



Laboratuvar 04

BİL 210-Olasılık ve İstatistik

Regresyon

Regresyon Doğrusu

- Antropologlar, insan kalıntılarından boy tahmini yapabilmek için belirtilen **regresyon doğrusunu** kullanmaktadır:

$$y' = 61.4 + 2.4x$$

- Burada,
- **y'** : beklenen boy uzunluğunu ve
- **x** : uyluk uzunluğunu belirtmektedir.

Regresyon Doğrusu

- Regresyon doğrusunun y ile kesiştiği nokta: 61.4
- Eğimi: 2.4
- $y' = 61.4 + 2.4(0) = 61.4$
- Uyluk uzunluğu 50 cm olan bir kişinin beklenen boy uzunluğu
- $x=50$ için: $y' = 61.4 + 2.4(50) = 181.4$
- Regresyon doğrusunu $(0, 61.4)$ ve $(50, 181.4)$ noktalarını birleştirerek çizebiliriz.

Regresyon Doğrusu

- Regresyon doğrusunu Minitab'te oluşturmak için belirtilen verileri giriniz:

C1	C2
Uyluk Uzunluğu	Boy Uzunluğu
0	61,4
50	181,4

- Graph>Scatter Plot>Simple adımlarını takip ediniz.

Regresyon Doğrusu

Scatterplot - Simple

C1 Uyluk Uzunluğu
C2 Boy Uzunluğu

	Y variables	X variables
1	'Boy Uzunluğu	'Uyluk Uzunlu
2		
3		
4		
5		
6		
7		

Scale... Labels... Data View...

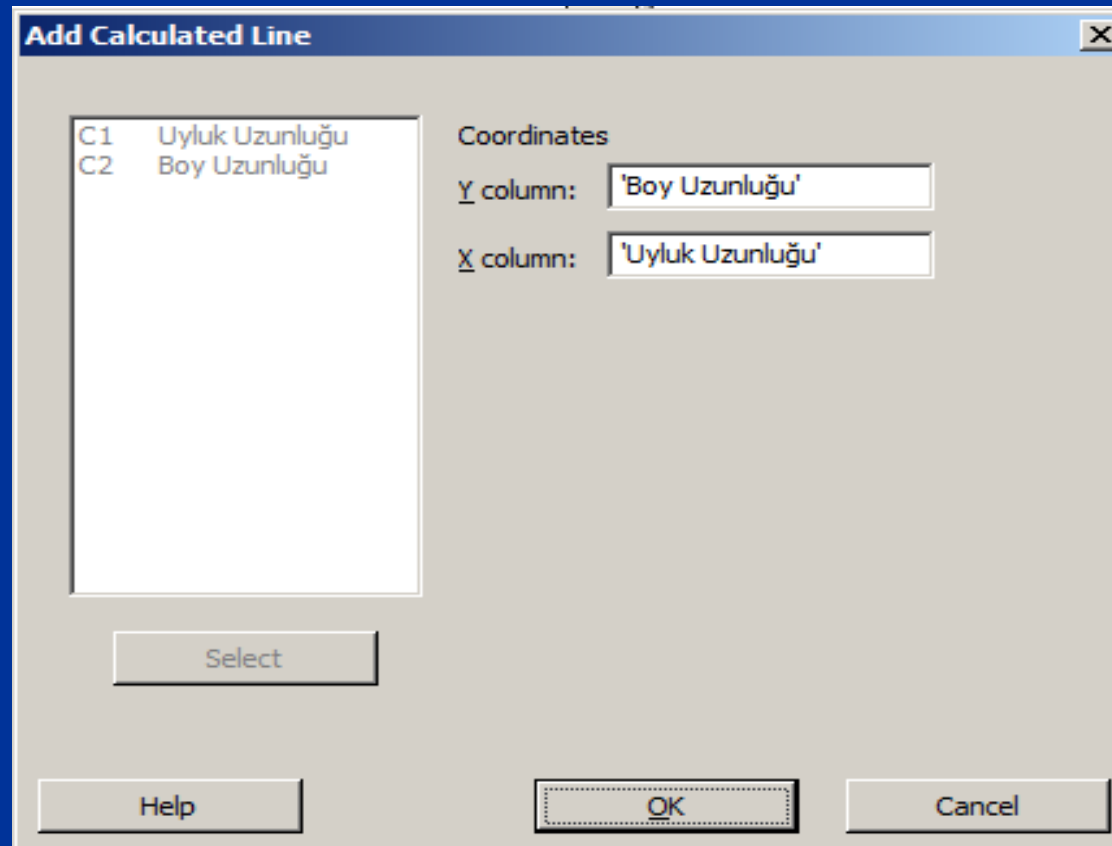
Multiple Graphs... Data Options...

Select

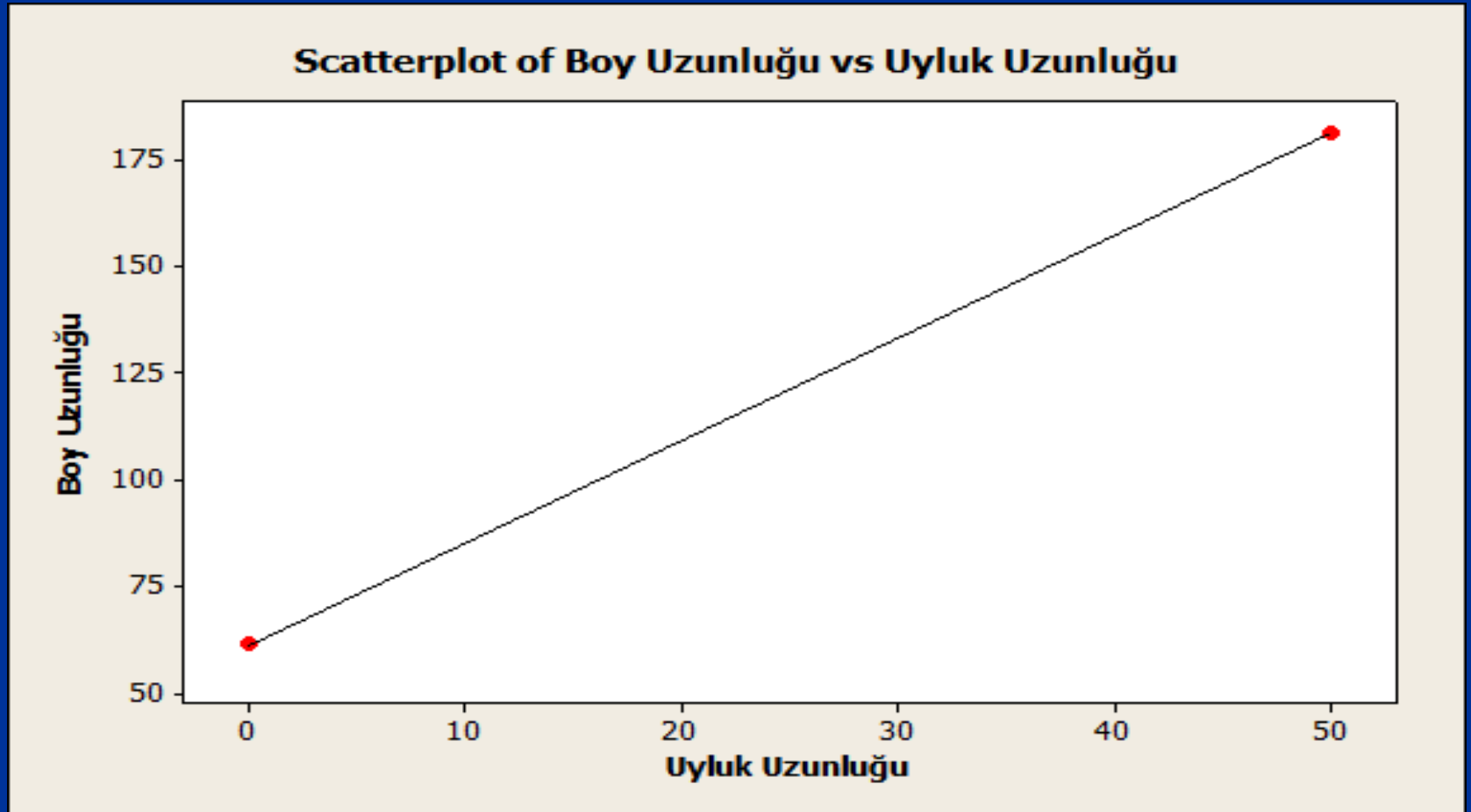
Help OK Cancel

Regresyon Doğrusu

- Ardından, oluşturduğunuz grafik üzerine sağ tıklayarak **Add>Calculated line** adımlarını izleyiniz.

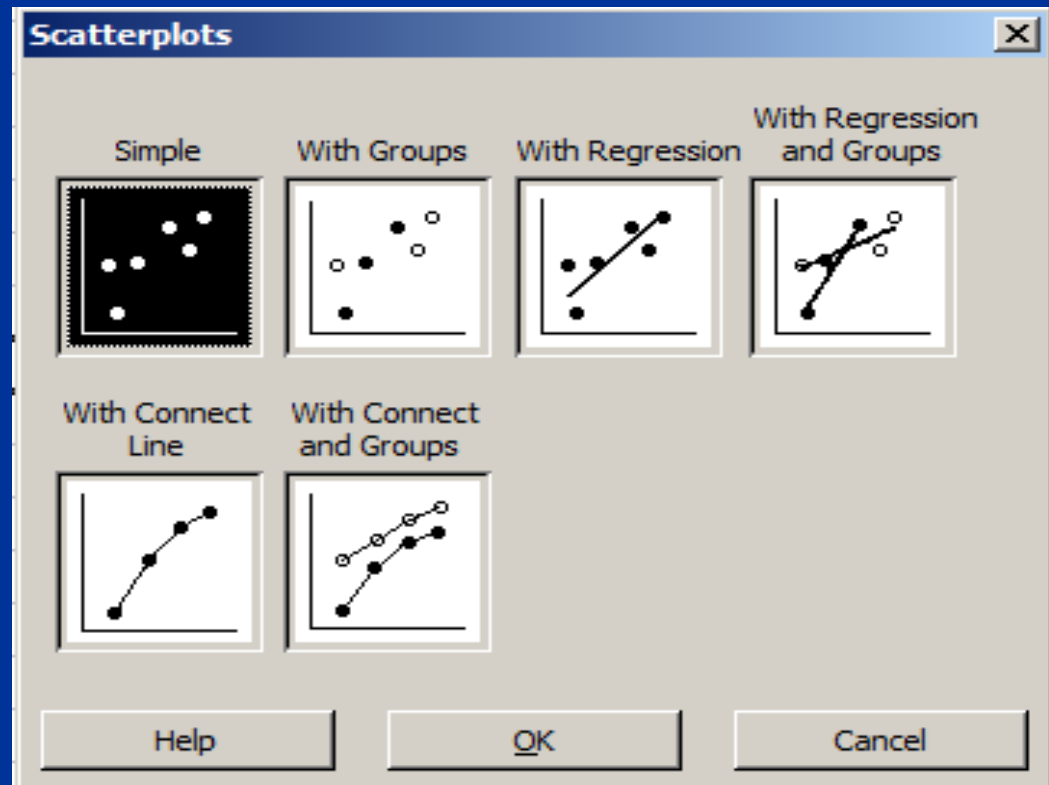
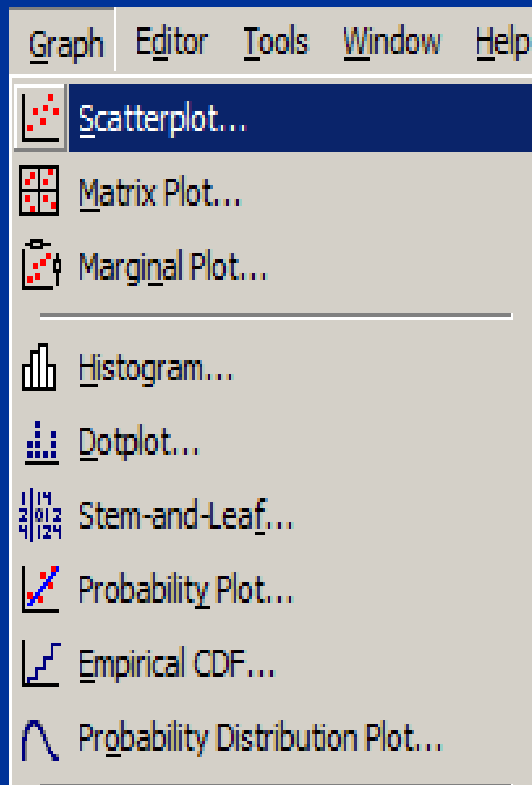


Regresyon Doğrusu



Örnek: Beyzbol Skor Tahmini

- Laboratuvar Sayfasından “Takım İstatistikleri” veri setini indiriniz.



Örnek: Beyzbol Skor Tahmini

Scatterplot - Simple

C2 Ortalama Vuruş Sayısı
C3 Takım Skoru

	Y variables	X variables
1	'Takım Skoru'	'Ortalama Vur
2		
3		
4		
5		
6		
7		

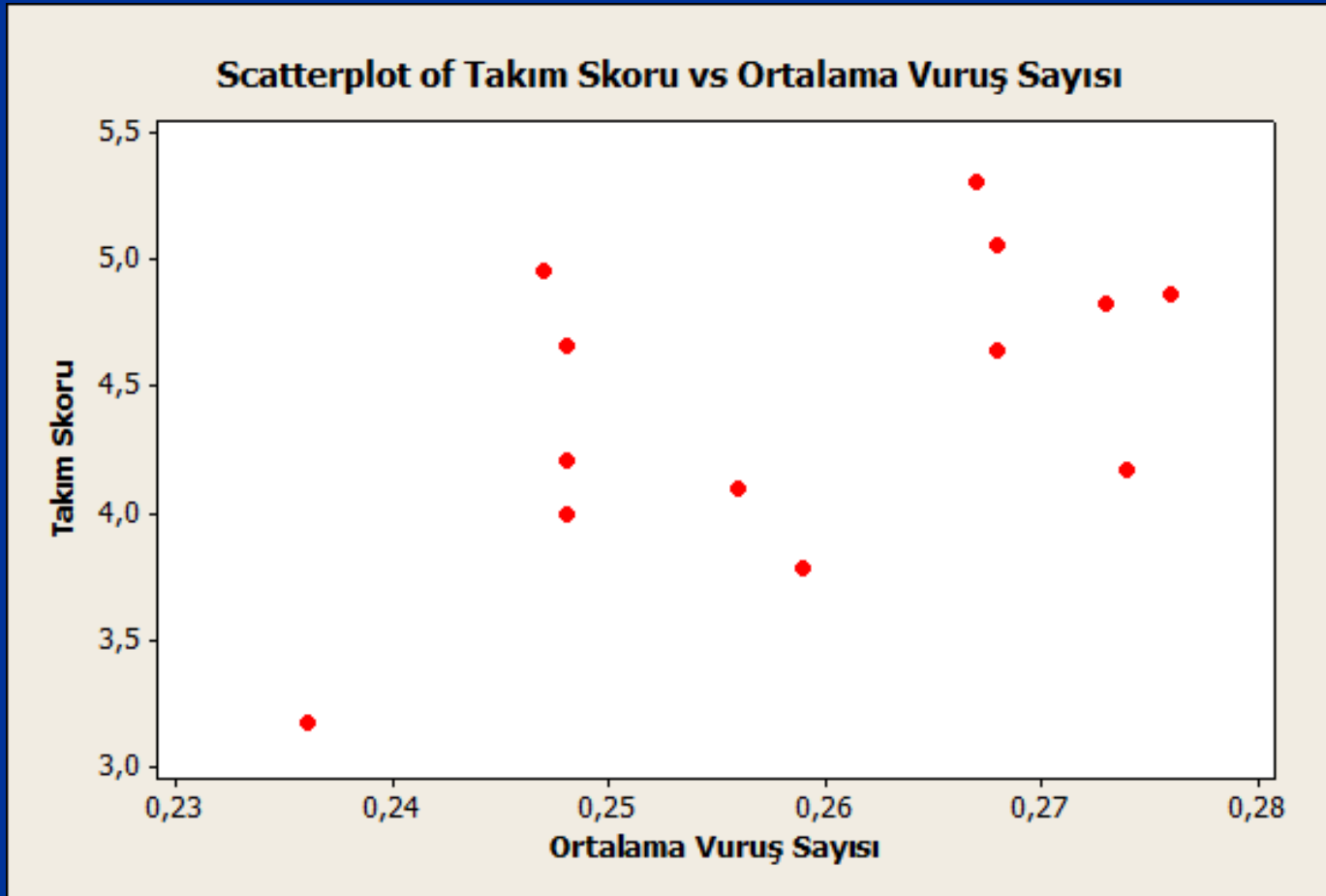
Scale... Labels... Data View...

Multiple Graphs... Data Options...

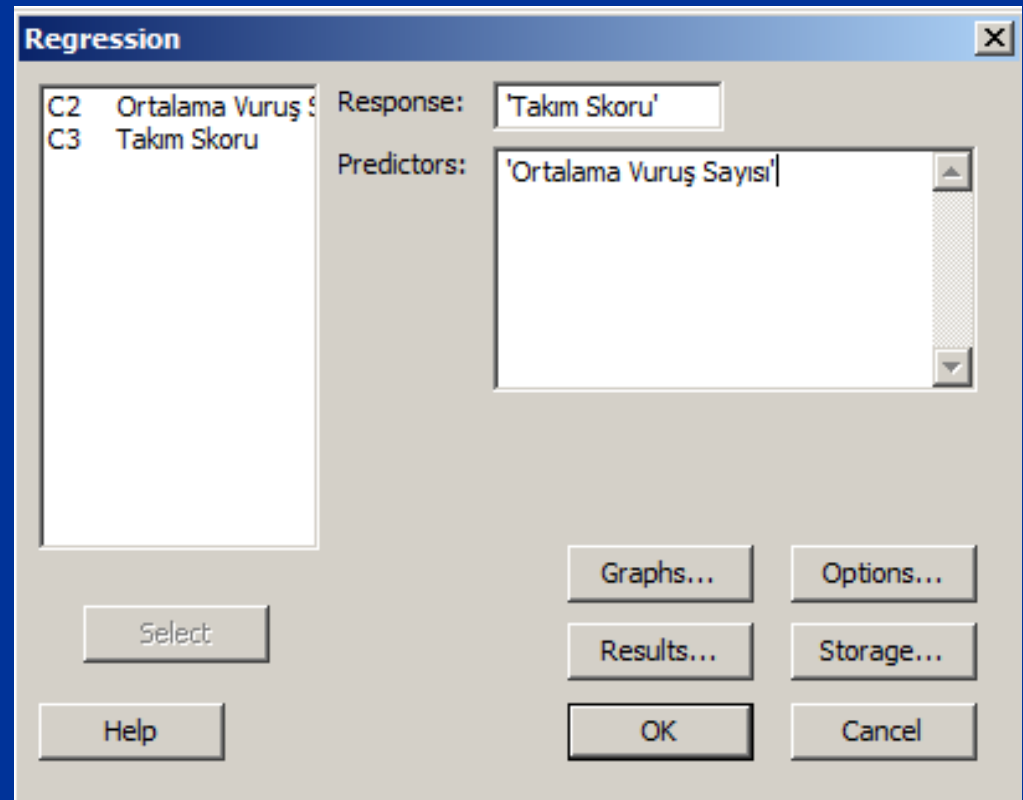
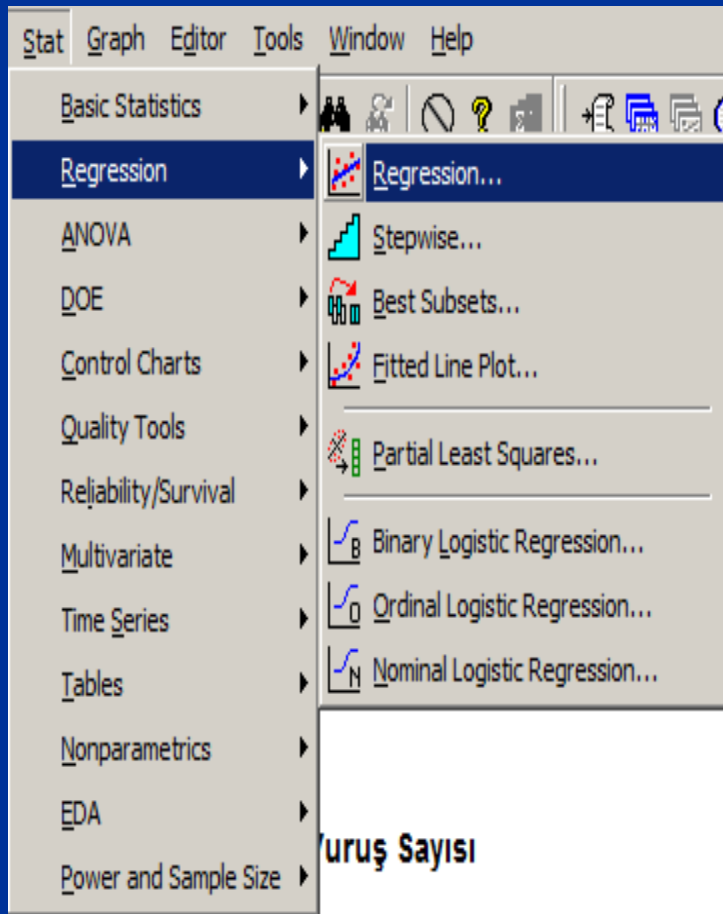
Select

Help OK Cancel

Örnek: Beyzbol Skor Tahmini



Beyzbol Skor Tahmini İçin Regresyon Doğrusunun Hesaplanması



Beyzbol Skor Tahmini İçin Regresyon Doğrusunun Hesaplanması

Regression Analysis: Takım Skoru versus Ortalama Vuruş Sayısı

The regression equation is

Takım Skoru = - 2,32 + 26,1 Ortalama Vuruş Sayısı

Predictor	Coef	SE Coef	T	P
Constant	-2,320	2,832	-0,82	0,429
Ortalama Vuruş Sayısı	26,07	10,89	2,39	0,034

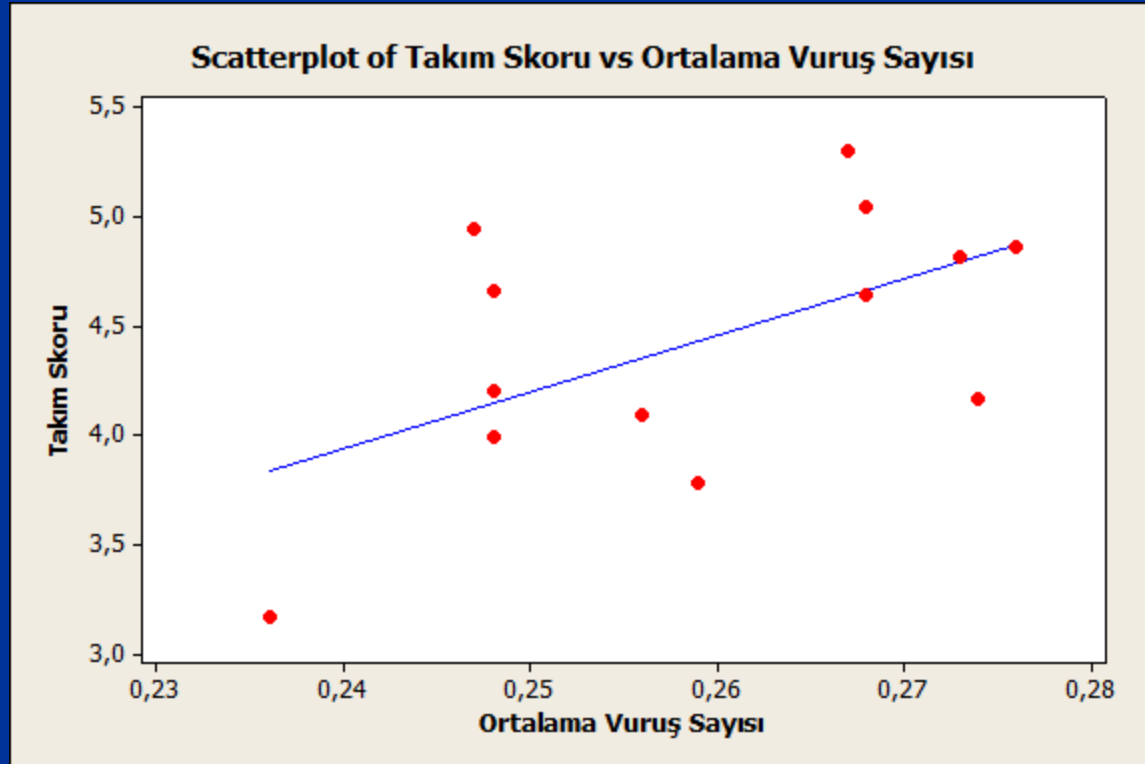
S = 0,493821 R-Sq = 32,3% R-Sq(adj) = 26,7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	1,3969	1,3969	5,73	0,034
Residual Error	12	2,9263	0,2439		
Total	13	4,3232			

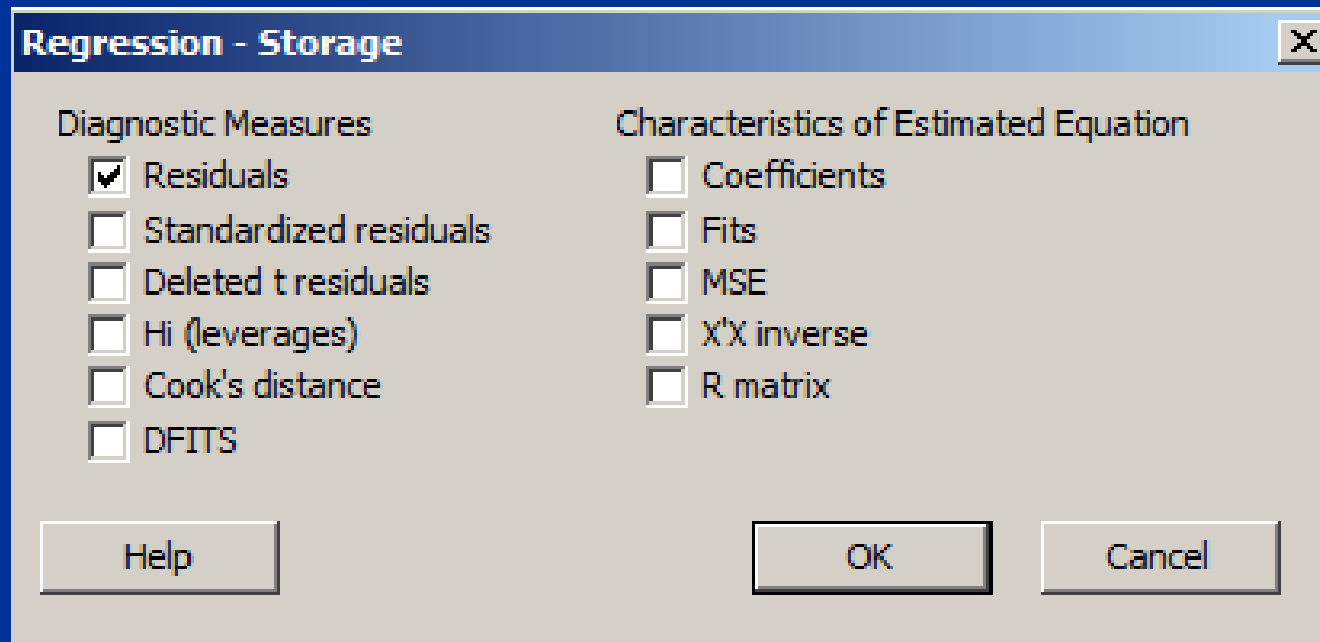
Beyzbol Skor Tahmini İçin Regresyon Doğrusunun Hesaplanması

- Daha önce oluşturduğumuz Scatterplot üzerinde, sağ tıklayarak **Add>Regression fit>Linear** adımlarını izleyiniz.



Beyzbol Skor Tahmini İçin Hesaplama Hatalarının Belirlenmesi

- Hesaplama hatalarının belirlenmesi için, **Stat>Regression>Regression>Storage** adımlarını takip ediniz.

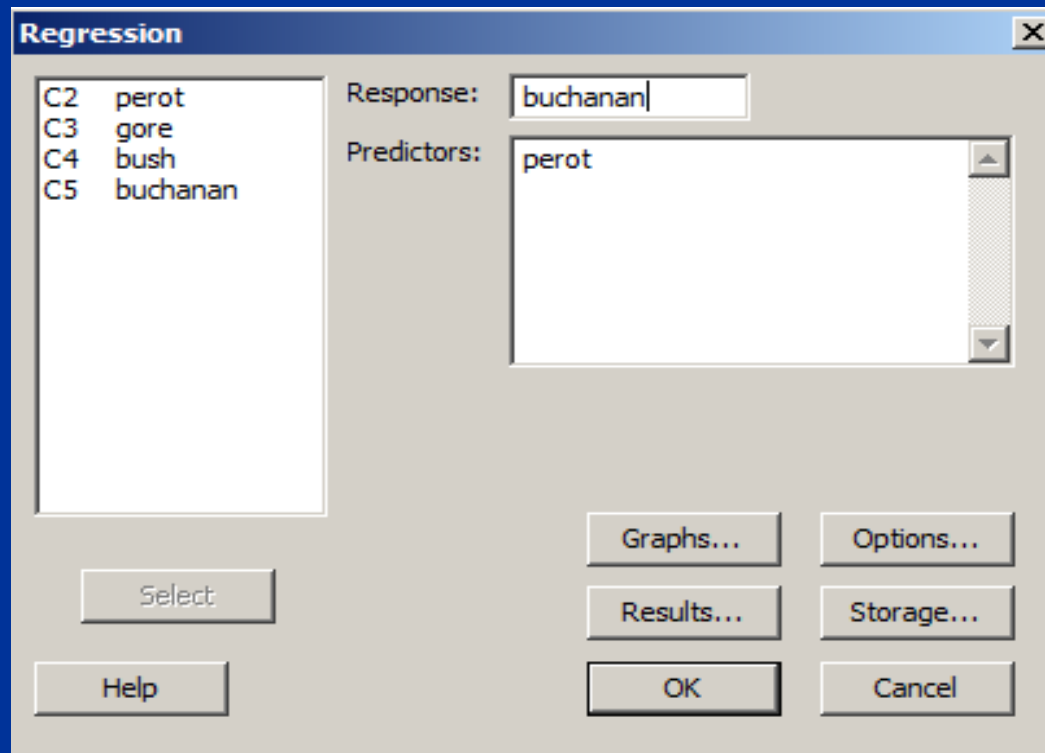


Beyzbol Skor Tahmini İçin Hesaplama Hatalarının Belirlenmesi

C1-T	C2	C3	C4
Takım	Ortalama Vuruş Sayısı	Takım Skoru	RESI1
NY Yankees	0,267	5,30	0,658613
Boston	0,268	5,05	0,382540
Tampa Bay	0,247	4,95	0,830067
Texas	0,276	4,86	-0,016041
Minnesota	0,273	4,82	0,022177
Toronto	0,248	4,66	0,513995
Chicago Sox	0,268	4,64	-0,027460
Detroit	0,268	4,64	-0,027460
LA Angeles	0,248	4,20	0,053995
Kansas City	0,274	4,17	-0,653896
Oakland	0,256	4,09	-0,264587
Cleveland	0,248	3,99	-0,156005
Baltimore	0,259	3,78	-0,652805
Seattle	0,236	3,17	-0,663133

Örnek: Olağandışı Oy Toplamının Belirlenmesi

- Laboratuvar Sayfasından “Buchanan and Butterfly” veri setini indiriniz.
- **Stat>Regression>Regression** adımlarını takip ediniz.



Örnek: Olağandışı Oy Toplamının Belirlenmesi

- **Regression>Storage** adımlarını takip ediniz.

Regression - Storage

Diagnostic Measures

- ☒ Residuals
- ☐ Standardized residuals
- ☐ Deleted t residuals
- ☐ Hi (leverages)
- ☐ Cook's distance
- ☐ DFITS

Characteristics of Estimated Equation

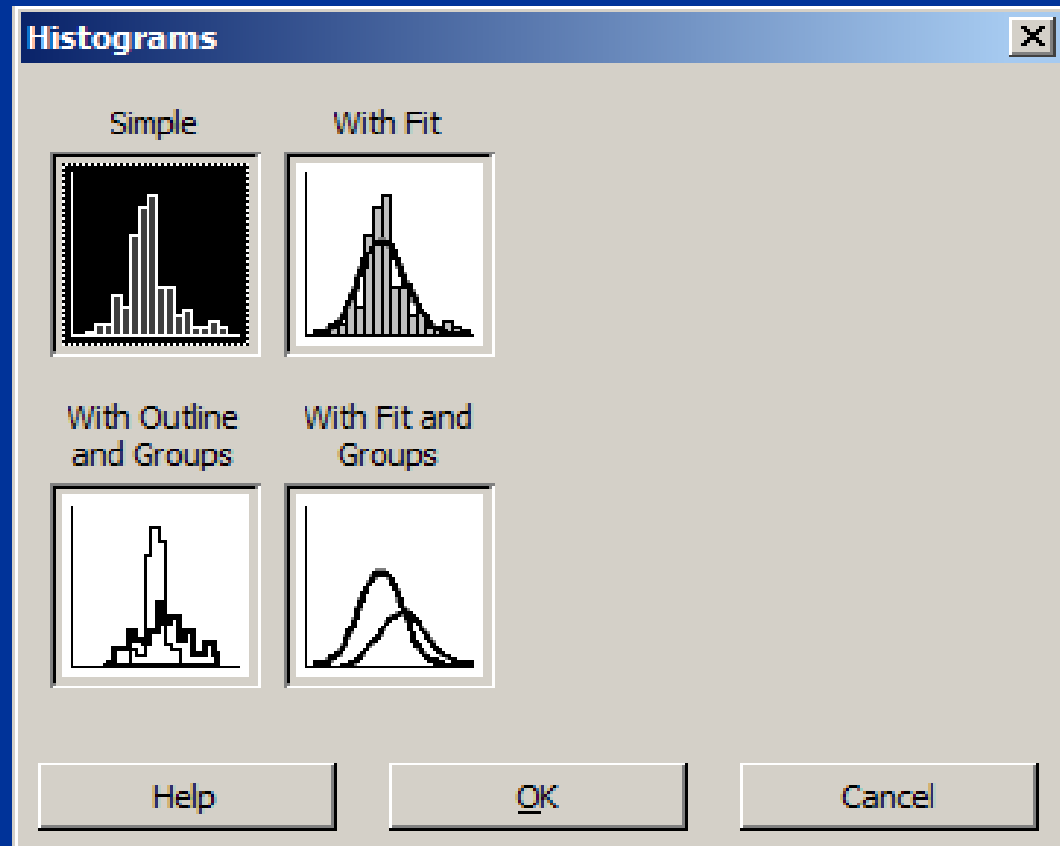
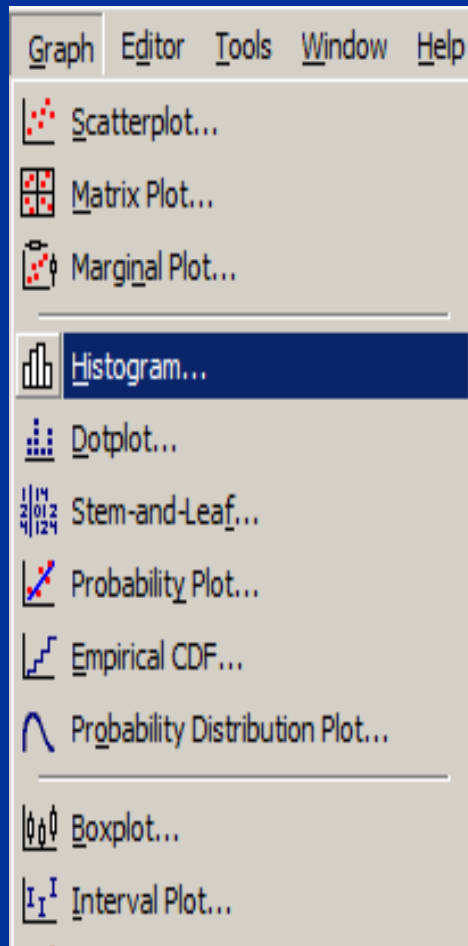
- ☐ Coefficients
- ☐ Fits
- ☐ MSE
- ☐ X'X inverse
- ☐ R matrix

Help OK Cancel

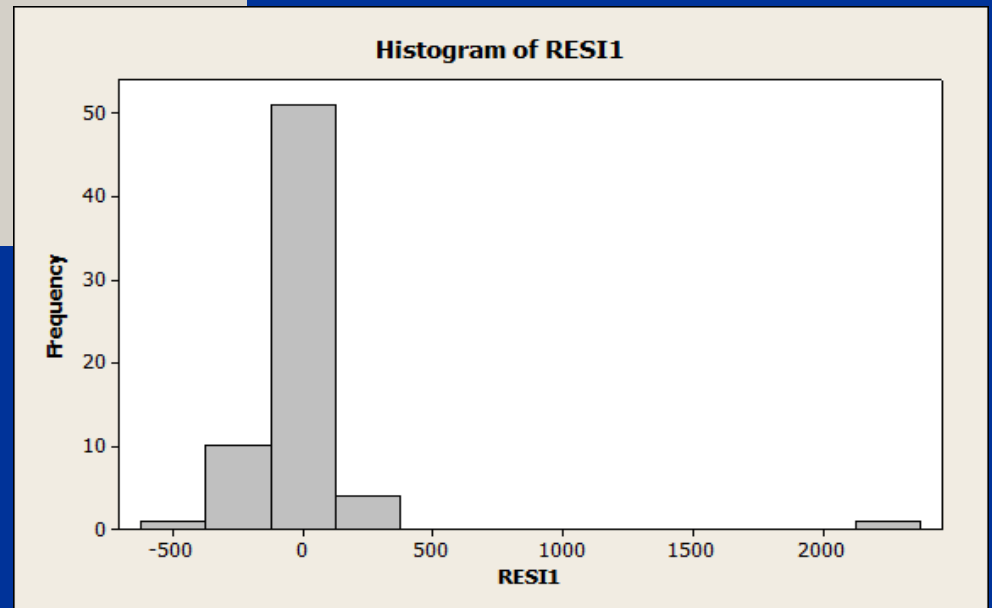
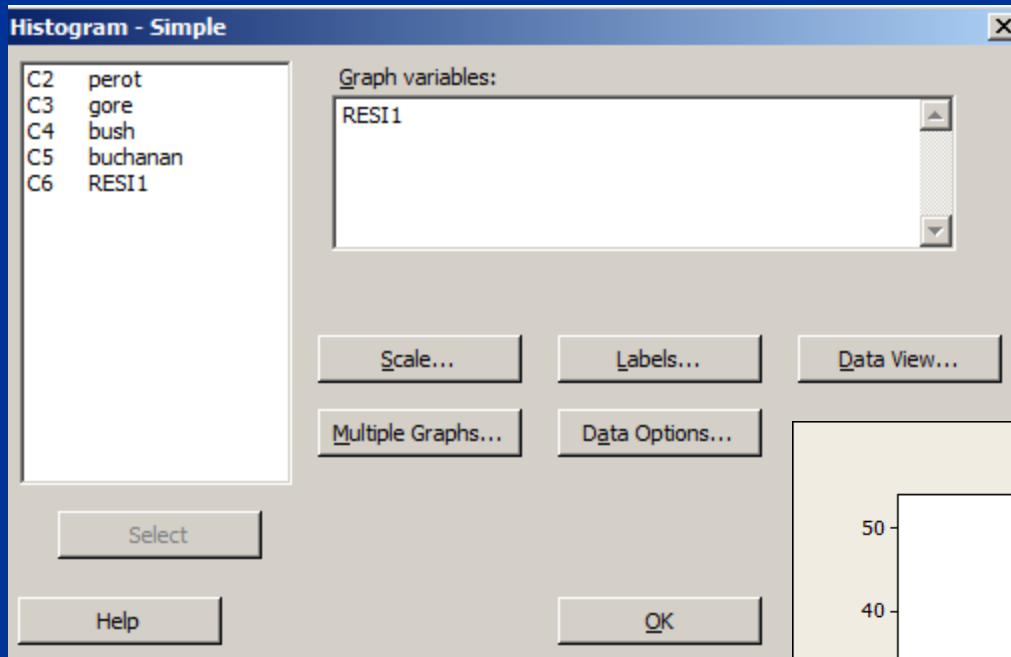
BUCHANAN_AND_THE_BUTTERFLY_BALLOT.MTW ***

↓	C1-T	C2	C3	C4	C5	C6
	county	perot	gore	bush	buchanan	RESI1
1	Alachua	8072	47300	34062	262	-27.53
2	Baker	667	2392	5610	73	48.08
3	Bay	5922	18850	38637	248	35.30
4	Bradford	819	3075	5414	65	34.65
5	Brevard	25249	97318	115185	570	-333.34
6	Broward	38964	386518	177279	789	-604.44

Olağandışı Oy Toplamı İçin Histogram Çizilmesi

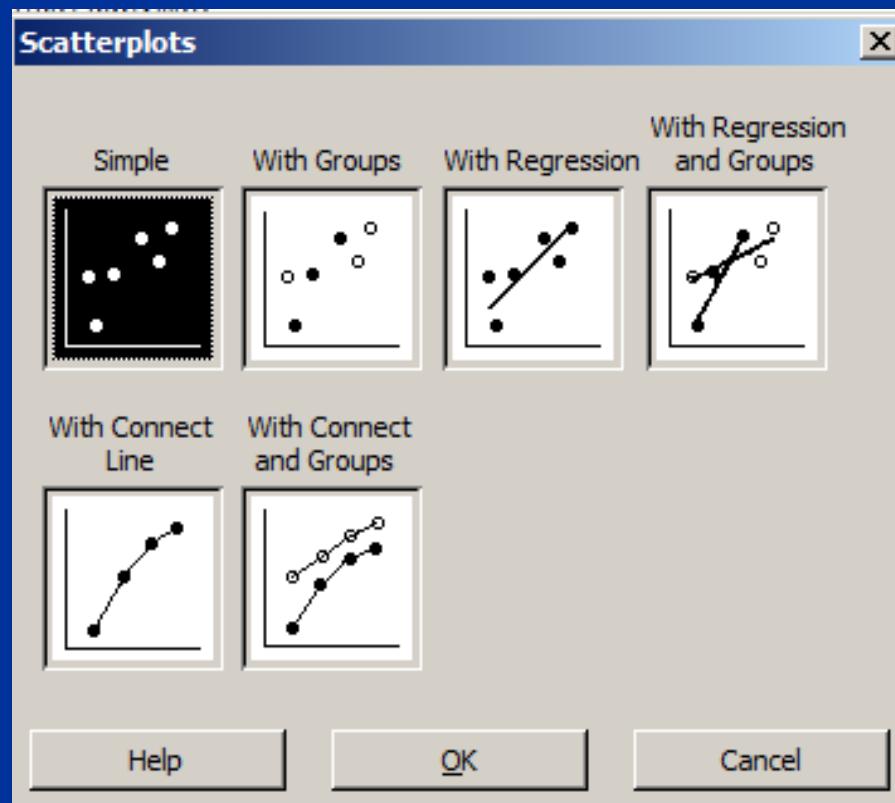
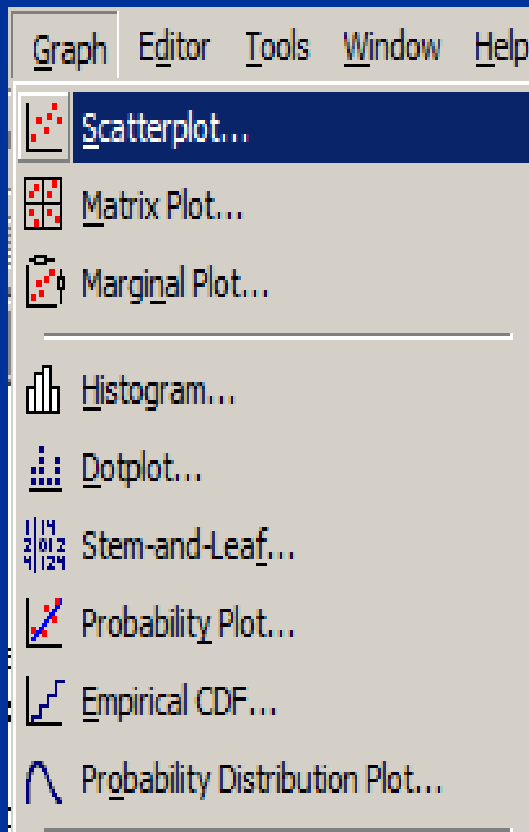


Olağandışı Oy Toplamı İçin Histogram Çizilmesi



Örnek: Olimpik Yüksek Atlama

- Laboratuvar Sayfasından “High Jump” veri setini indiriniz.



Örnek: Olimpik Yüksek Atlama

Scatterplot - Simple [X]

C1 Men_Meters
C2 Women_Meters
C3 Year_Men
C4 Year_Women
C6 Winning Height (m)
C8 Year

	Y variables	X variables	
1	'Men_Meters'	'Year_Men'	
2	'Women_Met	'Year_Women'	
3			
4			
5			
6			
7			

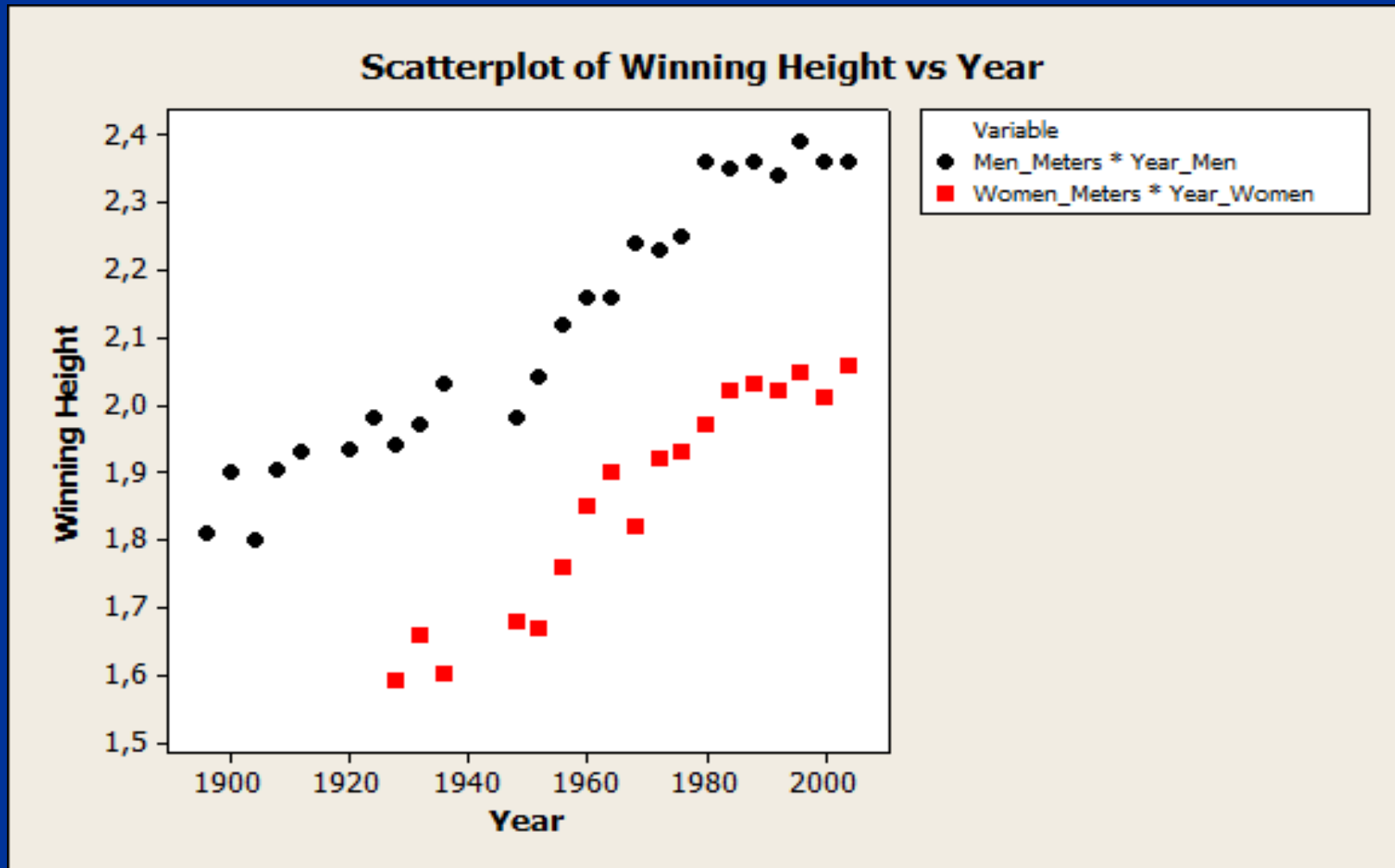
Scale... Labels... Data View...

Multiple Graphs... Data Options...

Select

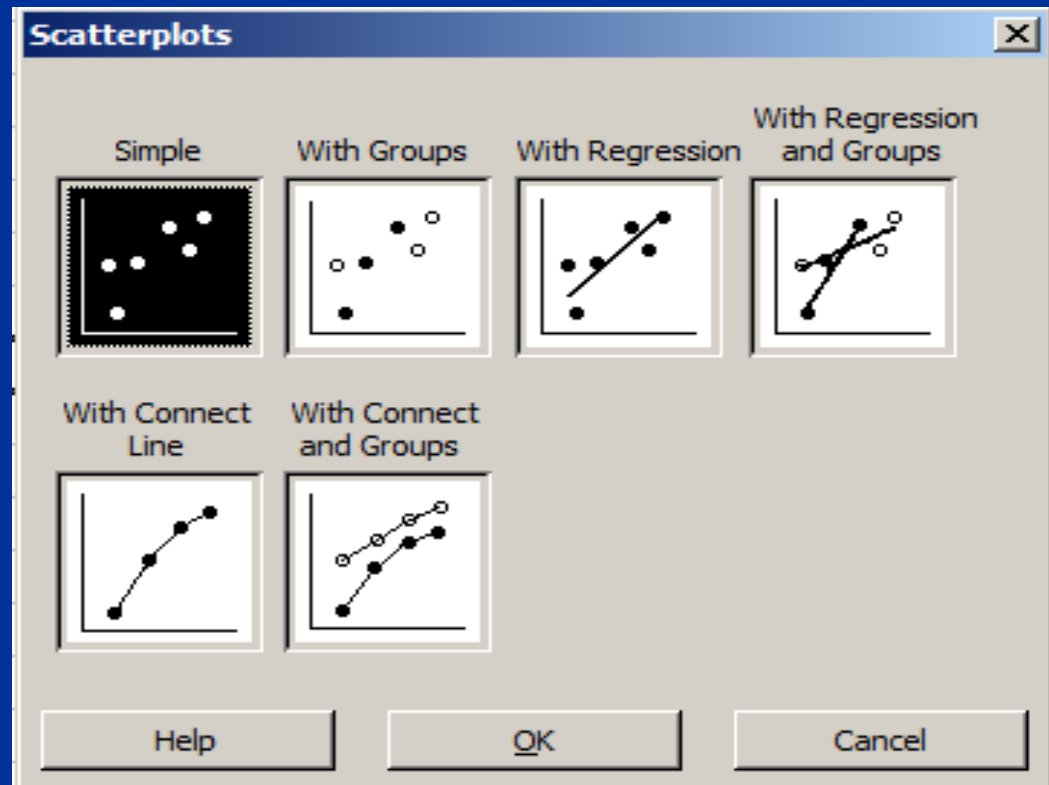
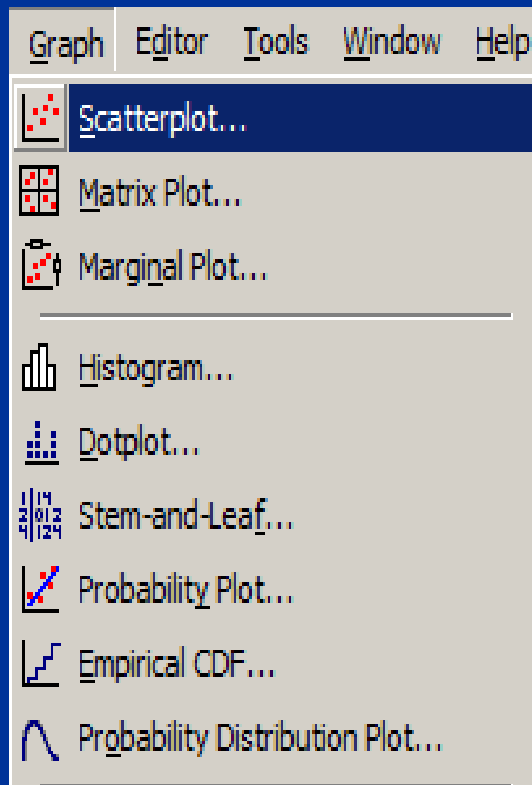
Help OK Cancel

Örnek: Olimpik Yüksek Atlama



Örnek: Eğitim ve Cinayet Oranları İlişkisi

- Laboratuvar Sayfasından “US Statewide Crime” veri setini indiriniz.



Örnek: Eğitim ve Cinayet Oranları İlişkisi

Scatterplot - Simple

C2	violent crime rate
C3	murder rate
C4	poverty
C5	high school
C6	college
C7	single parent
C8	unemployed
C9	metropolitan

	Y variables	X variables
1	'murder rate'	college
2		
3		
4		
5		
6		
7		

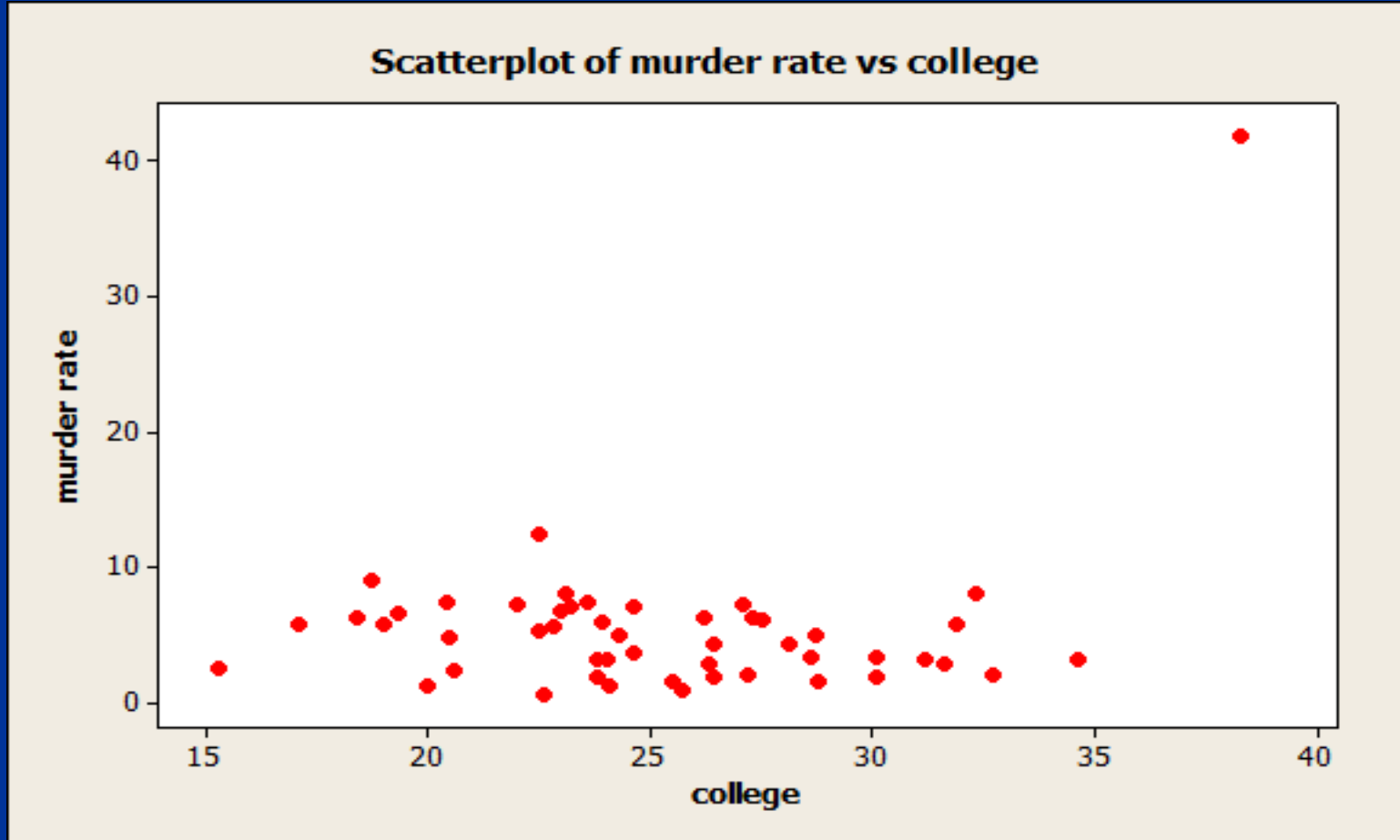
Select

Scale... Labels... Data View...

Multiple Graphs... Data Options...

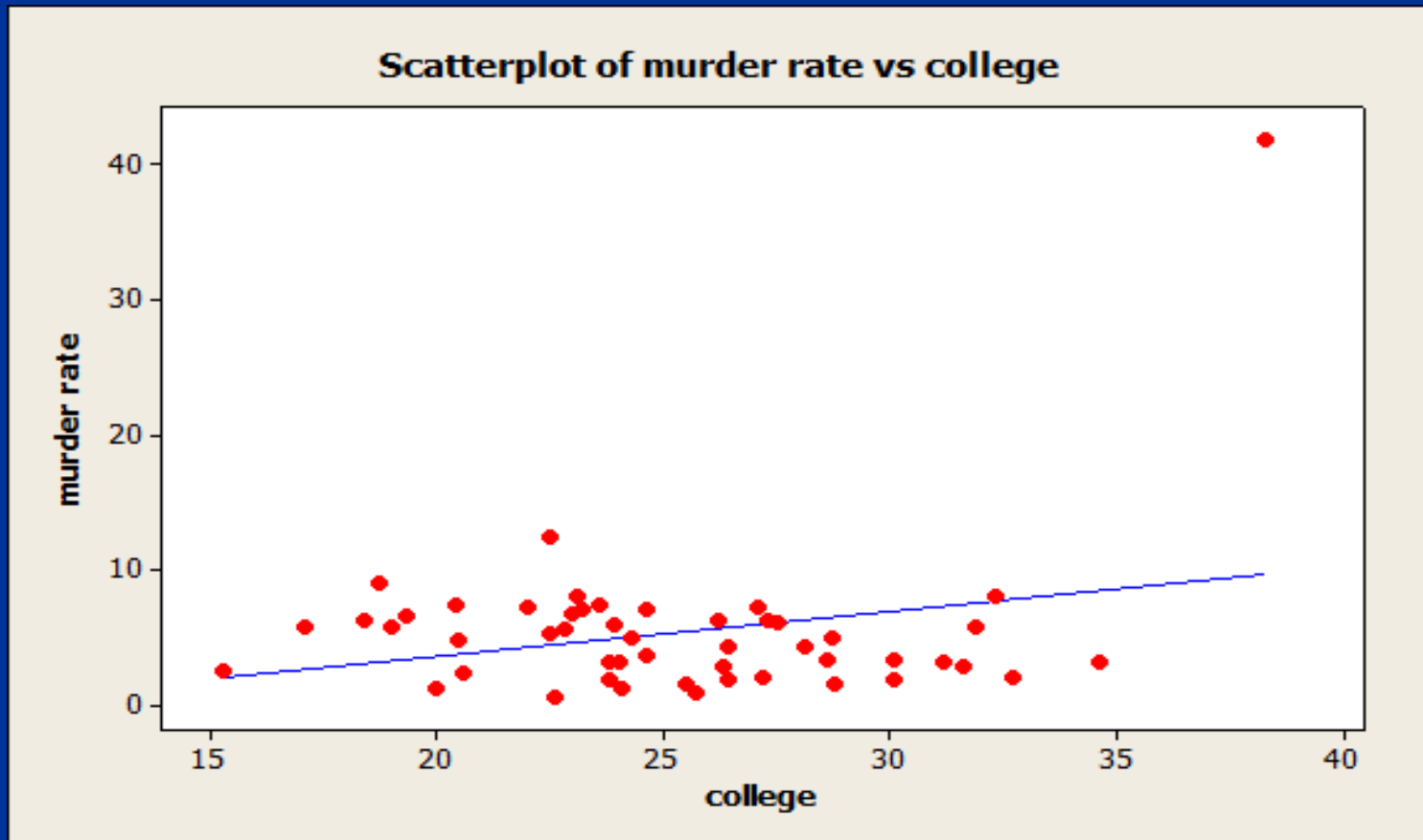
Help OK Cancel

Örnek: Eğitim ve Cinayet Oranları İlişkisi



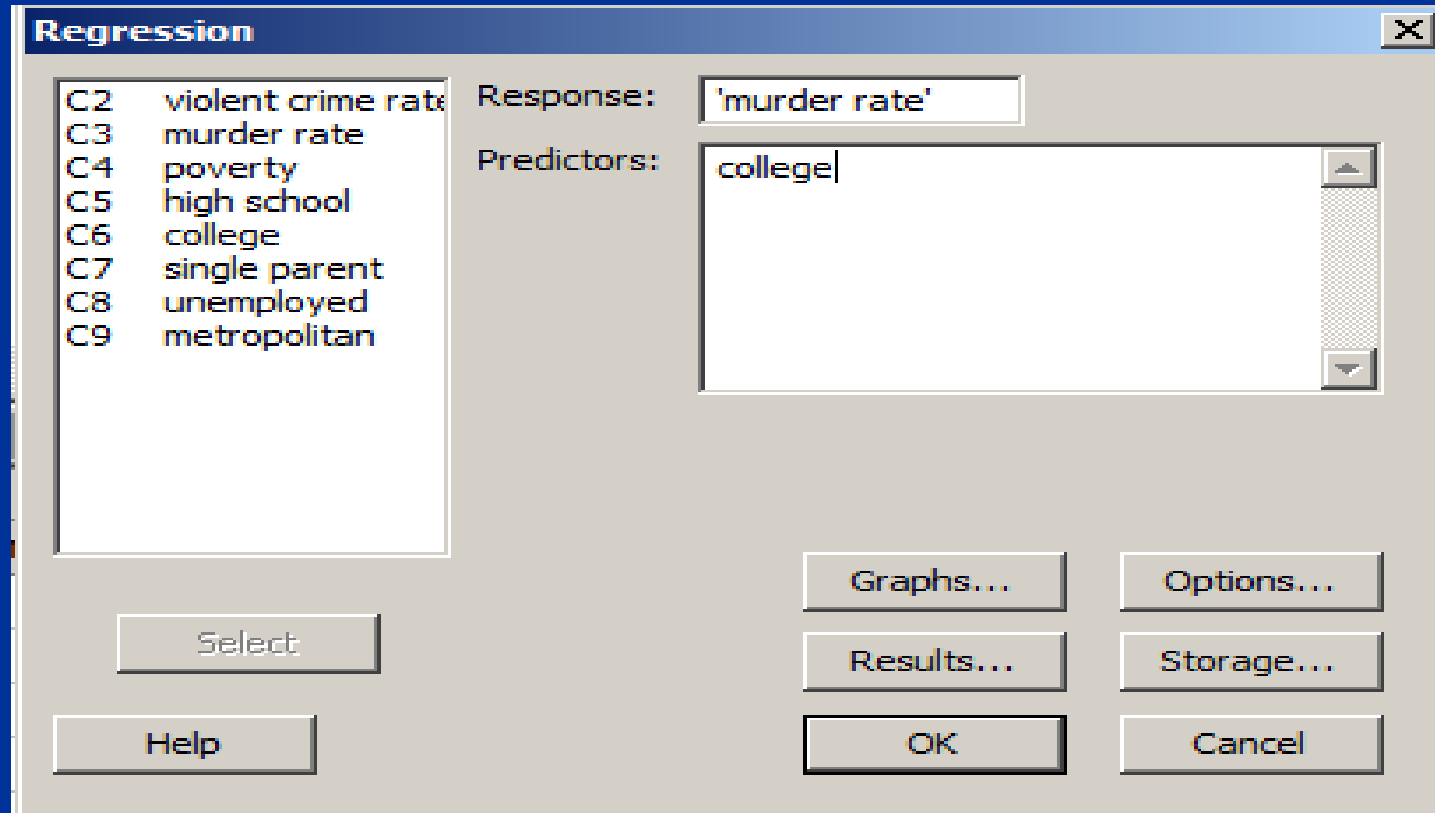
Örnek: Eğitim ve Cinayet Oranları İlişkisi

- Ardından, **Add>Regression Fit>Linear>OK** adımlarını takip ediniz.



Örnek: Eğitim ve Cinayet Oranları İlişkisi

- Veri setinin regresyon doğrusunu belirlemek için **Stat>Regression>Regression** adımlarını takip ediniz.



Örnek: Eğitim ve Cinayet Oranları İlişkisi

Regression Analysis: murder rate versus college

The regression equation is
murder rate = - 3,06 + 0,333 college

Predictor	Coef	SE Coef	T	P
Constant	-3,058	4,363	-0,70	0,487
college	0,3331	0,1703	1,96	0,056

|
S = 5,61458 R-Sq = 7,2% R-Sq(adj) = 5,3%