ISBN 978-634-255-788-4



STATISTIKA PENELITIAN PENDIDIKAN

Dr. Afifah Latip Rasyid Jauhari, M.Pd.
Anwar Sadat, M.Pd.

STATISTIKA PENELITIAN PENDIDIKAN

Dr. Afifah Latip Rasyid Jauhari, M.Pd.
Anwar Sadat, M.Pd.

GET PRESS INDONESIA

STATISTIKA PENELITIAN PENDIDIKAN

Penulis :
Dr. Afifah Latip Rasyid Jauhari, M.Pd.
Anwar Sadat, M.Pd.

ISBN:978-634-255-788-4

Editor : Arabela Sinta, S.Si
Desain Sampul dan Tata Letak : Amisha Hartono, S.Pd.

PENERBIT GET PRESS INDONESIA
Anggota IKAPI No. 033/SBA/2022
Jl. Palarik RT 01 RW 06 Kelurahan Air Pacah
Kecamatan Koto Tangah Padang Sumatera Barat
website: www.getpress.co.id
email: adm.getpress@gmail.com

Cetakan pertama, Oktober 2026

Hak cipta dilindungi undang-undang. Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa izin tertulis dari penerbit.

KATA PENGANTAR

Puji dan syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga buku **Statistika Penelitian Pendidikan** ini dapat diselesaikan dengan baik. Buku ini disusun sebagai panduan komprehensif yang dirancang untuk membantu mahasiswa, pendidik, dan peneliti pendidikan dalam memahami serta mengaplikasikan konsep-konsep statistika secara praktis, jujur, dan terstruktur. Dalam dunia pendidikan modern, penguasaan analisis data merupakan fondasi mendasar untuk mengambil keputusan yang tepat, mengevaluasi pembelajaran, serta melahirkan penelitian yang valid dan dapat dipertanggungjawabkan.

Pembahasan dalam buku ini dirancang secara sistematis, dimulai dari penyajian data dasar, ukuran pemusatan dan penyebaran, teknik penyamplingan, hingga pengujian hipotesis parametrik dan non-parametrik. Selain menyajikan landasan teoritis dan formula matematis, buku ini dilengkapi dengan langkah-langkah praktis pengolahan data menggunakan program SPSS, panduan membaca luaran (*output*), serta latihan soal di setiap bab. Harapan kami, buku ini dapat mengikis stigma bahwa statistika adalah subjek yang rumit, sekaligus menjadi rujukan utama yang memudahkan pembaca dalam mengolah dan mengkaji data pendidikan secara mandiri dan efektif.

Jakarta, September 2026

Penulis

DAFTAR ISI

KATA PENGANTAR i
DAFTAR ISI ii
DAFTAR GAMBAR viii
BAB I KONSEP DASAR STATISTIKA PENDIDIKAN 1
A. Pendahuluan 1
B. Statistik dan Statistika 1
C. Ciri, Fungsi, dan Kegunaan Statistika 2
D. Pembagian Statistika 3
E. Alur Mengolah dan Mengkaji Data Pendidikan 4
F. Peta Pemilihan Uji Statistik 6
BAB II DATA STATISTIK DAN PENYAJIAN DATA 7
A. Pengertian dan Sumber Data 7
B. Klasifikasi Data 8
C. Skala Pengukuran 9
D. Tabel Distribusi Frekuensi 10
E. Penyajian Grafik 11
F. Prinsip Penyajian Data yang Jujur 13
BAB III UKURAN PEMUSATAN, LETAK DAN PENYEBARAN 15
A. Ukuran Pemusatan 15
B. Rata-rata Aritmetik 15
C. Rata-rata Geometrik dan Harmonik 17
D. Median dan Modus 17
E. Kuartil, Desil, dan Persentil 18

F. Rentang, Varians, dan Simpangan Baku....18
G. Koefisien Variasi dan Pemilihan Ukuran....19
LATIHAN....21
BAB IV POPULASI, SAMPEL, TEKNIK SAMPLING, DAN UJI NORMALITAS....23
A. Populasi, Sampel, Parameter, dan Statistik....23
B. Probability Sampling....24
C. Nonprobability Sampling....24
D. Menentukan Ukuran Sampel....25
E. Distribusi Normal....26
F. Pemeriksaan Normalitas....27
G. Langkah SPSS untuk Uji Normalitas....28
LATIHAN....37
BAB V PENGUJIAN HIPOTESIS DESKRIPTIF SATU SAMPEL....39
A. Logika Pengujian Hipotesis....39
B. Taraf Signifikansi, P-value, dan Kesalahan Keputusan....40
C. Uji Satu Sisi dan Dua Sisi....40
D. Uji t Satu Sampel....41
E. Interval Kepercayaan dan Ukuran Efek....42
F. Langkah SPSS dan Penyajian Hasil....42
LATIHAN....47
BAB VI UJI BINOMIAL, CHI-KUADRAT SATU SAMPEL, DAN RUN TEST....49
A. Uji Binomial....49

B. Chi-Kuadrat Goodness-of-Fit 50
C. Run Test 50
D. Pemilihan Uji dan SPSS 51
LATIHAN 61

BAB VII KOMPARATIF DUA SAMPEL BERPASANGAN: PAIRED T DAN WILCOXON 65

A. Desain Berpasangan 65
B. Paired Samples t-Test 66
C. Wilcoxon Signed-Rank Test 67
D. Ukuran Efek 68
E. Langkah SPSS dan Penyajian Hasil 68
LATIHAN 74

BAB VIII KOMPARATIF DUA SAMPEL INDEPENDEN: HOMOGENITAS, UJI T, WELCH, DAN MANN-WHITNEY 77

A. Desain Dua Kelompok Independen 77
B. Homogenitas Varians 78
C. Independent Samples t-Test 79
D. Mann-Whitney U 79
E. Ukuran Efek dan Penyajian Hasil 80
F. Langkah SPSS 80
LATIHAN 86

BAB IX UJI CHI-KUADRAT, MEDIAN TEST, DAN FISHER EXACT 89

A. Tabel Kontingensi dan Hipotesis 89
B. Contoh Ketuntasan Dua Kelas 90

C. Median Test untuk Sampel Independen 90
D. Asumsi dan Frekuensi Harapan 93
E. Fisher Exact Test 94
F. Ukuran Efek 94
G. SPSS dan Penyajian Hasil 95
LATIHAN 108

BAB X PERBANDINGAN RATA-RATA BEBERAPA KELOMPOK (ANOVA) 113

A. Mengapa Menggunakan ANOVA? 113
B. Asumsi ANOVA Satu Arah Independen 114
C. Contoh ANOVA Independen 114
D. Post Hoc dan Koreksi Multipel 115
E. Repeated-Measures ANOVA 115
F. Ukuran Efek 116
G. Langkah SPSS 116
LATIHAN 123

BAB XI UJI KRUSKAL-WALLIS DAN FRIEDMAN 125

A. Kruskal-Wallis untuk k Kelompok Independen 125
B. Friedman untuk k Kondisi Berpasangan 126
C. Penanganan Ties dan Post Hoc 126
D. Ukuran Efek 127
E. Langkah SPSS 127
LATIHAN 135

BAB XII UJI COCHRAN Q 139

A. Tujuan dan Struktur Data 139
B. Contoh Tiga Metode Mengajar 140

C. Asumsi dan Catatan 141
D. Post Hoc 141
E. Langkah SPSS 141
LATIHAN 147
H. LATIHAN PENERAPAN SPSS BAGIAN XII 147
BAB XIII KORELASI PEARSON DAN SPEARMAN 149
A. Hipotesis Asosiatif 149
B. Korelasi Product Moment Pearson 150
C. Contoh Pearson 151
D. Korelasi Rank Spearman 152
E. Signifikansi dan Interval Kepercayaan 152
F. Langkah SPSS dan Penyajian Hasil 153
LATIHAN 159
BAB XIV REGRESI LINEAR SEDERHANA, BERGANDA, DAN LOGISTIK 161
A. Konsep Dasar Regresi 161
B. Estimasi Intersep dan Koefisien Arah 162
C. Residual dan Asumsi 163
D. Keberartian Model dan R Square 164
E. Prediksi dan Bahaya Ekstrapolasi 165
F. Langkah SPSS Regresi Linear Sederhana 165
G. Regresi Linear Berganda 166
H. Regresi Logistik Biner 167
I. Pemilihan Model dan Prinsip Penyajian Hasil 168
LATIHAN 202
LAMPIRAN A MATRIKS PEMILIHAN UJI STATISTIK 205

LAMPIRAN B PANDUAN MEMBACA KELUARAN SPSS..... 208

LAMPIRAN C DATASET LATIHAN..... 210

C.1 Dataset Nilai 40 Peserta.....210

C.2 Dataset Berpasangan 210

C.3 Dataset Dua Kelompok.....211

C.4 Dataset Korelasi dan Regresi.....211

LAMPIRAN D KUNCI JAWABAN TERPILIH..... 213

LAMPIRAN E CARA MENYUSUN CATATAN PENGOLAHAN DATA 214

LAMPIRAN F RUBRIK EVALUASI PENERAPAN PENGOLAHAN DAN PENGKAJIAN DATA 216

GLOSARIUM 217

DAFTAR PUSTAKA..... 220

BIODATA PENULIS..... 223

DAFTAR GAMBAR

Gambar 1. 1. Peta awal pemilihan uji statistik berdasarkan tujuan, jumlah kelompok, dan hubungan sampel 6
Gambar 2. 1. Klasifikasi umum data statistik 8
Gambar 2. 2. Histogram dan poligon frekuensi nilai ujian 12
Gambar 2. 3. Ogive frekuensi kumulatif 12
Gambar 4. 1. Kurva normal dan persentase area berdasarkan simpangan baku 26
Gambar 4. 2. Metode pemeriksaan normalitas 27
Gambar 7. 1. Garis yang menghubungkan skor subjek pada dua kondisi 66
Gambar 8. 1. Boxplot dua kelompok independent 78
Gambar 13. 1. Contoh pola korelasi positif, negatif, dan mendekati nol 150
Gambar 14. 1. Scatterplot dan garis regresi nilai Qurdis terhadap Aqidah Akhlak 163
Gambar 14. 2. Contoh plot residual terhadap nilai prediksi 164

BAB I
KONSEP DASAR STATISTIKA PENDIDIKAN

A. Pendahuluan

Dalam kehidupan sehari-hari, keputusan sering didasarkan pada informasi numerik seperti rata-rata nilai kelas, persentase ketuntasan, jumlah siswa yang hadir, perubahan capaian belajar, atau hubungan antara motivasi dan prestasi. Angka-angka tersebut belum otomatis menjadi pengetahuan. Agar bermakna, data harus dikumpulkan dengan cara yang tepat, ditata, diringkas, dikaji, dan ditafsirkan sesuai konteks (Field, 2024; Triola, 2022).

Istilah statistik berasal dari kata Latin status yang pada awalnya berkaitan dengan pencatatan keadaan negara, seperti jumlah penduduk dan penghasilan. Penggunaannya kemudian meluas ke pendidikan, psikologi, ekonomi, kesehatan, dan bidang lain. Dalam pendidikan, statistika membantu peneliti menilai proses belajar, membandingkan metode pembelajaran, menguji ketercapaian kriteria, dan menilai hubungan antarvariabel.

B. Statistik dan Statistika

Statistik dapat diartikan sebagai kumpulan fakta berbentuk angka yang disusun dalam daftar, tabel, atau diagram sehingga menggambarkan suatu persoalan. Statistik juga dipakai untuk menyebut nilai ringkasan tertentu, misalnya rata-rata sampel, simpangan baku sampel, atau koefisien korelasi sampel.

Statistika adalah ilmu yang mempelajari cara mengumpulkan, menabulasi, menggolongkan, menyajikan, mengkaji, menafsirkan data, serta menarik kesimpulan yang

beralasan. Karena itu, statistik adalah produk atau hasil pengolahan data, sedangkan statistika adalah ilmu dan metode yang menghasilkan statistik. Statistika pendidikan secara khusus menerapkan prinsip tersebut pada data dan masalah bidang pendidikan (Subana *et al*., 2000; Sudijono, 2001).

C. Ciri, Fungsi, dan Kegunaan Statistika

Statistika memiliki tiga ciri utama yaitu bersifat kuantitatif, objektif dan universal. Bersifat kuantitatif artinya dapat disajikan dalam bentuk angka, objektif artinya data yang diolah sesuai fakta di lapangan tanpa memihak, dan universal artinya dapat dipakai di berbagai bidang ilmu.

Fungsi statistika meliputi menggambarkan data dalam bentuk yang komunikatif, menyederhanakan data kompleks, membuat perbandingan, mengukur besaran gejala, menilai hubungan, dan mendukung pembuatan keputusan. Dalam penelitian, statsitika membantu menentukan ukuran sampel, menilai validitas dan reliabilitas instrumen, menyajikan data, menguji hipotesis, dan mengukur ketidakpastian (Nurgiyantoro *et al*., 2000: Sugiyono, 2012).

Kegunaan Statistika membantu proses pengambilan keputusan yang lebih akurat, memprediksi atau meramalkan tren di masa depan, dan menyederhanakan data yang kompleks agar mudah dipahami.

D. Pembagian Statistika

Cabang-cabang statistika beserta contoh penerapannya dalam pendidikan dirangkum pada Tabel 1.1.

Cabang	Fokus	Contoh dalam pendidikan
Statistika teoretis	Pengembangan dalil, model, dan sifat matematis	Penurunan distribusi sampling dan pembuktian sifat estimator
Statistika terapan	Penggunaan metode statistik pada bidang tertentu	Pengolahan hasil tes dan evaluasi pembelajaran
Statistika deskriptif	Merangkum data tanpa generalisasi ke populasi	Tabel frekuensi, mean, median, grafik
Statistika inferensial	Menarik kesimpulan dari sampel ke populasi	Uji t, ANOVA, korelasi, regresi
Parametrik	Menggunakan asumsi distribusi dan parameter populasi	Uji t, ANOVA, Pearson
Nonparametrik	Lebih sedikit bergantung pada bentuk distribusi tertentu	Wilcoxon, Mann-Whitney, Kruskal-Wallis

Statistika deskriptif dan inferensial saling melengkapi (Sugiyono, 2017). Penerapan statsitika inferensial yang baik selalu didahului oleh deskripsi data. Tanpa mengetahui

jumlah observasi, pusat data, penyebaran, pencilan, dan pola grafik, peneliti beresiko menerapkan uji yang tidak sesuai atau menafsirkan hasil secara keliru (Gravetter & Wallnau, 2017; Triola, 2022).

Istilah parametrik dan nonparametrik tidak identik dengan data normal dan tidak normal saja. Pemilihan uji juga ditentukan skala data, bentuk desain, independensi observasi, ukuran sampel, ada atau tidaknya pencilan, dan tujuan pengolahan data (Field, 2024; Pallant, 2026).

E. Alur Mengolah dan Mengkaji Data Pendidikan

Alur mengolah dan mengkaji data Pendidikan terdiri dari beberapa tahapan sistematis agar data mentah berubah menjadi informasi yang akurat untuk kebijakan sekolah atau dinas Pendidikan.

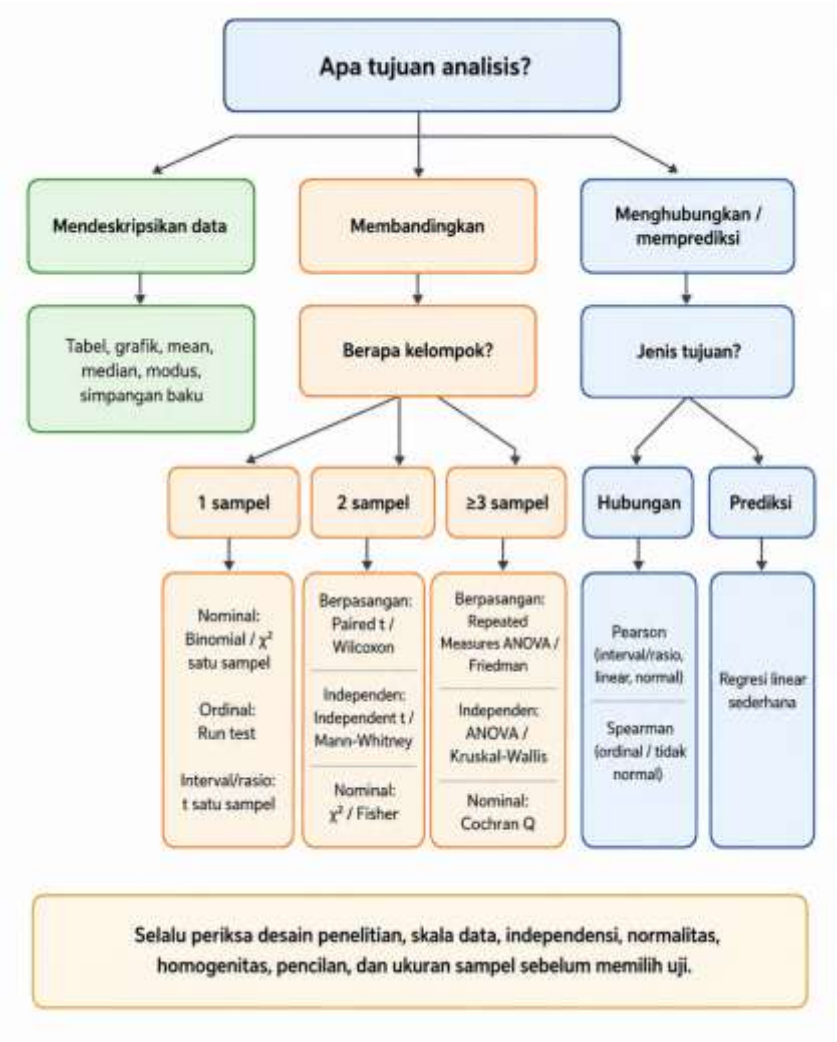
Tahapan tersebut terdiri dari 5 tahapan, diantaranya :

1. Pengumpulan Data
   Pengumpulan data dapat dilakukan dengan cara sensus, survei maupun dokumentasi.
2. Pengolahan Data
   Pengolahan data dapat dilakukan dengan pengkodean atau coding maupun tabulasi
3. Penyajian Data
   Data dapat disajikan dalam beberapa bentuk, yaitu Tabel, grafik atau diagram dan lain sebagainya.
4. Analisis Data
5. Analisis data terdiri dari tiga macam, yaitu : desktiptif, inferensial dan komparatif. Deskriptif adalah menghitung rata-rata nilai ujian, median dan persentase kelulusan. Inferensial yaitu menguji pengaruh fasilitas sekolah terhadap prestasi belajar

siswa sedangkan komparatif yaitu membandingkan capaian nilai antarsekolah atau antarwilayah.

6. Interpretasi Data dan Pelaporan
   Pada interpretasi data terdapat beberapa hal yang perlu diperhatikan yaitu penarikan Kesimpulan dan rekomendasi kebijakan. Penarikan Kesimpulan menjelaskan arti dari hasil analisis data yang ditemukan sedangkan rekomendasi kebijakan yaitu menyusun usulan perbaikan mutu atau alokasi bantuan dana. Setelah data diinterpretasikan, kemudian data dilaporkan dalam bentuk tertulis atau karya ilmiah.

F. Peta Pemilihan Uji Statistik



Gambar 1. 1. Peta awal pemilihan uji statistik berdasarkan tujuan, jumlah kelompok, dan hubungan sampel

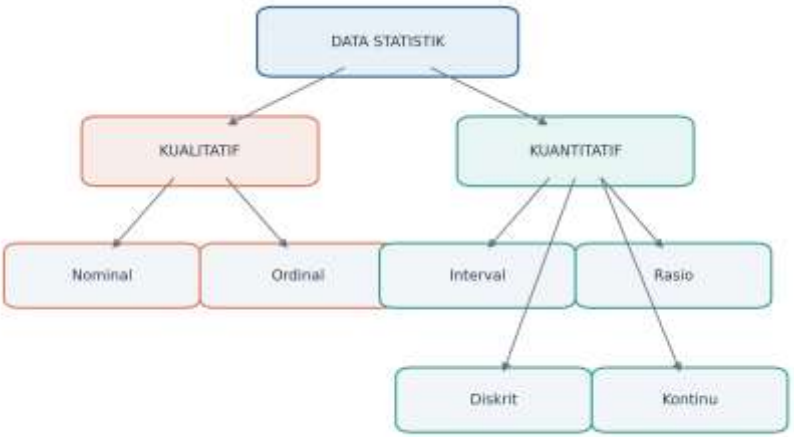
BAB II
DATA STATISTIK DAN PENYAJIAN DATA

A. Pengertian dan Sumber Data

Data adalah bentuk jamak dari datum, yaitu informasi tentang objek, individu, kejadian, atau variabel. Dalam penelitian pendidikan, data dapat berupa skor tes, kategori ketuntasan, tingkat motivasi, waktu menyelesaikan tugas, jumlah kehadiran, tanggapan wawancara, maupun catatan observasi (Sugiyono, 2012).

Berdasarkan sumbernya, terlihat bahwa data dapat dibedakan menjadi data primer dan sekunder. Data primer dikumpulkan langsung oleh peneliti melalui tes, angket, observasi, atau wawancara. Data sekunder diperoleh dari dokumen yang sudah tersedia, seperti rapor, basis data sekolah, atau laporan institusi. Keduanya harus dinilai dari segi ketepatan definisi, kelengkapan, validitas, dan kesesuaian waktu.

B. Klasifikasi Data



Gambar 2. 1. Klasifikasi umum data statistik

Data kualitatif menggambarkan kategori atau kualitas, misalnya jenis kelamin, status sekolah, kategori motivasi, dan tingkat kepuasan. Data kuantitatif menyatakan jumlah atau ukuran dalam bentuk angka, misalnya skor, usia, tinggi badan, waktu belajar, dan jumlah kehadiran. Pemberian kode angka pada data kualitatif tidak otomatis mengubahnya menjadi data kuantitatif. Kode laki-laki = 1 dan perempuan = 0, misalnya, tetap berskala nominal karena angka hanya digunakan sebagai label (Sudjana, 2005; Triola, 2022).

Data kuantitatif dapat bersifat diskrit atau kontinu. Data diskrit biasanya diperoleh melalui penghitungan dan berbentuk bilangan bulat, seperti jumlah siswa. Data kontinu diperoleh melalui pengukuran dan secara teoritis dapat mengambil nilai pada suatu rentang, seperti waktu, berat, atau tinggi badan.

C. Skala Pengukuran

Skala pengukuran terdiri atas nominal, ordinal, interval, dan rasio, yang sering disingkat NOIR. Karakteristik setiap skala disajikan pada Tabel 2.1.

Tabel 2. 1. Skala pengukuran dan implikasi pengolahan data

Skala	Karakteristik	Contoh	Operasi yang bermakna
Nominal	Kategori tanpa urutan	Jenis sekolah, program studi	Frekuensi, proporsi, modus
Ordinal	Kategori dapat diurutkan, jarak tidak pasti sama	Peringkat, kategori Likert	Median, persentil, rank
Interval	Jarak antarnilai bermakna, nol tidak absolut	Skor standar, suhu Celsius	Mean, simpangan baku, selisih
Rasio	Jarak bermakna dan memiliki nol absolut	Waktu belajar, jumlah hadir, usia	Semua operasi aritmetik yang relevan

Catatan Skala Likert

Satu butir Likert bersifat ordinal. Skor total dari banyak butir sering diperlakukan mendekati interval bila instrumen valid, reliabel, dan distribusinya memadai; keputusan ini perlu dijelaskan dalam metode.

D. Tabel Distribusi Frekuensi

Tabel distribusi frekuensi digunakan ketika data cukup banyak sehingga daftar nilai mentah tidak komunikatif. Tabel ini mengelompokkan data ke dalam interval kelas dan menunjukkan banyaknya observasi pada setiap interval. Distribusi frekuensi dapat berbentuk absolut, relatif, atau kumulatif (Sudjana, 2005).

Untuk membuat distribusi frekuensi, tentukan nilai minimum dan maksimum, hitung rentang, pilih jumlah kelas, tentukan panjang kelas, lalu hitung frekuensinya. Aturan Sturges dapat dipakai sebagai titik awal, tetapi interval akhir tetap perlu dibuat agar mudah dibaca.

$$k \approx 1 + 3{,}3 log_{10} n$$

Tabel 2. 2. Distribusi frekuensi nilai 40 peserta ujian

Interval nilai	Frekuensi	Frekuensi relatif (%)	Kumulatif
37-45	3	7,5	3
46-54	5	12,5	8
55-63	7	17,5	15

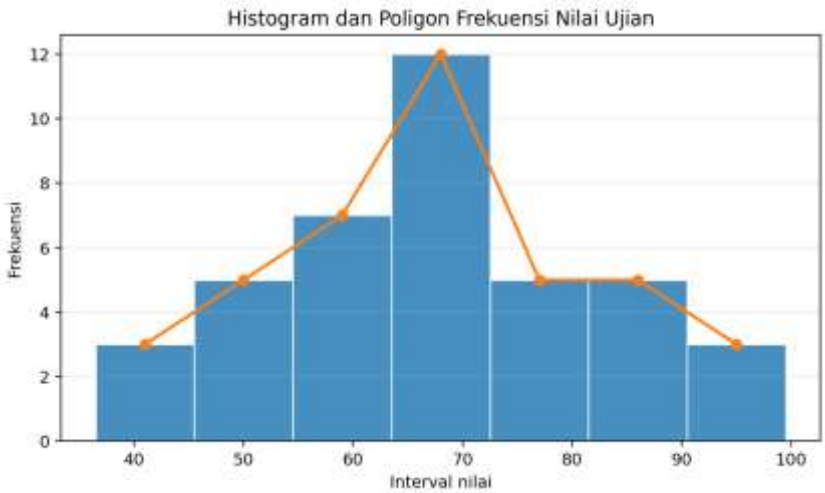
Interval nilai	Frekuensi	Frekuensi relatif (%)	Kumulatif
64-72	12	30,0	27
73-81	5	12,5	32
82-90	5	12,5	37
91-99	3	7,5	40
Jumlah	40	100,0	-

Sumber: Diolah dari bahan ajar Statistika Deskriptif

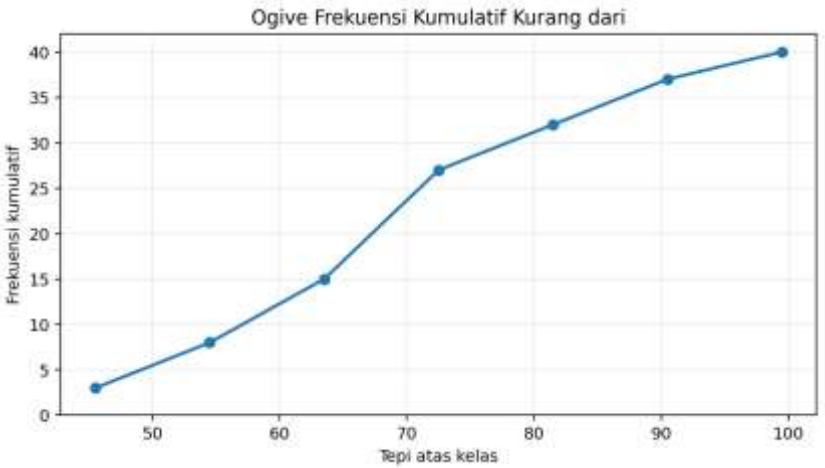
E. Penyajian Grafik

Grafik membantu pembaca melihat bentuk data, pusat, penyebaran, kecenderungan, dan pencilan. Histogram digunakan untuk data kuantitatif dan batangnya saling menempel. Grafik batang digunakan untuk data kategori dan batangnya biasanya terpisah. Poligon frekuensi menghubungkan titik tengah kelas, sedangkan ogive menunjukkan frekuensi kumulatif.

Diagram lingkaran dapat dipakai untuk komposisi sederhana dengan sedikit kategori, tetapi kurang efektif bila kategori banyak atau proporsinya mirip. Untuk perbandingan yang presisi, grafik batang sering lebih mudah dibaca. Selalu cantumkan judul, label sumbu, satuan, sumber data, dan skala yang tidak menyesatkan (Field, 2024; Triola, 2022).



Gambar 2. 2. Histogram dan poligon frekuensi nilai ujian



Gambar 2. 3. Ogive frekuensi kumulatif

F. Prinsip Penyajian Data yang Jujur

- Gunakan skala sumbu yang konsisten dan jelaskan bila sumbu tidak dimulai dari nol.
- Hindari efek tiga dimensi yang mengubah persepsi tinggi atau luas.
- Jangan memilih interval kelas hanya untuk menonjolkan pola tertentu.
- Tampilkan jumlah data dan data hilang bila relevan.
- Bedakan persentase dari perubahan persentase poin.
- Jangan menyimpulkan sebab-akibat hanya dari grafik deskriptif (Appelbaum *et al*., 2018).

BAB III
UKURAN PEMUSATAN, LETAK DAN PENYEBARAN

A. Ukuran Pemusatan

Ukuran pemusatan merangkum sekumpulan data dengan satu nilai yang dianggap mewakili pusat data. Tiga ukuran yang paling umum adalah mean, median, dan modus. Pemilihan ukuran tidak hanya ditentukan kemudahan hitung, tetapi juga skala data, bentuk distribusi, dan ada atau tidaknya pencilan (Gravetter & Wallnau, 2017; Triola, 2022).

Mean menggunakan seluruh nilai sehingga informatif, tetapi sensitif terhadap pencilan. Median membagi data terurut menjadi dua bagian dan lebih tahan terhadap pencilan. Modus menunjukkan nilai atau kategori yang paling sering muncul dan dapat digunakan pada data nominal.

B. Rata-rata Aritmetik

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

Rata-rata aritmetik data tunggal

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

Rata-rata aritmetik data berfrekuensi

Rata-rata sampel dilambangkan $\bar{x}$. Sementara itu, rata-rata populasi dilambangkan μ. Contoh data 2, 7, 3, 4, 5, 6, 9, dan 8 memiliki jumlah 44 sehingga mean-nya 44/8 = 5,5.

Untuk data berkelompok, x_i adalah titik tengah kelas dan f_i adalah frekuensinya. Pada distribusi berat badan dengan titik tengah 44, 49, 54, 59, 64, 69, dan 74 serta total frekuensi 50, jumlah $f_i x_i$ adalah 3.000 sehingga mean = 60(Sudjana, 2005).

Tabel 3. 1. Tabel penolong mean data berkelompok

Interval	f_i	x_i	$f_i x_i$
42-46	2	44	88
47-51	4	49	196
52-56	11	54	594
57-61	12	59	708
62-66	10	64	640
67-71	8	69	552
72-76	3	74	222
Jumlah	50	-	3.000

Sifat Transformasi Mean

Jika setiap nilai ditambah konstanta c, mean baru menjadi $\bar{x}$ + c. Jika setiap nilai dikali konstanta a, mean baru menjadi a $\bar{x}$. Sifat ini membantu pemeriksaan hasil dan pengubahan satuan.

C. Rata-rata Geometrik dan Harmonik

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}$$

Rata-rata geometrik data positif

$$H = \frac{n}{\sum(1/x_i)}$$

Rata-rata harmonik

Rata-rata geometrik cocok untuk pertumbuhan majemuk, perubahan berbasis rasio, indeks, atau laju pertumbuhan antarperiode. Semua nilai harus positif. Untuk data 2, 7, 3, 4, 5, 6, 9, dan 8, rata-rata geometrik sekitar 4,954.

Rata-rata harmonik tepat untuk rata-rata laju ketika pembilang atau jarak tetap. Jika seseorang menempuh jarak yang sama dengan kecepatan 30 km/jam saat pergi dan 20 km/jam saat pulang, kecepatan rata-rata bukan 25 km/jam, tetapi 24 km/jam. Hal ini diperoleh dari rata-rata harmonik dua kecepatan.

Untuk data positif berlaku hubungan mean aritmetik ≥ mean geometrik ≥ mean harmonik. Ketiganya akan sama hanya jika semua nilai identik (Triola, 2022).

D. Median dan Modus

$$Me = Tb + k \cdot \frac{(n/2) - Fk}{fMe}$$

Median data berkelompok

$$Mo = Tb + k \cdot \frac{d_1}{d_1 + d_2}$$

Modus data berkelompok

Median dicari setelah data diurutkan. Jika jumlah data ganjil, median adalah nilai di tengah. Jika jumlah data genap, median adalah rata-rata dari dua nilai tengah. Pada data

berkelompok, kelas median adalah kelas yang memuat posisi n/2. Pada Tabel 3.1, median berada pada kelas 57-61 dan nilainya sekitar 59,83.

Modus adalah nilai dengan frekuensi terbesar. Untuk data berkelompok, kelas modus adalah kelas berfrekuensi tertinggi. Dengan tepi bawah 56,5, panjang kelas 5, $d_1 = 1$, dan $d_2 = 2$, modus distribusi tersebut sekitar 58,17 (Sudjana, 2005; Triola, 2022).

E. Kuartil, Desil, dan Persentil

Kuartil membagi data terurut menjadi empat bagian, desil menjadi sepuluh bagian, dan persentil menjadi seratus bagian. Q_2 sama dengan median. Q_1 dan Q_3 sering digunakan untuk menggambarkan 50% bagian tengah data serta mendeteksi pencilan melalui rentang antarkuartil.

Untuk data berkelompok, posisi kuartil ke-i adalah in/4, posisi desil ke-i adalah in/10, dan posisi persentil ke-i adalah in/100. Perhitungan dilanjutkan dengan interpolasi pada kelas yang memuat posisi tersebut (Triola, 2022).

$$IQR = Q_3 - Q_1$$

Aturan Pencilan

Batas bawah umum = Q_1 - 1,5(IQR) dan batas atas = Q_3 + 1,5(IQR). Observasi di luar batas perlu diperiksa, tetapi tidak boleh otomatis dihapus.

F. Rentang, Varians, dan Simpangan Baku

$$R = xmaks - xmin$$

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$$

Varians sampel

$$s = \sqrt{\sum(x_i - \bar{x})^2/(n-1)}$$

Simpangan baku sampel

Rentang adalah selisih nilai maksimum dan minimum. Ukuran ini sederhana, tetapi hanya menggunakan dua nilai. Varians mengukur rata-rata kuadrat penyimpangan dari pusat. Sementara itu, simpangan baku adalah akar varians sehingga kembali ke satuan data asli.

Penyebut n - 1 digunakan untuk varians sampel sebagai koreksi derajat kebebasan. Pada data 60, 70, 65, 80, 70, 65, 75, 80, 70, dan 75, mean = 71, jumlah kuadrat penyimpangan = 390, varians sampel = 43,33, dan simpangan baku = 6,58(Gravetter & Wallnau, 2017).

G. Koefisien Variasi dan Pemilihan Ukuran

$$KV = \frac{s}{\bar{x}} \times 100\%$$

Koefisien variasi

Koefisien variasi menyatakan simpangan baku relatif terhadap mean dan berguna untuk membandingkan variasi dua variabel yang memiliki skala atau satuan berbeda. Penggunaannya memerlukan skala rasio dengan mean positif dan bermakna (Triola, 2022).

Untuk distribusi simetris tanpa pencilan, mean dan simpangan baku adalah pilihan umum. Untuk distribusi menceng atau mengandung pencilan, median dan IQR sering lebih representatif. Untuk data nominal, modus dan proporsi lebih sesuai (Field, 2024).

Tabel 3. 2. Panduan pemilihan ukuran deskriptif

Kondisi data	Ukuran pusat	Ukuran variasi
Nominal	Modus / proporsi	Rasio variasi kategori
Ordinal	Median	IQR / rentang
Interval/rasio simetris	Mean	Simpangan baku
Interval/rasio menceng	Median	IQR
Pertumbuhan majemuk	Mean geometrik	Rentang logaritmik
Laju untuk jarak tetap	Mean harmonik	Sesuai konteks

LATIHAN

1. Hitung mean, median, modus, rentang, varians sampel, dan simpangan baku data 7, 6, 8, 5, 7, 4, 5, 3, 4, 3, 4, 5, 6, 6, 7!
2. Jelaskan kapan mean geometrik lebih tepat daripada mean aritmetik!
3. Seseorang menempuh jarak yang sama dengan kecepatan 40 km/jam dan 60 km/jam. Hitung kecepatan rata-rata perjalanan!
4. Tentukan Q_1, Q_2, Q_3, dan IQR dari dataset latihan pada Lampiran C!
5. Bandingkan dua kelas yang memiliki mean sama tetapi simpangan baku berbeda. Apa maknanya bagi guru!

Dataset Latihan Penerapan SPSS

Ukuran pemusatan dan penyebaran

ID	Skor
1	55
2	60
3	62
4	65
5	68
6	70
7	72
8	72
9	74
10	76
11	78
12	80

ID	Skor
13	81
14	83
15	85
16	87
17	88
18	90
19	94
20	98

Status Hipotesis Statistik

Pengolahan data pada Bab 3 bertujuan merangkum data, bukan menguji dugaan populasi. H_0 dan H_a tidak digunakan.

1. Gunakan Explore untuk memperoleh mean, median, simpangan baku, kuartil, minimum, maksimum, histogram, dan boxplot.
2. Hitung rentang dan IQR berdasarkan keluaran SPSS.
3. Bandingkan mean dengan median. Jelaskan apakah pusat data tampak seimbang.
4. Identifikasi kemungkinan pencilan dan jelaskan pengaruhnya terhadap mean serta simpangan baku.
6. Susun satu paragraf penyajian deskriptif yang sesuai untuk karya ilmiah.

BAB IV
POPULASI, SAMPEL, TEKNIK SAMPLING, DAN UJI NORMALITAS

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:
membedakan populasi dan sampel; memilih teknik sampling; menjelaskan pertimbangan ukuran sampel; mengevaluasi normalitas secara grafik dan uji statistik.

Kata Kunci

Populasi, sampel, representatif, probability sampling, nonprobability sampling, normalitas, Shapiro-Wilk.

A. Populasi, Sampel, Parameter, dan Statistik

Populasi adalah keseluruhan objek atau subjek yang memiliki karakteristik tertentu sesuai tujuan penelitian. Sampel adalah bagian dari populasi yang dipilih untuk dipelajari. Kesimpulan dari sampel dapat digeneralisasikan hanya bila proses sampling, pengukuran, dan pengolahan data mendukung keterwakilan (Sugiyono, 2012).

Populasi target adalah kelompok yang menjadi sasaran generalisasi. Sementara itu, populasi terjangkau adalah bagian yang dapat diakses peneliti. Parameter adalah ukuran numerik populasi, seperti μ dan σ. Sementara itu, statistik sampel adalah ukuran yang dihitung dari sampel, seperti $\bar{x}$ dan s.

B. Probability Sampling

Probability sampling memberikan peluang terpilih yang diketahui bagi setiap anggota populasi. Teknik ini mendukung estimasi kesalahan sampling dan generalisasi yang lebih kuat, selama kerangka sampel lengkap dan pelaksanaan sesuai rancangan (Sugiyono, 2012).

Tabel 4. 1. Teknik probability sampling

Teknik	Kapan digunakan	Catatan
Simple random sampling	Populasi relatif homogen dan tersedia daftar anggota	Pemilihan dilakukan secara acak
Systematic sampling	Tersedia daftar berurutan	Pilih titik awal acak lalu interval k; waspadai pola periodik
Stratified random sampling	Populasi berstrata dan setiap strata perlu terwakili	Dapat proporsional atau tidak proporsional
Cluster sampling	Populasi tersebar luas dan unit alami berupa kelompok	Perlu memperhitungkan kemiripan anggota dalam klaster
Multistage sampling	Populasi besar dan bertingkat	Pemilihan dilakukan dalam beberapa tahap

C. Nonprobability Sampling

Tabel 4. 2. Teknik nonprobability sampling

Teknik	Karakteristik	Keterbatasan utama
Convenience/ incidental	Memilih responden yang mudah dijangkau	Rentan bias seleksi

Teknik	Karakteristik	Keterbatasan utama
Purposive	Memilih berdasarkan kriteria atau pertimbangan	Generalisasi terbatas pada konteks
Quota	Menetapkan jumlah untuk kategori tertentu	Pemilihan dalam kuota tidak acak
Jenuh/sensus	Semua anggota populasi terjangkau dijadikan responden	Tetap dapat mengalami nonresponse
Snowball	Responden awal merekomendasikan responden lain	Jaringan awal dapat mendominasi sampel

Nonprobability sampling tidak selalu salah. Teknik ini dapat relevan untuk studi eksploratif, populasi sulit dijangkau, penelitian kualitatif, atau evaluasi lokal. Namun peneliti harus membatasi klaim dan menjelaskan alasan serta potensi bias (Sugiyono, 2012).

D. Menentukan Ukuran Sampel

Ukuran sampel tidak ditentukan hanya oleh besar populasi. Pertimbangan penting meliputi desain, tingkat kepercayaan, ketelitian estimasi, variabilitas, ukuran efek yang diharapkan, jumlah kelompok, jumlah prediktor, tingkat kehilangan data, dan rencana pengolahan data.

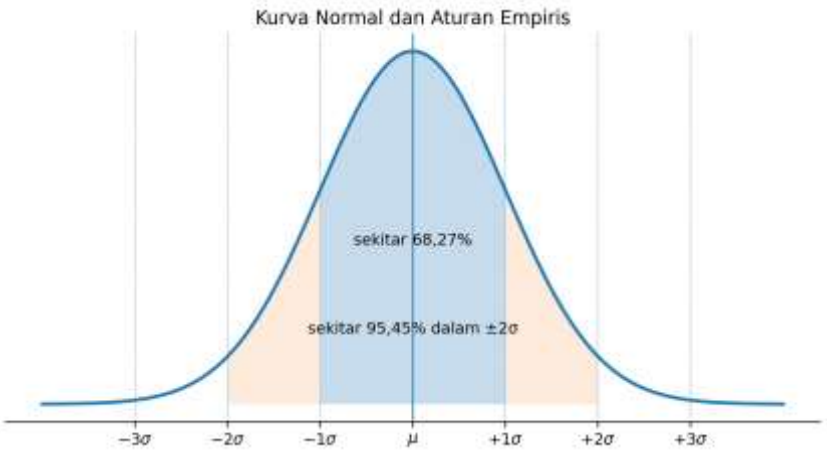
Tabel Isaac dan Michael atau rumus sederhana dapat menjadi perkiraan awal, tetapi power analysis lebih informatif untuk pengujian hipotesis. Sampel besar tidak memperbaiki bias sampling atau instrumen yang buruk. Sebaliknya, sampel kecil dapat menghasilkan interval

kepercayaan lebar dan daya uji rendah (Cohen, 1988; Gravetter & Wallnau, 2017).

Prinsip Penting

Representativitas tidak dijamin oleh jumlah sampel saja. Sampel besar yang dipilih secara bias tetap dapat menghasilkan kesimpulan bias.

E. Distribusi Normal



Gambar 4. 1. Kurva normal dan persentase area berdasarkan simpangan baku

Distribusi normal berbentuk lonceng, simetris, dan memiliki mean, median, serta modus yang berimpit. Sekitar 68,27% observasi berada dalam μ ± 1σ, 95,45% dalam μ ± 2σ, dan 99,73% dalam μ ± 3σ. Distribusi normal adalah model teoretis; data empiris jarang benar-benar sempurna (Sundayana, 2010; Triola, 2022).

Dalam banyak uji parametrik, asumsi normalitas berlaku pada residual, skor perbedaan, atau distribusi dalam setiap kelompok, bukan selalu pada seluruh data yang digabung. Untuk paired t-test, yang diperiksa adalah distribusi

perbedaan. Untuk regresi, yang diperiksa terutama residual (Field, 2024).

F. Pemeriksaan Normalitas

Gambar 4. 2. Metode pemeriksaan normalitas

Cara	Apa yang diperiksa	Kelebihan / keterbatasan
Histogram	Bentuk distribusi, kemencengan, multimodal	Mudah, tetapi dipengaruhi pilihan bin
Q-Q plot	Kedekatan kuantil data dengan garis normal	Sangat informatif untuk pola penyimpangan
Boxplot	Pencilan dan asimetri	Tidak secara khusus menguji normalitas
Shapiro-Wilk	Hipotesis distribusi normal	Umum untuk sampel kecil-menengah; sensitif pada sampel besar
Kolmogorov-Smirnov/ Lilliefors	Perbedaan distribusi empiris dan teoretis	Daya dapat lebih rendah daripada Shapiro-Wilk
Skewness/ kurtosis	Derajat kemencengan dan keruncingan	Perlu ditafsirkan relatif terhadap ukuran sampel

Pada uji normalitas, H_0 biasanya menyatakan data mengikuti distribusi normal. Jika $p > \alpha$, tidak ditemukan bukti yang cukup untuk menolak normalitas. Pernyataan ini bukan bukti bahwa distribusi pasti normal. Untuk sampel besar, penyimpangan kecil dapat menghasilkan $p < 0{,}05$ meskipun uji parametrik tetap cukup robust. Karena itu,

gabungkan uji formal dengan grafik, ukuran sampel, dan sifat uji yang akan digunakan (Field, 2024; Pallant, 2026; Shapiro & Wilk, 1965).

G. Langkah SPSS untuk Uji Normalitas

1. Siapkan variabel dan masukkan data pada Data View.
2. Akses Analyze > Descriptive Statistics > Explore.
3. Isikan variabel ke Dependent List; masukkan variabel kelompok ke Factor List bila diperlukan.
4. Buka Plots dan centang Normality plots with tests serta histogram.
5. Baca tabel Tests of Normality, histogram, Q-Q plot, dan boxplot.
7. Laporkan metode, statistik uji, p-value, dan keputusan yang relevan dengan pengujian lanjutan (Pallant, 2026).

Contoh Panduan membaca keluaran

Uji Shapiro-Wilk menunjukkan skor perbedaan berdistribusi normal, W = 0,96, p = 0,42. Karena p > 0,05 dan Q-Q plot tidak memperlihatkan penyimpangan berat, asumsi normalitas untuk paired t-test dianggap memadai.

CONTOH PENERAPAN BAB 4

Memeriksa Normalitas Data dengan Shapiro-Wilk dan Q-Q Plot

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: peneliti akan menggunakan uji parametrik terhadap nilai kemampuan komunikasi matematis 30 siswa. Sebelum pengujian, distribusi data diperiksa.

Tabel 4. 3. Ilustrasi 4.1 Data fiktif untuk pemeriksaan normalitas

ID	Nilai
1	72
2	75
3	78
4	80
5	81
6	83
7	84
8	76
9	79
10	82
11	85
12	87
13	74
14	77
15	80
16	82

ID	Nilai
17	86
18	88
19	73
20	75
21	79
22	81
23	83
24	84
25	85
26	76
27	78
28	80
29	82
30	89

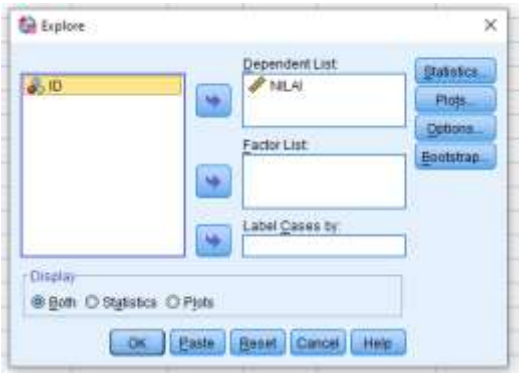
Rumusan Hipotesis pada Ilustrasi

H_0: Data nilai berasal dari populasi yang berdistribusi normal.

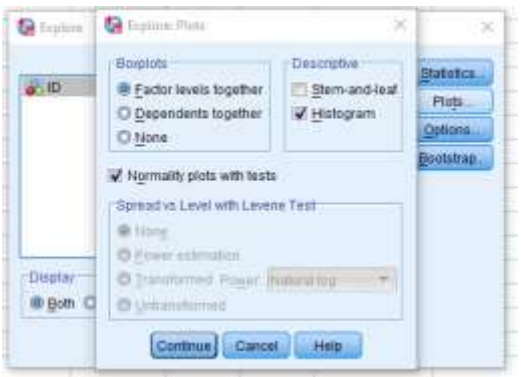
H_a: Data nilai berasal dari populasi yang tidak berdistribusi normal.

Langkah Mengolah dan Mengkaji Data

1. Masukkan variabel id dan nilai.
2. Akses Analyze > Descriptive Statistics > Explore.
3. Masukkan nilai ke Dependent List.



4. Pada Plots, aktifkan Factor levels together, Histogram dan Normality plots with tests.



5. Klik OK, keluaran akan muncul sebagai berikut:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
NILAI	30	100.0%	0	0.0%	30	100.0%

Pada tabel keluaran Case Processing Summary, seluruh data NILAI yang berjumlah 30 observasi dinyatakan valid. Tidak ada data yang hilang, sehingga seluruh data dapat digunakan dalam pemeriksaan normalitas. Tabel ini hanya menunjukkan kelengkapan data dan belum dapat dipakai

untuk menyimpulkan apakah distribusi data normal atau tidak.

Descriptives

			Statistic	Std. Error
NILAI	Mean		80.47	.822
	95% Confidence Interval for Mean	Lower Bound	78.79	
		Upper Bound	82.15	
	5% Trimmed Mean		80.46	
	Median		80.50	
	Variance		20.257	
	Std. Deviation		4.501	
	Minimum		72	
	Maximum		89	
	Range		17	
	Interquartile Range		7	
	Skewness		-.025	.427
	Kurtosis		-.718	.833

Tabel Descriptives menunjukkan rata-rata nilai 80,47, median 80,50, dan 5% trimmed mean 80,46. Ketiga nilai ini hampir sama, sehingga pusat data cukup seimbang dan tidak banyak dipengaruhi nilai ekstrem. Nilai berada antara 72 dan 89, dengan simpangan baku 4,501 dan rentang 17. Skewness -0,025 menunjukkan bentuk yang hampir simetris. Kurtosis -0,697 menunjukkan distribusi sedikit lebih datar daripada distribusi normal. Secara deskriptif, pola data mendekati normal.

Percentiles

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average (Definition 1)	NILAI	72.55	74.10	76.75	80.50	84.00	86.90	88.45
Tukey's Hinges	NILAI			77.00	80.50	84.00		

Tabel Percentiles menunjukkan median 80,50. Tukey's Hinges menghasilkan Q1 = 77,00 dan Q3 = 84,00, sehingga

50% nilai bagian tengah berada antara 77,00 dan 84,00. Jarak median ke Q1 dan Q3 sama-sama 3,50. Pola ini menunjukkan bahwa bagian tengah data relatif simetris. Namun, keputusan normalitas tetap harus mempertimbangkan uji dan grafik lain.

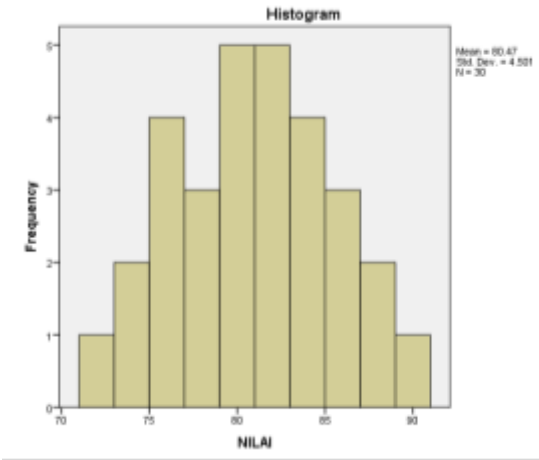
Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
NILAI	.073	30	.200*	.983	30	.904

*. This is a lower bound of the true significance.

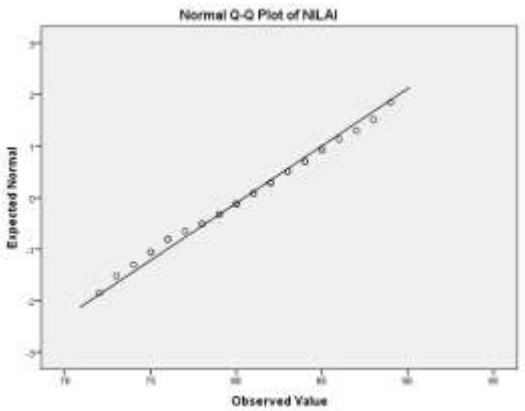
a. Lilliefors Significance Correction

Tabel Tests of Normality menunjukkan bahwa nilai signifikansi Kolmogorov–Smirnov sebesar 0,200 dan Shapiro–Wilk sebesar 0,904. Kedua nilai signifikansi tersebut lebih besar dari 0,05. Karena itu, H_0 gagal ditolak dan data NILAI dapat dianggap berdistribusi normal. Kesimpulan ini terutama didasarkan pada uji Shapiro–Wilk karena jumlah data yang dikaji sebanyak 30 observasi.

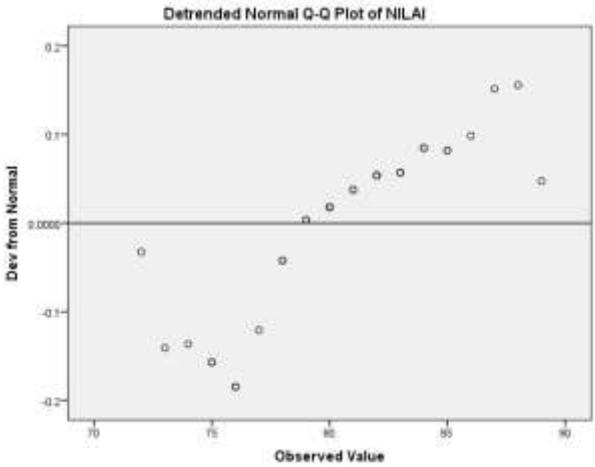


Histogram menunjukkan bahwa nilai 30 peserta cukup simetris dan mendekati bentuk lonceng. Sebagian besar nilai

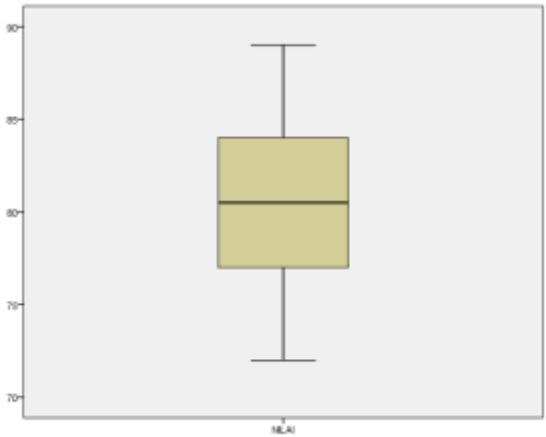
berada di sekitar rata-rata 80,47. Frekuensi menurun secara bertahap ke arah nilai yang lebih rendah dan lebih tinggi. Tidak terlihat pencilan yang mencolok. Pola ini mendukung hasil Shapiro-Wilk yang menunjukkan data mendekati normal.



Berdasarkan Normal Q-Q Plot, titik-titik data berada relatif dekat dan mengikuti arah garis diagonal. Tidak terlihat penyimpangan sistematis atau titik yang menjauh secara ekstrem dari garis acuan. Pola tersebut menunjukkan bahwa distribusi NILAI secara visual mendekati distribusi normal. Hasil ini sejalan dengan uji Shapiro-Wilk yang memperoleh nilai signifikansi sebesar 0,904 atau lebih besar dari 0,05. Karena itu, asumsi normalitas data dinyatakan terpenuhi.



Detrended Normal Q-Q plot menunjukkan bahwa titik-titik data berada relatif dekat dengan garis nol. Meskipun ada pola ringan, yaitu titik pada nilai rendah cenderung berada di bawah garis nol dan titik pada nilai tinggi berada di atas garis nol, besar penyimpangannya relatif kecil. Karena itu, penyimpangan tersebut tidak menunjukkan pelanggaran normalitas yang serius. Hasil ini tetap mendukung kesimpulan bahwa data NILAI mendekati distribusi normal.



Boxplot menunjukkan bahwa nilai peserta tersebar antara 72 sampai 89, dengan median sebesar 80,50. Kuartil pertama sebesar 77 dan kuartil ketiga sebesar 84, sehingga 50% data bagian tengah berada dalam rentang 77 sampai 84. Posisi median yang berada di tengah kotak serta panjang whisker yang relatif seimbang menunjukkan bahwa distribusi data cukup simetris. Tidak ada titik di luar whisker, sehingga tidak ditemukan pencilan pada data nilai.

Catatan: Periksa tabel Tests of Normality, Q-Q plot, histogram, dan boxplot. Jangan menentukan keputusan hanya dari satu indikator.

Cara Membaca Keluaran

Pada dataset ilustratif, uji Shapiro-Wilk menghasilkan W = 0,983 dan p = 0,904. Karena p lebih besar dari 0,05, data tidak menunjukkan penyimpangan yang berarti dari distribusi normal. Histogram, Q-Q plot, detrended Q-Q plot, dan boxplot juga menunjukkan pola yang cukup simetris tanpa pencilan yang berarti.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p > 0,05$, H_0 gagal ditolak dan H_a belum didukung. Jadi, data nilai dapat dianggap memenuhi asumsi normalitas untuk pengujian parametrik yang relevan, dengan tetap mempertimbangkan hasil pemeriksaan grafik.

Normalitas yang relevan untuk paired t-test adalah normalitas skor selisih. Untuk regresi, yang diperiksa adalah residual, bukan sekadar variabel mentah.

LATIHAN

1. Bedakan populasi target, populasi terjangkau, sampel, parameter, dan statistik sampel!
2. Pilih teknik sampling untuk survei siswa SMA di seluruh kabupaten dan jelaskan alasannya!
3. Mengapa purposive sampling tidak otomatis buruk?
4. Uraikan mengapa p > 0,05 pada Shapiro-Wilk bukan bukti mutlak normalitas!
5. Untuk paired t-test, variabel manakah yang seharusnya diuji normalitasnya!

Dataset Latihan Penerapan SPSS

Pemeriksaan normalitas dan pencilan

ID	Nilai
1	42
2	48
3	52
4	55
5	58
6	60
7	62
8	64
9	65
10	66
11	67
12	68
13	69
14	70
15	71

ID	Nilai
16	72
17	73
18	74
19	76
20	78
21	80
22	82
23	85
24	92
25	100

Rumusan Hipotesis pada Ilustrasi

H_0: Data berasal dari populasi yang berdistribusi normal.

H_a: Data berasal dari populasi yang tidak berdistribusi normal.

1. Gunakan prosedur Explore dengan Shapiro-Wilk, histogram, Q-Q plot, dan boxplot.
2. Catat W dan p-value.
3. Tentukan apakah data layak diperlakukan sebagai normal. Gunakan bukti dari uji dan grafik.
4. Identifikasi nilai ekstrem atau pencilan.
5. Uraikan pilihan pengujian jika data akan dibandingkan dengan nilai standar 75.

BAB V
PENGUJIAN HIPOTESIS DESKRIPTIF SATU SAMPEL

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: merumuskan H_0 dan H_a; membedakan uji satu sisi dan dua sisi; menjelaskan kesalahan tipe I dan II; melakukan uji t satu sampel dan menyajikan dan menafsirkan hasilnya.

Kata Kunci

Hipotesis, H_0, H_a, alpha, p-value, uji satu sisi, uji dua sisi, one-sample t-test.

A. Logika Pengujian Hipotesis

Hipotesis deskriptif adalah dugaan mengenai nilai satu variabel populasi tanpa membandingkan kelompok atau menguji hubungan. Contohnya adalah dugaan bahwa rata-rata hasil belajar telah mencapai KKM 70 atau proporsi siswa yang tuntas adalah 80% (Sugiyono, 2012).

Pengujian dimulai dengan H_0 sebagai pernyataan acuan dan H_a sebagai pernyataan yang ingin didukung. Data sampel digunakan untuk menilai apakah hasil yang diamati cukup tidak lazim bila H_0 benar. Keputusan statistik bersifat probabilistik dan tidak menyatakan kepastian absolut (Gravetter & Wallnau, 2017).

B. Taraf Signifikansi, P-value, dan Kesalahan Keputusan

Taraf signifikansi α adalah batas risiko kesalahan tipe I yang ditetapkan sebelum pengolahan data, umumnya 0,05. Kesalahan tipe I terjadi ketika H_0 benar tetapi ditolak. Kesalahan tipe II terjadi ketika H_0 salah tetapi tidak ditolak. Daya uji adalah peluang mendeteksi efek yang benar-benar ada.

P-value adalah probabilitas memperoleh hasil setidaknya se-ekstrem data yang diamati dengan asumsi H_0 benar. P-value bukan peluang bahwa H_0 benar dan bukan ukuran besarnya efek. Keputusan $p < \alpha$ berarti menolak H_0; $p \geq \alpha$ berarti gagal menolak H_0, bukan membuktikan H_0 (Wasserstein & Lazar, 2016).

Tabel 5. 1. Kemungkinan keputusan dalam pengujian hipotesis

Keadaan sebenarnya	Keputusan: tidak menolak H_0	Keputusan: menolak H_0
H_0 benar	Keputusan benar	Kesalahan tipe I (α)
H_0 salah	Kesalahan tipe II (β)	Keputusan benar / daya = 1-β

C. Uji Satu Sisi dan Dua Sisi

Uji dua sisi digunakan ketika perbedaan ke dua arah sama-sama relevan: H_0: $\mu = \mu_0$ dan H_a: $\mu \neq \mu_0$. Uji sisi kanan digunakan untuk dugaan peningkatan: H_0: $\mu \leq \mu_0$ dan H_a: $\mu > \mu_0$. Uji sisi kiri digunakan untuk dugaan penurunan: H_0: $\mu \geq \mu_0$ dan H_a: $\mu < \mu_0$.

Arah uji harus ditetapkan berdasarkan rumusan masalah sebelum melihat data. Mengubah uji dua sisi menjadi satu sisi setelah melihat hasil meningkatkan risiko kesimpulan keliru (Field, 2024).

D. Uji t Satu Sampel

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

Statistik uji t satu sampel, df = n - 1

Uji t satu sampel digunakan untuk membandingkan mean sampel dengan nilai yang dihipotesiskan ketika data berskala interval/rasio, observasi independen, dan distribusi tidak mengalami pelanggaran berat. Untuk sampel kecil, normalitas dan pencilan perlu diperiksa lebih cermat (Field, 2024; Triola, 2022).

Dari 20 siswa dengan frekuensi skor 60 sebanyak 2, 65 sebanyak 4, 70 sebanyak 7, dan 75 sebanyak 7, diperoleh mean 69,75 dan simpangan baku 4,99. Dibandingkan nilai acuan 70, t = -0,224 dengan df = 19 dan p = 0,825. Tidak ada bukti bahwa mean berbeda dari 70.

Tabel 5. 2. Data contoh uji t satu sampel

Skor	Frekuensi
60	2
65	4
70	7
75	7
Jumlah	20

E. Interval Kepercayaan dan Ukuran Efek

Interval kepercayaan memberikan rentang nilai populasi yang konsisten dengan data dan model. Jika interval kepercayaan 95% untuk selisih mean terhadap μ_0 mencakup nol, hasil uji dua sisi pada $\alpha = 0{,}05$ tidak signifikan.

Ukuran efek one-sample Cohen d dapat dihitung sebagai $(\bar{x} - \mu_0)/s$. Ukuran efek membantu menilai besarnya perbedaan dalam satuan simpangan baku, tetapi penjelasan makna harus mempertimbangkan konteks pendidikan (Cohen, 1988; Lakens, 2013).

$$d = \frac{\bar{x} - \mu_0}{s}$$

F. Langkah SPSS dan Penyajian Hasil

1. Buka Analyze > Compare Means > One-Sample T Test.
2. Masukkan variabel ke Test Variable(s).
3. Tempatkan nilai pembanding pada Test Value.
4. Baca mean, mean difference, t, df, Sig. (2-tailed), dan confidence interval.
5. Untuk hipotesis satu sisi, gunakan setengah p dua sisi hanya bila arah hasil sesuai hipotesis dan arah ditetapkan sebelumnya (Field, 2024; Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Rata-rata skor siswa (M = 69,75; SD = 4,99) tidak berbeda secara signifikan dari KKM 70, t(19) = -0,22, p = 0,825, d = -0,05. Dengan demikian, data belum memberikan bukti bahwa rata-rata populasi berada di atas atau di bawah 70.

CONTOH PENERAPAN BAB 5

Uji t Satu Sampel terhadap Kriteria Ketuntasan

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: nilai hasil belajar 20 mahasiswa dibandingkan dengan standar kompetensi 75. Hipotesis dua sisi dipakai untuk menilai apakah rata-rata populasi berbeda dari 75.

Tabel 5. 3. Ilustrasi 5.1. Data fiktif uji t satu sampel

ID	Nilai
1	78
2	82
3	80
4	76
5	85
6	79
7	81
8	84
9	77
10	83
11	80
12	86
13	75
14	82
15	79
16	81

ID	Nilai
17	88
18	76
19	84
20	80

Rumusan Hipotesis pada Ilustrasi

H_0: μ = 75 (rata-rata nilai populasi sama dengan 75).

H_a: μ ≠ 75 (rata-rata nilai populasi tidak sama dengan 75).

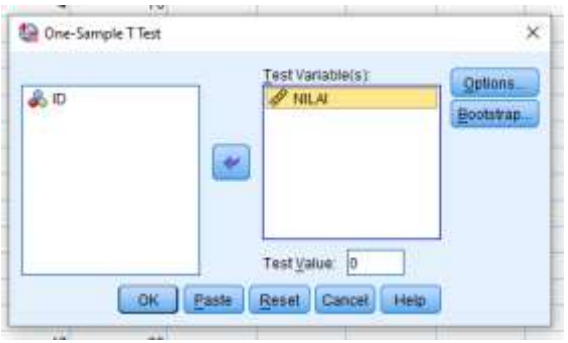
Taraf signifikansi: 0,05.

Langkah Mengolah dan Mengkaji Data

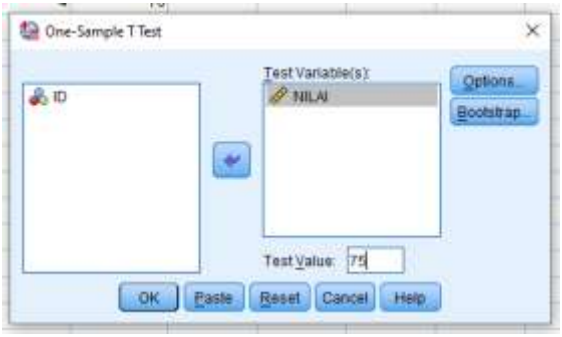
1. Periksa normalitas dan pencilan nilai.

   Pada contoh ini, data dianggap telah memenuhi asumsi normalitas.

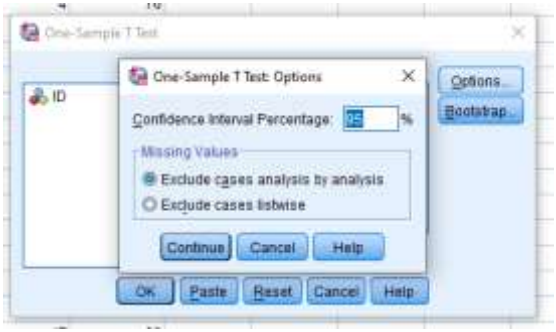
2. Akses Analyze > Compare Means > One-Sample T Test.
3. Masukkan nilai ke Test Variable(s).



4. Isi Test Value dengan 75.



5. Pada Options, gunakan confidence interval 95%.



6. Pilih OK, didapatkan keluaran berikut:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
NILAI	20	80.80	3.518	.787

One-Sample Test

	Test Value = 75					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
NILAI	7.372	19	.000	5.800	4.15	7.45

Cara Membaca Keluaran

Pada dataset ilustratif, rata-rata nilai peserta sebesar 80,80 dengan simpangan baku 3,52, atau 5,80 poin lebih tinggi daripada nilai acuan 75. Uji t satu sampel menghasilkan $t(19) = 7{,}37$, $p < 0{,}001$, dengan interval kepercayaan 95% untuk selisih rata-rata [4,15, 7,45]. Cohen's d sebesar 1,65 menunjukkan ukuran efek yang sangat besar.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Rata-rata nilai populasi berbeda secara signifikan dari 75 dan, dilihat dari arah selisih, lebih tinggi daripada nilai acuan tersebut. Perbedaannya juga bermakna secara praktis karena memiliki ukuran efek yang sangat besar.

LATIHAN

1. Rumuskan H_0 dan H_a dua sisi untuk menguji apakah mean literasi digital berbeda dari 75!
2. 2.Jelaskan perbedaan kesalahan tipe I, tipe II, dan daya uji!
3. Hitung t satu sampel untuk data pada Lampiran C dengan nilai acuan 70!
4. Mengapa "p > 0,05 membuktikan H_0 benar" adalah pernyataan keliru?
5. Tulis contoh penyusunan kalimat statistik one-sample t-test yang lengkap!

Dataset Latihan Penerapan SPSS

Uji t satu sampel

ID	Nilai
1	70
2	74
3	78
4	80
5	76
6	82
7	75
8	79
9	81
10	77
11	73
12	85

ID	Nilai
13	76
14	78
15	80
16	72
17	84
18	79

Rumusan Hipotesis pada Ilustrasi

H_0: $\mu = 75$ (rata-rata nilai populasi sama dengan KKM 75).

H_a: $\mu \neq 75$ (rata-rata nilai populasi berbeda dari KKM 75).

1. Uji apakah rata-rata nilai berbeda dari KKM 75 menggunakan taraf 0,05.
2. Tinjau normalitas dan pencilan terlebih dahulu.
3. Catat N, mean, SD, t, df, p-value, mean difference, dan 95% confidence interval.
4. Hitung Cohen's d secara manual dari mean dan SD.
5. Tulis keputusan statistik dan penjelasan makna pendidikan dalam satu paragraf.

BAB VI
UJI BINOMIAL, CHI-KUADRAT SATU SAMPEL, DAN RUN TEST

Capaian Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu: memilih uji untuk data nominal dan ordinal satu sampel; melakukan uji binomial dan goodness-of-fit; menilai kerandoman urutan dengan run test; menafsirkan hasil secara tepat.

Kata Kunci

binomial, chi-kuadrat, goodness-of-fit, frekuensi harapan, run test, kerandoman.

A. Uji Binomial

Uji binomial digunakan untuk data nominal dikotomi, seperti lulus-gagal, setuju-tidak setuju, atau memilih A-B. Uji ini menilai apakah proporsi keberhasilan populasi sama dengan proporsi yang dihipotesiskan. Pengamatan harus independen, setiap pengamatan hanya memiliki dua kemungkinan, dan probabilitas acuan ditetapkan (Triola, 2022).

Pada contoh yang digunakan, dari 24 responden ada 15 yang memilih mobil bensin dan 9 memilih solar. Dengan H_0: $p = 0{,}50$, probabilitas hasil yang setidaknya sama ekstrem dinilai menggunakan distribusi binomial. Hasil $p > 0{,}05$ menunjukkan belum cukup bukti bahwa kecenderungan dua pilihan berbeda dari 50:50.

$$P(X = x) = C(n, x)p^{x}(1 - p)^{n-x}$$

Peluang binomial untuk tepat x keberhasilan

B. Chi-Kuadrat Goodness-of-Fit

Uji chi-kuadrat satu sampel atau goodness-of-fit digunakan untuk menilai apakah frekuensi kategori mengikuti proporsi teoretis tertentu. Data berupa hitungan frekuensi, kategori saling lepas, dan observasi independen. Frekuensi harapan biasanya diupayakan tidak terlalu kecil (Triola, 2022).

Contoh pilihan warna mobil menghasilkan frekuensi observasi 1.000 silver, 900 hitam, 600 putih, dan 500 warna lain. Bila keempat warna dihipotesiskan sama, setiap frekuensi harapan adalah 750. Nilai chi-kuadrat 226,67 jauh lebih besar daripada nilai kritis df = 3, sehingga H_0 ditolak.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Statistik chi-kuadrat goodness-of-fit

Tabel 6. 1. Perhitungan chi-kuadrat satu sampel

Warna	O_i	E_i	Kontribusi χ^2
Silver	1.000	750	83,33
Hitam	900	750	30,00
Putih	600	750	30,00
Lain	500	750	83,33
Jumlah	3.000	3.000	226,67

C. Run Test

Run adalah rangkaian simbol sejenis yang berurutan dan dibatasi simbol berbeda atau ujung deret. Run test menilai apakah urutan dua kategori bersifat acak. Uji ini dapat

dipakai untuk data urutan yang dikodekan menjadi dua kelompok, misalnya di atas dan di bawah median (Triola, 2022).

Urutan GGAGGGAAGAAAGAGGGAAA memiliki 10 run. Terlalu sedikit run menunjukkan pengelompokan. Sementara itu, terlalu banyak run menunjukkan pergantian yang sangat teratur. Untuk sampel kecil, keputusan menggunakan tabel batas run; untuk sampel besar dapat digunakan pendekatan normal.

$$E(R) = \frac{2n_1n_2}{n_1 + n_2} + 1$$

Nilai harapan banyak run

$$Var(R) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

Penafsiran Run Test

Hasil signifikan menunjukkan urutan tidak konsisten dengan kerandoman, tetapi tidak menjelaskan penyebab pola. Urutan waktu, prosedur pencatatan, atau pengelompokan responden perlu diperiksa.

D. Pemilihan Uji dan SPSS

Tabel 6. 2. Pemilihan uji satu sampel nonparametrik

Masalah	Data	Uji
Proporsi satu kategori dari dua pilihan	Nominal dikotomi	Binomial
Kesesuaian beberapa kategori dengan proporsi teoretis	Nominal ≥2 kategori	Chi-kuadrat goodness-of-fit
Kerandoman urutan dua symbol	Ordinal/urutan dikotomi	Run test

- Uji binomial: Analyze > Nonparametric Tests > Legacy Dialogs > Binomial.
- Chi-kuadrat satu sampel: Analyze > Nonparametric Tests > Legacy Dialogs > Chi-Square.
- Run test: Analyze > Nonparametric Tests > Legacy Dialogs > Runs.
- Pastikan nilai kode kategori dan proporsi acuan didefinisikan dengan benar (Pallant, 2026).

CONTOH PENERAPAN BAB 6

Uji Data Kategorik Satu Sampel dan Pola Urutan

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Bab ini memuat tiga contoh penerapan singkat karena struktur hipotesisnya berbeda: binomial untuk dua kategori, chi-kuadrat goodness-of-fit untuk lebih dari dua kategori, dan run test untuk menilai keacakan urutan.

A. Uji binomial: preferensi penggunaan GeoGebra

Tabel 6. 3. Ilustrasi 6.1. Preferensi fiktif, 1 = memilih GeoGebra; 0 = tidak

ID	Preferensi
1	1
2	1
3	1
4	0
5	1
6	1
7	0

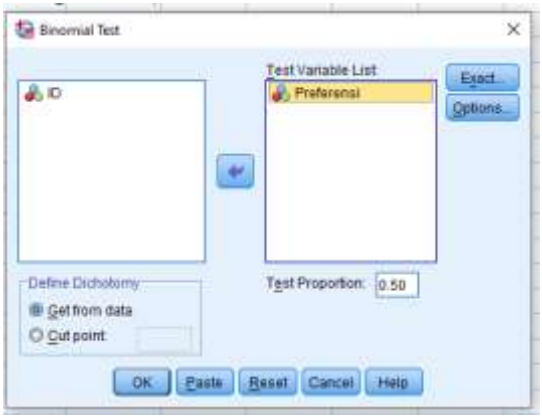
ID	Preferensi
8	1
9	1
10	1
11	0
12	1
13	1
14	0
15	1
16	1
17	0
18	1
19	0
20	1

Rumusan Hipotesis pada Ilustrasi

H_0: p = 0,50 (proporsi mahasiswa yang memilih GeoGebra sama dengan 50%).

H_a: p ≠ 0,50 (proporsi mahasiswa yang memilih GeoGebra tidak sama dengan 50%).

1. Pilih Analyze > Nonparametric Tests > Legacy Dialogs > Binomial.
2. Isikan preferensi ke Test Variable List.



3. Gunakan Test Proportion = 0,50. Pilih Get from data agar SPSS membandingkan kategori 0 dan 1; jangan menggunakan Cut point = 1 karena seluruh nilai 0 dan 1 akan masuk ke satu kelompok.
4. Klik OK keluaran akan muncul sebagai berikut

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
Preferensi	Group 1	memilih Geogebra	14	.70	.50	.115
	Group 2	tidak memilih	6	.30		
	Total		20	1.00		

Cara Membaca Keluaran

Pada dataset ilustratif, sebanyak 14 dari 20 mahasiswa (70%) memilih GeoGebra. Sementara itu, 6 mahasiswa (30%) tidak memilihnya. Uji binomial dua sisi dengan proporsi pembanding 0,50 menghasilkan p = 0,115. Walaupun proporsi sampel yang memilih GeoGebra tampak lebih tinggi, perbedaannya belum signifikan secara statistik.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p > 0,05$, H_0 gagal ditolak dan H_a belum didukung. Data belum memberikan bukti yang cukup bahwa proporsi mahasiswa yang memilih GeoGebra berbeda dari 50%.

B. Chi-kuadrat goodness-of-fit: pilihan media

Tabel 6. 4. Ilustrasi 6.2. Frekuensi pilihan media belajar

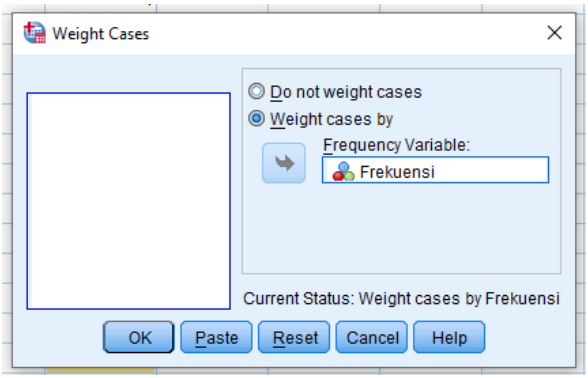
Media	Frekuensi
GeoGebra	12
Video	8
Modul	6
LMS	4

Rumusan Hipotesis pada Ilustrasi

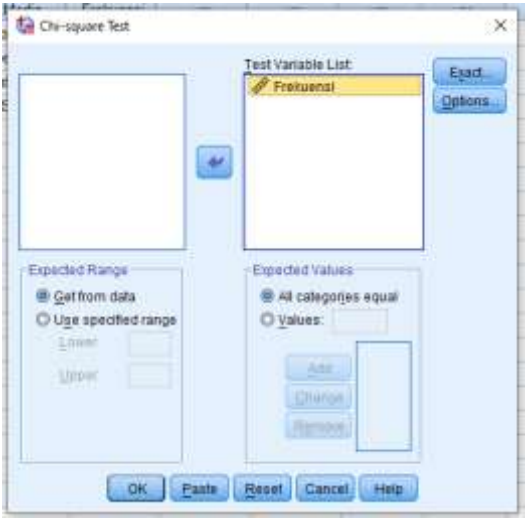
H_0: Proporsi pilihan keempat media belajar sama.

H_a: setidaknya satu proporsi pilihan media belajar berbeda.

1. Tempatkan data dalam bentuk satu baris per mahasiswa, atau gunakan Data > Weight Cases jika memakai tabel frekuensi.



2. Pilih Analyze > Nonparametric Tests > Legacy Dialogs > Chi-Square.
3. Isikan variabel media dan pilih semua kategori memiliki expected values yang sama.



4. Klik OK, keluaran akan muncul sebagai berikut:

Chi-Square Test

Frequencies

Frekuensi

	Observed N	Expected N	Residual
4	4	7.5	-3.5
6	6	7.5	-1.5
8	8	7.5	.5
12	12	7.5	4.5
Total	30		

Test Statistics

	Frekuensi
Chi-Square	4.667[a]
df	3
Asymp. Sig.	.198

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 7.5.

Cara Membaca Keluaran

Pada dataset ilustratif, frekuensi pilihan terhadap empat media belajar berturut-turut adalah 12, 8, 6, dan 4. Sementara itu, frekuensi harapan untuk setiap kategori adalah 7,5. Uji chi-kuadrat goodness-of-fit menghasilkan $\chi^2(3, N = 30) = 4{,}667$ dan $p = 0{,}198$. Perbedaan antara frekuensi observasi dan frekuensi harapan belum cukup besar untuk dinyatakan signifikan.

Arti Keputusan Statistik

Pada ilustrasi ini, karena p > 0,05, H_0 gagal ditolak dan H_a belum didukung. Proporsi pilihan keempat media belajar belum terbukti berbeda secara signifikan dari distribusi yang sama rata.

C. Run test: keacakan urutan nilai

Tabel 6. 5. Ilustrasi 6.3. Urutan fiktif nilai yang akan diuji

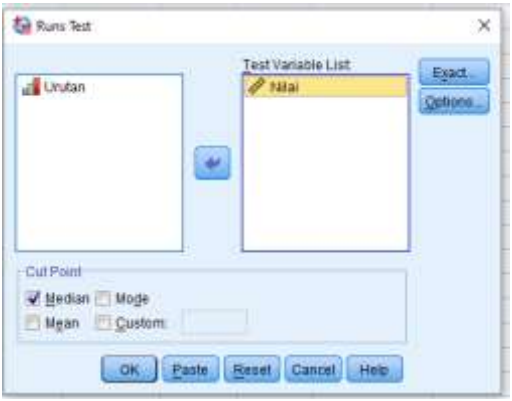
Urutan	Nilai
1	72
2	74
3	75
4	73
5	76
6	78
7	77
8	79
9	81
10	80
11	82
12	83
13	81
14	84
15	86
16	85
17	87
18	88
19	86
20	89

Rumusan Hipotesis pada Ilustrasi

H_0: Urutan nilai bersifat acak.

H_a: Urutan nilai tidak bersifat acak.

1. Pertahankan urutan data sesuai waktu pengumpulan.
2. Pilih Analyze > Nonparametric Tests > Legacy Dialogs > Runs.
3. Isikan nilai ke Test Variable List dan gunakan median sebagai cut point.



6. Klik OK muncul keluaran berikut (baca bagian Asymp. Sig. (2-tailed)).

Runs Test

	Nilai
Test Value[a]	81
Cases < Test Value	9
Cases >= Test Value	11
Total Cases	20
Number of Runs	4
Z	-2.972
Asymp. Sig. (2-tailed)	.003

a. Median

Cara Membaca Keluaran

Pada dataset ilustratif, run test menghasilkan z = -2,972 dan p = 0,003. Urutan data menunjukkan bahwa nilai yang lebih rendah cenderung muncul pada bagian awal. Sementara itu, nilai yang lebih tinggi cenderung muncul

pada bagian berikutnya. Pola tersebut menunjukkan adanya susunan yang sistematis dan tidak konsisten dengan urutan acak.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0,05$, H_0 ditolak dan H_a didukung. Urutan nilai tidak bersifat acak. Peneliti perlu memeriksa kemungkinan tren waktu, urutan pengumpulan, atau prosedur pencatatan yang menyebabkan pola tersebut.

LATIHAN

1. Dari 30 siswa, 21 memilih pembelajaran daring dan 9 memilih luring. Uji apakah proporsi pilihan berbeda dari 50:50!
2. Sebuah sekolah mencatat pilihan empat kegiatan ekstrakurikuler. Susun H_0 untuk goodness-of-fit proporsi sama!
3. Hitung jumlah run pada urutan AABBABAABBBBAA!
4. Jelaskan perbedaan tujuan uji binomial dan chi-kuadrat satu sampel!
5. Tuliskan contoh penyusunan kalimat statistik run test!

Dataset Latihan Penerapan SPSS Bab 6.

Binomial dan run test

ID	Pilihan GeoGebra
1	1
2	0
3	1
4	1
5	1
6	0
7	1
8	0
9	1
10	1
11	1
12	0
13	1

ID	Pilihan GeoGebra
14	1
15	0
16	1
17	0
18	1

Dataset Latihan 6.2. Urutan nilai untuk run test

Urutan	Nilai
1	70
2	72
3	68
4	75
5	71
6	77
7	69
8	78
9	73
10	80
11	74
12	82
13	76
14	84
15	79
16	85
17	81
18	87

Rumusan Hipotesis pada Ilustrasi Uji Binomial

H_0: p = 0,50 (proporsi pilihan GeoGebra sama dengan 50%).

H_a: p ≠ 0,50 (proporsi pilihan GeoGebra tidak sama dengan 50%).

Rumusan Hipotesis pada Ilustrasi Run Test

H_0: Urutan nilai bersifat acak.

H_a: Urutan nilai tidak bersifat acak.

1. Jalankan uji binomial dengan proporsi pembanding 0,50. Kode 1 berarti memilih GeoGebra.
2. Tentukan apakah proporsi pilihan berbeda dari 50%.
3. Susun file terpisah untuk data urutan pada tabel berikut dan jalankan run test dengan median sebagai cut point.
4. Bandingkan tujuan uji binomial dengan run test. Jelaskan mengapa keduanya tidak dapat saling menggantikan.

BAB VII
KOMPARATIF DUA SAMPEL BERPASANGAN: PAIRED T DAN WILCOXON

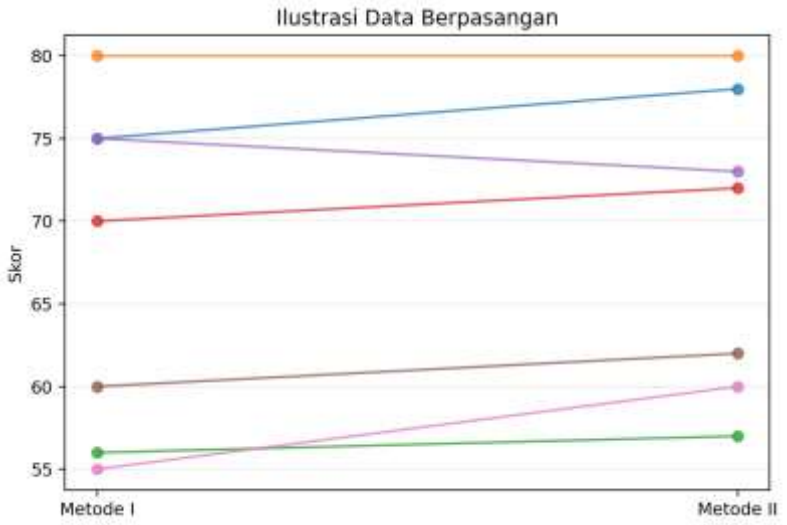
Capaian Pembelajaran
Setelah mempelajari bab ini, pembaca diharapkan mampu: mengidentifikasi desain berpasangan; melakukan paired t-test; melakukan Wilcoxon signed-rank test; menghitung ukuran efek dan menyajikan dan menafsirkan hasil.

Kata Kunci
Sampel berpasangan, paired t-test, skor perbedaan, Wilcoxon, effect size.

A. Desain Berpasangan

Dua sampel berpasangan berasal dari subjek yang sama pada dua waktu atau kondisi, atau dari pasangan yang dicocokkan. Contohnya adalah skor sebelum dan sesudah pelatihan, hasil dua metode pada siswa yang sama, atau pasangan siswa yang disetarakan (Field, 2024; Pallant, 2026).

Satuan yang dikaji utama adalah perbedaan dalam setiap pasangan. Karena itu, memasukkan data berpasangan ke independent t-test akan mengabaikan korelasi dalam pasangan dan dapat menghasilkan galat baku yang tidak tepat.



Gambar 7. 1. . Garis yang menghubungkan skor subjek pada dua kondisi

B. Paired Samples t-Test

$$t = \frac{\bar{d}}{s_d/\sqrt{n}}$$

Statistik paired t-test, dengan d adalah skor perbedaan

Asumsi paired t-test meliputi pasangan yang benar, pasangan independen satu sama lain, variabel berskala interval/rasio, dan distribusi skor perbedaan tidak memiliki penyimpangan berat. Normalitas diperiksa pada d, bukan pada masing-masing kondisi secara terpisah (Field, 2024).

Pada contoh tujuh siswa, skor metode I adalah 75, 80, 56, 70, 75, 60, 55 dan metode II adalah 78, 80, 57, 72, 73, 62, 60. Perbedaan metode II - metode I adalah 3, 0, 1, 2, -2, 2, dan 5. Diperoleh $\bar{d}$ = 1,57, s_d = 2,23, t(6) = 1,87, p = 0,111. Hasil tidak signifikan pada α = 0,05.

Tabel 7. 1. Data berpasangan dan skor perbedaan

Siswa	Metode I	Metode II	d
1	75	78	3
2	80	80	0
3	56	57	1
4	70	72	2
5	75	73	-2
6	60	62	2
7	55	60	5

C. Wilcoxon Signed-Rank Test

Wilcoxon signed-rank digunakan untuk membandingkan dua kondisi berpasangan ketika data ordinal atau distribusi perbedaan tidak mendukung paired t-test. Uji ini merangking nilai absolut perbedaan, memberikan tanda positif atau negatif, kemudian membandingkan jumlah rank bertanda (Wilcoxon, 1945).

Perbedaan nol dikeluarkan dari perankingan. Asumsi pentingnya adalah pasangan yang benar, pasangan independen, dan distribusi perbedaan yang kurang lebih simetris bila uji ditafsirkan sebagai perbedaan median. Untuk data contoh, hasil exact Wilcoxon W = 3, p = 0,156 (Field, 2024).

D. Ukuran Efek

$$d_z = \frac{\bar{d}}{s_d}$$

Cohen dz untuk data berpasangan

$$r = \frac{Z}{\sqrt{N}}$$

Ukuran efek untuk Wilcoxon

Ukuran efek paired t dapat dihitung dari mean dan simpangan baku perbedaan. Untuk contoh, dz sekitar 0,71, yang menunjukkan perbedaan terstandar cukup besar meskipun p tidak signifikan. Hal ini dapat terjadi karena sampel sangat kecil sehingga interval kepercayaan lebar (Cohen, 1988; Lakens, 2013).

Ukuran efek harus dilaporkan bersama arah perubahan dan statistik deskriptif. Perbedaan praktis juga perlu ditinjau terhadap batas yang bermakna secara pendidikan (Appelbaum *et al*., 2018).

E. Langkah SPSS dan Penyajian Hasil

1. Paired t: Analyze > Compare Means > Paired-Samples T Test; masukkan pasangan variabel.
2. Wilcoxon: Analyze > Nonparametric Tests > Legacy Dialogs > 2 Related Samples; centang Wilcoxon.
3. Cermati tabel Paired Samples Statistics, Paired Samples Correlations, dan Paired Samples Test.
5. Untuk Wilcoxon, baca positive ranks, negative ranks, ties, Z, dan Asymp. Sig./Exact Sig (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Skor setelah metode II (M = 68,86; SD = 9,14) lebih tinggi daripada metode I (M = 67,29; SD = 10,16), tetapi perbedaannya tidak signifikan, t(6) = 1,87, p = 0,111, dz = 0,71. Karena sampel kecil, interval kepercayaan dan relevansi praktis perlu dipertimbangkan.

CONTOH PENERAPAN BAB 7

Membandingkan Pretest dan Posttest pada Peserta yang Sama

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: kemampuan pemecahan masalah 20 mahasiswa diukur sebelum dan sesudah penggunaan lembar kerja berbasis masalah.

Tabel 7. 2. Ilustrasi 7.1. Data fiktif pretest dan posttest

ID	Pretest	Posttest
1	62	65
2	65	73
3	70	72
4	68	75
5	72	77
6	66	76
7	74	75
8	69	75
9	71	75
10	67	76
11	75	78
12	64	71

ID	Pretest	Posttest
13	73	75
14	70	78
15	68	73
16	72	83
17	66	67
18	69	75
19	71	75
20	65	74

Rumusan Hipotesis pada Ilustrasi

H_0: μd = 0 (tidak ada perbedaan rata-rata antara skor pretest dan posttest).

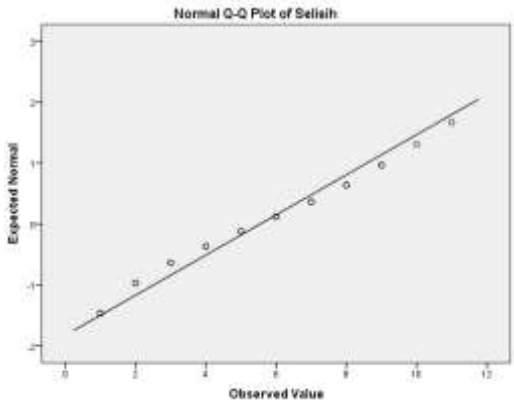
H_a: μd ≠ 0 (ada perbedaan rata-rata antara skor pretest dan posttest).

Langkah Mengolah dan Mengkaji Data

1. Buat variabel pretest dan posttest pada kolom berbeda untuk peserta yang sama.
2. Hitung selisih melalui Transform > Compute Variable, misalnya selisih = posttest - pretest.



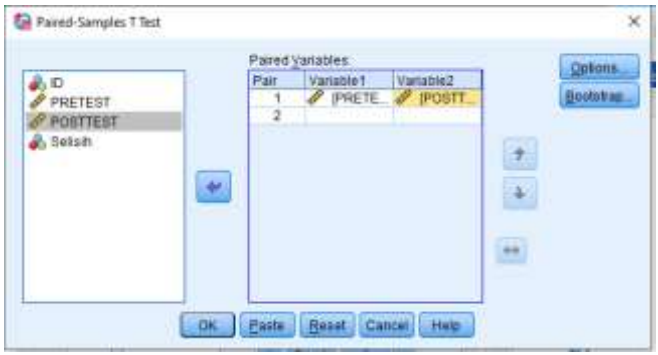
3. Periksa normalitas selisih menggunakan Explore.



Dari grafik Normal Q-Q plot, terlihat bahwa titik-titik data menyebar di sekitar dan mengikuti garis diagonal. Jadi, secara visual data selisih memenuhi asumsi normalitas atau berdistribusi mendekati normal.

Catatan: Jika selisih sangat tidak normal atau mengandung pencilan berat, gunakan Analyze > Nonparametric Tests > Legacy Dialogs > 2 Related Samples > Wilcoxon.

4. Buka Analyze > Compare Means > Paired-Samples T Test.
5. Pasangkan pretest dengan posttest, lalu klik OK.



6. Klik OK, keluaran akan muncul sebagai berikut:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PRETEST	68.85	20	3.514	.786
	POSTTEST	74.40	20	3.831	.857

Rata-rata nilai pretest adalah 68,85. Sementara itu, rata-rata posttest menjadi 74,40. Artinya, setelah perlakuan atau pembelajaran diberikan, nilai rata-rata meningkat sebesar:

$$74{,}40 - 68{,}85 = 5{,}55$$

Jadi, secara deskriptif ada peningkatan nilai sebesar 5,55 poin.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	PRETEST & POSTTEST	20	.662	.001

Nilai korelasi antara pretest dan posttest sebesar 0,662 dengan nilai signifikansi 0,001.

Kondisi ini menunjukkan bahwa nilai pretest dan posttest mempunyai hubungan positif yang cukup kuat dan signifikan. Responden yang memiliki nilai pretest lebih tinggi cenderung juga memiliki nilai posttest yang lebih tinggi.

Akan tetapi, tabel korelasi ini bukan dasar utama untuk menentukan ada atau tidaknya perbedaan. Keputusan hipotesis tetap mengacu pada tabel Paired Samples Test.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PRETEST - POSTTEST	-5.550	3.034	.679	-6.973	-4.130	-8.180	19	.000

Pretest−Posttest

Karena nilai posttest lebih tinggi daripada pretest, hasil pengurangannya menjadi −5,550. Tanda negatif tidak menunjukkan adanya penurunan. Justru, hasil tersebut menunjukkan bahwa nilai posttest meningkat sebesar 5,55 poin dibandingkan pretest.

Nilai Sig. (2-tailed) sebesar 0,000 harus ditulis sebagai $p < 0{,}001$, bukan $p = 0{,}000$. Karena $p < 0{,}05$, maka:

Karena $p < 0{,}05$, H_0 ditolak.

Cara membacanya, ada perbedaan yang signifikan antara nilai pretest dan posttest.

Cara Membaca Keluaran

Pada dataset ilustratif, skor posttest ($M = 74{,}40$; $SD = 3{,}83$) lebih tinggi daripada skor pretest ($M = 68{,}85$; $SD = 3{,}51$), dengan peningkatan rata-rata 5,55 poin. Uji t berpasangan menghasilkan $t(19) = 8{,}18$, $p < 0{,}001$, dan Cohen's $dz = 1{,}83$. Uji Shapiro-Wilk terhadap skor selisih menghasilkan $W = 0{,}960$ dan $p = 0{,}553$, sehingga asumsi normalitas skor selisih terpenuhi.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Skor setelah pembelajaran secara signifikan lebih tinggi daripada skor sebelum pembelajaran. Peningkatan tersebut memiliki ukuran efek yang sangat besar, sehingga bermakna secara statistik dan praktis pada dataset ilustratif ini.

LATIHAN

1. Paparkan mengapa pretest dan posttest pada siswa yang sama tidak boleh dikaji dengan independent t-test!
2. Hitung skor perbedaan dan mean perbedaan pada data latihan berpasangan di Lampiran C!
3. Kapan Wilcoxon lebih sesuai daripada paired t-test!
4. Paparkan peran ties dan perbedaan nol pada Wilcoxon!
5. Susun penyajian hasil paired t-test lengkap dengan ukuran efek!

Dataset Latihan Penerapan SPSS

Paired t-test atau Wilcoxon

ID	Pretest	Posttest
1	60	68
2	63	70
3	67	73
4	70	77
5	65	72
6	72	78
7	68	75
8	66	71
9	74	80
10	69	74
11	71	78
12	64	70
13	73	79
14	62	69

ID	Pretest	Posttest
15	68	74
16	70	76
17	66	72
18	75	82

Rumusan Hipotesis pada Ilustrasi

H_0: μd = 0 (tidak ada perbedaan rata-rata antara pretest dan posttest).
H_a: μd ≠ 0 (ada perbedaan rata-rata antara pretest dan posttest).

1. Hitung variabel selisih = posttest - pretest.
2. Periksa normalitas selisih dan pencilan.
3. Gunakan prosedur paired samples t-test jika asumsi layak. Jalankan Wilcoxon sebagai pengujian pembanding.
4. Catat mean pretest, mean posttest, mean selisih, statistik uji, p-value, dan ukuran efek.
5. Tulis kesimpulan yang membedakan signifikansi statistik dan besarnya perubahan.

BAB VIII
KOMPARATIF DUA SAMPEL INDEPENDEN: HOMOGENITAS, UJI T, WELCH, DAN MANN-WHITNEY

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: membedakan sampel independen dan berpasangan; menilai homogenitas varians; memilih pooled t, Welch, atau Mann-Whitney; melaporkan perbandingan dua kelompok.

Kata Kunci

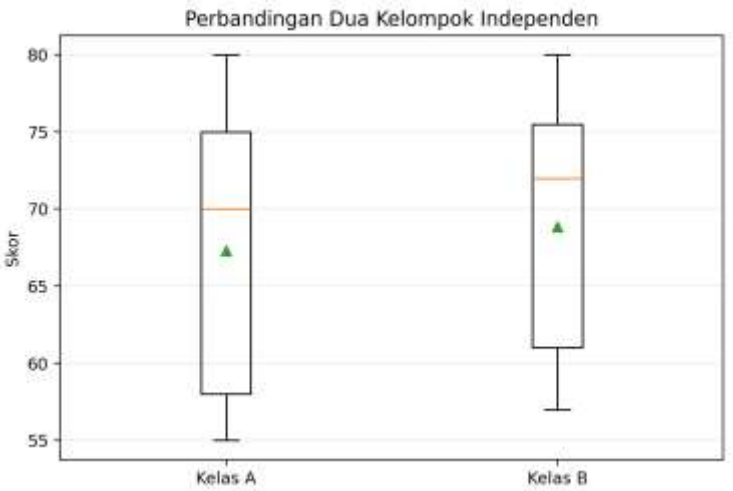
independen, homogenitas, Levene, independent t-test, Welch, Mann-Whitney.

A. Desain Dua Kelompok Independen

Dua sampel independen terdiri dari subjek berbeda dan keanggotaan satu kelompok tidak menentukan keanggotaan kelompok lain. Contohnya adalah kelas A menggunakan metode I dan kelas B menggunakan metode II. Observasi dalam dan antar-kelompok harus independen (Field, 2024; Pallant, 2026).

Tujuan pengkajian biasanya membandingkan parameter lokasi dua populasi. Independent-samples t-test berfokus pada mean. Sementara itu, Mann-Whitney berfokus pada distribusi rank dan hanya dapat ditafsirkan sebagai

perbedaan median bila bentuk distribusi kedua kelompok serupa.



Gambar 8. 1. Boxplot dua kelompok independent

B. Homogenitas Varians

Homogenitas berarti varians populasi dua kelompok dianggap sama. Pemeriksaan dapat dilakukan melalui perbandingan varians, grafik, atau uji Levene. Uji F sederhana sensitif terhadap ketidaknormalan, sehingga uji Levene lebih sering digunakan dalam perangkat lunak.

Jika p Levene > 0,05, baris Equal variances assumed dapat digunakan. Jika p Levene ≤ 0,05, gunakan baris Equal variances not assumed atau uji Welch. Keputusan tidak seharusnya hanya bergantung pada Levene; ukuran sampel yang sangat tidak seimbang dan pencilan perlu diperhatikan (Field, 2024).

C. Independent Samples t-Test

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

Uji t dengan varians gabungan

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

Pada contoh ini, kelas A memiliki n = 7, M = 67,29, s^2 = 103,24 dan kelas B memiliki n = 7, M = 68,86, s^2 = 83,48. Varians relatif serupa. Independent-samples t-test menghasilkan t(12) = -0,304, p = 0,766. Tidak ada bukti perbedaan mean yang signifikan.

Ketika varians tidak homogen, uji Welch menggunakan galat baku berdasarkan varians masing-masing dan derajat kebebasan yang disesuaikan. Welch umumnya lebih aman daripada memaksakan varians gabungan (Welch, 1947).

D. Mann-Whitney U

Mann-Whitney U digunakan untuk dua kelompok independen dengan data ordinal atau ketika distribusi tidak mendukung uji t. Semua nilai digabung, dirangking, lalu jumlah rank kelompok dihitung. Nilai yang sama memperoleh rank rata-rata (Mann & Whitney, 1947).

Pada contoh ini, kelompok A berjumlah 8 dan kelompok B berjumlah 9. SPSS menghasilkan U = 23,5 dan p asimtotik sekitar 0,225. Hasil tidak signifikan. Perbedaan antara p exact dan asimtotik dapat muncul pada sampel kecil atau banyak ties; metode yang digunakan perlu disebutkan.

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

E. Ukuran Efek dan Penyajian Hasil

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s_p}$$

Cohen d dua kelompok

Untuk independent t-test, Cohen d atau Hedges g dapat digunakan. Hedges g memberikan koreksi bias pada sampel kecil. Untuk Mann-Whitney, ukuran efek r = Z/√N atau rank-biserial correlation dapat dilaporkan (Lakens, 2013).

Selain mean dan simpangan baku, tampilkan jumlah sampel. Untuk data menceng, median dan IQR dapat lebih informatif. Jangan menyimpulkan "metode tidak berpengaruh" hanya karena p > 0,05; nyatakan bahwa data belum menunjukkan perbedaan yang cukup meyakinkan (Appelbaum *et al*., 2018).

F. Langkah SPSS

1. Siapkan variabel nilai dan variabel grup; definisikan kode grup.
2. Independent t: Analyze > Compare Means > Independent-Samples T Test.
3. Isikan nilai ke Test Variable(s), grup ke Grouping Variable, lalu Define Groups.
4. Baca Levene dan pilih baris equal variances assumed/not assumed.
5. Mann-Whitney: Analyze > Nonparametric Tests > Legacy Dialogs > 2 Independent Samples (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Skor kelas A (M = 67,29; SD = 10,16) tidak berbeda signifikan dari kelas B (M = 68,86; SD = 9,14), t(12) = -0,30, p = 0,766, d = -0,16. Selisih rata-rata kecil dan interval kepercayaan mencakup nol.

CONTOH PENERAPAN BAB 8

Membandingkan Dua Kelompok Independen

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: nilai kemampuan representasi matematis dibandingkan antara kelas Problem-Based Learning dan kelas pembelajaran konvensional. Setiap siswa hanya berada pada satu kelompok.

Tabel 8. 1. Ilustrasi 8.1. Data fiktif dua kelompok independen, 1 = PBL; 2 = Konvensional

ID	Kelompok	Nilai
1	1	78
2	1	74
3	1	82
4	1	77
5	1	84
6	1	73
7	1	81
8	1	79
9	1	76
10	1	83
11	1	80

ID	Kelompok	Nilai
12	1	75
13	1	82
14	1	78
15	1	77
16	2	74
17	2	72
18	2	78
19	2	70
20	2	76
21	2	75
22	2	73
23	2	77
24	2	71
25	2	79
26	2	74
27	2	72
28	2	76
29	2	75
30	2	73

Rumusan Hipotesis pada Ilustrasi

H_0: μPBL = μKonvensional (rata-rata kedua kelompok sama).

H_a: μPBL ≠ μKonvensional (rata-rata kedua kelompok berbeda).

Langkah Mengolah dan Mengkaji Data

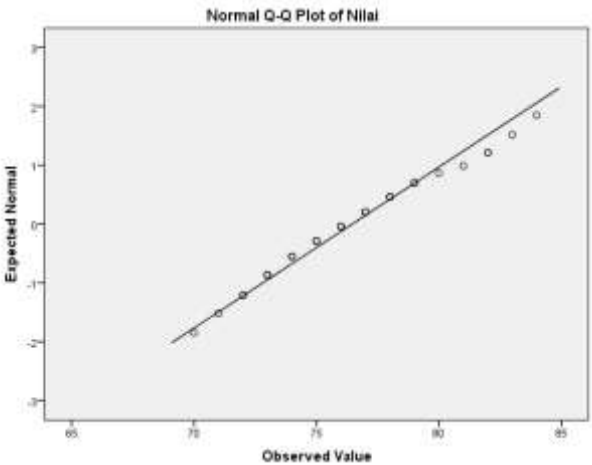
1. Beri value labels pada kelompok: 1 = PBL dan 2 = Konvensional.
2. Tinjau boxplot dan normalitas nilai per kelompok melalui Explore.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Nilai	.089	30	.200*	.974	30	.646

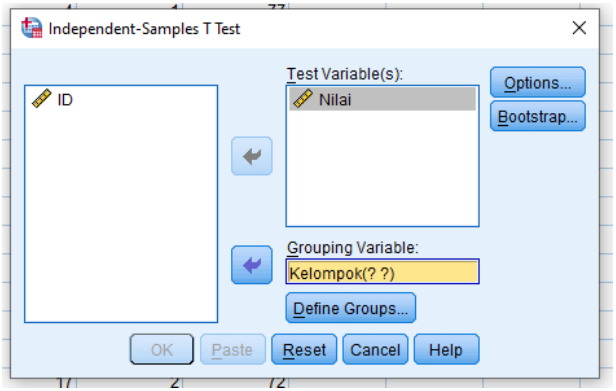
*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

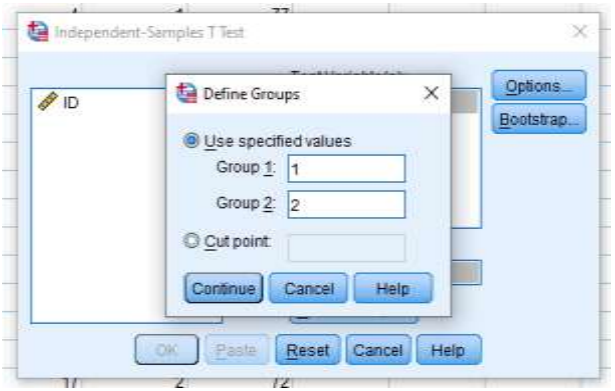


Pembacaan terhadap uji Shapiro-Wilk menunjukkan bahwa diperoleh nilai signifikansi sebesar 0,646. Karena nilai tersebut lebih besar dari 0,05, maka data dinyatakan berdistribusi normal. Hasil ini juga didukung oleh Normal Q-Q plot yang menunjukkan bahwa titik-titik data menyebar di sekitar dan mengikuti garis diagonal.

3. Buka Analyze > Compare Means > Independent-Samples T Test.
4. Isikan nilai sebagai Test Variable dan kelompok sebagai Grouping Variable.



5. Klik Define Groups, masukkan 1 dan 2.



6. Klik ok, keluaran akan muncul sebagai berikut:

Catatan: Baca Levene Test terlebih dahulu. Jika p > 0,05, gunakan baris Equal variances assumed. Jika p <= 0,05, gunakan baris Equal variances not assumed atau Welch.

T-Test

Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
Nilai	PBL	15	78.60	3.355	.866
	Konvensional	15	74.33	2.582	.667

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Nilai	Equal variances assumed	1.403	.246	3.883	28	.001	4.267	1.099	2.026	6.508
	Equal variances not assumed			3.883	26.276	.001	4.267	1.099	2.021	6.512

- Levene's Test Sig. = 0,295.
- Karena 0,295 > 0,05, varians kelompok PBL dan konvensional dinyatakan homogen.
- Karena itu, hasil uji t dibaca pada baris Equal variances assumed.

Cara Membaca Keluaran

Pada dataset ilustratif, kelompok PBL memiliki rata-rata nilai 78,60 (SD = 3,36). Sementara itu, kelompok konvensional memiliki rata-rata 74,33 (SD = 2,58). Uji Levene menghasilkan p = 0,295, sehingga asumsi homogenitas varians terpenuhi dan hasil dibaca pada baris Equal variances assumed. Uji t independen menghasilkan $t(28) = 3,90$, $p < 0,001$, dengan selisih rata-rata 4,27 poin, interval kepercayaan 95% [2,03, 6,51], dan Cohen's d = 1,43.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0,05$, H_0 ditolak dan H_a didukung. Rata-rata nilai kelompok PBL secara signifikan lebih tinggi daripada kelompok konvensional. Besarnya perbedaan tergolong sangat besar berdasarkan Cohen's d.

LATIHAN

1. Bedakan desain independen dan berpasangan dengan dua contoh!
2. Jelaskan kapan memilih baris Equal variances not assumed pada SPSS!
3. Hitung rasio varians untuk contoh kelas A dan B!
4. Mengapa Mann-Whitney tidak selalu dapat disebut "uji perbedaan median"?
5. Susun penyajian hasil independent t-test yang tidak signifikan tanpa menyatakan H_0 terbukti!

Dataset Latihan Penerapan SPSS

Independent-samples t-test, Welch, atau Mann-Whitney

ID	Kelompok	Nilai
1	1	76
2	1	80
3	1	78
4	1	84
5	1	75
6	1	82
7	1	79
8	1	81
9	1	77
10	1	83
11	1	74
12	1	85
13	2	72
14	2	75
15	2	70

ID	Kelompok	Nilai
16	2	78
17	2	74
18	2	76
19	2	73
20	2	77
21	2	71
22	2	79
23	2	75
24	2	74

Rumusan Hipotesis pada Ilustrasi

H_0: μDiscovery = μEkspositori (rata-rata kedua kelompok sama).

H_a: μDiscovery ≠ μEkspositori (rata-rata kedua kelompok berbeda).

1. Beri label kelompok 1 = Discovery Learning dan 2 = Ekspositori.
2. Tinjau distribusi, pencilan, dan homogenitas varians.
3. Lakukan independent samples t-test dan tentukan baris hasil yang tepat.
4. Jalankan Mann-Whitney U sebagai pembanding, lalu jelaskan perbedaan fokus penjelasan maknanya.
5. Hitung Cohen's d dan susun penyajian hasil lengkap.

BAB IX
UJI CHI-KUADRAT, MEDIAN TEST, DAN FISHER EXACT

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: menyusun tabel kontingensi; melakukan chi-kuadrat independensi; menjalankan Median Test untuk dua atau lebih kelompok independen; memeriksa frekuensi harapan; memilih Fisher exact; serta menghitung dan menafsirkan ukuran asosiasi.

Kata Kunci

tabel kontingensi, chi-kuadrat, independensi, median gabungan, Median Test, expected count, Fisher exact, phi, Cramer V.

A. Tabel Kontingensi dan Hipotesis

Uji chi-kuadrat independensi digunakan untuk menilai hubungan antara dua variabel kategorik atau perbedaan proporsi antar-kelompok. Data disusun dalam tabel kontingensi. H_0 menyatakan kedua variabel independen atau proporsi kategori sama antar-kelompok; H_a menyatakan ada asosiasi (Field, 2024; Triola, 2022).

Frekuensi harapan setiap sel dihitung dari total baris dan kolom. Uji membandingkan frekuensi observasi dengan frekuensi yang diharapkan bila H_0 benar.

E_{ij} = (total baris i × total kolom j) / total keseluruhan

$X^2 = \Sigma[(O_{ij} - E_{ij})^2 / E_{ij}]$

B. Contoh Ketuntasan Dua Kelas

Tabel 9.1. Ketuntasan belajar klasikal dua kelas

Kelas	Tuntas	Tidak tuntas	Jumlah
A	29	8	37
B	20	18	38
Jumlah	49	26	75

Untuk tabel 2 × 2 tersebut, chi-kuadrat tanpa koreksi sekitar 5,49 dengan p = 0,019. Dengan koreksi kontinuitas Yates, χ^2 = 4,41 dan p = 0,036. Pada contoh ini digunakan nilai terkoreksi 4,41. Kedua pendekatan pada α = 0,05 menghasilkan kesimpulan ada perbedaan proporsi ketuntasan.

Koreksi Yates dirancang untuk pendekatan kontinu terhadap data diskrit pada tabel 2 × 2, tetapi dapat terlalu konservatif. Banyak laporan modern menampilkan Pearson chi-square tanpa koreksi dan Fisher exact ketika frekuensi harapan kecil. Metode yang dipakai harus disebutkan (Fisher, 1922; Field, 2024).

C. Median Test untuk Sampel Independen

Median Test berfungsi untuk menguji apakah dua atau lebih sampel independen berasal dari populasi yang memiliki median yang sama. Prosedur ini sesuai untuk variabel ordinal atau numerik ketika fokus penelitian terletak pada median dan observasi antar-kelompok bersifat independen. Hipotesis nol menyatakan median seluruh populasi sama. Sementara itu, hipotesis alternatif menyatakan setidaknya ada satu median yang berbeda (Field, 2024; IBM, 2026).

Semua skor digabung untuk mencari median gabungan. Setelah itu, setiap skor dimasukkan ke kategori kurang dari atau sama dengan median, atau lebih besar dari median. Frekuensi kedua kategori pada setiap kelompok kemudian diuji dengan chi-kuadrat. Karena skor asli diubah menjadi dua kategori, sebagian informasi urutan hilang. Karena itu, Median Test biasanya kurang kuat daripada Kruskal-Wallis. Uji ini tetap berguna jika tujuan utama adalah membandingkan median.

Cramer's V berfungsi untuk mengukur seberapa kuat hubungan antara metode pembelajaran dan posisi skor terhadap median.

Pada Median Test, skor sudah diubah menjadi dua kategori:

- skor > median
- skor ≤ median

Kemudian kategori tersebut dibandingkan dengan kelompok metode, misalnya:

- PBL
- langsung
- proyek

Jadi, tabelnya berbentuk 2 × 3. Karena datanya sudah berbentuk kategorik, ukuran kekuatan hubungan yang sesuai adalah Cramer's V. Nilai signifikansi hanya menjawab ada perbedaan/hubungan yang signifikan atau tidak.

Perhitungan Cramer's V

$$V = \sqrt{\frac{x^2}{N(k-1)}}$$

Keterangan:

V = nilai Cramer's V

x^2= nilai Pearson Chi-Square

N = jumlah seluruh sampel atau kasus yang dikaji

k = jumlah katgori paling sedikit antara baris dan kolom (k = min(r, c) dengan r adalah jumlah baris dan c = jumlah kolom)

k – 1 = derajat penyesuaian berdasarkan dimensi tabel

Ilustrasi kasus: seorang dosen membandingkan skor kemampuan pemecahan masalah mahasiswa yang belajar melalui *Problem-Based Learning* (PBL), pembelajaran langsung, dan pembelajaran berbasis proyek. Setiap kelompok terdiri atas 12 mahasiswa dan tidak ada mahasiswa yang mengikuti lebih dari satu metode.

Tabel 9. 1. Data fiktif Median Test tiga metode pembelajaran

Metode	Skor fiktif	Skor <= 75,50	Skor > 75,50
PBL	72, 74, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85	2	10
Langsung	60, 62, 64, 66, 68, 70, 71, 72, 73, 74, 76, 77	10	2
Proyek	69, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81	6	6

Median gabungan seluruh skor adalah 75,50. Tabel kategorisasi menghasilkan chi-kuadrat sebesar 10,67 dengan df = 2 dan p = 0,005. Karena p < 0,05, hipotesis nol ditolak. Ada bukti bahwa median ketiga populasi tidak seluruhnya sama. Nilai Cramer V yaitu:

$$V = \sqrt{\frac{10{,}67}{36(2-1)}} = \sqrt{\frac{10{,}67}{36}} = \sqrt{0{,}2964} = 0{,}544$$

Nilai Cramer's V sebesar 0,54 menunjukkan perbedaan yang kuat pada data contoh. Dari 12 skor pada setiap kelompok, PBL memiliki 10 skor di atas median gabungan, pembelajaran langsung 2 skor, dan pembelajaran proyek 6 skor.

Hasil signifikan tidak langsung menunjukkan pasangan kelompok mana yang berbeda. Lakukan perbandingan berpasangan yang direncanakan dan terapkan koreksi pengujian multipel, misalnya Holm, bila pengujian lanjutan diperlukan.

Langkah Mengolah Data dengan Median Test

1. Susun satu variabel skor dan satu variabel kelompok. Setiap baris mewakili satu peserta.
2. Buka Analyze > Nonparametric Tests > Legacy Dialogs > K Independent Samples.
3. Tempatkan skor ke Test Variable List dan metode ke Grouping Variable.
4. Klik Define Range, lalu masukkan kode minimum 1 dan maksimum 3.
5. Pada Test Type, aktifkan Median. Gunakan median sampel gabungan sebagai nilai pemisah.
6. Baca tabel Frequencies dan Test Statistics. Laporkan median gabungan, chi-square, df, p-value, serta ukuran efek yang dihitung dari tabel 2 x k.

D. Asumsi dan Frekuensi Harapan

- Data berupa frekuensi, bukan persentase atau skor rata-rata.
- Kategori saling lepas dan setiap unit hanya dihitung sekali.
- Observasi independen.
- Frekuensi harapan tidak terlalu kecil; pedoman umum: tidak ada expected count < 1 dan tidak lebih dari 20% sel < 5.

- Jika tabel 2 × 2 memiliki expected count kecil, gunakan Fisher exact (Field, 2024).
-

E. Fisher Exact Test

Fisher exact menghitung probabilitas exact semua tabel yang konsisten dengan margin yang sama dan setidaknya se-ekstrem tabel observasi. Uji ini sangat sesuai untuk sampel kecil tabel 2 × 2 dan tidak bergantung pada pendekatan distribusi chi-kuadrat (Fisher, 1922).

Pada contoh 8 birokrat dan 7 akademisi dengan pilihan mobil gelap-terang, tabelnya adalah 5, 3, 2, dan 5. Fisher exact dua sisi menghasilkan p sekitar 0,315 sehingga belum ada bukti perbedaan preferensi warna.

Tabel 9. 2. Contoh data untuk Fisher exact

Kelompok	Gelap	Terang	Jumlah
Birokrat	5	3	8
Akademisi	2	5	7
Jumlah	7	8	15

F. Ukuran Efek

Koefisien phi untuk tabel 2 × 2

$\Phi = \sqrt{(\chi^2 / N)}$

Cramer's V untuk tabel lebih besar

$$V = \sqrt{\frac{x^2}{N(k-1)}}$$

Phi digunakan pada tabel 2 × 2 dan Cramer V pada tabel lebih besar. Odds ratio juga informatif untuk tabel 2 × 2.

Ukuran efek menunjukkan kekuatan asosiasi, sedangkan p-value dipengaruhi ukuran sampel (Field, 2024).

G. SPSS dan Penyajian Hasil

1. Buka Analyze > Descriptive Statistics > Crosstabs.
2. Tempatkan satu variabel ke Row dan variabel lain ke Column.
3. Buka Statistics: Chi-square dan Phi and Cramer's V.
4. Pilih Cells: Observed, Expected, Row %, dan Column %.
5. Pilih Exact untuk Fisher bila tersedia.
6. Periksa catatan expected count sebelum memilih hasil (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Terdapat asosiasi antara kelas dan ketuntasan, $\chi^2(1, N = 75) = 5{,}49$, $p = 0{,}019$, $\varphi = 0{,}27$. Proporsi tuntas lebih tinggi di kelas A (78,4%) daripada kelas B (52,6%).

CONTOH PENERAPAN BAB 9

Contoh A: Menguji Hubungan Metode Pembelajaran dan Ketuntasan

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: peneliti ingin mengetahui apakah ketuntasan belajar berkaitan dengan metode pembelajaran. Data disusun satu baris untuk setiap siswa.

Tabel 9. 3. Ilustrasi 9.1. Data fiktif, metode 1 = PBL; 2 = Konvensional, ketuntasan 1 = Tuntas; 0 = Belum

ID	Metode	Ketuntasan
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	1
13	1	1
14	1	1
15	1	1
16	1	1
17	1	0
18	1	0
19	1	0
20	1	0
21	2	1
22	2	1
23	2	1
24	2	1

ID	Metode	Ketuntasan
25	2	1
26	2	1
27	2	1
28	2	1
29	2	1
30	2	1
31	2	0
32	2	0
33	2	0
34	2	0
35	2	0
36	2	0
37	2	0
38	2	0
39	2	0
40	2	0

Rumusan Hipotesis pada Ilustrasi

H_0: Metode pembelajaran dan ketuntasan belajar saling independen atau tidak berhubungan.

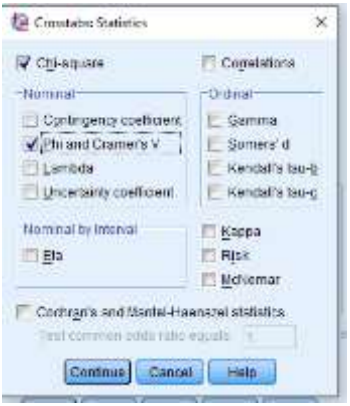
H_a: Metode pembelajaran dan ketuntasan belajar tidak independen atau berhubungan.

Langkah Mengolah dan Mengkaji Data

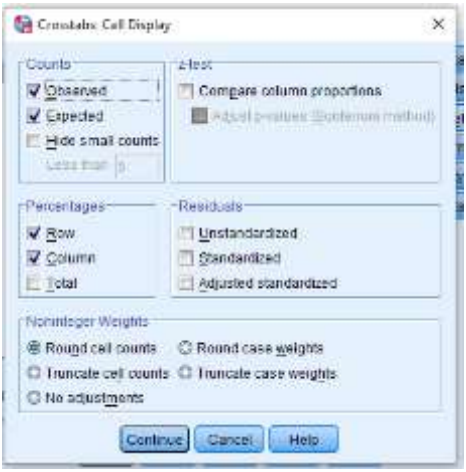
1. Beri value labels pada metode dan ketuntasan.
2. Akses Analyze > Descriptive Statistics > Crosstabs.
3. Masukkan metode ke Rows dan ketuntasan ke Columns.



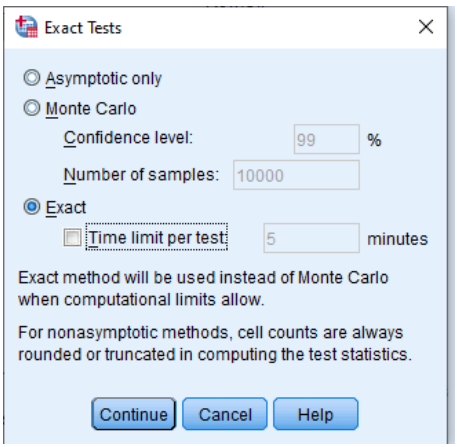
4. Pilih Statistics, aktifkan Chi-square serta Phi and Cramer's V.



5. Klik Cells, aktifkan Observed, Expected, Row percentages, dan Column percentages.



6. Pilih Exact jika tersedia dan Fisher exact diperlukan.



7. Klik OK. Keluaran akan muncul sebagai berikut:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Metode * Ketuntasan	40	100.0%	0	0.0%	40	100.0%

Pembacaan terhadap tabel Case Processing Summary menunjukkan bahwa seluruh data sebanyak 40 responden

dinyatakan valid dan dapat dikaji. Tidak ada data yang hilang atau missing.

Metode * Ketuntasan Crosstabulation

			Ketuntasan		Total
			Belum Tuntas	Tuntas	
Metode	PBL	Count	4	16	20
		Expected Count	7.0	13.0	20.0
		% within Metode	20.0%	80.0%	100.0%
		% within Ketuntasan	28.6%	61.5%	50.0%
	Konvensional	Count	10	10	20
		Expected Count	7.0	13.0	20.0
		% within Metode	50.0%	50.0%	100.0%
		% within Ketuntasan	71.4%	38.5%	50.0%
Total		Count	14	26	40
		Expected Count	14.0	26.0	40.0
		% within Metode	35.0%	65.0%	100.0%
		% within Ketuntasan	100.0%	100.0%	100.0%

Tabel crosstabulation menunjukkan bahwa 16 dari 20 peserta kelompok PBL (80%) tuntas, sedangkan 4 peserta (20%) belum tuntas. Pada kelompok konvensional, 10 dari 20 peserta (50%) tuntas dan 10 peserta (50%) belum tuntas. Jadi, persentase ketuntasan kelompok PBL lebih tinggi 30 poin persentase. Uji chi-kuadrat diperlukan untuk menilai apakah perbedaan tersebut cukup kuat secara statistik.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3.956[a]	1	.047	.096	.048	
Continuity Correction[b]	2.747	1	.097			
Likelihood Ratio	4.054	1	.044	.096	.048	
Fisher's Exact Test				.096	.048	
Linear-by-Linear Association	3.857[c]	1	.050	.096	.048	.039
N of Valid Cases	40					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.00.

b. Computed only for a 2x2 table

c. The standardized statistic is -1.964.

Pada keluaran contoh Pearson Chi-Square diperoleh nilai χ^2 sebesar 3,956 dengan df = 1 dan nilai signifikansi sebesar 0,047. Seluruh sel memiliki expected count lebih dari 5, dengan nilai minimum expected count sebesar 7,00, sehingga asumsi uji Chi-Square telah terpenuhi. Karena nilai signifikansi lebih kecil dari 0,05, maka H_0 ditolak. Karena itu, ada hubungan yang signifikan antara metode pembelajaran dan ketuntasan belajar.

Symmetric Measures

		Value	Approx. Sig.	Exact Sig.
Nominal by Nominal	Phi	-.314	.047	.096
	Cramer's V	.314	.047	.096
N of Valid Cases		40		

Keluaran Pearson Chi-Square menunjukkan $\chi^2(1) = 3{,}956$ dan $p = 0{,}047$. Karena p lebih kecil dari 0,05, ada hubungan yang signifikan antara metode pembelajaran dan ketuntasan. Cramer's V = 0,314 menunjukkan kekuatan hubungan sedang. Ketuntasan pada kelompok PBL adalah 80%, sedangkan pada kelompok konvensional 50%. Hasil ini menunjukkan hubungan, tetapi tidak langsung membuktikan bahwa metode pembelajaran menjadi satu-satunya penyebab ketuntasan.

Cara Membaca Keluaran

Pada dataset ilustratif, proporsi ketuntasan pada kelompok PBL adalah 80%. Sementara itu, pada kelompok konvensional 50%. Uji Pearson chi-square menghasilkan $\chi^2(1, N = 40) = 3{,}956$ dan $p = 0{,}047$. Seluruh expected count memenuhi syarat dengan nilai minimum 7,00. Koefisien phi sebesar 0,314 menunjukkan kekuatan hubungan sedang.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Metode pembelajaran dan ketuntasan belajar

memiliki hubungan yang signifikan dengan kekuatan sedang. Persentase ketuntasan lebih tinggi pada kelompok PBL, tetapi hasil ini menunjukkan asosiasi dan tidak membuktikan hubungan sebab-akibat.

Untuk tabel 2 x 2 dengan expected count kecil, prioritaskan Fisher exact. Jangan memilih Fisher hanya karena p Pearson tidak signifikan.

Contoh B: Median Test untuk Tiga Metode Pembelajaran

Skenario penelitian: peneliti ingin mengetahui apakah median skor kemampuan pemecahan masalah berbeda pada mahasiswa yang belajar menggunakan PBL, pembelajaran langsung, dan pembelajaran berbasis proyek. Data disusun dalam format panjang, yaitu satu baris untuk setiap mahasiswa.

Tabel 9. 4. Ilustrasi 9.2. Data fiktif Median Test, metode 1 = PBL; 2 = Langsung; 3 = Proyek

ID	Metode	Skor
1	1	72
2	1	74
3	1	76
4	1	77
5	1	78
6	1	79
7	1	80
8	1	81
9	1	82
10	1	83
11	1	84
12	1	85

ID	Metode	Skor
13	2	60
14	2	62
15	2	64
16	2	66
17	2	68
18	2	70
19	2	71
20	2	72
21	2	73
22	2	74
23	2	76
24	2	77
25	3	69
26	3	71
27	3	73
28	3	74
29	3	75
30	3	76
31	3	77
32	3	78
33	3	79
34	3	80
35	3	72

ID	Metode	Skor
36	3	81

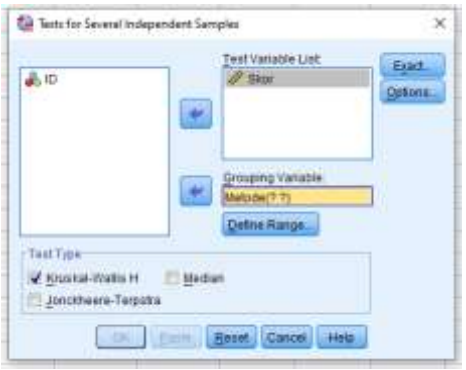
Rumusan Hipotesis pada Ilustrasi

H_0: Proporsi skor di atas dan di bawah median gabungan sama pada ketiga metode.

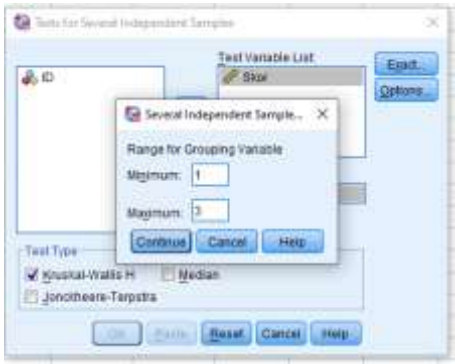
H_a: setidaknya satu metode memiliki proporsi skor yang berbeda terhadap median gabungan.

Langkah Mengolah dan Mengkaji Data

1. Pada Variable View, buat variabel id, metode, dan skor. Beri value labels pada metode.
2. Buka Analyze > Nonparametric Tests > Legacy Dialogs > K Independent Samples.
3. Tempatkan skor ke Test Variable List dan metode ke Grouping Variable.



4. Tekan Define Range, masukkan Minimum = 1 dan Maximum = 3.



5. Aktifkan Median pada Test Type, kemudian klik OK maka akan muncul keluaran berikut:

Median Test

Frequencies

		Metode		
		PBL	Langsung	Proyek
Skor	> Median	10	2	6
	<= Median	2	10	6

Tabel ini menunjukkan jumlah skor pada setiap metode yang berada di atas median gabungan dan sama dengan atau di bawah median gabungan. Median gabungan seluruh responden adalah 75,50 (dilihat dari tabel test statsitics)

Test Statistics[a]

	Skor
N	36
Median	75.50
Chi-Square	10.667[b]
df	2
Asymp. Sig.	.005

a. Grouping Variable: Metode

b. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 6.0.

nilai signifikansi:

$$0{,}005 < 0{,}050$$

Karena p = 0,005 < 0,05, H_0 ditolak.

Maknanya, terlihat perbedaan yang signifikan pada distribusi skor terhadap median gabungan antara metode PBL, pembelajaran langsung, dan metode proyek.

Hitung Cramer's V untuk mengukur seberapa kuat hubungan antar merode pembelajaran dan posisi skor terhadap median. Cramer's V = 0,54 menunjukkan perbedaan yang kuat pada dataset ilustratif.

Cara Membaca Keluaran

Pada dataset ilustratif, median gabungan seluruh skor adalah 75,50. Proporsi skor di atas median mencapai 83,3% pada kelompok PBL, 16,7% pada pembelajaran langsung, dan 50,0% pada metode proyek. Median Test menghasilkan $\chi^2(2, N = 36) = 10{,}67$, p = 0,005, dengan Cramer's V = 0,54 yang menunjukkan kekuatan hubungan besar.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Distribusi skor terhadap median gabungan berbeda secara signifikan di antara ketiga metode. Uji lanjut dengan koreksi pengujian multipel diperlukan untuk menentukan pasangan kelompok yang berbeda.

Median Test mengubah skor menjadi dua kategori sehingga informasi rinci berkurang. Bandingkan hasilnya dengan Kruskal-Wallis dan jelaskan alasan pemilihan uji dalam catatan pengolahan data.

LATIHAN

1. Susun tabel kontingensi untuk jenis kelamin dan status ketuntasan!
2. Hitung frekuensi harapan Tabel 9.1!
3. Paparkan kapan Fisher exact lebih tepat daripada Pearson chi-square!
4. Mengapa persentase baris atau kolom perlu dilaporkan?
5. Susun penyajian hasil chi-square yang mencantumkan ukuran efek!
6. Paparkan perbedaan tujuan Median Test dan Kruskal-Wallis serta akibat pengelompokan skor berdasarkan median gabungan!

Dataset Latihan Penerapan SPSS Bab 9.

Chi-kuadrat atau Fisher exact

ID	Metode	Ketuntasan
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	0

ID	Metode	Ketuntasan
13	1	0
14	1	0
15	1	0
16	2	1
17	2	1
18	2	1
19	2	1
20	2	1
21	2	1
22	2	1
23	2	1
24	2	0
25	2	0
26	2	0
27	2	0
28	2	0
29	2	0
30	2	0

Rumusan Hipotesis pada Ilustrasi Uji Chi-Kuadrat

H_0: Metode pembelajaran dan ketuntasan belajar saling independen.

H_a: Metode pembelajaran dan ketuntasan belajar saling berhubungan.

1. Susun crosstab metode x ketuntasan dengan observed count, expected count, dan row percentage.
2. Lakukan Pearson chi-square dan Fisher exact.

3. Cermati expected count untuk menentukan statistik yang paling tepat.

4. Hitung dan penjelasan maknakan Phi atau Cramer's V.

5. Tulis kesimpulan tentang hubungan metode dan ketuntasan tanpa menyatakan hubungan kausal.

Tabel 9. 5. Dataset Latihan 9.2. Median Test tiga kelompok independen

ID	Metode	Skor
1	1	72
2	1	74
3	1	76
4	1	77
5	1	78
6	1	79
7	1	80
8	1	81
9	1	82
10	1	83
11	2	65
12	2	67
13	2	69
14	2	70
15	2	71
16	2	72
17	2	73
18	2	74
19	2	76

ID	Metode	Skor
20	2	78
21	3	70
22	3	72
23	3	74
24	3	75
25	3	76
26	3	77
27	3	78
28	3	79
29	3	73
30	3	80

Rumusan Hipotesis pada Ilustrasi Median Test

H_0: Proporsi skor terhadap median gabungan sama pada ketiga metode.

H_a: setidaknya satu metode memiliki proporsi skor yang berbeda terhadap median gabungan.

1. Tempatkan data ke SPSS dengan kode metode 1 = PBL, 2 = Langsung, dan 3 = Proyek.
2. Gunakan prosedur Median Test menggunakan median gabungan yang dihitung SPSS.
3. Rumuskan median gabungan dan frekuensi skor di bawah atau sama dengan serta di atas median untuk setiap kelompok.
4. Laporkan chi-square, df, p-value, dan hitung Cramer V.
5. Jalankan Kruskal-Wallis pada data yang sama, lalu bandingkan keputusan dan informasi yang diberikan oleh kedua uji.
6. Susun kesimpulan yang sesuai konteks tanpa menyatakan hubungan sebab-akibat.

BAB X
PERBANDINGAN RATA-RATA BEBERAPA KELOMPOK (ANOVA)

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: menjelaskan tujuan ANOVA; membedakan ANOVA independen dan repeated measures; menafsirkan F dan post hoc; melaporkan ukuran efek.

Kata Kunci

ANOVA, between groups, within groups, F ratio, post hoc, Scheffe, Tukey, eta squared, repeated measures.

A. Mengapa Menggunakan ANOVA?

ANOVA membandingkan mean tiga atau lebih kondisi dengan satu pengujian omnibus. Melakukan banyak uji t meningkatkan peluang kesalahan tipe I. ANOVA menguji H_0 bahwa semua mean populasi sama melawan H_a bahwa setidaknya satu mean berbeda (Field, 2024; Gravetter & Wallnau, 2017).

ANOVA bekerja dengan membandingkan variasi antar-kelompok dengan variasi dalam kelompok. Jika variasi antar-kelompok jauh lebih besar daripada variasi dalam kelompok, nilai F meningkat.

$$F = \frac{MSantar}{MSdalam}$$

B. Asumsi ANOVA Satu Arah Independen

- Variabel dependen berskala interval/rasio.
- Kelompok independen dan observasi independen.
- Residual dalam setiap kelompok tidak mengalami penyimpangan normalitas berat.
- Varians antar-kelompok relatif homogen.
- tidak ada pencilan ekstrem yang tidak dapat dijelaskan.

ANOVA cukup robust terhadap penyimpangan moderat bila ukuran kelompok seimbang. Jika varians tidak homogen, Welch ANOVA dapat digunakan. Jika data ordinal atau distribusi sangat bermasalah, Kruskal-Wallis adalah alternatif (Field, 2024; Welch, 1947).

C. Contoh ANOVA Independen

Tabel 10. 1. Contoh tiga kelompok independen

Asal pelayan	Data lama berdiri (jam)
Kota	4, 5, 4, 6, 5, 3, 4, 3, 5, 6
Desa	4, 5, 6, 7, 4, 6, 4, 5, 3
Gunung	7, 4, 5, 6, 7, 5, 6, 7, 6, 7, 6

Proses pengolahan data dimulai dengan statistik deskriptif, boxplot, dan uji homogenitas. Jika uji omnibus signifikan, pengujian dilanjutkan dengan uji post hoc untuk mengetahui pasangan kelompok yang berbeda. Jika tidak signifikan, post hoc biasanya tidak diperlukan (Field, 2024).

D. Post Hoc dan Koreksi Multipel

Tabel 10. 2. Pilihan uji lanjut ANOVA

Metode	Penggunaan umum	Catatan
Tukey HSD	Semua perbandingan pasangan; varians homogen	Mengendalikan familywise error
Scheffe	Perbandingan fleksibel dan konservatif	Dapat dipakai untuk kontras kompleks
Bonferroni	Jumlah perbandingan terbatas	Sederhana, dapat konservatif
Games-Howell	Varians tidak homogen dan n berbeda	Alternatif setelah Welch
LSD	Eksploratif dengan perlindungan terbatas	Risiko tipe I lebih tinggi

E. Repeated-Measures ANOVA

Jika orang yang sama diukur tiga kali atau lebih, gunakan repeated-measures ANOVA, bukan one-way ANOVA independen. Contohnya adalah produktivitas 15 karyawan sebelum alat baru, tiga bulan sesudahnya, dan enam bulan sesudahnya. Karena orang yang sama diukur berulang, hubungan antar-pengukuran harus diperhitungkan.

Pada data tersebut, mean produktivitas meningkat dari 12,47 menjadi 15,73 dan 17,53. Repeated-measures ANOVA menghasilkan $F(2,28)$ sekitar 33,75 dengan $p < 0,001$.

Pengujian lanjutan pasangan waktu perlu menggunakan koreksi multipel. Asumsi tambahan adalah sphericity; bila dilanggar, gunakan koreksi Greenhouse-Geisser atau Huynh-Feldt (Field, 2024).

Tabel 10. 3. Rata-rata produktivitas pada tiga waktu

Waktu	Mean produktivitas
Sebelum	12,47
3 bulan	15,73
6 bulan	17,53

Klarifikasi Metodologis

Contoh pengukuran berulang pada orang yang sama tidak boleh diperlakukan sebagai tiga kelompok independen. Pemilihan repeated-measures ANOVA mempertahankan struktur pasangan dan menghasilkan galat uji yang lebih tepat.

F. Ukuran Efek

$$\eta^2 = \frac{SSantar}{SStotal}$$

Eta squared

Eta squared dan partial eta squared menunjukkan proporsi variasi yang dijelaskan faktor. Nilai harus ditafsirkan bersama desain dan bidang kajian. Ukuran efek tidak membuktikan kausalitas jika desain tidak mendukung inferensi sebab-akibat (Cohen, 1988; Lakens, 2013).

G. Langkah SPSS

1. One-way independen: Analyze > Compare Means > One-Way ANOVA.

2. Tempatkan variabel dependen ke Dependent List dan grup ke Factor.
3. Options: Descriptives dan Homogeneity of variance test.
4. Post Hoc: Tukey/Scheffe bila homogen, Games-Howell bila tidak homogen.
5. Repeated measures: Analyze > General Linear Model > Repeated Measures; definisikan faktor waktu dan masukkan variabel tiap waktu (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Terdapat perbedaan produktivitas antarwaktu, $F(2, 28) = 33{,}75$, $p < 0{,}001$. Mean meningkat dari sebelum penggunaan alat ($M = 12{,}47$) ke 3 bulan ($M = 15{,}73$) dan 6 bulan ($M = 17{,}53$). Perbandingan berpasangan dengan koreksi Bonferroni digunakan untuk mengidentifikasi waktu yang berbeda.

CONTOH PENERAPAN BAB 10

ANOVA Satu Arah dan Uji Lanjut Tukey

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: tiga metode pembelajaran dibandingkan dilihat dari nilai kemampuan berpikir kritis matematis. Kelompok bersifat independen.

Tabel 10. 4. Ilustrasi 10.1. Data fiktif ANOVA, 1 = Langsung; 2 = PBL; 3 = Proyek

ID	Metode	Nilai
1	1	72
2	1	75
3	1	78

ID	Metode	Nilai
4	1	74
5	1	76
6	1	73
7	1	77
8	1	79
9	1	75
10	1	74
11	2	80
12	2	82
13	2	85
14	2	78
15	2	84
16	2	81
17	2	83
18	2	86
19	2	79
20	2	82
21	3	86
22	3	88
23	3	90
24	3	85
25	3	89
26	3	87
27	3	91
28	3	84

ID	Metode	Nilai
29	3	88
30	3	90

Rumusan Hipotesis pada Ilustrasi

H_0: μLangsung = μPBL = μProyek (rata-rata ketiga metode sama).

H_a: setidaknya satu rata-rata metode berbeda.

Langkah Mengolah dan Mengkaji Data

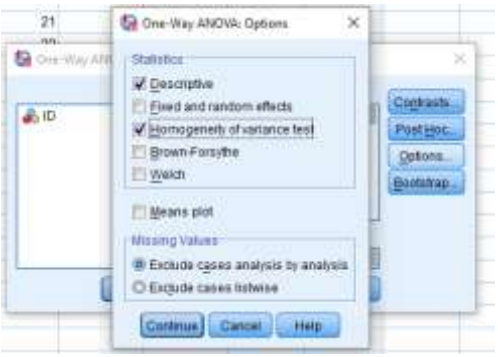
1. Cermati normalitas nilai per metode dan pencilan.

   Diasumsikan normal

2. Akses Analyze > Compare Means > One-Way ANOVA.
3. Isikan nilai ke Dependent List dan metode ke Factor.



4. Pada Options, aktifkan Descriptive dan Homogeneity of variance test.



5. Pada Post Hoc, pilih Tukey jika varians homogen. Pilih Games-Howell jika varians tidak homogen.



6. Tekan OK, kemudian muncul keluaran berikut:

Oneway

Descriptives

Nilai

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
PBL	10	75.30	2.214	.700	73.72	76.88	72	79
Langsung	10	82.00	2.582	.816	80.15	83.85	78	86
Proyek	10	87.80	2.300	.727	86.15	89.45	84	91
Total	30	81.70	5.676	1.036	79.58	83.82	72	91

Secara deskriptif, rata-rata tertinggi ada pada metode proyek, diikuti metode PBL. Sementara itu, rata-rata terendah ada pada metode pembelajaran langsung.

Test of Homogeneity of Variances

Nilai

Levene Statistic	df1	df2	Sig.
.086	2	27	.918

Keluaran contoh Levene menunjukkan nilai Sig. = 0,918. Karena 0,918 > 0,05, varians ketiga kelompok dinyatakan homogen. Jadi, asumsi homogenitas varians untuk One-Way ANOVA terpenuhi.

ANOVA

Nilai

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	782.600	2	391.300	69.645	.000
Within Groups	151.700	27	5.619		
Total	934.300	29			

Karena $p < 0,05$, H_0 ditolak dan H_a didukung. Artinya, ada perbedaan rata-rata nilai yang signifikan di antara metode pembelajaran langsung, PBL, dan proyek.

Walaupun begitu, tabel ANOVA belum menunjukkan kelompok mana yang berbeda secara signifikan. Untuk mengetahui perbedaan antar pasangan kelompok, perlu dilakukan uji post hoc, Tukey HSD.

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Nilai

Tukey HSD

(I) Metode	(J) Metode	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
PBL	Langsung	-6.700*	1.060	.000	-9.33	-4.07
	Proyek	-12.500*	1.060	.000	-15.13	-9.87
Langsung	PBL	6.700*	1.060	.000	4.07	9.33
	Proyek	-5.800*	1.060	.000	-8.43	-3.17
Proyek	PBL	12.500*	1.060	.000	9.87	15.13
	Langsung	5.800*	1.060	.000	3.17	8.43

*. The mean difference is significant at the 0.05 level.

Semua interval kepercayaan 95% tidak melewati angka nol. Jadi, seluruh pasangan metode memiliki perbedaan rata-rata yang signifikan.

Urutan rata-rata nilai dari tertinggi hingga terendah adalah:

Proyek (87,80) > PBL (82,00) > pembelajaran langsung (75,30).

Homogeneous Subsets

Nilai

Tukey HSD[a]

Metode	N	Subset for alpha = 0.05		
		1	2	3
PBL	10	75.30		
Langsung	10		82.00	
Proyek	10			87.80
Sig.		1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.

Tabel ini mengelompokkan metode yang rata-ratanya tidak berbeda secara signifikan. Metode yang berada dalam subset yang sama berarti tidak memiliki perbedaan signifikan.

Cara Membaca Keluaran

Pada dataset ilustratif, rata-rata nilai pada pembelajaran langsung adalah 75,30 (SD = 2,21), PBL 82,00 (SD = 2,58), dan metode proyek 87,80 (SD = 2,30). Uji Levene menghasilkan $p = 0{,}918$, sehingga asumsi homogenitas varians terpenuhi. ANOVA satu arah menghasilkan $F(2, 27) = 69{,}65$, $p < 0{,}001$, dengan $\eta^2 = 0{,}84$. Uji Tukey menunjukkan bahwa seluruh pasangan metode berbeda secara signifikan, dengan urutan rata-rata metode proyek, PBL, kemudian pembelajaran langsung.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. setidaknya satu rata-rata metode pembelajaran berbeda, dan hasil uji lanjut menunjukkan bahwa ketiga metode berbeda satu sama lain. Metode proyek menghasilkan rata-rata tertinggi. Sementara itu, pembelajaran langsung menghasilkan rata-rata terendah. Ukuran efek ANOVA tergolong sangat besar.

LATIHAN

1. Jelaskan mengapa lima kelompok tidak sebaiknya dibandingkan dengan sepuluh uji t tanpa koreksi!
2. Bedakan ANOVA satu arah independen dan repeated-measures ANOVA!
3. Kapan menggunakan Games-Howell?
4. Jelaskan arti F yang besar!
5. Tulis contoh penyusunan kalimat statistik ANOVA dengan post hoc!

Dataset Latihan Penerapan SPSS Bab 10.
ANOVA satu arah

ID	Metode	Nilai
1	1	68
2	1	72
3	1	70
4	1	74
5	1	69
6	1	73
7	1	71
8	1	75
9	2	74
10	2	78
11	2	76
12	2	80
13	2	75

ID	Metode	Nilai
14	2	79
15	2	77
16	2	81
17	3	80
18	3	84
19	3	82
20	3	86
21	3	81
22	3	85
23	3	83
24	3	87

Rumusan Hipotesis pada Ilustrasi

H_0: μCeramah = μPBL = μProyek (rata-rata ketiga metode sama).

H_a: setidaknya satu rata-rata metode berbeda.

1. Beri label metode 1 = Ceramah, 2 = PBL, dan 3 = Proyek.
2. Periksa normalitas per kelompok, pencilan, dan homogenitas varians.
3. Jalankan One-Way ANOVA.
4. Buka uji lanjut yang sesuai dari hasil homogenitas.
6. Laporkan mean, SD, F, df, p-value, ukuran efek, serta hasil perbandingan pasangan.

BAB XI
UJI KRUSKAL-WALLIS DAN FRIEDMAN

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: memilih uji nonparametrik untuk tiga atau lebih sampel; melakukan perankingan gabungan dan dalam subjek; menafsirkan Kruskal-Wallis dan Friedman; merencanakan post hoc.

Kata Kunci

Kruskal-Wallis, Friedman, rank, k sampel, independen, berpasangan, post hoc nonparametrik.

A. Kruskal-Wallis untuk k Kelompok Independen

Kruskal-Wallis adalah alternatif nonparametrik ANOVA satu arah independen. Uji digunakan untuk data ordinal atau ketika setidaknya satu kelompok tidak memenuhi asumsi parametrik secara cukup. Semua observasi digabung dan dirangking; rank kemudian dikembalikan ke kelompok masing-masing (Kruskal & Wallis, 1952).

H_0 menyatakan distribusi rank kelompok sama. Jika bentuk distribusi serupa, perbedaan dapat ditafsirkan sebagai perbedaan lokasi atau median. Jika bentuk berbeda, hasil menunjukkan perbedaan distribusi secara lebih umum.

$$H = \frac{12}{N(N+1)} \sum \frac{R_j{}^2}{n_j} - 3(N+1)$$

Pada contoh tiga kelompok prestasi, jumlah rank adalah 57; 140,5; dan 102,5. Diperoleh H = 4,466, df = 2, p = 0,107. Data belum menunjukkan perbedaan signifikan antar-kelompok.

B. Friedman untuk k Kondisi Berpasangan

Uji Friedman adalah alternatif nonparametrik repeated-measures ANOVA. Setiap subjek dinilai pada tiga atau lebih kondisi. Rank diberikan di dalam setiap subjek, bukan pada seluruh data sekaligus. Hal ini mempertahankan struktur berpasangan (Friedman, 1937).

Pada contoh sepuluh siswa yang mengikuti tiga metode, jumlah rank metode I, II, dan III adalah 17,5; 19,5; dan 23. Statistik Friedman χ^2 = 1,879, df = 2, p = 0,391. Tidak ada bukti perbedaan hasil belajar antar-metode.

$$\chi^2{}_F = \frac{12}{nk(k+1)} \sum R_j{}^2 - 3n(k+1)$$

C. Penanganan Ties dan Post Hoc

Ties atau nilai sama memerlukan rank rata-rata dan koreksi tertentu yang dilakukan otomatis oleh perangkat lunak. Bila uji omnibus signifikan, Kruskal-Wallis dapat diikuti perbandingan Dunn dengan koreksi Holm/Bonferroni. Friedman dapat diikuti Wilcoxon berpasangan dengan koreksi multipel (Field, 2024).

Penyajian hasil post hoc harus menyatakan metode koreksi. Melakukan semua pasangan tanpa koreksi dapat meningkatkan peluang temuan positif palsu (Appelbaum *et al.*, 2018).

D. Ukuran Efek

$$\varepsilon^2 = \frac{H - k + 1}{N - k}$$

Epsilon squared untuk Kruskal-Wallis

$$W = \frac{\chi^2{}_F}{n(k-1)}$$

Kendall's W untuk Friedman (Field, 2024).

E. Langkah SPSS

1. Kruskal-Wallis: Analyze > Nonparametric Tests > Legacy Dialogs > K Independent Samples.
2. Isikan skor ke Test Variable List dan grup ke Grouping Variable; tentukan rentang kode grup.
3. Friedman: Analyze > Nonparametric Tests > Legacy Dialogs > K Related Samples.
4. Isikan semua kondisi ke Test Variables dan centang Friedman.
5. Baca mean rank, statistik chi-square, df, dan p-value (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Tidak ada perbedaan rank prestasi antar-tiga kelompok, H(2) = 4,47, p = 0,107, ε^2 = 0,11. Karena uji omnibus tidak signifikan, perbandingan post hoc tidak dilakukan.

CONTOH PENERAPAN BAB 11

Kruskal-Wallis dan Friedman untuk Data yang Tidak Memenuhi Asumsi Parametrik

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka*

yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.

A. Kruskal-Wallis untuk tiga kelompok independen

Ilustrasi kasus: skor pemahaman konsep dari tiga kelompok memiliki skala ordinal atau tidak layak dikaji dengan ANOVA parametrik.

Tabel 11. 1. Ilustrasi 11.1. Data fiktif Kruskal-Wallis

ID	Kelompok	Skor
1	1	62
2	1	65
3	1	67
4	1	70
5	1	68
6	1	64
7	1	66
8	1	69
9	2	72
10	2	74
11	2	76
12	2	73
13	2	75
14	2	71
15	2	77
16	2	74
17	3	80
18	3	82
19	3	85

ID	Kelompok	Skor
20	3	81
21	3	83
22	3	79
23	3	84
24	3	82

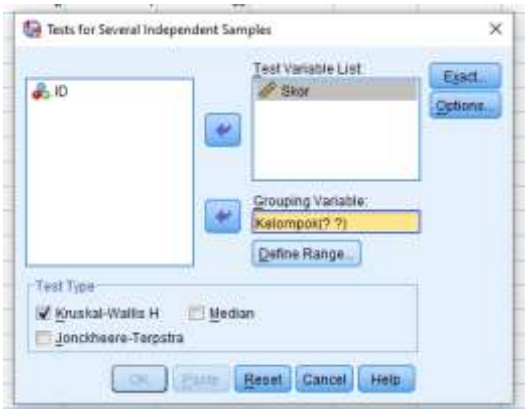
Rumusan Hipotesis pada Ilustrasi

H_0: Distribusi skor pada ketiga kelompok sama.

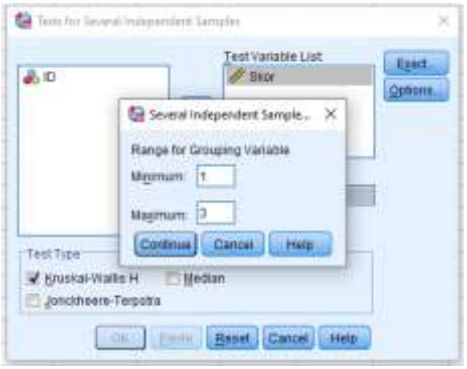
H_a: setidaknya satu kelompok memiliki distribusi skor yang berbeda.

Langkah Mengolah dan Mengkaji Data

1. Pilih Analyze > Nonparametric Tests > Legacy Dialogs > K Independent Samples.
2. Tempatkan skor ke Test Variable List dan kelompok ke Grouping Variable.



3. Define Range: minimum 1 dan maksimum 3.



4. Buka Kruskal-Wallis H, dan akan muncul keluaran berikut:

Kruskal-Wallis Test

Ranks

	Kelompok	N	Mean Rank
Skor	1	8	4.50
	2	8	12.50
	3	8	20.50
	Total	24	

Secara deskriptif urutan skor dari tertinggi sampai terendah adalah:

$$Kelompok3 > Kelompok2 > Kelompok1$$

Perlu dicatat bahwa mean rank bukan rata-rata skor asli, melainkan rata-rata peringkat setelah seluruh skor diurutkan.

Test Statistics[a,b]

	Skor
Chi-Square	20.498
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Kelompok

$$p < 0{,}001 < 0{,}05$$

Karena p < 0,001, H_0 ditolak. Artinya, setidaknya satu kelompok memiliki distribusi skor yang berbeda.

Cara Membaca Keluaran

Pada dataset ilustratif, median skor berturut-turut adalah 66,50 pada kelompok 1, 74,00 pada kelompok 2, dan 82,00 pada kelompok 3. Uji Kruskal-Wallis menghasilkan H(2) = 20,498 dan p < 0,001. Nilai ε^2 = 0,881 menunjukkan ukuran efek yang sangat besar.

Arti Keputusan Statistik

Pada ilustrasi ini, karena p < 0,05, H_0 ditolak dan H_a didukung. setidaknya satu kelompok memiliki distribusi skor yang berbeda. Uji pasangan dengan koreksi pengujian multipel diperlukan untuk menentukan kelompok yang berbeda secara spesifik.

B. Friedman untuk tiga kondisi berpasangan

Tabel 11. 2. Ilustrasi 11.2. Data fiktif tiga kondisi pada peserta yang sama

ID	Kondisi 1	Kondisi 2	Kondisi 3
1	68	74	80
2	70	76	82
3	72	75	81
4	65	72	78

ID	Kondisi 1	Kondisi 2	Kondisi 3
5	69	73	79
6	71	77	83
7	67	71	77
8	73	78	84
9	66	72	79
10	70	75	81
11	68	74	80
12	69	76	82

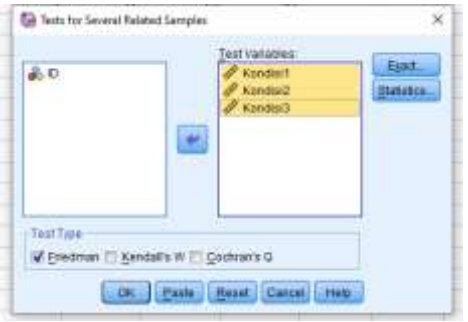
Rumusan Hipotesis pada Ilustrasi

H_0: Distribusi skor pada Kondisi 1, Kondisi 2, dan Kondisi 3 sama.

H_a: setidaknya satu kondisi memiliki distribusi skor yang berbeda.

Langkah Mengolah dan Mengkaji Data

1. Akses Analyze > Nonparametric Tests > Legacy Dialogs > K Related Samples.
2. Masukkan ketiga variabel kondisi.



3. Pilih Friedman, lalu klik OK. Keluaran SPSS

Friedman Test

Ranks

	Mean Rank
Kondisi1	1.00
Kondisi2	2.00
Kondisi3	3.00

Urutan skor secara deskriptif adalah:

$$Kondisi3 > Kondisi2 > Kondisi1$$

Mean rank bukan rata-rata skor asli, melainkan rata-rata peringkat skor setiap responden pada ketiga kondisi.

Test Statistics[a]

N	12
Chi-Square	24.000
df	2
Asymp. Sig.	.000

a. Friedman Test

Karena p < 0,001, H_0 ditolak. Artinya, ada perbedaan distribusi skor yang signifikan di antara ketiga kondisi pengukuran.

Kekuatan perbedaan

Ukuran efek uji Friedman dapat dihitung menggunakan **Kendall's W**:

$$w = \frac{x^2}{N(k-1)} = \frac{24}{12(3-1)} = 1{,}00$$

Nilai Kendall's W = 1,00 menunjukkan perbedaan atau kesepakatan peringkat yang sangat kuat atau sempurna. Ini menunjukkan pola peringkat yang sangat konsisten pada seluruh responden.

Cara Membaca Keluaran

Pada dataset ilustratif, median skor meningkat dari 69,00 pada Kondisi 1 menjadi 74,50 pada Kondisi 2 dan 80,50 pada Kondisi 3. Uji Friedman menghasilkan $\chi^2(2) = 24{,}00$ dan $p < 0{,}001$. Kendall's W sebesar 1,00 menunjukkan kekuatan perbedaan atau konsistensi peringkat yang sempurna pada dataset ilustratif.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Distribusi skor berbeda secara signifikan di antara ketiga kondisi. Uji pasangan dengan koreksi pengujian multipel diperlukan untuk mengetahui kondisi yang berbeda secara spesifik.

LATIHAN

1. Paparkan perbedaan cara perankingan Kruskal-Wallis dan Friedman!
2. Kapan hasil Kruskal-Wallis dapat ditafsirkan sebagai perbedaan median?
3. Susun H_0 dan H_a untuk membandingkan tiga metode independen!
4. Jelaskan mengapa post hoc tidak dilakukan bila uji omnibus tidak signifikan!
5. Susun penyajian hasil Friedman lengkap dengan Kendall W!

Dataset Latihan Penerapan SPSS Bab 11: Kruskal-Wallis

ID	Kelompok	Skor
1	1	60
2	1	64
3	1	66
4	1	62
5	1	68
6	1	65
7	1	63
8	2	70
9	2	73

ID	Kelompok	Skor
10	2	72
11	2	75
12	2	71
13	2	74
14	2	76
15	3	78
16	3	82
17	3	80
18	3	84
19	3	79
20	3	83
21	3	81

Rumusan Hipotesis pada Ilustrasi

H_0: Distribusi skor ketiga kelompok sama.

H_a: setidaknya satu kelompok memiliki distribusi skor yang berbeda.

1. Gunakan prosedur Kruskal-Wallis untuk membandingkan tiga kelompok.
2. Dokumentasikan median setiap kelompok, H, df, dan p-value.
3. Hitung epsilon squared.
4. Jika hasil signifikan, rancang uji lanjut (post hoc) pasangan yang disertai koreksi multipel.

5. Paparkan kondisi yang membuat Kruskal-Wallis lebih tepat daripada ANOVA satu arah.

BAB XII
UJI COCHRAN Q

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: mengidentifikasi data dikotomi berpasangan; melakukan Cochran Q; menafsirkan hasil dan post hoc; menyajikan pengujian data nominal berulang.

Kata Kunci

Cochran Q, dikotomi, k sampel berpasangan, lulus-gagal, McNemar.

A. Tujuan dan Struktur Data

Cochran Q digunakan untuk membandingkan tiga atau lebih kondisi berpasangan ketika hasil setiap kondisi dikotomi, misalnya lulus-gagal, ya-tidak, berhasil-tidak berhasil, atau setuju-tidak setuju. Uji ini adalah perluasan McNemar untuk k kondisi (Cochran, 1950).

Setiap baris mewakili subjek yang sama dan setiap kolom mewakili kondisi. Nilai diberi kode 1 untuk kejadian yang ditetapkan sebagai berhasil dan 0 untuk tidak berhasil. Observasi antar-subjek independen, sedangkan pengukuran dalam subjek berpasangan (Field, 2024).

B. Contoh Tiga Metode Mengajar

Tabel 12 1. Data dikotomi keberhasilan tiga metode (1 = skor ≥70)

Siswa	Metode I	Metode II	Metode III
1	0	1	1
2	1	0	1
3	0	0	0
4	0	0	1
5	1	1	1
6	0	1	0
7	0	0	1
8	1	1	0
9	1	0	1
10	1	1	1

Jumlah keberhasilan pada metode I, II, dan III adalah 5, 5, dan 7. Cochran Q menghasilkan Q = 1,143 dengan df = 2 dan p = 0,565. Hasil ini tidak menunjukkan perbedaan proporsi keberhasilan yang signifikan antar-metode.

$$Q = \frac{(k-1)[k\sum C_j^2 - T^2]}{kT - \sum R_i^2}$$

Salah satu bentuk rumus Cochran Q

C. Asumsi dan Catatan

1. Variabel respons dikotomi dan dikodekan konsisten.
2. Subjek yang sama dinilai pada semua kondisi.
3. Baris atau subjek independen satu sama lain.
4. Tidak ada kondisi yang adalah duplikasi sempurna.
5. Kehilangan data perlu ditangani dengan hati-hati karena uji membutuhkan set kondisi lengkap.

D. Post Hoc

Jika Cochran Q signifikan, lakukan uji McNemar untuk pasangan kondisi. Karena ada beberapa perbandingan, gunakan koreksi Holm atau Bonferroni. Laporkan proporsi keberhasilan setiap kondisi agar arah perbedaan jelas (Appelbaum *et al*., 2018).

Ukuran efek dapat dinyatakan melalui selisih proporsi berpasangan atau odds ratio berpasangan untuk setiap post hoc. Kesimpulan tidak cukup hanya menyebut nilai Q.

E. Langkah SPSS

1. Isikan setiap kondisi sebagai variabel kolom yang berisi 0 dan 1.
2. Buka Analyze > Nonparametric Tests > Legacy Dialogs > K Related Samples.
3. Tempatkan semua variabel ke Test Variables dan centang Cochran Q (Pallant, 2026).
4. Baca frekuensi setiap kondisi, Q, df, dan p-value.
5. Jika signifikan, jalankan McNemar per pasangan dengan koreksi multipel.

Contoh Penyusunan Kalimat Statistik

Proporsi keberhasilan pada metode I, II, dan III masing-masing 50%, 50%, dan 70%. Perbedaan tersebut tidak signifikan, Cochran Q(2) = 1,14, p = 0,565. Dengan demikian, data belum menunjukkan perbedaan peluang keberhasilan antar-metode.

CONTOH PENERAPAN BAB 12

Cochran Q untuk Tiga Pengukuran Dikotomi Berpasangan

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Skenario penelitian: 15 mahasiswa dinilai tuntas atau belum tuntas setelah menggunakan tiga bentuk bantuan belajar. Setiap mahasiswa mengalami seluruh metode.

Tabel 12. 1. Ilustrasi 12.1. Data fiktif dikotomi berpasangan, 1 = Tuntas; 0 = Belum

ID	Metode A	Metode B	Metode C
1	1	1	1
2	0	1	1
3	1	1	1
4	0	0	1
5	1	1	1
6	0	1	1
7	1	1	1
8	0	0	1
9	0	1	1
10	1	1	1
11	0	1	0
12	0	0	1

ID	Metode A	Metode B	Metode C
13	1	1	1
14	0	1	1
15	0	0	0

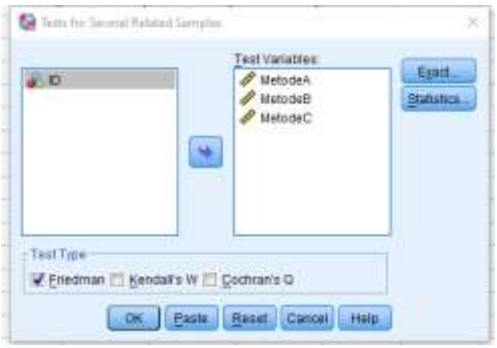
Rumusan Hipotesis pada Ilustrasi

H_0: Proporsi keberhasilan pada Metode A, Metode B, dan Metode C sama.

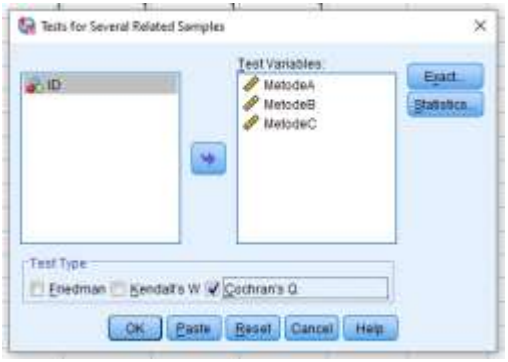
H_a: setidaknya satu metode memiliki proporsi keberhasilan yang berbeda.

Langkah Mengolah dan Mengkaji Data

1. Masukkan setiap metode pada kolom terpisah dan setiap mahasiswa pada satu baris.
2. Berikan value labels setiap metode
3. Akses Analyze > Nonparametric Tests > Legacy Dialogs > K Related Samples.
4. Isikan Metode A, Metode B, dan Metode C ke Test Variables.



5. Pilih Cochran Q, lalu klik OK.



Keluaran SPSS akan muncul sebagai berikut:

Cochran Test

Frequencies

	Value	
	0	1
MetodeA	9	6
MetodeB	4	11
MetodeC	2	13

Secara deskriptif, Metode C memiliki jumlah keberhasilan tertinggi, yaitu 13 orang atau 86,7%. Metode B menghasilkan keberhasilan pada 11 orang atau 73,3%. Sementara itu, Metode A hanya 6 orang atau 40%.

Urutan proporsi keberhasilan adalah:

$$MetodeC > MetodeB > MetodeA$$

Test Statistics

N	15
Cochran's Q	9.750[a]
df	2
Asymp. Sig.	.008

a. 1 is treated as a success.

$$0{,}008 < 0{,}050$$

Karena p = 0,008 < 0,05, H_0 ditolak. Artinya, terlihat perbedaan proporsi keberhasilan yang signifikan di antara ketiga metode.

Walaupun begitu, uji Cochran hanya menunjukkan bahwa ada perbedaan secara umum. Uji ini belum menunjukkan pasangan metode mana yang berbeda secara signifikan. Untuk mengetahuinya, diperlukan uji lanjut McNemar pada setiap pasangan:

- Metode A dengan Metode B
- Metode A dengan Metode C
- Metode B dengan Metode C

Karena ada tiga perbandingan, koreksi Bonferroni dapat digunakan:

$$\alpha = \frac{0{,}05}{3} = 0{,}0167$$

Jadi, pada uji lanjut, pasangan metode dinyatakan berbeda signifikan jika $p < 0{,}0167$

Cara Membaca Keluaran

Pada dataset ilustratif, proporsi ketuntasan pada Metode A adalah 40,0%, Metode B 73,3%, dan Metode C 86,7%. Uji Cochran Q menghasilkan Q(2) = 9,75 dan p = 0,008. Secara

deskriptif, ketuntasan tertinggi ada pada Metode C dan terendah pada Metode A.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0,05$, H_0 ditolak dan H_a didukung. Proporsi ketuntasan berbeda secara signifikan di antara ketiga metode. Uji lanjut McNemar pada setiap pasangan dengan koreksi Bonferroni diperlukan untuk menentukan pasangan metode yang berbeda.

LATIHAN

1. Berikan tiga contoh data pendidikan yang cocok dikaji dengan Cochran Q!
2. Ubah skor tiga metode pada Lampiran C menjadi kode 0-1 menggunakan batas 75!
3. Apa perbedaan Cochran Q dan Friedman?
4. Mengapa McNemar memerlukan koreksi bila dilakukan berulang?
5. Tulis contoh penyusunan kalimat statistik Cochran Q yang tidak signifikan!

H. LATIHAN PENERAPAN SPSS BAGIAN XII

Dataset Latihan Penerapan SPSS Bab 12.

Cochran Q

ID	Metode A	Metode B	Metode C
1	1	1	1
2	0	1	1
3	0	1	1
4	1	1	1
5	0	0	1
6	1	1	1
7	0	1	0
8	0	1	1
9	1	1	1
10	0	0	1
11	0	1	1
12	1	1	1

ID	Metode A	Metode B	Metode C
13	0	0	0
14	0	1	1
15	1	1	1

Rumusan Hipotesis pada Ilustrasi

H_0: Proporsi ketuntasan pada ketiga metode sama.

H_a: setidaknya satu metode memiliki proporsi ketuntasan yang berbeda.

1. Jalankan Cochran Q terhadap tiga metode.
2. Hitung persentase ketuntasan setiap metode.
3. Catat Q, df, dan p-value.
4. Jika signifikan, jalankan McNemar untuk tiga pasangan dengan taraf Bonferroni 0,0167.
5. Tulis penjelasan makna yang memperhatikan bahwa pengukuran berasal dari peserta yang sama.

BAB XIII
KORELASI PEARSON DAN SPEARMAN

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: menjelaskan arah dan kekuatan hubungan; menghitung serta menafsirkan Pearson r; memilih Spearman rho; membedakan korelasi dan kausalitas.

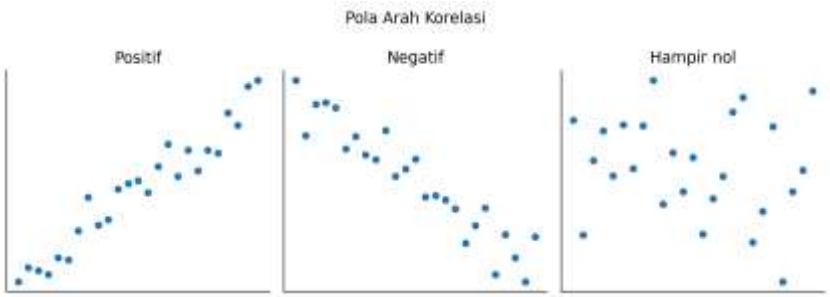
Kata Kunci

korelasi, Pearson, Spearman, scatterplot, koefisien determinasi, kausalitas.

A. Hipotesis Asosiatif

Hipotesis asosiatif menyatakan dugaan tentang hubungan antarvariabel dalam populasi. Koefisien korelasi menunjukkan arah dan kekuatan hubungan. Nilai positif berarti kedua variabel cenderung bergerak searah; nilai negatif berarti bergerak berlawanan; nilai mendekati nol menunjukkan hubungan linear atau monoton yang lemah (Field, 2024; Triola, 2022).

Korelasi tidak membuktikan sebab-akibat. Hubungan dapat dipengaruhi variabel perancu, arah kausal terbalik, bias pengukuran, atau kebetulan. Inferensi kausal memerlukan desain dan asumsi tambahan (Field, 2024).



Gambar 13. 1. Contoh pola korelasi positif, negatif, dan mendekati nol

B. Korelasi Product Moment Pearson

Pearson r digunakan untuk dua variabel interval/rasio yang memiliki hubungan linear. Observasi harus berpasangan dan independen antar-subjek. Pencilan dapat sangat memengaruhi nilai r, sehingga scatterplot harus diperiksa sebelum pengujian (Field, 2024; Triola, 2022).

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

Koefisien korelasi Pearson

Tabel 13. 1. Pedoman penjelasan makna kekuatan korelasi

|r|	Panduan membaca keluaran umum
0,00-0,19	Sangat rendah
0,20-0,39	Rendah
0,40-0,59	Sedang
0,60-0,79	Tinggi
0,80-1,00	Sangat tinggi

Gunakan Secara Kontekstual

Batas kategori korelasi bersifat pedoman. Dalam bidang tertentu, korelasi 0,30 dapat memiliki makna praktis penting, sedangkan korelasi tinggi belum tentu berguna bila pengukuran bias.

C. Contoh Pearson

Tabel 13. 2. Nilai Matematika dan IPA

Siswa	Matematika	IPA
1	6	6
2	6	7
3	7	7
4	8	9
5	7	8
6	8	8
7	6	7
8	7	8

Data delapan siswa menghasilkan r = 0,832 dan p = 0,010. Hubungan positif sangat tinggi menurut pedoman tabel. Koefisien determinasi r^2 = 0,692, yang berarti sekitar 69,2% variasi linear bersama dapat dinyatakan oleh kedua variabel. Pernyataan ini bukan berarti nilai Matematika menyebabkan 69,2% nilai IPA (Field, 2024).

$$KD = r^2 \times 100\%$$

Koefisien determinasi

D. Korelasi Rank Spearman

Spearman rho digunakan untuk data ordinal atau hubungan monoton yang tidak harus linear. Data diubah menjadi rank, lalu korelasi Pearson dihitung pada rank. Spearman lebih tahan terhadap pencilan daripada Pearson, tetapi tetap dapat dipengaruhi pengamatan ekstrem dan ties (Spearman, 1904).

Pada contoh motivasi dan peringkat kelas, skor motivasi yang tinggi cenderung terkait dengan nomor peringkat yang lebih kecil. Karena nomor 1 adalah peringkat terbaik, korelasinya negatif kuat: rho sekitar -0,939, $p < 0{,}001$. Tanda negatif perlu ditafsirkan sesuai arah skala.

$$\rho = 1 - \frac{6\sum {d_i}^2}{n(n^2 - 1)}$$

Spearman rho tanpa ties

E. Signifikansi dan Interval Kepercayaan

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$$

Uji signifikansi Pearson, df = n - 2

Nilai p korelasi dipengaruhi oleh ukuran sampel. Korelasi kecil dapat signifikan pada sampel besar, sedangkan korelasi besar dapat tidak signifikan pada sampel sangat kecil. Interval kepercayaan membantu menunjukkan presisi estimasi (Wasserstein & Lazar, 2016).

Pengujian banyak korelasi tanpa rencana meningkatkan risiko temuan kebetulan. Peneliti perlu menetapkan hipotesis, mengendalikan pengujian multipel, dan tidak memilih hanya korelasi yang signifikan (Appelbaum *et al.*, 2018).

F. Langkah SPSS dan Penyajian Hasil

1. Buat scatterplot melalui Graphs > Chart Builder untuk memeriksa pola dan pencilan.
2. Pilih Analyze > Correlate > Bivariate.
3. Tempatkan variabel, centang Pearson atau Spearman, dan pilih uji dua sisi/satu sisi sesuai hipotesis awal.
4. Baca koefisien, Sig., dan jumlah pasangan N.
5. Laporkan arah, kekuatan, p-value, interval kepercayaan bila tersedia, dan keterbatasan kausal (Appelbaum *et al*., 2018; Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Terdapat korelasi positif yang kuat antara nilai Matematika dan IPA, r(6) = 0,83, p = 0,010, r^2 = 0,69. Siswa dengan nilai Matematika lebih tinggi cenderung memiliki nilai IPA lebih tinggi; hasil ini tidak dengan sendirinya membuktikan hubungan sebab-akibat.

CONTOH PENERAPAN BAB 13

Menguji Hubungan Motivasi dan Hasil Belajar

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: peneliti menguji hubungan antara skor motivasi belajar dan nilai matematika 20 mahasiswa.

Tabel 13. 3. Ilustrasi 13.1. Data fiktif korelasi

ID	Motivasi	Nilai
1	58	67
2	62	64
3	65	73
4	67	69

ID	Motivasi	Nilai
5	70	78
6	72	71
7	74	82
8	76	75
9	78	85
10	80	79
11	82	88
12	84	83
13	86	91
14	88	86
15	90	94
16	60	70
17	69	72
18	73	80
19	77	77
20	85	89

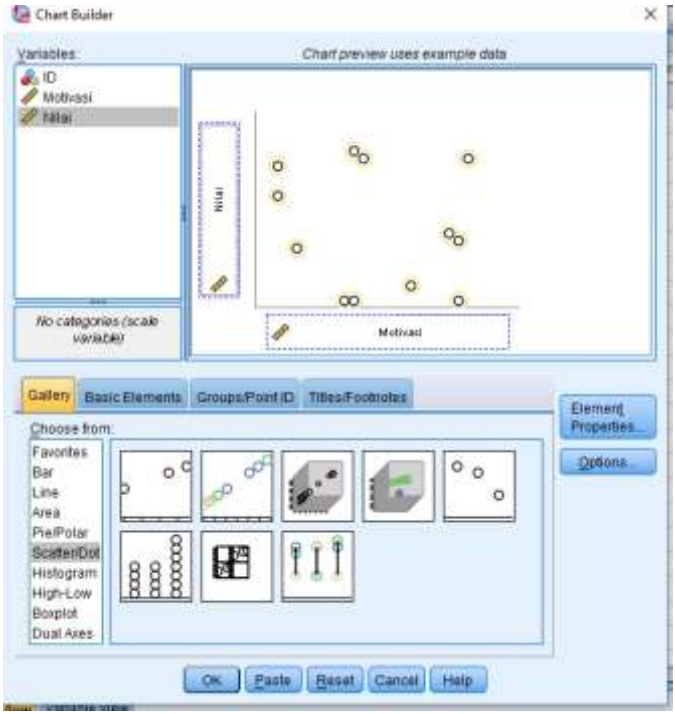
Rumusan Hipotesis pada Ilustrasi

H_0: $\rho = 0$ (tidak ada hubungan linear antara motivasi dan nilai).

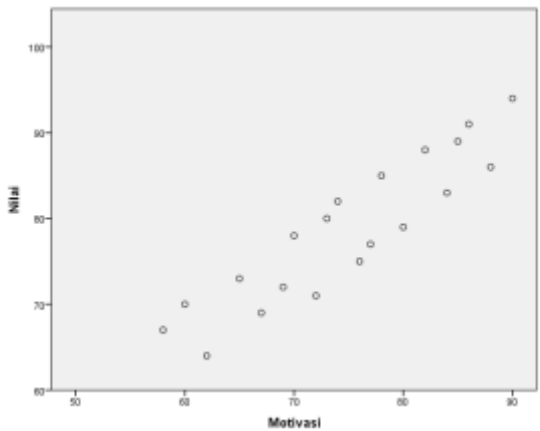
H_a: $\rho \neq 0$ (ada hubungan linear antara motivasi dan nilai).

Langkah Mengolah dan Mengkaji Data

1. Siapkan scatterplot melalui Graphs > Chart Builder untuk menilai bentuk hubungan dan pencilan.



Komponen Keluaran SPSS



Scatter plot menunjukkan bahwa terlihat bahwa titik-titik data membentuk pola yang cenderung naik dari kiri bawah

ke kanan atas. Hal ini menunjukkan adanya hubungan positif antara motivasi dan nilai. Semakin tinggi motivasi peserta, semakin tinggi pula nilai yang diperoleh. Pola hubungan tampak cukup linear dan tidak menunjukkan adanya pencilan ekstrem. Namun, untuk mengetahui kekuatan dan signifikansi hubungan tersebut, diperlukan uji korelasi.

2. Tinjau skala data serta distribusi jika Pearson akan digunakan.

Korelasi Pearson berfungsi untuk dua variabel numerik berskala interval atau rasio. Karena Motivasi dan Nilai berbentuk skor, keduanya dapat diatur sebagai Scale.

Asumsikan data berdistribusi normal

3. Pilih Analyze > Correlate > Bivariate.

4. Tempatkan motivasi dan nilai ke kotak Variables



5. Akses Pearson untuk hubungan linear data skala. Pilih Spearman untuk data ordinal atau hubungan monoton yang tidak memenuhi syarat Pearson.

6. Aktifkan Two-tailed dan Flag significant correlations.

Tekan OK sehingga muncul keluaran berikut:

Correlations

Correlations

		Motivasi	Nilai
Motivasi	Pearson Correlation	1	.910**
	Sig. (2-tailed)		.000
	N	20	20
Nilai	Pearson Correlation	.910**	1
	Sig. (2-tailed)	.000	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Tabel tersebut menunjukkan hasil uji korelasi Pearson antara variabel Motivasi dan Nilai pada 20 responden.

1. Nilai Pearson Correlation

   Nilai korelasi antara Motivasi dan Nilai yaitu r = 0,910

   Nilai positif menunjukkan bahwa hubungan kedua variabel **searah**. Artinya, semakin tinggi motivasi, nilai cenderung semakin tinggi. Sebaliknya, semakin rendah motivasi, nilai juga cenderung lebih rendah. Karena nilai korelasi mendekati 1, kekuatan hubungan antara Motivasi dan Nilai tergolong sangat kuat.

2. Nilai signifikansi

   Pada tabel diperoleh nilai Sig. (2-tailed) < 0,000, Dalam penulisan ilmiah, nilai tersebut ditulis sebagai:

   $$p < 0{,}001$$

   Karena p < 0,001, H_0 ditolak. Jadi, ditemukan hubungan linear positif yang signifikan antara motivasi dan nilai.

3. Jumlah sampel

   Nilai N=20 berarti pengolahan data dilakukan terhadap 20 responden.

4. Koefisien determinasi

   Besarnya variasi yang dimiliki bersama oleh kedua variabel dapat dihitung dengan:

$$r^2 = (0{,}910)^2 = 0{,}8281$$

$$0{,}8281 \times 100\% = 82{,}81\%$$

   Dalam praktik, motivasi dan nilai memiliki variasi bersama sekitar 82,81%. Sementara itu, sekitar 17,19% berkaitan dengan faktor lain atau variasi yang tidak dijelaskan oleh hubungan linear tersebut. Angka ini tidak boleh langsung ditafsirkan sebagai bukti bahwa motivasi menyebabkan 82,81% perubahan nilai.

Cara Membaca Keluaran

Pada dataset ilustratif, uji korelasi Pearson menghasilkan r(18) = 0,910, p < 0,001, dengan N = 20. Koefisien positif menunjukkan hubungan searah dan besarnya koefisien menunjukkan hubungan yang sangat kuat. Koefisien determinasi r^2 = 0,829 berarti sekitar 82,9% variasi kedua variabel saling berbagi secara statistik.

Arti Keputusan Statistik

Pada ilustrasi ini, karena p < 0,05, H_0 ditolak dan H_a didukung. Ada hubungan linear positif yang sangat kuat antara motivasi dan nilai. Peserta dengan motivasi lebih tinggi cenderung memperoleh nilai lebih tinggi, tetapi korelasi ini tidak membuktikan hubungan sebab-akibat.

LATIHAN

1. Paparkan perbedaan hubungan linear dan monoton!
2. Mengapa korelasi antara motivasi dan nomor peringkat dapat bernilai negatif?
3. Hitung koefisien determinasi untuk r = 0,65!
4. Berikan tiga alasan mengapa korelasi tidak membuktikan kausalitas!
5. Susun penyajian hasil korelasi Spearman yang lengkap!

Dataset Latihan Penerapan SPSS Bab 13.

Korelasi kecemasan dan nilai

ID	Kecemasan	Nilai
1	42	88
2	45	84
3	48	82
4	50	85
5	52	79
6	55	80
7	58	76
8	60	78
9	62	72
10	65	75
11	68	68
12	70	70
13	72	64
14	75	67

ID	Kecemasan	Nilai
15	78	60
16	80	62
17	82	58
18	85	55
19	88	52
20	90	50

Rumusan Hipotesis pada Ilustrasi

H_0: $\rho = 0$ (tidak ada hubungan antara kecemasan dan nilai).

H_a: $\rho \neq 0$ (ada hubungan antara kecemasan dan nilai).

1. Buat scatterplot dan identifikasi arah serta bentuk hubungan.
2. Jalankan Pearson dan Spearman.
3. Bandingkan koefisien dan p-value keduanya.
4. Tentukan teknik yang lebih tepat berdasarkan skala, linearitas, distribusi, dan pencilan.
5. Tulis penjelasan makna yang tidak mengubah korelasi menjadi klaim sebab-akibat.

BAB XIV
REGRESI LINEAR SEDERHANA, BERGANDA, DAN LOGISTIK

Capaian Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu: menentukan variabel bebas dan terikat; mengestimasi regresi linear sederhana dan berganda; menilai asumsi, keberartian, serta daya jelas model; menafsirkan koefisien regresi logistik dan odds ratio; menggunakan model untuk prediksi secara bertanggung jawab.

Kata Kunci

regresi, prediktor, variabel hasil, intersep, koefisien, R square, adjusted R square, residual, multikolinearitas, logit, odds ratio, klasifikasi.

A. Konsep Dasar Regresi

Regresi digunakan untuk mengkaji ketergantungan variabel terikat Y pada variabel bebas X. Regresi linear sederhana menggunakan satu prediktor dan satu variabel hasil. Tujuannya dapat berupa penjelasan hubungan linear, estimasi perubahan rata-rata Y untuk setiap satu unit X, atau prediksi (Field, 2024; Pallant, 2026).

Persamaan regresi tidak otomatis menyatakan sebab-akibat. Istilah variabel bebas dan terikat menggambarkan posisi dalam model, bukan jaminan manipulasi eksperimental (Field, 2024).

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

Persamaan regresi linear sederhana

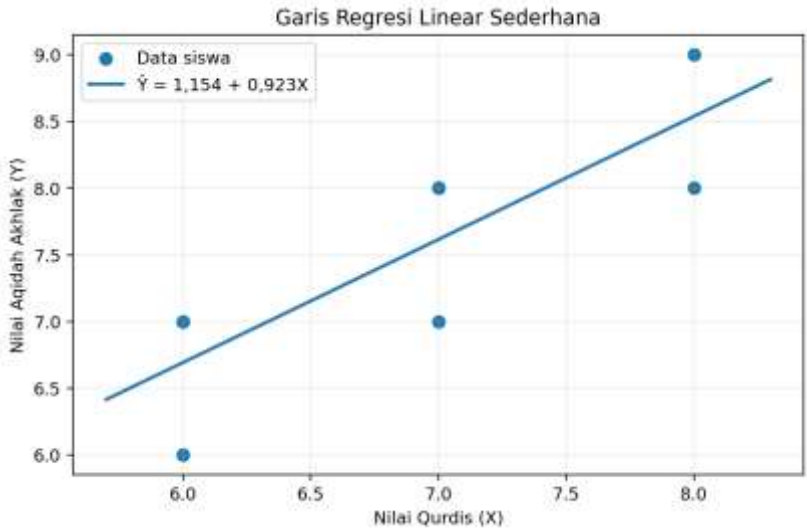
B. Estimasi Intersep dan Koefisien Arah

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

$$a = \bar{y} - b\bar{x}$$

Koefisien b adalah perubahan rata-rata Y yang diprediksi untuk kenaikan satu unit X. Intersep a adalah nilai prediksi Y ketika X = 0. Intersep hanya bermakna dalam konteks bila X = 0 berada dalam rentang yang relevan (Field, 2024).

Pada contoh nilai Qurdis dan Aqidah Akhlak, diperoleh a = 1,154 dan b = 0,923 sehingga $\hat{Y}$ = 1,154 + 0,923X. Untuk X = 7,2, nilai prediksi Y adalah 7,80.



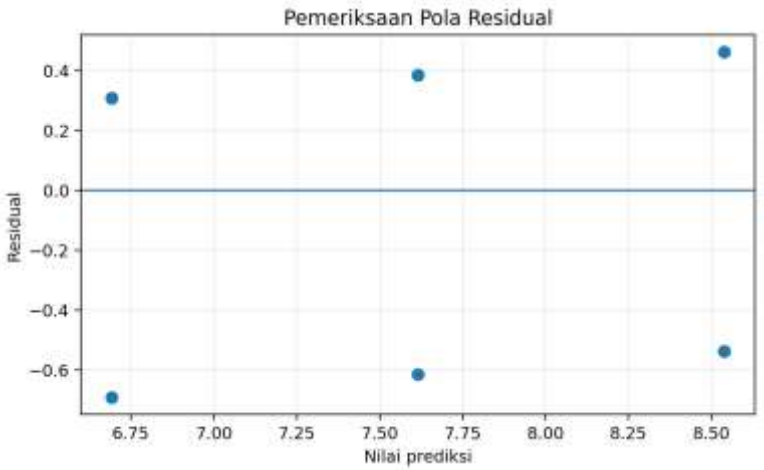
Gambar 14. 1. Scatterplot dan garis regresi nilai Qurdis terhadap Aqidah Akhlak

C. Residual dan Asumsi

Residual adalah selisih nilai observasi dan nilai prediksi, e = Y - Ŷ. Model ordinary least squares memilih garis yang meminimalkan jumlah kuadrat residual. Pemeriksaan residual penting untuk menilai apakah model linear cukup (Field, 2024).

1. Linearitas: hubungan rata-rata Y terhadap X mendekati garis lurus.
2. Independensi residual: observasi tidak saling bergantung.
3. Homoskedastisitas: variasi residual relatif konstan sepanjang nilai prediksi.
4. Normalitas residual: terutama penting untuk interval dan uji pada sampel kecil.

5. Tidak ada pengamatan sangat berpengaruh yang mendominasi garis.



Gambar 14. 2. Contoh plot residual terhadap nilai prediksi

D. Keberartian Model dan R Square

Uji F pada tabel ANOVA menilai apakah model dengan prediktor memberikan peningkatan penjelasan dibandingkan model tanpa prediktor. Untuk satu prediktor, uji F setara dengan uji t koefisien slope. H_0 menyatakan b = 0.

R^2 menyatakan proporsi variasi Y yang dijelaskan secara linear oleh X dalam sampel. Pada regresi sederhana, $R^2 = r^2$. Nilai tinggi tidak menjamin model benar, kausal, atau mampu memprediksi data baru. Adjusted R^2 lebih relevan ketika jumlah prediktor bertambah (Field, 2024; Triola, 2022).

$$R^2 = 1 - \frac{SSresidual}{SStotal}$$

E. Prediksi dan Bahaya Ekstrapolasi

Prediksi pada nilai X yang berada dalam rentang data disebut interpolasi dan umumnya lebih aman. Ekstrapolasi di luar rentang data berisiko karena pola linear mungkin tidak berlanjut. Prediksi individu memiliki interval prediksi yang lebih lebar daripada interval untuk mean respons (Field, 2024).

Model perlu divalidasi pada data baru atau melalui teknik cross-validation jika tujuan utama adalah prediksi. Model yang sangat cocok pada sampel dapat mengalami overfitting (Field, 2024).

Peringatan Ekstrapolasi

Jangan menggunakan persamaan dari rentang skor 6-8 untuk memprediksi hasil pada X = 0 atau X = 20 tanpa dasar yang jelas dan data pendukung.

F. Langkah SPSS Regresi Linear Sederhana

1. Pilih Analyze > Regression > Linear.
2. Masukkan variabel hasil ke Dependent dan prediktor ke Independent(s).
3. Pada Statistics, pilih Estimates, Model fit, Confidence intervals, dan Descriptives.
4. Pada Plots, masukkan ZPRED ke X dan ZRESID ke Y; minta normal probability plot.
5. Cermati Model Summary, ANOVA, Coefficients, residual plot, dan pengamatan berpengaruh.
6. Untuk uji linearitas eksploratif, dapat menggunakan Analyze > Compare Means > Means > Test for linearity, tetapi residual plot dan model regresi tetap penting (Pallant, 2026).

Contoh Penyusunan Kalimat Statistik

Nilai Qurdis memprediksi nilai Aqidah Akhlak secara signifikan, $F(1, 6) = 13{,}85$, $p = 0{,}010$, $R^2 = 0{,}69$. Persamaan model adalah $\hat{Y} = 1{,}154 + 0{,}923X$. Setiap kenaikan satu poin Qurdis berkaitan dengan peningkatan prediksi Aqidah Akhlak sebesar 0,923 poin.

G. Regresi Linear Berganda

Regresi linear berganda digunakan ketika satu variabel hasil numerik diprediksi oleh dua atau lebih prediktor. Bentuk umumnya adalah $\hat{Y} = a + b_1X_1 + b_2X_2 + \ldots + b_kX_k$. Setiap koefisien b menunjukkan perubahan nilai prediksi Y untuk kenaikan satu unit prediktor terkait ketika prediktor lain dipertahankan konstan.

Selain asumsi regresi linear, peneliti perlu memeriksa multikolinearitas. Tolerance yang sangat kecil atau Variance Inflation Factor (VIF) yang besar menunjukkan bahwa prediktor membawa informasi yang terlalu tumpang tindih. Sebagai pedoman praktis, VIF di bawah 5 umumnya tidak menunjukkan masalah berat, tetapi keputusan tetap mempertimbangkan konteks dan desain penelitian (Field, 2024; Pallant, 2026).

Langkah Menjalankan Regresi Linear Berganda

1. Pilih Analyze > Regression > Linear.
2. Tempatkan variabel hasil numerik ke Dependent dan seluruh prediktor ke Independent(s).
3. Pada Statistics, aktifkan Estimates, Model fit, Confidence intervals, Descriptives, Collinearity diagnostics, dan Durbin-Watson.
4. Pada Plots, tempatkan ZPRED pada X dan ZRESID pada Y, lalu aktifkan Histogram dan Normal probability plot.

5. Pada Save, simpan Standardized residuals dan Cook's distance untuk menilai pengamatan berpengaruh.
6. Baca Model Summary, ANOVA, Coefficients, Tolerance/VIF, plot residual, dan Cook's distance.

H. Regresi Logistik Biner

Regresi logistik biner digunakan ketika variabel hasil hanya memiliki dua kategori, misalnya tuntas = 1 dan belum tuntas = 0. Model tidak memprediksi skor secara langsung, melainkan probabilitas kejadian. Hubungan prediktor dengan probabilitas dinyatakan melalui logit: $\log[p/(1-p)] = a + b_1X_1 + ... + b_kX_k$.

Koefisien B menunjukkan perubahan log odds. Sementara itu, Exp(B) adalah odds ratio. Exp(B) lebih besar dari 1 menunjukkan kenaikan odds kejadian ketika prediktor meningkat satu unit; Exp(B) kurang dari 1 menunjukkan penurunan odds. Penjelasan makna harus menyebut kategori variabel hasil yang dikodekan sebagai 1.

Asumsi penting meliputi variabel hasil dikotomi, observasi independen, tidak ada multikolinearitas berat, hubungan linear antara prediktor kontinu dan logit, jumlah kejadian yang cukup, serta tidak terjadi complete separation. Kecocokan model dapat ditinjau melalui Omnibus Test, -2 Log Likelihood, pseudo R^2, Hosmer-Lemeshow, tabel klasifikasi, dan Variables in the Equation (Hosmer *et al*., 2013).

Langkah Menjalankan Regresi Logistik Biner

1. Pastikan variabel hasil dikodekan 0 dan 1 serta kategori kejadian yang dikaji adalah 1.
2. Pilih Analyze > Regression > Binary Logistic.

3. Masukkan variabel hasil ke Dependent dan prediktor ke Covariates.
4. Gunakan Method = Enter untuk memasukkan seluruh prediktor secara simultan; klik Categorical jika ada prediktor kategorik.
5. Pada Options, aktifkan Hosmer-Lemeshow goodness-of-fit dan Confidence interval for Exp(B) sebesar 95%.
6. Baca Omnibus Tests, Model Summary, Hosmer-Lemeshow Test, Classification Table, dan Variables in the Equation.

I. Pemilihan Model dan Prinsip Penyajian Hasil

Gunakan regresi linear sederhana jika satu variabel digunakan untuk memprediksi variabel hasil numerik. Gunakan regresi linear berganda jika ada beberapa prediktor. Gunakan regresi logistik biner jika variabel hasil hanya memiliki dua kategori. Pilih model berdasarkan tujuan, skala data, desain, asumsi, dan kualitas data, bukan karena menunya paling mudah.

CONTOH PENERAPAN BAB 14

Regresi Linear Sederhana untuk Prediksi Nilai

Catatan pembelajaran: *Seluruh data dan keluaran pada bagian ini bersifat fiktif dan digunakan untuk memperagakan langkah pengolahan data. Angka yang ditampilkan bukan laporan penelitian empiris dan tidak mewakili individu, kelas, sekolah, perguruan tinggi, atau populasi tertentu.*

Ilustrasi kasus: peneliti memprediksi nilai matematika berdasarkan jumlah jam belajar mandiri per minggu.

Tabel 14. 1. Ilustrasi 14.1. Data fiktif regresi linear sederhana

ID	Jam belajar	Nilai
1	2	58
2	3	64
3	4	61
4	5	70
5	6	66
6	7	75
7	8	71
8	9	81
9	10	76
10	11	86
11	12	82
12	3	59
13	5	72
14	7	69
15	9	84
16	11	79
17	4	67
18	6	65
19	8	78
20	10	74
21	12	90
22	5	63
23	8	73
24	11	88

ID	Jam belajar	Nilai
25	13	85

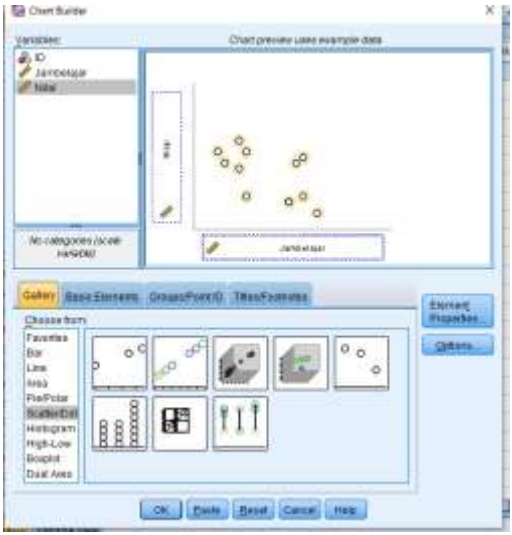
Rumusan Hipotesis pada Ilustrasi

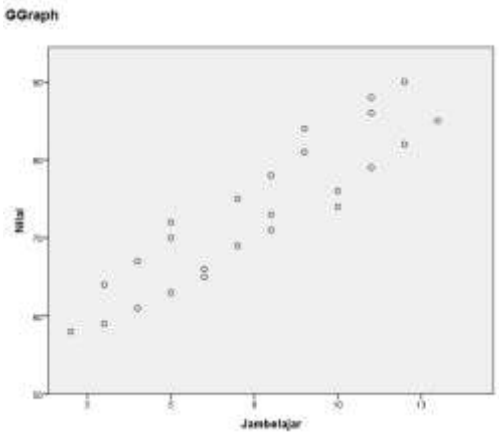
H_0: $\beta_1 = 0$ (jam belajar tidak menjadi prediktor linear nilai).

H_a: $\beta_1 \neq 0$ (jam belajar menjadi prediktor linear nilai).

Langkah Mengolah dan Mengkaji Data

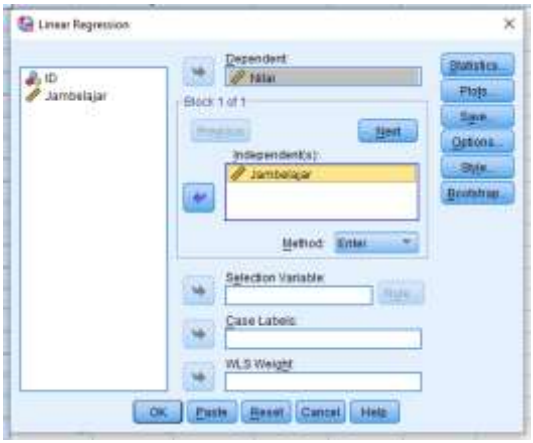
1. Buat scatterplot untuk menilai linearitas dan pencilan.



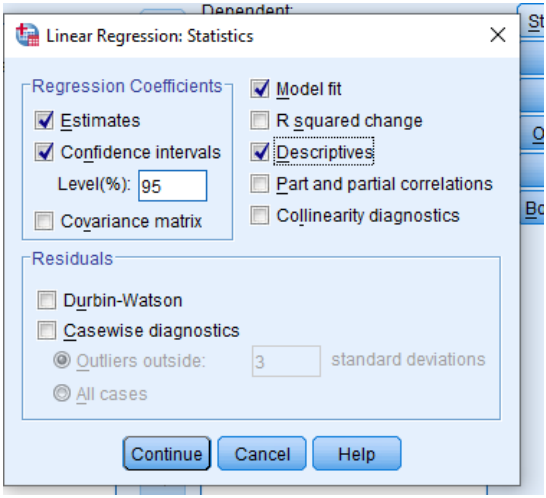


Dari scatterplot, terlihat bahwa ada kecenderungan hubungan positif dan linear antara jam belajar dan nilai. Semakin banyak jam belajar, semakin tinggi nilai yang diperoleh. Namun, kepastian kekuatan dan signifikansi hubungan perlu dilihat dari nilai koefisien korelasi dan nilai signifikansinya.

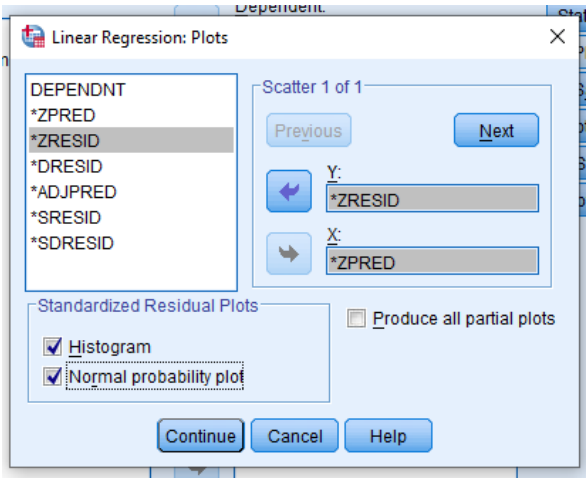
2. Pilih Analyze > Regression > Linear.
3. Masukkan nilai sebagai Dependent dan jam_belajar sebagai Independent.



4. Pada Statistics, aktifkan Estimates, Model fit, Confidence intervals, dan Descriptives.



5. Pada Plots, masukkan ZPRED pada X dan ZRESID pada Y. Aktifkan histogram serta normal probability plot.



6. Pada Save, simpan standardized predicted values dan standardized residuals jika diperlukan.



7. Tekan Ok, dan muncul keluaran berikut:

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
Nilai	73.44	9.328	25
Jambelajar	7.56	3.190	25

Secara umum responden memiliki nilai rata-rata 73,44 dan waktu belajar rata-rata 7,56 jam. Standar deviasi menunjukkan tingkat penyebaran data dari nilai rata-ratanya.

Correlations

		Nilai	Jambelajar
Pearson Correlation	Nilai	1.000	.909
	Jambelajar	.909	1.000
Sig. (1-tailed)	Nilai	.	.000
	Jambelajar	.000	.
N	Nilai	25	25
	Jambelajar	25	25

Nilai korelasi sebesar 0,909 menunjukkan hubungan positif yang sangat kuat. Artinya, semakin lama jam belajar, nilai cenderung semakin tinggi.

Karena nilai signifikansi lebih kecil dari 0,05, hubungan tersebut signifikan secara statistik. Namun, karena output menggunakan Sig. (1-tailed), hasil ini tepat digunakan jika sejak awal hipotesis penelitian bersifat searah, misalnya:

Semakin tinggi jam belajar, semakin tinggi nilai.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Jambelajar[b]	.	Enter

a. Dependent Variable: Nilai

b. All requested variables entered.

Metode Enter berarti variabel Jam Belajar dimasukkan langsung ke dalam model regresi. Karena hanya ada satu variabel bebas, prosedur ini termasuk regresi linear sederhana.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.909[a]	.826	.818	3.978

a. Predictors: (Constant), Jambelajar

b. Dependent Variable: Nilai

Dari tabel Model Summary, terlihat bahwa diperoleh nilai R sebesar 0,909 yang menunjukkan hubungan sangat kuat antara jam belajar dan nilai. Nilai R Square sebesar 0,826 menunjukkan bahwa jam belajar dapat menjelaskan 82,6% variasi nilai. Sementara itu, 17,4% sisanya dijelaskan oleh faktor lain di luar model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1724.252	1	1724.252	108.977	.000[b]
	Residual	363.908	23	15.822		
	Total	2088.160	24			

a. Dependent Variable: Nilai

b. Predictors: (Constant), Jambelajar

Tabel ANOVA menunjukkan bahwa model regresi signifikan karena nilai p lebih kecil dari 0,05. Jadi, jam belajar dapat dipakai untuk memprediksi nilai pada rentang data yang dikaji.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	53.350	2.082		25.619	.000	49.042	57.658
	Jambelajar	2.657	.255	.909	10.439	.000	2.131	3.184

a. Dependent Variable: Nilai

Koefisien B sebesar 2,657 berarti setiap penambahan 1 jam belajar diperkirakan meningkatkan nilai sebesar 2,657 poin, dengan asumsi faktor lain tetap. Karena $p <$ 0,001, koefisien tersebut signifikan. Interval kepercayaan

juga tidak melewati angka nol, sehingga memperkuat adanya hubungan positif yang signifikan. Nilai Beta standar sebesar 0,909 menunjukkan bahwa kenaikan satu standar deviasi jam belajar berkaitan dengan kenaikan nilai sebesar 0,909 standar deviasi.

Persamaan regresinya adalah:

$$Y = 53{,}350 + 2{,}657X$$

Keterangan:

- Y = nilai yang diprediksi
- X = jam belajar
- 53,350 = konstanta
- 2,657 = koefisien regresi jam belajar

Contoh, jika seorang peserta belajar selama 8 jam:

$$Y = 53{,}350 + 2{,}657(8)$$

Jadi, nilai yang diprediksi sekitar 74,61.

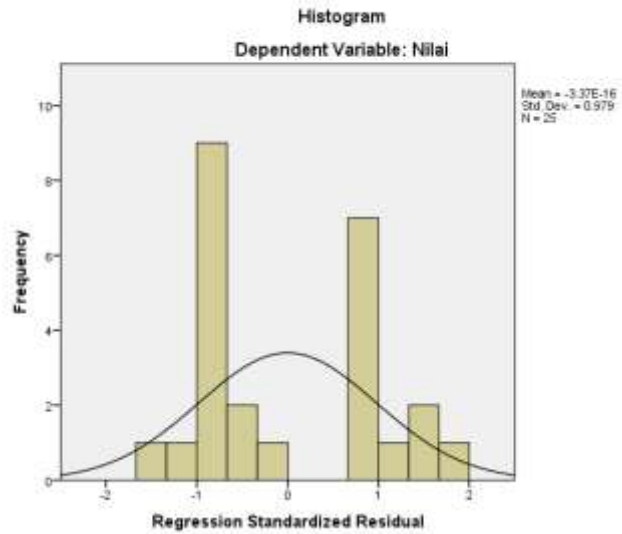
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	58.66	87.90	73.44	8.476	25
Residual	-5.924	6.733	.000	3.894	25
Std. Predicted Value	-1.743	1.706	.000	1.000	25
Std. Residual	-1.489	1.693	.000	.979	25

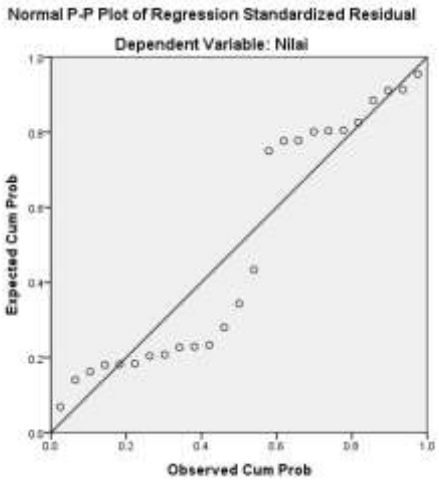
a. Dependent Variable: Nilai

Tabel Residuals Statistics menunjukkan rata-rata residual 0,000. Standardized residual berada antara -1,489 dan 1,693. Semua nilainya masih berada dalam batas ±3, sehingga tidak terlihat residual ekstrem. Secara umum, kesalahan prediksi model masih dalam batas yang wajar.

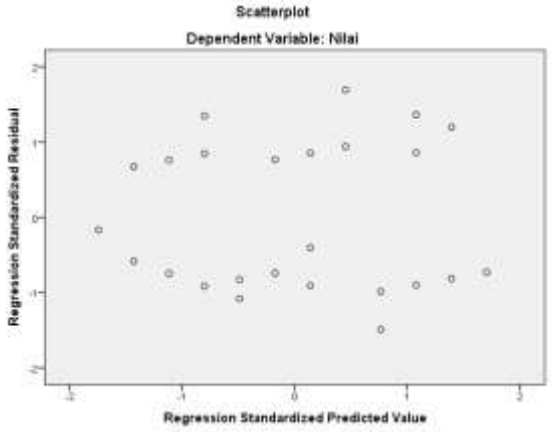
Charts



Histogram menunjukkan bahwa residual memiliki rata-rata yang mendekati nol dan tidak menunjukkan pencilan ekstrem. Namun, penyebarannya belum membentuk pola kurva normal yang jelas dan cenderung memiliki dua puncak. Karena itu, asumsi normalitas residual belum dapat dipastikan hanya melalui histogram. Pemeriksaan perlu dilengkapi dengan Normal P-P plot atau Q-Q plot serta uji Shapiro-Wilk terhadap residual.



Normal P-P plot menunjukkan bahwa titik-titik residual tidak menyebar secara konsisten di sekitar garis diagonal dan menunjukkan penyimpangan yang cukup besar, terutama pada bagian tengah grafik. Jadi, secara visual residual belum memenuhi asumsi normalitas.



Pada scatterplot ilustratif, titik-titik residual tersebar di atas dan di bawah angka nol serta tidak membentuk pola corong yang jelas. Karena itu, tidak ada indikasi heteroskedastisitas yang kuat dan asumsi homoskedastisitas

relatif terpenuhi. Meskipun demikian, adanya dua kelompok residual menunjukkan bahwa penyebaran residual belum sepenuhnya acak, sehingga model perlu ditinjau bersama hasil uji normalitas residual dan kemungkinan adanya variabel lain yang belum dimasukkan dalam model.

Cara Membaca Keluaran

Pada dataset ilustratif, model regresi linear sederhana signifikan, $F(1, 23) = 108{,}98$, $p < 0{,}001$. Model menjelaskan 82,6% variasi nilai ($R^2 = 0{,}826$). Koefisien jam belajar adalah $B = 2{,}657$ dengan $p < 0{,}001$ dan interval kepercayaan 95% [2,131, 3,184]. Persamaan modelnya adalah $\hat{Y} = 53{,}350 + 2{,}657X$. Dalam rentang data ini, tambahan satu jam belajar berkaitan dengan kenaikan nilai sekitar 2,657 poin. Hasil prediksi tetap perlu dibaca hati-hati karena model tidak boleh digunakan di luar rentang data.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0{,}05$, H_0 ditolak dan H_a didukung. Jam belajar adalah prediktor linear positif yang signifikan terhadap nilai, dan model menjelaskan 82,6% variasi nilai pada sampel. Namun, hasil prediksi perlu ditafsirkan secara hati-hati karena normalitas residual belum sepenuhnya terpenuhi dan model tidak boleh diekstrapolasi di luar rentang data.

Jangan melakukan ekstrapolasi jauh di luar rentang 1 sampai 13 jam pada data. Intersep juga tidak selalu memiliki makna pendidikan yang kuat jika nilai X = 0 tidak teramati.

B. Regresi Linear Berganda untuk Memprediksi Nilai

Ilustrasi kasus: peneliti memprediksi nilai matematika berdasarkan motivasi belajar, persentase kehadiran, dan jam belajar mandiri per minggu.

Tabel 14. 2. Ilustrasi 14.2. Data fiktif regresi linear berganda

ID	Motivasi	Kehadiran (%)	Jam belajar	Nilai
1	61	100	4,2	71
2	63	84	5,5	75
3	62	96	7,3	78
4	69	98	5,5	77
5	89	72	4,6	74
6	59	97	5,4	77
7	84	87	5,6	73
8	88	80	5,2	81
9	71	84	6,9	79
10	68	92	10,0	81
11	80	91	7,3	81
12	87	91	10,1	87
13	95	83	8,1	86
14	84	97	8,3	90
15	91	99	7,8	85
16	70	100	5,8	80
17	66	95	7,4	82
18	58	100	7,2	79
19	63	100	4,1	78
20	71	93	8,1	80
21	77	94	7,8	85
22	62	92	7,3	86
23	83	92	5,3	79
24	74	94	1,0	67
25	69	76	9,7	79
26	67	86	6,3	76
27	95	91	1,7	72
28	77	73	7,9	76
29	66	78	5,0	63
30	65	99	8,0	76

Rumusan Hipotesis pada Ilustrasi

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ (motivasi, kehadiran, dan jam belajar secara simultan tidak memprediksi nilai).

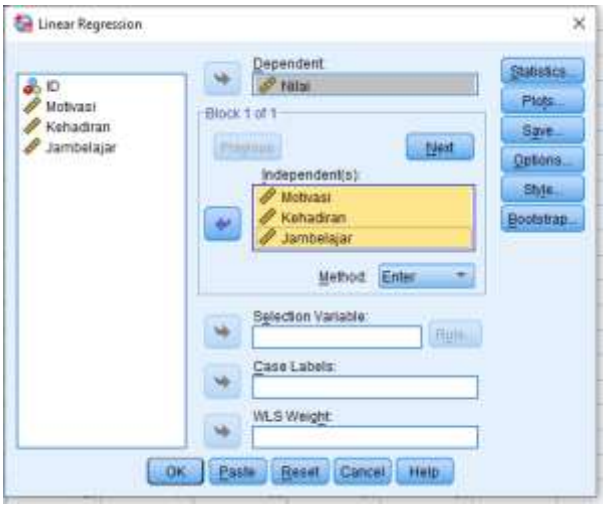
H_a: setidaknya satu koefisien regresi tidak sama dengan 0.

Langkah Mengolah dan Mengkaji Data

1. Atur Nilai, Motivasi, Kehadiran, dan Jam belajar sebagai Scale; ID sebagai Nominal.
2. Tinjau scatterplot, pencilan, dan korelasi antarprediktor. Pastikan tidak ada hubungan yang menunjukkan multikolinearitas berat.

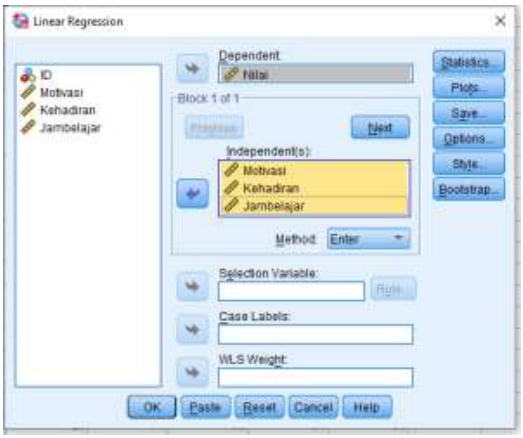
Cara membuat scatterplot: analyze > regression > linear

- ❖ Masukkan:
  - **Nilai** ke kotak **Dependent**
  - **Motivasi, Kehadiran, dan Jam belajar** ke kotak **Independent(s)**
- ❖ Klik **Plots**
- ❖ Masukkan:
  - *ZPRED ke kotak**X**
  - *ZRESID ke kotak**Y**
- ❖ Klik **Continue**
- ❖ Klik **OK**

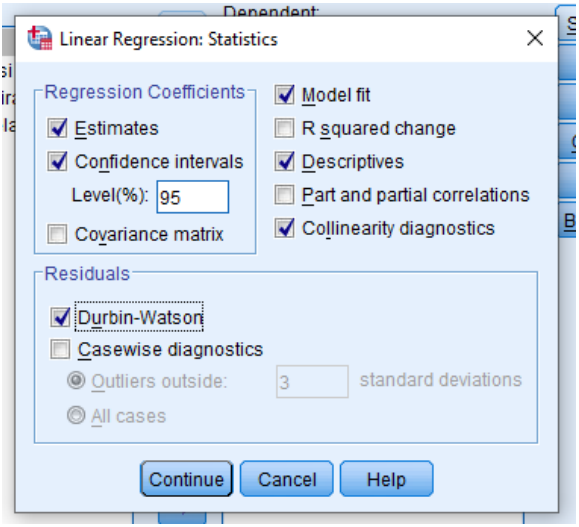


Akan muncul di bagian data view ZPRED dan ZRESID

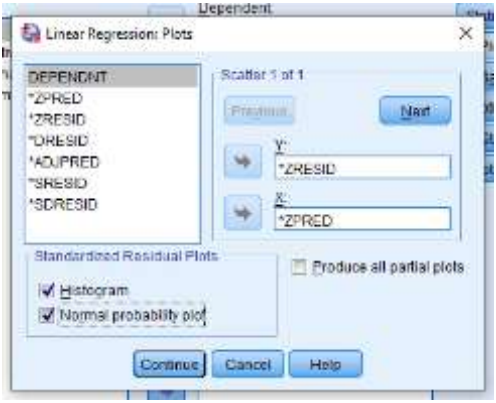
3. Pilih Analyze > Regression > Linear.
4. Tempatkan Nilai ke Dependent serta Motivasi, Kehadiran, dan Jam belajar ke Independent(s). Gunakan Method = Enter.



5. Pada Statistics, aktifkan Estimates, Model fit, Confidence intervals, Descriptives, Collinearity diagnostics, dan Durbin-Watson.



6. Pada Plots, masukkan ZPRED ke X dan ZRESID ke Y, lalu aktifkan Histogram dan Normal probability plot.



7. Pada Save, aktifkan Standardized residuals dan Cook's distance. Klik OK.



8. Baca Model Summary, ANOVA, Coefficients, Tolerance/VIF, serta plot residual sebelum menyusun kesimpulan.

Diperoleh keluaran sebagai berikut:

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Jambelajar, Motivasi, Kehadiran[b]	.	Enter

a. Dependent Variable: Nilai

b. All requested variables entered.

Seluruh variabel bebas, yaitu jam belajar, motivasi, dan kehadiran, dimasukkan secara simultan untuk memprediksi nilai menggunakan metode Enter.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.832[a]	.693	.657	3.431

a. Predictors: (Constant), Jambelajar, Motivasi, Kehadiran

b. Dependent Variable: Nilai

Tabel Model Summary menunjukkan R = 0,832. Nilai ini menunjukkan hubungan yang sangat kuat antara seluruh prediktor dan nilai. Adjusted R^2 = 0,657 berarti motivasi, kehadiran, dan jam belajar bersama-sama menjelaskan 65,7% variasi nilai. Sisanya, 34,3%, berkaitan dengan faktor lain yang tidak masuk dalam model.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	689.332	3	229.777	19.521	.000[b]
	Residual	306.035	26	11.771		
	Total	995.367	29			

a. Dependent Variable: Nilai

b. Predictors: (Constant), Jambelajar, Motivasi, Kehadiran

$$p < 0{,}001 < 0{,}05$$

Karena p < 0,001, H_0 ditolak.

Cara membacanya, jam belajar, motivasi, dan kehadiran secara bersama-sama adalah prediktor yang signifikan terhadap nilai. Model regresi secara umum layak digunakan.

Penulisannya dapat diringkas sebagai:

$$F(3,26) = 19{,}521, p < 0{,}001$$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	24.140	10.063		2.399	.024
	Motivasi	.215	.060	.411	3.569	.001
	Kehadiran	.284	.081	.405	3.516	.002
	Jambelajar	1.976	.299	.721	6.612	.000

a. Dependent Variable: Nilai

1. Tabel Coefficients

Tabel ini berfungsi untuk mengetahui pengaruh masing-masing variabel independen terhadap Nilai, setelah pengaruh variabel independen lainnya dikendalikan.

Persamaan regresi

Dari kolom Unstandardized Coefficients B, terlihat bahwa persamaan regresinya adalah:

$$\hat{Y} = 24{,}140 + 0{,}215X_1 + 0{,}284X_2 + 1{,}976X_3$$

Keterangan:

- $\hat{Y}$= Nilai yang diprediksi
- X_1= Motivasi
- X_2= Kehadiran
- X_3= Jam belajar

Konstanta

Nilai konstanta sebesar 24,140, dengan p = 0,024

Dalam praktik, jika motivasi, kehadiran, dan jam belajar bernilai nol, nilai yang diprediksi adalah 24,140. Penjelasan makna ini bersifat matematis dan tidak selalu memiliki makna praktis jika nilai nol tidak termasuk dalam rentang data.

Motivasi

Hasil variabel motivasi:

- B=0,215
- Beta = 0,411
- t=3,569
- p=0,001

Karena p < 0,05, motivasi berpengaruh positif dan signifikan terhadap nilai. Setiap peningkatan satu satuan motivasi diperkirakan meningkatkan nilai sebesar 0,215 poin, dengan kehadiran dan jam belajar dianggap tetap.

Kehadiran

Hasil variabel kehadiran:

- B=0,284
- Beta = 0,405
- t=3,516
- p=0,002

Karena p < 0,05, kehadiran berpengaruh positif dan signifikan terhadap nilai. Setiap peningkatan satu satuan kehadiran diperkirakan meningkatkan nilai sebesar 0,284 poin, dengan motivasi dan jam belajar dianggap tetap.

Jam belajar

Hasil variabel jam belajar:

- B = 1,976.
- Beta = 0,721.
- t = 6,612.
- p < 0,001.

Jam belajar berpengaruh positif dan signifikan terhadap nilai. Setiap penambahan satu jam belajar diperkirakan meningkatkan nilai sebesar 1,976 poin, dengan motivasi dan kehadiran dikendalikan.

Pembacaan terhadap nilai Standardized Beta menunjukkan bahwa jam belajar adalah prediktor yang kontribusi relatifnya paling besar:

$$Jambelajar(0{,}721) > Motivasi(0{,}411) > Kehadiran(0{,}405)$$

Pernyataan bahwa jam belajar adalah prediktor dominan harus dipastikan setelah asumsi multikolinearitas diperiksa melalui nilai Tolerance dan VIF.

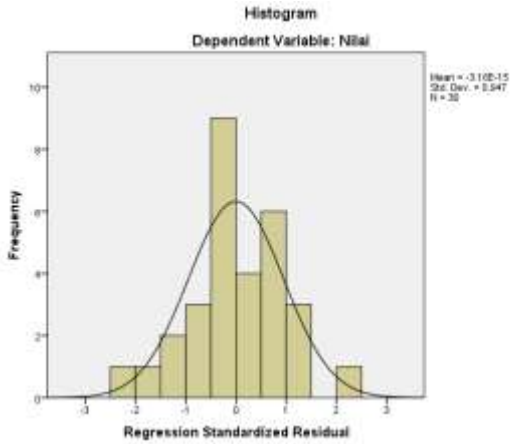
Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	68.65	88.57	78.43	4.875	30
Std. Predicted Value	-2.006	2.079	.000	1.000	30
Standard Error of Predicted Value	.784	1.972	1.210	.331	30
Adjusted Predicted Value	69.24	88.95	78.51	4.847	30
Residual	-7.300	8.044	.000	3.249	30
Std. Residual	-2.128	2.345	.000	.947	30
Stud. Residual	-2.362	2.440	-.010	1.013	30
Deleted Residual	-8.992	8.709	-.073	3.723	30
Stud. Deleted Residual	-2.613	2.724	-.010	1.067	30
Mahal. Distance	.549	8.612	2.900	2.134	30
Cook's Distance	.000	.323	.037	.062	30
Centered Leverage Value	.019	.297	.100	.074	30

a. Dependent Variable: Nilai

Tabel Residuals Statistics tidak menunjukkan residual ekstrem karena standardized residual, studentized residual, dan studentized deleted residual masih berada dalam batas ±3. Mahalanobis Distance juga tidak menunjukkan pencilan multivariat yang ekstrem. Namun, Cook's Distance maksimum 0,323 dan leverage maksimum 0,297 menunjukkan bahwa ada data yang mungkin cukup berpengaruh. Data tersebut perlu diperiksa, tetapi tidak boleh langsung dihapus tanpa alasan.

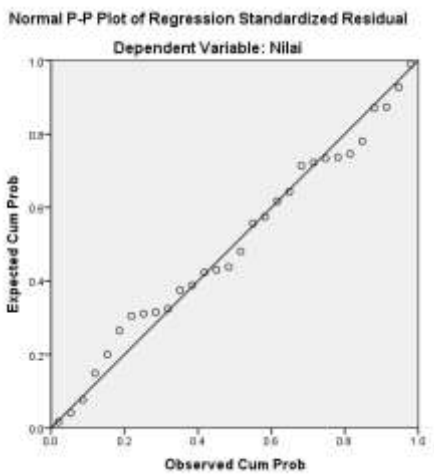
Tabel ini belum cukup untuk menilai normalitas dan heteroskedastisitas. Normalitas perlu diperiksa melalui histogram atau P-P plot residual. Sementara itu, heteroskedastisitas diperiksa melalui scatterplot antara nilai prediksi dan residual.



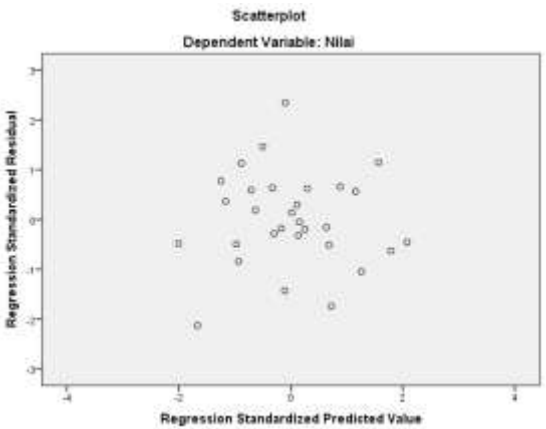
Informasi pada grafik menunjukkan:

- Mean residual = $-3{,}16 \times 10^{-15}$, atau secara praktis mendekati 0.
- Standar deviasi = 0,947, mendekati 1.
- Jumlah data = 30 responden.

Walaupun bentuk histogram tidak sepenuhnya simetris karena jumlah sampel relatif terbatas, secara umum distribusinya masih dapat dianggap mendekati distribusi normal. Hasil ini memperkuat kesimpulan dari grafik P–P Plot bahwa asumsi normalitas residual terpenuhi.



Pada grafik Normal P–P Plot, titik-titik residual terlihat mengikuti dan berada cukup dekat dengan garis diagonal. Walaupun ada sedikit penyimpangan pada beberapa bagian, penyimpangannya tidak terlalu jauh atau membentuk pola yang ekstrem. Artinya, residual regresi dapat dianggap berdistribusi normal. Jadi, asumsi normalitas residual pada model regresi telah terpenuhi.



Scatterplot digunakan untuk melihat ada atau tidaknya gejala heteroskedastisitas. Pada grafik, titik-titik residual:

- Tersebar di atas dan di bawah nilai 0;
- tidak membentuk pola bergelombang;
- Tidak membentuk pola mengerucut atau melebar seperti corong;
- Tidak menunjukkan pola tertentu yang sistematis.

Makna keputusan pada ilustrasi ini, secara statistik dapat dinyatakan bahwa tidak tampak gejala heteroskedastisitas. Varians residual cenderung konstan pada berbagai nilai prediksi, sehingga asumsi homoskedastisitas terpenuhi.

Cara Membaca Keluaran

Pada dataset ilustratif, model regresi linear berganda signifikan, $F(3, 26) = 19{,}521$, $p < 0{,}001$. Model menjelaskan 69,3% variasi nilai ($R^2 = 0{,}693$; adjusted $R^2 = 0{,}657$). Motivasi, kehadiran, dan jam belajar memiliki koefisien positif serta signifikan. Persamaan modelnya adalah $\hat{Y} = 24{,}140 + 0{,}215X_1 + 0{,}284X_2 + 1{,}976X_3$. Nilai beta standar terbesar ada pada jam belajar. Pemeriksaan residual tidak menunjukkan masalah berat, tetapi data yang berpengaruh tetap perlu diperiksa.

Arti Keputusan Statistik

Pada ilustrasi ini, karena $p < 0,05$, H_0 ditolak dan H_a didukung. Motivasi, kehadiran, dan jam belajar secara simultan memprediksi nilai secara signifikan serta menjelaskan 69,3% variasi nilai pada sampel. Ketiga prediktor juga signifikan secara parsial. Jam belajar memiliki beta standar terbesar, tetapi hasil tersebut tidak dengan sendirinya membuktikan pengaruh kausal.

C. Regresi Logistik Biner untuk Memprediksi Ketuntasan

Skenario penelitian: peneliti memprediksi peluang mahasiswa mencapai ketuntasan berdasarkan motivasi dan jam belajar. Variabel hasil dikodekan 0 = belum tuntas dan 1 = tuntas.

Tabel 14. 3. Ilustrasi 14.3. Data fiktif regresi logistik biner

ID	Motivasi	Jam belajar	Tuntas
1	86	4,6	0
2	78	9,4	1
3	70	4,5	0
4	62	4,3	1
5	56	6,3	1
6	75	7,0	1
7	67	9,1	1
8	66	4,3	0
9	73	2,6	1
10	74	6,2	1
11	52	8,1	1
12	84	2,8	0
13	55	5,3	0
14	94	7,3	1
15	81	7,1	1
16	70	6,7	1
17	88	5,9	1
18	75	4,5	0
19	82	4,2	0
20	76	9,4	1
21	86	3,7	0
22	86	9,6	1
23	66	3,2	0
24	69	12,0	1

ID	Motivasi	Jam belajar	Tuntas
25	78	2,2	0
26	73	10,5	1
27	52	4,3	0
28	95	6,2	1
29	53	6,3	0
30	54	4,9	0
31	62	12,0	1
32	81	6,7	0
33	93	3,2	1
34	73	8,1	0
35	71	8,3	0
36	77	3,6	0
37	71	7,8	0
38	76	6,7	0
39	64	8,6	1
40	79	8,3	1
41	79	6,5	1
42	51	5,5	0
43	74	8,8	1
44	77	7,3	1
45	81	2,8	1
46	67	5,1	0
47	62	5,3	0
48	65	5,8	0
49	80	12,0	1
50	66	5,6	0

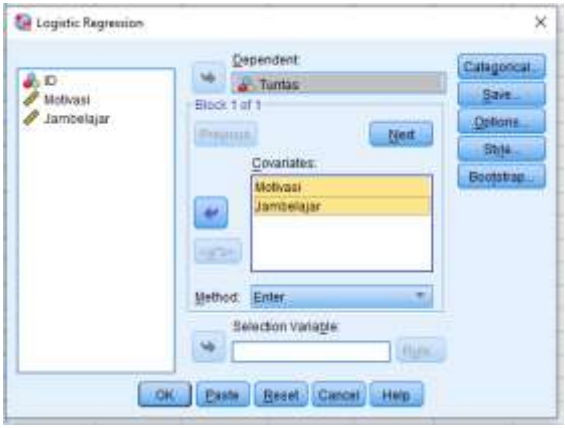
Rumusan Hipotesis pada Ilustrasi Uji Omnibus

H_0: βMotivasi = βJam belajar = 0 (model dengan prediktor tidak lebih baik daripada model konstanta).

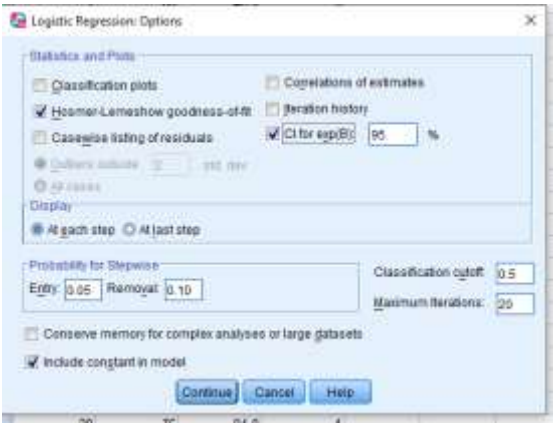
H_a: setidaknya satu koefisien regresi logistik tidak sama dengan 0.

Langkah Mengolah dan Mengkaji Data

1. Pada Variable View, atur Tuntas sebagai Nominal dan beri value labels 0 = belum tuntas, 1 = tuntas. Motivasi dan Jam belajar diatur sebagai Scale.
2. Pilih Analyze > Regression > Binary Logistic.
3. Isikan Tuntas ke Dependent serta Motivasi dan Jam belajar ke Covariates. Gunakan Method = Enter.



4. Pada Options, aktifkan Hosmer-Lemeshow goodness-of-fit dan CI for Exp(B) sebesar 95%.



5. Pilih OK. Pastikan kategori kejadian yang dimodelkan adalah Tuntas = 1.

Diperoleh keluaran sebagai berikut:

Logistic Regression

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	50	100.0
	Missing Cases	0	.0
	Total	50	100.0
Unselected Cases		0	.0
Total		50	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
Belum Tuntas	0
Tuntas	1

Tabel Case Processing Summary menunjukkan bahwa seluruh 50 data digunakan dan tidak ada data yang hilang. Pada tabel Dependent Variable Encoding, kategori Belum Tuntas diberi kode 0 dan Tuntas diberi kode 1. Kode ini menentukan kategori yang diprediksi oleh model. Karena Tuntas diberi kode 1, koefisien positif menunjukkan kenaikan odds untuk menjadi tuntas.

Block 0: Beginning Block

Classification Table[a,b]

Observed			Predicted		
			Tuntas		Percentage Correct
			Belum Tuntas	Tuntas	
Step 0	Tuntas	Belum Tuntas	0	24	.0
		Tuntas	0	26	100.0
	Overall Percentage				52.0

a. Constant is included in the model.

b. The cut value is .500

Tabel ini menunjukkan kemampuan model awal (tanpa variabel prediktor) dalam mengklasifikasikan responden ke dalam kategori variabel dependen.

Dari 50 responden:

- Belum Tuntas = 24 orang (48%)
- Tuntas = 26 orang (52%)

Karena kategori Tuntas lebih banyak, maka model kosong memprediksi semua responden sebagai Tuntas.

Nilai:

Overall Percentage=52%

Artinya:

Tanpa memasukkan variabel independen, model sudah mampu mengklasifikasikan data dengan tingkat ketepatan sebesar 52%.

Nilai 52% ini menjadi nilai dasar (baseline accuracy) yang nantinya dibandingkan dengan model setelah variabel independen dimasukkan

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.080	.283	.080	1	.777	1.083

Nilai B menunjukkan koefisien regresi logistik dari konstanta, persamaan awal Logit (P) = 0,080

Karena belum ada variabel bebas, maka model hanya memperkirakan peluang berdasarkan konstanta

Nilai signifikansi konstanta adalah 0,777. Karena p > 0,05, konstanta pada model awal belum signifikan. Artinya, pada model awal tanpa variabel independen, konstanta saja belum mampu menjelaskan status ketuntasan secara signifikan.

Nilai Exp(B) atau odds ratio pada model awal adalah 1,083, yang menggambarkan odds tuntas dibandingkan belum tuntas sebesar 1,083. Namun, karena konstanta tidak signifikan, nilai ini belum memiliki makna pengaruh

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	Motivasi	3.456	1	.063
		Jambelajar	12.236	1	.000
	Overall Statistics		16.269	2	.000

Tabel ini menunjukkan variabel independen yang belum dimasukkan ke model, tetapi diuji apakah memiliki potensi untuk meningkatkan model.

Variabel:

1. Motivasi

   Nilai signifikansi motivasi pada tahap awal adalah 0,063. Karena nilainya lebih besar dari 0,05, motivasi belum signifikan pada tahap ini. Keputusan memasukkan prediktor tetap perlu mengikuti kerangka teori dan tujuan model, bukan hanya kedekatannya dengan batas 0,05.

2. Jam belajar

   Jam belajar menghasilkan p < 0,001 pada pengujian awal. Hasil ini menunjukkan bahwa jam belajar berpotensi meningkatkan kemampuan model dalam memprediksi status ketuntasan.

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	18.818	2	.000
	Block	18.818	2	.000
	Model	18.818	2	.000

Tabel Omnibus Test berfungsi untuk mengetahui apakah model regresi logistik yang sudah memasukkan variabel independen lebih baik dibandingkan model awal (Block 0) yang hanya menggunakan konstanta.

Karena $p < 0{,}001$, H_0 ditolak. Model yang memasukkan motivasi dan jam belajar secara simultan lebih baik daripada model yang hanya memuat konstanta.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	50.417[a]	.314	.418

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Tabel Model Summary digunakan untuk mengetahui seberapa besar kemampuan variabel independen dalam menjelaskan variasi variabel dependen.

Nilai -2 Log Likelihood pada model akhir adalah 50,417. Ukuran ini dipakai untuk membandingkan kecocokan model, dengan nilai yang lebih kecil menunjukkan model yang relatif lebih baik. Penurunan -2 Log Likelihood dari Block 0 ke Block 1 menunjukkan bahwa penambahan motivasi dan jam belajar memperbaiki model.

Nilai Cox & Snell $R^2 = 0{,}314$ dan Nagelkerke $R^2 = 0{,}418$ adalah ukuran pseudo-R^2 yang menunjukkan peningkatan daya jelas relatif model setelah motivasi dan jam belajar dimasukkan. Kedua ukuran ini tidak ditafsirkan sebagai persentase variasi yang dijelaskan secara identik dengan R^2 pada regresi linear.

Uji Hosmer-Lemeshow berfungsi untuk mengetahui apakah model regresi logistik sudah sesuai dengan data (goodness of fit).

Rumusan Hipotesis pada Ilustrasi Uji Kecocokan Model (Hosmer-Lemeshow)

H_0: Model sesuai dengan data (fit).

H_a: Model tidak sesuai dengan data.

Karena p = 0,533 > 0,05, H_0 gagal ditolak. Tidak ada bukti bahwa prediksi model berbeda secara bermakna dari data yang diamati; jadi, kecocokan model dapat diterima.

Contingency Table for Hosmer and Lemeshow Test

		Tuntas = Belum Tuntas		Tuntas = Tuntas		
		Observed	Expected	Observed	Expected	Total
Step 1	1	4	4.411	1	.589	5
	2	3	4.130	2	.870	5
	3	5	3.772	0	1.228	5
	4	4	3.452	1	1.548	5
	5	3	2.836	2	2.164	5
	6	1	1.956	4	3.044	5
	7	2	1.443	3	3.557	5
	8	2	1.106	3	3.894	5
	9	0	.636	5	4.364	5
	10	0	.258	5	4.742	5

Tabel tersebut merinci hasil Hosmer-Lemeshow dengan membandingkan frekuensi yang diamati dan frekuensi yang diprediksi model pada setiap kelompok probabilitas.

Observed (nilai sebenarnya)→jumlah responden berdasarkan data asli.

Expected (nilai yang diprediksi model)→jumlah responden yang diperkirakan oleh model regresi logistik.

Data dibagi menjadi 10 kelompok (desil) berdasarkan probabilitas prediksi model.

Contoh pembacaan pada kelompok pertama adalah sebagai berikut:

Data aktual mencatat 4 siswa belum tuntas dan 1 siswa tuntas.

Model memperkirakan 4,411 siswa belum tuntas dan 0,589 siswa tuntas.

Selisih antara frekuensi aktual dan frekuensi prediksi pada kelompok tersebut relatif kecil.

Classification Table[a]

Observed			Predicted		
			Tuntas		Percentage Correct
			Belum Tuntas	Tuntas	
Step 1	Tuntas	Belum Tuntas	18	6	75.0
		Tuntas	6	20	76.9
	Overall Percentage				76.0

a. The cut value is .500

Tabel ini menunjukkan kemampuan model regresi logistik dalam mengklasifikasikan responden dari hasil prediksi.

Dari 24 siswa yang sebenarnya belum tuntas, model mengklasifikasikan 18 siswa dengan benar sebagai belum tuntas dan 6 siswa secara keliru sebagai tuntas. Perhitungan:

Artinya, model mengenali siswa yang belum tuntas dengan ketepatan 75%.

Dari 26 siswa yang sebenarnya tuntas, model mengklasifikasikan 20 siswa dengan benar sebagai tuntas dan 6 siswa secara keliru sebagai belum tuntas. perhitungan:

Artinya, model mengenali siswa yang tuntas dengan ketepatan 76,9%.

Overall percentage sebesar 76,0% berarti bahwa model regresi logistik mengklasifikasikan status ketuntasan dengan akurasi keseluruhan 76%.

Perbandingan dengan Block 0

Sebelumnya pada Block 0:

Model tanpa variabel independen: Akurasi = 52%

Setelah variabel motivasi dan jam belajar dimasukkan ke dalam model:

Pada Block 1: akurasi = 76%.

Terjadi peningkatan:

$$76\% - 52\% = 24\%$$

Artinya:

Penambahan variabel motivasi dan jam belajar meningkatkan akurasi klasifikasi model dari 52% menjadi 76%, yaitu naik 24 poin persentase.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	Motivasi	.066	.032	4.178	1	.041	1.069	1.003	1.139
	Jambelajar	.566	.176	10.333	1	.001	1.761	1.247	2.486
	Constant	-8.255	2.805	8.661	1	.003	.000		

a. Variable(s) entered on step 1: Motivasi, Jambelajar.

Pembacaan terhadap nilai koefisien B menunjukkan bahwa persamaan regresi logistik dapat dituliskan:

Logit(P) = -8,255 + 0,066(Motivasi) + 0,566(Jam Belajar)

Keterangan:

P: probabilitas siswa berada pada kategori tuntas.

Motivasi: skor motivasi belajar

Jam belajar: lama waktu belajar

Motivasi berhubungan positif dan signifikan dengan odds ketuntasan karena p = 0,041 < 0,05. Artinya, setelah jam

belajar dikendalikan, peningkatan motivasi berkaitan dengan kenaikan odds siswa untuk tuntas.

Koefisien motivasi sebesar B = 0,066 dan bernilai positif. Arah positif menunjukkan bahwa kenaikan motivasi diikuti oleh kenaikan log-odds ketuntasan.

Nilai Exp(B) atau odds ratio motivasi sebesar 1,069. Setiap kenaikan satu poin motivasi berkaitan dengan kenaikan odds ketuntasan sekitar 6,9%, dengan jam belajar dikendalikan.

Perhitungan perubahan odds: (1,069 − 1) × 100% = 6,9%.

Nilai Exp(B) atau odds ratio jam belajar sebesar 1,761. Setiap tambahan satu jam belajar berkaitan dengan kenaikan odds ketuntasan sekitar 76,1%, dengan motivasi dikendalikan.

Perhitungan perubahan odds: (1,761 − 1) × 100% = 76,1%.

Cara Membaca Keluaran

Pada dataset ilustratif, uji Omnibus menunjukkan bahwa model yang memuat motivasi dan jam belajar lebih baik daripada model yang hanya berisi konstanta, $\chi^2(2) = 18{,}818$, $p < 0{,}001$. Nagelkerke $R^2 = 0{,}418$ menunjukkan bahwa model memiliki daya jelas yang cukup, tetapi nilainya tidak dibaca sama seperti R^2 pada regresi linear. Uji Hosmer-Lemeshow menghasilkan $p = 0{,}533$, sehingga tidak ada bukti bahwa model tidak cocok. Akurasi klasifikasi meningkat dari 52% menjadi 76%. Motivasi memiliki OR = 1,069 dan jam belajar OR = 1,761. Artinya, kenaikan satu poin motivasi berkaitan dengan kenaikan odds ketuntasan sekitar 6,9%, sedangkan tambahan satu jam belajar berkaitan dengan kenaikan odds sekitar 76,1%.

Arti Keputusan Statistik

Pada uji Omnibus, $p < 0,05$ sehingga H_0 ditolak dan H_a didukung. Artinya, motivasi dan jam belajar membantu meningkatkan prediksi ketuntasan. Setiap kenaikan satu poin motivasi berkaitan dengan kenaikan odds ketuntasan sekitar 6,9%. Setiap tambahan satu jam belajar berkaitan dengan kenaikan odds sekitar 76,1%. Pada uji Hosmer-Lemeshow, H_0 gagal ditolak karena tidak ada bukti bahwa model tidak cocok dengan data.

LATIHAN

1. Tentukan X dan Y pada penelitian hubungan waktu belajar dan skor ujian!
2. Uraikan makna slope b = 2,5 dan intersep a = 40!
3. Hitung prediksi jika $\hat{Y} = 40 + 2{,}5X$ dan X = 6!
4. Paparkan perbedaan residual, interval kepercayaan mean, dan interval prediksi individu!
5. Mengapa R^2 tinggi tidak membuktikan model kausal Kapan regresi linear berganda lebih tepat daripada regresi linear sederhana?
6. Rumuskan persamaan regresi berganda dari koefisien a = 20, $b_1 = 0{,}30$, dan $b_2 = 2{,}00$.
7. Apa arti VIF = 1,20 dan Adjusted R^2 = 0,65 pada regresi berganda?

Dataset Latihan Penerapan SPSS Bab 14.

Regresi penggunaan GeoGebra

ID	Jam penggunaan	Skor visualisasi
1	1	52
2	2	57
3	3	60
4	4	63
5	5	68
6	6	66
7	7	72
8	8	75
9	9	73
10	10	80

ID	Jam penggunaan	Skor visualisasi
11	11	82
12	12	85
13	3	59
14	5	65
15	7	70
16	9	78
17	11	81
18	13	88
19	6	69
20	8	76

Rumusan Hipotesis pada Ilustrasi

H_0: $\beta_1 = 0$ (durasi penggunaan GeoGebra tidak memprediksi skor visualisasi secara linear).

H_a: $\beta_1 \neq 0$ (durasi penggunaan GeoGebra memprediksi skor visualisasi secara linear).

1. Buat scatterplot dan periksa linearitas.
2. Jalankan regresi linear sederhana dengan skor visualisasi sebagai variabel dependen.
3. Periksa R, R Square, ANOVA, coefficients, confidence interval, dan residual plot.
4. Susun persamaan regresi dan prediksi skor untuk penggunaan 8 jam.
5. Paparkan keterbatasan prediksi dan larangan melakukan ekstrapolasi di luar rentang data.

LAMPIRAN A
MATRIKS PEMILIHAN UJI STATISTIK

Tabel A.1. Matriks pemilihan uji statistik

Tujuan	Kelompok / variabel	Hubungan sampel	Skala / kondisi	Uji utama
Menguji mean terhadap nilai acuan	1 sampel	-	Interval/rasio, normal	One-sample t
Menguji proporsi dua kategori	1 sampel	-	Nominal dikotomi	Binomial
Menguji distribusi kategori	1 sampel	-	Nominal	Chi-square goodness-of-fit
Menguji kerandoman urutan	1 deret	-	Dikotomi/ordinal	Run test
Membandingkan 2 kondisi	2	Berpasangan	Interval/rasio, normalitas perbedaan	Paired t
Membandingkan 2 kondisi	2	Berpasangan	Ordinal/tidak normal	Wilcoxon
Membandingkan 2 kelompok	2	Independen	Interval/rasio, normal	Independent t / Welch
Membandingkan 2 kelompok	2	Independen	Ordinal/tidak normal	Mann-Whitney

Tujuan	Kelompok / variabel	Hubungan sampel	Skala / kondisi	Uji utama
Membandingkan proporsi	2 kategori	Independen	Nominal, expected memadai	Chi-square independensi
Membandingkan proporsi	2 kategori	Independen	Nominal, expected kecil	Fisher exact
Membandingkan ≥3 mean	≥3	Independen	Interval/rasio, asumsi memadai	One-way ANOVA / Welch
Membandingkan ≥3 rank	≥3	Independen	Ordinal/tidak normal	Kruskal-Wallis
Membandingkan ≥3 kondisi	≥3	Berpasangan	Interval/rasio, asumsi memadai	Repeated-measures ANOVA
Membandingkan ≥3 rank	≥3	Berpasangan	Ordinal/tidak normal	Friedman
Membandingkan ≥3 proporsi	≥3	Berpasangan	Nominal dikotomi	Cochran Q
Menguji hubungan	2 variabel	Berpasangan per subjek	Interval/rasio, linear	Pearson
Menguji hubungan	2 variabel	Berpasangan per subjek	Ordinal/monoton	Spearman
Memprediksi variabel hasil	1 X dan 1 Y	Berpasangan per subjek	Linear dan residual memadai	Regresi linear sederhana

Tujuan	Kelompok / variabel	Hubungan sampel	Skala / kondisi	Uji utama
Memprediksi variabel hasil numerik	Beberapa prediktor	Observasi independen	variabel hasil interval/rasio; residual memenuhi asumsi	Regresi linear berganda
Memprediksi variabel hasil dua kategori	Satu atau beberapa prediktor	Observasi independen	variabel hasil dikotomi 0/1	Regresi logistik biner

LAMPIRAN B
PANDUAN MEMBACA KELUARAN SPSS

Tabel B.1. Komponen penting keluaran SPSS

Komponen keluaran	Pertanyaan yang dijawab
N / Valid N	Berapa observasi yang dikaji? Apakah ada data hilang?
Mean / Median	Di mana pusat data?
SD / IQR	Seberapa besar variasi data?
Levene Sig.	Apakah asumsi homogenitas varians memadai?
Test statistic	Seberapa besar deviasi hasil dari H_0 dalam satuan uji?
Df	Berapa derajat kebebasan yang digunakan?
Sig. / p-value	Seberapa kompatibel data dengan H_0?
Confidence interval	Rentang nilai parameter yang konsisten dengan data?
Effect size	Seberapa besar perbedaan atau hubungan?
Expected count	Apakah asumsi chi-square terpenuhi?
Residual plot	Apakah pola linear, varians konstan, dan pencilan bermasalah?

Komponen keluaran	Pertanyaan yang dijawab
Adjusted R Square	Berapa proporsi variasi variabel hasil yang dijelaskan setelah mempertimbangkan jumlah prediktor?
Tolerance / VIF	Apakah terdapat multikolinearitas berat pada regresi berganda?
Exp(B) / Odds Ratio	Berapa kali perubahan odds kejadian untuk kenaikan satu unit prediktor pada regresi logistik?
Hosmer-Lemeshow	Apakah data menunjukkan ketidaksesuaian model logistik yang signifikan?
Classification Table	Seberapa baik model logistik mengklasifikasikan kategori variabel hasil pada cut-off tertentu?

Komponen Penyajian Statistik

Cantumkan jenis uji, statistik deskriptif, statistik uji, df bila ada, p-value, ukuran efek, interval kepercayaan bila tersedia, arah hasil, dan kesimpulan yang sesuai konteks.

LAMPIRAN C
DATASET LATIHAN

C.1 Dataset Nilai 40 Peserta

67, 65, 60, 66, 77, 64, 54, 37, 86, 62, 70, 83, 71, 46, 57, 52, 45, 47, 52, 55, 72, 79, 73, 66, 40, 70, 73, 95, 68, 84, 61, 60, 99, 92, 71, 90, 66, 86, 56, 75.

C.2 Dataset Berpasangan

Tabel C.1. Dataset pretest-posttest

Subjek	Pretest	Posttest
1	62	70
2	68	74
3	75	78
4	55	64
5	80	82
6	71	73
7	66	69
8	59	65
9	77	80
10	64	72

C.3 Dataset Dua Kelompok

Tabel C.2. Dataset dua kelompok independen

Subjek	Kelas A	Kelas B
1	65	68
2	72	75
3	70	74
4	61	69
5	77	80
6	68	72
7	74	76
8	66	71
9	69	73
10	71	78

C.4 Dataset Korelasi dan Regresi

Tabel C.3. Dataset waktu belajar dan skor ujian

Siswa	Waktu belajar (jam/minggu)	Skor ujian
1	2	58
2	3	62
3	4	66

Siswa	Waktu belajar (jam/minggu)	Skor ujian
4	5	69
5	5	72
6	6	74
7	7	78
8	8	83
9	9	85
10	10	89

LAMPIRAN D
KUNCI JAWABAN TERPILIH

Bab 3 No. 1: Mean = 5,33; median = 5; modus = 4; rentang = 5; varians sampel sekitar 2,81; simpangan baku sekitar 1,68.

Bab 3 No. 3: Mean harmonik dua kecepatan untuk jarak sama = 2/(1/40 + 1/60) = 48 km/jam.

Bab 4 No. 5: Normalitas diperiksa pada skor perbedaan posttest-pretest.

Bab 5 No. 1: H_0: $\mu = 75$; H_a: $\mu \neq 75$.

Bab 7 No. 1: Karena skor berasal dari orang yang sama dan berkorelasi; independent t mengasumsikan kelompok terpisah.

Bab 9 No. 2: Expected count diperoleh dari total baris × total kolom / 75.

Bab 13 No. 3: $r^2 = 0{,}4225$ atau 42,25%.

Bab 14 No. 3: $\hat{Y} = 40 + 2{,}5(6) = 55$.

LAMPIRAN E
CARA MENYUSUN CATATAN PENGOLAHAN DATA

Gunakan susunan berikut untuk mencatat proses pengolahan dan pengkajian data. Setiap keputusan perlu dapat ditelusuri dari data, prosedur, dan keluaran SPSS.

Tabel E.1. Struktur catatan pengolahan data

Bagian	Isi minimum
Identitas	Nama, NIM, kelas, bab pembelajaran, tanggal, dan nama file data.
Tujuan	Pertanyaan penelitian dan tujuan pengolahan data.
Variabel dan desain	Definisi variabel, skala pengukuran, kelompok, dan hubungan sampel.
Persiapan data	Codebook, penanganan missing values, pemeriksaan kode, dan pencilan.
Prosedur SPSS	Menu atau sintaks yang digunakan.
Asumsi	Bukti numerik dan grafik yang relevan.
Hasil	Statistik deskriptif, statistik uji, df, p-value, confidence interval, dan ukuran efek.
Panduan membaca keluaran	Jawaban terhadap pertanyaan penelitian dalam konteks pendidikan.
Makna keputusan statistik	Ringkasan satu paragraf, termasuk batasan pengolahan data.
Lampiran	Keluaran terpilih dan sintaks, bukan seluruh tampilan yang tidak relevan.

Contoh penyusunan kalimat statistik

Deskriptif: Rata-rata skor sebesar M =..., SD =..., dengan rentang... sampai....

Komparatif: Kelompok A memiliki skor... daripada Kelompok B, t(df) =..., p =..., d =....

ANOVA: ada perbedaan di antara kelompok, F(df1, df2) =..., p =..., η^2 =....

Korelasi: ada hubungan positif/negatif... antara X dan Y, r(df) =..., p =....

Regresi: X memprediksi Y secara signifikan, F(df1, df2) =..., p =..., R^2 =..., b =....

LAMPIRAN F
RUBRIK EVALUASI PENERAPAN PENGOLAHAN DAN PENGKAJIAN DATA

Tabel F.1. Rubrik penilaian

Komponen	Bobot	Kriteria utama
Persiapan dan codebook	15%	Nama variabel, label, kode, missing values, dan measure tepat.
Ketepatan prosedur	20%	Uji sesuai tujuan, desain, skala, dan hubungan sampel.
Pemeriksaan asumsi	15%	Asumsi diperiksa dengan indikator yang tepat dan tidak mekanis.
Ketepatan membaca keluaran	20%	Tabel, baris, statistik, df, dan p-value dipilih dengan benar.
Panduan membaca keluaran dan ukuran efek	20%	Makna keputusan statistik menjawab masalah dan membedakan makna statistik serta praktis.
Kerapian dan replikasi	10%	File tersusun, sintaks tersedia, tabel diberi judul, dan catatan pengolahan data mudah ditelusuri.

GLOSARIUM

Tabel G.1. Glosarium istilah utama

Istilah	Pengertian
Alpha (α)	Taraf risiko kesalahan tipe I yang ditetapkan sebelum pengolahan data.
Asumsi	Kondisi yang diperlukan agar model dan inferensi statistik dapat digunakan secara tepat.
Data	Informasi hasil pengamatan, pengukuran, pencatatan, atau kategorisasi.
Derajat kebebasan	Jumlah informasi bebas yang digunakan dalam estimasi statistik.
Efek	Besarnya perbedaan, hubungan, atau perubahan yang menjadi fokus pembahasan.
Frekuensi harapan	Jumlah kasus yang diperkirakan pada sel tabel bila H_0 benar.
Hipotesis alternatif	Pernyataan yang dipertimbangkan ketika H_0 ditolak.
Hipotesis nol	Pernyataan acuan mengenai tidak adanya perbedaan, hubungan, atau kesesuaian tertentu.
Homoskedastisitas	Varians residual yang relatif konstan sepanjang nilai prediksi.
Independensi	Keadaan ketika satu observasi tidak menentukan atau bergantung pada observasi lain.
Interval kepercayaan	Rentang estimasi parameter yang mencerminkan ketidakpastian sampling.

Istilah	Pengertian
Mean	Jumlah nilai dibagi banyaknya observasi.
Median	Nilai tengah data yang telah diurutkan.
Modus	Nilai atau kategori yang paling sering muncul.
Normalitas	Kedekatan distribusi data atau residual dengan distribusi normal.
P-value	Probabilitas hasil setidaknya se-ekstrem observasi dengan asumsi H_0 benar.
Parameter	Ukuran numerik yang menggambarkan populasi.
Pencilan	Observasi yang jauh dari pola sebagian besar data.
Populasi	Keseluruhan unit yang menjadi sasaran penelitian.
Power	Peluang menolak H_0 ketika H_0 memang salah.
Prediktor	Variabel penjelas atau X dalam model regresi.
Rank	Urutan relatif observasi setelah data disusun.
Residual	Selisih nilai observasi dengan nilai prediksi.
Sampel	Bagian populasi yang diamati.
Simpangan baku	Akar varians dan ukuran penyebaran dalam satuan data.
Statistik	Ringkasan numerik data sampel atau kumpulan fakta angka.

Istilah	Pengertian
Statistika	Ilmu tentang pengumpulan, pengolahan, penafsiran, dan penarikan kesimpulan dari data.
Ukuran efek	Ukuran kuantitatif besarnya perbedaan atau hubungan.
Varians	Rata-rata kuadrat penyimpangan dari mean dengan penyebut sesuai populasi atau sampel.
Variabel	Karakteristik yang nilainya dapat berbeda antar-unit.
Adjusted R^2	R^2 yang disesuaikan terhadap jumlah prediktor dan ukuran sampel.
Logit	Logaritma natural dari odds, yaitu log[p/(1-p)].
Odds ratio	Perbandingan odds; pada regresi logistik ditunjukkan oleh Exp(B).
VIF	Variance Inflation Factor, indikator tingkat multikolinearitas suatu prediktor.
Hosmer-Lemeshow test	Uji kecocokan regresi logistik berdasarkan perbandingan kejadian observasi dan prediksi dalam kelompok risiko.

DAFTAR PUSTAKA

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika, 37*(3–4), 256–266. https://doi.org/10.1093/biomet/37.3-4.256

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Field, A. (2024). *Discovering statistics using IBM SPSS Statistics* (6th ed.). SAGE Publications.

Fisher, R. A. (1922). On the interpretation of χ^2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society, 85*(1), 87–94. https://doi.org/10.1111/j.2397-2335.1922.tb00768.x

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association, 32*(200), 675–701. https://doi.org/10.1080/01621459.1937.10503522

Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the behavioral sciences* (10th ed.). Cengage Learning.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (3rd ed.). Wiley.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics, 18*(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Nurgiyantoro, B., Gunawan, & Marzuki. (2000). *Statistik terapan untuk penelitian ilmu-ilmu sosial*. Gadjah Mada University Press.

Pallant, J. (2026). SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS (8th ed.). Routledge.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology, 15*(1), 72–101. https://doi.org/10.2307/1412159

Subana, M., Rahadi, M., & Sudrajat. (2000). *Statistik pendidikan*. Pustaka Setia.

Sudijono, A. (2001). *Pengantar statistik pendidikan*. RajaGrafindo Persada.

Sudjana. (2005). *Metoda statistika*. Tarsito.

Sugiyono. (2012). *Statistika untuk penelitian*. Alfabeta.

Sundayana, R. (2010). *Statistika penelitian pendidikan*. STKIP Garut Press.

Triola, M. F. (2022). *Elementary statistics* (14th ed.). Pearson.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician, 70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Welch, B. L. (1947). The generalization of "Student's" problem when several different population variances are involved. *Biometrika, 34*(1–2), 28–35. https://doi.org/10.1093/biomet/34.1-2.28

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83. https://doi.org/10.2307/3001968

American Psychological Association. (2020). Publication manual of the American Psychological Association (7th ed.). American Psychological Association.

IBM Corp. (2026). IBM SPSS Statistics for Windows, Version 32.0. IBM Corp.

IBM. (2026). IBM SPSS Statistics 32 documentation. IBM Documentation.

BIODATA PENULIS



Dr. Afifah Latip Rasyid Jauhari, M.Pd.
Program Studi Pendidikan Matematika
Universitas Mandiri, Subang.

Riwayat pendidikan: Magister Pendidikan Matematika, Universitas Pasundan (2016); Doktor Pendidikan Matematika, Universitas Pendidikan Indonesia (2023).

Bidang keahlian: pendidikan statistika, penalaran statistis, self-regulated learning, pengembangan e-modul penggunaan SPSS dalam pembelajaran statistika.

Penulis aktif dalam pendidikan, penelitian, dan publikasi ilmiah di bidang pendidikan matematika dan statistika. Dalam buku ini, penulis menyusun contoh penerapan, panduan SPSS, dan latihan yang membantu mahasiswa menentukan keputusan berdasarkan data.

BIODATA PENULIS



Anwar Sadat, M.Pd.
Program Studi Pendidikan Matematika, Universitas Mandiri, Subang.

Anwar Sadat adalah Ketua Program Studi Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Mandiri, Subang. Materi dalam buku ini dikembangkan dari bahan ajar Statistika Penelitian Pendidikan untuk membantu mahasiswa memahami pengolahan dan pengkajian data pendidikan.

Bidang yang dikembangkan meliputi statistika deskriptif, sampling, uji asumsi, pengujian hipotesis parametrik dan nonparametrik, korelasi, regresi, serta penggunaan SPSS. Kontribusi penulis dalam buku ini terutama pada penyusunan alur belajar, latihan berbasis data, dan contoh pengolahan data dalam pendidikan matematika.