

Fallacies in Interpreting Data

Kecohan dalam Interpretasi Data

Fathul Wahid

fathul.wahid@uii.ac.id



UNIVERSITAS
ISLAM
INDONESIA

Program Studi Informatika, Universitas Islam Indonesia

Be aware

I just want to share stories of books and literature that I have read and learned



Opening illustrations

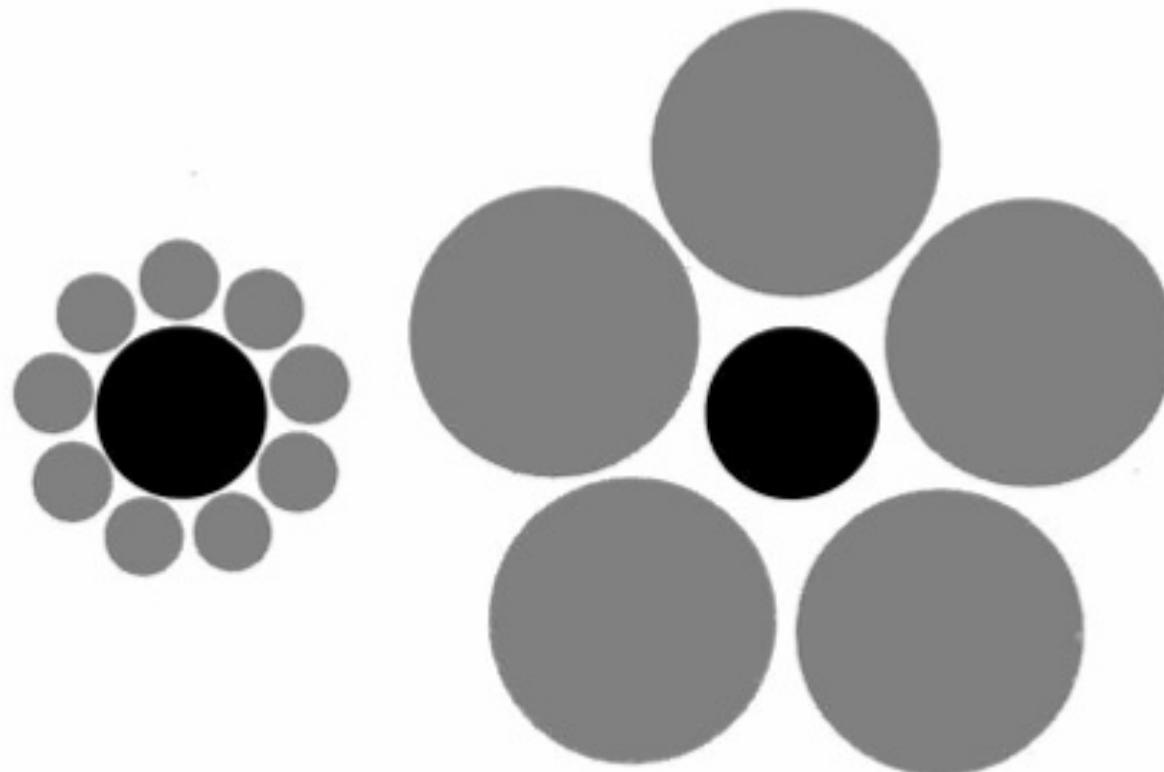
Which is the biggest number?

56

83

19

The truth of relativity



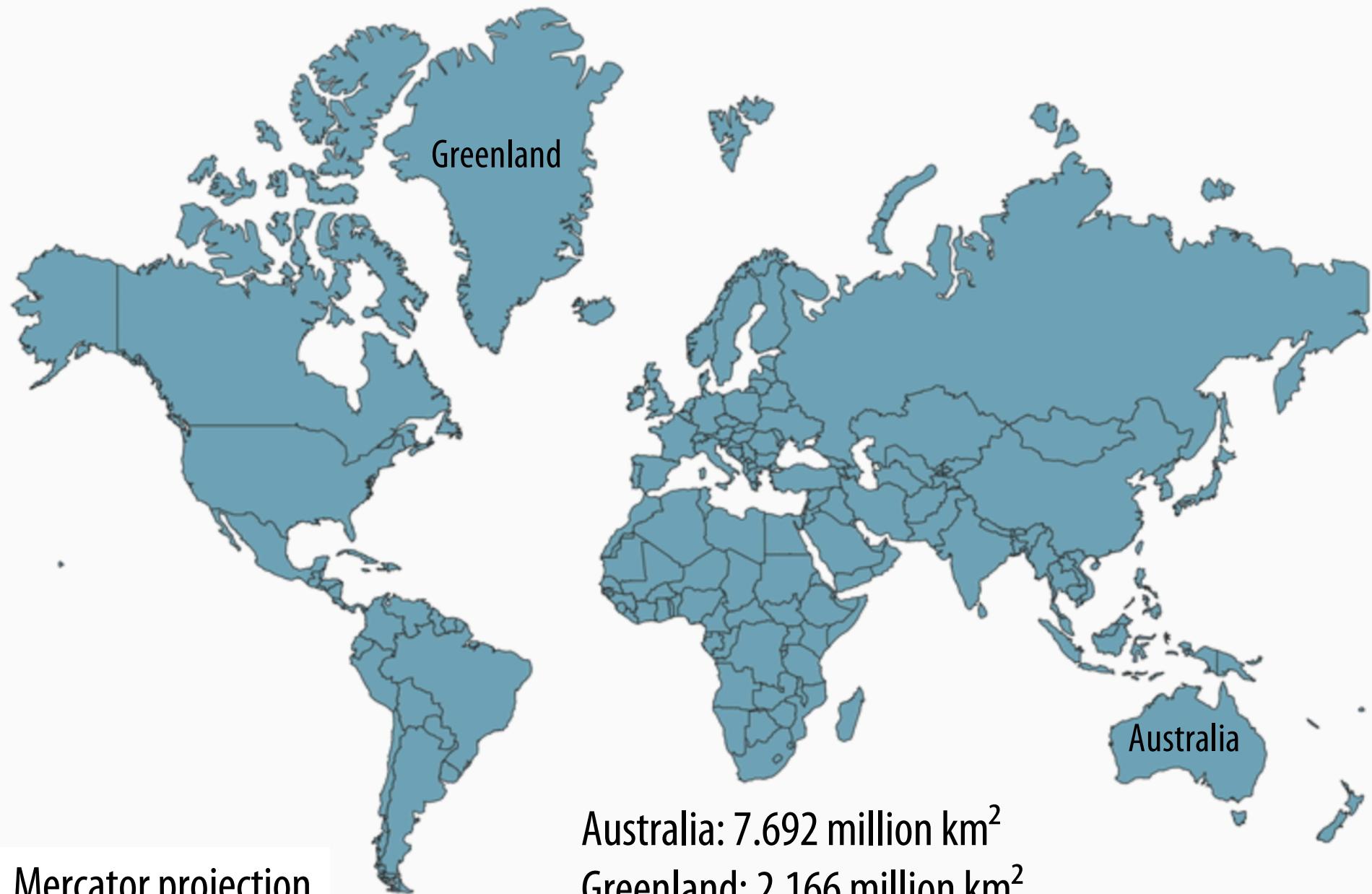
COVID-19 - Confirmed cases and deaths

Country, Other	Total Cases	New Cases	Total Deaths	New Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop
China	80,796	+3	3,169		62,836	14,791	4,257	56.1
Italy	15,113	+2,651	1,016	+189	1,258	12,839	1,153	250.0
Iran	10,075	+1,075	429	+75	3,276	6,370		120.0
S. Korea	7,869	+114	66	+6	333	7,470	93	153.5
Spain	3,146	+869	86	+31	189	2,871	190	67.3
France	2,876	+595	61	+13	12	2,803	129	44.1
Germany	2,745	+779	6	+3	25	2,714	9	32.8
USA	1,672	+371	40	+2	30	1,602	10	5.1
Switzerland	868	+216	7	+3	4	857		100.3
Norway	800	+171	1	+1	1	798		147.6

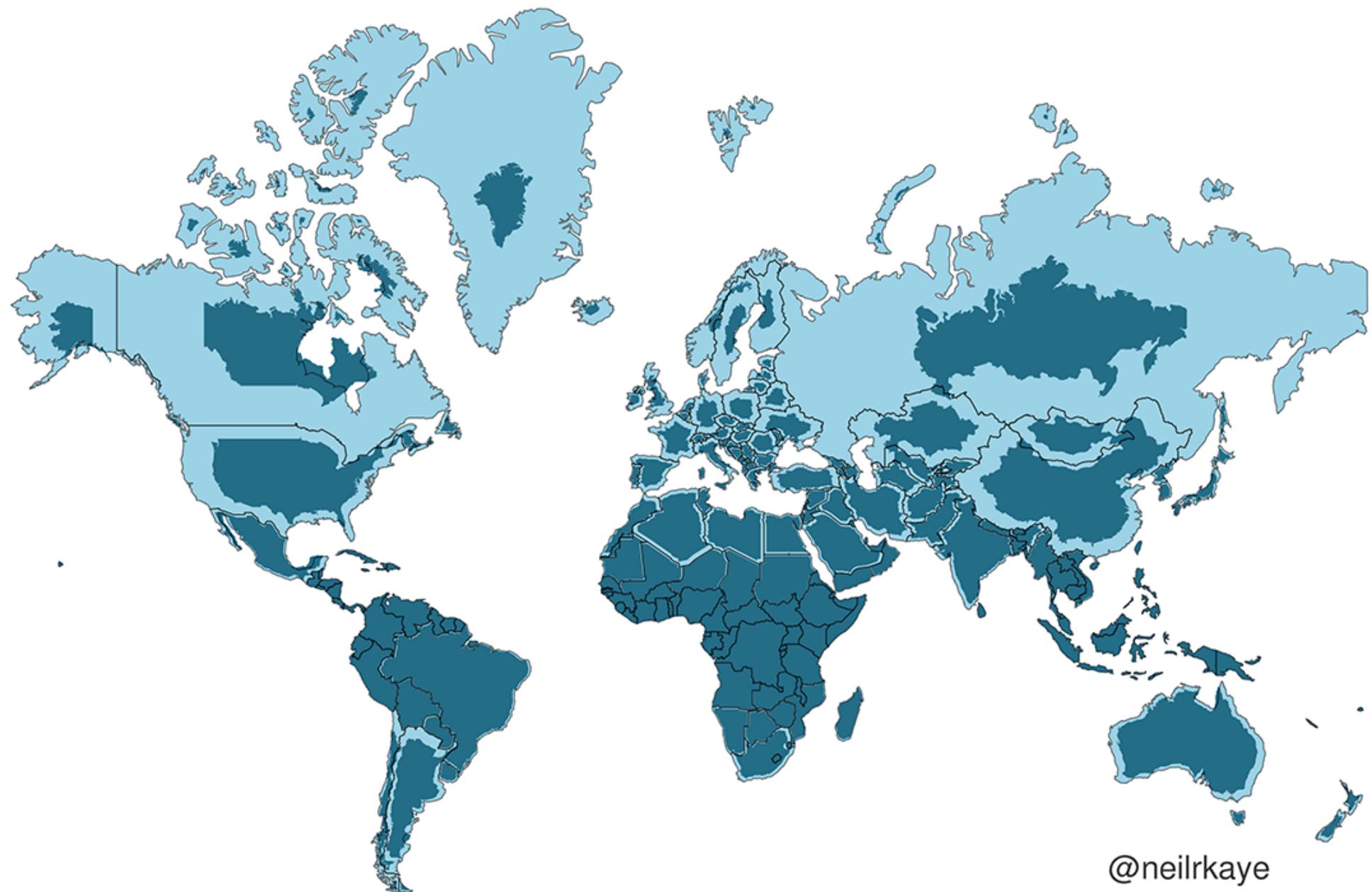
COVID-19 – Cont'd.

Country, Other	Total Cases	New Cases	Total Deaths	New Deaths	Total Recovered	Active Cases	Serious, Critical	Tot Cases/ 1M pop
Romania	59	+12			6	53	1	3.1
Philippines	52	+3	2		2	48	1	0.5
Poland	51	+20	1	+1		50	3	1.3
Taiwan	49	+1	1		20	28		2.1
Saudi Arabia	45				1	44		1.3
Vietnam	39				16	23		0.4
Indonesia	34		1		3	30		0.1
Russia	34	+6			3	31		0.2
Chile	33	+10				33		1.7
Palestine	31	+1				31		6.1

Data visualization that misleads people

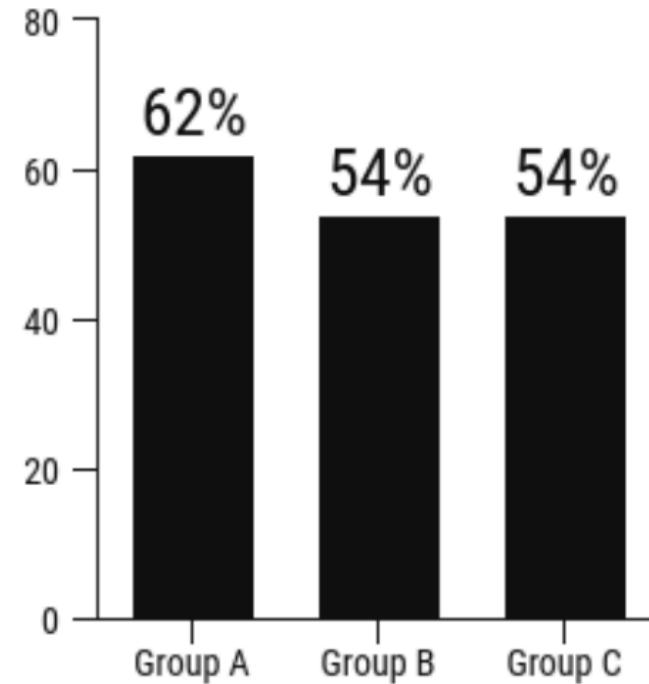
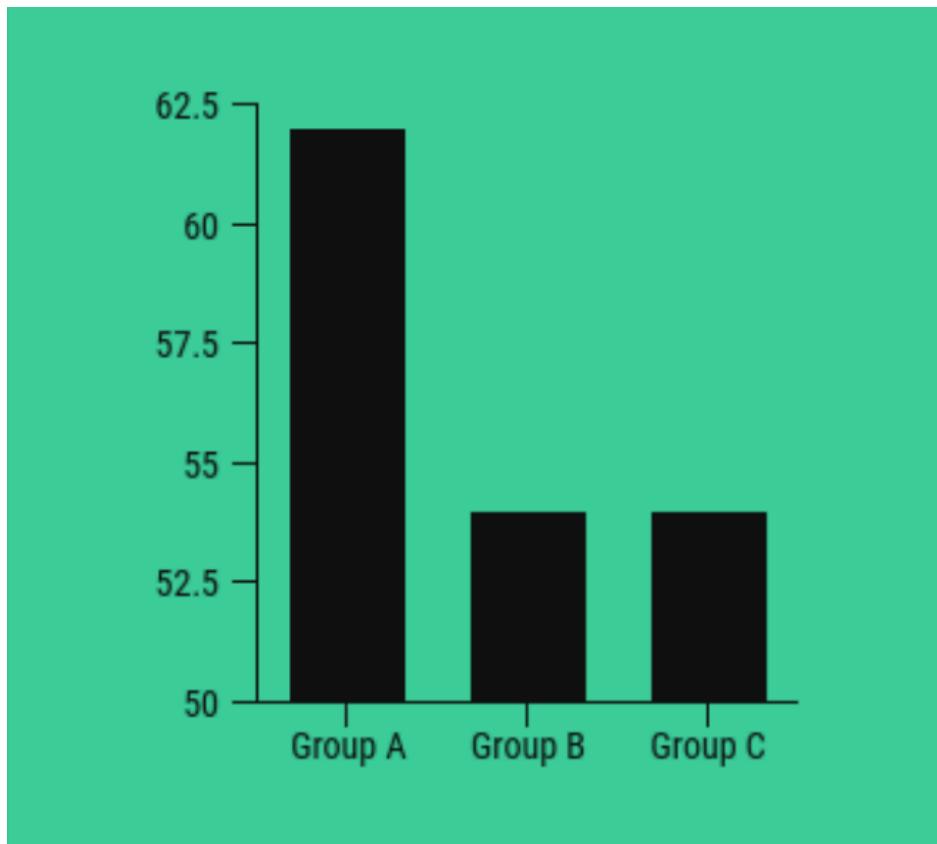


MERCATOR PROJECTION VS THE TRUE SIZE OF COUNTRIES

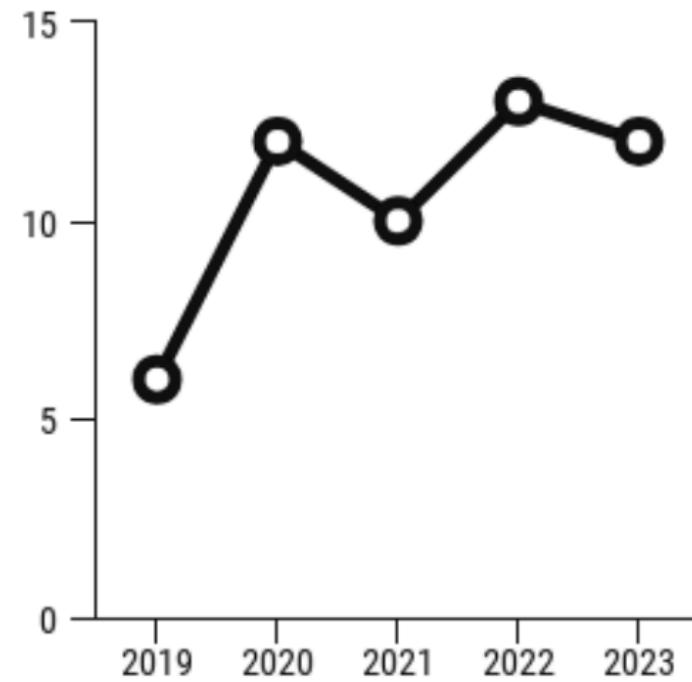
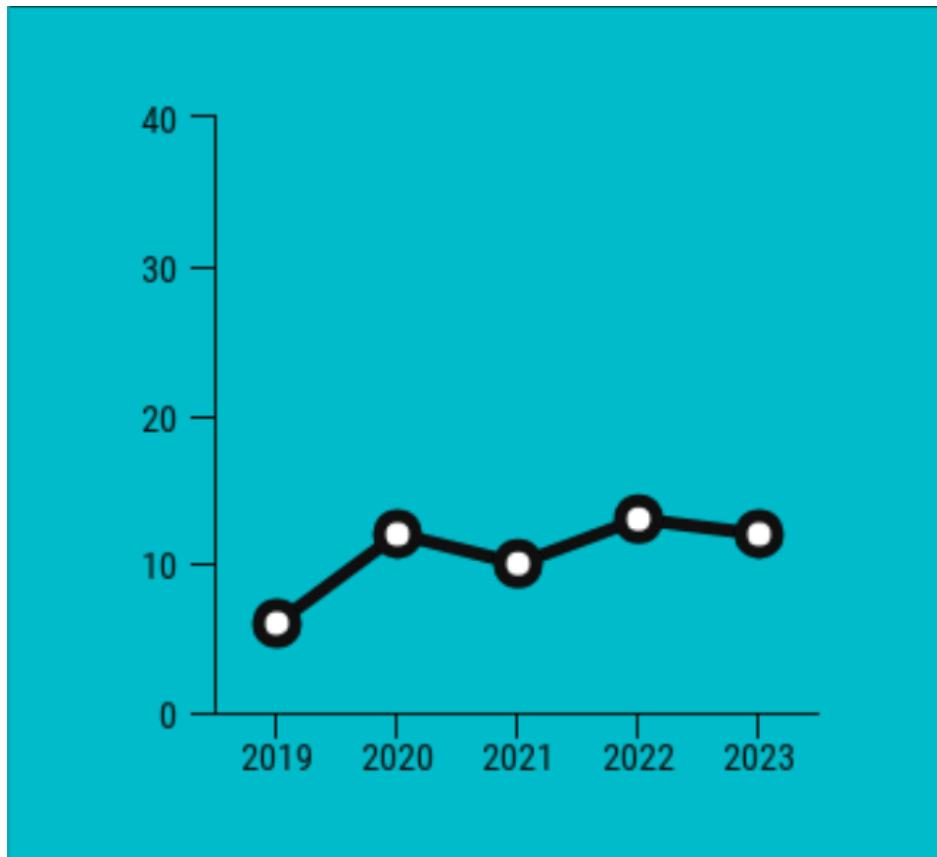


@neilrkaye

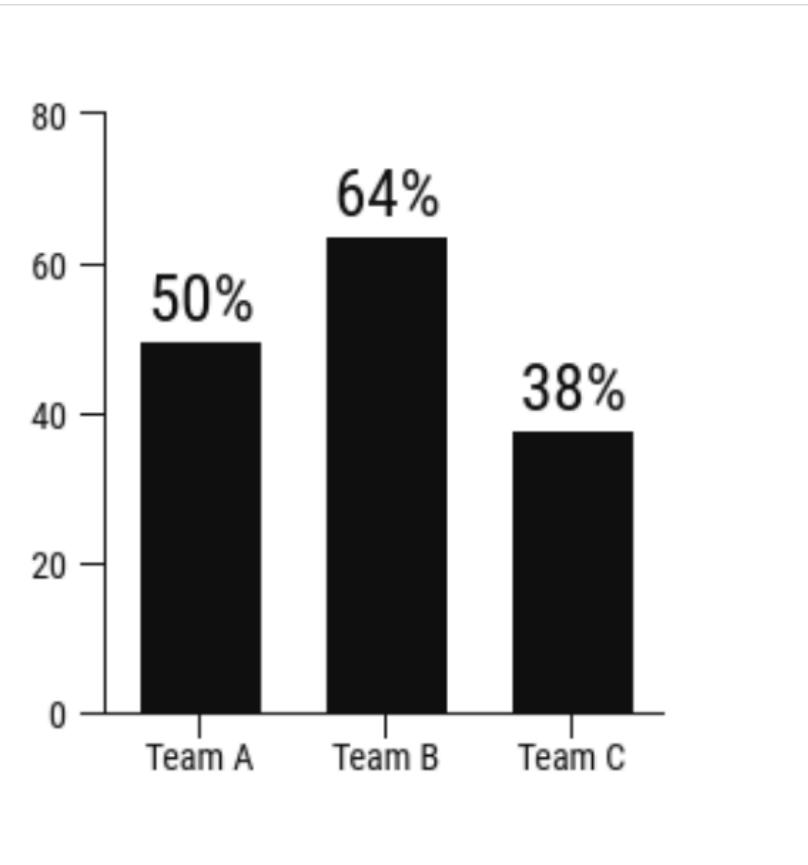
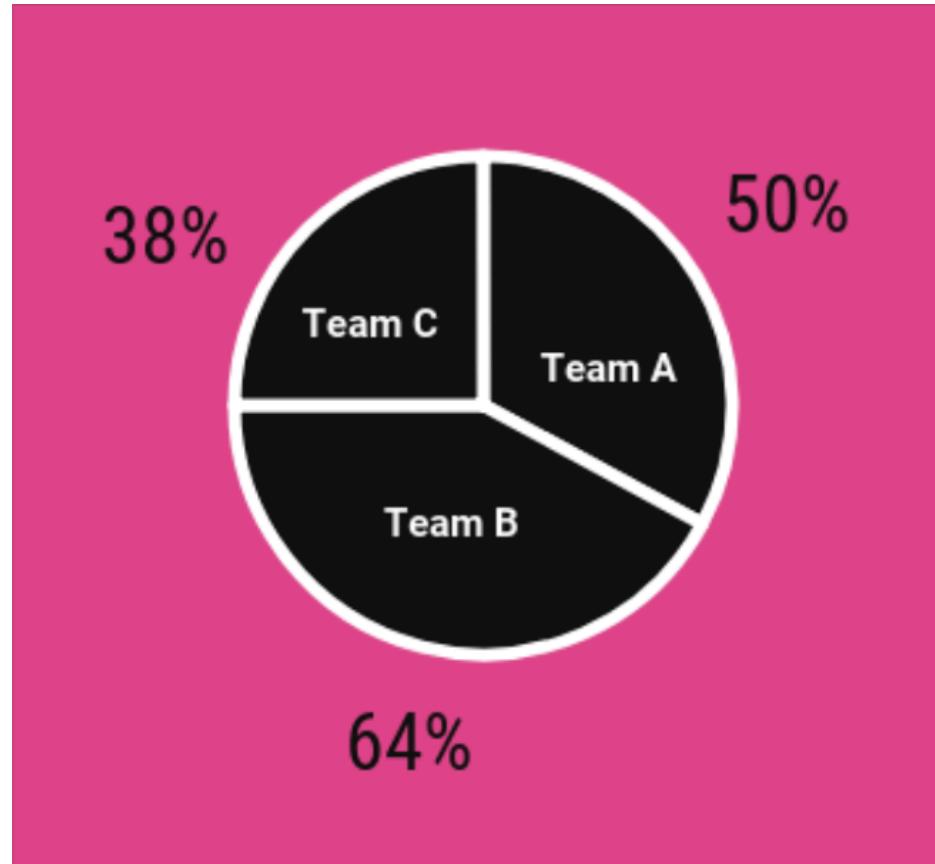
Omitting the baseline



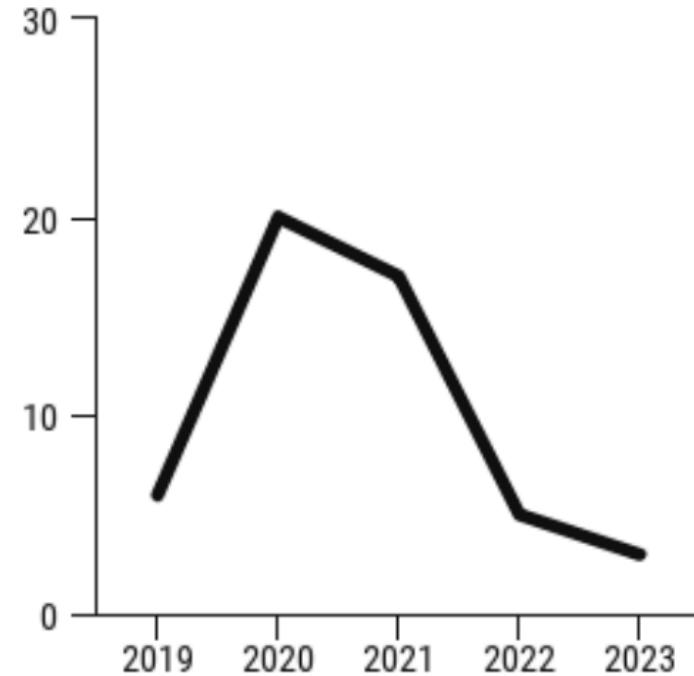
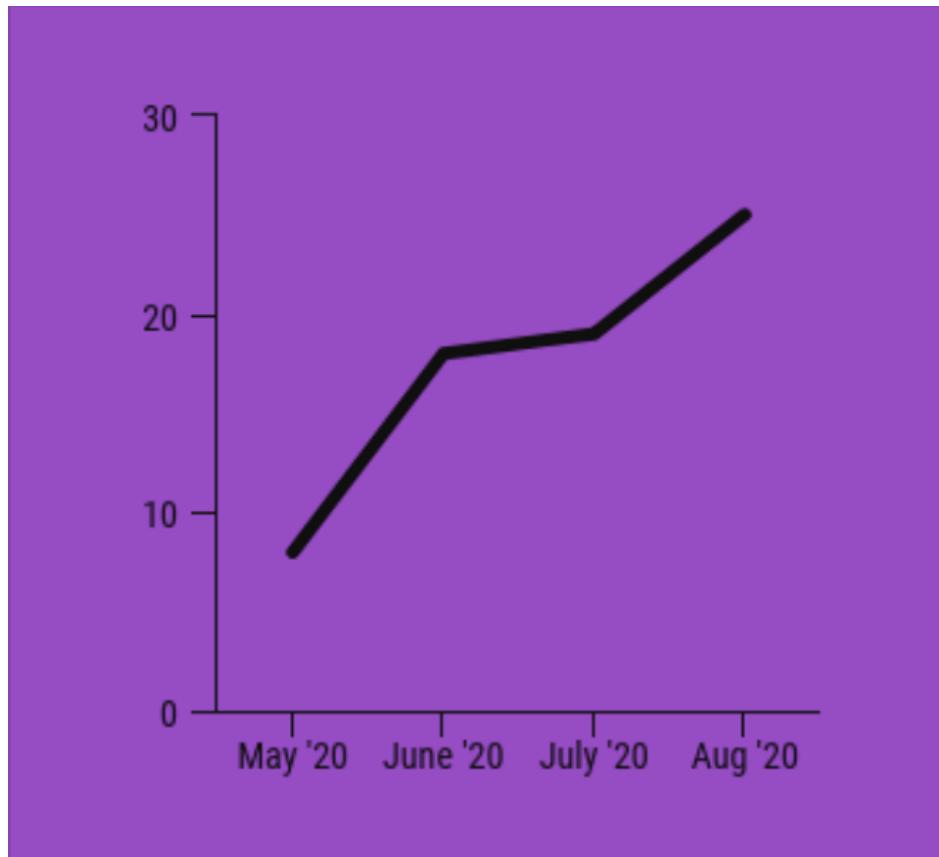
Manipulating the y-axis



Using the wrong graph



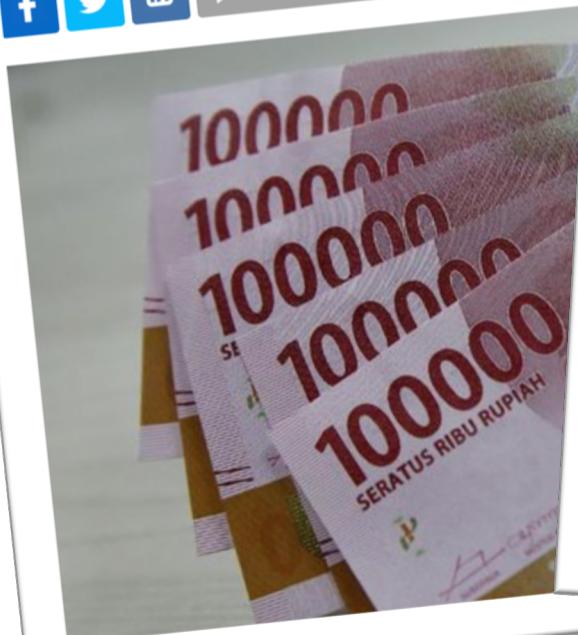
Cherry picking data



Rabu, 11 Mar 2020 18:24 WIB

Pajak Ditanggung Pemerintah Gajian Full!

Trio Hamdani - detikFinance



Follow detikFinance [f](#) [t](#) [in](#) [o](#)

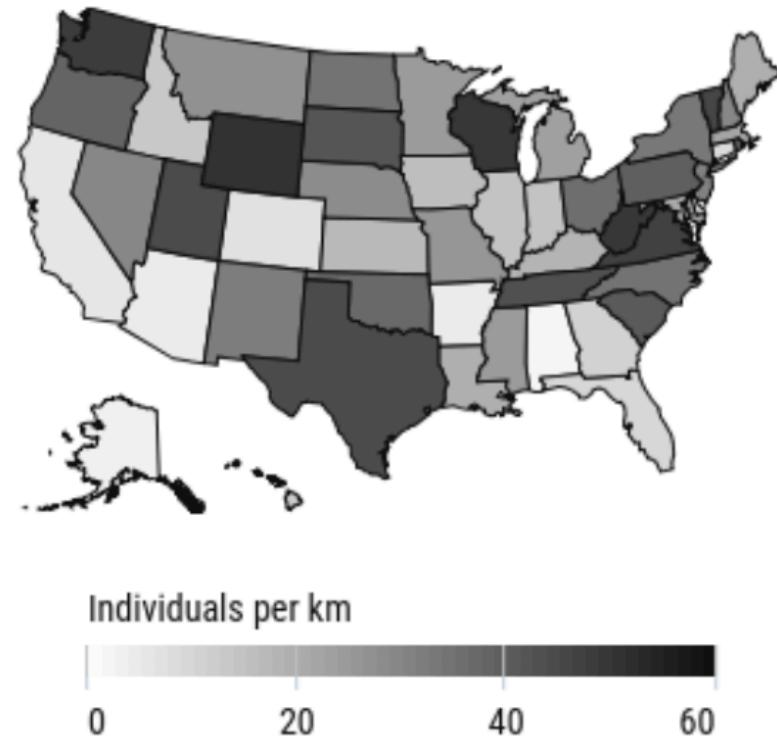
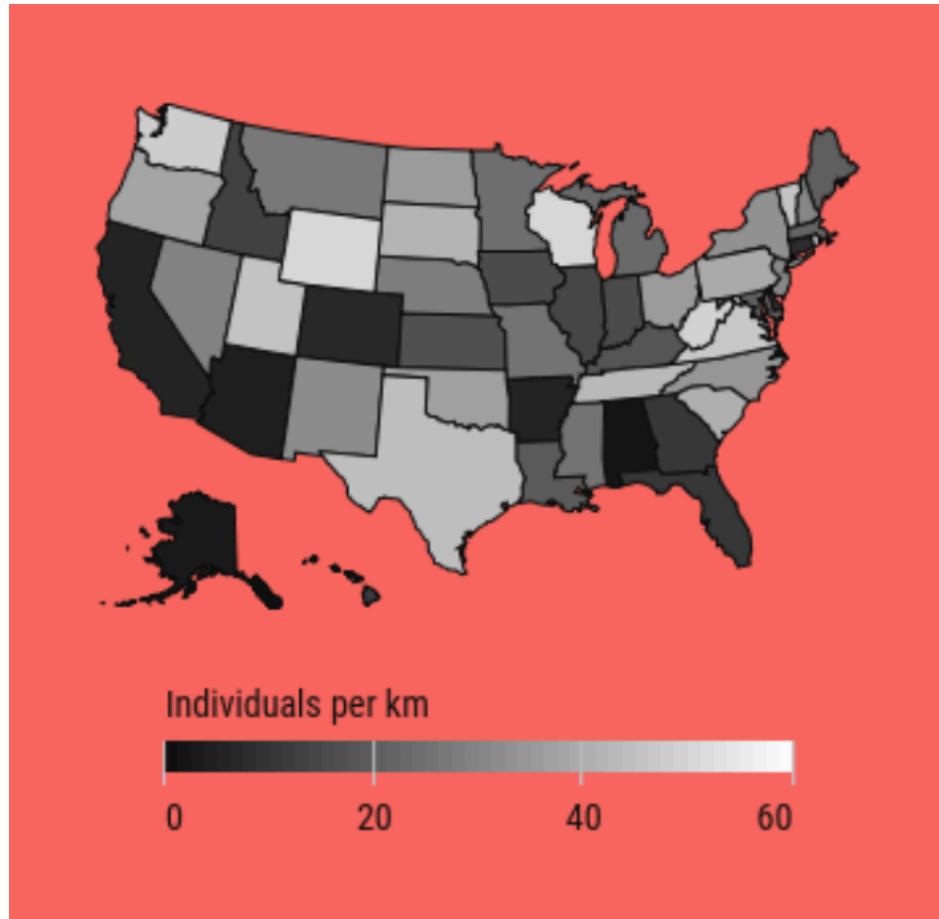
Kamis, 12 Mar 2020 14:30 WIB

Perhatian! Gajian Bebas Pajak Buat Pekerja Manufaktur

Danang Sugianto - detikFinance



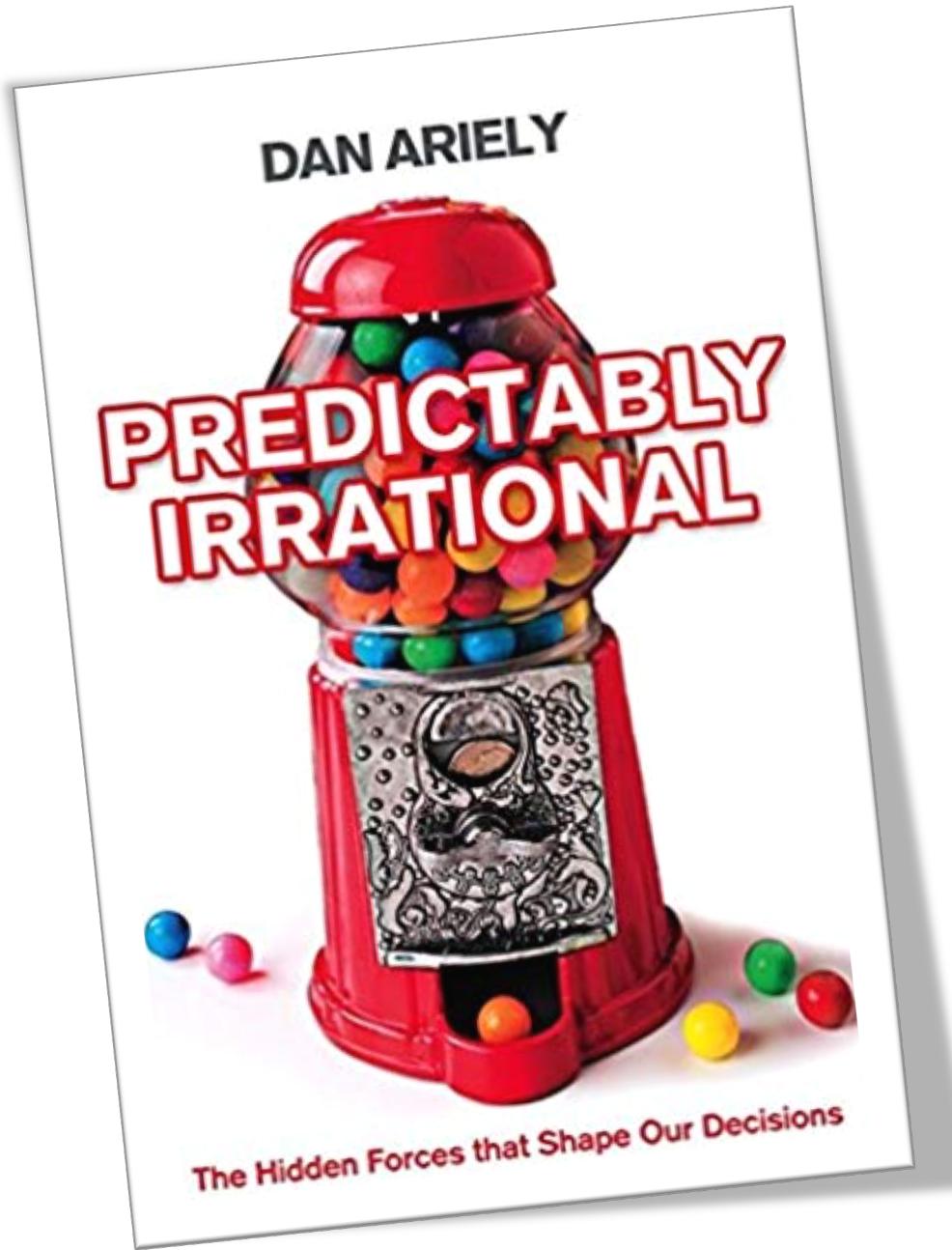
Going against convention



Fallacies

Biases

Irrationalities



Are we
rational or
irrational?

Simulasi #1

TEMPO



Majalah Tempo
Daring
Rp100.000
perbulan



Majalah Tempo
Cetak
Rp150.000
perbulan



Majalah Tempo
Daring + Cetak
Rp150.000
per bulan

Simulasi #2

TEMPO



Majalah Tempo
Daring
Rp100.000
per bulan



Majalah Tempo
Daring + Cetak
Rp150.000
per bulan



Bias 1 – Ease of recall

(based on vividness and recency)

Which of the following list was the cause of more premature deaths in the US in 1999?

- A. Tobacco use, obesity/inactivity, and alcohol
- B. Cancer, heart disease, and auto accidents

Many decisions are affected by the vividness of information

Perception: more easily recalled = more frequent

Kematian
akibat
Covid-19:
2 orang

Easy to recall

Kemenkes: 104 Orang Meninggal akibat DBD, Paling Banyak di NTT

Kompas.com - 11/03/2020, 18:57 WIB

BAGIKAN:



Komentar

Lihat Foto



Bias 2 – Retrievability (#1)

(based on memory structure)

In four pages of novel (about 2,000 words), how many words should you expect to find that have the form _____ ing (seven-letter words that end with “ing”)? Indicate your best estimate by circling one of the following values

- | | | | |
|--------|--------|----------|--------|
| A. 0 | C. 3-4 | E. 8-19 | G. 16+ |
| B. 1-2 | D. 5-7 | F. 11-15 | |

Bias 2 – Retrievability (#2)

(based on memory structure)

In four pages of novel (about 2,000 words), how many words should you expect to find that have the form _____ n _____ (seven-letter words that have “n” in the sixth position)? Indicate your best estimate by circling one of the following values

- A. 0
- C. 3-4
- E. 8-19
- G. 16+
- B. 1-2
- D. 5-7
- F. 11-15



Bias 3 – Presumed association

- Likelihood of two events occurring together.
- At least four separate situations to consider when assessing association between two dichotomous events.





"Preman" jadi imam



Pemulung berkorban seekor sapi

In sum

**frequent events versus
infrequent events**

**likely events versus
unlikely events**



The former ones are more easily to recall

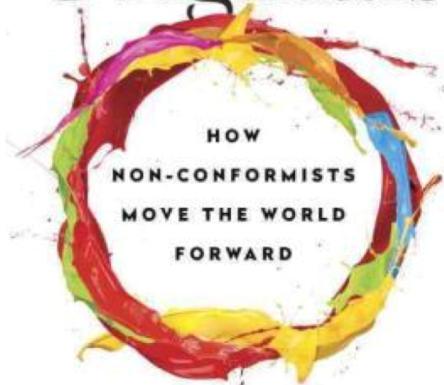
THE
BLACK SWAN



The Impact of the
HIGHLY IMPROBABLE

Nassim Nicholas Taleb

Originals



Adam Grant

New York Times bestselling author of
GIVE AND TAKE

Preface by SHERYL SANDBERG

DAVID
AND
GOLIATH



UNDERDOGS, MISFITS AND
THE ART OF BATTLING GIANTS

MALCOLM
GLADWELL

Bias 4 – Insensitivity to base rates

Mark is finishing MBA at a prestigious university. He is very interested in the arts and at one time considered a career as a musician. Is he more likely to take a job:

- A. In arts management
- B. With an Internet start-up

Description of Mark versus characteristics of workers (arts management and Internet start-ups)

Bias 5 – Insensitivity to sample size

Example: “Four of five dentists surveyed recommend sugarless gum for their patients who chew gum”

Issue: The number of dentist surveyed (i.e. sample size)

Bias 6 – Misconception of chance

You have started stocks on the Internet, beginning with five different stocks. Each stock goes down soon after your purchase. As you prepare to make a sixth purchase, you reason that it should be more successful, since the last five were “lemons”. After all, the odds favor making at least one successful pick in six decisions. This thinking is:

- A. Correct
- B. Incorrect

Issue: Do the past events have “connection” to the upcoming one?

Bias 7 – Regression to the mean

Examples of regression to the mean:

- Brilliant students frequency have less successful siblings
- Short parents tend to have taller children
- Firms that achieve outstanding profits one year tend to perform less well the next year

Another example

Nine stores,
similar size and
merchandise
selection

Prediction 2002:
10% increase from
2000

Store	Sales in 2000 (USD)	Sales in 2002 (USD)
1	12,000,000	?
2	11,500,000	?
3	11,000,000	?
4	10,550,000	?
5	10,000,000	?
6	9,500,000	?
7	9,000,000	?
8	8,500,000	?
9	8,000,000	?
Total	90,000,000	99,000,000

Bias 8 – The conjunction fallacy

Individuals falsely judge that conjunctions (two events co-occurring) are more probable than a more global set of occurrences of which the conjunction is a subset.



Look alike?

In sum

- Our experience tells us that the likelihood of a specific occurrence is related to the likelihood of a group of occurrences which the specific occurrence represents.
- Systematic irrationalities, when we are unaware of this tendency.

Bias 9 – The confirmation trap

Individuals tend to seek confirmatory information for what they think is true and fail to search for disconfirmatory evidence.



Bias 10 – Hindsight and the curse of knowledge

After finding out whether or not an event occurred, individuals tend to overestimate the degree to which they would have predicted the correct outcome.

Familiar comments?

“I **knew** that was a bad play”

“I knew you should have turned left at the fork”

“There were a plenty of evidence that he was the wrong man for the job”

Another problem

Information framing - Problem 1

A large care manufacturer has recently been hit with a number of economic difficulties, and it appears as if three plants need to be closed and 6,000 employees laid off. The vice president of production has been exploring alternative ways to avoid this crisis. She has developed two plans:

- A. This plan will save one of the three plants and 2,000 jobs.
- B. This plan has a 1/3 probability of saving all three plants and all 6,000 jobs, but has a 2/3 probability of saving no plants and no jobs.

Which plan would you select?

Information framing - Problem 2

A large care manufacturer has recently been hit with a number of economic difficulties, and it appears as if three plants need to be closed and 6,000 employees laid off. The vice president of production has been exploring alternative ways to avoid this crisis. She has developed two plans:

- C. This plan will result in the loss of two of the three plants and 4,000 jobs.
- D. This plan has a 2/3 probability of resulting in the loss of all three plants and all 6,000 jobs, but has a 1/3 probability of losing no plants and no jobs.

Which plan would you select?

The framing of information

- People evaluate rewards and losses relative to a neutral reference point
- People think about potential outcomes as gains or losses relative to this fixed, neutral reference point

References (+ sources)

- Various relevant webpages that are not always traceable (e.g. visualcapitalist.com)
- Bazerman, M. H. (2002). *Judgment in Managerial Decision Making*. New York: Wiley.
- Eugene Levine, E. (1959). Interpreting Statistical Data. *The American Journal of Nursing*, 59 (2), 230-233.
- Areily, D. (2008). *Predictably Irrational*. London: Harper.