

COVID-19 em São Paulo

Tendências e Qualidade dos Dados

James R. Hunter, PhD

12 de agosto de 2020

DIPA EPM UNIFESP

- Primeiros casos em Brasil
- Provavelmente provenientes da Itália
- Em algum momento de fevereiro
- Primeiro caso “confirmado” - 26 de fevereiro
 - Quando realmente ocorreu?

- Ministério de Saúde
 - Trocas de ministros
 - Política intrometendo nos dados
 - Interrupções nos relatórios no início de Pazuello
 - Mudanças no formato dos relatórios
 - Confiabilidade ???
- Johns Hopkins University
 - Geralmente usado para dados comparativos internacionais
 - Confiáveis
 - Tem Brasil como país mas não subdividido

- FAPESP, Einstein, Sírio, Fleury
 - Publicado uma vez - 30 de junho
 - Pode ter análises interessantes dentro
 - Qual é a definição da população?
 - Difícil manusear
- Grupo dos Meios de Comunicação
 - Globo, Estadão, Folha
 - Falta de transparência
 - Arquivos não disponíveis

- Publicação no site
- Arquivos disponíveis no repo de GitHub
 - <https://github.com/seade-R/dados-covid-sp>
- Arquivo com dados demográficos e comorbidades
 - Gênero
 - Idade
 - Internação
 - 18 Comorbidades
- Mais fácil a usar
- Tem problemas também, mas menos

Qualidade de Dados

Falta de informação sobre geolocalização

- Dados limitados ao nível de municipalidade
- Precisamos informação sobre o local dos casos
 - Entender os locais onde infecções ocorrem
 - Relação entre densidade de população e infecção
- Teoria atual: infecções ocorrem nos meios de transporte público e dentro das residências
 - Verdade?
 - Precisa dados geográficos para tirar inferências sobre isso
 - Agora, chutamos

- 3 Categorias Mais Importantes
 - Idosos, crianças, comorbidades
 - Idosos mais susceptíveis?
 - com ou sem comorbidades
 - Crianças como “shedders” ou não-susceptíveis?
 - Políticas públicas baseadas em hipóteses sobre essas questões
- Outras categorias de interesse
 - Imigrantes recentes
 - Negros e outras pessoas de cor
 - Indígenas

- Estudos observacionais
 - Estudo do grupo de Pedro Hallal
- Estudos longitudinais
 - Quais grupos se infeccionam com qual frequência?
 - Grau de severidade para subgrupos
 - Existe tendência ou aleatória?
 - Evolução das Comorbidades

- Valores negativas em casos novos e mortes novas
- Problemas com as idades
 - Número de pessoas com > 110 anos
 - Max 135 anos
 - Única variável com problemas óbvios
- Cidade de “Ignorado”
 - Parece uma cidade num faroeste de John Wayne
 - Nome aparece quando o nome da cidade não foi entrado no registro

Tabela de “Ignorado”

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 |

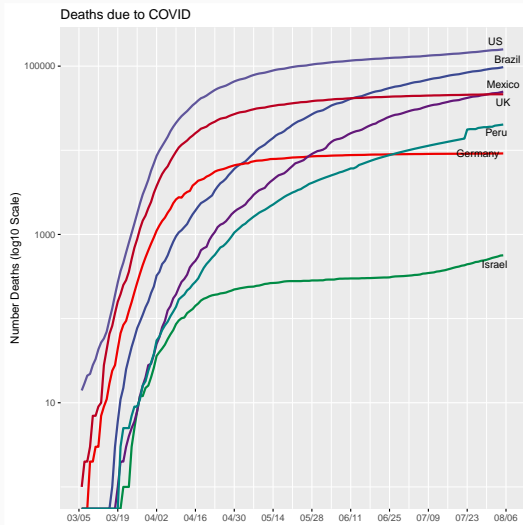
- Valores negativos variam de dia em dia
 - Baseado em quando as correções estão incorporados na base
 - Apesar os totais são sempre certos
 - Os detalhes podem variar
- Estratégia segura de usar os dados: faça um novo download do arquivo quando quer atualizar os dados.
- Use: **`dados_covid_sp.csv`**

- Afetam todos os fontes de dados
1. SRAG - diferenciar entre COVID-19 e outros problemas respiratórios sem RT-PCR
 2. Demora na entrega dos resultados dos testes diagnósticos
 3. Diagnóstico via exame clínica x teste PCR
 4. Como preservar anonimato dos pacientes enquanto coleccionar os dados demográficos

- As bases de dados de SEADE são confiáveis
 - Dentro das limitações de todos os dados desta doença
- Dados do Ministério sofreram das intervenções políticas em abril e maio
 - Voltaram a ser confiáveis? (Não sei)
- Johns Hopkins - dashboard e arquivos de alta qualidade

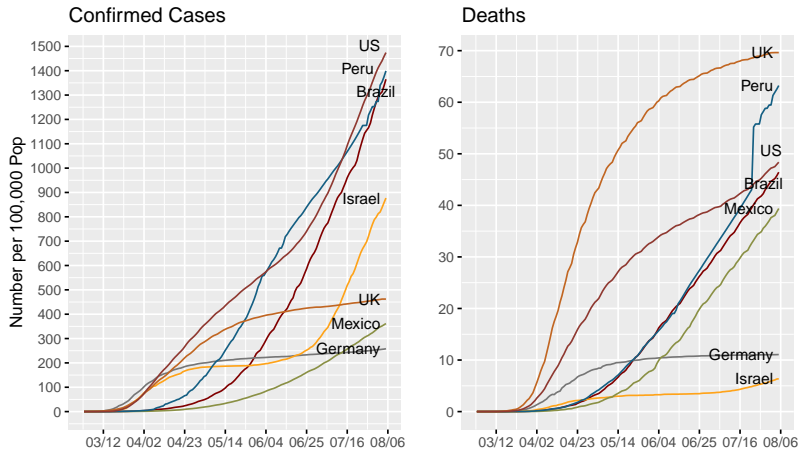
Em qual fase da crise estamos aqui
em São Paulo?

Zoom out - Brasil no Mundo - 1



Zoom out - Brasil no Mundo - 2

COVID-19 Cases & Deaths per 100,000 Population



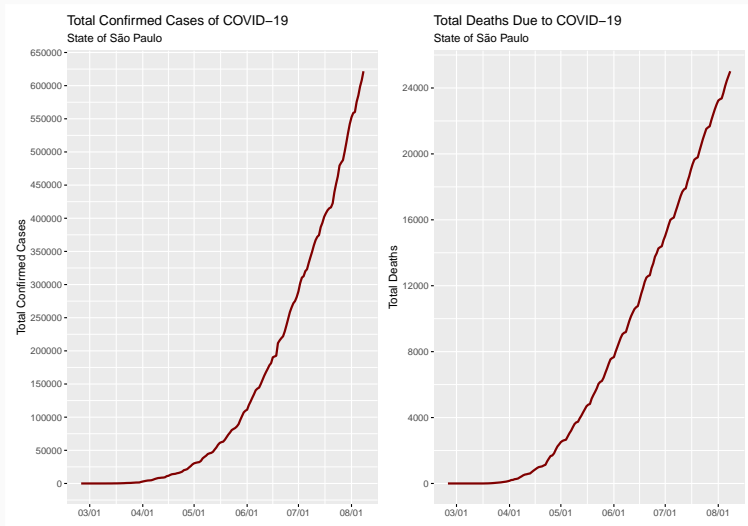
Primeiros Casos em São Paulo

- Primeiro caso confirmado: 26/02
 - arquivo principal de SEADE
- Primeiros casos de sintomas: 8 casos no 04/02
 - arquivo comorbidades

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 |

Evolução dos Casos e Óbitos no Estado de São Paulo

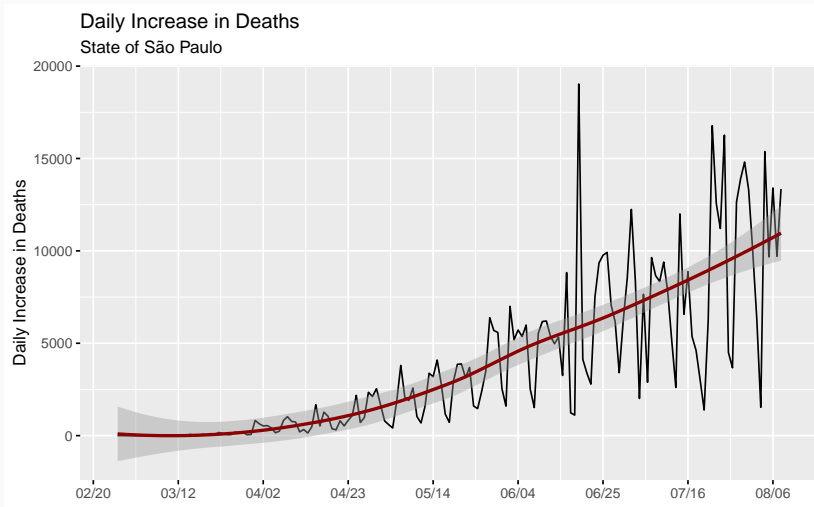


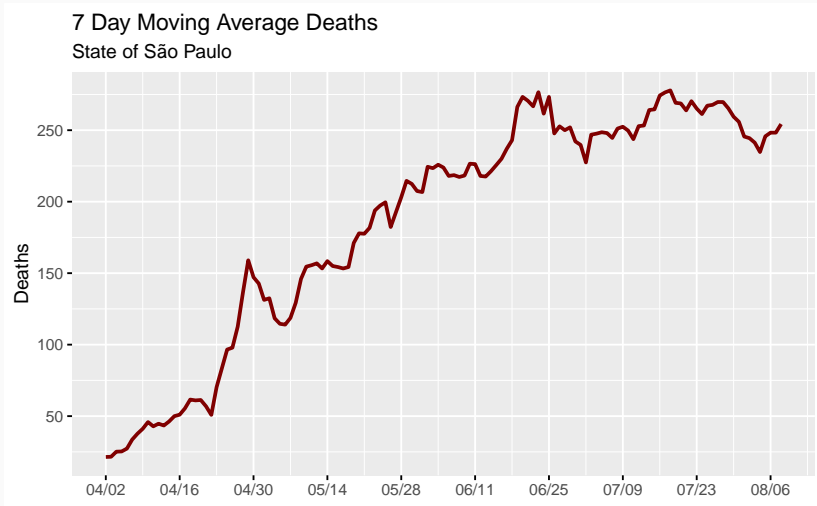
- Mostra crescimento exponencial típico da primeira fase de qualquer doença nova
- Mas, não mostra a evolução típica de uma epidemia: formato de curva sigmoide ou Gompertz (*ainda*)
 - Curva sigmoide achata quando aproxima a um limite.
- Curvas cumulativas são lentas para mostrar mudanças nas tendências
 - Redução na taxa de óbitos na cidade de SP

Tendências: Contagem Diária, Média Móvel, Ajustado por População

Estado de São Paulo

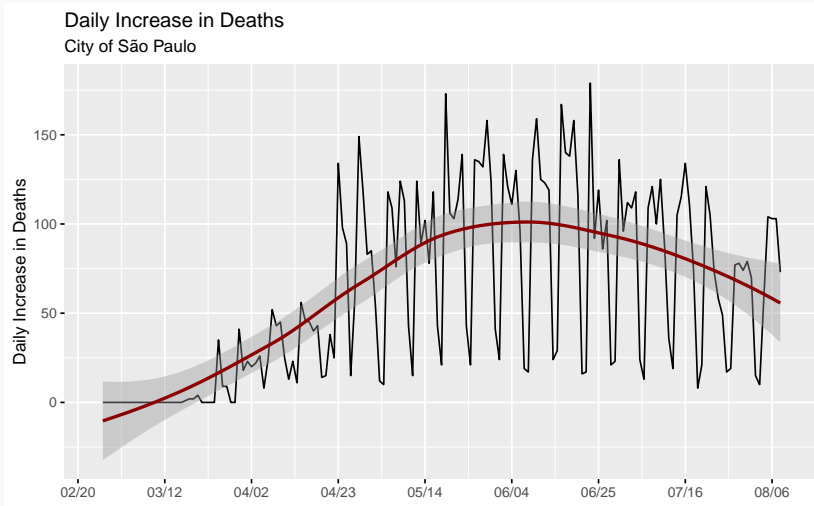
Tendência dos Óbitos no Estado de São Paulo



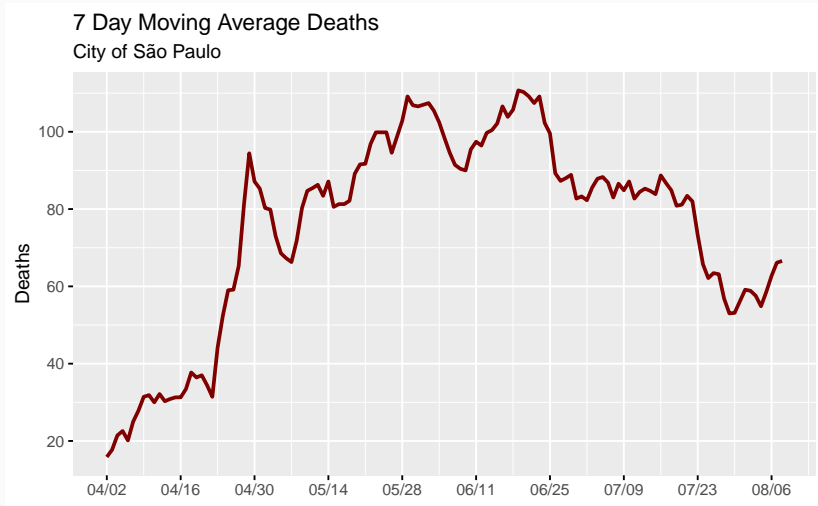


Cidade de São Paulo

Óbitos na Cidade de São Paulo



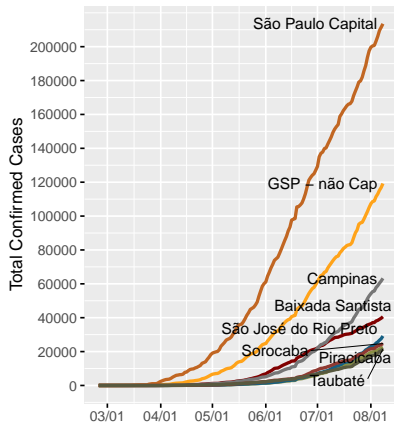
Cidade de São Paulo – Óbitos – 7 Dia Média Móvel



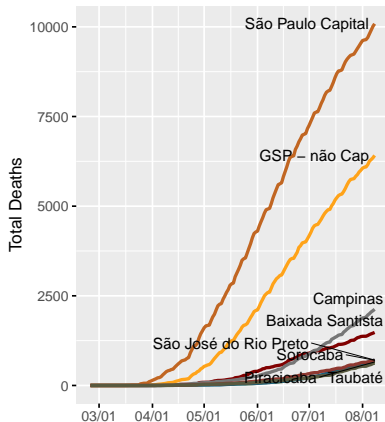
Regiões do Estado

Totais de Casos e Óbitos

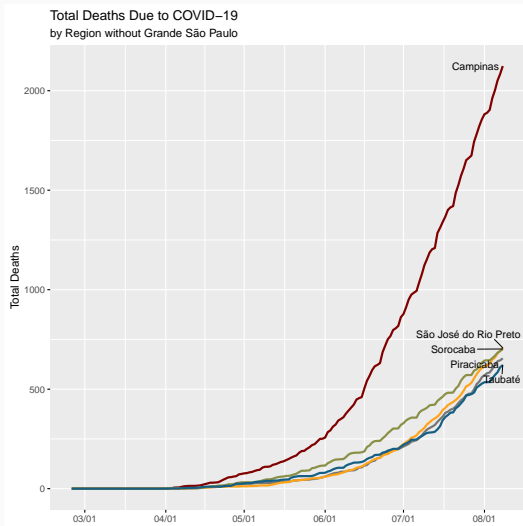
Total Confirmed Cases of COVID-19
by Region



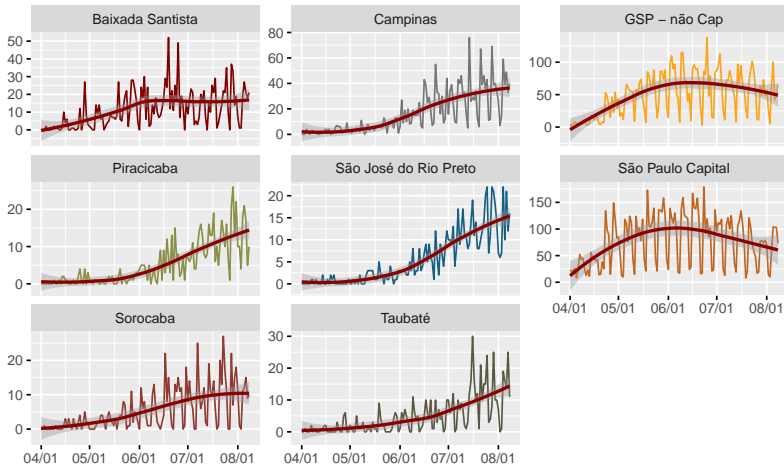
Total Deaths Due to COVID-19
by Region



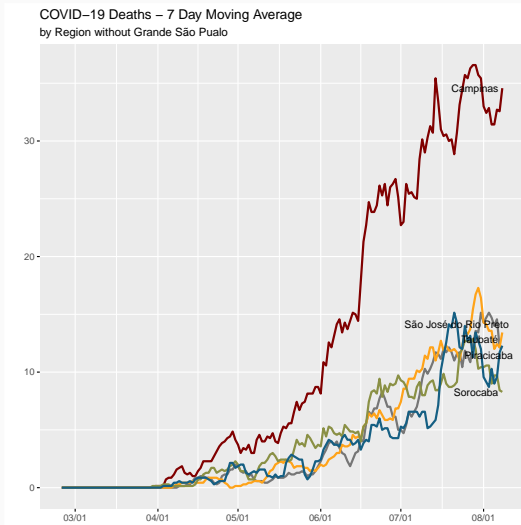
Óbitos por Região sem Grande São Paulo



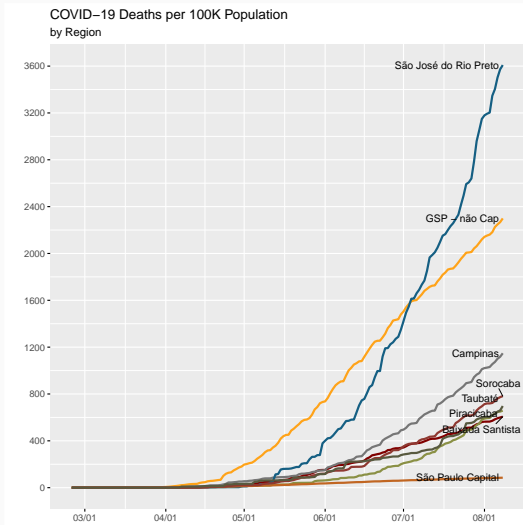
Daily Deaths by Region



Regiões – 7 Dia Média Móvel



Regiões por 100.000 População



Grande São Paulo

8 Cidades com Número Maior de Casos

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

Gráfico dos Casos nos Top 8

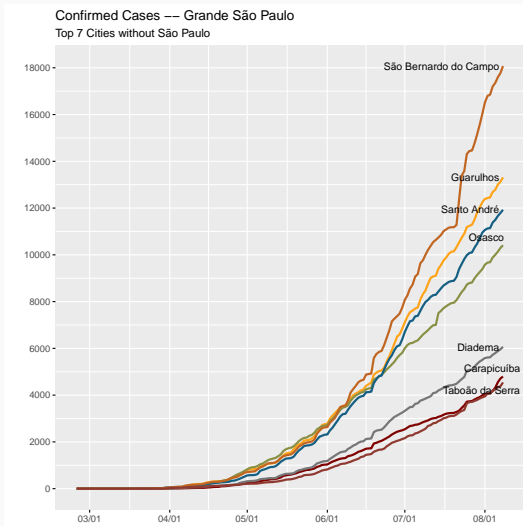
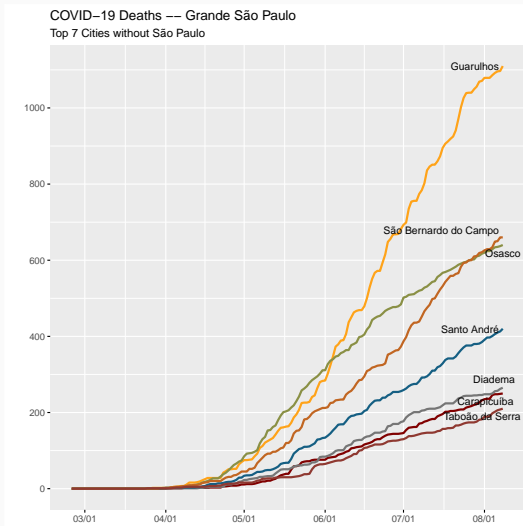


Gráfico dos Mortos nos Top 8



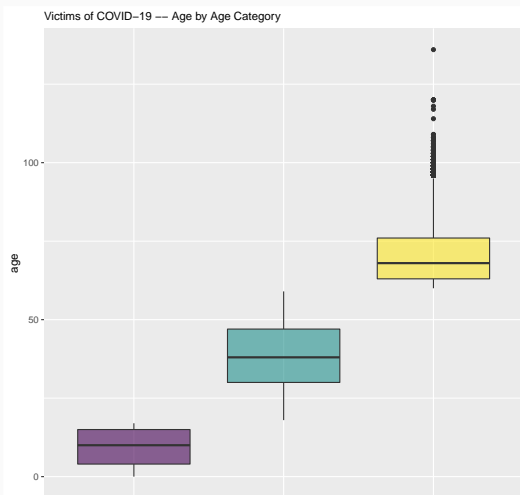
Demografia das Vítimas do COVID-19 em São Paulo

- Arquivo SEADE: *casos_obitos_doencas_preexistentes.csv*
- 8 de Agosto
- disponível no repositório GitHub
- 627.101 casos (depois de 1º grau de limpeza)

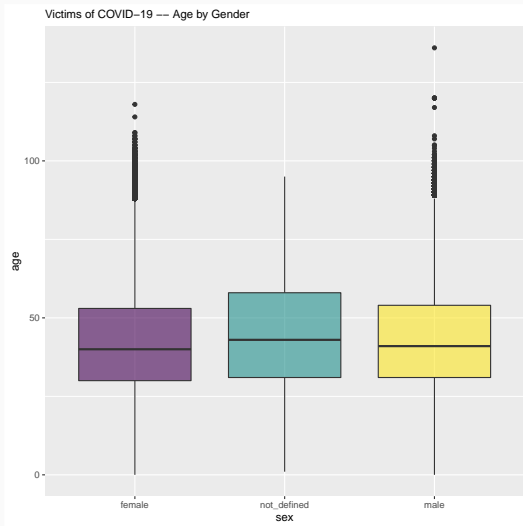
- Média: 42.76, Mediana: 41, Desvio Padrão: 17.34

Boxplot das Idades por Categoria

- Criança (< 18 anos)
- Adulto (18 - 59 anos)
- Idoso (60 ou mais)



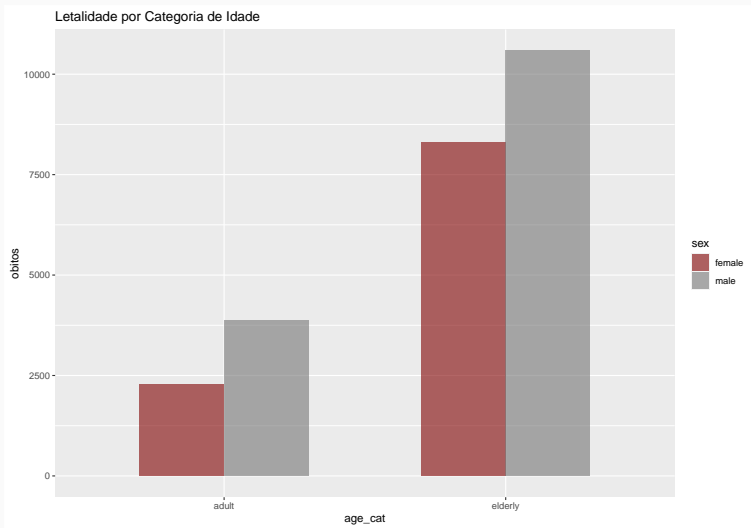
Idade por Gênero



Letalidade da COVID-19

- Taxa de Letalidade: 4.02%
- Baseado em
 - 25016 óbitos
 - 621731 casos

Letalidade por Grupo de Idade



Mortes por Idade

- Os idosos morrem numa proporção acima da população em geral?
- Contraste entre adultos (18 - 60) e idosos (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
```

Comorbidades

Os mais frequentes entre os casos

- Mais de 3.000 casos
 - Não categoria “outro”
1. Cardiopatia – 32.852
 2. Diabetes – 24.230
 3. Obesidade – 4.459
 4. Doença Neurológica – 4.195
 5. Doença Renal – 3.846
 6. Doença Pulmonar – 3.543
- Comorbidades relatadas em +/- 5% dos casos

- Questão do critério para relato de uma comorbidade
 - Tem consistência de medico em medico, hospital em hospital?
- Questão importante para a relação entre comorbidades e óbitos