

Sosyolojide Veri Analizi 1

Teori ve Uygulama

Başlarken



Prof.Dr. Ünal H. ÖZDEN

İstatistik Olmadan Olmaz

- Araştırmadaki sayıların yararlı bilgiler olup olmadığını belirlemek
- Belirsizlik altında karar vermek
- Sınıflama ve kümeleme
- Nedensellik iddialarını doğrulamak (Değişimin ve farklılaşmanın sebebi)
- Büyük miktarda verilerin ortaya koyduğu kalıpları görmek
- Tahmin

İstatistik nedir?

- Üç çeşit yalan vardır:
 - Yalan
 - Kuyruklu yalan
 - İstatistik

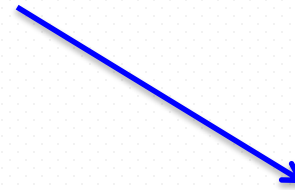
Benjamin Disraeli

Doğru Veri Toplama İstatistiksel Analiz İçin Çok Kritiktir



İstatistiğin Konusu Olan Olaylar

İstatistik olaylarla ilgilenir.
Olayları ikiye ayırmak mümkündür.



Toplu olaylar

Bir çok faktör tarafından etkilenen olaylardır. İstatistiğin konusu kapsamındadır. (Enflasyon, başarıya etki eden faktörler...)

Tekil olaylar

Tek bir faktör tarafından etkilenen olaylardır. İstatistiğin konusu kapsamında değildir. Belirli şartlar birleştğinde daima aynı sonucu verir. (Kimyasal olaylar...)

İstatistiğin İki Dalı

İstatistik

Karar vermede verileri yararlı bilgilere dönüştürmeye yardımcı yöntemler bütünü. Ham verilerden bilgi üretme.



Tanımsal İstatistik

Verileri toplama, düzenleme, görselleştirme, analiz etme ve yorumlamadan oluşan süreci kapsar.



Çıkarımsal İstatistik

Küçük bir gruptan (örnekten) toplanan verileri kullanarak daha büyük bir grup (anakitle) hakkında sonuçlara varmak için kullanılır.

Tanımsal İstatistik

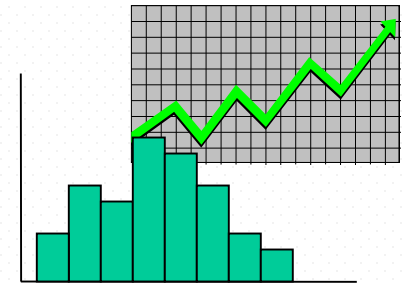
- Verilerin toplanması

- Ör. Anket



- Verilerin düzenlenmesi ve sunulması

- Ör. Tablolar ve grafikler



- Karakteristik değerlerin hesaplanması

- Ör. Örnek ortalaması =
$$\frac{\sum X_i}{n}$$

Çıkarımsal İstatistik

- Tahmin
 - Ör. Anakitle ağırlık ortalamasının örnek ortalamasından yararlanarak tahmin edilmesi.
- Hipotez testleri
 - Ör. Anakitle ortalama ağırlığının 75 kg olduğu iddasının testi.



Örnekten hesaplanan sonuçlara göre anakitle hakkında karar verilir.

İstatistiğin Kullanım Alanları

- İşletmelerde; insan kaynakları, finansal analiz, Pazar arařtırmaları, tedarik zinciri gibi...
- Psikoloji
- Sosyoloji
- Ekonomi
- Tıp
- Biyoloji
- Fizik
- Mühendislik
- vs.

Software (Bilg. Paket Programı) ve İstatistik

- Software, istatistiksel yöntemleri uygularken hesaplamalarda size yardımcı olacak programlardır.
- Microsoft Excel ile istatistiksel veri analizi yapabilirsiniz.
- Bir çok istatistik paket programı vardır. En bilinenleri;
 - SPSS
 - Minitab
 - R
 - Eviews
 - SAS

Verilerin Tanımlanması ve Toplanması



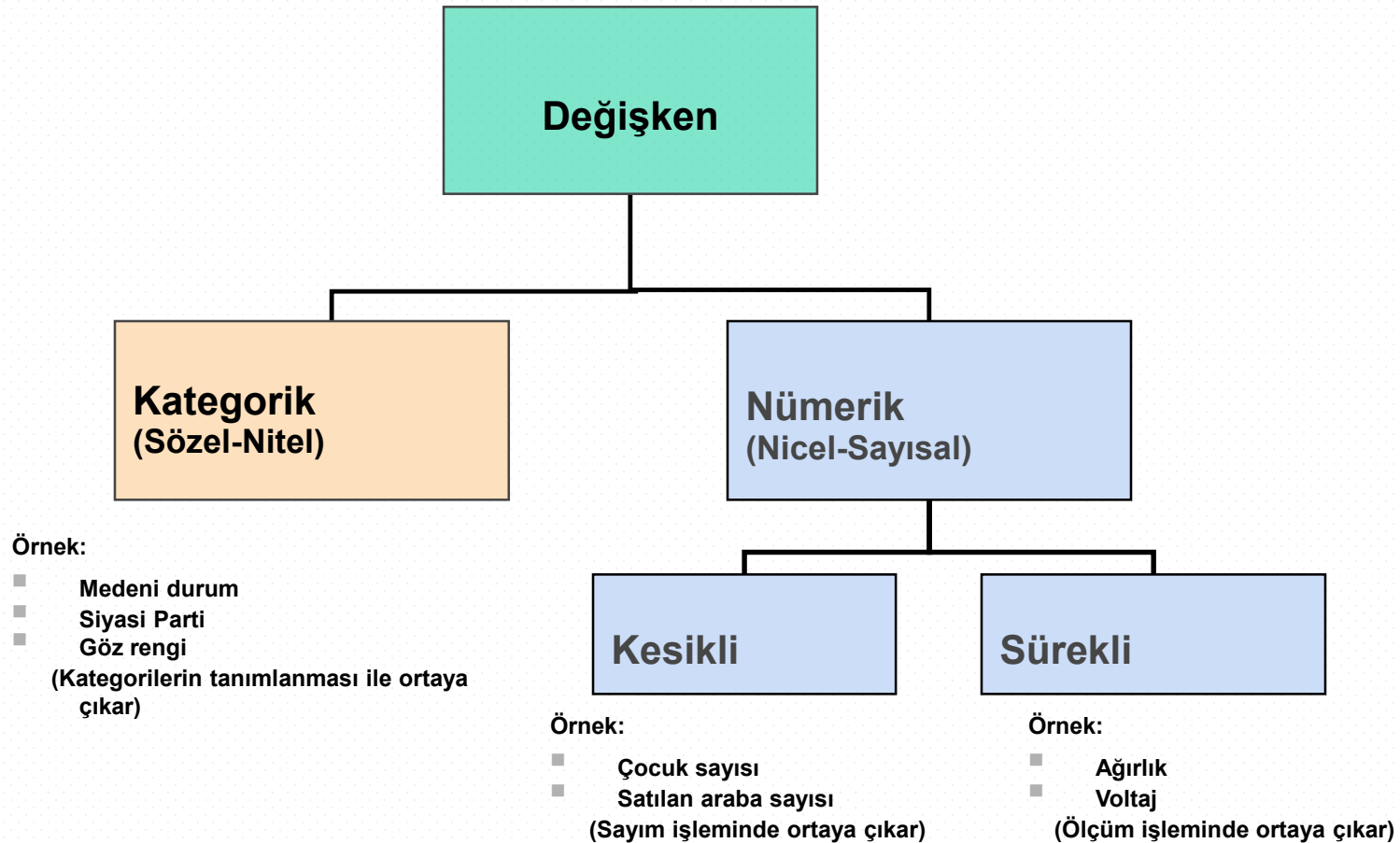
Temel Kavramlar

Değişken: Herhangi bir ögenin veya birimin herhangi bir özelliği

Veri (Data): Herhangi bir değişkenin birimlerine ilişkin değerler kümesi

İstatistik: Karar vermede verilerden yararlanarak, yararlı bilgiler üretmeye yardımcı yöntemler bütünü veya ham verilerden bilgi üretme süreci

Değişken Türleri



Değişken

Değişken : Gözlemden gözleme farklı değerler alabilen objelere, niteliklere ya da durumlara değişken denir. İstatistik birimlerinin sahip oldukları özellikler birer değişken olarak görülebilir.

- **Sürekli değişken** : Matematiksel olarak herhangi iki değeri arasında daima bir başka değeri bulunabilen değişken. (Örneğin: Uzunluk, ağırlık, yaş)
- **Süreksiz değişken** : Ölçüm birimleri daha küçük bölümlere bölünemediğinden ölçek üzerinde ayrı ayrı noktalar halinde yer alan değişken. (Örneğin Pekiyi 5, İyi 4 Orta 3 gibi)
- **Bağımsız değişken** : Başka bir değişkene bağlı olmadan değerler alabilen değişken.
- **Bağımlı değişken** : Başka bir değişkene bağlı olarak değerler olabilen değişken.

Ölçüm Düzeyleri ve Ölçekler

Oransal

Bu ölçme düzeyi, aralıklı ölçme düzeyinin bütün özelliklerine sahiptir. Aralıklı ölçme düzeyinden farklı olarak; oransal ölçekte sıfır gerçek yokluğu ifade eder ve iki sayı arasında oransal ilişki vardır.

Örnek:

Boy, yaş, haftalık tüketilen gıda miktarı...



Aralıklı

Bütün sıralı veri türlerini kapsar, değerler arasındaki uzaklık sabit büyüklüktedir, sayılar arasında oransal ilişki yoktur ve sıfırın gerçek bir yokluğu ifade etmez.

Hava sıcaklığı, standartlaştırılmış sınav skoru...



Sıralı

Veriler farklı sıralı kategorilere göre sınıflandırılır. Nominal ölçme düzeyi ile sıralı ölçme düzeyi arasındaki temel farklılık, sıralı ölçme düzeyi sınıfları arasında '... den daha iyi' ilişkisinin olmasıdır.

Hizmet kalite puanı, ürün memnuniyeti, akademik ünvan, S & P derecelendirmesi, Öğrenci bağıl notu (harf olarak)...



İsimsel

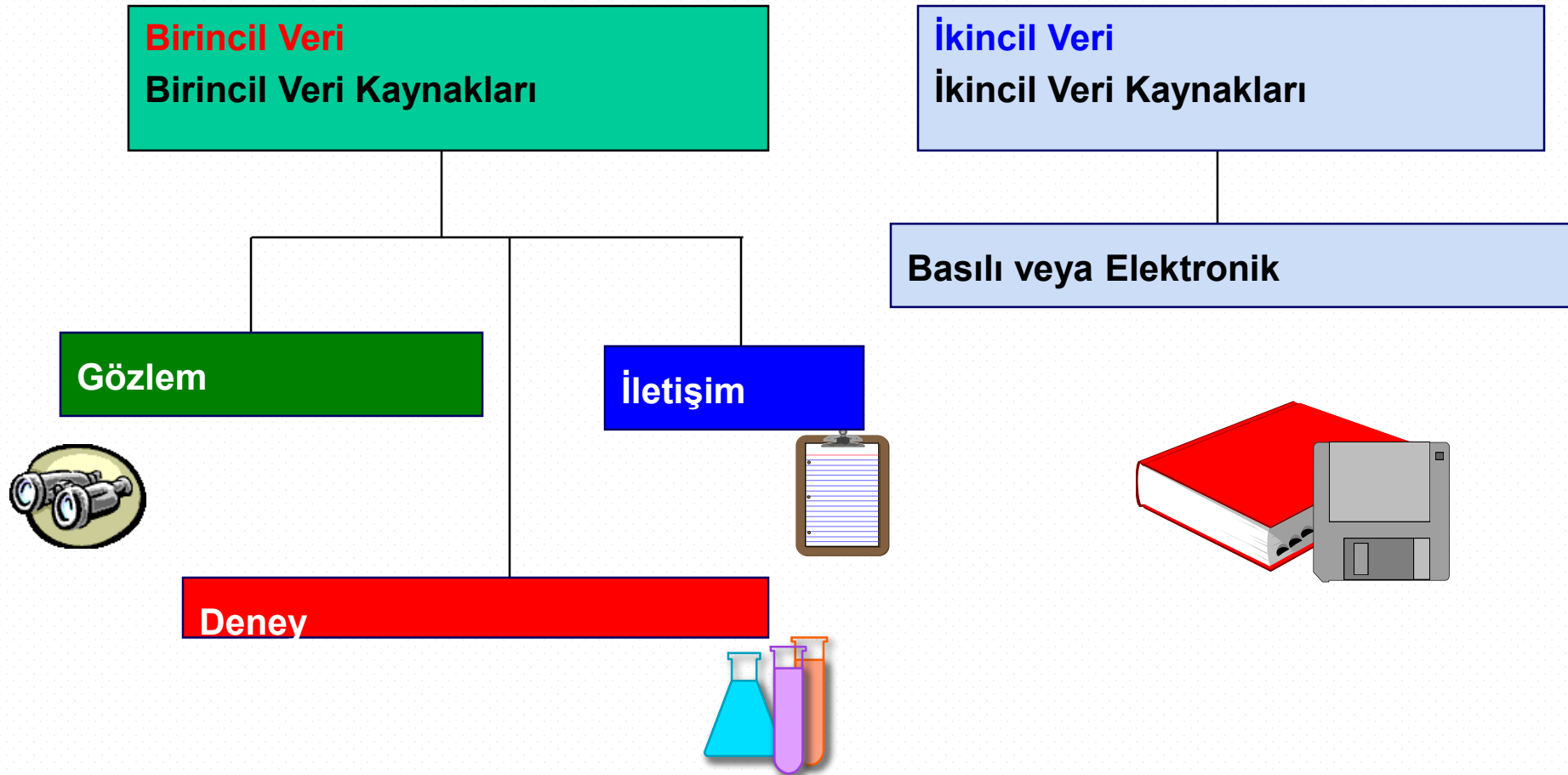
Nominal ölçekte veriler için hiçbir sıralama yoktur. Veriler farklı kategorilere göre sınıflandırılır.

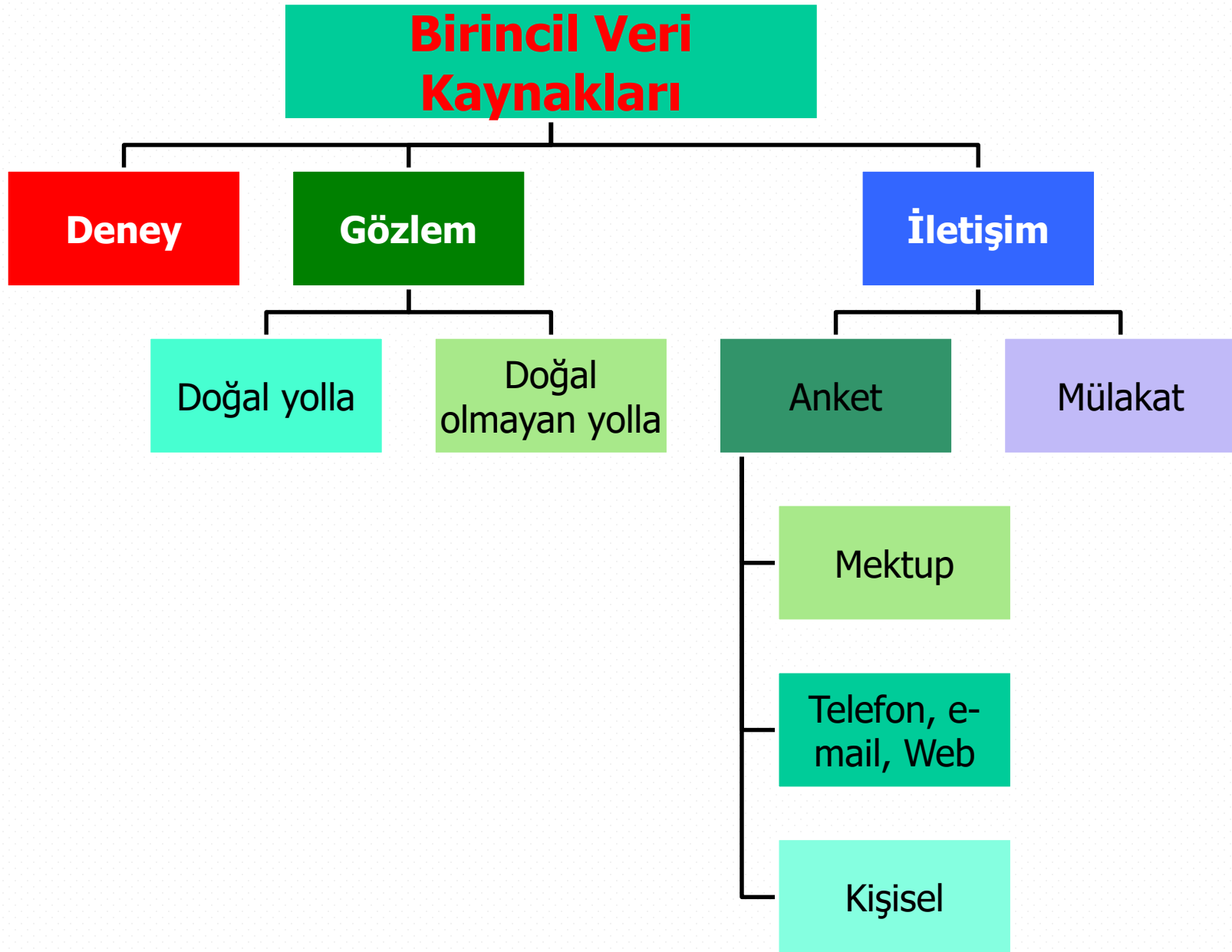
Medeni durum, araba markası, facebook profili sahipliği, yatırım türü...

Veri Kaynakları

- **Birincil Veriler:** Veri analizi yapacak kişi/kişiler tarafından toplanmış veriler
 - Siyasetle ilgili anketlerden elde edilen veriler
 - Deneylerden elde edilen veriler
 - Gözlemlerden elde edilen veriler
- **İkincil Veriler:** Veri analizi yapacak kişi(ler)den farklı kişiler tarafından toplanmış veriler
 - Nüfus sayımı verileri
 - İnternet veya basılı yayınlardaki yer alan veriler

Veri Toplama





Anakitle-Örnekleme

İstatistik

Teori ve Uygulama

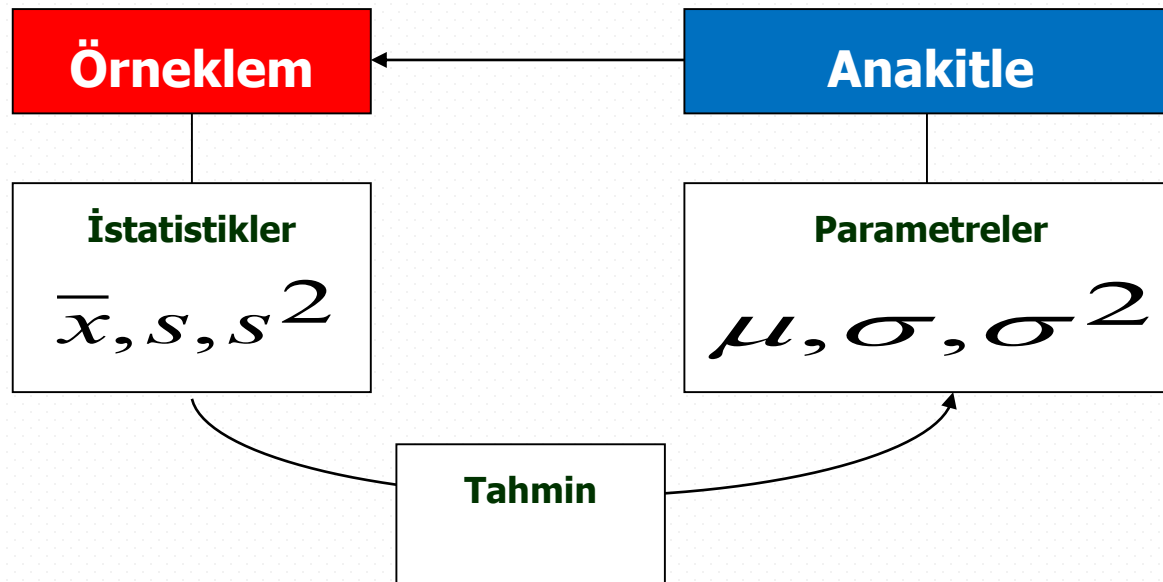


Prof. Dr. Ünal H. ÖZDEN

Tanımlar

- **Anakitle:** Araştırmaya konu olan birimlerin oluşturduğu kümeye denir.
- **Örnekleme:** Belli kurallara göre, belli bir anakitleden seçilmiş ve seçildiği anakitleyi temsil yeterliliği olan alt kümedir. (temsil gücü ve yeterlilik)
- **Parametre:** Anakitleyi tanımlamak için hesaplanan karakteristik değerler
- **İstatistik:** Örnekten hesaplanan karakteristik değerler
- **Tamsayım:** Anakitleyi oluşturan birimlerin tamamının sayılması
- **Örnekleme:** Bir araştırmanın konusunu oluşturan anakitlenin bütün özelliklerini yansıtan bir parçasının seçilmesi ve seçilen bu örneklemden yararlanarak hesaplanan karakteristik değerlerden (istatistik) yararlanarak anakitle karakteristik değerlerinin (parametre) tahmin edilmesi
- **Birim:** Anakitleyi oluşturan en küçük parça. Birim tekil olmak zorunda değildir.
- **Karakteristik Değer:** Herhangi bir verinin veya değişkenin özelliklerini tanımlamak için hesaplanan değerlerdir (aritmetik ortalama, mod, medyan, standart sapma vb...)

Parametre ve İstatistik



Parametre Ve İstatistik Simgeleri

DEĞER	PARAMETRE	İSTATİSTİK
Birim Sayısı	N	n
Aritmetik Ortalama	μ	\bar{x}
Standart Sapma	σ	S
Varyans	σ^2	S^2
Standart Hata	σ_{μ}	$s_{\bar{x}}$
Oran	π	p

Anakitle - Örneklem

Anakitle



Araştırmaya konu olan birimlerin oluşturduğu kümedir

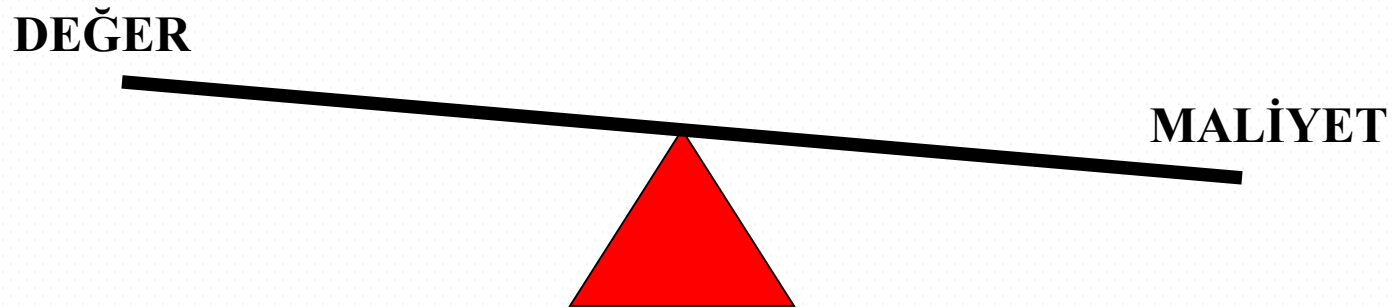
Örneklem



Bir anakitleden seçilmiş ve seçildiği anakitleyi temsil yeterliliği olan alt kümedir.

Niçin TAMSAYIM?

- Kesin sonuç



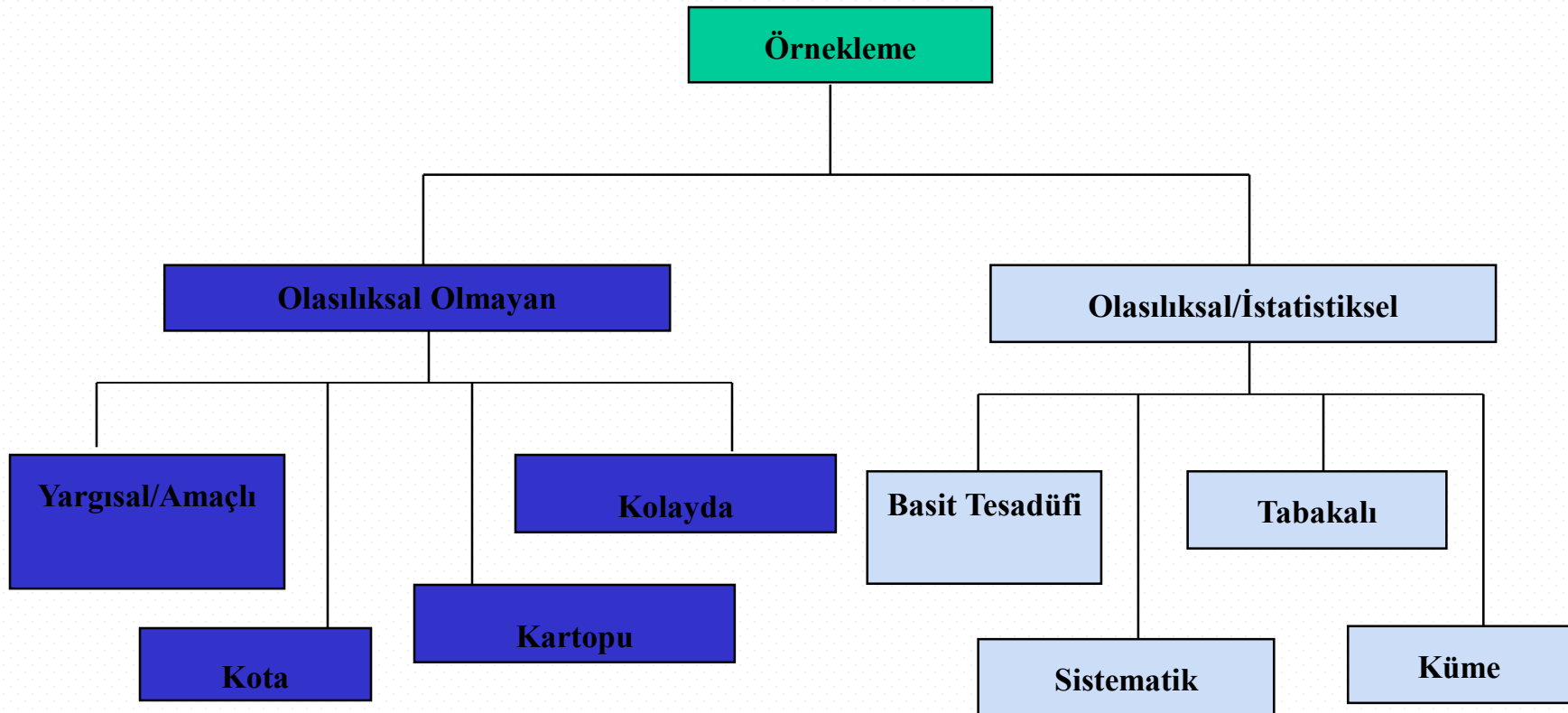
Niçin Örnekleme?

- Anakütleye ulaşılammaması
- Zaman
- Maliyet
- Kolaylık

Veri Kaynakları ve Veriler ile İlgili Bilinmesi Gerekenler

- Veri kaynağı yapılandırılmış veya yapılandırılmamış mı?
 - Yapılandırılmamış/Düzenlenmiş
 - Yapılandırılmış/Düzenlenmemiş
- Elektronik veriler hangi formatta yer almaktadır?
- Veriler nasıl kodlanmış?
 - Veriler kodlanmış mı?
 - Kodlanmış verilerin tekrar orijinal hale dönüştürülmesi gerekir mi?
- Veri temizlemesi yapılmış mı?
 - Veri yanlışlıklar, Kayıp veriler, Uç değerler...
 - Tanımlanamayan veriler vs.

Örnekleme Yöntemleri



Olasılıksal Olmayan Örneklemeye Yöntemleri

Olasılıksal olmayan örnekleme, birimlerin seçiminde keyfi seçim yönteminin uygulandığı örnekleme yöntemleridir.

Kolayda (Gelişigüzel) Örneklemeye: Kolayca ulaşılabilir birimleri seçmek suretiyle bir örnek oluşturulmaya çalışılır. Örneklemede birimlerinin seçimi görüşmeci tarafından doğru zamanda doğru yerde bulunan birimler, gönüllü katılımcılar arasından yapılır. Herhangi bir fakülteye gidip saptanacak sayıda rastlanan öğrenciyi örnekleme alma

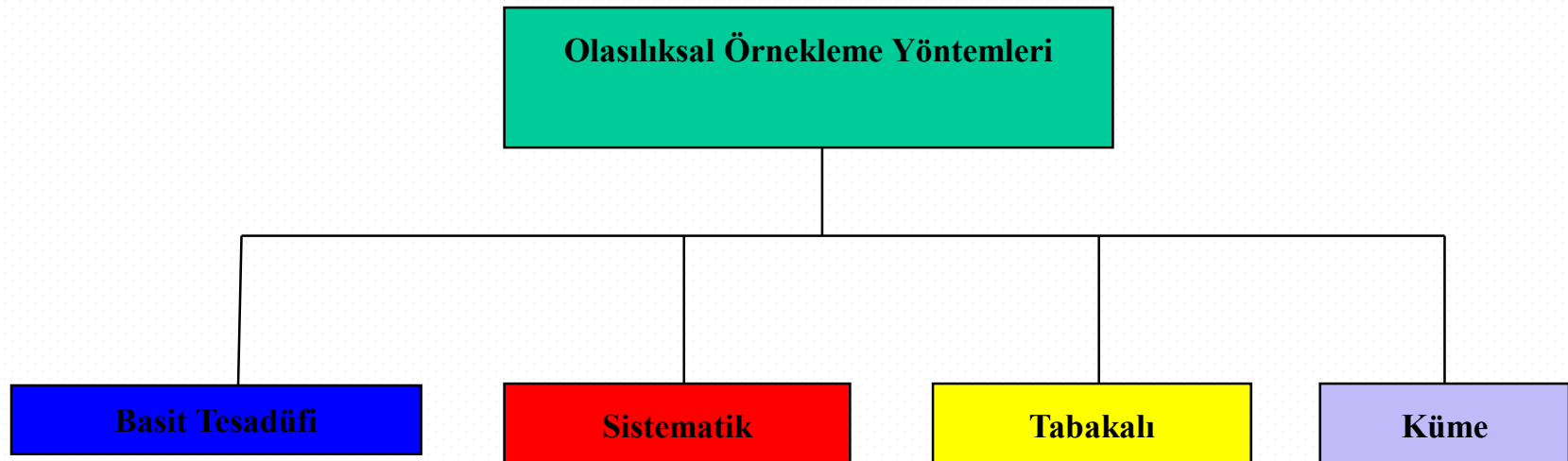
Yargısal Örneklemeye: Birimlerin seçiminin araştırmacının amacına, arzu, düşünce ve deneyimlerine dayanarak yapılmasıdır. Meslek hastalıklarıyla ilgili yapılacak bir araştırmada örneklemin, meslek hastalıklarının tüm anakitle içinden değil, özellikle belli bir hizmet süresini aşmış ya da belli bir yaş sınırının üstündekiler arasından seçmesi gibi.

Kota Örneklemesi: Bu yöntemde tabakalı örnekleme yönteminde olduğu gibi anakitle alt tabakalara ayrılır. Her alt tabakanın temsili için kota konulur. Bu kota belirlenen tabakanın anakütleyle oranına göre belirlenir. Kota örneklemede örneğe girecek elemanlar tesadüfen değil araştırmacını kendi isteğine göre belirlenir.

Kartopu Örneklemesi: Anakitleye ulaşmak mümkün olmadığında, ulaşabilen ilk birim belirlenir. Bu birimden elde edilen bilgilerle diğer birimlere ve bu şekilde zincirleme olarak anakitleyi temsil eden örneğe ulaşılmaya çalışılır.

Olasılıksal Örneklemeye Yöntemleri

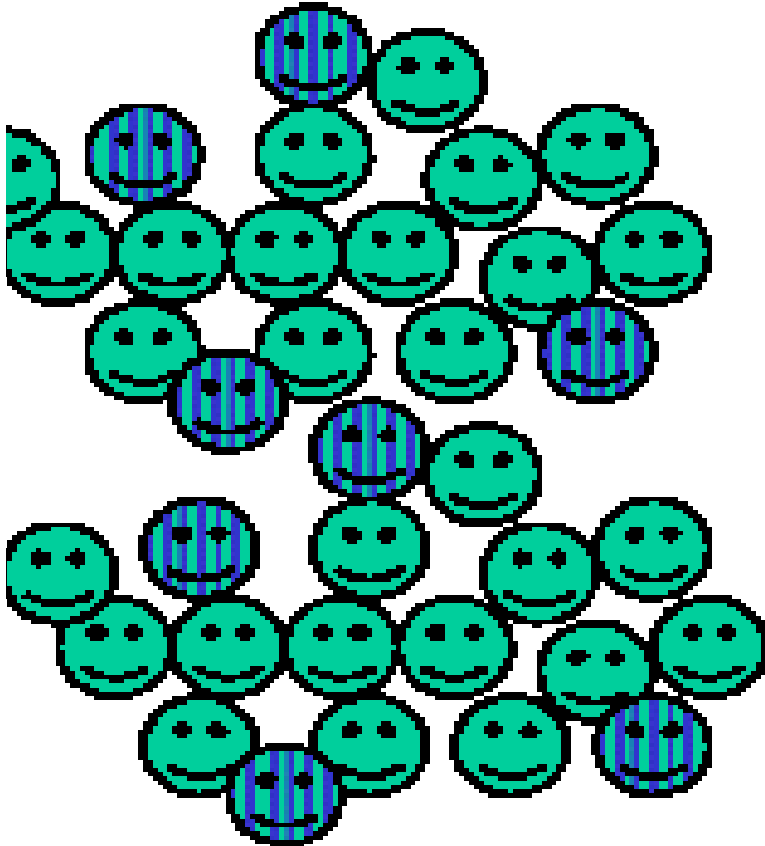
- Olasılık örnekleme yöntemlerinde, birimler bilinen olasılıklara bağlı olarak seçilir.



Basit Tesadüfi Örnekleme

- Anakitlede yer alan her bir birimin örneklem kümesine girme şansı var ve bu şanslar eşit
- Seçimler iadeli olarak yapılabilir.
- Birimler tesadüfi sayılar tablosu veya bilgisayar yardımı ile çekilebilir.
- Anakütle incelenen konu açısından HOMOJEN yapıda olduğunda iyi sonuç verir
- Anakitleyi oluşturan birimlere birer numara verilir ve rasgele bu numaralar çekilir.

Basit Tesadüfi Örnekleme



Rasgele Sayılar Tablosu

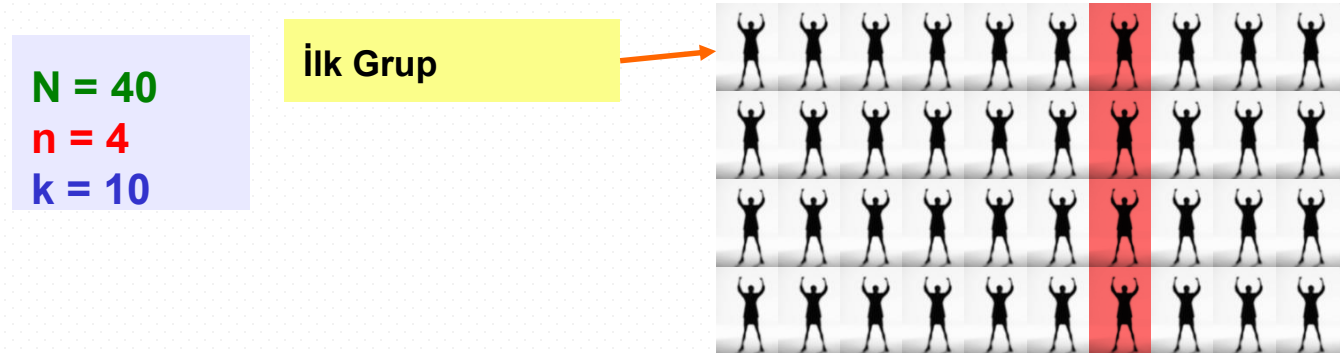
49280 88924 35779 00283 81163 07275
 11100 02340 12860 74697 96644 89439
 09893 23997 20048 49420 88872 08401

Örneklemeye seçilen ilk 5 birim

Item # 492
 Item # 808
 Item # 892 -- iptal böyle bir gözlem yok
 Item # 435
 Item # 779
 Item # 002

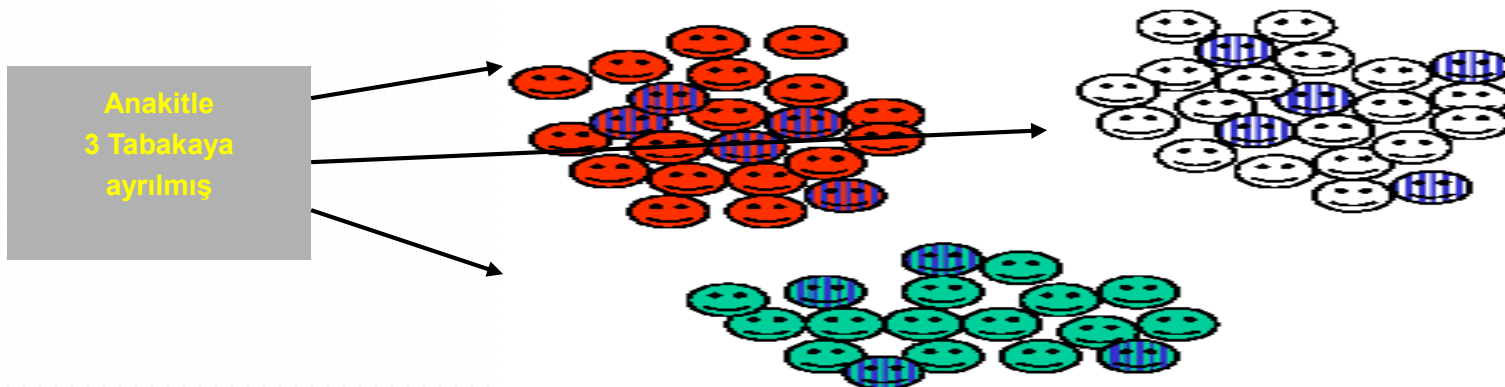
Sistematik Örnekleme

- Anakitle birimlerini kurala göre numaralandırılır (1...N) ve örneklem büyüklüğünü (n) belirlenir
- Örnekleme oranı k'yı ($k=N/n$) hesaplanır ve anakitle sıra numarasına göre her biri k birimden oluşan n gruba ayırılır.
- 1 ile k arasında rasgele bir rakam (s) seçilir.
- Her gruptaki s'inci sıradaki birim örneklem kümesine dahil edilir.



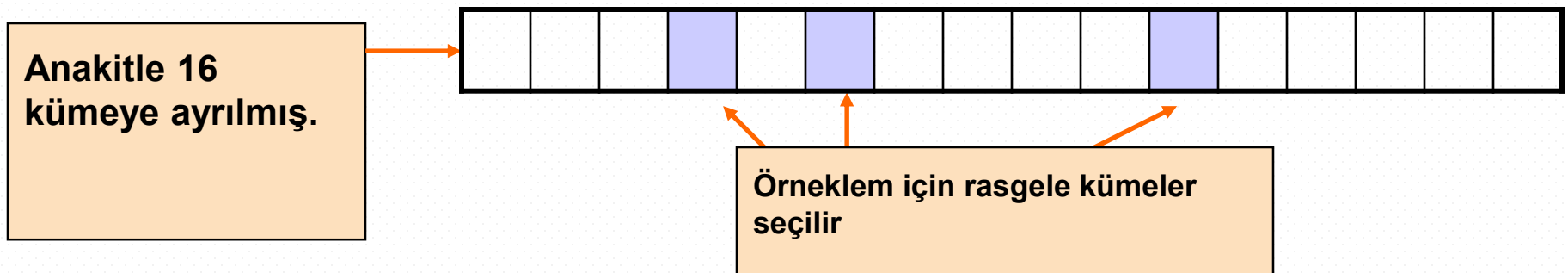
Tabakalı Örnekleme

- Homojen olmayan anakitle birimleri, karakteristik özelliklerine göre tabaka denilen homojen alt gruplara ayrıştırılır
- Her tabakadan anakitle içindeki oranına bağlı olarak basit tesadüfi örnekleme yöntemi ile birimler seçilir
- Bu tabakalardan seçilen birimler birleştirilerek örneklem oluşturulur
- Çok yaygın kullanılan bu teknikte tabakalar kendi içinde homojen birbirleri arasında heterojendir.



Küme Örnekleme

- Anakitle, anakitleyi temsil eden birden fazla “küme”ye bölünür
- Kümeler arasından basit tesadüfi örnekleme ile rasgele seçim yapılır
- Seçilen küme içindeki tüm birimler örnekleme içinde yer alır veya seçilen kümelerdeki birimler başka bir örnekleme tekniğinde kullanılabilir
- Kümeler kendi içinde heterojen, kümeler arasında homojendir.



Örnekleme Yöntemlerinin Karşılaştırılması

- Basit tesadüfi örnekleme ve sistematik örnekleme
 - Kullanımı kolay
 - Anakitle özelliklerini için temsil sorunu yaşanabilir.
- Tabakalı örnekleme
 - Anakitleyi oluşturan ve farklı karakteristiklere sahip tüm birimlerin temsil edilmesini sağlar.
- Küme örnekleme
 - Daha düşük maliyetlidir.
 - Daha az etkindir. Etkinliğin ve temsiliyetin diğerleri kadar olabilmesi için daha yüksek örnekleme büyüklüğüne ihtiyaç vardır.

Hata Türleri

- Kapsam hatası



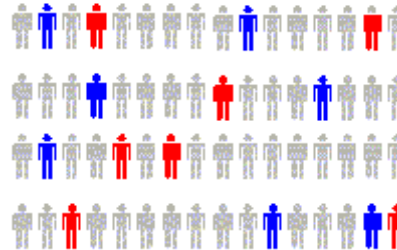
Listeden dışlananlar

- Tepki hatası



Cevaplamayanları takip

- Örnekleme hatası



Tesadüflik var

- Ölçme hatası



Kötü veya yönlendirici soru

Hata Türleri

- Kapsama hatası veya seçim yanlılığı
 - Bazı gruplar çerçeve dışında kalması nedeniyle seçilme şansları yoktur
- Cevaplamama hatası
 - Cevap vermeyen insanlar cevap verenlerden farklı kanaate sahip olabilir
- Örneklem hatası
 - Her zaman var olur ve örneklemden örnekleme değişkenlik gösterir.
- Ölçme hatası
 - Yanlış ve yönlendirici hazırlanmış sorular nedeniyle yanlı cevaplar olacaktır.

Örneklem Hataları

- Örneklem yöntemlerine göre yapılan tahminlerde iki çeşit hata vardır. Tesadüfi hatalar, örnek sayısı artırılarak giderilirken, sistematik hatalar örnekleme sürecinde ortaya çıkar ve sonradan giderilmesi zordur. Bu hatalar:
 1. Örneklem yönteminin yanlış seçilmesi
 2. Populasyonun yanlış tanımlanması
 3. Örnek çerçevesinin yanlış belirlenmesi
 4. Örnek birimlerinin doğru alınmayışından
 5. Örnek büyüklüğünün yanlış belirlenmesinden kaynaklanır.

Örnekleme Süreci



Örneklem Büyüklüğünün Saptanması

%95 güven aralığında %3, %5, %10 örnekleme hataları için karşılık gelen örneklem büyüklükleri yanda verilmiştir.



Hedef Kitle Büyüklüğü (N)	$\alpha = 0.05$ için örneklem büyüklükleri					
	\pm %3 örnekleme hatası (d)		\pm %5 örnekleme hatası (d)		\pm %10 örnekleme hatası (d)	
	p=0.5 q=0.5	p=0.8 q=0.2	p=0.5 q=0.5	p=0.8 q=0.2	p=0.5 q=0.5	p=0.8 q=0.2
100	92	87	80	71	49	38
250	203	183	152	124	70	49
500	341	289	217	165	81	55
750	441	358	254	185	85	57
1.000	516	406	278	198	88	58
2.500	748	537	333	224	93	60
5.000	880	601	357	234	94	61
10.000	964	639	370	240	95	61
25.000	1023	665	378	244	96	61
50.000	1045	674	381	245	96	61
100.000	1056	678	383	245	96	61
1.000.000	1066	682	384	246	96	61
100.000.000	1067	683	384	246	96	61

İstatistik

Teori ve Uygulama



Prof. Dr. Ünal H. ÖZDEN

Araştırma Sürecinin Aşamaları

Arařtırma Süreci

- Arařtırma Konusunun Belirlenmesi
- Problemin Ortaya Konması
- Konuyla İliřkili Kaynakların Taranması
- Hipotezlerin Yazılması
- Arařtırma Yönteminin Belirlenmesi
- Süre ve Olanakların Belirlenmesi
- Arařtırmanın Sonuçlandırılması

Araştırmanın Raporlaştırılması

- Araştırma planlanan şekilde gerçekleştirildikten sonra, araştırmanın verilerinin analizi sonucunda elde edilen bulgular yazılır ve bu bulguların yorumları yapılır.
- Bilimsel araştırma sürecinin son aşamasında ise araştırma raporu hazırlanır. Sosyal bilim araştırmaları genellikle dört ana bölümden ve çeşitli alt bölümlerden oluşmaktadır. Son yıllarda en yaygın kullanılan raporlaştırma biçimi şöyledir:
- I. GİRİŞ
Problem
- Kaynak Taraması Önem Hipotezler
- II. YÖNTEM
Evren ve Örneklem
Araştırma Modeli
- Verilerin Toplanması ve Analizi
- III. BULGULAR
- IV. SONUÇ (TARTIŞMA) Bulguların Yorumu Sınırlılıklar
Öneriler

Verilerin Düzenlenmesi

- Veriler hangi yöntemle toplanırsa toplan-sın, elde edilen veriler genellikle istatistik-sel analize hazır değildir (bu veriler ham veri olarak adlandırılır)
- Bu verilerin analize uygun hale getirilmesi için düzenlenmeleri gerekir

Basit Seriler

■ Tanım

Elde edilecek ham verilerin küçükten büyüğe doğru sıralanması ile elde edilen serilere **basit seri** denir

- Basit bir seri birden fazla birimden oluşur ve birim sayısı n ile gösterilir
- Serinin i 'ninci elemanı X_i değişkeni ile gösterilir

Basit Seriler

- Basit serinin toplam değeri X_i 'lerin toplamına eşit olacağından,

$$\text{Serinin toplam değeri} = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$$

ÖRNEK:

Bir işletmede 25 işçiye verilecek çocuk paraları ile ilgili bir araştırmaya yapılmaktadır. İşçilerin çocuk sayıları aşağıda verilmiştir.

1,3,2,2,3,1,4,5,3,6,0,5,2,3,2,4,8,0,1,2,3,3,1,0,4

Basit Seriler

- Verilen değerleri basit seri şeklinde düzenleyelim:

0,0,0,1,1,1,1,2,2,2,2,2,3,3,3,3,3,3,4,4,4,5,5,6,8

Frekans Serileri (Tasnif Edilmiş Seriler)

- İncelenecek birim sayısı arttıkça basit seriler çok uzun olacaktır
- Bu durumda çalışma kolaylığı sağladığı için frekans serileri kullanılır
- Frekans serilerinin değişkenin çok sayıda farklı değer almadığı durumlarda kullanılması daha uygun olur

Frekans Serileri (Tasnif Edilmiş Seriler)

- Düzenlenen bir frekans serisi iki sütundan oluşur:
 - **Birinci Sütun:** Değişkenin aldığı farklı değerler yer alır
 - **İkinci Sütun:** f_i ile gösterilir ve değişkenin aldığı değerlerin tekrar sayısı gösterilir

X_i	f_i
x_1	f_1
x_2	f_2
x_3	f_3
.	.
.	.
.	.
x_k	f_k

Değişkenin k sayıda farklı değer aldığı bir frekans serisi

Frekans Serileri (Tasnif Edilmiş Seriler)

- Frekans serisinin toplam değerini aşağıdaki gibi hesaplayabiliriz:

$$f_1X_1 + f_2X_2 + \dots + f_kX_k = \sum_{i=1}^k f_iX_i$$

ÖRNEK: Daha önce basit seri olarak düzenlenen seriyi frekans serisi olarak düzenleyiniz

Frekans Serileri (Tasnif Edilmiş Seriler)

- Örnekteki basit serinin frekans serisi olarak düzenlenmiş hali:

x_i	f_i
0	3
1	4
2	5
3	6
4	3
5	2
6	1
8	1
	<hr/>
	25

Gruplandırılmış Seriler (Sınıflı Seriler)

- İncelenecek birim sayısının çok fazla olması durumunda frekans serileri de uzun sayı dizilerine dönüşür
- Bu durumda gruplandırılmış seriler düzenlenir
- Gruplandırılmış seriler iki sütundan oluşur:
 - Birinci Sütun: Sınıflar sütunudur
 - İkinci Sütun: frekans sütunudur

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış serilerde en önemli nokta sınıf sayısının kaç tane veya sınıf aralığının ne olacağına belirlenmesidir.
- Sınıf aralığının ne olması gerektiği konusunda bazı yazarlar çeşitli formüller önermektedir.
- Ancak bu formüller sadece birer öneridir, yani kesin değildir.

Gruplandırılmış Seriler (Sınıflı Seriler)

- Sınıf sayısının az olması serinin verdiği bilgilerin kaybına yol açacağından sınıf sayısının dörtten az olmaması
- Diğer yandan, çok fazla sınıf sayısının ise işlem zorluğu ve serinin yorumlanmasını zorlaştıracığı için sekizden fazla olmaması tavsiye edilir
- Önerilen kurallardan biri, sınıf sayısının birim serideki sayısının kare kökü olarak seçilmesidir

Gruplandırılmış Seriler (Sınıflı Seriler)

- Bir diğer kural ise **Sturges** kuralıdır:

$$k = 1 + 3,322 \log(\sum f_i)$$

k: Minimum Sınıf Sayısı

- Sınıf Genişliğinin hesaplanması

$$S = \frac{X_{\max} - X_{\min}}{k} \quad (\text{Eşit sınıf aralığı})$$

S: Sınıf Genişliği

Gruplandırılmış Seriler (Sınıflı Seriler)

- Sınıf aralığı belirlendikten sonra
 1. İlk sınıfın alt sınır değeri sınıflar sütununa yazılır
 2. Alt sınır değere sınıf aralığı eklenerek üst sınır değeri elde edilir
 3. Her sınıfın üst sınıf değeri bir sonraki sınıfın alt sınır değerini oluşturur

Sınıflar değişkeninin tüm değerlerini kapsayana dek işleme devam edilir

Gruplandırılmış Seriler (Sınıflı Seriler)

- **ÖRNEK:** Daha önce frekans serisi olarak düzenlenen örneği 2 eşit aralıklı sınıflı seri olarak düzenleyiniz

Frekans Serisi

X_i	f_i
0	3
1	4
2	5
3	6
4	3
5	2
6	1
8	1
	<hr/> 25

Sınıflandırılmış Seri

Sınıflar	f_i
0-2	7
2-4	11
4-6	5
6-8	1
8-10	1

⇐ Sınıflardaki frekanslar belirlenirken alt Sınıf değeri dahil üst sınıf değeri **hariç** tutulur

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış seriler ile işlem yapılırken frekansların her bir sınıf içerisinde eşit biçimde dağıldığı düşünülür
- Bu nedenle sınıf ortaları X_i olarak düşünülerek serinin toplamı frekans serilerinde olduğu gibi hesaplanır

$$\text{Serinin toplam değeri} = f_1 m_1 + f_2 m_2 + \dots + f_k m_k = \sum_{i=1}^k f_i m_i$$

$$m_i = \frac{X_{\text{alt}}^i + X_{\text{üst}}^i}{2}$$

m_i i'nci sınıf ortası m_i sınıf alt sınır ve üst sınır toplamının ikiye bölünmesi ile bulunur

Gruplandırılmış Seriler (Sınıflı Seriler)

ÖRNEK: Bir önceki örnekteki gruplandırılmış serinin toplamını hesaplayınız

$$f_1 m_1 + f_2 m_2 + \dots + f_k m_k = \sum_{i=1}^k f_i m_i$$

Sınıflar	f_i	m_i	$f_i m_i$
0-2	7	$(0+2)/2=1$	7
2-4	11	$(2+4)/2=3$	33
4-6	5	$(4+6)/2=5$	25
6-8	1	$(6+8)/2=7$	7
8-10	1	$(8+10)/2=9$	9

$$\sum_{i=1}^k f_i m_i = 81$$

Gruplandırılmış Seriler (Sınıflı Seriler)

- Gruplandırılmış serilerde sınıflardaki verilerin sınıf aralığında düzgün dağıldığının varsayılması hesaplanan ölçülerde sapmalara neden olacaktır.

- Basit seride, $\sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$
 $= 0 + 0 + 0 + \dots + 5 + 6 + 8 = 68$

- Frekans serisinde,

$$\sum_{i=1}^n f_i X_i = f_1 X_1 + f_2 X_2 + \dots + f_k X_k$$

$$= 0(3) + 1(4) + 2(5) + \dots + 6(1) + 8(1) = 68$$

Birikimli Frekanslar

- Bazı istatistiksel çalışmalarda bir frekans serisinde veya gruplandırılmış seride belirli bir değerden daha küçük veya daha büyük değer alan alan birim sayısının belirlenmesi gerekebilir
- Bu durumlarda birikimli frekanslar hesaplanır

Artan Birikimli Frekanslar

- Artan birikimli frekanslar hesaplanırken frekanslar sütununun ilk frekans değeri aynen alınır
- İkinci frekans değeri ilk frekans değeri ile toplanarak artan birikimli frekanslar sütununun ikinci değeri olarak yazılır
- Bu işleme frekanslar bitene dek devam edilir

Artan Birikimli Frekanslar

ÖRNEK: Daha önce frekans serisi olarak düzenlenen serinin artan birikimli frekansını hesaplayınız

X_i	f_i	Artan Birikimli Frekanslar
0	3	3
1	4	3+4=7
2	5	7+5=12
3	6	12+6=18
4	3	18+3=21
5	2	21+2=23
6	1	23+1=24
8	1	24+1=25
	<u>25</u>	

Seride 4 ve daha küçük birimlerin sayısı 21'dir

Azalan Birikimli Frekanslar

- Artan birikimli frekanslar için yaptığımız işlemleri bu kez aşağıdan yukarıya doğru yaparız
- Diğer bir yol ise frekanslar toplamından başlanarak çıkarma işlemi yapılmasıdır:
 - Azalan frekanslar sütununa ilk olarak frekanslar toplamı yazılır
 - İkinci frekans değeri frekanslar toplamından çıkarılarak azalan frekanslar sütununa yazılır
 - Bu işleme serinin son değerine dek devam edilir

Azalan Birikimli Frekanslar

ÖRNEK: Daha önce artan birikimli frekansı hesaplanan frekans serisinin azalan birikimli frekanslarını açıkladığı gibi iki şekilde hesaplayınız

x_i	f_i	1. Yol Azalan Birikimli Frekanslar	2. Yol Azalan Birikimli Frekanslar
0	3	25	25
1	4	22	$25-3=22$
2	5	18	$22-4=18$
3	6	13	$18-5=13$
4	3	7	$13-6=7$
5	2	$2+2=4$	$7-3=4$
6	1	$1+1=2$	$4-2=2$
8	1	1	$2-1=1$
	<hr/> 25		

28

Azalan Birikimli Frekanslar

ÖRNEK: Daha önce gruplandırılmış seri olarak düzenlediğimiz serinin azalan birikimli frekanslarını hesaplayınız

Sınıflar	f_i	Azalan Birikimli Frekanslar
0-2	7	
2-4	11	
4-6	5	
6-8	1	
8-10	1	

Oransal Frekanslar

- Bazı durumlarda birimlerin sayısı yerine birimlerin toplam birim sayısına oranının hesaplanması gerekir
- Değişkenin aldığı değerlerin veya sınıf frekanslarının toplam frekansa bölünmesi ile elde edilen frekanslara **oransal frekans** denir

Oransal Frekanslar

ÖRNEK: Daha önce ele aldığımız frekans serisi ve sınıfsal serinin oransal frekanslarını hesaplayınız

X_i	f_i	Oransal Frekanslar	Sınıflar	f_i	Oransal Frekanslar
0	3	$3/25 = 0,12$	0-2	7	
1	4	$4/25 = 0,16$	2-4	11	
2	5	$5/25 = 0,20$	4-6	5	
3	6	$6/25 = 0,24$	6-8	1	
4	3	$3/25 = 0,12$	8-10	1	
5	2	$2/25 = 0,08$			
6	1	$1/25 = 0,04$			
8	1	$1/25 = 0,04$			
		25			1,00

GRAFİKLER

Grafikler

■ Tanım

Araştırma sonucunda elde edilen ve düzenlenen verilerin daha kolay anlaşılabilmesi için gösterildiği şekillere **grafik** denir.

■ Grafikler göze hitap ettikleri için, toplanan verilerin daha açık bir şekilde görülmesine ve yorumlanmasına yardımcı olur. Buradaki en önemli nokta grafiklerin açık ve anlaşılır biçimde çizilmeleridir.

Zaman Serilerinin Grafikleri

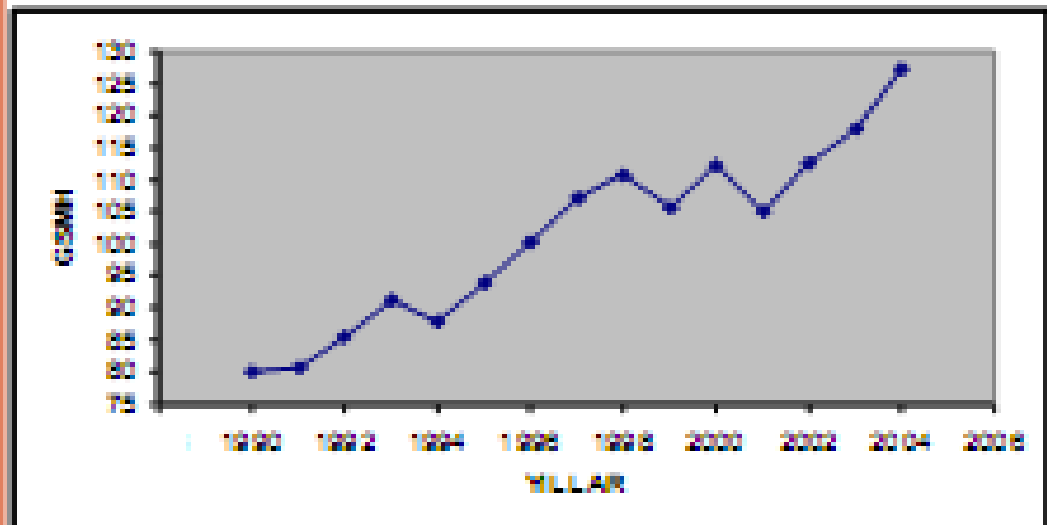
- Zaman serileri, değişkenin aldığı değerlerin zaman birimlerine göre dağılımını gösterir; bu nedenle bunların grafikleri bir koordinat sistemi üzerine çizilebilir.
- Bu koordinat sisteminde yatay ekseninde zaman birimleri ve dikey ekseninde ise değişkenin aldığı değerler yer alır.
- Eksenler, zaman birimleri ve değişkenin aldığı değerler dikkate alınarak ölçeklendirilir.

3

Zaman Serilerinin Grafikleri

ÖRNEK: Aşağıda verilen serinin grafiğini çiziniz.

Yıllar	Sabit Fiyatlar ile GSMH (1987) (YTL)
1990	80,1247
1991	80,7768
1992	85,4189
1993	91,3018
1994	87,8921
1995	94,0664
1996	100,1083
1997	107,1451
1998	110,6840
1999	105,5264
2000	112,2314
2001	104,9710
2002	112,4851
2003	117,9757
2004	127,2194



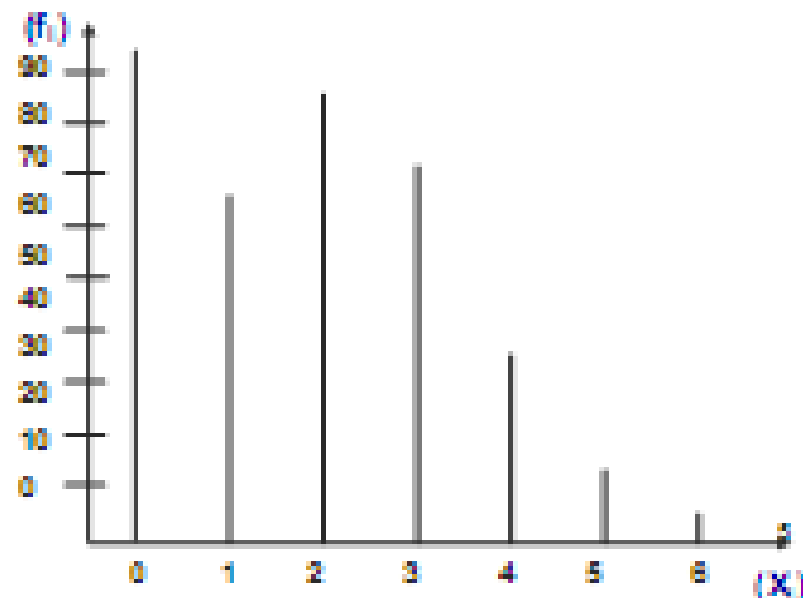
Dağılım Serilerinin Grafikleri

Tanım: Değişkenlerin aldıkları değerlerin dağılımını gösteren ve serilerin türlerine göre çizilen grafiklerdir.

- Basit serilerde birim sayısı az olduğundan ve frekanslar bulunmadığından bu serilerin grafikleri, değişkenin aldığı değerlerin büyüklüklerini gösterecek şekilde çizilir.

ÖRNEK:

Çocuk Sayısı (X)	Aile Sayısı (f)
0	94
1	65
2	87
3	71
4	34
5	12
6	5



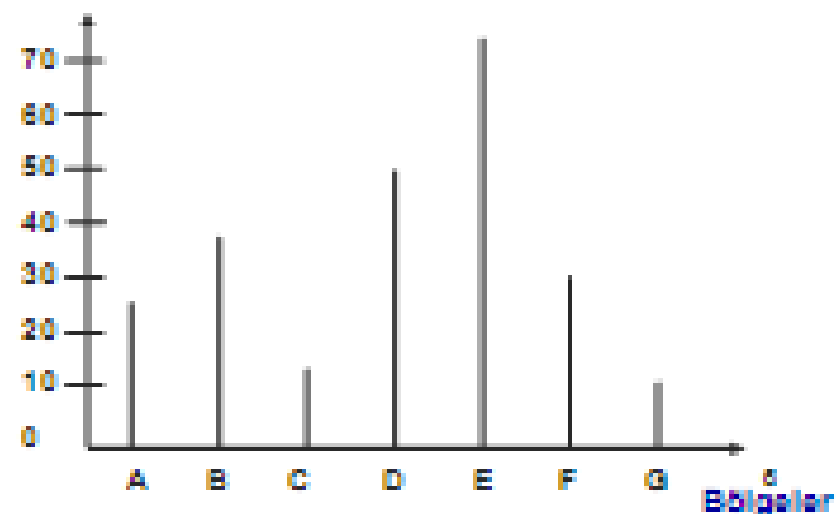
Dağılım Serilerinin Grafikleri (Frekans Serisi)

- Frekans serilerinin grafikleri, değişkenin aldığı değerlere göre frekansların dağılımlarını gösterir ve bir koordinat sistemi üzerine çizilebilir.

ÖRNEK: Bölgelere göre televizyon sayılarının dağılımı aşağıda verilmiştir. Bu dağılımın grafiğini çiziniz.

Bölgeler	Televizyon Sayısı (1=10.000)
A	25
B	36
C	12
D	49
E	71
F	28
G	12

TV Sayısı



Dağılım Serilerinin Grafikleri (Sınıflı Seri)

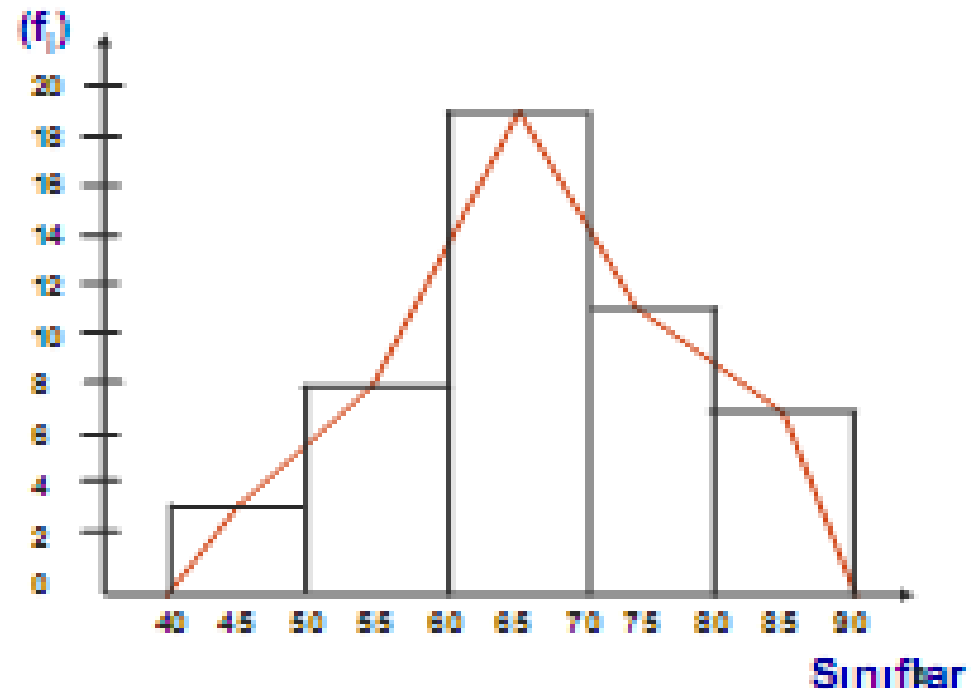
Tanım: Sınıf aralıklarının alt ve üst sınırlarından frekans değerlerine kadar çizilen dikmelerin, yatay eksene paralel çizgiler ile birleştirilmesi ile elde edilen dikdörtgenlerin tümüne histogram denir.

Tanım: Histogramı oluşturan bu dikdörtgenlerin üst orta noktalarının birleştirilmesi ile elde edilen grafiğe de frekans poligonu denir.

Gruplandırılmış Serilerin Grafikleri

Örnek: Bir sınıftaki 48 öğrencinin ağırlıklarına göre dağılımı aşağıdadır. Bu serinin histogram ve frekans poligonunu çiziniz.

Sınıflar	(f_i)
40-50	3
50-60	8
60-70	19
70-80	11
80-90	7
	<u>48</u>



Gruplandırılmış Serilerin Grafikleri (Minitab)

1

The image illustrates the steps to create a histogram for grouped data in Minitab. The main window shows a data table with columns C1 and C2. The 'Histogram' menu option is highlighted. A 'Histogram of C1' window displays a histogram with a frequency axis from 0 to 10 and a C1 axis from 40 to 60. A 'Histogram with' dialog box is open, showing four histogram styles and the 'OK' button highlighted. Another 'Histogram with' dialog box is open, showing the 'Binned' option selected and the 'OK' button highlighted.

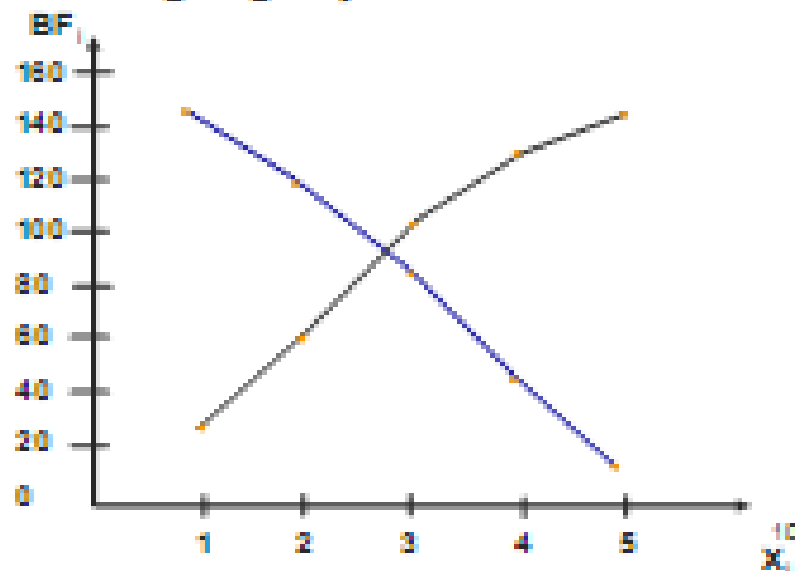
C1	C2
40	
40	
40	
41	
41	
41	
42	
42	
42	
43	
43	
43	
44	
44	
44	
45	
45	
45	
46	
46	
46	
47	
47	
47	
48	
48	
48	
49	
49	
49	
50	
50	
50	
51	
51	
51	
52	
52	
52	
53	
53	
53	
54	
54	
54	
55	
55	
55	

Dağılım Serilerinin Grafikleri (Artan-Azalan Birikimli Frekans Serileri)

- Artan ve azalan birikimli frekans serilerinin grafikleri de koordinat sistemi üzerine çizilebilir. Değişkenin aldığı değerler ile birikimli frekansların kesiştiği noktaların birleştirilmesiyle birikimli serilerin grafiği çizilir.

ÖRNEK: Aşağıda verilen bir bölgedeki binaların kat sayısına göre dağılımı gösteren serinin artan ve azalan frekanslarının grafiğini çiziniz.

X_i	f_i	Artan Birikimli Frekanslar	Azalan Birikimli Frekanslar
1	26	26	145
2	34	60	119
3	43	103	85
4	29	132	42
5	13	145	13

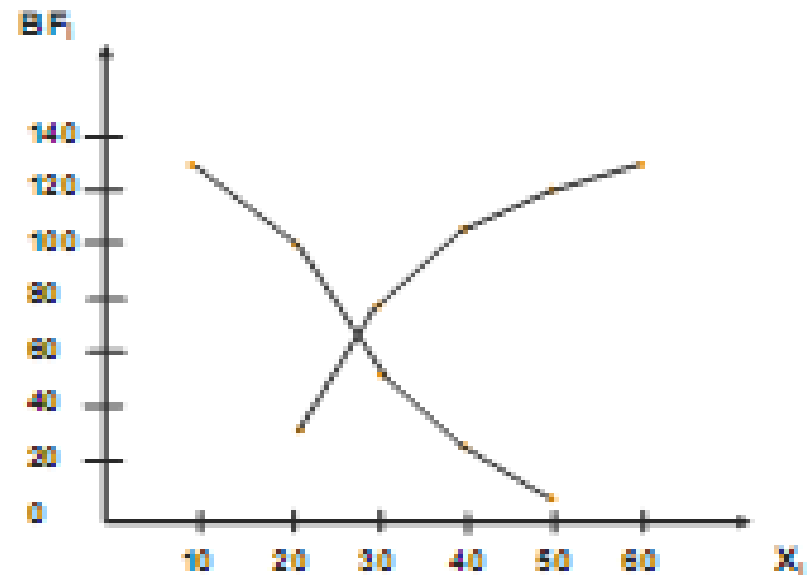


(Artan-Azalan Birikimli Sınıflı Frekans Serileri)

ÖRNEK: Dayanıklı tüketim malları satan bir mağazadaki aylık kredili satış miktarlarının kişilere göre dağılımı yansıda verilmiştir. Serinin artan ve azalan birikimli frekanslarının grafiğini çiziniz.

Sınıflı serilerde artan birikimli frekansların grafikleri çizilirken **sınıfın üst sınır değeri**, azalan birikimli frekanslar çizilirken ise **alt sınır değeri** alınır.

Sınıflar (1=1000TL)	(f)	Artan Birikimli Frekanslar	Azalan Birikimli Frekanslar
10-20	32	32	128
20-30	44	76	96
30-40	29	105	52
40-50	15	120	23
50-60	8	128	8



Bileşik Serilerin Grafikleri

- **Tanım:** Birimlerin birden fazla karaktere göre dağılımlarını gösteren serilere bileşik seriler denir.
- Bileşik serilerin grafikleri iki karakter söz konusu olduğunda bir karakter yatay ekseninde, diğer karakter dikey ekseninde gösterilerek çizilebilir. Karakter sayısı ikiden fazla ise çizim güçleşir.

ÖRNEK: Aynı iş kolunda faaliyet gösteren 6 işletmenin yıllık satış miktarları ve karları aşağıda verilmiştir. Verilen serinin grafiğini çiziniz.

İşletmeler	Satış Miktarı (1=10000TL)	Kar Miktarı (1=10000TL)
A	22	5
B	38	7
C	14	2
D	8	1
E	33	8
F	27	6



Merkezi Eğilim Ölçüleri (Ortalamalar)

Ortalamlar

- **Tanım:** Bir serideki tüm gözlem değerlerini temsil eden tek bir rakama ortalama denir.
- Ortalamalar özellikle tek maksimumlu serilerde gözlemlerin hangi değer etrafında toplandığını gösterir.
- Ortalama değer daima serinin minimum ve maksimum değerleri arasında yer alır.

$$X_{\min} \leq \text{Ortalama} \leq X_{\max}$$

Ortalamaların Faydaları:

1. Ortalamalar çoğu zaman serinin normal değerini gösterir. Tabii bunun için serinin dağılımının da aşırı çarpık olmaması gerekir.
2. İstatistik analiz işleminin temel elemanlarından biridir.
3. Aynı birimle ölçmek kaydıyla farklı serileri karşılaştırmaya imkan tanır.
4. Tek bir rakam olması sebebiyle hatırdan tutulması kolaydır.

Ortalamlar

Ortalamlar hesaplanmalarında kullanılan birimlere göre iki ana gruba ayrılabilir.

1. Grup (Analitik)

- Aritmetik Ortalama
- Kareli Ortalama
- Geometrik Ortalama
- Harmonik Ortalama

2. Grup (Analitik Olmayan)

- Mod
- Medyan

- 1. Gruptaki ortalamaların değeri, serinin herhangi bir biriminin değeri değiştiğinde değişir.
- 2. gruptaki ortalamaların değerinin değişmesi için bu ortalamaların hesabında kullanılan birimlerin değerinin değişmesi gerekir.

Aritmetik Ortalama

- Tanım:** Serideki gözlem değerleri toplamının, toplam gözlem sayısına oranı şeklinde hesaplanır.

Basit seride

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{N} = \frac{\sum_{i=1}^n X_i}{N}$$

Frekans serisinde

$$\bar{X} = \frac{f_1 X_1 + f_2 X_2 + \dots + f_n X_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_i X_i}{\sum_{i=1}^n f_i}$$

Gruplandırılmış seride

$$\bar{X} = \frac{f_{1m_1} + f_{2m_2} + \dots + f_{km_k}}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^k f_i m_i}{\sum_{i=1}^k f_i}$$

X_i : i . gözlem değeri,

f_i : i . gözlemin tekrar sayısı (frekansı)

m_i : i . sınıfın sınıf orta noktası

Aritmetik Ortalama

Örnek 1: Bir işletmede aynı parçayı üreten işçilerin bu parçayı üretim sürelerinin dağılımı aşağıdaki gibi gözlenmiştir. Parça üretim süresinin aritmetik ortalamasını bulunuz.

Parça üretim süresi (dakika) (X_i)	İşçi sayısı (f_i)	$f_i X_i$
12	2	24
13	4	52
14	7	98
15	6	90
16	1	16
Toplam	$\sum f_i = 20$	$\sum f_i X_i = 280$

$$\bar{X} = \frac{\sum_{i=1}^5 f_i X_i}{\sum_{i=1}^5 f_i} = \frac{280}{20} = 14 \text{ dk/parça}$$

Aritmetik Ortalama

Örnek 2: Bir işyerinde yapılan telefon görüşmelerinin süresinin dağılımı için aşağıdaki gruplanmış seri verilmiştir. Buna göre görüşme süresinin aritmetik ortalamasını bulunuz.

Görüşme süresi (dakika)	Görüşme sayısı (f_i)	m_i	$f_i m_i$
0 - 2 den az	5	1 $(0+2)/2$	5
2 - 4 " "	10	3 $(2+4)/2$	30
4 - 6 " "	40	5 $(4+6)/2$	200
6 - 10 " "	30	8	240
10 - 20 " "	25	15	375
Toplam	$\Sigma f_i = 110$		$\Sigma f_i m_i = 850$

$$\bar{X} = \frac{\sum_{i=1}^k f_i m_i}{\sum_{i=1}^k f_i} = \frac{850}{110} \cong 7.73 \text{ dk./görüşme}$$

Gruplanmış serilerde ortalamayı hesaplayabilmek için her sınıfı temsil eden tek bir değere ihtiyaç vardır. Bu değer o sınıfı en iyi temsil etmesi muhtemel olan sınıf orta noktasıdır.

$$m_i = \frac{l_{i1} + l_{i2}}{2}$$

Tartılı Aritmetik Ortalama

Tanım: Bir serideki değerler arasında önem derecesi farklı oluyorsa, bu tür serilerin aritmetik ortalaması tartılı olarak hesaplanır. Bunun için önem düzeyini gösteren katsayılar (tartılar) kullanılır.

Basit seride

$$X_T = \frac{\sum_{i=1}^k t_i X_i}{\sum t_i}$$

Frekans serilerinde

$$X_T = \frac{\sum_{i=1}^k t_i f_i X_i}{\sum_{i=1}^k t_i f_i}$$

Gruplanmış seride

$$X_T = \frac{\sum_{i=1}^k t_i f_i m_i}{\sum_{i=1}^k t_i f_i}$$

Tartılı Aritmetik Ortalama

Örnek 1: Aşağıda bir öğrencinin almış olduğu dersler, notları ve kredileri verilmiştir. Not ortalamasını tartılı aritmetik ortalama cinsinden hesaplayınız.

Dersler	Notlar (X_i)	Kredi (t_i)	$t_i X_i$
İstatistik	70	3	210
Matematik	60	4	240
Fizik	50	3	150
Kimya	80	2	160
Toplam	260	$\sum t_i = 12$	$\sum t_i X_i = 760$

Tartılı Ortalama

$$\bar{X}_T = \frac{\sum_{i=1}^n t_i X_i}{N} = \frac{760}{12} = 63.33$$

Basit Ortalama

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{260}{4} = 65$$

Tartılı Aritmetik Ortalama

Örnek 2: Bir işletmede işçilerin saat ücretleri çalıştıkları süre (kıdem) dikkate alınarak belirlenmektedir. Veriler aşağıdaki gibi olduğuna bu işletmede ortalama saat ücretini tartılı aritmetik ortalama cinsinden hesaplayınız.

Saat ücreti (milyon TL)	İşçi sayısı (f _i)	Ortalama kıdem (t _i)	m _i	f _i t _i	f _i t _i m _i	f _i m _i
1.00 – 1.40	10	2.5	1.20	25	30.0	12.00
1.40 – 1.60	30	5.0	1.50	150	225.0	45.00
1.60 – 1.80	50	9.5	1.70	475	807.5	85.00
1.80 – 2.00	15	13.0	1.90	195	370.5	16.90
2.00 – 2.50	5	18.0	2.25	90	202.5	11.25
Toplam	110			935	1635.5	170.15

Tartılı Ortalama

$$\bar{X}_T = \frac{\sum f_i t_i m_i}{\sum f_i t_i} = \frac{1635.5}{935} = 1.75 \text{ milyon TL/saat}$$

Basit Ortalama

$$\bar{X} = \frac{\sum f_i m_i}{N} = \frac{170.15}{110} = 1.547 \text{ milyon TL/saat}$$

20

Tartılı Aritmetik Ortalama

Tartılı aritmetik ortalamamanın kullanıldıđı yerler:

- Veriler arasında önem farkı bulunması halinde kullanılır.
- Oranların ve ortalamaların ortalaması hesaplanırken kullanılır.

Örnek: Bir işletmede bulunan üç tezgahın belli bir günde ürettikleri malların sayısı ve üretimlerindeki kusurlu oranları aşağıdaki tabloda verilmiştir. Buna göre bu tezgahlarının ürettiđi mamul kütlesinin ortalama kusurlu oranını bulunuz.

Tezgahlar	Üretim miktarı (t_i)	Kusurlu oranı (X_i)	$t_i X_i$
A	100	0.03	3
B	200	0.05	10
C	50	0.01	0.5
	$\Sigma t_i = 350$	$\Sigma X_i = 0.09$	$\Sigma t_i X_i = 13.5$

Tartılı Aritmetik Ortalama

Basit aritmetik ortalama ile kütlenin kusurlu oranı;

$$\bar{X} = \frac{\sum X_i}{N} = \frac{0.09}{3} = 0.03$$

Tartılı aritmetik ortalama ile kütlenin kusurlu oranı;

$$X_T = \frac{\sum_{i=1}^n t_i X_i}{\sum_{i=1}^n t_i} = \frac{13.5}{350} = 0.03857$$

Kütlenin gerçek kusurlu oranı 0.03857 dir

Aritmetik Ortalamamanın Özellikleri

1. Aritmetik ortalama hassas (diğer ortalamalara nazaran) bir ortalama olup serideki aşırı değerlerden etkilenir ve aşırı değere doğru kayma gösterir.

2. Serinin gözlem sayısı ile aritmetik ortalaması çarpılırsa serinin toplam değeri elde edilir.

$$N \cdot \bar{X} = \sum X_i$$

3. Serideki gözlem değerlerinin aritmetik ortalamadan sapmaları toplamı sıfır olur.

$$\sum (X_i - \bar{X}) = 0$$

Aritmetik Ortalamamanın Özellikleri

4. Serideki değerlerin aritmetik ortalamadan sapmalarının kareleri toplamı minimum dur.

$$\sum(X_i - \bar{X})^2 \rightarrow \text{minimum olur.}$$

5. Bir serinin değerleri, diğer iki serinin değerleri toplamından düşüyorsa bu serinin aritmetik ortalaması da diğer iki serinin aritmetik ortalamaları toplamına eşit dur.

$$\bar{X} = \bar{Y} + \bar{Z}$$

Geometrik Ortalama(G)

Tanım: Bir serideki gözlem değerlerinin birbirleri ile çarpımlarının, gözlem sayısı derecesinde kökünden alınması ile elde edilir.

$$G = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \cdots X_N}$$

$$G = \sqrt[N]{\prod_{i=1}^N X_i}$$

Geometrik Ortalama(G)

- Bu yoldan geometrik ortalamayı bulmak için gözlem sayısının az olması gerekir.
- Gözlem sayısı arttıkça bu yoldan geometrik ortalamayı hesaplamak güçleşmektedir.
- Bunun yerine logaritmik dönüşüm uygulanarak geometrik ortalama hesaplanır.

Basit seride

$$G = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_N}$$

$$\log G = \frac{\sum_{i=1}^N \log X_i}{N}$$

28

Geometrik Ortalama(G)

Frekans serileri için

$$G = \sqrt[f_1 + f_2 + f_3 + \dots + f_k]{X_1^{f_1} \cdot X_2^{f_2} \cdot X_3^{f_3} \cdot \dots \cdot X_k^{f_k}}$$

$$\log G = \frac{f_1 \log X_1 + f_2 \log X_2 + f_3 \log X_3 + \dots + f_k \log X_k}{f_1 + f_2 + f_3 + \dots + f_k} = \frac{\sum_{i=1}^k f_i \log X_i}{\sum_{i=1}^k f_i}$$

Gruplanmış seri için

$$G = \sqrt[f_1 + f_2 + f_3 + \dots + f_k]{m_1^{f_1} \cdot m_2^{f_2} \cdot m_3^{f_3} \cdot \dots \cdot m_k^{f_k}}$$

$$\log G = \frac{f_1 \log m_1 + f_2 \log m_2 + f_3 \log m_3 + \dots + f_k \log m_k}{f_1 + f_2 + f_3 + \dots + f_k} = \frac{\sum_{i=1}^k f_i \log m_i}{\sum_{i=1}^k f_i}$$

logG'yi G'ye çevirmek için; $G = 10^{\log G}$ dönüşümü yapılır

27

Harmonik Ortalama

Tanım: Harmonik ortalama bir serideki gözlem değerlerinin terslerinin aritmetik ortalamasının tersine eşittir.

Basit Seri için;

$$H = \frac{\frac{1}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \dots + \frac{1}{X_N}}}{N} = \frac{N}{\frac{1}{X_1} + \frac{1}{X_2} + \frac{1}{X_3} + \dots + \frac{1}{X_N}} \Rightarrow H = \frac{N}{\sum_{i=1}^N \frac{1}{X_i}}$$

Harmonik Ortalama

Frekans serisi için

$$H = \frac{\sum f_i}{\sum \frac{f_i}{X_i}}$$

Gruplanmış seride

$$H = \frac{\sum f_i}{\sum \frac{f_i}{m_i}}$$

Kareli Ortalama (K)

Tanım: Kareli ortalama serideki değerlerin karelerinin aritmetik ortalamasının kareköküdür

Basit seride

$$K = \sqrt{\frac{X_1^2 + X_2^2 + X_3^2 + \dots + X_N^2}{N}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^N X_i^2}{N}}$$

Tasnif edilmiş seride

$$K = \sqrt{\frac{f_1 X_1^2 + f_2 X_2^2 + f_3 X_3^2 + \dots + f_n X_n^2}{f_1 + f_2 + f_3 + \dots + f_n}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^n f_i X_i^2}{\sum_{i=1}^n f_i}}$$

Gruplanmış seride

$$K = \sqrt{\frac{f_1 m_1^2 + f_2 m_2^2 + f_3 m_3^2 + \dots + f_n m_n^2}{f_1 + f_2 + f_3 + \dots + f_n}} \Rightarrow K = \sqrt{\frac{\sum_{i=1}^n f_i m_i^2}{\sum_{i=1}^n f_i}}$$

Kareli Ortalama (K)

Örnek: Bir şehirdeki konutlarda elektrik enerjisi tüketimi üzerine yapılan araştırmada, 200 konut rasgele seçilmiş ve aşağıdaki sonuçlar elde edilmiştir.

Aylık elektrik tüketimi (Kwh)	Konut sayısı (f_i)	m_i	$f_i m_i$
0 – 60	10	30	9000
60 – 100	20	80	128000
100 – 120	40	110	484000
120 – 140	50	130	845000
140 – 180	45	160	1152000
180 – 250	35	215	1617875
Toplam	200		4235875

$$K = \sqrt{\frac{\sum_{i=1}^k f_i m_i^2}{\sum_{i=1}^k f_i}} = \sqrt{\frac{4235875}{200}} = \sqrt{21179,375} \Rightarrow K \approx 145,53 \text{ Kwh/ay}$$

Analitik Ortalamalar Arasındaki İlişkiler

Normal bir seride ortalamalar arasında aşağıdaki gibi bir büyüklük ilişkisi vardır.

$$K > \bar{X} > G > H$$

Mod

Tanım: Bir seride en çok tekrarlanan değere mod adı verilir.

İstatistikte nispeten az kullanılan bu ölçü özellikle verilerin simetrik bir dağılışı göstermediği durumlarda iyi bir ölçü olarak düşünülebilir.

Tanım: Eğer seride en çok tekrarlanan birden fazla eleman varsa bu tür seriler çok modlu seriler olarak isimlendirilir.

Böyle serilerde modun tek bir değerle ifade edilmesi istenirse seri gruplanmış hale dönüştürülerek modu hesaplanabilir. Gruplama sonrasında da en yüksek frekansa sahip tek bir sınıf bulunamazsa sınıflar birleştirilerek mod hesaplanabilir.

Mod

Örnek: Aşağıda basit ve tasnif edilmiş iki seri verilmiştir.
Bu serilerin modlarını bulunuz

x_i
10
12
14
15
15
15
16
16
16
18
20

Mod=15

x_i	(f _i)
10	1
12	1
14	1
15	3
16	2
18	1
20	1

Mod=15

Gruplanmış Seride Modun Bulunması

- Gruplanmış seride modu bulmak için serinin frekanslar sütunundan hareketle en çok tekrarlanan sınıf belirlenir.
- Belirlenen sınıf içerisinde mod'a karşılık gelen değeri elde etmek için aşağıdaki formül kullanılır

l_1 : mod sınıfı alt sınırı

Δ_1 : mod sınıfı ile bir önceki sınıf frekansı arasındaki fark

Δ_2 : mod sınıfı ile bir sonraki sınıf frekansı arasındaki fark

s : serinin sabit sınıf aralığı

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s$$

- Gruplanmış serilerde modu bulabilmek serinin sınıf aralıklarının eşit olmasına dikkat edilmelidir. Eğer sınıf aralıkları eşit verilmemişse, aralıkları eşit hale getirebilmek için sınıfları birleştirme yoluna gidilebilir. Buna rağmen sınıf aralıkları eşit hale getirilemiyorsa mod hesaplanamaz.
- Bazen seri iki, üç modlu olabilir. Böyle serilerde tek modlu hale getirmek için tasnif edilmiş seriyi sınıflandırmak, sınıflandırılmış serinin de sınıflarını birleştirmek gerekebilir.

Gruplanmış Seride Modun Bulunması

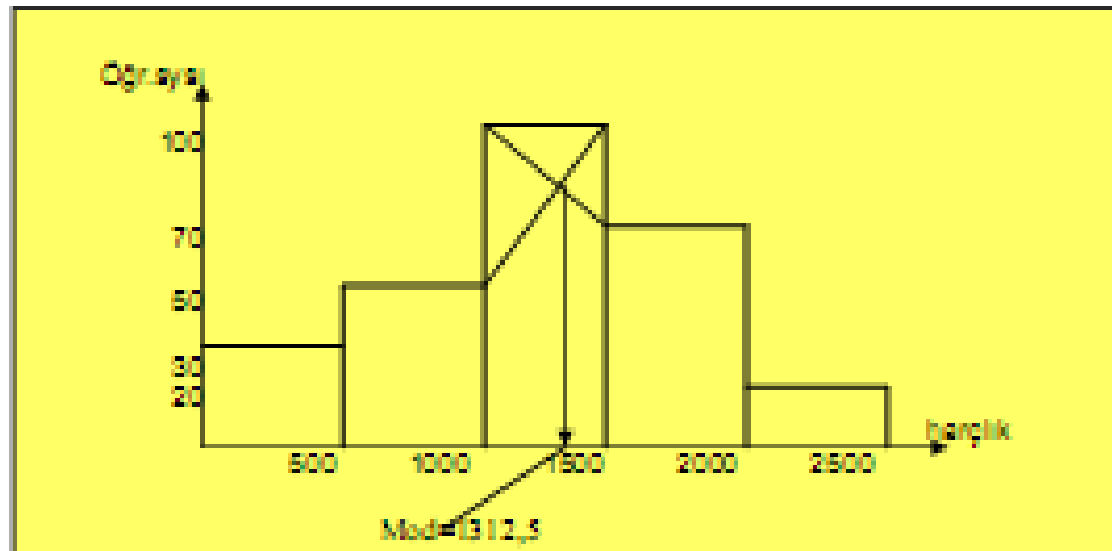
Örnek: Bir ilköğretim okulunda öğrencilerin günlük olarak aldıkları harçlıkların dağılımı aşağıda verilmiştir. Öğrencilerin aldıkları günlük harçlık miktarının ortalamasını mod ile belirleyiniz.

Harçlık (10 ³)TL/gün	Öğrenci sayısı
0 - 500	30
500 - 1000	50
1000 - 1500	100
1500 - 2000	70
2000 - 2500	20

$$Mod = l_1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \cdot s \Rightarrow Mod = 1000 + \frac{50}{50 + 30} \cdot 500 \Rightarrow Mod = 1312,5 \Rightarrow Mod = 1312500TL / gün$$

Modun Grafikle Gösterilmesi

- Modun grafikte gösterilebilmesi için serinin histogramı çizilir.
- Histogramda en yüksek sütun mod sınıfına karşılık gelir.
- Modun yerini tayin etmek için en yüksek sütunun üst köşegenleri ile komşu sütunların bitişik üst köşeleri çapraz olarak birleştirilir.
- İki doğrunun kesim noktasından yatay eksene çizilen doğrunun ekseni kestiği nokta mod olarak tespit edilir.



Modun Özellikleri

- Ortalamalar oranında en temsili alanıdır.
- Pratik hayatta çok kullanılan ortalamalardandır
- Özellikle kalitatif (niteliksel) serilerin ortalaması mod ile ifade edilir.
- Mod serideki aşırı değerlere karşı hassas değildir.
- Yukarıdaki avantajlarının yanında analitik olmaması sebebi ile matematik işlemlere elverişli değildir.
- J, ters J ve U tipi serilerde mod temsili alma özelliğini kaybeder. Böyle serilerde mod ya en küçük veya en büyük değere karşılık gelir

Medyan (Ortanca)

Tanım: Serideki değerler küçükten büyüğe sıralandığında tam ortaya düşen ve seriyi iki eşit parçaya bölen değere medyan adı verilir.

Basit ve tasnif edilmiş seride medyanın bulunuşu:

- Bunun için serideki değerler küçükten büyüğe sıralanır daha sonra medyana karşılık gelen değerın sıra değeri belirlenir.
- $\frac{N+1}{2}$ işlemi ile medyanın hangi sıradaki eleman olduđu belirlenir.
- Eğer bu işlemin sonucu tam sayı ise bu sıradaki eleman medyan olarak belirlenmiş olur. Eğer bu işlemin sonucu kesirli çıkarsa medyan iki değerin tam ortasına düşeceğinden bu iki değerin ortalaması alınarak medyan bulunur.

Medyan (Ortanca)

Örnek: $X_i: 15, 8, 12, 23, 45, 32, 5, 18, 16, 28, 39, 51$

Yukarıdaki serinin medyanını bulunuz.

Önce seri büyüklük sırasına göre dizilir

$X_i: 5, 8, 12, 15, 16, 18, 23, 28, 32, 39, 45, 51$

$$\frac{N+1}{2} = \frac{12+1}{2} = 6.5. \text{ sıradaki değer medyandır.}$$

Bu değer 18-23 arasında olduğundan

$$\text{Medyan} = \frac{18+23}{2} \Rightarrow \text{Medyan} = 20.5 \text{ olarak bulunur}$$

Gruplanmış Seride Medyanın Hesaplanması

- Gruplanmış seriler sürekli karakterde seriler olduğu için medyanın sıra değeri $N/2$ şeklinde bulunur.
- Bu değer toplam frekansın yarısına eşit olup serinin orta noktasını gösterir.
- Ortaya düşen sınıf yani medyan sınıfı kolayca belirlenebilir.
- Belirlenmiş olan medyan sınıfından hareketle aşağıdaki formül yardımı ile medyan değeri hesaplanır.

$$\text{Medyan} = l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{m-1} N_i}{N_m} \cdot s_m$$

l_1 : Medyan sınıfının alt sınırı

$N/2$: Medyanın sıra değeri

$m-1$

$\sum_{i=1}^{m-1} N_i$: Medyandan önceki frekanslar toplamı

$i=1$

N_m : Medyan sınıfının frekansı

s_m : Medyan sınıfının sınıf aralığı

Gruplanmış Seride Medyanın Hesaplanması

Örnek: Bir işletmede işçilere ödenen saat ücretlerinin dağılımı aşağıda verilmiştir. Bu verilere göre medyan saat ücretini hesaplayınız.

Saat ücreti (Bin TL)	İşçi sayısı
500 – 600	10
600 – 700	50
700 – 800 ← Medyan sınıfı	40 ←
800 – 1000	30
1000 – 1500	20
Toplam	150

$\frac{N}{2} = \frac{150}{2} = 75$ değer medyan dır. Bu değer 800 – 1000 ücret sınıfında yer almaktadır.

$$Medyan = l_1 + \frac{\frac{N}{2} - \sum_{i=1}^{k-1} N_i}{N_k} \cdot s_k \Rightarrow Medyan = 700 + \frac{75 - 60}{40} \cdot 100 \Rightarrow Medyan = 737.5 \text{ TL/saat}$$

Medyanın Grafikle Belirlenmesi

- Medyanın grafik üzerinde gösterilebilmesi için kümülatif ve ters kümülatif frekans serilerin oluşturulması gerekir.
- Bu serilerin grafiği birlikte çizildiğinde iki eğrinin birbirini kestiği noktadan yatay eksene çizilen doğrunun ekseni kestiği nokta medyan olarak tespit edilir. Bu işlem sadece eğrilerden birinin çizilmesiyle de yapılabiliriz.
- Eğrilerden biri çizildiğinde Y ekseninde $N/2$ değerine karşılık gelen noktadan X eksenine paralel çizildiğinde, bu doğrunun kümülatif eğriye temas ettiği noktadan X eksenine çizilen doğrunun ekseni kestiği noktada medyanı gösterecektir.

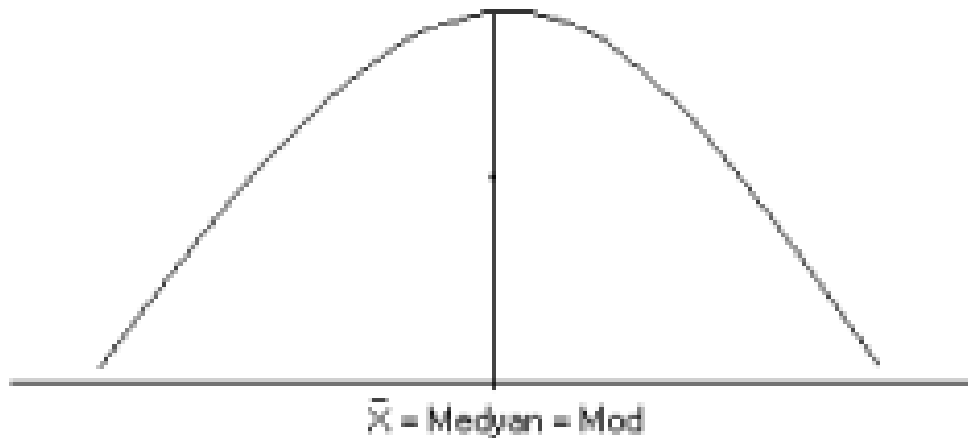
Medyanın Özellikleri

1. Pratik bir ortalamadır, sadece basit bir sıralama işlemi gerektirir.
2. Özellikle açık sınıflı seriler için medyan daha bir önem kazanır. Medyan böyle serilerin ortalamasında problemsiz olarak hesaplanabilir.
3. Serideki aşın değerlere karşı hassas değildir. Çünkü medyan serinin ortasına rastladığında, uçlarda oluşan aşın değerler medyanı etkilemez.
4. Serideki değerlerin medyandan mutlak farkları toplamı minimum olur. $\sum |X_i - \text{medyan}| \rightarrow \text{minimum}$
5. Medyanın zayıf tarafı serideki bütün değerleri dikkate almaması sebebi ile matematik işlemlere elverişli değildir.

Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

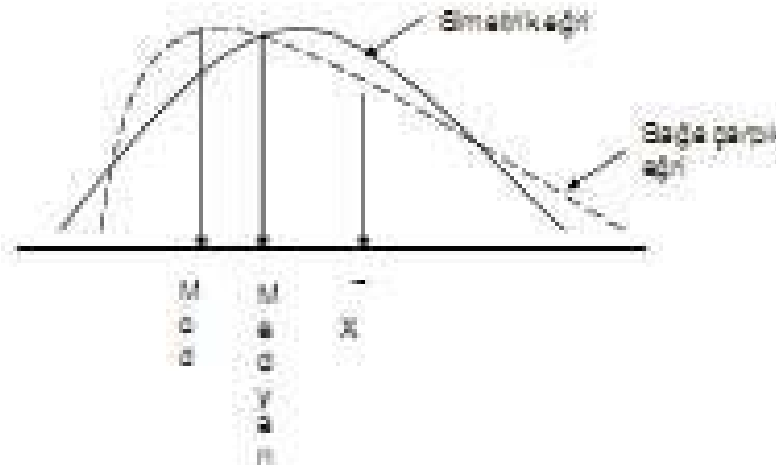
1- Simetrik serilerde her üç ortalama birbirine eşit olur.

$$\bar{X} = \text{medyan} = \text{mod}$$



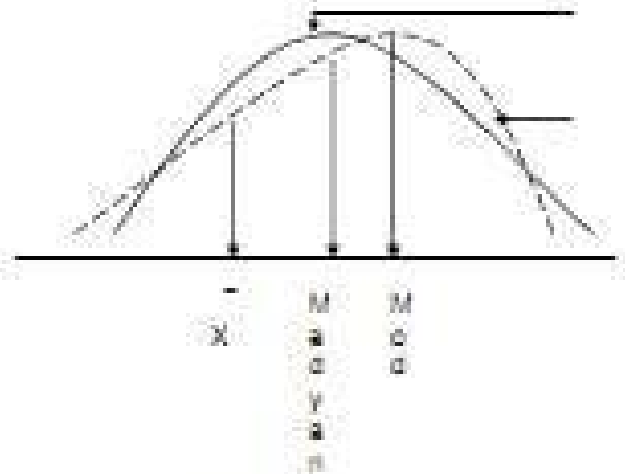
Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

2- Sağa çarpık serilerde $\bar{X} > \text{Medyan} > \text{Mod}$ olur.



Mod, Medyan ve Aritmetik Ortalama Arasındaki İlişkiler

3- Sola çarpık seride $\bar{X} < \text{Medyan} < \text{Mod}$ olur.



4- Asimetrisi hafif seriler için yaklaşık olarak aşağıdaki eşitlik geçerlidir.

$$(\bar{X} - \text{Mod}) \cong 3(\bar{X} - \text{Medyan})$$

DEĞİŞKENLİK ÖLÇELERİ

Giriş

Ortalamalar, serilerin karşılaştırılmasında her zaman yeterli ölçüler değildir. Aynı ortalamayı sahip seriler farklı dağılım gösterebilirler. Bu nedenle serilerin karşılaştırılmasında, değişkenlik ve asimetri ölçülerine bakılır.

Değişkenlik

- X_i ve Y_i aşağıdaki gibi iki seri verilmiş olsun:

X_i	Y_i
0	2
1	2
2	3
3	4
3	4

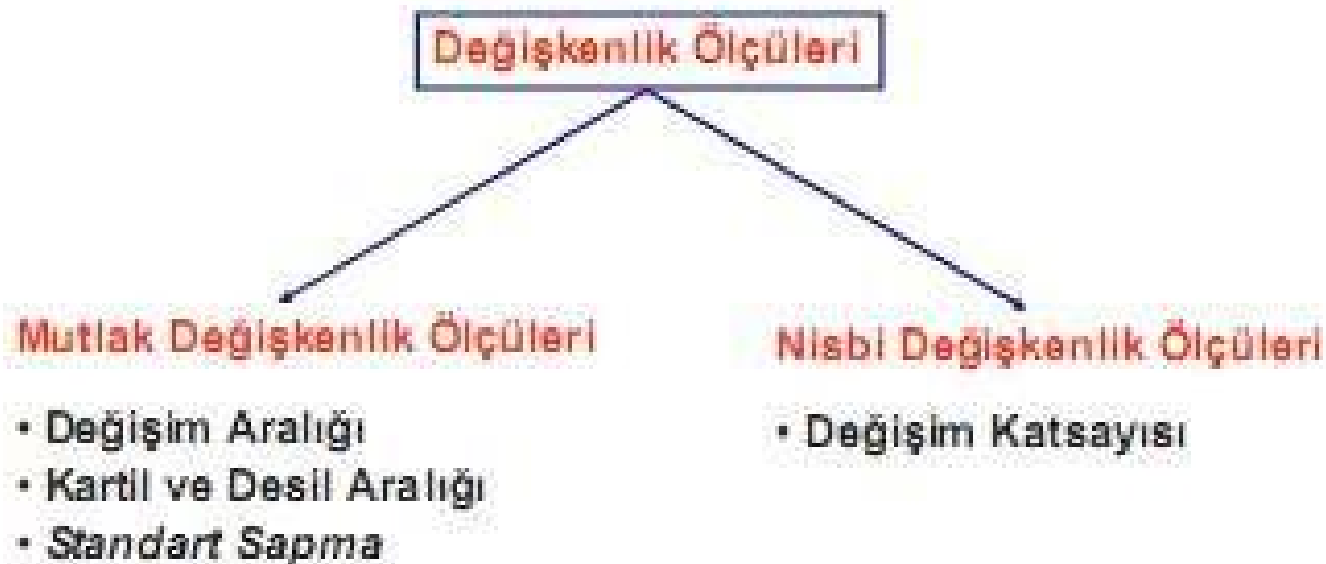
Serilerin aritmetik ortalamaları

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{15}{5} = 3$$

$$\bar{Y} = \frac{\sum_{i=1}^N Y_i}{N} = \frac{15}{5} = 3$$

Değişkenlik Ölçüleri

Değişkenliği az olan serilerin ortalamaları daha temsili oldukları halde, değişkenliği fazla olanların ortalamaları seriyi daha az temsil eder .



4

Değişim Aralığı

Tanım: Gözlem değerlerinin maksimum ve minimumu arasındaki fark olup, verilerin ne kadarlık bir aralıkta değiştiğini gösterir.

Değişim aralığı frekans serilerinde X_i sütununun maksimum ve minimum değerleri arasındaki farka, gruplandırılmış serilerde ise ilk grubun alt sınır değeri ile, son sınıfın üst sınır değeri arasındaki farka eşittir.

$$R = X_{\max} - X_{\min}$$

Örnek: X_i : 12,15,20,30,50,52,58,70,90 olan bir serinin değişim aralığı

$$R=90-12 =78$$

Yani gözlem değerleri 78 birimlik bir aralıkta değişim göstermektedir .

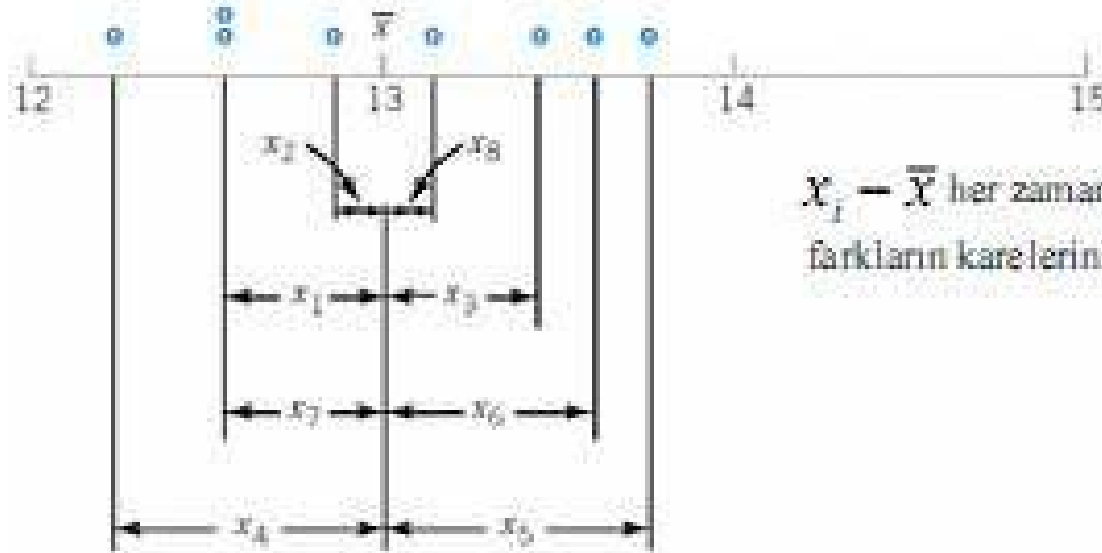
Standart Sapma

Tanım: Seri değerlerinin aritmetik ortalamadan farklarının kareli ortalamasına standart sapma denir.
Standart sapma σ ile gösterilir.

Tanım: Standart sapmanın karesine varyans adı verilir.
Varyans $V(X)$ yada σ^2 ile ifade edilir.

Standart Sapma

- Varyans serideki deęişimi nasıl ölçer?



$x_i - \bar{x}$ her zaman sıfır olacağından farkların karelerini alırız



Standart Sapma

Basit Seride

$$\sigma = \sqrt{\frac{\sum (X_i - \bar{X})^2}{N}}$$

Frekans Serisinde

$$\sigma = \sqrt{\frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i}}$$

Gruplanmış Seride

$$\sigma = \sqrt{\frac{\sum f_i (m_i - \bar{X})^2}{\sum f_i}}$$

■

Standart Sapma

Örnek: Doğru , yanlış şeklinde cevap şıkları olan 10 soruya öğrencilerin verdikleri doğru cevap sayılarının dağılımı aşağıda verilmiştir . Bu serinin standart sapmasını ve varyansını bulunuz .

Doğru Cev.Sayı	Öğ.Sayı
2	2
3	4
4	5
5	10
6	20
7	30
8	20
9	10
10	3

$$\sigma = \sqrt{\frac{\sum f(x) - \bar{X}}{\sum f}}$$

$$\bar{X} = \frac{696}{104} \approx 6,69$$

Standart Sapma

$$\sigma = \sqrt{\frac{\sum f(x) - \bar{X}}{\sum f}} = \sqrt{\frac{296,15}{104}} \Rightarrow \sigma = 1,69$$

Varyans $\sigma^2 = 2,847$ olarak bulunur.

2

Standart Sapma

Örnek: Bir liseden mezun olan ve ÖSS sınavına giren öğrencilerin puanlarının dağılımı aşağıda verilmiştir. Buna göre öğrenci puanlarının standart sapmasını bulunuz.

ÖSS Puanları	Öğr.Sayısı
90-110	10
110-130	30
130-150	50
150-170	25
170-210	5
Toplam	120

$$\bar{X} = \frac{16550}{120} \approx 137,9 \text{ puan}$$

$$\sigma = \sqrt{\frac{\sum f_i(m_i - \bar{X})^2}{\sum f_i}} = \sqrt{\frac{49979,2}{120}} = \sigma \approx 20,4 \text{ puan}$$

STANDART SAPMANIN KISA YOLDAN HESAPLANMASI

$$\sigma^2 = \frac{\sum (X_i - \bar{X})^2}{N} \quad \text{formülü açılarak yazılırsa}$$

$$\sigma^2 = \frac{\sum (X_i^2 - 2\bar{X} \cdot X_i + \bar{X}^2)}{N} \quad \text{şeklinde yazılabilir.}$$

Tanım: Standart sapma kareli ortalamasının (K) karesinden aritmetik ortalamasının karesi farkının kareköküdür.

$$\sigma^2 = \frac{\sum X_i^2}{N} - 2\bar{X} \cdot \frac{\sum X_i}{N} + \frac{N \bar{X}^2}{N} \quad \text{şeklinde ayrı ayrı ifade edilebilir}$$

$$\frac{\sum X_i^2}{N} = K^2, \quad \frac{\sum X_i}{N} = \bar{X} \quad \text{olduğuna göre}$$

$$\sigma^2 = K^2 - 2\bar{X} \cdot \bar{X} + \bar{X}^2 \quad \text{olur. Böylece } \sigma^2 = K^2 - \bar{X}^2 \quad \text{Varyans elde edilir. Buradan;}$$

$$\sigma = \sqrt{K^2 - \bar{X}^2}$$

□

STANDART SAPMANIN KISA YOLDAN HESAPLANMASI

Örnek:

ÖSS Puanları	Ögr.Sayı	m_i	$f_i.m_i$	$f_i.m_i^2$
90-110	10	100	1000	100000
110-130	30	120	3600	432000
130-150	50	140	7000	980000
150-170	30	160	4800	768000
170-190	20	180	3600	648000
190-210	10	200	2000	400000
	150		22000	3 328 000

$$\bar{X} = \frac{22000}{150} = 146,67 \text{ puan}$$

$$K^2 = \frac{3328000}{150} = 22186,67$$

$$\sigma = \sqrt{K^2 - \bar{X}^2} \cong 25,97 \text{ puan}$$

SHEPPARD DÜZELTMESİ

Tanım: Gruplanmış serilerde bir sınıfa düşen gözlem değerleri sınıfın orta değerine eşitmiş gibi kabul edilerek işlemler yapılması bir hataya sebep olmaktadır. Bu hatayı düzeltmek amacıyla Sheppard şu düzeltme formülünü geliştirmiştir.

$$\sigma' = \sqrt{\sigma^2 - \frac{s^2}{12}}$$

Buna göre $\sigma' \leq \sigma$ olacaktır.

σ' : Düzeltilmiş standart sapma
 σ^2 : Düzeltilmemiş varyans
 s : Sınıf aralığı

Örnek: Önceki slayttaki örnek için elde edilen standart sapmaya Sheppard düzeltmesini uygulayınız .

$$\sigma = 25,97 \quad s = 20$$

$$\sigma' = \sqrt{(25,97)^2 - \frac{20^2}{12}} = \sqrt{641,1} \Rightarrow \sigma' = 25,32 \text{ puan olur}$$

Standart Sapmanın Özellikleri

- 1) Matematik işlemler için uygun bir dağılım ölçüsüdür . Bu sebeple en yaygın kullanılan ölçüdür
- 2) Standart sapmada aritmetik ortalama gibi istatistik analiz için temel ölçülerden birisidir .
- 3) N_1 ve N_2 gözlemden oluşan iki serinin ortalamaları aynı ve sırayla varyansları σ_1 ve σ_2 olsun . Bu iki serinin birleştirilmiş ortak varyansı

$$\sigma^2 = \frac{N_1 \cdot \sigma_1^2 + N_2 \cdot \sigma_2^2}{N_1 + N_2}$$

şeklinde bulunur.

Değişim Katsayısı

Tanım: Standart sapmanın ortalamasının bir yüzdesi olarak ifade edilmesine değişim katsayısı adı verilir . Bu tanıma göre standart sapmanın büyüklüğü ortalamaya göre ifade edilmektedir .

$$D.K = \frac{\sigma}{\bar{X}} . 100$$

- Bu hesaplama ile ölçü birimlerinin etkisi giderilmiş olmaktadır. Bu nedenle bu karşılaştırılmak istenen serilerin değerleri farklı ölçü birimleri ile ifade edildiği durumlarda değişim katsayısı kullanılabilir.
- Değişim katsayısı küçük olan serilerde, birimlerin ortalama etrafında daha uygun dağıldıkları sonucuna varılır.

15

Değişim Katsayısı

Örnek: Konutlarda tüketilen aylık elektrik ve su miktarları için aşağıdaki veriler elde edilmiştir. Değişim katsayılarını bularak hangi grupta değişkenliğin daha fazla olduğunu araştırın.

Elektrik Tük. (kw/ş)	Konut Sayı
50-100	10
100-150	20
150-200	30
200-300	15
300-500	5
	80

Su Tük. (ton/ş)	Konut Sayı
5-15	20
15-25	30
25-35	40
35-45	20
45-65	10
	130

Değişim Katsayısı

Elektrik Tüketimi İçin

$$\bar{X} = \frac{14250}{80} = 178,125$$

$$K^2 = \frac{3025000}{80} = 37812,5$$

$$\sigma = \sqrt{K^2 - \bar{X}^2} = \sqrt{37812,5 - 178,125^2} = \sqrt{6083,98} \approx 78$$

$$DK = \frac{\sigma}{\bar{X}} \cdot 100 \Rightarrow DK = \frac{78}{178,125} \cdot 100 \approx 44,8$$

Su Tüketimi İçin

$$\bar{X} = \frac{3250}{110} = 29,55$$

$$K^2 = \frac{111250}{110} = 1011,4$$

$$\sigma = \sqrt{K^2 - \bar{X}^2} = \sqrt{1011,4 - 29,55^2} = \sqrt{138,16} \approx 11,75$$

$$DK = \frac{\sigma}{\bar{X}} \cdot 100 \Rightarrow DK = \frac{11,75}{29,55} \cdot 100 \approx 39,76$$

TT

ASİMETRİ ÖLÇELERİ

Asimetri Ölçüleri

Tanım: Serilerin dağılma şekillerini belirlemek için hesaplanan ölçülere asimetri ölçüleri denir.

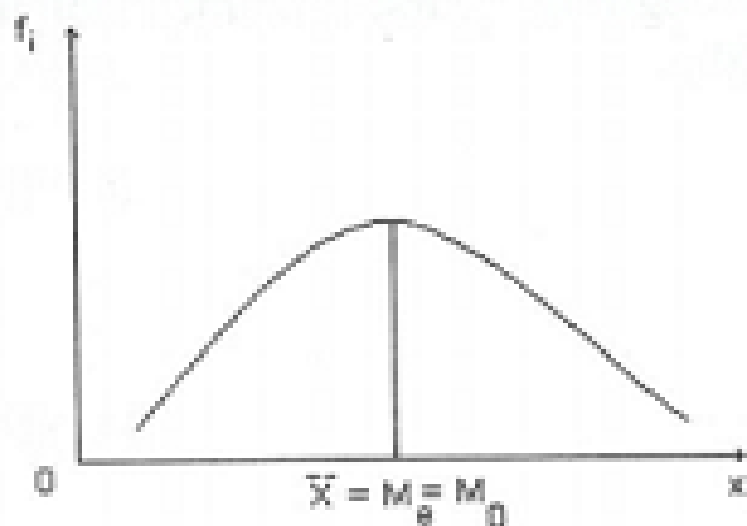
■ Değişkenlik ölçüleri serilerin değişkenliğini ölçebilmesine rağmen, onların dağılma şekilleri hakkında bilgi vermez. Oysa ortalamaları ve değişkenlik ölçüleri birbirine eşit veya yakın olan serilerin dağılımları birbirinden farklı olabilir. Bu nedenle serilerin sadece değişkenlik ölçüleri değil, dağılma şekilleri de araştırılmalıdır.

Asimetri Ölçüleri

- Dağılımı çan eğrisine benzeyen, tek maksimumlu ve simetrik olan dağılım istatistikte önemli bir yere sahiptir ve bazı yöntemlerin uygulanması dağılımların normal dağılım olduğu varsayımına dayanmaktadır.
- Bir dağılımın normal dağılım olarak kabul edilebilmesi için simetri durumunun bulunması ve basıklık ölçüsünün hesaplanması gerekir. Basıklık ölçüsü ile bir dağılımın basıklığının, normal dağılımın basıklığından farklı olup olmadığı belirlenmektedir.

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Simetrik dağılım gösteren tek modlu serilerde aritmetik ortalama, mod ve medyan birbirine eşittir.

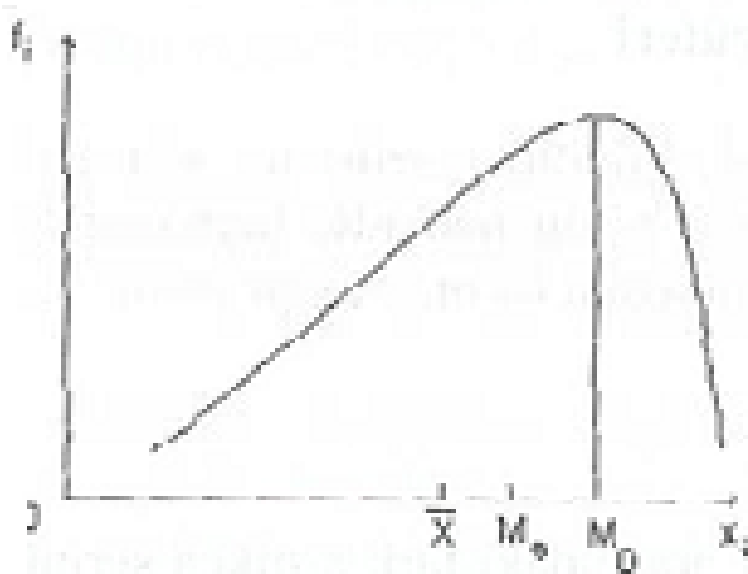


$$\bar{X} = M_e = M_o$$

İse seri simetriktir. Bu değerler arasındaki fark arttıkça serinin eğikliği artar.

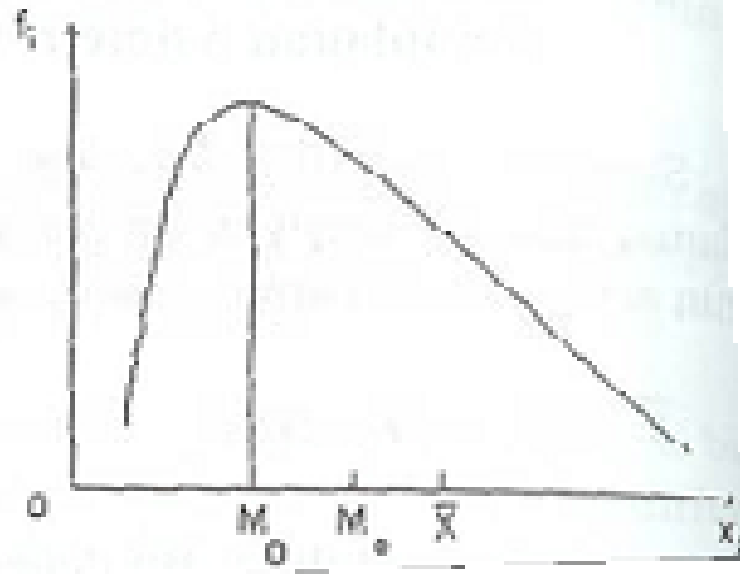
23

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri



$$\bar{X} < M_e < M_o$$

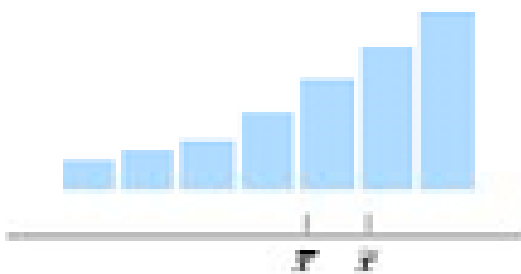
ise seri sola çarpıktır.



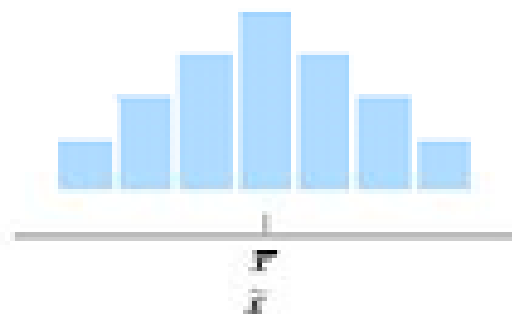
$$\bar{X} > M_e > M_o$$

ise seri sağa çarpıktır.

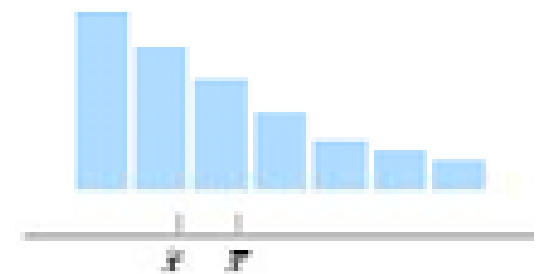
Asimetri Ölçüleri



Sola Eğik Seri



Simetrik Seri



Sağa Eğik Seri



Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Sağa ve sola eğikliği fazla olmayan serilerde

$$\bar{X} - M_o \cong 3(\bar{X} - M_e)$$

- Bu eşitlikten yararlanarak Pearson asimetri ölçüleri (α_1 ve α_2) hesaplanabilir.
- Bu ölçülerin hesaplanmasında farklı ölçü birimlerinin etkisini gidermek ve oransal bir ölçü etmek için aritmetik ortalama ile mod yada medyan arasındaki farklar standart sapmaya bölünür.

$$\alpha_1 = \frac{\bar{X} - M_o}{\sigma}$$

$$\alpha_2 = \frac{3(\bar{X} - M_e)}{\sigma}$$

Ortalamalar Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Seri simetrik ise,

$$\bar{X} = M_e = M_o \quad \text{olacağından} \quad \alpha_1 = 0 \quad \alpha_2 = 0$$

Seri sola eğikse,

$$\bar{X} < M_e < M_o \quad \text{olacağından} \quad \alpha_1 < 0 \quad \alpha_2 < 0$$

Seri sağa eğikse,

$$\bar{X} > M_e > M_o \quad \text{olacağından} \quad \alpha_1 > 0 \quad \alpha_2 > 0$$

Kartiller Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Asimetri ölçüleri ortalamalar arası farka dayanarak hesaplanabildikleri gibi, kartiller arası farka dayanarak da hesaplanabilir.

a) Simetrik serilerde

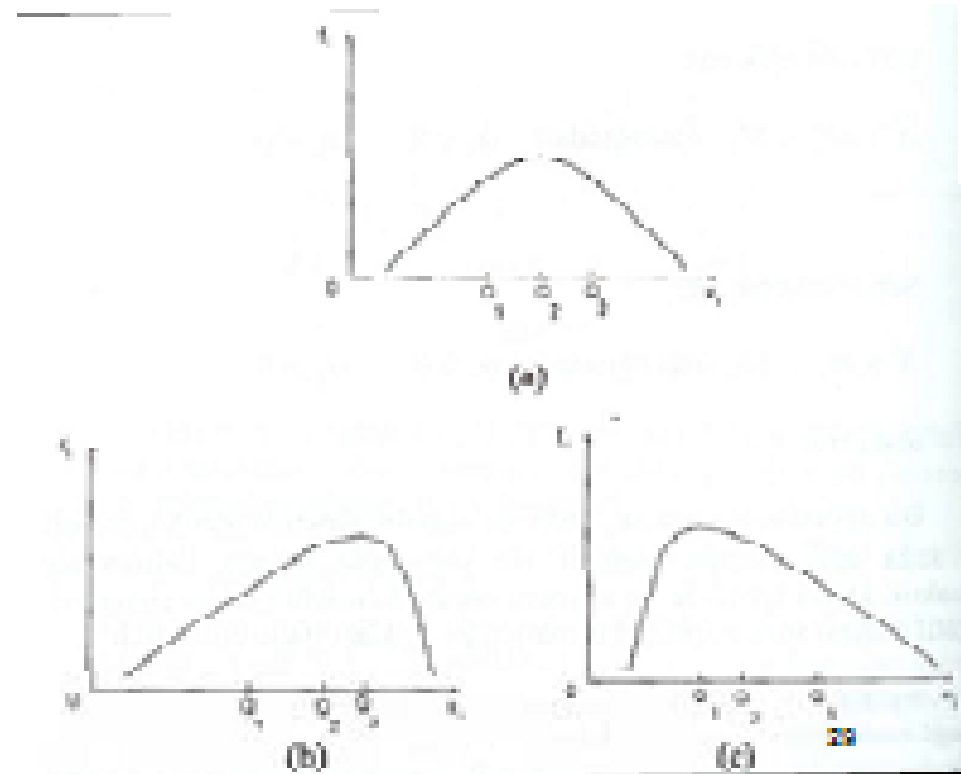
$$Q_3 - Q_2 = Q_2 - Q_1$$

b) Seri sola eğikse,

$$Q_3 - Q_2 < Q_2 - Q_1$$

c) Seri sağa eğikse,

$$Q_3 - Q_2 > Q_2 - Q_1$$



Kartiller Arası Farklar ile Hesaplanan Asimetri Ölçüleri

Tanım: Bu ölçülerin hesaplanmasında farklı ölçü birimlerinin etkisini gidermek ve oransal bir ölçü etmek için kartiller arası farklar, kartiller arası farklar toplamına bölünür. Bu şekilde elde edilen ölçü **Bowley asimetri** ölçüsü olarak adlandırılır.

$$A_s = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)} \quad A_s = \frac{(Q_3 - 2Q_2 + Q_1)}{Q_3 - Q_1}$$

Bowley simetri ölçüsüne göre elde edilen sonuç (+1) ile (-1) arasında değişmektedir. Sonuç:

- (0) ise seri simetrik,
- (-) ise sola asimetric,
- (+) ise sağa asimetrictir.

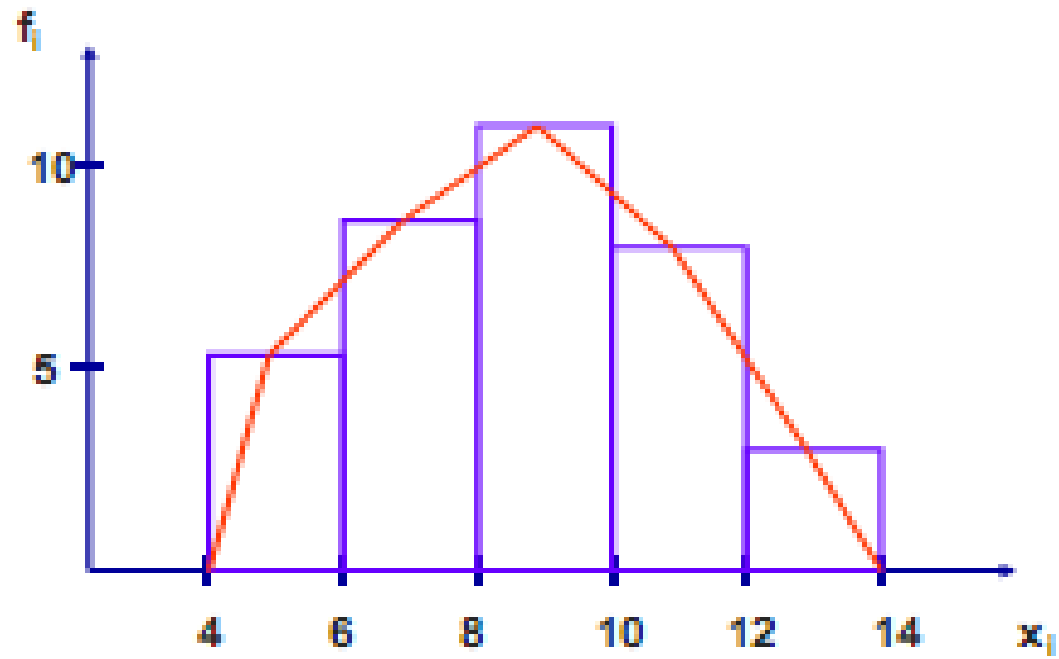
Örnek: Aşağıda verilen gruptandırılmış serinin simetrik olup olmadığını

- Grafiğini çizerek,
- Mod, medyan ve aritmetik ortalamasını karşılaştırarak,
- Pearson katsayılarını ve Bowley asimetri ölçüsünü hesaplayarak inceleyiniz.

Sınıflar	f	ArtanBirkimli Frekanslar	X_i	f X_i	f X_i^2
4-6	5	5	5	25	125
6-8	8	13	7	56	392
8-10	12	25	9	108	972
10-12	7	32	11	77	847
12-14	3	35	13	39	507
	35			305	2843

31

a)



32

$$b) M_o = I_o + \frac{\nabla_1}{\nabla_1 + \nabla_2} \cdot c = 8 + \frac{4}{4 + 5} \cdot 2 = 8,88$$

$$M_e = I_o + \frac{\frac{n}{2} - \sum_{i=1}^{k-1} f_b}{f_m} \cdot c$$

$$\frac{n}{2} = \frac{35}{2} = 17,5$$

$$M_e = 8 + \frac{17,5 - 13}{12} \cdot 2 = 8,75$$

$$\bar{X} = \frac{\sum fX_i}{\sum f_i} = \frac{305}{35} = 8,71$$

$$\bar{X} < M_e < M_o$$

olduğundan seri sola asimetriktir.

33

$$c) \quad \sigma = \sqrt{\frac{\sum f(x - \bar{X})^2}{\sum f}}$$

$$\sigma = \sqrt{\frac{2843}{35} - (8,71)^2} = 2,31$$

$$\alpha_1 = \frac{\bar{X} - M_e}{\sigma}$$

$$\alpha_1 = \frac{8,71 - 8,88}{2,31} = -0,07$$

$$\alpha_2 = \frac{3(\bar{X} - M_e)}{\sigma}$$

$$\alpha_2 = \frac{3(8,71 - 8,75)}{2,31} = -0,05$$

$\alpha_1 < 0$ $\alpha_2 < 0$ olduğundan seri sola asimetriktir. ³⁴

$$c) Q_{h/r} = I_a + \frac{n \frac{h}{r} - \sum_{i=1}^{k-1} f_i}{f_b} \cdot c$$

$$\frac{n}{4} = \frac{35}{4} = 8,75$$

$$\frac{3n}{4} = \frac{3(35)}{4} = 26,25$$

$$Q_1 = 6 + \frac{8,75 - 5}{8} \cdot 2 = 6,93$$

$$Q_2 = M_x = 8,75$$

$$Q_3 = 10 + \frac{26,25 - 25}{7} \cdot 2 = 10,35$$

$$A_x = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$$

$$A_x = \frac{(10,35 - 8,75) - (8,75 - 6,93)}{(10,35 - 8,75) + (8,75 - 6,93)} = \frac{1,6 - 1,82}{1,6 + 1,82} = -0,06$$

Sonuç negatif olduğundan seri sola asimetriktir. Ancak asimetri katsayısının değeri sifıra yakın olduğundan seri simetrik kabul edilebilir.

25

Momentler ile Asimetri ve Basıklık Ölçüleri

Tanım: Seri değerlerinin belirli bir orijine göre çeşitli derecelerden kuvvetlerinin ortalamalarına **moment** adı verilir.

Basit Seriler için

$${}_a m_q = \frac{\sum_{i=1}^n (X_i - a)^q}{n}$$

Frekans ve Gruplanmış Seriler için

$${}_a m_q = \frac{\sum_{i=1}^k f_i (X_i - a)^q}{\sum_{i=1}^k f_i}$$

$a=0$ ise orijine yada 0'a göre momentler (m)

$a = \bar{X}$ ise ortalamaya göre momentler (μ)

Momentler

Örnek: Aşağıda verilen frekans serisinin orijine ve ortalamaya göre momentlerini hesaplayınız.

X_i	f_i
1	12
2	17
3	20
4	21
5	10
	80

Orijine göre momentler:

$$m_1 = \frac{\sum_{i=1}^5 f X_i}{\sum_{i=1}^5 f_i} = \frac{240}{80} = 3 = \bar{X}$$

$$m_2 = \frac{\sum_{i=1}^5 f X_i^2}{\sum_{i=1}^5 f_i} = \frac{846}{80} = 10,575 = K^2$$

$$m_3 = \frac{\sum_{i=1}^5 f X_i^3}{\sum_{i=1}^5 f_i} = \frac{3282}{80} = 41,025$$

$$m_4 = \frac{\sum_{i=1}^5 f X_i^4}{\sum_{i=1}^5 f_i} = \frac{13530}{80} = 169,125$$

Momentler

X_i	f_i
1	12
2	17
3	20
4	21
5	10
	80

Ortalamaya göre momentler:

$$\mu_1 = \frac{\sum_{i=1}^5 f_i (X_i - \bar{X})}{\sum_{i=1}^5 f_i} = \frac{0}{80} = 0$$

$$\mu_2 = \frac{\sum_{i=1}^5 f_i (X_i - \bar{X})^2}{\sum_{i=1}^5 f_i} = \frac{126}{80} = 1,575 = \sigma^2$$

$$\mu_3 = \frac{\sum_{i=1}^5 f_i (X_i - \bar{X})^3}{\sum_{i=1}^5 f_i} = \frac{-12}{80} = -0,15$$

$$\mu_4 = \frac{\sum_{i=1}^5 f_i (X_i - \bar{X})^4}{\sum_{i=1}^5 f_i} = \frac{390}{80} = 4,875$$

Momentler Yardımıyla Hesaplanan Asimetri Ölçüsü

Momentler ile asimetri ölçüsü α_3 aşağıdaki şekilde hesaplanır.

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}}$$

$$\mu_2 = \sigma^2$$

$$\alpha_3 = \frac{\mu_3}{\sigma^3}$$

$\alpha_3 = 0$ ise seri simetrik

$\alpha_3 < 0$ ise seri sola çarpık

$\alpha_3 > 0$ ise seri sağa çarpıktır

Momentler Yardımıyla Hesaplanan Asimetri Ölçüsü

Örnek:

Sınıf	X_i	f_i
0-4	2	15
4-8	6	44
8-12	10	30
12-16	14	18
16-20	18	13
		120

$$\bar{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{1080}{120} = 9$$

$$\mu_2 = \sigma^2 = \frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i} = \frac{2664}{120} = 22,2$$

$$\sigma = \sqrt{22,2} = 4,71$$

$$\mu_3 = \frac{\sum f_i (X_i - \bar{X})^3}{\sum f_i} = \frac{5424}{120} = 45,2$$

$\alpha_3 > 0$
seri sağa çarpıktır

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3} = \frac{45,2}{(4,71)^3} = 0,43$$

Momentler ile Hesaplanan Basıklık Ölçüsü

Momentler ile basıklık ölçüsü α_4 aşağıdaki şekilde hesaplanır.

$$\alpha_4 = \frac{\mu_4}{\mu_2^2}$$

$$\mu_2 = \sigma^2$$

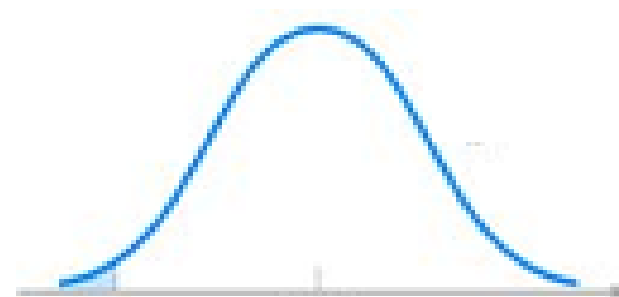
$$\alpha_4 = \frac{\mu_4}{\sigma^4}$$

Normal dağılım için $\alpha_4=3$ olduğundan

$\alpha_4 = 3$ ise seri normal

$\alpha_4 < 3$ ise seri basık

$\alpha_4 > 3$ ise seri sivri yada yüksektir.



Normal dağılım

Basıklık Ölçüsü

Örnek: Verilen serinin basıklığını inceleyiniz.

Sınıf	X_i	f_i
0-4	2	15
4-8	6	44
8-12	10	30
12-16	14	18
16-20	18	13
		120

$$\bar{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{1080}{120} = 9$$

$$\mu_2 = \sigma_2^2 = \frac{\sum f_i (X_i - \bar{X})^2}{\sum f_i} = \frac{2664}{120} = 22,2$$

$$\mu_4 = \frac{\sum f_i (X_i - \bar{X})^4}{\sum f_i} = \frac{136152}{120} = 1134,6$$

$\alpha_4 < 3$ seri basık

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4} = \frac{1134,6}{(22,2)^2} = \frac{1134,6}{492,84} = 2,230$$

ENDEKSLER

ENDEKSLER

- *Endeks Kavramı*
- *Mekan İndeksleri*
- *Zaman İndeksleri*
- *Bileşik İndeksler*

ENDEKS KAVRAMI

İstatistik oranlarının bir türü olan indeksler kısaca, sayısal değerlerin birbirine oranlanması ile elde edilir ve hesaplanabilmeleri için en az iki değer bilinmesi gerekmektedir.

Bir olaya ait bir veya birden fazla değişkenin, farklı yerlerdeki veya zaman içindeki oransal değişimini belirlemek için hesaplanan sayılar, “indeks sayılar” olarak tanımlanır. İndeks sayılar, kısaca “indeksler” diye adlandırılmaktadır. İndeksler, fiyat, üretim, yatırım, ücret, satış değişimlerinin belirlenmesi gibi farklı olaylarda kullanılabilir.

İndeksler, hesaplandıkları seri türüne göre zaman ve mekan indeksleri; hesaplandıkları madde sayısına göre basit ve bileşik indeksler olmak üzere ikiye ayrılırlar. Zaman ve mekan indeksleri aynı zamanda basit veya bileşik indeksler olabilir.

Zaman indeksleri, hesaplanmalarında temel alınan devreye göre sabit ve değişken esaslı indeksler olarak, bileşik indeksleri ise hesaplanmalarına konu olan maddelerin önem derecelerine göre tartılı ve tartısız bileşik indeksler olarak ikiye ayrılırlar.

Mekan İndeksleri

Mekan serilerinin mekan içindeki oransal değişimini belirlemek için hesaplanan indekslere mekan indeksleri denir. Mekan indeksleri, nüfus, üretim, tüketim, fiyat gibi çeşitli değişkenlerin bölgeler, iller gibi mekan birimlerine göre oransal değişimlerinin belirlenmesi için kullanılırlar.

Mekan indeksleri seriyi oluşturan değerlerden her birinin serinin aritmetik ortalamasına oranlanması ile hesaplanır.

$$I_i = \frac{X_i}{\bar{X}} \cdot 100$$

I_i = Hesaplanan
indeks

X_i = Seri değeri

\bar{X} = Seri değerinin
aritmetik
ortalaması

Örnek: Beş ayrı bölgenin 1992 yılı buğday üretim miktarları aşağıda verilmiştir. (1=10.000 ton)
Bu bölgelerin mekan indekslerini hesaplayınız.

Bölgeler	Buğ.Üret.(X _i)
----------	----------------------------

A	15
B	30
C	9
D	27
E	39
	120

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{120}{5} = 24$$

Mekan İndeksleri;

$$I_A = \frac{X_A}{\bar{X}} \cdot 100 = \frac{15}{24} \cdot 100 = 62,5$$

$$I_B = \frac{X_B}{\bar{X}} \cdot 100 = \frac{30}{24} \cdot 100 = 125$$

$$I_C = \frac{X_C}{\bar{X}} \cdot 100 = \frac{9}{24} \cdot 100 = 37,5$$

$$I_D = \frac{X_D}{\bar{X}} \cdot 100 = \frac{27}{24} \cdot 100 = 112,5$$

$$I_E = \frac{X_E}{\bar{X}} \cdot 100 = \frac{39}{24} \cdot 100 = 162,5$$

Bu sonuçlara göre, 1992 yılı buğday üretim miktarı, A, B, C, D, E bölgeleri ortalamasına göre, A bölgesinde %37,5 ($=100-62,5$), C bölgesinde %62,5 ($=100-37,5$) daha az, B bölgesinde %12,5 ($=112,5-100$) ve E bölgesinde %62,5 ($=162,5-100$) daha fazladır.

Zaman İndeksleri

Zaman serilerinin oransal deęişimini belirlemek için hesaplanan indekslere zaman indeksleri denir.

Herhangi bir zaman birimine ait indeksin hesaplanmasında, zaman serisinin herhangi bir zaman biriminin deęeri temel alınarak, dięerleri bu deęere oranlanabileceęi gibi, her zaman biriminin deęeri bir önceki zaman biriminin deęerine de oranlanabilir. Buna dayanarak zaman indeksleri “sabit esaslı” ve “deęişken esaslı” olarak ikiye ayrılır.

Sabit esaslı indeksler

Zaman indekslerinde hesaplanacak zaman birimlerini devre olarak adlandırırız; belirli bir devrenin değerini temel alıp, diğer devrelerdeki oransal değişimleri bu değere göre belirlemek için hesaplanan indekslere, “sabit esaslı indeksler” denir.

Sabit esaslı indekslerde hesaplanmaya temel alınan devre, “temel devre” veya “esas devre” olarak adlandırılır ve değeri 100’e eşittir.

$$I_i = \frac{X_i}{X_0} \cdot 100$$

I_i = i'nci devrenin indeks sayısı

X_i = indeksi hesaplanacak devrenin değeri

X_0 = esas devre değeri

Not: İndekslerde değişken genellikle X_i ile gösterilse de, fiyat indekslerinde P_i , miktar indekslerinde ise q_i değişkenleri kullanılmaktadır.

Örnek: Aşağıda bir malın 6 yıllık satış fiyatları verilmiştir. Bu verilerden yararlanarak 1994 yılı temel yıl ve 1997 yılı temel yıl olarak alındığına göre 6 yıllık sabit esaslı indeksleri hesaplayınız.

Yıllar	p (A malı fiyatı)	I (1994=100)	I (1997=100)
1994	35	100,0	125,0
1995	40	114,2	142,8
1996	32	91,4	114,2
1997	28	80,0	100,0
1998	50	142,8	178,5
1999	73	208,5	260,7

1994=100 (1994 temel yıl ise)

$$I_i = \frac{p_i}{p_0} \cdot 100$$

$$I_{94} = \frac{p_{94}}{p_{94}} \cdot 100 = \frac{35}{35} \cdot 100 = 100,0$$

$$I_{95} = \frac{p_{95}}{p_{94}} \cdot 100 = \frac{40}{35} \cdot 100 = 114,2$$

$$I_{96} = \frac{p_{96}}{p_{94}} \cdot 100 = \frac{32}{35} \cdot 100 = 91,4$$

$$I_{97} = \frac{p_{97}}{p_{94}} \cdot 100 = \frac{28}{35} \cdot 100 = 80,0$$

$$I_{98} = \frac{p_{98}}{p_{94}} \cdot 100 = \frac{50}{35} \cdot 100 = 142,8$$

$$I_{99} = \frac{p_{99}}{p_{94}} \cdot 100 = \frac{73}{35} \cdot 100 = 208,5$$

1997=100 (1997 temel yıl ise)

$$I_i = \frac{p_i}{p_0} \cdot 100$$

$$I_{94} = \frac{p_{94}}{p_{97}} \cdot 100 = \frac{35}{28} \cdot 100 = 125,0$$

$$I_{95} = \frac{p_{95}}{p_{97}} \cdot 100 = \frac{40}{28} \cdot 100 = 142,8$$

$$I_{96} = \frac{p_{96}}{p_{97}} \cdot 100 = \frac{32}{28} \cdot 100 = 114,2$$

$$I_{97} = \frac{p_{97}}{p_{97}} \cdot 100 = \frac{28}{28} \cdot 100 = 100,0$$

$$I_{98} = \frac{p_{98}}{p_{97}} \cdot 100 = \frac{50}{28} \cdot 100 = 178,5$$

$$I_{99} = \frac{p_{99}}{p_{97}} \cdot 100 = \frac{73}{28} \cdot 100 = 260,7$$

1994 yılı temel yıl olarak alındığında yapılabilecek yorum:

Malın fiyatları 1994 yılına göre 1995 yılında %14,2, 1998 yılında %42,8, 1999 yılında %108,5 artarken; 1996 yılında %8,6, 1997 yılında %20 azalmıştır.

1995 yılı temel yıl olarak alındığında yapılabilecek yorum:

Malın fiyatları 1997 yılına göre 1994 yılında %25, 1995 yılında %42,8, 1996 yılında %14,2, 1998 yılında %78,5 ve 1999 yılında %160,7 oranında yükselmiştir.

Sabit esaslı indekslerin hesaplanmasında, bir devre yerine, birkaç devrelik bir dönem de temel alınabilir. Bu durumda birkaç devrelik dönemin aritmetik ortalaması alınarak, bu ortalama esas devre değeri kabul edilir ve diğer devre değerleri hesaplanan bu ortalamaya oranlanır.

Örnek: Aşağıda bir bölgenin yıllık kömür üretim miktarları verilmiştir. Bölgenin yıllık kömür üretimine ait sabit esaslı indeksleri 1993-1995 dönemi temel alarak hesaplayınız.

Yıllar	Kömür üretimi (1=10000ton)
1993	11
1994	18
1995	16
1996	19
1997	13
1998	21

$$\bar{q}_0 = \frac{q_{93} + q_{94} + q_{95}}{3} = \frac{11+18+16}{3} = \frac{45}{3} = 15$$

$$I_i = \frac{P_i}{P_0} \cdot 100$$

$$I_{93} = \frac{q_{93}}{q_{93-95}} \cdot 100 = \frac{11}{15} 100 = 73,3$$

$$I_{94} = \frac{q_{94}}{q_{93-95}} \cdot 100 = \frac{18}{15} 100 = 120,0$$

$$I_{95} = \frac{q_{95}}{q_{93-95}} \cdot 100 = \frac{16}{15} 100 = 106,7$$

$$I_{96} = \frac{q_{96}}{q_{93-95}} \cdot 100 = \frac{19}{15} 100 = 126,6$$

$$I_{97} = \frac{q_{97}}{q_{93-95}} \cdot 100 = \frac{13}{15} 100 = 86,66$$

$$I_{98} = \frac{q_{98}}{q_{93-95}} \cdot 100 = \frac{21}{15} 100 = 140,0$$

Yorum: Bölgenin yıllık kömür üretim miktarı 1993-1995 dönemine (1993-1995 yılları ortalamasına) göre; 1993'de %16,7 daha az, 1994'de %20, 1995'de %6,7, 1996'da %26,6, 1998'de %40 daha az olmuştur.

Değişken esaslı indeksler

Zaman indekslerinde, indeksi hesaplanacak devre değerinin bir önceki devre değerine oranlanması ile hesaplanacak indeksler “değişken esaslı indeksler” olarak adlandırılır.

Değişken esaslı indeksler her devre indeksinin bir önceki devreye göre oransal artışını veya azalışını gösterecektir. Bu indekslerde hesaplanacak her devre için bir önceki devre 100 kabul edilmektedir.

I_i = i'nci devrenin indeks
sayısı

= indeksi hesaplanacak
devrenin değeri

$$I_i = \frac{X_i}{X_{i-1}} \cdot 100$$

X_i
= bir önceki devre değeri

X_{i-1}

Örnek: Aşağıda bir fabrikada üretilen yıllık buzdolabı miktarları verilmiştir. Bu verilerden yararlanarak değişken esaslı indeksleri hesaplayınız.

Yıllar	Üretim mik	I_i (1=1000ad)
1995	5	-
1996	4	80
1997	3	75
1998	6	200
1999	8	133

$$I_i = \frac{q_i}{q_{i-1}} 100$$

$$I_{96} = \frac{4}{5} 100 = 80$$

$$I_{97} = \frac{3}{4} 100 = 75$$

$$I_{98} = \frac{6}{3} 100 = 200$$

$$I_{99} = \frac{8}{6} 100 = 133$$

Yıllar	Üretim mik	I_i
1995	5	-
1996	4	80
1997	3	75
1998	6	200
1999	8	133

Yorum: Fabrikanın yıllık üretim miktarı, 1996 yılında 1995 yılına göre %20 azalmış; 1997 yılında 1996 yılına göre %25 azalmış; 1998 yılında 1997 yılına göre %100 artmış; 1999 yılında 1998 yılına göre %33 artmıştır.

Sabit esaslı indekslerden deęişken esaslı indekslerin hesaplanması

Yapılan araştırma için deęişken esaslı indeksler gerekiyorsa, ancak mevcut kaynaklardan sabit esaslı indeksler elde edilmişse, bu sabit esaslı indekslerin deęişken esaslı indekslere çevrilmesi gerekecektir.

Sabit esaslı indeksler,

$$I_{i,D} = \frac{I_{i,S}}{I_{i-1,S}} 100$$

formülü ile deęişken esaslı indekslere dönüştürülür.

Örnek: Bir önceki örnekte hesaplanan sabit esaslı indeksler yardımı ile değişken esaslı indeksleri hesaplayınız.

Devreler	$I_{(1=100)}$
1	100
2	80
3	90
4	125
5	140
6	175

$$I_{1,D} = -$$

$$I_{2,D} = \frac{80}{100} 100 = 80$$

$$I_{3,D} = \frac{90}{80} 100 = 112,5$$

$$I_{4,D} = \frac{125}{90} 100 = 138,88$$

$$I_{5,D} = \frac{140}{125} 100 = 112$$

$$I_{6,D} = \frac{175}{140} 100 = 125$$

Değişken esaslı indekslerden sabit esaslı indekslerin hesaplanması

Yapılan araştırma için sabit esaslı indeksler gerekiyorsa, ancak mevcut kaynaklardan değişken esaslı indeksler elde edilmemişse, bu değişken esaslı indekslerin sabit esaslı indekslere çevrilmesi gerekecektir.

Değişken esaslı indeksler,

$$I_i^S = \frac{I_{1,D}}{100} \cdot \frac{I_{2,D}}{100} \cdot \dots \cdot \frac{I_{i,D}}{100} 100$$

formülü ile sabit esaslı indekslere dönüştürülür.

Örnek: Bir önceki örnekte hesaplanan değişken esaslı indeksler yardımı ile sabit esaslı indeksleri hesaplayınız.

Devreler	$I_i(\text{D.E})$
1	--
2	80
3	112,5
4	138,88
5	112
6	125

$$I_1^S = 100$$

$$I_2^S = \frac{I_{1,D}}{100} \cdot \frac{I_{2,D}}{100} 100 = I_1^S \frac{I_{2,D}}{100} = 100 \frac{80}{100} = 80$$

$$I_3^S = \frac{I_{1,D}}{100} \cdot \frac{I_{2,D}}{100} \cdot \frac{I_{3,D}}{100} 100 = I_2^S \frac{I_{3,D}}{100} = 80 \frac{112,5}{100} = 90$$

$$I_4^S = \frac{I_{1,D}}{100} \cdot \frac{I_{2,D}}{100} \cdot \frac{I_{3,D}}{100} \cdot \frac{I_{4,D}}{100} 100 = I_3^S \frac{I_{4,D}}{100} = 90 \frac{138,88}{100} = 125$$

$$I_5^S = I_4^S \frac{I_{5,D}}{100} = 125 \frac{112}{100} = 140$$

$$I_6^S = I_5^S \frac{I_{6,D}}{100} = 140 \frac{125}{100} = 175$$

Sabit esaslı indekslerde temel devrenin değiştirilmesi

Herhangi bir devrenin esas alınması ile hesaplanan sabit esaslı indekslerde temel devre değiştirilebilir. Temel devreyi değiştirmek için kullanılacak formül aşağıdaki gibidir.

$$I_{i,y} = \frac{I_i}{I_{0,y}} \cdot 100$$

indeks

= yeni temel devreye
göre hesaplanacak

$I_{i,y}$ sabit esaslı indeks
= yeni temel devrenin

I_i
 $I_{0,y}$

Örnek: Aşağıda 1995 yılı temel yıl olarak düzenlenmiş indeksleri, temel yıl 1997 olarak düzenleyiniz.

Yıllar	$I_i(1995=100)$
1995	100
1996	120
1997	200
1998	160
1999	240

$$I_{95,y} = \frac{I_{95}}{I_{97}} 100 = \frac{100}{200} 100 = 50$$

$$I_{96,y} = \frac{I_{96}}{I_{97}} 100 = \frac{120}{200} 100 = 60$$

$$I_{97,y} = \frac{I_{97}}{I_{97}} 100 = \frac{200}{200} 100 = 100$$

$$I_{98,y} = \frac{I_{98}}{I_{97}} 100 = \frac{160}{200} 100 = 80$$

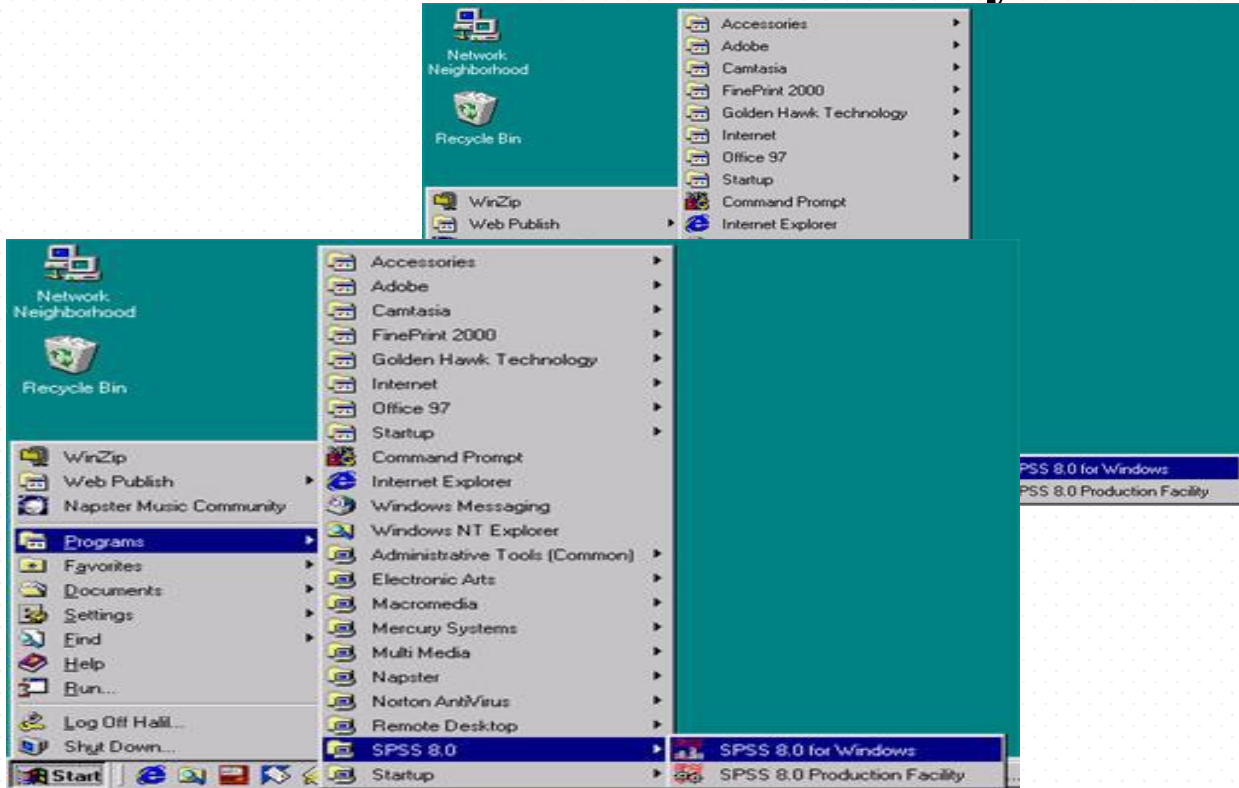
$$I_{99,y} = \frac{I_{99}}{I_{97}} 100 = \frac{240}{200} 100 = 120$$

SOSYOLOJİDE VERİ ANALİZİ 2 (SPSS Uygulamaları)

Prof.Dr. Ünal H. Özden

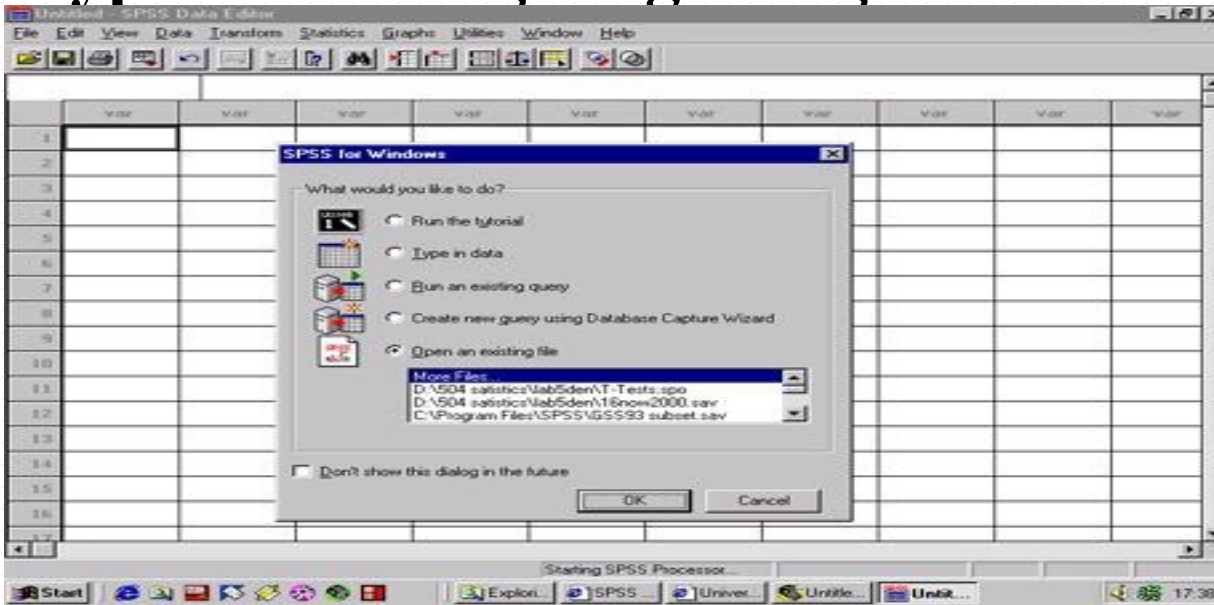
SPSS'i Bařlatmak

- SPSS'i Bařlatma Windows'da Start butonu ile aılan menuden Program Files'ı sein; yeni aılan menuden de SPSS X.0 seeneęini sein; son aılan menuden de SPSS X.0 For Windows'u tıcklayın.SPSS aılacaktır.



SPSS İlk Ekran

- İlk Ekran Karşınızda ilk göreceğiniz ekran, yan tarafta
- gördüğünüz pencere olacaktır. SPSS for Windows başlığı ile açılan küçük pencere, sizin ne yapmak istediğinizi seçmenizi bekliyor. İlk kez girdiğiniz için Type in data seçeneğini seçin. Yeni bir data dosyası



SPSS Ekran Görünümü

- SPSS'in 5 farklı ekran görünümü vardır.
 1. Data Editör (Soyadı "sav")
 2. Output (Soyadı "spo")
 3. Chart (Soyadı "sgt")
 4. Syntax (Soyadı "sps") (SPSS Makro dosyası)
 5. Script (Soyadı "sbs") (Basic dilinde yazılmış ek kodlar)
 6. Help

Data Editör Penceresi

- Data View: Değerlerin gösterildiği ekran
- Variable View: Değişkenlere ilişkin bilgilerin gösterildiği

	aday	oy_veren	v3	v4	var
1	1	79	1	18000	
2	2	32	1	23000	
3	2	50	2	21000	
4	2	56	2	24000	
5	2	51	2	19500	
6	2	48	1	24000	
7	2	29	2	20000	
8	2	40	2	19000	
9	1	46	2	16500	
10	1	37	2	25000	
11	1	43	2	23000	
12	2	45	2	20000	
13	2	53	2	14000	
14	1	34	1	15000	
15	1	32	1	18000	
16	2	53	1	14000	
17	1	27	1	32000	
18	2	42	2	28000	
19	2	27	2	26500	
20	1	40	2	19000	
21	2	74	1	17500	
22	2	41	1	31500	
23	1	26	1	25000	
24	1	49	2	28000	
25	1	49	2	32500	
26	2	37	2	17000	
27	2	55	2	14500	
28	1	26	2	16500	
29	1	25	1	22500	
30	2	47	1	20000	

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measur
1	aday	Numeric	11	0		None	None	8	Right	Scale
2	oy_veren	Numeric	11	0	Oy veren kisinin yasi	None	None	8	Right	Scale
3	v3	Numeric	11	0	Oy veren kisinin cinsiyeti	None	None	8	Right	Scale
4	v4	Numeric	11	0	Oy veren kisinin yıllık gelir durumu	None	None	8	Right	Scale
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										

Output Ekranı

Output1 - SPSS Viewer

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

Output

- Output
 - Frequencies
 - Title
 - Notes
 - Statistics
 - cinsiyet

→ Frequencies

Statistics

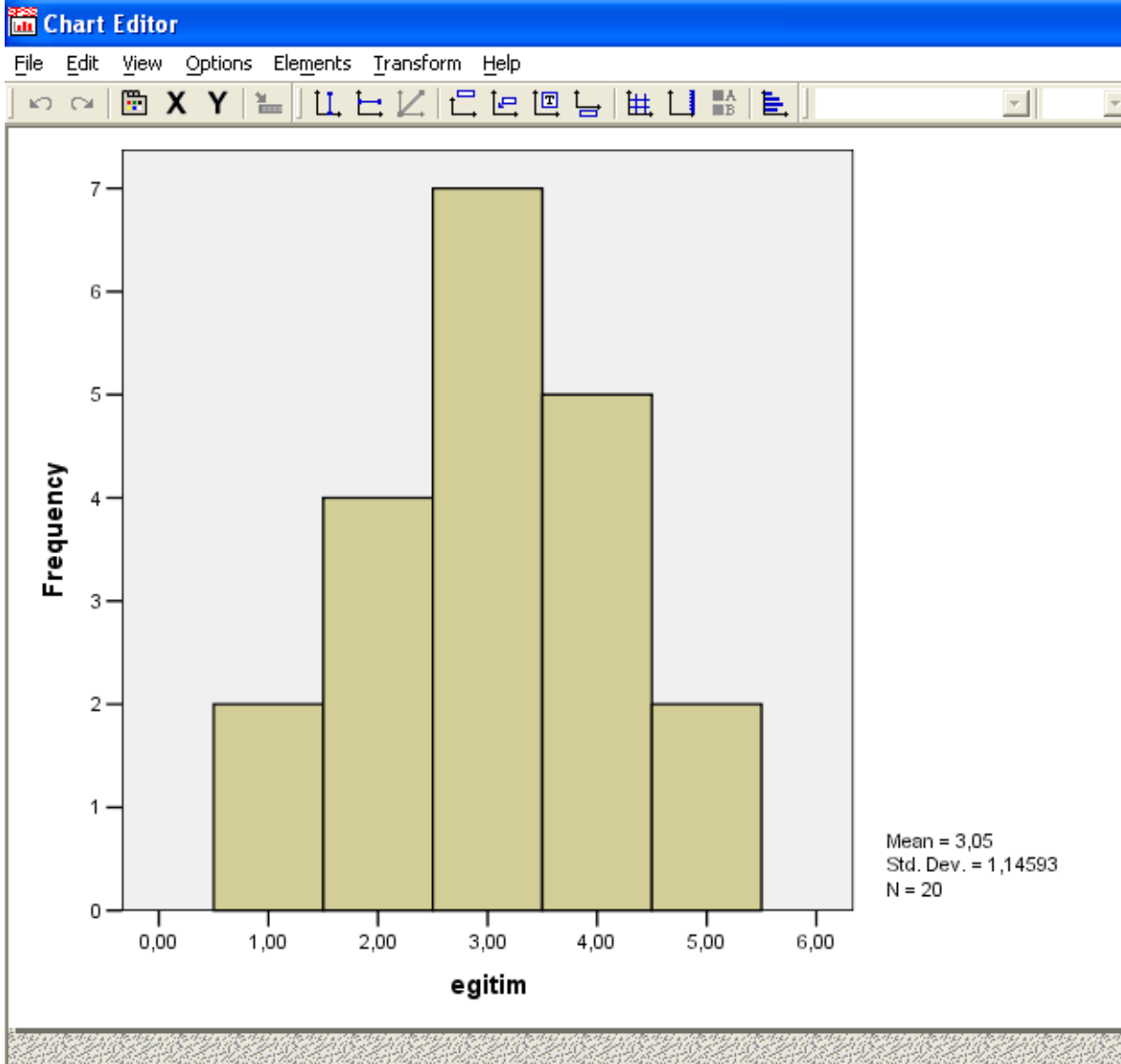
cinsiyet

N	Valid	20
	Missing	0

cinsiyet

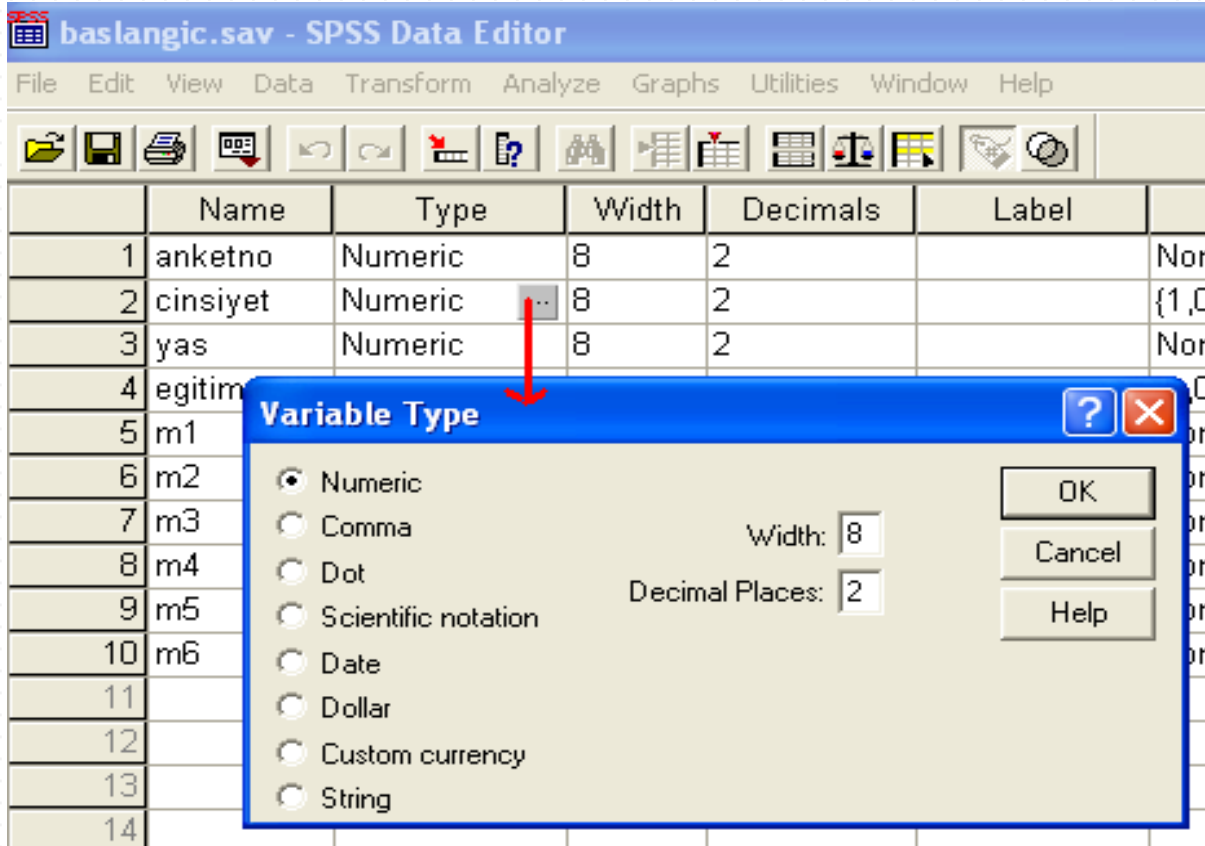
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid kiz	9	45,0	45,0	45,0
erkek	11	55,0	55,0	100,0
Total	20	100,0	100,0	

Chart Editör Ekranı



Değişken Tanımlama

- **Name:** Bu kutucuğa en fazla 8 karakter olacak şekilde değişkeninizin ismini girebilirsiniz.
- **Type...:** Değişkeninizin tipini belirlemenizi sağlar.



The screenshot shows the SPSS Data Editor window for a file named 'baslangic.sav'. The main window displays a table with columns for Name, Type, Width, Decimals, and Label. The table contains the following data:

	Name	Type	Width	Decimals	Label
1	anketno	Numeric	8	2	
2	cinsiyet	Numeric	8	2	{1,00
3	yas	Numeric	8	2	Non
4	egitim				
5	m1				on
6	m2				on
7	m3				on
8	m4				on
9	m5				on
10	m6				on
11					
12					
13					
14					

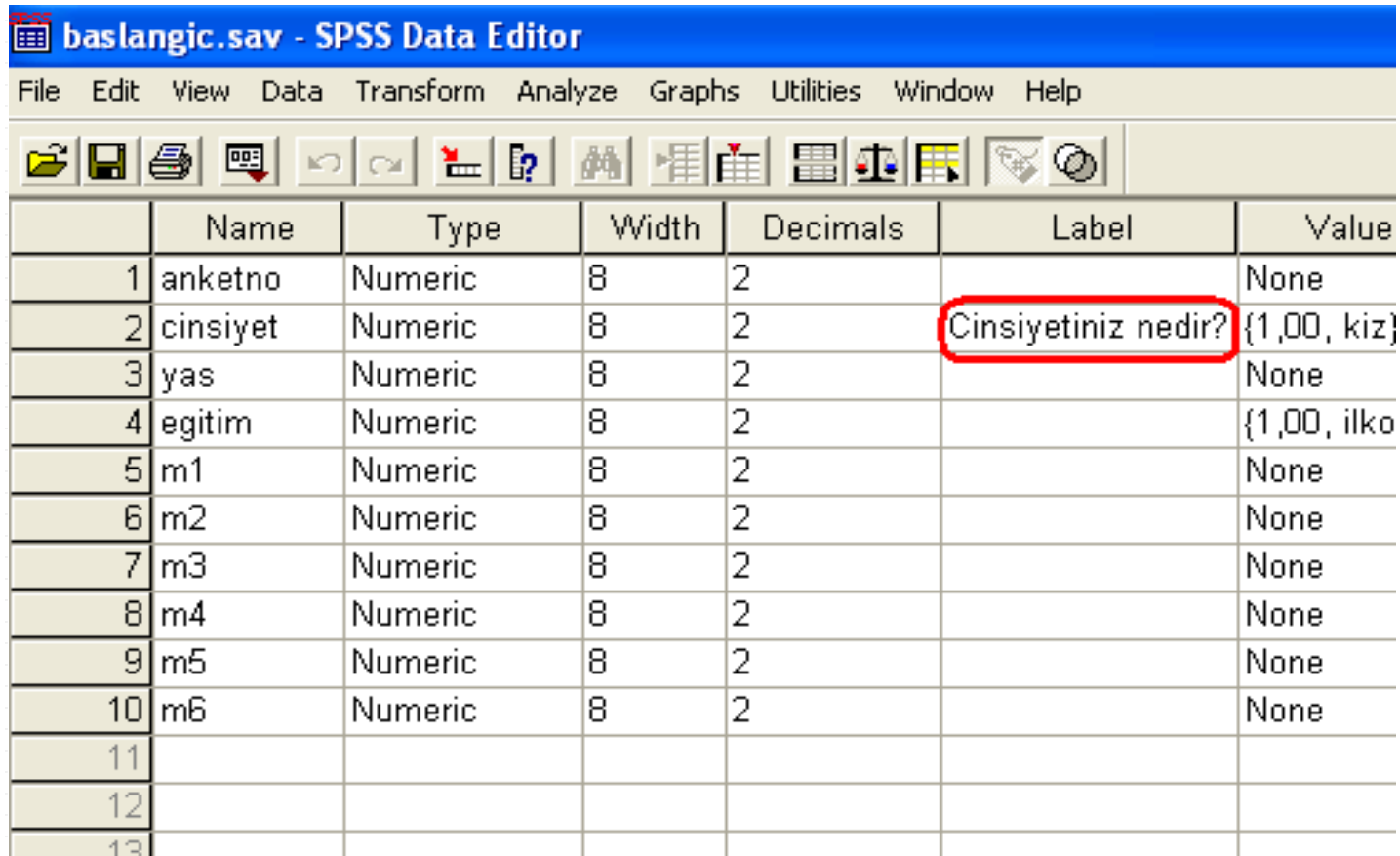
A red arrow points to the 'Type' column for the 'cinsiyet' variable, which has a small '...' icon next to it. A dialog box titled 'Variable Type' is open, showing the following options:

- Numeric
- Comma
- Dot
- Scientific notation
- Date
- Dollar
- Custom currency
- String

The dialog box also shows 'Width: 8' and 'Decimal Places: 2'. Buttons for 'OK', 'Cancel', and 'Help' are visible.

Labels

- **Labels...: Değişkeninizin değerlerine etiket (label) vermenizi sağlar.**



The screenshot shows the SPSS Data Editor window for a file named 'baslangic.sav'. The 'Labels' dialog box is open for the variable 'cinsiyet', and the label 'Cinsiyetiniz nedir?' is entered in the 'Label' field. The variable list below shows the following details:

	Name	Type	Width	Decimals	Label	Value
1	anketno	Numeric	8	2		None
2	cinsiyet	Numeric	8	2	Cinsiyetiniz nedir?	{1,00, kız}
3	yas	Numeric	8	2		None
4	egitim	Numeric	8	2		{1,00, ilko
5	m1	Numeric	8	2		None
6	m2	Numeric	8	2		None
7	m3	Numeric	8	2		None
8	m4	Numeric	8	2		None
9	m5	Numeric	8	2		None
10	m6	Numeric	8	2		None
11						
12						
13						

Value Labels

- Value

baslangic.sav - SPSS Data Editor

	Name	Type	Width	Decimals	Label	Values	Missing
1	anketno	Numeric	8	2		None	None
2	cinsiyet	Numeric	8	2	Cinsiyetiniz nedir?	{1,00, kız}..	None
3	yas	Numeric	8	2		None	None
4	egitim	Numeric	8	2		{1,00, ilkokul}	None
5	m1	Numeric					
6	m2	Numeric					
7	m3	Numeric					
8	m4	Numeric					
9	m5	Numeric					
10	m6	Numeric					
11							
12							
13							
14							
15							

Value Labels

Value Labels

Value: 1

Value Label: Kız

1,00 = "Kız"
2,00 = "Erkek"

Buttons: Add, Change, Remove, OK, Cancel, Help

Column Format-Measurement

- **Column Format...: Değişkeninizin bulunduğu sütunun kaç karakter genişliğinde olacağına, içindeki verilerin sağa, sola ya da ortaya yazılmasına karar verirsiniz. Orjinal hali 8 karakter ve sağ taraftır.**
- **Measurement Değişkeninizde kullanacağınız değerler sadece rakamdan oluşuyorsa (mesela yaş, gelir, notlandırmalar,vs..) Scale (Oransal);**
- **Hem rakam hem de metin içeriyorsa (kardeş sayısı: "1","2","3", "4 ve üzeri" vs..) Ordinal (sıralı);**
- **Sadece metin içerecekse (Sınıflar: A, B, C, D, vs...) Nominal seçeneklerini kullanın.**
- **Bu özellikler, değişkeniniz ile yapabileceğiniz**

işlemlerin sınırını belirler. Mesela Ordinal ve Nominal ölçekli değişkenlerin ortalaması bulunamaz, ya da aritmetik işlemlerde kullanılamaz. Onlar sadece gruplama yapmak için ve grafiklerde verileri temsil etmek için kullanılabilirler.

Scale düzeyindeki değişkenler ise her türlü hesaplamayı kabul eder. SPSS, siz belirlemediğiniz sürece girdiğiniz rakamlardan oluşan değişkenleri Scale yapacaktır.

Değişkenin düzeyini seçtikten sonra uygun olan tipi de (Type) belirlemeniz gerekir. Eğer Nominal değişken tanımlayıp, tipini numara olarak bırakmışsanız, o değişkene yazı yazamazsınız.String yapmanız gerekir.!

Columns	Align	Measure
3	Right	Scale
3	Right	Scale
3	Right	Scale
3	Right	Ordinal
3	Right	Nominal
3	Right	Scale
3	Right	Scale
3	Right	Scale
3	Right	Scale
3	Right	Scale

Veri Girişİ

baslangic.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

2 : cinsiyet 1 Split File

Değişken Adları

	anketno	cinsiyet	yas	egitim	m1	m2	m3	m4
1	1,00	kiz	20,00	ilkokul	2,00	1,00	3,00	
2	2,00	kiz	25,00	ortaokul	3,00	3,00	3,00	
3	3,00	kiz	23,00	ortaokul	2,00	2,00	2,00	
4	4,00	kiz	35,00	lise	3,00	3,00	3,00	
5	5,00	erkek	40,00	lise	4,00	4,00	4,00	
6	6,00	erkek	35,00	üniversite	5,00	4,00	5,00	
7	7,00	erkek	37,00	lise	4,00	3,00	4,00	
8	8,00	erkek	35,00	ortaokul	3,00	3,00	3,00	
		erkek	41,00	üniversite	4,00	4,00	4,00	
		erkek	55,00	lisans ustü	5,00	5,00	5,00	
		erkek	60,00	lisans ustü	4,00	5,00	4,00	
		erkek	23,00	lise	3,00	3,00	3,00	

Cinsiyet Değişkeninin ikinci sıradaki değeri

SPSS için Bazı Menüler / File Menüsü

- **Open**
- Open alt menüsünde; Data, Syntax, Output, Script, Other olmak üzere farklı formatlarda, kayıtlı olan dosyalar açılabilir. **Data** komutu kayıtlı bir veri dosyası, **Output** komutu ise kayıtlı bir çıktı dosyası açmak için kullanılır. **Syntax** tekrarlanan rutin işlemlerin yapılması için oluşturulan SPSS makrolarıdır. **Script** ise ek yazılım parçalarıdır.
- **Open Database:** SPSS programında açabileceğimiz diğer veri tabanı programlarındaki veri dosyalarını açmak amacıyla kullanılır.
- **Save:** Aktif penceredeki dokümanın kaydedilmesi için kullanılır
- **Save As:** Üzerinde çalışılan dosyayı, farklı bir yere, farklı bir isimle ve/veya farklı bir formatla kaydetmek için kullanılır. (“Variables” düğmesine tıklanarak yalnızca istenen değişkenler belirlenerek kaydedilebilir)
- **Display Data Information:** Working File seçeneği ile üzerinde çalışılan dosya, External File seçeneği ile HD, CD veya diğer fiziki ortamlarda daha önce kaydedilmiş “sav” uzantılı data dosyaları hakkında bilgi verir.
- **Read Data Text:** Farklı uzantılara (SPSS dosya uzantıları dışında) sahip daha önce oluşturulmuş dosyalardaki verileri SPSS programına aktarmada kullanılır.
- **Display Data File Information:** Sav uzantılı dosyalar hakkında bilgi verir. (Working file: üzerinde çalışılan dosya, External File: Fiziki ortamlarda kayıtlı -cd, hd, vs.- dosyalar.)
- **Cache Data:** Veri dosyasının geçici bir kopyasını oluşturmak amacıyla kullanılır.
- **Switch Server:** Kullanıcı adı ve şifre girmek suretiyle server üzerinde kayıtlı dosyalar üzerinde çalışmak için kullanılır.
- **Print:** Yazıcıdan çıktı alma.
- **Print Preview:** Sayfa önizleme yapmak için. Kullanılır.
- **Exit**
- Exit komutu ile SPSS programı kapatılır.
- **Recently Used Data:** Son kullanılan “sav” uzantılı data dosyalarının isimleri belirtilir ve bu dosyalara hızlı erişim imkanı sağlanır.
- **Recently Used File:** Son kullanılan ancak “sav” uzantılı olmayan dosyaların isimleri listelenir ve bu dosyalara hızlı erişim imkanı sağlanır.

Edit Menüsü

- **Undo:** Undo komutu, son yapılan işlemin iptal edilmesini sağlar.
- **Redo:** Undo komutu ile geri alınan bir işlevin tekrar gerçekleştirilmesini sağlar.
- **Cut:** Aktif pencerede seçilen bölümün kesilip panoya alınmasını sağlar.
- **Copy:** Aktif pencerede seçilen bölümün kopyalanıp panoya alınmasını sağlar.
- **Paste:** Daha önce kesilip ya da kopyalanıp panoya alınmış olan bölümün, seçilen alana yapıştırılmasını sağlar.
- **Paste Variable:** Değişken sayfasında Copy komutu ile kopyalanan bir değişkeni, seçilecek olan bir veya daha fazla değişken satırına yapıştırmak için kullanılır.
- **Clear:** Seçilen hücre satır veya sütunun silinmesini/boşaltılmasını sağlar.
- **Find:** Aktif pencerede, belirli bir metnin ya da değer bulunmasını sağlar.
- **Options:** SPSS'in varsayılan özelliklerinin kullanıcı tarafından değiştirilmesinin sağlandığı seçenektir.

View Menüsü

SPSS penceresinin bilgisayar ekranındaki görüntüsü ile ilgili değişiklikler yapılmasını sağlayan seçeneklerin bulunduğu menüdür.

- **Status Bar:** Durum çubuğunun pencere üzerinde bulunup bulunmayacağı belirler.
- **Toolbars:** Araç çubuğunun pencere üzerinde bulunup bulunmayacağı belirler.
- **Value Labels:** Veri sayfasında (Data View) kodlanmış olarak görülen verilerin etiketlerinin görüntülenmesini sağlar.
- **Grid Lines:** Veri sayfasındaki çizgilerin (Grid Lines) görüntülenip görüntülenmeyeceği bu menüden belirlenir.
- **Fonts:** Verilerin girildiği Data View, değişkenlerin tanımlandığı Variable View ve çıktı ekranı olan Output sayfalarındaki karakterlerin yazı tipi, stili ve boyutunu belirlemek için kullanılır.

Data Menüsü

- **Transpose**

Bir SPSS veri dosyasında her satır bir birime ve her sütun bir değişkene karşılık gelmektedir. Bu düzenin ters olduğu dosyalarda, veriyi doğru formata getirmek için bu alt menü kullanılabilir. Verilerin yer aldığı sütunların satırlara, satırların ise sütunlara dönüşmesini sağlar.

- **Sort Cases**

Sort Cases alt menüsü veri dosyasında kayıtlı olan birimlerin seçilen bir veya daha fazla değişkene göre küçükten büyüğe ya da büyükten küçüğe doğru sıralanmasını sağlar.

- **Split File**

Split File alt menüsü bir ya da daha fazla grup değişkenine göre, veri dosyasını gruplara ayırır. Böylece analizlerin her grup için ayrı ayrı yapılmasını sağlar.

- **Select Cases**

Farklı değerler alan değişken(ler)in istenilen değerlerinin analize katılması için kullanılır.(Cinsiyeti bayan olanlar üzerinde analiz yapmak istenmesi gibi)

- **Weight Cases:** Frekanslı serilerde veri sayfasına girilmiş değerlerin ilgili değişkene ait ağırlıkları temsil ettiğini tanımlamak için kullanılır.(bir sınıftaki notları ve bu notları alan öğrenci sayıları veri sayfasına girilirse öğrenci sayısı Weight Cases komutu ile mutlaka tanıtılmalıdır.)

Data Menüsü-devam

- **Insert Cases:** Data View sayfasında istenilen yere satır eklemek için kullanılır. Etkin olan hücrenin bulunduğu satırdan bir öncesine yeni bir satır ekler.
- **Insert Variable:** Data View veya Variable View sayfasında istenilen yere yeni bir değişken eklemek için kullanılır.
- **Go To Case:** Data View sayfasında istenilen satırdaki verilerin olduğu satıra gitmek için kullanılır.
- **Merge File:** İki ayrı dosyada ortak değişken ismi ile belirtilen verileri çalışılan dosyada birleştirmek ya da iki ayrı dosyada farklı değişkenlerle ifade edilen verileri tek bir dosyada toplamak için kullanılır. Merge Files alt menüsünde Add Cases ve Add Variables olmak üzere iki seçeneği vardır.
- **Identify Duplicated Cases:** (SPSS 12 ve Sonrası için) Herhangi bir değişkene ilişkin veriler içerisinde tekrarlanan verileri tespit etmek ve bu verileri yeni bir değişkende bazı değerler ile ifade etmek için kullanılır.

Transform Menüsü

- **Compute:** Compute alt menüsünde, nümerik ve string değişkenler için hesaplamalar yapılarak tüm birimler için yeni değişkenler oluşturulabilir ya da varolan bir değişkenin değerleri değiştirilebilir. Mantıksal koşullar altında verinin alt grupları için hesaplamalar yapılabilir. Aritmetik fonksiyonlar, istatistiksel fonksiyonlar, dağılım fonksiyonları ve string fonksiyonları ile 70'in üzerinde farklı fonksiyon oluşturulabilir.
- **Count:** Bir veya daha fazla değişkene ait veriler arasından kullanıcının belirleyeceği bir değer(ler)e eşit olanları yeni bir değişkende tespit etmek için kullanılır.
- **Recode:** Recode alt menüsündeki **Into Same Variables** alt menüsü, bir değişkenin belirlenen aralıktaki değerlerinin yeni kod değerlerine dönüştürülmesini ve bu kod değerlerinin aynı değişken üzerine yazdırılmasını sağlar. **Into Different Variables** alt menüsü, bir değişkenin belirlenen aralıktaki değerlerinin yeni kod değerlerine dönüştürülmesini ve bu kod değerlerinin farklı bir değişken olarak yazdırılmasını sağlar.
- **Automatic Recode** ile yazı (string) ile tanımlanan değerler 1, 2, 3, ... Şeklinde sırayla kodlanır.
- **Rank Cases:** Gözlem değerlerine büyüklüklerine göre sıra numarası verilir.
- **Categorize Variables** ile metrik değişkenler sınıflı seri haline getirilebilir.
- **Replace Missing Values:** Herhangi bir değişkene ilişkin sütunda yer alan veriler bazı birimler için eksik ya da cevapsız ise hücreleri değişkene ait verilerin ortalaması ile veya verileri kullanıcının belirleyeceği daha farklı bir fonksiyonla işleyerek doldurmak için kullanılır.

Değişkenlerin Özelliklerini Görüntüleme

- Utilities Menüsü altında Variable seçeneği işaretlenirse değişkenlerin özellikleri (nasıl kodlandığı vs.) görüntülenebilir.

The screenshot shows the SPSS Data Editor interface. The 'Utilities' menu is open, and the 'Variables...' option is selected. The 'Variables' dialog box is also open, showing the 'Variable Information' for the 'Yaş' variable. The 'Variables' dialog box has a list of variables on the left, with 'Yaş' selected. The 'Variable Information' section shows the following details for 'Yaş':

Variable	Label	Type	Missing Values	Measurement Level
Yaş	Kaç yaşındasınız?	F8	none	Scale

The 'Value Labels' section is empty. The 'Variables' dialog box has buttons for 'Go To', 'Paste', 'Close', and 'Help' at the bottom.

Ölçek Türlerine Göre Kullanılabilecek İstatistikler

Ölçü	Nominal	Ölçek			
	<i>Dikotom</i>	Nominal	Sıralı	Aralık	Oransal
Grafik Türü	Pasta, sütun	Pasta, sütun	Sütun	Histogram, Poligon, 2 değişken için X-Y	
Frekans Dağılımı	X	X	X	X	X
Mod		X	X	X	X
Medyan			X	X	X
Aritmetik/ Kareli Ortalama				X	X
Geometrik Ortalama					X
Minimum, maksimum, DA			X	X	X
Standart Sapma, Varyans				X	X
Asimetri Ölç. Skewness Beta1				X	X
Basıklık Ölç. Kurtosis Beta2				X	X
Bölenler ve Kartillerarası Sapma				X	X
Oran	X	X	X	XS	X

Ölçek Türlerine Göre İstatistiksel Testler

ÖLÇEK TÜRÜ	TEK DEĞİŞKEN DURUMU	FARKLI DURUMLARDAKİ UYGULAMALAR			
		2 BAĞIMSIZ ÖRNEK (Grup)	2 BAĞIMLI ÖRNEK	k BAĞIMSIZ ÖRNEK	k BAĞIMLI ÖRNEK
ARALIK ORANSAL	- μ , σ^2 ; durum parametreleri Anakütle testleri; z-, t- ve χ^2 -Asimetri ve basıklık testleri	-Ortalama farkı testleri (büyük örnekte z; küçük örnekte t) -Varyans Oranı F testi	-Eşlenik örnek ortalama testi (t dağılımı) -Pearson Çarpım Moment Korelasyon Katsayısı (çok boyutlu normal dağılım sağlanmalı bunun için → Mardia Uygunluk testi)	-ANOVA (her tipi) -Levene Homojenlik testi -Tek gözlemlili tesadüfi blok tasarımı*	-Mardia Testi (çoklu asimetri ve basıklık ölçülerine dayanır) -Bartlett küresel ilişki Testi (Çok değişkenli analizlerde kullanılır) -*
SIRALI	-Medyan (İşaret) Testi - χ^2 uygunluk testi -Kolmogorov- Smirnof Uygunluk Testi	-Wilcoxon Sıra Toplam Testi -Mann-Whitney-U	-Wilcoxon İşaretili Sıralar Testi -Spearman Sıra Korelasyon testi -Kolmogorov-Smirnof-Z Uygunluk testi -McNemar testi	-Medyan farkları kullanılan Levene testi de uygulanabilir. -Kruskal-Wallis-H Testi -Medyan testi	-Friedman Testi (pratikte sıralı ölçekte kullanılır ve Tek gözlemlili tesadüfi blok tasarımına alternatiftir.)
NOMİNAL	- χ^2 Uygunluk testi (çoklu oran testi)	- χ^2 Homojenlik testi	- χ^2 Bağımsızlık testi -Kappa Agreement (anlaşma) Testi [daha çok tıbbi çalışmalarda kullanılır]	- χ^2 Homojenlik testi	
İKİLİ (DICHOTOM)	-II oran parametresi testleri (z dağılımı) -Örnek hacmi küçükse Binom testi (medyan testi)	-Oran farkı testi (z dağılımlıdır)	-Phi Korelasyon Katsayısı -Fischer Exact Testi (özellikle küçük örnekte kullanılır çünkü asimptotik bir dağılıma sahiptir)	-Oran farkı testi (z dağılımlıdır) Bu amaçla χ^2 Bağımsızlık testi kullanılır.	

Frekans Tabloları

SPSS Data Editor

Data Transform Analyze Graphs Utilities Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Missing Value Analysis...
Complex Samples

Frequency...
Descriptives...
Explore...
Crosstabs...
Ratio...

AnketNo	Yaş	dirm	var	var	var	var
1	2					
38	4					
43	5					
54	3					
40	4					
32	6					
12	3					
23	4					
33	3					
11	3					

Frequency

Variable(s):

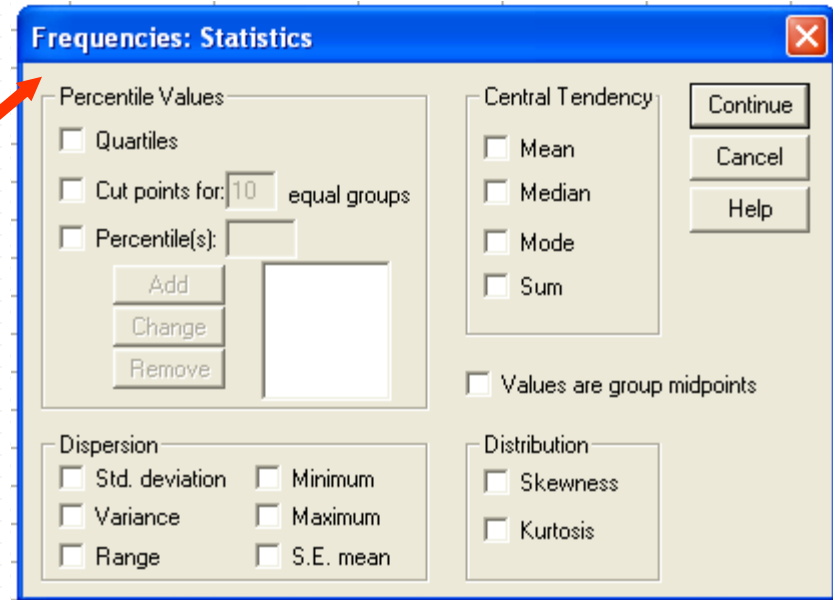
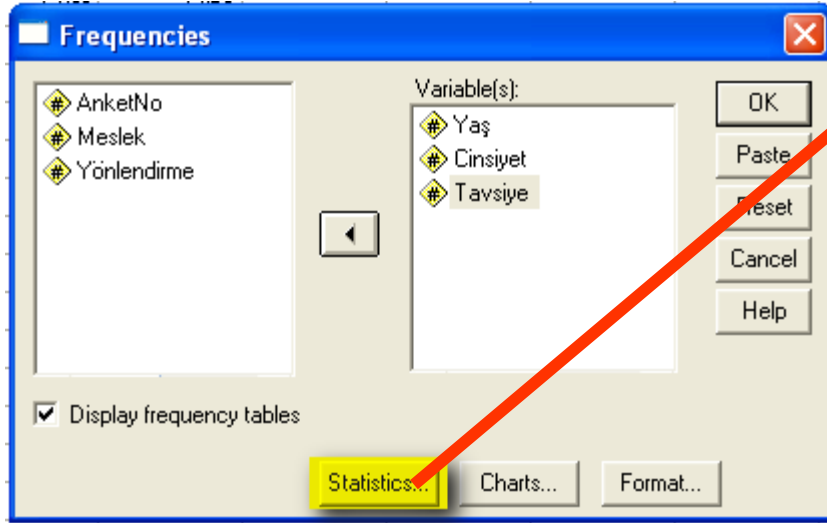
AnketNo
Yaş
Meslek
Cinsiyet
Tavsiye
Yönlendirme

Display frequency tables

OK
Paste
Reset
Cancel
Help

Statistics... Charts... Format...

Frekans Tabloları-devam



Frekans Tabloları

Frequencies

Statistics

		Kaç yapýndasýnýz?	Cinsiyetiniz nedir?	Hastanemizi babkalarýna önerir misiniz?
N	Valid	10	10	10
	Missing	0	0	0
Mean		40,50	1,50	1,5000
Std. Error of Mean		3,851	,167	,16667
Median		37,50	1,50	1,5000
Mode		34	1 ^a	1,00 ^a
Std. Deviation		12,177	,527	,52705
Variance		148,278	,278	,278
Skewness		,884	,000	,000
Std. Error of Skewness		,687	,687	,687
Kurtosis		,582	-2,571	-2,571
Std. Error of Kurtosis		1,334	1,334	1,334
Range		41	1	1,00
Minimum		24	1	1,00
Maximum		65	2	2,00
Sum		405	15	15,00

a. Multiple modes exist. The smallest value is shown

Frequency Table

Kaç yapýndasýnýz?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	24	1	10,0	10,0	10,0
	30	1	10,0	10,0	20,0
	34	2	20,0	20,0	40,0
	35	1	10,0	10,0	50,0
	40	1	10,0	10,0	60,0
	43	1	10,0	10,0	70,0
	45	1	10,0	10,0	80,0
	55	1	10,0	10,0	90,0
	65	1	10,0	10,0	100,0
Total		10	100,0	100,0	

Cinsiyetiniz nedir?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Kadýn	5	50,0	50,0	50,0
	Erkek	5	50,0	50,0	100,0
Total		10	100,0	100,0	

Hipotez Testleri

- Hipotez testleri, örnek verilerinden hareketle ana kitle parametreleri hakkındaki iddaların belirli bir hata düzeyinde araştırılması konularını içerir. Bu testlerde ,örnek birim değerleri kullanılarak hesaplanan istatistiğin değeriyle, bu istatistiğin bilgi ürettiği ana kitle parametresinin önceden bilinen değeri arasında farklılığın istatistiksel olarak anlamlı olup olmadığı belirlenir. Farklılık varsa ,bu farkın öneminin, sıfır hipotezini reddetmek için yeterli olup olmadığına karar verilir. Söz konusu farklılığın anlamlı olması durumunda sıfır hipotezi red, tersi durumda kabul edilir.
- Hipotez testlerinde daima sıfır hipotezi test edilir.

Parametrik Hipotez Testleri

- **Parametrik hipotez testlerinin varsayımları:**
 1. Veriler oransal veya aralıklı olmalıdır.
 2. Veriler normal dağılıma uymalıdır.(Basıklık ve çarpıklık değerleri -1 ve +1 arasında olmalıdır.) (K-S Testi: Non-parametrik testlerde anlatılacaktır.
 3. Grup varyansları eşit olmalıdır.

Arařtırmada Hangi Analiz Kullanılacak?

Bir arařtırma yaparken, arařtırmada hangi analizin kullanılacağına karar verebilmek için ařağıdaki üç sorunun cevaplanması gerekir.

1. Elinizde kaç tane grup var?
2. Gruplar birbirleri ile ilişkili mi (Bağımlı grup-bağımsız grup (örnek)?
 1. Bağımsız grup: İki grubun üyeleri birbirinden ayrıdır. Gruplar arasında kesinlikle ortak üye olamamalıdır. (Cinsiyet: bayan-erkek)
 2. Bağımlı grup:
3. Hangi varsayımlar karşılanmaktadır?

Arařtırmada Hangi Analiz Kullanılacak?-devam

Grup Sayısı	Grupların Durumu	Gerçeklenen Varsayımlar	Kullanılacak Test
2	Bağımsız gruplar	İlk iki varsayım karşılanıyorsa	Bağımsız örneklem t Testi
2	Bağımsız gruplar	1. Ya da 2. varsayım ihlal edilmişse	Mann-Whitney U Testi
2	Bağımlı gruplar	İlk iki varsayım karşılanıyorsa	Bağımlı (Eşlenik) örneklem t Testi
2	Bağımlı gruplar	1. Ya da 2. varsayım ihlal edilmişse	Wilcoxon Testi
2		Nominal veri kullanılmışsa	Ki Kare Testi
3 ve üzeri	Bağımsız gruplar	İlk iki varsayım karşılanıyorsa	ANOVA
3 ve üzeri	Bağımsız Gruplar	1. Ya da 2. varsayım ihlal edilmişse	Kruskal Wallis Testi

Parametrik Hipotez Testleri / t - Testi

- İki örneklem grubu arasında ortalamalar açısından fark olup olmadığını araştırmak için kullanılır.
- 1. Independent Samples t Testi: Bağımsız iki örnek t testi
- 2. Paired Samples t Test: Bağımlı iki örnek t testi
- 3. One Sample t Test: Tek örnek t testi

Bağımsız İki Örnek t Testi

- Independent –Samples t Test
- Burada bağımsız iki örnek grubunun ortalamaları karşılaştırılır.
- Örnek: 5 Likert ölçeği kullanılarak (5: Tamamen Katılıyorum, 4: Katılıyorum, 3: Kararsızım, 2: Katılmıyorum, 1: Hiç Katılmıyorum) yapılan bir anket çalışmasında katılımcılardan çalıştıkları kurumun kendileri için bir prestij kaynağı olup olmadığını belirtmeleri istenmiştir. Katılımcılar erkek ve kadın olarak iki gruba ayrılarak, sorulan bu soru ile ilgili yaptıkları yorumlar karşılaştırılmak istenmektedir. Cinsiyete göre yapılan yorumların farklı olup olmadığını test ediniz. ($\alpha=0,05$)

Cinsiyet	1	2	1	2	1	1	2	1	2	2	1	1	2	1	1
Yorum	3	4	3	4	3	4	4	1	4	4	3	3	5	4	3
Cinsiyet	2	2	2	1	1	1	2	1	2	2	2	2	2	1	2
Yorum	4	4	5	2	3	2	3	3	4	5	4	5	4	3	4

Örnek-devam

Hipotezler

- H_0 : İki grubun ortalamaları arasında fark yoktur
- H_1 : İki grubun ortalamaları arasında fark vardır

Örnek-devam

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

12 :

	Cinsiyet	Yorum
1	Erkek	Kararsızdır
2	Kadın	Katılıyor
3	Erkek	Kararsızdır
4	Kadın	Katılıyor
5	Erkek	Kararsızdır
6	Erkek	Katılıyor

Independent-Samples T Test

Cinsiyet
Yorum

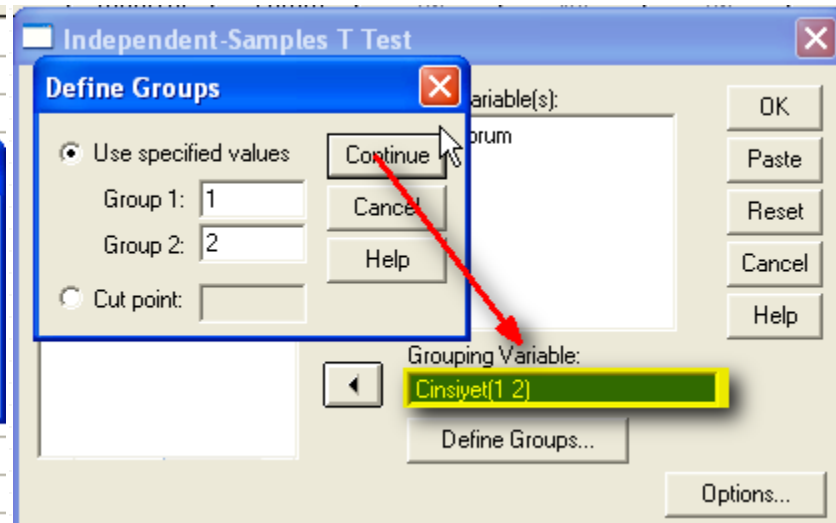
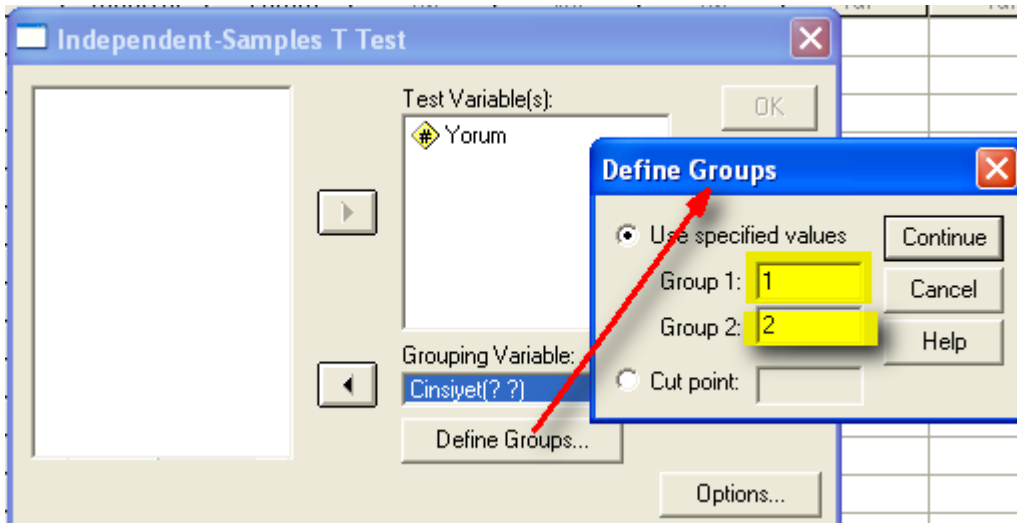
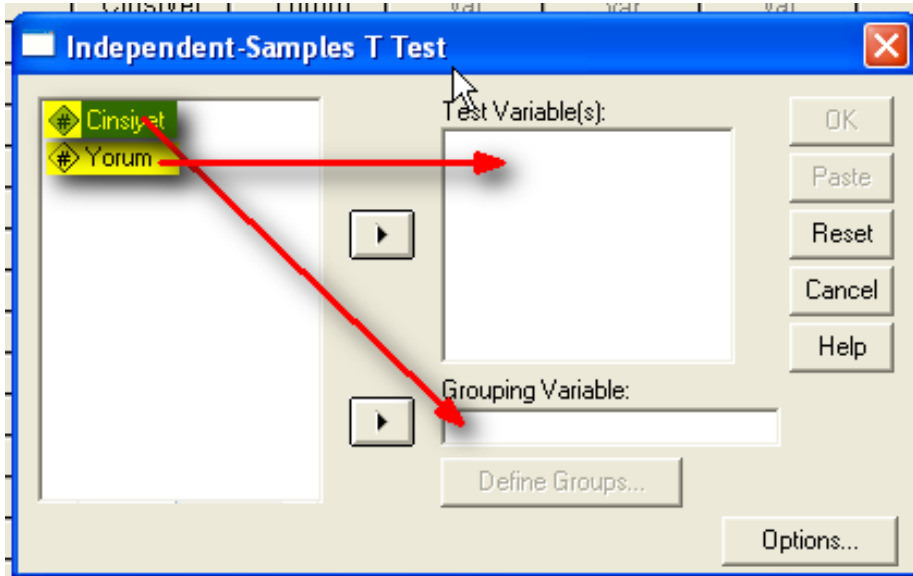
OK
Paste
Reset
Cancel
Help

Grouping Variable:
Define Groups...
Options...

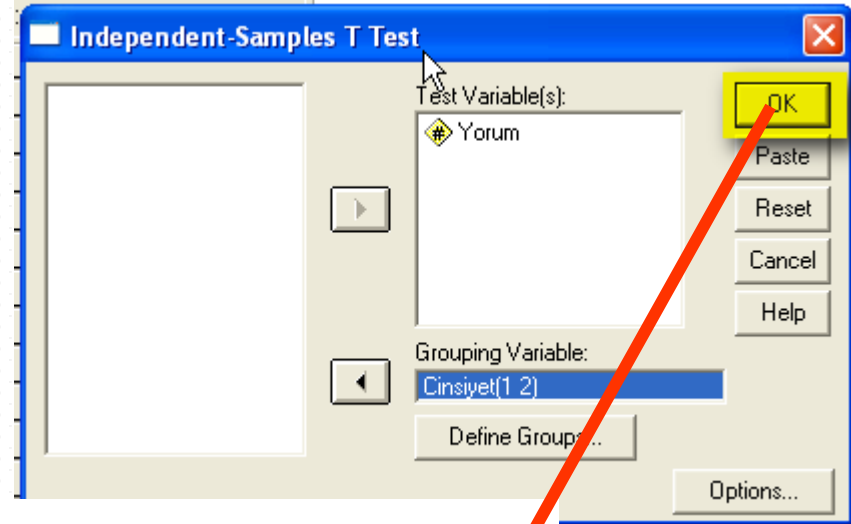
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Missing Value Analysis...
Complex Samples

Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...

Örnek-devam



Örnek-devam



► T-Test

Group Statistics

	Cinsiyet	N	Mean	Std. Deviation	Std. Error Mean
Yorum	Erkek	14	2,8571	,77033	,20588
	Kadıym	16	4,1875	,54391	,13598

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Yorum	Equal variances assumed	,385	,540	-5,518	28	,000	-1,33036	,24109	-1,82421	-,83651
	Equal variances not assumed			-5,392	23,019	,000	-1,33036	,24673	-1,84073	-,81998

Örnek-devam

Karar Kriteri

$$p < \alpha=0,05 \Rightarrow H_0 \text{ red}$$

Karar

$$0,00 < \alpha=0,05 H_0 \text{ red}$$

Yorum: Kadın ve Erkekler çalıştıkları kurumun kendilerine prestij kaynağı olması konusunda farklı düşündükleri ve kadınların çalıştıkları kurumu kendileri için bir prestij kaynağı olarak gördükleri yorumu yapılabilir.

Örnek-devam

Analiz sonuçlarında Levene's Test for Equality of Variances kısmının yorumunun Equal Variance Assumed ve Equal Variance Not Assumed alanlarına göre yapılması gerekir. Buradaki Equal Variance Assumed kısmındaki p değeri (sig.) 0,540 olduğundan varyansların eşit olduğunu işaret etmektedir. Bu durumda “Equal varvances assumed”ın ver aldığı kısımdaki p değerine

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Yorum					
Equal variances assumed	,385	,540	5,518	28	,000
Equal variances not assumed			-5,392	23,019	,000

Bağımlı İki Örnek t Testi

- Paired Samples t Test

Bağımlı iki örnek t testinde de yine ortalamaları karşılaştırılmaktadır. Ancak burada iki ayrı örneklem grubu yoktur. Aynı örneklem grubu üzerinde analizler yapılır. (aynı grubun farklı zaman dilimlerindeki beklentilerini, başarılarını, hızlarını vb. ölçülür.)

Örnek: Üniversite öğrencilerinin vize ve final notları arasındaki başarı durumunu ölçmek isteyen bir öğretim elemanı 15 kişilik öğrenci grubunun vize ve final notları ortalamaları arasında fark olup olmadığını araştırmak istemektedir.

Vize	45	67	60	55	48	62	48	63	72	50
Final	75	73	85	72	56	73	76	80	95	82
Vize	77	81	56	45	68					
Final	92	90	70	60	87					

Örnek-devam

SPSS Data Editor

File Edit View Transform Analyze Graphs Utilities Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction

Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...

	Final
10	75,0
10	73,0
10	85,0
10	72,0
10	56,0
10	73,0

Paired-Samples T Test

Paired Variables:
Vize -- Final

Current Selections
Variable 1:
Variable 2:

OK
Paste
Reset
Cancel
Help
Options...

Örnek-devam

- Hipotezler
 - H_0 : Vize ve final notlarının ortalamaları arasında fark yoktur
 - H_1 : Vize ve final notlarının ortalamaları arasında fark vardır

Örnek-devam

T-Test

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean	
Pair 1	Vize	59,8000	15	11,58324	2,99078
	Final	77,7333	15	11,17693	2,88587

Paired Samples Correlations

	N	Correlation	Sig.	
Pair 1	Vize & Final	15	,745	,001

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Vize - Final	-17,93333	8,12814	2,09868	-22,43455	-13,43212	-8,545	14	,000

Örnek-devam

- Karar

$p < \alpha=0,05$ olduğu için H_0 red.

- Yorum

Öğrecilerin vize ve final notlarının ortalmaları arasında fark olduğunu %95 söyleyebiliriz.

Tek Örnek t Testi

Tek örnek t testi herhangi bir örneklem grubuna ait ortalamanın, daha önceden belirlenmiş bir değerden önemli derecede farklı olup olmadığını belirlemek için kullanılır. Analizi yapacak kişinin grup ortalamasına ilişkin belirlediği veya istediği değerle grubun ortalaması karşılaştırılır. (örneğin performans değerlendirmeleri, bir gruba ait başarı oranının belirlenmesi, sporcuların beklenenin altında veya üstünde efor göstermeleri vs.)

Örnek: Yukarıda verilen ikinci örnekle ilgili dersin hocası öğrencilerin final ortalamasının 90 olup olmadığını test etmektedir.

Final	75	73	85	72	56	73	76	80	95	82
Final	92	90	70	60	87					

Örnek-devam

- Hipotezler
 - H_0 : Final not ortalaması 90'a eşittir.
 - H_1 : Final not ortalaması 90'a eşit değildir.

Örnek-devam

t testi12.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

16 : Final

	Vize	Final
1	45,00	75,0
2	67,00	73,0
3	60,00	85,0
4	55,00	72,0
5	48,00	56,0

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify

Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...

One-Sample T Test

Test Variable(s):
Final

Test Value: 90

OK
Paste
Reset
Cancel
Help
Options...

Örnek-devam

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Final	15	77,7333	11,17693	2,88587

One-Sample Test

	Test Value = 90					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Final	-4,251	14	,001	-12,28667	-18,4562	-6,0771

Karar

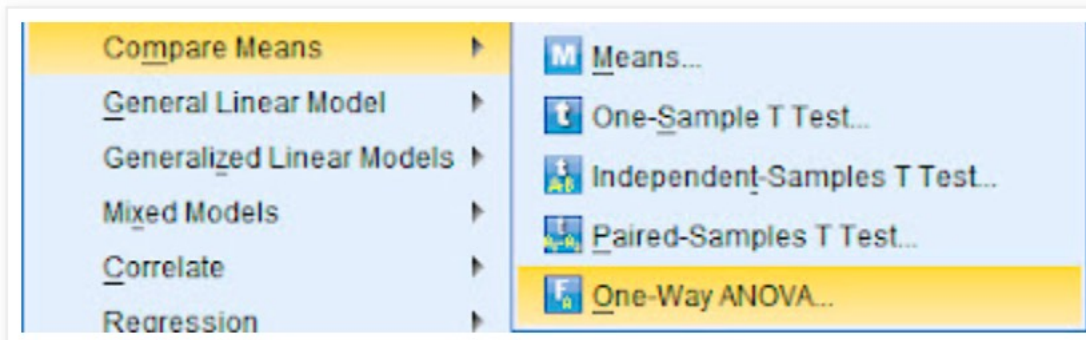
$p < \alpha$ H_0 red

Yorum: Öğrencilerin final not ortalaması 90'dan farklıdır.

Öğrencilerin final not ortalaması 77,73'tür. © 27 Şubat 2023 Pazartesi

Tek Yönlü Varyans Analizi

- Eğer grup değişkenindeki test değişkeni normal dağılıyorsa, grup değişkenindeki grup sayısı 2'den fazla ise ve gruplar bağımsızsa:
 - Böyle bir durumda grup ortalamaları arasında fark olup olmadığına ilişkin karar alabilmek için tek yönlü varyans analizi kullanılır.
 - Tek yönlü varyans analizi için SPSS'te Analyz/Compare Means/One Way Anova/ komut bileşeni kullanılır.

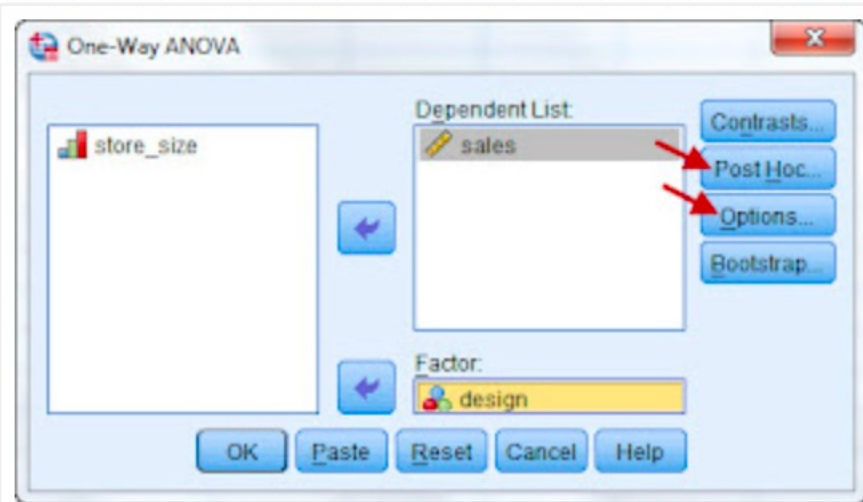
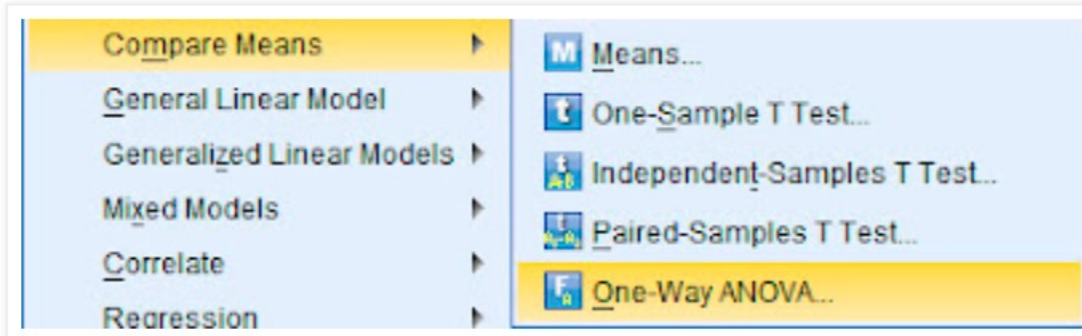


Örnek

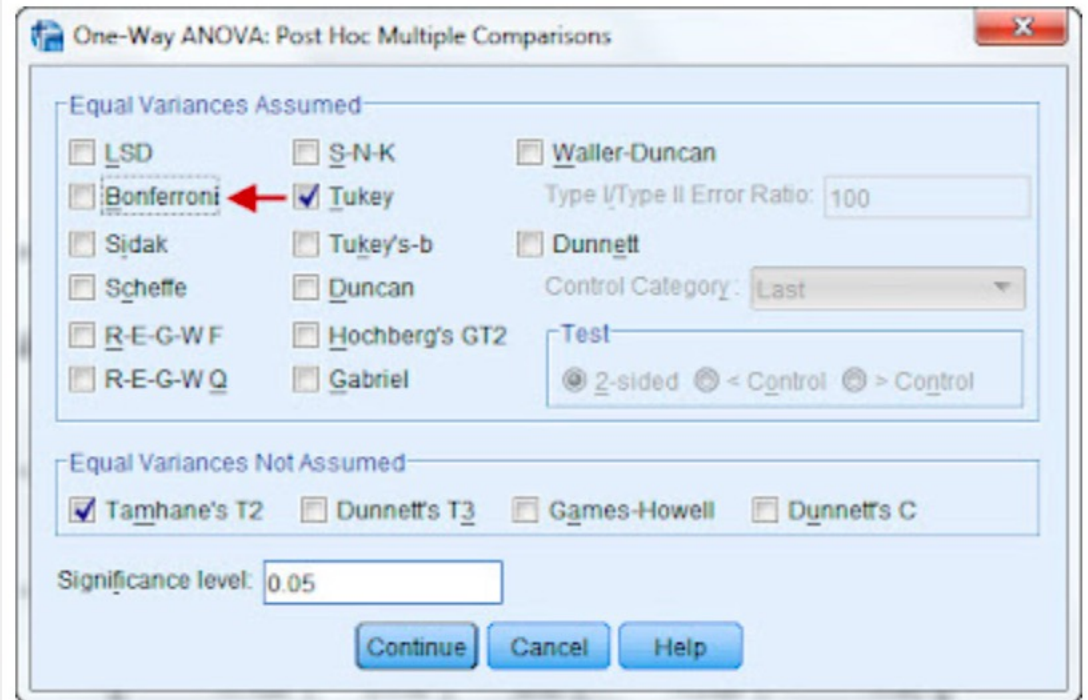
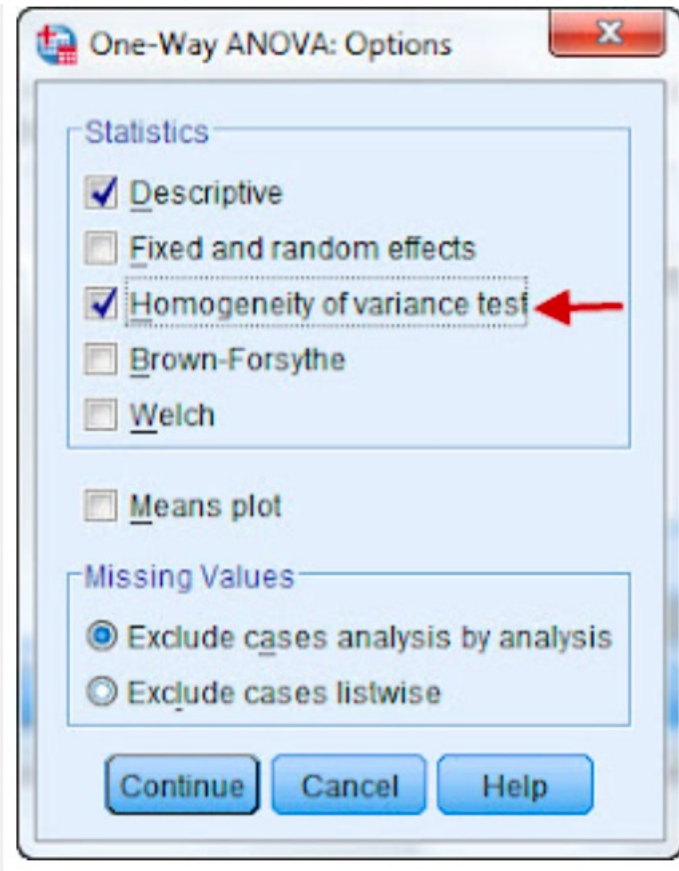
Varyansların homojenliği testi için hipotezler;
H0: %95 güvenle, grup varyansları homojendir.
H1: %95 güvenle, grup varyansları homojen değildir.

Tek yönlü varyans analizi için hipotezler;
H0: %95 güvenle, grupların ortalamaları arasında istatistiksel olarak anlamlı bir farklılık yoktur.
H1: %95 güvenle, grupların ortalamaları arasında istatistiksel olarak anlamlı bir farklılık vardır.

sales	design
11	1
17	1
16	1
14	1
15	1
12	2
10	2
15	2
19	2
11	2
23	3
20	3
18	3
17	3
27	4
33	4
22	4
26	4
28	4



Örnek



Örnek

Test of Homogeneity of Variances

sales

Levene Statistic	df1	df2	Sig.
.438	3	15	.729

ANOVA

sales

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	588.221	3	196.074	18.591	.000
Within Groups	158.200	15	10.547		
Total	746.421	18			

Örnek

Multiple Comparisons

Dependent Variable: sales

	(I) design	(J) design	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1	2	1.200	2.054	.935	-4.72	7.12
		3	-4.900	2.179	.155	-11.18	1.38
		4	-12.600*	2.054	.000	-18.52	-6.68
	2	1	-1.200	2.054	.935	-7.12	4.72
		3	-6.100	2.179	.058	-12.38	.18
		4	-13.800*	2.054	.000	-19.72	-7.88
	3	1	4.900	2.179	.155	-1.38	11.18
		2	6.100	2.179	.058	-.18	12.38
		4	-7.700	2.179	.014	-13.98	-1.42
	4	1	12.600*	2.054	.000	6.68	18.52
		2	13.800*	2.054	.000	7.88	19.72
		3	7.700*	2.179	.014	1.42	13.98
Tamhane	1	2	1.200	1.929	.992	-5.87	8.27
		3	-4.900	1.676	.147	-11.31	1.51
		4	-12.600*	2.049	.004	-20.26	-4.94
	2	1	-1.200	1.929	.992	-8.27	5.87
		3	-6.100	2.100	.130	-13.71	1.51
		4	-13.800*	2.408	.003	-22.16	-5.44
	3	1	4.900	1.676	.147	-1.51	11.31
		2	6.100	2.100	.130	-1.51	13.71
		4	-7.700	2.211	.062	-15.76	.36
	4	1	12.600*	2.049	.004	4.94	20.26
		2	13.800*	2.408	.003	5.44	22.16
		3	7.700	2.211	.062	-.36	15.76

* The mean difference is significant at the 0.05 level.

İSTATİSTİK PAKET PROGRAMLARI

SPSS Uygulamaları

NON PRAMETRİK TESTLER

Non Parametrik Testler

- Non parametrik testler, parametrik testlere alternatif testlerdir. Her parametrik testin alternatifi olan bir non parametrik test vardır. Uygulanması düşünülen herhangi bir parametrik test için koşulların sağlanamadığı durumlarda onun yerine alternatifi olan non arametrik test uygulanır. En çok kullanılanları birer örnekle anlatılacaktır.

Tek Örneklem Kolmogorov-Smirnov Testi

- Gözlenen bir frekans dağılımının kuramsal bir dağılıma uyup uymadığını test etmek için kullanılır. Ki kare testi için koşulların sağlanmadığı durumlarda bu test uygulanır.

Örnek 1

- Bir toplumdaki seçilen 4 çocuklu 100 ailenin sahip oldukları kız çocuk sayıları aşağıda verilmiştir.
- Kız çocuğu Sayısı 0 1 2 3 4
- Aile Sayısı 4 26 38 28 4
- H_0 : Örnekte gözlenen frekans dağılımı, normal dağılımına uyuyor.
- H_1 : Örnekte gözlenen frekans dağılımı, normal dağılımına uymaz.
- Yukarıdaki tablo yardımıyla hipotezleri test edelim. Ailelerin sahip oldukları kız çocuklarının sayısına ilişkin veriler, **Variable View** penceresinde “ç_say” isimli değişken olarak tanımlandıktan sonra, **Data View** penceresinde veri girişi aşağıdaki gibi yapılır. 4 ailenin sahip oldukları kız çocuğu sayısı 0 olduğu için 4 tane 0 değeri, 26 ailenin sahip oldukları kız çocuğu sayısı 1 olduğu için 26 tane 1 değeri, 38 ailenin sahip oldukları kız çocuğu sayısı 2 olduğu için 38 tane 2 değeri, 28 ailenin sahip oldukları kız çocuğu sayısı 3 olduğu için 28 tane 3 değeri, 4 ailenin sahip oldukları kız çocuğu sayısı 4 olduğu için ise 4 tane 4 değeri **Data View** penceresinde tek sütun halinde girilir.

Örnek 1-devam

Untitled - SPSS Da	
File Edit View Data	
12:	
	ç_say
1	0
2	0
3	0
4	0
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1

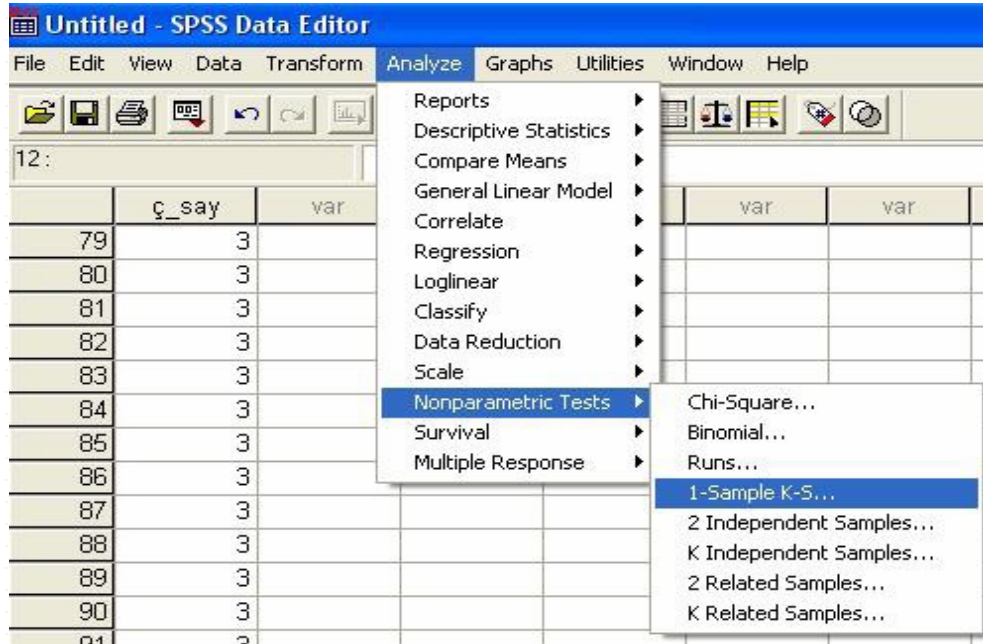
31	2
32	2
33	2
34	2
35	2
36	2
37	2
38	2
39	2
40	2
41	2
42	2
43	2
44	2
45	2
46	2
47	2
48	2
49	2
50	2
51	2
52	2
53	2
54	2

55	2
56	2
57	2
58	2
59	2
60	2
61	2
62	2
63	2
64	2
65	2
66	2
67	2
68	2
69	3
70	3
71	3
72	3
73	3
74	3
75	3
76	3
77	3
78	3


79	3
80	3
81	3
82	3
83	3
84	3
85	3
86	3
87	3
88	3
89	3
90	3
91	3
92	3
93	3
94	3
95	3
96	3
97	4
98	4
99	4
100	4

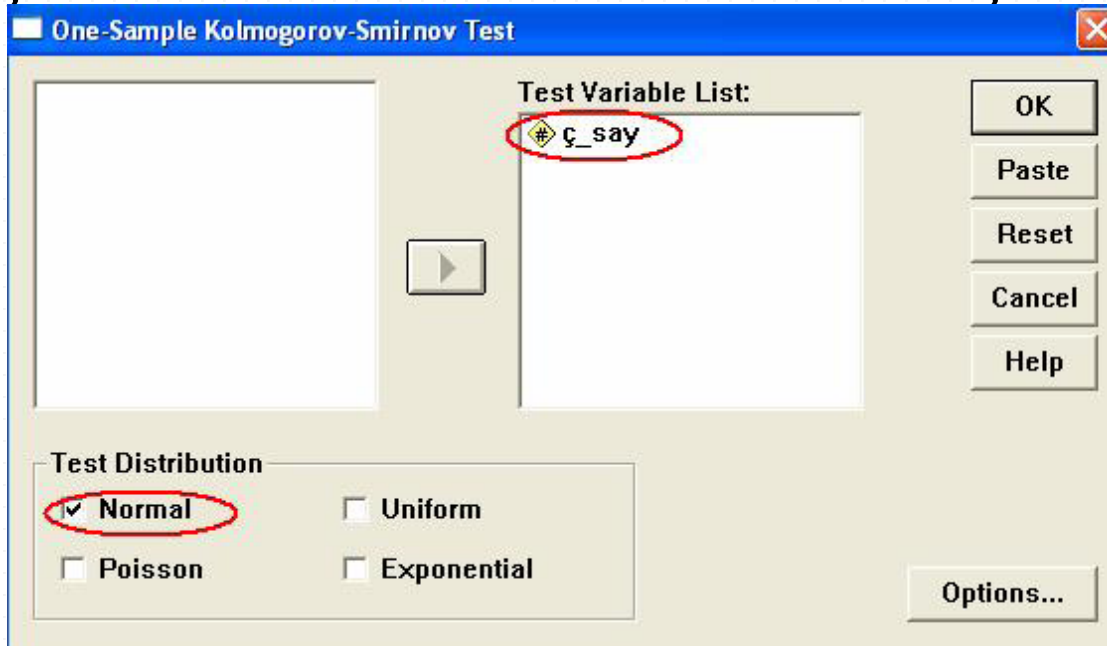
Çözüm 1-devam

- Örneğimizde, ilgili hipotezleri test etmek için “Tek Örneklem Kolmogorov Smirnov Testi” kullanılacaktır. SPSS’de Tek Örneklem Kolmogorov Smirnov Testi’ni uygulamak için **Analyze menüsünün** bir alt menüsü olan **Nonparametrics**’den, **1-Sample K-S** seçilir.



Çözüm 1-devam

- Ekrana gelecek aşağıdaki pencerede, “ç_say” değişkeni **Test Variable List** bölümüne  simgesi tıklanarak aktarılır. **Test Distribution** alanından istenilen dağılım işaretlenir ve **OK** tıklanarak sonuçlara ulaşılır.



Çözüm 1-devam

- İlgili SPSS çıktısı ve yorumu aşağıdaki gibidir.
- **NPar Tests**

One-Sample Kolmogorov-Smirnov Test

		Ç SAY
N		100
Normal Parameters ^{a,b}	Mean	2,02
	Std. Deviation	,93
Most Extreme Differences	Absolute	,191
	Positive	,189
	Negative	-,191
Kolmogorov-Smirnov Z		1,914
Asymp. Sig. (2-tailed)		,001

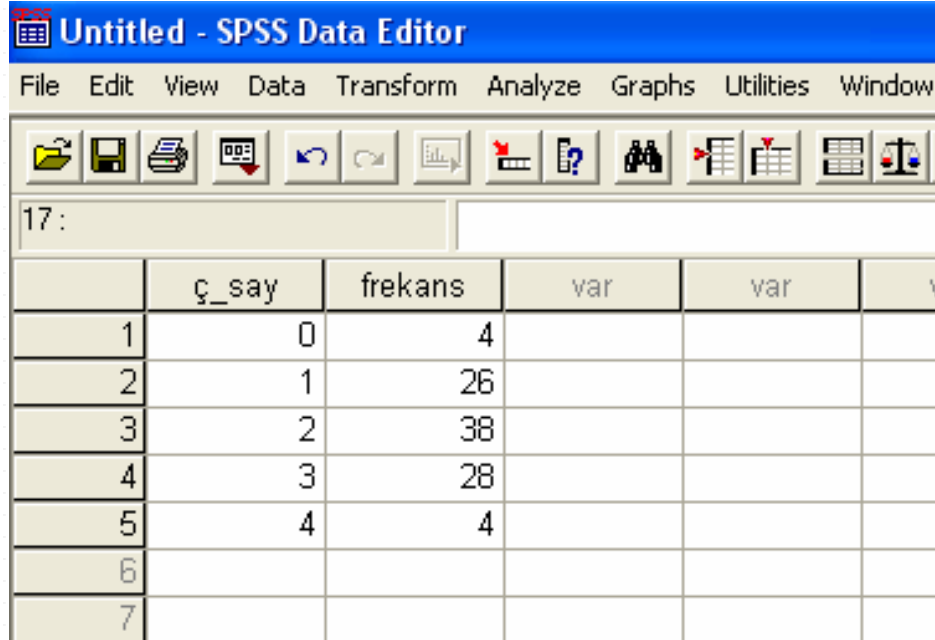
a. Test distribution is Normal.

b. Calculated from data.

Çözüm 1-devam

- Uygulanan test sonucunda p değeri 0,001 olarak bulunmuştur. $p < 0,05$ olduğu için H_0 hipotezi reddedilir. Buna göre 0,05 hata düzeyinde, sahip olunan kız çocuk sayısına göre dağılım, Normal dağılımdan farklıdır.
- SPSS’de Tek Örneklem Kolmogorov Smirnov Testi uygulanacağı zaman, verilerin **Variable View** penceresine girişleri ikinci bir yolla da yapılabilir. Ailelerin sahip oldukları çocuk sayılarına ilişkin veriler “ç_say” isimli değişken olarak, her bir çocuk sayısına karşılık gelen aile sayıları ise “frekans” isimli değişken olarak tanımlandıktan sonra, **Data View** penceresinde veri girişi aşağıdaki gibi yapılır.

Çözüm 1-devam

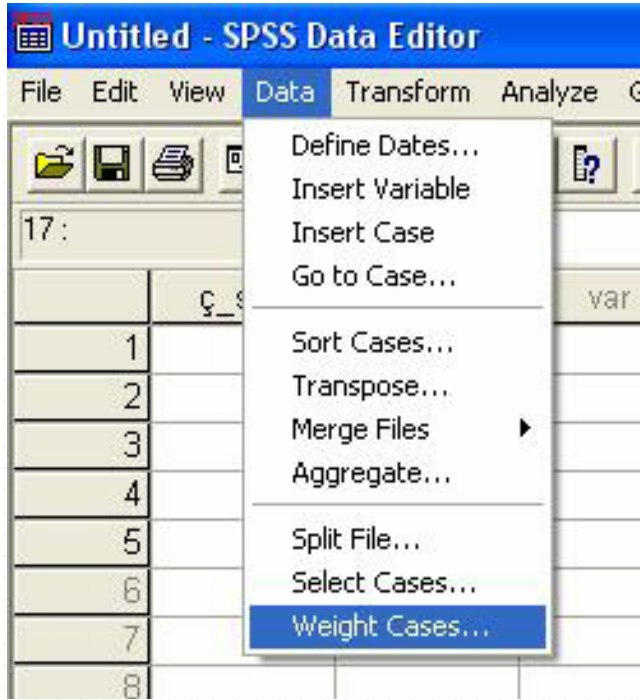



The screenshot shows the SPSS Data Editor interface. The title bar reads 'Untitled - SPSS Data Editor'. The menu bar includes 'File', 'Edit', 'View', 'Data', 'Transform', 'Analyze', 'Graphs', 'Utilities', and 'Window'. The toolbar contains various icons for file operations and data manipulation. Below the toolbar, a text box contains the number '17:'. The main data grid has the following columns: 'ç_say', 'frekans', 'var', 'var', and 'v'. The data rows are as follows:

	ç_say	frekans	var	var	v
1	0	4			
2	1	26			
3	2	38			
4	3	28			
5	4	4			
6					
7					

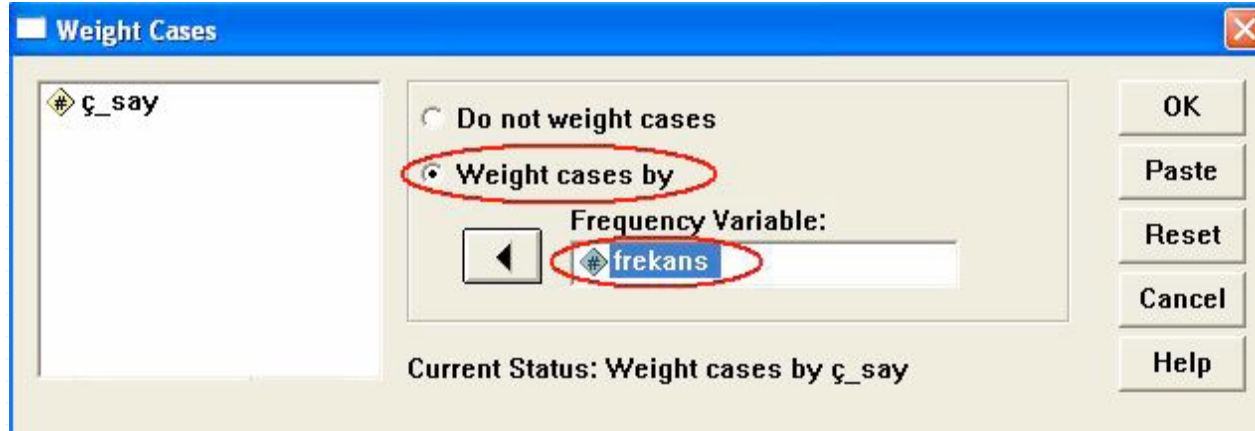
- Verilerin frekanslarını gösteren sütunun birim veriler olduğunun SPSS'e belirtilmesi için **Data** menüsünün bir alt menüsü olan **Weight Cases** seçilir.

Çözüm 1-devam



Ekranı gelecek aşağıdaki pencerede, **Weight cases by** seçeneği işaretlendikten sonra, “frekans” değişkeni **Frequency Variable** bölümüne  simgesi tıklanarak aktarılır ve **OK** tıklanır.

Çözüm 1-devam



Daha sonra, SPSS’de Tek Örneklem Kolmogorov Smirnov Testi’ni uygulamak için, ilk olarak gösterilen veri girişi sonrasında olduğu gibi, **Analyze menüsünün** bir alt menüsü olan **Nonparametrics**’den, **1-Sample K-S** seçilir.

Çözüm 1-devam

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

12:

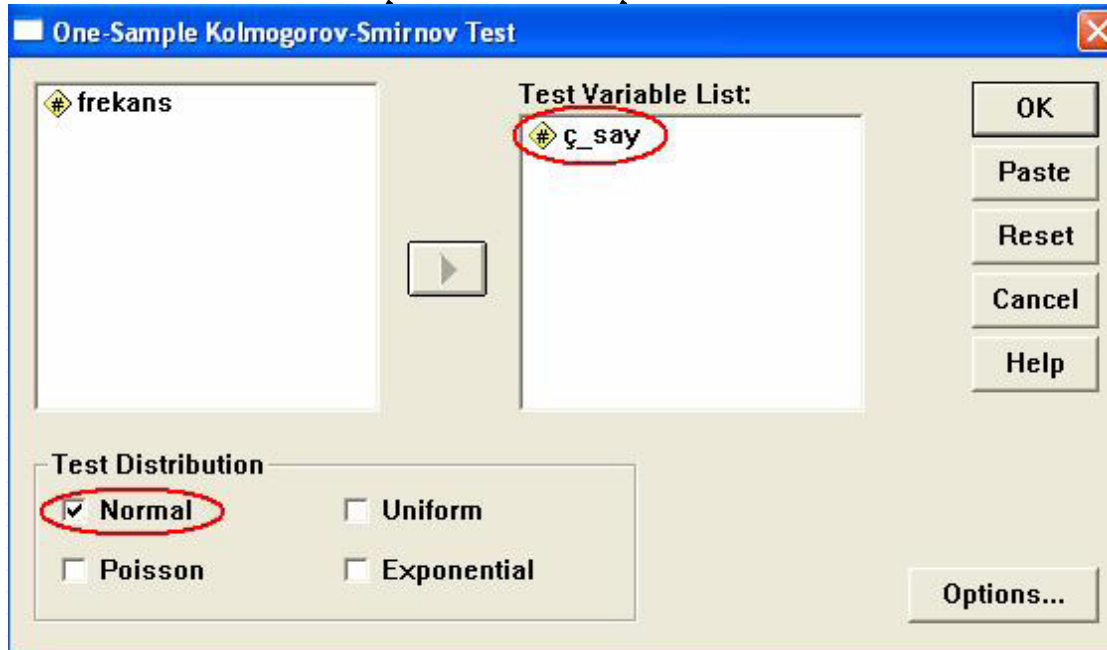
	ç_say	var
79	3	
80	3	
81	3	
82	3	
83	3	
84	3	
85	3	
86	3	
87	3	
88	3	
89	3	
90	3	
91	3	

Reports
Descriptive Statistics
Compare Means
General Linear Model
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Survival
Multiple Response

Chi-Square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...

Çözüm 1-devam

- Ekranı gelecek aşağıdaki pencerede, “ç_say” değişkeni **Test Variable List** bölümüne simgesi tıklanarak aktarılır. **Test Distribution** alanından istenilen dağılım işaretlenir ve **OK** tıklanarak sonuçlara ulaşılır.



- Bu yol kullanılarak veri girişi yapıldıktan sonra elde edilecek SPSS çıktısı, ilk olarak gösterilen yolla elde edilenle aynı olacaktır.

İşaret Testi

- Bir toplumdaki seçilen n birimlik örneklemin, ortanca değeri M_0 olan bir toplumdaki gelip gelmediğinin test edilmesi durumunda kullanılır. İşaret testi, n birimlik bir örnekleme, ortanca değerin altında ve üstünde olan değerlerin gözlenme sıklığını Binom dağılımını kullanarak test eder. Eşleştirilmiş işaret testinde, birbirine karşılık gelen gözlemlerin farkları pozitif ya da negatiftir. Testte, pozitif işaret elde etme olasılığının negatif işaret elde etme olasılığına eşitliği hipotezi test edilir. Eşleştirilmiş işaret testi, eşleştirilmiş örneklem t testinin non parametrik alternatifidir.

Örnek 2

- Fen Bilimleri eğitimi alan bireyler ile Sosyal Bilimler eğitimi alan bireylerin toplumsal sorunlara eğilimleri arasında farklılık bulunduğu ve Sosyal Bilimler eğitimi alan bireylerin toplumsal sorunlara daha fazla ilgi duydukları savı ileri sürülmektedir. Bu savı denetlemek amacıyla toplumdan ikiz olarak doğan ve ikizlerden birinin Fen Bilimleri eğitimi aldığı, diğerlerinin ise Sosyal Bilimler eğitimi aldığı 12 çift seçiliyor. Bu çiftlerin sosyal sorunlara bakış açılarını değerlendiren bir test yardımı ile sosyal sorunları değerlendirme puanları belirleniyor. Bulgular aşağıdaki tabloda verilmiştir. Fen Bilimleri eğitimi ile Sosyal Bilimler eğitiminin bakış açısını etkileyip etkilemediğini $\alpha=0,05$ anlamlılık düzeyinde kontrol edelim.

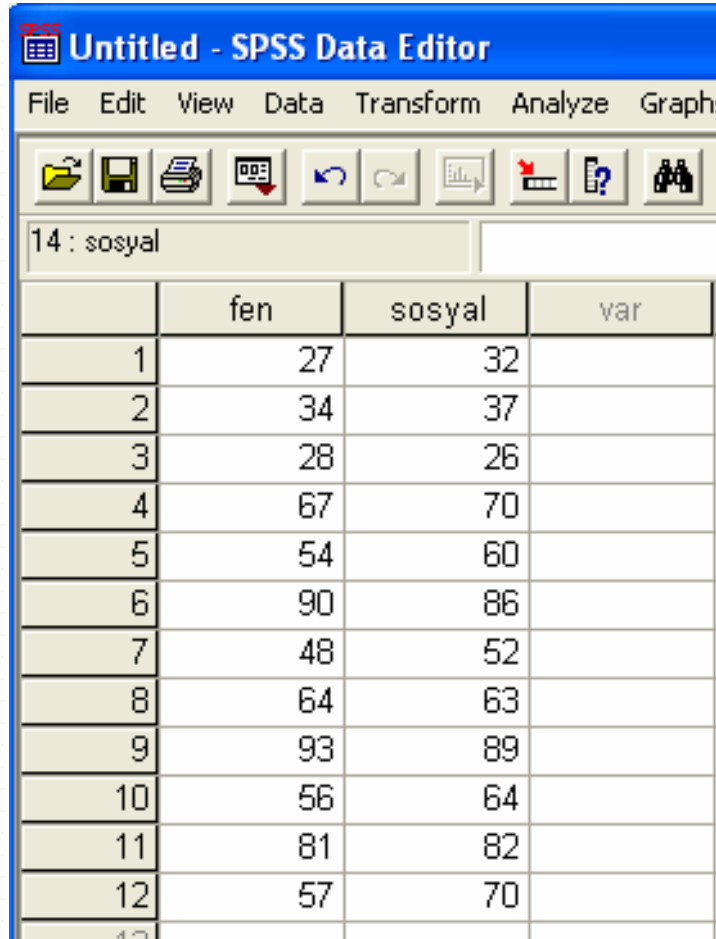
İkiz no	1	2	3	4	5	6	7	8	9	10	11	12
Fen Bilim	27	34	28	67	54	90	48	64	93	56	81	57
Sosyal Bilim	32	37	26	70	60	86	52	63	89	64	82	70

Çözüm 2

- Hipotezlerin Kurulması
- H_0 : Eğitim türünün bireylerin toplumsal sorunlara eğilimleri üzerinde bir etkisi yoktur.
- H_1 : Eğitim türünün bireylerin toplumsal sorunlara eğilimleri üzerinde bir etkisi vardır.

Çözüm 2-devam

- Bireylerin test puanlarına ilişkin veriler, **Variable View** penceresinde “fen” ve “sosyal” isimli iki ayrı değişken olarak tanımlandıktan sonra, **Data View** penceresinde veri girişi aşağıdaki gibi yapılır.

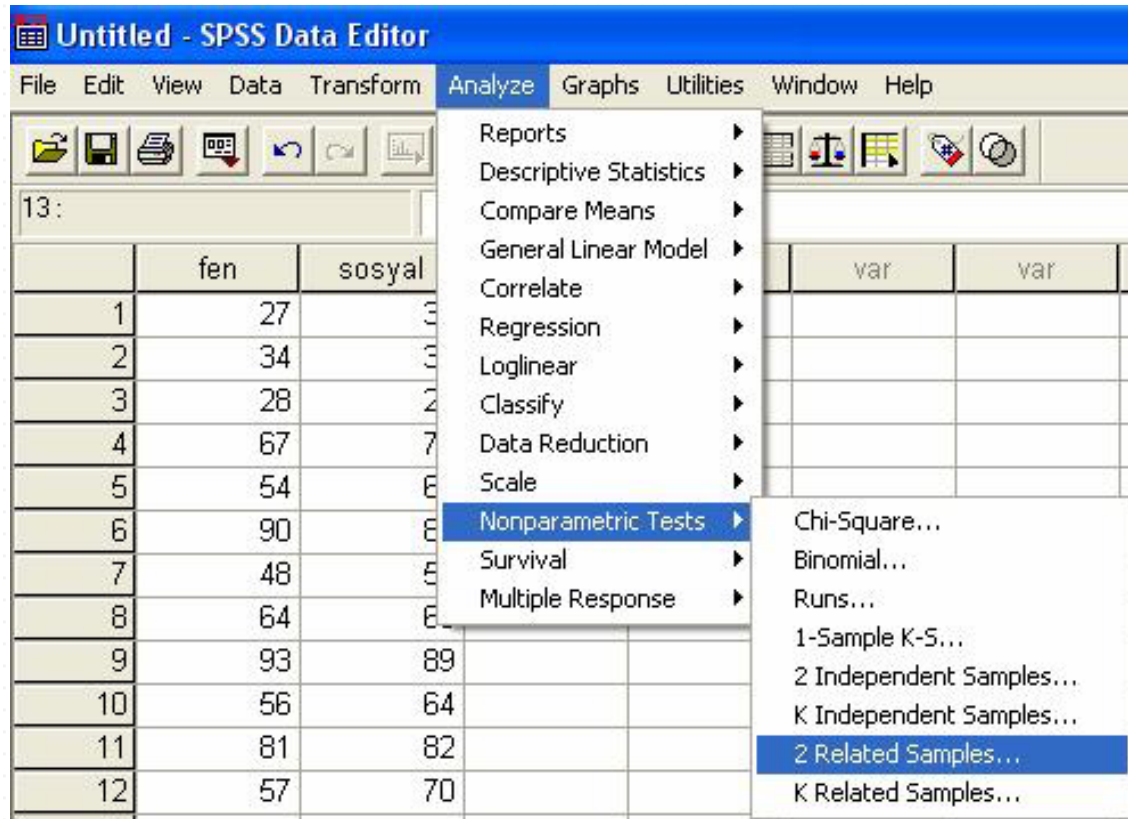


The screenshot shows the SPSS Data Editor window titled "Untitled - SPSS Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, and Graphs. The toolbar contains icons for file operations, undo, redo, and help. The active window is labeled "14 : sosyal". The data table below has four columns: an unlabeled column for case numbers, "fen", "sosyal", and "var". The data is as follows:

	fen	sosyal	var
1	27	32	
2	34	37	
3	28	26	
4	67	70	
5	54	60	
6	90	86	
7	48	52	
8	64	63	
9	93	89	
10	56	64	
11	81	82	
12	57	70	

Çözüm 2 devam


- Örneğimizdeki ilgili hipotezi test etmek için istatistiksel test olarak “Eşleştirilmiş İşaret Testi” kullanılmalıdır. SPSS’de tek örneklem işaret testi uygulaması bulunmamaktadır. Eşleştirilmiş örneklerde işaret testi uygulanmaktadır. Bunu gerçekleştirmek için, **Analyze** menüsünün alt menüsü olan **Nonparametrics**’den **2 Related Samples** seçilir.

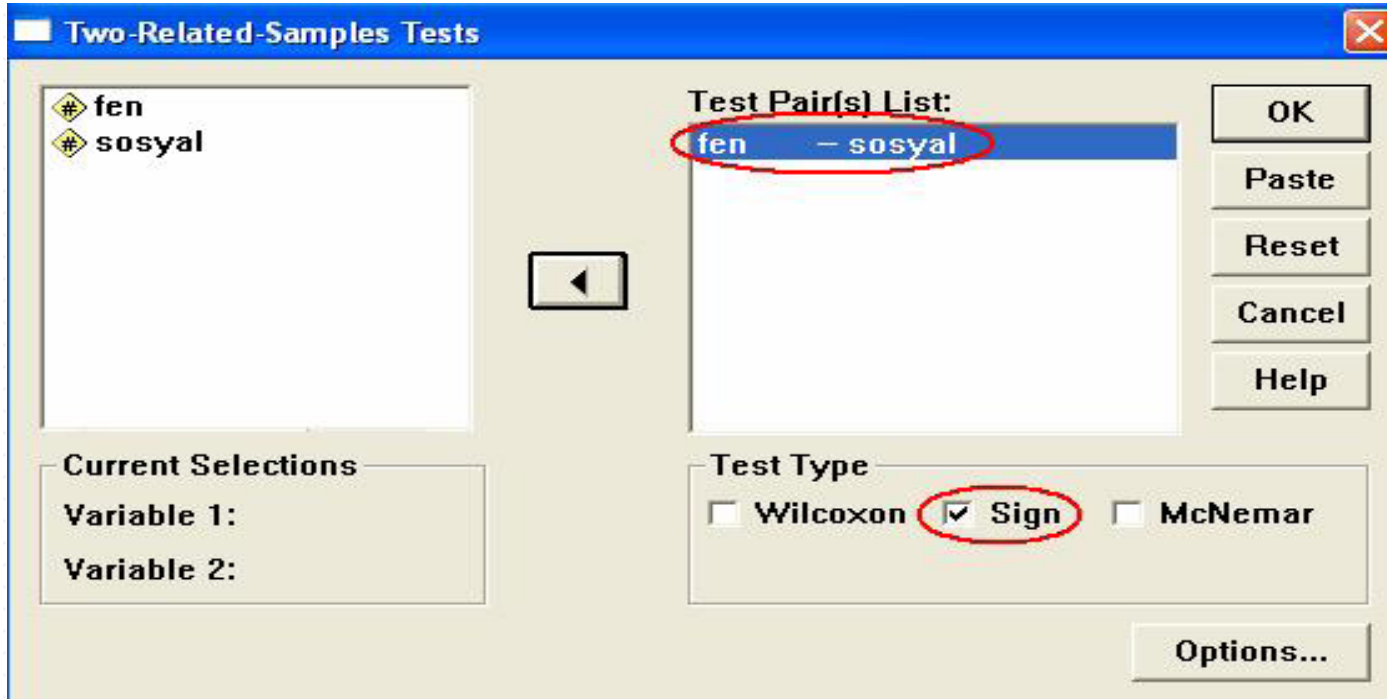


The screenshot shows the SPSS Data Editor interface with the 'Analyze' menu open. The 'Nonparametric Tests' option is selected, and the '2 Related Samples...' option is highlighted in the submenu. The data table below shows the variables 'fen' and 'sosyal' with 12 rows of data.

	fen	sosyal
1	27	3
2	34	3
3	28	2
4	67	7
5	54	8
6	90	8
7	48	8
8	64	8
9	93	89
10	56	64
11	81	82
12	57	70

Çözüm 2-devam

- Ekrana gelecek aşağıdaki pencerede **Variable 1** için “fen”, **Variable 2** için ise “sosyal” seçilir. Daha sonra bu değişken çifti,  simgesi tıklanarak **Test Pair(s) List** kısmına aktarılır. **Test Type** alanındaki **Sign** seçeneği işaretlenir ve **OK** tıklanır.



Çözüm-2 devam

- İlgili SPSS çıktısı aşağıdaki gibidir.

Sign Test

Frequencies

	N
SOSYAL - FEN Negative Differences ^a	4
Positive Differences ^b	8
Ties ^c	0
Total	12

a. SOSYAL < FEN

b. SOSYAL > FEN

c. FEN = SOSYAL

Test Statistics^d

	SOSYAL - FEN
Exact Sig. (2-tailed)	,388 ^a

a. Binomial distribution used.

b. Sign Test

- Uygulanan test sonucunda p değeri 0,388 olarak bulunmuştur. $p > 0,05$ olduğu için H_0 hipotezi kabul edilir. Buna göre $\alpha = 0,05$ için, eğitim türünün bireylerin toplumsal sorunlara eğilimleri üzerinde bir etkisi olmadığı saptanmıştır.

Wilcoxon Sıra Toplamı Testi

- Bağımlı iki örneklem testidir. Eşleştirilmiş t testinin non parametrik alternatifidir. Wilcoxon testi, iki A ve B örneğinin çiftleştirilmiş farkları dikkate alınarak yapılır. İşaret testi yalnız farkların yönü hakkındaki bilgiyi kullanırken, Wilcoxon testi farkların yönü kadar onların miktarı hakkındaki bilgiyi de kullandığı için daha kuvvetli bir testtir.

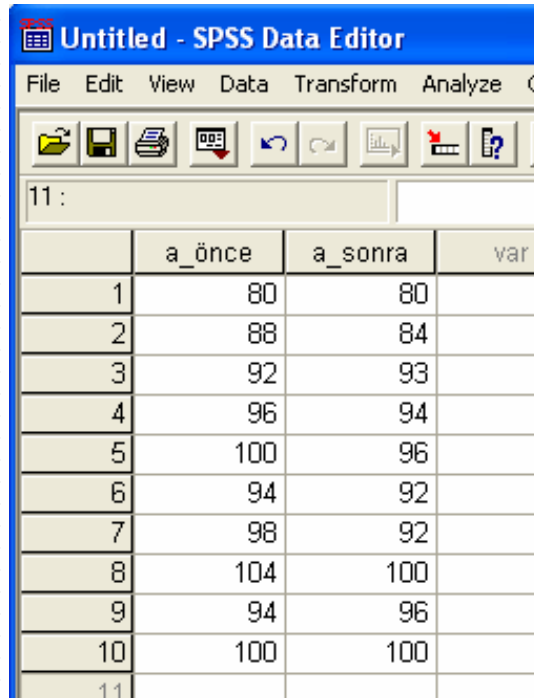
Örnek 3

- Pentothal ile indüksiyon yapılan hastalarda anestezi öncesi ve anestezi sırasında (10'ncu dakikada) elde edilen nabız sayıları aşağıda verilmiştir. t testi için varsayımların gerçekleşmediğini göz önüne alarak, “Anestezi öncesi ve sonraki nabız sayıları arasında fark yoktur” hipotezini $\alpha = 0,05$ anlamlılık düzeyinde kontrol edelim.

Nabız Sayıları	
Anestezi öncesi	Anestezi sonra
80	80
88	84
92	93
96	94
100	96
94	92
98	92
104	100
94	96
100	100

Çözüm 3

- H_0 : Anesteziden önceki ve sonraki nabız sayıları arasında fark yoktur.
- H_1 : Anesteziden önceki ve sonraki nabız sayıları farklıdır.
- Burada, aynı hastalara ait iki farklı zamanda ölçüm yapılmıştır (“önce-sonra karşılaştırması”). Hastaların anesteziden önce ve sonraki nabız sayılarına ilişkin veriler, Variable View penceresinde “a_önce” ve “a_sonra” isimli iki ayrı değişken olarak tanımlandıktan sonra Data View penceresinde veri girişi aşağıdaki gibi yapılır.

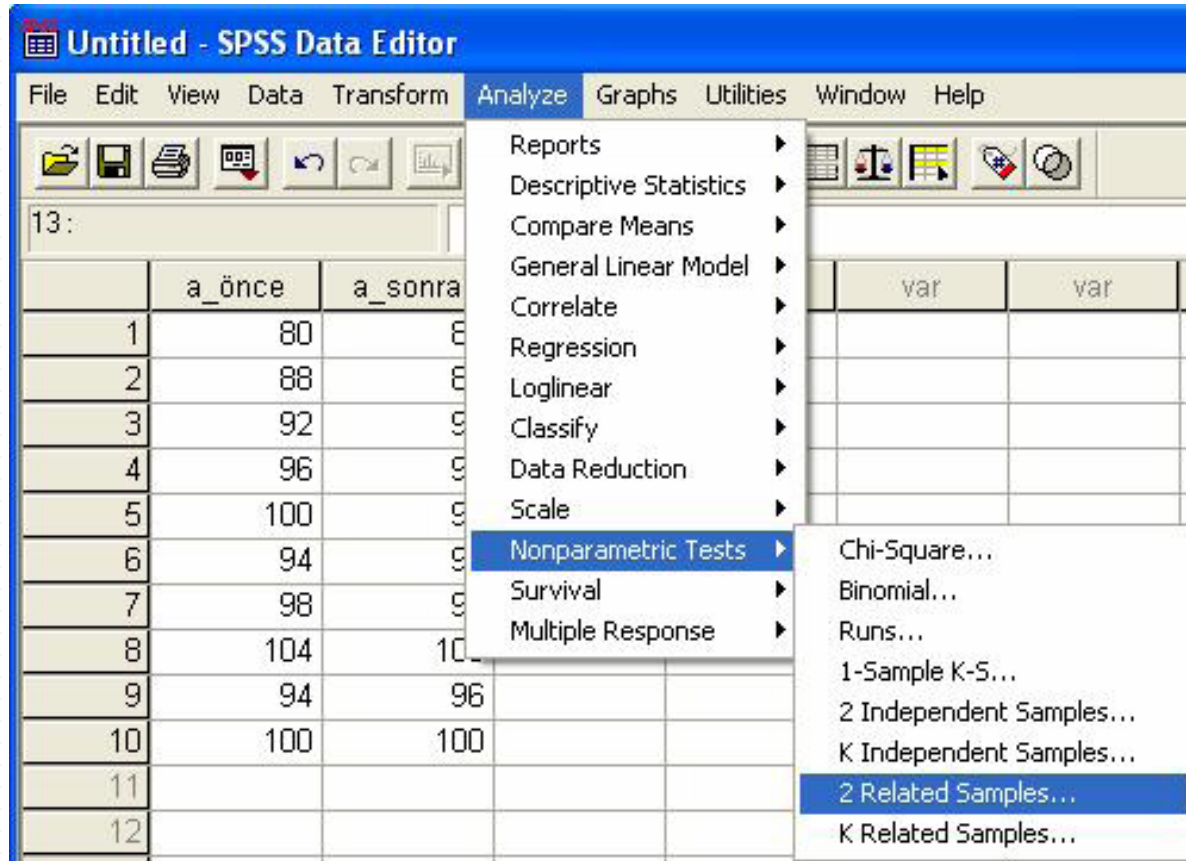


The screenshot shows the SPSS Data Editor interface. The title bar reads "Untitled - SPSS Data Editor". The menu bar includes "File", "Edit", "View", "Data", "Transform", "Analyze", and "G". The toolbar contains icons for file operations and data manipulation. The data grid shows 11 rows of data. The first column is labeled "11:" and contains row numbers 1 through 11. The second column is labeled "a_önce" and the third column is labeled "a_sonra". The fourth column is labeled "var".

	a_önce	a_sonra	var
1	80	80	
2	88	84	
3	92	93	
4	96	94	
5	100	96	
6	94	92	
7	98	92	
8	104	100	
9	94	96	
10	100	100	
11			

Çözüm 3-devam

- Örneğimizdeki ilgili hipotezi test etmek için istatistiksel test olarak “Wilcoxon Sıra Toplamı Testi” kullanılmalıdır. SPSS’te Wilcoxon Sıra Toplamı Testi’ni uygulamak için Analyze menüsünün alt menüsü olan Nonparametrics’ den 2 Related Samples seçilir.

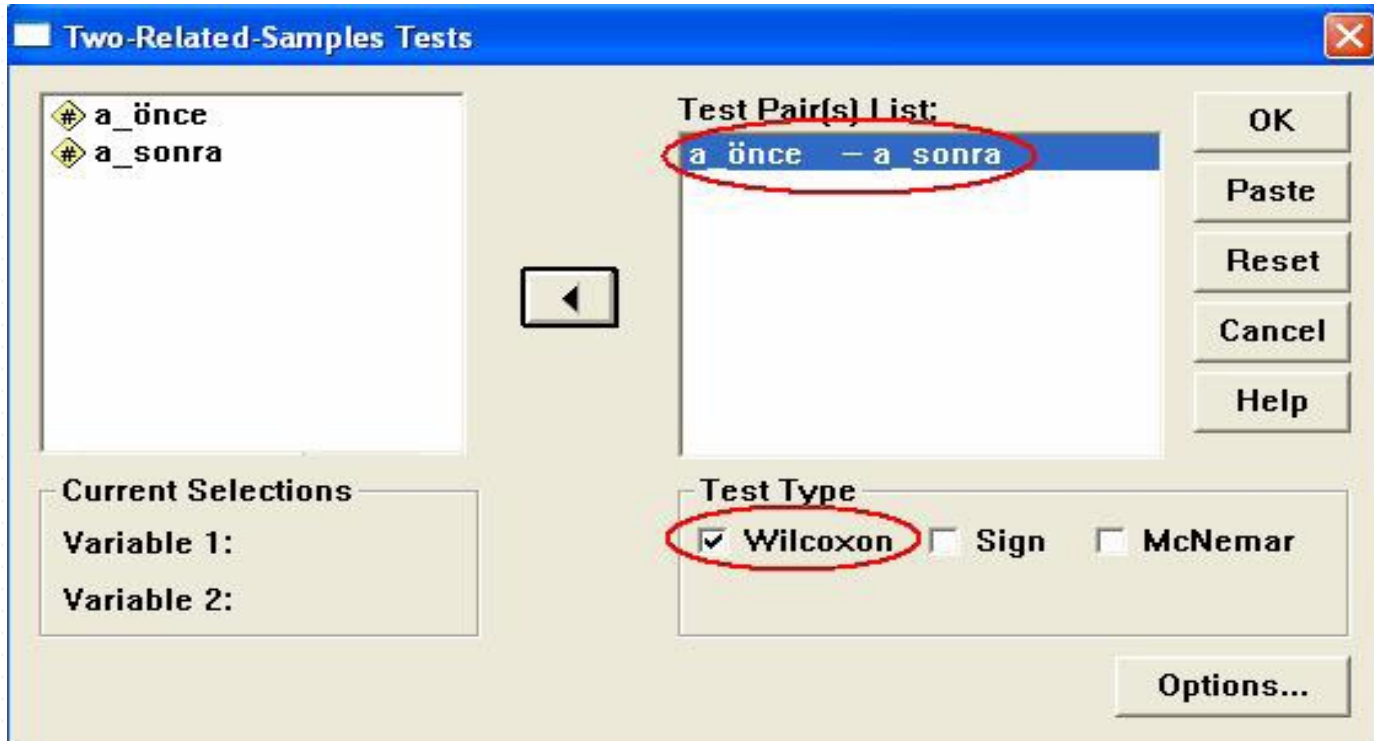


The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, and the 'Nonparametric Tests' option is selected, which has opened a sub-menu. In this sub-menu, the '2 Related Samples...' option is highlighted. The data table in the background has columns 'a_önce' and 'a_sonra' with values for rows 1 through 12.

	a_önce	a_sonra
1	80	8
2	88	8
3	92	9
4	96	9
5	100	9
6	94	9
7	98	9
8	104	10
9	94	96
10	100	100
11		
12		

Çözüm 3-devam

- “Variable 1” için “a_önce”, “Variable 2” için ise “a_sonra” seçildikten sonra, bu değişken çifti, simgesi tıklanarak **Test Pair(s) List** kısmına aktarılır. **Test Type** alanındaki **Wilcoxon** seçeneği işaretlenir ve **OK** tıklanır.



Çözüm 3-devam

- İlgili SPSS çıktısı aşağıdaki gibidir

Wilcoxon Signed Ranks Test

Ranks

	N	Mean Rank	Sum of Ranks
A_SONRA - A_ÖNCE Negative Ranks	6 ^a	5,33	32,00
Positive Ranks	2 ^b	2,00	4,00
Ties	2 ^c		
Total	10		

a. A_SONRA < A_ÖNCE

b. A_SONRA > A_ÖNCE

c. A_ÖNCE = A_SONRA

Test Statistics^b

	A_SONRA - A_ÖNCE
Z	-1,980 ^a
Asymp. Sig. (2-tailed)	,048

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

- Yapılan test sonucunda p değeri 0,048 olarak bulunmuştur. $p < 0,05$ olduğu için H_0 hipotezi reddedilir. Buna göre $\alpha = 0,05$ için, anestezi öncesi ve sonrası nabız sayıları arasında fark olduğu saptanmıştır.

Mann-Whitney U Testi

- İki gruba ait gözlemlerin karşılaştırılmasında kullanılır. Parametrik testlerden t testinin non parametrik alternatifidir. Mann-Whitney U testi, gözlemlerden elde edilen bilgilerin en azından sıralı ölçme ile ölçülenebildiği iki bağımsız örneğin, ait oldukları sıra toplamalarının dağılımlarının aynı olup olmadığını test eder.

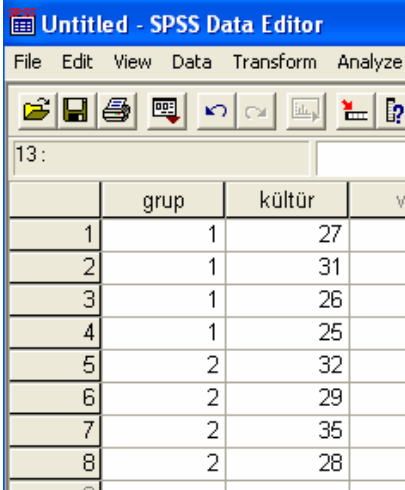
Örnek 4

- İki tip A ve B kültürlerinde her birim hacim için bakteri sayıları aşağıda verilmiştir. Bu iki kültürü $\alpha = 0,05$ düzeyinde birbirleriyle karşılaştıralım.

A Kültürü	B Kültürü
27	32
31	29
26	35
25	28

Cevap-4

- H_0 : A ve B kültürleri arasında, bakteri sayıları bakımından anlamlı bir fark yoktur.
- H_1 : A ve B kültürleri arasında, bakteri sayıları bakımından anlamlı bir fark vardır.
- **Variable View** penceresinde “kültür” ve “grup” isimli iki ayrı değişken tanımlandıktan sonra **Data View** penceresinde veri girişi aşağıdaki gibi yapılır. Burada iki kültür grubunun kodları “grup” yazılı sütuna girilir (örnekte A kültürü grubuna “1” kodu ve B kültürü grubuna “2” kodu verilmiş olduğuna dikkat ediniz). Sonra, birim hacimdeki bakteri sayıları “kültür” yazılı sütuna, gruplara uygun olarak sırasıyla girilir.

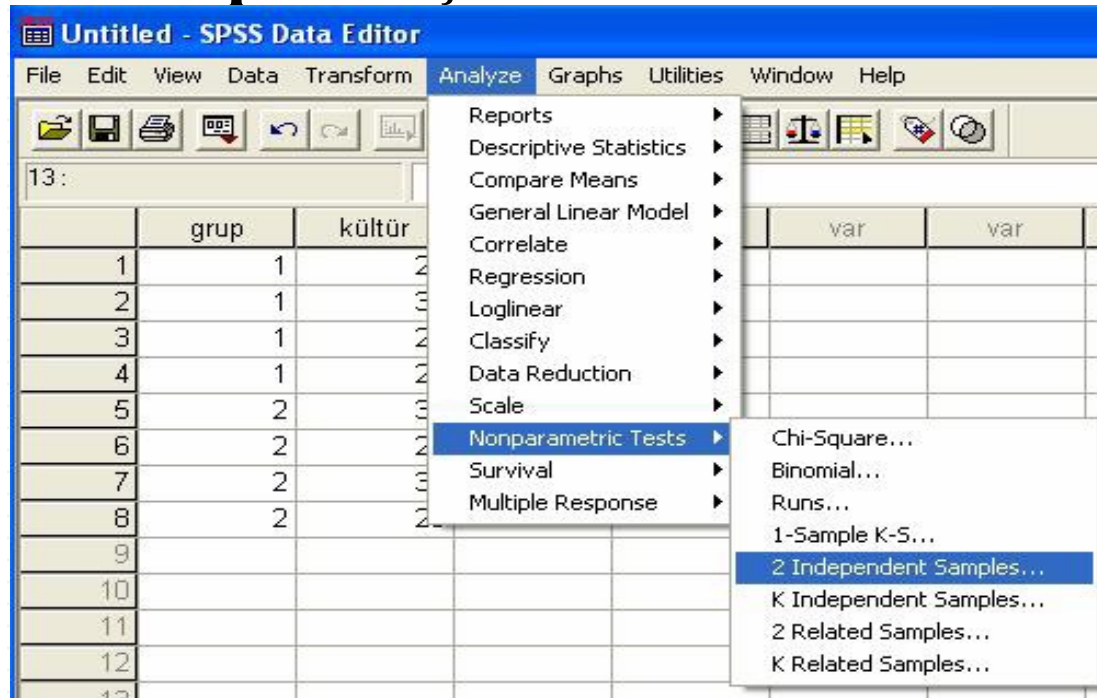


The screenshot shows the SPSS Data Editor window titled "Untitled - SPSS Data Editor". The menu bar includes File, Edit, View, Data, Transform, and Analyze. Below the menu bar is a toolbar with various icons. The main window displays a data table with the following content:

	grup	kültür	v
1	1	27	
2	1	31	
3	1	26	
4	1	25	
5	2	32	
6	2	29	
7	2	35	
8	2	28	

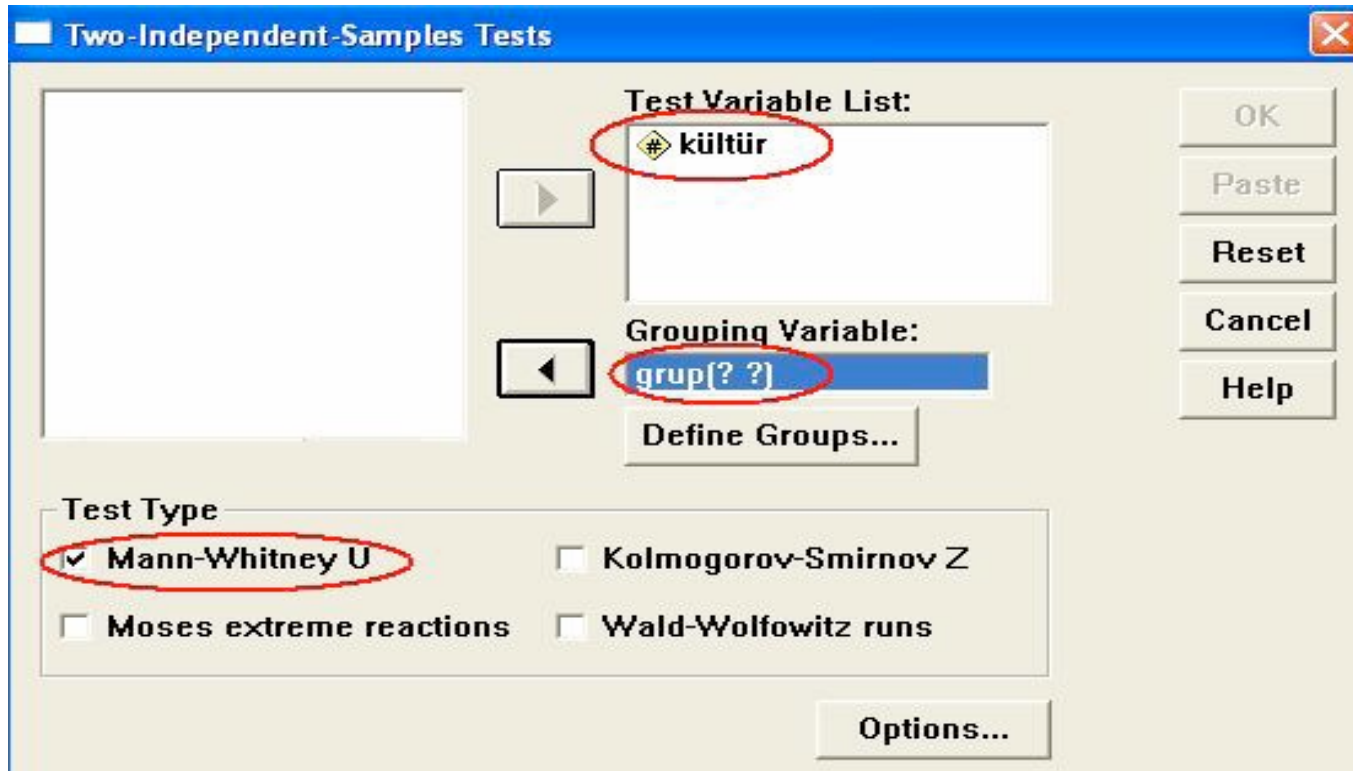
Cevap-4

- Örneğimizdeki ilgili karşılaştırmayı yapabilmek için “Mann-Whitney U Testi” kullanılmalıdır. SPSS’de Mann-Whitney U Testi’ni uygulamak için **Analyze** menüsünün bir alt menüsü olan **Nonparametrics**’den, **2 Independent Samples** seçilir.



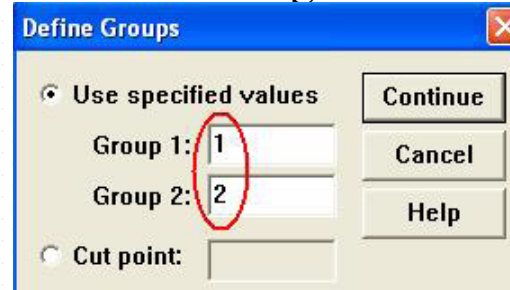
Cevap-4

- Ekrana gelecek aşağıdaki pencerede **Test Variable List** bölümüne “kültür” değişkeni ve **Grouping Variable** bölümüne ise “grup” değişkeni simgeleri tıklanarak aktarılır.

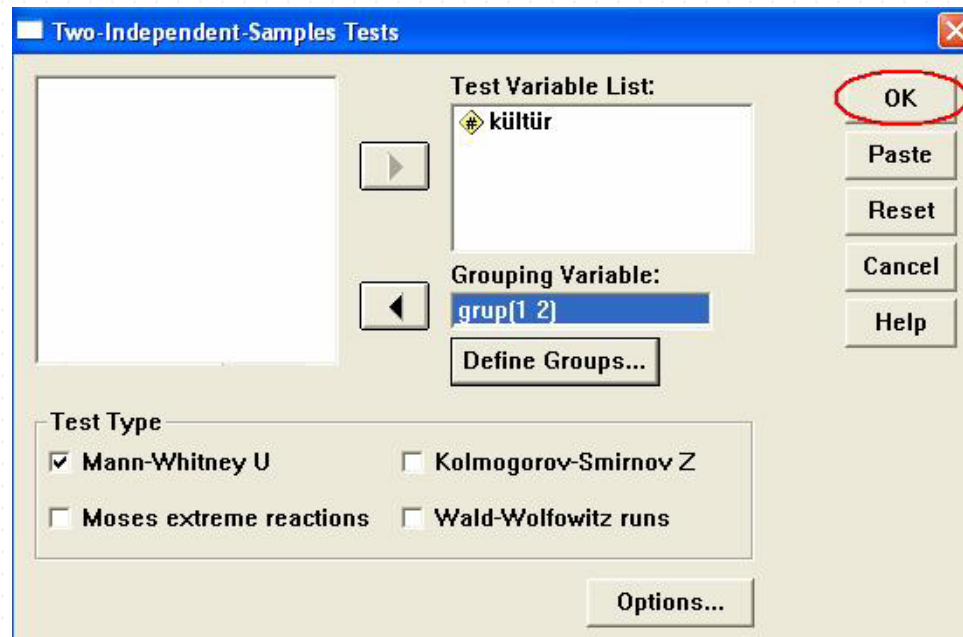


Cevap-4

- **Test Type** alanındaki **Mann-Whitney U** seçeneği işaretlenir ve iki grubun kodları **Define Groups** tıklanarak tanımlanır. Örnekte iki grubun kodları aşağıdaki gibi girilerek **Continue** tıklanır.



- **Grouping Variable** kutusu, aşağıdaki pencerede görülen durumunu alır. Bu pencerede **OK** tıklanarak sonuçlara ulaşılır.



Cevap-4

- İlgili SPSS çıktısı aşağıdaki gibidir.

•SPSS’de Mann Whitney U testi uygulandıktan sonra elde edilen çıktıda, Asymptotic Significance ve Exact Significance olmak üzere iki farklı p değeri elde edilir. Ancak Exact Significance değeri gruptaki denek sayısı az olduğunda ortaya çıkar. İki gruptan herhangi birinde bulunan denek sayısı yaklaşık olarak 20’den az olduğunda, Asymptotic Significance ve Exact Significance değerlerinin her ikisi de çıktı olarak görülmektedir. Bu durumda Exact Significance değeri kullanılmalıdır.

Mann-Whitney Test

Ranks

	GRUP	N	Mean Rank	Sum of Ranks
KÜLTÜR	1	4	3,00	12,00
	2	4	6,00	24,00
	Total	8		



Test Statistics^b

	KÜLTÜR
Mann-Whitney U	2,000
Wilcoxon W	12,000
Z	-1,732
Asymp. Sig. (2-tailed)	,083
Exact Sig. [2*(1-tailed Sig.)]	,114 ^a

a. Not corrected for ties.

b. Grouping Variable: GRUP

- Örnekte, her iki gruptaki denek sayısı 4 olduğu için Exact Significance değeri göz önüne alınır. $p = 0,114 > 0,05$ olduğu için H_0 hipotezi kabul edilir. Buna göre $\alpha = 0,05$ anlamlılık düzeyinde, A ve B kültürleri arasında, bakteri sayıları bakımından anlamlı bir fark yoktur.

Kruskal Wallis Testi

- k bağımsız örneklemden elde edilen verilerin, aynı toplumdaki gelip gelmediğinin test edilmesinde kullanılır. Tek Yönlü Varyans Analizinin non parametrik karşılığıdır.

Örnek 5


- Bir araştırma için alınan 3 grubun yaş ortalamaları arasında fark olup olmadığı test edilmek isteniyor. Gruplara göre yaş dağılımı aşağıda verildiğine göre Tek Yönlü Varyans Analizi için gerekli varsayımların sağlanmadığını göz önüne alarak bu gruplardaki yaşların farklı olup olmadığını $\alpha = 0,05$ anlamlılık düzeyinde kontrol edelim.

Grup 1	Grup 2	Grup 3
15	18	23
18	20	20
12	22	25
10	24	24
13	25	20
12		26
10		

Çözüm 5

- H_0 : Gruplar arasında yaş bakımından fark yoktur.
- H_1 : En az bir grup diğerlerinden farklıdır.

Variable View penceresinde “yaş” ve “grup” isimli iki ayrı değişken tanımlandıktan sonra **Data View** penceresinde veri girişi aşağıdaki gibi yapılır. Burada üç grubun kodları “grup” yazılı sütuna girilir (örnekte grup 1’e “1” kodu, grup 2’ye “2” kodu ve grup 3’e “3” kodu verilmiş olduğuna dikkat ediniz). Sonra, her gruba ait yaş değerleri “yaş” yazılı sütuna, gruplara uygun olarak sırasıyla girilir.

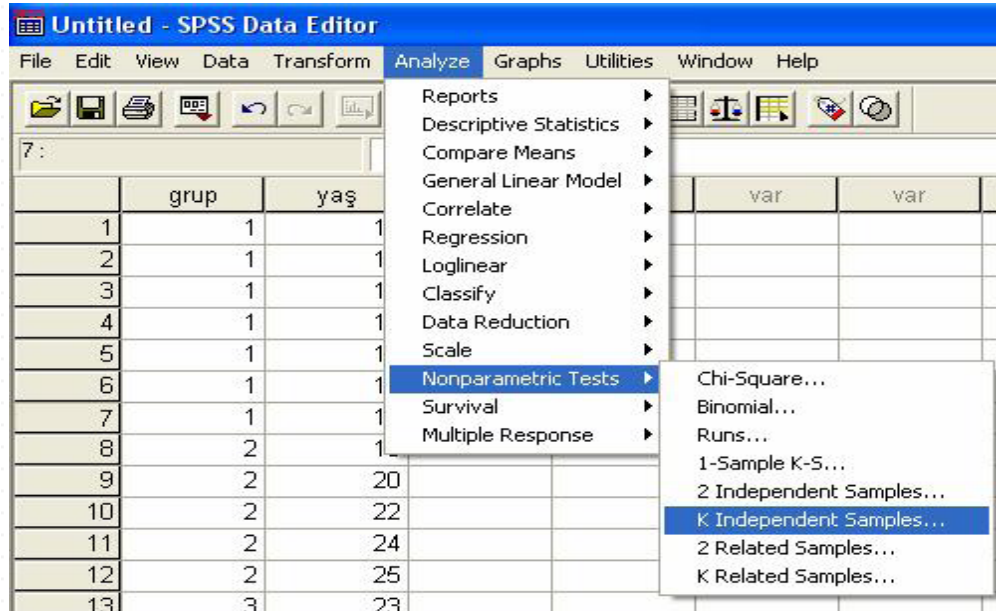


12:

	grup	yaş	var	
1	1	15		
2	1	18		
3	1	12		
4	1	10		
5	1	13		
6	1	12		
7	1	10		
8	2	18		
9	2	20		
10	2	22		
11	2	24		
12	2	25		
13	3	23		
14	3	20		
15	3	25		
16	3	24		
17	3	20		
18	3	26		
19				

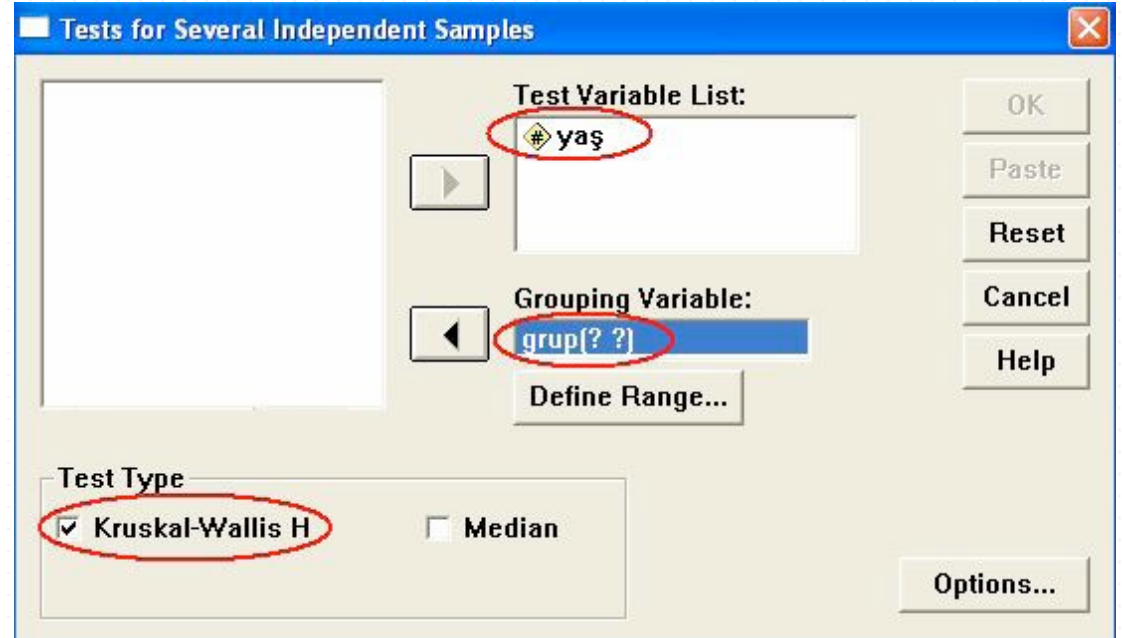
Çözüm 5

- Örneğimizdeki ilgili hipotezi test etmek için istatistiksel test olarak “Kruskal Wallis Testi” kullanılmalıdır. SPSS’de Kruskal Wallis Testi’ni uygulamak için, **Analyze** menüsünün alt menüsü olan **Nonparametrics**’den **K Independent Samples** seçilir.

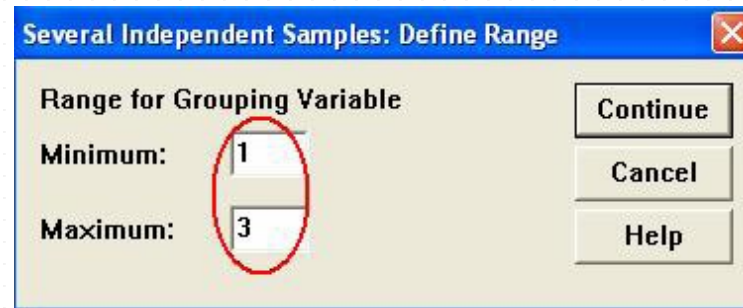


Çözüm 5

Ekrana gelecek aşağıdaki pencerede **Test Variable List** bölümüne “yaş” değişkeni ve **Grouping Variable** bölümüne ise “grup” değişkeni simgeleri tıklanarak aktarılır.

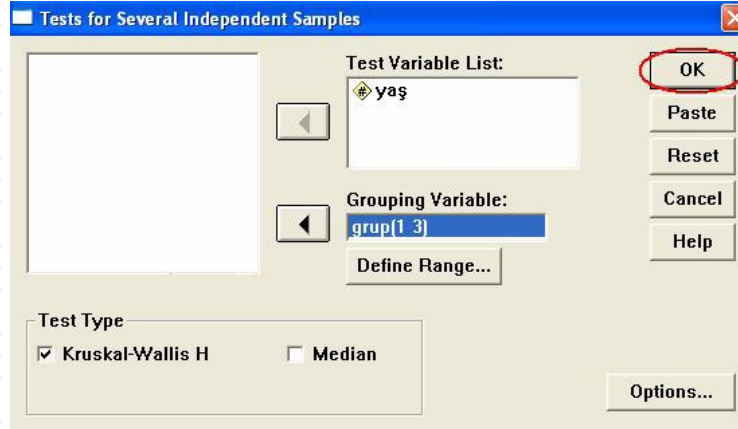


Test Type alanındaki **Kruskal-Wallis H** seçeneği işaretlenir. **Define Range** tıklanarak işleme giren en küçük (minimum) ve en büyük (maksimum) grup numaraları aşağıdaki gibi tanımlandıktan sonra **Continue** tıklanır.



Çözüm 5

Grouping Variable kutucuğu, aşağıdaki pencerede görülen durumunu alır. Bu pencerede **OK** tıklanarak sonuçlara ulaşılır.



İlgili SPSS çıktısı aşağıdaki gibidir.
Kruskal-Wallis Test

Ranks			
	GRUP	N	Mean Rank
YAŞ	1	7	4,07
	2	5	12,10
	3	6	13,67
Total		18	

Test Statistics^{a,b}

	YAŞ
Chi-Square	12,192
df	2
Asymp. Sig.	,002

a. Kruskal Wallis Test

b. Grouping Variable: GRUP

Uygulanan test sonucunda p değeri 0,002 olarak bulunmuştur. $p < 0,05$ olduğu için H_0 hipotezi reddedilir. Buna göre yaş bakımından en az bir grup $\alpha = 0,05$ anlamlılık düzeyinde diğerlerinden farklıdır. Bu örneğimizde olduğu gibi, eğer gruplar arasında anlamlı farklılık olduğu sonucu elde ediliyorsa, farkın hangi gruplardan kaynaklandığını görmek için Mann-Whitney U testi ile gruplar ikili olarak karşılaştırılmalıdır. Örneğimizde, farkın hangi gruplardan kaynaklandığını görmek için, 1-2, 1-3 ve 2-3 grupları Mann-Whitney U testi ile karşılaştırılır.

İki Örneklem Kolmogorov Smirnov Testi

- Bağımsız iki örneklemin dağılımları aynı olan toplumlardan gelip gelmediğini test etmek için kullanılır.
- **Örnek 6** : Primer kanser özofagusta olan 41 hastanın ve primer kanser kardiada olan 42 hastanın yaş gruplarına göre dağılımı aşağıda verilmiştir. İki grubun yaş gruplarına göre dağılımı, $\alpha = 0,05$ anlamlılık düzeyine göre aynı olup olmadığını kontrol edelim.

Yaş Grupları	Kardia	Özofagus
<50	4	3
50-59	4	6
60-69	10	12
70-79	20	17
80-+	4	3

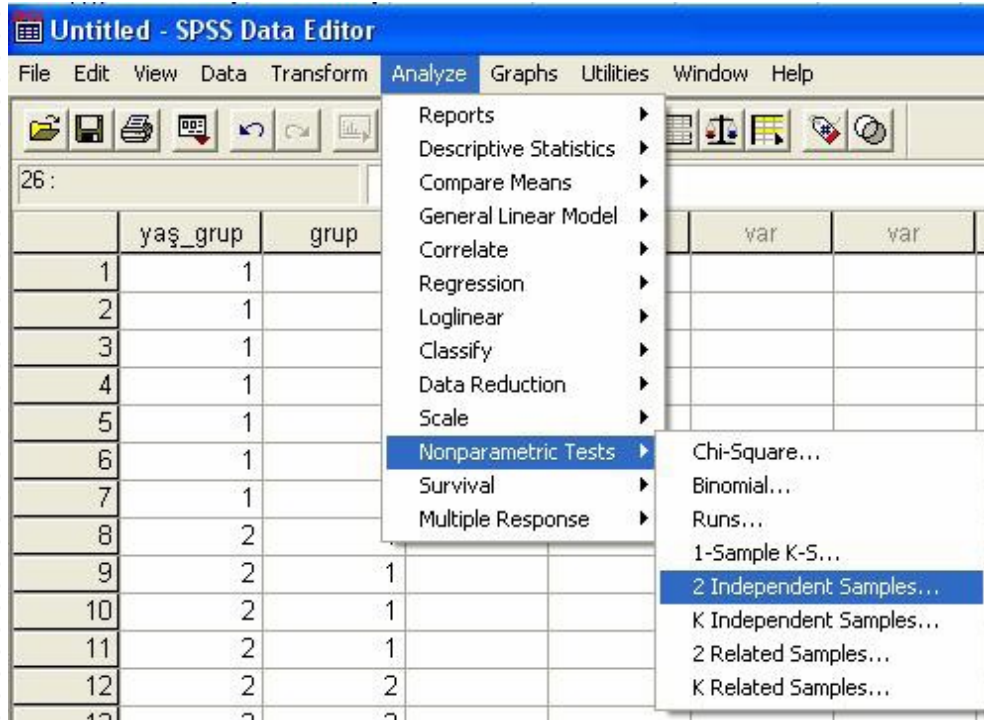
Cevap 6

- H0: İki grubun yaş gruplarına göre dağılımları arasında fark yoktur.
- H1: İki grubun yaş gruplarına göre dağılımları birbirinden farklıdır.

Hasta grubundaki bireylerin yaş gruplarını gösteren kod değerleri “yaş_grup” ismi ile; iki hasta grubunu gösteren kod değerleri ise “grup” ismi ile **Variable View** penceresinde iki ayrı değişken olarak tanımlandıktan sonra, **Data View** penceresinde veri girişi aşağıdaki gibi yapılır. Yaş gruplarının kodları “yaş_grup” yazılı sütuna (<50 için “1” kodu, 50-59 için “2” kodu, 60-69 için “3” kodu, 70-79 için “4” kodu ve 80 ve üstü için “5” kodu verilmiş olduğuna dikkat ediniz); hasta gruplarının kodları ise “grup” yazılı sütuna girilir (kardia grubuna “1” kodu ve özofagus grubuna “2” kodu verilmiş olduğuna dikkat ediniz).

Cevap 6

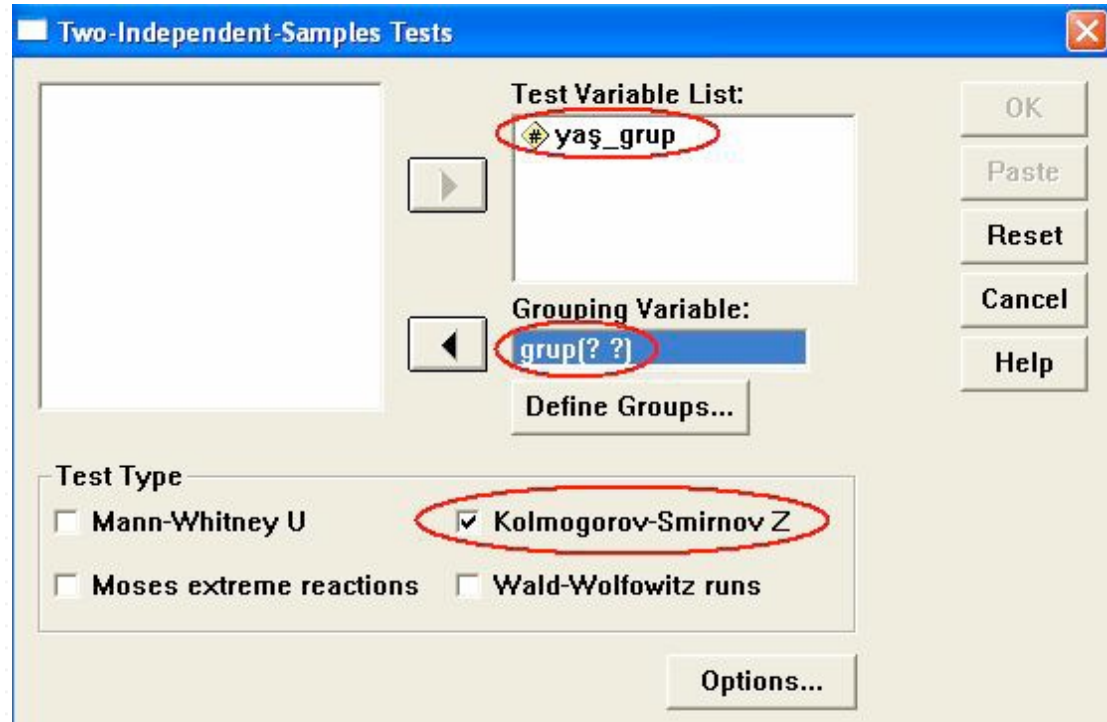
Örneğimizde, ilgili hipotezleri test etmek için “İki Örneklem Kolmogorov Smirnov Testi” uygulanacaktır. Bunun için SPSS’de, **Analyze** menüsünün bir alt menüsü olan **Nonparametric Tests**’den **2 Independent Samples** seçeneği tıklanır.



	yaş_grup	grup
1	1	1
2	1	1
3	1	1
4	1	1
5	1	2
6	1	2
7	1	2
8	2	1
9	2	1
10	2	1
11	2	1
12	2	2
13	2	2
14	2	2
15	2	2
16	2	2
17	2	2
18	3	1
19	3	1
20	3	1
21	3	1
22	3	1
23	3	1
24	3	1
25	3	1
26	3	1
27	3	1
28	3	2
29	3	2
an	3	2

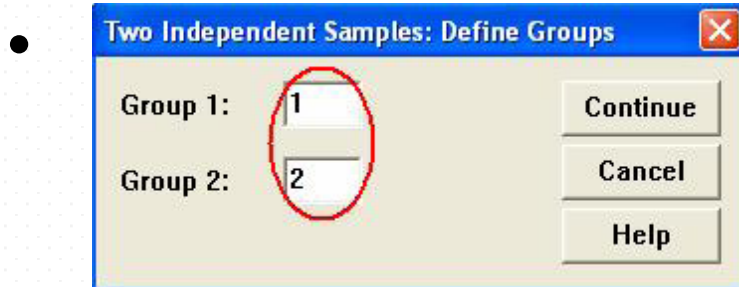
Cevap 6

- Ekranı gelecek ařağıdaki pencerede, “yař_grup” deęiřkeni **Test Variable List** bölümüne, “grup” deęiřkeni ise **Grouping Variable** bölümüne > simgesi tıklanarak aktarılır.

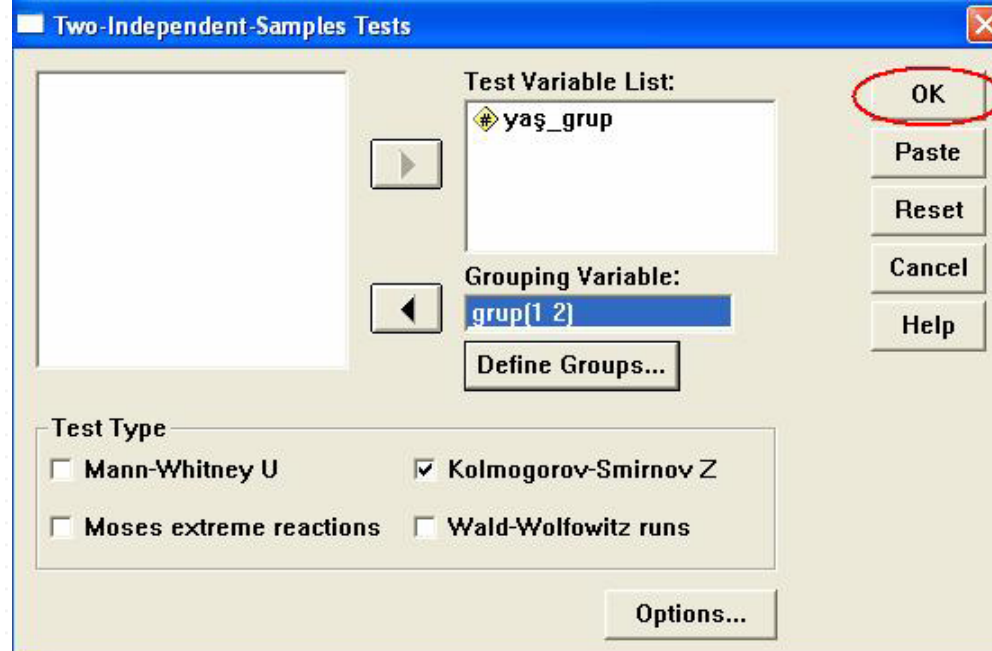


- **Test Type** alanında **Kolmogorov-Smirnov Z** iřaretlenir ve iki grubun kodları **Define Groups** tıklanarak tanımlanır. Örnekte iki grubun kodları ařağıdaki gibi girilerek **Continue** tıklanır.

Cevap 6



- **Grouping Variable** kutusu, aşağıdaki pencerede görülen durumunu alır. Bu pencerede **OK** tıklanarak sonuçlara ulaşılır.



Cevap 6

- İlgili SPSS çıktısı aşağıdaki gibidir.

Frequencies

	GRUP	N
YAŞ_GRUP	1	42
	2	41
	Total	83

Test Statistics^a

		YAŞ_GRUP
Most Extreme Differences	Absolute	,084
	Positive	,022
	Negative	-,084
Kolmogorov-Smirnov Z		,381
Asymp. Sig. (2-tailed)		,999

a. Grouping Variable: GRUP

- Uygulanan test sonucunda p değeri 0,999 olarak bulunmuştur. $p > 0,05$ olduğu için H_0 hipotezi kabul edilir. Buna göre 0,05 anlamlılık düzeyinde kanser hücresi kardiada ve özofagusta olan hastaların yaş gruplarına göre dağılımları arasında fark yoktur.

Ki-Kare Testi-Uygunluk Testi

- Örneklem grubundaki değerlerinin dağılımının hipotez testinde ileri sürülen (normal dağılım) ana kitle dağılımıyla uyumlu olup olmadığını test etmek için kullanılır.
- **Örnek:** Bir otomobil firması bayilerden aldığı sipariş miktarının aylara göre değişip değişmediğini öğrenmek istemektedir. Bunu ki-kare uygunluk testi ile test ediniz.

	Aylar	Sip. Mik.
1	1,00	60,00
2	2,00	68,00
3	3,00	63,00
4	4,00	70,00
5	5,00	80,00
6	6,00	95,00
7	7,00	98,00
8	8,00	46,00
9	9,00	75,00
10	10,00	51,00
11	11,00	120,00
12	12,00	125,00

led - SPSS Data Editor

View Data Transform Analyze Gr



- Define Variable Properties...
- Copy Data Properties...
- Define Dates...
- Insert Variable
- Insert Cases
- Go to Case...

- Sort Cases...
- Transpose...
- Restructure...
- Merge Files ▶
- Aggregate...
- Identify Duplicate Cases...
- Orthogonal Design ▶

- Split File...
- Select Cases...
- Weight Cases...**

Weight Cases

Aylar

Do not weight cases

Weight cases by

Frequency Variable:

◀ Sip_Mik

Current Status: Do not weight cases

OK

Paste

Reset

Cancel

Help

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Mixed Models
- Correlate
- Regression
- Loglinear
- Classify
- Data Reduction
- Scale
- Nonparametric Tests**
 - Chi-Square...**
 - Binomial...
 - Runs...
 - 1-Sample K-S...
 - 2 Independent Samples...
 - K Independent Samples...
 - 2 Related Samples...
 - K Related Samples...
- Time Series
- Survival
- Multiple Response
- Missing Value Analysis...
- Complex Samples

Chi-Square Test

Test Variable List:
Sip_Mik

Expected Range:
 Get from data
 Use specified range
Lower:
Upper:

Expected Values:
 All categories equal
 Values:
Add Change Remove

Buttons: OK, Paste, Reset, Cancel, Help, Exact..., Options...

- H_0 : Aylara göre sipariş miktarları arasında fark yoktur.
- H_1 : Aylara göre sipariş miktarları arasında fark vardır.

Test Statistics

	Sip_Mik
Chi-Square ^a	89,871
df	11
Asymp. Sig.	,000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 79,3.

- H_0 red, Aylara göre sipariş miktarları arasında fark vardır.

Ki-Kare-Bağımsızlık Testi

İki veya daha fazla değişken grubu arasında ilişki bulunup bulunmadığını incelemek için kullanılır.

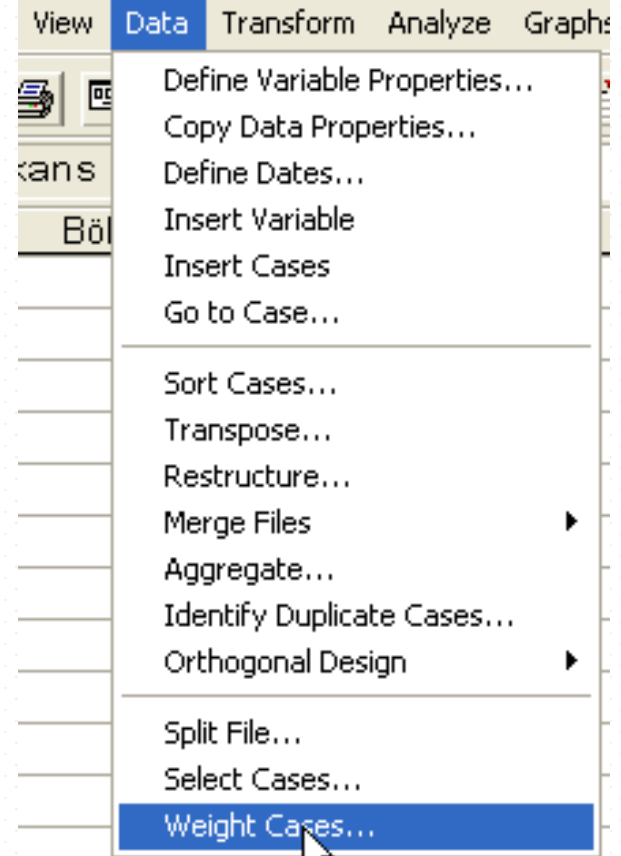
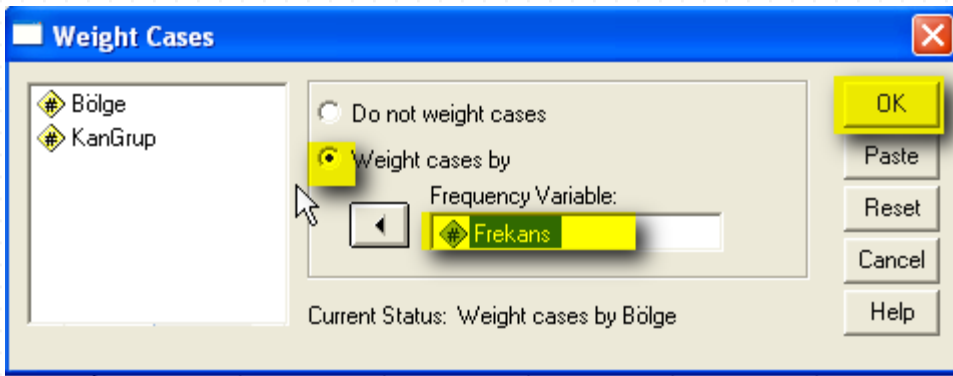
H_0 : Değişkenler birbirinden bağımsızdır.

H_1 : Değişkenler birbirinden bağımsız değildir.

- **Örnek:** İki farklı bölgeye ait kişiler, kan gruplarına göre sınıflandırılmış ve aşağıdaki sonuçlar elde edilmiştir. Buna göre bölgeler ve kan grupları arasındaki ilişkiyi $\alpha=0,05$ hata düzeyine göre test ediniz.

Bölgeler	Kan Grupları				Toplam
	0	A	B	AB	
Batı	30	145	68	37	280
Doğu	60	115	32	13	220
Toplam	90	260	100	50	500

	Bölge	KanGrup	Frekans
1	1,00	1,00	30,00
2	1,00	2,00	145,00
3	1,00	3,00	68,00
4	1,00	4,00	37,00
5	2,00	1,00	60,00
6	2,00	2,00	115,00
7	2,00	3,00	32,00
8	2,00	4,00	13,00



H_0 : Bölgeler ve kan grupları arasında ilişki yoktur

H_1 : Bölgeler ve kan grupları arasında ilişki vardır

Analyze Graphs Utilities Window Help

- Reports
- Descriptive Statistics**
 - Frequencies...
 - Descriptives...
 - Explore...
 - Crosstabs...**
 - Ratio...
- Tables
- Compare Means
- General Linear Model
- Mixed Models
- Correlate
- Regression
- Loglinear
- Classify
- Data Reduction
- Scale
- Nonparametric Tests
- Time Series
- Survival
- Multiple Response
- Missing Value Analysis...
- Complex Samples

Crosstabs

Row(s): Bölge

Column(s): KanGrup

Layer 1 of 1

Previous Next

Display clustered bar charts

Suppress tables

Exact... **Statistics...** Cells... Format...

OK Paste Reset Cancel Help

Crosstabs: Statistics

Chi-square

Nominal

- Contingency coefficient
- Phi and Cramér's V
- Lambda
- Uncertainty coefficient

Nominal by Interval

- Eta

Cochran's and Mantel-Haenszel statistics

Test common odds ratio equals:

Correlations

Ordinal

- Gamma
- Somers' d
- Kendall's tau-b
- Kendall's tau-c

Kappa

Risk

McNemar

Continue

Cancel

Help

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Bölge * KanGrup	500	100,0%	0	,0%	500	100,0%

Bölge * KanGrup Crosstabulation

Count

	KanGrup	KanGrup				Total
		1,00	2,00	3,00	4,00	
Bölge 1,00		30	145	68	37	280
2,00		60	115	32	13	220
Total		90	260	100	50	500

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	31,191 ^a	3	,000
Likelihood Ratio	31,710	3	,000
Linear-by-Linear Association	28,126	1	,000
N of Valid Cases	500		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 22,00.

H_0 red Bölgeler ve kan grupları arasında ilişki olduğu söylenir.

Ki-Kare Homojenlik testi

- Birbirinden bağımsız olarak seçilen iki veya daha fazla örneklemin aynı anakitleden çekilip çekilmediğinin belirlenmesinde kullanılır. Hipotezler

H_0 : Örneklemeler aynı ana kitleden seçilmiştir

H_1 : Örneklemeler aynı ana kitleden seçilmemiştir

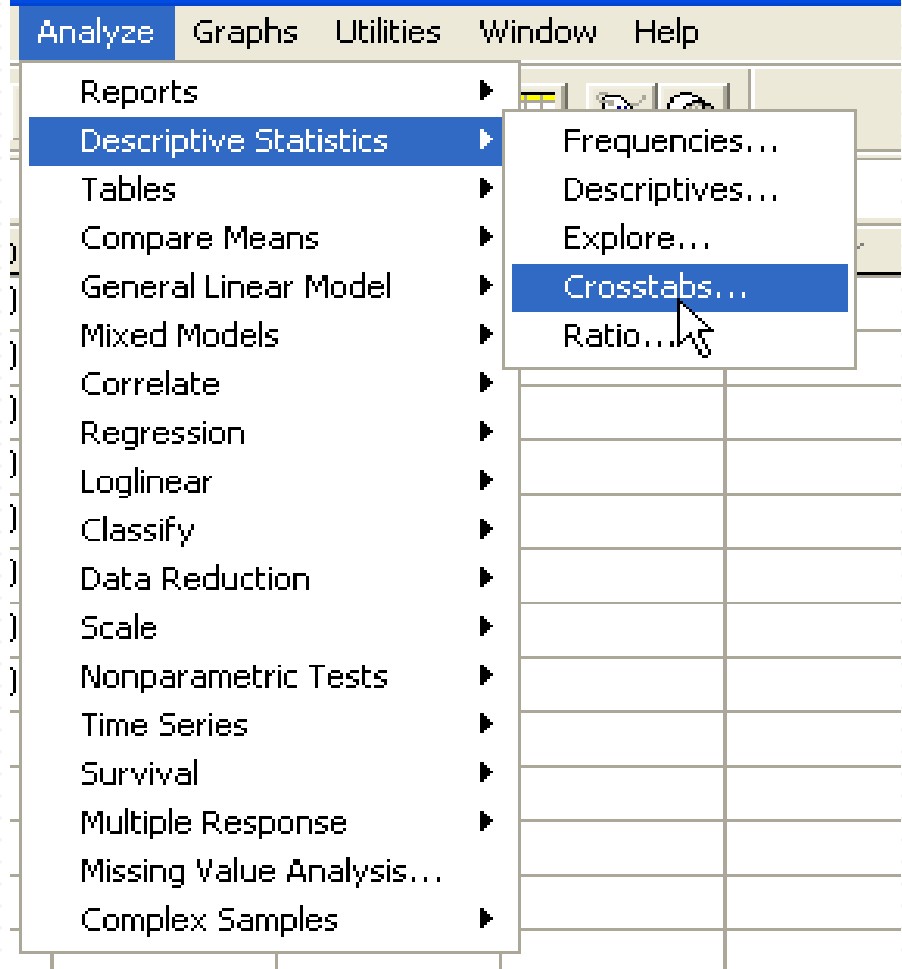
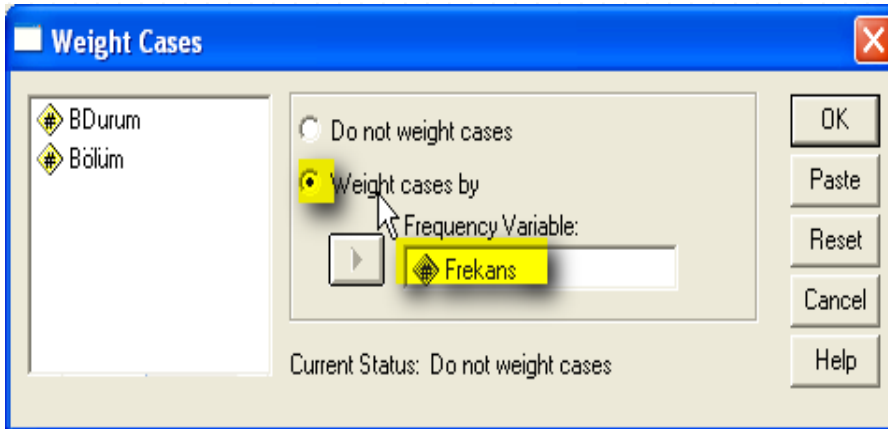
Örnek: Bir bankanın açmış olduğu sınava giren öğrencilerin başarı durumlarının bölümlere göre aşağıda verilmiştir.

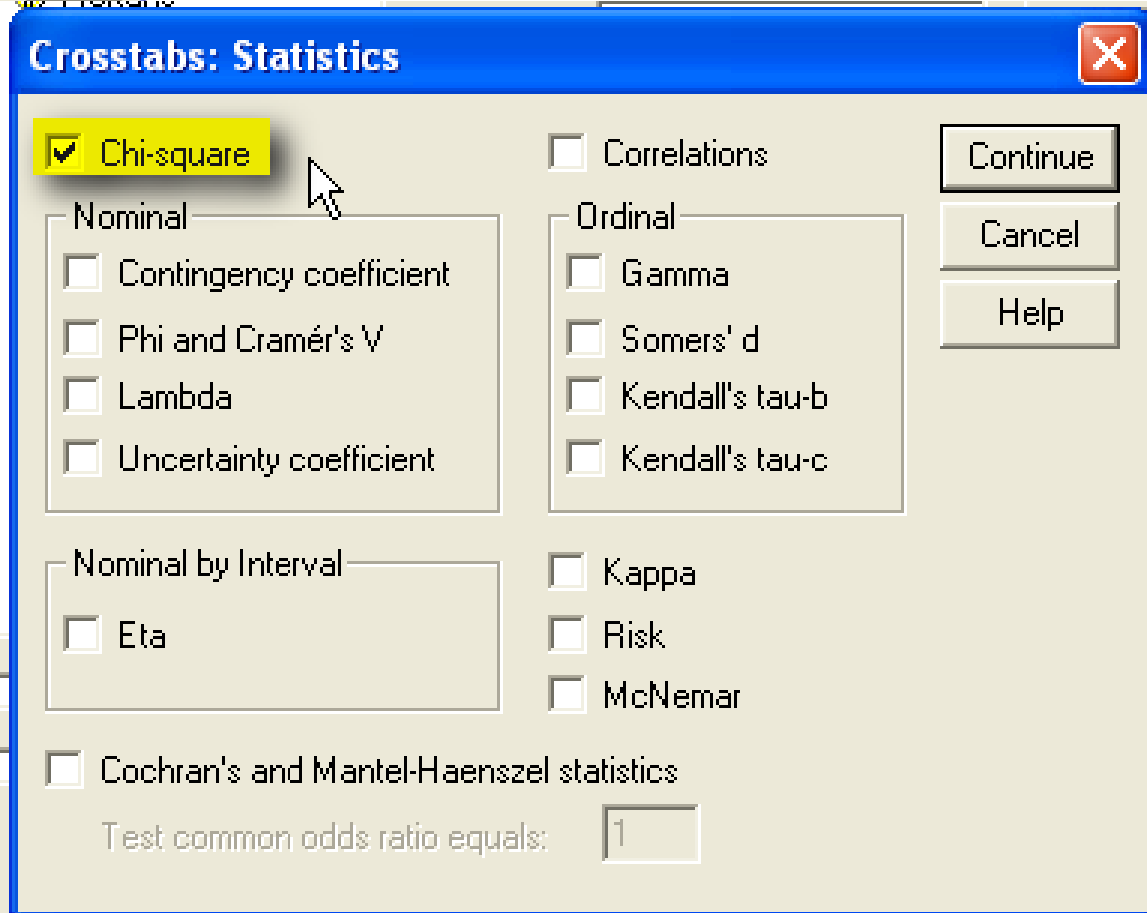
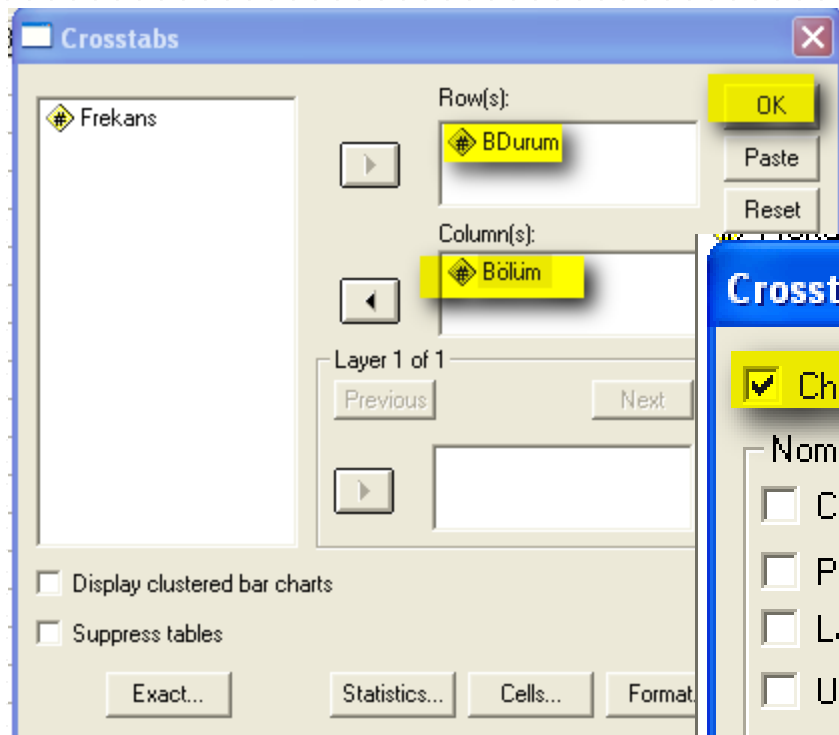
Başarı Durumu	Bölümler				Toplam
	İşletme	İktisat	Maliye	Kamu Yönetimi	
Başarılı	30	36	24	20	110
Başarısız	24	20	18	28	90
Toplam	54	56	42	48	200

H_0 : Bölümler başarı açısından homojendir

H_1 : Bölümler başarı açısından homojen değildir

/Data/Weight Cases





Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
BDurum * Bölüm	200	100,0%	0	,0%	200	100,0%

BDurum * Bölüm Crosstabulation

Count

		Bölüm				Total
		1,00	2,00	3,00	4,00	
BDurum	1,00	30	36	24	20	110
	2,00	24	20	18	28	90
Total		54	56	42	48	200

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,483 ^a	3	,140
Likelihood Ratio	5,500	3	,139
Linear-by-Linear Association	2,368	1	,124
N of Valid Cases	200		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 18,90.

H_0 kabul, Bölümlere göre başarı durumları arasında fark yoktur. Başarı durumu bölümlere göre homojendir.