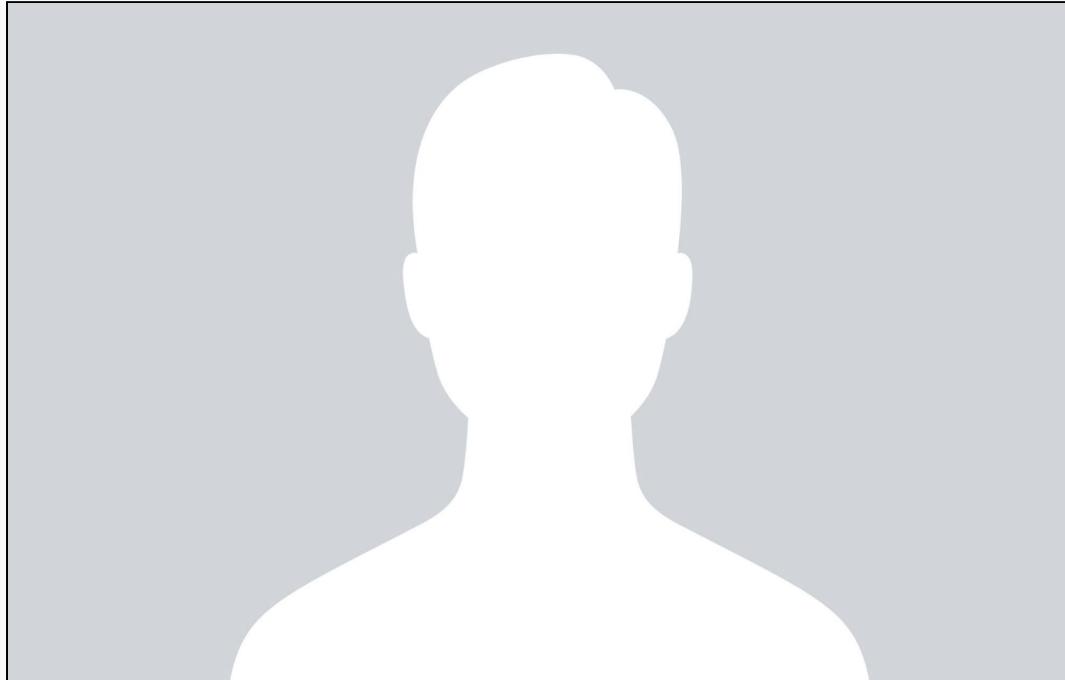


Data Science in the Time of Coronavirus

Storytelling with Data Visualizations

Srikanth Aravamuthan
September 12, 2022

Background & Research

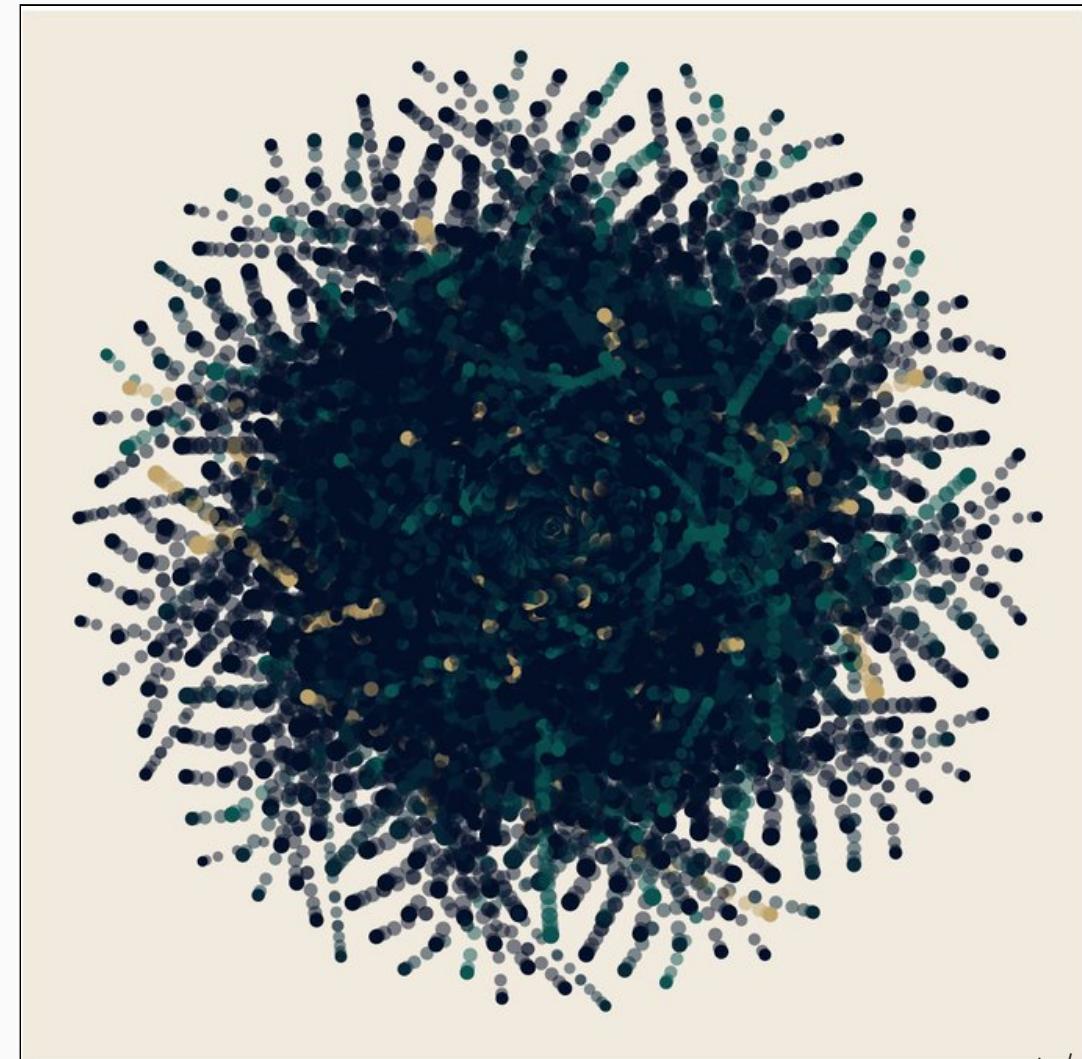
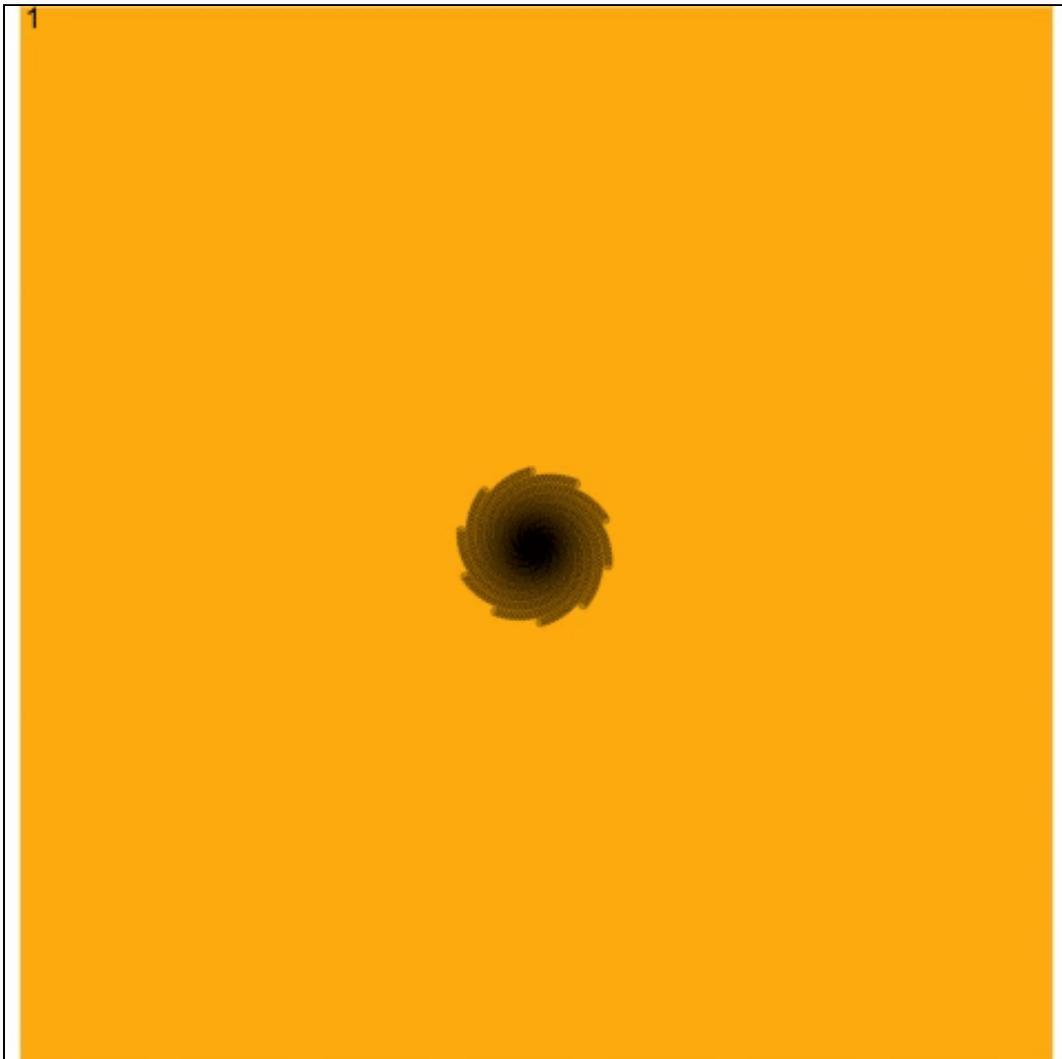


- Work for Prof. Dörte Döpfer
- Background in Math & Science
- Experience in Wet & Dry Lab Research

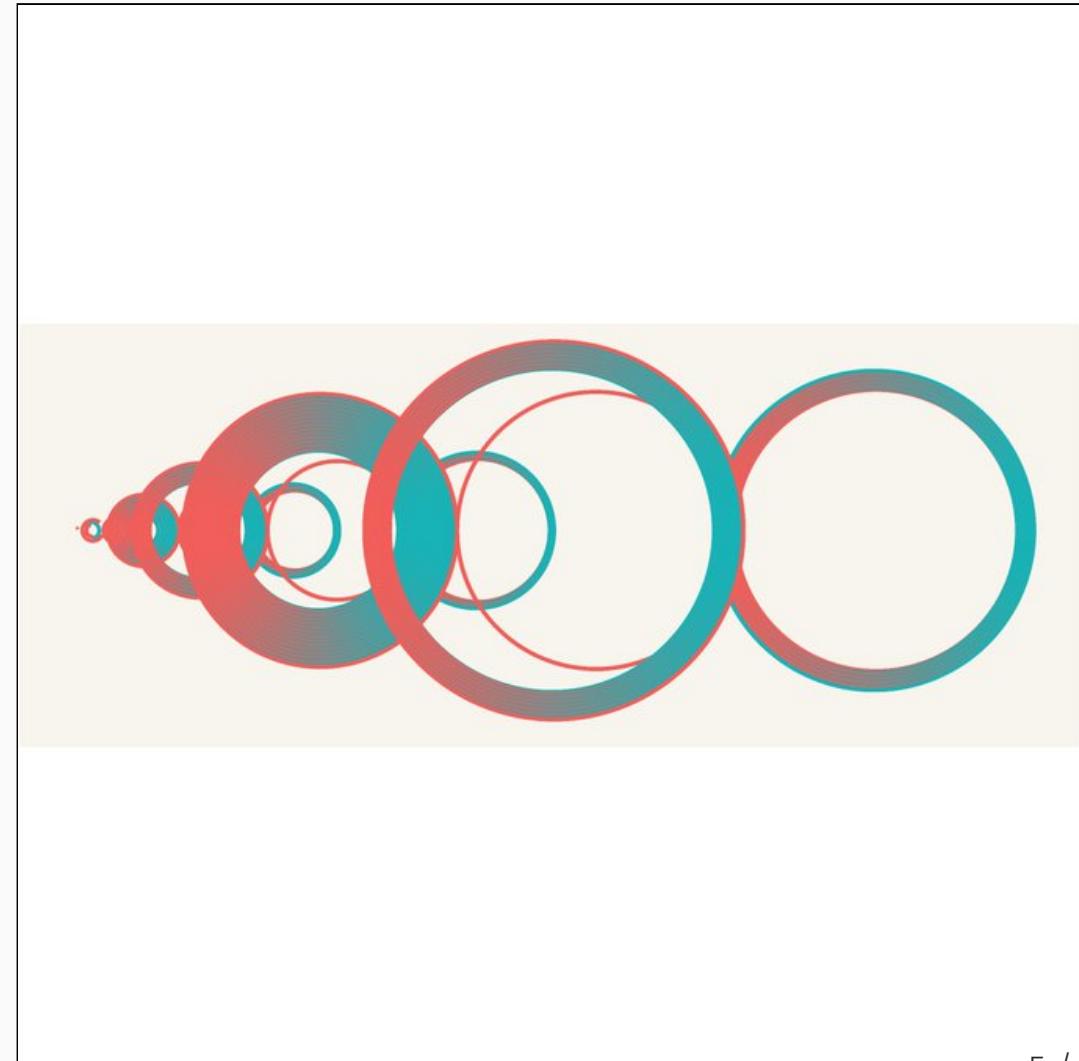
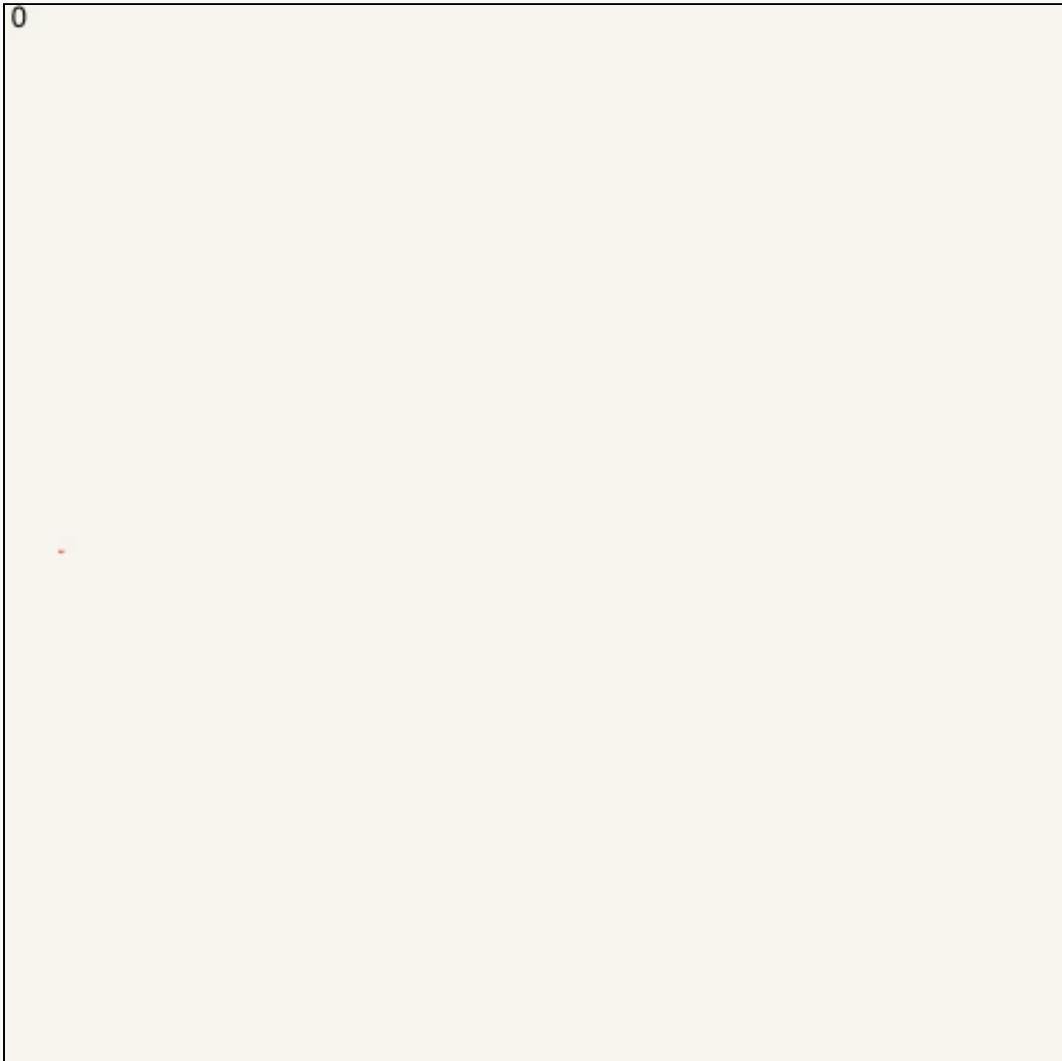
Are you an artist?



Generative aRt



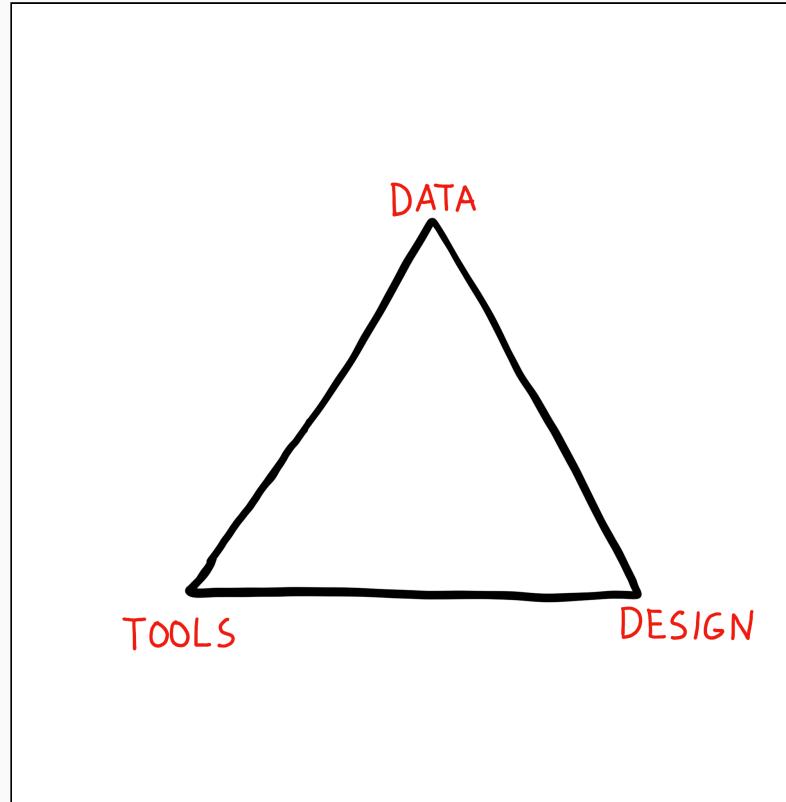
Generative aRt



Where do we go from here?

"The simple graph has brought more information to the data analyst's mind than any other device." - **John Tukey**

Data Visualization Triangle



Where are we going?

"The greatest value of a picture is when it forces us to notice what we never expected to see." - **John Tukey**

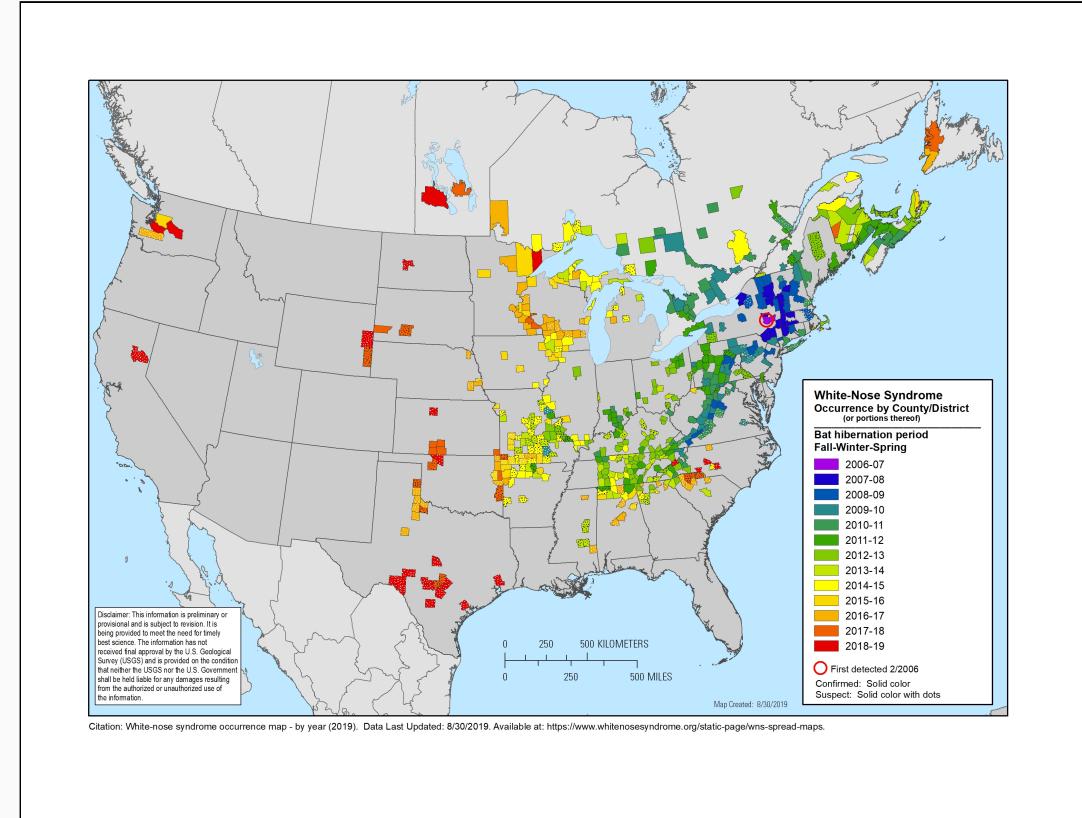
Ecological Diffusion of White Nose Syndrome

Daniel Walsh & Robin Russell (USGS)

Ting Fung Ma & Jun Zhu (University of Wisconsin-Madison)

White-Nose Syndrome (WNS)

- A disease that has killed millions of hibernating bats in North America.
- Caused by *Pseudogymnoascus destructans* (Pd), a fungus which damages skin so much that bats warm up and become active, wasting energy they need to make it through the winter.
- Found in 33 states and 7 Canadian provinces and the fungus without the syndrome had been found in 3 more states.



Goal

We aim to detect the time such that $p_{s,t} > \eta$, where η is a given decision threshold and future resource (sample) allocation will be based on the latest set of estimators $\{p_{s,t}\}_{s=1}^S$.

- To provide a list of recommended counties/ecoregions to sample in the upcoming active surveillance season
- To forecast the spread of the disease over the long-run in the North America

Recommendation

Table: High prevalence cells for sampling by growth rate

id	x	y	long	lat	COUNTY	STATE	SECTION	lambda
125	-1038045	1615005	-107.804	36.997	La Plata	Colorado	Navajo Canyon- lands Section	-0.012
232	-1068045	1705005	-108.273	37.760	Dolores	Colorado	South- Central Highlands	-0.012
24	-1008045	1705005	-107.591	37.827	San Juan	Colorado	Section South- Central Highlands	-0.011
32	-1008045	1945005	-107.936	39.952	Rio Blanco	Colorado	Tavaputs Plateau Section	-0.010
137	-1038045	1975005	-108.333	40.184	Rio Blanco	Colorado	Tavaputs Plateau Section	-0.010
22	-1008045	1645005	-107.508	37.296	La Plata	Colorado	South- Central Highlands Section	-0.008

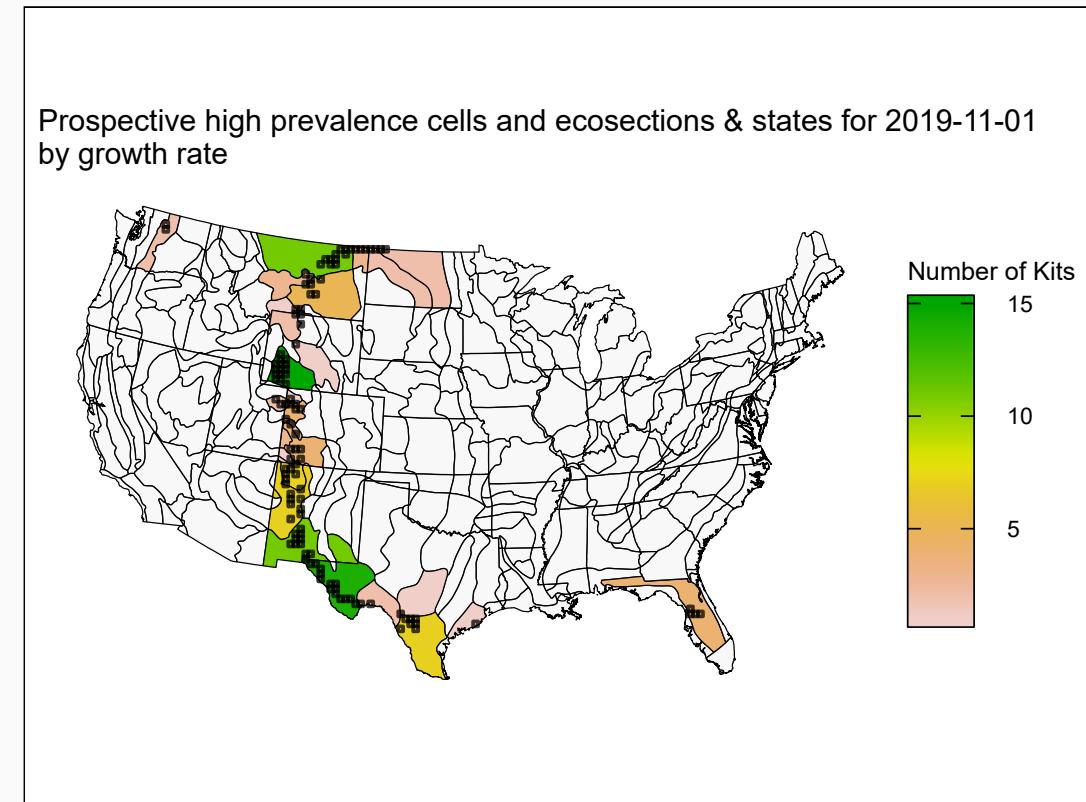
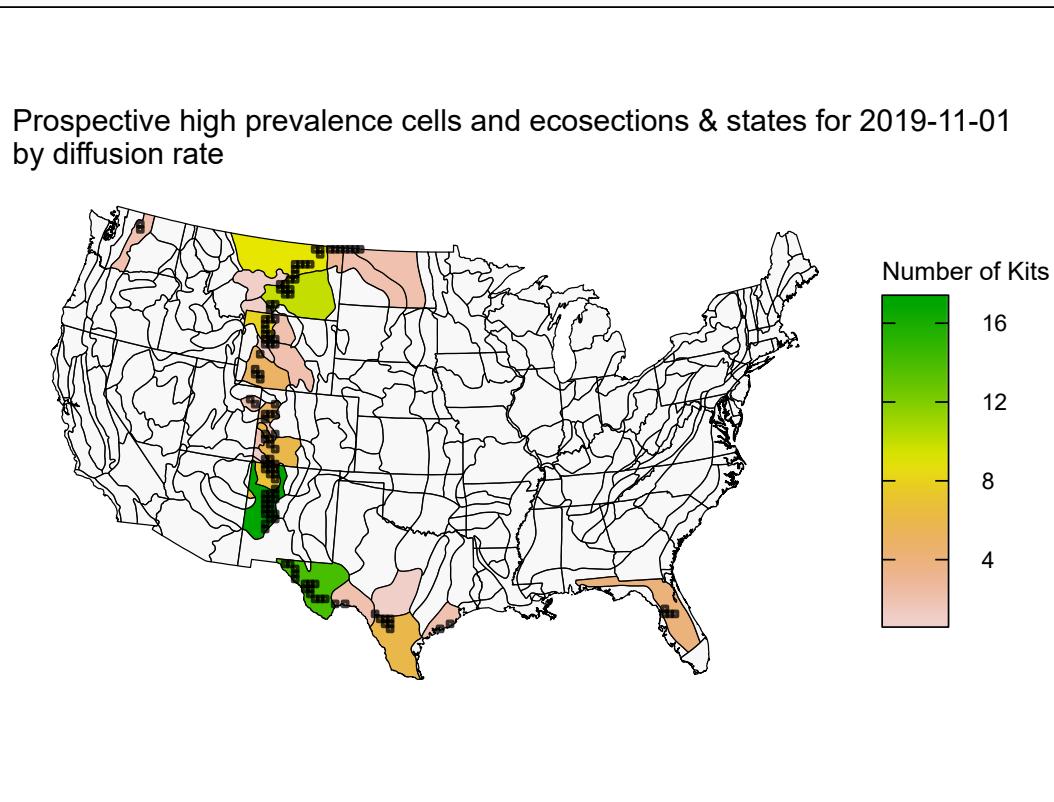
Recommendation

Table: High prevalence states and ecoregions for sampling by growth rate

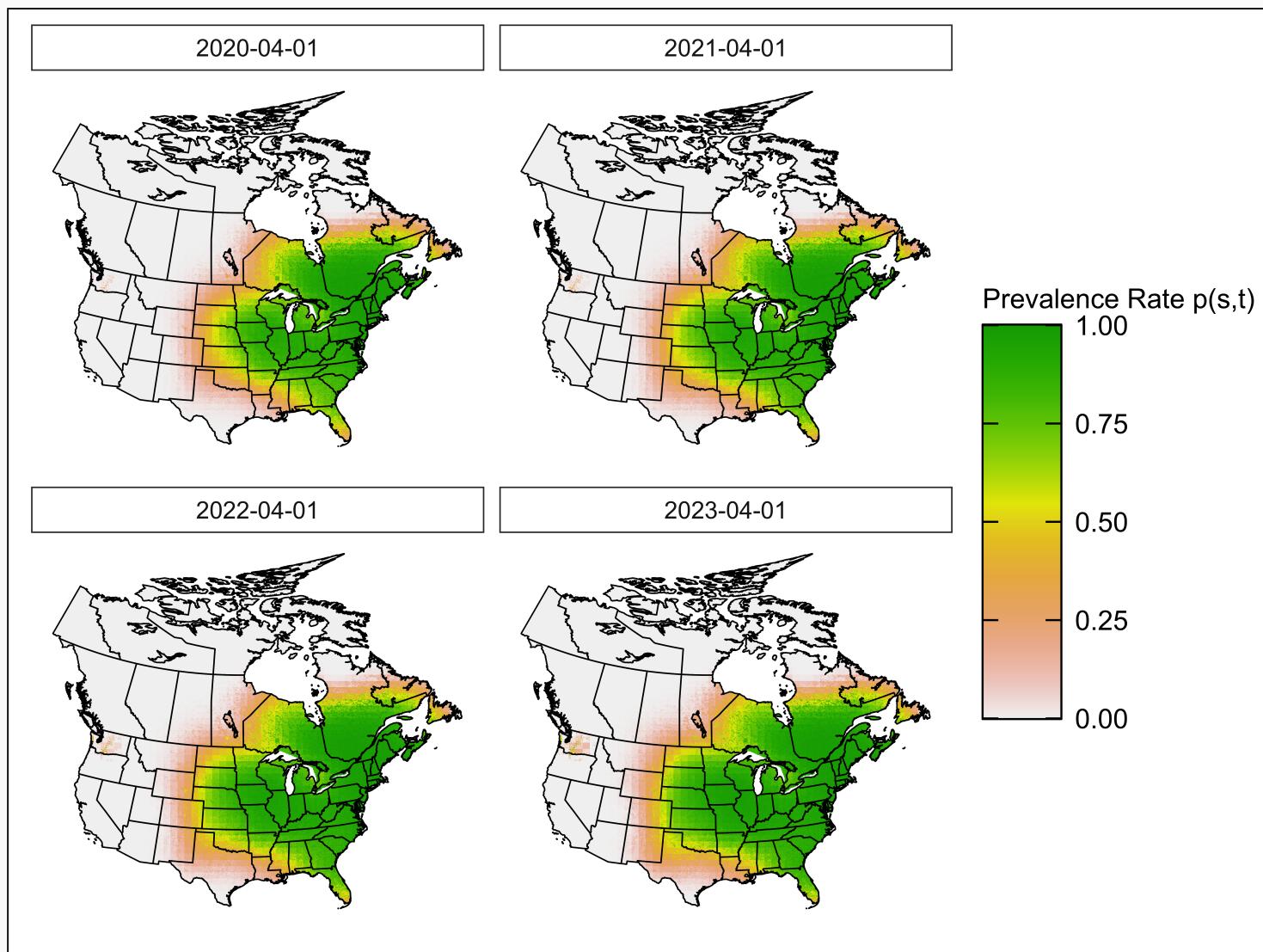
STATE	SECTION	k	x	y	long	lat	lambda
Colorado	Navajo Canyonlands	1	-1038045	1615005	-107.8044	36.99729	-0.0121115
Colorado	Section South- Central Highlands	4	-1030545	1690005	-107.8260	37.66930	-0.0085590
Colorado	Tavaputs Plateau Section	4	-1038045	1960005	-108.3101	40.05062	-0.0064228
North Dakota	Northeastern Glaciated Plains	2	-513045	2905005	-102.9291	48.98229	-0.0040968
Colorado	Section Grand Canyon Lands Section	1	-1068045	1645005	-108.1854	37.22919	-0.0000769
Colorado	Uinta Mountains Section	1	-1068045	2005005	-108.7325	40.41391	0.0024110

Recommendation

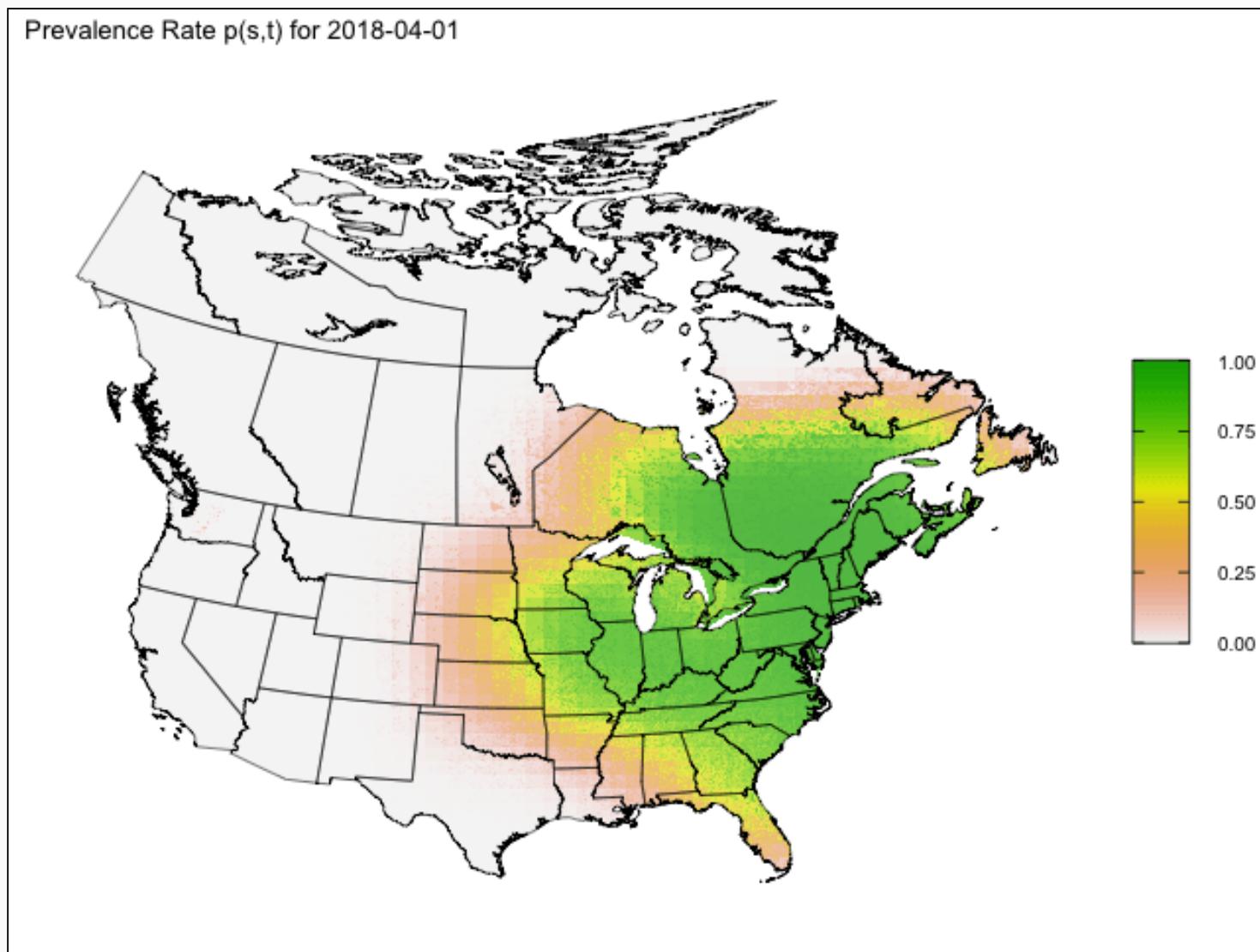
High prevalence cells for sampling by state and ecoregion



Forecasting of WNS



Forecasting of WNS



I'm *not* Batman



COVID-19 Visualizations of Descriptive Statistics

Sean Kent, Steve Goldstein, and Brian Yandell (University of Wisconsin-Madison)

Francisco Mandujano & Dörte Döpfer (University of Wisconsin-Madison)

New York Times Data Visualizations

The Upshot

How Severe Are Coronavirus Outbreaks Across the U.S.? Look Up Any Metro Area

By Josh Katz, Kevin Quealy and Margot Sanger-Katz Updated April 28, 2020, 2:11 PM E.T.

Daily growth rate of confirmed cases
once reaching 200 cases

+80%

+60%

+40%

+20%

-20%

Mar. 15 Mar. 22 Mar. 29 Apr. 5 Apr. 12 Apr. 19 Apr. 20

Sterling, Colo.

Sioux City, Iowa

Worthington, Minn.

Green Bay, Wis.

Providence, R.I.

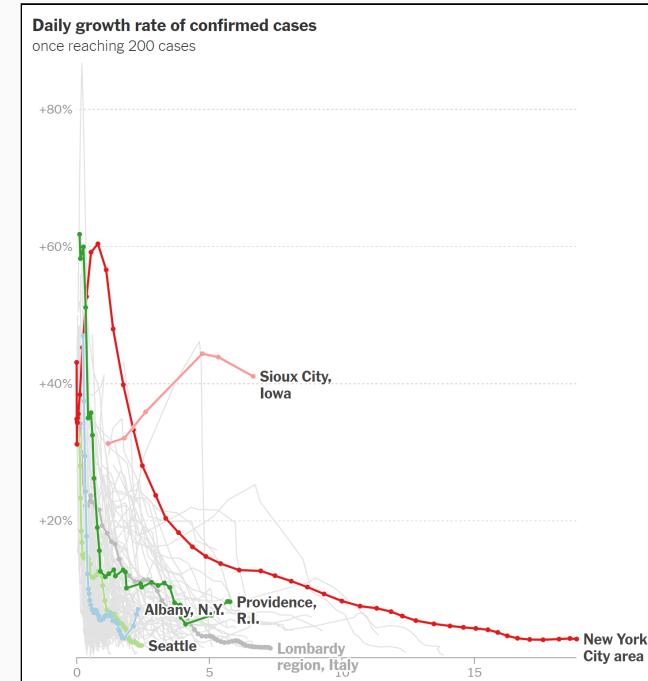
Lombardy region, Italy

Notes: Growth rates are averaged over the previous week.

Growth Rates by Case Count

The charts below show the growth rate by the number of cases or deaths in a given metropolitan area.

Choose a metro area



<https://www.nytimes.com/interactive/2020/04/03/upshot/coronavirus-metro-area-tracker.html>

COVID-19 Growth Rates in Wisconsin

"Wisconsin leaders need information about the extent and impact of COVID-19 to make good decisions." -

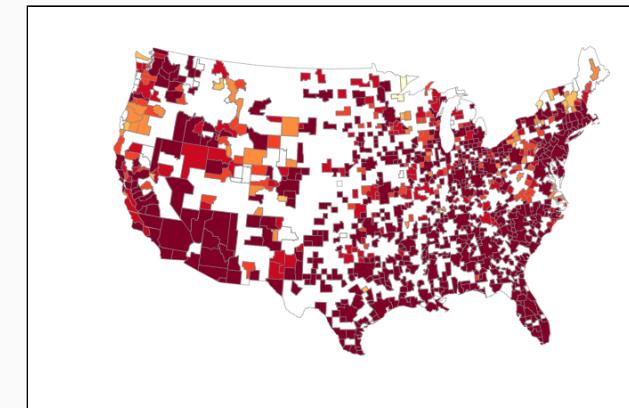
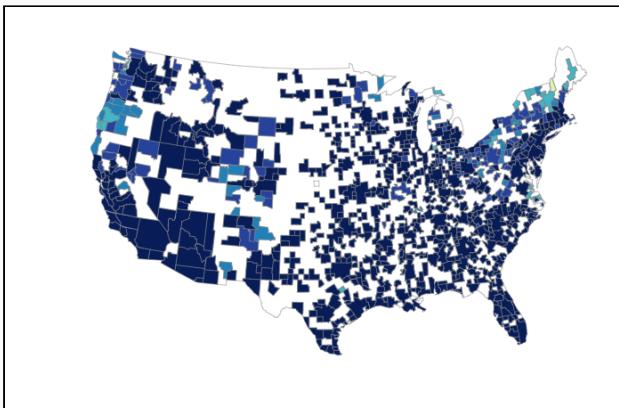
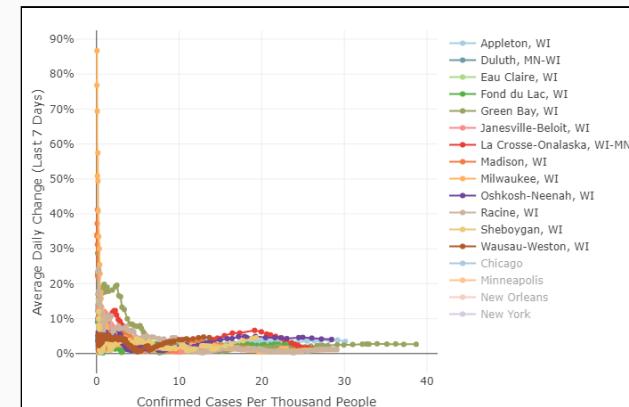
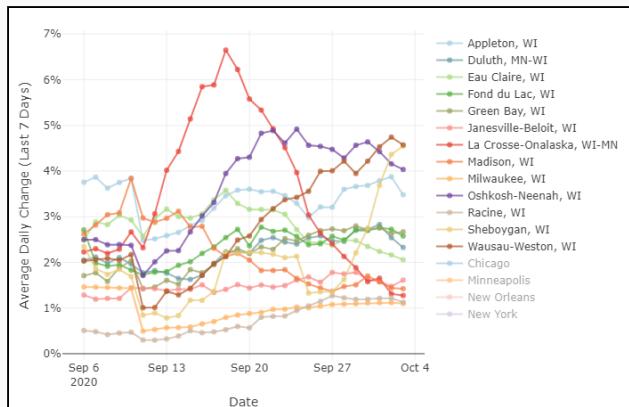
AFI DSI COVID-19 Research Group Charter

"Can we add value with visualizations of growth rates in Wisconsin metro areas?" - **Steve Goldstein**

Audience:

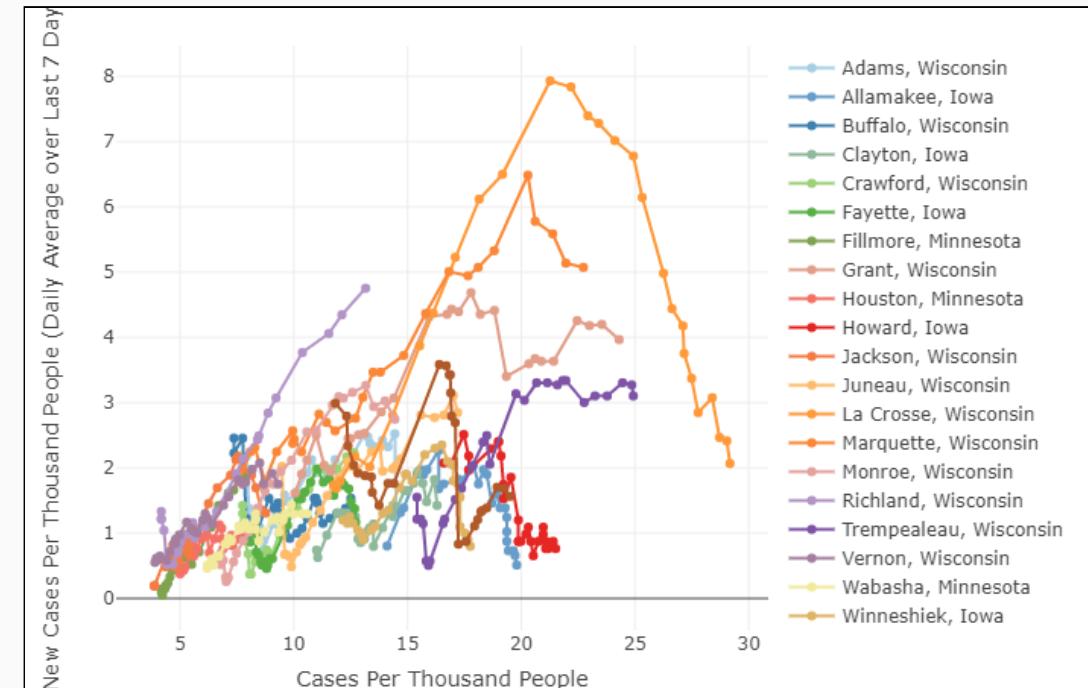
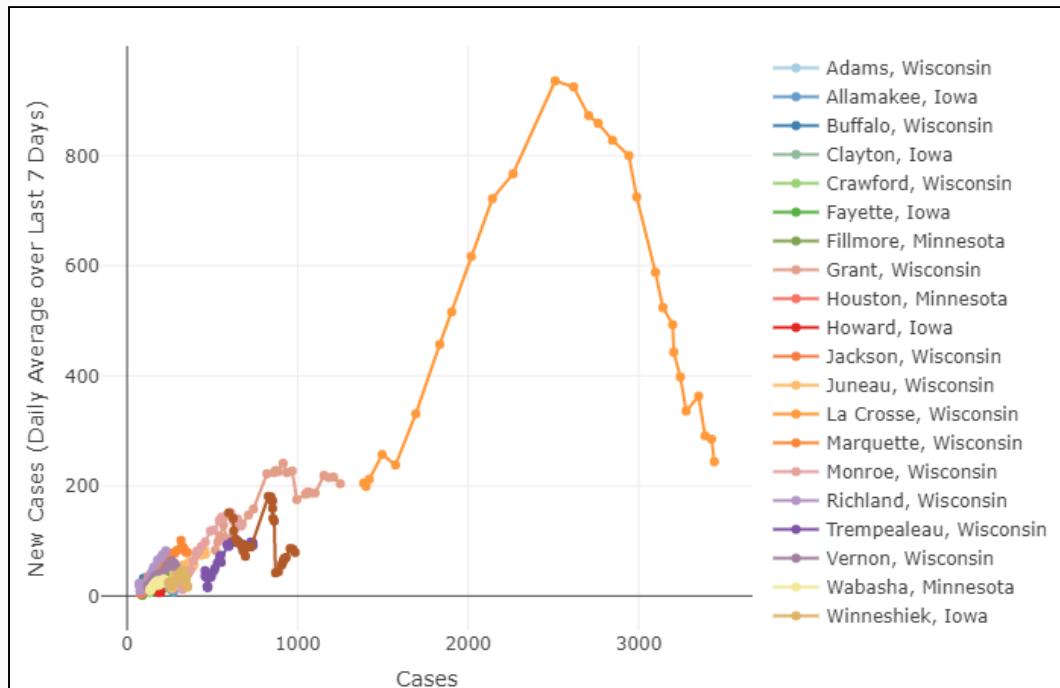
- Wisconsin public health officials
- Health system officials
- Interested modelers and citizens

COVID-19 Growth Rates in Wisconsin



data-viz.it.wisc.edu/wi-metro-growth-rate/

COVID-19 Growth Rates in Wisconsin



Website

- data-viz.it.wisc.edu/wi-metro-growth-gundersen/
- data-viz.it.wisc.edu/wi-metro-growth-marshfield/
- Daily updates to data
- Customized to the needs of health system



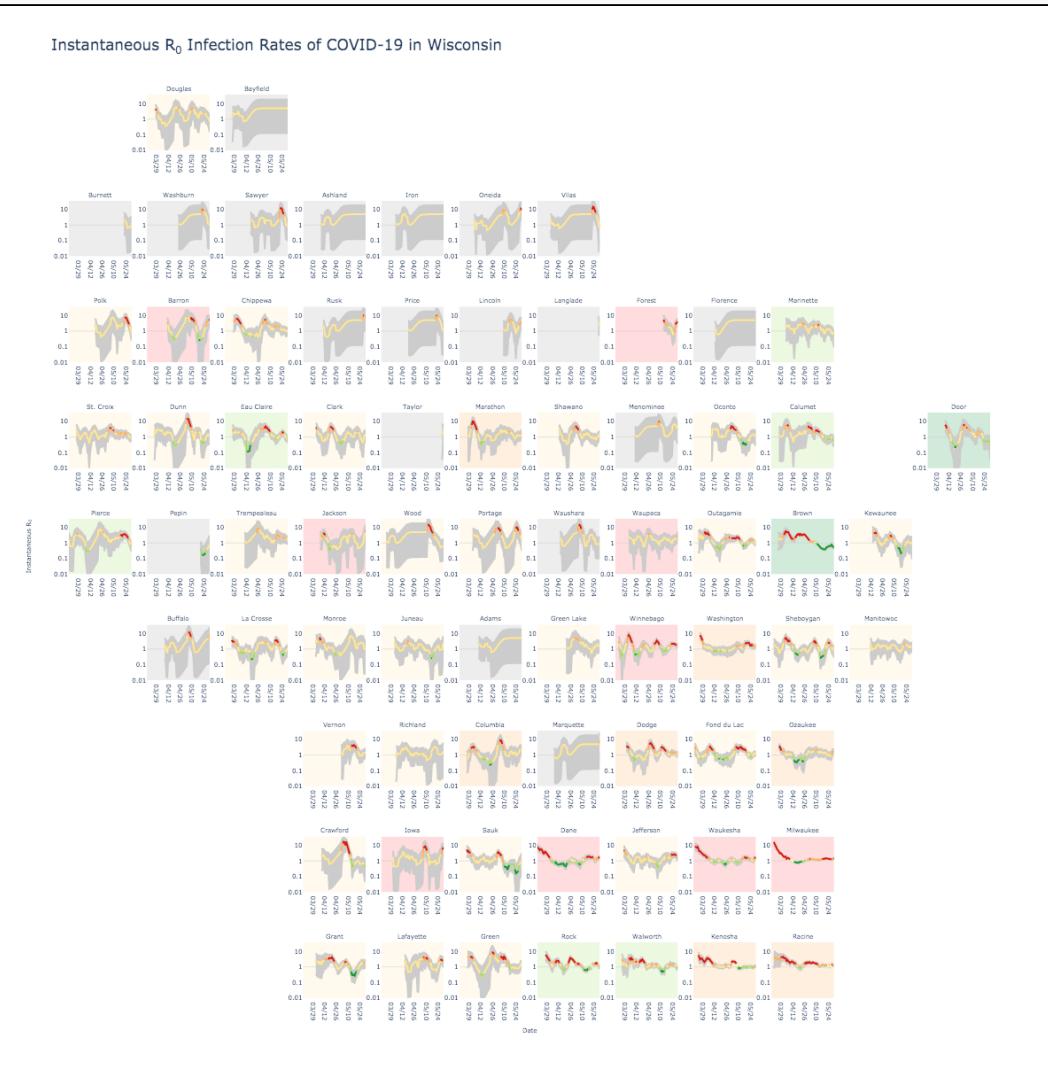
Geofaceting Instantaneous R_0

Project Idea: "Can we summarize R_0 spatially and temporally for all Wisconsin counties over all time points."

Audience:

- Wisconsin public health officials
- Interested modelers and citizens

Geofaceting Instantaneous R_0



- Preserves the geographical orientation
- Maintains the time dependence
- Won honorable mention in John Hunter Excellence in Plotting Competition at SciPy 2020
- Website: data-viz.it.wisc.edu/instantaneous-r0-geofacet-wi-county/

COVID-19 Activity Level in Wisconsin

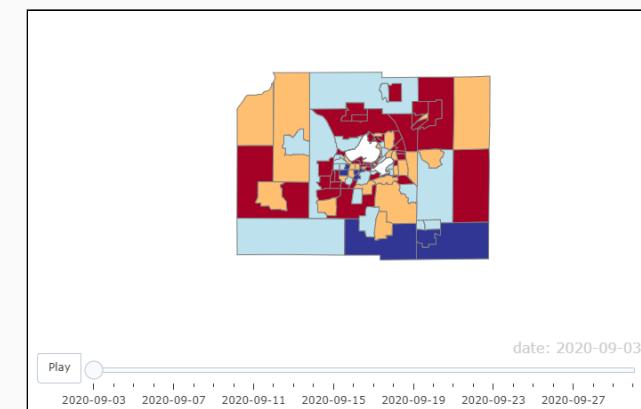
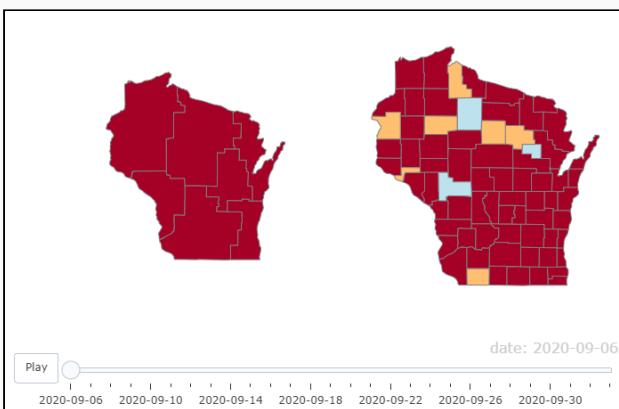
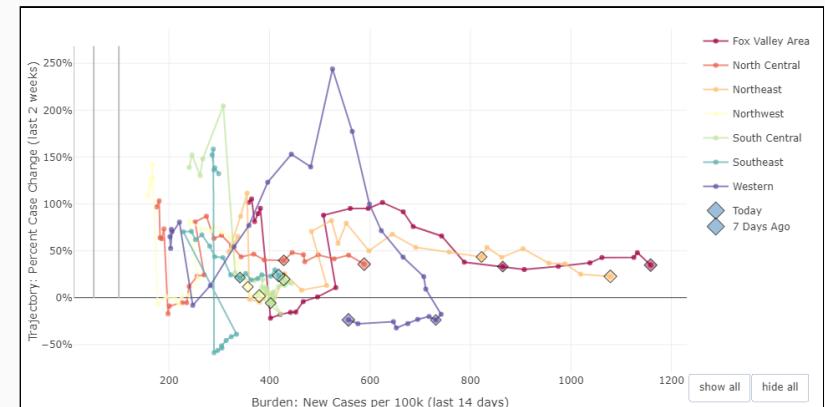
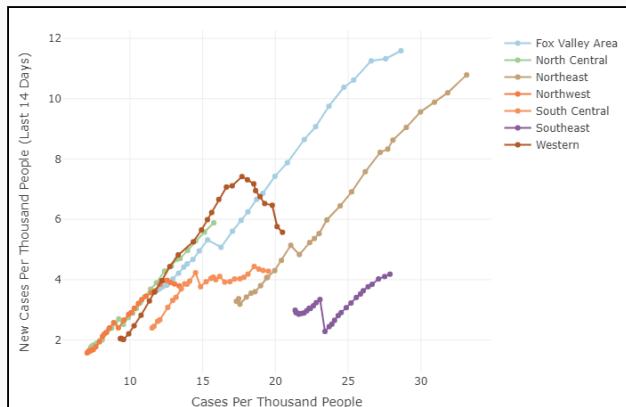
Project Idea:

- Overlay HERC regions on the county map
- View the census tract level map
- View time series

Audience:

- Wisconsin public health officials
- Interested modelers

COVID-19 Activity Level in Wisconsin



data-viz.it.wisc.edu/dashboard/

SARS-CoV-2 Screening Strategies

JAMA Network Open™

Original Investigation | Public Health

Assessment of SARS-CoV-2 Screening Strategies to Permit the Safe Reopening of College Campuses in the United States

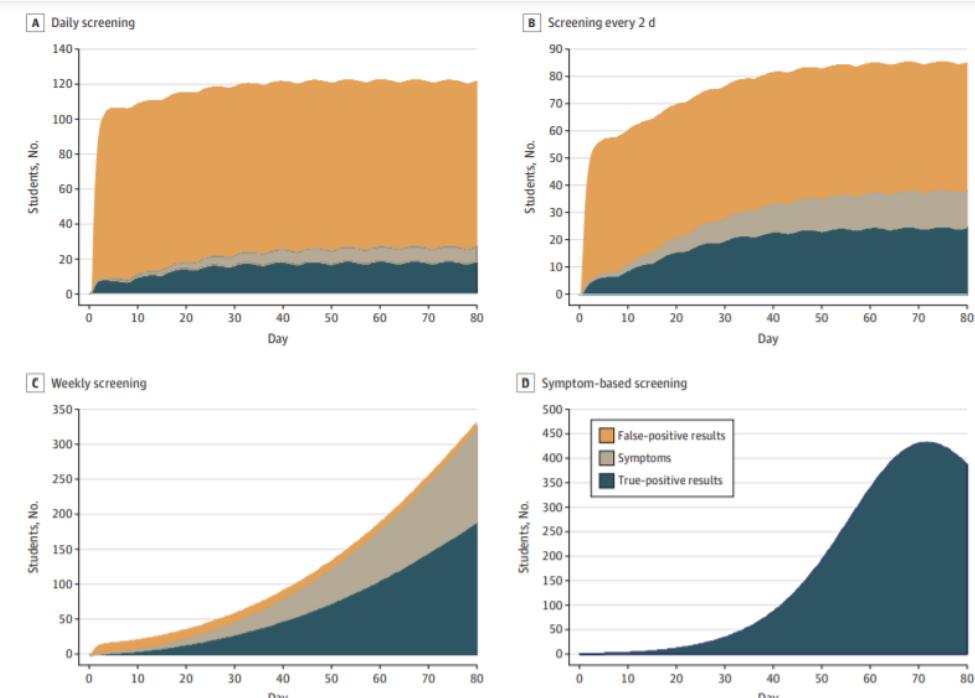
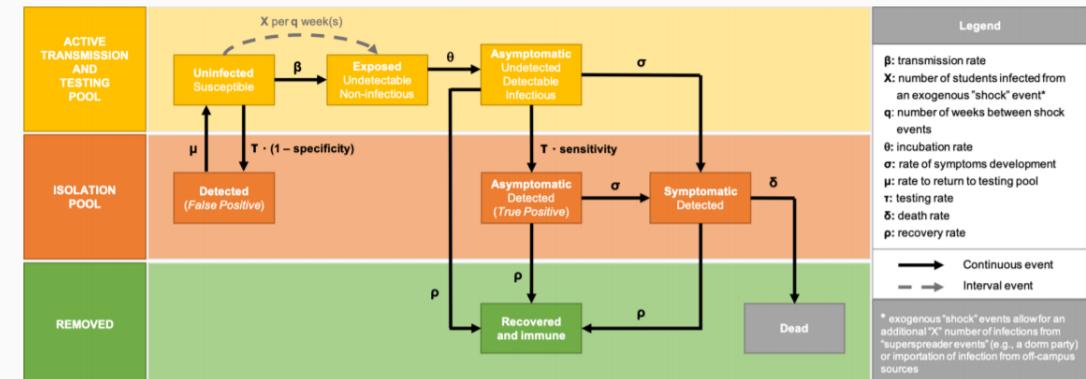
A. David Palatiel, PhD; Amy Zheng, BA; Rochelle P. Walensky, MD, MPH

medRxiv preprint doi: <https://doi.org/10.1101/2020.06.22.20136309>; this version posted September 8, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance

Daniel B. Larremore^{†1,2}, Bryan Wilder³, Evan Lester^{6,5}, Soraya Shehata^{4,5}, James M. Burke⁶, James A. Hay^{7,8}, Milind Tambe³, Michael J. Mina^{†7,8,9,*}, and Roy Parker^{§4,6,10,2,*}

<https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2768923>



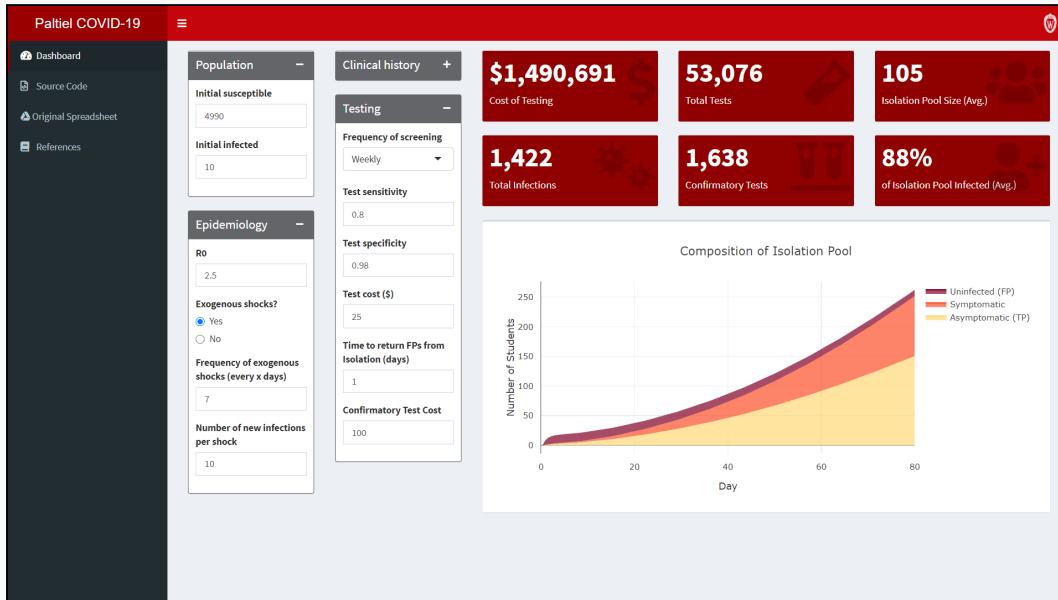
SARS-CoV-2 Screening Strategies

Project Idea: *"Can we implement an interactive dashboard with improved accessibility and ease of use."*

Audience:

- University of Wisconsin-Madison campus leaders
- Interested modelers

SARS-CoV-2 Screening Strategies



- Calculate minimum performance attributes of a SARS-CoV-2 monitoring program
- Suggest what isolation and treatment capacity would need to be in place
- Forecast what testing might cost
- Recommended by the authors of the paper
- Website: data-viz.it.wisc.edu/covid-19-screening/

Nowcast COVID-19 Cases and Hospitalizations

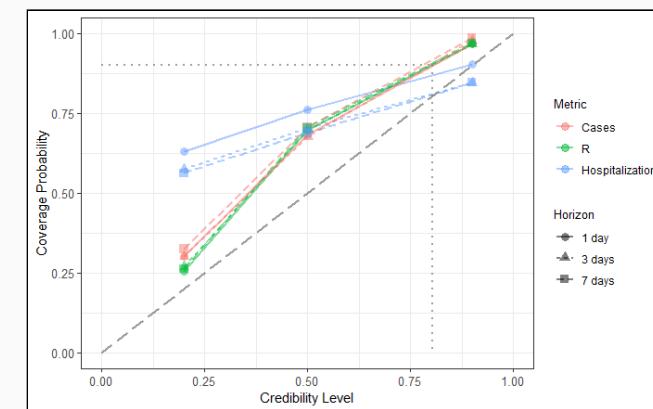
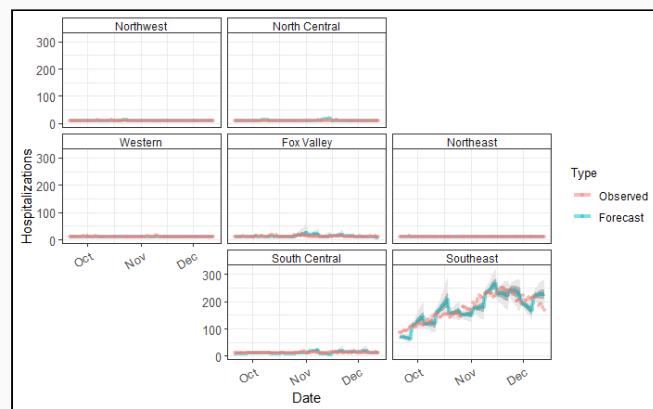
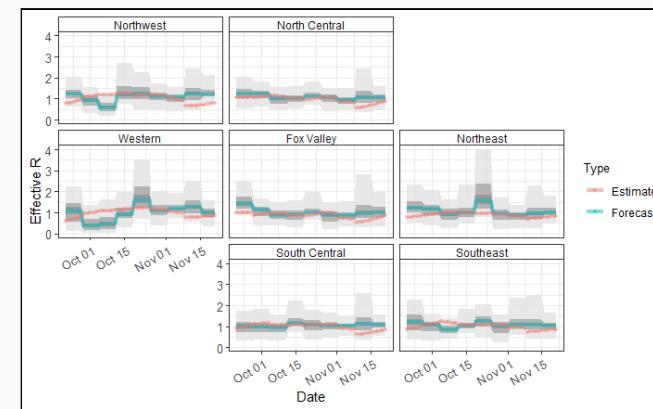
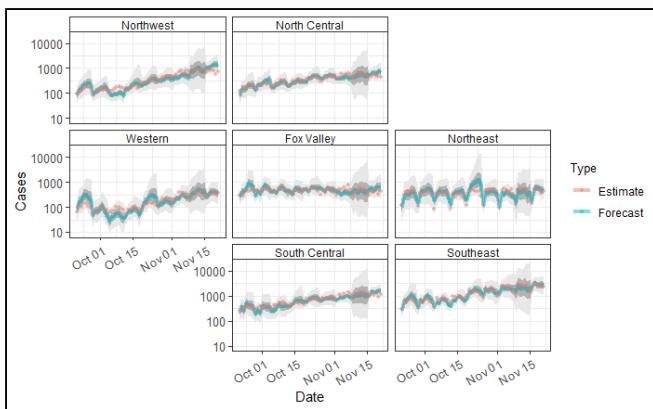
Project Idea:

- Nowcast cases by HERC regions
- Nowcast hospitalizations by HERC regions
- View time series
- View coverage of credible intervals

Audience:

- Wisconsin public health officials
- Interested modelers

Nowcast COVID-19 Cases and Hospitalizations



<https://data-viz.it.wisc.edu/cases-r-hosp-geofacet-wi-region/>

Nowcast COVID-19 Cases and Hospitalizations

Real-Time Estimation and Forecasting of COVID-19 Cases
and Hospitalizations in Wisconsin HERC Regions for Public
Health Decision Making Processes

Srikanth Aravamuthan ^{1 + 2 *}, Juan Francisco Mandujano Reyes ^{1 + 2}, Dörte
Döpfer ¹

1 Department of Medical Science, School Of Veterinary Medicine, University of
Wisconsin, Madison, WI, 53706

2 Department of Statistics, University of Wisconsin, Madison, WI, 53706

* Corresponding author: aravamuthan@wisc.edu

Automate DairyCOMP 305

Emil Walleser & Dörte Döpfer (University of Wisconsin-Madison)

DairyCOMP 305



Automate DairyCOMP 305

Project Idea:

- Automate DairyCOMP 305 write out
- Apply tools to support data-driven decisions
- Customize reports and graphs for a better management experience

Audience:

- Dairy Skills I instructors & students
- DairyCOMP 305 users

DairyCOMP 305

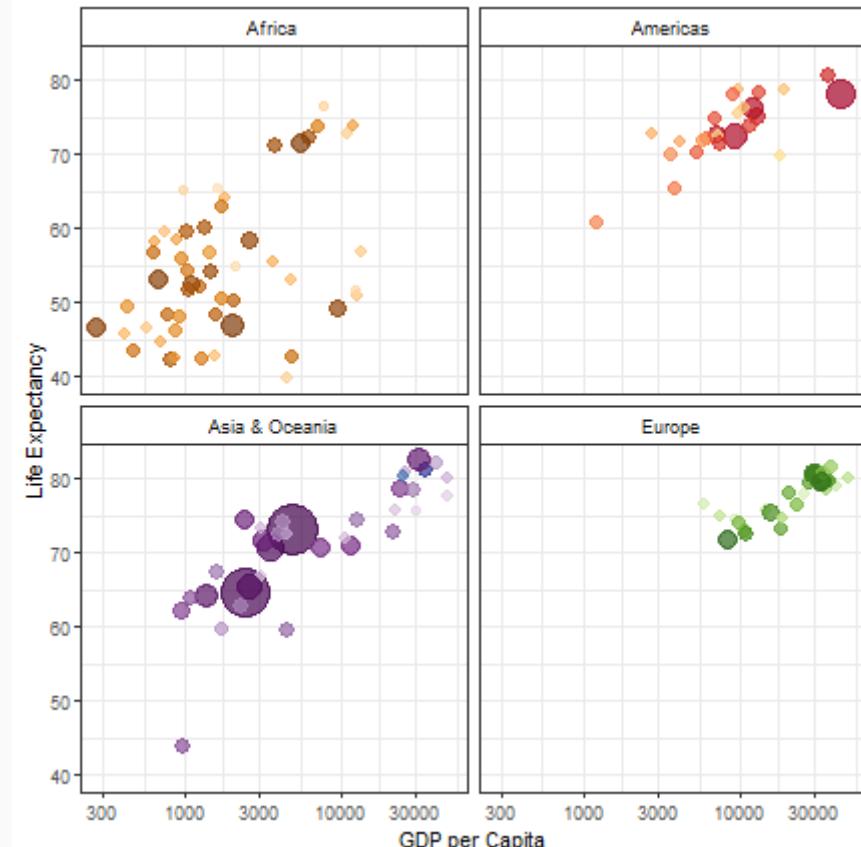
Automate DairyCOMP 305



```

gapminder %>%
  filter(year == max(year)) %>%
  mutate(continent = ifelse(continent %in% c("Asia", "Oceania"), "Asia & Oceania", continent))
  ggplot() +
  aes(x = gdpPercap,
      y = lifeExp) +
  geom_point(aes(size = pop,
                 color = country),
             alpha = 0.7,
             show.legend = FALSE) +
  scale_color_manual(values = country_colors) +
  scale_size(range = c(2, 12)) +
  scale_x_log10() +
  facet_wrap(vars(continent)) +
  labs(x = "GDP per Capita",
       y = "Life Expectancy") +
  theme_bw() +
  theme(strip.background = element_rect(fill = "white"))

```

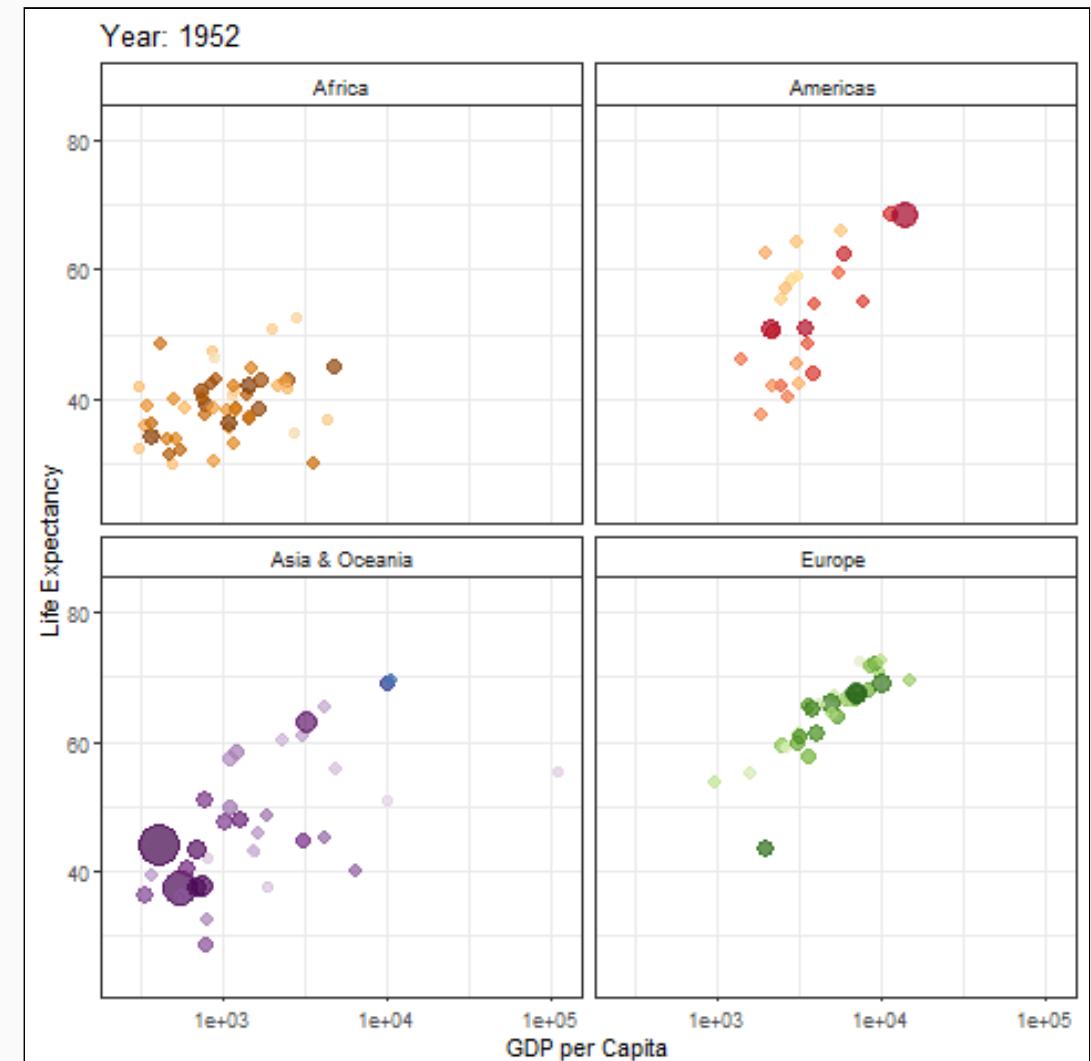


◀ ▶

```

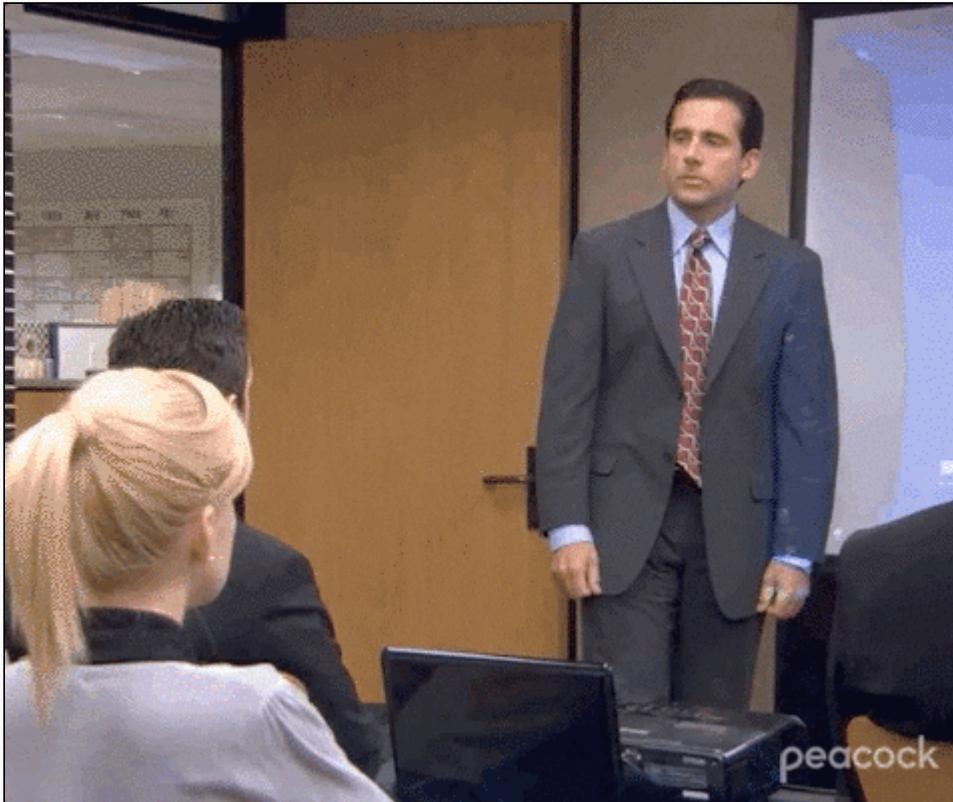
gapminder %>%
  mutate(continent = ifelse(continent %in% c("Asia", "Oceania"), "Asia & Oceania", continent))
  ggplot() +
    aes(x = gdpPercap,
        y = lifeExp) +
    geom_point(aes(size = pop,
                   color = country),
               alpha = 0.7,
               show.legend = FALSE) +
    scale_color_manual(values = country_colors) +
    scale_size(range = c(2, 12)) +
    scale_x_log10() +
    facet_wrap(vars(continent)) +
    labs(title = "Year: {frame_time}",
         x = "GDP per Capita",
         y = "Life Expectancy") +
    theme_bw() +
    theme(strip.background = element_rect(fill = "white"))
  transition_time(year) +
  ease_aes('linear')

```



Presenting Data Visualizations with Data Visualizations

Presentation Application



Presentation Application

The screenshot shows the RStudio interface with a presentation application open. The left pane displays the R Markdown code for the presentation:

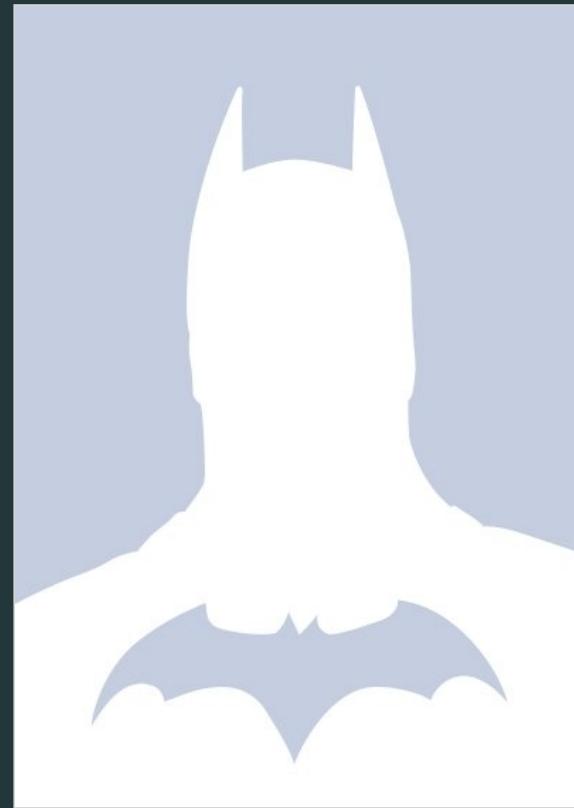
```
1 ---  
2 title: "Data science in the Time of Coronavirus"  
3 subtitle: "Storytelling with data Visualizations"  
4 author: "Srikanth Aravamuthan"  
5 date: "September 12, 2022"  
6 output:  
7   xaringan::moon_reader:  
8     self_contained: true  
9     lib_dir: "lib"  
10    css: default, metropolis, metropolis-fonts]  
11    fig_caption: true  
12    nature:  
13      ratio: 16:9  
14      highlightStyle: github  
15      highlightLines: true  
16      countIncrementalSlides: false  
17 ---  
18  
19```{r setup, include=FALSE}  
20 knitr::opts_chunk$set(comment=NA,  
21   warning=FALSE,  
22   message=FALSE,  
23   Cache = FALSE,  
24   # Cache = TRUE,  
25   # fig.pos="H",  
26   fig.height = 6,  
27   fig.width = 6,  
28   fig.align = "center")  
29 options(tibble.print_max = 20)  
30  
31 library(tidyverse)  
32 library(runnr)  
8021  "Thank You!!" :  
33  
34
```

The right pane shows the presentation slide, which includes the title, subtitle, author, date, and a "Thank You!!" message.

Data Science in the Time of Coronavirus
Storytelling with Data Visualizations

Srikanth Aravamuthan
September 12, 2022

Thank You!



Thank You!



