

1 Predictive model of COVID-19 incidence and socioeconomic description of
2 municipalities in Brazil

3

4 Extended Title (or short abstract):

5 The present study aims to describe the demographic and socioeconomic
6 characteristics of cities with cases of COVID-19 in Brazil, as well as to
7 determine a predictive model for the number of cases. We analyzed data from
8 672 cities where 73.1% of the Brazilian population lives. Our model predicts, if
9 the conditions are maintained, 2,358,703 (2,172,930 to 2,544,477) cumulative
10 cases on July 25, 2020.

11

12 Isadora C. R. Carneiro^{1#},

13 Eloiza D. Ferreira^{1#},

14 Janaina C. da Silva^{2#},

15 Guilherme Soares^{1,3},

16 Daisy M. Strottmann⁴,

17 Guilherme F. Silveira^{*1}

18

19 1. Instituto Carlos Chagas - Fiocruz/PR, Curitiba, PR

20 2. Laboratório de Biologia de Tumores, Universidade Estadual do Oeste do
21 Paraná, UNIOESTE, Francisco Beltrão-PR.

22 3. Laboratórios De Referência Em Virose Emergentes E Reemergentes,
23 Instituto Carlos Chagas - Fiocruz/PR, Curitiba, PR

24 4. Laboratório de Virologia Molecular, Instituto Carlos Chagas - Fiocruz/PR,
25 Curitiba, PR

26

27 * Corresponding author: gfsilveira@gmail.com, guilherme.silveira@fiocruz.br

28 # These authors contributed equally to the work.

29

30

31

32

33

34 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

35 ABSTRACT:

36

37 Coronaviruses are enveloped viruses that can cause respiratory,
38 gastrointestinal, hepatic, and neurological diseases. In December 2019, a new
39 highly contagious coronavirus termed severe acute respiratory syndrome
40 coronavirus 2 (SARS-CoV-2) emerged in China. SARS-CoV-2 causes a
41 potentially lethal human respiratory infection, COVID-19, that is associated with
42 fever and cough and can progress to pneumonia and dyspnea in severe cases.
43 Since the virus emerged, it has spread rapidly, reaching all continents around
44 the world. A previous study has shown that, despite being the best alternative in
45 the current pandemic context, social distancing measures alone may not be
46 sufficient to prevent COVID-19 spread, and the overall impact of the virus is of
47 great concern. The present study aims to describe the demographic and
48 socioeconomic characteristics of 672 cities with cases of COVID-19, as well as
49 to determine a predictive model for the number of cases. We analyzed data
50 from cities with at least 1 reported case of COVID-19 until June 26, 2020. It was
51 observed that cities with confirmed cases of the disease are present in all
52 Brazilian states, affecting 36.5% of the municipalities in Rio de Janeiro State.
53 The inhabitants in cities with reported cases of COVID-19 represent more than
54 73.1% of the Brazilian population. Stratifying the age groups of the inhabitants
55 and accounting for the percentage of women and men does not affect COVID-
56 19 incidence (confirmed cases/100,000 inhabitants). The demographic density,
57 the MHDl and the per capita income of the municipalities with cases of COVID-
58 19 do not affect disease incidence. In addition, if conditions are maintained, our
59 model predicts 2,358,703 (2,172,930 to 2,544,477) cumulative cases on July
60 25, 2020.

61

62

63

64

65

66

67 Keywords: COVID-19, COVID-19 in Brazil, Exploratory data analysis, ARIMA,
68 computational model.

69 INTRODUCTION:

70

71 A pandemic is defined as any epidemic disease widely distributed
72 geographically that affects different regions simultaneously. Over the years,
73 humanity has faced many instances when health and science are put to the test
74 and need to present answers. The Spanish flu ravaged the whole world and
75 was often confused with several other diseases, such as cholera, dengue, and
76 typhus (GOULART, 2005). At the end of the pandemic, approximately 50 million
77 people had died. In 2009, a novel influenza A (H1N1) virus emerged in Mexico,
78 where the transmission remained geographically contained for at least one
79 month. Migratory events and population movement caused the virus to spread
80 to other countries, leading the World Health Organization (WHO) to declare
81 H1N1 a global pandemic (LIPSITCH et al., 2011). In official WHO records, the
82 H1N1 pandemic resulted in the deaths of 18,500 people around the world.

83 On March 11, 2020, after a declaration by the WHO, the world population
84 again lived under a shadow of fear of a new pandemic, this time caused by a
85 coronavirus. Coronaviruses (CoV) are enveloped viruses that are part of a large
86 family of single-stranded RNA viruses with a positive-sense genome and can
87 cause respiratory, gastrointestinal, hepatic, and neurological diseases. CoV can
88 infect many animal species, including birds, cows, pigs, and humans, causing
89 acute and chronic diseases (Chang et al., 2012; Weiss, 2011). Most infections
90 caused by viruses from the family *Coronaviridae* induce a mild form of the
91 disease in humans, usually causing flu-like symptoms. However, after the
92 discovery of severe acute respiratory syndrome (SARS), a greater contagion
93 capacity and lethality potential of this viral family was evidenced (Weiss et al.,
94 2011). The etiological agent of SARS, SARS-CoV, was identified in mid-2003
95 after an outbreak of the disease in November 2002 in Guangdong Province,
96 China, where 8,700 cases and 774 deaths were confirmed (Contini et al., 2020).
97 The so-called novel CoV, initially referred to as 2019-nCoV, was first described
98 when a group of patients reported symptoms of pneumonia of unknown cause
99 in Wuhan City, Hubei Province, China, in December 2019 (ZHU, 2020).

100 On February 11, 2020, after phylogenetic and pathophysiological
101 analyses, 2019-nCoV was officially named SARS-CoV-2 due to its similarity to
102 SARS-CoV, as announced by the Coronavirus Study Group (CSG) of the

103 International Committee on Taxonomy of Viruses (ICTV), according to the 2015
104 WHO nomenclature guidelines (Gorbalenya et al., 2020). The pathology caused
105 by SARS-CoV-2 infection was termed COVID-19, characterized by a flu-like
106 condition associated with fever and cough that can progress to pneumonia and
107 dyspnea in more severe cases (CHAN et al., 2020). The incubation period of
108 the disease varies from 2 to 14 days, and in approximately 80% of cases,
109 infected individuals remain asymptomatic. However, unlike patients with
110 influenza, viral transmission from asymptomatic individuals is possible (Contini
111 et al., 2020). In addition, according to Contini (2020), the mechanism of
112 contagion is direct, that is, through contact with respiratory fomites of infected
113 people. Other studies show that SARS-CoV-2 can survive in the air for more
114 than 3 hours and on surfaces such as plastics and metals for up to 3 days (Van
115 Doremalen et al., 2020). Currently, there are no vaccines to prevent the
116 disease, reinforcing the need for prophylactic measures, namely, correct hand,
117 environment and surface hygiene and social distancing.

118 A previous study has shown that social distancing and other preventive
119 measures alone may not be sufficient to prevent the spread of COVID-19, and
120 the overall impact of the virus is of great concern (Sohrabi et al., 2020). It is also
121 noteworthy that additional research is needed to help define the exact rates and
122 mechanisms of person-to-person transmission, as well as to determine
123 additional factors that can guide containment actions.

124 The internal and external logistic and transitory movements, as well as
125 several other socioeconomic factors, can not only contribute to the
126 understanding of viral spread but also assist in surveillance measures and
127 competent decision-making for regional health systems, where such analysis
128 can (and should) be implemented to reduce the exponential growth rate of
129 positive cases. Different approaches are being used to better understand the
130 transmission dynamics of SARS-CoV-2 to inform pandemic prevention and
131 control measures. In this context, the present study aims to analyze the
132 demographic and socioeconomic characteristics of cities with COVID-19 cases,
133 as well as to adjust a predictive model for the cumulative number of disease
134 cases thus expanding the possibilities of decision-making at the micro- and
135 macroregional levels.

136

137 MATERIALS AND METHODS:

138

139 In the present work, an ecological study design was used; this method of
140 epidemiological study helped us to generate hypotheses about possible
141 associations between socioeconomic characteristics of the Brazilian
142 municipalities and the COVID-19 incidence and fatality rate.

143 For the exploratory data analysis (EDA) and the predictive model
144 adjustment, the Python programming language was used with several libraries
145 specifically for this purpose. Along with Python, it was necessary to import
146 different packages and libraries, the most used ones being pandas, NumPy and
147 SciPy, with the function of organizing and structuring the data. For statistical
148 calculations, the statsmodels package was imported to analyze the time series
149 and the autoregressive integrated moving average (ARIMA) model for the
150 forecasts. The purpose of these methods is to fit the model to the data as well
151 as possible. Matplotlib and Seaborn were used to generate two-dimensional
152 (2D) graphics. The project can be accessed through the GitHub page,
153 <https://github.com/gfsilveira/covid>. The analyzed database contains data from
154 672 municipalities (12,1% of the cities in Brazil) in the 26 states of the
155 federation plus the Federal District.

156 The records of COVID-19 cases at the municipal level were obtained
157 through daily updates from the Health Departments of the Federative Units
158 compiled by Álvaro Justen and his collaborators until June 26, 2020, available
159 at <https://brasil.io/dataset/covid19/caso>. Demographic and socioeconomic
160 characteristics publicly available at the municipal level, such as population
161 density, Municipal Human Development Index (MHDI), total area in km² and per
162 capita income, were obtained from the Brazilian Institute of Geography and
163 Statistics (IBGE) from the demographic census conducted in 2010. Data on the
164 age range and sex distribution of the population were obtained from the 2015
165 census.

166 A data structure containing different age groups between 0 and 80+
167 years old from the IBGE 2015 census database was used. The age of residents
168 in Brazilian municipalities was separated into groups of young (0 to 29 years),
169 adults (30 to 64 years) and seniors (65+ years). This distribution grouped ages
170 according to low risk (young), comorbidities risk (adults) and complications risk

171 (seniors) from SARS-CoV-2. The percentage of declared sex data
172 (female/male) was also analyzed in cities with cases of COVID-19. In addition to
173 data from the population, the number of cities with COVID-19 cases within each
174 state was analyzed by the percentage of Brazilian municipalities/states with
175 confirmed cases until June 26, 2020.

176 Incidence (cases per 100,000 inhabitants) was analyzed according to
177 demographic density (inhabitants/km²), Municipal Human Development Index
178 (MHDI), socioeconomic data (per capita income), age groups and sex ratios.

179 For the development of the predictive model, we used the ARIMA model
180 proposed by Box & Jenkins (1970), which consists of developing and adjusting
181 stationary or nonstationary linear models relative to an observed time series.
182 The autoregressive (AR) component indicates that the variable of the time
183 series is regressed on its own lagged values. The I (for "integrated") indicates
184 that the data values have been replaced with the difference between $n+1$ and n
185 values, performed more than once. The moving average (MA) is a calculation of
186 data points by creating a series of averages of different subsets of full data. In
187 the moving average, the regression error is a linear combination of various
188 times in the past and error terms whose values occurred contemporaneously.
189 The construction of the model was based on daily COVID-19 cases. To analyze
190 the stationary condition of the time series in different orders, the augmented
191 Dickey-Fuller (ADF) test was used. The null hypothesis of the ADF test is that
192 the time series is nonstationary. The p-value of the test was 0.000234, on the
193 order of differentiation 2, less than the significance level (0.05), rejecting the null
194 hypothesis and indicating that the time series is indeed stationary (Figure 1G).
195 To determine which predictive model to use, the autocorrelation function (ACF)
196 (Figure 1B, E, H) and partial autocorrelation function (PACF) (Figure 1C, F, I)
197 were analyzed to determine the ARIMA p , d , and q parameters. The ACF is the
198 correlation of a variable with itself at differing time lags, and the PACF partial
199 autocorrelation at lag 1 is very high (it equals the ACF at lag 1), but the other
200 values are correlated when lag > 1 . The PACF does not include any value for
201 lag 0 because it is impossible to remove any intermediate autocorrelation
202 between t and $t-k$ when $k=0$, and therefore the PACF does not exist at lag 0.
203 The ACF for an AR(p) process approaches zero very slowly, but the PACF goes
204 to zero for values of lag $> p$. The ACF for an MA(q) process goes to zero for

205 values of lag $> q$, but the PACF approaches zero very slowly. As defined by the
206 ADF test, $d = 2$ was a second order of differentiation, making the series
207 stationary (Figure 1G). In the second order of differentiation, the PACF with lag
208 2 is already below significance (Figure 1I), $p = 1$. The same occurs in the ACF
209 in the second order, where lag 2 is below significance (Figure 1H). Then, the
210 parameters were tested by the minor Akaike Information Criteria (AIC), four
211 ARIMA models, (1,2,0), (1,2,1), (2,2,0), (2,2,1) and (2,2,2). With these data, the
212 ARIMA parameters (2,2,1) showed the best adjustment (Figure 1J).

213

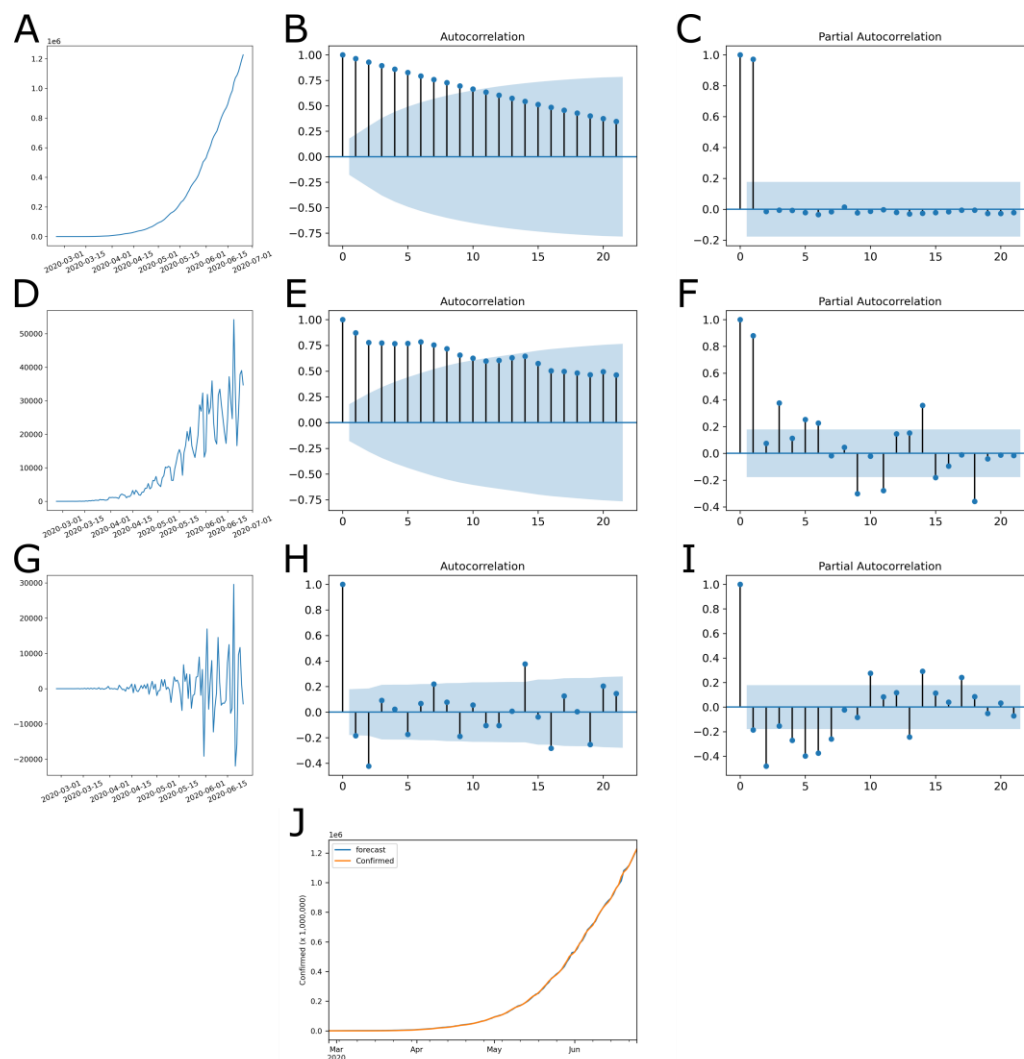


Figure 1. ACF and PACF plots for determination of model parameters. (A-C) Original series, (D-F) 1st-order differencing and (G-I) 2nd-order differencing. (A, D and G) Time series, with number of cases by date; (B, E and H) autocorrelation function; and (C, F and I) partial autocorrelation function. With this data, a model ARIMA ($p=1$, $d=2$, $q=1$) was determined. (J) Adjustment of forecast data with observed outcome data.

214 RESULTS:

215

216 1) A total of 73.1% of the Brazilian population lives in cities with confirmed
217 cases of COVID-19

218

219 Since COVID-19 is a pathology caused by SARS-CoV-2 that is transmitted
220 directly from person to person, in the present work, we seek to observe the
221 characteristics of the affected cities. The analyzed database contains data from
222 672 municipalities (12,1% of the 5,570 cities in Brazil) from the 26 states of the
223 federation, plus the Federal District (Figure 2), which had at least 1 confirmed
224 infection as of June 26, 2020, totaling 1,225,993 (0.58% of the Brazilian
225 population) cases of COVID-19, which resulted in 54,918 deaths. The most
226 affected state was Rio de Janeiro, reaching 36.5% of the municipalities reported
227 cases of COVID-19. Until the last analyzed date, we observed an incidence rate
228 of 63.31/100,000 inhabitants and a mortality rate of 2575 (per 100,000
229 inhabitants). The data analyzed in this work have been updated daily since
230 February 25, 2020. The first 100 cases had been confirmed by March 14, and
231 after one week, on March 21, there were already 1,000 confirmed cases. Brazil
232 reached 10,000 and 100,000 confirmed cases on April 4 and May 3,
233 respectively. As of Jun 19, the number of confirmed cases reached 1,039,339.

234

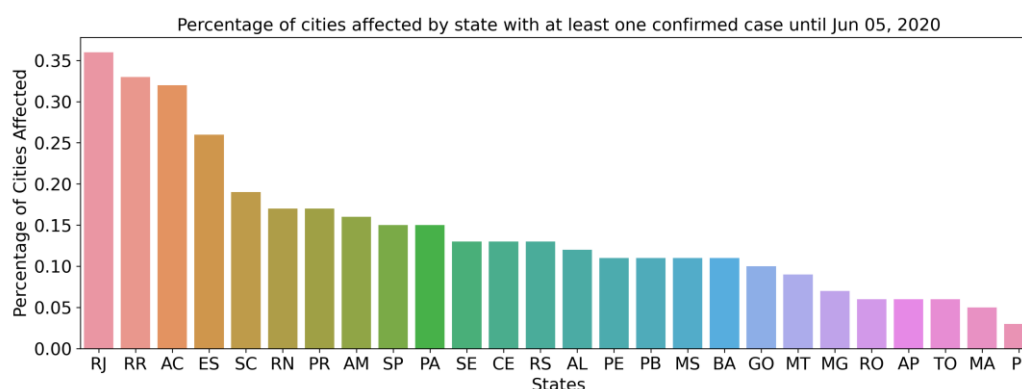


Figure 2. All Brazilian states have confirmed cases of COVID-19. Distribution of confirmed cases of COVID-19 by state. Percentage of cities with at least 1 case of positive infection.

235

236 The population of Brazil in 2018 (last available data) was 210,147,125
237 inhabitants, with the largest number of inhabitants between 10 and 34 years old

238 (Figure 3A). The cities where cases of COVID-19 were observed have
239 153,528,953 inhabitants, representing 73.1% of the Brazilian population, and as
240 expected, the distribution of age groups is the same as the general distribution
241 in Brazil (Figure 3B). In the present study, the age ranges of the population in
242 the affected cities were grouped into seniors over age 65, with a higher risk;
243 adults between 30 and 64 with a greater likelihood of comorbidity; and the
244 young, aged 0 to 29, with lower risk. The percentage of inhabitants in each age
245 group (Figure 3C) was used for the analysis in relation to COVID-19
246 incidence. The city with the greatest incidence, Santo Antônio do Içá
247 (Amazonas) (4592.17 cases/100,000 inhabitants), has 70% of inhabitants in the
248 young group (0-29 years old), 27% in the adult group (30-64 years old) and 3%
249 in the seniors group (65+ years old). In contrast, the city of Utinga (Bahia), with
250 a lower COVID-19 incidence in the database (15,64 cases/100,000 inhabitants),
251 has 55% of inhabitants in the young group (0-29 years old), 38% in the adult
252 group (30-64 years old) and 7% in the seniors group (65+ years old). The
253 quartiles of the percentage of inhabitants in each age group were analyzed to
254 determine if cities with a large percentage of senior, adult or young people have
255 a greater incidence. For the young group, the percentage of inhabitants was
256 divided into 34 – 44% (Figure 3D, blue), 45 – 48% (Figure 3D, orange), 49 –
257 51% (Figure 3D, green), and 52 – 71% (Figure 3D, red) of the total population.
258 The adult group composed 26 – 41% (Figure 3E, blue), 42 – 45% (Figure 3E,
259 orange), 46% (Figure 3E, green), and 47 – 52% (Figure 3E, red) of the total
260 population. The senior group included 2 – 6% (Figure 3F, blue), 7 – 8% (Figure
261 3F, orange), 9 – 10% (Figure 3F, green), and 11 – 16% (Figure 3F, red) of the
262 total population. These groups were analyzed against the incidence of COVID-
263 19 (Figure 3D-F). In conclusion, in the affected cities, the age groups do not
264 show a relationship with the incidence of the disease. Municipalities with a
265 greater or lesser percentage of inhabitants in the young, adult, or senior age
266 groups do not have a greater or lesser incidence of COVID-19.

267

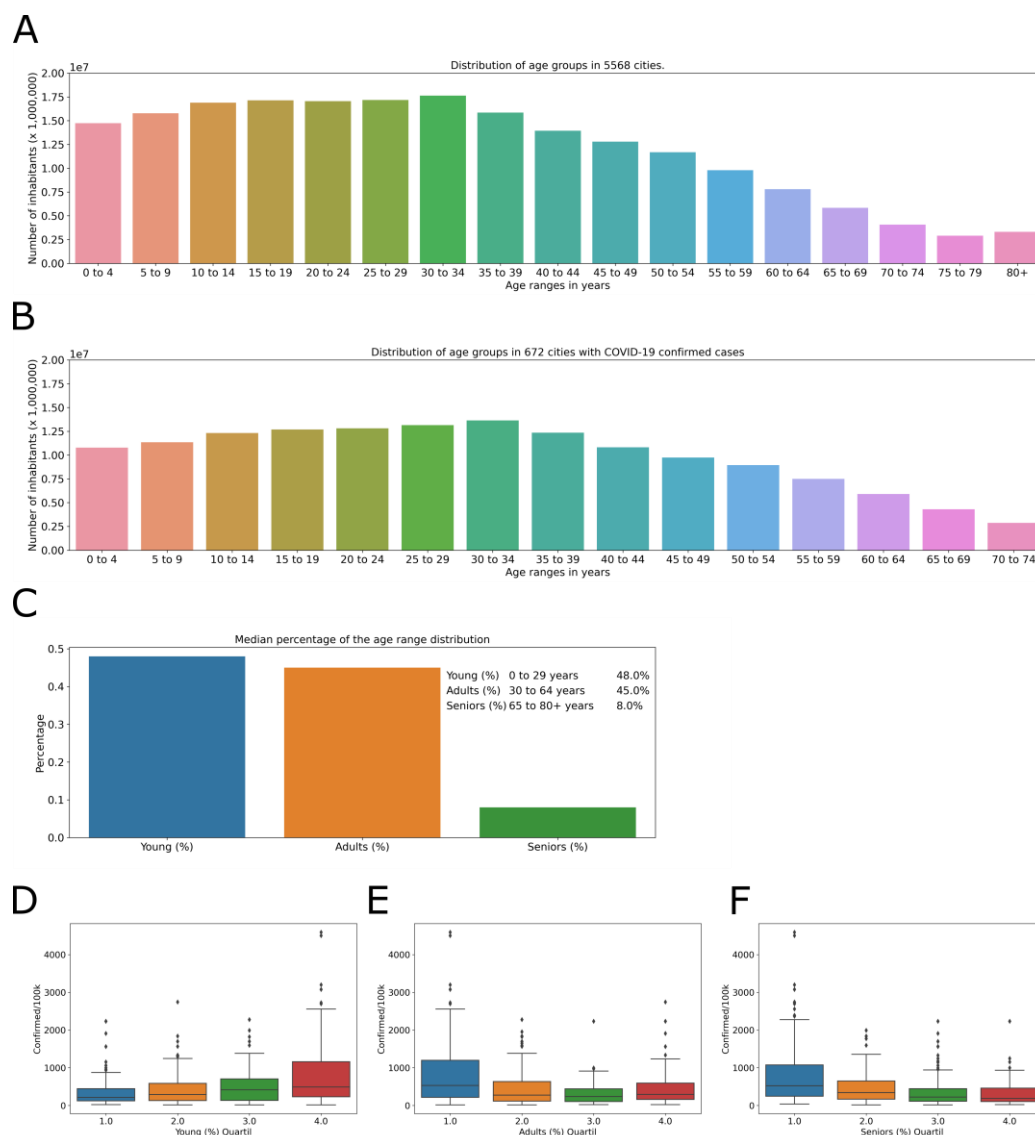


Figure 3. Age distribution in cities with cases of COVID-19. (A) Number of inhabitants by age group (increments of 5 years) in Brazil. (B) Number of inhabitants by age group in cities with cases of COVID-19. Stratification of age groups by age in the municipalities studied, represented by the percentage of the number of inhabitants for each city (C). Distribution of quartiles of the number of inhabitants per age groups of (D) young, (E) adults, (F) seniors, in relation to the COVID-19 incidence, in the affected cities.

268

269 Another demographic factor analyzed was the declared sex of the
 270 inhabitants. In Brazil, the population distribution percentage is 51.7% for women
 271 and 48.3% for men. We observed a similar distribution, with 50.11% (2,979,950)
 272 women and 49.89% (2,965,800) men, in cities with COVID-19 cases. The
 273 quartiles of the number of women and men inhabitants in the affected cities do
 274 not show a relationship with the incidence of the disease. For the male group,
 275 46 – 49% (Figure 4A, blue), 50% (Figure 4A orange), 51% (Figure 4A green),

276 and 52 – 69% (Figure 4A red) of the total population were included. For the
277 female group, the percentage of inhabitants was divided into 31 – 49% (Figure
278 4B, blue), 50% (Figure 4B, orange), 51% (Figure 4B, green), and 52 – 54%
279 (Figure 4B, red) of the total population. Municipalities with a greater or lesser
280 number of men (Figure 4A) or women (Figure 4B) inhabitants do not have a
281 greater or lesser incidence.
282

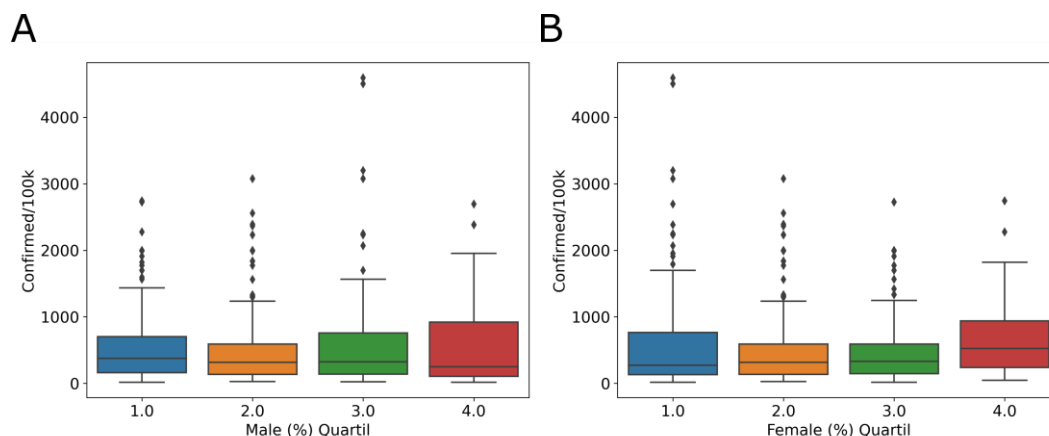


Figure 4. Declared sex of inhabitants in cities with cases of COVID-19.
Distribution of quartiles of the percentage of (A) male and (B) female
inhabitants, in relation to the COVID-19 incidence, in the affected cities.

283

284 Since SARS-CoV-2 is transmitted from person to person, we describe
285 characteristics, such as demographic density (inhabitants/km²), of the
286 dispersion and agglomeration of people in the municipalities. Of the 672 cities
287 analyzed, 577 (86%) are more densely populated than the national average
288 density, 23.9 hab/km². In the database, the mean population density was
289 683.76 hab/km² in all cities with COVID-19 cases. The city with the lowest
290 demographic density was Novo Airão (Amazonas), with 0.39 hab/km² and
291 1074.33 cases/100,000 inhabitants. The highest demographic density was in
292 São João de Meriti (Rio de Janeiro), with 13024.6 hab/km² and 334.881
293 cases/100,000 inhabitants. These apparent differences have no significance in
294 the grouped data. The COVID-19 incidence was observed for the demographic
295 density (hab/km²) quantiles, (Figure 5A) 0.39 to 38.86; (Figure 5B) 38.86 to
296 127.27; (Figure 5C) 127.27 to 540.95; and (Figure 5D) 540.95 to 13024.56. The
297 analyses of the quartiles of demographic density show no relation with COVID-

298 19 incidence (Figure 5E). Municipalities with a greater or lesser demographic
299 density do not have a greater or lesser incidence of COVID-19.

300

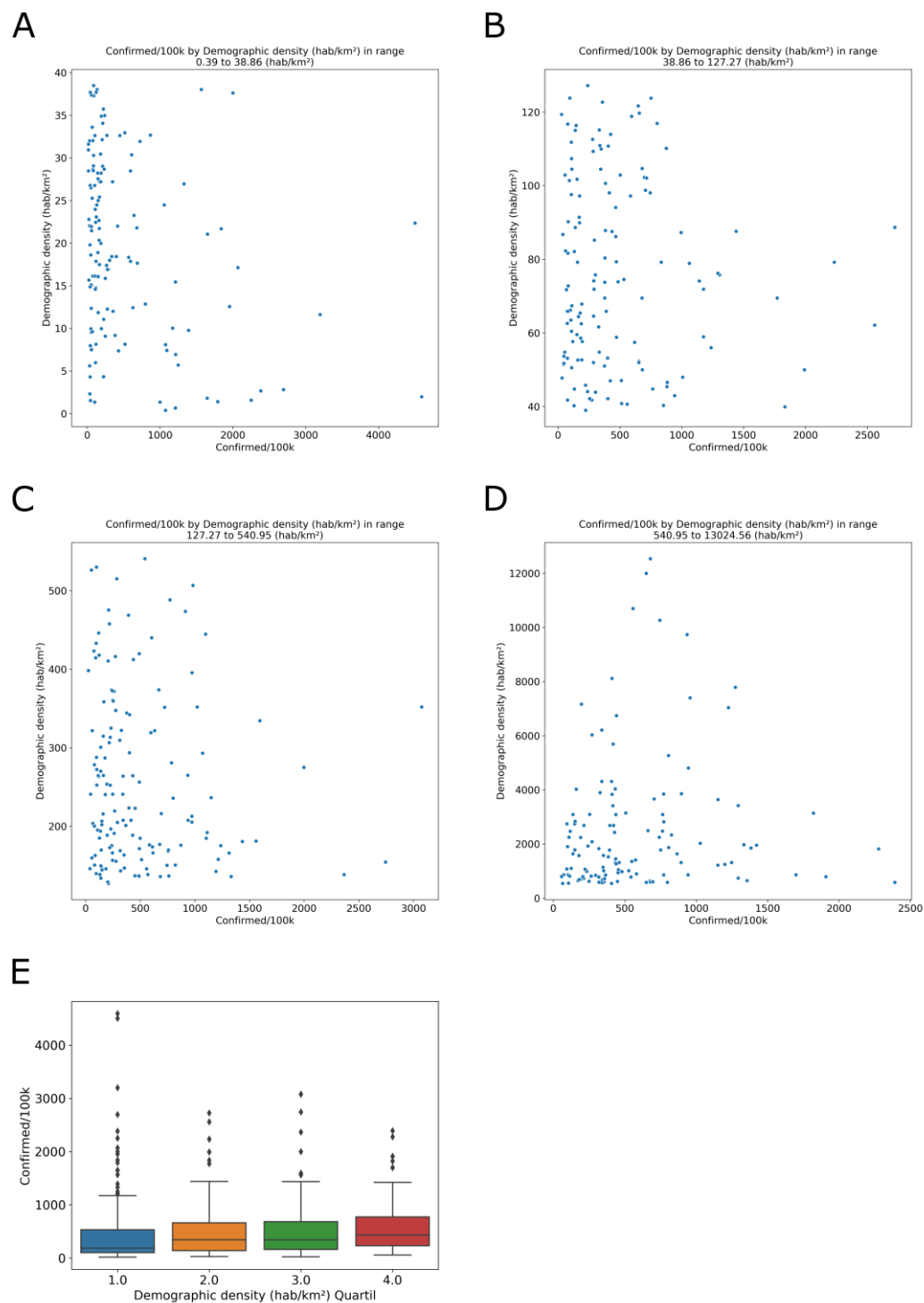


Figure 5. The demographic density in cities with cases of COVID-19. Number of inhabitants (population) by area (km²) resulting in demographic density (hab/km²) in four quartiles, (A) 0.39 to 38.86; (B) 38.86 to 127.27; (C) 127.27 to 540.95; and (D) 540.95 to 13024.56. (E) The relationship between the demographic density quartiles and COVID-19 incidence.

301

302 The Municipal Human Development Index (MHDI) is related to life
303 expectancy, educational level and income distribution. The global Brazilian HDI
304 for 2013 was 0.744, the 79th position in the world, ranking among the 187
305 countries and territories recognized by the United Nations. In the Global HDI for
306 HDR 2014, the three dimensions have the same weight, and the human
307 development ranges are fixed as follows: low human development, less than
308 0.550; average, between 0.550 and 0.699; high, between 0.700 and 0.799; and
309 very high, above 0.800. In Brazil, the per capita income was 1.48 (R\$ 1,443.10)
310 for 2017 (last year with available data). In this context, the MHDI (Figure 6A)
311 and per capita income (Figure 6B) were compared with COVID-19 incidence.
312

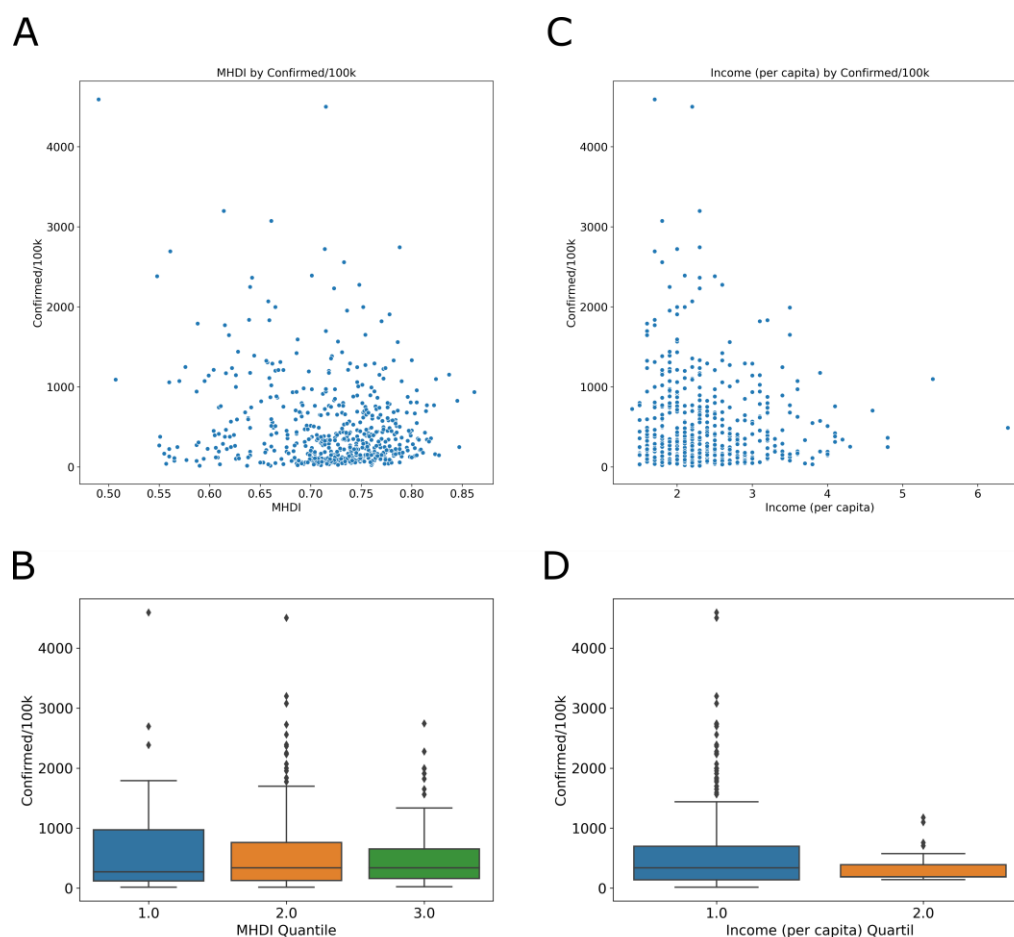


Figure 6. The Municipal Human Development Index (MHDI) and per capita income in cities with cases of COVID-19. (A) MHDI and (C) income (number of times the minimum wage [R\$ 975.00] is earned per month for formal workers) in cities with cases of COVID-19. (B) The relationship between the MHDI-low 0.49 to 0.613 (blue), MHDI-mean 0.614 to 0.737 (orange) and MHDI-high 0.738 to 0.862 (green) groups. (D) The income was divided into low 1.4 to 3.8 (blue) and high 3.9 to 6.4 (orange).

313

314 The results show that most cities with positive cases for COVID-19 are
315 above the national average for both MHDl and per capita income. With the goal
316 to observe the relationships in the data sets in greater detail, we analyzed the
317 MHDl divided into low 0.49 to 0.613 (Figure 6B, blue), mean 0.614 to 0.737
318 (Figure 6B, orange) and high 0.738 to 0.862 (Figure 6B, green) groups. The
319 income was divided into low 1.4 to 3.8 (Figure 6B, blue) and high 3.9 to 6.4
320 (Figure 6B, orange). The socioeconomic index does not show a relationship
321 with the incidence of the disease.

322 Once some of the characteristics of the cities with cases of COVID-19 were
323 described, we sought to determine a model for predicting infection using the
324 time series of confirmed cases.

325

326 2) For July 25, 2020, the evolution model predicts 2,358,703 (2,172,930 to
327 2,544,477) confirmed cases.

328

329 Due to the current level of infection in the cities analyzed, the scarcity of
330 data does not allow the development of a robust predictive model for cases
331 confirmed at the municipal level. To understand the condition of the infection at
332 the national level, we analyzed the time series of accumulated data for
333 confirmed cases. There was a clear upward trend in the number of cases (data
334 not shown). To suggest a prediction for the evolution of COVID-19 cases in
335 Brazil, we use computational modeling in the time series. The best adjusted
336 model for the forecast was ARIMA(2,2,1) using data from the last 30 days,
337 which forecasts 2,358,703 cumulative cases on July 25, 2020, with a 95%
338 confidence interval of 2,544,477 to 2,172,930 (Figure 7).

339

340

341

342

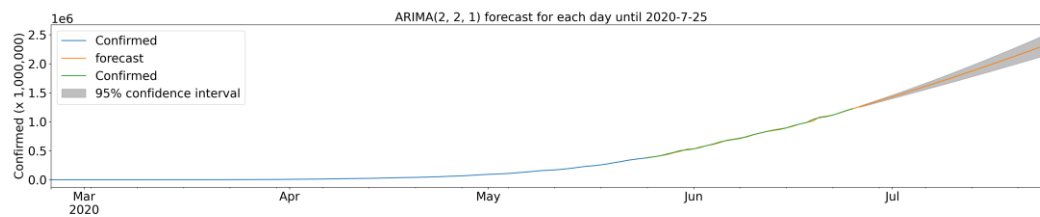


Figure 7. Average estimate of 2,358,703 cumulative confirmed cases in 30 days. ARIMA model of the forecast of confirmed cases until July 25, 2020. Confirmed cases (blue), forecast (orange), model fit analysis (green) and forecast with 95% confidence interval (gray). Up to the end date, between 2,544,477 and 2,172,930 cases are expected.

343

344 CONCLUSION:

345

346 Altogether, 672 cities, accounting for 73.1% of the Brazilian population, had
347 at least 1 case of COVID-19 by June 26, 2020. The age distribution of the
348 inhabitants in those cities, which include the most populous cities of Brazil,
349 remains the same as the average age distribution of the Brazilian population.
350 The average distribution of women and men in the cities studied also
351 corresponds to the national average. The demographic density, the MHDl and
352 the per capita income of the municipalities with cases of COVID-19 are above
353 the national average. However, there seems to be no relationship between the
354 indexes analyzed and the incidence of COVID-19 in these cities, suggesting
355 that other factors (virulence, immune background) may influence the spread of
356 the disease. Our model predicts 2,358,703 cumulative cases (2,172,930 to
357 2,544,477) on July 25, 2020.

358

359

360

361

362

363

364

365

366

367

368 DISCUSSION:

369

370 In this work, we studied the recent occurrence of COVID-19, a respiratory
371 disease caused by the coronavirus, SARS-CoV-2, which originated in the city of
372 Wuhan, China, and analyzed the correlation of transmission and death rates,
373 through confirmed cases using Brazilian demographic and socioeconomic data.
374 Knowledge of the demographic distribution and socioeconomic situation of the
375 population becomes significant when comparing transmission and death rates
376 across the country.

377 Brazil currently has 5,570 municipalities in 26 states distributed
378 disproportionately over a total area of 8,511,000 km² (IBGE, 2019). As
379 described in the results section, municipalities with a population larger than
380 295,955 inhabitants showed a positive correlation between the size of the
381 population and the number of confirmed cases of the disease. However, of all
382 the municipalities, only 95 have a population of over 295,955 inhabitants,
383 including all state capitals and the Federal District. The 95 municipalities most
384 populous represent 1.71% of the country's total cities, with the majority, 94.22%
385 (5,245 municipalities), of Brazilian cities having a population less than or equal
386 to 100,000 inhabitants (IBGE, 2019).

387 Therefore, based on our results, it is possible to assert that the
388 transmission of the disease is more likely to impact less than 2% of Brazilian
389 municipalities. However, it is essential to reinforce that the 95 most populous
390 cities in the country, together, are home to 83,951,535 inhabitants, which
391 represents 40% of the total population of Brazil, with a current demographic
392 density of 205.5 million people (IBGE, 2018). Thus, neglecting the
393 recommendations of the WHO about isolation and social conduct, in a time of a
394 pandemic, is an attack on public health in Brazil.

395 When analyzing the data in relation to the states of the federation, we
396 noted similar observations as those obtained in relation to the municipalities.
397 The state of São Paulo is the most populous in the country and in turn is the
398 state with the highest number of confirmed cases of COVID-19, with 11,043
399 cases confirmed to date, according to the Ministry of Health (2020). It is
400 important to note that the actual numbers of cases and deaths from the disease
401 may be different from the official data, taking into account the impact that the

402 delay in reporting has on the estimates and that reported cases depend on
403 hospitalization (FIOCRUZ, 2020).

404 When we take into account the data obtained, it is possible to show that
405 social isolation is a valid measure to be applied in municipalities that have a
406 resident population larger than 295,000 inhabitants. For these municipalities,
407 the more intense the measures, the flatter the transmission curve becomes;
408 thus, hospitals and health units can have greater control of the situation under
409 the demand of patients who require specialized care.

410 However, it is important to mention that it is not possible to conclude the
411 real importance of social isolation in municipalities with a population below the
412 aforementioned number; however, according to the results, there is a negative
413 correlation between demographic density and the number of cases in these
414 cities, which, in theory, would indicate a lack of connection between these
415 aspects.

416 According to Hellewell et al. (2012), in a study to evaluate measures to
417 contain the transmission of the disease, social isolation is insufficient to control
418 the outbreak, requiring new interventions to achieve control of the transmission
419 of the disease. However, isolation can contribute to spreading the overall size of
420 an outbreak over a longer period of time (Hellewell et al., 2012). Taking this into
421 account, it becomes possible to assess the importance of measures of social
422 isolation, even for municipalities with a small population, demonstrating the
423 great importance of such measures that should be intensified in the most
424 populous cities and not neglected in cities with fewer than 295,000 inhabitants.
425 Therefore, it is interesting to evaluate the average traffic of the Brazilian
426 population mainly in the forms of essential workers, e.g., truck drivers, who
427 supply basic necessities. This supply is primarily carried out in Brazil through
428 land transportation, often long distance, between the capitals and other
429 municipalities of the federation. During the month of March, an average of 1000
430 trucks arrived at the Supply Center of the Federal District every Monday and
431 Thursday (CEASA, 2020).

432 It would be interesting, as a future perspective, to further these studies in
433 order to develop a system of equations that could indicate a proportional factor
434 of the relationship between population and lethality rate to compare the fold-
435 change in the lethality in a more populous city to that of a less populous city.

436 Perhaps, as the number of inhabitants in a city doubles or triples, lethality does
437 not necessarily double or triple, indicating that there is no linearity in the cases.

438 Therefore, the information reported in this study allows us to highlight
439 that cities with a higher number of inhabitants who choose not to comply with
440 social isolation have a higher risk and probability of infection. However, it is
441 essential to show that infection with SARS-CoV-2 is not due to the simple fact of
442 living in a more or less populous city and that there are different regional
443 characteristics, both geographic and socioeconomic, that can influence
444 dispersion, not only for SARS-CoV-2 but also for many other pathogens (Mogi
445 and Spijker, 2020; Dowd et al., 2020).

446 Thus, it is worth emphasizing once again that in the current pandemic
447 context, without effective prophylactic and therapeutic treatments, social
448 isolation has proven to be an efficient measure to control outbreaks in the most
449 populous cities. Therefore, our findings suggest that this approach should
450 indeed not just be simulated but also applied to reduce transmission and avoid
451 hospital demands above service capacity, to provide care for all patients at the
452 micro- and macroregional levels (FIOCRUZ, 2020).

453

454 ACKNOWLEDGMENTS:

455

456 We would like to thank Hellen Geremias dos Santos, PhD, for the critical
457 reading of the work. We would like to thank the Creative Commons Attribution
458 ShareAlike and Álvaro Justen and collaborators from Brazil. We thank IO for the
459 confirmed case database. We would like to thank the scholarship funding
460 agencies CAPES and Fiocruz.

461

462

463

464

465

466

467

468

469

470 BIBLIOGRAPHY

471

472 BARATA, Rita Barradas. **A century of endemic and epidemic diseases.**

473 Departamento de Medicina Social/FCM, Santa Casa de São Paulo, rua Cesário

474 Motta Jr. 61 5^o andar, 01221-020 São Paulo, SP.

475

476 Chang HW, Egberink HF, Halpin R, Spiro DJ, Rottier PJ. Spike protein fusion
477 peptide and feline coronavirus virulence. *Emerg Infect Dis.* 2012;18(7):1089–
478 1095. doi:10.3201/eid1807.120143.

479 CHAN, J. F.-W. et al. A familial cluster of pneumonia associated with the 2019
480 novel coronavirus indicating person-to-person transmission: a study of a family
481 cluster. *The Lancet*, v. 395, n. 10223, p. 514–523, fev. 2020.

482

483 Ceasa orienta caminhoneiros sobre a COVID-19. Distrito Federal, mar. 26,
484 2020. Disponível em: <[http://ceasa.df.gov.br/ceasa-orienta-caminhoneiros-
485 sobre-a-covid-19/](http://ceasa.df.gov.br/ceasa-orienta-caminhoneiros-sobre-a-covid-19/)>. Acesso em: abr.1, 2020.

486 Contini C, Di Nuzzo M, Barp N, Bonazza A, De Giorgio R, Tognon M, Rubino S
487 (2020) The novel zoonotic COVID-19 pandemic: An expected global health
488 concern. *J Infect Dev Ctries* 14;254-264. doi: 10.3855/jidc.12671

489 Coronavirus disease (COVID-19) advice for the public: When and how to use
490 masks. Geneva. Disponível em:
491 <[https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-
492 public/when-and-how-to-use-masks](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks)>. Acesso em: mar. 30, 2020.

493 Coronavírus: 28.320 casos confirmados e 1.736 mortes. Ministério da Saúde.
494 Brasília, abr. 15, 2020. Disponível em :
495 <[https://www.saude.gov.br/noticias/agencia-saude/46738-coronavirus-28-320-
496 casos-confirmados-e-1-736-mortes](https://www.saude.gov.br/noticias/agencia-saude/46738-coronavirus-28-320-casos-confirmados-e-1-736-mortes)>. Acesso em: abr. 15, 2020.

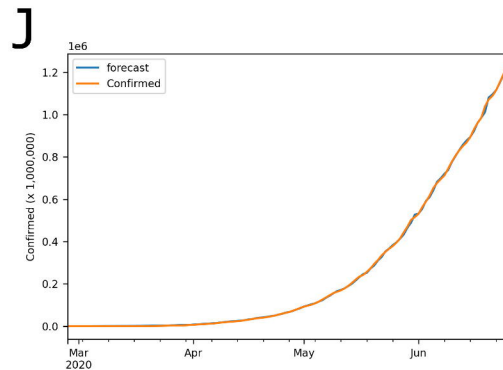
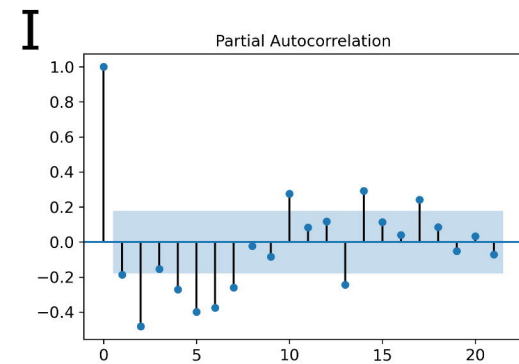
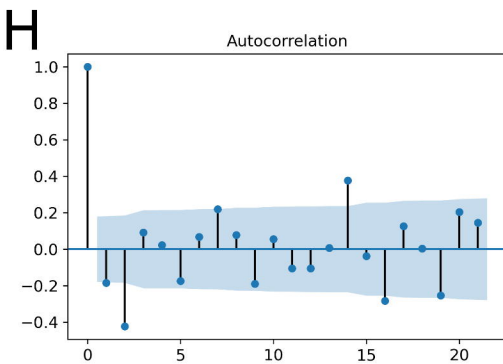
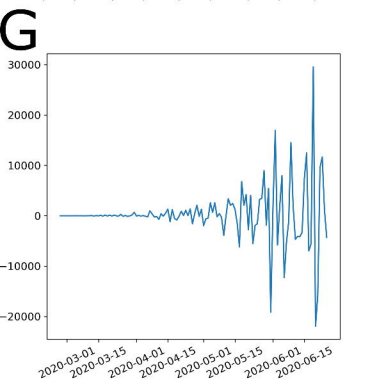
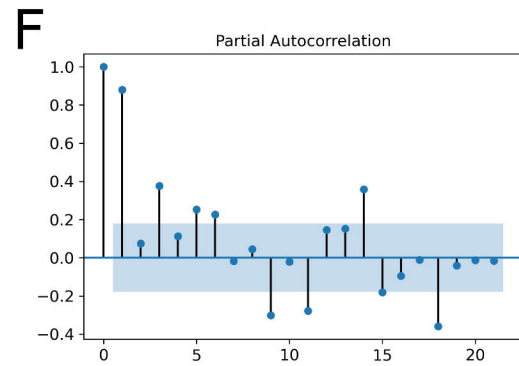
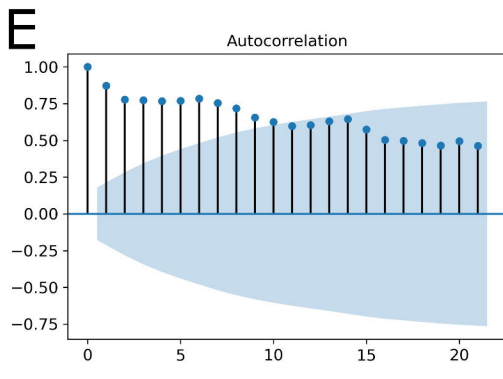
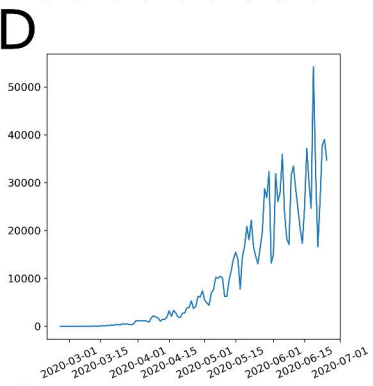
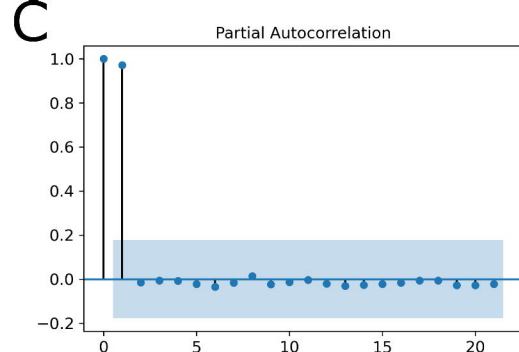
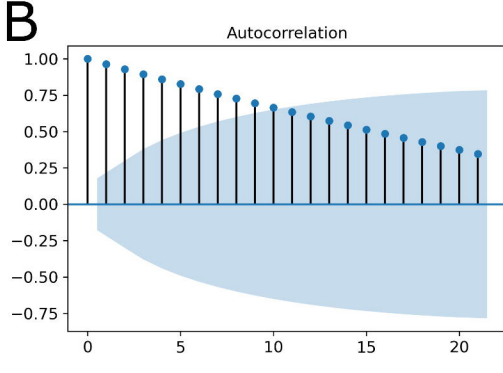
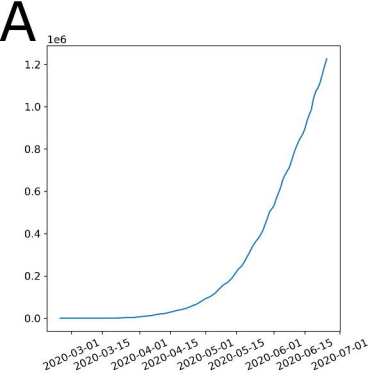
497 Coronavirus disease 2019 (COVID-19) Situation Report – 85. World Health
498 Organization. Geneva, abr. 15, 2020. Disponível em :
499 <[https://www.who.int/docs/default-source/coronaviruse/situation-
500 reports/20200415-sitrep-86-covid-19.pdf?sfvrsn=c615ea20_4](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200415-sitrep-86-covid-19.pdf?sfvrsn=c615ea20_4)>. Acesso em:
501 abr. 15, 2020.

502 Jennifer Beam Dowd, J.B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P.,
503 Ding, X., Liu, Y., Mills, M. C. Demographic science aids in understanding the
504 spread and fatality rates of COVID-19. *The Proceedings of the National
505 Academy of Sciences* – PNAS: abr, 16, 2020.

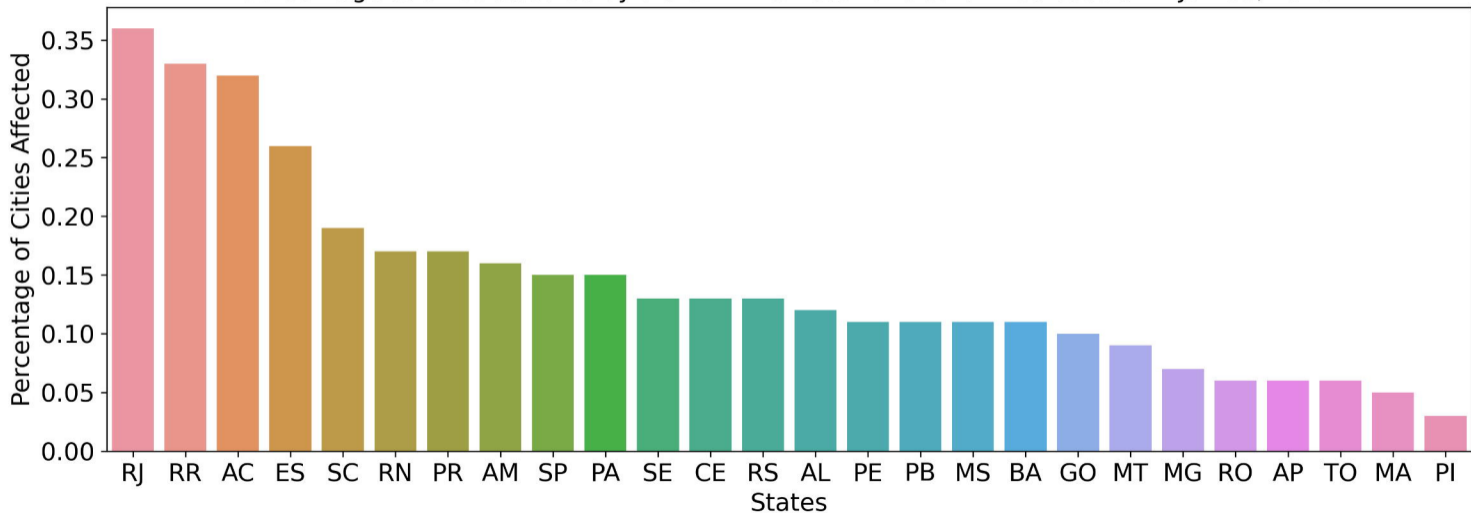
506

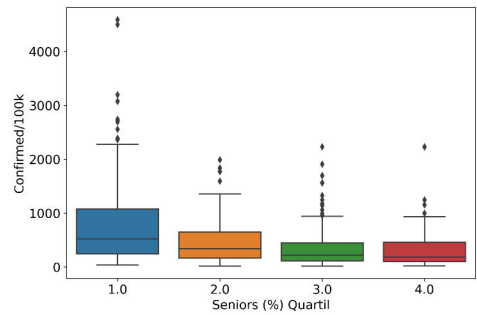
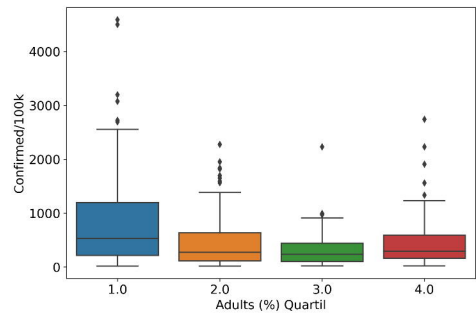
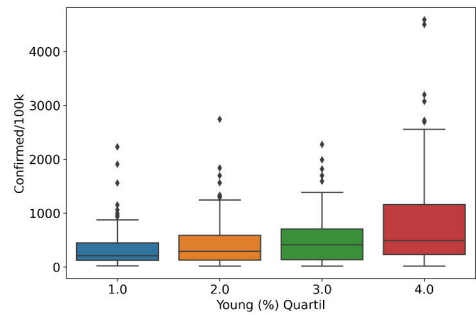
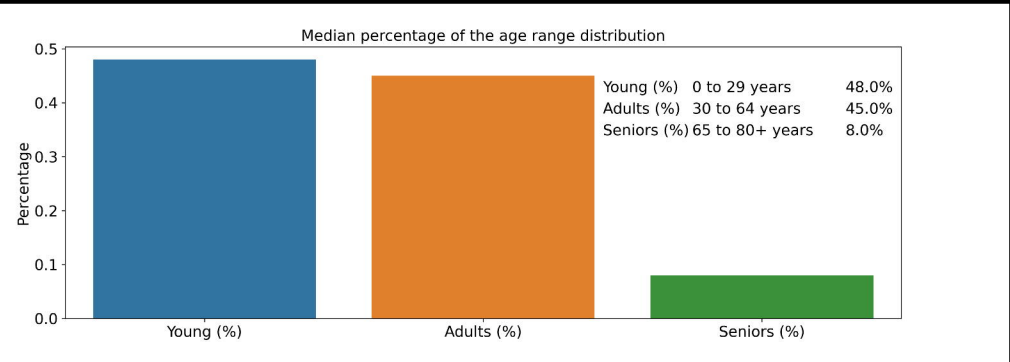
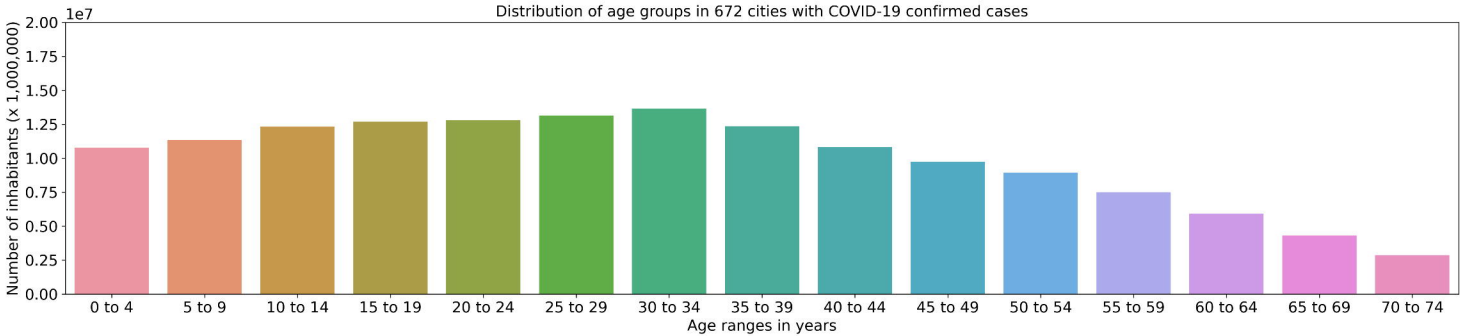
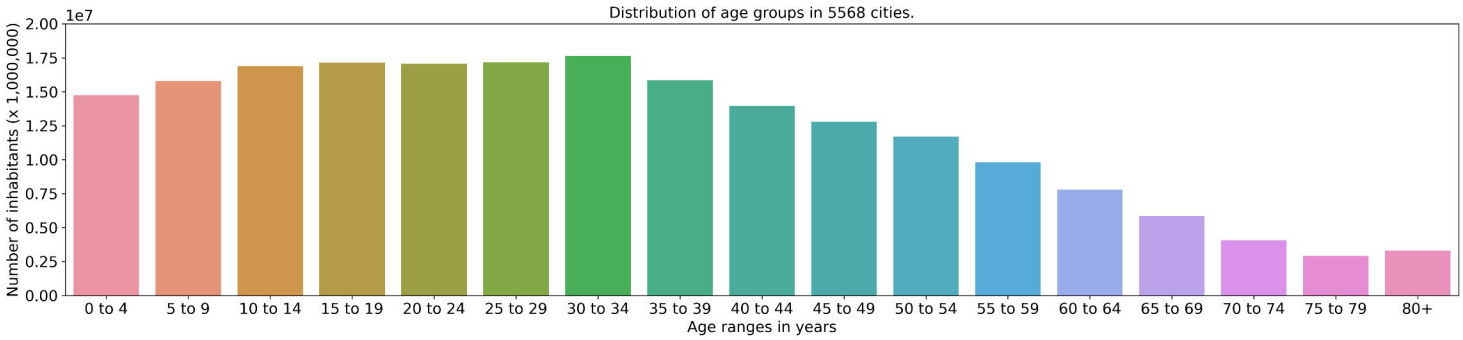
- 507 Gorbalenya, A.E., Baker, S.C., Baric, R.S. *et al.* The species *Severe acute*
508 *respiratory syndrome-related coronavirus*: classifying 2019-nCoV and naming it
509 SARS-CoV-2. *Nat Microbiol* **5**, 536–544 (2020). [https://doi.org/10.1038/s41564-](https://doi.org/10.1038/s41564-020-0695-z)
510 [020-0695-z](https://doi.org/10.1038/s41564-020-0695-z).
- 511 GOULART, A. da C.: **Revisiting the Spanish flu: the 1918 influenza**
512 **pandemic in Rio de Janeiro**. *História, Ciências, Saúde — Manguinhos*, v. 12,
513 n. 1, p. 101-42, Jan.-Apr. 2005.
- 514 HAYS, J. N. ***Epidemics and pandemics. Their impacts on human history.***
515 Austin, Texas. Fundação Kahle, 2005.
- 516
- 517 HELLEWELL, Joel; ABBOTT, Sam; GIMMA, Amy; BOSSE, Nikos ; JARVIS,
518 Christopher; RUSSELL , Timothy W; MUNDAY, James D; EDMUNDS, Adam J
519 W John - Centre for the Mathematical Modelling of Infectious Diseases COVID-
520 19 Working Group - FEASIBILITY OF CONTROLLING COVID-19
521 OUTBREAKS BY ISOLATION OF CASES AND CONTACTS - *Lancet Glob*
522 *Health* 2020; 8: e488–96.
- 523 InfoGripe destaca aceleração de internações por Síndrome Respiratória Aguda
524 Grave. Agência Fiocruz de Notícias, Rio de Janeiro, abr. 30, 2020. Disponível
525 em: [https://portal.fiocruz.br/noticia/infogripe-destaca-aceleracao-de-](https://portal.fiocruz.br/noticia/infogripe-destaca-aceleracao-de-internacoes-por-sindrome-respiratoria-aguda-grave)
526 [internacoes-por-sindrome-respiratoria-aguda-grave](https://portal.fiocruz.br/noticia/infogripe-destaca-aceleracao-de-internacoes-por-sindrome-respiratoria-aguda-grave)>. Acesso em: abr.1, 2020.
- 527 JUSTEN, Álvaro e colaboradores. Boletins informativos e casos do coronavírus
528 por município por dia. Brasil.io, 2020. Disponível em:
529 <https://brasil.io/dataset/covid19/caso>>. Acesso em: 04, abril de 2020.
- 530 MILANESI, Rafaela; CAREGNATO, Rita C. A.; WACHHOLZ, Neiva I.R:
531 **Pandemic Infl uenza A (H1N1): changing population health habits in**
532 **Cachoeira do Sul, Rio Grande do Sul State, Brazil, 2010**. *Cad. Saúde*
533 *Pública*, Rio de Janeiro, 27(4):723-732, abr, 2011.
- 534 Marc Lipsitch, Lyn Finelli, Richard T. Heffernan, Gabriel M. Leung, and Stephen
535 C. Redd; for the 2009 H1N1 Surveillance Group - **Improving the Evidence**
536 **Base for Decision Making During a Pandemic: The Example of 2009**
537 **Influenza A/H1N1- Biosecurity and Bioterrorism: Biodefense Strategy, Practice,**
538 **and Science**Vol. 9, No. 2.
- 539
- 540 Ministério da Economia - Instituto Brasileiro de Geografia e Estatística - IBGE
541 Diretoria de Pesquisas - DPE Coordenação de População e Indicadores

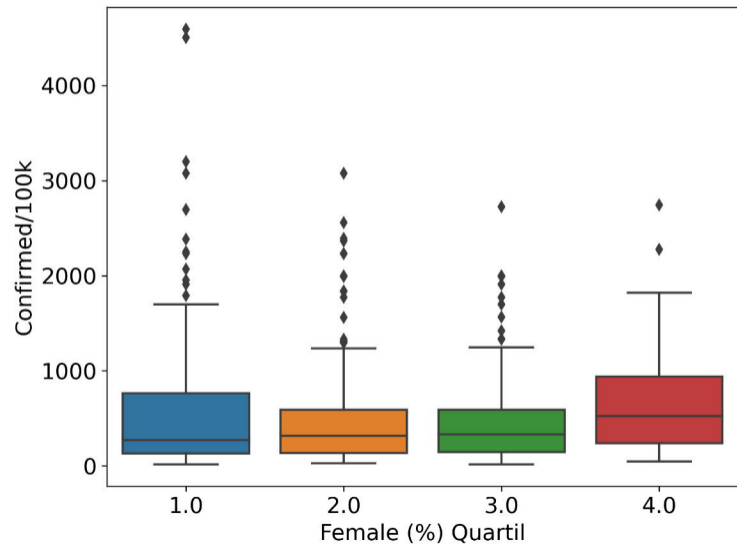
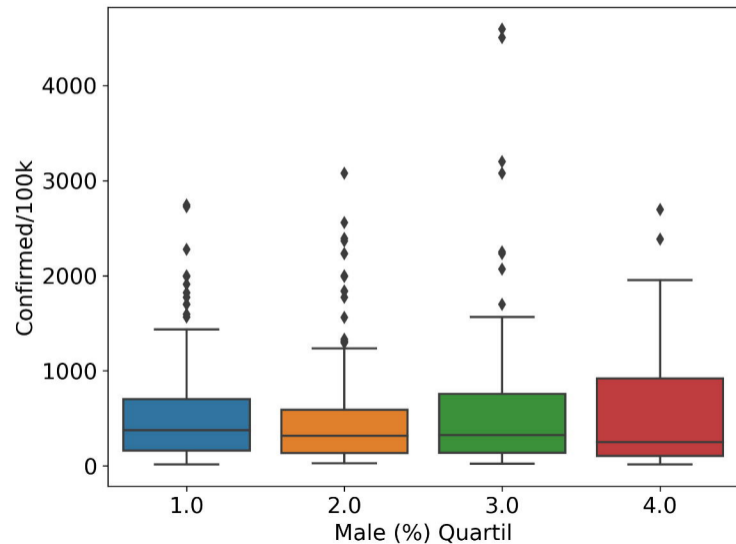
- 542 Sociais – COPIS - **ESTIMATIVAS DA POPULAÇÃO RESIDENTE NO BRASIL**
543 **E UNIDADES DA FEDERAÇÃO COM DATA DE REFERÊNCIA EM 1º DE**
544 **JULHO DE 2019** - *Instituto Brasileiro de Geografia e Estatística (IBGE)*. 28 de
545 agosto de 2019.
- 546 Ministério da Economia - Instituto Brasileiro de Geografia e Estatística - IBGE
547 Diretoria de Pesquisas - DPE Coordenação de População e Indicadores
548 Sociais – COPIS - **ESTIMATIVAS DA POPULAÇÃO RESIDENTE NO BRASIL**
549 **E UNIDADES DA FEDERAÇÃO COM DATA DE REFERÊNCIA EM 1º DE**
550 **JULHO DE 2018** - *Instituto Brasileiro de Geografia e Estatística (IBGE)* em
551 publicação em Diário Oficial da União – Imprensa Nacional – Governo Federal,
552 Brasil, 2018.
- 553 Mogi, R., Spijker, J. The influence of social and economic ties to the spread of
554 COVID-19 in Europe. Centre d'Estudis Demogràfics, Universitat Autònoma de
555 Barcelona. Disponível em: < <https://osf.io/preprints/socarxiv/sb8xn/>>. Acesso
556 em: mar. 30, 2020.
- 557 Sohrabi C, Alsafi Z, O'Neill N, et al. World Health Organization declares global
558 emergency: A review of the 2019 novel coronavirus (COVID-19). *Int J Surg*.
559 2020;76:71–76. doi:10.1016/j.ijsu.2020.02.034.
- 560 UJVARI, Stefan Cunha. **Pandemias: A humanidade em risco**. São Paulo:
561 Contexto, 2011. 220 p. ISBN: 978-8-7244-632-7.
- 562 Van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and Surface Stability
563 of SARS-CoV-2 as Compared with SARS-CoV-1. *N Engl J Med*.
564 2020;382(16):1564–1567. doi:10.1056/NEJMc2004973
- 565 Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with
566 Pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727–733.
567 doi:10.1056/NEJMoa2001017.
- 568 Weiss SR, Leibowitz JL. Coronavirus pathogenesis. *Adv Virus Res*.
569 2011;81:85–164. doi:10.1016/B978-0-12-385885-6.00009-2.
- 570

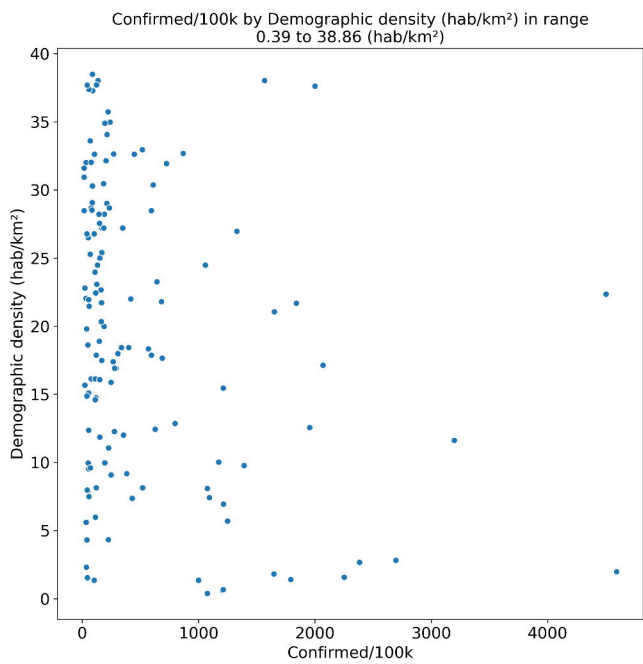
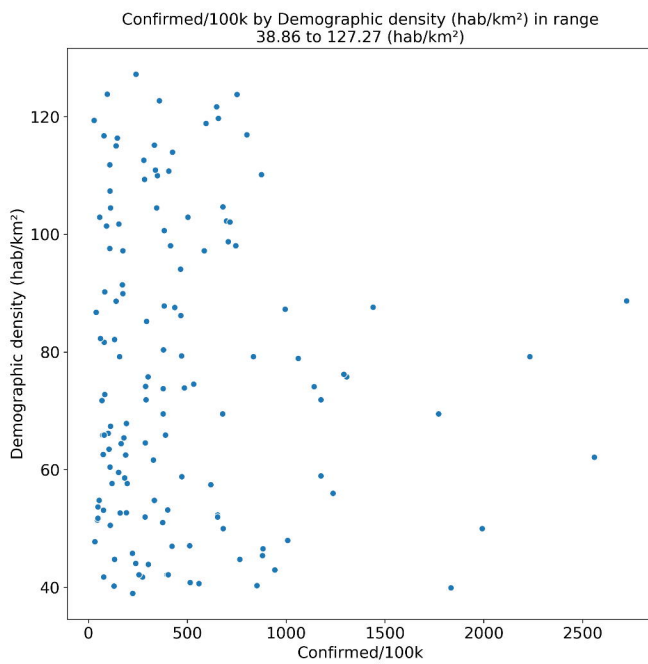
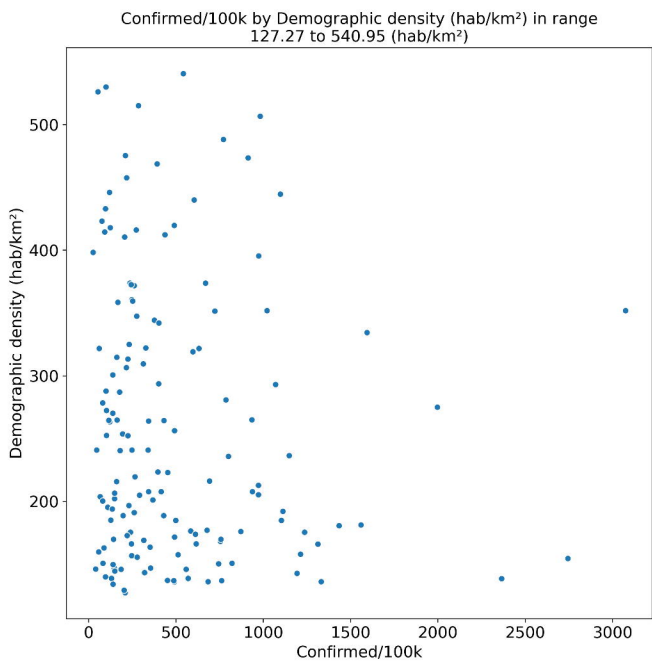
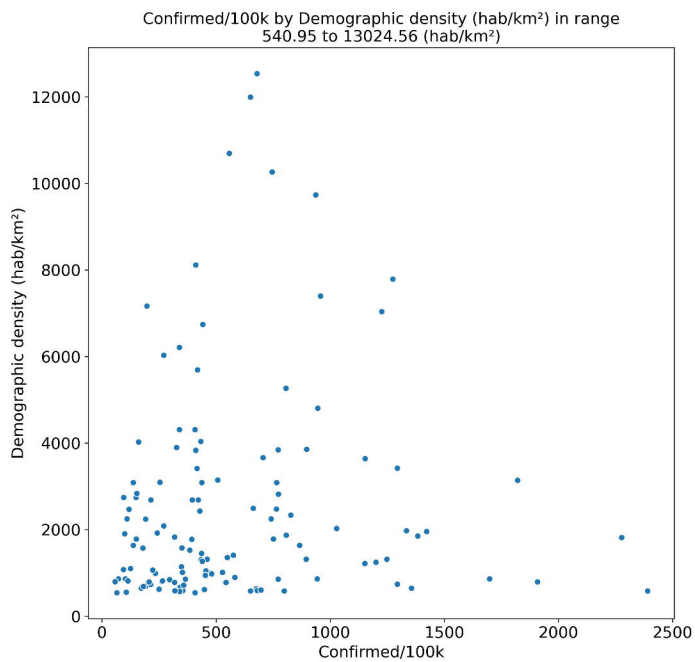
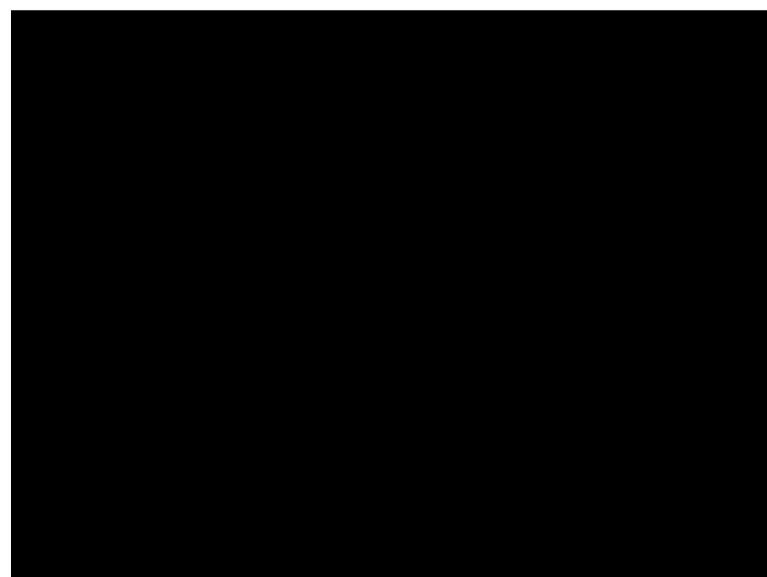
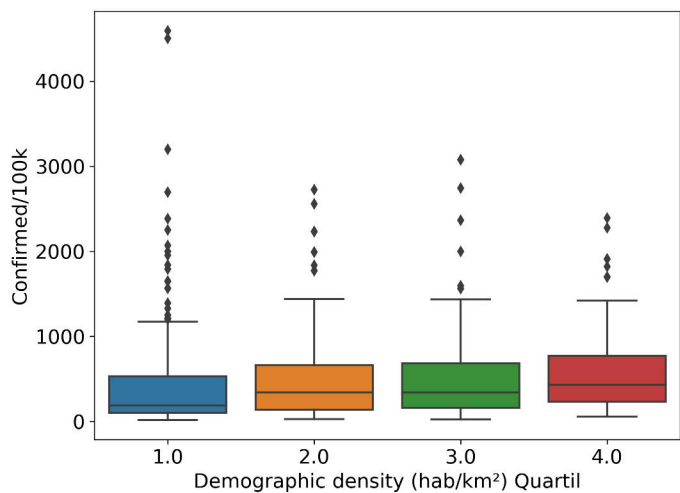


Percentage of cities affected by state with at least one confirmed case until Jun 05, 2020

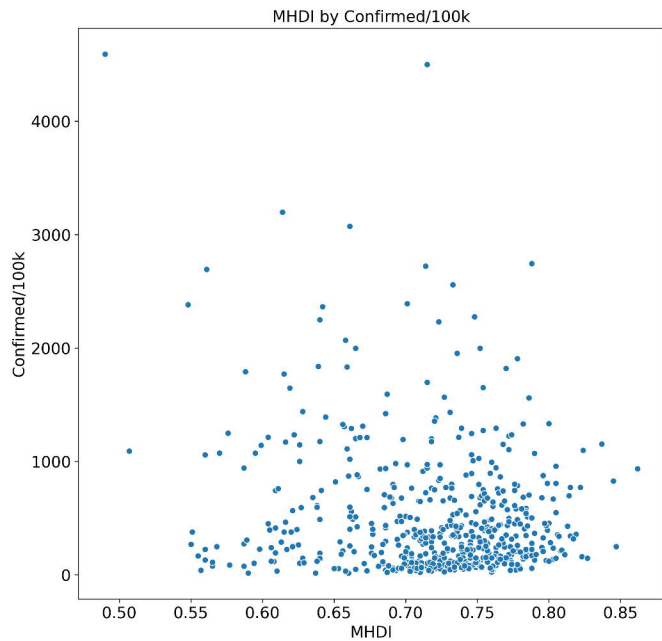




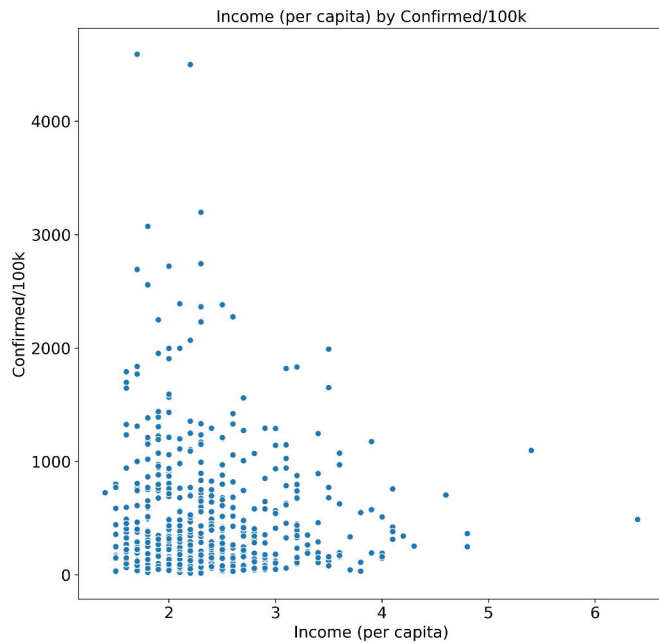


A**B****C****D****E**

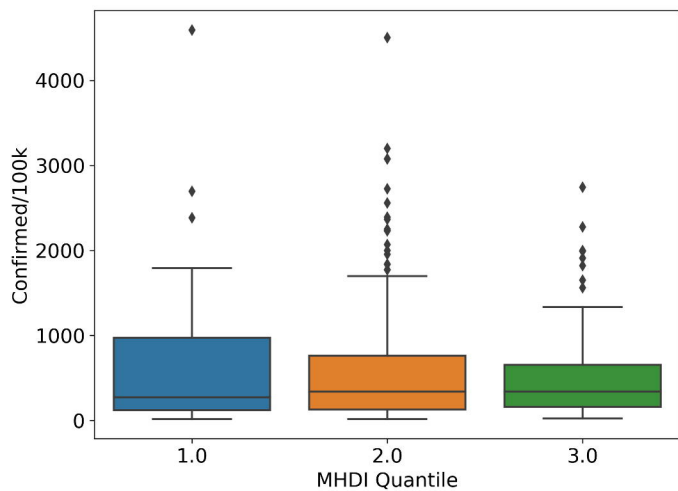
A



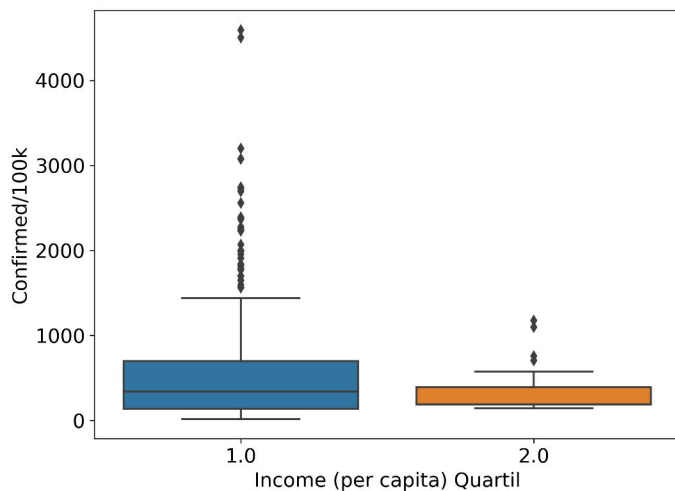
C



B



D



ARIMA(2, 2, 1) forecast for each day until 2020-7-25

