World Food Production

Vinodh Mallikarjun Chincholi | MDS202252

Abstract

In this project I attempt to understand the order of magnitude in which we grow food, how it has changed over the years as well as its variation with respect to geographical location over the world.

1 Introduction

Our world population is expected to grow from 7.3 billion today to 9.7 billion in the year 2050. Finding solutions for feeding the growing world population has become a hot topic for food and agriculture organizations, entrepreneurs and philanthropists. These solutions range from changing the way we grow our food to changing the way we eat. To make things harder, the world's climate is changing and it is both affecting and affected by the way we grow our food – agriculture.

Through this visualization I hope to uncover the magnitude of food's importance, the order in which we grow food and how it has changed over the years. The variation in time can be seen readily by plotting time series plots while the variation in geographical location has been presented as a choropleth map.

2 Dataset Description

The Food and Agriculture Organization of the United Nations provides free access to food and agriculture data for over 245 countries and territories, from the year 1961 to the most recent update (depends on the dataset). One dataset from the FAO's database is the Food Balance Sheets. It presents a comprehensive picture of the pattern of a country's food supply during a specified reference period, the last time an update was loaded to the FAO database was in 2013. The food balance sheet shows for each food item the sources of supply and its utilization.

The columns of interest in this project are:

- 'Area': which is the name of the country
- 'Item':
- 'Element':
  - Food: refers to the total amount of the food item available as human food during the reference period.
  - Feed: refers to the quantity of the food item available for feeding to the livestock and poultry during the reference period.

Due to `Feed` data missing for multiple countries, only `Food` data is used in graphs.

3 Visualising the data in R

3.1 Geographical variation in a year

Setting things up by including required packages:

```
library("ggplot2")
library("dplyr")
library("sf")
library("rnaturalearth")
library("rnaturalearthdata")

theme_set(theme_bw())
```

Importing the dataset and selecting relevant columns

```
fao <- read.csv('fao.csv')
fao <- fao %>% select(-c("Area.Abbreviation", "Area.Code",
                         "Item.Code", "Element.Code", "Unit"))
```

Setting up `world` dataset to plot choropleth.

```
world <- ne_countries(scale = "medium", returnclass = "sf")
world <- filter(world, name != 'Antarctica')
```

Cleaning dataset by changing country names to standard names.

```
my_fao <- fao

my_fao['Area'][my_fao['Area'] == 'United States of America'] <- 'United States'
my_fao['Area'][my_fao['Area'] == 'China, mainland'] <- 'China'
my_fao['Area'][my_fao['Area'] == 'Russian Federation'] <- 'Russia'
my_fao['Area'][my_fao['Area'] == 'Iran (Islamic Republic of)'] <- 'Iran'
my_fao['Area'][my_fao['Area'] == 'Sudan'] <- 'S. Sudan'
my_fao['Area'][my_fao['Area'] == "Democratic People's Republic of Korea"] <- 'Korea'
my_fao['Area'][my_fao['Area'] == "Bolivia (Plurinational State of)"] <- 'Bolivia'
my_fao['Area'][my_fao['Area'] == 'Venezuela (Bolivarian Republic of)'] <- 'Venezuela'
my_fao['Area'][my_fao['Area'] == "Lao People's Democratic Republic"] <- 'Lao PDR'
my_fao['Area'][my_fao['Area'] == "Central African Republic"] <- 'Central African Rep.'
my_fao['Area'][my_fao['Area'] == "Czechia"] <- "Czech Rep."
my_fao['Area'][my_fao['Area'] == "C<f4>te d'Ivoire"] <- "Côte d'Ivoire"
```

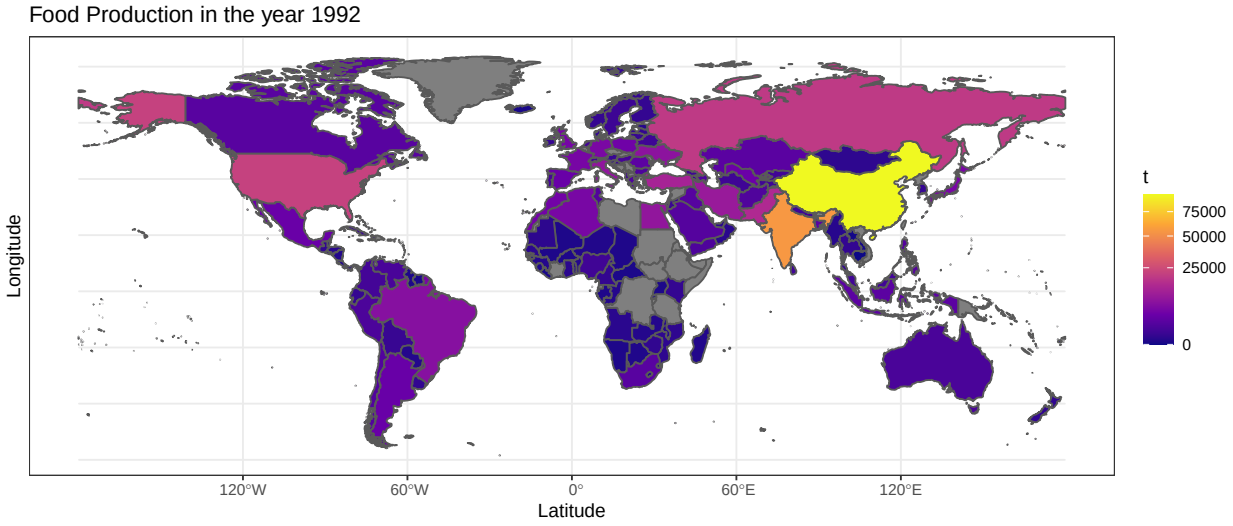
Finally, defining a function to plot the choropleth given a `year` between `1961` and `2013` and one of the many food item types `type`.

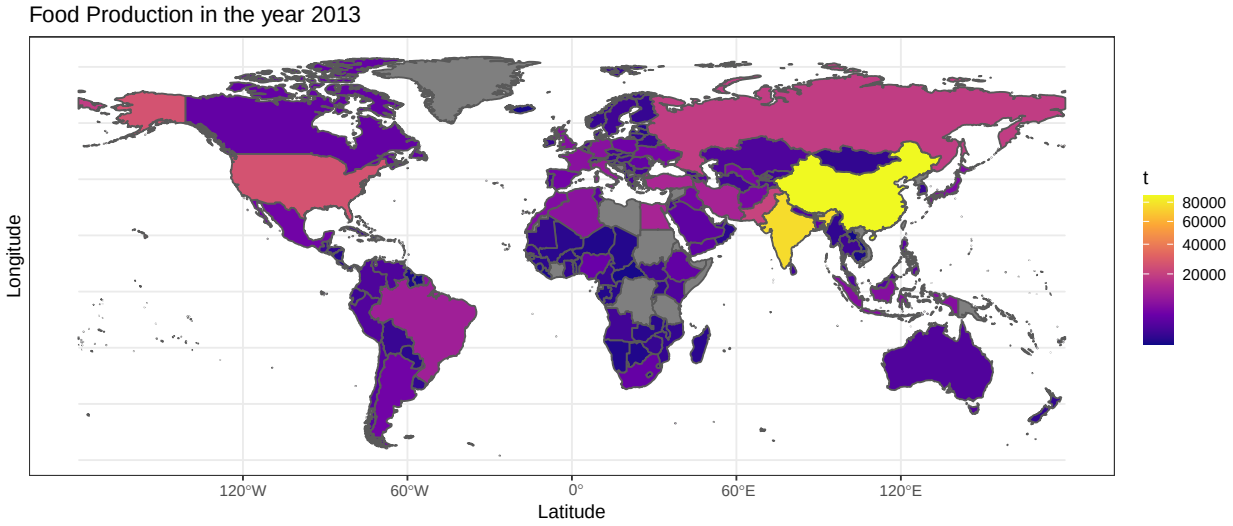
```
foodprod <- function(year = 1961, type = "Food") {

  my_fao <- my_fao %>%
    filter(Item == "Wheat and products", Element == type)

  df <- left_join(world, my_fao, by=c("name"="Area"))
  mydf <- select(df, t = paste("Y", year, sep=''))
```

```
  plot <- ggplot(data = mydf) +
    geom_sf(aes(fill = t)) +
    scale_fill_viridis_c(option = "plasma", trans = "sqrt") +
    xlab("Latitude") + ylab("Longitude") +
    ggtitle(paste("Food Production in the year", year))

  print(plot)

}
```

Plotting the choropleths,

```
foodprod(year = 1992)
foodprod(year = 2013)
```





3.2 Temporal variation at a location

Now that we have seen how the production rate is over the world in a given year, it will be interesting to see how the time series in India looks like.

Reload data.

```
fao2 <- read.csv('fao.csv')
```

Defining a function to plot wheat production as a function of country:

```
wheat <- function(country){

  df <- filter(fao2, Area == country)
  wheat <- filter(df, Item.Code == 2511)
  wheat <- data.frame(t(wheat))
  rownames(wheat) <- NULL


  wheat <- wheat[-c(1:10),]
  d <- seq(as.Date("1961-01-01"), by = 'year', length.out = 53)
  wheat <- cbind.data.frame(dat = d, wheat)

  plot <- ggplot() +
    geom_smooth(wheat, mapping = aes(x = dat, y = X1, group=1), se=F)+
    geom_point(wheat, mapping = aes(dat, X1)) +
    labs(x = "Year", y = "food production in tonnes of KG",
         title=paste("Graph of ", country, "wheat production over the years"))

  print(plot)
}
```

Finally plot production over the years.

```
wheat("India")
wheat("United States of America")
```