

UWI1B2 LITERASI DATA

ANALISIS DAN INTERPRETASI DATA II

Anisa Herdiani, S.T., M.T.



Capaian Pembelajaran

Mampu menganalisis dan menginterpretasikan data menggunakan uji statistik dengan benar.

Topik

Uji T

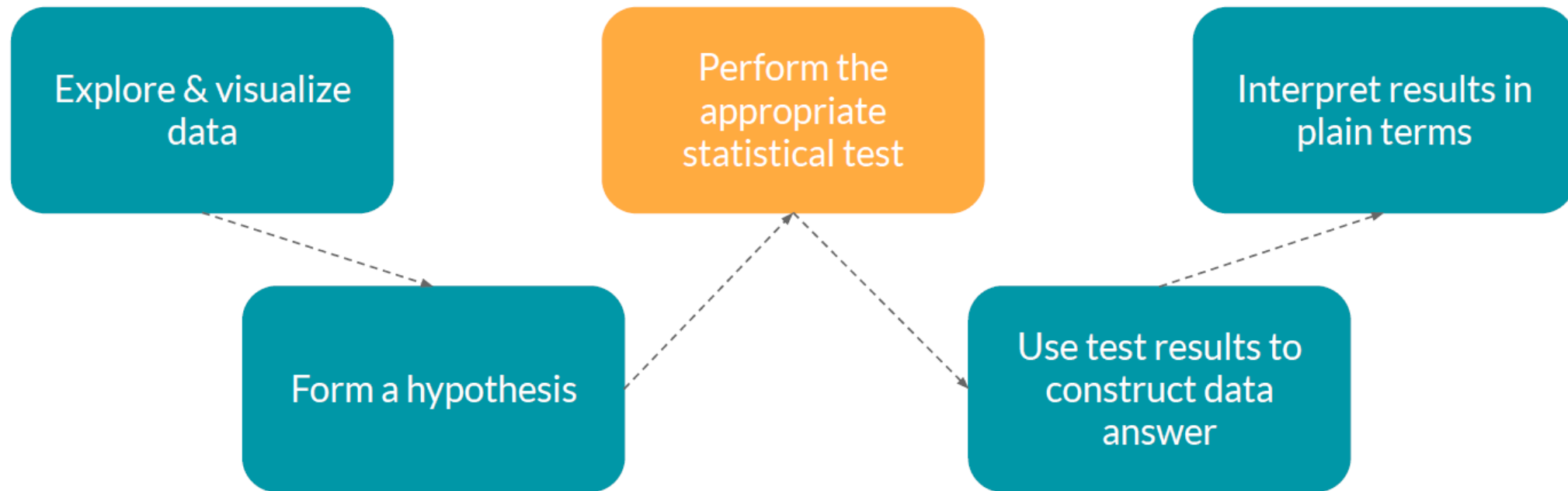
Korelasi

Chi-square test

Regresi

ANOVA

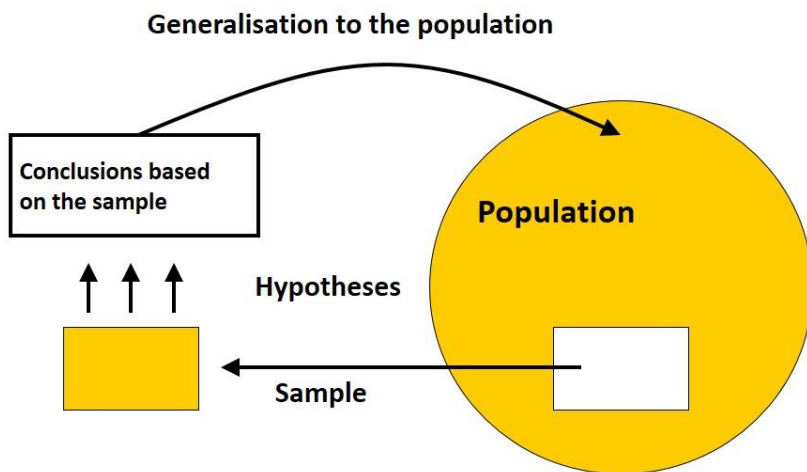
Pendekatan dalam Membandingkan Dua Variabel



Setiap pengujian memiliki **hipotesis nol**, **taraf signifikansi**, **uji statistik**, dan **nilai P**

Nilai P (*P-value*)

The idea of statistical inference



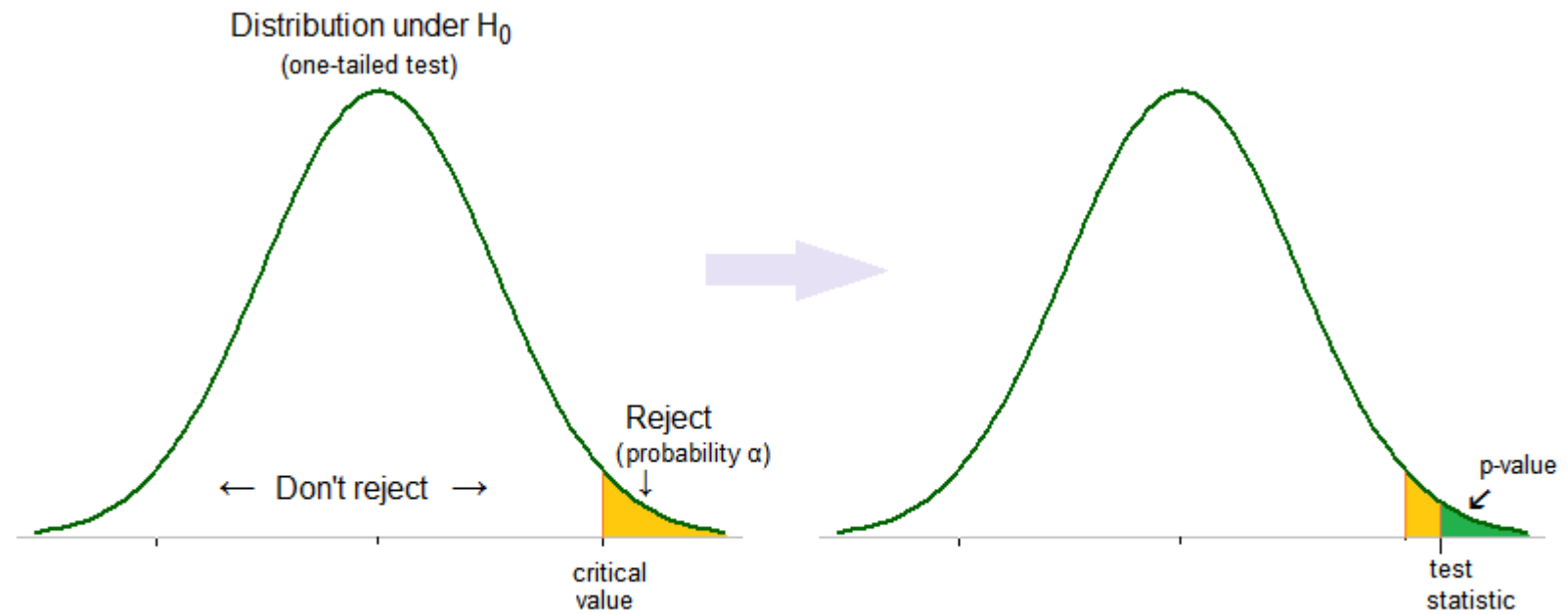
4

Nilai α (alpha) / taraf signifikansi

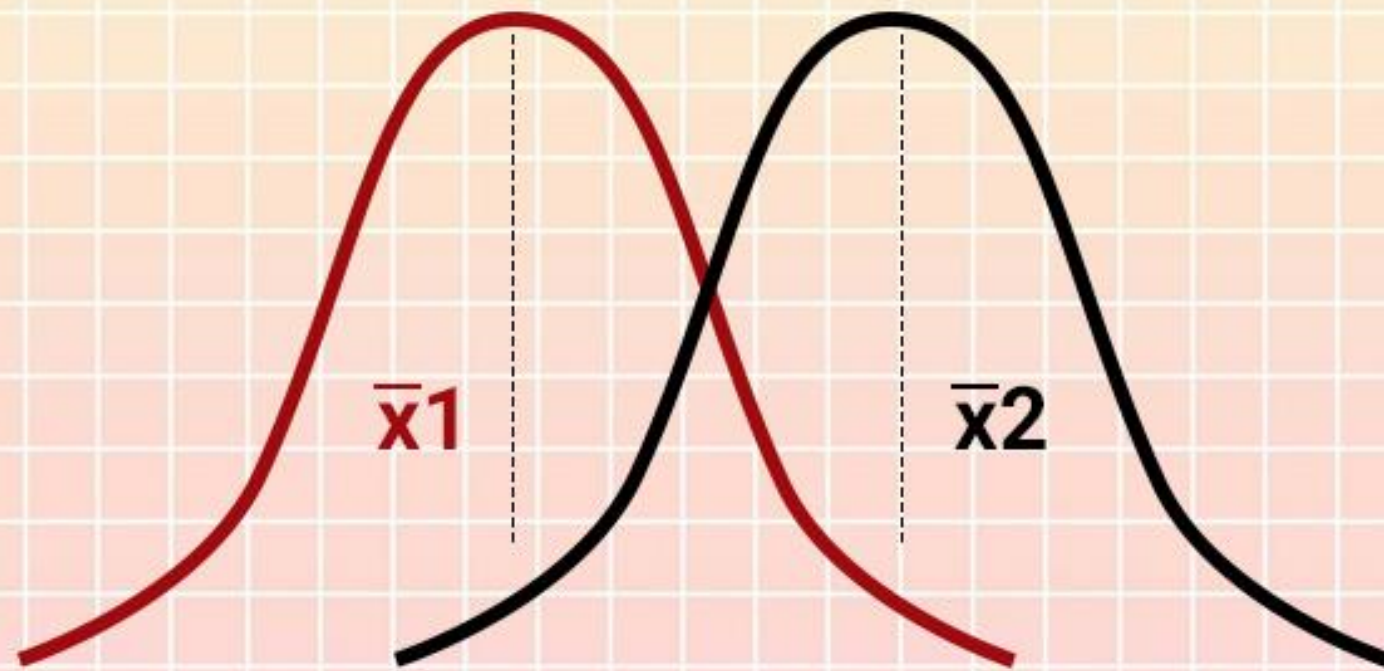
- Kesalahan maksimal yang ditentukan oleh peneliti.

p-value (sig.)

- Nilai kesalahan yang didapat peneliti dari hasil uji statistik.



- Nilai $p = 0.002$
 - Kesalahan penelitian secara statistik sebesar 0.2%
- Nilai $\alpha = 0.05$ (5%)
 - Dari 100 data sampel, paling banyak 5 kesalahan dapat ditolerir dalam penelitian.
- Nilai $p < \alpha$
 - peluang menolak hipotesis nol semakin besar



UJI-T





► Membandingkan Nilai Rata-rata Antar Dua Kelompok

- Untuk memahami perbedaan kelompok atau kondisi dapat dilakukan dengan cara membandingkan kecenderungan skor (khususnya rerata) antarkeduanya yang diperoleh dari amatan (sampel) → menggunakan uji T
- Uji T
 - **Uji T independen** untuk membandingkan rerata satu amatan dari dua kelompok subjek
 - **Uji T dependen** untuk rerata amatan dua kondisi dari kelompok subjek yang sama.





Langkah Uji T

1. Menentukan H_0 dan H_a

H_0 : tidak terdapat pengaruh yang signifikan antara kedua kelompok/kondisi

H_a : terdapat pengaruh yang signifikan antara kedua kelompok/kondisi

2. Menentukan taraf signifikansi

Taraf signifikansi yang umum digunakan sebesar 5% atau $(\alpha) = 0,05$

3. Menentukan nilai t (t hitung)

Melihat nilai t hitung dan membandingkannya dengan t tabel.

4. Menentukan kriteria penerimaan dan penolakan H_0


sebagai berikut :

Jika signifikansi $< 0,05$ maka H_0 ditolak

Jika signifikansi $> 0,05$ maka H_0 diterima









Contoh Uji T Independen

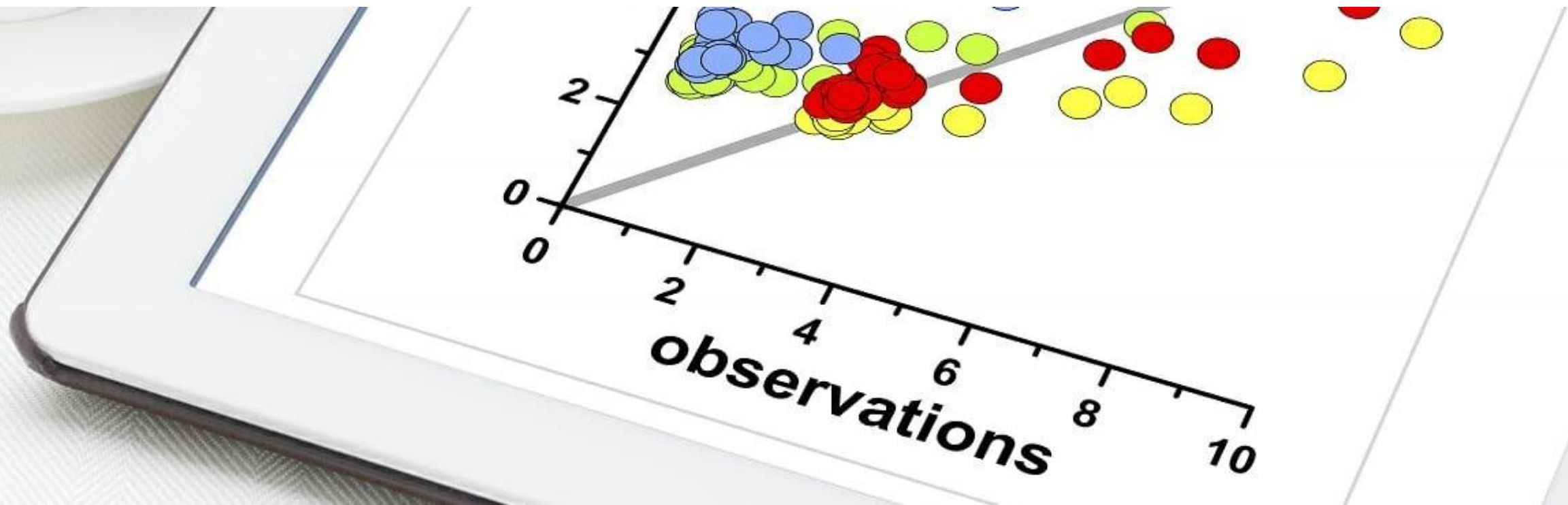
	Gender	Age group	Height (m)	Weight (Kg)
	Female	Adult	1.4	60
	Male	Child	1.2	15
	Male	Adult	1.5	85
	Female	Adult	1.3	74
	Male	Adult	1.6	77
	Female	Elderly	1.5	65

- Misalkan kita memiliki data historis terkait rata-rata tinggi manusia yaitu 1,4m, dan ingin mengetahui apakah rata-rata tinggi manusia pada masa ini berbeda dari masa sebelumnya?
- Definisikan H_0 : Rata-rata tinggi manusia saat ini sama dengan rata-rata tinggi manusia pada masa sebelumnya.
- Definisikan H_a : Rata-rata tinggi manusia saat ini berbeda dengan rata-rata tinggi manusia pada masa sebelumnya.
- Lakukanlah uji T, jika nilai P lebih rendah dari taraf signifikan 0,05 maka tolak hipotesis nol
- Kesimpulan : Rata-rata tinggi manusia saat ini berbeda signifikan dengan rata-rata tinggi manusia pada masa sebelumnya.

Contoh T Test Depend

	Gender	Age group	Height (m)	Weight (Kg)
	Female	Adult	1.4	60
	Male	Child	1.2	15
	Male	Adult	1.5	85
	Female	Adult	1.3	74
	Male	Adult	1.6	77
	Female	Elderly	1.5	65

- Misalkan kita ingin mengetahui apakah ada perbedaan tinggi badan antara pria dan wanita?
- Definisikan H_0 : Tidak terdapat perbedaan tinggi yang signifikan antara pria dan wanita.
- Definisikan H_a : Terdapat perbedaan tinggi yang signifikan antara pria dan wanita.
- Lakukanlah uji T dengan dua variable yaitu gender (jenis kelamin) dan height (tinggi).
- Jika nilai P lebih rendah dari taraf signifikansi 0,05 maka tolak hipotesis nol
- Kesimpulan : Terdapat perbedaan tinggi yang signifikan antara pria dan wanita.



KORELASI






Hubungan Antardua Variabel

Pertanyaan data:

- Apakah terdapat hubungan linear antara kedua variable?
- Apakah variasi skor pada serangkaian skor atau variable diikuti oleh variasi serangkaian skor atau variable yang lain?

Contoh:

- Apakah variasi skor kecerdasan berhubungan dengan variasi skor hasil belajar?
- 



Variasi Korelasi



Variasi searah

- Skor tinggi pada variable x diikuti oleh skor yang tinggi pula pada variable y. Sebaliknya skor yang rendah pada variable x diikuti oleh skor yang rendah pada variable y.
- **Korelasi Positif**




Variasi berlawanan arah

- Skor tinggi pada variable x diikuti skor rendah dari variable y. Sebaliknya, skor rendah pada variable x diikuti oleh skor tinggi pada variable y
- **Korelasi Negatif**



Variasi lainnya

- Variasi pada variable yang satu tidak diikuti secara konsisten oleh variasi skor pada variable yang lain, yakni skor tinggi pada variable X diikuti oleh skor yang tinggi, sedang, dan rendah pada variable y
 - **Tidak Ada Hubungan / Tidak Ada Korelasi**
- 

Koefisien Korelasi

r

-1

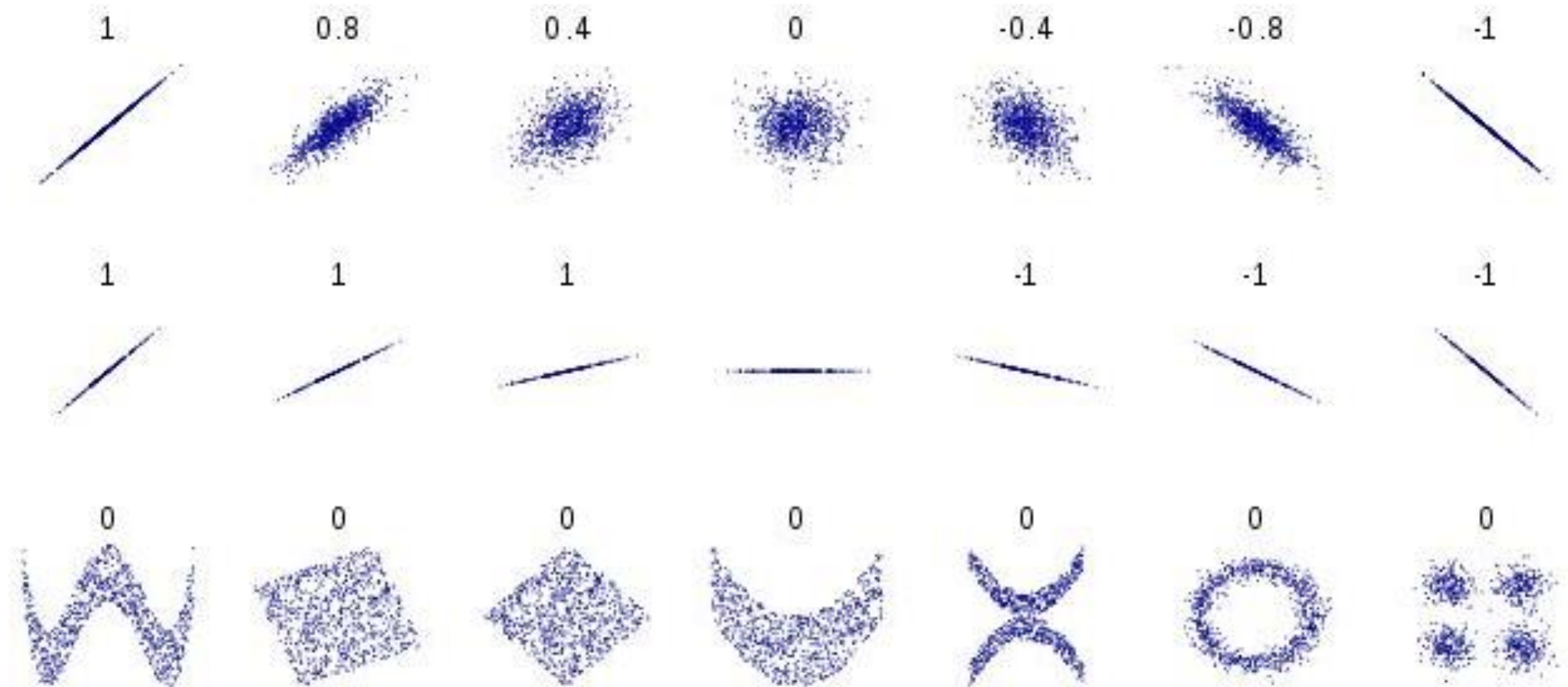
0

1

strong negative
correlation

strong positive
correlation

Bentuk Korelasi dalam Scatter Plot






Teknik Korelasi

Korelasi Pearson Product Moment

- Dua variable yang dihubungkan memiliki data kontinu

Korelasi Point Biserial

- Salah satu variable memiliki data kategorikal, dan variable lainnya berupa data kontinu
- 

Menginterpretasikan Hasil Uji Korelasi

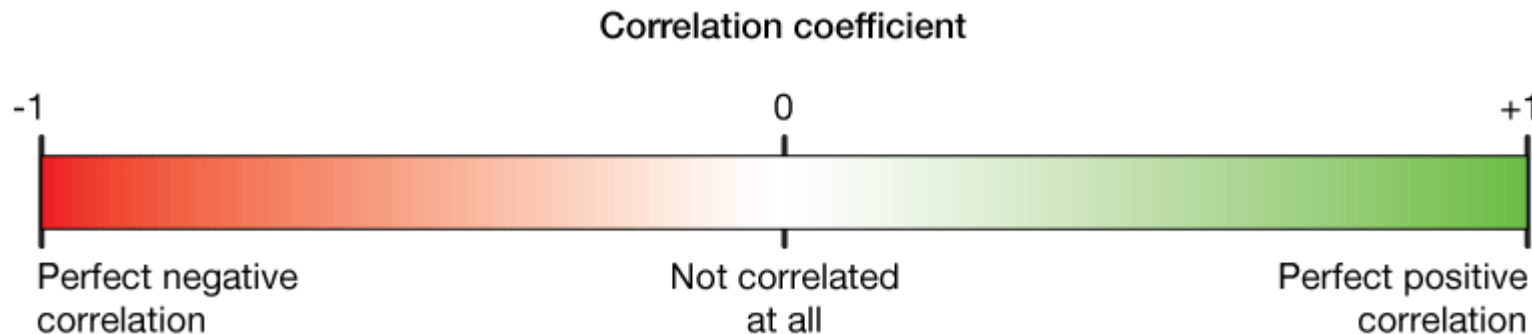
Keluaran dari uji korelasi meliputi dua nilai, yaitu koefisien korelasi dan nilai p

Perhatikan nilai koefisien korelasi (r)

Apakah *positif* / *negatif*?

Seberapa dekat ke angka 1 atau -1?

- Sangat Kuat ($\pm 0,81-1$)
- Kuat ($\pm 0,61-0,80$)
- Sedang ($\pm 0,41-0,60$)
- Lemah ($\pm 0,21-0,40$)
- Sangat lemah ($\pm 0,00-0,20$)





▶ Menginterpretasikan Hasil Uji Korelasi (2)




Perhatikan nilai-p

Apakah lebih rendah dari taraf signifikansi ?

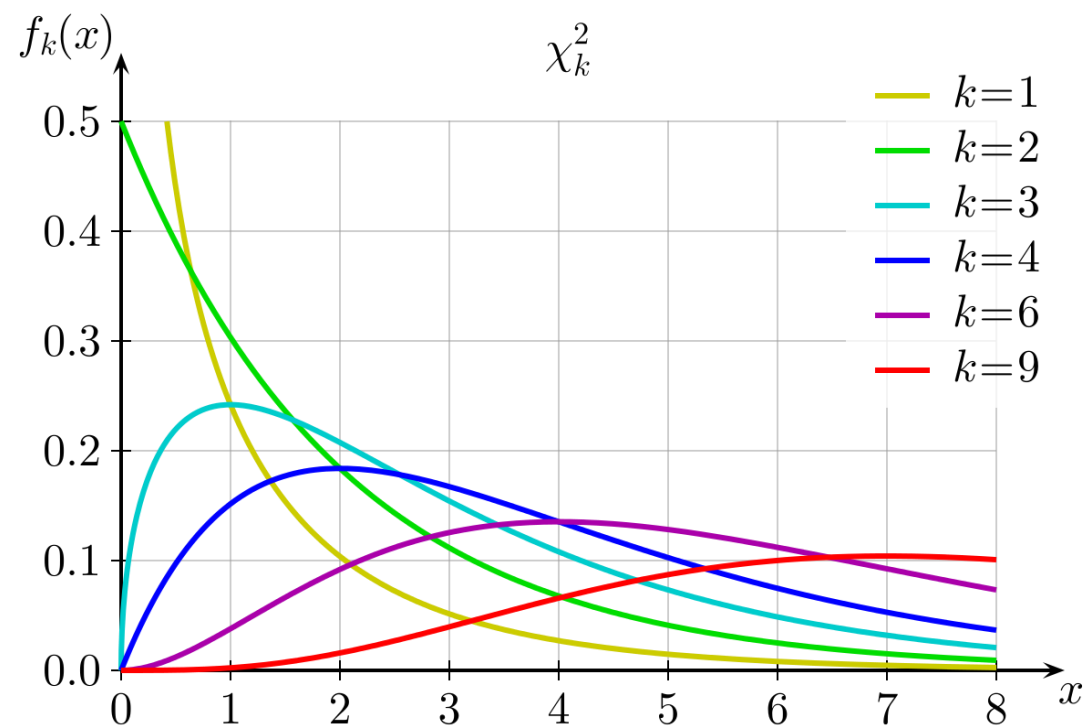
- Jika **ya** → signifikan, berarti ada hubungan.
- Jika **tidak** → tidak signifikan, berarti tidak ada hubungan.



Contoh

	Gender	Age group	Height (m)	Weight (Kg)
	Female	Adult	1.4	60
	Male	Child	1.2	15
	Male	Adult	1.5	85
	Female	Adult	1.3	74
	Male	Adult	1.6	77
	Female	Elderly	1.5	65

- Kita ingin mengetahui apakah ada hubungan antara height (tinggi badan) dan weight (berat badan)?
- Definisikan H_0 : Tidak terdapat hubungan antara tinggi badan dan berat badan.
- Definisikan H_a : Terdapat hubungan antara tinggi badan dan berat badan.
- Lakukanlah uji korelasi dengan dua variable yaitu tinggi dan berat badan.
- Didapat nilai $r = 0,75$,
- Jika nilai p lebih kecil dari taraf signifikansi $0,05$ maka tolak hipotesis nol,
- Kesimpulan : Terdapat hubungan **positif** yang **kuat** antara tinggi dan berat badan.



CHI SQUARE TEST




Hubungan antardua Variabel

Korelasi

Variabel target dan
variable penentu
bertipe numerik

Chi square

Variabel target dan
variable penentu
bertipe kategorikal






Menginterpretasikan Chi-Square Test







- ✓ Keluaran dari Chi-Square Test meliputi **chi square statistic** dan **nilai p**.
- ✓ Perhatikan **chi squared statistic**
- ✓ Perhatikan **nilai p**
 - Apakah nilai P lebih rendah dari taraf signifikansi?
 - Jika ya, maka simpulkan:


“Kedua variabel memiliki asosiasi satu sama lainnya”




Contoh



	Gender	Age group	Height (m)	Weight (Kg)
	Female	Adult	1.4	60
	Male	Child	1.2	15
	Male	Adult	1.5	85
	Female	Adult	1.3	74
	Male	Adult	1.6	77
	Female	Elderly	1.5	65


- Kita ingin mengetahui apakah proporsi pria dan wanita berbeda untuk setiap kelompok usia?
 - Definisikan H_0 : Proporsi pria dan wanita tidak bergantung pada kelompok usia.
 - Definisikan H_a : Proporsi pria dan wanita bergantung pada kelompok usia.
 - Lakukanlah χ^2 square dengan dua variable yaitu gender dan age group.
 - Jika nilai p lebih kecil dari taraf signifikansi 0,05 maka tolak hipotesis nol,
 - Kesimpulan : Proporsi pria dan wanita bergantung pada kelompok usia.
- 



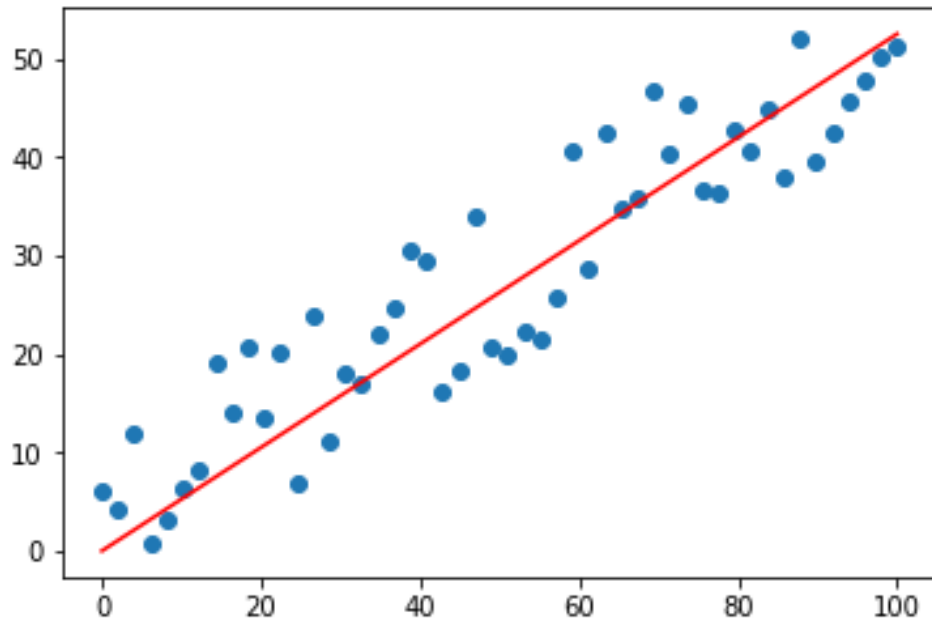


Tujuan utama penelitian ilmiah adalah untuk menjelaskan atau memprediksi fenomena yang teramati (Pedhazur, 1982:16)

Analisis regresi merupakan suatu teknik untuk menaksir atau memprediksi skor suatu variable yaitu kriteria/dependen (y), berdasarkan skor dari variable lain yaitu predictor/independent (x)



Regresi Linear



- Asumsi adanya hubungan linear antara variable kriteria dan variable predictor.
- Dalam bentuk grafis, variable dan konstan dihubungkan melalui suatu garis lurus yang menggambarkan perubahan nilai suatu variable pada sisi vertical sejalan dengan perubahan variable lain pada sisi horizontal.



Model Regresi Linier


$$Y' = a + bX$$

- a = intersep (nilai variable dependen Ketika pengaruh variable independent dihilangkan atau bernilai 0)
- b = slop (nilai perubahan Y sebagai konsekuensi dari perubahan nilai X)





Analisis Regresi

- Tujuan regresi adalah untuk menaksir skor y berdasarkan skor x , sehingga bila x diketahui, y dapat ditaksir.
 - Untuk mengetahui probabilitas dalam prediksi tersebut perlu diuji untuk menentukan apakah dapat diberlakukan generalisasi ke populasi.
 - Ada tiga yang diuji yaitu :
 - regresi y pada x ,
 - proporsi varian, dan
 - koefisien regresi.
- 

Contoh

Ice Cream Sales vs Temperature	
Temperature (in °C)	Ice Cream Sales (in \$)
14.2	215
16.4	325
11.9	185
15.2	332
18.5	406
22.1	522
19.4	412
25.1	614
23.4	544
18.1	421
22.6	445
17.2	408

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.958
R Square	0.917
Adjusted R Square	0.909
Standard Error	38.127
Observations	12

- R Square = 0.917
- → 92% variasi penjualan eskrim dapat ditentukan dari suhu udara di waktu itu.
- Semakin dekat dengan 1, maka garis regresi semakin sesuai dengan data.

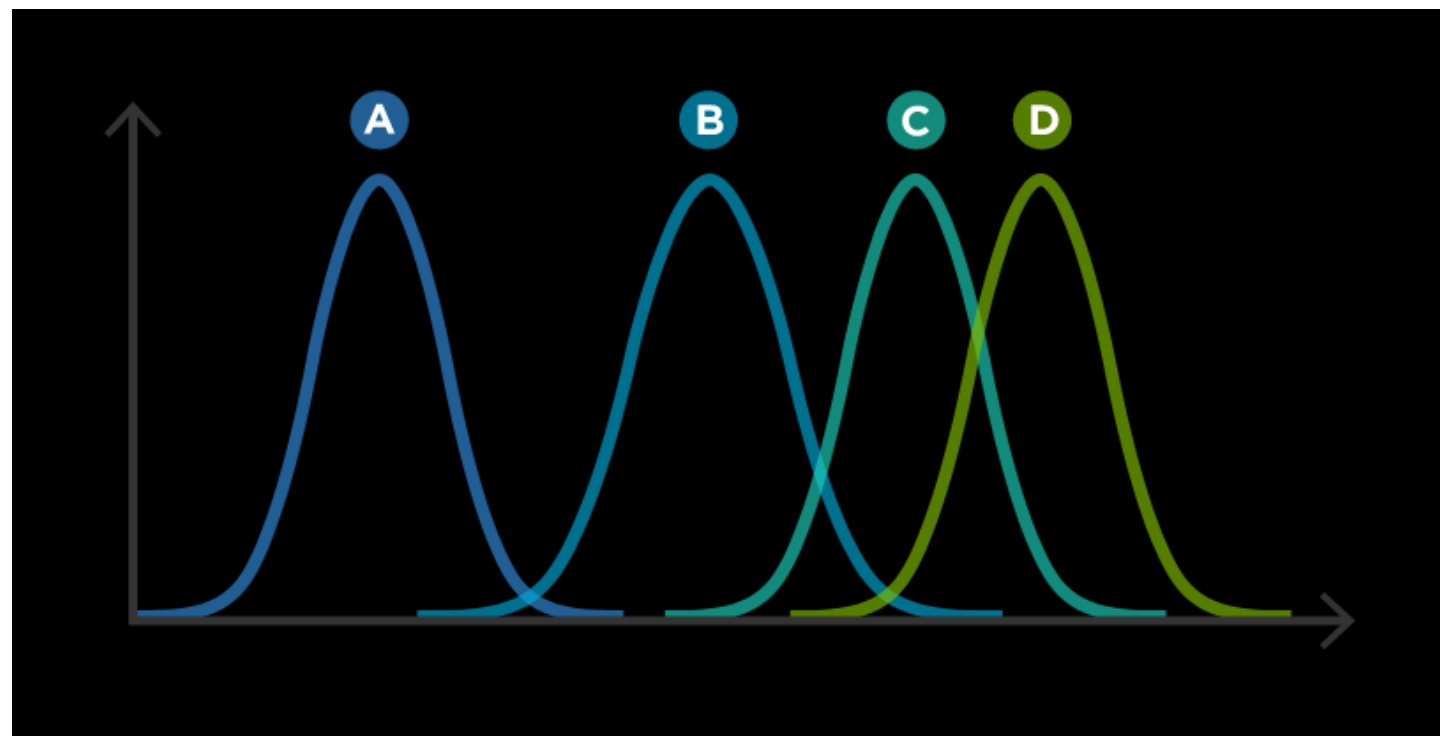
Periksa **nilai significance F**. Jika lebih kecil dari 0,05 maka dapat dilanjutkan.
Jika lebih dari 0,05 maka hentikan analisisnya.
Pastikan **nilai P** lebih kecil dari 0,05.

24									
25	ANOVA								
26		df	SS	MS	F	Significance F			
27	Regression	1	160218.616	160218.616	110.220	0.000			
28	Residual	10	14536.300	1453.630					
29	Total	11	174754.917						
30									
31		Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
32	Intercept	-159.474	54.641	-2.919	0.019	-281.221	-37.727	-281.221	-37.727
33	Temperature (°C)	30.088	2.866	10.499	0.000	23.702	36.473	23.702	36.473
34									



$$y = \text{ice cream sale} = -159,474 + 30.088 \times \text{Temperature of the day}$$

Temperature = 28 °C

$$\text{Ice cream sale} = -159,474 + 30.088 \times 28 = 682,99$$




ANOVA

- 
- Untuk membandingkan kecenderungan skor (khususnya rerata) antarkedua kelompok/kondisi → menggunakan uji T
 - Bagaimana jika kita memiliki lebih dari dua kelompok atau kondisi ? → gunakan uji f dengan ANOVA
- 




ANOVA

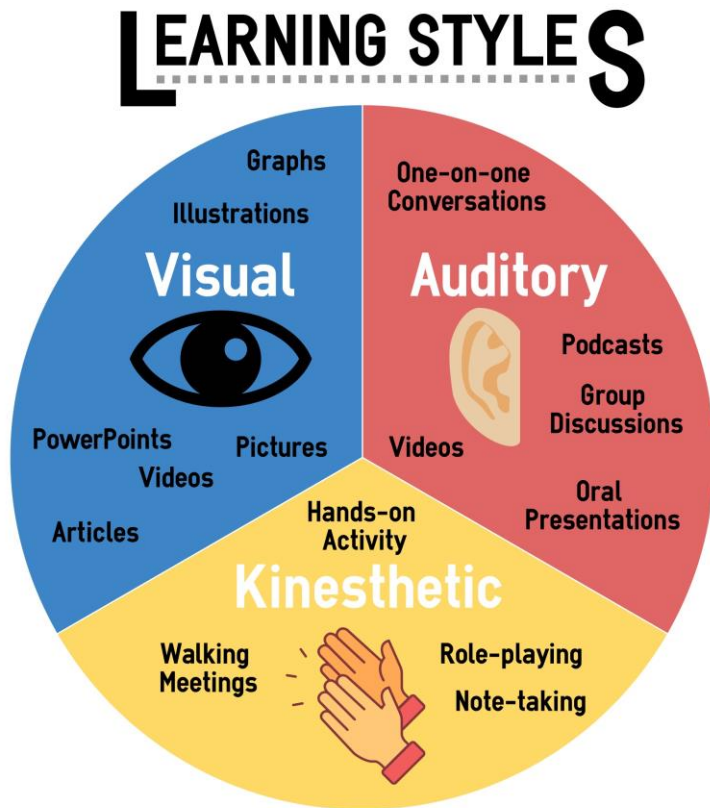
- Analisis varian (ANOVA) sederhana digunakan untuk membandingkan rerata dua kelompok atau lebih secara simultan dan sekaligus mengurangi galat.
 - Tujuan ANOVA adalah untuk menentukan probabilitas bahwa nilai rerata dari beberapa kelompok skor (level) menyimpang dari yang lain semata-mata karena galat sampling.
- 





Hipotesis ANOVA

- Hipotesis dua arah
 - $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$: Tidak ada perbedaan yang nyata antara rata-rata hitung dari n kelompok
 - $H_a : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_n$: Ada perbedaan yang nyata antara rata-rata hitung dari n kelompok
- 

Contoh




- Seorang peneliti ingin menguji apakah ada perbedaan hasil belajar kalkulus pada pokok bahasan integral antara tiga kelompok mahasiswa tingkat 1.
- Kelompok :
 - kelompok mahasiswa dengan gaya belajar visual,
 - kelompok mahasiswa dengan gaya belajar auditorial,
 - kelompok mahasiswa dengan gaya belajar kinestetik
- Sampel acak masing-masing sebanyak 30 mahasiswa.

- 
- $H_0 : \mu_1 = \mu_2 = \mu_3$: Tidak ada perbedaan yang nyata antara rata-rata hitung dari ketiga kelompok gaya belajar
 - $H_a : \mu_1 \neq \mu_2 \neq \mu_3$: Ada perbedaan yang nyata antara rata-rata hitung dari ketiga kelompok gaya belajar
 - Taraf signifikansi : 5% atau 0,05
 - Nilai $p = 0,036$
 - Keputusan : H_0 ditolak
 - Kesimpulan : Ada perbedaan yang nyata antara rata-rata hitung dari ketiga kelompok gaya belajar
- 



Referensi

- Ott, Lyman. (2001). An introduction to statistical methods and data analysis. 5th ed. Duxbury Thomson Learning.
 - Hadjar, Ibnu. (2019). Statistika. untuk ilmu Pendidikan, sosial, dan humaniora. Bandung: PT Remaja Rosdakarya.
- 



Terima kasih

