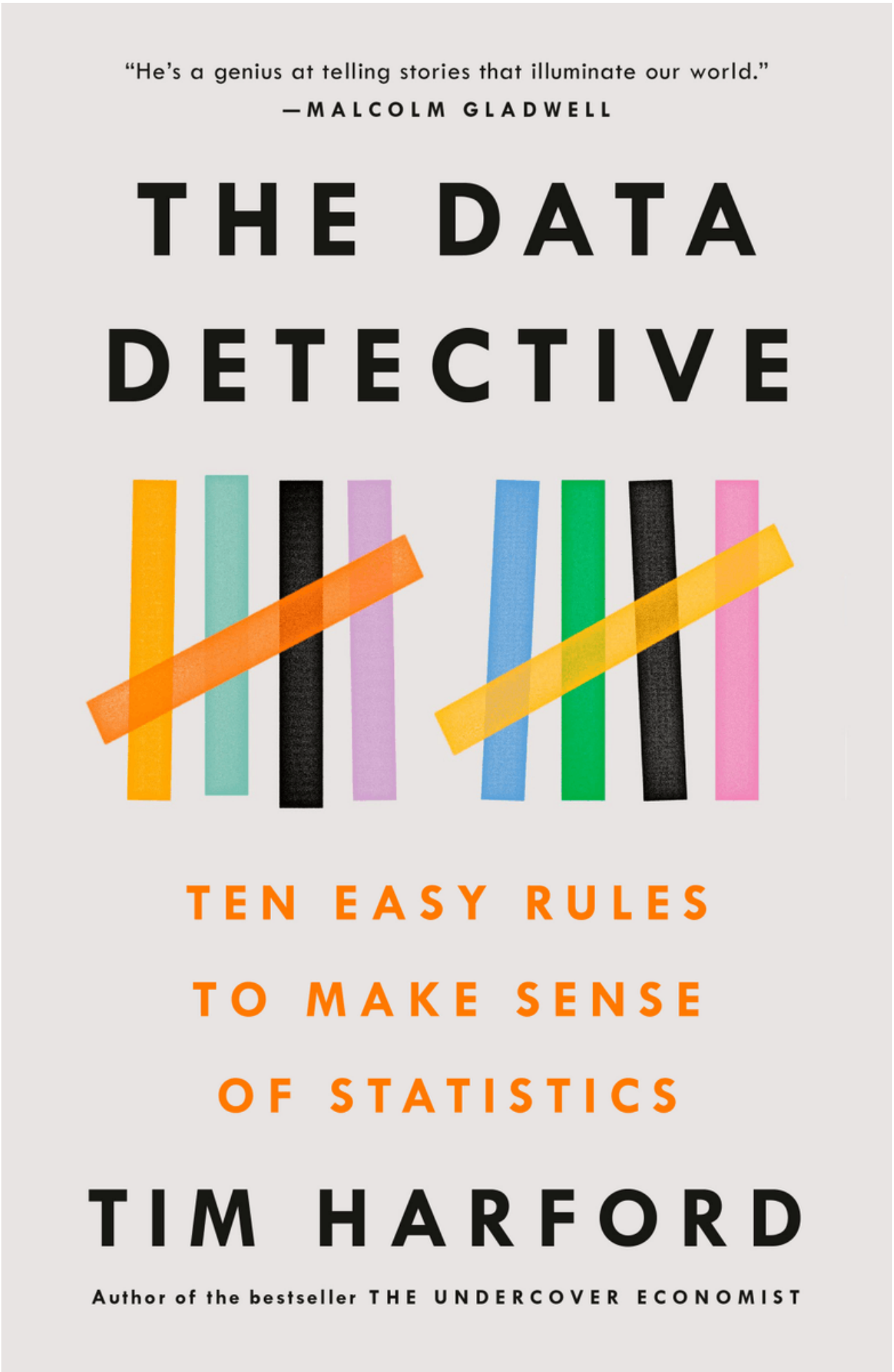
Twitter Thread by Cory Doctorow #BLM

Cory Doctorow #BLM
@doctorow

Publishing works on long schedules, which means that long-planned books can be overtaken by events...like covid.

2020 was tough for those of us with books in trail, especially nonfiction. But for a few lucky writers, covid imparted a terrible salience to their books.

1/



One such writer is @TimHarford, host of @BBCRadio4's @bbcmoreorless, which is hands-down the greatest statistical literacy program in the world, using the numbers in each week's headlines to impart statistical lessons and render the news

in perspective.

2/

Harford's latest book is THE DATA DETECTIVE (published as HOW TO MAKE THE WORLD ADD UP in the UK), which should really have been entitled HOW TO TRUTH WITH STATISTICS.

https://t.co/09AcSoFsyv

3/

You've likely heard of Darrell Huff's 1954 classic HOW TO LIE WITH STATISTICS, a classic, accessible guide to statistical malpractice in service to pushing an agenda. What I didn't know (until I heard Harford discourse on it) was that Huff himself had an ulterior motive.

4/

Huff was a shill for the tobacco industry, which, in 1954, was deliberately muddying the waters about the link between cancer and smoking. Huff's book isn't just a guide to statistical malpractice - it's a broadside against cancer science.

https://t.co/lurIujpAHd

5/

(Huff's followup volume, HOW TO LIE WITH CANCER STATISTICS, was never published)

Huff and his paymasters at Big Tobacco created the scientific doubt playbook - a system that weaponizes critical thinking by demonizing inconvenient science as junk science.

6/

Superficially, Harford's work - especially More or Less - is a critique of bad stats (certainly, it's grimly fun to listen to Harford and his colleagues reveal the bad math behind the UK government's covid spin), but as this book demonstrates, there's more going on there.

7/

What Harford and More or Less strive for isn't merely debunking - it's discovering truth. His brand of statistical sleuthing isn't merely a critique, it's a fact-finding mission, a campaign to find out, to the greatest possible extent, what is actually happening.

8/

That's where Harford and Huff's approaches diverge. Huff's book is a sharp critical toolsuite, but it's also a counsel of despair. Huff tells you how to lie with statistics - but he never tells you how to use statistics to discover the truth.

9/

To read Huff - and his successors in various denial and doubt movements - is to conclude that the world is unknowable. It is nihilism packaged up as mathematics.

Enter THE DATA DETECTIVE: a book about statistical best practice AND malpractice.

10/

DATA DETECTIVE proffers ten rules for getting stats right:

I. Vaccinate yourself against motivated reasoning by examining your emotional reaction to any statistical finding; use this check-in to adjust your intuition about whether a claim is true or false

11/

II. Neither the quantitative statistical picture ("bird's eye view") nor the qualitative lived experience ("worm's eye view") is sufficient: to understand a stat, you need both

12/

III. Check carefully to determine what a statistical label refers to (for example, in stories about links between "violent behavior and violent video games," how are "violent behavior" and "violent games" defined?)

13/

IV. Get "comparisons and context" - numbers with a lot of zeroes after them SOUND big, but are they? For example, if a bridge will cost $10m to build, is that an expensive bridge, or a cheap one?

14/

V. Understand survivor bias: if someone tells you that "every successful business does X," find out whether every failed business also does X.

15/

VI. Investigate statistical measurements for what's NOT being measured - are mortality stats broken out by age? Sex? Was a study conducted on a representative random sample, or just a bunch of grad students whose prof ordered them to fill in a survey?

16/

VII. Theory-free machine-learning algorithms have specific, pernicious, and possibly irredeemable failure modes. Between bad training data, spurious correlations, and opacity, they are a source of continuous mischief.

17/

VIII. Understand the "bedrock statistics" - censuses and other important, widely relied-upon measures, generally produced by governments, and defend the statisticians who produce them from political interference and retaliation.

18/

IX. Data visualization is a powerful tool for illuminating the meaning of data, but also for convincing (and misleading) people. Learn to parse visualizations and to spot their implicit arguments.

19/

X. Curiousity (and its correlate, humility about the things you think you know) is the best tonic against being deceived - by others, or yourself.

20/

Harford is a gifted science communicator, a wonderful and sprightly writer. He's spent many years teaching millions of people to consume statistics critically - but with this book, he's doing something nobler:

21/

He's teaching us to consume statistics WISELY.

eof/