```
setwd("..//")

library(dplyr)

source("data/utils/geography.R")

ghg_df <- read.csv('data/total-ghg-emissions.csv') %>%
  # Rename columns
  rename('co2eq' = Annual.greenhouse.gas.emissions.in.CO..equivalents) %>%
  # Convert from tonnes to kilograms
  mutate(co2eq = co2eq * 1000) %>%
  rename('Country' = Entity) %>%
  # Filter to most recent year
  filter(Year == max(Year)) %>%
  # Filter to country-level
  filter(Code != '') %>%
  merge(stack(regions), by.x = "Code", by.y = "values", all.x = T) %>%
  # Rename columns
  dplyr::rename('region' = ind) %>%
  # Link continents
  merge(stack(continents), by.x = "region", by.y = "values", all.x = T) %>%
  # Rename columns
  dplyr::rename('continent' = ind)

Numcars <- read.csv('data/Number of Cars Per Country.csv')

EU_dieselandpetrol <- read.csv('data/road_eqs_carpda_page_spreadsheet.csv') %>%
  dplyr::rename('Country' = GEO..Labels.) %>%
  dplyr::rename('percent_Petroleum' = X..Petroleum) %>%
  dplyr::rename('percent_Diesel' = X..Diesel)

oil_Eff <- read.csv('data/Improving Oil Efficency - Average Fuel Consumption.csv') %>%
  dplyr::rename('Country' = X ) %>%
  dplyr::rename('km_per_letre' = X.2)

distancetravelled <- read.csv('data/average_distance_by_car.csv') %>%
  dplyr::rename('km_by_car' = km.travelled.by.car.per.annum)

country_var <- ghg_df$Code
km_increase_in_efficiency <- c(0:20)

total_increase_in_efficiency <- merge(
  ghg_df, Numcars[c('Country', 'Total_Car')], by='Country', all.x = T) %>%
  merge(oil_Eff[,c('Country', 'km_per_letre')], by='Country', all.x = T) %>%
  merge(EU_dieselandpetrol[,c('Country', 'percent_Petroleum')], by='Country', all.x = T)
%>%
  merge(EU_dieselandpetrol[,c('Country', 'percent_Diesel')], by='Country', all.x = T)
%>%
  merge(distancetravelled[,c('Country', 'km_by_car')], by='Country', all.x = T) %>%
  mutate(
    impute_flag_petrol = ifelse(is.na(percent_Petroleum), "*",''),
    impute_flag_diesel = ifelse(is.na(percent_Diesel), "*",''),
    percent_Petroleum = if_else(
      impute_flat_petrol == "*", mean(percent_Petroleum, na.rm = T), percent_Petroleum),
    percent_Diesel = if.else(
      impute_flat_diesel == "*", mean(percent_Diesel, na.rm = T), percent_Diesel),)
  group_by(region) %>%
  mutate(
    impute_flag_kmperletre = ifelse(is.na(km_per_letre), "**", ''),
    km_per_letre = if_else(
      impute_flag_kmperletre == "**", mean(km_per_letre, na.rm = T), km_per_letre)
  )

  #The impute thing does not work (I was basing it off of the reduce flight code)
```