

1 **Title:** CovidCounties - an interactive, real-time tracker of the COVID-19 pandemic at the level of
2 US counties

3

4 **Authors:** Douglas Arneson¹, Matthew Elliott¹, Arman Mosenia^{1,2}, Boris Oskotsky¹, Rohit
5 Vashisht¹, Travis Zack^{1,3}, Paul Bleicher⁴, Atul J. Butte^{1,5,6}, Vivek A. Rudrapatna^{1,7}

6

7 **Affiliations:**

8 1. Bakar Computational Health Sciences Institute, University of California, San Francisco, CA

9 2. School of Medicine, University of California, San Francisco, CA

10 3. Department of Medicine, University of California, San Francisco, CA

11 4. Evident Health Strategies, West Newton, MA

12 5. Department of Pediatrics, University of California, San Francisco, CA

13 6. Center for Data-Driven Insights and Innovation, University of California Health, Oakland, CA

14 7. Division of Gastroenterology, Department of Medicine, University of California, San Francisco,
15 CA

16

17 **Corresponding Author:**

18 Vivek A. Rudrapatna: 550 16th Street, 4th Floor #4748, San Francisco, CA 94158-2549. tel:

19 415.514.0542. email: Vivek.Rudrapatna@ucsf.edu

20

21

22

23

24 **Abstract:**

25 Management of the COVID-19 pandemic has proven to be a significant challenge to policy
26 makers. This is in large part due to uneven reporting and the absence of open-access
27 visualization tools to present local trends and infer healthcare needs. Here we report the
28 development of CovidCounties.org, an interactive web application that depicts daily disease
29 trends at the level of US counties using time series plots and maps. This application is
30 accompanied by a manually curated dataset that catalogs all major public policy actions made
31 at the state-level, as well as technical validation of the primary data. Finally, the underlying
32 code for the site is also provided as open source, enabling others to validate and learn from this
33 work.

34

35 **Introduction:**

36

37 The disease known as COVID-19 was first reported in December of 2019 in Wuhan,
38 China¹. Three months later it was declared a pandemic by the WHO, and since then its death
39 toll has reached over 150,000 while infecting over 2 million people across 210 countries
40 worldwide². Additionally, the pandemic has disrupted the daily lives of billions and has incurred
41 significant socioeconomic costs at the global level.

42 In the US, the very assessment of the disease's impact has been challenged by
43 limitations in accurate data capture and analysis. Variable testing, uneven reporting, barriers to
44 data sharing, and a lack of easy-to-use analytic tools have all contributed to a lack of clarity in
45 establishing and trending the state of the pandemic. As a consequence, policy makers at all
46 levels have been forced to make decisions of great socioeconomic consequence in the face of
47 significant uncertainty.

48 To improve the accessibility of basic COVID-19-related information in the US, especially
49 by the general public and policymakers without a data science background, we report the
50 creation of a new interactive visualization tool that depicts daily disease trends at the level of
51 individual US counties. This web application features the novel reuse of several publicly
52 available sources of data while also introducing a new, manually curated dataset accompanying
53 this manuscript. This site features several unique views, including local doubling times and
54 estimated ICU bed requirements by county. Additionally, we report the technical validation of
55 the primary data (counts per county per day) against other official- and commonly used sources
56 of data.

57

58 **Methods:**

59

60 Data sources: Data on state-wide and county-level counts were obtained from The New York
61 Times³ via their *github* repository (<https://github.com/nytimes/covid-19-data>). County-wise
62 population data were obtained from the US Census⁴ using the *R* package *tidycensus*⁵. Data on
63 ICU bed availability per county was obtained from Kaiser Health News⁶.

64

65 As per The New York Times, cases and deaths reported from New York, Kings, Queens, Bronx
66 and Richmond counties were assigned to New York City. Similarly, Cass, Clay, Jackson and Platte
67 counties in Missouri were assigned to Kansas City. When a patient's county of residence was
68 unknown or pending many state departments reported these cases as coming from "unknown"
69 counties. Cases reported from unknown counties were only included at the state level.

70

71 Data related to state-wide implementation of social-distancing policies were manually curated
72 by web search and independently reviewed by a second author; disagreements were rare and
73 resolved by discussion. Government websites were prioritized as sources of truth where
74 feasible; otherwise, news reports covering state-wide proclamations were used. All citations
75 are captured in the open data file accompanying this manuscript.

76 [<https://datadryad.org/stash/share/whGecW9DWYmoAVMDdAHNF0z712Vbxrj9YwI5QKRAWUs>
77]. These data were up to date and confirmed as of the date of data deposit: April 19, 2020.

78

79 Ground truth data used for validation were manually curated from the websites of multiple
80 state departments of public health as well as Corona Data Scraper

81 [<https://coronadatascraper.com/>], a commonly used resource for aggregating county-level
82 tracking of COVID-19 over time. Citations of the validation data are included in the data file
83 accompanying this manuscript.

84 [<https://datadryad.org/stash/share/whGecW9DWYmoAVMDdAHNF0z712Vbxrj9YwI5QKRAWUs>
85]

86

87 Descriptive statistics on all datasets except that of the US Census and validation data are
88 reported in Table 1.

89

90 Doubling Time: Doubling time was calculated for each state and county by taking the reciprocal
91 of difference between the log (base 2) case counts corresponding to adjacent days, then
92 applying the *R* function *loess* for smoothing. The input of this model required a minimum of 8

93 days of data where the minimum number of cases was greater than 10. Regularization was
94 performed by replacing extreme doubling times (>500 days) with the average of the
95 surrounding values.

96

97 ICU Bed Occupancy Model: We incorporated parameters related to rates of hospitalization and
98 ICU admission from work previously published by Ferguson et al.⁷. Although simpler than other
99 models, it fit publicly available county-level ICU bed data in California well and was easier to
100 understand for the user than more complicated models proposed⁸⁻¹¹. This model assumed a
101 4.4% rate of hospitalization among all new cases, a 30% rate of intensive care unit admission
102 among hospitalized patients, and a 9-day average length of stay (time until discharge or death).

103

104 Web Application Development and Deployment: See Figure 1 for an overall schematic of the
105 web application. The source code was written in *R* (4.1.0)¹² using the *shiny*¹³, *shinyjs*¹⁴,
106 *tidyverse*¹⁵ and *plotly*¹⁶ packages. Software version control was achieved using Docker. The
107 entire software code for the site is publicly available on *github*
108 (<https://github.com/vivical/ButteLabCOVID>) and *dockerhub*
109 (https://hub.docker.com/r/pupster90/covid_tracker). The web hosting was organized as a
110 unified data share between all instances running *R shiny* code and controlled by a load balancer
111 using an auto-scaling mechanism. The web environment is hosted by Amazon Web Services and
112 is located at covidcounties.org.

113

114 **Results:**

115

116 CovidCounties derives a majority of its data from The New York Times Coronavirus github page
117 [<https://github.com/nytimes/covid-19-data>] which is updated daily with cases and deaths
118 reported in each state and county from the previous day. This time series dataset was derived
119 from a variety of governmental sources. However, to our knowledge this data has never been
120 formally validated against other reputed sources of COVID-19 reporting including state and
121 local departments of public health.

122

123 First, we demonstrate the high concordance of cumulative cases and deaths calculated and
124 displayed in CovidCounties at the county level by directly comparing these to numbers reported
125 by the Departments of Public Health in California and Connecticut (Figure 2A, 2B). These two
126 states were chosen because they both publicly report the daily counts of cases requiring
127 hospitalization or intensive care at the county level. R^2 rates corresponding to the concordance
128 between predicted and actual counts ranged from 0.86 to 1. To our knowledge, California is
129 only state in the US to report county-wide ICU bed utilization rates. We found a high degree of
130 concordance ($R^2 = 0.87$) with minimal model bias (Figure 2A), indicating a fairly high degree of
131 explained variation despite a relatively simplistic model.

132

133 An R^2 of 1 was specifically found with respect to cumulative cases and deaths in Connecticut
134 (Figure 2B), suggesting a shared common data source.

135

136 We compared the concordance of our data with that reported by Corona Data Scraper
137 [<https://coronadatascraper.com/>], another widely used source of aggregated publicly-available
138 COVID-19 timeseries data at the county level. We found very high concordance ($R^2 = 0.95-0.97$)
139 for deaths and cases respectively with no model bias (Figure 2C).

140

141 Lastly, we compared the concordance of our predicted hospitalizations, cases, and deaths from
142 our dataset against data reported by 8 different State Departments of Public Health (Figure 2D).
143 We noted some relative over-estimation of our estimated hospitalized counts against that
144 actual data reported by multiple State Departments of Public Health. Nonetheless concordance
145 was very high ($R^2 = 0.97-0.99$).

146

147 Descriptive statistics on the New York Times data are provided in Table 1. 23 states reported
148 cases with unknown counties of residence, however, in all states except Rhode Island these
149 cases made up less than 4% of the total cases in that state (Table 1). The inability to map these
150 cases to specific counties may explain some of the discrepancies between the New York Times

151 data used in CovidCounties and the curated data from state public health departments and the
152 Corona Data Scraper.

153

154 The data and tools incorporated into CovidCounties support the effectiveness of social
155 distancing measures, consistent with several events that have occurred following the initial
156 release of the website. South Dakota, one of six states which does not have a statewide shelter
157 in place order (as of April 15, 2020), has experienced rapid case growth following an exposure
158 at a meat plant (Figure 3A). This has accounted for more than half of the state's cases¹⁷ as of
159 April 15, 2020, with the fastest statewide doubling time of 4.5 days (Figure 3B). By contrast,
160 states with early shelter in place times like Arizona on March 11, 2020, California on March 19,
161 2020, and Connecticut on March 20, 2020 (Figure 3A) have much slower doubling times of 19.3
162 days, 22.8 days, and 12.7 days respectively (Figure 3B).

163

164 The web application located at covidcounties.org was first released to the public on April 3,
165 2020. It features two sections: a line plot depicting time-series trends in disease dynamics, and
166 a map depicting geospatial relationships (Figure 4). The site has had over 15 thousand unique
167 site in the first week as of April 11, 2020, most of whom accessed the website using a mobile
168 device.

169

170 **Discussion:**

171

172 The effective management of the COVID-19 pandemic has been hindered by both inaccurate
173 data collection and reporting, as well as relative inaccessibility by non-data scientists. Taken
174 together, these difficulties have impeded optimal policymaking by both government (imposing
175 social distancing policies) and health systems (anticipating ICU utilization) alike. Consequently,
176 responses across institutions have been highly variable and with varying degrees of success. To
177 help address these gaps we developed covidcounties.org and performed the technical
178 validation reported in this work.

179

180 The curation of COVID-19 case and death counts by The New York Times is an impressive effort
181 by over 60 reporters to collect, curate and analyze a constantly growing and evolving dataset³.
182 However, they acknowledge that the underlying data is extremely fragmented and comes from
183 thousands of different sources at both the state and county levels and thus is inherently limited
184 by accuracy, consistency, and timeliness. The New York Times notes that reported cases have
185 been corrected mere hours after the initial report and there have been numerous instances
186 where data has disappeared from databases without explanation. The New York Times has also
187 chosen to count patients where they were treated rather than their place of residence and
188 report on a number of geographic exceptions in their dataset
189 (<https://github.com/nytimes/covid-19-data>) including the treatment of cities like New York City
190 and Kansas City and the allocation of cases from cruise ships. Further, there are a subset of
191 cases where the patient's county of residence cannot or has not yet been identified which is
192 generally a small fraction of a state's total cases but can be a significant number in a small state
193 like Rhode Island (Table 1).

194

195 Taken together, these subtleties of the data collection process imply that the COVID-19 data
196 from The New York Times may not exactly agree with the numbers reported by various state
197 and county Departments of Public Health. We quantified the consistencies between The New
198 York Times COVID-19 data and county (Figure 2A, 2B) and state (Figure 2D) Department of
199 Public Health data and found the datasets to be largely comparable. Based on the exact
200 agreement, it seems likely that The New York Times is deriving their data for Connecticut
201 directly from the Connecticut Department of Public Health (Figure 2B).

202

203 The comparison of our estimated hospitalized cases based on the simple model from Ferguson
204 et al.⁷ with state (Figure 2C) and county (Figure 2B) reported hospitalizations revealed a
205 systematic bias towards increased hospitalizations in our model. We suspect that this bias is
206 due to a number of factors including time lags between the date of hospitalization and the
207 results of testing, as well as miscalibration of the assumed 4.4% rate of hospitalization taken
208 from the Ferguson model^{7,8,11}.

209
210 With the advent of the COVID-19 pandemic we have observed a trend towards government
211 agencies at the municipal, county, state, and national levels making their data increasingly
212 accessible for re-use and therefore provide potential value. However, many of the most popular
213 tools which are built upon this freely available data do not provide their source code for further
214 development. The Johns Hopkins dashboard², which receives more than 1.2 billion hits per day,
215 has made their data publicly available¹⁸, however, the source code for their dashboard is not
216 made available for further development by third parties. Similarly, the IHME dashboard¹⁹ which
217 has been referenced by the White House for making policy decisions²⁰ has had their dashboard
218 peer reviewed²¹, however, their epidemiological model has yet to be peer reviewed⁹. While
219 IHME provides open source code on their data aggregation process
220 (<https://github.com/beoutbreakprepared/nCoV2019>) and some features of their model
221 including the curve fitting of their projections (<https://github.com/ihmeuw-msca/CurveFit>), the
222 whole dashboard is not open source. Additionally, many states and counties are using *Tableau*,
223 a proprietary piece of software, to visualize COVID-19²² and as of 4/17/2020 there are 1,184
224 coronavirus dashboards on Tableau public²³. While Tableau facilitates powerful data
225 visualization, the software is not open source and requires a license for use. To promote further
226 development of CovidCounties and fully leverage the available data we have implemented our
227 website using the commonly used *R* and *Rshiny* frameworks, and made all of our source code
228 freely available on github (<https://github.com/vivical/ButteLabCOVID>).

229
230 CovidCounties represents an improvement over existing dashboards in terms of both scope and
231 granularity. Existing COVID-19 dashboards generally focus either on county level data within a
232 particular state (primarily at a static timepoint) or at the state level across the United States.
233 We have developed an intuitive tool that facilitates temporal comparisons between all counties
234 in the US. However, we are inherently limited by the availability of data. While CovidCounties'
235 estimation of ICU needs at the county level allows for higher resolution allocation of resources
236 compared to the widely used state level model from IHME
237 (<https://covid19.healthdata.org/united-states-of-america>), zip code level data would further

238 improve the value of these estimations for resource allocation. States like Maryland²⁴,
239 Arizona²⁵, and South Carolina²⁶ and counties like Johnson County, Kansas²⁷, San Diego County,
240 California²⁸, and King County, Washington²⁹ have already made zip code level data available.
241 However, there are many states and counties that are hesitant to provide data of this
242 granularity due to concerns over privacy thus highlighting the challenge of balancing privacy
243 with public good.

244

245 A limitation of CovidCounties is the inherent dependence on publicly available data. To date,
246 most states and counties are primarily providing case and death data with an increasing
247 number also providing hospitalization data. However, there is a severe lack of testing
248 information. The lack of testing data limits the ability to make inferences on the infection rate
249 in the population and the improvement of model trajectories. It has also been proposed that
250 there has been an under ascertainment of cases especially in the asymptomatic³⁰, which can
251 influence case rates. States and counties are continuously ramping up testing and this sudden
252 availability of tests can artificially distort counts by attributing individuals who were infected
253 previously to a later date due to an earlier shortage of tests. These numbers are further
254 complicated by the wide variety of commercially available tests that rely on different
255 technologies with varying sensitivity and specificity.

256

257 With its release, covidcounties.org represents a powerful open-source platform to empower
258 non-data scientists to track the current trends of the COVID-19 pandemic at the county level to
259 help facilitate policy and healthcare decisions which can help improve outcomes. We welcome
260 volunteers (both technical and non-technical) to help us to further develop CovidCounties
261 (<https://covidcounties.org/buttelabcovid/www/volunteers.html>).

262

263 **Usage Notes:**

264 A summary of the website features is available from the University of California, San Francisco
265 [<https://ucsf.app.box.com/v/Covid19Townhall041720>]. A detailed tutorial illustrating use of the
266 website is available on youtube.com (<https://youtu.be/5OHDSplV1kY>).

267

268 **Code Availability:**

269 The website source code is available on github (<https://github.com/vivical/ButteLabCOVID>). A
270 version-controlled *Docker* [ref] container is also available on dockerhub
271 (https://hub.docker.com/r/pupster90/covid_tracker).

272

273 **Data Availability:**

274 Curated data on the state-wide implementation of social-distancing policies and curated
275 validation data are hosted on datadryad.com
276 (<https://datadryad.org/stash/share/whGecW9DWYmoAVMDdAHNF0z712Vbxrj9YwI5QKRAWUs>
277).

278

279 **Contributions:**

280 PB and AJB jointly conceived the initial concept. DA, PB, ME, BO, and RV participated in data
281 acquisition, software development and website deployment. AM, VAR, and TZ participated in
282 manual collection and curation of the state governmental policy data. DA, AM, VAR, and TZ
283 contributed to the initial draft of the manuscript. All authors performed editing for important
284 intellectual content. PB, AJB, BO, and VAR performed project management and overall
285 supervision.

286

287 **Acknowledgements:** The authors kindly acknowledge Amazon Web Services for donating the
288 necessary computing resources for website hosting.

289

290 **Funding support:** Research reported in this publication was supported by funding from the
291 UCSF Bakar Computational Health Sciences Institute and the National Center for Advancing
292 Translational Sciences of the National Institutes of Health under award number UL1 TR001872.
293 VAR was supported by the National Center for Advancing Translational Sciences, National
294 Institutes of Health, through of the National Institutes of Health grant under award number TL1
295 TR001871. AJB was supported in part by the National Institute of Allergy and Infectious

296 Diseases (Bioinformatics Support Contract HHSN316201200036W). The content is solely the
297 responsibility of the authors and does not necessarily represent the official views of the
298 National Institutes of Health.

299

300 **Competing Interests:** The authors declare no relevant competing interests

301

302

303 **Figures/Table Legends:**

304

305 **Figure 1: Database schematic.**

306 Source data was obtained from The New York Times, US Census Bureau, Kaiser Health News,
307 and from a manual curation of state governmental websites and news outlets as described in
308 *Methods*. Data was processed to reflect case and death counts at the level of states and
309 counties. Functions were written to perform x- and y-axis rescaling, normalization by
310 population, doubling time estimation, and ICU bed utilization. Results were depicted using
311 interactive line plots and maps.

312

313 **Figure 2: Technical validation of the datasets.**

314 A. Comparison of estimated ICU bed occupancy, cumulative cases, and cumulative deaths
315 reported by CovidCounties against corresponding data reported by the California Department
316 of Public Health. Each point corresponds to a measurement from a given California county on a
317 particular date where both datasets report counts. Data is from 4/1/2020 - 4/9/2020.

318 B. Comparison of the estimated hospital bed occupancy, cumulative cases, and cumulative
319 deaths reported by CovidCounties against corresponding data reported by the Connecticut
320 Department of Public Health. Each point corresponds to a measurement from a given
321 Connecticut county on a particular date where both datasets report counts. Data is from
322 3/24/2020 - 4/9/2020.

323 C. Comparison of the estimated hospital bed occupancy, cumulative cases, and cumulative
324 deaths reported by CovidCounties against corresponding data reported by the website Corona
325 Data Scraper as of 4/10/2020 (includes data up to 4/9/2020). Each point corresponds to a
326 measurement from any US county in the dataset at a particular time where both datasets
327 report counts.

328 D. Comparison of the estimated hospital bed occupancy, cumulative cases, and cumulative
329 deaths reported by CovidCounties against corresponding data reported by 8 different state

330 Departments of Public Health. Data ranges vary by state; curated state data is available in the
331 data file accompanying this manuscript.

332

333 **Figure 3: Effect of shelter in place orders on doubling time**

334 A. States within the United States are color-coded by percentile of date to implement state
335 mandated shelter in place. White indicates earlier dates (among states) while dark orange
336 indicates later dates or no state mandate.

337 B. States within the United States are color-coded by percentile of case doubling time on April
338 15, 2020. Dark orange indicates a fast doubling time (among states), white indicates a slow
339 doubling time.

340

341 **Figure 4: Overview of CovidCounties.org.**

342 A. The primary view of CovidCounties.org is the line plot view, depicting time-series trends by
343 individual county. Depicted counties may be selected by single or double clicking the counties
344 displayed in the legend. They may also be selected by typing in counties (including from outside
345 of a given state) at the bottom.

346 B. User-selected individual states are color coded according to the variable of interest (e.g.
347 cumulative cases). Dark orange corresponds to the highest percentiles within the state, white
348 indicates the lowest percentile. Hovering functionality displays statistics corresponding to a
349 given county.

350 C. Line plot views can be extensively customized, with features to enable axis re-
351 centering/scaling, count normalization, depiction of doubling time, and predicted ICU bed
352 utilization. Individual state and United States plots update to reflect selected parameters where
353 appropriate.

354 D. States within the United States are color-coded by percentile according to the variable of
355 interest (e.g. cumulative cases). Dark orange indicates relatively high percentile (among states),
356 white indicates low percentile. Hovering functionality displays statistics corresponding to a
357 given state. The dropdown menu below allows the user to change the view to depict timing

358 that various social distancing policies were implemented: white indicates relatively early
359 adoption (by percentile), dark orange indicates late or no current adoption.

360 **Table 1: Descriptive statistics on the included data sets.**

New York Times

<i>% of counties with non-missing data†</i>	85.8%		
<i>States with greatest % of counties reporting</i>	17 counties tied at 100%*		
<i>States with lowest % of counties reporting</i>	Alaska (37.9% - 11/29)	Montana (50% - 28/56)	Nebraska (50.5% - 47/93)
<i>States with highest % of unknown cases</i>	Rhode Island (23.3%; 756; 715)	Connecticut (3.8%; 533; 149)	Arkansas (3.0%; 45; 15)
<i>Counties with highest cases per million</i>	Rockland, New York (25,591)	Westchester, New York (20,867)	Blaine, Idaho (20,265)
<i>Counties with the fastest doubling times</i>	Louisa, Iowa (1.13 days)	Walker, Texas (1.20 days)	Isle of Wight, Virginia (1.40 days)
<i>Counties with highest estimated ICU needs</i>	Rockland, New York (2,146% - 837/39)	Westchester, New York (1,166% - 2087/179)	Eagle, Colorado (985% - 49/5)

361

Kaiser Health News

<i>% of counties with non-missing ICU beds</i>	45.0%		
<i>Counties with most ICU beds</i>	Los Angeles, California (2,126)	Cook, Illinois (1,606)	New York City, New York (1,592)
<i>Counties with the most ICU beds per million</i>	Otero, Colorado (27,452)	Montour, Pennsylvania (2,303)	Emmet, Michigan (1,741)
<i>Counties with the least ICU beds per million</i>	Wright, Minnesota (22)	Clinton, Michigan (25.2)	Stafford, Virginia (26.7)

362

Policy Data

<i>% of states with non-missing data for all 4 policies</i>	60.8%		
<i>First to declare state of emergency</i>	Washington (2/29/2020)	California (3/4/2020)	Hawaii & Maryland (3/5/2020)
<i>First to close public schools</i>	Kentucky & Ohio (3/12/2020)	Delaware, Virginia & W Virginia (3/13/2020)	Arizona, Iowa, Nevada & NH (3/15/2020)
<i>First to declare shelter in place</i>	Arizona (3/11/2020)	California (3/19/2020)	Connecticut (3/20/2020)
<i>First to close restaurants and bars</i>	Ohio (3/15/2020)	12 states** on (3/16/2020)	9 states*** on (3/17/2020)

363

364 Data reported as of 4/16/2020. States with highest % of unknown cases shows the percent of
365 cases from unknown counties as a fraction of total cases in the state, the absolute number of
366 cases from unknown counties in the state, and the cases per million from unknown counties in
367 the state. States with lowest % of counties reporting shows the percentage of counties
368 reporting, the number of counties reporting, and total counties in the state. Counties with
369 highest estimated ICU needs shows the ICU needs as a percentage based on the estimated
370 number of ICU beds needed and KHN reported number of ICU beds.

371 † In the NY Times data all counties that reported cases also reported deaths (or were assumed
372 to be 0)

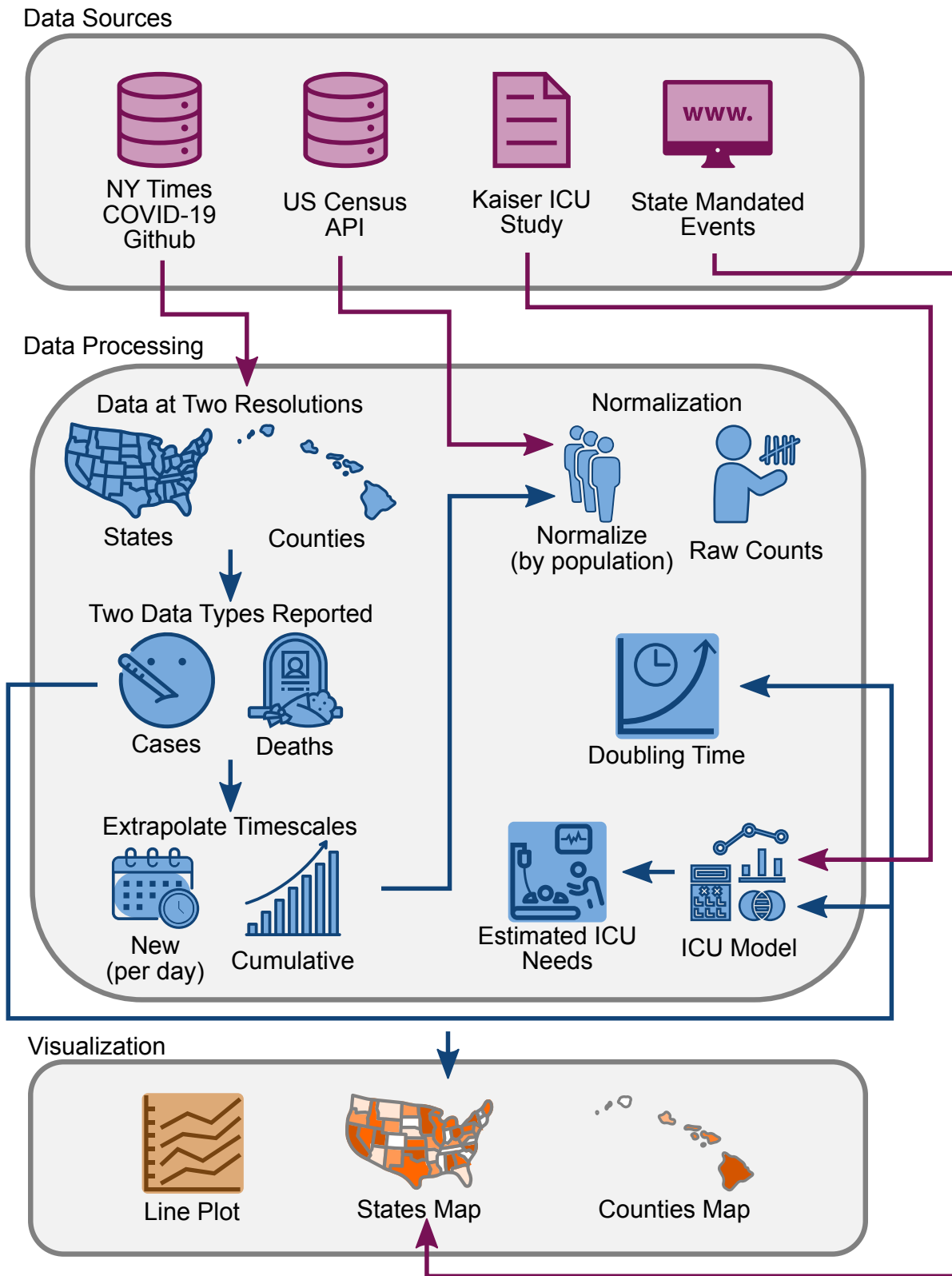
373 * AL, AZ, CT, DE, DC, FL, IN, LA, MD, MA, NH, NJ, NY, PA, RI, SC, VT

374 ** CA, CT, DE, DC, KY, LA, MI, NJ, NY, PA, RI, WA

375 *** CO, IL, IN, IA, MA, MN, NC, OR, VT

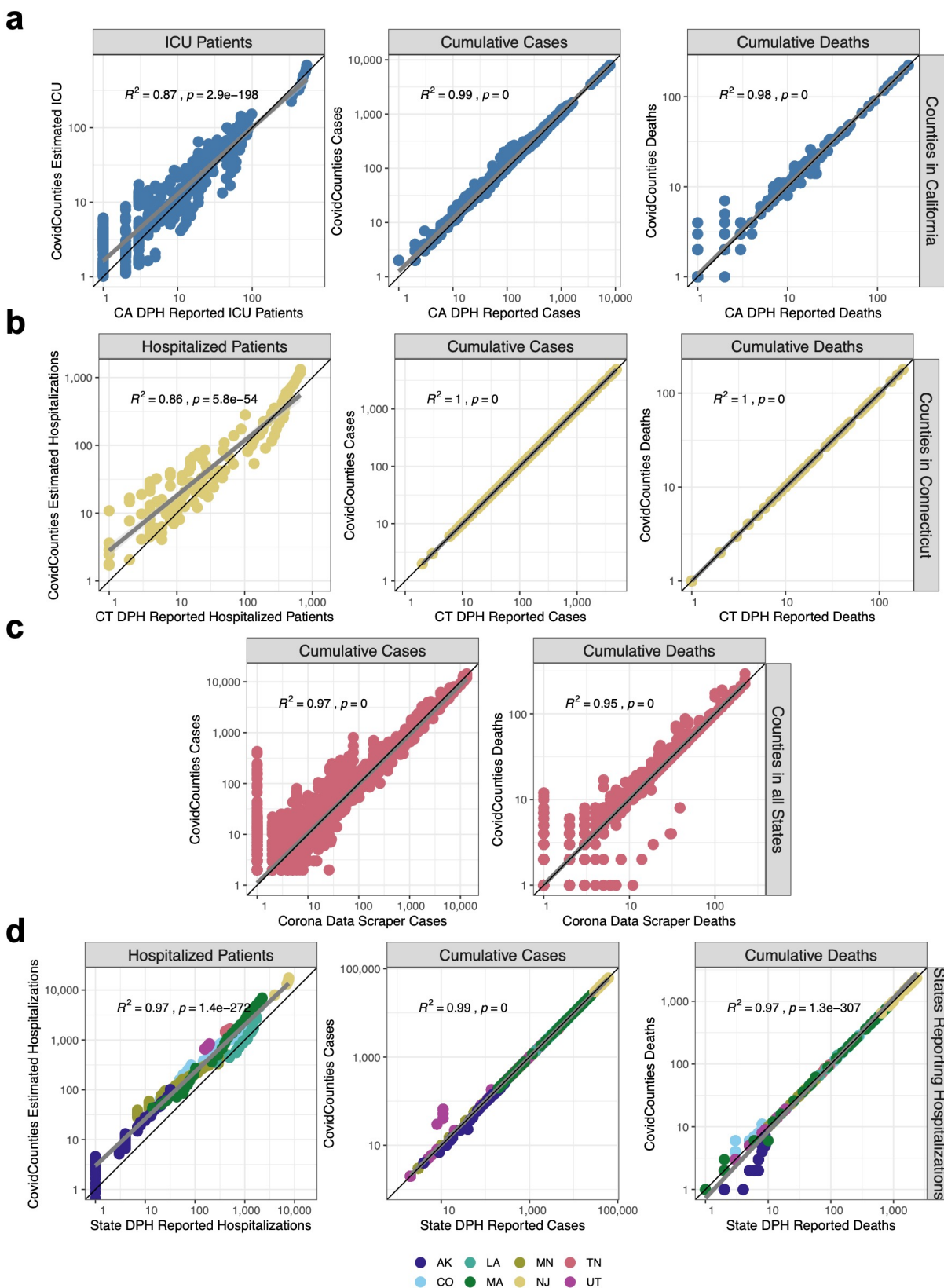
376

377 **Figure 1**



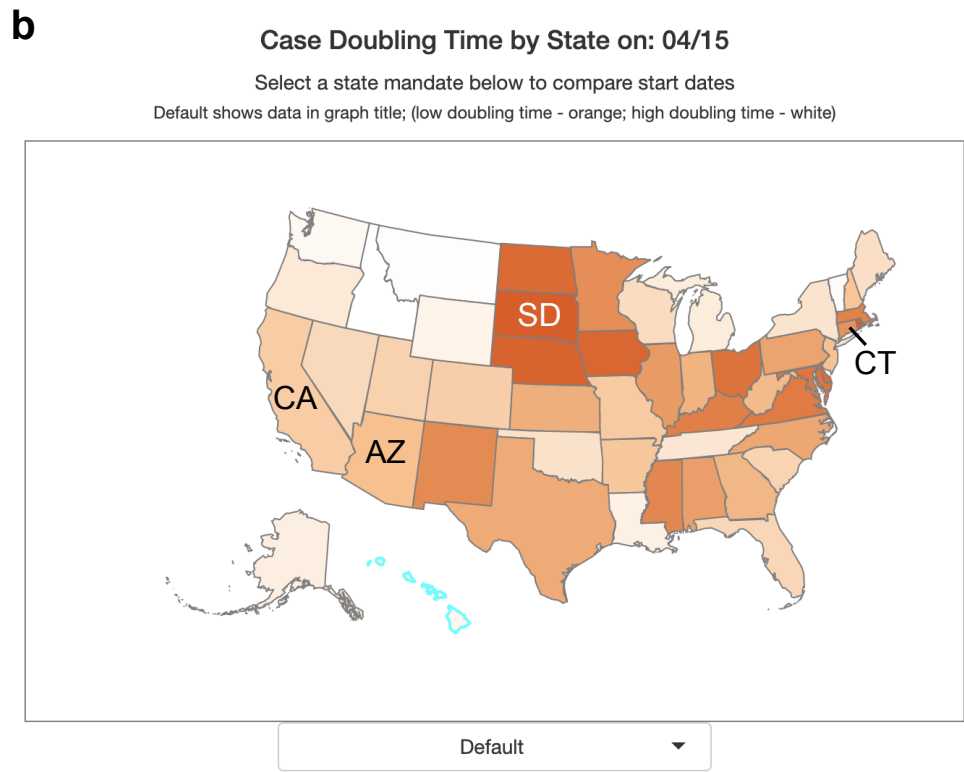
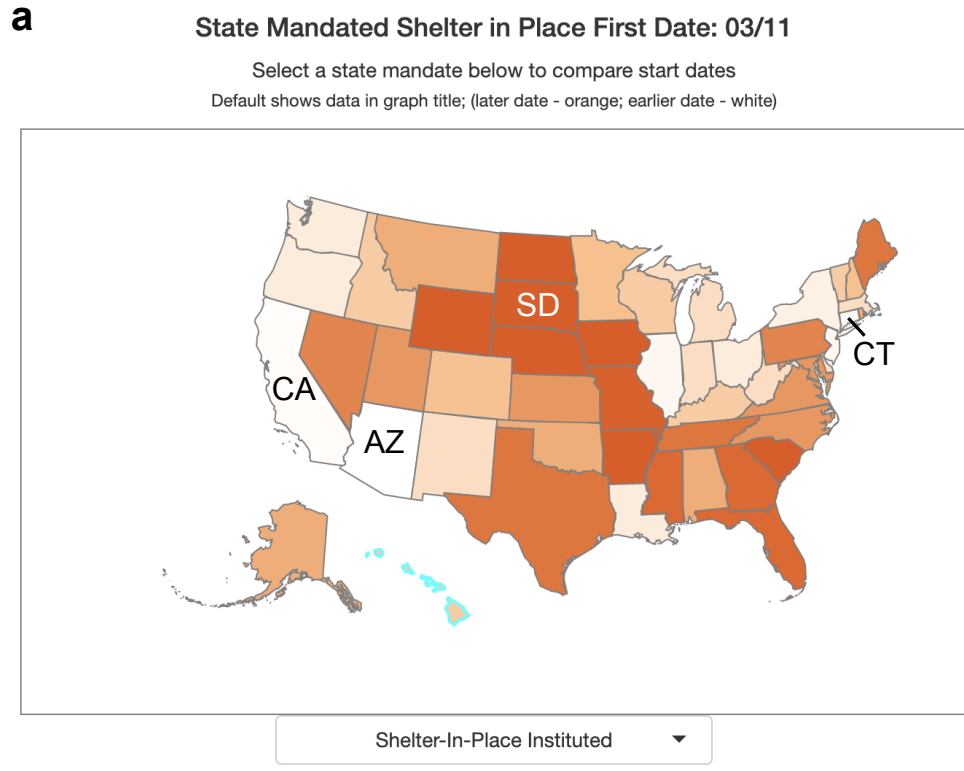
378

379 **Figure 2**

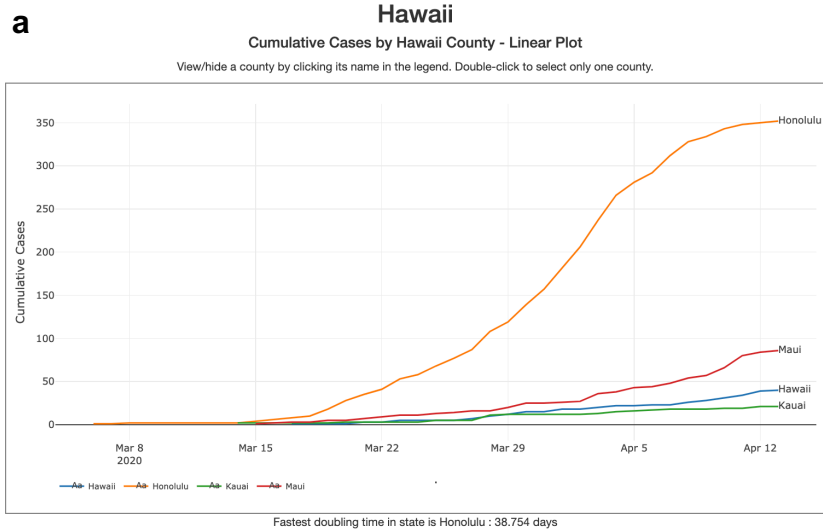


380

381 **Figure 3**



382



Select Counties:

Select a state:

c

Plot by:

- Total cases
- Total deaths
- New cases
- New deaths
- Doubling time
- Estimated ICU bed utilization

Time Scale

- Aligned (since first 10 cases)
- Actual dates

Scale by Population Density:

- True
- False

Y-Axis:

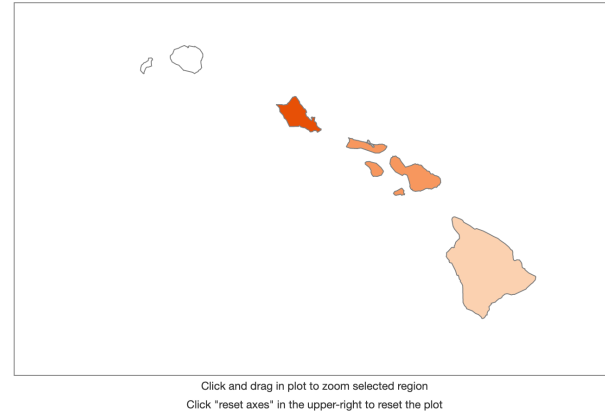
- Linear
- Log

Include counties with more cases than:

Date Range (some data streams begin as early as 1/21/2020): -

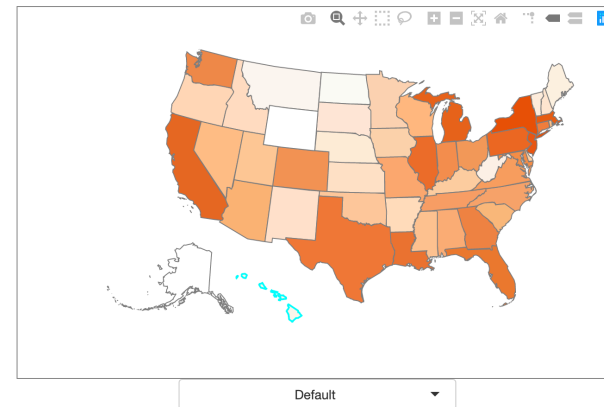
b Cumulative Cases by Hawaii County from: 03/01 - 04/13

Percentile ranks are relative to counties within the state
(high number of cases - orange; low number of cases - white)



d Cumulative Cases by State from: 03/01 - 04/13

Select a state mandate below to compare start dates
Default shows data in graph title; (high number of cases - orange; low number of cases - white)



385 **References:**

386

- 387 1. World Health Organization. Novel Coronavirus – China. [https://www.who.int/csr/don/12-](https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/)
388 [january-2020-novel-coronavirus-china/en/](https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/) (2020).
- 389 2. Center for Systems Engineering and Science at Johns Hopkins. COVID-19 Dashboard.
390 *Coronavirus Resource Center* <https://coronavirus.jhu.edu/map.html> (2020).
- 391 3. The New York Times. Coronavirus (Covid-19) Data in the United States.
392 <https://github.com/nytimes/covid-19-data> (2020).
- 393 4. The U.S. Census Bureau. U.S. Census Bureau’s 2014-2018 American Community Survey 5-year
394 estimates. <https://www.census.gov/data/developers/data-sets/acs-5year.html> (2020).
- 395 5. Walker, K. tidy census: Load US Census Boundary and Attribute Data as ‘tidyverse’ and ‘sf’-
396 Ready Data Frames. R package version 0.9.6. <https://CRAN.R-project.org/package=tidy census>
397 (2020).
- 398 6. Kaiser Health News. Millions Of Older Americans Live In Counties With No ICU Beds As
399 Pandemic Intensifies. [https://khn.org/news/as-coronavirus-spreads-widely-millions-of-older-](https://khn.org/news/as-coronavirus-spreads-widely-millions-of-older-americans-live-in-counties-with-no-icu-beds/)
400 [americans-live-in-counties-with-no-icu-beds/](https://khn.org/news/as-coronavirus-spreads-widely-millions-of-older-americans-live-in-counties-with-no-icu-beds/) (2020).
- 401 7. Ferguson, N. M. *et al.* Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce
402 COVID-19 mortality and healthcare demand. [https://www.imperial.ac.uk/media/imperial-](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf)
403 [college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf)
404 [2020.pdf](https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf) (2020).
- 405 8. Predictive Healthcare at Penn Medicine. COVID-19 Hospital Impact Model for Epidemics
406 (CHIME). <https://penn-chime.phl.io/> (2020).
- 407 9. IHME COVID-19 Health Service Utilization Forecasting Team. Forecasting COVID-19 impact on
408 hospital bed-days, ICU-days, ventilator-days and deaths by US state in the next 4 months.
409 *Medrxiv* 2020.03.27.20043752 (2020) doi:10.1101/2020.03.27.20043752.
- 410 10. Enns, E. A. *et al.* Modeling the Impact of Social Distancing Measures on the spread of SARS-
411 CoV-2 in Minnesota. https://mn.gov/covid19/assets/MNmodel_tech_doc_tcm1148-427724.pdf
412 (2020).
- 413 11. Biocomplexity Institute & Initiative University of Virginia. Estimation of COVID-19 Impact in
414 Virginia. [https://covid19.biocomplexity.virginia.edu/sites/covid19.biocomplexity/files/COVID-](https://covid19.biocomplexity.virginia.edu/sites/covid19.biocomplexity/files/COVID-19-UVA-MODEL-FINDINGS-Apr13-2020-FINAL%20PRESS.pdf)
415 [19-UVA-MODEL-FINDINGS-Apr13-2020-FINAL%20PRESS.pdf](https://covid19.biocomplexity.virginia.edu/sites/covid19.biocomplexity/files/COVID-19-UVA-MODEL-FINDINGS-Apr13-2020-FINAL%20PRESS.pdf) (2020).
- 416 12. R Core Team. R: A language and environment for statistical computing. R Foundation for
417 Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (2019).
- 418 13. Chang, W., Cheng, J., Allaire, J., Xie, Y. & McPherson, J. shiny: Web Application Framework
419 for R. R package version 1.4.0. <https://CRAN.R-project.org/package=shiny> (2019).
- 420 14. Attali, D. shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds. R
421 package version 1.1. <https://CRAN.R-project.org/package=shinyjs> (2020).
- 422 15. Wickham, H. *et al.* Welcome to the Tidyverse. *J Open Source Softw* **4**, 1686 (2019).
- 423 16. Sievert, C. plotly for R. <https://plotly-r.com> (2018).
- 424 17. AP News. South Dakota COVID cases top 1,100; meat plant worker dies.
425 <https://apnews.com/b20cd0c7c71b828eb164fbc069050aac> (2020).
- 426 18. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real
427 time. *Lancet Infect Dis* (2020) doi:10.1016/s1473-3099(20)30120-1.

- 428 19. IHME. COVID-19 Projections. <https://covid19.healthdata.org/united-states-of-america>
429 (2020).
- 430 20. Ripley, B. Our COVID-19 forecasting model, otherwise known as “the Chris Murray Model”.
431 [http://www.healthdata.org/acting-data/our-covid-19-forecasting-model-otherwise-known-](http://www.healthdata.org/acting-data/our-covid-19-forecasting-model-otherwise-known-chris-murray-model)
432 [chris-murray-model](http://www.healthdata.org/acting-data/our-covid-19-forecasting-model-otherwise-known-chris-murray-model) (2020).
- 433 21. Xu, B. *et al.* Epidemiological data from the COVID-19 outbreak, real-time case information.
434 *Sci Data* **7**, 106 (2020).
- 435 22. Anzilotti, E. How governments are using Tableau to keep you up-to-date on the coronavirus.
436 [https://www.tableau.com/about/blog/2020/4/how-governments-are-using-tableau-keep-you-](https://www.tableau.com/about/blog/2020/4/how-governments-are-using-tableau-keep-you-date-coronavirus)
437 [date-coronavirus](https://www.tableau.com/about/blog/2020/4/how-governments-are-using-tableau-keep-you-date-coronavirus) (2020).
- 438 23. Tableau Public. <https://public.tableau.com/search/all/%23coronavirus> (2020).
- 439 24. Maryland Department of Public Health. Coronavirus Disease 2019 (COVID-19) Outbreak.
440 <https://coronavirus.maryland.gov/> (2020).
- 441 25. Arizona Department of Health Services. Confirmed Covid-19 Cases by Zip Code.
442 [https://adhs.gis.maps.arcgis.com/apps/opsdashboard/index.html#/84b7f701060641ca8bd9ea0](https://adhs.gis.maps.arcgis.com/apps/opsdashboard/index.html#/84b7f701060641ca8bd9ea0717790906)
443 [717790906](https://adhs.gis.maps.arcgis.com/apps/opsdashboard/index.html#/84b7f701060641ca8bd9ea0717790906) (2020).
- 444 26. South Carolina Department of Health and Environmental Control. SC Cases by County & ZIP
445 Code (COVID-19). [https://scdhec.gov/infectious-diseases/viruses/coronavirus-disease-2019-](https://scdhec.gov/infectious-diseases/viruses/coronavirus-disease-2019-covid-19/sc-cases-county-zip-code-covid-19)
446 [covid-19/sc-cases-county-zip-code-covid-19](https://scdhec.gov/infectious-diseases/viruses/coronavirus-disease-2019-covid-19/sc-cases-county-zip-code-covid-19) (2020).
- 447 27. Johnson County Kansas AIMS. Johnson County, KS - COVID-19 Update.
448 [https://public.tableau.com/profile/mapper.of.the.day.mod.#/vizhome/covid19_joco_public/D](https://public.tableau.com/profile/mapper.of.the.day.mod.#/vizhome/covid19_joco_public/Dashboard)
449 [ashboard](https://public.tableau.com/profile/mapper.of.the.day.mod.#/vizhome/covid19_joco_public/Dashboard) (2020).
- 450 28. County of San Diego Health and Human Services Agency. County of San Diego Daily
451 Coronavirus Disease 2019 (COVID-19) Summary of Cases by Zip Code of Residence.
452 [https://www.sandiegocounty.gov/content/dam/sdc/hhsa/programs/phs/Epidemiology/COVID-](https://www.sandiegocounty.gov/content/dam/sdc/hhsa/programs/phs/Epidemiology/COVID-19%20Summary%20of%20Cases%20by%20Zip%20Code.pdf)
453 [19%20Summary%20of%20Cases%20by%20Zip%20Code.pdf](https://www.sandiegocounty.gov/content/dam/sdc/hhsa/programs/phs/Epidemiology/COVID-19%20Summary%20of%20Cases%20by%20Zip%20Code.pdf) (2020).
- 454 29. Seattle and King County Public Health. King County COVID-19 outbreak summary.
455 [https://kingcounty.gov/depts/health/communicable-diseases/disease-control/novel-](https://kingcounty.gov/depts/health/communicable-diseases/disease-control/novel-coronavirus/data-dashboard.aspx)
456 [coronavirus/data-dashboard.aspx](https://kingcounty.gov/depts/health/communicable-diseases/disease-control/novel-coronavirus/data-dashboard.aspx) (2020).
- 457 30. Omori, R., Mizumoto, K. & Nishiura, H. Ascertainment rate of novel coronavirus disease
458 (COVID-19) in Japan. *Medrxiv* 2020.03.09.20033183 (2020) doi:10.1101/2020.03.09.20033183.