

Making ~~beautiful clinical~~ meaningful graphics

LSE Data Science Institute

John Burn-Murdoch | [@jburnmurdoch](#)

2 July 2021

Online

John Burn-Murdoch

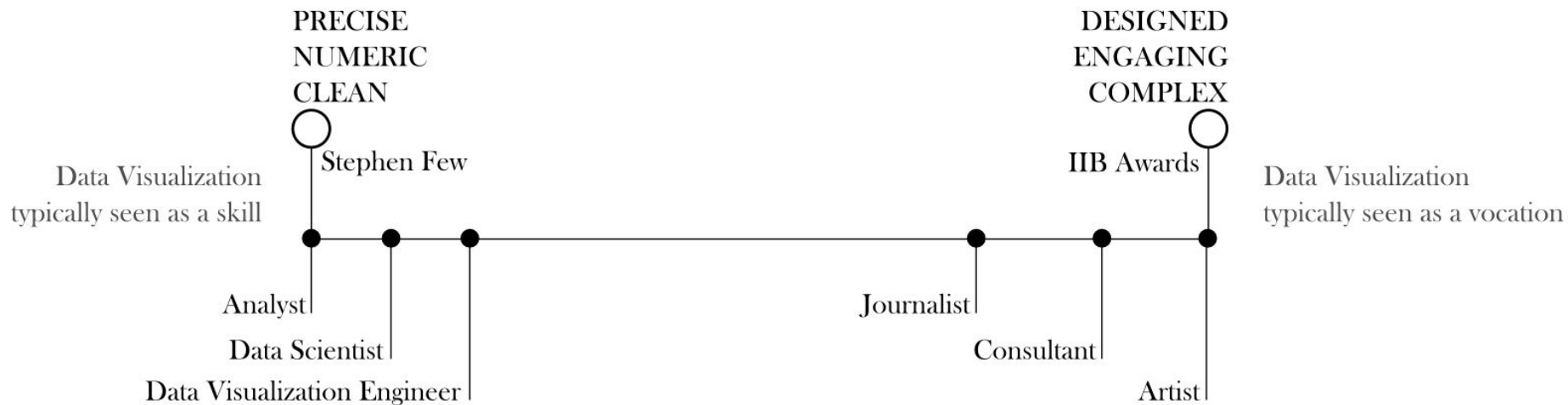
Senior data visualisation journalist

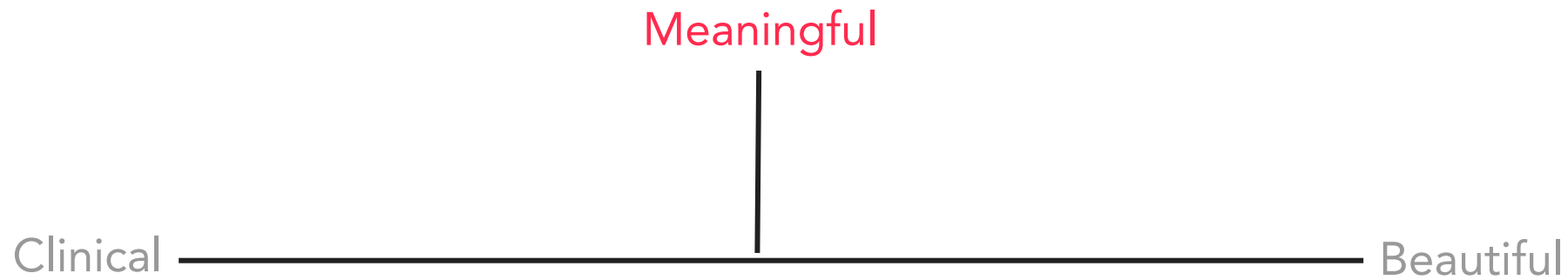
 FT Financial Times

 @jburnmurdoch

Data visualisation is a
communication skill

The data visualisation spectrum





What is effective data visualisation?

“You have to be like the worst tabloid newspaper in the front and the Academy of Science in the back.”

(Hans Rosling, 2014)

“Effective communication is getting messages across ... it suggests that we get this audience to understand something. To ensure that they understand it, we must first get them to pay attention.”

(Jean-Luc Doumont, 2009)

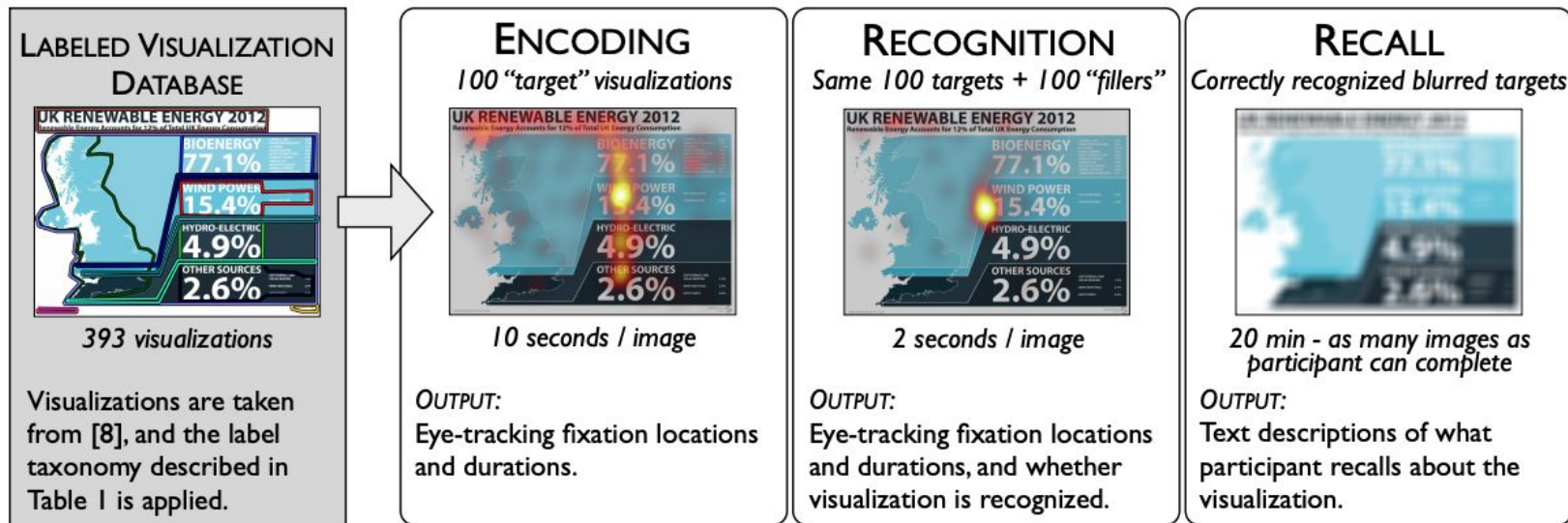
The best data visualisation paper you'll ever read

 *Beyond Memorability: Visualization Recognition and Recall*

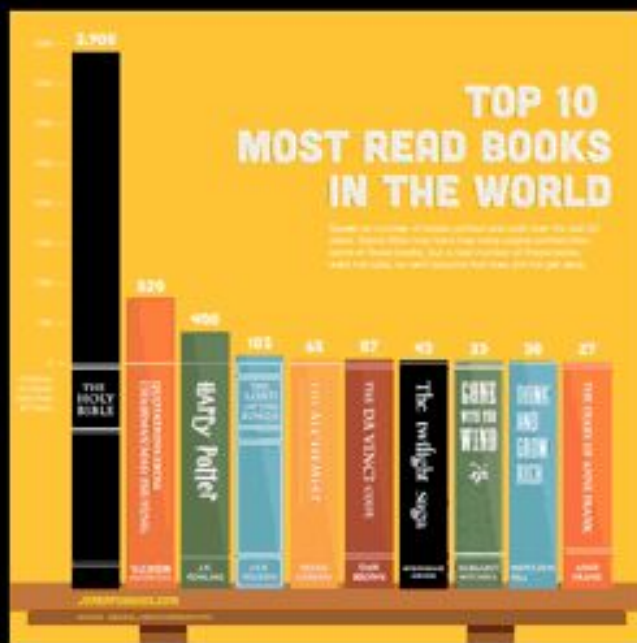
Beyond Memorability: Visualization Recognition and Recall

Michelle A. Borkin*, *Member, IEEE*, Zoya Bylinskii*, Nam Wook Kim, Constance May Bainbridge, Chelsea S. Yeh, Daniel Borkin, Hanspeter Pfister, *Senior Member, IEEE*, and Aude Oliva

EXPERIMENT DESIGN



STEP 1: Encoding

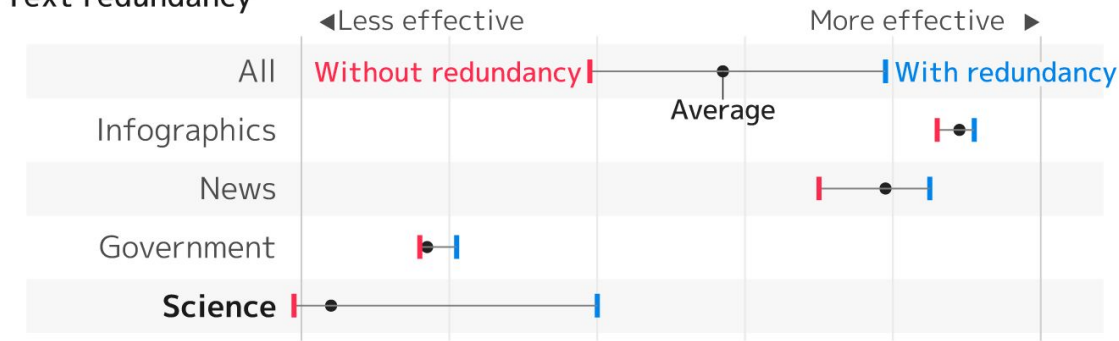


10 SECONDS

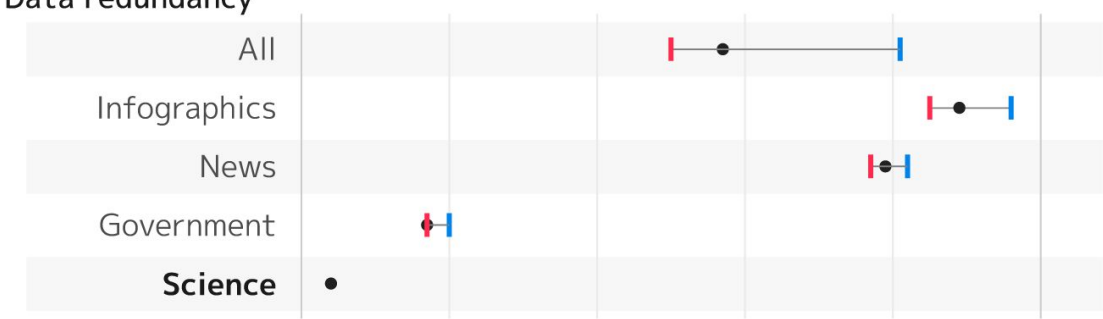
Graphics are more effective when they use redundancy in data or text to repeat the core message

Effectiveness of visualisation, by level of text and data redundancy

Text redundancy

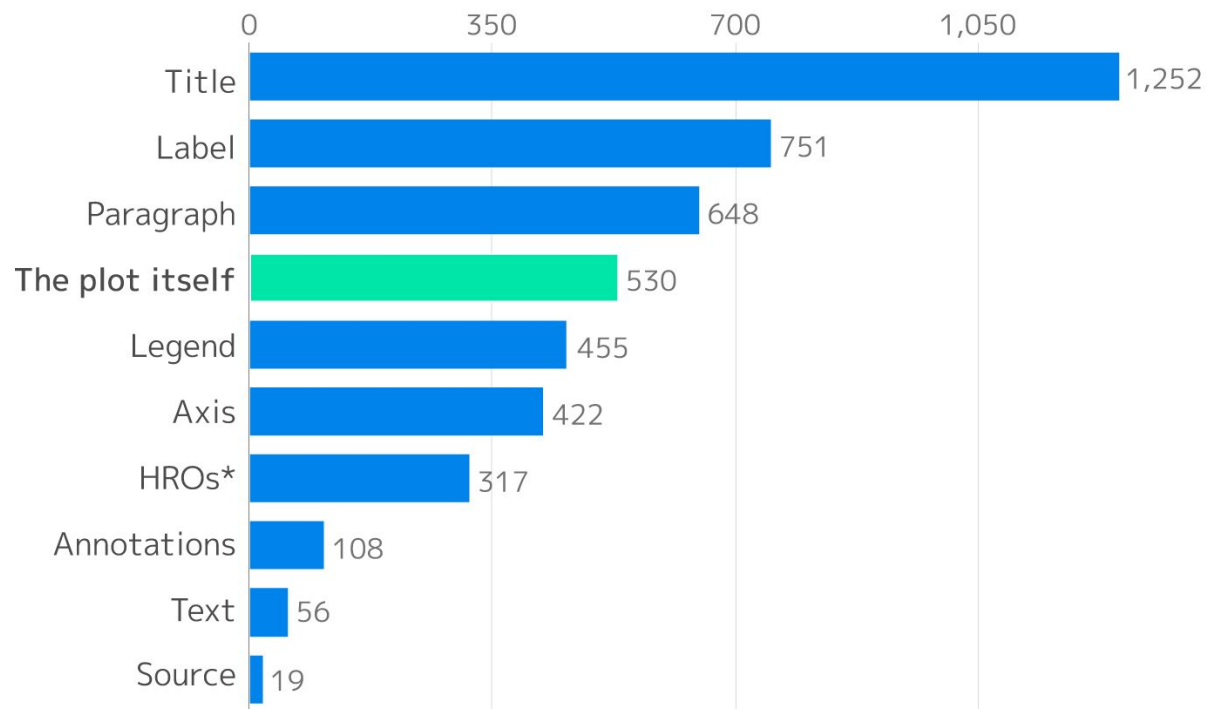


Data redundancy



Text elements are mentioned most often in descriptions of graphics

Total number of mentions of different graphical elements



*Human-recognisable objects, e.g images of people, places, animals

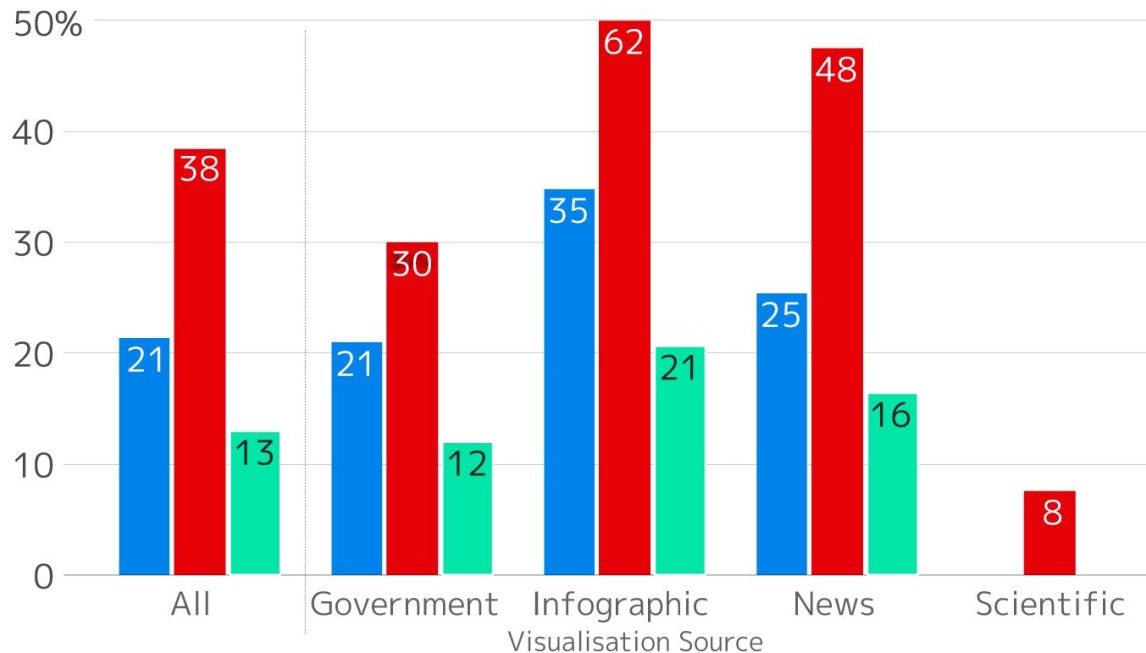
Source: Borkin, M et al., Beyond Memorability: Visualization Recognition and Recall

Graphic: @jburnmurdoch

Scientific visualisations very rarely exploit redundancy to emphasise their point

Share of graphics with data and or text redundancy

■ Data ■ Text ■ Both



Source: Borkin, M et al., Beyond Memorability: Visualization Recognition and Recall

Graphic: @jburnmurdoch

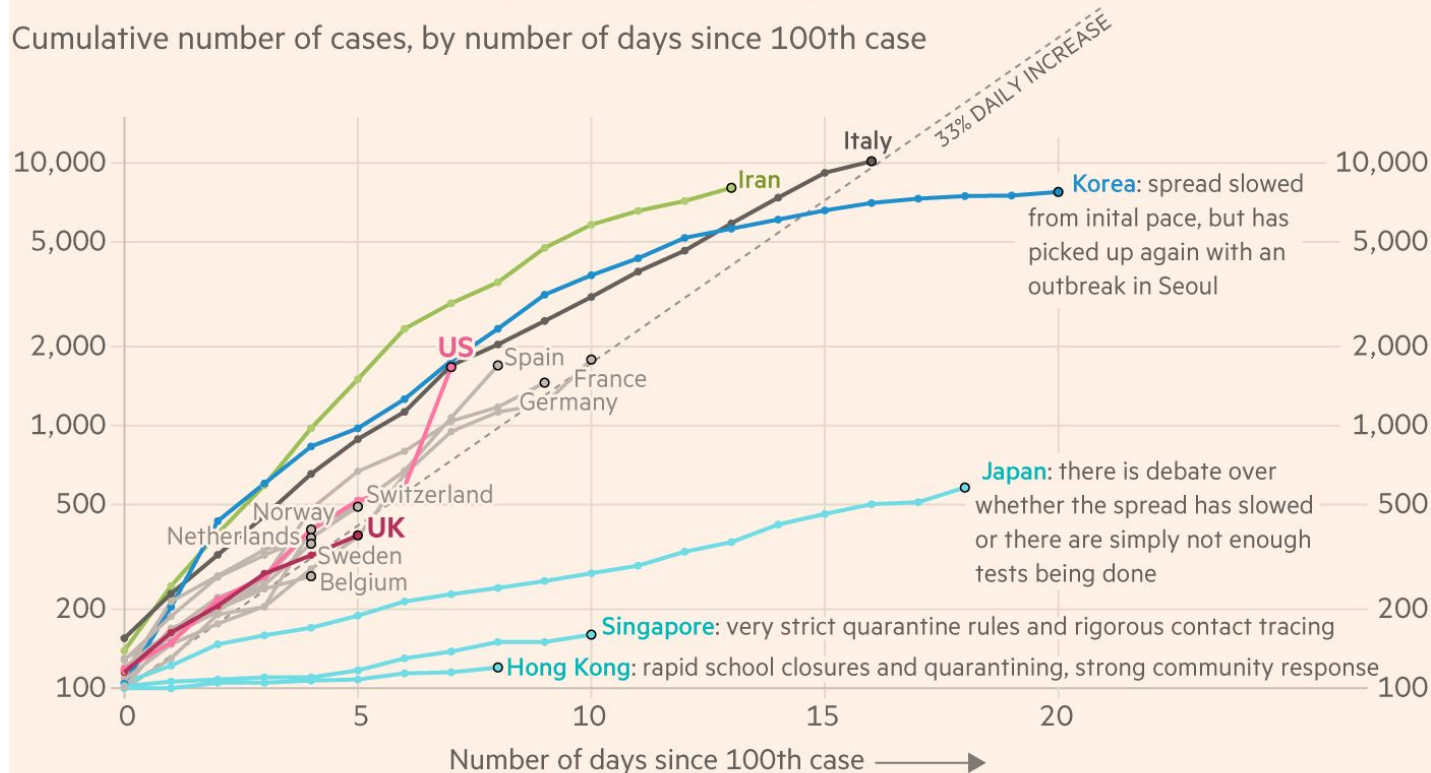
Guidelines

- Narrative titles. Make a point. Own the message
- Say it twice
- Exploit colour
- Minimum legible font size: $\text{graphic width} / 20$

In action

Most Western countries are on the same coronavirus trajectory. Hong Kong and Singapore have managed to slow the spread

Cumulative number of cases, by number of days since 100th case



Source: FT analysis of Johns Hopkins University, CSSE

FT graphic: John Burn-Murdoch / @jburnmurdoch

© FT

Table 1: **2010 Census Backward Projection**

Year	Age	Male	Female	Sex ratio	Expected at 105 SRB	Missing females
2010	0	7,461,199	6,325,235	118	7,105,904	780,669
2009	1	8,574,973	7,082,982	121	8,166,641	1,083,659
2008	2	8,507,697	7,109,678	120	8,102,569	992,891
2007	3	8,272,491	6,978,314	119	7,878,563	900,249
2006	4	8,246,206	6,973,835	118	7,853,530	879,695
2005	5	7,988,151	6,743,986	118	7,607,763	863,777
2004	6	8,034,452	6,770,018	119	7,651,859	881,841
2003	7	7,292,300	6,136,861	119	6,945,048	808,187
2002	8	7,423,559	6,243,397	119	7,070,056	826,659
2001	9	7,726,203	6,522,622	118	7,358,289	835,667
2000	10	7,830,808	6,623,549	118	7,457,912	834,363
1999	11	7,522,558	6,413,156	117	7,164,341	751,185
1998	12	8,288,987	7,110,572	117	7,894,273	783,701
1997	13	8,161,000	7,064,032	116	7,772,381	708,349
1996	14	8,463,924	7,429,876	114	8,060,880	631,004
1995	15	9,524,898	8,499,586	112	9,071,331	571,745
1994	16	9,795,181	8,995,340	109	9,328,744	333,404
1993	17	10,760,828	10,014,541	107	10,248,408	233,867
1992	18	10,744,556	10,010,718	107	10,232,910	222,192
1991	19	11,079,367	10,464,099	106	10,551,778	87,679
1990	20	14,201,091	13,825,863	103	13,524,849	-301,014
Total		185,900,429	163,338,260	115	177,048,028	13,709,768

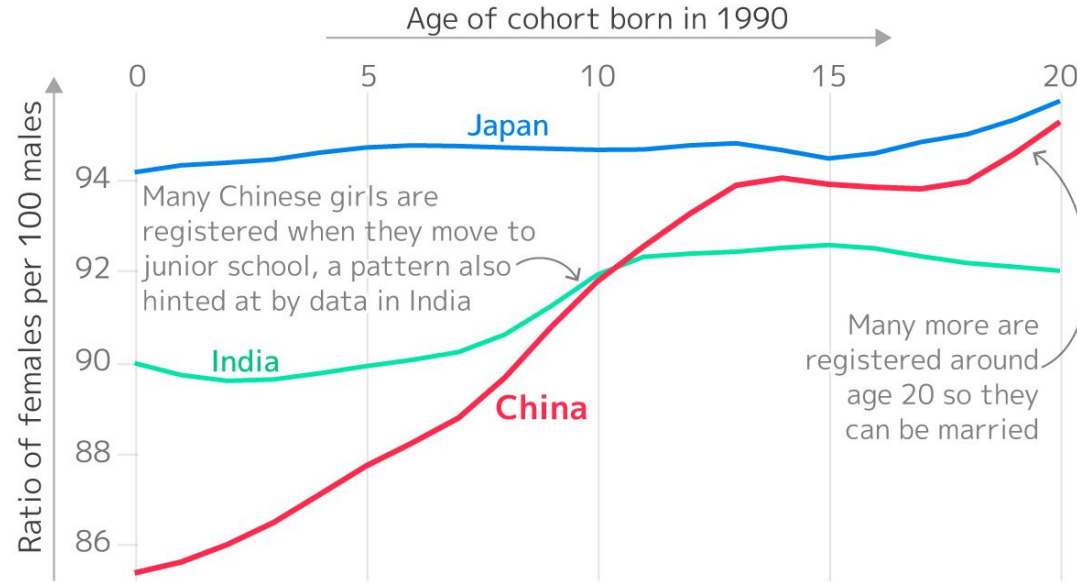
Source:

NBS [2011](#).

China's missing baby girls are re-appearing in the census years after they were born

Research suggests that rather than sex-selective abortion or infanticide, the leading cause of China's male-heavy sex ratios at birth is parents delaying official registration of their daughters

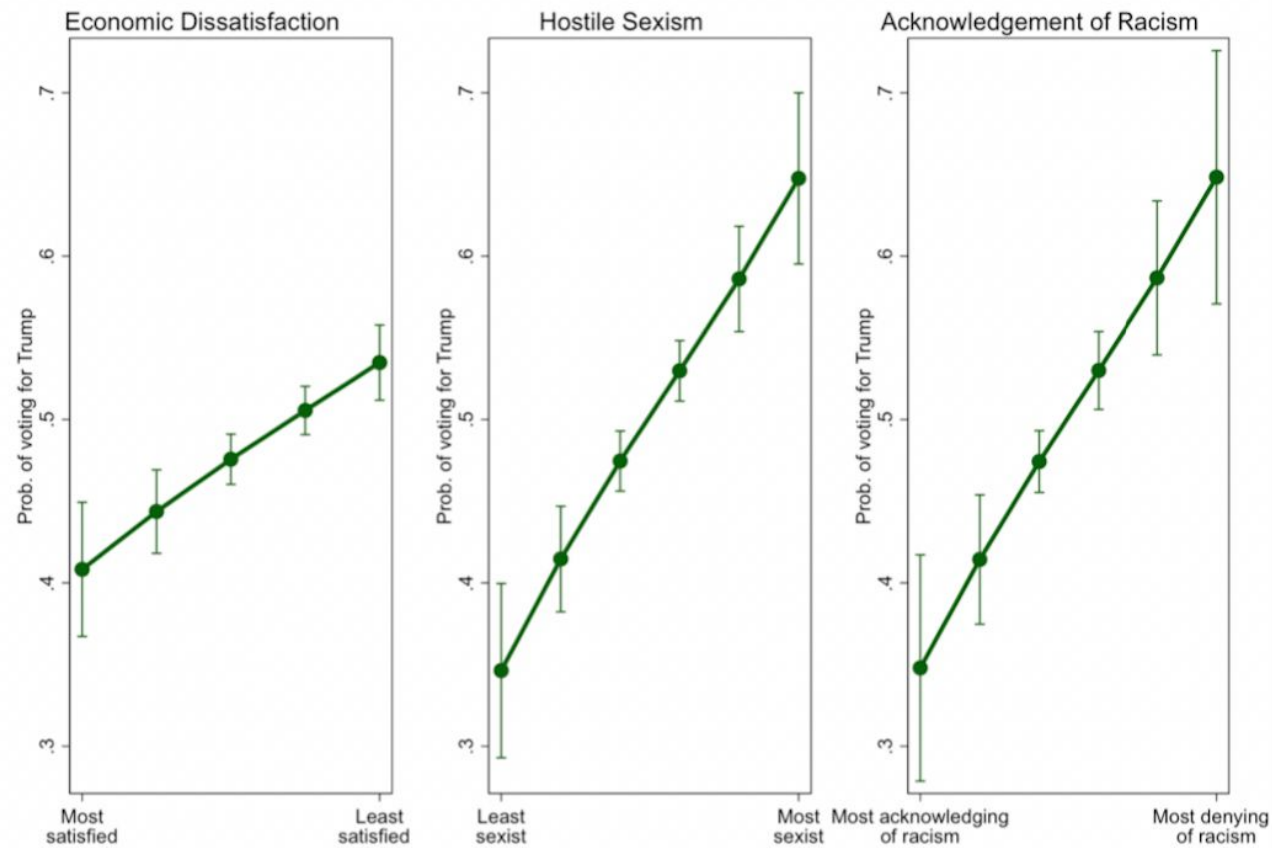
Ratio of females to males, by age



Sources: United Nations Population Division, Shi and Kennedy (2016)

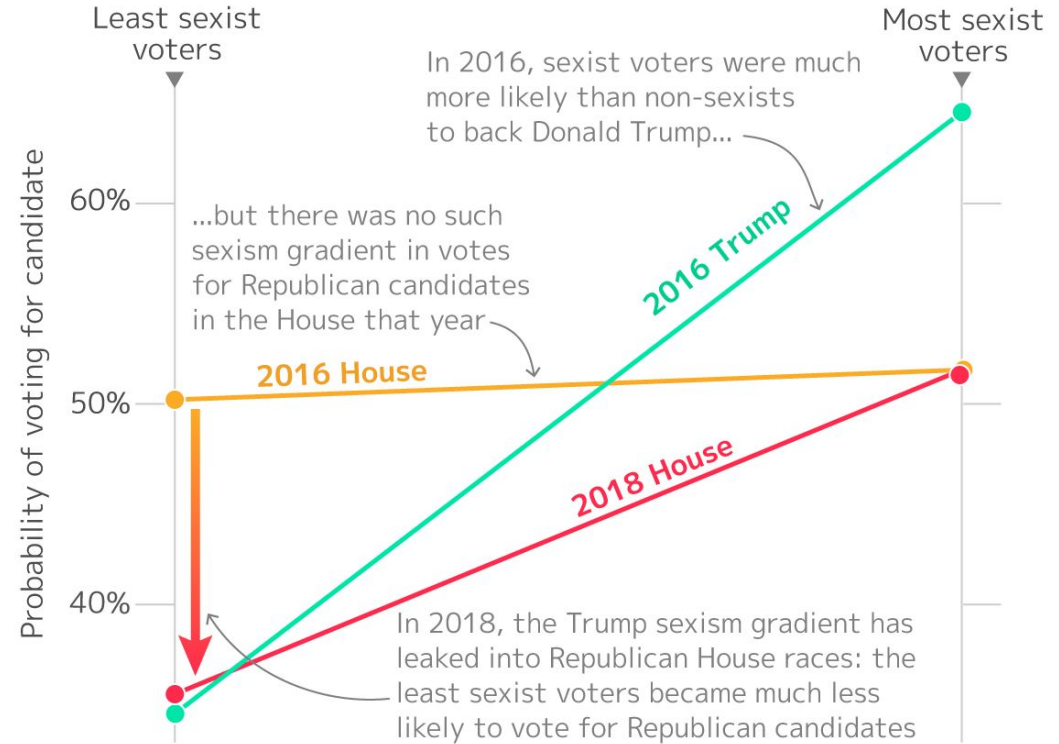
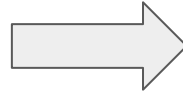
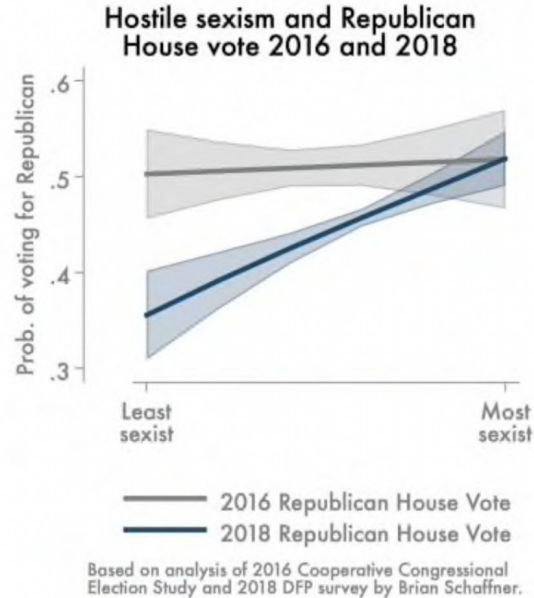
Graphic: @jburnmurdoch

Figure 3: Predicted probability of voting for Trump based on values of economic dissatisfaction, racism, and sexism



Trump's hostility towards women has tarnished the Republican party, costing votes in the 2018 midterms

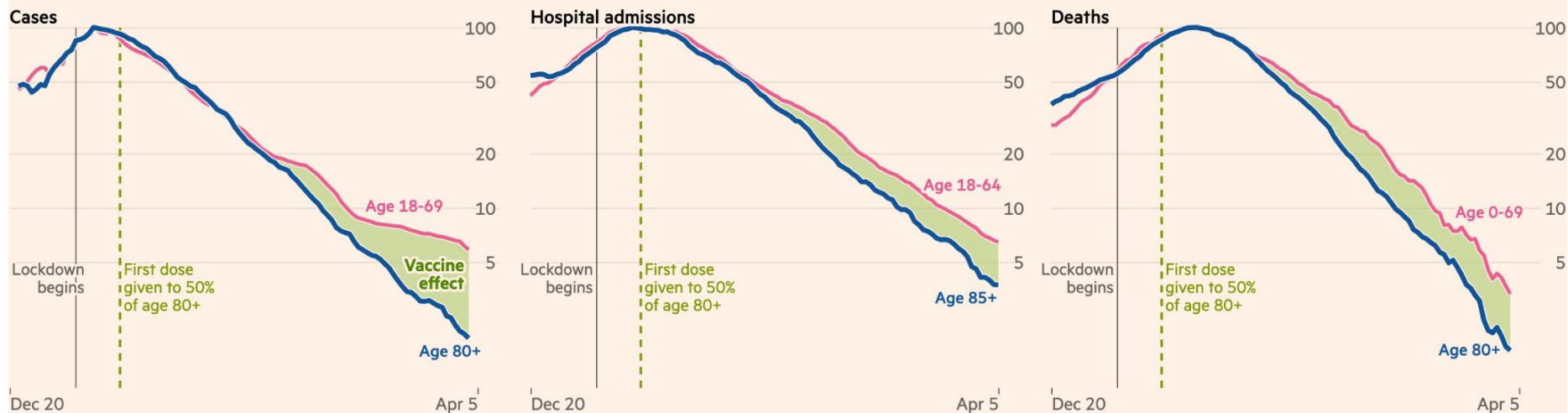
Probability of voting for Republican candidate, by level of sexism



Source: Brian Schaffner's analysis of Cooperative Congressional Election Study
Graphic: @jburnmurdoch

The UK's vaccine effect: cases, hospital admissions and deaths are now falling much faster in older groups than younger ones

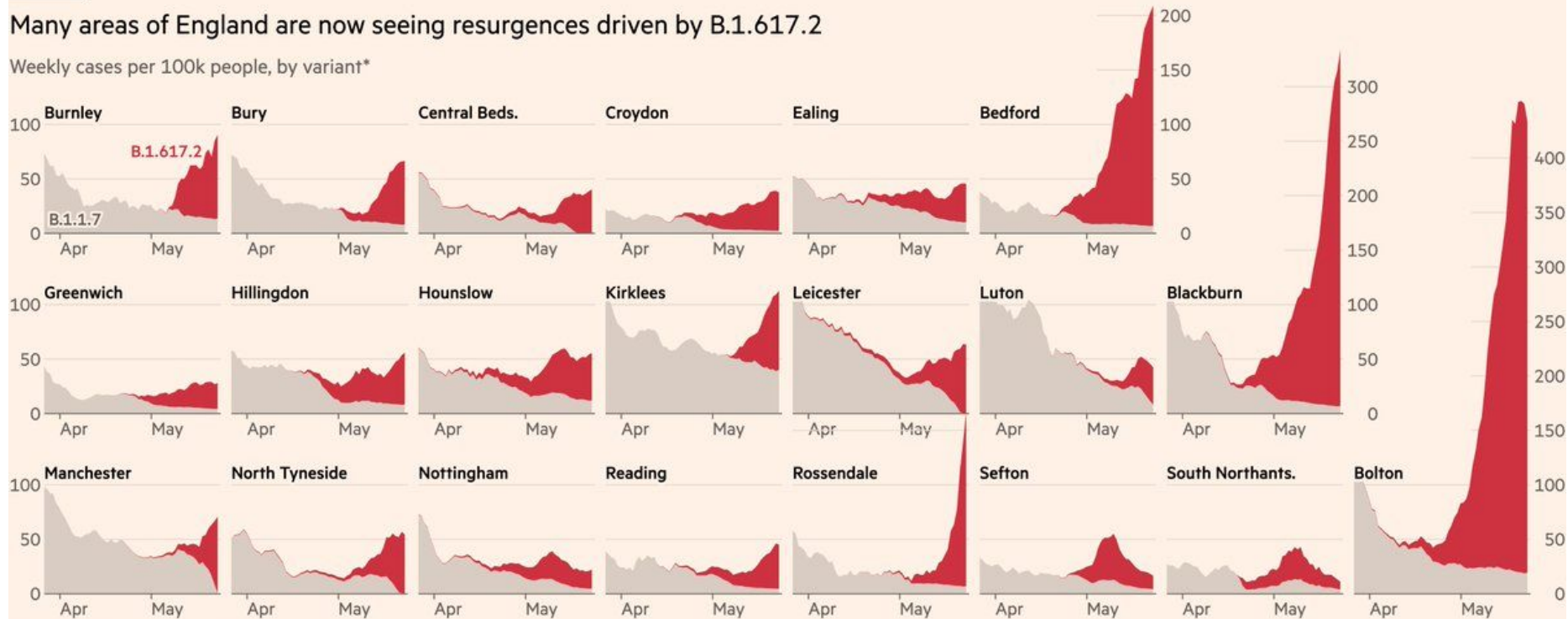
Cases, hospital admissions and deaths as a percentage of winter peak, by age group (log scale)



Sources: UK government Covid-19 dashboard
© FT

Many areas of England are now seeing resurgences driven by B.1.617.2

Weekly cases per 100k people, by variant*



*Based on applying proportions of sequenced samples to total cases

Sources: FT analysis of data from the Sanger Institute and UK government Covid-19 dashboard

© FT

Making ~~beautiful clinical~~ meaningful graphics

LSE Data Science Institute

John Burn-Murdoch | [@jburnmurdoch](#)

2 July 2021

Online