



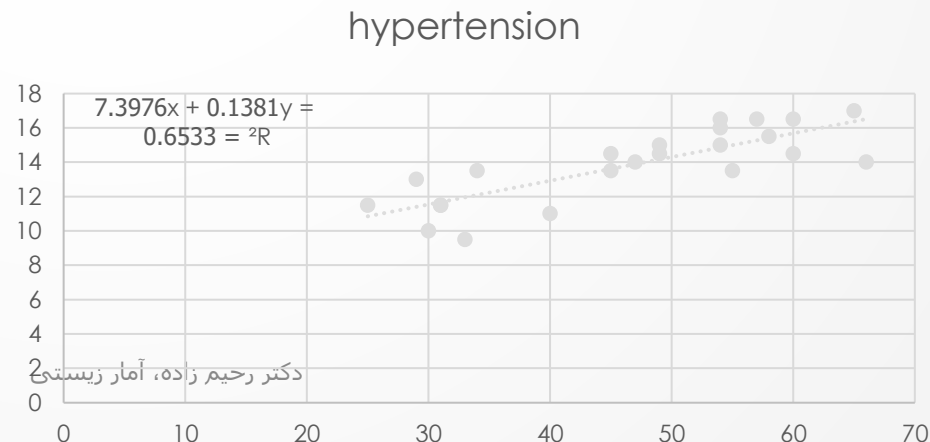
رگرسیون خطی با متغیر کیفی دو
حالته در اکسل

رگرسیون خطی

age	hypertension
31	11.5
29	13
34	13.5
40	11
33	9.5
49	14.5
54	15
60	14.5
57	16.5
45	13.5
66	14
30	10
47	14
55	13.5
31	11.5
58	15.5
60	16.5
65	17
54	16
45	14.5
54	16.5
49	15
25	11.5

در این مثال رابطه بین سن و فشار خون مورد نظر است. در نمونه ای به حجم 23 نفر را که فشارخون و سن افراد را اندازه گیری کرده ایم.

در صورتی که فقط یک متغیر مستقل داشته باشیم فقط کافی است متغیر مستقل را اول و سپس متغیر وابسته را بعد از آن وارد کنیم. سپس از نمودارها استفاده کرده و نمودار پراکنش را رسم نماییم با کلیک راست روی نقاط نمودار پراکنش روی گزینه Add Trendline کلیک کرده در منوی Format Chart Area می توان بر روی دو گزینه (Display R-square value on Chart و Display Equation on Chart) تیک زد تا هم معادله خط و هم ضریب تعیین روی نمودار نشان داده شود.



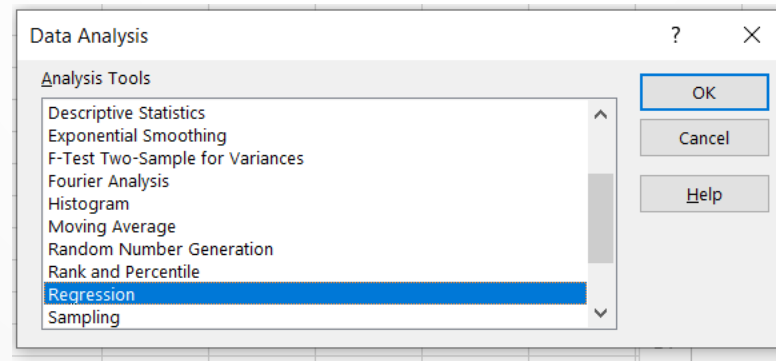
رگرسیون خطی با متغیر کیف دو حالت

age	gender	hyperten sion
31	0	11.5
29	0	13
34	0	13.5
40	0	11
33	0	9.5
49	0	14.5
54	0	15
60	0	14.5
57	0	16.5
45	0	13.5
66	0	14
30	1	10
47	1	14
55	1	13.5
31	1	11.5
58	1	15.5
60	1	16.5
65	1	17
54	1	16
45	1	14.5
54	1	16.5
49	1	15
25	1	11.5

وقتی یک متغیر مستقل دو حالت مثل جنسیت داریم امکان به دست آوردن معادله رگرسیون به روش قبلی و از روی نمودار امکان پذیر نمی باشد. در این حالت باید از منوی Data Analysis و گزینه رگرسیون استفاده کنیم.

در مثال قبلی یک ستون برای جنسیت اضافه کرده ایم.

برای برازش مدل رگرسیونی که در آن متغیر فشار خون وابسته و متغیرهای سن و جنسیت مستقل باشند. در منوی Data Analysis ، گزینه رگرسیون را انتخاب می کنیم.



منوی رگرسیون

با کلیک روی Regression پنجره زیر باز می شود. در این پنجره آدرس داده های را وارد می کنیم و دکمه Ok را می زنیم تا خروجی در Work Sheet جدید ایجاد شود.

Regression

Input

Input Y Range:

Input X Range:

☐ Labels ☐ Constant is Zero

☐ Confidence Level: %

Output options

☐ Output Range:

☒ New Worksheet Ply:

☐ New Workbook

Residuals

☐ Residuals ☐ Residual Plots

☐ Standardized Residuals ☐ Line Fit Plots

Normal Probability

☐ Normal Probability Plots

OK Cancel Help

Excel OUTPUT

SUMMARY OUTPUT	
<i>Regression Statistics</i>	
Multiple R	0.8222
R Square	0.676013
Adjusted R	0.643614
Standard E	1.284198
Observatic	23

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regressior	2	68.82104	34.41052	20.86541	1.27432E-05
Residual	20	32.98331	1.649166		
Total	22	101.8043			

تفسیر ضرایب در رگرسیون خطی

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	7.1857594	1.062273237	6.764511	1.4E-06	4.969896256	9.4016225
X Variable	0.135455114	0.021849617	6.199427	4.69E-06	0.089877612	0.1810326
X Variable	0.63792559	0.538780469	1.184018	0.250288	-0.485950774	1.761802

معنی داری متغیر
جنسیت از 0.05 کمتر
است، پس سن روی
فشار فشار خون اثر
معنی داری دارد. با یک
واحد افزایش سن فشار
خون 0.135 واحد افزایش
می یابد.

معنی داری متغیر جنسیت از 0.05 بیشتر است، پس فشار خون در زن و مرد با
هم تفاوت معنی داری ندارد. تفسیر متغیر کیفی کد یک در مقابل کد صفر اگر
کد یک مرد بودن باشد یعنی فشار خون مردان 0.638 واحد از زنان بیشتر است.