

گزارش ۳- نمودار پراکنش ، ضریب همبستگی (کمی) و رگرسیون (ساده)

برای متغیرهای زیر تعداد ۱۱۷۲ پرسشنامه تکمیل شده است (نتایج در CD).

متغیر	توضیحات	شرح	وضعیت ها
value	Property (Tax) Value	ارزش املاک- مالیاتی	--
view	view	چشم انداز	۰ = خیر ۱ = بله
constrql	House Construction Quality	کیفیت ساخت و ساز خانه	۱ = بدترین ۶ = بهترین
age	age (in years)	سن خانه (به سال)	--
roofstyl	roof style	جنس سقف	۱ = سرامیک ۲ = چوب ۳ = آسفالت
basement	Basement?	وضعیت زیر زمین	۱ = زیر زمین ۲ = فضای انباری
aircond	Air Conditioner type	نوع تهویه مطبوع	۱ = ندارد ۲ = پنجره ۳ = تهویه مرکزی
attic	Attic type	نوع اتاق زیرشیروانی	۱ = مبله کامل ۲ = بدون مبل ۳ = فاقد اتاق زیرشیروانی
rooms	Number of rooms	تعداد اتاق	--
bedrooms	Number of bedroms	تعداد اتاق خواب	--
bathroom	Number of bathrooms	تعداد حمام	--
fireplac	Number of fireplaces	تعداد شومینه	--
homesqft	Square footage of house	اندازه فضای خانه- متر مربع	--
lotsize	Lot size (square feet)	اندازه زمین	--
frontage	Lake frontage (in feet)	اندازه استخر	--
income	Average household income (by city)	میانگین درآمد خانوار	--
ac_dum	Air Conditioning Dummy Variable	داشتن تهویه مطبوع	۰ = خیر ۱ = بله
atticdum	Attic Dummy Variable	تنوع ساخت اتاق زیرشیروانی	۰ = خیر ۱ = بله
roof_dum	Asphalt Shingle Roof Dummy Variable	آسفالت سقف	۰ = خیر ۱ = بله
filter_\$	aircond = ۱ (FILTER)	فیلتر تهویه مطبوع	۰ = انتخاب نشده ۱ = انتخاب شده
slat_dum	slate dummy variable	تنوع ساخت سنگ	۰ = خیر ۱ = بله
wood_dum	wood dummy variable	تنوع ساخت چوب	۰ = خیر ۱ = بله
asph_dum	asph dummy variable	تنوع ساخت آسفالت	۰ = خیر ۱ = بله
ac_adj	A/C with reassigned values	تهویه مطبوع با ارزش معین	۰ = خیر ۱ = بله

در ادامه گزارش قبلی که اطلاعات مربوط به ۱۱۷۲ ساختمان براساس متغیرهای جدول فوق گردآوری شده است، به دنبال یافتن رابطه میان برخی از متغیرها هستیم؛ لذا سؤالاتی به شرح زیر مطرح می کنیم:

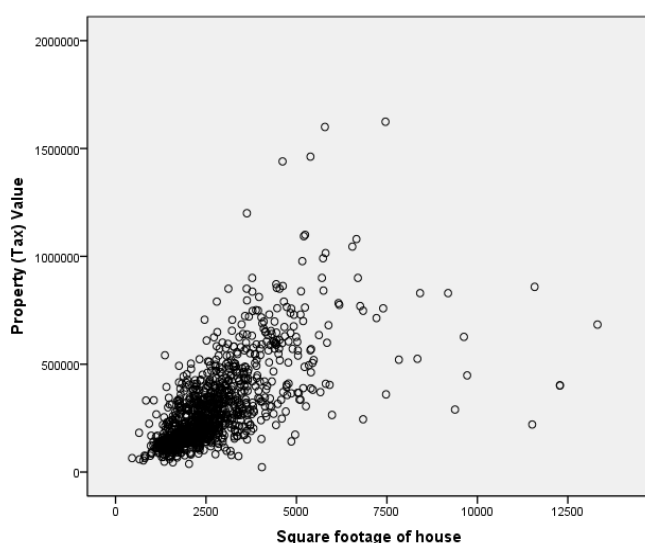
❖ سؤالات

۱. نمودار هندسی رابطه میان اندازه خانه (homesqft) و قیمت خانه (value) چگونه می باشد؟
۲. چه نوع رابطه ای میان اندازه خانه (homesqft) و قیمت خانه (value) وجود دارد؟
۳. آیا رابطه میان اندازه خانه (homesqft) و قیمت خانه (value) براساس مدل ریاضی قابل تعریف می باشد؟

❖ برای پاسخ به سوال ۱ از نمودار پراکنش، سوال ۲ ضریب همبستگی و سوال ۳ رگرسیون استفاده می کنیم.

بخش اول : نمودار پراکنش

برای رسم نمودار پراکنش گزینه Scatter/Dot را از منوی Graphs انتخاب می کنیم و پس از معرفی متغیرهای homesqft و Value نمودار زیر نمایش داده می شود:



تفسیر:

در نمودار پراکنش تعدادی نقاط مجزا از هم در صفحه رسم شده است. در نظر کلی این نقاط تشکیل رابطه خطی مستقیمی می دهند به طوری که می توان گفت با افزایش متراژ ساختمان (homesqft) قیمت ساختمان (value) نیز افزایش می یابد.

بخش دوم : ضریب همبستگی

ضریب همبستگی ابزاری آماری برای تعیین نوع و درجه رابطه یک متغیر کمی با متغیر کمی دیگر است. ضریب همبستگی، یکی از معیارهای مورد استفاده در تعیین همبستگی دو متغیر است. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می دهد. این ضریب بین ۱ تا -۱ است و در عدم وجود رابطه بین دو متغیر، برابر صفر است. ضریب همبستگی به دو روش پیرسون (داده های کمی) و اسپیرومن (داده های کیفی) قابل محاسبه می باشد.

❖ ضریب همبستگی پیرسون

ضریب همبستگی پیرسون (Pearson Correlation Coefficient)، روشی پارامتری است و برای داده‌هایی با توزیع نرمال یا تعداد داده‌های زیاد استفاده می‌شود. این ضریب برای داده‌های کمی کاربرد دارد. در پاسخ به سوال ۲ از این ضریب استفاده می‌شود. برای محاسبه ضریب همبستگی پیرسون مراحل زیر را انجام می‌دهیم:

۱. ابتدا از منوی **Analyze** به ترتیب گزینه‌های **Correlate** و **Bivariate** را انتخاب می‌کنیم.

۲. گزینه **Pearson** را انتخاب می‌کنیم.

۳. متغیرهای مورد نظر را در قسمت **Variables** به نرم افزار معرفی می‌کنیم. پس از تایید پنجره نتایج به

شرح زیر نمایش داده می‌شود:

Correlations			
		Property (Tax) Value	Square footage of house
Property (Tax) Value	Pearson Correlation	۱	.۶۵۱**
	Sig. (۲-tailed)		.۰۰۰
	N	۱۱۷۲	۱۱۷۲
Square footage of house	Pearson Correlation	.۶۵۱**	۱
	Sig. (۲-tailed)	.۰۰۰	
	N	۱۱۷۲	۱۱۷۲

** . Correlation is significant at the ۰.۰۱ level (۲-tailed).

تفسیر:

- ضریب همبستگی محاسبه شده برابر با ۰.۶۵۱ و $\text{Sig}=۰.۰۰۰$ می‌باشد.
- این مقدار همبستگی نشان می‌دهد که براساس اطلاعات داده شده بین دو متغیر اندازه ساختمان و قیمت ساختمان همبستگی مثبت وجود دارد.
- مقدار $\text{Sig}=۰.۰۰۰$ بیانگر این است که با ضریب همبستگی محاسبه شده از نظر آماری معنادار می‌باشد چرا که فرض H_0 رد می‌شود. به عبارت دیگر چون رقم Sig کمتر از سطح معناداری ۰.۰۱ می‌باشد لذا ۱٪ احتمال دارد که همبستگی محاسبه شده به سبب خطای نمونه گیری و تصادفی باشد.

رگرسیون از جمله آزمونهای رایج و معروف در آمار تحلیلی می باشد و هدف از آن بررسی و پیش بینی رفتار یک متغیر (متغیر وابسته) با استفاده از یک یا چندین متغیر (متغیر مستقل) می باشد. چنانچه در رگرسیون فقط از یک متغیر مستقل استفاده شود به آن رگرسیون ساده می گویند اما اگر تعداد متغیرهای مستقل از یک متغیر بیشتر باشد آن را رگرسیون چندگانه می گویند.

❖ در ادامه برای تعیین معادله رگرسیون خطی ساده مراحل زیر را انجام می دهیم:

۱. ازمونوی **Analyze** به ترتیب گزینه های **Regression** و **Linear** را انتخاب می کنیم.
۲. متغیر قیمت ساختمان (**Value**) را به عنوان متغیر وابسته در بخش **Dependent** و متغیر اندازه ساختمان (**homesqft**) را به عنوان متغیر مستقل در بخش **Independent** به نرم افزار معرفی می کنیم.
۳. پس از تایید پنجره نتایج آزمون به شرح زیر نمایش داده می شود که تفسیر مربوطه در ادامه بیان می شود.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
۱	.۶۵۱۸	.۴۲۳	.۴۲۳	۱۴۸۱۵۱,۰۱۰

a. Predictors: (Constant), Square footage of house

تفسیر:

- در جدول **Model Summary** مقدار $R=۰,۶۵۱$ می باشد که همان ضریب همبستگی میان دو متغیر است. همچنین مقدار $R^2=۴۲\%$ می باشد که بیانگر این است که متغیر اندازه ساختمان (متغیر پیش بینی کننده یا مستقل) توانسته است ۴۲٪ از تغییرات متغیر قیمت ساختمان (متغیر ملاک یا وابسته) را تبیین کند.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
۱	Regression	۱,۸۸۵E۱۳	۱	۱,۸۸۵E۱۳	۸۵۸,۵۹۶	.۰۰۰ ^a
	Residual	۲,۵۶۸E۱۳	۱۱۷۰	۲,۱۹۵E۱۰		
	Total	۴,۴۵۳E۱۳	۱۱۷۱			

تفسیر:

- در جدول **ANOVA** درمورد وجود رابطه خطی بین متغیرها آزمون صورت گرفته است. مقدار $\text{Sig}=۰,۰۰۰$ می باشد که بیانگر رد فرض H_0 و وجود رابطه خطی بین متغیر وابسته و مستقل می باشد.

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	۳۹۹۸۳,۵۳۲	۹۶۴۴,۲۴۱		۴,۱۴۶	.۰۰۰
Square footage of house	۹۳,۸۲۸	۳,۲۰۲	.۶۵۱	۲۹,۳۰۲	.۰۰۰

a. Dependent Variable: Property (Tax) Value

تفسیر:

- در جدول **Coefficientsa** محاسبات مقداری رگرسیون ارائه شده است.
- مقدار ثابت برابر با ۳۹۹۸۳.۵۳۲ می باشد و با توجه به مقدار **Sig=۰,۰۰۰** می توان نتیجه گرفت فرض **H**. رد می شود و معناداری مقدار ثابت تایید می گردد.
- ضریب رگرسیون متغیر مستقل برابر با ۹۳,۸۲۸ می باشد که با توجه به **Sig=۰,۰۰۰** معناداری آن تایید می گردد.
- با استفاده از نتایج آزمون می توان مدل خط رگرسیون را به شرح زیر نوشت:

$$Y = 39983.532 + 93.828 \times X$$

Y = مقدار متغیر وابسته - قیمت ساختمان

X = مقدار متغیر مستقل - اندازه ساختمان

پایان