

607 Fall 2021 Project 2 - Energy

Mark Schmalfeld

10/1/2021

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr   0.3.4
## v tibble  3.1.4      v dplyr   1.0.7
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   2.0.1      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(curl)
```

```
## Using libcurl 7.64.1 with LibreSSL/2.8.3
```

```
##
```

```
## Attaching package: 'curl'
```

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

```
##
```

```
##   parse_date
```

```
library(purrr)
```

```
library(lubridate)
```

```
##
```

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

```
## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union
```

Load energy data from github.

Convert characters that are numbers into numeric value. Clean up data set to eliminate NA rows, adjust to update column names, and setup data so it can be transformed into a tidy data setup.

```
urlfile<-"https://raw.githubusercontent.com/schmalmr/607_Fall_2021_Project_2_Energy/main/Temp_primaryen
energy <- read_csv(url(urlfile))
```

```
## Rows: 108 Columns: 59
```

```
## -- Column specification -----
## Delimiter: ","
## chr (59): Gigajoule_per_capita, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 19...
```

```
##
## 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.
```

```
energy<-as_tibble(energy)
energy<- energy[-which(is.na(energy)), ]
energy<-as.data.frame(energy)
energy<-type_convert((energy))
```

```
##
## -- Column specification -----
## cols(
##   .default = col_character()
## )
## i Use 'spec()' for the full column specifications.
```

```
#dataset %>% mutate_if(is.character,as.numeric)
options(warn=-1)
energy_data<- energy %>% mutate_if(is.character,as.numeric)
energy_data<-bind_cols(energy$`Percent Change 2020`,energy_data)
```

```
## New names:
## * NA -> ...1
```

```
energy_data<-bind_cols(energy$Gigajoule_per_capita,energy_data)
```

```
## New names:
## * NA -> ...1
## * ...1 -> ...2
```

```

energy_data<-rename(energy_data, "2020_percent_change"= 2)
energy_data<-rename(energy_data, "Country_Region"=1)
energy_data <- energy_data[ -c(3)]
#energy_data<-select(energy_data,-Gigajoule_per_capita)
energy_data <- energy_data[ -c(59:60) ]
glimpse(energy)

```

```

## Rows: 100
## Columns: 59
## $ Gigajoule_per_capita      <chr> "Canada", "Mexico", "US", "Total North Ameri~
## $ '1965'                   <chr> "250.9", "23.8", "259.5", "219.4", "53.3", "~
## $ '1966'                   <chr> "262.3", "24.4", "271.1", "228.5", "54.2", "~
## $ '1967'                   <chr> "269.6", "24.1", "277.7", "233.2", "55.0", "~
## $ '1968'                   <chr> "282.3", "25.5", "291.6", "244.0", "56.0", "~
## $ '1969'                   <chr> "294.2", "27.1", "303.9", "253.6", "57.0", "~
## $ '1970'                   <chr> "309.9", "28.1", "311.9", "260.0", "53.2", "~
## $ '1971'                   <chr> "314.5", "28.6", "315.6", "262.3", "55.6", "~
## $ '1972'                   <chr> "332.9", "30.7", "328.4", "272.6", "55.4", "~
## $ '1973'                   <chr> "356.0", "32.2", "339.2", "281.7", "56.9", "~
## $ '1974'                   <chr> "361.6", "34.4", "327.9", "273.3", "58.2", "~
## $ '1975'                   <chr> "351.6", "35.8", "316.3", "263.6", "55.9", "~
## $ '1976'                   <chr> "363.4", "37.1", "330.6", "274.3", "57.8", "~
## $ '1977'                   <chr> "379.9", "38.1", "336.1", "278.9", "59.4", "~
## $ '1978'                   <chr> "382.5", "41.8", "340.2", "282.0", "58.6", "~
## $ '1979'                   <chr> "389.5", "45.1", "340.6", "282.8", "61.9", "~
## $ '1980'                   <chr> "392.8", "48.0", "324.4", "271.3", "60.5", "~
## $ '1981'                   <chr> "386.0", "51.7", "311.8", "262.0", "57.8", "~
## $ '1982'                   <chr> "370.6", "52.7", "296.1", "249.3", "57.9", "~
## $ '1983'                   <chr> "364.5", "50.9", "292.3", "245.2", "59.4", "~
## $ '1984'                   <chr> "383.5", "52.1", "304.4", "254.9", "58.5", "~
## $ '1985'                   <chr> "391.3", "53.5", "300.8", "252.7", "56.9", "~
## $ '1986'                   <chr> "389.9", "50.9", "300.7", "251.5", "59.2", "~
## $ '1987'                   <chr> "397.1", "51.5", "308.0", "256.8", "60.9", "~
## $ '1988'                   <chr> "407.7", "50.9", "319.1", "264.7", "61.0", "~
## $ '1989'                   <chr> "406.5", "52.7", "323.6", "267.8", "58.0", "~
## $ '1990'                   <chr> "385.6", "54.4", "321.3", "264.5", "56.3", "~
## $ '1991'                   <chr> "382.0", "55.4", "317.7", "261.7", "56.5", "~
## $ '1992'                   <chr> "387.6", "55.4", "319.4", "262.9", "59.8", "~
## $ '1993'                   <chr> "392.8", "54.7", "322.9", "265.1", "60.7", "~
## $ '1994'                   <chr> "403.4", "56.9", "325.5", "267.9", "63.3", "~
## $ '1995'                   <chr> "409.0", "55.1", "328.8", "269.9", "65.0", "~
## $ '1996'                   <chr> "415.7", "55.9", "335.9", "275.2", "67.7", "~
## $ '1997'                   <chr> "414.0", "56.0", "334.2", "273.7", "68.1", "~
## $ '1998'                   <chr> "404.6", "57.6", "332.2", "271.8", "69.9", "~
## $ '1999'                   <chr> "413.5", "57.4", "334.0", "273.6", "69.8", "~
## $ '2000'                   <chr> "420.8", "59.3", "337.7", "276.9", "69.2", "~
## $ '2001'                   <chr> "408.5", "57.8", "326.4", "267.7", "66.7", "~
## $ '2002'                   <chr> "418.3", "58.8", "328.0", "269.6", "63.4", "~
## $ '2003'                   <chr> "416.5", "58.5", "326.1", "267.8", "67.0", "~
## $ '2004'                   <chr> "418.5", "60.8", "329.7", "270.7", "70.3", "~
## $ '2005'                   <chr> "417.3", "63.8", "326.9", "269.2", "73.9", "~
## $ '2006'                   <chr> "410.0", "64.8", "321.2", "264.8", "76.9", "~
## $ '2007'                   <chr> "419.8", "64.2", "322.6", "266.1", "79.7", "~

```

```
## $ '2008' <chr> "410.6", "64.6", "311.6", "257.8", "79.8", "~
## $ '2009' <chr> "389.5", "63.1", "293.4", "243.4", "76.8", "~
## $ '2010' <chr> "388.8", "64.1", "300.7", "248.2", "79.1", "~
## $ '2011' <chr> "398.9", "66.2", "295.4", "245.7", "80.7", "~
## $ '2012' <chr> "395.5", "65.7", "285.4", "238.4", "82.9", "~
## $ '2013' <chr> "400.7", "65.1", "290.9", "242.1", "85.2", "~
## $ '2014' <chr> "398.0", "64.0", "291.8", "242.0", "84.2", "~
## $ '2015' <chr> "395.9", "63.1", "287.0", "238.2", "84.9", "~
## $ '2016' <chr> "387.7", "63.1", "284.7", "235.9", "83.5", "~
## $ '2017' <chr> "387.8", "63.3", "283.8", "235.1", "82.9", "~
## $ '2018' <chr> "389.4", "62.1", "292.4", "240.5", "80.9", "~
## $ '2019' <chr> "386.3", "59.2", "288.4", "236.6", "75.5", "~
## $ '2020' <chr> "361.1", "50.2", "265.2", "216.8", "69.7", "~
## $ 'Percent Change 2020' <chr> "-6.8%", "-15.4%", "-8.3%", "-8.6%", "-8.0%", "~
## $ 'Percentchange_2009 to19' <chr> "-0.1%", "-0.6%", "-0.2%", "-0.3%", "-0.2%", "~
```

```
glimpse(energy_data)
```

```
## Rows: 100
## Columns: 58
## $ Country_Region <chr> "Canada", "Mexico", "US", "Total North America",~
## $ '2020_percent_change' <chr> "-6.8%", "-15.4%", "-8.3%", "-8.6%", "-8.0%", "--
## $ '1965' <dbl> 250.9, 23.8, 259.5, 219.4, 53.3, 11.6, 29.0, 17.~
## $ '1966' <dbl> 262.3, 24.4, 271.1, 228.5, 54.2, 12.3, 30.5, 18.~
## $ '1967' <dbl> 269.6, 24.1, 277.7, 233.2, 55.0, 12.4, 30.6, 18.~
## $ '1968' <dbl> 282.3, 25.5, 291.6, 244.0, 56.0, 13.9, 30.0, 19.~
## $ '1969' <dbl> 294.2, 27.1, 303.9, 253.6, 57.0, 14.7, 31.8, 18.~
## $ '1970' <dbl> 309.9, 28.1, 311.9, 260.0, 53.2, 16.7, 33.0, 20.~
## $ '1971' <dbl> 314.5, 28.6, 315.6, 262.3, 55.6, 17.8, 35.1, 19.~
## $ '1972' <dbl> 332.9, 30.7, 328.4, 272.6, 55.4, 20.0, 35.6, 21.~
## $ '1973' <dbl> 356.0, 32.2, 339.2, 281.7, 56.9, 23.3, 33.7, 20.~
## $ '1974' <dbl> 361.6, 34.4, 327.9, 273.3, 58.2, 24.9, 32.8, 22.~
## $ '1975' <dbl> 351.6, 35.8, 316.3, 263.6, 55.9, 25.6, 29.1, 22.~
## $ '1976' <dbl> 363.4, 37.1, 330.6, 274.3, 57.8, 27.5, 30.2, 23.~
## $ '1977' <dbl> 379.9, 38.1, 336.1, 278.9, 59.4, 28.7, 30.8, 23.~
## $ '1978' <dbl> 382.5, 41.8, 340.2, 282.0, 58.6, 30.9, 32.1, 24.~
## $ '1979' <dbl> 389.5, 45.1, 340.6, 282.8, 61.9, 32.7, 32.7, 25.~
## $ '1980' <dbl> 392.8, 48.0, 324.4, 271.3, 60.5, 32.5, 33.3, 24.~
## $ '1981' <dbl> 386.0, 51.7, 311.8, 262.0, 57.8, 31.0, 33.7, 24.~
## $ '1982' <dbl> 370.6, 52.7, 296.1, 249.3, 57.9, 31.8, 30.4, 25.~
## $ '1983' <dbl> 364.5, 50.9, 292.3, 245.2, 59.4, 31.7, 30.8, 26.~
## $ '1984' <dbl> 383.5, 52.1, 304.4, 254.9, 58.5, 33.2, 31.8, 26.~
## $ '1985' <dbl> 391.3, 53.5, 300.8, 252.7, 56.9, 34.6, 31.7, 26.~
## $ '1986' <dbl> 389.9, 50.9, 300.7, 251.5, 59.2, 36.2, 32.7, 27.~
## $ '1987' <dbl> 397.1, 51.5, 308.0, 256.8, 60.9, 36.5, 33.7, 28.~
## $ '1988' <dbl> 407.7, 50.9, 319.1, 264.7, 61.0, 37.2, 37.0, 28.~
## $ '1989' <dbl> 406.5, 52.7, 323.6, 267.8, 58.0, 37.2, 40.2, 29.~
## $ '1990' <dbl> 385.6, 54.4, 321.3, 264.5, 56.3, 36.0, 41.7, 26.~
## $ '1991' <dbl> 382.0, 55.4, 317.7, 261.7, 56.5, 36.7, 45.9, 27.~
## $ '1992' <dbl> 387.6, 55.4, 319.4, 262.9, 59.8, 37.8, 50.0, 28.~
## $ '1993' <dbl> 392.8, 54.7, 322.9, 265.1, 60.7, 38.8, 52.0, 30.~
## $ '1994' <dbl> 403.4, 56.9, 325.5, 267.9, 63.3, 40.1, 53.8, 31.~
## $ '1995' <dbl> 409.0, 55.1, 328.8, 269.9, 65.0, 41.4, 57.1, 31.~
## $ '1996' <dbl> 415.7, 55.9, 335.9, 275.2, 67.7, 43.1, 59.7, 32.~
```

```
## $ '1997' <dbl> 414.0, 56.0, 334.2, 273.7, 68.1, 44.8, 67.4, 32.~
## $ '1998' <dbl> 404.6, 57.6, 332.2, 271.8, 69.9, 45.8, 66.9, 32.~
## $ '1999' <dbl> 413.5, 57.4, 334.0, 273.6, 69.8, 45.9, 68.9, 29.~
## $ '2000' <dbl> 420.8, 59.3, 337.7, 276.9, 69.2, 46.0, 72.0, 28.~
## $ '2001' <dbl> 408.5, 57.8, 326.4, 267.7, 66.7, 44.4, 71.6, 27.~
## $ '2002' <dbl> 418.3, 58.8, 328.0, 269.6, 63.4, 44.9, 73.3, 26.~
## $ '2003' <dbl> 416.5, 58.5, 326.1, 267.8, 67.0, 45.1, 72.2, 27.~
## $ '2004' <dbl> 418.5, 60.8, 329.7, 270.7, 70.3, 46.7, 75.1, 27.~
## $ '2005' <dbl> 417.3, 63.8, 326.9, 269.2, 73.9, 47.5, 77.2, 26.~
## $ '2006' <dbl> 410.0, 64.8, 321.2, 264.8, 76.9, 48.2, 81.6, 29.~
## $ '2007' <dbl> 419.8, 64.2, 322.6, 266.1, 79.7, 50.6, 82.4, 29.~
## $ '2008' <dbl> 410.6, 64.6, 311.6, 257.8, 79.8, 52.3, 80.8, 31.~
## $ '2009' <dbl> 389.5, 63.1, 293.4, 243.4, 76.8, 51.3, 78.3, 30.~
## $ '2010' <dbl> 388.8, 64.1, 300.7, 248.2, 79.1, 56.0, 77.9, 31.~
## $ '2011' <dbl> 398.9, 66.2, 295.4, 245.7, 80.7, 58.0, 83.8, 32.~
## $ '2012' <dbl> 395.5, 65.7, 285.4, 238.4, 82.9, 58.5, 84.9, 34.~
## $ '2013' <dbl> 400.7, 65.1, 290.9, 242.1, 85.2, 60.2, 84.9, 34.~
## $ '2014' <dbl> 398.0, 64.0, 291.8, 242.0, 84.2, 61.1, 82.4, 35.~
## $ '2015' <dbl> 395.9, 63.1, 287.0, 238.2, 84.9, 59.7, 83.2, 35.~
## $ '2016' <dbl> 387.7, 63.1, 284.7, 235.9, 83.5, 57.7, 86.2, 37.~
## $ '2017' <dbl> 387.8, 63.3, 283.8, 235.1, 82.9, 57.9, 85.5, 37.~
## $ '2018' <dbl> 389.4, 62.1, 292.4, 240.5, 80.9, 57.8, 87.8, 37.~
## $ '2019' <dbl> 386.3, 59.2, 288.4, 236.6, 75.5, 58.9, 88.9, 37.~
## $ '2020' <dbl> 361.1, 50.2, 265.2, 216.8, 69.7, 56.5, 84.1, 34.~
```

```
view(energy_data)
view(energy)
```

Gather data set into a tidy database with country-region, year and Giga Joules per capita. (retained the 2020 % change as not clear we can derive it)

```
energy_tidy<-energy_data
energy_tidy<-gather(energy_tidy,year,GJoules_per_capita,3:58)
energy_tidy<-as.data.frame(energy_tidy)
glimpse(energy_tidy)
```

```
## Rows: 5,600
## Columns: 4
## $ Country_Region <chr> "Canada", "Mexico", "US", "Total North America",~
## $ '2020_percent_change' <chr> "-6.8%", "-15.4%", "-8.3%", "-8.6%", "-8.0%", "--
## $ year <chr> "1965", "1965", "1965", "1965", "1965", "1965", ~
## $ GJoules_per_capita <dbl> 250.9, 23.8, 259.5, 219.4, 53.3, 11.6, 29.0, 17.~
```

```
#energy_tidy<-as.character.factor(energy_tidy$Country_Region)
#energy_world <- spread(energy_tidy,GJ_per_capita,County_Region)
#select(energy_tidy, ("Total World","Total North America"))
#view(energy_world)
#iris %>% select(starts_with("Sepal"))
```

```
energy_world<-filter(energy_data,Country_Region=="Total World")
energy_NA<-filter(energy_data,Country_Region=="Total North America")
energy_Asia<-filter(energy_data,Country_Region=="Total Asia Pacific")
energy_LA<-filter(energy_data,Country_Region=="Total S. & Cent. America")
```

```

energy_ME<-filter(energy_data,Country_Region=="Total Middle East")
energy_Africa<-filter(energy_data,Country_Region=="Total Africa")
energy_Europe<-filter(energy_data,Country_Region=="Total Europe")
energy_CIS<-filter(energy_data,Country_Region=="Total CIS")
Regions<-bind_rows(energy_world,energy_Africa,energy_Asia,energy_CIS,energy_Europe,energy_LA,energy_ME)
Regions <- Regions[ -c(2)]
Regions<-gather(Regions,year,GJoules_per_capita,2:57)
Regions<-as.data.frame(Regions)

```

In general the 25% of the energy demand is supplied by oil/ refineries. One Gigajoule is equivalent to 0.16 BBL of Oil. A refinery operates to produce approx 65-70% of the production for transportation fuels and the fuel for ICE (internal combustion engines - gasoline) is approx 50% of the refinery production.

Use the 70% as an estimate of the total refinery production converted for transport fuels in gallons. (ignore the impact of volume swell due to density drop) to estimate gallons per capita. Also 42 gal per BBL is nominal for oil.

Assumption for this is production is in the given region. This is not 100% true - but currently between regions within approx 10% or so of each regions own production (ie neglect trying to figure out export mix between regions)

```

Regions<-mutate(Regions,BBL_per_Capita_Refinery=GJoules_per_capita*0.16*0.25)
Regions<-mutate(Regions, Gal_Fuel_per_capita= BBL_per_Capita_Refinery*42*0.7)

```

Including Plots

Plot the Giga Joules per capita and the refinery BBL per capita and the Gals fuel for transportation per capita. All plots are similar in this simplified example and the next step to highlight the more dramatic changes are to convert with the population increase.

The general trends are still highlighted: North America with the highest per capita consumption of energy. The 1980 drop developing as reformulated fuel standards developed, elimination of lead occurred and the goal to increase fuel economy standards became important following our oil shortage. [The drop in the CIS is likely related to improved fuel standards as communism declined in the early 1990's - and increased market acceptance and use of Western improvements to fuel standards occurred for transport]

Asia and Middle East shows the dramatic and continuing increase in the per capita consumption of energy and transport fuel as a rising middle class uses more transportation. Asia has the majority of the population and the largest total fuel and oil demand globally today in spite of the North America having a large per capita demand.

Europe has maintained better public transport and alternative transport to cars - and maintained a relatively high tax on fuel / cars holding demand per capita down.

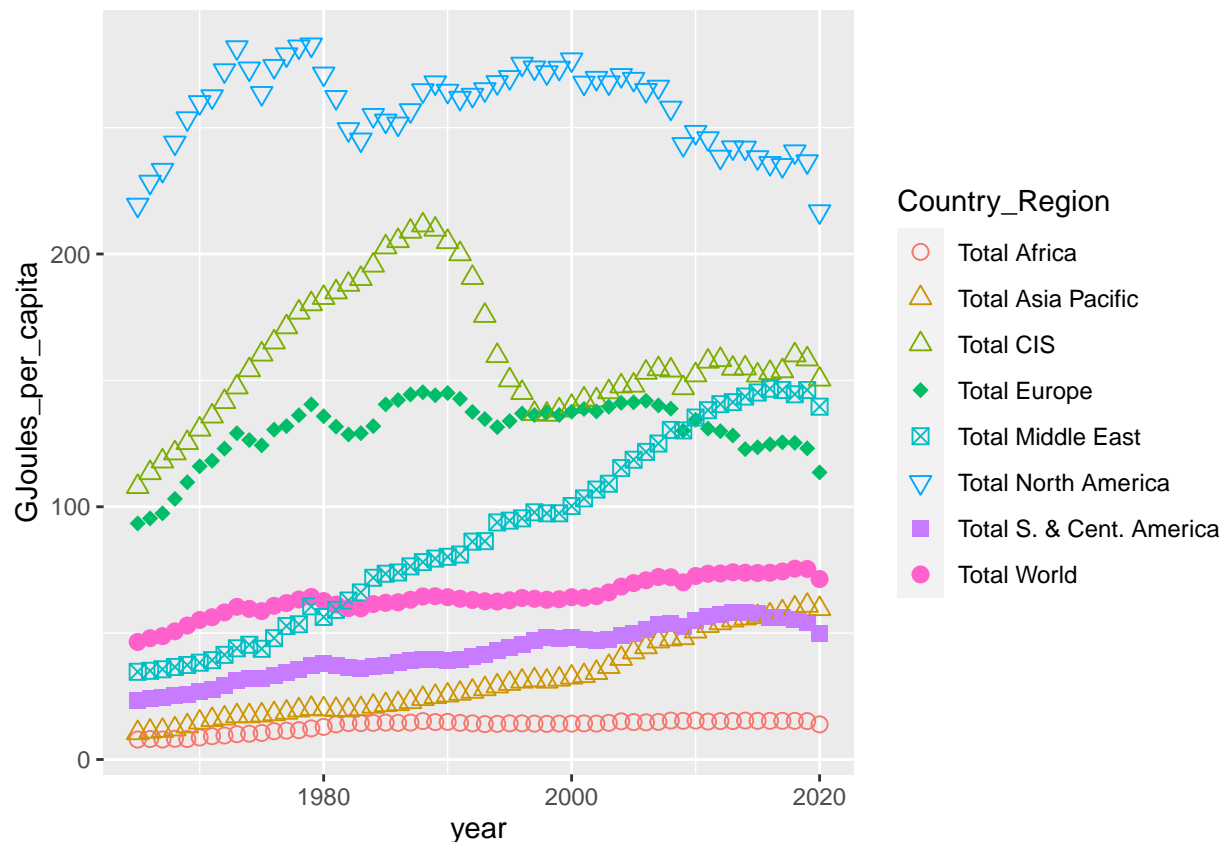
Africa remains the next big region to experience a larger middle income growth and would be expected to show an increase in transport fuel demand / and general energy demand. [It has remained very low and it shows in this analysis how depressed the growth has been in the region]

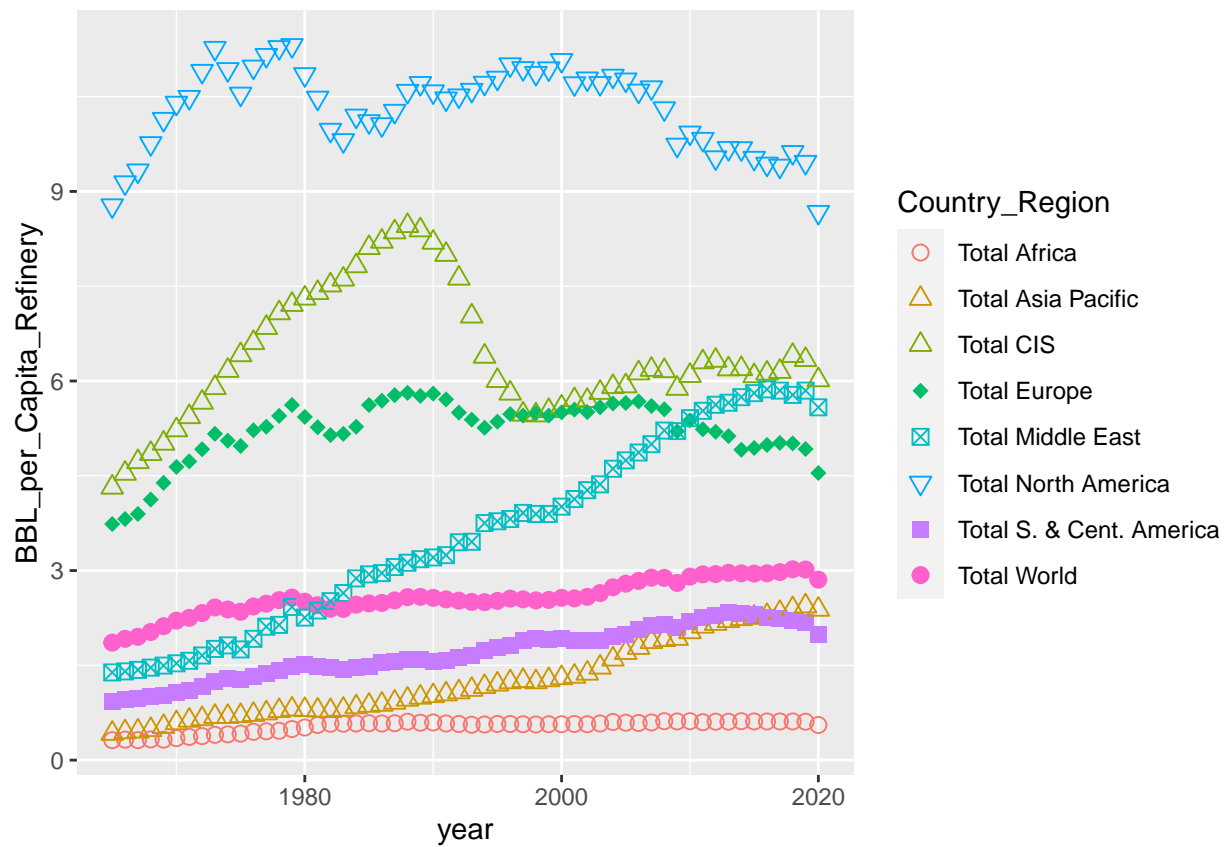
My final question concerning is this sufficient data to decide how all of these changes are impacting greenhouse gas emissions is more of a look at the data graphs vs a quick internet search on what has changed.

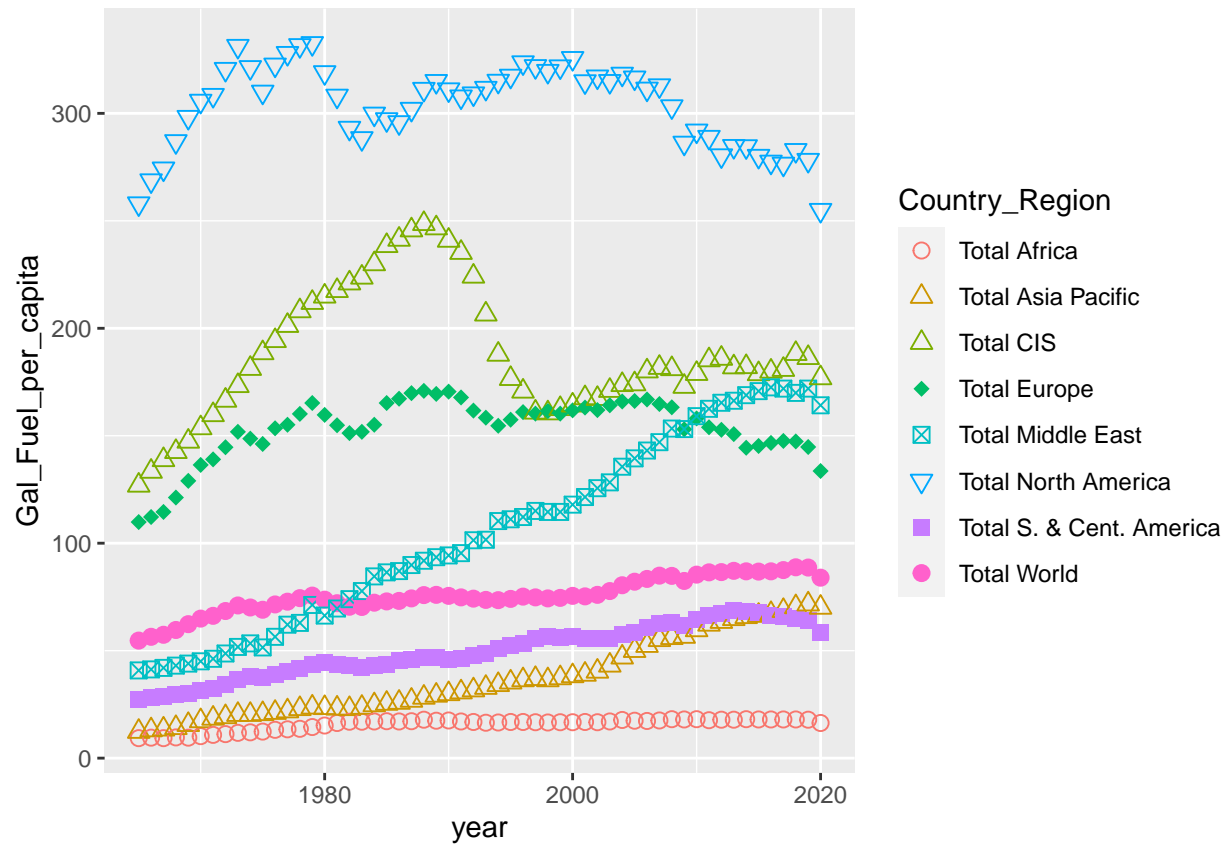
The demand growth for energy in the transport segment is approx 55 gal per capita to 84 gal per capita from 1965 to 2020. Is this sufficient to indicate how the use of transport fuels are impacting the environment? and are changing Green House Gas emissions due to fuel?

- Actually, I do not think this is nearly sufficient data to make a claim that it is worse than it was in 1965. Here is why:

- 1) Since 1965 we have significantly changed the fuel we use.
 - In 1975 we began to eliminate lead from the fuel banned by 1996 in US (globally followed)
 - Sulfur in fuel was reduced globally and very dramatically reduces GHG emissions
 - Refining NOX, SOX controls required
- 2) Introduction of the catalytic converter in 1975 with relatively full use by 1985.
 - Dramtically reduces hydrocarbon emissions and NOx emissions - but does not change CO2 emissions.
- 3) Reformulated fuel standards in 2005 in US required improvements and the use of cleaner biofuel (Ethanol) Globally this has followed and reduces the emissions and is in the long cycle more sustainable than oil. (even though a small %)
- 4) We have additionally improved our MPG (car miles per gallon) from 13.5 MPG in 1965 to approx 22-23 MPG in 2020. While not changing the direct consumption we see - it does dramatically reduce the demand for fuel. The behavior pattern to improve this is the most significant near term factor to reduce GHG in next 5-10 yr.
- 5) It was a fun simple data set to try and see if we can identify any overall trends. More work can be done on this data set.
- 6) Data set is from the IEA (International Energy Agency based in France. Data is free on website in exchange for an email/name.)







Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.