

# Case 1

Matheus Barreto Alves de Almeida

07/07/2020

Primeiramente, vamos iniciar a bibliotecas importantes

```
library(ggplot2)
library(dplyr)
library(readxl)
library(plotly)
library(purrr)
library(cluster)
library(forecast)
library(fpp2)
```

Carregando e visualizando o arquivo.

```
df <- read.csv('TFP.csv')
head(df)
```

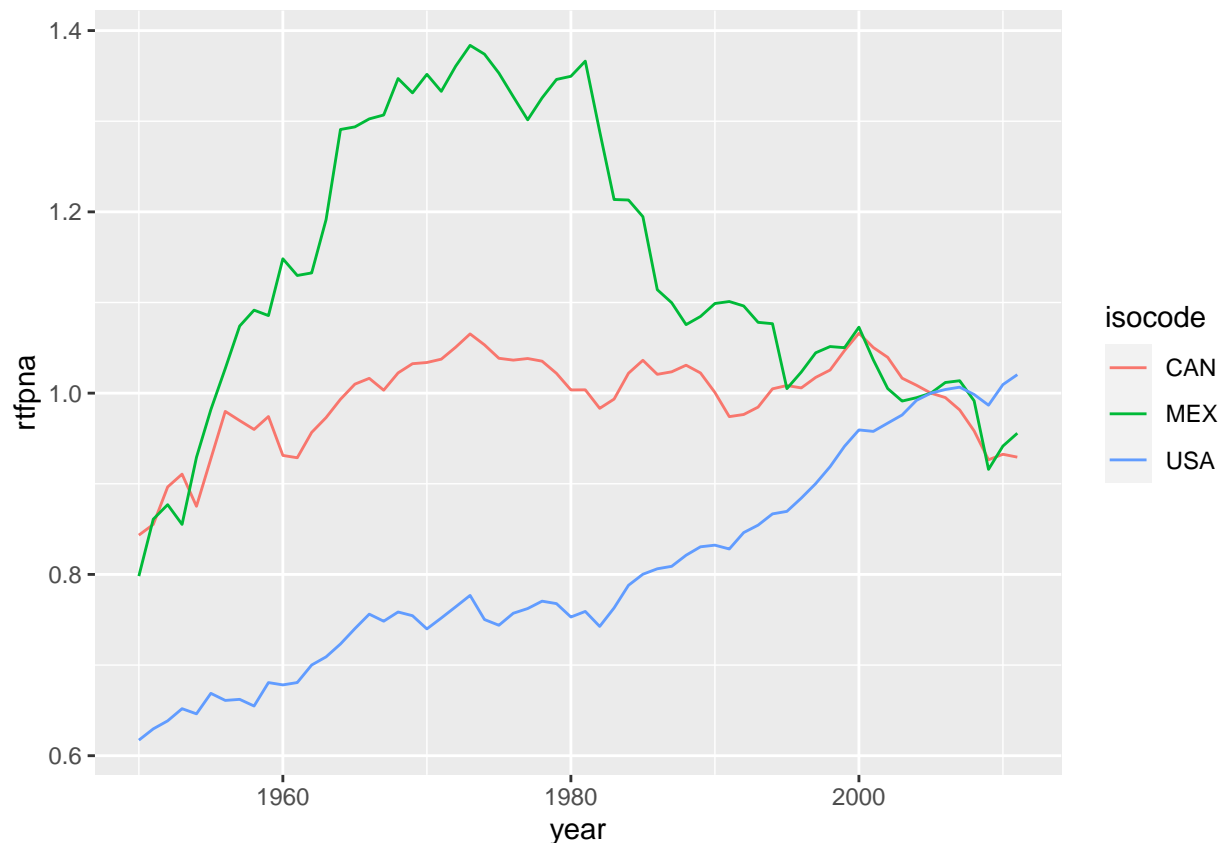
```
##   isocode year   rtfpna
## 1     USA 1950 0.6171479
## 2     USA 1951 0.6295884
## 3     USA 1952 0.6384513
## 4     USA 1953 0.6518582
## 5     USA 1954 0.6461794
## 6     USA 1955 0.6687729
```

Transformando a variável isocode em um fator

```
df$isocode <- factor(df$isocode)
```

Agora, vamos agrupar pela variável isocode e colocar em um gráfico a evolução temporal do TFP dos três países:

```
df %>% group_by(isocode) %>% ggplot(aes(year, rtfpna, color= isocode)) + geom_line()
```



A Principio, nota-se que o TFP dos Estados Unidos difere dos demais pois apresenta um padrão de crescimento ao longo dos anos. Tanto o México quanto o Canadá apresentam uma tendência de decrescimento do TPF, no entanto, o México apresenta uma maior volatilidade.

Uma previsão pode ser feita utilizando a biblioteca “forecast”. Para tanto, vamos separar nosso dataframe em subsets:

```
mex <- df %>% filter(isocode=='MEX') # %>% select(-isocode)
can <- df %>% filter(isocode=='CAN') # %>% select(-isocode)
usa <- df %>% filter(isocode=='USA') # %>% select(-isocode)
```

Para usar a biblioteca do forest, precisamos a principio deixar nossa variável de tempo no formado de time series:

```
x <- ts(mex[,3],start= 1950)
y <- ts(can[,3],start= 1950)
z <- ts(usa[,3],start= 1950)
```

Feito isso, podemos agora gerar um modelo para tentar prever os proximos 10 anos:

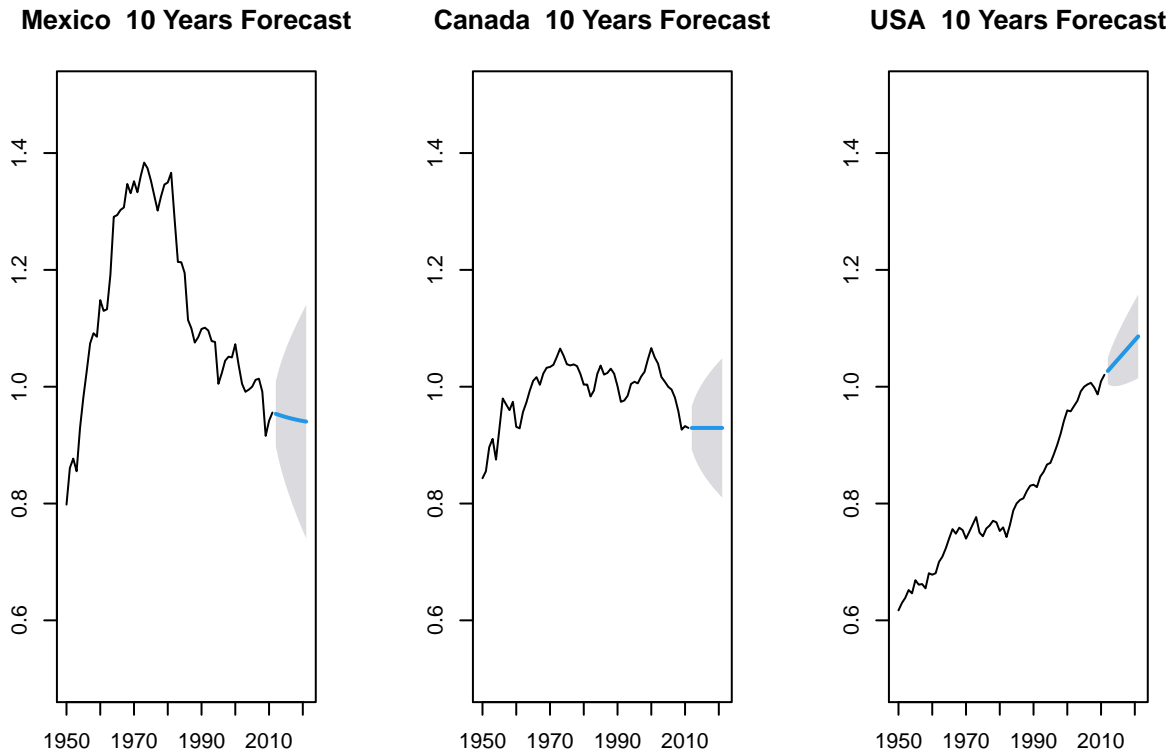
```
f.x <- forecast(x, level = c(95), h = 10)
f.y <- forecast(y, level = c(95), h = 10)
f.z <- forecast(z, level = c(95), h = 10)
```

Agora, vamos criar os plots para cada país:

```
mex.plot <- autoplot(f.x,alpha=0.5,xlab = "Year", ylab = "rtfpna",main="Mexico TFP Projection", colour = 
can.plot <- autoplot(f.y,alpha=0.5,xlab = "Year", ylab = "rtfpna",main="Canada TFP Projection", colour = 
usa.plot <- autoplot(f.z,alpha=0.5,xlab = "Year", ylab = "rtfpna",main="USA TFP Projection", colour = T)
```

Por fim, utilizando os gráficos básicos do R, podemos gerar as seguintes visualizações:

```
#Plotando os gráficos lado a lado para comparação  
par(mfrow= c(1,3))  
plot(f.x, main = 'Mexico 10 Years Forecast',ylim = c(0.5,1.5))  
plot(f.y, main = 'Canada 10 Years Forecast',ylim = c(0.5,1.5))  
plot(f.z, main = 'USA 10 Years Forecast',ylim = c(0.5,1.5))
```



Como esperado, a projeção para o México possui um spread maior, por conta da alta volatilidade de seu TFP, mantendo a tendência de diminuição.

O Canadá embora também possua uma alta volatilidade, seus extremos não são acentuados como no México. A tendência é de que seu TFP se mantenha.

Os Estados Unidos, como esperado, possui uma projeção de aumento progressivo.

O fator de produtividade tem correlação com muitas variáveis econômicas que podem ser vistas em <https://www.investopedia.com/terms/s/solow-residual.asp>.

Portanto, poderiam ter uma visão multivariável do problema utilizando “emp”, “avh”, “cda”, “ck”, seriam algumas das variáveis mais óbvias a serem analisadas a princípio.