



**NAVÜ**

Nemzeti  
Adatvagyon Ügynökség

# **A NEM VÉNYKÖTELES GYÓGYSZEREK FORGALMÁNAK VIZSGÁLATA A MAGYARORSZÁGI GYÓGYSZERTÁRAKBAN ONLINE PÉNZTÁRGÉPEKBŐL SZÁRMAZOTT ADATOK ALAPJÁN**

**DR. BOZSONYI KÁROLY – Vezető szakértő**  
**RÁTI CSABA – IT Igazgató**

QR





**NAVÜ**

Nemzeti  
Adatvagyon Ügynökség

# TARTALOM

**Célok**

**Adatok**

**Preprocessing**

**Leíró statisztikák**

**Tanulságok**

**Térökonometriai regresszió**

**Eredmények/tanulságok**

**További irányok**



## A pilot projekt célja

### Az OTC termékek árazásának vizsgálata robusztus adatbázisban, big data design mellett

---

Hipotézis 1: A piaci versenyhelyzet hatással van az árakra

Hipotézis 2: Az általános gazdasági-társadalmi környezet hatással van az árakra

Hipotézis 3: Markáns téri dependenciák figyelhetőek meg az árak vonatkozásában

Hipotézis 4: Markáns területi mintázatok figyelhetőek meg

A későbbi teljeskörű modell kereteinek optimalizálása





**ADATOK**





# ALAPADATOK

## NAV ONLINE PÉNZTÁRGÉPEK

IDŐSZAK: 2023-24

ÖSSZES PÉNZTÁRGÉP  
AZONOSÍTÓ:

2286 db

ÖSSZES SOR:

22 528 693 db

EBBŐL VOLTAREN TERMÉKEK

TRANZAKCIÓKBÓL ÖSSZESEN:

2 795 363 db

12,4%

19db

KÜLÖNBÖZŐ  
TERMÉKNÉV

3474 db

KÜLÖNBÖZŐ  
TERMÉKNÉV





## A PILOT SORÁN VIZSGÁLT GYÓGYSZEREK

- Algopyrin
- No-spa
- Voltaren
- Nasivin
- Normaflore
- Neocitran





# ADATOK ELŐKÉSZÍTÉSE

## ADATOK TISZTÍTÁSA

Az adatok elemzéséhez először adattisztítást volt szükség elvégezni, hogy a termékek összehasonlíthatóak legyenek, mivel termékenként több mint 1000 féle gyógyszernev volt

NEGATÍV ÉRTÉKEK KISZŰRÉSE

AGGREGÁLÁS

Adatkonzisztencia vizsgálat  
Az outlier értékek vizsgálata.  
Túl nagy és túl kicsi értékek.  
Nem egész számú doboz  
darabok kiszűrése.

Mutatóképzések  
(MIN , MAX , AVG)



## ADATKÉPZÉS





# ADATOK ELŐKÉSZÍTÉSE

## AGGREGÁLÁSI FOLYAMAT





# MÓDSZERTAN

## ADATTRANSZFORMÁCIÓ

AGGREGÁLÁS (HÓNAP)

GEOMETRIAI ÁTLAG  
KISZÁMÍTÁSA

MINIMUM ÉS MAXIMUM  
SOROK MEGHATÁROZÁSA



NAVÜ

Nemzeti  
Adatvagyon Ügynökség





# MÓDSZERTAN

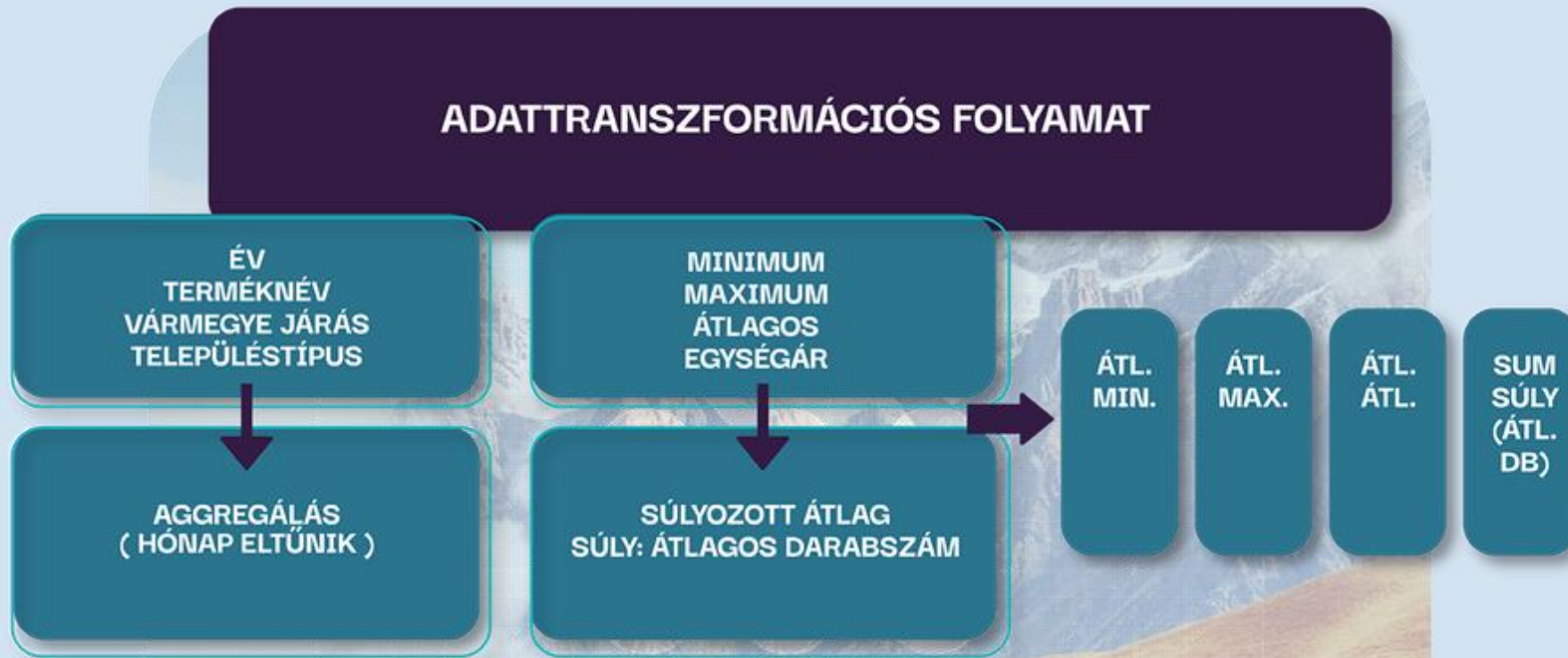
## ADATTRANSZFORMÁCIÓ

AGGREGÁLÁS (HÓNAP)





# MÓDSZERTAN





# MÓDSZERTAN

## ADATTRANSZFORMÁCIÓ

GEOMETRIAI ÁTLAG  
KISZÁMÍTÁSA



NAVÜ

Nemzeti  
Adatvagyon Ügynökség





# MÓDSZERTAN

## ADATTRANSZFORMÁCIÓS FOLYAMAT





# MÓDSZERTAN

## ADATTRANSZFORMÁCIÓ

MINIMUM ÉS MAXIMUM  
SOROK MEGHATÁROZÁSA



## Minimumok

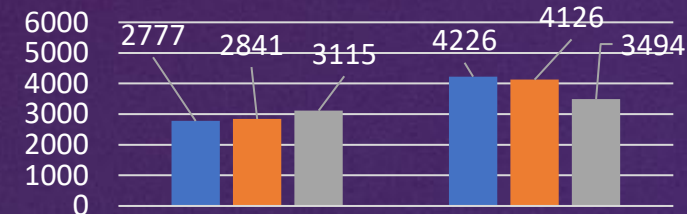
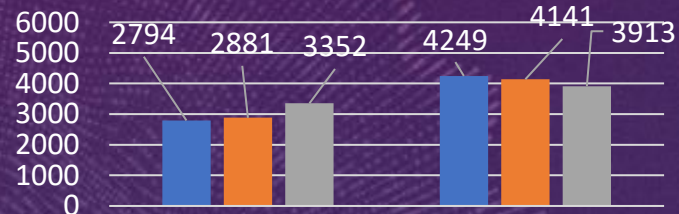
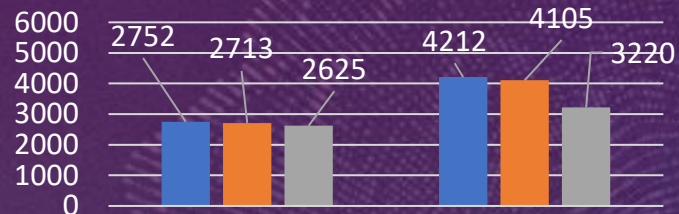
## Maximumok

## Átlagok

2023

2023

2023



Minimuma

Maximuma

■ Községek

■ Városok

■ Megyei jogú városok és kerületek

Minimuma

Maximuma

■ Községek

■ Városok

■ Megyei jogú városok és kerületek

Minimuma

Maximuma

■ Községek

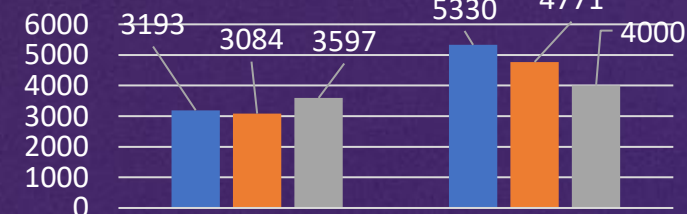
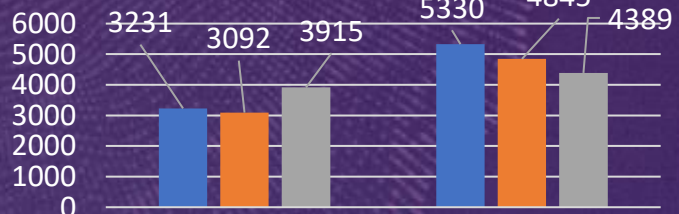
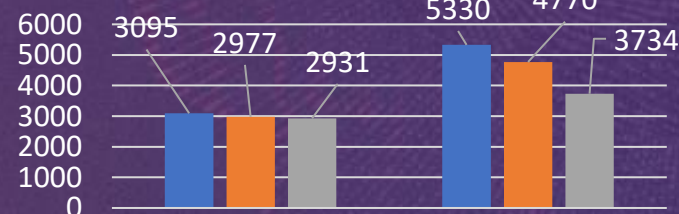
■ Városok

■ Megyei jogú városok és kerületek

2024

2024

2024



Minimuma

Maximuma

■ Községek

■ Városok

■ Megyei jogú városok és kerületek

Minimuma

Maximuma

■ Községek

■ Városok

■ Megyei jogú városok és kerületek

Minimuma

Maximuma

■ Községek

■ Városok

■ Megyei jogú városok és kerületek



# EREDMÉNYEK





## FÜGGŐ VÁLTOZÓK

### Járásra aggregált községi ár

- Minimum ár
- Átlag ár
- Maximum ár

### Járásra aggregált városi (város, vármegyeszékhely, vármegyei jogú város, budapesti kerület) ár

- Minimum ár
- Átlag ár
- Maximum ár





## VIZSGÁLT FÜGGETLEN VÁLTOZÓK

- Járás létszáma (1000 fő) - népszámlálás
- Férfiak/nők száma - népszámlálás
- Iskolai végzettség – népszámlálás
- Életkori csoportok – népszámlálás
- Adófizetők aránya - jövedelem (NAV – 2022)
- 100 főre jutó átlagos SZJA adóalapot képező jövedelem (NAV – 2022)
- Városi rétegen belül mekkora a vármegyeszékhelyek/vármegyei jogú városok aránya  
(BP kerület vármegyeszékhelynek számít a számolásban)





## MODELLEZÉSBEN FELHASZNÁLT FÜGGETLEN VÁLTOZÓK

Települési adatból kiválasztjuk a rétegeket (község/város)  
és a kiválasztott rétegbe tartozó településeket aggregáljuk járási szintre

A független változók között erős a multikollinearitás  
(végzettség és NAV adatok között elsősorban) –  
ezért szűkített változó készlet a modellekben

- Járás létszáma (1000 fő)
- Férfiak aránya
- 60+ személyek aránya
- Jövedelem





## MODELLEZÉS

**Téri dependencia tesztelése – Moran I statisztika**

**Téri regressziós modell**

- Error model  $y = X\beta + u, u = \lambda W u + \varepsilon$

**Községi adatok esetében három járásnál volt adathiány a függő változóban (két olyan járás van, amiben nem volt községi adat )**

- A szomszédos járások átlagával pótoltuk





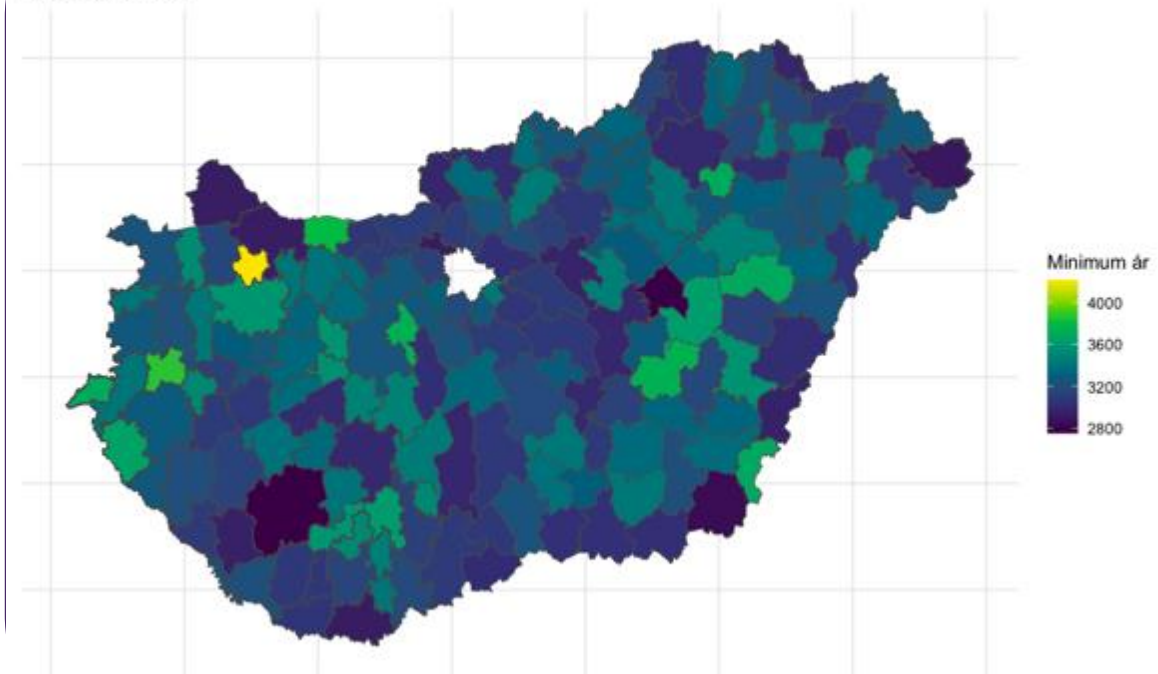
# KÖZSÉGI RÉTEG



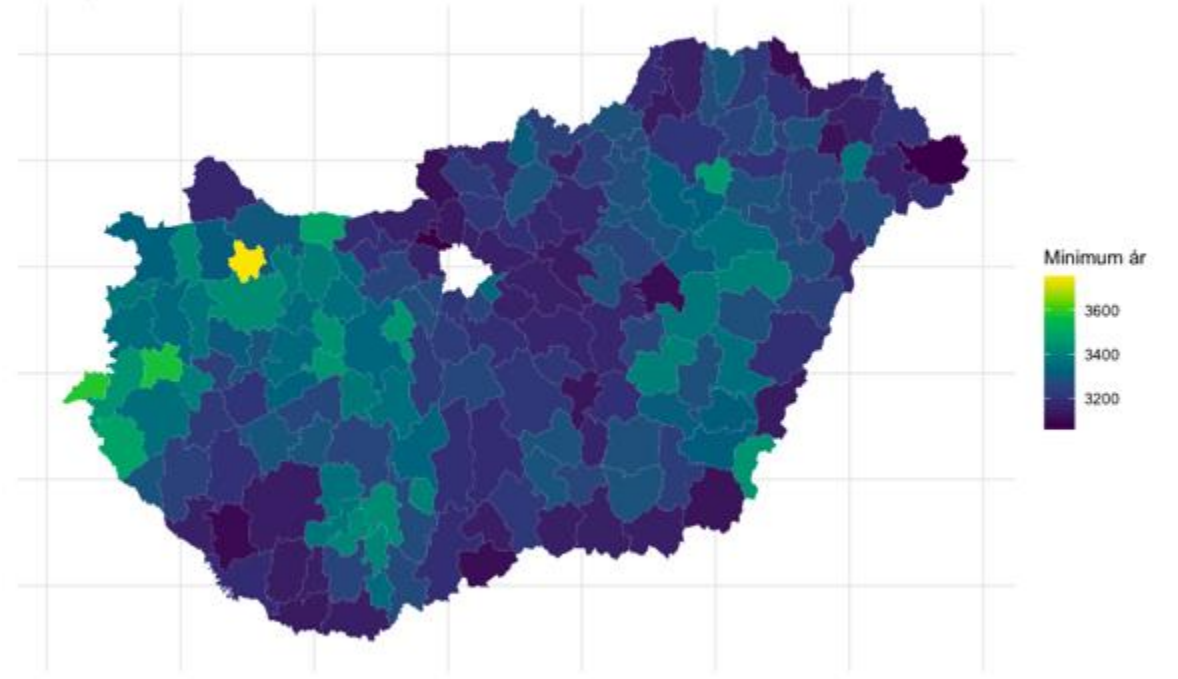


# MINIMUM ÁR

Községi adatok - minimum ár  
Moran I stat: 0.048



Községi adatok - Minimum ár - Simított





## MINIMUM ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	4569.9753	1204.8761	3.7929	0.0001489
lakossag	-10.7126	1.4987	-7.1481	8.8e-13
ffi_arany	-29.9278	23.1681	-1.2918	0.1964382
idos_arany	3.7557	5.1459	0.7298	0.4654884
szja	13.8013	3.9283	3.5133	0.0004426

Lambda: 0.22511, LR test value: 3.4687, p-value: 0.062541

Asymptotic standard error: 0.10893

z-value: 2.0665, p-value: 0.038781

Wald statistic: 4.2704, p-value: 0.038781

Log likelihood: -1150.511 for error model

ML residual variance (sigma squared): 37419, (sigma: 193.44)

Nagelkerke pseudo-R-squared: 0.25532

Number of observations: 172

Number of parameters estimated: 7

AIC: 2315, (AIC for lm: 2316.5)

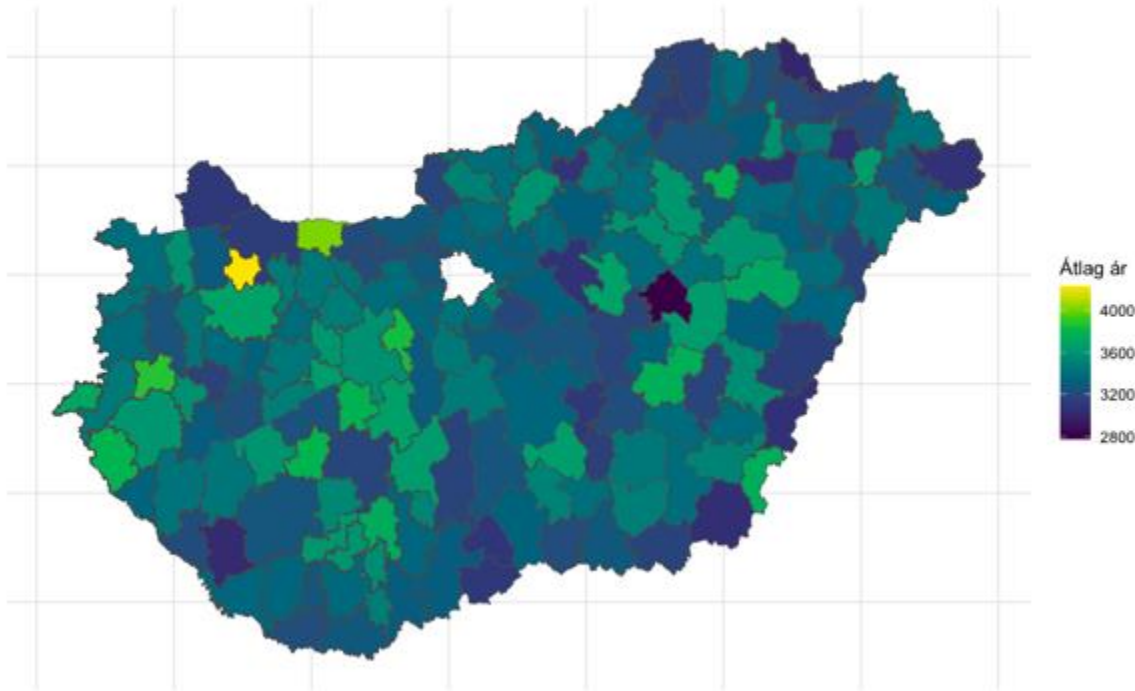
Hausman test: 5.5601, df: 5, p-value: 0.3514



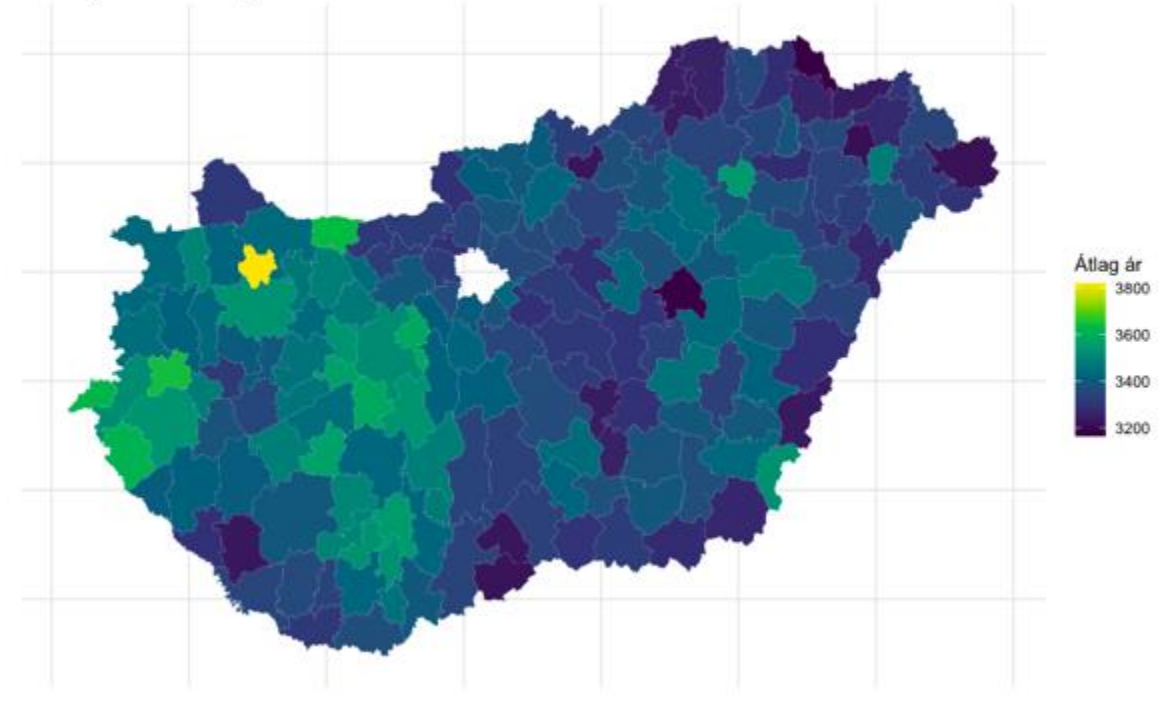


# ÁTLAGOS ÁR

Községi adatok - átlag ár  
Moran I stat: 0.05



Községi adatok - Átlag ár - Simitott





## ÁTLAGOS ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	4173.3658	1162.9967	3.5885	0.0003326
lakossag	-5.1307	1.4820	-3.4620	0.0005361
ffi_arany	-24.0902	22.4465	-1.0732	0.2831689
idos_arany	9.2596	4.8403	1.9130	0.0557421
szja	13.8826	3.6487	3.8048	0.0001419

Lambda: 0.1374, LR test value: 1.1524, p-value: 0.28304

Asymptotic standard error: 0.11396

z-value: 1.2057, p-value: 0.22793

Wald statistic: 1.4537, p-value: 0.22793

Log likelihood: -1147.229 for error model

ML residual variance (sigma squared): 36253, (sigma: 190.4)

Nagelkerke pseudo-R-squared: 0.14192

Number of observations: 172

Number of parameters estimated: 7

AIC: 2308.5, (AIC for lm: 2307.6)

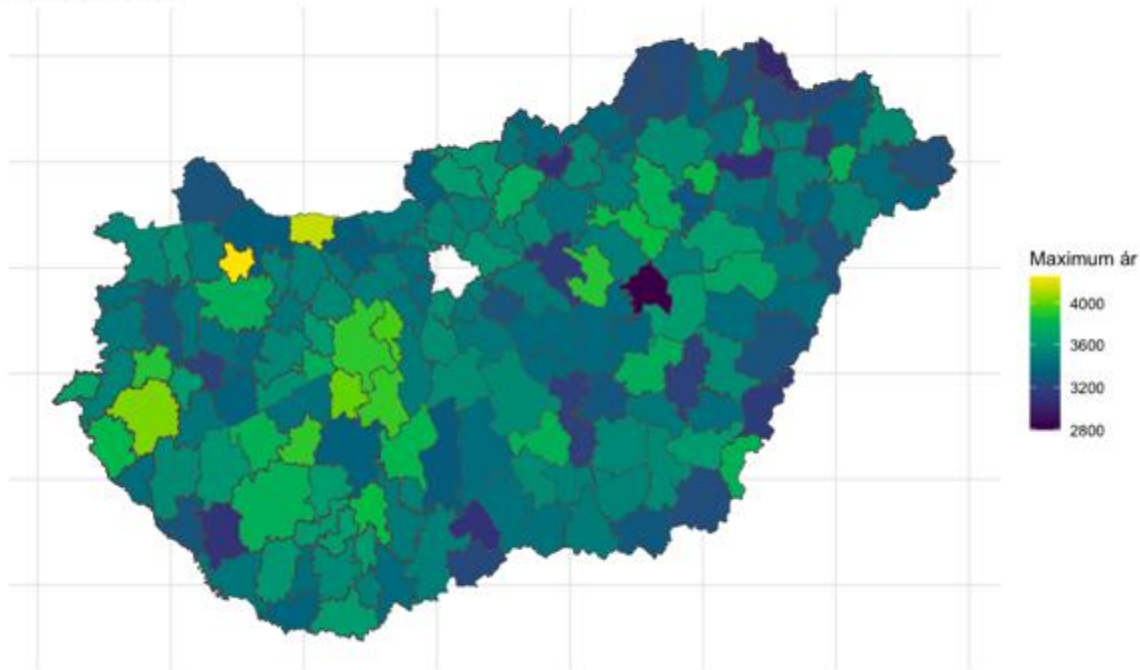
Hausman test: 5.3538, df: 5, p-value: 0.37425



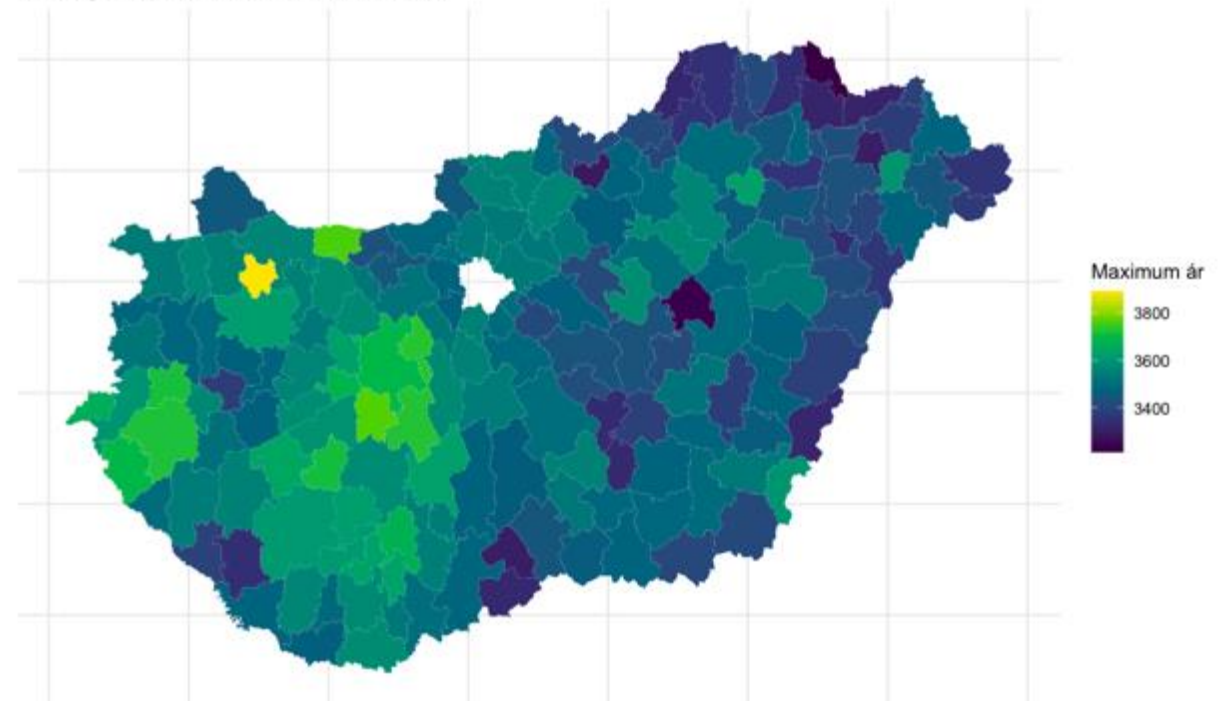


# MAXIMUM ÁR

Községi adatok - maximum ár  
Moran I stat: 0.105



Községi adatok - Maximum ár - Simított





## MAXIMUM ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	4071.97683	1258.80394	3.2348	0.0012173
lakossag	0.09287	1.59026	0.0584	0.9534308
ffi_arany	-22.99339	24.26374	-0.9476	0.3433106
idos_arany	11.83067	5.28775	2.2374	0.0252620
szja	14.24821	4.00305	3.5593	0.0003718

Lambda: 0.16903, LR test value: 1.8038, p-value: 0.17926

Asymptotic standard error: 0.11224

z-value: 1.5059, p-value: 0.1321

Wald statistic: 2.2677, p-value: 0.1321

Log likelihood: -1159.778 for error model

ML residual variance (sigma squared): 41867, (sigma: 204.61)

Nagelkerke pseudo-R-squared: 0.11792

Number of observations: 172

Number of parameters estimated: 7

AIC: 2333.6, (AIC for lm: 2333.4)

Hausman test: 4.8409, df: 5, p-value: 0.43561





# VÁROSI RÉTEG

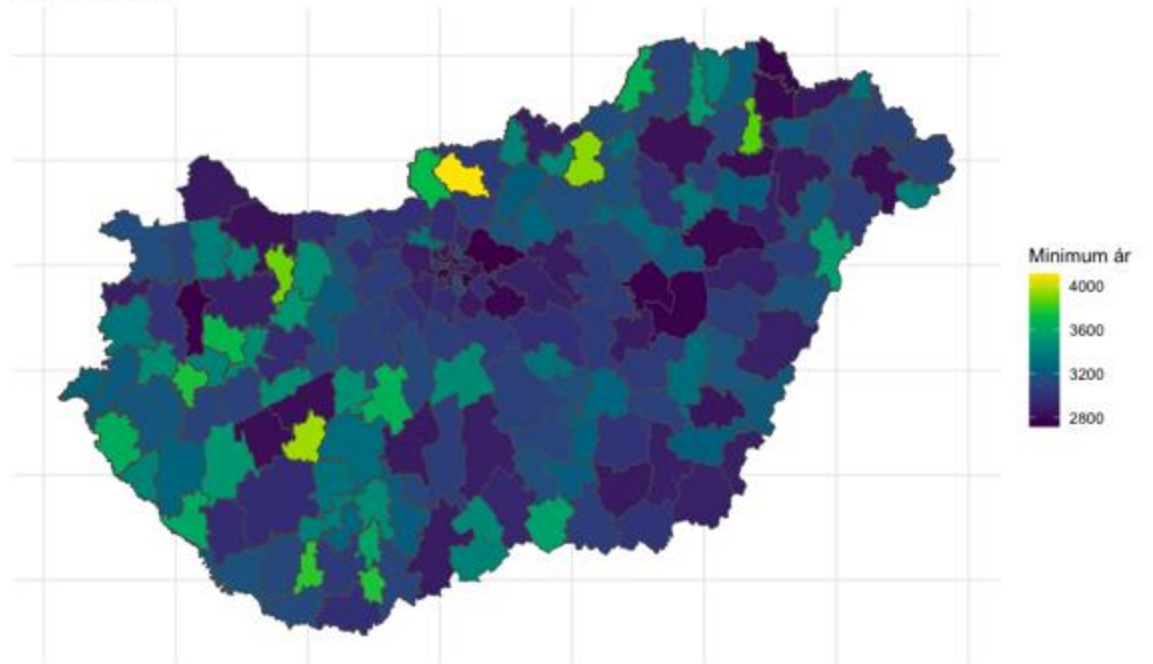




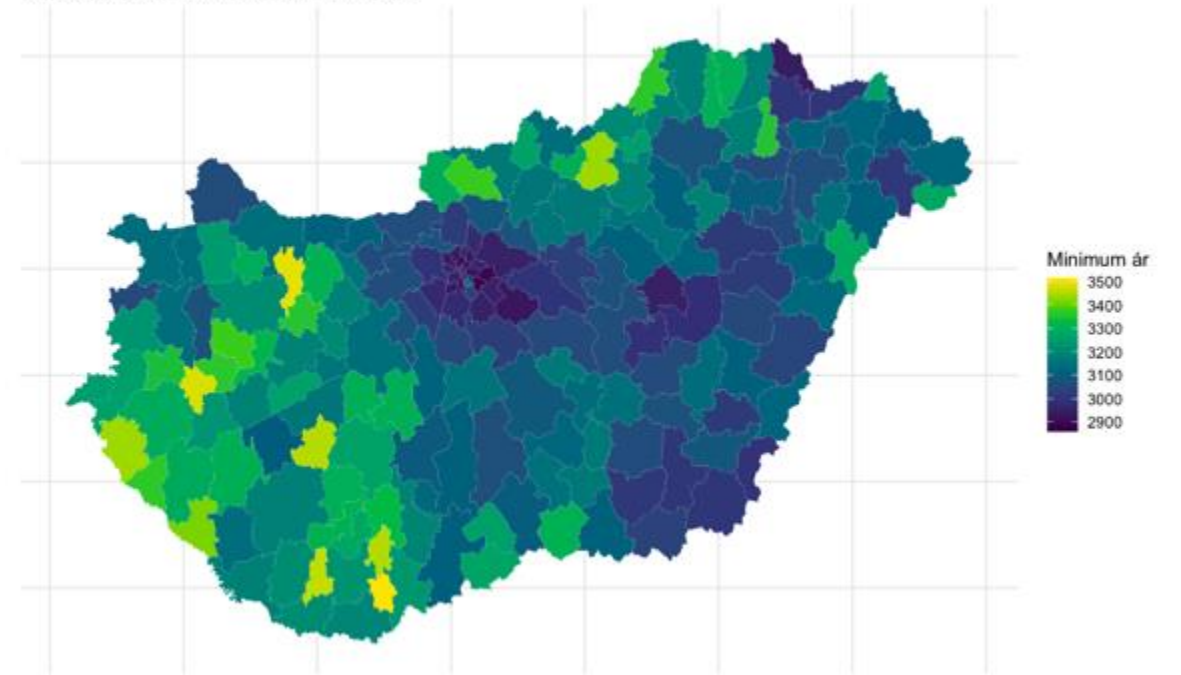
# MINIMUM ÁR

Községi adatok - minimum ár

Moran I stat: 0.116



Városi adatok - Minimum ár - Simított





## MINIMUM ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	1694.638497	1423.240934	1.1907	0.2338
lakossag	-2.661177	0.546924	-4.8657	1.14e-06
ffi_arany	26.699493	26.559014	1.0053	0.3148
idos_arany	9.508083	7.051648	1.3483	0.1775
szja	-0.022957	0.022227	-1.0328	0.3017

Lambda: 0.096194, LR test value: 0.58511, p-value: 0.44432

Asymptotic standard error: 0.11089

z-value: 0.8675, p-value: 0.38567

Wald statistic: 0.75255, p-value: 0.38567

Log likelihood: -1355.162 for error model

ML residual variance (sigma squared): 55182, (sigma: 234.91)

Nagelkerke pseudo-R-squared: 0.23082

Number of observations: 197

Number of parameters estimated: 7

AIC: 2724.3, (AIC for lm: 2722.9)

Hausman test: 10.47, df: 5, p-value: 0.062953

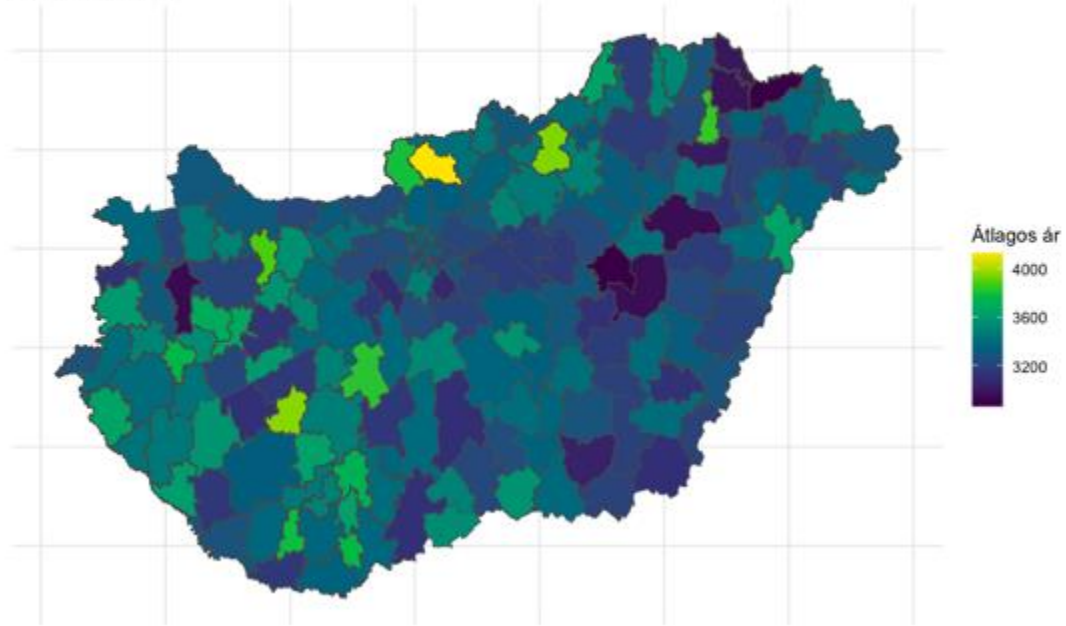




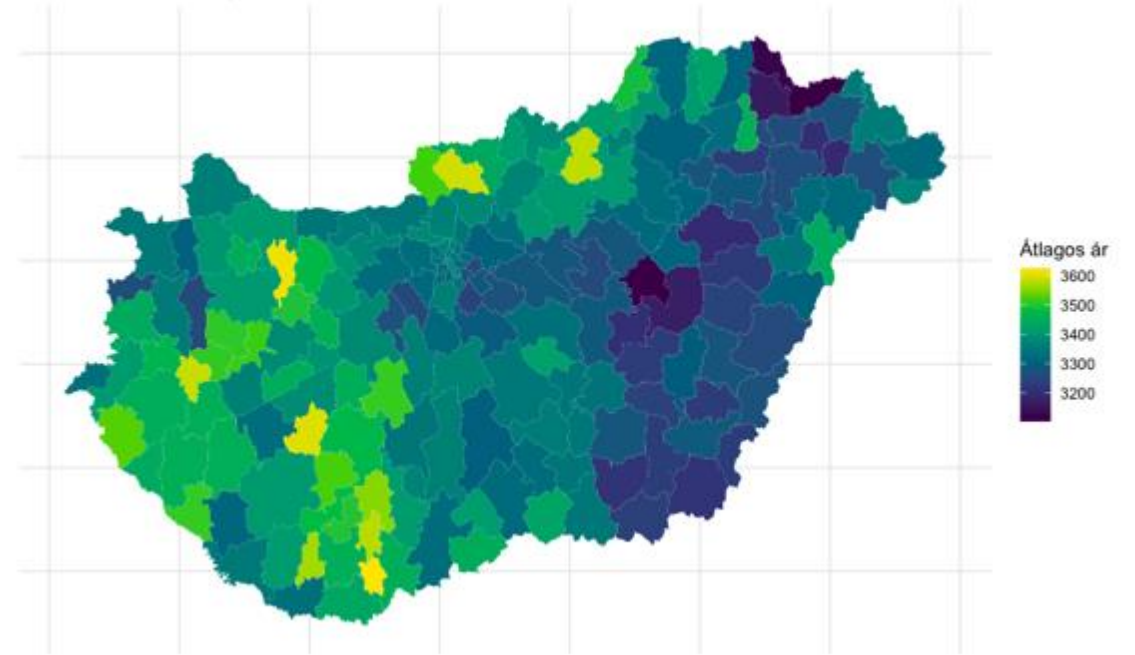
# ÁTLAGOS ÁR

Községi adatok - átlagos ár

Moran I stat: 0.143



Városi adatok - Átlagos ár - Simított





## ÁTLAGOS ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	2.7644e+03	1.1861e+03	2.3306	0.01978
lakossag	-6.4355e-01	4.3772e-01	-1.4702	0.14149
ffi_arany	9.4920e+00	2.2068e+01	0.4301	0.66711
idos_arany	5.7475e+00	6.1134e+00	0.9402	0.34714
szja	3.6905e-03	1.9955e-02	0.1849	0.85328

Lambda: 0.2868, LR test value: 6.5149, p-value: 0.010697

Asymptotic standard error: 0.10019

z-value: 2.8624, p-value: 0.004204

Wald statistic: 8.1935, p-value: 0.004204

Log likelihood: -1315.44 for error model

ML residual variance (sigma squared): 36336, (sigma: 190.62)

Nagelkerke pseudo-R-squared: 0.070992

Number of observations: 197

Number of parameters estimated: 7

AIC: 2644.9, (AIC for lm: 2649.4)

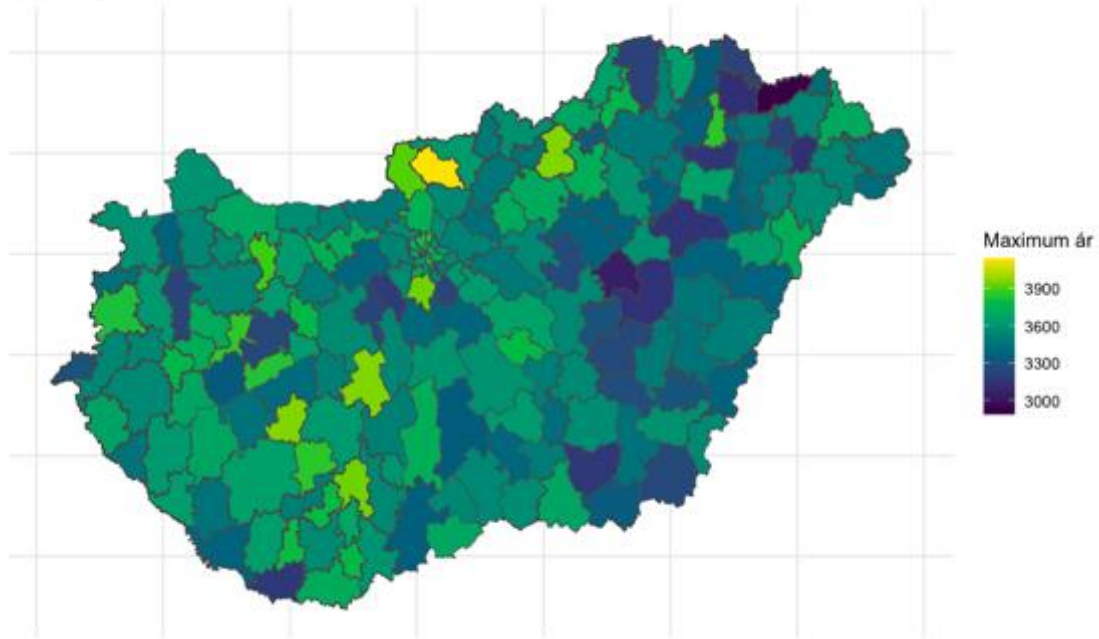
Hausman test: 10.168, df: 5, p-value: 0.070616



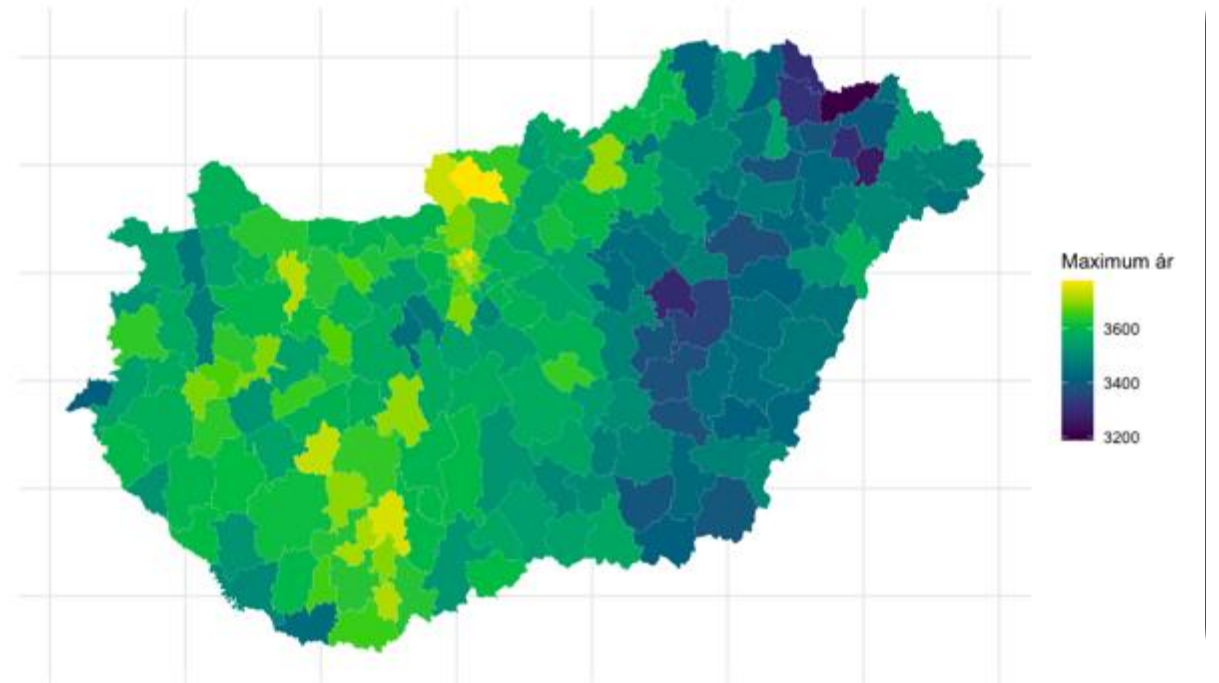


# MAXIMUM ÁR

Községi adatok - maximum ár  
Moran I stat: 0.208



Városi adatok - Maximum ár - Simitott





## MAXIMUM ÁR

Type: error

Coefficients: (asymptotic standard errors)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	4269.872636	1141.427092	3.7408	0.0001834
lakossag	0.631114	0.419251	1.5053	0.1322374
ffi_arany	-17.813111	21.227453	-0.8392	0.4013827
idos_arany	3.682219	5.911922	0.6228	0.5333855
szja	0.028647	0.019394	1.4771	0.1396374

Lambda: 0.30881, LR test value: 8.6611, p-value: 0.0032507

Asymptotic standard error: 0.098707

z-value: 3.1286, p-value: 0.0017567

Wald statistic: 9.7879, p-value: 0.0017567

Log likelihood: -1307.615 for error model

ML residual variance (sigma squared): 33469, (sigma: 182.95)

Nagelkerke pseudo-R-squared: 0.12706

Number of observations: 197

Number of parameters estimated: 7

AIC: 2629.2, (AIC for lm: 2635.9)

Hausman test: 7.6522, df: 5, p-value: 0.17647





## EREDMÉNYEK / TANULSÁGOK

- A falvakban - ahol tipikusan egy patika van - nagyobb az árak volatilitása mint a városokban
- A falvak esetén erős társadalmi/gazdasági hatások mozgatják az árakat.
- Megfigyelhető a kelet – nyugat hatás





## TOVÁBBI IRÁNYOK

- A pilot folytatása a többi kiválasztott termékkel
- Az adat anomáliák további vizsgálata/megértése
- Hosszabb időszak (minimum 4 év) vizsgálata a szezonális hatások feltárása érdekében
- Teljes körű modell (minden OTC termék) megalkotása --- ajánlások az áruk volatilitásának ésszerű csökkentése érdekében





NAVÜ

Nemzeti  
Adatvagyon Ügynökség

**KÖSZÖNÖM  
A FIGYELMET!**

QR

