# Analysis of Ecological Footprint and Biocapacity

*Hsin-Yu,Ku (Matrikel Num. : 3038591), Feb. 2020*

Out of love for our earth, the latest edtion of National Footprint and Biocapacity Accounts 2019 (NFA 2019), an dataset published by Global Footprint Network is chosen for analysis.

NFA 2019 is a collected dataset(see A guide book of NFA 2019) to measure Ecological Footprint(EF) and biocapacity(BCP) of 193 countries (including "world") for over 50 years (1961-2016).

> BCP measures how much nature we have, while EF shows how much nature we use. (The latter can be viewed as consumption. EF & consumption mean the same.)

PS. In this report, *Global hectares(gha)*, — a globally comparable, standardized hectares – is the only unit of measure. It is a unit of land normalized by biological productivity across landtype, country, and year.

(See related key terms, full glossary, as well as Data and method for more.)

## Create and prepare the dataset

For region-wise analysis, I combined 2 files of NFA 2019 with the all.csv file on github by ISO 3166 standard country code.

### Read in 3 datasets

```
data <- read_csv("https://query.data.world/s/fbvknrt2w5nmufj6iue6cv24v4utl2")
country <- read_csv("https://query.data.world/s/mgysbl7mcnisg7akii776dangvktnx")
  names(country) <- c("country","country_code","alpha2","alpha3")
region <- read_csv("all.csv") %>%
  select(c(country="name",country_code = "country-code", "region", "sub-region",
           alpha2 = "alpha-2", alpha3 = "alpha-3"))
```

### Combine 3 datasets from different sources (keep only the useful variables)

```
data_coun <- merge(data, country, by=c("country","country_code"))
data_regi <- merge(data_coun, region, all.x=T, all.y=F,
                   by=c("alpha2","alpha3")) %>% rename(country="country.x") %>%
  select(-c("alpha2","alpha3","country_code.x","country_code.y","country.y","QScore")) %>% glimpse
```

```
## Observations: 72,186
## Variables: 12
## $ country        <chr> "United Arab Emirates", "United Arab Emirates",...
## $ year           <dbl> 1999, 2000, 2000, 2001, 2001, 2002, 2002, 2002,...
## $ record         <chr> "AreaPerCap", "EFProdPerCap", "EFProdTotGHA", "...
## $ crop_land      <dbl> 7.228529e-02, 2.083519e-01, 6.573355e+05, 7.877...
## $ grazing_land   <dbl> 1.020695e-01, 3.598003e-03, 1.135145e+04, 9.170...
## $ forest_land    <dbl> 1.015675e-01, 1.952233e-03, 6.159158e+03, 9.332...
## $ fishing_ground <dbl> 1.709413e+00, 3.944098e-01, 1.244335e+06, 1.535...
## $ built_up_land  <dbl> 3.169650e-02, 0.000000e+00, 0.000000e+00, 3.169...
## $ carbon         <dbl> 0.000000e+00, 9.201178e+00, 2.902907e+07, 0.000...
## $ total          <dbl> 2.017032e+00, 9.809490e+00, 3.094825e+07, 1.831...
## $ region         <chr> "Asia", "Asia", "Asia", "Asia", "Asia", "Asia",...
## $ `sub-region`   <chr> "Western Asia", "Western Asia", "Western Asia",...
```

### Overview of the variables

- record: measurement of **EF** ("EFCons...") & **BCP**("Biocap...") in gha.
- From *crop_land* to *carbon* : 6 columns including 5 landtyes and carbon. The number (in gha) of a landtype either required to support EFc, or that are supported by BP. (column "total" is the sum of these 6 columns.)

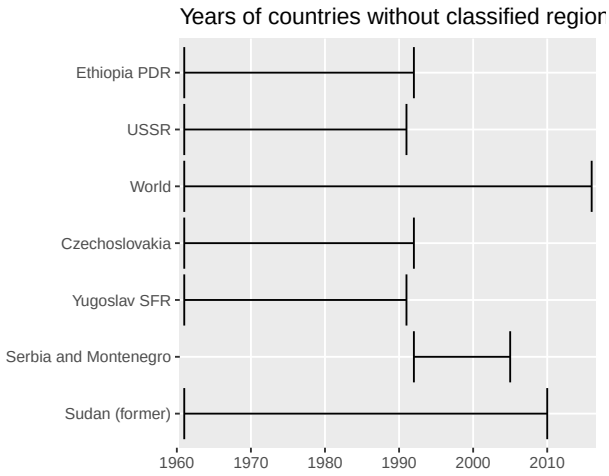
## Examine missing data and rearrange dataset

```
# See numbers of missing values for each variables
apply(is.na(data_regi),2,sum)
```

```
##        country           year         record      crop_land   grazing_land
##              0              0              0          20472          20472
##    forest_land fishing_ground  built_up_land         carbon          total
##          20473          20473          20473          20473              9
##         region     sub-region
##           1968           1968
```

### Deal with missing values in variable "region"

```
data_regi %>%
  filter(is.na(region)) %>% group_by(country) %>%
  summarize(begin = min(year), end = max(year)) %>% ungroup() %>%
  mutate(region_country = paste(region, country),
         country = fct_reorder(as.factor(country), desc(region_country))) %>%
  ggplot(aes(y = country)) + geom_errorbarh(aes(xmin = begin, xmax = end)) +
  scale_x_discrete(limits=seq(1960, 2016, 10))+
  labs(x="", y="", title="Years of countries without classified regions")
```




After research, it shows that these countries are not classified with a corresponding region due to historical reasons :

- Exist before 1992: Czechoslovakia / Ethiopia PDR / USSR / Yugoslav SFR
- Exist between 1992-2003: Serbia and Montenegro
- Exist before 2010: Sudan (former)

Besides, "World" doesn't correspond to any region in nature.

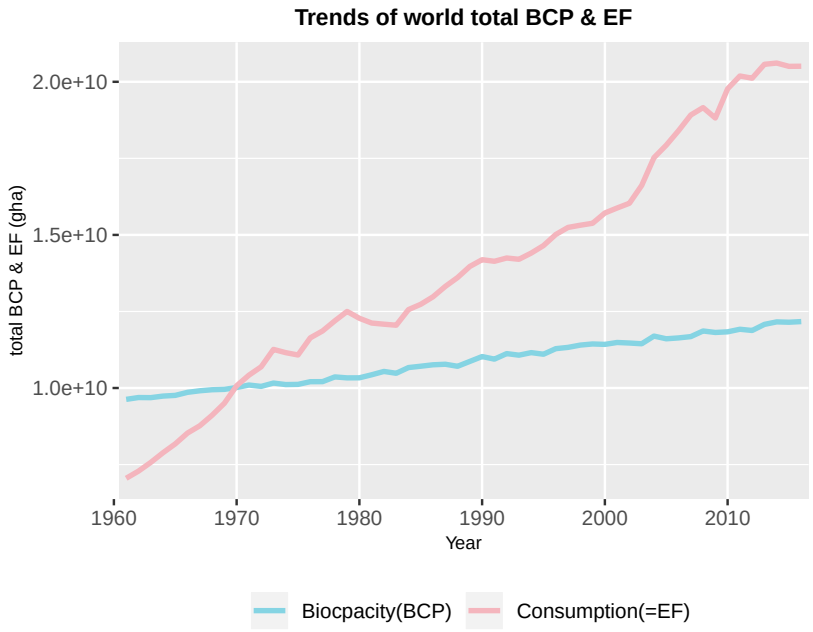
**Manually assign regions to countries**

```
assign_region <- function(Country, Region){
  data_no_coun <- data_regi %>% filter(!country==Country)
  data_coun <- data_regi %>% filter(country==Country) %>%
    mutate_at(vars(c("region","sub-region")), ~replace(.,is.na(.), Region))
  data_regi <<- bind_rows(data_no_coun , data_coun)
}
assign_region("World", "World")
assign_region("USSR", "Europe")
assign_region("Czechoslovakia", "Europe")
assign_region("Yugoslav SFR", "Europe")
assign_region("Serbia and Montenegro", "Europe")
assign_region("Sudan (former)", "Africa")
assign_region("Ethiopia PDR", "Africa")
```

Let's start to look into some trends !

## Global trends in biocapacity(BCP) & consumption(=EF)

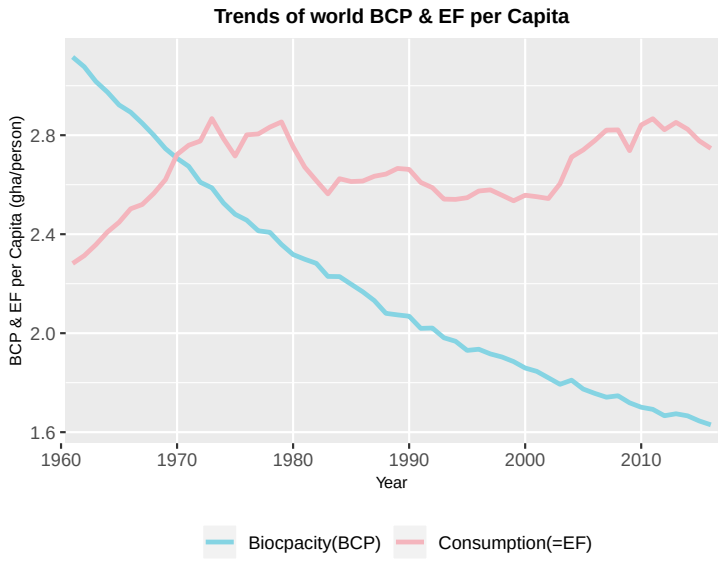
```
trends("BiocapTotGHA", "EFConsTotGHA", "World") +
  labs(x = "Year", y = "total BCP & EF (gha)", title = "Trends of world total BCP & EF")
```




At the scale of the world, both BCP and consumption(=EF) are growing.

- Since 1970, consumption overrides biocapacity
- increase rate : consumption >> biocapacity

We know that the world population keeps growing. What are EF and BCP per capita?

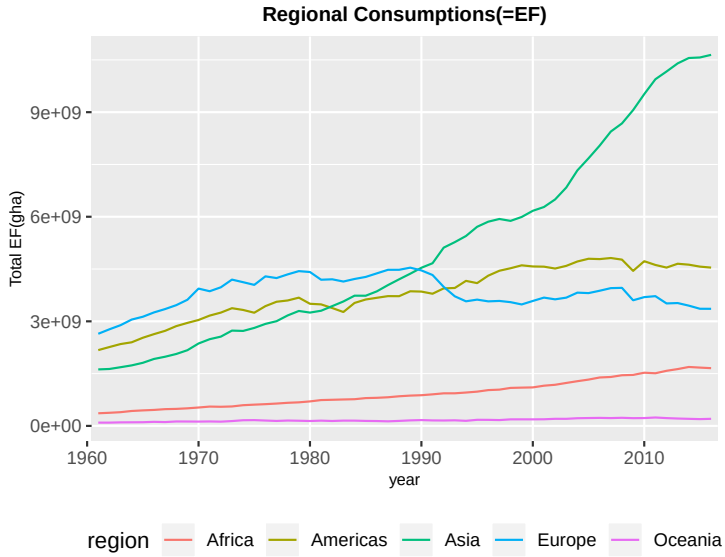



The former graph shows that **worldwide BCP grew slightly**, however this graph hightlights **BCP per capita slumped all the way down due to growing population.** On the other hand, EF per Capita flunctated with an overall up trend.

What about regional trends? Each continent might have different patterns. Let's dig into it.

## Regional trends in biocapacity(BCP) & consumption(=EF)

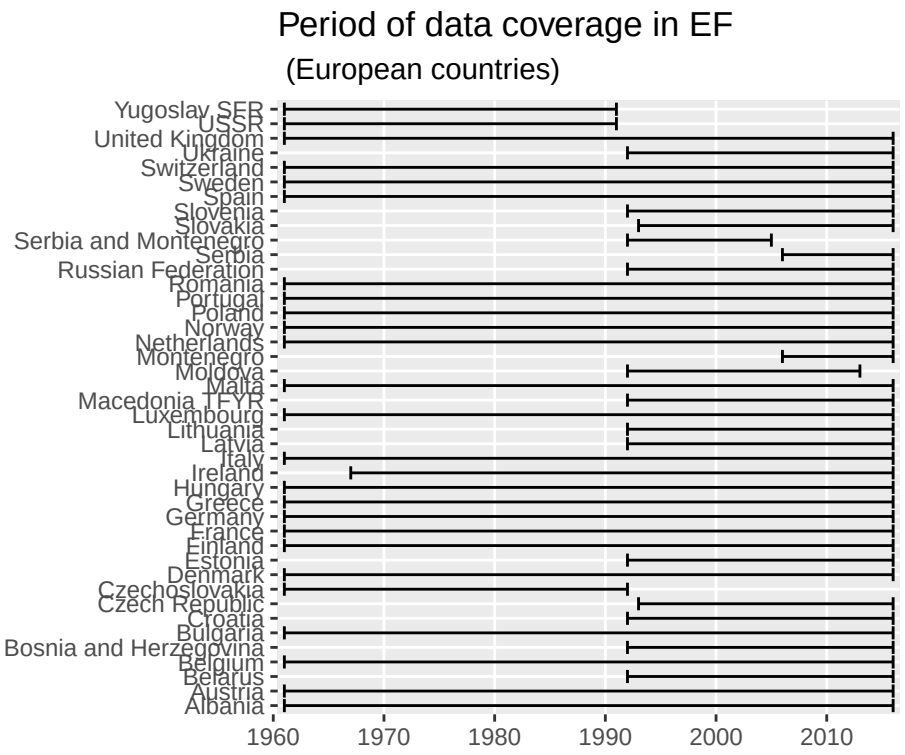
```
Region_plot <- function(Data, Record){
Data %>% filter(record == Record & region !='World') %>%
  ggplot(aes(year, total, col=region)) + geom_line()+ scale_x_discrete(limits=seq(1960, 2016, 10)) +
  theme(plot.title = element_text(size=10, hjust=0.5, face="bold"),
        axis.title=element_text(size=8), legend.position="bottom")
} # write a function for making graphs

data_record <- data_regi %>% filter(record %in% c('BiocapTotGHA', 'EFConsTotGHA')) %>%
  group_by(region, year, record) %>% summarize(total=sum(total))

Region_plot(data_record,'EFConsTotGHA') + ylab('Total EF(gha)') + ggtitle('Regional Consumptions(=EF)')
```



Abviously, Asia is the region with highest growth rate in comsumptions. Notice that *there was a drop in Europe around 1992*. Supposedly it has something to do with *the disolvement of USSR in 1992*.

**Examine the difference of EF around 1992**

```
data_Eur <- data_regi %>% filter(region == "Europe", record=="EFConsTotGHA") %>%
  select(c(country, year, total)) %>% group_by(country) %>%
  summarize(begin = min(year), end = max(year), total=sum(total))

data_Eur %>%
  ggplot(aes(y = country)) + geom_errorbarh (aes(xmin = begin, xmax = end))+
  labs(y="", title='Period of data coverage in EF', subtitle=" (European countries)") +
  scale_x_discrete(limits=seq(1960,2016,10))
```




The graph reveals that around 1992, Yugoslav SFR, USSR, and Czechoslovakia ended. Several countries (ex:Belarus, Slovenia...) that were dissolved from them appeared afterward.

**A closer look in countries that ended and started around 1992:**

```
europe_before1992 <- data_Eur %>%  filter(end < 1993) %>% pull(country)
europe_before1992
```

```
## [1] "Czechoslovakia" "USSR"           "Yugoslav SFR"
```

```
europe_after1992 <- data_Eur %>% filter(begin > 1991) %>% pull(country)
europe_after1992
```

```
##  [1] "Belarus"                "Bosnia and Herzegovina"
##  [3] "Croatia"                "Czech Republic"
```

```
##  [5] "Estonia"                "Latvia"
##  [7] "Lithuania"              "Macedonia TFYR"
##  [9] "Moldova"                "Montenegro"
## [11] "Russian Federation"     "Serbia"
## [13] "Serbia and Montenegro"  "Slovakia"
## [15] "Slovenia"               "Ukraine"
```

```r
Record_Eur <- function(Country, Year, Record){
  data_regi %>% filter(country %in% Country, year==Year, record==Record)%>%
  group_by(country) %>% summarize(Total=sum(total)) %>% pull(Total) %>% sum()
}
```

```r
EF_1992 <- Record_Eur(europe_after1992, 1992, "EFConsTotGHA")
EF_1991 <- Record_Eur(europe_before1992, 1991, "EFConsTotGHA")

EF_1992          # 1434085366
EF_1991          # 1794505148
EF_1992-EF_1991 # -360419782
```

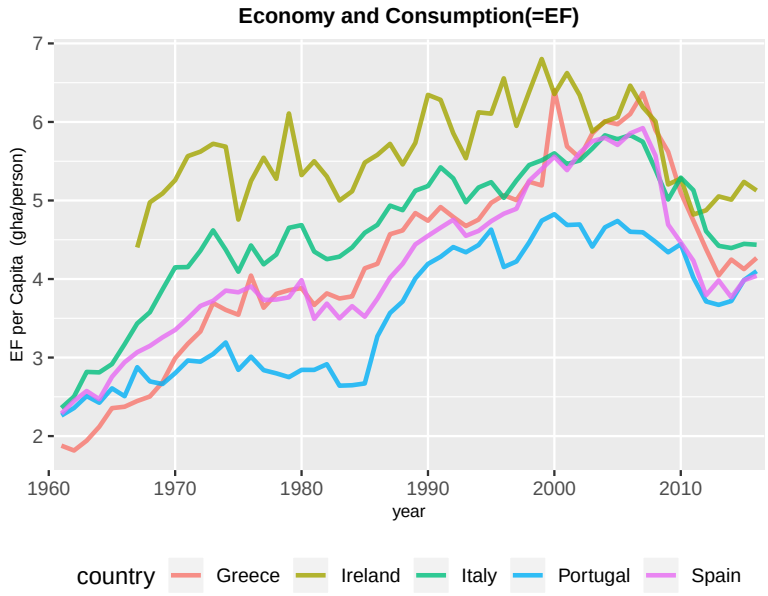
Leaving out the countries that always exist, we see 16 countries existing only after 1991 have less EF (360 M) than 3 countries exisiting only before 1993. This corresponds to the drop of EF in Europe around 1992 in the previous plot.

## Economy and Consumption

Consumption(=EF) is connected to wealth. Could we see *economic flunctuation can in EF*?

For example: Around the financial crisis in the late 2000's, was there changes in EF for *PIIGS*?

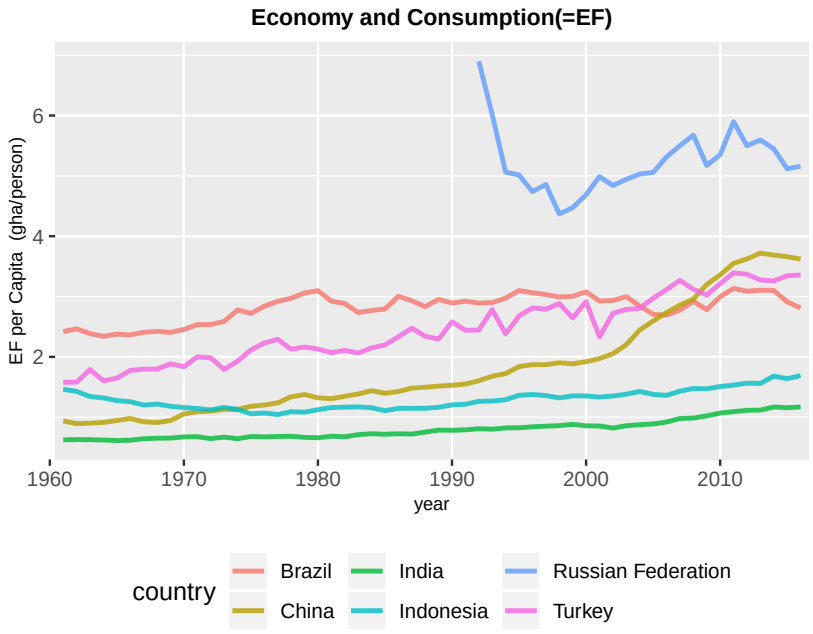
```r
PIIGS <- c('Spain','Greece','Italy','Portugal','Ireland')

eco_EF_plot <- function(Country){
  data_regi %>% filter(country %in% Country & record =='EFConsPerCap') %>%
    ggplot(aes(x=year, y=total, col=country)) + geom_line(alpha=.8,size=1) +
    ylab('EF per Capita  (gha/person)') + ggtitle('Economy and Consumption(=EF)') +
    theme(plot.title = element_text(size=10, hjust=0.5, face="bold"),
          axis.title=element_text(size=8), legend.position = "bottom") +
    scale_x_discrete(limits=seq(1960,2016,10))
}
eco_EF_plot(PIIGS)
```



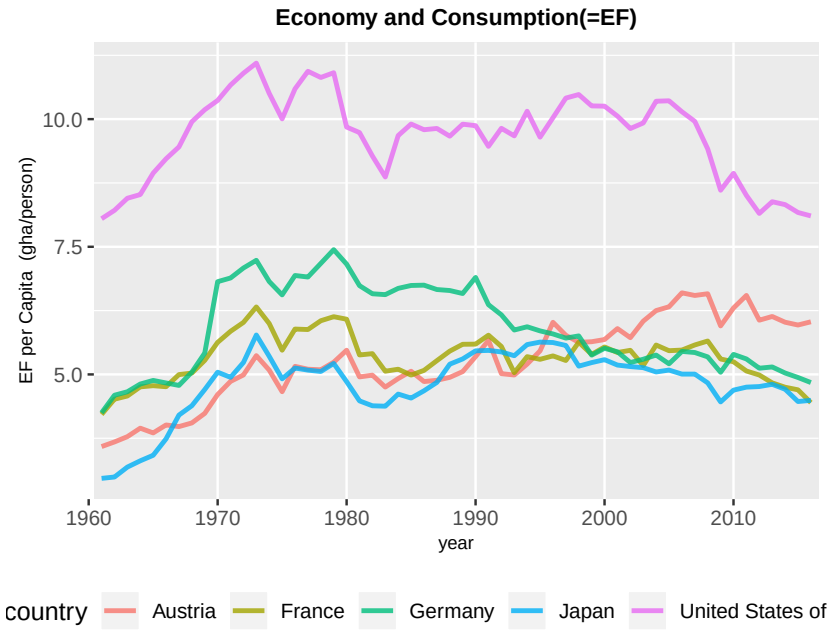
We can see there was an obvious drop in comsumption after the financial crisis in 2008.

Does the financial crisis have different impacts on EF between developed and developing countries?

```
COUN_ing <- c('China','Brazil','India','Russian Federation',"Turkey","Indonesia")
COUN_ed <- c('United States of America','Japan','Germany','Austria',"France")
eco_EF_plot(COUN_ing)
```



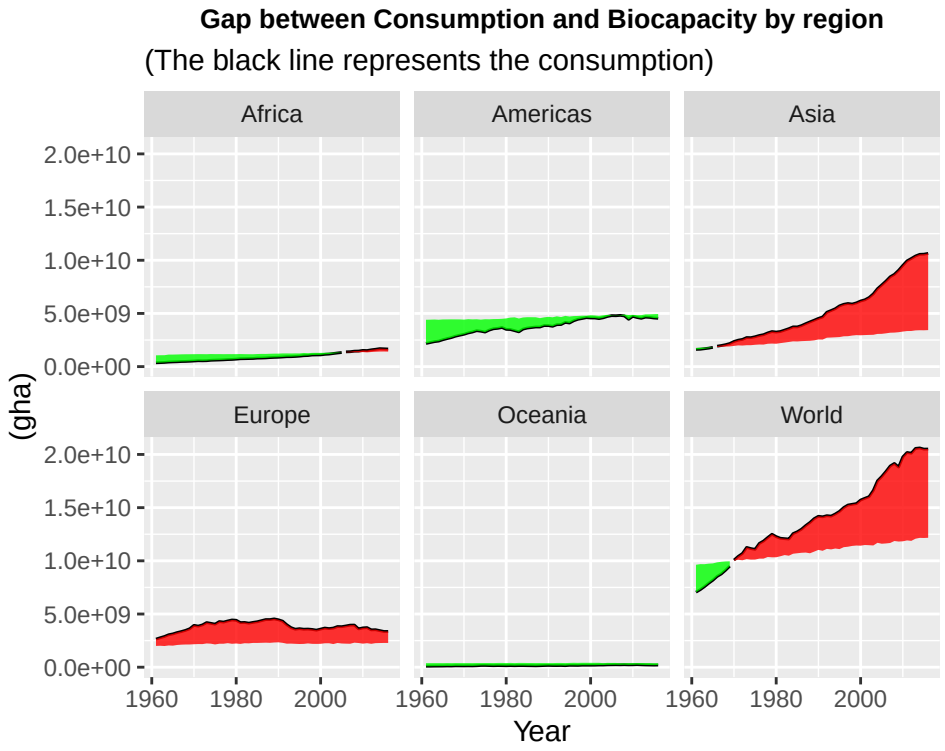

```
eco_EF_plot(COUN_ed)
```




EF in developing countries seems to keep a growing trend regardless of the crisis. On the countrary, EF in developed countries (except Austria) shows overall descending trend in EF after 2000.

We have seen trends for EF so farth. Let's look into comparison between EF and BCP.

```
record_spread <- data_regi %>%
  select(c("country", "region", "year", "total", "record"))%>%
  spread(key = record, value = total) %>%
  select(c("country","region","year","BiocapTotGHA","EFConsTotGHA","EFProdTotGHA")) %>%
  group_by(region, year) %>%
  summarise(BiocapTotGHA = sum(BiocapTotGHA),
            EFConsTotGHA = sum(EFConsTotGHA),
            EFProdTotGHA = sum(EFProdTotGHA))

record_spread %>%
  ggplot(aes(year, y=EFConsTotGHA, ymin = BiocapTotGHA, ymax = EFConsTotGHA,
             fill = BiocapTotGHA < EFConsTotGHA))+
  geom_line(size=.5)+ geom_ribbon(alpha = 0.8) + facet_wrap(~region, scales = "fixed") +
  scale_fill_manual(values=c("green","red"), name="fill")+
  theme(legend.position = "none", plot.title = element_text(size=10, hjust=0.5, face="bold"))+
  labs(x = "Year",y ="(gha)",title = "Gap between Consumption and Biocapacity by region",
       subtitle = "(The black line represents the consumption)")
```




The **green area above the black line(=consumption/EF) represents *ecological reserve* (EF < BCP)**, which is the case for Americas, Oceania all the time, and for Africa until around 2005.

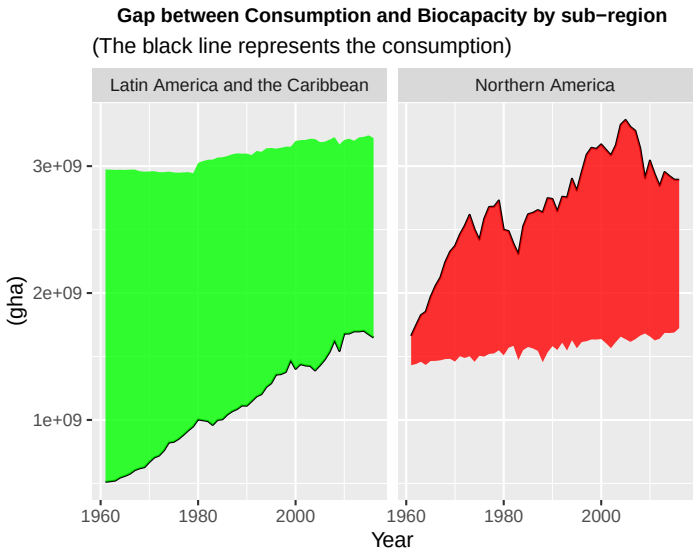
In contrast, the black line(=consumption/EF) above a **red area means *ecological deficit* (EF > BCP)**. This is the case for Europe during all the period, and for Asia since long ago.

The whole world, as seen before, has *ecological deficit* since around 1970.

Let's look deeper into each sub-region (for example: dividing Americas into sub-parts)

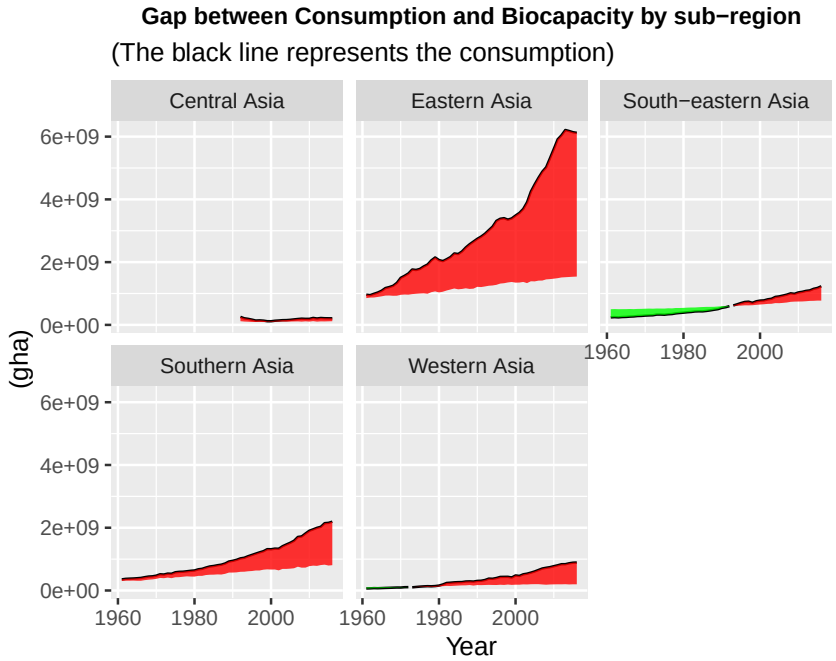
```
subregion_plot <- function(Region){
  record_spread_sub %>% filter(`sub-region` != "World", region==Region) %>%
  ggplot(aes(year, y=EFConsTotGHA, ymin = BiocapTotGHA, ymax = EFConsTotGHA,
```

```
            fill = BiocapTotGHA < EFConsTotGHA))+
  geom_line(size=.5)+ geom_ribbon(alpha = 0.8) +
  facet_wrap(~`sub-region`, scales = "fixed") +
  scale_fill_manual(values=c("green","red"), name="fill")+
  theme(legend.position = "none", plot.title = element_text(size=10, hjust=0.5, face="bold"))+
  labs(x="Year",y="(gha)",title ="Gap between Consumption and Biocapacity by sub-region",
       subtitle = "(The black line represents the consumption)")
}
subregion_plot("Americas")
```




It seems clearer that even Americas as a whole has *ecological reserve*, north america itself is with huge *ecological deficit*.

```
subregion_plot("Asia")
```




```
data_regi %>% filter(`sub-region`=="Eastern Asia", record=="EFConsTotGHA") %>%
  arrange(country, desc(total)) %>% pull(country) %>% unique() %>% head(1)
```

```
## [1] "China"
```

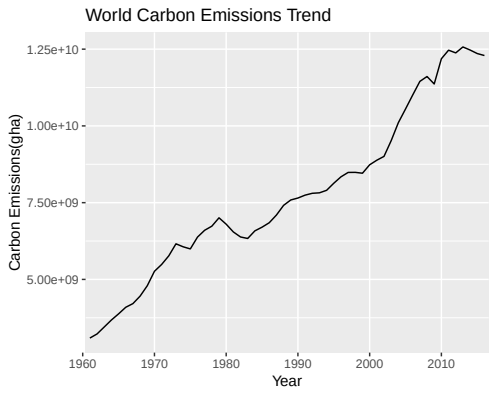
Unsurprisingly, China is the number one with *ecological deficit* in Eastern Asia.

## Carbon emisions(CE)

We have seen overall EF, which is composed by 5 components. Let's look now into trends from one component.

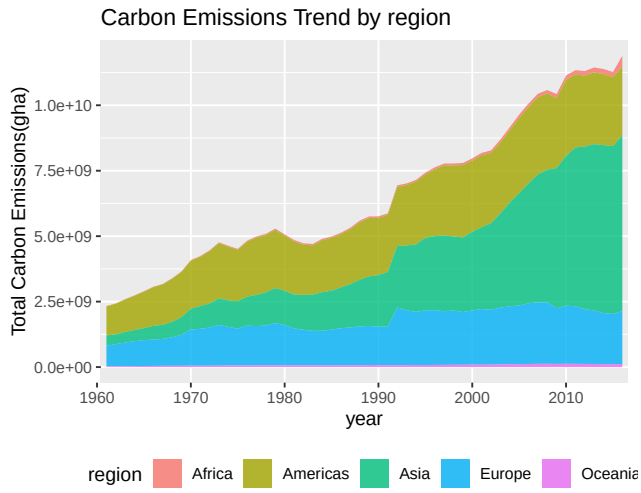
### Trend of Carbon Emissions - Worldwide

```
per_year_continent <- data_regi %>% filter(record=="EFConsTotGHA") %>%
  group_by(year, region) %>% summarise(total_carbon = sum(carbon, na.rm = T))
per_year_continent %>% filter(region =='World') %>%
  ggplot(aes(x=year, y=total_carbon)) + scale_x_discrete(limits=seq(1960,2016,10))+
  geom_line()+ labs(x = "Year", y = "Carbon Emissions(gha)", title = "World Carbon Emissions Trend")
```




### Trend of Carbon Emissions - Per region

```
per_year_continent %>% filter(region != "World")%>%
  ggplot(aes(year, total_carbon)) + geom_area(aes(fill=region), alpha=0.8) +
  ggtitle('Carbon Emissions Trend by region')+ theme(legend.position = "bottom")+
  ylab('Total Carbon Emissions(gha)') +  scale_x_discrete(limits=seq(1960,2016,10))
```




The huge green area depicts the fact that Asia accounts for the largest part among total carbon emission. Also, notice that with time Asia takes up more and more percentage. In comparison, Oceania & Africa have very less carbon emission.