

## Regressziószámítás

Nézzük meg, hogy Európa néhány országában az egy főre jutó GDP hogyan befolyásolja a gépjárművek számát, adjuk meg a [lineáris regresszió](#) egyenletét.

ország		X GDP/fő (EUR)	Y Gépjárművek (db / 10 000 fő)
Ausztria	AT	50 380	5500
Belgium	BE	46 237	5030
Csehország	CZ	23 539	5020
Franciaország	FR	41 897	4790
Görögország	GR	19 570	4790
Hollandia	NL	52 646	4810
Lengyelország	PL	15 601	5710
Magyarország	HU	16 470	3380
Németország	DE	46 473	5550
Svájc	CH	82 484	5390

[Megnézem a kapcsolódó epizódot](#)

Egy [statisztika](#) vizsgára történő készüléskor a tanulók saját tapasztalatain alapuló felmérés szerint a tanulással töltött órák száma és az elért pontszám között az alábbi összefüggéseket lehet kimutatni.

Tanulással töltött órák $x$	Pontszám (max 100) $y$
3	5
4	6
5	8
6	9
9	16
10	20
12	24
16	56
20	81
24	96

Adjuk meg a lineáris, a hatványkitevős, és az [exponenciális regresszió](#) egyenletét, és döntsük el, hogy melyik [regresszió](#) illeszkedik-e jobban.

[Megnézem a kapcsolódó epizódot](#)

Az egy főre jutó GDP és az egymillió lakosra jutó orvosok számának kapcsolatát vizsgáljuk.

ország	X	Y
	GDP/fő (USD)	Egymillió lakosra jutó orvosok száma
Ausztria	50 380	5183
Belgium	46 237	3083
Dánia	59 770	3998
Franciaország	41 897	3158
Norvégia	75 294	4659
Hollandia	52 646	3583
Svédország	51 404	4117
Olaszország	33 159	3990
Németország	46 473	4249
Svájc	82 484	4298

Számoljuk ki az elaszticitást 50 ezer dolláros és 60 ezer dolláros egy főre jutó GDP-nél.

[Megnézem a kapcsolódó epizódot](#)

Az alábbi táblázat néhány ország egy főre jutó GDP-jét és a nők életkorát tartalmazza első házasságkötésük idején. Készítsünk lineáris regressziót, ahol a magyarázó változó az egy főre jutó GDP. Értelmezzük a modell paramétereit, készítsünk varianciánálízis táblázatot, adjuk meg a modell magyarázó erejét!

ország		X	Y
		GPD/fő (EUR)	Nők életkora házasságkötéskor
Ausztria	AT	28 978	26,6
Belgium	BE	30 349	29,8
Csehország	CZ	15 216	28,9
Franciaország	FR	26 656	31,6
Görögország	GR	17 941	26,9
Hollandia	NL	28 669	26,9
Lengyelország	PL	10 135	25,3
Magyarország	HU	13 767	29,7
Németország	DE	28 232	31
Svájc	CH	31 987	29,4

[Megnézem a kapcsolódó epizódot](#)

Hőmérséklet (°C)	Kőolaj hordónkénti ára (USD)	Hétvége van?	Eladott gombócok száma
25	100	Igen	760
28.	96	Nem	746
32	98	Nem	796
12	100	Igen	658
7	102	Igen	466
16	96	Igen	642
24	92	Nem	724
5	94	Igen	412
31	98	Nem	756
27	104	Nem	710
25	108	Igen	678
18	110	Nem	655

Készítsünk lineáris regressziót, majd értelmezzük a modell paramétereit. Számítsuk ki a [korreláció](#)-mátrixát.

[Megnézem a kapcsolódó epizódot](#)

Hőmérséklet (°C)	Átlagos levegőminőség (%)	Front van?	Halálozások száma
8	100	Nem	50
12.	64	Nem	43
16	56	Nem	38
25	38	Nem	36
28	85	Igen	42
30	96	Igen	50
5	120	Nem	56
16	68	Nem	40
26	93	Nem	46
27	104	Nem	52
30	24	Igen	48
8	35	Igen	41

Készítsünk lineáris regressziót, majd értelmezzük a modell paramétereit. Végezzük el a [regresszió](#) becsléseit.

[Megnézem a kapcsolódó epizódot](#)

Egy városban a naponta elhalálozottak száma és különböző meteorológiai hatások közötti összefüggést szeretnénk felderíteni, ezért 12 napon vizsgáljuk a hőmérsékletet, a levegőminőséget, valamint, hogy érkezik-e front.

Hőmérséklet (°C)	Átlagos levegőminőség (%)	Front van?	Halálozások száma
8	100	Nem	50
12.	64	Nem	43
16	56	Nem	38
25	38	Nem	36
28	85	Igen	42
30	96	Igen	50
5	120	Nem	56
16	68	Nem	40
26	93	Nem	46
27	104	Nem	52
30	24	Igen	48
8	35	Igen	41

Készítsünk lineáris regressziót, majd értelmezzük a modell paramétereit. Elemezzük a regressziós modellt hipotézisvizsgálatokkal, készítsünk varianciaanalízis táblázatot.

[Megnézem a kapcsolódó epizódot](#)

Napi középhőmérséklet (°C)	Víz hőmérséklete (°C)	Strand napi forgalma
22	21	765
23	21	1572
18	18	510
25	20	1967
22	21	1142
16	19	576
24	22	986
20	21	1216
24	22	1267
26	24	1686
19	19	981
20	21	1412

Készítsünk lineáris regressziót, majd értelmezzük a modell paramétereit. Vizsgáljuk a multikollinearitást és autokorrelációt.

[Megnézem a kapcsolódó epizódot](#)

Néhány ország középfokú iskolai képzésének egy diákra jutó oktatási ráfordítása illetve az éves egy főre jutó GDP adatai láthatóak az alábbi táblázatban. Adjuk meg a [lineáris regresszió](#) modellt, a reziduális szórást, határozzuk meg a modell magyarázó erejét.

		X	Y
ország		GDP/fő (EUR)	Oktatási ráfordítás (középfokú képzés; diák/EUR)
Ausztria	AT	28 978	76 900
Belgium	BE	30 349	61 000
Csehország	CZ	15 216	33 800
Franciaország	FR	26 656	57 600
Görögország	GR	17 941	59 200
Hollandia	NL	28 669	61 500
Lengyelország	PL	10 135	30 700
Magyarország	HU	13 767	33 000
Németország	DE	28 232	65 300
Svájc	CH	31 987	60 400

[Megnézem a kapcsolódó epizódot](#)

Egy strand forgalmának alakulása a napi középhőmérséklettől függően 12 megfigyelt nap alapján az alábbi volt:

nap	napi középhőmérséklet (C°)	forgalom (fő)
1.	22	765
2.	23	1572
3.	18	510
4.	25	1967
5.	22	1142
6.	16	576
7.	24	986
8.	20	1216
9.	24	1267
10.	26	1686
11.	19	981
12.	20	1412

Adjuk meg a [lineáris regresszió](#) egyenletét, adjuk meg a korrelációs és a determinációs együtthatót és döntsük el, hogy a lineáris vagy a [hatványkitevős regresszió](#) illeszkedik-e jobban, ha ismeretes, hogy

$$\sum d^2x = 100,91 \quad \sum d^2y = 2155847 \quad \sum dx \cdot dy = 10894,67 \quad \hat{y} = 1,43 \cdot x^{2,17}$$

[Megnézem a kapcsolódó epizódot](#)

Az alábbi táblázat néhány ország egy főre jutó GDP-jét és a nők életkorát tartalmazza első házasságkötésük idején. Készítsünk lineáris regressziót, ahol a magyarázó változó az egy főre jutó GDP. Értelmezzük a modell paramétereit, készítsünk varianciánálízis táblázatot, adjuk meg a modell magyarázó erejét!

ország		X	Y
		GPD/fő (EUR)	Nők életkora házasságkötéskor
Ausztria	AT	28 978	26,6
Belgium	BE	30 349	29,8
Csehország	CZ	15 216	28,9
Franciaország	FR	26 656	31,6
Görögország	GR	17 941	26,9
Hollandia	NL	28 669	26,9
Lengyelország	PL	10 135	25,3
Magyarország	HU	13 767	29,7
Németország	DE	28 232	31
Svájc	CH	31 987	29,4

[Megnézem a kapcsolódó epizódot](#)

Néhány ország adatai alapján vizsgáljuk meg az átlagos iskolázottsági szint és a születéskor várható élettartam közti kapcsolatot. Adjuk meg a lineáris és az exponenciális regressziós modellt, amiben magyarázó változó az átlagos iskolázottsági szint. Melyik modell illeszkedik jobban?

	Átlagos iskolázottsági szint (év)	Születéskor várható élettartam (év)
1.	12,6	81,1
2.	12,4	78,5
3.	11,6	75,4
4.	10,4	74
5.	4,4	65,4
átlag	10,3	74,9

[Megnézem a kapcsolódó epizódot](#)

Egy cégnél 30 alkalmazottat vizsgáltak meg, hogy miként magyarázza az életkor, illetve az, hogy az illető férfi-e vagy nő (férfi=0, nő=1) a fizetés nagyságát. A kapott regressziós modell a havi fizetés nagyságát ezer forintban adja meg, ahol  $x_1$  jelenti az életkort és  $x_2$  jelenti azt, hogy az illető férfi-e vagy nő.

$$\hat{y} = 64 + 7,6x_1 - 16,7x_2 \quad s_{\hat{\beta}_1} = 4,2 \quad s_{\hat{\beta}_2} = 10,83 \quad SSE = 81,2 \quad SST = 105,7$$

Adjuk meg a modell paramétereinek jelentését. Szignifikánsnak tekinthető-e a modell alapján az életkor, illetve a nem, az alkalmazott fizetése szempontjából 10%-os szignifikanciaszinten? Teszteljük teljes modellt 10%-os [szignifikanciaszint](#) mellett.

[Megnézem a kapcsolódó epizódot](#)

Egy strand forgalmának modellezésére két magyarázó változót használunk, a napi középhőmérsékletet ( $x_1$ ) illetve azt, hogy hétvége van vagy sem ( $x_2 = 0$  ha nincs hétvége és  $x_2 = 1$  ha igen).

Egy 12 megfigyelés alapján készített modellről az alábbiakat tudjuk:

$$\hat{y} = 396 + 12,6x_1 + 18x_2 \quad s_{\hat{\beta}_1} = 2,19 \quad s_{\hat{\beta}_2} = 38,15$$

$$R = \begin{pmatrix} 1 & & \\ 0,92 & 1 & \\ -0,57 & -0,67 & 1 \end{pmatrix}$$

Adjuk meg a lineáris regressziós modell paramétereinek jelentését. Szignifikánsnak tekinthető-e a modell alapján a napi középhőmérséklet a strand forgalmának szempontjából 10%-os szignifikanciaszinten? Adjuk meg a fogalom és a hőmérséklet kapcsolatát leíró parciális [korrelációs együttható](#) értékét. Adjuk meg a többszörös determinációs hányados értékét.

[Megnézem a kapcsolódó epizódot](#)

---