

Pengantar Statistik

Nanang Erma Gunawan
nanang_eg@uny.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.

Statistik sebagai ilmu penunjang,
disebut STATISTIKA

1. Konsep Statistika

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)

METODE ILMIAH :

Adalah cara mencari kebenaran dengan resiko keliru paling kecil.

Berikut adalah LANGKAH-LANGKAH METODE ILMIAH :

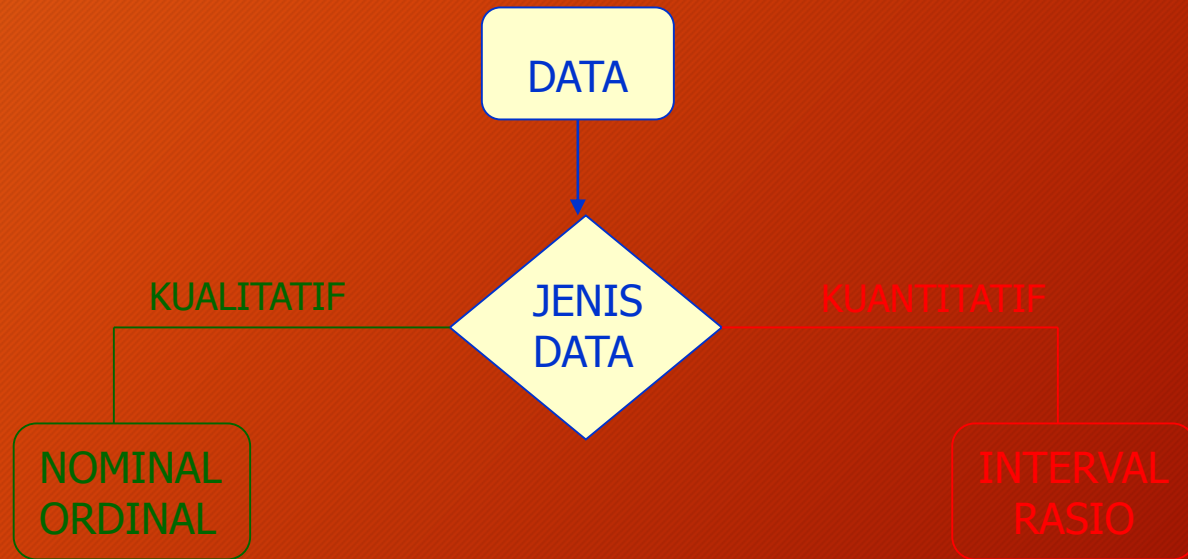
1. Merumuskan masalah
2. Melakukan studi literature atau kajian pustaka
3. Membuat dugaan-dugaan, pertanyaan-pertanyaan atau hipotesis
- 4. Peran Statistika: Mengumpulkan dan mengolah data, menguji hipotesis, atau menjawab pertanyaan**
5. Mengambil kesimpulan



DATA KUALITATIF dan DATA KUANTITATIF

DATA KUALITATIF :
Data yang dinyatakan dalam bentuk **bukan angka**.
Contoh : jenis pekerjaan, status marital, tingkat kepuasan kerja

DATA KUANTITATIF :
Data yang dinyatakan dalam bentuk **angka**
Contoh : lama bekerja, jumlah gaji, usia, hasil ulangan



DATA NOMINAL :

Data berskala nominal adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi.

CIRI : posisi data setara
tidak bisa dilakukan operasi matematika

(+, -, x, :)

CONTOH : jenis kelamin, jenis pekerjaan

DATA ORDINAL :

Data berskala ordinal adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi, tetapi di antara data tersebut terdapat hubungan

CIRI : posisi data tidak setara
tidak bisa dilakukan operasi matematika

(+, -, x, :)

CONTOH : kepuasan kerja, motivasi

DATA INTERVAL :

Data berskala interval adalah data yang diperoleh dengan cara pengukuran, di mana jarak antara dua titik skala sudah diketahui.

CIRI : Tidak ada kategorisasi
bisa dilakukan operasi matematika

CONTOH : temperatur yang diukur berdasarkan $^{\circ}\text{C}$ dan $^{\circ}\text{F}$, sistem kalender

DATA RASIO :

Data berskala rasio adalah data yang diperoleh dengan cara pengukuran, di mana jarak antara dua titik skala sudah diketahui dan mempunyai titik 0 absolut.

CIRI : tidak ada kategorisasi
bisa dilakukan operasi matematika

CONTOH : gaji, skor ujian, jumlah buku

PENGOLAHAN DATA :

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.

TABEL

Tabel 1.1 Bidang Pekerjaan berdasarkan Latar Belakang Pendidikan

Count

| | | pendidikan | | | Jumlah |
|------------------|--------------|------------|---------|---------|--------|
| | | SMU | Akademi | Sarjana | |
| bidang pekerjaan | administrasi | 1 | 8 | 6 | 15 |
| | personalia | | 1 | 7 | 8 |
| | produksi | 4 | 3 | 5 | 12 |
| | marketing | 2 | 14 | 11 | 27 |
| | keuangan | 3 | 4 | 6 | 13 |
| Jumlah | | 10 | 30 | 35 | 75 |

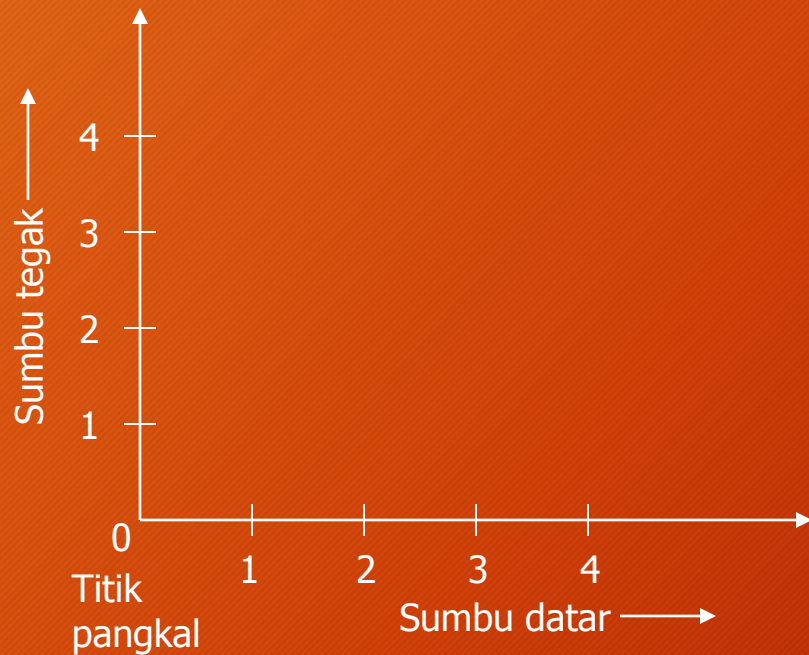
GRAFIK



GRAFIK : memberikan informasi dengan benar dan cepat, tetapi tidak rinci.

Syarat :

1. Pemilihan sumbu (sumbu tegak dan sumbu datar), kecuali grafik lingkaran
2. Penetapan skala (skala biasa, skala logaritma, skala lain)
3. Ukuran grafik (tidak terlalu besar, tinggi, pendek)

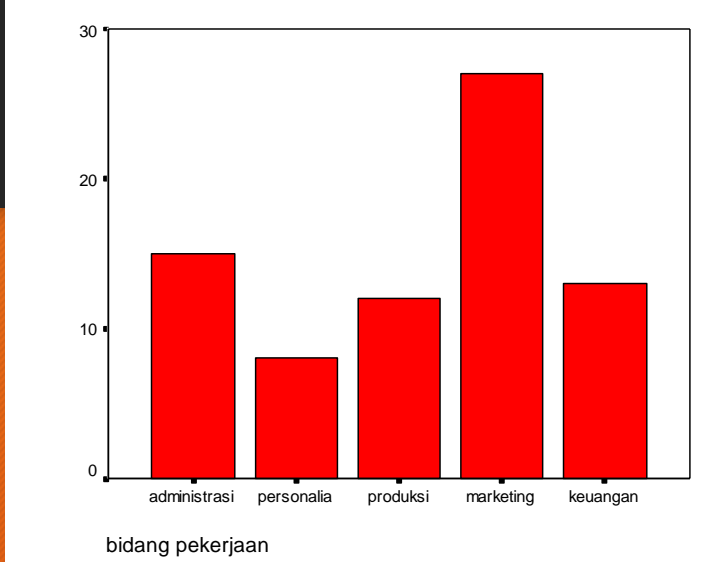


Jenis Grafik :

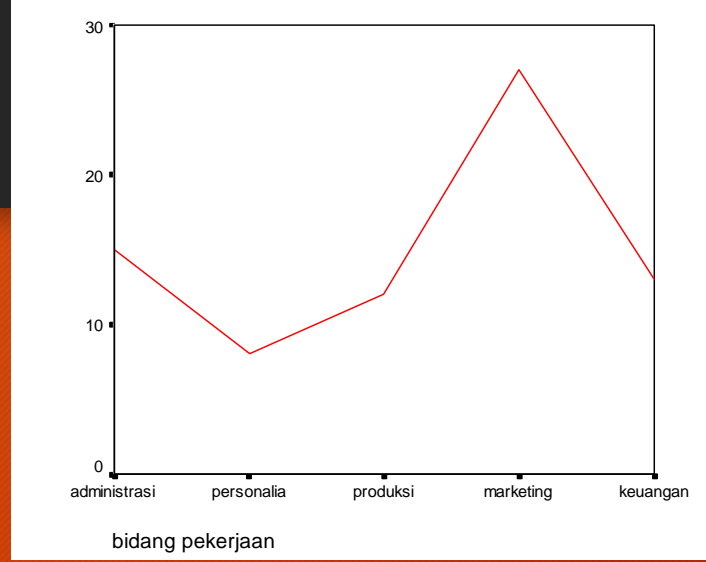
- **Grafik Batang (Bar)**
- **Grafik Garis (line)**
- **Grafik Lingkaran (Pie)**
- **Grafik Interaksi (Interactive)**

10. Jenis Grafik

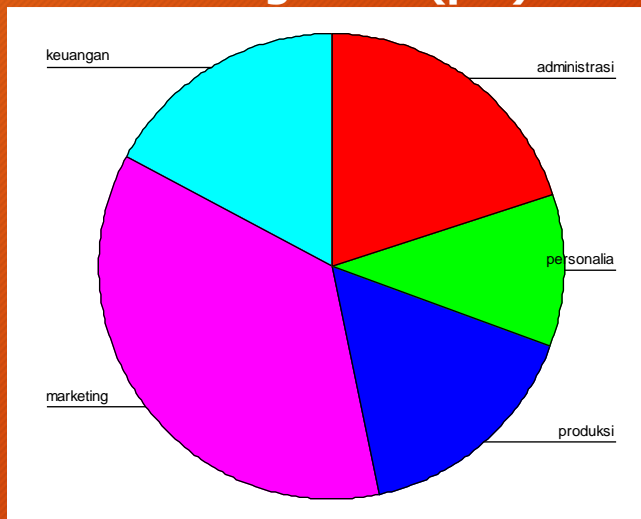
Grafik Batang (Bar)



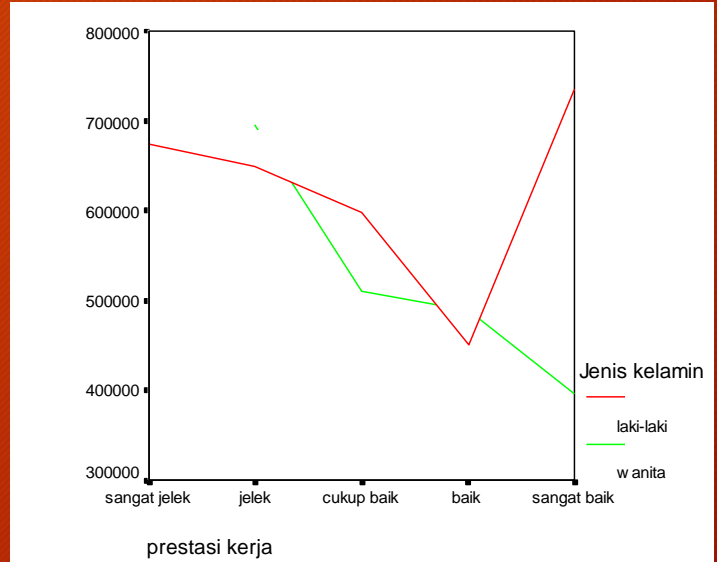
Grafik Garis (line)



Grafik lingkaran (pie)



Grafik Interaksi (interactive)



FREKUENSI : banyaknya data untuk satu kelompok/klasifikasi

| KELOMPOK | FREKUENSI |
|---------------|------------------------|
| Kelompok ke-1 | f1 |
| Kelompok ke-2 | f2 |
| Kelompok ke-3 | f3 |
| Kelompok ke-i | f _i |
| Kelompok ke-k | f _k |
| | $n = \sum_{i=1}^k f_i$ |

| Pendidikan | Frekuensi |
|------------|-----------|
| S1 | 34 |
| S2 | 29 |
| S3 | 37 |
| | 100 |



$$n = \sum_{i=1}^k f_i = f_1 + f_2 + f_3 + \dots + f_i + \dots + f_k$$

12. Distribusi Frekuensi

DISTRIBUSI FREKUENSI : mengelompokkan data interval/rasio dan menghitung banyaknya data dalam satu kelompok/klasifikasi

| USIA | FREKUENSI |
|------|-----------|
| 20 | 5 |
| 21 | 6 |
| 22 | 13 |
| 23 | 4 |
| 24 | 7 |
| 25 | 7 |
| 26 | 7 |
| 27 | 5 |
| 28 | 3 |
| 29 | 4 |
| 30 | 15 |
| 31 | 3 |
| 33 | 5 |
| 35 | 1 |

Membuat distribusi frekuensi :

1. Mencari sebaran (range) yakni selisih antara data paling besar dengan data paling kecil) + 1 $\rightarrow 35 - 20 + 1 = 16$
2. Menentukan banyak kelas dengan rumus $k = 1 + 3,3 \log n$
 $\rightarrow 7$
1. Menentukan panjang kelas dengan rumus
 $p = \text{sebaran} / \text{banyak kelas} \rightarrow 16/7 = 2$

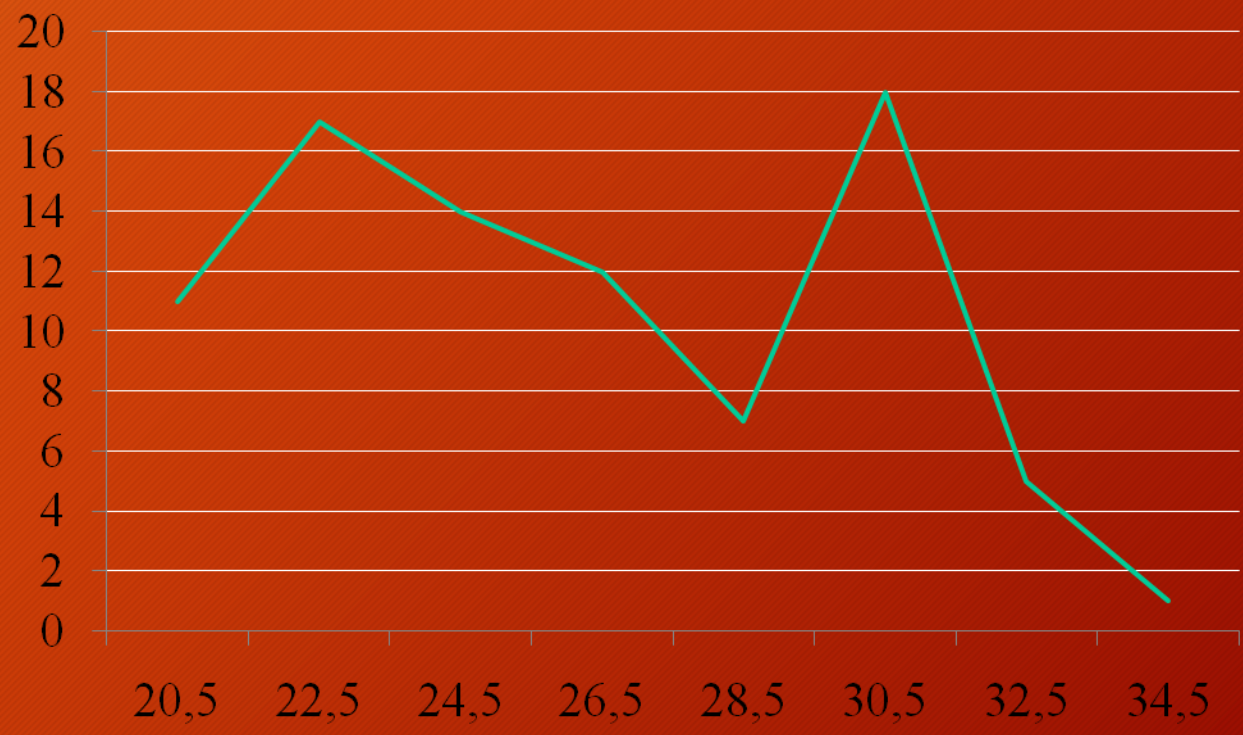


| KELOMPOK USIA | FREKUENSI |
|---------------|-----------|
| 20 – 21 | 11 |
| 22 – 23 | 17 |
| 24 – 25 | 14 |
| 26 – 27 | 12 |
| 28 – 29 | 7 |
| 30 – 31 | 18 |
| 32 - 33 | 5 |
| 34 - 35 | 1 |

13. Grafik Poligon

| KELOMPOK USIA | FREKUENSI | NILAI TENGAH |
|---------------|-----------|--------------|
| 20-21 | 11 | 20,5 |
| 22-23 | 17 | 22,5 |
| 24-25 | 14 | 24,5 |
| 26-27 | 12 | 26,5 |
| 28-29 | 7 | 28,5 |
| 30-31 | 18 | 30,5 |
| 32-33 | 5 | 32,5 |
| 34-35 | 1 | 34,5 |

Grafik poligon



14. Grafik Histogram

| KELOMPOK USIA | FREKUENSI | NILAI NYATA |
|---------------|-----------|-------------|
| 20-21 | 11 | 19,5-21,5 |
| 22-23 | 17 | 21,5-23,5 |
| 24-25 | 14 | 23,5-25,5 |
| 26-27 | 12 | 25,5-27,5 |
| 28-29 | 7 | 27,5-29,5 |
| 30-31 | 18 | 29,5-31,5 |
| 32-33 | 5 | 31,5-33,5 |
| 34-35 | 1 | 33,5-35,5 |

13. Ukuran Tendensi Sentral

a. Mean

RATA-RATA : suatu bilangan yang bertindak mewakili sekumpulan bilangan
RATA-RATA HITUNG (RERATA/mean) : jumlah bilangan dibagi banyaknya

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$$

$$\frac{\sum_{i=1}^n X_i}{n}$$

Bila terdapat sekumpulan bilangan di mana masing-masing bilangannya memiliki frekuensi, maka rata-rata hitung menjadi :

$$\bar{X} = \frac{X_1 f_1 + X_2 f_2 + X_3 f_3 + \dots + X_k f_k}{f_1 + f_2 + f_3 + \dots + f_k}$$

$$\frac{\sum_{i=1}^k X_i f_i}{\sum_{i=1}^k f_i}$$

Cara menghitung :

| Bilangan (X_i) | Frekuensi (f_i) | $X_i f_i$ |
|--------------------|---------------------|------------|
| 70 | 3 | 210 |
| 63 | 5 | 315 |
| 85 | 2 | 170 |
| Jumlah | 10 | 695 |

Maka :

$$\bar{X} = \frac{695}{10} = 69.5$$

b. Median

MEDIAN : nilai tengah dari sekumpulan data setelah diurutkan yang fungsinya membantu memperjelas kedudukan suatu data.

Contoh : diketahui rata-rata hitung/mean nilai ulangan dari sejumlah siswa adalah 6.55. Pertanyaannya adalah apakah siswa yang memperoleh nilai 7 termasuk istimewa, baik, atau biasa-biasa saja ?

Jika nilai ulangan tersebut adalah : 10 10 8 7 7 6 5 5 5 5 4,
maka rata-rata hitung = 6.55, median = 6
Kesimpulan : nilai 7 termasuk kategori baik sebab berada di atas rata-rata hitung dan median (kelompok 50% atas)

Jika nilai ulangan tersebut adalah : 8 8 8 8 8 7 5 5 4 3,
maka rata-rata hitung = 6.55, median = 8
Kesimpulan : nilai 7 termasuk kategori kurang sebab berada di bawah median (kelompok 50% bawah)

Jika sekumpulan data banyak bilangannya genap (tidak mempunyai bilangan tengah)
Maka mediannya adalah rerata dari dua bilangan yang ditengahnya.
Contoh : 1 2 3 4 5 6 7 8 8 9 maka median $(5+6) : 2 = 5.5$

c. Modus

MODUS : bilangan yang paling banyak muncul dari sekumpulan bilangan, yang fungsinya untuk melihat kecenderungan dari sekumpulan bilangan tersebut.

Contoh : nilai ulangan 10 10 8 7 7 6 5 5 5 5 4

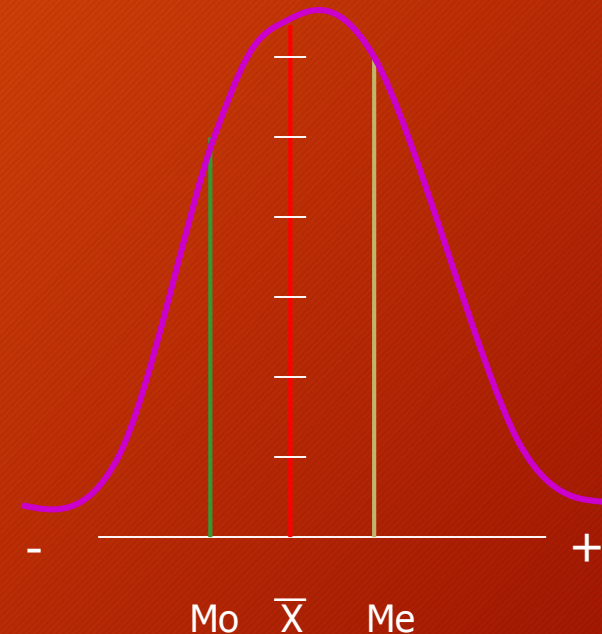
Maka : $s = 6$; $k = 3$; $p = 2$

rata-rata hitung/mean = 6.55 ; median = 6

modus = 5 ; kelas modus = 5 - 7

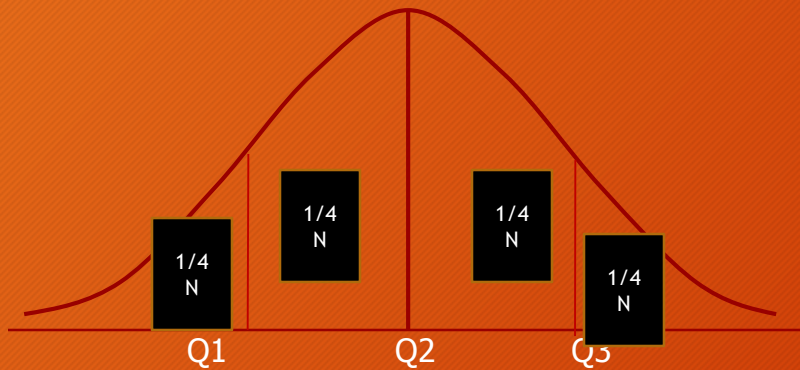
| Nilai | Frekuensi |
|--------|-----------|
| 10 | 2 |
| 8 | 1 |
| 7 | 2 |
| 6 | 1 |
| 5 | 4 |
| 4 | 1 |
| Jumlah | 11 |

| Nilai | Frekuensi |
|--------|-----------|
| 8 - 10 | 3 |
| 5 - 7 | 7 |
| 2 - 4 | 1 |
| Jumlah | 11 |



Kurva **positif** apabila rata-rata hitung $>$ modus / median
Kurva **negatif** apabila rata-rata hitung $<$ modus / median

Quartile: titik/skor/nilai yang membagi seluruh distribusi frekuensi ke dalam empat bagian sama besar, yakni masing-masing $1/4N$.



Untuk data tunggal

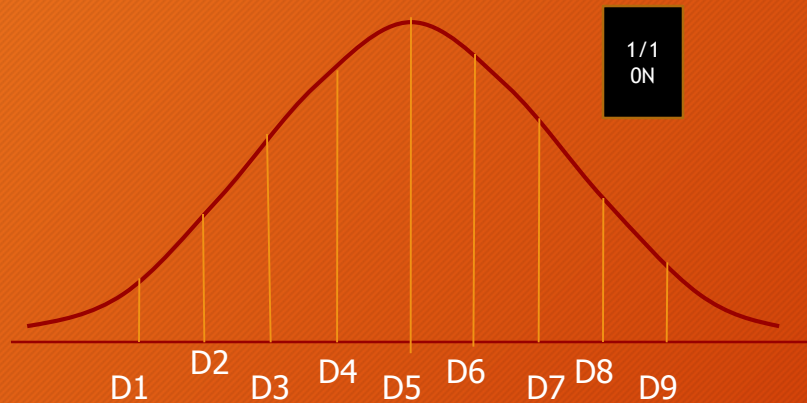
$$Q_n = \ell + \left[\frac{n/4N - f_{kb}}{f_i} \right]$$

Untuk data berkelompok

$$Q_n = \ell + \left[\frac{n/4N - f_{kb}}{f_i} \times i \right]$$

ℓ = batas bawah nyata dari skor yang mengandung Q_n

Desile: titik/skor/nilai yang membagi seluruh distribusi frekuensi ke dalam sepuluh bagian sama besar, yakni masing-masing $1/10N$.



Untuk data tunggal

$$D_n = \ell + \left[\frac{n/10N - f_{kb}}{f_i} \right]$$

Untuk data berkelompok

$$D_n = \ell + \left[\frac{n/10N - f_{kb}}{f_i} \right] \times i$$

ℓ = batas bawah nyata dari skor yang mengandung D_n

Percentile: titik/skor/nilai yang membagi seluruh distribusi frekuensi ke dalam seratus bagian sama besar, yakni masing-masing $1/100N$.



Untuk data tunggal

$$P_n = \ell + \left[\frac{n/100N - f_{kb}}{f_i} \right]$$

Untuk data berkelompok

$$P_n = \ell + \left[\frac{n/100N - f_{kb}}{f_i} \right] \times i$$

ℓ = batas bawah nyata dari skor yang mengandung P_n

14. Ukuran Penyebaran

UKURAN YANG MENYATAKAN HOMOGENITAS / HETEROGENITAS :

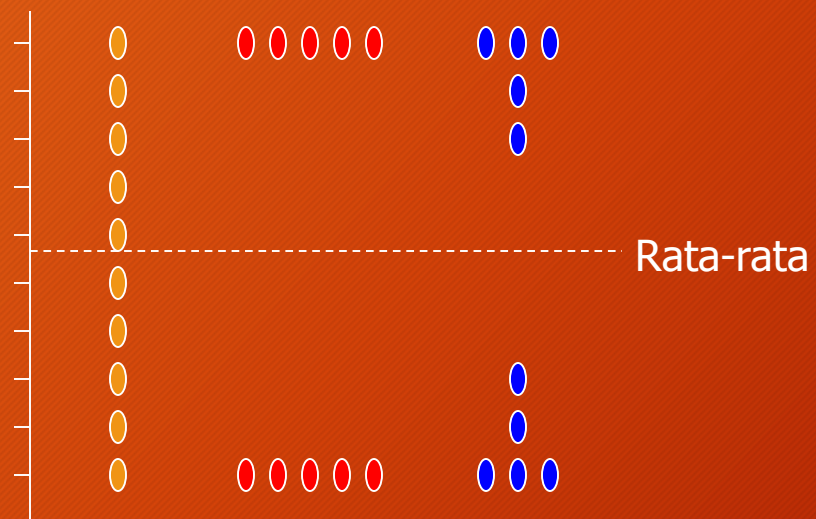
1. RENTANG (*Range*)
2. DEVIASI RATA-RATA (*Average Deviation*)
3. VARIANS (*Variance*)
4. DEVIASI STANDAR (*Standard Deviation*)

Rentang (range) : selisih bilangan terbesar dengan bilangan terkecil.
Sebaran merupakan ukuran penyebaran yang sangat kasar, sebab hanya bersangkutan dengan bilangan terbesar dan terkecil.

Contoh :

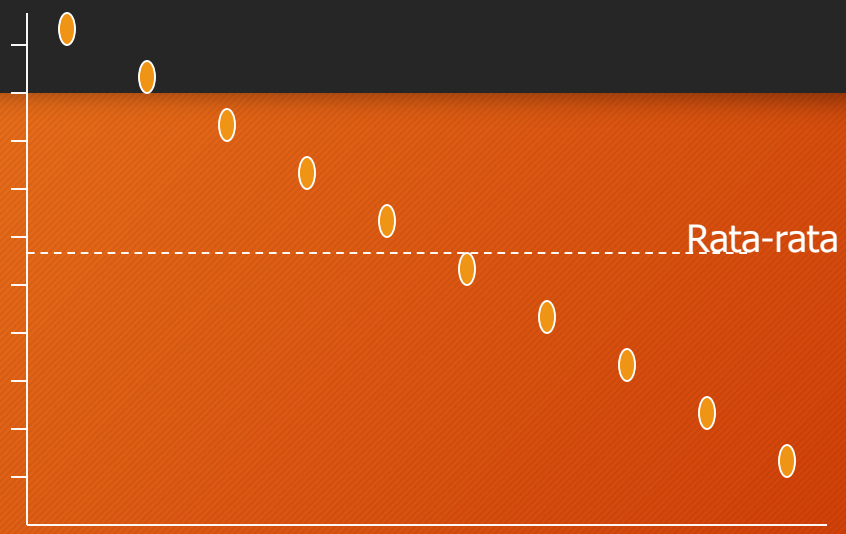
| | | | | | | | | | | | |
|---|---|-----|-----|-----|-----|-----|----|----|----|----|----|
| A | : | 100 | 90 | 80 | 70 | 60 | 50 | 40 | 30 | 20 | 10 |
| B | : | 100 | 100 | 100 | 100 | 100 | 10 | 10 | 10 | 10 | 10 |
| C | : | 100 | 100 | 100 | 90 | 80 | 30 | 20 | 10 | 10 | 10 |

$$\bar{X} = 55$$
$$r = 100 - 10 = 90$$



17. Deviasi rata-rata

Deviasi Rata-rata : penyebaran Berdasarkan harga mutlak simpangan bilangan-bilangan terhadap rata-ratanya.

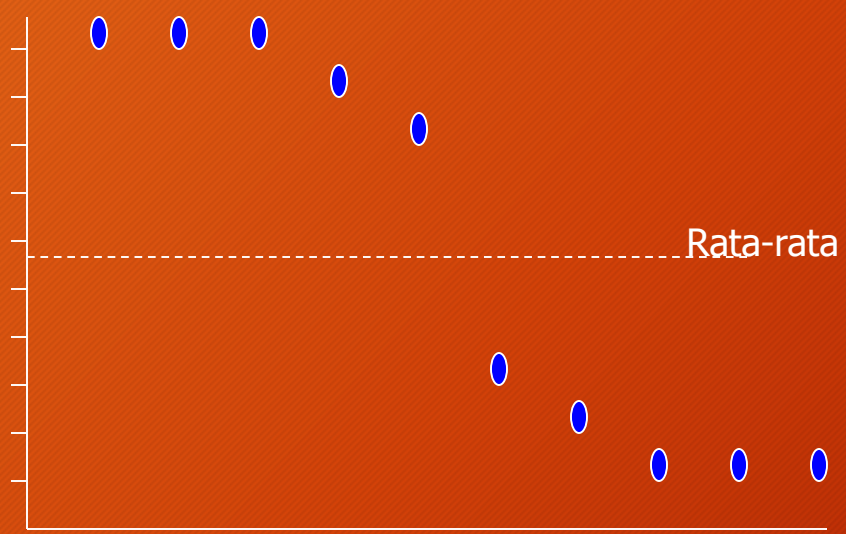


| Kelompok A | | |
|------------|---------------|-----------------|
| Nilai X | $X - \bar{X}$ | $ X - \bar{X} $ |
| 100 | 45 | 45 |
| 90 | 35 | 35 |
| 80 | 25 | 25 |
| 70 | 15 | 15 |
| 60 | 5 | 5 |
| 50 | -5 | 5 |
| 40 | -15 | 15 |
| 30 | -25 | 25 |
| 20 | -35 | 35 |
| 10 | -45 | 45 |
| Jumlah | 0 | 250 |

$$DR = \frac{250}{10} = 25$$

| Kelompok B | | |
|------------|---------------|-----------------|
| Nilai X | $X - \bar{X}$ | $ X - \bar{X} $ |
| 100 | 45 | 45 |
| 100 | 45 | 45 |
| 100 | 45 | 45 |
| 90 | 35 | 35 |
| 80 | 25 | 25 |
| 30 | -25 | 25 |
| 20 | -35 | 35 |
| 10 | -45 | 45 |
| 10 | -45 | 45 |
| 10 | -45 | 45 |
| Jumlah | 0 | 390 |

$$DR = \frac{390}{10} = 39$$



$$DR = \frac{\sum_{i=1}^n |X_i - \bar{X}|}{n}$$

Makin besar simpangan, makin besar nilai deviasi rata-rata

18. Varians & Deviasi Standar

Varians : penyebaran berdasarkan jumlah kuadrat simpangan bilangan-bilangan terhadap rata-ratanya ; melihat ketidaksamaan sekelompok data

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

Deviasi Standar : penyebaran berdasarkan akar dari varians ; menunjukkan keragaman kelompok data

$$s = \sqrt{\sum_{i=1}^n \frac{(X_i - \bar{X})^2}{n-1}}$$

Kelompok A

| Nilai X | X - \bar{X} | (X - \bar{X}) ² |
|---------|---------------|-------------------------------|
| 100 | 45 | 2025 |
| 90 | 35 | 1225 |
| 80 | 25 | 625 |
| 70 | 15 | 225 |
| 60 | 5 | 25 |
| 50 | -5 | 25 |
| 40 | -15 | 225 |
| 30 | -25 | 625 |
| 20 | -35 | 1225 |
| 10 | -45 | 2025 |
| Jumlah | | 8250 |

Kelompok B

| Nilai X | X - \bar{X} | (X - \bar{X}) ² |
|---------|---------------|-------------------------------|
| 100 | 45 | 2025 |
| 100 | 45 | 2025 |
| 100 | 45 | 2025 |
| 90 | 35 | 1225 |
| 80 | 25 | 625 |
| 30 | -25 | 625 |
| 20 | -35 | 1225 |
| 10 | -45 | 2025 |
| 10 | -45 | 2025 |
| 10 | -45 | 2025 |
| Jumlah | | 15850 |

$$s = \sqrt{\frac{8250}{9}} = 30.28 \quad s = \sqrt{\frac{15850}{9}} = 41.97$$

Kesimpulan :

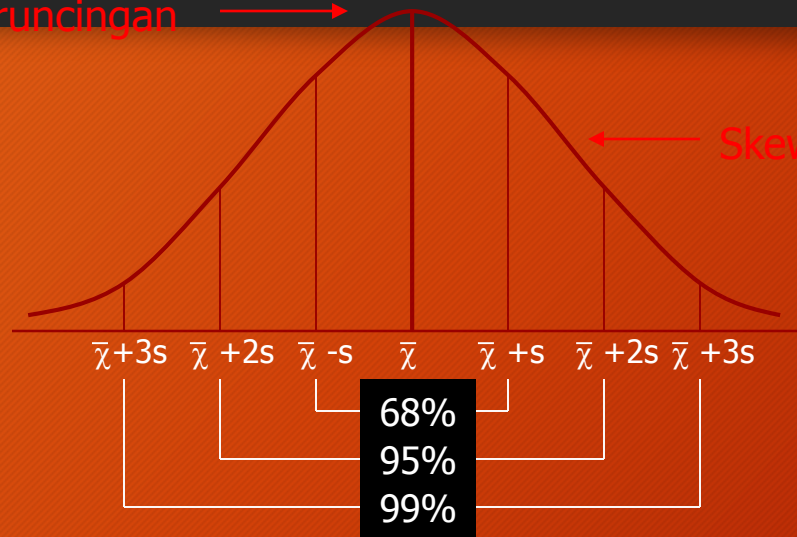
Kelompok A : rata-rata = 55 ; DR = 25 ; s = 30.28

Kelompok B : rata-rata = 55 ; DR = 39 ; s = 41.97

Maka data kelompok B lebih tersebar daripada kelompok A

Distribusi Normal : kurva berbentuk bel, simetris, simetris terhadap sumbu yang melalui nilai rata-rata

Kurtosis = keruncingan



- Lakukan uji normalitas
- Rasio Skewness & Kurtosis berada -2 sampai $+2$
Rasio = $\frac{\text{nilai}}{\text{Standard error}}$
- Jika tidak berdistribusi normal, lakukan uji normalitas non parametrik (Wilcoxon, Mann-White, Tau Kendall)

20. Normalitas, Hipotesis, Pengujian

Hipotesis : uji signifikansi (keberartian) terhadap hipotesis yang dibuat ; berbentuk hipotesis penelitian dan hipotesis statistik (H_0) ; hipotesis bisa terarah, bisa juga tidak terarah ; akibat dari adanya H_0 , maka akan ada H_a (hipotesis alternatif) yakni hipotesis yang akan diterima seandainya H_0 ditolak

| HIPOTESIS | TERARAH | TIDAK TERARAH |
|--------------------------------------|--|--|
| Hipotesis Penelitian | Siswa yang belajar bahasa lebih serius daripada siswa yang belajar IPS | Ada perbedaan keseriusan siswa antara yang belajar bahasa dengan yang belajar IPS |
| Hipotesis Nol (Yang diuji) | Siswa yang belajar bahasa tidak menunjukkan kelebihan keseriusan daripada yang belajar IPS $H_0 : b < i$ $H_a : b > i$ | Tidak terdapat perbedaan keseriusan belajar siswa antara bahasa dan IPS $H_0 : b = i$ $H_a : b \neq I$ |

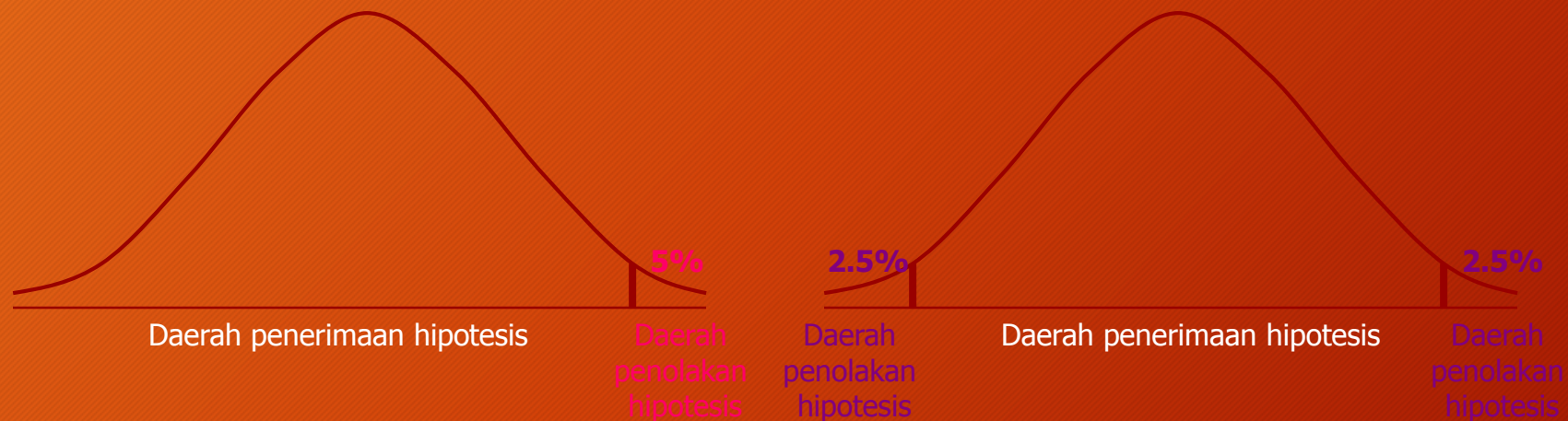
21. Normalitas, Hipotesis, Pengujian

Pengujian : bila H_0 terarah, maka pengujian signifikansi satu pihak
bila H_0 tidak terarah, maka pengujian signifikansi dua pihak

Pengujian signifikansi satu arah (hipotesis terarah):

Siswa yang belajar bahasa tidak menunjukkan kelebihan keseriusan daripada yang belajar IPS □ $H_0 : b < i$

Jika H_0 ditolak, maka H_a diterima ; daerah penolakan berada di sebelah kanan



Pengujian signifikansi dua arah (hipotesis tidak terarah):

Tidak terdapat perbedaan keseriusan belajar siswa antara bahasa dan IPS

□ $H_0 : b = i$

Jika H_0 ditolak, maka H_a diterima ; daerah penolakan bisa berada di sebelah kiri atau kanan

Uji t : menguji apakah rata-rata suatu populasi sama dengan suatu harga tertentu atau apakah rata-rata dua populasi sama/berbeda secara signifikan.

1. Uji t satu sampel

Menguji apakah satu sampel sama/berbeda dengan rata-rata populasinya

- hitung rata-rata dan std. dev (s)
- $df = n - 1$
- tingkat signifikansi ($\alpha = 0.025$ atau 0.05)
- pengujian apakah menggunakan 1 ekor atau 2 ekor
- diperoleh t hitung ; lalu bandingkan dengan t tabel : jika t hitung $>$ t tabel H_0 ditolak

$$t = \frac{(\bar{x} - \mu)}{s / \sqrt{n}}$$

Contoh :

Peneliti ingin mengetahui apakah guru yang bekerja selama 8 tahun memang berbeda dibandingkan dengan guru lainnya.

$H_0 : p_1 = p_2$

Diperoleh rata2 = 17.26 ; std. Dev = 7.6 ; $df = 89$; t hitung = 11.55

Berdasarkan tabel $df=89$ dan $\alpha = 0.05$ diperoleh t tabel = 1.987

Kesimpulan : t hitung $>$ t tabel sehingga H_0 ditolak

guru yang bekerja selama 8 tahun secara signifikan berbeda dengan guru lainnya

2. Uji t dua sampel bebas

Menguji apakah rata-rata dua kelompok yang tidak berhubungan sama/berbeda

$$t = \frac{(\bar{X} - \bar{Y})}{S_{\bar{X}-\bar{Y}}}$$

Di mana

$$S_{\bar{X}-\bar{Y}} = \sqrt{\frac{(\sum x^2 + \sum y^2) (1/n_x + 1/n_y)}{(n_x + n_y - 2)}}$$

Contoh :

Peneliti ingin mengetahui apakah ada perbedaan penghasilan (sebelum sertifikasi) antara guru yang lulusan S1 dengan yang lulusan S3

Ho : $\mu_x = \mu_y$

Diperoleh : rata2 x = 1951613 ; y = 2722222 ; t hitung = - 7.369

Berdasarkan tabel df=69 dan $\alpha = 0.025$ diperoleh t tabel = 1.994

Kesimpulan : t hitung > t tabel sehingga Ho ditolak

Rata-rata penghasilan guru yang S1 berbeda secara signifikan dengan penghasilan guru yang S3

3. Uji t dua sampel berpasangan

Menguji apakah rata-rata dua sampel yang berpasangan sama/berbeda

$$t = \frac{\bar{D}}{s_D}$$

Di mana \bar{D} = rata-rata selisih skor pasangan

$$s_D = \sqrt{\frac{\sum d^2}{N(N-1)}}$$

$$\sum d^2 = \frac{\sum D^2 - (\sum D)^2}{N}$$

Contoh :

Seorang guru ingin mengetahui efektivitas model pembelajaran diskusi. Setelah selesai pembelajaran pertama, ia memberikan tes dan setelah selesai pembelajaran kedua kembali ia memberikan tes. Kedua hasil tes tersebut dibandingkan dengan harapan adanya perbedaan rata-rata tes pertama dengan kedua.

$H_0 : \mu_d = \mu_c$

Diperoleh rata_{2d} = 66.28 ; rata_{2c} = 73.84 ; t hitung = -8.904

Berdasarkan tabel df=163 dan $\alpha = 0.05$ diperoleh t tabel = 1.960

Kesimpulan : t hitung > t tabel sehingga H_0 ditolak

Terdapat perbedaan yang signifikan antara hasil tes pertama dengan hasil tes kedua, sehingga ia menyimpulkan model diskusi efektif meningkatkan hasil belajar siswanya

Korelasi : hubungan keterkaitan antara dua atau lebih variabel.
Angka koefisien korelasi (r) bergerak $-1 \leq r \leq +1$

POSITIF

makin besar nilai variabel 1 menyebabkan **makin besar** pula nilai variabel 2
Contoh : makin banyak waktu belajar, makin tinggi skor Ulangan → **korelasi positif** antara waktu belajar dengan nilai ulangan

NEGATIF

makin besar nilai variabel 1 menyebabkan **makin kecil** nilai variabel 2
contoh : makin banyak waktu bermain, makin kecil skor Ulangan → **korelasi negatif** antara waktu bermain dengan nilai ulangan

NOL

tidak ada atau tidak menentunya hubungan dua variabel
contoh : pandai matematika dan jago olah raga ; pandai matematika dan tidak bisa olah raga ; tidak pandai matematika dan tidak bisa olah raga
→ **korelasi nol** antara matematika dengan olah raga

1. KORELASI PEARSON/PRODUCT MOMENT :

apakah di antara kedua variabel terdapat hubungan, dan jika ada hubungan bagaimana arah hubungan dan berapa besar hubungan tersebut.

Digunakan jika data variabel **kontinyu** dan **kuantitatif**

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

Di mana : ΣXY = jumlah perkalian X dan Y
 ΣX^2 = jumlah kuadrat X
 ΣY^2 = jumlah kuadrat Y
 N = banyak pasangan nilai

Contoh :

10 orang siswa yang memiliki waktu belajar berbeda dites dengan tes IPS

| | | | | | | | | | | | |
|-----------|---|---|---|---|---|---|---|---|---|---|---|
| Siswa | : | A | B | C | D | E | F | G | H | I | J |
| Waktu (X) | : | 2 | 2 | 1 | 3 | 4 | 3 | 4 | 1 | 1 | 2 |
| Tes (Y) | : | 6 | 6 | 4 | 8 | 8 | 7 | 9 | 5 | 4 | 6 |

Apakah ada korelasi antara waktu belajar dengan hasil tes ?

| Siswa | X | X ² | Y | Y ² | XY |
|-------|------------|----------------|------------|----------------|-------------|
| A | | | | | |
| B | | | | | |
| | ΣX | ΣX^2 | ΣY | ΣY^2 | ΣXY |

2. KORELASI SPEARMAN (rho) dan Kendall (tau) :

Digunakan jika data variabel **ordinal** (berjenjang atau peringkat). Disebut juga korelasi non parametrik

$$r_p = 1 - \frac{6\sum d^2}{N(N^2 - 1)}$$

Di mana : N = banyak pasangan
d = selisih peringkat

Contoh :

10 orang siswa yang memiliki perilaku (sangat baik, baik, cukup, kurang) dibandingkan dengan tingkat kerajinannya (sangat rajin, rajin, biasa, malas)

| | | | | | | | | | | | |
|-----------|---|---|---|---|---|---|---|---|---|---|---|
| Siswa | : | A | B | C | D | E | F | G | H | I | J |
| Perilaku | : | 2 | 4 | 1 | 3 | 4 | 2 | 3 | 1 | 3 | 2 |
| Kerajinan | : | 3 | 2 | 1 | 4 | 4 | 3 | 2 | 1 | 2 | 3 |

Apakah ada korelasi antara perilaku siswa dengan kerajinannya ?

| Siswa | A | B | C | D | |
|----------------|---|---|---|---|-----------------|
| Perilaku | | | | | |
| Kerajinan | | | | | |
| d | | | | | |
| d ² | | | | | Σd ² |