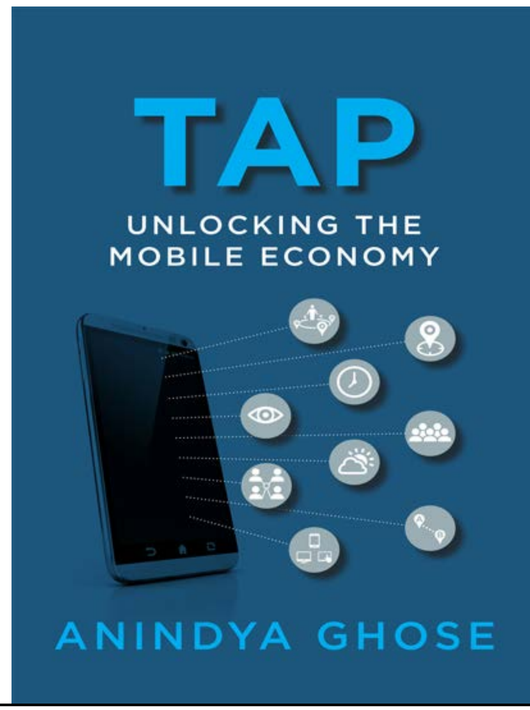
Unlocking Technology and Data Science to Combat Health Pandemics

Anindya Ghose
Heinz Riehl Chair Professor of Business
New York University



Agenda

- How tech and data were used in the early days of Covid 19?
- Insights from mobile location data
- Relevant lessons from epidemiology
- How should we open the economy?
- Privacy and data cultures

Technology For Predictive Analytics in Pandemics

- Canadian AI startup BlueDot spotted COVID-19 nine days before the WHO alerted people.
- Predicted the extent of the outbreak internationally using airline ticketing data.
- Google's DeepMind used deep learning to analyze the structure of proteins associated with COVID 19.
- Baidu used infrared and AI-powered facial recognition to screen people for fever.

Technology For Consumer Empowerment

- Alibaba and Tencent devised color coding via mobile QR codes to stratify users.
- Apps let people check if they have taken the same flight or train as Covid 19 patients.
- FaceMask App in Taiwan shows inventory of masks
- TraceTogether App in Singapore informs users about infected carriers.
- Robots performed contactless delivery of food and medicine in hospitals in lieu of nurses.

Data Science Modeling in Health Pandemics

- Significant transformation in the ability to collect massive datasets and harness them using AI and machine learning.

- Aggregate level tools include 'shelter-in-place' + 'social distancing analyses'.

  - **Italy?**

- Individual user level tools are 'active contact tracing' and 'individual quarantine analyses'.

- Geo-location data from smartphones, CCTV surveillance footage, GPS data from cars, offline credit card transactions, ATM banking records, etc.

- Successful use cases in Taiwan, China, Singapore, Israel, South Korea

Contact Tracing For Pandemic Surveillance

- Location data accuracy + tech sophistication in Covid (2020) vs. SARS/H1N1/MERS (2003-2012)

- Covid 19 transmission occurs between people who are within 1m of each other for 15 minutes or more.

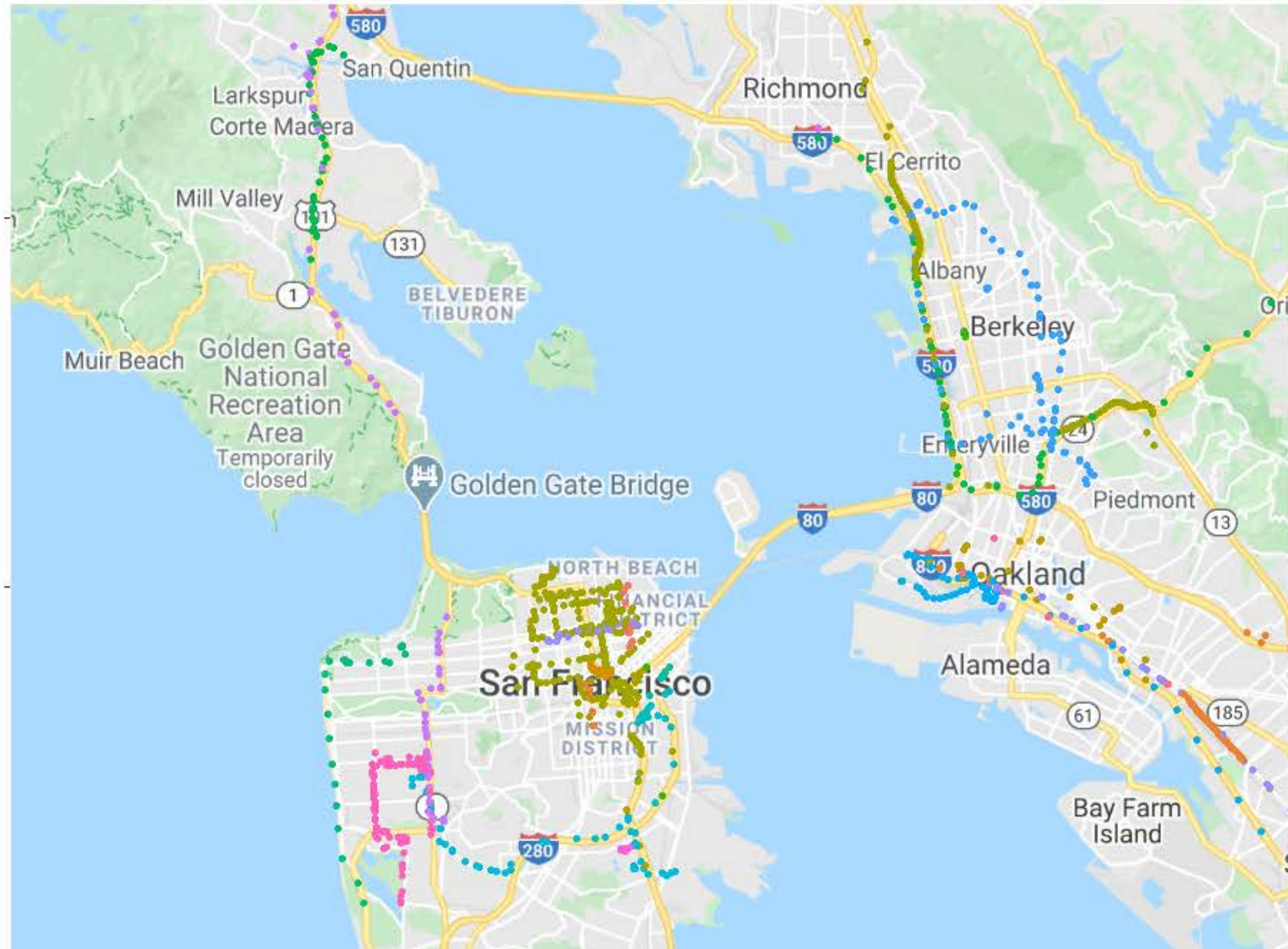
- My work with telecom providers, digital platforms, wearable tech firms, mobile app developers and mobile advertisers. (**TAP, MIT Press 2017**)

- Using machine learning techniques on atomic and granular consumers' location and trajectory data. (Ghose et al. 2019 *Management Science*)

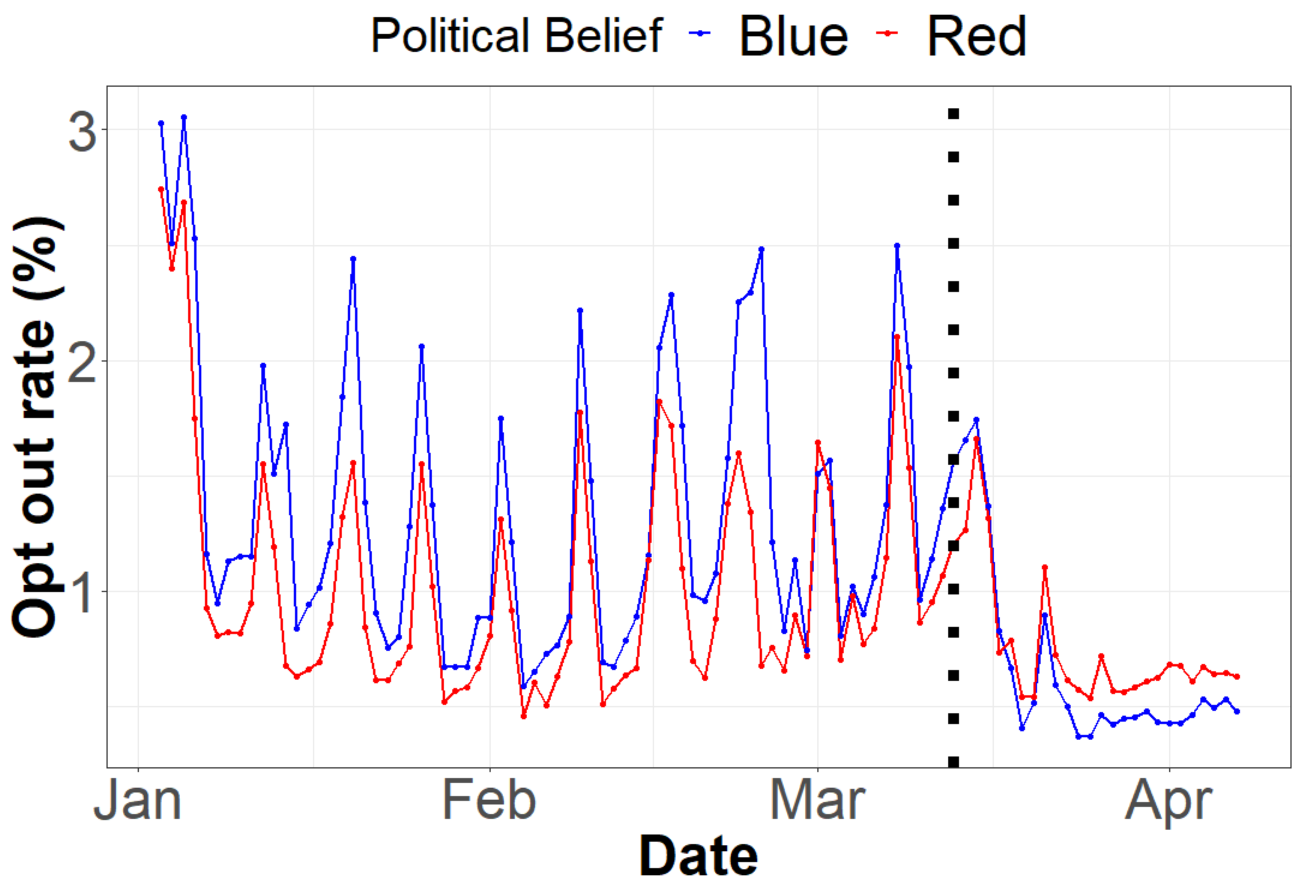
Agenda

- How tech and data were used in the early days of Covid 19?
- Insights from mobile location data
- Relevant lessons from epidemiology
- How should we open the economy?
- Privacy and data cultures

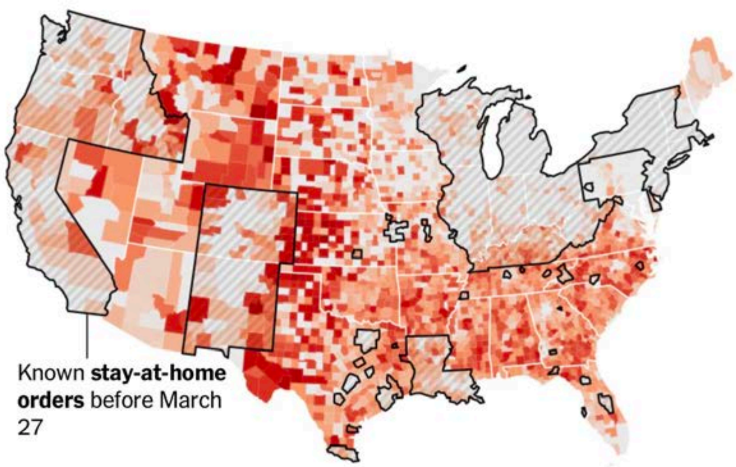
Mobile location data (Ghose, Li, Macha, Sun & Foutz 2020)



Did America Unite in Trading Privacy for Greater Social Good?



How mobile phones help measure compliance with social distancing

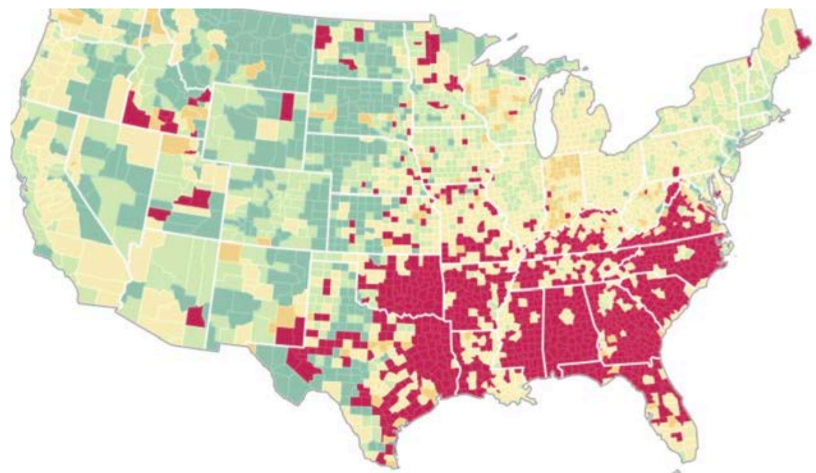


Where people were still traveling last week

Percent change in average travel for the week of March 23, compared with travel before the coronavirus outbreak.

Closer to normal travel →

No travel | Half of normal | Normal travel



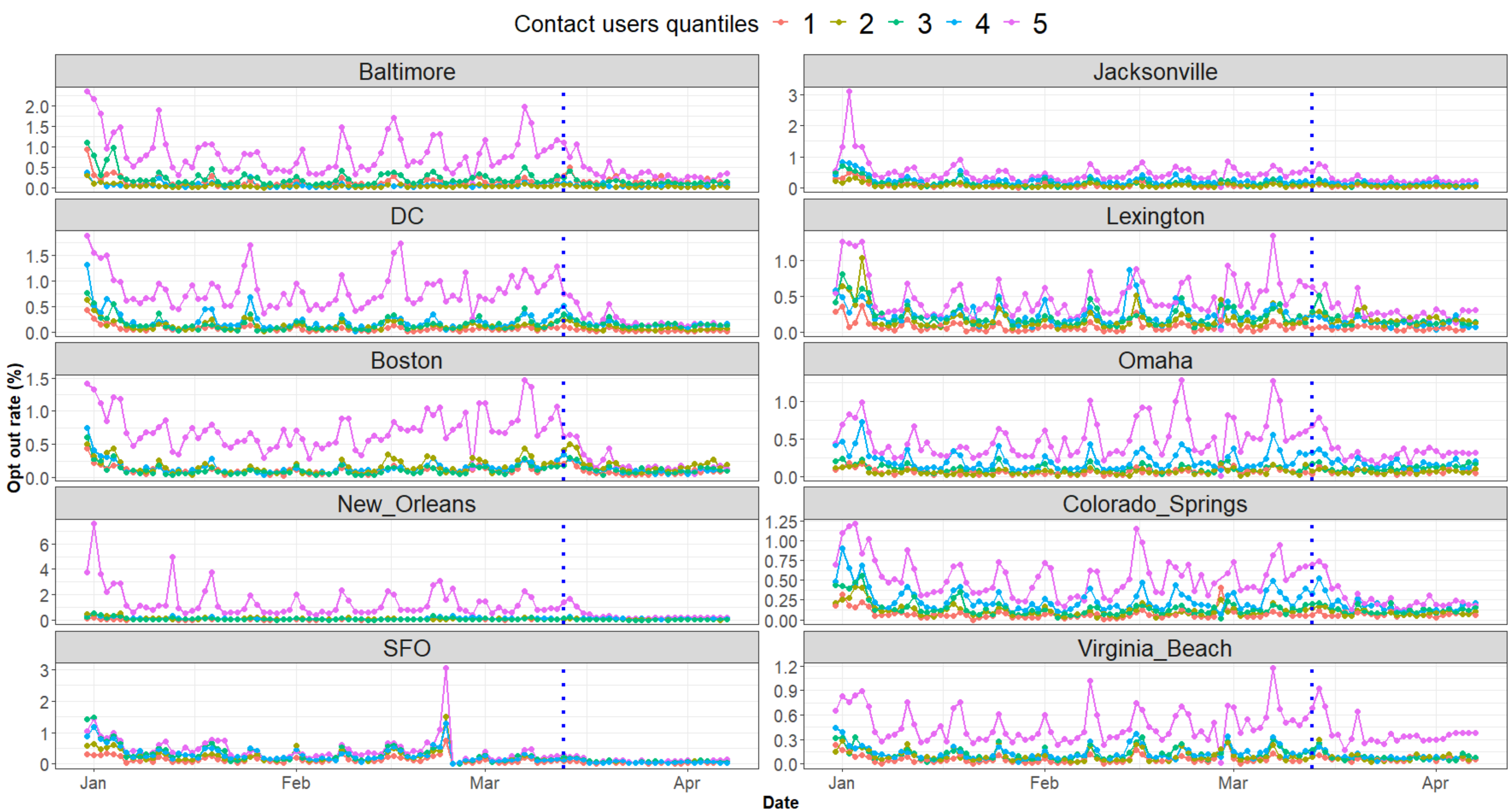
When average distance traveled first fell below 2 miles

By Mar. 16 | Mar. 19 | Mar. 24 | Mar. 26 | Not by Mar. 26

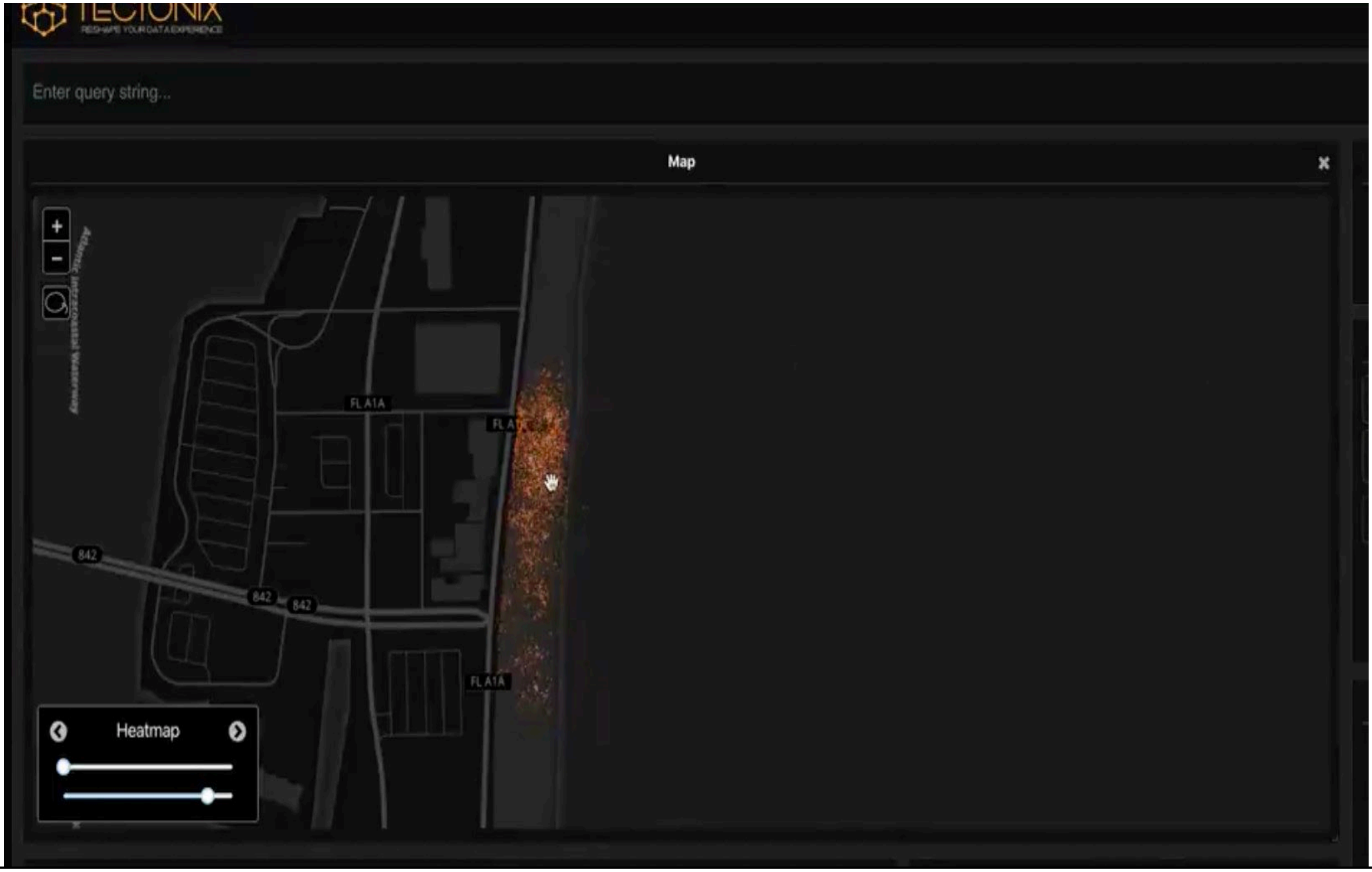
Data is through March 26. Only

@NYT April 2020

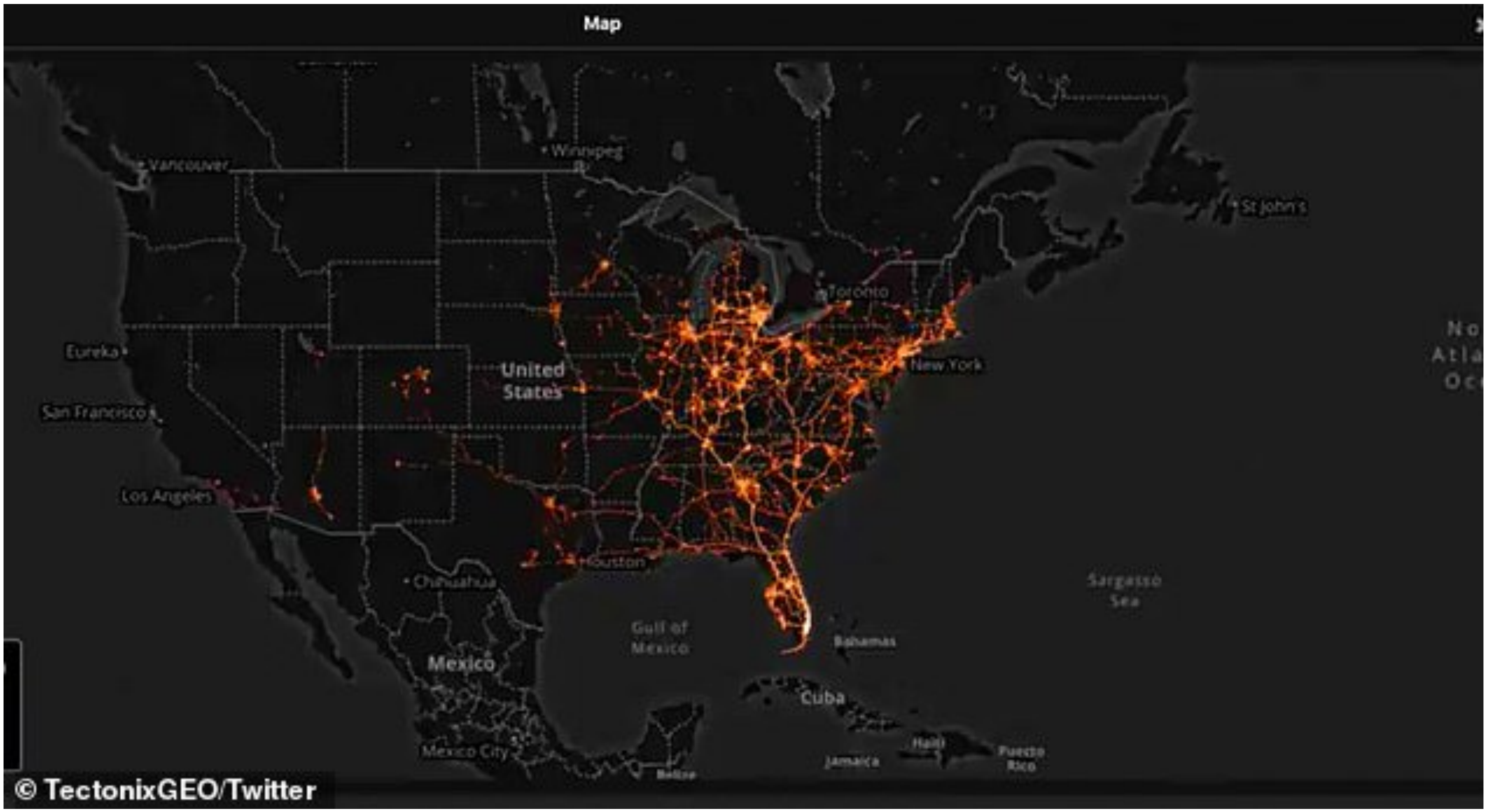
Location Tracking vs. Social Distancing

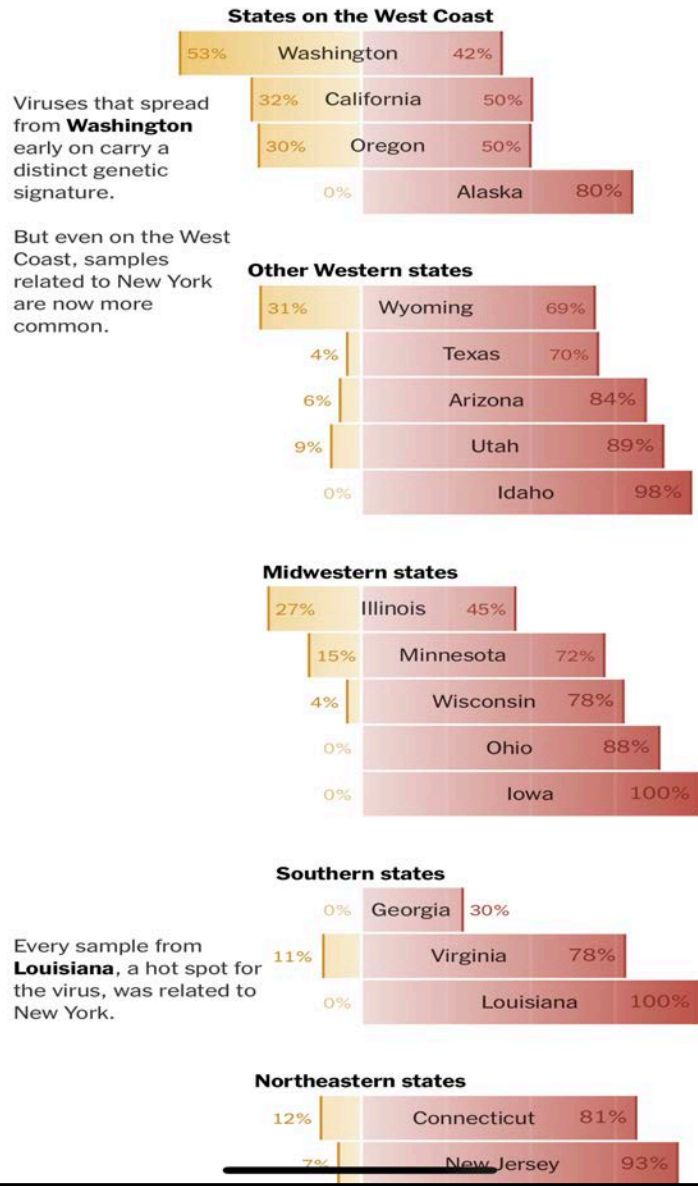


Florida 2020 Spring Break Beachgoers



Florida 2020 Spring Break Beachgoers





States on the West Coast
53% Washington 42%
32% California 50%
30% Oregon 50%
0% Alaska 80%
Viruses that spread from **Washington** early on carry a distinct genetic signature.
But even on the West Coast, samples related to New York are now more common.
Other Western states
31% Wyoming 69%
4% Texas 70%
6% Arizona 84%
9% Utah 89%
0% Idaho 98%
Midwestern states
27% Illinois 45%
15% Minnesota 72%
4% Wisconsin 78%
0% Ohio 88%
0% Iowa 100%
Southern states
0% Georgia 30%
11% Virginia 78%
0% Louisiana 100%
Every sample from **Louisiana**, a hot spot for the virus, was related to New York.
Northeastern states
12% Connecticut 81%
New Jersey 93%

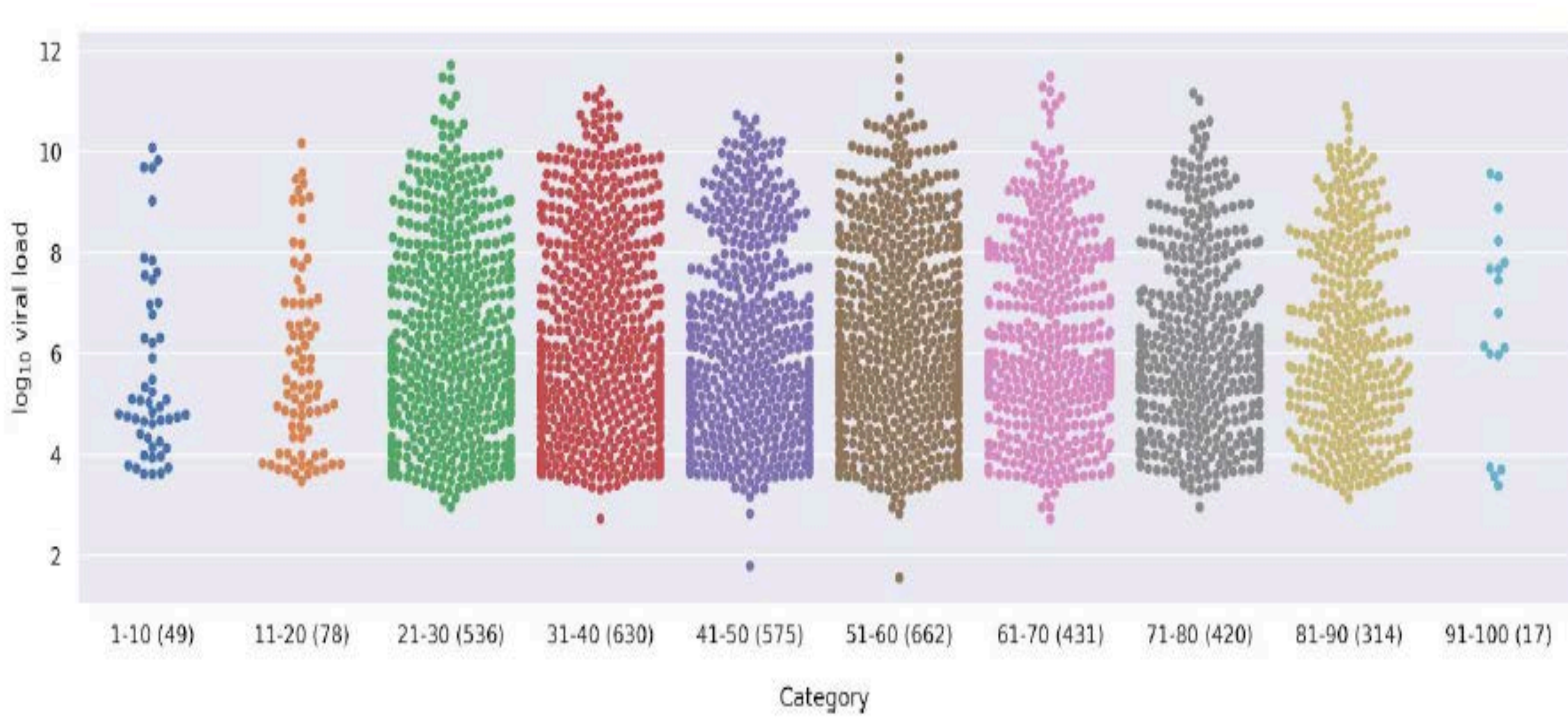
Agenda

- How tech and data were used in the early days of Covid 19?
- Insights from mobile location data
- Relevant lessons from epidemiology
- How should we open the economy?
- Privacy and data cultures

What about undetected cases?

- Three reasons for lack of detection:
    1. Asymptomatic cases
    2. Limited testing
    3. Restricted testing criteria.
- Asymptomatic users constitute between 52% and 86% of cases.
- Examples:
    - Diamond Princess cruise
    - Vo Euganeo (Italy)
    - Hubei (China).

Infectious people come in all ages, and they all shed different amounts of virus!



“Those who are dying would have died anyway?” Wrong!

	Men				Women			
Multimorbidity count	50-59	60-69	70-79	80+	50-59	60-69	70-79	80+
0	35.81	26.78	18.43	11.02	35.28	25.50	17.70	10.42
1	35.03	26.09	17.58	10.05	34.83	25.59	17.13	8.92
2	29.67	22.07	14.72	8.15	29.06	21.35	14.20	7.19
3	25.01	19.05	12.50	6.59	26.27	18.08	11.98	5.85
4	23.55	16.28	10.64	4.95	20.44	15.58	9.97	4.52
5	19.39	13.43	8.61	3.51	16.88	11.61	8.23	3.54
6	-	6.24	7.04	2.42	17.67	10.09	6.44	2.70
7	-	7.99	6.32	2.03	-	7.96	4.83	2.32
8	-	6.60	4.79	1.65	-	6.23	3.94	1.85
9	-	5.97	3.95	1.40	-	-	3.04	1.58
10	-	-	2.62	1.17	-	2.81	2.55	1.22
11	-	-	-	1.40	-	-	2.05	1.20

- Years of life lost?
- 13 years for Men
- 11 years for Women
- Hanlon P, Chadwick F, Shah A et al. April 2020.

Agenda

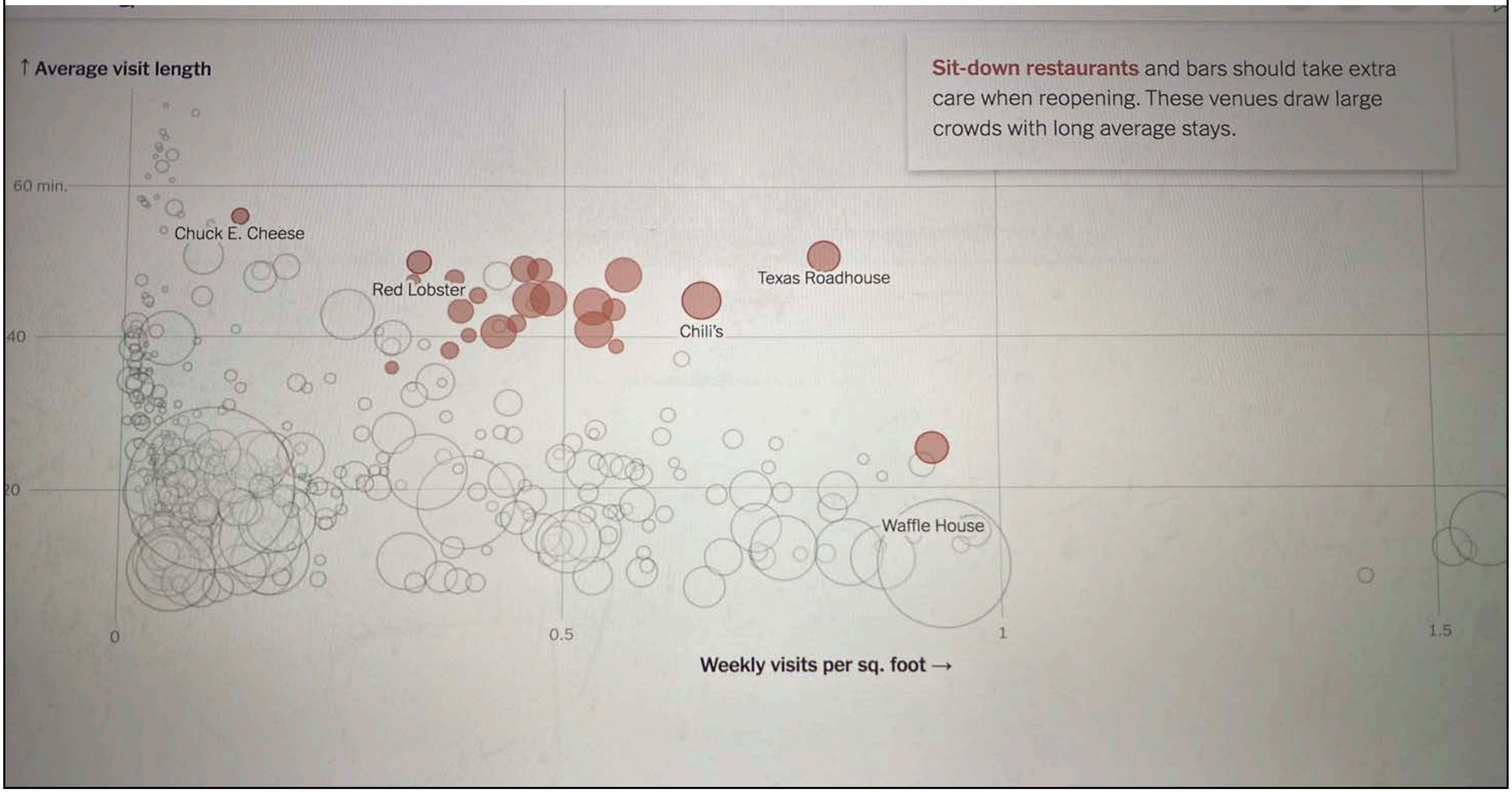
- How tech and data were used in the early days of Covid 19?
- Insights from mobile location data
- Relevant lessons from epidemiology
- How should we open the economy?
- Privacy and data cultures

How should we open the economy?

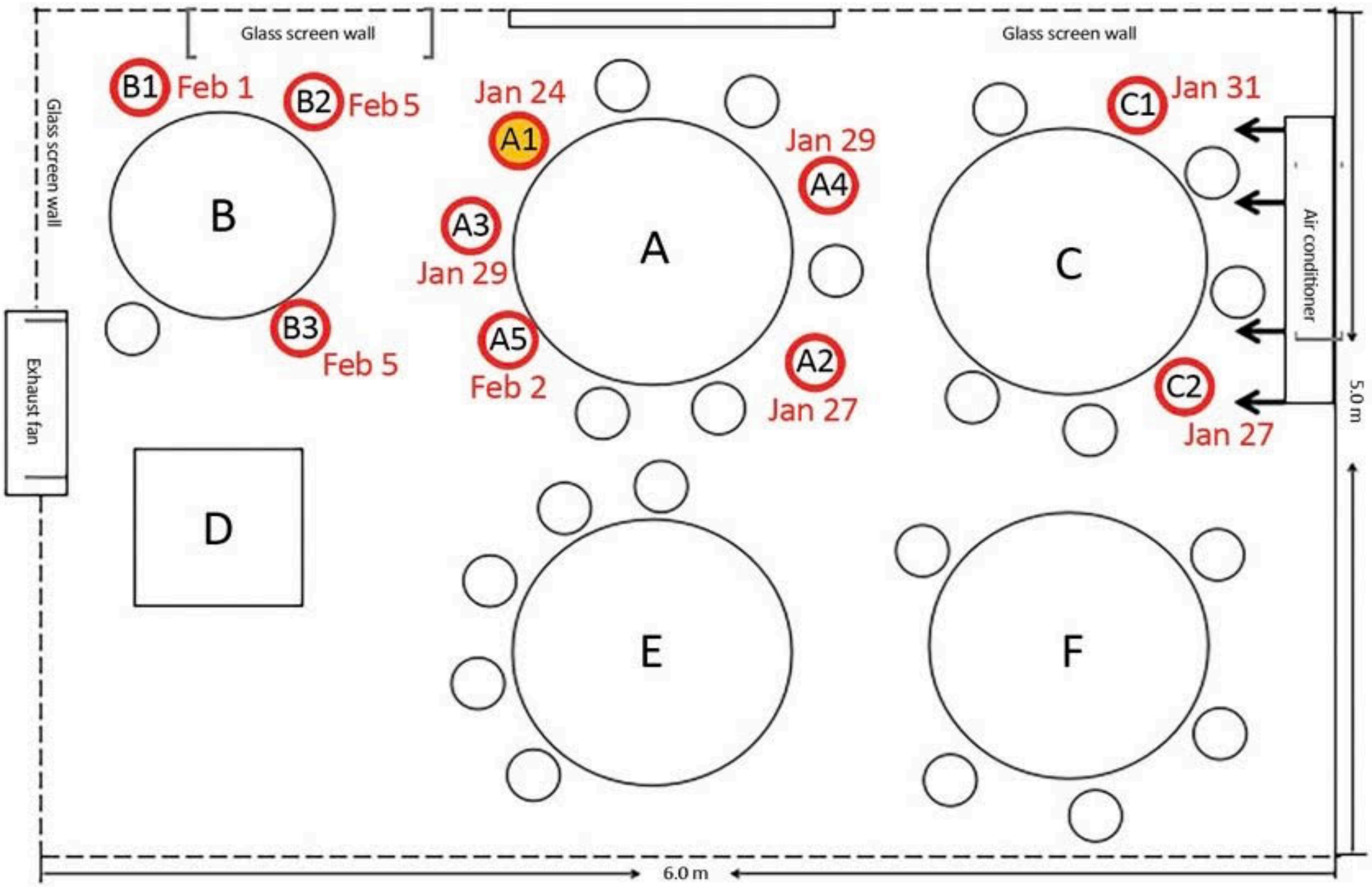
Which businesses to open first and how should we assess the risk?

Successful Infection = Exposure to Virus x Time

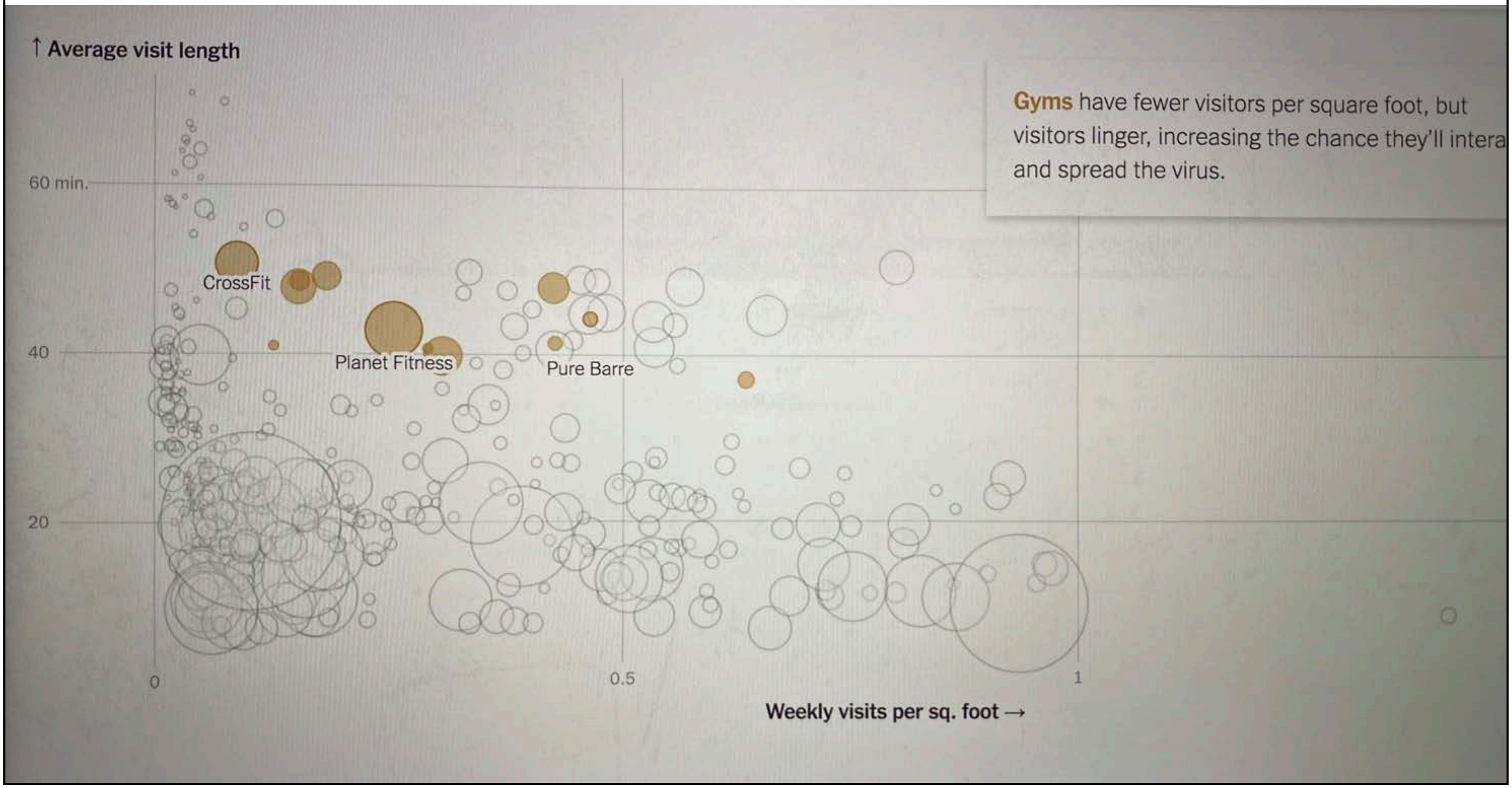
Sit down restaurants



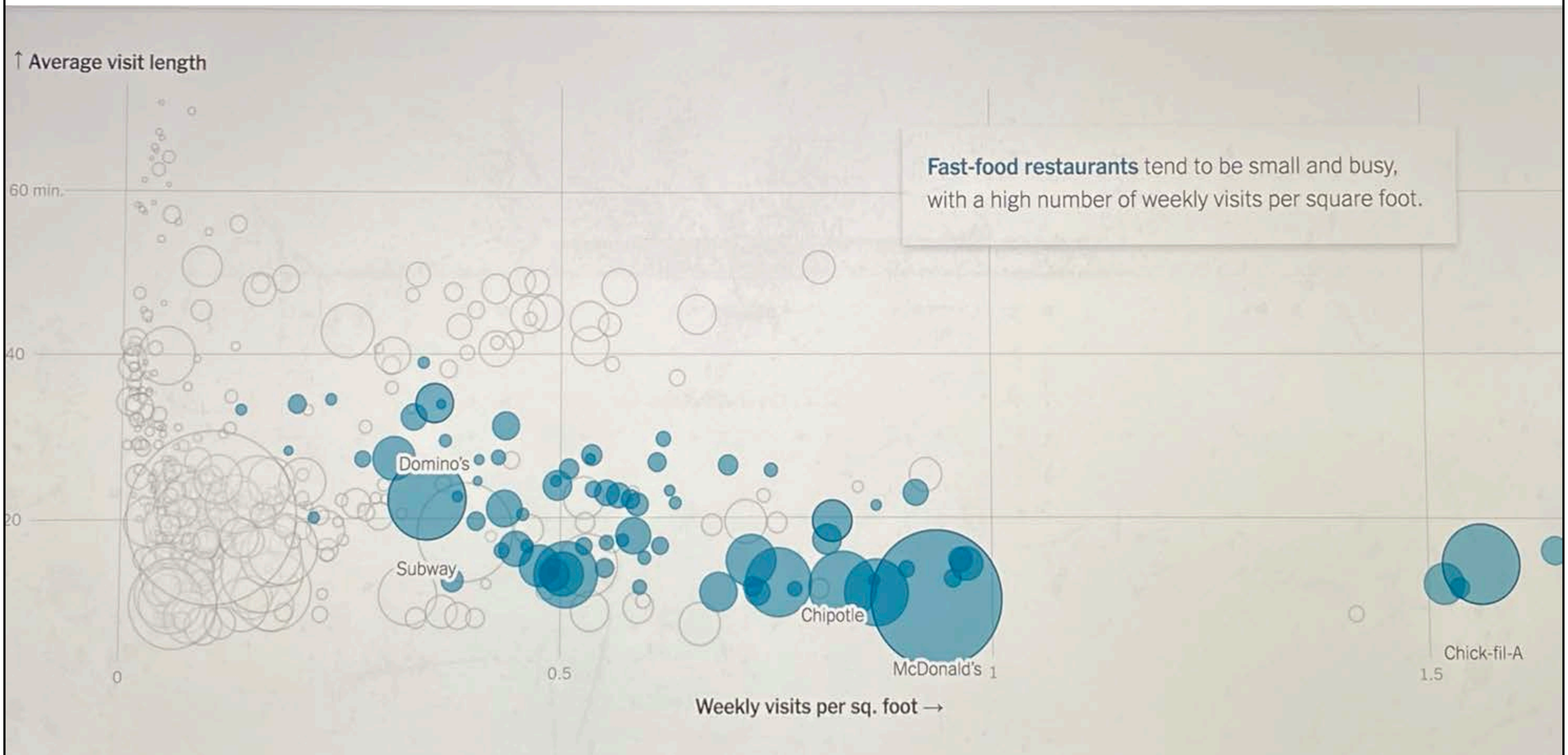
Super spreaders (Lu et al. 2020)



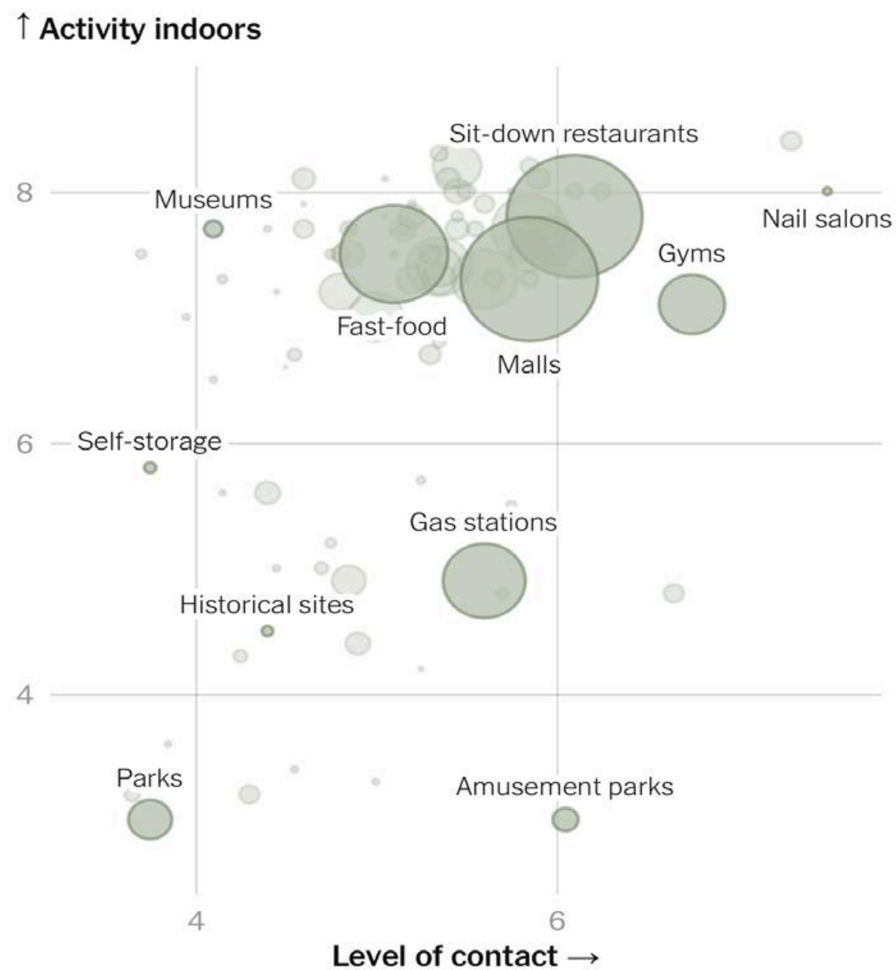
Gyms


↑ Average visit length
60 min.
40
20
0
0.5
1
Weekly visits per sq. foot →
CrossFit
Planet Fitness
Pure Barre
Gyms have fewer visitors per square foot, but visitors linger, increasing the chance they'll intera
and spread the virus.

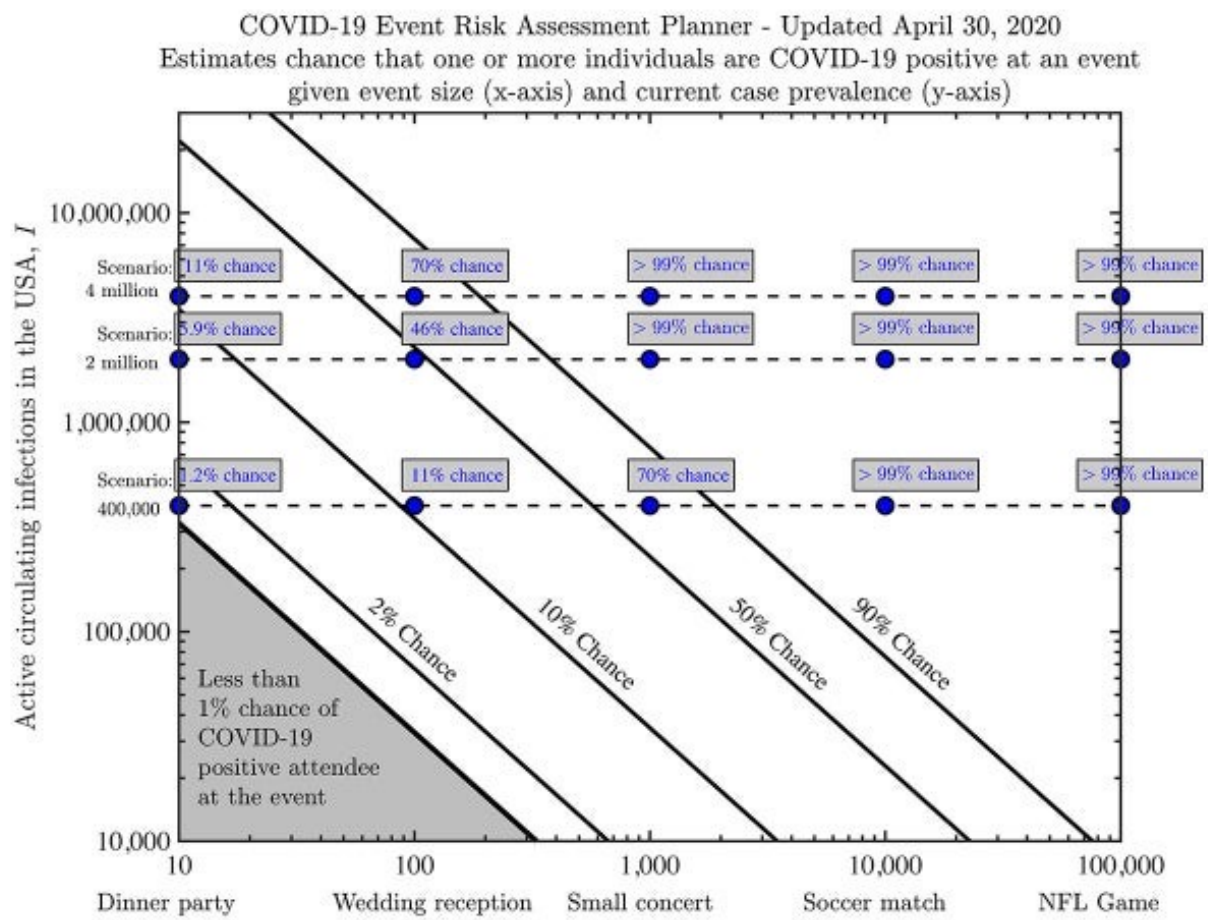
Fast food restaurants



Which businesses to patronize?



@NYT May 2020



Calculation note - J.S.Weitz - jsweitz@gatech.edu - 4/30/20 - Risk is $\epsilon \approx 1-(1-p_I)^n$ where $p_I = I/\left(330\times 10^6\right)$ and n is event size
Updated: April 30, 2020, License: Creative Commons BY-SA 4.0, i.e., Share, Adapt, Attribute
Assumes incidence homogeneity, uses last 2 weeks cumulative as baseline with 5x and 10x undercounts for alternative scenarios
Code https://github.com/jsweitz/covid-19-event-risk-planner

Agenda

- How tech and data were used in the early days of Covid 19?
- Insights from mobile location data
- Relevant lessons from epidemiology
- How should we open the economy?
- Privacy and data cultures

Regulations, Exemptions and Privacy

- To enable coordination between public and private sectors:
  - Assurances that data sharing will be exempt from any adverse regulatory action or private lawsuits.
  - Creating narrow exceptions to data sharing, just like competition laws.
  - Tech platforms are going out of their way to help us.
  - Our health privacy system was created at a time in which bioethicists worked within a 1970s framework.
  - Caveats needed such that there is zero tolerance for the misuse of data.

Four Data Cultures

- America's data model maybe too elitist
- Mainland China's data model maybe too stringent
- Germany's data protection angst maybe too technophobic
- South Korea's model maybe too collectivist for the West.
- **Best role model? Taiwan's hybrid information-collecting model (participatory self-surveillance)**

A. Kluth @Bloomberg

Thank you & Stay Safe!

Anindya Ghose
New York University

aghose@stern.nyu.edu