



UNIVERSITAS AIRLANGGA  
Excellence with Morality

# DATA DAN PUSTAKA (SIP107)

Pertemuan 12: Gaya Bahasa yang Digunakan dalam Menginterpretasi  
Data dan Melaporkannya





# Topik

1. Prinsip-prinsip dasar melaporkan interpretasi atas hasil analisis data melalui bahasa tulis
2. Menarasikan *uncertainty* dengan *hedging language*
3. Menarasikan rentang kepercayaan (*confidence interval*)





# Data telah dianalisis, lalu?

- Setelah data dianalisis, maka temuan-temuan penting dari proses analisis data tersebut harus dikomunikasikan ke publik. Ada tiga jenis audiens, yaitu:
  - **Komunitas akademik** yang latar belakangnya **serumpun** dengan anda.
  - ***Informed lay public***, yaitu orang awam terdidik. Meskipun awam, audiens ini memiliki kapasitas mencerna informasi yang kompleks tetapi terbatas pada keahlian/spesialisasi tertentu. Misalnya, komunitas akademik di luar disiplin ilmu anda atau orang awam kebanyakan tetapi memiliki minat sains yang tinggi.
  - ***General lay public***, yaitu orang awam yang tidak memiliki latar belakang dalam bidang sains.
- Cara mengkomunikasikan temuan dari analisis data sangat **bergantung pada karakteristik audiens** yang anda tuju.





# Komunitas akademik serumpun

- Dalam mengkomunikasikan temuan penelitian kepada sejawat dengan latar belakang keilmuan yang sama (secara tertulis), maka anda perlu *memperhatikan norma-norma/kelaziman* yang berlaku di komunitas akademik anda.
- Norma-norma ini secara umum mencakup bagaimana cara menata letak tulisan, aturan pengutipan dan penulisan daftar pustaka, cara mempresentasikan analisis data dan mendiskusikannya, dsb.
- Misalnya,
  - Bidang Medis menggunakan gaya [American Medical Association \(AMA\)](#)
  - Bidang Sosiologi mengikuti gaya [American Sociological Association \(ASA\)](#)
  - Bidang Teknik menggunakan [Institute of Electrical and Electronics Engineers \(IEEE\)](#)
  - Bidang Psikologi menggunakan gaya [American Psychological Association \(APA\)](#)





# Komunitas akademik serumpun

- Ketika berkomunikasi dengan sejawat serumpun, maka anda **diperkenankan** (dan mungkin, diharuskan) untuk **menggunakan istilah-istilah teknis**, yang lazim dan dipahami secara khusus oleh sejawat anda.
- Dalam konteks ini, menggunakan *jargon* atau istilah-istilah teknis akan menambah **akurasi** dari informasi yang anda ingin komunikasikan.
- Sebisa mungkin, *pelajari dan ikuti dengan baik* norma/kelaziman tersebut agar sejawat anda dapat secara aktif memberikan masukan pada gagasan/temuan penelitian anda.





## Untuk Publik Awam (*general public*)

- Berkomunikasi dengan publik awam umumnya membutuhkan usaha yang lebih besar daripada dengan sejawat serumpun.
- Ketika berkomunikasi dengan publik awam, maka anda harus *menyeimbangkan* antara *ketepatan* (presisi) dan *keterbacaan* (kemudahan informasi tersebut dipahami awam).
- Ilmuwan pada umumnya sangat berhati-hati untuk menjelaskan hasil analisis datanya karena mereka *menghindari* untuk membuat kesimpulan (*ekstrapolasi*) yang berlebihan dari data yang dianalisis.





## Untuk Publik Awam (*general public*)

- Bagian yang tersulit dari hal ini adalah menarasikan **ketidakpastian**.
  - Contohnya, ketika seorang ilmuwan mencoba untuk menjelaskan hubungan antara dua variabel; ketika mereka *melebih-lebihkan* hubungan tersebut, maka kemungkinan publik (dan sejawat) kurang mempercayai/mencurigai klaim tersebut.
  - Namun apabila ilmuwan tersebut *mengecilkan* atau menyederhanakan hubungan tersebut, maka mungkin ilmuwan tersebut melewatkan suatu temuan yang penting.
  - Ingat! Statistik sejatinya bicara mengenai probabilitas sehingga anda sebaiknya **menghindari diksi** yang menyiratkan realitas yang rigid/pasti.
- Untuk mengatasi hal ini, maka ilmuwan menggunakan teknik yang disebut dengan *hedging language*.





## ***Hedging language***

- *Hedging language* adalah gaya bahasa yang biasa digunakan ilmuwan untuk mengekspresikan kehati-hatian ketika mendiskusikan ketidakpastian dalam menginterpretasi data.
- Ilmuwan terkadang tidak terlalu yakin untuk membuat klaim yang kuat berdasarkan data yang ia miliki, sehingga interpretasi mereka atas data biasanya dideskripsikan dengan *tone* berhati-hati.

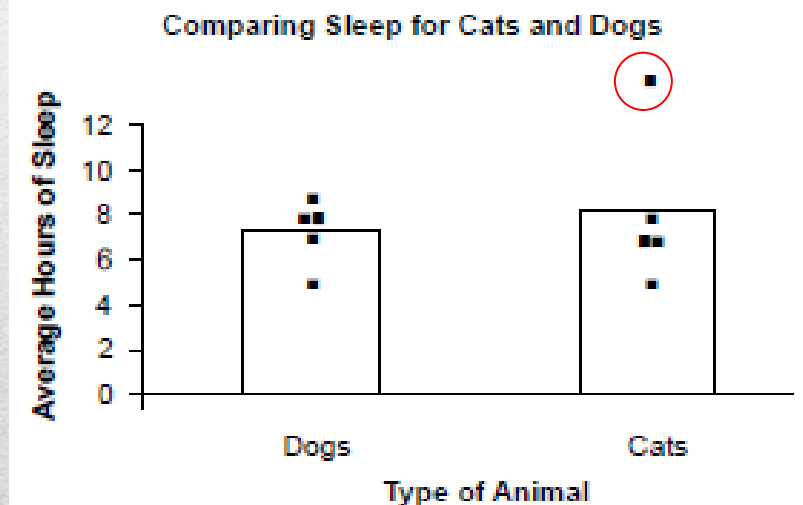




## Contohnya...

Pada grafik di sebelah kanan terlihat bahwa kucing rata-rata tidur lebih lama daripada anjing. Namun, pada kelompok sampel kucing, ada **satu kucing yang tidur jauh lebih lama** daripada kucing-kucing lainnya dan mempengaruhi rerata kelompok kucing. Oleh karena itu, *sangat mungkin* sebenarnya hanya ada sedikit perbedaan rerata durasi tidur antara kucing dan anjing, namun perbedaan ini akan terlihat lebih jelas apabila pengambilan data diulang menggunakan jumlah sampel yang lebih besar.

Comparing sleep times for cats and dogs with nominal data



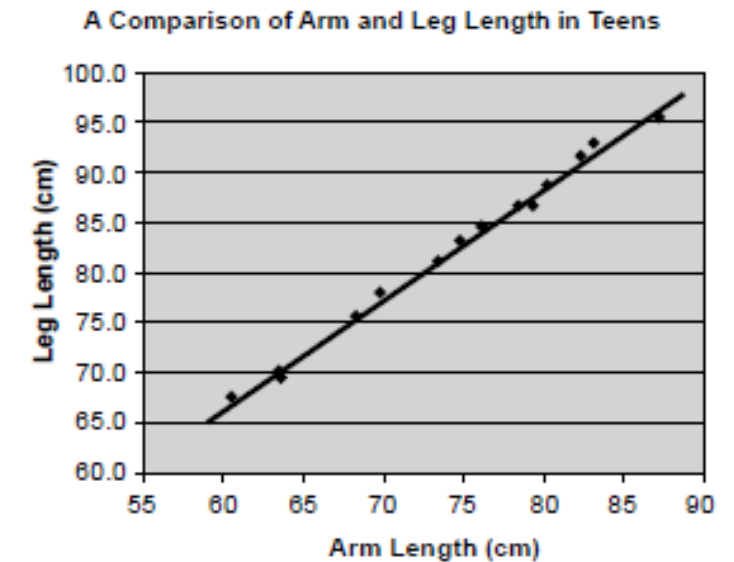




## Contohnya..

Grafik di sebelah kanan menunjukkan hubungan antara panjang lengan dengan tungkai pada remaja. Dari grafik tersebut, terlihat ada **hubungan yang sangat kuat** antara panjang lengan dan tungkai pada remaja, dimana semua partisipan posisinya mendekati garis diagonal yang menggambarkan hubungan antara kedua variabel tersebut. Tidak tampak adanya partisipan yang memiliki nilai ekstrim (*outlier*). Informasi ini mungkin berguna bagi desainer baju terusan (*one-piece*) yang dapat mendesain baju dengan proporsi lengan dan tungkai yang presisi dan akan selalu pas digunakan untuk remaja.

Comparing arm- and leg-length data in teens with interval ratio data







UNIVERSITAS AIRLANGGA  
Excellence with Morality

NASIONAL

## Generasi Z yang Ogah Divaksin

Dua hasil survei menyebutkan anak-anak muda enggan divaksin dan meragukan kemanjuran vaksin Covid-19.

EDISI, 22 FEBRUARI 2021

DIKO OKTARA



## Covid-19: Dua survei menyebut banyak anak muda menolak vaksin, bagaimana meyakinkan mereka?

Pijar Anugerah  
BBC News Indonesia

23 Februari 2021



Pemerintah Indonesia diminta memperbaiki strategi komunikasi tentang vaksin Covid-19 setelah dua survei menyebut banyak anak muda, terutama Generasi Z, menolak untuk divaksinasi.

Sumber:

BBC News, 23 Februari 2021

Koran Tempo, 22 Februari 2021





# Mari kita lihat datanya...

- Ringkasan eksekutif (bukan hasil analisis lengkapnya) [dapat diakses di sini](#).
- Gambaran singkat surveinya..
  - Survei melibatkan 1200 partisipan yang diambil secara acak dari *participant pool* milik lembaga survei tersebut, lalu partisipan yang terpilih menjadi sampel, dihubungi via telepon untuk wawancara.
  - Dari 1200 partisipan tersebut, kurang dari 10 persennya (sekitar 120 orang) berusia antara 22-25 tahun.
  - Dasar dari klaim tersebut adalah pada kelompok usia 22-25 tahun tersebut, hanya 45.1 persen yang bersedia divaksinasi, paling rendah di antara kelompok usia lainnya. Sisanya, (yaitu 55 persen atau 54 partisipan) tidak bersedia divaksinasi.





## KESEDIAAN DIVAKSIN & KESEDIAAN MEMBAYAR JIKA BERSEDIA DIVAKSIN MENURUT DEMOGRAFI

		BERSEDIA DIVAKSIN			BERSEDIA MEMBAYAR VAKSIN (JIKA BERSEDIA DIVAKSIN)		
	Base	Bersedia	Tidak bersedia	TT/TJ	Bersedia	Tidak bersedia	TT/TJ
<b>GENDER</b>							
Laki-laki	50.1	56.3	38.9	4.8	24.4	70.7	5.0
Perempuan	49.9	53.5	42.9	3.6	22.9	69.4	7.7
<b>USIA</b>							
<= 21 tahun	12.0	51.5	45.5	0.0	23.7	63.4	12.9
22 - 25 tahun	9.8	45.1	43.9	10.9	25.1	74.9	0.0
26 - 40 tahun	37.0	55.9	40.4	3.8	23.7	68.7	7.6
41 - 55 tahun	25.4	53.6	42.6	3.7	21.0	75.2	3.8
> 55 tahun	15.8	61.0	34.1	4.9	26.6	67.9	5.4
<b>AGAMA</b>							
Islam	87.8	52.6	43.6	3.9	23.1	71.4	5.5
Lainnya	12.2	71.6	21.9	6.5	26.5	63.0	10.5
<b>ETNIS</b>							
Jawa	41.8	56.8	39.1	4.0	24.3	71.5	4.2
Lainnya	58.2	53.5	42.2	4.3	23.2	69.0	7.8
<b>PENDIDIKAN</b>							
<= SD	25.8	50.6	42.9	6.5	19.8	76.4	3.8
SLTP	20.9	46.4	48.6	5.0	22.2	69.4	8.4
SLTA	39.7	60.0	37.0	2.9	21.3	72.7	6.0
Kuliah	13.6	60.8	36.8	2.3	39.0	52.9	8.1
<b>PENDAPATAN</b>							
< 1.2 juta	40.7	50.8	43.8	5.3	17.7	76.4	5.9
1.2- 2 juta	23.3	54.1	43.0	2.9	21.3	70.4	8.3
>= 2 juta	36.0	60.5	35.8	3.8	31.0	64.1	4.9
<b>DESA/KOTA</b>							
Pedesaan	50.5	51.7	42.8	5.5	23.4	69.2	7.4
Perkotaan	49.5	58.2	39.0	2.9	23.9	70.8	5.2

INDIKATOR





## Persoalannya...

- Membandingkan proporsi orang yang bersedia/tidak bersedia divaksin sifatnya *relative to the group*, **tidak tepat apabila dibandingkan dengan kelompok usia lainnya** karena jumlah partisipannya bervariasi di tiap kelompok.
- Peneliti menyimpulkan partisipan Gen Z banyak yang menolak vaksinasi tetapi dasarnya hanya proporsi di kelompok usia 22-25 tahun, **tanpa melibatkan** partisipan yang **berusia 21 tahun ke bawah**.
- Ketika dilakukan analisis korelasi antara usia (dengan seluruh partisipan dari semua kelompok usia dilibatkan) dengan kesediaan partisipan divaksinasi, **hubungan tersebut ternyata sangat lemah**.





Coefficients<sup>a</sup>

reg\_iv6\_income, reg\_iv2\_usia, re  
reg\_iv10\_percayaVaksin, reg\_iv7\_

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1	(Constant)	,857		3,298	,001
	reg_iv1_gender	,080	,031	1,141	,254
	reg_iv2_usia	,001	,015	,567	,571
	reg_iv3_etnis	,165	,063	2,294	,022
	reg_iv4_agama	-,277	-,070	-2,486	,013
	reg_iv5_educ	,033	,094	3,489	,001
	reg_iv6_income	-9,203E-5	-,001	-,028	,977
	reg_iv7_UR	-,018	-,007	-,252	,801
	reg_iv9_takutCOVID19	,141	,114	4,238	,000
	reg_iv10_percayaVaksin	,711	,491	17,804	,000
	reg_iv11_nilaiAgama	,040	,030	1,118	,264
	reg_iv12_Pilpres19	,083	,031	1,138	,255

a. Dependent Variable: reg\_dv1\_V29





## Kesimpulannya..

- Kasus ini merupakan contoh konkrit dari Simpson's Paradox, yaitu korelasi nampaknya terdeteksi pada satu-dua kelompok, namun korelasi ini tidak nampak ketika seluruh kelompok digabungkan.
- Artinya, ketika peneliti menyimpulkan ada hubungan antarvariabel **hanya dari** analisis subsampel bukanlah keputusan yang tepat dan merupakan praktik penyalahgunaan statistik, yang sayangnya, cukup sering terjadi.
- *Moral of the story*, selalu cek kesesuaian antara klaim dengan data. Kalau klaim cenderung *too good to be true*, mungkin hal tersebut adalah indikasi bahwa anda perlu mengecek data dengan lebih teliti dan hati-hati.





# Rentang kepercayaan

- Ketika melaporkan hasil statistik, anda sangat disarankan untuk selalu menyertakan rentang kepercayaan (*confidence interval*) bersama dengan *point of estimatenya*.
- Rentang kepercayaan memberikan gambaran yang lebih realistis mengenai data dan ketidakpastian yang menyertainya.
- Melaporkan rentang kepercayaan adalah keharusan di beberapa komunitas akademik (misalnya, American Psychological Association), meskipun menerjemahkan rentang kepercayaan dalam bahasa sehari-hari yang dimengerti awam bukan pekerjaan mudah.





# Bagaimana melaporkan CI?

- Mari kita gunakan contoh sebelumnya, yaitu survei mengenai vaksinasi yang dilakukan oleh suatu lembaga survei/konsultan politik,
  - Peneliti sekadar melaporkan *point of estimate* (bahwa hanya 45.1 persen partisipan berusia 22-25 tahun) yang bersedia divaksinasi.
  - Dalam laporan, disebutkan *sampling error* (atau kadang-kadang disebut sebagai *margin of error*) adalah 2.9 persen.
- Bagaimana cara yang tepat melaporkannya?





## Mari kita perbaiki...

“Setelah melakukan survei pada 1200 responden yang diambil secara acak, kami menemukan bahwa partisipan yang berusia 22-25 tahun, yang bersedia divaksinasi kurang dari separuhnya. Dari sekitar 120 partisipan yang berusia 22-25 tahun, hanya 45.1 persen yang bersedia divaksinasi, sedangkan sisanya tidak bersedia. Namun apabila pengambilan data diulang dalam jangka panjang, maka ada peluang sebesar 95 persen, persentase partisipan 22-25 tahun yang bersedia divaksinasi yang sebenarnya adalah antara 42.2 - 48.0 persen.

Tetapi, setelah peneliti mencoba memastikan ada tidaknya keterkaitan antara usia dengan kesediaan divaksinasi, peneliti tidak menemukan hubungan yang kuat. Artinya, meskipun nampaknya kesediaan divaksinasi pada partisipan berusia 22-25 cukup mengkhawatirkan, namun hal ini tidak bermakna apa-apa.”





UNIVERSITAS AIRLANGGA  
Excellence with Morality

**TERIMA KASIH**