Regional investment climate estimations with application of multidimensional statistical analysis

Aleksandr Izotov, and Olga Rostova

***Abstract*—** The investment decisions are one the most difficult when viewed from the selection procedure. They are based on the multivariate and multi-criteria estimations of many factors and trends, often having mixed dynamics. That's why the investment attractiveness estimation for the territory is the most important aspect of investment decision making. The accuracy of estimation affects the consequences for the investor and for targeted the economic system.

The investment decisions should be based on professional estimations of investment climate in countries and regions. In this connection the article solves the problem of selection of the most indicative factors for the purpose of estimating of investment climate and application of multidimensional statistical methods for prompt classification of the regions.

***Keywords*—** multidimensional statistical methods, main components method, discriminant analysis, investment climate, regional rating, factors of investment attractiveness.

I. Introduction

THE changes induced by the globalizing processes in the economy make the competition between the federal subjects stronger, which is also valid on the regional level. Here one of the main problems of effective regional development is limited investment resources needed to achieve its strategic goals and tasks. The scarcity of resources makes its negative impact on the economic growth and won't let the desired structural changes in the economy. The influence of foreign direct investments on the regional economics is substantial [1], [2]. The most important are the following advantages: increase of investments into fixed assets, implementation of modern industrial technologies and managerial methods, stimulating of national economic growth. The investments in the social sphere enable to improve the level of workforce training and to increase the number of high-qualified specialists [3].

The important component of regional investment activity is support of innovation processes. As it is pointed out in [4] innovations in the modern world serve as the basis for effective and expedient economic growth. It enables gaining from effective competition for resources and markets, technological improvements, implementation of new organizational decisions. To the large degree it concerns the regional economics that have become autonomous in the investment aspect, which entails also responsibility.

While competing for the resources positive credit (investment) ratings of various rating agencies play a significant role. Much research have been done and many research papers was dedicated to the investment climate appraisals [5], [6]. The performed analysis of the references allows to make up the comparative characteristics of the investment attractiveness appraisal methods.

A. Izotov is with Information Systems in Economics and Management Department, Institute of Industrial Economics and Management, St. Petersburg Polytechnical University, Politekhnicheskaya 29, 195251, Saint-Petersburg, Russia, (e-mail: izotovs@gmail.com).

O. Rostova is with Information Systems in Economics and Management Department, Institute of Industrial Economics and Management, St. Petersburg Polytechnical University, Politekhnicheskaya 29, 195251, Saint-Petersburg, Russia, (e-mail o.rostova_isem@mail.ru).

II. Problem Formulation

A. Validity of ratings

Recently in the Russian mass media the question of a not-unbiased ratings attributed by the leading rating agencies has widely been discussed. The initiative of creating of our own international rating agencies has become a matter of concern. The task of gaining confidence on the international level is a tough one. The international rating agencies have been earning their good name for decades.

One of the methods of testing the adequacy of the existing ratings and determining the factors exerting the major influence on the investment attractiveness of the region is discriminant analysis. It allows to find out the most significant factors of some given rating agency. The discriminant analysis can be applied only for factors that can be quantified. This poses some limitations, because one can't consider, say, the changes of political environment with the method. However, the quantifiable indicators as a rule make a substantial contribution to the final result.

B. Finding the set of significant factors

In the economic literature one can find many methods of estimation of the investment attractiveness of the region. This methods vary depending on the goals of analysis according to the number of analyzed indicators and their qualitative characteristics. Some researchers suggest using up to 200 of various indicators [7]. Many of the indicators are interrelated

and, consequently, are duplicating each other. Moreover, the significance level of each indicator is pretty much a subjective estimation. According to this, the task is to examine the possibilities of discriminant analysis application for classification of regions of the Russian Federation in the investment attractiveness aspect and to determine the most indicative factors.

C. Polysemantic grouping of factors and subjective significance level

In the majority of the examined methodical approaches for estimating the investment attractiveness the indices estimating the influence of factors and indicators are used. But frequently the investment climate estimations employ indicators that cannot be uniquely attributed to this or that group. Besides the calculations of the aggregate index imply the subjective selection of the significance level based on the expert judgements. It is suggested that the main components are used instead of indices. The main components being as summarized factors do not cause the decrease of the informative value.

III. The Application of Discriminant Analysis for Identifying of the Indicative Factors

The first stage of research is determining the set of indicators influencing the investment attractiveness of the region. The analysis of various methods of investment attractiveness estimation has been performed [8]. It should be pointed out that the features of each method depend on different characteristics of investment climate given by the methods. The latter can be explained with the target of analysis as the guideline for the developers of the method. Each method has its own user, i.e. the investor that will make the calculations in order to make an investment decision. The goals of investment climate appraisals have stipulated the different approaches for information sources (statistical data, scientific research data or expert appraisals), for determining of the main factors and indicators, for the organization of the research itself, distinguishing between the directions and stages of research. As a result, the regional ratings from one research are not valid for decision making in other circumstances [9], [10].

The accomplished research allows to identify the groups of factors influencing the investment climate of the region:

- The factors of economic development;
- The factors of economic growth dynamics;
- The factors of the social sphere;
- The technological factors;
- The factors of the institutional sphere.

For comparability of the indicators between the regions the factors were calculated in per capita or in percents of the total format.

The regional ratings for the Russian Federation attributed by the international rating agencies Standard & Poor's, Moody's, Fitch and by the Russian rating agency Expert were usen in the research.

The discriminant analysis was carried out on the basis of more than 50 indicators of socio-economic environment of the Russian regions [11], [12]. The hierarchical structure of the groups of factors influencing the investment climate is shown on the Figure 1.

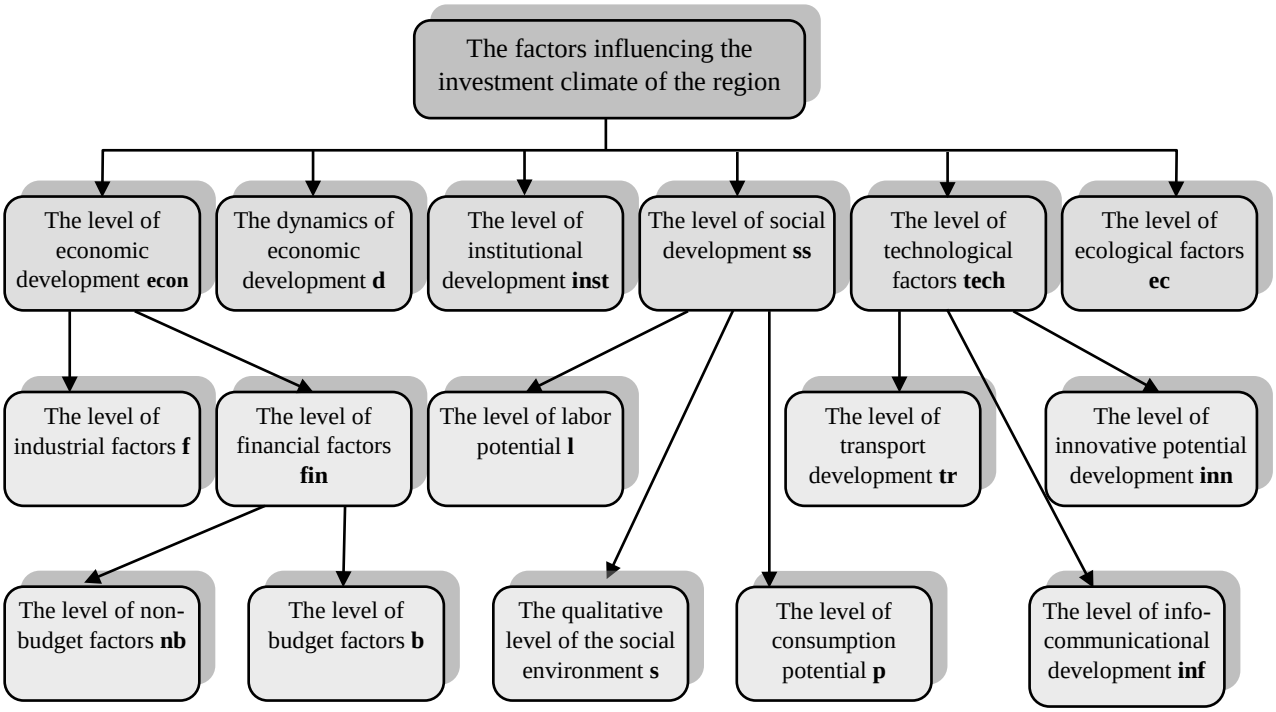


Fig. 1. The hierarchical structure of the regional investment climate factors

The list of indicators included into the models are outlined below:

Production factors:

f_2 – production of the manufacturing sector, thsd. RUR/pers.;
f_4 – production of the agriculture, thsd. RUR/pers.;
f_5 – mining output, thsd. RUR/pers.;
f_6 – fixed assets investments, thsd. RUR/pers.;
f_7 – fixed assets in the economy, thsd. RUR/pers.;
f_8 – regional inflation, %;
f_9 – building, thsd. RUR/pers.;
f_{10} – production of electricity, gas and water, thsd. RUR/pers.;

The economic growth dynamics factors:

d_2 – industrial production index, %;
d_5 – retail sales turnover index, in % to the previous year;

Non-budget factors:

nb_1 – total financial result (profit minus loss) of the organizational economic activity, thsd. RUR/pers.;
nb_2 – the share of companies with losses, %;
nb_3 – the share of the outstanding accounts payable to the total, %;
nb_4 – the share of the outstanding accounts receivable to the total, %;

The budget factors:

b_1 – the income of the regional budgets, thsd. RUR/pers.;

The factors of labor potential:

l_4 – the number of students per 10000 of population;

The factors of consumption:

p_1 - the regional subsistence wage, RUR/mnth.;
p_2 – the ratio of the average income to the subsistence wage, %;
p_3 – the average wage, RUR;
p_7 – retail sales turnover, thsd. RUR/pers.;

The factors of social environment:

s_1 – the unemployment level, %;
s_2 – the housing level, sq.m./pers.;

The factors of infrastructure:

tr_1 – the density of the roads, km./10000sq.km.;
tr_2 – the automobile cargo transportation, thsd. tn./pers.;
tr_3 – the density of the railroads, km./10000sq.km.;
tr_4 – the railroad cargo transportation, thsd. tn./pers.;

Information and communication factors:

inf_2 – the share of companies using LAN, %;
inf_3 – the share of companies using special software, %;
inf_4 – the share of companies using WAN, %;

Innovative factors:

inn_5 – the innovation activity of companies, %

Institutional factors:

$inst_1$ – number of organizations per 100000 pers.;
$inst_2$ – number of credit organizations per 100000 pers.;
$inst_4$ – number of organizations with foreign capital per 1000 pers.;
$inst_5$ – number of small enterprises per 10000 pers.

The analysis is being carried out with the help of STATISTICA using the stepwise discriminant analysis [13]. The method of stepwise inclusion of indicators was used (menu option – forward stepwise). The research of international ratings was performed with the figures of 2012. It should be stressed, that the discriminant analyses gives no opportunity to estimate the influence of factors on the final result. The results of the analysis help to discover the indicators enabling the correct classification of the objects between the groups.

The results below are obtained with the Standard & Poor's rating. Table 1 illustrates the results of the 5th step of the discriminant analysis. The discrimination of the regions is highly significant (Wilks' Lambda =0,0008; F=34; p<0,0000). With the 5% error probability all the variables within the model are statistically significant (the p-level column). The percentage of correct forecasts equals 100% (table 2). Involvement of new variables on the further steps of the discriminant analysis has led to appearance of insignificant factors and didn't improve the quality of classification.

Table 1. Discriminant Function Analysis Summary (Standard & Poor's) Step 5

N of vars in model: 5; Grouping: SP (4 grps) Wilks' Lambda: 0000008 approx. F (15,16)=33,969 p< 0.0000						
	Wilks' - Lambda	Partial - Lambda	F-remove - (3,6)	p-level	Toler.	1-Toler. - (R-Sqr.)
tr_1	0.0972	0.0008	2550.4	0.000000	0.015	0.985
tr_3	0.0063	0.0121	163.9	0.000004	0.008	0.992
d_5	0.0024	0.0315	61.5	0.000067	0.033	0.967
nb_4	0.0013	0.0546	34.7	0.000348	0.097	0.903
p_1	0.0004	0.1981	8.1	0.015683	0.093	0.907

The results of discriminant analysis enable to estimate the contribution of each factor to the total discrimrnation of the regions (table 1, column Partial Lambda). The less the value of Wilks' partial statistics the greater is the contribution of the indicator. In the table 1 the indicators are sorted according to their descending significance for the correct classification.

The most contribution is given by the infrastructure factors. The application of only this factors allows to classify correctly 78% of the regions. The next significant factors are trade dynamics and the share of the outstanding accounts receivable. Adding up of these factors improves the classification up to 98%. Table 3 gives the coefficients of the built discrimination functions.

Table 2. Classification Matrix (Standard & Poor's)

Rows: Observed classifications Columns: Predicted classifications					
	Percent - Correct	BBB p=0.143	BB p=0.357	BB+ p=0.357	B+ p=0.1429
BBB	100	2	0	0	0
BB	100	0	5	0	0
BB+	100	0	0	5	0
B+	100	0	0	0	2
Total	100	2	5	5	2

Table 3. Classification Functions; grouping: Standard & Poor's

	BBB p=0.1429	BB p=0.3571	BB+ p=0.3571	B+ p=0.1429
tr_1	-4.8	-10.3	-10.9	-11.3
tr_3	9.1	18.9	20.0	20.7
d_5	198.9	391.7	414.9	428.9
nb_4	105.1	210.1	223.1	231.1
p_1	0.2	0.4	0.5	0.5
Constant	-6449	-23025	-25828	-27610

The following discriminant functions are obtained:

For the group BBB

$$d_1 = -6449 - 5 \cdot tr_1 + 9 \cdot tr_3 + 199 \cdot d_5 + 105 \cdot nb_4 + 0.2 \cdot p_1$$

For the group BB

$$d_2 = -23025 - 10 \cdot tr_1 + 19 \cdot tr_3 + 392 \cdot d_5 + 210 \cdot nb_4 + 0.4 \cdot p_1$$

For the group BB+

$$d_3 = -25828 - 11 \cdot tr_1 + 20 \cdot tr_3 + 415 \cdot d_5 + 223 \cdot nb_4 + 0.5 \cdot p_1$$

For the group B+

$$d_4 = -27610 - 11 \cdot tr_1 + 21 \cdot tr_3 + 429 \cdot d_5 + 231 \cdot nb_4 + 0.5 \cdot p_1$$

Into the obtained discriminant functions one substitutes the values of the indicators and the object is being attributed to the group with the maximum value of the discriminant function. Using the obtained functions one can classify the other regions or forecast the changes of the regional rating with the changes of any indicator.

The results below are obtained with the Moody's rating. The model that enables correct classification is obtained on the 9th step of discriminant analysis (table 4). The discrimination of the regions is highly significant (Wilks' Lambda =0,000; F=124; p<0,0000). With the 5% error probability all the variables within the model are statistically significant (the p-level column). The percentage of correct forecasts equals 100% (table 5). Involvement of new variables on the further steps of the discriminant analysis has led to appearance of insignificant factors and didn't improve the quality of classification.

Table 4. Discriminant Function Analysis Summary (Moody's) Step 9

N of vars in model: 9; Grouping: Moodys (4 grps) Wilks' Lambda: 0.00000 approx. F (27,9)=123.58 p< 0.0000						
	Wilks' - Lambda	Partial - Lambda	F-remove - (3.3)	p-level	Toler.	1-Toler. - (R-Sqr.)
tr_1	0.001379	0.000026	38942.62	0.000000	0.000027	0.999973
tr_3	0.001057	0.000034	29841.20	0.000000	0.000031	0.999969
d_5	0.000079	0.000446	2242.59	0.000016	0.000371	0.999629
inf_4	0.000029	0.001217	820.89	0.000072	0.000080	0.999920
inf_3	0.000019	0.001841	542.26	0.000134	0.000184	0.999816
p_2	0.000024	0.001482	673.95	0.000097	0.000649	0.999351
f_4	0.000010	0.003642	273.59	0.000373	0.004134	0.995866
nb_2	0.000003	0.012259	80.57	0.002296	0.004204	0.995796
$inst_4$	0.000001	0.060448	15.54	0.024767	0.013582	0.986418

Table 5. Classification Matrix (Moody's)

Rows: Observed classifications Columns: Predicted classifications					
	Percent – Correct	Ba3 p=0.200	Baa1 p=0.133	Ba1 p=0.333	Ba2 p=0.333
Ba3	100	3	0	0	0
Baa1	100	0	2	0	0
Ba1	100	0	0	5	0
Ba2	100	0	0	0	5
Total	100	3	2	5	5

Let's analyze the significance of the indicators for the correct classification. The most contribution is given by the infrastructure factors. The usage of only these factors allows to classify correctly 60% of the regions. The next significant factor is trade dynamics. Adding up of this factor improves the classification up to 80%. The next significant factors are information and communication factors. Adding up of these factors improves the classification up to 93%. Table 6 gives the coefficients of the built discrimination functions.

On the next stage of the research the analysis is made based on the Fitch rating. The model that enables correct

classification is obtained on the 18th step of discriminant analysis (table 7).

Table 6. Classification Functions; grouping: Moody's

	Ba3 p=0.200	Baa1 p=0.133	Ba1 p=0.333	Ba2 p=0.333
tr_1	-2145	201	-2101	-2053
tr_3	2788	-260	2731	2669
d_5	21824	-2023	21376	20892
inf_4	45218	-4256	44293	43289
inf_3	-36241	3427	-35501	-34696
p_2	1532	-143	1501	1466
f_4	3405	-320	3336	3260
nb_2	10748	-994	10532	10291
$inst_4$	-243439	22614	-238729	-233101
Constant	-2136971	-20025	-2050454	-1958424

The discrimination of the regions is highly significant (Wilks' Lambda =0,000; F=11,9; p<0,0000). The usage of 16 indicators in the model has led to statistical insignificance of two of them (inst2, S2). However the exclusion of them from the model deteriorated the quality of classification. The percentage of correct forecasts equals 100% (table 8). Table 9 gives the coefficients of the built discriminant functions.

classification. The most contribution is given by the infrastructure factors. The next significant factors are consumption potential regional economic growth dynamics. These factors enable classification of 70% of the regions.

Table 7. Discriminant Function Analysis Summary (Fitch) Step 18

N of vars in model: 16; Grouping: Fitch (5 grps) Wilks' Lambda: 0.00000 approx. F (72,33)=11.933 p< 0.0000						
	Wilks' - Lambda	Partial - Lambda	F-remove - (4,8)	p-level	Toler.	1-Toler. - (R-Sqr.)
tr_1	0.000220	0.011688	169.1221	0.000000	0.033896	0.966104
tr_3	0.000024	0.108522	16.4294	0.000633	0.032274	0.967726
p_3	0.000023	0.110187	16.1510	0.000672	0.015443	0.984557
p_1	0.000017	0.148986	11.4241	0.002170	0.013766	0.986234
d_5	0.000012	0.213440	7.3703	0.008605	0.151093	0.848907
s_3	0.000055	0.046400	41.1032	0.000022	0.008836	0.991164
$inst_2$	0.000004	0.593265	1.3712	0.325420	0.205785	0.794215
f_{10}	0.000009	0.299546	4.6768	0.030609	0.053066	0.946935
f_6	0.000017	0.149790	11.3521	0.002215	0.043783	0.956217
inn_5	0.000010	0.248302	6.0547	0.015231	0.147960	0.852040
f_3	0.000039	0.065292	28.6316	0.000086	0.006758	0.993242
s_2	0.000006	0.395216	3.0605	0.083417	0.180648	0.819352
tr_4	0.000013	0.195994	8.2044	0.006221	0.071740	0.928260
d_3	0.000040	0.064572	28.9730	0.000082	0.019700	0.980300
nb_4	0.000018	0.144697	11.8220	0.001938	0.052423	0.947577
f_8	0.000015	0.175261	9.4115	0.004056	0.081494	0.918505

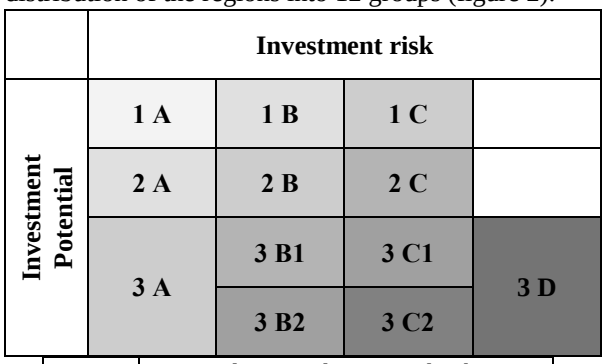
Table 8. Classification Matrix (Fitch)

Rows: Observed classifications Columns: Predicted classifications						
	Percent – Correct	B+	BB	BB-	BB+	BBB
B+	100	4	0	0	0	0
BB	100	0	7	0	0	0
BB-	100	0	0	9	0	0
BB+	100	0	0	0	8	0
BBB	100	0	0	0	0	2
Total	100	4	7	9	8	2

Table 9. Classification Functions; grouping: Fitch

	B+	BB	BB-	BB+	BBB
tr_1	-13	-13	-13	-13	-8
tr_3	-14	-15	-14	-14	-17
p_3	0	0	0	0	0
p_1	-1	-1	-1	-1	-1
d_5	-114	-118	-112	-113	-141
s_3	-10762	-10805	-10533	-10530	-9183
$inst_2$	24846	24232	23524	22548	9866
f_{10}	104	104	101	101	87
f_6	-29	-30	-29	-28	-29
inn_5	-124	-128	-121	-120	-148
f_3	31894	31973	31224	31250	26673
s_2	-206	-202	-202	-204	-109
tr_4	10161	10269	9938	9936	9387
d_3	-241	-242	-236	-237	-205
nb_4	-467	-469	-457	-458	-401
f_8	2714	2725	2665	2668	2395
Constant	-136443	-137119	-131699	-131831	-107177

Let's select the most significant indicators for correct In the method of Expert rating agency the investment climate is consisting of: the investment potential (the sum of the objective prerequisites for making effective investments, that are dependent on the existence and the variety of the investment objects); the investment risk (the probability of loosing either the investments or the interest on the investments). The rating actually is distribution of the regions into 12 groups (figure 2).



1A	Maximal potential – minimal risk
2A	Medium potential – minimal risk
3A	Low potential – minimal risk
1B	High potential - medium risk
2B	Medium potential – medium risk
3B1	Reduced potential – medium risk
3B2	Small potential – medium risk
1C	High potential – high risk
2C	Medium potential – high risk
3C1	Reduced potential – high risk
3C2	Small potential – high risk
3D	Low potential – extreme risk

Fig.2. Distribution of the regions into groups

According to the method, the regions, attributed to the 1A group, are the most preferable for investors, while the regions in the 3A group are the less desired investment targets. The classification enables the potential investors to focus their attention only on the regions that are satisfying the needs of the investor most of all when viewed from the expected risk and return [14].

The discriminant analysis was done with the application of the Russian agency Expert data for the period 2006-2012. The results for each year of this period were approximately the same. The differences were only in the percentage of the correctly predicted results, which were between 94% an 100%. Therefore we present only the results for the last step for the year 2012 (table 10):

The discrimination of the regions is significant (Wilks' Lambda =0,00422; F=3,98; p<0,0000). Not all the indicators involved in the model appeared to be significant. However the reduction of number of indicators from the model deteriorated the quality of classification. The percentage of correct forecasts equals 97% (table 11).

Table 10. Classification Matrix

	Percent	3B1	3B2	2A	3A1	3C2	2B	1A
3B1	100	34	0	0	0	0	0	0
3B2	80	2	8	0	0	0	0	0
2A	100	0	0	2	0	0	0	0
3A1	100	0	0	0	4	0	0	0
3C2	100	0	0	0	0	6	0	0
2B	100	0	0	0	0	0	9	0
1A	100	0	0	0	0	0	0	3
Total	97	36	8	2	4	6	9	3

Two regions (Amur and Pskov regions) are attributed to the group 3B1 by mistake. Let's illustrate the fragment of the table with the distances from these objects to the centers of the each group (table 12). We can see from the table that the distance from these objects to the centers of the groups 3B1 and 3B2 is almost the same. That leads to an incorrect classification. In the table 13 one can see the coefficients of the obtained discriminant functions.

Table 11. Discriminant Function Analysis Summary, Step 17

Step 17, N of vars in model: 17; Grouping: r (7 grps) Wilks' Lambda: 0.00422 approx. F (102,252)=3.9784 p<0.0000						
	Wilks' - Lambda	Partial - Lambda	F-remove - (6,43)	p-level	Toler.	1-Toler. - (R-Sqr.)
$inst_1$	0.013214	0.319418	15.26998	0.000000	0.491336	0.508664
f_4	0.009612	0.439082	9.15527	0.000002	0.321723	0.678277
tr_4	0.007620	0.553903	5.77182	0.000179	0.648170	0.351830
l_4	0.005335	0.791097	1.89248	0.104072	0.464392	0.535609
s_1	0.004945	0.853521	1.22993	0.310030	0.613314	0.386686
p_7	0.005740	0.735321	2.57965	0.031828	0.363983	0.636017
$inst_2$	0.006027	0.700337	3.06650	0.013770	0.449365	0.550635
nb_3	0.004876	0.865683	1.11196	0.371323	0.592620	0.407380
f_5	0.005986	0.705091	2.99751	0.015496	0.162681	0.837319
f_9	0.008316	0.507532	6.95397	0.000033	0.044788	0.955212
f_6	0.006729	0.627241	4.25903	0.001881	0.049192	0.950808
d_2	0.004885	0.863945	1.12861	0.362126	0.679495	0.320506
$inst_5$	0.005412	0.779895	2.02261	0.083286	0.512698	0.487302
nb_2	0.005403	0.781231	2.00689	0.085565	0.466530	0.533470
b_2	0.005011	0.842216	1.34263	0.259553	0.664757	0.335243
d_1	0.005434	0.776753	2.05978	0.078133	0.416595	0.583405
inn_1	0.005266	0.801504	1.77486	0.127107	0.540670	0.459330

Table 12. Squared Mahalanobis Distances from Group Centroids

Incorrect classifications are marked with *								
	Observed	3B1	3B2	2A	3A1	3C2	2B	1A
Novgorod region	3B2	17.6999	14.4514	45.0583	22.3191	23.3459	51.2070	131.1173
* Amur region	3B2	25.0270	27.6361	69.1655	61.2027	37.7799	40.4197	146.0429
Kostroma region	3B2	19.6512	10.8463	76.7756	55.1213	22.3769	62.2044	170.5927
* Pskov region	3B2	11.8408	13.3354	72.7242	45.3653	42.5260	42.5257	151.5080

Table 13. Classification Functions

	3B1	**3B2**	**2A**	**3A1**	**3C2**	**2B**	**1A**
$inst_1$	0.02	0.01	0.03	0.02	0.02	0.04	0.08
f_4	-3.64	-3.57	-2.89	-3.21	-3.34	-3.67	-3.08
tr_4	28.88	25.30	42.24	32.48	21.56	44.53	41.56
l_4	0.13	0.12	0.11	0.11	0.10	0.14	0.13
s_1	9.49	9.94	10.72	8.91	11.07	10.11	10.28
p_7	0.45	0.40	0.65	0.60	0.32	0.56	0.58
$inst_2$	-372.72	-194.98	-348.96	-709.21	191.98	-432.64	-682.59
nb_3	3.68	3.71	3.30	3.53	4.02	3.59	3.20
f_5	0.23	0.23	0.19	0.20	0.21	0.21	0.16
f_9	0.19	0.13	0.38	-0.05	0.37	0.40	1.36
f_6	-0.56	-0.55	-0.61	-0.40	-0.58	-0.64	-0.86
d_2	2.84	2.73	2.66	3.14	2.90	2.78	2.97
$inst_5$	0.17	0.16	0.23	0.22	0.13	0.17	0.07
nb_2	3.46	3.49	3.51	3.23	3.75	3.34	4.54
b_2	659.04	639.35	627.22	625.14	639.28	661.62	707.34
d_1	31.89	32.15	30.88	31.10	31.23	32.19	30.06
inn_1	-883.63	-897.63	-780.33	-811.92	-844.34	-871.05	-722.86
Constant	-2254.33	-2245.46	-2179.74	-2194.97	-2194.74	-2317.12	-2272.64

Let's enumerate the most significant indicators for correct classification (from high significance to the low): institutional, industrial, infrastructure factors and factors of labor and consumer potential. These factors enable classification of 81% of the regions.

IV. Building of the Aggregate Factors with the Usage of the Main Components Method

Another possible approach is preliminary selection of the main components which are used as the basis for the further discriminant analysis. Because of the large number of indicators used in the Fitch rating, it was decided to apply here the main components method.

Table 14 gives the eigenvalues and the percentage of the explained variance for the first 11 main components of the Fitch rating.

Table 14. The eigenvalues and the percentage of the explained variance of the main components.

Eigenvalues (2012 исход) Extraction: Principal components				
	Eigenvalue	**% Total - variance**	**Cumulative - Eigenvalue**	**Cumulative - %**
1	16,2	29,9	16,2	29,9
2	10,2	18,2	26,4	48,1
3	4,4	13,4	30,8	61,5
4	2,9	9,1	33,7	70,6
5	1,6	7,5	35,3	78,1
6	0,9	2,3	36,2	80,5
7	0,9	2,0	37,0	82,4
8	0,8	1,4	37,8	83,9
9	0,5	0,9	38,3	84,8
10	0,3	0,8	38,6	85,6
11	0,3	0,4	38,9	86,0

According to the Kaiser criterion one should retain the main components with eigenvalues greater than 1 (the Eigenvalue column). This criterion selects five factors that explain 78.1% of the variance.

In order to explain the selected main components we'll study their correlation with the initial indicators. Table 15 gives the fragment of the table acquired in "Statistica" software.

The first main component is most closely connected with the indicators:

f1- gross regional product, thsd.RUR/pers.

f9 – production of electricity, gas and water, thsd. RUR/pers.;

f5 – mining output, thsd. RUR/pers.;

f6 – fixed assets investments, thsd. RUR/pers.;

f7 – fixed assets in the economy, thsd. RUR/pers.;

nb1 – the receipts of the regional budget , thsd. RUR/pers.;

p3 – the average wage per month, RUR/pers.;

p4 – the average personal income, RUR.

Table 15. The correlation between the initial indicators and the main components

	Factor Loadings (Varimax raw) (2012 исход) Extraction: Principal components (Marked loadings are >,700000)				
	Factor - 1	**Factor - 2**	**Factor - 3**	**Factor - 4**	**Factor - 5**
f1	***0,934***	0,086	0,076	-0,009	-0,032
f9	***0,897***	0,098	0,075	0,117	0,015
f5	***0,910***	0,060	0,045	-0,054	-0,102
f6	***0,937***	0,073	0,024	0,076	-0,072
f7	***0,940***	-0,015	0,025	-0,134	-0,101
d1	-0,127	-0,089	-0,245	***0,829***	0,237
d6	-0,158	-0,133	-0,235	***0,749***	0,268
d3	0,014	-0,152	-0,203	***0,747***	-0,264
nb1	***0,863***	-0,028	0,001	-0,009	-0,094
nb2	-0,065	0,299	-0,130	-0,552	0,032
nb3	-0,140	-0,230	-0,168	-0,676	0,114
nb4	-0,122	-0,338	-0,044	***-0,713***	0,032
p3	***0,757***	0,028	0,114	-0,118	-0,124
p4	***0,718***	0,015	0,102	-0,041	0,134
p6	0,580	0,015	0,183	0,029	0,173
p7	0,522	-0,014	0,155	0,054	0,275
s2	-0,035	0,062	-0,049	0,063	***0,784***
tr1	-0,313	0,048	0,050	0,116	***0,766***
tr2	0,215	-0,100	0,056	-0,041	***0,713***
tr3	-0,254	0,266	0,011	0,180	***0,775***
inf2	0,243	-0,063	***0,832***	-0,064	0,042
inf3	0,068	0,054	***0,888***	-0,030	-0,031
inf4	0,064	-0,051	***0,902***	-0,113	0,011
inst1	0,103	***0,720***	0,065	0,095	0,100
inst4	0,078	***0,868***	-0,062	0,067	0,257
inst5	0,248	***0,726***	-0,036	-0,146	-0,110

Also there is interdependence (correlation of 0.58 and 0.52 correspondingly) with the indicators:

p6 – average personal consumption, RUR/month;

p7 – turnover of the retail sector, RUR/pers.

The first main component concerns the groups of industrial, non-budget and consumption factors. It explains about 30% of total variance.

The second main component is closely connected with the indicators within the group of institutional factors:

inst1 – number of organizations per 100000 pers.;

inst4 – number of organizations with foreign capital per 1000 pers.;

inst5 – number of small enterprises per 10000 pers.

We'll call it the institutional factor. It explains 18.2% of the total variance.

The third main component is tightly connected with three indicators of the info-communication group of factors:

inf2 – the share of companies using LAN, %;

inf3 – the share of companies using special software, %;

inf4 – the share of companies using WAN, %;

We'll call it the info-communicational component. It explains 13.4% of the total variance.

The forth main component is closely connected with the following indicators:

d1 – growth rate of gross regional product index, in % to the previous year;

d3 – growth rate of investments, in % to the previous year;

d6 – productivity of labor index, in % to the previous year;

nb4 – the share of the outstanding accounts receivable to the total, %;

The less tight interdependence can be seen with the indicators:

nb2 – the share of companies with losses, %;

nb3 – the share of the outstanding accounts payable to the total, %;

It explains 9.1% of the total variance. The strong linear relationship with the first three indicators, that are within the group characterizing the development dynamics of the region. The next three indicators from the group on non-budget factors have the opposite correlation. We'll call it development dynamics component.

The fifth component is closely connected with the following indicators:

s2 – the housing level, sq.m./pers.;

tr1 – the density of the roads, km./10000sq.km.;

tr2 – the automobile cargo transportation, thsd. tn./pers.;

tr3 – the density of the railroads, km./10000sq.km.;

The first of the indicators is attributed to the group of factors, characterizing the social environment. The next three indicators are from the group of transport factors. All of them can be classified as infrastructural factors, so we'll call it the infrastructure factor. It explains 7.5% of the total variance.

In "Statistica" software one can obtain the values of each of the selected main components for every region. Based on the main components the discriminant analysis was performed. The results correspond to those of the discriminant analysis on the initial data.

V. Conclusion

The procedure of searching for vast information sets is a time-taking one because the data needed for rating calculations is usually published with significant delays. The obtained discriminant functions allows to make an operative estimations of regional rating with the usage of several indicators. The research shows that the discriminant analysis enables pretty much precise classifications of the regions.

The most informative indicators for regional classification according to the methods of the international rating agencies appeared to be the infrastructure factors, the economic growth dynamics and the consumer potential factors. As for the method used in the Russian agency Expert the most informative indicators are institutional, industrial and infrastructure factors and the factors of labor and consumer potential. It can be explained with, the fact that each of the analyzed ratings is oriented on their users that are supposed to apply the rating while making their investment decisions.

The results of the fulfilled discriminant analysis enable to classify correctly the regions on the basis of the ratings of the international rating agencies which confirms the high quality of the research.

The main components method gives the five aggregate factors, which can be used for composing the aggregate indicator of the regional investment attractiveness and for classification of the regions with application of discriminant and cluster analysis.

References

[1] Jude C., Silaghi M.I.P. *Foreign direct investment, employment creation and economic growth in CEE countries. An open issue*, Proceedings of the WSEAS International Conference on Development, Energy, Environment, Economics, Tenerife 2010, pp. 352-356.

[2] Hamplová E., Kacetl J., Kovárník J. Investment Activities of Czech Companies by Means of Foreign Direct Investment and Their Development, *International Journal of Mathematical Models and Metods in Applied Sciences.* Issue 8, Vol. 5, 2011, pp. 1326-1333.

[3] Teslya A.B. Social Investments of Russian Business: Problems and Prospects, *Proc. of the International Conference on Economics and Applied Statistics*. Saint Petersburg, Russia, 2014, pp. 382-388.

[4] Silkina G.U. Mathematical Modelling as Analytical Instrument of Research of Innovative Processes, *Proc. of the International Conference on Economics and Applied Statistics*. Saint Petersburg, Russia, 2014, pp. 263-272.

[5] Burtseva T.A., Problems of statistics, *Indicative model of monitoring of investment attractiveness of the region*, No.6, 2009, pp. 37 - 45.

[6] Gradov A.P., Kuzin B.I., Mednikov M.D., Sokolitsyn A.S., *Regional economy*, Peter, 2005.

[7] Maksimov I.B., *Investment climate: estimation technique,* Publishing house of BSUEL, 2011.

[8] Suloeva S.B, Rostova O.V., *Management of the investment process in the region (concepts, methods and tools),* Publishing house of Polytechnical University, 2009.

[9] Guskova T.N., Ryabtsev V.N., *Estimation of investment attractiveness of objects by statistical methods,* Publishing house of RGUTS, 2009.

[10] Kuznetsov S.V., *Investment potential of the region: estimation and implementation mechanisms,* Publishing house of IPE RAS, 2003.

[11] Surinov A.E., *Russia in figures. 2013: Summary statistical compilation*, Rosstat, 2013.

[12] *The regions of Russia.2013: Socio-economic indicator,* Rosstat, 2013.

[13] Dubrov A.M., Mhitaryn B.C., Troshin L.I., *Multivariate statistical metods*, Finansy i statistika, 2011.

[14] Maksimov I.B., *Investment climate: estimation technique***,** Publishing house of BSUEL, 2011.

Alexandr Izotov is the Senior Lecturer at the Department of Information Systems in Economics and Management Saint-Petersburg State Polytechnical University (Institute of Industrial Economics and Management), Russia His activities in the research are focused on mathematical methods in economics, statistics, econometrics.

Olga Rostova is the Candidate of Economic Sciences, Associate Professor at the Department of Information Systems in Economics and Management Saint-Petersburg State Polytechnical University (Institute of Industrial Economics and Management), Russia. Her activities in the research are focused on management of the investment process in the region, investment designing..