

COVID-19 in Sao Paulo State and City

Trends and data Quality

James R. Hunter, PhD

September 11, 2020

DIPA EPM UNIFESP

- First cases in Brazil
- Probably coming from Italy
- Some time in early February
- First confirmed case - February 26
 - When did it really happen?

Data Sources

- Changes of ministers
 - Luiz Henrique Mandetta (- 4/16)
 - Nelson Teich (- 5/15)
 - Eduardo Pazuello (6/2 -) (interim)
- Political interference in data collection and reporting
 - Interruptions in the reports at the beginning of Pazuello
 - Changes in report formats
 - Confidence in numbers ???

- Generally used for international comparisons
- High degree of confidence
- Has Brazilian national numbers but not subdivisions

- Joint effort of FAPESP and private sector
 - Albert Einstein Hospital,
 - Sirio-Libanes Hospital,
 - Fleury Laboratory
- Published once - 6/30
 - Interesting analyses of tests and evolution
 - What is their definition of the population?
 - Difficult to work with and inconsistent data

- Globo, Estado de Sao Paulo, Folha de Sao Paulo
- Lack of transparency
 - Files not available for inspection

- Data agency of State of Sao Paulo
- Published on their site
- Files available on a GitHub repo
- File with demographic data and comorbidities
 - Sex
 - Age
 - Hospitalization
 - 18 Comorbidities
- Easy to use
- Data quality problems but fewer

Data Quality

Lack of geographic information

- Data limited to city level
- We need information about the precise location of cases
 - Understand the locations where cases occur
 - Relation between population density and infections
- Current theory: infections occur on public transport or within residences
 - True?
 - Geographic data needed to make inferences
 - Currently, we guess

- 3 Most important categories
 - Elderly, children, comorbidities
 - Are elderly really more likely to become infected?
 - With or without other conditions
 - Are children “shedders” or are they non-susceptible?
 - Have or not ACE-2 receptor?
 - Public policies are being based on hypotheses about these issues

- Recent immigrants
 - Venezuela
 - Haiti
 - Sub-Saharan Africa
- Blacks and other people of color
- Indigenous Brazilians

- Observational studies
 - Pedro Hallal Group
 - 120 cities; antibody tests; determine real R_1 for city
- Longitudinal Studies
 - Which groups become infected with what frequency?
 - Degree of severity for subgroups
 - Trends exist or is it by chance?
 - Evolution of comorbidities

- Negative values in new cases and new deaths
- Problems with ages
 - Number of people > 110 years
 - Max 135 year old
 - Only variable with obvious problems
- The City called “Ignorado”
 - Bit like some town from a Western film
 - Name appears when the city was not entered in the data record

Table 1: Casos Mais Recentes da 'Ignorado'

munic	date	cases	deaths	cases_new	deaths_new
Ignorado	2020-07-30	157	20	-366	15
Ignorado	2020-07-31	141	9	-16	-11
Ignorado	2020-08-01	145	12	4	3
Ignorado	2020-08-02	155	14	10	2
Ignorado	2020-08-03	152	10	-3	-4
Ignorado	2020-08-04	189	13	37	3
Ignorado	2020-08-05	154	13	-35	0
Ignorado	2020-08-06	148	13	-6	0
Ignorado	2020-08-07	138	3	-10	-10
Ignorado	2020-08-08	136	0	-2	-3

- Negative values vary from day to day
 - Based on when corrections were incorporated into the database
 - This leaves the total numbers always correct
 - But details continue to vary
- Best strategy for using the data:
 - Make a new download when you want to update the data

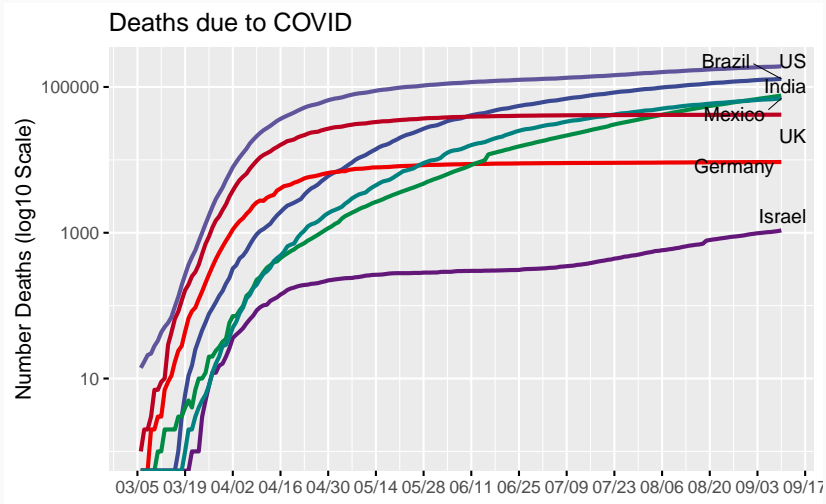
- Affect all the data sources
1. Differentiate between COVID-19 and other respiratory problems without RT-PCR
 2. Delay in delivery of diagnostic test results
 3. Diagnosis via clinical symptoms x PCR test
 4. How to protect the anonymity of patients while collecting demographic data

Conclusion about Quality

- The databases of SEADE are trustworthy
 - Within limits outlined here
- Data from the Health Ministry suffer from political intervention in April and May
 - Have they returned to be trustable? (Who knows?)
- Johns Hopkins - dashboard and high quality data files

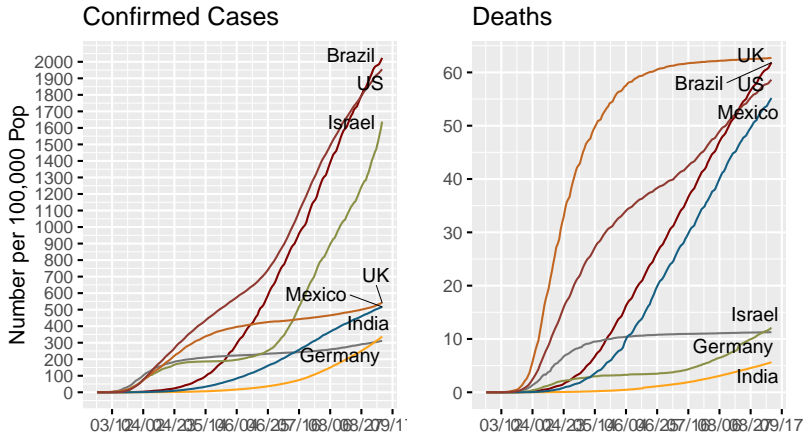
Where Are We in the Crisis in Brazil?

Zoom out - Brazil in the World - 1



Zoom out - Brazil in the World - 2

COVID-19 Cases & Deaths per 100,000 Population



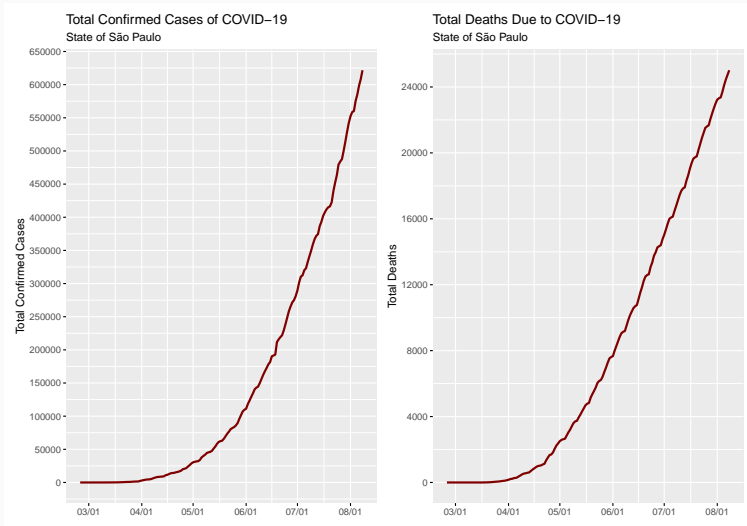
First Cases in Sao Paulo

- First confirmed case: 02/26
 - Principal SEADE file
- First symptomatic cases: 8 cases on 02/04
 - Comorbidities file

Table 2: First Cases – Reported Symptoms 4 February

city	age	sex	death
Limeira	44	female	FALSE
Morungaba	21	female	FALSE
Sao Paulo	57	male	FALSE
Ilhabela	45	female	FALSE
Sao Paulo	44	female	FALSE
Conchas	49	male	FALSE
Sao Paulo	58	female	FALSE

Evolution of Cases and Deaths in the State of São Paulo



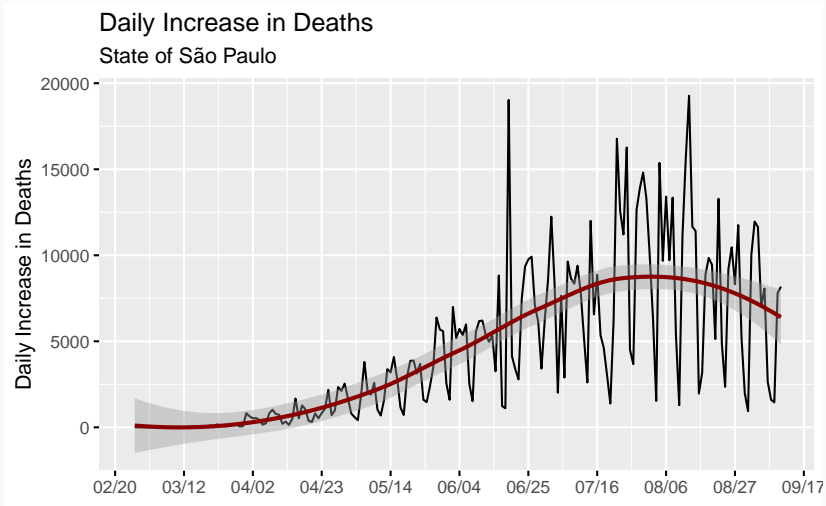
What Does the Curve Show?

- Typical exponential growth for the initial phase of a new disease
- Does not show typical evolution of an epidemic
 - Shape of a Sigmoid or Gompertz curve (yet)
 - Sigmoid curve will flatten as it approaches a given limit
- Cumulative curves are slow to show changes in trends
 - Reduction in the rate of deaths in the City of SP

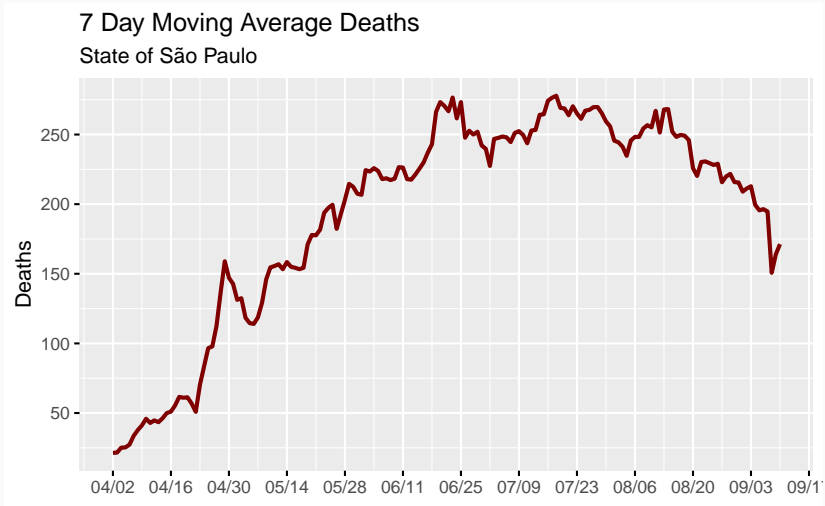
Trends: Daily Counts, Moving Averages, Adjustments for Population

Estado de São Paulo

Trend of Deaths in the State of Sao Paulo

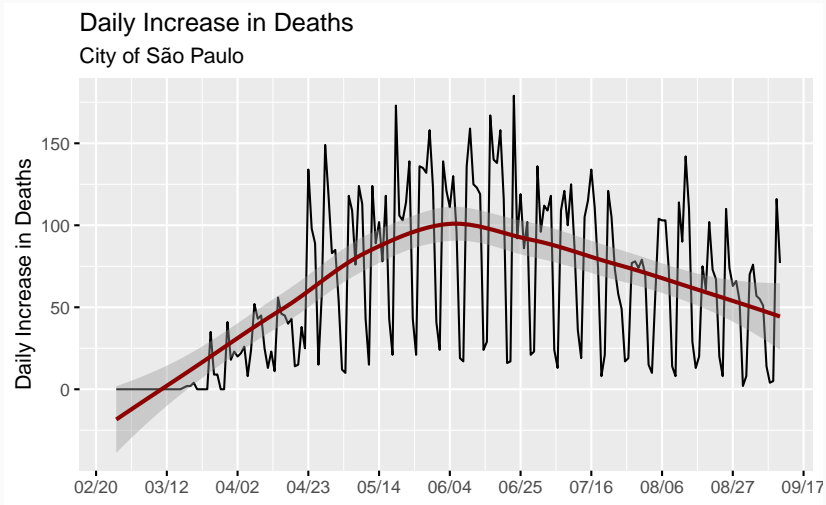


State – Deaths – 7 Day Moving Average

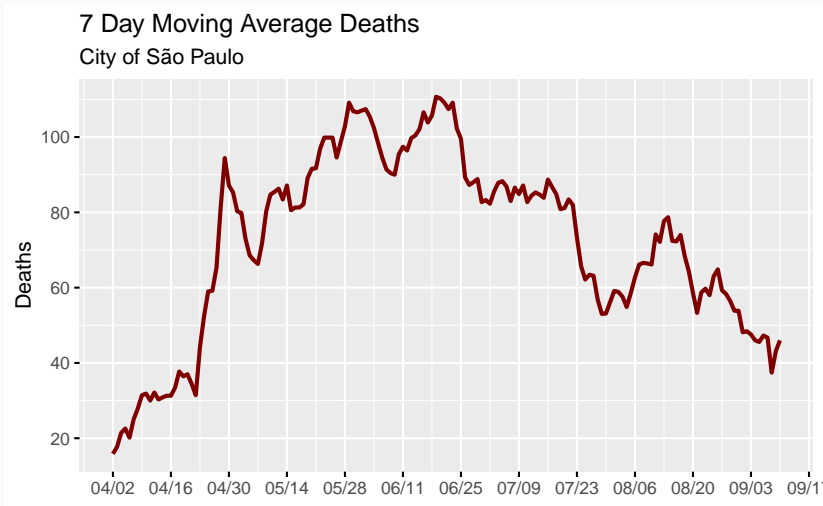


City of Sao Paulo

Deaths in the City of Sao Paulo



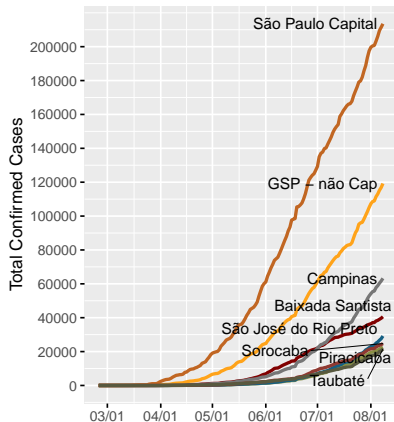
City of Sao Paulo – Deaths – 7 Day Moving Average



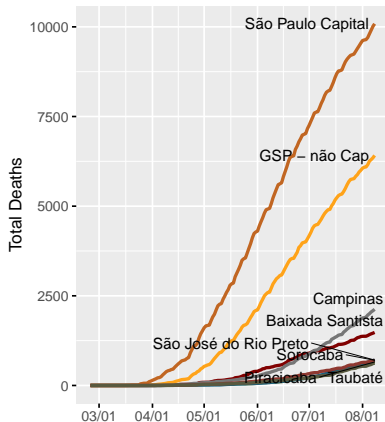
Regions of the State

Totals of Cases and Deaths

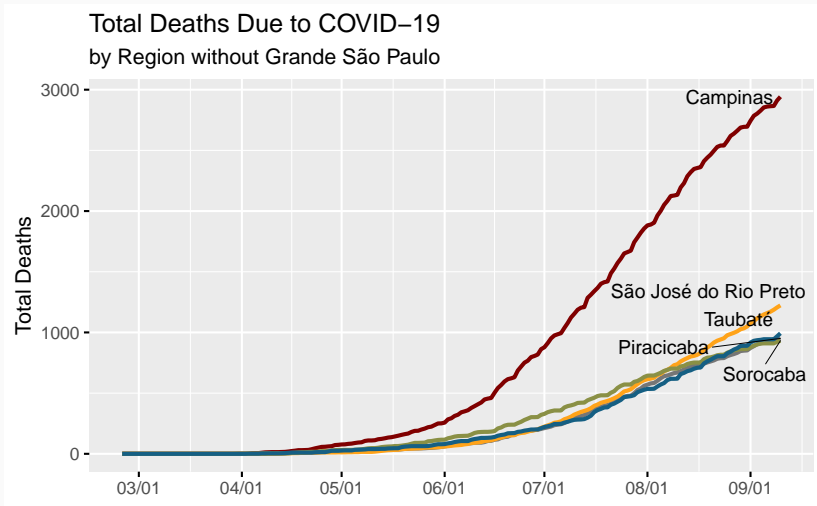
Total Confirmed Cases of COVID-19
by Region



Total Deaths Due to COVID-19
by Region

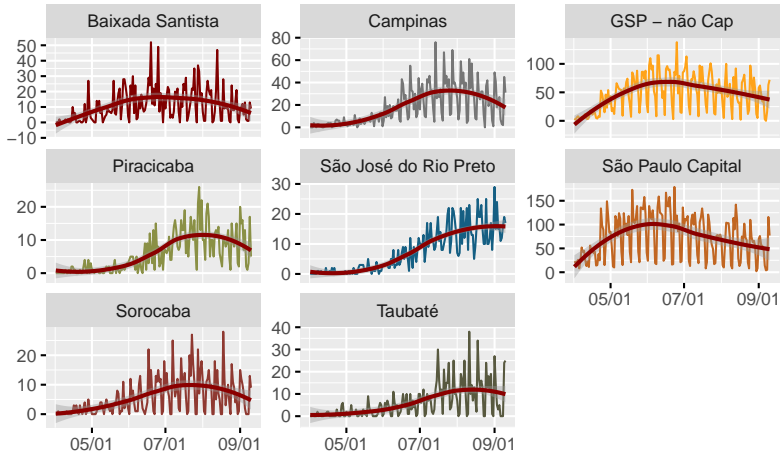


Deaths by Region (without Metro Sao Paulo)



Regions – Daily Deaths

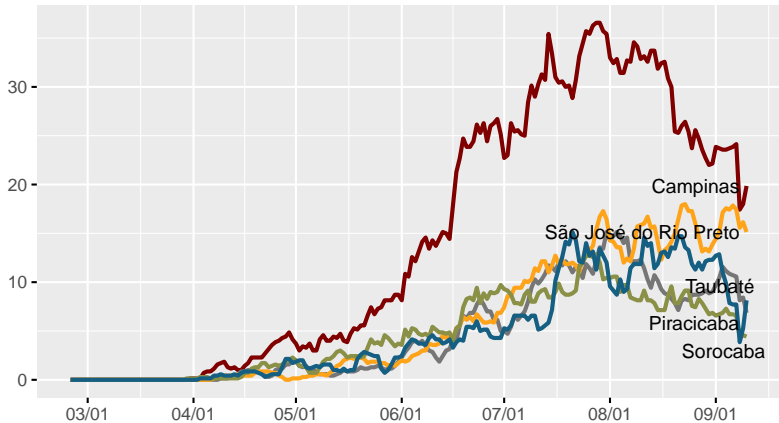
Daily Deaths by Region



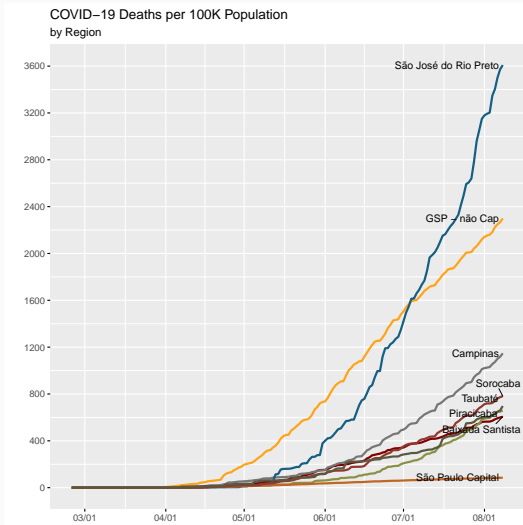
Regions – 7 Day Moving Average

COVID-19 Deaths – 7 Day Moving Average

by Region without Grande São Paulo



Regions per 100,000 Population



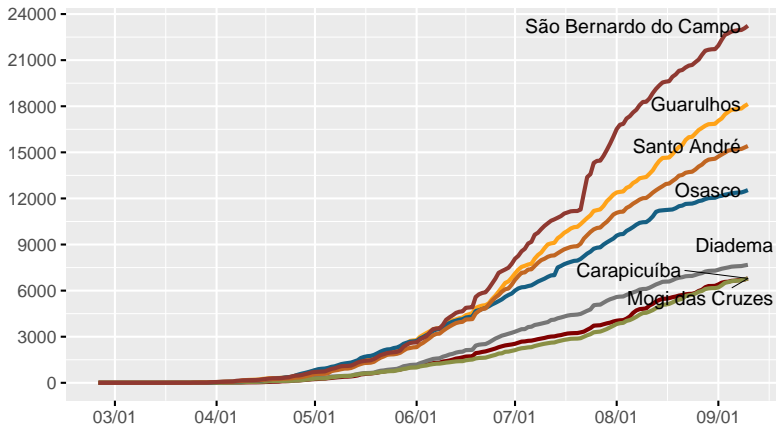
Metropolitan Sao Paulo

8 Cities with Greatest Number of Cases

1. São Paulo
2. São Bernardo do Campo
3. Guarulhos
4. Santo André
5. Osasco
6. Diadema
7. Carapicuíba
8. Taboão da Serra

Cases in the Top 8

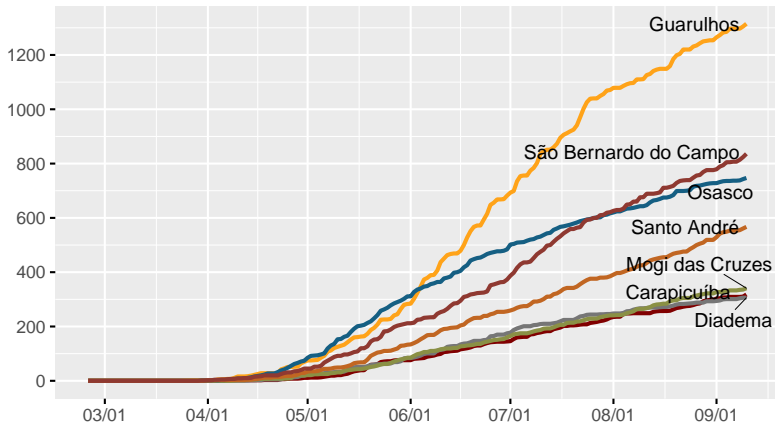
Confirmed Cases -- Grande São Paulo
Top 7 Cities without São Paulo



Deaths in the Top 8

COVID-19 Deaths -- Grande São Paulo

Top 7 Cities without São Paulo



Demographics of the Victims of COVID-19 in Sao Paulo

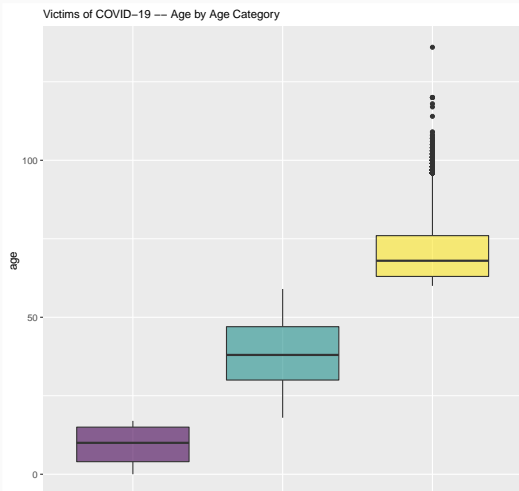
- Arquivo SEADE: *casos_obitos_doencas_preexistentes.csv*
- August 8
- available in GitHub repository
- 627,101 cases (after 1st level of cleaning)

Victims' Ages

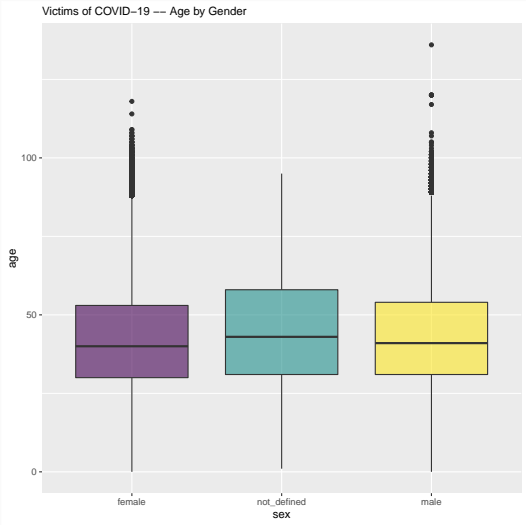
- Mean: 42.76
- Median: 41,
- Standard Deviation: 17.34

Boxplot of Ages by Category

- Child (< 18)
- Adult (18 - 59)
- Elderly (60 or older)



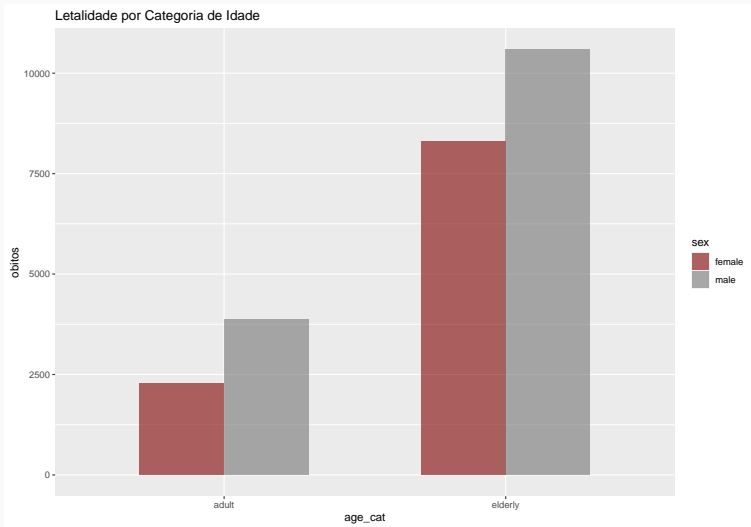
Idade por Gênero



Lethality of COVID-19

- Lethality: 4.02%
- Based on
 - 25016 deaths
 - 621731 cases

Lethality by Age Group



Deaths by Age Group

- Elderly die at a higher rate than the general population?
- Contrast between adults (18 - 60) and elderly (60 +)

```
##           obito
##           TRUE  FALSE
##  adult      6143 485158
##  elderly  18897 87618

##
## 2-sample test for equality of proportions with continuity correction
##
## data:  table(data_co$age_cat, obito)[2:3, ]
## X-squared = 59314, df = 1, p-value < 2.2e-16
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.1672289 -0.1625873
## sample estimates:
##      prop 1      prop 2
## 0.01250354 0.17741163
```

Comorbidities

The most frequently appearing among the cases

- Minimum of 3,000 cases
 - Omitting “outro” (other) category
1. Cardiopathy – 32,852
 2. Diabetes – 24,230
 3. Obesity – 4,459
 4. Neurological Disease – 4,195
 5. Kidney Disease – 3,846
 6. Lung Disease – 3,543
- Comorbidities reported in +/- 5% of all cases

- Criterion for including a report of comorbidity
 - Consistency from doctor to doctor, hospital to hospital?
- Has impact on the relation between comorbidities and date rates from COVID-19

Presentation Available:

www.github.com/jameshunterbr/BR_Covid-Analysis
