Beadandó

Kornél Bánhegyi

2025-01-15

# Előszó

Ezt a beadandót R markdown fájlként írom, hogy a beadandó dolgozat replikálható legyen a fájlokkal és a kiegészíés esélye fennálljon. A következő csomagokat használom: readr dplyr knitr tidyr ggplot2 tidyverse ggrepel

## **Bevezető:**

Ebben a beadandó dolgozatomban Magyarország népességét, illetve népességének változását szeretném megvizsgáni, azzal kapcsolatos okokat, összefüggéseket és esetleges következtetéseket levonni.

## **Adatgyűjtés és rendezés**

Forrásként főleg a KSH adatait használtam.

Források:

[KSH - Népességi adatok területileg](https://www.ksh.hu/stadat_files/nep/hu/nep0034.html)

[KSH - Korcsoportok régió szerint](https://www.ksh.hu/stadat_files/nep/hu/nep0035.html)

[KSH - Munkanélküliségi adatok](https://www.ksh.hu/stadat_files/mun/hu/mun0171.html)

Először fontosnak tartom az adatokat beolvasni és szétszedni, hogy területegységileg láthassuk a releváns adatokat.

library(dplyr)  
adat1 <- read\_delim('stadat-nep0034-22.1.2.1-hu\_utf.csv', delim = ";",   
 escape\_double = FALSE, trim\_ws = TRUE, show\_col\_types = FALSE)  
adat\_munka <- read\_delim('stadat-mun0171-20.2.2.5-hu\_utf.csv', delim = ";",   
 escape\_double = FALSE, trim\_ws = TRUE, show\_col\_types = FALSE)  
adat\_oreg <- read\_delim('stadat-nep0035-22.1.2.2-hu\_utf.csv', delim = ";",   
 escape\_double = FALSE, trim\_ws = TRUE, show\_col\_types = FALSE)  
orszag<- adat1 %>%  
 filter(`Területi egység szintje` == "ország")

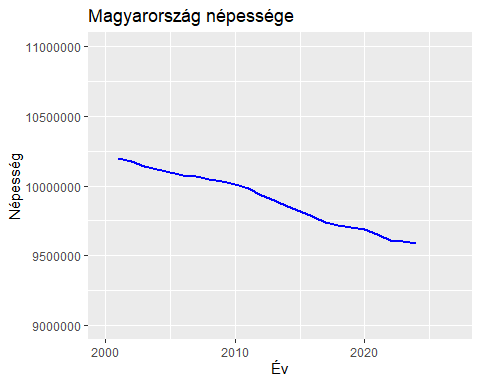
## Személyes érdeklődés

Ez a téma hazánkban szerintem kifejezetten fontos a társadalmi, gazdasági és politikai döntéshozatal szempontjából. A népesség változása nemcsak az egyének életét befolyásolja, hanem hatással van az ország munkaerőpiacára, gazdasági növekedésére, szociális ellátórendszerére és általános fejlődési kilátásaira is, és szerintem ez a probléma generációkat aggasztja.

Az alábbi ábrán is látható, hogy az elmúlt több, mint 20 évben mennyire láthatóan csökken Magyarország népessége:

library(ggplot2)  
library(dplyr)  
library(tidyr)  
orszag\_plot <- orszag %>%  
 pivot\_longer(cols = `2001`:`2024`, names\_to = "Year", values\_to = "Population")  
  
orszag\_plot2 <- orszag\_plot %>%  
 mutate(Year = as.numeric(Year),  
 Population = as.numeric(gsub(" ", "", Population)))  
ggplot(orszag\_plot2, aes(x = Year, y = Population)) +  
 geom\_line(color = "blue", size = 1) +  
 labs(title = "Magyarország népessége",  
 x = "Év",  
 y = "Népesség") +  
 scale\_y\_continuous(limits = c(9000000, 11000000)) +  
 scale\_x\_continuous(limits = c(2000, 2027)) +   
 theme\_gray()

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.  
## ℹ Please use `linewidth` instead.  
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last\_lifecycle\_warnings()` to see where this warning was  
## generated.



Személyes hipotézisem az, hogy a foglalkoztatottság szoros kapcsolatban áll a népességcsökkenéssel, ezt próbálom meg bizonyítani a későbbiekben. Szemléltetésként a vármegyei népességi adatok táblázatát választottam, ezzel fogom összehasonlítani a munkanélküliségi adatokkal 2011 és 2022 között. Az első táblázat az alap táblázatot szemlélteti, a második a népességkülönbözetet 2011 és 2022 között, vármegyékre bontva.

library(dplyr)  
library(knitr)  
varmegye<- adat1 %>%  
 filter(`Területi egység szintje` == "vármegye")  
regio<- adat1 %>%  
 filter(`Területi egység szintje` == "régió")  
nagyregio<- adat1 %>%  
 filter(`Területi egység szintje` == "nagyrégió")  
varmegye\_kulonb <- varmegye %>%  
 mutate(  
 `2011` = as.numeric(gsub("[^0-9]", "", `2011`)),   
 `2022` = as.numeric(gsub("[^0-9]", "", `2022`)),   
 Difference = `2022` - `2011`  
 ) %>%  
 select(`Területi egység neve`, `Területi egység szintje`, `2011`, `2022`, Difference) %>%  
 arrange(Difference)   
  
kable(regio)

| Területi egység neve | Területi egység szintje | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Közép-Dunántúl | régió | 1 120 730 | 1 120 610 | 1 113 671 | 1 112 984 | 1 110 897 | 1 108 124 | 1 107 453 | 1 104 841 | 1 103 132 | 1 098 654 | 1 094 104 | 1 079 589 | 1 075 027 | 1 069 800 | 1 064 521 | 1 062 306 | 1 057 816 | 1 057 449 | 1 060 306 | 1 062 982 | 1 061 979 | 1 058 909 | 1 059 399 | 1 057 303 |
| Nyugat-Dunántúl | régió | 1 003 854 | 1 002 959 | 1 004 328 | 1 003 185 | 1 000 348 | 1 000 142 | 999 361 | 997 939 | 998 187 | 996 390 | 994 698 | 984 399 | 982 654 | 979 146 | 975 754 | 973 166 | 970 910 | 971 397 | 973 587 | 977 086 | 979 459 | 979 744 | 982 924 | 981 569 |
| Dél-Dunántúl | régió | 997 671 | 993 466 | 989 408 | 983 612 | 977 465 | 970 700 | 967 677 | 960 088 | 952 982 | 947 986 | 940 585 | 931 215 | 924 900 | 916 856 | 908 400 | 900 042 | 893 263 | 885 800 | 878 402 | 873 097 | 869 726 | 863 057 | 858 675 | 852 420 |
| Észak-Magyarország | régió | 1 302 835 | 1 296 504 | 1 288 960 | 1 280 040 | 1 271 111 | 1 261 489 | 1 251 441 | 1 236 690 | 1 223 238 | 1 209 142 | 1 194 697 | 1 200 831 | 1 189 094 | 1 176 263 | 1 163 818 | 1 152 668 | 1 142 892 | 1 133 848 | 1 125 185 | 1 117 255 | 1 111 193 | 1 101 369 | 1 093 790 | 1 083 327 |
| Észak-Alföld | régió | 1 563 714 | 1 559 073 | 1 554 177 | 1 547 003 | 1 541 818 | 1 533 162 | 1 525 317 | 1 514 020 | 1 502 409 | 1 492 502 | 1 481 922 | 1 498 795 | 1 489 460 | 1 480 007 | 1 473 223 | 1 466 389 | 1 459 130 | 1 450 191 | 1 439 991 | 1 430 692 | 1 423 489 | 1 412 540 | 1 405 012 | 1 398 574 |
| Dél-Alföld | régió | 1 380 387 | 1 373 194 | 1 367 064 | 1 360 214 | 1 354 938 | 1 347 294 | 1 342 231 | 1 334 506 | 1 325 527 | 1 318 214 | 1 308 470 | 1 296 278 | 1 287 203 | 1 276 680 | 1 266 904 | 1 257 756 | 1 246 137 | 1 237 501 | 1 230 022 | 1 223 386 | 1 214 723 | 1 206 377 | 1 200 150 | 1 191 955 |

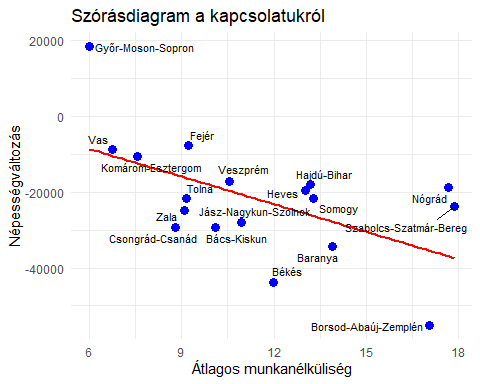
kable(varmegye\_kulonb)

| Területi egység neve | Területi egység szintje | 2011 | 2022 | Difference |
| --- | --- | --- | --- | --- |
| Borsod-Abaúj-Zemplén | vármegye | 684793 | 629621 | -55172 |
| Békés | vármegye | 361802 | 318081 | -43721 |
| Baranya | vármegye | 391455 | 357066 | -34389 |
| Csongrád-Csanád | vármegye | 421827 | 392621 | -29206 |
| Bács-Kiskun | vármegye | 524841 | 495675 | -29166 |
| Jász-Nagykun-Szolnok | vármegye | 386752 | 358883 | -27869 |
| Zala | vármegye | 287043 | 262370 | -24673 |
| Szabolcs-Szatmár-Bereg | vármegye | 555496 | 531855 | -23641 |
| Tolna | vármegye | 231183 | 209569 | -21614 |
| Somogy | vármegye | 317947 | 296422 | -21525 |
| Heves | vármegye | 307985 | 288430 | -19555 |
| Nógrád | vármegye | 201919 | 183318 | -18601 |
| Hajdú-Bihar | vármegye | 539674 | 521802 | -17872 |
| Veszprém | vármegye | 356573 | 339498 | -17075 |
| Komárom-Esztergom | vármegye | 311411 | 300869 | -10542 |
| Vas | vármegye | 257688 | 248909 | -8779 |
| Fejér | vármegye | 426120 | 418542 | -7578 |
| Győr-Moson-Sopron | vármegye | 449967 | 468465 | 18498 |

varmegye\_munka <- adat\_munka %>%  
 filter(grepl("vármegye", `Területi egység szintje`)) %>%   
 mutate(  
 Avg\_2011 = rowMeans(select(., starts\_with("2011")), na.rm = TRUE) / 10   
 ) %>%  
 select(`Területi egység neve`, `Területi egység szintje`, Avg\_2011)  
  
kable(varmegye\_munka)

| Területi egység neve | Területi egység szintje | Avg\_2011 |
| --- | --- | --- |
| Pest | vármegye, régió | 7.575 |
| Fejér | vármegye | 9.225 |
| Komárom-Esztergom | vármegye | 7.550 |
| Veszprém | vármegye | 10.550 |
| Győr-Moson-Sopron | vármegye | 6.000 |
| Vas | vármegye | 6.725 |
| Zala | vármegye | 9.100 |
| Baranya | vármegye | 13.900 |
| Somogy | vármegye | 13.300 |
| Tolna | vármegye | 9.150 |
| Borsod-Abaúj-Zemplén | vármegye | 17.075 |
| Heves | vármegye | 13.025 |
| Nógrád | vármegye | 17.675 |
| Hajdú-Bihar | vármegye | 13.200 |
| Jász-Nagykun-Szolnok | vármegye | 10.925 |
| Szabolcs-Szatmár-Bereg | vármegye | 17.875 |
| Bács-Kiskun | vármegye | 10.100 |
| Békés | vármegye | 11.975 |
| Csongrád-Csanád | vármegye | 8.800 |

library(tidyverse)  
library(ggplot2)  
library(dplyr)  
library(ggrepel)  
  
kulonb <- varmegye\_kulonb %>%  
 select(`Területi egység neve`, Difference) %>%  
 mutate(Source = "Population Change")  
  
munkat <- varmegye\_munka %>%  
 select(`Területi egység neve`, Avg\_2011) %>%  
 rename(Difference = Avg\_2011) %>%  
 mutate(Source = "Avg Employment")  
osszes1 <- bind\_rows(kulonb, munkat)  
  
kulonb2 <- osszes1 %>% filter(Source == "Population Change")  
munka2 <- osszes1 %>% filter(Source == "Avg Employment")  
  
egybe <- merge(kulonb2, munka2, by = "Területi egység neve")  
  
ggplot(egybe, aes(x = Difference.y, y = Difference.x, label = `Területi egység neve`)) +  
 geom\_point(color = "blue", size = 3) +  
 geom\_smooth(method = "lm", se = FALSE, color = "red") +  
 geom\_text\_repel(size = 3) +   
 labs(  
 title = "Szórásdiagram a kapcsolatukról",  
 x = "Átlagos munkanélküliség",  
 y = "Népességváltozás"  
 ) +  
 theme\_minimal()



A fenti ábrán ábrázoltam a népességváltozást és a munkanélküliséget, már ezen az ábrán is látszik, hogy az elemek nagy része közel fekszik el a vonalhoz.

library(tidyverse)  
library(ggplot2)  
library(dplyr)  
library(ggrepel)  
  
kulonb2 <- osszes1 %>% filter(Source == "Population Change")  
munka2 <- osszes1 %>% filter(Source == "Avg Employment")  
  
merged\_data <- merge(kulonb2, munka2, by = "Területi egység neve")  
colnames(merged\_data) <- c("Területi\_egysegek", "Population\_Change", "Source1", "Avg\_Employment", "Source2")  
correlation <- cor(merged\_data$Avg\_Employment, merged\_data$Population\_Change, use = "complete.obs")  
print(paste("Korrelációs érték: ", round(correlation, 2)))

## [1] "Korrelációs érték: -0.56"

A kapott korrelációs érték −0.56 azt mutatja, hogy mérsékelt negatív kapcsolat áll fenn az átlagos foglalkoztatás és a népességváltozás között. Ez azt jelenti, hogy a magasabb foglalkoztatási szintek általában kisebb (vagy akár pozitív) népességváltozással járnak, míg az alacsonyabb foglalkoztatottságú területeken nagyobb népességcsökkenés figyelhető meg.

Annak ellenére, hogy a korreláció nem feltétlenül jelenti azt, hogy egyértelmű kapcsolat van, ez az eredmény arra utalhat, hogy a jobb foglalkoztatási lehetőségek hozzájárulhatnak a népesség megtartásához, a magas munkanélküliséggel küzdő területekről meg nagyobb elvándorlás jelentkezhet.

regression\_model <- lm(Population\_Change ~ Avg\_Employment, data = merged\_data)  
summary(regression\_model)

##   
## Call:  
## lm(formula = Population\_Change ~ Avg\_Employment, data = merged\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20651 -8265 1683 7641 27084   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5958.4 10671.8 0.558 0.5843   
## Avg\_Employment -2424.1 891.2 -2.720 0.0151 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 13210 on 16 degrees of freedom  
## Multiple R-squared: 0.3162, Adjusted R-squared: 0.2734   
## F-statistic: 7.398 on 1 and 16 DF, p-value: 0.01514

A fenti regressziós modell azt mutatja, hogy ahol magasabb az átlagos foglalkoztatottság, ott kisebb a népességcsökkenés. A negatív összefüggés (-242.41) azt jelzi, hogy minden egyes foglalkoztatási egység növekedése átlagosan 242 fővel csökkenti a népességcsökkenést. Bár az összefüggés statisztikailag szignifikáns (p = 0.015), a modell csupán a népességváltozás körülbelül ~32%-át magyarázza, tehát más tényezők is befolyásolják a helyzetet.

## Utószó:

Nyilvánvalóan végső következtetéseket nehéz levonni egy ilyen rövid dolgozat alatt, azonban máris felfedezhető a szoros kapcsolat a tényezők között. A munkanélküliség arra kényszeríti sok esetben az embereket, hogy elvándoroljanak más területekre, legyen az belföld (pl.: pozitív Győr-Moson-Sopron) vagy akár külföld. Sokszor a munkanélküliséggel küzdő vármegyék más problémákkal is küzdhetnek, legyen az oktatás, vagy infrastruktúra szinten.