



**T.C.  
ONDOKUZ MAYIS ÜNİVERSİTESİ**

**FEN EDEBİYAT FAKÜLTESİ  
İSTATİSTİK BÖLÜMÜ**

**İST.398 SOSYAL BİLİMLER İÇİN  
BİLGİSAYAR UYGULAMALI İSTATİSTİK**

**PROF. DR. YÜKSEL ÖNER**

**5. Hafta**

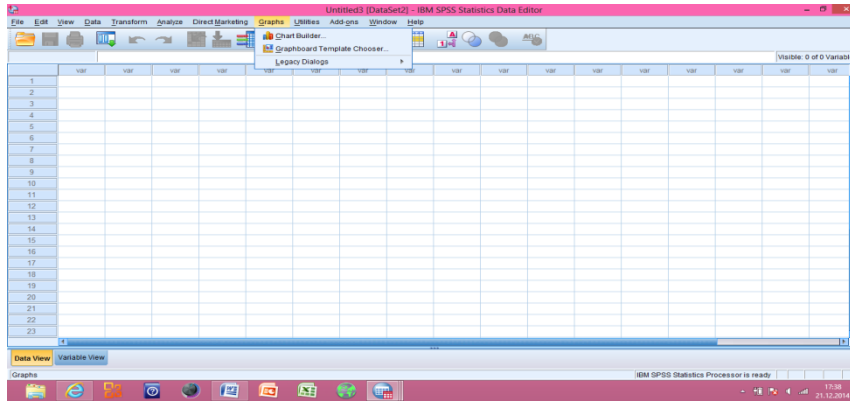
### III.4 GRAFİK ÇİZİMİ

İstatistiksel verilerin özetlenmesinde, değişkenlerin yada bu değişkenlere ait ölçüm değerlerinden oluşan veri kümelerinin dağılım biçimi, eğilimi hakkında ayrıntılı bilgi edinmede yardımcı araç olarak grafik çizimlerinden yararlanılır. Verilerle ilgili analitik çalışmalara geçmeden önce grafiklerin izlenmesi doğru analizi seçme bakımından faydalı bir yoldur.

Paket programlarda değişken tipi ve çizim amacına bağlı olarak değişik türde grafikler çizilebilir. SPSS’de çizilebilecek grafik türlerini; genel olarak, çizimlerinde yararlanılan tekniklere göre 6 gruba ayırmak mümkündür.

1. Çizgi Grafikleri
2. Çubuk Grafikleri
3. Alan Grafikleri
4. Resimli Grafikler
5. Birikimli (Eklemeli) Grafikler
6. İlişki Grafikleri

Bu genel grupta içinde alt gruplarda yer alan birçok grafik türü de mevcuttur. SPSS’de grafik çizimi için Ekran 3.34 da görünen **Graphs** menüsü kullanılır.



**Ekran 3.34** SPSS Grafik menüsü

Bu menüde üç ana başlık yer almaktadır. Bunlar:

1. Chart Builder (Grafik Oluşturucu)
2. Graphboard Template Chooser (Grafik Şablon Seçici)
3. Legacy Dialogs (Eski/Yerleşik Diyaloglar)

Bu ana başlıklardan özellikle üçüncüsü kullanılarak amaca uygun grafik türü belirlenerek çizilebilir. İstatistik verileri özetlemede yaygın olarak kullanılan grafik türleri ise:

1. Frekans Histogramı Grafiği
2. Frekans Eğrisi (Poligonu) Grafiği
3. Eklemeli Frekans Eğrisi Grafiği
4. Daire Grafiği
5. Kutu (Boxplot) Grafiği
6. İlişki (Scatter Plot=Yayılım) Grafiği

olarak verilebilir.

### III.4.1 Frekans Histogramı Grafiği

Frekans histogramı nicel türden, sürekli/kesikli ve en az eşit aralıklı ölçme düzeyine sahip değişkenlere ait verileri özetlemede kullanılan bir çubuk grafik türüdür. Öncelikle orijinal veriler, kodlama işlemi uygulanarak sınıflandırılmış veya gruplandırılmış frekans verisine dönüştürülür. Koordinat düzleminde yatay ekseninde sınıf değerlerinin ( $x_j, j = 1, 2, \dots, m$ )/sınır sınırlarının ve dikey ekseninde frekansların ( $f_j, j = 1, 2, \dots, m$ ) ya da yüzde değerlerinin (oransal frekansların,  $p_j, j = 1, 2, \dots, m$ ) yer aldığı yan yana dikdörtgenlerden oluşan bir grafik türüdür. Her bir sınıf/kategori bir dikdörtgen ile temsil edilir. Bir dikdörtgenin boyu/yüksekliği temsil ettiği sınıfın frekansı/yüzde değeri ile eşlenir. Grafikte sınıflandırılmış/gruplandırılmış verinin sınıf sayısı kadar dikdörtgen yer alır. Dikdörtgenlerin boy uzunluklarının toplamı; dikey ekseninde frekanslar gösterilmiş ise sınıf frekanslarının toplamına ( $\sum_{j=1}^m f_j = n$ ) eşit olacağından toplam gözlem sayısını ( $n$ ), yüzde değerleri gösterilmiş ise yüzde değerlerinin toplamını ( $\sum_{j=1}^m p_j = 1$ ) verir.

SPSS’de frekans histogramı çizim algoritması:

**Adım:1** Değişken/ değişkenlerin tanımlamaları ve özellikleri (Variable View) girilir.

**Adım:2** Değişken/değişkenlere ait veriler veri sayfasına (Data View) girilir.

**Adım:3** Eğer gerekiyorsa (nitel türden verilerde gerekli olmasa bile, nicel türden verilerde gerekebilir), orijinal veriler kodlama işlemi ile sınıflandırılmış ya da gruplandırılmış frekans verisine dönüştürülür.

**Adım:4** **Graphs** menüsünden Legacy Dialogs seçeneği seçilir. Açılan seçim ekranından **Histogram** seçilir.

**Adım:5** Ekranın sağ tarafında yer alan değişkenler listesinden ilgili değişken seçilerek (kategorik bir değişken olmalı) **Variable** işlem kutusuna aktarılır. **OK** tuşu tıklanır. Grafik çıktısı çıktı sayfasında sunulur.

**Örnek 3.16** 100 bireyin boy uzunlukları (cm) aşağıdaki gibidir. Bu verilerin bir grafik özetlemesini, verileri gruplandırılmış frekans verisine dönüştürerek frekans histogramı ile gösteriniz?

179	166	191	181	174	174	172	191	188	163
167	163	194	169	176	190	173	173	158	180
171	165	180	185	170	179	170	174	179	158
179	163	164	172	167	157	171	164	177	172
158	160	178	170	168	168	161	156	159	163
156	161	157	176	190	169	175	176	172	176
161	171	170	170	160	164	158	154	193	185
184	159	158	182	157	159	175	173	187	188
161	173	163	184	154	154	165	178	170	187
173	164	170	182	179	154	160	194	173	147

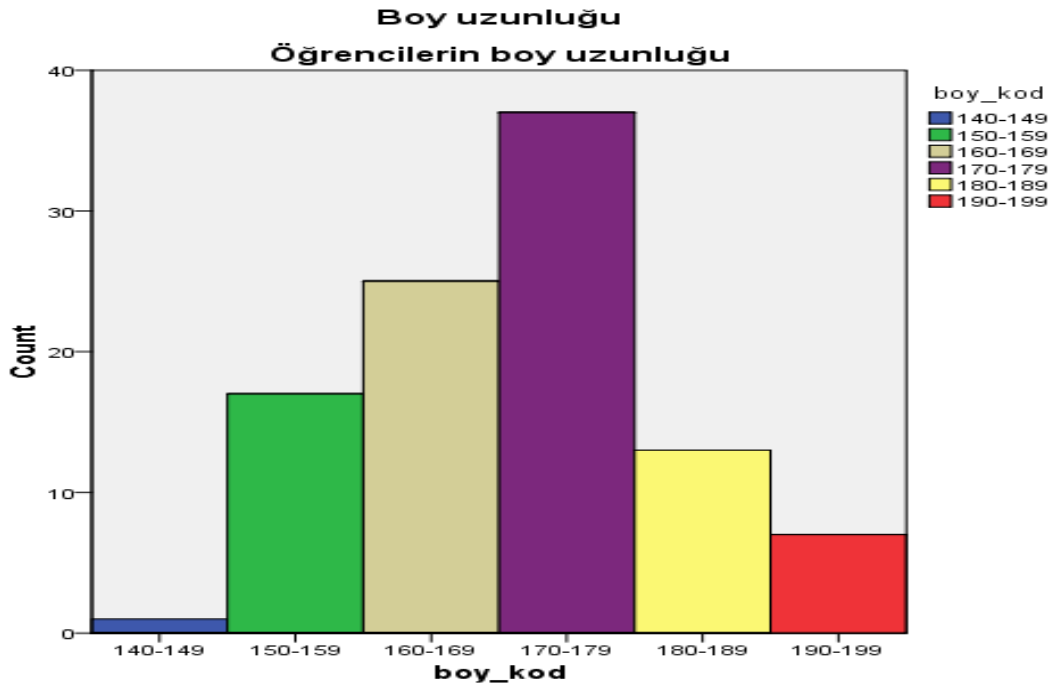
**Çözüm:** Bu verilerin gruplandırılmış frekans verisi dönüşümünü aşağıdaki tabloyu dikkate alarak yapalım.

Sınıf Sınırları	Kod değeri	Sınıf Sınırları	Kod değeri	Sınıf Sınırları	Kod değeri
140-149	1	160-169	3	180-189	5
150-159	2	170-179	4	190-199	6

Verilere ait frekans tablosu Tablo 3.45 ve frekans histogramı ise Grafik 3.1 ile verildi.

**Tablo 3.43** Öğrencilerin boy uzunluğu

Boy uzunluğu	Frequency	Percent	Cumulative Percent
140-149	1	1,0	1,0
150-159	17	17,0	18,0
160-169	25	25,0	43,0
170-179	37	37,0	80,0
180-189	13	13,0	93,0
190-199	7	7,0	100,0
Total	100	100,0	



**Grafik 3.1** Öğrencilerin boy uzunluğu frekans histogramı

#### III.4.2 Frekans Eğrisi, Eklemeli Frekans Eğrisi Grafiği

Bu grafikler nicel türden, sürekli/kesikli ve en az eşit aralıklı ölçme düzeyine sahip değişkenlere ait verileri özetlemede kullanılan bir çizgi grafik türüdür. Öncelikle orijinal veriler, kodlama işlemi uygulanarak sınıflandırılmış veya gruplandırılmış frekans verisine dönüştürülür. Koordinat düzleminde yatay ekseninde sınıf değerleri/orta noktaları ( $x_j$  veya  $\bar{x}_j$   $j = 1, 2, \dots, m$ )

veya sınıf sınırları ve düşey ekseninde frekans eğrisi için frekanslar ( $f_j, j = 1, 2, \dots, m$ ) ya da yüzde değerleri (oransal frekanslar  $p_j, j = 1, 2, \dots, m$ ), eklemeli frekans eğrisi için eklemeli frekanslar(-den daha az) ya da eklemeli yüzde değerleri (-den daha az) yer alır. Frekans eğrisi için her bir sınıfı temsil eden değer ile karşılık gelen frekans/yüzde değerinin oluşturduğu ( $x_j, f_j$ ) veya ( $x_j, p_j$ ),  $j = 1, 2, \dots, m$  nokta ikilileri, eklemeli frekans eğrisi için her bir sınıfın sınıf sınırları ile karşılık gelen eklemeli frekans/yüzde değerinin oluşturduğu nokta ikilileri düzlemde belirlenir. Belirlenen noktalar doğru parçaları ile birleştirilerek grafik tamamlanır. SPSS’de frekans eğrisi ve eklemeli frekans eğrisi çizim algoritması:

**Adım:1** Değişken/ değişkenlerin tanımlamaları ve özellikleri (Variable View) girilir.

**Adım:2** Değişken/değişkenlere ait veriler veri sayfasına (Data View) girilir.

**Adım:3** Orijinal veriler kodlama işlemi ile sınıflandırılmış ya da gruplandırılmış frekans verisine dönüştürülür.

**Adım:4** **Graphs** menüsünden **Legacy Dialogs** seçeneği seçilir. Bu seçim sonrasında Ekran 3.35 görüntülenir. Bu ekrandaki listeden **Line** seçeneği seçilir.

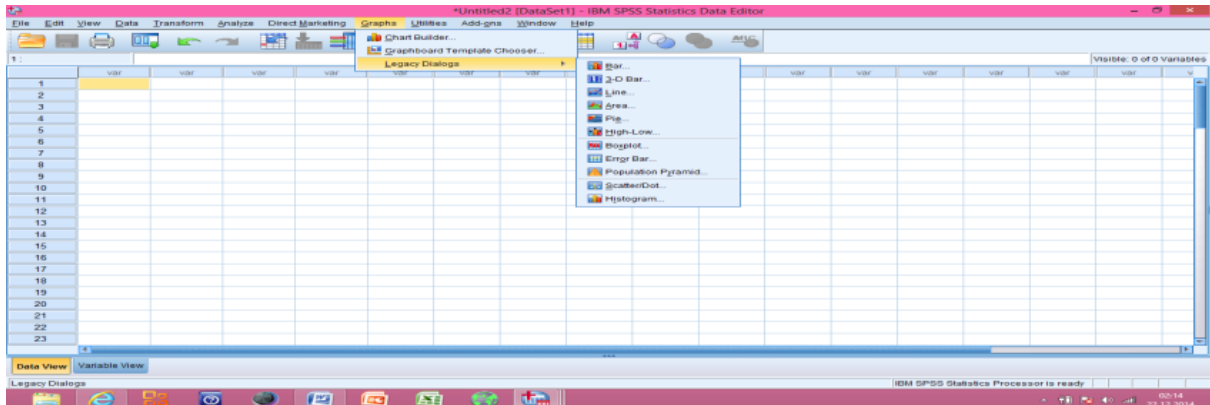
**Adım:5** Açılan yeni ekran penceresinden **Simple** grafik türü seçildikten sonra **Define** tuşuna basılır ve Ekran 3.36 görüntülenir.

**Adım:6** Değişkenler listesinden frekans eğrisi çizilecek değişkenin kategorik (kodlanmış) değişkeni **Category axis** işlem kutusuna aktarılır. **Line Represents** bölümünden ilgili seçenek işaretlenir. Bu bölümdeki seçenekler:

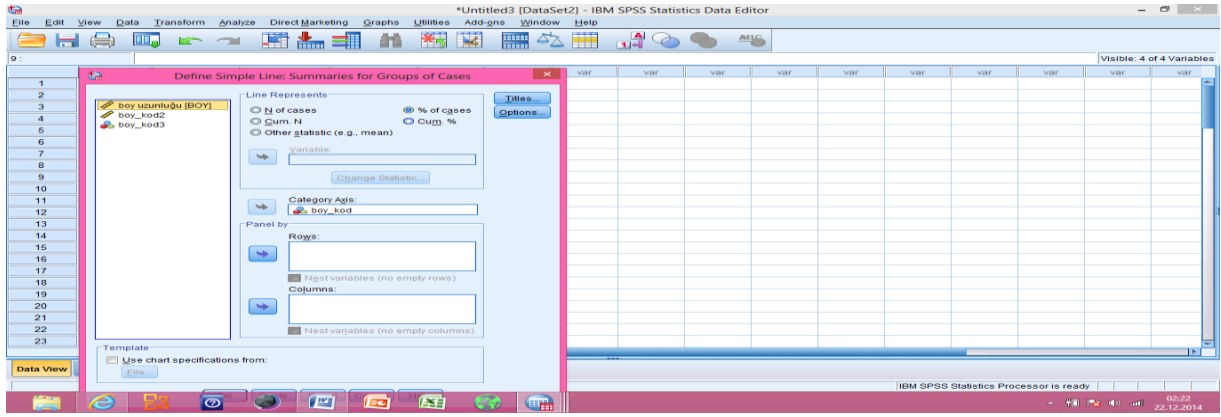
- **N of cases:** düşey ekseninde frekansları
- **% of cases:** düşey ekseninde yüzde değerlerini
- **Cum.N:** düşey ekseninde eklemeli frekansları
- **Cum.%:** düşey ekseninde eklemeli yüzde değerlerini

gösterir.

**Adım:7** Gerekirse grafiğe başlık adı vermek için **Titles** seçeneği ile açılan ekranda başlık (**Title**) bölümüne, istenirse alt başlık (**subtitle**) bölümüne yazılır. **Continue/OK** tıklanarak işlem bitirilir.



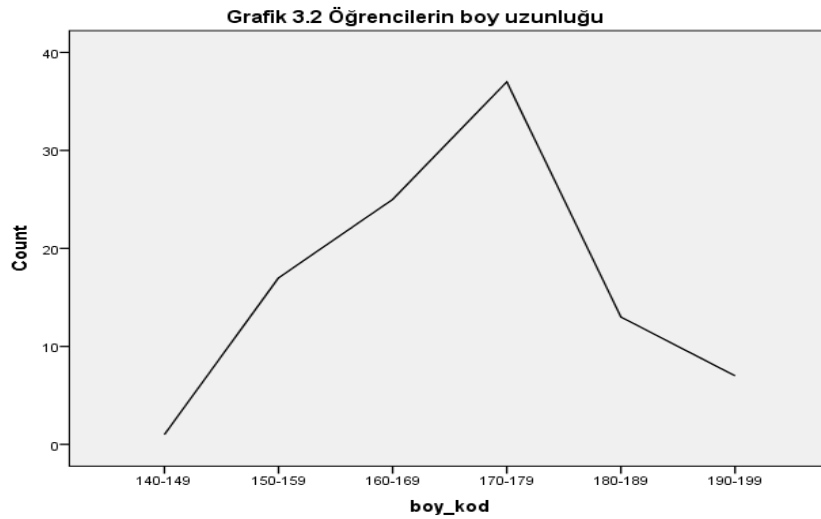
**Ekran 3.35** Legacy Dialogs işlem ekranı



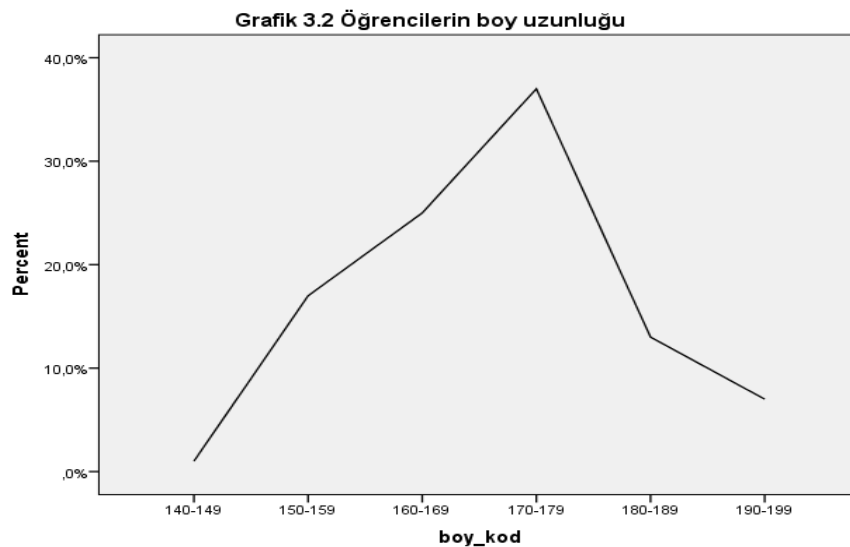
**Ekran 3.36** Frekans eğrisi çizimi işlem ekranı

**Örnek 3.17** Örnek 3.16'daki veri için a) Frekans eğrisi b) Oransal frekans eğrisi c) Eklemeli frekans eğrisi d) Eklemeli oransal frekans eğrisi grafiklerini çiziniz.

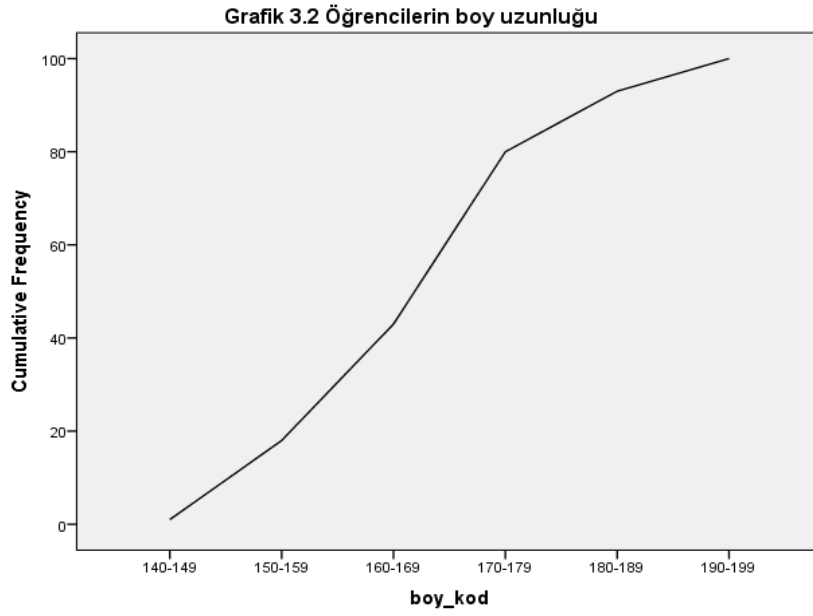
**Çözüm: a)**



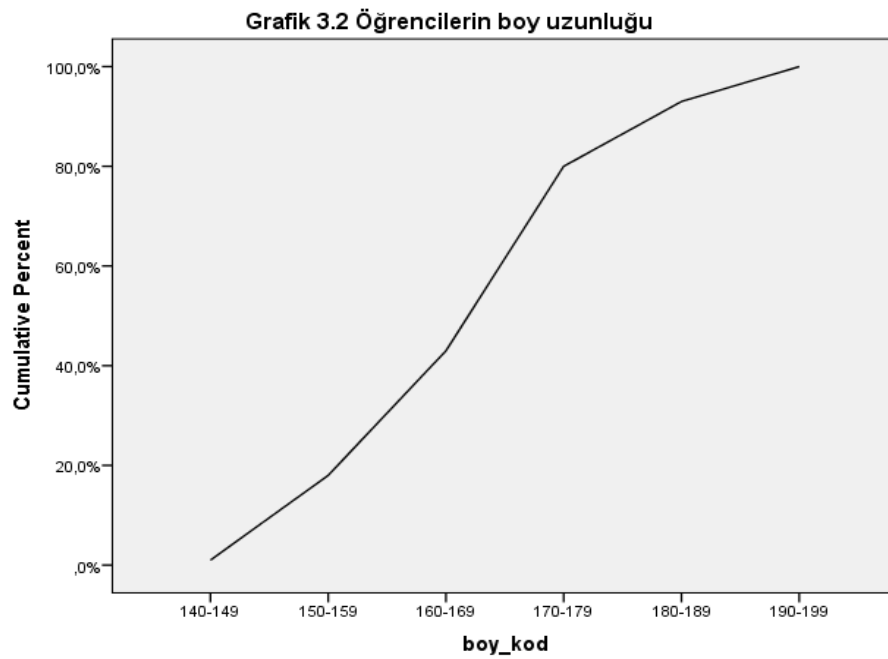
b)



c)



d)



### III.4.3 Daire Grafiği

Nitel türden kategorik verileri veya nicel türden kategorik hale getirilmiş gruplandırılmış frekans verilerini özetlemede kullanılan bir alan grafik türüdür. Daire grafiği  $m$  tane sınıf ya da kategoriye ait frekansların dağılımını daire içinde açısız yoğunluk biçiminde gösteren bir grafik türüdür. Her bir sınıf/kategori daire içinde temsil edilme derecelerine göre dilimlere ayrılarak

gösterilir. Daire grafiğinde her bir dilimin yanına sayı ya da yüzde açıklamalarından herhangi birisi yazılabilir.  $j = 1, 2, \dots, m$  olmak üzere,  $j$ .nci sınıf/kategorinin dilim payı  $= 360^\circ \times p_j$  ile hesaplanır. SPSS’de daire grafiği çizim algoritması:

**Adım:1** Değişken/ değişkenlerin tanımlamaları ve özellikleri (Variable View) girilir.

**Adım:2** Değişken/değişkenlere ait veriler veri sayfasına (Data View) girilir.

**Adım:3** Orijinal veriler kategorik (sınıflama ya da sıralama ölçekli) değilse, önce kodlama işlemi ile sınıflandırılmış ya da gruplandırılmış frekans verisine dönüştürülür.

**Adım:4** **Graphs** menüsünden **Legacy Dialogs** seçeneği seçilir. Bu seçim sonrasında Ekran 3.35 görüntülenir. Bu ekrandaki listeden **Pie...** seçeneği seçilir.

**Adım:5** Açılan yeni ekran penceresinden **birinci seçenek** seçildikten sonra **Define** tuşuna basılır ve Ekran 3.37 görüntülenir.

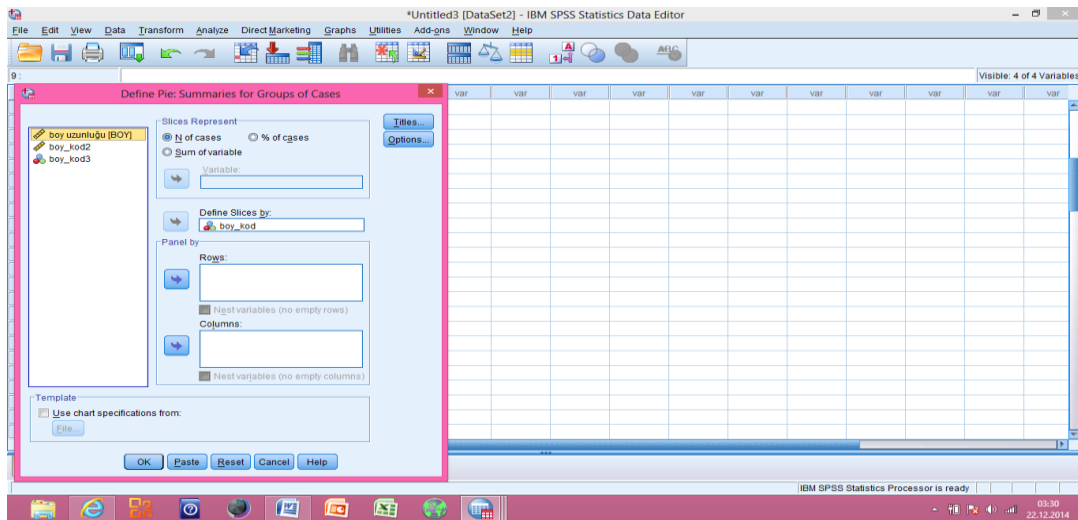
**Adım:6** Değişkenler listesinden daire grafiği çizilecek değişkenin kategorik değişkeni **Define slices by** işlem kutusuna aktarılır. **Slices Represents** bölümünden ilgili seçenek işaretlenir. Bu bölümdeki seçenekler:

- **N of cases:** daire dilimlerine düşen frekansları
- **% of cases:** daire dilimlerine düşen yüzde değerlerini

gösterir.

**Adım:7** Gerekirse grafiğe başlık adı vermek için **Titles** seçeneği ile açılan ekranda başlık (**Title**) bölümüne, istenirse alt başlık (**subtitle**) bölümüne yazılır. **Continue/OK** tıklanarak işlem bitirilir.

**Adım:8** Daire grafiği çıktı sayfasında görüntülenir. Görüntülenen grafik üzerine çift tıklayarak kullanıcı tanımlı daire grafikleri elde edilebilir. Yüzdelerin açılabilir alanlarda (daire dilimlerinde) yer almasını sağlamak için **Chart Editor** ekranında **Properties** seçeneği tıklanır ve **Data Values Labels** seçeneği ile gerekli tanımlamalar yapılır. **Fill&Border** seçenekleri aracılığı ile gerekli tanımlamalar yapılır. **Depth&Angle** seçeneği tıklanır ve 3-D seçeneği seçilir, **Apply** tıklanarak grafik tamamlanır.

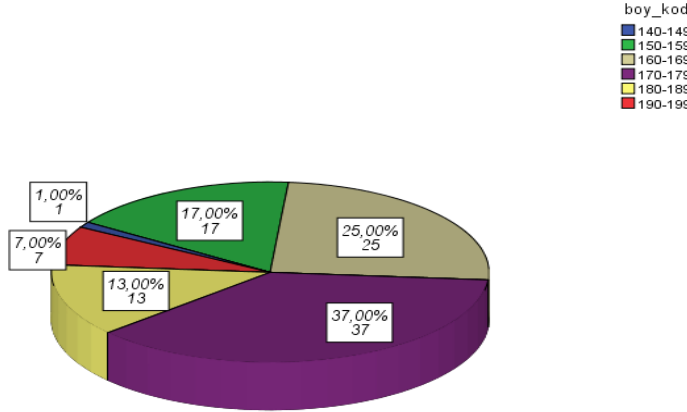


**Ekran 3.37** Daire grafiği çizimi işlem ekranı



**Örnek 3.18** Örnek 3.16'daki veri için ilgili değişkenin kategorik değişkenini (gruplandırılmış frekans verisini) kullanarak daire grafiğini çiziniz.

**Çözüm**



**Grafik 3.6** Boy uzunluğu daire grafiği

### III.4.4 Kutu (Boxplot) Grafiği

İstatistikte kutu grafiği, bir betimsel istatistik ve istatistiksel grafik türü olup, nicel türden verileri görsel şekilde özetlemek için kullanılan bir açıklayıcı veri analizi aracıdır. Kutu grafiği, ilgili değişken bakımından veri için hazırlanan **beş sayılı özetleme tablosu** [ $X_{enk}$ (en küçük gözlem değeri),  $Q_1$ (birinci çeyrek),  $X_{med} = Q_2$ (medyan veya ikinci çeyrek),  $Q_3$ (üçüncü çeyrek),  $X_{enb}$  (en büyük gözlem değeri)] gösterimini grafiksel olarak özetlemeye dayalıdır. Özellikle bu grafik türü ile verileri merkezsiz konum, yayılma, çarpıklık ve basıklık yönünden özetlemek ve aykırı değerleri tanımlamak esas amaçtır. Bu sebeple kutu grafiği, verilerin ortanca(medyan) etrafında yayılışlarını ve verilerin %75'inin hangi değerler arasında yer aldığını, aykırı (sapan) değerler (outliers) varsa bunların konumlarını (aykırı değer olup olmadığı gibi) belirlemeye yarayan bir grafik türüdür. Analitik çalışmalara başlamadan önce verileri tanıma bakımından çizilmesi uygun bir grafik türüdür.

Kutu grafiği iskeleti, bir veri grubunun en küçük ile en büyük değerleri arasında üzeri ölçekli bir çizgiden oluşur; yani  $X_{enk}$  ile  $X_{enb}$  noktaları arasında ya yatay ya da dikey bir çizgi halindedir. Bu çizgi üzerinde ölçeğe göre bir kutu çizilir. Kutunun en küçük değeri birinci çeyrek  $Q_1$  ve en büyük değeri üçüncü çeyrek  $Q_3$  olur. Kutunun uç noktalarına Tukey tarafından "menteşe" adı verilmiştir; yani  $Q_1$  ve  $Q_3$  kutunun uç noktaları "menteşe" noktalarıdır. Bu

kutunun uzunluğu  $QA = Q_3 - Q_1$  ile hesaplanır ve bu  $QA$  değeri veri grubunun **çeyrekler açıklığı** ölçüsünü gösterir. . **Medyan** "kutu"nın içinde ölçeğe göre yerine konulup işaretlenir.

Kutu grafiğinin yorumlanması: Kutu grafiği görsel olarak veri grubunun dağılımı ile ilgili merkezsiz konum, yayılma, şekil itibarıyla çarpıklık ve basıklık karakteri hakkında ön bilgi verebilir. Veri grubuna ait dağılımın karakteri ile ilgili bu ölçülerle hakkında kesin bilgi ancak bu ölçülerin hesaplanması yolu ile elde edilebilir.

**Merkezsiz konum:** Veri grubunun merkezsiz konumu kutunun içinde işaretlemiş olan **medyan** noktası ile gösterilir.

**Yayılma:** Veri grubunun yayılmasının **açıklık** ( $R = X_{enb} - X_{enk}$ ) ölçüsüyle ölçülen niteliği, "kutu grafiği"nin (yatay veya dikey) tüm çizgisinde görülür. **Çeyrekler açıklığı** ( $QA = Q_3 - Q_1$ ) ise "kutu grafiği"nde kutunun büyüklüğü ile görülür.

**Çarpıklık:** Verilerinin dağılımının "çarpıklık" ölçüsü kutu grafiği ile bulunamaz. Ancak dağılımın çarpıklık karakteri kutunun tüm çizginin  $X_{enk}$  tarafına mı yoksa  $X_{enb}$  tarafına mı yakın olması ile belirlenir. Eğer kutu, çizginin  $X_{enk}$  tarafında ise, kutu içindeki medyan noktası  $X_{enk}$  değerine yakın olur ve kutu dışında kalan çizginin uzun kısmı (yani  $X_{enb}$ 'e doğru olan kısmı) yukarıya (sağa) doğru uzunca bir kuyruk gösterir. Bu durum veri grubunun dağılımının sağa *çarpıklık* gösterdiği anlamına gelir. Eğer kutu, çizginin  $X_{enb}$  tarafında ise, yani medyan noktası  $X_{enb}$  değerine yakın ise, kutu dışında kalan çizginin uzun kısmı (yani  $X_{enk}$ 'e doğru olan kısmı) aşağıya (sola) doğru uzunca bir kuyruk gösterir. Bu durum veri grubunun dağılımının sola *çarpıklık* gösterdiği anlamına gelir. Ancak "çarpıklık" karakterini doğrudan doğruya ölçen bir sayısal ölçü "kutu grafiği"nden elde edilemez. Bu ölçü ile ilgili kesin bilgi hesaplama yolu ile elde edilir.

**Basıklık:** Veri grubunun dağılımına ait *basıklık* ölçüsü de kutu grafiğinden kesin bir bilgi olarak ölçülemez. Ancak dağılımın basıklık niteliği hakkında bu grafikten bir kanaat edinilebilir. Dikkat edilirse bu bir *kanaat edinme* olup, epeyce zayıf, bir intiba edinme şeklindedir ve tam olarak basıklık niteliği öğrenilmez. Basıklık için bir kanaat sağlama kutunun genişliği ile çizginin genişliğinin karşılaştırılması ile edinilebilir. Eğer kutu çizginin çoğunu (yani  $X_{enk}$  ile  $X_{enb}$  arasının çoğunu) kapsıyorsa genellikle dağılımın basık olduğu, kutu çizginin çok az bir kısmını kapsıyorsa dağılımın sivri olduğu hakkında bir ilk intiba ortaya çıkabilir. Bu ölçü ile ilgili kesin bilgi yine hesaplama yolu ile elde edilir.

SPSS'de kutu grafiği çizim algoritması:

**Adım:1** Değişken/ değişkenlerin tanımlamaları ve özellikleri (Variable View) girilir.

**Adım:2** Değişken/değişkenlere ait veriler veri sayfasına (Data View) girilir.

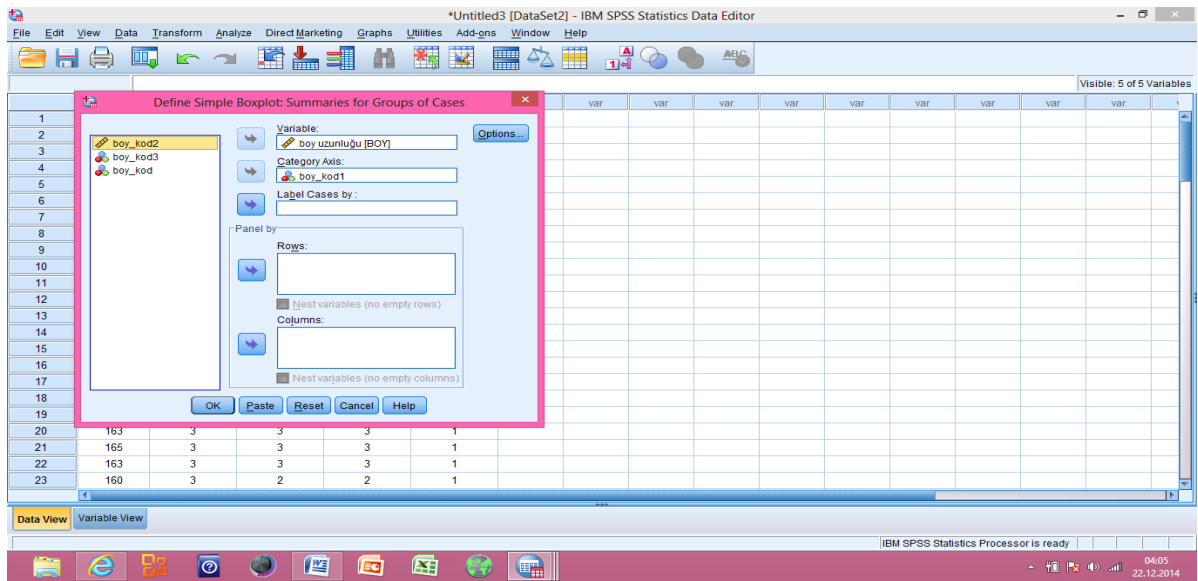
**Adım:3** Orijinal verilerin tek bir grubun verisi olduğunu belirtmek için bütün değerler "1" ile kodlanarak kategorik değişkene dönüştürülür. Bunun için **Transform > Recode into different variables** seçeneklerinden sonra, değişkenler listesinden ilgilenilen değişken seçilerek **Numeric Variable- Output variable** işlem kutusuna aktarılır ve **Output variable** bölümünde

yeni deęiřken tanımlanır. Böylece **Numeric Variable- Output variable** eřleřtirmesi yapılır. Sonra **Old and new values** iřlem penceresi aılır. Bu penceredeki seeneklerden **All other values** seeneęi seilir ve **New value**'de **value**'ye "1" yazılarak **Add** tuřu ile **Old—New** kutusuna aktarılır. **Continue** > **OK** ile iřlem bitirilir. Bu iřlem sonucunda, yeni deęiřken ve aldıęı deęerler hep "1" olarak **Data View** sayfasında yeni bir stunda grntlenir.

**Adım:4** **Graphs** mensnden **Legacy Dialogs** seeneęi seilir. Bu seim sonrasında Ekran 3.35 grntlenir. Bu ekrandaki listeden **Boxplot...**seeneęi seilir.

**Adım:5** Aılan yeni ekran penceresinden **simple** seeneęi seildikten sonra **Define** tuřuna basılır ve Ekran 3.38 grntlenir.

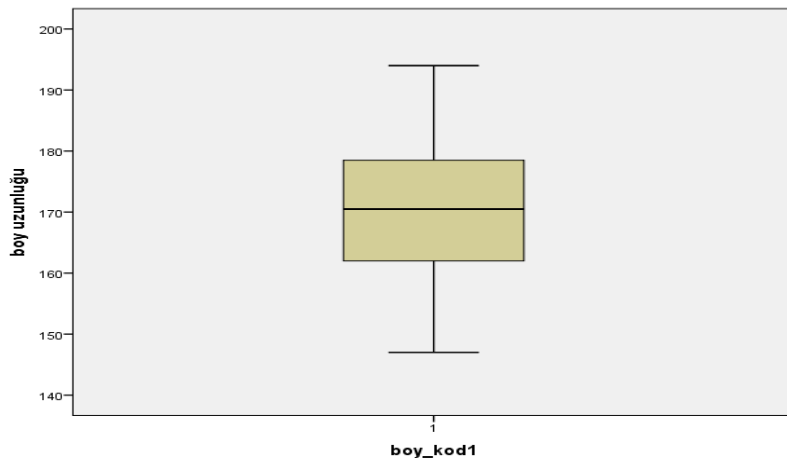
**Adım:6** Deęiřkenler listesinden kutu grafięi izilecek orijinal deęiřken **Variable** iřlem kutusuna, bu deęiřkenin kategorik deęiřkeni ise **Category Axis** iřlem kutusuna aktarılır. **OK** tıklanarak iřlem bitirilir.



**Ekran 3.38** Kutu grafięi izimi iřlem ekranı

**rnek 3.19** rnek 3.16'deki veri iin boy deęiřkeninin kutu grafięini iziniz ve deęerlendiriniz.

### **zm**



**Grafik 3.7** Bireylerin boy uzunluęu daęılımı kutu grafięi

Bu grafiğe göre bireylerin boy uzunlukları dağılımında en küçük gözlem değeri 140 ile 150 arasında iken, en büyük gözlem değerinin 190 ile 200 arasında olduğu söylenebilir. Bu durumda veriler yaklaşık olarak  $R = 200 - 140 = 60$  birimlik bir aralıkta dağılım göstermektedir. Ayrıca birinci çeyrek yaklaşık olarak  $Q_1 = 162$  ve üçüncü çeyrek yine yaklaşık olarak  $Q_3 = 179$  alınırsa çeyrekler arası açıklık (kutu genişliği)  $Q.A = 179 - 162 = 17$  birimdir. Boy uzunluğu verisi için merkezsel konum ölçüsü olan medyan değeri kutunun içinde çizili olan doğru olup hemen kutunun ortasında yer olmaktadır. Bu durum boy uzunluğu verisinin dağılımının simetrik dağılıma çok yakın olduğunu gösterir. Bu veri için medyan değeri yaklaşık olarak  $m = 170$  cm alınabilir.

Boy uzunluğu verisine ait dağılımın çarpıklık katsayısı her ne kadar kutu grafiği ile kestirilemez ise de, çarpıklık karakteri hakkında değerlendirme yapabiliriz. Grafikten kutunun hemen hemen veri grubuna ait en küçük gözlem değeri ile en büyük gözlem değerinin ortasında yer aldığı, bu sebeple de kutunun içerisindeki medyan çizgisinin kutunun ortasında yer aldığı görülmektedir. Bu durum boy uzunluğu verisine ait dağılımın simetrik olduğunun bir göstergesidir. Kutunun ve böylece medyan çizgisinin en küçük ya da en büyük gözleme yakın olması durumuna göre sağa ya da sola çarpıklık söz konusu olacaktır.

Boy uzunluğu verisine ait dağılımın basıklık katsayısı her ne kadar kutu grafiği ile kestirilemez ise de, basıklık karakteri hakkında değerlendirme yapabiliriz. Grafikte kutu çizgi üzerinde normal büyüklükte görülmektedir. Ne çok ince bir çizgi halinde ne de çizginin tamamına yakını kapsayacak büyüklükte olmadığından, boy uzunluğu verisinin dağılımının normal olduğu söylenebilir. Eğer çok ince bir çizgi halinde olsa idi dağılımın sivri olduğu, tersine çizginin tamamına yakını kapsıyor olsa idi dağılımın basık olduğu söylenebilir.