

Verilerin Temizlenmesi ve Varsayımların Kontrolü



Arş. Gör. Sinem Şenferah

Kastamonu Üniversitesi Eğitim Fakültesi

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı

16 Mayıs 2021

Veri Taramasında Önemli Konular

- Verilerin toplandıktan sonra genel analizlere hazır hale getirilmesi gerekmektedir. Düzgün ve doğru bir veri analizi için veri tarama oldukça önemlidir.
- Veri taramada dikkat edilmesi gereken konular;
 - Veri dosyasına verinin doğru girilip girilmediği
 - Kayıp verinin miktarı ve dağılımı
 - Tek değişkenli ve çok değişkenli uç değerler
 - Varsayımlar (Normallik, Dorusallık, Eş varyanslılık)
 - Çoklu Bağlantı problemi (Multicollinearity) ve Tekillik

Tabachnick & Fidell (2012)



Veri Dosyasının Doğruluğu

- Veri dosyasının doğruluğu için, ilk olarak bilgisayarda veri girişi esnasında hata yapıp yapılmadığı kontrol edilmelidir. Veri girişinde yapılan hatalar, analizleri önemli ölçüde etkileyebilmektedir.
 - Örn; ‘2’ yerine ‘23’ girilmesi bir korelasyon analizinin sonuçlarını saptırabilir (Atar, 2016)
- Veri dosyasının doğruluğu;
 - Küçük veri dosyalarında orijinal veri ile bilgisayara girilen veri karşılaştırılarak kontrol edilir
 - Büyük veri dosyalarında ise her bir değişken için betimsel istatistikler ve grafikler yoluyla kontrol edilir
 - Örn; Ortalama ve standart sapmalar mantıklı mı? Sürekli değişkenler için bütün değerler ranjın içinde mi? Kategorik bir değişken için girilen değer, kategorileri tanımlayan sayıların dışında mı?



Veri Dosyasının Doğruluğu

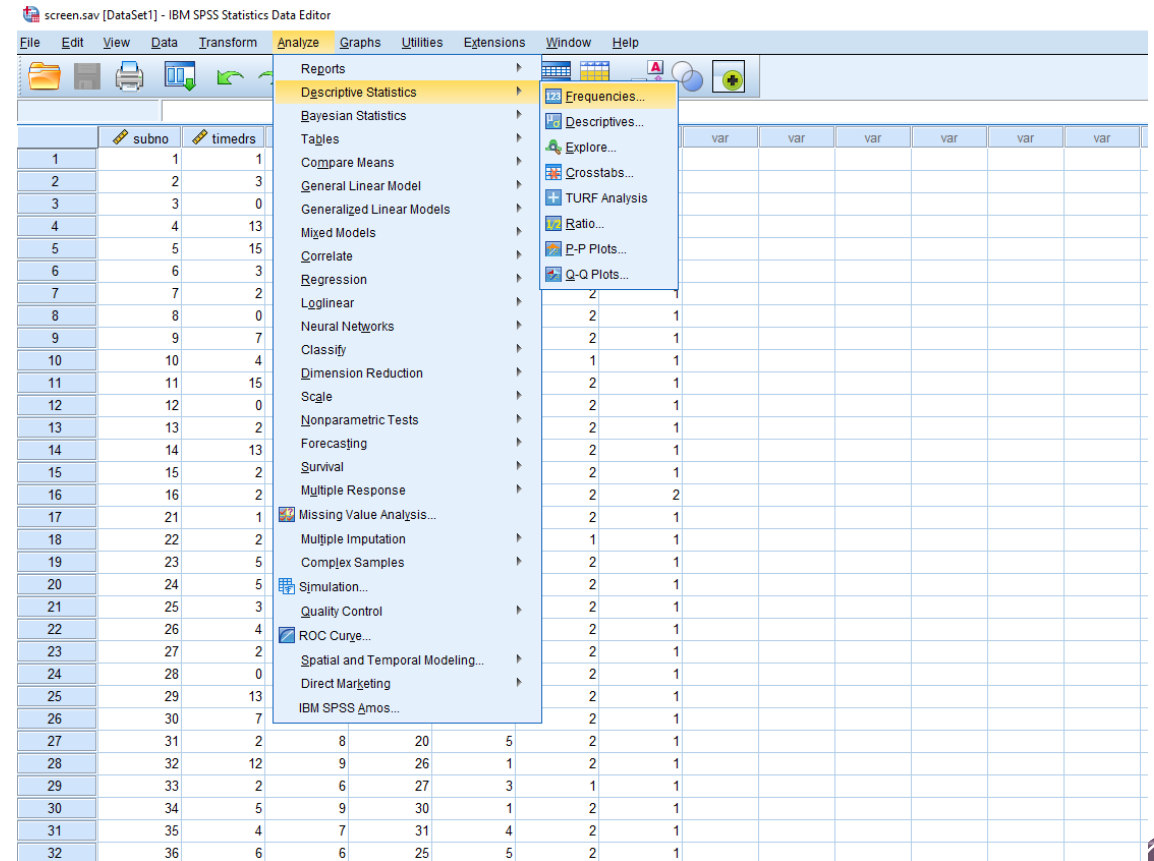
- Herhangi bir değişkene ait **aralık dışı** bir değer bulunduğunda, bu değer hangi bireye ait olduğunu bulmak için değişken **küçükten büyüğe veya büyükten küçüğe sıralanabilir**.
- Hatalı değer bulunduktan sonra,
 - Eğer **orijinal veriye** ulaşmak mümkünse **bu değer düzeltilir**,
 - Eğer orijinal veriye ulaşmak mümkün değilse ilgili değer boş bırakılır ve kayıp veri olarak değerlendirilir .
- Hatalar düzeltildikten sonra veri dosyasının doğruluğu tekrar kontrol edilir, bazen bir hata düzeltilirken yanlışlıkla başka bir hata yapılabilir.

(Tabachnick & Fidell, 2012; Atar, 2016)



Veri Dosyasının Doğruluğu SPSS Uygulama

- Verinin doğruluğunu kontrol etmek amacıyla SPSS programı kullanılarak betimleyici istatistikler şu şekilde elde edilebilir;
- Önce ana menüden “**Analyze**” sonra “**Descriptive Statistics**” daha sonra “**Frequencies**”e tıklayınız.



Veri Dosyasının Doğruluğu SPSS Uygulama

- Kontrol etmek istediğiniz değişkenleri seçiniz

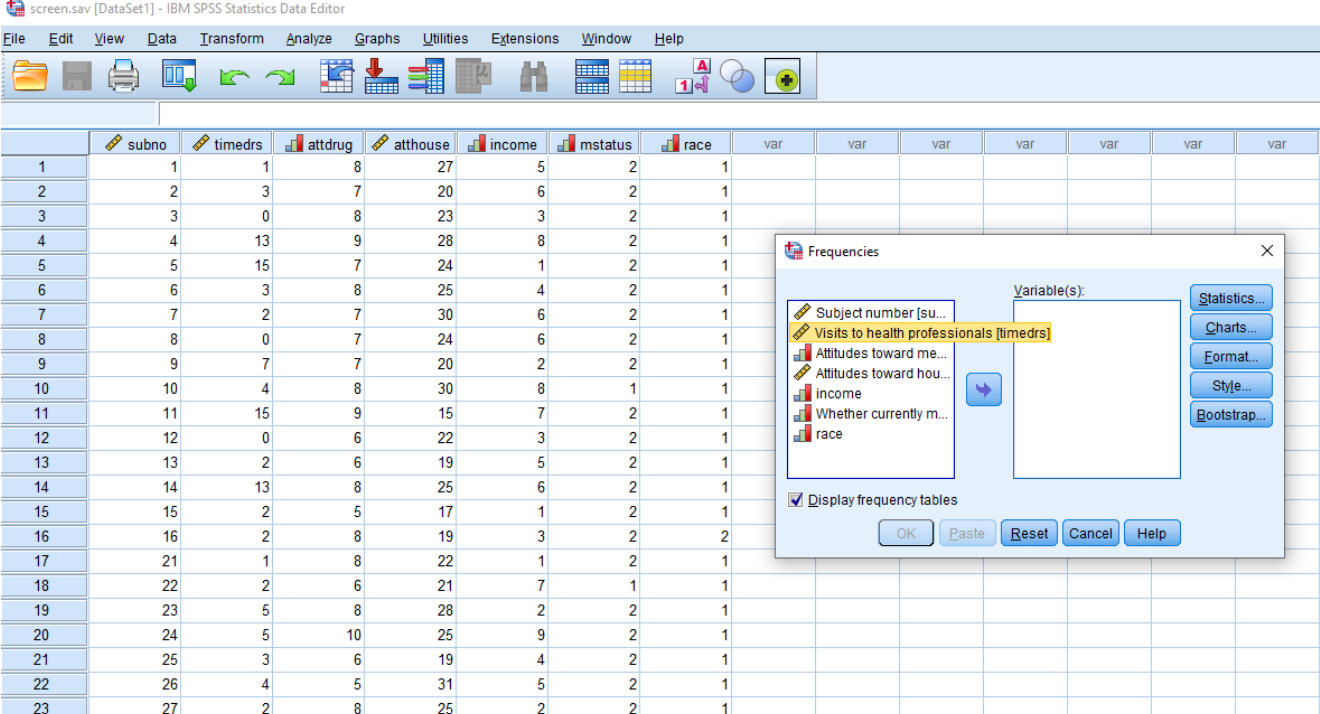
(Örn; timedrs, attdrug, atthouse, income, mstatus veya race)

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



Veri Dosyasının Doğruluğu SPSS Uygulama

- Seçilen değişkeni “Variable(s)” kutusuna aktarmak için ok işaretini tıklayınız.



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

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

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

Frequencies

Variable(s):

Subject number [su...]
Visits to health professionals [timedrs]
Attitudes toward me...
Attitudes toward hou...
income
Whether currently m...
race

Display frequency tables

OK Paste Reset Cancel Help



Veri Dosyasının Doğruluğu SPSS Uygulama

- **Statistics**'e tıklayınız. “Central Tendency” bölümünde “Mean”i, “**Dispersion**” bölümündeysse “Std. deviation”, “Minimum” ve “Maximum”u seçiniz.
- Önce “Continue”, sonra “OK”e tıklayınız.

The screenshot displays the IBM SPSS Statistics Data Editor interface. The main window shows a dataset with the following columns: subno, timedrs, attdrug, atthouse, income, mstatus, race, and several empty 'var' columns. The data is organized into rows, with the first row containing values for each column. Overlaid on the main window are two dialog boxes. The 'Frequencies' dialog box is open, showing a list of variables on the left and 'Visits to health profe...' selected in the 'Variable(s):' box. The 'Display frequency tables' checkbox is checked. The 'Frequencies: Statistics' sub-dialog box is also open, showing options for Central Tendency (Mean, Median, Mode, Sum) and Dispersion (Std. deviation, Minimum, Maximum, S.E. mean, Variance, Range). The 'Std. deviation', 'Minimum', and 'Maximum' options are checked under the 'Dispersion' section. The 'Continue' button is highlighted in the 'Frequencies: Statistics' dialog.



Veri Dosyasının Doğruluğu SPSS Uygulama

*Output3 [Document3] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Log
Frequencies
Title
Notes
Statistics
Frequency Table
Title
Visits to health
Attitudes toward
Attitudes toward
income
Whether currently
race

FREQUENCIES VARIABLES=timedrs attdrug atthouse income mstatus race
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN
/ORDER=ANALYSIS.

➔ Frequencies

Statistics

		Visits to health professionals	Attitudes toward medication	Attitudes toward housework	income	Whether currently married	race
N	Valid	465	465	464	439	465	465
	Missing	0	0	1	26	0	0
Mean		7,90	7,69	23,54	4,21	1,78	1,09
Std. Deviation		10,948	1,156	4,484	2,419	,416	,284
Minimum		0	5	2	1	1	1
Maximum		81	10	35	10	2	2

- Statistics” tablosunda her bir değişken için **ortalama**, **standart sapma**, **minimum ve maksimum** değerleri incelenir.



Kayıp Veri

- Kayıp veri, veri analizinde en yaygın sorunlardan biridir.
- Kayıp veri, örüntüsüne ve miktarına bağlı olarak ciddi sorunlar oluşturabilir.
- Bir değişkene ait beklenmeyen miktarda kayıp veri varsa, ilk olarak bunun nedeni araştırılmalıdır. Daha sonra kayıp verinin örüntüsüne bakılarak, **rastgele mi** yoksa **sistematik** bir örüntü mü gösterdiği belirlenmelidir (Atar, 2016).
 - Örneğin, katılımcılar gelir ile ilgili soruları cevaplamak istemeyebilir.
- **Kayıp verinin örüntüsü, miktarından daha önemlidir.** Veri matrisi içinde random dağılan kayıp veri daha az sorun oluştururken, ne kadar az olursa olsun random dağılmayan kayıp veri sonuçların genellenebilirliğini etkileyebileceğinden daha ciddi problemlere yol açar (Tabachnick & Fidell, 2012).



Kayıp Veri Özellikleri

Kayıp veri

1. Tamamen rastlantısal olarak kayıp (missing completely at random, **MCAR**),
 2. Rastlantısal olarak kayıp (missing at random, **MAR**) veya
 3. missing not at random or nonignorable, **MNAR** şeklinde sınıflandırmıştır.
- Kayıp verinin dağılımı, MCAR'da tahmin edilemezdir. Bu durum kayıp veri için en iyi durumdur.
 - MAR'da ise, kayıp verinin örüntüsü, veri setindeki diğer değişkenlerden tahmin edilebilir. MNAR'da kayıp veri, değişkenin kendisiyle ilişkilidir ve bu nedenle göz ardı edilemez.



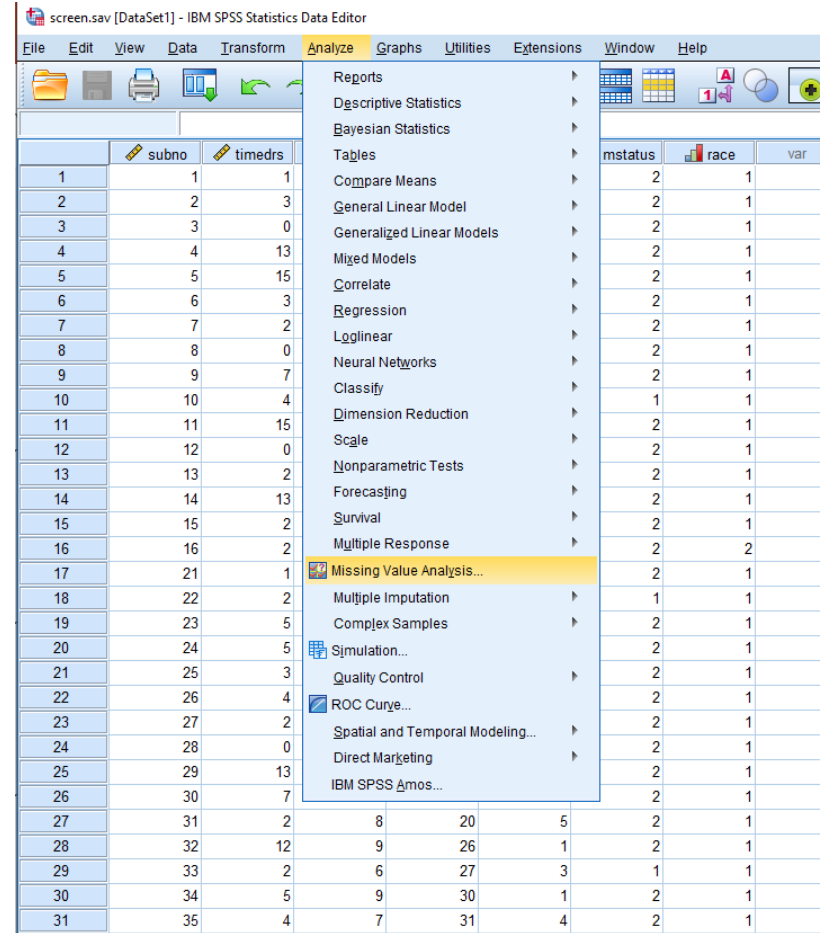
Kayıp Veri Analizi

- Büyük bir veri setinde rastgele bir desende yalnızca birkaç veri noktası (**örneğin, %5 veya daha az**) kayıpsa, sorunlar daha az ciddidir ve kayıp değerlerle baş etmek için hemen hemen her yöntem benzer sonuçlar verir.
- Bununla birlikte, küçük ve orta ölçekli bir veri kümesinden çok fazla veri kayıpsa, sorunlar çok ciddi olabilir. Ne yazık ki, belirli bir boyuttaki örneklem için ne kadar kayıp verinin tolere edilebileceğine dair henüz kesin bir kılavuz bulunmamaktadır.
- Eldeki veriler kullanılarak kayıp verideki örüntüler test edilir.
 - Örn; kayıp verinin bulunduğu gelir değişkeni için kayıp veriye sahip olan ve olmayan bireylerden 2 grup oluşturulup bu gruplar arasındaki gelir ile ilgili başka bir değişken için ortalama farklılıklara yönelik bir t testini yapılabilir. Hiçbir fark yoksa, kayıp verilerle nasıl başa çıkılacağı ile ilgili kararlar çok kritik değildir. Separate Variance t testleri, income değişkeni üzerindeki kayıp ile diğer değişkenlerden herhangi biri arasında sistematik bir ilişki olmadığını göstermektedir.



Kayıp Veri Analizi SPSS Uygulama

- **SPSS MISSING VALUE ANALYSIS** (MVA) ile kayıp değerlerin örüntüsü incelenebilir.
- Ana menüden önce **“Analyze”**, sonra **“Missing Value Analysis”**e tıklayınız.



Kayıp Veri Analizi SPSS Uygulama

- İncelemek istediğiniz değişkenleri seçiniz (Örn; timedrs, attdrug, atthouse, income, mstatus ve race)
- Seçilen değişkeni “**Quantitative Variables**” kutusuna aktarmak için ok işaretini tıklayınız.

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

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

	subno	timedrs	attdrug	athouse	income	mstatus	race	var	var	var	var	var	var	var
1	1	1	8	27	5	2	1							
2	2	3	7	20	6	2	1							
3	3	0	8	23	3	2	1							
4	4	13	9	28	8	2	1							
5	5	15	7	24	1	2	1							
6	6	3	8	25	4	2	1							
7	7	2	7	30	6	2	1							
8	8	0	7	24	6	2	1							
9	9	7	7	20	2	2	1							
10	10	4	8	30	8	1	1							
11	11	15	9	15	7	2	1							
12	12	0	6	22	3	2	1							
13	13	2	6	19	5	2	1							
14	14	13	8	25	6	2	1							
15	15	2	5	17	1	2	1							
16	16	2	8	19	3	2	2							
17	21	1	8	22	1	2	1							
18	22	2	6	21	7	1	1							
19	23	5	8	28	2	2	1							
20	24	5	10	25	9	2	1							
21	25	3	6	19	4	2	1							
22	26	4	5	31	5	2	1							
23	27	2	8	25	2	2	1							
24	28	0	8	26	1	2	1							
25	29	13	9	26	2	2	1							
26	30	7	9	33	1	2	1							
27	31	2	8	20	5	2	1							
28	32	12	9	26	1	2	1							

Missing Value Analysis

Quantitative Variables:

Categorical Variables:

Maximum Categories: 25

Case Labels:

Use All Variables

OK Paste Reset Cancel Help



Kayıp Veri Analizi SPSS Uygulama

- “Patterns” tıklayınız.
- “Display” bölümünde “Cases with missing values, sorted by missing value patterns”ı seçiniz.
- “Continue”a tıklayınız.

The screenshot shows the SPSS Missing Value Analysis dialog box and its Patterns sub-dialog box. The main dialog box has the following settings:

- Quantitative Variables: (empty)
- Categorical Variables: (empty)
- Maximum Categories: 25
- Case Labels: (empty)
- Buttons: Use All Variables, OK, Paste, Reset, Cancel, Help

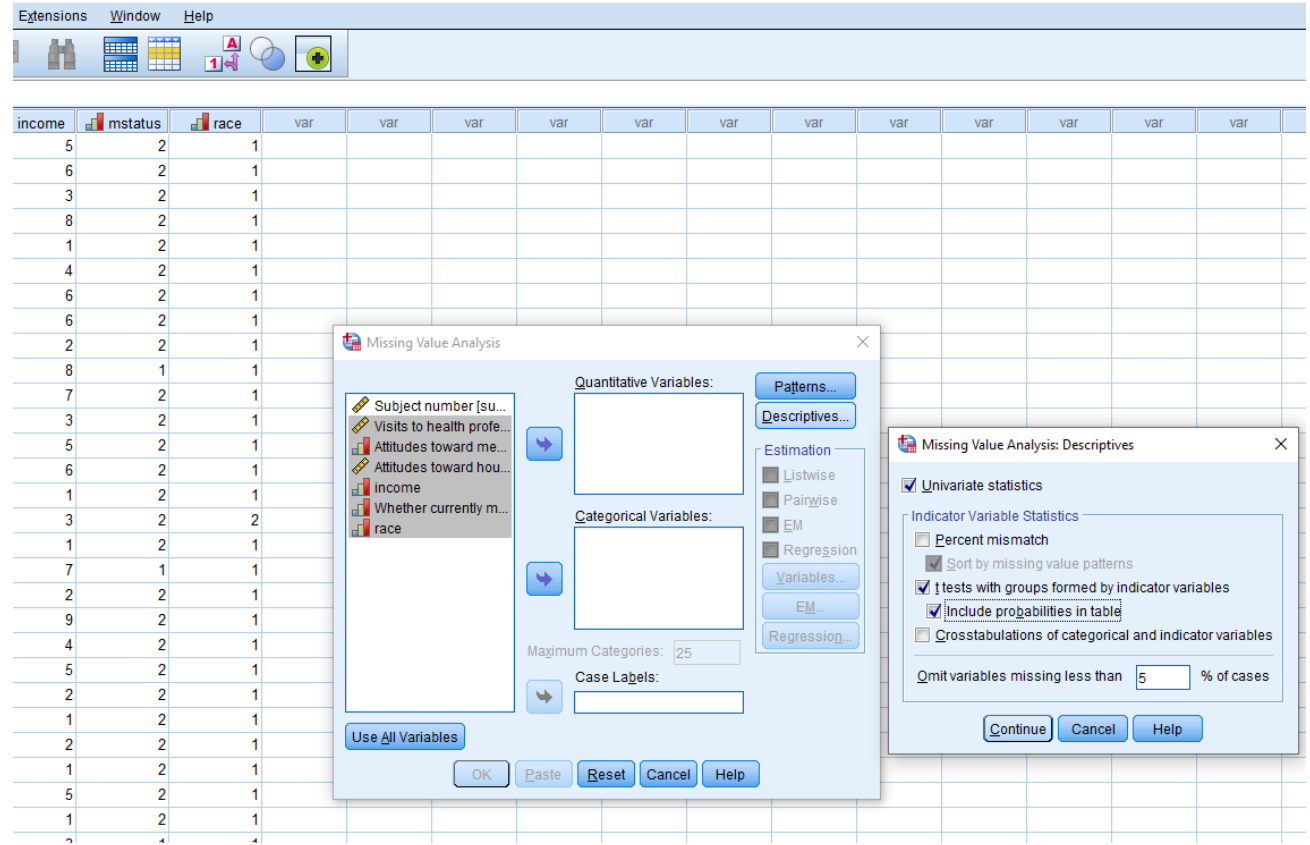
The Patterns sub-dialog box has the following settings:

- Display: ☒ Tabulated cases, grouped by missing value patterns. ☐ Sort variables by missing value pattern.
- Variables: ☒ Cases with missing values, sorted by missing value patterns. ☒ Sort variables by missing value pattern. ☐ All cases, optionally sorted by selected variable.
- Buttons: Continue, Cancel, Help



Kayıp Veri Analizi SPSS Uygulama

- “**Descriptives**” tıklayınız.
- “**Indicator Variable Statistics**” bölümünde “**t tests with groups formed by indicator variables**” ve “**Include probabilities in table**”¹ seçiniz.
- “Continue”a tıklayınız.



Kayıp Veri Analizi SPSS Uygulama

- Estimation” bölümünde “**EM**”i seçiniz.
- “OK”e tıklayınız.

Data Editor

Analyze Graphs Utilities Extensions Window Help

attdrug	atthouse	income	mstatus	race	var	var	var	var	var	var	var	var
8	27	5	2	1								
7	20	6	2	1								
8	23	3	2	1								
9	28	8	2	1								
7	24	1	2	1								
8	25	4	2	1								
7	30	6	2	1								
7	24	6	2	1								
7	20	2	2	1								
8	30	8	1	1								
9	15	7	2	1								
6	22	3	2	1								
6	19	5	2	1								
8	25	6	2	1								
5	17	1	2	1								
8	19	3	2	2								
8	22	1	2	1								
6	21	7	1	1								
8	28	2	2	1								
10	25	9	2	1								
6	19	4	2	1								
5	31	5	2	1								
8	25	2	2	1								
8	26	1	2	1								
9	26	2	2	1								
9	33	1	2	1								
8	20	5	2	1								
9	26	1	2	1								
6	27	3	1	1								
9	30	1	2	1								

Missing Value Analysis

Subject number [su...]

Quantitative Variables:

- Visits to health pr...
- Attitudes toward ...
- Attitudes toward ...
- Income
- Whether currentv...

Categorical Variables:

Maximum Categories: 25

Case Labels:

Use All Variables

Estimation

- ☐ Listwise
- ☐ Pairwise
- ☒ EM
- ☐ Regression

Variables...

EM...

Regression...

OK Paste Reset Cancel Help



Kayıp Veri Analizi SPSS Uygulama

*Output3 [Document3] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Log
Frequencies
Title
Notes
Statistics
Frequency Table
Title
Visits to health profess
Attitudes toward medic
Attitudes toward house
income
Whether currently marr
+ race
Log
MVA
Title
Notes
Univariate Statistics
Summary of Estimated Mea
Summary of Estimated Sta

```
/TTEST PROB PERCENT=5  
/EM(TOLERANCE=0.001 CONVERGENCE=0.0001 ITERATIONS=25) .
```

MVA

Univariate Statistics

	N	Mean	Std. Deviation	Missing		No. of Extremes ^a	
				Count	Percent	Low	High
timedrs	465	7,90	10,948	0	,0	0	34
attdrug	465	7,69	1,156	0	,0	0	0
atthouse	464	23,54	4,484	1	,2	4	0
income	439	4,21	2,419	26	5,6	0	0
mstatus	465	1,78	,416	0	,0	.	.
race	465	1,09	,284	0	,0	.	.

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

- Univariate Statistics” tablosu incelendiğinde veri setinde atthouse değişkeni için 1, income değişkeni içinse 26 kayıp değer olduğu ve atthouse değişkeni için kayıp veri yüzdesinin %5 i aşmadığı; income değişkeni için ise %5’i aştığı görülmektedir.



Kayıp Veri Analizi SPSS Uygulama

*Output3 [Document3] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output

- Log
- Frequencies
 - Title
 - Notes
 - Statistics
- Frequency Table
 - Title
 - Visits to health profess
 - Attitudes toward medic
 - Attitudes toward house
 - income
 - Whether currently mar
 - race
- Log
- MVA
 - Title
 - Notes
 - Univariate Statistics
 - Summary of Estimated Mea
 - Summary of Estimated Sta
 - Separate Variance t Tests
 - EM Estimated Statistics
 - Title

All Values	10,948	1,156	4,484	2,419	,416	,284
EM	10,948	1,156	4,484	2,416	,416	,284

Separate Variance t Tests^a

	times/r	attdrug	athouse	income	mstatus	race
t	,2	-1,1	-,2	.	-1,0	-,4
df	32,2	29,6	28,6	.	29,0	27,3
P(2-tail)	,846	,289	,851	.	,346	,662
# Present	439	439	438	439	439	439
# Missing	26	26	26	0	26	26
Mean(Present)	7,92	7,67	23,53	4,21	1,77	1,09
Mean(Missing)	7,62	7,88	23,69	.	1,85	1,12

For each quantitative variable, pairs of groups are formed by indicator variables (present, missing).

a. Indicator variables with less than 5% missing are not displayed.

- “Separate Variance t Tests” tablosu kayıp değerlerin yüzdesi %5’ten fazla olan income değişkeni için oluşturulmuştur. Bu tablo income değişkenindeki **veri kaybının kritik olmadığını** göstermektedir.



Kayıp Veri Analizi SPSS Uygulama

EM Correlations^a

	timeds	attdrug	atthouse	income	mstatus	race
timeds	1					
attdrug	,104	1				
atthouse	,128	,023	1			
income	,050	-,007	,002	1		
mstatus	-,065	-,006	-,030	-,466	1	
race	-,035	,019	-,038	,102	-,035	1

a. Little's MCAR test: Chi-Square = 18,675, DF = 10, Sig. = ,045

- EM Correlations” tablosunun altında verinin tamamen **rastlantısal olarak kayıp olup olmadığının test edildiği Little’ın MCAR** testi sonuçları yer alır.
- Bu testte istatistiksel olarak **anlamsız sonuç istenir**. Yani kayıp veride örüntü yoktur hipotezinin kabul edilmesi beklenir.
- *Örnekte* $p = 0.045$, kayıp verinin Rastlantısal olmadığını belirtir.



Kayıp Veri ile Başa Çıkma Yolları

- Liste Bazında Veri Silme (Listwise Deletion)
- Çiftler Bazında Veri Silme (Pairwise Deletion)
- Kayıp Verinin Tahmin Edilmesi (Estimating Missing Data)
 1. Ortalama Yerleştirme (Mean Substitution)
 2. Regresyon Yerleştirme (Regression Substitution)
 3. Beklenti Maksimizasyonu (Expectation-Maximization(EM))
 4. Çoklu hesaplama/yükleme (Multiple Imputation)



Liste Bazında Veri Silme (Listwise Deletion)

- Bir veya daha fazla değişkende kayıp veriye sahip **bireylerin analizden çıkarılmasıdır.**
- Liste bazında veri silme yöntemine göre income değişkenindeki kayıp değerlere sahip 584, 241, 512, ..., subno.lu bireyler analiz dışı bırakılacak, analiz diğer bireylerle gerçekleştirilecektir.

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

	subno	timeds	attdrug	atthouse	income	mstatus	race	var
1	-584	0	8	11	.	2	2	
2	-241	2	7	17	.	2	1	
3	-512	4	8	19	.	2	1	
4	239	8	7	20	.	2	1	
5	156	1	6	21	.	1	1	
6	321	1	9	21	.	2	1	
7	84	7	8	22	.	2	1	
8	453	13	9	22	.	2	1	
9	568	3	7	22	.	2	1	
10	570	4	8	22	.	1	1	



Çiftler Bazında Veri Silme (Pairwise Deletion)

- Bir değişkende kayıp değere **sahip bir bireyin, kayıp değere sahip olmadığı diğer değişkenlerde kullanılmasıdır.**
- Çiftler bazında veri silme verinin daha fazla kısmının kullanılmasına izin verir. Ancak hesaplanan her istatistik farklı **gözlem setine dayanabilir ve bu problem olabilir.**

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

	subno	timedrs	attdrug	atthouse	income	mstatus	race	var
250	335	12	8	23	4	1	1	
251	336	5	7	26	1	2	1	
252	337	2	9	22	1	2	1	
253	338	2	6	.	5	1	2	
254	339	4	7	25	6	2	2	
255	340	2	7	19	5	2	2	
256	341	5	9	22	3	2	1	
257	342	0	9	18	5	2	1	
258	343	3	8	25	.	2	2	
259	344	7	9	23	5	2	1	



Çiftler Bazında Veri Silme (Pairwise Deletion)

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

	subno	timedrs	attdrug	atthouse	income	mstatus	race	var
250	335	12	8	23	4	1	1	
251	336	5	7	26	1	2	1	
252	337	2	9	22	1	2	1	
253	338	2	6	.	5	1	2	
254	339	4	7	25	6	2	2	
255	340	2	7	19	5	2	2	
256	341	5	9	22	3	2	1	
257	342	0	9	18	5	2	1	
258	343	3	8	25	.	2	2	
259	344	7	9	23	5	2	1	

- Çiftler bazında veri silme yöntemine göre 10 bireye ait **korelasyon matrisi hesaplanırken, atthouse ve income** değişkenleri için korelasyon kayıp değerler içermeyen 335,336,337,339,340,341,342,344 subno.lubireylere ait gözlemlerden hesaplanır. Ancak **attdrug ve atthouse** değişkenleri için korelasyon 335,336,337,339,340,341,342,**343** ve 344 subno.lubireylere ait gözlemlerden hesaplanacaktır.



Kayıp Verinin Tahmin Edilmesi

Kayıp verinin kestirilmesinde kullanılan yöntemler aşağıdaki gibidir:

1. **Ortalama değerlerin yerleştirilmesi:** Bu yöntemde mevcut veriden ilgili değişkenlere ait ortalamalar hesaplanır ve analize başlamadan önce kayıp değer yerine ortalama değer koyulur (Atar, 2016).
 - Örneğin, atthouse değişkeni için ortalama değer 23,54, kayıp veri için 338 subno.lu bireyin tutum değeri 23,54 olarak kabul edilir.
2. **Regresyon kullanılması:** Bu yöntemde kayıp veriye sahip değişken bağımlı değişken, diğer değişkenlerse bağımsız değişkenler olarak ele alınarak bir regresyon eşitliği oluşturulur. Daha sonra eşitlik kayıp değerleri tahmin etmek için kullanılır (Atar, 2016).
3. **Beklenti maksimizasyonu (EM):** Beklenti – maksimizasyon algoritmasında, kayıp değerlerin gözlenen değerler üzerinden kurulan bir regresyon denklemiyle tahmin edildiği beklenti adımı ve bu sayede ulaşılan eksiksiz veri seti üzerinden regresyon denkleminin yeniden kurulduğu maksimizasyon adımı tekrarlı (iterative) bir süreçte birbirini izlemektedir (Enders, 2010 akt. Akbaş, 2014).
4. **Çoklu değer atama (MI):** Veri setindeki kayıpların yerine iki ya da daha fazla değer ataması söz konusudur. Bu teknikte, verinin kayıp olmasından kaynaklanan belirsizliğe bağlı olarak tek ve kesin bir değer ataması yapılmamaktadır. Çoklu değer atama tekniği, $m > 1$ sayıda eksiksiz veri seti elde edilecek şekilde atama işleminin yapılması, elde edilen m farklı veri setinin standart tekniklerle analiz edilmesi ve elde edilen sonuçların birleştirilmesinden oluşan üç aşamalı bir süreçtir (Schafer ve Graham, 2002 akt. Akbaş, 2014).



Kayıp Veri için Geliştirilen Çözümler (screen verisi örneği)

- “screen.sav” veri setinde 1 kayıp değere sahip **atthouse** değişkeni için liste bazında veri silme yöntemi kullanılarak kayıp değere sahip birey veri setinden çıkarılabilir veya kayıp verinin tahmin edilmesi yöntemlerinden ortalamanın yerleştirilmesi kullanılarak kayıp değer yerine değişkenin ortalama değeri (23,54 değeri) yerleştirilebilir.
- Veri setinde income değişkeni 26 kayıp değere sahiptir ve bu sayı örneklemin %5’inden fazladır. Eğer bu değişken araştırma açısından öneme sahip değilse, veri setinden çıkarılabilir, aksi halde kayıp verinin tahmin edilmesi yöntemlerinden biri kullanılabilir.

