

# Economic and statistical modeling of the value of urban real property (using the example of Tyumen city, Russia)

Marina Podkovyrova<sup>1\*</sup>, Leonid Popkov<sup>2</sup>, and Olga Volobueva<sup>2</sup>

<sup>1</sup>Industrial University of Tyumen, Tyumen, Russia

**Abstract.** The article presents the results on the derivation of an empirical formula of the relationship between the price of apartments and their area in different planning districts of the city of Tyumen, Russia. Economic and mathematical modeling was performed on the basis of analysis of variance, analysis of variational series, regression and correlation analyses. In a specific mathematical form, methods of mathematical statistics allowed answering a wide range of questions about the influence of various factors on the value of primary real property.

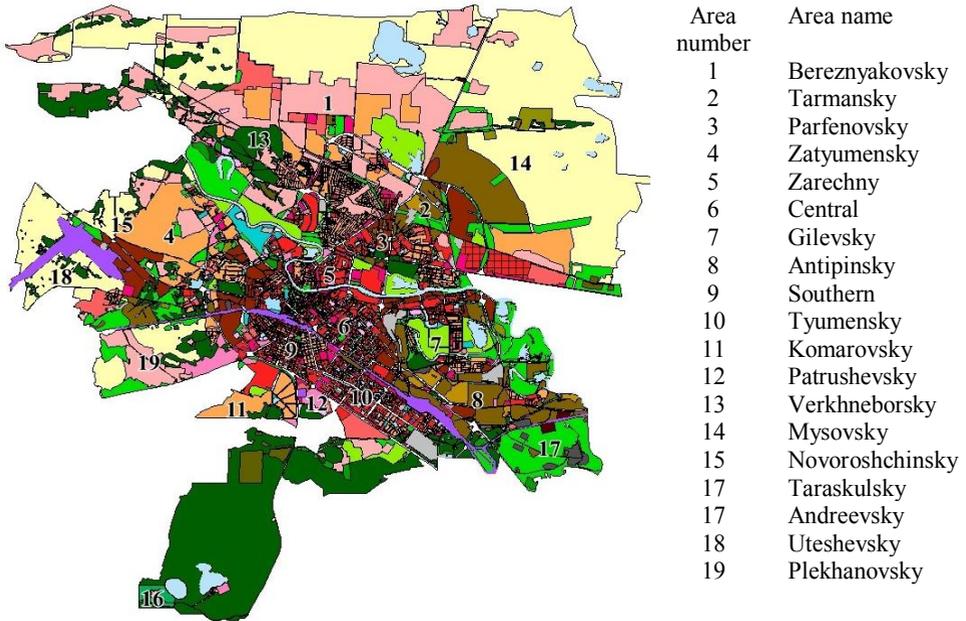
## 1 Introduction

The real property market is a means of redistributing land, buildings, structures and other property between owners and users by economic methods based on competitive demand and supply [1-5].

The main aim of the study was to derive an empirical formula for the relationship between the price of an apartment and its area (as a determining factor). The assumption of a determining factor will be axiomatic in this study, and it is from this factor that one will have to make the necessary calculations. However, in order to broaden the range of the problem under study, in our opinion, at the initial stage it is advisable to carry out a variance analysis of how the location of the studied real property impacts its value, since there is no consensus whether this factor will be significant in modeling the primary housing in the city of Tyumen (Fig. 1).

---

\*Corresponding author: [podkovyrovama@tyuiu.ru](mailto:podkovyrovama@tyuiu.ru)



**Fig. 1.** Planning areas of the city district of Tyumen.

The primary real property market is represented by real property that has not yet been made into ownership. As a rule, those are buildings under construction or newly constructed buildings. The primary real property market will always be formed by the supply of various developers. When it comes to the supply in the primary market, then it should be understood as all the supply in the market of new buildings. The main advantage of the primary real property market is that the apartments are "clean", that is, they have no so-called history [1-6]. It should also be borne in mind that the primary real property market involves the sale of houses, the construction of which took into account all the imperfections of design, and applied the latest technology and modern building materials.

## 2 Study areas and analyses

In our study, four urban areas are presented. At the end of 2017, the construction of new residential complexes was carried out there: Olympia Residential Complex, Vostochny-2, Zarechny-3 and Yamalsky-2 areas.

In these residential complexes, prices for one-room apartments in completed construction projects of approximately the same type of layout and area were analyzed (30 – 60 m<sup>2</sup>) on the basis of economic and statistical modeling [5, 7, 8].

### 2.1 Analysis of variance

The results of the analysis of variance are presented in Table 1 (the analysis was carried out for five apartments in each area considered).

**Table 1.** Grouping of new buildings by residential areas in Tyumen in 2017.

Name of the residential complex	Price of apartments, mln. rub.				
	1	2	3	4	5
Olympia	2050	1979	2008	1850	1935

Vostochny-2	2570	2245	2550	2350	2460
Zarechny-3	2700	2290	3190	2430	2150
Yamalsky-2	1880	2300	2350	2220	1925

The analysis took into account that the residential complexes Vostochny-2 and Zarechny-3 were in the last stages of construction, while the residential complexes Olympia and Yamalsky-2 were already built up and are characterized by a completion level of 50–60 %.

As samples, apartments which 70–80 % are suitable for living were selected (some furniture is missing, but communications are installed and sanitary equipment is fully installed, walls and ceilings are finished). We have put forward a statistical hypothesis: the location of the apartment does not have a significant impact on its value.

It was necessary to find out if the hypothesis was correct. Also, the significance level was taken as the probability of accepting an incorrect hypothesis:  $\alpha = 0.05$ . The results of the calculations are presented in Table 2.

**Table 2.** Results of variance analysis of the influence of the geographical location of the apartments of Tyumen on their value in 2017.

Variance components	Sums of areas	Number of degrees of freedom	Average
Intergroup	1091775.6	3	363925.200
Intragroup	962673.2	16	60167.075
Common	2054448.8	19	

The actually observed value of  $F$  is 6.05, which exceeds the critical value of the Fisher-Snedecor criterion at a significance level of  $\alpha = 0.05$  with  $k_1 = 3$  and  $k_2 = 16$  degrees of freedom  $F_{0.05; 3; 16} = 3.24$ . Since  $F > F_{0.05; 3; 16}$ , we nullified the null hypothesis, that is, at a significance level of  $\alpha = 0.05$  (with a reliability of 0.95), it was assumed that the difference between the areas has a significant impact on the price of the primary one-room housing in Tyumen. Returning to Table 1, we concluded that apartments in Vostochny-2 and Zarechny-3 will have a higher price than apartments in Olympia and Yamalsky-2 residential complexes. This can be explained by the level of completeness of residential buildings, as well as the mass scale of development, since the built-up area is larger in two last districts, therefore, developers have the opportunity to reduce the price due to the scale of the services provided.

## 2.2 Analysis of variational series

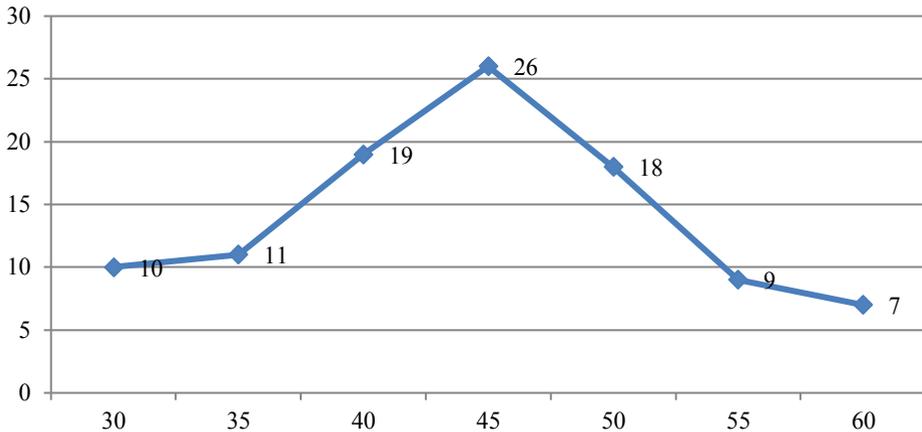
Analysis of variance performed above suggests that apartments in residential complexes Zarechny-3 and Vostochny-2 are more expensive because of the economically inexpedient distribution of one-room apartments by their area. This fact led the study to its main phase: the derivation of an empirical formula of the relationship between the price of an apartment and its area. For this purpose, it was necessary at the initial stage to apply the main method of descriptive statistics – the method of analysis of variational series.

For the study, 100 apartments in Olympia and Yamalsky-2 residential complexes were selected by this method and were distributed across the average area (interval – 5 m<sup>2</sup>) in order to determine for which of the given average areas the demand is greatest (since developers are developing areas of apartments under construction according to the level of demand for each of the given area intervals). The data are presented in Table 3.

**Table 3.** The variation series of the distribution of the number of one-room apartments for sale by their area in Tyumen in 2017.

Apartments areas (m <sup>2</sup> )	30	35	40	45	50	55	60
Number of apartments	10	11	19	26	18	9	7
Specific weight	0.1	0.11	0.19	0.26	0.18	0.09	0.07
Cumulative frequency	0.1	0.21	0.4	0.66	0.84	0.93	1

The table shows that in the considered area interval, there is a gradual increase in the number of apartments to 45 m<sup>2</sup> and its decrease after passing through this area. It is also obvious that apartments with an area of 45 m<sup>2</sup> have the largest share of 26 % in the existing structure, which demonstrates a large percentage of the accumulation of specific weight in the interval from 40 to 50 m<sup>2</sup> (from 40 % to 80 %). For a more convenient presentation of the data of Table 3, a frequency polygon is constructed (Fig. 2).



**Fig. 2.** Frequency polygon of distribution of apartments by their area.

The polyline of the constructed polygon shows the presence of a distinct peak – 26 apartments out of 100 considered (more than 40 %) have an average area of 45 m<sup>2</sup>. For the interval grouping of the available data, it was necessary to determine the number of the group interval and the step size of the interval. Our calculations show that these figures will be equal to six intervals, respectively, and the interval spacing corresponds to 5 m<sup>2</sup>.

**Table 4.** Interval variational series of representation of distribution of the number of apartments by the area.

Apartments areas (m <sup>2</sup> )	30-35	35-40	40-45	45-50	50-55	55-60
Number of apartments	21	19	26	18	9	7

When analyzing the interval series constructed in the form of Table 4, it becomes obvious that apartments with a smaller area are still more popular than apartments with a higher price (more than 45 m<sup>2</sup>) that may be explained by the desire of potential buyers to save on amenities, due to the lower price.

The second stage of the variational analysis is the calculation of the numerical characteristics obtained in Tables 3 and 4. Characteristics are divided into two main groups: average values and indices of variation. The results of the calculation of average values are given in Table 5.

**Table 5.** The results of the calculation of the average values of the discrete and interval variational series.

Index	Average value (m <sup>2</sup> )	Mode (m <sup>2</sup> )	Median (m <sup>2</sup> )
The result of the calculation	44.3	45	45

The value of the sample average 44.3 m<sup>2</sup> shows that the average area of the proposed apartments is very close to 45 m<sup>2</sup>. The same fact is confirmed by the same (also 45 m<sup>2</sup>) modal and median values. The second group of calculated characteristics is the variation indicators; the results of the calculations are presented in Table 6.

**Table 6.** The results of the calculation of the variation indices of the discrete and interval variational series.

Index	Average deviation	Asymmetry	Excess	Coefficient of variation
The result of the calculation	8.19	0.03	- 0.64	18%

The values of the indicators allow us to conclude that: firstly, the empirical distribution of apartments by their area is almost symmetrical (the asymmetry factor is only 3 %, which indicates a slight advantage in demand for apartments up to 45 m<sup>2</sup>); secondly, the value of excess less than 0 (- 0.64) shows that this series has no evident peak, i.e. mode does not have much advantage over the other variants of the series. And finally, the value of the variation coefficient of 18 % allows to make a key conclusion that the distribution of the number of new one-room apartment buildings of Tyumen, depending on their area, is subject to a negative-parabolic law. That in turn created the prerequisite for the regression analysis.

### 2.3 Regression analysis

At the initial stage of regression analysis, one should take the first two lines of a discrete variational series of distribution of the number of proposed new one-room apartment buildings in Vostochny-2 and Zarechny-3 residential complexes in Tyumen as the initial tabular dependence between the considered parameters (Table 7).

**Table 7.** Tabular representation of the expected relationship between the area of apartments and the volume of supply for one-room apartments in new buildings in Tyumen.

Area of apartments (m <sup>2</sup> )	30	35	40	45	50	55	60
Number of apartments	10	11	19	26	18	9	7

Frequency polygon built according to the above table in Fig. 2 suggests that there is a dependence of the parabolic shape between the values under study. In order to check the hypothesis put forward, it is necessary to use the method of least squares. The calculations are given in the form of Table 8 (the area of apartments is taken for  $x$ , and the number of apartments for  $y$ ).

**Table 8.** Calculation table for determining the system of linear equations, expressing a parabolic dependence between the area of apartments and their number.

I	$x_i$	$y_i$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i y_i$	$x_i^2 y_i$
1	30	10	900	27000	810000	300	9000
2	35	11	1225	42875	1500625	385	13475
3	40	19	1600	64000	2560000	760	30400
4	45	26	2025	91125	4100625	1170	52650
5	50	18	2500	125000	6250000	900	45000
6	55	9	3025	166375	9150625	495	27225
7	60	7	3600	216000	12960000	420	25200
$\Sigma$	315	100	14875	732375	37331875	4430	202950

Substituting the values obtained in the table above into the system of linear equations given by formula, we obtain the following system:

$$\begin{cases} 37331875a + 732375b + 14875c = 202950 \\ 732375a + 14875b + 315c = 4430 \\ 14875a + 315b + 7c = 100 \end{cases}$$

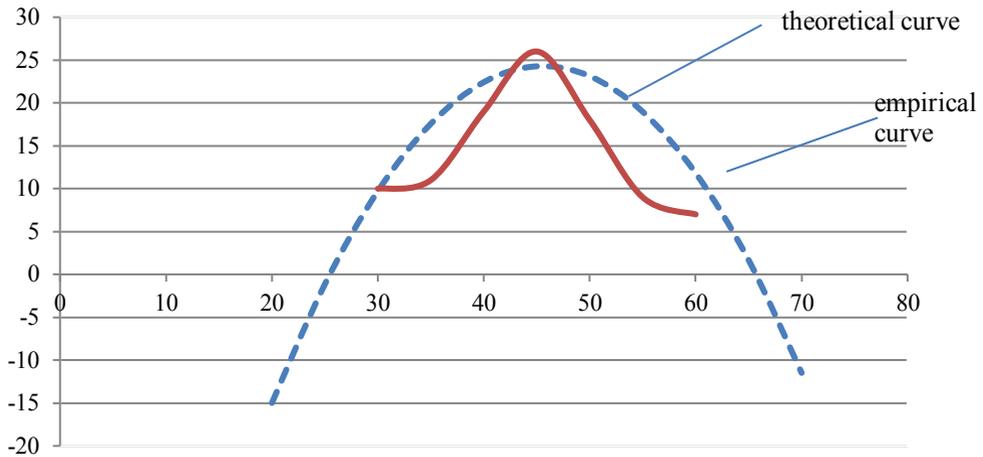
This system was decided to be solved using the Kramer method, as a result of which the values of the desired coefficients were obtained:  $a = -0.06$ ;  $b = 5.47$ ;  $c = -100.38$ .

The final step was to obtain the desired empirical formula:

$$y = -0.06 + 5.47x - 100.38.$$

The resulting formula allows you to determine the following key points:

- $a = -0.06$ ; sets the degree of steepness of the empirical parabola, i.e. shows how much the number of apartments will vary with a square change of  $1 \text{ m}^2$ ;
- $b = 5.47$ ; this value shows that with an increase in the area of an apartment by  $1 \text{ m}^2$ , the number of apartments corresponding to this area will increase by 5.47 times;
- $c = -100.38$ ; shows that if the apartment's area is  $0 \text{ m}^2$ , the volume of supply of these apartments will be:  $-100.38$ ;
- equating the obtained empirical formula to 0, we obtain the values of  $x$  and  $y$ . They mean that if an apartment is larger than  $25.46 \text{ m}^2$ , apartments will have a positive supply of up to  $65.7 \text{ m}^2$ ;
- the highest point = 45.58 will be the value corresponding to the highest supply of apartments.



**Fig. 3.** Empirical and theoretical curves depending on the number of supply of one-room apartments in new buildings in Tyumen.

The final stage of the regression analysis was the comparison of empirical and theoretical curves of the relationship between the studied characteristics. It is shown in Figure 3.

Comparison of the curves presented in Figure 3 confirms that the distribution of the proposed apartments in Zarechny-3 and Vostochny-2 complexes has a parabolic dependence on their area.

## 2.4 Correlation analysis

This analysis involves a comparison of the influence of two main factors on the value of an apartment: price and area. An initial correlation table was compiled, where various options for price and area are given, as well as the number of matching pair frequencies, for each of the possible options. These frequencies were the number of apartments, each of which corresponds to some combination of price and area.

Apartments are located in similar areas (new buildings) and have identical physical characteristics. The area range of apartments is set from 30 to 60 m<sup>2</sup>, the price range is from 1,800 to 2,800 million rubles. We present these data in Table 9.

**Table 9.** Comparison of the number of apartments with their price and area.

Area (m <sup>2</sup> )	Price (mln. rub.)						Total
	1800	2000	2200	2400	2600	2800	
30	4	3	2	1			10
35	5	3	3				11
40	2	7	6	3	1		19
45	3	8	8	4	2	1	26
50		4	6	3	3	2	18
55		1	2	2	3	1	9
60			1	1	3	2	7
Total	14	26	28	14	12	6	100

The basis for the study was 100 apartments, which made it possible to comprehensively consider this issue, eliminating the possible error associated with the sample size. In this table, there is a tendency: the larger area apartments have, the more expensive they are. The

task was to identify, on the basis of this obvious fact, the extent to which the indicators under consideration are interrelated.

To test the hypothesis about the close relationship between area and price, you can calculate the correlation coefficient (Table 10).

**Table 10.** Calculation of the correlation coefficient of the relationship between price and area of apartments in Tyumen.

Group average of prices and areas	Average area (m <sup>2</sup> )	Average price (mln. rub.)	Deviation from the average price	Deviation from the average area	The correlation coefficient between price and area
98.960	44.3	2.204	8.19	281.4	0.57

The obtained correlation coefficient is 0.57, which indicates that the degree of relationship between these indicators is characterized as average.

### 3 Conclusion

1. The hypothesis put forward at the beginning of the study is rejected on the basis that the actually observed value of F exceeds the critical value of the Fisher-Snedecor criterion.

2. The location of apartments has a significant impact on the value of apartments, as evidenced by the data in Table 1. As a conclusion, it follows that that apartments in Vostochny-2 and Zarechny-3 districts will have a higher price than apartments in Olympia and Yamal-2 residential complexes.

3. Methods of mathematical statistics allowed in a specific mathematical form to answer a wide range of questions about the influence of various factors on the value of primary real property: location, urban planning solutions, the presence and condition of communication networks, the technical condition of the building, environmental factors [2, 3, 5].

### References

1. A. Asaul, *Real property economics: textbook* (St. Petersburg, Peter, 2013)
2. N. Fastovets, M. Popov, *Mathematical statistics (examples, tasks and typical tasks)* (Moscow, 2012)
3. N. Kremer, *Probability theory and mathematical statistics* (Moscow, 2010)
4. E. Malyshev, M. Podkovyrova, I. Skripa, *Workshop on real property economics* (Omsk, 2005)
5. T. Kasyanenko, G. Makhovikova, V. Esipov, S. Mirzazhanov, *Real property appraisal* (2018)
6. M. Podkovyrova, A. Oleinik, *Scientific and methodological foundations of the functioning of the land and property complex* (Tyumen, 2018)
7. V. Tsyganenko, *Economy of real property market* (St. Petersburg, 2013)
8. L. Popkov, V. Almieva, *Actual problems of geodesy, cadastre, rational land and nature management* (Tyumen, 2018)