

Többváltozós statisztika, ANOVA és adatredukciós módszerek (BMNPS07700M)

Előadó: Soltész-Várhelyi Klára

Többszemponos ANOVA

Első rész

Elmélet és független szempontok

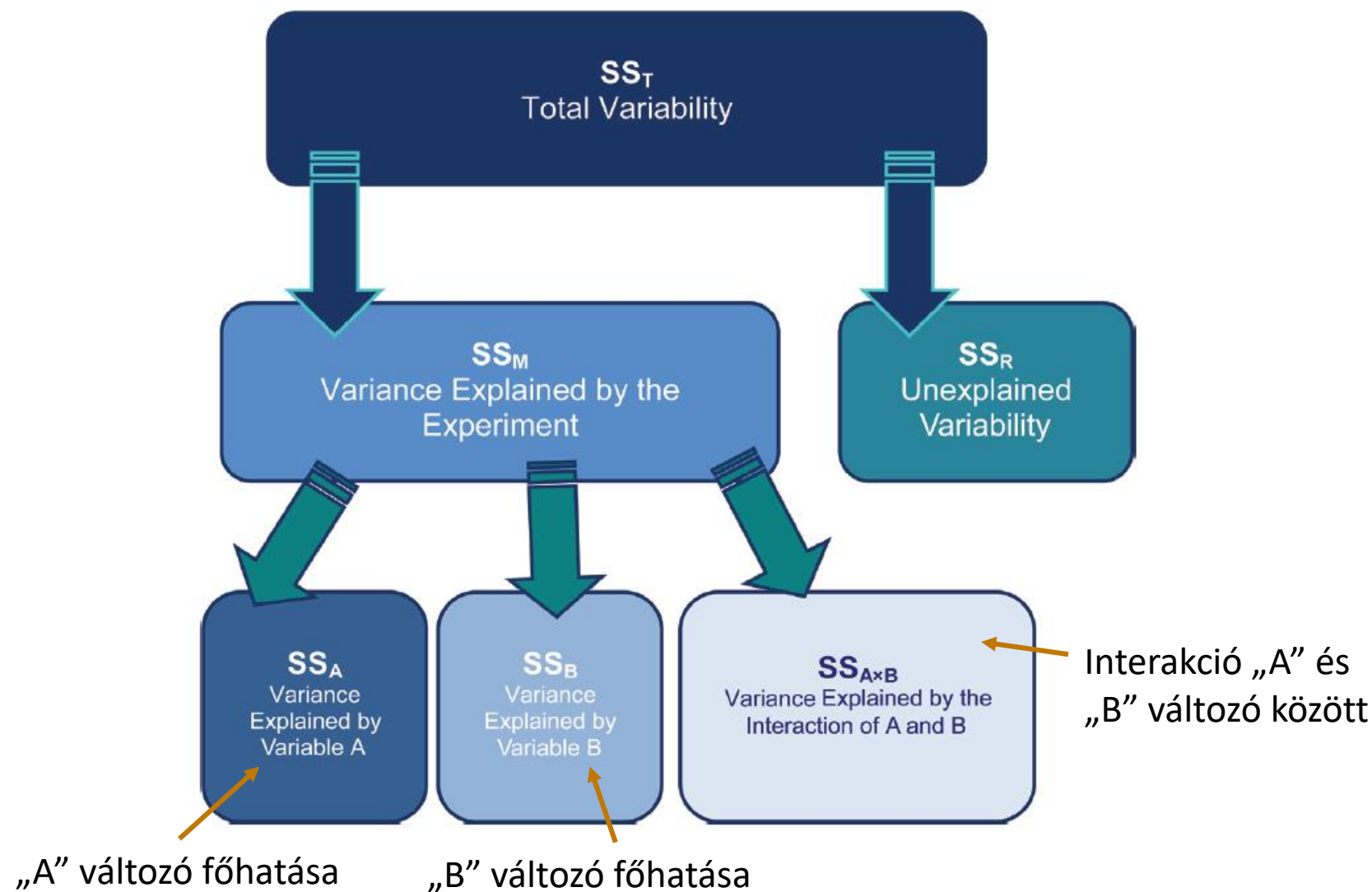
Elmélet

„However, independent variables often get lonely and want to have friends. Scientists are obliging individuals and often put a second (or third) independent variable into their designs to keep the others company”

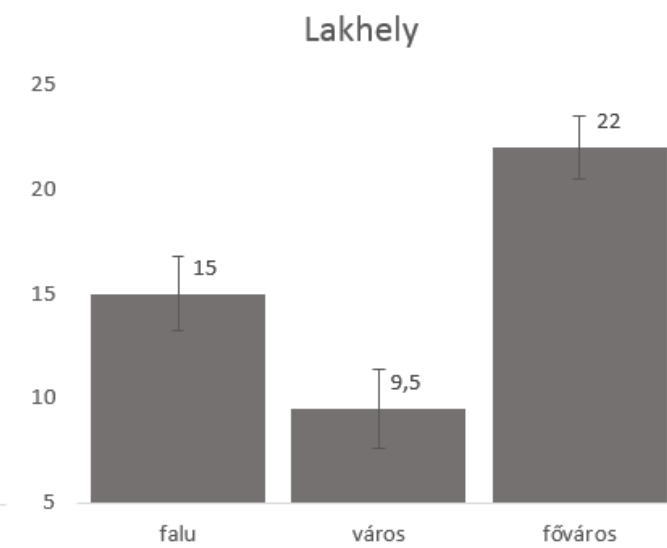
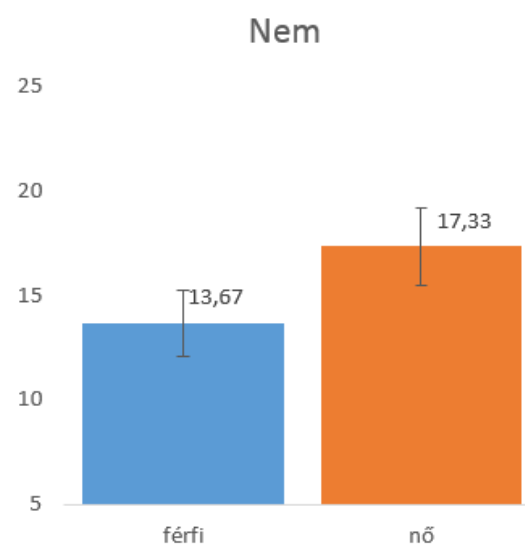
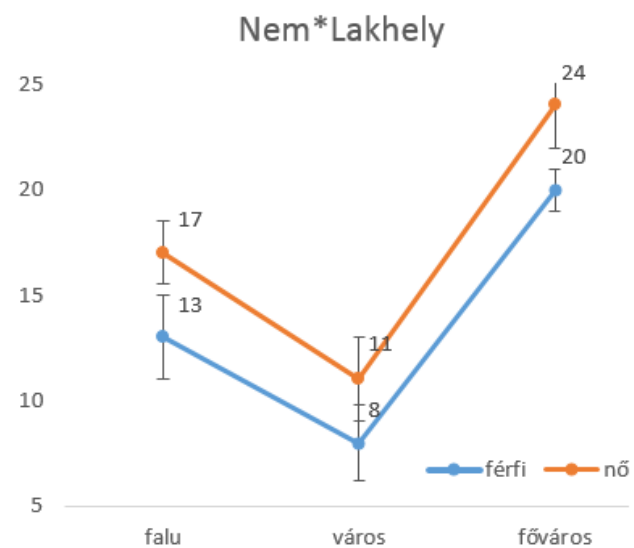
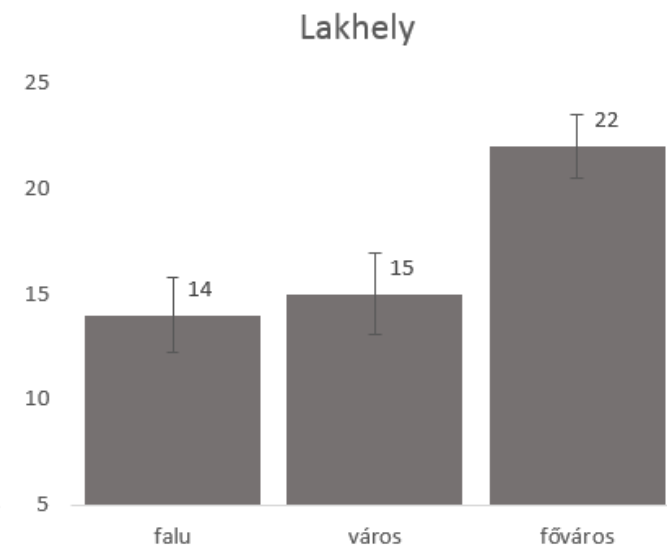
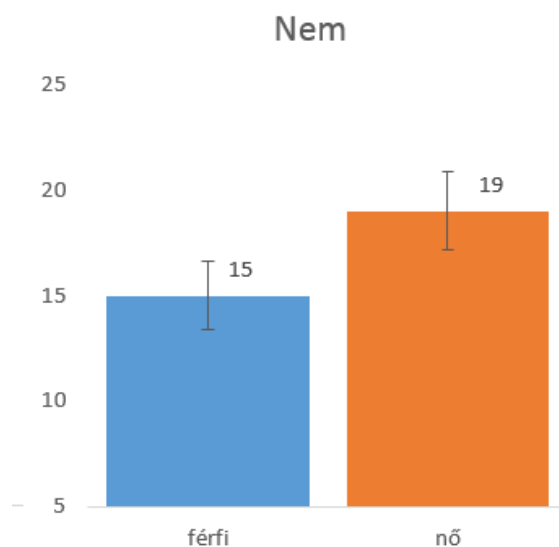
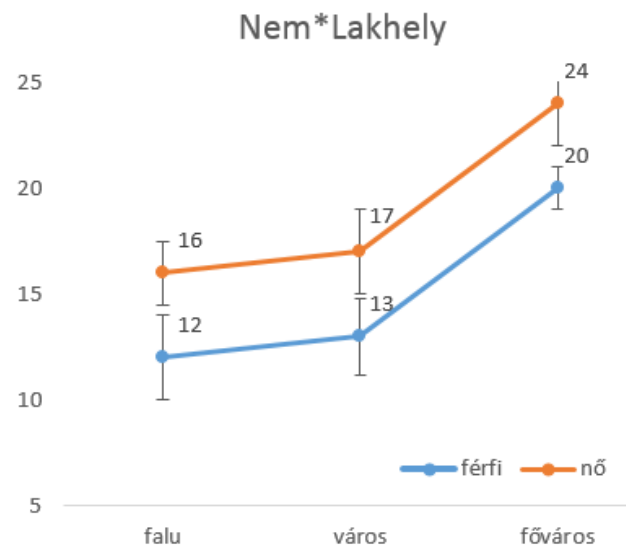
Andy Field

- 3 design:
 - Független-mintás többszemponos (faktoriális) ANOVA
 - Összefüggő mintás többszemponos (faktoriális) ANOVA
 - Kevert mintás ANOVA
- Folytonos változót is bevéve: ANCOVA

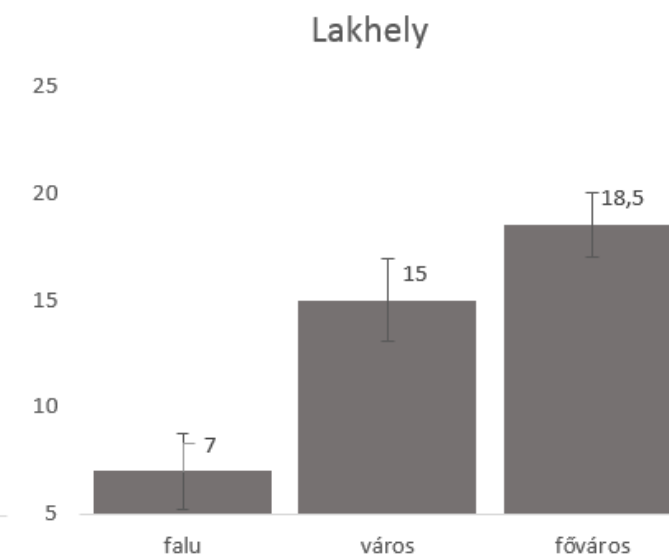
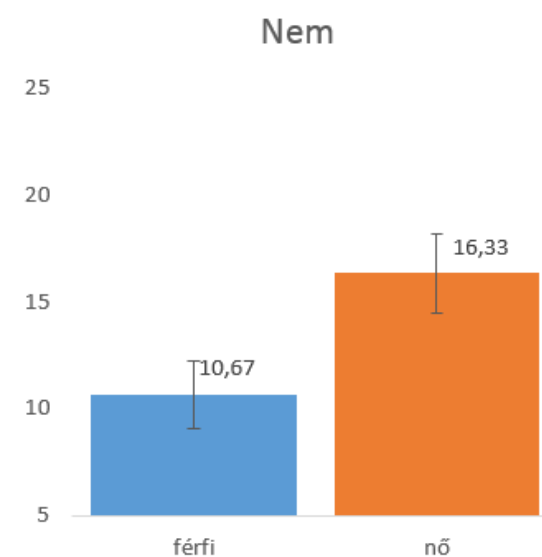
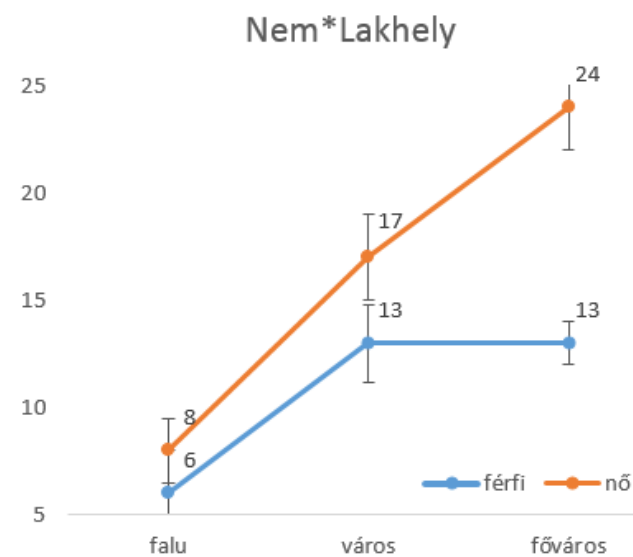
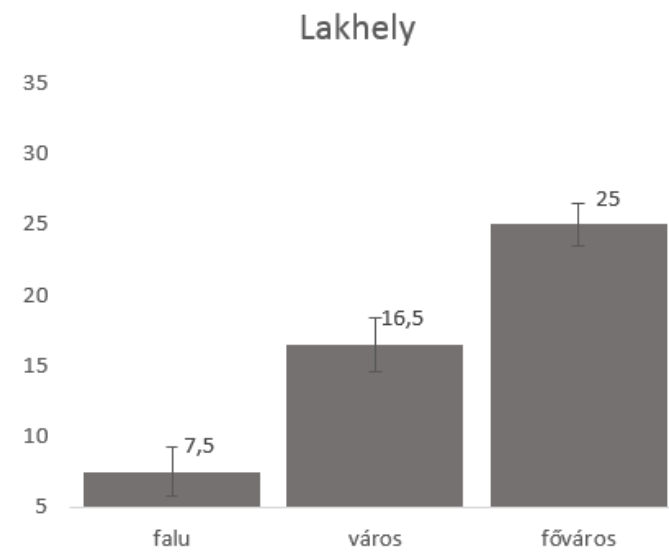
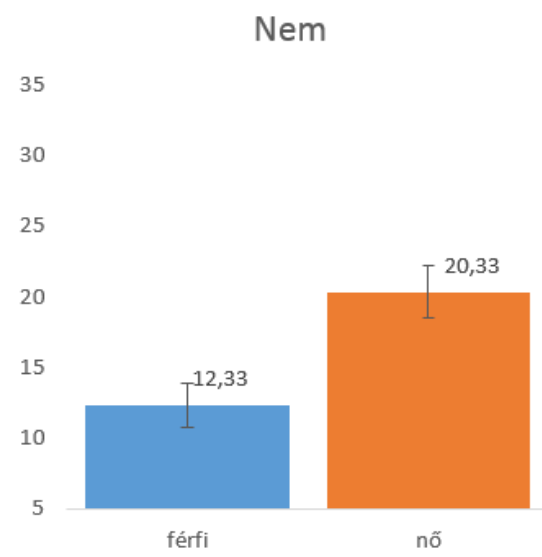
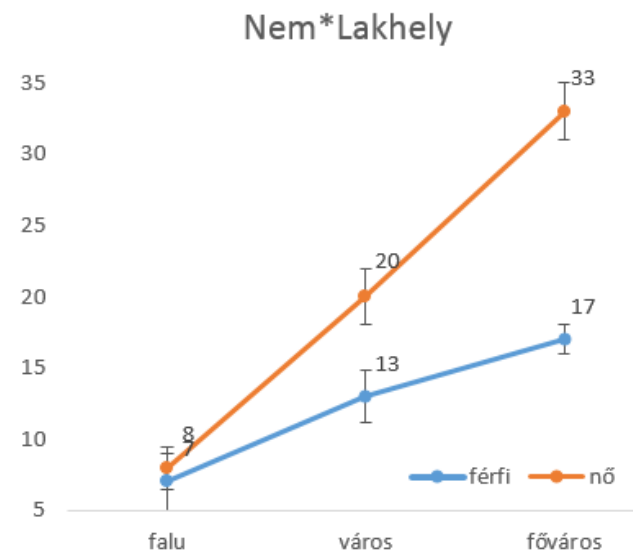
A variancia felosztása egy több szempontos modellben



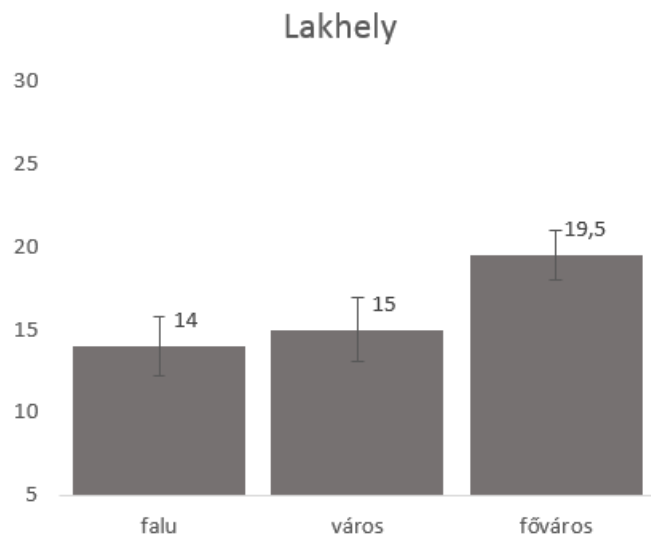
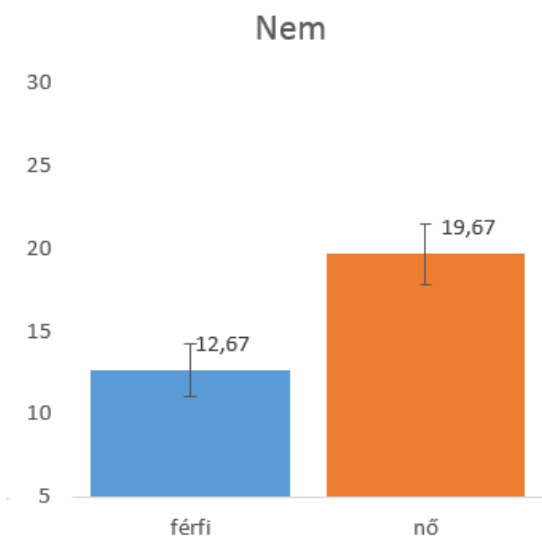
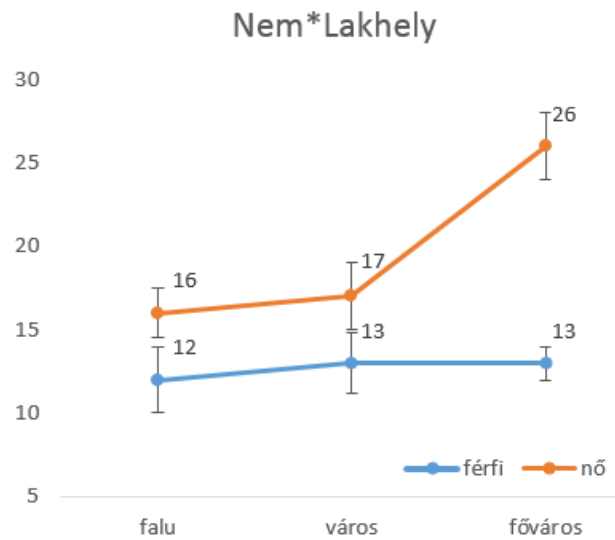
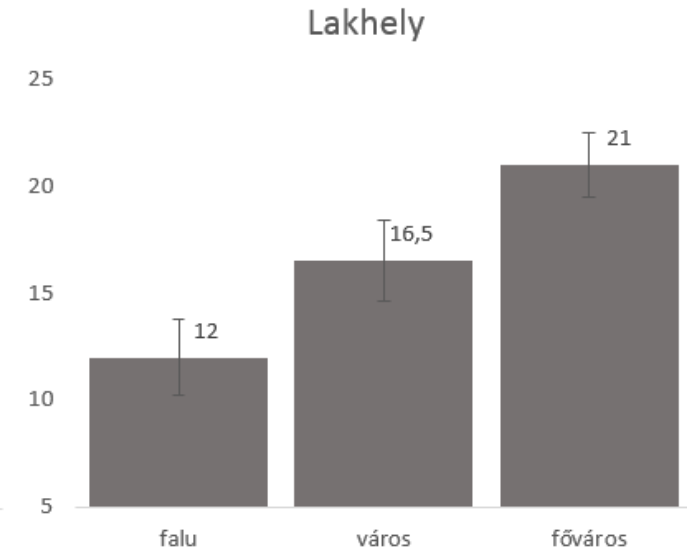
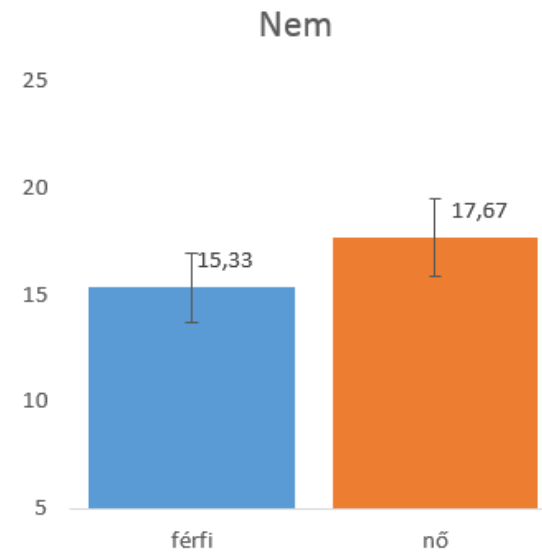
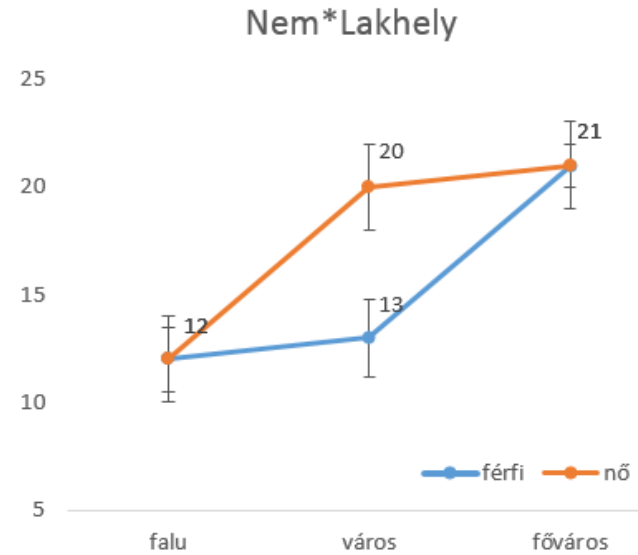
Főhatás? Interakció?



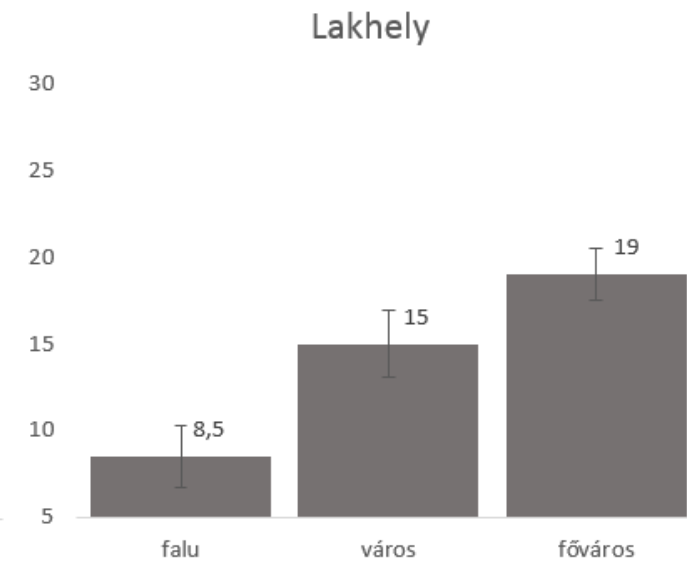
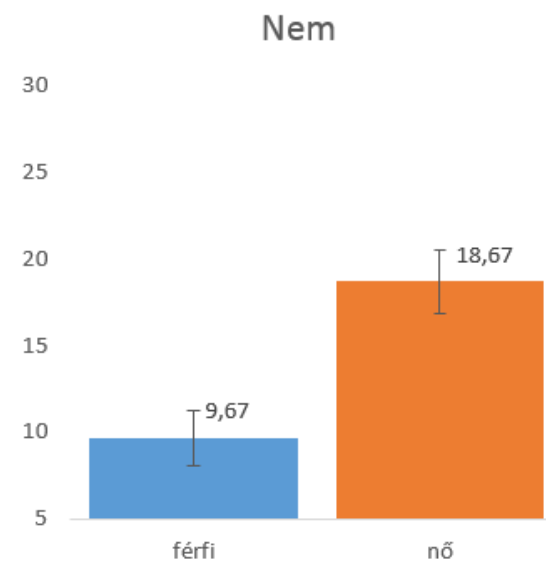
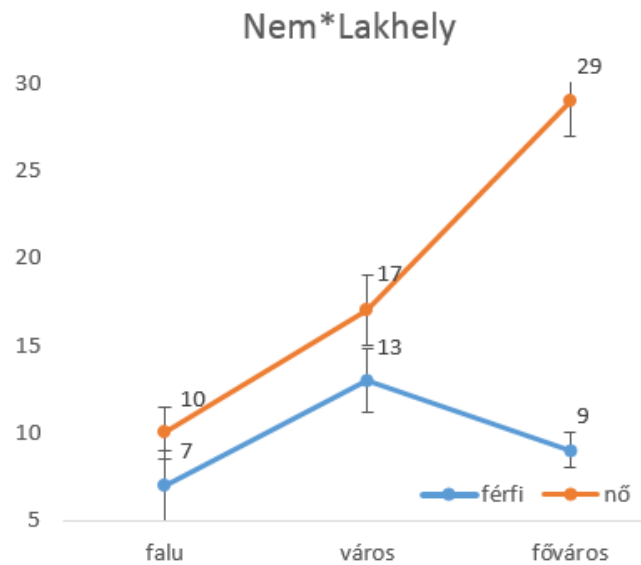
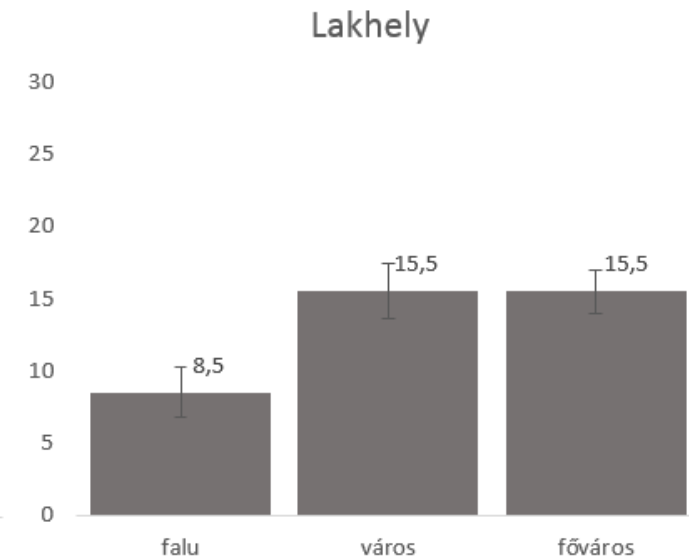
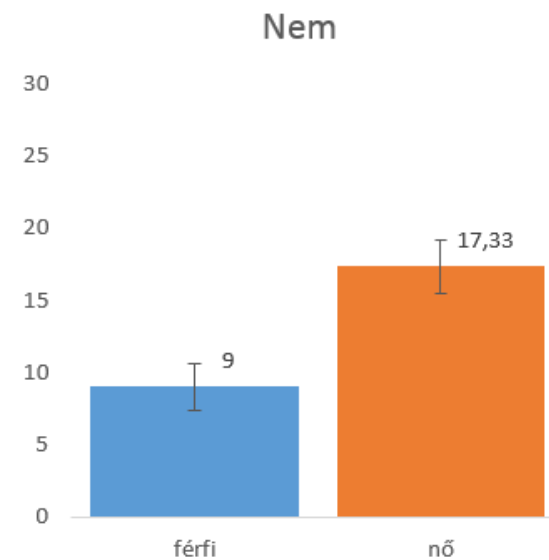
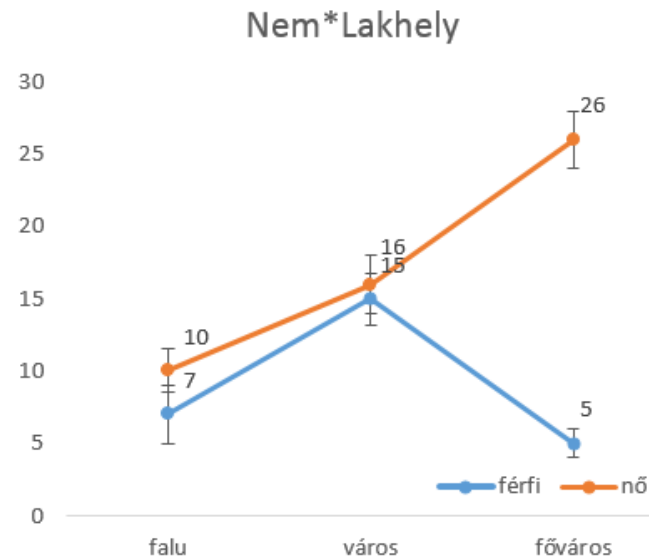
Főhatás? Interakció?



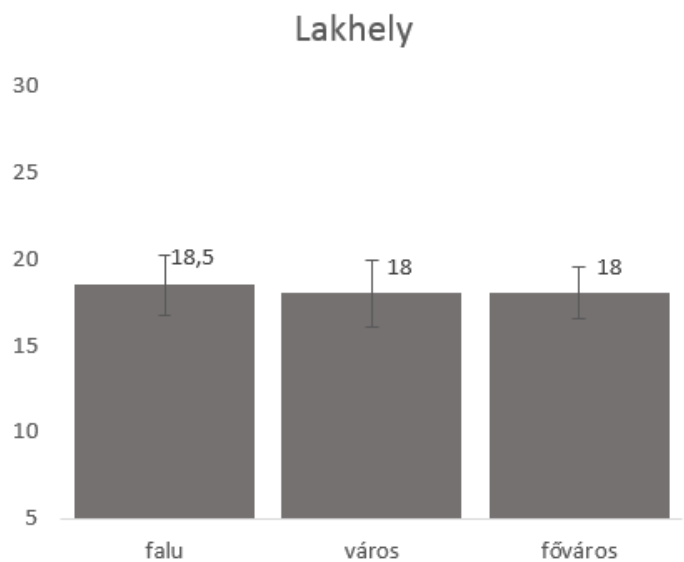
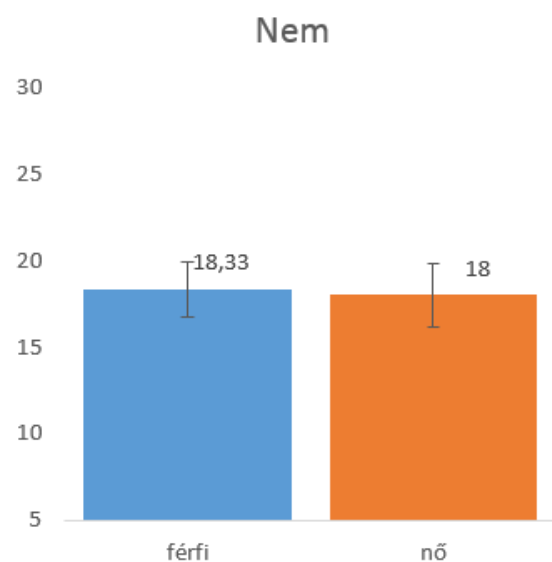
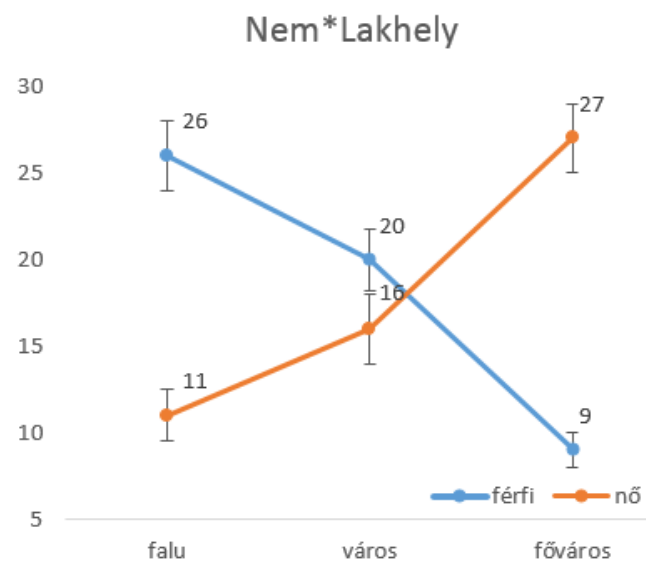
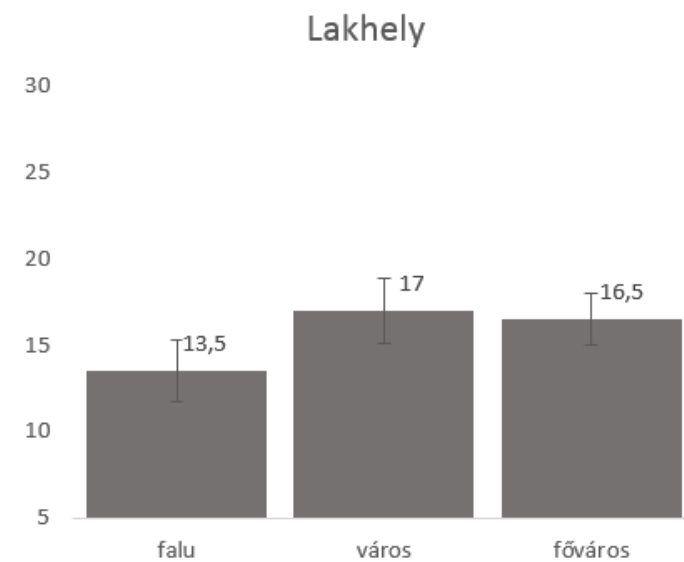
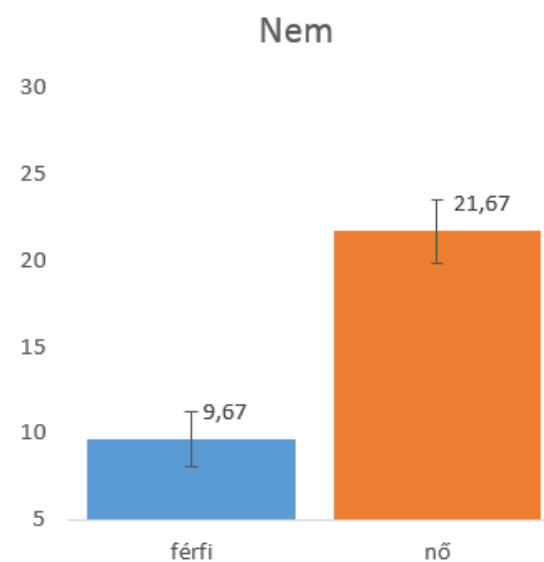
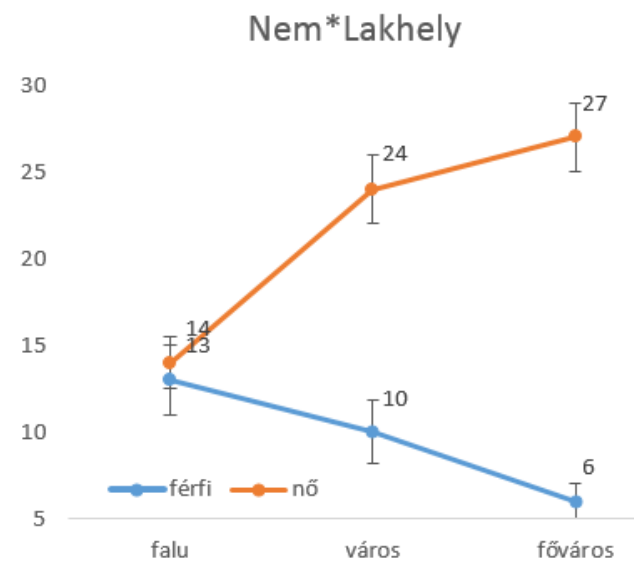
Főhatás? Interakció?



Főhatás? Interakció?

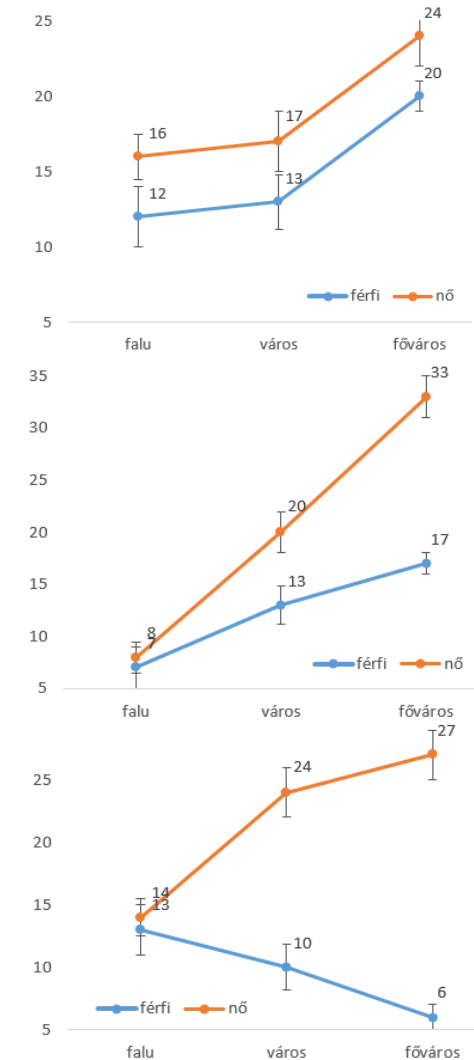


Főhatás? Interakció?



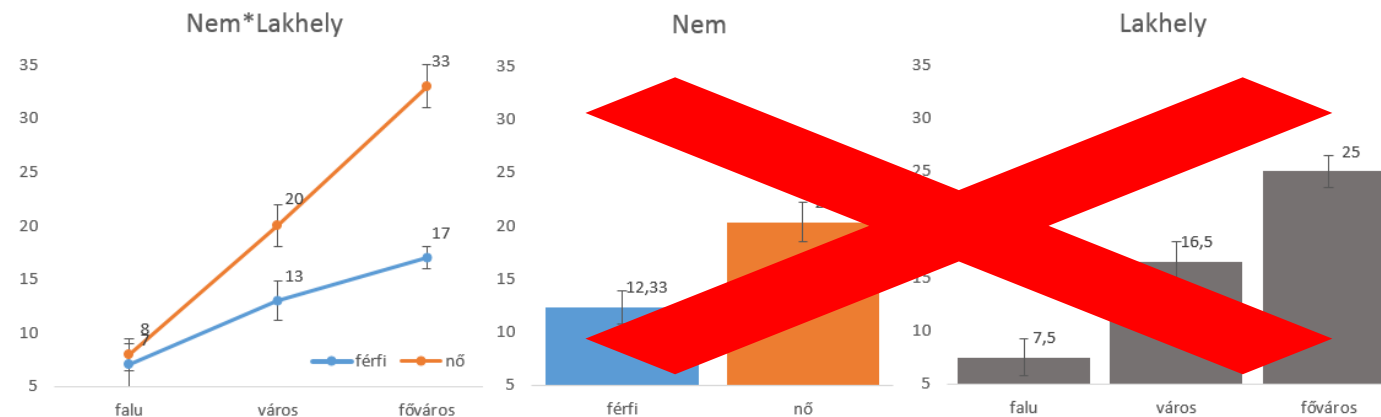
Interakció?

- Nincs
 - *Férfiakra és nőkre ugyanúgy hat a lakhely*
 - A főhatások értelmezhetők, A főhatásokon kell utóvizsgálatot végezni
- Ordinális
 - *Férfiakra és nőkre azonos irányban hat a lakhely, de a nőkre jobban*
 - A főhatások az interakció figyelembe vételével értelmezhetőek
 - A főhatásokon és az interakciókon is utóvizsgálatot kell végezni
- Diszordinális
 - *Férfiakra és nőre máshogy, más irányban hat a lakhely*
 - A főhatások NEM értelmezhetők, csak az interakciók utóvizsgálatát kell végezni
 - A változók hatásáról a Simple effects teszteléssel szerezhető információ
 - Az egyik változó minden szintjén külön kell megnézni a másik változó hatását (külön nők és férfiak esetében a lakhely hatását vagy külön falun, városban és fővárosban a nemi különbséget)
 - Ilyenkor „kézzel” kell korrigálni a familywise hibát



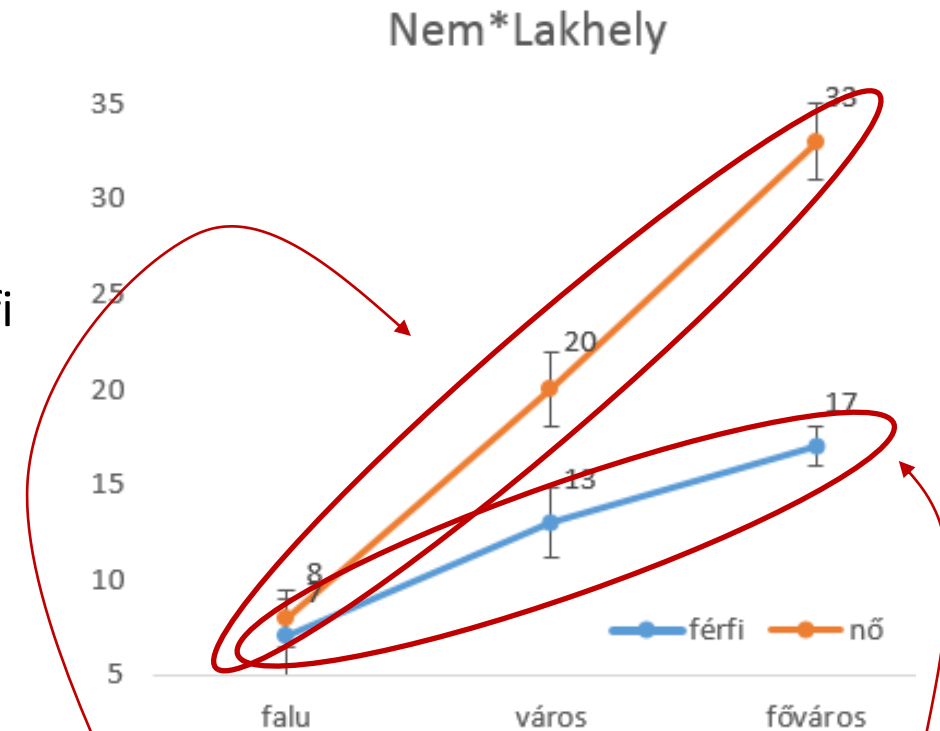
Simple effects tesztelése

- Az ANOVA tábláknak eddig 3 lehetséges részét ismerjük
 - Fő ANOVA tábla, Post hoc vagy Kontraszt tábla
- Egyszempontos elemzésben
 - egyetlen főhatás volt, ha az szignifikáns, akkor valamelyik utóvizsgálattal megnéztük, mely minták különböznek egymástól
- Többszempontos elemzésben
 - több főhatás van és köztük lehet interakció is
 - Ha az interakció szignifikáns (főleg ha diszordinális) az interakcióban résztvevő változókhoz tartozó főhatás vizsgálat és utóvizsgálatok nem értelmezhetőek.
 - Főhatás vizsgálat és utóvizsgálatok HELYETT **Simple effects** tesztelést végzünk



Simple effects tesztelése

- Interakció esetén a szempontok vizsgálatára két megoldás van:
 - 1. **One-Way ANOVÁk sora**
 - Elterjedt, kevésbé jó módszer
 - Szűrjük a mintát az egyik szempont adott szintjére, és One-Way ANOVÁkat végzünk a másik szemponton
 - Például férfiakra szűrve ANOVA a lakhelyre csak a férfi adatokon, majd nőkre szűrve másik ANOVA a lakhelyre csak a női adatokon
 - Másik példa: szűrés falura, majd nemi különbség ellenőrzése, majd szűrés városra, végül szűrés fővárosra
 - 2. **Simple effects elemzés**
 - Ugyanúgy az egyik szempont vizsgálata a másik szempont minden szintjén külön-külön, de egyetlen GLM részeként
 - Estimated Marginal Means részeként kérhető ki (később)
 - Azért jobb, mintha nagyon sok One-Way ANOVÁt végeznénk, mert a Simple effects a reziduális hibát pontosabban számolja (egyesített/pooled reziduális hibával dolgozik).



1. ANOVA és a hozzá tartozó utóvizsgálat

2. ANOVA és a hozzá tartozó utóvizsgálat

Simple effects tesztelése és a familywise error

- A Simple effects tesztelése során több ANOVÁt is végezni fogunk, ez megnöveli a familywise error valószínűségét.
- Kontrollálni kell:
 - **Bonferroni korrekció**
 - Az elemzések számával osztod a szignifikancia szintet – túl szigorú
 - Pl. 4 ANOVA esetén $0.05/4 = 0.0125$, tehát az szignifikáns, ami $p < .0125$
 - **Holm procedúra vagy Holm-Bonferroni módszer**
 - 1. Legyen az elvégzett ANOVÁk száma „k”
 - 2. Növekvő sorba kell rendezni a p értékeket (egyenlő p értékeknél nem számít a sorrend)
 - 3. Minden i-dik p értékhez hozzá kell rendelni egy $\alpha_H = \alpha/(k-i+1)$ értéket
 - 4. A szignifikancia szint az a legkisebb α_H érték lesz, ahol nem teljesül $p_i < \alpha_{Hi}$
 - Az alábbi példában a szignifikancia szint .025 lenne

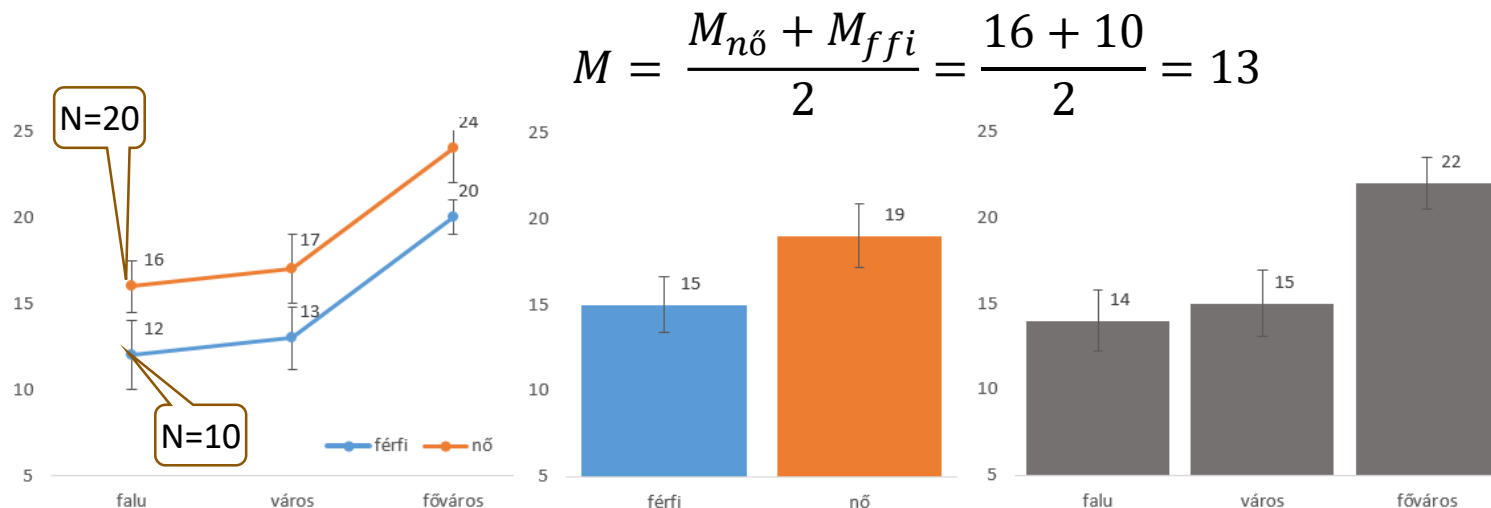
i	p	$\alpha_H = \alpha/(k-i+1)$	
1	p = .001	$0.05/(4-1+1) = 0.0125$	✓ $p_i < \alpha_{Hi}$
2	p = .014	$0.05/(4-2+1) = 0.016$	✓ $p_i < \alpha_{Hi}$
3	p = .026	$0.05/(4-3+1) = 0.025$	× $p_i > \alpha_{Hi}$
4	p = .032	$0.05/(4-4+1) = 0.05$	-

Leíró statisztikák kérdése

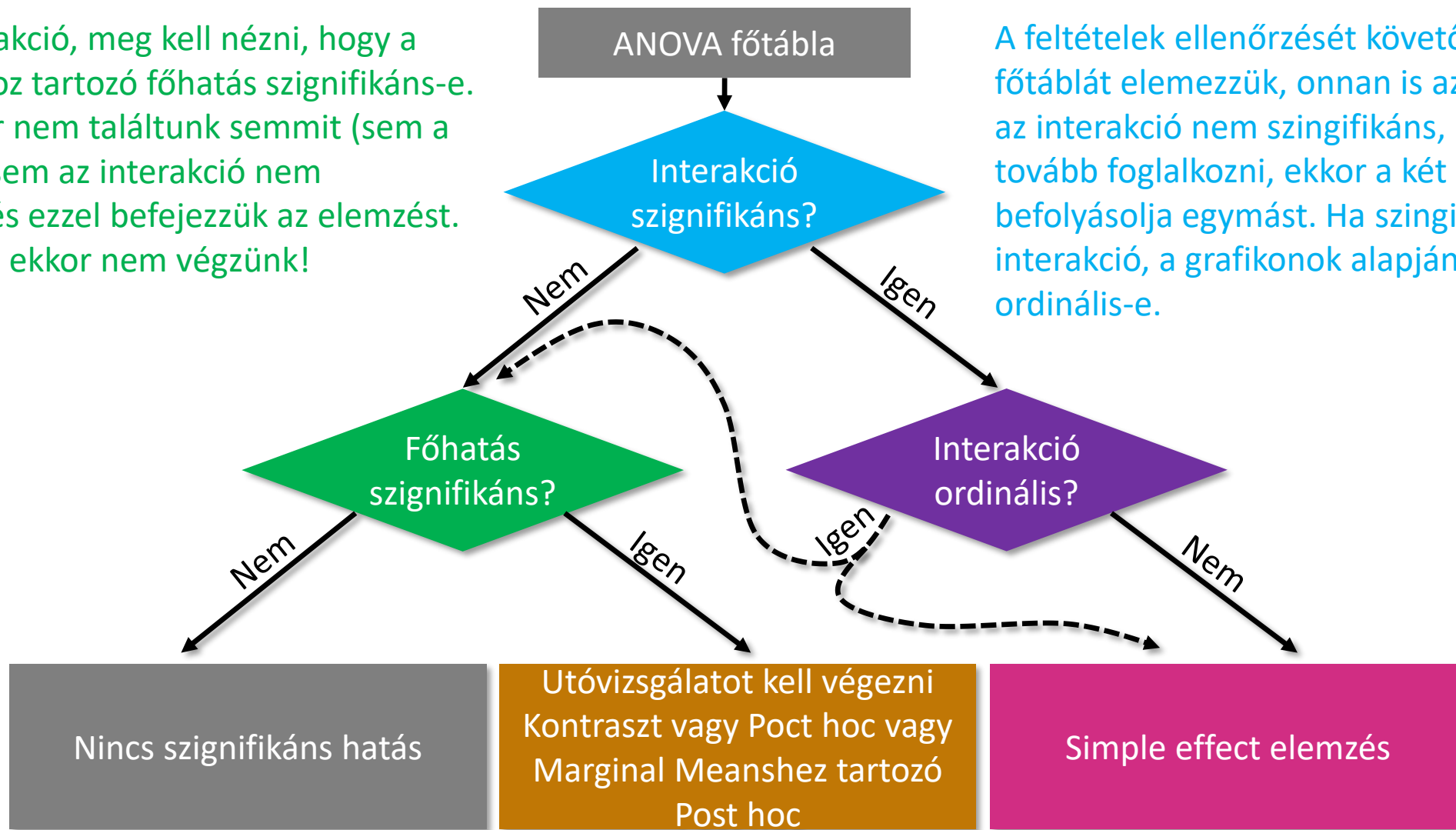
- Kétféleképp lehet egy szempont egy szintjének átlagát kiszámolni:
- **Az elemszámmal súlyozottan:**
 - a nagyobb minta több értékkel kerül be az átlagba
 - Például az összes falusi (függetlenül attól, hogy nő vagy férfi) átlaga

$$M = \frac{n_{nő} * M_{nő} + n_{ffi} * M_{ffi}}{n_{nő} + n_{ffi}} = \frac{20 * 16 + 10 * 10}{20 + 10} = 14$$

- **Súlyozatlanul:**
 - A férfi falusi minta átlagát és a női falusi minta átlagát, és az elemszámot figyelmen kívül hagyva átlagoljuk őket. Ez az **Estimated Marginal Means**
 - Ez történik az ANOVÁban (bár a modell paraméterei között beállítható más módszer)
 - Ezért előnyös, ha egyenlő elemszámmal dolgozol



Ha nincs interakció, meg kell nézni, hogy a szempontokhoz tartozó főhatás szignifikáns-e. Ha nem, akkor nem találtunk semmit (sem a szempontok, sem az interakció nem szignifikáns), és ezzel befejezzük az elemzést. Utóvizsgálatot ekkor nem végzünk!



A feltételek ellenőrzését követően először a főttáblát elemezzük, onnan is az interakciót. Ha az interakció nem szignifikáns, nem kell vele tovább foglalkozni, ekkor a két szempont nem befolyásolja egymást. Ha szignifikáns, az interakció, a grafikonok alapján meg kell nézni, ordinális-e.

Ha egy főhatás szignifikáns, meg kell nézni, az adott szemponton belül mely csoportok között van különbség. Ez történhet Kontraszttal, Post hoc-kal vagy a Marginal Meanshez tartozó Post hoc-kal is.

Ha az interakció szignifikáns, meg kell nézni, hogy ordinális-e. Ha nem, akkor a főhatások (és ezzel együtt a Post hoc és kontraszt táblák) nem értelmezhetők, helyette a Simple effect tábla eredményét kell nézni. Ha ordinális az interakció, akkor dönteni lehet: a Simple effect tábla vagy a főhatások kerülnek elemzésre (esetleg mindkettő).

Több szempontos ANOVA

Adatbázis: tobbval2_02_ANOVA02_spiderxsneake.sav

	azonosito	pok_foto	pok_elo	pok_erint	kigyo_foto	kigyo_elo	kigyo_erint	GAD	lakhely	kor
1	4	13	22	26	7	10	15	0	1	25,00
2	7	13	20	19	8	8	16	0	1	23,00
3	19	8	14	14	10	10	11	0	1	38,00
4	20	13	19	23	13	9	13	0	1	28,00
5	24	11	15	25	8	10	16	0	1	33,00
6	27	13	20	19	9	6	15	0	1	19,00
7	39	7	12	15	9	15	10	0	1	48,00
8	40	13	19	21	14	9	14	0	1	19,00
9	44	10	14	25	7	15	18	0	1	33,00
10	47	13	22	19	7	11	17	0	1	25,00
11	59	6	11	13	9	16	12	0	1	34,00
12	60	14	15	20	13	10	13	0	1	44,00
13	67	15	17	20	7	11	18	0	1	36,00
14	79	7	10	14	10	15	10	0	1	46,00
15	80	15	12	19	14	11	13	0	1	31,00
16	87	17	20	20	5	11	19	0	1	48,00
17	99	8	11	13	10	10	9	0	1	34,00
18	100	17	13	20	14	7	13	0	1	42,00
19	107	18	20	19	5	11	21	0	1	33,00
20	119	10	14	14	11	8	11	0	1	50,00

- Összefüggő mintás rész:
 - Szempont1: közelség (3 szint)
 - Szempont2: állat (2 szint)
- Független mintás rész:
 - Csoportosító változó1: GAD (2 szint)
 - Csoportosító változó2: lakhely (3 szint)
- Folytonos változó:
 - Kor



Több független szempont

*Van-e különbség klinikai szorongó és kontrollcsoport
valamint különböző kontinensről származó emberek
között abban, hogy mennyire szoronganak egy pókról
készült képre?*

The ONLY thing
we have to fear is
FEAR ITSELF...
and spiders



- Nézzük meg az adatainkat grafikusán: Graphs / Chart Builder.
- Válasszuk a Clustered oszlopdiagramot, és jelenítsük meg X tengelyen, hogy honnan származik (lakhely), színnel pedig azt, hogy valaki klinikai szorongó-e vagy sem (GAD)

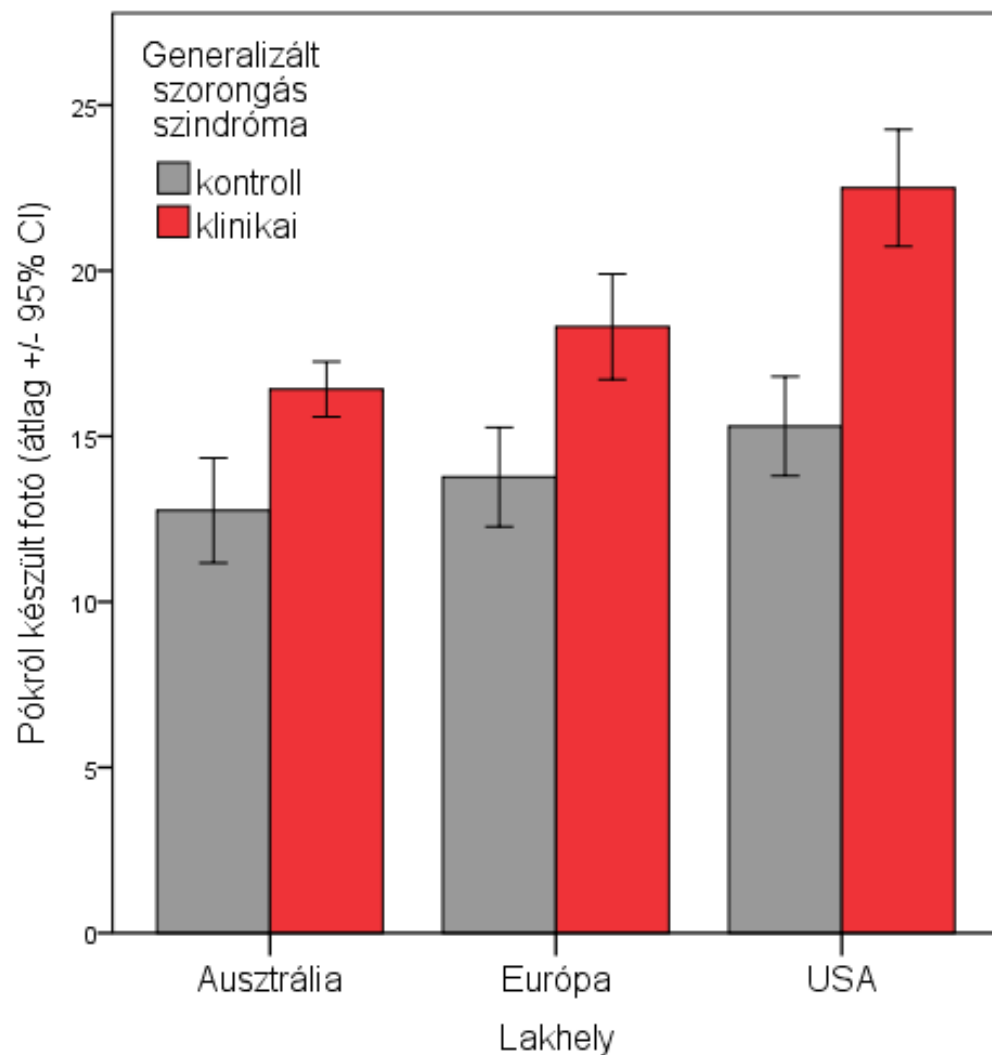
The image shows two windows from the SPSS software interface: the **Chart Builder** window and the **Element Properties** dialog box.

Chart Builder Window:

- Variables:** A list of variables is shown on the left, including `azonosito`, `pok_foto`, `pok_elo`, `pok_erint`, `kigyo_foto`, `kigyo_elo`, `kigyo_erint`, `GAD`, and `lakhely`.
- Chart preview:** A preview of a clustered bar chart is shown. The Y-axis is labeled "Y-Axis?" and the X-axis is labeled "X-Axis?". The chart has three groups of bars, each with three bars of different colors (blue, green, yellow). Annotations point to the chart: "Csop. változó: GAD" points to the top of the chart, "Függő változó: pok_foto" points to the Y-axis, and "Csop. változó: lakhely" points to the X-axis. A red circle highlights the "Cluster on X: set color" button in the top right corner of the preview area.
- Gallery:** A section at the bottom left shows various chart types. The "Bar" chart type is selected, and a red circle highlights the "Clustered" bar chart icon.
- Buttons:** At the bottom of the window are buttons for "OK", "Paste", "Reset", "Cancel", and "Help".

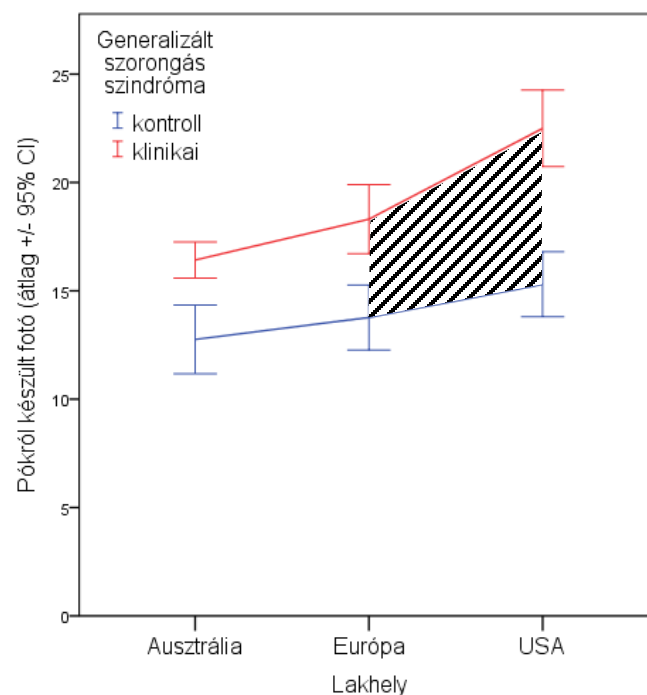
Element Properties Dialog Box:

- Edit Properties of:** A list of elements is shown, including "Bar1", "X-Axis1 (Bar1)", "Y-Axis1 (Bar1)", and "GroupColor (Bar1)".
- Statistics:** The "Variable" is set to "Pókról készült fotó". The "Statistic" is set to "Mean". A "Set Parameters..." button is visible.
- Display error bars:** This option is checked.
- Error Bars Represent:** The "Confidence intervals" radio button is selected. The "Level (%)" is set to "95". The "Multiplier" is set to "2".
- Bar Style:** The "Bar" style is selected.
- Buttons:** At the bottom of the dialog are buttons for "Apply", "Cancel", and "Help".



Mire számíthatunk?

- Interakció: lehet, hogy lesz ordinális interakció. Bár mindkét csoportban $AUS < EU < USA$, és minden kontinensen $kontroll < klinikai$, de a klinikai csoportban nagyobbak a lakhelyek közötti különbségek, mint a kontrollban
- Főhatások: csoport szerint $kontroll < klinikai$ és lakhely szerint $AUS < EU < USA$



Független csoportok közötti különbséget oszlopdiagramon illik ábrázolni (és a dolgozatodba is ezt tedd be), azonban az interakciókat gyakran könnyebb vonaldiagrammon látni. Az ábrán berajzoltam egy potenciális interakciót – majd meglátjuk, szignifikáns-e.

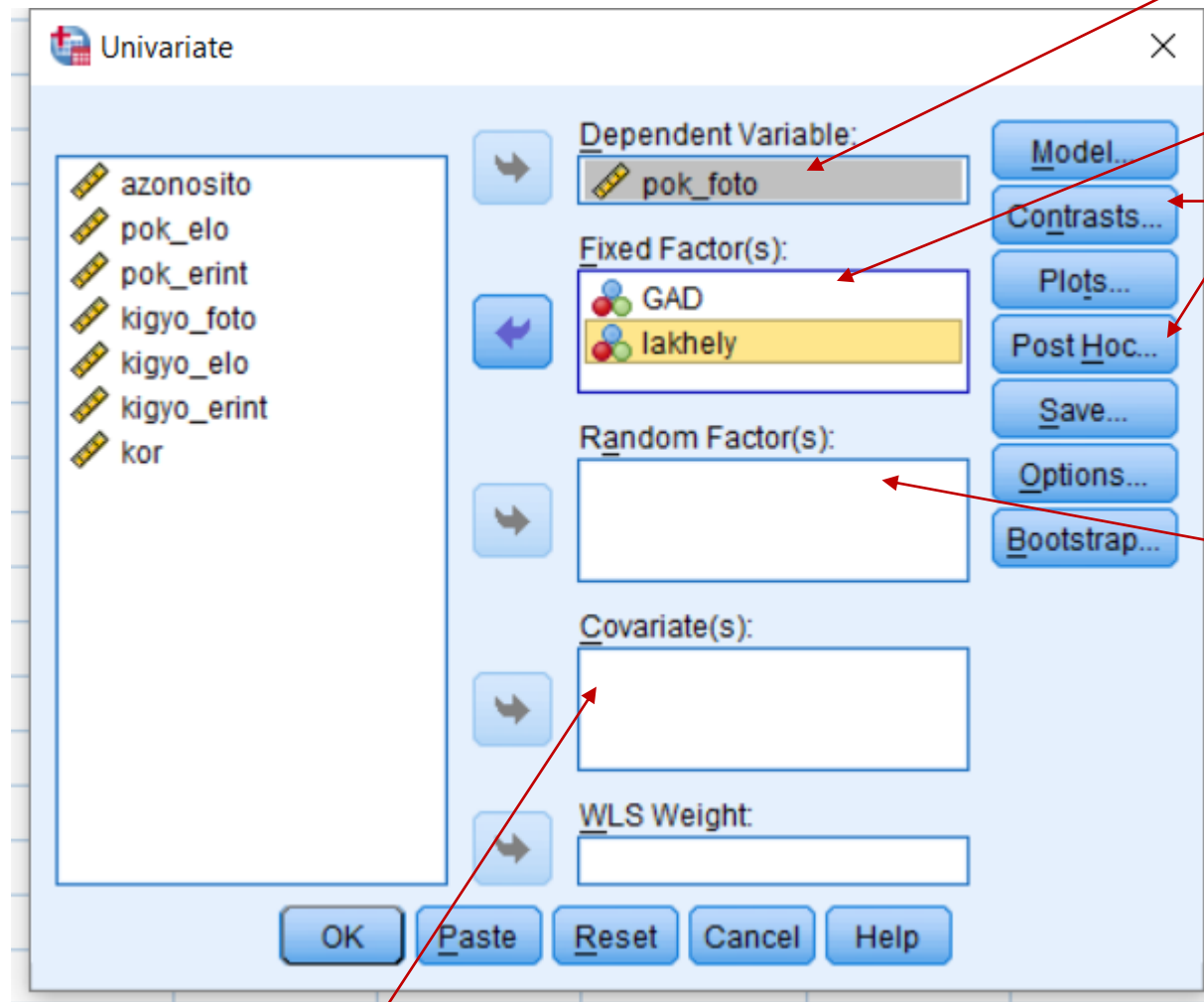
- **Feltételek**

- Megegyezik az egyszempontos független mintás ANOVA feltételeivel (részletek ott):
 - A kitöltők függetlenek
 - A független változók csoportjai függetlenek
 - Független változó folytonos
 - Normalitás
 - Szóráshomogenitás
 - Nincsenek (többdimenziós) outlierok

- **Robusztusság**

- Az ANOVA robusztus egyes feltételek sérülésére (főleg normalitás és szóráshomogenitás sérülésére vannak vizsgálatok)
- (Részletek az egyszempontos ANOVA diáiban)

- Analyze / General Linear Models / Univariate
(Az Univariate itt azt jelenti, hogy egy függő változó lesz)



Függő változó

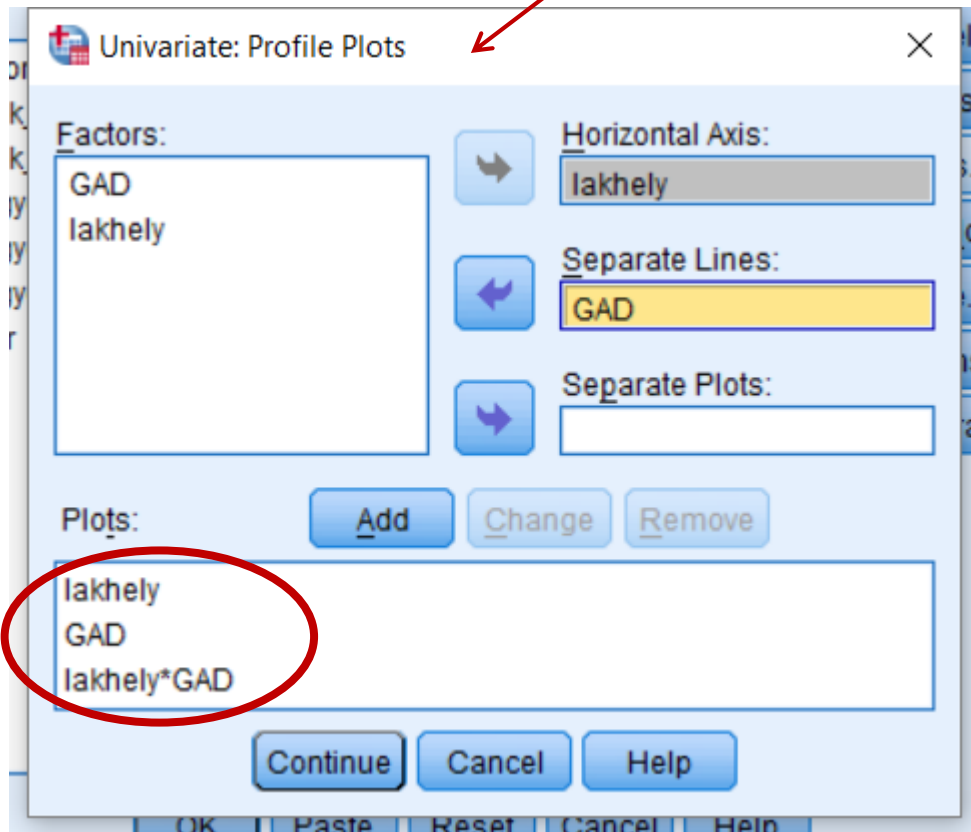
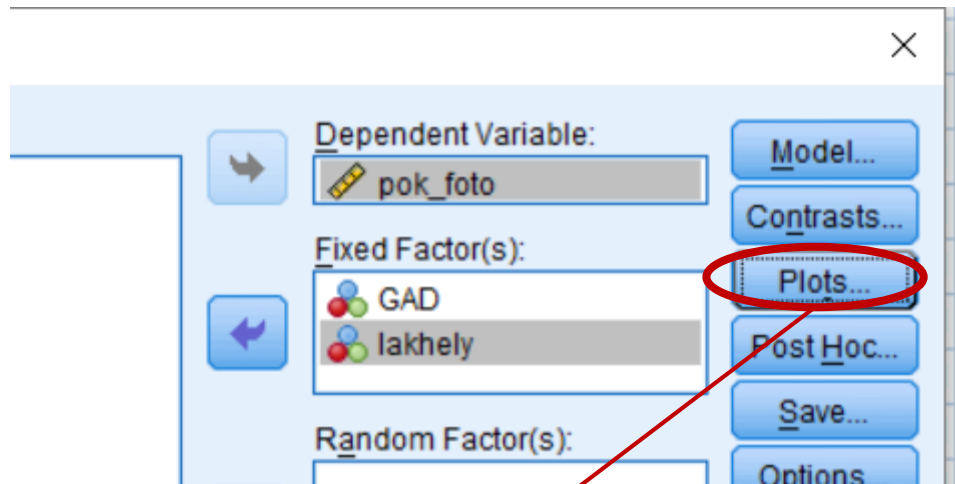
Független változók

Ha elég magabiztos vagy, beállíthatsz rögtön utóvizsgálatot is, és a fő ANOVA tábla eredményei alapján utólag is eldöntheted, mit értelmezel belőle, de így nagyon gyorsan nagyon nagyra duzzadhat az outputod, ezért a kezdetekben nem ajánlom.

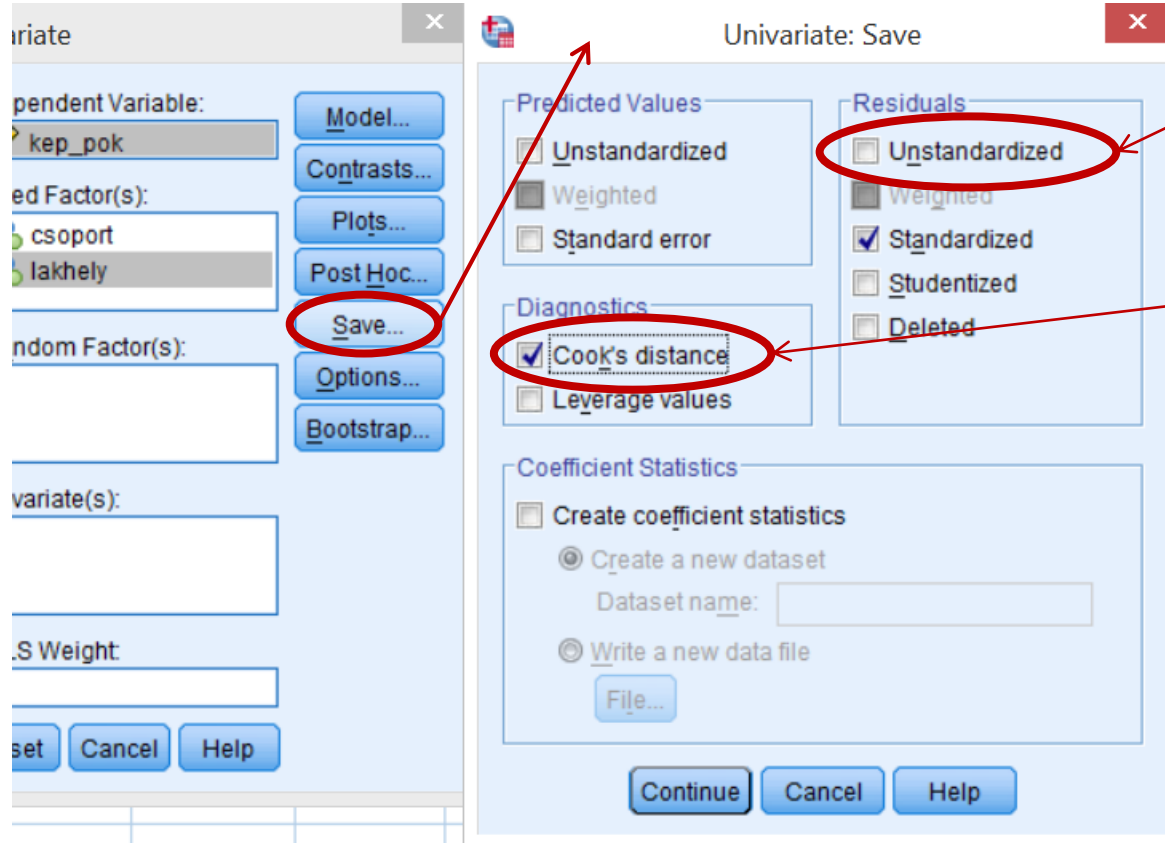
A független változók általában Fixed Factorként jelennek meg. A Random Faktort akkor használjuk, ha a független változó egy nagyon sok csoportos változóból véletlenszerűen kiválasztott kisszámú csoportból áll

Például random faktor lehetne, ha a világ 196 országából véletlenszerűen kiválasztottunk volna tízet.

A kovariáns olyan folytonos változó, mely befolyásolja a független és függő változók kapcsolatát
A kovariánst is tartalmazó elemzésről (ANCOVA) később beszélünk

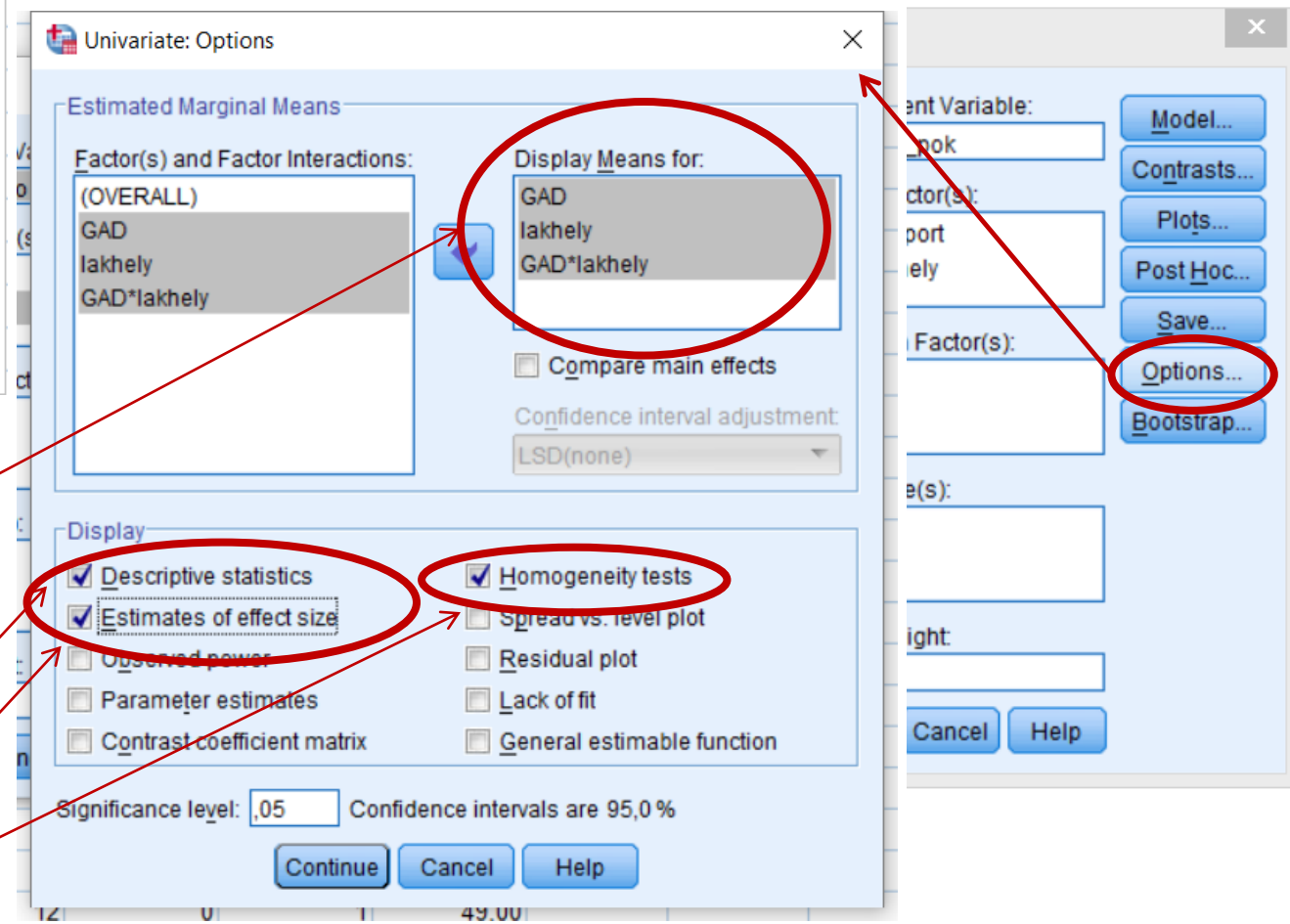


- Három Plotot kérül ki:
 - Egyszempontos plotok:
 - GAD a horizontális tengelyre – klinikai szorongó és kontroll személyek közötti különbség ábrázolására (lakhely szerint összeátlagolja az embereket)
 - Lakhely horizontális tengelyre – AUS, EU, USA közötti különbségek ábrázolására (GAD szerint összeátlagolva)
 - Többszepontos plot:
 - A két szempont interakcióját ábrázolhatjuk, ha a horizontális tengelyre a lakhelyet, separate line-ra a GAD-ot tesszük
- Az első kettő más lesz, mintha mi kérnénk ki egyszempontos plotokat, mert az átlagolás során nem veszi figyelembe a csoportok elemszámát (súlyozatlan átlagot számol, úgy tesz, mintha mindegyik csoportban ugyanannyi ember lenne). Ezt nevezzük **Estimated Marginal Means**-nek, és az ANOVA ezen az elven dolgozik.
 - Amit mi tudnánk egyszempontos plotokat kikérni, azok az adatokat ábrázolnák pontosabban
 - Ezek a plotok az ANOVA eredményét ábrázolják pontosabban



Reziduális hibák normáloszlásának teszteléséhez fogunk egy új változót létrehozni

Többdimenziós outlierok teszteléséhez fogunk egy új változót létrehozni



Becsült mintaátlagok (akkor érdekes, ha az elemszámok eltérőek a mintákban) Estimated Marginal Means elemzéshez kell

Leíró statisztikák

Effect-size mutató (parciális eta²-et számol)

Szóráshomogenitás vizsgálat

- Feltételek ellenőrzése:
 - A kitöltők függetlenek ✓**
 - A független változók csoportjai függetlenek ✓**
 - Függő változó folytonos ✓**
 - Többdimenziós outlierok ✓**
 - Ehhez létrehoztunk a Save opcióban egy új változót (COO_1)
 - Egy egyed akkor többdim outlier, ha a Cook érték > 1
 - Normalitás**
 - mivel kicsi a minta, ezért nem csoportonként ellenőriztük, hanem a reziduális hibák normál eloszlását nézzük meg.
 - Ehhez létrehoztunk a Save opcióban egy új változót (ZRE_1 = z-scores of residuals).
 - A szokott módon ellenőrizzük, normál eloszlást követnek-e a reziduális hibák (következő dia)
 - Szóráshomogenitás**
 - Levene teszt eredményét kell nézni (következő dia)

ZRE_1	COO_1
,07	,00
,07	,00
-1,37	,01
,07	,00
-,51	,00
,07	,00
-1,66	,02
,07	,00
-,79	,00
,07	,00
-1,95	,03
,36	,00
,65	,00
-1,66	,02
,65	,00
1,22	,01
-1,37	,01
1,22	,01
1,51	,02
-,79	,00
1,22	,01
1,51	,02
-1,08	,01
,65	,00
1,80	,02
,07	,00
,35	,00

Descriptive Statistics

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	Lakhely	Mean	Std. Deviation	N
kontroll	Ausztrália	12,76	3,844	25
	Európa	13,77	3,713	26
	USA	15,30	3,470	23
	Total	13,91	3,779	74
klinikai	Ausztrália	16,42	2,277	31
	Európa	18,31	3,947	26
	USA	22,50	3,552	18
	Total	18,53	3,998	75
Total	Ausztrália	14,79	3,556	56
	Európa	16,04	4,432	52
	USA	18,46	5,005	41
	Total	16,23	4,519	149

Levene's Test of Equality of Error Variances^a

Dependent Variable: Pókról készült fotó

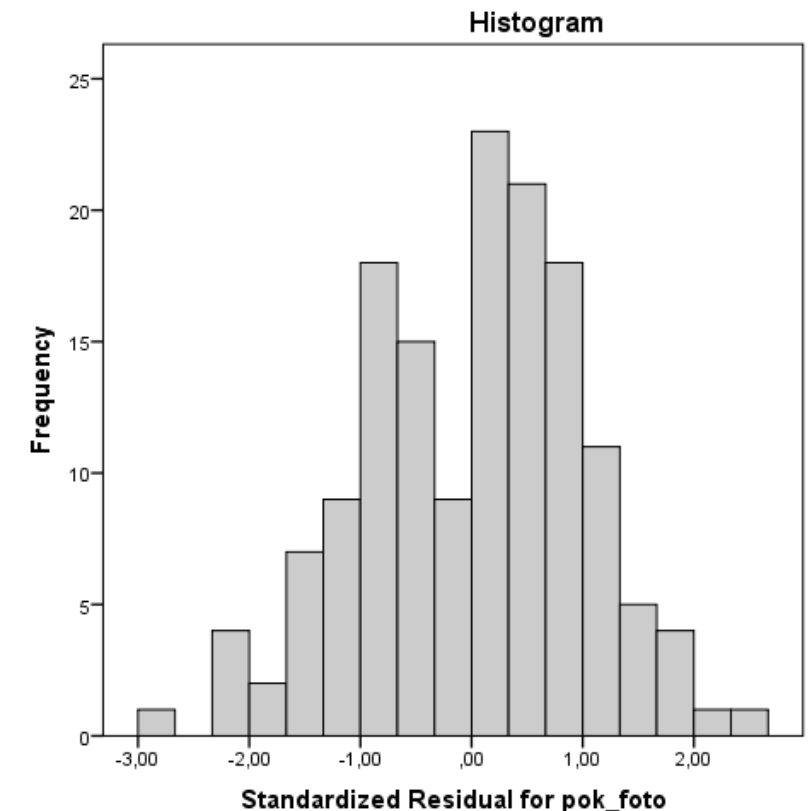
F	df1	df2	Sig.
1,548	5	143	,179

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + GAD + lakhely + GAD * lakhely

A normalitás vizsgálata S-W teszt nem szignifikáns, a K-S viszont (valószínűleg az elemszám miatt az), de ez nem gond, mert a szóráshomogenitás teljesül, tehát az ANOVA robosztus a normalitás esetleges sérülésér, és így értelmezhetőek az eredmények

El lehetne gondolkodni azon az eseten, akihez a majdnem -3-as standardizált reziduális hiba tartozik – őt nem tudtuk jól bejósolni



Leíró statisztikák

Két leíró statisztika van az outputban:

Between-Subjects Factors				
		Value Label	N	
Generalizált szorongás szindróma	0	kontroll	74	
	1	klinikai	75	
Lakhely	1	Ausztrália	56	
	2	Európa	52	
	3	USA	41	

Descriptive Statistics				
Dependent Variable: Pókról készült fotó				
Generalizált szorongás szindróma	Lakhely	Mean	Std. Deviation	N
kontroll	Ausztrália	12,760	3,844	25
	Európa	13,769	3,713	26
	USA	15,304	3,470	23
	Total	13,905	3,779	74
klinikai	Ausztrália	16,419	2,277	31
	Európa	18,308	3,947	26
	USA	22,500	3,552	18
	Total	18,533	3,998	75
Total	Ausztrália	14,786	3,556	56
	Európa	16,038	4,432	52
	USA	18,463	5,005	41
	Total	16,235	4,519	149

Az elején lévő **deskriptív statisztika** a csoportok elemszámát figyelembe veszi az átlagolásnál.

Az **Estimated Marginal Means** úgy átlagol, mintha minden csoportban egyforma elemszám lenne (súlyozatlan átlag) – az ANOVA számításai ezen alapulnak, van folyóirat, ahol ezt kéri közlésre, nem a tényleges értékeket.

Hasonlítsd össze, hogy milyen különbségek vannak a súlyozott és súlyozatlan értékek között. Minél nagyobb az elemszámok közötti eltérés, annál nagyobb lesz e két átlagszámítás eredményének eltérése is.

Estimated Marginal Means

1. Generalizált szorongás szindróma

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
kontroll	13,94	,404	13,146	14,743
klinikai	19,08	,411	18,263	19,889

2. Lakhely

Dependent Variable: Pókról készült fotó

Lakhely	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Ausztrália	14,590	,467	13,667	15,512
Európa	16,038	,481	15,087	16,990
USA	18,902	,546	17,822	19,982

3. Generalizált szorongás szindróma * Lakhely

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	Lakhely	Mean	Std. Error	95% Confidence ...	
				Lower Bound	Upper Bound
kontroll	Ausztrália	12,760	,694	11,387	14,133
	Európa	13,769	,681	12,423	15,115
	USA	15,304	,724	13,873	16,735
klinikai	Ausztrália	16,419	,624	15,187	17,652
	Európa	18,308	,681	16,962	19,654
	USA	22,500	,818	20,882	24,118

Az ANOVA vizsgálat fő táblája

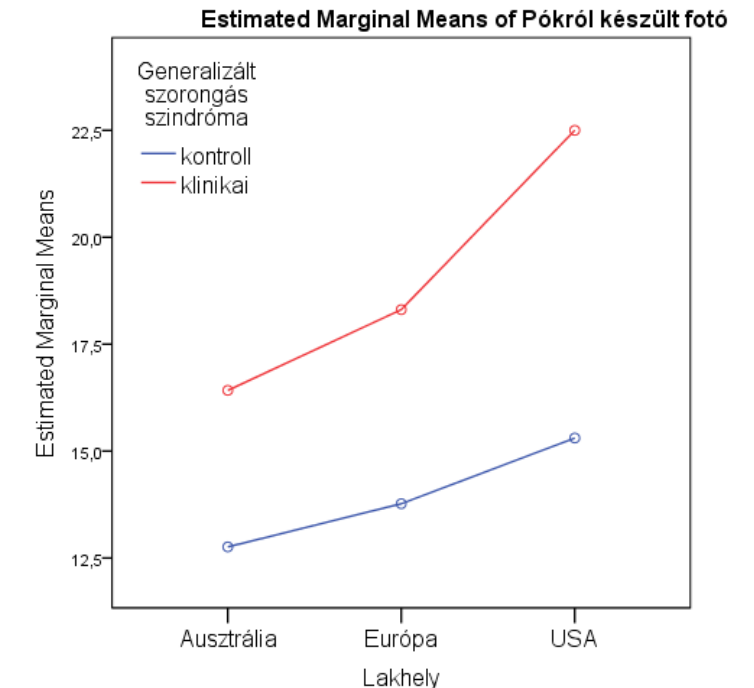
- A tábla alatti jegyzet szerint, $R^2 = .430$, és ez viszonylag közel van az Adjusted R^2 - hez
- A táblázatban látjuk a csoport és a lakhely főhatását valamint a kettő közötti interakciót
 - Először az **interakciót** nézzük meg:
 - $F(2, 143) = 3.160$ $p = .045$ *part.* $\eta^2 = .042$ ami éppen csak szignifikáns, és gyenge hatás. A grafikon alapján látjuk, hogy az interakció ordinális
 - Mivel ordinális (ráadásul gyenge) interakciót találtunk csak, a főhatások értelmezhetőek, de azért érdemes észben tartani, hogy nem biztos, hogy mindkét csoport szorongására ugyanúgy hat a lakhely (vagy megfordítva, minden kontinensen ugyanolyan mértékű a két csoport közötti különbség)

Tests of Between-Subjects Effects

Dependent Variable: Pókról készült fotó

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1299,147 ^a	5	259,829	21,557	,000	,430
Intercept	39534,312	1	39534,312	3279,939	,000	,958
GAD	954,652	1	954,652	79,202	,000	,356
lakhely	438,914	2	219,457	18,207	,000	,203
GAD * lakhely	76,168	2	38,084	3,160	,045	,042
Error	1723,632	143	12,053			
Total	42295,000	149				
Corrected Total	3022,779	148				

a. R Squared = ,430 (Adjusted R Squared = ,410)



Az ANOVA vizsgálat fő táblája

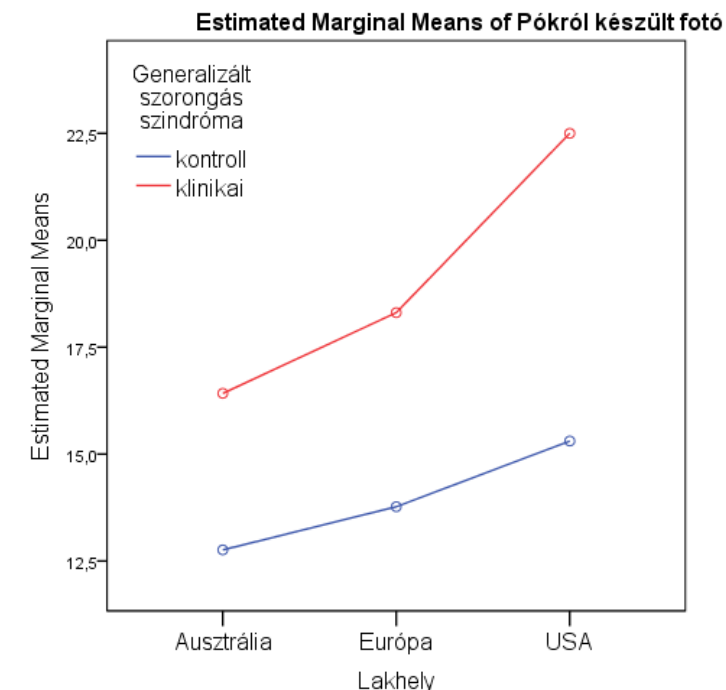
- **Főhatások:**
 - **EMLÉKEZTETŐ: CSAK AKKOR ÉRTELMEZHETŐEK A FŐHATÁSOK, HA VAGY NINCS AZ ADOTT SZEMPONTTAL SZIGNIFIKÁNS INTERKAKCIÓ, VAGY HA VAN, AKKOR AZ CSAK ORDINÁLIS!**
 - GAD: hatása szignifikáns, $F(1, 143) = 79.202$ $p < .001$ *part.* $\eta^2 = .356$, a grafikonról, vagy a leíró statisztikákból leolvashatjuk, hogy a klinikai szorongó csoport mutat nagyobb szorongást.
 - Lakhely hatása szignifikáns, $F(2, 143) = 18.207$ $p < .001$ *part.* $\eta^2 = .203$, azt tudjuk, hogy AUS, EU és USA között szignifikáns különbség van, de hogy mely csoportok különböznek egymástól szignifikánsan, azt nem. Erre utóvizsgálatot kell végeznünk. A főhatások utóvizsgálatára továbbra is a Post hoc VAGY a Kontraszt használható.

Tests of Between-Subjects Effects

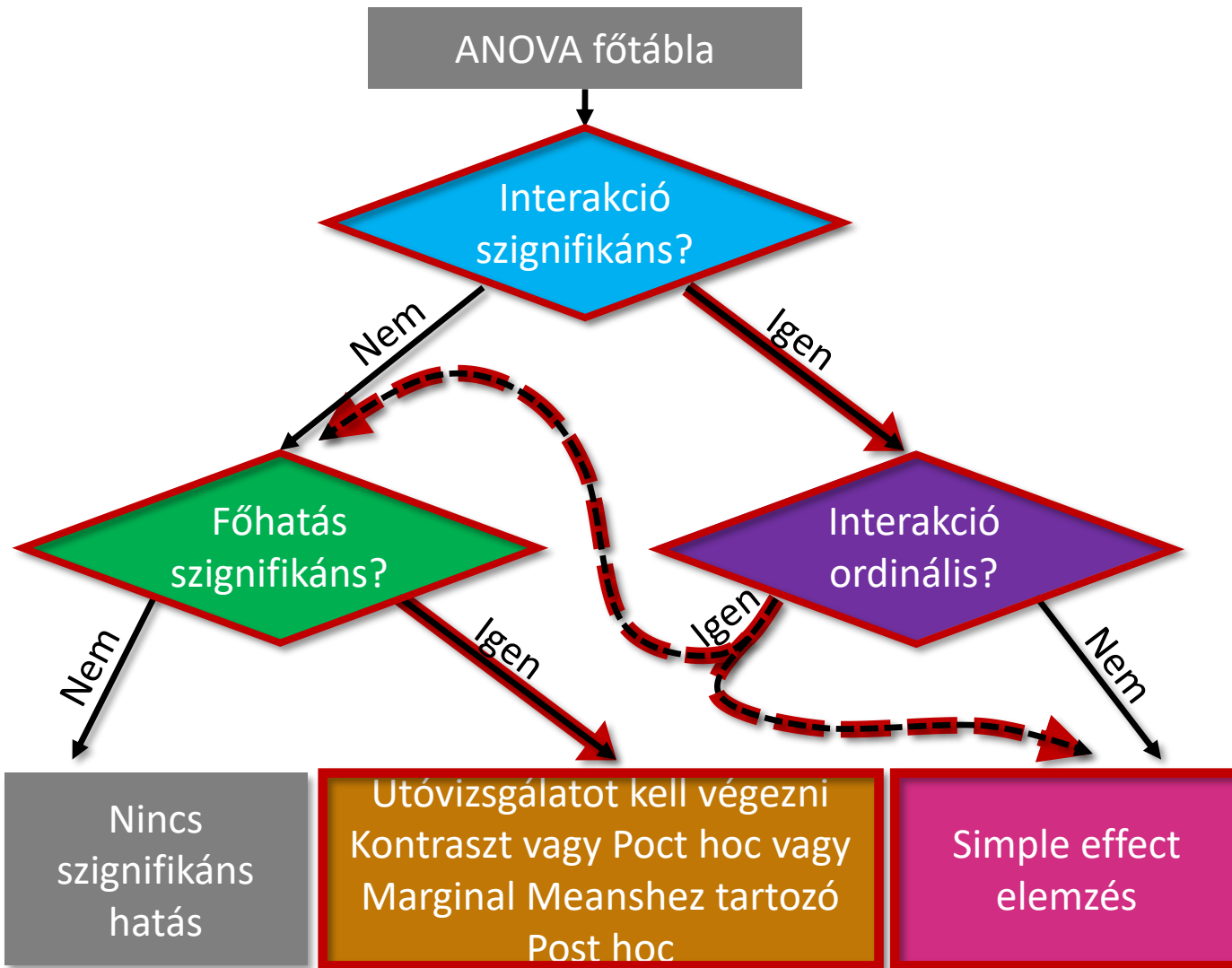
Dependent Variable: Pókról készült fotó

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1299,147 ^a	5	259,829	21,557	,000	,430
Intercept	39534,312	1	39534,312	3279,939	,000	,958
GAD	954,652	1	954,652	79,202	,000	,356
lakhely	438,914	2	219,457	18,207	,000	,203
GAD * lakhely	76,168	2	38,084	3,160	,045	,042
Error	1723,632	143	12,053			
Total	42295,000	149				
Corrected Total	3022,779	148				

a. R Squared = ,430 (Adjusted R Squared = ,410)



Hogyan tovább?



Eddigi lépések összefoglalása:

1. Fő táblában láttuk, hogy az interakció szignifikáns
2. A grafikon alapján megállapítottuk, hogy az interakció ordinális.

Innen két út áll előttünk:

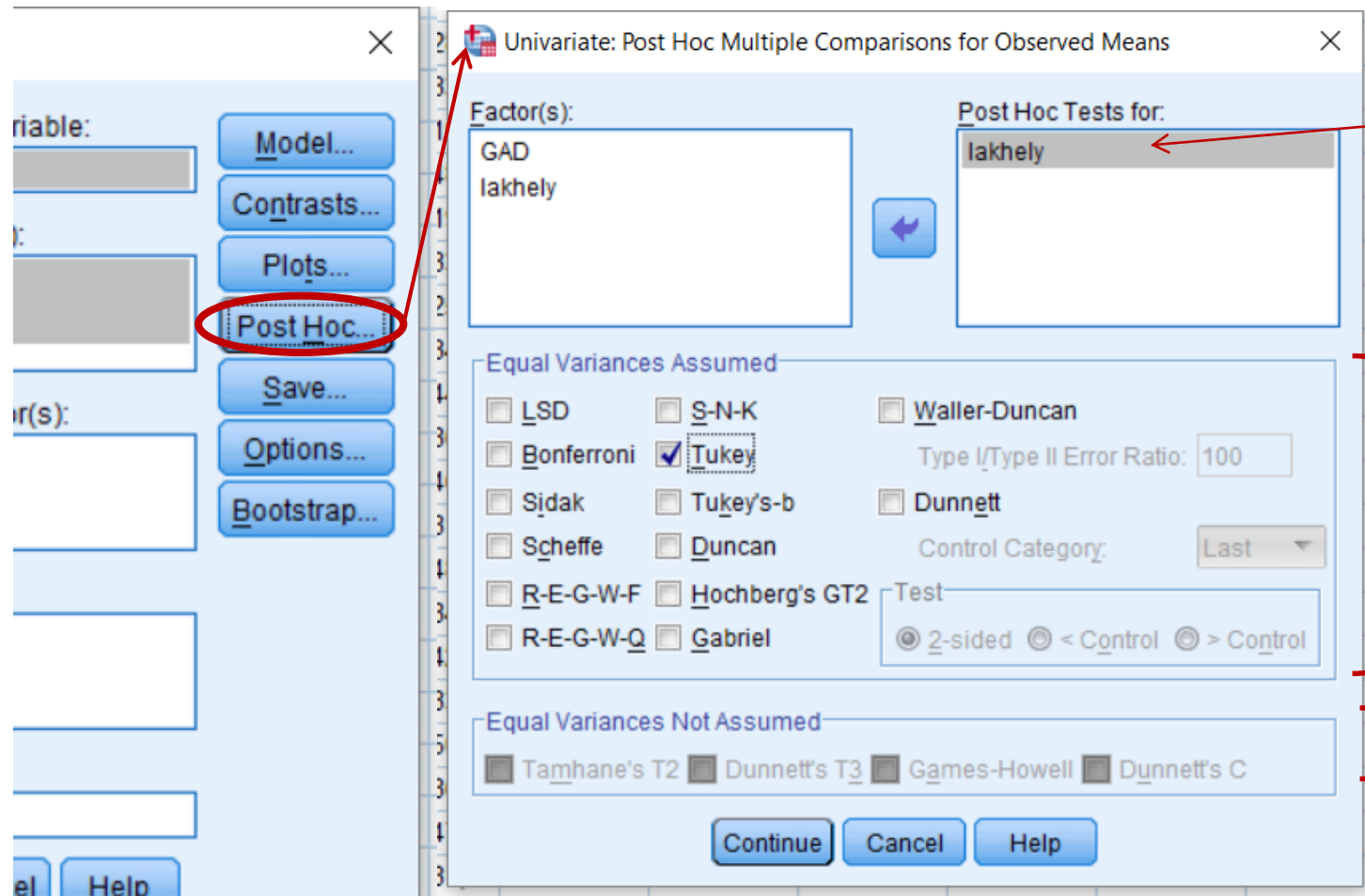
- a. Értelmezzük a főhatásokat.
 - GAD szignifikáns, a klinikai csoport magasabb szorongást mutat, mint a kontrol. Ezzel készen vagyunk.
 - Lakhely szignifikáns, de mivel 3 csoportot hasonlítottunk össze, ebből még nem tudjuk, mely csoportok különböznek, ezért utóvizsgálatot kell végezzünk rá.
- b. Simple effect tesztelést végzünk.

Következő lépések:

- A lakhely főhatásához egyszempontos utóvizsgálat, mely lehet Post hoc VAGY Kontraszt
- Simple effect elemzés az interakció értelmezésére

- **Post hoc vizsgálat**

- Post hoc-ot csak olyan szempontra kérjük mely nem vesz rész szignifikáns, diszordinális interakcióban, és a szempont főhatása a fő ANOVA táblában szignifikáns. Nézzük végig a szempontokat:
 - Csoport: csak ordinális interakcióban vesz rész, a főhatás szignifikáns, de fölösleges rá utóvizsgálatot kérni, mert csak 2 szintje van, ha van szignifikáns különbség, az a klinikai és kontroll csoport között van.
 - Lakhely: csak ordinális interakcióban vesz rész, a főhatás szignifikáns, kikérjük rá az utóvizsgálatot.



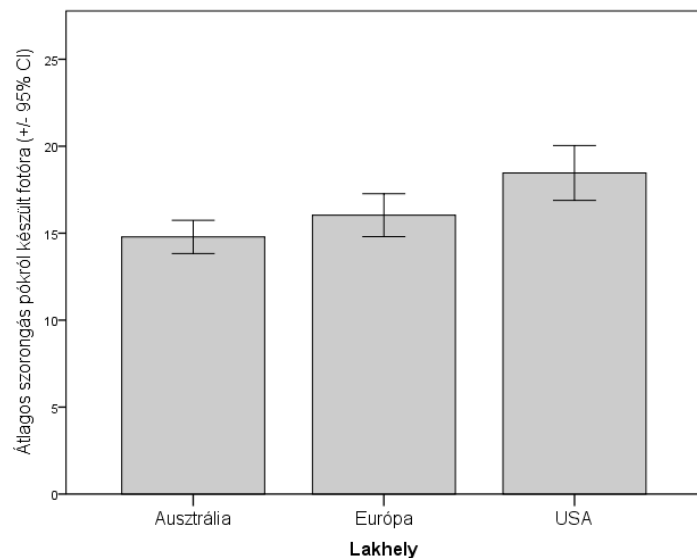
Utóvizsgálatot csak a lakhelyre kell végezni. A csoportra azért nem kell post hocot kikérni, mert csak 2 szintje van (ha van különbség, az a főhatásból is látszik)

Szóráshomogenitás teljesülése estén használandó post hoc tesztek

SPSS-ben többszemponos ANOVÁhoz nincs implementálva korrekció szóráshomogenitás sérülésére ☹

- **Post hoc vizsgálat**

- A **Post Hoc tábla** a lakhely különböző szintjeit hasonlítja össze (ebben össze van vonva a kontroll és klinikai) Tukey HSD korrekciót használva. Ez az egyetlen az utóvizsgálatok vizsgálatok közül, ami a súlyozott átlagokat használja
 - AUS és EU között a különbség 1.25 pont, és ez nem szignifikáns $p = .150$
 - AUS és USA között a különbség 3.68 pont, és ez szignifikáns $p < .001$, a grafikonról vagy a leíró statisztikákból le tudod olvasni, hogy USÁban nagyobb a szorongás a pókképre.
 - EU és USA között a különbség 2.42 pont, és ez szignifikáns $p = .003$, az USÁban nagyobb a szorongás a pókképre.
- A **Homogeneous Subsets** megadja, mely csoportok hasonlíthatnak egymásra, ez sokszor segít az eredmények megfogalmazásában. Itt AUS és EU alkot egy alcsoportot és USA egy másikat.



Post Hoc Tests

lakhely

Multiple Comparisons

Dependent Variable: Pókról készült fotó

Tukey HSD

(I) Lakhely	(J) Lakhely	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Ausztrália	Európa	-1,25	,669	,150	-2,84	,33
	USA	-3,68*	,714	,000	-5,37	-1,99
Európa	Ausztrália	1,25	,669	,150	-,33	2,84
	USA	-2,42*	,725	,003	-4,14	-,71
USA	Ausztrália	3,68*	,714	,000	1,99	5,37
	Európa	2,42*	,725	,003	,71	4,14

Based on observed means.

The error term is Mean Square(Error) = 12,053.

*. The mean difference is significant at the 0,05 level.

Homogeneous Subsets

Pókról készült fotó

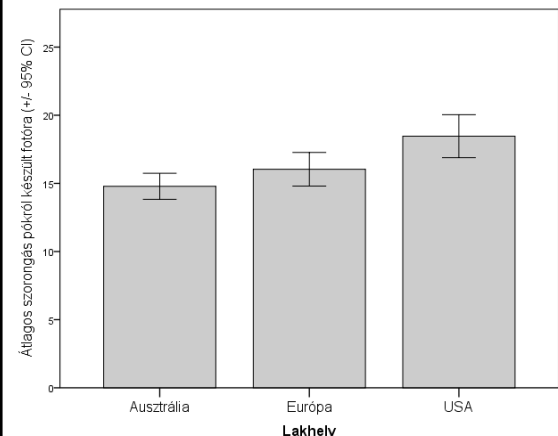
Tukey HSD^{a,b,c}

Lakhely	N	Subset	
		1	2
Ausztrália	56	14,79	18,46
Európa	52	16,04	
USA	41		
Sig.		,179	1,000

Means for groups in homogeneous subsets are displayed.

Based on observed means.

- **Kontraszt vizsgálat**
 - A kontraszt vizsgálathoz az Estimated Marginals Means átlagait használja, tehát a Post hoc-kal ellentétben súlyozatlan átlagokkal dolgozik)
 - Két táblát kapunk eredményül:
 - Az alsó tábla (Test Results) csak megismétli a szemponthoz tartozó főhatás vizsgálatát (megegyezik a fő ANOVA tábla ide vonatkozó részével)
 - A kontraszt tábla (Contrast Results) a lakhely különböző szintjeit hasonlítja össze (ebben össze van vonva a kontroll és klinikai).
 - Mivel Repeated kontrasztot kértünk, így az első szint lesz a másodikkal, a második a harmadikkal összehasonlítva.



AUS és EU között a különbség 1.449, és ez szignifikáns $p = .032$ (ennyivel érzékenyebb a kontraszt a Post hocnál)

EU és USA között a különbség 2.864, mely szignifikáns $p < .001$

Custom Hypothesis Tests

Contrast Results (K Matrix)

		Dependent Variable
		Pókról készült fotó
Lakhely Repeated Contrast		
Level 1 vs. Level 2	Contrast Estimate	-1,449
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-1,449
	Std. Error	,670
	Sig.	,032
	95% Confidence Interval for Difference	Lower Bound: -2,774 Upper Bound: -,123
Level 2 vs. Level 3	Contrast Estimate	-2,864
	Hypothesized Value	0
	Difference (Estimate - Hypothesized)	-2,864
	Std. Error	,728
	Sig.	,000
	95% Confidence Interval for Difference	Lower Bound: -4,303 Upper Bound: -1,424

Test Results

Dependent Variable: Pókról készült fotó

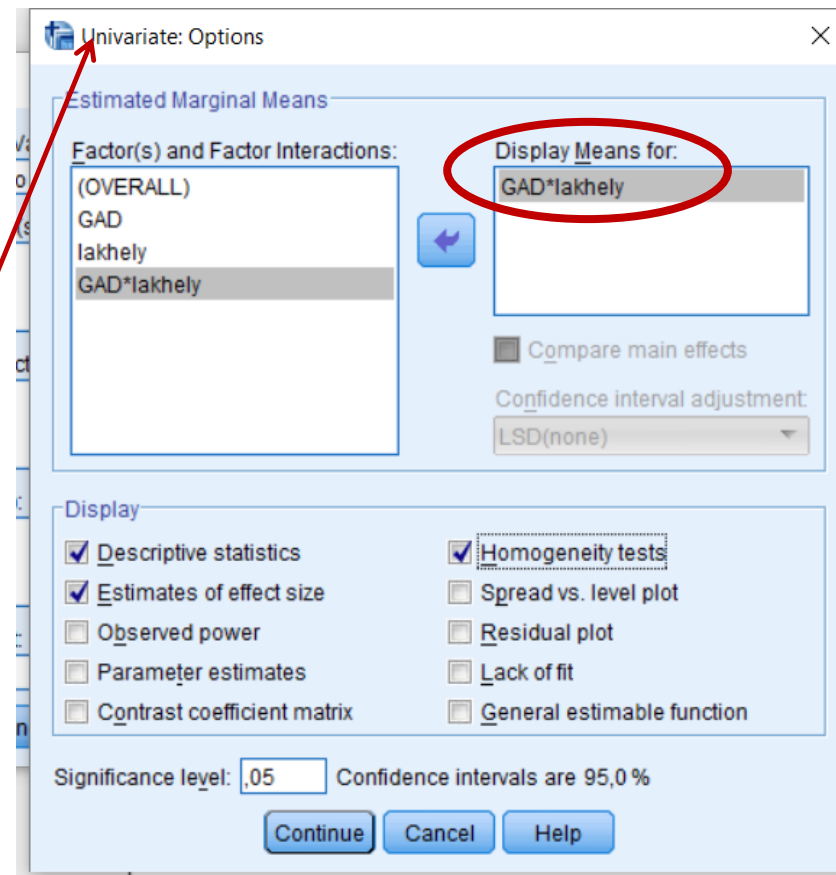
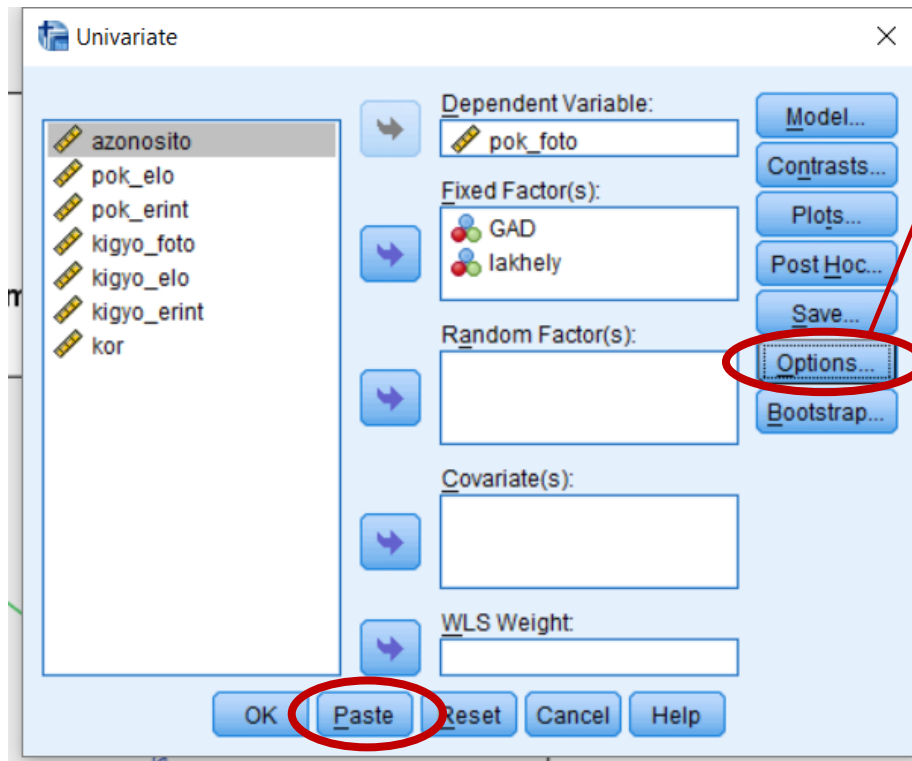
Source	Sum of Squares	df	Mean Square	F	Sig.
Contrast	438,914	2	219,457	18,207	,000
Error	1723,632	143	12,053		

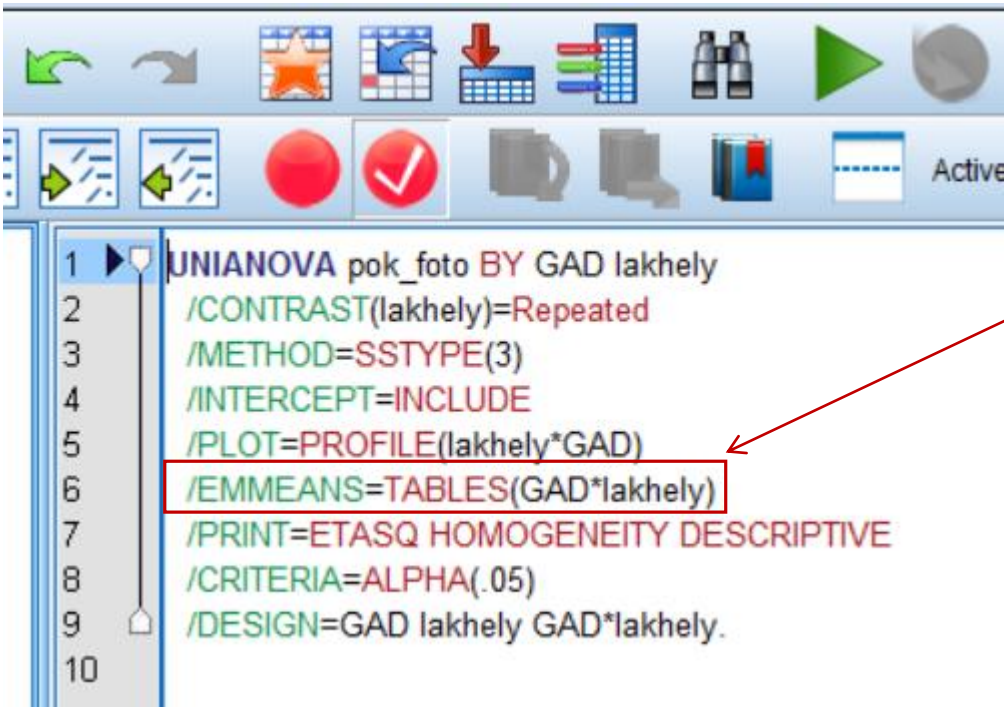
- **Simple effects tesztelése:**

- Az egyik szempont vizsgálata a másik szempont minden szintjén külön-külön
 - Pl. külön kontroll és klinikai esetében van-e a lakhelyek között különbség

- Kikérése:

- Options-ban az Estimated Marginal Means az interakciós tagra (GAD*lakhely) kell kikérni
- Az eljárás menübe nincs beépítve, ezért szintaxot használva kell kikérni -> OK helyett a **Paste**-ra kattintunk





The image shows the top toolbar of the SPSS Syntax Editor, which includes icons for undo, redo, save, print, and other standard functions. Below the toolbar is the Syntax Editor window, which contains a list of lines on the left and a text area on the right. The text area contains the following code:

```
1 UNIANOVA pok_foto BY GAD lakhely
2 /CONTRAST(lakhely)=Repeated
3 /METHOD=SSTYPE(3)
4 /INTERCEPT=INCLUDE
5 /PLOT=PROFILE(lakhely*GAD)
6 /EMMEANS=TABLES(GAD*lakhely)
7 /PRINT=ETASQ HOMOGENEITY DESCRIPTIVE
8 /CRITERIA=ALPHA(.05)
9 /DESIGN=GAD lakhely GAD*lakhely.
10
```

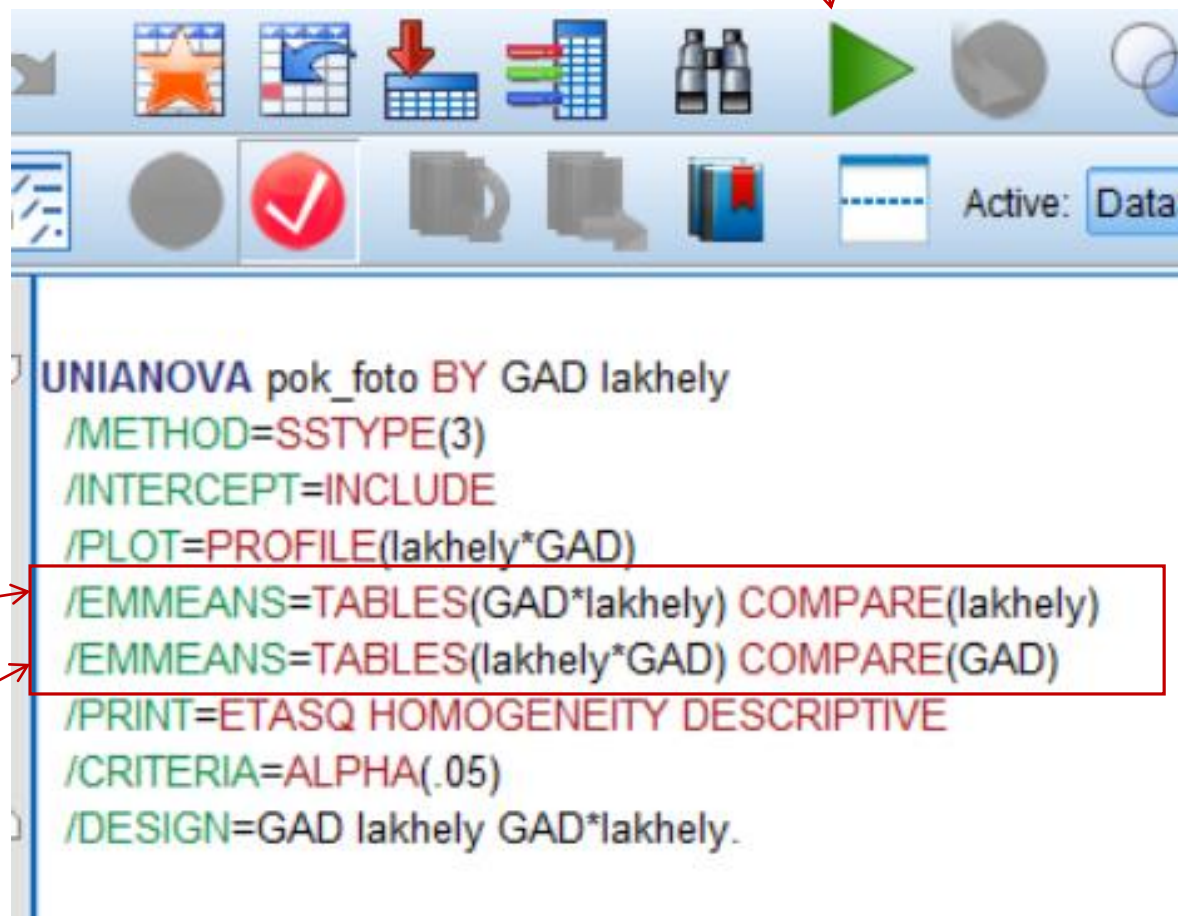
Ezt a kódrészt keresd meg!

Az összes kód kijelölése után
ezzel futtatható

Erre írd át!

Lakhelyek összehasonlítva a csoport
minden szintjén

Csoportok összehasonlítva a lakhely
minden szintjén



The image shows the top toolbar of the SPSS Syntax Editor, which includes icons for undo, redo, save, print, and other standard functions. Below the toolbar is the Syntax Editor window, which contains a list of lines on the left and a text area on the right. The text area contains the following code:

```
UNIANOVA pok_foto BY GAD lakhely
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PLOT=PROFILE(lakhely*GAD)
/EMMEANS=TABLES(GAD*lakhely) COMPARE(lakhely)
/EMMEANS=TABLES(lakhely*GAD) COMPARE(GAD)
/PRINT=ETASQ HOMOGENEITY DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=GAD lakhely GAD*lakhely.
```



Ez a dia a szintax megértését segíti, de nem kell végrehajtani. Beállítottam mindent, amit eddig kikértünk, és elmagyarázom, a szintax különböző sorai milyen parancsot tartalmaznak. (a diát egy régebbi verzióból tettem át itt még csoportnak neveztük a GAD változót)

Univariate ANOVA – az elemzés elnevezése

függő változó

független változók



A lakhely szempontra Repeated kontrasztot kértünk

Modelben állítható be a hiba becslésének módja, de mindig ezt használjuk

A regresszió számoljon b_0 -val is

A standardizált reziduálisokat és a Cook távolságot mentjük ki

A lakhelyre Tukey Post hocot kértünk

Marginal Means számítások – ezért nyúltunk bele a syntaxba

Kikértük az effect size-ot, a szóráshomogenitás tesztet és a leíró statisztikákat

Plotok Szignifikancia szintet $p < .05$ -re állítjuk

A modellben a két főhatást és a kettejük közötti interakciót vizsgáljuk

- **Simple effect tesztelése**

- A Simple effects tesztelés során azt történik, hogy a mintát szűrjük az egyik szempont egy szintjére, és ANOVÁt végzünk a másik szemponton, de ezt egyetlen GLM részeként valósítjuk meg.
- Ha a szempontok között ordinális interakciót találunk, akkor érdemes, ha diszordinális interakciót találunk, akkor pedig kötelező az utóvizsgálatot Simple effects teszteléssel megoldani.

- **Simple effects tesztelés lakhely szerint**

- **Mi fog történni:**

- 1. az elemzés leszűri a mintát a kontroll csoportra, és ANOVÁval ellenőrzi, hogy van-e különbség a lakhelyek között. Ha ez az ANOVA szignifikáns, akkor, Post hoc utóvizsgálattal ellenőrzi, mely kontinensek között van különbség
- 2. ezt követően leszűri a mintát a klinikai csoportra, és itt is végez egy ANOVÁt és szükség esetén Post hoc utóvizsgálatot a lakhelyre

- **3 táblát kapunk minden Simple effect tesztelés során:**

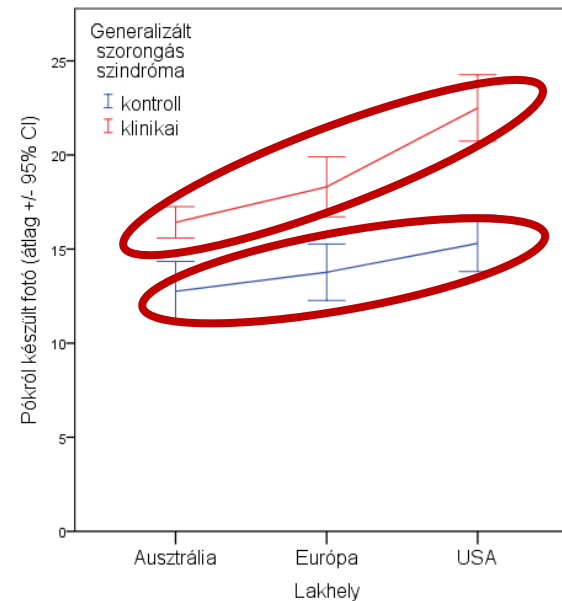
- **Leíró statisztikákat tartalmazó tábla (ez Estimated Marginal Means tábla)**

- **ANOVÁk tábláját**

- A táblázat két ANOVA vizsgálatot tartalmaz, egyet a kontroll, egyet a klinikai
- Mivel két ANOVÁt végeztünk, kontrollálni kell a Family wise hibára. Ezt megtehetjük Bonferroni vagy Holm procedúra segítségével

- **Post hoc utóvizsgálatok tábláját**

- Az ANOVA táblán végzett korrekció annyira szigorú, hogy a Post hoc táblában további korrekcióra nincs szükség (LSD korrekciót használ)



Estimated Marginal Means

1. Generalizált szorongás szindróma * Lakhely

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	Lakhely	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
kontroll	Ausztrália	12,760	,694	11,387	14,133
	Európa	13,769	,681	12,423	15,115
	USA	15,304	,724	13,873	16,735
klinikai	Ausztrália	16,419	,624	15,187	17,652
	Európa	18,308	,681	16,962	19,654
	USA	22,500	,818	20,882	24,118

Dependent Variable: Pókról készült fotó

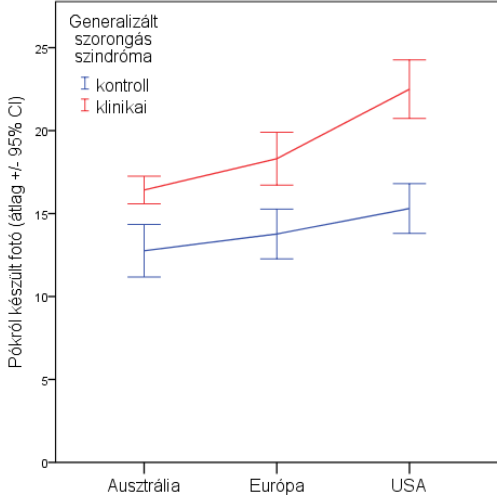
Generalizált szorongás szindróma	(I) Lakhely	(J) Lakhely	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
kontroll	Ausztrália	Európa	-1,009	,972	,301	-2,932	,913
		USA	-2,544 [*]	1,003	,012	-4,527	-,562
	Európa	Ausztrália	1,009	,972	,301	-,913	2,932
		USA	-1,535	,994	,125	-3,500	,429
	USA	Ausztrália	2,544 [*]	1,003	,012	,562	4,527
		Európa	1,535	,994	,125	-,429	3,500
klinikai	Ausztrália	Európa	-1,888 [*]	,923	,043	-3,713	-,063
		USA	-6,081 [*]	1,029	,000	-8,114	-4,047
	Európa	Ausztrália	1,888 [*]	,923	,043	,063	3,713
		USA	-4,192 [*]	1,065	,000	-6,297	-2,088
	USA	Ausztrália	6,081 [*]	1,029	,000	4,047	8,114
		Európa	4,192 [*]	1,065	,000	2,088	6,297

Based on estimated marginal means
*. The mean difference is significant at the ,05 level.
b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Dependent Variable: Pókról készült fotó

Univariate Tests		Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
kontroll	Contrast	78,293	2	39,146	3,248	,042	,043
	Error	1723,632	143	12,053			
klinikai	Contrast	423,080	2	211,540	17,550	,000	,197
	Error	1723,632	143	12,053			

Each F tests the simple effects of Lakhely within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.



Simple effect tesztelése **lakhely** szerint

ANOVA tábla:

- Az alul lévő Univariate Test tábla két ANOVA vizsgálatot tartalmaz, egyet a kontroll, egyet a klinikai személyekre szűrve.
- Mivel két ANOVÁt végeztünk, kontrollálni kell a **Family wise hibára**.
 - Ha **Bonferroni korrekciót** alkalmazunk, akkor a az új kritérium szint $\alpha = .025$, tehát a kontroll személyek esetén nincs szignifikáns különbség a kontinensek között $p = .042$, a klinikai személyek esetén viszont van $p < .001$
 - De ha **Holm procedúrát** alkalmazunk (mely optimálisabb), akkor kiderül, hogy a szignifikancia szint maradhat $\alpha < .05$, tehát mind a kontroll, mind a klinikai személyek esetén szignifikáns a kontinensek közötti különbség.

Estimated Marginal Means

1. Generalizált szorongás szindróma * Lakhely

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	Lakhely	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
kontroll	Ausztrália	12,760	,694	11,387	14,133
	Európa	13,769	,681	12,423	15,115
	USA	15,304	,724	13,873	16,735
klinikai	Ausztrália	16,419	,624	15,187	17,652
	Európa	18,308	,681	16,962	19,654
	USA	22,500	,818	20,882	24,118

Dependent Variable: Pókról készült fotó

Generalizált szorongás szindróma	(I) Lakhely	(J) Lakhely	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
kontroll	Ausztrália	Európa	-1,009	,972	,301	-2,932	,913
		USA	-2,544 [*]	1,003	,012	-4,527	-,562
	Európa	Ausztrália	1,009	,972	,301	-,913	2,932
		USA	-1,535	,994	,125	-3,500	,429
	USA	Ausztrália	2,544 [*]	1,003	,012	,562	4,527
		Európa	1,535	,994	,125	-,429	3,500
klinikai	Ausztrália	Európa	-1,888 [*]	,923	,043	-3,713	-,063
		USA	-6,081 [*]	1,029	,000	-8,114	-4,047
	Európa	Ausztrália	1,888 [*]	,923	,043	,063	3,713
		USA	-4,192 [*]	1,065	,000	-6,297	-2,088
	USA	Ausztrália	6,081 [*]	1,029	,000	4,047	8,114
		Európa	4,192 [*]	1,065	,000	2,088	6,297

Based on estimated marginal means

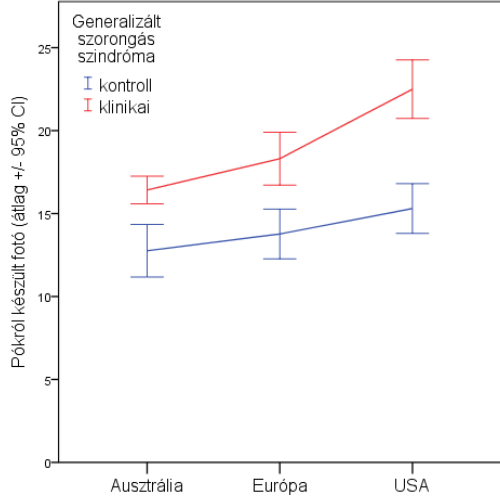
*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Dependent Variable: Pókról készült fotó

Univariate Tests		Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
kontroll	Contrast	78,293	2	39,146	3,248	,042	,043
	Error	1723,632	143	12,053			
klinikai	Contrast	423,080	2	211,540	17,550	,000	,197
	Error	1723,632	143	12,053			

Each F tests the simple effects of Lakhely within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.



Simple effect tesztelése **lakhely** szerint

Post hoc utóvizsgálat:

- Mely kontinensek különböznek egymástól?
- Az utóvizsgálatot a Pairwise Comparisons tábla tartalmazza.
- Az ANOVA táblán már megtörtént a korrekció, ezért a Post hoc táblában **további korrekcióra nincs szükség** (a tábla alatti szöveg említi is, hogy **LSD korrekció** történt, ami olyan, mintha nem lenne korrekció)
- Kontroll személyek: a Holm procedúra után szignifikáns az ANOVA, ezért értelmezhető az utóvizsgálat, AUS és EU között nincs szignifikáns különbség ($p = .301$), AUS és USA között szignifikáns ($p = .012$), EU és USA között nem szignifikáns ($p = .125$)
- Klinikai személyek: a Holm procedúra után szignifikáns az ANOVA, ezért értelmezhető az utóvizsgálat, AUS és EU között szignifikáns a különbség ($p = .043$), AUS és USA között szignifikáns a különbség ($p < .001$), EU és USA között szignifikáns a különbség ($p < .001$)

2. Lakhely * Generalizált szorongás szindróma

Estimates

Dependent Variable: Pókról készült fotó

Lakhely	Generalizált szorongás szindróma	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Ausztrália	kontroll	12,760	,694	11,387	14,133
	klinikai	16,419	,624	15,187	17,652
Európa	kontroll	13,769	,681	12,423	15,115
	klinikai	18,308	,681	16,962	19,654
USA	kontroll	15,304	,724	13,873	16,735
	klinikai	22,500	,818	20,882	24,118

Pairwise Comparisons

Dependent Variable: Pókról készült fotó

Lakhely	(I) Generalizált szorongás szindróma	(J) Generalizált szorongás szindróma	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
Ausztrália	kontroll	klinikai	-3,659 ^a	,933	,000	-5,504	-1,815
	klinikai	kontroll	3,659 ^a	,933	,000	1,815	5,504
Európa	kontroll	klinikai	-4,538 ^a	,963	,000	-6,442	-2,635
	klinikai	kontroll	4,538 ^a	,963	,000	2,635	6,442
USA	kontroll	klinikai	-7,196 ^a	1,093	,000	-9,355	-5,036
	klinikai	kontroll	7,196 ^a	1,093	,000	5,036	9,355

Based on estimated marginal means

*. The mean difference is significant at the ,05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Univariate Tests

Dependent Variable: Pókról készült fotó

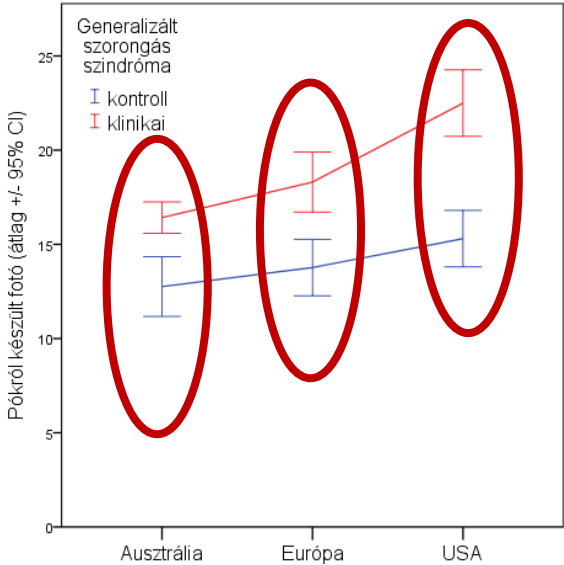
Lakhely		Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Ausztrália	Contrast	185,320	1	185,320	15,375	,000	,097
	Error	1723,632	143	12,053			
Európa	Contrast	267,769	1	267,769	22,215	,000	,134
	Error	1723,632	143	12,053			
USA	Contrast	522,826	1	522,826	43,376	,000	,233
	Error	1723,632	143	12,053			

Each F tests the simple effects of Generalizált szorongás szindróma within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

Simple effect tesztelése GAD szerint

ANOVA tábla:

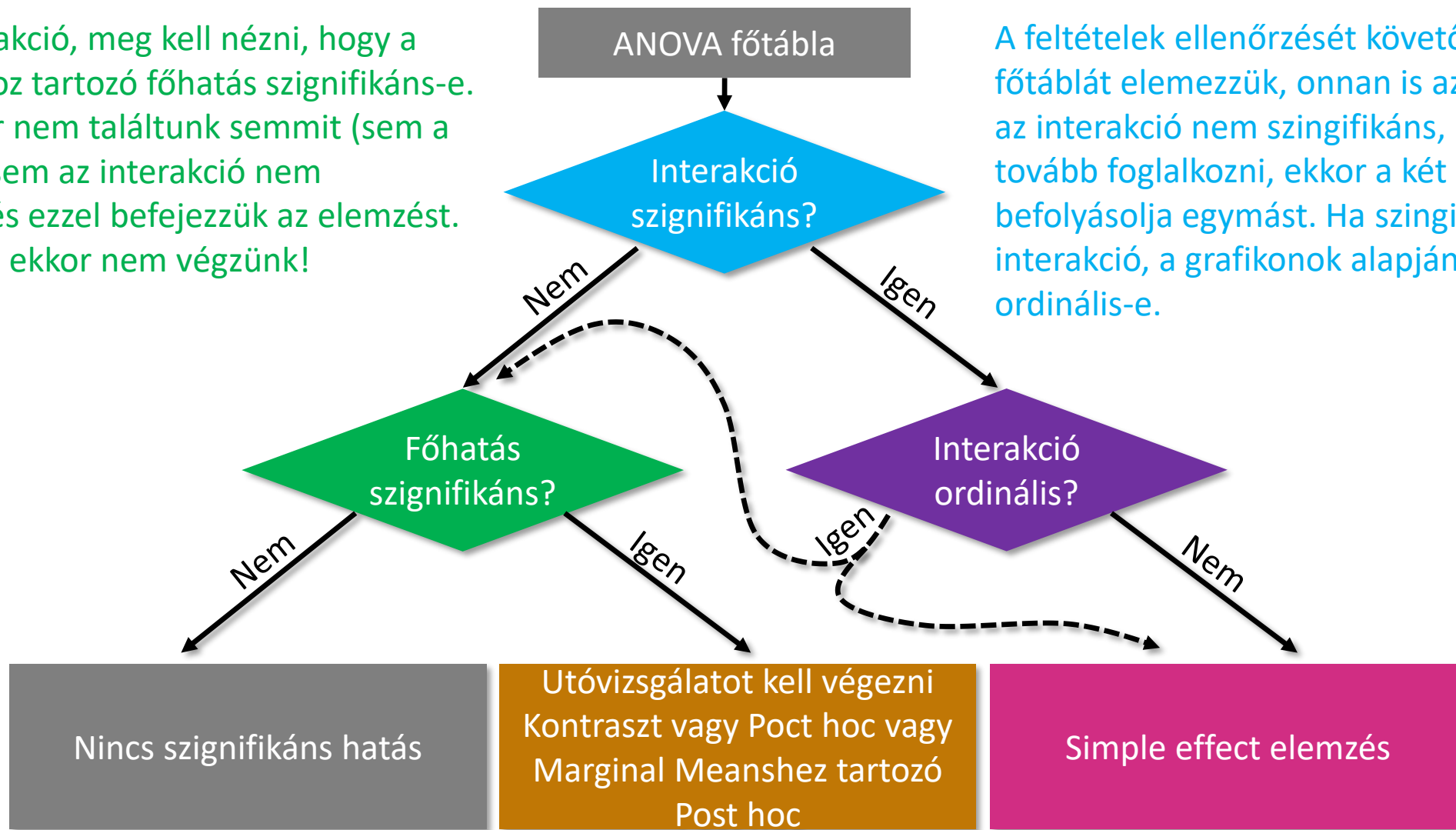
- Az Univariate Tests táblában látható, hogy három ANOVÁt végeztünk: megnéztük, hogy van-e különbség kontroll és klinikai csoport között AUS-ban, EU-ban és az USÁban.
- Ha Bonferroni korrekciót végzünk, az új szignifikancia szint $\alpha = .0166$, és már így is szignifikáns a különbség mind a három kontinensen. Ilyen esetben felesleges a Holm procedúrát használni (de ha mégis megtennéd, $\alpha = .05$ -öt írna elő)
- Mivel a GAD dichotóm változó, így ha az ANOVA szignifikáns, akkor tudod, hogy a különbség kontroll és klinikai között van, tehát nincs szükség a Post hoc tábla értelmezésére.
- Tipp: A grafikonra az SPSS nem tesz hibasávot, de az Estimatesben található átlag és konfidencia intervallum felhasználható a grafikon készítésére (pl. Excelben)



Összegezve, milyen táblák lehetnek?

- **Feltételeket ellenőrzését segítő táblák**
- **Leíró statisztika** táblákból kettő is van
 - Súlyozott átlagon alapuló, „hagyományos” deskriptív táblák
 - Estimated Marginal Means, súlyozatlan átlagon alapuló táblák (az ANOVA ezzel számol)
- **Fő ANOVA tábla**
 - A szempontok főhatásait és a közöttük lévő interakciót tartalmazza. **A főhatások diszordinális interakció esetén egyáltalán nem, ordinális esetén óvatosan értelmezhetők**
- **Kontraszt vizsgálat**
 - Főhatások egyfajta utóvizsgálata, a súlyozatlan átlagokat használja. **Diszordinális interakció esetén egyáltalán nem, ordinális interakció esetén óvatosan értelmezhető**
- **Post hoc vizsgálat**
 - Főhatások egyfajta utóvizsgálata, Post hoc vizsgálat, mely a súlyozott átlagot használja. **Diszordinális interakció esetén egyáltalán nem, ordinális interakció esetén óvatosan értelmezhető**
- **Estimated Marginal Means**
 - (Itt vannak a fent említett leíró statisztika táblák is)
 - **Simple effect táblák**
 - Egyik szempont vizsgálata a másik szempont minden szintjén külön
 - **Ha bármilyen interakció van, ezt kell értelmezni a főhatás és post hoc és kontraszt táblák helyett**

Ha nincs interakció, meg kell nézni, hogy a szempontokhoz tartozó főhatás szignifikáns-e. Ha nem, akkor nem találtunk semmit (sem a szempontok, sem az interakció nem szignifikáns), és ezzel befejezzük az elemzést. Utóvizsgálatot ekkor nem végzünk!



A feltételek ellenőrzését követően először a főttáblát elemezzük, onnan is az interakciót. Ha az interakció nem szignifikáns, nem kell vele tovább foglalkozni, ekkor a két szempont nem befolyásolja egymást. Ha szignifikáns, az interakció, a grafikonok alapján meg kell nézni, ordinális-e.

Ha egy főhatás szignifikáns, meg kell nézni, az adott szemponton belül mely csoportok között van különbség. Ez történhet Kontraszttal, Post hoc-kal vagy a Marginal Meanshez tartozó Post hoc-kal is.

Ha az interakció szignifikáns, meg kell nézni, hogy ordinális-e. Ha nem, akkor a főhatások (és ezzel együtt a Post hoc és kontraszt táblák) nem értelmezhetők, helyette a Simple effect tábla eredményét kell nézni. Ha ordinális az interakció, akkor dönteni lehet: a Simple effect tábla vagy a főhatások kerülnek elemzésre (esetleg mindkettő).