



مرکز مطالعات و تحقیقات پیشرفته  
تهران

پروژه درس اقتصاد سنجی ۱

## بررسی عوامل موثر بر قیمت مسکن در شهر تهران

تهیه و تنظیم:

مسعود فهمیده - زهرا حسنی - فائزه اسکندری

استاد راهنما

دکتر لیلا شیوا

زمستان ۹۹

## صفحه

## فهرست مطالب

فصل اول مقدمه و مرور ادبیات موضوع.....	۳
مقدمه و مروری بر ادبیات موضوع:.....	۴
فصل دوم داده های مورد مطالعه.....	۵
۱. داده های مورد مطالعه.....	۶
۲. انجام عملیات Clean data:.....	۶
۳. نکاتی در مورد متغیرهای توضیحی:.....	۷
فصل سوم برآورد مدل و نتایج.....	۱۰
۱. مدل اولیه.....	۱۱
۲. نتایج اولیه.....	۱۱
۳. تفسیر ضرایب.....	۱۲
فصل چهارم تست ها و آزمون های آماری.....	۱۴
۱. آزمون های آماری:.....	۱۵
۱.۱ آزمون معنی داری:.....	۱۵
۲.۱ آزمون واریانس نا همسانی:.....	۱۵
۳.۱ آزمون endogeneity:.....	۱۶
فصل پنجم نتیجه گیری و تحلیل نتایج.....	۱۷
تحلیل نتایج:.....	۱۸
منابع و مراجع:.....	۱۹

## فصل اول

### مقدمه و مرور ادبیات موضوع

## مقدمه و مروری بر ادبیات موضوع:

یکی از کلیدی ترین و اصلی ترین بخش های هر اقتصاد، بخش مسکن است که می تواند بخش های زیادی را با خود همراه کند. مطالعه بازار مسکن هم از جهت اینکه یک دارایی است و هم از لحاظ خدماتی که به عنوان سرپناه ارائه می دهد حائز اهمیت است. قیمت مسکن از مهم ترین عواملی است که در تصمیم گیری برای خانه دار شدن و خرید خانه تاثیرگذار است. آگاهی از روند تغییرات بازار مسکن و همچنین شناخت اثر عوامل مختلف بر قیمت مسکن به پیش بینی و سنجش رفتار بازار مسکن و در نتیجه تصمیم گیری در ست می انجامد. تحرک و رونق در بخش مسکن، باعث رونق بسیاری از بخش های یک اقتصاد می شود. در اقتصاد ایران، بازار مسکن دارای جایگاه ویژه ای است و این بازار دارای نوسانات خاصی می شود که می توان با شناخت عوامل موثر بر قیمت مسکن، این نوسانات را پیش بینی کرد. بخش مسکن و قیمت آن از عوامل بسیاری تاثیر پذیر است. افزایش سریع قیمت مسکن در سال های اخیر در ایران به عنوان یک دغدغه ملی برای مردم و دولتمردان مطرح بوده و به همین دلیل، مطالعه عوامل موثر بر افزایش قیمت مسکن بسیار با اهمیت است. در این بین تحولات قیمت مسکن در شهر تهران بیشترین توجهات را به سمت خود معطوف ساخته است. انتخاب محدوده مکانی شهر تهران به لحاظ روش شناختی نیز صحیح تر است، زیرا بازار مسکن در هر منطقه علاوه بر اثرپذیری از فاکتورهای ملی بطور وسیعی تحت تاثیر عوامل محلی نیز قرار می گیرد که تحلیل عملکرد بازار مسکن در سطح منطقه ای را موجه تر می سازد. با توجه به اینکه بازار مسکن ارتباط و پیوند عمیقی با متغیرهای اقتصاد کلان دارد در این پروژه ضمن ثابت فرض کردن کلیه متغیرهای اقتصادی مربوط به اقتصاد کلان، صرفاً به عنوان یک پروژه درسی مربوط به اقتصاد سنجی اثر برخی متغیرهای توضیحی بر قیمت مسکن مورد مطالعه و بررسی قرار گرفته است.

ضمن اینکه در سال های اخیر مطالعات زیادی در داخل و خارج از کشور جهت بررسی عوامل موثر بر قیمت مسکن انجام گرفته است که موضوع مسکن را از جوانب مختلف مورد تحلیل قرار داده اند. با توجه به تعدد مطالعات انجام شده و ترجیح بر اختصار صرفاً از کلیات مربوط به نتایج بدست آمده از این مطالعات در نگارش پروژه بهره گرفته شده است.

## فصل دوم

### داده های مورد مطالعه

## داده های مورد مطالعه

یکی از بزرگترین مشکلات بازار مسکن این است که به لحاظ اطلاعاتی مملو از عدم تقارن ها و نواقص است. بخشی از این عدم شفافیت و کاستی های اطلاعاتی به نبود یک بازار عمومی حاوی اطلاعات مربوط به قیمت ها و مبادلات (مشابه بازار بورس) باز می گردد. فقدان یک بازار عمومی حاوی اطلاعات روز آمد، سیال، قابل اطمینان، در دسترس و وسیع پیرامون مبادلات و مشخصا قیمت ها، منجر به هزینه های بالای جمع آوری اطلاعات به منظور تصمیم گیری و زمینه ساز عدم تقارن اطلاعاتی بین عاملان فعال در بازار است.

در این پروژه، با استفاده از داده های مربوط به معاملات مسکن وزارت راه و شهرسازی اثرات برخی متغیرهای مهم توضیحی از جمله مساحت، عمر بنا، نوع اسکلت ساختمان، موقعیت مکانی و ... بر قیمت مسکن در شهر تهران مورد بررسی قرار گرفته است. برای تجزیه و تحلیل تاثیر متغیرهای فوق بر قیمت مسکن، از داده های فصلی مربوط به تابستان ۱۳۹۸، استفاده شده است.

جدول زیر لیست متغیرهای موجود در داده های خام را نشان می دهد که با توجه به نیاز پروژه عملیات data cleaning بر روی این متغیر ها انجام گرفت.

جدول ۱-۲ متغیر های توضیحی داده های خام

کد قرارداد	نوع قرارداد	استان	شهرستان	منطقه شهرداری	نوع کاربری	قیمت یک مترمربع
قیمت	عمر بنا	نوع اسکلت	مساحت	نوع ملک	تاریخ ثبت قرارداد	شش رقم نخست کد پستی

## انجام عملیات Clean data :

- حذف برخی از متغیرهای که نقشی در پروژه نداشتند مانند کدپستی، کد قرارداد، نوع قرارداد و ...
- همگن کردن دیتا ها
- تعریف data type مربوط به هر متغیر توضیحی
- حذف و اصلاح داده های مشکل دار
- تعریف متغیرهای جدید

در نهایت لیست دیتاهای مورد استفاده در پروژه به شرح زیر می باشد:

جدول ۲-۲ متغیرهای توضیحی مورد استفاده در پروژه

Variable name	Storage type	Display format	variable label
city	str6	%9s	Name of city
region	float	%8.0g	Number of Region
sqrft	float	%8.0g	size of house in square feet
price	double	%8.0g	the total price of the house, 1000s
price_prsm	double	%8.0g	price of one square meter, 1000s
age	float	%8.0g	building age
houseframe	byte	%8.0g	=1 if material of building frame be steal =0 if material of building frame be concrete
ipollution	int	%8.0g	the amount of air pollution
traficpr	byte	%8.0g	=1 if region in Traffic plan range
lprice	float	%8.0g	log(price)
big_house	byte	%9.0g	=1 if sqrft>80
new_house	byte	%8.0g	=1 if age<5   age==5
middle_house	byte	%8.0g	=1 if age>5   age<15   age==15
old_house	byte	%8.0g	=1 if age>15

### نکاتی در مورد متغیرهای توضیحی:

#### • متغیر traficpr:

این متغیر دامی قیمت خانه های قرار گرفته در منطقه ویژه طرح ترافیک را با خانه هایی که در طرح ترافیک قرار نگرفته اند مورد مقایسه قرار می دهد. با توجه به نقشه طرح ترافیک تهران مناطق ۶، ۱۱، ۱۲ در محدوده ویژه طرح ترافیک قرار گرفته اند.

#### • متغیر ipollution:

برای تعریف این متغیر توضیحی، با توجه به میانگین سالانه آلودگی هوا در مناطق مختلف، شهر تهران را به چهار دسته مختلف:

۱- پاک      ۲- نسبتا سالم      ۳- ناسالم      ۴- خیلی ناسالم

دسته بندی نمودیم و مناطق مختلف را در این چهار گروه قرار دادیم.

• متغیر big house :

با توجه به اینکه تهران بصورت تقریبی گران ترین شهر ایران از نظر قیمت مسکن است ، گزارش های داده شده نشان می دهد که تقاضا برای خانه های کوچک و زیر ۸۰ متر با درصد قابل توجهی بیشتر از خانه های بزرگ است. به همین دلیل قیمت خانه ها در مترهای کوچک به دلیل تقاضای زیاد بیشتر است. این متغیر دامی برای بررسی قیمت خانه ها بزرگ تر از ۸۰ متر و یا کوچکتر از ۸۰ متر هستند ایجاد شده است.

• متغیر age :

این متغیر جهت بررسی اثر عمر بنا بر قیمت مسکن در شهر تهران ایجاد شده است. با توجه به اینکه افزایش تقاضا برای خانه های نوساز بیشتر از خانه های قدیمی ساز است مجددا این متغیر نیز به سه دسته: new house , middle age house , old house تقسیم شده است و قیمت خانه را در هرکدام از این ۳ حالت مورد بررسی قرار می دهد.

سایر متغیرهای توضیحی با توجه به variable label کاملا واضح هستند و در جدول ۲-۲ بصورت کامل توضیح داده شده اند.

خلاصه ای از مقادیر برخی از متغیرهای توضیحی به شرح زیر می باشد.

جدول ۲-۳ summarize data

Variable	Obs	Mean	Std. Dev.	Min	Max
sqrft	2,421	94.20414	312.8636	20	13297
price	2,421	1.31e+07	2.00e+07	600000	4.69e+08
price_prsm	2,421	129978.7	72273.76	10571.43	512661.6
age	2,421	9.757951	9.698431	0	69
ipollution	2,421	2.116481	.9725897	1	4

- میانگین معاملات مسکن در شهر تهران در شهریور ۱۳۹۸ ، ۹۴,۲ متر مربع بوده است.
- میانگین قیمت معاملات مسکن در شهر تهران در شهریور ۱۳۹۸ ، ۱۳,۱ میلیارد ریال بوده است.
- میانگین قیمت هر متر مربع معاملات مسکن در شهر تهران در شهریور ۱۳۹۸ ، ۱۲۹ میلیون ریال بوده است.



- میانگین عمر بنا در معاملات مسکن معاملات مسکن در شهر تهران در شهریور ۱۳۹۸ ، ۹,۷۵ سال می باشد.

- با توجه به شاخص بندی انجام شده مناطق بر اساس آلودگی هوا میانگین آلودگی هوا در شهر تهران ۲۰۱۱ می باشد که نشان می دهد تهران شهری نسبتاً آلوده است.

## فصل سوم

### برآورد مدل و نتایج

## مدل اولیه

روش استفاده شده در این پروژه، حداقل سازی مربعات یا همان OLS است که با استفاده از این روش می توان به تخمینگرهای درستی دست یافت.

$$\log(\widehat{\text{price}}) = \beta_0 + \beta_1 \text{sqrft} + \beta_2 \text{taficpr} + \beta_3 \text{ipollution} + \beta_4 \text{housefram} + \beta_5 \text{age} + u$$

مدل اصلی ما در این پروژه، مدل ذکر شده در سطر بالا می باشد که شامل ۵ متغیر توضیحی اصلی می باشد و در ادامه پروژه مدل extend شده و همچنین تست های مختلف با توجه به این متغیرها مورد بررسی قرار می گیرد.

## برقراری فروض رگرسیون:

- ۱- مدل نسبت به پارامترهای مدل دارای رابطه خطی است.
- ۲- نمونه کاملاً بصورت تصادفی انتخاب شده است.
- ۳- هیچگونه همخطی کامل بین متغیرها وجود ندارد
- ۴- مقدار انتظاری خطای مدل به شرط آنکه پارامترهای مدل در مدل وجود داشته باشد برابر ۰ است.
- ۵- فرض واریانس ناهمسانی در ادامه مورد بررسی قرار می گیرد.

## نتایج اولیه

در این قسمت ما سه رگرسیون مختلف را محاسبه نمودیم. مدل (۱) همان مدل اصلی و ساده است که اثر متغیرهای توضیحی مختلف را بر قیمت مسکن در شهر تهران نشان می دهد. در مدل (۲) متغیرمساحت تبدیل به متغیر دامی شده است و نشان دهنده افزایش قیمت در خانه های بالای ۸۰ متر نسبت به خانه های زیر ۸۰ متر و کوچک است. در مدل (۳) در مورد عمر بنا ۳ متغیر دامی در نظر گرفته شده است که شامل خانه های نوساز زیر ۵ سال، خانه های میان رده بین ۵ سال تا ۱۵ سال و خانه های قدیمی بالای ۱۵ سال است و ما با آوردن این متغیرها در مدل با در نظر گرفتن خانه های میان رده به عنوان base group تغییرات قیمت را در حالت های مختلف مورد بررسی قرار دادیم و در جدول صفحه بعد گزارش ضرایب بدست آمده به همراه standard error های هر ضریب گزارش شده است.

جدول ۱-۳ نتایج رگرسیون

**dependent variable: log(price)**

Independent variable	(1)	(2)	(3)
sqrft	0.00024 (0.00010)	-----	.0002479 .0000189
trafficpr	0.3201 (0.02470)	.2208082 .0178214	.3163538 .0215319
Ipollution	-0.2226 (0.00683)	-.1605263 .0059115	-.2222849 .0069661
houseframe	-0.0702 (0.13387)	-.04400 (.0105284)	-.0578026 .0128186
age	-0.0051 (0.00083)	-.0027521 .000529	-0.0051 (0.00083)
Big_house	-----	.4168299 .0102462	-----
new_house	-----	-----	.1535913 .0139354
Middle_age_house	-----	-----	-----
old_house	-----	-----	-.0294918 .015856
intercept	7.426 (0.02039)	7.115075 .015717	7.313904 .0177343
Observation	2421	2421	2421
R-squared	0.36	0.60	0.40

**تفسیر ضرایب**

- تفسیر ضریب متغیر  $\text{sqrft}$  :

در صورت کنترل کردن سایر متغیرها با افزایش یک مترمربع مساحت ساختمان، قیمت  $0.00024$  درصد افزایش می یابد.

- تفسیر ضریب متغیر  $\text{trafficpr}$

در صورت کنترل کردن سایر متغیرها افزایش قیمت مسکن در ساختمان هایی که در محدوده طرح ترافیک قرار دارند  $0.32$  درصد بیشتر است.

• تفسیر ضریب متغیر ipolltion :

در صورت ثابت بودن سایر متغیر ها با افزایش یک واحد شاخص آلودگی هوا ، احتمال افزایش قیمت ۲۲,۲ درصد کاهش می یابد .

• تفسیر ضریب houseframe :

در صورت ثابت بودن سایر متغیر ها ، افزایش قیمت در ساختمان هایی با جنس بدنه فلزی نسبت به ساختمان هایی با جنس بدنه بتونی ۷,۰ درصد کمتر است.

• تفسیر ضریب age :

در صورت ثابت بودن سایر متغیر ها با افزایش یک واحد سن ساختمان ، قیمت ۰,۵ درصد کاهش میابد .

• تفسیر ضریب big \_ house :

در صورت ثابت بودن سایر متغیر ها ، افزایش قیمت در ساختمان هایی با متراژ بالا تر از ۸۰ متر مربع ۴۱,۰ درصد نسبت به ساختمان هایی با متراژ زیر ۸۰ متر بیشتر است .

• تفسیر ضریب new \_house :

در صورت ثابت بودن سایر متغیر ها افزایش قیمت در ساختمان هایی با سن بنای زیر ۵ سال ۱۵,۳ درصد نسبت به ساختمان هایی با سن بنای بین ۵ تا ۱۵ سال بیشتر است.

• تفسیر ضریب old \_house :

در صورت ثابت بودن سایر متغیر ها افزایش قیمت در ساختمان هایی با سن بنای بالاتر از ۱۵ سال ۲,۹ درصد نسبت به ساختمان هایی با سن بنای زیر ۱۵ سال کمتر است.

## فصل چهارم

### تست ها و آزمون های آماری

## آزمون های آماری:

### آزمون معنی داری:

این آزمون برای معنی دار بودن ضرایب در سطوح اطمینان مختلف مورد استفاده قرار می گیرد. ساده ترین نوع آزمون معناداری t-test می باشد. تست های دیگری مانند F تست نیز برای حالتی که کورولیشن زیاد بین متغیرهای توضیحی برقرار باشد مورد استفاده قرار می گیرد. با توجه به مقادیر ضرایب و همچنین se هر ضریب می توان برای جدول شماره ۳ معناداری ضرایب را مورد بررسی قرار داد. (در سطح معنی داری ۰.۵٪) پس از بررسی مشخص می گردد که تمام ضرایب در سطح ۵ درصد معنی دار هستند.

$$t = \frac{(\text{estimate} - \text{hypothesized value})}{\text{standard error}}$$

### آزمون واریانس ناهمسانی:

برای بررسی وجود واریانس ناهمسانی در مدل از white test استفاده می کنیم. بدین منظور باید به بررسی اینکه بین  $u^2$  و متغیرهای توضیحی مدل و همچنین intraction های بین آن ها correlation دارد ، پردازیم.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}$$

$$\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + \text{error}$$

$$H_0: \delta_1 = 0, \delta_2 = 0$$

بدین صورت رگرسیون متغیر وابسته (lprice) را روی متغیرهای توضیح دهنده (trafficpr, sqrfitt, age, ipollution houseframe) انجام می دهیم و همان طور که ذکر شد lpricehat مدل را ذخیره کرده و رگرسیون  $\text{residual}^2$  را روی lpricehat و  $\text{lpricehat}^2$  می نویسیم .

برای وجود واریانس ناهمسانی در مدل باید نشان دهیم ضرایب متغیرهای فوق الذکر در این رگرسیون معنادار بوده، بدین صورت f test را محاسبه می کنیم:

$$F = \frac{(R^2)/k}{1 - R^2/n - k - 1} = \frac{(0.1699)/2}{(1 - 0.1699)/2424 - 2 - 1} = 247.7580412$$

پس ضرایب معنادار بوده و مشکل واریانس ناهمسانی در مدل داریم.

در ادامه با استفاده از Weighted Least Squares به ضرایبی از متغیرها دست پیدا می کنیم که مشکل واریانس ناهمسانی ندارند.

### آزمون endogeneity :

برای تست این موضوع که آیا در مدل Functional Form Misspecification داریم یا خیر، از این RESET تست استفاده می کنیم:

بدین منظور برای شروع برای اینکه نمی دانیم این مشکل ناشی از عدم وجود فرم quadratic متغیر است یا ناشی از interaction متغیرها و تست این موضوع با درجه آزادی زیادی برای f test مربوطه همراه است ، به شرح زیر عمل می کنیم:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \delta_1 \hat{y}^2 + \delta_2 \hat{y}^3 + \text{error}$$

$$\delta_1 = \delta_2 = 0$$

ابتدا lpricehat<sup>2</sup> و lpricehat<sup>3</sup> که fitted values مدل هستند را ذخیره می کنیم و مجدد رگرسیون مدل را این بار با وجود این دو متغیر run می کنیم و با f test معناداری این دو ضریب را تست می کنیم:

$$F = \frac{(R_{ur}^2 - R_r^2)/q}{1 - R_{ur}^2/n - q - 1} = \frac{(0.4082 - 0.3570)/2}{(1 - 0.4082)/2424 - 5 - 1} = 104.59$$

و با توجه به معنی داری ضرایب پس مشکل endogeneity داریم.



## فصل پنجم

### نتیجه گیری و تحلیل نتایج

## تحلیل نتایج :

با توجه به تفسیر ضرایب بدست آمده در پروژه میتوان اینگونه برداشت کرد که در بین پارامترها و متغیرهای توضیحی حاضر در پروژه همانگونه که انتظار هم داشتیم بین مترآژ و قیمت مسکن رابطه مستقیم وجود دارد ، از منظر نوع سازه خانه های بتنی دارای قیمت بالاتری هستند ، قرارگرفتن خانه در مرکز شهر و محدوده طرح ترافیک باعث افزایش قیمت مسکن می شود که میتوان این امر را اینگونه توجیه کرد که این منازل از نظر دسترسی و همچنین عدم نیاز به خرید طرح ترافیک دچار مزیت هستند در کنار اینکه ممکن است از نظر شلوغی و ترافیک زیاد درگیر مشکل باشند ، همچنین خانه هایی که در مناطق آلوده شهر از نظر کیفیت شاخص هوا قراردارند قیمت هایی پایینتر نسبت به خانه های مناطق خوش آب و هوا تر دارند. نکته ای که قابل ذکر است در عمل و واقعیت به دلیل قیمت های بالا در شهر تهران تقاضا برای خانه های مترآژ پایین بیشتر از خانه های مترآژ بالا است و نکته ای که ما در داده ها به آن برخورد کردیم این بود که به عنوان مثال قیمت واحد ۷۰ متری در منطقه ۳ از نظر هر متر مربع با خانه ۲۵۰ متری در منطقه ۱ شهر تهران به عنوان گرانترین منطقه برابری می کرد. در مورد عمر بنا نیز طبیعتا با افزایش سن بنا قیمت مسکن دچار افت می شود ولی برای بررسی دقیق تر ما ضمن دسته بندی عمر بنا در ۳ بخش خانه های نوساز و میان سال و قدیمی به این نتیجه رسیدیم که افت قیمت در خانه های نوساز نسبت به خانه های قدیمی کمتر است و معمولا خانه های بین ۰ تا ۵ سال در یک رده سنی و قیمتی در هر منطقه و خانه های بین ۵ تا ۱۵ سال در یک رده سنی و قیمتی در هر منطقه قرار می گیرند ، برای خانه های بالای ۱۵ سال و قدیمی ساز که شامل وام مسکن هم نمی شوند این افت قیمت شدت بیشتری پیدا می کند.

## منابع و مراجع:

- [1] Campbell, John Y and Cocco, Joao F (2004). How Do House Price Affect Consumption? Evidence from Micro Data. Harvard Institute of Economic Research .Discussion Paper No.2045
- [2] De La paz ,PalomaTaltavul (2003), Determinants of Housing Prices in Spanish Cities, Journal Investment and Finance, vol.21(2),109-137.
- [3] Abelson, P, Joyeux, R., Milunovich, G. and Chung, D. (2005). Explaining House Prices in Australia: 1970-2003 , Economic Record, 81(1).
- [۴]- میرحسین موسوی ، حسین درودیان - تحلیل عوامل موثر بر قیمت مسکن در شهر تهران- فصلنامه مدلسازی اقتصادی، شماره ۳۱
- [۵]- سید منصور خلیلی عراقی ، محسن مهرآرا ، سید رضا عظیمی - بررسی عوامل موثر بر قیمت مسکن در ایران با استفاده از داده های ترکیبی - فصلنامه پژوهش ها و سیاست های اقتصاد ، پاییز ۹۱
- [۶]- خیابانی، ناصر ، " عوامل تعیین کننده قیمت مسکن در ایران "، فصلنامه اقتصاد مسکن ، شماره ۳۸