

ENM 5210

**İSTATİSTİK VE YAZILIMLA
UYGULAMALARI**

Ders 1

Minitab' da Grafiksel Analiz-I

İstatistik Nedir?



- İstatistik kelimesi ilk olarak Almanyada devlet anlamına gelen **status** kelimesine dayanılarak kullanılmaya başlanmıştır.

Tanım

Verilerin toplanması, düzenlenmesi ve uygun yöntemlerle analiz edilmesine istatistik denir.

İstatistiğin Konusu Olan Olaylar



- Olayları ikiye ayırmak mümkündür:

Tipik Olaylar az sayıda faktör tarafından etkilenen olaylardır

Olay

Toplu Olay ise çok sayıdaki faktör tarafından belirlenir. Dolayısı ile birkaç deneme ile olayların genel karakterleri belirlenemez. Tüm faktörler belirlenebilse bile Hepsi kontrol edilemez

Tipik Olay

Toplu Olay

ÖRNEK: A ve B kimsayalarının Belirli şartlar altında birleştiğinde ne olacağı kesin olarak söylenebilir

ÖRNEK: Bir futbol topuna vurulduğunda topun alacağı mesafe

İstatistiğin Konusu Olan Olaylar



- İstatistik tipik olaylarla ilgilenmez
- İstatistiğin ilgilendiği olaylar toplu ve sonuçları sayısal olarak ifade edilebilecek olaylar ile ilgilenir

Tanım 2:

İstatistik sayısal değerlerin toplanarak analiz edildiği bilimsel bir yöntem veya yöntemler toplamıdır

Verilerin Toplanması ve Analiz Edilmesi



■ Tanım:

İstatistiksel araştırmalarda incelenmesi amaçlanan birimlerin oluşturduğu topluluğa **anakütle** denir.

Birim ise anakütleyi oluşturan en küçük parçadır.

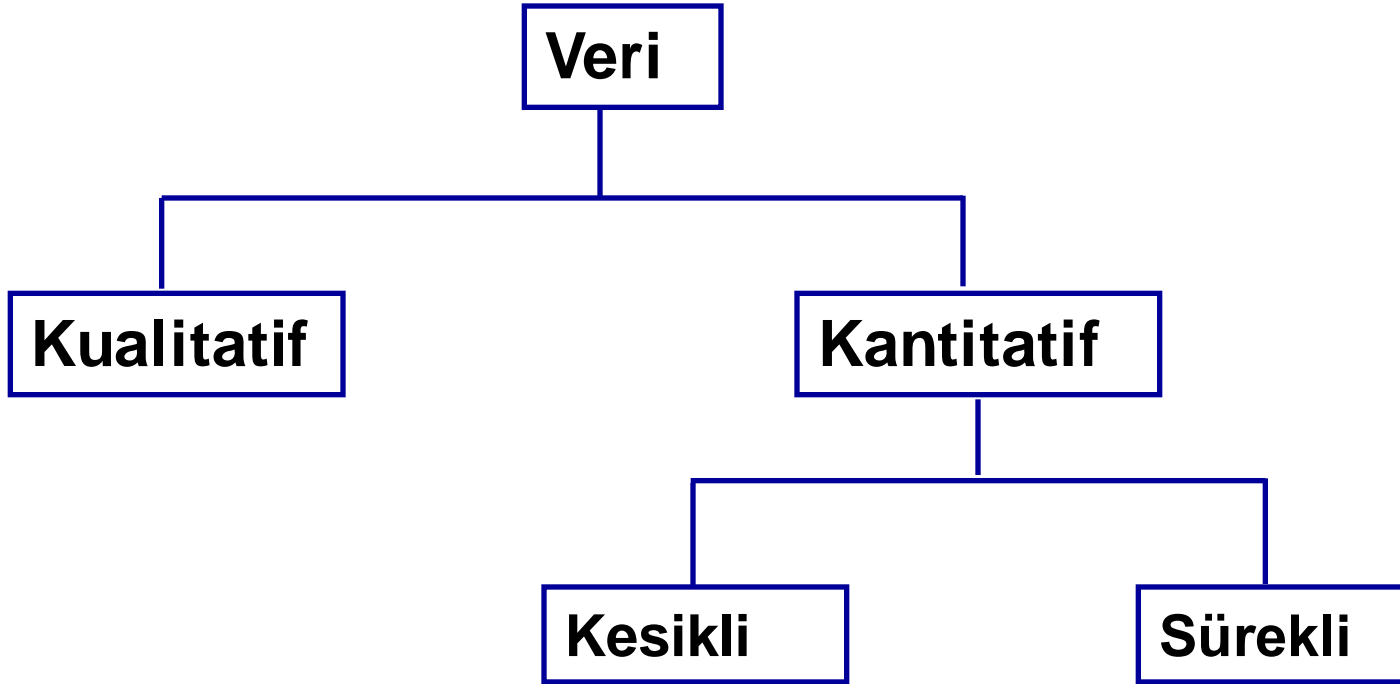
■ **ÖRNEK:** Bir bölgede yaşayan ailelerin gelir dağılımları ile ilgili bir araştırmayı ele alalım:

- Anakütle:
- Birim:

Değişken Türleri

- **Kantitatif (Quantitative) Değişken:**
 - Sayısal ölçekte ifade edilir.
 - Miktar hakkında bilgi verir.
 - Hesabınızdaki bakiye, pilin ömrü, sınıftaki öğrenci sayısı vb.
- **Kualitatif (Qualitative) Değişken:**
 - Sayısal olmayan değişkenlerdir.
 - Doğum yeri, göz rengi, ırk, vb.

Değişkenlerin Sınıflandırılması



Bir ailedeki çocuk sayısı
Bir yılda satılan TV adedi

Bir kargonun ağırlığı
Boy uzunluğu

Tam Sayım ve Örnekleme



- Anakütle ile ilgili bilgi toplamak istendiğinde tüm birimler teker teker inceleniyor ise bu işleme **tam sayım** denir.
- Tam sayımda anakütlenin özellikleri kesin olarak öğrenilebilir
- Dezavantajları ise
 - ✓ Zaman Alıcıdır
 - ✓ Masraflıdır

Tam Sayım ve Örnekleme



- Anakütleden anakütle birim sayısından *daha az sayıda* birimin seçilerek; bu birimler yardımıyla anakütle parametrelerinin tahmin edilmesi işlemine *örneklemeye* denir
- Anakütleden seçilen büyüklüğü anakütleden daha küçük olan bu topluluğa ise örnek denir

Tam Sayım ve Örnekleme



- **Örneklemenin Üstünlükleri**
 - Kısa sürede ve daha az masrafla yapılabilir
 - Veri toplama ve analizde yapılabilecek hatalar daha azdır
- **Örneklemenin Sakıncaları**
 - Parametre tahminlerinde sapmalar ortaya çıkar

Grafik Türleri



Amaçlarına göre grafikleri 3 gruba sınıflandırılabilir:

- Değişkenlik analizi
- Frekans karşılaştırması için verileri sınıflandırma
- Değişkenler arasındaki ilişkileri gösterme

Grafik Türleri (devam)



Ver setinin değişkenliğinin analizi için:

- Histogramlar
- Nokta Grafikleri
- Kutu Diyagramları

Grafik Türleri (devam)



Verileri sınıflarına ayırmak ve frekanslarını göstermek için:

- Çubuk Grafikleri
- Pasta Grafikleri
- Pareto Diyagramları

Grafik Türleri (devam)



Değişkenler arasındaki ilişkileri analiz için:

- Saçılım Diyagramları
- Zaman Serisi Grafikleri
- 3 Boyutlu Grafikler

Veri Manüplasyonları



a) Her bir bölgede ayrı bir satış takımı görev yapmaktadır. İlk 4 aydaki satışların %10' nun takım üyelerine verilmesine karar verilmiştir.

Her bir takımın alacağı miktarı hesaplatıp, C7 sütununa yazdırın.

b) Calculator ve Row Statistics' leri kullanarak bölgelere göre ortalama aylık satış miktarını hesaplatın ve C8 sütununa yazdırın.

Tanımlayıcı İstatistikler

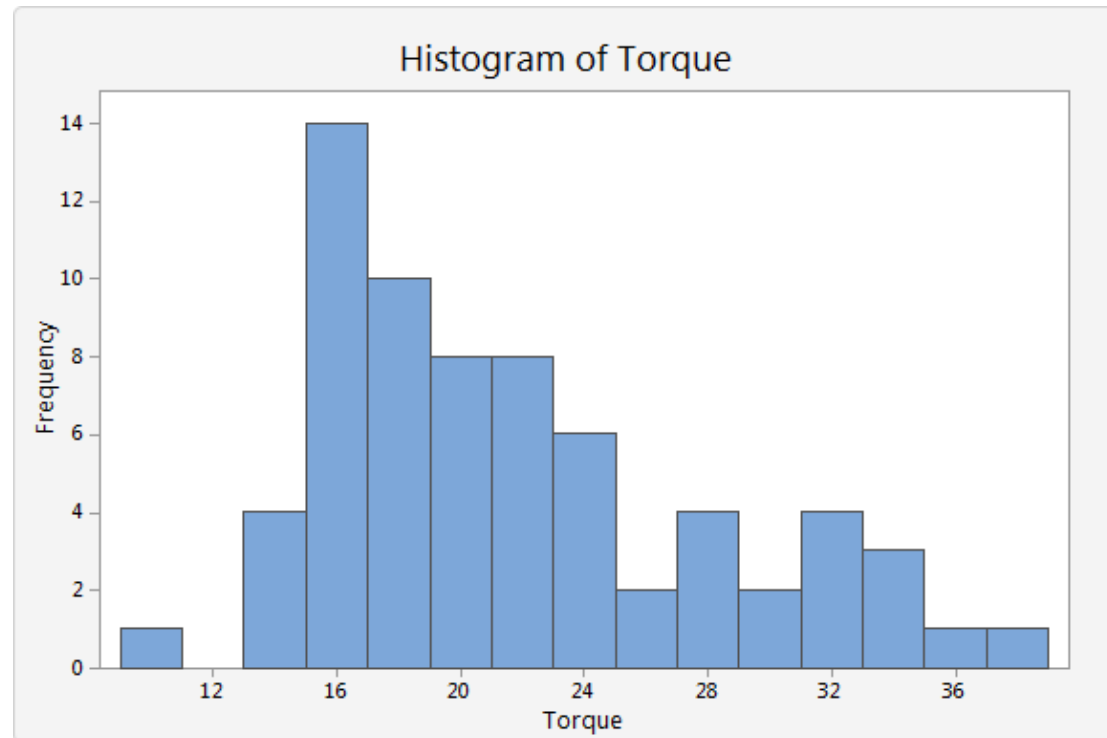


Örnek 2: Detergent.mtw dosyasında C1 sütununda, nominal ağırlığı 4 kg. olan deterjan paketlerinin 500 tanesinin ağırlık ölçümleri yer almaktadır. C2 sütunu ise kullanılan iki dolum hattını göstermektedir.

Dolum hatlarına göre deterjan ağırlıklarının dağılımını inceleyin.

Histogramlar

Verilerin şekli ve dağılımının analizinde kullanılan histogramlar örneklem verilerini gruplandırır ve her bir gruptaki verilerin frekansını çubukla gösterir.



Histogram İçin Şartlar



Histogram için checklist:

- Histogram için örneklem büyüklüğü en az 20 olmalıdır.

Alternatif: Eğer örneklem büyüklüğü 20' den azsa, histogram yerine nokta diyagramı (dotplot) kullanabilirsiniz.

- Veriler rassal örneklem olmalıdır. Aksi takdirde örneklem verisi, ana kitleyi temsil etmez.

Histogram Örnek



Bir kalite kontrol mühendisi şampuan şişesi üzerindeki kapakların doğru bir şekilde tutturulduğuna emin olmak istemektedir. Eğer kapak bağlantısı zayıfsa nakliye sırasında kapaklar açılabilmekte; çok sıkıysa açmak güçleşmektedir. Kapakların hedeflenen tork değeri 18' dir. Mühendisler 68 şişelik rassal bir örneklem almış ve kapakları açmak için gerekli tork miktarlarını test etmiştir.

Analizlerin ilk aşaması olarak mühendisler verilerin merkezini ve dağılımını değerlendirmek için histogram oluşturacaktır.

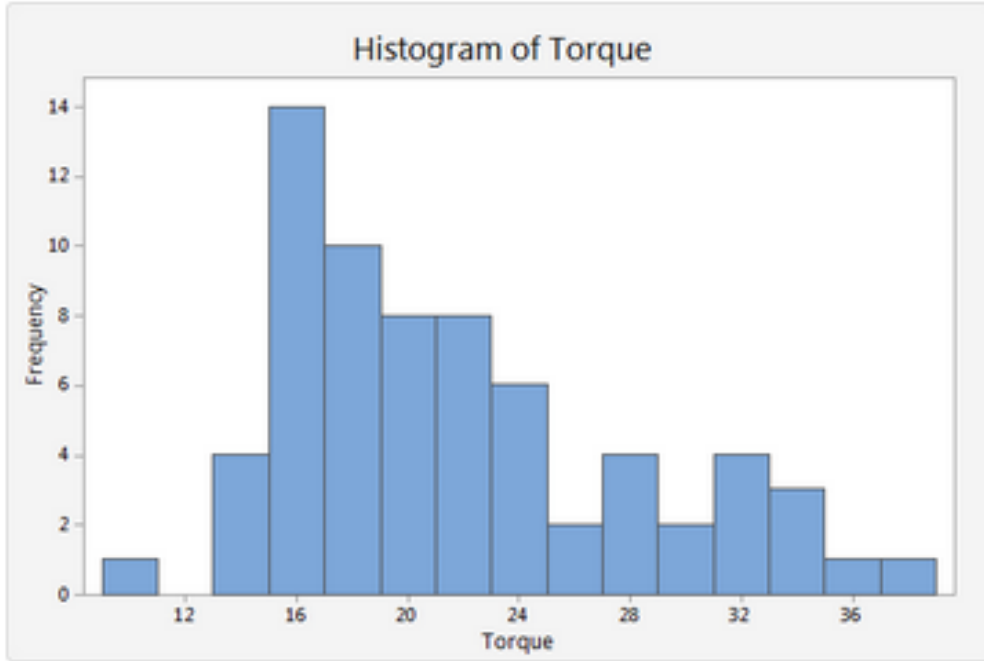
Örneklem veri dosyası Cap.MTW' yi aç.

Graph > Histogram > Simple

Variables, enter *Torque*.

Click **OK**.

Histogram Örnek



- Kapakların çoğu 14-24 tork güç uygulanmasıyla açılmaktadır.
- Yalnızca 1 kapak çok zayıftır ve bu kapak 11' den daha düşük tork değeriyle açılmaktadır.
- Dağılım sağa çarpıktır.
- Önemli miktarda kapak 24' ten fazla torka açılabilir. 5 kapak 33' ten daha fazla bir tork değeriyle açılabilir ki bu değer hedef değerinin neredeyse 2 katıdır.

Descriptive Statistics

N	Mean	StDev	Minimum	Maximum
68	21.2647	6.4220	10	37

Histogram Örnek



Kapakların şişelere bağlantıları 2 ayrı makine de yapılmaktadır. Rassal olarak seçilen şişelerin bağlantılandığı makineler C2 sütununa kaydedilmiştir.

Mühendisler kapakların çıkarılması için gerekli tork güçlerinin makinelere göre farklılık gösterip, göstermediğini araştırmaktadır.

Örnekleme veri dosyası Cap.MTW' yi aç.

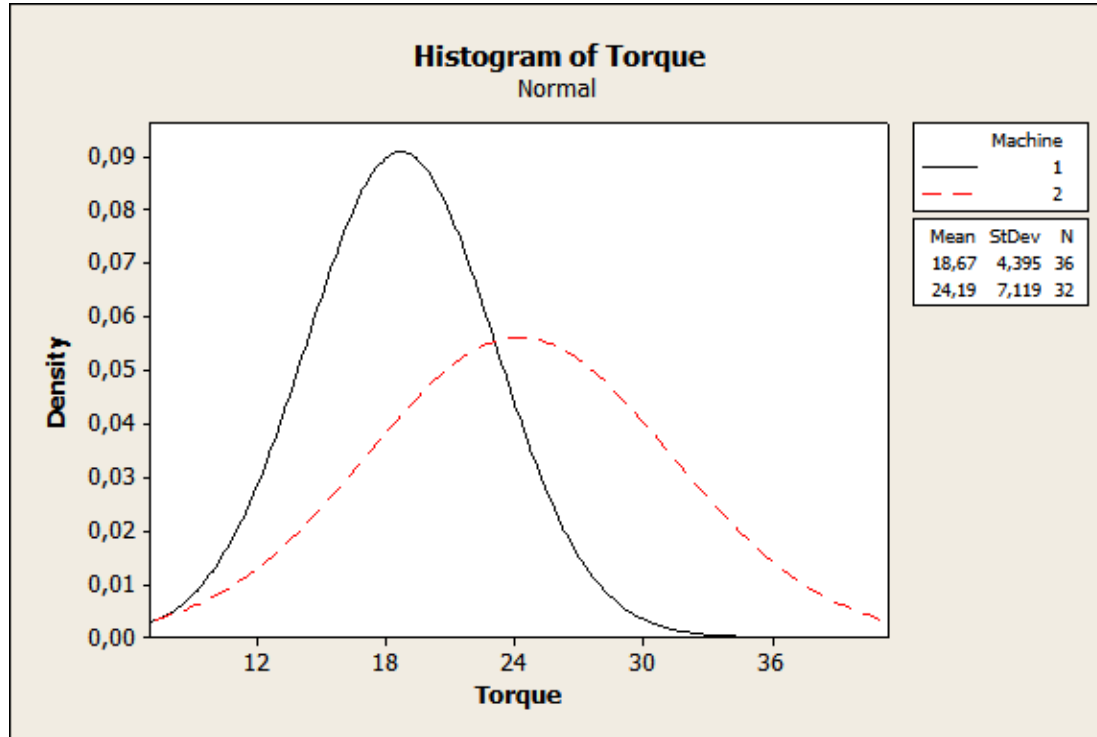
GRAPH > Histogram > With Fit and Groups

Variable, enter *Torque*.

Group variable, enter *Machine*.

Click **OK**.

Histogram Örnek

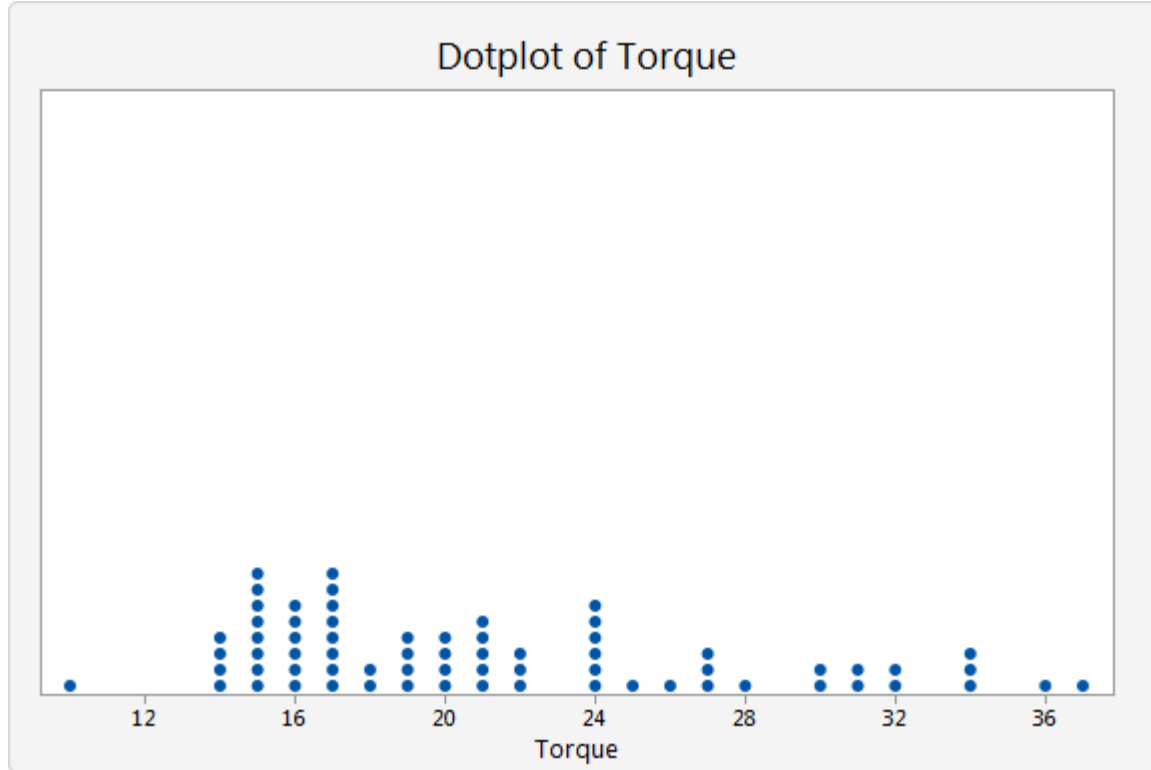


- Histogramlardan makine 1 'de tutturulan kapaklar için gerekli tork miktarlarınının makine 2' den küçük olduğu açıkça görülmektedir.
- Makine 1' in değişkenliği, makine 2' ye göre daha düşüktür.
- Makine 1' in üretimi hedeflenen değere daha uygundur.

Nokta Grafikleri



Nokta grafikleri örnek veri dağılımlarını değerlendirmek ve karşılaştırmak için kullanılır. Bir nokta grafiği örneklem verilerini gruplandırır ve bir çizgi boyunca her bir değeri noktayla gösterir.



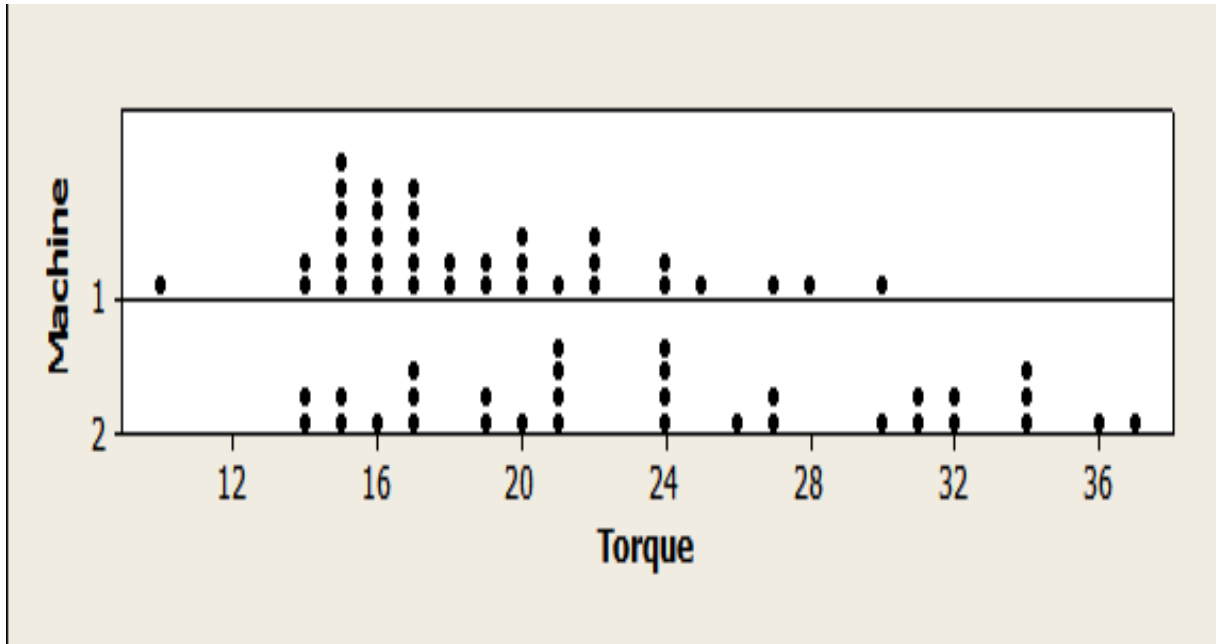
Nokta Grafikleri İçin Koşullar



Nokta Grafikleri İçin Checklist:

- Örneklem büyüklüğü 50' den az olmalı. Eğer örneklem büyüklüğü 50' den fazla ise, nokta grafiği yerine kutu grafiği yada histogram ile dağılım özelliklerini daha kolay gözleyebilirsiniz.
- Veriler, rassal örneklem olmalıdır. Aksi takdirde örneklem verisi, ana kitleyi temsil etmez.

Nokta Grafikleri



Örneklem veri dosyası Cap.MTW' yi aç.

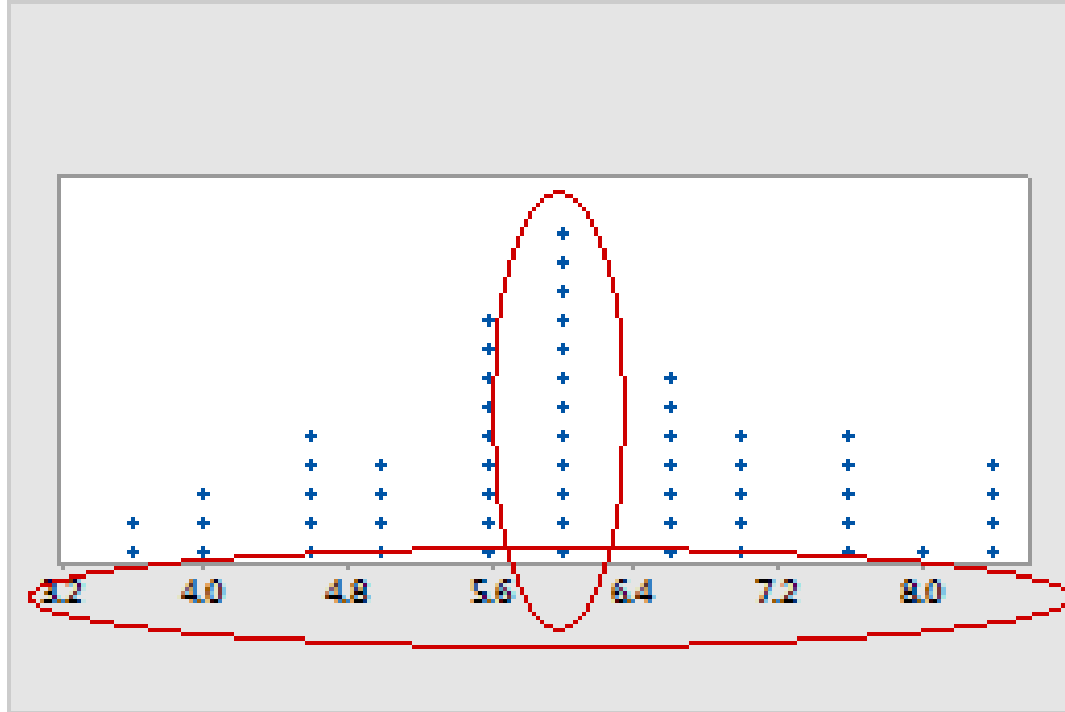
GRAPH > Dotplot > With Groups

Graph variables, *Torque*.

Categorical variables, *Machine*.

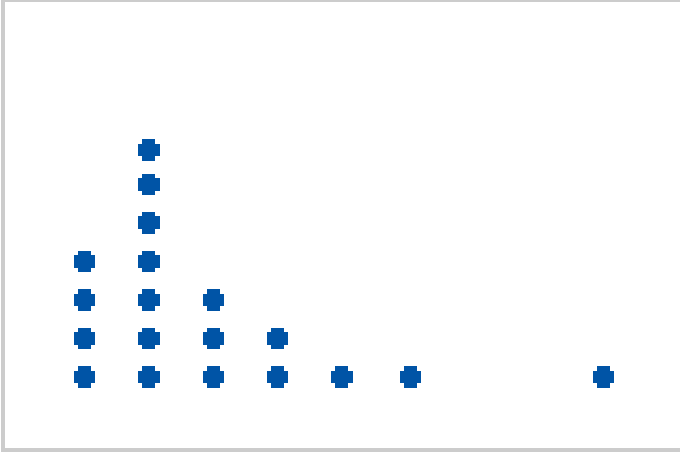
Click **OK**

Nokta Grafiklerin Yorumlanması

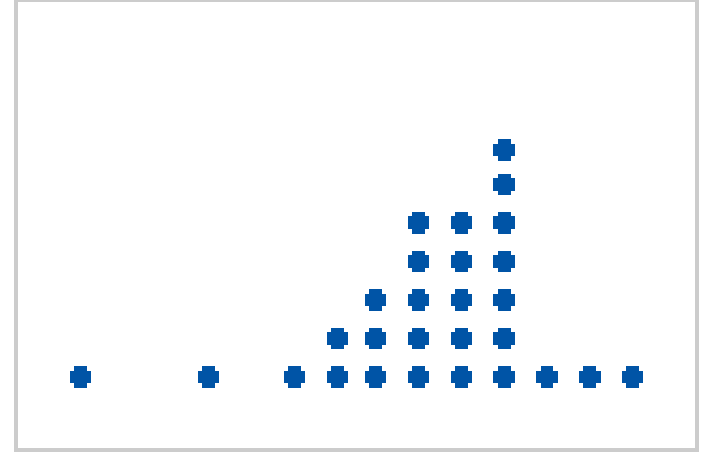


En çok noktanın bulunduğu bölümü tanımla. Burası örneklemede en çok tekrarlanan değeri (mod) gösterir. Verilerin ne kadar değiştiğini anlamak için örneklemin yayılımını değerlendir.

Nokta Grafiklerin Yorumlanması



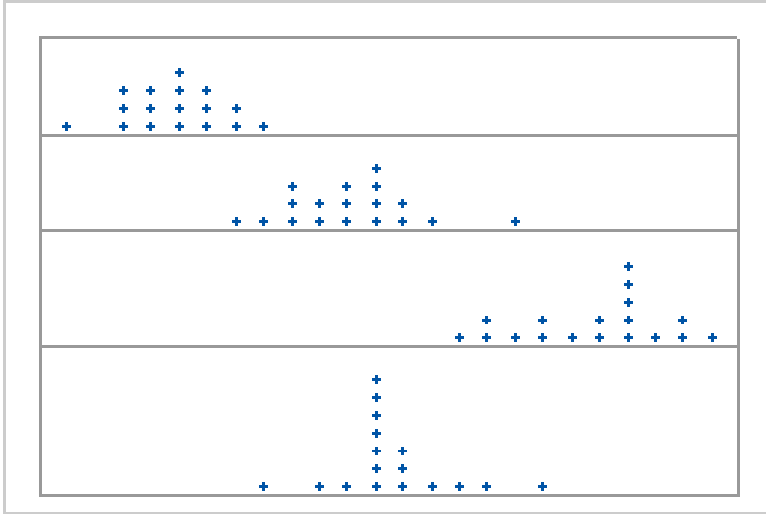
Sağa çarpık



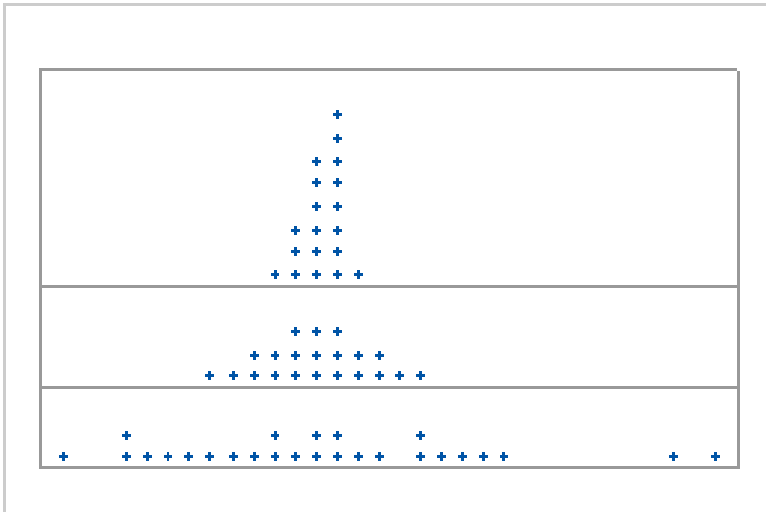
Sola çarpık

Veriler çarpık olması, verilerin çoğunun grafiğin yüksek yada düşük değerli tarafında olduğunu ve verilerin normal dağılmayabileceğini gösterir.

Nokta Grafiklerin Yorumlanması



Grupların merkezleri arasındaki farklılıklara dikkat edin.

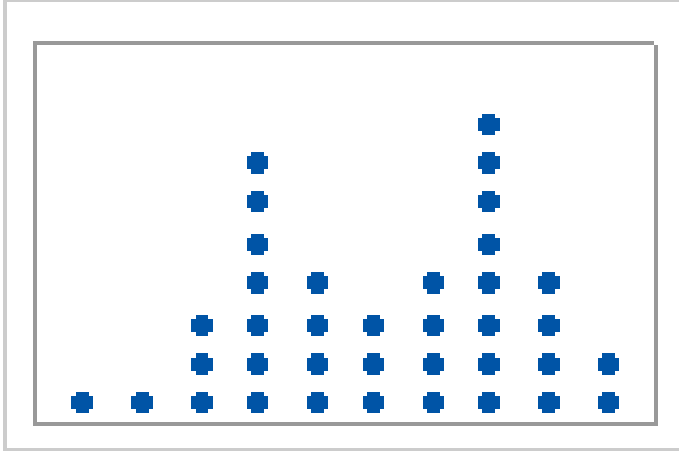


Grupların yayılımları arasındaki farklılıklara dikkat edin.

Nokta Grafiklerin Yorumlanması

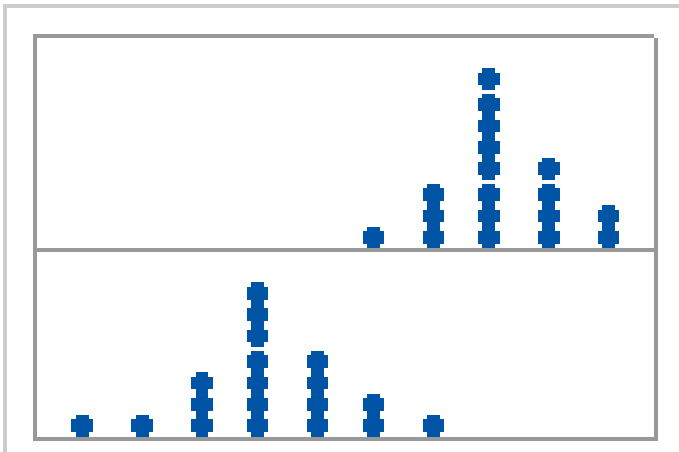


Simple



Çok modlu veriler birden fazla tepe noktasına sahiptir ve genellikle birden fazla süreç yada koşul altında veri toplandığında açığa çıkar.

With Groups



(Örneğin makine 1 ve makine 2 de üretilen kapak bağlantıları için gereken tork güçleri gibi.)