

Data Engineering Presentation Slides

*A User-Configurable COVID-19 Case and
Death Rate Dashboard for the U.S.*

George Pappy - 23 February 2022

Introduction

COVID-19 Dashboards

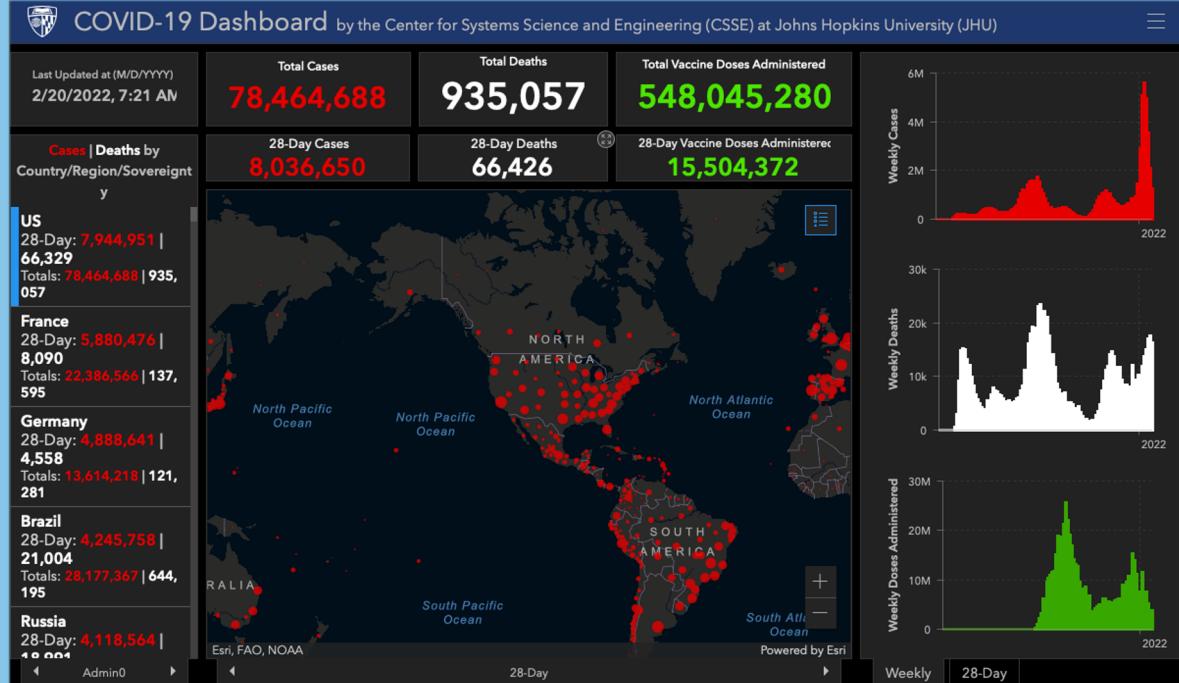
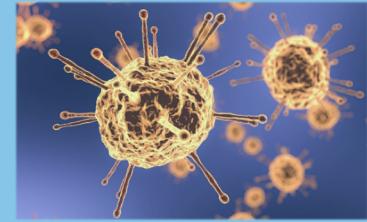
- Around since early in the pandemic



Introduction

COVID-19 Dashboards

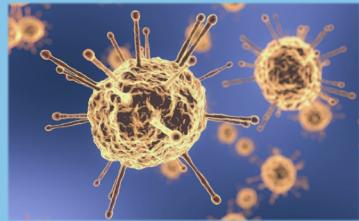
- Around since early in the pandemic
- Over time, they've gotten more complex



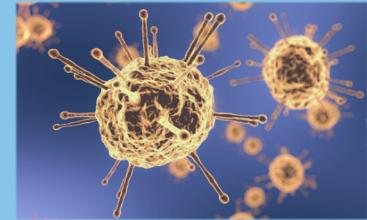
Introduction

COVID-19 Dashboards

- Around since early in the pandemic
- Over time, they've gotten more complex
- Too much information for the average user!



Introduction (con't.)

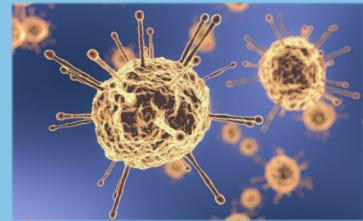


Goal: Develop a much simpler dashboard

- Provide only basic information for a user-specified U.S. state & county
- Avoid flooding the user with excessive data and visualizations



Dataset



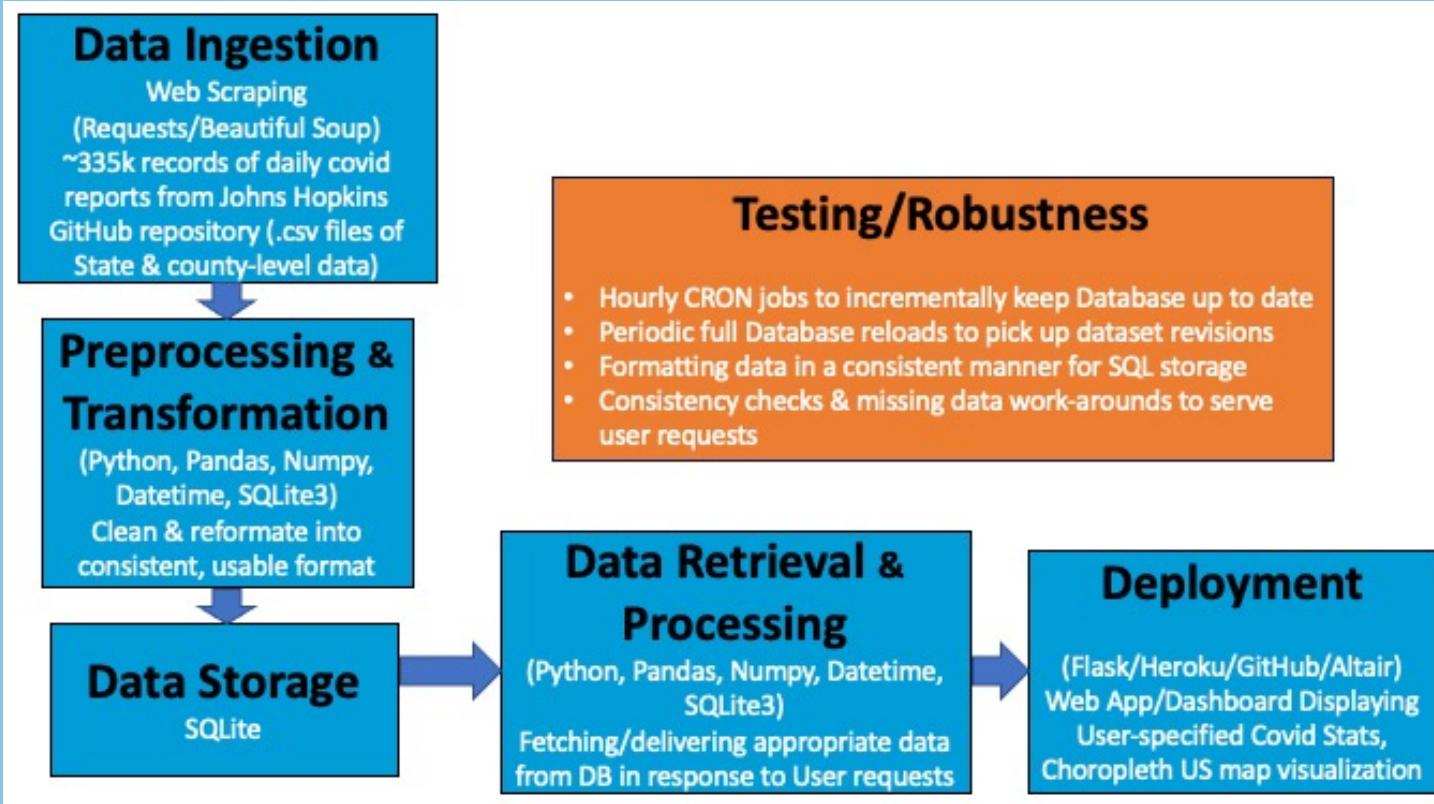
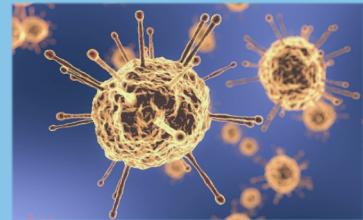
From Johns Hopkins Center for Systems Science & Engineering

- Available on GitHub and updated daily
- U.S. state & county-level cumulative daily COVID case data since April 2020

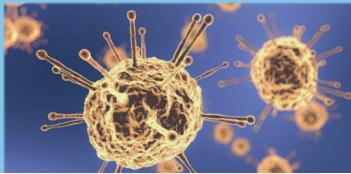
master COVID-19 / csse_covid_19_data / csse_covid_19_daily_reports_us / Go to file Add file

CSSEGISandData Automated update	4a818ce 12 hours ago	History
..		
01-01-2021.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago
01-01-2022.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago
01-02-2021.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago
01-02-2022.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago
01-03-2021.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago
01-03-2022.csv	patch california data from 04-12-2020 to 02-17-2022	2 days ago

Data Pipeline



Results: The Dashboard!



U.S. COVID-19 DASHBOARD

Data Provided by Johns Hopkins University & School of Medicine, Center for Systems Science and Engineering
 (UPDATED DAILY): https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data

Choose State Choose County

Reset Page

California Alameda

Submit

Time Window for Map

Since Pandemic Start
 Past 7 Days

Reported Map Statistic

Cases per 100k
 Total Deaths
 Deaths per 100k
 Death-to-Case Rate (%)

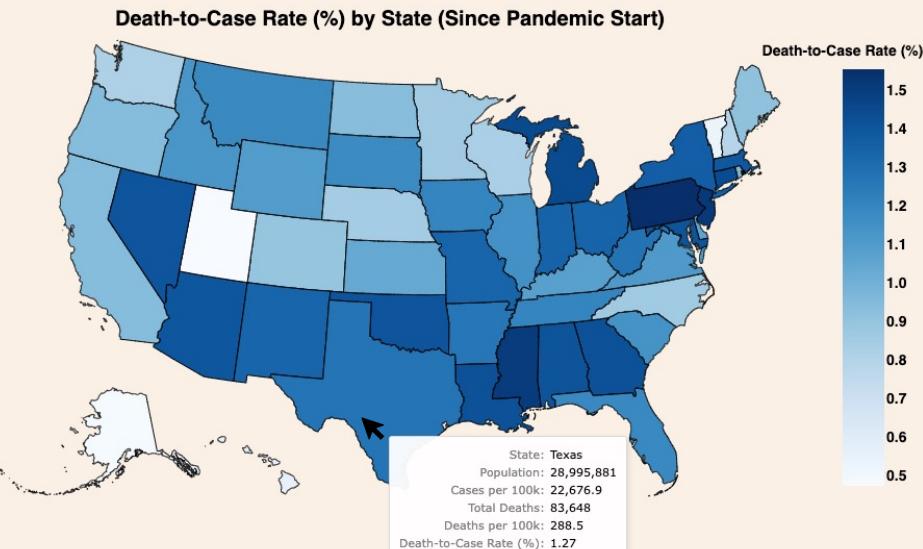
California (Since Pandemic Start)				California (Past 7 Days)			
Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)	Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)
22,503.7	83,946	212.5	0.94%	225.3	1,444	3.7	1.62%

Note: "per 100k" denotes the rate per 100,000 residents (state or county)

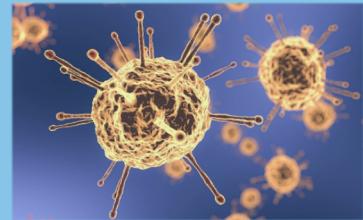
Los Angeles County (Since Pandemic Start)				Los Angeles County (Past 7 Days)			
Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)	Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)
27,617.7	30,146	300.3	1.09%	256.0	456	4.5	1.77%

U.S. FLU:	2014-15	2015-16	2016-17	2017-18	2018-19
Death-to-Case Rate:	0.17%	0.10%	0.13%	0.13%	0.10%

(Flu Data Source: <https://www.cdc.gov/flu/about/burden/index.html>)

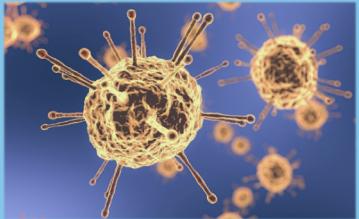


Conclusions



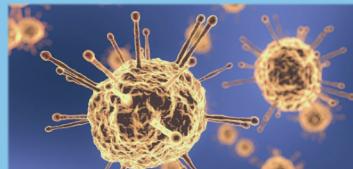
Recommendations/Future Work

- Look into more robust deployment options
 - Faster response times
 - Ability to serve many simultaneous users
- Add more counties (currently only 478 represented)
- Add state-level choropleth map option (with counties shown)



Appendix

Erroneous Data Robustness



covid-tracker-22.herokuapp.com

U.S. COVID-19 DASHBOARD

Data Provided by Johns Hopkins University & School of Medicine, Center for Systems Science and Engineering
(UPDATED DAILY): https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data

Choose State **Choose County**

Reset Page Washington Benton Submit

Time Window for Map **Reported Map Statistic**

Since Pandemic Start Cases per 100k
 Past 7 Days Total Deaths
 Deaths per 100k
 Death-to-Case Rate (%)

Washington (Since Pandemic Start)				BAD DATA (Daily Reporting Error) (Past 7 Days)			
Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)	Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)
18,506.5	11,615	152.5	0.82%	-	-	-	%

Note: "per 100k" denotes the rate per 100,000 residents (state or county)

Benton County (Since Pandemic Start)				BAD DATA (Daily Reporting Error) (Past 7 Days)			
Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)	Cases per 100k	Total Deaths	Deaths per 100k	Death-to-Case Rate (%)
24,858.8	441	215.8	0.87%	-	-	-	%

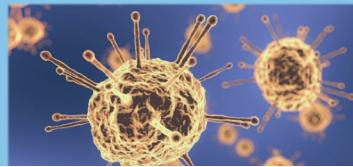
U.S. FLU:	2014-15	2015-16	2016-17	2017-18	2018-19
Death-to-Case Rate:	0.17%	0.10%	0.13%	0.13%	0.10%

Erroneous data replaced with user warning

Death-to-Case Rate (%)

1.5
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5

Erroneous Data Robustness (con't.)



← → ⌂ covid-tracker-22.herokuapp.com

U.S. COVID-19 DASHBOARD

Data Provided by Johns Hopkins University & School of Medicine, Center for Systems Science and Engineering
(UPDATED DAILY): https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data

