

A comparison of UN and Gapminder on refugee statistics

Lausanne, 2022.04.15

15 min reading

In the last decade, the number of refugees has been increasing dramatically, as shows the chart below from the United Nations Refugee Agency:

Data often gives important insights into the world. Nevertheless, it is important to keep in mind that data does not properly convey values such as human life and it is very hard to measure emotions like happiness or suffering. To address complex problems like this an important question is often how large is this situation? What is the share of persons affected?

A simple search brings us to gapminder.org which is often good source of data and data analysis. This organization claims trying to fight misconceptions and provide a factual view of the world. On this topic they've issued a study stating that the share of refugees in the world is very small and that the effort to support is smaller than most people believe: Gapminder on refugees

Knowing gapminder and his founder ethics, the objective here is to convey that the situation can be improved and the problem is not so large that it has no solution. Additionally to article it is possible to get a more detailed view on the topic by obtaining data and analyze it directly. The required numbers can be taken from the UN in the links below:

UN refugee population data

UN world population data

In the next paragraphs, the analysis is done with the programming language R, which has very effective tools to handle this type of datasets.

```
library(tidyr)
library(readr)
library(dplyr)
library(janitor)
```

```
##
## Attaching package: 'janitor'
## The following objects are masked from 'package:stats':
##
##   chisq.test, fisher.test
```

```
library(ggplot2)
library(scales)
```

```
##
## Attaching package: 'scales'
## The following object is masked from 'package:readr':
##
##   col_factor
```

```
world_population <- read_csv("data/worldpop2019.csv")
```

```
## Rows: 280932 Columns: 10
## -- Column specification -----
## Delimiter: ","
## chr (2): Location, Variant
## dbl (8): LocID, VarID, Time, MidPeriod, PopMale, PopFemale, PopTotal, PopDensity
##
## i Use `spec()` to retrieve the full column specification for this data.
```

```
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
world_totals <- world_population %>%  
  filter(Location == "World") %>%  
  select(Time, PopTotal) %>%  
  mutate(PopTotal = PopTotal * 1000)
```

```
refugee_population <- read_csv("data/refugeepop.csv") %>%  
  clean_names(. , "upper_camel") %>%  
  select(  
    -CountryOfOrigin,  
    -CountryOfOriginIso,  
    -CountryOfAsylum,  
    -CountryOfAsylumIso) %>%  
  replace_na(list(VenezuelansDisplacedAbroad = 0))
```

```
## Rows: 71 Columns: 11
```

```
## -- Column specification -----
```

```
## Delimiter: ","
```

```
## chr (4): Country of origin, Country of origin (ISO), Country of asylum, Country of asylum (ISO)
```

```
## dbl (7): Year, Refugees under UNHCR's mandate, Asylum-seekers, IDPs of concern to UNHCR, Venezuelans
```

```
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
```

```
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

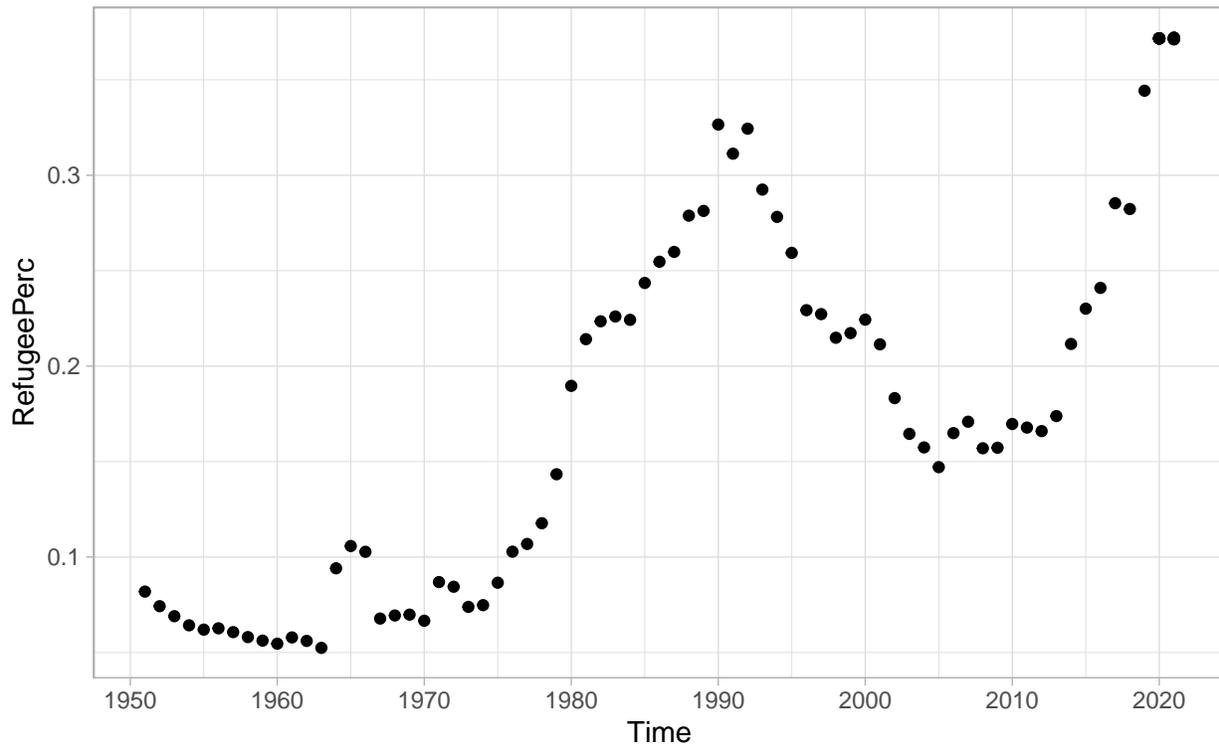
```
refugee_totals <- refugee_population %>%  
  rowwise() %>%  
  mutate(  
    RefugeeTotal =  
    RefugeesUnderUnhcrsMandate +  
    #AsylumSeekers +  
    #IdPsOfConcernToUnhcr +  
    #VenezuelansDisplacedAbroad +  
    #StatelessPersons +  
    OthersOfConcern) %>%  
  select(Year, RefugeeTotal)
```

```
refugee_prop <- world_totals %>%  
  left_join(refugee_totals, by = c("Time" = "Year")) %>%  
  mutate(RefugeePerc = 100 * RefugeeTotal/PopTotal) %>%  
  drop_na(RefugeePerc)
```

```
refugee_prop %>%  
  ggplot(aes(x = Time)) +  
  geom_point(aes(y = RefugeePerc)) +  
  #geom_point(aes(y = PopTotal)) +  
  #scale_y_continuous(labels = label_number(scale = 1)) +  
  scale_x_continuous(n.breaks = 10) +  
  theme_light() +  
  labs(  
    title = "Refugees",  
    subtitle = "Share of world population [%]"  
  )  
)
```

Refugees

Share of world population [%]



The analysis confirms gapminder numbers for 2019 when taking only the Refugee status. On the other hand the UN data provides much higher numbers by including internal displacements and other, which is not plotted here. The final plot shows that in the last decade the proportion is rising fast again.

As often high level statistics like this are useful but incomplete. For the specific countries where people have to flee their homes, the share can be very high. A focus on such situations is to be done on an upcoming article.