Eurostat EDA

Sindre H. Øveraas, Alen Colakovic & Sebastian M. Fløysand

Eurostat EDA

Meta-data description

Nama 10r

Regional accounts are based on the same definitions and concepts as national accounts, this is because they are a specification of the national accounts.

Gross domestic product (GDP) is the standard measure of the value added created through the production of goods and services in a country during a certain period. As such, it also measures the income earned from that production, or the total amount spent on final goods and services (less imports). While GDP is the single most important indicator to capture economic activity, it falls short of providing a suitable measure of people's material well-being for which alternative indicators may be more appropriate.
@GDPSpendingGross

Our goal with this task is to compare the GDP´s of the different regions within nations. We will dig deeper into the evelution of the GDP to each of our given nations aswell as some of their most intriguing regions.

Sub-national GDP

Country codes: BE - Belgium, BG - Bulgaria, HR - Croatia, IT - Italy, AT - Austria, SE - Sweden, RS - Serbia

These are the RStudio packages we install and load to do our calculations.

```
library(tidyverse)
library(vtable)
library(dineq)
library(dplyr)
```

We downloaded from Eurostat our subset of countries sub-regional GDP (nama_10r_3gdp___custom_356493 and population (demo_r_pjanaggr3___custom_3579517_linear.csv) for the years

2000 - 2020. We then calculated the GDP per capita, and named the new data set "GDP_Per_Capita".

```r
library(readr)
GDP <- read.csv('nama_10r_3gdp__custom_3564935_linear.csv')
Population <- read_csv("demo_r_pjanaggr3__custom_3579517_linear.csv")

gdpdata <- GDP %>%
  rename(Year = TIME_PERIOD, GDP = OBS_VALUE, Region = geo)

populationdata <- Population %>%
  rename(Year = TIME_PERIOD, Population = OBS_VALUE, Region = geo)

GDP_Per_Capita <- gdpdata %>%
  left_join(populationdata, by=c("Region", "Year")) %>%
  select(Region, Year, GDP, Population) %>%
  mutate(
    GDP_capita = (GDP * 1000000)/Population)
```

To report our descriptive statistics on GDP per capita, we used the summary command. This gave us the following descriptive statistics on the variables GDP, population and GDP capita.

```
      GDP                  Population          GDP_capita
Min.   :      74.55   Min.   :  20320   Min.   : 1087
1st Qu.:   1738.28   1st Qu.: 164518   1st Qu.:17180
Median :   5614.05   Median : 273920   Median :25185
Mean   :  10238.24   Mean   : 406217   Mean   :24191
3rd Qu.:  10640.23   3rd Qu.: 429030   3rd Qu.:31351
Max.   :181212.88   Max.   :4355725   Max.   :72062
                     NA's   :771        NA's   :771
```

To calculate the population watertight GDP Ginie coefficient, we used the following command.

```r
gini.wtd(GDP_Per_Capita$GDP_capita, weights = GDP_Per_Capita$Population)
```

```
[1] 0.2603924
```

To calculate the population watertight GDP Ginie coefficents for the European NUTS2 level, we used the following commands.

```
GDP_Per_Capita <- GDP_Per_Capita %>%
  mutate(NUTS2 = substr(GDP_Per_Capita$Region,1,4))

GDP_Per_Capita <- GDP_Per_Capita %>%
  mutate(NUTS = substr(GDP_Per_Capita$Region,1,2))

GDP_Per_Capita  %<>%
  group_by(NUTS2) %>%
  na.exclude(GDP_Per_Capita) %>%
  mutate(GINI_NUTS2 = gini.wtd(GDP_capita, weights = Population)) %>%
  ungroup()
```
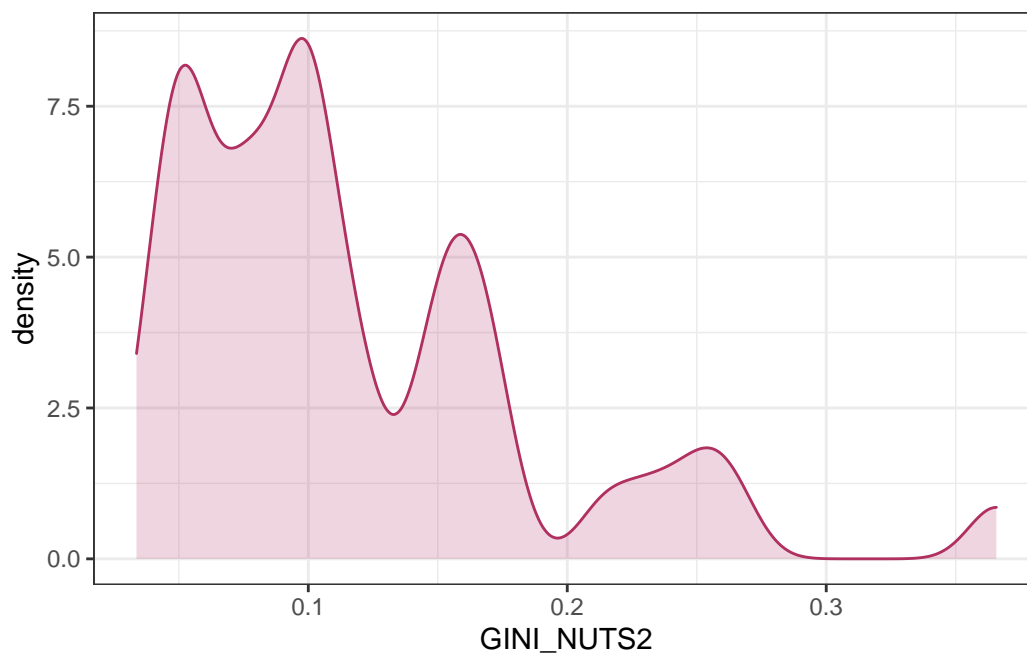
To summarize the data of the Ginie coefficients, we once more used the summarize command.
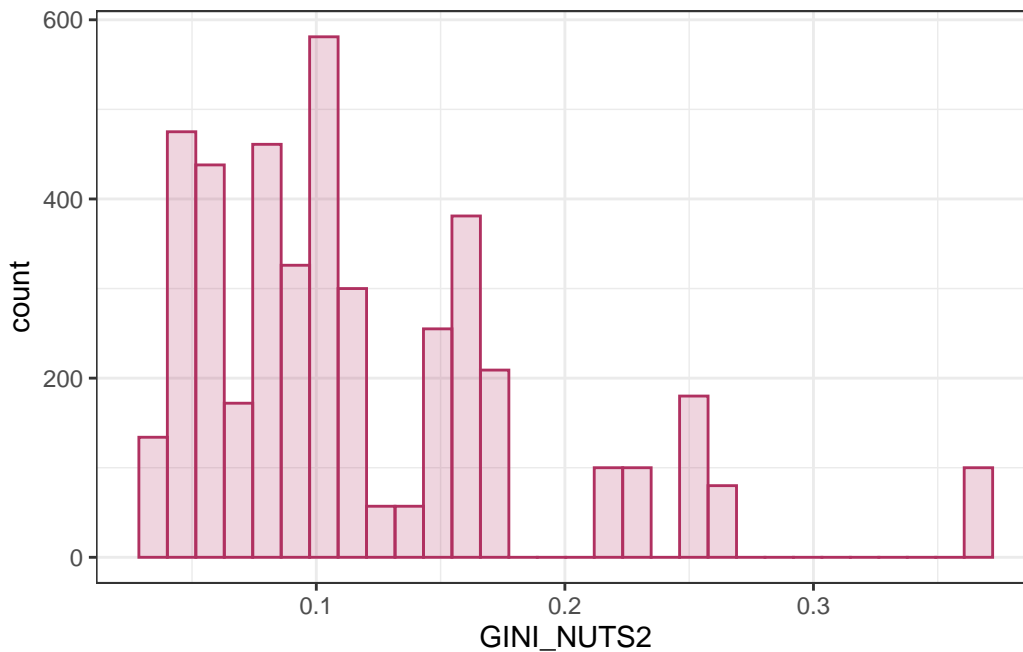
```
  GINI_NUTS2
Min.    :0.03367
1st Qu.:0.07065
Median :0.09839
Mean    :0.11800
3rd Qu.:0.15525
Max.    :0.36569
```

For visualization of our data (the Ginie coefficients for the European NUTS2 level of our selected countries), we produced a density plot and histogram, by the use of the ggplot2 package.

Looking at the plot above there is one outlier up against 0.4 with around 100 observations. The same result also seem to occur in the density plot.

We made a variety of diagrams to show the difference between some countries and regions.

For this diagram we picked and divided the countries in to their respected sub-regions and catagorized them, before comparing.

Central and Eastern Europe: Croatia, Serbia and Bulgaria

South Europe: Italy

Northern Europe: Sweden

Western Europe: Austria and Belgium

(Catagorized collected from wikipedia : (**FileEuropeanSubregions?**))

With this diagram, we can see the GDP of the different areas and their rise and fall from the 2000 to 2020.

explain what we see

Further on we wish to highlight some of the regions within countries to see differences in GDP. There are some countries with wider gaps between them, and makes this a insightful visual representation. We dig deeper into Sweden where we used the regions: Stockholm, Uppsala and Gotland. With these we se some big gaps when it comes to GDP and how Stockholm alone raises the country´s collective GDP. In the diagram below we also show the GDP Per Capita in each region.*** The simple reason behind the difference is the location of Stockholm and that it is the capital of sweden, as well as being one of the major industrial cities in Sweden also called industrial cluster area.

Do more calulations on differences on regional levels and give further explanations

**** show 2 diagrams of gdp aswell as gdp per capita ****