Projeto Final - Covid

Pedro Lealdino Filho

24/05/2020

## Estudo sobre o Covid no Mundo

Neste trabalho realizaremos uma Análise Exploratória de Dados Simples em dados reais sobre a pandemia do novo coronavírus, Covid19.

### Carregando as Bibliotecas Necessárias

library(ggplot2)  
library(modeest)

### Importando os Dados

dados <- read.table("coronavirus\_situacao.csv", sep = ",", header = TRUE)  
  
dados <- dados[1:10, ]

### Visualizando a Estrutura dos Dados

str(dados)

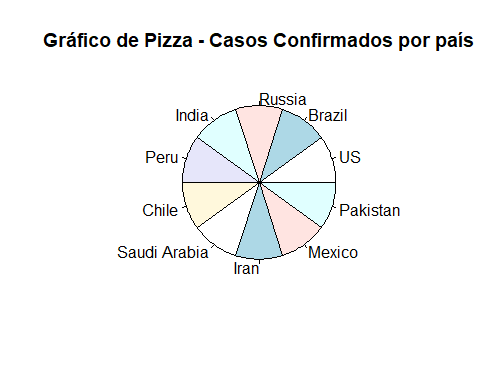
## 'data.frame': 10 obs. of 5 variables:  
## $ X : int 1 2 3 4 5 6 7 8 9 10  
## $ country : Factor w/ 188 levels "Afghanistan",..: 180 24 141 80 135 36 148 82 113 131  
## $ confirmed: int 23285 19694 8764 5553 4537 4038 2691 2346 2248 1932  
## $ death : int 1518 876 135 132 110 35 10 64 424 46  
## $ recovered: int 4920 9889 9262 3113 5444 997 1844 1635 1094 612

## Gráficos

dados

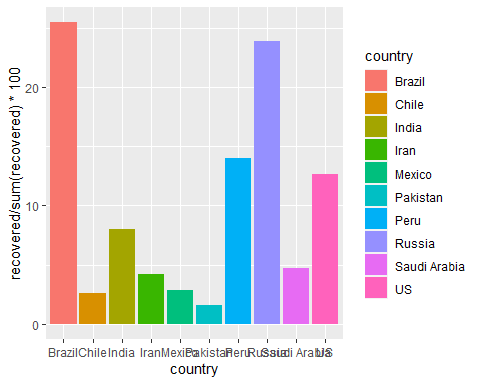
## X country confirmed death recovered  
## 1 1 US 23285 1518 4920  
## 2 2 Brazil 19694 876 9889  
## 3 3 Russia 8764 135 9262  
## 4 4 India 5553 132 3113  
## 5 5 Peru 4537 110 5444  
## 6 6 Chile 4038 35 997  
## 7 7 Saudi Arabia 2691 10 1844  
## 8 8 Iran 2346 64 1635  
## 9 9 Mexico 2248 424 1094  
## 10 10 Pakistan 1932 46 612

pie(table(dados$confirmed), labels = dados$country, main = "Gráfico de Pizza - Casos Confirmados por país")



Podemos perceber pelo gráfico acima que Estados Unidos, Rússia e Brasil lideram o ranking de casos confirmados no mundo. No próximo gráfico, visualizaremos a porcentagem de pessoas recuperadas por país.

ggplot(dados, aes(x = country, y = recovered/sum(recovered) \* 100, fill = country)) +   
 geom\_bar(stat = "identity")



## Medidas de Tendência Central

Vamos agora verificar a média de pessoas contaminadas, recuperadas e mortas dos países

### Média

mean(dados$confirmed)

## [1] 7508.8

mean(dados$death)

## [1] 335

mean(dados$recovered)

## [1] 3881

### Mediana

median(dados$confirmed)

## [1] 4287.5

median(dados$death)

## [1] 121

median(dados$recovered)

## [1] 2478.5

## Medidas de Variabilidade

### Variância

var(dados$confirmed)

## [1] 59143192

var(dados$death)

## [1] 242723.6

var(dados$recovered)

## [1] 11680957

### Desvio Padrão

sd(dados$confirmed)

## [1] 7690.461

sd(dados$death)

## [1] 492.6698

sd(dados$recovered)

## [1] 3417.741

## Resumo dos dados

summary(dados)

## X country confirmed death recovered   
## Min. : 1.00 Brazil :1 Min. : 1932 Min. : 10.0 Min. : 612   
## 1st Qu.: 3.25 Chile :1 1st Qu.: 2432 1st Qu.: 50.5 1st Qu.:1229   
## Median : 5.50 India :1 Median : 4288 Median : 121.0 Median :2478   
## Mean : 5.50 Iran :1 Mean : 7509 Mean : 335.0 Mean :3881   
## 3rd Qu.: 7.75 Mexico :1 3rd Qu.: 7961 3rd Qu.: 351.8 3rd Qu.:5313   
## Max. :10.00 Pakistan:1 Max. :23285 Max. :1518.0 Max. :9889   
## (Other) :4