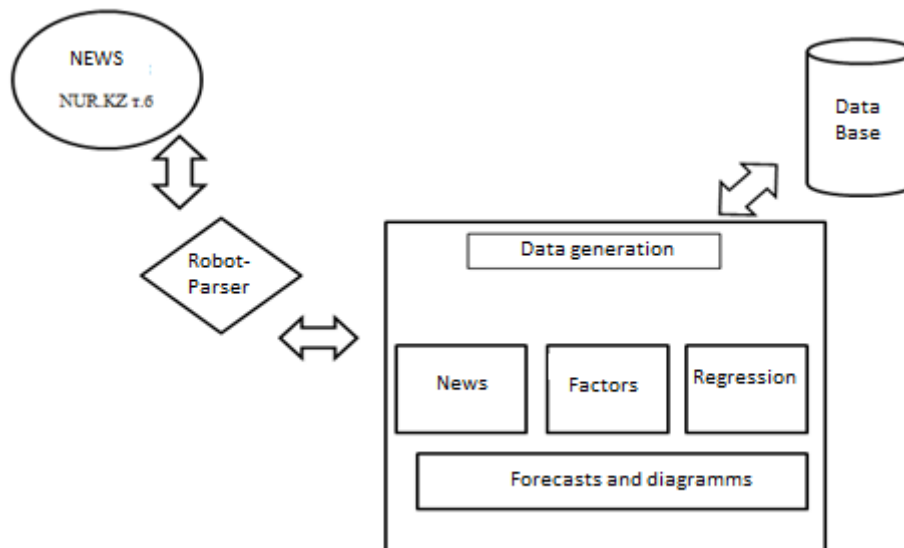


Fig. 12.12. Information system operation scheme

Source: own work

Data processing consists of four operations. Information on factors that affect the processing of data in the process of urbanization factors is the processing of tabular data and diagrams.

Regression data processing is an important part of the system, news processing is necessary for forecasting. These data are carried out with the help of robot parsing. For more information about database factors, the parsing table of data consists sets of data, in the form of news. The information obtained from the syntactic analysis of the factor is processed and used to predict the factors identified in the assessment. Processing technology, the architecture is shown in Fig. 12.13.

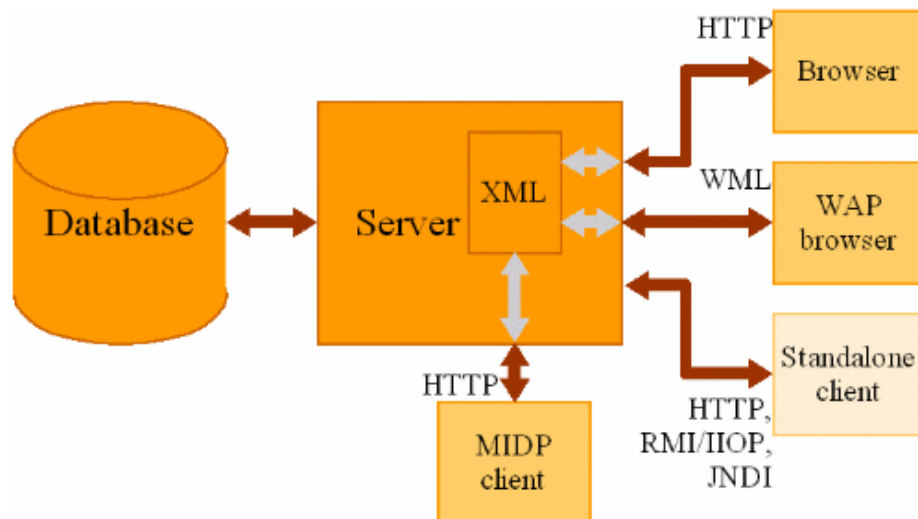


Fig. 12.13. Syntactic analysis of architecture

Source: [12]

The syntactic architecture for other sites using the HTTP protocol, using the interface of the JNDI application and the WML data of the minimum unit, to send information to the server [13].

The database of the system consists of the relationships between the main five tables. The database model is shown in Fig. 12.14:

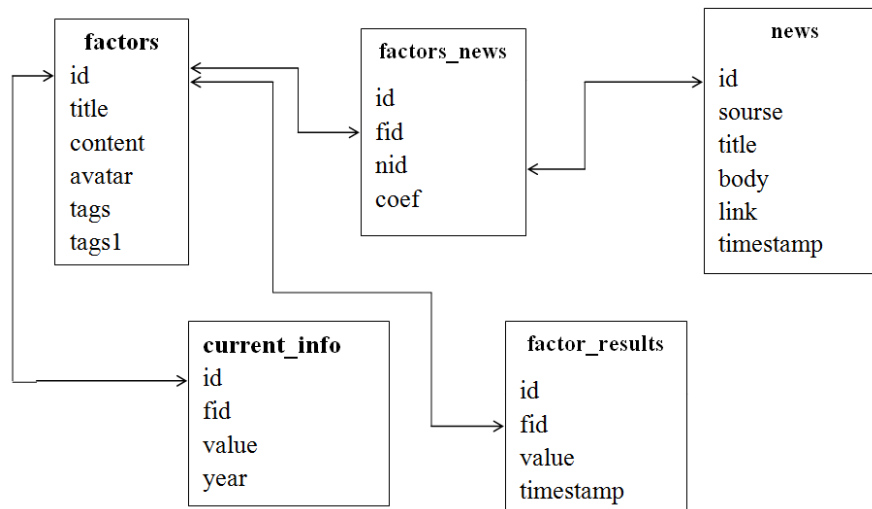


Fig. 12.14. The database schema

Source: own work

The table in the database contains information about the factors that influence of urbanization process. The **current_info** table in the database contains information about factors.

Web-based information system interface

An information system is an organized system for the collection, organization, storage and communication of information. More specifically, it is the study of complementary networks that people and organizations use to collect, filter, process, create and distribute data. The information system of the urbanization process model was created in PHP, in the text editor PphStorm [12].

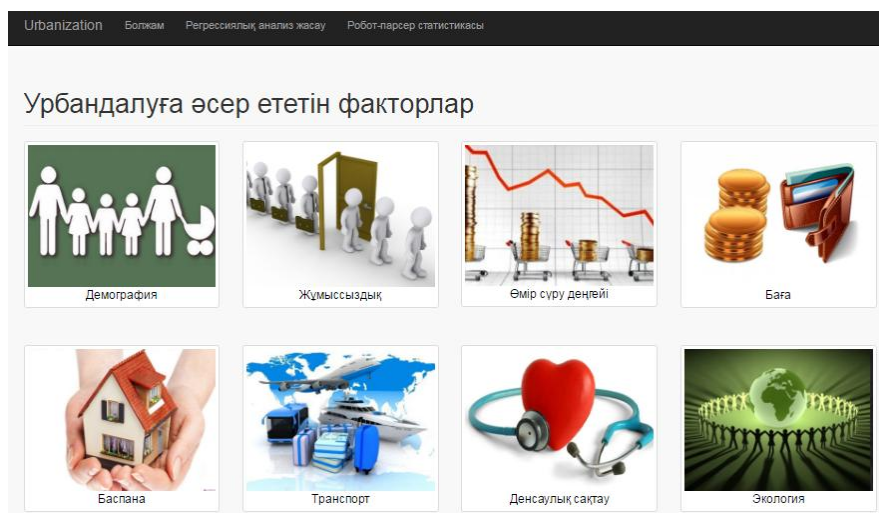


Fig. 12.15. Main page of the system

Source: own work

The forecast of information is in the main function of the system menu. Regression, correlation analysis is carried out in this menu. The analysis is carried out on the basis of a mathematical model.

№ (t)	Бақылау саны	Демографиялық көрсеткіш	Орташа айлық көлемі	Жұмыссыздық деңгейі	Баға индексі	Тркіелген вүру адам саны	Баспанамен қамтамасыз ету	Жұқ айналым	Зиянды заттар
		(мың адам)	(тг)	(%)	(%)	1000 адамға шаққанда	шаршы метр	млн. тонна	мың тонна
		Y	X1	X2	X3	X4	X5	X6	X7
1	2005	1247,9	39614	8,8	109,6	695,9	16,7	5843,6	6
2	2006	1278,3	49201	8,4	112	624,8	18,4	6260,9	8
3	2007	1324,7	59240	8,2	127,1	781,3	18,5	7418,9	8,7
4	2008	1361,9	78021	7,8	110,7	778,5	18,6	8024,8	10,1
5	2009	1390,7	90239	7,4	107,5	757	19,1	8212	10,6
6	2010	1413,5	95139	7,7	108	733,5	19,3	8591,3	18,2
7	2011	1449,8	106597	6,3	106,8	780,1	19,5	11476,8	11
8	2012	1475,4	121674	5,6	106,1	752,7	20,3	16130,5	11,7
9	2013	1507,5	134378	5,6	104,4	753,6	20,8	17957,6	12,1
10	2014	1642,3	142310	5,6	107	725,2	24,4	18928,4	12,4
11	2015	1703,5	155242	5,5	114,8	728,8	28	21780,3	43,5
12	2016	1716,7	168724	5,3	123,3	732,9	31,5	21845,6409	47,2

Fig. 12.16. Analysis window

Source: own work

By pressing the button, performing the regression analysis calculates a complete regression analysis [13]. The final indicator of the correlation analysis determines factors and factor relations.

1. Факторы 2. Корреляция 3. Факторы №1 4. Факторы №2 5. Автокорреляция 6. Прогноз								
	Y	X1	X2	X3	X4	X5	X6	X7
Y	1							
X1	0.96502	1						
X2	-0.8983	-0.963	1					
X3	-0.2208	-0.3726	0.41247	1				
X4	0.22095	0.31292	-0.3153	0.08273	1			
X5	0.94981	0.85188	-0.7606	-0.025	0.03591	1		
X6	0.94928	0.95072	-0.9451	-0.2632	0.16923	0.88899	1	
X7	0.72965	0.63332	-0.4882	0.11451	0.0219	0.8448	0.6417	1

Fig. 12.17. Results of the correlation analysis

Source: own work

The regression analysis developed by the program after creation the final regression equation. The regression equation is shown in Fig. 12.18:

Уравнения регрессии:
 $Y = 812.002402473 + 0.00208702574461 * X1 + 20.6927638572 * X2$

	Y	X1	X2
Коэффициенты	812.002402473	0.00208702574461	20.6927638572
Стандартная ошибка	41.0049923188	0.000238571026199	2.88696201831
t-статистика	19.8025254135	8.7480268575	7.16766058089
P-значение	4.40426151105E-8	2.2822227735E-5	9.54237040386E-5

Fig. 12.18. Regression equation

Source: own work

The program created as a result of the regression analysis model performs the forecasting function based on 2016 [8].

№(t)	Бақылау саны № жыл	Демографиялық көрсеткіш (мың адам) Y	Орташа айлық көлемі (тг) X1	Баспанамен қамтамасыз ету шаршы метр X5	Болжамды Y*	Ауытқу Ерә.
1	2005	1247.9	39614	16.7	1228.295455	19.6045455
2	2006	1278.3	49201	18.4	1269.827273	8.47272727
3	2007	1324.7	59240	18.5	1322.32123	2.37877
4	2008	1361.9	78021	18.6	1352.928	8.972
5	2009	1390.7	90239	19.1	1394.46	-3.76
6	2010	1413.5	95139	19.3	1416.21424	-2.71424
7	2011	1449.8	106597	19.5	1452.72214	-2.92214
8	2012	1475.4	121674	20.3	1482.5146	-7.1146
9	2013	1507.5	134378	20.8	1511.3104	-3.8104
10	2014	1642.3	142310	24.4	1641.2657	1.0343
11	2015	1703.5	155242	28	1700.5104	2.9896
12	2016				1724.2416	-
13	2017				1726.677273	-

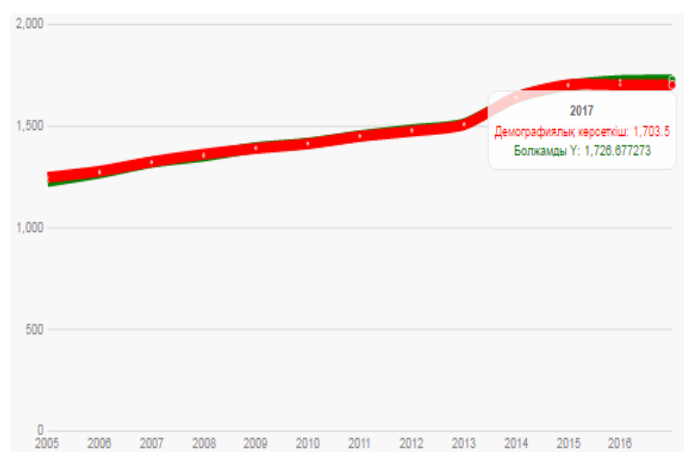


Fig. 12.19. Forecast based on formula PREDICT

Source: own work

The program allows you to automatically create a regression analysis. If you need a number of factors and changes in the number of checks, the program automatically creates a regression and correlation analysis.

Регрессиялық анализ жасау

Y:

X0:

X1:

X2:

X3:

X4:

Fig. 12.20. Automatic regression analysis

Source: own work

Робот-парсер статистикасы

[Статистиканы жаңарту](#)

Деректер қорындағы жалпы жаңалықтар саны: 290

Урбанизация факторларына әсер ететін жаңалықтар саны: 2

Show entries

Жаңалық	Фактор	Уақыт	Әрекет
В рамках НаскДау-2016 в Алматы выбраны 10 лучших интернет-проектов...	Өмір сүру деңгейі	2016-05-03 19:37:00	Көру
Жители подтопленных районов Алматы недовольны помощью акимата...	Баспана	2016-05-03 21:41:48	Көру

Showing 1 to 2 of 2 entries

Fig. 12.21. Robot-parser's statistics

Source: own work

In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called *linear models*. The advantage of PHP programming language is multi-functioning, one of which is the technology of the robot parser. Due to the use of technology, the function processes information from other sites. Parsing algorithms for

natural language cannot rely on the grammar having 'nice' properties as with manually designed grammars for programming languages. As mentioned earlier some grammar formalisms are very difficult to parse computationally; in general, even if the desired structure is not context-free, some kind of context-free approximation to the grammar is used to perform a first pass.

12.6. CONCLUSIONS

Urbanization is to organize the gradual and continuous transformation of information system to simplify it, to optimize its added value and to make it more responsive and flexible towards strategic business changes, while relying on technical opportunities of the market. Urban planning defines rules and a coherent, stable and modular context, in which different stakeholders are referring to any investment decision in the Information System.

We created Informational model of the urbanization process in this work. The urbanization mathematical model was based on regression analysis. The coefficients of the model determined by the method of least squares. The significance of the specified factors and coefficients were tested using Fisher and Student criteria. Demographic indicators and calculated values of the autocorrelation of Darbin Watson's criteria were determined on the basis of the forecast using the linear trend model.

Enterprises are now living in a fast changing business environment, subject to commercial, financial and technical evolutions that can occurs in only one or two years. So, this information system is useful for use by enterprises that want to implement long-term projects in Almaty. The new urban Multipolar paradigm provides an interesting metaphor to help us moving from a function-based static urbanization to a flow-based dynamic one.

REFERENCES

- [1] Akhiezer A.C., Kogan L., and Yanitsky O. (1969) *Urbanization, society and the scientific and technological revolution*. 2, pp. 41-44.
Demidenko E.S. (1992) *Urbanization: the concept and policy of urban development*. Moscow.
- [2] Pivovarov Yu.L. (1999) *Basics of geourbanistics: Urbanization and urban systems*. Moscow.
- [3] *The national composition of the population of the Republic of Kazakhstan (2009)*. 1, Population Censuses in the Republic of Kazakhstan.
- [4] *The proportion of the urban population in the total number of the population (1011)* The United Nations Department of Economic and Social Affairs.
- [5] *Site of statistical data of the city of Almaty (2017)* <http://almaty.stat.gov.kz/index.php/en> –
- [6] *Econometrics for beginners. Basic concepts, elementary methods, limits of applicability, interpretation of results (2000)*. Nosko, Moscow, pp. 210-216.
- [7] Draper N. and Smith G.(1996) *Applied regression analysis*. Economic analysis: theory and practice, 2, pp 60-63.
- [8] Shashkov V. (2003) *Applied regression analysis (multifactorial regression)*. Orenburg, pp. 363-65.
- [9] Seber J. (2003) *Linear Regression Analysis in Economic Analysis*. Sankt-Petersburg.
- [10] Gaidyshev I. (2010) *Analysis and data processing*. Fundamentals of Econometric, pp. 650-710.
- [11] Welling L. and Thomson L. (2008) *Web application development using PHP and MySQL*. Publishing house "Williams.
- [12] Lomov A. (2006) *HTML, CSS, scripts: the practice of creating websites*. BHV-Petersburg, pp. 71-72.
- [13] Nixon R. (2011) *Create dynamic websites using PHP, MySQL and JavaScript*. Peter, pp. 230- 243

- [14] Ross V. (2010) *Creating websites: HTML, CSS, PHP, MySQL. Textbook, Part 1*. Moscow, pp. 93-103.
- [15] *Robot parser technology* (2017) <https://pymotw.com/2/robotparser/>