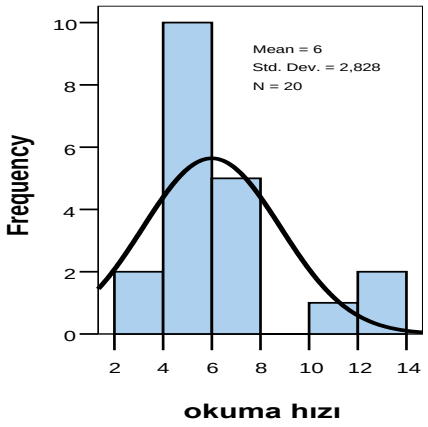
Uygulama-7: Aşağıdaki tabloda 20 öğrencinin dakika cinsinden bir metni okuma hızı verilmiştir. Deneme-3.sav dosyasına bu verileri girip okuma hızı dağılımının normal bir dağılım olup olmadığını **a)** merkezi eğilim, ortalamadan ve normallikten sapma ölçütlerini gösteren tabloya **b)** histogram grafiği **c)** kutu grafiğine ve **d)** Kolomogrow-Simirnov/Shapiro-Wilk testine göre belirleyiniz. **f)** Uç değerleri çıkararak aynı işlemleri yineleyiniz.

Öğrenci	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Okuma hızı (dk)	5	13	4	5	6	11	4	4	5	4	5	6	5	5	6	7	3	3	7	12

a) Merkezi eğilim, ortalamadan ve normallikten sapma ölçütlerini gösteren tabloya elde etmek için Uygulama 6 (a)'daki adımlar, **b)** Histogram grafiği için Uygulama 6 (b)'deki gibi adımları takip edilir ve aşağıdaki sonuçlar elde edilir:

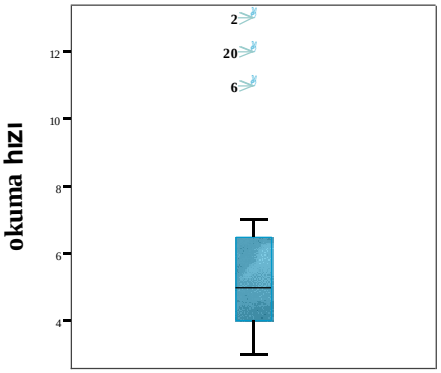
Descriptives

			Statistic	Std. Error
okuma hizi	Mean		6,00	,632
	95% Confidence Interval for Mean	Lower Bound	4,68	
		Upper Bound	7,32	
	5% Trimmed Mean		5,78	
	Median		5,00	
	Variance		8,000	
	Std. Deviation		2,828	
	Minimum		3	
	Maximum		13	
	Range		10	
	Interquartile Range		3	
	Skewness		1,535	,512
	Kurtosis		1,603	,992



Tablo incelendiğinde ÇK > +1 ve BK > 1 dir. Yanı dağılım normal bir dağılım değildir. Ortalama ortancadan büyük olduğundan sağa doğru bir pozitif çarpıklık vardır. Daha kesin hüküm verebilmek için standart çarpıklık değeri hesaplanır: SÇD=1.535/.512=2.998 bulunur. 2.998 > 1.96 olduğu için veriler normal dağılmamıştır.

c) Uygulama 6 (d)'deki adımlar takip edilerek kutu grafiği şekildeki gibi elde edilir. Grafik incelendiğinde kutunun dışında 3 tane uç değer gözlenir. 2., 20. ve 6. veriler uç değerlerdir. Veri setine bakıldığında bu değerler 13, 12 ve 11dk'lık okuma hızlarına tekabul eder.**d)** Gözlem sayısı



Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
okuma hizi	,250	20	,002	,795	20	,001

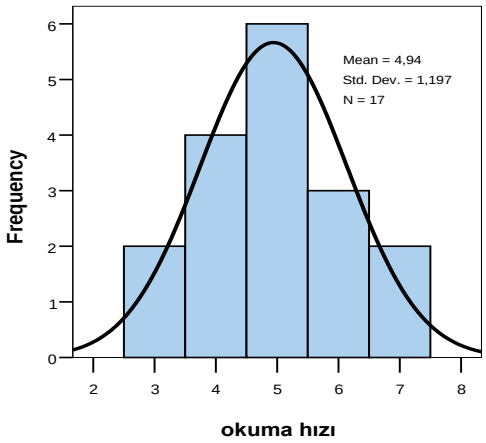
a. Lilliefors Significance Correction

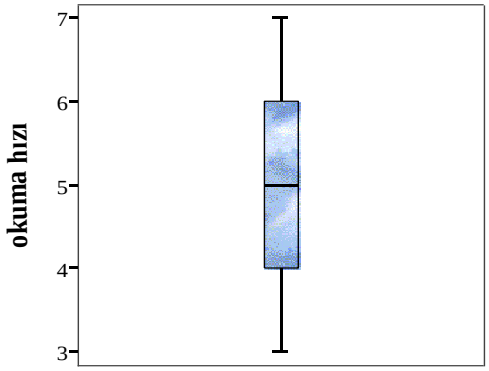
Sig. Değeri (p=0,01) %5'den küçük olduğu için %5 anlamlılık düzeyinde dağılım normal değildir.

f) Uç değerleri çıkararak aynı işlemleri yinelendiğinde aşağıdaki tablo ve grafikler elde edilir:

Descriptives

			Statistic	Std. Error
okuma hizi	Mean		4,94	,290
	95% Confidence Interval for Mean	Lower Bound	4,33	
		Upper Bound	5,56	
	5% Trimmed Mean		4,93	
	Median		5,00	
	Variance		1,434	
	Std. Deviation		1,197	
	Minimum		3	
	Maximum		7	
	Range		4	
	Interquartile Range		2	
	Skewness		,126	,550
	Kurtosis		-,501	1,063





Tabloya göre -1 < ÇK < +1 arasına düşmüştür, yani dağılımın artık normale dönüştüğü söylenebilir. Ayrıca SÇD=.126/.550=.229 bulunur. -1.96 < .229 < 1.96 olduğu için veriler normal dağılmıştır. Ancak ÇK değeri pozitif ve ortalama ortancadan biraz büyük olduğundan SAĞA doğru çok hafif bir çarpıklık vardır. BK negatif olduğundan normal eğrisi normalden biraz basıktır. Kutu grafiğinde uç değerler artık kalmamıştır. Kutu ortasındaki çizgi alt ve üst çizginin hemen hemen ortasına (biraz altında) düşmektedir. Bu da dağılımın çok hafif sağa çarpık olmakla birlikte normal bir dağılım olduğunu gösterir.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
okuma hizi	,186	17	,120	,930	17	,215

a. Lilliefors Significance Correction

Shapiro-Wilk normallik testinde ise Sig. Değeri (0.215) 0.05'den büyük olduğu için %5 anlamlılık düzeyinde dağılım normal değildir.

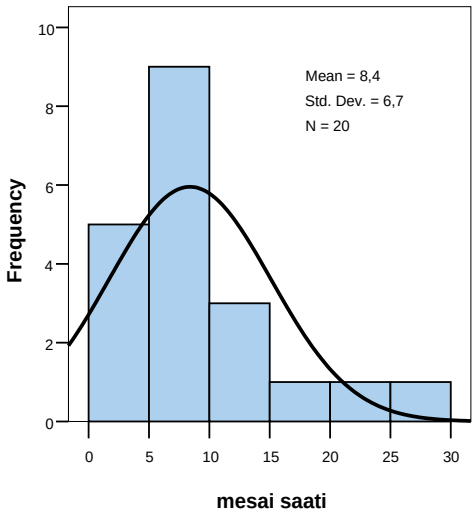
Normal Dağılıma Dönüştürme İşlemleri: Normal dağılım göstermeyen bir veri seti normal dağılıma belirli yöntemlerle dönüştürülebilir. _Orta derecede bir pozitif çarpıklık durumunda ve veriler ondalık sayı değilse *karekök dönüşümü*, aşırı pozitif çarpıklık durumunda ise *logaritmik dönüşüm* yapılabilir. Orta derecede negatif bir çarpıklığın dönüşümü için birinci alternatif; önce pozitif çarpık bir dağılıma ardından da karekök dönüşümü ile normal dağılıma dönüştürmektir. İkinci alternatif ise x^2 veya x^3 ya da Log[x/(1-x)] dönüşümü yapmaktır. Aşırı derecede negatif bir çarpıklık $Log_e[p/(1-p)]$ dönüşümü (p=1/x) ile normal bir dağılıma dönüştürülebilir.

Uygulama-8: Aşağıdaki tabloda 20 işçinin haftalık mesai saati verilmiştir. Deneme-4.sav dosyasına bu verileri girip mesai saati dağılımının normal bir dağılım olup olmadığını **a)** merkezi eğilim, ortalamadan ve normallikten sapma ölçütleri tablosunu **b)** histogram grafiği **c)** kutu grafiğine ve **d)** Kolomogrow-Simirnov/Shapiro-Wilk testine göre belirleyiniz. **e)** Eğer veriler normal dağılımda değilse uygun bir yöntemle verileri normal dağılıma dönüştürerek aynı işlemleri yineleyiniz.

Öğrenci	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Okuma hızı (dk)	5	13	4	1	6	11	2	4	5	21	5	6	5	19	6	7	26	3	7	12

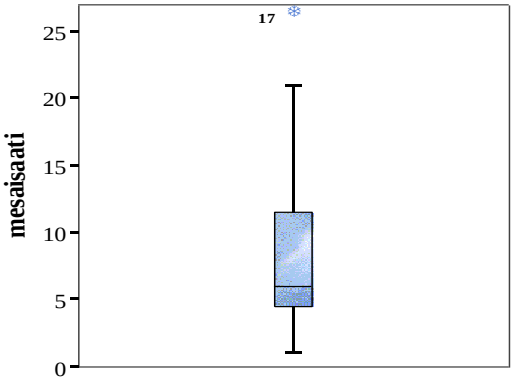
a) Merkezi eğilim, ortalamadan ve normallikten sapma ölçütlerini gösteren tabloya elde etmek için Uygulama 6 (a)'daki adımlar, **b)** Histogram grafiği için Uygulama 6 (b)'deki gibi adımları takip edilir ve aşağıdaki sonuçlar elde edilir:

			Statistic	Std. Error
mesai saati	Mean		8,40	1,498
	95% Confidence Interval for Mean	Lower Bound	5,26	
		Upper Bound	11,54	
	5% Trimmed Mean		7,83	
	Median		6,00	
	Variance		44,884	
	Std. Deviation		6,700	
	Minimum		1	
	Maximum		26	
	Range		25	
	Interquartile Range		8	
	Skewness		1,479	,512
	Kurtosis		1,530	,992



Tablo incelendiğinde ÇK > +1 ve BK > 1 dir. Yanı dağılım normal bir dağılım değildir. Ayrıca SÇD=1.479/.512=2.889 bulunur. 2.889 > 1.96 olduğu için veriler normal dağılmamıştır. Ortalama ortancadan büyük olduğundan sağa doğru bir pozitif çarpıklık vardır ve bu grafikten de net bir şekilde görülmektedir.

c) Uygulama 6 (d)'deki adımlar takip edilerek kutu grafiği şekildeki gibi elde edilir. Grafik incelendiğinde kutunun dışında 1 tane uç değer gözlenir. Kutu ortasındaki çizgi alt ve üst çizginin çok altına düşmektedir. Bu da dağılımın oldukça sağa pozitif çarpık olduğunu gösterir.



d) Shapiro-Wilk normallik testinde Sig. Değeri (0,002) %5'den küçük olduğu için %5 anlamlılık düzeyinde dağılım normal değildir

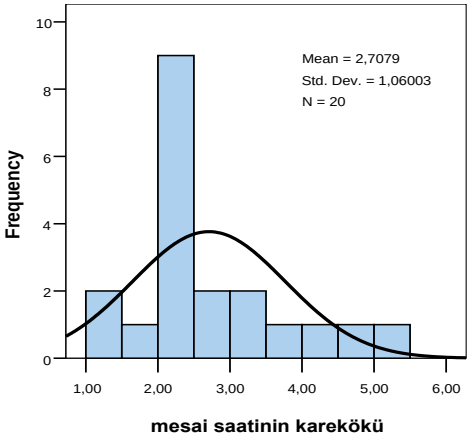
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
mesai saati	,283	20	,000	,820	20	,002

a. Lilliefors Significance Correction

e) Mesai saati verileri normal dağılımda değildir. Çarpıklık orta düzey kabul edilirse karekök dönüşümü yapılabilir. Bunun için şu adımlar takip edilir:
Transform → Compute (Açılan sayfada "target variable kısmına "karekök_mesai saati" yazılır, eşitliğin sağ kısmındaki pencereye mesai saati değişkeni aktarılır **0,5 yazılır) → OK. Veri sayfasında mesai saatinin karekök değerlerini gösteren sütün görülür. Mesai saati için yapılan normallik testleri mesai saatinin karekök değerleri için de yapılırsa aşağıdaki sonuçlar elde edilir:

			Statistic	Std. Error
karekök mesai saati	Mean		2,7079	,23703
	95% Confidence Interval for Mean	Lower Bound	2,2118	
		Upper Bound	3,2040	
	5% Trimmed Mean		2,6699	
	Median		2,4495	
	Variance		1,124	
	Std. Deviation		1,06003	
	Minimum		1,00	
	Maximum		5,10	
	Range		4,10	
	Interquartile Range		1,37	
	Skewness		,829	,512
	Kurtosis		,267	,992

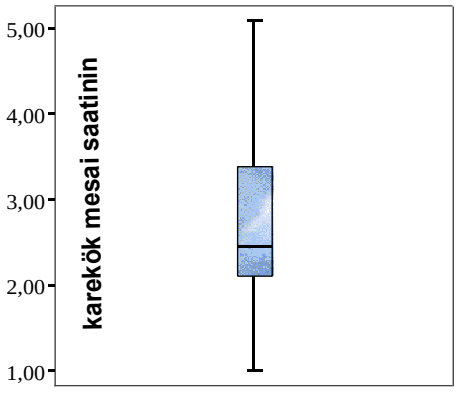


Tablo incelendiğinde -1 < ÇK < +1 arasına düşmüştür, yani dağılımın artık normale dönüştüğü söylenebilir. Ayrıca SÇK=.829/.512=1.619 bulunur. -1.96 < 1.619 < 1.96 olduğu için veriler normal dağılmıştır. Ancak ÇK değeri pozitif ve ortalama ortancadan biraz büyük olduğundan grafikten de görüldüğü gibi sağa doğru pozitif bir çarpıklık vardır. BK negatif olduğundan normal eğrisi normalden biraz basıktır. Kutu grafiğinde uç değer artık kalmamıştır. Kutu ortasındaki çizgi alt ve üst çizginin altına düşmektedir. Bu da dağılımın sağa çarpık olmakla birlikte normal bir dağılım olduğunu gösterir. Shapiro-Wilk normallik testinde ise Sig. Değeri (0,109) %5'den büyük olduğu için %5 anlamlılık düzeyinde dağılım normal değildir.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
mesai saatinin karekökü	,223	20	,010	,922	20	,109

a. Lilliefors Significance Correction



Çarpıklık aşırı ise logaritmik dönüşüm yapılabilir. Bunun için şu adımlar takip edilir: Transform → Compute (Açılan sayfada "target variable kısmına dönüşümün yeni adı yazılır, functions kısmından lg10 bulunup yukarı yani Numeric Expression kısmına aktarılır. Lg10 (?) de ? kısmına dönüştürülecek değişken aktarılır) → OK. Veri sayfasında ilgili değişkenin logaritmik değerlerini gösteren sütün görülür.