

Varsayımlar (Normallik)

- **Çok değişkenli normallik** her bir değişkenin ve değişkenlerin bütün doğrusal kombinasyonlarının normal dağılıma sahip olmasıdır.
- Çok değişkenli normallik varsayımı, kısmen tek tek değişkenlerin normalliği, doğrusallığı ve homoscedasticity incelenerek veya yordama içeren analizlerde artıkların incelenmesi yoluyla kontrol edilebilir.
- **Gruplanmamış** verilerde çok değişkenli normallik varsa, her değişkenin kendisi normal dağılıma sahiptir ve varsa değişken çiftleri arasındaki ilişkiler doğrusal ve homoscedastiktir (yani, bir değişkenin varyansı diğer değişkenin tüm değerlerinde aynıdır).
- **Gruplanmış** veriler için, değişkenlerin ortalamalarının örneklem dağılımlarının normal olması gerekir.



Varsayımlar (Normallik)

- Hemen hemen her çok değişkenli analiz için sürekli değişkenlerin normallik varsayımını sağlayıp sağlamadığını incelemek ilk önemli adımdır.
- Değişkenlerin normalliği her zaman analiz için gerekli olmasa da, değişkenlerin hepsi normal olarak dağılırsa çözüm genellikle biraz daha iyi olarak kabul edilir.
- Değişkenler normal olarak dağılmazsa (örneğin, pozitif ve negatif çarpık) ise, analiz sonuçları hatalı olabilmektedir.
- Değişkenlerin normalliği istatistiksel veya grafiksel yöntemlerle değerlendirilir.
 - İstatistiksel yöntemler; normallik için hipotez testlerini içerir.
 - Grafiksel yöntemler; histogram ve normal olasılık grafiklerini kapsar.



Varsayımlar (Normallik)

- Normalliğin iki bileşeni vardır: çarpıklık ve basıklıktır.
- **Dağılım normal olduğunda**, değişkenin çarpıklık ve basıklık **değerleri sıfırdır**.
- İstatistiksel olarak çarpıklık ve basıklık değerleri incelenir. Bu değerler -1 ile 1 aralığında ise dağılım normal kabul edilir.
- Elde edilen çarpıklık ve basıklık değerlerinin null hipotez olan sıfır değerine karşı test eden hem çarpıklık hem de basıklık için anlamlılık testleri vardır. Bu hipotez testlerinde z dağılımı kullanılmaktadır.

$$z = \frac{S-0}{s_s} \quad s_s = \sqrt{\frac{6}{N}}$$

$$z = \frac{K-0}{s_k} \quad s_k = \sqrt{\frac{24}{N}}$$

- Burada S çarpıklık katsayısı, s_s çarpıklık için standart hata ve N gözlem sayısıdır. K basıklık katsayısı, s_k basıklık için standart hata ve N gözlem sayısıdır.



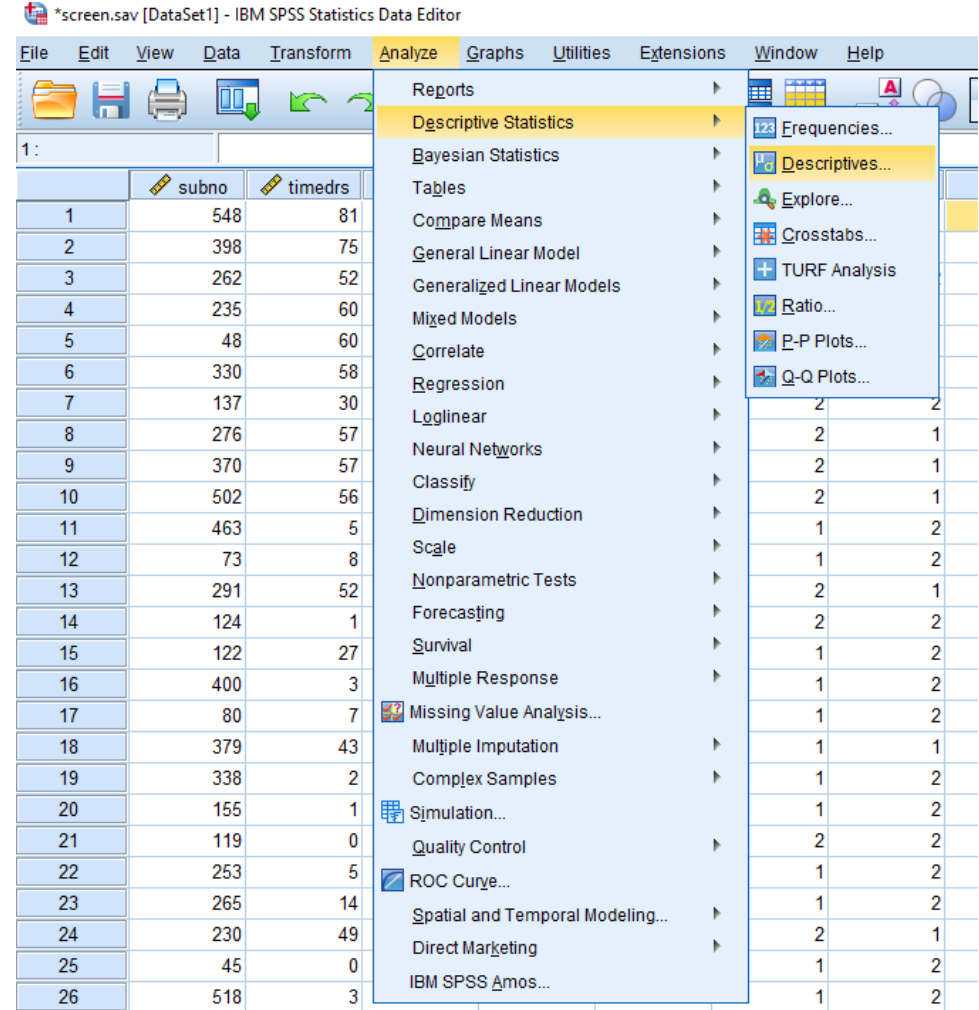
Varsayımlar (Normallik)

- Küçük ve orta büyüklükteki örneklem için çarpıklık ve basıklık değerlerini test etmede 0.01 veya 0.001 daha tutucu alfa düzeyleri kullanılır.
- Eğer örneklem büyükse dağılımın şekline bakmak daha iyi olacaktır çünkü standart hata, büyük veride küçülmekte bu da çarpıklık ve basıklık için belirlenen yokluk hipotezinin reddedilmesine yol açacaktır.
- Büyük bir örneklemde, istatistiksel olarak anlamlı çarpıklığa sahip bir değişken, normallikten yeterince sapmaz ve bu analizlerde önemli bir fark oluşturmaz. Başka bir deyişle, **büyük örneklemelerde, çarpıklığın anlamlılık düzeyi**, gerçek büyüklüğü ve dağılımın görsel görünümünü kadar önemli değildir.
- Büyük bir örneklemde, basıklık değerinin sıfırdan uzaklaşmasının etkisi de azalmaktadır.
 - Örneğin, pozitif basıklık ile ilişkili varyansın olduğundan daha az tahmini, 100 veya daha büyük örneklemelerde; negatif basıklık ile varyansın olduğundan daha az tahmini ise 200 veya daha büyük örneklemelerde tolere edilebilir (Waternaux, 1976).



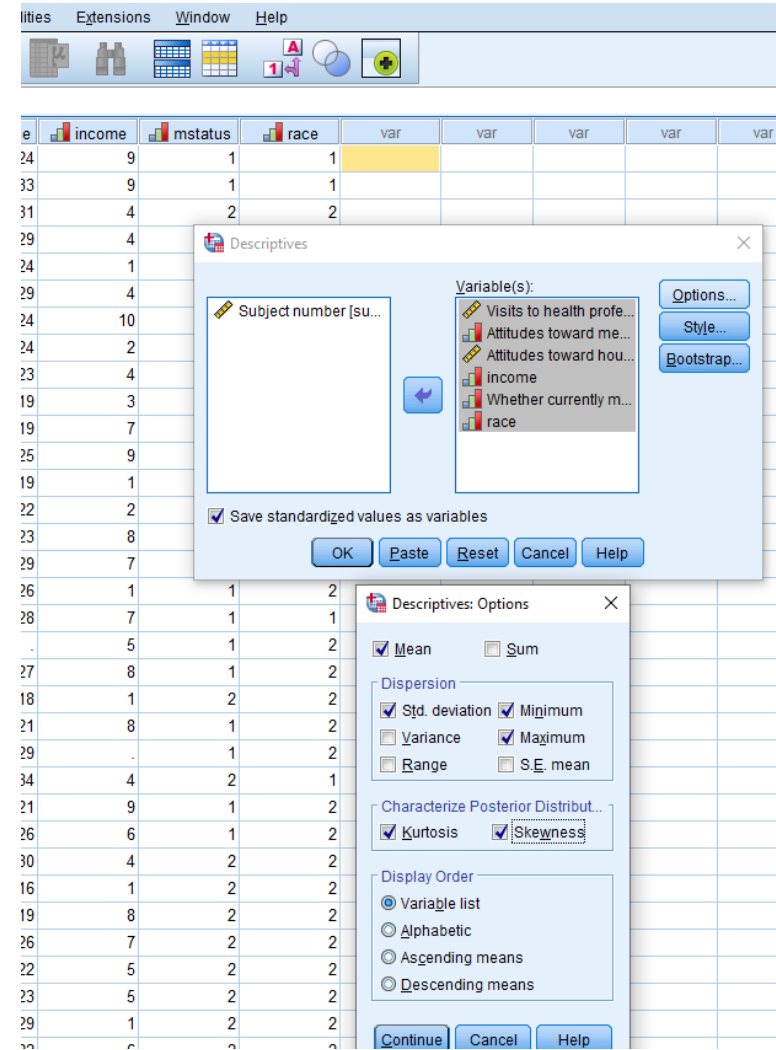
Varsayımlar (Normallik) SPSS Uygulama

- Çarpıklık ve basıklık değerlerini test etmek üzere şu adımlar izlenebilir;
- Ana menüden önce “Analyze”, sonra “Descriptive Statistics”, daha sonra da “Descriptives”e tıklayınız.



Varsayımlar (Normallik) SPSS Uygulama

- İncelemek istediğiniz değişkenleri seçiniz
- Seçilen değişkeni “Variable(s)” kutusuna aktarmak için ok işaretine tıklayınız.
- “Options” a tıklayınız.
- “Distribution” bölümünde “**Kurtosis**” ve “**Skewness**”ı seçiniz.
- Önce “Continue”, sonra “OK”e tıklayınız.



Varsayımlar (Normallik) SPSS Uygulama

Descriptives

Descriptive Statistics									
	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Visits to health professionals	465	0	81	7,90	10,948	3,248	,113	13,101	,226
Attitudes toward medication	465	5	10	7,69	1,156	-,123	,113	-,447	,226
Attitudes toward housework	464	2	35	23,54	4,484	-,457	,113	1,556	,226
income	439	1	10	4,21	2,419	,582	,117	-,359	,233
Valid N (listwise)	438								

Diğer sürekli değişkenlerin çarpıklık ve basıklık değerleri beklenen aralıkta yer almaktadır.

- Sürekli değişkenlerden «Visits to health professionals» değişkeni için çarpıklık ve basıklık değerleri beklenen aralığın dışında yer almaktadır. Çarpıklık ve basıklık değerlerinin anlamlılığını istatistiksel olarak test etmek için çarpıklık ve basıklık değerleri standart hatalarına bölünerek z değeri elde edilir. Elde edilen değerler 0.001 alfa anlamlılık düzeyinde kritik z değeri ($z = 3,29$) ile karşılaştırılır.

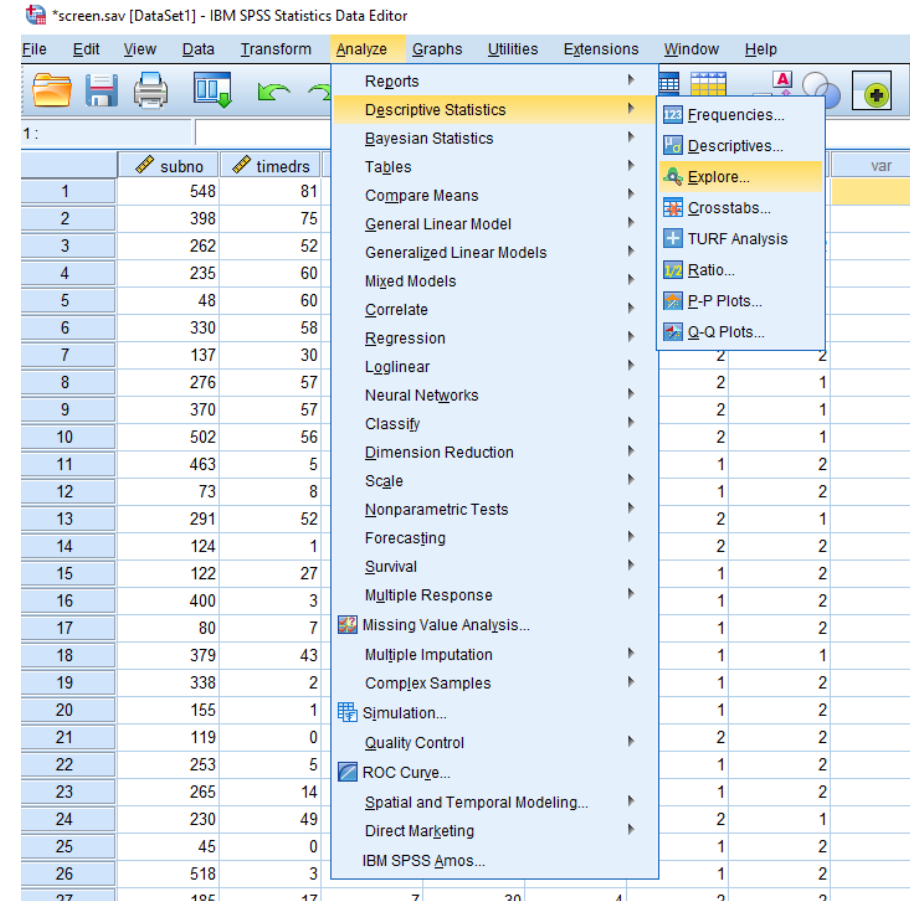
$$z = \frac{3,241}{0,113} = 28,68$$

$$z = \frac{13,041}{0,226} = 57,70$$



Varsayımlar (Normallik)

- Buna ek olarak tek değişkenli normallik varsayımı Kolmogorov-Smirnov testi ile de test edilebilir. Bu testin anlamlı olmayan sonuçları dağılımın normal olduğunun göstergesidir.
- Ancak hipotez testlerinde olduğu gibi bu testin **büyük örneklemelerde anlamlı çıkması kuvvetle muhtemeldir.**
- Analyze' menüsünden 'Descriptive Statistics' Ardından 'Explore' seçilir



Varsayımlar (Normallik)

- İncelemek istediğiniz değişkenleri seçiniz
- Seçilen değişkeni “Dependent List” kutusuna aktarmak için ok işaretine tıklayınız.
- “Plots” a tıklayınız.
- “Explore: Plots” bölümünde “Histogram” ve “**Normality plots with tests**”i seçiniz.
- Önce “Continue”, sonra “OK”e tıklayınız.

SPSS Statistics Data Editor

Transform Analyze Graphs Utilities Extensions Window Help

medrs	attdrug	atthouse	income	mstatus	race	var	var	var	var	var	var
81	8	24	9	1	1						
75	9	33	9	1	1						
52	9	31	4	2	2						
60	10	29									
60	7	24									
58	7	29									
30	5	24									
57	9	24									
57	8	23									
56	8	19									
5	10	19									
8	10	25									
52	8	19									
1	5	22									
27	8	23									
3	6	29									
7	6	26									
43	6	28									
2	6		5	1	2						
1	9	27	8	1	2						
0	10	18	1	2	2						
5	9	21	8	1	2						
14	7	29		1	2						
49	8	34	4	2	1						
0	8	21	9	1	2						
3	8	26	6	1	2						
17	7	30	4	2	2						
0	9	16	1	2	2						
1	9	19	8	2	2						

Explore

Subject number [su...]
Attitudes toward me...
income
Whether currently m...
race

Dependent List:
Visits to health profe...
Attitudes toward hou...

Factor List:

Label Cases by:

Display
☒ Both ☐ Statistics ☐ Plots

OK Paste Reset Cancel Help

Explore: Plots

Boxplots
☒ Factor levels together
☐ Dependents together
☐ None

Descriptive
☒ Stem-and-leaf
☒ Histogram

☒ Normality plots with tests

Spread vs Level with Levene Test
☐ None
☐ Power estimation
☐ Transformed Power: Natural log
☐ Untransformed

Continue Cancel Help



Varsayımlar (Normallik)

Descriptives

		Statistic	Std. Error
Visits to health professionals	Mean	7,91	,509
	95% Confidence Interval for Mean	Lower Bound	6,91
		Upper Bound	8,91
	5% Trimmed Mean	6,21	
	Median	4,00	
	Variance	120,053	
	Std. Deviation	10,957	
	Minimum	0	
	Maximum	81	
	Range	81	
	Interquartile Range	8	
	Skewness	3,245	,113
	Kurtosis	13,071	,226
Attitudes toward housework	Mean	23,54	,208
	95% Confidence Interval for Mean	Lower Bound	23,13
		Upper Bound	23,95
	5% Trimmed Mean	23,61	
	Median	24,00	
	Variance	20,102	
	Std. Deviation	4,484	
	Minimum	2	
	Maximum	35	
	Range	33	
	Interquartile Range	6	
	Skewness	-,457	,113
	Kurtosis	1,556	,226



Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Visits to health professionals	,235	464	,000	,638	464	,000
Attitudes toward housework	,058	464	,001	,978	464	,000

a. Lilliefors Significance Correction



Öncelikle çarpıklık ve basıklık değerleri incelenir Bu değerler -1 ile 1 aralığında olduğu zaman dağılım normal kabul edilir.

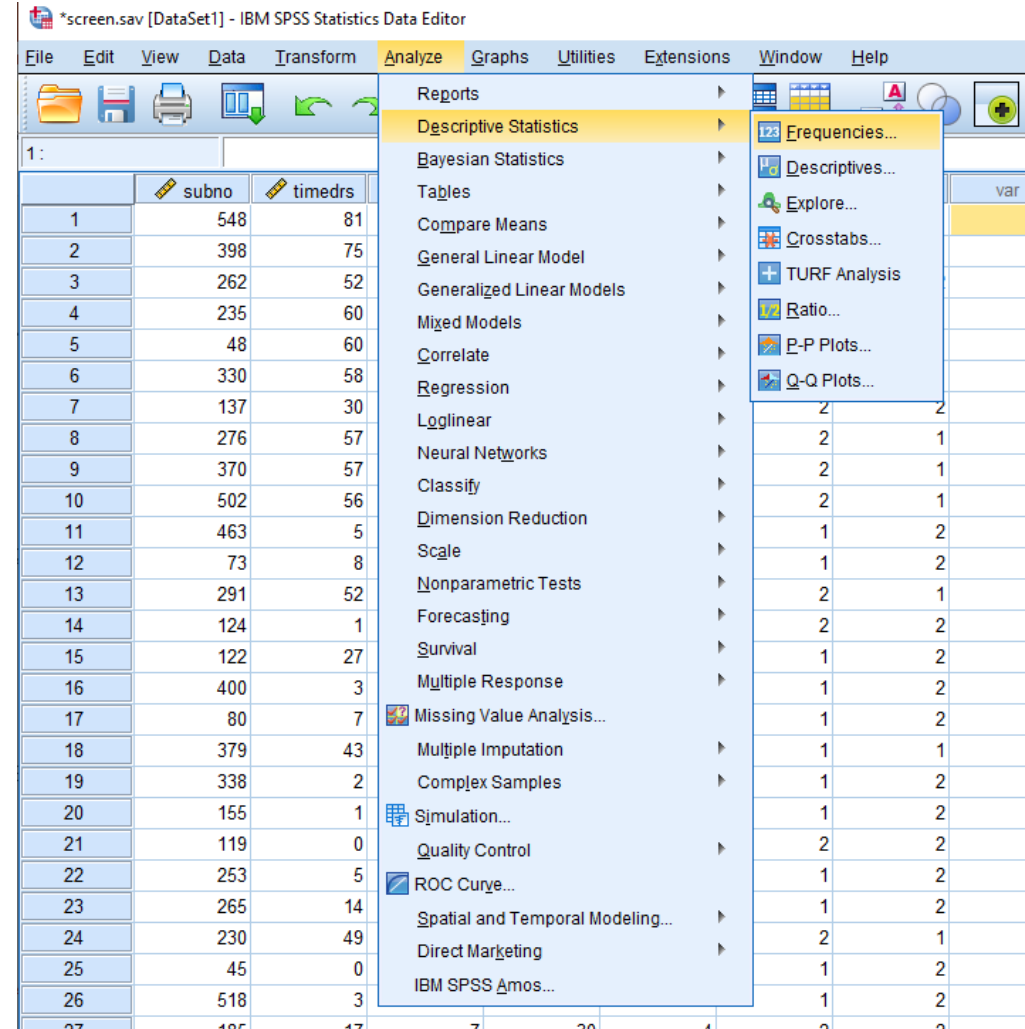
- Grup büyüklüğünün 50'den küçük olması durumunda Shapiro- Wilk; büyük olması durumunda Kolmogorov-Smirnov testi anlamlılık değerleri incelenir.
- K-S testine ilişkin p değerinin $p < 0.05$ küçük olması dağılımın normallikten uzaklaştığının göstergesi olarak yorumlanabilir.



Varsayımlar (Normallik) Grafiksel Yöntemler

Histogram

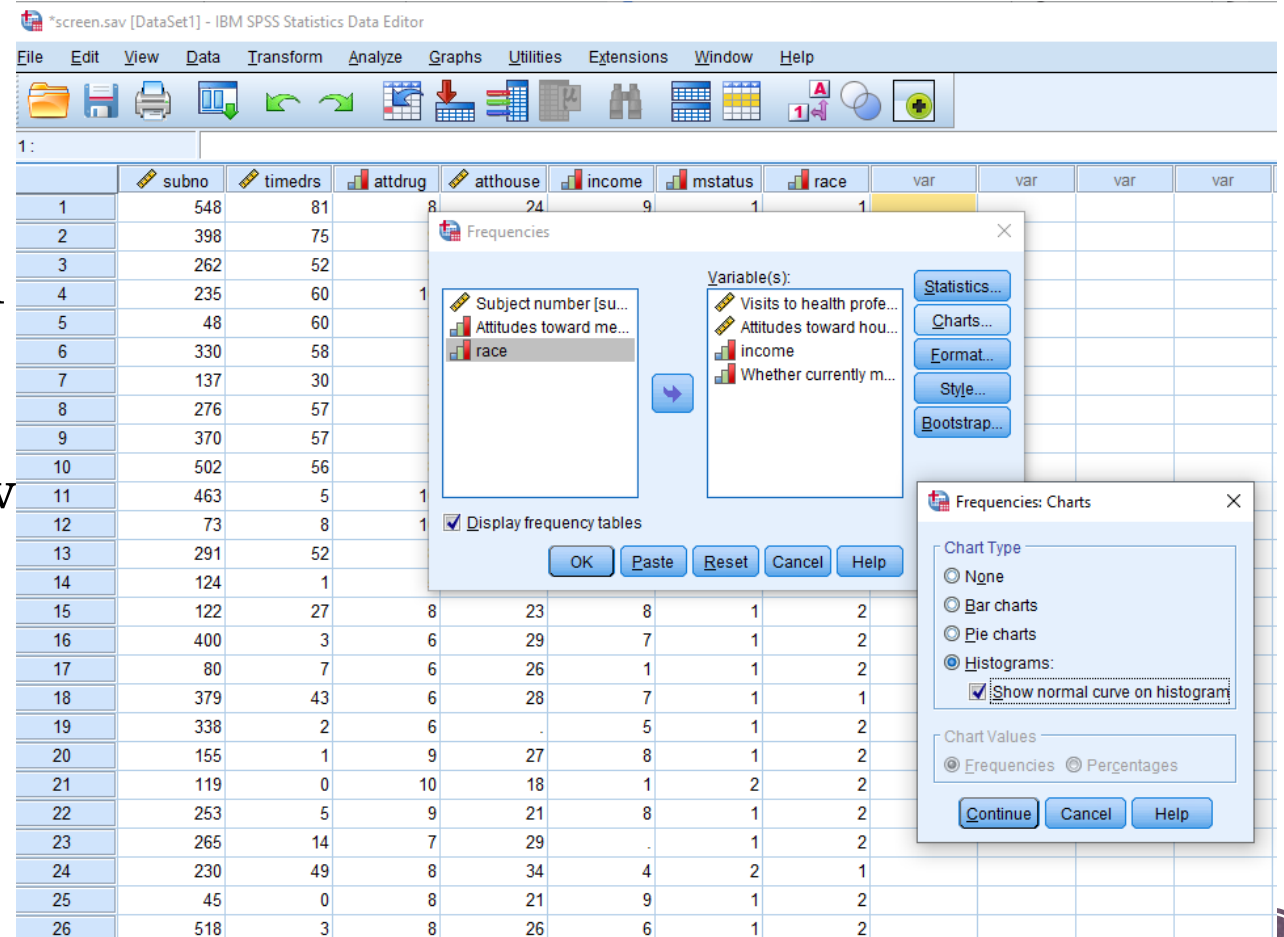
- “**Analyze**” menüsünde önce “**Descriptive Statistics**”, daha sonra da “**Frequencies**”e tıklayınız.



Varsayımlar (Normallik) Grafiksel Yöntemler

Histogram

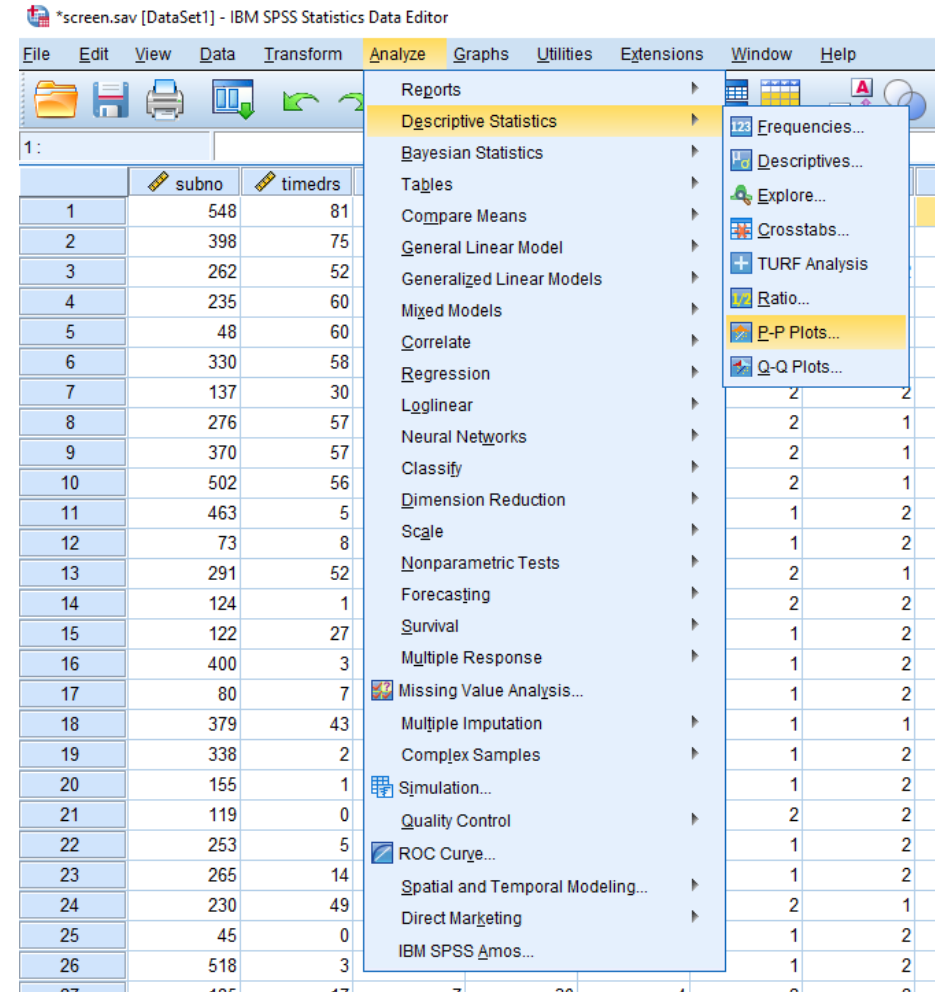
- İncelemek istediğiniz değişkenleri seçiniz. Seçilen değişkeni “Variables” kutusuna taşımak için ok düğmesine tıklayınız.
- “Charts”a tıklayınız. “Chart Type” bölümünde “Histograms” ve “Show normal curve on histogram”ı seçiniz.
- Continue’a ve daha sonra “OK”a tıklayınız.



Varsayımlar (Normallik) Grafiksel Yöntemler

Normal P-P plot grafiđi

- “Analyze” menüsünde önce “Descriptive Statistics”, daha sonra da “P-P Plots”a tıklayınız.
- İncelemek istediđiniz deđişkenleri seçiniz.
- Seçilen deđişkeni “Variables” kutusuna aktarmak için ok işaretine tıklayınız.
- “OK”a tıklayınız.



Varsayımlar (Normallik) Grafiksel Yöntemler

Normal P-P plot grafiđi

- İncelemek istediđiniz deđişkenleri seçiniz.
- Seçilen deđişkeni “Variables” kutusuna aktarmak için ok işaretine tıklayınız.
- “OK”a tıklayınız.

*screen.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

1:

	subno	timedrs	attdrug	athouse	income	mstatus	race	var	var	var	var
1	548	81	8	24	9	1	1				
2	398	75	9	33	9	1	1				
3	262	52	9	31	4	2	2				
4	235	60	10	29	4	1	1				
5	48	60	7	24	1	2	1				
6	330	58									
7	137	30									
8	276	57									
9	370	57									
10	502	56									
11	463	5									
12	73	8									
13	291	52									
14	124	1									
15	122	27									
16	400	3									
17	80	7									
18	379	43									
19	338	2									
20	155	1									
21	119	0									
22	253	5									
23	265	14									
24	230	49									
25	45	0									

P-P Plots

Variables:

- Subject number [su...
- Whether currently m...
- race
- Visits to health profess...
- Attitudes toward medic...
- Attitudes toward house...
- income

Test Distribution

Normal

df:

Distribution Parameters

☒ Estimate from data

Location: 0

Scale: 1

Proportion Estimation Formula

☒ Blom's ☐ Rankit ☐ Tukey's

☐ Van der Waerden's

Rank Assigned to Ties

☒ Mean ☐ High ☐ Low

☐ Break ties arbitrarily

Transform

☐ Natural log transform

☐ Standardize values

☐ Difference: 1

☐ Seasonally difference: 1

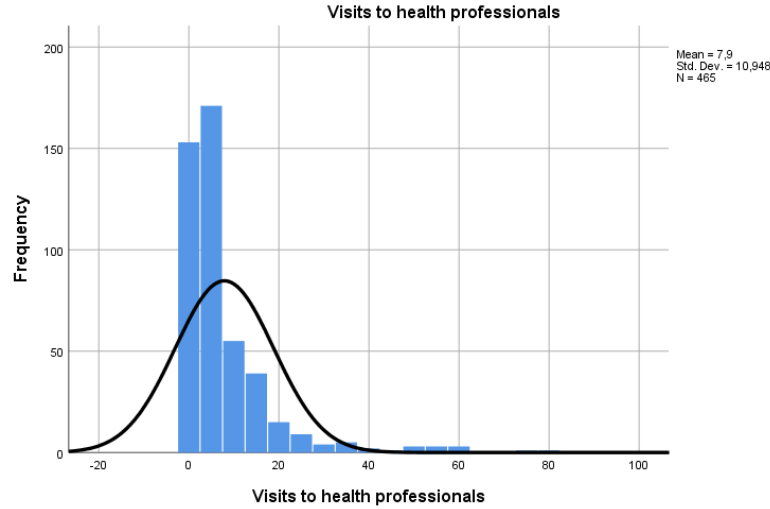
Current Periodicity: None

OK Paste Reset Cancel Help



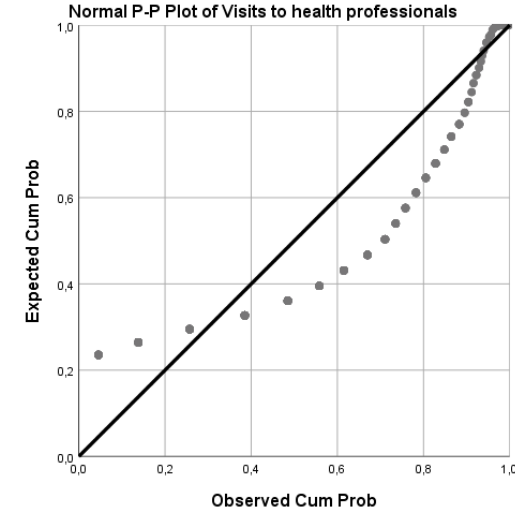
Varsayımlar (Normallik) Grafiksel Yöntemler

Histogram



Histogram

Visits to health professionals



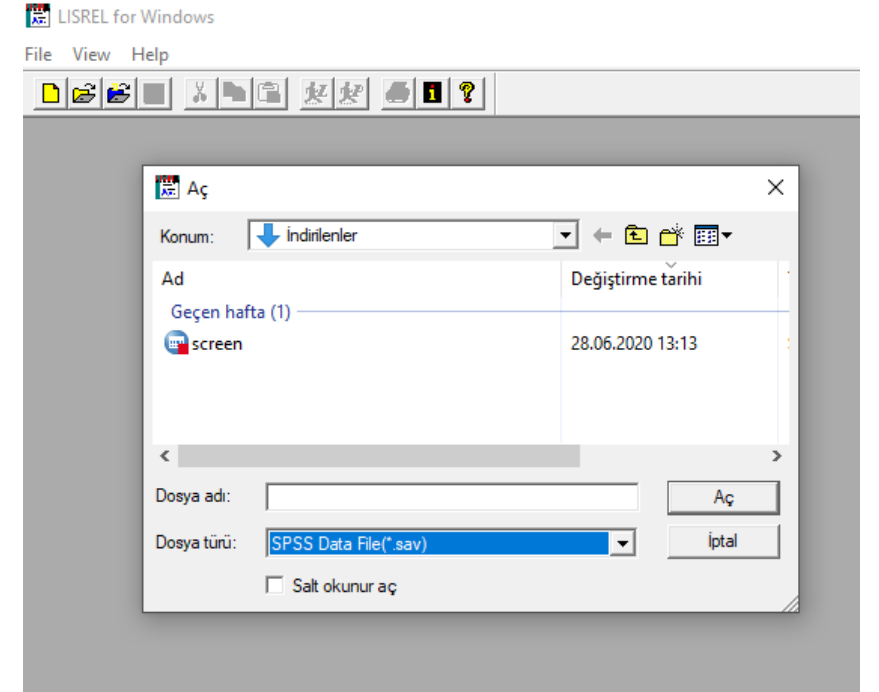
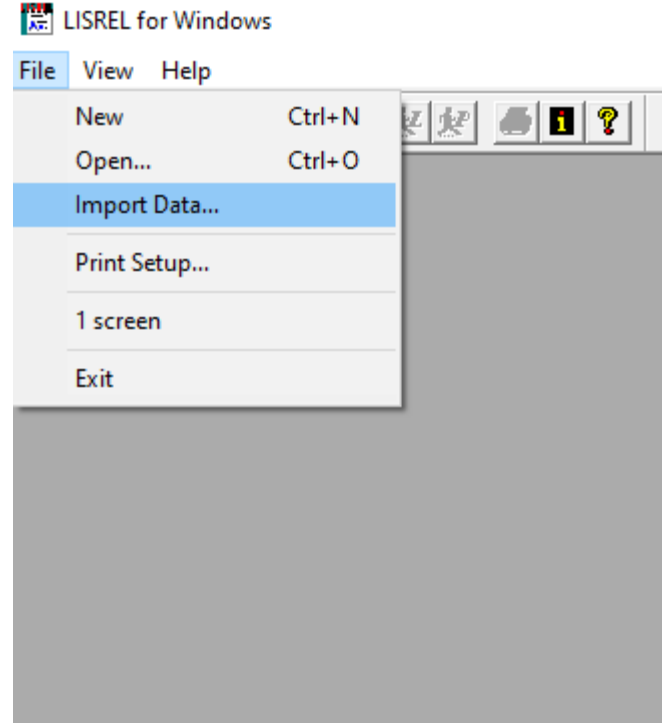
Normal P-P plot grafiğı

İncelenen değişken için histogram incelendiğinde dağılımın sağa çarpık ve sivri olduğu görülmektedir. P-P grafiğı incelendiğinde gözlemlerin köşegenden uzaklaştığı görülmektedir. Bu durumda dağılımın normal dağılımdan uzaklaştığı söylenebilir.



Çok değişkenli normallik

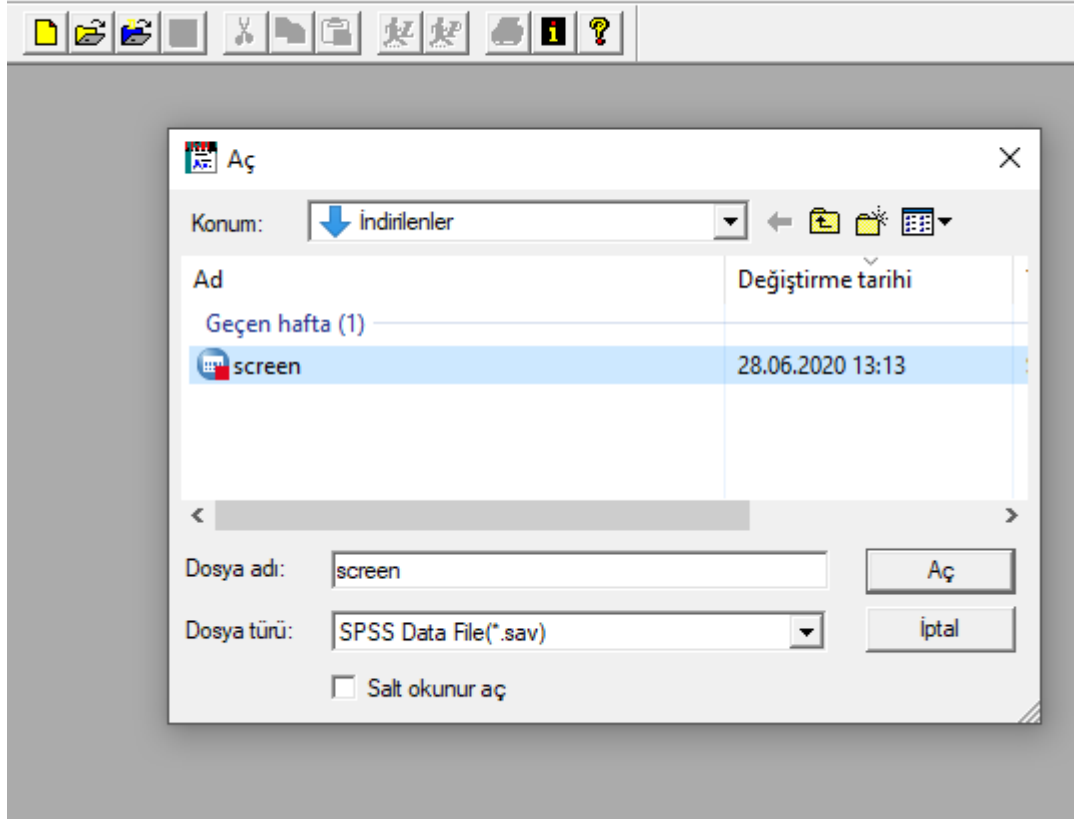
- LISREL, R gibi programlarda çok değişkenli normallik testi tek bir analizle yapılabilir.



Çok değişkenli normallik

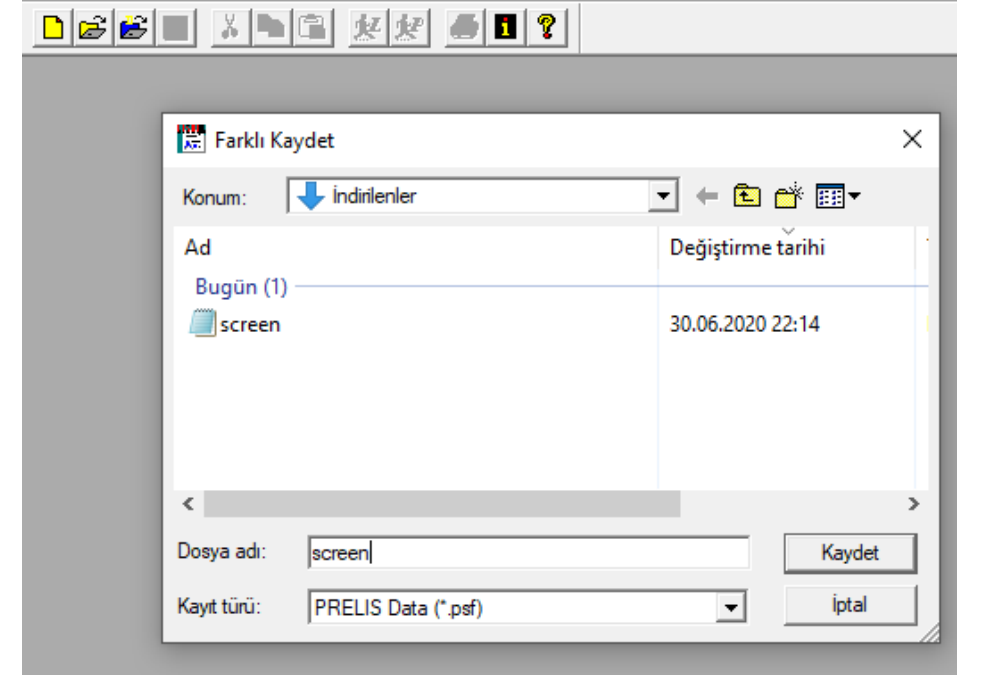
LISREL for Windows

File View Help



LISREL for Windows

File View Help



Çok değişkenli normallik

LISREL for Windows - screen

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Help

screen

	subno	timedrs	attdrug	atthouse	income	mstatus	race
1	1,00	1,00	8,00	27,00	5,00	2,00	1,00
2	2,00	3,00	7,00	20,00	6,00	2,00	1,00
3	3,00	0,00	8,00	23,00	3,00	2,00	1,00
4	4,00	13,00	9,00	28,00	8,00	2,00	1,00
5	5,00	15,00	7,00	24,00	1,00	2,00	1,00
6	6,00	3,00	8,00	25,00	4,00	2,00	1,00
7	7,00	2,00	7,00	30,00	6,00	2,00	1,00
8	8,00	0,00	7,00	24,00	6,00	2,00	1,00
9	9,00	7,00	7,00	20,00	2,00	2,00	1,00
10	10,00	4,00	8,00	30,00	8,00	1,00	1,00
11	11,00	15,00	9,00	15,00	7,00	2,00	1,00
12	12,00	0,00	6,00	22,00	3,00	2,00	1,00
13	13,00	2,00	6,00	19,00	5,00	2,00	1,00
14	14,00	13,00	8,00	25,00	6,00	2,00	1,00
15	15,00	2,00	5,00	17,00	1,00	2,00	1,00
16	16,00	2,00	8,00	19,00	3,00	2,00	2,00
17	21,00	1,00	8,00	22,00	1,00	2,00	1,00
18	22,00	2,00	6,00	21,00	7,00	1,00	1,00
19	23,00	5,00	8,00	28,00	2,00	2,00	1,00
20	24,00	5,00	10,00	25,00	9,00	2,00	1,00
21	25,00	3,00	6,00	19,00	4,00	2,00	1,00
22	26,00	4,00	5,00	31,00	5,00	2,00	1,00

LISREL for Windows - screen

File Edit Data Transformation Statistics Graphs Multilevel SurveyGLIM View Window Hel

Data

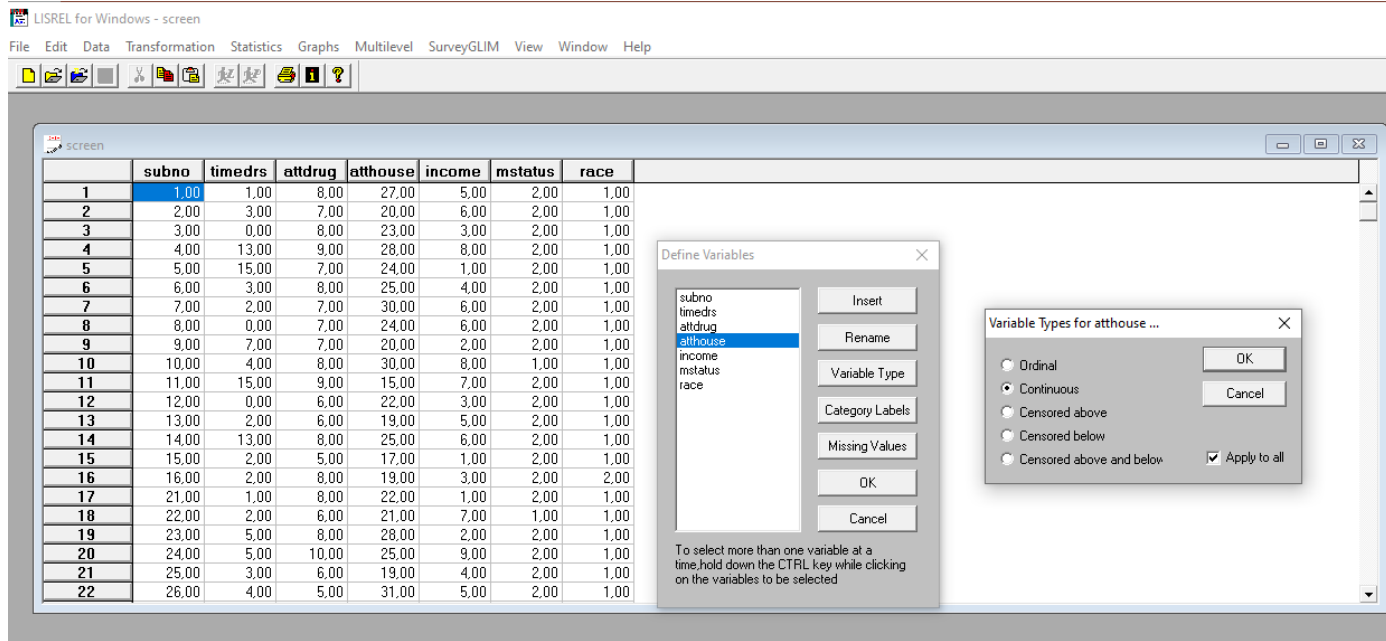
- Define Variables ...
- Select Variables/Cases...
- Sort Case...
- Insert Variable
- Insert Cases
- Delete Variable
- Delete Case
- Weight Cases...
- Survey Design...

Data menüsünden 'Define Variables' Seçilir

	subno	timedrs	attdrug	atthouse	income	mstatus	race
6	6,00	3,00	8,00	25,00	4,00	2,00	1,00
7	7,00	2,00	7,00	30,00	6,00	2,00	1,00
8	8,00	0,00	7,00	24,00	6,00	2,00	1,00
9	9,00	7,00	7,00	20,00	2,00	2,00	1,00
10	10,00	4,00	8,00	30,00	8,00	1,00	1,00
11	11,00	15,00	9,00	15,00	7,00	2,00	1,00
12	12,00	0,00	6,00	22,00	3,00	2,00	1,00
13	13,00	2,00	6,00	19,00	5,00	2,00	1,00
14	14,00	13,00	8,00	25,00	6,00	2,00	1,00
15	15,00	2,00	5,00	17,00	1,00	2,00	1,00
16	16,00	2,00	8,00	19,00	3,00	2,00	2,00
17	21,00	1,00	8,00	22,00	1,00	2,00	1,00
18	22,00	2,00	6,00	21,00	7,00	1,00	1,00
19	23,00	5,00	8,00	28,00	2,00	2,00	1,00
20	24,00	5,00	10,00	25,00	9,00	2,00	1,00
21	25,00	3,00	6,00	19,00	4,00	2,00	1,00
22	26,00	4,00	5,00	31,00	5,00	2,00	1,00



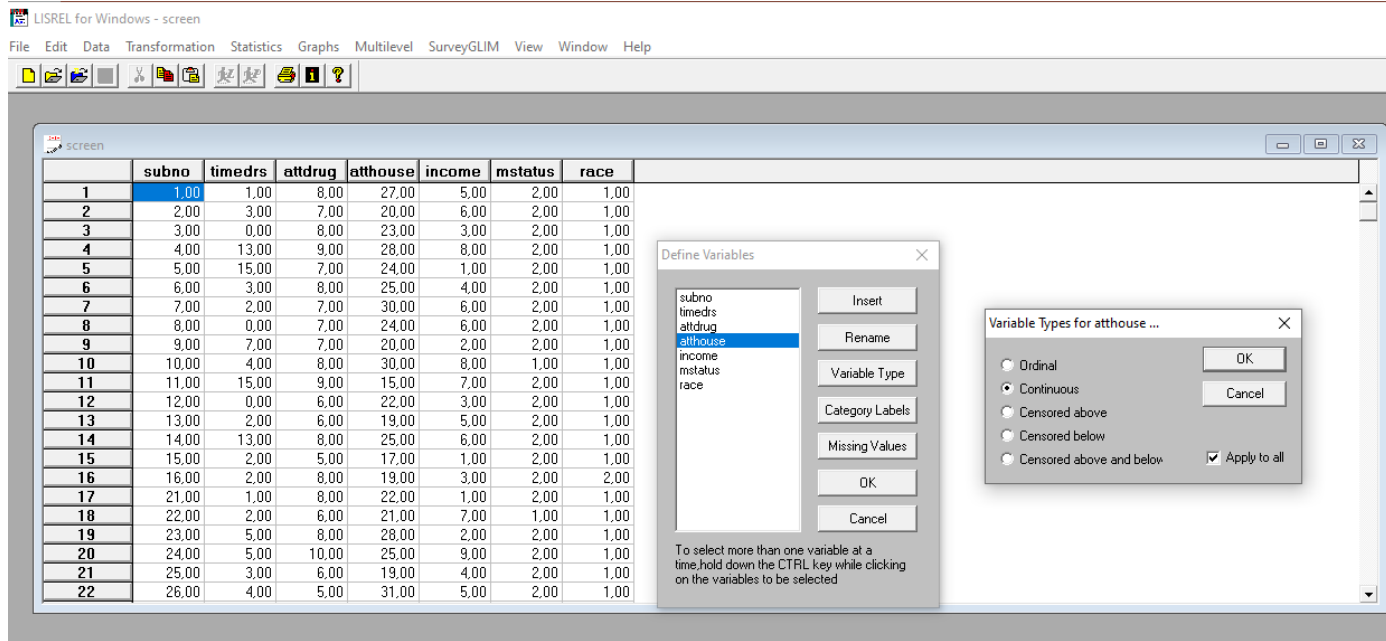
Çok değişkenli normallik



- 'Define Variables' penceresinden bir değişken seçiniz.
- 'Variable Type' a tıklayınız.
- 'Continuous' işaretleyiniz.
- Tüm verilere uygulamak için 'Apply to all' işaretleyiniz.
- Daha sonra 'OK' a tıklayınız



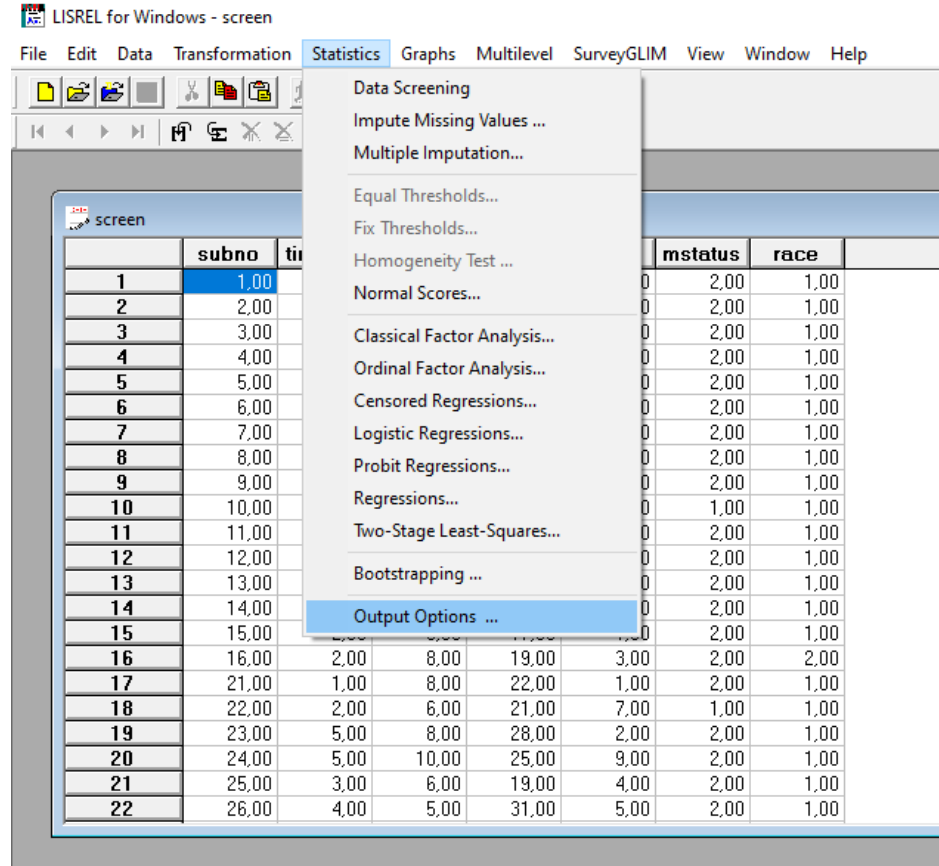
Çok değişkenli normallik



- ‘Define Variables’ penceresinden bir değişken seçiniz.
- ‘Variable Type’ a tıklayınız.
- ‘Continuous’ işaretleyiniz.
- Tüm verilere uygulamak için ‘Apply to all’ işaretleyiniz.
- Daha sonra ‘OK’ a tıklayınız



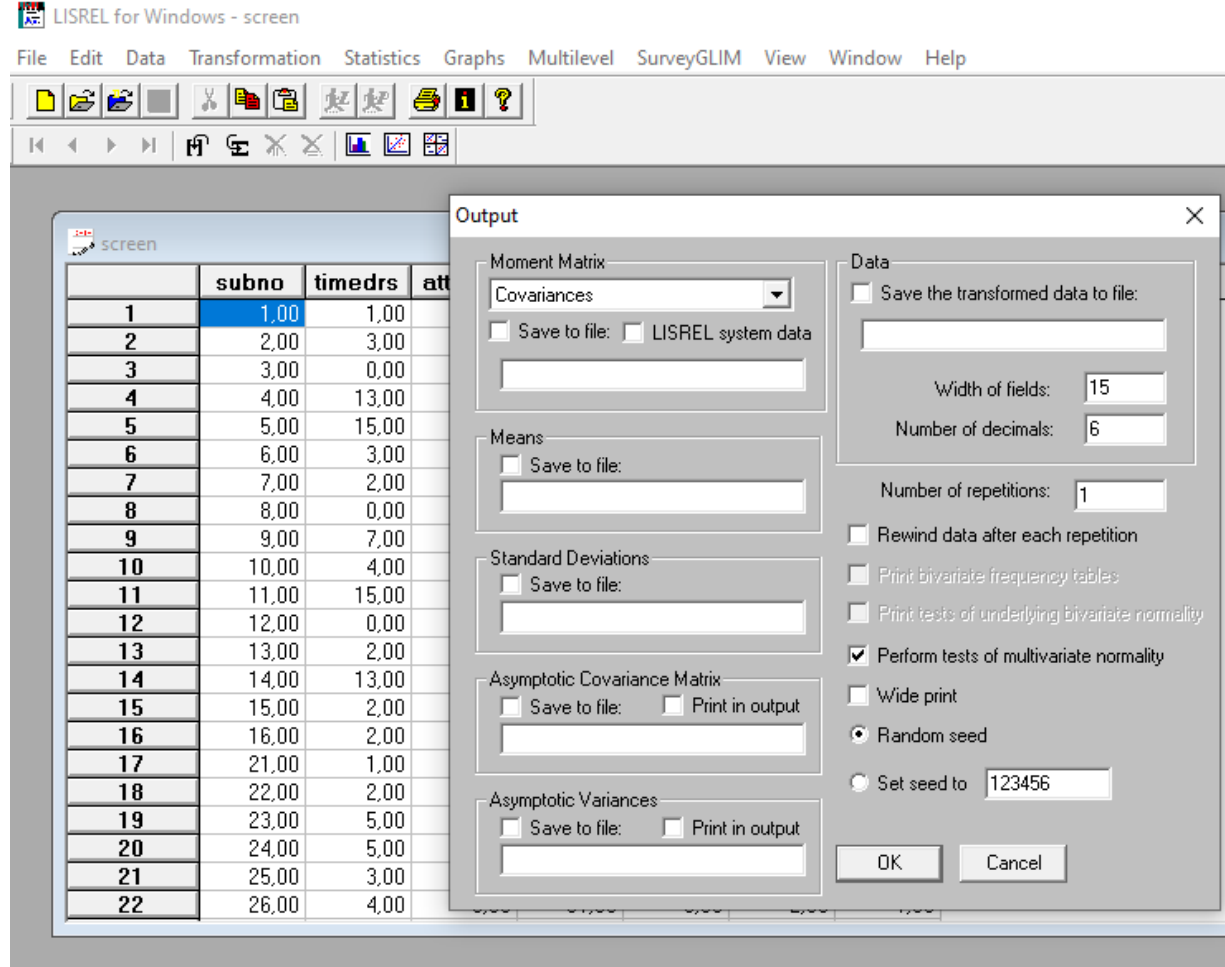
Çok değişkenli normallik



- 'Statistics' menüsünden 'Output Options'a tıklayınız.



Çok değişkenli normallik



- Açılan 'Output' penceresinden **'Perform tests of multivariate normality'** seçiniz.



Çok değişkenli normallik

LISREL for Windows - screen.OUT

File Edit Options Window Help

screen

screen.OUT

|Test of Univariate Normality for Continuous Variables

	Skewness		Kurtosis		Skewness and Kurtosis		
Variable	Z-Score	P-Value	Z-Score	P-Value	Chi-Square	P-Value	
subno	1.254	0.210	-9.018	0.000	82.904	0.000	
timedrs	14.987	0.000	10.055	0.000	325.778	0.000	
attdrug	-1.010	0.310	-2.531	0.011	7.439	0.024	
atthouse	-3.595	0.000	4.168	0.000	30.296	0.000	
income	4.711	0.000	-1.818	0.069	25.504	0.000	
mstatus	-9.018	0.000	-1.030	0.303	82.386	0.000	
race	14.431	0.000	8.412	0.000	279.014	0.000	

Relative Multivariate Kurtosis = 1.319

Test of Multivariate Normality for Continuous Variables

	Skewness		Kurtosis		Skewness and Kurtosis		
Value	Z-Score	P-Value	Value	Z-Score	P-Value	Chi-Square	P-Value
22.491	32.981	0.000	83.124	11.030	0.000	1209.373	0.000

- Açılan out dosyasında hem de tek değişkenli normallik hem de çok değişkenli normallik için istatistikler mevcuttur.
- Normallığın sağlanıp sağlanmadığını değerlendirmek için p değerleri incelenir.
- Anlamlı olmayan p değerleri normal dağılım varsayımının sağlandığını göstermektedir.
- Bununla birlikte RMK değerinin 1 civarında olması çok değişkenli normalliğin sağlandığına işaret eder.



Varsayımlar (Doğrusallık)

- Doğrusallık varsayımı, iki değişken arasında doğrusal bir ilişki olmasıdır.
- Doğrusallık pratik anlamda önemlidir, çünkü Pearson'un r değişkenler arasındaki doğrusal ilişkileri yakalar; değişkenler arasında önemli ölçüde doğrusal olmayan ilişkiler varsa, bunlar göz ardı edilir.
- İki değişken arasındaki doğrusallık kabaca iki değişkenli saçılım grafiklerinin incelenmesi ile değerlendirilir.
 - Her iki değişken de normal olarak dağılmışsa ve doğrusal olarak ilişkiliyse, saçılım grafiği oval şeklindedir.
 - Değişkenlerden biri normal değilse, bu değişken ile diğeri arasındaki saçılım grafiği oval değildir.



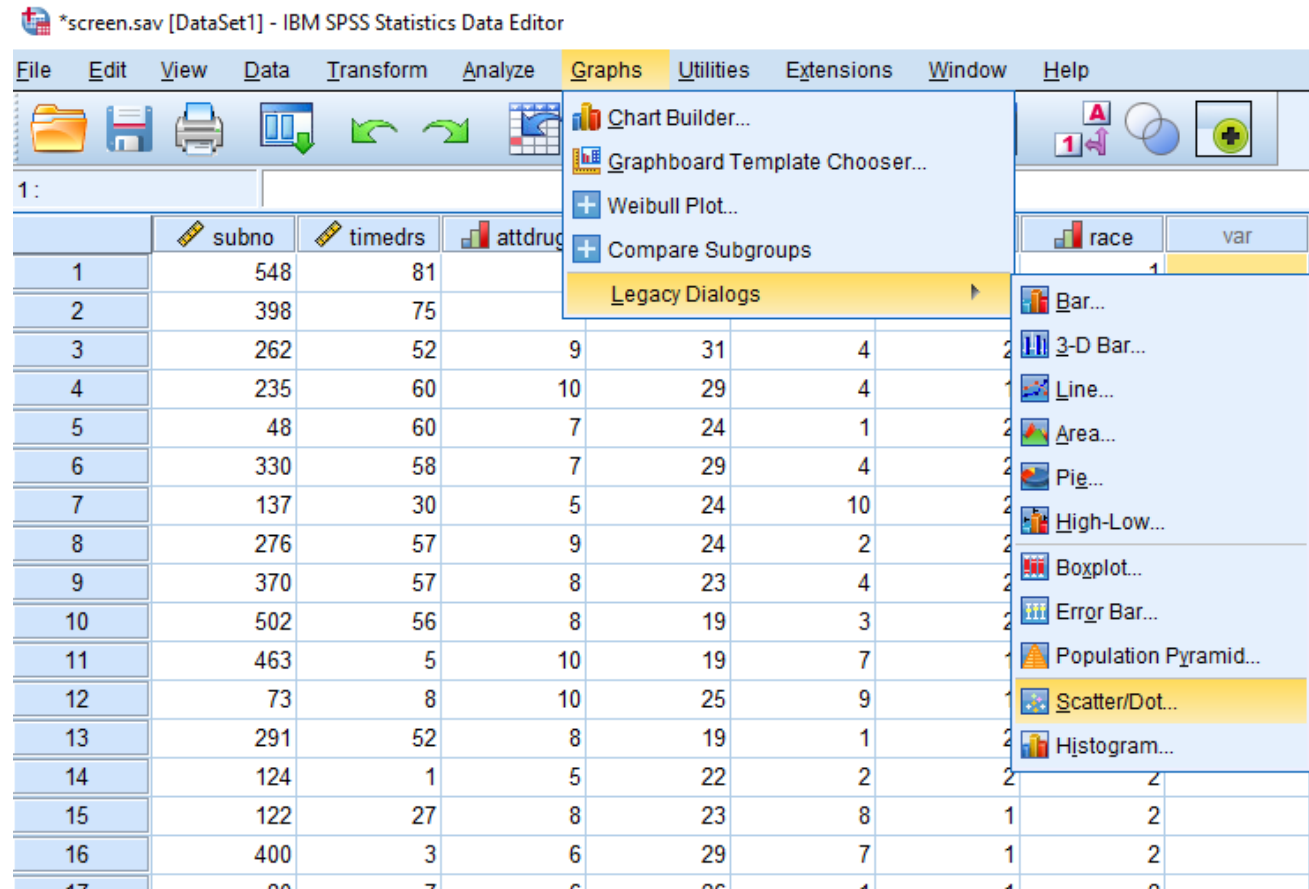
Varsayımlar (Doğrusallık)

- Doğrusallığı iki değişkenli saçılım grafikleri ile değerlendirmek, özellikle küçük örneklerle kolay olabilir ancak birkaç değişken varsa ve olası tüm çiftler incelendiğinde, özellikle veri gruplandırılmışsa ve analiz her grup içinde ayrı ayrı yapıldığında, işler daha karmaşıklaşır.
- Yalnızca birkaç değişken varsa, olası tüm çiftleri taramak çok külfetli değildir; ama çok sayıda değişken varsa, yalnızca doğrusallıktan uzaklaşma olasılığı olan çiftleri taramak için çarpıklık değerleri incelenebilir.



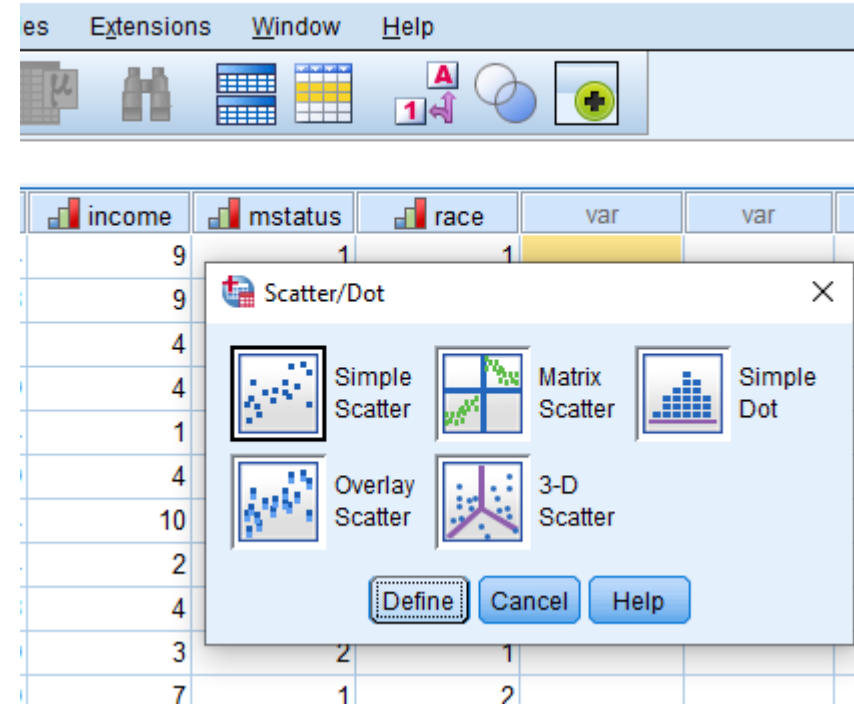
Varsayımlar (Doğrusallık) SPSS Uygulama

- “**Graphs**” ana menüsünden önce “**Legacy Dialogs**”, daha sonra da “**Scatter/Dot**”a tıklayınız.



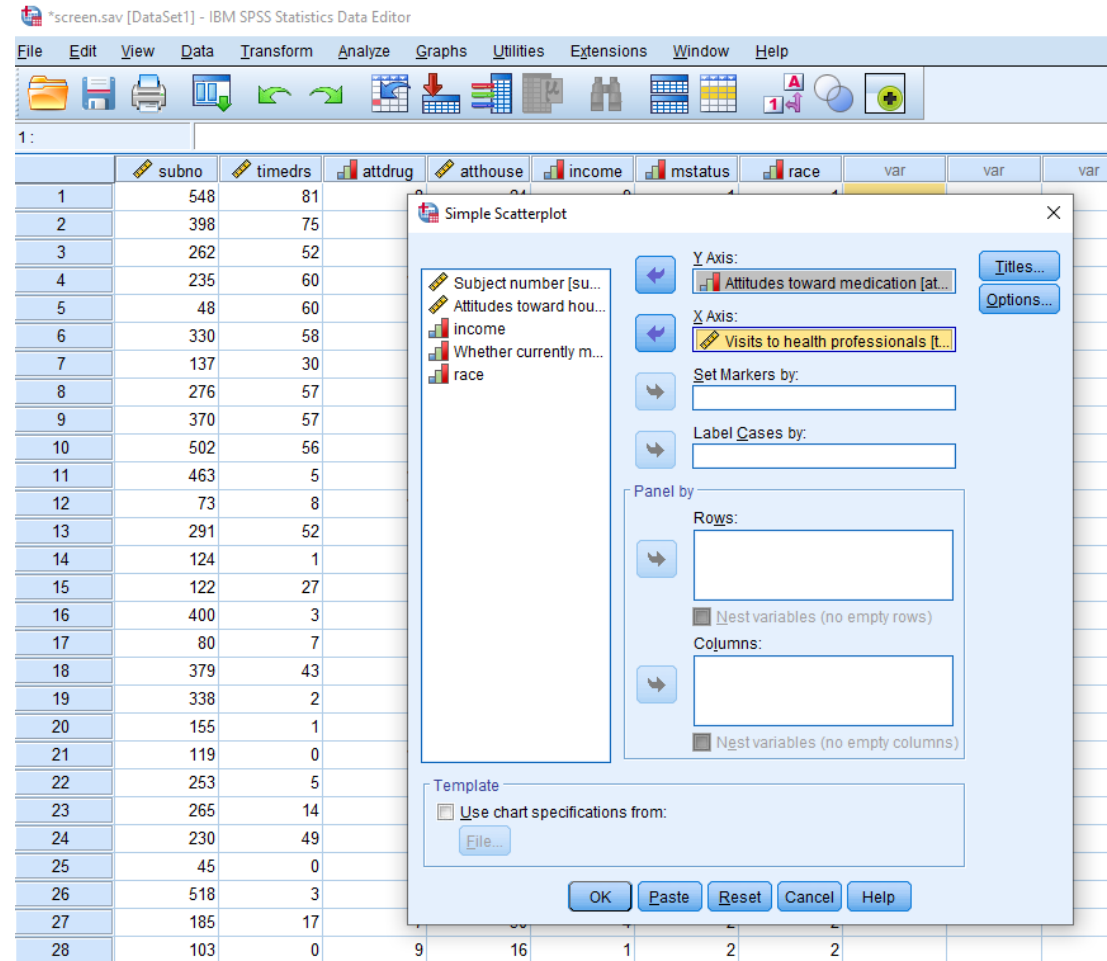
Varsayımlar (Doğrusallık) SPSS Uygulama

- “Simple Scatter” kutusunu seçiniz.
- “Define”a tıklayınız.



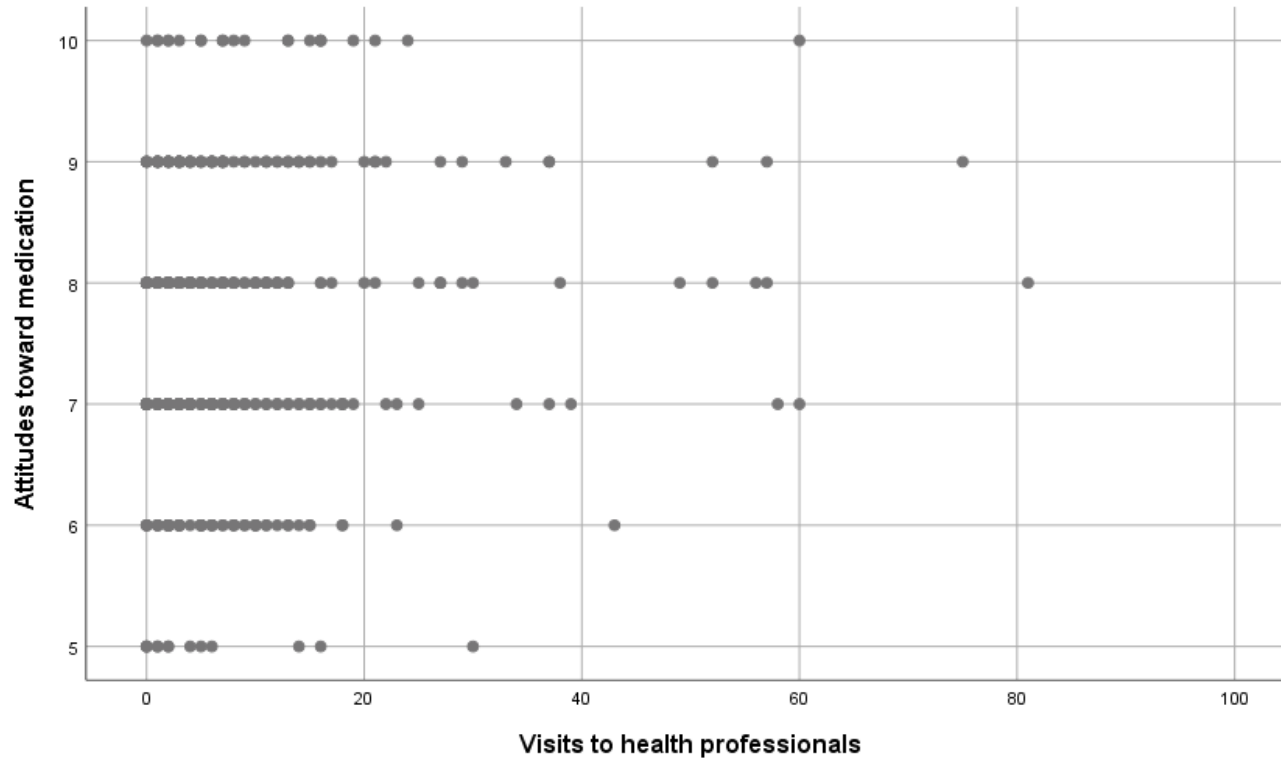
Varsayımlar (Doğrusalık) SPSS Uygulama

- İncelemek istediğiniz iki değişkenden birisini (attdrug) seçiniz ve seçilen değişkeni “Y Axis” kutusuna aktarınız
- Diğer değişkeni (timedrs) “X Axis” kutusuna taşıyınız
- “OK”e tıklayınız.



Varsayımlar (Doğrusallık) SPSS Uygulama

Graph



- Saçılma grafiğinin şekli oval değildir. Bu durumda ilgili iki değişken arasındaki ilişkinin doğrusal olmadığı ifade edilebilir

