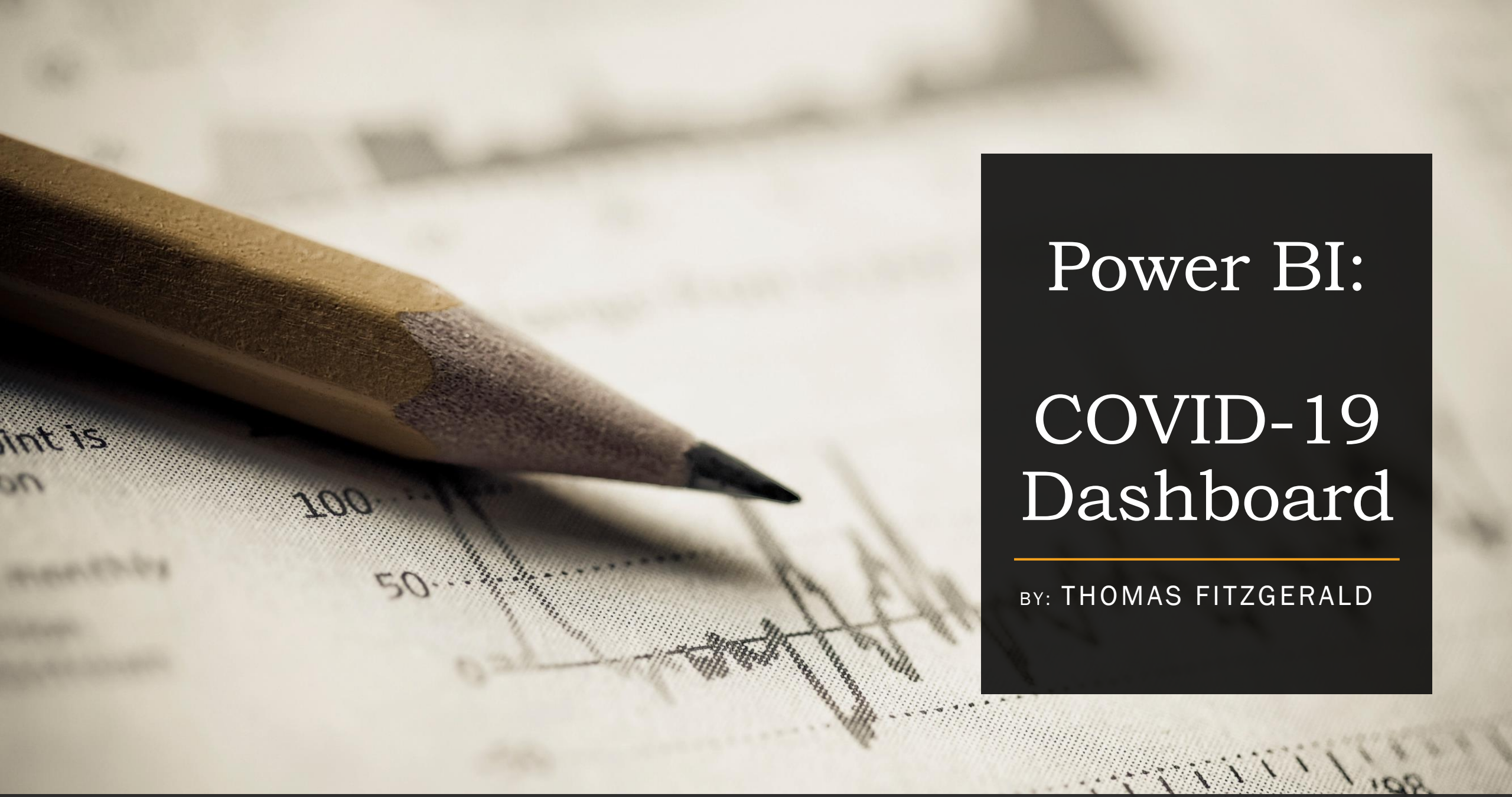


Power BI:

COVID-19 Dashboard

BY: THOMAS FITZGERALD

Table of Contents:

PowerPoint Overview Goal:

- Enable file sharing for this Power BI project
- Quick enough to compare John Hopkins & Tom's COVID-19 Dashboards
- Detailed enough to analyze beyond surface-level

- John Hopkins' Dashboard (has been updated several times):

  https://coronavirus.jhu.edu/map.html

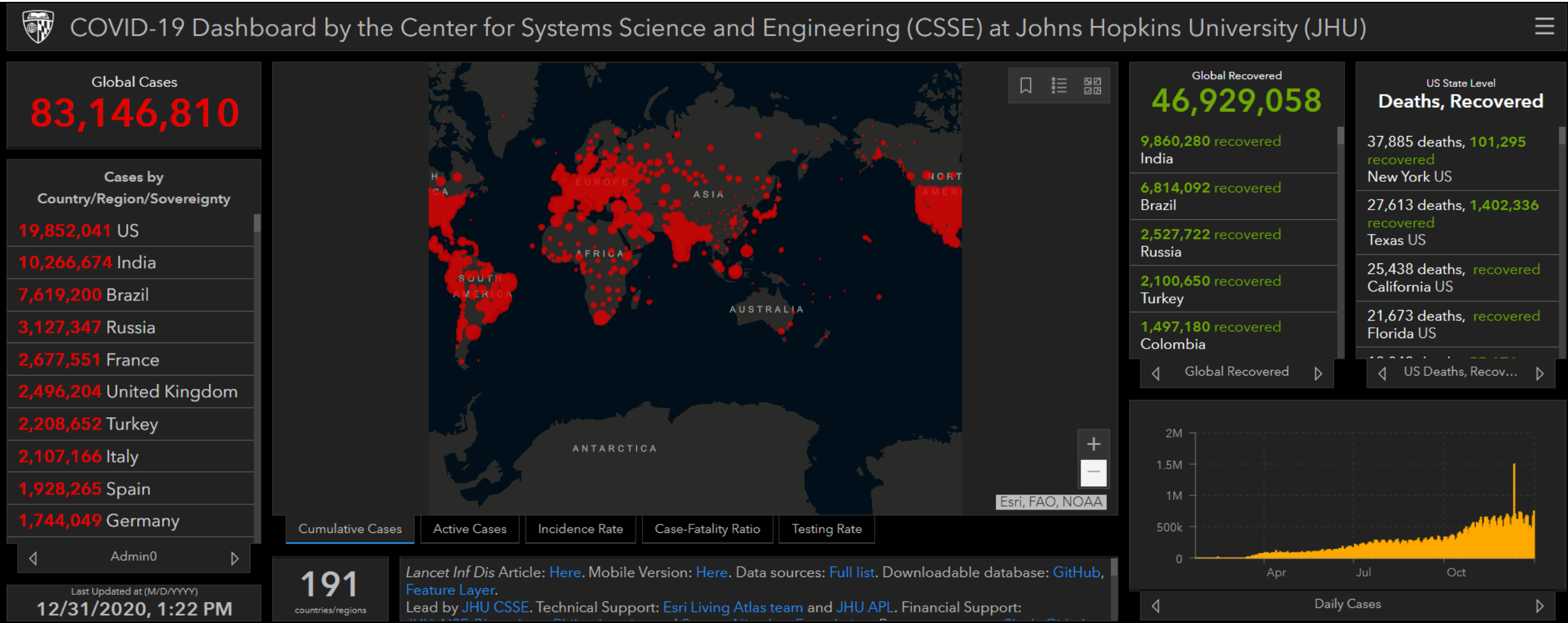
Motivation / Background Information

- Watched weekly stats of COVID-19 around April 2020 on worldometer.info/coronavirus (basic graphs & tables)
- Came across John Hopkins' dashboard & enjoyed the dynamic visualizations
- **Challenged myself to learn Power BI & recreate their dashboard**
- Fun project to analyze the global pandemic while the world is impacted by it

- Raw time series data provided on GitHub by the University
- https://github.com/CSSEGISandData/COVID-19

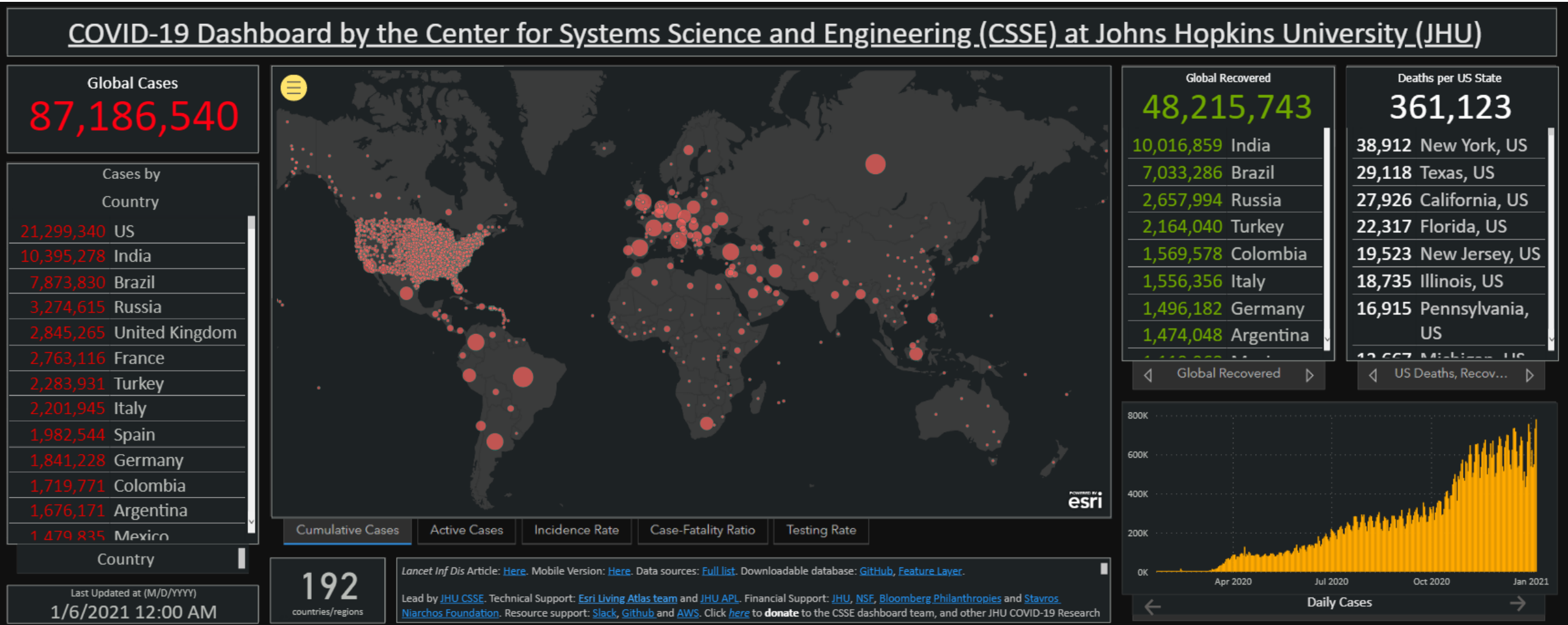
John Hopkins' Initial Dashboard

Screenshot Date: 12/31/2020, created by CSSE at John Hopkins



Tom Fitzgerald's Power BI Dashboard

Screenshot Date: 1/6/2021, created by Tom



T-Chart Dashboard Comparison

John Hopkins:

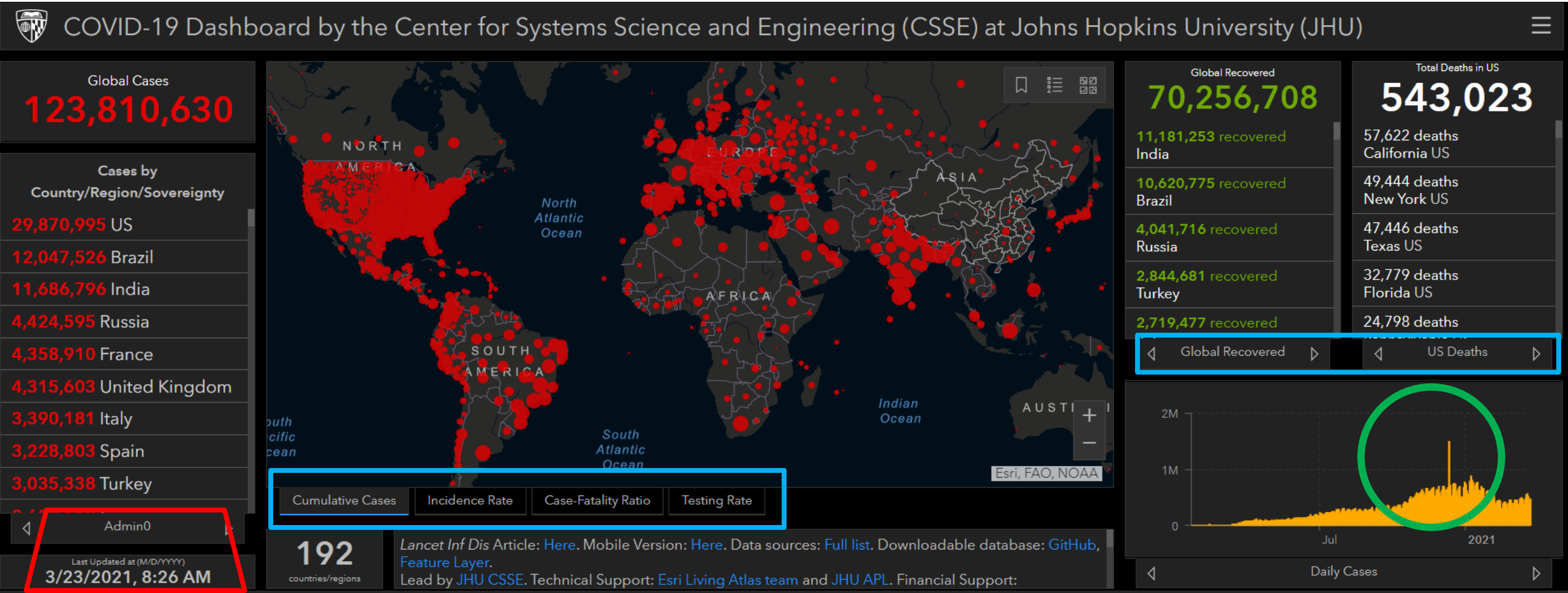
- **ArcGIS Dashboard (not free)**
- Includes additional data like testing rate (blue box, next slide)
- Known dirty data not cleaned (green circle)
- Stopped providing US recovery data during Winter 2020 Holiday
- Map data points = Red

Tom:

- **Power BI Dashboard** (Microsoft, free)
- Did not incorporate additional data / 85% completed (blue boxes)
- Cleaned daily cases data (green)
- Dashboard dependent on JHU data
- **ArcGIS has features Power BI may not have (i.e. Formatting, requires coding to manipulate the "look")**
- **Map data points = Pink & Red**

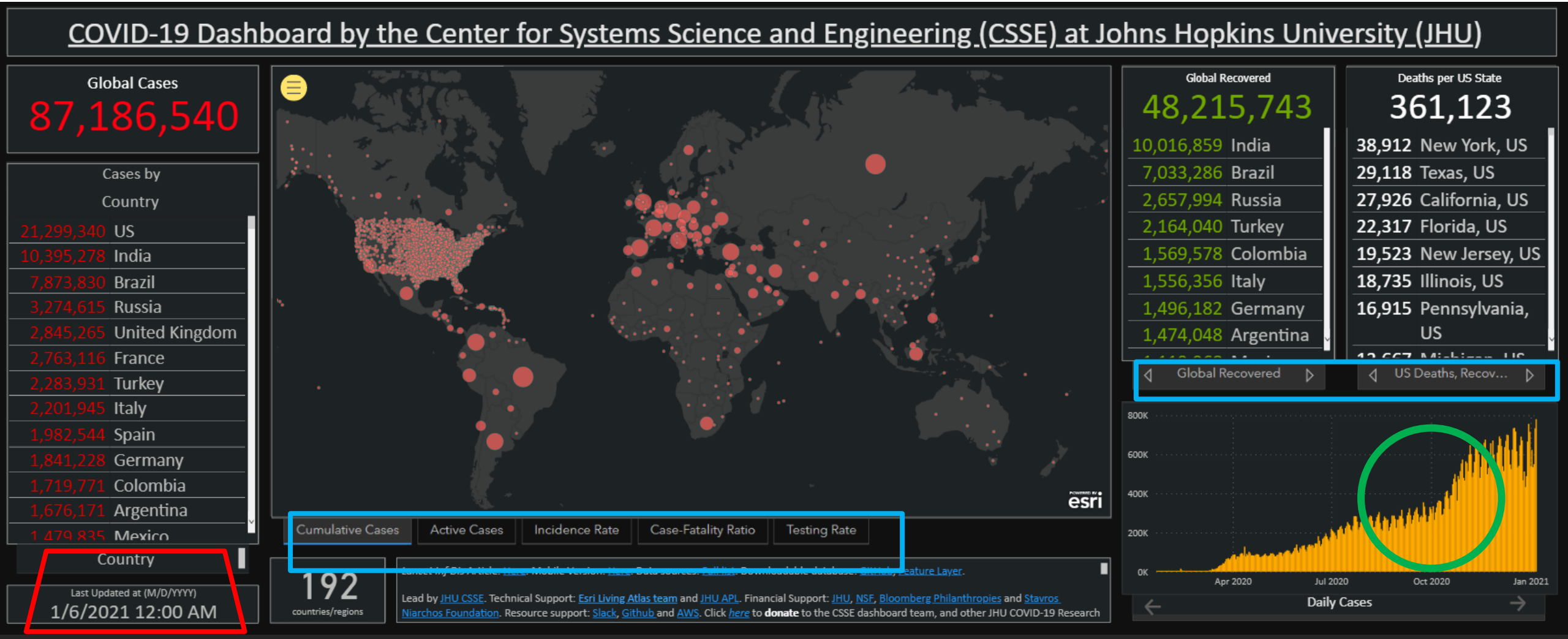
John Hopkins' Comparison:

1. ArcGIS Dashboard / Updated Data From March (Bottom Left)
2. Includes Incidence Rate, Case-Fatality Ratio, Testing Rate (Blue)
3. Daily Cases Data Not Cleaned (Green Circle on Bottom Right)



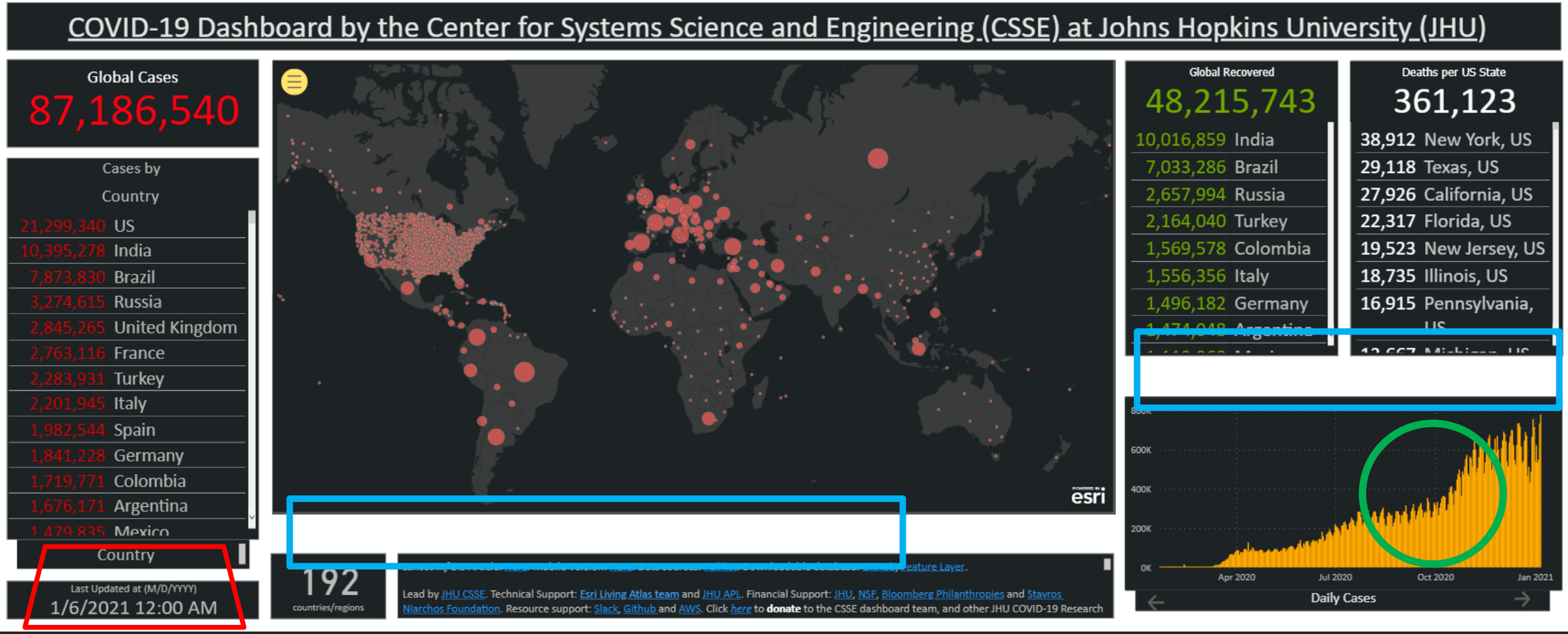
Tom's Comparison:

1. Power BI Dashboard (Jan 2021 Data)
2. Does not include additional data, "Bookmark Tabs" (Blue Rectangles)
3. Daily Cases Graph; Turkey Data Error Cleaned (Green Circle)



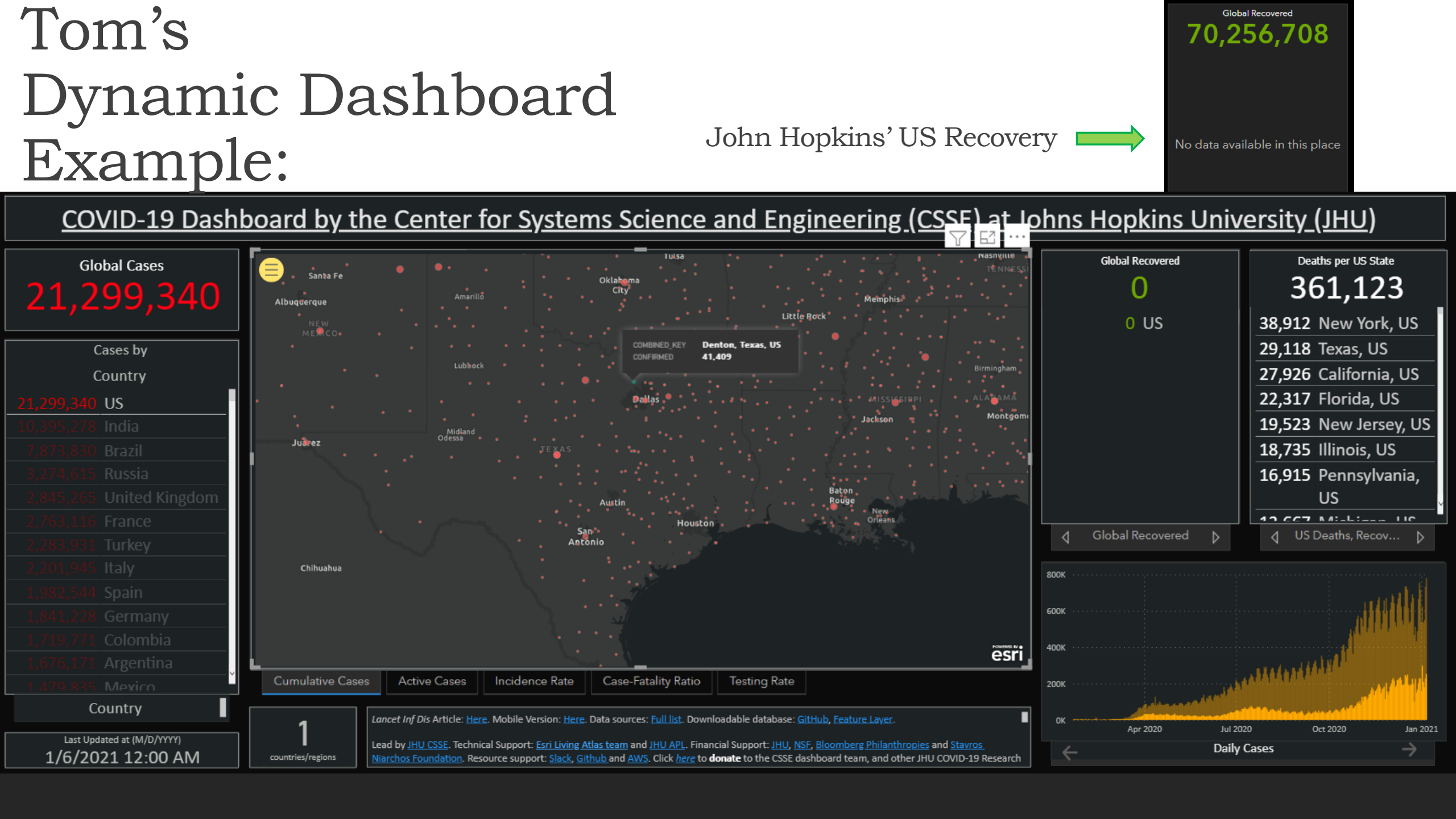
Tom's Dashboard – No Background

- Blue Rectangles show the "missing tabs"
- These tabs require Power BI "bookmarks" (i.e bottom right graph arrows to change the graph type)



Tom's Dynamic Dashboard Example:

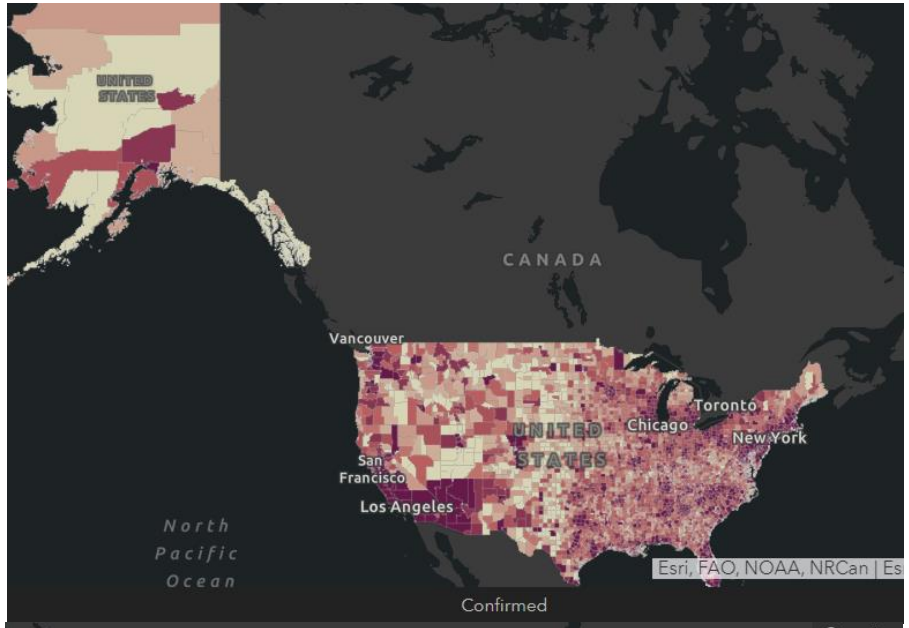
John Hopkins' US Recovery



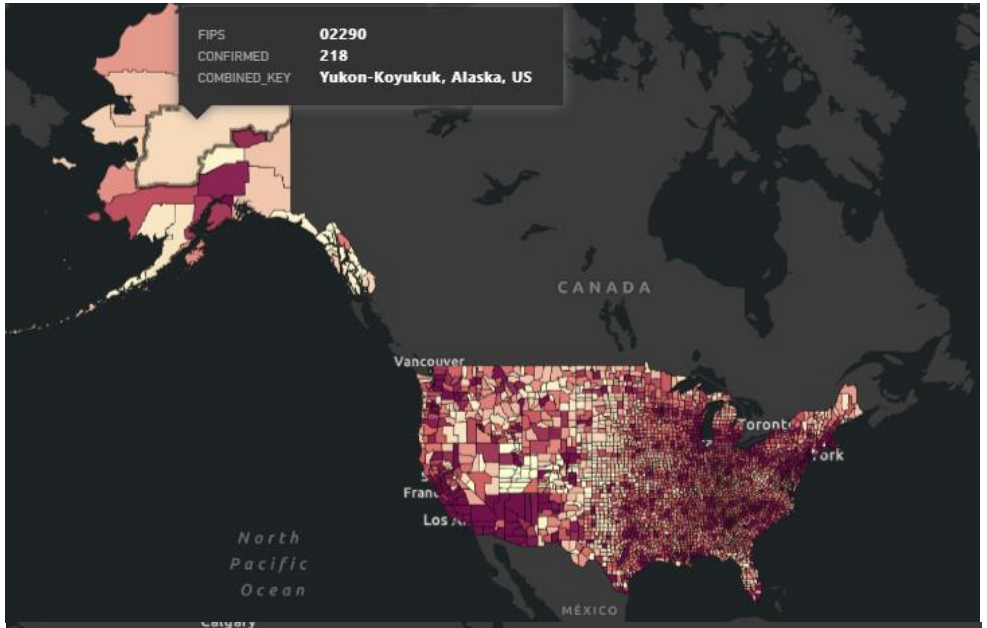
US Only Maps (By County/ Pop)

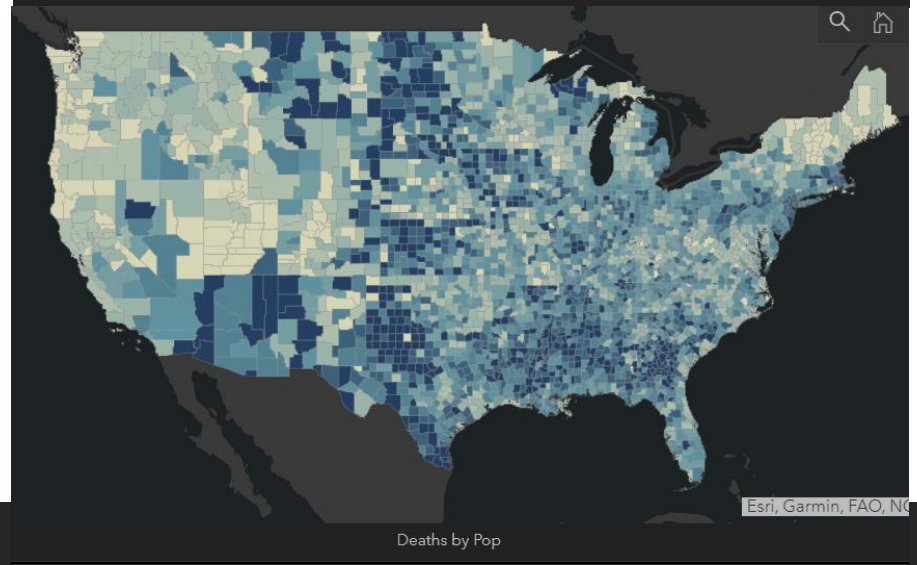
John Hopkins'
March 2021 Data

Tom's
Jan 2021 Data

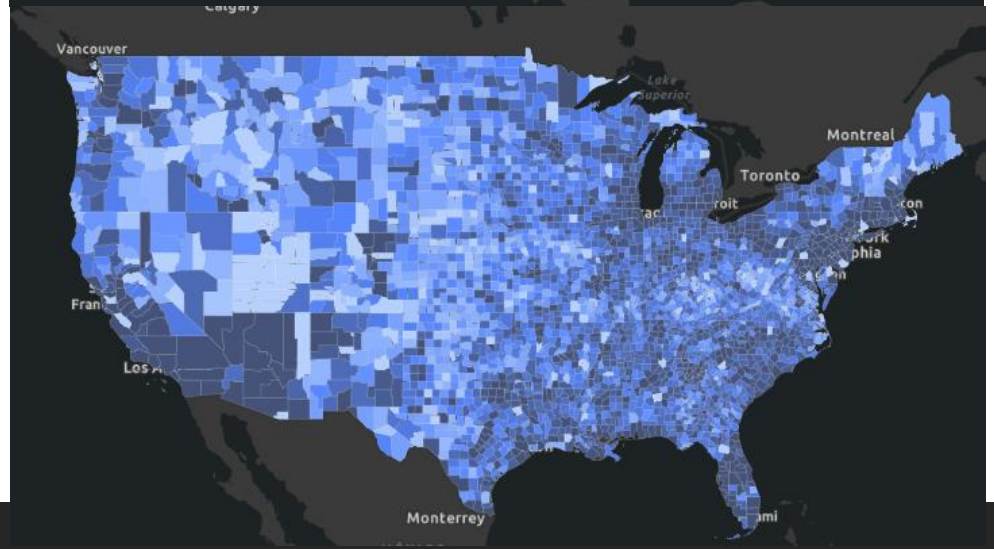


Confirmed Cases





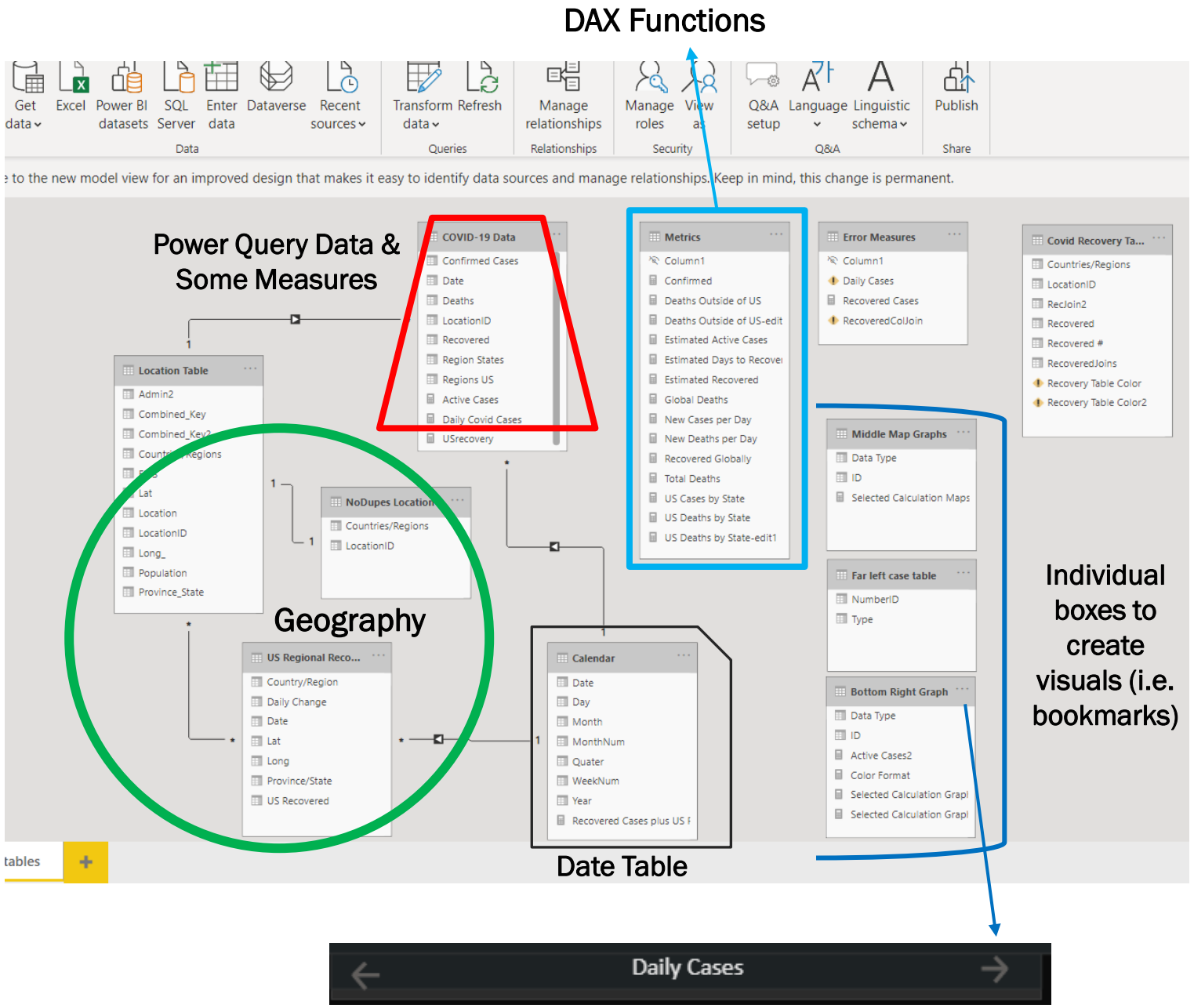
Deaths



Model View

- Shows all tables, columns, & relationships in the model
- This is how we can link data between multiple data sets or tables
- One to Many Relationships:
  - COVID-19 Data to Calendar via Date
  - Location Table to COVID-19 Data via LocationID
  - Calendar to US Regional Recov… via Date
- The “Metrics” table enables dynamic visuals as the functions are calculated in each visual interaction
- All Data Sourced From John Hopkins’s GitHub Page

https://github.com/CSSEGISandData/COVID-19



Power BI Challenges:

- Troubleshooting DAX / Measure errors & verifying correct values
- Starting over 3 times after identifying a better way to model & add data (increasing speeds)
- Including US County data is more involved to combine it with global data (3 location tables were used for this)
  - Otherwise there would be 1 pink dot displayed in the US or 1 for each US state / territory (i.e. shown on pages 9,10,11)
- Format and aesthetics were **trial and error & a 3rd of time** spent
- Power BI has limits while constantly adding new features (i.e. not including a green recovered column label)
- Default (left) table cannot create the middle table by itself

Default Power BI

Countries/Regions	Recovered Globally
India	10,016,859
Brazil	7,033,286

Tom's Attempt

John Hopkins Uni

Appendix A. DAX Sample Calculation

```
Confirmed COVID-19 Cases =

var LatestDate = MAX('COVID-19 Data'[Date])

return

CALCULATE(SUM('COVID-19 Data'[Confirmed Cases]),'COVID-19 Data'[Date] = LatestDate)
```

= In English =

```
CALCULATE(SUM('Table Name'[Table Column]),'Table Name'[Column] = Filtered Data)
```

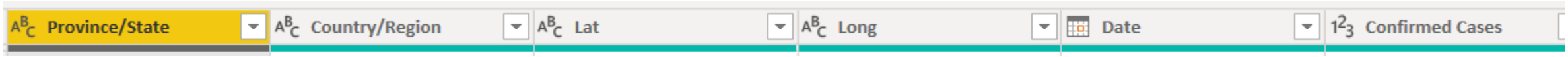
Data we want to examine

How we want to filter the data

- The calculate function is one way to filter data for:
  - Total Cases, Total Deaths, Total Recoveries…
  - Otherwise, our data would not be dynamic and contain inaccurate values

Appendix B. Gathering Data Setup

- Import time series data for global (confirmed, deaths, recovered) & ISO FIPS Lookup (this is the location table query to index countries / regions)
- This raw data is not efficient to model & must be reformatted
- Unpivot / pivot all columns that include date and COVID data to create 2 columns instead of having 100s
- Each individual confirmed, death, recovered query should look like this:

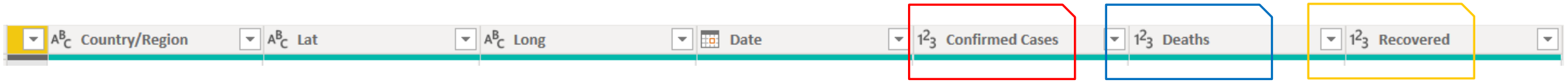


Time Series Data:

https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series

Appendix B. Gathering Data Cont.

- At this point we can merge all 3 queries into one given identical format while extracting the columns of interest (COVID-19 Data)



- On the Location Query, create an index column starting with 1, merge onto COVID-19 Query, then remove duplicate geographic data on the COVID query as it can be linked to the Location Query via the index column

LocationID	Date	Confirmed Cases	Deaths	Recovered
1	1/22/2020	0	0	0

- Location Query / Table Example:

FIPS	LocationID	Combined_Key	Admin2	Province_State	Countries/Regions	Lat
	1	Afghanistan			Afghanistan	3