Plastic Pollution Analysis

Vaibhav

5/21/2019

```
library(janitor)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(tidyverse)
library(countrycode)
```

```
## Warning: package 'countrycode' was built under R version 3.5.3
```

```
library(ggplot2)
library(ggalt)
library(ggthemes)
library(viridis)
```

```
## Loading required package: viridisLite
```

```
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
##
## Attaching package: 'magrittr'
```

```
## The following object is masked from 'package:purrr':
##
##     set_names
```

```
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
##
## Attaching package: 'ggpubr'
```

```
## The following objects are masked from 'package:tidylog':
##
##     group_by, mutate
```

```
library(CoordinateCleaner)
```

```
## Warning: package 'CoordinateCleaner' was built under R version 3.5.3
```

```
coast_vs_waste <- readr::read_csv("https://raw.githubusercontent.com/rfordatascience/tidytuesda
y/master/data/2019/2019-05-21/coastal-population-vs-mismanaged-plastic.csv") %>% clean_names()
```

```
## `curl` package not installed, falling back to using `url()`
```

```
## Parsed with column specification:
## cols(
##   Entity = col_character(),
##   Code = col_character(),
##   Year = col_integer(),
##   `Mismanaged plastic waste (tonnes)` = col_integer(),
##   `Coastal population` = col_integer(),
##   `Total population (Gapminder)` = col_integer()
## )
```

```
mismanaged_vs_gdp <- readr::read_csv("https://raw.githubusercontent.com/rfordatascience/tidytues
day/master/data/2019/2019-05-21/per-capita-mismanaged-plastic-waste-vs-gdp-per-capita.csv") %>%
 clean_names()
```

```
## `curl` package not installed, falling back to using `url()`
```

```
## Parsed with column specification:
## cols(
##   Entity = col_character(),
##   Code = col_character(),
##   Year = col_integer(),
##   `Per capita mismanaged plastic waste (kilograms per person per day)` = col_double(),
##   `GDP per capita, PPP (constant 2011 international $) (Rate)` = col_double(),
##   `Total population (Gapminder)` = col_integer()
## )
```

```
waste_vs_gdp <- readr::read_csv("https://raw.githubusercontent.com/rfordatascience/tidytuesday/m
aster/data/2019/2019-05-21/per-capita-plastic-waste-vs-gdp-per-capita.csv") %>% clean_names()
```

```
## `curl` package not installed, falling back to using `url()`
```

```
## Parsed with column specification:
## cols(
##   Entity = col_character(),
##   Code = col_character(),
##   Year = col_integer(),
##   `Per capita plastic waste (kilograms per person per day)` = col_double(),
##   `GDP per capita, PPP (constant 2011 international $) (constant 2011 international $)` = col
_double(),
##   `Total population (Gapminder)` = col_integer()
## )
```

Exploring the dataset

```
coast_vs_waste_final <- coast_vs_waste %>%
  filter(!is.na(entity)) %>%
  mutate(iso3 = countrycode(entity, "country.name", "iso3c"))
```

```
## Warning in countrycode(entity, "country.name", "iso3c"): Some values were not matched unambig
uously: Channel Islands, Micronesia (country), Netherlands Antilles, World
```

```
codes <- codelist %>%
  select(iso3c, country.name.en, region, continent) %>%
  janitor::clean_names() %>%
  filter(!is.na(continent)) %>%
  filter(!is.na(region)) %>%
  rename(iso3 = iso3c) %>%
  left_join(CoordinateCleaner::countryref %>% select(iso3, capital.lon, capital.lat)) %>%
  distinct() %>%
  filter(!is.na(capital.lon)) %>%
  filter(!is.na(capital.lat))
```

```
## Joining, by = "iso3"
```

```
## Warning: Column `iso3` joining character vector and factor, coercing into
## character vector
```

```
coast_vs_waste_final <- coast_vs_waste_final %>%
  left_join(codes)
```

```
## Joining, by = "iso3"
```

Plotting mismanaged plastic pollution on world map

```
world <- map_data("world")
```

```
##
## Attaching package: 'maps'
```

```
## The following object is masked from 'package:purrr':
##
##     map
```

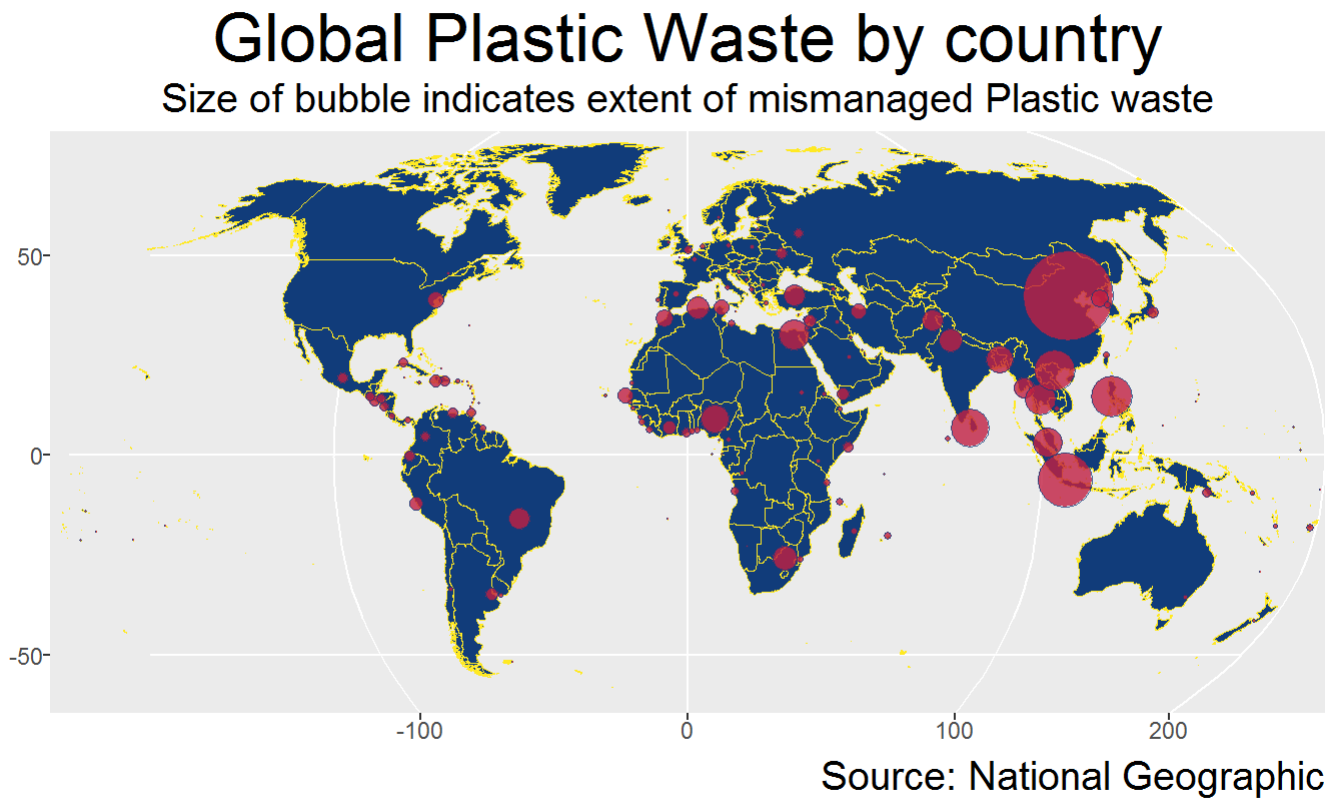
```
world <- world[world$region != "Antarctica", ]
names_dif <- anti_join(coast_vs_waste_final, world, by = c("country_name_en" = "region"))

ggplot() +
  geom_cartogram(
    data = world, map = world,
    aes(x = long, y = lat, map_id = region),
    color = "#ffe923", fill = "#113c7a", size = 0.125
  ) +
  geom_point(
    data = coast_vs_waste_final, aes(capital.lon, capital.lat, size = mismanaged_plastic_waste_t
onnes), fill = "#c11f42",
    shape = 21, alpha = 0.8, stroke = 0.25, color = "#113c7a"
  ) +
  coord_proj("+proj=robin") +
  scale_size_area(name = "Global Plastic Waste", breaks = c(10, 50, 100, 200), max_size = 30, la
bels = scales::comma) +
  labs(
    x = NULL, y = NULL,
    title = "Global Plastic Waste by country",
    subtitle = "Size of bubble indicates extent of mismanaged Plastic waste",
    caption = "Source: National Geographic"
  ) +
  theme(plot.title = element_text(hjust = 0.5, size = 25)) +
  theme(plot.subtitle = element_text(hjust = 0.5, size = 15)) +
  theme(plot.caption = element_text(size = 15)) +
  theme(legend.position = "none") +
  theme(legend.title = element_text(size = 18)) +
  theme(legend.text = element_text(size = 18))
```

```
## Warning: range backtransformation not implemented in this coord; results
## may be wrong.
```

```
## Warning: range calculation not implemented in this coord; results may be
## wrong.
```

```
## Warning: Removed 19923 rows containing missing values (geom_point).
```



Plotting Coastal population Vs Total population on world map

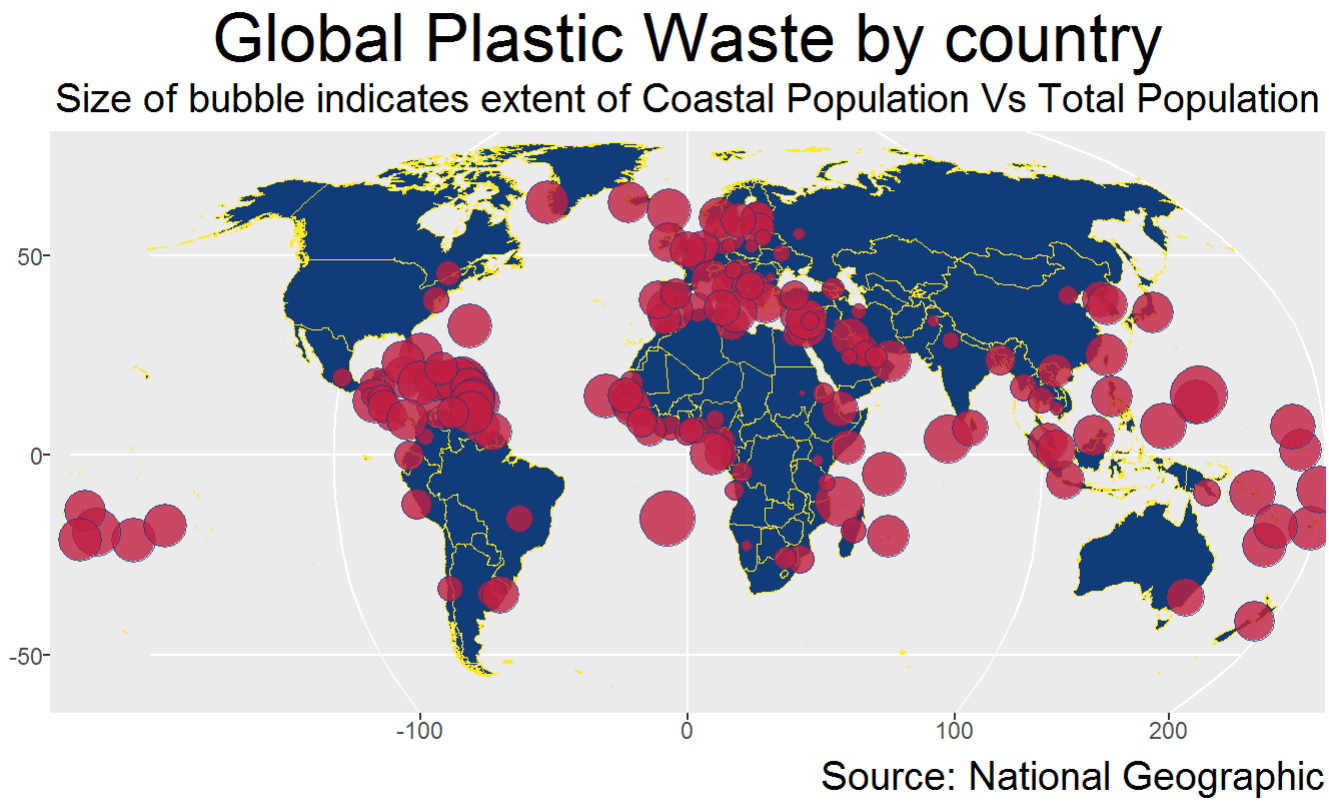
```
coast_vs_waste_final_plot <- coast_vs_waste_final %>% mutate(coast_by_total = coastal_population
/ total_population_gapminder * 100)

ggplot() +
  geom_cartogram(
    data = world, map = world,
    aes(x = long, y = lat, map_id = region),
    color = "#ffe923", fill = "#113c7a", size = 0.125
  ) +
  geom_point(
    data = coast_vs_waste_final_plot, aes(capital.lon, capital.lat, size = coast_by_total), fill
= "#c11f42",
    shape = 21, alpha = 0.8, stroke = 0.25, color = "#113c7a"
  ) +
  coord_proj("+proj=robin") +
  scale_size_area(name = "Global Plastic Waste", breaks = c(10, 20, 50, 70, 100), max_size = 10,
labels = scales::comma) +
  labs(
    x = NULL, y = NULL,
    title = "Global Plastic Waste by country",
    subtitle = "Size of bubble indicates extent of Coastal Population Vs Total Population",
    caption = "Source: National Geographic"
  ) +
  theme(plot.title = element_text(hjust = 0.5, size = 25)) +
  theme(plot.subtitle = element_text(hjust = 0.5, size = 15)) +
  theme(plot.caption = element_text(size = 15)) +
  theme(legend.position = "none") +
  theme(legend.title = element_text(size = 18)) +
  theme(legend.text = element_text(size = 18))
```

```
## Warning: range backtransformation not implemented in this coord; results
## may be wrong.
```

```
## Warning: range calculation not implemented in this coord; results may be
## wrong.
```

```
## Warning: Removed 19928 rows containing missing values (geom_point).
```



Mismanaged Vs GDP

GDP Vs Mismanaged waste

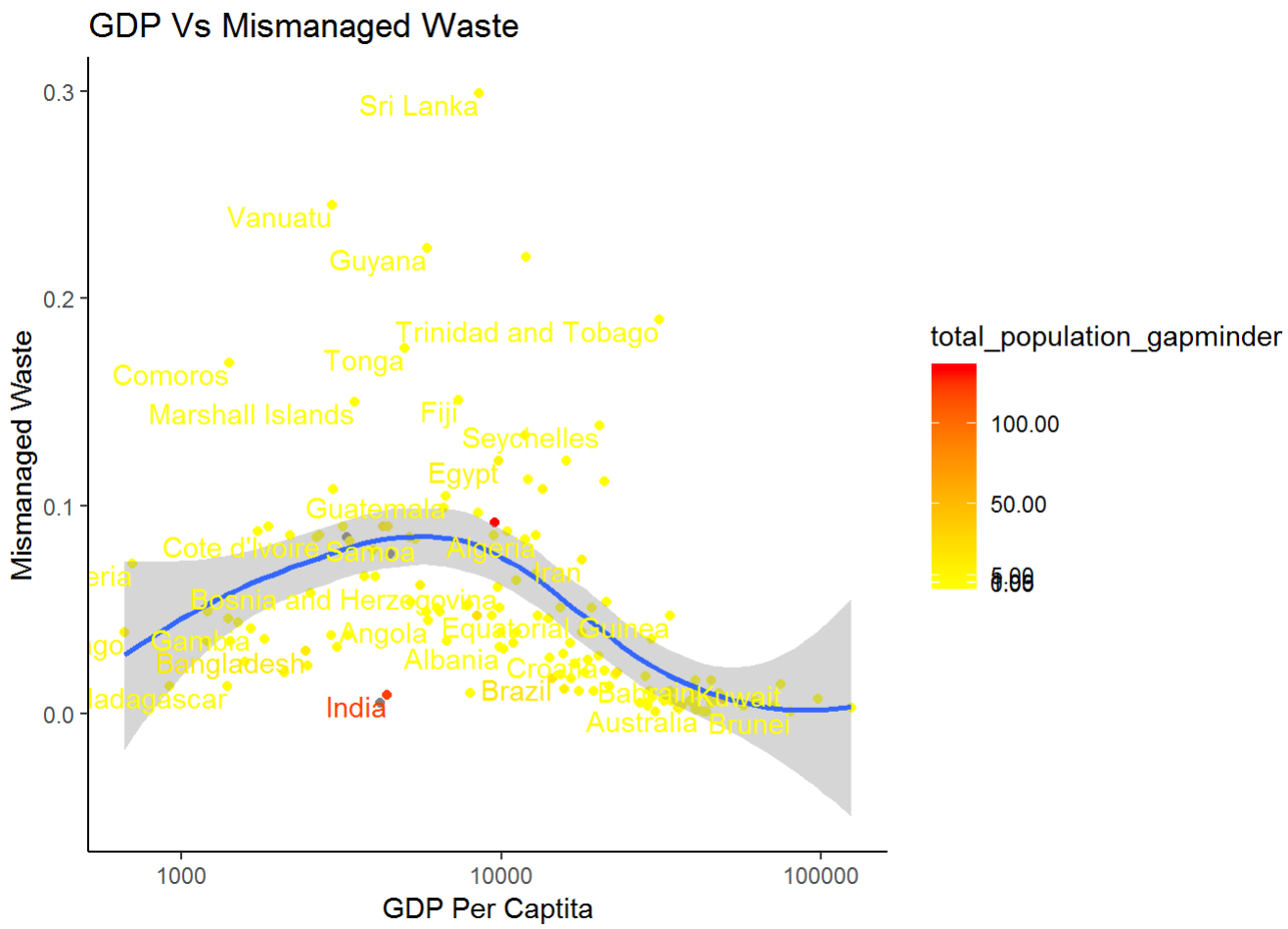
```
breaks <- c(0,0.05,1,5,50,100)
mismanaged_vs_gdp %>%mutate(total_population_gapminder=total_population_gapminder/10^7) %>%
filter(year == 2010) %>%
ggplot(aes(y=per_capita_mismanaged_plastic_waste_kilograms_per_person_per_day,
           x=gdp_per_capita_ppp_constant_2011_international_rate, color = total_population_gapmi
nder)) +  geom_point() +  scale_x_log10() + scale_colour_gradient2(low = "black", mid = "yellow"
, high = "red", midpoint = 0.003, breaks = breaks) + geom_smooth()+
geom_text(aes(label = entity), vjust = 1, hjust = 1, check_overlap = TRUE)+
  labs(title="GDP Vs Mismanaged Waste",x="GDP Per Captita", y="Mismanaged Waste")+
  theme_classic()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 134 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 134 rows containing missing values (geom_point).
```

```
## Warning: Removed 134 rows containing missing values (geom_text).
```



Richer Nations vs Plastic they produce

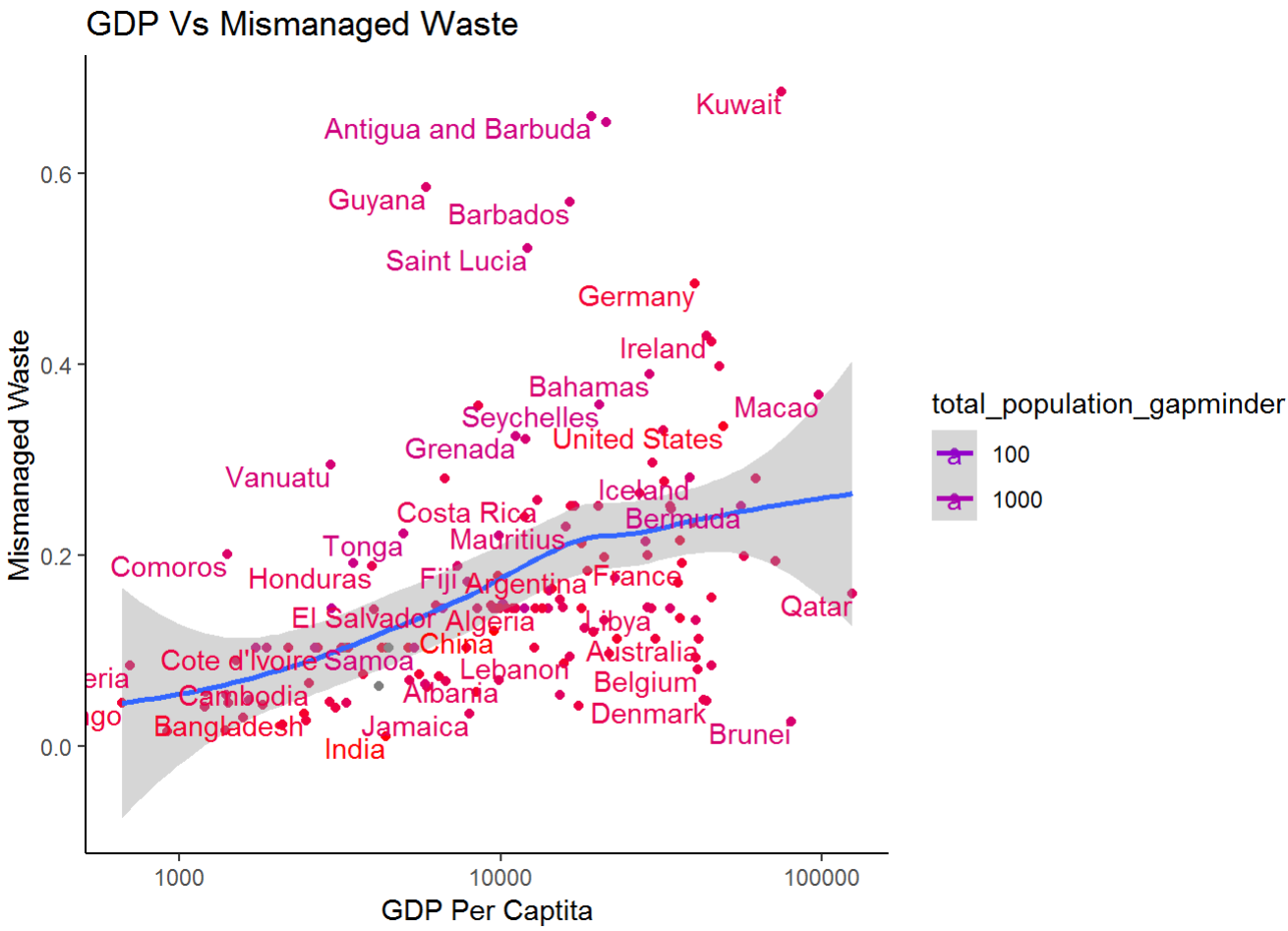
```
my_breaks <- c(10,1.5,15,100,1000)
waste_vs_gdp %>%
filter(year == 2010 & entity != "Trinidad and Tobago") %>%
ggplot(aes(y=per_capita_plastic_waste_kilograms_per_person_per_day,
           x=gdp_per_capita_ppp_constant_2011_international_constant_2011_international, color =
total_population_gapminder)) +  geom_point() +  scale_x_log10() + scale_colour_gradient2(low =
"green", mid = "blue", high = "red", trans = "log", guide="legend",
 breaks = my_breaks, labels = my_breaks, ) + geom_smooth()+
geom_text(aes(label = entity), vjust = 1, hjust = 1, check_overlap = TRUE)+
  labs(title="GDP Vs Mismanaged Waste",x="GDP Per Captita", y="Mismanaged Waste")+
  theme_classic()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
## Warning: Removed 134 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 134 rows containing missing values (geom_point).
```

```
## Warning: Removed 134 rows containing missing values (geom_text).
```


GDP Vs Mismanaged Waste
Mismanaged Waste
0.0
0.2
0.4
0.6
1000
10000
100000
GDP Per Captita
Kuwait
Antigua and Barbuda
Guyana
Barbados
Saint Lucia
Germany
Ireland
Bahamas
Seychelles
Macao
United States
Grenada
Vanuatu
Iceland
Costa Rica
Bermuda
Tonga
Mauritius
Comoros
Honduras
Fiji
Argentina
France
Qatar
El Salvador
Algeria
Libya
China
Australia
Cote d'Ivoire
Samoa
Lebanon
Belgium
Cambodia
Albania
Denmark
Bangladesh
Jamaica
Brunei
India
total_population_gapminder
100
1000