



**Kampus
Merdeka**
INDONESIA JAYA

Statistik Pendidikan

Sri Lestari Handayani, M.Pd



Instagram
@pgsduhamka



Youtube
pgsd fkip uhamka

AKREDITASI INSTITUSI
TERAKREDITASI-A
SK BAN-PT Nomor :
3128/SK BAN-PT/Akred/PT/XII/2016





**Pengertian Statistik &
Statistika**



**Fungsi dan
Kegunaan
Statistik**



**Pengertian &
Jenis-jenis
Data**



Integrity, Trust, Compassion

<http://fkip.uhamka.ac.id>

Sekilas tentang sejarah Statistik

Statistik: pada awal zaman Masehi, bangsa-bangsa mengumpulkan data untuk mendapatkan informasi mengenai pajak, perang, hasil pertanian, bahkan pertandingan atletik. Sekarang, berkembangnya statistik dapat digunakan untuk memprediksi masa depan dengan data yang sekarang dimiliki. Data tersebut dikumpulkan melalui generalisasi dan peramalan.

DEFINISI

- Statistik berasal dari bahasa Latin, yaitu *status* yang berarti negara —> menggambarkan keadaan dan menyelesaikan masalah yang berhubungan dengan kenegaraan
- Statistik dalam arti sempit : data yang disajikan dalam bentuk angka (data kuantitatif).
- Statistik dalam arti luas :
 - (1) pengumpulan
 - (2) penyusunan
 - (3) analisis
 - (4) interpretasi
 - (5) pengumuman

STATISTIK DALAM ARTI LUAS

1. Pengumpulan data : cara sensus dan cara sampling
 2. Penyusunan data : mengedit, mengklasifikasi, dan tabulasi data
 3. Pengumuman data : disajikan dalam bentuk tabel, grafik, diagram
 4. Analisis data : memperoleh gambaran keseluruhan dari data yg terkumpul
 5. Interpretasi data : untuk memperoleh kesimpulan yg benar
-

Statistik merupakan data mengenai suatu masalah, yang biasanya disusun dalam tabel dan atau diagram, yang melukiskan atau menggambarkan suatu persoalan.

Statistika adalah ilmu (metode ilmiah) yang mempelajari cara-cara mengumpulkan, menyusun, menyajikan, dan menganalisis data serta cara mengambil kesimpulan yang logis sehingga dapat diambil keputusan yang akurat.

Statistik dapat berfungsi sebagai :

1. Bank data untuk menyediakan data untuk diolah dan diinterpretasikan agar dapat digunakan untuk menerangkan keadaan yang perlu diketahui atau diungkap.
2. Alat *quality control* untuk membantu standarisasi dan sekaligus sebagai alat pengawasan
3. Alat analisis, merupakan suatu metode penganalisaan data

KEGUNAAN STATISTIK

1. Memperoleh gambaran tentang kejadian, gejala, atau keadaan
2. Mengikuti perkembangan kejadian, gejala, atau keadaan dari waktu ke waktu
3. Dapat menyusun laporan berupa data kuantitatif dengan teratur, ringkas, dan jelas
4. Dapat mengetahui hubungan antar gejala satu dengan gejala lainnya
5. Dapat meramalkan hal-hal yang akan terjadi di masa mendatang



Kegiatan STATISTIKA :

- mengumpulkan
- menyusun
- menyajikan, dan
- menganalisis data dengan metode tertentu
- menginterpretasikan hasil analisis

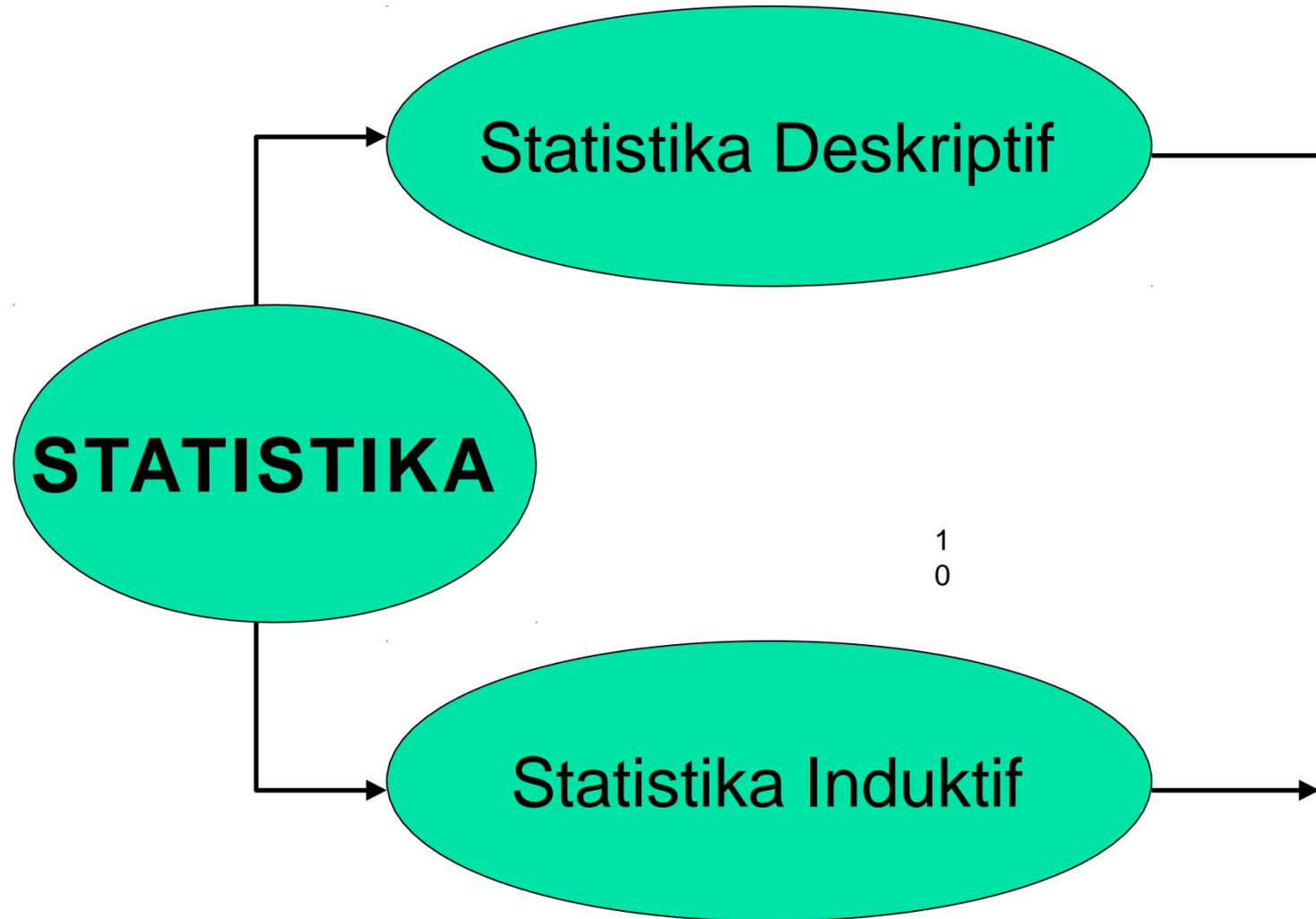
Jenis Statistik

STATISTIKA DESKRIPTIF :

Statistik yang berkenaan dengan pengumpulan, pengolahan, dan penyajian sebagian atau seluruh data (pengamatan) untuk memberikan informasi: rerata, prosentase, dll **tanpa pengambilan kesimpulan**

STATISTIKA INFERENSI :

Setelah data dikumpulkan, dilakukan analisis data, dilakukan interpretasi **serta pengambilan kesimpulan**. Statistika inferensi akan menghasilkan generalisasi hasil penelitian (jika sampel representatif)



Materi:

1. Penyajian data
2. Ukuran pemusatan
3. Ukuran penyebaran
4. Angka indeks
5. Deret berkala dan peramalan

Materi:

1. Probabilitas dan teori keputusan
2. Metode sampling
3. Teori pendugaan
4. Pengujian hipotesa
5. Regresi dan korelasi
6. Statistika nonparametrik

2. Statistika & Metode Ilmiah

METODE ILMIAH :

Adalah cara mencari kebenaran dengan resiko keliru paling kecil.

Berikut adalah LANGKAH-LANGKAH METODE ILMIAH :

1. Merumuskan masalah
2. Melakukan studi literature atau kajian pustaka
3. Membuat dugaan-dugaan, pertanyaan-pertanyaan atau hipotesis
- 4. Peran Statistika: Mengumpulkan dan mengolah data, menguji hipotesis, atau menjawab pertanyaan**
5. Mengambil kesimpulan

INSTRUMEN

SAMPEL

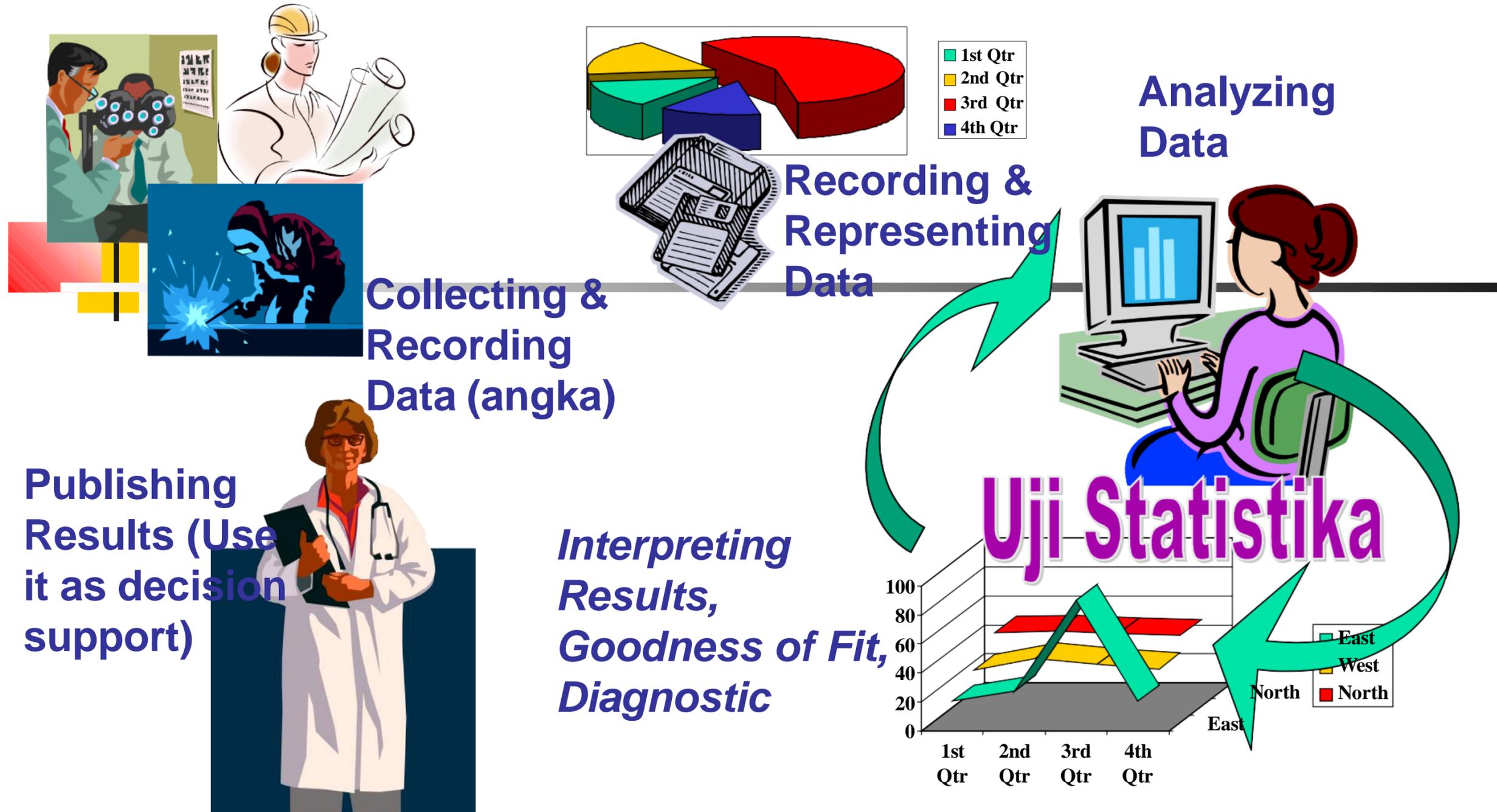
SIFAT DATA

VARIABEL

METODE ANALISIS



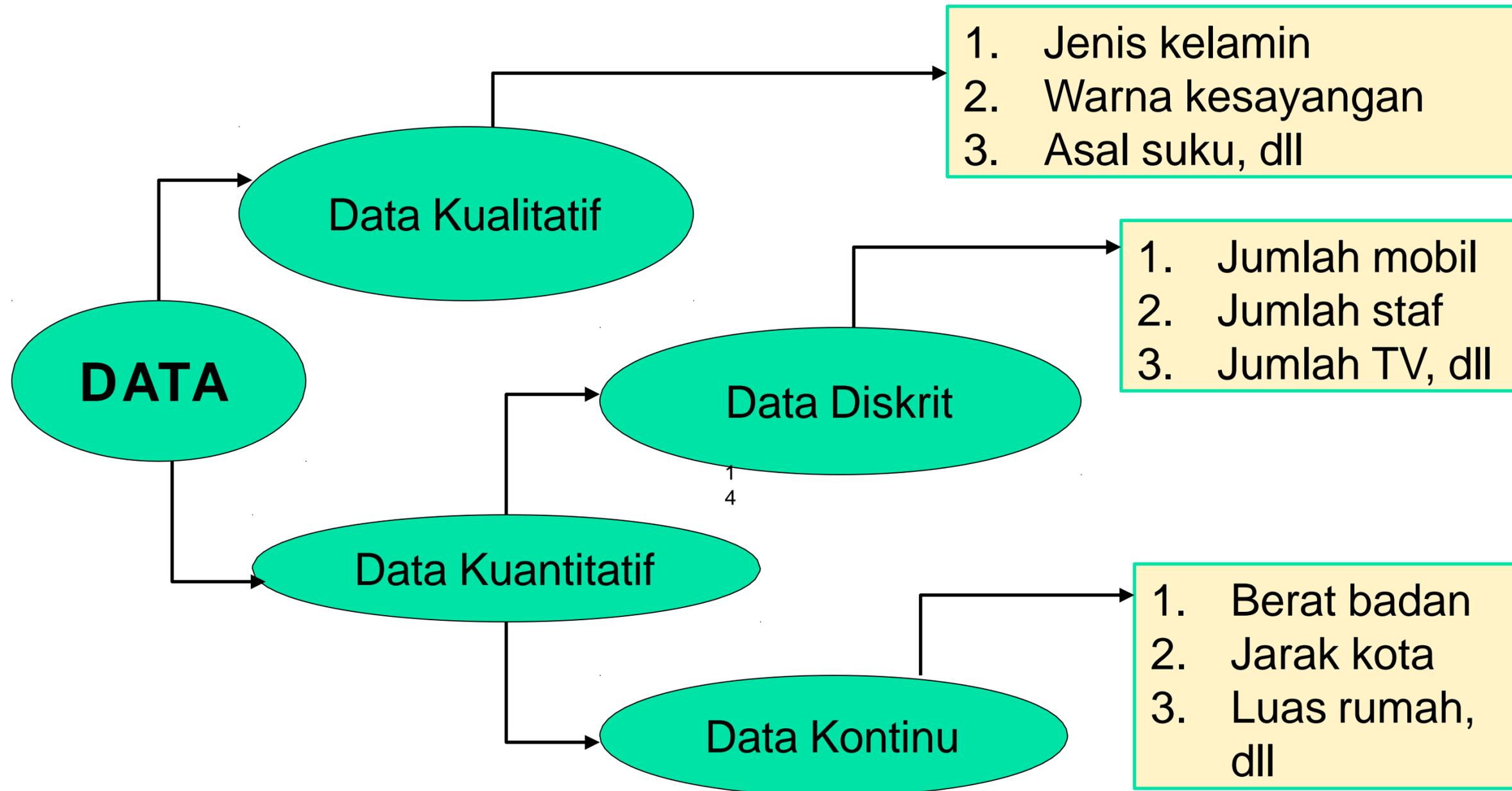
LANGKAH-LANGKAH PENTING



PENGERTIAN DATA

- Data adalah bentuk jamak dari *datum*.
- Data merupakan kumpulan fakta atau angka atau segala sesuatu yang dapat dipercaya kebenarannya sehingga dapat digunakan sebagai dasar menarik suatu kesimpulan.
- Tidak semua angka dapat disebut data statistik. Angka dapat disebut data statistik apabila dapat menunjukkan suatu ciri dari suatu penelitian yang bersifat agregatif, serta mencerminkan suatu kegiatan lapangan tertentu.
- Agregatif :
 - a) mengenai 1 individu, namun penelitian dilakukan berulang kali
 - b) Dilakukan satu kali, namun yang diteliti lebih dari satu individu

JENIS-JENIS DATA



A. Berdasarkan Sifat

Data kualitatif yaitu data yang berbentuk kalimat, kata atau gambar.

Data kuantitatif yaitu data berupa angka.

Data Kuantitatif

Data diskrit yaitu data data yang tidak dikonsepsikan adanya nilai- nilai di antara data (bilangan) lain yang terdekat.

Data kontinu yaitu data yang didapat dari hasil pengukuran. Data hasil pengukuran diperoleh dari tes, kuesioner ataupun alat ukur lain yang sudah terstandar misalnya timbangan, panjang ataupun data psikologis yang lain.

JENIS DATA BERDASARKAN SKALA PENGUKURAN

Skala Nominal

Angka yang diberikan hanya sebagai label saja.

Contoh: pria = 1, wanita = 2

Skala Ordinal

Angka mengandung pengertian tingkatan.

Contoh: ranking 1, 2, dan 3.

Ranking 1 menunjukkan lebih tinggi dari ranking 2 dan 3.

Skala Interval

Angka mengandung sifat ordinal ¹/₇ dan mempunyai jarak atau interval.

Contoh: 1. Saham sangat prospektif dengan harga saham Rp736-878,
2. saham prospektif Rp592-735.

Skala Rasio

Angka mempunyai sifat nominal, ordinal dan interval serta mempunyai nilai absolut dari objek yang diukur.

Contoh: bunga BCA 7% dan bunga Mandiri 14%, maka bunga Mandiri 2 kali bunga BCA.

B. Berdasarkan Skala

Data nominal adalah data mengelompokkan obyek atau peristiwa dalam berbentuk kategori

Data ordinal adalah data yang menunjukkan perbedaan tingkatan subjek secara kuantitatif.

Data interval adalah data yang memiliki jarak yang sama antar datanya akan tetapi tidak memiliki nol mutlak. Nol mutlak artinya tidak dianggap ada.

- **Data rasio** adalah data yang bersekala rasio hampir sama dengan data interval, yakni keduanya memiliki ketiga sifat di atas (menunjukkan klasifikasi dan kedudukan subjek dalam suatu kelompok)

C. Berdasarkan Sumbernya

Data Primer adalah data yang diperoleh secara langsung dengan melakukan sendiri pengumpulan terhadap obyek.

Data Sekunder adalah data yang diperoleh dari olahan data primer.

Data Tersier adalah data yang diperoleh dari olahan data sekunder.

Data Kuarter adalah data yang diperoleh dari olahan data primer.

D. Berdasarkan Bentuk Angkanya

Data tunggal adalah data statistik yang angka-angkanya merupakan satu unit atau satu kesatuan, tidak dikelompokkan.

Data kelompok adalah data statistik tiap unitnya terdiri dari sekelompok angka, 51-55, 56-60, 61-65, dst.

E. Berdasarkan Waktu Pengambilannya

Data seketika adalah data statistik yg mencerminkan keadaan pada suatu waktu saja.
Contoh : Pada semester gazal 2013/2014.

Data urutan waktu adalah data statistik yg mencerminkan keadaan dari waktu ke waktu secara berurutan.
Contoh: Jumlah mahasiswa FKIP Unsri yang lulus dari tahun 2010-2013



A. PARAMETER :

- Statistik **PARAMETRIK** : berhubungan dengan inferensi statistik yang membahas parameter-parameter populasi; jenis data interval atau rasio; distribusi data normal atau mendekati normal.
- Statistik **NONPARAMETRIK** : inferensi statistik membahas parameter- parameter populasi; jenis data nominal atau ordinal; distribusi data tidak diketahui atau tidak normal

B. JUMLAH VARIABEL :

- Analisis **UNIVARIAT** : hanya ada 1 pengukuran (variabel) untuk n sampel atau beberapa variabel tetapi masing-masing variabel dianalisis sendiri-sendiri..
- Analisis **BIVARIAT**
- Contoh : korelasi motivasi dengan pencapaian akademik
- Analisis **MULTIVARIAT** : dua atau lebih pengukuran (variabel) untuk n sampel di mana analisis antar variabel dilakukan bersamaan. Contoh : pengaruh motivasi terhadap pencapaian akademik yang dipengaruhi oleh faktor latar belakang pendidikan orang tua, faktor sosial ekonomi, faktor sekolah.





AKREDITASI INSTITUSI
TERAKREDITASI-A
SK BAN-PT Nomor :
3128/SK BAN-PT/Akred/PT/XII/2016



Terimakasih



Integrity, Trust, Compassion

<http://fkip.uhamka.ac.id>

**Penyajian Data
&
Distribusi Frekuensi**



Pengertian Penyajian Data

Penyajian data adalah komunikatif dan lengkap, dalam arti data yang disajikan dapat menarik perhatian pihak lain untuk membacanya dan mudah memahami isinya.

Ada beberapa penyajian data yang akan dikemukakan di sini adalah penyajian dengan tabel (daftar), grafik, dan diagram.



Tabel

Tabel adalah yang berisi ikhtisar sejumlah data-data informasi yang biasanya berupa huruf maupun angka.

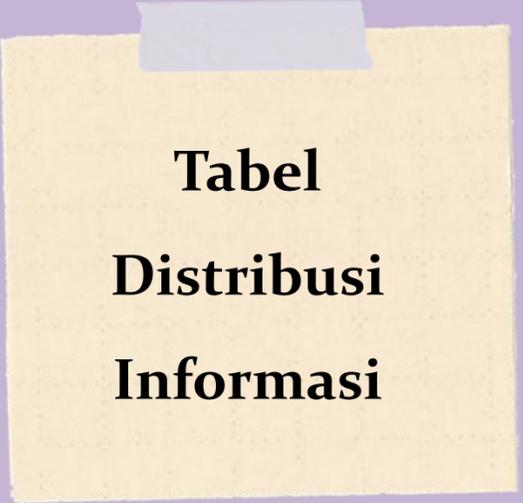
Jenis-jenis tabel :



**Tabel
Biasa**



**Tabel
Kontingensi**



**Tabel
Distribusi
Informasi**

Tabel Biasa

Tabel biasa sering digunakan untuk bermacam keperluan baik dalam bidang ekonomi, sosial, budaya, dan lain-lain untuk menginformasikan data dari hasil penelitian atau penyelidikan.

Contoh

Jumlah mahasiswa Bahasa Inggris per angkatan

NO	Angkatan	Jumlah
1	2016	40
2	2017	55
3	2018	60
4	2019	70
5	2020	90

Judul Kolom

Badan tabel yang berisi data dalam bentuk angka

Judul Baris

Sumber : data fiktif

Sumber data diperoleh

Tabel Kontingensi

Tabel kontingensi adalah tabel yang menunjukkan atau memuat data sesuai dengan rinciannya dan digunakan khusus data yang terletak antara baris dan kolom berjenis kategori

Contoh

Tabel produksi batu bara OPEC, Uni Soviet, dan Dunia tahun 2001, 2002, 2003, 2004, dan 2005

Tahun	OPEC	Uni Soviet	Dunia	Jumlah
2001	9.934	3.600	20.174	33.708
2002	11.240	3.822	21.831	36.893
2003	11.468	4.013	22.672	38.153
2004	10.914	4.204	22.897	38.015
2005	11.205	4.307	23.666	39.170
Jumlah	54. 761	19.946	111.240	185.947

Tabel Distribusi Frekuensi

Tabel distribusi frekuensi adalah pengelompokan data menjadi tabulasi data dengan memakai kelas-kelas data dan dikaitkan dengan masing-masing frekuensinya.

Istilah-istilah dalam distribusi frekuensi :

1. Kelas adalah penggolongan data yang dibatasi oleh nilai terendah dan nilai tertinggi dalam suatu kelas.
2. Interval Kelas adalah Lebar dari sebuah kelas dan dihitung dari perbedaan antara kedua tepi kelasnya.
3. Batas Kelas (class limit), nilai batas tiap kelas dalam sebuah distribusi frekuensi dan dipergunakan sebagai pedoman guna memasukkan angka-angka hasil observasi ke dalam kelas-kelas yang sesuai.
4. Batas Kelas Bawah (lower class limit) adalah angka pada kolom kelas yang letaknya disebelah kiri.
5. Batas Kelas Atas (upper class limit) adalah angka pada kolom kelas yang letaknya disebelah kanan.
6. Tepi Kelas (class boundaries/true limits) :
 1. Tepi Kelas Bawah (lower class bounderis), Batas kelas pertama yang benar-benar dimiliki oleh distribusi frekuensi tersebut, yaitu batas kelas bawah dikurangi 1 digit dibelakang koma.
 2. Tepi Kelas Atas (upper class bounderis), Batas kelas kedua yang benar-benar dimiliki oleh distribusi frekuensi tersebut, yaitu batas kelas atas ditambah 1 digit dibelakang koma.
$$\text{Tepi atas} = \text{batas atas} + 0,5$$
$$\text{tepi bawah} = \text{batas bawah} - 0,5$$
7. Lebar kelas = tepi atas-tepi bawah
8. Mid Point (titik tengah), rata-rata dari kedua batas kelasnya/kelas limitnya. Titik tengan = $\frac{1}{2}$ (batas atas + batas bawah)

Teknik pembuatan distribusi frekuensi

- a. Urutkan data dari terkecil sampai terbesar
- b. Hitung jarak atau rentangan (R). $(R) = \text{data tertinggi} - \text{data terendah}$
- c. Hitung jumlah kelas (K) dengan struger. $K = 1 + 3,3 \log n$
- d. Hitung panjang kelas interval (P). $P = R/K$
- e. Tentukan batas data terendah (ujung data pertama) dan menentukan batas atas dan bawah kelas.
- f. Buat table sementara (tabulasi table) dengan cara dihitung satu demi satuyang sesuai dengan urutan interval kelas.
- g. Membuat table distribusi frekuensi dengan cara memindahkan semua angka frekuensi.

Bentuk-bentuk distribusi frekuensi, yaitu :

**Distribusi
Frekuensi
Relatif**

**Distribusi
Frekuensi
Kumulatif**

**Distribusi
Frekuensi
kumulatif
Relatif**

Contoh Tabel Distribusi Frekuensi

Tabel Distribusi Frekuensi Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
60-64	2
65-69	6
70-74	15
75-79	20
80-84	16
85-89	7
90-94	4
Jumlah	70

Distribusi Frekuensi Relatif

Distribusi frekuensi relatif adalah distribusi frekuensi yang nilai frekuensinya tidak dinyatakan dalam bentuk angka mutlak

(nilai mutlak), akan tetapi setiap kelasnya dinyatakan dalam bentuk angka persentase (%) atau angka relatif.

Contoh

Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
60-64	2,857 %
65-69	2,571 %
70-74	21,429 %
75-79	28,571 %
80-84	22,857 %
85-89	10,000 %
90-94	5,714 %
Jumlah	100 %

Cara perhitungan

$$f_{\text{relatif ke-i}} = f_{\text{ke-i}}/n \times 100\%$$

$$\text{Contoh : } f_{\text{relatif ke-2}} = 2/70 \times 100\% = 2,857\%$$

Distribusi Frekuensi Kumulatif

Distribusi frekuensi kumulatif adalah distribusi frekuensi yang nilai frekuensinya diperoleh dengan cara menjumlahkan

frekuensi (berdasarkan tabel distribusi frekuensi mutlak)

Contoh

Tabel Distribusi Kumulatif (kurang dari)

Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
Kurang dari 60	0
Kurang dari 65	2
Kurang dari 70	8
Kurang dari 75	23
Kurang dari 80	43
Kurang dari 85	59
Kurang dari 90	66
Kurang dari 95	70

Tabel Distribusi Kumulatif (atau lebih)

Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
60 atau lebih	70
65 atau lebih	68
70 atau lebih	62
75 atau lebih	47
80 atau lebih	27
85 atau lebih	11
90 atau lebih	4
95 atau lebih	0

Distribusi Frekuensi Relatif Kumulatif

Distribusi frekuensi relatif kumulatif adalah distribusi frekuensi yang nilai frekuensi kumulatif diubah menjadi nilai frekuensi relatif atau dalam bentuk presentase (%) atau dengan rumus

$$f_{\text{kum}}(\%) \text{ ke-}i = \frac{f_{\text{kum}} \text{ ke-}i}{n} \times 100\%$$

Contoh : $f_{\text{kum}}(\%) \text{ ke-}0 = \frac{0}{70} \times 100\% = 0,000\%$

Contoh

Tabel Distribusi Kumulatif Relatif (kurang dari)
Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
Kurang dari 60	0,000%
Kurang dari 65	2,857%
.....
Kurang dari 90	94,286%
Kurang dari 95	100,00%

Tabel Distribusi Kumulatif Relatif (atau lebih)
Nilai Ujian Statistik UHAMKA Tahun 20016

Nilai Interval	f
60 atau lebih	100,00%
65 atau lebih	97,143%
.....
90 atau lebih	5,714%
95 atau lebih	0,000%

Grafik

Grafik adalah lukisan pasang surutnya suatu keadaan dengan garis atau gambar (tentang turun naiknya hasil statistik)

Jenis-jenis grafik :



Histogram



**Poligon
Frekuensi**



Ogive

Histogram

Histogram adalah grafik yang menggambarkan suatu distribusi frekuensi dengan bentuk segi empat.

Langkah-langkah :

a. Buatlah absis dan ordinat

Absis : sumbu mendatar (X) menyatakan nilai

Ordinat : sumbu tegak (Y) menyatakan frekuensi

b. Berilah nama pada masing-masing sumbu dengan cara sumbu absis diberi nama nilai dan ordinat diberi nama frekuensi.

c. Buatlah skala absis dan ordinat.

d. Buatlah batas kelas dengan cara :

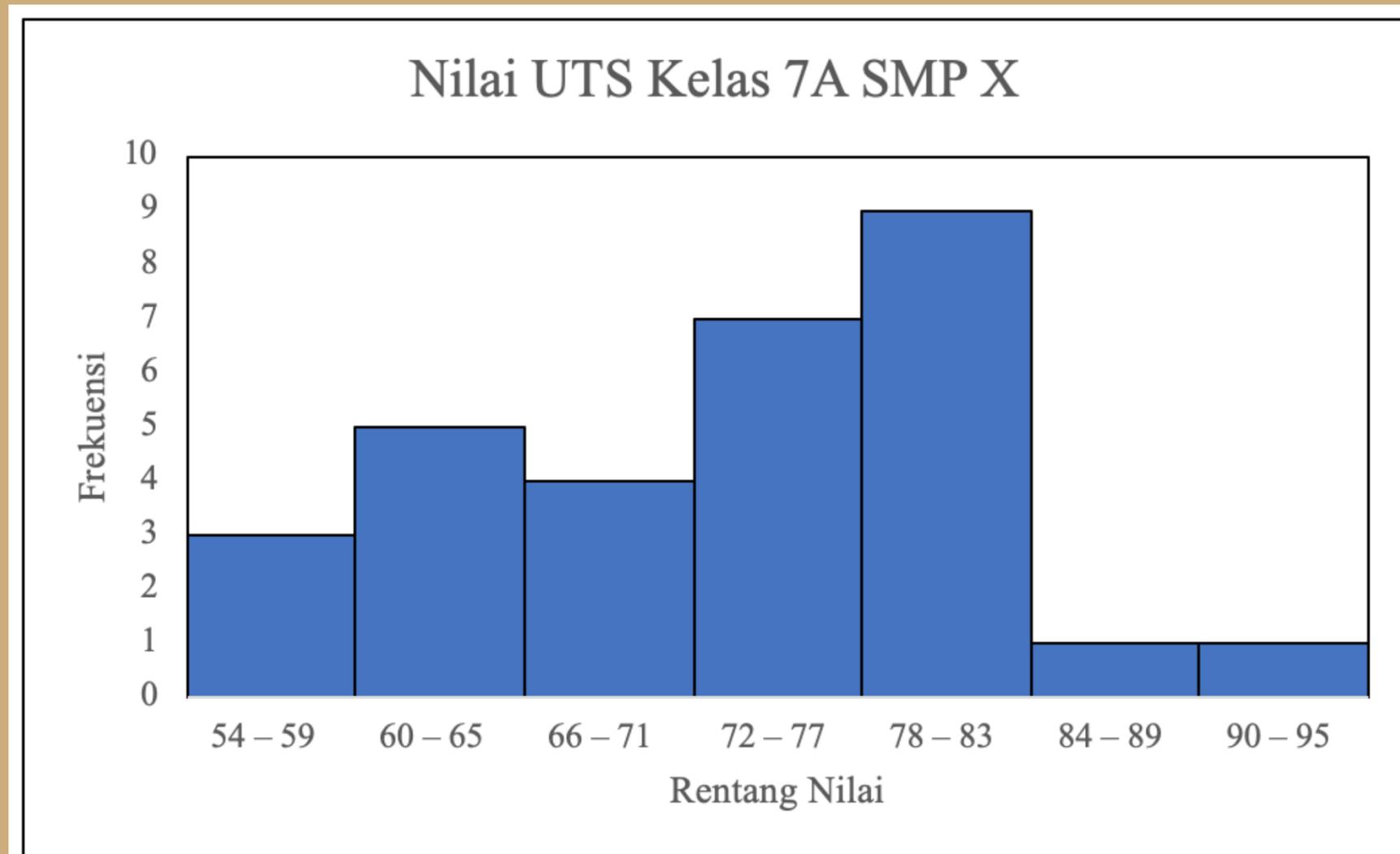
- ujung bawah interval kelas di kurangi 0,5.

- ujung atas interval kelas ditambah 0,5 atau $(U_b + U_a) : 2$

e. Membuat tabel distribusi frekuensi untuk membuat histogram.

f. Membuat grafik histogram.

Contoh Grafik Histogram



Gambar 2. Grafik Histogram.

Poligon Frekuensi

Poligon frekuensi adalah grafik garis yang menghubungkan nilai tengah tiap sisi atas yang berdekatan dengan nilai tengah jarak frekuensi mutlak masing-masing.

Perbedaan antara histogram dan poligon :

- a. Histogram menggunakan batas kelas sedangkan poligon menggunakan titik tengah.
- b. Grafik histogram berwujud segi empat sedang grafik poligon berwujud garis-garis atau kurva yang saling berhubungan satu dengan yang lainnya.

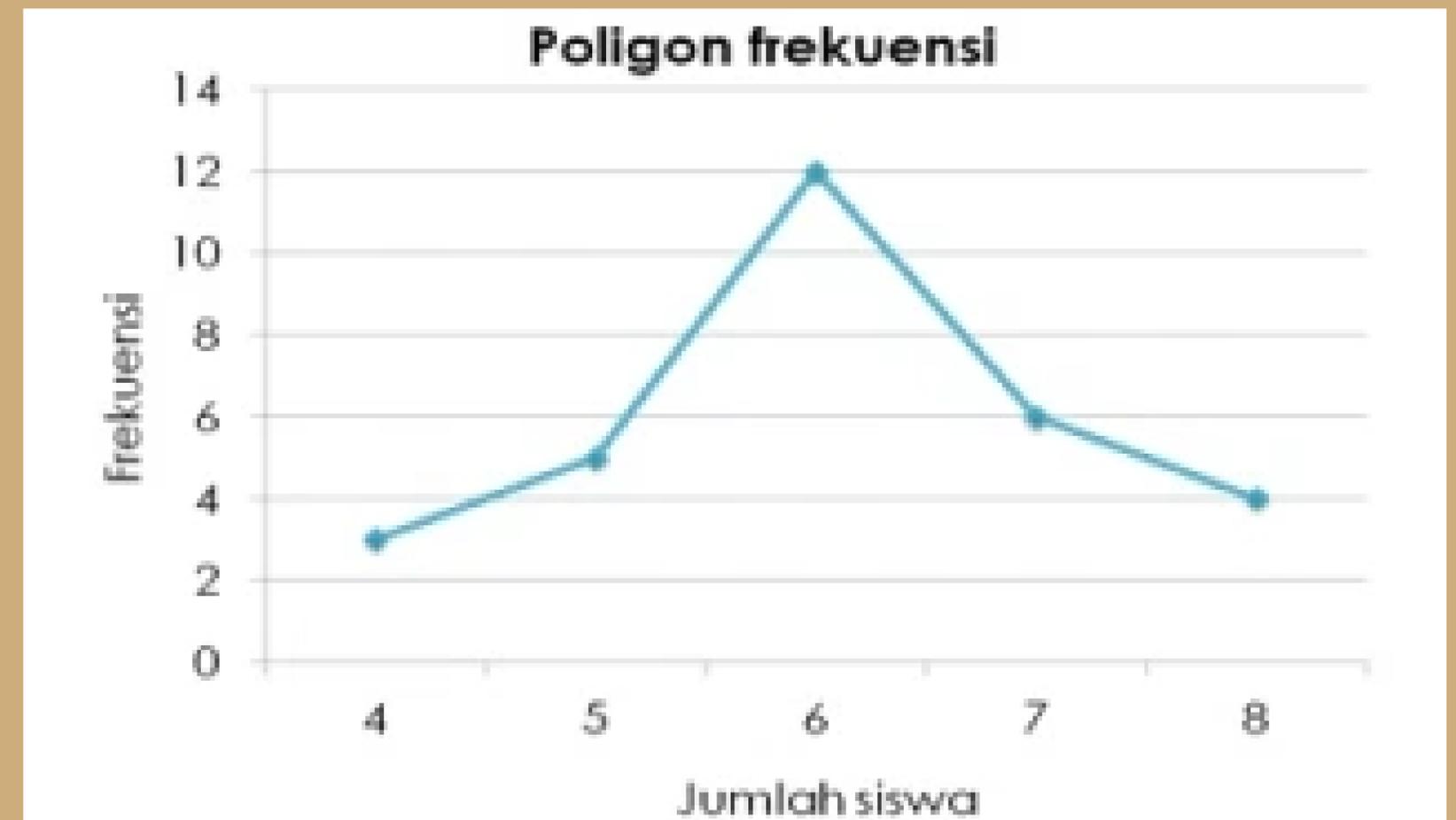
Contoh Grafik Poligon Frekuensi

Contoh pada data tunggal

Tentukan poligon frekuensi dari data di bawah ini :

Jumlah siswa terlambat	4	5	6	7	8	Jumlah
Frekuensi	3	5	12	6	4	30
Frekuensi relatif	0.1	0.17	0.4	0.2	0.13	1

Dan Poligon Frekuensinya, seperti di bawah ini :



Ogive

Ogive adalah distribusi frekuensi kumulatif yang menggambarkan diagramnya dalam sumbu tegak dan mendatar (eksponensial).

Perbedaan antara poligon frekuensi dan Ogive :

- Ogive menggunakan batas kelas dan poligon menggunakan titik tengah
- Grafik ogive menggambarkan distribusi frekuensi kumulatif kurang dari dan distribusi frekuensi kumulatif atau lebih, sedangkan grafik poligon mencantumkan nilai frekuensi tiap-tiap variabel.

Grafik ogive berguna untuk sensus penduduk, perancang mode, perkembangan dan penjualan saham, dan lain-lain.

Cara membuat grafik ogive :

- a. Grafik ogive diambil dari distribusi kumulatif kurang dari dan distribusi kumulatif atau lebih
- b. Cara ogive diambil dari table distribusi frekuensi ditambah satu kolom frekuensi meningkat dengan menggunakan batas kelas

Diberikan data kelompok untuk suatu nilai

Matematika yang terdiri dari enam kelas.

Data tersebut diberikan dalam bentuk bentuk tabel data kelompok. Dengan panjang setiap

kelas adalah 5 (lima).

Rentang	f_i
71 – 75	2
76 – 80	4
81 – 85	12
86 – 90	13
91 – 95	6
96 – 100	3

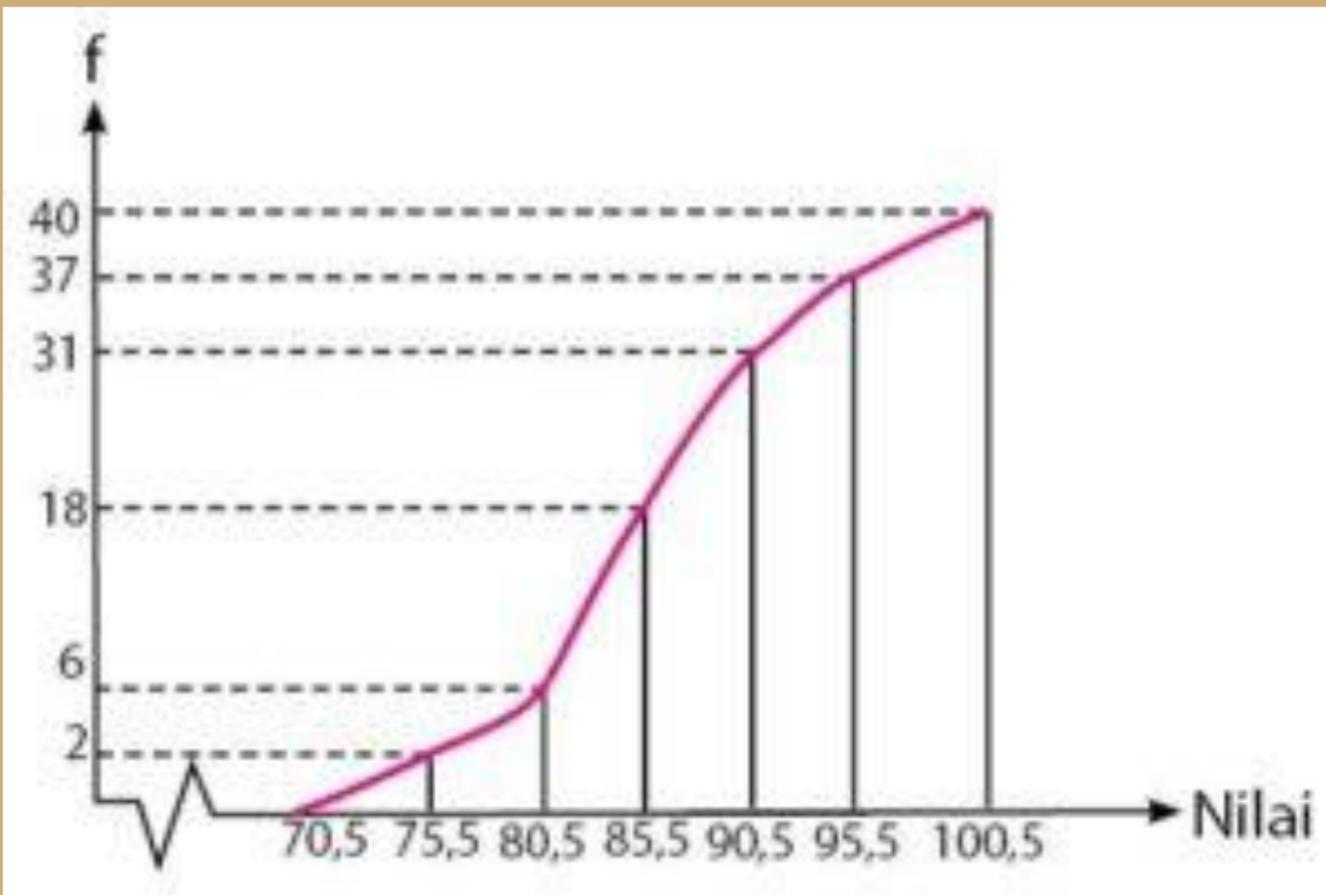
Langkah pertama untuk membuat ogive positif adalah menentukan batas atas masing-masing kelas dan data frekuensi kumulatif kurang dari sesuai dengan data pada tabel kelompok yang diberikan sebelumnya.

Batas atas kelas diperoleh dengan cara menambahkan nilai tertinggi pada kelas tersebut dengan 0,5. Misalnya pada kelas pertama, nilai tertingginya adalah 75. Batas atas untuk kelas pertama adalah $75 + 0,5 = 75,5$.

Nilai frekuensi komulatif kurang dari diperoleh dengan cara menjumlahkan frekuensi setiap kelas dengan semua frekuensi semua kelas di atasnya. Misalnya, akan dicari nilai frekuensi komulatif kurang dari untuk kelas dengan rentang 81 – 85. Frekuensi (f_i) kelas tersebut adalah 12, dan dua kelas sebelumnya memiliki nilai frekuensi 2 dan 4. Sehingga, nilai frekuensi komulatif kurang dari untuk kelas 81 – 85 adalah $2 + 4 + 12 = 18$. Selengkapnya lihat pada tabel di bawah.

Rentang	Batas Atas Kelas	f_i	Frek komulatif kurang dari (f_{kk})
71 – 75	75,5	2	2
76 – 80	80,5	4	6
81 – 85	85,5	12	18
86 – 90	90,5	13	31
91 – 95	95,5	6	37
96 – 100	100,5	3	40

Selanjutnya menyiapkan bidang kartesius yang memuat nilai pada sumbu x dan data frekuensi pada sumbu y. Dan Pasangkan data yang sesuai antara data nilai batas atas setiap kelas dan frekuensi kurang dari. Kemudian hubungan setiap titik sehingga membentuk sebuah kurva. Sehingga akan diperoleh ogive positif di bawah ini.



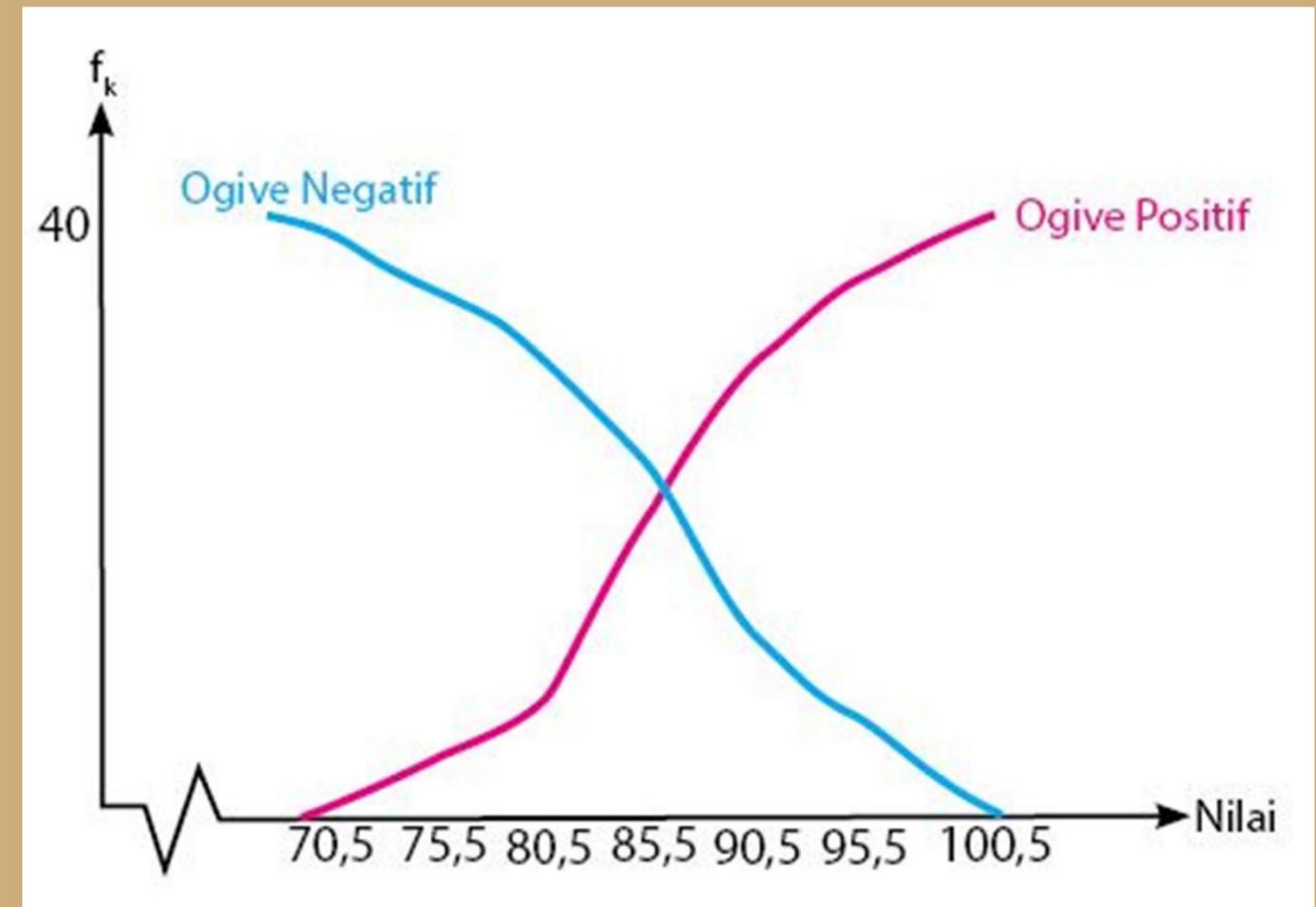
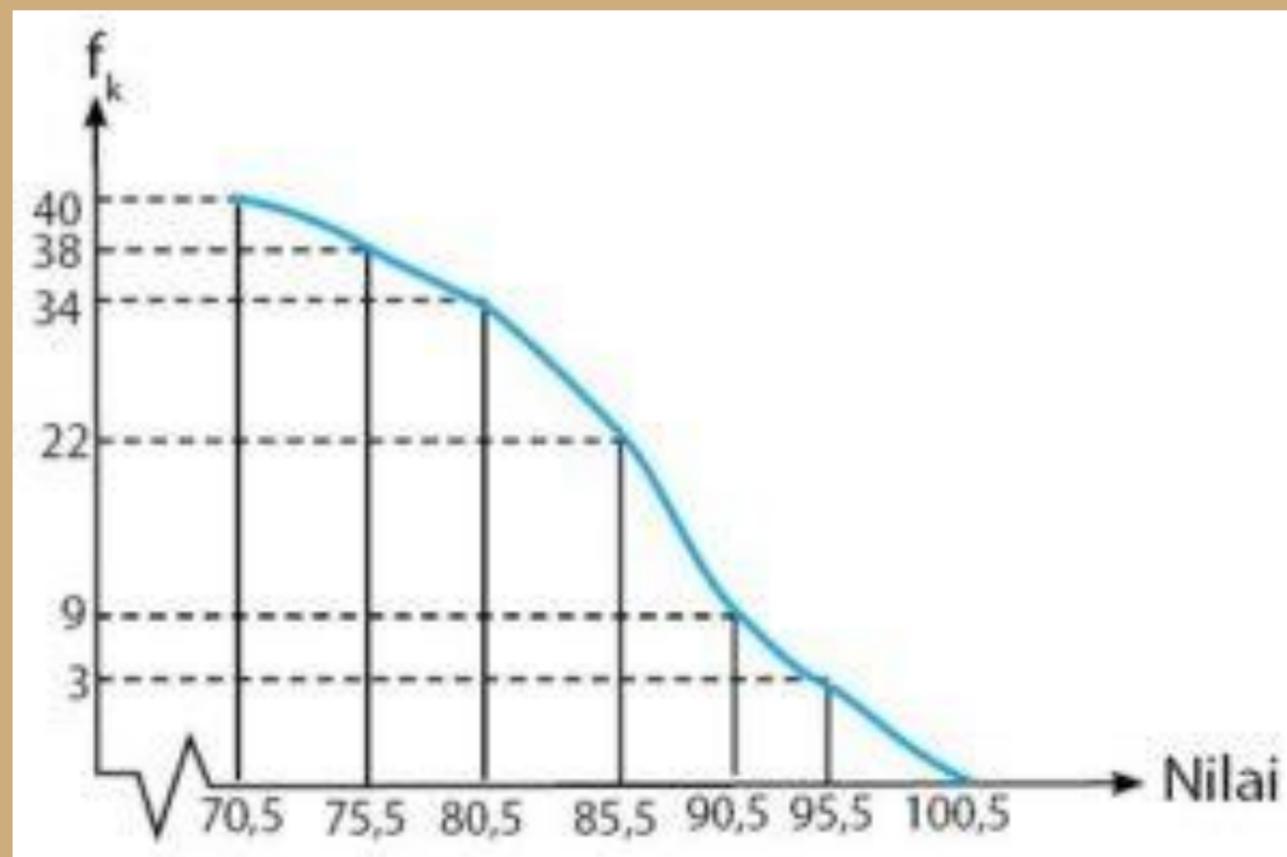
Langkah-langkah yang digunakan untuk membuat ogive negatif sama dengan cara membuat ogive positif. Perbedaannya terletak pada data batas kelas dan frekuensi komulatif yang digunakan. Pada ogive positif menggunakan batas atas kelas dan frekuensi komulatif kurang dari. Sedangkan pada ogive negatif menggunakan batas bawah kelas dan frekuensi komulatif lebih dari.

Berikut ini adalah data batas bawah untuk setiap kelas dan frekuensi komulatif lebih dari.

Rentang	Batas Bawah Kelas	F_i	Frek komulatif lebih dari
71 – 75	70,5	2	40
76 – 80	85,5	4	38
81 – 85	80,5	12	34
86 – 90	85,5	13	22
91 – 95	90,5	6	9
96 – 100	99,5	3	3

Selanjutnya dengan langkah sama dengan cara menggambar ogive positif sebelumnya. Yaitu menyesuaikan data antara batas bawah masing-masing kelas dan frekuensi kumulatif lebih dari. Akan diperoleh gambar ogive negatif seperti gambar berikut ini.

Maka hasil ogive positif dan negatif, seperti di bawah ini



Diagram

Diagram adalah gambaran untuk memperlihatkan atau menerangkan sesuatu data yang akan disajikan.

Diagram Batang

Digunakan untuk menyajikan data yang bersifat kategori atau data distribusi.

Cara menggambar diagram batang yaitu diperlukan sumbu tegak (vertikal) dan sumbu mendatar (horizontal) yang berpotongan tegak lurus.

Apabila diagram dibentuk berdiri (tegak lurus), maka sumbu mendatar digunakan untuk menyatakan atribut atau waktu, sedangkan nilai data (kuantum) ditunjukkan dengan sumbu tegak.

Adapun letak batang satu dengan lainnya harus terpisah dan serasi mengikuti tempat diagram yang ada.

Penyajian data berbentuk diagram batang ini banyak variasinya, tergantung pada keahlian pembuat diagram.

Contoh Diagram Batang



Diagram Garis

Digunakan untuk menggambarkan keadaan yang serba terus menerus (berkesinambungan), misalnya pergerakan indeks bursa saham, grafik kurs valuta, dan lain-lain.



Diagram Lambang (Simbol)

Diagram lambang adalah diagram yang menggambarkan simbol-simbol dari data sebagai alat visual untuk orang awam.

Lambang yang digunakan harus sesuai dengan objek yang diteliti, misalnya data angkatan kerja dilambangkan orang, hutan produksi digambarkan pohon, data listrik digambarkan bola lampu, dan lain-lain.

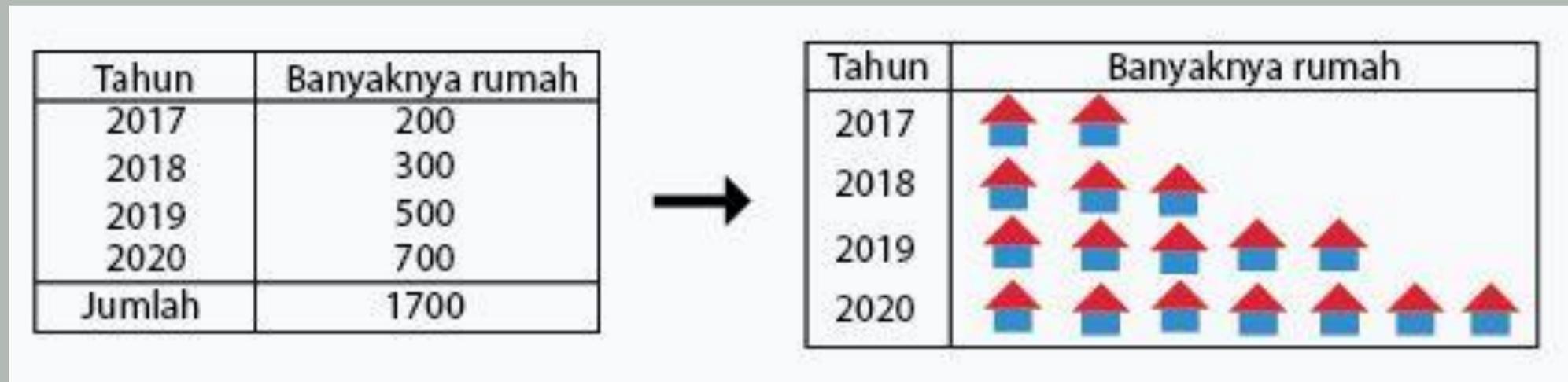


Diagram Lingkaran dan Pastel

Diagram lingkaran adalah diagram yang didasarkan pada sebuah lingkaran yang dibagi-bagi dalam beberapa bagian sesuai dengan macam data dan perbandingan frekuensi masing-masing data yang disajikan.

Langkah-langkah membuatnya :

- a. Ubahlah setiap perubahan nilai data ke dalam derajat
- b. Buatlah lingkaran (360°), kemudian bagilah lingkaran tersebut menjadi beberapa bidang
- c. Setiap bidang menggambarkan kategori data

Diagram pastel adalah perubahan wujud dari model diagram lingkaran versi terpotong yang disajikan dalam bentuk tiga dimensi.

Contoh Diagram Lingkaran dan Pastel

Diagram lingkaran dan diagram pastel :

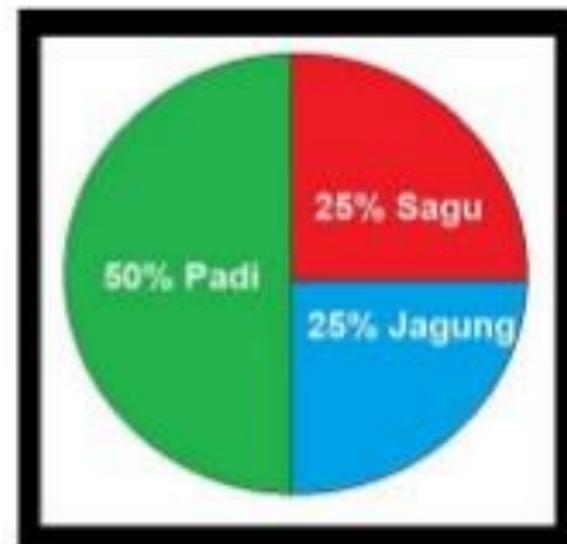


Diagram Lingkaran



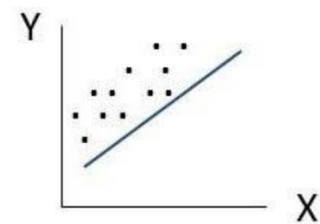
Diagram Pastel



Diagram Pencar

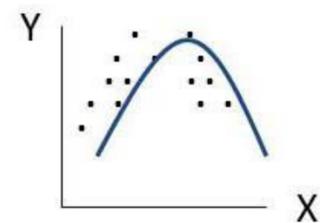
Diagram pencar adalah diagram yang menunjukkan gugusan titik-titik setelah garis koordinat sebagai penghubung yang dihapus.

Biasanya diagram ini digunakan untuk menggambarkan titik data dan korelasi atau regresi yang terdiri dari variabel



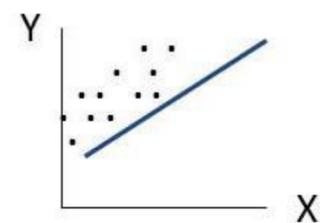
Korelasi Linear:

Jika semua titik (X,Y) pada diagram pencar mendekati bentuk garis lurus.



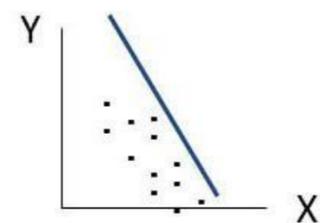
Korelasi Non-linear:

Jika semua titik (X,Y) pada diagram pencar tidak membentuk garis lurus.



Korelasi Positif:

Jika jika arah perubahan kedua variabel sama \Rightarrow Jika X naik, Y juga naik.



Korelasi Negatif:

Jika jika arah perubahan kedua variabel tidak sama \Rightarrow Jika X naik, Y turun.

Diagram Campuran

Diagram campuran adalah diagram yang disajikan dalam bentuk gabungan dari beberapa dimensi dalam satu penyaji data.

Contoh : diagram pastel dengan diagram lambang, diagram pastel dengan diagram batang, diagram lambang dengan tabel, dan sebagainya

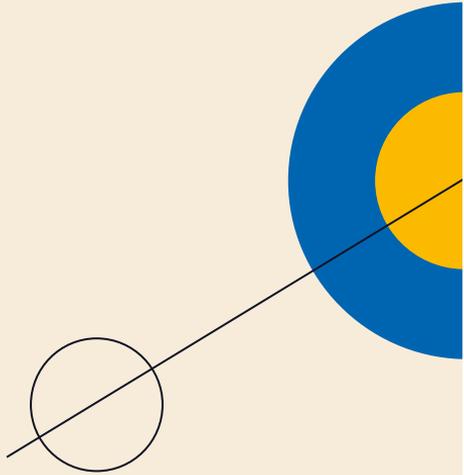


Thank You



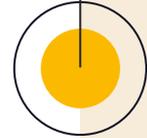
UKURAN

PEMUSATAN DATA



UKURAN PEMUSATAN

Ukuran pemusatan merupakan penyederhanaan data untuk mempermudah peneliti membuat interpretasi dan mengambil suatu keputusan.



Mean

Kuartil

Ukuran
Pemusatan Data

Median

Desil

Modus

Persentil

1. RATA-RATA (MEAN)



Rata-rata (mean) yaitu nilai yang menunjukkan pusat dari nilai data dan merupakan nilai yang dapat mewakili dari sekelompok data yang ada.

Rumus dari Mean:

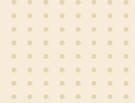
$$\bar{x} = \frac{\sum x_i}{n}$$

Keterangan:

\bar{x} : nilai rata-rata

x_i : nilai data ke-i

n : banyaknya data



CONTOH SOAL

Rata-rata tinggi badan siswa di suatu kelas. Kita bisa mengambil sampel misalnya sebanyak 10 siswa dan kemudian diukur tinggi badannya. Dari hasil pengukuran diperoleh data tinggi badan kesepuluh siswa tersebut dalam ukuran sentimeter (cm) sebagai berikut. 172, 167, 180, 170, 169, 160, 175, 165, 173, 170. Rata-rata tinggi badan siswa adalah...

Jawaban:

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x} = \frac{172+167+180+170+169+160+175+165+173+170}{10}$$

$$\bar{x} = \frac{1701}{10}$$

$$\bar{x} = 170,1$$

$$\bar{X} = 170,1$$

Jadi rata-rata tinggi siswa adalah 170,1 cm.

2. MEDIAN

Median adalah nilai pusat yang terletak di tengah-tengah kumpulan data.

Rumus dari Median Data Tunggal:

$$Me = \frac{1}{2} (n+1)$$

Keterangan :

n = jumlah data

Contoh

1. Data ganjil Diketahui data setelah diurutkan sbb:
4, 5, 7, 8, 10, 10, 12

Dengan rumus :

$$Me = \frac{1}{2} (7+1) = 4.$$

maka posisi Me terletak pada data ke 4, yakni 8.

MEDIAN DATA KELOMPOK

$$me = B + P \left(\frac{1/2n + Jf}{f} \right)$$

Me = Median

b = Batas bawah kelas median, yakni kelas dimana median akan terletak

p = Panjang kelas median

n = ukuran sampel atau banyaknya data

Jf = Jumlah semua frekuensi sebelum kelas median

f = frekuensi kelas median



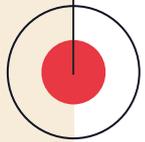
3. MODUS

Modus didefinisikan sebagai bilangan yang paling banyak muncul atau bilangan yang frekuensi kemunculannya paling besar dari suatu data.

Modus Data Tunggal

ada sampel memiliki data nilai sbb:
12, 34, 14, 34, 28, 34, 34, 28, 14

Dari data tersebut bisa diketahui bahwa nilai 34 merupakan modulusnya, karena muncul 4 kali (frekuensinya = 4)



Modus Data Kelompok

Jika data yang kita miliki sudah disusun dalam tabel distribusi frekuensi, maka cara mencari modus adalah dengan rumus:

$$mo = b + p (f1 / f1+f2)$$

Mo = Modus

b = Batas bawah kelas modal, yakni kelas interval dengan frekuensi terbanyak

p = Panjang kelas modal

F1 = frekuensi kelas modal dikurangi kelas intervalterdekat sebelumnya

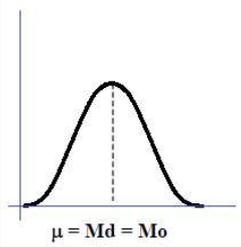
F2 = frekuensi kelas modal dikurangi kelas intervalterdekat sesudahnya.

Data Kelompok

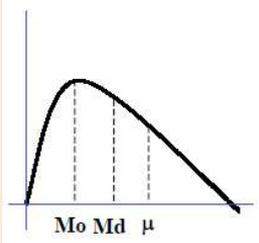
Kelas Interval	Frekuensi
27-38	7
39-50	5
52-62	7
63-74	10
75-86	6
87-98	5
Jumlah	40

$$\begin{aligned}
 Mo &= 62,5 + 12 \left(\frac{3}{3+4} \right) \\
 &= 62,5 + 12 (3/7) \\
 &= 62,5 + 12 (0,43) \\
 &= 62,5 + 5,14 \\
 &= 67,64
 \end{aligned}$$

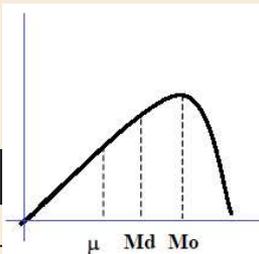
HUBUNGAN MEAN, MEDIAN DAN MODUS



1. Jika rata-rata, median dan modus memiliki nilai yang sama, maka nilai rata rata, median dan modus akan terletak pada satu titik dalam kurva distribusi frekuensi. Kurva distribusi frekuensi tersebut akan terbentuk simetris.



2. Jika rata-rata lebih besar dari median, dan median lebih besar dari modus, maka pada kurva distribusi frekuensi, nilai rata-rata akan terletak di sebelah kanan, sedangkan median terletak di tengahnya dan modus di sebelah kiri. Kurva distribusi frekuensi yang terbentuk adalah menceng kanan atau kemencengan positif.



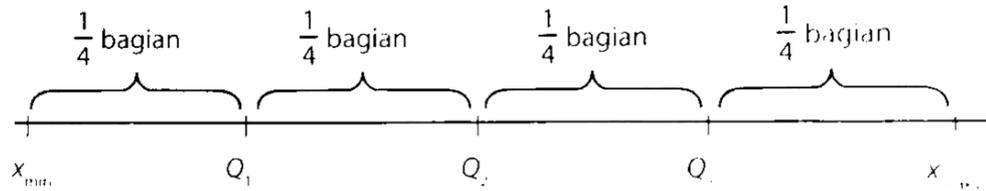
3. Jika rata-rata lebih kecil dari median, dan median lebih kecil dari modus, maka pada kurva distribusi frekuensi, nilai rata-rata akan terletak di sebelah kiri, sedangkan median terletak di tengahnya dan modus di sebelah kanan. Kurva distribusi frekuensi yang terbentuk adalah menceng kiri atau kemencengan negatif.

4. KUARTIL

Kuartil adalah ukuran letak yang membagi data yang telah diurutkan menjadi empat bagian yang sama besar.

Untuk empat bagian tersebut dibutuhkan 3 titik kuartil yaitu:

Dimana masing-masing kuartil diberi nama kuartil satu (Q_1), kuartil dua (Q_2), kuartil tiga (Q_3)



■ **Gambar 4.2** Bagian Kuartil

KUARTIL UNGROUPED DATA

Langkah awal untuk mencari kuartil adalah dengan mencari di mana letak kuartilnya setelah itu baru mencari nilai kuartil tersebut.

Rumus mencari **Letak Kuartil:**

Ungrouped Data

$$Q = \frac{q(n + 1)}{4}$$

Untuk data yang belum dikelompokkan langkah awal adalah mengurutkan data tersebut dari yang terendah sampai yang tertinggi selanjutnya mencari letak kuartil berdasarkan pada rumus tersebut.

Contoh Soal:

Berikut adalah data nilai matematika dari 11 orang siswa, yaitu:

70, 62, 46, 52, 48, 61, 53, 44, 50, 54, 57

Dari data tersebut carilah Q_1 , Q_2 , Q_3 ?

Jawab:

a. Mencari nilai Q_1

1. Mengurutkan data dari yang terkecil sampai yang terbesar.
2. Mencari letak kuartil dengan rumus letak kuartil.

$$Q = \frac{q(n + 1)}{4} = \frac{1(11 + 1)}{4} = 3$$

3. Letak kuartil untuk Q_1 adalah di data ke-3 seperti tampilan tersebut.

44, 46, 48, 50, 52, 53, 54, 57, 61, 62, 70



Q_1

4. Langkah ke-3, mencari nilai Q_1 yang terletak di data ke-3 yaitu 48

b. Mencari nilai Q_2

1. Mengurutkan data dari yang terkecil sampai yang terbesar.
2. Mencari letak kuartil dengan rumus letak kuartil.

$$Q = \frac{q(n+1)}{4} = \frac{2(11+1)}{4} = 6$$

3. Letak kuartil untuk Q_1 adalah di data ke-3 seperti tampilan tersebut.

44, 46, 48, 50, 52, 53, 54, 57, 61, 62, 70

↑ ↑
 Q_1 Q_2

4. Langkah ke-3, mencari nilai Q_1 yang terletak di data ke-6 yaitu 53

c. Mencari nilai Q_3

1. Mengurutkan data dari yang terkecil sampai yang terbesar.
2. Mencari letak kuartil dengan rumus letak kuartil.

$$Q = \frac{q(n+1)}{4} = \frac{3(11+1)}{4} = 9$$

3. Letak kuartil untuk Q_1 adalah di data ke-3 seperti tampilan tersebut.

44, 46, 48, 50, 52, 53, 54, 57, 61, 62, 70

↑ ↑ ↑
 Q_1 Q_2 Q_3

4. Langkah ke-3, mencari nilai Q_1 yang terletak di data ke-9 yaitu 61

KUARTIL GOUPED DATA

Untuk mencari data yang dikelompokkan maka langkah awal adalah mencari letak kuartil dengan rumus berikut:

$$Q = \frac{q \cdot n}{4}$$

Setelah letak kuartil diketahui, langkah selanjutnya adalah mencari nilai kuartil dengan rumus seperti berikut. Rumus mencari Nilai Kuartil grouped data :

$$Q = Tb + S \times \frac{fq}{f}$$

Keterangan: q = kuartil fq = Frekuensi kuartil n = banyaknya data

i = interval Tb = Tepi batas bawah kelas s = selisih anatar letak kuartil dengan frekuensi kumulatif sebelum kelas .

Contoh Soal

Berikut adalah tabel distribusi frekuensi dan nilai statistika pendidikan dari 60 mahasiswa di UHAMKA jurusan PGSD. Dari data berikut carilah Q1, Q2 dan Q3!

Interval	Frekuensi	Frekuensi Kumulatif
20-29	3	3
30-39	6	9
40-49	12	21
50-59	15	36
60-69	12	48
70-79	9	57
80-89	3	60

Penyelesaian:

a. Langkah awal, yaitu mencari di mana letak kuartil dengan rumus

$$Q = \frac{q \cdot n}{4}$$

4

$$Q1 = \frac{1 \cdot 60}{4} = 15$$

4

$$Q2 = \frac{2 \cdot 60}{4} = 30$$

4

$$Q3 = \frac{3 \cdot 60}{4} = 45$$

4

b. Q1 terletak pada data ke-15 yang berada pada kelas ke-3,

Q2 pada data ke-30 yang terletak di kelas ke-4 dan

Q3 pada data ke-45 yang terletak di kelas-5

c. Masukkan ke dalam rumus kuartil untuk data yang dikelompokkan.

Seperti berikut :

$$Q1 = 39,5 + \frac{15 - 9}{12} \times 9 = 44$$

12

$$Q2 = 55,5 + \frac{25 - 21}{15} \times 9 = 57,9$$

15

$$Q3 = 59,5 + \frac{45 - 36}{12} \times 9 = 66,25$$

12

5. DESIL

Desil merupakan titik atau nilai yang membagi seluruh distribusi frekuensi dari data yang kita selidiki ke dalam 10 bagian yang sam besar.

Lambang dari desil adalah D. Jadi, 9 buah titik desil yang dimaksud tersebut adalah titik-titik: D1, D2, D3, D4, D5, D6, D7, D8 dan D9.

Rumus mencari Letak Desil:

$$D = \frac{d(n+1)}{10}$$

Ukuran letak

Berikut adalah data nilai ujian statistika pendidikan 13 orang mahasiswa UHAMKA

No	Nama	Nilai	No	Nama	Nilai
1	Aji	55	8	Iqbal	65
2	Asep	35	9	Neneng	70
3	Ayu	40	10	Ratih	80
4	Budi	45	11	Ratna	85
5	Dwi	50	12	Siti	95
6	Imas	30	13	Ujang	100
7	Ismail	60			

Jawaban :

$$D1 = \frac{d(n+1)}{10} = \frac{1(13+1)}{10} = \frac{14}{10}$$

$$D2 = \frac{d(n+1)}{10} = \frac{2(13+1)}{10} = \frac{28}{10}$$

$$D3 = \frac{d(n+1)}{10} = \frac{9(13+1)}{10} = \frac{126}{10}$$

Ukuran letak

Data tersebut adalah data ganjil yang belum dikelompokkan. Langkah awal, yaitu mengurutkan data terendah sampai tertinggi.

Letak D1 berada pada data ke- $\frac{14}{10}$ atau 1,4 (data ke-1 + 0,4)

Nilai D1 adalah data ke-1 yang ditambah $\frac{4}{10}$ lalu dikali (data ke-2 – data ke-1)

$$\text{Nilai D1} = 30 + \frac{4}{10} (35 - 30) = 32$$



6. PERSENTIL

Pengertian-pengertian pada median, kuartil dan desil dapat digunakan untuk memahami pengertian yang terdapat pada persentil.

Bedanya, jika median distribusinya dibagi menjadi 2 kategori, kuartil dibagi menjadi 4 kategori dan desil dibagi menjadi 10 kategori. Maka, persentil distribusinya dibagi menjadi 100 kategori.

Dengan demikian , dalam perhitungannya nanti akan dijumpai sebanyak 99 titik persentil.



Ukuran letak

Rumus mencari Letak Persentil adalah sebagai berikut:

$$P = \frac{p(n+1)}{100}$$

Berikut Daftar siswa dan nilainya

No	Nama	Nilai
1	Aji	55
2	Asep	35
3	Ayu	40
4	Budi	45
5	Dwi	50
6	Imas	30
7	Ismail	60

No	Nama	Nilai
8	Iqbal	65
9	Neneng	70
10	Ratih	80
11	Ratna	85
12	Siti	95
13	Ujang	100

Ukuran letak

Carilah P26 dan P75?

Langkah 1: Mencari letak persentil dengan rumus

$$P_{26} = \frac{p(n+1)}{100} = \frac{26(13+1)}{100} = 3,64$$

$$P_{75} = \frac{p(n+1)}{100} = \frac{75(13+1)}{100} = 10,50$$

Langkah 2: Mencari nilai data dari persentil tersebut.

Letak P26 berada pada data ke -3

Nilai P26 adalah data ke-3

$$P_{26} = 40 = 64 - 100(45 - 40) = 43,2$$

Letak P75 berada pada data ke -10

$$\text{Nilai } P_{75} = 80 = 50$$

Ukuran letak

Ukuran Penyebaran Data





Ukuran penyebaran data

Ukuran penyebaran data adalah ukuran yang menunjukkan seberapa jauh data suatu menyebarkan dari rata-ratanya. Pada ukuran penyebaran data, kita akan mempelajari: jangkauan (range), simpangan, ragam (variasi), ukuran penyebaran pada nilai kuartil dan pencilan (outlier).





01

Jangkauan (*Range*)

02

Simpangan (*Deviasi*)

03

Ragam (*Variasi*)

04

Quartil

05

Pencilan (*Outlier*)



01

Jangkauan
(Range)





Range yang biasa diberi lambang "R" adalah salah satu ukuran statistik yang menunjukkan jarak penyebaran antara skor (nilai) yang terendah (lowest score) samapi skor (nilai) yang tertinggi (highest score). Dengan singkat dapat dirumuskan:

$$R = x_{maks} - x_{min}$$

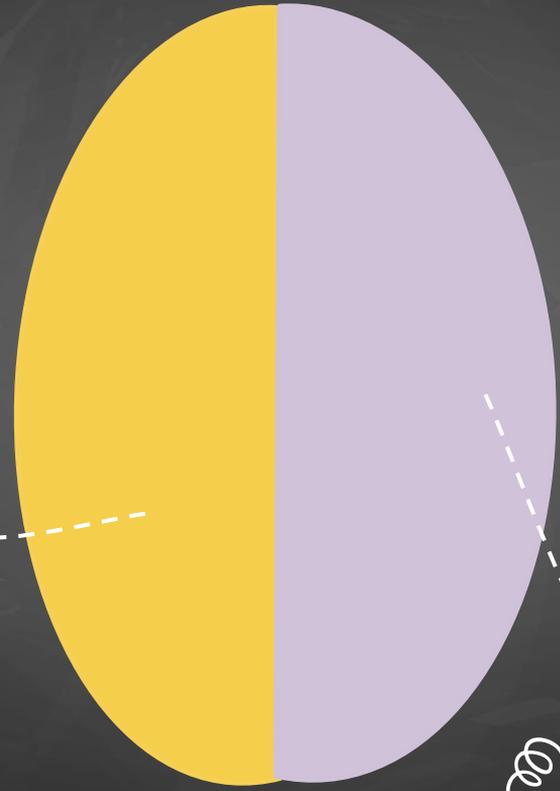
Keterangan :

R = jangkauan atau range

x_{min} = nilai atau data terkecil

x_{maks} = nilai atau data terbesar

Jangkauan
Data Tunggal



Jangkauan
Data Kelompok





a. Jangkauan Data Tunggal

Jangkauan data tunggal langsung tentukan nilai terbesar dan terkecil nya, lalu dikurang;

Contoh :

- Tentukanlah jangkauan (range) dari data-data dibawah ini :
6,7,3,4,8,3,7,6,10,15,20

Penyelesaian :

Dik : $X_{maks} = 20$, dan $X_{min} = 3$
menentukan jangkauan nya :

$$R = X_{maks} - X_{min} = 20 - 3 = 17$$

Jadi jangkauan data tersebut adalah 17



b. Jangkauan Data Berkelompok

Untuk data bergolong, nilai tertinggi diambil dari nilai tengah kelas tertinggi dan nilai terendah diambil dari nilai kelas yang terendah

Contoh soal:

Tentukan range dari table berikut ini:

Nilai	Frekuensi
3 – 5	3
6 – 8	6
9 – 11	16
12 – 14	8
15 – 17	7
18 – 20	10

*) nilai tengah kelas terendah:

$$X_{\min} = \frac{3 + 5}{2} = 4$$

*) nilai tengah kelas tertinggi:

$$X_{\max} = \frac{18 + 20}{2} = 19$$

*) menentukan jangkauannya:

$$R = X_{\max} - X_{\min} = 19 - 4 = 15$$





02

Simpangan (Deviasi)





Simpangan atau deviasi ialah selisih satu simpangan dari masing-masing skor atau interval, dari nilai rata-rata hitungannya.

Deviasi merupakan salah satu ukuran variabilitas data yang biasa dilambangkan dengan huruf kecil dari huruf yang digunakan bagi lambing skornya.

Jika apabilaskor nya dibagi lambing X maka deviasinya berlambang x;

jika skornya Y makadeviasinya y;

jika skornya Z maka lambing deviasinya z.

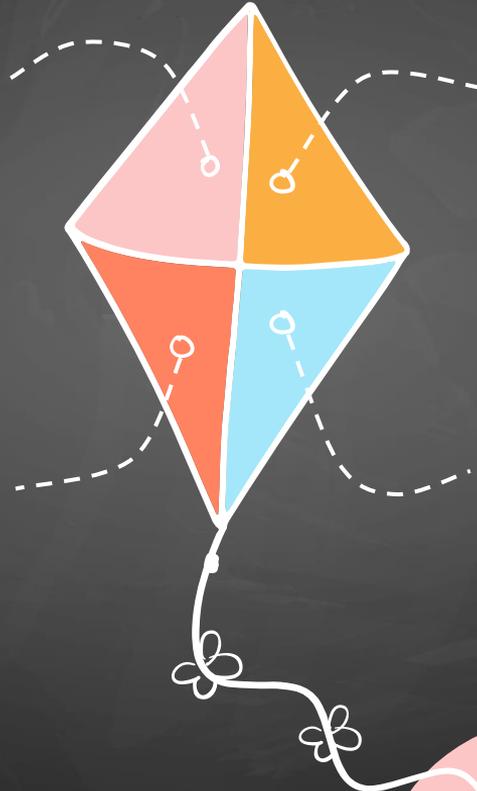
Karena deviasi merupakan simpangan atau selisih dari masing-masing skor terhadap mean groupnya.



Skor (X)	Banyaknya (f)	Deviasi ($x=X-M_x$)
8	1	$8-6 = +2$
7	1	$7-6 = +1$
6	1	$6-6 = 0$
5	1	$5-6 = -1$
4	1	$4-6 = -2$
$30 = \sum X$	$5 = N$	$0 = \sum x$

Deviasi
Rata-rata
"TUNGGAL"

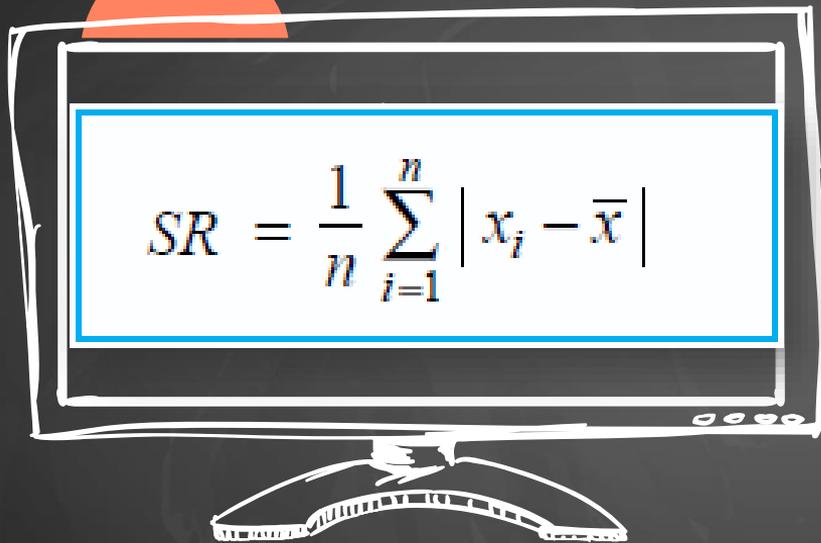
Deviasi
Rata-rata
"KELOMPOK"



Deviasi
Standar
"TUNGGAL"

Deviasi
Standar
"KELOMPOK"

a. Simpangan Rata-rata Data Tunggal



$$SR = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$



Keterangan :

SR = simpangan rata-rata

N = ukuran data (total frekuensi)

X_i = data ke i dari data $x_1 x_2 x_3 \dots x_n$

\bar{x} = rata-rata hitung

Σ = notasi sigma yang artinya jumlah

$|x_i - \bar{x}|$ = harga mutlak dari $x_i - \bar{x}$ yang

hasilnya selalu positif

Contoh : $|3| = 3$ dan $3| = 3$

Contoh soal:

Diketahui data 7,6,8,7,6,10,5. tentukan simpangan rata-ratanya!

Penyelesaian:

*) Menentukan rata-ratanya,

$$\bar{x} = \frac{7 + 6 + 8 + 7 + 6 + 10 + 5}{7} = \frac{49}{7} = 7$$

*) Menentukan simpangan rata-ratanya :

$$SR = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

$$= \frac{1}{7} \sum_{i=1}^7 |x_i - 7|$$

$$= \frac{1}{7} (|7 - 7| + |6 - 7| + |8 - 7| + |7 - 7| + |6 - 7| + |10 - 7| + |5 - 7|)$$

$$= \frac{1}{7} (|0| + |-1| + |1| + |0| + |-1| + |3| + |-2|)$$

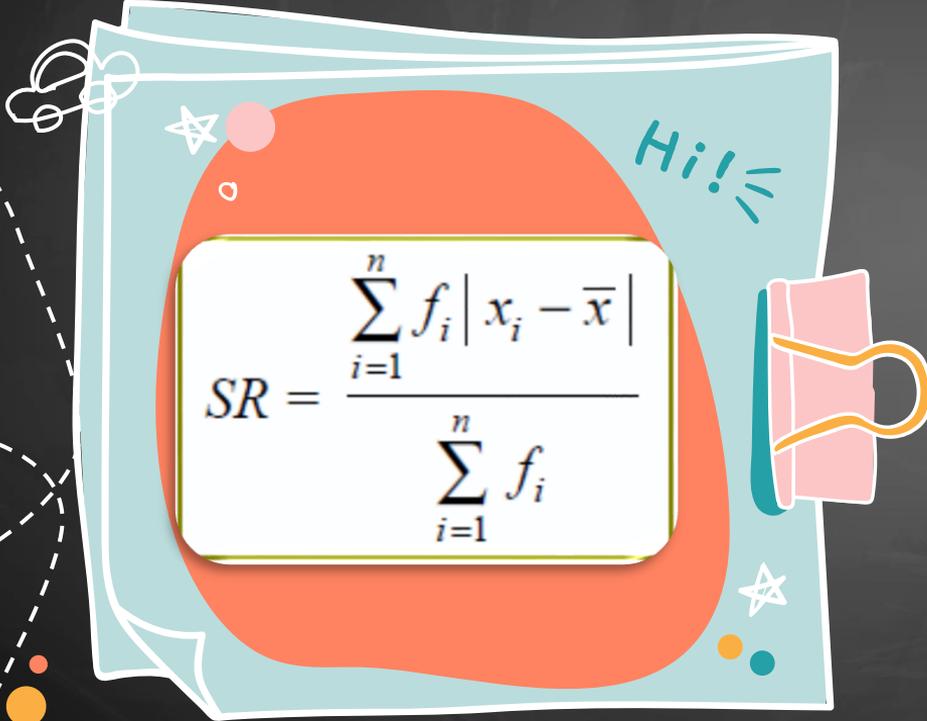
$$= \frac{1}{7} (0 + 1 + 1 + 0 + 1 + 3 + 2)$$

$$= \frac{1}{7} (8) = \frac{8}{7}$$

Jadi, simpangan rata-ratanya adalah $\frac{8}{7}$



b. Simpangan Rata-rata Data Kelompok


$$SR = \frac{\sum_{i=1}^n f_i |x_i - \bar{x}|}{\sum_{i=1}^n f_i}$$

Keterangan :

SR = simpangan rata-rata

N = banyak kelas

x_i = nilai tengah kelas I

\bar{x} = rata-rata hitung

$\sum_{i=1}^n F_i$ = total frekuensi



Contoh soal:

Tentukan simpangan rata-rata pada tabel tersebut!

Nilai	Frekuensi
141 – 145	2
146 – 150	4
151 – 155	8
156 – 160	12
161 – 165	10
166 – 170	4

*) Melengkapi isi table

Nilai	f_i	x_i	$f_i \cdot x_i$	$ x_i - \bar{x} $	$f_i x_i - \bar{x} $
141 – 145	2	143	286	14,5	29
146 – 150	4	148	592	9,5	38
151 – 155	8	153	1.224	4,5	36
156 – 160	12	158	1.896	0,5	6
161 – 165	10	163	1.630	5,5	55
166 – 170	4	168	672	10,5	42
Jumlah	40		6.300		260

*) Menentukan rata-rata:

$$\bar{x} = \frac{\sum_{i=1}^n f_i \cdot x_i}{\sum_{i=1}^n f_i} = \frac{6300}{40} = 157,5$$

*) Menentukan simpangan rata-ratanya :

$$SR = \frac{\sum_{i=1}^n f_i \cdot |x_i - \bar{x}|}{\sum_{i=1}^n f_i} = \frac{260}{40} = 5,15.$$

Jadi, simpangan rata-ratanya adalah 5,15.

c. Simpangan Baku (Standar Deviasi) Data Tunggal

Diketahui sekumpulan data kuantitatif yang tidak dikelompokkan dan dinyatakan oleh $X_1, X_2 \dots x_n$ dari data tersebut, dapat diperoleh nilai simpang baku (s) yang ditentukan oleh rumus berikut :



$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

untuk sampel

dan

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}}$$

untuk populasi

Data sampel berlaku untuk $n < 30$ dan data populasi untuk $n > 30$

Contoh soal :

Dari 40 siswa kelas XI IPA diperoleh nilai yang mewakili adalah 7,9,6,3 dan 5. tentukan simpangan baku dari data tersebut!

Penyelesaian :

$$\bar{X} = \frac{7 + 9 + 6 + 3 + 5}{5} = \frac{30}{5} = 6$$

*) melengkapi table

Nilai (x)	$x_i - \bar{x}$	$(x - \bar{x})^2$
3	-3	9
5	-1	1
6	0	0
7	1	1
9	3	9
30		20

*) menentukan simpangan baku

$$s = \sqrt{\frac{\sum_{i=1}^5 (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{20}{5 - 1}} = \sqrt{5} = 2,24$$

Jadi, simpangan bakunya adalah 2,24.



d. Simpangan Baku (Standart Deviasi) Data Berkelompok

Sekumpulan data kuantitatif yang dikelompokkan, dapat dinyatakan oleh $X_1, X_2, x_3 \dots x_n$ dan masing-masing data mempunyai frekuensi $F_1, F_2, F_3 \dots F_n$ simpangan baku (S) dari data tersebut diperoleh dengan rumus :

$$S = \sqrt{\frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{n-1}}$$

untuk sampel

dan

$$\sigma = \sqrt{\frac{\sum_{i=1}^n f_i (x_i - \mu)^2}{n}}$$

untuk populasi

Data sampel berlaku untuk $n < 30$ dan data populasi untuk $n > 30$

Contoh:

hasil matematika 30 siswa kelas XI IPA pada table:

nilai	Frekuensi
5-9	3
10-14	8
15-19	11
20-24	6
25-29	2

*) melengkapi table

Nilai	f_i	Titik Tengah (x_i)	$f_i \cdot x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i (x_i - \bar{x})^2$
5 - 9	3	7	21	-9,33	87,05	261,15
10 - 14	8	12	96	-4,33	18,75	150
15 - 19	11	17	187	0,67	0,45	4,95
20 - 24	6	22	132	5,67	32,15	192,9
25 - 29	2	27	54	10,67	113,85	227,7
Jumlah	30		490			836,7

*) menentukan rata-rata :

$$\bar{x} = \frac{\sum_{i=1}^n f_i \cdot x_i}{\sum_{i=1}^n f_i} = \frac{490}{30} = 16,33$$

*) Menentukan simpangan bakunya

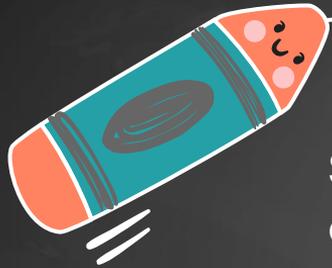
$$s = \sqrt{\frac{\sum_{i=1}^n f_i \cdot (x_i - \bar{x})^2}{n}} = \sqrt{\frac{836,7}{30}} = \sqrt{27,89} = 5,28$$

Jadi, simpangan bakunya adalah 5,28.



03
Ragam
(Variasi)





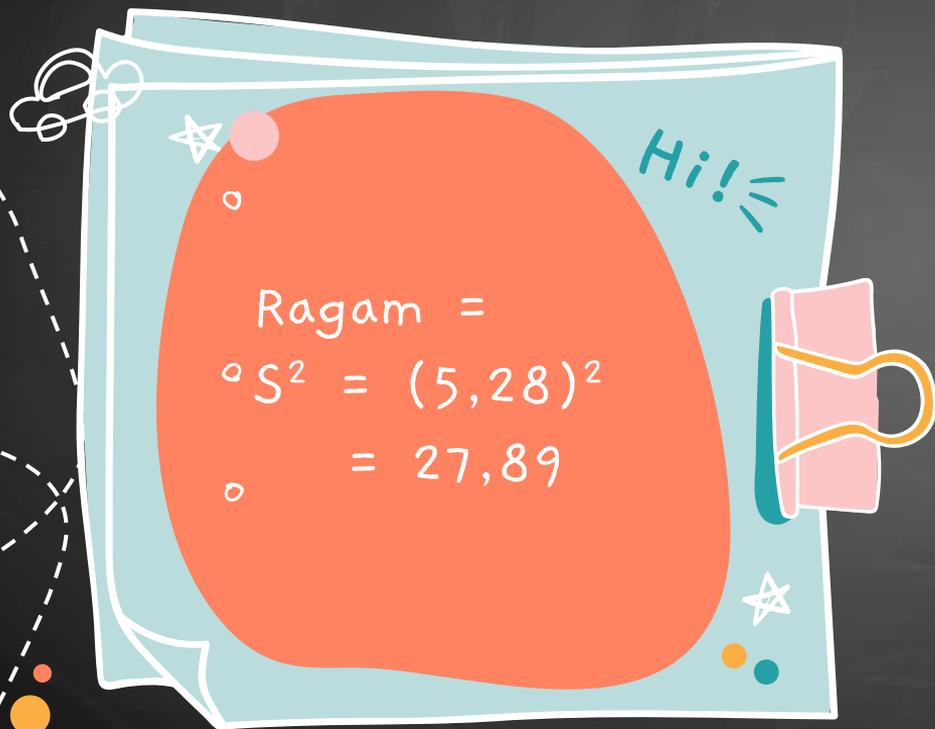
Simpangan Ragam atau Variasi adalah rata-rata dari jumlah kuadrat simpangan tiap data.

Ragam dirumuskan sebagai :

$$\text{Ragam} = S^2$$

Artinya, ragam diperoleh dari nilai simpangan baku dikuadratkan.

Contoh :



Diperoleh simpangan bakunya adalah 5,28. Sehingga nilai ragamnya (variasi)



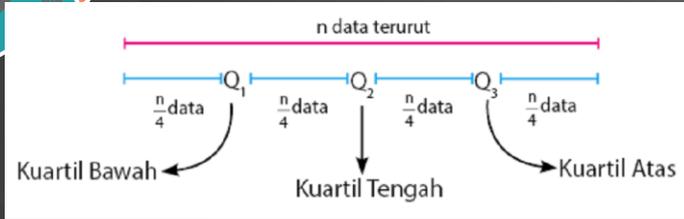




Dari data kita bisa menentukan Jenis kuartil yaitu Kuartil bawah atau Q_1 , kuartil tengah atau Q_2 , dan kuartil atas atau Q_3
Kelompok kuartil berdasarkan jenis data yaitu kuartil data tunggal dan kuartil data kelompok

Kuartil Data Tunggal

Untuk Data Ganjil



Untuk Data Genap

- Carilah nilai yang menjadi nilai tengahnya (median atau Q_2).
- Dengan Membagi data di sebelah kiri median menjadi dua bagian yang sama dan juga menghasilkan kuartil bawah atau Q_1 .
- Dengan membagi data di sebelah kanan median menjadi dua bagian yang sama dan juga menghasilkan kuartil atas atau Q_3 .

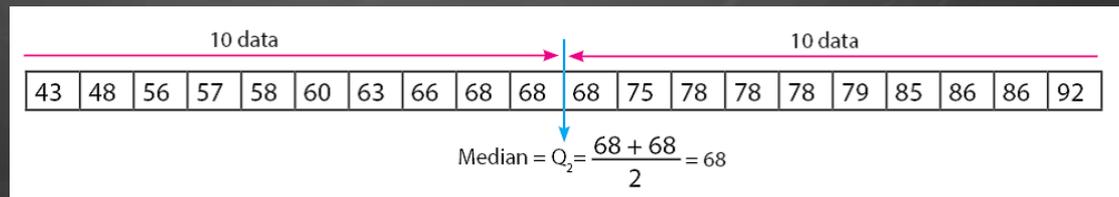
Jenis Kuartil	Rumus Kuartil Data Tunggal
Kuartil Bawah	$Q_1 = x_{\frac{1}{4}(n+1)}$
Kuartil Tengah	$Q_2 = x_{\frac{1}{2}(n+1)}$
Kuartil Atas	$Q_3 = x_{\frac{3}{4}(n+1)}$

Contoh soal :

Berikut tabel data nilai matematika yang diperoleh sekelompok siswa antara lain.

78	86	57	68	56	86	78	92	68	75
63	58	66	78	43	48	60	68	79	85

Mengurutkan data dan mencari kuartil bawah (Q_1), median (Q_2), dan Kuartil atas (Q_3)



Kuartil Data Kelompok



$$Q_i = Tb_i + (((i/4)n - F_i)/f_i)c$$

Keterangan :

- Tb_i = Tepi bawah kuartil ke- i
- F_i = Jumlah frekuensi sebelum frekuensi kuartil ke- i
- f_i = Frekuensi kuartil ke- i . $i = 1, 2, 3$
- n = Jumlah seluruh frekuensi
- C = Panjang interval kelas

Contoh soal :

Perhatikan tabel di bawah berikut ini dan tentukan kuartil atas pada tabel tersebut

Berat Badan	Frekuensi
50 - 54	4
55 - 59	6
60 - 64	8
65 - 69	10
70 - 74	8
75 - 79	4

Pembahasan :

Kuartil atas ialah disimbolkan Q_3

Jumlah data yaitu:

$$4 + 6 + 8 + 10 + 8 + 4 = 40$$

Letak kuartil atas berada di $\frac{3}{4}$ bagian data. Maka, letak kuartil atas tersebut berada di data ke-30. Maka caranya yaitu sebagai berikut:

$$\frac{3}{4} \times 40 = 30$$



Kemudian, perhatikanlah tabel yang telah dilengkapi dengan frekuensi kumulatif kurang dari (fkk) dan juga letak kuartil atas, adalah:



Sehingga, nilai kuartis atasnya ialah:

Berat Badan	Frekuensi	fkk
50 - 54	4	4
55 - 59	6	10
60 - 64	8	18
65 - 69	10	28
70 - 74	8	36
75 - 79	4	40

panjang kelas $p = 5$

fkk sebelum kelas Q_3

Letak kuartil atas = Q_3
 $f_i = 36$

$$\begin{aligned}T_b &= 70 - 0,5 \\ &= 69,5\end{aligned}$$

$$Q_3 = 69,5 + \left(\frac{\frac{3}{4} \cdot 40 - 28}{36} \right) \times 5$$

$$Q_3 = 69,5 + \left(\frac{30 - 28}{36} \right) \times 5$$

$$Q_3 = 69,5 + \left(\frac{2}{36} \right) \times 5$$

$$Q_3 = 69,5 + 0,28 = 69,78$$



Ulasan Mengenai Data Kuartil

01

Simpangan Kuartil

nilai yang menunjukkan setengah kali dari hamparan. Didapat dengan cara mengurangkan kuartil bawah dengan kuartil atas kemudian membagi dengan 2 (dua).

03

Rataan Tiga

Kuartil

Rata-rata dari tiga nilai kuartil yang terdiri dari kuartil atas, kuartil tengah, dan kuartil bawah. Cara mendapatkan rata-rata kuartil adalah dengan menjumlahkan ketiga kuartil kemudian membaginya dengan 2 (dua).

02

Rataan Kuartil

Rata-rata dari kuartil atas dan kuartil bawah. Cara mendapatkan rata-rata kuartil adalah dengan menjumlahkan kuartil atas dan kuartil bawah kemudian membaginya dengan 2 (dua).

04

Statistika

Terdiri atas lima nilai, yaitu nilai tertinggi X_{max} , nilai terendah X_{min} , kuartil atas (Q1), kuartil tengah (Q2), dan kuartil bawah (Q3).

Rumus Ukuran Penyebaran Data Kuartil

Jenis Ukuran Penyebaran	Rumus
Simpangan Kuartil	$Q_d = \frac{1}{2}(Q_3 - Q_1)$
Rataan Kuartil	$\frac{1}{2}(Q_1 + Q_3)$
Rataan Tiga Kuartil	$\frac{1}{4}(Q_1 + 2Q_2 + Q_3)$
Statistika Lima Serangkai	$X_{\max}; X_{\min}; Q_1; Q_2; Q_3$



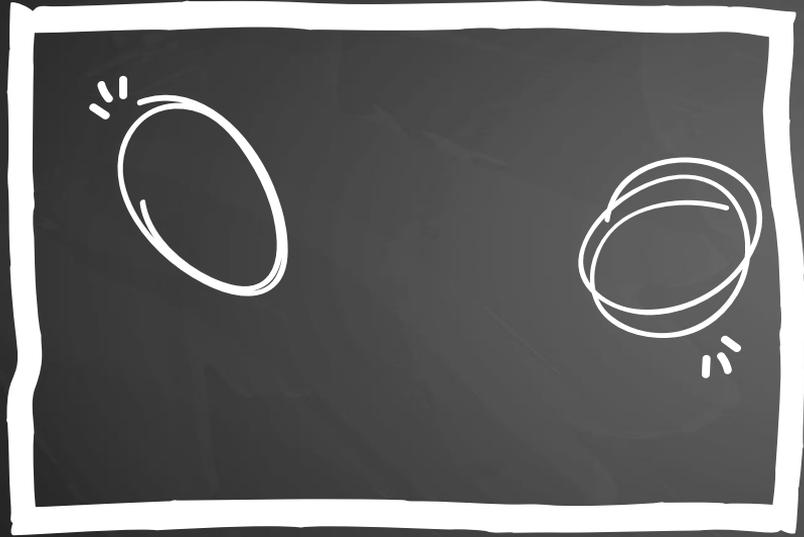
Dari nilai-nilai kuartil tsb juga berlaku ukuran penyebaran yaitu JAngkauan antarkuartil(hamparan) yg kita simbolkan dgn JK dan Jangkauan semi antarkuartil (Simpangan kuartil) dengan SK.

Rumus masing-masing :

$$JK = Q3 - Q1 \text{ dan } SK = \frac{1}{2} (JK) = \frac{1}{2} (Q3 - Q1)$$

Dari nilai penyelesaian pada kuartil dikenal juga istilah Langkah (L), yg dirumuskan:

$$L = \frac{3}{2} (JK) = \frac{3}{2} (Q3 - Q1)$$



Thanks!

CREDITS: This presentation template was created by **Slidesgo**, including icon by **Flaticon**, and infographics & images from **Freepik**



UJI NORMALITAS DATA

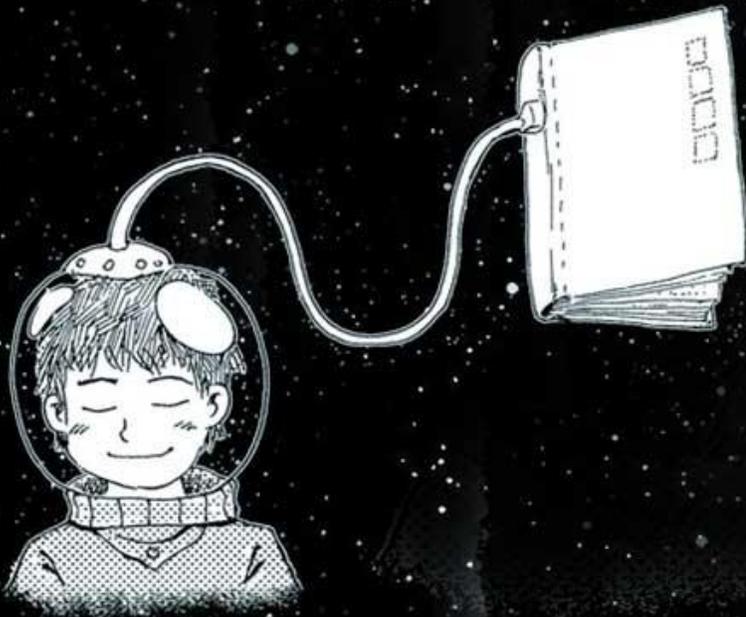


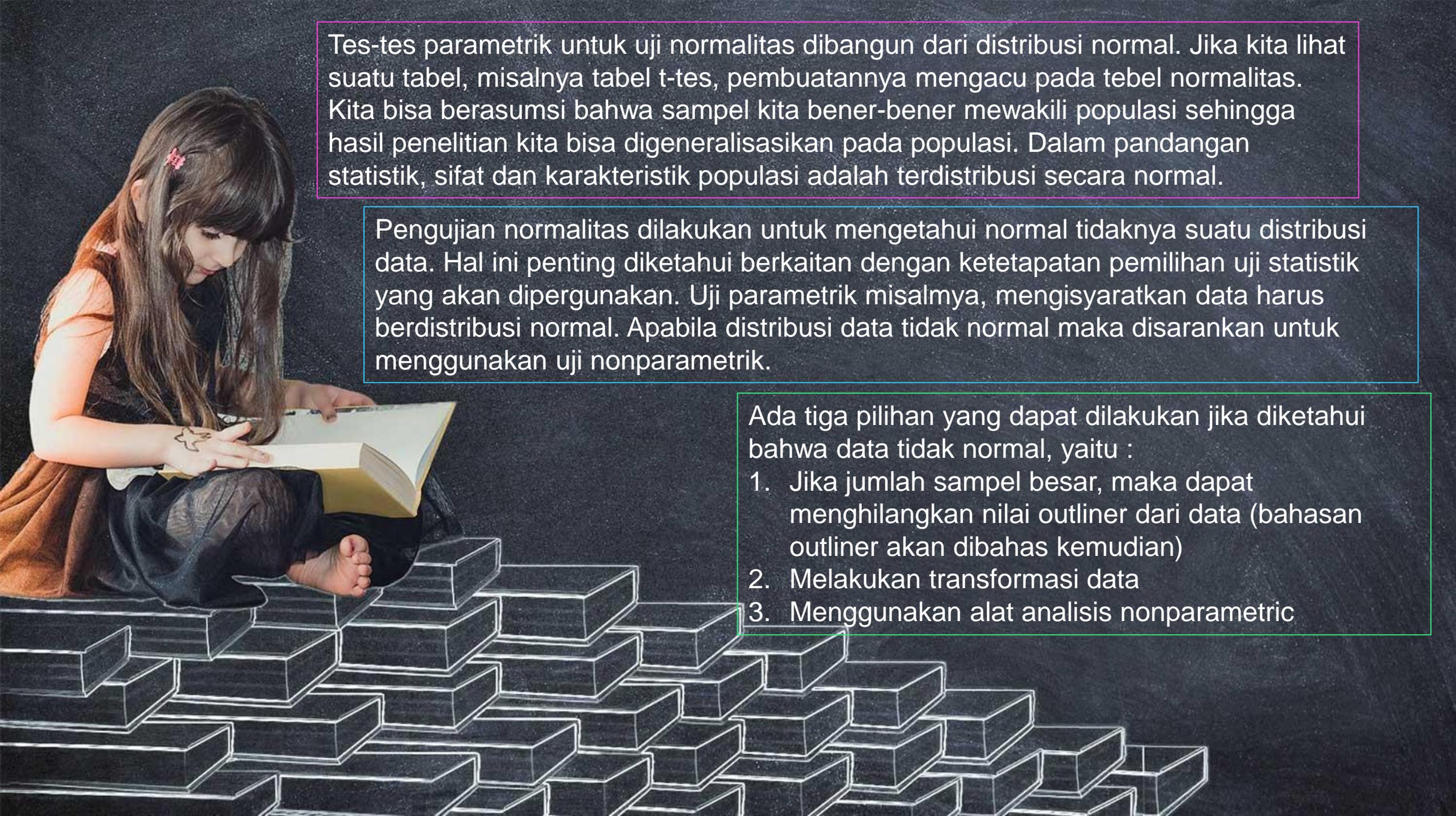
Definisi Uji Normalitas

Uji distribusi normal adalah uji untuk mengukur apakah data yang didapatkan memiliki distribusi normal sehingga dapat dipakai dalam statistik parametrik (statistik inferensial).

Dengan kata lain, uji normalitas adalah uji untuk mengetahui apakah data empirik yang didapatkan dari lapangan itu sesuai dengan distribusi teoritik tertentu.

Dalam kasus ini, distribusi normal. Dengan kata lain, apakah data yang diperoleh berasal dari populasi yang berdistribusi normal.





Tes-tes parametrik untuk uji normalitas dibangun dari distribusi normal. Jika kita lihat suatu tabel, misalnya tabel t-tes, pembuatannya mengacu pada tabel normalitas. Kita bisa berasumsi bahwa sampel kita benar-bener mewakili populasi sehingga hasil penelitian kita bisa digeneralisasikan pada populasi. Dalam pandangan statistik, sifat dan karakteristik populasi adalah terdistribusi secara normal.

Pengujian normalitas dilakukan untuk mengetahui normal tidaknya suatu distribusi data. Hal ini penting diketahui berkaitan dengan ketetapan pemilihan uji statistik yang akan dipergunakan. Uji parametrik misalnya, mengisyaratkan data harus berdistribusi normal. Apabila distribusi data tidak normal maka disarankan untuk menggunakan uji nonparametrik.

Ada tiga pilihan yang dapat dilakukan jika diketahui bahwa data tidak normal, yaitu :

1. Jika jumlah sampel besar, maka dapat menghilangkan nilai outlier dari data (bahasan outlier akan dibahas kemudian)
2. Melakukan transformasi data
3. Menggunakan alat analisis nonparametric

Macam-Macam Uji Normalitas

Uji Normalitas Dengan Liliefors Test

1 Kelebihan Liliefors test adalah penggunaan/perhitungannya yang sederhana, serta cukup kuat (power full) sekalipun dengan ukuran sampel kecil ($n = 4$).

Uji Normalitas Dengan Chi Square

2 Salah satu fungsi dari chi square adalah uji kecocokan. Dalam uji kecocokan akan dibandingkan antara frekuensi hasil observasi dengan frekuensi harapan/teoritis

Pengolahan data dengan Uji normalitas Liliefors

1. Mengurutkan data sampel dari kecil ke besar dan menentukan frekuensi tiap-tiap data

2. Menentukan nilai Z_i dari tiap-tiap data dengan rumus : $Z_i = \frac{X_i - \bar{X}}{S}$

Keterangan :

Z_i : simpangan baku kurva standard
 X_i : data ke i dari suatu kelompok data
 \bar{X} : rata-rata kelompok
 S : simpangan baku.

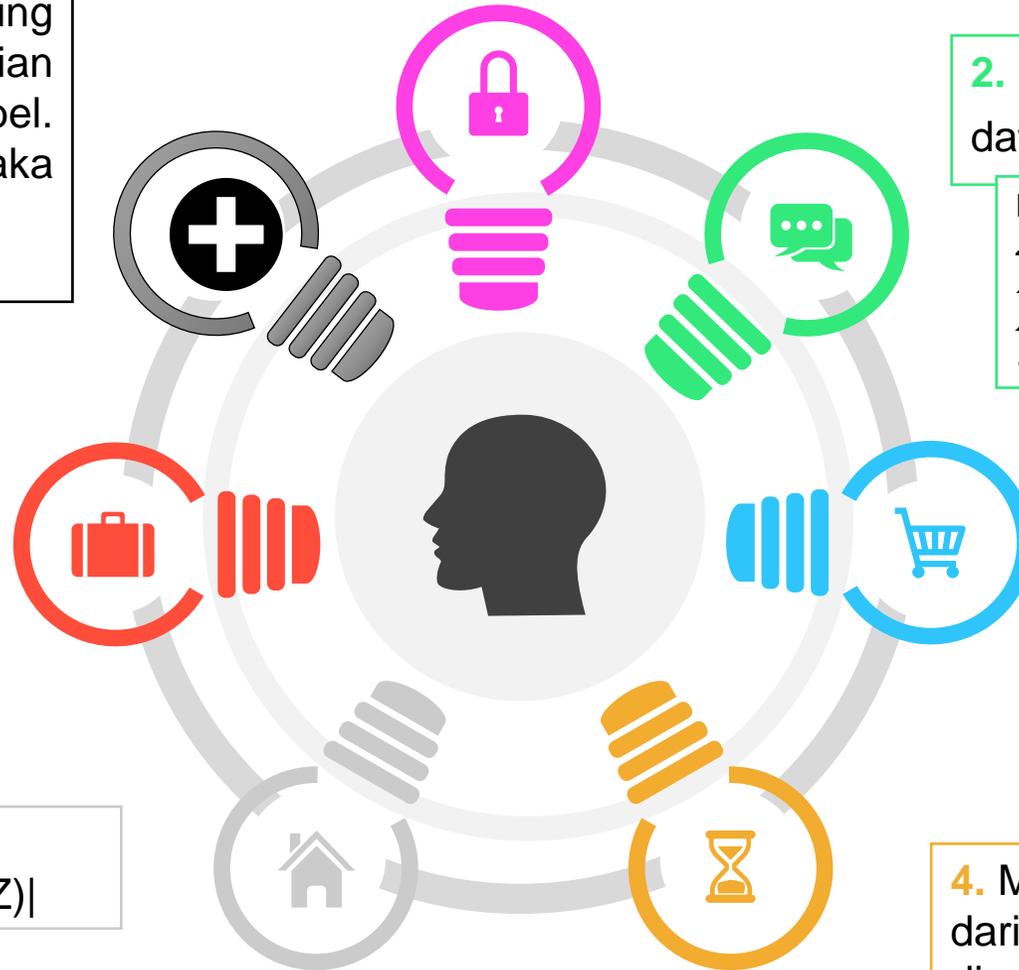
3. Menentukan besar peluang untuk masing-masing nilai Z berdasarkan tabel Z yang disebut $F(Z)$.

4. Menghitung frekuensi kumulatif dari masing-masing nilai Z , dan disebut $S(Z)$.

7. Mengambil harga L_{hitung} yang paling besar kemudian dibandingkan dengan L_{tabel} . Jika $L_{hitung} < L_{tabel}$ maka sampel berdistribusi normal.

6. Menentukan L_{tabel} untuk $n > 30$ dengan taraf signifikansi 5% melalui Tabel Lilliefors. Maka $L_{tabel} = \frac{0,886}{\sqrt{n}}$ dengan n adalah jumlah.

5. Menentukan nilai $L_{hitung} = |F(Z) - S(Z)|$



Seorang mahasiswa program studi Penjas kesrek melakukan penelitian tentang minat siswa terhadap mata pelajaran penjas di sekolah. Skor siswa dalam menyelesaikan angket minat disajikan dalam tabel.

Hipotesis statistik yang digunakan:

H_0 : sampel berdistribusi normal

H_1 : sampel data berdistribusi tidak normal



Tabel 7.8 Skor Nilai Angket Minat

No	Skor	No	Skor	No	Skor	No	Skor
1	2	11	4	21	5	31	7
2	3	12	4	22	5	32	7
3	3	13	4	23	6	33	7
4	3	14	5	24	6	34	7
5	3	15	5	25	6	35	8
6	4	16	5	26	6		
7	4	17	5	27	6		
8	4	18	5	28	6		
9	4	19	5	29	6		
10	4	20	5	30	6		

Langkah - langkah

1. Mengurutkan data sampel dari kecil ke besar dan menentukan frekuensi tiap-tiap data, menentukan nilai Z, menentukan F(Z), menghitung S(Z), dan menentukan nilai Lhitung = $|F(Z) - S(Z)|$

X	f	Z	F(Z)	S(Z)	$ F(Z) - S(Z) $
2	1	-2,15	0,022	0,029	0,01
3	4	-1,44	0,090	0,143	0,05
4	8	-0,72	0,251	0,371	0,12
5	9	0,00	0,500	0,629	0,13
6	8	0,72	0,764	0,857	0,09
7	4	1,44	0,910	0,971	0,06
8	1	2,15	0,978	1,000	0,02
	35				

$$Sd = \sqrt{\frac{\sum (Xi - \bar{X})^2}{(n-1)}} = 1,3932$$

$$\bar{X} = \frac{\sum fX}{\sum f} = 5$$

Hitungan data ke-1:

$$Zi = \frac{2 - \bar{X}}{1,3932} = 2,153$$

F(Z) melihat peluang Z1 pada tabel Z (Z1 = -2,15 maka peluang pada tabel Z adalah 0,022). Apabila Z berharga positif maka F(Z) = 1- harga Z pada tabel. Misalnya untuk Z = 0,72 pada tabel tertera 0,2358 maka F(Z) = 1- 0,2358 = 0,764.

S(Z) menghitung frekuensi kumulatif nilai Z. S(Z1) = 1/35 = 0,029 Menghitung $|F(Z) - S(Z)| = [0,022 - 0,029] = 0,01$

$$L_{tabel} = \frac{0,886}{\sqrt{n}} = \frac{0,886}{\sqrt{35}} = 0,15$$

2. Mengambil harga L hitung yang paling besar kemudian dibandingkan dengan Ltabel.

Lhitung terbesar = 0,13 (Lo)

Diperoleh Lo = 0,13 < Lf = 0,15. Maka diterima Ho yang berarti data terdistribusi **normal**.

Uji Normalitas dengan Chi Square

langkah-langkah pengujian sebagai berikut :

- Membuat tabel distribusi frekuensi yang dibutuhkan
- Menentukan rata-rata dan standar deviasi
- Menentukan batas kelas, yaitu angka skor kiri kelas interval pertama dikurangi 0.5 dan angka skor kanan kelas interval ditambah 0.5
- Mencari nilai z skor untuk batas kelas interval dengan rumus: $z =$
- Mencari luas 0 - Z dari tabel kurva normal dari 0 - Z dengan menggunakan angka-angka untuk batas kelas

- Mencari luas tiap kelas interval dengan jalan mengurangkan angka-angka 0 - Z, yaitu angka baris pertama dikurangi baris kedua, angka baris kedua dikurangi baris ketiga, dan seterusnya. Kecuali untuk angka yang berbeda arah (tanda "min" dan "plus", bukan tanda aljabar atau hanya merupakan arah) angka-angka 0 - Z dijumlahkan.
- Mencari frekuensi harapan (E) dengan cara mengalikan luas tiap interval dengan jumlah responden.
- Menentukan nilai Chi-Kuadrat (χ^2)
- Membandingkan nilai uji χ^2 dengan nilai χ^2 tabel, dengan kriteria perhitungan: Jika nilai uji $\chi^2 \leq$ nilai χ^2_{tabel} , maka data tersebut berdistribusi normal. Dengan $dk = (1 - a)$ ($dk = k - 1$), Dimana $dk =$ derajat kebebasan (degree of freedom), dan $k =$ banyak kelas pada distribusi frekuensi.

• Contoh uji normalitas Chi Square

Tabel 10.1

Nilai hasil ujian Pengantar Statistika 80 mahasiswa S1 Jurusan Sistem Komputer STMIK Raharja



Interval Kelas	Frekuensi (f)	Titik Tengah (X)	Fx
45-49	6	47	282
40-44	10	42	420
35-39	12	37	444
30-34	17	32	544
25-29	14	27	378
20-24	9	22	198
15-19	6	17	102
10-14	6	12	72
-	80	-	2.440

Guna memperoleh luas daerah dibawah kurva normal, kita memerlukan satu tabel lagi, yaitu tabel frekuensi harapan.

Berdasarkan tabel 11.1 , setelah melalui perhitungan yang panjang diperoleh χ^2 sebesar 6,06954.

Langkah selanjutnya adalah menguji χ^2 yang dibandingkan dengan χ^2 tabel. Terdapat perbedaan dalam menentukan derajat kebebasan untuk uji normalitas. Menurut Soedjano (2008), banyaknya kelas dikurangi tiga, sedangkan Hadi (2007), sel fh dikurangi satu.

Menurut Soedjana, derajat kebebasannya adalah $8-3 = 5$, sedangkan menurut Hadi, $8-1 = 7$. Marilah kita uji dengan pendapat ahli tersebut.

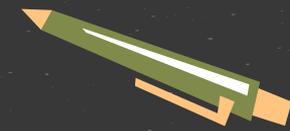
Pertama dengan $dk = 5$ pada taraf signifikansi 1% chi-kuadrat tabel sebesar 15,086.

Kedua dengan $dk = 7$ chi-kuadrat sebesar 18,475. Ternyata chi-kuadrat tabel hitung , yakni 6,06954 berada jauh dibawah chi-kuadrat tabel baik menurut pendekatan Soedjana maupun Hadi.

UJI HOMOGENITAS
DATA

2

3



1

01

2

Pengertian Uji
Homogenitas

3

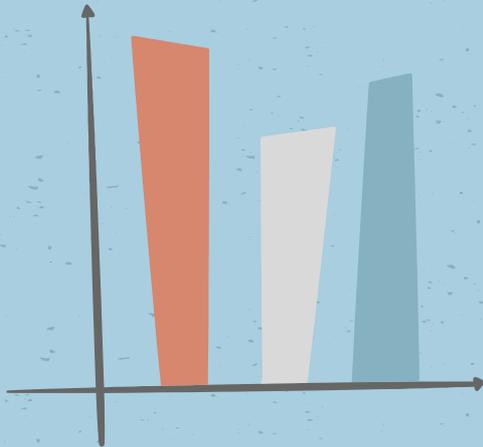
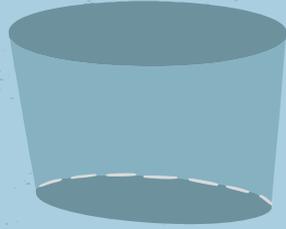


Uji Homogenitas

Uji homogenitas adalah suatu prosedur uji statistik yang dimaksudkan untuk memperlihatkan bahwa dua atau lebih kelompok data sampel berasal dari populasi yang memiliki variansi yang sama.

02

**Metode Analisis
Homogenitas**



METODE ANALISIS HOMOGENITAS

```
graph TD; A[METODE ANALISIS HOMOGENITAS] --- B[UJI FISHER]; A --- C[UJI BARTLETT];
```

UJI FISHER

(Digunakan hanya pada 2 kelompok data)

UJI BARTLETT

(Digunakan pada data > 2 kelompok data)

1

A

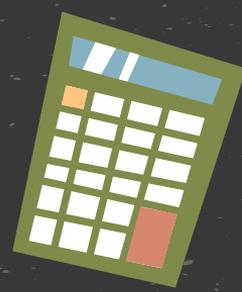
2

Uji Fisher

3



Langkah-langkah pada Uji Fisher menurut Supardi (2013:142)



01

Tentukan taraf signifikansi (α) untuk menguji hipotesis:

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

02

Menghitung varians tiap kelompok data

$$S_i^2 = \frac{n \cdot \sum X_i^2 - [\sum Xi]^2}{n(n-1)}$$

03

Tentukan varian yaitu

$$F_{hitung} = \frac{\text{varians terbesar}}{\text{varians terkecil}}$$

04

Tentukan F_{tabel} untuk taraf signifikansi α

$$dk_1 = dk_{pembilang} = n_a - 1$$

$$dk_2 = dk_{penyebut} = n_b - 1$$

Contoh Penggunaan Uji Fisher:



Untuk data terdiri dari 2 kelompok (Uji F)

Suatu data penelitian untuk Mengetahui kinerja guru berdasarkan golongan kepangkatannya. Kemudian dibuat suatu alat ukur kinerja guru. Dengan menggunakan alat tersebut diperoleh skor kinerja guru dari sebanyak 70 orang responden. Adapun ringkasan data dari kinerja guru tersebut berdasarkan golongan seperti pada tabel berikut:

Golongan	Jumlah resp.	Jumlah skor kinerja	Rata-rata skor	Varians data
II	20	1894	92,45	8,23
III	50	4634	92,68	8,46

Langkah Pengujian

1. Varians dari setiap kelompok sampel:
Varians dari Golongan II
 $S_1^2 = 8,23$; dengan dk =
 $20-1 = 19$.

Varians dari Golongan III
 $S_2^2 = 8,46$; dengan dk =
 $50-1 = 49$.

2. Menghitung nilai F,
yaitu:
 $F = S_1^2 / S_2^2 = 8,23 / 8,46 =$
 $0,973$

3. Melihat nilai F_{tabel} ,
dengan $dk_1 = 19$ dan $dk_2 =$
 49 pada $\alpha = 5\%$ yaitu:
 $F_{\text{tabel}}(0,05; 19; 49) =$
 $1,803$

Karena $F_{\text{hitung}} < F_{\text{tabel}}(0,05; 19; 49)$ yaitu $0,973 < 1,803$, maka Terima H_0 .

Hal ini bermakna, bahwa varians skor data kinerja guru kelompok golongan II dengan kelompok golongan III homogeny pada taraf kepercayaan 95%

B

UJI BARTLETT



langkah-langkah Uji Bartlett menurut Supardi adalah (2013:145) :

1. Sajikan data semua kelompok sampel
2. Menghitung rerata (mean) dan varian serta derajat kebebasan (dk) setiap kelompok data yang akan diuji homogenitasnya
3. Sajikan $dk = (n-1)$ dan varian tiap kelompok sampel dalam table, serta sekaligus hitung nilai logaritma dari setiap varian kelompok dan hasil kali dk dengan logaritma varian dari tiap kelompok sampel
4. Hitung varian gabungan dari semua kelompok sampel: $S^2 = \frac{\sum(n_i - 1) s_i^2}{\sum(n_i - 1)}$
5. Hitung harga logaritma varian gabungan dan harga satuan Bartlett (B), dengan rumus:

$$B = (\log s^2) \sum (n_i - 1)$$

6. Hitung nilai chi kuadrat (x^2 *hitung*), dengan rumus:

$$x^2 \text{ hitung} = (\ln 10) (B - \sum dk \cdot \log s_i^2)$$

7. Tentukan harga chi kuadrat tabel (x^2 *tabel*), pada taraf nyata misal $\alpha = 0,05$ dan derajat kebebasan (dk) = $k - 1$, yaitu:

$$x^2 \text{ tabel} = x_{(1-\alpha)(k-1)}$$

(dalam hal ini k = banyaknya kelompok sampel)

8. Menguji hipotesis homogenitas data dengan cara membandingkan nilai x^2 *hitung* dengan x^2 *tabel* Kriteria pengujian adalah:

Tolak H_0 jika $x^2 \text{ hitung} > x^2 \text{ tabel}$

Terima H_0 jika $x^2 \text{ hitung} < x^2 \text{ tabel}$

Contoh penggunaan Uji Bartlett:

	Jenis Makanan			
	1	2	3	4
Pertambahan Berat Badan	12	14	6	9
	20	15	16	14
	23	10	16	18
	10	19	20	19
	17	22		

Dengan varians masing-masing kelompok sampel sebagai berikut:

$$s_1^2 = 29,3$$

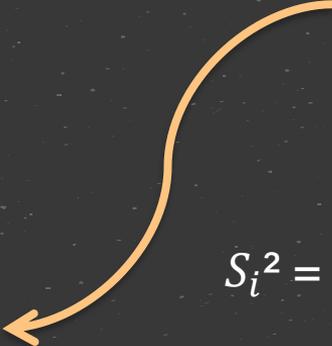
$$s_2^2 = 21,5$$

$$s_3^2 = 35,7$$

$$s_4^2 = 20,7$$

masukkan angka ke dalam table

Sampel ke	Dk	1/dk	si ²	log si ²	dk (log si ²)
1	4	0,25	29,3	1,4669	5,8676
2	4	0,25	21,5	1,3324	5,3296
3	3	0,33	35,7	1,5527	4,6581
4	3	0,33	20,7	1,3160	3,9460
	16				19,8033


$$S_i^2 = \frac{n \cdot \sum X_i^2 - [\sum Xi]^2}{n(n-1)}$$

Langkah-Langkah

4. Hitung varians gabungan dari empat sampel:

$$S^2 = \frac{4(29,3)+4(21,5)+3(35,7)+3(20,7)}{4+4+3+3} = 26,6$$

$$\text{Log } S^2 = \log 26,6 = 1,4249$$

5. Hitung harga satuan Barlett

$$B = (\log s^2) \cdot \sum dk$$

$$B = (1,4249) \cdot (14)$$

$$B = 19,9486$$

6. Hitung nilai χ^2_{hitung}

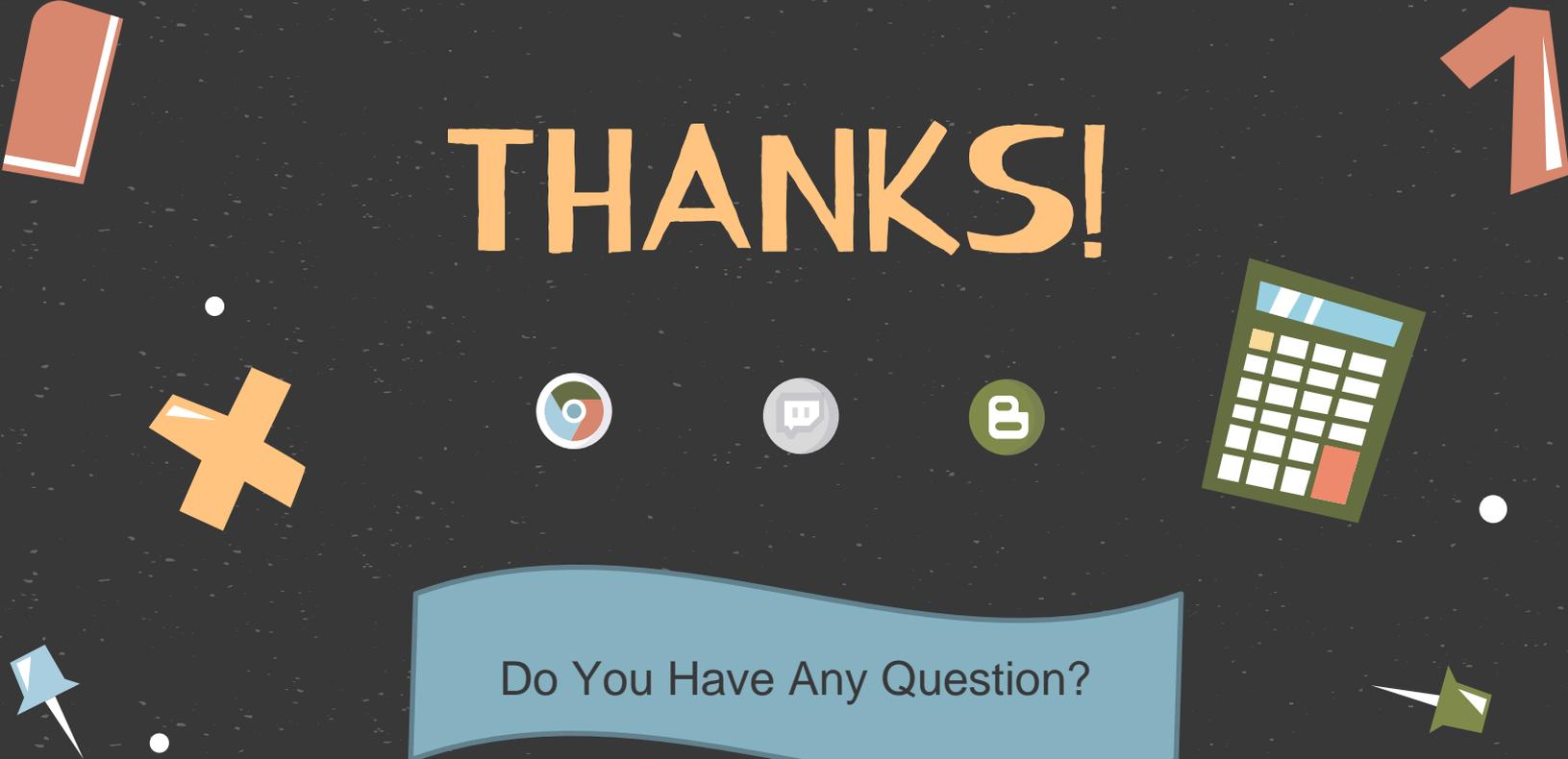
$$\chi^2_{\text{hitung}} = (\ln 10) \{B - \sum dk \cdot \log s^2\}$$

$$\chi^2_{\text{hitung}} = (2,3026) \{19,9486 - 19,8033\}$$

$$\chi^2_{\text{hitung}} = 0,3346$$

8. Jika $\alpha=0,05$, dari daftar distribusi chi kuadrat dengan $dk = 3 (k-1)$ didapat harga $\chi^2_{\text{tabel}(0,95)(3)} = 0,3518$. Karena $\chi^2_{\text{hitung}} <$ dari χ^2_{tabel} ($0,3346 < 0,3518$) berarti data diperoleh dari populasi yang homogen.

7. Lihat nilai χ^2_{tabel}



THANKS!

Do You Have Any Question?



HIPOTESIS, POPULASI DAN SAMPEL

HIPOTESIS

Dalam statistik dan penelitian terdapat dua macam hipotesis, yaitu hipotesis nol dan alternatif. Pada statistik, hipotesis nol diartikan sebagai tidak adanya perbedaan antara parameter dengan statistik, atau tidak adanya perbedaan antara ukuran populasi dengan ukuran sample. Hipotesis alternative adalah lawannya hipotesis nol, yang berbunyi adanya perbedaan antara data populasi dan data sampel.

Dalam penelitian, hipotesis nol juga menyatakan “tidak ada”, tetapi bukan tidak adanya perbedaan antara populasi dan data sampel, tetapi bisa berbentuk tidak adanya hubungan antara satu variabel dengan variabel lain, tidak adanya perbedaan antara satu variabel atau lebih pada populasi atau sampel yang berbeda, dan tidak adanya perbedaan antara yang diharapkan dengan kenyataan pada satu variabel atau lebih untuk populasi atau sampel yang sama.

TIGA BENTUK RUMUSAN HIPOTESIS

I. Hipotesis Deskriptif

Hipotesis Deskriptif adalah dugaan tentang nilai suatu variabel mandiri, tidak membuat perbandingan atau hubungan. Sebagai contoh, bila rumusan masalah penelitian sebagai berikut, maka hipotesis (jawaban sementara) yang dirumuskan adalah hipotesis deskriptif.

- a. Seberapa tinggi daya tahan lampu merk X?
- b. Seberapa tinggi produktivitas padi di Kabupaten Klaten?
- c. Berapa lama daya tahan lampu merk A dan B?
- d. Seberapa baik gaya kepemimpinan di lembaga X?

Dari empat pernyataan tersebut antara lain dapat dirumuskan hipotesis seperti berikut?

- a. Daya tahan lampu merk X = 800 jam.
- b. Produktivitas padi di Kabupaten Klaten 8 ton/ha.
- c. Daya tahan lampu merk A = 450 jam dan merk B = 600 jam.
- d. Gaya kepemimpinan di lembaga X telah mencapai 70% dari yang diharapkan

Berikut ini diberikan contoh berbagai pernyataan yang dapat dirumuskan hipotesis deskriptif-statistiknya.

1. Suatu perusahaan minuman harus mengikuti ketentuan, bahwa salah satu unsur kimia hanya boleh dicampurkan paling banyak 1%. (Paling banyak berarti lebih kecil atau sama dengan \leq)

Dengan demikian rumusan hipotesis statistiknya adalah:

- $H_0 : \mu \leq 0,01 ; \leq$ (Lebih kecil atau sama dengan)
- $H_a : \mu > 0,01 ; >$ (Lebih besar)

Dapat dibaca: Hipotesis nol untuk parameter populasi berbentuk proporsi (1% proporsi) lebih kecil atau sama dengan 1%, dan hipotesis alternatifnya, untuk populasi yang berbentuk proporsi lebih besar dari 1%.

2. HIPOTESIS KOMPARATIF

Hipotesis komparatif adalah pernyataan yang menunjukkan dugaan nilai dalam satu variabel atau lebih pada sample yang berbeda. Contoh rumusan masalah komparatif dan hipotesisnya.

- a. Apakah ada perbedaan daya tahan lampu merk A dan B?
 - b. Apakah ada perbedaan produktivitas kerja antara pegawai golongan I, II, III?
- Rumusan Hipotesis adalah:
 1. Tidak terdapat perbedaan daya tahan lampu antara lampu merk A dan B.
 2. Daya tahan lampu merk B paling kecil sama dengan lampu merk A.
 3. Daya tahan lampu merk B paling tinggi sama dengan lampu merk A.

Hipotesis statistiknya adalah:

- $H_0 : \mu_1 = \mu_2$
- $H_a : \mu_1 \neq \mu_2$

Rumusan uji hipotesis dua pihak

- $H_0 : \mu_1 \geq \mu_2$
- $H_a : \mu_1 < \mu_2$

Rumusan hipotesis uji satu pihak

- $H_0 : \mu_1 \leq \mu_2$
- $H_a : \mu_1 > \mu_2$
- Rumusan Hipotesis uji satu pihak

Tidak terdapat perbedaan (ada persamaan) produktivitas kerja antara Golongan I, II, III.

Rumusan hipotesis statistiknya adalah:

- $H_0 : \mu_1 = \mu_2 = \mu_3$
- $H_a : \mu_1 \neq \mu_2 = \mu_3$ (Salah satu berbeda sudah merupakan H_a)

3. HIPOTESIS HUBUNGAN (ASOSIATIF)

Hipotesis asosiatif adalah suatu pernyataan yang lebih. Contoh rumusan masalahnya adalah “Apakah ada hubungan antara Gaya Kepemimpinan dengan Efektifitas Kerja?”. Rumus dan hipotesis nolnya adalah: Tidak ada hubungan antar gaya kepemimpinan dengan efektivitas kerja.

Hipotesis statistiknya adalah:

- $H_0 : \rho = 0$
- $H_a : \rho \neq 0$ (ρ = simbol yang menunjukkan kuatnya hubungan)

Dapat dibaca: Hipotesis nol, yang menunjukkan tidak adanya hubungan (nol = tidak ada hubungan) antara Gaya Kepemimpinana dengan Efektifitas Kerja dalam populasi. Hipotesis alternatifnya menunjukkan ada hubungan (tidak sama dengan nol, mungkin lebih besar dari 0 atau lebih kecil dari nol).

DUA KESALAHAN DALAM PENGUJIAN HIPOTESIS

Dalam menaksir parameter populasi berdasarkan data sampel, kemungkinan akan terdapat dua kesalahan yaitu:

1. Kesalahan Tipe I adalah suatu kesalahan bila menolak hipotesis nol (H_0) yang benar (seharusnya diterima). Dalam hal ini tingkat kesalahan dinyatakan dengan α . (baca alpha)
2. Kesalahan Tipe II adalah kesalahan bila menerima hipotesis yang salah (seharusnya ditolak). Tingkat kesalahan untuk ini dinyatakan dengan β (baca betha).

POPULASI

Populasi berasal dari kata bahasa Inggris population, yang berarti jumlah penduduk. Populasi adalah wilayah generalisasi yang terdiri atas obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulan (Sugiyono, 2013 : 117).

Populasi adalah wilayah generalisasi yang terdiri atas : obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan peneliti untuk dipelajari dan kemudian ditarik kesimpulannya (Sugiyono (2006:117)

Menurut Muri (2007:182) Secara Umum Dapat Dikatakan Beberapa Karakteristik Populasi Adalah:

- a. Merupakan keseluruhan dari unit analisis sesuai dengan informasi yang akan diinginkan.
- b. Dapat berupa manusia/individu, hewan, tumbuh tumbuhan, benda benda atau objek maupun kejadian kejadian yang terdapat dalam suatu area/daerah tertentu yang telah ditetapkan.
- c. Merupakan batas batas (boundary) yang mempunyai sifat sifat tertentu yang memungkinkan peneliti menarik kesimpulan dari keadaan itu.
- d. Memberikan pedoman kepada apa atau siapa hasil penelitian itu dapat digeneralisasikan.

SAMPEL

Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi. Bila populasi besar, dan peneliti tidak mungkin mempelajari semua yang ada pada populasi, misalnya karena keterbatasan dana, tenaga, dan waktu, maka penelitian dapat menggunakan sampel yang diambil dari populasi itu. Kesimpulan yang dipelajari dari sampel akan dapat diberlakukan untuk populasi.

- Ciri ciri sampel yang baik :

- 1) **Sampel dipilih dengan cara hati hati**, dengan menggunakan cara tertentu dan benar.
- 2) **Sampel harus mewakili populasi**, sehingga gambaran yang diberikan mewakili keseluruhan karakteristik yang terdapat pada populasi.
- 3) **Besarnya ukuran sampel**, hendaknya mempertimbangkan tingkat kesalahan sampel yang dapat ditolerir dan tingkat kepercayaan yang dapat diterima secara statistik.

CARA PENGAMBILAN SAMPEL ATAU TEKNIK SAMPLING

PROBABILITY SAMPLING

- Simple Random sampling
- Proportionate stratified Random Sampling
- Dispproportionate Stratified Random Sampling
- Cluster Sampling (Area Sampling)

NON PROBABILITY SAMPLING

- Sampling Sistematis
- Sampling Kuota
- Sampling Aksidental
- Sampling Purposive
- Sampling Totall
- Snowball Sampling

PROBABILITY SAMPLING

1) Simple Random sampling

Dikatakan simple (sederhana) karena pengambilan anggota sampel dari populasi dilakukan secara acak tanpa memperhatikan strata yang ada dalam populasi itu. Pengambilan sampel acak yaitu dengan cara undian, memilih bilangan dari daftar bilangan secara acak.

2) Proportionate stratified Random Sampling

Teknik ini digunakan bila populasi mempunyai anggota/unsure yang tidak homogen dan bersastra secara proposional. Suatu organisasi yang mempunyai pegawai dari latar belakang pendidikan yang berstrata, maka populasi pegawai itu bersrata. Jumlah sample yang harus diambil meliputi strata pendidikan tersebut.

PROBABILITY SAMPLING

3) Dispproportionate Stratified Random Sampling

Teknik ini digunakan untuk menentukan jumlah sampel, bila populasi berstrata tetapi kurang proporsional. Misalnya pegawai dari unit kerja tertentu mempunyai; 3 orang lulusan S3, 4 orang lulusan S2, 90 orang S1, 800 orang SMU, 700 orang SMP, maka 3 orang lulusan S3 dan 4 orang S2, itu diambil semua sebagai sample karena dua kelompok ini terlalu kecil bila dibandingkan dengan kelompok S1, SMU, dan SMP.

4) Cluster Sampling (Area Sampling)

Teknik sampling daerah digunakan untuk menentukan sampel bila obyek yang akan diteliti atau sumber data sangat luas, misal penduduk dari suatu negara, provinsi atau kabupaten. Untuk menentukan penduduk mana yang akan dijadikan sumber data, maka pengambilan sampel ditetapkan secara bertahap dari wilayah yang luas (negara) sampai ke wilayah terkecil (kabupaten). Setelah terpilih sampel terkecil, kemudian baru dipilih sampel secara acak.

NON PROBABILITY SAMPLING

1) Sampling Sistematis

Sampling sistematis adalah teknik pengambilan sampel berdasarkan urutan dari anggota populasi yang telah diberi nomor urut. Misalnya anggota populasi yang terdiri dari 100 orang. Dari semua anggota itu diberi nomor urut, yaitu nomor 1 sampai dengan nomor 100. Pengambilan sampel dapat dilakukan dengan mengambil nomor ganjil saja, genap saja, atau kelipatan dari bilangan tertentu.

2) Sampling Kuota

Sampling kuota adalah teknik untuk menentukan sampel dari populasi yang mempunyai ciri-ciri tertentu sampai jumlah (kuota) yang diinginkan. Sebagai contoh, akan melakukan penelitian tentang pendapat masyarakat terhadap pelayanan masyarakat dalam urusan Ijin Mendirikan Bangunan (IMB). Jumlah sampel yang ditentukan 500 orang. Kalau pengumpulan data belum memenuhi kuota 500 orang tersebut, maka penelitian dipandang belum selesai.

NON PROBABILITY SAMPLING

3) Sampling Aksidental

Sampling Aksidental adalah teknik penentuan sampel berdasarkan kebetulan, yaitu siapa saja yang secara kebetulan/insidental bertemu dengan peneliti dapat digunakan sebagai sampel, bila dipandang orang yang kebetulan ditemui itu cocok sebagai sumber data.

4) Sampling Purposive

Sampling purposive adalah teknik penentuan sampel dengan pertimbangan tertentu. Misalnya akan melakukan penelitian tentang Kualitas makanan, maka sampel sumber datanya adalah orang yang ahli makanan, atau penelitian tentang kondisi politik di suatu daerah, maka sampel sumber datanya adalah orang yang ahli politik.

NON PROBABILITY SAMPLING

5) Sampling Total

Sampling total adalah teknik penentuan sampel bila semua anggota populasi digunakan sebagai sampel. Hal ini sering dilakukan bila jumlah populasi relatif kecil, kurang dari 30 orang, istilah lain sampel total adalah sensus, dimana semua anggota populasi dijadikan sampel.

6) Snowball Sampling

Snowball sampling adalah teknik penentuan sampel yang mula-mula jumlahnya kecil, kemudian membesar. ibarat bola salju yang menggelinding yang lama-lama menjadi besar. Dalam penentuan sampel, pertama-tama dipilih satu atau dua orang, tetapi karena dengan dua Orang ini belum merasa lengkap terhadap data yang diberikan, maka peneliti mencari orang lain yang dipandang lebih tahu dan dapat melengkapi data yang diberikan oleh dua orang sebelumnya. Begitu seterusnya, sehingga jumlah sampel semakin banyak.

UKURAN SAMPEL

- Rumus Slovin

Rumusan slovin untuk menentukan sampel minimal (n) jika diketahui ukuran populasi (N) pada taraf signifikansi α adalah:

$$n = \frac{N}{1+N\alpha^2}$$

- Contoh:

Berapa ukuran sampel minimum yang harus diambil dari populasi yang berukuran 1000 dengan taraf signifikansi $\alpha = 0,05$?

$$n = \frac{N}{1+N\alpha^2}$$

$$n = \frac{1000}{1+1000(0,05)^2}$$

$$n = 285,713$$

$$n = 286$$

- Jadi untuk ukuran populasi 1000 dengan taraf signifikansi 0,05 ukuran sampel yang harus diambil adalah 286 sampel.



STATISTIK PARAMETRIK DAN NON-PARAMETRIK

Statistik Parametrik



Statistik Parametrik yaitu ilmu statistik yang mempertimbangkan jenis sebaran atau distribusi data, yaitu apakah data menyebar secara normal atau tidak. Dengan kata lain, data yang akan dianalisis menggunakan statistik parametrik harus memenuhi asumsi normalitas. Pada umumnya, jika data tidak menyebar normal, maka data seharusnya dikerjakan dengan metode statistik non-parametrik, atau setidaknya dilakukan transformasi terlebih dahulu agar data mengikuti sebaran normal, sehingga bisa dikerjakan dengan statistik parametrik

metode statistik parametrik .



- a. Uji-z (1 atau 2 sampel)
- b. Uji-t (1 atau 2 sampel)
- c. Korelasi pearson
- d. Perancangan percobaan (*one or two-way* anova parametrik), dll.

Ciri-ciri statistik parametrik

- Data dengan skala interval dan rasio
- Data menyebar/berdistribusi normal

Keunggulan dan kelemahan statistik parametrik

Keunggulan :

1. Syarat syarat parameter dari suatu populasi yang menjadi sampel biasanya tidak diuji dan dianggap memenuhi syarat, pengukuran terhadap data dilakukan dengan kuat.
2. Observasi bebas satu sama lain dan ditarik dari populasi yang berdistribusi normal serta memiliki varian yang homogen.

Kelemahan :

1. Populasi harus memiliki varian yang sama.
2. Variabel-variabel yang diteliti harus dapat diukur setidaknya dalam skala interval.
3. Dalam analisis varian ditambahkan persyaratan rata-rata dari populasi harus normal dan bervarian sama, dan harus merupakan kombinasi linear dari efek-efek yang ditimbulkan

Distribusi Normal



Salah satu asumsi yang harus dipenuhi dalam statistik parametris adalah data yang dianalisis harus berdistribusi normal. Dalam banyak hal distribusi normal dapat dipandang sebagai model atau dasar bagi teori statistika modern. Distribusi normal banyak digunakan untuk menghampiri distribusi data hasil penelitian. Distribusi normal memegang peranan yang sangat penting dalam statistik inferensial, yaitu sebagai model distribusi peluang (probability distribution).

KARAKTERISTIK DISTRIBUSI NORMAL

- 1) Unimodal, terdiri dari dua kata yaitu Uni = satu dan modal = modus, distribusi normal memiliki hanya satu modus.
- 2) Simetrik, yaitu jika data dibagi menjadi dua pada bagian median, maka distribusi frekuensi skor yang berada di atas median sama dengan distribusi frekuensi skor di bawah median.
- 3) Identik, yaitu nilai modus, median dan rata-rata pada distribusi normal adalah sama. (modus = median = rata rata)
- 4) Asimtotik, yaitu kurva distribusi normal tidak akan pernah menyentuh absisnya, yaitu distribusi normal terbentuk dari perangkat dari skor yang bersifat kontinu dari mulai data yang tak hingga sampai dengan nilai yang tak hingga pula.

Distribusi Normal Baku



Distribusi normal baku adalah distribusi normal yang memiliki empat ciri-ciri sebagaimana Distribusi Normal dengan ditambah syarat rata-rata $\mu = 0$ dan simpangan baku $\sigma \in 1$, sehingga syarat-syarat Distribusi Normal baku adalah sebagai berikut:

- 1). Unimodal
- 2). Simetrik
- 3). Identik
- 4). Asimtotik
- 5). Rata-rata nilai = 0
- 6). Simpangan Baku nilai = 1

Di Bawah Kurva Normal

Distribusi normal dimanfaatkan sebagai rujukan dalam menafsirkan data apabila distribusi data itu dapat dihampiri oleh model distribusi normal. Daerah di bawah kurva normal, luasan daerah itu menunjukkan peluang munculnya nilai perubah acak yang memiliki distribusi normal baku pada interval 0 sampai dengan z untuk $z = 0,0; 0,01; 0,02 \dots 0,09$ dst. Oleh karena distribusi normal bersifat simetrik terhadap rata-ratanya, maka kita tidak perlu menghitung luas daerah dari 0 ke skor z yang bertanda negatif.

Macam-macam Uji Statistic Parametrik

1. One-sample t test
One-Sample T Test digunakan untuk menguji perbedaan rata-rata suatu sampel dengan suatu nilai hipotesis.
2. Independent sampel t test
Independent Sampel T Test untuk menguji signifikansi beda rata-rata 2 kelompok. Biasanya digunakan untuk menguji pengaruh 1 variabel independen terhadap 1 variabel dependen atau lebih.
3. Paired-sampel t test
Paired-Sampel T Test adalah 2 pengukuran pada subjek yang sama terhadap suatu pengaruh. Yaitu pengaruh dari sebelum dan sesudah mengalami perngaruh (perlakuan)
4. Analisis variance (one-way anova)
Analisis Variance (One-Way Anova) : Analisis Varian untuk satu variabel Independen digunakan untuk menentukan apakah rata-rata dua/lebih kelompok berbeda secara nyata.
5. Analisis general linear model (glm)- univariate
Analisis General Linear Model (GLM) - Univariate merupakan analisis regresi dan varian variabel dependen dengan 2/lebih variabel faktor atau variabel lainnya.

Statistik Non Parametrik

yaitu statistik bebas sebaran (tidak mensyaratkan bentuk sebaran parameter populasi, baik normal atau tidak). Selain itu, statistik non-parametrik biasanya menggunakan skala pengukuran sosial, yakni nominal dan ordinal yang umumnya tidak berdistribusi normal.

Ciri-ciri Statistik Non Parametrik

1. Data tidak berdistribusi normal
2. Umumnya data berskala nominal dan ordinal
3. Umumnya dilakukan pada penelitian sosial
4. Umumnya jumlah sampel kecil

01

02

03

04

05

06



Metode-metode Statistik Non Parametrik

a. Uji tanda (sign test)

Uji tanda digunakan untuk membandingkan data yang berpasangan. Adapun syarat-syarat pada uji tanda yang harus dipenuhi, yaitu:

- v Pasangan hasil pengamatan yang sedang dibandingkan bersifat independent
- v Masing-masing pengamatan dalam tiap pasang terjadi karena pengaruh kondisi yang serupa.
- v Pasangan yang berlainan terjadi karena kondisi yang berbeda.

Rumus Uji Tanda $\frac{1}{2} (n - 1) - K$

b. Rank sum test (wilcoxon)

Merupakan suatu uji yang menghitung tanda dan besarnya selisih dari dua buah rata-rata populasi. Uji ini lebih peka dari pada uji tanda dalam menemukan perbedaan antara populasi. Dengan kata lain, uji peringkat bertanda wilcoxon digunakan jika besaran maupun arah perbedaan relevan

Untuk menentukan apakah terdapat perbedaan yang sesungguhnya antara pasangan data yang diambil dari dua sampel yang berkait.

Prosedur Uji wilcoxon Untuk Pengamatan Berpasangan

1. Menyatakan hipotesis nol dan hipotesis alternatif (H_0 dan H_1).
2. Memilih tarap keberartian.
3. Menentukan daerah kritis W (bila dist Z digunakan).
4. Menyusun peringkat tanpa memperhatikan tanda.
5. Pemberian tanda atas peringkat yang telah ditetapkan.
6. Menjumlahkan peringkat dengan jumlah terkecil sebagai W .
7. Penarikan kesimpulan statistic tentang hipotesis nol (tolak H_0 atau terima H_0)



Koefisien Korelasi Peringkat .

Keofisien korelasi pangkat atau koefisien Spearman adalah derajat hubungan yang mengukur korelasi pangkat. Yang dimana korelasi pangkat bersimbol r' (baca: er aksen).

Misalkan pasangan data hasil pengamatan $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ kita susun menurut urutan besar nilainya dalam tiap Variabel. Nilai X_1 disusun menurut urutan besarnya, yang terbesar diberi nomor urut atau peringkat 1, terbesar kedua diberi peringkat 2, terbesar ketiga diberi peringkat 3, dan seterusnya sampai kepada nilai X_1 terkecil diberi peringkat n demikian pula untuk Variabel Y_1 . Sekarang kita bentuk selisih atau beda peringkat X_1 dan Y_1 yang data aslinya berpasangan. Sebutlah beda ini b_1 .

Maka Koefisien korelasi peringkat r' antara serentetan pasangan X_1 dan Y_1 dihitung dengan rumus :

Harga r' bergerak dari -1 sampai dengan $+1$, sebagaimana halnya koefisien korelasi r biasa. Harga $r' = +1$ berarty terdapat persesuaian yang sempurna antara X_i dan Y_i sedangkan $r' = -1$ menyatakan penilaian yang betul-betul bertentangan antara X_i dan Y_i .

Uji RUNTUN



Uji Runtun adalah pengujian yang berdasarkan adanya keruntunan.

- Runtun adalah barisan huruf-huruf atau tanda-tanda yang identic yang didahului atau didikuti oleh sebuah huruf atau sebuah tanda yang berbeda. Untuk runtun permulaan, barisan dimaksud tidak tidak didahului oleh huruf atau tanda apapun. Demikian pula untuk runtun terakhir, barisan itu tidak diakhiri oleh huruf atau tanda yang berbeda. Panjang runtun ditentukan oleh banyak huruf atau tanda dalam setiap runtun.

Dengan adanya runtun ini , kita dapat menguji hipotesis tentang :

- Data pengamatan telah diambil secara acak dari sebuah populasi, atau sampel yang diambil dari sebuah populasi adalah acak.
- Dua sampel acak berasal dari populasi yang sama atau dua populasi mempunyai distribusi yang sama.

Uji median

Hipotesis yang dihadapi :

H₀ : Dua sampel acak telah diambil dari dua populasi dengan median yang sama atau telah diambil dari populasi yang sama.

H₁ : kedua sampel itu berasal dari dua populasi dengan median yang berlainan atau dari dua populasi yang berlainan.

Uji kenormalan

Uji kenormalan secara parametric dengan menggunakan penaksiran rata-rata dan simpangan baku. Misalkan kita mempunyai sampel acak dengan hasil pengamatan X_1, X_2, \dots, X_n . Berdasarkan sampel ini akan diuji hipotesis nol bahwa sampel tersebut berasal dari populasi berdistribusi normal melawan hipotesis tandingan bahwa distribusi tidak normal. Untuk pengujian hipotesis nol tersebut kita tempuh prosedur berikut :

- Pengamatan x_1, x_2, \dots, x_n dijadikan bilangan baku z_1, z_2, \dots, z_n dengan menggunakan rumus $z_i = \frac{x_i - \bar{x}}{s}$ (\bar{x} dan s masing-masing merupakan rata-rata dan simpangan baku sampel).
- Untuk tiap bilangan baku ini dan menggunakan daftar distribusi normal baku, kemudian dihitung peluang $F(z_i) = P(Z \leq z_i)$.
- Selanjutnya dihitung proporsi z_i , z_2, \dots, z_n yang lebih kecil atau sama dengan z_i . jika proporsi ini dinyatakan oleh $S(z_i)$, maka $S(z_i) =$
- Hitung selisih $F(z_i) - S(z_i)$ kemudian tentukan harga mutlaknya.

Ambil harga yang paling besar di antara harga-harga mutlak selisih tersebut. Harga terbesar di beri lambing L_0 .



KEUNGGULAN DAN KELEMAHAN STATISTIK NONPARAMETRIK

Keunggulan

1. Tidak membutuhkan asumsi normalitas.
2. Secara umum metode statistik non-parametrik lebih mudah dikerjakan dan lebih mudah dimengerti jika dibandingkan dengan statistik parametrik karena ststistika non-parametrik tidak membutuhkan perhitungan matematik yang rumit seperti halnya statistik parametrik.
3. Statistik non-parametrik dapat digantikan data numerik (nominal) dengan jenjang (ordinal).
4. Kadang-kadang pada statistik non-parametrik tidak dibutuhkan urutan atau jenjang secara formal karena sering dijumpai hasil pengamatan yang dinyatakan dalam data kualitatif.
5. Pengujian hipotesis pada statistik non-parametrik dilakukan secara langsung pada pengamatan yang nyata.
6. Walaupun pada statistik non-parametrik tidak terikat pada distribusi normal populasi, tetapi dapat digunakan pada populasi berdistribusi normal.

Kelemahan

1. Statistik non-parametrik terkadang mengabaikan beberapa informasi tertentu.
2. Hasil pengujian hipotesis dengan statistik non-parametrik tidak setajam statistik parametrik.
3. Hasil statistik non-parametrik tidak dapat diekstrapolasikan ke populasi studi seperti pada statistik parametrik. Hal ini dikarenakan statistik non-parametrik mendekati eksperimen dengan sampel kecil dan umumnya membandingkan dua kelompok tertentu.

01

02

03

04

05

06





UJI KORELASI



Analisi Korelasi Korelasi

Analisi Korelasi Korelasi adalah istilah statistik yang menyatakan derajat hubungan linear antara dua variabel atau lebih, (Usman,2006:197). Hubungan antara dua variabel di dalam teknik korelasi bukanlah dalam arti hubungan sebab akibat (timbal balik), melainkan hanya merupakan hubungan searah saja. Misalnya tinggi badan menyebabkan berat badan bertambah tetapi berat badan bertambah belum tentu menyebabkan tinggi badan bertambah pula. Sehingga dalam korelasi dikenal penyebab dan akibatnya.

Data penyebab atau yang mempengaruhi disebut variabel bebas, disebut juga dengan independen yang biasa dilambangkan dengan huruf X atau X_1 X_2 X_3 ,... X_n . Sedangkan data akibat atau yang dipengaruhi disebut variabel terikat, disebut juga dependen yang biasa dilambangkan dengan huruf Y, (Usman,2006:197) Variabel-variabel yang akan dihubungkan terdiri atas berbagai tingkatan data meliputi data nominal, ordinal, interval, dan rasio. Tingkatan data tersebut menentukan analisis korelasi mana yang paling tepat digunakan.

TEKNIK KORELASI DUA VARIABEL BIVARIANT
UNTUK BERBAGAI VARIABEL

Teknik korelasi	Simbol	Variabel 1	Variabel 2	Keterangan
Product	R	Interval	interval	Teknik yang paling banyak dipakai, khususnya untuk mendapatkan standar kesalahan terkecil
Rank	P	Ordinal	ordinal	Sering dipakai sebagai pengganti produk momen terutama jika sampel kurang dari 30
Tan Kendal	T	Ordinal	ordinal	Untuk pengganti jika sampel kurang dari 10
Biserial	r_{bis}	Interval	interval	Kadang-kadang lebih dari 1 = standar kesalahan lebih

				besar dari r umumnya dipakai untuk analisis item
Biserial Widespread	r_{wbs}	Interval	interval	Khususnya dipakai untuk perseorangan yang ekstrem dalam dikotomi variabel
Point Biserial	r_{pbis}	Interval	interval	Hasilnya lebih tendah daripada r_{bis}
Tetrachoris	S^s	Dikotomi Artifisial Buatan	Dikotomi Artifisial Buatan	Digunakan jika kedua variabel dapat dipecah pada titik kritis
Phi	Φ	Dikotomi sebenarnya	Dikotomi sebenarnya	Digunakan pada perhitungan antara analisis item
Contingensi	E	2 kategori atau lebih	2 kategori atau lebih	Ialah kondisi khusus dapat dibandingkan dengan r_t – berhubungan erat dengan chi kuadrat
Rasio otomatis	H	Interval	interval	Digunakan untuk mengetahui hubungan nonlinear

(Disadur dari Usman, H, 2006, hal 199)

Analisis Korelasi bertujuan untuk mengetahui keeratan hubungan (kuat- lemahnya) hubungan antara variabel bebas X dengan variabel terikat Y, tanpa melihat bentuk hubungannya, apakah linear atau tidak linear. Kuat-lemahnya hubungan antara dua variabel dilihat dari koefisien korelasinya.

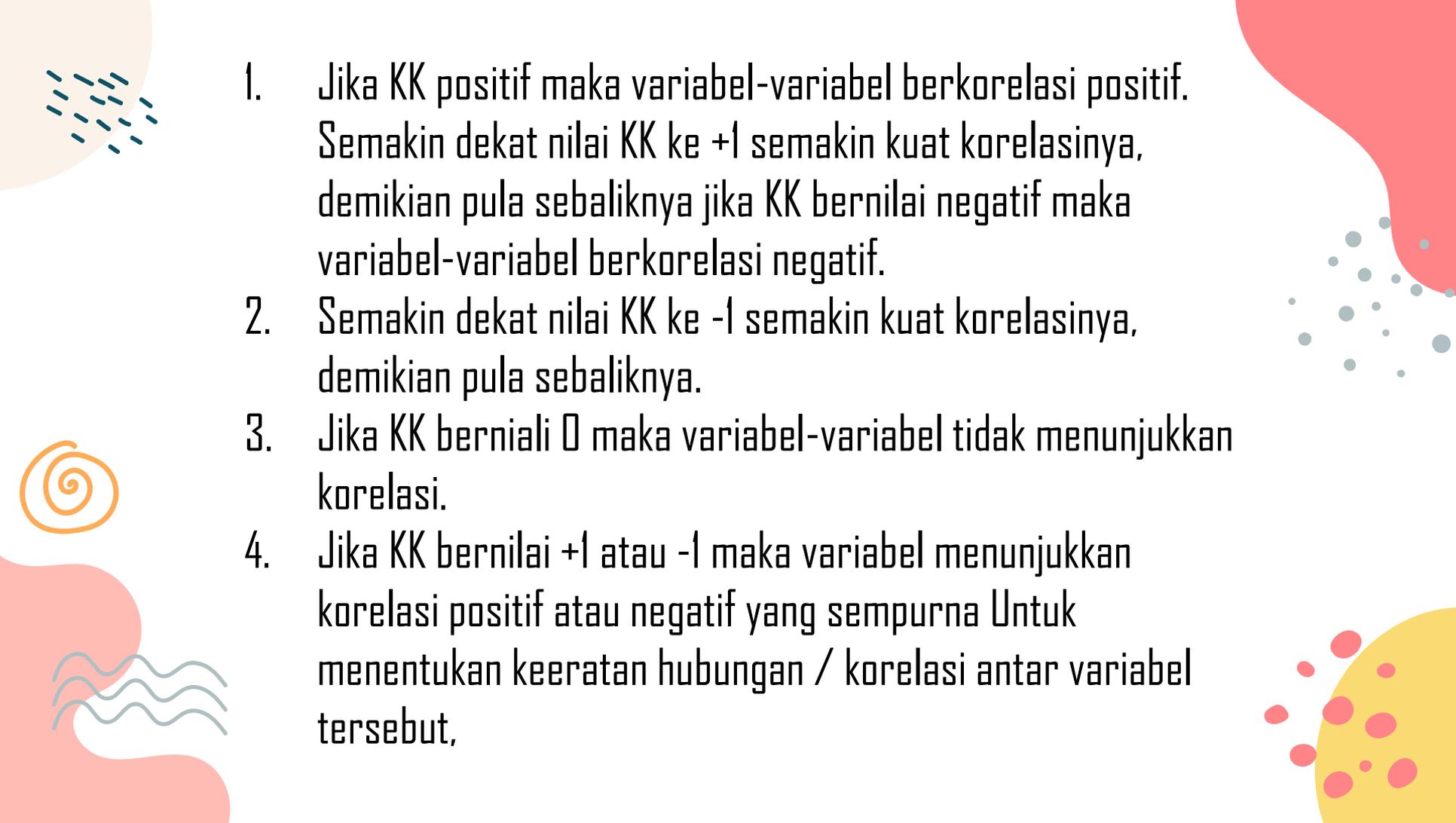
2. Koefisien Korelasi

Koefisien korelasi merupakan indeks atau bilangan yang digunakan untuk mengukur keeratan (kuat, lemah, atau tidak ada) hubungan antar variabel

Koefisien Korelasi	variabel yang diukur
1. Produk Momen Pearson	kedua variabelnya berskala interval
2. Order Rank Sperman	kedua variabelnya berskala ordinal

(Disadur dari Usman, H, 2006, hal 199)

Koefisien korelasi memiliki nilai antara -1 dan +1 ($-1 \leq KK \leq +1$), (Hasan, 2008: 234)

- 
1. Jika KK positif maka variabel-variabel berkorelasi positif. Semakin dekat nilai KK ke +1 semakin kuat korelasinya, demikian pula sebaliknya jika KK bernilai negatif maka variabel-variabel berkorelasi negatif.
 2. Semakin dekat nilai KK ke -1 semakin kuat korelasinya, demikian pula sebaliknya.
 3. Jika KK bernilai 0 maka variabel-variabel tidak menunjukkan korelasi.
 4. Jika KK bernilai +1 atau -1 maka variabel menunjukkan korelasi positif atau negatif yang sempurna Untuk menentukan keeratan hubungan / korelasi antar variabel tersebut,



Berikut ini diberikan nilai-nilai dari KK sebagai patokan, (Hasan, 2008: 234).

1. $KK = 0$ tidak ada korelasi
 2. $0 < KK \leq 0,20$ korelasi sangat rendah / lemah sekali
 3. $0,20 < KK \leq 0,40$ korelasi rendah / lemah tapi pasti
 4. $0,40 < KK \leq 0,70$ korelasi yang cukup berarti
 5. $0,70 < KK \leq 0,90$ korelasi yang tinggi; kuat
 6. $0,90 < KK \leq 1,00$ korelasi yang sangat tinggi; kuat sekali, dapat diandalkan.
 7. $KK = 1$ korelasi sempurna.
- 
- 

3. Jenis-jenis Koefisien/analisis Korelasi

a. Analisis Korelasi Person Product Moment (r) Teknik analisis Korelasi Product moment termasuk teknik statistik para metrik yang menggunakan interval dan ratio dengan persyaratan tertentu. Misalnya: data dipilih secara acak (random); datanya berdistribusi normal; data yang dihubungkan berpola linier; dan data yang dihubungkan mempunyai pasangan yang sama sesuai dengan subjek yang sama. Kalau salah satu tidak terpenuhi persyaratan tersebut analisis korelasi tidak dapat dilakukan.

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}} \quad \text{Sudjana (2002:369)}$$

Langkah-Langkah Menghitung Korelasi Product Moment

1. Membuat H_a dan H_o dalam bentuk kalimat
2. Membuat Tabel
3. Mencari r hitung
4. Mencari besarnya sumbangan variabel X terhadap variabel Y
5. Menghitung signifikansi dengan rumus thitung
6. Membuat kesimpulan

Analisis Korelasi Rank Spearman (P) Korelasi rank dipakai apabila:

1. kedua variabel yang akan dikorelasikan itu mempunyai tingkatan data ordinal,
2. jumlah anggota sampel di bawah 30 (sampel kecil),
3. data tersebut memang diubah dari interval ke ordinal, dan
4. data interval tersebut ternyata tidak berdistribusi normal.

Korelasi rank ini ditemukan oleh Spearman, sehingga disebut juga sebagai korelasi Spearman. Korelasi ini dapat juga disebut sebagai korelasi bertingkat, korelasi berjenjang, korelasi berurutan, atau korelasi berpangkat. Besarnya hubungan antara dua variabel atau derajat hubungan yang mengukur korelasi berpangkat disebut koefisien korelasi berpangkat atau koefisien korelasi Spearman yang dinyatakan dengan lambang r_s . Makna dan kelayakan nilai r seperti halnya dengan yang diuraikan dalam korelasi Product moment.

Korelasi Spearman

$$r_s = 1 - \frac{6 \sum d^2}{n^3 - n}$$

d = selisih ranking X dan Y

n = banyak pasangan data

Korelasi Spearman—Brown

$$r_{sb} = 1 - \frac{2r_s}{1 + r_s}$$

4. Koefisien Penentu (KP) atau Koefisien Determinasi (R)

Apabila koefisien korelasi dikuadratkan, akan menjadi koefisien penentu (KP) atau koefisien determinasi, yang artinya penyebab perubahan pada variabel Y yang datang dari variabel X, sebesar kuadrat koefisien korelasinya. Koefisien penentu ini menjelaskan besarnya pengaruh nilai suatu variabel (variabel X) terhadap naik turunnya (variasi) nilai variabel lainnya (variabel Y).
Dirumuskan: $KP = R = (K)^2 \times 100\%$ Nilai koefisien penentu ini terletak antara 0 dan +1 ($0 \leq KP \leq +1$). Jika koefisien korelasinya adalah koefisien korelasi Pearson (r) maka koefisien penentunya $KP = R = (r)^2 \times 100\%$

5. Pendugaan Koefisien Korelasi Populasi

- Pendugaan Koefisien Korelasi Populasi (interval keyakinan ρ) menggunakan distribusi Z.
- Pendugaannya dapat dilakukan dengan terlebih dahulu mengubah koefisien korelasi sampel r menjadi nilai Z_r , yang dalam bentuk persamaan dituliskan

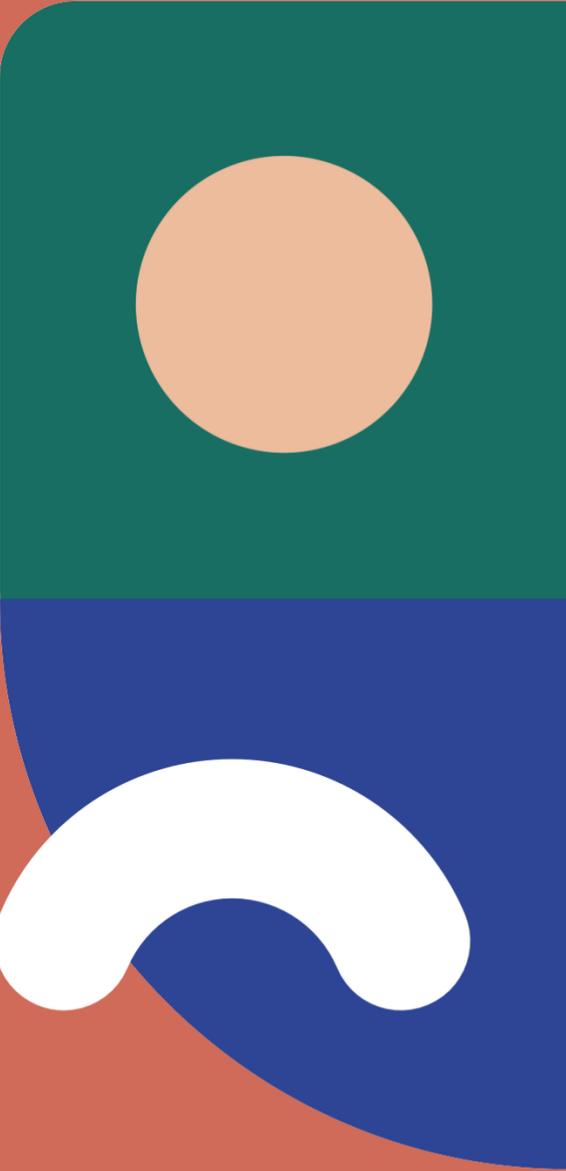
$$Z_r = \frac{1}{2} \ln \frac{1+r}{1-r}$$

Variabel Z_r akan mendekati distribusi normal dengan rata-rata dan varians sebagai berikut

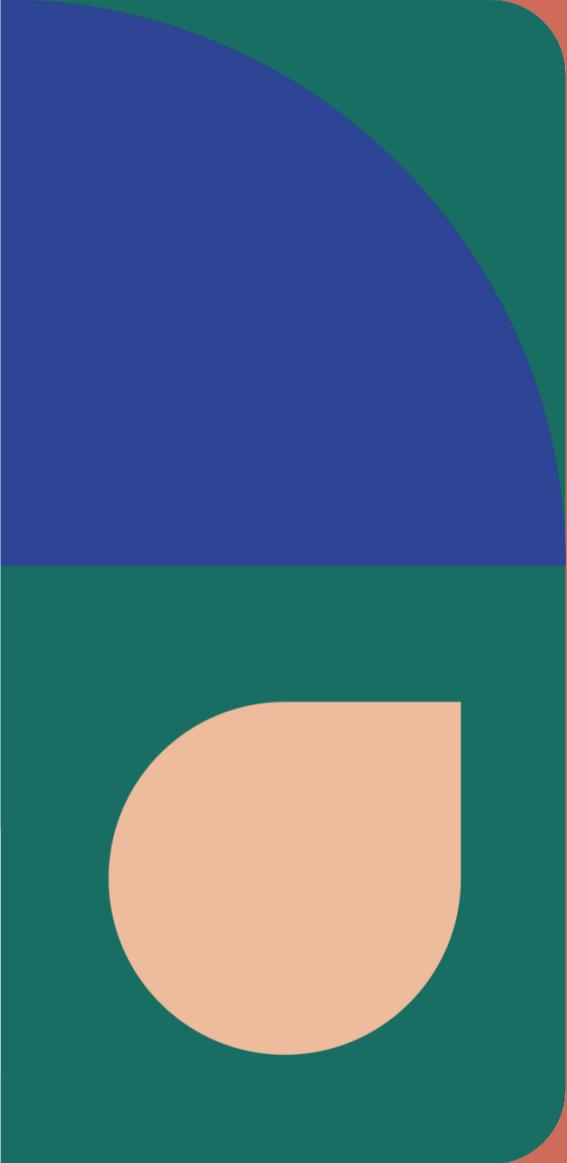
$$\mu_{Z_r} = (1,1513) \log \frac{1+\rho}{1-\rho}$$



Thanks!



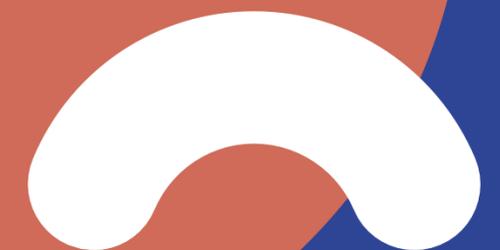
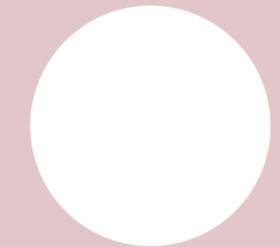
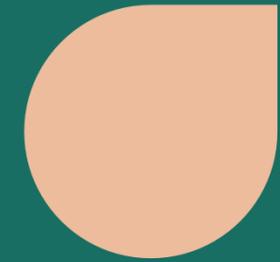
**UJI REGRESI
LINIER
SEDERHANA**



A. Definisi Regresi Linier

Regresi merupakan suatu alat ukur yang digunakan untuk mengukur ada atau tidaknya hubungan antar variabel. Dalam analisis regresi, suatu persamaan regresi atau persamaan penduga dibentuk untuk menerangkan pola hubungan variabel-variabel apakah ada hubungan antara 2 variabel atau lebih. Hubungan yang didapat pada umumnya menyatakan hubungan fungsional antara variabel-variabel.

Regresi linier sederhana merupakan suatu prosedur untuk mendapatkan hubungan matematis dalam bentuk persamaan antara variabel tak bebas dengan variabel bebas tunggal. Dalam regresi linier sederhana hanya ada satu variabel bebas x yang dihubungkan dengan satu variabel tak bebas y .



B. Persamaan Regresi

Persamaan regresi (regression equation) adalah suatu persamaan matematis yang mendefinisikan hubungan antara dua variabel. Persamaan regresi yang digunakan untuk membuat taksiran mengenai variabel dependent disebut persamaan regresi estimasi, yaitu suatu formula matematis yang menunjukkan hubungan antara satu atau beberapa variabel yang nilainya sudah diketahui, dengan satu variabel yang nilainya belum diketahui. Sifat hubungan antar variabel dalam persamaan regresi merupakan hubungan sebab akibat (causal relationship).

C. Analisis Regresi

Analisis regresi merupakan studi ketergantungan satu atau lebih X (variabel bebas) terhadap Y (variabel terikat), dengan maksud untuk meramalkan nilai Y . Tujuan analisis regresi adalah mendapatkan pola hubungan secara matematis antara X dan Y , mengetahui besarnya perubahan variabel X terhadap Y , dan memprediksi Y jika nilai X diketahui. Sehingga dalam suatu persamaan regresi terdapat dua macam variabel, yaitu dependent variabel (variabel terikat, respon) dan independent variabel (variabel bebas, prediktor).

Syarat-syarat regresi antara lain data harus berbentuk interval atau rasio, data berdistribusi normal, adanya korelasi antarvariabel, dan tidak terdapat korelasi antarvariabel bebasnya untuk regresi ganda.

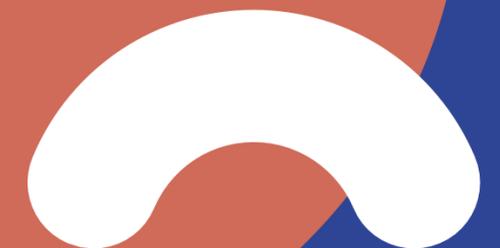
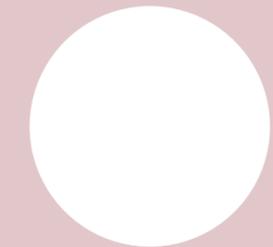
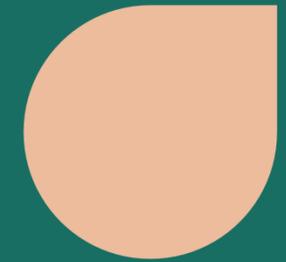
Berdasarkan banyak dan jenisnya data, analisis regresi dapat dibedakan atas :

1. Regresi linier
2. Regresi non linier

D. Kriteria Data Regresi Linier

Terdapat dua syarat yang harus dipenuhi oleh data dalam menggunakan analisis regresi linier yaitu:

1. Data, harus terdiri dari dua jenis variabel, yaitu dependen dan independen. Selain itu data berupa kuantitatif dan variabel berupa kategori, seperti SD, SMA, SMK, dll.
2. Asumsi, setiap data diasumsikan variabel dependen terdistribusi secara normal. Selain itu, antara variabel dependen dan independen harus memiliki hubungan linier dengan observasi harus saling bebas



E. Langkah-langkah Analisis dan Uji Regresi Linier Sederhana

Adapun langkah-langkah yang perlu dilakukan untuk melakukan analisis dan uji regresi linier sederhana adalah sebagai berikut :

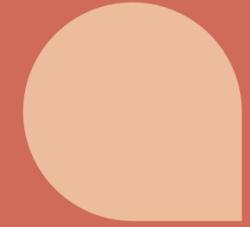
- Menentukan tujuan dari Analisis Regresi Linear Sederhana
- Mengidentifikasi variabel predictor dan variabel response
- Melakukan pengumpulan data dalam bentuk tabel frekuensi

E. Langkah-langkah Analisis dan Uji Regresi Linier Sederhana

- Menghitung X^2 , XY dan total dari masing-masingnya (Dalam tabel)
- Menghitung a dan b menggunakan rumus yang telah ditentukan
- Menghitung persamaan regresi
- Menghitung Koefisien korelasi (r): Untuk mengukur kekuatan hubungan antar variable predictor X dan response Y, dilakukan analisis korelasi yang hasilnya dinyatakan oleh suatu bilangan yang dikenal dengan koefisien korelasi. Biasanya analisis regresi sering dilakukan bersama-sama dengan analisis korelasi.
- Menghitung koefisien determinasi (r^2): Koefisien determinasi dapat ditentukan dengan mengkuadratkan koefisien korelasi.
- Uji signifikansi menggunakan Uji-t dan menentukan Taraf Signifikan

•

F. Contoh Kasus



Berikut contoh kasus untuk menganalisis regresi linier sederhana. Persamaan regresi linier sederhana secara umum yaitu: $\bar{Y} = a + bX$

Keterangan:

\bar{Y} = Respon (variabel terikat)

a = Constanta

b = Koefisien regresi variabel terikat

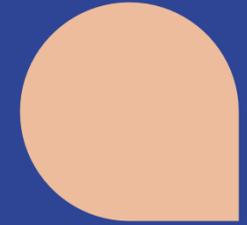
X = Prediktor (variabel bebas)

Di mana :

$$b = \frac{n(\sum XY) - (\sum X) \cdot (\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$a = \frac{(\sum Y) - b(\sum X)}{n}$$

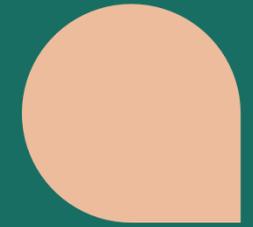
F. Contoh Kasus



Berikut ini adalah contoh data 10 responden yang berasal dari mahasiswa, untuk mengetahui pengaruh minat mahasiswa terhadap matakuliah data mining.

Minat (X)	18	16	20	18	14	15	16	18	17	15
Mata kuliah Data mining (Y)	21	18	23	21	16	20	21	17	19	17

Contoh Kasus



1. Perhitungan secara manual

a. Langkah 1 : Tujuan

Mengetahui pengaruh minat mahasiswa terhadap mata kuliah data mining.

b. Langkah 2 : Variabel

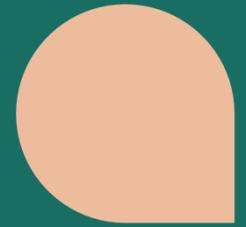
X (variable bebas/predictor) = Minat responden

Y (variable tak bebas/response) = mata kuliah data mining

Data:

Minat (X)	18	16	20	18	14	15	16	18	17	15
Mata kuliah Data mining (Y)	21	18	23	21	16	20	21	17	19	17

Contoh Kasus



c. Langkah 3 & 4 : Membuat tabel distribusi frekuensi

No.	X	Y	XY	X ²	Y ²
1	18	21	378	324	441
2	16	18	288	256	324
3	20	23	460	400	529
4	18	21	378	324	441
5	14	16	224	196	256
6	15	20	300	225	400
7	16	21	336	256	441
8	18	17	306	324	289
9	17	19	323	289	361
10	15	17	255	225	289
Σ	167	193	3248	2819	3771

Contoh Kasus



d. Langkah 5 : mencari nilai b dan a
Mencari nilai b ;

$$b = \frac{n(\sum XY) - (\sum X) \cdot (\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

$$b = \frac{10(3248) - (167) \cdot (193)}{10(2819) - (167)^2}$$

$$b = \frac{32480 - 32231}{28190 - 27889}$$

$$b = \frac{249}{301}$$

$$b = 0,8272425 / 0,827$$

Mencari nilai a ;

$$a = \frac{(\sum Y) - b(\sum X)}{n}$$

$$a = \frac{(193) - 0,8272425(167)}{10}$$

$$a = \frac{54,8505025}{10}$$

$$a = 5,48505025 / 5,485$$

Contoh Kasus

e. Langkah 6 : Menentukan persamaan regresi

$$\bar{Y} = a + bX$$

$$\bar{Y} = 5,485 + 0,827X$$

f. Langkah 7 : Menguji persamaan regresi dengan menghitung nilai r (koefisien korelasi).

$$r = \frac{N \sum XY - \sum X \sum Y}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}}$$

$$r = \frac{10 \times 3248 - 167 \times 193}{\sqrt{(10 \times 2819 - 167^2)(10 \times 3771 - 193^2)}}$$

$$r = \frac{32480 - 32231}{\sqrt{(28190 - 27889)(37710 - 37249)}}$$
$$\frac{249}{\sqrt{(301)(461)}}$$

$$r = \frac{249}{\sqrt{138761}}$$

$$r = \frac{249}{372,5063758}$$
$$r = 0,6684449 / 0,668$$

Contoh Kasus

g. Langkah 8 : Menghitung koefisien determinasi (r^2).

$$r^2 = 0,668^2$$

$$r^2 = 0,447$$

h. Langkah 9 : Uji signifikansi menggunakan Uji-t dan menentukan Taraf Signifikan

Uji Nilai Signifikansi (F hitung);

$$F = \frac{R^2 (n - m - 1)}{m(1 - R^2)}$$

$$F = \frac{0,6684449^2 (10 - 1 - 1)}{1(1 - 0,6684449^2)}$$

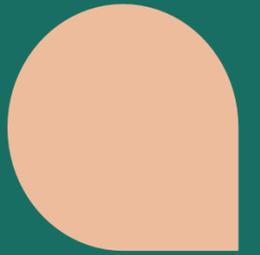
$$F = \frac{0,4468186 \times 8}{1(1 - 0,4468186)}$$

$$F = \frac{3,5745490}{0,5531814}$$

$$F = 6,4618030 / 6,462$$

(ket: F = uji nilai signifikansi, n = jumlah data, m = jumlah variabel bebas)

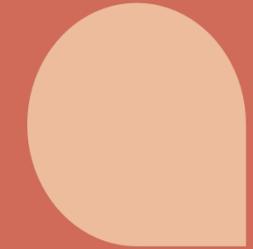
Contoh Kasus



- Menentukan Ftabel pada tabel, yaitu 5,32.
- Menentukan Hipotesis
 - Ha = Terdapat pengaruh yang signifikan
 - H0 = Tidak terdapat pengaruh yang signifikan
- Menentukan tingkat signifikansi (α)
 - Tingkat signifikansi yang sering digunakan adalah $\alpha = 5\%$ ($\alpha = 0,05$)
- Menghitung Derajat kebebasan
 - $db = n - m - 1$
 - $db = 10 - 1 - 1$
 - $db = 8$
 - (ket: db = banyaknya variabel bebas yang diikutseratakan, m = jumlah variable bebas)

Dikonsultasikan dengan Tabel Nilai F 0,05. Pada taraf signifikansi 5%,
Fhitung = 6,462 dan Ftabel = 5,32.
Jadi Fhitung > Ftabel, sehingga Ha diterima dan H0 ditolak.

Contoh Kasus



- i. Langkah 9 : Kesimpulan Terdapat pengaruh yang signifikan antara minat mahasiswa terhadap matakuliah data mining.
2. Perhitungan menggunakan SPSS (program computer yang dipakai untuk analisis statistika)
 - a. Langkah 1 : Masukkan definisi variabel pada variable view dan data ke data view
 - b. Langkah 2 : Klik menu analyze → regression → linier
 - c. Langkah 3 : Masukkan variabel minat ke dalam kotak independent dan variabel matakuliah data mining ke dalam kotak dependent → ok
 - d. Hasil :

Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.668 ^a	.447	.378	1.78542

a. Predictors: (Constant), X_Minat

Tabel ini menggambarkan derajat keeratan hubungan antar variabel.

Penjelasan

Penjelasan :

1) Dalam Sarwono (2006), kekuatan hubungan antara dua variabel memberikan kriteria sebagai berikut.

- a) 0 : Tidak ada korelasi antara dua variabel
- b) 0 – 0,25 : Korelasi sangat lemah
- c) 0,25 – 0,5 : Korelasi cukup
- d) 0,5 – 0,75 : Korelasi kuat
- e) 0,75 – 0,99 : Korelasi sangat kuat
- f) 1 : Korelasi sempurna

Angka R yang diperoleh dari perhitungan SPSS sebesar 0,668 menunjukkan bahwa hubungan antara minat mahasiswa dengan matakuliah data mining adalah kuat, karena besarnya $R > 0,5$.

2. R Square atau Koefisien Determinasi (KD) sebesar 0,447 menunjukkan besarnya peran variabel minat terhadap variabel matakuliah sebesar 44,7%, sedangkan 55,3% dipengaruhi oleh faktor-faktor lain.

3. Std. Error of the Estimate sebesar 1,78542 menggambarkan tingkat ketepatan prediksi regresi. Semakin kecil angkanya, maka prediksinya semakin baik.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	20.598	1	20.598	6.462	.035 ^a
	Residual	25.502	8	3.188		
	Total	46.100	9			

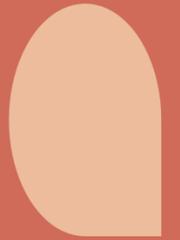
a. Predictors: (Constant), X_Minat

b. Dependent Variable: Y_Matakuliah

Tabel di atas menggambarkan tingkat signifikansi.

Penjelasan :

Nilai uji F atau nilai signifikansi (Sig.) sebesar $0,035 < \text{kriteria signifikansi } 0,05$, maka model persamaan regresi berdasarkan data penelitian adalah signifikan, karena model regresi linier memenuhi kriteria linieritas.



Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	5.485	5.464		1.004	.345
X_Minat	.827	.325	.668	2.542	.035

a. Dependent Variable: Y_Matakuliah

Tabel di atas menggambarkan besaran koefisien regresi.

Penjelasan :

Model persamaan regresi yang diperoleh dari dengan koefisien konstanta dan koefisien variabel, diperoleh model persamaan regresi : $\bar{Y} = 5,485 + 0,827 X_Minat$.



Terimakasih

Any Question?

UJI-T & Aplikasinya





01

Pengertian & Fungsi Uji-T

Pengertian Uji- T

T - test atau uji t adalah uji statistik yang digunakan untuk menguji kebenaran atau kepalsuan hipotesis nol. Uji t pertama kali dikembangkan oleh William Seely Gosset pada tahun 1915. Uji t dapat dibagi menjadi 2 , yaitu uji t yang digunakan untuk pengujian hipotesis 1 sampel dan uji t yang digunakan untuk pengujian hipotesis 2 sampel. Bila dihubungkan dengan kebebasan (independency) sampel yang digunakan (khusus bagi uji t dengan 2 sampel), maka uji t dibagi lagi menjadi 2, yaitu uji t untuk sampel bebas (independent) dan uji t untuk sampel berpasangan (paired).

Fungsi Uji – T

Fungsi nya adalah menguji hipotesis yang membandingkan dua rata-rata apakah hipotesis nol tersebut terjadi secara kebetulan. Alat analisis untuk menguji satu atau dua populasi. Dan mempunyai Syarat yaitu Distribusi data normal dan Varians pada kedua kelompok sama.

02

Pembagian Uji- T



Uji- T

01

**Uji-T Sample Berpasangan
(Dependent / Paired)**

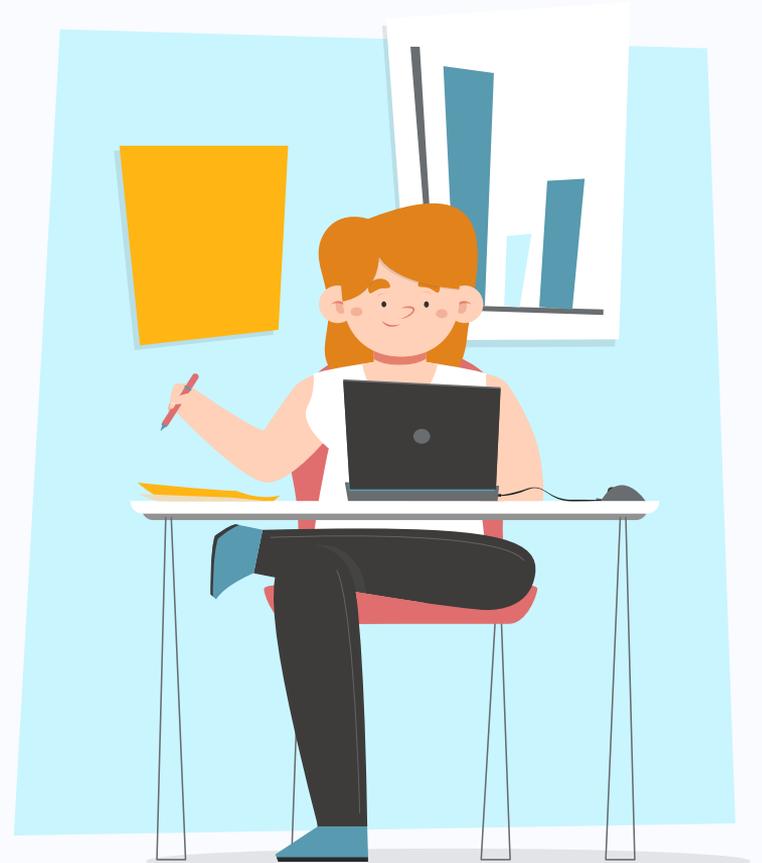


02

**Uji-T Sample Bebas
(Independent)**

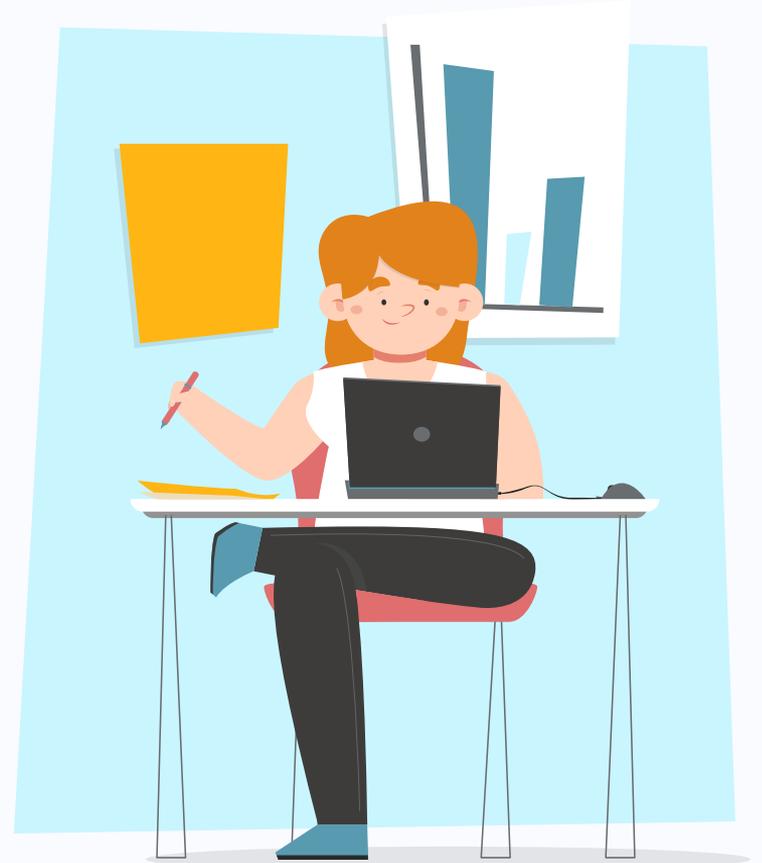
Uji-T Sample Berpasangan

1. Pengertian Uji T sample Berpasangan
2. Fungsi Uji T sample Berpasangan
3. Syarat-Syarat Penggunaan Sample Berpasangan
4. Jenis Hipotesis Sample Berpasangan
5. Rumus Sample Berpasangan
6. Langkah Menggunakan



Uji-T Sample Bebas

1. Pengertian Sample Bebas
2. Langkah-Langkah Menentukan Sample Bebas



Pengertian Uji T Sample Berpasangan

T-test dependent atau sering diistilakan dengan Paired Sampel t-Test, adalah jenis uji statistika yang bertujuan untuk membandingkan rata-rata dua grup yang saling berpasangan. Sampel berpasangan dapat diartikan sebagai sebuah sampel dengan subjek yang sama namun mengalami 2 perlakuan atau pengukuran yang berbeda, yaitu pengukuran sebelum dan sesudah dilakukan sebuah treatment

Sample berpasangan dapat berupa :

1. Satu Sample yang diukur dua kali
2. Dua sample berpasangan diukur Bersama

Fungsi Uji T Sample Berpasangan

Fungsi dari t-test dependent adalah untuk

- membandingkan rata-rata dua grup yang saling berpasangan.
- Sampel berpasangan dapat diartikan sebagai sebuah sampel dengan subjek yang sama namun mengalami 2 perlakuan atau pengukuran yang berbeda, yaitu pengukuran sebelum dan sesudah dilakukan sebuah perlakuan.
- Selain itu untuk menguji efektifitas suatu perlakuan terhadap suatu besaran variabel yang ingin ditentukan, misalnya untuk mengetahui efektifitas metode penyuluhan terhadap peningkatan pengetahuan dari responden

Syarat Penggunaan Uji T Sample Berpasangan

Syarat – syarat penggunaan uji t – test dependent, terdiri dari :

- a. Uji komparasi antar dua nilai pengamatan berpasangan, misalnya:
sebelum dan sesudah
- b. Digunakan pada uji parametrik dimana syaratnya sebagai berikut:
 - satu sampel
 - merupakan data kuantitatif
 - Data berdistribusi normal

Jenis Hipotesis Sample Berpasangan

- Uji dua arah. Pada hipotesis awal tidak terdapat perbedaan yang signifikan antara rata-rata 1 dan rata-rata 2, sedangkan pada hipotesis alternatif sebaliknya yaitu terdapat perbedaan rata-rata 1 dan rata-rata 2.
- Uji satu arah dimana pada hipotesis awal kelompok atau sampel 1 memiliki rata-rata sama dengan atau lebih besar dengan rata-rata kelompok 2. sedangkan hipotesis alternatif rata-rata kelompok 1 lebih kecil dibandingkan dengan rata-rata kelompok 2.
- Uji satu arah ini kebalikan pada hipotesis kedua, dimana pada hipotesis awal kelompok atau sampel 1 memiliki rata-rata sama dengan atau lebih kecil dengan rata-rata kelompok 2. sedangkan hipotesis alternatif rata-rata kelompok 1 lebih besar dibandingkan dengan rata-rata kelompok 2.

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

$$H_0 : \mu_1 \leq \mu_2$$

$$H_1 : \mu_1 > \mu_2$$

Menurut Sugiyono (2010), rumus uji t-test dependent, yaitu : Statistik hitung (t hitung):

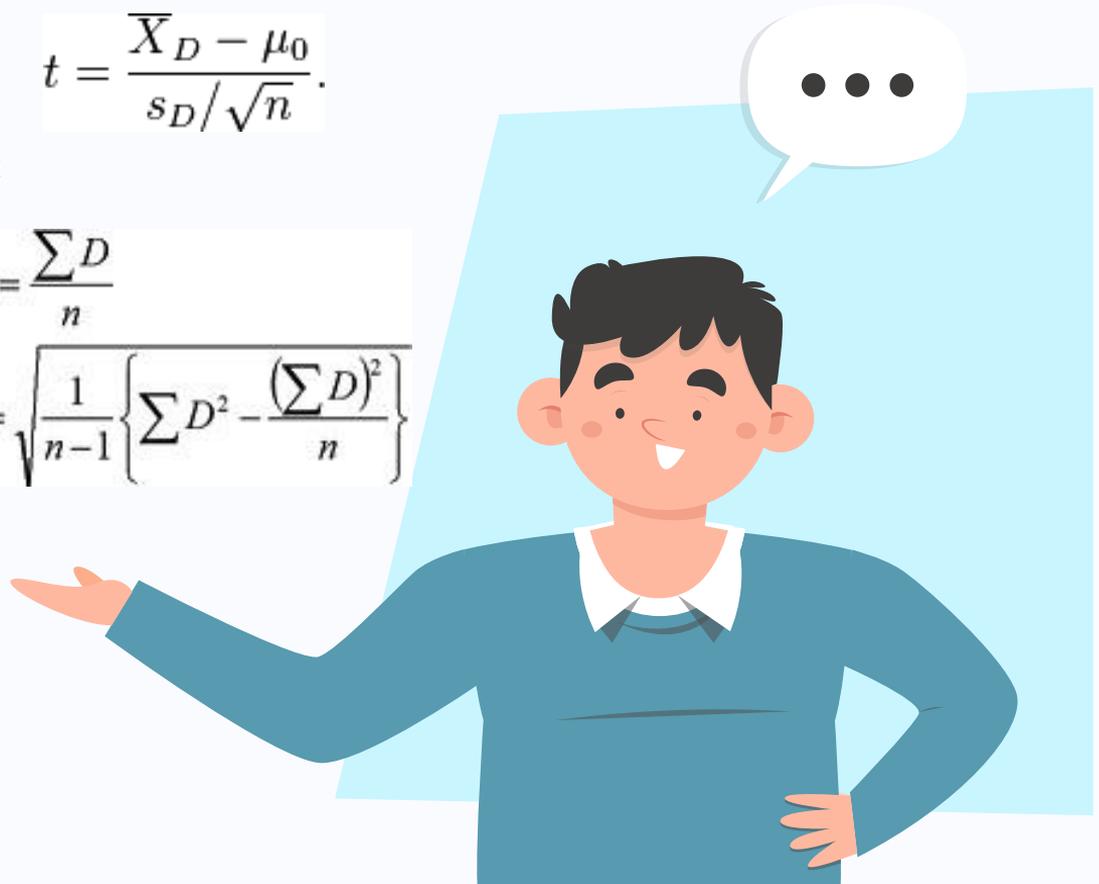
$$t = \frac{\bar{X}_D - \mu_0}{s_D / \sqrt{n}}$$

Dimana

$$\bar{X}_d = \frac{\sum D}{n}$$
$$s_d = \sqrt{\frac{1}{n-1} \left\{ \sum D^2 - \frac{(\sum D)^2}{n} \right\}}$$

Keterangan

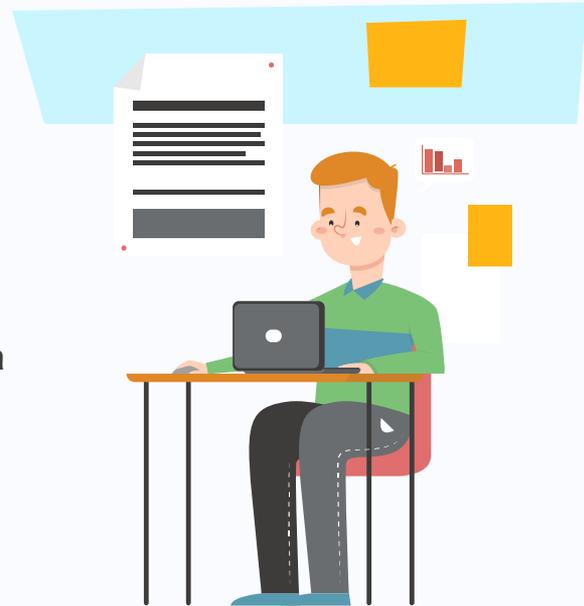
- D** = Selisih x1 dan x2 (x1-x2)
- n** = Jumlah Sampel
- X bar** = Rata-rata
- S d** = Standar Deviasi dari d.



Langkah Menggunakan Uji T – Test Dependent (Sample Berpasangan)

Menurut Ratih (2014), Langkah-langkah pengujian signifikansi (hipotesis) dalam Pengujian Perbedaan Rata-rata Dua kelompok berpasangan:

- **Tetapkan H0 dan H1**
- **Tetapkan titik kritis (tingkat kepercayaan 95 %) atau (tingkat kepercayaan 99 %) yang terdapat pada tabel “t”.**
 - **Tentukan daerah kritis, dengan $db = n - 1$**
 - **Tentukan t hitung dengan menggunakan rumus.**
- **Lakukan uji signifikansi dengan membandingkan besarnya “t” hitung dengan “t” tabel.**

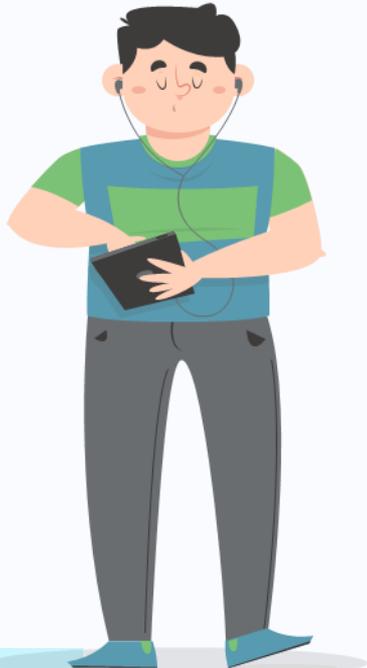


Pengertian Uji T – Dua Sample Bebas

Uji-t dua sampel bebas merupakan uji statistik parametrik yang membandingkan dua kelompok independen untuk menentukan apakah ada bukti bahwa rata-rata populasi secara statistik signifikan berbeda. Variabel yang digunakan dalam uji ini yaitu variabel terikat dan variabel bebas.



Syarat Data Pada Uji T- Dua Sample Bebas



1. **Variabel dependen numerik.**
2. **Variabel independen kategorikal.**
3. **Tidak ada hubungan antara subjek dalam setiap sampel atau kelompok.**
4. **Pengambilan sampel pada populasi secara acak.**
5. **Variabel dependen memiliki distribusi normal pada setiap kelompok**
6. **Varian pada kedua kelompok sama**
7. **Tidak ada outliers**

Langkah Uji T – Test Independent (Dua Sample Bebas)



1. Menentukan H_0 dan H_1
2. Tentukan tingkat signifikan
3. Uji varian
4. Hitung nilai t dan df
5. Bandingkan nilai t hitung dengan t table
6. Pengambilan keputusan hipotesis

1. Menentukan H_0 & H_1

Hipotesis pada Uji-t dua sampel bebas yaitu Hipotesis nol (H_0) dan hipotesis alternatif (H_1), yang dapat dinyatakan dalam dua cara yang berbeda tetapi setara :

$H_0: \mu_1 = \mu_2$ (rata-rata dua kelompok sama)

$H_1: \mu_1 \neq \mu_2$ (rata-rata dua kelompok tidak sama)

Atau

$H_0: \mu_1 - \mu_2 = 0$ (perbedaan rata-rata dua kelompok sama dengan 0)

$H_1: \mu_1 - \mu_2 \neq 0$ (perbedaan rata-rata dua kelompok tidak sama dengan 0)

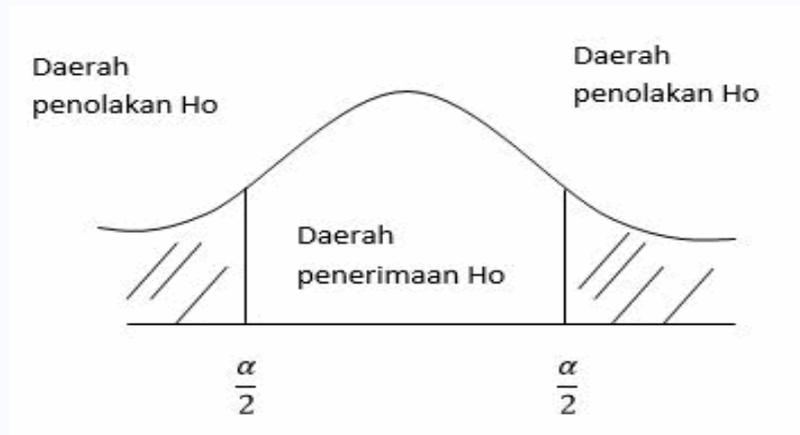
Keterangan:

μ_1 adalah rata-rata populasi kelompok 1 μ_2 adalah rata-rata populasi kelompok 2

2. Menentukan Tingkat Signifikansi

Nilai α adalah peluang untuk membuat kesalahan tipe I. Kesalahan tipe I adalah kesalahan menolak H_0 , padahal H_0 benar. Penentuan tingkat signifikansi ini bervariasi sesuai keinginan peneliti. Nilai α yang umum digunakan adalah 0,05 (5%) dan 0,01 (1%).

Nilai α merupakan batasan dalam menentukan pengambilan keputusan uji hipotesa.



3. Uji Varian

Homogenitas varian menguji keragaman atau varian kedua kelompok untuk menentukan metode uji-t dua sampel bebas yang akan digunakan, yaitu:

1. Asumsi varian sama, nilai $p > \alpha > H_0$ diterima
2. Asumsi varian tidak sama, nilai $p > \alpha > H_0$ ditolak

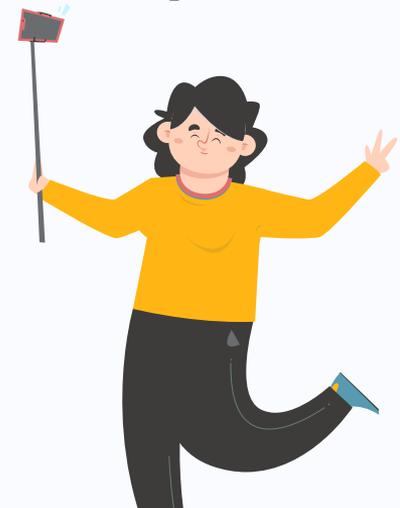
Hipotesis untuk uji Levene adalah:

- $H_0: \sigma_1^2 - \sigma_2^2 = 0$ (varian populasi kelompok 1 dan 2 sama)
- $H_1: \sigma_1^2 - \sigma_2^2 \neq 0$ (varian populasi kelompok 1 dan 2 tidak sama)

Keterangan:

σ_1^2 adalah varian populasi kelompok 1

σ_2^2 adalah varian populasi kelompok 2



4. Menghitung Nilai t dan df

Uji statistik yang digunakan sesuai dengan asumsi varian.
Hipotesis nol dan hipotesis alternatif sama pada kedua uji statistik yang sesuai asumsi varian.

Menghitung Nilai t dan df ini ada dua asumsi yaitu :

1. Asumsi Varian Sama
2. Asumsi Varian tidak sama

Asumsi Varian Sama

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}}$$

Ctt : derajat bebas = $n_1 + n_2 - 2$

Keterangan

X_1 : rata-rata sampel 1

X_2 : rata-rata sampel 2

N_1 : jumlah sampel 1

N_2 : jumlah sampel 2

S_1 : standar deviasi sampel 1

S_2 : standar deviasi sampel 2

S_p : gabungan standar deviasi

Nilai t yang dihitung kemudian dibandingkan dengan nilai t kritis pada tabel distribusi t dengan derajat kebebasan (df) = $n_1 + n_2 - 2$ dan tingkat signifikansi yang dipilih.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$df = \frac{[\left(\frac{S_a^2}{n_a}\right) + \left(\frac{S_b^2}{n_b}\right)]^2}{\left[\left(\frac{S_a^2}{n_a}\right)^2 / (n_a - 1) + \left(\frac{S_b^2}{n_b}\right)^2 / (n_b - 1)\right]}$$

Keterangan

S_a : rata-rata sampel 1

S_b : rata-rata sampel 2

N_a : jumlah sampel 1

N_b : jumlah sampel 2

S_1 : standar deviasi sampel 1

S_2 : standar deviasi sampel 2

Nilai t yang dihitung kemudian dibandingkan dengan nilai t kritis pada tabel distribusi t dengan derajat kebebasan (df) dan tingkat signifikansi yang dipilih.

Asumsi Varian Tidak Sama



5. Pengambilan Keputusan

Dasar pengambilan keputusan uji-t dua sampel bebas untuk mengukur ada tidaknya perbedaan rata-rata dua kelompok yang diuji berdasarkan :

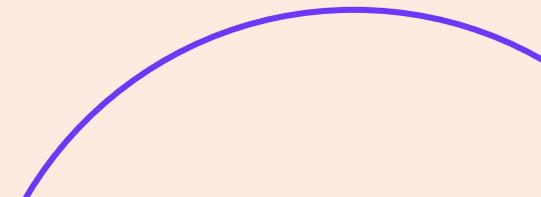
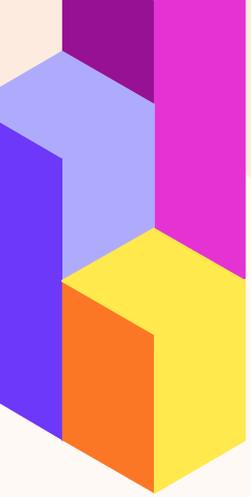
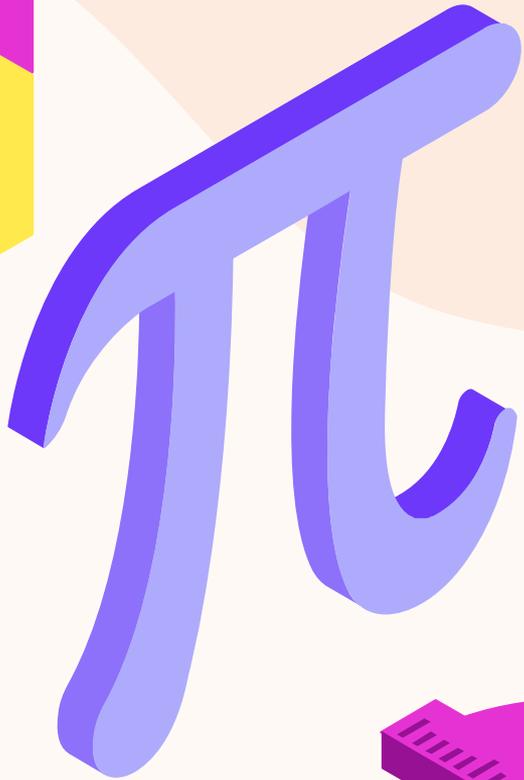
Membandingkan t hitung dengan t tabel.

- Nilai t hitung $>$ nilai t tabel ,Maka H_0 ditolak.
- Nilai t hitung $<$ nilai t tabel ,Maka H_0 diterima.



Terimakasih

Uji T dan Aplikasinya (Lanjutan)



Pengertian uji t

Tes t atau uji t adalah uji statistik yang digunakan untuk menguji kebenaran atau kepalsuan hipotesis nol (H_0).

Uji t (t-test) merupakan statistik uji yang sering kali ditemui dalam masalah-masalah praktis statistika. Uji t merupakan dalam golongan statistika parametrik. Statistik uji ini digunakan dalam pengujian hipotesis, uji t digunakan ketika informasi mengenai nilai variance (ragam) populasi tidak diketahui.

Uji t adalah salah satu uji yang digunakan untuk mengetahui ada atau tidaknya perbedaan yang signifikan (menyakinkan) dari dua mean sampel (dua buah variabel yang dikomparasikan).

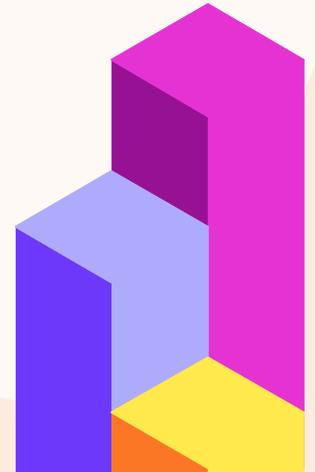
Hipotesis adalah dugaan sementara dalam penelitian dimana hipotesis dibedakan menjadi dua yaitu H_0 atau hipotesis nol dan H_1 atau biasa disebut hipotesis kerja.

H_0 adalah hipotesis yang menyatakan tidak adanya hubungan antara variabel independen (X) dan variabel dependen (Y). Artinya, dalam rumusan hipotesis, yang diuji adalah ketidakbenaran variabel (X) mempengaruhi (Y).



—Klasifikasi T - Satu Sampel (*one sampel t-test*)

Uji t satu sampel digunakan untuk menguji nilai rata-rata dari suatu sampel tunggal dengan suatu nilai acuan. Dalam uji t satu sampel terdapat asumsi yang harus dipenuhi sebelum masuk ke analisis, yaitu data sampel berdistribusi normal. Uji t satu sampel (*one sample t-test*) digunakan untuk menguji apakah suatu nilai berbeda secara nyata atau tidak dengan rata-rata sebuah sampel. Uji T sebagai teknik pengujian hipotesis deskriptif memiliki tiga kriteria yaitu uji pihak kanan, pihak kiri, dan dua pihak.



Uji Pihak Kanan (1-tailed)

Dikatakan sebagai uji pihak kanan karena t-tabel dibagi dua dan diletakkan dibagian kanan kurva Hipotesis statistiknya:

$$H_0: \mu \leq \mu_0$$

$$H_a: \mu > \mu_0$$

Misalnya :

Ho: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **paling tinggi** 70% dari kriteria yang diharapkan.

Ha: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **lebih** 70% dari kriteria yang diharapkan.

Uji Pihak Kiri (1-tailed)

Dikatakan sebagai uji pihak kiri karena t-tabel dibagi dua dan diletakkan dibagian kiri kurva. Hipotesis statistiknya:

$$H_0: \mu \geq \mu_0$$

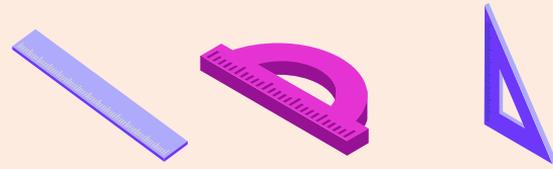
$$H_a: \mu < \mu_0$$

Misalnya :

Ho: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **paling sedikit** 70% dari kriteria yang diharapkan.

Ha: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **kurang** 70% dari kriteria yang diharapkan.

Uji Dua Pihak (2-tailed)



Dikatakan sebagai uji dua pihak karena t-tabel dibagi dua dan diletakkan dibagian kiri dan kanan kurva.
Hipotesis statistiknya:

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

Misalnya :

Ho: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **mencapai** 70% dari kriteria yang diharapkan.

Ha: Tingkat kualitas pelayanan pramuniaga toko di Yogyakarta **tidak mencapai** 70% dari kriteria yang diharapkan.

Rumus statistik yang digunakan dalam Uji T – Satu Sampel, yaitu:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Dengan:

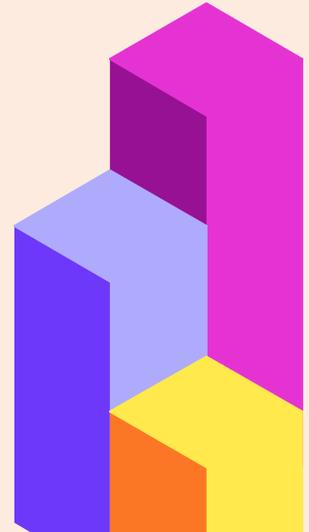
t = t hitung

x = rata-rata dari data

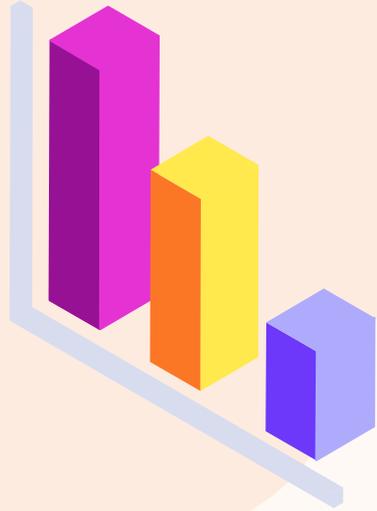
μ_0 = rata-rata nilai yang dihipotesiskan

s = standar deviasi

n = jumlah sampel



Kapan Uji T Sample di Gunakan?



Pengujian rata-rata sampel tunggal digunakan ketika kita ingin tahu apakah sampel kita berasal dari populasi tertentu tetapi kita tidak memiliki informasi populasi yang tersedia bagi kita.

Penggunaan SPSS Pada Uji T – Satu Sampel

SPSS (Statistical Product and Service Solutions) adalah sebuah program aplikasi yang memiliki kemampuan analisis statistik cukup tinggi serta sistem manajemen data pada lingkungan grafis dengan menggunakan menu-menu deskriptif dan kotak-kotak dialog yang sederhana sehingga mudah untuk dipahami cara pengoperasiannya. dapat membaca berbagai jenis data atau memasukkan data secara langsung ke dalam SPSS Data Editor. Bagaimanapun struktur dari file data mentahnya, maka data dalam data editor SPSS harus dibentuk dalam bentuk baris (cases) dan kolom (variables). Case berisi informasi untuk satu unit analisis, sedangkan variable adalah informasi yang dikumpulkan dari masing-masing kasus.

Contoh Penggunaan

Seminggu setelah diadakan Ujian Tengah Semester mata kuliah Statistika, nilai hasil UTS sudah bisa dilihat oleh mahasiswa. Dari hasil nilai tersebut seorang mahasiswa kelas F ingin mengetahui apakah rata-rata nilai kelas F lebih tinggi dari pada nilai rata-rata kelas G. Dari informasi diketahui bahwa nilai rata-rata kelas G yaitu 75. Untuk mengetahuinya, seorang mahasiswa tersebut mengambil sampel acak sebanyak 25 dari nilai teman satu kelasnya. Nilai yang diperoleh adalah sebagai berikut : 85, 73, 62, 75, 77, 79, 73, 80, 70, 69, 80, 90, 60, 73, 70, 75, 75, 79, 65, 80, 70, 65, 78, 78, 77. Dengan tingkat kepercayaan yang digunakan dalam pengujian adalah 95%.

Contoh Penggunaan

Langkah – langkah untuk melakukan uji T satu sampel adalah sebagai berikut:

1. Memasukkan data kedalam data editor dengan mendefinisikan variable pada variable view. Dengan format sebagai berikut:

Nama : Nilai

Decimal : 0

Label : Nilai Kelas F

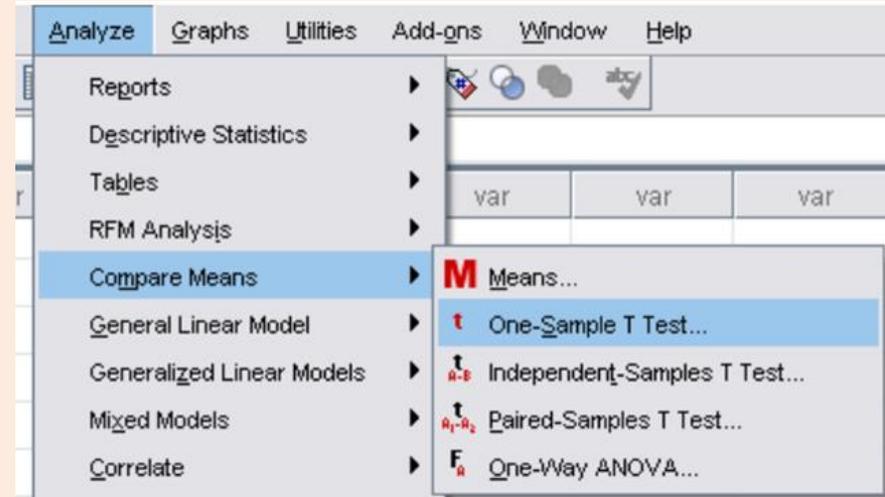
Measure : scale

properti lainnya tidak diatur.

	Nilai
1	85
2	73
3	62
4	75
5	77
6	79
7	73
8	80
9	70
10	69
11	80
12	90
13	60
14	73
15	70
16	75
17	75
18	79
19	65
20	80
21	70
22	65
23	78
24	78
25	77

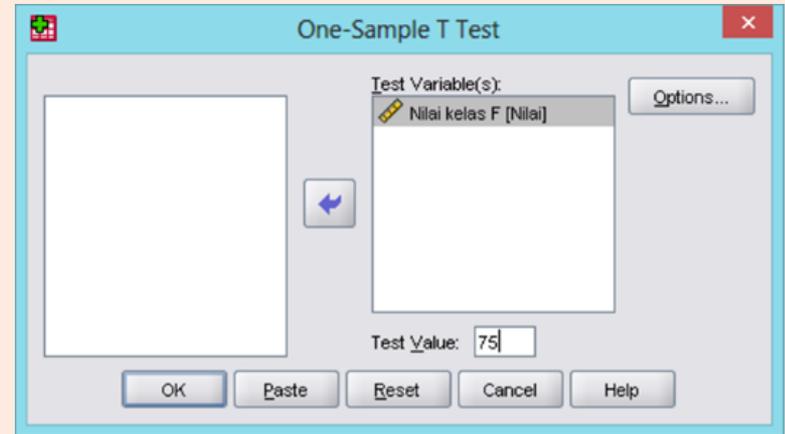
Contoh Penggunaan

2. Pada menu SPSS, memilih menu **Analyze** → **Compare Means** → **One-Sample T Test**.



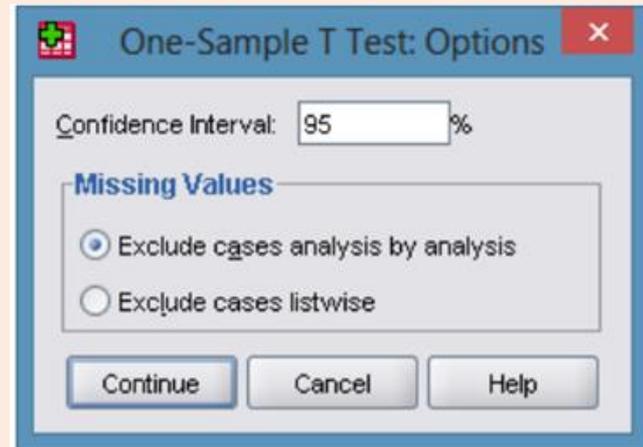
Contoh Penggunaan

3. Akan tampil jendela One-Sample T Test. Kemudian memindahkan variable Nilai ke kotak Test Variable(s) dengan memilih tanda panah. Mengisikan Test Value dengan angka pembanding yaitu 75.



Contoh Penggunaan

4. Mengklik Options pada jendela One-Sample T Test, maka akan tampil jendela Options. Pada pengujian ini tingkat kepercayaan yang digunakan yaitu 95%. Mengisikan 95 pada kolom Confidence Interval. Setelah itu mengklik continue.



Contoh Penggunaan

- Selanjutnya memilih tombol OK pada jendela One-Sample T Test.
- Maka hasil pada jendela Output SPSS akan muncul seperti berikut:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Nilai kelas F	25	74.32	6.963	1.393

One-Sample Test

	Test Value = 75					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Nilai kelas F	-.488	24	.630	-.680	-3.55	2.19

Prosedur Pengujian Hipotesis

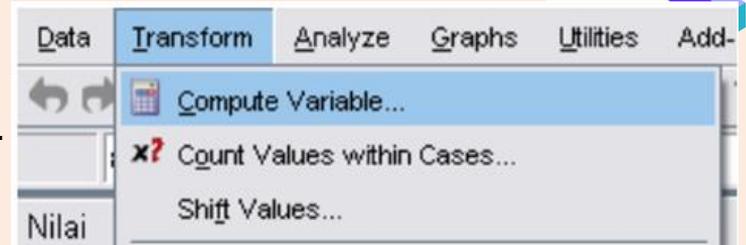
$$H_0 : \mu \leq 75$$

$$H_0 : \mu > 75$$

1. Hipotesis pengujiannya yaitu :
2. Statistik uji: uji t.
3. $\alpha = 0.05$.
4. Statistik tabel : $t(\alpha, n-1) = t(0.05, 24) = 1.711$.

Untuk menentukan statistik tabel menggunakan SPSS yaitu:

- a. Pada menu SPSS memilih **Transform** → **Compute Variable**.
- b. Akan tampil jendela **Compute Variable**.
- c. Memasukkan nama **Target Variable**, yaitu T_TABEL.
- d. Memilih fungsi **Inverse DF** pada **Function Group** dan memilih **Idf.T** pada **Function and Special Variables**. Kemudian meletakkan fungsi **Idf.T** ke **Numeric Expression** dengan melakukan double klik pada **Idf.T**.
- e. Memasukkan nilai $p = 1 - \alpha = 1 - 0.05 = 0.95$ dan nilai $df = n - 1 = 25 - 1 = 24$.
- f. Setelah itu klik **OK**.



g. Hasilnya, pada data editor akan muncul variabel baru dengan nama T_TABEL.

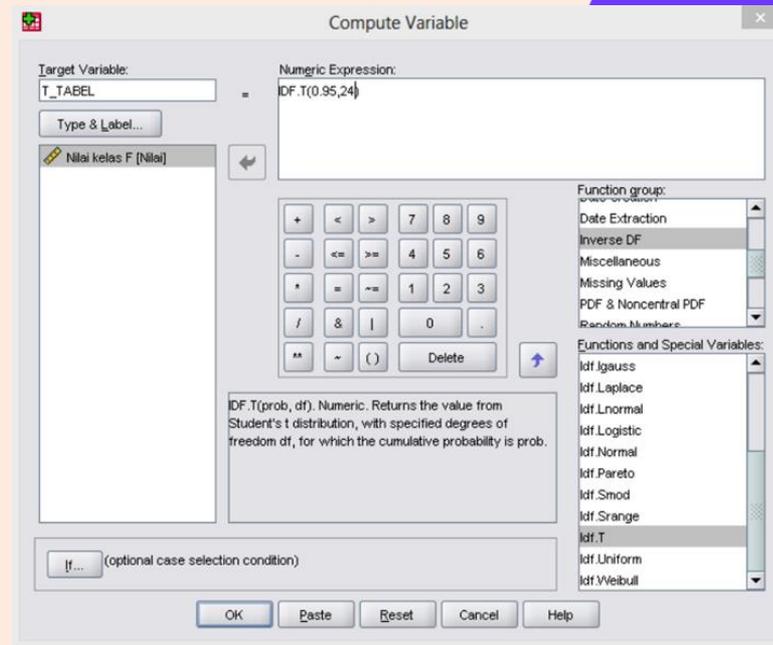
5. Daerah kritis : H_0 ditolak jika **t hitung** > **t tabel**; atau H_0 ditolak jika **Sign.** < α .

6. Dari hasil pengolahan dengan SPSS, diperoleh :

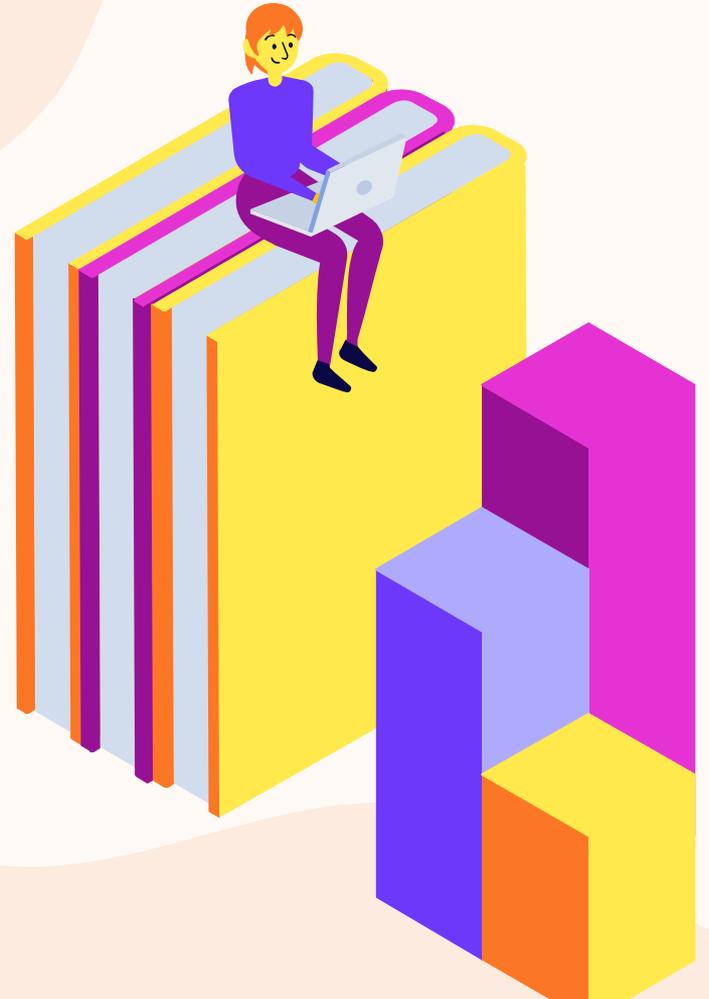
T hitung = - 0.488 dan Sign. = 0.630

7. Karena **t hitung** < **t tabel** ($-0.488 < 1.71$) maka H_0 diterima atau **Sign.** α ($0.630 > 0.05$).

	Nilai	T_TABEL
1	85	1.71
2	73	1.71
3	62	1.71
4	75	1.71
5	77	1.71
6	79	1.71
7	73	1.71
8	80	1.71
9	70	1.71
10	69	1.71
11	80	1.71
12	90	1.71



Thanks



$$2+2=4$$

$$\sqrt[n]{X}$$

-



$$x/2y$$

Anova Satu Jalur

+

x

$$42:9$$

%



a

PEMBAHASAN

01

Pengertian Anova
& Anava

03

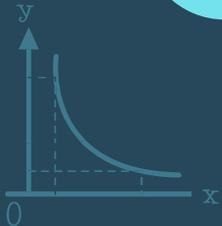
Anova satu arah

02

Perbedaan rata-rata
antar kelompok

04

Contoh dan soal
pembahasan



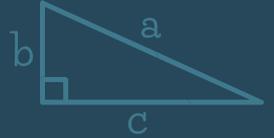
$2+2=4$

$42:9$

 x

01

Pengertian Anova & Anava



+

%

x + x
% y

Anava atau Anova adalah sinonim dari analisis varians terjemahan dari *analysis of variance*, sehingga banyak orang menyebutnya dengan anova. Anova merupakan bagian dari metoda analisis statistika yang tergolong analisis komparatif lebih dari dua rata-rata



Jika kita menguji hipotesis nol bahwa rata-rata dua buah kelompok tidak berbeda, teknik ANAVA dan uji-t (uji dua pihak) akan menghasilkan kesimpulan yang sama; keduanya akan menolak atau menerima hipotesis nol. Dalam hal ini, statistik F pada derajat kebebasan 1 dan n-k akan sama dengan kuadrat dari statistic T. ANAVA digunakan untuk menguji perbedaan antara sejumlah rata-rata populasi dengan cara membandingkan variansinya. Pembilang pada rumus variansi tidak lain adalah jumlah kuadrat skor simpangan dari rata-ratanya, yang secara sederhana dapat ditulis sebagai $\sum(X_i - \mu)^2$. Istilah jumlah kuadrat skor simpangan sering disebut jumlah kuadrat (sum of squares). Jika jumlah kuadrat tersebut dibagi dengan n atau n-1 maka akan diperoleh rata-rata kuadrat yang tidak lain dari variansi suatu distribusi. Rumus untuk menentukan varians sampel yaitu,

$$s^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1}$$



$$2+2=4$$

$$\sqrt[n]{X}$$



$$x/2y$$

Seandainya kita mempunyai suatu populasi yang memiliki variansi σ^2 dan rata-rata μ . Dari populasi tersebut misalkan diambil tiga buah sampel secara independent, masing-masing dengan n_1 , n_2 , dan n_3 . Dari setiap sampel tersebut dapat ditentukan rata-rata dan variansinya, sehingga akan diperoleh tiga buah rata-rata dan variansi sampel yang masing-masing merupakan statistik (penaksir) yang tidak bias bagi parameterunya. Dikatakan demikian karena, dalam jumlah sampel yang tak hingga, rata-rata dari rata-rata sampel akan sama dengan rata-rata populasi (μ) dan rata-rata dari variansi sampel juga akan sama dengan variansi populasi (σ^2).



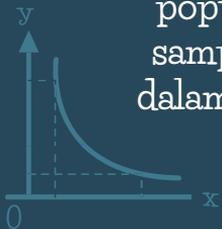
Ada dua hal yang perlu diperhatikan, yaitu

$$2+2=4$$



Kita memiliki 3 buah variansi sampel (S_i^2) yang masing-masing merupakan penaksir yang tidak bias bagi variansi populasinya. Jika $n_1=n_2=n_3=.....=n_k$, maka seluruh variansi sampel tersebut dapat dijumlahkan dan kemudian dibagi dengan banyaknya sampel (k) sehingga akan diperoleh rata-rata variansi sampel yang dalam jangka panjang akan sama dengan variansi populasi. Dalam bahasa ANAVA, rata-rata variansi sampel ini dikenal dengan rata-rata jumlah kuadrat dalam kelompok (RJKD) atau mean of squares within groups (MS_w).

Kita memiliki 3 buah rata-rata sampel yang dapat digunakan untuk menentukan rata-rata dari rata-rata sampel. Simpangan baku distribusi rata-rata sampel ($S_{\bar{X}}$) atau galat baku rata-rata adalah simpangan baku distribusi skor dibagi dengan akar pangkat dua dari besarnya sampel



$$2+2=4$$

$$s_{\bar{y}} = \frac{s_y}{\sqrt{n}}$$

+

x

Sejalan dengan itu, variansi distribusi rata-rata sampel $s_{\bar{y}}^2$ dapat ditulis sebagai berikut.

$$s_{\bar{y}}^2 = \frac{s^2}{n}$$

$$s_{\bar{y}}^2 = \frac{s^2}{n}$$

$$n\sqrt{x}$$

Demikian, s^2 sebagai penaksir yang tidak bias bagi variansi populasi akan ekuivalen dengan variansi distribusi rata-rata dikalikan dengan besarnya sampel (n) yang secara aljabar dapat ditulis sebagai berikut.

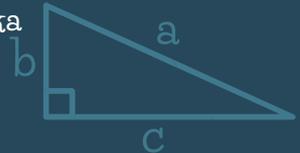
$$\text{Dengan } n s_{\bar{y}}^2 = s^2$$

Dalam konteks ANAVA, $n s_{\bar{y}}^2$ dikenal dengan sebutan rata-rata jumlah kuadrat antar kelompok (RJKA) atau mean of squares between groups (MS_B).

Jika seluruh sampel diambil secara acak dari populasi yang sama, maka

$$MS_B = MS_W \text{ atau } RJKA = RJKD,$$

$$\text{Sehingga, } F = MS_B / MS_W = \frac{\sigma^2}{\sigma^2} = 1$$



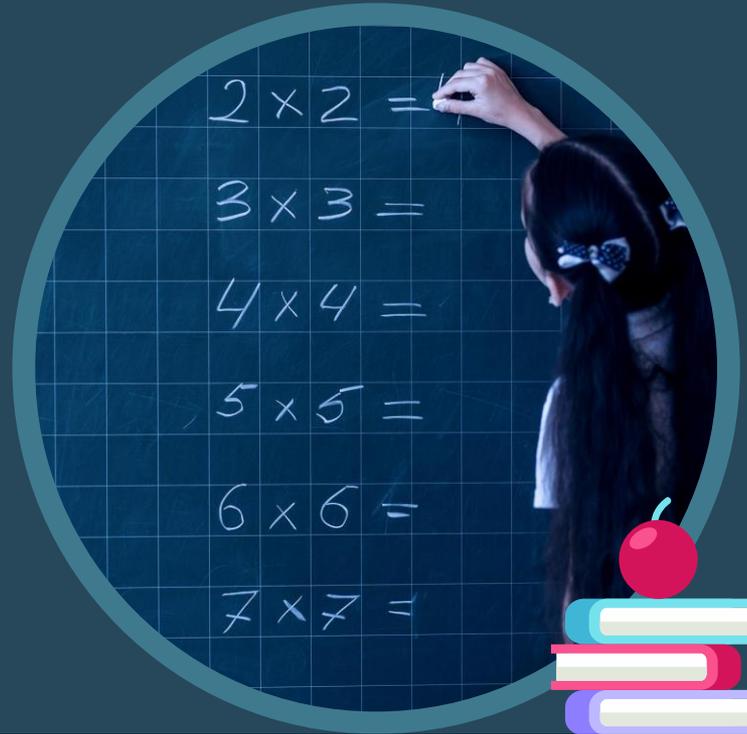
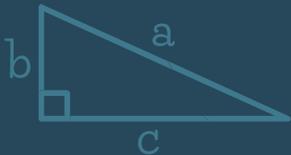
$$x/2y \quad \%$$

ANOVA digunakan untuk menguji hipotesis nol tentang perbedaan dua buah rata-rata atau lebih. Secara formal, hipotesis tersebut dapat ditulis sebagai berikut.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots =$$

μ_k H_1 : Paling tidak salah satu tanda sama dengan (=) tidak berlaku

Hipotesis nol di atas mengatakan bahwa rata-rata populasi pertama sama dengan rata-rata populasi ke dua dan seterusnya yang berarti bahwa seluruh sampel diambil dari populasi yang sama. Jika demikian maka, rata-ratanya akan mirip satu sama lain



+

x

02

$$2+2=4$$

$$\sqrt[n]{x}$$

%



Perbedaan rata-rata antar kelompok

-

3



$$2+2=4$$

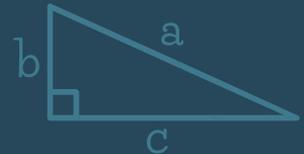
+

x

Keragaman galat dan keragaman yang berkaitan perbedaan pada peubah bebas. Oleh karena keragaman di dalam kelompok (MS_W) merupakan penaksir yang tidak bias atas variansi populasi dan keragaman antara kelompok (MS_B) terdiri atas MS_W dan keragaman yang berkaitan dengan perlakuan, maka hubungan antara keduanya dapat dituliskan sebagai berikut:

$$MS_W = \sigma^2$$

$$n\sqrt{X}$$



$$2+2=4$$

$$MS_B = \sigma^2 + \text{dampak perlakuan}$$

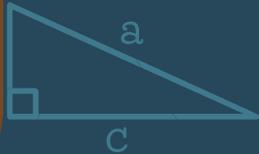
Dengan demikian, F dapat juga dituliskan:

$$F = MS_B / MS_W$$

$$F = (\sigma^2 + \text{dampak perlakuan}) / \sigma^2$$

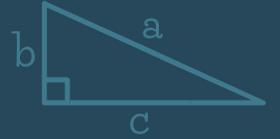
Jika dampak perlakuan sama dengan nol, maka

$$F = \frac{\sigma^2}{\sigma^2} = 1$$



$$2+2=4$$

x



03

$$\sqrt[n]{X}$$

42:9

Anova Satu

$$x/2y$$

Arah

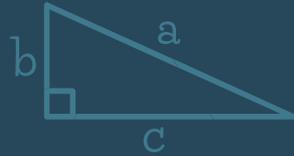


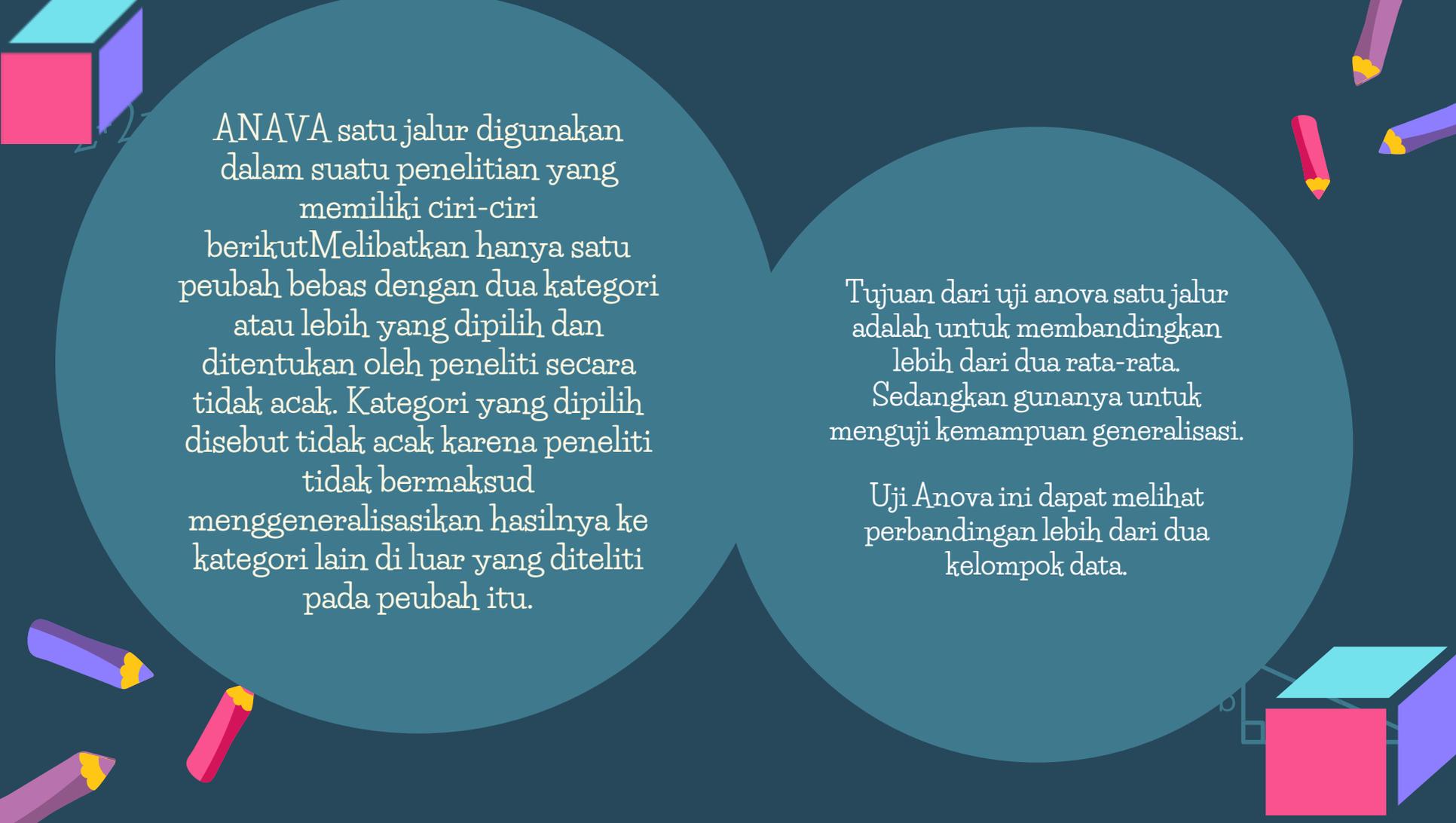
$$2+2=4$$

Dinamakan analisis varians satu arah, karena analisisnya menggunakan varians dan data hasil pengamatan merupakan pengaruh satu faktor. Dari tiap populasi secara independen kita ambil sebuah sampel acak, berukuran n_1 dari populasi kesatu, n_2 dari populasi kedua dan seterusnya berukuran n_k dari populasi ke k . Data sampel akan dinyatakan dengan Y_{ij} yang berarti data ke- j dalam sampel yang diambil dari populasi ke- i . (

Sudjana.1996. Metoda Statistika. Bandung: Tarsito Bandung).

ANAVA satu jalur yaitu analisis yang melibatkan hanya satu peubah bebas





ANAVA satu jalur digunakan dalam suatu penelitian yang memiliki ciri-ciri berikut

- Melibatkan hanya satu peubah bebas dengan dua kategori atau lebih yang dipilih dan ditentukan oleh peneliti secara tidak acak. Kategori yang dipilih disebut tidak acak karena peneliti tidak bermaksud menggeneralisasikan hasilnya ke kategori lain di luar yang diteliti pada peubah itu.

Tujuan dari uji anova satu jalur adalah untuk membandingkan lebih dari dua rata-rata. Sedangkan gunanya untuk menguji kemampuan generalisasi.

Uji Anova ini dapat melihat perbandingan lebih dari dua kelompok data.

Anova lebih dikenal dengan uji-F (Fisher Test), sedangkan arti variasi atau varian itu asalnya dari pengertian konsep ^{aa} Mean Square ^{vv} atau kuadrat rerata (KR).



$$2+2=4$$

$$\sqrt[n]{X}$$

-

$$x/2y$$

+

x

42:9

$$2+2=4$$



Rumusnya :

$$KR = \frac{JK}{db}$$

Dimana: JK = jumlah kuadrat (*some of square*)
 db = derajat bebas (*degree of freedom*)

Menghitung nilai Anova atau F (F_{hitung}) dengan rumus :

$$F_{hitung} = \frac{V_A}{V_D} = \frac{KR_A}{KR_D} = \frac{JK_A : db_A}{JK_D : db_D} = \frac{\text{varian antar group}}{\text{varian antar group}}$$



Varian dalam group dapat juga disebut Varian Kesalahan (Varian Galat). Dapat dirumuskan :

$$JK_A = \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_{\tau})^2}{N} \text{ untuk } db_A = A - 1$$

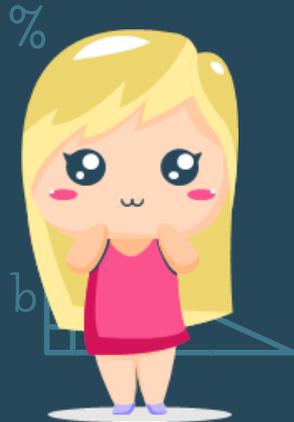
$$JK_D = (\sum X_{\tau})^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} \text{ untuk } db_D = N - A$$

Dimana

$$\frac{(\sum X_{\tau})^2}{N} = \text{sebagai faktor koreksi}$$

N = Jumlah keseluruhan sampel (jumlah kasus dalam penelitian).

A = Jumlah keseluruhan group sampel.



Langkah-langkah Anova Satu Arah

Prosedur Uji Anova Satu Arah

1. Sebelum anova dihitung, asumsikan bahwa data dipilih secara random, berdistribusi normal, dan variannya homogen.
2. Buatlah hipotesis (H_a dan H_0) dalam bentuk kalimat.
3. Buatlah (H_a dan H_0) dalam bentuk statistik.
4. Buatlah daftar statistik induk.
5. Hitunglah jumlah kuadrat antar group (JK_A) dengan rumus :

$$JK_A = \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_\tau)^2}{N} = \left(\frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right) - \frac{(\sum X_\tau)^2}{N}$$

6. Hitunglah derajat bebas antar group dengan rumus : $db_A = A - 1$
7. Hitunglah kudrat rerata antar group (KR_A) dengan rumus : $KR_A = \frac{JK_A}{db_A}$
8. Hitunglah jumlah kuadrat dalam antar group (JK_D) dengan rumus :

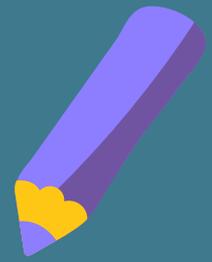
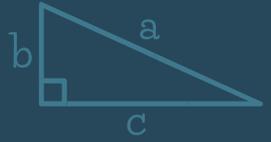
$$JK_D = (\sum X_\tau)^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} \\ = \sum X^2_{A1} + \sum X^2_{A2} + \sum X^2_{A3} - \left(\frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right)$$

9. Hitunglah derajat bebas dalam group dengan rumus : $db_D = N - A$
10. Hitunglah kuadrat rerata dalam antar group (KR_D) dengan rumus : $KR_D = \frac{JK_D}{db_D}$
11. Carilah F_{hitung} dengan rumus : $F_{hitung} = \frac{KR_A}{KR_D}$
12. Tentukan taraf signifikansinya, misalnya $\alpha = 0,05$ atau $\alpha = 0,01$
13. Cari F_{tabel} dengan rumus : $F_{tabel} = F_{(1-\alpha)(db_A, db_D)}$
14. Buat Tabel Ringkasan Anova.



$2+2=4$

x



04

$\sqrt[n]{X}$

Contoh Soal & Pembahasan

42:9

$x/2y$



1. Seorang ingin mengetahui perbedaan prestasi belajar untuk mata kuliah dasar-dasar statistika antara mahasiswa tugas belajar, izin belajar dan umum. Data diambil dari nilai UTS sebagai berikut :

Tugas belajar (A_1) = 6,8,5,7,7,6,6,8,7,6,7 = 11 orang
Izin belajar (A_2) = 5,6,6,7,5,5,5,6,5,6,8,7 = 12 orang
Umum (A_3) = 6,9,8,7,8,9,6,6,9,8,6,8 = 12 orang

Buktikan apakah ada perbedaan atau tidak?

LANGKAH-LANGKAH MENJAWAB :

1. Diasumsikan bahwa data dipilih secara random, berdistribusi normal, dan variannya homogen.

2. Hipotesis (H_a dan H_0) dalam bentuk kalimat.

H_a = Terdapat perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

H_0 = Tidak ada perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

3. Hipotesis (H_a dan H_0) dalam bentuk statistik

$H_a : A_1 \neq A_2 = A_3$ $H_0 : A_1 = A_2 = A_3$

4. Daftar statistik induk

NILAI UTS			
NO	A_1	A_2	A_3
1	6	5	6
2			
3	8	6	9
4	5	6	8
5			
6	7	7	7
7			
8	7	5	8
9			
10	6	5	9
11	6	5	6
12			
	8	6	6
	7	5	9
	6	6	8
	7	8	6
	-	7	8



1. Menghitung jumlah kuadrat antar group (JK_A) dengan rumus :

$$JK_A = \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_{\tau})^2}{N}$$
$$= \left(\frac{(73)^2}{11} + \frac{(71)^2}{11} + \frac{(90)^2}{12} \right) - \frac{(234)^2}{35} = 1579,53 - 1564,46 = 15,07$$

2. Hitunglah derajat bebas antar group dengan rumus :

$$db_A = A - 1 = 3 - 1 = 2 \quad A = \text{jumlah group } A$$

3. Hitunglah kudrat rerata antar group (KR_A) dengan rumus :

$$KR_A = \frac{JK_A}{db_A} = \frac{15,07}{2} = 7,54$$

4. Hitunglah jumlah kuadrat dalam antar group (JK_D) dengan rumus :

$$JK_D = (\sum X_{\tau})^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} = (493 + 431 + 692) - \left(\frac{(73)^2}{11} + \frac{(71)^2}{11} + \frac{(90)^2}{12} \right)$$
$$= 1616 - 1579,53 = 36,47$$

5. Hitunglah derajat bebas dalam group dengan rumus :

$$db_D = N - A = 35 - 3 = 32$$

6. Hitunglah kudrat rerata dalam antar group (KR_D) dengan rumus :

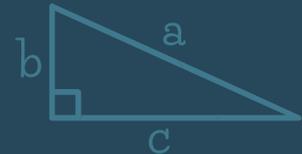
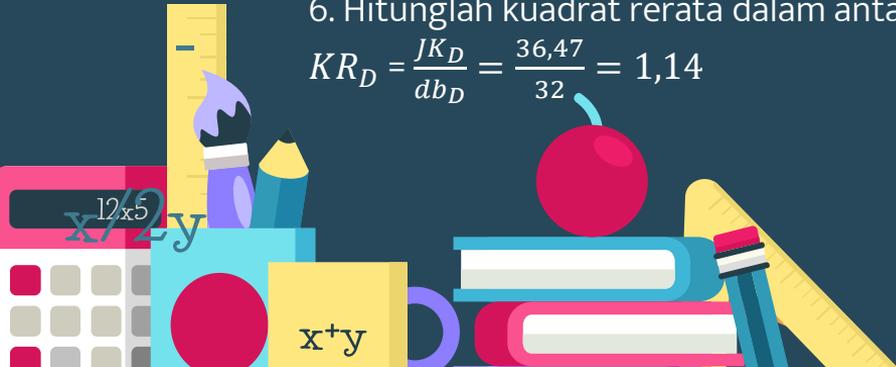
$$KR_D = \frac{JK_D}{db_D} = \frac{36,47}{32} = 1,14$$

+

x



$n\sqrt{x}$



7. Carilah F_{hitung} dengan rumus :

$$F_{hitung} = \frac{KR_A}{KR_D} = \frac{7,54}{1,14} = 6,61$$

8. Tentukan taraf signifikansinya, misalnya $\alpha = 0,05$

9. Cari F_{tabel} dengan rumus :

$$F_{tabel} = F_{(1-\alpha)(db_A, db_D)}$$

$$F_{tabel} = F_{(1-0,05)(2,32)}$$

$$F_{tabel} = F_{(0,95)(2,32)}$$

$$F_{tabel} = 3,30$$

Cara mencari : Nilai $F_{tabel} = 3,30$ dan arti angka $F_{tabel} = F_{(0,95)(2,32)}$
0,95 = Taraf kepercayaan 95% atau taraf signifikan 5%.

Angka 2 = pembilang atau hasil dari db_A

Angka 32 = penyebut atau hasil dari db_D

Apabila angka 2 dicari ke kanan dan angka 32 ke bawah maka akan bertemu dengan nilai $F_{tabel} = 3,30$. Untuk taraf signifikansi 5% dipilih pada bagian atas dan 1% dipilih pada bagian bawah.

10. Buat Tabel Ringkasan Anova

x/2y

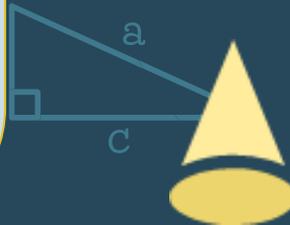


TABEL

RINGKASAN ANOVA SATU JALUR

Sumber	Jumlah Kuadrat	Derajat bebas (db)	Kuadrat Rerata (KR)	F_{hitung}	Taraf Signifikan (ρ)
Antar group (A)	15,07	2	7,54	6,61	< 0,05 $F_{tabel} = 3,30$
Dalam group (D)	36,47	32	1,14	-	-
Total	51,54	54	-	-	-

$$2+2=4$$



x

+

x

%

y

11. Tentukan kriteria pengujian : jika $F_{hitung} \geq F_{tabel}$, maka tolak H_0 berarti signifikan.

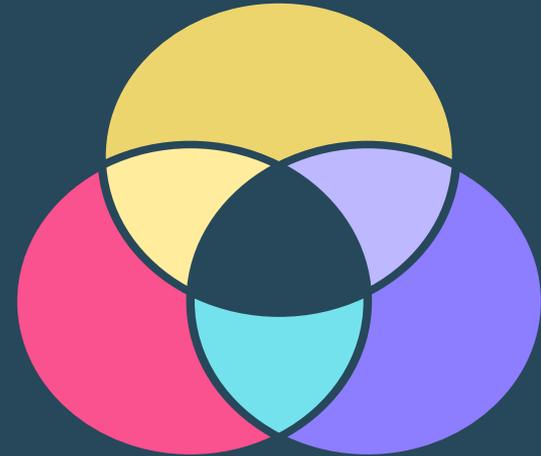
Setelah konsultasikan dengan tabel F kemudian

bandingkan antara F_{hitung} dengan F_{tabel} , ternyata :

$F_{hitung} > F_{tabel}$ atau $6,61 > 3,30$ maka tolak H_0 berarti signifikan.

12. Kesimpulan

H_0 ditolak dan H_a diterima. Jadi, terdapat perbedaan yang signifikan antara mahasiswa tugas belajar, izin belajar dan umum.

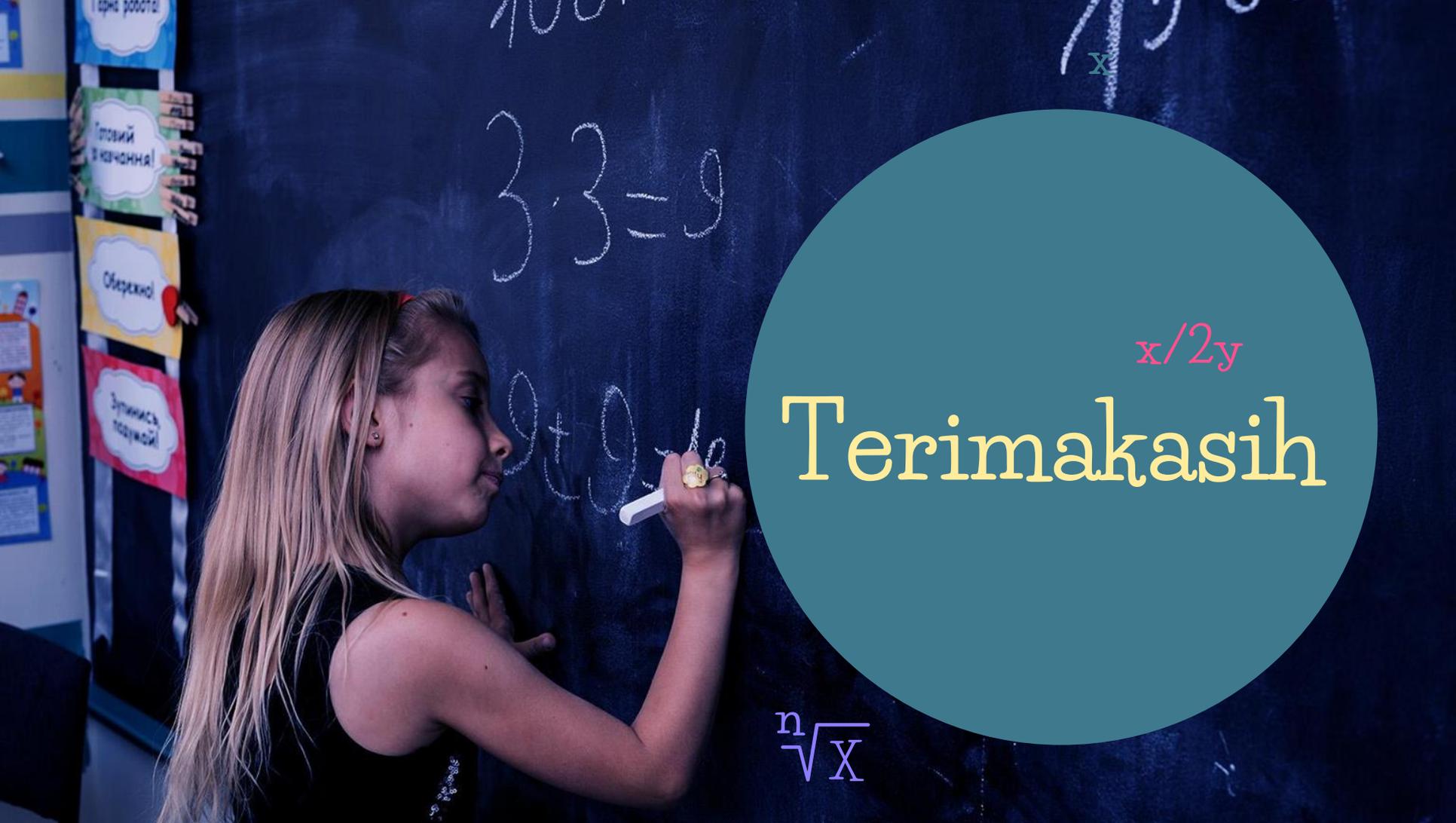


3



2





$x/2y$

Terimakasih

$\sqrt[n]{x}$