Personal Projects

John Goldin

Thursday, March 18

Motto for this talk:

What can you do when you're retired...

...anything you want

Topics

Having a personal presence on the web

An abundance of data via API's

My current deep dive into the Apple Health Export

4

Having a Presence on the Web

Why?

"Publication precipitates reality." -- Lila Freedman, former editor of the Yale Blue Book

My blog: johngoldin.com

Feeling part of a larger world

- Google American walker in Britain
- Google "Apple Health Export"

Professional visibility (not so relevant for a retiree, and maybe a risk for others)

Caveats:

8

Publishing to the Web

I've created some notes to accompany this talk. They are available at

https://goldin-projects-2021.netlify.app/

The notes point you to a video by Alison Hill & Desirée De Leon that describe a number of ways to use RMarkdown to publish to the web. The first and simplest technique was used to create the notes page.

1. Knit an RMarkdown document
2. Sign into Netlify.com
3. Drag the folder containing the project onto the deploy page in Netlify

When first deployed, Netlify creates an arbitray name for the URL. But I have the option to change it, provided I can supply a name that hasn't already been used.

There are lots of other ways to have a presence on the Web, such as LinkedIn, Twitter, Instagram, Flickr, and so on.

Where To Find Some Data

Getting data via web API's

- Connecticut and CDC Covid-19 data (via RSocrata)

```
dph_counties <- read.socrata("https://data.ct.gov/resource/bfnu-rgqt.json",
                             app_token = Sys.getenv("CTDATA_APP1_TOKEN")) %>
  as_tibble() %>%
  rename(cases = totalcases, deaths = totaldeaths) %>%
  mutate(date = as_date(dateupdated),
         cases = as.numeric(cases), deaths = as.numeric(deaths),
         tests = NA_real_, tests_positive = NA_real_,
         hospital = as.numeric(hospitalization),
         confirmedcases = as.numeric(confirmedcases),
         confirmeddeaths = as.numeric(confirmeddeaths),
         probablecases = as.numeric(probablecases),
         probabledeaths = as.numeric(probabledeaths)) %>%
  select(-dateupdated, -hospitalization)
```

An API is much easier to use if there is an R package that manages it

Some examples:

US Census: tidycensus and tigris

Twitter: twitteR or rtweet

FRED economic stats: fredr

More links in the notes at goldin-projects-2021.netlify.app

11

I don't need no stinkin' guvment databases to satisfy my data anlysis habit.

I've got almost three million rows of data in my pocket!

I did a blog post last winter describes how to export Apple Health Export and load it into R.

What's in the Apple Health Export dataset?

Frequency Data by Category

category	n	percent
Energy	1,483,322	49.600%
Heart Rate	514,203	17.200%
Distance	417,707	14.000%
Steps	240,152	8.000%
Exercise	190,119	6.400%
Dietary	72,541	2.400%
Mobility	25,497	0.900%
Audio	24,593	0.800%
Blood Pressure	8,688	0.300%
Sleep	5,978	0.200%

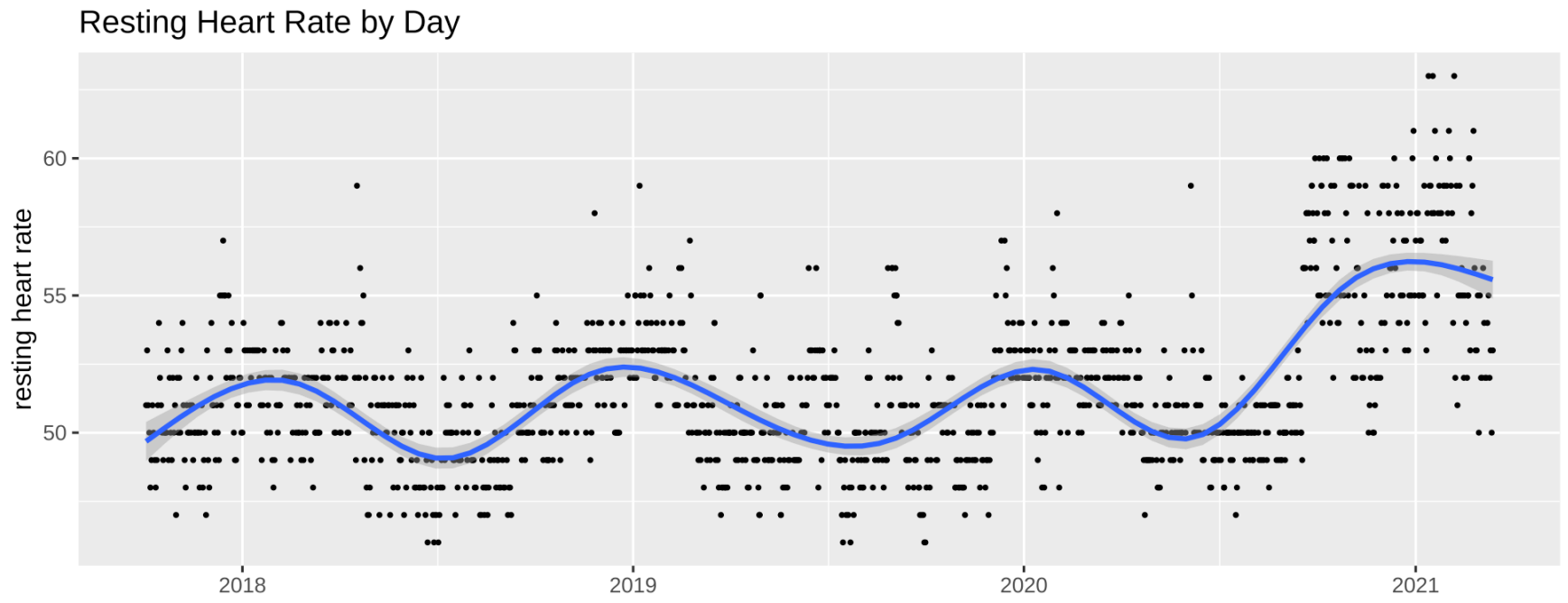
Frequency Data by Category

	category	n	percent
11	Oxygen Saturation	2,962	0.100%
12	Resting Heart Rate	1,318	0.000%
13	Body Metric	659	0.000%
14	VO2Max	489	0.000%
15	Mindful	145	0.000%
16	Symptoms	2	0.000%
17	Times Fallen	2	0.000%
18	ECG	1	0.000%
19	Unknown	1	0.000%

14


Resting Heart Rate by Day
resting heart rate
60
55
50
2018
2019
2020
2021

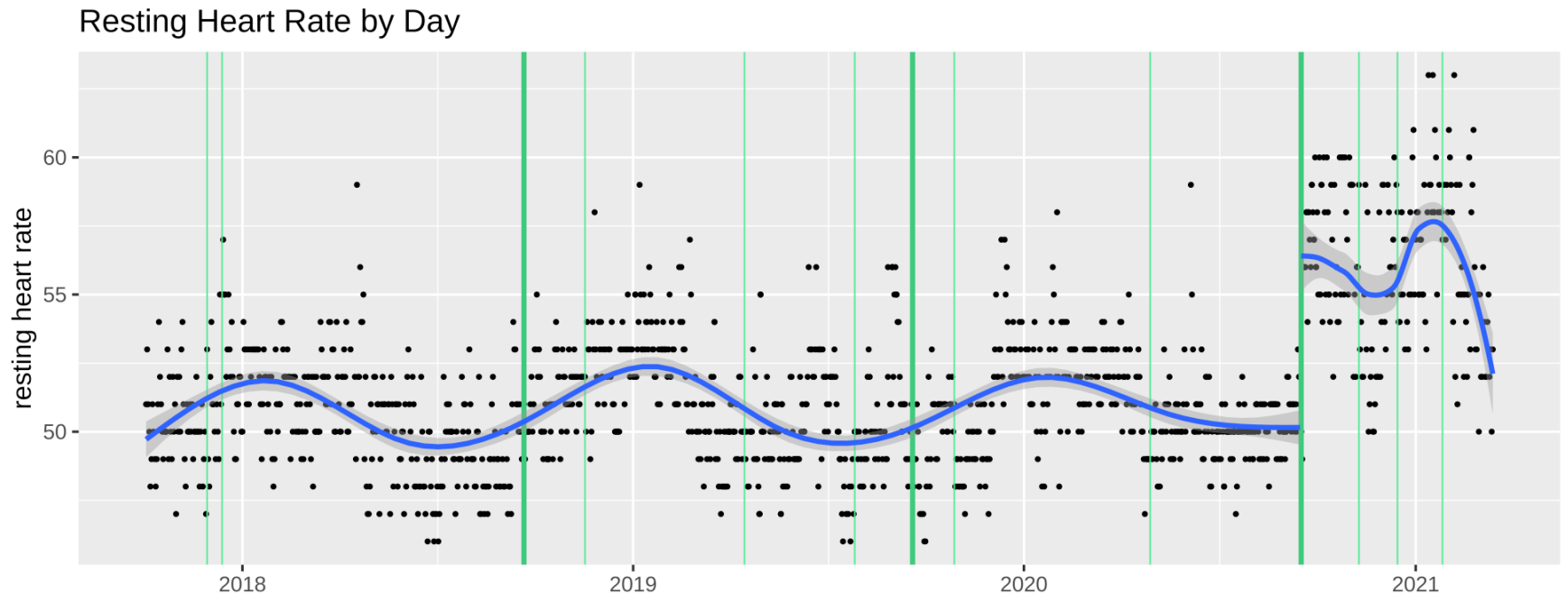
15


Resting Heart Rate by Day
resting heart rate
60
55
50
2018
2019
2020
2021

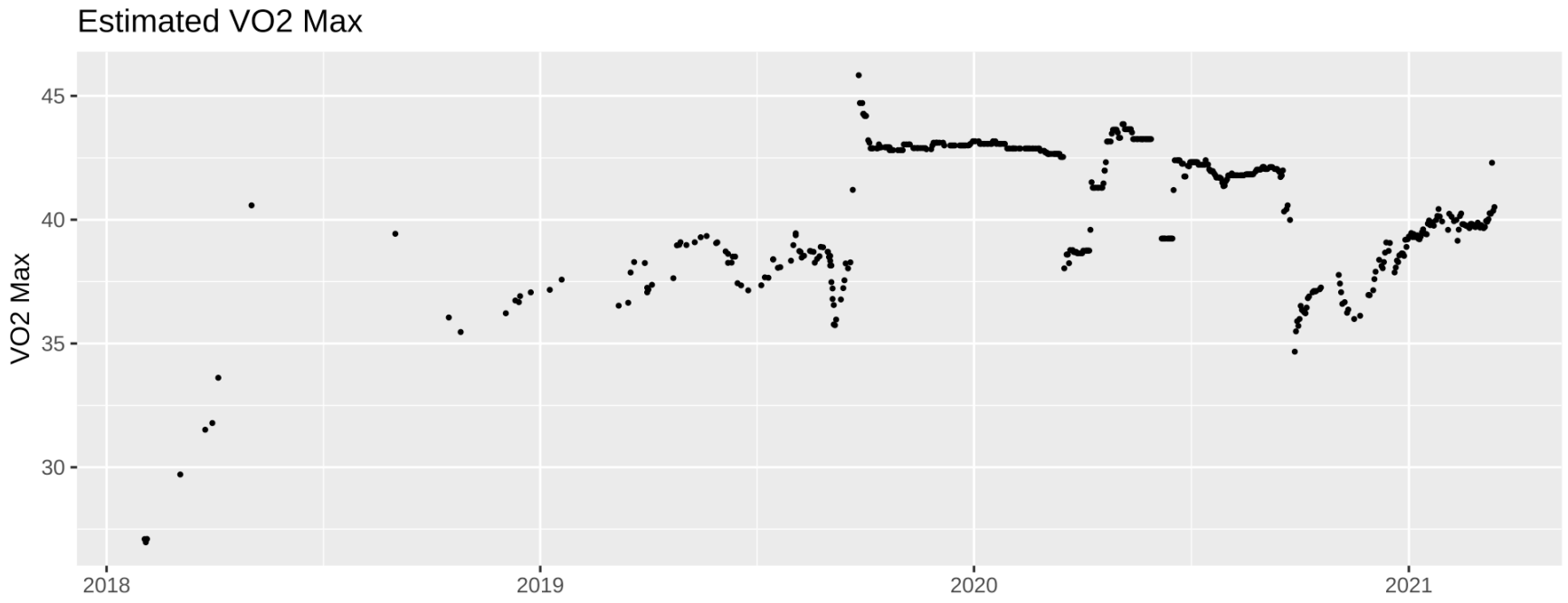
16


Resting Heart Rate by Day
resting heart rate
60
55
50
2018
2019
2020
2021

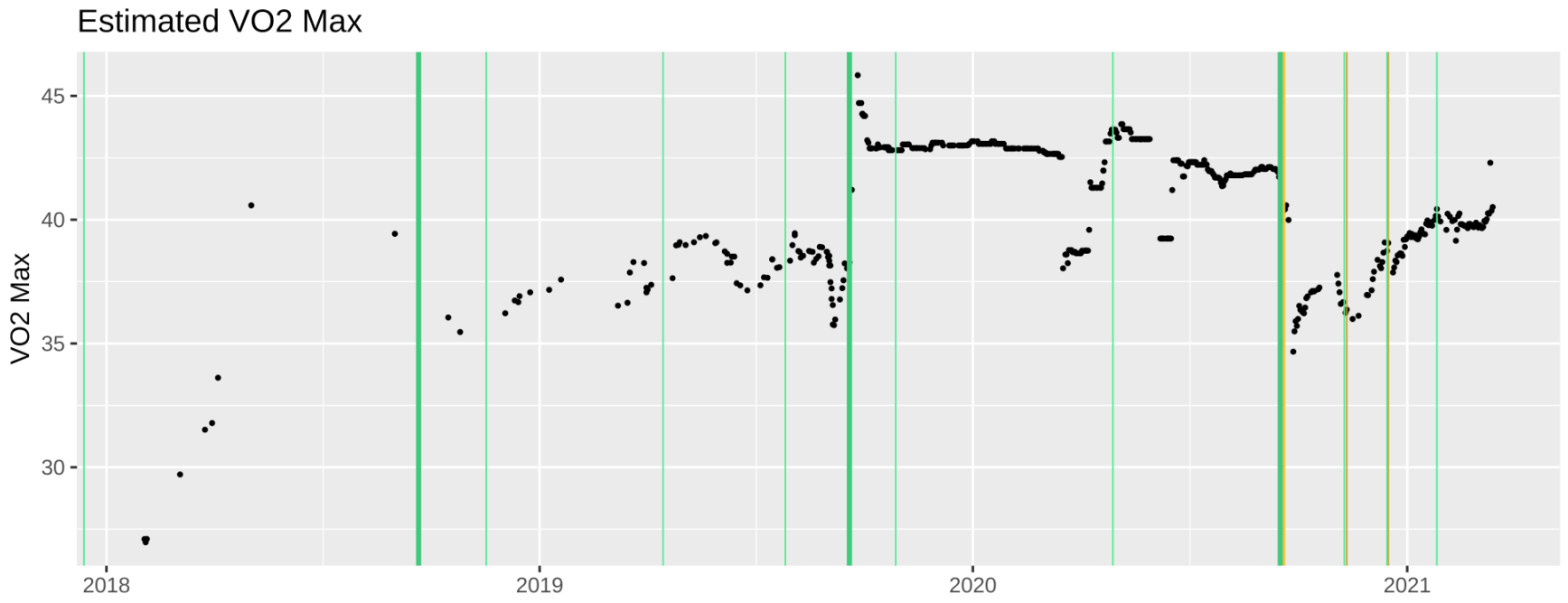
17


Resting Heart Rate by Day
resting heart rate
60
55
50
2018
2019
2020
2021

18



19



20