Carregando os dados

Out[5]:

A primeira etapa vai ser juntar os 5 csvs diferentes em um dataframe só

```
In [1]:
          import pandas as pd
In [2]:
         df = pd.concat([pd.read_csv('HIST_PAINEL_COVIDBR_2020_Parte1_27jan2022.csv', delimiter=';'),
                           pd.read_csv('HIST_PAINEL_COVIDBR_2020_Parte2_27jan2022.csv', delimiter=';'),
                           pd.read_csv('HIST_PAINEL_COVIDBR_2021_Parte1_27jan2022.csv', delimiter=';'),
                           pd.read_csv('HIST_PAINEL_COVIDBR_2021_Parte2_27jan2022.csv', delimiter=';')
                           pd.read_csv('HIST_PAINEL_COVIDBR_2022_Partel_27jan2022.csv', delimiter=';')],
                           ignore index=True)
In [3]:
                   regiao estado municipio coduf
                                                    codmun codRegiaoSaude nomeRegiaoSaude
                                                                                                   data semanaEpi populacaoTCU20
                                                                                                  2020-
                                                                                                                            21014712
                    Brasil
                             NaN
                                        NaN
                                                 76
                                                         NaN
                                                                          NaN
                                                                                             NaN
                                                                                                  02-25
                                                                                                  2020-
                1
                    Brasil
                              NaN
                                        NaN
                                                 76
                                                         NaN
                                                                          NaN
                                                                                             NaN
                                                                                                                  9
                                                                                                                            21014712
                                                                                                  02-26
                                                                                                  2020-
                    Brasil
                                                76
                                                                                                                  9
                                                                                                                            21014712
                2
                             NaN
                                        NaN
                                                         NaN
                                                                          NaN
                                                                                             NaN
                                                                                                  02-27
                                                                                                  2020-
                                                 76
                                                                                                                  9
                                                                                                                            21014712
                3
                    Brasil
                                        NaN
                                                         NaN
                                                                          NaN
                                                                                                  02-28
                                                                                                  2020-
                                                                                                                  9
                                                                                                                            21014712
                4
                                        NaN
                                                76
                                                                          NaN
                    Brasil
                             NaN
                                                         NaN
                                                                                             NaN
                   Centro-
                                                                                                  2022-
         3776831
                               DF
                                      Brasília
                                                53 530010.0
                                                                       53001.0
                                                                                 DISTRITO FEDERAL
                                                                                                                  4
                                                                                                                              301526
                    Oeste
                   Centro-
                                                                                                  2022-
         3776832
                                      Brasília
                                                                       53001.0
                                                                                 DISTRITO FEDERAL
                                                                                                                              301526
                               DF
                                                53 530010.0
                                                                                                                  4
                                                                                                  01-24
                   Centro-
                                                                                                  2022-
         3776833
                                      Brasília
                                                                       53001.0
                                                                                 DISTRITO FEDERAL
                                                                                                                  4
                                                                                                                              301526
                                                    530010.0
                                                                                                  01-25
                    Oeste
                   Centro-
         3776834
                                                                                 DISTRITO FEDERAL
                               DF
                                      Brasília
                                                    530010.0
                                                                       53001.0
                                                                                                                              301526
                    Oeste
                                                                                                  01-26
                   Centro-
                                                                                                  2022-
         3776835
                               DF
                                      Brasília
                                                53 530010.0
                                                                       53001.0
                                                                                 DISTRITO FEDERAL
                                                                                                                  4
                                                                                                                              301526
                    Oeste
                                                                                                  01-27
        3776836 rows × 17 columns
In [4]:
          df.describe()
Out[4]:
                                                                semanaEpi populacaoTCU2019 casosAcumulado
                                   codmun codRegiaoSaude
                                                                                                                   casosNovos obito
         count 3.776836e+06 3.757152e+06
                                               3.743040e+06 3.776836e+06
                                                                                 3.762724e+06
                                                                                                  3.776836e+06
                                                                                                                 3.776836e+06
                                                                                                                                   3
         mean 3.236002e+01 3.252580e+05
                                               3.240312e+04 2.843944e+01
                                                                                 1.160557e+05
                                                                                                  6.132050e+03
                                                                                                                 1.967110e+01
                                                                                                                                   1
               9.874906e+00 9.853504e+04
                                               9.836343e+03 1.474039e+01
                                                                                 3.001490e+06
                                                                                                  1.950728e+05
                                                                                                                 6.805280e+02
                                                                                                                                   5
           min
               1.100000e+01 1.100000e+05
                                               1.100100e+04 1.000000e+00
                                                                                 7.810000e+02
                                                                                                  0.000000e + 00
                                                                                                                 -5.942900e+04
                                                                                                                                   0
               2.500000e+01 2.512000e+05
                                               2.501000e+04 1.700000e+01
                                                                                 5.474000e+03
                                                                                                   7.900000e+01
                                                                                                                 0.000000e+00
                                                                                                                                   1
               3.100000e+01 3.146100e+05
                                               3.105900e+04 2.900000e+01
                                                                                 1.169500e+04
                                                                                                  3.760000e+02
                                                                                                                 0.000000e+00
          75% 4.100000e+01 4.119150e+05
                                               4.101500e+04 4.100000e+01
                                                                                 2.576500e+04
                                                                                                  1.182000e+03
                                                                                                                 2.000000e+00
                                                                                                                                   2
                                                                                                  2.476484e+07
                                                                                                                                   6
          max 7.600000e+01 5.300100e+05
                                               5.300100e+04 5.300000e+01
                                                                                 2.101471e+08
                                                                                                                 2.289540e+05
In [5]:
         len(df['estado'].unique())
```

```
In [6]: len(df['coduf'].unique())
```

Tratando os dados

Data Cleaning

Removendo informações não úteis

Primeiramente vamos remover os dados duplicados ou que não servem para fazer análises. (coduf, codmun, codRegiaoSaude, nomeRegiaoSaude). Isso é importante por se tratar de uma database com um grandíssimo número de dados, remover dados duplicados ou desnecessários melhora a perfomance na hora de fazer as análises.

```
In [7]:
          df1 = df.drop(['coduf', 'codmun', 'codRegiaoSaude', 'nomeRegiaoSaude'], axis = 1)
                                                data semanaEpi populacaoTCU2019 casosAcumulado casosNovos obitosAcumulado o
                    regiao estado municipio
Out[7]:
                                                2020-
                                                                           210147125.0
                                                                                                                     0
                                                                                                                                        0
                 0
                     Brasil
                               NaN
                                           NaN
                                                                 9
                                                                                                      0.0
                                                02-25
                                                2020-
                                                                           210147125.0
                                                                                                                     1
                      Brasil
                                                                                                      1.0
                                                                                                                                         0
                 1
                               NaN
                                           NaN
                                                02-26
                                                2020-
                      Brasil
                               NaN
                                           NaN
                                                                 9
                                                                           210147125.0
                                                                                                      1.0
                                                                                                                     0
                                                                                                                                         0
                                                02-27
                                                2020-
                                                                           210147125.0
                                                                                                                     0
                      Brasil
                               NaN
                                           NaN
                                                                                                      1.0
                                                02-28
                                                2020-
                                                                 9
                                                                                                                                         0
                     Brasil
                               NaN
                                           NaN
                                                                           210147125.0
                                                                                                      2.0
                                                                                                                     1
                                                02-29
                    Centro-
                                                2022-
         3776831
                                DF
                                        Brasília
                                                                 4
                                                                             3015268.0
                                                                                                 561457.0
                                                                                                                     0
                                                                                                                                    11139
                     Oeste
                                                01-23
                                                2022-
                    Centro-
         3776832
                                        Brasília
                                                                             3015268.0
                                                                                                 568433.0
                                                                                                                  6976
                                                                                                                                    11143
                                                01-24
                     Oeste
                    Centro-
                                                2022-
         3776833
                                 DF
                                        Brasília
                                                                 4
                                                                             3015268.0
                                                                                                 579130.0
                                                                                                                 10697
                                                                                                                                    11147
                                                01-25
                     Oeste
                    Centro-
                                                2022-
         3776834
                                 DF
                                        Brasília
                                                                             3015268.0
                                                                                                 584615.0
                                                                                                                  5485
                                                                                                                                    11152
                     Oeste
                                                01-26
                                                2022-
                    Centro-
         3776835
                                 DF
                                        Brasília
                                                                 4
                                                                             3015268.0
                                                                                                 589858.0
                                                                                                                  5243
                                                                                                                                    11156
                                                01-27
                     Oeste
         3776836 rows × 13 columns
```

```
In [8]:
         df1['Recuperadosnovos'].dropna()
        54
                       22130.0
Out[8]:
        55
                       22991.0
         56
                       24325.0
         57
                       25318.0
                       26573.0
         58
         3625145
                    21849545.0
        3625146
                    21834758.0
         3625147
                    21926277.0
         3625148
                    22036168.0
         3625149
                    22098157.0
        Name: Recuperadosnovos, Length: 649, dtype: float64
```

```
3625145 1571613.0
3625146 1669481.0
3625147 1761197.0
3625148 1875303.0
3625149 2041596.0
```

Name: emAcompanhamentoNovos, Length: 649, dtype: float64

Essas duas colunas, 'Recuperadosnovos' e 'emAcompanhamentoNovos' só possuem 649 dados úteis em uma planilha com mais de 3 milhões, portanto vamos removê-las por interferir em quase nada.

In [10]: df['interior/metropolitana'].dropna().astype(bool).describe()

Out[11]

Name: interior/metropolitana, dtype: object

In [11]: df[(df['municipio'] == 'São Paulo')]

:		regiao	estado	municipio	coduf	codmun	codRegiaoSaude	nomeRegiaoSaude	data	semanaEpi	populacaoTCU2
	492739	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2020- 03-27	13	1225207
	492740	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2020- 03-28	13	1225202
	492741	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2020- 03-29	14	1225202
	492742	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2020- 03-30	14	1225207
	492743	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2020- 03-31	14	1225207
	3729716	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2022- 01-23	4	1225202
	3729717	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2022- 01-24	4	1225202
	3729718	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2022- 01-25	4	1225202
	3729719	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2022- 01-26	4	1225202
	3729720	Sudeste	SP	São Paulo	35	355030.0	35016.0	SAO PAULO	2022- 01-27	4	1225207
6	72 rows	× 17 colu	umns								

In [12]: df[df['regiao'] == 'Brasil']

Out[12]:		regiao	estado	municipio	coduf	codmun	codRegiaoSaude	nomeRegiaoSaude	data	semanaEpi	populacaoTCU20
	0	Brasil	NaN	NaN	76	NaN	NaN	NaN	2020- 02-25	9	210147125
	1	Brasil	NaN	NaN	76	NaN	NaN	NaN	2020- 02-26	9	210147125
	2	Brasil	NaN	NaN	76	NaN	NaN	NaN	2020- 02-27	9	210147125
	3	Brasil	NaN	NaN	76	NaN	NaN	NaN	2020- 02-28	9	210147125
	4	Brasil	NaN	NaN	76	NaN	NaN	NaN	2020- 02-29	9	210147125
	3625145	Brasil	NaN	NaN	76	NaN	NaN	NaN	2022- 01-23	4	210147125
	3625146	Brasil	NaN	NaN	76	NaN	NaN	NaN	2022- 01-24	4	210147125
									2022		

```
2022-
          3625148
                                                                                                                               210147125
                     Brasil
                                          NaN
                                                                            NaN
                                                                                                     01-26
                                                                                                     2022-
          3625149
                     Brasil
                               NaN
                                          NaN
                                                  76
                                                          NaN
                                                                            NaN
                                                                                                NaN
                                                                                                                     4
                                                                                                                               210147125
                                                                                                     01-27
         703 rows \times 17 columns
In [13]:
           df removed
                       = df1.drop(['Recuperadosnovos', 'emAcompanhamentoNovos'], axis = 1)
           df removed
Out[13]:
                    regiao estado municipio
                                                data semanaEpi populacaoTCU2019 casosAcumulado casosNovos
                                                                                                                    obitosAcumulado (
                                               2020-
                                                                                                                  0
                                                                                                                                     0
                 0
                      Brasil
                               NaN
                                          NaN
                                                                         210147125.0
                                                                                                    0.0
                                               02-25
                                                2020-
                                                               9
                 1
                      Brasil
                               NaN
                                          NaN
                                                                         210147125.0
                                                                                                    1.0
                                                                                                                  1
                                                                                                                                    0
                                               02-26
                                               2020-
                 2
                                                               9
                                                                         210147125.0
                                                                                                    1.0
                                                                                                                  0
                                                                                                                                    0
                      Brasil
                                          NaN
                               NaN
                                               02-27
                                               2020-
                      Brasil
                               NaN
                                          NaN
                                                                         210147125.0
                                                                                                    1.0
                                                                                                                  0
                                               02-28
                                                2020-
                                                                                                                                    0
                                                               9
                                                                         210147125.0
                                                                                                    2.0
                                                                                                                  1
                 4
                      Brasil
                               NaN
                                          NaN
                                                02-29
                    Centro-
                                               2022-
          3776831
                                        Brasília
                                 DF
                                                               4
                                                                           3015268.0
                                                                                              561457.0
                                                                                                                  0
                                                                                                                                11139
                     Oeste
                                               01-23
                    Centro-
                                                2022-
          3776832
                                                                           3015268.0
                                                                                                               6976
                                DF
                                        Brasília
                                                               4
                                                                                              568433.0
                                                                                                                                11143
                     Oeste
                                               01-24
                                               2022-
                    Centro-
          3776833
                                 DF
                                        Brasília
                                                                4
                                                                            3015268.0
                                                                                              579130.0
                                                                                                              10697
                                                                                                                                11147
                                               01-25
                     Oeste
                                               2022-
          3776834
                                 DF
                                        Brasília
                                                                            3015268.0
                                                                                              584615.0
                                                                                                               5485
                                                                                                                                11152
                                               01-26
                      Oeste
                    Centro-
                                                2022-
          3776835
                                        Brasília
                                 DF
                                                               4
                                                                            3015268.0
                                                                                              589858.0
                                                                                                               5243
                                                                                                                                11156
                     Oeste
                                               01-27
         3776836 rows × 11 columns
In [14]:
           df_removed[df_removed['regiao'] == 'Brasil']['casosNovos'].sum()
          24764838
Out[14]:
In [15]:
           df_removed[df_removed['regiao'] != 'Brasil']['casosNovos'].sum()
          49529676
Out[15]:
In [16]:
           df_removed['casosNovos'].sum()
          74294514
Out[16]:
In [17]:
           df_removed[df_removed['estado'] == 'AC']['casosNovos'].sum()
          193668
Out[17]:
In [18]:
           df_removed[df_removed['estado'] == 'AC']
                                                data semanaEpi populacaoTCU2019 casosAcumulado casosNovos obitosAcumulado
Out[18]:
                    regiao estado municipio
                                               2020-
               316
                                          NaN
                                                                            881935.0
                                                                                                                  0
                                                                                                                                    0
                     Norte
                                AC
                                                               9
                                                                                                   0.0
                                               02-25
```

3625147

Brasil

NaN

NaN

76

NaN

NaN

NaN 01-25

210147125

317	Norte	AC	NaN	2020- 02-26	9	881935.0	0.0	0	0
318	Norte	AC	NaN	2020- 02-27	9	881935.0	0.0	0	0
319	Norte	AC	NaN	2020- 02-28	9	881935.0	0.0	0	0
320	Norte	AC	NaN	2020- 02-29	9	881935.0	0.0	0	0
3627899	Norte	AC	Porto Acre	2022- 01-23	4	18504.0	1611.0	14	38
3627900	Norte	AC	Porto Acre	2022- 01-24	4	18504.0	1615.0	4	38
3627901	Norte	AC	Porto Acre	2022- 01-25	4	18504.0	1618.0	3	38
3627902	Norte	AC	Porto Acre	2022- 01-26	4	18504.0	1652.0	34	38
3627903	Norte	AC	Porto Acre	2022- 01-27	4	18504.0	1654.0	2	38

15487 rows × 11 columns

		regiao	estado	municipio	data	semanaEpi	populacaoTCU2019	casosAcumulado	casosNovos	obitosAcumulado	O
362	5145	Brasil	NaN	NaN	2022- 01-23	4	210147125.0	24044255.0	135080	623097	
362	5146	Brasil	NaN	NaN	2022- 01-24	4	210147125.0	24127595.0	83340	623356	
362	5147	Brasil	NaN	NaN	2022- 01-25	4	210147125.0	24311317.0	183722	623843	
362	5148	Brasil	NaN	NaN	2022- 01-26	4	210147125.0	24535884.0	224567	624413	
362	5149	Brasil	NaN	NaN	2022- 01-27	4	210147125.0	24764838.0	228954	625085	
4											

Após essa análise, foi percebido que a planilha fornecida é separada em três tipos de dados: o primeiro é o número de casos de COVID no Brasil, o segundo é o n de casos por estado e o terceiro por município. Esses dados são iguais em número de casos/óbitos. Portanto, ao calcular o número total de casos, obtemos o triplo do número de casos. Para evitar isso, iremos considerar apenas o terceiro tipo de dados e remover o primeiro e segundo.

In [20]: df_fully_removed = df_removed.dropna()

fn [21]: df_fully_removed

ut[21]:		regiao	estado	municipio	data	semanaEpi	populacaoTCU2019	casosAcumulado	casosNovos	obitosAcumulado
	4551	Norte	RO	Alta Floresta D'Oeste	2020- 03-27	13	22945.0	0.0	0	0
	4552	Norte	RO	Alta Floresta D'Oeste	2020- 03-28	13	22945.0	0.0	0	0
	4553	Norte	RO	Alta Floresta D'Oeste	2020- 03-29	14	22945.0	0.0	0	0
	4554	Norte	RO	Alta Floresta D'Oeste	2020- 03-30	14	22945.0	0.0	0	0
	4555	Norte	RO	Alta Floresta D'Oeste	2020- 03-31	14	22945.0	0.0	0	0
	3776831	Centro- Oeste	DF	Brasília	2022- 01-23	4	3015268.0	561457.0	0	11139

```
Centro-
                                                2022-
                                        Brasília
          3776833
                                 DF
                                                                4
                                                                            3015268.0
                                                                                               579130.0
                                                                                                               10697
                                                                                                                                 11147
                                                01-25
                                                2022-
                    Centro-
          3776834
                                        Brasília
                                                                            3015268.0
                                                                                               584615.0
                                                                                                                5485
                                                                                                                                 11152
                                                01-26
                      Oeste
                    Centro-
                                                2022-
                                        Brasília
          3776835
                                 DF
                                                                4
                                                                            3015268.0
                                                                                               589858.0
                                                                                                                5243
                                                                                                                                 11156
                                                01-27
                      Oeste
         3743040 rows × 11 columns
In [22]:
           df_fully_removed['casosNovos'].sum()
          24712902
Out[22]:
In [23]:
           df_fully_removed = df_fully_removed.reset_index(drop=True)
In [24]:
           df_fully_removed
Out[24]:
                     regiao estado municipio
                                                data semanaEpi populacaoTCU2019 casosAcumulado casosNovos obitosAcumulado c
                                           Alta
                                                2020-
                 0
                      Norte
                                RO
                                       Floresta
                                                               13
                                                                              22945.0
                                                                                                    0.0
                                                                                                                   0
                                                                                                                                      0
                                                03-27
                                       D'Oeste
                                           Alta
                                                2020-
                                RO
                                       Floresta
                                                               13
                                                                              22945.0
                                                                                                     0.0
                                                                                                                   0
                                                                                                                                      0
                      Norte
                                                03-28
                                       D'Oeste
                                           Alta
                                                2020-
                                                                                                                   0
                                                                                                                                      0
                 2
                      Norte
                                RO
                                       Floresta
                                                               14
                                                                              22945.0
                                                                                                    0.0
                                                03-29
                                       D'Oeste
                                           Alta
                                                2020-
                                RO
                                                                              22945.0
                                                                                                                   0
                                                                                                                                      0
                                       Floresta
                                                               14
                                                                                                    0.0
                 3
                      Norte
                                                03-30
                                       D'Oeste
                                           Alta
                                                2020-
                                RO
                                                               14
                                                                              22945.0
                                                                                                     0.0
                                                                                                                   0
                                                                                                                                      0
                      Norte
                                       Floresta
                                                03-31
                                       D'Oeste
                                                2022-
                    Centro-
          3743035
                                 DF
                                        Brasília
                                                                4
                                                                            3015268.0
                                                                                               561457.0
                                                                                                                   0
                                                                                                                                 11139
                                                01-23
                      Oeste
                    Centro-
                                                2022-
          3743036
                                 DF
                                        Brasília
                                                                            3015268.0
                                                                                               568433.0
                                                                                                                6976
                                                                                                                                 11143
                                                01-24
                                                2022-
                    Centro-
          3743037
                                                                            3015268.0
                                        Brasília
                                                                                               579130.0
                                                                                                               10697
                                                                                                                                 11147
                                                01-25
                      Oeste
                                                2022-
                    Centro-
          3743038
                                 DF
                                        Brasília
                                                                            3015268.0
                                                                                               584615.0
                                                                                                                5485
                                                                                                                                 11152
                                                01-26
                     Oeste
                    Centro-
                                                2022-
          3743039
                                 DF
                                        Brasília
                                                                4
                                                                            3015268.0
                                                                                               589858.0
                                                                                                                5243
                                                                                                                                 11156
                      Oeste
                                                01-27
         3743040 rows × 11 columns
In [25]:
           import time
In [26]:
           # start time = time.time()
           # df fully removed.to csv('DadosCovidTratados.csv')
           # print("--- %s seconds ---" % (time.time() - start_time))
         Convert numbers to use less memory space
In [27]:
           df fully removed = df fully removed.astype({'interior/metropolitana' : bool, 'casosAcumulado' : int,
                                                            'populacaoTCU2019' : int});
In [28]:
           df_fully_removed
```

3015268.0

568433.0

6976

11143

Centro-

Oeste

DF

3776832

Out[28]:

2022-

01-24

Brasília

	regiao	estado	municipio	data	semanaEpi	populacaoTCU2019	casosAcumulado	casosNovos	obitosAcumulado
0	Norte	RO	Alta Floresta D'Oeste	2020- 03-27	13	22945	0	0	0
1	Norte	RO	Alta Floresta D'Oeste	2020- 03-28	13	22945	0	0	0
2	Norte	RO	Alta Floresta D'Oeste	2020- 03-29	14	22945	0	0	0
3	Norte	RO	Alta Floresta D'Oeste	2020- 03-30	14	22945	0	0	0
4	Norte	RO	Alta Floresta D'Oeste	2020- 03-31	14	22945	0	0	0
3743035	Centro- Oeste	DF	Brasília	2022- 01-23	4	3015268	561457	0	11139
3743036	Centro- Oeste	DF	Brasília	2022- 01-24	4	3015268	568433	6976	11143
3743037	Centro- Oeste	DF	Brasília	2022- 01-25	4	3015268	579130	10697	11147
3743038	Centro- Oeste	DF	Brasília	2022- 01-26	4	3015268	584615	5485	11152
3743039	Centro- Oeste	DF	Brasília	2022- 01-27	4	3015268	589858	5243	11156
3743040 ı	ows × 1	.1 colum	ns						

```
In [29]: # start_time = time.time()
# df_fully_removed.to_csv('DadosCovidTratados2.csv')
# print("--- %s seconds ---" % (time.time() - start_time))
```

In [30]: df_no_accumulated = df_fully_removed.drop(['casosAcumulado', 'obitosAcumulado'], axis=1)

In [31]: df_no_accumulated

out[31]:		regiao	estado	municipio	data	semanaEpi	populacaoTCU2019	casosNovos	obitosNovos	interior/metropolitana
	0	Norte	RO	Alta Floresta D'Oeste	2020- 03-27	13	22945	0	0	False
	1	Norte	RO	Alta Floresta D'Oeste	2020- 03-28	13	22945	0	0	False
	2	Norte	RO	Alta Floresta D'Oeste	2020- 03-29	14	22945	0	0	False
	3	Norte	RO	Alta Floresta D'Oeste	2020- 03-30	14	22945	0	0	False
	4	Norte	RO	Alta Floresta D'Oeste	2020- 03-31	14	22945	0	0	False
	3743035	Centro- Oeste	DF	Brasília	2022- 01-23	4	3015268	0	0	True
	3743036	Centro- Oeste	DF	Brasília	2022- 01-24	4	3015268	6976	4	True
	3743037	Centro- Oeste	DF	Brasília	2022- 01-25	4	3015268	10697	4	True
	3743038	Centro- Oeste	DF	Brasília	2022- 01-26	4	3015268	5485	5	True
	3743039	Centro- Oeste	DF	Brasília	2022- 01-27	4	3015268	5243	4	True

```
In [32]:
         brasil_pop = int(df_removed['populacaoTCU2019'][0])
         brasil pop
        210147125
Out[32]:
In [33]:
         df_removed[df_removed['estado'] == 'RO']['populacaoTCU2019'].iloc[0]
        1777225.0
Out[33]:
In [34]:
         no brasil = df removed[df removed['regiao'] != 'Brasil']
         all_states = no_brasil['estado'].dropna().unique()
         states_pop = []
         for state in all_states:
             states_pop.append(int(no_brasil[no_brasil['estado'] == state]['populacaoTCU2019'].iloc[0]))
In [35]:
         all_states, states_pop
        Out[35]:
          [1777225,
          881935,
          4144597,
          605761,
          8602865.
          845731,
          1572866,
           7075181,
          3273227,
          9132078,
           3506853,
           4018127,
          9557071,
           3337357,
           2298696,
           14873064,
           21168791,
           4018650,
          17264943
          45919049,
           11433957,
           7164788,
           11377239
           2778986,
           3484466,
           7018354,
          3015268])
In [36]:
         states_pop_df = pd.DataFrame(states_pop, index = all_states, columns=['Populacao'])
In [37]:
         states_pop_df
             Populacao
         RO
               1777225
         AC
                881935
         ΑM
               4144597
         RR
                605761
               8602865
         PA
                845731
         AP
               1572866
         то
         MΑ
               7075181
          ΡI
               3273227
         CE
               9132078
```

```
RN
      3506853
РΒ
      4018127
      9557071
PE
      3337357
AL
SE
      2298696
BA
     14873064
MG
     21168791
ES
      4018650
RJ
     17264943
     45919049
SP
PR
     11433957
SC
      7164788
RS
     11377239
      2778986
MS
MT
      3484466
      7018354
GO
DF
      3015268
```

Agora população para cada município

```
In [38]:
          start time = time.time()
          hundredthousand_time = time.time()
          actual_city = ''
          cities_population = {}
          for index in df_no_accumulated.index:
              if(df_no_accumulated['municipio'].loc[index] != actual_city):
                  actual_city = df_no_accumulated['municipio'].loc[index]
                  actual_city_population = df_no_accumulated.loc[index]['populacaoTCU2019']
                  cities_population[actual_city] = actual_city_population
              if(index % 100000 == 0):
                  print(index)
                  print("--- %s seconds ---" % (time.time() - hundredthousand_time))
                  hundredthousand time = time.time()
          print("--- %s seconds ---" % (time.time() - start_time))
         --- 0.0006864070892333984 seconds ---
         100000
         --- 0.8014662265777588 seconds ---
         200000
         --- 0.7769520282745361 seconds ---
         300000
          --- 0.7794997692108154 seconds ---
         400000
         --- 0.7835428714752197 seconds ---
         500000
         --- 0.7789051532745361 seconds ---
         600000
         --- 0.778876781463623 seconds ---
         700000
         --- 0.7804036140441895 seconds ---
         800000
         --- 0.7785589694976807 seconds ---
         900000
         --- 0.7704124450683594 seconds ---
         1000000
         --- 0.7704691886901855 seconds ---
         --- 0.7694189548492432 seconds ---
         1200000
         --- 0.7714345455169678 seconds ---
         1300000
         --- 0.7754995822906494 seconds ---
         1400000
          --- 0.7715840339660645 seconds ---
         1500000
         --- 0.7808136940002441 seconds ---
         1600000
         --- 0.7829239368438721 seconds ---
         1700000
          -- 0.8046474456787109 seconds ---
```

```
1800000
--- 0.8541445732116699 seconds ---
1900000
--- 0.8634541034698486 seconds ---
2000000
--- 0.7602965831756592 seconds ---
2100000
--- 0.7591595649719238 seconds ---
2200000
--- 0.7630362510681152 seconds ---
2300000
--- 0.7518496513366699 seconds ---
2400000
 -- 0.7526981830596924 seconds ---
2500000
--- 0.7519717216491699 seconds ---
2600000
--- 0.7616653442382812 seconds ---
2700000
--- 0.7607686519622803 seconds ---
2800000
--- 0.777000904083252 seconds ---
2900000
--- 0.768500804901123 seconds ---
3000000
--- 0.7945687770843506 seconds ---
3100000
--- 0.7716050148010254 seconds ---
3200000
--- 0.7628188133239746 seconds ---
3300000
--- 0.7658531665802002 seconds ---
3400000
--- 0.7669014930725098 seconds ---
3500000
--- 0.76566481590271 seconds ---
3600000
--- 0.8105733394622803 seconds ---
3700000
--- 1.0465588569641113 seconds ---
--- 29.54274892807007 seconds ---
```

In [39]:

cities_population

```
Out[39]: {"Alta Floresta D'Oeste": 22945,
           'Ariquemes': 107863,
           'Cabixi': 5312,
'Cacoal': 85359,
           'Cerejeiras': 16323,
           'Colorado do Oeste': 15882,
           'Corumbiara': 7391,
           'Costa Marques': 18331,
           "Espigão D'Oeste": 32374,
           'Guajará-Mirim': 46174,
           'Jaru': 51775,
           'Ji-Paraná': 128969,
           "Machadinho D'Oeste": 39991,
"Nova Brasilândia D'Oeste": 20474,
           'Ouro Preto do Oeste': 36035,
           'Pimenta Bueno': 36660,
           'Porto Velho': 529544,
           'Presidente Médici': 7015,
           'Rio Crespo': 3764,
           'Rolim de Moura': 55058,
           "Santa Luzia D'Oeste": 6495,
           'Vilhena': 99854,
           'São Miguel do Guaporé': 23005,
           'Nova Mamoré': 30583,
           "Alvorada D'Oeste": 14411,
           'Alto Alegre dos Parecis': 13241,
           'Alto Paraíso': 2741,
           'Buritis': 24841,
           'Novo Horizonte do Oeste': 8538,
           'Cacaulândia': 6230,
           'Campo Novo de Rondônia': 14139,
           'Candeias do Jamari': 26693,
           'Castanheiras': 3052,
           'Chupinguaia': 11182,
           'Cujubim': 25215,
           'Governador Jorge Teixeira': 7767,
           'Itapuã do Oeste': 10458,
           'Ministro Andreazza': 9660,
           'Mirante da Serra': 10947,
           'Monte Negro': 15852,
           'Nova União': 5725,
```

```
'Parecis': 6074,
'Pimenteiras do Oeste': 2169,
'Primavera de Rondônia': 2856,
"São Felipe D'Oeste": 5172,
'São Francisco do Guaporé': 20266,
'Seringueiras': 11856,
'Teixeirópolis': 4308,
'Theobroma': 10444,
'Urupá': 11467,
'Vale do Anari': 11204,
'Vale do Paraíso': 6825,
'Acrelândia': 15256,
'Assis Brasil': 7417,
'Brasiléia': 26278,
'Bujari': 10266,
'Capixaba': 11733,
'Cruzeiro do Sul': 12348,
'Epitaciolândia': 18411,
'Feijó': 34780,
'Jordão': 8317,
'Mâncio Lima': 18977,
'Manoel Urbano': 9459,
'Marechal Thaumaturgo': 18867,
'Plácido de Castro': 19761,
'Porto Walter': 11982,
'Rio Branco': 5156,
'Rodrigues Alves': 18930,
'Santa Rosa do Purus': 6540,
'Senador Guiomard': 23024,
'Sena Madureira': 45848,
'Tarauacá': 42567,
'Xapuri': 19323,
'Porto Acre': 18504,
'Alvarães': 16041,
'Amaturá': 11536,
'Anamã': 13614,
'Anori': 21010,
'Apuí': 21973,
'Atalaia do Norte': 19921,
'Autazes': 39565,
'Barcelos': 27502
'Barreirinha': 32041,
'Benjamin Constant': 42984,
'Beruri': 19679,
'Boa Vista do Ramos': 19207,
'Boca do Acre': 34308,
'Borba': 41161,
'Caapiranga': 13081,
'Canutama': 15629,
'Carauari': 28294,
'Careiro': 37869,
'Careiro da Várzea': 30225,
'Coari': 85097,
'Codajás': 28637,
'Eirunepé': 35273,
'Envira': 20033,
'Fonte Boa': 17609,
'Guajará': 16678,
'Humaitá': 4762,
'Ipixuna': 29689
'Iranduba': 48296
'Itacoatiara': 101337,
'Itamarati': 7851,
'Itapiranga': 16872,
'Japurá': 9425,
'Juruá': 14712,
'Jutaí': 14317,
'Lábrea': 46069.
'Manacapuru': 97377,
'Manaquiri': 32105,
'Manaus': 2182763,
'Manicoré': 55751,
'Maraã': 18224,
'Maués': 63905,
'Nhamundá': 21173,
'Nova Olinda do Norte': 37378,
'Novo Airão': 19454,
'Novo Aripuanã': 25644,
'Parintins': 114273,
'Pauini': 19426,
'Presidente Figueiredo': 36279,
'Rio Preto da Eva': 33347,
'Santa Isabel do Rio Negro': 25156,
'Santo Antônio do Içá': 21602,
'São Gabriel da Cachoeira': 45564,
'São Paulo de Olivença': 39299,
'São Sebastião do Uatumã': 14020,
'Silves': 9171,
'Tabatinga': 16496,
```

```
'Tapauá': 17156,
'Tefé': 59849,
'Tonantins': 18755,
'Uarini': 13540,
'Urucará': 16256,
'Urucurituba': 23065,
'Amajari': 12796,
'Alto Alegre': 1638,
'Boa Vista': 7051,
'Bonfim': 6868,
'Cantá': 18335,
'Caracaraí': 21926,
'Caroebe': 10169,
'Iracema': 14297,
'Mucajaí': 17853,
'Normandia': 11290,
'Pacaraima': 17401,
'Rorainópolis': 30163,
'São João da Baliza': 8201,
'São Luiz': 7986,
'Uiramutã': 10559,
'Abaetetuba': 157698,
'Abel Figueiredo': 7434,
'Acará': 55591,
'Afuá': 39218,
'Água Azul do Norte': 27430,
'Alenquer': 56789,
'Almeirim': 34109,
'Altamira': 114594,
'Anajás': 29277,
'Ananindeua': 530598,
'Anapu': 27890,
'Augusto Corrêa': 45998,
'Aurora do Pará': 31338,
'Aveiro': 16388,
'Bagre': 30673,
'Baião': 47446,
'Bannach': 3286,
'Barcarena': 124680,
'Belém': 4344,
'Belterra': 17732,
'Benevides': 62737,
'Bom Jesus do Tocantins': 4894,
'Bonito': 21976,
'Bragança': 127686,
'Brasil Novo': 15086,
'Brejo Grande do Araguaia': 7380,
'Breu Branco': 66046,
'Breves': 102701,
'Bujaru': 29132,
'Cachoeira do Piriá': 33900,
'Cachoeira do Arari': 23767,
'Cametá': 137890,
'Canaã dos Carajás': 37085,
'Capanema': 19124,
'Capitão Poço': 54303,
'Castanhal': 200793,
'Chaves': 23717, 'Colares': 12085,
'Conceição do Araguaia': 47864,
'Concórdia do Pará': 33318,
'Cumaru do Norte': 13473,
'Curionópolis': 17929,
'Curralinho': 34448,
'Curuá': 14393,
'Curuçá': 40066
'Dom Eliseu': 59719,
'Eldorado do Carajás': 33808,
'Faro': 7194,
'Floresta do Araguaia': 20304,
'Garrafão do Norte': 26066,
'Goianésia do Pará': 40475,
'Gurupá': 33376,
'Igarapé-Açu': 38807,
'Igarapé-Miri': 62698,
'Inhangapi': 11711,
'Ipixuna do Pará': 64053,
'Irituia': 32550,
'Itaituba': 101247,
'Itupiranga': 53269,
'Jacareacanga': 8239,
'Jacundá': 59155,
'Juruti': 57943,
'Limoeiro do Ajuru': 28935,
'Mãe do Rio': 30077,
'Magalhães Barata': 8548,
'Marabá': 279349,
'Maracanã': 29473,
'Marapanim': 28336,
```

```
'Marituba': 131521,
'Medicilândia': 31597,
'Melgaço': 27654,
'Mocajuba': 31136,
'Moju': 82094,
'Mojuí dos Campos': 16084,
'Monte Alegre': 22451,
'Muaná': 40349,
'Nova Esperança do Piriá': 21368,
'Nova Ipixuna': 16678,
'Nova Timboteua': 15363,
'Novo Progresso': 25762,
'Novo Repartimento': 75919,
'Óbidos': 52137,
'Oeiras do Pará': 32512,
'Oriximiná': 73096,
'Ourém': 17842,
'Ourilândia do Norte': 32832,
'Pacajá': 47706,
'Palestina do Pará': 7589,
'Paragominas': 113145,
'Parauapebas': 208273,
"Pau D'Arco": 4849,
'Peixe-Boi': 8077,
'Piçarra': 12981,
'Placas': 30982,
'Ponta de Pedras': 31082,
'Portel': 62043,
'Porto de Moz': 41135,
'Prainha': 29866,
'Primavera': 14966
'Quatipuru': 13608,
'Redenção': 29053,
'Rio Maria': 18193
'Rondon do Pará': 52357,
'Rurópolis': 50510,
'Salinópolis': 40675,
'Salvaterra': 23752,
'Santa Bárbara do Pará': 21079,
'Santa Cruz do Arari': 10128,
'Santa Izabel do Pará': 70801,
'Santa Luzia do Pará': 19848,
'Santa Maria das Barreiras': 21449,
'Santa Maria do Pará': 24861,
'Santana do Araguaia': 72817,
'Santarém': 304589,
'Santarém Novo': 6709,
'Santo Antônio do Tauá': 31482,
'São Caetano de Odivelas': 18050,
'São Domingos do Araguaia': 25557,
'São Domingos do Capim': 31989,
'São Félix do Xingu': 128481,
'São Francisco do Pará': 15882
'São Geraldo do Araguaia': 24847,
'São João da Ponta': 6139,
'São João de Pirabas': 23045,
'São João do Araguaia': 13996,
'São Miguel do Guamá': 58986,
'São Sebastião da Boa Vista': 26640,
'Sapucaia': 18228,
'Senador José Porfírio': 11658,
'Soure': 25374,
'Tailândia': 106339,
'Terra Alta': 11720,
'Terra Santa': 18769,
'Tomé-Açu': 63447,
'Tracuateua': 30959,
'Trairão': 18989,
'Tucumã': 39602,
'Tucuruí': 113659
'Ulianópolis': 59210,
'Uruará': 45476,
'Vigia': 53686,
'Viseu': 61403,
'Vitória do Xingu': 15134,
'Xinguara': 44751,
'Serra do Navio': 5397,
'Amapá': 9109,
'Pedra Branca do Amapari': 16502,
'Calçoene': 11117,
'Cutias': 5983,
'Ferreira Gomes': 7780,
'Itaubal': 5503,
'Laranjal do Jari': 50410,
'Macapá': 503327,
'Mazagão': 21632,
'Oiapoque': 27270,
'Porto Grande': 21971,
'Pracuúba': 5120,
```

```
'Santana': 26614,
'Tartarugalzinho': 17315,
'Vitória do Jari': 15931,
'Abreulândia': 2579,
'Aquiarnópolis': 6733.
'Aliança do Tocantins': 5390,
'Almas': 7055,
'Alvorada': 210305,
'Ananás': 9549,
'Angico': 3433,
'Aparecida do Rio Negro': 4795,
'Aragominas': 5758,
'Araguacema': 7086,
'Araguaçu': 8517,
'Araguaína': 180470,
'Araguanã': 15426,
'Araguatins': 35761,
'Arapoema': 6643,
'Arraias': 10567,
'Augustinópolis': 18412,
'Aurora do Tocantins': 3757,
'Axixá do Tocantins': 9756,
'Babaçulândia': 10664,
'Bandeirantes do Tocantins': 3553,
'Barra do Ouro': 4591,
'Barrolândia': 5632,
'Bernardo Sayão': 4459,
'Brasilândia do Tocantins': 2201,
'Brejinho de Nazaré': 5497,
'Buriti do Tocantins': 11348,
'Cachoeirinha': 130293,
'Campos Lindos': 10116,
'Cariri do Tocantins': 4382,
'Carmolândia': 2580,
'Carrasco Bonito': 4095,
'Caseara': 5369,
'Centenário': 2891,
'Chapada de Areia': 1406,
'Chapada da Natividade': 3333,
'Colinas do Tocantins': 35424,
'Combinado': 4852,
'Conceição do Tocantins': 4105,
'Couto Magalhães': 5588,
'Cristalândia': 7289,
'Crixás do Tocantins': 1722,
'Darcinópolis': 6097,
'Dianópolis': 22139,
'Divinópolis do Tocantins': 6900,
'Dois Irmãos do Tocantins': 7198,
'Dueré': 4686,
'Esperantina': 39737,
'Fátima': 17890,
'Figueirópolis': 5263,
'Filadélfia': 16377,
'Formoso do Araguaia': 18440,
'Tabocão': 2589,
'Goianorte': 5123,
'Goiatins': 13019,
'Guaraí': 25923,
'Gurupi': 86647,
'Ipueiras': 38166,
'Itacajá': 7433,
'Itaguatins': 5864,
'Itapiratins': 3777,
'Itaporã do Tocantins': 2427,
'Jaú do Tocantins': 3849,
'Juarina': 2193,
'Lagoa da Confusão': 13357,
'Lagoa do Tocantins': 4316,
'Lajeado': 84014,
'Lavandeira': 1923,
'Lizarda': 3740,
'Luzinópolis': 3108,
'Marianópolis do Tocantins': 5175,
'Mateiros': 2684,
'Maurilândia do Tocantins': 3426,
'Miracema do Tocantins': 18248,
'Miranorte': 13434,
'Monte do Carmo': 7947,
'Monte Santo do Tocantins': 2279,
'Palmeiras do Tocantins': 6658,
'Muricilândia': 3551,
'Natividade': 15317,
'Nazaré': 28525,
'Nova Olinda': 5949,
'Nova Rosalândia': 4260,
'Novo Acordo': 4342,
'Novo Alegre': 2332,
'Novo Jardim': 2722,
```

```
'Oliveira de Fátima': 1112,
'Palmeirante': 6026,
'Palmeirópolis': 7659,
'Paraíso do Tocantins': 51252,
'Paranã': 10449.
'Pedro Afonso': 13578,
'Peixe': 11749,
'Pequizeiro': 5477,
'Colméia': 8205,
'Pindorama do Tocantins': 4447,
'Piraquê': 3028,
'Pium': 7654,
'Ponte Alta do Bom Jesus': 4596,
'Ponte Alta do Tocantins': 8039,
'Porto Alegre do Tocantins': 3139,
'Porto Nacional': 53010,
'Praia Norte': 8432,
'Presidente Kennedy': 11574,
'Pugmil': 2688,
'Recursolândia': 4293,
'Riachinho': 8136,
'Rio da Conceição': 2130,
'Rio dos Bois': 2834,
'Rio Sono': 6478,
'Sampaio': 4711,
'Sandolândia': 3375,
'Santa Fé do Araguaia': 7512,
'Santa Maria do Tocantins': 3434,
'Santa Rita do Tocantins': 2365,
'Santa Rosa do Tocantins': 4829,
'Santa Tereza do Tocantins': 2866
'Santa Terezinha do Tocantins': 2529,
'São Bento do Tocantins': 5324,
'São Félix do Tocantins': 1585,
'São Miguel do Tocantins': 12139,
'São Salvador do Tocantins': 3082,
'São Sebastião do Tocantins': 4805,
'São Valério': 3960,
'Silvanópolis': 5403,
'Sítio Novo do Tocantins': 9029,
'Sucupira': 1966,
'Taguatinga': 16683,
'Taipas do Tocantins': 2148,
'Talismã': 2793,
'Palmas': 50986,
'Tocantínia': 7545,
'Tocantinópolis': 22870,
'Tupirama': 1891,
'Tupiratins': 2671
'Wanderlândia': 11683,
'Xambioá': 11540,
'Açailândia': 112445,
'Afonso Cunha': 6524,
'Água Doce do Maranhão': 12571,
'Alcântara': 22097,
'Aldeias Altas': 26532,
'Altamira do Maranhão': 8128,
'Alto Alegre do Maranhão': 27053,
'Alto Alegre do Pindaré': 31919,
'Alto Parnaíba': 11190,
'Amapá do Maranhão': 6962,
'Amarante do Maranhão': 41435,
'Anajatuba': 26803,
'Anapurus': 15732,
'Apicum-Açu': 17239,
'Araioses': 46440,
'Arame': 32701,
'Arari': 29848,
'Axixá': 12130,
'Bacabal': 104949,
'Bacabeira': 17055,
'Bacuri': 18582,
'Bacurituba': 5644,
'Balsas': 94887,
'Barão de Grajaú': 18820,
'Barra do Corda': 88212,
'Barreirinhas': 62528,
'Belágua': 7469,
'Bela Vista do Maranhão': 11209,
'Benedito Leite': 5632,
'Bequimão': 21280,
'Bernardo do Mearim': 6043,
'Boa Vista do Gurupi': 9287,
'Bom Jardim': 27446,
'Bom Jesus das Selvas': 34028,
'Bom Lugar': 16294,
'Brejo': 36397,
'Brejo de Areia': 9188,
'Buriti': 28678,
```

```
'Buriti Bravo': 23884,
'Buriticupu': 72358,
'Buritirana': 15430,
'Cachoeira Grande': 9431,
'Cajapió': 11177,
'Cajari': 19379,
'Campestre do Maranhão': 14374,
'Cândido Mendes': 20178,
'Cantanhede': 21995,
'Capinzal do Norte': 10934,
'Carolina': 24322,
'Carutapera': 23807,
'Caxias': 164880,
'Cedral': 9237,
'Central do Maranhão': 8673,
'Centro do Guilherme': 13458,
'Centro Novo do Maranhão': 21622,
'Chapadinha': 79675,
'Cidelândia': 14697,
'Codó': 122859,
'Coelho Neto': 49435,
'Colinas': 2438,
'Conceição do Lago-Açu': 16237,
'Coroatá': 65296,
'Cururupu': 32695,
'Davinópolis': 2094,
'Dom Pedro': 23350,
'Duque Bacelar': 11349,
'Esperantinópolis': 17241,
'Estreito': 41946,
'Feira Nova do Maranhão': 8504,
'Fernando Falcão': 10360,
'Formosa da Serra Negra': 19089,
'Fortaleza dos Nogueiras': 12631,
'Fortuna': 15552,
'Godofredo Viana': 11819,
'Gonçalves Dias': 17934,
'Governador Archer': 10840,
'Governador Edison Lobão': 18296,
'Governador Eugênio Barros': 16828,
'Governador Luiz Rocha': 7807,
'Governador Newton Bello': 10180,
'Governador Nunes Freire': 25577,
'Graça Aranha': 6262,
'Grajaú': 69527,
'Guimarães': 12030,
'Humberto de Campos': 28717,
'Icatu': 27113,
'Igarapé do Meio': 14177,
'Igarapé Grande': 11320,
'Imperatriz': 258682,
'Itaipava do Grajaú': 16057,
'Itapecuru Mirim': 68203,
'Itinga do Maranhão': 26000,
'Jatobá': 14796,
'Jenipapo dos Vieiras': 16515,
'João Lisboa': 23632,
'Joselândia': 16168,
'Junco do Maranhão': 3432,
'Lago da Pedra': 50266,
'Lago do Junco': 10840.
'Lago Verde': 16257,
'Lagoa do Mato': 11250
'Lago dos Rodrigues': 8873,
'Lagoa Grande do Maranhão': 11394,
'Lajeado Novo': 7550,
'Lima Campos': 11893,
'Loreto': 12157.
'Luís Domingues': 6951,
'Magalhães de Almeida': 19826,
'Maracaçumé': 21395,
'Marajá do Sena': 7792,
'Maranhãozinho': 16265,
'Mata Roma': 16829,
'Matinha': 23370,
'Matões': 33782,
'Matões do Norte': 16745,
'Milagres do Maranhão': 8464,
'Mirador': 2213,
'Miranda do Norte': 28381,
'Mirinzal': 14962,
'Monção': 33434,
'Montes Altos': 9160,
'Morros': 19433,
'Nina Rodrigues': 14454,
'Nova Colinas': 5384,
'Nova Iorque': 4683,
'Nova Olinda do Maranhão': 20928,
"Olho d'Água das Cunhãs": 19505,
```

```
'Olinda Nova do Maranhão': 14701,
'Paço do Lumiar': 122197,
'Palmeirândia': 19722,
'Paraibano': 21386,
'Parnarama': 34907,
'Passagem Franca': 19019,
'Pastos Bons': 19472,
'Paulino Neves': 16035,
'Paulo Ramos': 21040,
'Pedreiras': 39229,
'Pedro do Rosário': 25144,
'Penalva': 38470,
'Peri Mirim': 14318,
'Peritoró': 23196,
'Pindaré-Mirim': 32941,
'Pinheiro': 83387,
'Pio XII': 21485
'Pirapemas': 18625,
'Poção de Pedras': 17873,
'Porto Franco': 23885,
'Porto Rico do Maranhão': 5975,
'Presidente Dutra': 15142,
'Presidente Juscelino': 3641,
'Presidente Sarney': 18918,
'Presidente Vargas': 11193,
'Primeira Cruz': 15315,
'Raposa': 30761,
'Riachão': 3588,
'Ribamar Fiquene': 7791,
'Rosário': 42740,
'Sambaíba': 5671,
'Santa Filomena do Maranhão': 7773,
'Santa Helena': 2223,
'Santa Inês': 1596,
'Santa Luzia': 219134,
'Santa Luzia do Paruá': 25254,
'Santa Quitéria do Maranhão': 25642,
'Santa Rita': 136586,
'Santana do Maranhão': 13386,
'Santo Amaro do Maranhão': 15846,
'Santo Antônio dos Lopes': 14528,
'São Benedito do Rio Preto': 18663,
'São Bento': 34031,
'São Bernardo': 28507,
'São Domingos do Azeitão': 7392,
'São Domingos do Maranhão': 34376,
'São Félix de Balsas': 4585,
'São Francisco do Brejão': 11798,
'São Francisco do Maranhão': 12210,
'São João Batista': 37424,
'São João do Carú': 15808,
'São João do Paraíso': 23618,
'São João do Soter': 18543,
'São João dos Patos': 25929
'São José de Ribamar': 177687,
'São José dos Basílios': 7641,
'São Luís': 1101884,
'São Luís Gonzaga do Maranhão': 18856,
'São Mateus do Maranhão': 41529,
'São Pedro da Água Branca': 12690,
'São Pedro dos Crentes': 4668,
'São Raimundo das Mangabeiras': 18868,
'São Raimundo do Doca Bezerra': 5237,
'São Roberto': 6719,
'São Vicente Ferrer': 22247,
'Satubinha': 13914,
'Senador Alexandre Costa': 11141,
'Senador La Rocque': 14293,
'Serrano do Maranhão': 10253,
'Sítio Novo': 5522,
'Sucupira do Norte': 10636,
'Sucupira do Riachão': 5660,
'Tasso Fragoso': 8521,
'Timbiras': 29124,
'Timon': 169107,
'Trizidela do Vale': 21998,
'Tufilândia': 5840,
'Tuntum': 41832,
'Turiaçu': 35604,
'Turilândia': 25619,
'Tutóia': 58860,
'Urbano Santos': 33122,
'Vargem Grande': 56510,
'Viana': 78239,
'Vila Nova dos Martírios': 13392,
'Vitória do Mearim': 32764,
'Vitorino Freire': 31523,
'Zé Doca': 51714,
'Acauã': 7084,
```

```
'Agricolândia': 5139,
'Água Branca': 20196,
'Alagoinha do Piauí': 7651,
'Alegrete do Piauí': 4915,
'Alto Longá': 14304,
'Altos': 40524,
'Alvorada do Gurguéia': 5419,
'Amarante': 17598,
'Angical do Piauí': 6788,
'Anísio de Abreu': 9880,
'Antônio Almeida': 3164,
'Aroazes': 5835,
'Aroeiras do Itaim': 2551,
'Arraial': 4727,
'Assunção do Piauí': 7846,
'Avelino Lopes': 11289,
'Baixa Grande do Ribeiro': 11586,
"Barra D'Alcântara": 3951,
'Barras': 47066,
'Barreiras do Piauí': 3348,
'Barro Duro': 7032,
'Batalha': 18232,
'Bela Vista do Piauí': 4015,
'Belém do Piauí': 3566,
'Beneditinos': 10467,
'Bertolínia': 5501,
'Betânia do Piauí': 6205,
'Boa Hora': 6781,
'Bocaina': 12329,
'Bom Jesus': 11349,
'Bom Princípio do Piauí': 5630,
'Bonfim do Piauí': 5670,
'Boqueirão do Piauí': 6407,
'Brasileira': 8329,
'Brejo do Piauí': 3875,
'Buriti dos Lopes': 19781,
'Buriti dos Montes': 8244,
'Cabeceiras do Piauí': 10586,
'Cajazeiras do Piauí': 3559,
'Cajueiro da Praia': 7642,
'Caldeirão Grande do Piauí': 5776,
'Campinas do Piauí': 5613,
'Campo Alegre do Fidalgo': 5045,
'Campo Grande do Piauí': 5943,
'Campo Largo do Piauí': 7279,
'Campo Maior': 46833,
'Canavieira': 3950,
'Canto do Buriti': 21187,
'Capitão de Campos': 11417,
'Capitão Gervásio Oliveira': 4100,
'Caracol': 6116,
'Caraúbas do Piauí': 5868,
'Caridade do Piauí': 5067,
'Castelo do Piauí': 19716,
'Caxingó': 5424,
'Cocal': 27787,
'Cocal de Telha': 4891,
'Cocal dos Alves': 6153,
'Coivaras': 4020,
'Colônia do Gurguéia': 6484,
'Colônia do Piauí': 7656,
'Conceição do Canindé': 4803,
'Coronel José Dias': 4682,
'Corrente': 26644,
'Cristalândia do Piauí': 8294,
'Cristino Castro': 10423,
'Curimatá': 11388,
'Currais': 4954.
'Curralinhos': 4443,
'Curral Novo do Piauí': 5343,
'Demerval Lobão': 13817,
'Dirceu Arcoverde': 7011,
'Dom Expedito Lopes': 6904,
'Domingos Mourão': 4355,
'Dom Inocêncio': 9556,
'Elesbão Veloso': 14602,
'Eliseu Martins': 4915,
'Fartura do Piauí': 5307,
'Flores do Piauí': 4463,
'Floresta do Piauí': 2558,
'Floriano': 59935,
'Francinópolis': 5348,
'Francisco Ayres': 4348
'Francisco Macedo': 3184,
'Francisco Santos': 9319,
'Fronteiras': 11625,
'Geminiano': 5442,
'Gilbués': 10690,
'Guadalupe': 10499,
```

```
'Guaribas': 4562,
'Hugo Napoleão': 3877,
'Ilha Grande': 9426,
'Inhuma': 15308,
'Ipiranga do Piauí': 9811,
'Isaías Coelho': 8549,
'Itainópolis': 11530,
'Itaueira': 11019,
'Jacobina do Piauí': 5741,
'Jaicós': 19104,
'Jardim do Mulato': 4504,
'Jatobá do Piauí': 4865,
'Jerumenha': 4452,
'João Costa': 3008,
'Joaquim Pires': 14354,
'Joca Marques': 5443,
'José de Freitas': 39208,
'Juazeiro do Piauí': 5478,
'Júlio Borges': 5627,
'Jurema': 15378,
'Lagoinha do Piauí': 2845,
'Lagoa Alegre': 8542,
'Lagoa do Barro do Piauí': 4655,
'Lagoa de São Francisco': 6758,
'Lagoa do Piauí': 4064,
'Lagoa do Sítio': 5177,
'Landri Sales': 5295,
'Luís Correia': 30311,
'Luzilândia': 25486,
'Madeiro': 8310,
'Manoel Emídio': 5349,
'Marcolândia': 8492,
'Marcos Parente': 4549,
'Massapê do Piauí': 6441,
'Matias Olímpio': 10936,
'Miguel Alves': 33760,
'Miguel Leão': 1246,
'Milton Brandão': 6613,
'Monsenhor Gil': 10564,
'Monsenhor Hipólito': 7749,
'Monte Alegre do Piauí': 10613,
'Morro Cabeça no Tempo': 4532,
'Morro do Chapéu do Piauí': 6796,
'Murici dos Portelas': 9159,
'Nazaré do Piauí': 7339,
'Nazária': 8570,
'Nossa Senhora de Nazaré': 4870,
'Nossa Senhora dos Remédios': 8692,
'Novo Oriente do Piauí': 6565,
'Novo Santo Antônio': 2640,
'Oeiras': 37029,
"Olho D'Água do Piauí": 2459,
'Padre Marcos': 6868,
'Paes Landim': 4129,
'Pajeú do Piauí': 3389,
'Palmeira do Piauí': 5043,
'Palmeirais': 14539,
'Paquetá': 3945,
'Parnaguá': 10791
'Parnaíba': 153078,
'Passagem Franca do Piauí': 4313,
'Patos do Piauí': 6392,
"Pau D'Arco do Piauí": 4045,
'Paulistana': 20523,
'Pavussu': 3677,
'Pedro II': 38742,
'Pedro Laurentino': 2536,
'Nova Santa Rita': 29300,
'Picos': 78222,
'Pimenteiras': 12127,
'Pio IX': 18425,
'Piracuruca': 28791,
'Piripiri': 63742,
'Porto': 12568,
'Porto Alegre do Piauí': 2710,
'Prata do Piauí': 3151,
'Queimada Nova': 8992,
'Redenção do Gurquéia': 8778,
'Regeneração': 17978,
'Riacho Frio': 4312,
'Ribeira do Piauí': 4477,
'Ribeiro Gonçalves': 7341,
'Rio Grande do Piauí': 6432,
'Santa Cruz do Piauí': 6240,
'Santa Cruz dos Milagres': 4019,
'Santa Filomena': 14477,
'Santa Luz': 5860,
'Santana do Piauí': 4634
'Santa Rosa do Piauí': 5257,
```

```
'Santo Antônio de Lisboa': 6415,
'Santo Antônio dos Milagres': 2161,
'Santo Inácio do Piauí': 3798,
'São Braz do Piauí': 4448,
'São Félix do Piauí': 2942.
'São Francisco de Assis do Piauí': 5755,
'São Francisco do Piauí': 6423,
'São Gonçalo do Gurguéia': 3041,
'São Gonçalo do Piauí': 5015,
'São João da Canabrava': 4608
'São João da Fronteira': 6042,
'São João da Serra': 6139,
'São João da Varjota': 4840,
'São João do Arraial': 7989,
'São João do Piauí': 20601,
'São José do Divino': 3860,
'São José do Peixe': 3745,
'São José do Piauí': 6705,
'São Julião': 6363,
'São Lourenço do Piauí': 4573,
'São Luis do Piauí': 2644,
'São Miguel da Baixa Grande': 2452,
'São Miguel do Fidalgo': 3039,
'São Miguel do Tapuio': 17662,
'São Pedro do Piauí': 14291,
'São Raimundo Nonato': 34710,
'Sebastião Barros': 3469,
'Sebastião Leal': 4294,
'Sigefredo Pacheco': 10041,
'Simões': 14633,
'Simplício Mendes': 12711,
'Socorro do Piauí': 4569,
'Sussuapara': 6730,
'Tamboril do Piauí': 2919,
'Tanque do Piauí': 2765,
'Teresina': 864845,
'União': 44485,
'Uruçuí': 21558,
'Valença do Piauí': 20918,
'Várzea Branca': 4947,
'Várzea Grande': 284971,
'Vera Mendes': 3077,
'Vila Nova do Piauí': 2971,
'Wall Ferraz': 4462,
'Abaiara': 11737,
'Acarape': 14929,
'Acaraú': 62641,
'Acopiara': 54270,
'Aiuaba': 17399,
'Alcântaras': 11714,
'Altaneira': 7586,
'Alto Santo': 17146,
'Amontada': 43452,
'Antonina do Norte': 7353,
'Apuiarés': 14600,
'Aquiraz': 80271,
'Aracati': 74547,
'Aracoiaba': 26469,
'Ararendá': 10935,
'Araripe': 21600,
'Aratuba': 11847,
'Arneiroz': 7840,
'Assaré': 23417,
'Aurora': 5679,
'Baixio': 6288,
'Banabuiú': 18197,
'Barbalha': 60781,
'Barreira': 22425,
'Barro': 22680,
'Barroquinha': 15017,
'Baturité': 35750,
'Beberibe': 53573,
'Bela Cruz': 32591
'Boa Viagem': 54470,
'Brejo Santo': 49477,
'Camocim': 63661,
'Campos Sales': 27426,
'Canindé': 76997,
'Capistrano': 17738,
'Caridade': 22547,
'Cariré': 18448,
'Caririacu': 26965,
'Cariús': 18699,
'Carnaubal': 17606,
'Cascavel': 328454,
'Catarina': 20698,
'Catunda': 10342,
'Caucaia': 361400,
```

'Cedro': 11807,

```
'Chorozinho': 20264,
                         'Coreaú': 23136,
                         'Crateús': 75074,
                         'Crato': 132123,
                         'Croatá': 18063,
                         'Cruz': 24827,
                         'Deputado Irapuan Pinheiro': 9625,
                         'Ereré': 7198,
                         'Eusébio': 53618,
                         'Farias Brito': 19450,
                         'Forquilha': 24218,
                         'Fortaleza': 2669342,
                         'Fortim': 16480,
                         'Frecheirinha': 14072,
                         'General Sampaio': 7618,
                         'Graça': 14399,
                         'Granja': 54748,
                          'Granjeiro': 4844,
                         'Groaíras': 11068,
                          'Guaiúba': 26064,
                          'Guaraciaba do Norte': 40642,
                         'Guaramiranga': 5193,
                          'Hidrolândia': 21706,
                         'Horizonte': 67337,
                          'Ibaretama': 13353,
                          'Ibiapina': 24997,
                         'Ibicuitinga': 12525,
                         'Icapuí': 19934,
                          'Icó': 68018,
                         'Iguatu': 2256,
                          'Independência': 6167,
                          'Ipaporanga': 11593,
                         'Ipaumirim': 12463,
                         'Ipu': 41964,
                         'Irauçuba': 24156,
                         'Itaiçaba': 7827,
                          'Itaitinga': 37980,
                         'Itapajé : 52675,
                         'Itapipoca': 129358,
                         'Itapiúna': 20382,
                         'Itarema': 41826,
                         'Itatira': 21647
                          'Jaguaretama': 18162,
                         'Jaguaribara': 11401,
                         'Jaguaribe': 34682,
'Jaguaruana': 33705,
                         'Jardim': 26097,
                          'Jati': 8111,
                         'Jijoca de Jericoacoara': 19816,
                         'Juazeiro do Norte': 274207,
                          'Jucás': 24833,
                         'Lavras da Mangabeira': 31508,
                         'Limoeiro do Norte': 59540,
                         'Madalena': 19691,
'Maracanaú': 227886,
                         'Maranguape': 128978,
                          'Marco': 27361,
                         'Martinópole': 11233,
                         'Massapê': 38737,
'Mauriti': 47962,
                         'Meruoca': 15057,
                          'Milagres': 10950,
                          'Milhã': 13155,
                         'Miraíma': 13818,
                          'Missão Velha': 35393,
                         'Mombaça': 43797,
                         'Monsenhor Tabosa': 17234,
                         ...}
In [40]:
                        len(cities_population)
Out[40]: 5297
In [41]:
                        cities_population_df = pd.DataFrame(cities_population.values(), index=cities_population.keys(), columns=['Population_values(), index=cities_population_values(), index=cities_population_values(), columns=['Population_values(), index=cities_population_values(), index=cities_po
In [42]:
                        cities population df
Out[42]:
                                                                       Populacao
```

'Chaval': 13069, 'Choró': 13521,

Alta Floresta D'Oeste	22945
Ariquemes	107863
Cabixi	5312
Cacoal	85359
Cerejeiras	16323
Vianópolis	13863
Vicentinópolis	8743
Vila Boa	6171
Vila Propício	5821
Brasília	3015268

5297 rows × 1 columns

Agora que já temos uma tabela separada para a população das cidades, vamos remover a população da planilha principal

In [43]:
 df_final = df_no_accumulated.drop('populacaoTCU2019', axis=1)

In [44]: df f

df_final

Out[44]: regiao estado municipio data semanaEpi casosNovos obitosNovos interior/metropolitana Alta Floresta 2020-03-0 Norte RO 13 0 0 False D'Oeste Alta Floresta 2020-03-Norte RO 13 False D'Oeste Alta Floresta 2020-03-2 RO 0 0 Norte 14 False D'Oeste Alta Floresta 2020-03-RO 14 Norte False D'Oeste Alta Floresta 2020-03-0 0 4 Norte RO 14 False D'Oeste 31 2022-01-Centro-3743035 DF Brasília 4 0 0 True Oeste Centro-2022-01-DF 3743036 Brasília 6976 True 2022-01-Centro-3743037 DF Brasília 10697 4 True Oeste 2022-01-Centro-3743038 DF Brasília 5485 True Oeste Centro-2022-01-3743039 DF Brasília 5243 4 True Oeste

3743040 rows × 8 columns

Export CSVS

O ultimo passo é exportar os CSVs para análise no Qlik Sense ou PowerBI.

Temos 2 tabelas para exportar como csv, os dados totais de covid, e a populacao de cada município.

```
In [45]: df_final.to_csv('DadosCovidTratados3.csv')
In [46]: cities_population_df.to_csv('DadosPopulacao3.csv')
```