



# MODEL REGRESI LINEAR PADA DATA PENCEMARAN UDARA WILAYAH DKI JAKARTA TAHUN 2021

Eky Fernanda  
3322600025

Sains Data Terapan

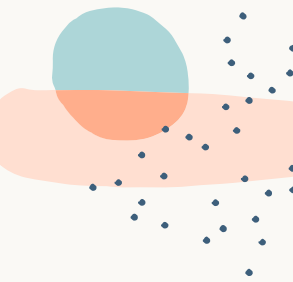
Dosen Pengampu : Ronny Susetyoko  
Politeknik ELEktronika Negeri Surabaya



# TUJUAN

Tujuan laporan ini ialah memaparkan atau menyajikan kondisi kualitas udara DKI Jakarta dengan pemantauan tahun 2021 menggunakan regresi linear dan mendapati perbandingan kadar karbonsida (co) dan nitrodioksida (no2) di Provinsi DKI Jakarta tahun 2021.

Selain itu penelitian uji regresi linear adalah untuk mengetahui kualitas category udara wilayah DKI Jakarta menciptakan antisuiipasi kepada masyarakat dengan category tersebut apakah bisa dibilang aman atau tidak.



# METODOLOGI

Metode yang digunakan dalam status ini adalah metode regresi linear berganda. Suatu model regresi linear serupa mengaitkan satu variable bebas atau predictor, model ini dapat dituliskan sebagai:

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n + e$$

Keterangan:

$Y$  = Variabel terikat atau variabel response.

$X$  = Variabel bebas atau variabel predictor.

$a$  = Konstanta.

$\beta$  = Slope atau Koefisien estimate.



# SUMBER DATA

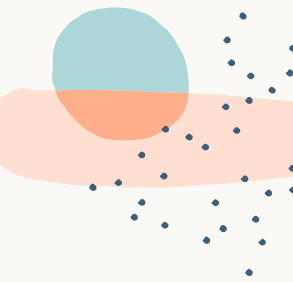


sumber data yang didapatkan dari  
data.jakarta.go.id dengan dataset Indeks  
Pencemaran Udara (ISPU) Provinsi DKI Jakarta  
Tahun 2021

# TIPE VARIABLE

Tipe variable pada data ini adalah

terdapat variable berdasarkan sifatnya dibagi menjadi dua yaitu variable kategorisk (kualitatif) dan variable numerik (kuantitatif). pada kali ini variabel yang digunakan adalah variabel kategorik (kualitatif)



# PENETAPAN VARIABLE

variable static merupakan karakteristik yang sedang diamati yang dapat mengasumsikan lebih dari satu kumpulan nilai yang dapat diberikan ukuran numerik atau kategori dari klasifikasi.

padqa data pencemaran udara DKI Jakarta Tahun 2021

Terdapat 11 variable yaitu :

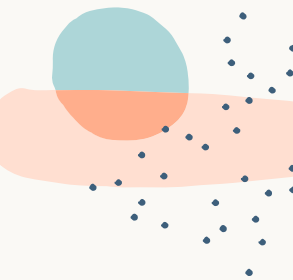
1. Tanggal : Tanggal pengukuran kualitas udara
2. Pm10 : salah satu partikulat yang diukur
3. Pm25 : salah satu partikulat parameter yang diukur
4. So2 : Sulfida parameter salah satu yang diukur
5. Co : karbon Monoksida
6. O3 : Ozon salah satu parameter yang diukur
7. No2 : Nitrogen Dioksida
8. Max : nilai ukur paling tinggi dari suatu parameter diukur
9. Critical : parameter dengan hasil pengukuran paling tinggi
10. Category : kategori hasil pengukuran indeks (ISPU)
11. Location : lokasi pengukuran



# PREPROCESSING

data preprocessing merupakan teknik untuk mengubah data mentah dalam kedalam format yang statistic. inisiatif ini diperlukan karena data mentah seringkali memiliki format yang tidak sesuai, manfaat preprocessing antara lain :

- mengurangi durasi data mining
- membuat data lebih mudah dimengerti
- mempermudah proses analisis data



# TAMPILAN DATA SETELAH DIOLAH

Out[5]:

	Source.Name	tanggal	pm10	so2	co	o3	no2	max	critical	categori	location
0	a.csv	1/1/2021	43	58	29	35	65.0	65	O3	SEDANG	DKI2
1	a.csv	1/2/2021	58	86	38	64	80.0	86	PM25	SEDANG	DKI3
2	a.csv	1/3/2021	64	93	25	62	86.0	93	PM25	SEDANG	DKI3
3	a.csv	1/4/2021	50	67	24	31	77.0	77	O3	SEDANG	DKI2
4	a.csv	1/5/2021	59	89	24	35	77.0	89	PM25	SEDANG	DKI3
...	...	...	...	...	...	...	...	...	...	...	...
360	l.csv	12/27/2021	75	61	23	40	47.0	121	PM25	TIDAK SEHAT	NaN
361	l.csv	12/28/2021	59	53	16	34	33.0	89	PM25	SEDANG	NaN
362	l.csv	12/29/2021	61	54	15	37	29.0	98	PM25	SEDANG	NaN
363	l.csv	12/30/2021	60	53	17	38	44.0	102	PM25	TIDAK SEHAT	NaN
364	l.csv	12/31/2021	64	52	44	37	53.0	90	PM25	SEDANG	NaN

365 rows × 11 columns

