HAYVANCILIK DENEMELERİNİ DEĞERLENDİRME TEKNİKLERİ

Prof. Dr.Yavuz AKBAŞ
Ege Üniversitesi Ziraat Fakültesi Zootekni Bölümü
Biyometri ve Genetik Anabilim Dalı Bornova, İzmir
yavuz@ziraat.ege.edu.tr

Dersin Amacı ve Hedefi

Lisansüstü düzeyde araştırma yapacak öğrencilerin çalışmalarının dizaynında ve denemelerinin yürütülmesinde nelere dikkat edecekleri temel istatistik kavramlarla verilmeye çalışılmaktadır.

Ayrıca öğrencilerin lisansüstü çalışmalarında toplayacakları verileri kendilerinin değerlendirebilmesi için gereksinim duyacakları istatistik yöntemler anlatılmakta ve bu yöntemleri kolayca uygulayabilecekleri bilgisayar yazılımları bu ders kapsamında tanıtılmaktadır.

Amaç istatistik problemlere sistematik ve analitik bir şekilde yaklaşma yeteneğinin lisansüstü öğrencilere verilmesidir. Temel kavram ve yöntemler tarım alanından seçilmiş problemlere uygulanmaktadır.

İçerik

- Araştırma ve Deneme Kavramları
- Hipotezler ve Hipotez Testleri
- Deneme Tipleri
- Veri nedir
- Hayvancılık verilerinin yapısı
- Veri toplamada dikkat edilecek konular
- Veri tipleri
- Xedit, Notepad, Edit, SPSS, SAS gibi çeşitli programlar kullanarak veri girişi ve editleme

İçerik

- Model kavramı ve modelleme
- SPSS ve SAS gibi genel istatistik programlara giriş
- Hayvancılık verilerine yönelik istatistik programları (LSML, DFREML, VCE, ASREML)
- Deneme Tipleri
  - Tek gruplu denemeler
  - İki gruplu Denemeler
  - Çok gruplu denemeler
- Tanımlayıcı İstatistikler
- Verilerin kontrol edilmesi

İçerik

- Temel Deneme Desenleri
  - Tam şansa bağlı deneme deseni ve analizi
    - Varyans analizi (ANOVA) ve varsayımları
    - Çoklu karşılaştırma testleri
  - Şansa bağlı tam bloklar deneme deseni ve analizi
  - Latin karesi deneme deseni ve analizi
- Çok faktörlü denemeler ve analizi
- ANOVA'ya karşı alternatifler

İçerik

- Doğrusal regresyon ve korelasyon
- Kovaryans analizi
- Tekrarlamalı ölçümlü denemeler
- Hayvan besleme denemelerinde veri analiz
- Süt sığırlığı verilerinin analizi
- Kanatlı verilerinin analizi
- Koyun, keçi ve diğer türlere ait verilerinin analizi.

Elinizin Altında Bulunması Gerekenler

- İstatistik kitabı
- Araştırma ve Deneme Metotları kitabı
- İstatistik Tablolar

Sınıf çalışmaları için bir disket getirilecek (üzerine adınız, soyadınız yazılacak)

Dersin Değerlendirilmesi

Bir yarıyılda öğrencilere verilen uygulamalar arasınav olarak değerlendirilmekte (% 40) ve yarıyılsonu sınavı (% 60) bilgisayar başında sözlü olarak yapılmaktadır.

Önşart/Önerilen

Lisans düzeyinde:

1. İstatistik
2. Araştırma ve Deneme Metotları

derslerinin almış olması tavsiye edilir.

Kitaplar

İstatistik

- Püskülcü, H., İkiz, F. 1986. İstatistiğe Giriş. Ege Üniversitesi Basımevi, Bornova, İzmir.
- Atıl, H. 1998. İstatistik. Ege Üniversitesi Ziraat Fakültesi Yayınları No: 531, İzmir.
- Miran, B. 2003. İstatistik.

Araştırma ve Deneme Metotları

- Yızdız, N.; Bircan, H. 1994. Araştırma ve Deneme Metotları. Atatürk Üniversitesi Yayınları No: 697. Ziraat Fakültesi No: 305, Ders Kitapları Serisi No: 57. Erzurum.
- Düzgüneş, O.; Kesici, T.; Kavuncu, O.; Gürbüz, F. 1987. Araştırma ve Deneme Metotları. Ankara Üniversitesi Ziraat Fakültesi Yayınları: 1021, Ders Kitabı: 295, Ankara.
- Ott, Lyman, 1988. An introduction to statistical methods and data analysis. PWS-Kent Publishing Company, Boston.

İstatistik Paket Programlar

- İstatistikte değerlendirmelerde kullanılan çok sayıda paket program vardır. Bunlardan bazıları;
- SPSS
- SAS
- MINITAB
- S-PLUS
- STATISTIKA
- SYSTAT

Bu ders kapsamında konular SPSS üzerinden anlatılacaktır.

Neden SPSS?

- Kullanımının Kolay olması (özellikle öğrenciler için)
- Menü üzerinden yönetim
- İçerdiği istatistik analiz modüllerinin sıklıkla güncelleştirilmesi.
- Popüler bir program olması.

SPSS

- SPSS sözcüğü "sosyal bilimler için istatistik paketi (Statistical Packages for Social Science) sözcüklerinin baş harflerinden oluşmaktadır.
- SPSS günümüzde artık sadece sosyal bilimlerde kullanılan bir program değildir.
- Birçok istatistiksel analizi yapabilen oldukça geniş kapsamlı bir istatistiksel paket program haline gelmiştir.
- www.spss.com adresinden daha ayrıntılı bilgi alınabilir.

SPSS Hakkında Genel Bazı Bilgiler

- İlk versiyonları DOS tabanlı idi.
- Windows tabanlı olarak çalışmaktadır.

•	1967 Dev ilkel bilgisayarlar	... Delikli kartlar !
•	1983 SPSS/PC+	... DOS
•	1991 SPSS/PC+ 4.0	
•	1993 SPSS/Win 5.0	... Win 2.0
•	1994 SPSS/Win 6.1	... Win 3.1
•	1995 SPSS/Win 6.1.3	... Win 3.1, Win 95, NT
•	1996 SPSS/Win 7.0	... Win 95, NT
•	1997 SPSS/Win 7.5	
•	1998 SPSS/Win 8.0	
•	1999 SPSS/Win 9.0	
•	2000 SPSS/Win 10.0	
•	2001 SPSS/Win 11.0	
•	2003 SPSS/Win 12.0	
•	2004 SPSS/Win 13.0	

Bu ders SPSS 11 üzerinden anlatılacaktır.

SPSS 13.0

Meta Brown
Systems Engineer
August 2004

"I think the new feature set is a major improvement in an already outstanding product."

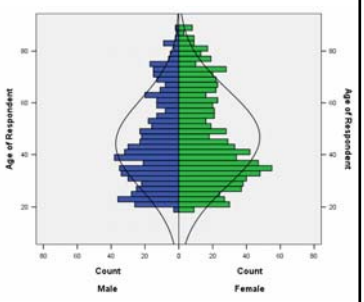
***James W. Golden, PhD , U of Arkansas at Little Rock,* SPSS 13.0 Beta Tester**

Themes of SPSS 13.0

- SPSS Base
  - Enhanced reporting
  - More powerful data management
- Add-on modules
  - Better identification of groups
  - Predictive analytics for survey research

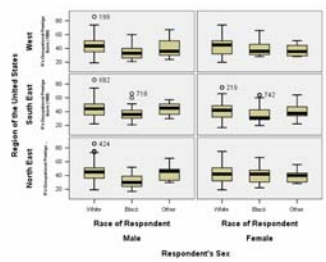
Enhanced reporting

- New chart types
  - 3D bar (Simple, clustered, and stacked)
  - Population pyramids/mirror plots
  - Dot charts/dot density



Enhanced reporting

- More chart features/options
  - Paneled charts
  - Error bars on categorical charts (bars, lines)
  - Sort categories display option
  - More control over data value labels
  - More control over reference lines



Enhanced reporting

- Export charts and tables to PowerPoint
- Pivot table output for numerous procedures
- Improvements to SPSS Tables

More powerful data management

- Very long text string variables (No more 255 byte limit!)
- Save time by saving aggregated values back to the active data file with the improved Aggregate command
- Improved Autorecode (strings to numeric)
- Splitter controls in the Data Editor

What else is new in SPSS 13.0: Predictive Analytics for Survey Research

- Improvements to SPSS Complex Samples add-on module:
  - Complex Samples General Linear Models to analyze and to predict numerical outcomes like customer satisfaction
  - Complex Samples Logistic Regression to analyze and to predict categorical outcomes like who will respond to my mailing

System Requirements

- Windows Me/XP/2000 are the preferred platforms. Windows 98 is supported as well.
- Pentium-class processor
- 128 MB RAM minimum
- 220MB hard disk space
- SVGA monitor

Questions and Answers

- **Sales:**
  sales@spss.com
- **Technical support:**
  http://support.spss.com/
  (312) 651-3410
- **"See it in SPSS" events:**
  www.spss.com/seeit

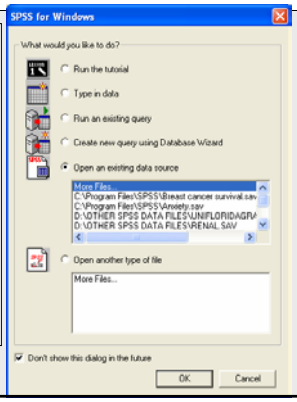
SPSS'i çalıştıralım.



SPSS ortamını (Pencere, menü, toolbar'ları) inceliyelim.

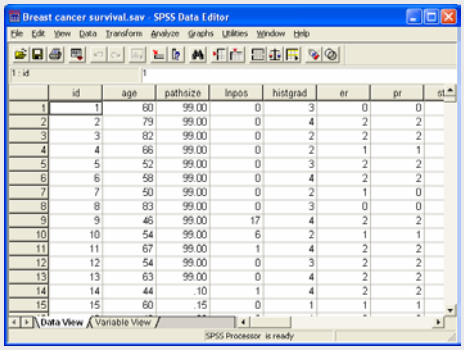
Başlangıç Ekranı

- Listeyi inceleyelim…
- "Don't show this dialog in the future."
- OK



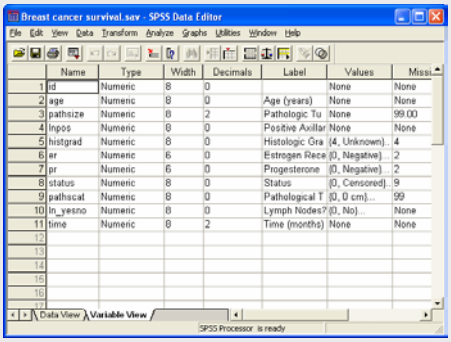
Örnek Veri: "Breast Cancer Survival"

Veri View



Veri Setindeki Değişkenlerin Yapısı

Değişkenler



Dosya Bilgisi

- Utilities Menusü
- "File Info…" seçeneği
- Output penceresine bak



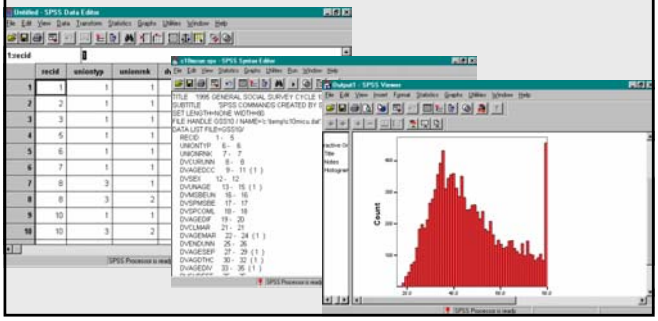
Kullanıcı Pencereleri

SPSS'in kullanımında üç farklı pencere ile karşılaşırız. Bunlar:

- Data Editor (veri işleme) penceresi
- Syntax Editor (komut işleme) penceresi
- Viewer (görüntüleme) penceresi

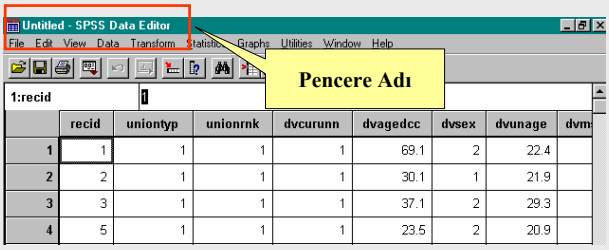
Pencereler

Hangi pencerede çalışıldığı önemlidir. Çünkü her pencere özel fonksiyonlara sahiptir.



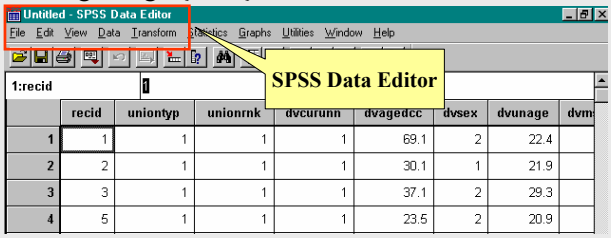
Pencereler

Pencerelerin ismi, pencerenin sol üst köşesinde yazılıdır.



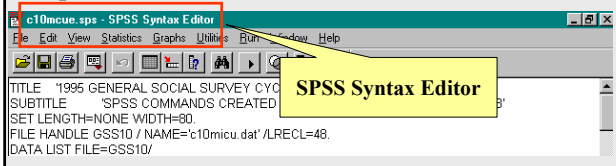
Pencereler

Data Editor penceresi, SPSS tarafından verilerin okunduğu, aktif veri dosyamızın bir kopyasını içerir. Temel veri yönetimi ve analizi bu pencere aracılığı ile gerçekleştirilir.



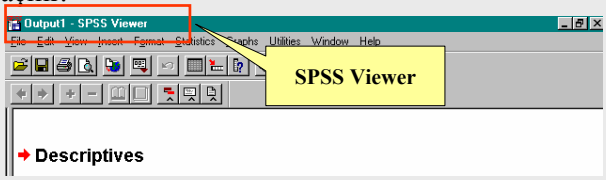
Pencereler

Syntax Editor penceresi, analizlerin menüler aracılığı ile değil, SPSS komutları yazılarak gerçekleştirildiği penceredir. Benzer şekilde daha önce yazılıp kaydedilmiş komutlar, bu pencerede açılıp çalıştırılabilir.



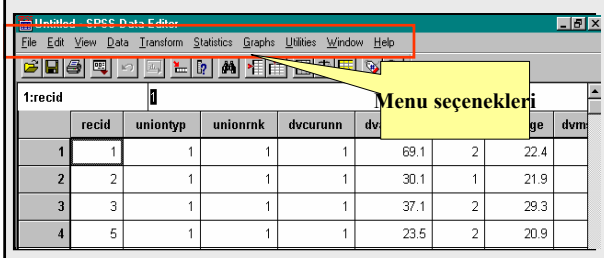
Pencereler

Viewer penceresi, analizler sonunda elde ettiğimiz bütün istatistik sonuçları ile tablo ve grafiklerin (charts) görüntülendiği penceredir. Bu pencere sistem ilk defa çalıştırıldığında otomatik olarak açılır.



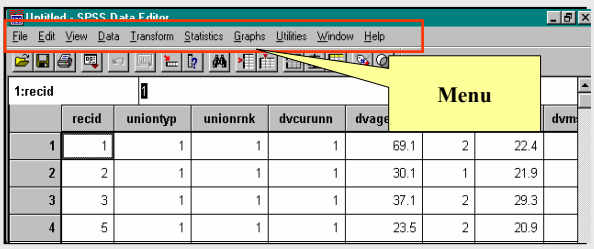
Menu

Kullanıcılar için ikinci önemli arayüz, SPSS'in menu sistemidir.



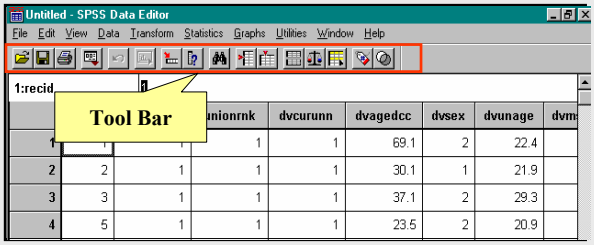
Menu

SPSS'de gerçekleştirilen işlerin çoğu menulerden seçme şeklinde gerçekleştirilir. Alternatif yol ise SPSS komut dilinin bilinmesidir.



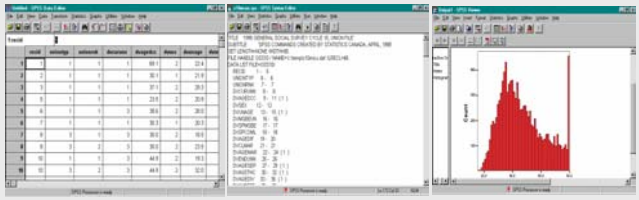
Tool Bar

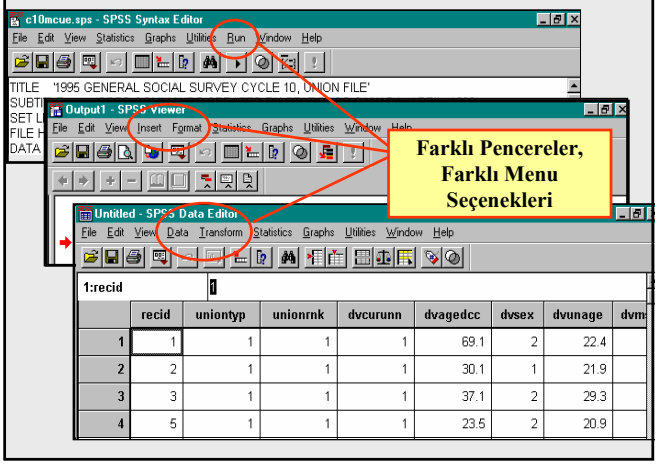
Toolbar üzerindeki ikonlar aracılığı ile sık kullanılan bazı özelliklere daha **hızlı ve kolay** ulaşabiliriz.

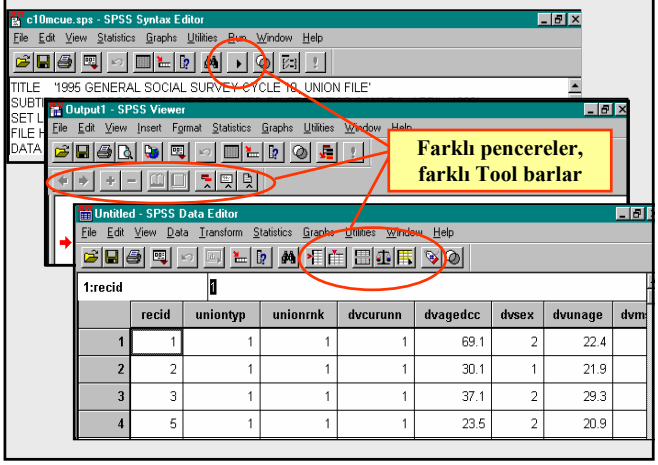


Menu ve Tool bar

Menu ve tool bar içindeki seçenekler bulunduğumuz pencereye göre değişmektedir.

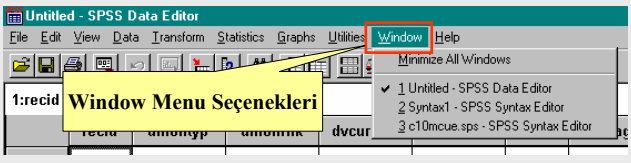






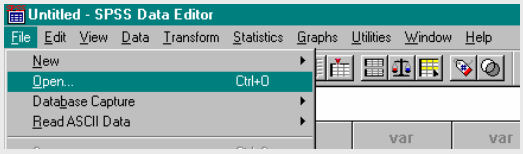
Pencereler Arası Geçiş

SPSS pencereleri arasında geçiş yapmada "Window" menu seçeneklerini kullanabiliriz.



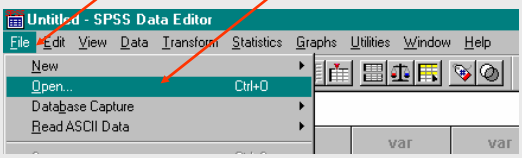
Dosya işlemleri

DLI dosyalarından SPSS komutlarının kullanımı için, File menüsünden dosyayı (File) seç ve aç (Open).



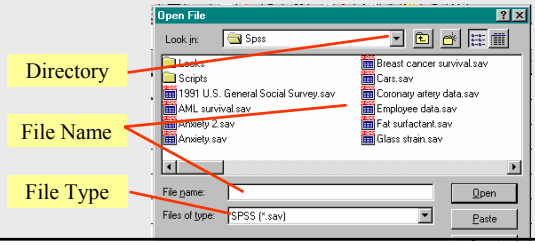
Dosya İşlemleri

Örneğin DLI dosyalarından SPSS komutlarının kullanımı için, File menüsünden dosyayı (File) seç ve aç (Open).



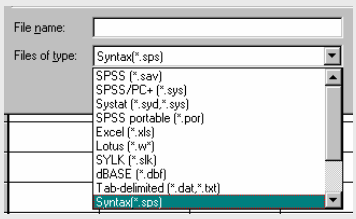
Dosya Açma

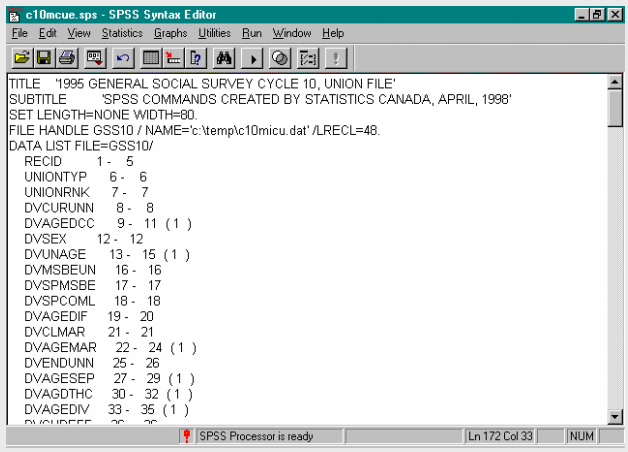
Açılan diyalog penceresinden dizin (directory), dosya adı (file name) ve dosya tipi (file type) seçilebilir.



DLI Komut Dosyaları

Bu durumda dosya tipi olarak "Syntax" seçilir. Bu dosyaların dosya adı uzantısı ".sps"dir.





Farklı Tip Dosyalardan Veri Alımı

- Excel, SAS, dBase
- İlk satır değişken adı
- "File" Menusünden Open
  - Data…
- Dosya Tipi
  - Excel

Etiketleme

- Labels
- Değişken tipi
- Düzey etiketleri (Value Labels)

SPSS Tutorials

- Help Mensüsünde
- Web sayfalarında
- Kitaplarda

SPSS'in Özellikleri

SPSS istatistik yazılımı aşağıdaki işlemlerde bize yardımcı olur:

- Veri yönetimi
- Verileri dönüştürme (transforme etme)
- Verileri analiz etme
- Grafik ve tablolar ile verileri sunma

Veri Yönetimi

Veri yönetiminde:

- Veri ekleme, veri birleştirmek veya dosya yazmak
- Verileri sıralamak
- Koşullu veya şansa bağlı olarak veri seçmek
- birleştirilmiş veya özet dosyalar oluşturmak

Veri dönüştürme

Veri dönüştürme ile:

- numeric dönüşümler
  - koşullu veya koşulsuz yeni değişkenler oluşturma
  - mevcut verileri tekrar kodlama
- string dönüşümleri
- tarih ve zaman dönüşümleri

Verileri Analiz Etme

İstatistik prosedürler:

- tanımlayıcı
- karşılaştırıcı
- veri-azaltma (data-reduction)
- modelleme

Verileri Sunma

İki farklı şekilde verileri görsel olarak sunabiliriz.

- Grafikler
- Tablolar

İstatistik, Veri &
İstatistiksel Düşünme

Amaç

- 1. İstatistiğin tanımlanması
- 2. İstatistiğin kullanımının anlatımı
- 3. Tanımlayıcı ve yorumlayıcı istatistikler
- 4. Populasyon, örnek, değişken, parametre, ve istatistik terimleri
- 5. Veri tipleri

Şimdi düşünelim

•Acaba Türkiye'deki Siyah Alaca ineklerin 305 günlük süt verim ortalaması ne kadardır?

•Bu soruya cevap verebilmek için ne yapmalıyız?



Tek Grup Sınıf

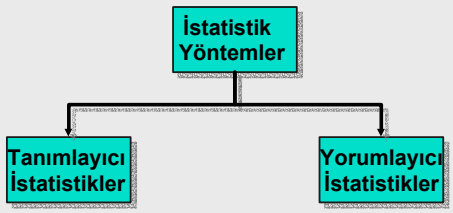
İstatistiğe Giriş

İstatistik Nedir?

- **1. Veri Bilimi**
- **2. Neyi içeriyor:**
  - Verilerin toplanması
  - Verilerin Sınıflandırılması
  - Verilerin özetlenmesi
  - Verilerin organize edilmesi
  - Verilerin analizi
  - Yorumlanması



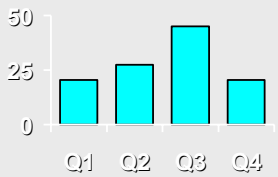
İstatistik Tipleri



Tanımlayıcı İstatistikler

Aşağıdaki basamakları içerir

- Verilerin toplanması
- Verilerin sunumu
- Verilerin karakterize edilmesi



Amaç

- Verileri tanımlamak

$\bar{X} = 30.5 \quad S^2 = 113$

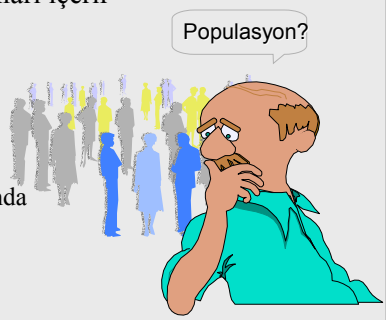
Yorumlayıcı İstatistikler

Aşağıdaki basamakları içerir

- Tahminleme
- Hipotez testi

Amaç

- Populasyonun özellikleri hakkında karar verme



Temel Terimler

İstatistikte analiz ve yorumlamada aşağıdaki terimlerin bilinmesi oldukça önemlidir.

Populasyon: Üzerinde durulan özellikle ilgili bütün elemanları içine alan elemanlar topluluğuna populasyon denir.

Parametre: Populasyonların elemanları üzerinde hesaplanan ortalama, varyans, regresyon katsayısı gibi ölçülere parametre denir.

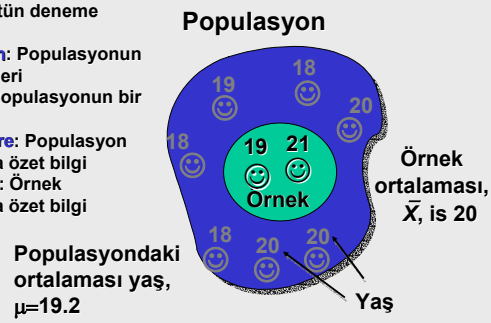
Örnek: Populasyondan şansa bağlı olarak çekilen ve populasyonu temsil etme özelliklerine sahip elemanlar topluluğuna örnek denir.

İstatistik: Örnek verilerinden hesaplanan ortalama, varyans, regresyon katsayısı gibi tahminleyicilere istatistik denir.

Parametreler Yunan alfabesi ile istatistikler ise Latin alfabesi ile sembolize edilirler.

Temel Terimler

1. **Populasyon:** **Populasyon: İlgi alanına giren bütün deneme üniteleri**
2. **Değişken: Populasyonun karakterleri**
3. **Örnek: Populasyonun bir bölümü**
4. **Parametre: Populasyon hakkında özet bilgi**
5. **İstatistik: Örnek hakkında özet bilgi**

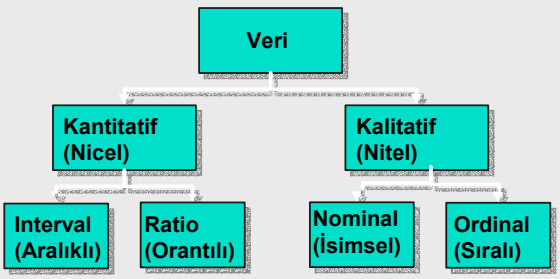


	İstatistik	Parametre
Ortalama	$\bar{X}$	μ (mü)
Varyans	S^2	σ^2 (sigma kare)
Standart sapma	S	σ (sigma)
Korelasyon katsayısı	r	ρ (rho, ro)

Veri Nedir?

- Gözlem, sayım ve ölçüm sonucu elde edilen bilgi, sembol ve rakamlara veri denir.
- Değişken: Gözlem, sayım ve ölçüm sonucu elde edilen verilerin atandığı terime değişken adı verilir.

Veri Tipleri



Değişkenlerin gruplandırılması ve özellikleri

Değişken	Alt seçenek	Değer Belirleme
Kalitatif	İsimsel (nominal)	Özelliğin isimsel sınıflara göre belirlenmesi
	Sıralı (Ordinal)	Özelliğin sıraya konulmuş biçimde elde edilmesi
Kantitatif	Aralık (Interval)	Değerlerin belirli aralıklara göre belirlenmesi
	Orantılı (Ratio)	Fiziksel ölçümlemeye göre değerlerin saptanması

Kalitatif
İsimsel (nominal) ölçeği

- Verilere isim verilerek sınıflandırma ve değerlendirme yapılıyorsa,
- Bu durumda aritmetik işlemler kullanılamaz
- İstatistik tekniklerin kullanımı sınırlıdır.
- **Örnek:** eşey bilgisi (erkek, dişi); Sigara içme durumu (içen, içmeyen); Saç rengi(siyah, sarı, kumral)

Kalitatif
Sıralı (ordinal) ölçeği

- Verilerdeki sınıflama **sıralama** şeklinde yapılıyorsa, bu tip ölçekler sıralı ölçeği olarak tanımlanır.
- Bu durumda sınıfın birinin diğerinden daha düşük veya daha önemli olduğu söylenebilir.
- İsimsel olana göre daha ileri bir tanımlama içerir fakat istatistik teknikler hala sınırlıdır.
- **Örnek:** Başarı durumu(zayıf, orta,iyi,pekiyi); gelir düzeyi(düşük, orta, yüksek); öğrenim sınıfları(1.,2.,3.ve 4. sınıf)

Kantitatif
Aralık (interval) ölçeği

- Bu tip verilerde yapı ordinal veriler gibidir. Aralarındaki fark interval verilerin eşit aralıklı sınıflandırılmasıdır.
- Veriler gerçek olmayan bir başlangıç noktasına göre belirlenirler. Aralık ölçeği, iki noktası belirlenen bir aralığı eşit bölümlere ayırarak oluşturulduğu gibi, gerçek olmayan bir noktadan itibaren sabit birimler bölünerek de oluşturulabilir. Fakat veriler birbirinin katı şeklinde ifade edilemezler.
- **Örnek:** Sıcaklık ölçümleri(Celcius, Fahrenheit, Reaumer termometreler); Tarih belirlemede kullanılan takvimler (miladi, hicri, rumi takvimler)

Kantitatif
Oran (Ratio) ölçeği

- Bu tip ölçeğin oluşturulmasında gerçek bir başlangıç noktası (sıfır noktası) vardır. Veriler birbirinin katı olarak ifade edilebilir.
- Diğer ölçeklere göre daha ileri bir tanımlamadır.
- İstatistik tekniklerin en yaygın bu ölçekteki verilerde kullanılır
- **Örnek:** Uzunluk ölçüleri; Ağırlık ölçüleri; Alan ölçüleri; Hacim ölçüleri; Zaman ölçüleri(sn, d, s, gün, yıl); Yoğunluk, basınç, ses ölçüleri.

Veri tipleri neden önemlidir?

- İstatistik hesaplamalar ve analizler, değişkenlerin belirli bir ölçekte almasını gerektirir.
- Saç renginin ortalamasını hesaplayabilir miyiz?
- Eğitim deneyimlerinin ortalamasını hesaplamak ne derece mantıklı olur?
- Ortalama, değişkenin aralık veya oran ölçekli olmasını gerektirir.

Değişken Tipleri

Kesikli: Ölçülen özellik kalitatif ise veya değerler sayı doğrusu üzerinde sadece belirli noktalara atanabiliyorsa bu tip değişkenlere kesikli değişken denir.

Örnekler:

Sağlık durumu (hasta-sağlam); Cinsiyet(kız-erkek);Başarı durumu(zayıf-orta-iyi-pekiyi);ürün kalitesi(1.kalite,2.kalite,vb)

Değişken Tipleri

Sürekli: Ölçülen özellik kantitatif ise ve değerler sayı doğrusu üzerinde belirli bir aralıktaki bütün noktalara atanabiliyorsa bu tip değişkenlere sürekli değişken denir.

Örnekler:

Boy uzunluğu, vücut ağırlığı, kan Ph değerleri, kan kolesterol ve şeker düzeyleri, hacim veya alan değerleri vb…

Şimdi karar verelim

Aşağıdaki özellikler kantitatif mi yoksa kalitatif mi?

- **1. Cinsiyet**
  - **Erkek, dişi**
- **2. Ağırlık**
  - **123, 140.2 vb.**
- **3. Arabanın hızı**
  - **78, 64, 45 vb.**
- **4. Sıcaklık**
  - **78, 64, 85 vb.**
- **5. kardeş sayısı**
  - **0-2, 3-5, 6+**
- **6. Alınan not**
  - **A, B, C vb.**
- **7. Alınan not**
  - **40, 70, 100 vb.**

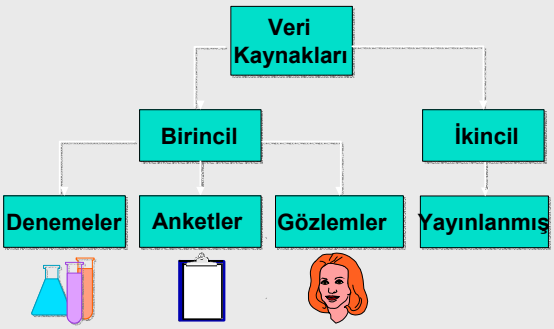
Tek Grup Sınıf

SORU

- Hayvancılıkta incelenen özellikleri ölçekler bakımından değerlendiriniz.

Verilerin Toplanması

Veri Kaynakları



Sorular

- 1. İstatistik nedir?
- 2. İstatistiğin kullanım alanları hakkında bilgi veriniz.
- 3. Tanımlayıcı ve yorumlayıcı istatistikler nelerdir? Farkını söyleyiniz.
- 4. Populasyon, örnek, değişken, parametre ve istatistik nedir?
- 5. Veri tipleri hangileridir?

Bugün...

- 1.Öğrendiğiniz en önemli bilgi hangisidir?
- 2.Öğretilenlerle ilgili bir sorunuz var mı?
  3.Bugün anlatılanları geliştirmek için herhangi bir öneriniz var mı?

Beklenenler

- Herkes çalışma alanı ile ilgili olarak bir veri seti bulacak.
- Bu veri seti sınıfta tanıtılacak
- Anlatılanlar çerçevesinde veri seti tartışılacak
- Hipotezler ortaya konacak
- Veri seti için analiz önerileri yapılacak.

HAFTAYA

Konu: Tanımlayıcı istatistikler hatırlanacak ve SPSS'te elde edilmeleri uygulanacak.

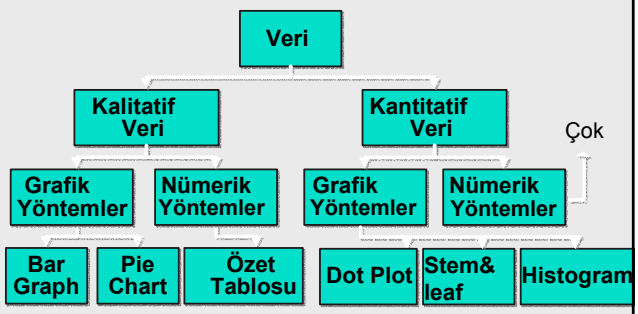
İstenen: İstatistik kitaplarından bu konu okunacak

Değişkenlerin Tanımlanmasında Kullanılan İstatistikler
(Descriptive Statistics)

Amaç

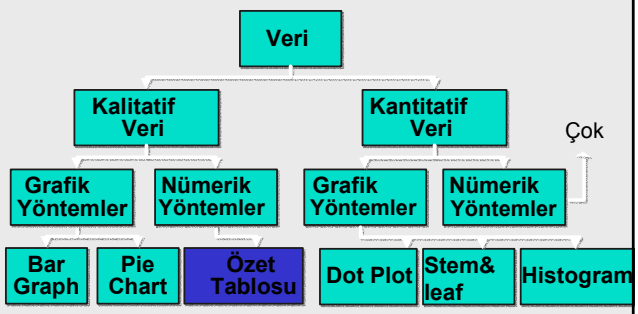
- 1. Grafik ile tanımlama ve yorumlama
- 2. Verileri tanımlamada kullanılan nümerik yöntemler
- 3. Nümerik yöntemlerle analiz
- 4. Gerçek değer, tanımlayıcı tekniklerle ne düzeyde tanımlanabiliyor?

Verilerin Tanımlanması



Kalitatif Verilerin Tanımlanması

Verilerin Tanımlanması



Özet Tablosu

•1.Grupları listele & gruptaki eleman sayısını belirle

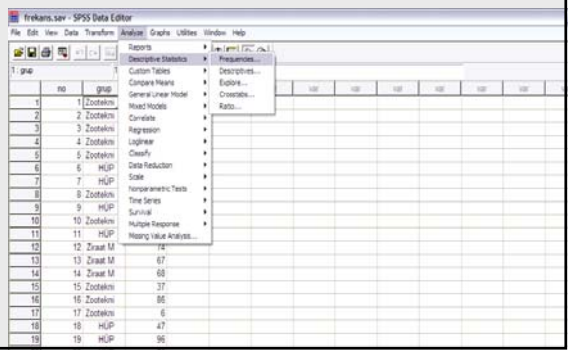
•2.Gruba ait verilerin çetelesini tut

•3.Verilerin frekanslarını (tekrarlanma sayılarını), toplam veriler içindeki %'sini veya her ikisini birden kullanarak özetle

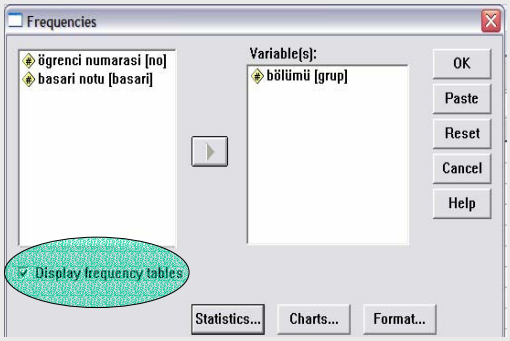
Sınıflar	Frekans		Nisbi frekans
	# öğrenci		Oran
Ziraat Müh.	130	Çetele	.65
Zootekni	20	𝍸 𝍸	.10
HUP	50	𝍸 𝍸	.25
Total	200		1.00

- Frekans.sav dosyasında farklı öğrenci grupları ve başarı notları bulunmaktadır.
- Gruplarda kaçar öğrenci vardır? SPSS ile bulunuz.

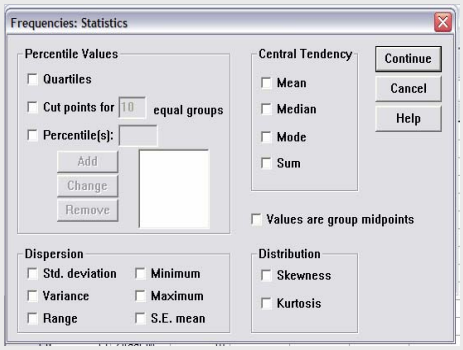
Frekansların Belirlenmesi



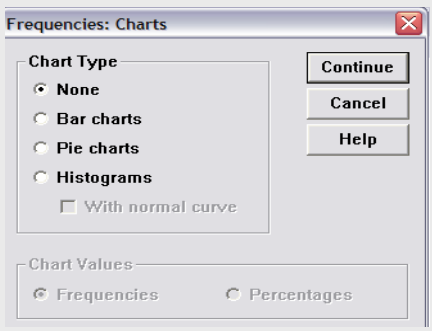
Değişken atama

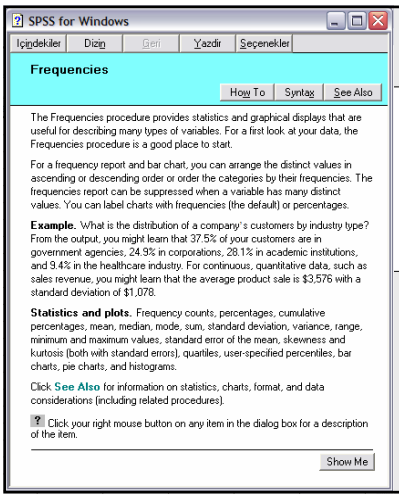


"Statistics" Seçenekleri



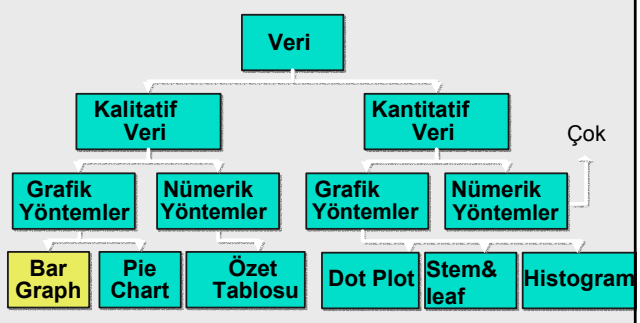
"Charts" Seçenekleri



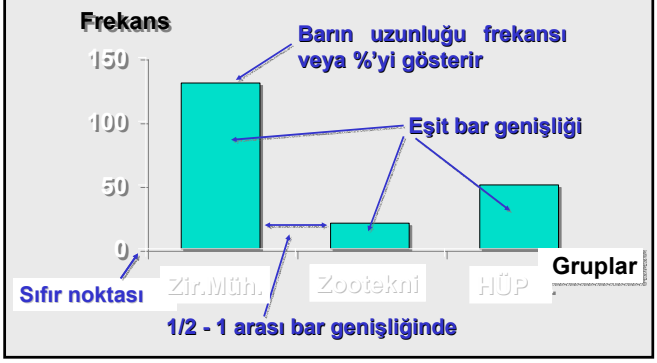


Frekanslar penceresinden "HELP"

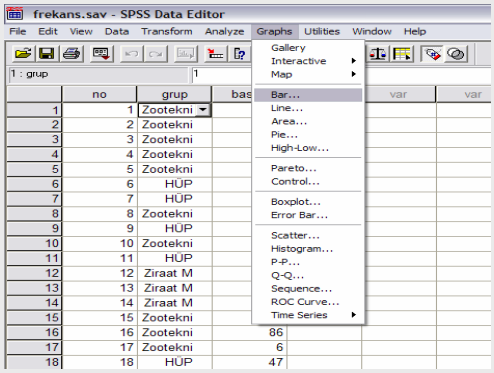
Verilerin Tanımlanması

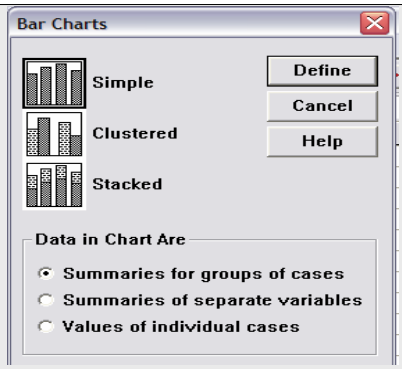


Bar Grafik

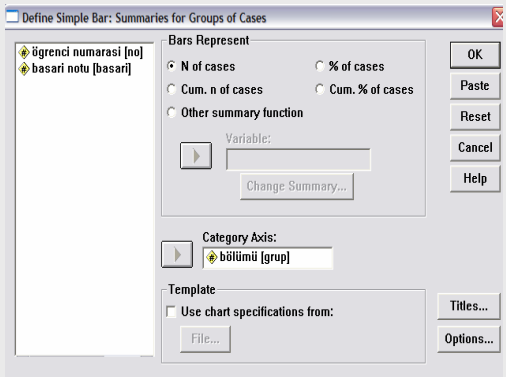


Bar Grafik

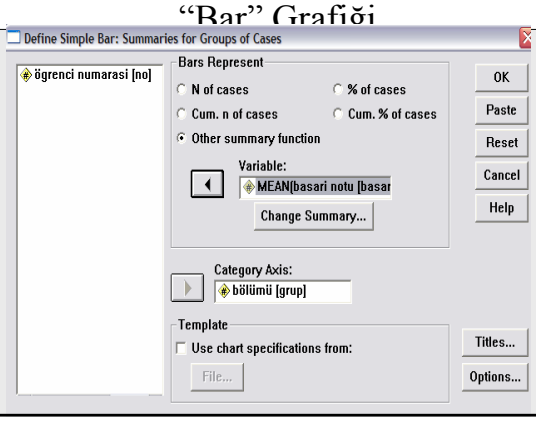




Bar Grafik Tanımlamaları



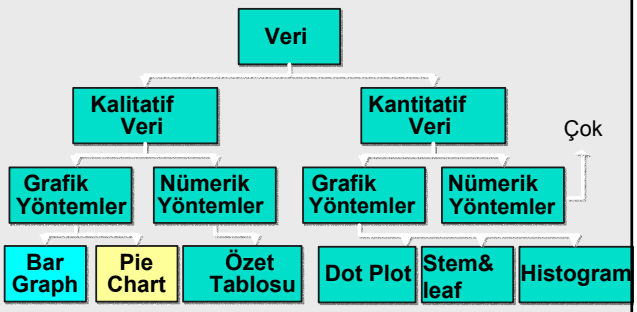
Başarı Ortalamalarının Gruplara Göre "Bar" Grafiği



Başarı durumlarının gruplara göre değişimi

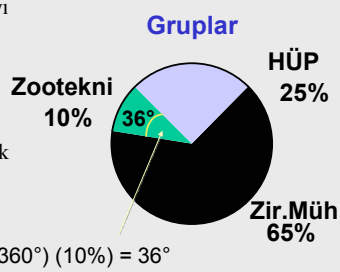


Verilerin Tanımlanması



Pie Chart

- 1.Grup içindeki toplam payı belirler
- 2.Nisbi farklılıkları göstermek için kullanışlıdır.
- 3.Toplam açı 360°'dır. Buna göre açılar belirlenerek yüzde olarak miktar gösterilir)

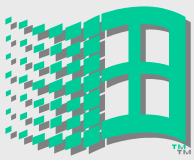


(360°) (10%) = 36°

Şimdi düşünelim

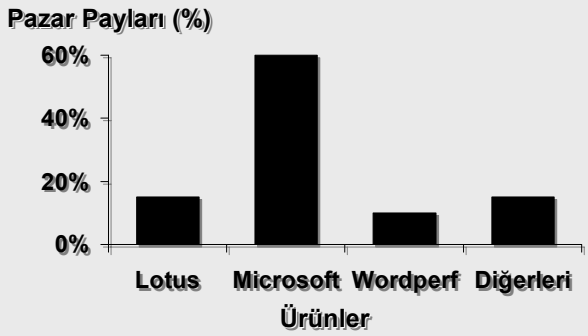
- Bir piyasa analizi yapan kişi olsaydınız, belirli bir yılda Windows program üreticilerinin pazar paylarını göstermek isteseydik bu bilgileri nasıl sunardık?
- Verileri bir **bar** grafik ve **pie** chart ile tanımlayınız.

	Pazar payı **(%)**
Lotus	15
Microsoft	60
WordPerfect	10
Diğerleri	15

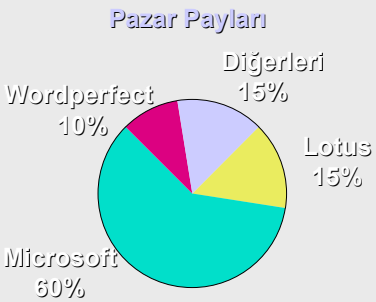


Tek Grup Sınıf

Bar Chart Çözümü



Pie Chart Çözüm



Pie Grafik

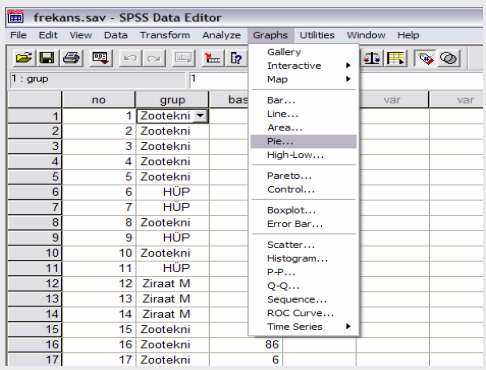
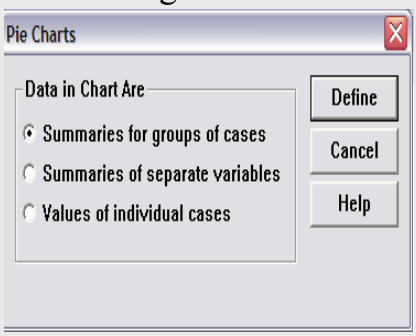
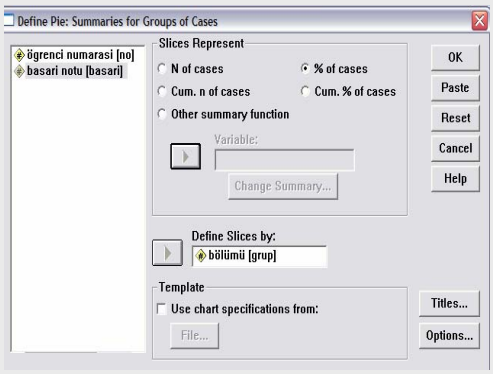
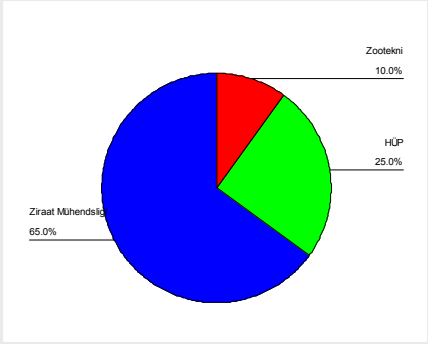


Chart Alanındaki Verilerin Yapısının Sorgulanması



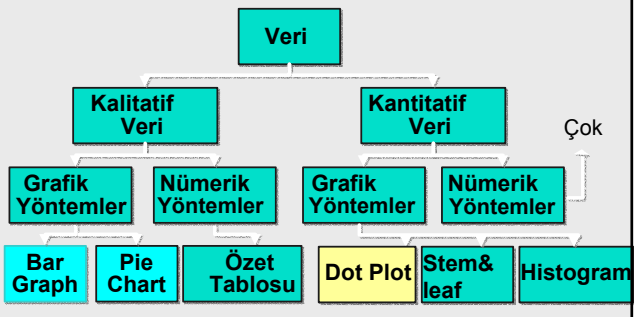
Tanımlamalar





Kantitatif Verilerin Tanımlanması

Verilerin Tanımlanması

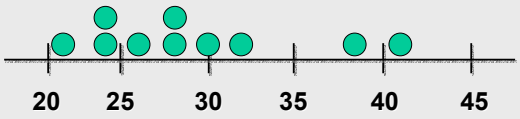


Dot Plot

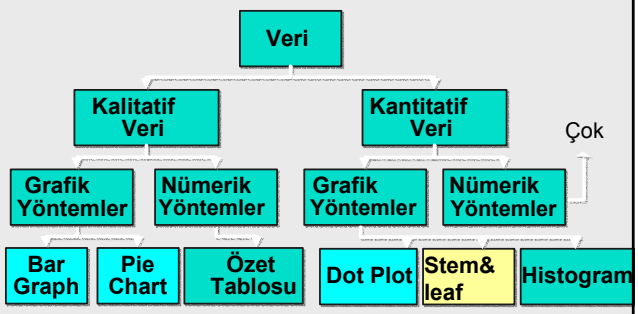
- 1. Aynı değerleri gruplayarak bir araya getirme sonucu verilerin özetlenmesi
- 2. Veriler yatay eksende noktalar ile gösterilir.

Dot Plot

- 1. Aynı değerleri gruplayarak bir araya getirme sonucu verilerin özetlenmesi
- 2. Veriler yatay eksende noktalar ile gösterilir.
- 3. Veri: 21, 24, 24, 26, 27, 27, 30, 32, 38, 41



Verilerin Tanımlanması



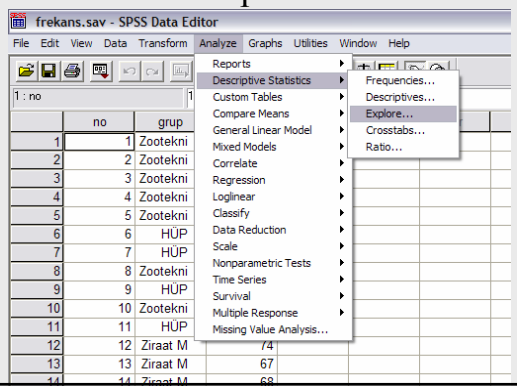
Stem-and-Leaf Display

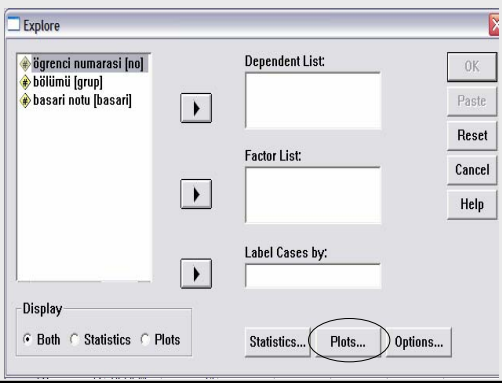
- 1. Her gözlemi gövde ve yaprak değeri olarak ikiye ayır.
  - Stem (gövde) değeri sınıf değerini tanımlar
  - Yaprak değeri ise frekans değerini tanımlar

2 | 144677
3 | 028
4 | 1

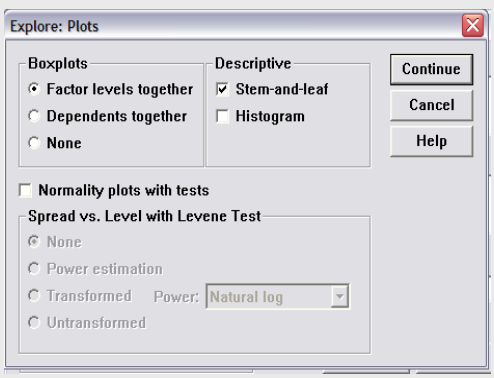
2. Data: 21, 24, 24, 26, 27, 27, 30, 32, 38, 41

Tanımlayıcı İstatistiklerden "Explore"





Stem and Leaf

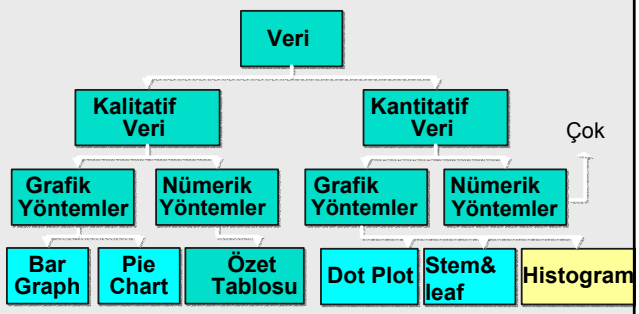


basari notu Stem-and-Leaf Plot

Frequency	Stem &	Leaf
18.00	0 .	233444556667888999
16.00	1 .	1112234666779999
21.00	2 .	000001222344555666799
23.00	3 .	00222223344456778888899
22.00	4 .	0011123345566667788899
12.00	5 .	133556677789
21.00	6 .	001113333444567888999
24.00	7 .	011112334444556667778899
23.00	8 .	001123333455556667778889
19.00	9 .	0023344466778889999
1.00	10 .	0

Stem width: 10
Each leaf: 1 case(s)

Verilerin Tanımlanması



Histogram

- 1. Benzer değerleri aynı sınıf içine alıp, verileri grafikle göstererek özetlemeye denir.
- 2. Frekansları (sayıları) veya nisbi frekansları (oranları) da kullanabiliriz.
- 3. Öncelikle frekans dağılış tablosunun oluşturulması gerekir.

Frekans Dağılış Tablosunun Oluşturulması

- 1. "Range"i (maksimum ile minimum arası farkı) belirle
- 2. Sınıf sayısını seç
  - *Genellikle* 5 - 15 arasındadır
- 3. Sınıf aralığını (genişliğini) hesapla
4. Sınıfların alt ve üst sınır noktalarını belirle
5. Sınıf orta değerlerini hesapla
- 6. Sınıflara ait gözlemleri say ve sınıflara ata

Frekans Dağılış Tablosu Örnek

Ham veriler: **24**, **26**, **24**, **21**, **27**, **27**, **30**, 41, **32**, 38

Sınıflar	Orta noktalar	Frekanslar
15 ≤ X < 25	20	3
25 ≤ X < 35	30	5
35 ≤ X < 45	40	2

Sınırlar

(Üst + Alt Sınır) / 2

Nisbi Frekans ve % Dağılış Tabloları

Nisbi Frekans Dağılışı

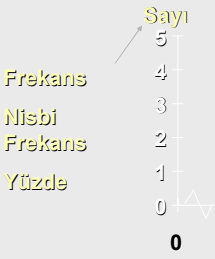
Sınıflar	Oran
15 ≤ X < 25	.3
25 ≤ X < 35	.5
35 ≤ X < 45	.2

Yüzde Dağılış

Sınıflar	%
15 ≤ X < 25	30.0
25 ≤ X < 35	50.0
35 ≤ X < 45	20.0

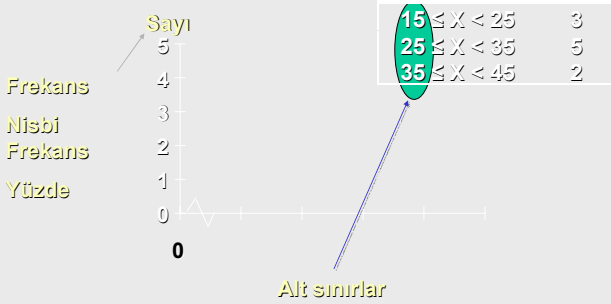
Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2



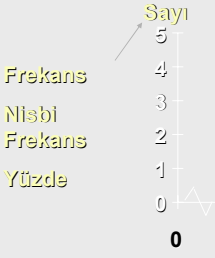
Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2



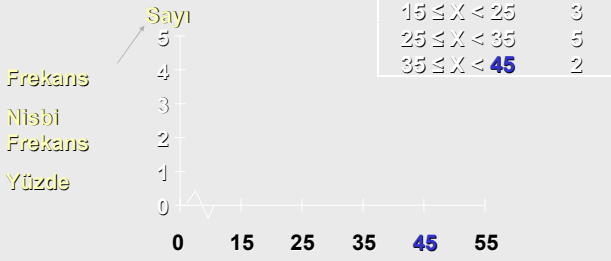
Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2



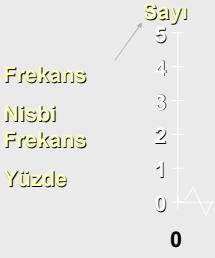
Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2



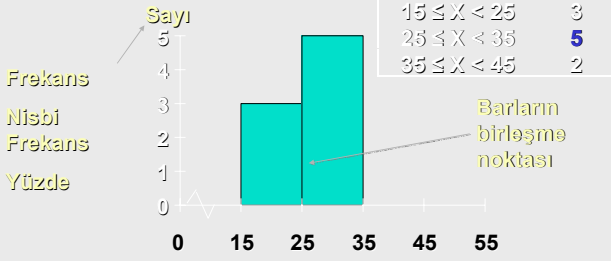
Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2

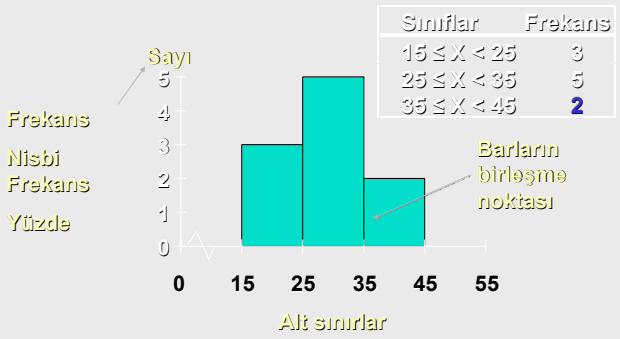


Histogram

Sınıflar	Frekans
15 ≤ X < 25	3
25 ≤ X < 35	5
35 ≤ X < 45	2

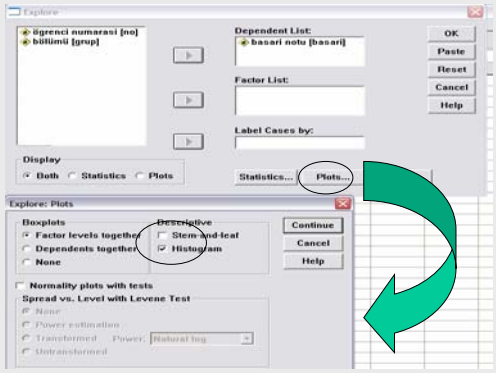


Histogram

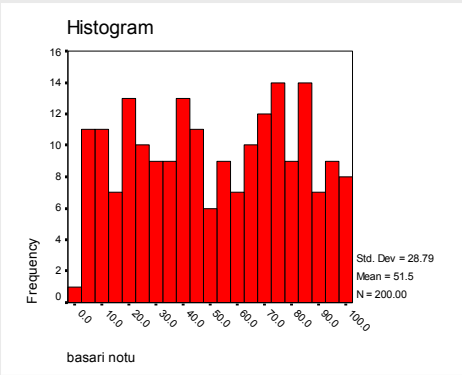


Histogramın Tanımlanması

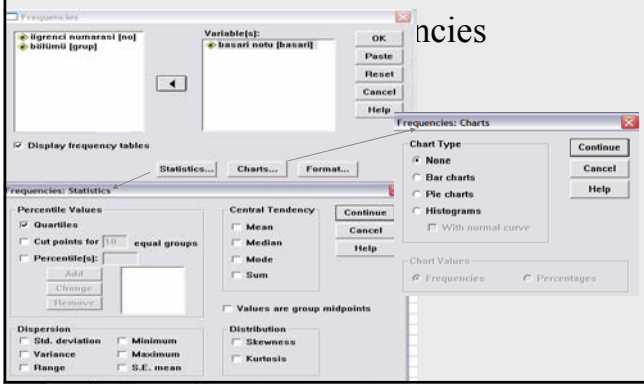
Desc.Stat./Explore/Plot/Histogram



Oluşturulan Histogram

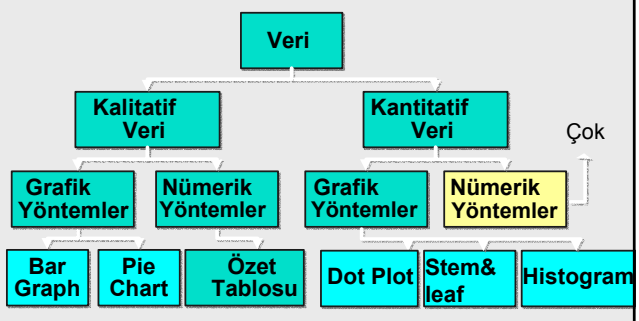


Bar, Pie ve Histogramlar aşağıdaki şekilde de elde edilebilir



Kantitatif Veriler için Nümerik Yöntemler

Verilerin Tanımlanması



Düşünelim



... İşçiler, çalışanların düşün ücret aldığını, çoğunun sadece 20,000 $ kazandığını belirtmişler

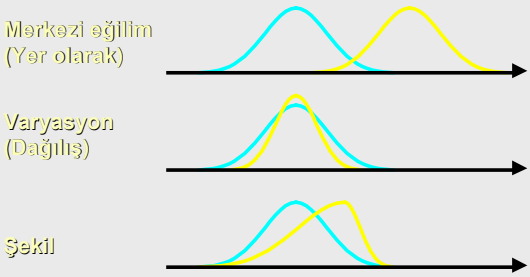
... Patron ise ortalama ücretin 70,000 $ olduğunu iddia etmiştir.

Tek → Grup → Sınıf

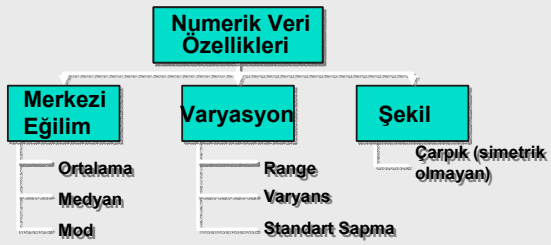
Standart Notasyon

Ölçüm	Örnek	Populasyon
Ortalama(Mean)	$\bar{X}$	μ
Standart Sapma Stand. Deviation	S	σ
Varyans	S^2	σ^2
Gözlem sayısı	n	N

Kantitatif Verilerin Özellikleri



Kantitatif Verilere Uygulanan Numerik Yöntemler



Merkezi Eğilim

Ortalama

- 1. Merkezi eğilimi ölçer
- 2. En yaygın ölçüdür
- 3. 'Denge noktası' rolü oynar
- 4. Ekstrem değerlerden ('outliers') etkilenir
- 5. Hesaplanışı (örnek ortalaması)

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

Ortalama: Örnek

- Veriler: 10.3 4.9 8.9 11.7 6.3 7.7

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6}{6}$$

$$= \frac{10.3 + 4.9 + 8.9 + 11.7 + 6.3 + 7.7}{6}$$

$$= 8.30$$

Median

1. Merkezi eğilimin bir ölçüsüdür
2. Gözlemler büyüklüğüne göre sıralandığında ortada kalan değerdir.
   - Gözlem sayısı tek ise sıralamada ortada kalan değerdir
   - Gözlem sayısı çift ise sıralamada ortada kalan iki değerin ortalamasıdır.
3. Sıralamada medyanın pozisyonu

$$\text{Pozisyon noktası} = \frac{n+1}{2}$$

4. 4. Ekstrem değerlerden etkilenmez

Medyanın Hesaplanması Gözlem sayısı tek

- Ham Veri: 24.1 22.6 21.5 23.7 22.6
- Sıralanmış: 21.5 22.6 **22.6** 23.7 24.1
- Pozisyon: 1 2 **3** 4 5

♦

$$\text{Pozisyon Noktası} = \frac{n+1}{2} = \frac{5+1}{2} = 3.0$$

Medyan= 22.6

Medyanın Hesaplanması Gözlem sayısı çift

- Ham Veri: 10.3 4.9 8.9 11.7 6.3 7.7
- Sıralanmış: 4.9 6.3 **7.7** **8.9** 10.3 11.7
- Pozisyon: 1 2 **3** **4** 5 6

♦

$$\text{Pozisyon Noktası} = \frac{n+1}{2} = \frac{6+1}{2} = 3.5$$

$$\text{Medyan} = \frac{7.7 + 8.9}{2} = 8.30$$

Mod

- 1. Merkezi eğilimin bir ölçüsüdür.
3. Genellikle elde edilen bir değerdir.
4. Ekstrem değerlerden etkilenmez.
- 4. Verilerde hiç mod olmayabileceği gibi çok sayıda mod değeri de olabilir.
- 5. Kantitatif ve kalitatif veriler için kullanılabilir.

Mod Örnekler

•Mode Yok

Ham Veriler: 10.3 4.9 8.9 11.7 6.3 7.7

•Tek Mod

Raw Data: 6.3 4.9 8.9 6.3 11.7 11.7

•1'den fazla mod

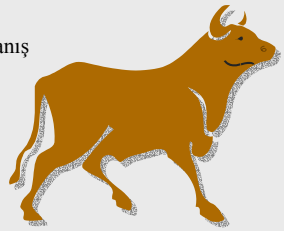
Ham Veri: 21 28 28 41 42 42

Düşünelim

•Mali analizler yapan bir kişi olduğunuzu kabul edelim.

•Borsadaki damızlık fiyatlarının kapanış değerlerinin aşağıdaki şekilde topladığınızı kabul edelim:
•17, 16, 21, 18, 13, 16, 12, 11.

•Merkezi eğilim olarak damızlık fiyatlarını tanımlayınız.



Alone Group Class

Merkezi Eğilim

- **Ortalama**

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + \cdots + X_8}{8}$$

$$= \frac{17 + 16 + 21 + 18 + 13 + 16 + 12 + 11}{8}$$

$$= 15.5$$

Merkezi Eğilim

- **Medyan**
- **Ham Veri**: 17 16 21 18 13 16 12 11
- Sıralanmış: 11 12 13 16 16 17 18 21
- Pozisyon: 1 2 3 4 5 6 7 8

$$\text{Pozisyon Noktası} = \frac{n+1}{2} = \frac{8+1}{2} = 4.5$$

$$\text{Medyan} = \frac{16 + 16}{2} = 16$$

Merkezi Eğilim

Mod

Ham Veri: 17 16 21 18 13 16 12 11

Sıralanmış: 11 12 13 16 16 17 18 21

Merkezi Eğilim Ölçülerinin Özeti

Ölçü	Eşitlik	Tanımlama
Ortalama	$\Sigma X_i / n$	Denge noktası
Medyan	$\frac{(n+1)}{2}$ Pozisyon	Sıralandığında ortanca değer
Mod	yok	En çok tekrarlanan

Varyasyon Değişkenlik

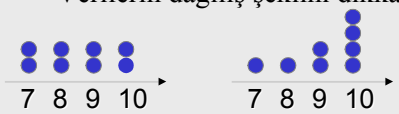
Range

- 1. Dağılışın ölçüsü

1. En büyük değer ile en küçük değer arasındaki farklılık

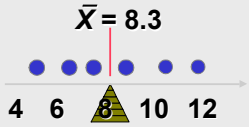
- $\text{Range} = X_{\text{En büyük}} - X_{\text{En küçük}}$

3. Verilerin dağılış şeklini dikkate almaz



Varyans ve Standart Sapma

- 1. Dağılışın ölçüleridir
- 2. En yaygın ölçülerdir
- 3. Dağılışın şeklini dikkate alır
- 4. Ortalama ($\bar{X}$ veya μ) etrafındaki varyasyonu gösterin



Örnek Varyansı

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

$$= \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1}$$

Örnek Varyansı

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Payda n - 1!
(Populasyon varyansı için ***N*** kullanın)

$$= \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1}$$



Örnek Standart Sapma

$$S = \sqrt{S^2}$$

$$= \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

$$= \sqrt{\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1}}$$

Varyans Örnek:

- Ham Veri: 10.3 4.9 8.9 11.7 6.3 7.7

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \qquad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = 8.3$$

$$S^2 = \frac{(10.3 - 8.3)^2 + (4.9 - 8.3)^2 + \cdots + (7.7 - 8.3)^2}{6-1}$$

$$= 6.368$$

Şimdi düşünelim

- Mali analizler yapan bir kişi olduğunuzu kabul edelim.
- Borsadaki damızlık fiyatlarının kapanış değerlerinin aşağıdaki şekilde topladığınızı kabul edelim:
- 17, 16, 21, 18, 13, 16, 12, 11.
- Damızlık fiyatlarının varyans ve standart sapması nedir?



Tek Grup Sınıf

Varyasyon

Örnek varyansı

Ham Veri: 17 16 21 18 13 16 12 11

$$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1} \qquad \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = 15.5$$

$$S^2 = \frac{(17-15.5)^2 + (16-15.5)^2 + \cdots + (11-15.5)^2}{8-1}$$

$$= 11.14$$

Varyasyon

- **Örnek Standart Sapması**

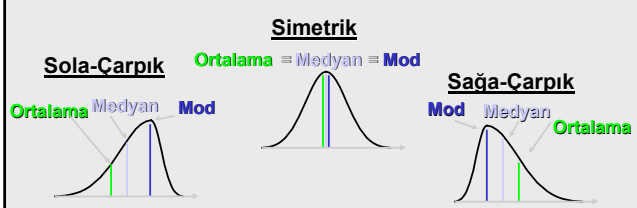
$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}} = \sqrt{11.14} = 3.34$$

Varyasyon Ölçülerinin Özeti

Ölçüler	Eşitlik	Tanımlama
Range	$X_{\text{Enbüyük}} - X_{\text{En küçük}}$	Toplam yayılma
Interquartile Range	$Q_3 - Q_1$	Ortadaki % 50'nin yayılımı
Standart Sapma (Örnek)	$\sqrt{\frac{\sum (X_i - \bar{X})^2}{n-1}}$	Örnek ortalamasına ait dağılış
Standart Sapma (Populasyon)	$\sqrt{\frac{\sum (X_i - \mu)^2}{N}}$	Populasyon ortalamasına Ait dağılış
Varyans (Örnek)	$\frac{\Sigma (X_i - \bar{X})^2}{n-1}$	Örnek ortalamasına ait karesi alınmış dağılış

Şekil

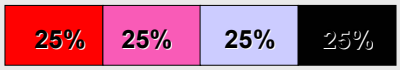
- 1. Verilerin nasıl dağıldığını tanımlar
- 2. "Skewness" (simetri) ile ölçülür.



Kartiller ve Box Plots

Kartiller

1. Merkezi olmayan eğilimi ölçer
2. Sıralanmış verileri 4 kartile ayıralım.



- 3. i'nci kartilin pozisyonu

Pozisyon noktası $Q_i = \frac{i \cdot (a+1)}{4}$

Kartil (Q_1) Örnek

- Ham veri: 10.3 4.9 8.9 11.7 6.3 7.7
- Sıralanmış: 4.9 **6.3** 7.7 8.9 10.3 11.7
- Pozisyon: 1 **2** 3 4 5 6

Q_1 Pozisyon $= \frac{1 \cdot (a+1)}{4} = \frac{1 \cdot (a+1)}{4} = 1.75 \cong 2$

$Q_1 = 6.3$

Kartil (Q_2) Örnek

- Ham veri: 10.3 4.9 8.9 11.7 6.3 7.7
- Sıralanmış: 4.9 6.3 **7.7** **8.9** 10.3 11.7
- Pozisyon: 1 2 **3** **4** 5 6

Q_2 Pozisyon $= \frac{2 \cdot (a+1)}{4} = \frac{2 \cdot (a+1)}{4} = 3.5$

$Q_2 = \frac{7.7 + 8.9}{2} = 8.3$

Kartil (Q_3) Örnek

- Ham veri: 10.3 4.9 8.9 11.7 6.3 7.7
- Sıralanmış: 4.9 6.3 7.7 8.9 **10.3** 11.7
- Pozisyon: 1 2 3 4 **5** 6

Q_3 Pozisyon $= \frac{3 \cdot (a+1)}{4} = \frac{3 \cdot (a+1)}{4} = 5.25 \cong 5$

$Q_3 = 10.3$

Kartiller arası (Interquartile) Range

- 1. Dağılışın ölçüsü
- 2. Ortadaki dağılış ölçüsü olarak da bilinir
3. Üçüncü ile birinci kartik arasındaki farktır

   Interquartile Range $= Q_3 - Q_1$
4. Orta % 50'nin dağılışıdır.
5. Ekstrem değerlerden etkilenmez

Düşünelim



- Mali analizler yapan bir kişi olduğunuzu kabul edelim.
- Borsadaki damızlık fiyatlarının kapanış değerlerinin aşağıdaki şekilde topladığınızı kabul edelim:
- **17, 16, 21, 18, 13, 16, 12, 11.**

- Damızlık fiyatlarının **Q_1 ve Q_3** kartilleri ile **interquartile** range'ni bulunuz?

Tek Grup Sınıf

Kartiller

Q_1

Ham veri:	17	16	21	18	13	16	12	11
Sıralanmış:	11	12	13	16	16	17	18	21
Pozisyon:	1	2	3	4	5	6	7	8

$$Q_1 \text{Pozisyon} = \frac{1 \cdot (a) + 1(f)}{4} = \frac{1 \cdot (a) + 1(f)}{4} = 2.5$$

$$Q_1 = 12.5$$

Kartil

Q_3

Ham veri:	17	16	21	18	13	16	12	11
Sıralanmış:	11	12	13	16	16	17	18	21
Pozisyon:	1	2	3	4	5	6	7	8

$$Q_3 \text{ Pozisyon} = \frac{3 \cdot (a) + 1(f)}{4} = \frac{3 \cdot (a) + 1(f)}{4} = 6.75 \cong 7$$

$$Q_3 = 18$$

Interquartile Range

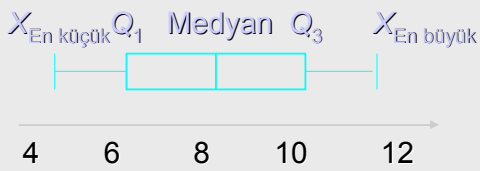
Interquartile Range

Ham veri:	17	16	21	18	13	16	12	11
Sıralanmış:	11	12	13	16	16	17	18	21
Pozisyon:	1	2	3	4	5	6	7	8

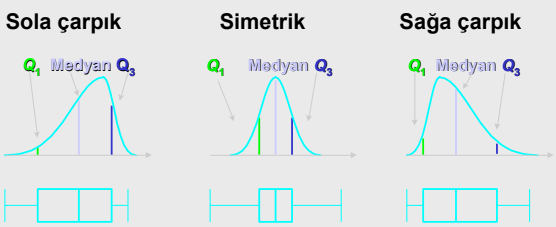
$$\text{Interquartile Range} = Q_3 - Q_1 = 18.0 - 12.5 = 5.5$$

Box Plot

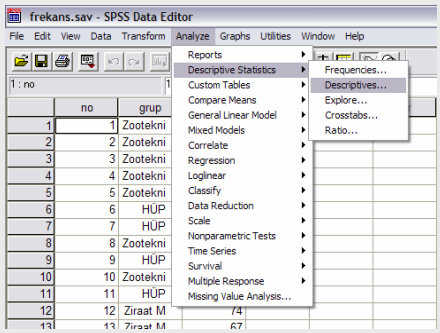
- 1. 5 özet bilgiyi kullanarak verilerin grafikle gösterilmesidir.

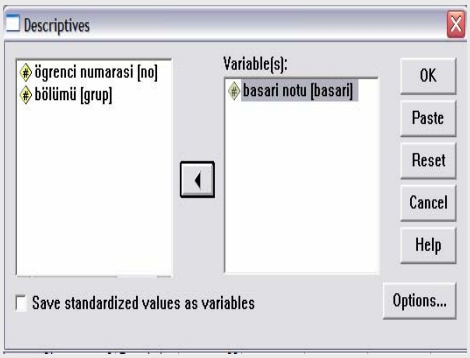


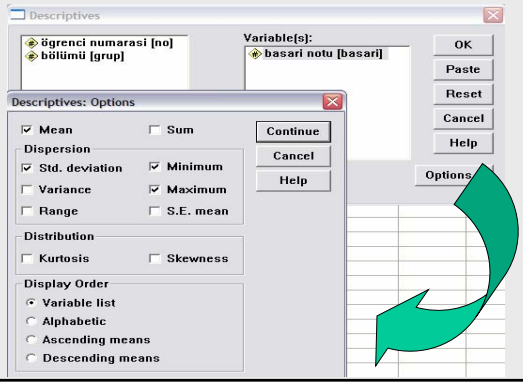
Şekil & Box Plot



Tanımlayıcı İstatistikler Desc.Stat./Descriptives





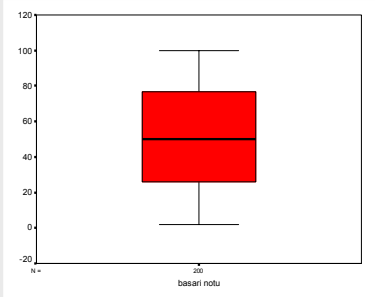


Sonuçlar

Descriptives

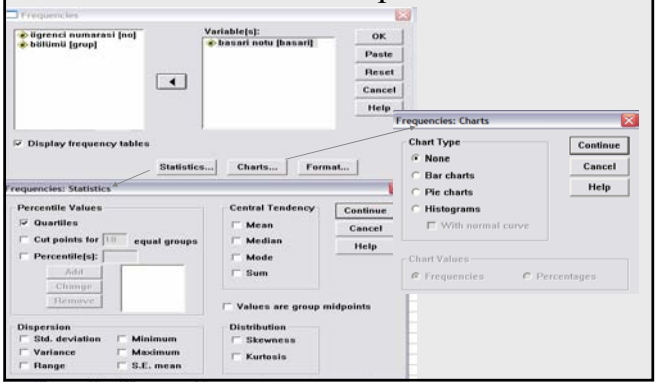
			Statistic	Std. Error
basari notu	Mean		51.53	2.035
	95% Confidence Interval for Mean	Lower Bound	47.51	
		Upper Bound	55.54	
	5% Trimmed Mean		51.55	
	Median		50.00	
	Variance		828.633	
	Std. Deviation		28.786	
	Minimum		2	
	Maximum		100	
	Range		98	
	Interquartile Range		50.75	
	Skewness		-.030	.172
	Kurtosis		-1.237	.342

Box-Plot Çizimi
Desc.Stat./Explore/Plot/Histogram



Histogram istenince sunuluyor.

Kartiller ve diğer bilgiler
Desc.Stat./Frequencies



Tanımlayıcı Tekniklerle Gerçeklerin Çarpıtılması

Verilerin Sunulmasında Yapılan Hatalar

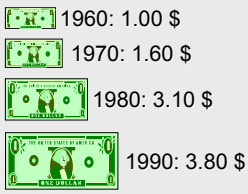
- 1. Chart'ların hatalı kullanımı
- 2. Veri gruplarına ait sonuçların karşılaştırılabilir olmaması
- 3. Dikey eksenin sıkıştırılması
- 4. Yatay eksende sıfır noktasının bulunmaması



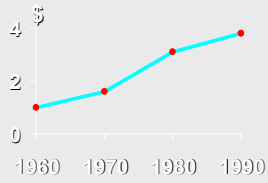
Chart'ların hatalı kullanımı

Kötü sunum

Minimum Ücret

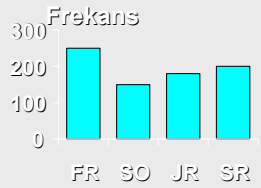


İyi Sunum

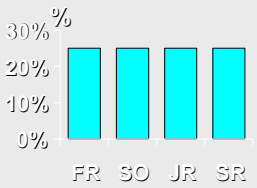


No Relative Basis

Kötü Sunum



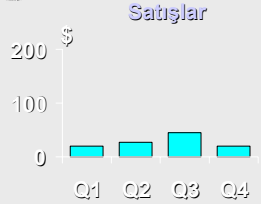
İyi Sunum



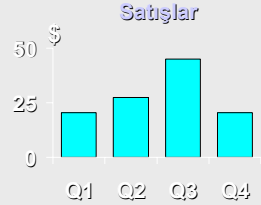
Her sınıftaki gözlem sayısının farklı olduğunu kabul edersek

Dikey eksenin Sıkıştırılması

Kötü Sunum

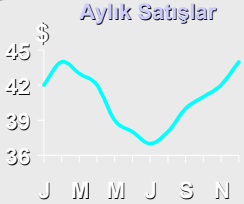


İyi Sunum

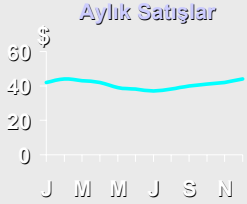


Yatay eksende sıfır noktasının olmaması

Kötü Sunum



İyi Sunum



Bu derste ele alınanlar:

- 1. Grafik sunuşların oluşturulması ve yorumlanması
- 2. Verileri tanımlanmasında kullanılan numerik yöntemler
- 3. Numerical yöntemleri kullanarak verilerin analizi
- 4. Tanımlayıcı teknikleri kullanarak gerçeklerin nasıl çarpıtıldığı

Hipotez Testi

Karar:	*Gerçek*	
	H_0 Doğru	***H_0 Yanlış***
H_0 Red	I Tip Hata (alfa hata)	Doğru Karar
H_0 Kabul	Doğru Karar	II Tip Hata (beta hata)

Farklılıkların Testi

Tek Örnek Testleri

Amaç

- SPSS kullanarak bir grubun ortalamasını, hipotezde belirtilen ve populasyona ait olduğu düşünülen değer ile karşılaştırılması;
- Serbestlik derecesi kavramının açıklanması ve hesabı;
- t-dağılışı ile z-dağılışının incelenmesi ve hipotez testinde kullanımı;
- SPSS sonuçlarının yorumlanması;

Farklılıklar için Tek – örnek testi

- Problem: Yazın çalışmalarda öğrenciler 3000 $'dan fazla kazanıyorlar mı?
- Sıfır hipotezi: Bir öğrenci genellikle 3000 $ veya daha az para kazanabilir.
  H_0: $\mu \leq 3{,}000$ $
- Acaba 3000 $'dan daha fazla para kazanılmış mıdır?
- H_1: $\mu > 3{,}000$ $
- Hipotez testi, örnek istatistiği ile populasyon parametresi arasındaki fark için yapılmaktadır.
- Anket yapılan öğrenci sayısı, n = 100
  - Kazanca ait örnek istatistikleri :
    - x = 3,150 $
    - s_x = 825 $
  - Dikkate alacağımız risk düzeyi: α = 5% (0.05)
- Testimiz: **Z test** veya one-sample t-test

Tek örnek Z testinin Formülü:

- Ortalama için: $z = \dfrac{\bar{x} - \mu_H}{s_{\bar{x}}}$

$\bar{x}$ = örnek ortalaması

μ_H = hipotezde beklenen ortalama

$s_{\bar{x}}$ = ortalamanın standart hatası = $\sigma / \sqrt{n}$

Fakat biz ortalamanın standart hatasını bilmiyoruz.
Bunun yerine örneğin standart sapmasını kullanıyoruz..

Hesaplama:

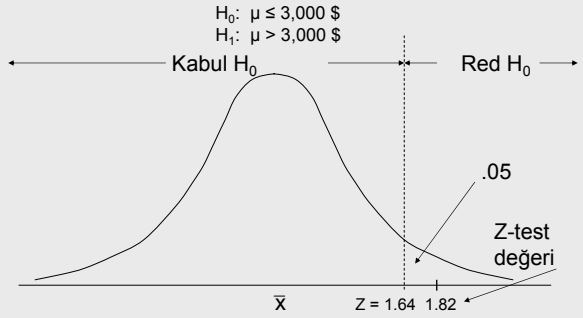
$$z = \frac{\bar{x} - \mu_H}{s_{\bar{x}}}$$

z - değeri:

$$\frac{3150 - 3000}{\frac{825}{\sqrt{100}}} = \frac{150}{82.5} = 1.82 \text{ (test istatistiği)}$$

Hesaplanan bu *z* - değeri normal dağılışa uyar mı?
Verilen z-değerine göre H_0 'ı kabul mu edeceğiz?
Yoksa red edip H_1'ri mi kabul edeceğiz?

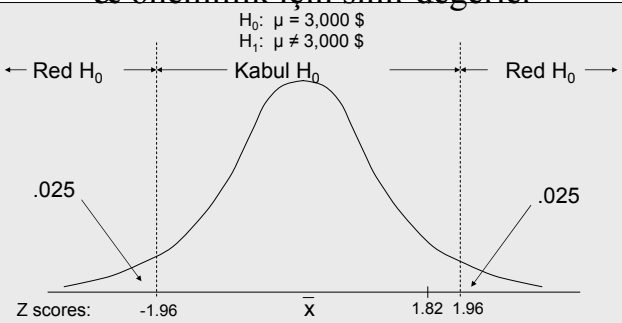
Red ve kabul alanlarının belirlenmesi ve yorum



Tek yönlü veya Çift yönlü hipotezler?

- Problem: Bir önceki örneği çift yönlü hale getirirsek:
- H_0: μ = 3,000 $
- H_1: μ ≠ 3,000 $
- Riskimiz hala 5% olsun. (Yani 100 karardan 5'inde hata yapmayı normal kabul edersek),
- Fakat bu durumda %5'in yarısı bir kuyrukta diğer yarısı ise diğer kuyrukta olacaktır. Çünkü biz 3000 $'dan sapmanın hem pozitif hem negatif yöndeki sapmaları ile ilgileniyoruz.

Çift yönlü hipotez, p-value, & önemlilik için sınır değerler



Z-testi veya bir örnek t-testi?

- t-testi, t-skorlarını kullanırken z-teski z-skorlarını dikkate alır. Aralarındaki fark dağılışlarının farkıdır. T-skorları t-dağılışına , z-skorları Z-dağılışına sahiptir).
- n < 30 olduğunda t-dağılışı ve t-testi
- n > 30 olduğunda ise t-dağılışı normal dağılışa çok yaklaşır ve büyük n değerlerinde aynıdır.
- Bu nedenle SPSS sadece t-skorlarına göre analiz eder.

SPSS test sonuçları:

One-Sample Statistics — Örnek ortalaması

	N	Mean	Std. Deviation	Std. Error Mean
TV violence opinion	94	3.45	1.284	.132

SPSS karşılaştırılan değeri gösterir. Burada 3

Güven aralıkları

One-Sample Test

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
TV violence opinion	3.375	93	.001	.45	.18	.71

Hesaplanan T-test istatistiği

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Actual hours	94	14.29	9.101	.939

One-Sample Test

	Test Value = 16					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
Actual hours	-1.825	93	.071	-1.71	-3.58	.15

Yorumlayıcı İstatistikler

Çok sayıda yöntem vardır.
Herbiri belirli koşullarda önerilen ve aşağıdaki veri yapılarına göre önerilmektedir:

- Ölçümün *Skala'sına*
  *(*nominal, ordinal, interval, ratio)
- Örnek dağılışının *Şekli*
- Denemenin *Dizaynına*
  Bağımsız değişken sayısı
  Her bağımsız değişkenin düzey sayısı
  Bağımlı değişken sayısı

Gruplar arası *t* test

Ne zaman uygundur ?
Bağımsız iki grup ortalamasının karşılaştırılması

- En azından "interval skala"da ölçümler
- Deneme üniteleri şansa bağlı örneklerdir.
- Populasyon değerleri normal dağılış göstermektedir.

Gruplar arası *t* test

Hesaplamalarda hangi değerler dikkate alınmaktadır?

- Ortalama farklılığı
- Varyans

Hesaplanışı:

$$t = \frac{\overline{Y}_1 - \overline{Y}_2}{SE_{\overline{Y}_1 - \overline{Y}_2}}$$

Örnek ortalama farklılığı
Değişkenlik (Standart hata)

Değişkenlik ve örnek büyüklüğünü dikkate alınarak, büyük değerlerin, gruplar arasında daha büyük farklılıklar olduğunu gösterdiğini söyleyebiliriz.

Gruplar arası *t* test

Standart hatanın hesaplanışı:

1) Pooled (Ortak varyanstan) – Her iki örneğin eşit varyanslı populasyonlardan örneklendiği kabul edilir. Ortak varyans, iki örnek varyansının tartılı ortalamasıdır.

$$SE_{\overline{Y}_1 - \overline{Y}_2} = \sqrt{\frac{S_p^2}{N_1} + \frac{S_p^2}{N_2}}$$

S_p^2 ortak varyans

Gruplar arası *t* test

Ortak varyansın bulunuşu

$$S_p^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

$$S_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

Gruplar arası *t*-test

$$t = \frac{\overline{Y}_1 - \overline{Y}_2}{SE_{\overline{Y}_1 - \overline{Y}_2}} \qquad SE_{\overline{Y}_1 - \overline{Y}_2} = \sqrt{\frac{S_p^2}{N_1} + \frac{S_p^2}{N_2}}$$

$$S_p^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

Gruplar arası *t* test
Standart hatanın hesaplanışı

2) Unpooled (Varyansın ortak alınmaması)- İki örneğin hata terimlerinin toplamıdır. Bu yaklaşım varyansların homojenliği varsayımının geçerli olmadığı yani örnekler önemli düzeyde farklı varyanslara sahip olduğunda durumda kullanılır.

$$SE_{\bar{Y}_1-\bar{Y}_2} = \sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}$$

Örnek: Hafıza denemesi
30 örnek şansa bağlı olarak 15'erli iki gruba ayrılmıştır

- **Deneme grubu**
  - Erkek denekler
  - Bütün uygulamalar diğer grupla aynı
  - A uygulaması (Bağımsız değişken)
  - Ölçümlar (bağımsız değişken)

- **Kontrol grubu**
  - Erkek denekler
  - Bütün uygulamalar diğer grupla aynı
  - B uygulaması (Bağımsız değişken)
  - Ölçümlar (bağımsız değişken)

Ölçülen hafıza değerleri bakımından İstatistik farklılıkları karşılaştırınız

Örnek: Hafıza denemesi

	Şok	Kontrol
Ortalama	6	8
S.D.	2	3
N	15	15

t değerini elle hesaplayalım

t değerinin önemliliğinin belirlenmesi

- Hangi büyüklükteki değerler önemlilik için yeterince büyüktür?
- *t*-dağılışının şeklini biliyoruz
- Büyük örnekler oluştur ve dağılışın şeklinin normale döndüğünü izle. Belirli bir t-değerine (veya ondan daha yüksek bir t-değerine) şansa bağlı olarak ulaşma şansı nedir?

Evaluating the Significance of *t*

- Şans olarak hangi düzeyi dikkate almaktayız?
- I-*tip* hata yapma olasılığı kaçtır?
- Bu değeri siz belirleyiniz (genel olarak 0.05'dir)
- Kritik değer hangi düzeydedir?
- Önemliliğe ait olasılık değerleri

t-testinin SPSS ile uygulanışı

- Bağımsız örneklerde t-testi için "*stacked (alt alta)*" verileri gir.
- *Bağımlı örneklerde ise "unstacked (sütun halinde)"* verileri gir.

- ***Eğer*** varyanslar *eşit değilse* (Levene testinde H_0 $p<.05$ ile red edilmişse) "equal variances not assumed" bölümündeki sonuçlara bak…
- ***Eğer*** varyanslar *eşit* (Levene testinde $p>.05$) ise "equal variances assumed" bölümüne bak…
- ***Eğer*** p-değeri, alfa değerinden küçük ise, H_0 hipotezini red et.

SPSS t test çıktısı: Yorum? (re H_0)

Group Statistics

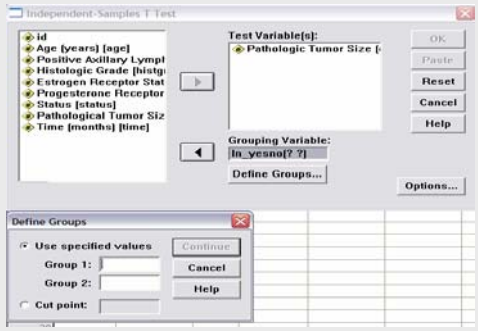
	GROUP	N	Mean	Std. Deviation	Std. Error Mean
SHOCK	Shock	10	4.1000	3.0350	.9597
	Control	10	5.2000	2.6583	.8406

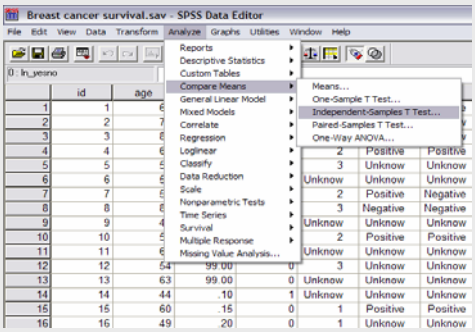
Eğer olasılık değeri (Sig.) <.05 ise H_0 hipotezini red edilir. Eğer burada olduğu gibi p >.05 ise, H_0 kabul edilir.

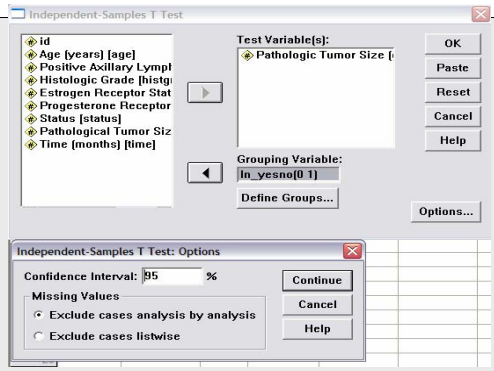
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% CI Interv Diffe Lower
SHOCK	Equal variances assumed	.266	.612	-.862	18	.400	-1.1000	1.2758	-3.7804
	Equal variances not assumed			-.862	17.693	.400	-1.1000	1.2758	-3.7838

Eğer Levene testinin olasılık değeri buradaki gibi >.05 ise, varyansların eşit olduğu anlaşılır ve üst satırdaki bilgiler kullanılır.







SPSS t test çıktısı: *testi yorumlayınız.*

Group Statistics

	Memshock	N	Mean	Std. Deviation	Std. Error Mean
SCORE	Shock	10	4.1000	3.2813	1.0376
	Control	10	6.5000	1.3540	.4282

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95 I Lo
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
SCORE	Equal variances assumed	18.294	.000	-2.138	18	.046	-2.4000	1.1225	-4.7
	Equal variances not assumed			-2.138	11.979	.054	-2.4000	1.1225	-4.8

SPSS t test çıktısı: *testi yorumlayınız.*

Group Statistics

	Groups	N	Mean	Std. Deviation	Std. Erro Mean
memory score	Shock	40	3.6000	1.6140	.255
	.00	40	5.1500	2.9574	.467

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Erro Differenc
memory score	Equal variances assumed	15.948	.000	-2.910	78	.005	-1.5500	.532
	Equal variances not assumed			-2.910	60.340	.005	-1.5500	.532

Ünite için *t*-test

Eşleştirilmiş örneklerde t-test (matched samples *t*-test veya paired *t*-test)
Correlated (ilişkili) örneklerdir

Ne zaman uygundur?
İki grubun karşılaştırılması:

İki değer alabilen "within-subject" ünite içi denemeler üniteler karşılıklı (matched) olduğunda
İkizlerde, kardeşlerde, aynı batında doğanlarda
Evli çiftler gibi eş "pairs" ünitelerde

Ünite için *t*-test

- Neyi hesaplıyoruz?
- Ortalama farklılığı
- Varyans
- Hesaplanışı:

$$t = \frac{\overline{D}}{SE_{\overline{D}}}$$

Belirlenen Ortalama farklılık
Şansla beklenen farklılık

Ünite için *t*-test

$$t = \frac{\overline{D}}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{(N-1)}}}$$

serbestlik derecesi = d.f.= N-1
N= ünite veya eş sayısı

Eşleştirilmiş örneklerde *t*-testi

•Eşler arasında sorumluluk bakımından yapılan değerlendirmede yanda verilen puanlar elde edilmiştir....

Eş	bay	bayan
1	6	8
2	4	3
3	3	8
4	2	4
5	3	6
6	1	0
7	1	4

Eşleştirilmiş örneklerde *t*-testi

Eş	bay	bayan	D	D^2
1	6	8		
2	4	3		
3	3	8		
4	2	4		
5	3	6		
6	1	0		
7	1	4		

Eşleştirilmiş örneklerde *t*-testi

Eş	bay	bayan	D	D^2
1	6	8	**2**	
2	4	3	**-1**	
3	3	8	**5**	
4	2	4	**2**	
5	3	6	**3**	
6	1	0	**-1**	
7	1	4	**3**	

Eşleştirilmiş örneklerde *t*-testi

Eş	Bay	Bayan	D	D^2
1	6	8	2	4
2	4	3	-1	1
3	3	8	5	25
4	2	4	2	4
5	3	6	3	9
6	1	0	-1	1
7	1	4	3	9

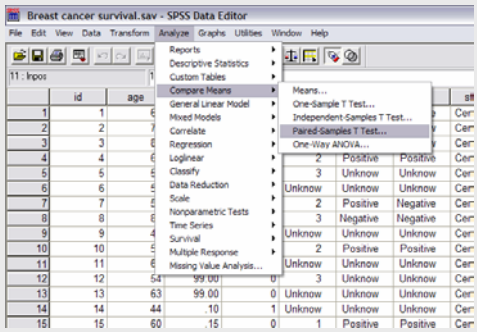
$\Sigma D=13$ $\Sigma D^2=53$

$\bar{D}=1.86$

$(\Sigma D)^2=169$

Ünite için *t*-test

$$t=\frac{\bar{D}}{\sqrt{\frac{\Sigma D^2-\frac{(\Sigma D)^2}{N}}{N-1}}}=\frac{1.86}{\sqrt{\frac{53-\frac{169}{7}}{6}}}=\frac{1.86}{2.19}=0.85$$



SPSS çıktısı: eşleştirilmiş örneklerde *t*- testi

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MALES	2.8571	7	1.7728	.6701
	FEMALES	4.7143	7	2.8702	1.0848

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	MALES & FEMALES	7	.646	.117

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	MALES - FEMALE	-1.8571	2.1931	.8289	-3.8854	.1711	-2.240	6	.066

Ünite içi Denemeler (Within-subjects designs)

Bu denemeler ne zaman gündemdedir?

- Ünite sayısı az olduğunda
- Ünite ile ilişkili faktörler nedeniyle hata varyansı yüksek olduğunda
- Deneme gruplarının etkisi düşük düzeyde olduğunda
- Zamana bağlı etkiler üzerinde çalışıldığında