T412006-17320268-assignment2

17320268

2022-11-02

Set-up

```
rm(list = ls())
library(tidyverse)
library(knitr)
library(ggmosaic)
library(knitr)
setwd("D:/MAPEC/Applied methods/assignment 2")
data_qog <- read_csv("2022-10-31-T412006-assignment2-data.csv")
```

Q1

```
dim(data_qog)
```

```
## [1] 12156    20
```

```
ncol(data_qog)
```

```
## [1] 20
```

```
nrow(data_qog)
```

```
## [1] 12156
```

Q2

```
names(data_qog)
```

```
##  [1] "ccode"             "cname"             "year"
##  [4] "ht_region"         "wdi_area"          "wdi_pop"
##  [7] "wdi_popden"        "br_dem"            "br_elect"
## [10] "chga_hinst"        "ht_regtype"        "p_polity2"
## [13] "wdi_gnicon2010"    "wdi_gnicapcon2010" "wdi_gdpcapcon2010"
## [16] "wdi_lifexp"        "wdi_litrad"        "undp_hdi"
## [19] "wdi_expmil"        "wdi_internet"
```

Q3

Variable name	Short description	Information
ccode	Country code	Country code with the ISO-3166-1 standard
cname	Country name	Country name
year	Year	
ht_region	Region of the Country	Tenfold politico-geographic classification of world region from 1 to 10
wdi_area	Land area (sq.km)	Country's total area excluding area under inland water bodies, exclusive economic zones and claims to continental shelf
wdi_pop	Total population	All residents regardless of legal status or citizenship (midyear estimates)
wdi_popden	Population density	People per sq. km of land area. Midyear population divided by land area in square kilometers
br_dem	Is the country a democracy	Dummy variable indicator of democracy based on minimalist definition (if there is free and fair election, peaceful turnover of officers)
br_elect	Typology of political institutions	Alternative democracy indicator capturing degree of multi-party competition.
chga_hinst	Regime Institutions	Six-fold classification of political regimes
ht_regtype	Regime Type	Qualitative variable representing the political regime of each country based on 26 levels
p_polity2	Revised Combined Polity Score	Ordinal variable: range from -10 (strongly autocratic) to +10 (strongly democratic)
wdi_gnicon2010	GNI (constant 2010 US dollar)	Gross national income: sum of value added by all resident producers plus any product taxes less subsidies
wdi_gnicapcon2010	GNI per capital (constant 2010 US dollar)	GNI divided by midyear population
wdi_gdpcapcon2010	GDP per capita (constant 2010 US dollar)	Gross domestic product divided by midyear population. GDP is the sum of gross value added by residents producers plus product taxes minus subsidies
wdi_lifexp	Life expectancy at birth, total (years)	Number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life

Variable name	Short description	Information
wdi_litrad	Literacy rate, adult total (% of people ages 15 and above)	Percentage of the population above 15 who can understand, read and write a short simple statement
undp_hdi	Human Development Index	Summary measure of average achievement in health, knowledge and decent standards of living
wdi_expmil	Military expenditure (% of GDP)	All current and capital expenditures on the armed forces
wdi_internet	Individuals using the Internet (% of population)	Internet users who have used Internet in the last 3 months

Q4

```
data_qog %>%
  select(year) %>%
  summarise(
    minium_year = min(year),
    max_year = max(year)
  ) %>%
  kable(caption = "Minimum and maximum year")
```

Table 2: Minimum and maximum year

minium_year	max_year
1960	2020

Q5

```
sort(unique(data_qog$cname)) %>%
  kable(caption = "Countries of the dataset sorted alphabetically")
```

Table 3: Countries of the dataset sorted alphabetically

x
Afghanistan
Albania
Algeria
Andorra
Angola
Antigua and Barbuda
Argentina
Armenia
Australia
Austria

x
Azerbaijan
Bahamas (the)
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belize
Benin
Bhutan
Bolivia (Plurinational State of)
Bosnia and Herzegovina
Botswana
Brazil
Brunei Darussalam
Bulgaria
Burkina Faso
Burundi
Cabo Verde
Cambodia
Cameroon
Canada
Central African Republic (the)
Chad
Chile
China
Colombia
Comoros (the)
Congo (the Democratic Republic of the)
Congo (the)
Costa Rica
Côte d'Ivoire
Croatia
Cuba
Cyprus
Czechia
Czechoslovakia
Denmark
Djibouti
Dominica
Dominican Republic (the)
Ecuador
Egypt
El Salvador
Equatorial Guinea
Eritrea
Estonia
Eswatini
Ethiopia
Fiji
Finland
France

x
Gabon
Gambia (the)
Georgia
German Democratic Republic
Germany
Ghana
Greece
Grenada
Guatemala
Guinea
Guinea-Bissau
Guyana
Haiti
Honduras
Hungary
Iceland
India
Indonesia
Iran (Islamic Republic of)
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Kiribati
Korea (the Democratic People's Republic of)
Korea (the Republic of)
Kuwait
Kyrgyzstan
Lao People's Democratic Republic (the)
Latvia
Lebanon
Lesotho
Liberia
Libya
Liechtenstein
Lithuania
Luxembourg
Madagascar
Malawi
Malaysia
Maldives
Mali
Malta
Marshall Islands
Mauritania
Mauritius
Mexico

x
Micronesia (Federated States of)
Moldova (the Republic of)
Monaco
Mongolia
Montenegro
Morocco
Mozambique
Myanmar
Namibia
Nauru
Nepal
Netherlands (the)
New Zealand
Nicaragua
Niger (the)
Nigeria
North Macedonia
Norway
Oman
Pakistan
Palau
Panama
Papua New Guinea
Paraguay
Peru
Philippines (the)
Poland
Portugal
Qatar
Romania
Russian Federation (the)
Rwanda
Saint Kitts and Nevis
Saint Lucia
Saint Vincent and the Grenadines
Samoa
San Marino
Sao Tome and Principe
Saudi Arabia
Senegal
Serbia
Serbia and Montenegro
Seychelles
Sierra Leone
Singapore
Slovakia
Slovenia
Solomon Islands
Somalia
South Africa
South Sudan
Spain

x
Sri Lanka
Sudan
Sudan (the)
Suriname
Sweden
Switzerland
Syrian Arab Republic (the)
Taiwan (Province of China)
Tajikistan
Tanzania, the United Republic of
Thailand
Tibet
Timor-Leste
Togo
Tonga
Trinidad and Tobago
Tunisia
Turkey
Turkmenistan
Tuvalu
Uganda
Ukraine
United Arab Emirates (the)
United Kingdom of Great Britain and Northern Ireland (the)
United States of America (the)
Uruguay
USSR
Uzbekistan
Vanuatu
Venezuela (Bolivarian Republic of)
Viet Nam
Vietnam, North
Vietnam, South
Yemen
Yemen Democratic
Yugoslavia
Zambia
Zimbabwe

Data transformations

Q6

```
typeof(data_qog$ht_region)
```

```
## [1] "double"
```

```
is.numeric(data_qog$ht_region)
```

```
## [1] TRUE
```

Variable ht_region is stored as a numerical variable. This does not make sense, because the variable is a geographical classification of world region from 1 to 10 each number representing a region, it should hence be stored as a nominal variable.

```
data_qog$ht_region <-  factor(data_qog$ht_region, levels = c(1:10))
unique(data_qog$ht_region)
```

```
##  [1] 8    1    <NA> 3    5    4    10   2    9    7    6
## Levels: 1 2 3 4 5 6 7 8 9 10
```

```
head(data_qog)
```

```
## # A tibble: 6 x 20
##   ccode cname        year ht_region wdi_area wdi_pop wdi_popden br_dem br_elect
##   <dbl> <chr>       <dbl> <fct>        <dbl>   <dbl>      <dbl>  <dbl>    <dbl>
## 1     4 Afghanistan  1960 8               NA 8996967       NA        0        2
## 2     4 Afghanistan  1961 8           652860 9169406       14.0      0        2
## 3     4 Afghanistan  1962 8           652860 9351442       14.3      0        2
## 4     4 Afghanistan  1963 8           652860 9543200       14.6      0        2
## 5     4 Afghanistan  1964 8           652860 9744772       14.9      0        2
## 6     4 Afghanistan  1965 8           652860 9956318       15.3      0        2
## # ... with 11 more variables: chga_hinst <dbl>, ht_regtype <dbl>,
## #   p_polity2 <dbl>, wdi_gnicon2010 <dbl>, wdi_gnicapcon2010 <dbl>,
## #   wdi_gdpcapcon2010 <dbl>, wdi_lifexp <dbl>, wdi_litrad <dbl>,
## #   undp_hdi <dbl>, wdi_expmil <dbl>, wdi_internet <dbl>
```

Q7

```
data_cntr_reg <- data_qog %>%
  filter(year == 2015) %>%
  select(cname, ht_region)
head(data_cntr_reg)
```

```
## # A tibble: 6 x 2
##   cname              ht_region
##   <chr>              <fct>
## 1 Afghanistan        8
## 2 Albania            1
## 3 Algeria            3
## 4 Andorra            5
## 5 Angola             4
## 6 Antigua and Barbuda 10
```

Q8

```
data_cntr_reg %>%
  arrange(ht_region, cname) %>%
  kable(caption = "Countries and their region", align = "c")
```

Table 4: Countries and their region

cname	ht_region
Albania	1
Armenia	1
Azerbaijan	1
Belarus	1
Bosnia and Herzegovina	1
Bulgaria	1
Croatia	1
Czechia	1
Czechoslovakia	1
Estonia	1
Georgia	1
Hungary	1
Kazakhstan	1
Kyrgyzstan	1
Latvia	1
Lithuania	1
Moldova (the Republic of)	1
Montenegro	1
North Macedonia	1
Poland	1
Romania	1
Russian Federation (the)	1
Serbia	1
Serbia and Montenegro	1
Slovakia	1
Slovenia	1
Tajikistan	1
Turkmenistan	1
Ukraine	1
USSR	1
Uzbekistan	1
Argentina	2
Bolivia (Plurinational State of)	2
Brazil	2
Chile	2
Colombia	2
Costa Rica	2
Cuba	2
Dominican Republic (the)	2
Ecuador	2
El Salvador	2
Guatemala	2
Haiti	2
Honduras	2
Mexico	2

cname	ht_region
Nicaragua	2
Panama	2
Paraguay	2
Peru	2
Uruguay	2
Venezuela (Bolivarian Republic of)	2
Algeria	3
Bahrain	3
Cyprus	3
Egypt	3
Iran (Islamic Republic of)	3
Iraq	3
Israel	3
Jordan	3
Kuwait	3
Lebanon	3
Libya	3
Morocco	3
Oman	3
Qatar	3
Saudi Arabia	3
Syrian Arab Republic (the)	3
Tunisia	3
Turkey	3
United Arab Emirates (the)	3
Yemen	3
Angola	4
Benin	4
Botswana	4
Burkina Faso	4
Burundi	4
Cabo Verde	4
Cameroon	4
Central African Republic (the)	4
Chad	4
Comoros (the)	4
Congo (the Democratic Republic of the)	4
Congo (the)	4
Côte d'Ivoire	4
Djibouti	4
Equatorial Guinea	4
Eritrea	4
Eswatini	4
Ethiopia	4
Gabon	4
Gambia (the)	4
Ghana	4
Guinea	4
Guinea-Bissau	4
Kenya	4
Lesotho	4
Liberia	4

cname	ht_region
Madagascar	4
Malawi	4
Mali	4
Mauritania	4
Mauritius	4
Mozambique	4
Namibia	4
Niger (the)	4
Nigeria	4
Rwanda	4
Sao Tome and Principe	4
Senegal	4
Seychelles	4
Sierra Leone	4
Somalia	4
South Africa	4
South Sudan	4
Sudan (the)	4
Tanzania, the United Republic of	4
Togo	4
Uganda	4
Zambia	4
Zimbabwe	4
Andorra	5
Australia	5
Austria	5
Belgium	5
Canada	5
Denmark	5
Finland	5
France	5
Germany	5
Greece	5
Iceland	5
Ireland	5
Italy	5
Liechtenstein	5
Luxembourg	5
Malta	5
Monaco	5
Netherlands (the)	5
New Zealand	5
Norway	5
Portugal	5
San Marino	5
Spain	5
Sweden	5
Switzerland	5
United Kingdom of Great Britain and Northern Ireland (the)	5
United States of America (the)	5
China	6
Japan	6

cname	ht_region
Korea (the Democratic People's Republic of)	6
Korea (the Republic of)	6
Mongolia	6
Taiwan (Province of China)	6
Brunei Darussalam	7
Cambodia	7
Indonesia	7
Lao People's Democratic Republic (the)	7
Malaysia	7
Myanmar	7
Philippines (the)	7
Singapore	7
Thailand	7
Timor-Leste	7
Viet Nam	7
Afghanistan	8
Bangladesh	8
Bhutan	8
India	8
Maldives	8
Nepal	8
Pakistan	8
Sri Lanka	8
Tibet	8
Fiji	9
Kiribati	9
Marshall Islands	9
Micronesia (Federated States of)	9
Nauru	9
Palau	9
Papua New Guinea	9
Samoa	9
Solomon Islands	9
Tonga	9
Tuvalu	9
Vanuatu	9
Antigua and Barbuda	10
Bahamas (the)	10
Barbados	10
Belize	10
Dominica	10
Grenada	10
Guyana	10
Jamaica	10
Saint Kitts and Nevis	10
Saint Lucia	10
Saint Vincent and the Grenadines	10
Suriname	10
Trinidad and Tobago	10

Q9

```
data_qog %>%
  filter(ht_region == 5) %>%
  distinct(sort(cname))
```

```
## # A tibble: 27 x 1
##    `sort(cname)`
##    <chr>
##  1 Andorra
##  2 Australia
##  3 Austria
##  4 Belgium
##  5 Canada
##  6 Denmark
##  7 Finland
##  8 France
##  9 Germany
## 10 Greece
## # ... with 17 more rows
```

Q10

```
data_cntr_reg %>%
  mutate(
    my_region = ifelse(ht_region %in% c(2, 10), "Americas",
                       ifelse(cname %in% c("Canada", "United States of America (the)"), "Americas",
                         ifelse(ht_region %in% c(6, 7, 8) | cname %in% c("Tajikistan", "Turkmenistan",
                             ifelse(ht_region %in% c(1, 5), "Europe",
                                    ifelse(ht_region == 3 | cname %in% c("Turkey", "Israel", "Cyprus")
                                           ifelse(ht_region == 4, "Sub-Saharan Africa",
                                                  ifelse(ht_region == 9 | cname %in% c("Australia", "N
  ) -> data_cntr_reg

data_cntr_reg %>%
  arrange(ht_region, cname) %>%
  kable(caption = "Countries and their region", align = "c")
```

Table 5: Countries and their region

cname	ht_region	my_region
Albania	1	Europe
Armenia	1	Europe
Azerbaijan	1	Europe
Belarus	1	Europe
Bosnia and Herzegovina	1	Europe
Bulgaria	1	Europe
Croatia	1	Europe
Czechia	1	Europe

cname	ht_region	my_region
Czechoslovakia	1	Europe
Estonia	1	Europe
Georgia	1	Europe
Hungary	1	Europe
Kazakhstan	1	Asia
Kyrgyzstan	1	Asia
Latvia	1	Europe
Lithuania	1	Europe
Moldova (the Republic of)	1	Europe
Montenegro	1	Europe
North Macedonia	1	Europe
Poland	1	Europe
Romania	1	Europe
Russian Federation (the)	1	Europe
Serbia	1	Europe
Serbia and Montenegro	1	Europe
Slovakia	1	Europe
Slovenia	1	Europe
Tajikistan	1	Asia
Turkmenistan	1	Asia
Ukraine	1	Europe
USSR	1	Europe
Uzbekistan	1	Asia
Argentina	2	Americas
Bolivia (Plurinational State of)	2	Americas
Brazil	2	Americas
Chile	2	Americas
Colombia	2	Americas
Costa Rica	2	Americas
Cuba	2	Americas
Dominican Republic (the)	2	Americas
Ecuador	2	Americas
El Salvador	2	Americas
Guatemala	2	Americas
Haiti	2	Americas
Honduras	2	Americas
Mexico	2	Americas
Nicaragua	2	Americas
Panama	2	Americas
Paraguay	2	Americas
Peru	2	Americas
Uruguay	2	Americas
Venezuela (Bolivarian Republic of)	2	Americas
Algeria	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Bahrain	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Cyprus	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Egypt	3	North Africa & Middle East (including Israel, Turkey & Cyprus)

cname	ht_region	my_region
Iran (Islamic Republic of)	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Iraq	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Israel	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Jordan	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Kuwait	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Lebanon	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Libya	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Morocco	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Oman	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Qatar	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Saudi Arabia	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Syrian Arab Republic (the)	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Tunisia	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Turkey	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
United Arab Emirates (the)	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Yemen	3	North Africa & Middle East (including Israel, Turkey & Cyprus)
Angola	4	Sub-Saharan Africa
Benin	4	Sub-Saharan Africa
Botswana	4	Sub-Saharan Africa
Burkina Faso	4	Sub-Saharan Africa
Burundi	4	Sub-Saharan Africa
Cabo Verde	4	Sub-Saharan Africa
Cameroon	4	Sub-Saharan Africa
Central African Republic (the)	4	Sub-Saharan Africa
Chad	4	Sub-Saharan Africa
Comoros (the)	4	Sub-Saharan Africa
Congo (the Democratic Republic of the)	4	Sub-Saharan Africa
Congo (the)	4	Sub-Saharan Africa
Côte d'Ivoire	4	Sub-Saharan Africa
Djibouti	4	Sub-Saharan Africa
Equatorial Guinea	4	Sub-Saharan Africa
Eritrea	4	Sub-Saharan Africa
Eswatini	4	Sub-Saharan Africa
Ethiopia	4	Sub-Saharan Africa
Gabon	4	Sub-Saharan Africa
Gambia (the)	4	Sub-Saharan Africa

cname	ht_region	my_region
Ghana	4	Sub-Saharan Africa
Guinea	4	Sub-Saharan Africa
Guinea-Bissau	4	Sub-Saharan Africa
Kenya	4	Sub-Saharan Africa
Lesotho	4	Sub-Saharan Africa
Liberia	4	Sub-Saharan Africa
Madagascar	4	Sub-Saharan Africa
Malawi	4	Sub-Saharan Africa
Mali	4	Sub-Saharan Africa
Mauritania	4	Sub-Saharan Africa
Mauritius	4	Sub-Saharan Africa
Mozambique	4	Sub-Saharan Africa
Namibia	4	Sub-Saharan Africa
Niger (the)	4	Sub-Saharan Africa
Nigeria	4	Sub-Saharan Africa
Rwanda	4	Sub-Saharan Africa
Sao Tome and Principe	4	Sub-Saharan Africa
Senegal	4	Sub-Saharan Africa
Seychelles	4	Sub-Saharan Africa
Sierra Leone	4	Sub-Saharan Africa
Somalia	4	Sub-Saharan Africa
South Africa	4	Sub-Saharan Africa
South Sudan	4	Sub-Saharan Africa
Sudan (the)	4	Sub-Saharan Africa
Tanzania, the United Republic of	4	Sub-Saharan Africa
Togo	4	Sub-Saharan Africa
Uganda	4	Sub-Saharan Africa
Zambia	4	Sub-Saharan Africa
Zimbabwe	4	Sub-Saharan Africa
Andorra	5	Europe
Australia	5	Europe
Austria	5	Europe
Belgium	5	Europe
Canada	5	Americas
Denmark	5	Europe
Finland	5	Europe
France	5	Europe
Germany	5	Europe
Greece	5	Europe
Iceland	5	Europe
Ireland	5	Europe
Italy	5	Europe
Liechtenstein	5	Europe
Luxembourg	5	Europe
Malta	5	Europe
Monaco	5	Europe
Netherlands (the)	5	Europe
New Zealand	5	Europe
Norway	5	Europe
Portugal	5	Europe
San Marino	5	Europe
Spain	5	Europe

cname	ht_region	my_region
Sweden	5	Europe
Switzerland	5	Europe
United Kingdom of Great Britain and Northern Ireland (the)	5	Europe
United States of America (the)	5	Americas
China	6	Asia
Japan	6	Asia
Korea (the Democratic People's Republic of)	6	Asia
Korea (the Republic of)	6	Asia
Mongolia	6	Asia
Taiwan (Province of China)	6	Asia
Brunei Darussalam	7	Asia
Cambodia	7	Asia
Indonesia	7	Asia
Lao People's Democratic Republic (the)	7	Asia
Malaysia	7	Asia
Myanmar	7	Asia
Philippines (the)	7	Asia
Singapore	7	Asia
Thailand	7	Asia
Timor-Leste	7	Asia
Viet Nam	7	Asia
Afghanistan	8	Asia
Bangladesh	8	Asia
Bhutan	8	Asia
India	8	Asia
Maldives	8	Asia
Nepal	8	Asia
Pakistan	8	Asia
Sri Lanka	8	Asia
Tibet	8	Asia
Fiji	9	The Pacific (including Australia & New-Zealand
Kiribati	9	The Pacific (including Australia & New-Zealand
Marshall Islands	9	The Pacific (including Australia & New-Zealand
Micronesia (Federated States of)	9	The Pacific (including Australia & New-Zealand
Nauru	9	The Pacific (including Australia & New-Zealand
Palau	9	The Pacific (including Australia & New-Zealand
Papua New Guinea	9	The Pacific (including Australia & New-Zealand
Samoa	9	The Pacific (including Australia & New-Zealand
Solomon Islands	9	The Pacific (including Australia & New-Zealand
Tonga	9	The Pacific (including Australia & New-Zealand
Tuvalu	9	The Pacific (including Australia & New-Zealand
Vanuatu	9	The Pacific (including Australia & New-Zealand
Antigua and Barbuda	10	Americas
Bahamas (the)	10	Americas
Barbados	10	Americas
Belize	10	Americas
Dominica	10	Americas
Grenada	10	Americas
Guyana	10	Americas
Jamaica	10	Americas
Saint Kitts and Nevis	10	Americas

cname	ht_region	my_region
Saint Lucia	10	Americas
Saint Vincent and the Grenadines	10	Americas
Suriname	10	Americas
Trinidad and Tobago	10	Americas

Q11

```
data_qog <- data_qog %>%
  mutate(
    my_region = ifelse(ht_region %in% c(2, 10), "Americas",
                       ifelse(cname %in% c("Canada", "United States of America (the)"), "Americas",
                           ifelse(ht_region %in% c(6, 7, 8) | cname %in% c("Tajikistan", "Turkmenistan
                             ifelse(ht_region %in% c(1, 5), "Europe",
                                    ifelse(ht_region == 3 | cname %in% c("Turkey", "Israel", "Cyprus")
                                           ifelse(ht_region == 4, "Sub-Saharan Africa",
                                                  ifelse(ht_region == 9 | cname %in% c("Australia", "N
  )
```

Q12

```
my_gnicon2010 <- data_qog %>%
  select(cname, wdi_pop, wdi_gnicon2010, wdi_gnicapcon2010) %>%
  mutate(my_gnicapcon = wdi_gnicon2010/wdi_pop)
head(my_gnicon2010)
```

```
## # A tibble: 6 x 5
##   cname       wdi_pop wdi_gnicon2010 wdi_gnicapcon2010 my_gnicapcon
##   <chr>         <dbl>          <dbl>              <dbl>        <dbl>
## 1 Afghanistan 8996967             NA                 NA           NA
## 2 Afghanistan 9169406             NA                 NA           NA
## 3 Afghanistan 9351442             NA                 NA           NA
## 4 Afghanistan 9543200             NA                 NA           NA
## 5 Afghanistan 9744772             NA                 NA           NA
## 6 Afghanistan 9956318             NA                 NA           NA
```

Q13

```
data_gnipc <-
    my_gnicon2010 %>%
    mutate(check = ifelse(my_gnicapcon == wdi_gnicapcon2010, 0, 1))
head(data_gnipc)
```

```
## # A tibble: 6 x 6
##   cname       wdi_pop wdi_gnicon2010 wdi_gnicapcon2010 my_gnicapcon check
##   <chr>         <dbl>          <dbl>              <dbl>        <dbl> <dbl>
```

```
## 1 Afghanistan 8996967            NA               NA          NA    NA
## 2 Afghanistan 9169406            NA               NA          NA    NA
## 3 Afghanistan 9351442            NA               NA          NA    NA
## 4 Afghanistan 9543200            NA               NA          NA    NA
## 5 Afghanistan 9744772            NA               NA          NA    NA
## 6 Afghanistan 9956318            NA               NA          NA    NA
```

Q14

```
data_gnipc %>%
  count(check) %>%
  mutate(percentage = n/sum(n))
```

```
## # A tibble: 3 x 3
##   check     n percentage
##   <dbl> <int>      <dbl>
## 1     0   219     0.0180
## 2     1  5117     0.421
## 3    NA  6820     0.561
```

Q15

It does not look like, when check is equal to one, that the values are different. The reason is perhaps because of how the values are rounded: observations with check = 0 are rounded at the fourth decimal and observations with check = 1 are rounded at the third decimal.

```
data_gnipc <- data_gnipc %>%
  mutate(check2 = ifelse(round(my_gnicapcon, digits = 3) == round(wdi_gnicapcon2010, digits = 3), 0, 1))
```

```
data_gnipc %>%
  count(check2) %>%
  kable()
```

check2	n
0	5155
1	181
NA	6820

Describing variation

Q16

```
data_qog %>%
  summarize(mean = mean(wdi_internet, na.rm = TRUE),
            median = median(wdi_internet, na.rm = TRUE),
```

```
            max = max(wdi_internet, na.rm = TRUE),
            min = min(wdi_internet, na.rm = TRUE),
            sd = sd(wdi_internet, na.rm = TRUE),
            IQR = IQR(wdi_internet, na.rm = TRUE)) %>%
  kable(caption = "descriptive statistics for wdi_internet")
```

Table 7: descriptive statistics for wdi_internet

mean	median	max	min	sd	IQR
21.75882	5.793264	100	0	28.65702	37.86554

```
data_qog %>%
  filter(wdi_internet == 100)
```

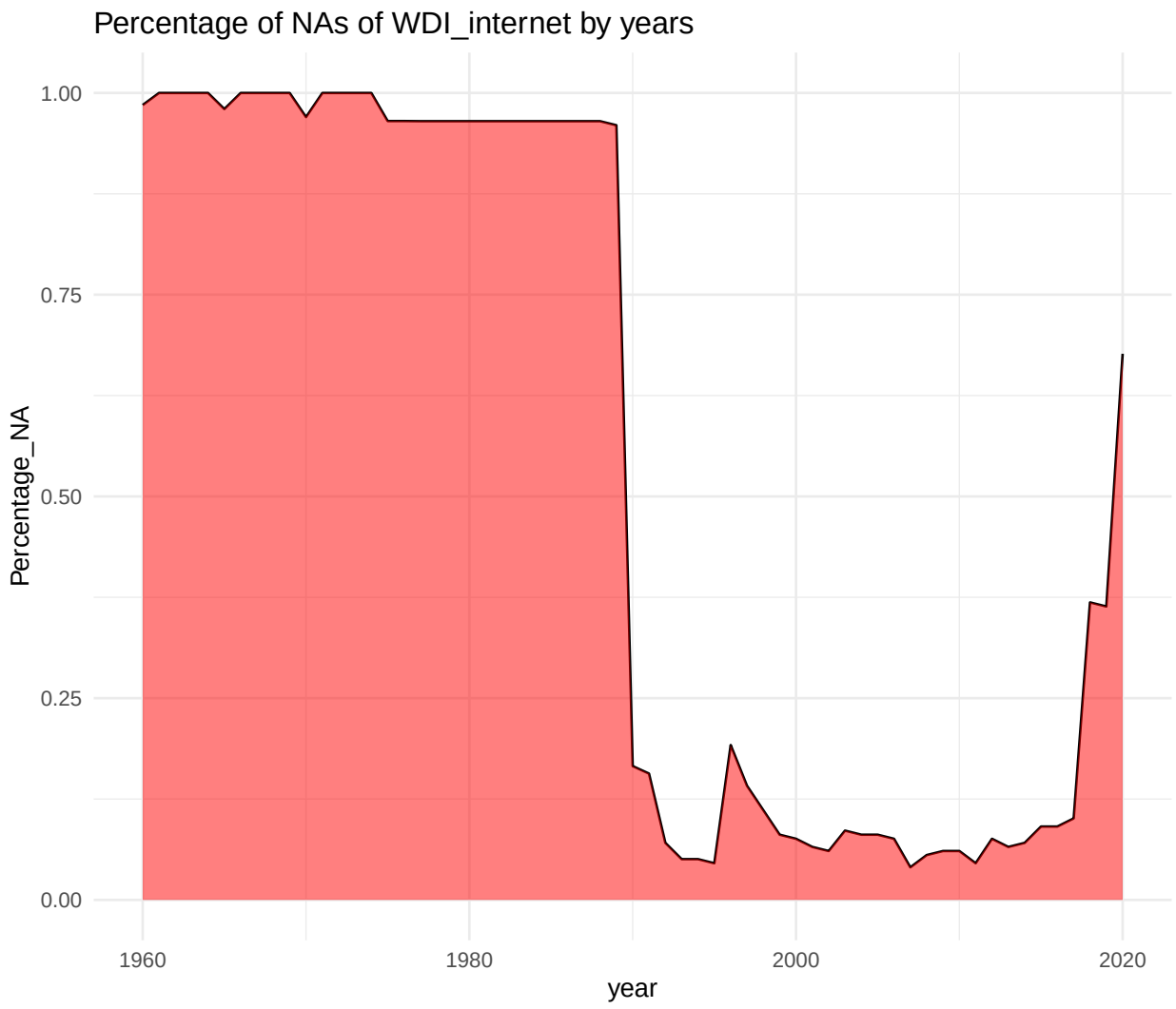
```
## # A tibble: 1 x 21
##   ccode cname         year ht_region wdi_area wdi_pop wdi_popden br_dem br_elect
##   <dbl> <chr>        <dbl> <fct>        <dbl>   <dbl>      <dbl>  <dbl>    <dbl>
## 1   784 United Arab~  2020 3            71020 9890400       139.      0        0
## # ... with 12 more variables: chga_hinst <dbl>, ht_regtype <dbl>,
## #   p_polity2 <dbl>, wdi_gnicon2010 <dbl>, wdi_gnicapcon2010 <dbl>,
## #   wdi_gdpcapcon2010 <dbl>, wdi_lifexp <dbl>, wdi_litrad <dbl>,
## #   undp_hdi <dbl>, wdi_expmil <dbl>, wdi_internet <dbl>, my_region <chr>
```

```
data_qog %>%
  filter(wdi_internet == 0 & cname == "Afghanistan" & year == 1990)
```

```
## # A tibble: 1 x 21
##   ccode cname        year ht_region wdi_area  wdi_pop wdi_popden br_dem br_elect
##   <dbl> <chr>       <dbl> <fct>        <dbl>    <dbl>      <dbl>  <dbl>    <dbl>
## 1     4 Afghanistan  1990 8           652860 12412311       19.0      0        2
## # ... with 12 more variables: chga_hinst <dbl>, ht_regtype <dbl>,
## #   p_polity2 <dbl>, wdi_gnicon2010 <dbl>, wdi_gnicapcon2010 <dbl>,
## #   wdi_gdpcapcon2010 <dbl>, wdi_lifexp <dbl>, wdi_litrad <dbl>,
## #   undp_hdi <dbl>, wdi_expmil <dbl>, wdi_internet <dbl>, my_region <chr>
```

Q17

```
data_qog %>%
  group_by(year) %>%
  count(Na = (is.na(wdi_internet))) %>%
  mutate(Percentage_NA = n/sum(n)) %>%
  ungroup() %>%
  filter(Na == TRUE) %>%
  ggplot(aes(x = year, y = Percentage_NA))+
  geom_line()+
  geom_area(fill = "red", alpha = 0.5)+
  theme_minimal()+
  labs(title = "Percentage of NAs of WDI_internet by years")
```

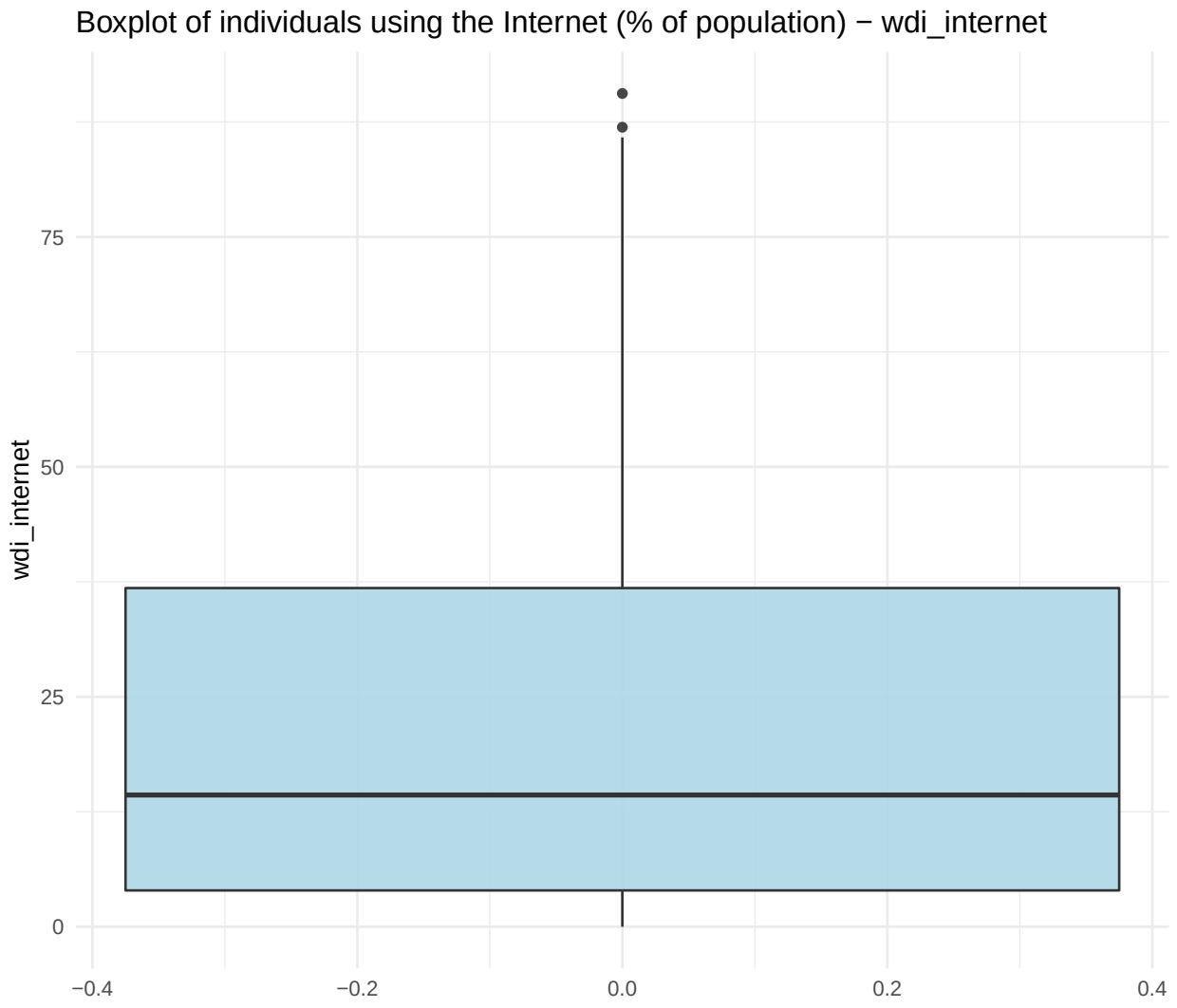



```
data_qog %>%
  group_by(year) %>%
  count(Na = (is.na(wdi_internet))) %>%
  mutate(Percentage_NA = n/sum(n)) %>%
  ungroup() -> na_analysis_internet

na_analysis_internet %>%
  filter(Na == TRUE) %>%
  summarize(best_coverage = min(Percentage_NA),
            best_coverage_year = year[which(Percentage_NA == min(Percentage_NA))]) %>%
  kable()
```

best_coverage	best_coverage_year
0.040404	2007

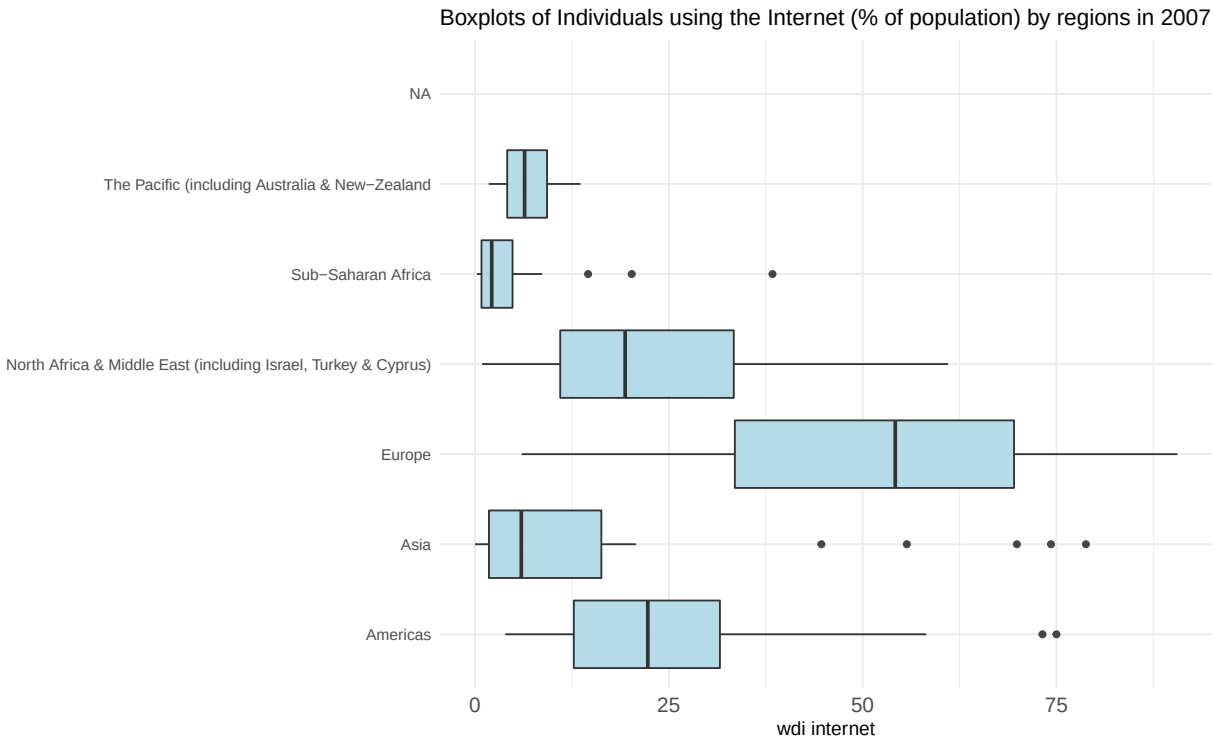
Q18

```
data_qog %>%
  filter(year == 2007) %>%
  ggplot()+
  aes(y = wdi_internet)+
  geom_boxplot(fill = "lightblue", alpha = 0.9)+
  theme_minimal()+
  labs(title = "Boxplot of individuals using the Internet (% of population) - wdi_internet")
```




```
data_qog %>%
  filter(year == 2007) %>%
  ggplot()+
  aes(y = my_region, x = wdi_internet)+
  geom_boxplot(fill = "lightblue", alpha = 0.9)+
  theme_minimal()+
  theme(axis.text.x = element_text(size = 12))+
```

```
  labs(title = "Boxplots of Individuals using the Internet (% of population) by regions in 2007")+
  xlab("wdi internet")+
  ylab("")
```



Q19

Analysis of variable life expectancy across countries and time (wdi_lifexp) We want: mean, median, range, variance, sd, IQR and percentage of NA for all dataset

```
data_qog %>%
  summarize(
    mean_life_exp = mean(wdi_lifexp, na.rm = TRUE),
    median_life_exp = median(wdi_lifexp, na.rm = TRUE),
    variance_life_exp = sd(wdi_lifexp, na.rm = TRUE)^2,
    standard_deviation_life_exp = sd(wdi_lifexp, na.rm = TRUE),
    IQR_life_exp = IQR(wdi_lifexp, na.rm = TRUE),
    min_life_exp = min(wdi_lifexp, na.rm = TRUE),
    max_life_exp = max(wdi_lifexp, na.rm = TRUE)
  ) %>%
  kable(caption = "Descriptive statistics of wdi_lifexp")
```

Table 9: Descriptive statistics of wdi_lifexp

mean_life_exp	median_life_exp	variance_life_exp	standard_deviation_life_exp	IQR_life_exp	min_life_exp	max_life_exp
64.31728	67.4765	130.9105	11.44161	17.30323	18.907	85.41707

Percentage of NAs

```
data_qog %>%
  select(wdi_lifexp) %>%
  count(is.na(wdi_lifexp)) %>%
  mutate(percentage_NAs = n/sum(n)) %>%
  kable(caption = "NAs of wdi_lifexp")
```

Table 10: NAs of wdi_lifexp

is.na(wdi_lifexp)	n	percentage_NAs
FALSE	9460	0.7782165
TRUE	2696	0.2217835

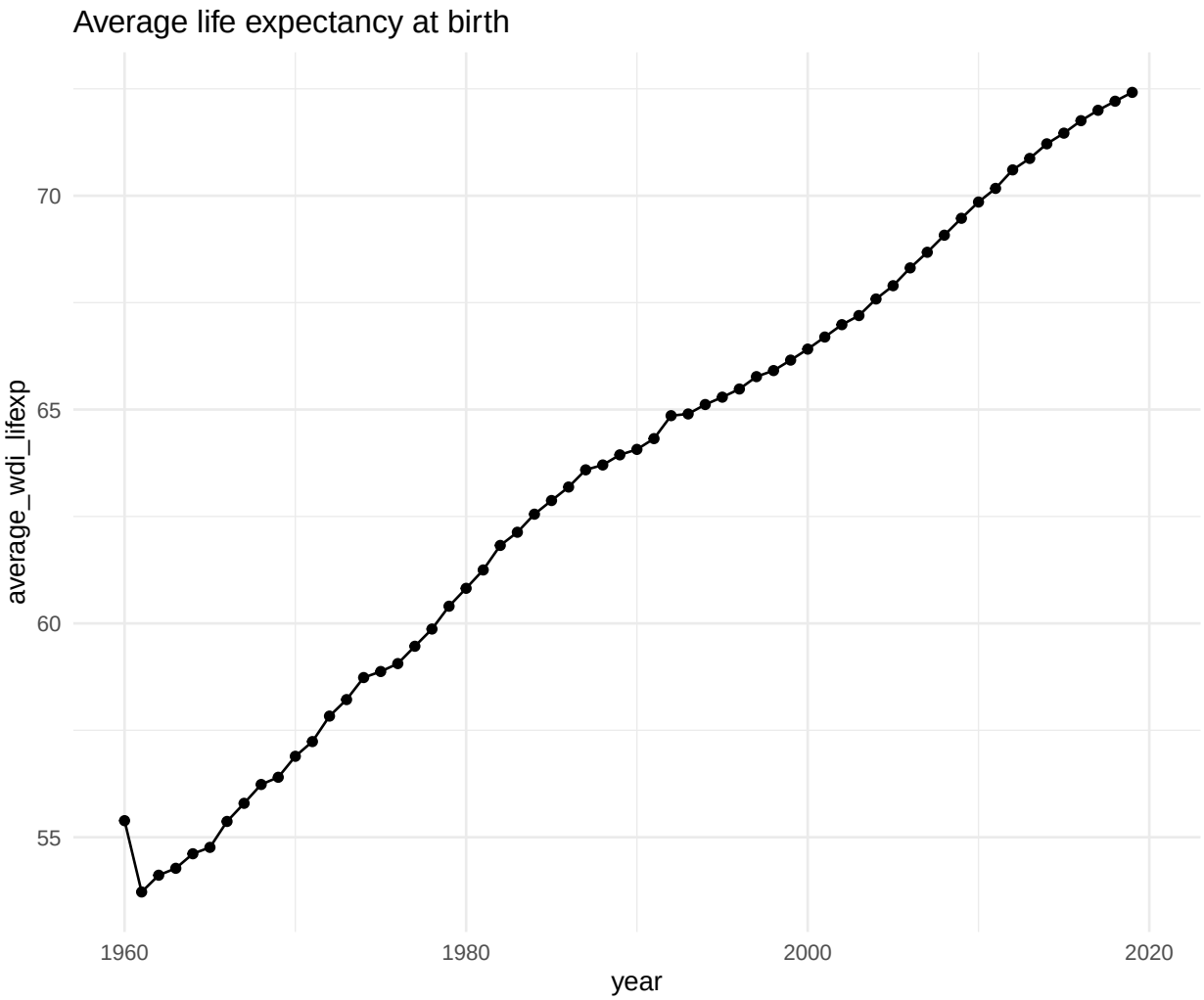
Q20

```
data_qog %>%
  group_by(year) %>%
  summarize(average_wdi_lifexp = mean(wdi_lifexp, na.rm = TRUE)) %>%
  ungroup() -> Q20_tibble
head(Q20_tibble)
```

```
## # A tibble: 6 x 2
##    year average_wdi_lifexp
##   <dbl>              <dbl>
## 1  1960               55.4
## 2  1961               53.7
## 3  1962               54.1
## 4  1963               54.3
## 5  1964               54.6
## 6  1965               54.8
```

Plotting average wdi_lifexp with year

```
data_qog %>%
  group_by(year) %>%
  summarize(average_wdi_lifexp = mean(wdi_lifexp, na.rm = TRUE)) %>%
  ungroup() %>%
  ggplot()+
  aes(x = year, y = average_wdi_lifexp)+
  geom_point()+
  geom_line()+
  theme_minimal()+
  labs(title = "Average life expectancy at birth",
       caption = "Average number of years a newborn infant would live if prevailing patterns of mortali
```

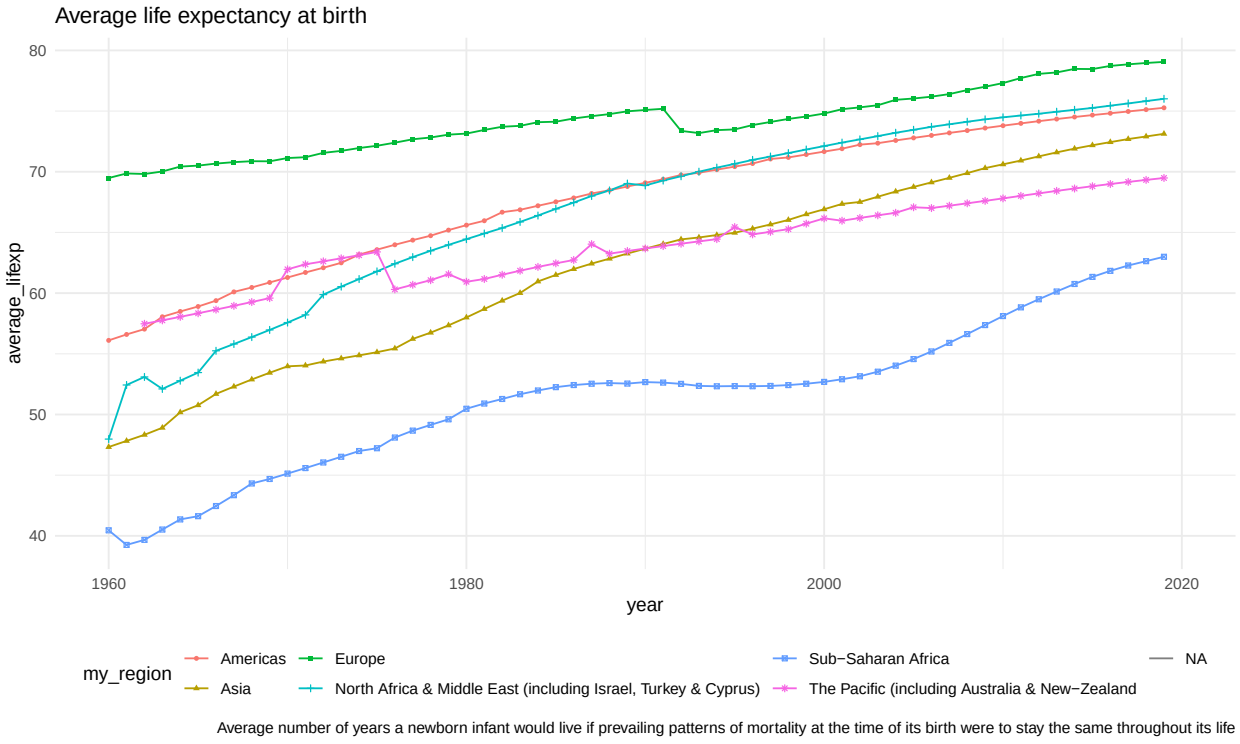


ars a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life

Q21

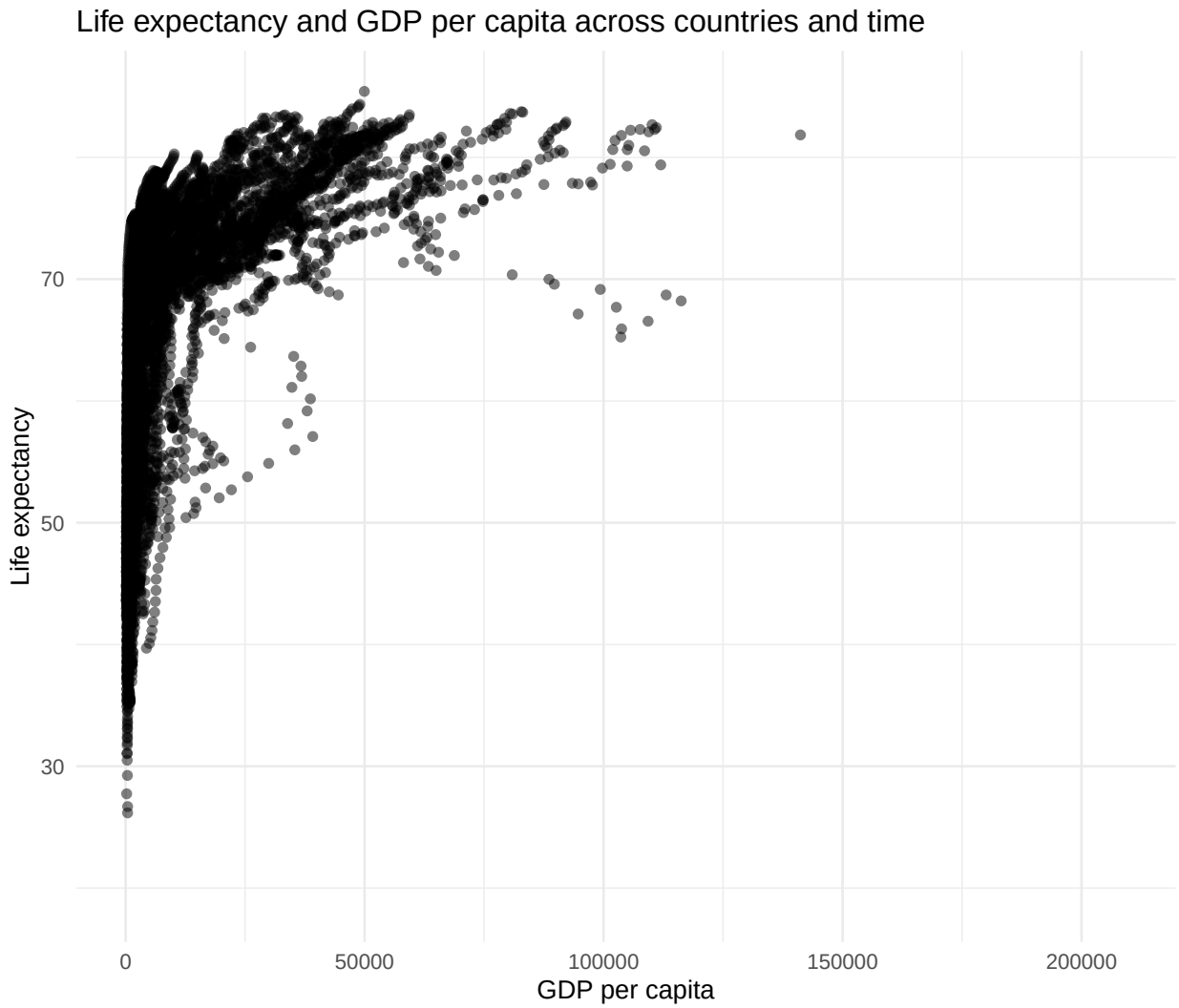
```
data_qog %>%
  group_by(my_region, year) %>%
  summarize(average_lifexp = mean(wdi_lifexp, na.rm = TRUE)) %>%
  ungroup() -> Q21_tibble


data_qog %>%
  group_by(my_region, year) %>%
  summarize(average_lifexp = mean(wdi_lifexp, na.rm = TRUE)) %>%
  ungroup() %>%
  ggplot()+
  aes(x = year, y = average_lifexp, color = my_region, shape = my_region)+
  geom_line()+
  geom_point(size = 1)+
```

```
  theme_minimal()+
  theme(legend.position = "bottom")+
  labs(title = "Average life expectancy at birth",
       caption = "Average number of years a newborn infant would live if prevailing patterns of mortali
```



Q22

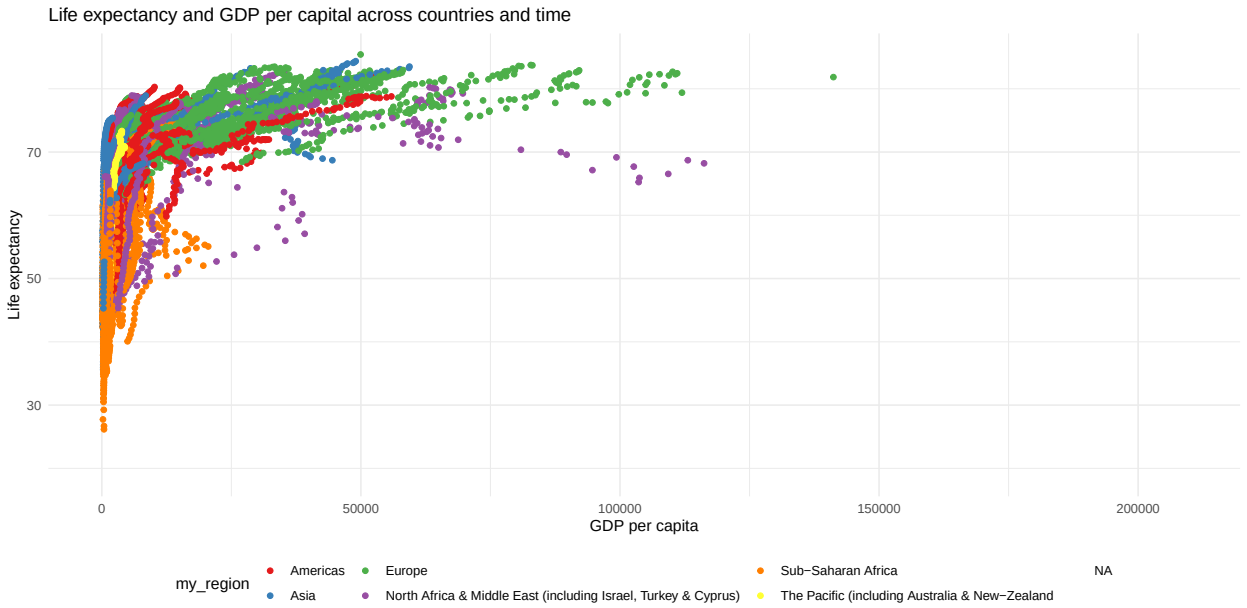
```
data_qog %>%
  ggplot()+
  aes(x = wdi_gdpcapcon2010, y = wdi_lifexp)+
  geom_point(alpha = 0.5)+
  theme_minimal()+
  labs(title = "Life expectancy and GDP per capita across countries and time")+
  xlab("GDP per capita")+
  ylab("Life expectancy")
```



There seems to be a positive correlation between GDP per cap and life expectancy.

Q23

```
data_qog %>%
  ggplot()+
  aes(x = wdi_gdpcapcon2010, y = wdi_lifexp, color = my_region)+
  geom_point(alpha = 1)+
  theme_minimal()+
  theme(legend.position = "bottom")+
  scale_color_brewer(palette = "Set1")+
  labs(title = "Life expectancy and GDP per capital across countries and time")+
   xlab("GDP per capita")+
  ylab("Life expectancy")
```

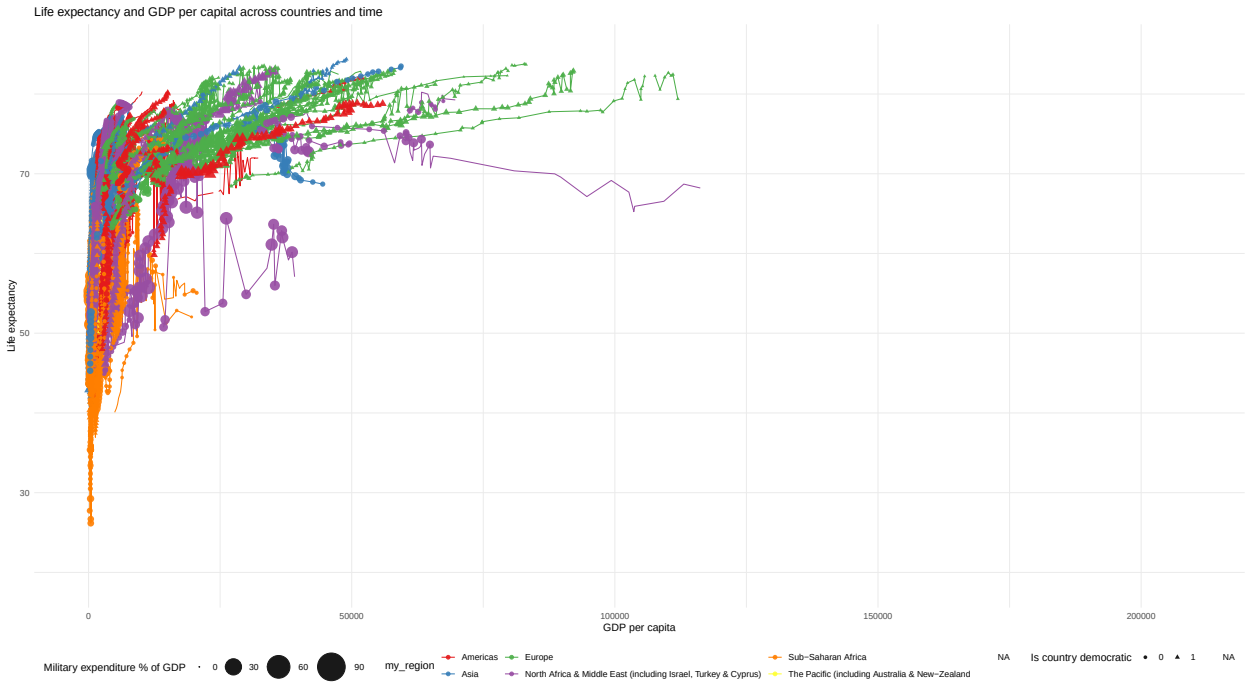


- A region with low GDP per cap and low life expectancy is Sub-Saharan Africa
- Regions with high GDP per cap and high life expectancy are Europe and America
- A region for which life expectancy tend to be relatively lower than their GDP per cap is North Africa

North Africa and Middle East is an interesting case whose observations stand out in the graph above. Around 50000 gdp per cap, life expectancy seems to even decrease with gdp. Thus, I am going to include several variable to see any pattern.

I reproduce the graph by adding two variables: br_dem and wdi_expmil, my guess is that countries which are democratic and with few military expenditures are expected to have high gdp and high life expectancy (for example because less money in invested in the military and more in health care) and conversely. I expect the countries which had a declining trend in the previous graph to have relatively high military expenditure. I also add lines aesthetics to distinguish countries within the regions.

```
data_qog %>%
  ggplot()+
  aes(x = wdi_gdpcapcon2010, y = wdi_lifexp, color = my_region, shape = factor(br_dem), size = wdi_expmi
  geom_line(aes(x = wdi_gdpcapcon2010, y = wdi_lifexp, fill = cname, color = my_region), inherit.aes = F
  geom_point(alpha = 0.9)+
  theme_minimal()+
  theme(legend.position = "bottom")+
  scale_color_brewer(palette = "Set1")+
  scale_size_continuous(name = "Military expenditure % of GDP",
                        range = c(0, 15))+
  scale_shape_discrete(name = "Is country democratic")+
  labs(title = "Life expectancy and GDP per capital across countries and time")+
  xlab("GDP per capita")+
  ylab("Life expectancy")
```



I can see now that European countries tend to have relatively low military expenditure, have a positive correlation between life expectancy and GDP per capita and be democratic. Sub-Saharan Africa countries tend to have increasing life expectancy over the years, but no important GDP per capita growth. For Asia, There is also a positive correlation between life expectancy and gdp per cap, except for one country. The difference with Europe is that some Asian countries are not democratic (circle = non democratic). North Africa and Middle East is the most intringuing case, with a non democratic country with very high military expenditure and high fluctuation of life expectancy and another non democratic country with negative correlation between life expectancy and GDP per cap.

Q24

```
typeof(data_qog$br_dem)
```

```
## [1] "double"
```

```
typeof(data_qog$p_polity2)
```

```
## [1] "double"
```

br_dem (is the country a democracy) is stored as a double, it is however a dummy variable (nominal variable) taking value 0 is country is not democratic and 1 if country is a democracy. Thus, it does not make sense to compute its mean, I should make a table with the frequencies or relative frequencies or do a barplot.

p_polity2 is also stored as a double. It also does not make sense because p_polity2 (revised combined polity score) is an ordinal variable ranging from -10 (strongly autocratic) to +10 (strongly democratic). To give information about this variable, I would do a table and a barplot (sorted according to the range).

Q25

I first transform those two variables into factors

```
data_qog$br_dem <- factor(data_qog$br_dem, levels = c(0,1), labels = c("Not democratic", "democratic"))

data_qog$p_polity2 <- factor(data_qog$p_polity2, levels = (-10:10))
```

I then do some univariate and bivariate analysis of the variables through tables and a cross table

```
data_qog %>%
  count(p_polity2) %>%
  kable(caption = "Is the country democratic", align = "c")
```

Table 11: Is the country democratic

p_polity2	n
-10	287
-9	606
-8	328
-7	1311
-6	419
-5	190
-4	231
-3	240
-2	148
-1	177
0	224
1	114
2	95
3	123
4	252
5	334
6	469
7	386
8	593
9	516
10	1577
NA	3536

```
data_qog %>%
  count(br_dem) %>%
  kable(caption = "Is the country democratic", align = "c")
```

Table 12: Is the country democratic

br_dem	n
Not democratic	5140
democratic	4915

br_dem	n
NA	2101

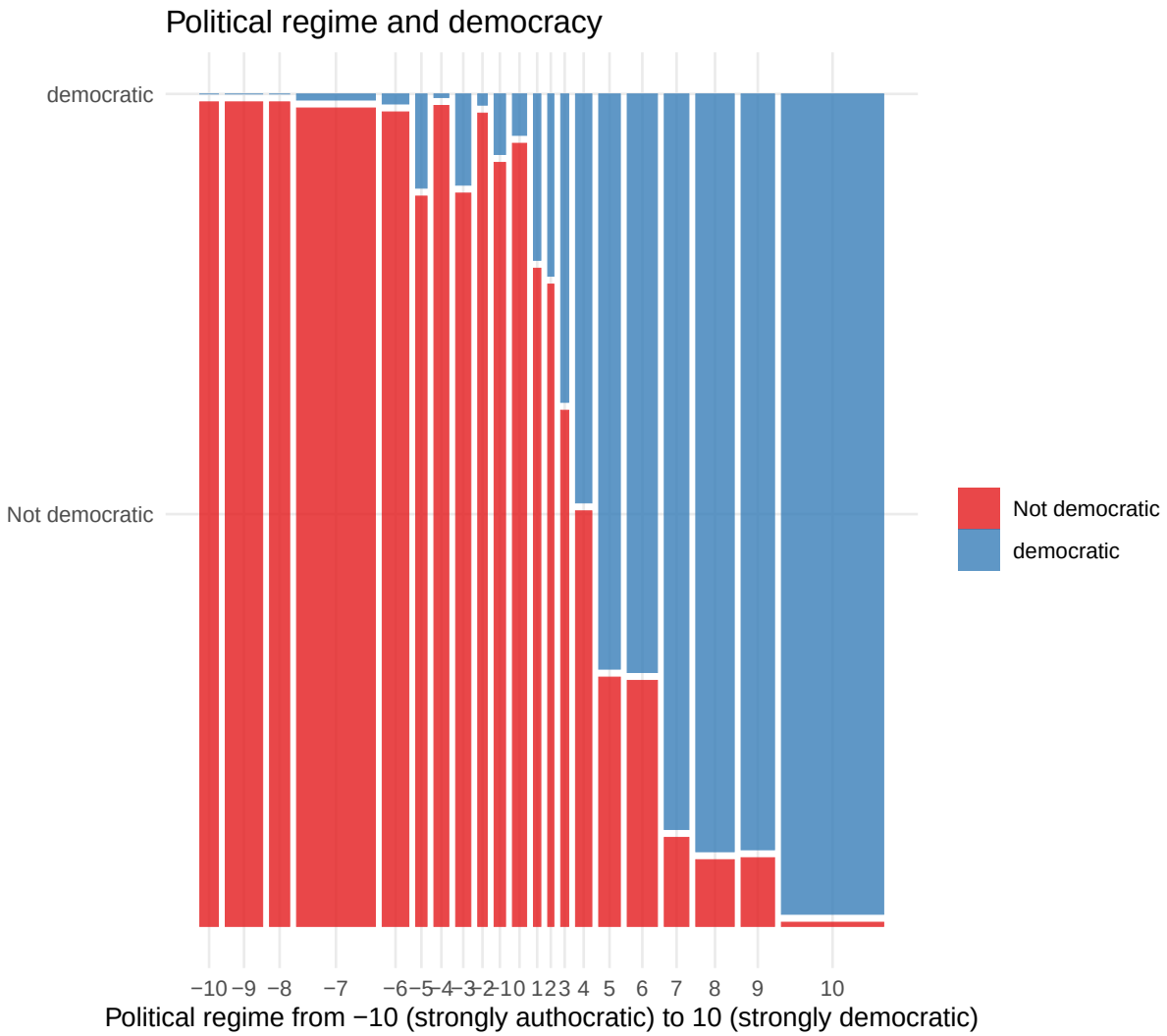
```
kable(table(data_qog$p_polity2, data_qog$br_dem), caption = "Revised combined polity score according to
```

Table 13: Revised combined polity score according to democracy or not

	Not democratic	democratic
-10	287	0
-9	577	0
-8	311	0
-7	1206	9
-6	406	5
-5	155	20
-4	229	1
-3	209	26
-2	145	2
-1	164	13
0	208	11
1	91	23
2	74	21
3	77	46
4	127	125
5	101	233
6	139	327
7	41	339
8	48	545
9	43	473
10	8	1569

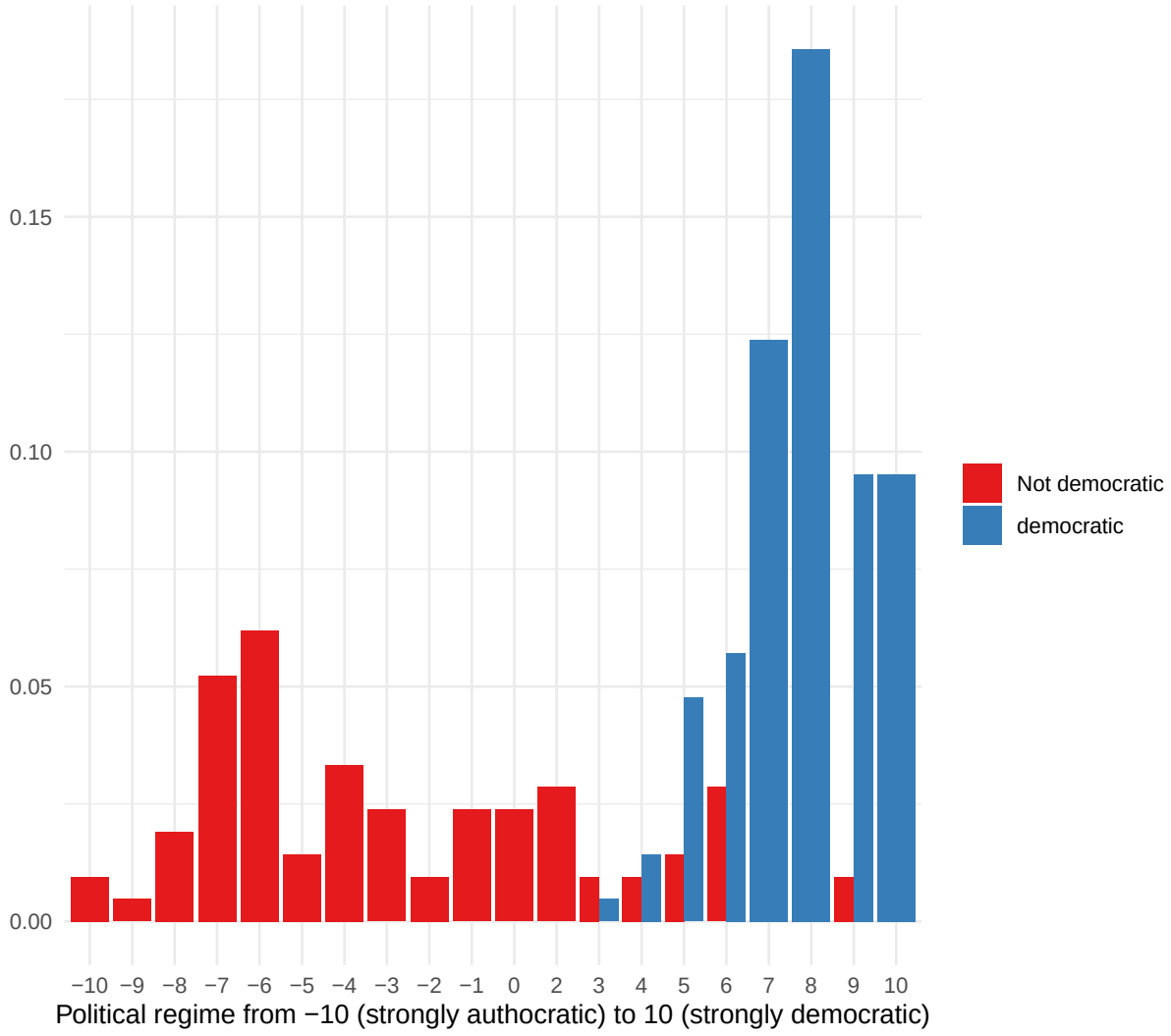
I choose then to do a mosaic plot of p_polity according to br_dem to represent the association of the two variable visually

```
data_qog %>% ggplot()+
  geom_mosaic(aes(x = product(br_dem, p_polity2), fill = br_dem), na.rm = TRUE)+
  theme_minimal()+
  scale_fill_brewer(palette = "Set1", name = "")+
  ylab("")+
  xlab("Political regime from -10 (strongly authocratic) to 10 (strongly democratic)")+
  ggtitle("Political regime and democracy")
```

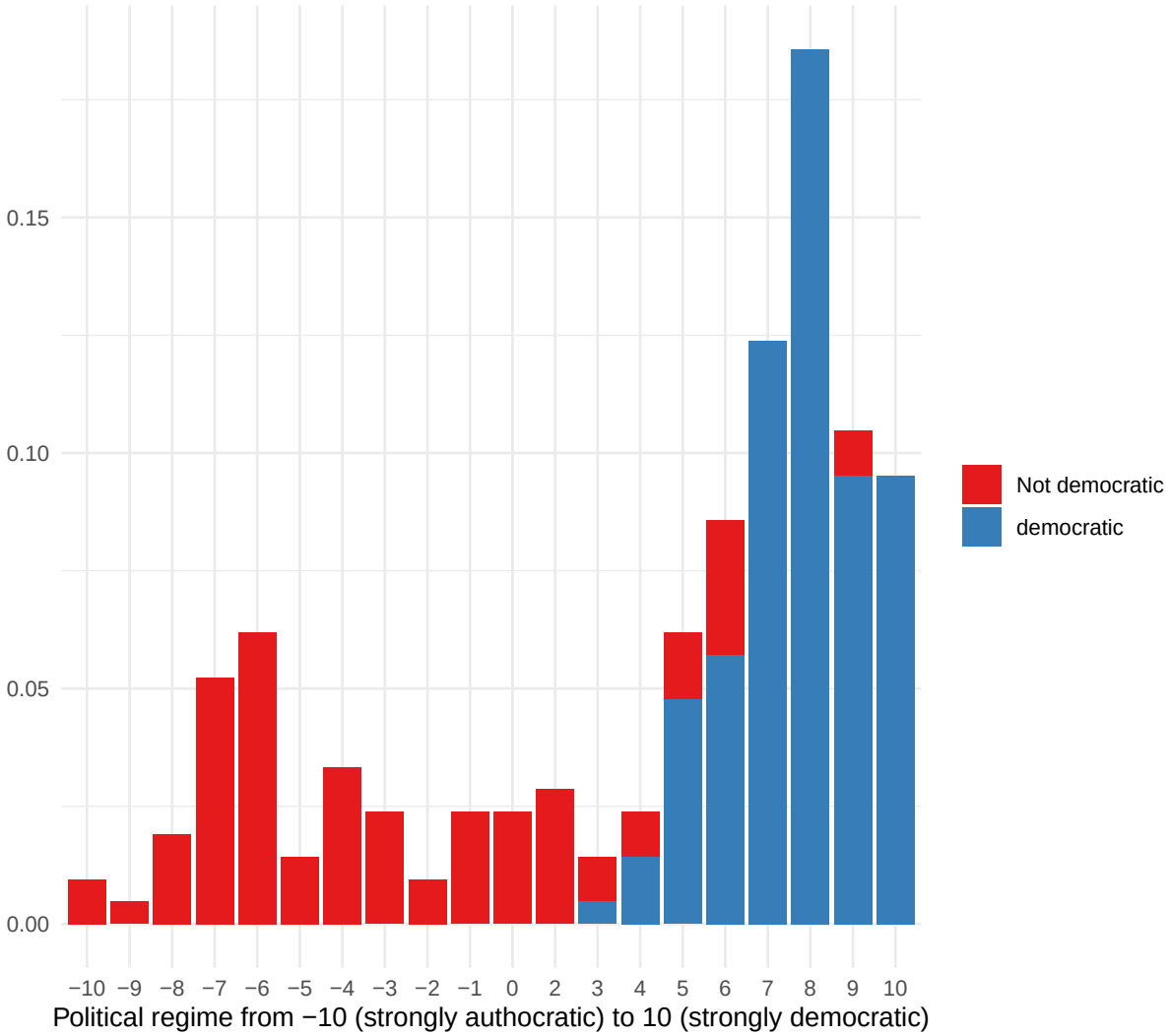


Another possiblity is to do a stacked or dodge barplot of p_polity and br_dem

```
data_qog %>%
  drop_na() %>%
  ggplot()+
  aes(x = p_polity2, y = (..count..)/sum(..count..), fill = br_dem)+
  geom_bar(position = "dodge")+
  scale_fill_brewer(palette = "Set1", name = "")+
  ylab("")+
  xlab("Political regime from -10 (strongly authocratic) to 10 (strongly democratic)")+
  theme_minimal()
```




```
data_qog %>%
  drop_na() %>%
  ggplot()+
  aes(x = p_polity2, y = (..count..)/sum(..count..), fill = br_dem)+
  geom_bar()+
  scale_fill_brewer(palette = "Set1", name = "")+
  ylab("")+
  xlab("Political regime from -10 (strongly authocratic) to 10 (strongly democratic)")+
  theme_minimal()
```



Bonus task

For the bonus task, I would like to try to do some maps to visualize some of the dataset variables

I first load the packages required to do maps on r

```
library(rgdal)
library(countrycode)
library(gganimate)
library(lubridate)
library(sf)
library(viridis)
library("rnaturalearth")
library("rnaturalearthdata")
library(maps)
library(maptools)
library(rgeos)
library(biscale)
```

```
library(cowplot)

map <- readOGR("D:/MAPEC/Applied methods/assignment 2/world map/ne_50m_admin_0_countries.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "D:\MAPEC\Applied methods\assignment 2\world map\ne_50m_admin_0_countries.shp", layer: "ne_5
## with 242 features
## It has 168 fields
## Integer64 fields read as strings:  NE_ID
```

I then have to "fortify" the map data to make it tidyverse compatible

```
map_fortify <- fortify(map, region = "SOVEREIGNT")
unique(map_fortify$id)
```

```
##   [1] "Afghanistan"                    "Albania"
##   [3] "Algeria"                        "Andorra"
##   [5] "Angola"                         "Antarctica"
##   [7] "Antigua and Barbuda"            "Argentina"
##   [9] "Armenia"                        "Australia"
##  [11] "Austria"                        "Azerbaijan"
##  [13] "Bahrain"                        "Bangladesh"
##  [15] "Barbados"                       "Belarus"
##  [17] "Belgium"                        "Belize"
##  [19] "Benin"                          "Bhutan"
##  [21] "Bolivia"                        "Bosnia and Herzegovina"
##  [23] "Botswana"                       "Brazil"
##  [25] "Brunei"                         "Bulgaria"
##  [27] "Burkina Faso"                   "Burundi"
##  [29] "Cabo Verde"                     "Cambodia"
##  [31] "Cameroon"                       "Canada"
##  [33] "Central African Republic"       "Chad"
##  [35] "Chile"                          "China"
##  [37] "Colombia"                       "Comoros"
##  [39] "Costa Rica"                     "Croatia"
##  [41] "Cuba"                           "Cyprus"
##  [43] "Czechia"                        "Democratic Republic of the Congo"
##  [45] "Denmark"                        "Djibouti"
##  [47] "Dominica"                       "Dominican Republic"
##  [49] "East Timor"                     "Ecuador"
##  [51] "Egypt"                          "El Salvador"
##  [53] "Equatorial Guinea"              "Eritrea"
##  [55] "Estonia"                        "eSwatini"
##  [57] "Ethiopia"                       "Federated States of Micronesia"
##  [59] "Fiji"                           "Finland"
##  [61] "France"                         "Gabon"
##  [63] "Gambia"                         "Georgia"
##  [65] "Germany"                        "Ghana"
##  [67] "Greece"                         "Grenada"
##  [69] "Guatemala"                      "Guinea"
##  [71] "Guinea-Bissau"                  "Guyana"
##  [73] "Haiti"                          "Honduras"
```

```
##  [75] "Hungary"                           "Iceland"
##  [77] "India"                             "Indonesia"
##  [79] "Iran"                              "Iraq"
##  [81] "Ireland"                           "Israel"
##  [83] "Italy"                             "Ivory Coast"
##  [85] "Jamaica"                           "Japan"
##  [87] "Jordan"                            "Kashmir"
##  [89] "Kazakhstan"                        "Kenya"
##  [91] "Kiribati"                          "Kosovo"
##  [93] "Kuwait"                            "Kyrgyzstan"
##  [95] "Laos"                              "Latvia"
##  [97] "Lebanon"                           "Lesotho"
##  [99] "Liberia"                           "Libya"
## [101] "Liechtenstein"                     "Lithuania"
## [103] "Luxembourg"                        "Madagascar"
## [105] "Malawi"                            "Malaysia"
## [107] "Maldives"                          "Mali"
## [109] "Malta"                             "Marshall Islands"
## [111] "Mauritania"                        "Mauritius"
## [113] "Mexico"                            "Moldova"
## [115] "Monaco"                            "Mongolia"
## [117] "Montenegro"                        "Morocco"
## [119] "Mozambique"                        "Myanmar"
## [121] "Namibia"                           "Nauru"
## [123] "Nepal"                             "Netherlands"
## [125] "New Zealand"                       "Nicaragua"
## [127] "Niger"                             "Nigeria"
## [129] "North Korea"                       "North Macedonia"
## [131] "Northern Cyprus"                   "Norway"
## [133] "Oman"                              "Pakistan"
## [135] "Palau"                             "Panama"
## [137] "Papua New Guinea"                  "Paraguay"
## [139] "Peru"                              "Philippines"
## [141] "Poland"                            "Portugal"
## [143] "Qatar"                             "Republic of Serbia"
## [145] "Republic of the Congo"             "Romania"
## [147] "Russia"                            "Rwanda"
## [149] "Saint Kitts and Nevis"             "Saint Lucia"
## [151] "Saint Vincent and the Grenadines" "Samoa"
## [153] "San Marino"                        "São Tomé and Principe"
## [155] "Saudi Arabia"                      "Senegal"
## [157] "Seychelles"                        "Sierra Leone"
## [159] "Singapore"                         "Slovakia"
## [161] "Slovenia"                          "Solomon Islands"
## [163] "Somalia"                           "Somaliland"
## [165] "South Africa"                      "South Korea"
## [167] "South Sudan"                       "Spain"
## [169] "Sri Lanka"                         "Sudan"
## [171] "Suriname"                          "Sweden"
## [173] "Switzerland"                       "Syria"
## [175] "Taiwan"                            "Tajikistan"
## [177] "Thailand"                          "The Bahamas"
## [179] "Togo"                              "Tonga"
## [181] "Trinidad and Tobago"               "Tunisia"
```

```
## [183] "Turkey"                          "Turkmenistan"
## [185] "Tuvalu"                          "Uganda"
## [187] "Ukraine"                         "United Arab Emirates"
## [189] "United Kingdom"                  "United Republic of Tanzania"
## [191] "United States of America"        "Uruguay"
## [193] "Uzbekistan"                      "Vanuatu"
## [195] "Vatican"                         "Venezuela"
## [197] "Vietnam"                         "Western Sahara"
## [199] "Yemen"                           "Zambia"
## [201] "Zimbabwe"
```

The biggest problem to plot a map will the names of all the countries in the two datasets to successfully join the dataset to the map-dataset, I have to recode the name in one dataset. The package countrycode will do most of the work, but some countries will still need to be recoded

```
data_qog$cname <- countryname(data_qog$cname, destination = "cldr.short.en")
data_qog$cname[data_qog$cname == "US"] <- "United States of America"
data_qog$cname[data_qog$cname == "UK"] <- "United Kingdom"
data_qog$cname[data_qog$cname == "Congo - Kinshasa"] <- "Democratic Republic of the Congo"
data_qog$cname[data_qog$cname == "Côte d'Ivoire"] <- "Ivory Coast"
data_qog$cname[data_qog$cname ==  "Tanzania"] <- "United Republic of Tanzania"
```

I joine the datasets by the country names

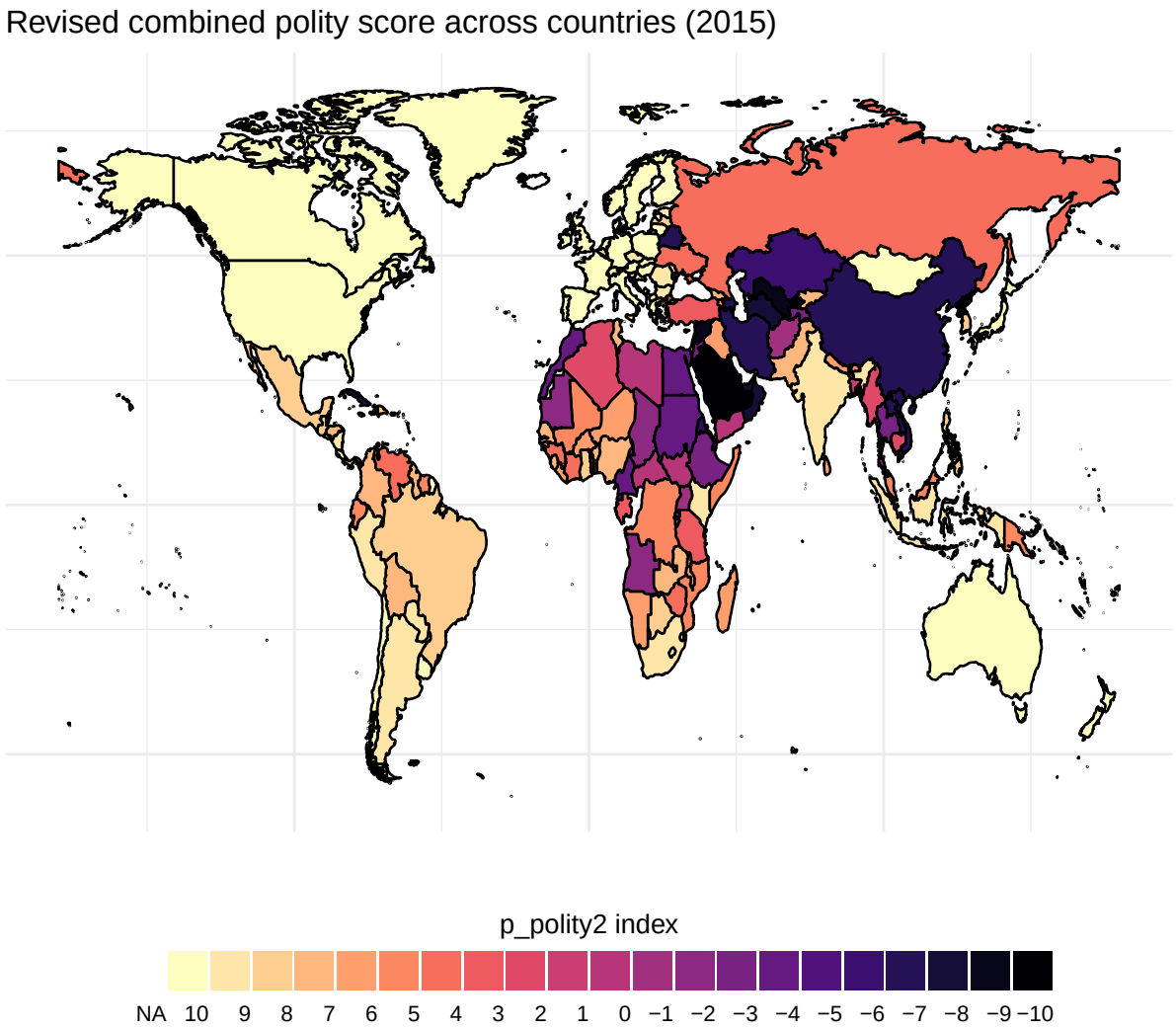
```
map_join <- map_fortify %>% left_join(data_qog, by = c("id" = "cname"))
```

Are military expenditure geographically related to the probability of a country to be autocratic? I will explore with two maps the relation between p_polity and wdi_expmil. I expect countries with high military expenditure to be more authorcratic.

I then plot the first map with the variable p_polity2

```
map_join %>%
  filter(year == 2015) %>%
  ggplot()+
  aes(x = long, y = lat, group = group, fill = factor(p_polity2), color = I("black"))+
  geom_polygon()+
  theme_minimal()+
  scale_fill_ordinal(
          guide = guide_legend(
            direction = "horizontal",
            title.position = 'top',
            title.hjust = 0.5,
            label.hjust = 1,
            nrow = 1,
            byrow = TRUE,
            reverse = TRUE,
            label.position = "bottom"
          ),
          option = "magma",
          name = "p_polity2 index"
      )+
  theme(legend.position = "bottom",
```

```
        axis.text.x = element_blank(),
        axis.text.y = element_blank()
        )+
  labs(title = "Revised combined polity score across countries (2015)")+
  ylab("")+
  xlab("") -> map_plot

map_plot
```

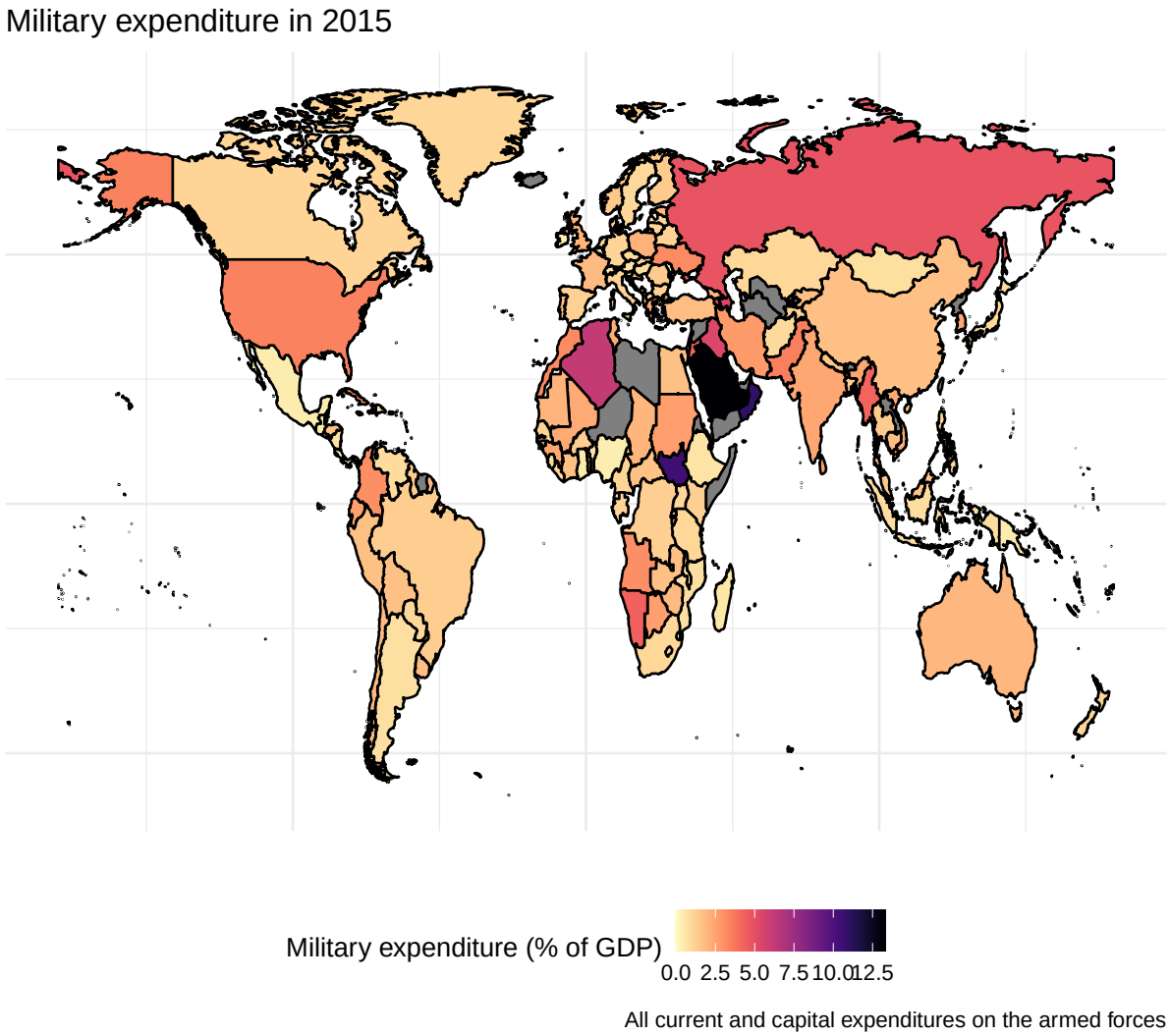


The more democratic countries in 2015 are concentrated in western Europe, America, East and South East Asia, South Pacific with some exception for example India and Mongolia.

I now plot the second map military expenditure wdi_expmil:

```
map_join %>%
  filter(year == 2015) %>%
  ggplot()+
  aes(x = long, y = lat, group = group, fill = wdi_expmil, color = I("black"))+
```

```
geom_polygon()+
theme_minimal()+
ylab("")+
xlab("")+
scale_fill_viridis(
        option = "magma",
        direction = -1,
        name = "Military expenditure (% of GDP)"
    )+
theme(legend.position = "bottom",
      axis.text.x = element_blank(),
      axis.text.y = element_blank()
      )+
labs(title = "Military expenditure in 2015",
     caption = "All current and capital expenditures on the armed forces")
```



Seeing the latter map on military expenditure, I would conclude that there is no clear inverse correlation between degree of democracy and military expenditure: there are countries which are strongly democratic

which have relatively high military expenditure such as the US. A lot of African countries are autocratic, but do not have a really high military expenditure except for Algeria and South Soudan. The only region for which there is a correlation between autocracy and military expenditure seems to be Middle East, with Saudi Arabia being the best illustration.