



# Matemática computacional

Análise estatística dos dados da COVID-19

19 de Junho de 2020

São José dos Campos - SP

**Fernando Cossetin**

**Felipe Menino**

**Felipe Perin**

# Conjunto de dados



Niterói



Estado de Minas Gerais

# Conjunto de dados



Brasil



Canadá



Rússia

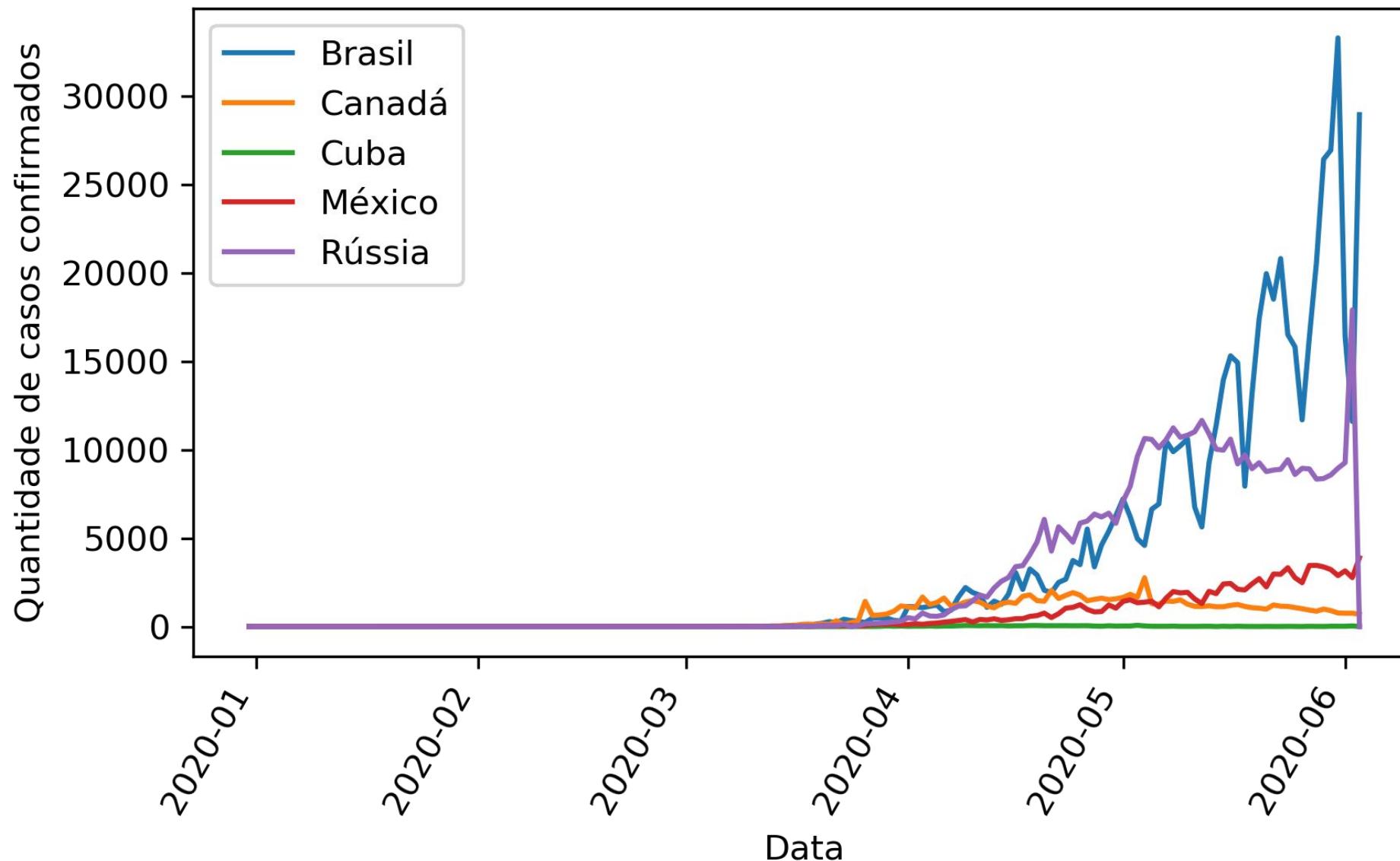


Cuba



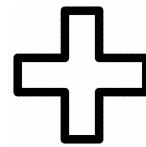
México

# Casos de COVID-19 confirmados diariamente



# Conjunto de dados

**Confiabilidade das fontes**



**Automatização do acesso**

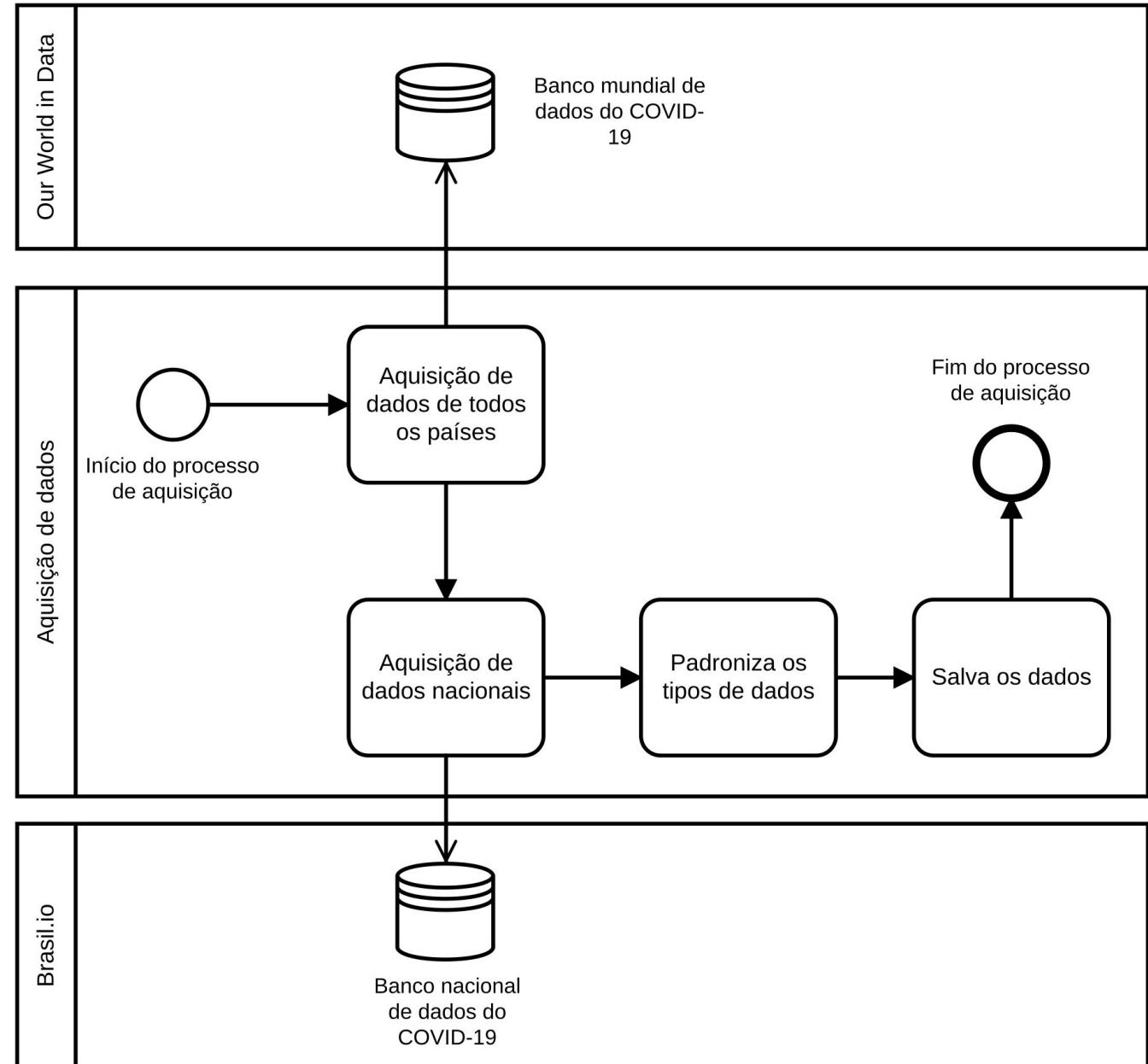
# Aquisição de dados

Dados Nacionais



Dados coletados até 03/06/2020

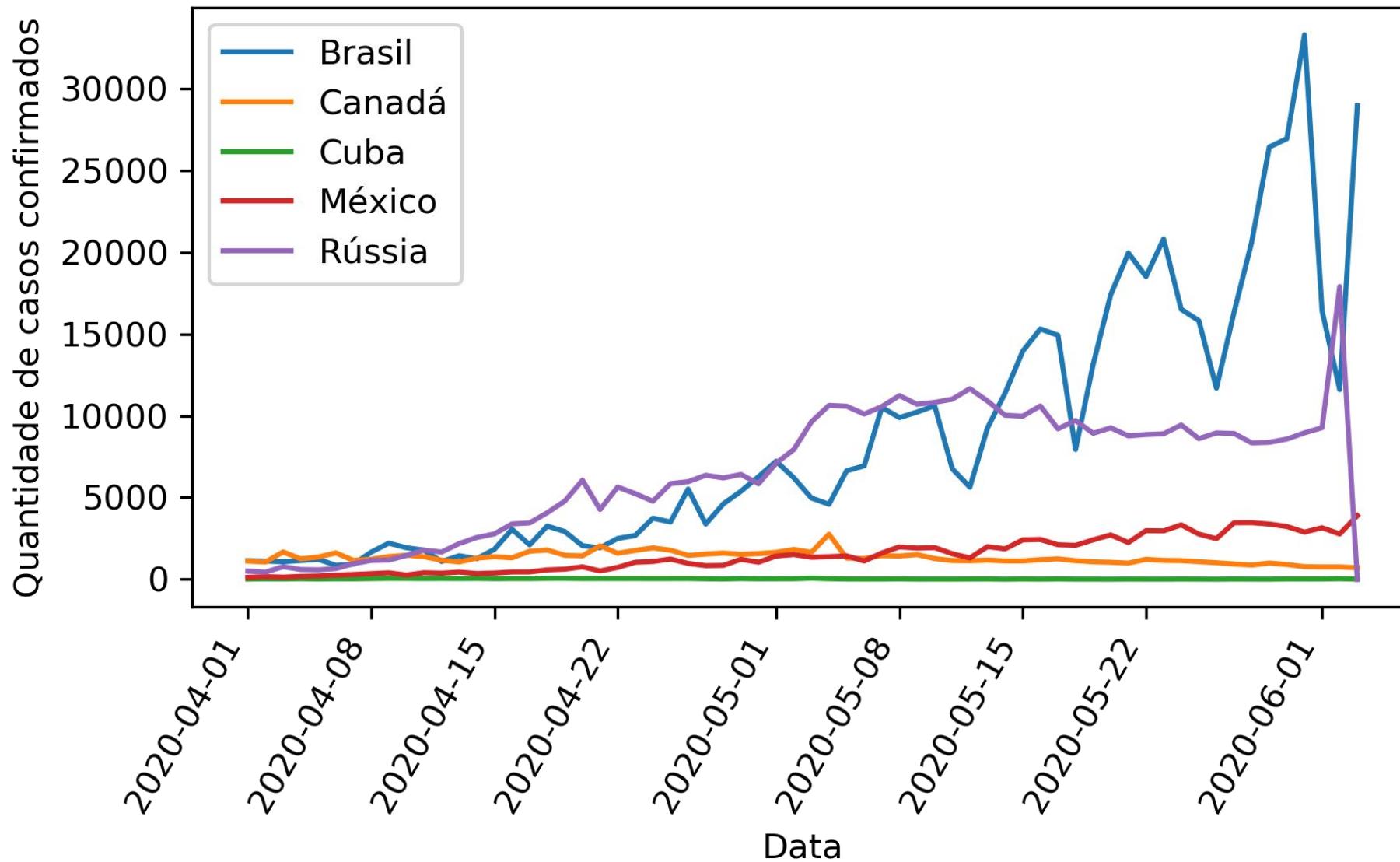
Dados internacionais



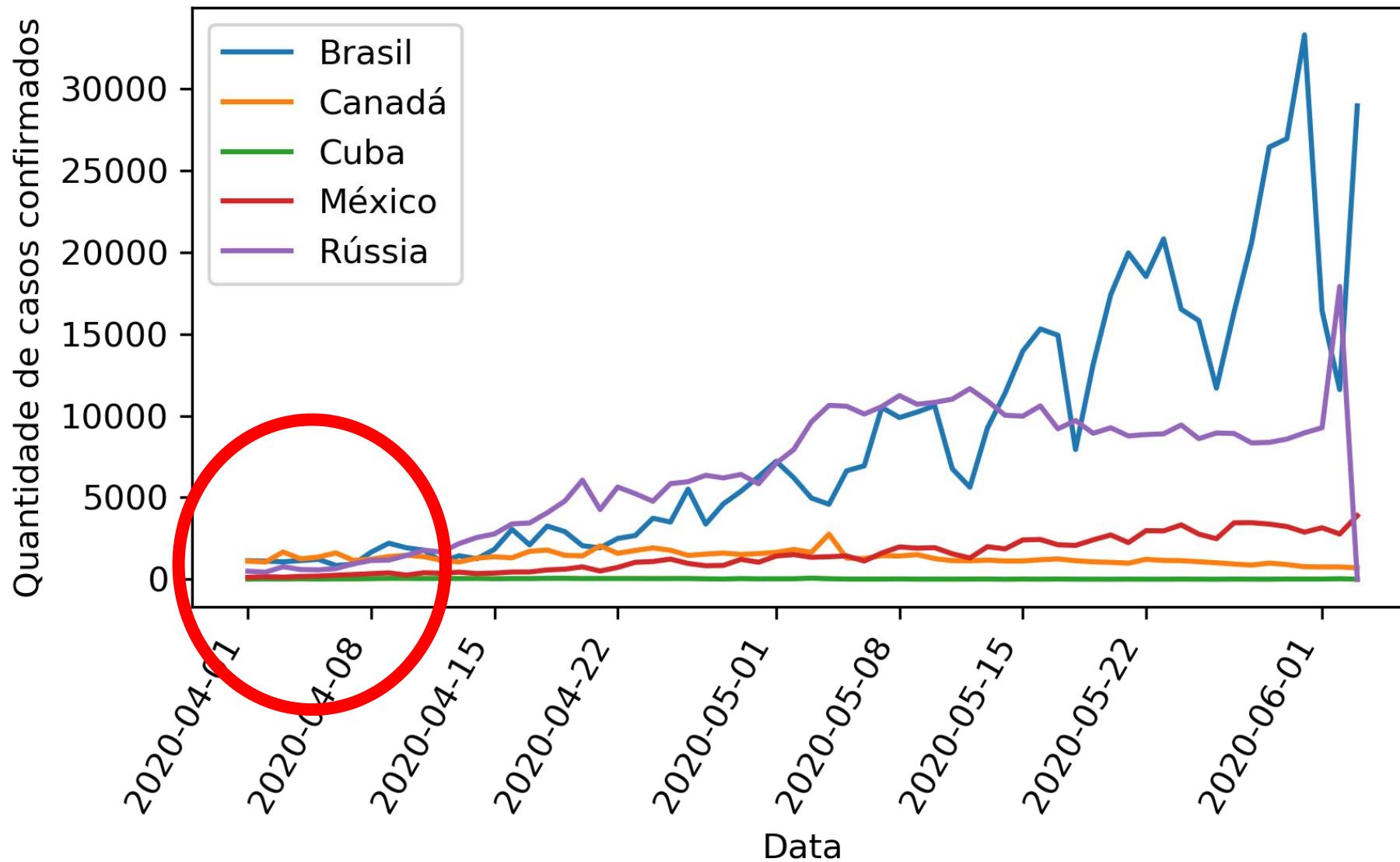
# Tratamento dos dados

- Padronização dos tipos de datas;
- Filtro para as variáveis de interesse;
- Redução da quantidade de dias antes dos primeiros registros de infectados.

## Casos de COVID-19 confirmados diariamente

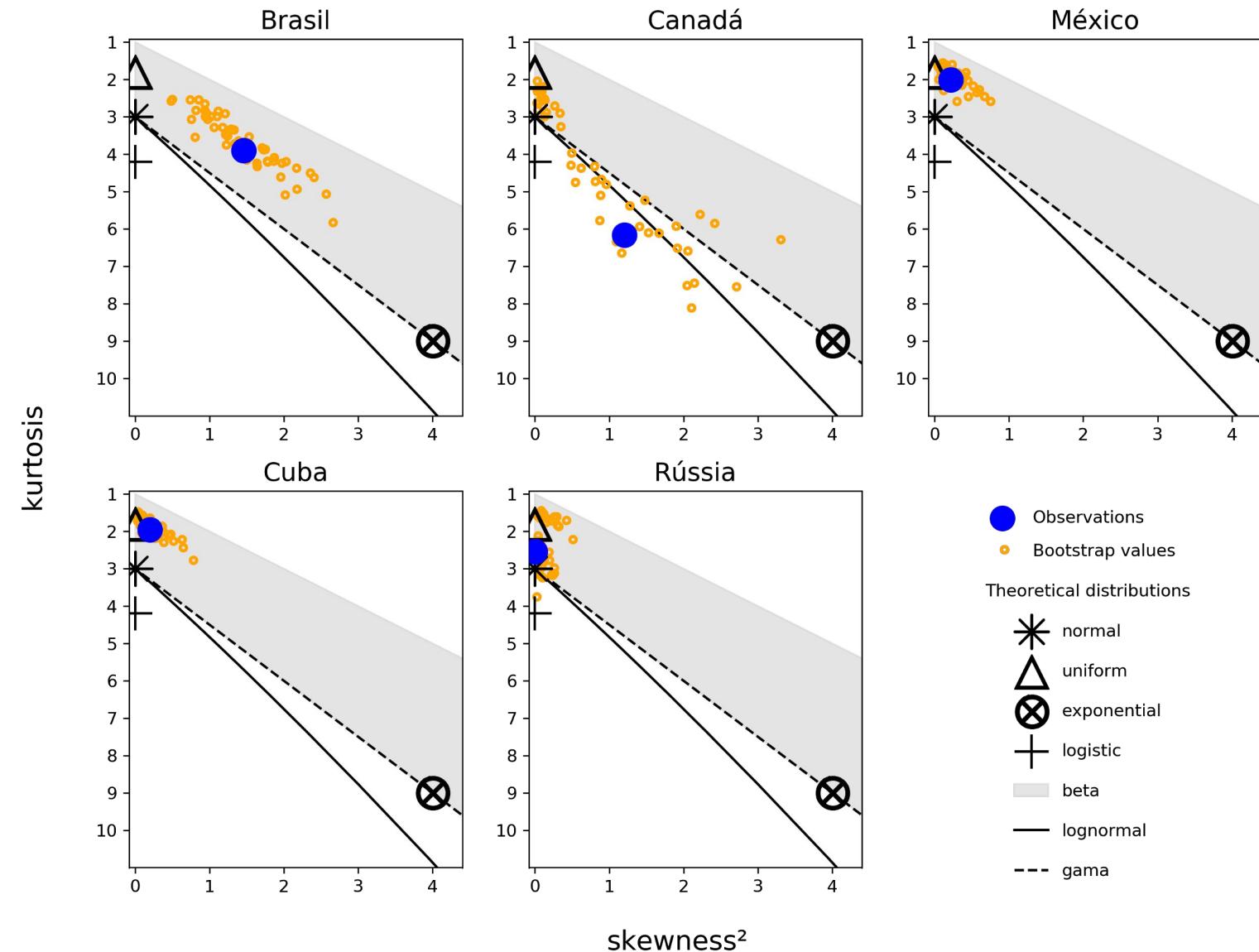


## Casos de COVID-19 confirmados diariamente

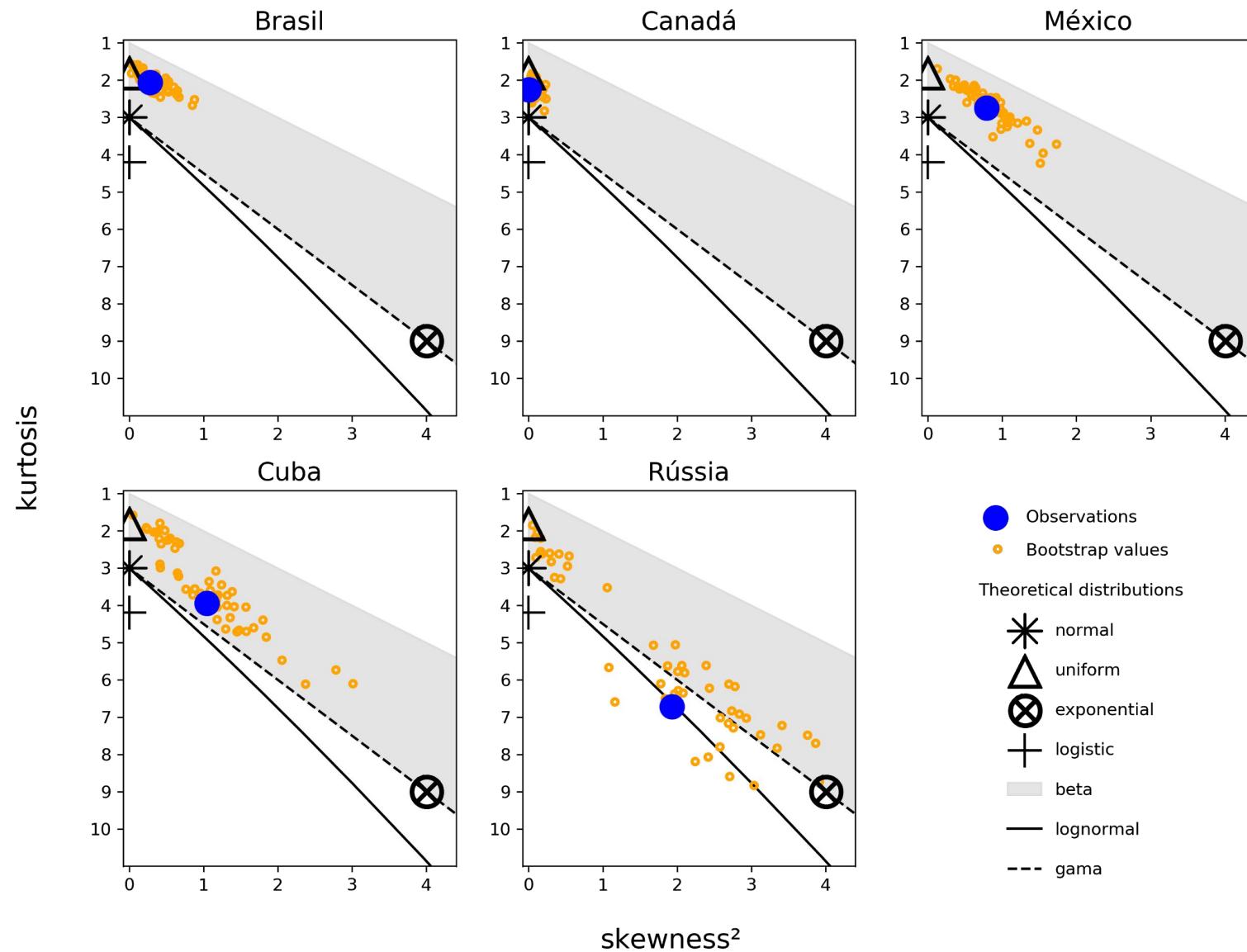


# **Caracterização dos dados**

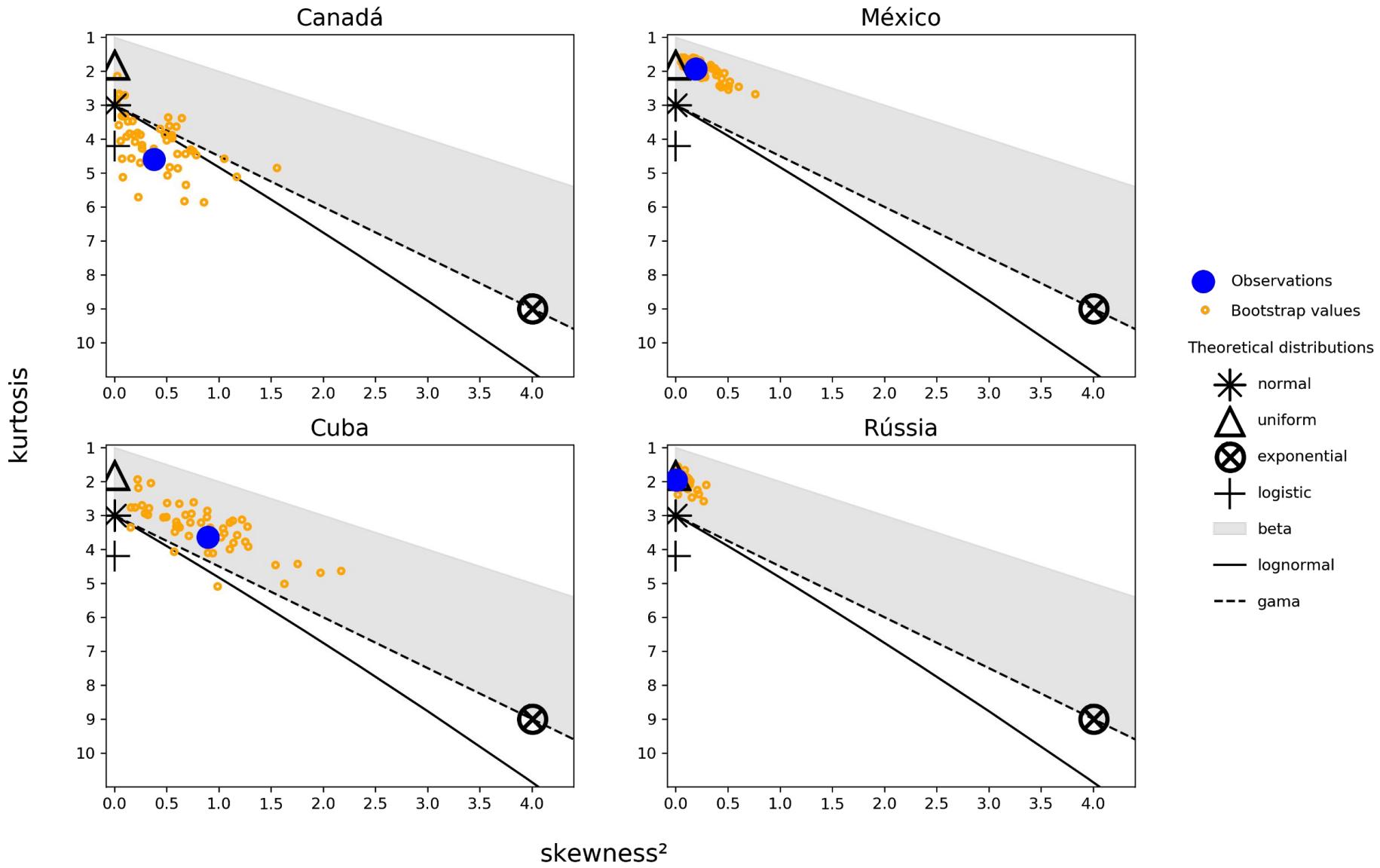
## Número de casos diários



## Número diário de mortes

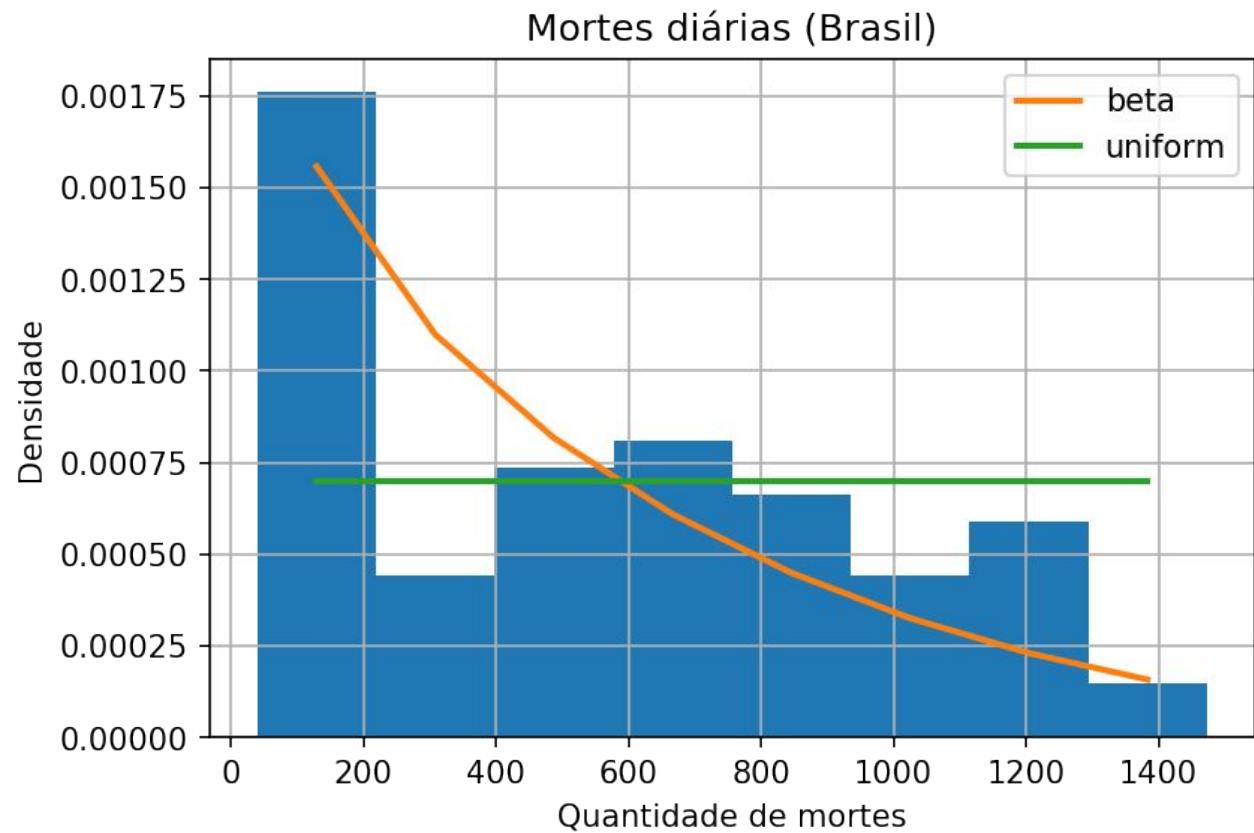
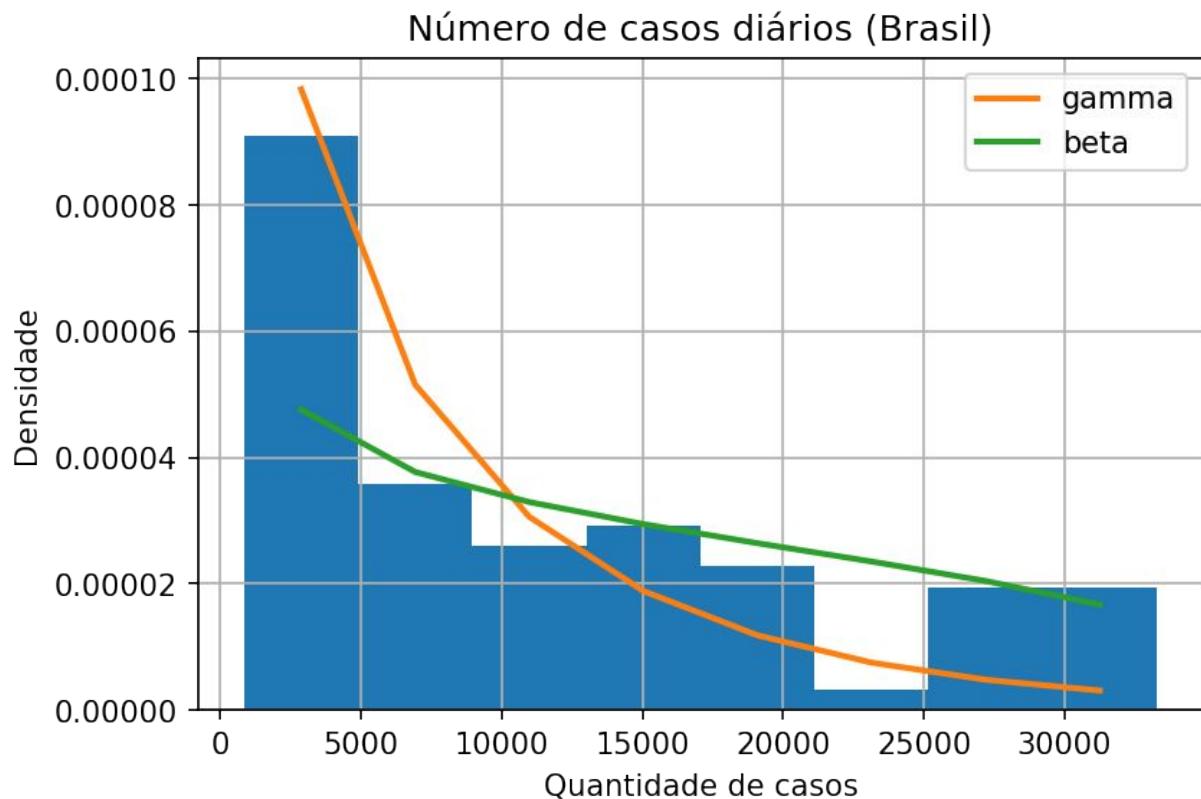


## Número diário de testes

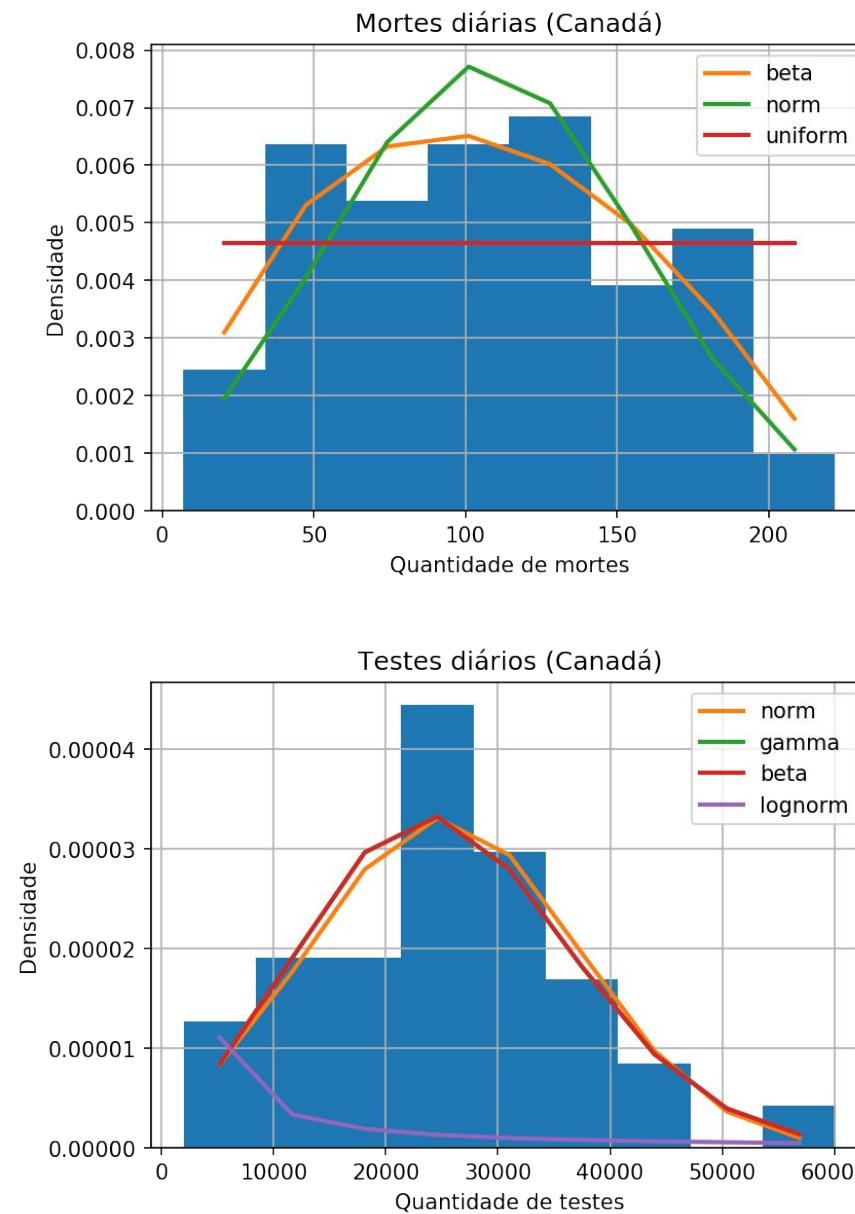
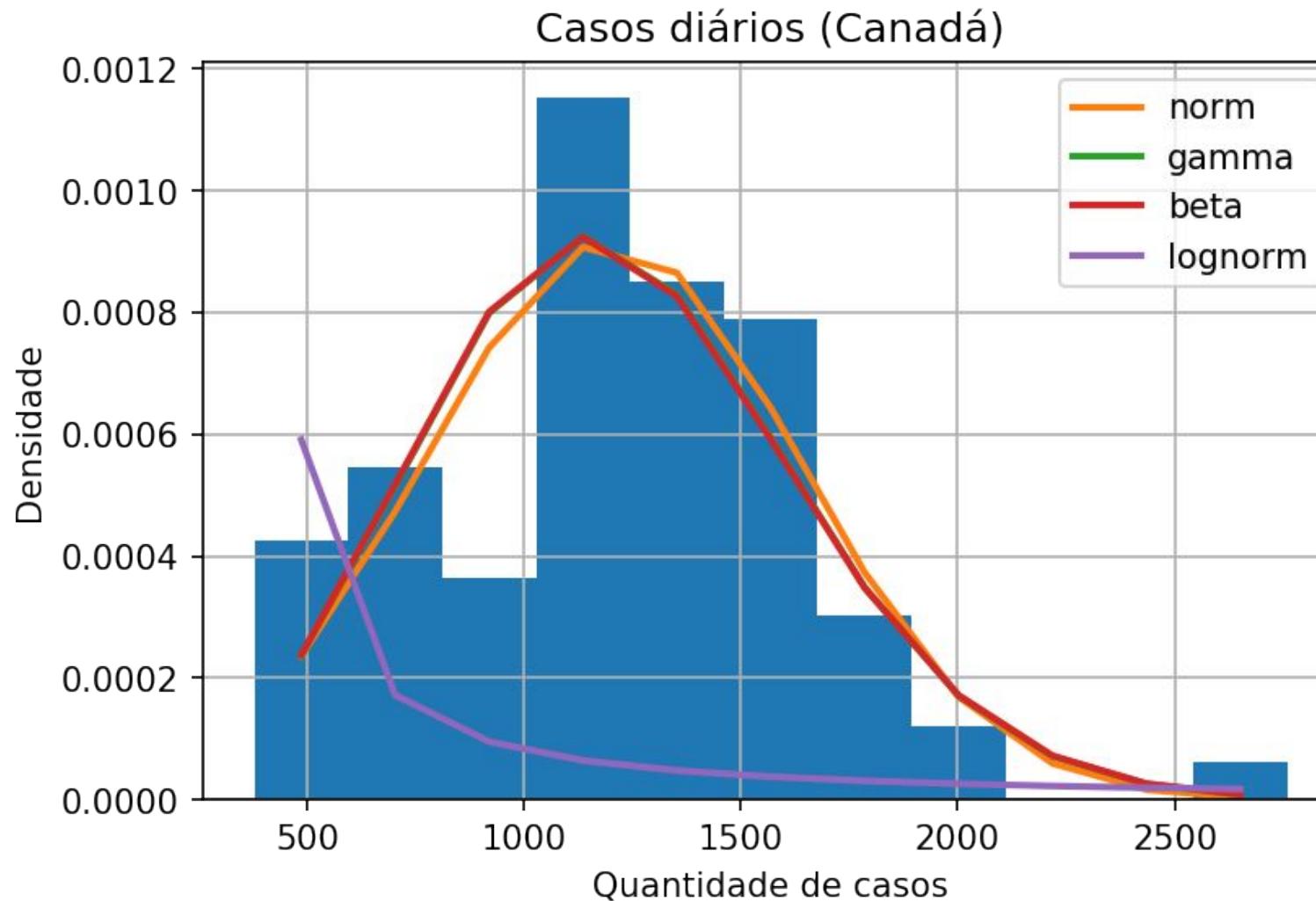


# Histogramas e PDFs

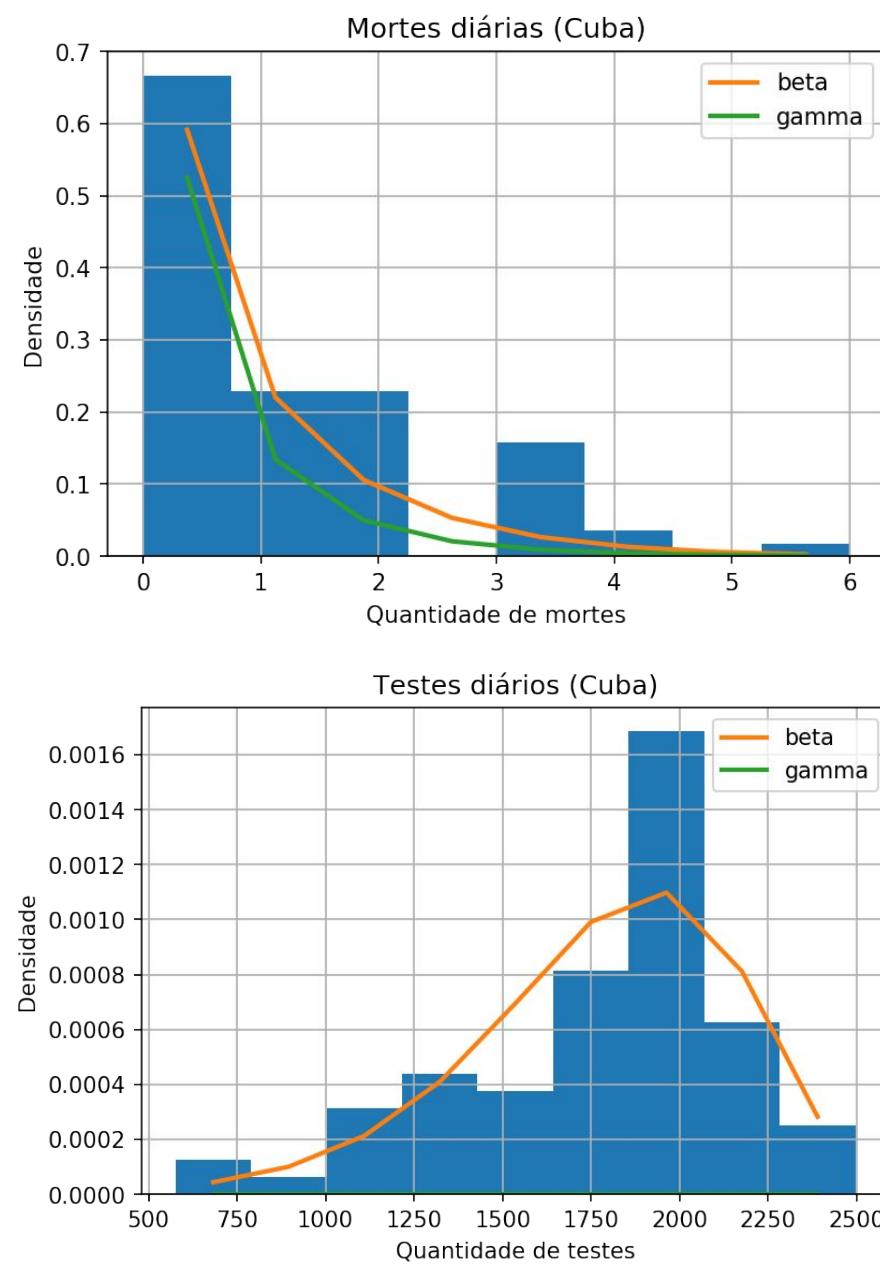
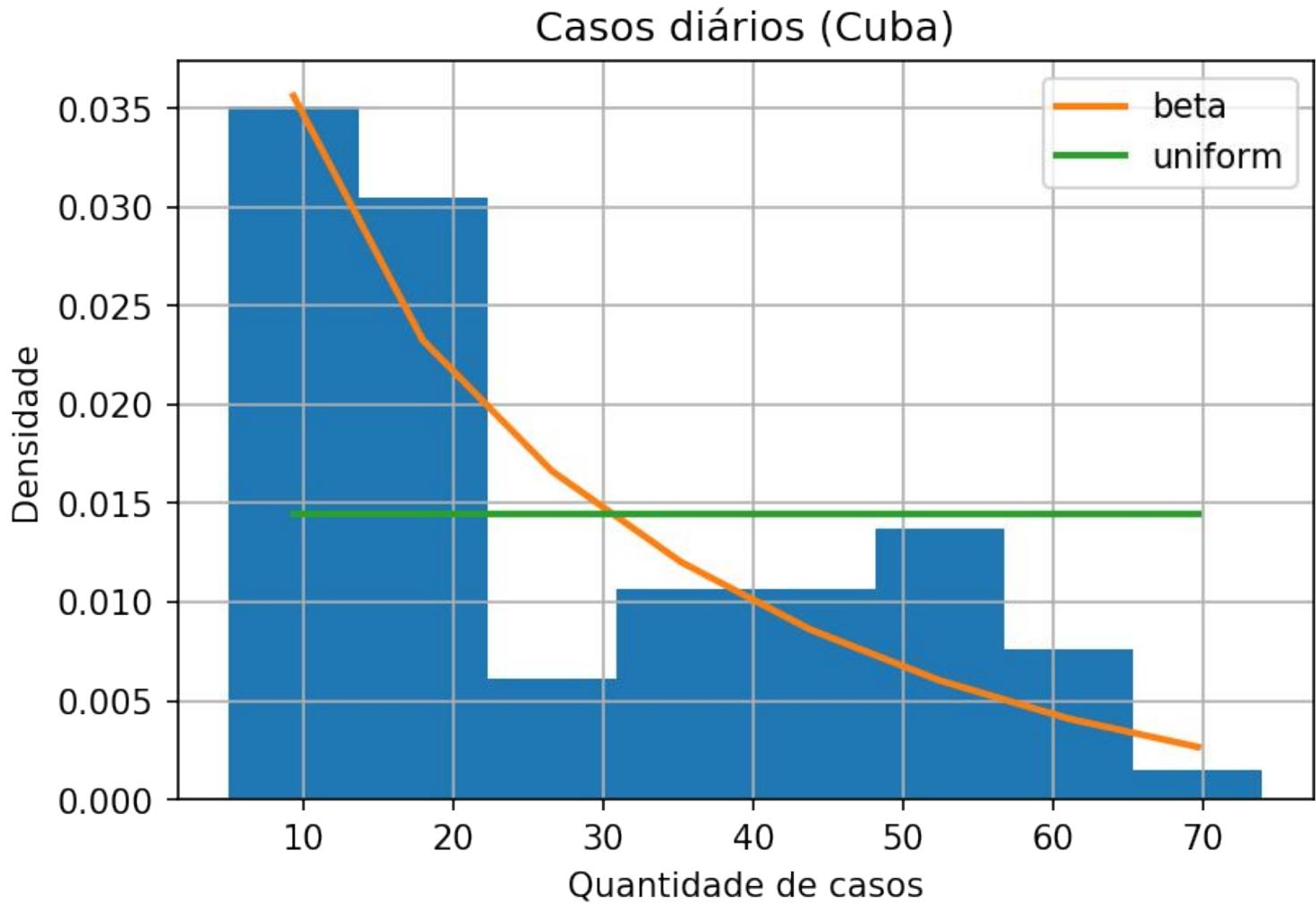
# Histogramas e PDFs - Brasil



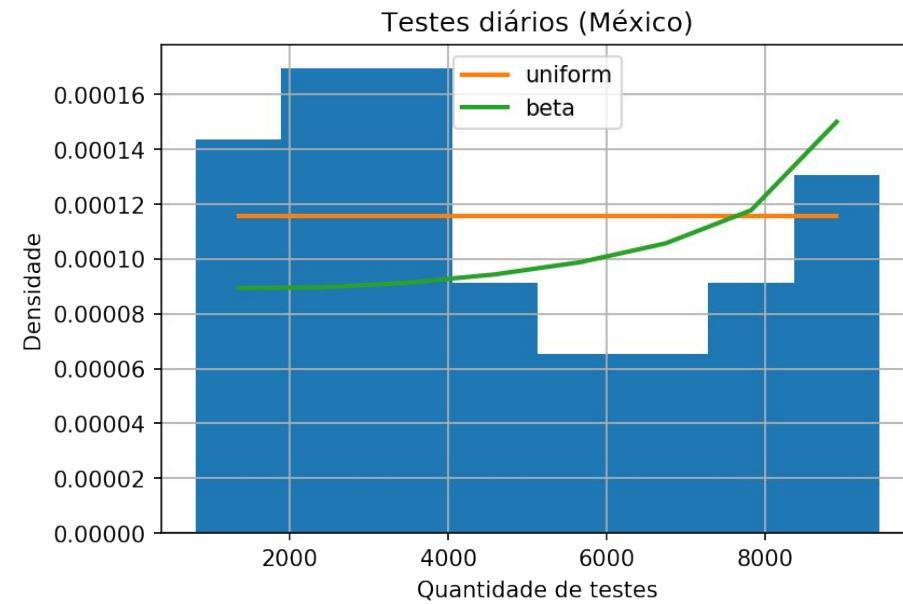
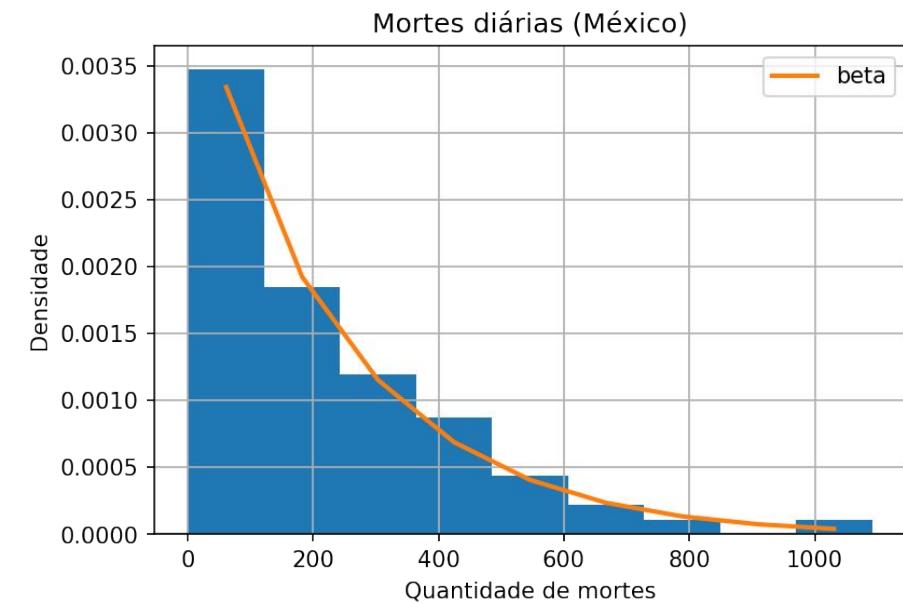
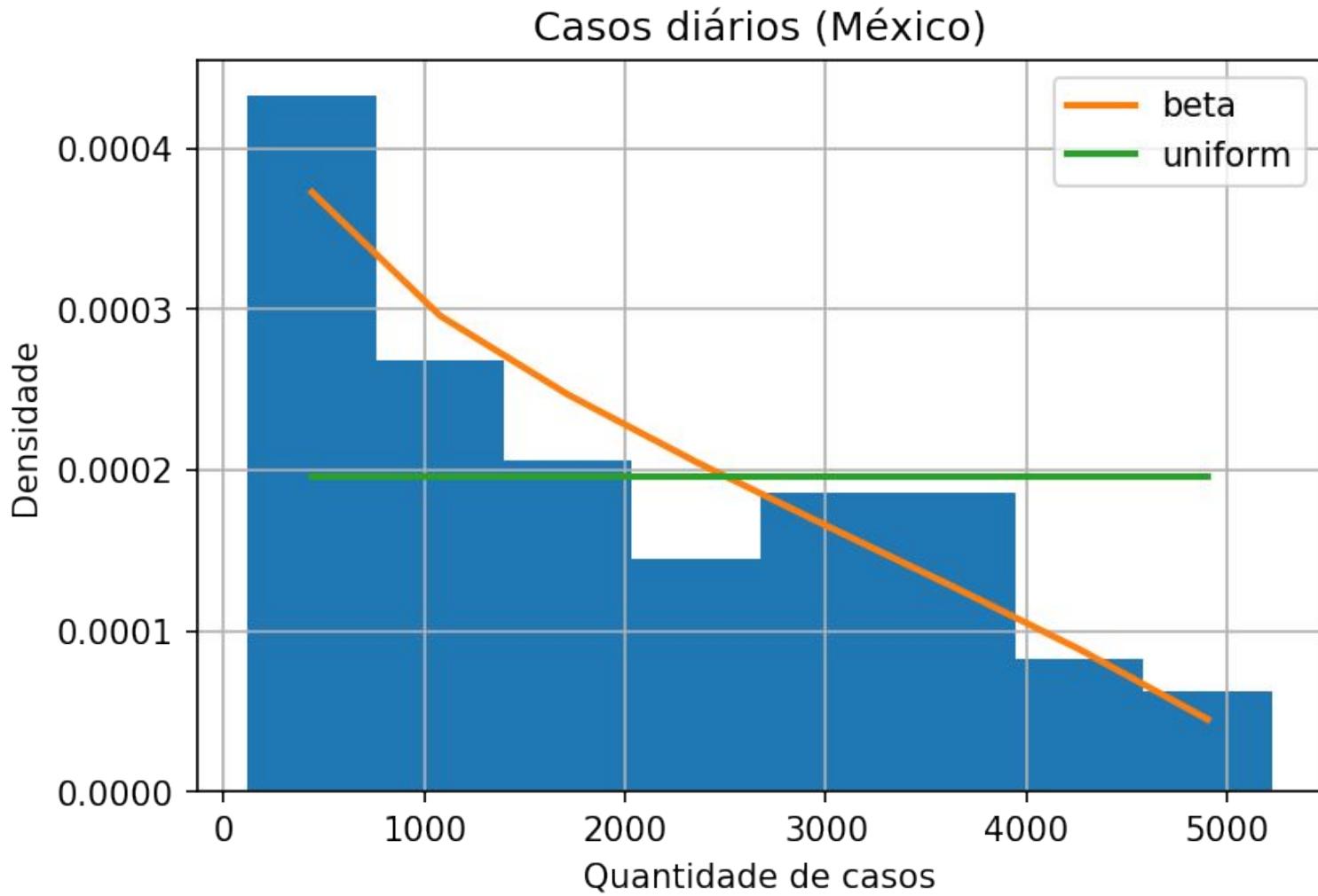
# Histogramas e PDFs - Canadá



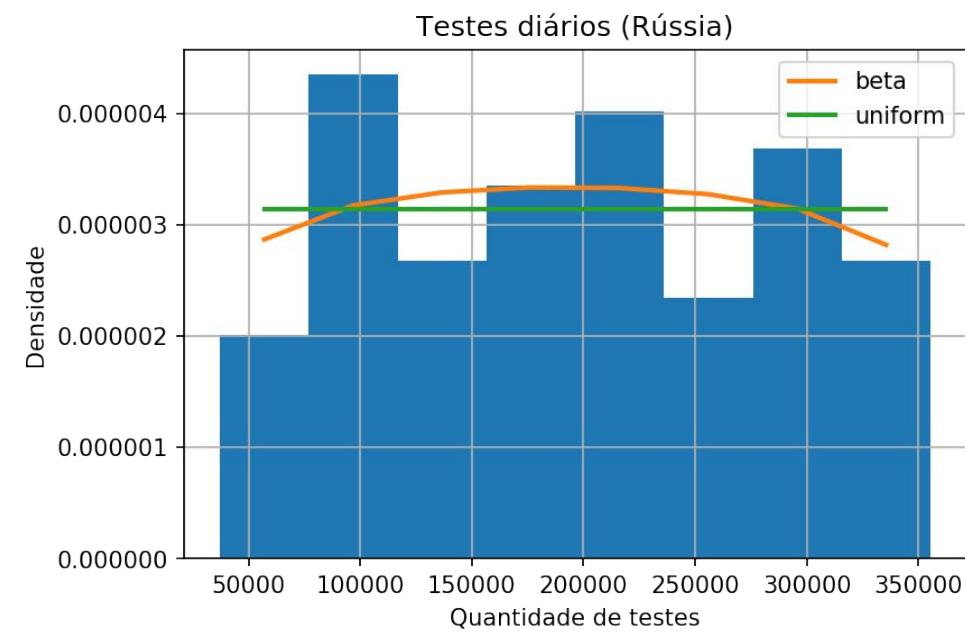
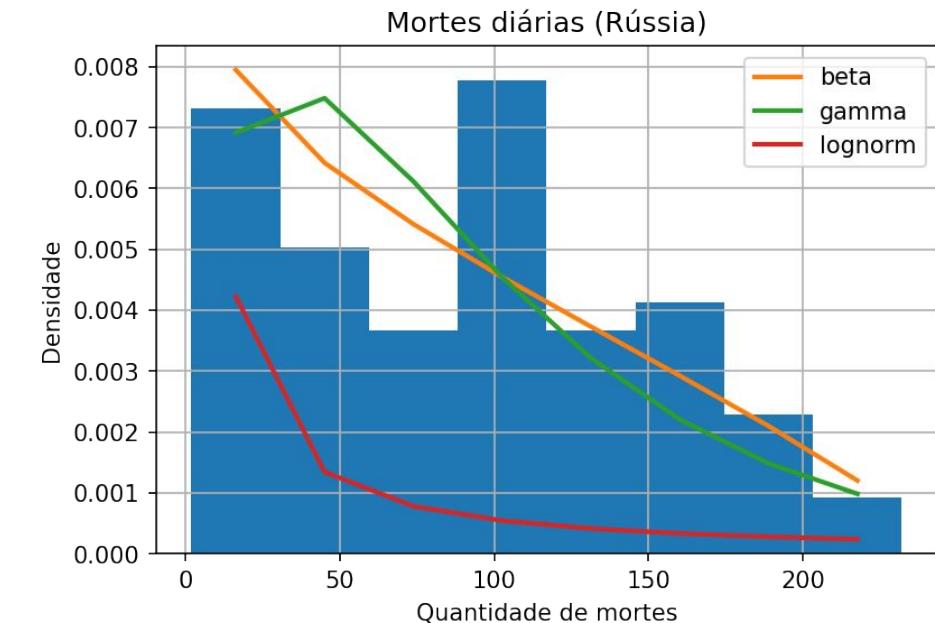
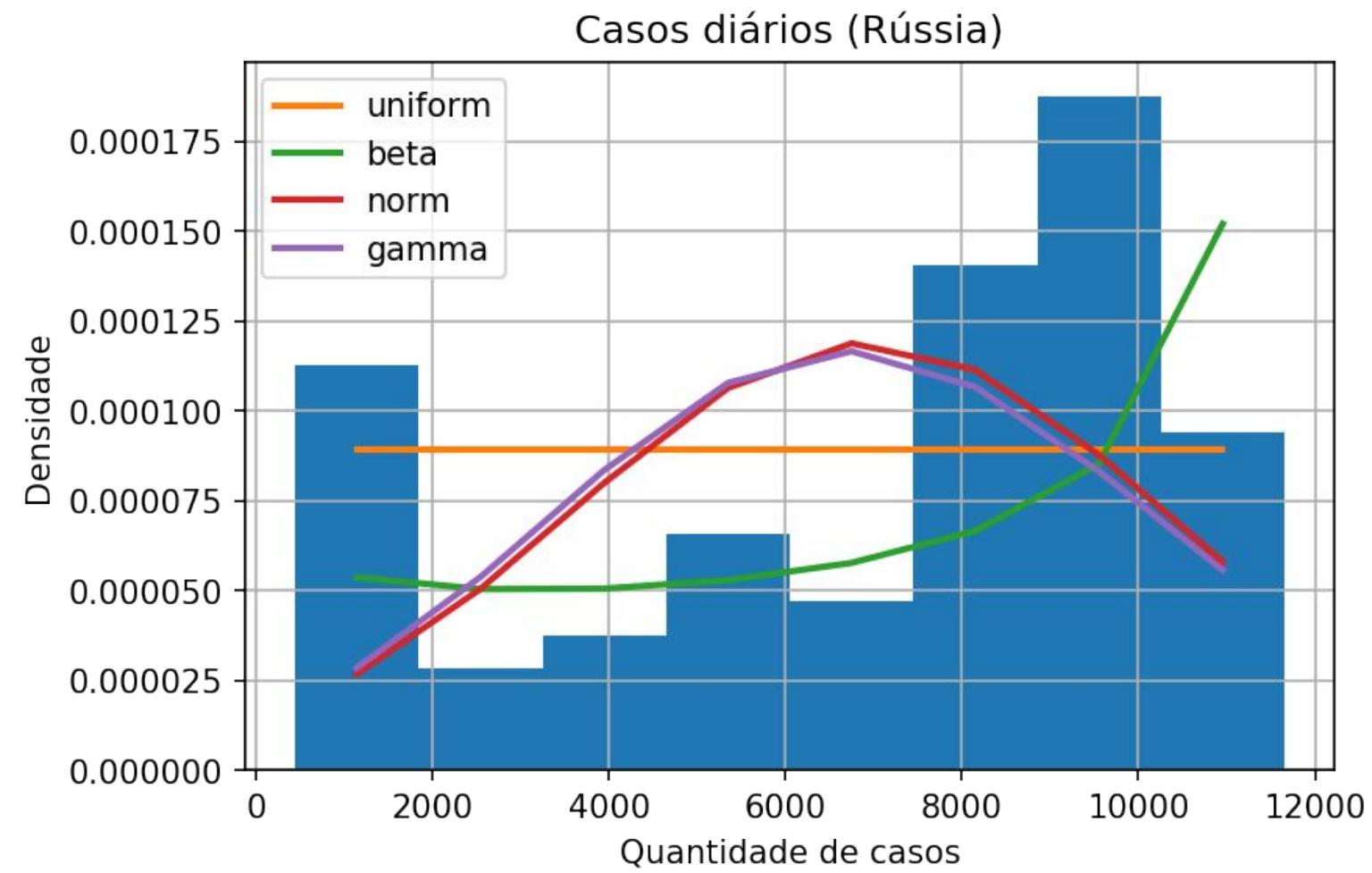
# Histogramas e PDFs - Cuba



# Histogramas e PDFs - México



# Histogramas e PDFs - Rússia

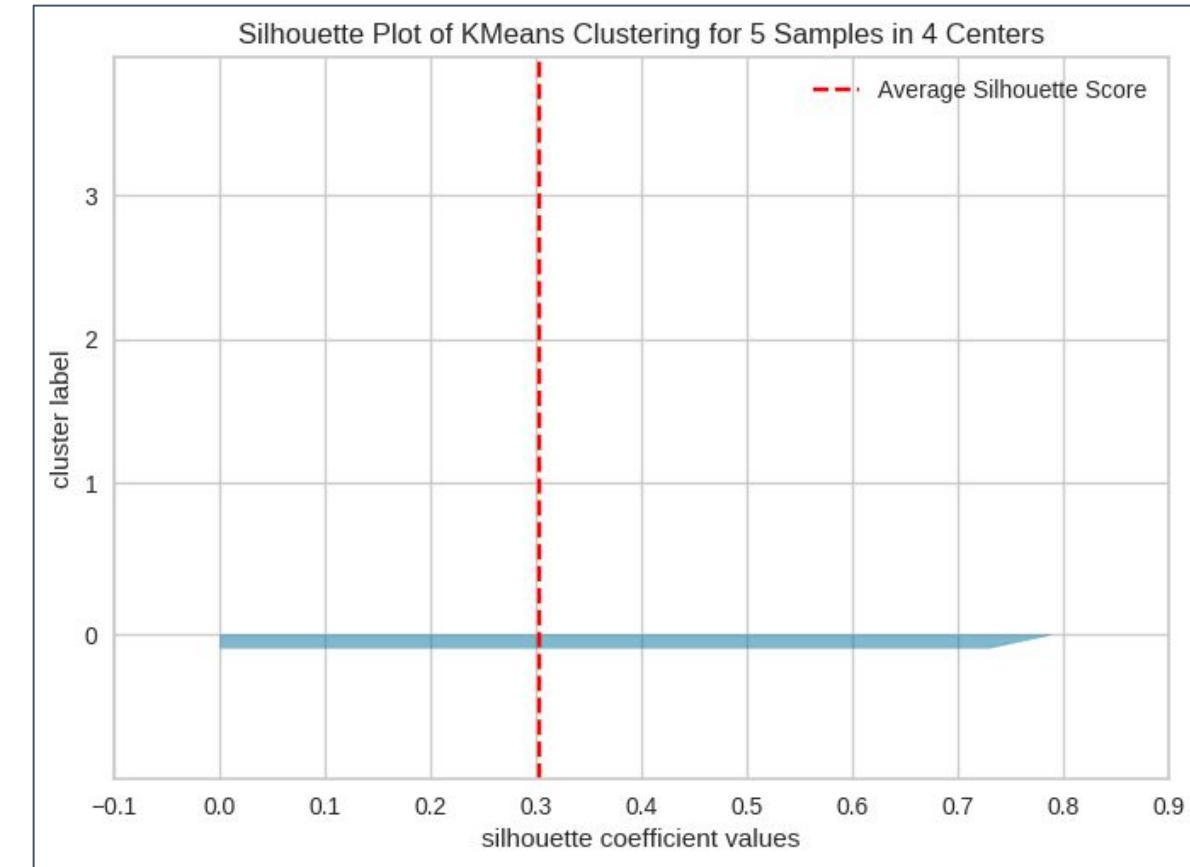
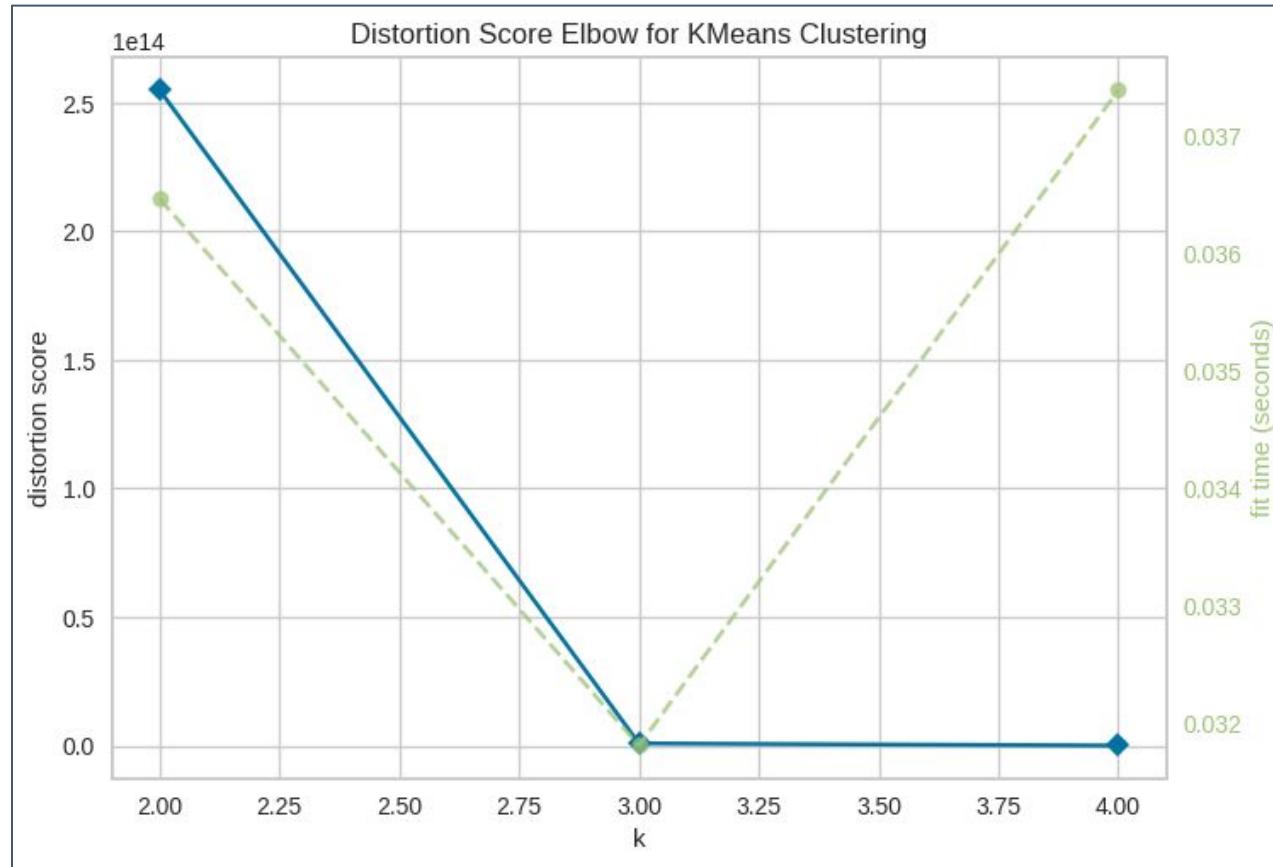


# **Há algum comportamento implícito dos dados ?**

# **Há algum comportamento implícito dos dados ?**

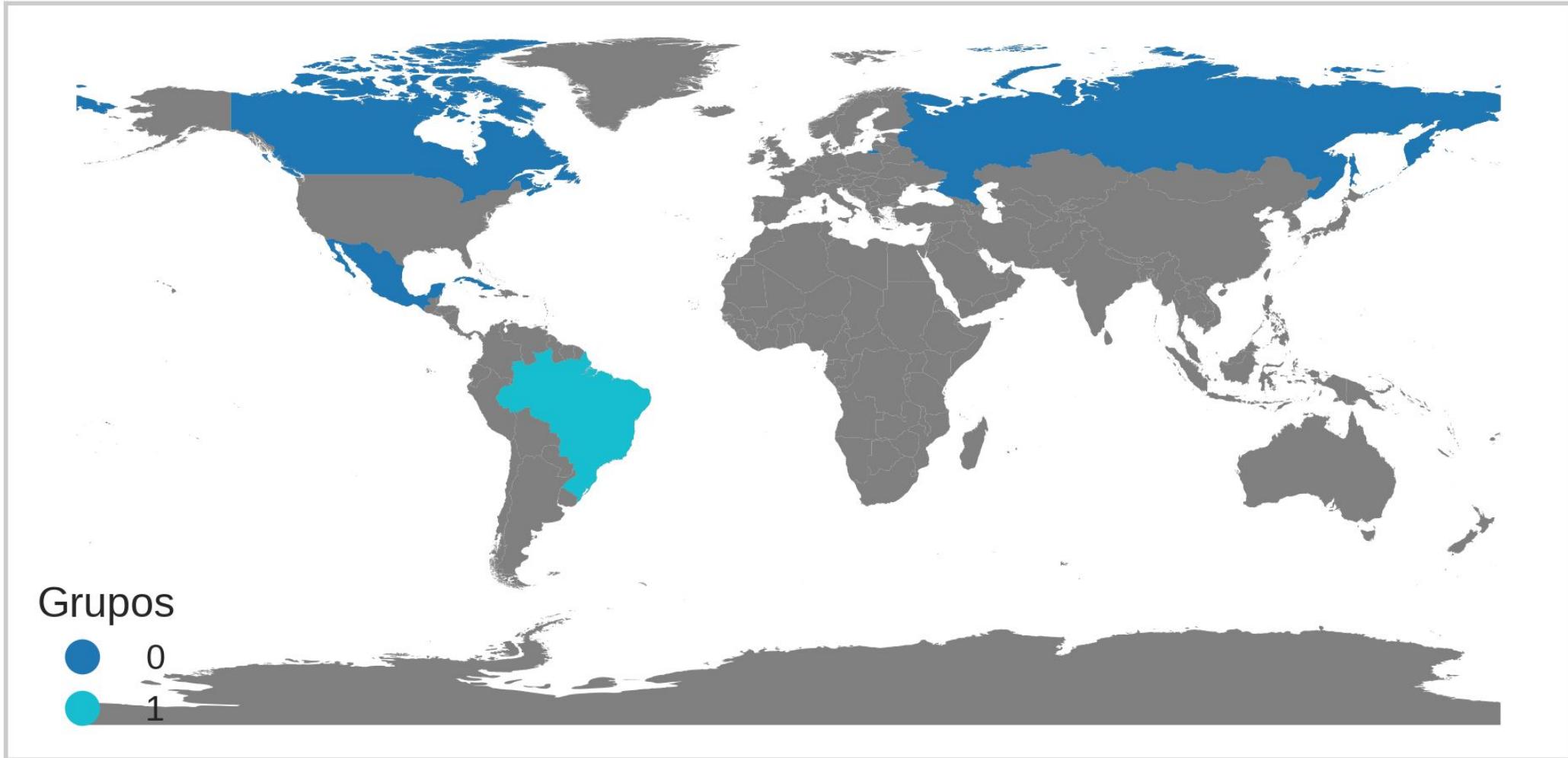
## **Análise de agrupamento**

# Problemas com os métodos de verificação de $K$



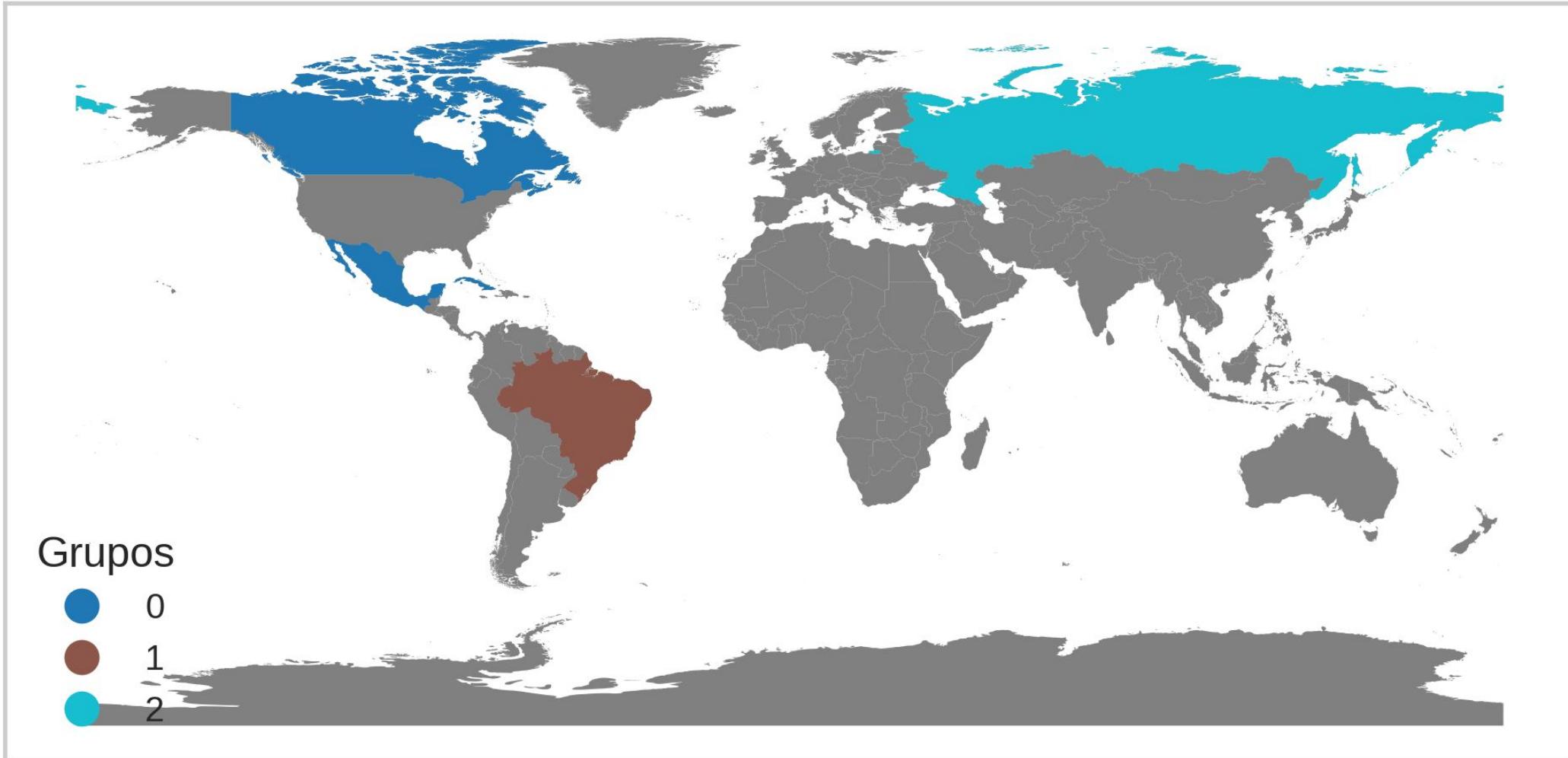
**K = 2**

Países agrupados pela quantidade diária de casos e mortes



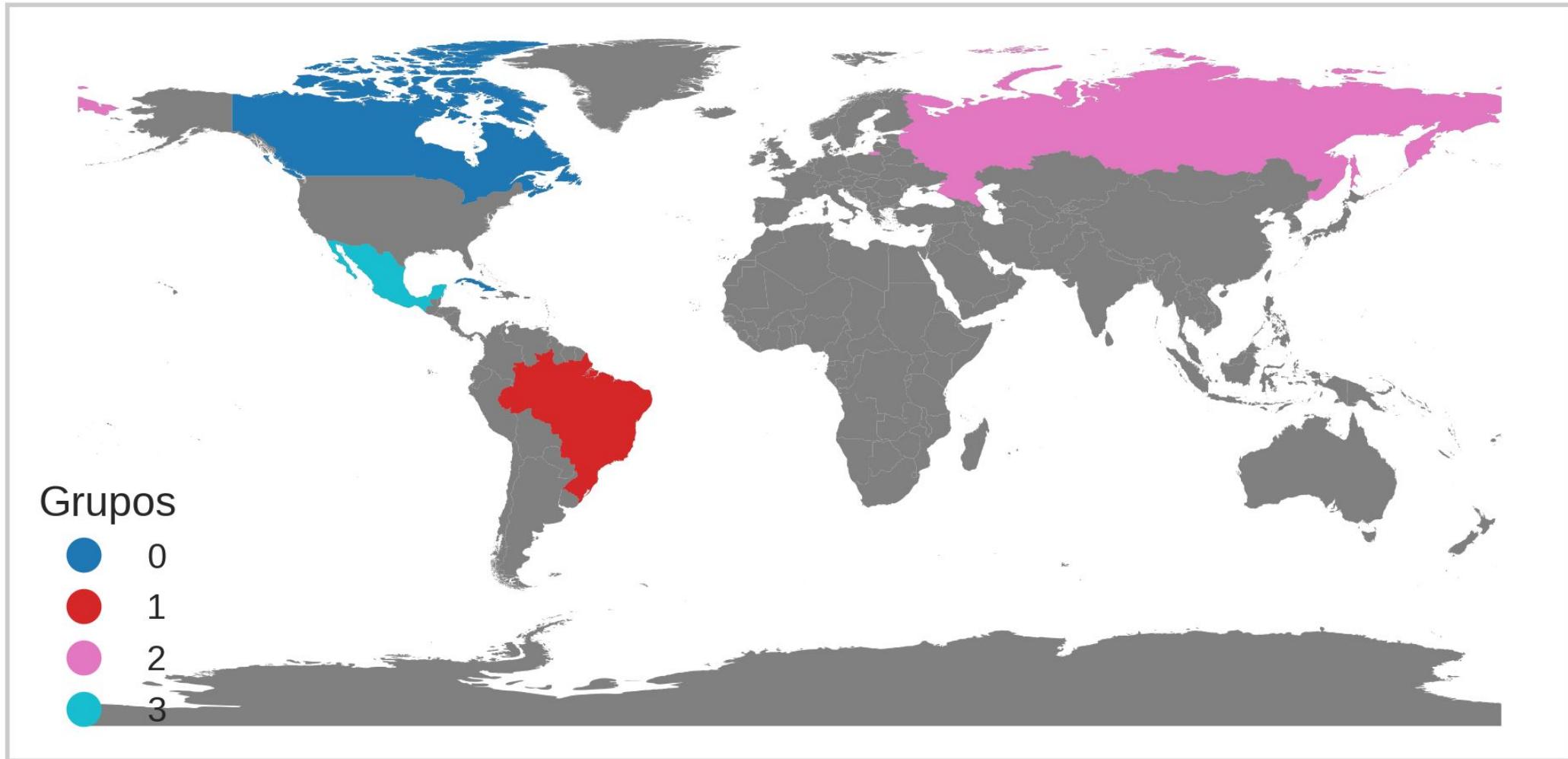
**K = 3**

Países agrupados pela quantidade diária de casos e mortes



**K = 4**

Países agrupados pela quantidade diária de casos e mortes



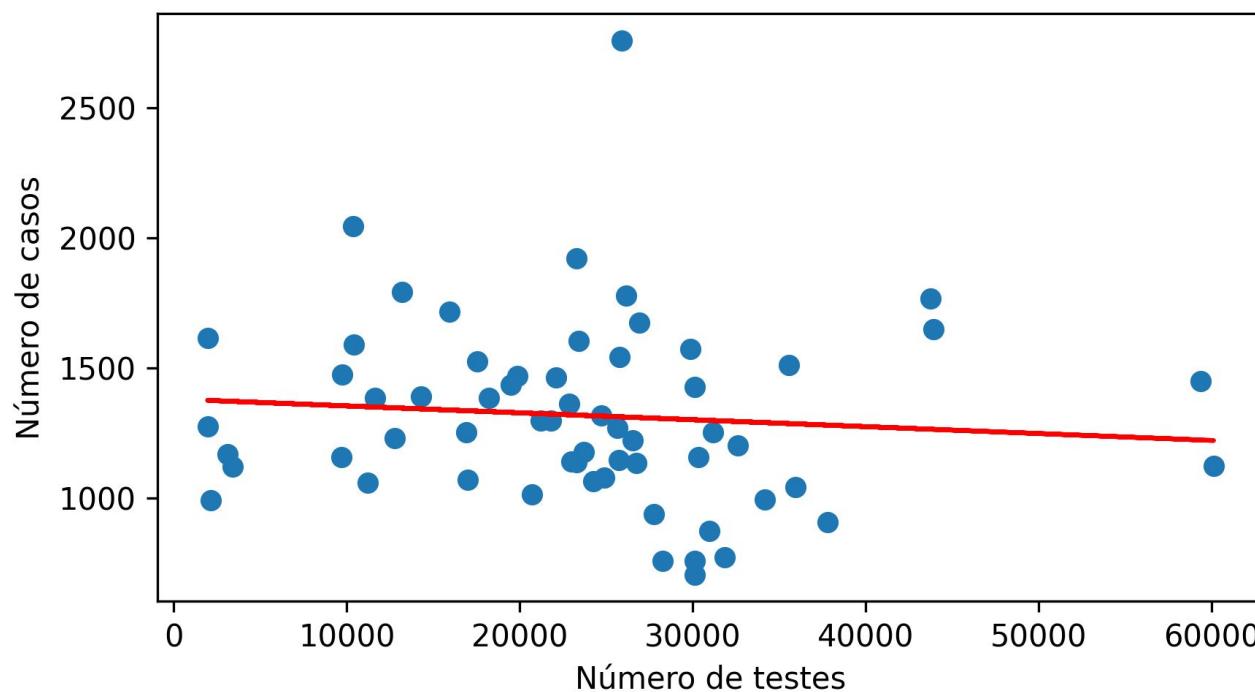
# Análise de regressão

# Análise de regressão

Regressão Linear (Canadá)

Coef. Corr. = -0.0899

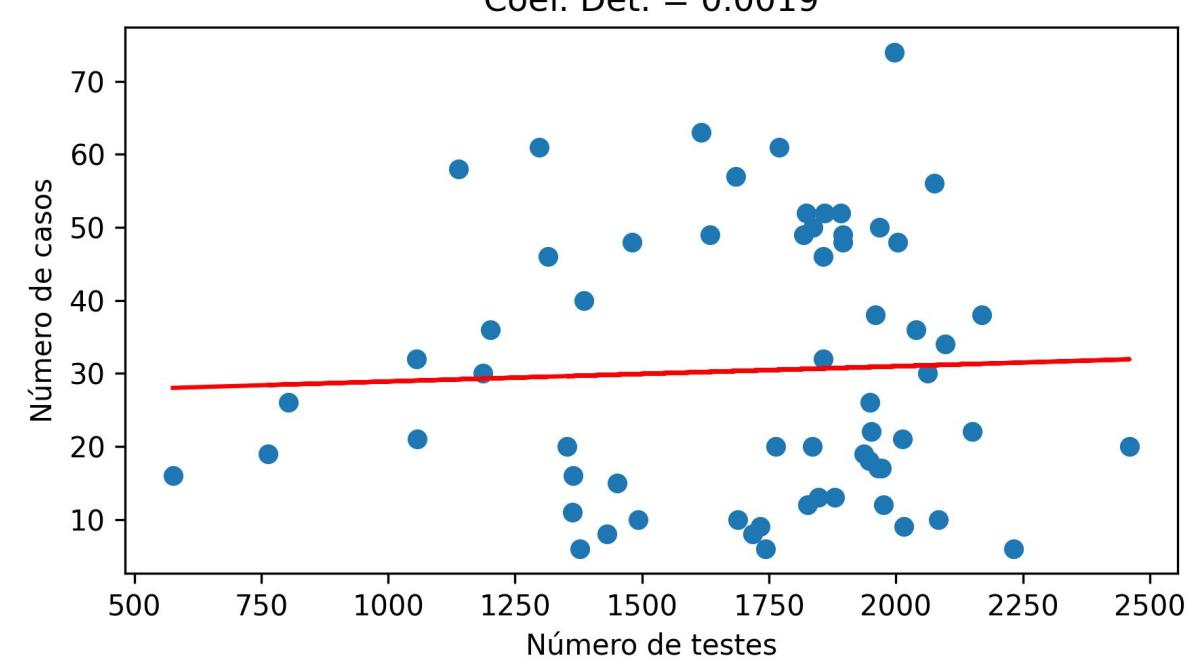
Coef. Det. = 0.0081



Regressão Linear (Cuba)

Coef. Corr. = 0.0433

Coef. Det. = 0.0019

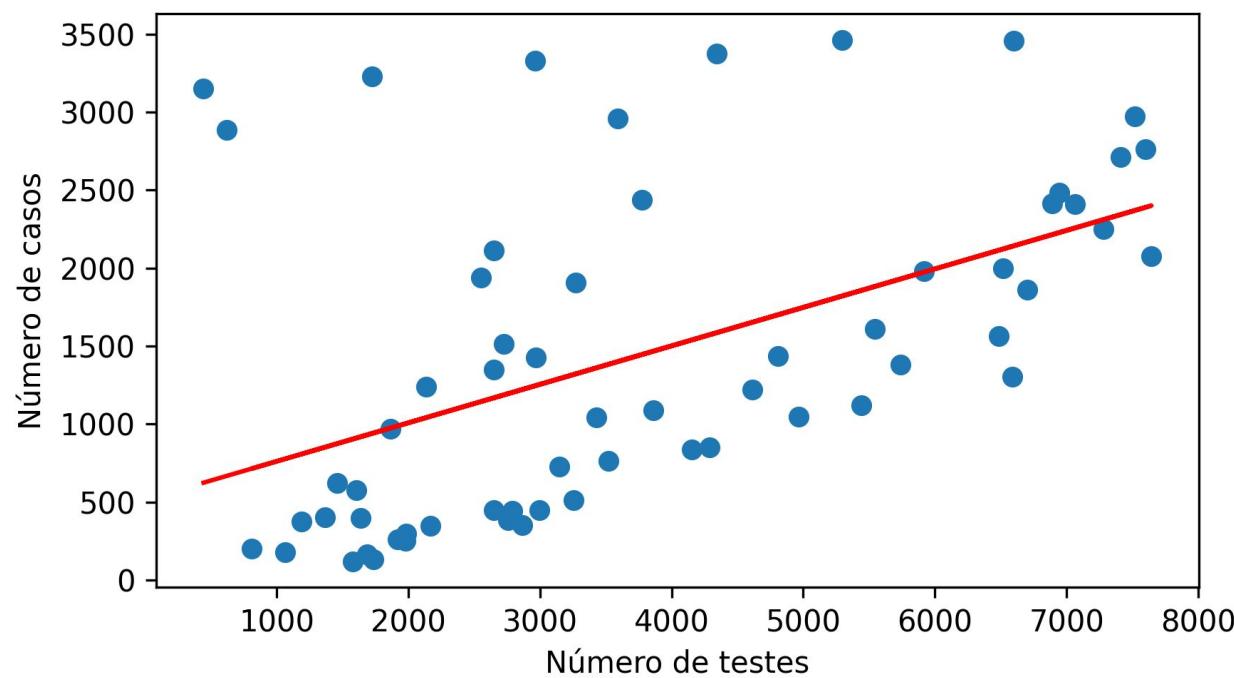


# Análise de regressão

Regressão Linear (México)

Coef. Corr. = 0.5036

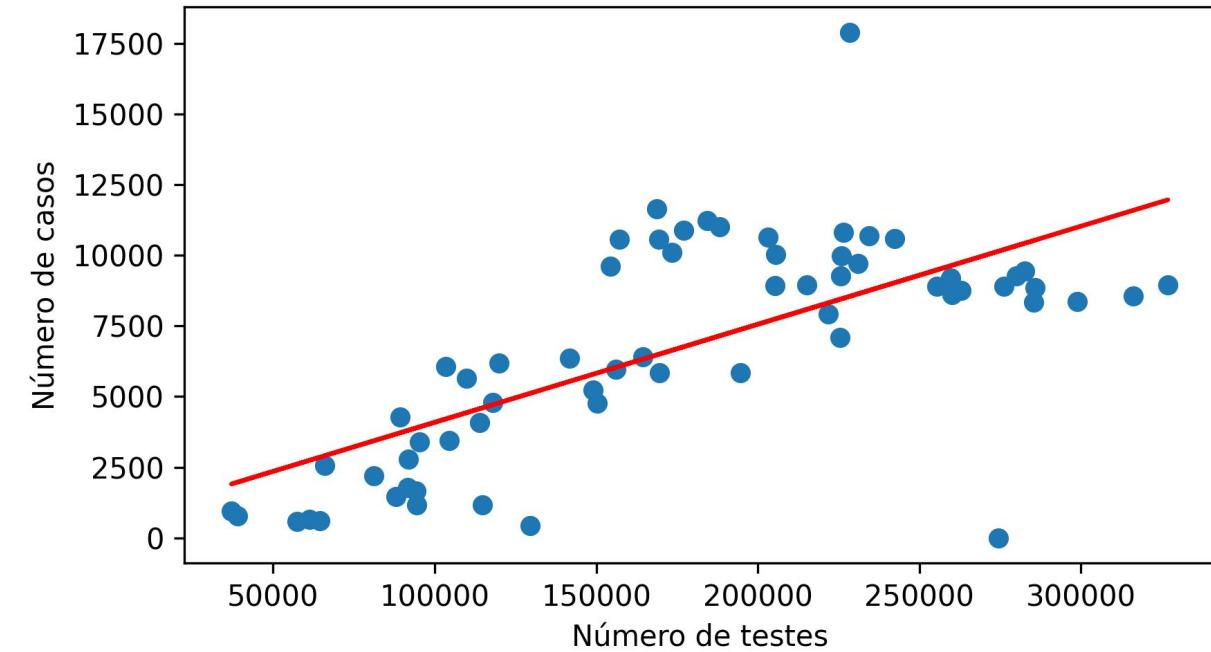
Coef. Det. = 0.2536



Regressão Linear (Rússia)

Coef. Corr. = 0.6881

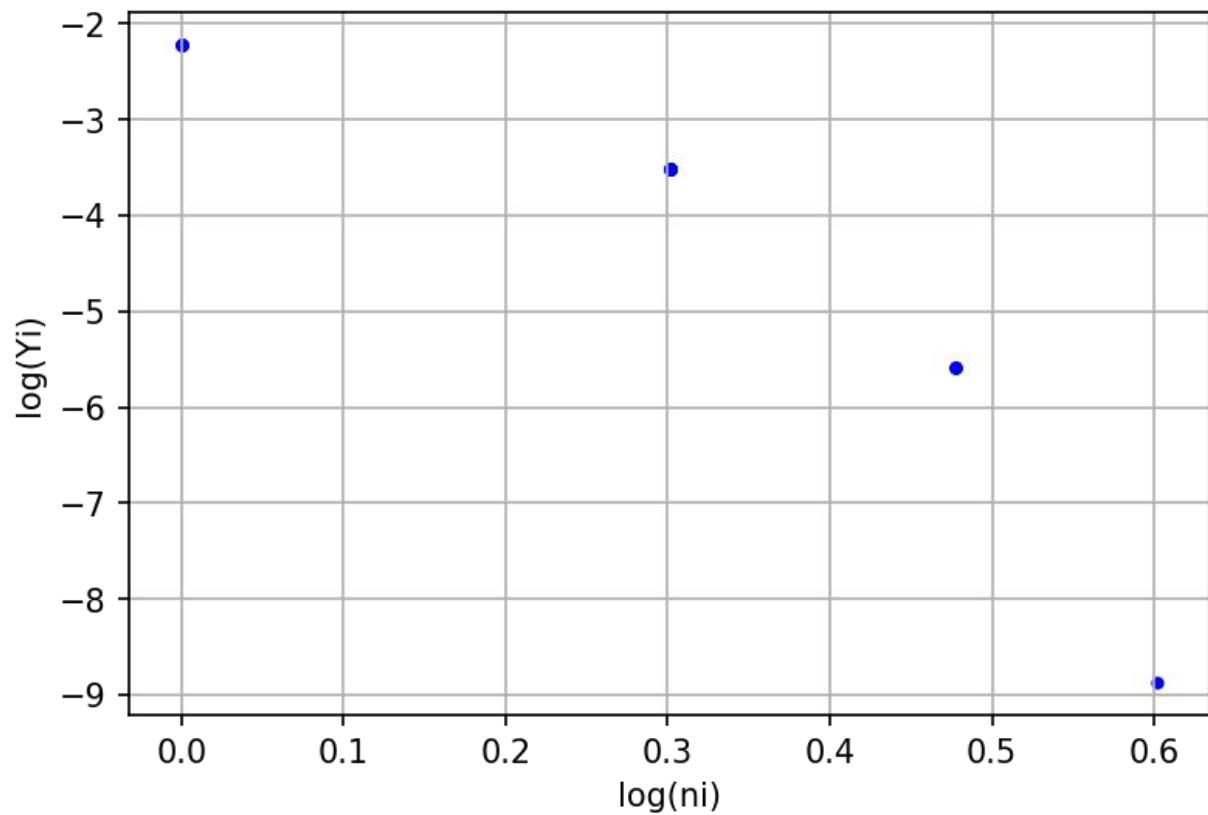
Coef. Det. = 0.4734



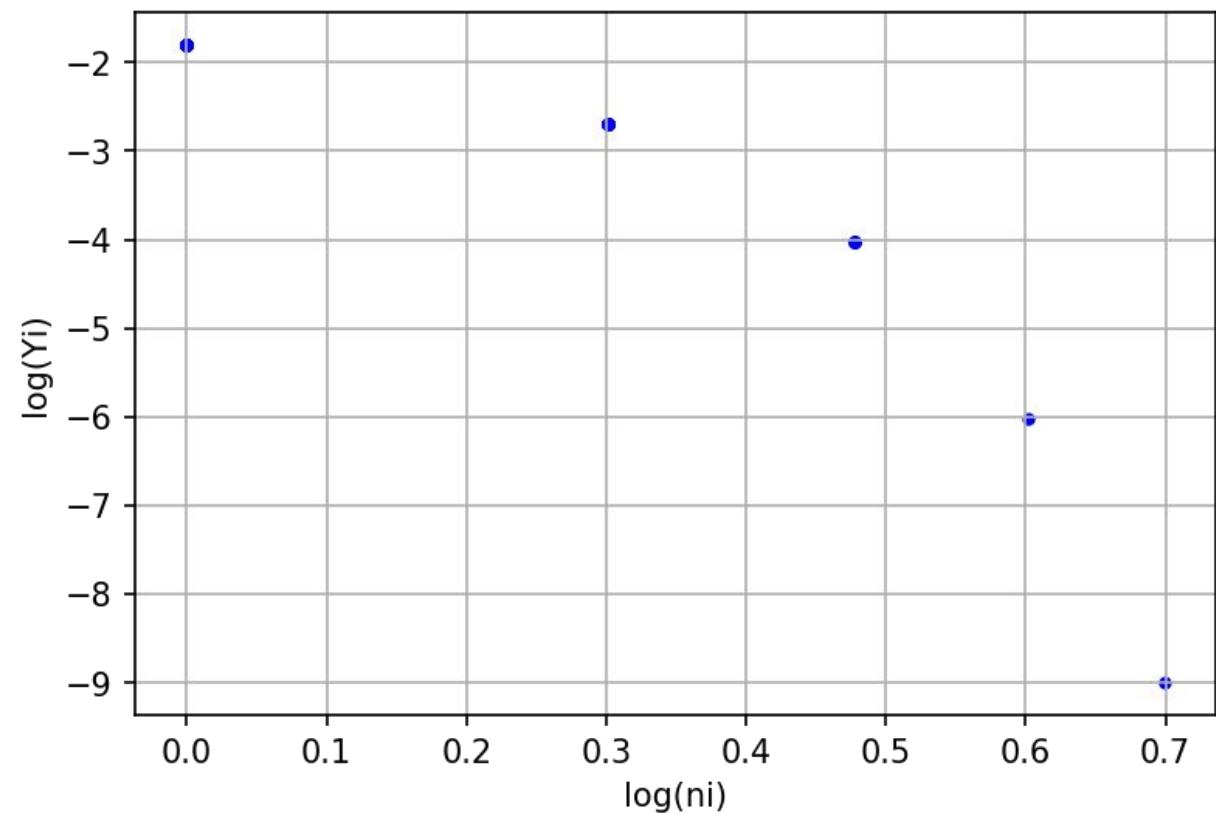
# **Criticalidade auto-organizada (soc)**

# Assinatura SOC

SOC de casos diários - Cuba



SOC de casos diários - México

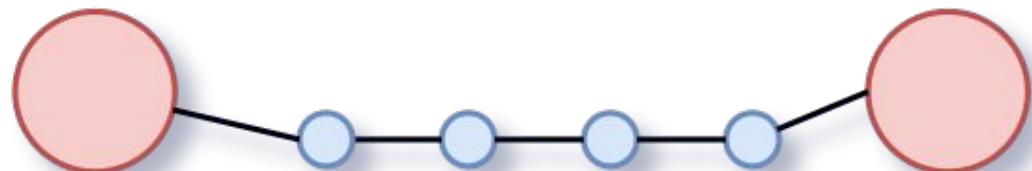


# Interpolação aleatória

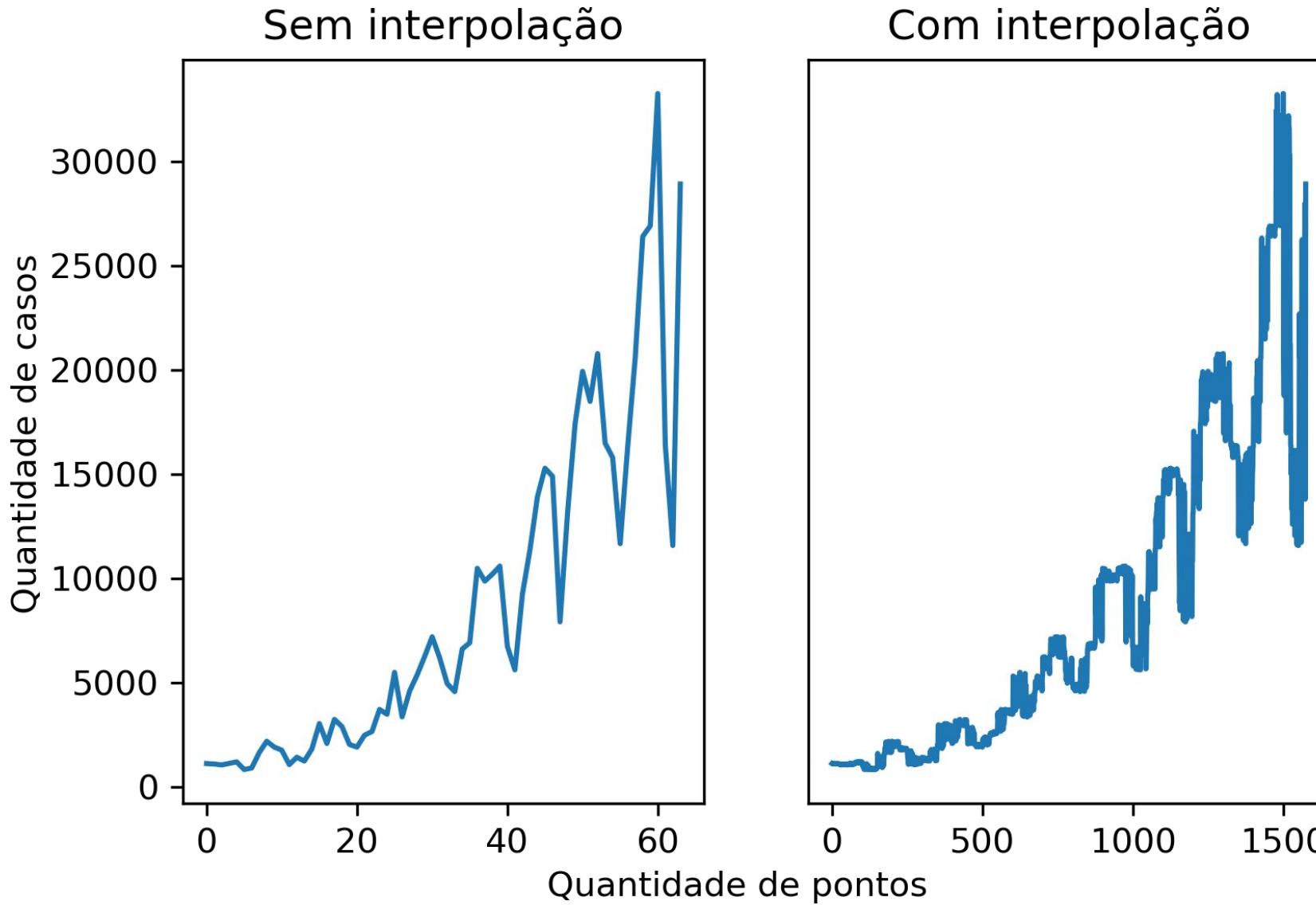
(A) Método por inteiros



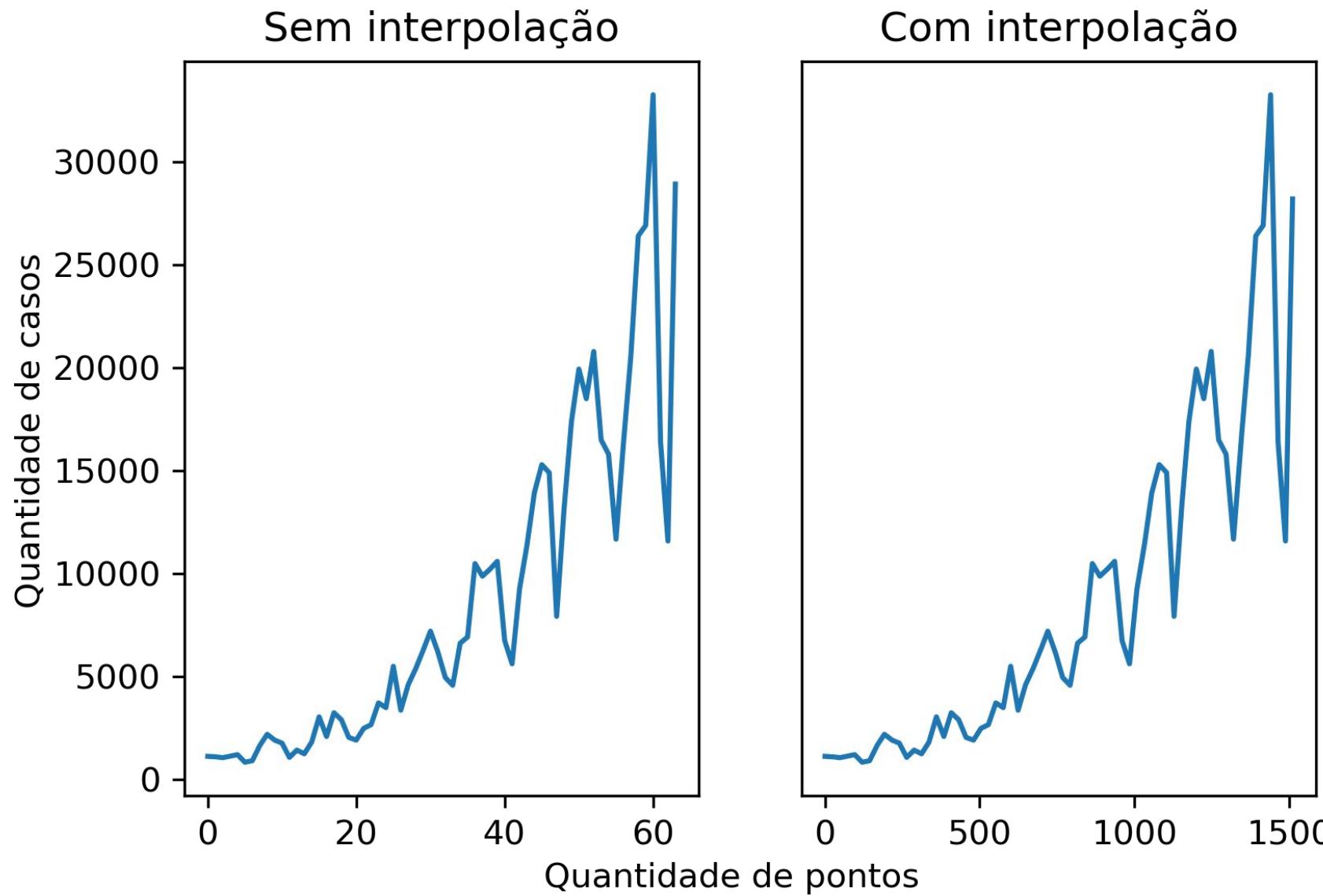
(B) Método por polinômios



# Interpolação aleatória com 23 pontos - Método de inteiros

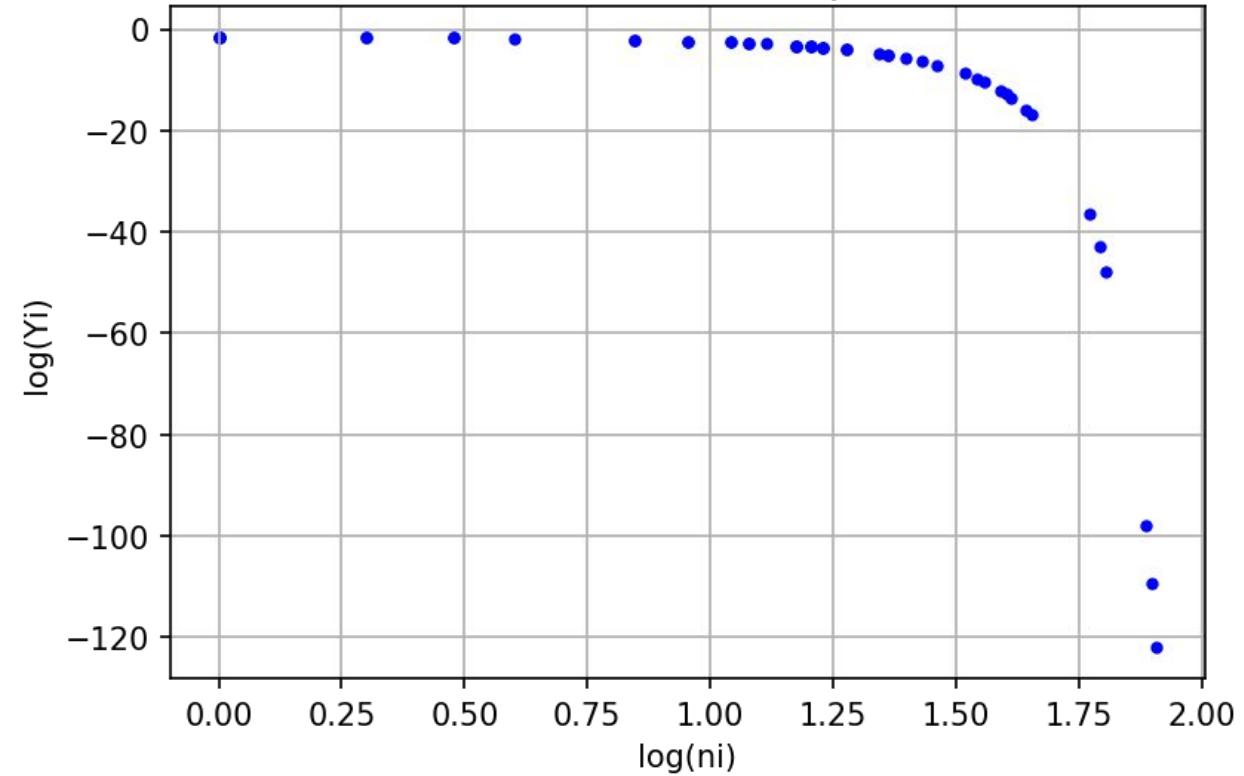


# Interpolação aleatória com 23 pontos - Método dos polinômios

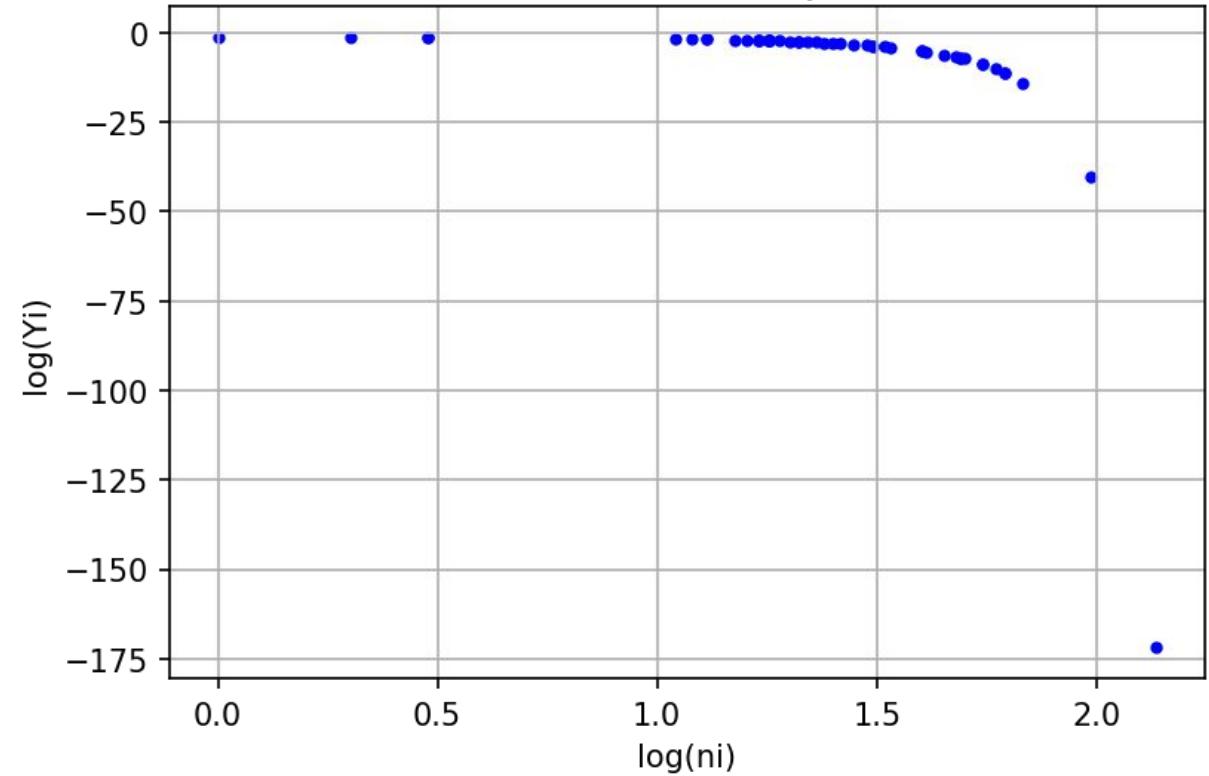


# Assinatura SOC

SOC de casos diários interpolados - Cuba

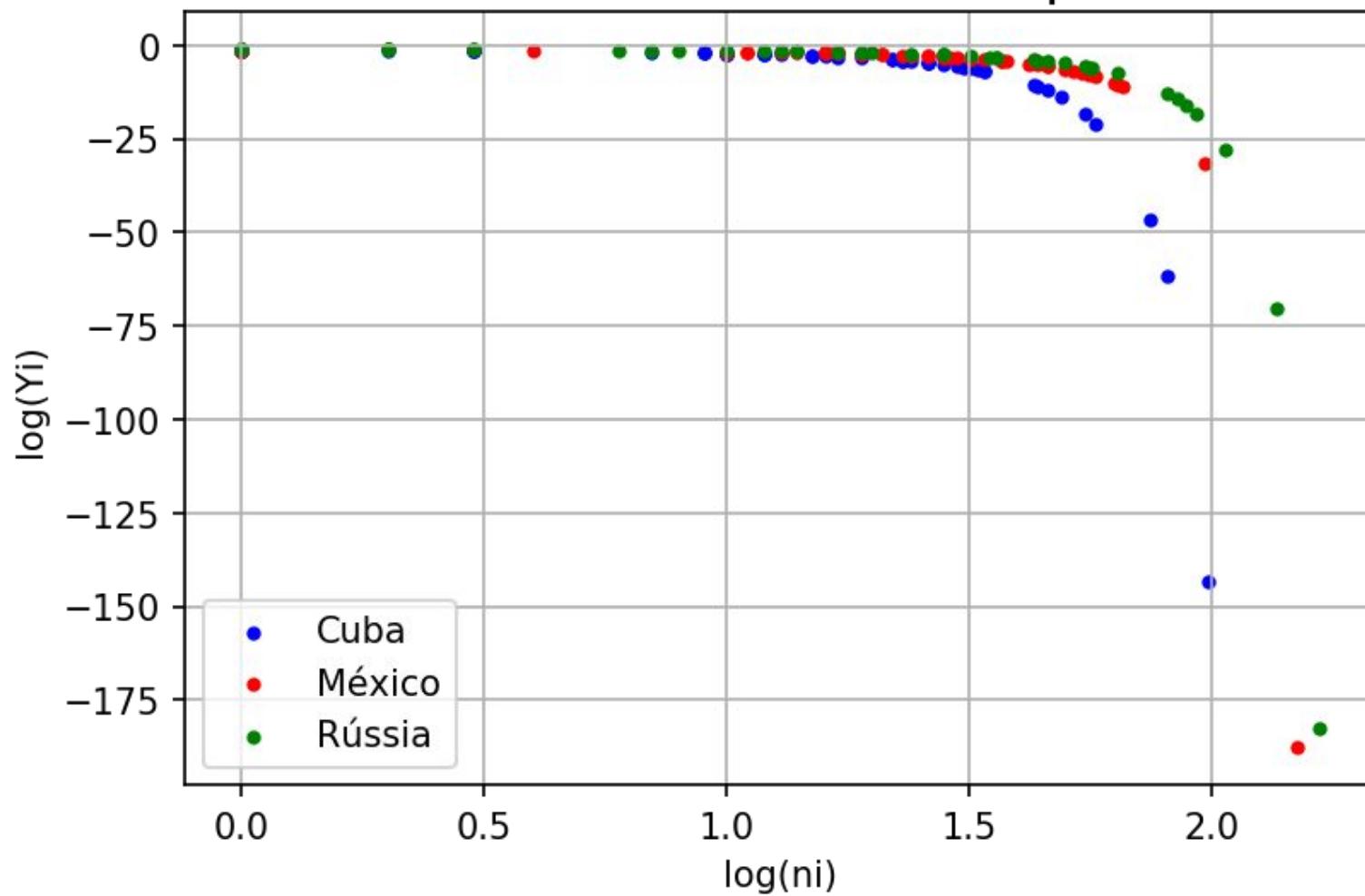


SOC de casos diários interpolados - México



# Assinatura SOC

SOC de casos diários interpolados



# **Previsão de casos diários**

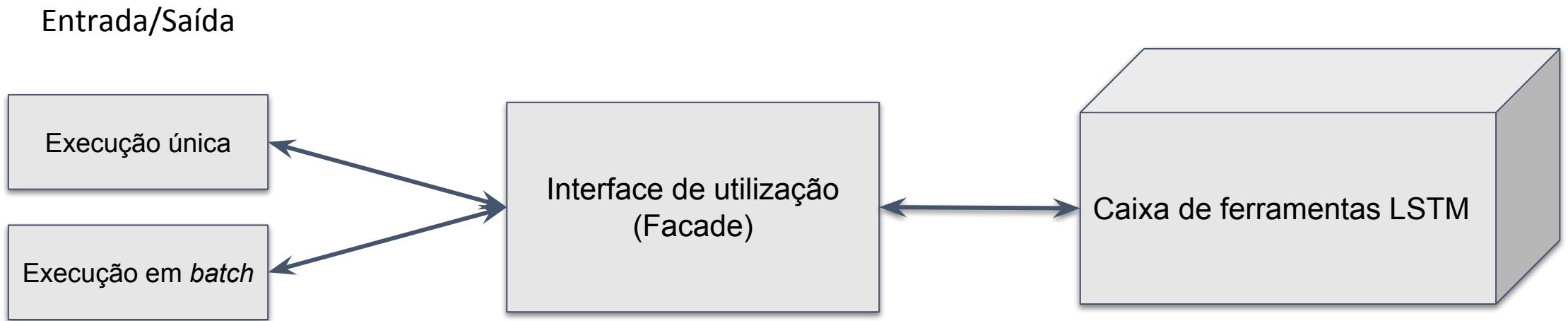
# **Long Short-Term Memory**

# Long Short-Term Memory

- Rede neural recorrente (RNN);
- Trabalha com unidades de memória;
- Dados do passado para gerar comportamentos futuros.

\*Modelo disponibilizado por Luis Ricardo Arantes e Marcos Rodrigues.

# Long Short-Term Memory



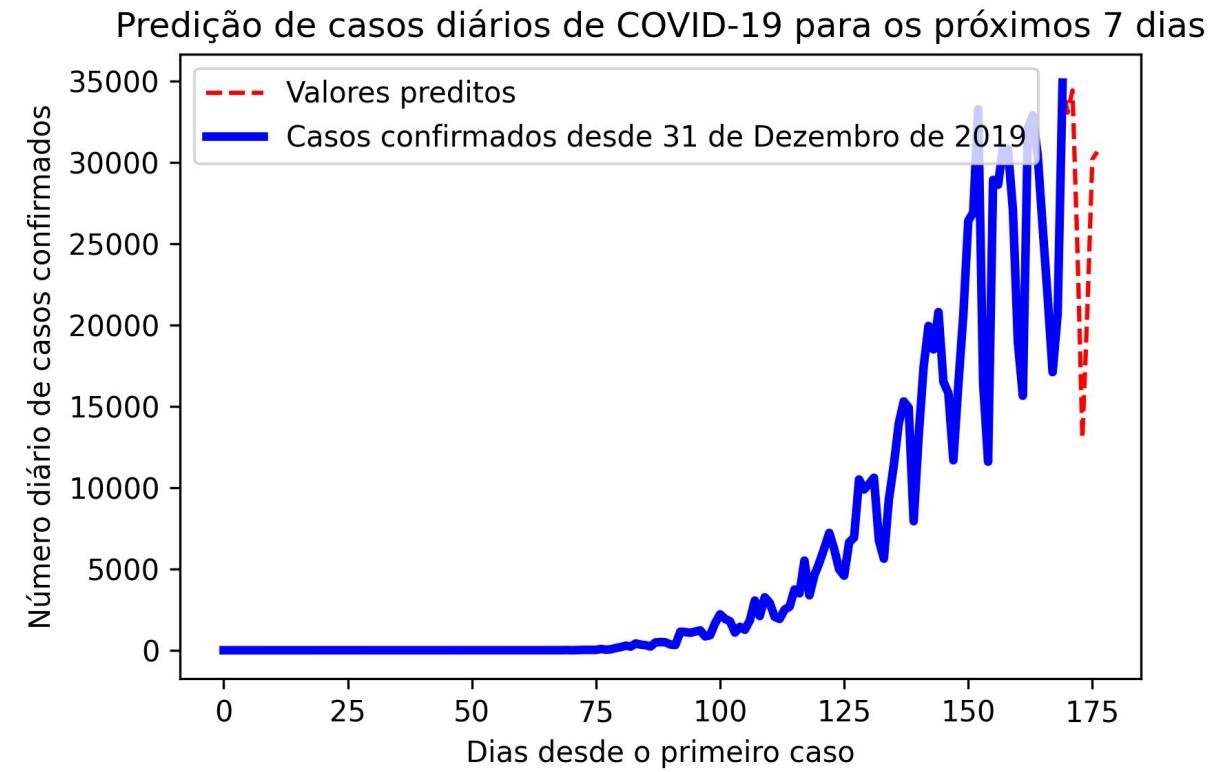
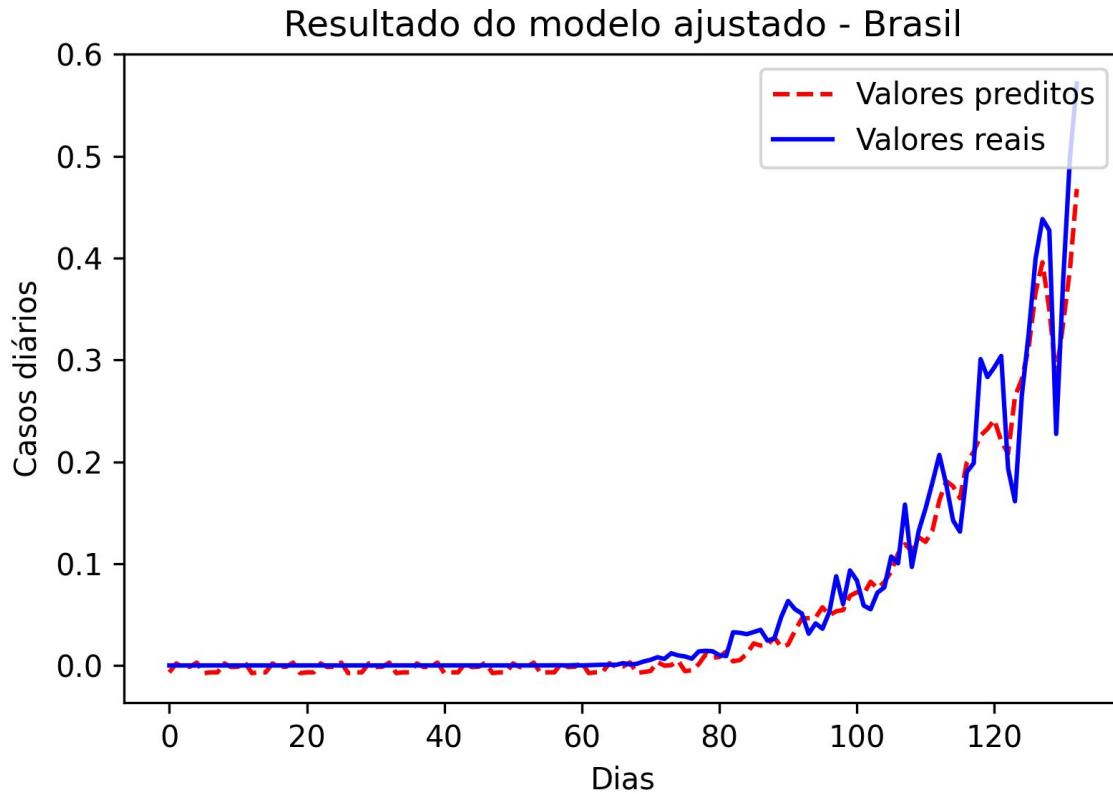
# Long Short-Term Memory

- Geração de boletins semanais sobre os casos de COVID-19;
- Bateria de testes para verificação dos melhores parâmetros;
- O melhor modelo é definido pelo menor RMSE.

# Long Short-Term Memory

## Previsão de casos semanais para o Brasil

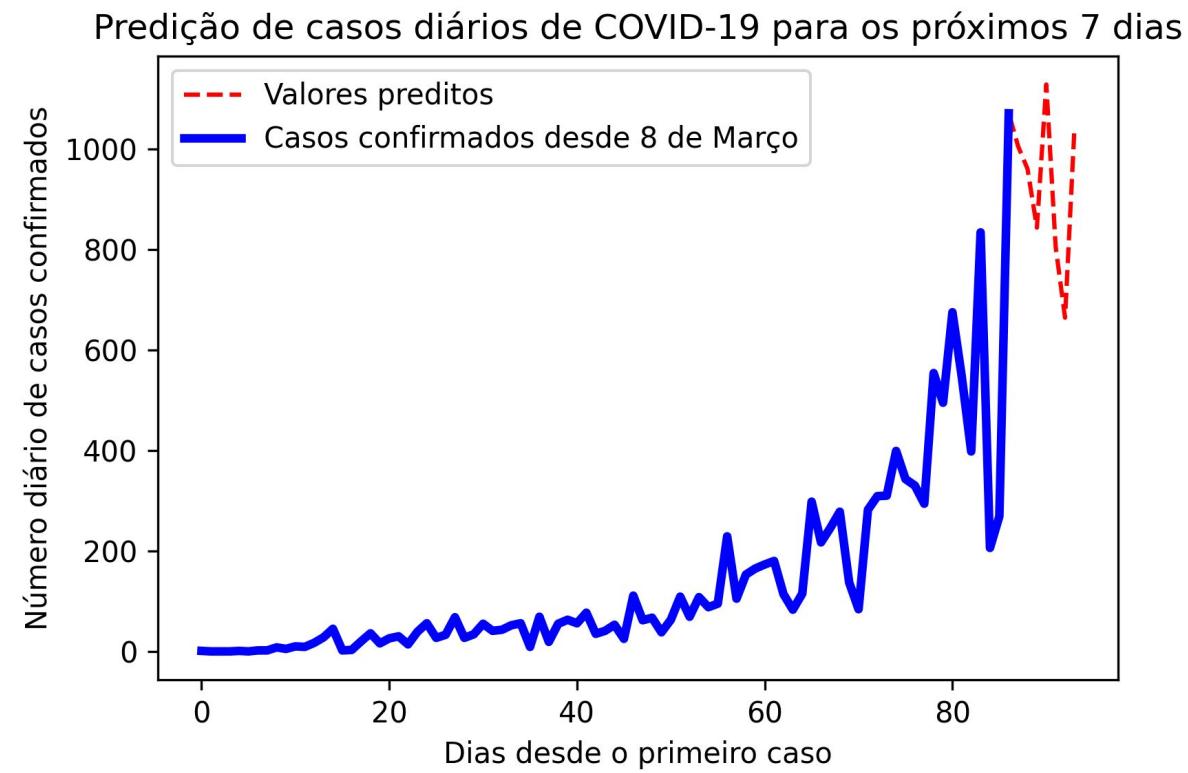
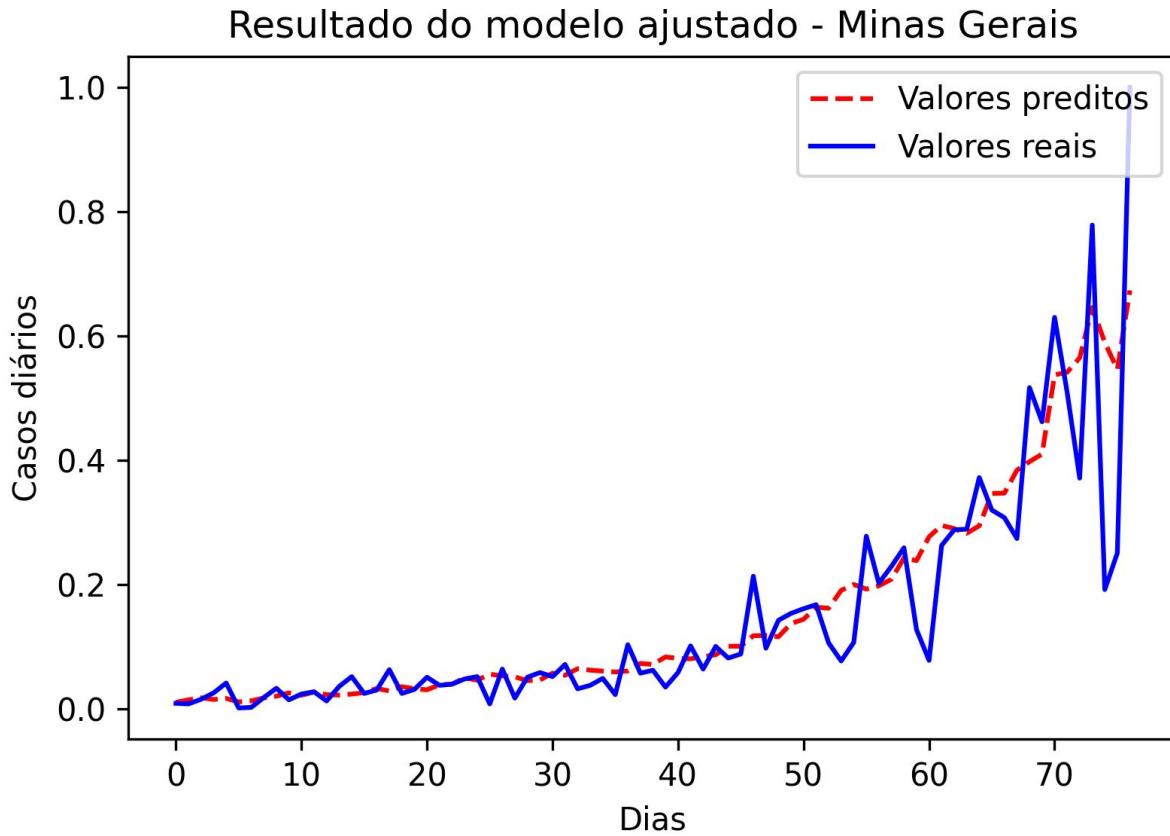
**Entrada:** 10 dias  
**Saída:** 7 dias  
**Treinado com:** 100 epochs



# Long Short-Term Memory

## Previsão de casos semanais para o estado de MG

Entrada: 10 dias  
Saída: 7 dias  
Treinado com: 50 epochs



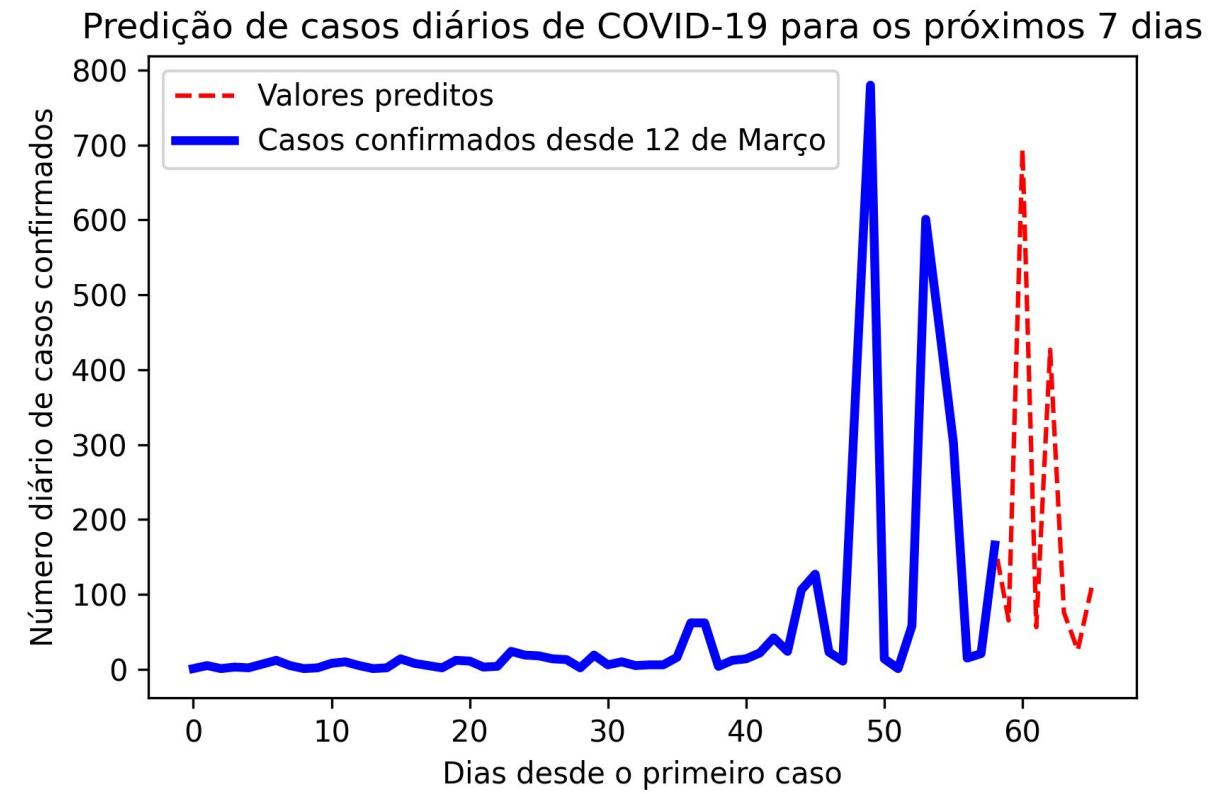
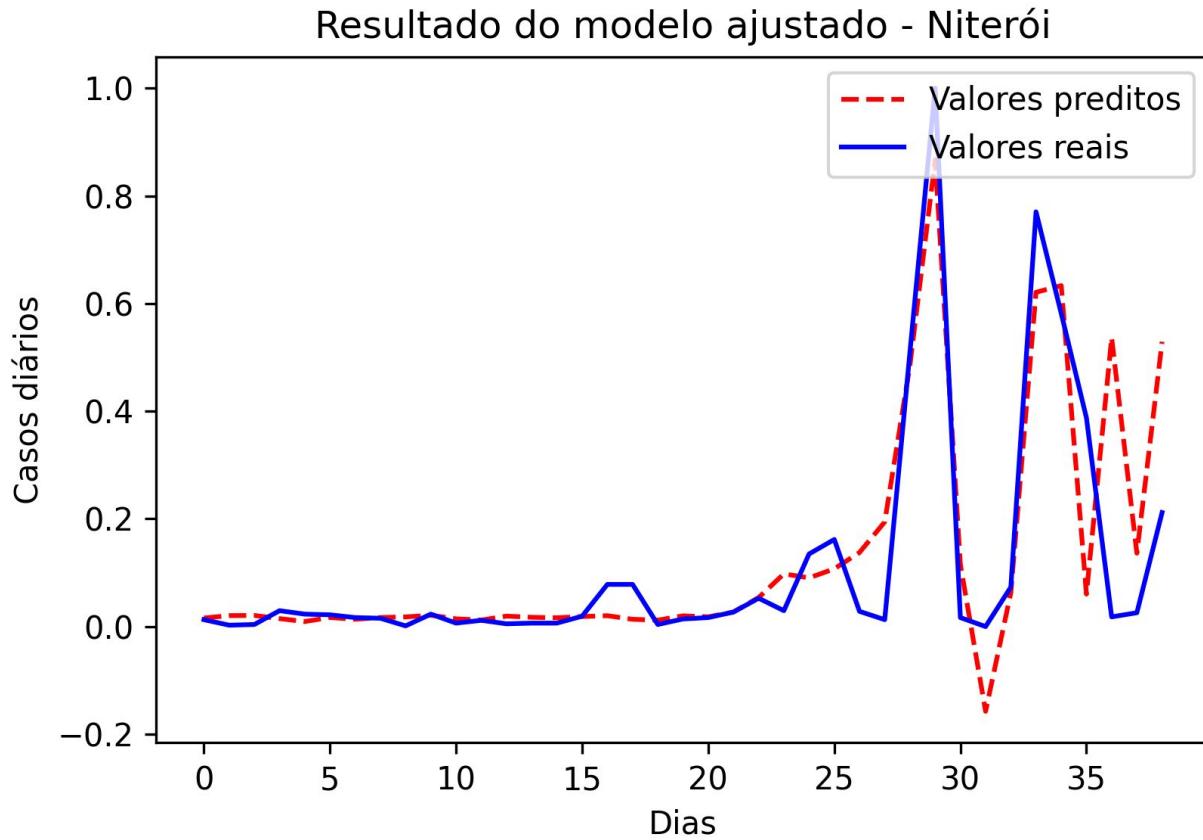
# Long Short-Term Memory

## Previsão de casos semanais para o município de Niterói

Entrada: 20 dias

Saída: 7 dias

Treinado com: 100 epochs



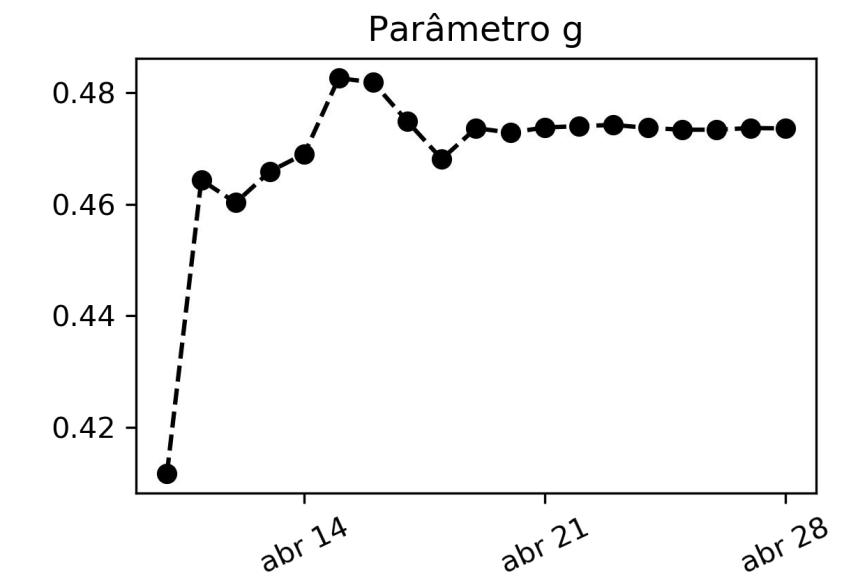
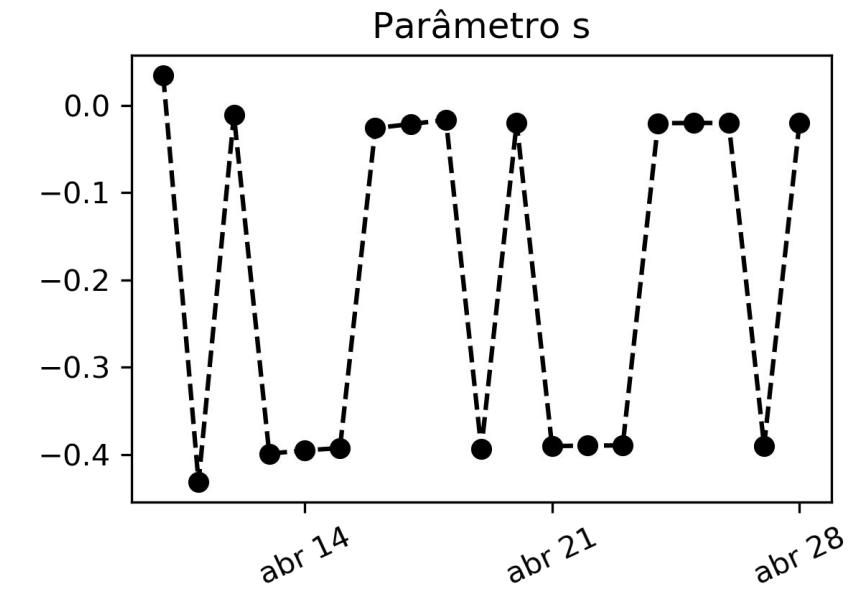
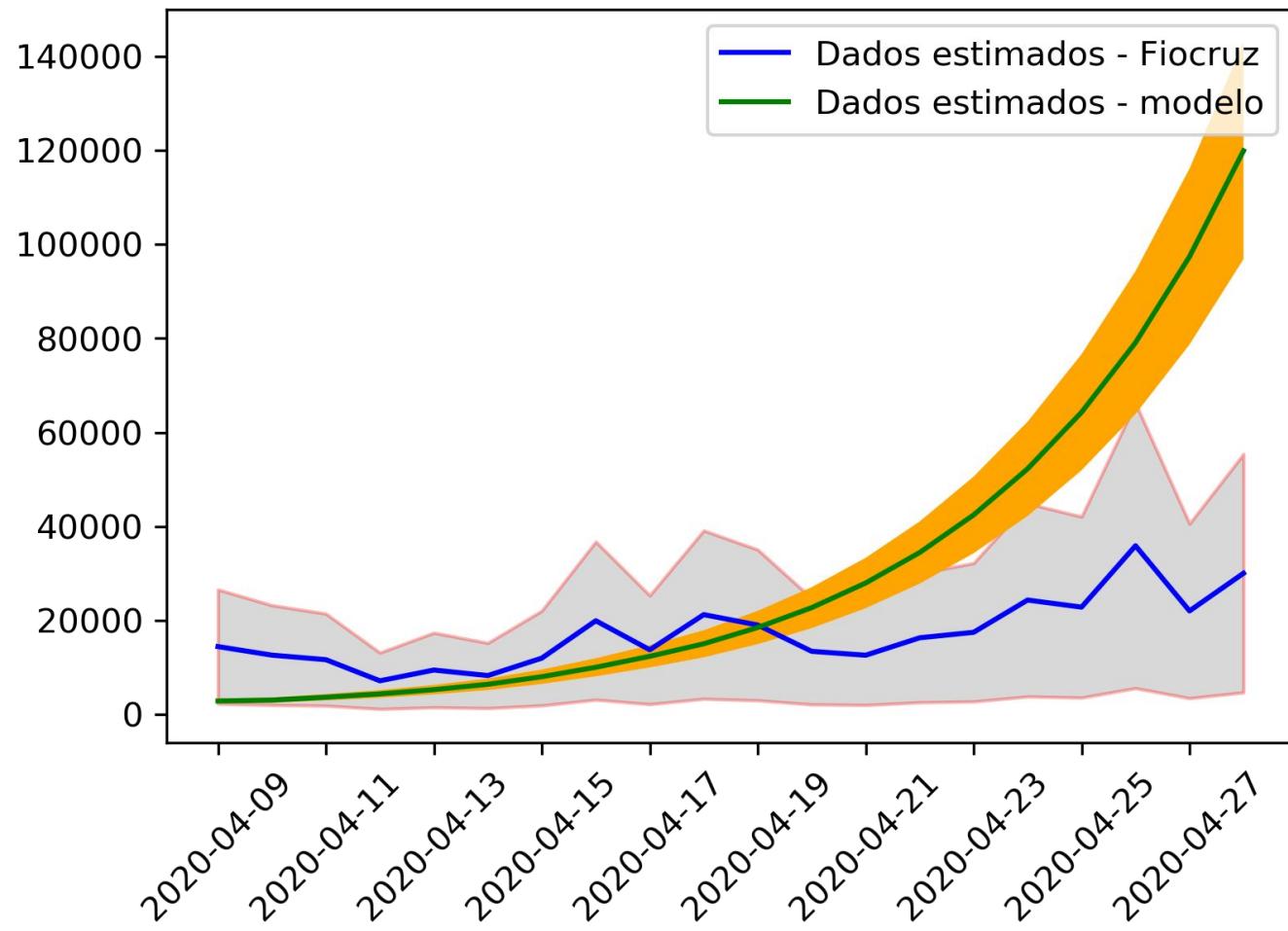
# **IMCSF COVID-19**

- Modelo de cascatas não-homogênea;
- Modelo de transmissão;
- Identificação de sub-notificações nos dados.

\*Modelo desenvolvido e disponibilizado pelo Professor Doutor Reinaldo Rosa

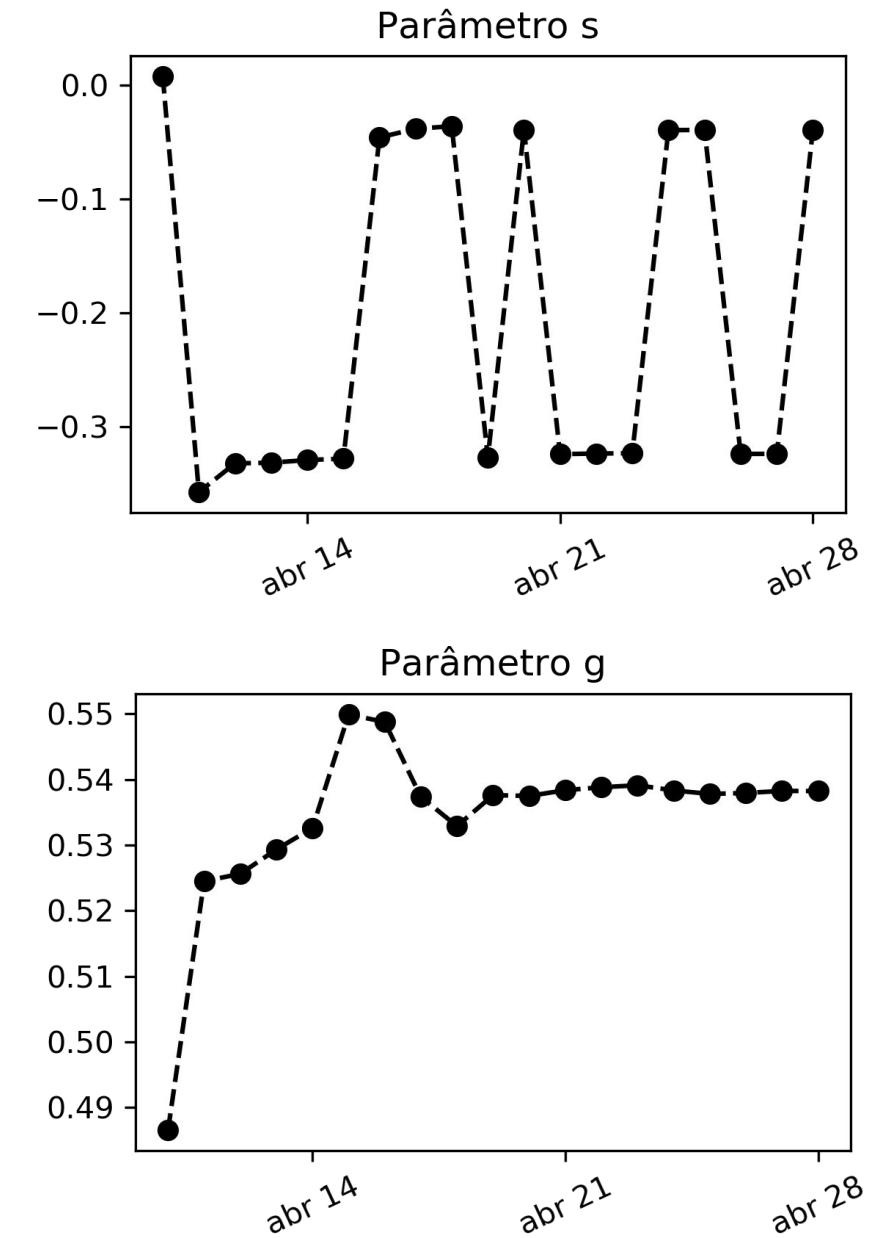
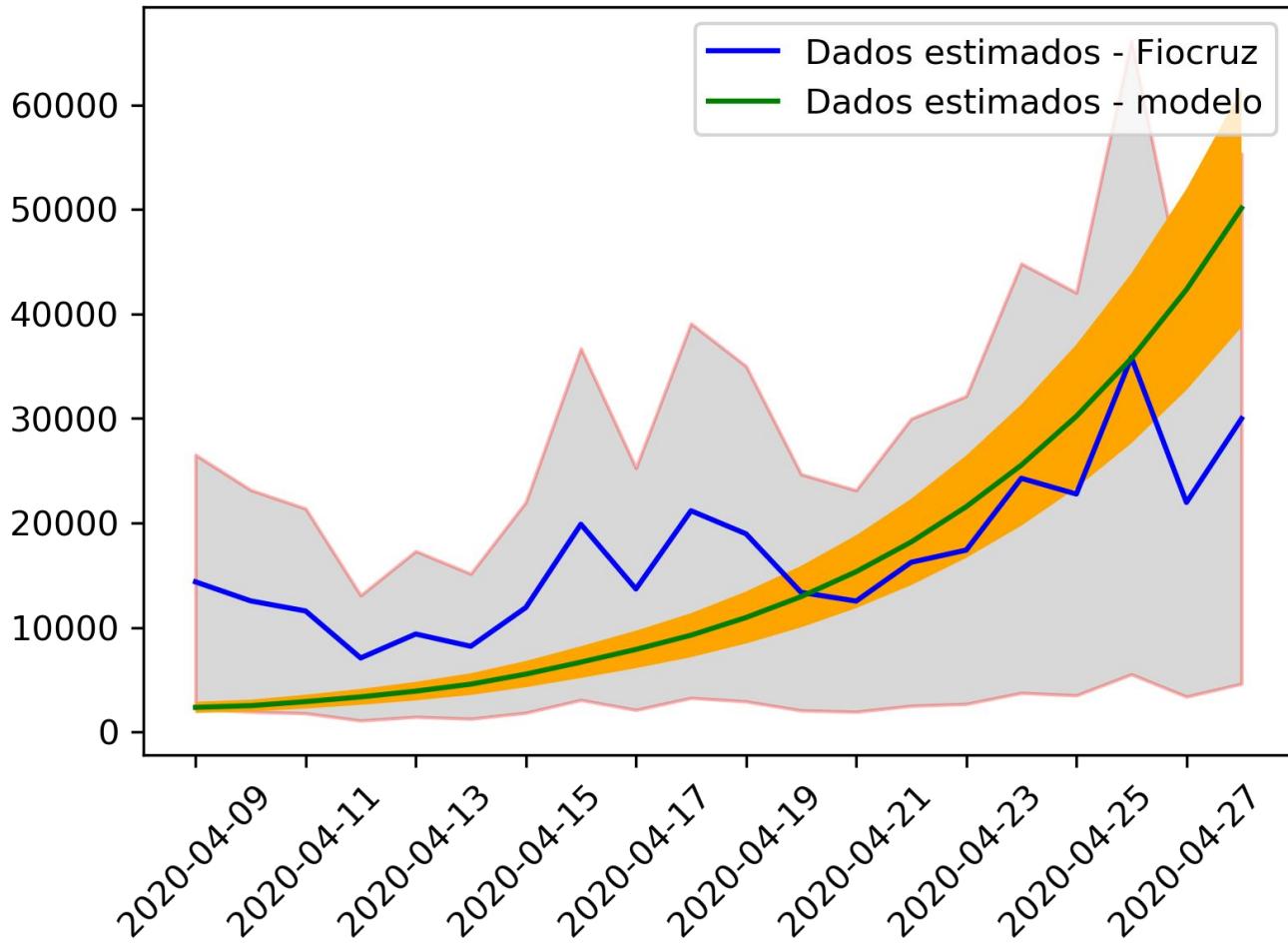
# IMCSF COVID-19

Previsão de casos para o Brasil (08/04/2020 - 28/04/2020)  
Espectro de peso 1



# IMCSF COVID-19

Previsão de casos para o Brasil (08/04/2020 - 28/04/2020)  
Espectro de peso 2



# Conclusão