

HACETTEPE ÜNİVERSİTESİ

**BARIŞ KARATAŞ**

**22003290..**

**İSTATİSTİK BÖLÜMÜ**

**1. SINIF**

PROF.DR. SERPİL AKTAŞ ALTUNAY

ARŞ.GÖR. LEYLA BAKACAK KARABENLİ

- ❖ Ödev için ECUADOR ve URUGUAY ülkelerinden, ortak karar olarak ECUADOR ülkesini seçip 2 farklı değişken belirledik. Değişkenlerimizi EXCEL dosyasına aktardık.

</

➤ Daha sonrasında X ve Y değişkenlerimizi belirledik:

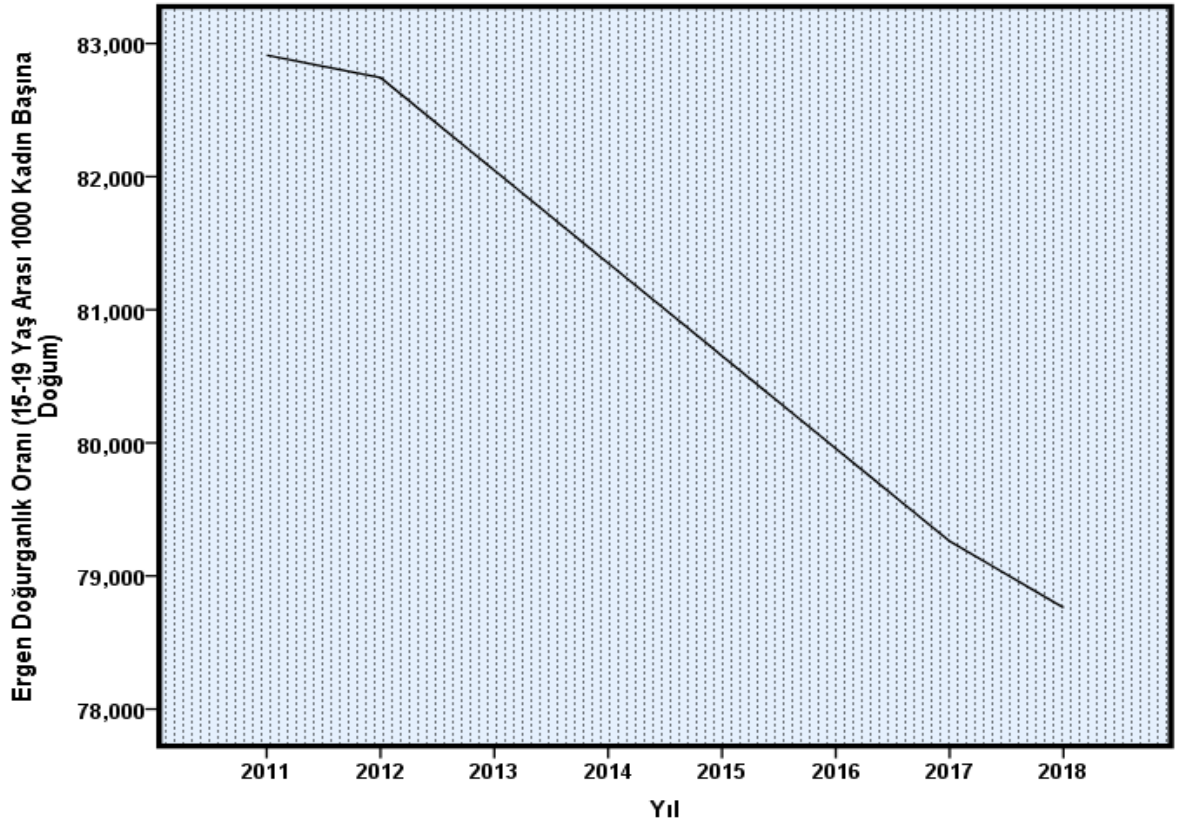
**X değişkeni:** Ergen doğurganlık oranı (15-19 yaş arası 1000 kadın başına doğum sayısı)

**Y değişkeni:** Nüfus yoğunluğu (km kare arazi alanı başına düşen kişi)

X i bağımsız Y i bağımlı değişkenler olarak düşünüp bu seçimi yaptık. Doğurganlık oranı kendi içinde bağımsız olarak değişkenlik gösteriyorken, Nüfus yoğunluğu bu doğurganlık oranına göre değişecektir, bundan dolayı bağımlı bir değişken olacaktır şeklinde düşündük.

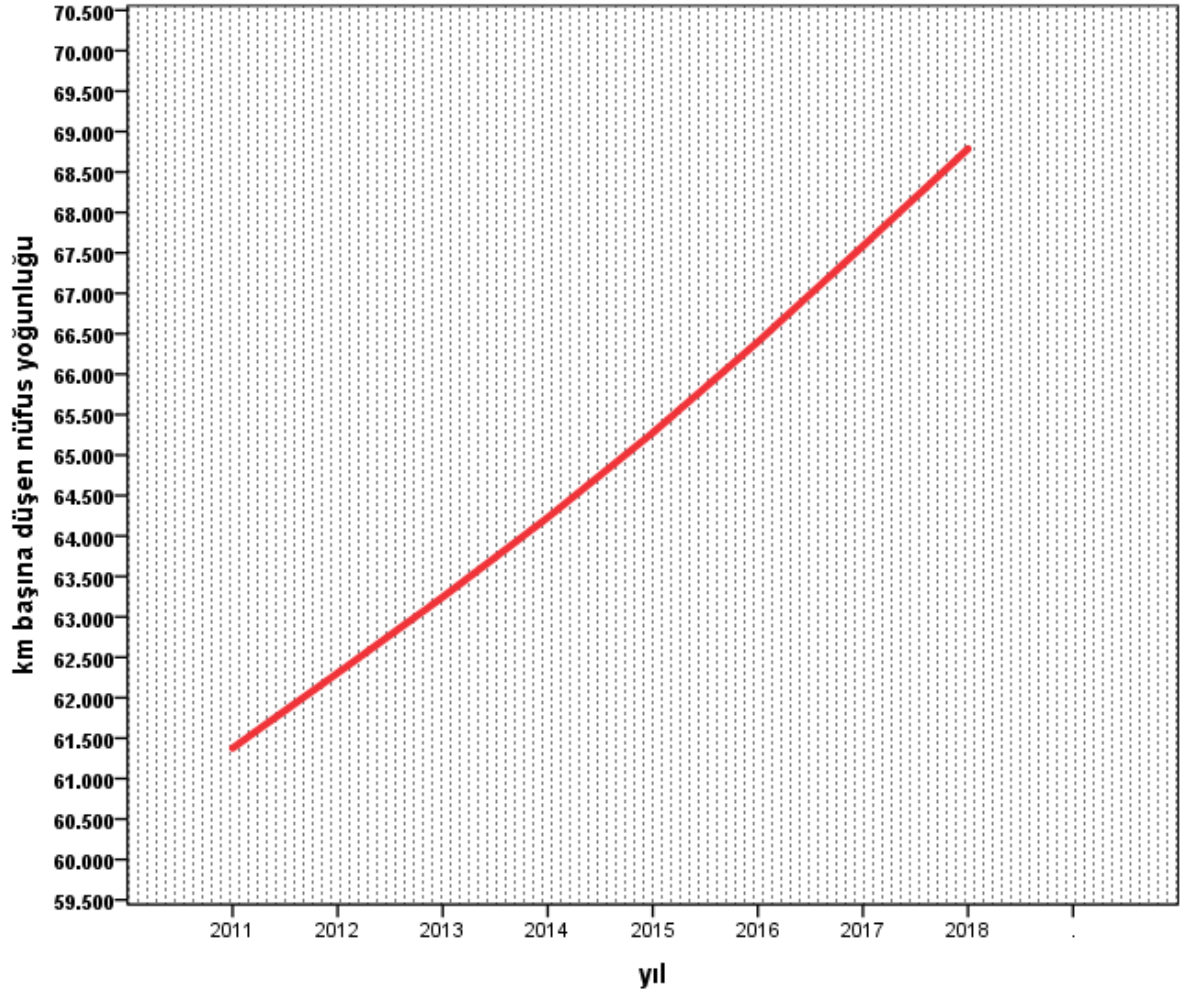
**NOT:** Verilerin daha doğru ve dengeli bir sonuç vermesi için işlemlerimizde [2011,2018] yılları arasını baz aldık ve bu verileri kullandık.

**SORU 1:** X ve Y deęiřkeni iin uygun bir veri grselleřtirme (grafik) oluřturunuz.



- 2011 ve 2018 yılları arasında Ecuador lkesine ait Ergen Doęurganlık Oranı (15-19 Yař Arası 1000 Kadın Bařına Doęum Sayısı) trend grafięi ile grselleřtirildi.

**YORUM:** Grafik incelendięinde ergen doęurganlık oranının negatif ynl ilerledięi grlmektedir. Bu yn kimi noktalarda daha dik bir eęime sahiptir, bu da bizlere dřřn daha sert olduęunu gstermektedir. Sonu olarak 2011-2018 yılları arasında geen her yılda ergen doęurganlık oranının azaldıęını syleyebiliriz.



2011 ve 2018 yılları arasında Ekvator ülkesinin km başına düşen insan sayısı

**YORUM:** Burada görüldüğü üzere, trend çizgimiz pozitif yönlü ilerlemekte.  
Bununda anlamı: Ekvator ülkesinde 2011-2018 Yılları arasında geçen her yıl,  $\text{km}^2$  başına düşen kişi sayısı artmıştır diyebiliriz.

**SORU 2:** X değişkenini kullanarak kitle ortalaması için %95 güven aralığını hesaplayınız ve yorumlayınız.

T-TEST

/TESTVAL=0

/MISSING=ANALYSIS

/VARIABLES=Ergen\_Dogurganlik\_Orani

/CRITERIA=CI (.95) .

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Ergen_Dogurganlik_Orani	8	80,96087	1,564977	,553303

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Ergen_Dogurganlik_Orani	146,323	7	,000	80,960875	79,65252	82,26923

Kitle varyansı ( $\sigma^2$ ) bilinmediği ve  $n < 30$  olduğu için, t dağılımından yararlandık. X değişkenimizi Ergen Doğurganlık Oranı (15-19 Yaş Arası 1000 Kadın Başına Doğum) olarak belirlemiştik.

**YORUM:** Hesaplanan güven aralığının kitle ortalamasını kapsama olasılığı %95'tir. 15-19 Yaş Arası 1000 Kadın Başına Ergen Doğurganlık Oranının [2011-2018] yılları aralığında ortalama değerlerinin [79.65252, 82.26923] arasında olduğunu %95 güven düzeyinde söyleyebiliriz.

3.SORU: X değişkenini kullanarak kitle varyansı için %90 güven aralığını hesaplayınız ve yorumlayınız.

```
1 ##KİTLE VARYANSI İÇİN GÜVEN ARALIĞI ELDE ETME##
2 attach(odev_2_donem)
3 n=8
4 alpha=0.10
5 ##Kikare tablo değerleri
6 tablo=qchisq(c(1-alpha/2,alpha/2),n-1)
7 tablo
8 ##Güven aralığı hesabı
9 aralik=(n-1)*var(ergen_dogurganlik_orani)/tablo
10 aralik
11 ##kısa yol
12 library(DescTools)
13 VarCI(ergen_dogurganlik_orani,conf.level = 0.90)
14 |
```

```
> n=8
> alpha=0.10
> ##Kikare tablo değerleri
> tablo=qchisq(c(1-alpha/2,alpha/2),n-1)
> tablo
[1] 14.06714 2.16735
> ##Güven aralığı hesabı
> aralik=(n-1)*var(ergen_dogurganlik_orani)/tablo
> aralik
[1] 1.218778 7.910454
> ##kısa yol
> library(DescTools)
> VarCI(ergen_dogurganlik_orani,conf.level = 0.90)
      var   lwr.ci   upr.ci
2.449246 1.218778 7.910454
```

alpha	0.1
aralik	num [1:2] 1.22 7.91
n	8
tablo	num [1:2] 14.07 2.17

Bu değerlere bakarak güven aralığının 1.22 ile 7.91 arasında olduğunu görebiliyoruz.

**YORUM:** Buna göre Ekvator ülkesinin 2011-2018 yılları arasında geçen her yılda (15-19) yaş aralığındaki her 1000 ergen kadının doğurganlık oranındaki değişim(varyans) yaklaşık 1 ila 8 kişi arasında olacağını %90 güvenlilikle söyleyebiliriz.

**SORU 4:** Y deęiřkenini kullanarak belirledięiniz bir  $\mu_0$  deęeri iin  $H_0:\mu=\mu_0$  hipotezini  $H_1:\mu\neq\mu_0$  hipotezine karřı %10 anlamlılık dzeyinde test ediniz ve yorumlayınız.

- Y deęiřkenimizi Nfus Yoęunluęu (Km kare arazi alanı bařına dřen kiři) olarak belirlemiřtik.
- Yıllık ortalama nfus yoęunluęu oranının 65 olup olmadıęını incelemek istedik, (bunun iin hipotez testi yapmamız gerekli) dolayısıyla  $\mu_0$  deęerimizi  $\mu_0=65$  olacak řekilde belirlemiř olduk.

Hipotez testinde de kitle varyansı bilinmedięi ve rneklem hacmi 30'dan kk olduęu iin t testi kullanılacak ve

**$H_0: \mu=65$**

**$H_1: \mu\neq 65$**

hipotezi test edilecektir.

- Kitle varyansı ( $\sigma^2$ ) bilinmedięi ve  $n<30$  olduęu iin t testi kullanılır.

#### T-TEST

```
/TESTVAL=65  
/MISSING=ANALYSIS  
/VARIABLES=Nufus_Yogunlugu  
/CRITERIA=CI(.90).
```

#### T-Test

##### One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Nufus_Yogunlugu	8	64,90038	2,592412	,916556

##### One-Sample Test

	Test Value = 65					
	t	df	Sig. (2-tailed)	Mean Difference	90% Confidence Interval of the Difference	
					Lower	Upper
Nufus_Yogunlugu	-,109	7	,916	-,099625	-1,83611	1,63686

- Hipotez testinin sonucuna ulaşabilmek için t test istatistiği, p değeri veya güven aralığı kullanılabilir.

**T test istatistiği:** Hipotez iki yönlü olduğu için mutlak t test istatistiğinin,  $\alpha/2$  tablo değerinden büyük olup olmadığı kontrol edilir.

Eğer  $|t| \geq t_{\alpha/2, n-1}$  ise  $H_0$  reddedilir. Tam tersi için ise kabul edilir.

$|t| = |-0.109| < t_{0.05, 7} = 1.895$  olduğu için  $H_0$  kabul edilir.

**P değeri:**  $P = 0.916 > \alpha = 0.100$  olduğu için  $H_0$  kabul edilir.

**Güven aralığı:** Güven aralığının, hipotezde test edilen değeri içerip içermediğini kontrol etmeden önce SPSS çıktısındaki alt ve üst sınıra hipotez test değerinin eklenmesi gerekmektedir.

Güven Aralığı:  $(65 + (-1.83611), 65 + (1.63686)) \gg (63.16389, 66.63686)$

**Bu işlemden sonra güven aralığı 65 değerini içerdiği için  $H_0$  kabul edilir.**

**YORUM:** %10 anlamlılık düzeyinde  $H_0$  kabul edildiği için yıllık ortalama nüfus yoğunluğu (km kare arazi alanı başına düşen kişi) oranının 65 olduğunu söyleyebiliriz (65 olduğuna dair yeterli kanıt mevcuttur).





## 5.SORU: X ve Y değişkenleri arasında Pearson, Spearman ve Kendall'in korelasyon katsayılarından en uygun olanı hesaplayarak yorumlayınız.

İlk olarak korelasyon ilişkilerinden olan Pearson, Spearman, Kendall's Tau arasından birini seçmek için dağılımımızın normal dağılıma sahip mi değil mi kontrol etmemiz gerekiyor.

Bunu da SPSS üzerinden Kolmogorov-smirnov | Shapiro-Wilk testleriyle gerçekleştirdik. Normal dağılım olması için iki testin p değeri  $p > 0,05$  olmalıdır. Buradaki bağımlı değişkenimizin Normal dağılıma sahip olduğunu görebiliyoruz.

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
km_nufus_yogunlugu	.114	8	.200*	.973	8	.919
*. This is a lower bound of the true significance.						
a. Lilliefors Significance Correction						

Değişkenlerimizin normal dağılıma sahip olduğunu karar verdiğimizize göre artık Pearson, Spearman, Kendall's Tau seçmemiz gerekir. Bizim için en uygun olan ilişki katsayısı Pearson'un Çarpım Moment Korelasyon Katsayısıdır.

Correlations			
		ergen_dogur ganlik_orani	km_nufus_yo gunlugu
ergen_dogurganlik_orani	Pearson Correlation	1	-.996**
	Sig. (2-tailed)		.000
	N	8	8
km_nufus_yogunlugu	Pearson Correlation	-.996**	1
	Sig. (2-tailed)	.000	
	N	8	8
**. Correlation is significant at the 0.01 level (2-tailed).			

Buradan bakarak ergen doğurganlık oranı ile  $\text{km}^2$  başına düşen nüfus yoğunluğu arasındaki Pearson Korelasyon katsayısı değeri =  $-.996$  olduğundan;

**YORUM:** Ergen doğurganlık oranı ile  $\text{km}^2$  başına düşen nüfus yoğunluğu arasında %99,6 lık negatif yönlü bir ilişki söz konusudur diyebiliriz. Yani Ergen doğurganlık oranı azaldıkça  $\text{km}^2$  başına düşen nüfus yoğunluğu artar.

### ÖNEM KONTROLÜ:

$H_0: \rho = 0$  ( $H_0$ : Ergen doğurganlık oranı ile  $\text{km}^2$  başına düşen nüfus yoğunluğu arasında ilişki yoktur.)

H1:  $p \neq 0$  (H1: Ergen doğurganlık oranı ile  $\text{km}^2$  başına düşen nüfus yoğunluğu arasında ilişki vardır.)

$p = 0.000 < 0.05$  olduğu için  $H_0$  reddedilir. Ergen doğurganlık ve  $\text{km}^2$  başına düşen nüfus yoğunluğu arasında anlamlı ilişki olduğu %95 güven düzeyinde (%5 anlamlılık düzeyinde) söylenebilir.

6.SORU: Y'yi bağımlı, X'i bağımsız değişken olarak alarak, X ve Y değişkenleri arasında basit doğrusal regresyon denklemini kurunuz ve denklemin anlamlılığını % 5 anlamlılık düzeyinde test ediniz ve yorumlayınız.

İlk olarak Normallik incelemesi yapmamız gerekir. Bağımlı değişken olan  $\text{km}^2$  nüfus yoğunluğu üzerinden yapacağız.

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
km_nufus_yogunlugu	.114	8	.200 <sup>*</sup>	.973	8	.919
*. This is a lower bound of the true significance.						
a. Lilliefors Significance Correction						

$n=8$

$n < 50$  olduğu için Shapiro-Wilk testi kullanılacaktır.

$H_0$ : Boy uzunluklarının dağılımı ile normal dağılım arasında fark yoktur.

$H_1$ : Boy uzunluklarının dağılımı ile normal dağılım arasında fark vardır.

$p=0,919 > \alpha=0,05$  olduğu için  $H_0$  reddedilemez ve verilerin dağılımının % 5 anlamlılık düzeyinde normal dağılıma uygun dağılış gösterdiği söylenebilir.

$\beta_0$  ve  $\beta_1$  yorumları ve anlamlılığının testi:

Coefficients <sup>a</sup>							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
1 (Constant)	198.505	4.754		41.759	.000	186.874	210.137
ergen_dogurganlik_orani	-1.650	.059	-.996	-28.111	.000	-1.794	-1.507

a. Dependent Variable: km\_nufus\_yogunlugu

$\beta_0$ =Kesim noktası (Genel ortalama): Ergen doğurganlık oranının etkisi olmadığında  $\text{km}^2$  başına düşen nüfus yoğunluğu ortalaması = 198.505

$\beta_1$ =Eğim: Ergen doğurganlık oranındaki 1 birimlik değişimin  $\text{km}^2$  başına düşen nüfus yoğunluğuna etkisi ortalaması = -1.650

Dolayısıyla basit doğrusal regresyon denklemimiz şöyledir:  $y_i = 198.505 - 1.650x_i$

Basit doğrusal regresyon denkleminin anlamlılık düzeyi:

**DİP-NOT:** Basit doğrusal regresyon modelinde modelin anlamlılığı ile  $\beta_1$ 'in anlamlılığı aynı şeyi ifade etmektedir.

Coefficients <sup>a</sup>							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
1 (Constant)	198.505	4.754		41.759	.000	186.874	210.137
ergen_dogurganlik_orani	-1.650	.059	-.996	-28.111	.000	-1.794	-1.507

a. Dependent Variable: km\_nufus\_yogunlugu

Buradan bakarak;

$H_0: \beta_1 = 0$  (Ergen doğurganlık oranının,  $\text{km}^2$  başına düşen nüfus yoğunluğu üzerindeki etkisi anlamlı değildir.)

$H_1: \beta_1 \neq 0$  (Ergen doğurganlık oranının,  $\text{km}^2$  başına düşen nüfus yoğunluğu üzerindeki etkisi anlamlıdır.)

p değeri üzerinden test ettim:

$p = .000 < \alpha = 0.05$  olduğu için  $H_0$  reddedilir.

Yani ergen doğurganlık oranının,  $\text{km}^2$  başına düşen nüfus yoğunluğu üzerindeki etkisi anlamlıdır deriz.

# Kaynakça:

- Hacettepe İstatistik Bölümünde Huzem platformundaki istatistiğe giriş I-II klasörlerine aktarılan 1.sınıf ders anlatım PDF lerinden yararlanılmıştır.
- <https://databank.worldbank.org/reports.aspx?source=2&country=ECU>
- Normallik testi;  
➤ <https://www.tezyardimplatformu.com/post/normal-da%C4%9Filim-nedi%CC%87r-normalli%CC%87k-testi%CC%87-nasil-yapilir>

Teşekkür ederiz 😊.