

## МОДЕЛИРОВАНИЕ РЕГИОНАЛЬНОЙ СИСТЕМЫ ЦЕН НА ЖИЛУЮ НЕДВИЖИМОСТЬ

*Е. Кузнецова*

В условиях рыночной экономики цена является результатом взаимодействия факторов спроса и предложения. Предложение на рынке недвижимости характеризуется показателем ввода жилья в эксплуатацию. Спрос на недвижимость можно выразить при помощи косвенных показателей, таких как реальные доходы населения, уровень безработицы, а также параметров отражающих экономическую ситуацию в стране.

С помощью корреляционно-регрессионного анализа определим степень воздействия на стоимость жилой недвижимости в течение периода с января 2006 по январь 2011 года. Для этого необходимо построить регрессионные модели зависимости стоимости квадратного метра от различных факторов.

Регрессионная модель – экономико-статистическая модель, основанная на уравнении регрессии или системе регрессионных уравнений, связывающих величины экзогенных (объясняющих) и эндогенных (выходных) переменных.

В регрессионных моделях зависимая (объясняемая) переменная  $Y$  может быть представлена в виде функции  $f(X_1, X_2, \dots, X_n)$ , где  $X_1, X_2, \dots, X_n$  – независимые (объясняющие) переменные или факторы. В зависимости от вида функции  $f(X_1, X_2, \dots, X_n)$  модели делятся на линейные и нелинейные. В зависимости от количества включенных в модель факторов  $X$  модели делятся на однофакторные и многофакторные.

Основными этапами построения регрессионной модели являются:

Построение системы показателей (факторов).

Сбор и предварительный анализ исходных данных.

Построение матрицы коэффициентов парной корреляции.

Выбор вида модели и численная оценка ее параметров.

Проверка качества модели.

Оценка влияния отдельных факторов на основе модели.

Прогнозирование на основе модели регрессии.

Выбор факторов, влияющих на исследуемый показатель, производится на основании качественного и количественного анализа исследуемых явлений.

Исключение части факторов осуществляется на основе анализа парных коэффициентов корреляции и оценкой их значимости. Коэффициент парной корреляции определяется по формуле

$$r_{y,x} = \frac{\sum (y - \bar{y})(x - \bar{x})}{\sqrt{\sum (y - \bar{y})^2 \sum (x - \bar{x})^2}},$$

где  $\bar{x}$  – среднее значение факторного признака,  
 $\bar{y}$  – среднее значение результативного признака.

Значение коэффициентов парной корреляции лежит в интервале от -1 до +1. Его положительное значение свидетельствует о прямой связи, отрицательное – об обратной, т. е. когда растет одна переменная, другая уменьшается. Коэффициент корреляции выражает количественную характеристику тесноты связи между признаками (табл. 1).

Таблица 1

Характер тесноты связи

Диапазон изменения $ r $	0,1–0,3	0,3–0,5	0,5–0,7	0,7–0,9	0,9–0,99
Характер тесноты связи	Слабая	Умеренная	Заметная	Высокая	Весьма высокая

Для оценки значимости коэффициента корреляции применяется  $t$  – критерий Стьюдента. При этом фактическое значение этого критерия ( $t_{\text{набл}}$ )

$$t_{\text{набл}} = \sqrt{\frac{r^2}{1-r^2}}(n-2)$$

сравнивается с критическим значением  $t_{\text{кр}}$  которое берется из таблицы значений  $t$  с учетом заданного уровня значимости ( $\alpha = 0,05$ ) и числа степеней свободы ( $n - 2$ ).

Если  $t_{\text{набл}} > t_{\text{кр}}$ , то полученное значение коэффициента парной корреляции признается значимым.

Одним из условий регрессионной модели является предположение о функциональной независимости объясняющих переменных. Связь между факторами называется мультиколлинеарностью, которая делает вычисление параметров модели либо невозможным, либо затрудняет содержательную интерпретацию параметров модели. Считают явление мультиколлинеарности в исходных данных установленным, если коэффициент парной корреляции между двумя переменными больше 0,8. Чтобы избавиться от мультиколлинеарности, в модель включают лишь один из функционально связанных между собой факторов, причем тот, который в большей степени связан с зависимой переменной.

Итак, построим регрессионную модель для первичного рынка жилья города Челябинска:

$y$  – стоимость 1 квадратного метра жилья на первичном рынке г. Челябинска;

$x_1$  – биржевой индекс ММВБ;

Таблица 2

Матрица коэффициентов парной корреляции

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	Уп	Ув
X1	1																
X2	0,972	1															
X3	-0,2531	-0,2771	1														
X4	-0,0612	0,0045	-0,0127	1													
X5	-0,0423	-0,1048	0,9016	-0,0664	1												
X6	-0,5887	-0,6823	-0,0803	-0,2512	0,1936	1											
X7	0,5788	0,6757	0,2949	0,0234	0,3306	0,6327	1										
X8	0,1026	0,1902	-0,1130	-0,0123	0,1649	0,4036	0,0941	1									
X9	-0,4907	-0,3775	-0,5028	0,1021	0,6627	0,3476	-0,4058	0,1598	1								
X10	-0,0255	-0,0703	0,9235	-0,0044	0,9513	0,2864	0,3948	-0,0616	-0,7050	1							
X11	-0,3482	-0,4756	0,4189	-0,1971	0,2694	0,7054	-0,2839	-0,3913	-0,1954	0,2525	1						
X12	0,4438	0,4863	-0,1861	0,0750	0,1156	0,3923	0,4347	0,0135	-0,1915	-0,0436	-0,3164	1					
X13	-0,3528	-0,4380	0,8753	-0,1248	0,8085	0,2456	0,0298	-0,1973	-0,4080	0,8110	0,6702	-0,3190	1				
X14	0,0475	0,0847	0,8004	0,0575	0,7950	0,3925	0,6262	-0,0663	-0,5742	0,8422	0,0797	0,1845	0,5878	1			
X15	-0,1059	-0,2159	0,8673	-0,1050	0,8794	0,0525	0,1797	-0,2786	-0,7390	0,8980	0,5479	-0,1624	0,8591	0,6715	1		
Уп	-0,1393	-0,0563	0,6564	0,1691	0,5713	0,4430	0,3598	0,2492	-0,0749	0,6366	-0,0851	-0,0607	0,5356	0,6411	0,3112	1	
Ув	-0,0038	0,1707	0,2083	0,3057	0,1430	0,6666	0,4529	0,4282	0,2101	0,2245	-0,5568	0,1303	-0,0229	0,4143	-0,1876	0,81	1

- x2 – биржевой индекс РТС;
- x3 – среднемесячный доход на душу населения, рублей;
- x4 – миграция, человек;
- x5 – среднемесячная номинальная начисленная заработная плата, рублей;
- x6 – средневзвешенная процентная ставка, процент;
- x7 – стоимость барреля нефти (марка Brent), USD/ баррель;
- x8 – уровень инфляции, процент;
- x9 – ставка рефинансирования, процент;
- x10 – денежная масса, миллиардов рублей;
- x11 – уровень безработицы, тысяч человек;
- x12 – сводные индексы цен строительной продукции по Челябинской области;
- x13 – стоимость строительства квадратного метра, рублей;
- x14 – потребительская стоимость 1 литра бензина АИ-92, рублей;
- x15 – потребительская стоимость 1 м<sup>3</sup> природного газа, рублей;
- уп – стоимость квадратного метра первичного жилья, рублей;
- ув – стоимость квадратного метра вторичного жилья, рублей.

Первоначально построим матрицу коэффициентов парной корреляции, для чего воспользуемся средствами MS Excel. Необходимо выбрать команду меню *Сервис/Анализ данных/Корреляция*.

Анализ матрицы коэффициентов парной корреляции показывает, что существенное влияние на зависимую переменную оказывают факторы x3, x5, x10, ув, поэтому для исключения явления мультиколлинеарности из модели следует исключить фактор среднемесячный доход на душу населения, среднемесячная номинальная начисленная заработная плата, денежная масса.

Для проведения регрессионного анализа выберем пункт меню *Сервис/Анализ данных/Регрессия*. После заполнения полей ввода нажимаем кнопку *ОК* и получаем следующие результаты:

Таблица 3

Регрессионная статистика уравнения для стоимости первичного жилья г. Челябинска

<i>Регрессионная статистика</i>	
Множественный R	0,977810716
R-квадрат	0,956113795
Нормированный R-квадрат	0,945142244
Стандартная ошибка	1123,718464
Наблюдения	61

Множественный R (множественный коэффициент корреляции) – множественный коэффициент корреляции является показателем тесноты линейной связи между результативным признаком и совокупностью факторных признаков. Множественный коэффициент корреляции изменяется в пределах от 0 до 1. Равенство его нулю говорит об отсутствии линейной

связи, равенство единице говорит о функциональной связи. Указаний на то, является ли связь прямой или обратной, коэффициент не дает.

Квадрат множественного коэффициента корреляции называется множественным коэффициентом детерминации (R-квадрат). Он показывает, какая доля дисперсии результативного признака объясняется влиянием независимых переменных.

Нормированный R-квадрат – скорректированный индекс множественной детерминации, содержащий поправку на число степеней свободы.

Таблица 4

Коэффициенты уравнения регрессии для стоимости первичного жилья г. Челябинска

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	12	1,32E+09	1,1E+08	87,14482	2,49E-28
Остаток	48	60611673	1262743		
Итого	60	1,38E+09			

*df* – число степеней свободы; *SS* – сумма квадратов (в строке «Регрессия» – сумма квадратов отклонений вычисленных значений отклика от среднего значения откликов, в строке «Остаток» – сумма квадратов разностей наблюдаемых и вычисленных значений отклика);

Значимость уравнения в целом оценивается с помощью *F* – критерия Фишера.

*F* – критерий Фишера используют для сравнения дисперсий двух вариационных рядов.

Альтернативным способом проверки значимости гипотезы является критерий значимость *F*, который представляет собой вероятность ошибки в случае принятия данной модели.

Столбец «Коэффициенты» содержит значения коэффициентов уравнения регрессии.

Таблица 5

Коэффициенты уравнения регрессии для стоимости первичного жилья г. Челябинска

	Коэффициенты	Стандартная ошибка	<i>t</i> -статистика	<i>p</i> -значение	Нижние 95 %	Верхние 95 %
Y–пересечение	75648,31	17110,82	4,42	0,00	41244,7	110051
Переменная X 1	-7,84	5,99	-1,31	0,20	-19,89	4,22
Переменная X 2	6,55	5,17	1,27	0,21	-3,85	16,95
Переменная X 3	-0,62	1,59	-0,39	0,70	-3,81	2,58
Переменная X 4	-5411,29	626,26	-8,64	0,00	-6670,48	-4152,11
Переменная X 5	-45,48	26,77	-1,70	0,10	-99,30	8,34
Переменная X 6	-408,33	245,94	-1,66	0,10	-902,82	86,16
Переменная X 7	1102,30	386,40	2,85	0,01	325,40	1879,21
Переменная X 8	1,67	72,16	0,02	0,98	-143,42	146,76
Переменная X 9	-33,08	127,73	-0,26	0,80	-289,89	223,74
Переменная X 10	1,38	0,12	11,92	0,00003	1,14	1,61
Переменная X 11	525,82	199,09	2,64	0,01	125,51	926,12
Переменная X 12	-3872,10	1032,01	-3,75	0,00001	-5947,10	-1797,11

Столбец «Стандартная ошибка» содержит выборочные стандартные отклонения по каждому коэффициенту уравнения регрессии, стандартные ошибки коэффициентов.

Если стандартная ошибка больше абсолютной величины коэффициента, это коэффициент незначимый. Этот коэффициент (свободный член или регрессор) нужно исключить из уравнения регрессии и пересчитать таблицы. Но это грубый анализ. Столбец  $t$ -статистика дает более точную оценку значимости коэффициентов.

Столбец  $t$ -статистика содержит значения  $t$ -критерия Стьюдента, позволяющий оценить значимость отдельных коэффициентов путем проверки гипотезы о равенстве нулю параметра уравнения (кроме свободного члена).

В столбце  $p$ -значение приводится достоверность отличия соответствующих коэффициентов от нуля. В случае, когда  $p > 0,05$ , коэффициент может считаться нулевым.

Интервальные оценки каждого коэффициента регрессии приведены в столбцах «Нижние 95 %» и «Верхние 95 %».

Оценка качества модели по критериям Стьюдента и Фишера будет проводиться путём сравнения расчетных значений с табличными.

Для оценки качества модели по критерию Стьюдента фактическое значение этого критерия ( $t_{\text{набл}}$ )  $t_{\text{набл}} = \sqrt{\frac{r^2}{1-r^2}}(n-2)$  сравнивается с критическим значением  $t_{\text{кр}}$  которое берется из таблицы значений  $t$  с учетом заданного уровня значимости ( $\alpha = 0,05$ ) и числа степеней свободы ( $n - 2$ ).

Если  $t_{\text{набл}} > t_{\text{кр}}$ , то полученное значение коэффициента парной корреляции признается значимым. Критическое значение при  $\alpha = 0,05$  и  $v=48$  равно 2,010635.

Таблица 6

Сравнение по критерию Стьюдента

Фактор	$t_{\text{набл}}$	$t_{\text{кр}}$	Значимость
Переменная X 1	-1,31	2,0106347	не существенна
Переменная X 2	1,27	2,0106347	не существенна
Переменная X 3	-0,39	2,0106347	не существенна
Переменная X 4	-8,64	2,0106347	существенна
Переменная X 5	-1,70	2,0106347	не существенна
Переменная X 6	-1,66	2,0106347	не существенна
Переменная X 7	2,85	2,0106347	не существенна
Переменная X 8	0,02	2,0106347	не существенна
Переменная X 9	-0,26	2,0106347	не существенна
Переменная X 10	11,92	2,0106347	существенна
Переменная X 11	2,64	2,0106347	существенна
Переменная X 12	-3,75	2,0106347	существенна

Также необходимо, чтобы  $p$ -значение было ниже указанного нами уровня существенности (0,05), в нашем случае  $p$ -значение выше для переменной 1,2,3,5,6,7,9.

Проверим значимость коэффициента детерминации, используя  $F$ -критерий Фишера.

Математической моделью статистического распределения  $F$ -статистики является распределение Фишера, Критическое значение этой статистики при  $\alpha = 0,05$  и  $\nu_1=12$  и  $\nu_2=48$  степенях свободы равно 1,96012106.

Таблица 7

Проверка адекватности модели

$F_{\text{расч}}$	$F_{\text{кр}}$	Уравнение регрессии
87,14482	1,96012106	адекватно

Таким образом, из регрессионной модели необходимо исключить факторы: биржевой индекс РТС, биржевой индекс ММВБ, миграция, стоимость барреля нефти, уровень инфляции, ставка рефинансирования, уровень безработицы, сводные индексы цен строительной продукции.

Построим новую модель на основании оставшихся факторов:

Таблица 8

Регрессионная статистика уравнения для стоимости первичного жилья г. Челябинска

Регрессионная статистика	
Множественный R	0,9558664
R-квадрат	0,91368058
Нормированный R-квадрат	0,90751491
Стандартная ошибка	1459,06408
Наблюдения	61

Таблица 9

Коэффициенты уравнения регрессии для стоимости первичного жилья г. Челябинска

	$df$	$SS$	$MS$	$F$	Значимость $F$
Регрессия	4	1261893324	3,15E+08	148,1883019	4,32E-29
Остаток	56	119216606,9	2128868		
Итого	60	1381109931			

В новой регрессионной модели коэффициент детерминации равен 0,914, его значение меньше, чем в предыдущей модели.

Таблица 10

Коэффициенты уравнения регрессии для стоимости первичного жилья г. Челябинска

	Коэффициенты	Стандартная ошибка	t-статистика	p-Значение	Нижние 95 %	Верхние 95 %
Y–пересечение	69443,1479	6452,963998	10,76143	2,99244E–15	56516,31	82369,99
Переменная X 1	–4414,16447	370,1816741	–11,9243	5,32262E–17	–5155,73	–3672,6
Переменная X 2	1,61203872	0,093922806	17,16344	5,71979E–24	1,423889	1,800189
Переменная X 3	547,263884	157,9108067	3,465652	0,001022466	230,9305	863,5972
Переменная X 4	–7637,68885	571,3462071	–13,3679	4,54641E–19	–8782,23	–6493,14

Проверим новую модель на адекватность и значимость факторов.

Таблица 11

Проверка факторов на значимость

Фактор	$t_{\text{набл}}$	$t_{\text{кр}}$	Значимость
Переменная X 1	–11,9243	2,003241	существенна
Переменная X 2	17,16344	2,003241	существенна
Переменная X 3	3,465652	2,003241	существенна
Переменная X 4	–13,3679	2,003241	существенна

Все факторы, по которым строилась модель регрессии являются существенными, также это подтверждается тем, что  $p$ -значение ниже уровня существенности.

Таблица 12

Проверка регрессионной модели на адекватность

$F_{\text{расч}}$	$F_{\text{кр}}$	Уравнение регрессии
148,18	2,536	адекватно

Необходимо определить вид регрессионной модели. В MS Excel это можно сделать с помощью функций ЛИНЕЙН и ЛГРФПРИБЛ.

Для линейной аппроксимации в Excel существует функция *ЛИНЕЙН* (изв. зн. Y, изв. зн. X, константа, статистика) она возвращает массив значений, описывающих кривую вида:

$$Y = b + m_1 \cdot x_1 + m_2 \cdot x_2 + \dots + m_n \cdot x_n$$

где изв. зн. Y – это известные значения функции; изв. зн. X – это известные значения аргументов

*константа* – определяет чему должно равняться  $b$ , если *константа* имеет значение ЛОЖЬ то  $b$  полагается равным 1, иначе  $b$  вычисляется обычным образом.

*статистика* – если значение равно ИСТИНА, то будет представлена дополнительная регрессионная статистика, если ЛОЖЬ, то нет.

Данная функция возвращает массив значений, в данных ячейках будет полная статистическая информация



$m_n$	$m_{n-1}$	...	$b$
$S_{en}$	$S_{en-1}$	...	$S_{eb}$
$R^2$	$S_{ey}$		
$F$	$Df$		
$Ssreg$	$Ssresid$		

- $S_e$  – стандартная ошибка для коэффициента  $m$ ;  
 $S_{eb}$  – стандартная ошибка для свободного члена  $b$ ;  
 $R^2$  – коэффициент детерминированности, который показывает как близко уравнение описывает исходные данные. Чем ближе он к 1, тем больше сходится теоретическая зависимость и экспериментальные данные;  
 $S_{ey}$  – стандартная ошибка для  $y$ ;  
 $F$  – критерий Фишера определяет случайная или нет взаимосвязь между зависимой и независимой переменными;  
 $Df$  – степень свободы системы;  
 $Ssreg$  – регрессионная сумма квадратов;  
 $Ssresid$  – остаточная сумма квадратов.

Таблица 13

Результаты, полученные при помощи функции ЛИНЕЙН при аргументе Константа, имеющем значение ИСТИНА

-7637,68885	547,263884	1,61203872	-4414,16	69443,14786
571,3462071	157,910807	0,093922806	370,1817	6452,963998
0,913680581	1459,06408	#Н/Д	#Н/Д	#Н/Д
148,1883019	56	#Н/Д	#Н/Д	#Н/Д
1261893324	119216607	#Н/Д	#Н/Д	#Н/Д

Таблица 14

Результаты, полученные при помощи функции ЛИНЕЙН при аргументе Константа, имеющем значение ЛОЖЬ

-8095,595345	1948,34934	1,203478147	-631,505	0
989,1825701	155,140175	0,149147833	201,5811	#Н/Д
0,995367148	2533,14015	#Н/Д	#Н/Д	#Н/Д
3061,60878	57	#Н/Д	#Н/Д	#Н/Д
78582913082	365757545	#Н/Д	#Н/Д	#Н/Д

Для экспоненциальной аппроксимации в Excel существует функция ЛГРФПРИБЛ (изв. зн.  $Y$ , изв. зн.  $X$ , константа, статистика) она возвращает массив значений описывающих кривую вида:

$$Y = b \cdot m_1^{x_1} \cdot m_2^{x_2} \cdot \dots \cdot m_n^{x_n},$$

изв. зн.  $Y$  – это известные значения функции;

изв. зн.  $X$  – это известные значения аргументов;

*константа* – определяет чему должно равняться  $b$ , если *константа* имеет значение ЛОЖЬ, то  $b$  полагается равным 1, иначе  $b$  вычисляется обычным образом;

*статистика* – если значение равно ИСТИНА, то будет представлена дополнительная регрессионная статистика, если ЛОЖЬ то нет.

Таблица 15

Результаты, полученные при помощи функции ЛГРФПРИБЛ при аргументе Константа, имеющем значение ИСТИНА

0,804727	1,014203	1,000048	0,873206738	103430,2
0,016139	0,004461	2,65E-06	0,010456696	0,18228
0,923222	0,041215	#Н/Д	#Н/Д	#Н/Д
168,3447	56	#Н/Д	#Н/Д	#Н/Д
1,143845	0,095125	#Н/Д	#Н/Д	#Н/Д

Таблица 16

Результаты, полученные при помощи функции ЛГРФПРИБЛ при аргументе Константа, имеющем значение ЛОЖЬ

0,74573	1,280267	0,99998	1,637840053	1
0,135975	0,021326	2,05E-05	0,027709826	#Н/Д
0,998967	0,348212	#Н/Д	#Н/Д	#Н/Д
13780,98	57	#Н/Д	#Н/Д	#Н/Д
6683,848	6,911327	#Н/Д	#Н/Д	#Н/Д

Из представленных моделей необходимо выбрать ту, которая наиболее точно описывает фактические данные. Сравним коэффициенты детерминации в каждой модели.

Зависимость	Вид уравнения	$R^2$
Линейная	$y(x) = b_0 + b_1 \cdot x$	0,9314
Линейная	$y(x) = b \cdot x$	0,9954
Экспоненциальная	$y(x) = b_0 \cdot b_1^x$	0,9232
Экспоненциальная	$y(x) = b^x$	0,9989

Моделью наиболее точно описывающей фактические данные является экспоненциальная модель вида  $y(x)=b^x$ , так как для нее коэффициент детерминированности  $R^2$  имеет наибольшее значение.

Таким образом, модель  $y(x)=b_1^{x^1} \cdot b_2^{x^2} \cdot b_3^{x^3} \cdot b_4^{x^4}$  объясняет **99,8 %** общей дисперсии признака  $Y$ . Это указывает на то, что подобранная модель является адекватной.

В результате проведенного корреляционно–регрессионного анализа было выявлено, что стоимость квадратного метра на первичном рынке г. Челябинска зависит от средневзвешенной процентной ставки, стоимость

строительства квадратного метра жилья, потребительской стоимости бензина и потребительской стоимости газа.

Далее составим регрессионную модель для вторичного рынка жилья г. Челябинска, основываясь на тех же факторах.

Таблица 17

Регрессионная статистика уравнения для стоимости квадратного метра жилья вторичного рынка г. Челябинска

<i>Регрессионная статистика</i>	
Множественный R	0,980685138
R–квадрат	0,96174334
Нормированный R–квадрат	0,952179175
Стандартная ошибка	1388,20223
Наблюдения	61

Таблица 18

Показатели регрессионного анализа уравнения регрессии для стоимости квадратного метра жилья вторичного рынка г. Челябинска

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	12	2,33E+09	193783859	100,557	9,51E–30
Остаток	48	92501061	1927105,43		
Итого	60	2,42E+09			

Таблица 19

Коэффициенты уравнения регрессии для стоимости квадратного метра жилья для вторичного рынка г. Челябинска

	Коэффициенты	Стандартная ошибка	<i>t</i> -статистика	<i>p</i> -значение	Нижние 95 %	Верхние 95 %
Y–пересечение	126244,2938	21138,1008	5,9724	0,0000	83743,29	168745,293
Переменная X 1	–13,8004	7,4047	–1,8637	0,0685	–28,6885	1,0878
Переменная X 2	7,4425	6,3891	1,1649	0,2498	–5,4037	20,2887
Переменная X 3	1,3363	1,9646	0,6802	0,4996	–2,6137	5,2863
Переменная X 4	–7059,1000	773,6649	–9,1242	0,0000	–8614,66	–5503,5424
Переменная X 5	21,2563	33,0677	0,6428	0,5234	–45,2308	87,7433
Переменная X 6	–283,6754	303,8216	–0,9337	0,3551	–894,547	327,1988
Переменная X 7	1091,4756	477,3428	2,2866	0,0267	131,714	2051,2377
Переменная X 8	–112,1105	89,1474	–1,2576	0,2146	–291,353	67,1323
Переменная X 9	–171,7349	157,7919	–1,0884	0,2819	–488,997	145,5270
Переменная X 10	1,0864	0,1427	7,6117	0,0000	0,7995	1,3734
Переменная X 11	689,5893	245,9516	2,8038	0,0073	195,071	1184,1080
Переменная X 12	–6365,2525	1274,9093	–4,9927	0,0000	–8928,63	–3801,8755

Коэффициент детерминации (0,96) свидетельствует о высокой зависимости между ценой и выбранными факторами.

Факторы  $x_1, x_2, x_3, x_5, x_6, x_8, x_9$  являются не существенными, что также подтверждается тем, что  $p$ -значение выше уровня значимости.

Таблица 20

Проверка факторов регрессионной модели на значимость

Фактор	$t_{\text{набл}}$	$t_{\text{кр}}$	Значимость
Переменная X 1	-1,8637	2,0106347	не существенна
Переменная X 2	1,1649	2,0106347	не существенна
Переменная X 3	0,6802	2,0106347	не существенна
Переменная X 4	-9,1242	2,0106347	существенна
Переменная X 5	0,6428	2,0106347	не существенна
Переменная X 6	-0,9337	2,0106347	не существенна
Переменная X 7	2,2866	2,0106347	существенна
Переменная X 8	-1,2576	2,0106347	не существенна
Переменная X 9	-1,0884	2,0106347	не существенна
Переменная X 10	7,6117	2,0106347	существенна
Переменная X 11	2,8038	2,0106347	существенна
Переменная X 12	-4,9927	2,0106347	существенна

Таблица 21

Проверка регрессионной модели на адекватность

$F_{\text{расч}}$	$F_{\text{кр}}$	Уравнение регрессии
100,557	1,96014	адекватно

Несмотря на высокий коэффициент детерминации в модели присутствуют незначимые факторы, поэтому связь между независимыми переменными и результативной может не являться статистически значимой. Поскольку данная модель не является адекватной при составлении прогноза, то в данном случае будет целесообразно разработать альтернативную модель.

Таким образом, из регрессионной модели необходимо исключить следующие факторы: биржевой индекс РТС, биржевой индекс ММВБ, миграцию, стоимость барреля нефти, уровень инфляции, уровень безработицы, сводные индексы цен строительной продукции.

Построим новую регрессионную модель на основании значимых факторов.

Таблица 22

Регрессионная статистика уравнения для стоимости квадратного метра вторичного жилья г. Челябинска

<i>Регрессионная статистика</i>	
Множественный R	0,973267066
R-квадрат	0,947248782
Нормированный R-квадрат	0,942453217
Стандартная ошибка	1522,841588
Наблюдения	61

Таблица 23

Показатели дисперсионного анализа уравнения регрессии для стоимости вторичного жилья г. Челябинска

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	Значимость <i>F</i>
Регрессия	5	2,29E+09	458071962	197,526	7,79E-34
Остаток	55	1,28E+08	2319046,5		
Итого	60	2,42E+09			

Таблица 24

Коэффициенты уравнения регрессии для стоимости вторичного жилья г. Челябинска

	Коэффициенты	Стандартная ошибка	<i>t</i> -статистика	<i>p</i> -Значение	Нижние 95 %	Верхние 95 %
У–пересечение	81618,72168	6821,776	11,9644395	6,17E-17	67947,58	95289,87
Переменная X 1	-7657,873755	423,9917	-18,061376	8,71E-25	-8507,57	-6808,18
Переменная X 2	2751,039541	285,3525	9,64084627	2,03E-13	2179,18	3322,899
Переменная X 3	0,844454032	0,115999	7,27986294	1,31E-09	0,611988	1,07692
Переменная X 4	1173,122124	165,4563	7,09022454	2,68E-09	841,5403	1504,704
Переменная X 5	-3864,305305	1029,497	-3,7535843	0,000422	-5927,46	-1801,15

Таблица 25

Проверка факторов регрессионной модели на значимость

Фактор	$t_{\text{набл}}$	$t_{\text{кр}}$	Значимость
Переменная X 1	-18,061376	2,00405	существенна
Переменная X 2	9,64084627	2,00405	существенна
Переменная X 3	7,27986294	2,00405	существенна
Переменная X 4	7,09022454	2,00405	существенна
Переменная X 5	-3,7535843	2,00405	существенна

Таблица 26

Проверка регрессионной модели на адекватность

$F_{\text{расч}}$	$F_{\text{кр}}$	Уравнение регрессии
197,526	2,38223	адекватно

Табличное значение коэффициента Стьюдента при уровне значимости  $\alpha = 0,05$  и значением степени свободы 55 равно 2,004.

Расчётное значение 197,526 больше табличного критерия с уровнем с уровнем значимости  $\alpha = 0,05$  и значениями степеней свободы 5 и 55 равно 2,38223, поэтому предположение об уравнении регрессии значимо, то есть зависимость цены от выбранных факторов существует, а не является результатом случайного совпадения. Критерий значимость  $F$  представляет собой вероятность ошибки в случае принятия данной модели. В нашем случае показатель равен  $7,79 \cdot 10^{-34}$ , что меньше уровня существенности. В таблице приведены коэффициенты уравнения регрессии. Стандартная

ошибка позволяет судить о среднем отклонении факторов от среднего значения.

Далее необходимо выбрать тип регрессионной зависимости.

Таблица 27

Результаты, полученные при помощи функции ЛИНЕЙН при аргументе Константа, имеющем значение ИСТИНА

-3864,305305	1173,122124	0,844454	2751,03954	-7657,87	81618,72
1029,497418	165,4562724	0,115999	285,352495	423,9917	6821,776
0,947248782	1522,841588	#Н/Д	#Н/Д	#Н/Д	#Н/Д
197,5259926	55	#Н/Д	#Н/Д	#Н/Д	#Н/Д
2290359811	127547557,6	#Н/Д	#Н/Д	#Н/Д	#Н/Д

Таблица 28

Результаты, полученные при помощи функции ЛИНЕЙН при аргументе Константа, имеющем значение ЛОЖЬ

-2792,785494	2750,533019	0,258442	3293,76017	-3656,46	
1929,195896	188,0446345	0,197794	529,937286	490,2202	
0,995063242	2864,543067	#Н/Д	#Н/Д	#Н/Д	#Н/Д
2257,495204	56	#Н/Д	#Н/Д	#Н/Д	#Н/Д
92620592050	459513991,1	#Н/Д	#Н/Д	#Н/Д	#Н/Д

Таблица 29

Результаты, полученные при помощи функции ЛГРФПРИБЛ при аргументе Константа, имеющем значение ИСТИНА

0,889357348	1,023822719	1,000025	1,06424805	0,818164	145953,4
0,026590664	0,004273534	3E-06	0,00737031	0,010951	0,176198
0,943811278	0,039333144	#Н/Д	#Н/Д	#Н/Д	#Н/Д
184,7688243	55	#Н/Д	#Н/Д	#Н/Д	#Н/Д
1,429275712	0,08509029	#Н/Д	#Н/Д	#Н/Д	#Н/Д

Таблица 30

Результаты, полученные при помощи функции ЛГРФПРИБЛ при аргументе Константа, имеющем значение ЛОЖЬ

1,039618386	1,288342611	0,99994	1,15181336	1,465613	1
0,240331543	0,023425852	2,46E-05	0,06601748	0,06107	#Н/Д
0,998949304	0,35685337	#Н/Д	#Н/Д	#Н/Д	#Н/Д
10648,39897	56	#Н/Д	#Н/Д	#Н/Д	#Н/Д
6780,06602	7,131282333	#Н/Д	#Н/Д	#Н/Д	#Н/Д

Из представленных моделей необходимо выбрать ту, которая наиболее точно описывает фактические данные. Сравним коэффициенты детерминации в каждой модели.

Зависимость	Вид уравнения	$R^2$
Линейная	$y(x) = b_0 + b_1 \cdot x$	0,9314
Линейная	$y(x) = b \cdot x$	0,9954
Экспоненциальная	$y(x) = b_0 \cdot b_1^x$	0,9232
Экспоненциальная	$y(x) = b^x$	0,9989

Моделью наиболее точно описывающей фактические данные является экспоненциальная модель вида  $y(x)=b^x$ , так как для нее коэффициент детерминированности  $R^2$  имеет наибольшее значение.

Таким образом, модель  $y(x)=b_1^{x_1} * b_2^{x_2} * b_3^{x_3} * b_4^{x_4} * b_5^{x_5}$  объясняет **99,8 %** общей дисперсии признака  $Y$ . Это указывает на то, что подобранная модель является адекватной.

Таким образом, данная регрессионная модель является адекватной и основана на значимых факторах, поэтому, в результате проведенного корреляционно-регрессионного анализа было установлено, что цена квадратного метра на вторичном рынке жилья г. Челябинска зависит от средневзвешенной процентной ставки, ставки рефинансирования, стоимости строительства квадратного метра, потребительской стоимости бензина, потребительской стоимости природного газа.

Таким образом, был проведен корреляционный анализ и разработаны регрессионные модели для показателя стоимости первичного и вторичного жилья г. Челябинска. Статистические характеристики полученных уравнений свидетельствуют об их значимости, то есть возможности их использования в качестве основы для получения прогноза изменения стоимости недвижимости.