

COVID19_30_04_estados

April 30, 2020

1 Dados COVID19 em .ipynb 30/04/2020

Estou tentando importar e utilizar os dados do Ministério da Saúde em <https://covid.saude.org.br> para avaliar a situação divulgada oficialmente pelo MS; A url muda diariamente, há uma série de caracteres em primeira vista aleatórios que é atualizada também todo dia... como puxar o arquivo csv automaticamente?

O site não disponibiliza dados de SRAG, apenas uma tabela repetida e desatualizada do COVID19.

Nova tabela com dados do Brasil (população)

Dados de SRAG da fiocruz estão no padrão brasileiro (109,0009) como transformar o arquivo em padrão americano/internacional (109.0009) automaticamente pela URL?

1.1 Gráficos dos Dados

Mostrar infectados e mortes confirmados diariamente

Mostrar infectados e mortes confirmados acumulados

Mostrar dados separados por região e por estado

Mostrar dados de SRAG e comparar com o aumento dos números confirmados de COVID

Dados em semana epidemiológica

1.2 Estimativas (Regressão)

Estimar próximos números (até 7 dias?) utilizando 3 dias anteriores e regressão não linear

Estimar curvas com ajuda dos dados de SRAG inconclusivos (qual a porcentagem de testes incompletos espera-se que serão confirmados)

1.3 Modelo de Contágio

Utilizar dados adquiridos para rodar um modelo epidemiológico que

utiliza valores de reprodução base e evolução da COVID-19, causada pelo vírus SARS-COV-2

1.4 Creating the base of the code

```
[1]: import pandas as pd
import numpy as np
from datetime import datetime, timedelta
from sklearn.metrics import mean_squared_error
from scipy.stats import norm
from scipy.optimize import curve_fit
from scipy.optimize import fsolve
import matplotlib.pyplot as plt
%matplotlib inline

# Databases

url_B = "https://raw.githubusercontent.com/Yannngn/COVID19-04-20/master/brasil.
↪csv"
url_C = "https://mobileapps.saude.gov.br/esus-vepi/files/
↪unAFkcaNDeXajurGB7LChj8SgQYS2ptm/
↪1d2b944e065c7304b2754cc386635e38_Download_COVID19_20200430.csv"

[2]: brasil = pd.read_csv(url_B, sep = ';')
df = pd.read_csv(url_C, sep = ";")

[3]: FMT = '%Y-%m-%d'

date = df['data']
regi = df['regiao']

df['data'] = date.map(lambda x : (datetime.strptime(x, FMT) - datetime.
↪strptime("2020-01-30", FMT)).days)

[4]: df_BR = brasil.loc[range(27), ['UF', 'Unidade da Federação', 'Regiao',
↪'População']]
df_BR = df_BR.sort_values('UF')

df = df.sort_values('estado')

df1 = df.loc[:, ['data', 'casosAcumulados', 'casosNovos', 'obitosAcumulados',
↪'obitosNovos']]
dfs = df.loc[:, ['regiao', 'data', 'casosAcumulados', 'casosNovos',
↪'obitosAcumulados', 'obitosNovos', 'estado']]
df_UF = df.loc[:, ['estado', 'data', 'casosAcumulados', 'casosNovos',
↪'obitosAcumulados', 'obitosNovos']]
```

```
[5]: df1 = df1.groupby("data").sum()

# total of days since 2020-01-30 - x
days = list(df1.index)
x = days
```

1.5 Exponential Model

```
[6]: # exponential model for the evolution of the covid 19 in Brazil

def exponential_model (x, a, b, c) :
    return a * np.exp (b * (x - c))

def exp_param (R, X) :
    a = np.exp(X[2] * (((np.log(R[2]))/X[2]) - ((np.log(R[1]) - np.log(R[0]))/
    ↪(X[1] - X[0]))))
    b = (np.log(R[2])/X[2]) - (1/X[2])*(((np.log(R[0])/X[0]) - (np.log(R[1])/
    ↪X[1])))/((X[0]**(-1) - (X[1])**(-1)))
    return [a, b]

def exp_est (R, X) :
    np.random.seed(max(x))
    a = np.exp(X[2] * (((np.log(R[2]))/X[2]) - ((np.log(R[1]) - np.log(R[0]))/
    ↪(X[1] - X[0]))))
    b = (np.log(R[2])/X[2]) - (1/X[2])*(((np.log(R[0])/X[0]) - (np.log(R[1])/
    ↪X[1])))/((X[0]**(-1) - (X[1])**(-1)))
    return a * np.exp(b * (max(x) + 1)) + norm.rvs(size = 1, scale = 0.05 *
    ↪(max(x) + 1))
```

1.6 Logistic Model

```
[7]: # logistic model for the evolution of the covid 19 in Brazil

def logistic_model (x, a, b, c) :
    return c / (1 + np.exp ( - (x - b) / a))

a0 = 2.7

lb = [df1.index.size + 31, df1.index.size + 86, df1.index.size + 45, df1.index.
    ↪size + 45]

pBrasil = brasil.at[brasil['UF'].eq('BR').idxmax(), 'População']
lc = [0.2 * pBrasil, 0.001 * pBrasil, 0.008 * pBrasil, 0.00016 * pBrasil]
```

1.7 Data for every Brazilian State

1.7.1 Population

Source: IBGE

1.7.2 Last Confimed Numbers

Source: covid.saude.gov.br

1.7.3 Incidence of Cases and Deaths

Number of cases or deaths for million citizens

1.7.4 Speed of the Contamination

How many days does the COVID takes to double in every state

```
[8]: casosInc = []
    casosDia = []
    casosCres = []
    obitosInc = []
    obitosDia = []
    obitosCres = []
    dfs = dfs.sort_values('estado')
    estados = dfs["estado"].unique()
    x = days

    for e in estados :
        pop = brasil.at[brasil['UF'].eq(e).idxmax(), 'População']
        uf = brasil.at[brasil['UF'].eq(e).idxmax(), "Unidade da Federação"]

        # Values for daily cases
        df_e = df_UF.loc[df_UF.iloc[:,0] == e, :].iloc[:, [1, 2]].
        ↪sort_values("data")
        y = list(df_e.iloc[:, 1])
        casosDia.append(round(y[max(days)], 2))
        casosInc.append(round((10 ** 6) * y[max(days)] / pop, 2))

        exp_p = exp_param([y[-5], y[-3], y[-1]], [x[-5], x[-3], x[-1]])
        exponential_fit = curve_fit(exponential_model, x, y, p0 = [exp_p[0] * np.
        ↪exp(-1), exp_p[1], 1])
        casosCres.append(round(np.log(2) / exponential_fit[0][1], 1))

        # Values for daily deaths
        df_e = df_UF.loc[df_UF.iloc[:,0] == e, :].iloc[:, [1, 4]].
        ↪sort_values("data")
        y = list(df_e.iloc[:, 1])
        obitosInc.append(round((10 ** 6) * y[max(days)] / pop, 2))
```

```

obitosDia.append(round(y[max(days)], 2))

exp_p = exp_param([y[-5], y[-3], y[-1]], [x[-5], x[-3], x[-1]])
exponential_fit = curve_fit(exponential_model, x, y, p0 = [exp_p[0] * np.
→exp(-1), exp_p[1], 1])
obitosCres.append(round(np.log(2) / exponential_fit[0][1], 1))

df_BR['Casos'] = casosDia
df_BR['Incidência de Casos'] = casosInc
df_BR['Dias para dobrar (casos)'] = casosCres
df_BR['Mortes'] = obitosDia
df_BR['Incidência de Mortes'] = obitosInc
df_BR['Dias para dobrar (obitos)'] = obitosCres
df_BR = df_BR.sort_values('Incidência de Casos', ascending = False)
df_BR.reset_index(drop = True, inplace = True)

df_BR

```

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.

```

[8]:
UF  Unidade da Federação      Regiao  População  Casos  \
0   AP                    Amapá        Norte    842914   1080
1   AM                    Amazonas      Norte   4240210   5254
2   RR                    Roraima       Norte    546891    519
3   CE                    Ceará         Nordeste  9178363   7606
4   PE                    Pernambuco    Nordeste  9650604   6876
5   SP                    São Paulo     Sudeste  46064928  28698
6   ES                    Espírito Santo Sudeste   4138657   2465
7   RJ                    Rio de Janeiro Sudeste  16946541  9453
8   AC                    Acre          Norte    866811    404
9   MA                    Maranhão     Nordeste  7121156   3190
10  DF                    Distrito Federal Centro-Oeste 3223048   1356
11  PA                    Pará          Norte    8628901   2876
12  RN                    Rio Grande do Norte Nordeste 3598288   1177
13  AL                    Alagoas       Nordeste 3419689   1044
14  SC                    Santa Catarina Sul      7266193   2085
15  RO                    Rondônia      Norte    1857992    502
16  PB                    Paraíba       Nordeste 4097859    814
17  SE                    Sergipe        Nordeste 2352207    447
18  BA                    Bahia         Nordeste 15522855  2851
19  PI                    Piauí         Nordeste 3233891    513
20  RS                    Rio Grande do Sul Sul      11416895  1466
21  PR                    Paraná        Sul      11538518  1407
22  GO                    Goiás      Centro-Oeste 7017496    781
23  MS                    Mato Grosso do Sul Centro-Oeste 2800704    255

```

24	MT	Mato Grosso	Centro-Oeste	3455092	297
25	TO	Tocantins	Norte	1599316	137
26	MG	Minas Gerais	Sudeste	21451356	1827

	Incidência de Casos	Dias para dobrar (casos)	Mortes \
0	1281.27	7.0	34
1	1239.09	7.7	425
2	949.00	7.2	7
3	828.69	7.9	482
4	712.49	6.5	565
5	622.99	9.6	2375
6	595.60	7.4	83
7	557.81	9.3	854
8	466.08	8.2	19
9	447.96	6.5	184
10	420.72	12.9	30
11	333.30	5.5	208
12	327.10	9.2	56
13	305.29	4.1	47
14	286.95	10.4	46
15	270.18	5.6	16
16	198.64	5.9	62
17	190.03	4.5	12
18	183.66	8.4	104
19	158.63	5.7	24
20	128.41	12.2	51
21	121.94	12.1	83
22	111.29	9.4	29
23	91.05	11.4	9
24	85.96	10.5	11
25	85.66	6.9	3
26	85.17	11.5	82

	Incidência de Mortes	Dias para dobrar (obitos)
0	40.34	6.1
1	100.23	6.9
2	12.80	9.4
3	52.51	6.9
4	58.55	6.7
5	51.56	8.3
6	20.05	6.5
7	50.39	7.8
8	21.92	6.2
9	25.84	5.6
10	9.31	11.4
11	24.11	4.2
12	15.56	8.2

13	13.74	5.5
14	6.33	10.6
15	8.61	5.2
16	15.13	7.6
17	5.10	9.2
18	6.70	7.4
19	7.42	9.0
20	4.47	8.9
21	7.19	9.2
22	4.13	9.8
23	3.21	9.1
24	3.18	8.0
25	1.88	5.8
26	3.82	8.2

1.8 Plots

1.8.1 Daily Confirmed Number by Brazilian regions

Norte
Centro-Oeste
Nordeste
Sudeste
Sul

New Confirmed Cases
New Confirmed Deaths

```
[9]: plt.rc('font', size=14)
plt.rcParams['figure.figsize'] = [14, 7]

ttl = ["New Confirmed Cases", "New Confirmed Deaths"]
dfs = dfs.sort_values("data")
regioes = dfs["regiao"].unique()
estados = list(dfs["estado"].unique())

# Real data
# Brazilian areas separated
for r in regioes :
    for f in range(2) :
        for e in estados :
            plt.grid(True)

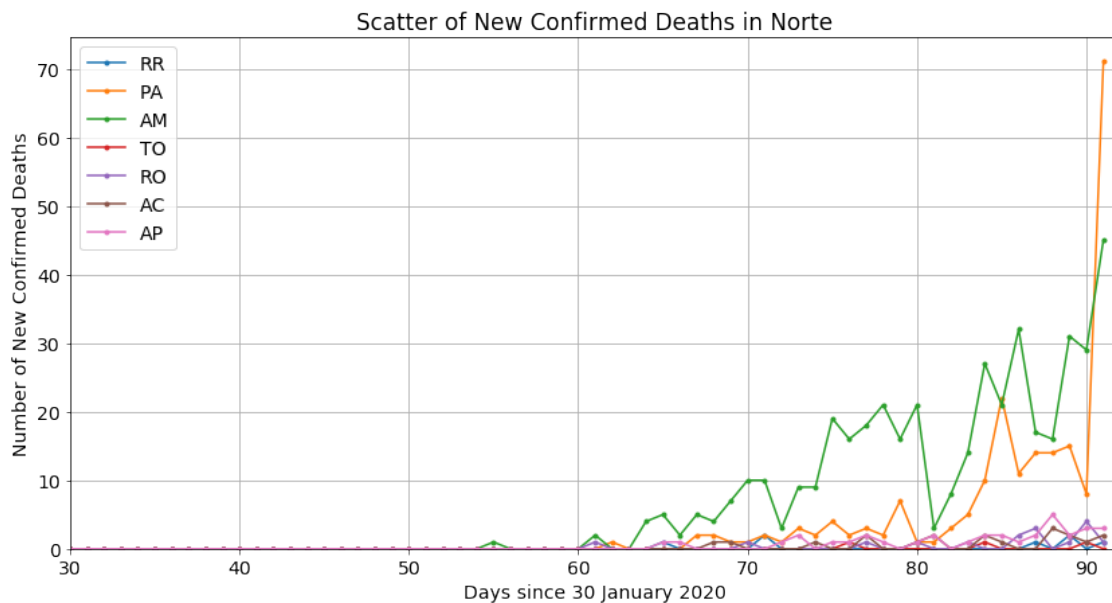
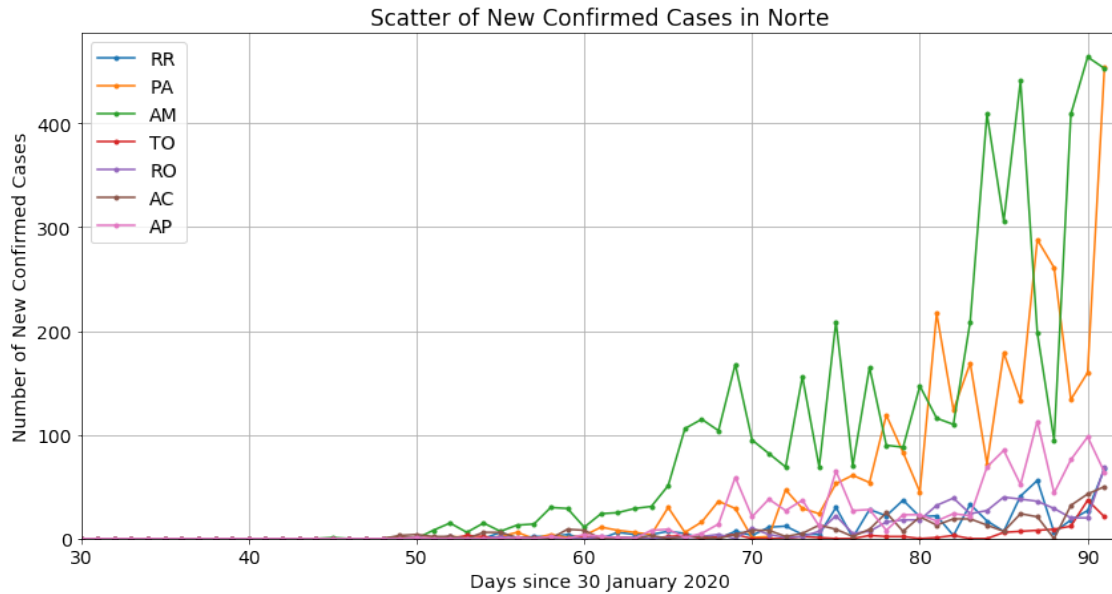
            uf = brasil.at[brasil['UF'].eq(e).idxmax(), "Regiao"]
            yx = max(dfs[dfs['regiao'] == r].iloc[:, 3 + (2 * f)])

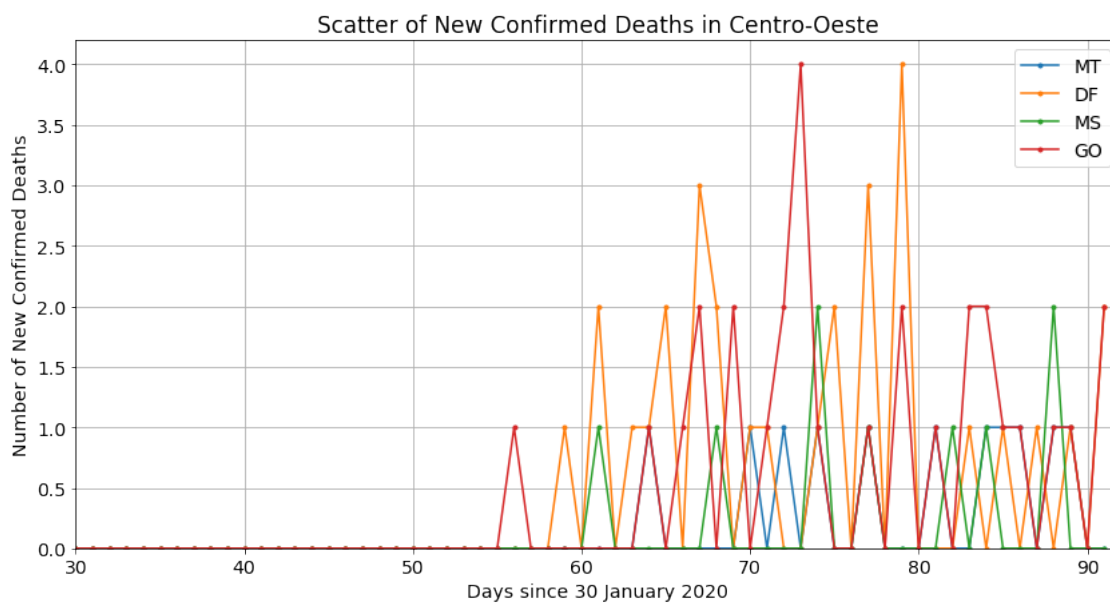
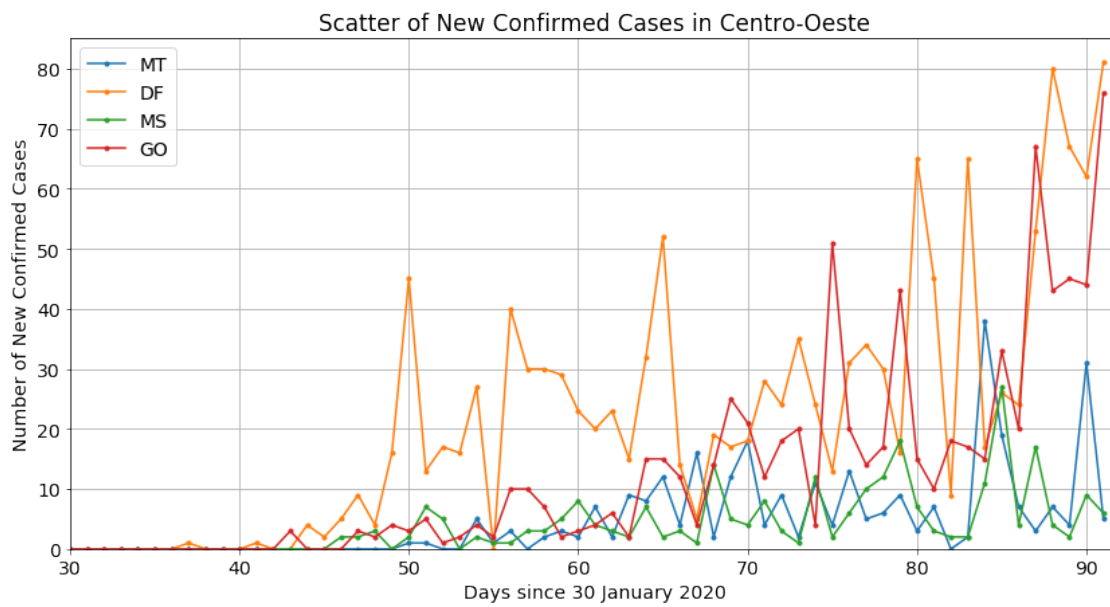
            if uf == r :
                xs = days
```

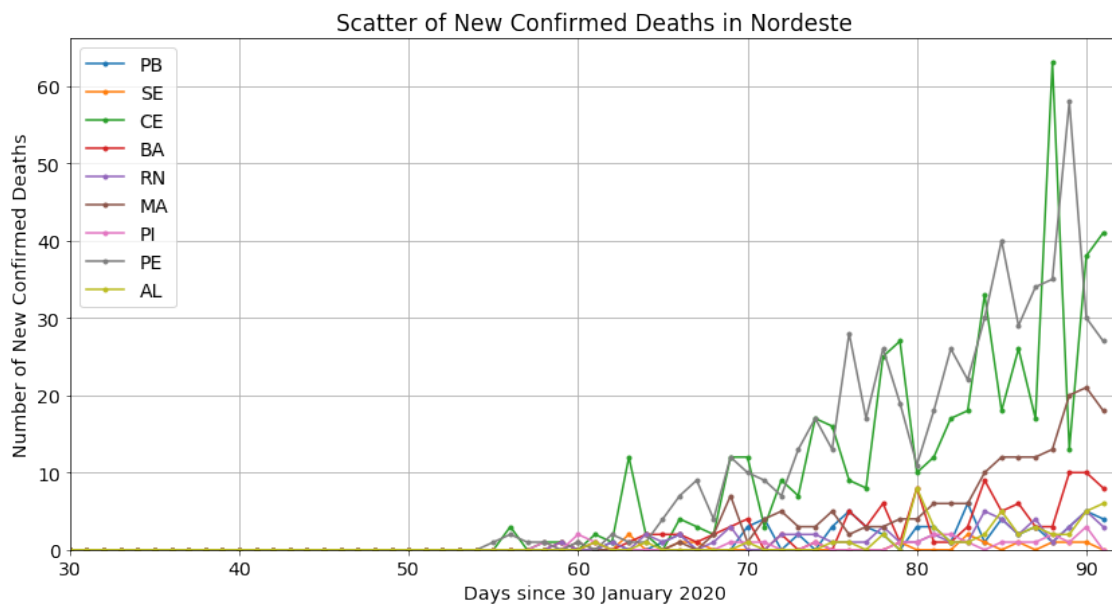
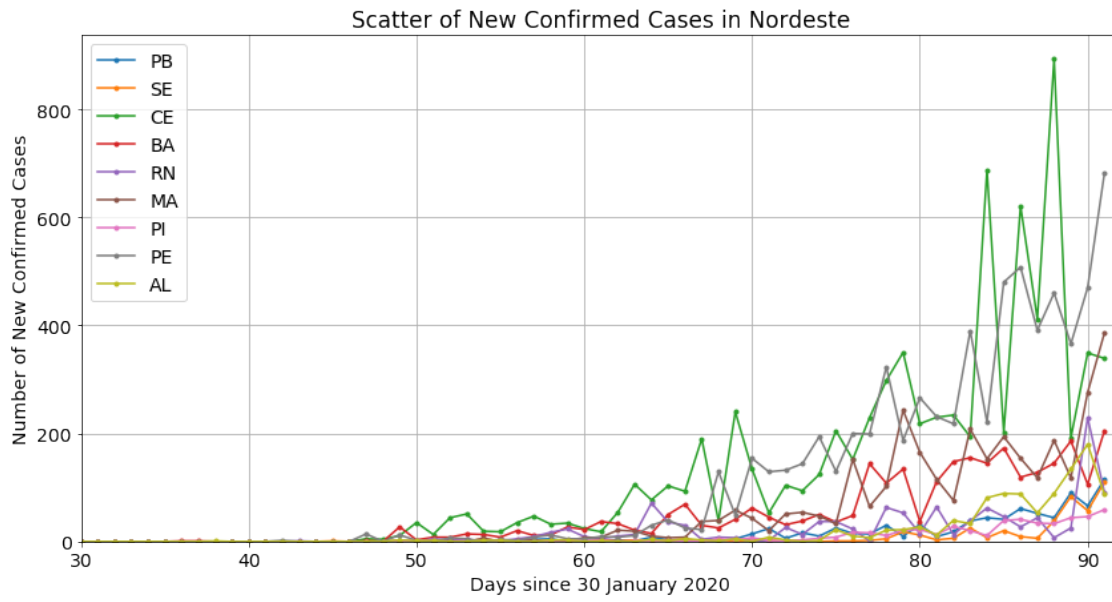
```

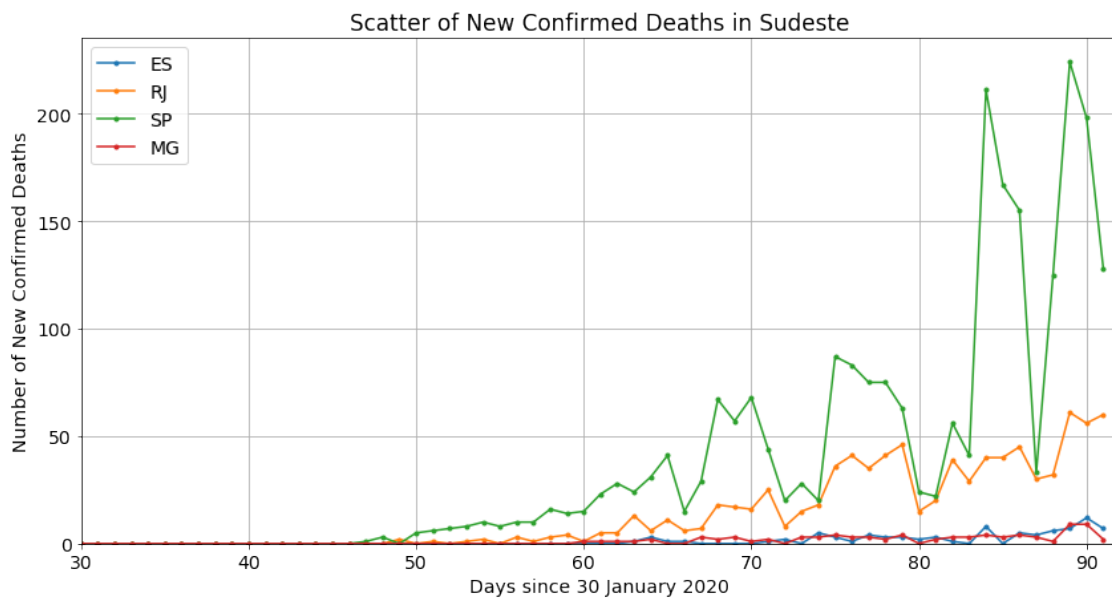
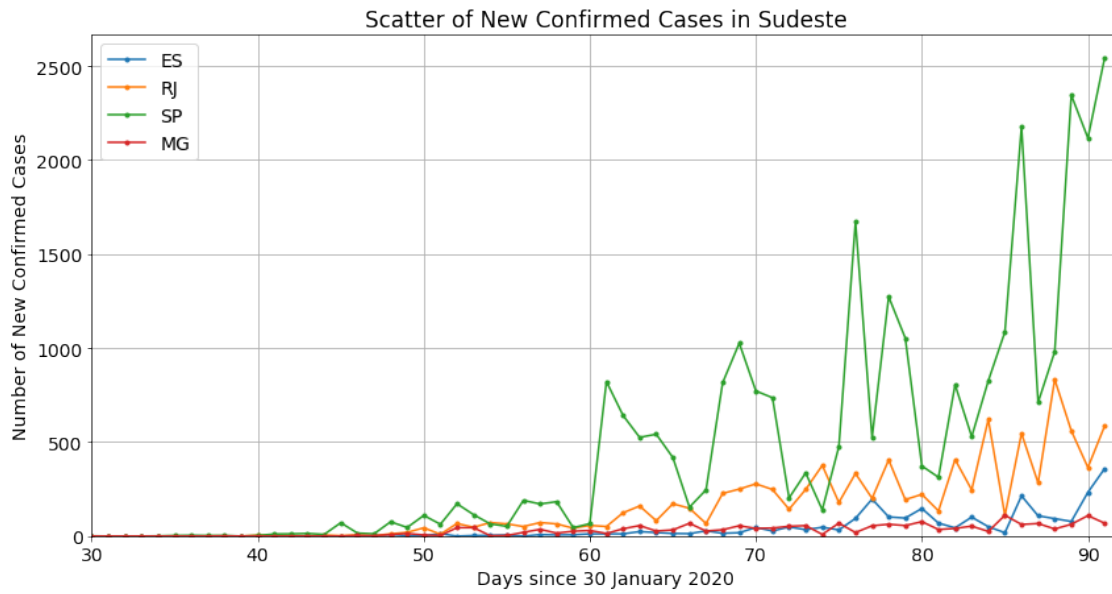
ys = list(dfs[dfs['estado'] == e].iloc[:, 3 + (2 * f)])
plt.plot(xs, ys, marker = '.', label = e)
plt.title("Scatter of " + ttl[f] + " in " + r)
plt.legend()
plt.xlabel("Days since 30 January 2020")
plt.ylabel("Number of " + ttl[f])
plt.xlim(30, max(xs) * 1.01)
plt.ylim(0, yx * 1.05)
plt.show()

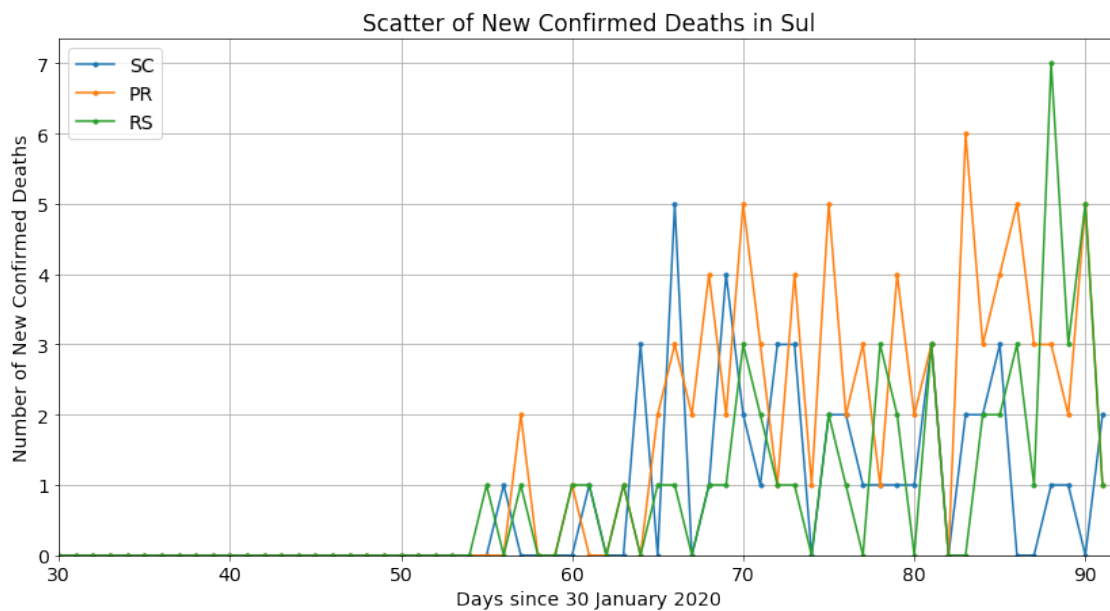
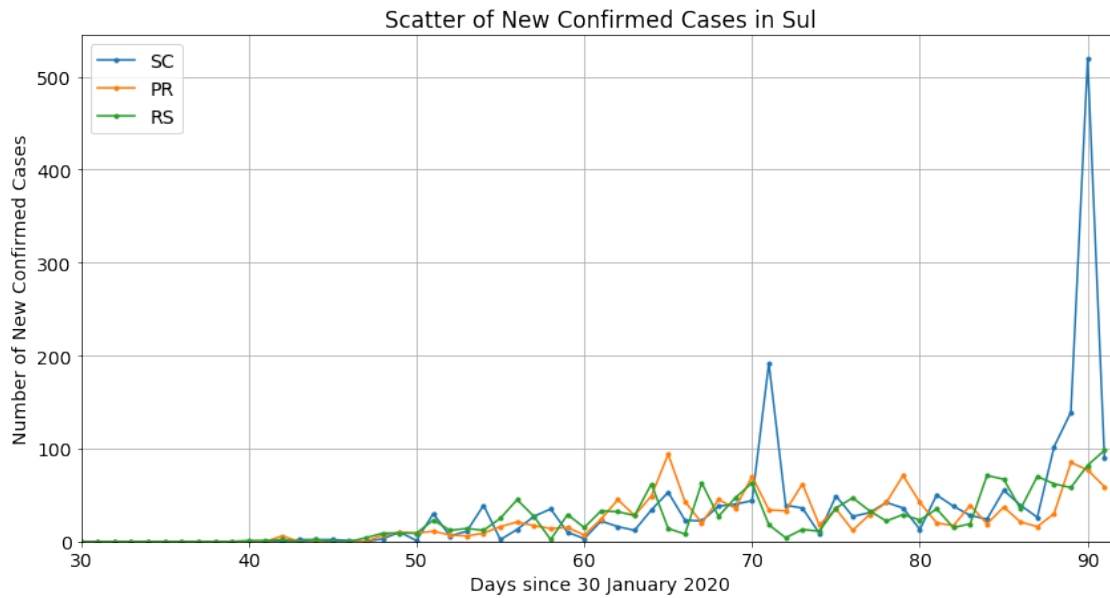
```











1.9 Casos Totais por estado

Casos confirmados x tempo

Dados reais, estimativa exponencial e em estados com mais de 1000 casos estimativa logistica

Os dados são subnotificados e possuem baixa confiabilidade

```

[10]: plt.rc('font', size = 14)
plt.rcParams['figure.figsize'] = [14, 7]

ttl = ["Confirmed Cases", "Confirmed Deaths"]
dfs = dfs.sort_values('estado')
estados = dfs["estado"].unique()
cor = ["#1d2b49", "#235174"]

for e in estados :
    df_e = df_UF.loc[df_UF.iloc[:, 0] == e, :].iloc[:, [2, 4, 1]].
    ↪sort_values("data")
    pop = brasil.at[brasil['UF'].eq(e).idxmax(), 'População']
    est = brasil.at[brasil['UF'].eq(e).idxmax(), 'Unidade da Federação']

    for f in range(2) :
        plt.grid(True)
        x = days

        # Real Numbers
        y = list(df_e.iloc[:, f])
        plt.plot(x, y, marker = '.', label = "Real numbers", color = cor[f])

        # Lines

        if max(y) > 1000 :
            logistic_fit = curve_fit(logistic_model, x, y, p0 = [a0, lb[f * 2],
            ↪0.1 * pop])
            logistic_sol = int(fsolve(lambda x : logistic_model(x,
            ↪logistic_fit[0][0], logistic_fit[0][1],
            ↪logistic_fit[0][2]) - int(logistic_fit[0][2]),
            ↪logistic_fit[0][1]))

            pred_x = list(range(max(x), logistic_sol))

            # Predicted logistic curve
            plt.plot(x + pred_x, [logistic_model(i, logistic_fit[0][0],
            ↪logistic_fit[0][1], logistic_fit[0][2])
            ↪for i in x + pred_x], label = "Logistic
            ↪model", color = "#c2cdd8")

            plt.xlim(0, max(pred_x) * 1.05)
            plt.ylim(0, logistic_fit[0][2] * 1.05)
        else :

            pred_x = list(range(max(x), max(x) + 15))

```

```

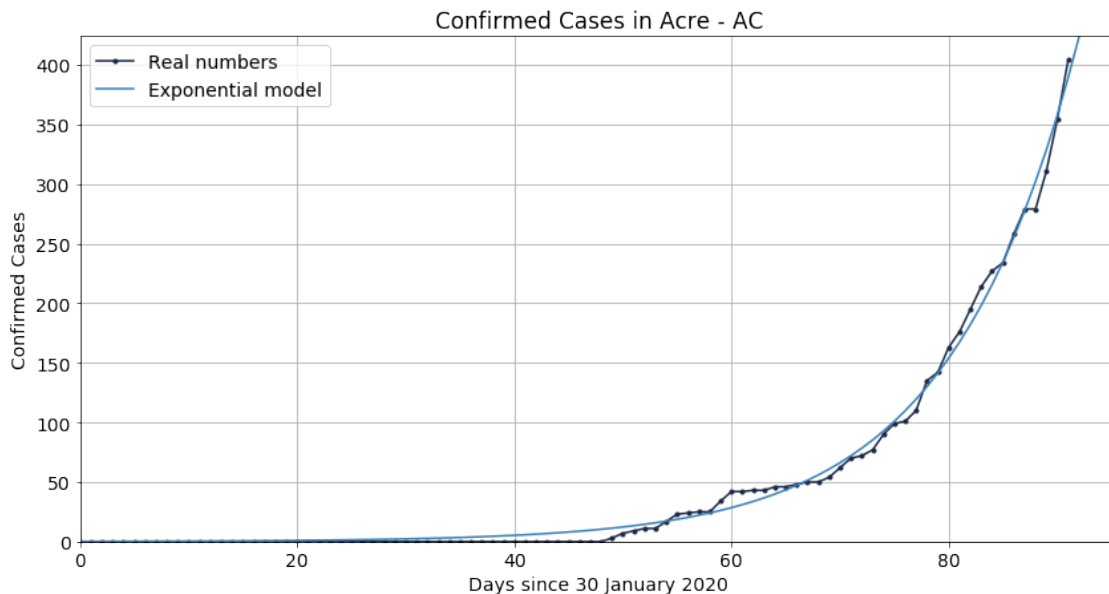
plt.xlim(0, max(x) * 1.05)
plt.ylim(0, max(y) * 1.05)

exp_p = exp_param([y[-5], y[-3], y[-1]], [x[-5], x[-3], x[-1]])
exponential_fit = curve_fit(exponential_model, x, y, p0 = [exp_p[0] *
↪np.exp(-1), exp_p[1], 1])

plt.plot(x + pred_x, [exponential_model(i, exponential_fit[0][0],
↪exponential_fit[0][1], exponential_fit[0][2])
                        for i in x + pred_x], label = "Exponential
↪model", color = "#3881b8")

plt.title(ttl[f] + " in " + est + " - " + e)
plt.legend()
plt.xlabel("Days since 30 January 2020")
plt.ylabel(ttl[f])
plt.show()
print("Next estimated number of confirmed cases in " + est + " is: " +
      str(round(exponential_model(max(x) + 1, exponential_fit[0][0],
↪exponential_fit[0][1], exponential_fit[0][2]), 3))
      + " or + " +
      str(round(exponential_model(max(x) + 1, exponential_fit[0][0],
↪exponential_fit[0][1], exponential_fit[0][2]) - max(y), 3)) )

```

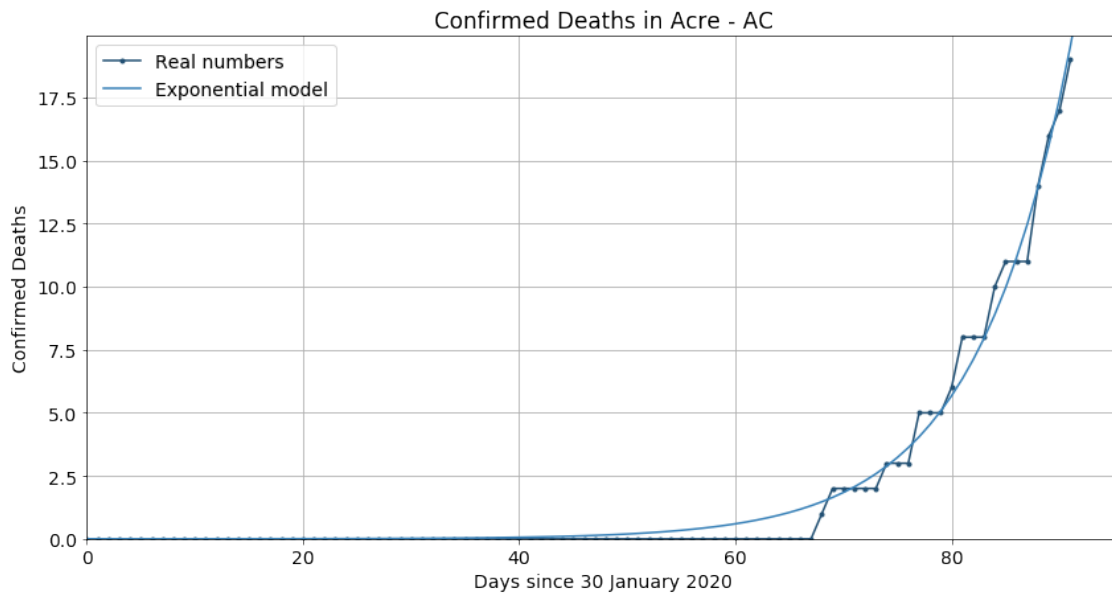


Next estimated number of confirmed cases in Acre is: 423.412 or + 19.412

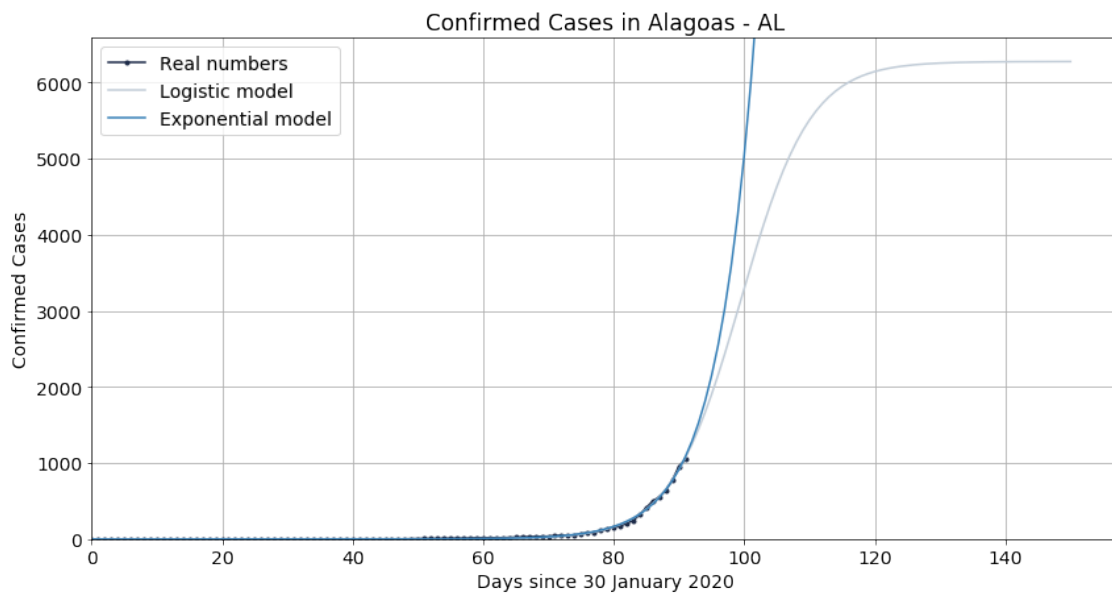
C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

RuntimeWarning: overflow encountered in exp

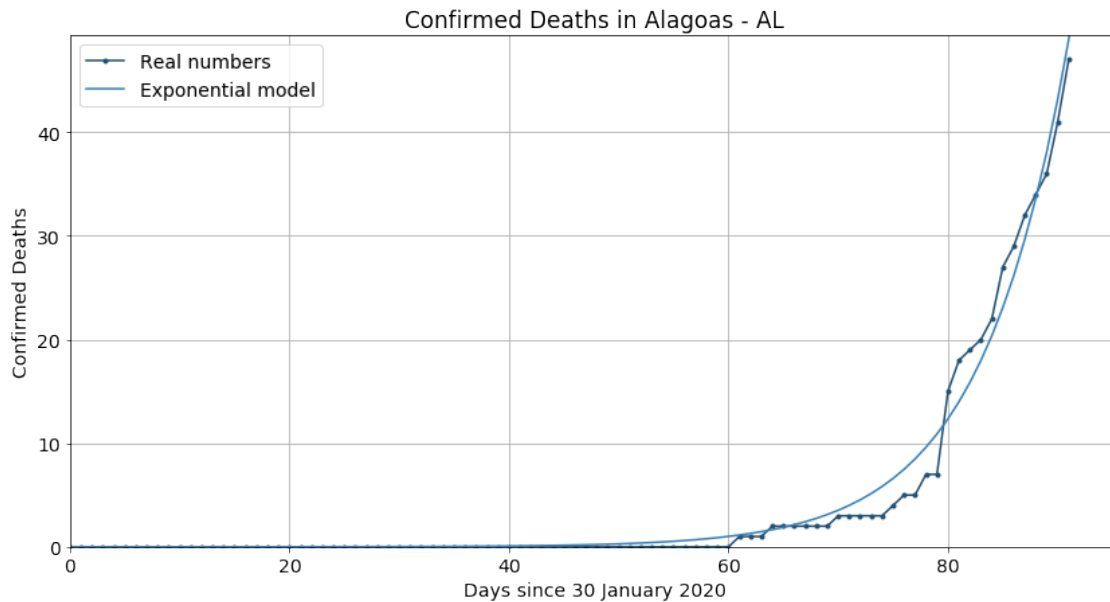
after removing the cwd from sys.path.



Next estimated number of confirmed cases in Acre is: 21.868 or + 2.868



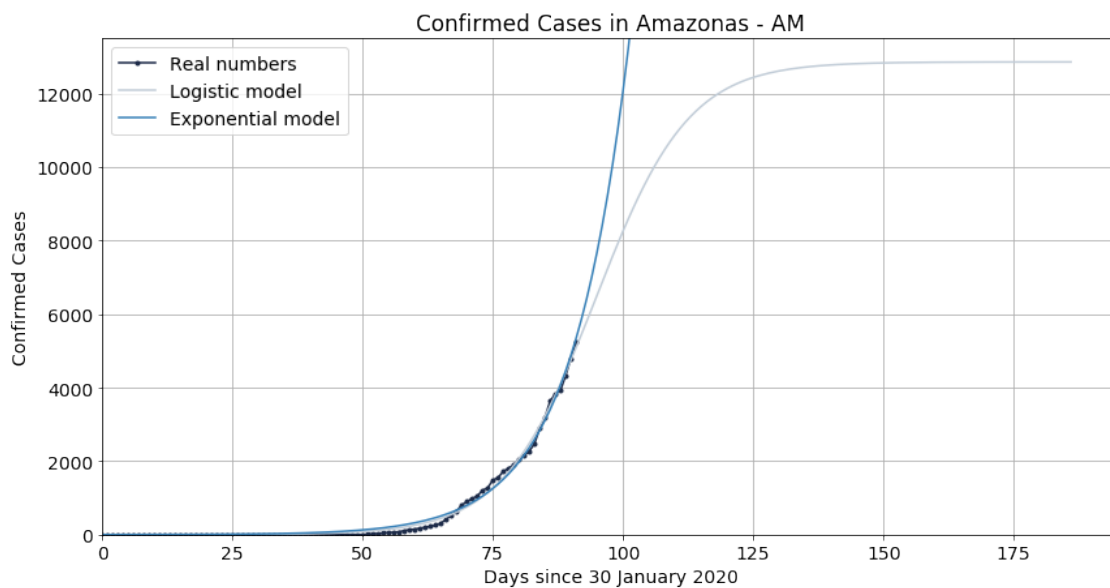
Next estimated number of confirmed cases in Alagoas is: 1287.105 or + 243.105



Next estimated number of confirmed cases in Alagoas is: 55.467 or + 8.467

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

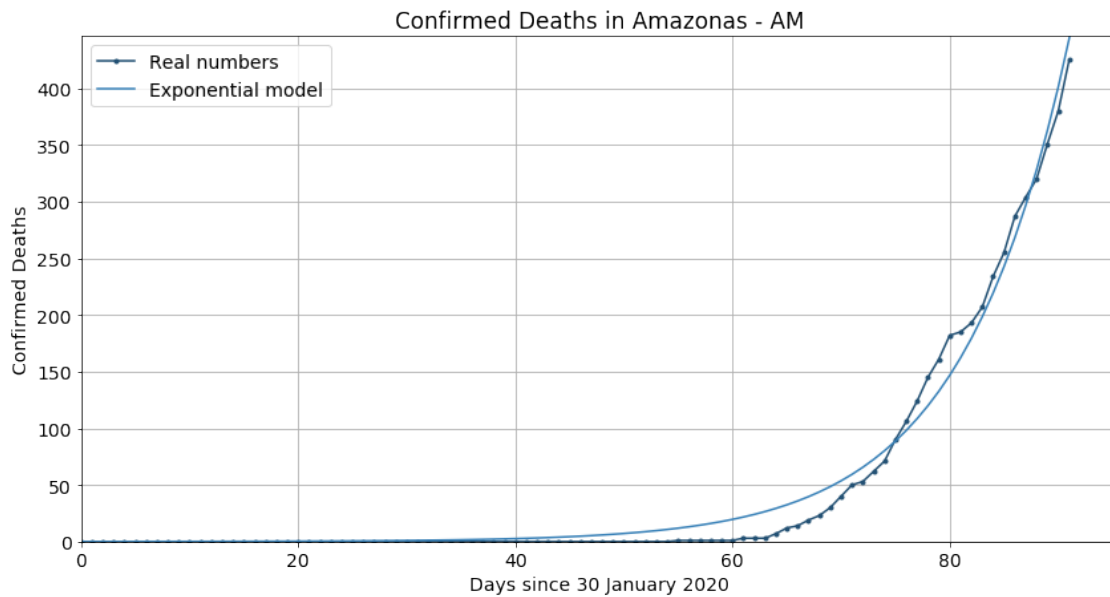
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



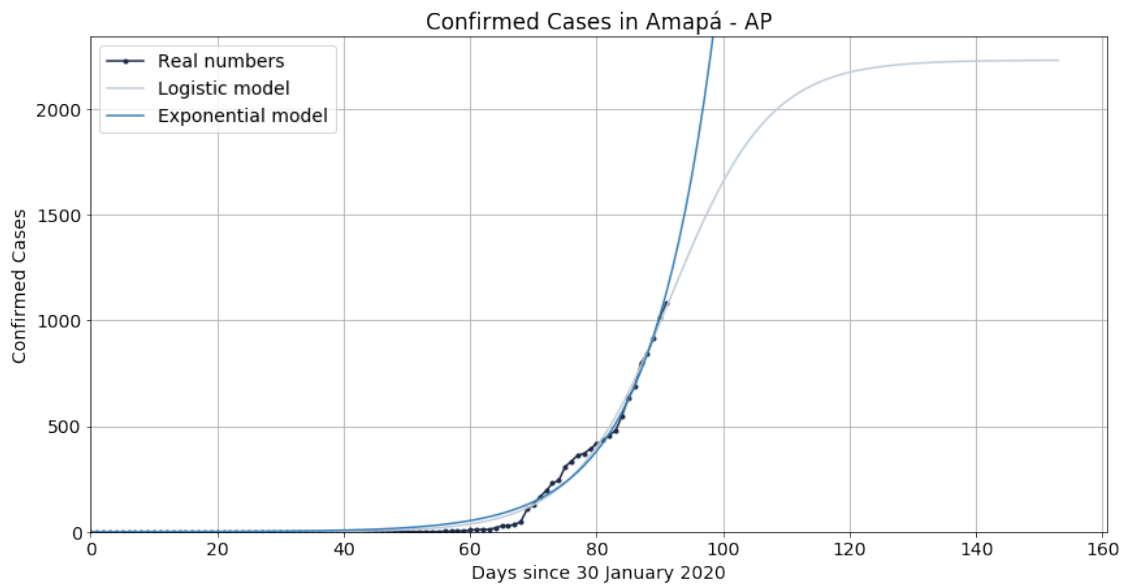
Next estimated number of confirmed cases in Amazonas is: 5855.522 or + 601.522

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

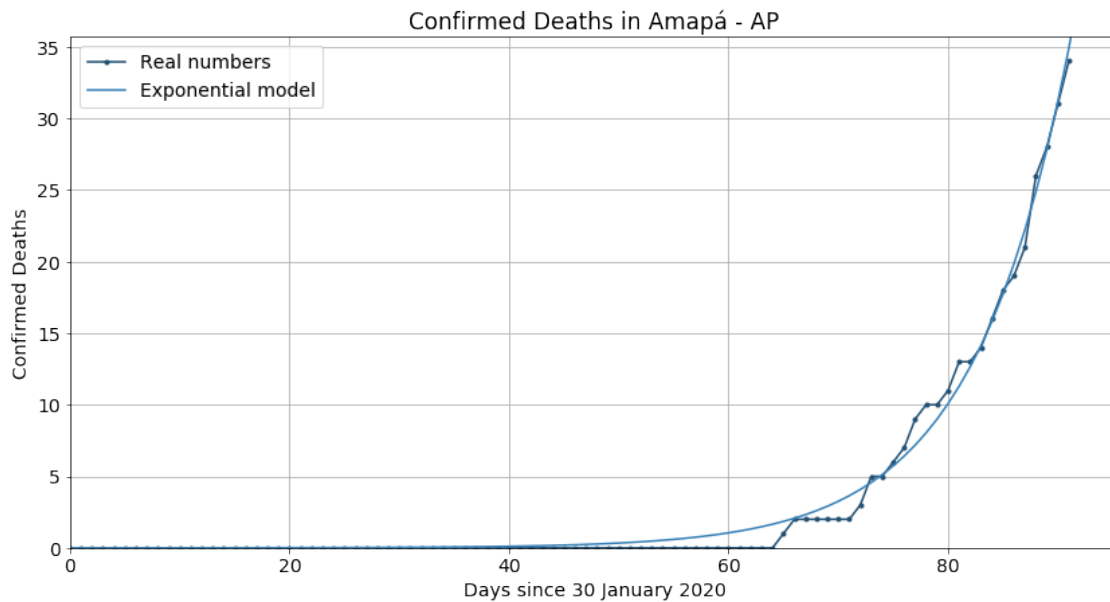
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



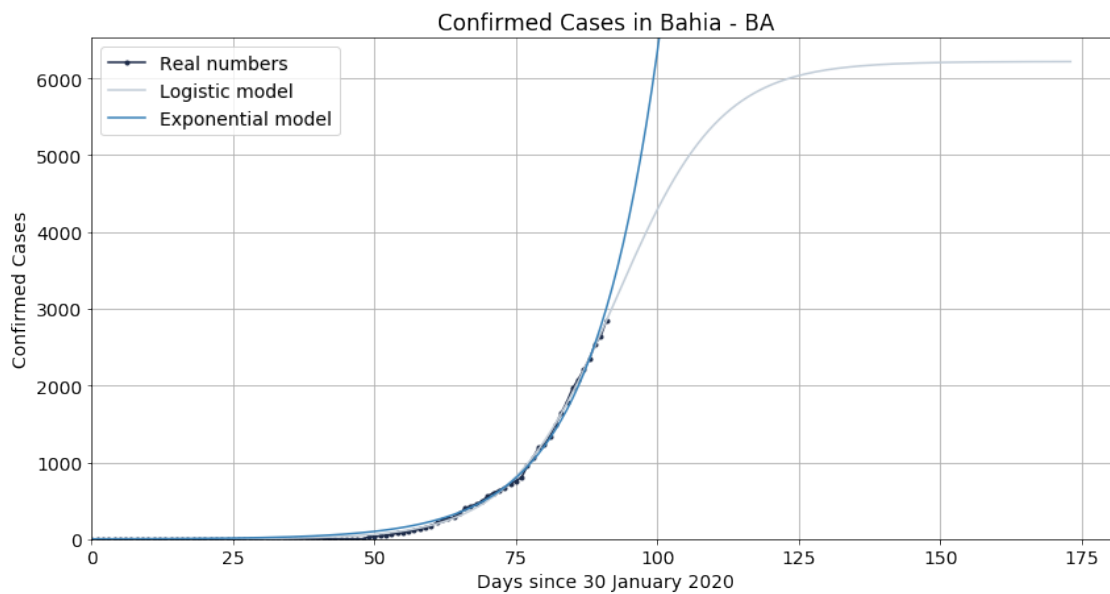
Next estimated number of confirmed cases in Amazonas is: 490.704 or + 65.704



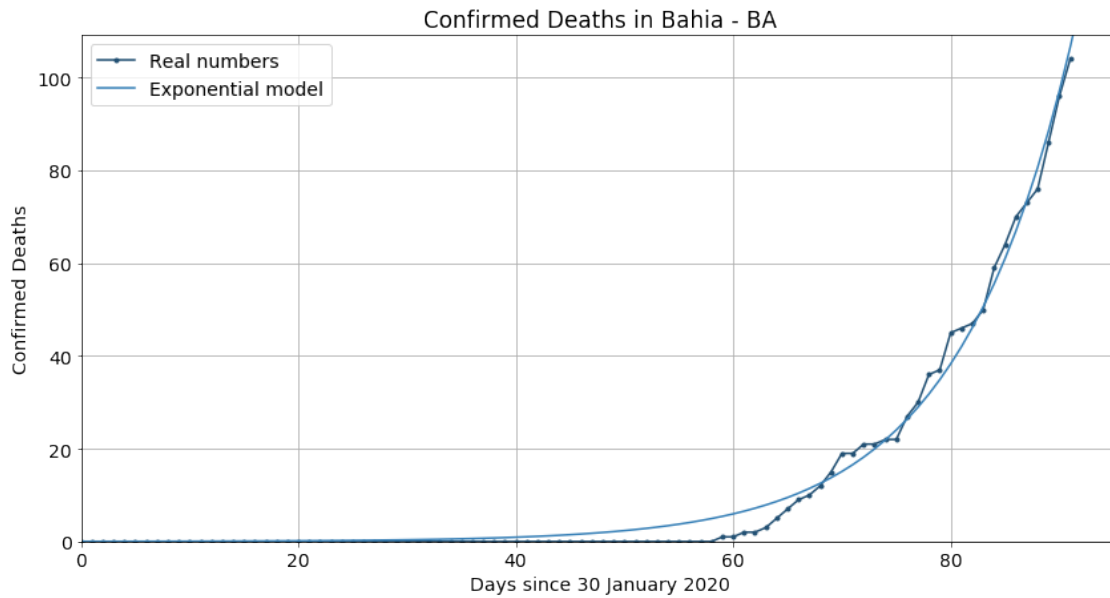
Next estimated number of confirmed cases in Amapá is: 1245.971 or + 165.971



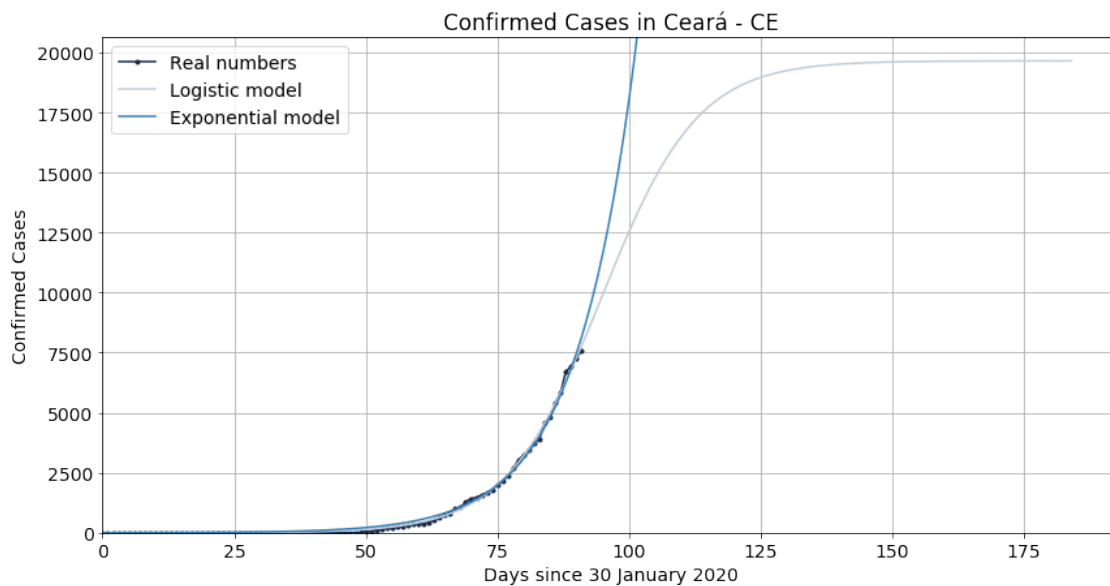
Next estimated number of confirmed cases in Amapá is: 39.076 or + 5.076



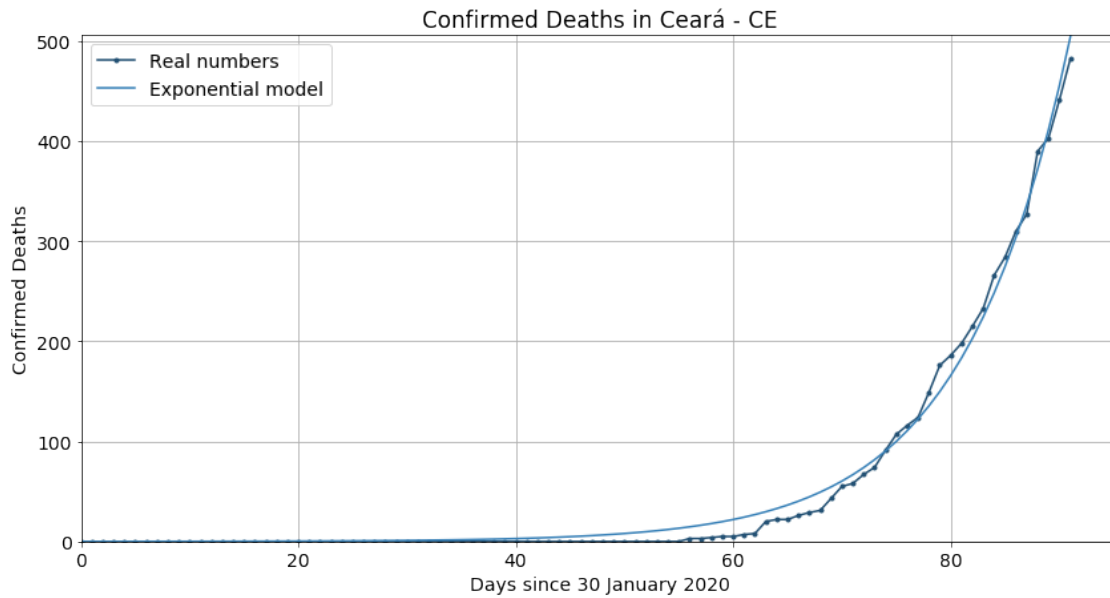
Next estimated number of confirmed cases in Bahia is: 3274.571 or + 423.571



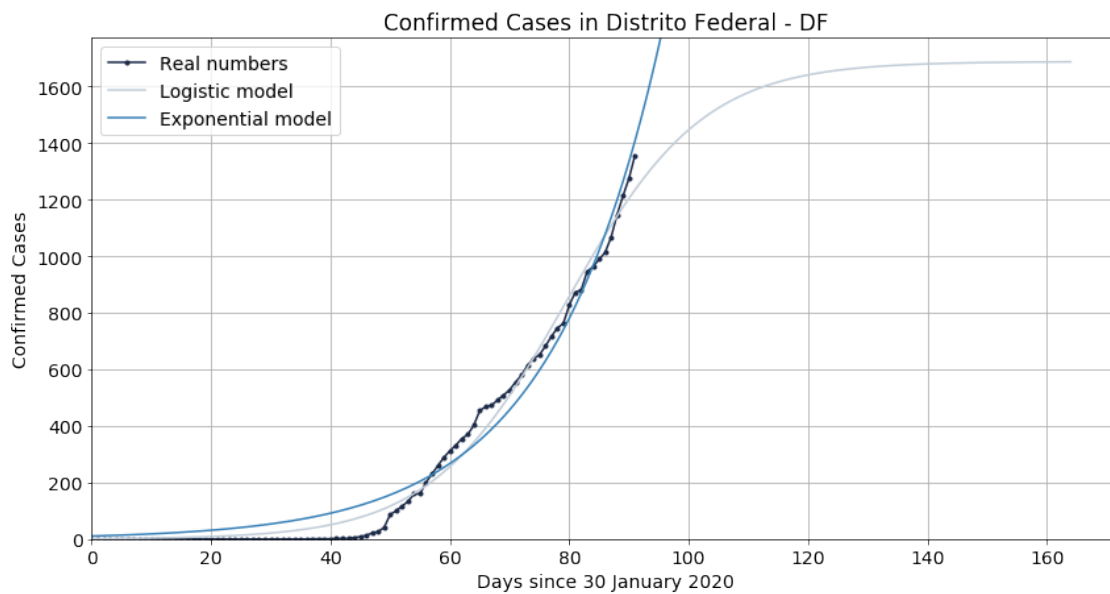
Next estimated number of confirmed cases in Bahia is: 116.981 or + 12.981



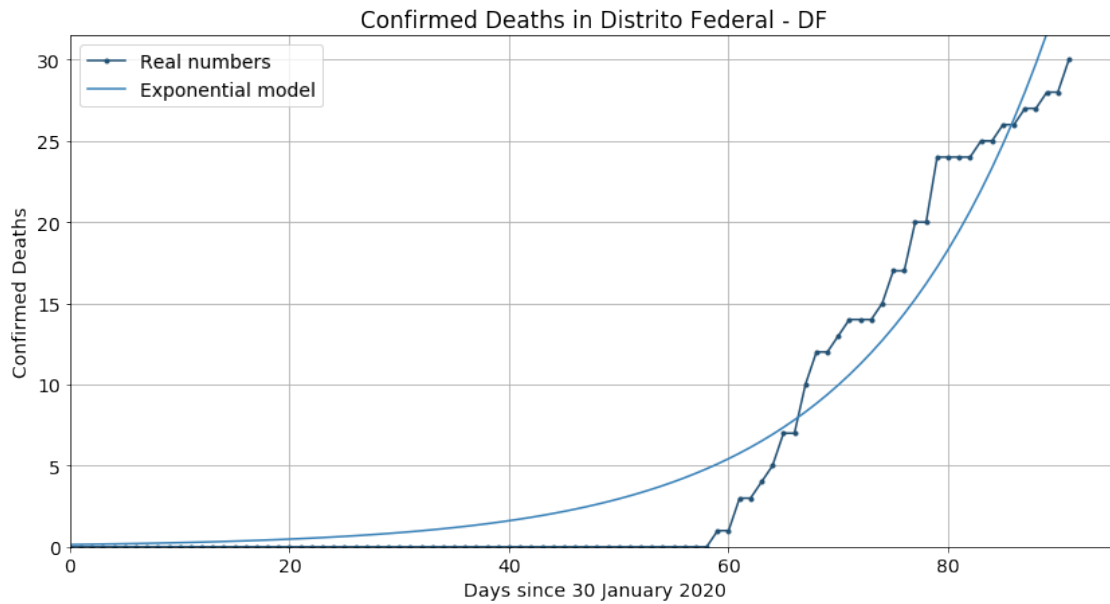
Next estimated number of confirmed cases in Ceará is: 8913.326 or + 1307.326



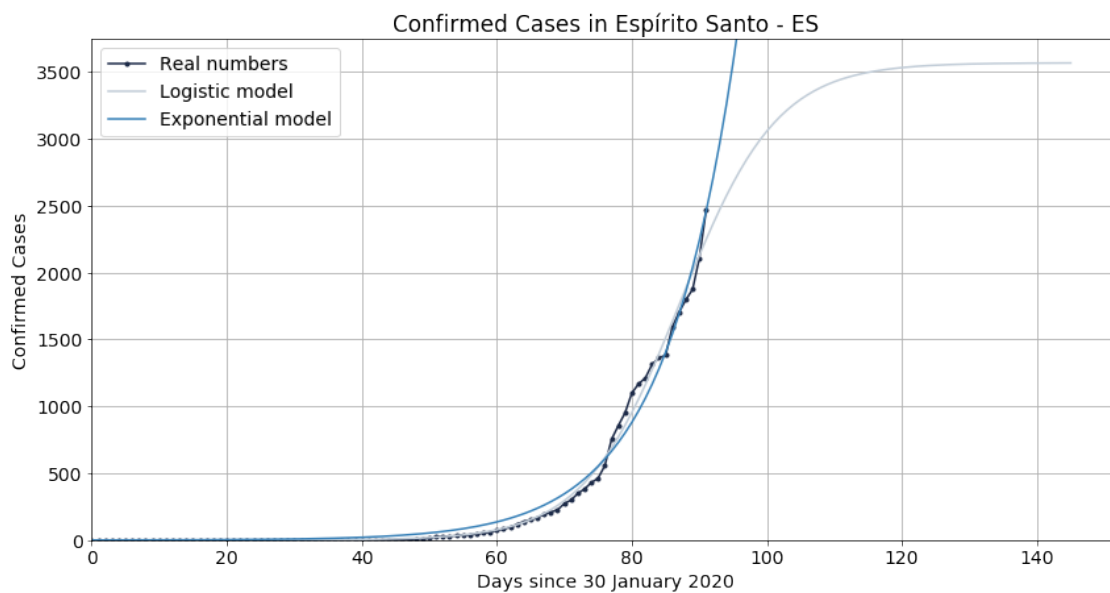
Next estimated number of confirmed cases in Ceará is: 557.208 or + 75.208



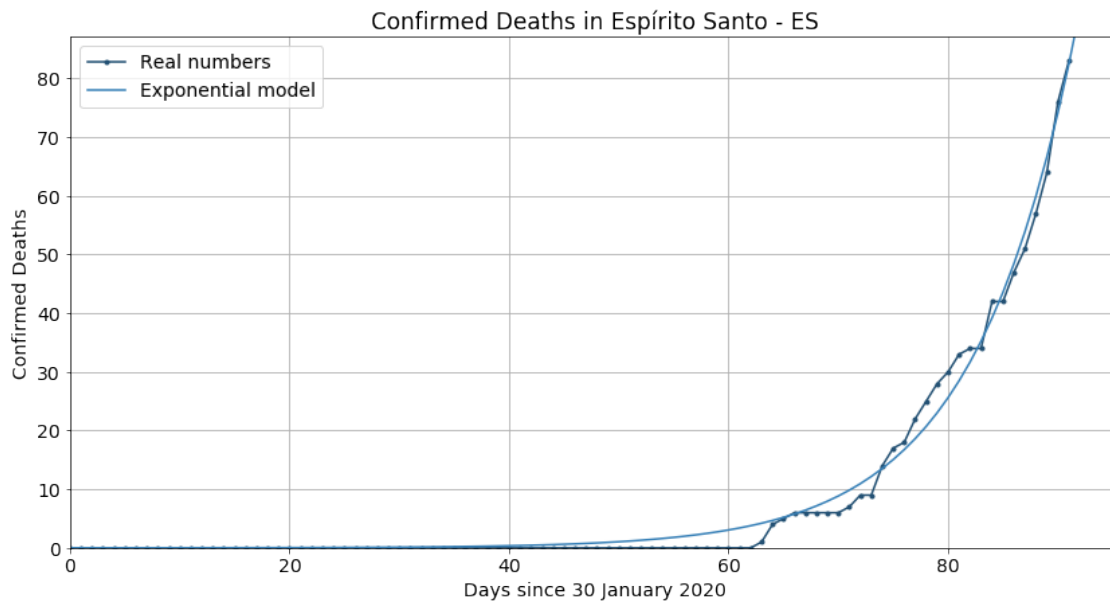
Next estimated number of confirmed cases in Distrito Federal is: 1484.329 or + 128.329



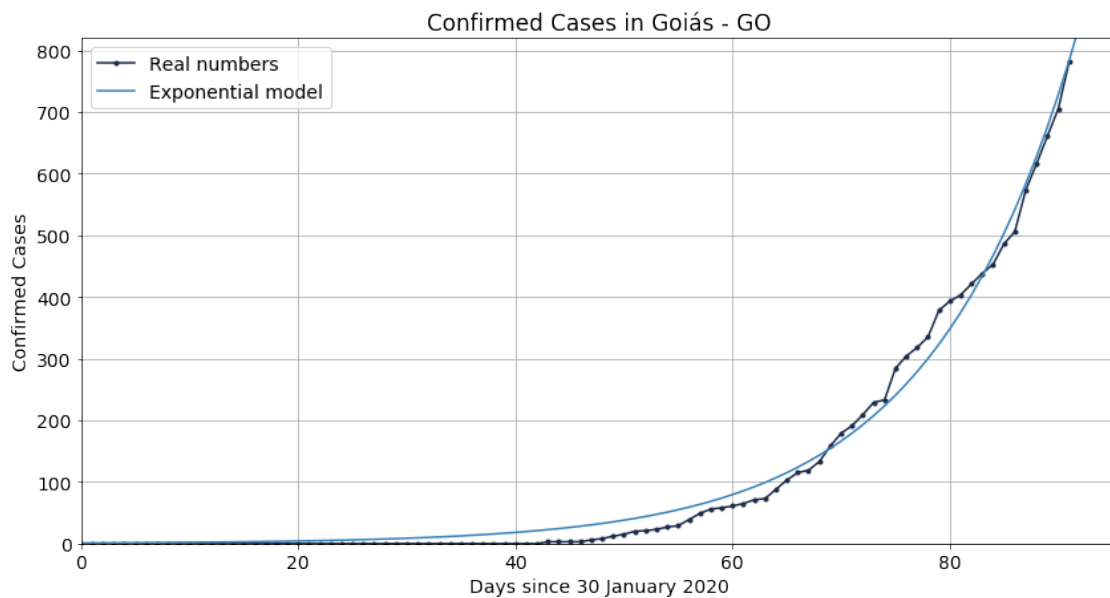
Next estimated number of confirmed cases in Distrito Federal is: 37.952 or + 7.952



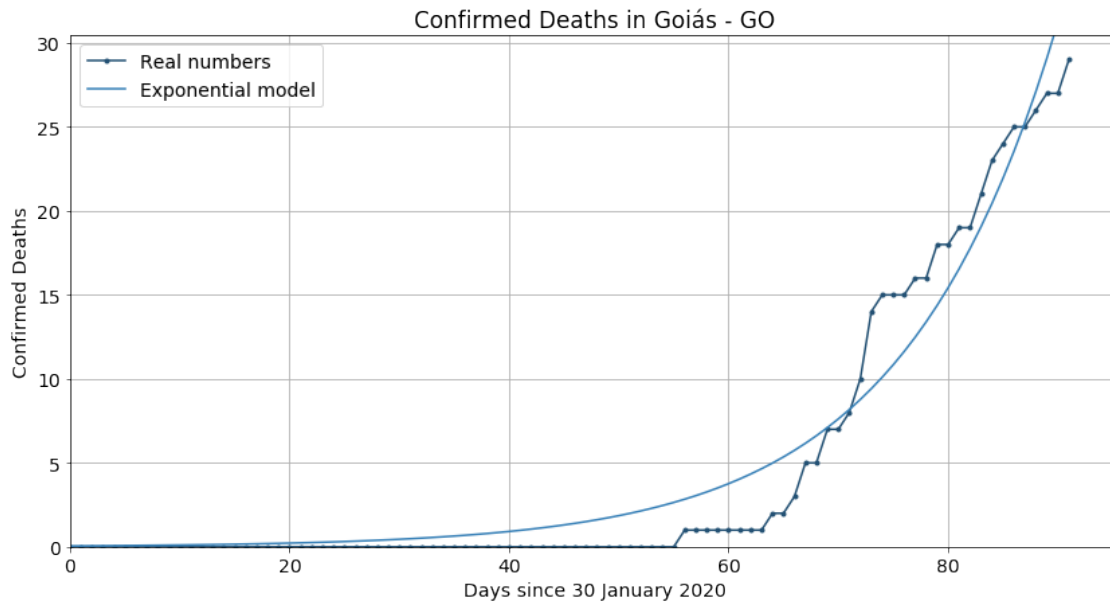
Next estimated number of confirmed cases in Espírito Santo is: 2698.314 or + 233.314



Next estimated number of confirmed cases in Espírito Santo is: 91.707 or + 8.707



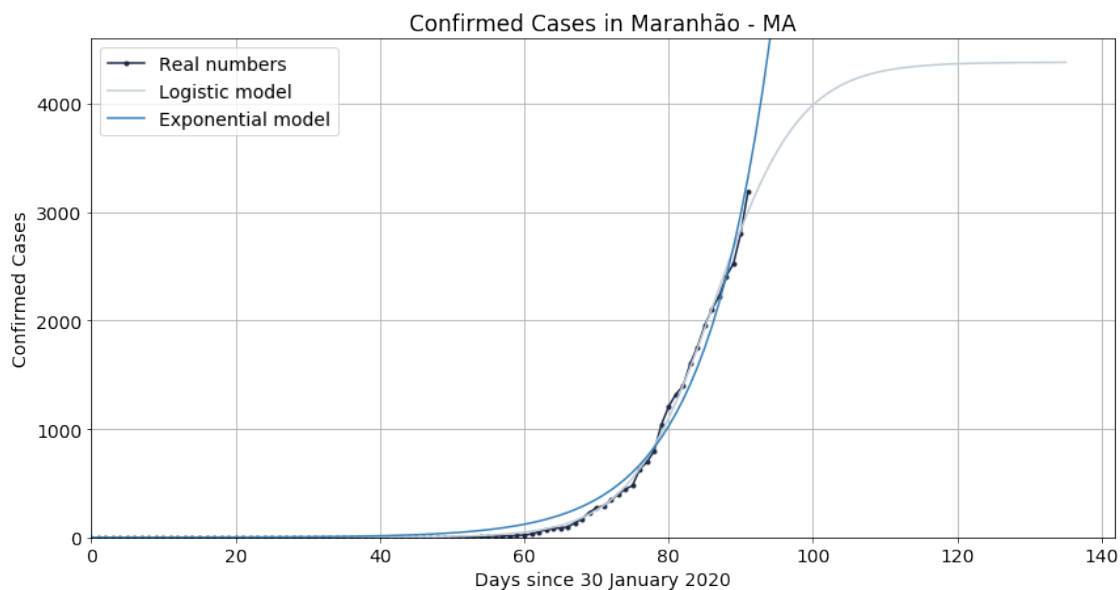
Next estimated number of confirmed cases in Goiás is: 843.598 or + 62.598



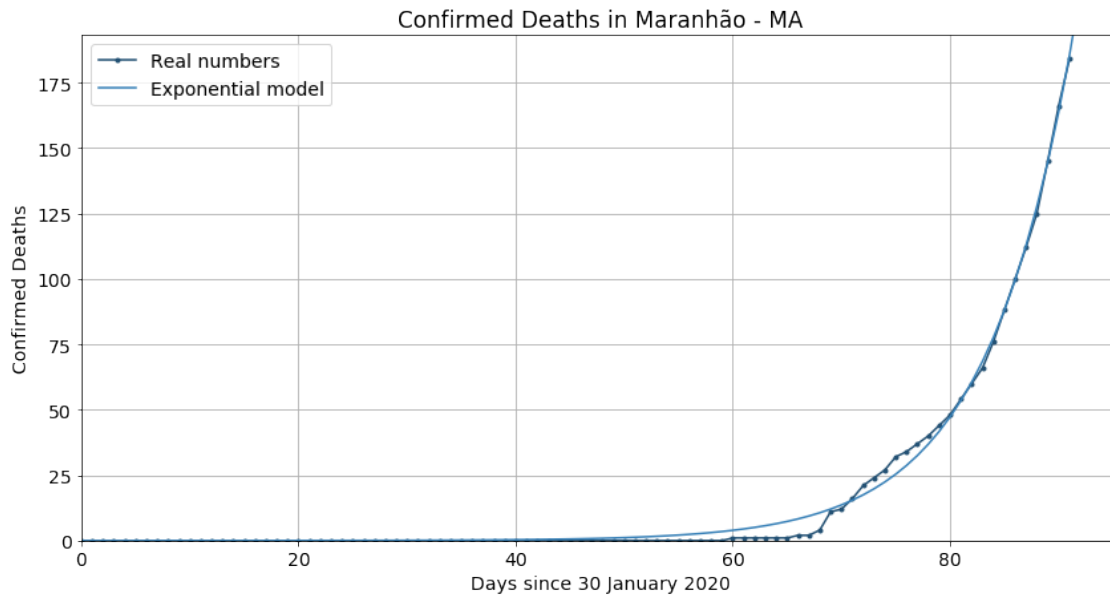
Next estimated number of confirmed cases in Goiás is: 35.953 or + 6.953

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

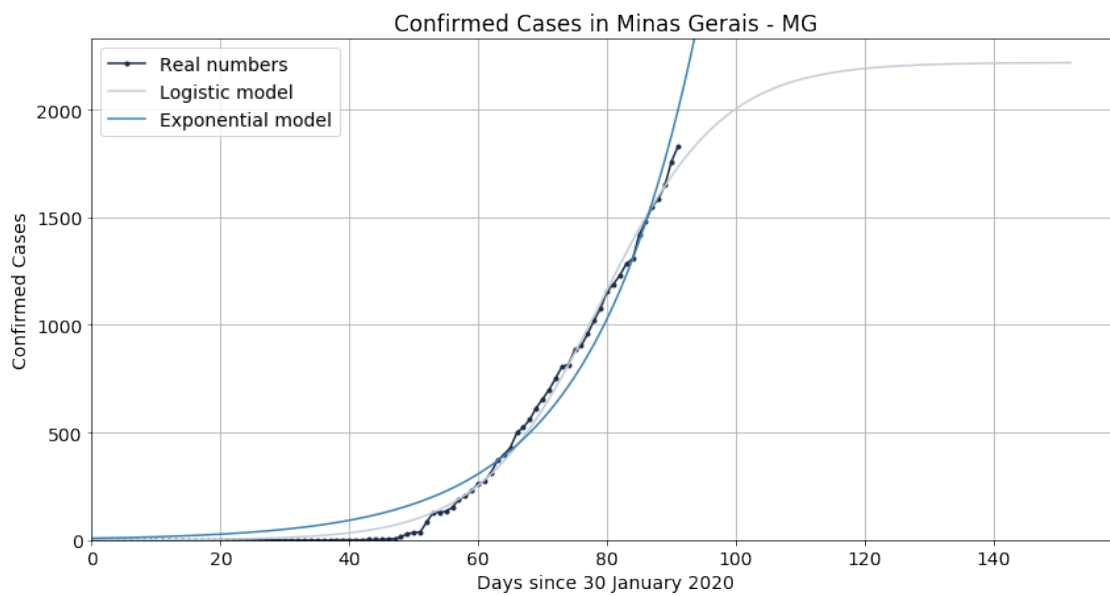
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



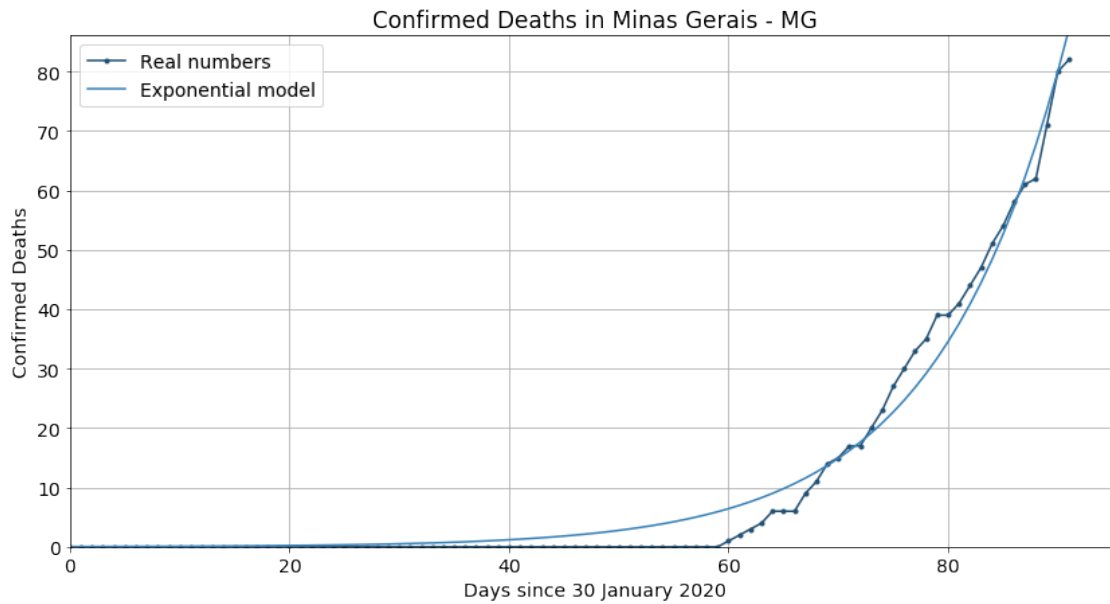
Next estimated number of confirmed cases in Maranhão is: 3693.448 or + 503.448



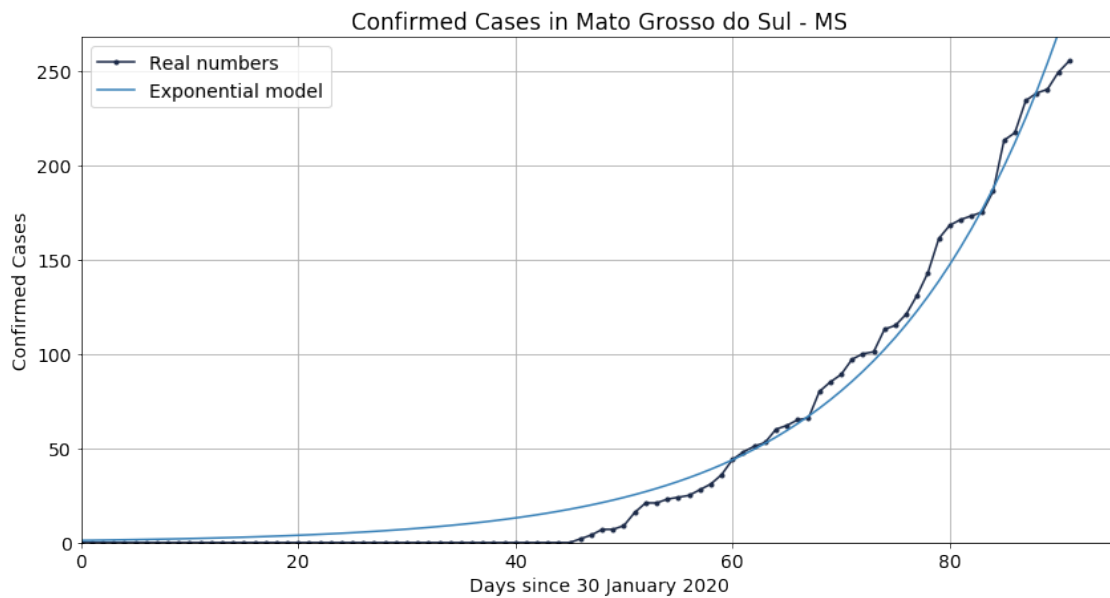
Next estimated number of confirmed cases in Maranhão is: 209.941 or + 25.941



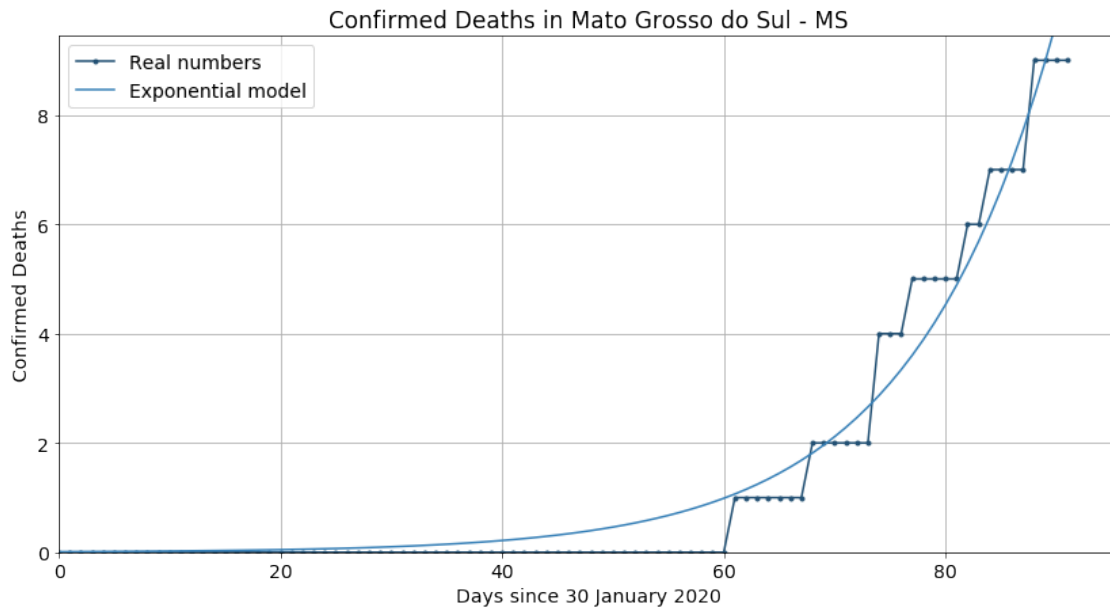
Next estimated number of confirmed cases in Minas Gerais is: 2118.478 or + 291.478



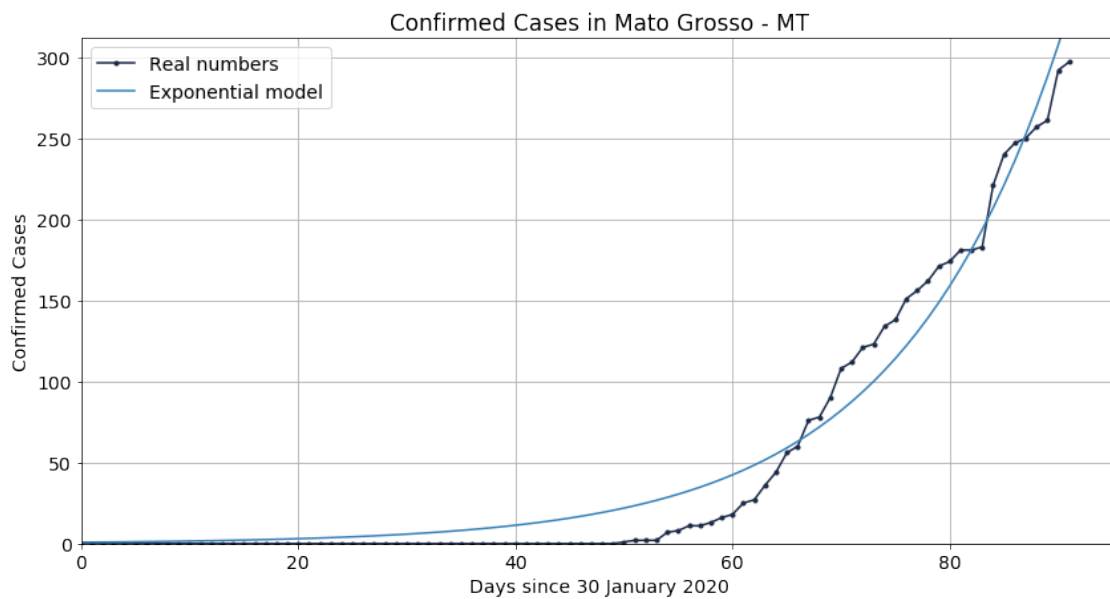
Next estimated number of confirmed cases in Minas Gerais is: 94.714 or + 12.714



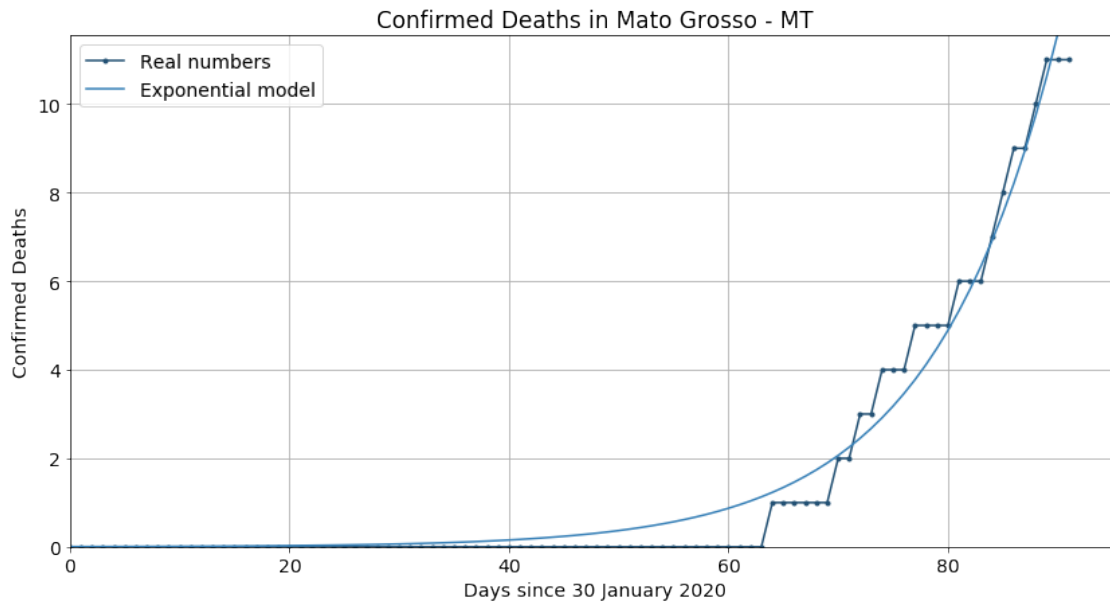
Next estimated number of confirmed cases in Mato Grosso do Sul is: 304.439 or + 49.439



Next estimated number of confirmed cases in Mato Grosso do Sul is: 11.294 or + 2.294



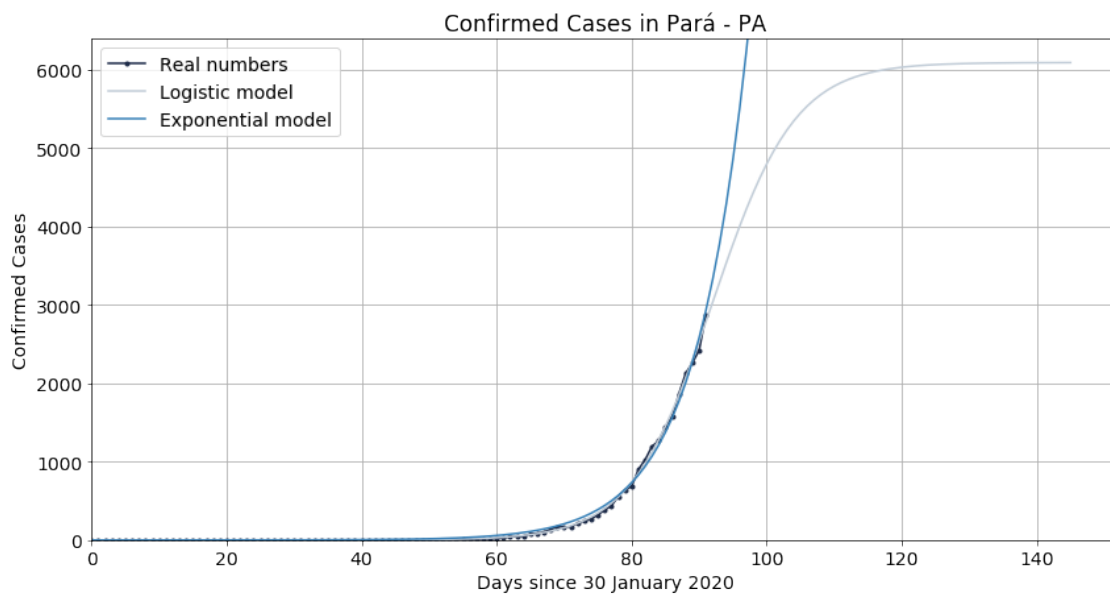
Next estimated number of confirmed cases in Mato Grosso is: 351.039 or + 54.039



Next estimated number of confirmed cases in Mato Grosso is: 13.759 or + 2.759

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

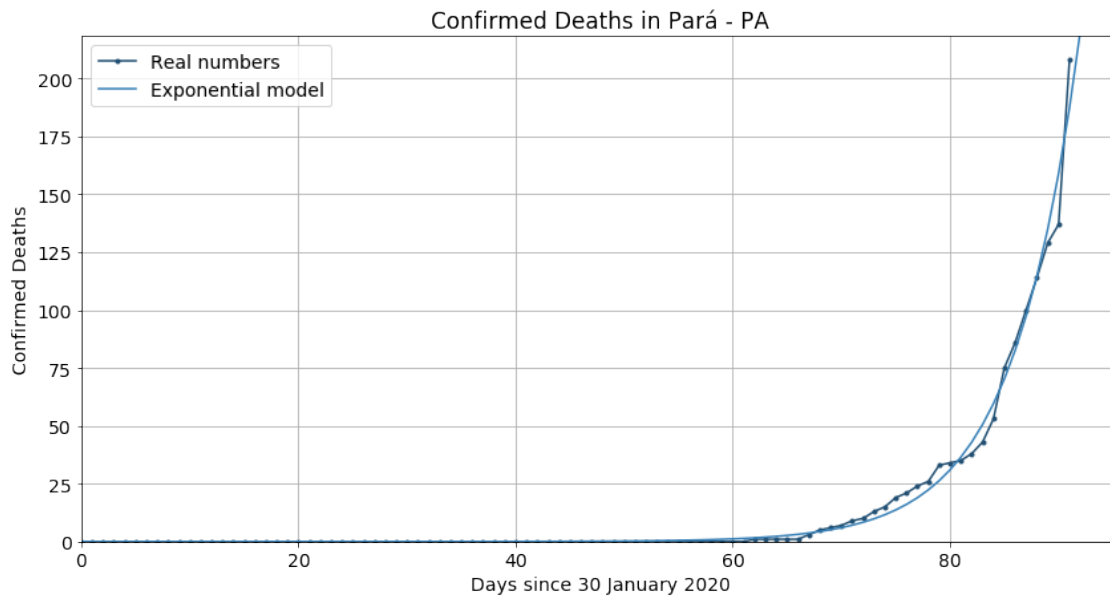
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



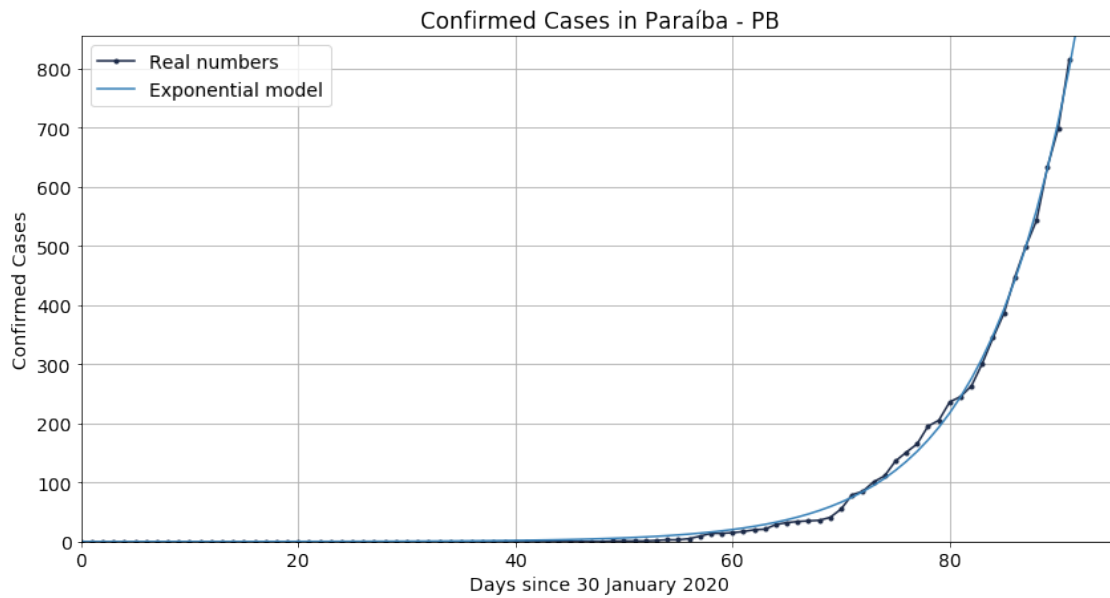
Next estimated number of confirmed cases in Pará is: 3326.416 or + 450.416

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

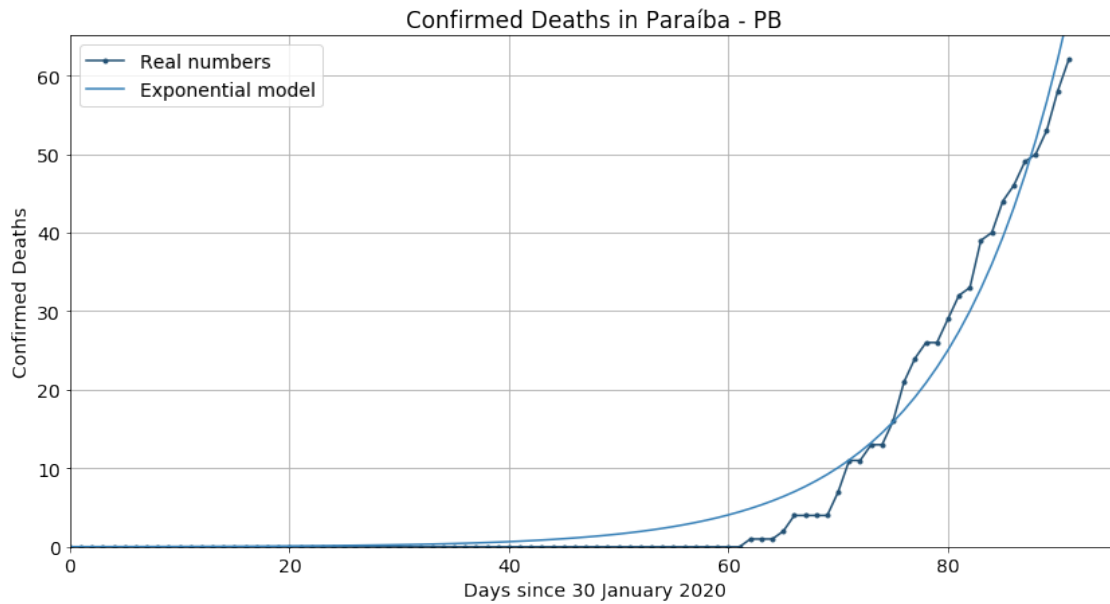
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



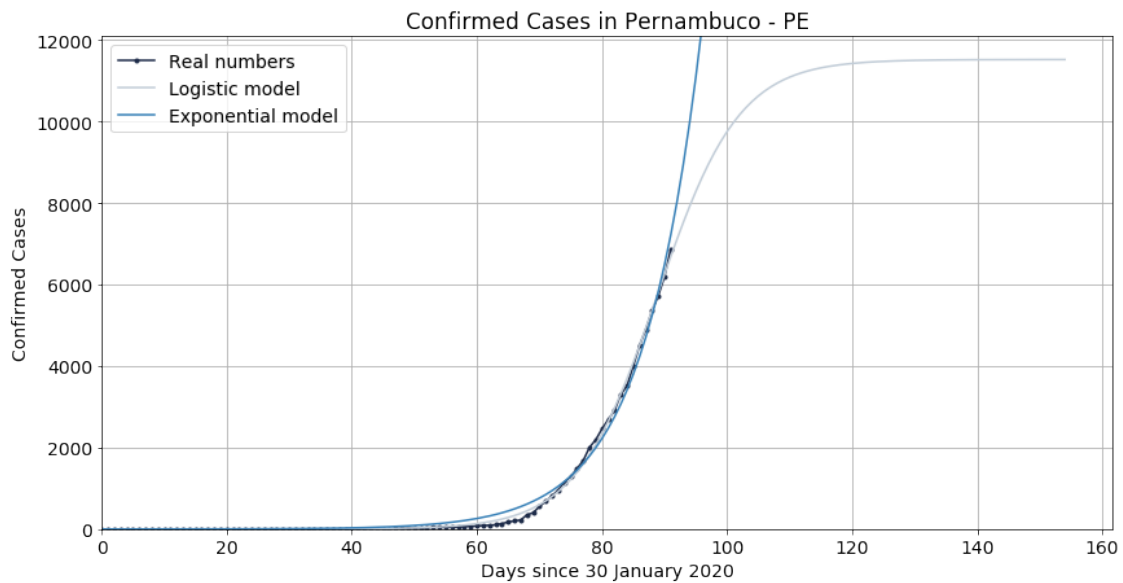
Next estimated number of confirmed cases in Pará is: 220.56 or + 12.56



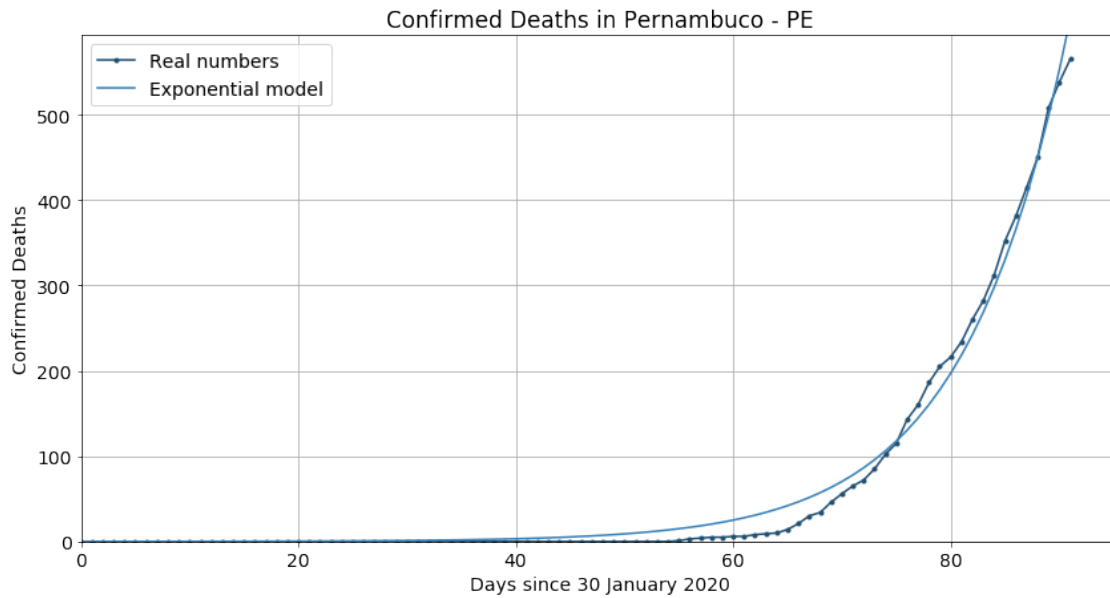
Next estimated number of confirmed cases in Paraíba is: 899.782 or + 85.782



Next estimated number of confirmed cases in Paraíba is: 74.425 or + 12.425



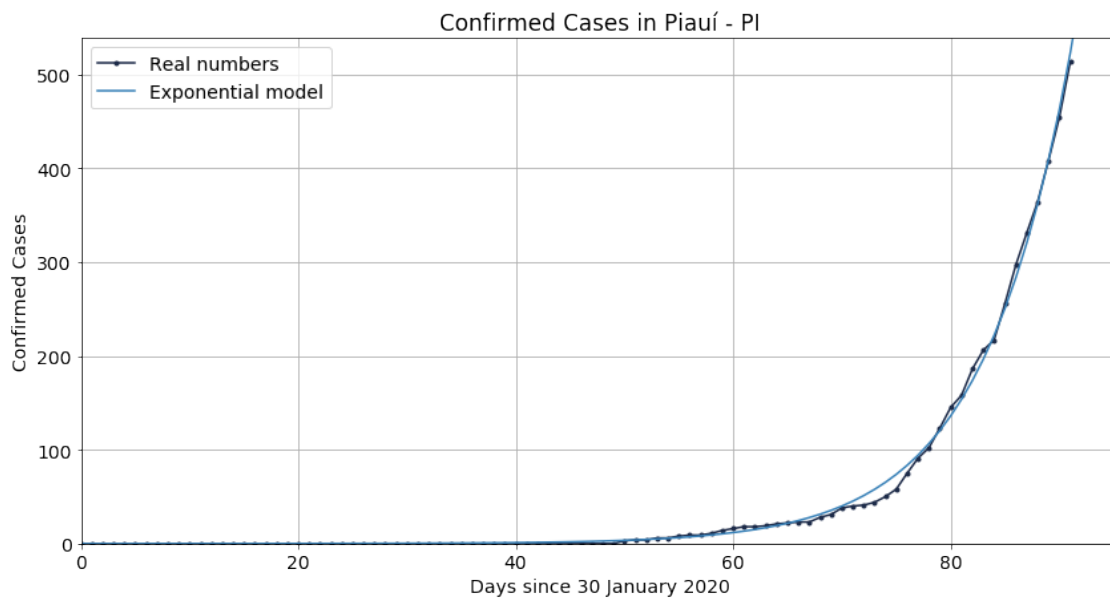
Next estimated number of confirmed cases in Pernambuco is: 8039.848 or + 1163.848



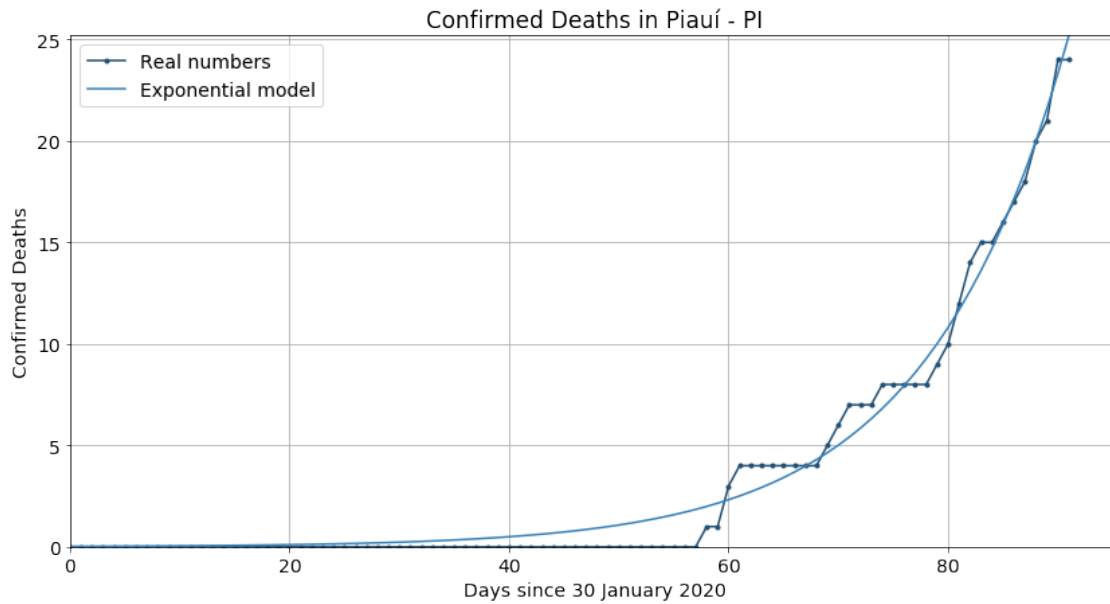
Next estimated number of confirmed cases in Pernambuco is: 678.992 or + 113.992

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

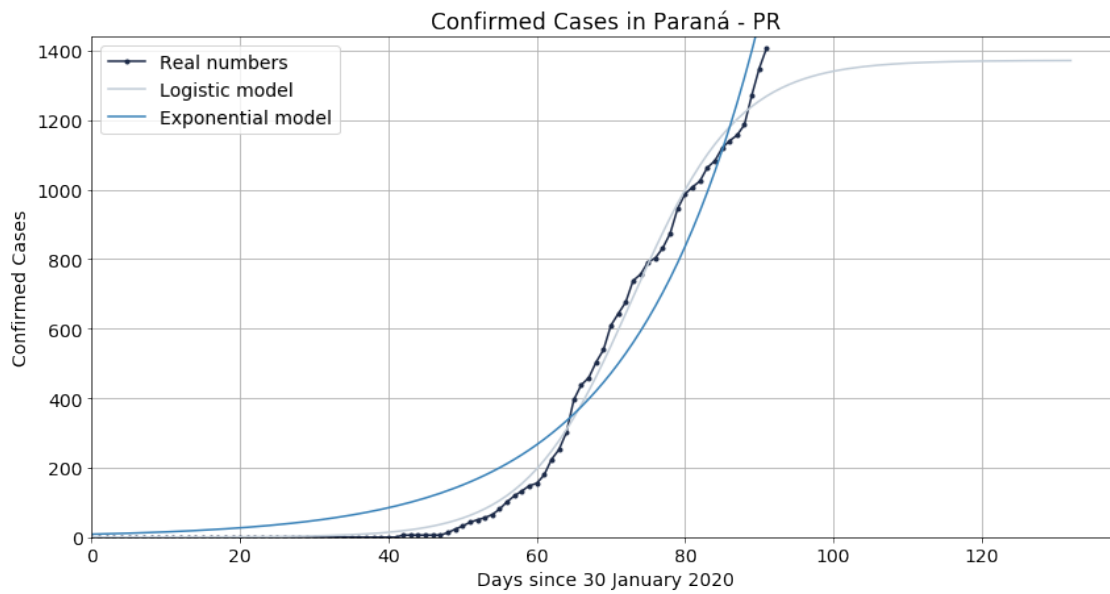
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



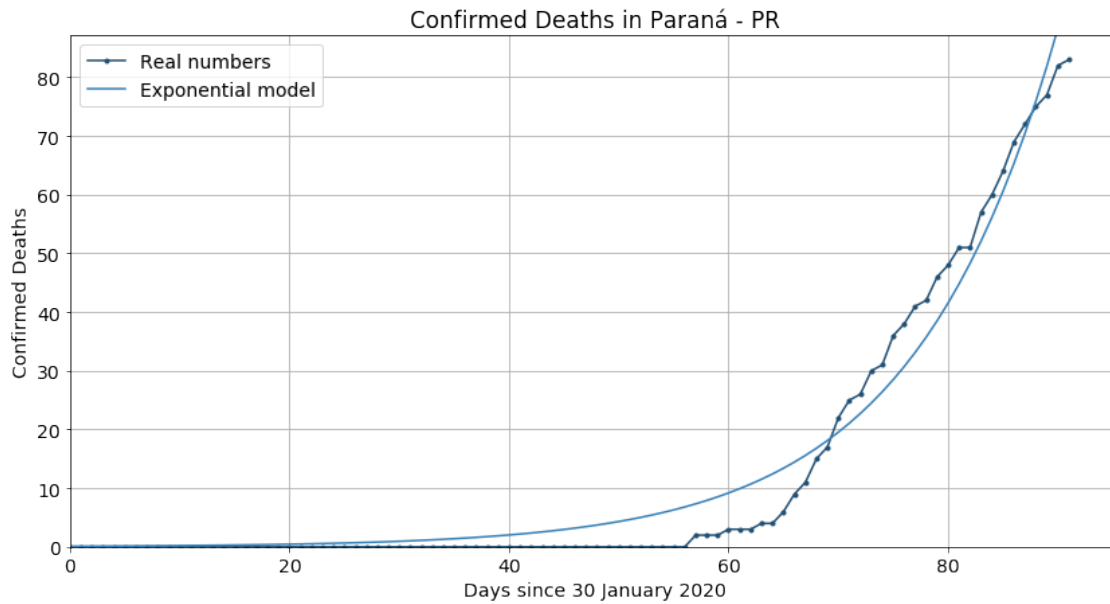
Next estimated number of confirmed cases in Piauí is: 590.351 or + 77.351



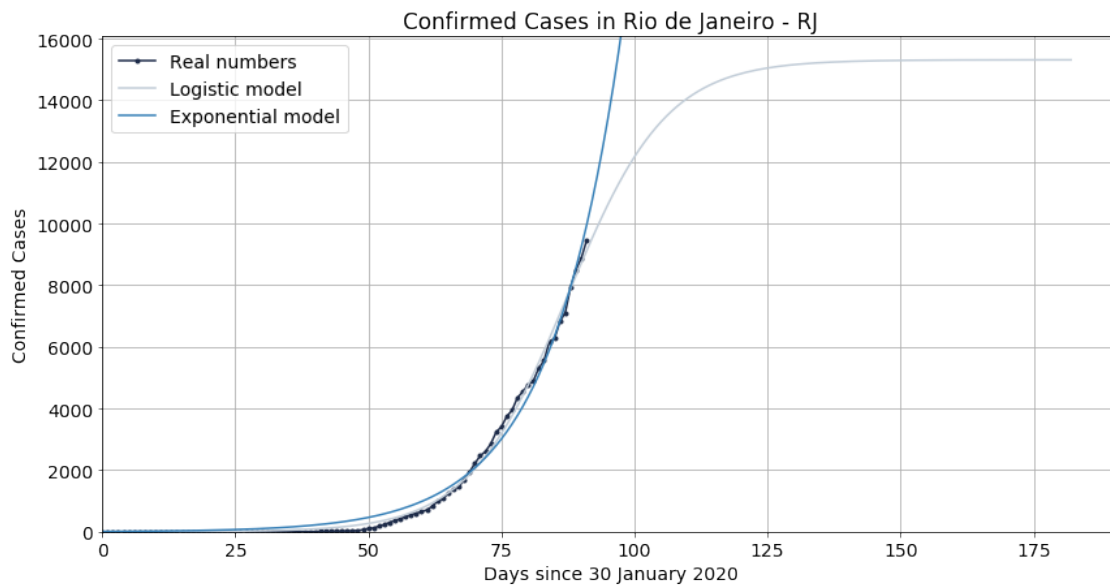
Next estimated number of confirmed cases in Piauí is: 27.192 or + 3.192



Next estimated number of confirmed cases in Paraná is: 1660.857 or + 253.857



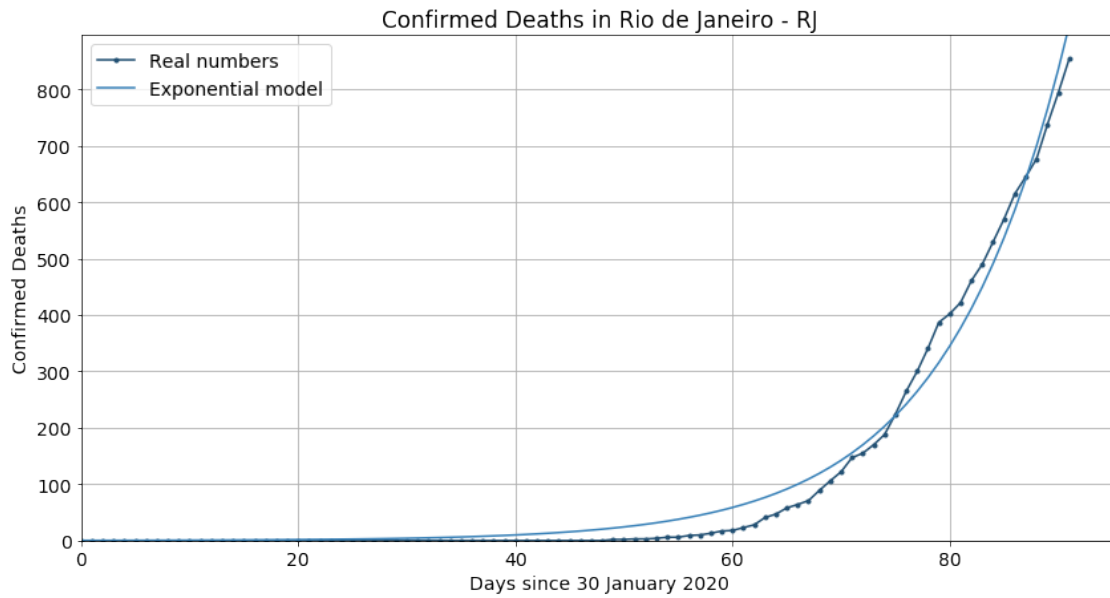
Next estimated number of confirmed cases in Paraná is: 102.717 or + 19.717



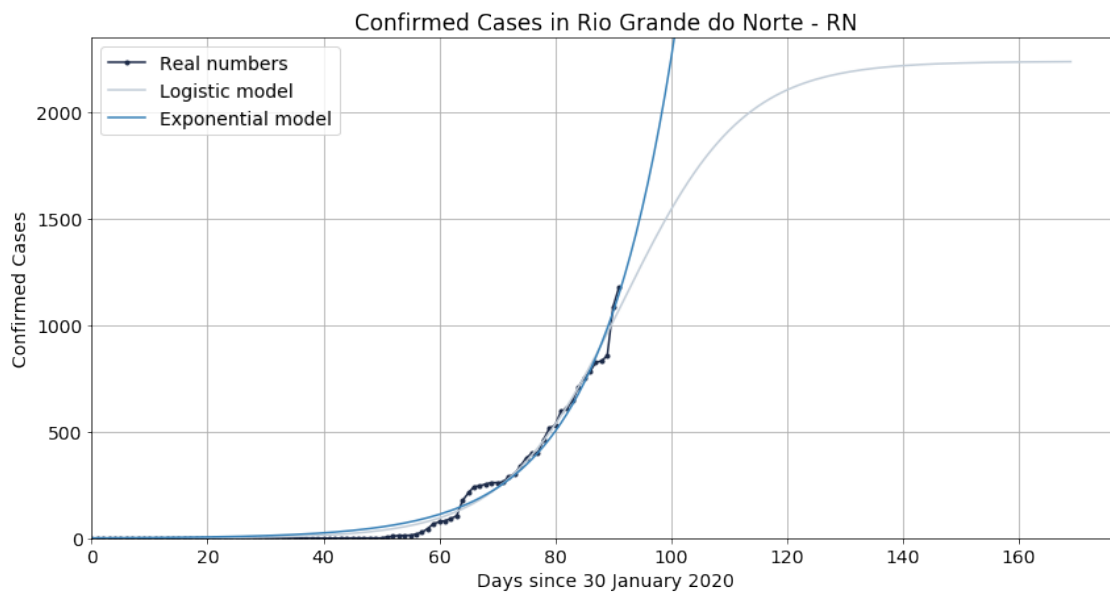
Next estimated number of confirmed cases in Rio de Janeiro is: 10714.042 or + 1261.042

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

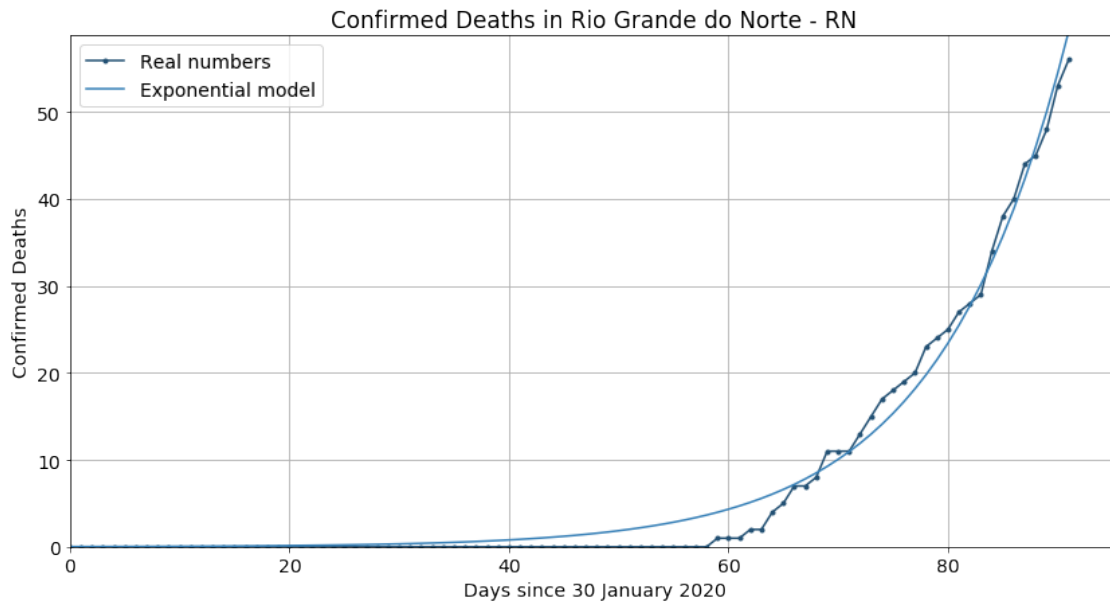
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



Next estimated number of confirmed cases in Rio de Janeiro is: 997.514 or + 143.514



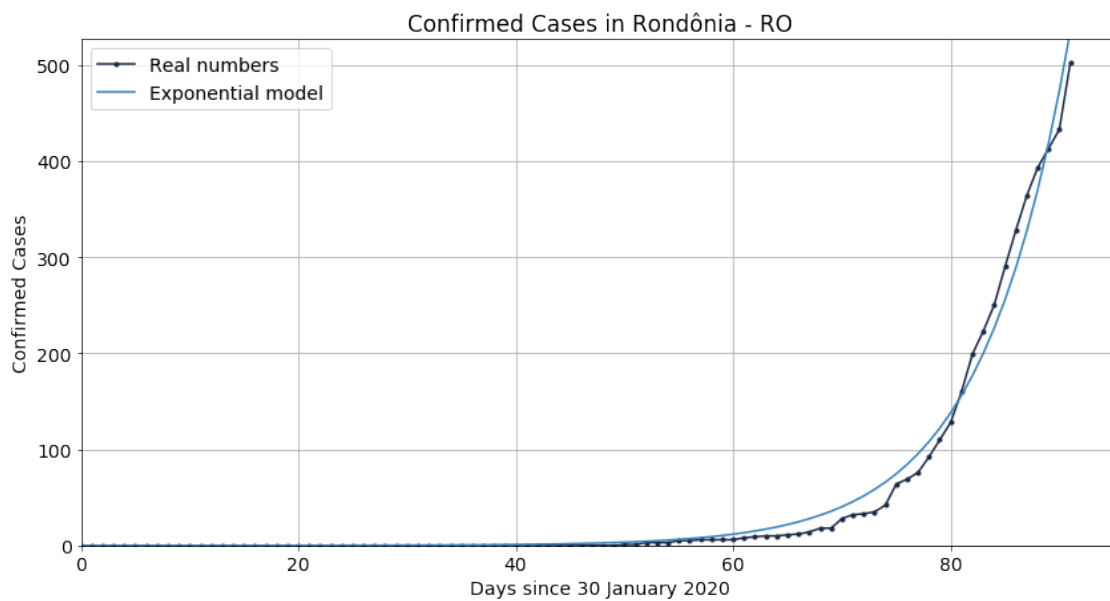
Next estimated number of confirmed cases in Rio Grande do Norte is: 1234.132 or + 57.132



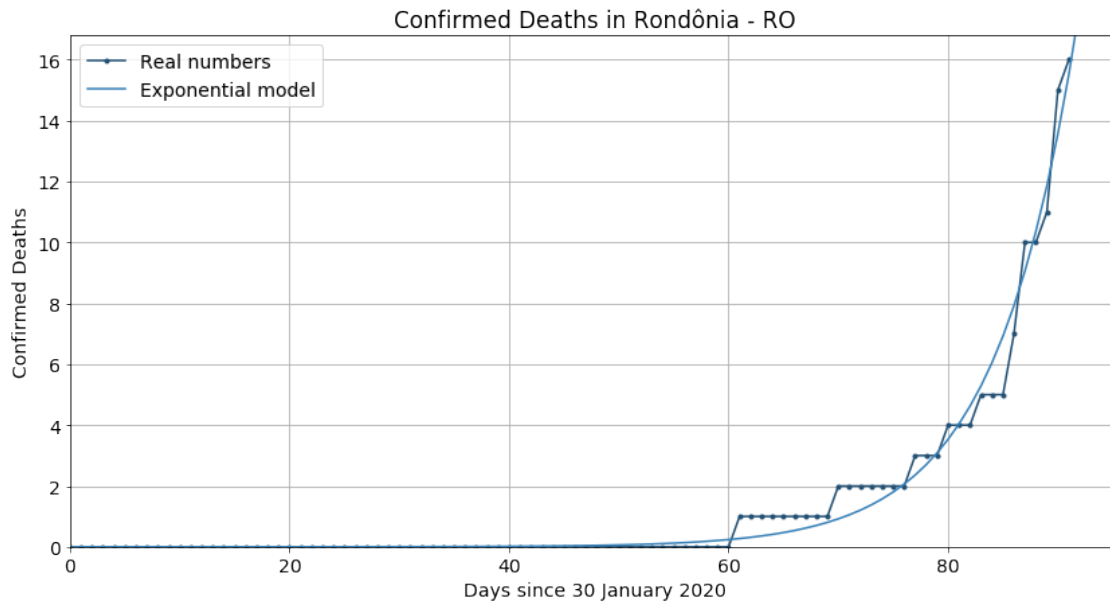
Next estimated number of confirmed cases in Rio Grande do Norte is: 64.477 or + 8.477

C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:

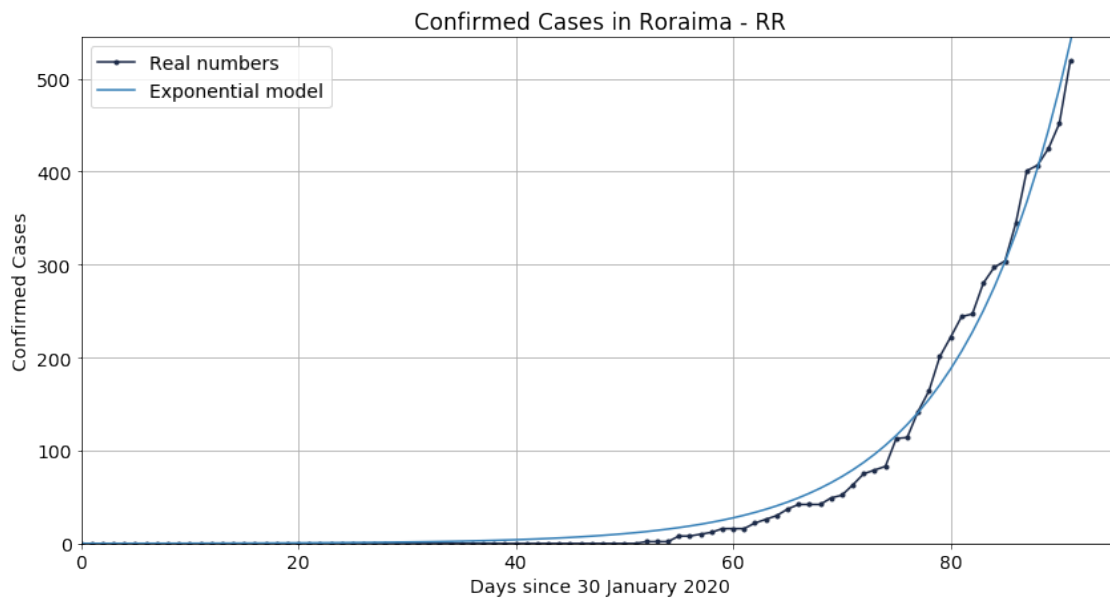
RuntimeWarning: overflow encountered in exp
after removing the cwd from sys.path.



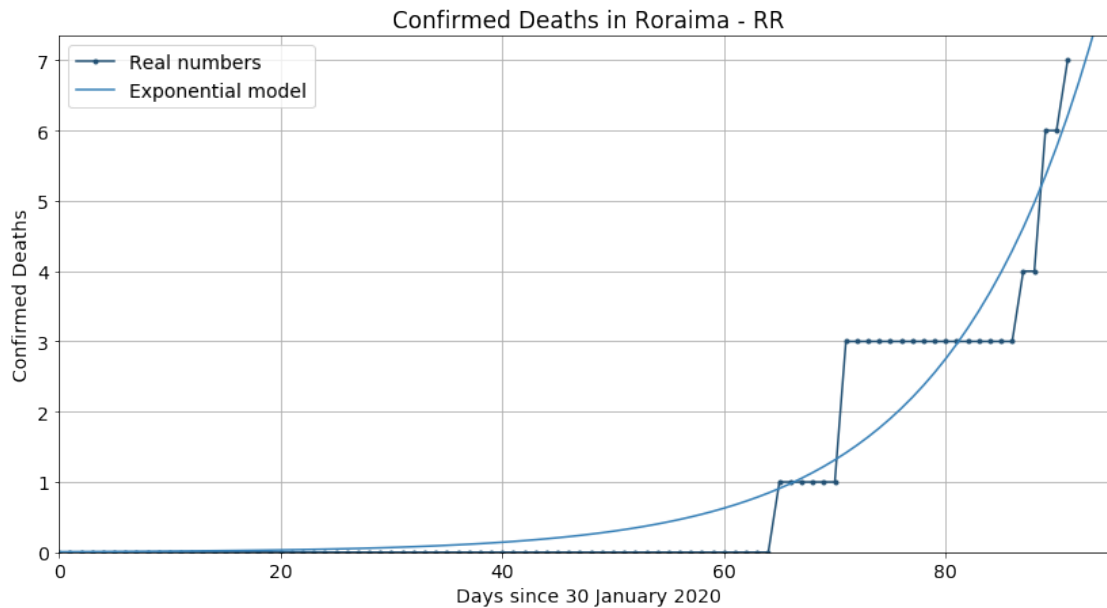
Next estimated number of confirmed cases in Rondônia is: 605.252 or + 103.252



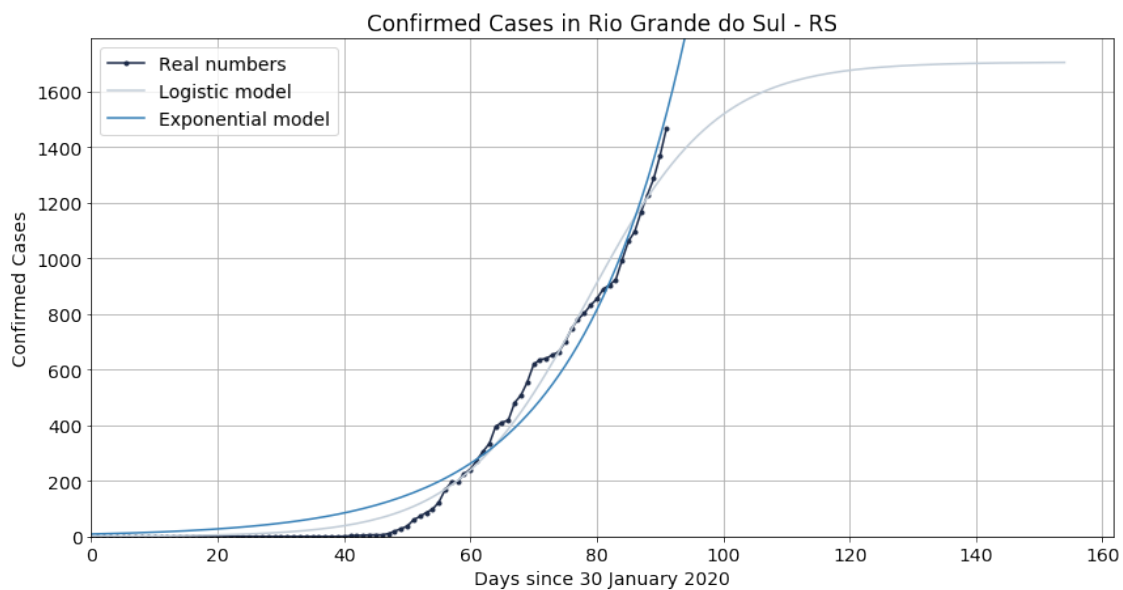
Next estimated number of confirmed cases in Rondônia is: 17.752 or + 1.752



Next estimated number of confirmed cases in Roraima is: 593.288 or + 74.288

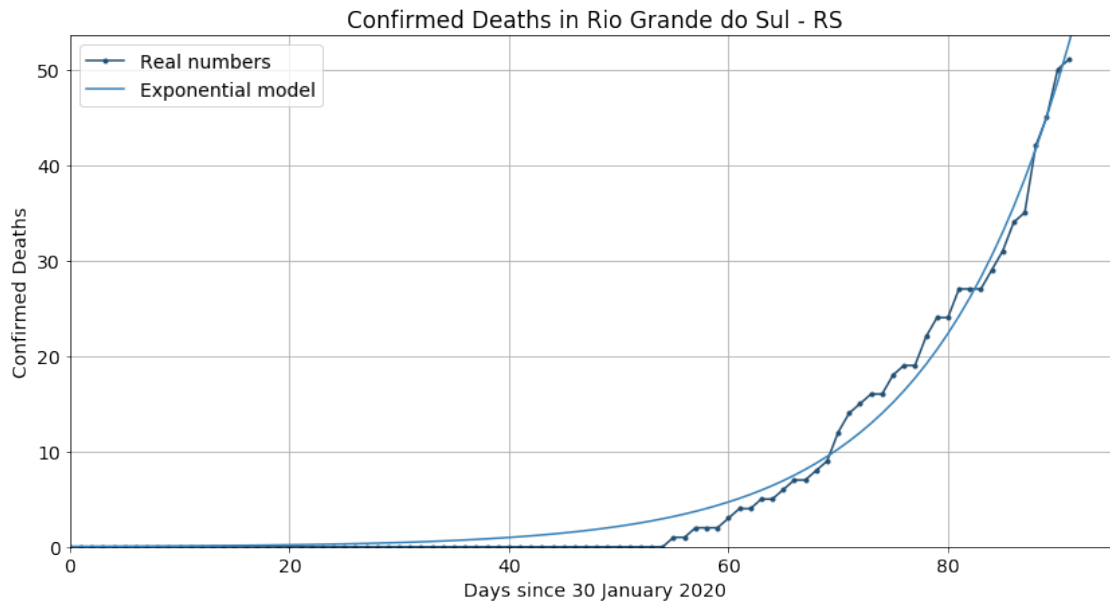


Next estimated number of confirmed cases in Roraima is: 6.685 or + -0.315

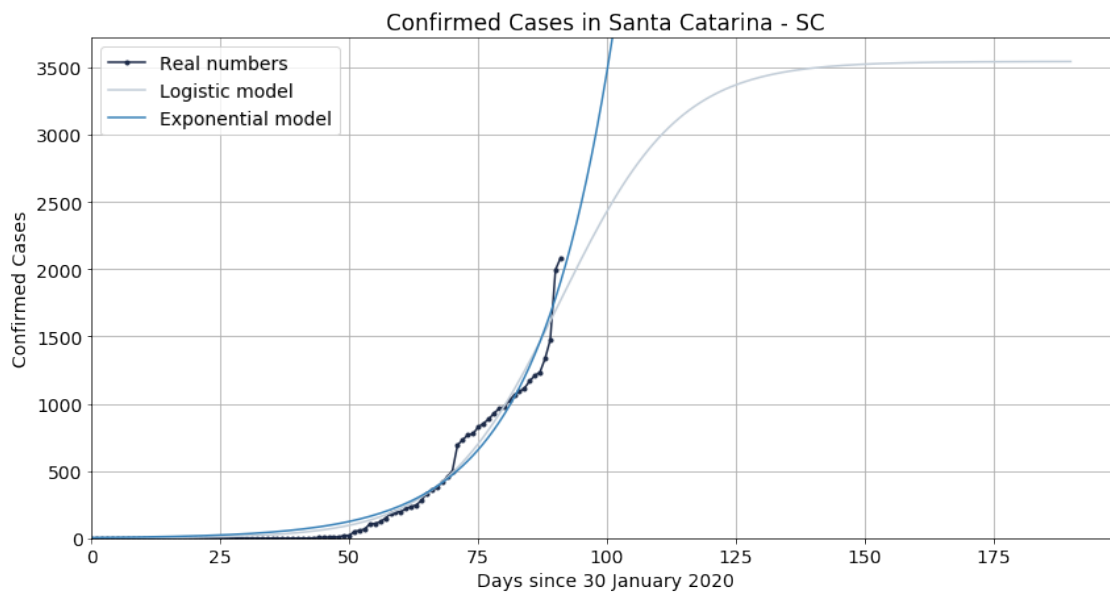


Next estimated number of confirmed cases in Rio Grande do Sul is: 1606.484 or + 140.484

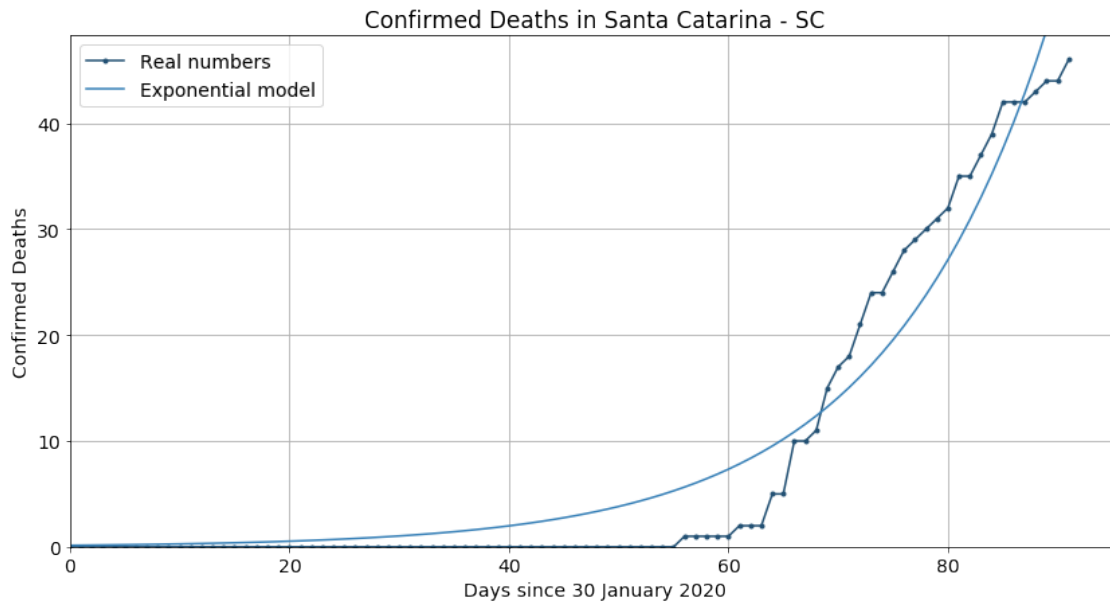
```
C:\Users\calad\anaconda3\lib\site-packages\ipykernel_launcher.py:4:
RuntimeWarning: overflow encountered in exp
  after removing the cwd from sys.path.
```



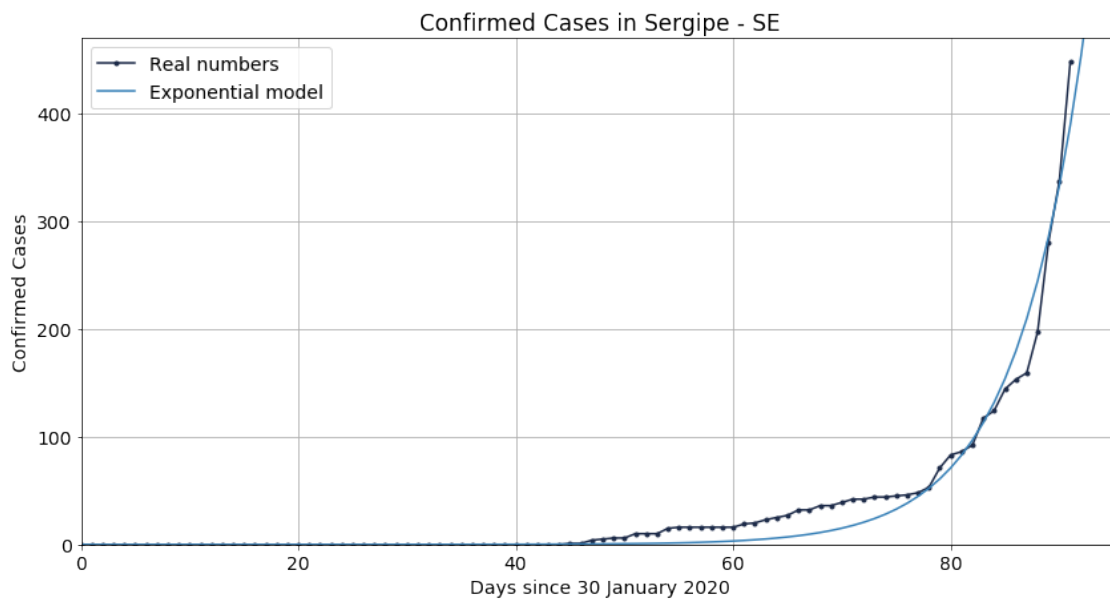
Next estimated number of confirmed cases in Rio Grande do Sul is: 56.911 or + 5.911



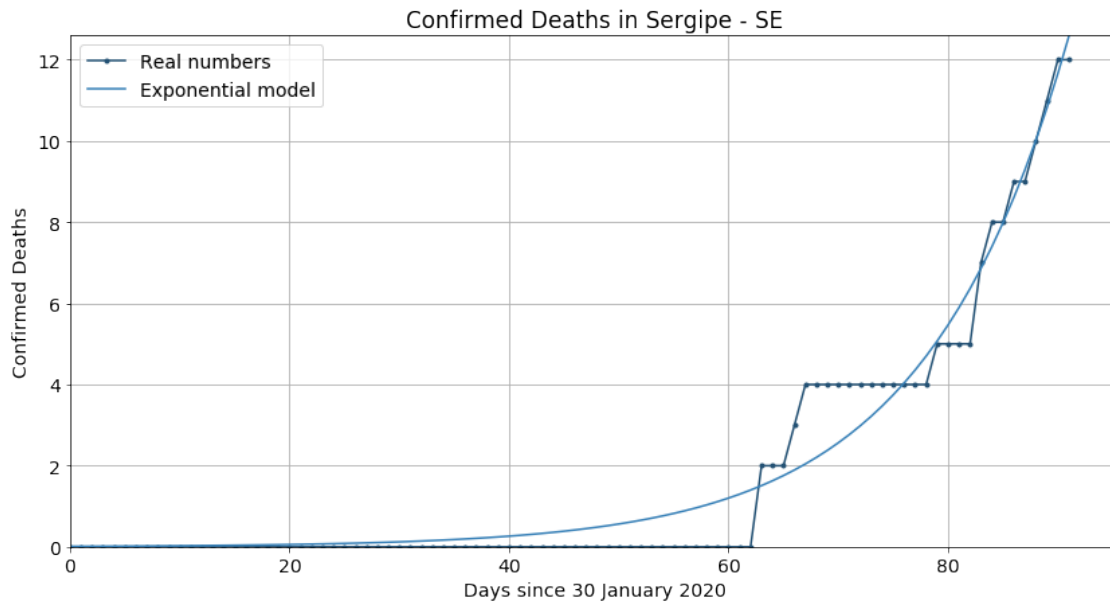
Next estimated number of confirmed cases in Santa Catarina is: 2038.004 or + -46.996



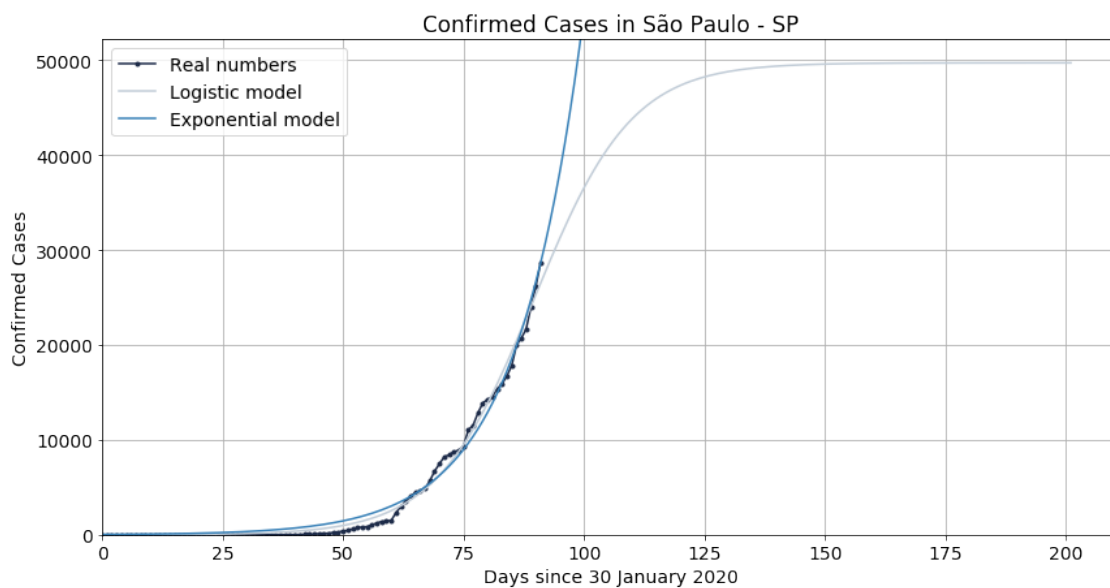
Next estimated number of confirmed cases in Santa Catarina is: 59.403 or + 13.403



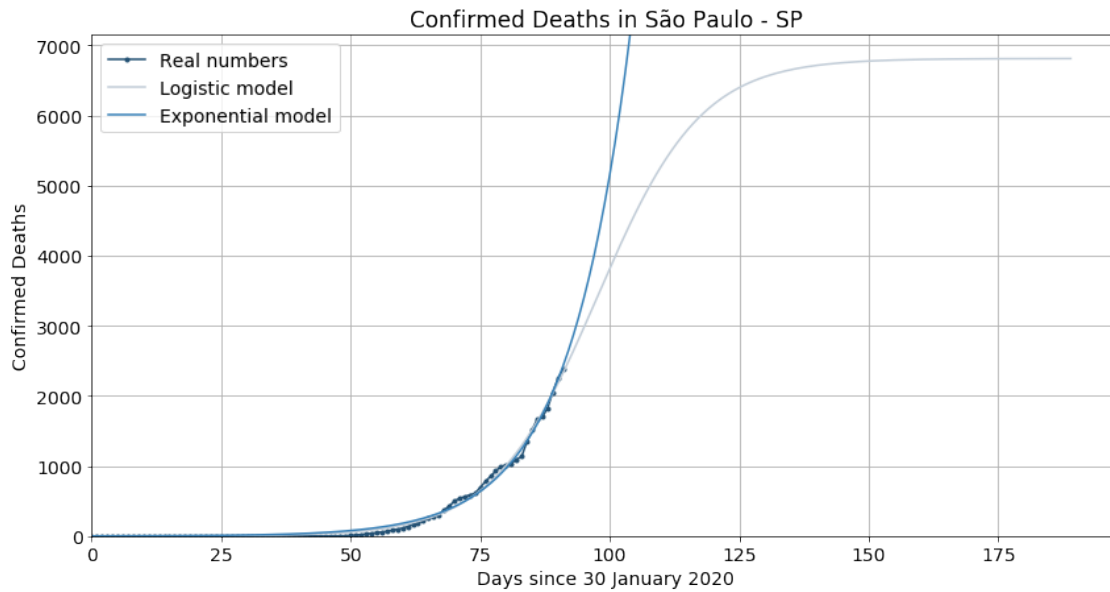
Next estimated number of confirmed cases in Sergipe is: 453.196 or + 6.196



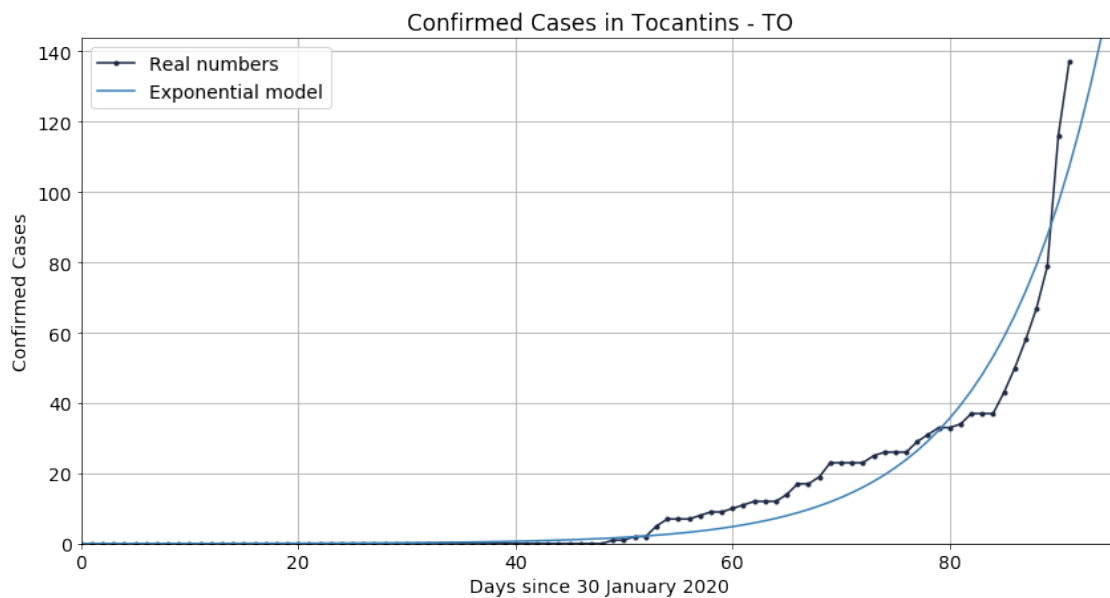
Next estimated number of confirmed cases in Sergipe is: 13.565 or + 1.565



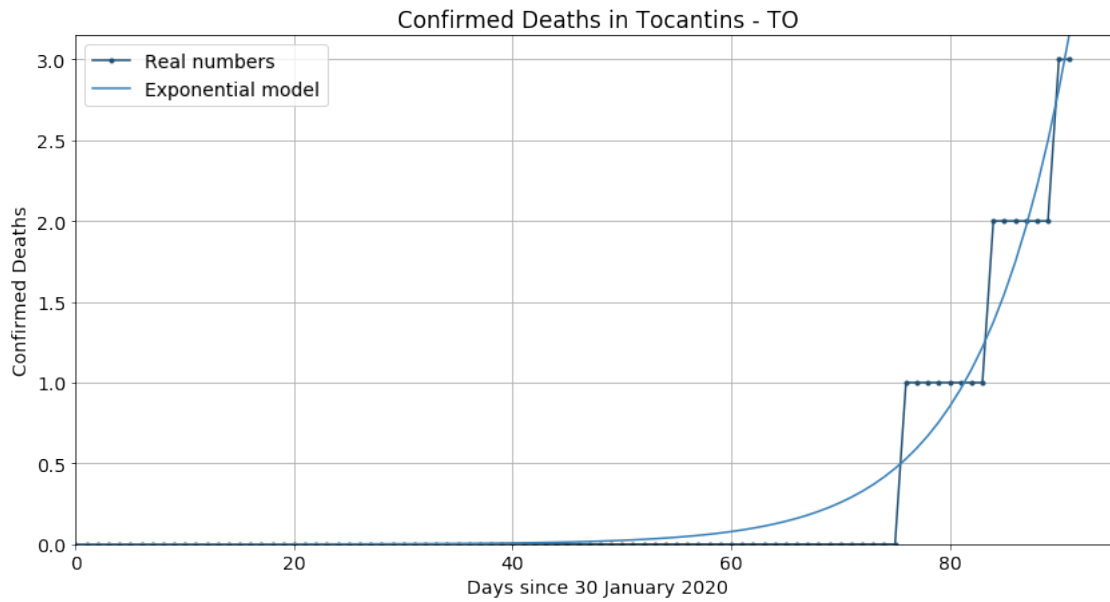
Next estimated number of confirmed cases in São Paulo is: 30813.583 or + 2115.583



Next estimated number of confirmed cases in São Paulo is: 2639.596 or + 264.596



Next estimated number of confirmed cases in Tocantins is: 118.367 or + -18.633



Next estimated number of confirmed cases in Tocantins is: 3.574 or + 0.574

[]: