

Assignment 1 — MSB104 — Group 3

Irjan & Magnus

Part A_ Sub-national GDP and GDP per capita

Data aquisition

```
Sys.setlocale("LC_ALL", "Norwegian")
```

```
[1] "LC_COLLATE=Norwegian_Norway.1252;LC_CTYPE=Norwegian_Norway.1252;LC_MONETARY=Norwegian_N
```

```
library(readxl)
library(dplyr)
library(tidyr)
library(janitor)

# Henter inn populasjons datasett fra excell
Populasjon <- read_excel("DEMO_Ass1.xlsx", sheet = 2, col_types = "text") %>%
  clean_names()

# Henter inn regional BNP datasett fra excell
BNP <- read_excel("GDP_Ass1.xlsx", sheet = 2, col_types = "text") %>%
  clean_names()
```

```
# Omgjør BNP og populasjonen til langt format
```

```
PopulasjonLang <- Populasjon %>%
  pivot_longer(
    cols = starts_with("x"),
    names_to = "aar",
    values_to = "befolknинг"
```

```

) %>%
mutate(
  aar = as.integer(str_remove(aar, "^x")),
  befolkning = as.numeric(befolkning)
)

BNPLang <- BNP %>%
pivot_longer(
  cols = starts_with("x"),
  names_to = "aar",
  values_to = "BNP"
) %>%
mutate(
  aar = as.integer(str_remove(aar, "^x")),
  BNP = as.numeric(BNP)
)

```

Kort gjennomgang av datasett og variabler

Datasette *demo_r_pjanggr3* som er hentet fra Eurostat inneholder årlige befolkningestimer på NUTS3-nivå for EU-, EFTA- og kandidatland. Variablene **values** viser totalt antall bosatte personer per 1. januar, målt i antall personer. Hver observasjon identifiseres ved regionkode(**geo**) og år (**time**), sammt kjønn (**sex**) og alder (**age**). I denne analysen benyttes kun total befolkning (**sex = T, age = TOTAL**), slik at dataene ikke er splittet etter kjønn eller alder.

GDP per Capita Calculation

```

# Kombinerer tabellene og filtrerer bort NA-verdiene

Kombinert <- PopulasjonLang %>%
  left_join(BNPLang, by = join_by(region, aar)) %>%
  mutate(bnp_per_innbygger = (BNP * 1e6) / befolkning)

Kombinert <- Kombinert %>%
  filter(!is.na(bnp_per_innbygger))

```

Descriptive Analysis

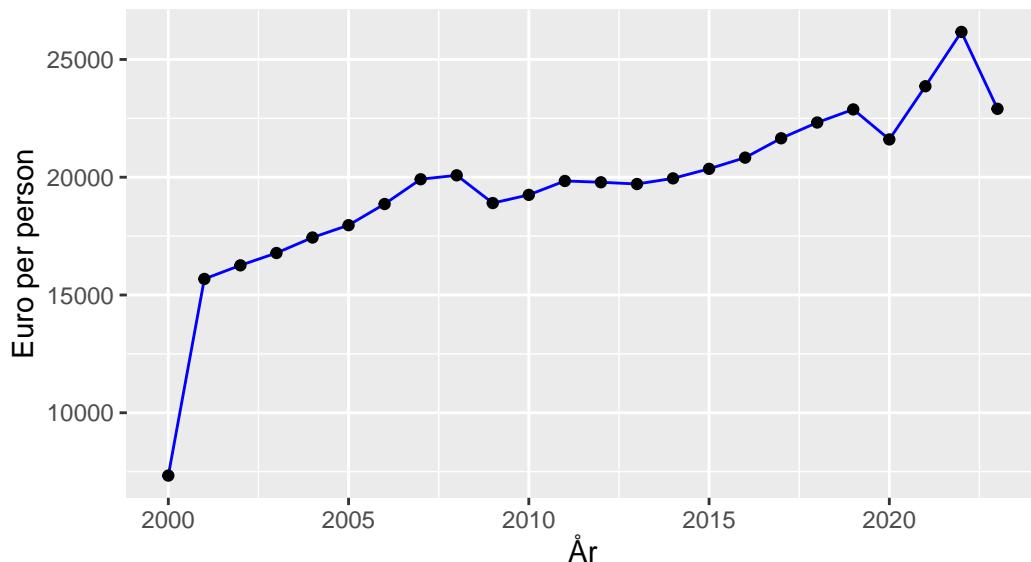
```
# A tibble: 1 x 5
  gjennomsnitt median standardavvik minimum maksimum
  <dbl>     <dbl>      <dbl>    <dbl>    <dbl>
1 19883.   20462.     11386.    825.    66633.
```

Basert på Eurostat-dataene, er BNP per innbygger beregnet som total BNP delt på antall innbyggere i hver NUTS3-region. Datasettet dekker perioden 2000-2023 for Italia, Romania, Finland og Litauen. BNP-verdiene er målt i millioner euro og er konvertert til euro per person for å gjøre regionene sammenlignbare både mellom og innenfor landene.

Gjennomsnittlig BNP per innbygger i utvalget er **19 883 euro**, med et standardavvik på **11 386 euro**, noe som viser en tydelig variasjon i økonomisk velstand mellom regionene. Minimumsverdien er **825 euro**, mens maksimum når opp mot **66 633 euro**, som reflekterer store forskjeller mellom utviklede regioner i Italia og Finland sammenlignet med mindre utviklede områder i Romania og Litauen. Medianverdien på **20 462 euro** ligger nær gjennomsnittet, noe som indikerer en tilnærmet symmetrisk fordeling, men noen enkelte regioner med svært høyt BNP per innbygger trekker gjennomsnittet litt opp.

Over tid viser dataene en jevn vekst i BNP per innbygger i alle fire land. Finland og Italia hadde de høyeste nivåene gjennom hele perioden, mens Romania og Litauen viser sterkt vekst og gradvis konvergens mot de rike landene etter 2010. Figurene nedenfor illustrerer både tidsutviklingen i gjennomsnittlig BNP per innbygger og fordelingen mellom regionene i det siste observasjonsåret, som danner grunnlaget for videre analyse av regional ulikhet i del B.

Utvikling i gjennomsnittlig BNP per innbygger (2000–2023)



Kilde: Eurostat – nama_10r_3gdp og demo_r_pjanaggr3

```
# Fordeling per land

# Samler alle regionene basert på de 2 første bokstavene i geo variabelen for å samle alle regionene
Kombinert <- Kombinert %>%
  mutate(land = substr(geo, 1, 2))

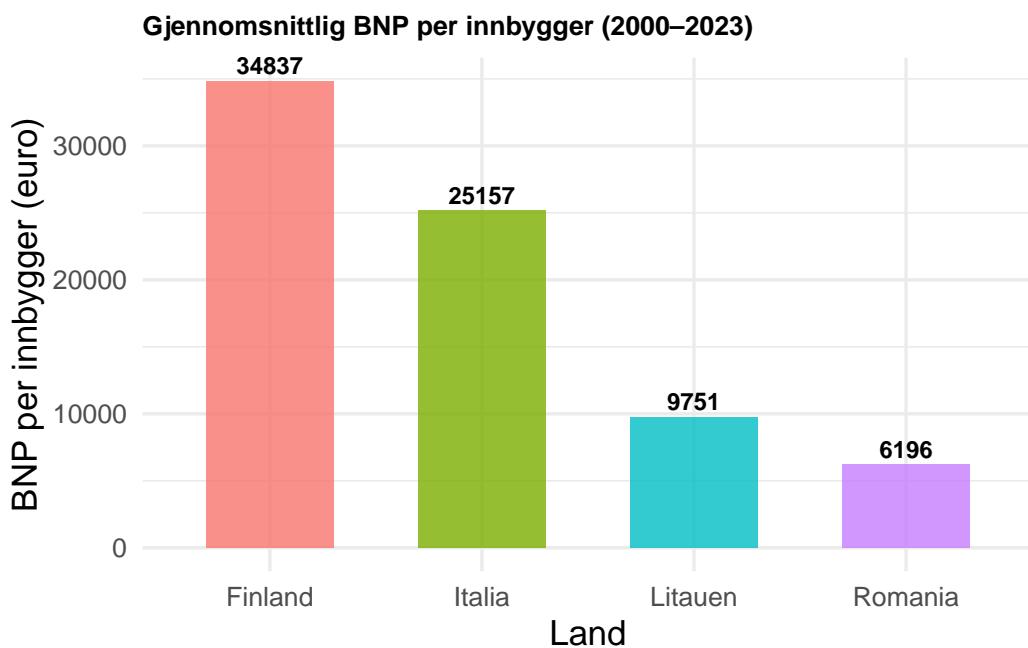
# Beregner gjennomsnittlig BNP per innbygger per land
SnittLand <- Kombinert %>%
  group_by(land) %>%
  summarise(gjennomsnitt_bnp = mean(bnp_per_innbygger, na.rm = TRUE))

# Lager en til variabel slik at det blir enklere å lese av grafen
SnittLand <- SnittLand %>%
  mutate(
    land_navn = recode(land,
      "IT" = "Italia",
      "FI" = "Finland",
      "LT" = "Litauen",
      "RO" = "Romania"
    )
  )
```

```

ggplot(SnittLand, aes(x = land_navn, y = gjennomsnitt_bnp, fill = land_navn)) +
  geom_col(alpha = 0.8, width = 0.6) +
  geom_text(aes(label = round(gjennomsnitt_bnp, 0)),
            vjust = -0.4, size = 3, fontface = "bold") +
  labs(
    title = "Gjennomsnittlig BNP per innbygger (2000–2023)",
    x = "Land",
    y = "BNP per innbygger (euro)"
  ) +
  theme_minimal(base_size = 13) +
  theme(
    legend.position = "none",
    plot.title = element_text(face = "bold", size = 10)
  )

```



```

library(tidyverse)
library(ggrepel)

# 1. Beregn gjennomsnitt per land og år
SnittTidLand <- Kombinert %>%
  group_by(land, aar) %>%
  summarise(
    gjennomsnitt_bnp = mean(bnp),
    gjennomsnitt_tid = mean(tid)
  )

```

```

gjennomsnitt_bnp = mean(bnp_per_innbygger, na.rm = TRUE),
  .groups = "drop"
)

# 2. Bytt ut landkodene med fulle navn
SnittTidLand <- SnittTidLand %>%
  mutate(
    land_navn = recode(land,
      "IT" = "Italia",
      "FI" = "Finland",
      "LT" = "Litauen",
      "RO" = "Romania"
    )
  )

# 3. Finn siste observasjon for etikettpllassering
etiketter <- SnittTidLand %>%
  group_by(land_navn) %>%
  filter(aar == max(aar))

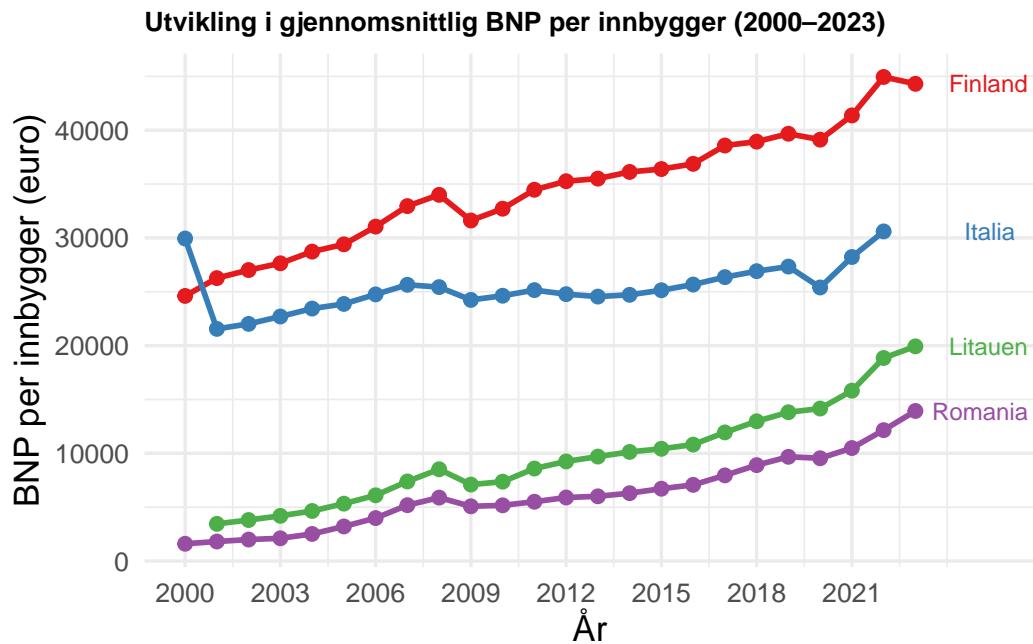
# 4. Lag linjediagram med etiketter
ggplot(SnittTidLand, aes(x = aar, y = gjennomsnitt_bnp, color = land_navn)) +
  geom_line(linewidth = 1) +
  geom_point(size = 2) +
  geom_text_repel(
    data = etiketter,
    aes(label = land_navn),
    nudge_x = 2.5,
    size = 3,
    direction = "y",
    hjust = 0,
    segment.color = NA
  ) +
  scale_x_continuous(
    breaks = seq(2000, 2023, 3)
  ) +
  scale_color_brewer(palette = "Set1") +
  labs(
    title = "Utvikling i gjennomsnittlig BNP per innbygger (2000-2023)",
    x = "År",
    y = "BNP per innbygger (euro)",
    color = NULL

```

```

) +
theme_minimal(base_size = 13) +
theme(
  plot.title = element_text(face = "bold", size = 10),
  legend.position = "none",
)

```



Part B: Regional Inequity

Litterature Review

- Read the paper by Lessmann & Seidel (2017). You will find a copy under Filer/pdf at the Canvas site of the course. Give a short summary of the paper (200 – 400 words).

Gini Coefficient Calculation

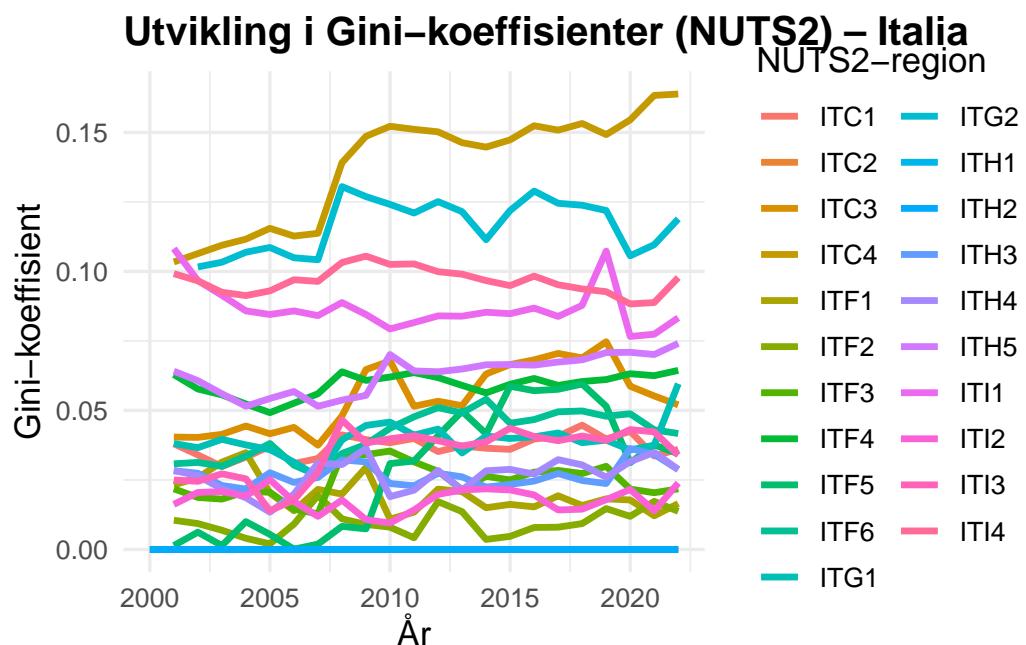
```

# A tibble: 4 x 2
  land  antall_nuts2
  <chr>    <int>
1 FI          5

```

2 IT	21
3 LT	2
4 RO	8

```
# A tibble: 464 x 3
  nuts2   aar   gini
  <chr> <int> <dbl>
1 ITC1    2001 0.038
2 ITC1    2002 0.034
3 ITC1    2003 0.0298
4 ITC1    2004 0.0332
5 ITC1    2005 0.0371
6 ITC1    2006 0.0309
7 ITC1    2007 0.0327
8 ITC1    2008 0.0412
9 ITC1    2009 0.0395
10 ITC1   2010 0.0384
# i 454 more rows
```



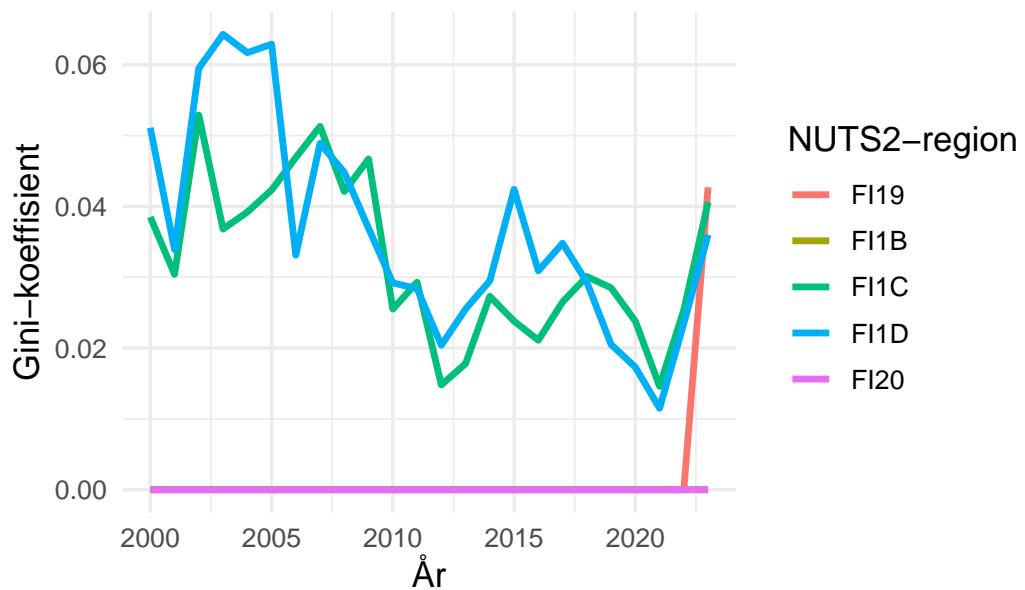
```
# A tibble: 120 x 3
  nuts2   aar   gini
  <chr> <int> <dbl>
```

```

1 FI19 2000 0
2 FI19 2001 0
3 FI19 2002 0
4 FI19 2003 0
5 FI19 2004 0
6 FI19 2005 0
7 FI19 2006 0
8 FI19 2007 0
9 FI19 2008 0
10 FI19 2009 0
# i 110 more rows

```

Utvikling i Gini-koeffisienter (NUTS2) – Finland



```

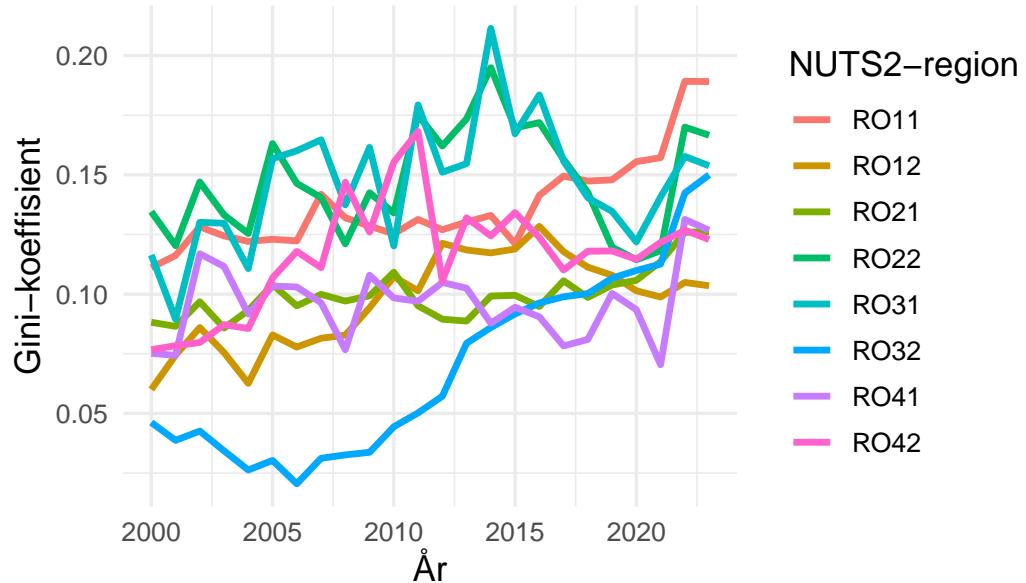
# A tibble: 192 x 3
  nuts2   aar   gini
  <chr> <int> <dbl>
1 R011   2000 0.111
2 R011   2001 0.116
3 R011   2002 0.128
4 R011   2003 0.124
5 R011   2004 0.122
6 R011   2005 0.123
7 R011   2006 0.122
8 R011   2007 0.122
9 R011   2008 0.122
10 R011  2009 0.122
11 R011  2010 0.122
12 R011  2011 0.122
13 R011  2012 0.122
14 R011  2013 0.122
15 R011  2014 0.122
16 R011  2015 0.122
17 R011  2016 0.122
18 R011  2017 0.122
19 R011  2018 0.122
20 R011  2019 0.122
21 R011  2020 0.122
22 R011  2021 0.122
23 R011  2022 0.122
24 R012   2000 0.111
25 R012   2001 0.116
26 R012   2002 0.128
27 R012   2003 0.124
28 R012   2004 0.122
29 R012   2005 0.123
30 R012   2006 0.122
31 R012   2007 0.122
32 R012   2008 0.122
33 R012   2009 0.122
34 R012   2010 0.122
35 R012   2011 0.122
36 R012   2012 0.122
37 R012   2013 0.122
38 R012   2014 0.122
39 R012   2015 0.122
40 R012   2016 0.122
41 R012   2017 0.122
42 R012   2018 0.122
43 R012   2019 0.122
44 R012   2020 0.122
45 R012   2021 0.122
46 R012   2022 0.122
47 R013   2000 0.111
48 R013   2001 0.116
49 R013   2002 0.128
50 R013   2003 0.124
51 R013   2004 0.122
52 R013   2005 0.123
53 R013   2006 0.122
54 R013   2007 0.122
55 R013   2008 0.122
56 R013   2009 0.122
57 R013   2010 0.122
58 R013   2011 0.122
59 R013   2012 0.122
60 R013   2013 0.122
61 R013   2014 0.122
62 R013   2015 0.122
63 R013   2016 0.122
64 R013   2017 0.122
65 R013   2018 0.122
66 R013   2019 0.122
67 R013   2020 0.122
68 R013   2021 0.122
69 R013   2022 0.122
70 R014   2000 0.111
71 R014   2001 0.116
72 R014   2002 0.128
73 R014   2003 0.124
74 R014   2004 0.122
75 R014   2005 0.123
76 R014   2006 0.122
77 R014   2007 0.122
78 R014   2008 0.122
79 R014   2009 0.122
80 R014   2010 0.122
81 R014   2011 0.122
82 R014   2012 0.122
83 R014   2013 0.122
84 R014   2014 0.122
85 R014   2015 0.122
86 R014   2016 0.122
87 R014   2017 0.122
88 R014   2018 0.122
89 R014   2019 0.122
90 R014   2020 0.122
91 R014   2021 0.122
92 R014   2022 0.122
93 R015   2000 0.111
94 R015   2001 0.116
95 R015   2002 0.128
96 R015   2003 0.124
97 R015   2004 0.122
98 R015   2005 0.123
99 R015   2006 0.122
100 R015   2007 0.122
101 R015   2008 0.122
102 R015   2009 0.122
103 R015   2010 0.122
104 R015   2011 0.122
105 R015   2012 0.122
106 R015   2013 0.122
107 R015   2014 0.122
108 R015   2015 0.122
109 R015   2016 0.122
110 R015   2017 0.122
111 R015   2018 0.122
112 R015   2019 0.122
113 R015   2020 0.122
114 R015   2021 0.122
115 R015   2022 0.122
116 R016   2000 0.111
117 R016   2001 0.116
118 R016   2002 0.128
119 R016   2003 0.124
120 R016   2004 0.122
121 R016   2005 0.123
122 R016   2006 0.122
123 R016   2007 0.122
124 R016   2008 0.122
125 R016   2009 0.122
126 R016   2010 0.122
127 R016   2011 0.122
128 R016   2012 0.122
129 R016   2013 0.122
130 R016   2014 0.122
131 R016   2015 0.122
132 R016   2016 0.122
133 R016   2017 0.122
134 R016   2018 0.122
135 R016   2019 0.122
136 R016   2020 0.122
137 R016   2021 0.122
138 R016   2022 0.122
139 R017   2000 0.111
140 R017   2001 0.116
141 R017   2002 0.128
142 R017   2003 0.124
143 R017   2004 0.122
144 R017   2005 0.123
145 R017   2006 0.122
146 R017   2007 0.122
147 R017   2008 0.122
148 R017   2009 0.122
149 R017   2010 0.122
150 R017   2011 0.122
151 R017   2012 0.122
152 R017   2013 0.122
153 R017   2014 0.122
154 R017   2015 0.122
155 R017   2016 0.122
156 R017   2017 0.122
157 R017   2018 0.122
158 R017   2019 0.122
159 R017   2020 0.122
160 R017   2021 0.122
161 R017   2022 0.122
162 R018   2000 0.111
163 R018   2001 0.116
164 R018   2002 0.128
165 R018   2003 0.124
166 R018   2004 0.122
167 R018   2005 0.123
168 R018   2006 0.122
169 R018   2007 0.122
170 R018   2008 0.122
171 R018   2009 0.122
172 R018   2010 0.122
173 R018   2011 0.122
174 R018   2012 0.122
175 R018   2013 0.122
176 R018   2014 0.122
177 R018   2015 0.122
178 R018   2016 0.122
179 R018   2017 0.122
180 R018   2018 0.122
181 R018   2019 0.122
182 R018   2020 0.122
183 R018   2021 0.122
184 R018   2022 0.122
185 R019   2000 0.111
186 R019   2001 0.116
187 R019   2002 0.128
188 R019   2003 0.124
189 R019   2004 0.122
190 R019   2005 0.123
191 R019   2006 0.122
192 R019   2007 0.122
193 R019   2008 0.122
194 R019   2009 0.122
195 R019   2010 0.122
196 R019   2011 0.122
197 R019   2012 0.122
198 R019   2013 0.122
199 R019   2014 0.122
200 R019   2015 0.122
201 R019   2016 0.122
202 R019   2017 0.122
203 R019   2018 0.122
204 R019   2019 0.122
205 R019   2020 0.122
206 R019   2021 0.122
207 R019   2022 0.122
208 R020   2000 0.111
209 R020   2001 0.116
210 R020   2002 0.128
211 R020   2003 0.124
212 R020   2004 0.122
213 R020   2005 0.123
214 R020   2006 0.122
215 R020   2007 0.122
216 R020   2008 0.122
217 R020   2009 0.122
218 R020   2010 0.122
219 R020   2011 0.122
220 R020   2012 0.122
221 R020   2013 0.122
222 R020   2014 0.122
223 R020   2015 0.122
224 R020   2016 0.122
225 R020   2017 0.122
226 R020   2018 0.122
227 R020   2019 0.122
228 R020   2020 0.122
229 R020   2021 0.122
230 R020   2022 0.122
231 R021   2000 0.111
232 R021   2001 0.116
233 R021   2002 0.128
234 R021   2003 0.124
235 R021   2004 0.122
236 R021   2005 0.123
237 R021   2006 0.122
238 R021   2007 0.122
239 R021   2008 0.122
240 R021   2009 0.122
241 R021   2010 0.122
242 R021   2011 0.122
243 R021   2012 0.122
244 R021   2013 0.122
245 R021   2014 0.122
246 R021   2015 0.122
247 R021   2016 0.122
248 R021   2017 0.122
249 R021   2018 0.122
250 R021   2019 0.122
251 R021   2020 0.122
252 R021   2021 0.122
253 R021   2022 0.122
254 R022   2000 0.111
255 R022   2001 0.116
256 R022   2002 0.128
257 R022   2003 0.124
258 R022   2004 0.122
259 R022   2005 0.123
260 R022   2006 0.122
261 R022   2007 0.122
262 R022   2008 0.122
263 R022   2009 0.122
264 R022   2010 0.122
265 R022   2011 0.122
266 R022   2012 0.122
267 R022   2013 0.122
268 R022   2014 0.122
269 R022   2015 0.122
270 R022   2016 0.122
271 R022   2017 0.122
272 R022   2018 0.122
273 R022   2019 0.122
274 R022   2020 0.122
275 R022   2021 0.122
276 R022   2022 0.122
277 R023   2000 0.111
278 R023   2001 0.116
279 R023   2002 0.128
280 R023   2003 0.124
281 R023   2004 0.122
282 R023   2005 0.123
283 R023   2006 0.122
284 R023   2007 0.122
285 R023   2008 0.122
286 R023   2009 0.122
287 R023   2010 0.122
288 R023   2011 0.122
289 R023   2012 0.122
290 R023   2013 0.122
291 R023   2014 0.122
292 R023   2015 0.122
293 R023   2016 0.122
294 R023   2017 0.122
295 R023   2018 0.122
296 R023   2019 0.122
297 R023   2020 0.122
298 R023   2021 0.122
299 R023   2022 0.122
300 R024   2000 0.111
301 R024   2001 0.116
302 R024   2002 0.128
303 R024   2003 0.124
304 R024   2004 0.122
305 R024   2005 0.123
306 R024   2006 0.122
307 R024   2007 0.122
308 R024   2008 0.122
309 R024   2009 0.122
310 R024   2010 0.122
311 R024   2011 0.122
312 R024   2012 0.122
313 R024   2013 0.122
314 R024   2014 0.122
315 R024   2015 0.122
316 R024   2016 0.122
317 R024   2017 0.122
318 R024   2018 0.122
319 R024   2019 0.122
320 R024   2020 0.122
321 R024   2021 0.122
322 R024   2022 0.122
323 R025   2000 0.111
324 R025   2001 0.116
325 R025   2002 0.128
326 R025   2003 0.124
327 R025   2004 0.122
328 R025   2005 0.123
329 R025   2006 0.122
330 R025   2007 0.122
331 R025   2008 0.122
332 R025   2009 0.122
333 R025   2010 0.122
334 R025   2011 0.122
335 R025   2012 0.122
336 R025   2013 0.122
337 R025   2014 0.122
338 R025   2015 0.122
339 R025   2016 0.122
340 R025   2017 0.122
341 R025   2018 0.122
342 R025   2019 0.122
343 R025   2020 0.122
344 R025   2021 0.122
345 R025   2022 0.122
346 R026   2000 0.111
347 R026   2001 0.116
348 R026   2002 0.128
349 R026   2003 0.124
350 R026   2004 0.122
351 R026   2005 0.123
352 R026   2006 0.122
353 R026   2007 0.122
354 R026   2008 0.122
355 R026   2009 0.122
356 R026   2010 0.122
357 R026   2011 0.122
358 R026   2012 0.122
359 R026   2013 0.122
360 R026   2014 0.122
361 R026   2015 0.122
362 R026   2016 0.122
363 R026   2017 0.122
364 R026   2018 0.122
365 R026   2019 0.122
366 R026   2020 0.122
367 R026   2021 0.122
368 R026   2022 0.122
369 R027   2000 0.111
370 R027   2001 0.116
371 R027   2002 0.128
372 R027   2003 0.124
373 R027   2004 0.122
374 R027   2005 0.123
375 R027   2006 0.122
376 R027   2007 0.122
377 R027   2008 0.122
378 R027   2009 0.122
379 R027   2010 0.122
380 R027   2011 0.122
381 R027   2012 0.122
382 R027   2013 0.122
383 R027   2014 0.122
384 R027   2015 0.122
385 R027   2016 0.122
386 R027   2017 0.122
387 R027   2018 0.122
388 R027   2019 0.122
389 R027   2020 0.122
390 R027   2021 0.122
391 R027   2022 0.122
392 R028   2000 0.111
393 R028   2001 0.116
394 R028   2002 0.128
395 R028   2003 0.124
396 R028   2004 0.122
397 R028   2005 0.123
398 R028   2006 0.122
399 R028   2007 0.122
400 R028   2008 0.122
401 R028   2009 0.122
402 R028   2010 0.122
403 R028   2011 0.122
404 R028   2012 0.122
405 R028   2013 0.122
406 R028   2014 0.122
407 R028   2015 0.122
408 R028   2016 0.122
409 R028   2017 0.122
410 R028   2018 0.122
411 R028   2019 0.122
412 R028   2020 0.122
413 R028   2021 0.122
414 R028   2022 0.122
415 R029   2000 0.111
416 R029   2001 0.116
417 R029   2002 0.128
418 R029   2003 0.124
419 R029   2004 0.122
420 R029   2005 0.123
421 R029   2006 0.122
422 R029   2007 0.122
423 R029   2008 0.122
424 R029   2009 0.122
425 R029   2010 0.122
426 R029   2011 0.122
427 R029   2012 0.122
428 R029   2013 0.122
429 R029   2014 0.122
430 R029   2015 0.122
431 R029   2016 0.122
432 R029   2017 0.122
433 R029   2018 0.122
434 R029   2019 0.122
435 R029   2020 0.122
436 R029   2021 0.122
437 R029   2022 0.122
438 R030   2000 0.111
439 R030   2001 0.116
440 R030   2002 0.128
441 R030   2003 0.124
442 R030   2004 0.122
443 R030   2005 0.123
444 R030   2006 0.122
445 R030   2007 0.122
446 R030   2008 0.122
447 R030   2009 0.122
448 R030   2010 0.122
449 R030   2011 0.122
450 R030   2012 0.122
451 R030   2013 0.122
452 R030   2014 0.122
453 R030   2015 0.122
454 R030   2016 0.122
455 R030   2017 0.122
456 R030   2018 0.122
457 R030   2019 0.122
458 R030   2020 0.122
459 R030   2021 0.122
460 R030   2022 0.122
461 R031   2000 0.111
462 R031   2001 0.116
463 R031   2002 0.128
464 R031   2003 0.124
465 R031   2004 0.122
466 R031   2005 0.123
467 R031   2006 0.122
468 R031   2007 0.122
469 R031   2008 0.122
470 R031   2009 0.122
471 R031   2010 0.122
472 R031
```

```

8 RO11    2007 0.142
9 RO11    2008 0.132
10 RO11   2009 0.128
# i 182 more rows

```

Utvikling i Gini-koeffisienter (NUTS2) – Romania

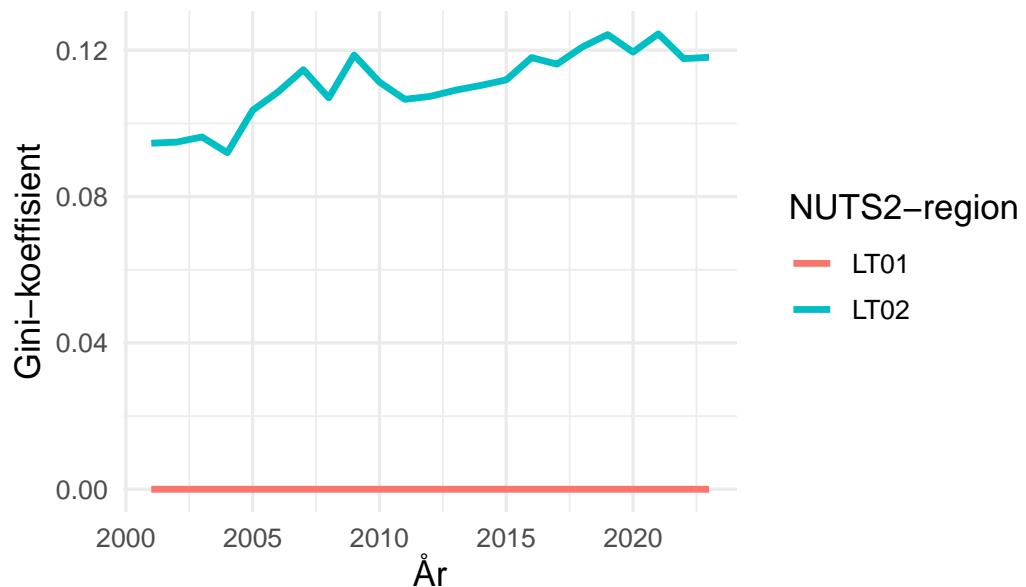


```

# A tibble: 46 x 3
  nuts2   aar   gini
  <chr> <int> <dbl>
1 LT01    2001    0
2 LT01    2002    0
3 LT01    2003    0
4 LT01    2004    0
5 LT01    2005    0
6 LT01    2006    0
7 LT01    2007    0
8 LT01    2008    0
9 LT01    2009    0
10 LT01   2010    0
# i 36 more rows

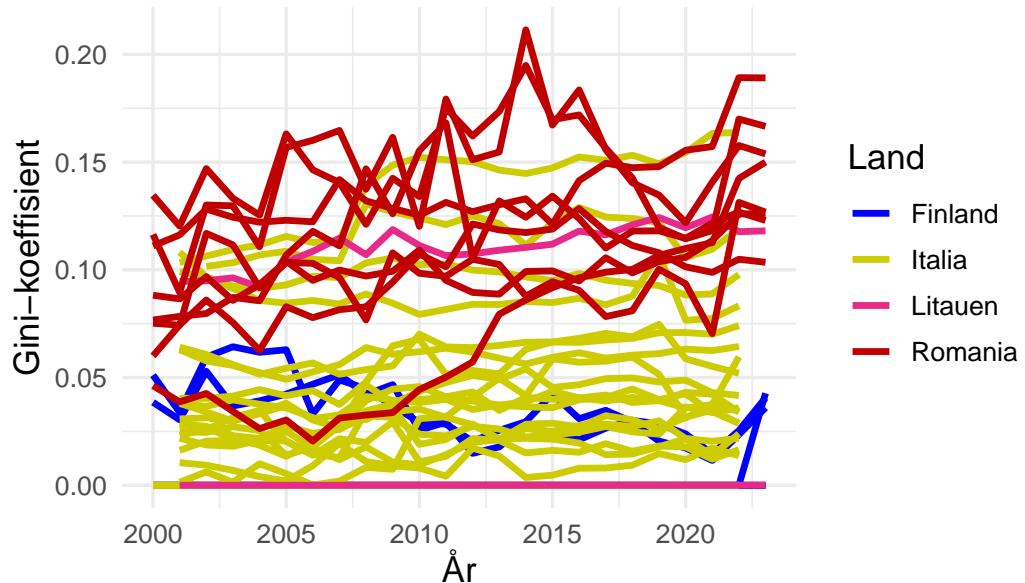
```

Utvikling i Gini-koeffisienter (NUTS2) – Litauen



```
# A tibble: 6 x 4
nuts2   aar   gini land
<chr> <int> <dbl> <chr>
1 ITC1    2001 0.038 Italia
2 ITC1    2002 0.034 Italia
3 ITC1    2003 0.0298 Italia
4 ITC1    2004 0.0332 Italia
5 ITC1    2005 0.0371 Italia
6 ITC1    2006 0.0309 Italia
```

Utvikling i Gini-koeffisienter (NUTS2)



```
# A tibble: 4 x 5
  land    gjennomsnitt_gini sd_gini min_gini max_gini
  <chr>          <dbl>   <dbl>    <dbl>    <dbl>
1 Romania        0.114    0.035    0.0205   0.211
2 Litauen        0.111    0.0095   0.092    0.124
3 Italia          0.051    0.0364   0.0001   0.164
4 Finland        0.0346   0.0132   0.0115   0.0643
```

Gjennomsnittlig Gini–koeffisient per land (2000–1)

Beregnet på NUTS2-nivå, uten nullverdier

