BDA - Assignment 9

Tomi Räsänen - student_no & Erik Husgafvel - 528867

Contents

1. Introduction

```r
library(rstan)
library(ggplot2)
library(reshape2)
library(gridExtra)
#library(aaltobda)
#data("factory")
```

```r
# Read data to data frame
data <- read.csv("../data_co2.csv")
# Our exploratory data analysis discovered some missing values. In order to
# avoid errors, we'll replace them with zeros.
data[is.na(data)] <- 0


# We discovered that the CO2-emissions difference between our selected countries is so vast
# that it's better to split the data into two different plots.

df_data1 <- data[, (data[dim(data)[1], ]) >= 100]
# Sorting the columns in ascending order based on the last row value
df_data1 <- df_data1[,order(df_data1[69,])]
df_data1_2 <- data.frame(years=seq(1950,2018), df_data1)
# Plot the result
df_plot1 <- melt(data = df_data1_2, id.vars = "years", variable.name = "country")

df_data2 <- data[, (data[dim(data)[1], ]) < 100]
# Sorting the columns in ascending order based on the last row value
df_data2 <- df_data2[,order(df_data2[69,])]
df_data2_2 <- data.frame(years=seq(1950,2018), df_data2)
# Plot the result
df_plot2 <- melt(data = df_data2_2, id.vars = "years", variable.name = "country")


# Plot the countries
plot1 <- ggplot(df_plot1, aes(x=years, y=value, colour=country), environment = ) +
        geom_line() +
        ggtitle("Selected countries' co2-emissions: \nbig emitters") +
        xlab("Year") +
        ylab("CO2-emission")

plot2 <- ggplot(df_plot2, aes(x=years, y=value, colour=country)) +
        geom_line() +
        ggtitle("Selected countries' co2-emissions: \nsmall emitters") +
        xlab("Year") +
        ylab("CO2-emission")

grid.arrange(plot1, plot2, nrow=2)
```
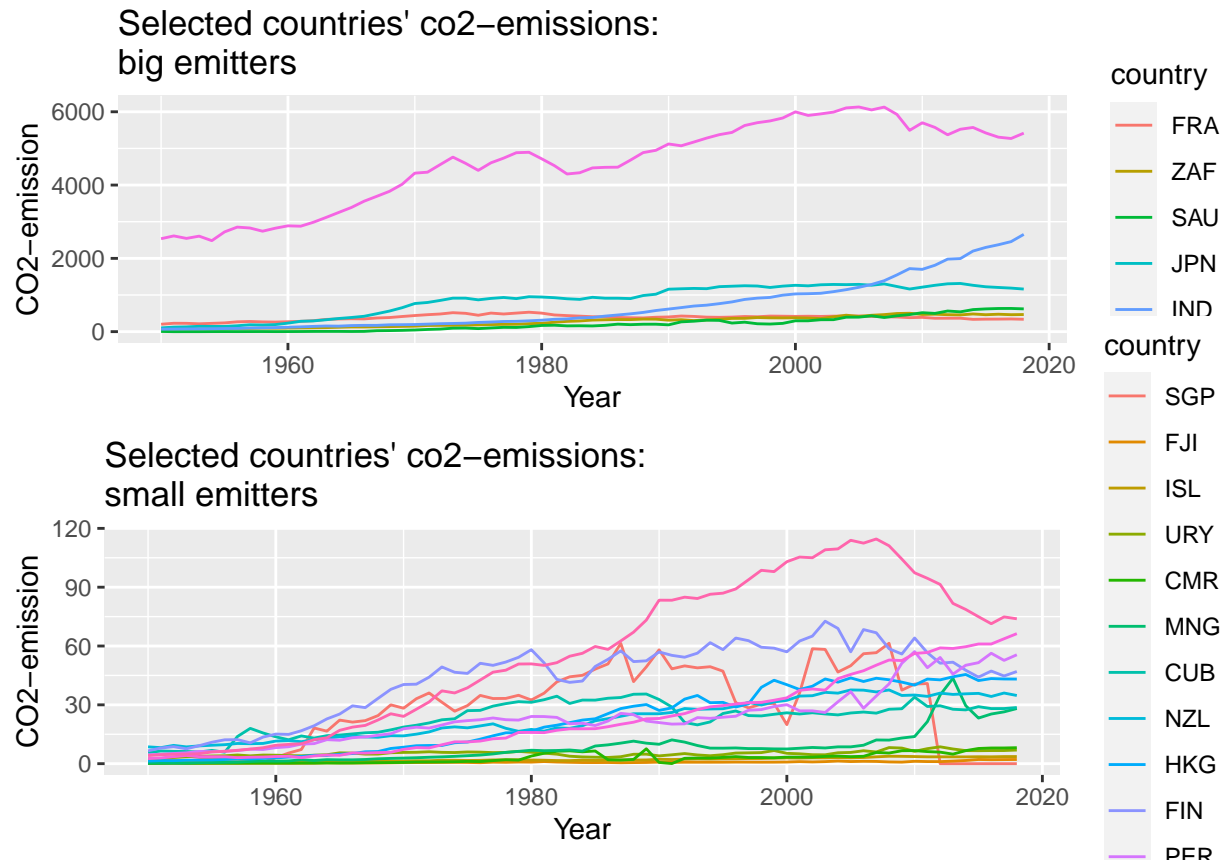
Selected countries' co2-emissions: big emitters

Selected countries' co2-emissions: small emitters

Example code for hierarchical model from assignment

```
# # Setting seed to get same "random" results
# SEED <- 12345
#
# # Printing out our hierarchical model
# writeLines(readLines("assignment9_hierarchical_model.stan"))
#
#
# num_of_chains = 10
# hier_model <- rstan::stan_model(file = "assignment9_hierarchical_model.stan")
# hier_data <- list(N = nrow(factory),
#                   J = ncol(factory),
#                   y = factory)
# hier_fit <- rstan::sampling(object = hier_model,
#                             data = hier_data,
#                             iter = 5000,
#                             warmup = 1000,
#                             chains = num_of_chains,
#                             verbose = FALSE,
#                             refresh = 0,
#                             show_messages = FALSE,
#                             seed = SEED)
```