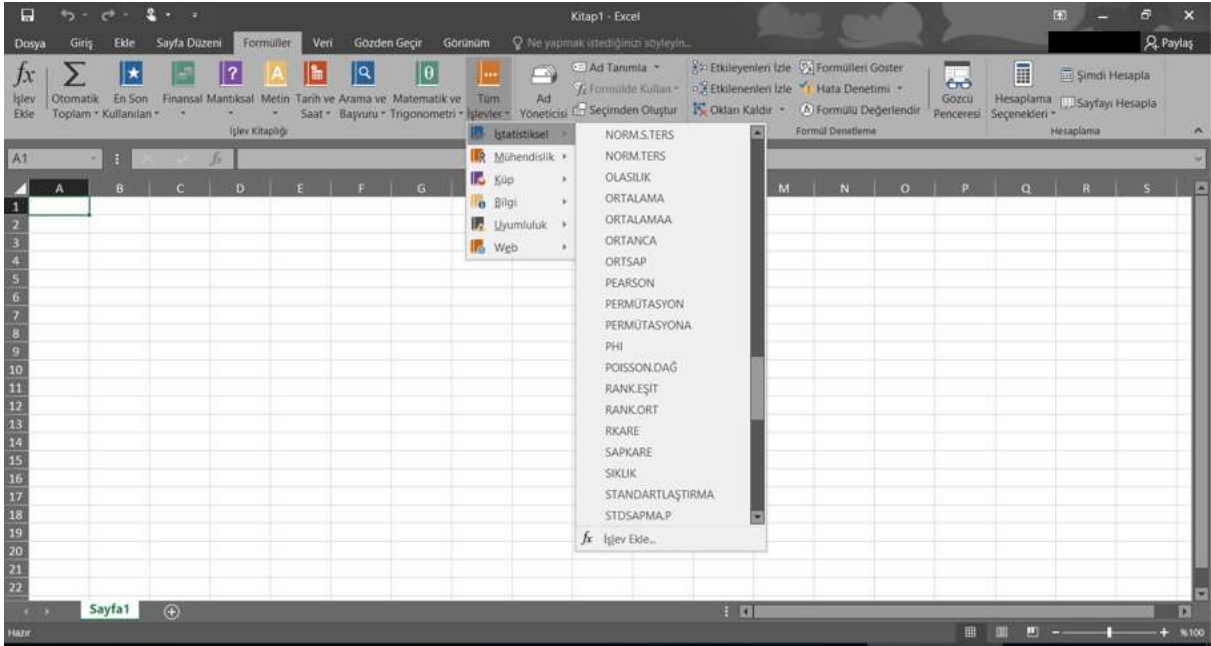


EXCEL'DE İSTATİSTİKSEL FONKSİYONLAR

Excel'in istatistiksel yönlerinin anlaşılması için istatistiksel fonksiyonlar ve veri analizi araçları önemli bir yere sahiptir. Çalışma sayfası fonksiyonları, Excel'de bulunan bir seri hesaplamalar yaparak kullanıcıların kendi kendine hesaplamalar yapması sırasında oluşabilecek hataları engelleyen hazır yerleşik formüller olarak yer almaktadır. Veri analizi araçları ise formüllerinde ötesinde olmakla birlikte kullanıcılara açıklayıcı sonuçlar sağlamaktadır.

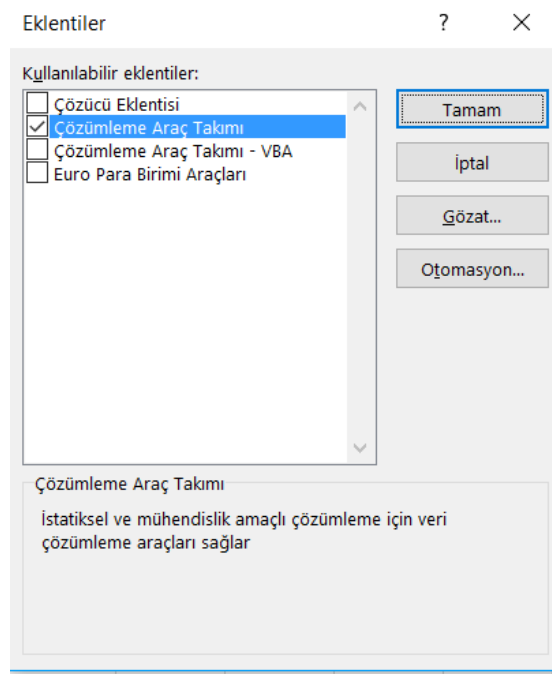
Bu bölümde istatistiksel fonksiyonlara hızlıca ulaşmak için kısayol bilgisi verilmektedir. Excel istatistiksel fonksiyonlarına Şekil 1'de görüldüğü gibi **Formüller** (Formulas) komut sekmesinde bulunan **İşlev kitaplığı** (Function Library) komut kümesinde bulunan **Tüm işlevler** (More Functions) simgesine tıklanması ile açılan listeden **İstatistiksel** (Statistical) seçeneğine tıklanması ile ulaşılabilir.



Şekil 1: İşlev Kitaplığı (Function Library) komut kümesi, Tüm İşlevler (More functions) simgesi, İstatistiksel (Statistical) fonksiyonlar.

VERİ ÇÖZÜMLEME ARAÇ TAKIMINI AKTİF HALE GETİRME

Excel veri çözümlenme araç takımını kullanmadan önce **Çözümleme Araç Takımı** (Analysis Toolpak) eklentisinin (add-ins) aktif hale getirilmesi gerekmektedir. Bunun için **Dosya** (File) komut sekmesinden **Seçenekler**'e (Options), **Excel Seçenekleri** (Excel Options) iletişim kutusunun **Eklentiler** (Add-Ins) bölümünün altında bulunan **Yönet** (Manage) kısmından **Excel Add-Ins** seçildikten sonra **Git** (Go) düğmesine tıklandıktan sonra Şekil 2'de görülen **Eklentiler** (Add-Ins) iletişim kutusu açılmaktadır.



Şekil 2: Çözümleme Araç Takımı (Analysis Toolpak) seçeneği ile Excel veri çözümleme araç takımı aktif hale getirilmektedir.

Excel programı veri analizi yapılması için çok karmaşık araçlara sahiptir. Excel veri çözümleme araç takımı aktif hale getirildikten sonra ileri düzeyde istatistiksel veri çözümleme araçları ile analizler gerçekleştirilebilmektedir. Fakat biz bu dersimizde sadece Tanımlayıcı İstatistik (Descriptive Statistics) ve Histogram analizlerini gerçekleştireceğiz.

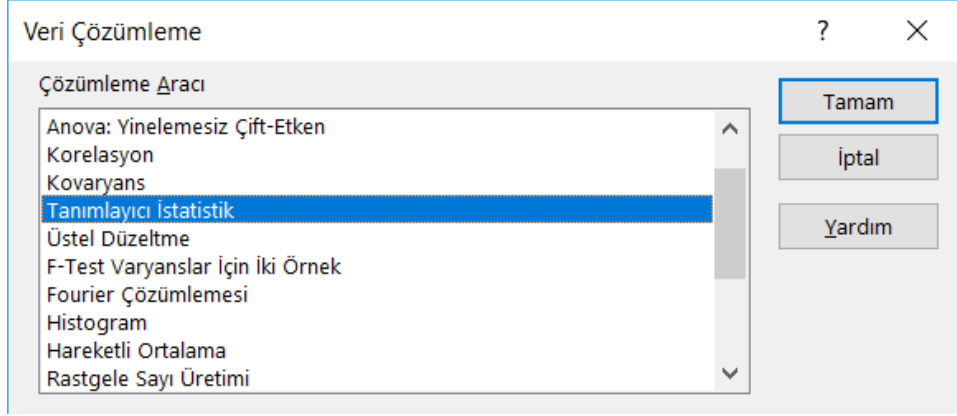
Tanımlayıcı İstatistik (Descriptive Statistics): Seçilen hücrelerde değerlerin merkezi eğilimini (central tendency), değişkenlik (variability) ve diğer karakteristik özelliklerini gösteren bir rapor oluşturmaktadır.

Histogram: Seçilen hücrelerdeki değerler için bireysel ve birikimli (kümülatif) frekansları çizelge haline getirmektedir.

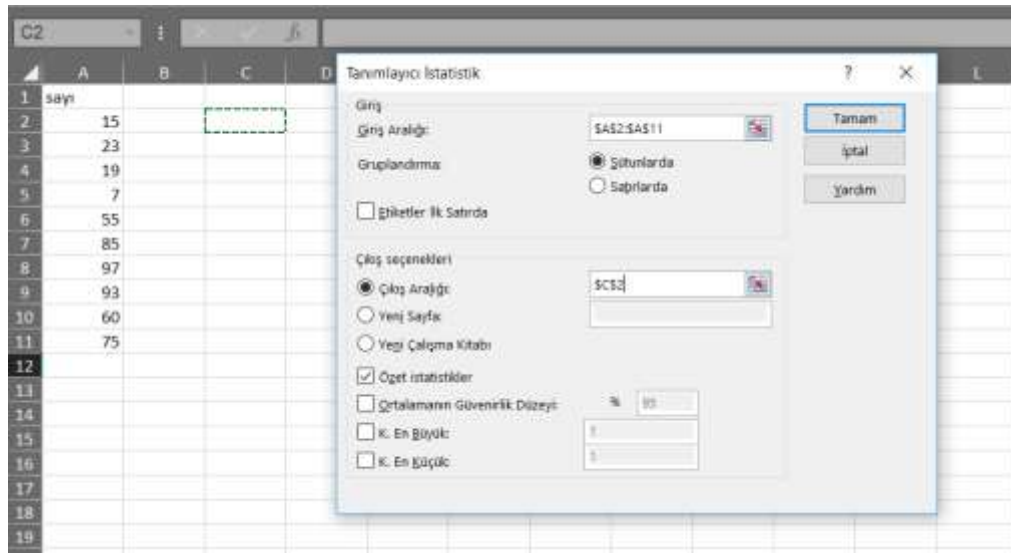
VERİ ÇÖZÜMLEME ARAÇ TAKIMINI KULLANMA

Veri çözümleme araç takımı kullanımını örneklemek için belirlenen değerlerin bazı istatistiklerini özetleyen Tanımlayıcı İstatistik aracını kullanmak için aşağıdaki adımlar izlenir:

1. Şekil 4'te görüldüğü gibi A2 ile A11 hücreleri arasına (A2:A11) tanımlayıcı istatistik değerleri için istenen veri değerlerini girin.
2. **Veri** (Data) komut sekmesi **Çözümleme** (Analysis) komut kümesi altında bulunan **Veri Çözümleme** (Data Analysis) simgesine tıklayın. Şekil 3'te görülen Veri Çözümleme iletişim kutusu ekrana gelmektedir.



Şekil 3: Veri çözümleme iletişim kutusunda bütün veri çözümleme (veri analizi araçları) listelenmektedir.



Şekil 4: Tanımlayıcı İstatistikler penceresi.

	A	B	C	D
1	sayı			
2	15		<i>Sütun1</i>	
3	23			
4	19		Ortalama	52,9
5	7		Standart Hata	10,91018892
6	55		Ortanca	57,5
7	85		Kip	#YOK
8	97		Standart Sapma	34,50104668
9	93		Örnek Varyans	1190,322222
10	60		Basıklık	-1,829562102
11	75		Çarpıklık	-0,075577453
12			Aralık	90
13			En Büyük	7
14			En Küçük	97
15			Toplam	529
16			Say	10
17				

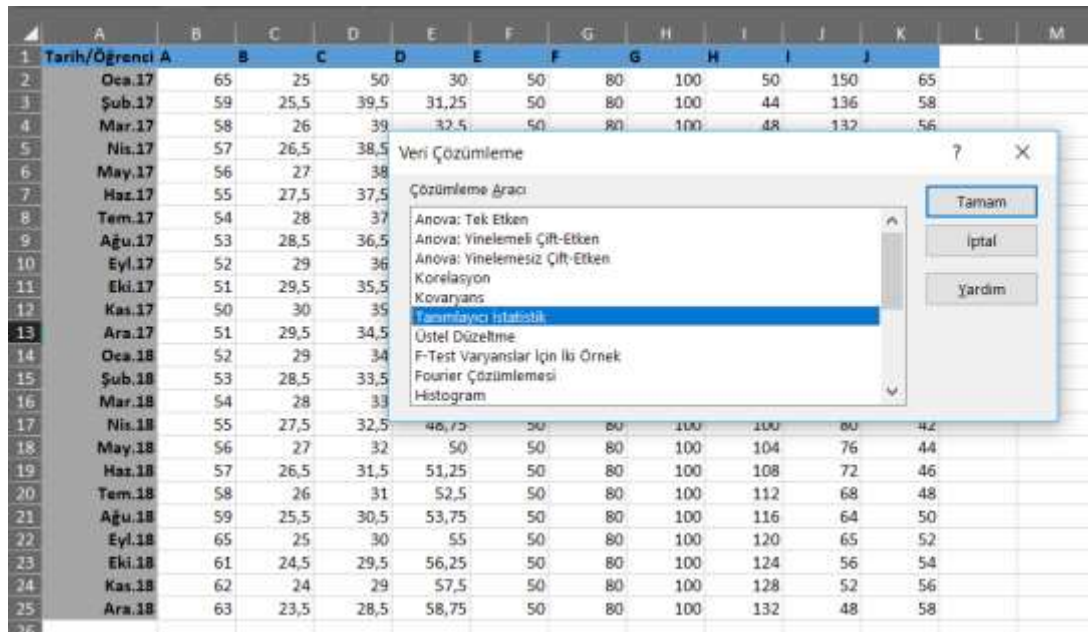
Şekil 5: Özet İstatistikler

Not: Microsoft office, veri çözümlene aracını Windows işletim sistemlerine sunmakta iken artık Mac sürümüne sahip sistemler için sunmamaktadır. Bu yüzden Excel ile İstatistiksel veri analizi için Mac işletim sistemlerinde Veri çözümlene aracı yerine StatPlus gibi (genellikle ücretli olan) eklentilerin (add-in) Excel programına yüklenmesi gerekmektedir.

VERİLERİ TANIMLAYICI İSTATİSTİKLER İLE ÖZETLEME

Excel'de veriler genellikle Grafikler, Histogramlar, Özet Tablolar (Pivot Tables) ve Tanımlayıcı İstatistik (Descriptive Statistics) gibi bazı araçlar kullanılarak özetlenebilmekte ve tanımlanabilmektedir. Veriler; Excel tarafından Tanımlayıcı İstatistik (Descriptive Statistics) olarak adlandırılan ortalama, medyan, standart sapma ve varyans gibi belirli özellikleri ile tanımlanabilmektedir.

Şekil 6'da Tanımlayıcı istatistikleri hesaplamak için her ay farklı testlere giren 10 öğrencinin 100 üzerinden aldıkları sınav sonuçlarını gösterebiliriz.



Tarih/Öğrenci	A	B	C	D	E	F	G	H	I	J	K	L	M
Oca.17	65	25	50	30	50	80	100	50	150	65			
Şub.17	59	25,5	39,5	31,25	50	80	100	44	136	58			
Mar.17	58	26	39	32,5	50	80	100	48	132	56			
Nis.17	57	26,5	38,5										
May.17	56	27	38										
Haz.17	55	27,5	37,5										
Tem.17	54	28	37										
Ağu.17	53	28,5	36,5										
Eyl.17	52	29	36										
Eki.17	51	29,5	35,5										
Kas.17	50	30	35										
Ara.17	51	29,5	34,5										
Oca.18	52	29	34										
Şub.18	53	28,5	33,5										
Mar.18	54	28	33										
Nis.18	55	27,5	32,5	46,75	50	80	100	100	80	42			
May.18	56	27	32	50	50	80	100	104	76	44			
Haz.18	57	26,5	31,5	51,25	50	80	100	108	72	46			
Tem.18	58	26	31	52,5	50	80	100	112	68	48			
Ağu.18	59	25,5	30,5	53,75	50	80	100	116	64	50			
Eyl.18	65	25	30	55	50	80	100	120	65	52			
Eki.18	61	24,5	29,5	56,25	50	80	100	124	56	54			
Kas.18	62	24	29	57,5	50	80	100	128	52	56			
Ara.18	63	23,5	28,5	58,75	50	80	100	132	48	58			

Şekil 6: Veri komut sekmesinden Veri çözümlene simgesine tıklanınca açılan listeden Tanımlayıcı İstatistik aracı seçilmektedir.

Tarih/Öğrenci	A	B	C	D	E	F	G	H	I	J	
2	Oca.17	65	25	50	30	50	80	100	50	150	65
3	Şub.17	59	25,5	39,5	31,25						
4	Mar.17	58	26	39	32,5						
5	Nis.17	57	26,5	38,5	33,75						
6	May.17	56	27	38	35						
7	Haz.17	55	27,5	37,5	36,25						
8	Tem.17	54	28	37	37,5						
9	Ağu.17	53	28,5	36,5	38,75						
10	Eyl.17	52	29	36	50						
11	Eki.17	51	29,5	35,5	41,25						
12	Kas.17	50	30	35	42,5						
13	Ara.17	51	29,5	34,5	43,75						
14	Oca.18	52	29	34	45						
15	Şub.18	53	28,5	33,5	46,25						
16	Mar.18	54	28	33	47,5						
17	Nis.18	55	27,5	32,5	48,75						
18	May.18	56	27	32	50						
19	Haz.18	57	26,5	31,5	51,25						
20	Tem.18	58	26	31	52,5						
21	Ağu.18	59	25,5	30,5	53,75						
22	Eyl.18	65	25	30	55						
23	Eki.18	61	24,5	29,5	56,25						
24	Kas.18	62	24	29	57,5						
25	Ara.18	63	23,5	28,5	58,75						

Tanımlayıcı İstatistik

Giriş:

Giriş Aralığı:

Gruplandırma: Sütunlarda Satırlarda

Etiketler ilk Satırda

Çılgı seçenekleri:

Çılgı Aralığı:

Yeni Sayfa:

Yeni Çalışma Kitabı

Özet İstatistikler

Ortalamanın Güvenirlik Düzeyi:

K. En Büyük:

K. En Küçük:

Tamam İptal Yardım

Şekil 7: 10 öğrencinin her ay aldıkları test sonuçları için tanımlayıcı istatistikler için giriş aralığı B2:K25 olarak ve sütunlar bazında gruplandırma yapılarak, yeni çalışma sayfasında görüntülenmek üzere belirlenmektedir.

Sütun1	Sütun2	Sütun3	Sütun4	Sütun5	Sütun6	Sütun7	Sütun8	Sütun9	Sütun10										
Ortalama	56,5	Ortalama	26,9583	Ortalama	34,6667	Ortalama	44,7917	Ortalama	50	Ortalama	80	Ortalama	100	Ortalama	86,625	Ortalama	94,625	Ortalama	49,0417
Standart Hata	0,89281	Standart Hata	0,37821	Standart Hata	0,95015	Standart Hata	1,8084	Standart Hata	0	Standart Hata	0	Standart Hata	0	Standart Hata	5,60403	Standart Hata	5,88163	Standart Hata	1,45298
Ortanca	56	Ortanca	27	Ortanca	34,25	Ortanca	45,625	Ortanca	50	Ortanca	80	Ortanca	100	Ortanca	86	Ortanca	94	Ortanca	50
Kip	85	Kip	25	Kip	#YOK	Kip	50	Kip	50	Kip	80	Kip	100	Kip	#YOK	Kip	#YOK	Kip	50
Standart S	4,37384	Standart S	1,85283	Standart S	4,65475	Standart S	8,85931	Standart S	0	Standart S	0	Standart S	0	Standart S	27,454	Standart S	28,814	Standart S	7,12301
Örnek Var	19,1204	Örnek Var	3,43297	Örnek Var	21,6667	Örnek Var	78,4873	Örnek Var	0	Örnek Var	0	Örnek Var	0	Örnek Var	753,723	Örnek Var	830,245	Örnek Var	50,7373
Basıklık	-0,58653	Basıklık	-0,97226	Basıklık	3,80645	Basıklık	-1,20151	Basıklık	#SAYI/DI	Basıklık	#SAYI/DI	Basıklık	#SAYI/DI	Basıklık	-1,23191	Basıklık	-0,58799	Basıklık	-0,17054
Çarpıklık	0,50107	Çarpıklık	-0,12999	Çarpıklık	3,45426	Çarpıklık	-0,13619	Çarpıklık	#SAYI/DI	Çarpıklık	#SAYI/DI	Çarpıklık	#SAYI/DI	Çarpıklık	0,06071	Çarpıklık	0,11378	Çarpıklık	0,07233
Aralık	15	Aralık	6,5	Aralık	21,5	Aralık	28,75	Aralık	0	Aralık	0	Aralık	0	Aralık	88	Aralık	102	Aralık	29
En Büyük	50	En Büyük	23,5	En Büyük	28,5	En Büyük	30	En Büyük	50	En Büyük	80	En Büyük	100	En Büyük	44	En Büyük	48	En Büyük	36
En Küçük	65	En Küçük	30	En Küçük	50	En Küçük	58,75	En Küçük	50	En Küçük	80	En Küçük	100	En Küçük	132	En Küçük	150	En Küçük	65
Toplam	1356	Toplam	647	Toplam	832	Toplam	1075	Toplam	1200	Toplam	1800	Toplam	2000	Toplam	2079	Toplam	2271	Toplam	1377
Say	24	Say	24	Say	24	Say	24	Say	24	Say	24	Say	24	Say	24	Say	24	Say	24

Şekil 8: Her ay farklı testlere katılan 10 öğrencinin 100 üzerinden aldıkları sınav sonuçları verileri için tanımlayıcı istatistikler görüntülenmektedir.

Tanımlayıcı İstatistik aracı ile Şekil 8’de görüldüğü gibi Ortalama, Standart Hata, Ortanca, Mod, Standart Sapma, Örneklem Varyansı, Basıklık, Çarpıklık, Aralık, En büyük, En küçük, Toplam ve Say olmak üzere 13 adet istatistik hesaplanmaktadır.

Tanımlayıcı istatistikler aracı sonuçlarının her bir ögesinin anlaşılması ve yorumlanması gerekmektedir. Öncelikle, verilerin burada öğrenci sonuçlarının merkezi eğiliminin belirlenmesi için Ortalama (Mean/Average), Ortanca (Median) ve Mod (Mode) olmak üzere üç ana ölçü kullanılmaktadır.

Ortalama, veri setinde bulunan değerlerin aritmetik ortalamasını ifade etmektedir. Excel’in ortalama fonksiyonu ile de uygun hücreler kullanılarak hesaplanabilmektedir.

Ortanca (Median), veri seti küçükten büyüğe sıralandığında en ortada bulunan değerdir.

Mod (Mode) veriler arasında en sık bulunan değer olup mod, medyan ve ortalamanın eşit olduğu simetrik veri setinin dışında genel olarak merkezi eğilimin belirlenmesi için hiç kullanılmamaktadır. Excel’in ENÇOK_OLAN (MODE) fonksiyonu ile de uygun hücreler kullanılarak hesaplanabilmektedir. Şekil 8’de görüldüğü gibi bazı test sonuçları değerleri için eğer birden fazla tekrarlayan hiç değer yok ise fonksiyon #YOK (#N/A) hata değerini vermektedir.

Eğer verinin Çarpıklık (Skewness) değeri yüksek ise veride bulunan sapan değerlerin ortalama değeri etkilemesi nedeniyle ortalama yerine medyan daha sağlıklı bir merkezi eğilim belirleyicisi olmakta iken Çarpıklık değeri düşük ise ortalama daha sağlıklı bir merkezi eğilim belirleyicisi olmaktadır. Çarpıklık, bir frekans dağılımının asimetriklik ölçüsüdür.

Şekil 8'de görülen bir diğer tanımlayıcı istatistik değeri olan çarpıklık değeri veri setinin çarpıklığını ölçmektedir eğer; +1'den yüksek ise yüksek düzeyde pozitif çarpıklığı göstermekte iken, -1'den düşük ise yüksek düzeyde negatif çarpıklığı göstermekte ve -1 ile +1 değerleri arasında ise simetrik dağılıma sahip veri setini göstermektedir. Ayrıca, Excel'in Çarpıklık (SKEW) fonksiyonu ile de veri setlerinin çarpıklığı hesaplanabilmektedir.

Şekil 8'de görülen bir diğer tanımlayıcı istatistik değeri olan basıklık (Kurtosis) değeri, veri setinin basıklık durumunu göstermekte olup 0'a yakın ise veri setinin normal ya da standart çan eğrisi dağılımına yakın olduğunu göstermektedir. Pozitif basıklık değeri normal dağılıma göre daha fazla tepeli bir dağılımı ifade etmekte iken negatif basıklık değeri normal dağılıma göre daha fazla basık bir dağılımı ifade etmektedir.

Şekil 8'de görülen bir diğer tanımlayıcı istatistik değeri olan Aralık (RANGE) değeri, veri setinin maksimum değerinden minimum değerinin çıkarılması ile hesaplanmakta olup verinin yayılım durumunu göstermektedir.

Kaynakça

İltir, Cenk (2018). Excel ile İstatistiksel Veri Analizi. Seçkin Yayıncılık, Ankara.