



**Hacettepe Üniversitesi**  
**Fen Fakültesi**  
**İstatistik Bölümü**  
**İST268 İstatistiksel Yöntemler - I**

**Genel Ödev**

**İstatistiksel Yöntemler ve Uygulamaları: İki Örneklem Analizi, Uyum  
İyiliği, Gruplar Arası Fark ve Bağımsızlık Kontrolü**

**Bünyamin AKIN – 2210329044**  
**Ünal Giray ERGÜN – 2210329023**

**Ders Sorumlusu – Prof. Dr. Sevil BACANLI**

# İST 268 İSTATİSTİKSEL YÖNTEMLER – I

## GENEL ÖDEV

### Giriş

Bu çalışmanın amacı, İST 268 İstatistiksel Yöntemler - I dersi kapsamında, iki örneklem hipotez testleri, güven aralıkları ve , ki-kare çözümlemeleri uygulamalı olarak incelemektir. Çalışmada iki veri kümesi kullanılarak, kitle ortalaması, kitle varyansı ve belirli bir özelliğe sahip bireylerin kitledeki oranı için hipotez testleri yapılacak, betimleyici istatistikler hesaplanacak ve yorumlanacaktır. Ayrıca, ki-kare çözümlemesi ile uyum iyiliği testi, gruplar arası fark ve bağımsızlık kontrolü gerçekleştirilecektir. Analizlerde SPSS veya MINITAB yazılımları kullanılarak sonuçlar elde edilecektir. Elde edilen veriler ve sonuçlar, istatistiksel yöntemlerin pratik uygulamalarla nasıl kullanılabileceğini göstermeyi amaçlamaktadır.

### 1.İki örneklem

İki örneklem hipotez testlerini, güven aralıklarını ve verilerin basit istatistiklerini elde ederek yorumlayınız (Kitle ortalaması, kitle varyansı ve belli bir özelliğe sahiplerin kitledeki oranının her biri için bir örnek).

➤ Verilerin nasıl elde edildiğini anlatınız. Mutlaka kaynak belirtiniz:

Veriler, DASL (The Data And Story Library) sitesinden elde edilmiştir. DASL, çok çeşitli konulardan gerçek verileri arşivleyen bir veri kütüphanesidir.

**Kitle ortalaması için kaynakça:** [https://dasl.datadescription.com/datafile/grocery-prices/?\\_sfm\\_methods=Comparing+Two+Groups&\\_sfm\\_cases=4+30](https://dasl.datadescription.com/datafile/grocery-prices/?_sfm_methods=Comparing+Two+Groups&_sfm_cases=4+30)

**Kitle varyansı için kaynakça:** [https://dasl.datadescription.com/datafile/cloud-seeding/?\\_sfm\\_methods=Comparing+Two+Groups&\\_sfm\\_cases=4+30](https://dasl.datadescription.com/datafile/cloud-seeding/?_sfm_methods=Comparing+Two+Groups&_sfm_cases=4+30)

**Kitledeki oran için kaynakça\*:** <https://online.stat.psu.edu/stat415/lesson/9/9.4>

\*Oran analizleri için kullanılan veri ise, PennState Eberly College of Science'ın STAT 415 dersindeki "9.4 - Comparing Two Proportions" başlıklı bölümden alınmıştır.

## Betimsleyici bilgiler

Kitle ortalaması için:

### Descriptives

Market			Statistic	Std. Error
Fiyat	WinCo	Mean	2.3021	.45592
		95% Confidence Interval for Mean	Lower Bound	1.3682
			Upper Bound	3.2360
		5% Trimmed Mean	1.9428	
		Median	1.5800	
		Variance	6.028	
		Std. Deviation	2.45521	
		Minimum	.42	
		Maximum	12.68	
		Range	12.26	
		Interquartile Range	1.58	
		Skewness	3.061	.434
		Kurtosis	11.312	.845
	Walmart	Mean	2.7444	.48458
		95% Confidence Interval for Mean	Lower Bound	1.7518
			Upper Bound	3.7370
		5% Trimmed Mean	2.3930	
		Median	1.9800	
		Variance	6.810	
		Std. Deviation	2.60955	
		Minimum	.56	
		Maximum	12.84	
		Range	12.28	
		Interquartile Range	1.58	
		Skewness	2.573	.434
		Kurtosis	7.676	.845

Genel yorum, WinCo ve Walmart marketlerinden alınan fiyatların betimsleyici istatistiklerinin SPSS çıktısı yukarıda verilmiştir.

WinCo marketindeki fiyatların ortalama 2,3 dolar civarındadır. Veri setindeki standart sapma 2,45'tir. Bu, verilerin ortalama fiyat etrafında 2,45\$ kadar yayıldığını gösterir. Veri setinde minimum fiyat 0.42 dolar ve maksimum fiyat ise 12.68 dolar olarak bulunmuştur.

Walmart marketindeki fiyatların ortalama 2.74 dolar civarındadır. Veri setindeki standart sapma 2.61'dir. Bu, verilerin ortalama fiyat etrafında 2.61\$ kadar yayıldığını gösterir. Veri setinde minimum fiyat 0.56 dolar ve maksimum fiyat ise 12.84 dolar olarak bulunmuştur.

### Kitle varyansı için:

#### Descriptives

Tohum			Statistic	Std. Error
YagisMiktari	Tohumlanmis	Mean	441.985	127.6299
		95% Confidence Interval for Mean	Lower Bound Upper Bound	179.126 704.843
		5% Trimmed Mean	351.720	
		Median	221.600	
		Variance	423523.942	
		Std. Deviation	650.7872	
		Minimum	4.1	
		Maximum	2745.6	
		Range	2741.5	
		Interquartile Range	365.3	
		Skewness	2.435	.456
		Kurtosis	6.008	.887
	Tohumlanmamis	Mean	164.588	54.6039
		95% Confidence Interval for Mean	Lower Bound Upper Bound	52.130 277.047
		5% Trimmed Mean	120.735	
		Median	44.200	
		Variance	77521.263	
		Std. Deviation	278.4264	
		Minimum	1.0	
		Maximum	1202.6	
		Range	1201.6	
		Interquartile Range	159.6	
		Skewness	2.789	.456
		Kurtosis	8.173	.887

Genel yorum, gümüş iyodür ile tohumlanmış ve tohumlanmamış bulutların betimleyici istatistiklerinin SPSS çıktısı yukarıda verilmiştir.

Gümüş iyodür ile tohumlanmış bulutların yağmur miktarı (dönüm – feet cinsinden) ortalama 441.98 dönüm-foot civarındadır. Veri setindeki standart sapma 650.78'tir. Bu, verilerin ortalama yağmur miktarında 650.78 birim yayıldığını gösterir. Veri setinde minimum yağmur miktarı 4.1 dönüm-foot, maksimum yağmur miktarı ise 2745.6 dönüm-foot olarak bulunmuştur.

Gümüş iyodür ile tohumlanmamış bulutların yağmur miktarı (dönüm – feet cinsinden) ortalama 164.58 dönüm-foot civarındadır. Veri setindeki standart sapma 278.42'dir. Bu, verilerin ortalama yağmur miktarında 278.42 birim yayıldığını gösterir. Veri setinde minimum yağmur miktarı 1 dönüm-foot, maksimum yağmur miktarı ise 1202.6 dönüm-foot olarak bulunmuştur.

### Kitledeki oran için:

Genel yorum, 800 yetişkin Amerikalının 392 tanesi sağlık reformu için kullanılmak üzere sigaradan alınan verginin artırılması gerektiğini belirtmiştir. 408 tanesi ise bu fikre karşı çıkmıştır. Sigara içen kişilerin ortalama %21'i, sigara içmeyen kişilerin %58'i vergilerin artırılması gerektiğini düşünmüştür.

## Kitle ortalaması:

Verilerimiz, WinCo ve Walmart market zincirlerinin ortalama ürün fiyatlarının karşılaştırılmasıdır.

$\mu_1$  = WinCo marketinin kitle ortalaması

$\mu_2$  = Walmart marketinin kitle ortalaması

**İDDİA:** Market zincirleri arasında fiyat farkı yoktur.

Kitlenin normal dağılım gösterdiği varsayımı altında ve kitle varyansının bilinmediği durumda, örneklem büyüklüğü 30'dan küçük olduğunda ( $n < 30$ ), %95 güven düzeyi ile t-testi kullanılır.

Öncelikle varyansların homojenliği test edilmelidir. Bunun için kurulacak hipotez,

$$H_0 = \sigma_1^2 = \sigma_2^2$$

$$H_1 = \sigma_1^2 \neq \sigma_2^2$$

biçimindedir.  $\alpha = 0.05$  alındığında, Levene's Test istatistiği sonucu p-değeri 0.771 elde edilir ve  $H_0$  reddedilemez. Bu durumda, marketlerin fiyat varyanslarının homojen olduğu %95 güven düzeyinde söylenebilir.

Bu durumda SPSS çıktısında "Equal variances assumed" satırını göz önünde bulunduracağız.

Ortalama fiyatlar ilişkin iddia için kurulan hipotezler,

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

biçimindedir.  $\alpha = 0.05$

Hipotezimizi test etmek için,

### 1.Yol: Güven aralığı ile,

Güven aralığı  $-1.77519 < x < 0.89050$  sıfırı kapsadığı için kitle ortalamaları arasında fark yoktur.

$H_0$  hipotezi, 0.05 anlamlılık düzeyinde reddedilemez.

### 2.Yol: P-değeri ile,

P-değeri .509 elde edilmiştir  $p = 0.509 > \alpha = 0.05$  olduğu için  $H_0$  hipotezi, 0.05 anlamlılık düzeyinde reddedilemez.

### 3.Yol: t istatistiği ile,

t-hesap değeri -0.665 dir. Serbestlik derecesi 56'dır. Hipotezimiz iki yönlüdür bu yüzden  $\alpha/2$  kullanılacaktır.

t-tablo değeri +/- 2.003241'dir.

$-2 < -0.665 < +2$  olduğundan  $H_0$ , 0.05 anlamlılık düzeyinde reddedilemez.

**YORUM:** Kitle ortalamaları arasında %95 güven düzeyi ile fark olmadığını söyleyebiliriz.

Group Statistics

Market	N	Mean	Std. Deviation	Std. Error Mean
Fiyat WinCo	29	2.3021	2.45521	.45592
Walmart	29	2.7444	2.60955	.48458

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% Confidence Interval of the Difference	
									Lower	Upper
Fiyat	Equal variances assumed	.085	.771	-.665	56	.509	-.44234	.66534	-1.77519	.89050
	Equal variances not assumed			-.665	55.793	.509	-.44234	.66534	-1.77530	.89061

### Kitle varyansı:

Verilerimiz, gümüş iyodür ile bulutlara tohumlama sonrası yağış miktarını artırıp artırmadığını belirlemek için bir deneyde, 52 bulut rastgele seçildi . Daha sonra bulutların yağış miktarı ölçüldü.

SPSS ile basit istatistikleri elde ettik.

$\sigma_1^2$  = Tohumlanmış bulutların kitle varyansı

$\sigma_2^2$  = Tohumlanmamış bulutların kitle varyansı

**İDDİA:** Tohumlanmış ve tohumlanmamış bulutların ürettiği yağış miktarlarının varyansları arasında fark yoktur.

### Tohum

Case Processing Summary

		Cases					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
YagisMiktari	Tohumlanmış	26	100.0%	0	0.0%	26	100.0%
	Tohumlanmamış	26	100.0%	0	0.0%	26	100.0%

Descriptives

Tohum			Statistic	Std. Error
YagisMiktari	Tohumlanmış	Mean	441.985	127.6299
		95% Confidence Interval for Mean	Lower Bound Upper Bound	179.126 704.843
		5% Trimmed Mean	351.720	
		Median	221.600	
		Variance	423523.942	
		Std. Deviation	650.7872	
		Minimum	4.1	
		Maximum	2745.6	
		Range	2741.5	
		Interquartile Range	365.3	
		Skewness	2.435	.456
		Kurtosis	6.008	.887
	Tohumlanmamış	Mean	164.588	54.6039
		95% Confidence Interval for Mean	Lower Bound Upper Bound	52.130 277.047
		5% Trimmed Mean	120.735	
		Median	44.200	
		Variance	77521.263	
		Std. Deviation	278.4264	
		Minimum	1.0	
		Maximum	1202.6	
		Range	1201.6	
		Interquartile Range	159.6	
		Skewness	2.789	.456
		Kurtosis	8.173	.887

Normal dağılım varsayımı altında ve örneklem büyüklüğü 30'dan küçük olduğunda ( $n < 30$ ), %95 güven düzeyi ile F-testi kullanılır.

Öncelikle varyansların homojenliği denetlenmelidir. Bunun için kurulacak hipotez,

$$H_0 = \sigma_1^2 = \sigma_2^2$$

$$H_1 = \sigma_1^2 \neq \sigma_2^2$$

biçimindedir.

Hipotezi test etmek için Minitab programından devam edebiliriz.

Tohumlanmış bulutların yağış miktarının standart sapması = 650.7872

Tohumlanmamış bulutların yağış miktarının standart sapması = 278.4264

## Test and CI for Two Variances

### Method

$\sigma_1$ : standard deviation of Sample 1

$\sigma_2$ : standard deviation of Sample 2

Ratio:  $\sigma_1/\sigma_2$

F method was used. This method is accurate for normal data only.

### Descriptive Statistics

Sample	N	StDev	Variance	95% CI for $\sigma$
Sample 1	26	650.787	423523.980	(510.385, 898.352)
Sample 2	26	278.426	77521.260	(218.358, 384.342)

### Ratio of Standard Deviations

Estimated Ratio	95% CI for Ratio using F
2.33738	(1.565, 3.491)

### Test

Null hypothesis  $H_0: \sigma_1 / \sigma_2 = 1$

Alternative hypothesis  $H_1: \sigma_1 / \sigma_2 \neq 1$

Significance level  $\alpha = 0.05$

Test				
Method	Statistic	DF1	DF2	P-Value
F	5.46	25	25	0.000

Two-Sample Variance

Sample standard deviations

Sample size: Sample 1: 26, Sample 2: 26

Standard deviation: Sample 1: 650.7872, Sample 2: 278.4264

Two-Sample Variance: Options

Ratio: (sample 1 standard deviation) / (sample 2 standard deviation)

Confidence level: 95.0

Hypothesized ratio: 1

Alternative hypothesis: Ratio  $\neq$  hypothesized ratio

☒ Use test and confidence intervals based on normal distribution

Hipotezimizi test etmek için,

**1.Yol: Güven aralığı ile,**

Güven aralığı  $1.565 < x < 3.491$  biri(1) kapsamıyor için kitle varyansları arasında fark vardır.

$H_0$  hipotezi, 0.05 anlamlılık düzeyinde reddedilir.

**2.Yol: P-değeri ile,**

P-değeri 0.000 elde edilmiştir  $p=0.0.000 < \alpha = 0.05$  olduğu için  $H_0$ , 0.05 anlamlılık düzeyinde reddedilir. Kitle varyansları arasında fark vardır.

**3.Yol: F istatistiği ile,**

F-hesap değeri 5.46'dır. Serbestlik dereceleri 25,25'tir.  $\alpha = 0.05$

T-tablo değeri 1.96'dır.

$5.46 > 1.96$  olduğundan  $H_0$  hipotezi, 0.05 anlamlılık düzeyinde reddedilir.

**YORUM:** Kitle varyansları fark olduğunu %95 güven düzeyi ile söyleyebiliriz.



### Kitle oranı:

Verilerimiz, Time dergisinin 800 yetişkin Amerikalıya sorduğu telefon anketinde, “Sağlık reformu için kullanılmak üzere sigaradan alınan vergi artırılmalı mıdır?” sorusuna verilen cevaplar:

605 sigara içmeyen vatandaşın 351’i evet demiştir.  $\hat{p}_1 = 351/605 = 0.58$

195 sigara içen vatandaşın 41’i evet demiştir.  $\hat{p}_2 = 41/195 = 0.21$

**İDDİA:** Sigara içenlerin/içmeyenlerin verdiği cevapların oranında fark var mıdır?

Normal dağılım varsayımı %95 güven düzeyi ile Z-testi kullanılır.

Hipotezlerimizi kuralım.

Sağlık reformu için sigaranın vergilendirmesinin destekleyenlerin/desteklemeyenlerin kitle oranlarının arasında fark olup/olmadığına ilişkin iddia için kurulan hipotezler,

$$H_0 = p_1 = p_2$$

$$H_1 = p_1 \neq p_2$$

biçimindedir.  $\alpha = 0.05$

## Test and CI for Two Proportions

### Method

$p_1$ : proportion where Sample 1 = Event

$p_2$ : proportion where Sample 2 = Event

Difference:  $p_1 - p_2$

### Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	605	351	0.580165
Sample 2	195	41	0.210256

### Estimation for Difference

Difference	95% CI for Difference
0.369909	(0.300499, 0.439319)

CI based on normal approximation

### Test

Null hypothesis  $H_0: p_1 - p_2 = 0$

Alternative hypothesis  $H_1: p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Normal approximation	8.99	0.000
Fisher's exact		0.000

The pooled estimate of the proportion (0.49) is used for the tests.

Hipotezimizi test etmek için,

**1.Yol: Güven aralığı ile,**

Güven aralığı  $0.3 < x < 0.43$  sıfırı kapsamıyor için kitle oranları arasında fark vardır.

$H_0$  hipotezi, 0.05 anlamlılık düzeyinde reddedilir.

**2.Yol: P-değeri ile,**

P-değeri 0.000 elde edilmiştir  $p=0.000 < \alpha = 0.05$  olduğu için  $H_0$ , 0.05 anlamlılık düzeyinde reddedilir. Kitle oranları arasında fark vardır.

**3.Yol: Z istatistiği ile,**

Z-hesap değeri 8.99 dur.  $\alpha = 0.05$  Hipotez iki yönlüdür,  $\alpha/2 = 0.025$ .

Z-tablo değeri 1.96'dır.

$8.99 > 1.96$  olduğundan  $H_0$ , 0.05 anlamlılık düzeyinde reddedilir.

**YORUM:** Kitle oranları arasında fark olduğu %95 güven düzeyi ile söylenebilir.

## **2. Uyum iyiliği, gruplar arası fark kontrolü ve bağımsızlık kontrolü**

Ki-kare çözümlemesinde verilen; uyum iyiliği testi, gruplar arası fark ve bağımsızlık kontrolüne ilişkin çözümleri yapınız (Her biri için bir örnek).

➤ Verilerin nasıl elde edildiğini anlatınız. Mutlaka kaynak belirtiniz:

Veriler, DASL (The Data And Story Library) sitesinden elde edilmiştir. DASL, çok çeşitli konulardan gerçek verileri arşivleyen bir veri kütüphanesidir.

**Uyum iyiliği için kaynakça:**

[https://dasl.datadescription.com/datafile/mens-weights/?\\_sfm\\_methods=Normal+model&\\_sfm\\_cases=4+59943](https://dasl.datadescription.com/datafile/mens-weights/?_sfm_methods=Normal+model&_sfm_cases=4+59943)

**Gruplar arası fark için kaynakça:**

[https://dasl.datadescription.com/datafile/covid\\_blood\\_type\\_and\\_age/?\\_sfm\\_cases=4+10](https://dasl.datadescription.com/datafile/covid_blood_type_and_age/?_sfm_cases=4+10)

**Bağımsızlık/ilişki kontrolü için kaynakça:**

[https://dasl.datadescription.com/datafile/pregnancies/?\\_sfm\\_methods=Chi+Square+Tests&\\_sfm\\_cases=4+12](https://dasl.datadescription.com/datafile/pregnancies/?_sfm_methods=Chi+Square+Tests&_sfm_cases=4+12)

## Uyum iyiliği :

Verilerimiz, Ulusal Sağlık ve Beslenme İnceleme Araştırması (NHANES), Amerika Birleşik Devletleri'ndeki yetişkinlerin ve çocukların sağlık ve beslenme durumlarını değerlendirmek ve zaman içindeki değişiklikleri izlemek amacıyla Ulusal Sağlık İstatistikleri Merkezi tarafından yürütülen bir anket araştırma programıdır. Bu anketin sonuçlarından çekilen 19 ve 24 yaş aralığındaki ortalama boya sahip 80 erkeğin kilolarıdır.

Hipotezler kurulur,

$H_0$  = Erkeklerin kilolarının dağılımı Normal dağılmaktadır.

$H_1$  = Erkeklerin kilolarının dağılımı Normal dağılmamaktadır.

SPSS programı ile normallik testi sonuçları aşağıda verilmiştir.

### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Kilo	80	100.0%	0	0.0%	80	100.0%

### Descriptives

			Statistic	Std. Error
Kilo	Mean		82.356	2.4897
	95% Confidence Interval for Mean	Lower Bound	77.401	
		Upper Bound	87.312	
	5% Trimmed Mean		80.697	
	Median		76.850	
	Variance		495.900	
	Std. Deviation		22.2688	
	Minimum		54.3	
	Maximum		161.5	
	Range		107.2	
	Interquartile Range		25.7	
	Skewness		1.318	.269
	Kurtosis		1.472	.532

### Tests of Normality

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Kilo	.174	80	.000	.876	80	.000

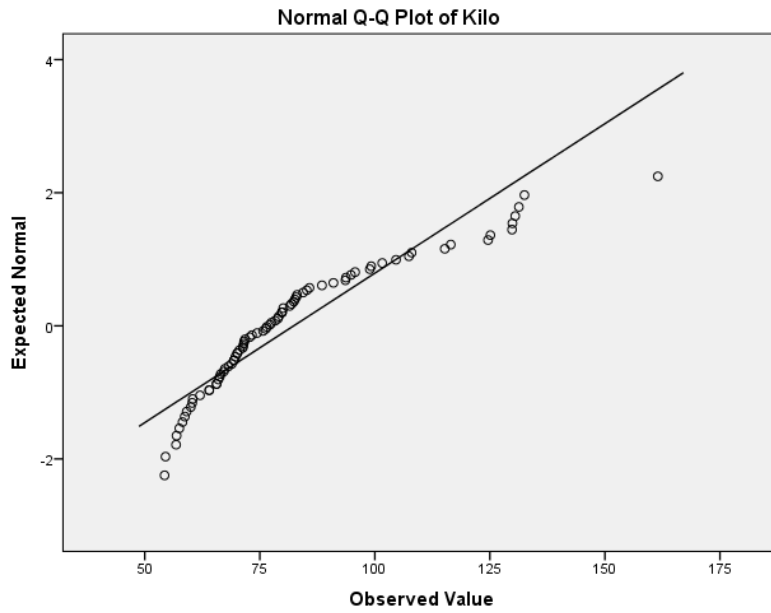
a. Lilliefors Significance Correction

Test sonucunda, örneklem büyüklüğümüz 50'den büyük olduğu için Kolmogorov-Smirnov testine bakıyoruz. Burada p-değerimiz  $0.000 < \alpha = 0.05$  olduğu için  $H_0$ , 0.05 anlamlılık düzeyinde reddedilir.

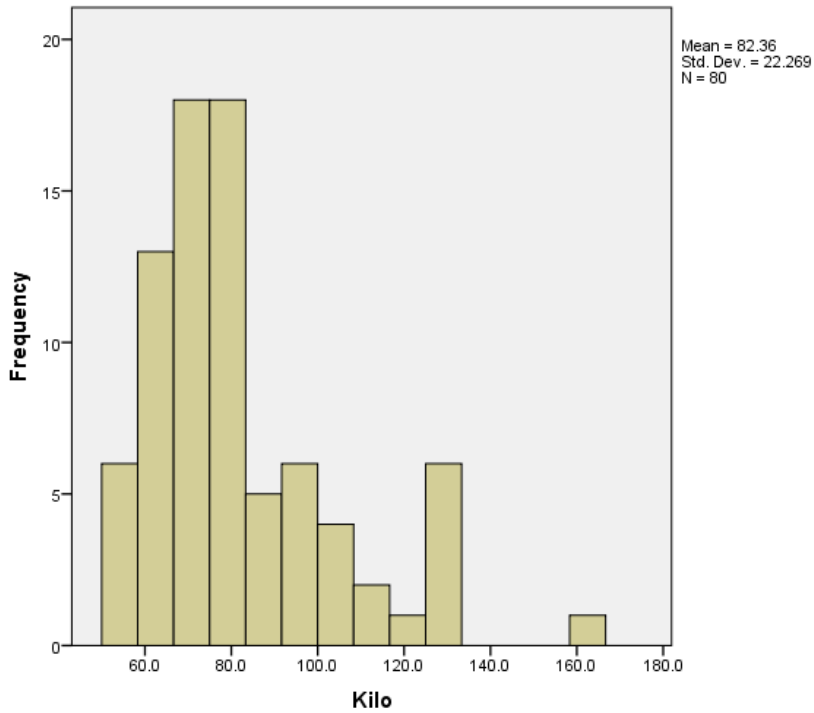
**YORUM:** Erkeklerin kilolarının dağılımı Normal dağılıma sahip olmadığı %95 güven düzeyi ile söylenebilir.

2.Yol olarak, Q-Q grafiğini çizdirebiliriz, Verilerimiz 45 derecelik doğru üzerinde değildir. Normal dağılıma sahip olmadığını buradan da yorumlayabiliriz.

#### Kilo



#### → Graph



## Gruplar Arası Fark Kontrolü:

Verilerimiz, COVID-19 hastalarının kan grupları ve yaşlarını göstermektedir.

**İDDİA:** COVID-19 hastaları yaşlarına ve kan gruplarına göre sınıflandırıldığında, kan grupları bakımından yaşlar arasında fark olup olmadığını %95 güven düzeyi ile test edelim.

Hipotezi kuralım,

$H_0$  = COVID-19 hastalarının kan grupları bakımından yaşlar arasında fark yoktur.

$H_1$  = COVID-19 hastalarının kan grupları bakımından yaşlar arasında fark vardır.

### → Crosstabs

[DataSet0]

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Yas * KanGrubu	1888	100.0%	0	0.0%	1888	100.0%

#### Yas \* KanGrubu Crosstabulation

			KanGrubu				Total
			A	B	AB	O	
Yas	Genc	Count	100	78	21	74	273
		Expected Count	103.4	71.4	27.9	70.3	273.0
	OrtaYas	Count	275	176	71	183	705
		Expected Count	267.0	184.5	72.1	181.5	705.0
	Yasli	Count	340	240	101	229	910
		Expected Count	344.6	238.1	93.0	234.2	910.0
	Total	Count	715	494	193	486	1888
		Expected Count	715.0	494.0	193.0	486.0	1888.0

#### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.158 <sup>a</sup>	6	.655
Likelihood Ratio	4.286	6	.638
Linear-by-Linear Association	.002	1	.961
N of Valid Cases	1888		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 27.91.

Beklenen sıklıklarda 5'ten küçük bir değer yok. Ki-kare Testi ile devam edebiliriz.

p-değeri = .655 >  $\alpha$  = 0.05 olduğu için  $H_0$ , 0.05 anlamlılık düzeyinde reddedilemez.

**YORUM:** COVID-19 hastalarının kan grupları bakımından yaşlar arasında fark olmadığını %95 güven düzeyi ile söyleyebiliriz.

## İlişki kontrolü :

Verilerimiz, annelerin erken veya geç doğum yapmalarını ve yaşlarını göstermektedir.

**İDDİA:** Annelerin yaşları ile erken/geç doğum yapmaları arasında bir ilişki olup olmadığını 0,05 yanılma düzeyi ile test edelim.

$H_0$  = Annelerin erken/geç doğum yapma sıklığı ile bulunduğu yaş arasında bir ilişki yoktur.

$H_1$  = Annelerin erken/geç doğum yapma sıklığı ile bulunduğu yaş arasında bir ilişki vardır.

### → Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Yas * Dogum	1900	100.0%	0	0.0%	1900	100.0%

Yas \* Dogum Crosstabulation

			Dogum		Total
			Erken Dogum	Gec Dogum	
Yas	0-20	Count	129	270	399
		Expected Count	116.6	282.5	399.0
	20-29	Count	243	612	855
		Expected Count	249.8	605.3	855.0
	30-39	Count	165	424	589
		Expected Count	172.1	417.0	589.0
	40+	Count	18	39	57
		Expected Count	16.7	40.4	57.0
Total	Count	555	1345	1900	
	Expected Count	555.0	1345.0	1900.0	

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	2.699 <sup>a</sup>	3	.440
Likelihood Ratio	2.665	3	.446
Linear-by-Linear Association	1.169	1	.280
N of Valid Cases	1900		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 16.65.

Symmetric Measures

		Value	Asymptotic Standardized Error <sup>a</sup>	Approximate T <sup>b</sup>	Approximate Significance
Ordinal by Ordinal	Kendall's tau-b	.024	.022	1.121	.262
	Kendall's tau-c	.025	.023	1.121	.262
	Gamma	.047	.042	1.121	.262
N of Valid Cases		1900			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Beklenen sıklıklarda 5'ten küçük bir değer yok. Ki-kare Testi ile devam edebiliriz.

p-değeri = 0.440 >  $\alpha$  = 0.05 olduğu için  $H_0$ , 0.05 anlamlılık düzeyinde reddedilemez. İlişki olmadığı için ilişki katsayılarına bakmıyoruz.

**YORUM:** Annelerin erken/geç doğum yapma sıklığı ile bulunduğu yaş arasında bir ilişki olmadığını %95 güven düzeyi ile söyleyebiliriz.