

stathelp.hu

Készítette: Soltész-Várhelyi Klára

Adattisztítás
Hibák és outlierok

Az órán használt adatbázishoz tartozó kérdőív

Jelige: _____

Szemek teszt pontszáma (a vizsgálat vezetője tölti ki): _____

Kérjük adja meg korát: _____

Nem:
férfi nő

Legmagasabb iskolai végzettség:
általános iskola
középiskola vagy gimnázium
főiskola, egyetem vagy magasabb

Családi állapot:
egyedülálló
párkapcsolatban élő
házas
özvegy
egyéb: _____

Facebook ismerősök száma: _____
Hányszor posztolt, kommentelt, like-olt a héten Facebookon: _____

Kérjük, tekintse át Facebook ismerőseit, és sorolja fel azon személyeket, akikre igaz a következő állítások valamelyike! Elég, ha a nevek Ön számára azonosíthatók. Olyan személyt is megemlíthet, aki nincs Facebook ismerősei között!

1. Merne tőle nagyobb összeget kérni. Elmondaná neki, ha súlyos betegsége lenne. Kérne tanácsot tőle egy súlyos probléma esetén:

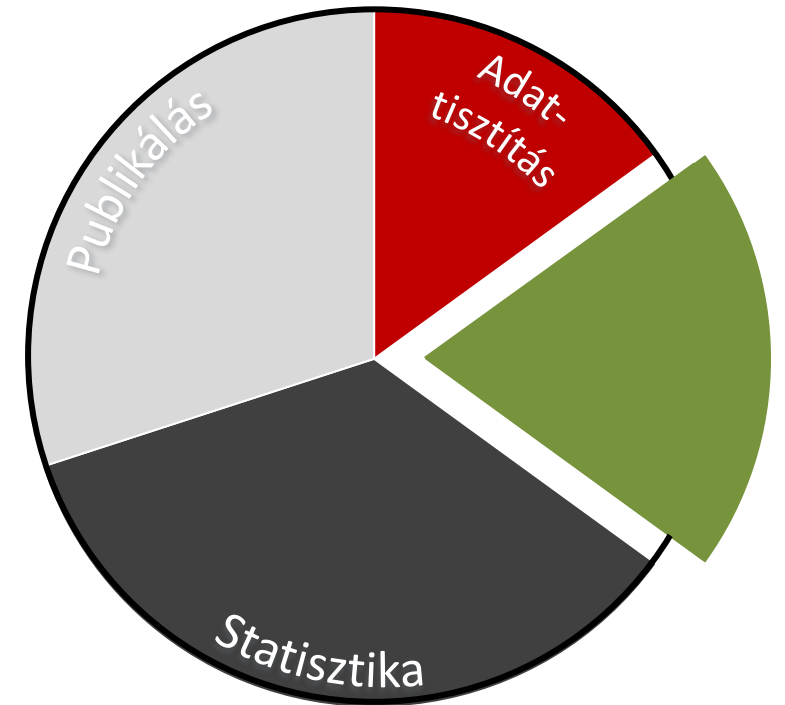
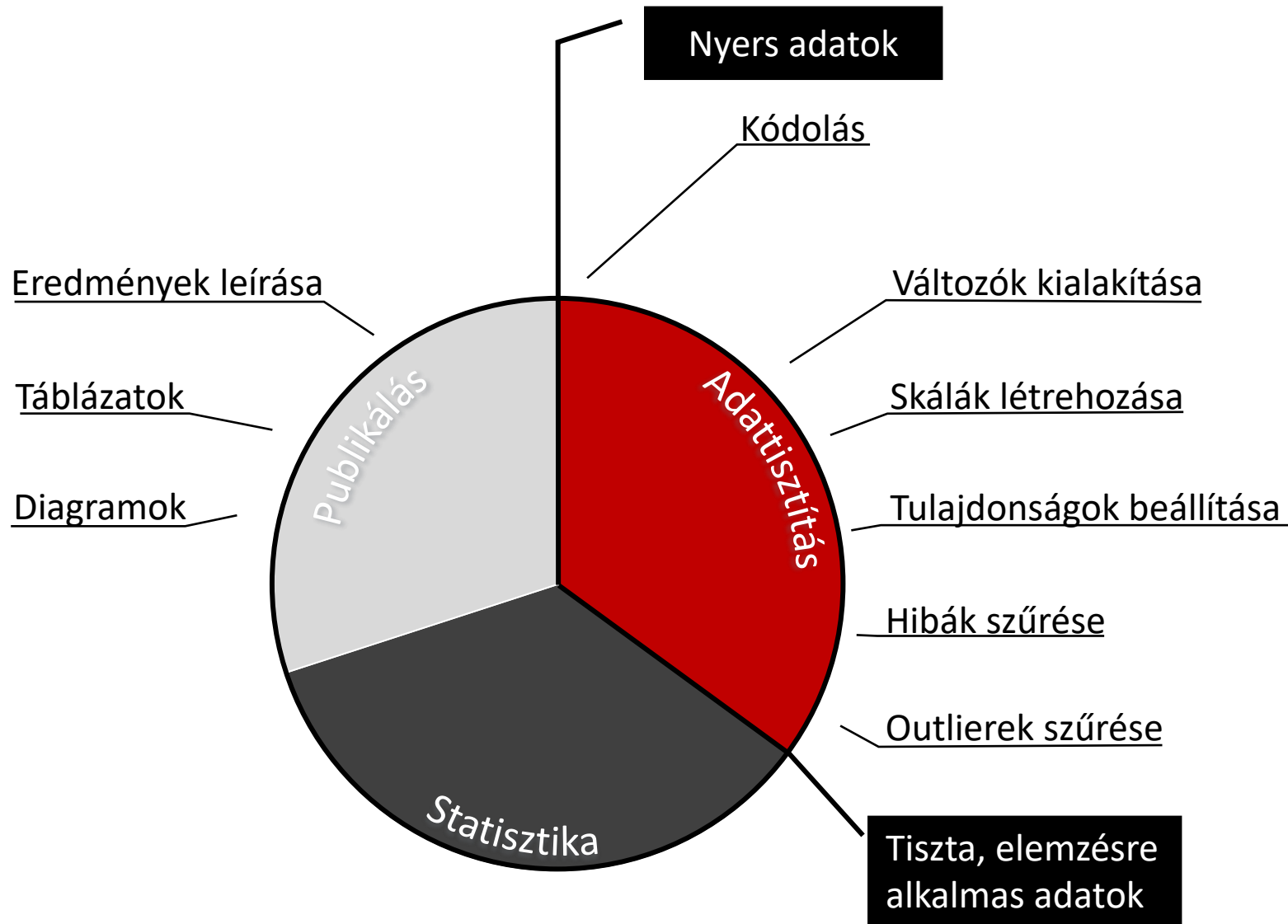
2. Legalább néhány hetente tartják a kapcsolatot, és életét viszonylag jól ismeri, illetve ők ismerik az ön életét:

Kérjük, karikázza be azt a válaszlehetőséget, mely legjobban leírja, hogyan érezte magát az elmúlt 2 hét során!

Az elmúlt két hét során érezte-e magát...	egyáltalán nem jellemző	alig jellemző	jellemző	teljesen jellemző
1. vidámnak és jókedvűnek?	0	1	2	3
2. nyugodtnak és ellazultnak?	0	1	2	3
3. aktívnak és élénknek?	0	1	2	3
4. ébredéskor frissnek és élénknek?	0	1	2	3
5. A napjai tele voltak számára érdekes dolgokkal?	0	1	2	3



Adattisztítás



Egyéni tapasztalataim alapján az adatok Google formról való letöltése és az első hipotézis tesztelése közötti munkák teszik ki általában egy statisztikai elemzés egyharmadát. Jól átgondolt kísérlettervezéssel ez az idő rövidíthető le.

Hibák & outlierek

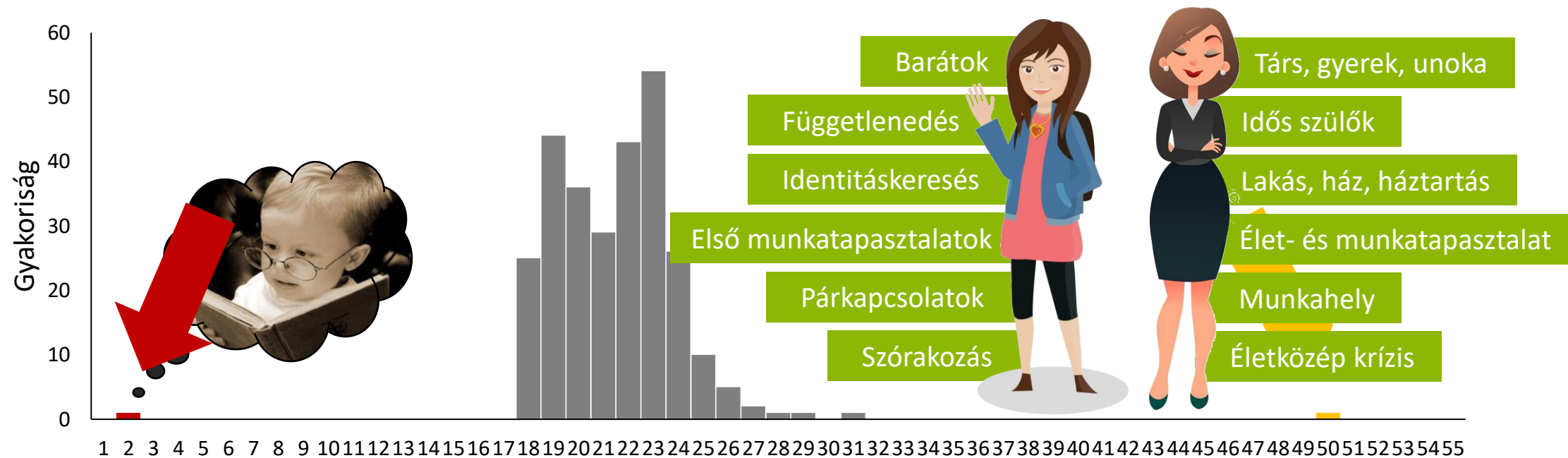


Fig. 1. A statisztika kurzus hallgatóinak életkori eloszlása

Hiba

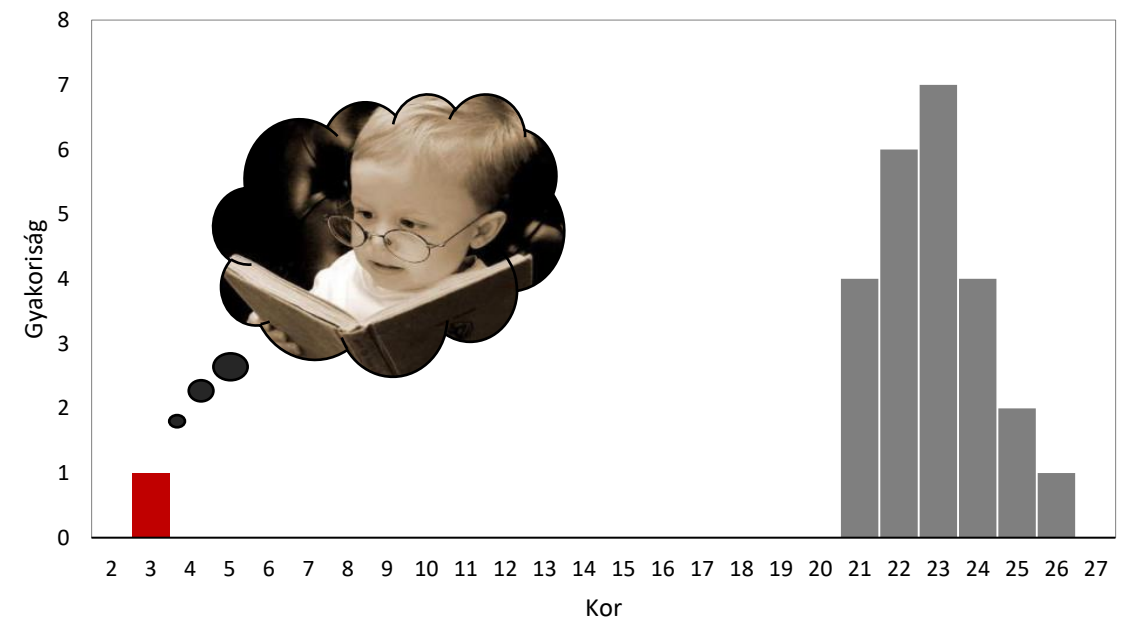
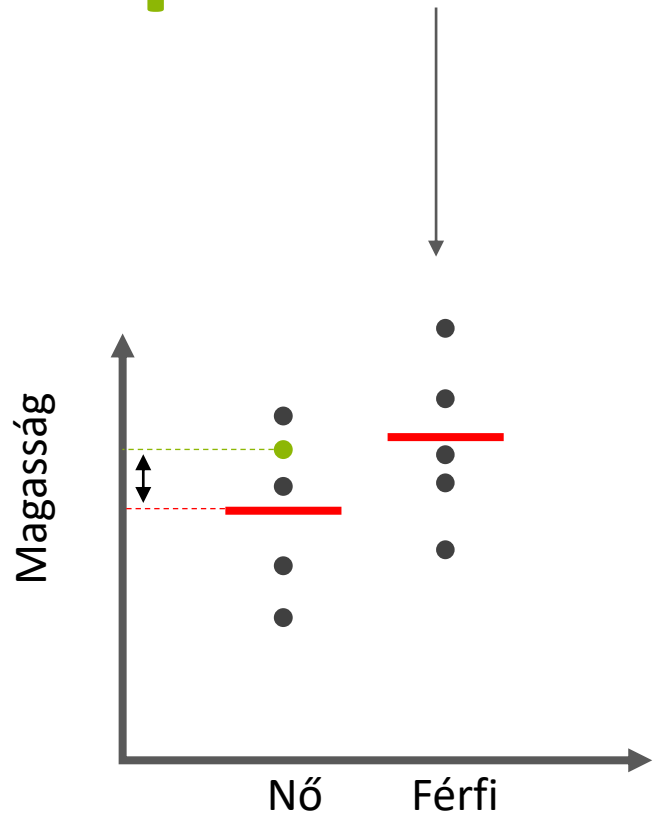


Fig. 1. A statisztika kurzus hallgatóinak életkori eloszlása

A hiba szó több értelemben is előfordul a kurzus során, de jelentésük teljesen más:

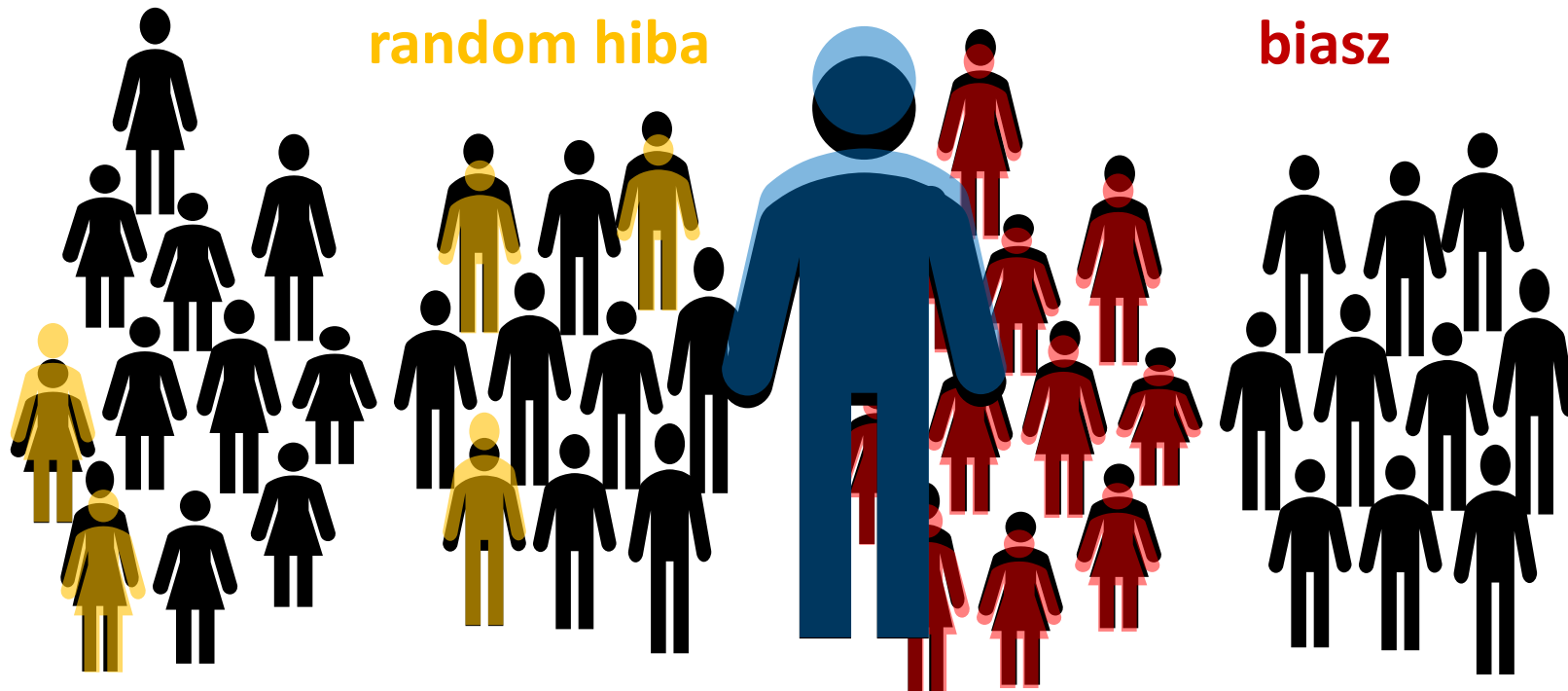
A predikció hibája



Adathiba / mérési hiba

Nem-szisztematikus,
random hiba

Szisztematikus
biasz



Predikció hibája

- az függő változó varianciájának azon része, amit a prediktorokkal nem tudunk magyarázni
- a zaj, a statisztikai érték számolásakor a nevező
- Ide tartozó fogalmak lesznek majd a SS_R (sum of squared residuals) és MS_R (mean squared residuals)

Mivel **többtenyezős rendszereket** vizsgálunk, és az összes tényezőt sosem vehetjük figyelembe, a modellnek egy jól megtervezett vizsgálatban is mindig lesz hibája.

- ettől statisztika, és nem matematika
- *Pl. a magasságot három tucat gén meg egy rák környezeti faktor befolyásolja, de mi csak a nem és a proteinfogyasztás hatását vizsgáljuk.*

Ezek a figyelembe nem vett tényezők **egyforma arányban fordulnak elő minden kísérleti mintánkban.**

- *Pl. egy reakcióidő vizsgálatban egyformán lesznek éberem és fáradtan érkezők minden kísérleti csoportunkban.*

A predikció hibája természetes része a statisztikának, de módszertanilag lehet (érdeemes) törekedni a csökkentésére, mert minél kisebb a zaj, annál kisebb hatások is kimutathatóak lesznek.

Szisztematikus mérési hiba / adathiba

Olyan hiba, ami **nem véletlen irányokban befolyásolja a mintáinkat**, így torzítja a mérést.

Több forrásból is származhat, leggyakoribbak:

- Forrása a mérőeszköz:
 - *Ha meg volt nyúlva a mérőszalag, és az egyik csoportot 2 centivel rövidebbnek mérted.*
 - *Szofisztikáltan megfogalmazott kérdőív alacsony iskolázottságúaknál (asszertív)*
- Forrása a környezet:
 - *Férfiak és nők teljesítménye közötti különbség & a szexi kísérletvezető nő.*
 - *Egyik csoport megszokott helyen, másik csoport idegen helyen.*
 - *Memóriateljesítmény – reggeli és esti mérés.*
- Forrása a mintaválasztás:
 - *Pszichológus-hallgató nők és mérnök-hallgató férfiak matematika tudása.*
 - *Pszichológus hallgató férfiak és nők empátiája.*

Törekedni kell arra, hogy **szisztematikus hiba ne legyen** az elemzésben.

- A vizsgálat megtervezésekor, minta kiválasztásakor & mérések lebonyolításakor ne emeljünk be szisztematikus hibát!
- Adatok áttekintésekor, tisztításakor figyelni kell a szisztematikus hibára utaló jelekre.

Nem-szisztematikus (random) mérési hiba / adathiba

Olyan hiba, ami **véletlen irányokban befolyásolja a mintáinkat.**

Gépeléskor félreütsz, és 2-es helyett 3-ast viszel be, de félregépelés egyforma arányban fordul elő minden kísérleti csoportokban.

Törekedni kell arra, hogy **random hiba alacsony legyen** az elemzésben, de teljesen úgyse lehet kizárni.

Adathiba, mérési hiba

Téves adatok. Az adathiba lehetnek szisztematikus és nem szisztematikus is.

Ha valaki 165 cm magas, de az adatok bevitelekor véletlenül 175-öt ütöttél be.

Ha meg volt nyúlva a mérőszalag, és mindenkit 2 centivel rövidebbnek mértél.

Adathiba sok forrásból származhat:

- **Félregépelések adatbevitelkor**
 - Félregépelés, tizedesvessző rossz helyen
- **(Számítási) hibák kiértékeléskor**
 - Elszámolások „fejben” számoláskor
 - Rosszul megírt kiértékelő, rossz megoldó kulcs
- **Rosszindulatú / komolytalan kitöltők**
 - Beleun, végére akar érni, és csupa 1-est nyom vagy mintázat szerint tölti ki
 - Visszajelzést adsz, ezért megpróbálja egy „alternatív személyisége” szerint is kitölteni
- Stb. ...

Félregépelések adatbevitelkor

Gyakori forrása:

- Félregépelés
- Billentyű dupla lenyomása
- Tizedesvessző rossz helyen

Hogy találd meg?

- Leíró statisztika minimumra és maximumra – ha 175 helyett 275 cm-t írtál, észre fogod venni.
- Ha tartományon belül van, nem igazán megtalálható, de a félregépelés viszonylag ritka, és elszórt, ezért a nagyon szélsőséges értékeket kivéve nem okoz nagy torzítást.

Példák:

- Kötődés kérdőív – két 44-es érték a négyfokú skálán – egész főkomponenselemzés és CFA újra
- Demográfiai adatok – szöveges infók a korban – gyanúsan sok hiányzó érték
- Vizuális kísérlet – Program rosszul mentette ki az eredményeket – gyanúsan kicsi szórások
- Utánkövetéses vizsgálat Szomáliában – összekevert azonosítók – eltérő életkorok, nevek
- Elcsúszó oszlopok – egy változónál gyanúsan sok hiányzó érték – ráadásul egymás alatt

Rosszindulatú / komolytalan kitöltők

Gyakori forrása:

- Hosszú a kérdőív, és elunja
- Kitöltés jutalomért, pénzért, csak „túl akar lenni rajta”
- „Humoros” kitöltők

Hogy találd meg?

- Komolytalan kitöltőkre utalhat, ha a válaszainak kicsi a szórása – ezt Excelben érdemes leellenőrizni – főleg internetes merítés esetén.
- Ha egy „humoros” válasszal találkozol, érdemes az adott személy összes adatát átnézni.
- Az „alternatív személyiség” ellen lehet védekezni, ha a kérdőíved végén megkérdezed, hogy ezek valós válaszok-e.

Példák:

- Arcok kódolása – komolytalan kitöltők – az adatokat átnézése a válaszok szórására
- Kulturális adaptáció skála – komolytalan kitöltők – kitöltési idők rövidek
- Egyedülállók a párkapcsolati kérdőívben

(Számítási) hibák kiértékeléskor

Gyakori forrása:

- Félrekódolt fordított itemek, illetve átkódolások
- Félreszámolt skálapontszámok

Hogyan előzd meg / találd meg?

- Ne számolj semmit fejben!
- Ellenőrizd a skála minimumát és maximumát, esetleg hozz létre tesztkitöltőket!
- Ha SPSS-ben dolgozol, a reliabilitás vizsgálat szintaxából másold át a változók listáját a változószámoláshoz!
- A hibás megoldó kulcs gyakoribb, mint gondolnád, fúsd át az itemeket!
- **Példák:**
 - Rossz automatikus kiértékelő – A skála minimuma kevesebb volt, mint kellene
 - Kézzel számolt skálapontszámok – Eltért az Excellel számolt értékektől
 - Rossz megoldó kulcs – Alacsony reliabilitás mutató

Jelige: sanci

Szemek teszt pontszáma (a vizsgálat vezetője tölti ki): 36

Kérjük adja meg korát: _____

Nem: férfi nő

Legmagasabb iskolai végzettség:
általános iskola
középiskola vagy gimnázium
főiskola, egyetem vagy magasabb

Családi állapot:
egyedülálló
párkapcsolatban élő
házas
özvegy
egyéb: _____

Facebook ismerősök száma: 1
Hányszor posztolt, kommentelt, like-olt a hétent Facebookon: 0

Kérjük, tekintse át Facebook ismerőseit, és sorolja fel azon személyeket, akikre igaz a következő állítások valamelyike! Elég, ha a nevek Ön számára azonosíthatók. Olyan személyt is megemlíthet, aki nincs Facebook ismerősei között!

1. Merne tőle nagyobb összeget kérni. Elmondaná neki, ha súlyos betegsége lenne. Kérne tanácsot tőle egy súlyos probléma esetén:

Párom, fiam

2. Legalább néhány hetente tartják a kapcsolatot, és életét viszonylag jól ismeri, illetve ők ismerik az ön életét:

Párom, fiam, kollégák (G.P., A.L., SZ-V.G., Z.O., A.T., A.G.) és a focista társak (Zoli, Robi, Laci, Laci2, Andris, Gáspár, János), Eszti, Zoltán

Kérjük, karikázza be azt a válaszlehetőséget, mely legjobban leírja, hogyan érezte magát az elmúlt 2 hét során!

Az elmúlt két hét során érezte-e magát...	egyáltalán nem jellemző	alig jellemző	jellemző	teljesen jellemző
1. vidámnak és jókedvűnek?	0	1	<u>2</u>	3
2. nyugodtnak és ellazultnak?	0	1	2	<u>3</u>
3. aktívnak és élénknek?	0	1	2	<u>3</u>
4. ébredéskor frissnek és élénknek?	0	1	2	<u>3</u>
5. A napjai tele voltak számára érdekes dolgokkal?	0	1	2	<u>3</u>

- Mit kezdjünk a hibás értékekkel?
 - Ha vissza tudjuk keresni a helyes adatot (pl. papírról), akkor javíthatjuk, minden más esetben **TÖRÖLNI KELL**.
 - Nézzünk a 27. sorban lévő „piramis” jeligét! Az adatbázisban hozzá 126 került, az adatlapról viszont látszik, hogy egyszerűen elgépelünk. Az ő **értékét javíthatjuk** az adatbázisban is 26-ra.
 - Nézzük a 38. sorban lévő „noname” jeligét! Látható, hogy nem elgépelésről van szó, a kitöltő maga írt 666-ot a korához. Ilyenkor érdemes megnézni, hogy a többi adata validnak tűnik-e, és ha nem, **az egész sort törölni** érdemes.
 - Ha nem tudjuk ehhez hasonlóan visszakeresni az adatlapot, vagy nincs okunk feltételezni, hogy a többi adata hamis lenne, akkor **nem kell az egész sort, elég csak a hibás értéket törölni**.

Jelige: PIRAMIS

Szemek teszt pontszáma (a vizsgálat nevezője tölti ki): 36

Kérjük adja meg korát: 26

Nem: férfi
(nő)

Legmagasabb iskolai végzettség: általános iskola
középiskola vagy gimnázium
főiskola, egyetem vagy magasabb

Családi állapot: egyedülálló
párkapcsolatban élő
házas
özvegy
egyéb:

Facebook ismerősök száma: 756

Kérjük, tekintse át Facebook ismerőseit, és sorolja fel azon személyeket, akikre igaz a következő állítások valamelyike! Elég, ha a nevek Ön számára azonosíthatók. Olyan személyt is megemlíthet, aki nincs Facebook ismerősei között!

1. Merne tőle nagyobb összeget kérni. Elmondaná neki, ha súlyos betegsége lenne. Kérne tanácsot tőle egy súlyos probléma esetén:
ANYU, APU, JULI, PETI

2. Legalább néhány hetente tartják a kapcsolatot, és életét viszonylag jól ismeri, illetve ők ismerik az ön életét:
ANYU, APU, JULI, PETI, ANNA, LACIKA, ANDRIS, BALU, BENCE, PISTA, ISTI, CSENGE, ESZTER, ROBI, ZSOLTI, FERI, GÖLYŐ, K. ANDRIS, KRISTI, F. ZOLI, DANÓ

Kérjük, karikázza be azt a válaszlehetőséget, mely legjobban leírja, hogyan érezte magát az elmúlt 2 hét során!

Az elmúlt két hét során érezte-e magát...	egyáltalán nem jellemző	alig jellemző	jellemző	teljesen jellemző
1. vidámnak és jókedvűnek?	0	1	☒	3
2. nyugodtnak és ellazultnak?	0	1	☒	3
3. aktívnak és élénknek?	0	1	2	☒
4. ébredéskor frissnek és élénknek?	0	1	☒	3
5. A napjai tele voltak számára érdekes dolgokkal?	0	☒	2	3

Jelige: noname

Szemek teszt pontszáma (a vizsgálat nevezője tölti ki): 36

Kérjük adja meg korát: 666

Nem: férfi
nő

Legmagasabb iskolai végzettség: általános iskola
középiskola vagy gimnázium
főiskola, egyetem vagy magasabb

Családi állapot: egyedülálló
párkapcsolatban élő
házas
özvegy
egyéb:

Facebook ismerősök száma: 69

Kérjük, tekintse át Facebook ismerőseit, és sorolja fel azon személyeket, akikre igaz a következő állítások valamelyike! Elég, ha a nevek Ön számára azonosíthatók. Olyan személyt is megemlíthet, aki nincs Facebook ismerősei között!

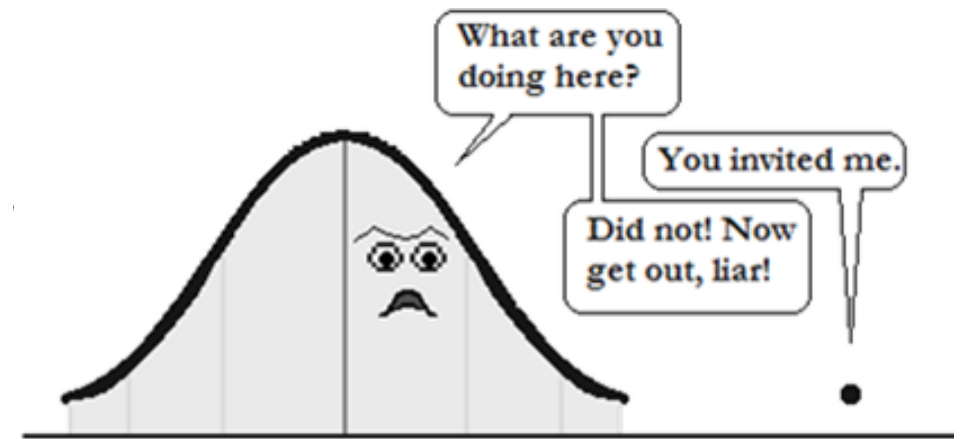
1. Merne tőle nagyobb összeget kérni. Elmondaná neki, ha súlyos betegsége lenne. Kérne tanácsot tőle egy súlyos probléma esetén:
Andi, Steve, Kata, Kótya, Laci, Feri

2. Legalább néhány hetente tartják a kapcsolatot, és életét viszonylag jól ismeri, illetve ők ismerik az ön életét:
Andi, Steve, Kata, Kótya, Laci, Feri, anyám, apám, nagyapám

Kérjük, karikázza be azt a válaszlehetőséget, mely legjobban leírja, hogyan érezte magát az elmúlt 2 hét során!

Az elmúlt két hét során érezte-e magát...	egyáltalán nem jellemző	alig jellemző	jellemző	teljesen jellemző
1. vidámnak és jókedvűnek?	0	1	2	☒
2. nyugodtnak és ellazultnak?	0	1	2	☒
3. aktívnak és élénknek?	0	1	2	☒
4. ébredéskor frissnek és élénknek?	0	1	2	☒
5. A napjai tele voltak számára érdekes dolgokkal?	0	1	2	☒

Outlier

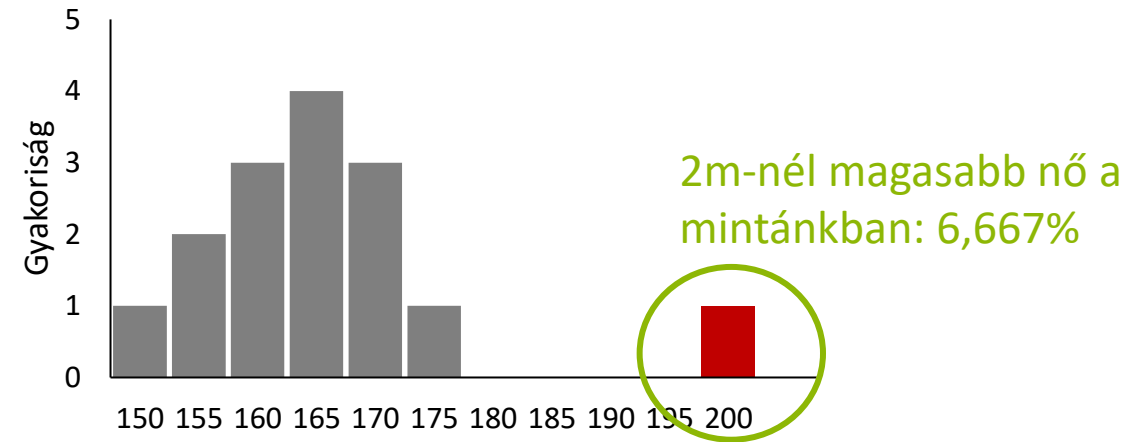
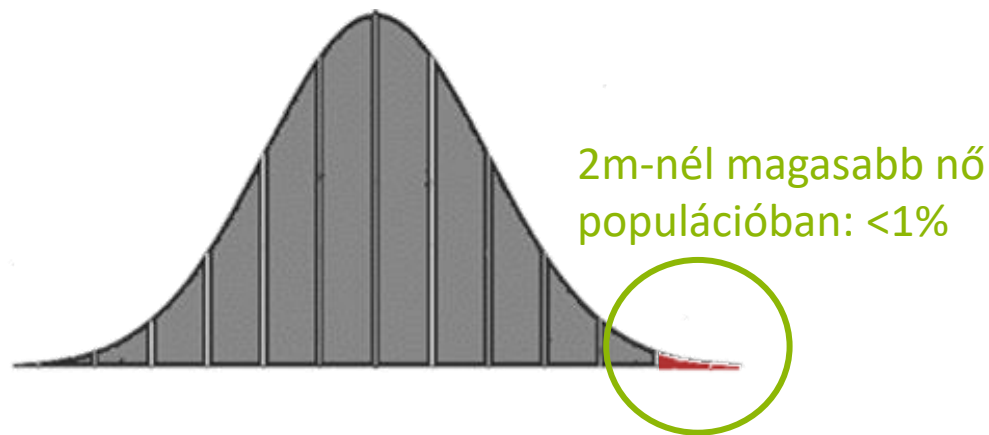


Outlier definíciója

Általános definíció: outlier az, „ami valid érték, de nagyon eltér a többi adatunktól”

Miért probléma:

Bár valid érték, de nem jól reprezentálja a populációt – ez főleg kis minták esetén probléma (egy 15 fős mintában van egy 2m magas nő – a valóságban nincs a nők közel 7%-a 2m körül)



Torzítja a statisztikai mutatókat: elhúzza az átlagot, megnöveli a szórást.

Skála típusú változók (nagyon ritkán ordinális változók) esetén van értelme outlierről beszélni, nominális változóban nem lehet outlier!

Hogy találjuk meg őket:

- **Outlier labelling rule**
- Olvasni lehet néha az „átlag $\pm$ 2 szórás” szabályról is, de az elavult, ne használjátok!

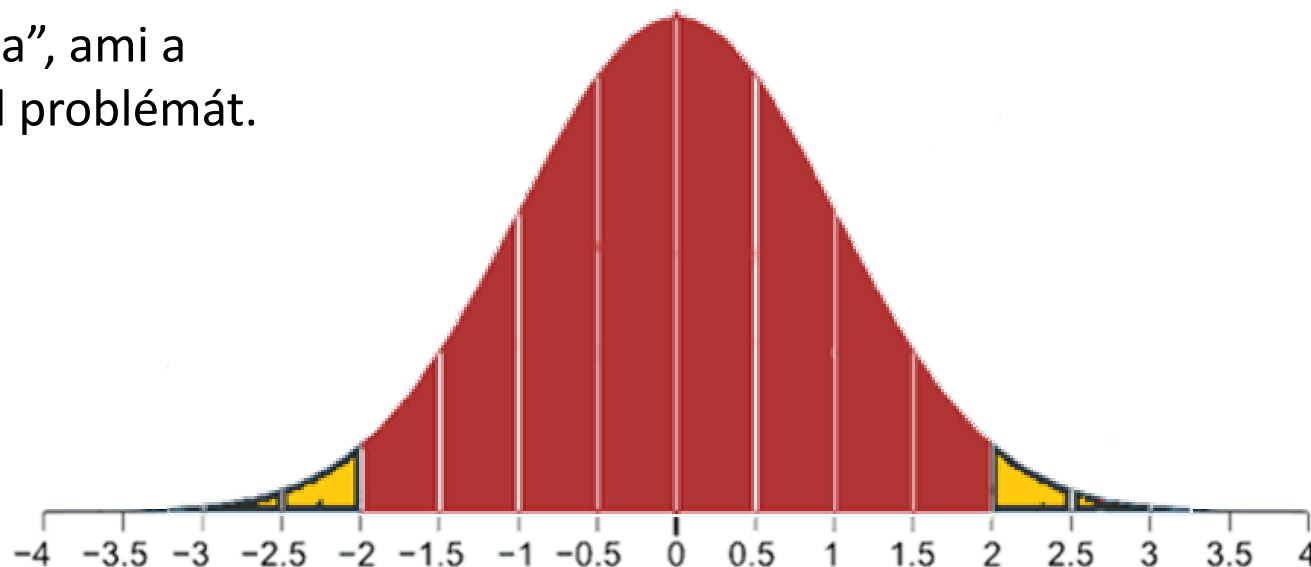
Átlag $\pm$ 2 vagy 3 SD szabály

ELAVULT, NE HASZNÁLD!!

Tekintsük outliernek azt az elemet, aki az átlagtól 2 (vagy 3) szórányira van!

Probléma:

- A szabály leginkább csak normálosztás esetén értelmezhető
- A ± 2 -es z-értékhez (a 2SD-hez) tartozó valószínűség 2,275% - tehát definíció szerint a minta majdnem 5%-a 2SD kívül fog elhelyezkedni.
- Ördögi kör – ha kiveszem a szélsőséges értékeket, csökken a szórás, akkor újabb értékek kerülnek a 2SD-n kívülre...
- Az elosztás görbét úgymond „csonkolja”, ami a normalitás szempontjából jelent majd problémát.



Outlier labelling rule

Outlier labelling rule

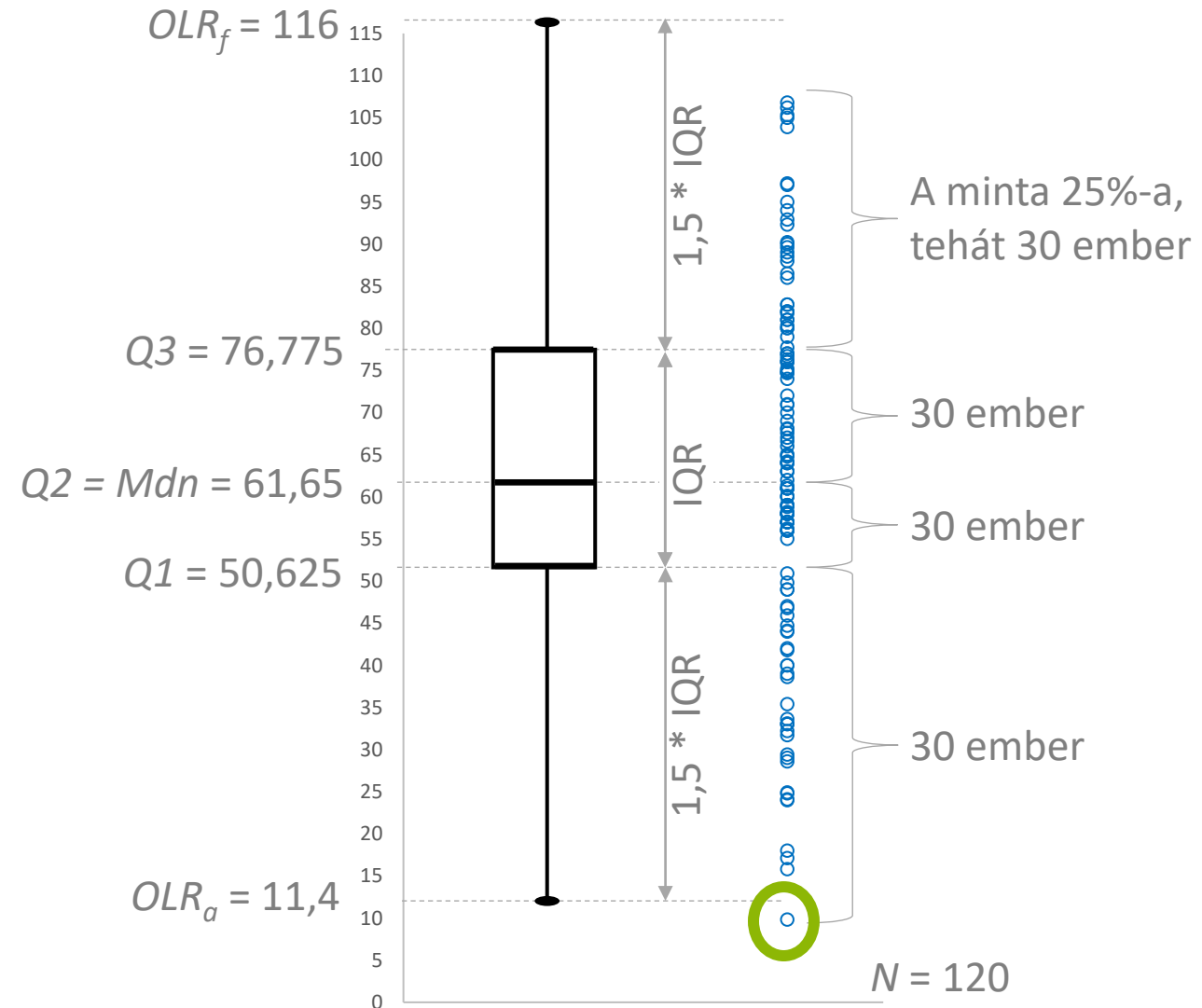
Outlier az, ami távolabb az alsó/felső negyedétől, mint a középső 50% nagyságának másfélszerese

Kvartiliseknek azokat az értékeket nevezzük, melyek a mintát pont 4 egyforma részre bontják.

- Q1 – alsó kvartilis alatt van a minta egynegyede, felette háromnegyede
- Q2 – a medián, azaz pont félbeveszi a mintát
- Q3 – felső kvartilis alatt van a minta háromnegyede, felette egynegyede.
- IQR – interkvartilis terjedelem, a Q1 és Q3 közötti rész, azaz a minta középső 50%-a

Azaz outliernek tekintjük azt, aki magasabb illetve alacsonyabb értéket vesz fel, mint a következő határok.

- **Felső határ:** $Q3 + 1,5 * IQR$
- **Alsó határ:** $Q1 - 1,5 * IQR$



Outlier



Eredeti, (SPSS-ben is használt) határok:

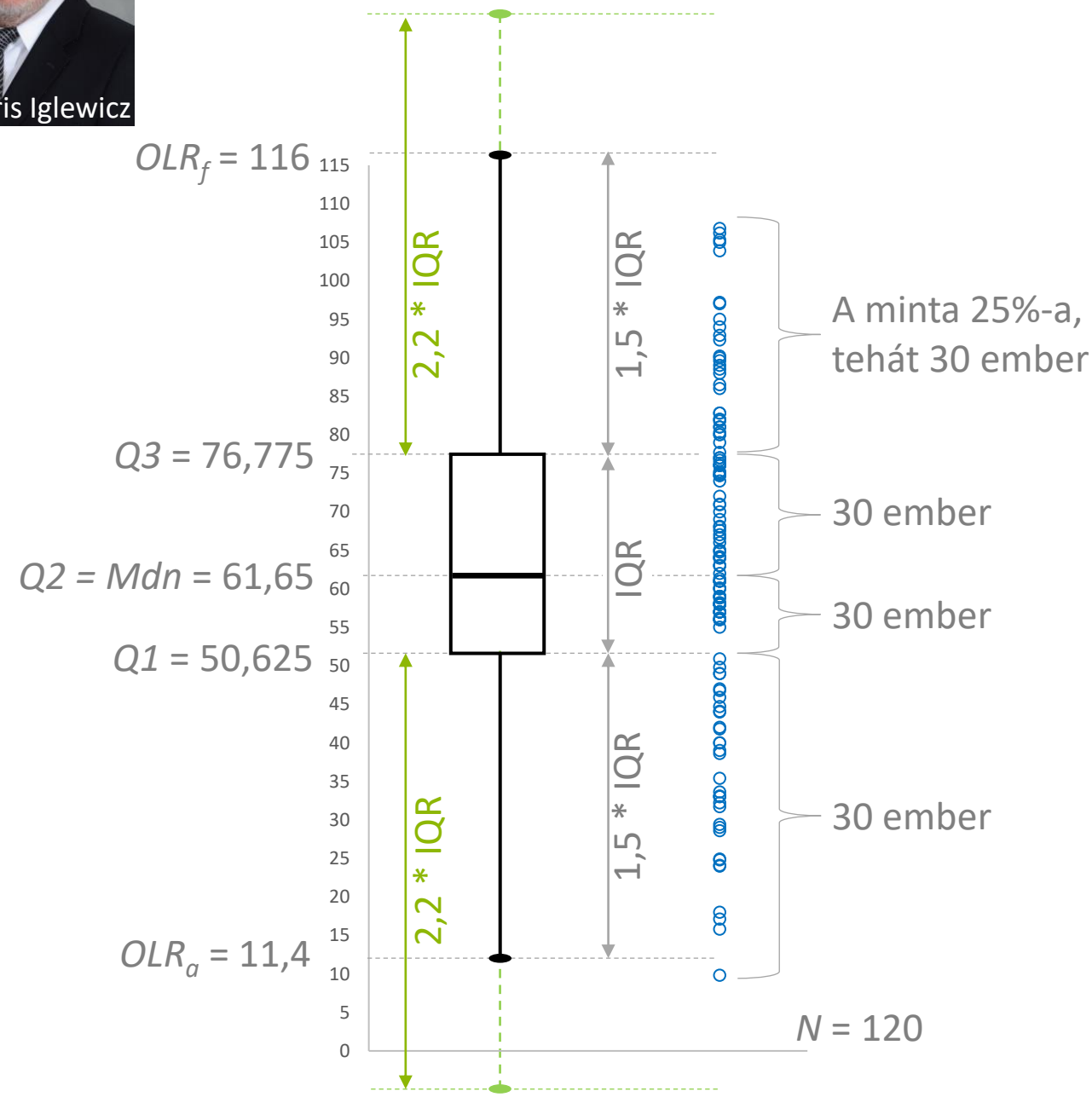
- $Q1 - 1,5 * IQR$ és $Q3 + 1,5 * IQR$

Módosított határok

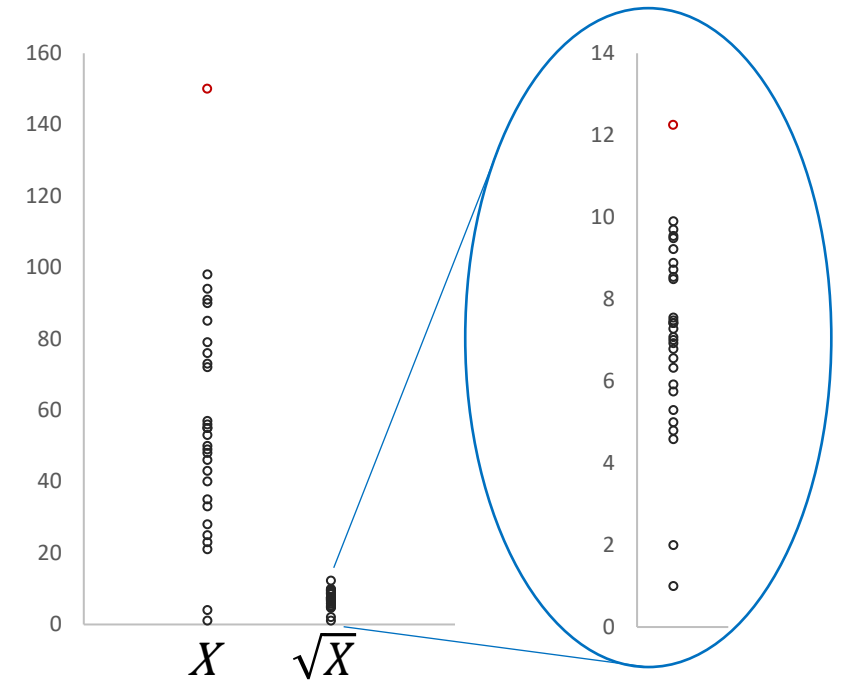
- $Q1 - 2,2 * IQR$ és $Q3 + 2,2 * IQR$

Forrás:

- Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley. – először veti fel az ötletet.
- Hoaglin, D.C., Iglewicz, B., and Tukey, J.W. (1986). Performance of some resistant rules for outlier labeling, *Journal of American Statistical Association*, 81, 991-999.
- Hoaglin, D. C., and Iglewicz, B. (1987), Fine tuning some resistant rules for outlier labeling, *Journal of American Statistical Association*, 82, 1147-1149. – felveti a 2,2-es szorzó használatát

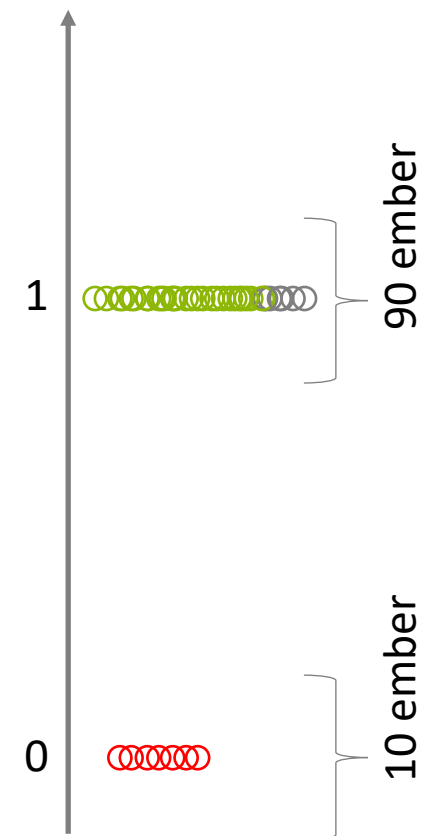


- Mit kezdünk velük?
 - **-1. Hiba vagy outlier?**
 - Ha hiba, akkor nincs helye a mérlegelésnek, mennie kell
 - **0. Elemezzük a helyzetet**
 - Pl. nem normális eloszlásnál mit jelentenek?
 - 2,2-es szorzóval is outliereknek tekinthetők?
 - **A. Töröljük őket**
 - Ésszel!!
 - PRO: torzító hatás megszűnik KON: információt veszünk
 - Leggyakrabban használt, kis minta esetén bajos
 - Gondoljuk át, miből jöhet egy szélsőséges érték
 - **B. Transzformáljuk az adatokat**
 - PRO: nincs információvesztés KON: nehéz értelmezni
 - Lineáris transzformációkkal (pl. gyökvonással) az outlierek hatásának csökkentése— elméletileg legjobb megoldás, gyakorlatilag minél jobban transzformáljuk az adatokat, annál nehezebb pszichológiailag releváns következtetéseket levonni majd
 - **C. Winzorizálás - Átírjuk az értéket valami olyanra, ami már nem outlier**
 - PRO: torzító hatás megszűnik kevés információvesztéssel KON: sokan (feleslegesen) idegenkednek tőle
 - Csalás lenne? De ha benne hagyjuk, az is torzítja az adatokat, a két rossz közül ez a kisebb
 - Mire? A következő legnagyobb -már nem outlier- érték plusz 1 egység (a következő legkisebb -1 egység)



Miért nincs outlier kategoriális változóban?

- Miért nincs outlier kategoriális változóban?
 - Mert kategoriális változó értékei nem sorba rendezhetők
 - Ha 1-gyel jelölöm a szőkét, 2-vel a vörösöket, 3-mal a barnákat, annak nincs olyan jelentése, hogy a barnák többek lennének, mint a szőkék.
 - Ezért a negyedelő pontok sem értelmezhetők (nincs olyan, hogy a minta alsó vagy felső negyede, mert a mintának nincs alsó és felső része)
- Miért jelölhet a program mégis outliert, ha kategoriális változóra kérünk ki boxplotot.
 - A program nem „tudja”, hogy milyen butaságra készülsz, és ha kéred, készít neked boxplotot kategoriális adatokhoz. Ahol a kategóriákat jelölő értékeket úgy veszi, mintha azok sorba rendezhetők lennének.
- Otthon gondold át, mi történik a következő esetben:
 - Van egy dichotóm változó, például az, hogy átmegy-e a hallható a statisztika vizsgán. 0-val jelölve a nem, 1-gyel jelölve az igen. Az évfolyam (a minta) 100 fős, ebből 10 bukik, 90 átmegy.
 - Ha kikéred a boxplotot, akkor a 10 bukók mind outlierként lesznek jelölve.
 - Miért? Gondold végig az outlier labelling rule számítását! Hol lesznek ebben a tévesen értelmezett helyzetben a „kvartilisek”



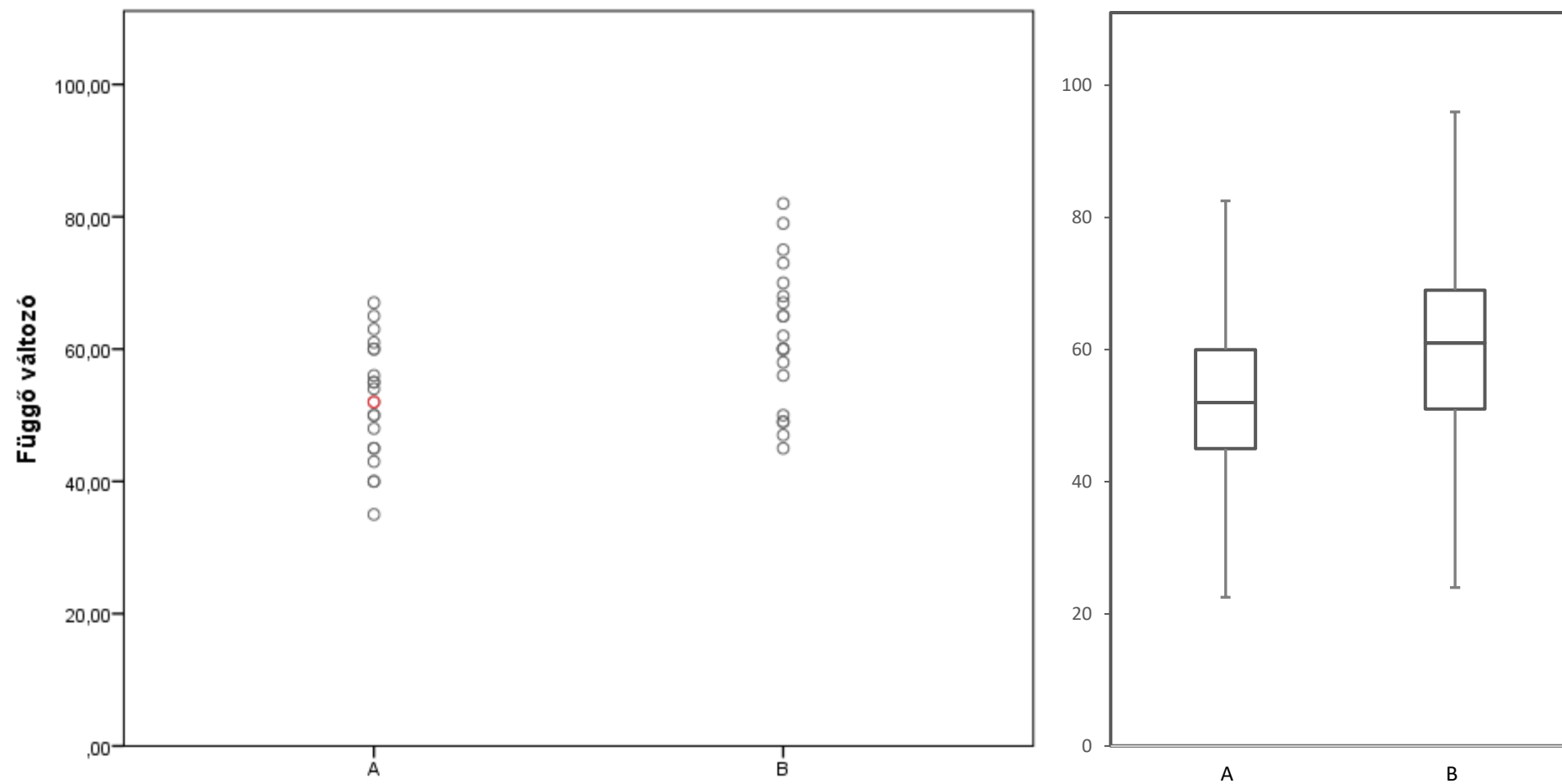
Lehetséges publikálás

A adatokban a hagyományos outlier labelling rule (Tukey, 1977) alapján két szélsőséges értéket is találtunk, a kor adatokban egy 77 éves személyt és a tág kapcsolati körben egy személyt, akinek 1910 Facebook ismerőse van.

Az így megtalált outlierok közül a felülvizsgált, 2.2-es szorzót használó outlier labelling szabály (Hoaglin és Iglewicz, 1987) alapján csak a tág kapcsolati körben talált érték minősül outlieroknak.

A megtalált szélsőséges értéket a winzorizálás szabályainak megfelelően átírtuk a rangsorban következő legnagyobb már nem szélsőséges értéknél egy egységgel nagyobbra, azaz 1218-ra.

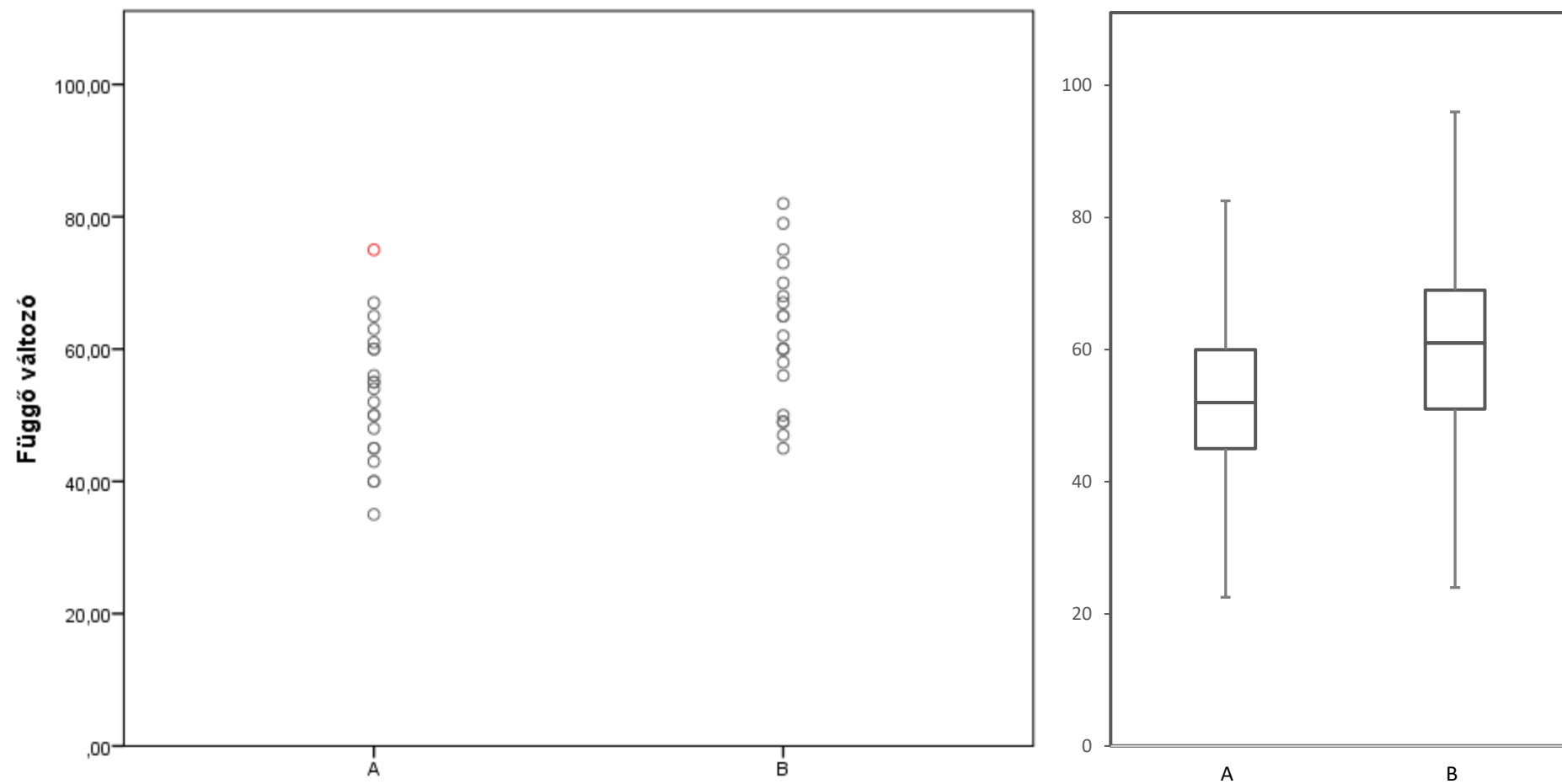
- Hivatkozások:
- Tukey, J. W. (1977). Exploratory data analysis.
- Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. Journal of the American Statistical Association, 82(400), 1147-1149.
- Hoaglin, D.C., Iglewicz, B., and Tukey, J.W. (1986). Performance of some resistant rules for outlier labeling, Journal of American Statistical Association, 81, 991-999.



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	985,738 ^a	1	985,738	10,325	,003	,209
Intercept	133575,006	1	133575,006	1399,165	,000	,973
group	985,738	1	985,738	10,325	,003	,209
Error	3723,238	39	95,468			
Total	137804,000	41				
Corrected Total	4708,976	40				

a. R Squared = ,209 (Adjusted R Squared = ,189)

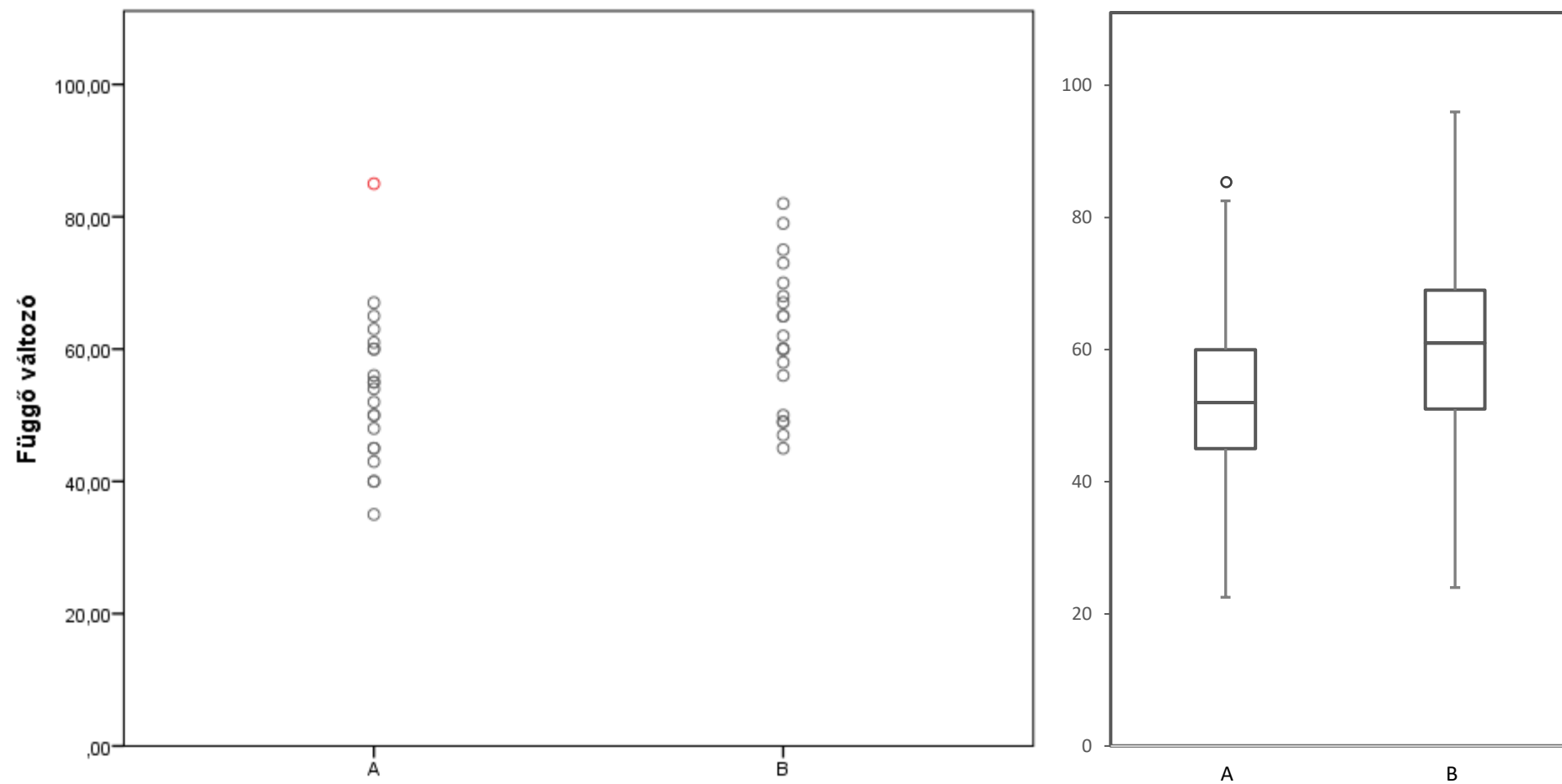
csoport	Mean	Std. Deviation	N
A	52,1905	8,77849	21
B	62,0000	10,71644	20
Total	56,9756	10,85009	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	777,909 ^a	1	777,909	7,192	,011	,156
Intercept	136149,617	1	136149,617	1258,766	,000	,970
group	777,909	1	777,909	7,192	,011	,156
Error	4218,286	39	108,161			
Total	140725,000	41				
Corrected Total	4996,195	40				

a. R Squared = ,156 (Adjusted R Squared = ,134)

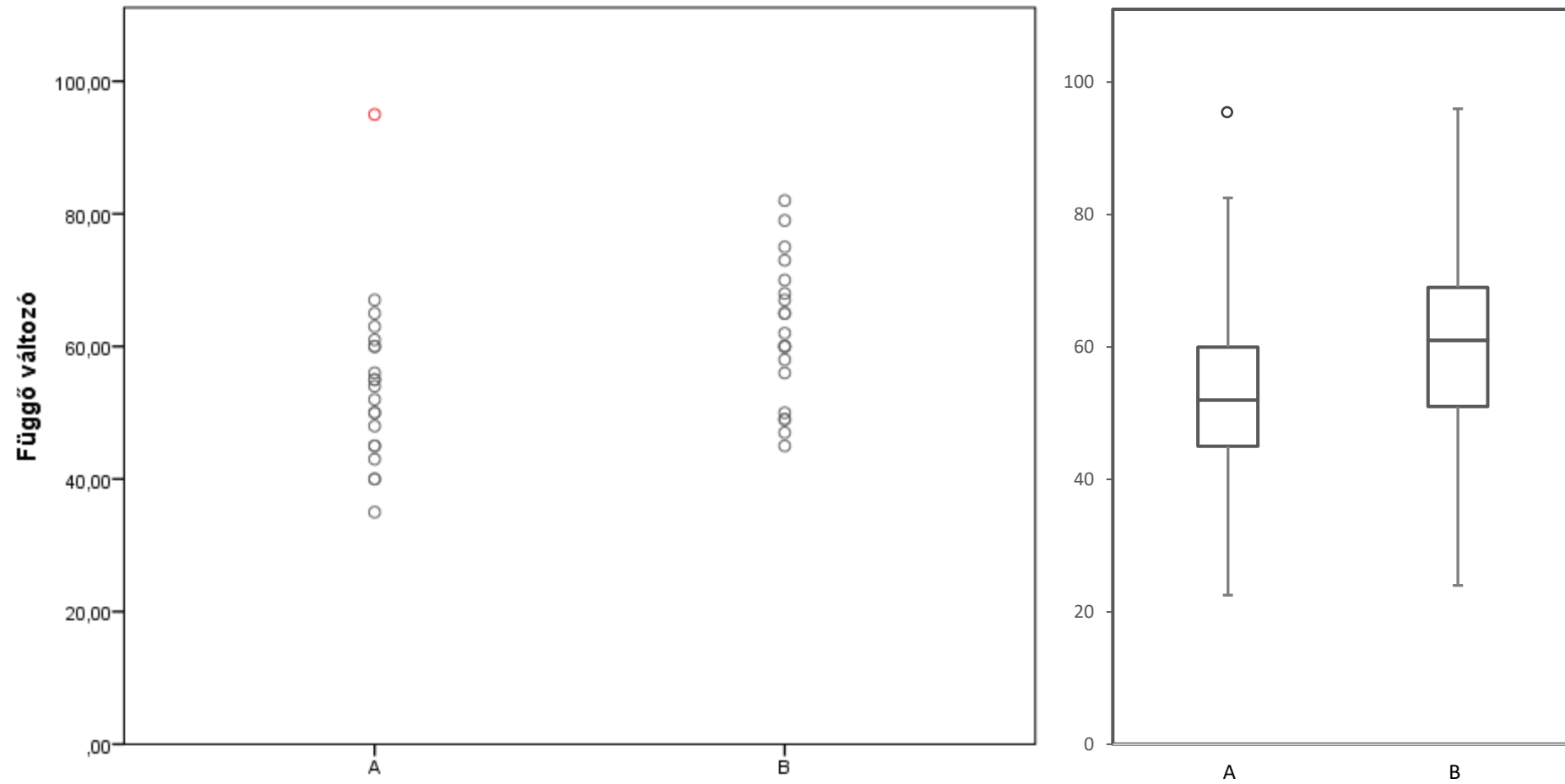
csoport	Mean	Std. Deviation	N
A	53,2857	10,09031	21
B	62,0000	10,71644	20
Total	57,5366	11,17609	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	695,215 ^a	1	695,215	5,711	,022	,128
Intercept	137276,678	1	137276,678	1127,634	,000	,967
group	695,215	1	695,215	5,711	,022	,128
Error	4747,810	39	121,739			
Total	142325,000	41				
Corrected Total	5443,024	40				

a. R Squared = ,128 (Adjusted R Squared = ,105)

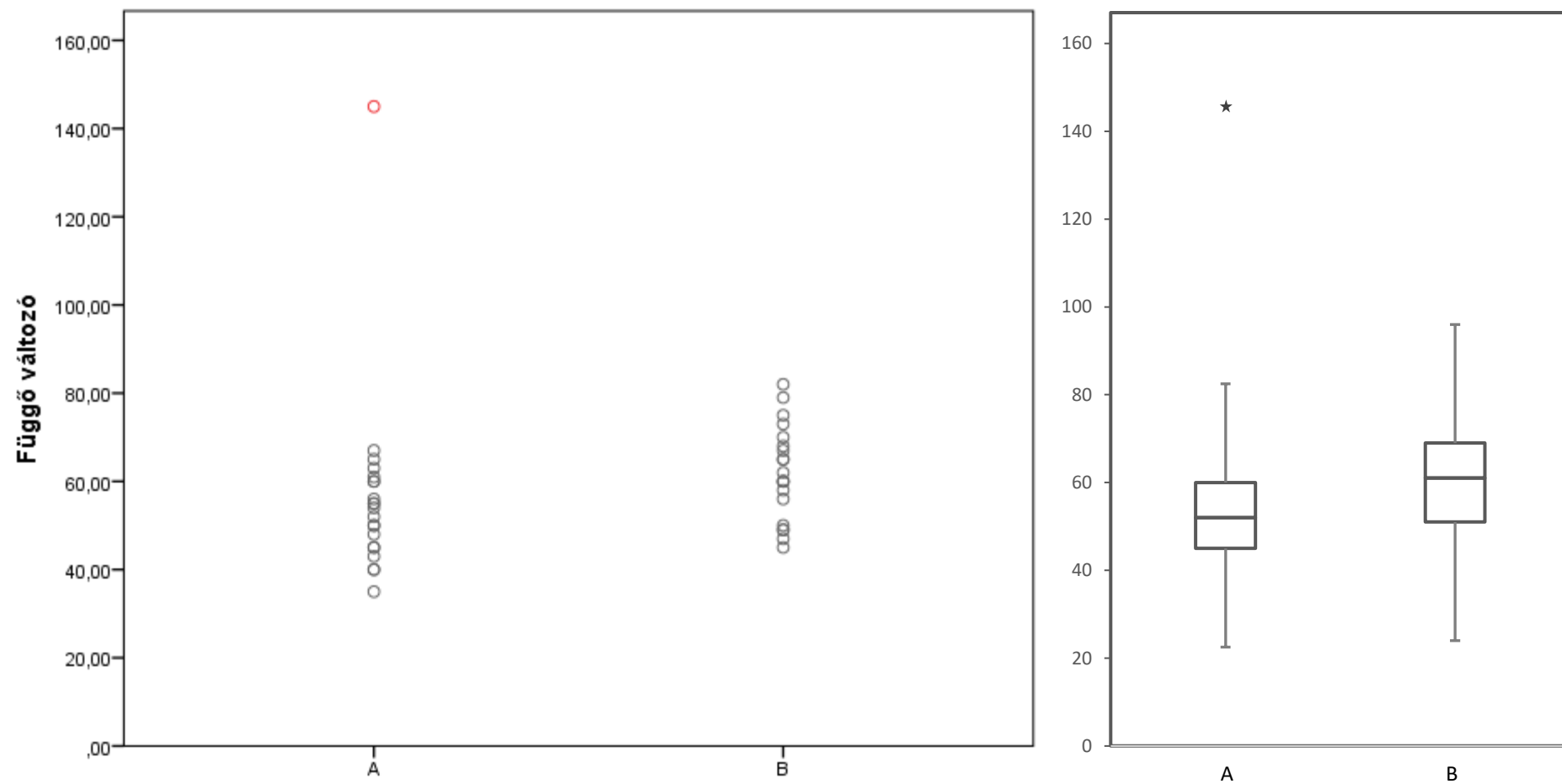
csoport	Mean	Std. Deviation	N
A	53,7619	11,32654	21
B	62,0000	10,71644	20
Total	57,7805	11,66515	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	617,166 ^a	1	617,166	4,402	,042	,101
Intercept	138408,386	1	138408,386	987,219	,000	,962
group	617,166	1	617,166	4,402	,042	,101
Error	5467,810	39	140,200			
Total	144125,000	41				
Corrected Total	6084,976	40				

a. R Squared = ,101 (Adjusted R Squared = ,078)

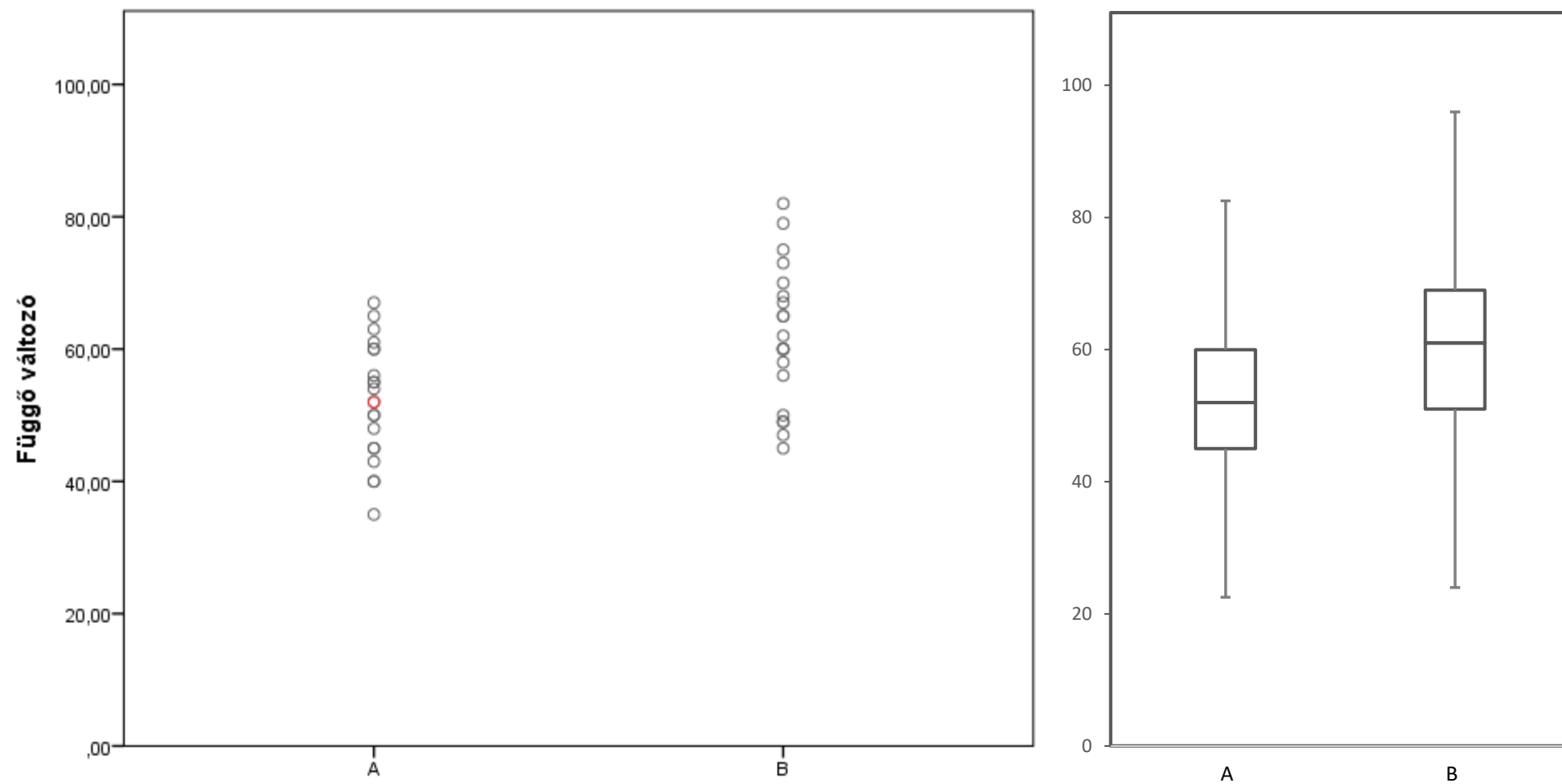
csoport	Mean	Std. Deviation	N
A	54,2381	12,81758	21
B	62,0000	10,71644	20
Total	58,0244	12,33387	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	296,609 ^a	1	296,609	,970	,331	,024
Intercept	144136,609	1	144136,609	471,392	,000	,924
group	296,609	1	296,609	,970	,331	,024
Error	11924,952	39	305,768			
Total	156125,000	41				
Corrected Total	12221,561	40				

a. R Squared = ,024 (Adjusted R Squared = -,001)

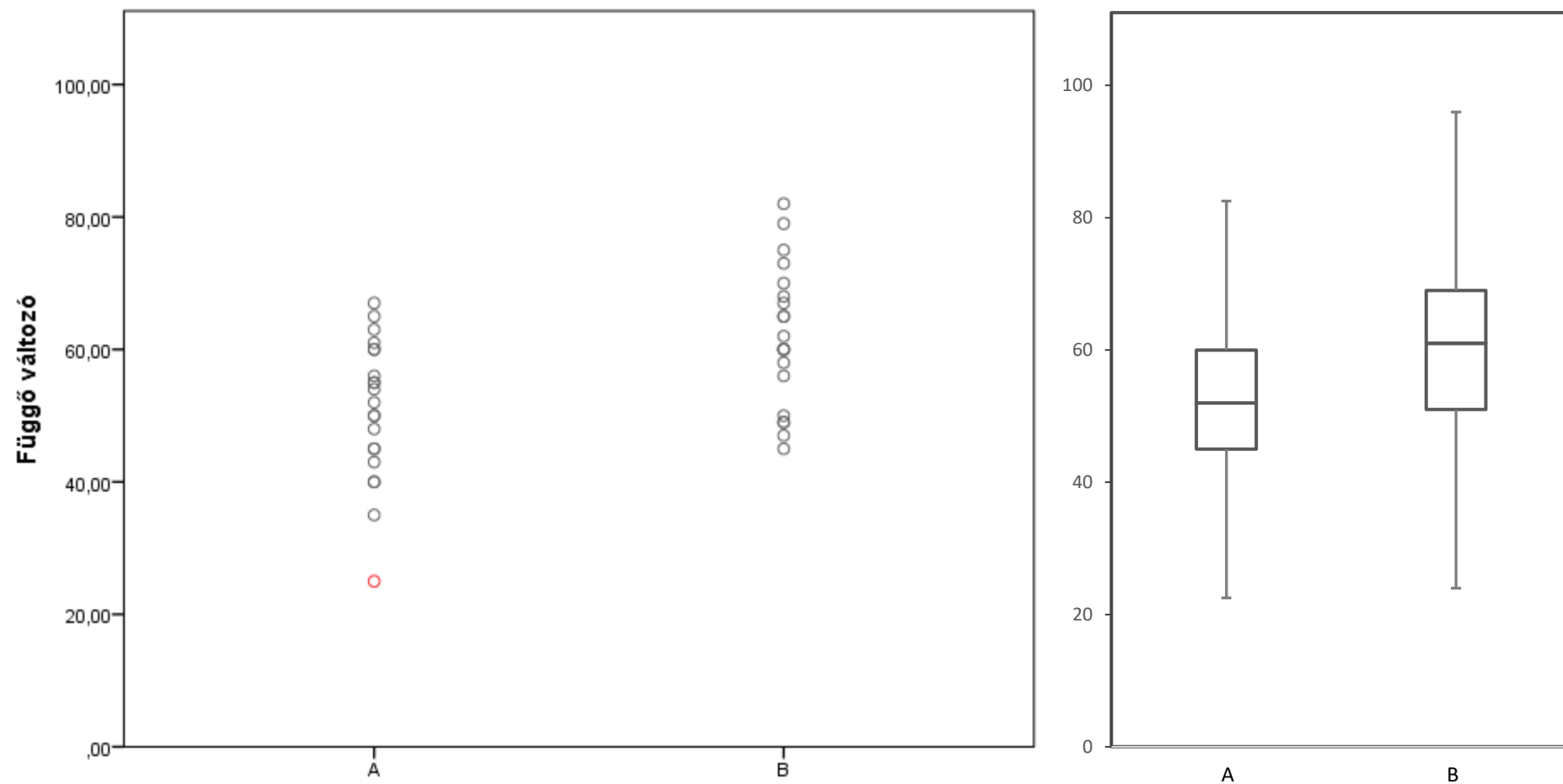
csoport	Mean	Std. Deviation	N
A	56,6190	22,07142	21
B	62,0000	10,71644	20
Total	59,2439	17,47967	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	985,738 ^a	1	985,738	10,325	,003	,209
Intercept	133575,006	1	133575,006	1399,165	,000	,973
group	985,738	1	985,738	10,325	,003	,209
Error	3723,238	39	95,468			
Total	137804,000	41				
Corrected Total	4708,976	40				

a. R Squared = ,209 (Adjusted R Squared = ,189)

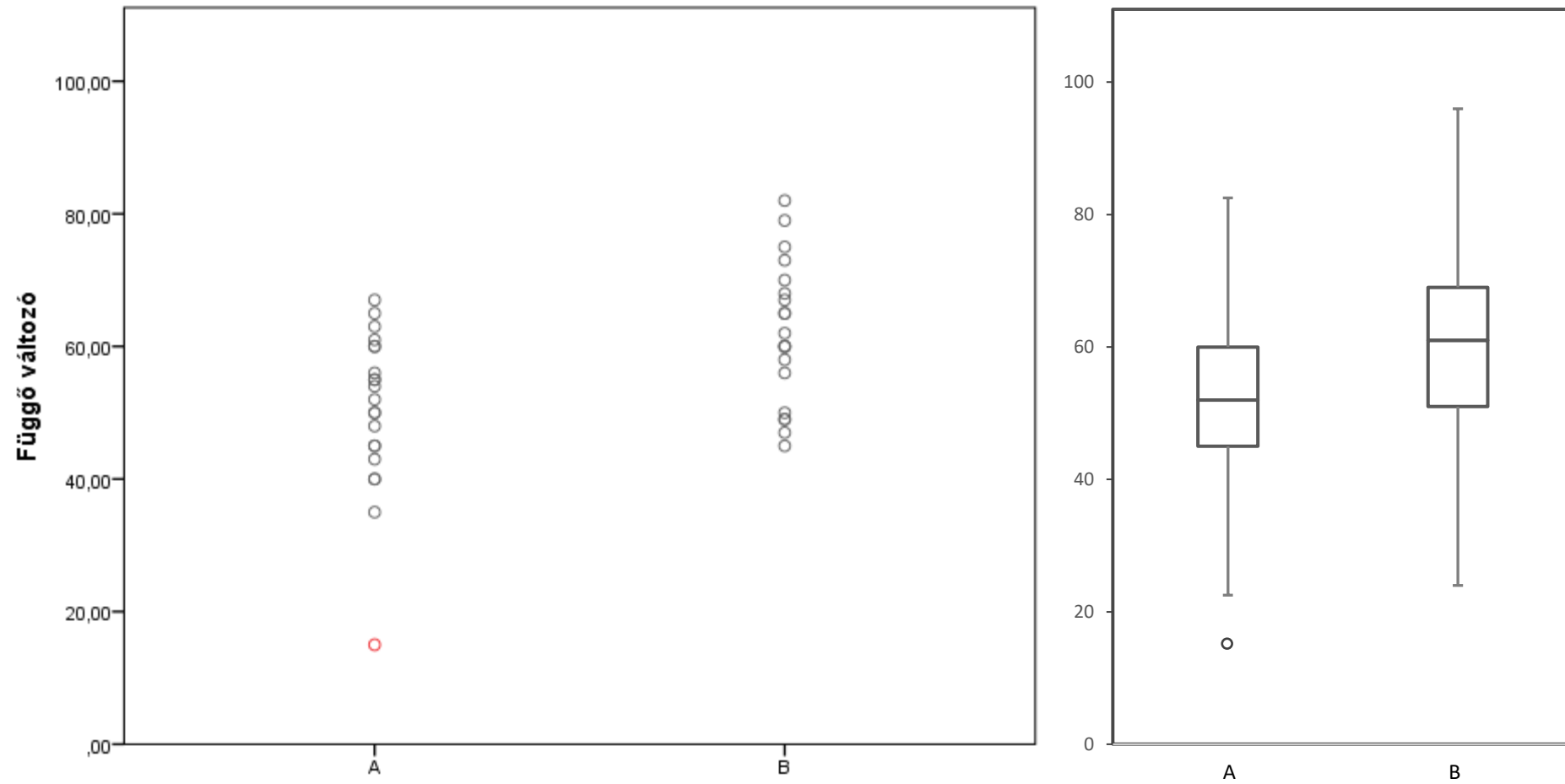
csoport	Mean	Std. Deviation	N
A	52,1905	8,77849	21
B	62,0000	10,71644	20
Total	56,9756	10,85009	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1261,069 ^a	1	1261,069	11,107	,002	,222
Intercept	130583,995	1	130583,995	1150,180	,000	,967
group	1261,069	1	1261,069	11,107	,002	,222
Error	4427,810	39	113,534			
Total	135725,000	41				
Corrected Total	5688,878	40				

a. R Squared = ,222 (Adjusted R Squared = ,202)

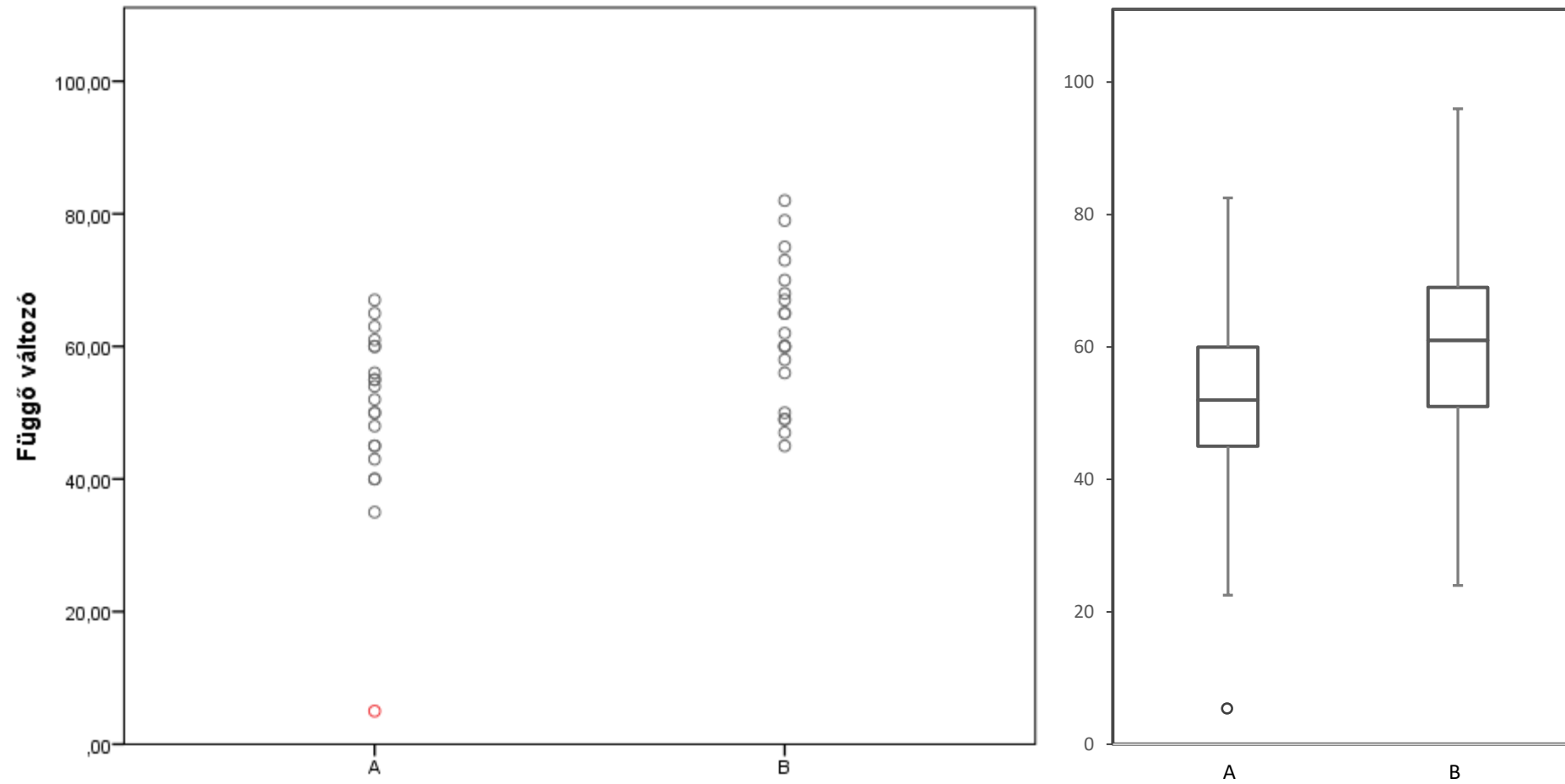
csoport	Mean	Std. Deviation	N
A	50,9048	10,59672	21
B	62,0000	10,71644	20
Total	56,3171	11,92568	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1371,638 ^a	1	1371,638	10,611	,002	,214
Intercept	129484,808	1	129484,808	1001,739	,000	,963
group	1371,638	1	1371,638	10,611	,002	,214
Error	5041,143	39	129,260			
Total	135325,000	41				
Corrected Total	6412,780	40				

a. R Squared = ,214 (Adjusted R Squared = ,194)

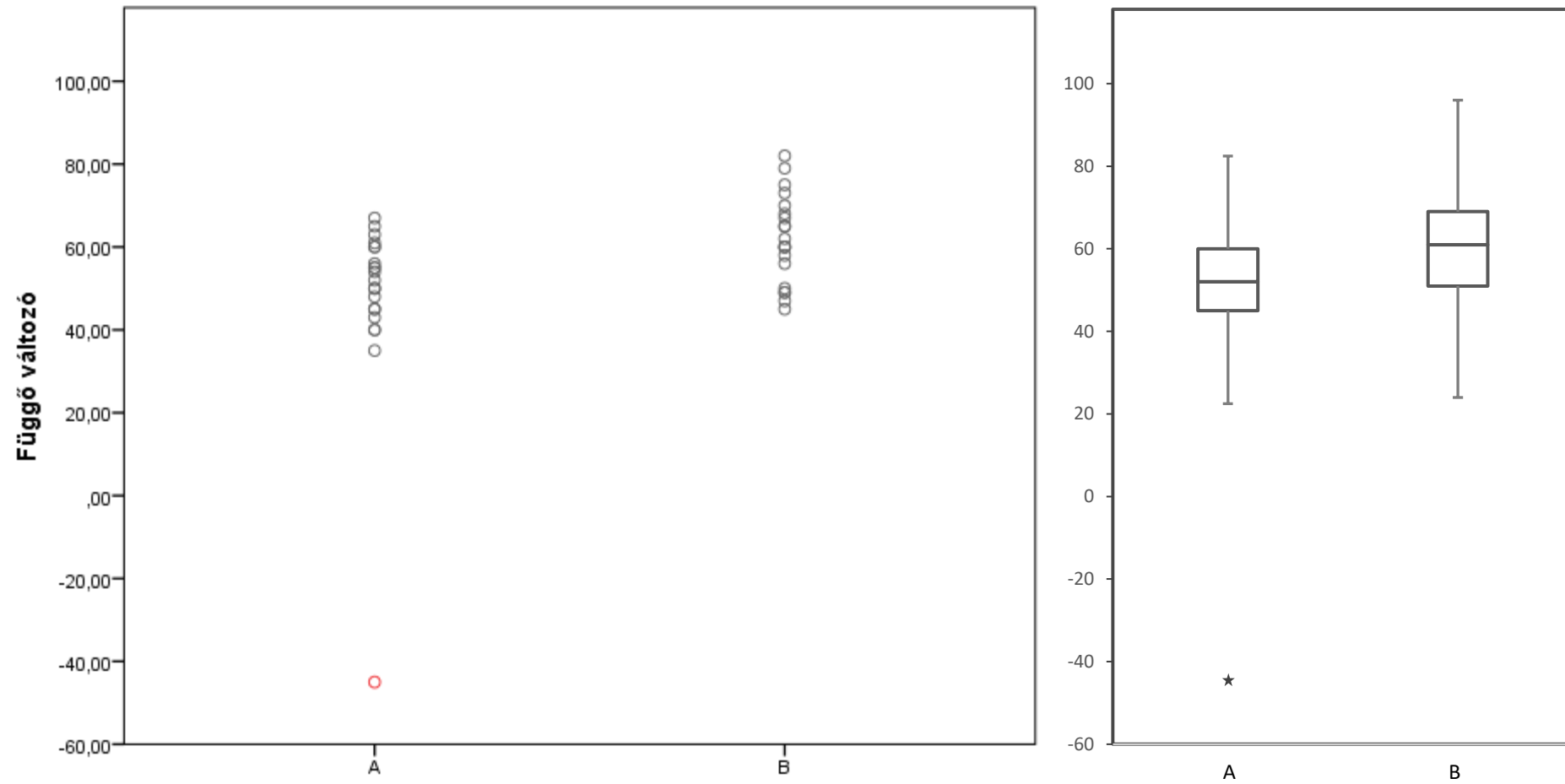
csoport	Mean	Std. Deviation	N
A	50,4286	11,95647	21
B	62,0000	10,71644	20
Total	56,0732	12,66173	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1486,852 ^a	1	1486,852	9,921	,003	,203
Intercept	128390,267	1	128390,267	856,674	,000	,956
group	1486,852	1	1486,852	9,921	,003	,203
Error	5844,952	39	149,871			
Total	135125,000	41				
Corrected Total	7331,805	40				

a. R Squared = ,203 (Adjusted R Squared = ,182)

csoport	Mean	Std. Deviation	N
A	49,9524	13,53320	21
B	62,0000	10,71644	20
Total	55,8293	13,53865	41



Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	csoport	Mean	Std. Deviation	N
Corrected Model	2132,613 ^a	1	2132,613	6,538	,015	,144	A	47,5714	22,95555	21
Intercept	122987,247	1	122987,247	377,050	,000	,906	B	62,0000	10,71644	20
group	2132,613	1	2132,613	6,538	,015	,144	Total	54,6098	19,27029	41
Error	12721,143	39	326,183							
Total	137125,000	41								
Corrected Total	14853,756	40								

a. R Squared = ,144 (Adjusted R Squared = ,122)