



Penerbit  
Gita Lentera

# PENGANTAR SPSS

Teori, Implementasi dan Interpretasi



# PENGANTAR SPSS: TEORI, IMPLEMENTASI DAN INTERPRETASI

## **Penulis:**

Ely Mulyati, S.T., M.T.

M. Rosihan Arsyad

Suryaningsih, S.E., M.M.

Sri Maryati, S.ST., M. Kes.

Leni Gustina, S.E., M.M.

Pilifus Junianto S.S, M.M.

Kiki Helencia, S.E., M.M.

Linda Prasetyaning Widayanti, S.KM., M.Kes.

Dr. Hwihanus, S.E., M.M., CMA.

Dr.Ir.La Ode Muhamad Nurrahmad Arsyad, S.T., M.T.

Sri Hidayati L, S.KM., M.Kes.

## **Editor:**

Nanda Harry Mardika, S.Pd., M.M.



Penerbit CV. Gita Lentera

[WWW.GITALENTERA.COM](http://WWW.GITALENTERA.COM)

# Pengantar Spss: Teori, Implementasi dan Interpretasi

## **Penulis:**

Ely Mulyati, S.T., M.T.

M. Rosihan Arsyad

Suryaningsih, S.E., M.M.

Sri Maryati, S.ST., M. Kes.

Leni Gustina, S.E., M.M.

Pilifus Junianto S.S, M.M.

Kiki Helencia, S.E., M.M.

Linda Prasetyaning Widayanti, S.KM., M.Kes.

Dr. Hwihanus, S.E., M.M., CMA.

Dr.Ir.La Ode Muhamad Nurrahmad Arsyad, S.T., M.T.

Sri Hidayati L, S.KM., M.Kes.

## **Editor:**

Nanda Harry Mardika, S.Pd., M.M.

Hak Cipta dilindungi oleh Undang-undang  
©All right reserved

ISBN: 978-634-7072-09-2

Layouter : Adnan, M.H.  
Desain Sampul : Sri Nursanti Sari, M.Pd.  
Penerbit : CV. Gita Lentera  
Perm. Permata Hijau Regency blok F/1 Kel. Pisang,  
Kec. Pauh, Padang  
Website: <https://gitalentera.com>  
Email: [gitalentera.publisher@gmail.com](mailto:gitalentera.publisher@gmail.com)  
Anggota IKAPI No. 042/SBA/2023

Cetakan Pertama, 3 Desember 2024

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh buku ini tanpa izin tertulis dari penerbit.

# KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa, karena berkat rahmat dan karunia-Nya, penulis dapat menyelesaikan penyusunan buku "Pengantar SPSS: Teori, Implementasi, dan Interpretasi" ini. Buku ini disusun sebagai upaya untuk memberikan pemahaman dasar dan praktis mengenai penggunaan SPSS (Statistical Package for the Social Sciences) yang merupakan salah satu perangkat lunak statistik yang banyak digunakan dalam analisis data, khususnya di bidang ilmu sosial, kesehatan, dan bisnis. Penulis berharap buku ini dapat bermanfaat bagi para mahasiswa, peneliti, serta praktisi yang ingin memahami dan menguasai SPSS untuk tujuan analisis data yang lebih efektif.

Penyusunan buku ini bertujuan untuk memberikan pengenalan menyeluruh mengenai SPSS, dari segi teori dasar statistik hingga langkah-langkah implementasi praktis yang dapat diterapkan dalam berbagai penelitian. Setiap bab disusun dengan pendekatan yang sistematis, mulai dari konsep dasar tentang pengolahan data, pengenalan antarmuka SPSS, hingga cara-cara dalam menginterpretasikan hasil analisis yang diperoleh. Penulis juga berusaha menghadirkan penjelasan yang sederhana, disertai dengan contoh-contoh praktis yang dapat membantu pembaca untuk lebih memahami cara kerja SPSS secara langsung.

Dalam buku ini, teori-teori yang mendasari penggunaan SPSS dijelaskan dengan tujuan agar pembaca tidak hanya mampu menggunakan perangkat lunak ini, tetapi juga memahami konsep-konsep statistik yang terlibat dalam setiap langkah analisis. Dengan pemahaman teori yang cukup, diharapkan pembaca dapat melakukan

analisis data dengan lebih tepat dan dapat menginterpretasikan hasilnya secara akurat. Buku ini juga membahas tentang bagaimana cara mengolah data penelitian secara efisien, serta memberikan solusi bagi masalah yang sering dihadapi dalam penggunaan SPSS.

Penulis menyadari bahwa dalam penyusunan buku ini, tidak lepas dari adanya kekurangan dan keterbatasan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan dan penyempurnaan buku ini di masa mendatang. Semoga buku ini dapat memberikan kontribusi yang positif bagi dunia pendidikan dan penelitian, serta mempermudah para pembaca dalam memanfaatkan SPSS untuk analisis data yang lebih baik dan terstruktur.

# SINOPSIS

*“Statistical analysis is one of the key elements of any research, and SPSS has been the go-to software for both students and professionals alike due to its versatility and user-friendly interface” (Pallant, 2020).*

Buku ini merupakan panduan komprehensif untuk memahami konsep-konsep dasar dalam analisis data menggunakan SPSS, serta memberikan langkah-langkah praktis dalam mengimplementasikan teknik-teknik statistik yang penting dalam penelitian. Dimulai dengan pengenalan mengenai SPSS sebagai perangkat lunak statistik yang populer, buku ini menjelaskan berbagai jenis variabel dan data yang dapat dikelola di dalam SPSS. Pembaca akan diajarkan bagaimana cara mentransformasikan data, baik itu untuk keperluan pengolahan data awal maupun untuk meningkatkan kualitas analisis. Selanjutnya, buku ini mengupas berbagai teknik analisis deskriptif, yang memungkinkan pengguna untuk menggambarkan dan menyimpulkan data dengan mudah.

Buku ini juga membahas berbagai teknik statistik lanjutan yang dapat digunakan untuk menguji validitas dan reliabilitas data, uji asumsi klasik, serta uji hipotesis parametrik dan non-parametrik. Untuk penelitian yang lebih mendalam, pembaca akan diberikan pengetahuan tentang regresi linier, regresi logistik, serta analisis multivariat seperti analisis faktor dan kluster. Setiap bab dilengkapi dengan ilustrasi praktis dan contoh kasus nyata, memungkinkan pembaca untuk mengimplementasikan setiap teknik dalam penelitian mereka. Dengan pendekatan yang sistematis dan mudah diikuti, buku ini sangat cocok bagi mahasiswa, peneliti, dan praktisi yang ingin

meningkatkan pemahaman dan kemampuan mereka dalam menggunakan SPSS untuk analisis data yang lebih tepat dan akurat.

# DAFTAR ISI

HALAMAN JUDUL.....	i
KATA PENGANTAR.....	iii
SINOPSIS.....	v
DAFTAR ISI.....	vii
PENGELANAN SPSS.....	1
1.1. Pendahuluan.....	1
1.2. Penggunaan SPSS di Berbagai Bidang.....	2
1.3. Hubungan Antara SPSS dengan Era Digital dan Data.....	9
1.4. Pentingnya Statistik dan SPSS.....	11
VARIABEL DAN JENIS DATA DALAM SPSS.....	18
2.1 Variabel dan Jenis Data dalam SPSS.....	18
2.2 Variabel dalam Konteks Penelitian dan SPSS.....	20
2.3 Jenis-Jenis Variabel dalam SPSS.....	23
2.4 Pengelompokan Variabel dalam Skala Pengukuran Data.....	26
2.5 Implementasi dalam SPSS: Memasukkan dan Mengatur Variabel.....	29
2.6 Pengaturan Label dan Nilai dalam SPSS.....	32
2.7 Skala Pengukuran Analisis Data dalam SPSS.....	35
2.8 Studi Kasus: Menggunakan Variabel dan Jenis Data dalam SPSS	39
2.9 Kesalahan Umum dalam Mengelola Variabel dan Data di SPSS ...	42
2.10 Kunci Sukses Pengelolaan Variabel dan Data dalam SPSS.....	46
TRANSFORMASI DATA DALAM SPSS.....	49
3.1 Pengertian Transformasi Data.....	49
3.1 Fungsi Transformasi Data.....	50
3.2 Jenis-Jenis Transformasi Data Dalam SPSS.....	51
3.3 Beberapa Langkah Umum yang digunakan dalam SPSS Untuk Transformasi Data.....	54
ANALISIS DESKRIPTIF.....	61
4.1 Pendahuluan.....	61
4.2 Teknik Analisis Deskriptif Kuantitatif.....	62
4.3 Analisis deskriptif dengan SPSS.....	66

4.4 Penyajian Data Statistik Deskriptif .....	75
UJI VALIDITAS DAN RELIABILITAS .....	80
5.1. Uji Validitas .....	80
5.2. Langkah-langkah Pengujian Uji Validitas .....	81
5.3. Interpretasi hasil Ouput Uji Validitas .....	82
5.4. Uji Reliabilitas .....	82
UJI ASUMSI KLASIK .....	86
6.1 Uji Normalitas .....	86
6.2 Multikolinearitas .....	87
6.3 Heterokedasdisitas .....	89
6.4 Autokorelasi .....	91
6.5 Linearitas .....	93
UJI HIPOTESIS PARAMETRIK .....	96
7.1 Uji t (t-test) .....	96
7.2 Uji Z (Z-test) .....	98
7.3 Uji ANOVA (Analysis of Variance) .....	100
7.4 Uji Chi-Square ( $\chi^2$ -test) .....	103
7.5 Uji regresi linier .....	105
UJI HIPOTESIS NON - PARAMETRIK .....	110
8.1 Uji Mann-Whitney .....	110
8.2 Uji Wilcoxon Signed-Rank Test .....	112
8.3 Uji Kruskal-Wallis .....	114
8.4 Uji Friedman .....	117
8.5 Uji Chi Square .....	119
8.6 Uji McNemar .....	121
8.7 Uji Cochran's Q .....	123
8.8 Uji Korelasi Spearman .....	125
8.9 Uji Koefisien Phi .....	127
REGRESI LINEAR .....	131
9.1. Pengertian dan Persamaan Regresi Linear .....	131
9.2. Penerapan Bidang Regresi Linear .....	132
9.3. Regresi Linear Sederhana .....	133
9.4. Interpretasi Koefisien .....	133
9.5. Langkah-langkah Melakukan Regresi Linear Sederhana .....	134
9.6. Metode Ordinary Least Squares (OLS) .....	135
9.7. Regresi Linear Berganda .....	138

9.8. Tujuan dan Fungsi Regresi Linear Berganda.....	139
9.9. Asumsi Regresi Linear Berganda .....	140
9.10. Koefisien Regresi dan Interpretasi.....	141
9.11. Analisis Signifikansi.....	143
9.12. Koefisien Determinasi ( $R^2$ ).....	145
9.13. Aplikasi Regresi Linear dalam SPSS.....	147
ANALISIS FAKTOR .....	152
10.1. Latar Belakang Penggunaan SPSS dalam Analisis Faktor .....	152
10.2. Dasar-Dasar Teori Analisis Faktor.....	154
10.3. Penggunaan SPSS dalam Melakukan Analisis Faktor .....	157
10.4. Implikasi Analisis Faktor Bagi Akademisi dan Praktisi.....	164
10.5. Arah Pengembangan Metode Analisis Faktor di Masa Depan..	166
ANALISIS KLUSTER .....	175
11.1 Pengantar.....	175
11.2 Metode Cluster Analisis.....	176
11.3 Langkah-Langkah SPSS Metode Cluster Analysis .....	179
BIODATA PENULIS .....	198

# **PENGELANAN SPSS**

*Oleh Ely Mulyati, S.T., M.T.*

## **1.1. Pendahuluan**

Di era modern ini, data telah menjadi bagian yang tidak terpisahkan dari kehidupan manusia. Dalam dunia akademik, penelitian berbasis data menjadi landasan bagi pengembangan ilmu pengetahuan. Kehadiran teknologi telah membawa dunia ke era revolusi data. Data dihasilkan setiap detik dari berbagai aktivitas manusia, seperti transaksi online, penggunaan media sosial, hingga sensor IoT (Internet of Things). Dalam era Big Data, analisis statistik menjadi lebih penting dari sebelumnya. Statistik tidak hanya digunakan untuk menganalisis data tradisional, tetapi juga untuk mengeksplorasi data dalam skala besar. SPSS, sebagai salah satu perangkat lunak statistik yang paling populer, memainkan peran penting dalam membantu pengguna menangani data dengan lebih mudah, bahkan bagi mereka yang tidak memiliki latar belakang statistik yang mendalam.

Di sektor bisnis, data menjadi dasar untuk memahami pasar dan mengambil keputusan strategis. Bahkan, di kehidupan sehari-hari, data membantu pemerintah dan organisasi memahami kebutuhan masyarakat. Namun, data yang melimpah tidak akan bernilai tanpa diolah dengan cara yang tepat, volume data yang besar ini hanya bermanfaat jika diolah menjadi informasi yang relevan dan dapat

ditindaklanjuti. Statistik sebagai disiplin ilmu memberikan metode untuk menganalisis dan menyimpulkan data secara objektif. Dalam hal ini, perangkat lunak statistik, seperti SPSS (*Statistical Package for the Social Sciences*), menjadi alat yang sangat penting. Dengan kemampuannya yang intuitif dan efisien, SPSS mendukung pengguna dari berbagai latar belakang untuk menganalisis data secara mendalam tanpa harus menjadi ahli statistik. SPSS mempermudah proses pengolahan data dengan menyediakan antarmuka yang ramah pengguna, berbagai metode analisis statistik, dan kemampuan untuk menangani data dalam skala besar. Dengan memanfaatkan SPSS, pengambilan keputusan dapat dilakukan dengan lebih cepat, akurat, dan berbasis bukti.

Salah satu kendala terbesar yang dihadapi oleh banyak peneliti dan praktisi adalah kesenjangan antara teori statistik dan implementasinya dalam perangkat lunak. Banyak yang memahami teori statistik secara konseptual, tetapi kesulitan menghubungkannya dengan proses analisis menggunakan alat seperti SPSS. Sebaliknya, ada juga yang dapat mengoperasikan SPSS tetapi tidak memahami dasar teori yang mendasari analisis yang dilakukan. Ketidakseimbangan ini sering kali menghasilkan interpretasi yang keliru atau analisis yang tidak valid. Hal ini menunjukkan perlunya panduan yang tidak hanya mengajarkan cara menggunakan SPSS, tetapi juga menjelaskan teori statistik di balik setiap langkah analisis. Di sinilah perangkat lunak statistik seperti SPSS hadir sebagai solusi yang memungkinkan analisis data dilakukan secara efisien dan akurat tanpa harus memahami setiap detail perhitungan manual.

## **1.2. Penggunaan SPSS di Berbagai Bidang**

Statistik adalah ilmu yang memberikan cara sistematis untuk mengumpulkan, menganalisis, dan menyajikan data. Data mentah seringkali sulit dipahami dan digunakan. Statistik membantu

menyaring data tersebut, mengidentifikasi pola, dan menghasilkan informasi yang dapat digunakan untuk membuat keputusan yang lebih baik (Seddaoui et al., 2025; Yildizhan et al., 2024). Di dunia yang semakin kompleks dan penuh dengan ketidakpastian, keputusan yang tepat tidak lagi hanya bergantung pada intuisi atau pengalaman. Statistik adalah cabang ilmu yang berperan penting dalam memberikan landasan untuk memahami pola, tren, dan hubungan dalam data. Dalam dunia akademik, penelitian yang dilakukan oleh mahasiswa dan dosen sangat bergantung pada kemampuan untuk menganalisis data secara objektif. Perangkat lunak SPSS (*Statistical Package for the Social Sciences*) pada awalnya dirancang untuk mendukung analisis data di bidang ilmu sosial. Namun, seiring waktu, kemampuannya yang luas dan fleksibilitasnya dalam mengolah data telah membuat SPSS diadopsi di berbagai bidang lainnya seperti ilmu sosial, kesehatan, pendidikan, ekonomi, sains dan teknologi. Statistik menjadi alat penting dalam pengambilan keputusan berbasis data, baik di dunia akademik maupun profesional. Di dunia profesional, statistik digunakan statistik memprediksi tren pasar dan pola konsumen, mengukur kinerja organisasi, hingga memprediksi tren di masa depan. Statistik tidak hanya memberikan wawasan, tetapi juga membantu membuktikan hipotesis secara ilmiah. Hal ini menunjukkan bahwa SPSS adalah alat statistik yang relevan untuk berbagai jenis penelitian dan aplikasi profesional.

#### **a. Penggunaan SPSS di Berbagai Bidang ilmu Sosial**

SPSS adalah alat yang sangat bermanfaat bagi ilmu sosial, memungkinkan peneliti untuk menjawab pertanyaan-pertanyaan penting tentang masyarakat, perilaku manusia, dan fenomena sosial lainnya (Su & Ju, 2025). Dengan kemampuannya untuk memproses data secara efisien dan memberikan hasil yang dapat diandalkan, SPSS membantu meningkatkan kualitas penelitian dan memberikan

kontribusi nyata bagi pengembangan ilmu sosial (Tavani et al., 2024). Adapun penggunaan SPSS pada bidang ilmu Sosial antara lain :

1) Analisis Survei

SPSS sering digunakan untuk menganalisis data hasil survei, seperti data dari kuesioner atau wawancara. Peneliti dapat melakukan uji statistik untuk menentukan hubungan antara variabel, seperti tingkat pendidikan dan partisipasi politik.

2) Evaluasi Program

Dalam studi kebijakan publik atau evaluasi program sosial, SPSS membantu mengevaluasi efektivitas program berdasarkan data kuantitatif. Misalnya, menilai dampak pelatihan keterampilan terhadap peningkatan pendapatan masyarakat.

3) Studi Perilaku

Psikolog dan sosiolog menggunakan SPSS untuk memahami pola perilaku individu atau kelompok, seperti faktor-faktor yang memengaruhi kepuasan kerja atau hubungan sosial di masyarakat.

4) Penelitian Akademik

Di bidang akademik, mahasiswa dan dosen memanfaatkan SPSS untuk mengolah data penelitian, seperti analisis regresi, analisis faktor, atau uji perbedaan rata-rata antar kelompok.

### **b. Penggunaan SPSS di Berbagai Bidang Kesehatan**

SPSS adalah perangkat lunak statistik yang sering digunakan dalam penelitian kesehatan untuk menganalisis data dan menarik kesimpulan berbasis bukti (Tavani et al., 2024). Dengan fitur-fitur yang fleksibel dan mudah digunakan, SPSS membantu profesional kesehatan, akademisi, dan peneliti dalam memahami pola data, mengidentifikasi faktor risiko, serta mengevaluasi efektivitas

intervensi kesehatan (Hoo et al., 2024; Jozpanahi et al., 2025; Salah et al., 2024).

1) Analisis Epidemiologi

SPSS digunakan untuk mempelajari distribusi penyakit, seperti menghitung prevalensi, insidensi, dan faktor risiko suatu penyakit. Misalnya, analisis data untuk mengetahui hubungan antara kebiasaan merokok dengan kejadian kanker paru-paru.

2) Penelitian Klinis

Dalam uji klinis, SPSS digunakan untuk mengevaluasi efektivitas obat atau terapi. Peneliti dapat membandingkan hasil antara kelompok perlakuan dan kelompok kontrol dengan uji statistik seperti *t-test* atau ANOVA.

3) Pengelolaan Data Rumah Sakit

SPSS membantu rumah sakit dalam mengolah data rekam medis untuk memahami tren penyakit, mengelola sumber daya, dan meningkatkan kualitas pelayanan kesehatan berdasarkan data (Duranteau et al., 2024).

4) Studi Kesehatan Masyarakat

Peneliti menggunakan SPSS untuk menganalisis data survei kesehatan masyarakat, seperti survei gizi, pola aktivitas fisik, atau tingkat akses ke fasilitas kesehatan

### c. Penggunaan SPSS di Berbagai Bidang Sains

Pada bidang sains manfaat penggunaan SPSS dapat memberikan manfaat pada Pengolahan Data Besar, SPSS memungkinkan analisis dataset besar dengan berbagai format, seperti data eksperimen, data sensor, atau data survei, sehingga mempermudah proses analisis dalam penelitian skala besar; Validasi Hipotesis, Dengan fitur statistik inferensial, SPSS membantu peneliti memvalidasi hipotesis, misalnya dengan uji *t-test*, ANOVA, atau regresi, sehingga mendukung pembuktian teori dalam sains; Peningkatan Akurasi Hasil, Algoritma

SPSS yang terstandarisasi menjamin hasil analisis yang akurat, membantu peneliti menghasilkan kesimpulan yang lebih terpercaya; Visualisasi yang Informatif, SPSS menyediakan berbagai pilihan visualisasi data seperti diagram batang, histogram, dan scatter plot, yang membantu menyederhanakan interpretasi data ilmiah yang kompleks; Fleksibilitas dalam Metodologi, SPSS mendukung berbagai metode statistik, mulai dari analisis deskriptif hingga multivariat, memungkinkan peneliti menggunakan pendekatan yang paling sesuai dengan jenis data yang dimiliki.

#### 1) Penelitian Eksperimen

SPSS digunakan untuk menganalisis hasil eksperimen di laboratorium, seperti pengujian hipotesis tentang efek perlakuan tertentu terhadap variabel yang diukur (Yildizhan et al., 2024), misalnya dalam studi biologi molekuler atau fisika material.

#### 2) Studi Lingkungan

Dalam penelitian lingkungan, SPSS membantu menganalisis data seperti kualitas air, polusi udara, atau distribusi spesies (Gandidzanwa & Togo, 2024). Analisis regresi dan korelasi sering digunakan untuk memahami hubungan antara faktor lingkungan.

#### 3) Analisis Data Observasi

Peneliti di bidang astronomi, ekologi, atau geologi menggunakan SPSS untuk mengolah data observasi dalam jumlah besar, seperti menganalisis pola pergerakan bintang atau distribusi populasi hewan.

#### 4) Studi Interdisipliner

Dalam studi sains yang melibatkan beberapa disiplin, seperti bioteknologi atau bioinformatika, SPSS membantu menganalisis data yang kompleks untuk menghasilkan temuan yang dapat diterapkan dalam pengembangan teknologi atau aplikasi lainnya.

#### **d. Penggunaan SPSS di Berbagai Bidang Teknik**

SPSS (Statistical Package for the Social Sciences) adalah perangkat lunak statistik yang banyak digunakan tidak hanya di bidang sosial, tetapi juga dalam berbagai bidang teknik (Seddaoui et al., 2025). SPSS membantu para insinyur dan peneliti teknik dalam menganalisis data, memvalidasi model, dan mendukung pengambilan keputusan berbasis data (Azmi et al., 2025).

##### **1) Teknik Industri**

- SPSS digunakan untuk menganalisis proses produksi, seperti evaluasi kualitas produk melalui analisis statistik (contohnya uji kontrol kualitas dan *Six Sigma*).
- Analisis efisiensi kerja dan produktivitas karyawan juga dapat dilakukan dengan SPSS menggunakan uji regresi atau analisis korelasi.

##### **2) Teknik Sipil**

- SPSS membantu dalam analisis data proyek konstruksi, seperti hubungan antara material yang digunakan dan durabilitas struktur.
- Studi risiko pada proyek besar, seperti identifikasi faktor risiko kegagalan proyek, dilakukan dengan uji statistik berbasis data historis .

##### **3) Teknik Mesin**

- SPSS digunakan untuk analisis data eksperimen dalam pengembangan mesin, seperti hubungan antara kecepatan rotasi mesin dengan konsumsi energi atau efisiensi operasional.
- Analisis keandalan (*reliability analysis*) pada komponen mekanis juga sering dilakukan untuk memprediksi masa pakai produk

#### 4) Teknik Lingkungan

- Dalam teknik lingkungan, SPSS digunakan untuk analisis kualitas air, udara, dan tanah, dengan tujuan memahami dampak pencemaran terhadap lingkungan.
- Analisis regresi digunakan untuk memprediksi dampak parameter lingkungan terhadap perubahan ekosistem.

#### 5) Teknik Elektro dan Elektronika

- Penelitian di bidang elektro memanfaatkan SPSS untuk menganalisis data eksperimen perangkat, seperti hubungan antara daya listrik yang digunakan dengan efisiensi alat.
- Dalam pengembangan teknologi komunikasi, SPSS membantu menganalisis data performa jaringan, misalnya, untuk optimasi sistem.

#### 6) Teknik Informatika dan Data Science

- SPSS sering digunakan untuk analisis dataset besar, seperti mempelajari pola penggunaan perangkat lunak atau tren dalam pengembangan teknologi (Duranteau et al., 2024; Wronkowski, 2024).
- Dengan kemampuan visualisasi datanya, SPSS mempermudah peneliti untuk memahami hasil analisis algoritma atau model pembelajaran mesin

Manfaat SPSS dalam Bidang Teknik antara lain : Meningkatkan Efisiensi dan Akurasi, SPSS membantu dalam pengolahan data besar dengan cepat dan akurat, sehingga mendukung keputusan teknis yang berbasis bukti; Memvalidasi Model Teknik, Dengan fitur statistik inferensial dan prediktif (Liu et al., 2024), SPSS mendukung validasi model, seperti model simulasi atau prediksi dalam berbagai aplikasi teknik, Mendukung Manajemen Risiko, Dalam proyek teknik, SPSS memungkinkan analisis faktor risiko secara statistik, membantu pengambilan keputusan untuk meminimalkan potensi kerugian; Visualisasi Data yang Informatif, Kemampuan SPSS untuk

menghasilkan grafik dan tabel mempermudah penyampaian informasi teknis kepada tim proyek atau pemangku kepentingan.

### **1.3. Hubungan Antara SPSS dengan Era Digital dan Data**

Dunia saat ini berada dalam era digital, di mana teknologi telah mengubah cara kita hidup, bekerja, dan berinteraksi. Setiap aktivitas manusia, baik secara individu maupun kolektif, menghasilkan data yang melimpah. Mulai dari transaksi online, media sosial, penggunaan perangkat IoT (Internet of Things), hingga penelitian ilmiah, data menjadi inti dari aktivitas manusia modern. Revolusi digital ini membawa tantangan baru yaitu bagaimana mengelola, menganalisis, dan memanfaatkan data dalam jumlah besar untuk mendukung pengambilan keputusan. Di sinilah SPSS (Statistical Package for the Social Sciences) menjadi alat yang relevan. SPSS hadir sebagai solusi bagi individu dan organisasi untuk mengolah data secara efisien, bahkan bagi mereka yang tidak memiliki latar belakang teknis yang mendalam. SPSS dirancang untuk membantu pengguna mengelola data, menjalankan analisis statistik, dan menghasilkan wawasan berbasis data. Dalam era digital, kemampuan ini sangat penting karena data telah menjadi "emas baru" yang menentukan keberhasilan di berbagai bidang. SPSS sebagai alat untuk mengolah data di Era Digital dalam hal

#### **(a) Pengelolaan Data dalam Skala Besar**

SPSS mampu menangani data dalam berbagai format digital, seperti CSV, Excel, atau database SQL. Kemampuan ini memudahkan pengguna untuk mengimpor, membersihkan, dan mengelola data yang dihasilkan dari perangkat digital. Era digital membawa tantangan analisis big data, yaitu kumpulan data yang sangat besar dan kompleks. SPSS mendukung proses ini dengan menyediakan alat-alat statistik yang kompatibel dengan analisis data dalam skala besar. Selain itu, SPSS juga digunakan dalam

proses eksplorasi awal data yang akan digunakan dalam model machine learning. SPSS dapat diintegrasikan dengan perangkat lunak lain yang sering digunakan dalam era digital, seperti Python, R, atau perangkat manajemen database. Hal ini memungkinkan SPSS untuk berfungsi sebagai bagian dari ekosistem analisis data yang lebih besar;

(b) Analisis Cepat dan Akurat

Dalam era di mana waktu sangat berharga, SPSS memungkinkan analisis data dilakukan dengan cepat dan akurat. Perangkat lunak ini dilengkapi dengan fitur otomatisasi untuk berbagai analisis statistik, mulai dari analisis deskriptif hingga inferensial. Dalam dunia bisnis dan penelitian, kemampuan untuk memprediksi tren masa depan adalah keunggulan kompetitif. SPSS menyediakan modul untuk analitik prediktif yang memungkinkan organisasi membuat keputusan berdasarkan data historis dan tren saat ini.

(c) Visualisasi Data yang Informatif

Era digital menuntut penyajian data yang menarik dan mudah dipahami. SPSS menawarkan berbagai jenis visualisasi, seperti grafik batang, histogram, scatter plot, dan diagram kotak, yang membantu pengguna memahami pola dan tren dalam data.

Di era digital, di mana data adalah inti dari setiap aktivitas, SPSS membuktikan dirinya sebagai alat yang relevan untuk mengolah, menganalisis, dan menginterpretasikan data. Dengan kemampuannya yang fleksibel, SPSS tidak hanya menjadi alat statistik tetapi juga mitra strategis dalam memanfaatkan potensi data untuk menciptakan peluang baru. SPSS membantu individu dan organisasi untuk menjawab tantangan era digital dengan memberikan kemampuan untuk:

- 1) Mengolah data digital dalam berbagai format.
- 2) Melakukan analisis statistik yang cepat dan akurat.

- 3) Menyampaikan wawasan yang dapat ditindaklanjuti melalui visualisasi yang efektif.

Dengan demikian, SPSS menjadi alat yang tidak tergantikan dalam upaya menghubungkan dunia digital yang penuh data dengan keputusan strategis yang lebih baik.

#### **1.4. Pentingnya Statistik dan SPSS**

Salah satu kendala terbesar yang dihadapi oleh banyak peneliti dan praktisi adalah kesenjangan antara teori statistik dan implementasinya dalam perangkat lunak. Banyak yang memahami teori statistik secara konseptual, tetapi kesulitan menghubungkannya dengan proses analisis menggunakan alat seperti SPSS. Sebaliknya, ada juga yang dapat mengoperasikan SPSS tetapi tidak memahami dasar teori yang mendasari analisis yang dilakukan. Ketidakseimbangan ini sering kali menghasilkan interpretasi yang keliru atau analisis yang tidak valid. Hal ini menunjukkan perlunya panduan yang tidak hanya mengajarkan cara menggunakan SPSS, tetapi juga menjelaskan teori statistik di balik setiap langkah analisis.

Statistik adalah cabang ilmu yang berperan penting dalam memberikan landasan untuk memahami pola, tren, dan hubungan dalam data (Sreenivasulu & Mamilla, 2024; Wang et al., 2024). Penggunaan statistik telah meluas ke berbagai bidang, seperti ilmu sosial, kesehatan, pendidikan, ekonomi, sains dan teknologi. Statistik menjadi alat penting dalam pengambilan keputusan berbasis data, baik di dunia akademik maupun profesional. Dalam dunia akademik, penelitian yang dilakukan oleh mahasiswa dan dosen sangat bergantung pada kemampuan untuk menganalisis data secara objektif. Di dunia profesional, statistik digunakan untuk memahami perilaku konsumen, mengukur kinerja organisasi, hingga memprediksi tren di masa depan (Herawati et al., 2023). Statistik tidak hanya memberikan wawasan, tetapi juga membantu membuktikan hipotesis

secara ilmiah. Kompleksitas analisis data seringkali menjadi tantangan bagi banyak individu. Di sinilah perangkat lunak statistik seperti SPSS hadir sebagai solusi yang memungkinkan analisis data dilakukan secara efisien dan akurat tanpa harus memahami setiap detail perhitungan manual.

Pentingnya statistik dan SPSS dalam hal meningkatkan keandalan analisis statistik memungkinkan penelitian dan pengambilan keputusan yang didukung oleh data. SPSS menyediakan platform untuk melakukan analisis statistik dengan akurasi tinggi, Relevansi SPSS dalam Dunia Profesional Banyak organisasi dan institusi mengadopsi SPSS sebagai alat utama mereka untuk menganalisis data. Oleh karena itu, memahami penggunaannya adalah keterampilan yang sangat berharga. SPSS mempermudah proses ini dengan menyediakan antarmuka yang ramah pengguna, berbagai metode analisis statistik, dan kemampuan untuk menangani data dalam skala besar (Ajalli, 2024; Sreenivasulu & Mamilla, 2024). Dengan memanfaatkan SPSS, pengambilan keputusan dapat dilakukan dengan lebih cepat, akurat, dan berbasis bukti. SPSS menawarkan berbagai fitur, seperti:

- a. Analisis Deskriptif: Menyediakan Ringkasan Statistik Seperti Mean, Median, dan Distribusi Data;
- b. Uji Statistik Inferensial: Meliputi Uji-T, ANOVA, Regresi, dan Analisis Korelasi;
- c. Visualisasi Data: Kemampuan Untuk Membuat Grafik Dan Tabel Yang Menarik dan Informatif;
- d. Kompatibilitas Data: memungkinkan Pengguna Untuk Mengimpor Data Dari Excel, CSV, Atau Basis Data Lainnya.

Meskipun SPSS telah digunakan secara luas, masih banyak tantangan yang dihadapi pengguna, antara lain: Keterbatasan Pemahaman Teori, Banyak pengguna yang mengandalkan SPSS tanpa

memahami logika statistik di balik analisis yang dilakukan; Kesalahan Interpretasi, Hasil yang dihasilkan oleh SPSS sering disalahartikan karena kurangnya pemahaman tentang makna statistik dari output yang dihasilkan; Kurangnya Panduan Praktis, Banyak panduan yang terlalu teknis atau hanya fokus pada salah satu aspek (teori atau implementasi), sehingga pengguna kesulitan menerapkan hasil analisis dalam konteks yang relevan. Buku ini hadir untuk menjawab kebutuhan tersebut dengan mengintegrasikan teori, implementasi, dan interpretasi dalam satu kesatuan panduan yang praktis dan aplikatif. Buku Pengantar SPSS: Teori, Implementasi, dan Interpretasi ditulis dengan tujuan untuk memberikan panduan yang komprehensif kepada pembaca. Statistik adalah fondasi yang kokoh untuk pengambilan keputusan di dunia modern (Sujianto et al., 2024; Tavani et al., 2024). Dengan memanfaatkan statistik, individu dan organisasi dapat mengambil keputusan yang lebih baik, lebih cepat, dan lebih bertanggung jawab. Sebagai pilar keputusan, statistik memberikan Keakuratan: Membantu menghasilkan hasil yang terpercaya berdasarkan data; Objektivitas: Mengurangi bias dalam proses pengambilan keputusan; Keberlanjutan: Mendukung keputusan yang berorientasi jangka panjang. Dalam era di mana data adalah kekuatan baru, kemampuan untuk memahami dan menggunakan statistik dengan benar akan menjadi keunggulan kompetitif yang signifikan. Dengan demikian, statistik tidak hanya menjadi alat analisis, tetapi juga panduan utama dalam menciptakan masa depan yang lebih baik.

Buku ini dirancang untuk: Membangun Pemahaman Teori Statistik, Menyediakan penjelasan dasar tentang konsep-konsep statistik yang relevan dengan analisis data; Mengajarkan Penggunaan SPSS secara Praktis: Membimbing pengguna melalui langkah-langkah teknis dalam menganalisis data menggunakan SPSS; Meningkatkan Kemampuan Interpretasi; Membantu pembaca memahami arti hasil analisis statistik dalam konteks penelitian atau aplikasi praktis. Buku

ini hadir untuk menjawab kebutuhan tersebut dengan mengintegrasikan teori, implementasi, dan interpretasi dalam satu kesatuan panduan yang praktis dan aplikatif. Dengan pendekatan yang terstruktur, buku ini cocok untuk mahasiswa yang baru mempelajari statistik, peneliti yang ingin meningkatkan keterampilan analisis data, hingga profesional yang menggunakan statistik dalam pekerjaan sehari-hari.

## DAFTAR PUSTAKA

- Ajalli, M. (2024). *The Optimization Model for Allocating Reward to Employees using GAHP and Cluster Analysis*. x(x), 414–426.
- Azmi, N. A. H. M., Ismail, F., Sarwin, N. M., Shafi, M. A., & Owaida, N. I. (2025). The Predictive Machine Learning Model Towards Effect of Human Resources Management Practices and Job Performances Among Academic Librarians in Malaysia. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 52(1), 259–270. <https://doi.org/10.37934/araset.52.1.259270>
- Duranteau, O., Blanchard, F., Popoff, B., van Etten-Jamaludin, F. S., Tuna, T., & Preckel, B. (2024). Mapping the landscape of machine learning models used for predicting transfusions in surgical procedures: a scoping review. *BMC Medical Informatics and Decision Making*, 24(1). <https://doi.org/10.1186/s12911-024-02729-3>
- Gandidzanwa, C. P., & Togo, M. (2024). Impacts of water, energy, and food nexus challenges on household vulnerability: the case of Harare City, Zimbabwe. *Environmental Research Letters*, 19(11). <https://doi.org/10.1088/1748-9326/ad7bcd>
- Herawati, S., Dadang, S., Ely, M., & Anang Sugeng, C. (2023). *Pengurangan Risiko Bencana*. Literasi Nusantara.
- Hoo, W. C., Lean, N. G., Ng, A. H. H., Liau, C. H., Ho, R. C., & Prompanyo, M. (2024). Determinants of Green Cosmetic Purchase Intentions in Malaysia. *Pakistan Journal of Life and Social Sciences (PJLSS)*, 22(2), 3811–3822. <https://doi.org/10.57239/pjlss-2024-22.2.00278>
- Jozpanahi, M., Mahmoudi, P., Moharrer, M., Taheri, H., & Kamali, K. (2025). Determination of vitamin D serum level in patients hospitalized due to herpes zoster. *Journal of Medicinal and Pharmaceutical Chemistry Research*, 7(2), 266–274. <https://doi.org/10.48309/jmpcr.2025.431297.1064>
- Liu, F., Yu, P., Wu, J., & Guo, L. (2024). The influence of exercise adherence on peace of mind among Chinese college students: a moderated chain mediation model. *Frontiers in Public Health*,

- 12(August), 1–11. <https://doi.org/10.3389/fpubh.2024.1447429>
- Salah, H. M., Hashem, A. A. R., Mustafa, T., Soliman, A. H., Khallaf, M., & Haddadeen, H. (2024). The impact of root end filling material type and the application of bone graft on healing of periapical tissues after endodontic microsurgery (a clinical randomized controlled trial). *Scientific Reports*, *14*(1), 1–13. <https://doi.org/10.1038/s41598-024-66033-w>
- Seddaoui, R., Abdullah, S. S., Ahmad, S., & Gorundutse, A. H. (2025). Influence of Marketing Innovation and Advanced Technologies on Firm Performance: The Case of Algeria. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *48*(2), 183–196. <https://doi.org/10.37934/araset.48.2.183196>
- Sreenivasulu, A., & Mamilla, R. (2024). Association Regarding S&P BSE Market and Specific Financial Parameters: A Statistical Analysis. *International Research Journal of Multidisciplinary Scope*, *5*(3), 331–336. <https://doi.org/10.47857/irjms.2024.v05i03.0720>
- Su, Y., & Ju, K. (2025). The current situation and strategies for cultivating college students' waste sorting behaviors through interactive media intervention. *Multidisciplinary Science Journal*, *7*(1). <https://doi.org/10.31893/multiscience.2025032>
- Sujianto, S., Hasibuan, A. M., Mahendri, I. G. A. P., Pribadi, E. R., Pujiharti, Y., Ardana, I. K., Ermianti, Sudjarmoko, B., Sukamto, & Gunawan, E. (2024). Drivers of Organic Rice Adoption Among Smallholder Farmers: Implications for Sustainable Development. *International Journal of Sustainable Development and Planning*, *19*(1), 289–299. <https://doi.org/10.18280/ijstdp.190127>
- Tavani, F. M., Behshid, M., Amini, A., Rahmani, A., Rahmani, P., & Mousavi, S. (2024). The effects of e-learning using educational multimedia on the ethical decision-making and professionalism of nursing students during the COVID-19 pandemic: a quasi-experimental study. *BMC Medical Education*, *24*(1), 1232. <https://doi.org/10.1186/s12909-024-05846-9>
- Wang, Y., Pan, B., & Zhang, Q. (2024). Rockery morphology based on quantitative analysis of shading. *Heritage Science*, *12*(1), 1–40. <https://doi.org/10.1186/s40494-024-01466-0>

- Wronkowski, A. (2024). Discovering the meaning of contemporary urban squares for its users—a case study of Poznan, Poland. *Humanities and Social Sciences Communications*, *11*(1), 1–11. <https://doi.org/10.1057/s41599-024-03497-1>
- Yildizhan, H., Udriștioiu, M. T., Pekdogan, T., & Ameen, A. (2024). Observational study of ground-level ozone and climatic factors in Craiova, Romania, based on one-year high-resolution data. *Scientific Reports*, *14*(1), 26733. <https://doi.org/10.1038/s41598-024-77989-0>

# **VARIABEL DAN JENIS DATA DALAM SPSS**

*Oleh M. Rosihan Arsyad*

## **2.1 Variabel dan Jenis Data dalam SPSS**

Konsep dasar variabel dan data adalah elemen fundamental dalam penelitian kuantitatif, karena variabel mewakili aspek atau karakteristik tertentu yang diukur atau diamati untuk menguji hipotesis penelitian. Dalam penelitian kuantitatif, variabel adalah komponen yang diidentifikasi, diukur, dan dianalisis untuk memahami hubungan atau pengaruh di antara mereka (SPSS, 2015). Secara sederhana, variabel bisa dianggap sebagai simbol yang merepresentasikan objek, orang, atau kejadian yang memiliki nilai yang dapat berubah, tergantung pada konteks dan kondisi penelitian (Herlina, 2019).

### **1. Variabel sebagai Dasar Analisis Kuantitatif**

Variabel adalah bagian yang tak terpisahkan dalam desain penelitian kuantitatif. Penelitian kuantitatif berfokus pada pengumpulan dan analisis data numerik untuk menemukan pola, hubungan, atau perbedaan antara variabel. Sebagai contoh, dalam studi tentang pengaruh lama belajar terhadap nilai ujian, lama belajar

dapat dikategorikan sebagai *variabel* bebas (*independen*) dan nilai ujian sebagai *variabel* terikat (*dependen*). Dalam hal ini, peneliti menguji hipotesis untuk melihat bagaimana variabel bebas berpengaruh terhadap variabel terikat.

## 2. Jenis-Jenis Variabel dalam Penelitian Kuantitatif

Dalam penelitian kuantitatif, variabel sering kali dikelompokkan menjadi berbagai jenis sesuai dengan perannya, yaitu:

- a) **Variabel Independen:** Variabel yang dianggap mempengaruhi atau menyebabkan perubahan pada variabel lain. Dalam konteks penelitian, variabel ini sering dimanipulasi atau dikontrol oleh peneliti.
- b) **Variabel Dependen:** Variabel yang hasilnya diukur untuk melihat pengaruh dari variabel independen.
- c) **Variabel Kontrol:** Variabel yang perlu dikendalikan agar tidak memengaruhi hasil penelitian, sehingga peneliti bisa memastikan bahwa hubungan antara variabel independen dan dependen tetap valid.
- d) **Variabel Moderator dan Mediasi:** Terkadang, ada variabel lain yang dapat memoderasi atau memediasi hubungan antara variabel utama. Misalnya, dalam penelitian tentang pengaruh motivasi belajar terhadap prestasi akademik, dukungan keluarga mungkin menjadi variabel moderator yang memperkuat atau melemahkan hubungan antara motivasi dan prestasi akademik.

## 3. Data Sebagai Representasi Variabel

Data adalah bentuk konkret dari variabel yang telah diukur. Data ini bisa berupa angka, kategori, atau skala, tergantung pada jenis variabelnya. Misalnya, jika variabelnya adalah usia, datanya bisa berupa angka seperti 20, 30, atau 40. Dalam konteks penelitian, data

dikumpulkan dari variabel yang sudah ditentukan dan dianalisis untuk menarik kesimpulan. Pengelolaan dan pengukuran data yang tepat akan memberikan informasi yang akurat dan valid untuk menjawab pertanyaan penelitian (Sitoayu et al., 2020).

#### **4. Peran Variabel dan Data dalam Menjawab Hipotesis Penelitian**

Hubungan antara variabel adalah inti dari penelitian kuantitatif. Dengan menganalisis data yang dikumpulkan dari variabel, peneliti dapat menguji hipotesis dan melihat apakah ada hubungan sebab-akibat atau korelasi antara variabel-variabel yang ada. Penggunaan variabel dan data dalam penelitian kuantitatif memungkinkan peneliti untuk menjelaskan fenomena, memperkirakan hasil, dan membuat keputusan berdasarkan bukti yang objektif.

Secara keseluruhan, variabel dan data adalah elemen dasar yang memberikan kerangka dan arah bagi penelitian kuantitatif. Tanpa pemahaman yang kuat tentang konsep-konsep ini, penelitian dapat kehilangan fokus dan akurasi. Oleh karena itu, penting bagi peneliti untuk memahami bagaimana mengidentifikasi, mengukur, dan menganalisis variabel, serta menginterpretasikan data dengan tepat untuk memperoleh hasil yang valid dan dapat diandalkan.

#### **2.2 Variabel dalam Konteks Penelitian dan SPSS**

Dalam konteks penelitian kuantitatif, variabel adalah elemen inti yang menjadi fokus analisis, karena variabel memungkinkan peneliti mengukur, mengamati, dan memahami fenomena yang diteliti. Konsep dasar variabel ini berkaitan erat dengan bagaimana fenomena penelitian dioperasionalkan dan dikonseptualisasikan untuk menjawab pertanyaan penelitian atau menguji hipotesis (Nugroho & Haritanto, 2022). SPSS (*Statistical Package for the Social Sciences*) sebagai perangkat lunak statistik memainkan peran besar

dalam mempermudah analisis variabel ini, memungkinkan peneliti untuk menyusun, mengelola, dan menganalisis data kuantitatif dengan akurat dan efisien.

### 1. Pengertian Variabel dalam Penelitian

Dalam penelitian kuantitatif, variabel adalah atribut atau karakteristik yang memiliki nilai atau kategori yang bervariasi dan dapat diukur. Variabel tersebut menjadi objek utama yang akan diukur atau diobservasi dalam proses pengumpulan data. Misalnya, dalam studi tentang pengaruh waktu belajar terhadap prestasi akademik, waktu belajar dan prestasi akademik adalah variabel yang akan diukur. Variabel ini memungkinkan peneliti mengidentifikasi pola atau hubungan di antara elemen-elemen yang diteliti (Ramadhani et al., 2021).

### 2. Jenis-Jenis Variabel dalam Penelitian

Ada beberapa jenis variabel yang biasa digunakan dalam penelitian kuantitatif:

- a. **Variabel Independen (Bebas):** Variabel yang dianggap sebagai penyebab atau faktor yang memengaruhi variabel lain. Dalam eksperimen, variabel ini sering kali dimanipulasi untuk melihat dampaknya terhadap variabel lain.
- b. **Variabel Dependen (Terikat):** Variabel yang dipengaruhi oleh variabel independen dan biasanya menjadi variabel yang diukur hasilnya dalam penelitian.
- c. **Variabel Kontrol:** Variabel tambahan yang dikendalikan atau diatur agar tidak memengaruhi hasil penelitian, menjaga keakuratan antara hubungan variabel independen dan dependen.
- d. **Variabel Moderator dan Mediasi:** Variabel yang dapat memoderasi atau memediasi hubungan antara variabel

independen dan dependen, memperkuat atau menengahi hubungan tersebut.

### 3. Peran Variabel dalam Analisis dengan SPSS

SPSS membantu peneliti mempermudah proses pengelolaan, pemrosesan, dan analisis variabel. SPSS memungkinkan peneliti mendefinisikan dan mengatur jenis variabel yang berbeda, serta menentukan jenis skala pengukurannya, seperti nominal, ordinal, interval, atau rasio. Berikut adalah beberapa cara bagaimana variabel dikelola dalam SPSS:

- a. **Definisi Variabel:** Dalam SPSS, peneliti dapat membuat definisi dan label variabel yang lebih jelas, sehingga lebih mudah dibedakan dalam hasil analisis.
- b. **Pemilihan Skala Pengukuran:** SPSS memungkinkan peneliti untuk menentukan skala pengukuran (nominal, ordinal, interval, rasio) yang sesuai, yang akan menentukan metode analisis yang digunakan.
- c. **Pengolahan Data:** SPSS mendukung berbagai teknik pengolahan data berdasarkan variabel yang ada, seperti uji korelasi untuk melihat hubungan antara variabel atau analisis regresi untuk melihat pengaruh variabel independen terhadap variabel dependen.

### 4. Peran Variabel dalam Menjawab Hipotesis Penelitian

Di dalam penelitian kuantitatif, variabel adalah komponen yang membantu peneliti menguji hipotesis dan menjawab pertanyaan penelitian secara empiris. Dengan mengidentifikasi hubungan atau pengaruh di antara variabel, peneliti dapat menentukan pola atau kesimpulan yang relevan. Misalnya, jika hipotesis menyatakan bahwa terdapat pengaruh positif antara waktu belajar dan prestasi akademik,

maka analisis variabel di SPSS dapat memberikan bukti apakah hipotesis ini benar atau tidak.

## **5. Signifikansi Variabel dalam Membangun Keterandalan Penelitian**

Pemahaman yang baik tentang variabel dan cara mengelolanya dalam SPSS memungkinkan peneliti untuk membuat keputusan yang lebih tepat dan analisis yang lebih akurat. Dengan menggunakan variabel sebagai komponen dasar analisis dalam SPSS, peneliti dapat memastikan bahwa hasil penelitian dapat diandalkan dan dapat diinterpretasikan dengan benar.

Secara keseluruhan, variabel memainkan peran kunci dalam proses analisis kuantitatif. Melalui SPSS, variabel tidak hanya diukur dan dianalisis secara numerik tetapi juga diproses secara sistematis untuk menghasilkan pemahaman yang lebih dalam tentang fenomena yang sedang dipelajari.

### **2.3 Jenis-Jenis Variabel dalam SPSS**

Dalam penelitian kuantitatif menggunakan SPSS, pemahaman tentang jenis-jenis variabel sangat penting untuk menentukan cara analisis yang tepat. SPSS menyediakan berbagai fitur untuk mengelola jenis variabel ini, sehingga peneliti dapat dengan mudah melakukan analisis statistik sesuai kebutuhan penelitian (Wahyuni, 2020). Berikut adalah jenis-jenis variabel utama dalam SPSS beserta perannya dalam penelitian kuantitatif:

#### **1. Variabel Independen (Bebas)**

Variabel independen adalah variabel yang dianggap sebagai penyebab atau faktor yang mempengaruhi variabel lain dalam penelitian. Dalam eksperimen atau studi kuantitatif, variabel ini dimanipulasi atau diukur untuk melihat pengaruhnya terhadap variabel dependen. Misalnya, dalam penelitian tentang pengaruh

metode belajar terhadap nilai ujian, metode belajar adalah variabel independen. Di SPSS, variabel independen ini dapat dianalisis untuk menguji hubungannya dengan variabel lain, seperti melalui regresi atau ANOVA (Analysis of Variance).

## **2. Variabel Dependen (Terikat)**

Variabel dependen adalah variabel yang dipengaruhi oleh variabel independen dan menjadi fokus hasil pengukuran dalam penelitian. Variabel ini menunjukkan hasil atau respon terhadap perubahan dalam variabel independen. Misalnya, jika variabel independennya adalah waktu belajar, maka variabel dependen dapat berupa nilai ujian yang diperoleh. Dalam SPSS, variabel dependen digunakan sebagai tolok ukur untuk melihat apakah perubahan pada variabel independen membawa dampak yang signifikan.

## **3. Variabel Kontrol**

Variabel kontrol adalah variabel tambahan yang dikendalikan atau dibuat tetap selama penelitian untuk memastikan bahwa hanya variabel independen yang mempengaruhi variabel dependen. Hal ini penting untuk mengurangi bias atau faktor eksternal yang bisa mempengaruhi hasil penelitian. Misalnya, dalam studi tentang efek latihan fisik pada berat badan, variabel seperti pola makan atau usia mungkin perlu dikontrol agar hasil yang diperoleh lebih akurat. Di SPSS, variabel kontrol sering diperhitungkan dalam model analisis dengan menambahkannya dalam analisis kovarian atau teknik statistik lain yang sesuai.

## **4. Variabel Moderator**

Variabel moderator adalah variabel yang mempengaruhi kekuatan atau arah hubungan antara variabel independen dan dependen. Variabel ini membantu peneliti memahami bagaimana atau

kan suatu hubungan terjadi. Misalnya, dalam hubungan antara motivasi dan prestasi akademik, dukungan keluarga bisa menjadi variabel moderator yang memperkuat hubungan tersebut. Dalam SPSS, variabel moderator dapat dianalisis melalui interaksi variabel pada analisis regresi atau moderasi untuk melihat pengaruhnya terhadap hubungan variabel utama.

### **5. Variabel Mediasi**

Variabel mediasi adalah variabel yang menjelaskan proses atau mekanisme melalui mana variabel independen mempengaruhi variabel dependen. Variabel ini berfungsi sebagai penghubung antara kedua variabel utama tersebut. Misalnya, jika dalam studi pengaruh pendidikan orang tua terhadap prestasi anak, motivasi belajar menjadi faktor mediasi yang menjelaskan bagaimana pendidikan orang tua berdampak pada prestasi anak. Di SPSS, variabel mediasi dapat dianalisis menggunakan teknik analisis jalur (path analysis) atau regresi hierarki untuk mengevaluasi apakah terdapat efek mediasi.

### **6. Variabel Dummy**

Dalam penelitian kuantitatif, variabel dummy adalah variabel kategori yang diubah menjadi bentuk numerik untuk keperluan analisis statistik. Biasanya, variabel dummy digunakan ketika variabel kategori (seperti jenis kelamin atau status pernikahan) perlu diikutsertakan dalam analisis regresi. Di SPSS, variabel dummy dibuat dengan mengubah kategori menjadi kode, seperti 0 dan 1, untuk memungkinkan perhitungan numerik.

### **7. Jenis Skala Pengukuran dalam Variabel**

Variabel dalam SPSS juga dikelompokkan berdasarkan skala pengukuran, yang meliputi:

- a. **Nominal**: Data kategori tanpa urutan tertentu (misalnya jenis kelamin, warna favorit).
- b. **Ordinal**: Data yang memiliki urutan tetapi tidak ada jarak konsisten antara kategori (misalnya peringkat kepuasan).
- c. **Interval**: Data dengan jarak tetap tetapi tanpa nilai nol absolut (misalnya suhu).
- d. **Rasio**: Data dengan jarak tetap dan nol absolut (misalnya berat, tinggi).

### Peran Jenis Variabel dalam Penelitian Kuantitatif

Jenis-jenis variabel ini sangat berperan dalam membangun hipotesis, memilih metode analisis, dan menginterpretasi hasil penelitian. Dengan menggunakan SPSS, peneliti dapat mengelola variabel dengan lebih mudah dan tepat, memilih analisis yang sesuai dengan jenis variabel, serta mendapatkan hasil yang akurat untuk mendukung atau menolak hipotesis. Pemahaman tentang variabel dan skala pengukuran memungkinkan analisis statistik yang efektif dan memperkuat validitas hasil penelitian.

### 2.4 Pengelompokan Variabel dalam Skala Pengukuran Data

Dalam penelitian kuantitatif, pengelompokan variabel berdasarkan skala pengukuran data merupakan hal penting untuk menentukan jenis analisis yang tepat. Skala pengukuran ini mengacu pada cara data diukur dan dikategorikan, yang akan mempengaruhi bagaimana data diolah dan diinterpretasikan. Ada empat jenis utama skala pengukuran data dalam penelitian kuantitatif, yaitu skala nominal, ordinal, interval, dan rasio. Setiap skala memiliki karakteristik unik yang menjelaskan bagaimana data dikelompokkan dan bagaimana hubungan antarvariabel diinterpretasikan (Jaya et al., 2022).

### 1. Skala Nominal

Skala nominal adalah skala pengukuran yang paling sederhana dan hanya melibatkan pengelompokan data ke dalam kategori tanpa urutan tertentu. Variabel nominal tidak memiliki nilai numerik atau peringkat, sehingga hanya mengidentifikasi kategori yang berbeda. Contoh variabel nominal adalah jenis kelamin (laki-laki, perempuan), warna favorit (merah, biru, hijau), atau daerah asal (kota A, kota B, kota C). Dalam SPSS, variabel nominal sering kali diberi kode numerik untuk kemudahan analisis, tetapi angka tersebut tidak memiliki arti kuantitatif; angka hanya sekadar label (Santoso, 2016).

### 2. Skala Ordinal

Skala ordinal memberikan peringkat atau urutan pada kategori data, namun jarak antarperingkat tidak selalu sama atau konsisten. Skala ini memungkinkan peneliti mengelompokkan data ke dalam tingkat yang terurut tetapi tidak mengukur seberapa besar perbedaan antara tingkatan tersebut. Contoh variabel ordinal adalah peringkat kepuasan pelanggan (sangat tidak puas, tidak puas, netral, puas, sangat puas) atau tingkat pendidikan (SMP, SMA, sarjana, pascasarjana). Dalam SPSS, variabel ordinal memungkinkan analisis data berdasarkan urutan tetapi terbatas pada analisis yang tidak membutuhkan informasi tentang interval yang konsisten (Muzawi et al., 2024).

### 3. Skala Interval

Skala interval mengukur data dengan jarak yang sama atau konsisten antara nilai, tetapi tidak memiliki nol mutlak. Artinya, pada skala interval, nol tidak menunjukkan ketiadaan variabel. Skala ini memungkinkan peneliti untuk melakukan operasi aritmatika dasar, seperti penjumlahan dan pengurangan, karena jarak antar-nilai tetap sama. Contoh variabel interval adalah suhu dalam derajat Celsius atau

Fahrenheit, di mana nol tidak berarti tidak ada suhu tetapi hanya titik pada skala. Dengan skala interval, analisis yang lebih kompleks seperti uji rata-rata atau regresi bisa dilakukan di SPSS karena jarak antar-nilai memiliki makna yang konsisten.

#### 4. Skala Rasio

Skala rasio adalah skala pengukuran yang paling lengkap, karena memiliki jarak yang konsisten antara nilai dan juga nol mutlak, yang menunjukkan tidak adanya variabel tersebut. Skala ini memungkinkan perbandingan yang lebih absolut dan penghitungan operasi matematika penuh, termasuk perkalian dan pembagian. Contoh variabel rasio adalah tinggi badan, berat badan, usia, dan pendapatan, di mana nol berarti ketiadaan tinggi, berat, usia, atau pendapatan. Data rasio dalam SPSS memungkinkan berbagai teknik analisis statistik, seperti analisis regresi, korelasi, dan statistik deskriptif, karena skala ini mencakup semua sifat skala pengukuran lainnya.

#### Pentingnya Pengelompokan Variabel dalam Skala Pengukuran

Pengelompokan variabel ke dalam skala pengukuran yang tepat penting untuk menentukan jenis analisis statistik yang relevan. Setiap skala pengukuran memiliki karakteristik dan batasannya sendiri, sehingga pemilihan skala yang sesuai akan memengaruhi keakuratan hasil analisis dan interpretasi data. Misalnya, skala nominal dan ordinal biasanya menggunakan analisis non-parametrik, sementara skala interval dan rasio dapat menggunakan analisis parametrik, seperti uji t atau ANOVA.

#### Peran Skala Pengukuran dalam Analisis di SPSS

Dalam SPSS, penentuan skala pengukuran memudahkan peneliti dalam memilih teknik analisis statistik yang sesuai. SPSS

memungkinkan peneliti untuk menetapkan jenis skala setiap variabel dalam tampilan *Variable View*, yang akan membantu perangkat lunak mengenali tipe analisis yang cocok. Dengan menetapkan skala pengukuran yang tepat, SPSS dapat memberikan hasil yang lebih valid dan interpretasi yang lebih mendalam terhadap hubungan antarvariabel dalam penelitian.

Secara keseluruhan, pengelompokan variabel dalam skala pengukuran data adalah langkah mendasar dalam penelitian kuantitatif. Pemahaman yang baik tentang skala pengukuran ini memungkinkan analisis yang lebih akurat, pengambilan keputusan yang lebih tepat, dan hasil penelitian yang lebih andal (Santoso, 2019).

## **2.5 Implementasi dalam SPSS: Memasukkan dan Mengatur Variabel**

Implementasi dalam SPSS untuk memasukkan dan mengatur variabel dalam penelitian kuantitatif adalah langkah awal yang sangat penting. Dengan memasukkan variabel secara tepat, peneliti dapat memastikan data yang dihasilkan lebih akurat dan siap untuk dianalisis menggunakan berbagai metode statistik (Priyastama, 2020). Berikut ini adalah narasi langkah demi langkah mengenai cara memasukkan dan mengatur variabel dalam SPSS untuk penelitian kuantitatif:

### **1. Membuka SPSS dan Memahami Tampilan Awal**

Saat membuka SPSS, peneliti akan melihat dua jendela utama: *Data View* dan *Variable View*. *Data View* adalah tempat peneliti akan memasukkan data numerik, sementara *Variable View* adalah tempat di mana peneliti mengatur dan mendefinisikan variabel yang akan digunakan dalam penelitian. Pemahaman tentang perbedaan kedua jendela ini sangat penting karena setiap variabel dan data harus disesuaikan dengan format dan tipe pengukuran yang sesuai (Sujarweni & Utami, 2019).

## 2. Memasukkan Variabel dalam Variable View

Dalam *Variable View*, peneliti perlu memasukkan variabel satu per satu pada kolom *Name*, yang merupakan nama variabel tanpa spasi atau karakter khusus. Setiap baris di *Variable View* mewakili satu variabel dalam penelitian. Misalnya, jika penelitian kuantitatif mengukur variabel Usia, Jenis Kelamin, dan Nilai Ujian, peneliti memasukkan setiap variabel tersebut dalam kolom *Name* pada baris yang berbeda.

## 3. Mengatur Tipe Data pada Kolom Type

Setelah memasukkan nama variabel, peneliti perlu menentukan jenis data variabel tersebut pada kolom *Type*. SPSS menyediakan beberapa tipe data, seperti *Numeric* untuk data numerik dan *String* untuk data teks. Sebagian besar penelitian kuantitatif menggunakan data numerik karena lebih mudah dianalisis secara statistik, tetapi data nominal atau kategori (seperti jenis kelamin) juga dapat diatur dalam tipe *String* atau *Numeric* dengan kode tertentu (misalnya, 1 untuk laki-laki, 2 untuk perempuan).

## 4. Menentukan Skala Pengukuran pada Kolom Measure

Pada kolom *Measure*, peneliti menentukan skala pengukuran dari setiap variabel, yaitu nominal, ordinal, atau scale (yang mencakup interval dan rasio). Misalnya, jika variabel adalah usia, maka skala pengukurannya adalah *Scale* (karena usia memiliki interval yang konsisten dan nol mutlak). Jika variabelnya adalah jenis kelamin, maka skala pengukurannya adalah *Nominal*.

- a. **Nominal**: untuk data kategori tanpa urutan, seperti jenis kelamin atau tempat tinggal.
- b. **Ordinal**: untuk data dengan urutan tetapi tanpa jarak yang tetap, seperti tingkat pendidikan.

- c. **Scale:** untuk data dengan interval atau rasio yang memiliki jarak tetap dan nol mutlak, seperti usia atau pendapatan.

## **5. Memberikan Label pada Variabel di Kolom Label**

Kolom *Label* digunakan untuk memberikan keterangan yang lebih deskriptif tentang variabel, sehingga peneliti dan orang lain dapat lebih memahami makna dari setiap variabel. Misalnya, variabel dengan nama “Usia” dapat diberi label “Usia responden dalam tahun.” Hal ini membantu peneliti saat melihat data output sehingga variabel lebih mudah dikenali.

## **6. Menetapkan Nilai Kategori pada Kolom Values**

Untuk variabel kategori, seperti jenis kelamin atau status pekerjaan, peneliti perlu menetapkan *Values* agar setiap kode numerik memiliki arti yang jelas. Misalnya, untuk variabel jenis kelamin, nilai 1 dapat diberi label “Laki-laki” dan nilai 2 sebagai “Perempuan.” Penetapan ini membuat data lebih mudah dibaca dan menghindari kebingungan saat melakukan analisis.

## **7. Mengatur Missing Values**

Dalam beberapa kasus, ada data yang hilang atau tidak diisi oleh responden. SPSS memungkinkan peneliti untuk mengatur data yang hilang atau *Missing Values* agar tidak memengaruhi hasil analisis. Peneliti dapat menandai data hilang dengan angka tertentu, misalnya -99 atau -1, yang menandakan bahwa data tersebut hilang atau tidak diisi. Dengan demikian, SPSS akan mengenali dan mengabaikan data yang hilang dalam analisis.

## **8. Menyimpan dan Memeriksa Kembali Pengaturan Variabel**

Setelah memasukkan semua variabel dan pengaturannya, peneliti perlu menyimpan file data dan memeriksa kembali

pengaturan yang telah dibuat. Hal ini untuk memastikan bahwa semua variabel telah diatur dengan benar, sehingga analisis statistik yang dilakukan akan memberikan hasil yang valid dan dapat diandalkan.

## **9. Masuk ke Data View dan Memasukkan Data**

Setelah semua variabel diatur di *Variable View*, peneliti dapat beralih ke *Data View* untuk memasukkan data yang diperoleh dari hasil survei atau eksperimen. Di sini, setiap baris mewakili satu responden atau unit observasi, dan setiap kolom mewakili variabel yang telah ditetapkan di *Variable View*.

## **Manfaat Mengatur Variabel dengan Tepat dalam SPSS**

Mengatur variabel dengan tepat dalam SPSS memberikan kemudahan dan keakuratan dalam analisis data. Dengan pengaturan yang benar, peneliti dapat memilih teknik analisis statistik yang sesuai, menginterpretasi data dengan lebih akurat, dan meminimalkan kesalahan dalam pengolahan data. Hal ini memastikan bahwa hasil yang diperoleh valid dan dapat digunakan untuk menarik kesimpulan yang tepat dalam penelitian kuantitatif.

Implementasi ini menjadi dasar yang kuat untuk melakukan analisis lanjut dan membuat laporan yang lebih terstruktur dan profesional.

## **2.6 Pengaturan Label dan Nilai dalam SPSS**

Dalam penelitian kuantitatif yang menggunakan SPSS, pengaturan label dan nilai adalah langkah penting untuk memastikan data dapat dibaca dan dianalisis dengan mudah. Dengan memberikan label dan nilai pada variabel, peneliti dapat membuat data lebih deskriptif dan terstruktur, sehingga interpretasi hasil menjadi lebih jelas dan akurat (Wibowo, 2023). Berikut adalah narasi mengenai pengaturan label dan nilai dalam SPSS untuk penelitian kuantitatif.

## 1 Pengaturan Label Variabel

Label variabel dalam SPSS adalah nama atau deskripsi yang memberikan informasi lebih rinci tentang suatu variabel. Dalam penelitian kuantitatif, nama variabel yang diinput biasanya singkat, sehingga pengaturan label diperlukan agar peneliti dan pembaca lainnya dapat memahami arti variabel dengan lebih baik (Chalimi & SE, 2021).

Misalnya, variabel dengan nama “Gaji” mungkin merupakan data tentang pendapatan bulanan, tetapi tanpa label yang lebih deskriptif, data ini mungkin membingungkan saat dianalisis. Dengan mengatur label menjadi “Pendapatan Bulanan dalam Rupiah”, peneliti menambahkan konteks yang lebih jelas, yang sangat membantu saat melihat hasil analisis. Proses pengaturan label variabel dilakukan di *Variable View*, tepatnya di kolom *Label*, di mana peneliti dapat mengetikkan deskripsi variabel dengan lebih lengkap (Ramadhayanti, 2019).

## 2. Pentingnya Pengaturan Label Variabel

Pengaturan label tidak hanya membantu peneliti dalam membaca data tetapi juga sangat berguna saat membuat laporan atau presentasi. Dalam tabel atau grafik hasil output SPSS, label variabel akan ditampilkan sebagai keterangan dari variabel terkait. Hal ini mempermudah peneliti dan audiens dalam memahami hasil tanpa kebingungan, terutama jika penelitian melibatkan banyak variabel atau variabel yang bersifat teknis.

## 3. Pengaturan Nilai Variabel (Value Labels)

Pengaturan nilai atau *value labels* dalam SPSS adalah proses memberikan deskripsi pada nilai-nilai kategori atau numerik dalam variabel tertentu, terutama pada variabel kategori atau nominal. Dengan memberi label pada nilai, peneliti dapat mengubah data

numerik menjadi kategori yang lebih bermakna, sehingga lebih mudah dibaca dan dianalisis.

Contoh: Dalam penelitian dengan variabel "Jenis Kelamin," data mungkin diinput sebagai angka 1 dan 2, di mana 1 mewakili "Laki-laki" dan 2 mewakili "Perempuan." Namun, angka-angka tersebut tidak informatif jika tidak diberi label. Dalam SPSS, peneliti dapat membuka kolom *Values* di *Variable View* dan mengatur nilai sebagai berikut:

- a. 1 diberi label "Laki-laki"
- b. 2 diberi label "Perempuan"

Proses ini membuat data lebih jelas dan terstruktur. Ketika hasil analisis ditampilkan di SPSS, nilai yang telah diberi label akan muncul sebagai "Laki-laki" dan "Perempuan" alih-alih hanya angka 1 dan 2, sehingga mempermudah interpretasi.

#### 4. Langkah-Langkah Mengatur Value Labels di SPSS

- a. Buka *Variable View* di SPSS.
- b. Pilih variabel yang akan diberi nilai pada kolom *Values*.
- c. Klik pada sel di kolom *Values* untuk membuka kotak dialog.
- d. Masukkan angka pada kolom *Value* dan deskripsi pada kolom *Value Label*.
- e. Tekan tombol *Add* setelah mengatur setiap nilai, kemudian klik *OK* untuk menyimpan pengaturan.

Langkah-langkah ini dapat diulang untuk setiap variabel kategori dalam penelitian.

#### 5. Keuntungan Mengatur Label dan Nilai dalam SPSS

Mengatur label dan nilai dalam SPSS memiliki banyak keuntungan:

- a. **Meningkatkan Keterbacaan Data:** Label membuat variabel lebih mudah dipahami, sementara value labels memberikan arti pada nilai numerik yang mungkin sulit dimengerti tanpa konteks.

- b. **Meminimalkan Kesalahan Interpretasi:** Dengan value labels, peneliti dan audiens dapat langsung mengerti arti nilai dalam hasil analisis, sehingga mengurangi kesalahan interpretasi.
- c. **Memperjelas Hasil Output:** Saat analisis selesai, hasil yang diperoleh dalam bentuk tabel atau grafik akan menampilkan label dan value labels, sehingga hasil lebih mudah dipresentasikan dan dipahami.

## 6. Penerapan Analisis Kuantitatif

Dalam penelitian kuantitatif yang menggunakan banyak variabel kategori (misalnya pendidikan, status pekerjaan, atau wilayah tempat tinggal), pengaturan value labels sangat penting agar hasil output dapat diinterpretasi dengan tepat. Value labels memungkinkan peneliti membedakan dan membandingkan kategori dengan mudah dalam analisis, terutama saat menggunakan teknik statistik seperti uji chi-square atau analisis varians (ANOVA).

Secara keseluruhan, pengaturan label dan nilai dalam SPSS adalah langkah sederhana namun sangat bermanfaat dalam meningkatkan keterbacaan dan kejelasan data penelitian kuantitatif. Dengan pengaturan ini, peneliti dapat memastikan hasil analisis lebih mudah dipahami, baik oleh diri sendiri maupun oleh audiens, yang merupakan hal penting dalam penelitian berbasis data.

### 2.7 Skala Pengukuran Analisis Data dalam SPSS

Dalam analisis data menggunakan SPSS, pemilihan skala pengukuran yang tepat sangat memengaruhi metode analisis yang digunakan dan kualitas hasil yang diperoleh. Skala pengukuran menentukan jenis data yang akan diolah dan teknik statistik yang cocok, sehingga penting bagi peneliti untuk memahami dampak dari pemilihan skala pengukuran serta kesalahan yang sering terjadi dan cara menghindarinya (Kusumawardhani & Astuti, 2024).

## 1. Dampak Pemilihan Skala Pengukuran terhadap Metode Analisis Data di SPSS

Pemilihan skala pengukuran memengaruhi metode analisis yang bisa diterapkan dalam SPSS karena setiap skala memiliki karakteristik khusus yang memengaruhi jenis teknik statistik yang cocok (Hidayat, 2021). Berikut ini adalah dampaknya berdasarkan skala pengukuran yang berbeda:

- a. **Skala Nominal:** Data nominal hanya mengelompokkan objek ke dalam kategori tanpa urutan atau perbedaan numerik. Metode analisis yang cocok adalah analisis frekuensi atau uji chi-square untuk melihat hubungan antar kategori. Analisis statistik parametrik, seperti uji t atau ANOVA, tidak sesuai untuk data nominal karena skala ini tidak memiliki urutan atau jarak yang konsisten antar kategori.
- b. **Skala Ordinal:** Data ordinal memungkinkan urutan tetapi tidak menyediakan jarak yang pasti antara kategori. Analisis deskriptif seperti median atau uji non-parametrik seperti uji Mann-Whitney dan uji Kruskal-Wallis adalah pilihan tepat untuk data ini. Menggunakan teknik parametrik pada skala ordinal dapat menghasilkan interpretasi yang tidak akurat karena skala ini tidak memiliki nilai interval yang konsisten.
- c. **Skala Interval:** Data interval memiliki jarak yang tetap, memungkinkan teknik statistik parametrik seperti korelasi Pearson dan ANOVA. Data interval tidak memiliki nol mutlak, sehingga tidak sesuai untuk perhitungan seperti rasio atau persentase.
- d. **Skala Rasio:** Dengan nol mutlak dan jarak tetap, data rasio dapat dianalisis menggunakan hampir seluruh metode statistik yang tersedia, termasuk regresi linier, uji t, dan analisis korelasi. Penggunaan skala rasio memungkinkan peneliti untuk melakukan analisis yang sangat mendetail,

seperti membandingkan perbedaan proporsi atau mengukur korelasi yang akurat antarvariabel.

Dengan memilih skala yang sesuai, SPSS dapat memberikan hasil analisis yang relevan dan interpretasi yang valid, sedangkan kesalahan dalam memilih skala dapat menyebabkan hasil analisis yang bias atau tidak valid.

## **2. Kesalahan Umum dalam Pemilihan Skala dan Konsekuensinya terhadap Hasil Analisis**

Kesalahan dalam memilih skala pengukuran sering terjadi, terutama ketika peneliti kurang memahami karakteristik data yang dikumpulkan. Beberapa kesalahan umum meliputi:

- a. **Menggunakan Teknik Parametrik pada Data Nominal atau Ordinal:** Teknik parametrik membutuhkan data dengan jarak interval yang konsisten, yang tidak dimiliki oleh data nominal dan ordinal. Jika teknik ini diterapkan pada data nominal atau ordinal, hasilnya bisa tidak akurat dan tidak valid.
- b. **Mengabaikan Nol Mutlak pada Data Interval:** Data interval tidak memiliki nol mutlak, sehingga mengukur rasio atau proporsi pada data ini tidak tepat. Jika peneliti menggunakan data interval seperti data rasio, hasilnya akan memberikan informasi yang menyesatkan.
- c. **Menganggap Data Ordinal sebagai Data Interval:** Dalam beberapa kasus, peneliti mungkin salah menganggap bahwa skala ordinal, seperti skala Likert, dapat dianalisis seperti data interval. Hal ini sering menyebabkan kesalahan interpretasi karena jarak antar pilihan pada skala ordinal tidak selalu konsisten.

Kesalahan-kesalahan ini berdampak pada keakuratan hasil analisis dan dapat menghasilkan kesimpulan yang salah, yang pada

gilirannya dapat menurunkan validitas dan keandalan penelitian secara keseluruhan.

### **3. Tips dalam Menentukan Skala yang Sesuai dengan Tujuan Penelitian dan Variabel yang Tersedia**

Agar analisis data yang dilakukan sesuai dengan tujuan penelitian, berikut adalah beberapa tips untuk menentukan skala pengukuran yang tepat:

- a. **Pahami Tujuan Pengukuran Variabel:** Sebelum menentukan skala, pahami tujuan pengukuran variabel. Jika tujuannya untuk membedakan kategori (misalnya jenis kelamin atau agama), skala nominal sudah cukup. Jika ingin mengurutkan variabel berdasarkan kategori, gunakan skala ordinal.
- b. **Perhatikan Sifat Data:** Tentukan apakah data memiliki nilai nol mutlak atau hanya urutan. Misalnya, data seperti usia atau pendapatan adalah data rasio karena memiliki nol mutlak, sedangkan skala Likert adalah skala ordinal karena hanya mengukur urutan atau tingkatan.
- c. **Periksa Kesesuaian dengan Analisis yang Diinginkan:** Pastikan skala pengukuran sesuai dengan analisis yang akan dilakukan. Jika analisis yang diinginkan membutuhkan data interval atau rasio (misalnya regresi linier), pastikan variabel diukur pada skala yang tepat.
- d. **Konsultasikan dengan Sumber Referensi Statistik:** Saat ragu, merujuk pada literatur statistik atau berkonsultasi dengan ahli statistik dapat membantu memastikan bahwa skala pengukuran sudah tepat.

Dengan mengikuti langkah-langkah ini, peneliti dapat memilih skala pengukuran yang sesuai untuk masing-masing variabel, meningkatkan kualitas data, dan memperoleh hasil analisis yang valid dan dapat diandalkan.

## 2.8 Studi Kasus: Menggunakan Variabel dan Jenis Data dalam SPSS

Studi kasus ini akan membahas bagaimana variabel dan jenis data digunakan dalam SPSS untuk analisis data dalam penelitian kuantitatif. Dengan pemilihan variabel dan penentuan jenis data yang tepat, peneliti dapat menggunakan SPSS untuk menghasilkan analisis yang akurat. Sebagai contoh, kita akan menggunakan sebuah studi kasus yang meneliti faktor-faktor yang memengaruhi kepuasan kerja karyawan di sebuah Perusahaan (Qudratullah, 2013).

### 1. Latar Belakang Penelitian

Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang memengaruhi tingkat kepuasan kerja di kalangan karyawan perusahaan. Beberapa variabel penting yang akan dianalisis meliputi Gaji, Lama Bekerja, Jenis Kelamin, Tingkat Pendidikan, dan Kepuasan Kerja. Setiap variabel diatur sesuai dengan jenis data yang tepat untuk memberikan hasil yang valid dalam SPSS.

### 2. Identifikasi dan Pengaturan Variabel dalam SPSS

Dalam SPSS, peneliti perlu terlebih dahulu menentukan jenis data untuk setiap variabel yang akan digunakan. Jenis data yang tepat memungkinkan SPSS untuk memilih metode analisis statistik yang sesuai. Berikut adalah rincian variabel dan jenis data dalam studi kasus ini:

- a. **Gaji:** Variabel ini diukur dalam mata uang dan dianggap sebagai data rasio karena memiliki nol mutlak dan jarak yang tetap antara nilai-nilai. Data ini memungkinkan perhitungan rata-rata, regresi, dan analisis parametrik lainnya.
- b. **Lama Bekerja:** Lama kerja diukur dalam tahun dan juga merupakan data rasio. Seperti *Gaji*, variabel ini memiliki nol mutlak, sehingga perbandingan dan analisis korelasi dapat dilakukan pada data ini.

- c. **Jenis Kelamin:** Data ini bersifat nominal dengan dua kategori: *Laki-laki* dan *Perempuan*. Di SPSS, variabel ini diatur sebagai data nominal dengan label nilai (1 untuk Laki-laki dan 2 untuk Perempuan). Teknik statistik non-parametrik, seperti uji chi-square, cocok untuk variabel nominal.
- d. **Tingkat Pendidikan:** Pendidikan dikategorikan ke dalam beberapa tingkatan, seperti *SMA*, *Diploma*, dan *S1*, sehingga merupakan data ordinal. Variabel ini diatur dalam SPSS sebagai data ordinal dan cocok untuk analisis non-parametrik seperti uji Mann-Whitney atau Kruskal-Wallis jika diperlukan.
- e. **Kepuasan Kerja:** Kepuasan kerja diukur melalui skala Likert (sangat tidak puas, tidak puas, netral, puas, sangat puas), yang merupakan data ordinal. Variabel ini diatur sebagai ordinal dalam SPSS, dan uji statistik yang sesuai dapat digunakan untuk mengukur pengaruh variabel lain terhadap kepuasan kerja.

### 3. Memasukkan dan Mengatur Data dalam SPSS

Setelah variabel dan jenis data ditentukan, langkah selanjutnya adalah memasukkan data ke dalam SPSS:

- a. Peneliti membuka *Variable View* di SPSS dan memasukkan setiap variabel pada kolom *Name*.
- b. Pada kolom *Type*, peneliti memilih jenis data: *Numeric* untuk *Gaji*, *Lama Bekerja*, dan variabel numerik lainnya, serta *String* jika diperlukan untuk data non-numerik.
- c. Kolom *Label* digunakan untuk memberikan deskripsi pada setiap variabel, misalnya, "Gaji bulanan karyawan dalam rupiah."
- d. Untuk variabel seperti *Jenis Kelamin* dan *Tingkat Pendidikan*, peneliti menetapkan *Value Labels* untuk memberikan keterangan pada setiap kategori (misalnya, 1=Laki-laki, 2=Perempuan).

- e. Kolom *Measure* diisi berdasarkan skala pengukuran setiap variabel: *Nominal*, *Ordinal*, atau *Scale* (untuk interval atau rasio).

#### 4. Melakukan Analisis di SPSS Berdasarkan Jenis Data

Setelah data dimasukkan, peneliti dapat mulai melakukan analisis statistik di SPSS:

- a. **Analisis Deskriptif:** Peneliti dapat menghasilkan statistik deskriptif untuk variabel rasio seperti Gaji dan Lama Bekerja, termasuk mean, median, dan standar deviasi, untuk mendapatkan gambaran umum dari data.
- b. **Uji Chi-Square untuk Jenis Kelamin:** Sebagai variabel nominal, peneliti dapat menggunakan uji *chi-square* untuk melihat apakah ada hubungan antara Jenis Kelamin dan Kepuasan Kerja.
- c. **Korelasi dan Regresi:** Untuk variabel Gaji dan Lama Bekerja (data rasio), SPSS memungkinkan analisis korelasi Pearson untuk melihat hubungan antarvariabel dan regresi linier untuk mengetahui pengaruh variabel-variabel ini terhadap Kepuasan Kerja.
- d. **Uji Non-Parametrik untuk Data Ordinal:** Variabel seperti Tingkat Pendidikan dan Kepuasan Kerja dapat dianalisis menggunakan uji non-parametrik jika distribusi data tidak memenuhi syarat normalitas. Misalnya, peneliti dapat menggunakan uji Kruskal-Wallis untuk membandingkan tingkat kepuasan berdasarkan tingkat pendidikan.

#### 5. Interpretasi Hasil Analisis

Dengan menggunakan variabel dan jenis data yang sesuai, peneliti dapat menginterpretasikan hasil secara akurat. Sebagai contoh, jika analisis menunjukkan korelasi positif antara Gaji dan Kepuasan Kerja, hal ini berarti semakin tinggi gaji, semakin tinggi pula tingkat kepuasan karyawan. Begitu pula, jika hasil uji chi-square

menunjukkan hubungan antara Jenis Kelamin dan Kepuasan Kerja, peneliti dapat mengeksplorasi lebih lanjut pengaruh demografis terhadap kepuasan.

## **6. Kesimpulan Studi Kasus**

Studi kasus ini menekankan pentingnya memahami jenis data dan variabel dalam analisis menggunakan SPSS. Dengan memilih skala pengukuran yang tepat dan metode analisis yang sesuai, peneliti dapat memastikan keakuratan hasil analisis dan menghindari kesalahan interpretasi. SPSS sebagai perangkat lunak statistik memberikan berbagai opsi analisis yang dapat disesuaikan dengan karakteristik data, sehingga sangat membantu peneliti dalam menghasilkan hasil yang dapat diandalkan untuk penelitian kuantitatif.

## **2.9 Kesalahan Umum dalam Mengelola Variabel dan Data di SPSS**

Kesalahan umum dalam mengelola variabel dan data di SPSS seringkali terjadi, terutama bagi pengguna yang baru memulai analisis statistik menggunakan perangkat lunak ini. Kesalahan tersebut bisa mempengaruhi validitas hasil analisis dan menyebabkan interpretasi yang keliru (Ikram & others, 2024). Berikut beberapa kesalahan umum dalam mengelola variabel dan data di SPSS serta dampaknya:

### **1. Memilih Skala Pengukuran yang Tidak Tepat**

Kesalahan ini sering terjadi ketika peneliti tidak benar-benar memahami skala pengukuran yang sesuai untuk data. Kesalahan dalam memilih skala pengukuran dapat menyebabkan pemilihan metode analisis yang tidak cocok. Misalnya, menggunakan uji parametrik pada data ordinal atau nominal yang seharusnya menggunakan uji non-parametrik akan menghasilkan hasil yang bias atau tidak valid.

- a. **Dampak:** Hasil analisis yang bias atau menyesatkan.

- b. **Solusi:** Pahami jenis data (nominal, ordinal, interval, atau rasio) dan sesuaikan dengan skala pengukuran yang tepat di SPSS.

## 2. Tidak Mengatur Label Variabel dan Nilai (Value Labels)

SPSS memungkinkan pengguna untuk menambahkan label pada variabel dan nilai kategori untuk membantu interpretasi data. Namun, beberapa peneliti seringkali mengabaikan pengaturan label, sehingga data menjadi sulit dibaca dan dipahami, terutama jika data kategori hanya terdiri dari angka atau kode.

- a. **Dampak:** Membingungkan saat interpretasi dan analisis, terutama jika melibatkan data kategori yang banyak.
- b. **Solusi:** Selalu tambahkan label pada variabel dan nilai kategori agar data lebih mudah dipahami.

## 3. Mengabaikan Penanganan Missing Value

Data yang hilang (missing value) adalah hal umum dalam penelitian, namun tidak menanganinya dengan tepat dapat mengurangi kualitas hasil analisis. Beberapa pengguna mengabaikan pengaturan missing value dalam SPSS atau tidak menyadari adanya data yang hilang, sehingga analisis dapat memberikan hasil yang tidak akurat.

- a. **Dampak:** Analisis yang bias dan hasil yang tidak valid karena missing value tidak diperhitungkan.
- b. **Solusi:** Identifikasi dan atur missing value di SPSS menggunakan fitur *Missing Value* di bagian *Variable View*. Lakukan imputasi data jika diperlukan atau analisis sensitivity untuk melihat dampak missing data.

## 4. Menggunakan Data Kategori sebagai Variabel Numerik

Terkadang peneliti memasukkan data kategori (misalnya, jenis kelamin atau tingkat pendidikan) sebagai data numerik tanpa mengubah skala pengukurannya menjadi nominal atau ordinal.

Kesalahan ini menyebabkan SPSS membaca data kategori tersebut sebagai data numerik, yang akan berdampak pada hasil analisis.

- a. **Dampak:** Hasil analisis menjadi tidak akurat karena SPSS memperlakukan data kategori sebagai numerik.
- b. **Solusi:** Untuk variabel kategori, pastikan memilih skala pengukuran sebagai *Nominal* atau *Ordinal* di SPSS. Jangan memasukkan data kategori sebagai variabel numerik.

### 5. Tidak Menggunakan Syntax untuk Dokumentasi Analisis

Banyak pengguna SPSS hanya menggunakan antarmuka menu (point-and-click) tanpa mencatat atau menyimpan syntax analisis yang telah dilakukan. Hal ini menyulitkan reproduksi analisis atau melihat kembali langkah-langkah yang dilakukan.

- a. **Dampak:** Kesulitan dalam mengulang analisis, terutama jika ada kesalahan yang perlu diperbaiki.
- b. **Solusi:** Gunakan *Syntax Editor* di SPSS untuk menyimpan semua langkah analisis. Syntax ini dapat disimpan dan digunakan kembali untuk analisis lanjutan atau referensi.

### 6. Tidak Mengubah Tipe Data yang Tidak Sesuai

Dalam beberapa kasus, data yang diimpor ke SPSS mungkin memiliki tipe data yang tidak sesuai (misalnya, data numerik yang terbaca sebagai string). Hal ini dapat menyebabkan kesalahan dalam perhitungan atau analisis lebih lanjut.

- a. **Dampak:** Data tidak bisa dianalisis dengan benar jika tipe data tidak sesuai.
- b. **Solusi:** Pastikan untuk memeriksa tipe data setiap variabel di *Variable View*. Ubah tipe data menjadi numerik atau string sesuai dengan kebutuhan analisis.

### 7. Mengabaikan Pembersihan dan Normalisasi Data

Pengguna seringkali langsung melakukan analisis tanpa membersihkan data terlebih dahulu. Data yang tidak konsisten,

duplikasi, atau kesalahan entri akan menghasilkan hasil yang tidak akurat atau bias.

- a. **Dampak:** Hasil analisis menjadi bias dan tidak mencerminkan data yang sebenarnya.
- b. **Solusi:** Lakukan pemeriksaan data awal, seperti mendeteksi outlier, mengidentifikasi duplikasi, dan membersihkan data dari kesalahan entri sebelum analisis.

## **8. Menggabungkan atau Membagi Data Tanpa Memperhatikan Variabel Penghubung**

Dalam penelitian yang menggunakan lebih dari satu set data, kesalahan umum terjadi saat menggabungkan data tanpa menggunakan variabel penghubung yang sesuai. Hal ini dapat menyebabkan data yang tidak cocok atau inkonsistensi dalam hasil analisis.

- a. **Dampak:** Data yang tidak sinkron dan hasil analisis yang tidak relevan.
- b. **Solusi:** Pastikan ada variabel kunci yang menghubungkan data saat menggabungkannya di SPSS, misalnya *ID Karyawan* atau *Nomor Subjek*.

## **9. Melewatkan Pemeriksaan Asumsi Statistik**

Beberapa teknik analisis memerlukan asumsi statistik tertentu, seperti normalitas data untuk uji parametrik. Kesalahan umum terjadi ketika pengguna tidak memeriksa asumsi ini sebelum melakukan analisis.

- a. **Dampak:** Hasil analisis bisa bias atau tidak valid jika asumsi statistik tidak dipenuhi.
- b. **Solusi:** Lakukan pemeriksaan asumsi, seperti uji normalitas, sebelum menerapkan teknik analisis statistik tertentu.

Mengelola variabel dan data dengan benar di SPSS adalah langkah penting dalam analisis statistik yang akurat. Dengan menghindari kesalahan-kesalahan umum ini, peneliti dapat

memastikan bahwa hasil analisis lebih valid dan interpretasi yang dihasilkan lebih andal.

## **2.10 Kunci Sukses Pengelolaan Variabel dan Data dalam SPSS**

Kesimpulannya, kunci sukses dalam pengelolaan variabel dan data di SPSS terletak pada pemahaman yang mendalam tentang jenis data, skala pengukuran, dan proses pengaturan yang sesuai di perangkat lunak. Dengan memilih skala pengukuran yang tepat, mengatur label dan nilai variabel dengan benar, serta melakukan pembersihan dan penanganan missing data secara menyeluruh, peneliti dapat memastikan kualitas analisis dan hasil yang valid (Rozak & Hidayati, 2019).

Memeriksa asumsi statistik dan mendokumentasikan langkah-langkah analisis melalui syntax juga merupakan langkah penting untuk menjaga transparansi dan keterulangan proses. Menghindari kesalahan umum, seperti salah memilih tipe data atau melewatkan pengaturan label, membantu mencegah interpretasi yang keliru. Dengan memperhatikan faktor-faktor ini, SPSS menjadi alat yang lebih efektif untuk menghasilkan analisis data kuantitatif yang akurat, relevan, dan dapat diandalkan.

## DAFTAR PUSTAKA

- Chalimi, A. N. F., & SE, M. A. (2021). *Aplikom statistik berbasis SPSS*. Lembaga Chakra Brahmana Lentera.
- Herlina, V. (2019). *Panduan praktis mengolah data kuesioner menggunakan SPSS*. Elex Media Komputindo.
- Hidayat, A. A. (2021). *Cara praktis uji statistik dengan SPSS*. Health Books Publishing.
- Ikram, F. Z., & others. (2024). Kesalahan Mahasiswa Fakultas Keguruan dan Ilmu Pendidikan dalam Menggunakan SPSS. *Jurnal MediaTIK*, 144–149.
- Jaya, A., Wartti, R., & others. (2022). *Statistik pendidikan: Teori dan aplikasi SPSS*. Penerbit NEM.
- Kusumawardhani, R., & Astuti, N. D. (2024). *PENGANTAR STATISTIKA DAN ANALISIS DENGAN SPSS*. CV. Ruang Tentor.
- Muzawi, R., Kom, M., & others. (2024). *Fundamental SPSS Dalam Pengolahan Data*. Serasi Media Teknologi.
- Nugroho, A. S., & Haritanto, W. (2022). *Metode penelitian kuantitatif dengan pendekatan statistika:(Teori, Implementasi & Praktik dengan SPSS)*. Penerbit Andi.
- Priyastama, R. (2020). *The Book of SPSS: Pengolahan \& Analisis Data*. Anak Hebat Indonesia.
- Qudratullah, M. F. (2013). *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS*. Penerbit Andi.
- Ramadhani, R., Bina, N. S., & others. (2021). *Statistika Penelitian Pendidikan: Analisis Perhitungan Matematis dan Aplikasi SPSS*. Prenada Media.
- Ramadhayanti, A. (2019). *Aplikasi SPSS untuk penelitian dan riset pasar*. Elex media komputindo.
- Rozak, A., & Hidayati, W. S. (2019). *Pengolahan data dengan SPSS*. Erhaka Utama.
- Santoso, S. (2016). *Panduan lengkap SPSS versi 23*. Elex Media Komputindo.
- Santoso, S. (2019). *Menguasai SPSS versi 25*. Elex Media Komputindo.
- Sitoayu, L., Nuzrina, R., & Rumana, N. A. (2020). *Aplikasi SPSS untuk analisis data kesehatan: Bonus analisis data dengan SEM*.

Penerbit NEM.

SPSS, M. O. D. (2015). *SPSS (Statistical Package for the Social Sciens)*.

Sujarweni, V. W., & Utami, L. R. (2019). *The master book of SPSS*. Anak Hebat Indonesia.

Wahyuni, M. (2020). *Statistik Deskriptif Untuk Penelitian Oleh Data Manual dan SPSS Versi 25*. Bintang Pustaka Madani.

Wibowo, F. X. P. (2023). *Pengolahan dan Analisa Data Statistika dengan SPSS*. Penerbit Andi.

# **TRANSFORMASI DATA DALAM SPSS**

**Oleh: Suryaningsih, S.E., M.M.**

## **3.1 Pengertian Transformasi Data**

Transformasi data dalam SPSS merujuk pada proses perubahan atau modifikasi data yang ada dalam dataset untuk mempersiapkan data tersebut agar lebih siap dianalisis atau untuk mengubahnya ke dalam format yang lebih sesuai dengan kebutuhan analisis tertentu. Proses ini dapat melibatkan berbagai teknik, seperti penambahan atau penghapusan variabel, penggabungan kategori, pembuatan variabel baru berdasarkan perhitungan atau kondisi tertentu, serta normalisasi atau standarisasi data. Dalam SPSS, transformasi data dilakukan menggunakan berbagai perintah seperti *Compute Variable*, *Recode*, dan *Visual Binning* untuk memodifikasi data. Selain itu, transformasi juga dapat mencakup perubahan data numerik menjadi data kategorikal, atau sebaliknya, serta penggabungan variabel yang dapat mengurangi kompleksitas dataset.

Transformasi data bertujuan untuk memastikan kualitas data yang optimal untuk analisis lebih lanjut dan meningkatkan interpretasi hasil analisis. Misalnya, transformasi dilakukan untuk mengatasi masalah distribusi data yang tidak normal atau untuk menciptakan

variabel yang lebih relevan dengan hipotesis penelitian yang diajukan. Hal ini penting dalam konteks analisis statistik, karena data yang tidak sesuai atau tidak terstandarisasi dapat menghasilkan hasil yang bias atau tidak valid.

Menurut Pallant (2020), transformasi data juga membantu dalam menyederhanakan proses analisis statistik yang lebih kompleks seperti regresi, analisis faktor, dan model-model statistik lainnya, yang sering memerlukan data dalam format tertentu untuk mencapai validitas yang optimal. Dengan demikian, transformasi data di SPSS bukan hanya sekadar modifikasi, tetapi juga bagian integral dari proses persiapan data yang lebih luas dalam penelitian kuantitatif.

### **3.1 Fungsi Transformasi Data**

Fungsi transformasi data dalam SPSS sangat penting untuk memastikan bahwa data yang digunakan dalam analisis statistik memenuhi persyaratan dan kriteria yang diperlukan untuk menghasilkan hasil yang valid dan dapat diandalkan. Proses ini melibatkan perubahan data mentah yang ada menjadi bentuk yang lebih sesuai dengan tujuan analisis. Salah satu fungsi utama dari transformasi data adalah untuk menangani masalah data yang tidak terstruktur atau tidak sesuai, seperti mengatasi nilai yang hilang, mendefinisikan kembali kategori data, atau mengubah data numerik menjadi data kategorikal, serta sebaliknya. Transformasi data ini memastikan bahwa data yang digunakan lebih tepat sasaran dan siap untuk digunakan dalam teknik analisis statistik yang lebih kompleks (Pallant, 2020).

Selain itu, transformasi data juga berfungsi untuk mempermudah interpretasi dan analisis. Misalnya, dalam kasus regresi linier, transformasi variabel dilakukan untuk memenuhi asumsi-asumsi dasar analisis seperti linearitas, normalitas, dan homogenitas varians. Teknik transformasi seperti standarisasi atau

normalisasi data juga dapat diterapkan untuk membuat perbandingan antarvariabel yang lebih adil, terutama ketika data memiliki satuan pengukuran yang berbeda. Selain itu, transformasi juga digunakan untuk membuat variabel baru berdasarkan rumus atau kondisi tertentu, yang dapat membantu dalam menguji hipotesis yang lebih kompleks atau dalam memperjelas hubungan antarvariabel (Field, 2018).

Fungsi transformasi data dalam SPSS tidak hanya terfokus pada persiapan data untuk analisis, tetapi juga pada peningkatan keakuratan dan efektivitas hasil yang diperoleh. Dengan melakukan transformasi yang tepat, peneliti dapat meningkatkan kualitas data dan mengurangi potensi kesalahan atau bias yang dapat muncul akibat ketidaksesuaian data dengan metode analisis yang digunakan. Transformasi data juga memungkinkan peneliti untuk mengoptimalkan penggunaan data yang ada agar dapat memaksimalkan kekuatan statistik dalam pengujian hipotesis dan model prediksi.

### **3.2 Jenis-Jenis Transformasi Data Dalam SPSS**

Dalam SPSS, transformasi data digunakan untuk memodifikasi data agar sesuai dengan tujuan analisis statistik. Beberapa jenis transformasi data yang umum dilakukan antara lain:

#### **1. Recode Data (Pengkodean Ulang)**

Pengkodean ulang adalah transformasi data yang digunakan untuk mengubah nilai variabel menjadi nilai yang berbeda. Hal ini sering dilakukan untuk mengubah data kategorikal menjadi bentuk yang lebih mudah dianalisis atau untuk menggabungkan beberapa kategori menjadi satu kategori baru. Misalnya, dalam sebuah survei, kategori usia dapat dikelompokkan menjadi beberapa rentang usia seperti "18-25", "26-35", dan seterusnya. SPSS menyediakan dua cara

untuk melakukan recoding, yaitu Recode Into Same Variables dan Recode Into Different Variables.

## 2. Compute Variable (Perhitungan Variabel Baru)

Fungsi ini memungkinkan peneliti untuk membuat variabel baru berdasarkan operasi matematika atau fungsi lainnya yang melibatkan variabel yang sudah ada. Misalnya, menghitung total skor dari beberapa item survei, atau membuat variabel yang mencerminkan perbedaan antara dua variabel. Proses ini dilakukan dengan menggunakan menu Transform > Compute Variable, dan dapat melibatkan operasi aritmatika (penjumlahan, pengurangan), statistik (mean, standar deviasi), atau logika kondisional (misalnya, jika nilai A lebih besar dari nilai B, maka variabel baru bernilai 1).

## 3. Visual Binning (Pengelompokan dengan Pemotongan Visual)

Visual Binning digunakan untuk mengubah variabel numerik menjadi kategorikal dengan cara membagi rentang nilai variabel tersebut menjadi beberapa kelompok. Misalnya, mengubah usia menjadi kategori seperti "Muda", "Tua", dan "Lansia". Pengguna dapat menentukan batas kategori secara manual atau berdasarkan distribusi data. Teknik ini sering digunakan ketika variabel numerik memiliki distribusi yang sangat lebar atau tidak normal.

## 4. Standardize Data (Standarisasi Data)

Transformasi ini digunakan untuk mengubah data ke dalam skala yang lebih konsisten, biasanya untuk memenuhi asumsi model statistik yang mengharuskan data dalam skala yang seragam. Salah satu teknik yang umum adalah standarisasi (z-score), yang mengubah setiap nilai dalam variabel menjadi deviasi standar dari rata-rata. Hal ini sangat berguna ketika variabel memiliki satuan pengukuran yang berbeda atau jika kita ingin membuat perbandingan antarvariabel yang tidak setara.

## 5. Normalisasi Data

Normalisasi data dilakukan untuk mengubah skala variabel agar berada dalam rentang tertentu, biasanya antara 0 dan 1. Teknik ini sering digunakan ketika data memiliki rentang nilai yang sangat besar atau variabel memiliki pengaruh yang sangat besar dalam analisis. Ini memungkinkan model untuk memperlakukan semua variabel dengan bobot yang sama.

## 6. Transformasi Logaritma atau Kuadrat

Salah satu jenis transformasi data adalah menggunakan logaritma atau akar kuadrat untuk mengatasi masalah distribusi data yang tidak normal. Misalnya, data dengan distribusi miring kanan (positively skewed) dapat diubah dengan mengambil logaritma untuk membuat distribusi data lebih mendekati normal. Transformasi ini dapat membantu dalam analisis regresi atau analisis statistik lainnya yang mengharuskan data normal.

## 7. Handling Missing Values (Penanganan Nilai yang Hilang)

Transformasi data juga melibatkan penanganan nilai yang hilang (missing values). SPSS menyediakan beberapa metode untuk menangani nilai hilang, seperti imputasi (mengganti nilai hilang dengan nilai yang dihitung berdasarkan statistik lainnya, seperti mean atau median), atau penghapusan observasi yang memiliki nilai hilang.

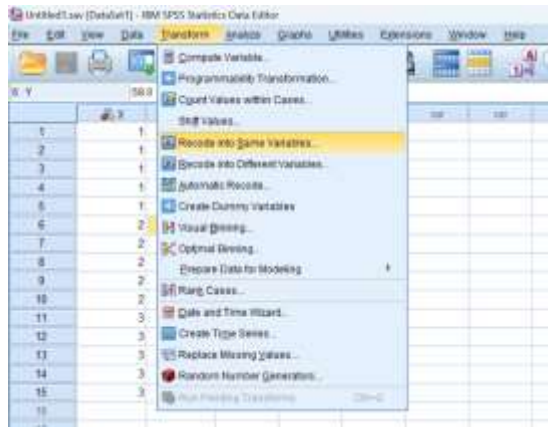
## 8. Categorical to Continuous (Mengubah Kategorikal Menjadi Kontinu)

Pada beberapa kasus, data kategorikal (misalnya, jenis kelamin, status perkawinan) dapat diubah menjadi data kontinu. Ini dilakukan melalui proses pengkodean numerik, di mana setiap kategori diberikan nilai numerik (misalnya, 0 untuk "Laki-laki" dan 1 untuk "Perempuan"), yang memungkinkan penggunaan teknik analisis statistik yang memerlukan variabel kontinu.

### 3.3 Beberapa Langkah Umum yang digunakan dalam SPSS Untuk Transformasi Data

#### 1. Recode into Same Variables atau Different Variables

- **Tujuan:** Mengubah nilai variabel, misalnya mengelompokkan nilai kontinu menjadi kategori.
- **Langkah-langkah:**
  1. Pilih menu **Transform > Recode into Same Variables** atau **Recode into Different Variables**.
  2. Pilih variabel yang akan diubah.
  3. Tentukan aturan perubahan (contoh: nilai 1 menjadi "Rendah", nilai 2 menjadi "Sedang").
  4. Untuk **Recode into Different Variables**, beri nama variabel baru untuk menyimpan hasil transformasi.

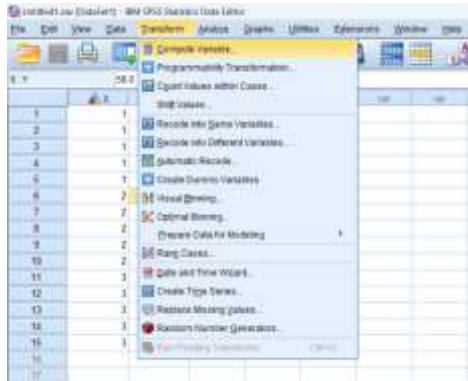


Gambar 1. Recode into Same Variables

#### 2. Compute Variable

- **Tujuan:** Membuat variabel baru berdasarkan perhitungan dari variabel yang ada.
- **Langkah-langkah:**
  1. Pilih menu **Transform > Compute Variable**.

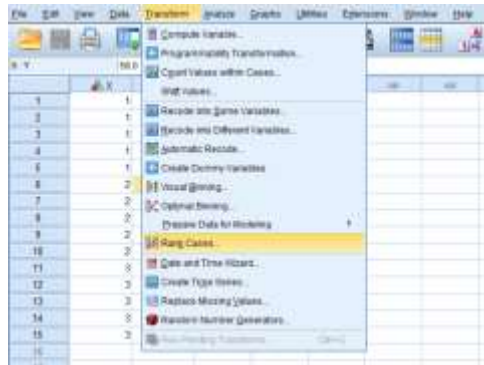
2. Masukkan nama variabel baru pada kotak *Target Variable*.
3. Tentukan ekspresi perhitungan, seperti penjumlahan (misal, VAR1 + VAR2) atau rata-rata (misal, (VAR1 + VAR2) / 2).
4. Klik **OK** untuk menyimpan variabel baru.



Gambar 2. Compute Variables

### 3. Rank Cases

- **Tujuan:** Memberikan peringkat pada data berdasarkan nilai suatu variabel.
- **Langkah-langkah:**
  1. Pilih menu **Transform > Rank Cases**.
  2. Pilih variabel yang ingin diranking.
  3. Atur opsi tambahan, seperti menentukan apakah peringkat akan menaik (ascending) atau menurun (descending).



Gambar 3. Rank Cases

#### 4. Log Transformation

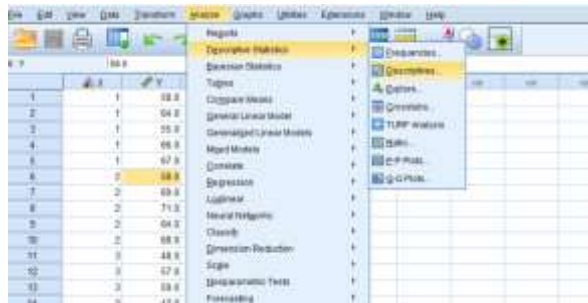
- **Tujuan:** Mengubah data ke skala logaritmik untuk normalisasi atau mengurangi heteroskedastisitas.
- **Langkah-langkah:**
  1. Gunakan menu **Transform > Compute Variable**.
  2. Pada kotak *Numeric Expression*, masukkan ekspresi logaritma (misal: LG10(VARIABLE) untuk logaritma basis 10 atau LN(VARIABLE) untuk logaritma natural).



Gambar 4. Numeric Expression

## 5. Standardize Variables (Z-Score)

- **Tujuan:** Mengubah data ke dalam bentuk skor baku (z-score) untuk analisis statistik.
- **Langkah-langkah:**
  1. Pilih menu **Analyze > Descriptive Statistics > Descriptives**.
  2. Centang opsi **Save standardized values as variables**.
  3. Klik **OK** untuk menyimpan variabel z-score.



Gambar 5. Descriptive

## 6. Split File

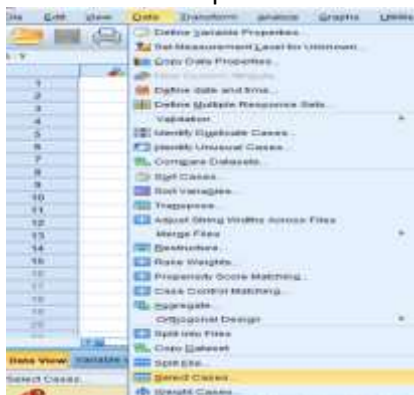
- **Tujuan:** Memisahkan dataset berdasarkan kategori tertentu untuk analisis terpisah.
- **Langkah-langkah:**
  1. Pilih menu **Data > Split File**.
  2. Pilih opsi **Compare groups** atau **Organize output by groups**.
  3. Pilih variabel pembagi.



Gambar 6. Split File

## 7. Select Cases

- **Tujuan:** Menyaring kasus tertentu untuk dianalisis.
- **Langkah-langkah:**
  1. Pilih menu **Data > Select Cases**.
  2. Tentukan kondisi penyaringan menggunakan ekspresi logika (misalnya,  $AGE > 30$ ).
  3. Klik **OK** untuk menerapkan filter.



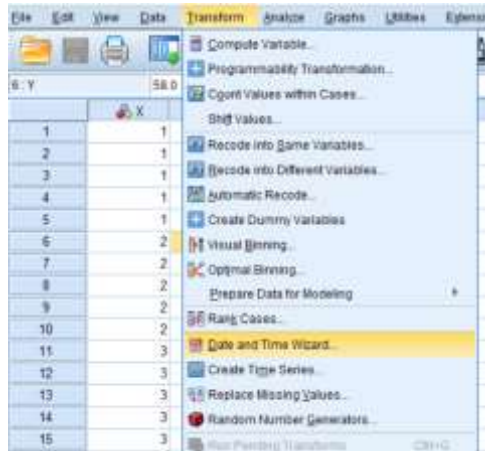
Gambar 7. Select Cases

## 8. Date and Time Transformations

- **Tujuan:** Mengubah format tanggal dan waktu menjadi variabel yang dapat dihitung.

- **Langkah-langkah:**

1. Gunakan menu **Transform > Date and Time Wizard**.
2. Pilih jenis transformasi, seperti menghitung durasi waktu atau mengekstrak bagian tertentu dari tanggal (tahun, bulan, dll.).



Gambar 8. Time Wizard

## 9. Visual Binning

- **Tujuan:** Mengelompokkan data kontinu menjadi kategori secara visual.
- **Langkah-langkah:**
  1. Pilih menu **Transform > Visual Binning**.
  2. Pilih variabel yang akan dikelompokkan.
  3. Tentukan batas kategori secara manual atau otomatis.



Gambar 9. Visual Binning

## DAFTAR PUSTAKA

- Pallant, J. (2020). SPSS Survival Manual: A Step by Step Guide to Data Analysis using IBM SPSS. 7th Edition. McGraw-Hill Education.
- Field, A. (2018). Discovering Statistics Using SPSS. 5th Edition. SAGE Publications Ltd.

# ANALISIS DESKRIPTIF

*Oleh Sri Maryati, SS.T., M.Kes.*

## 4.1 Pendahuluan

Statistika deskriptif adalah jenis analisis statistik yang digunakan untuk mendeskripsikan data yang telah dikumpulkan tanpa membuat kesimpulan umum tentang karakteristik serangkaian data. (Ghozali, 2016). Statistika deskriptif memungkinkan pengumpulan, peringkasan, dan penyajian data untuk membuat informasi lebih mudah dipahami (Muchson, 2017). Statistik deskriptif adalah statistic yang berfungsi untuk mendeskripsikan atau memberi gambaran terhadap obyek yang diteliti melalui data sampel atau populasi sebagaimana adanya, tanpa melakukan analisis dan membuat kesimpulan yang berlaku untuk umum.

Analisis deskriptif terbagi menjadi dua yaitu analisis data deskriptif kuantitatif dan kualitatif. Data kuantitatif dapat digambarkan, dirangkum, dan dianalisis melalui analisis statistik yang dikenal sebagai analisis deskriptif kuantitatif. Data kuantitatif termasuk data yang dapat diukur atau dihitung menggunakan angka seperti usia, berat badan, dan tinggi badan.

Analisis data deskriptif kuantitatif mengacu pada gambaran statistik yang membantu memahami detail data dengan meringkas dan menemukan pola dalam sampel data tertentu. Metode ini

membantu menggambarkan, menunjukkan, atau meringkas data secara konstruktif. Peneliti mengumpulkan angka absolut melalui sampel; namun, tidak selalu menjelaskan alasan di balik angka-angka tersebut. Oleh karena itu, untuk analisis tambahan, pendekatan inferensial diperlukan. Peneliti dapat membuat kesimpulan yang lebih akurat tentang fenomena yang mereka amati dengan memahami karakteristik dasar dari data dengan menggunakan statistika deskriptif kuantitatif.

Analisis statistika deskriptif kuantitatif sangat berguna dalam berbagai bidang seperti ilmu sosial, ekonomi, dan sains, di mana data kuantitatif seringkali merupakan bagian penting dari analisis dan pengambilan keputusan. Tujuan dari analisis ini adalah untuk memberikan gambaran yang jelas dan terperinci tentang data yang telah dikumpulkan, sehingga dapat memudahkan interpretasi dan pengambilan keputusan yang didasarkan pada data yang ada.

Berikut ini akan dijelaskan mengenai teknik analisis deskripsi berdasarkan skala data pada variabel penelitian:

## **4.2 Teknik Analisis Deskriptif Kuantitatif**

Beberapa teknik analisis deskriptif kuantitatif antara lain

### a. Pengukuran Pemusatan Data yaitu

#### 1) Mean (rata-rata)

Nilai mean adalah jumlah dari semua nilai dalam satu set data dibagi dengan jumlah total nilai. Dalam pembelajaran statistik tingkat dasar, rata-rata hitung selalu menjadi topik utama karena merupakan ukuran pemusatan yang sering digunakan. Saat data yang diamati mengandung data yang sangat ekstrim, rata-rata tidak dapat menggambarkan karakteristik data dengan baik. Misalnya, meskipun nilai rata-rata siswa di satu kelas adalah 70, rata-rata tersebut tidak sepenuhnya menggambarkan keadaan sebenarnya karena ada

siswa yang mendapatkan nilai 25 sementara yang lain mendapatkan nilai 100.

2) Median (nilai tengah)

Nilai tengah sebuah set data, yang diurutkan dari yang terkecil ke yang terbesar, disebut median. Ini memberikan gambaran tentang nilai tengah set data yang tidak dipengaruhi oleh nilai ekstrem. Selain itu, jika data diurutkan dari yang terkecil ke yang terbesar, median juga dapat diperoleh dengan membagi distribusi frekwensi menjadi dua sama besar.

3) Modus

Modus dikenal sebagai nilai yang sering muncul atau data yang paling tinggi frekuensinya. Mendapatkan metode ini tidak selalu mudah. Hal ini akan terjadi jika dihadapkan pada suatu nilai yang mempunyai frekuensi kemunculan yang sama dengan yang lainnya. Dalam suatu data yang disajikan, data dapat terjadi dengan beberapa modus (multi-modus), dan dapat juga terjadi data tanpa modus.

Contoh: Berat 5 orang bayi : 3.6 3.5 2.9 3.1 3.0

Data tanpa Modus

Usia Mahasiswa : 19 18 19 18 23 21 19 21 18 20 22 17

Modus : 18 dan 19

b. Pengukuran Persebaran Data yaitu:

1) Range

Range adalah selisih antara nilai maksimum dan minimum dalam satu set data. Hal ini menggambarkan sejauh mana data tersebar.

2) Variance (variansi)

Variansi merupakan pengukuran persebaran data yang mengukur sejauh mana data tersebar dari nilai rata-rata. Hal ini dapat dihitung dengan cara menghitung rata-rata selisih kuadrat dari setiap nilai dengan nilai rata-rata.

### 3) Standard deviation (deviasi standar)

Standar deviasi adalah akar kuadrat dari varian. Ini digunakan untuk mengukur sejauh mana data tersebar dari nilai rata-rata.

### 4) Kuartil

Kuartil adalah nilai yang membagi data menjadi empat bagian yang sama besar. Kuartil terbagi menjadi tiga yaitu kuartil pertama (Q1) adalah nilai tengah antara nilai minimum dan median. kuartil kedua (Q2) adalah nilai median, dan kuartil ketiga (Q3) adalah nilai tengah antara median dan nilai maksimum. Ini menggambarkan jumlah berapa banyak data yang berada di setiap bagian distribusi.

### 5) Desil

Desil adalah salah satu ukuran kuantitatif dalam statistik deskriptif yang mengukur posisi suatu data dalam urutan atau rangkaian data yang berurutan. Desil membagi data menjadi sepuluh bagian yang sama besar, di mana setiap bagian tersebut disebut desil. Dengan kata lain, desil adalah nilai yang membagi data menjadi sepuluh bagian yang sama besar.

Desil sering digunakan saat menganalisis data ekonomi dan keuangan, di mana data dapat memiliki banyak outlier atau nilai ekstrem. Dalam kasus seperti itu, penggunaan desil dapat membantu untuk mengurangi efek dari nilai ekstrem tersebut dan memberikan gambaran yang lebih akurat tentang data. Desil yang paling umum digunakan adalah desil ke-4 (D4) atau sering dikenal sebagai kuartil ke-2 (Q2) atau median, yang membagi data menjadi dua bagian sama besar. Desil ke-1 (D1) adalah nilai yang membagi 10% data terkecil, sementara desil ke-10 (D10) adalah nilai yang membagi 10% data terbesar. Desil lainnya, seperti desil ke-2 (D2) dan desil ke-3 (D3), dapat membagi data ke dalam 20% dan 30% bagian terkecil. Begitu

juga dengan desil ke-6 (D6) dan desil ke-7 (D7) membagi data ke dalam 60% dan 70% bagian terbesar. Desil juga bisa digunakan bersama dengan ukuran pemusatan data, seperti mean dan median, untuk memberikan gambaran yang menyeluruh tentang data.

#### 6) Persentil

Persentil adalah salah satu ukuran kuantitatif dalam statistik deskriptif yang digunakan untuk mengukur letak data dalam rangkaian data terurut. Persentil membagi data menjadi seratus bagian yang sama besar, di mana setiap bagian tersebut disebut persentil. Persentil juga dapat diartikan sebagai nilai yang membagi data ke dalam seratus bagian yang sama besar. Persentil sering digunakan untuk menganalisis data dan penelitian agar dapat memberikan gambaran tentang distribusi data, terutama pada data maksimum. Persentil yang paling sering digunakan adalah persentil ke-25 (P25), yang sering dikenal sebagai kuartil pertama (Q1), persentil ke-50 (P50) atau median, dan persentil ke-75 (P75), yang juga dikenal sebagai kuartil ketiga (Q3).

#### c. Pengukuran Kemiringan Data yaitu:

##### 1) Skewness (kemiringan)

Pengukuran kemiringan data dalam statistik deskriptif mengacu pada ukuran yang menggambarkan kecondongan atau asimetri dari distribusi data. Ada tiga jenis kemiringan data yang sering dikenal, yaitu simetri, kemiringan positif, dan kemiringan negatif.

##### 2) Kurtosis (keruncingan).

Keruncingan data atau dapat juga disebut kurtosis adalah ukuran statistik yang menggambarkan tingkat kecuraman dari distribusi data. Konsep ini menggambarkan bagaimana puncak

atau konsentrasi data terpusat pada nilai tengah dari data yang dianalisis.

d. Distribusi Frekuensi

Distribusi frekuensi adalah daftar nilai data (bisa nilai individual atau nilai data yang sudah dikelompokkan ke dalam selang interval tertentu) yang disertai dengan nilai frekuensi yang sesuai. Atau menurut definisi, distribusi frekuensi adalah penyusunan data dalam bentuk kelompok mulai dari yang terkecil sampai yang terbesar berdasarkan kelas-kelas interval dan kategori tertentu.\

### 4.3 Analisis deskriptif dengan SPSS

#### 1. Analisis Deskriptif Variabel Numerik dengan SPSS

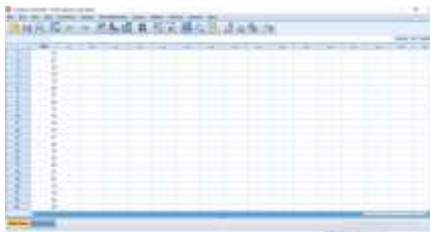
Berikut ini merupakan langkah membuat deskripsi variabel dengan skala pengukuran numerik:

**Kasus:**

Seorang peneliti sudah melakukan penelitian terkait karakteristik responden berupa indeks masa tubuh. Data yang sudah dikumpulkan kemudian dimasukkan ke dalam SPSS. Indeks massa tubuh adalah variabel numerik, maka deskripsi variabel yang dibuat adalah ukuran pemusatan dan ukuran penyebarannya serta penyajiannya dalam bentuk histogram.

Langkah-langkah yang dilakukan dengan program SPSS adalah sebagai berikut:

1. Buka software SPSS.
2. Lakukan input data ke dalam spss



3. Lakukan proses sebagai berikut.

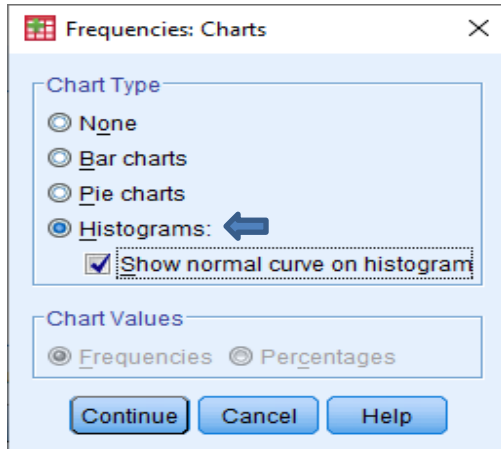
- ✓ *Analyze*→*Descriptive statistics*→*Frequencies*.
  - ✓ Masukkan Indeks Massa Tubuh ke dalam kotak *Variables*.
  - ✓ Pilihan *Display Frequency Tables* dinonaktifkan.
- Akan terlihat tampilan sebagai berikut.



- ✓ Klik kotak *Statistic*. Pilih *Mean*, *Median*, *Modus* pada *Central Tendency* (sebagai ukuran pemusatan), pilih *Std deviation*, *Variance*, *Minimum*, *Maksimum* pada *Dispersion*, pilih *Skewness* dan *Kurtosis* pada *Distribution* (sebagai ukuran penyebaran).
- Akan terlihat tampilan sebagai berikut.



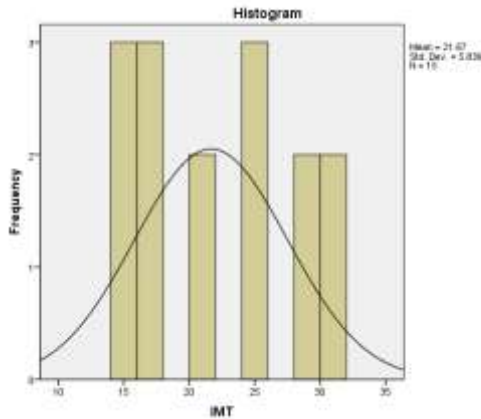
- ✓ Klik *Continue*, lalu aktifkan pilihan *Chart* pilih *Histogram* pada *Chart Type* dan aktifkan kotak *With normal curve*. Akan terlihat tampilan sebagai berikut.



- ✓ Proses telah selesai, klik *Continue*, klik *OK*.

Berikut ini merupakan Ouput SPSS yang didapatkan pada kasus di atas:

Statistics	
Jenis Measur Tubuh	
N	25
Missing	0
Mean	20.44
Median	20.00
Mode	15
Std. Deviation	2.181
Variance	4.757
Skewness	1.241
Std. Error of Skewness	.494
Kurtosis	1.594
Std. Error of Kurtosis	.807
Minimum	10
Maximum	25



Interpretasi Output SPSS di atas dapat dijelaskan sebagai berikut:

Metode ini dapat digunakan untuk memberikan penjelasan tentang variabel indeks massa tubuh. Parameter ukuran pemusatan, seperti mean, median, dan modus, serta ukuran penyebaran, seperti standar deviasi, variance, minimum-maksimum, skewness, dan kurtosis, dapat ditemukan di bagian statistik.

Distribusi data dengan kurva normal dapat digambarkan pada grafik histogram. Untuk menentukan apakah distribusi data normal, maka peneliti harus memahami karakteristik variabel numerik ini.

## 2. Membuat Deskripsi Variabel Kategorik

Berikut ini merupakan langkah-langkah membuat tabel frekuensi dan grafik:

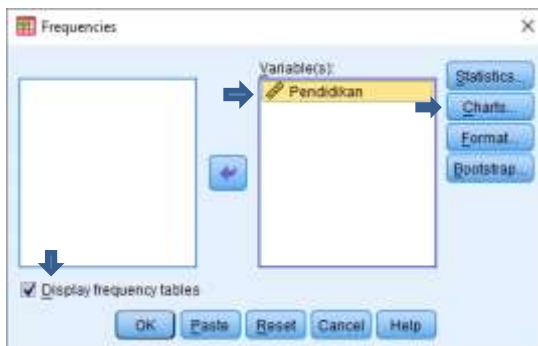
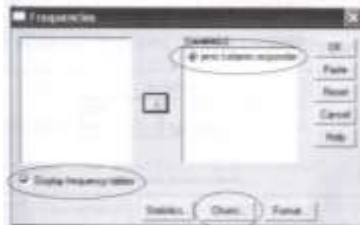
### Kasus:

Penelitian telah dilakukan, kemudian data yang dikumpulkan dimasukkan ke dalam SPSS. Salah satu variabel yang diteliti adalah variabel pendidikan. Pada penelitian ditentukan bahwa skala pengukuran variabel pendidikan adalah nominal. Berikut ini akan ditampilkan distribusi jenis kelamin dalam bentuk tabel dan grafik dengan menggunakan SPSS:

1. Masukkan data ke dalam software SPSS
2. Lakukan input data seperti gambar di bawah ini :

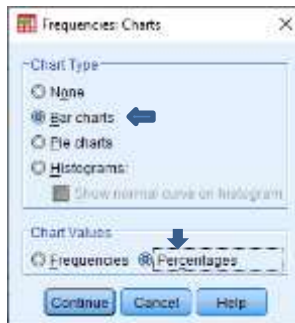


3. Lakukan proses sebagai berikut.
  - ✓ Analyze→Descriptive statistics→Frequencies.
  - ✓ Masukkan variabel pendidikan ke dalam kotak Variabel (s)
  - ✓ Aktifkan *Displays Frequency Tables*.Akan terlihat tampilan sebagai berikut.



- ✓ Klik kotak Charts.
- ✓ Pilih *bar* pada *Chart Type*.
- ✓ Pilih *Percentages* pada *Chart Values*.

Akan terlihat tampilan sebagai berikut:



- ✓ Klik *Continue*. Proses selesai, klik *OK*.

Berikut merupakan hasil output SPSS yang didapatkan antara lain:

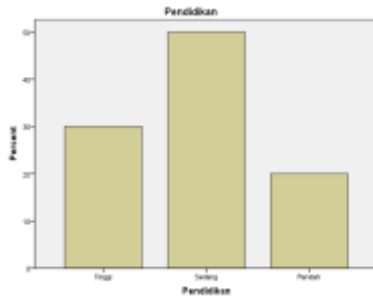
Statistics

Pendidikan		
N	Valid	50
	Missing	0

Pendidikan

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Tinggi	15	30.0	30.0	30.0
	Sedang	25	50.0	50.0	80.0
	Rendah	10	20.0	20.0	100.0
	Total	50	100.0	100.0	



Interpretasi Output SPSS di atas dapat dijelaskan sebagai berikut: Dengan prosedur ini, peneliti dapat mengenali karakteristik pendidikan. Dari output pertama, dapat diketahui jumlah total subjek sebanyak 50 orang dan tidak ada data yang *missing*. Distribusi pendidikan berdasarkan jumlah dan persentase dapat diketahui pada output kedua. Visualisasi dalam bentuk grafik disajikan pada output ketiga.

### 3. Perubahan Data dari Satu Skala ke Skala Lain dengan SPSS

Dalam membuat deskripsi variabel kategori dengan menggunakan software SPSS perlu dilakukan perubahan data. Perubahan data dari satu skala ke skala lain diperlukan dalam perhitungan distribusi frekuensi dengan software SPSS, karena pada praktiknya nanti, data awal yang dimasukkan ke dalam SPSS harus diubah untuk keperluan tertentu. Perubahan data juga biasanya dilakukan sebelum seorang peneliti menghitung distribusi frekuensi untuk data yang sebelumnya mempunyai skala data numerik.

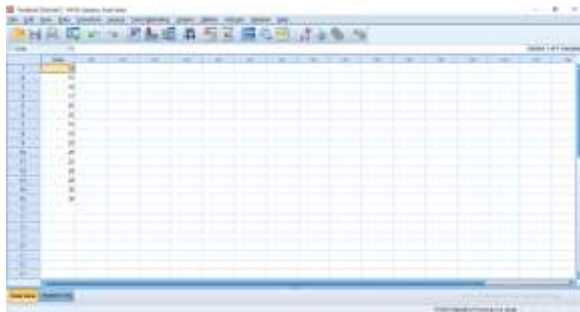
#### Kasus:

Peneliti akan melakukan penelitian dengan mengategorikan umur responden menjadi tiga kelompok, yaitu responden yang berumur <20 tahun, 20-22 tahun dan 22 tahun (dalam hal ini, Anda merubah

variabel numerik menjadi variabel ordinal).

Berikut merupakan langkah-langkah yang dilakukan dengan program SPSS:

1. Buka software SPSS
2. Lakukan input data seperti gambar di bawah ini:



3. Lakukan Langkah-langkah berikut ini:
  - ✓ *Transform*→*Recode*→*Recode into Different Variables*.
  - ✓ Masukkan variabel usia ke dalam *Input Variable*.
  - ✓ Ketik usia1 ke dalam *Output Variable*.
  - ✓ Ketik 'klasifikasi usia' ke dalam label.
  - ✓ Klik kotak *Change*. Setelah proses ini, Anda akan melihat tampilan sebagai berikut.



- ✓ Klik Old and New Values.
- ✓ Isilah kotak Old Values dan New Value (selanjutnya ikuti logika berpikir).

Logikanya adalah:

Semua data <19 tahun diubah menjadi kode 1.

Semua data 20-35 tahun diubah menjadi kode 2.

Semua data >35 tahun diubah menjadi kode 3.

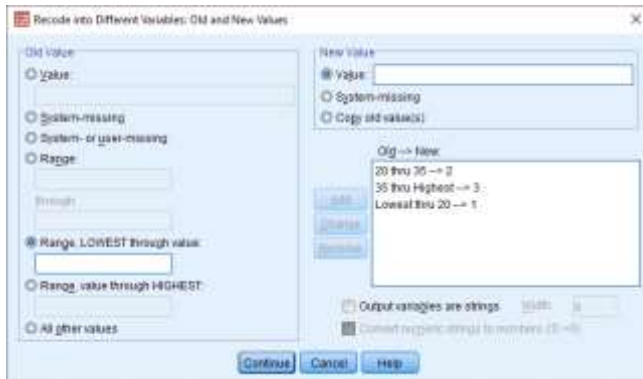
Dengan logika tersebut, isilah *Old Value* dan *New Value* sebagai berikut:

*Old Value: range lowest through 19, New Value: 1, klik Add.*

*Old Value: range 20 through 35, New Value: 2, klik Add.*

*Old Value: 35 through highest, New Value: 3, klik Add.*

Sampai tahap ini, Anda akan memperoleh tampilan sebagai berikut.



- ✓ Proses selesai, klik kotak *Continue*.
- ✓ Klik OK dan lihatlah hasilnya.

Apabila langkah di atas sudah selesai, maka akan muncul kolom baru (dengan nama usia1) dan nilai-nilai yang ada pada kolom tersebut, kemudian pada *Data View* akan terdapat kolom baru dengan variabel baru Bernama usia1.

	usia	usia1
1	1	1.00
2	1	1.00
3	1	1.00
4	1	1.00
5	1	1.00
6	1	1.00
7	1	1.00
8	1	1.00
9	1	1.00
10	1	1.00
11	1	1.00
12	1	1.00
13	1	1.00
14	1	1.00
15	1	1.00
16	1	1.00
17	1	1.00
18	1	1.00
19	1	1.00
20	1	1.00
21	1	1.00
22	1	1.00
23	1	1.00
24	1	1.00
25	1	1.00
26	1	1.00
27	1	1.00
28	1	1.00
29	1	1.00
30	1	1.00
31	1	1.00
32	1	1.00
33	1	1.00
34	1	1.00
35	1	1.00
36	1	1.00
37	1	1.00
38	1	1.00
39	1	1.00
40	1	1.00
41	1	1.00
42	1	1.00
43	1	1.00
44	1	1.00
45	1	1.00
46	1	1.00
47	1	1.00
48	1	1.00
49	1	1.00
50	1	1.00
51	1	1.00
52	1	1.00
53	1	1.00
54	1	1.00
55	1	1.00
56	1	1.00
57	1	1.00
58	1	1.00
59	1	1.00
60	1	1.00
61	1	1.00
62	1	1.00
63	1	1.00
64	1	1.00
65	1	1.00
66	1	1.00
67	1	1.00
68	1	1.00
69	1	1.00
70	1	1.00
71	1	1.00
72	1	1.00
73	1	1.00
74	1	1.00
75	1	1.00
76	1	1.00
77	1	1.00
78	1	1.00
79	1	1.00
80	1	1.00
81	1	1.00
82	1	1.00
83	1	1.00
84	1	1.00
85	1	1.00
86	1	1.00
87	1	1.00
88	1	1.00
89	1	1.00
90	1	1.00
91	1	1.00
92	1	1.00
93	1	1.00
94	1	1.00
95	1	1.00
96	1	1.00
97	1	1.00
98	1	1.00
99	1	1.00
100	1	1.00

Selanjutnya, lakukanlah pengisian value sesuai dengan kategori yang ditentukan pada Variable View untuk variabel usia1. Bila pada Value Anda mengisi sebagai berikut: kode 1 untuk usia beresiko; kode 2 untuk usia reproduksi; kode 3 untuk usia beresiko seperti pada gambar di bawah ini:



Apabila value sudah ditentukan maka data akan dapat dianalisis.

**4.4 Penyajian Data Statistik Deskriptif**

Berikut ini merupakan penyajian data statistic deskriptif berdasarkan jenis variabel:

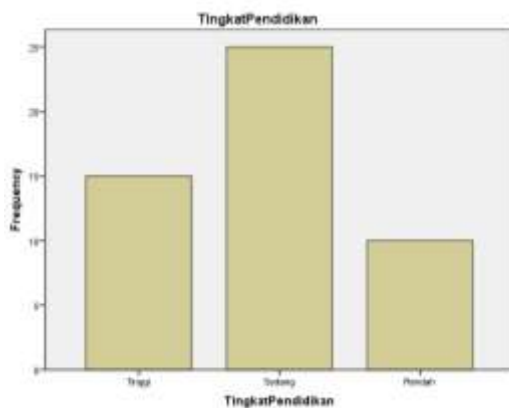
**a. Variabel Kategorik**

Penyajian data statistik deskriptif dengan variabel kategorik

merupakan jumlah atau frekuensi tiap kategori (n) dan persentase tiap kategori (%), yang biasanya disajikan dalam bentuk tabel atau grafik, dalam kaitannya dengan menunjukkan karakteristik satu set data dengan skala pengukuran kategorik. Berikut ini merupakan contoh penyajian variabel dengan skala pengukuran kategorik dalam bentuk grafik batang.

Tabel 1.1 Distribusi Frekuensi Jenis Kelamin dan Tingkat Pendidikan Responden

Variabel	n	%
Jenis Kelamin		
Laki-laki	22	44
Perempuan	28	56
Tingkat Pendidikan		
Rendah	10	20
Sedang	25	50
Tinggi	15	30
Total	50	100



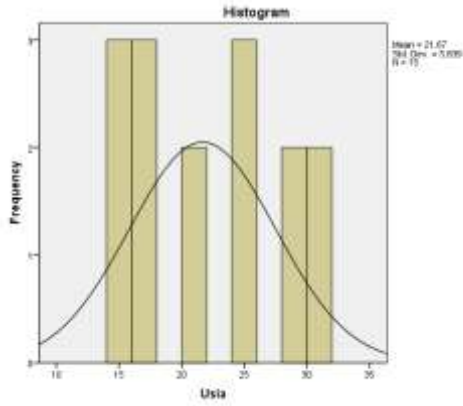
Gambar 1.1 Contoh penyajian variabel kategorik dalam bentuk diagram batang

## b. Variabel Numerik

Terdapat dua parameter ukuran pemusatan dan penyebaran yang biasa digunakan ketika peneliti melihat karakteristik satu set data dengan skala pengukuran numerik. Parameter ukuran pemusatan seperti mean, median, dan modus; parameter penyebaran seperti deviasi, varians, koefisien varians, interkuartil, range, dan minimum-maksimum. Data variabel dengan skala numerik biasanya disajikan dalam bentuk tabel atau grafik, juga dikenal sebagai histogram dan plots. Berikut adalah contoh penyajian data variabel dengan skala numerik dalam bentuk tabel dan histogram. Nilai median sebagai ukuran pemusatan dan standar deviasi (SD) sebagai ukuran penyebaran digunakan apabila distribusi data normal. Namun apabila distribusi data tidak normal, median akan digunakan sebagai ukuran pemusatan dan nilai minimum-maksimum akan digunakan sebagai ukuran penyebaran.

Table 1.2 Karakteristik Usia dan Berat Badan Responden

Variabel	Rerata	Median	Simpang Baku	Minimum	Maksimum
Usia	46,69	47	12,56	15	69
Berat Badan	50,4	50	8,33	45	64



Gambar 1.2 Contoh penyajian variabel numeric dalam bentuk histogram

## **DAFTAR PUSTAKA**

- Aziza, Nurul dkk. 2023. Metode Penelitian 1. Bandung: MEDIA SAINS INDONESIA
- Dahlan, Sopiudin. 2011. Statistik Untuk Kedokteran dan Kesehatan Edisi 5. Jakarta:Salemba Medika.
- Ghozali, I. 2016. Aplikasi Analisis Multivariete Dengan Program IBM SPSS 23. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro.
- Sugiyono. 2019. Statistika untuk Penelitian. Bandung : CV Alfabeta.
- Wahyuni, Moli. 2020. Statistik Deskriptif Untuk Penelitian Olah Data Manual Dan SPSS Versi 25. Yogyakarta: Bintang Pustaka Madani.

# UJI VALIDITAS DAN RELIABILITAS

*Oleh Leni Gustina, S.E., M.M.*

## 5.1. Uji Validitas

Sugiyono (2019:176) menjelaskan bahwa validitas adalah instrumen yang dapat digunakan untuk mengukur antara data yang terjadi pada objek dengan data yang dapat dikumpulkan oleh peneliti. Uji signifikansi dilakukan dengan cara nilai  $r$  hitung dibandingkan dengan nilai  $r$  tabel. Penentuan layak atau tidaknya suatu item ditentukan dengan uji signifikansi koefisien korelasi pada taraf signifikansi 0,1 yang berarti suatu item dapat dikatakan valid jika memiliki korelasi yang signifikan dengan skor total. Jika  $r$  hitung lebih besar dari  $r$  tabel dan nilainya positif maka item dinyatakan valid sedangkan jika  $r$  lebih kecil dari  $r$  tabel maka item dinyatakan tidak valid. Uji validitas terdiri dari dua macam, yaitu:

- Mengkorelasikan antar skor butir pertanyaan (item) dengan total item
- Mengkorelasikan antar masing-masing skor indikator item dengan total skor konstruk

Dengan kata lain suatu tes dikatakan valid jika sesuai dengan tujuan tes tersebut diadakan. Validitas terbagi tiga, yaitu validitas isi, validitas kriteria, dan validitas konstruk

a. Kriteria pengujian uji validitas

Dalam uji validitas ini, akan menjelaskan pengujian validitas yang mengkorelasikan antar masing-masing skor item indikator dengan total skor konstruk. Tingkat signifikansi yang digunakan yaitu 0,05. Kriteria pengujiannya yaitu:  $H_0$  diterima apabila  $r_{hitung} > r_{tabel}$ , (alat ukur yang digunakan valid atau sah)  $H_0$  ditolak apabila  $r_{statistik} \leq r_{tabel}$ . (alat ukur yang digunakan tidak valid). Cara menentukan besar nilai R tabel  $R_{tabel} = df (N-2)$ , tingkat signifikansi uji dua arah. Misalnya  $R_{tabel} = df (13-2, 0,05)$ . Untuk mendapatkan nilai R tabel kita harus melihat ditebal R.

## 5.2. Langkah-langkah Pengujian Uji Validitas

Langkah-langkah pengujian Uji Validitas menggunakan SPSS. Untuk praktek latihan menggunakan SPSS versi 22. Pertama yang harus diperhatikan adalah apa yang akan diuji. Variabelnya harus jelas. Adapun Langkah-langkah pengujiannya, yaitu sebagai berikut :

1. Buat skor total masing-masing variabel (Tabel perhitungan skor)
2. Klik Analyze -> Correlate -> Bivariate (Gambar/Output SPSS)
3. Masukkan seluruh item variabel x ke Variabels.
4. Cek list Pearson ; Two Tailed ; Flag.
5. Klik Ok.

Dalam penelitian ini, uji validitas dilakukan dengan menggunakan SPSS. Jumlah responden pada uji validitas ini ialah 40 orang. Menurut Singarimbun dan Effendi (2012, p. 137), jumlah responden pengujian minimal 30 orang agar distribusi nilai akan

mendekati kurva normal.

### **5.3. Interpretasi hasil Output Uji Validitas**

Interpretasi hasil output uji validitas adalah untuk mengetahui apakah instrumen yang digunakan valid atau tidak. Validitas berarti instrumen tersebut dapat digunakan untuk mengukur apa yang seharusnya diukur. Berikut adalah beberapa hal yang perlu diketahui tentang uji validitas:

1. Uji validitas dilakukan untuk mengukur ketepatan dan kecermatan suatu variabel terkait fungsinya dalam suatu penelitian.
2. Penelitian dikatakan valid apabila mampu mengukur apa yang diinginkan dan dapat mengungkapkan data dari variabel-variabel yang diteliti secara tepat.
3. Uji validitas tiap butir digunakan analisis item yaitu mengkorelasikan skor tiap butir skor total yang merupakan jumlah tiap skor butir.
4. Setelah menemukan hasil dari uji validitas, langkah berikutnya adalah melakukan uji reliabilitas

### **5.4. Uji Reliabilitas**

Uji reliabilitas adalah alat untuk mengukur tingkat keandalan, keakuratan, ketelitian, dan konsistensi dari indikator yang ada dalam kuesioner. Menurut para ahli, uji reliabilitas dapat diartikan sebagai: uji reliabilitas adalah alat untuk mengukur kuesioner yang merupakan indikator dari suatu variable ( Menurut Ghazali ; 2011). Sedangkan Menurut Sugiyono (2017: 130) menyatakan bahwa uji reliabilitas adalah sejauh mana hasil pengukuran dengan menggunakan objek yang sama, akan menghasilkan data yang sama.

## 2.2. Kriteria Pengujian Uji Reliabilitas

- a. Instrumen dikatakan reliabel atau terpercaya jika nilai koefisien reliabilitasnya lebih dari 0,6
- b. Instrumen dikatakan tidak reliabel jika nilai koefisien reliabilitasnya kurang dari 0,6

Adapun kriteria dari pengujian reliabilitas adalah: 1. Jika nilai koefisien reliabilitas  $> 0,6$  maka instrumen memiliki reliabilitas yang baik atau dengan kata lain instrumen adalah reliabel atau terpercaya. 2. Jika nilai koefisien reliabilitas  $< 0,6$  maka instrumen yang diuji tersebut adalah tidak reliabel.

Reliabilitas instrumen dapat diuji dengan beberapa uji reliabilitas. Beberapa uji reliabilitas suatu instrumen yang bisa digunakan antara lain test-retest, ekuivalen, dan internal consistency . Internal consistency sendiri memiliki beberapa teknik uji yang berbeda.

## 2.3. Langkah-langkah pengujian Uji Reliabilitas

1. Klik Analyze
2. Klik Scale
3. Klik Reliability Analysis
4. Masukkan seluruh item variabel X ke Items

Hasil perhitungan uji reliabilitas akan muncul pada Output dalam bentuk tabel, seperti Case Processing Summary, Reliability Statistics, Item-Total Statistics, dan Scale Statistics. Untuk mengetahui apakah kuesioner reliabel atau tidak, dapat dilihat dari nilai alpha:

1. Alpha  $> 0.7$ , maka kuesioner reliabel
2. Alpha  $> 0.80$ , maka kuesioner reliabel dan konsisten
3. Alpha  $> 0.90$ , maka kuesioner reliabel sempurna
4. Alpha  $0.70-0.90$ , maka kuesioner reliabel tinggi
5. Alpha  $0.50-0.70$ , maka kuesioner reliabel moderat

6.  $\text{Alpha} < 0.50$ , maka kuesioner reliabel rendah

#### 2.4. Interpretasi Hasil Output Uji reliabilitas

Interpretasi hasil output uji reliabilitas dapat dilakukan dengan melihat nilai koefisien reliabilitas, nilai Cronbach Alpha, dan nilai alpha Cronbach:

- **Koefisien reliabilitas**

Koefisien reliabilitas ditunjukkan dengan huruf "r" dan dinyatakan sebagai angka yang berkisar antara 0 dan 1,00. Nilai  $r = 0$  menunjukkan tidak ada reliabilitas, dan  $r = 1,00$  yang menunjukkan reliabilitas sempurna.

- **Nilai Cronbach Alpha**

Suatu variabel dapat dikatakan reliabel jika memberikan nilai cronbach alpha ( $\alpha$ )  $> 0,6$ . Semakin tinggi nilai Cronbach Alpha, maka tingkat reliabilitas data akan semakin baik.

- **Nilai alpha Cronbach**

Nilai alpha Cronbach yang negatif menunjukkan adanya reliabilitas negatif. Nilai negatif ini disebabkan oleh derivasi data dari item yang mewakili skala yang berbeda.

## **DAFTAR PUSTAKA**

- Martono, Nanang. 2010. Statistik Sosial: Teori dan Aplikasi Program SPSS. Yogyakarta: Gava Media
- Ghozali, I. (2015) Aplikasi Multivariete Dengan Program IBM SPSS 23 (Edisi 3). Semarang: Badan Penerbit Universitas Diponegoro.
- Sugiyono (2019). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung : Alfabeta
- Singarimbun, Masri & Sofian Effendi. 2008. Metode Penelitian Survei, Jakarta: LP3ES.
- Sugiyono. 2011. Statistika untuk Penelitian. Bandung: Alfabeta

# UJI ASUMSI KLASIK

*Oleh Pilifus Junianto, S.S., M.M.*

## 6.1 Uji Normalitas

Uji normalitas merupakan langkah penting dalam analisis statistik untuk menentukan apakah data yang digunakan berdistribusi normal atau tidak. Dalam konteks pengolahan data dengan SPSS, uji ini sering digunakan sebagai prasyarat untuk memilih jenis analisis statistik yang sesuai, baik parametrik maupun non-parametrik. Data yang berdistribusi normal memungkinkan penggunaan uji parametrik yang lebih kuat secara statistik, sementara data yang tidak berdistribusi normal memerlukan alternatif non-parametrik (Ghasemi & Zahediasl, 2012).

Dalam SPSS, uji normalitas dapat dilakukan melalui beberapa metode, termasuk Kolmogorov-Smirnov Test dan Shapiro-Wilk Test. Kolmogorov-Smirnov cocok digunakan untuk sampel besar, sedangkan Shapiro-Wilk lebih direkomendasikan untuk ukuran sampel kecil. Selain itu, grafik seperti histogram, Q-Q plot, dan boxplot juga dapat digunakan untuk memvisualisasikan distribusi data. Kedua uji ini memberikan nilai signifikansi yang menjadi dasar pengambilan keputusan (Razali & Wah, 2011).

Tahapan melakukan uji normalitas di SPSS melibatkan beberapa langkah sederhana. Pertama, masukkan data ke dalam variabel di SPSS. Kedua, buka menu "Analyze" dan pilih "Descriptive Statistics," kemudian klik "Explore." Setelah itu, pilih variabel yang akan diuji, kemudian centang opsi untuk uji normalitas dalam sub-menu "Plots." Hasilnya akan menunjukkan nilai signifikansi Kolmogorov-Smirnov dan Shapiro-Wilk, serta grafik pendukung.

Interpretasi hasil uji normalitas didasarkan pada nilai signifikansi (*p*-value). Jika nilai signifikansi lebih besar dari 0,05, data dianggap berdistribusi normal. Sebaliknya, jika *p*-value kurang dari 0,05, maka data tidak berdistribusi normal. Misalnya, jika nilai Shapiro-Wilk untuk suatu data adalah 0,03, maka data tersebut tidak memenuhi asumsi normalitas, sehingga perlu dipertimbangkan penggunaan uji non-parametrik (Field, 2013).

Ketidaknormalan distribusi data dapat disebabkan oleh berbagai faktor, seperti outlier, heterogenitas, atau karakteristik intrinsik data itu sendiri. Untuk mengatasi hal ini, metode transformasi data, seperti log atau square root transformation, dapat digunakan untuk mendekati data pada distribusi normal. Namun, jika transformasi tidak efektif, analisis non-parametrik menjadi solusi terbaik (Pallant, 2020).

Dengan memahami uji normalitas dalam SPSS, peneliti dapat memastikan validitas dan reliabilitas analisis statistik mereka. Pemilihan metode analisis yang sesuai akan meningkatkan kualitas interpretasi hasil penelitian, memberikan dasar yang kuat untuk pengambilan keputusan berbasis data.

## **6.2 Multikolinearitas**

Multikolinearitas adalah masalah yang sering muncul dalam analisis regresi linier ketika terdapat hubungan linear yang kuat antara dua atau lebih variabel independen dalam model. Kondisi ini

dapat mengganggu interpretasi koefisien regresi karena memperburuk akurasi estimasi parameter. Jika multikolinearitas terjadi, model regresi akan menghasilkan koefisien yang tidak stabil, yang berarti perubahan kecil dalam data dapat menyebabkan perubahan besar dalam hasil analisis (Gujrati, 2003).

Penyebab utama multikolinearitas adalah adanya hubungan yang sangat erat antara variabel independen yang digunakan dalam model. Sebagai contoh, jika dua variabel independen sangat berkorelasi satu sama lain, sulit untuk memisahkan dampak masing-masing terhadap variabel dependen. Dalam beberapa kasus, variabel-variabel ini mungkin mengukur aspek yang sangat mirip dari fenomena yang sama. Keadaan ini memperburuk estimasi koefisien dan meningkatkan varians, yang dapat menurunkan signifikansi statistik (Kennedy, 2008).

Untuk mendeteksi multikolinearitas, analisis regresi dapat dilakukan dengan mengamati beberapa indikator. Salah satu alat yang paling umum adalah Variance Inflation Factor (VIF). VIF mengukur seberapa besar varians dari estimasi koefisien regresi meningkat karena kolinearitas antar variabel independen. Nilai VIF yang lebih besar dari 10 umumnya dianggap sebagai indikasi adanya multikolinearitas yang serius (O'Brien, 2007). Selain itu, nilai korelasi Pearson yang tinggi antara variabel independen juga dapat menunjukkan masalah multikolinearitas.

Penyelesaian untuk masalah multikolinearitas dapat dilakukan dengan beberapa cara. Salah satunya adalah dengan menghapus satu atau lebih variabel yang sangat berkorelasi. Alternatif lain adalah dengan menggabungkan variabel yang berkorelasi tinggi menjadi satu variabel komposit, yang sering dilakukan dalam teknik analisis faktor atau principal component analysis (PCA). Dengan cara ini, hubungan antar variabel dapat disederhanakan tanpa mengurangi informasi yang penting dalam model (Tabachnick & Fidell, 2013).

Namun, tidak semua kasus multikolinearitas harus diatasi dengan menghapus atau menggabungkan variabel. Dalam beberapa situasi, jika variabel yang sangat berkorelasi secara teoritis penting, model regresi masih bisa digunakan meskipun ada sedikit multikolinearitas. Teknik regresi yang lebih canggih, seperti regresi ridge atau regresi lasso, dapat digunakan untuk mengatasi multikolinearitas tanpa menghilangkan variabel independen yang signifikan (Hoerl & Kennard, 1970).

Pada akhirnya, mengidentifikasi dan menangani multikolinearitas sangat penting untuk memastikan bahwa hasil dari model regresi dapat diinterpretasikan dengan benar. Peneliti perlu memeriksa multikolinearitas sebagai bagian dari langkah diagnostik setelah membangun model regresi. Dengan memahami dan mengelola multikolinearitas, penelitian dapat menghasilkan kesimpulan yang lebih akurat dan dapat diandalkan.

### **6.3 Heterokedasdisitas**

Heteroskedastisitas adalah masalah yang terjadi dalam analisis regresi ketika varians kesalahan tidak konstan di seluruh nilai variabel independen. Dalam regresi linier klasik, salah satu asumsi dasar adalah bahwa kesalahan (error) memiliki varians yang konstan, yang dikenal dengan istilah homoskedastisitas. Ketika asumsi ini dilanggar dan terdapat perbedaan varians kesalahan, fenomena ini disebut sebagai heteroskedastisitas (Wooldridge, 2016). Heteroskedastisitas dapat mengurangi efisiensi estimasi parameter dan mempengaruhi uji signifikansi.

Pada model regresi yang terpengaruh heteroskedastisitas, koefisien regresi tetap tidak bias, tetapi variansnya menjadi tidak efisien, yang berarti standar error yang dihasilkan mungkin terlalu kecil atau besar. Hal ini dapat menghasilkan uji t dan uji F yang tidak dapat diandalkan, karena nilai p yang dihasilkan bisa menyesatkan,

baik terlalu rendah atau terlalu tinggi. Oleh karena itu, penting untuk mendeteksi dan menangani masalah heteroskedastisitas agar hasil regresi tetap valid dan dapat dipertanggungjawabkan (Gujarati & Porter, 2009).

Untuk mendeteksi heteroskedastisitas dalam SPSS, beberapa metode dapat digunakan, salah satunya adalah dengan memvisualisasikan data menggunakan plot residual. Plot ini menggambarkan hubungan antara nilai prediksi (*fitted values*) dan residual (selisih antara nilai observasi dan prediksi). Jika terdapat pola atau pembentukan bentuk tertentu dalam plot residual, ini bisa menjadi indikasi adanya heteroskedastisitas. Selain itu, uji formal seperti Uji Breusch-Pagan atau Uji White juga dapat digunakan untuk mendeteksi masalah ini dengan lebih tepat (Breusch & Pagan, 1979).

Setelah mendeteksi adanya heteroskedastisitas, ada beberapa cara untuk mengatasinya. Salah satunya adalah dengan mentransformasi variabel independen atau dependen, seperti menggunakan logaritma, untuk menstabilkan varians. Metode lain yang sering digunakan adalah regresi dengan *robust standard errors*, yang memberikan estimasi koefisien yang lebih efisien meskipun heteroskedastisitas ada. Di SPSS, *robust standard errors* dapat diakses melalui menu Options dalam menu Linear Regression. Ini memungkinkan pengujian yang lebih akurat terhadap koefisien meskipun asumsi homoskedastisitas dilanggar (Huber, 1967).

Selain itu, regresi *generalized least squares (GLS)* dapat digunakan sebagai alternatif untuk regresi linier klasik dalam menangani heteroskedastisitas. GLS memberikan bobot yang berbeda pada setiap pengamatan, yang mengoreksi ketidakseimbangan varians dalam data. Dengan pendekatan ini, estimasi koefisien regresi menjadi lebih efisien dan valid meskipun heteroskedastisitas tetap ada (Kmenta, 1986). Teknik ini sering digunakan dalam model-model ekonometrika yang lebih kompleks.

Secara keseluruhan, heteroskedastisitas adalah masalah yang dapat memengaruhi validitas analisis regresi, tetapi dengan mendeteksi dan mengatasi masalah ini dengan menggunakan alat dan teknik yang tepat di SPSS, peneliti dapat memastikan hasil analisis regresi tetap dapat diandalkan. Oleh karena itu, penting untuk selalu memeriksa asumsi homoskedastisitas sebelum menarik kesimpulan dari model regresi.

#### **6.4 Autokorelasi**

Autokorelasi, juga dikenal sebagai serial correlation, adalah fenomena yang terjadi ketika residual dalam model regresi linier tidak independen satu sama lain. Dalam regresi klasik, asumsi yang diharapkan adalah bahwa kesalahan (error terms) pada satu pengamatan tidak berkorelasi dengan kesalahan pada pengamatan lainnya. Jika kesalahan pada suatu pengamatan berkorelasi dengan kesalahan pada pengamatan sebelumnya, maka terjadi autokorelasi, yang dapat merusak validitas hasil model (Wooldridge, 2016).

Masalah autokorelasi sering ditemukan dalam data time series, di mana nilai-nilai variabel dependen pada waktu tertentu bergantung pada nilai sebelumnya. Misalnya, dalam model prediksi ekonomi atau keuangan, nilai saat ini seringkali dipengaruhi oleh nilai sebelumnya. Namun, autokorelasi juga bisa muncul pada data cross-sectional jika ada pengaruh yang tertunda atau pola yang tidak terobservasi dalam data. Autokorelasi mengurangi efisiensi estimasi koefisien dan dapat membuat hasil uji signifikansi tidak valid (Gujarati & Porter, 2009).

Dalam SPSS, deteksi autokorelasi dapat dilakukan dengan menggunakan Durbin-Watson statistic yang dihasilkan dalam output regresi. Durbin-Watson adalah uji yang secara spesifik menguji adanya autokorelasi pada residual model regresi. Nilai Durbin-Watson yang mendekati 2 menunjukkan tidak adanya autokorelasi,

sedangkan nilai yang lebih rendah dari 1 atau lebih tinggi dari 3 menunjukkan adanya autokorelasi positif atau negatif, masing-masing. Pengujian ini penting untuk memeriksa apakah model regresi linier yang dibangun sesuai dengan asumsi independensi residual (Durbin & Watson, 1950).

Selain Durbin-Watson, teknik lain yang bisa digunakan untuk mendeteksi autokorelasi adalah dengan melihat plot residual. Jika residual menunjukkan pola yang teratur (misalnya, bentuk gelombang atau siklus), ini bisa menjadi indikasi adanya autokorelasi. Uji formal lainnya seperti uji Breusch-Godfrey atau uji Ljung-Box dapat digunakan untuk mendeteksi autokorelasi lebih lanjut, terutama pada data time series. Uji Breusch-Godfrey lebih fleksibel karena dapat digunakan untuk mendeteksi autokorelasi yang lebih tinggi (lebih dari satu lag) (Breusch & Godfrey, 1978).

Untuk mengatasi masalah autokorelasi, beberapa solusi dapat diterapkan. Salah satunya adalah transformasi data, seperti memasukkan variabel lag (nilai waktu sebelumnya) dalam model, yang dapat membantu mengurangi pengaruh autokorelasi. Alternatif lain adalah menggunakan model regresi ARIMA (AutoRegressive Integrated Moving Average) untuk data time series, yang dirancang khusus untuk menangani autokorelasi dalam data (Box, Jenkins, & Reinsel, 2015). Pendekatan ini sangat berguna dalam analisis prediktif yang melibatkan data dengan pola waktu.

Selain itu, regresi dengan standar error yang robust juga dapat digunakan untuk mengatasi masalah autokorelasi, terutama jika autokorelasi muncul dalam model regresi dengan data cross-sectional. Dengan menggunakan robust standard errors, koefisien regresi tetap dapat diestimasi dengan benar meskipun terdapat autokorelasi. Penggunaan robust standard errors ini dapat dengan mudah diterapkan di SPSS melalui menu Linear Regression dan memilih opsi Robust Standard Errors (Huber, 1967).

## 6.5 Linearitas

Linearitas adalah salah satu asumsi dasar dalam model regresi linier, yang menyatakan bahwa hubungan antara variabel independen dan variabel dependen bersifat linear. Artinya, perubahan satu unit pada variabel independen akan menghasilkan perubahan yang proporsional pada variabel dependen. Jika hubungan ini tidak linear, maka model regresi linier yang digunakan tidak akan memberikan hasil yang valid atau akurat (Gujarati & Porter, 2009). Asumsi linearitas sangat penting untuk memastikan estimasi koefisien regresi yang tepat dan interpretable.

Dalam analisis regresi, linearitas berarti bahwa hubungan antara variabel-variabel yang dianalisis harus digambarkan dengan garis lurus. Jika hubungan antara variabel independen dan dependen memiliki pola yang lebih kompleks, seperti bentuk kurva, maka model regresi linier sederhana mungkin tidak akan memberikan hasil yang memadai. Oleh karena itu, sebelum melakukan regresi, penting untuk memeriksa apakah hubungan antar variabel yang ada memang linear (Wooldridge, 2016).

Untuk mendeteksi linearitas dalam SPSS, salah satu metode yang umum digunakan adalah dengan memplot residual terhadap nilai prediksi (*fitted values*). Jika plot residual tidak menunjukkan pola tertentu dan tersebar acak di sekitar garis horizontal, maka asumsi linearitas cenderung terpenuhi. Sebaliknya, jika terlihat pola atau kurva dalam plot tersebut, ini mengindikasikan adanya hubungan non-linear antara variabel yang dapat mempengaruhi hasil analisis (Tabachnick & Fidell, 2013).

Selain plot residual, uji scatter plot juga dapat digunakan untuk memeriksa linearitas. Scatter plot menggambarkan hubungan langsung antara variabel independen dan dependen. Jika titik-titik dalam scatter plot membentuk pola linear, maka asumsi linearitas dapat dianggap terpenuhi. Namun, jika pola tersebut berbentuk

kurva atau tidak teratur, maka model regresi linier mungkin tidak sesuai, dan pendekatan alternatif seperti regresi polinomial atau transformasi data bisa dipertimbangkan (Field, 2013).

Jika hubungan antar variabel tidak bersifat linear, beberapa pendekatan dapat diambil untuk mengatasi masalah ini. Salah satu metode adalah dengan menggunakan transformasi data. Transformasi ini bisa berupa logaritma, akar kuadrat, atau pangkat dari variabel, yang bertujuan untuk mengubah hubungan menjadi linear. Misalnya, jika variabel dependen memiliki hubungan non-linear dengan variabel independen, logaritma dari variabel dependen dapat mengubah hubungan tersebut menjadi lebih linear dan lebih sesuai dengan asumsi model regresi linier (Gujarati, 2003).

Alternatif lain untuk menangani masalah linearitas adalah dengan menggunakan model regresi yang lebih kompleks, seperti regresi polinomial atau regresi spline. Dalam regresi polinomial, variabel independen dikembangkan menjadi bentuk kuadrat atau lebih tinggi, yang memungkinkan model untuk menangani hubungan yang tidak linear antara variabel. Sementara itu, regresi spline membagi data menjadi segmen-segmen dan memungkinkan model untuk menangani hubungan non-linear dengan lebih fleksibel (Kaufman & Rousseeuw, 2005). Kedua metode ini bisa digunakan ketika transformasi sederhana tidak cukup untuk mengatasi masalah linearitas.

## DAFTAR PUSTAKA

- Breusch, T. S., & Godfrey, L. G. (1978). A survey of recent advances in the econometrics of time series. *Econometrica*, 46(5), 1115–1129. <https://doi.org/10.2307/1911365>
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. *Biometrika*, 37(3/4), 409–428. <https://doi.org/10.1093/biomet/37.3-4.409>
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.
- Gujarati, D. N. (2003). *Basic econometrics* (4th ed.). McGraw-Hill.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 221–233). University of California Press.
- Kaufman, L., & Rousseeuw, P. J. (2005). *Finding groups in data: An introduction to cluster analysis*. Wiley-Interscience.
- Kmenta, J. (1986). *Elements of econometrics* (2nd ed.). Macmillan.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (6th ed.). Cengage Learning.

# UJI HIPOTESIS PARAMETRIK

*Oleh Kiki Helencia, S.E., M.M*

## 7.1 Uji t (t-test)

Uji t (t-test) adalah salah satu metode statistika yang digunakan untuk menguji apakah terdapat perbedaan yang signifikan antara rata-rata dari dua kelompok atau antara rata-rata sampel dengan nilai yang diketahui dalam populasi. Uji ini sangat berguna ketika ukuran sampel kecil dan variansi populasi tidak diketahui. Uji t pertama kali diperkenalkan oleh William Sealy Gosset pada tahun 1908 dengan nama samaran "Student," yang kemudian dikenal dengan istilah Student's t-test. Uji ini sering digunakan dalam berbagai disiplin ilmu, seperti psikologi, kedokteran, pendidikan, dan ilmu sosial (Biau & Kernéis, 2018).

Secara umum, terdapat tiga jenis uji t yang sering digunakan: one-sample t-test, independent two-sample t-test, dan paired sample t-test. Uji one-sample t-test digunakan untuk menguji apakah rata-rata sampel berbeda signifikan dari nilai rata-rata yang diharapkan atau nilai yang diketahui dalam populasi. Misalnya, seorang peneliti ingin mengetahui apakah rata-rata tinggi badan mahasiswa di universitasnya berbeda dari rata-rata tinggi badan yang umum di negara tersebut. Dalam uji ini, nilai rata-rata yang diharapkan menjadi pembanding untuk sampel yang diteliti (Cohen et al., 2013).

Pada uji independent two-sample t-test, perbandingan dilakukan antara dua kelompok yang independen satu sama lain. Uji ini digunakan untuk menguji apakah terdapat perbedaan yang signifikan antara rata-rata dua kelompok yang tidak saling berhubungan. Contoh penerapannya adalah untuk mengetahui apakah ada perbedaan rata-rata penghasilan antara pria dan wanita dalam suatu kota tertentu. Penting untuk memastikan bahwa kedua sampel independen satu sama lain dan tidak ada pengaruh eksternal yang memengaruhi hasil (Field, 2013).

Sementara itu, paired sample t-test digunakan ketika sampel yang diuji berpasangan, seperti dalam eksperimen yang dilakukan dengan pengukuran sebelum dan sesudah perlakuan terhadap subjek yang sama. Misalnya, untuk mengukur efek dari suatu obat, seorang peneliti dapat mengukur tekanan darah seseorang sebelum dan setelah pemberian obat. Uji ini akan menguji apakah perbedaan antara dua pengukuran tersebut cukup signifikan untuk menyimpulkan bahwa perlakuan tersebut memberikan pengaruh yang nyata (Gravetter & Wallnau, 2016).

Uji t membutuhkan asumsi dasar yang perlu dipenuhi agar hasilnya valid. Salah satu asumsi utama adalah bahwa data dalam setiap kelompok mengikuti distribusi normal. Untuk memastikan asumsi ini, peneliti dapat menggunakan uji normalitas seperti uji Shapiro-Wilk atau Kolmogorov-Smirnov. Selain itu, uji t juga mengasumsikan bahwa variansi dalam kedua kelompok yang dibandingkan adalah homogen atau sama (homoscedasticity), yang dapat diuji menggunakan uji Levene. Jika asumsi ini tidak terpenuhi, alternatif lain seperti uji non-parametrik (misalnya uji Mann-Whitney atau Wilcoxon) dapat dipertimbangkan (Field, 2013).

Selain itu, dalam melakukan uji t, penting untuk memilih tingkat signifikansi yang tepat, yang biasa disebut sebagai alpha ( $\alpha$ ), umumnya ditetapkan pada 0,05. Ini berarti bahwa peneliti menerima risiko 5%

untuk membuat kesalahan tipe I, yaitu menolak hipotesis nol yang sebenarnya benar. Jika nilai  $p$  yang diperoleh dari perhitungan lebih kecil dari  $\alpha$ , maka hipotesis nol ditolak, yang mengindikasikan bahwa terdapat perbedaan yang signifikan secara statistik antara kelompok yang diuji (Biau & Kernéis, 2018).

Meskipun uji  $t$  adalah alat yang sangat berguna dalam analisis statistik, ada beberapa keterbatasan yang perlu diperhatikan. Uji ini lebih efektif digunakan pada data yang berskala interval atau rasio, dan tidak cocok untuk data kategorikal. Selain itu, uji  $t$  tidak memberikan informasi tentang besar efek atau seberapa kuat hubungan antara dua variabel. Oleh karena itu, peneliti sering menggabungkan uji  $t$  dengan ukuran efek seperti Cohen's  $d$  untuk menilai besar perbedaan antara kelompok (Cohen, 1988).

Dalam kesimpulannya, uji  $t$  ( $t$ -test) merupakan alat yang sangat penting dalam statistik inferensial yang digunakan untuk menguji perbedaan rata-rata antara kelompok. Meskipun memiliki beberapa keterbatasan, penerapannya yang luas dalam berbagai bidang menjadikannya teknik yang sangat berguna dalam penelitian ilmiah. Pemahaman yang tepat tentang kapan dan bagaimana menggunakan uji  $t$  akan memastikan analisis yang valid dan hasil yang bermakna dalam penelitian.

## **7.2 Uji Z (Z-test)**

Uji Z ( $Z$ -test) adalah salah satu metode statistika yang digunakan untuk menguji hipotesis tentang rata-rata atau proporsi populasi, terutama ketika ukuran sampel besar dan variansi populasi diketahui atau diasumsikan diketahui. Berbeda dengan uji  $t$ , yang lebih sering digunakan pada sampel kecil dan variansi populasi tidak diketahui, uji  $Z$  lebih sering digunakan ketika ukuran sampel lebih besar ( $n > 30$ ) dan distribusi data dianggap normal atau dapat diperkirakan untuk mendekati distribusi normal (Zar, 2010). Uji  $Z$  pertama kali digunakan

untuk membandingkan nilai rata-rata sampel dengan nilai rata-rata populasi.

Pada dasarnya, uji Z digunakan untuk menguji apakah nilai rata-rata atau proporsi suatu sampel berbeda secara signifikan dari nilai rata-rata atau proporsi populasi yang diketahui. Misalnya, seorang peneliti mungkin ingin mengetahui apakah rata-rata berat badan siswa di suatu sekolah berbeda dari rata-rata berat badan nasional. Uji Z memungkinkan peneliti untuk menguji perbedaan tersebut dengan menghitung nilai Z, yang menunjukkan seberapa jauh rata-rata sampel berada dari rata-rata populasi dalam satuan deviasi standar (Gibbons, 2011).

Uji Z dilakukan dengan menghitung nilai Z berdasarkan rumus standar:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

- $\bar{X}$  adalah rata-rata sampel,
- $\mu$  adalah rata-rata populasi yang diketahui,
- $\sigma$  adalah deviasi standar populasi, dan
- $n$  adalah ukuran sampel (Field, 2013).

Uji Z juga digunakan untuk menguji hipotesis tentang proporsi populasi. Misalnya, jika suatu perusahaan ingin mengetahui apakah proporsi pelanggan yang puas dengan layanan mereka lebih dari 80%, uji Z untuk proporsi dapat digunakan untuk menguji apakah proporsi sampel pelanggan yang puas berbeda secara signifikan dari 0,80. Dalam hal ini, rumus untuk uji Z akan disesuaikan untuk menghitung selisih antara proporsi sampel dan proporsi populasi, yang kemudian dibandingkan dengan deviasi standar proporsi tersebut (Keller, 2011).

Salah satu alasan mengapa uji Z sering digunakan adalah karena distribusi sampel rata-rata cenderung mendekati distribusi normal menurut Teorema Limit Pusat ketika ukuran sampel cukup besar, meskipun populasi asalnya tidak terdistribusi normal. Oleh karena itu, uji Z banyak digunakan dalam situasi di mana data diambil dari

populasi besar dengan distribusi yang tidak diketahui, tetapi ukuran sampelnya cukup besar untuk memenuhi asumsi distribusi normal (Johnson & Bhattacharyya, 2010). Hal ini menjadikan uji Z sangat berguna dalam penelitian dengan data besar atau data yang diperoleh dari survei.

Namun, ada beberapa asumsi penting yang harus dipenuhi agar uji Z valid. Salah satu asumsi utama adalah bahwa variansi populasi ( $\sigma^2$ ) diketahui atau dapat diasumsikan diketahui. Dalam banyak kasus praktis, informasi tentang variansi populasi mungkin tidak tersedia, dan dalam situasi seperti itu, uji t lebih disarankan. Selain itu, ukuran sampel juga harus cukup besar untuk memenuhi asumsi normalitas atau untuk mengurangi bias dalam estimasi distribusi (Zar, 2010).

Selain itu, uji Z memiliki dua jenis pengujian, yaitu uji Z satu arah (one-tailed) dan uji Z dua arah (two-tailed). Uji Z satu arah digunakan ketika hipotesis mengarah pada satu sisi distribusi (misalnya, apakah nilai sampel lebih besar atau lebih kecil dari nilai populasi yang diharapkan), sedangkan uji Z dua arah digunakan untuk menguji apakah nilai sampel berbeda dari nilai populasi dalam kedua arah (misalnya, apakah nilai sampel lebih besar atau lebih kecil dari nilai rata-rata populasi) (Gibbons, 2011).

Secara keseluruhan, uji Z merupakan alat yang sangat berguna dalam statistik inferensial, terutama dalam konteks pengujian hipotesis dengan sampel besar dan ketika variansi populasi diketahui. Meskipun memiliki keterbatasan dalam hal asumsi yang diperlukan, uji Z tetap menjadi salah satu metode yang sangat relevan dalam berbagai penelitian dan analisis data.

### **7.3 Uji ANOVA (Analysis of Variance)**

Uji ANOVA (Analysis of Variance) adalah teknik statistika yang digunakan untuk menguji apakah ada perbedaan yang signifikan antara rata-rata lebih dari dua kelompok atau lebih. Uji ini bertujuan untuk menganalisis variasi data dalam satu atau lebih faktor (variabel independen) dan menentukan apakah variasi tersebut cukup besar untuk mengindikasikan perbedaan rata-rata antara kelompok yang

diuji. ANOVA pertama kali diperkenalkan oleh Ronald A. Fisher pada tahun 1920-an dan sejak itu telah menjadi alat penting dalam analisis data eksperimen (Fisher, 1935).

Secara umum, ANOVA membandingkan variansi antar kelompok (variansi antar perlakuan atau grup) dengan variansi dalam kelompok (variansi dalam perlakuan atau grup) untuk menentukan apakah perbedaan antar rata-rata kelompok lebih besar daripada perbedaan yang diharapkan berdasarkan kebetulan. Jika variasi antar kelompok lebih besar dibandingkan dengan variasi dalam kelompok, maka ada indikasi bahwa perbedaan antara kelompok tersebut signifikan secara statistik. Oleh karena itu, ANOVA dapat membantu menentukan apakah suatu perlakuan atau variabel independen mempengaruhi variabel dependen secara signifikan (Yates, Moore, & McCabe, 1999).

Jenis utama dari uji ANOVA adalah **One-way ANOVA** dan **Two-way ANOVA**. **One-way ANOVA** digunakan ketika hanya ada satu faktor yang diuji untuk mempengaruhi variabel dependen. Misalnya, jika seorang peneliti ingin menguji apakah ada perbedaan dalam hasil ujian antara tiga kelompok siswa yang mengikuti tiga metode pengajaran yang berbeda, maka One-way ANOVA dapat digunakan untuk menganalisis data ini (Field, 2013). Uji ini menguji hipotesis nol yang menyatakan bahwa semua rata-rata kelompok adalah sama, dan hipotesis alternatif yang menyatakan bahwa setidaknya ada satu perbedaan rata-rata antar kelompok.

Sementara itu, **Two-way ANOVA** digunakan ketika terdapat dua faktor atau variabel independen yang ingin diuji pengaruhnya terhadap variabel dependen. Selain menguji pengaruh masing-masing faktor, Two-way ANOVA juga dapat menguji interaksi antara kedua faktor tersebut. Misalnya, seorang peneliti ingin menguji efek dua faktor, yaitu jenis latihan fisik (seperti latihan aerobik dan angkat beban) dan jenis diet (diet rendah kalori dan diet tinggi kalori), terhadap penurunan berat badan. Two-way ANOVA akan memungkinkan peneliti untuk menguji pengaruh masing-masing faktor dan apakah ada interaksi antara faktor-faktor tersebut dalam mempengaruhi penurunan berat badan (Keller, 2011).

Asumsi dasar yang perlu dipenuhi dalam uji ANOVA meliputi independensi sampel, distribusi normal dalam setiap kelompok, dan kesetaraan variansi antar kelompok (homogenitas variansi). Untuk menguji apakah asumsi homogenitas variansi terpenuhi, uji **Levene's Test** atau uji **Bartlett's Test** dapat digunakan. Jika data tidak memenuhi asumsi normalitas atau homogenitas variansi, uji non-parametrik seperti **Kruskal-Wallis Test** dapat menjadi alternatif (Zar, 2010).

Jika uji ANOVA menghasilkan nilai p yang lebih kecil dari tingkat signifikansi yang telah ditentukan (misalnya  $p < 0,05$ ), maka hipotesis nol yang menyatakan bahwa semua rata-rata kelompok adalah sama ditolak. Namun, ANOVA hanya memberi tahu kita bahwa ada perbedaan yang signifikan antara kelompok, tetapi tidak memberitahukan kelompok mana yang berbeda secara spesifik. Oleh karena itu, setelah ANOVA menunjukkan hasil signifikan, **post hoc test** seperti **Tukey's HSD** atau **Bonferroni test** digunakan untuk menentukan kelompok mana yang secara signifikan berbeda satu sama lain (Keller, 2011).

ANOVA juga dapat digunakan dalam desain eksperimen yang lebih kompleks, seperti **repeated measures ANOVA**, di mana pengukuran dilakukan berulang kali pada subjek yang sama. Misalnya, dalam percobaan yang melibatkan pengukuran kinerja peserta pada beberapa titik waktu setelah diberikan perlakuan tertentu. Repeated measures ANOVA memungkinkan analisis perubahan dalam kelompok yang sama di berbagai waktu dan mengurangi pengaruh variansi antar individu (Gravetter & Wallnau, 2016).

Secara keseluruhan, uji ANOVA merupakan alat yang sangat berguna dalam analisis data eksperimen untuk memeriksa perbedaan antar lebih dari dua kelompok, baik itu dalam desain eksperimen sederhana maupun kompleks. Uji ini memungkinkan peneliti untuk menarik kesimpulan yang valid mengenai pengaruh faktor-faktor yang diuji terhadap variabel dependen, dan untuk memahami lebih dalam interaksi antara berbagai faktor yang mungkin berperan dalam fenomena yang diamati (Yates, Moore, & McCabe, 1999).

#### 7.4 Uji Chi-Square ( $\chi^2$ -test)

Uji Chi-Square ( $\chi^2$ -test) adalah metode statistika yang digunakan untuk menguji hubungan antara dua variabel kategorikal atau untuk mengevaluasi sejauh mana distribusi frekuensi yang diamati sesuai dengan distribusi yang diharapkan. Uji ini sangat berguna ketika data yang tersedia bersifat nominal atau ordinal, seperti data tentang jenis kelamin, status pernikahan, atau preferensi produk. Uji Chi-Square dapat membantu peneliti menentukan apakah ada asosiasi yang signifikan antara dua variabel atau apakah distribusi data sampel sesuai dengan distribusi teoritis yang diharapkan (Agresti, 2018).

Secara umum, terdapat dua jenis uji Chi-Square yang paling sering digunakan, yaitu **Uji Chi-Square untuk Goodness of Fit** dan **Uji Chi-Square untuk Independensi**. **Uji Chi-Square untuk Goodness of Fit** digunakan untuk menguji apakah distribusi frekuensi yang diamati sesuai dengan distribusi yang diharapkan. Misalnya, seorang peneliti ingin mengetahui apakah distribusi warna favorit di antara kelompok responden sesuai dengan distribusi yang diharapkan berdasarkan data populasi sebelumnya. Dalam uji ini, hipotesis nol menyatakan bahwa tidak ada perbedaan antara distribusi yang diamati dan yang diharapkan (Field, 2013).

Sementara itu, **Uji Chi-Square untuk Independensi** digunakan untuk menguji apakah dua variabel kategorikal bebas satu sama lain, yaitu apakah ada hubungan atau asosiasi antara dua variabel. Misalnya, seorang peneliti mungkin ingin mengetahui apakah ada hubungan antara jenis kelamin dan preferensi produk tertentu. Dalam hal ini, hipotesis nol menyatakan bahwa kedua variabel tersebut tidak saling memengaruhi atau tidak ada hubungan, sementara hipotesis alternatif menyatakan bahwa kedua variabel tersebut berhubungan (Agresti, 2013).

Untuk melakukan uji Chi-Square, langkah pertama adalah menghitung nilai Chi-Square menggunakan rumus berikut:

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

di mana  $O_i$  adalah frekuensi yang diamati pada kategori  $i$ , dan  $E_i$  adalah frekuensi yang diharapkan pada kategori tersebut. Nilai  $\chi^2$  yang dihitung kemudian dibandingkan dengan nilai kritis yang diambil dari distribusi Chi-Square dengan derajat kebebasan (df) yang sesuai, yang ditentukan berdasarkan jumlah kategori dan jumlah kelompok data yang dianalisis (Yates, 1984).

Asumsi utama dalam uji Chi-Square adalah bahwa data yang digunakan bersifat independen, artinya setiap individu hanya dapat masuk dalam satu kategori dan pengukuran antar individu tidak saling mempengaruhi. Selain itu, jumlah frekuensi yang diharapkan dalam setiap kategori harus cukup besar (umumnya lebih dari 5) untuk memastikan bahwa uji Chi-Square memberikan hasil yang valid. Jika frekuensi yang diharapkan terlalu kecil, alternatif lain seperti uji Fisher's Exact Test dapat digunakan (Agestri, 2018).

Uji Chi-Square adalah uji non-parametrik yang tidak memerlukan asumsi distribusi normal dalam data. Oleh karena itu, uji ini sangat berguna dalam analisis data kategorikal yang tidak memenuhi asumsi distribusi normal. Uji Chi-Square juga dapat digunakan untuk data dengan ukuran sampel besar dan dapat diadaptasi untuk berbagai jenis data kategorikal, baik dalam penelitian eksperimen maupun survei (Salkind, 2010).

Meskipun uji Chi-Square berguna dalam banyak situasi, uji ini memiliki beberapa keterbatasan. Salah satunya adalah ketidakmampuannya untuk menunjukkan kekuatan atau besar asosiasi antara dua variabel. Oleh karena itu, meskipun uji Chi-Square dapat memberitahukan kita apakah ada hubungan antara dua variabel, ia tidak memberikan informasi tentang seberapa kuat

hubungan tersebut. Untuk itu, peneliti sering menggunakan ukuran asosiasi lainnya, seperti koefisien kontingensi Cramer's  $V$ , untuk mengukur kekuatan hubungan antar variabel (Field, 2013).

## 7.5 Uji regresi linier

Uji regresi linier adalah metode statistik yang digunakan untuk menganalisis hubungan antara satu atau lebih variabel independen (prediktor) dengan variabel dependen (responden). Regresi linier bertujuan untuk memodelkan hubungan linear antara variabel-variabel tersebut, sehingga dapat digunakan untuk prediksi atau analisis sebab-akibat. Regresi linier sederhana melibatkan satu variabel independen dan satu variabel dependen, sedangkan regresi linier berganda melibatkan lebih dari satu variabel independen. Teknik ini sangat berguna dalam banyak bidang, seperti ekonomi, ilmu sosial, kedokteran, dan teknik (Montgomery et al., 2012).

Pada regresi linier sederhana, model yang dibangun adalah persamaan garis lurus yang menggambarkan hubungan antara variabel independen ( $X$ ) dan variabel dependen ( $Y$ ). Model regresi linier sederhana dapat ditulis dalam bentuk persamaan:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

di mana  $Y$  adalah variabel dependen,  $X$  adalah variabel independen,  $\beta_0$  adalah intersep (nilai  $Y$  saat  $X = 0$ ),  $\beta_1$  adalah koefisien regresi (kemiringan garis), dan  $\varepsilon$  adalah error term atau gangguan. Koefisien regresi  $\beta_1$  mengukur seberapa besar perubahan rata-rata pada  $Y$  untuk setiap unit perubahan pada  $X$ . Dengan kata lain, regresi linier sederhana digunakan untuk mengukur seberapa besar pengaruh variabel independen terhadap variabel dependen (Hair et al., 2010).

Regresi linier berganda digunakan ketika ada lebih dari satu variabel independen yang diuji untuk memprediksi variabel dependen. Model regresi linier berganda ditulis sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

di mana  $Y$  adalah variabel dependen,  $X_1, X_2, \dots, X_n$  adalah variabel independen,  $\beta_1, \beta_2, \dots, \beta_n$  adalah koefisien regresi masing-masing variabel independen, dan  $\beta_0$  adalah intersep. Regresi linier berganda memungkinkan peneliti untuk memahami bagaimana beberapa faktor berpengaruh terhadap suatu hasil secara bersamaan dan dapat digunakan untuk prediksi yang lebih akurat jika dibandingkan dengan regresi sederhana (Field, 2013).

Salah satu asumsi dasar dari regresi linier adalah adanya hubungan linear antara variabel dependen dan independen. Untuk memastikan model yang dibangun valid, regresi linier juga mengasumsikan bahwa data bersifat homoskedastis (varians residual tetap konstan untuk semua nilai  $X$ ), residual mengikuti distribusi normal, dan tidak terdapat multikolinearitas yang kuat antar variabel independen. Jika asumsi-asumsi ini tidak dipenuhi, maka hasil regresi bisa bias atau tidak dapat diinterpretasikan dengan benar (Gujarati & Porter, 2009).

Proses untuk menganalisis hasil regresi linier melibatkan perhitungan koefisien regresi dan uji signifikansi koefisien tersebut. Uji  $t$  digunakan untuk menguji apakah masing-masing koefisien regresi berbeda secara signifikan dari nol. Jika nilai  $p$  untuk koefisien regresi lebih kecil dari level signifikansi yang ditentukan (misalnya 0,05), maka koefisien tersebut dianggap signifikan dan variabel independen tersebut berpengaruh terhadap variabel dependen. Selain itu, uji  $F$

digunakan untuk menguji signifikansi keseluruhan model regresi, apakah model yang dibangun secara signifikan lebih baik daripada model tanpa variabel independen (Montgomery et al., 2012).

Koefisien determinasi ( $R^2$ ) adalah ukuran penting dalam regresi linier yang menunjukkan seberapa baik model dapat menjelaskan variasi dalam data. Nilai  $R^2$  berkisar antara 0 dan 1, di mana semakin tinggi nilai  $R^2$ , semakin baik model dalam menjelaskan variabilitas data. Namun,  $R^2$  yang tinggi tidak selalu berarti model yang baik jika ada masalah dengan asumsi atau adanya overfitting. Oleh karena itu, dalam analisis regresi, penting untuk memeriksa berbagai indikator statistik lainnya, seperti residual dan nilai  $p$ , untuk memastikan keandalan model (Field, 2013).

Selain itu, regresi linier juga dapat digunakan untuk prediksi. Setelah model yang sesuai dibangun, persamaan regresi yang diperoleh dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen yang baru. Misalnya, dalam analisis ekonomi, regresi linier berganda dapat digunakan untuk memprediksi pendapatan berdasarkan beberapa variabel independen seperti tingkat pendidikan, usia, dan pengalaman kerja. Prediksi ini dapat digunakan untuk pengambilan keputusan atau perencanaan (Keller, 2011).

## DAFTAR PUSTAKA

- Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Wiley.
- Agresti, A. (2018). *Statistical methods for the social sciences* (5th ed.). Pearson.
- Biau, D. J., & Kernéis, S. (2018). Statistics in clinical research: Part 1 - Descriptive statistics and hypothesis testing. *Clinical Orthopaedics and Related Research*, 476(8), 1553-1562. <https://doi.org/10.1097/CORR.0000000000000332>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cohen, J., Manion, L., & Morrison, K. (2013). *Research methods in education* (7th ed.). Routledge.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.
- Fisher, R. A. (1935). *The design of experiments*. Oliver and Boyd.
- Gravetter, F. J., & Wallnau, L. B. (2016). *Statistics for the behavioral sciences* (9th ed.). Cengage Learning.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Keller, G. (2011). *Statistics for management and economics* (9th ed.). Cengage Learning.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). Wiley.
- Salkind, N. J. (2010). *Encyclopedia of research design*. Sage.
- Yates, D., Moore, D. S., & McCabe, G. P. (1999). *The practice of statistics* (2nd ed.). W.H. Freeman.
- Zar, J. H. (2010). *Biostatistical analysis* (5th ed.). Pearson Education.



# UJI HIPOTESIS NON-PARAMETRIK

*Oleh Linda Prasetyaning Widayanti, S.KM., M.Kes.*

Statistik non-parametrik adalah metode analisis statistik yang tidak mengasumsikan distribusi tertentu pada data, seperti distribusi normal, sehingga cocok untuk data dengan skala ordinal atau data yang tidak memenuhi asumsi distribusi tertentu. Metode ini juga dikenal sebagai statistik bebas distribusi. Beberapa contoh statistika non parametrik adalah Mann-Whitney, Wilcoxon Sign Rank Test, Kruskall Wallis, Friedman, Chi Square, Mc Nemar, Chocran's Q, Korelasi Spearman dan Koefisien Phi (Pandriadi et al., 2023).

## 8.1 Uji Mann-Whitney

Analisis uji statistik Mann-Whitney, atau sering disebut sebagai uji Mann-Whitney U atau uji Wilcoxon Rank-Sum, adalah uji non-parametrik yang digunakan untuk membandingkan perbedaan antara dua kelompok independen. Uji ini menentukan apakah distribusi skor pada satu kelompok berbeda secara signifikan dari kelompok lainnya, tanpa mengharuskan data berdistribusi normal (Hulu & Kurniawan, 2021).

Uji Mann-Whitney digunakan ketika:

- a. Terdapat dua kelompok independen (misalnya, kelompok kontrol dan kelompok perlakuan).
- b. Data pada kedua kelompok berskala ordinal atau interval, tetapi tidak memenuhi asumsi normalitas.
- c. Jumlah sampel relatif kecil atau tidak cukup besar untuk menggunakan uji parametrik seperti uji t dua sampel

Langkah-langkah dalam uji Mann-Whitney:

- a. Menggabungkan dan merangking data
- b. Menghitung jumlah peringkat
- c. Menghitung statistic
- d. Menentukan nilai U yang lebih kecil
- e. Menentukan signifikansi

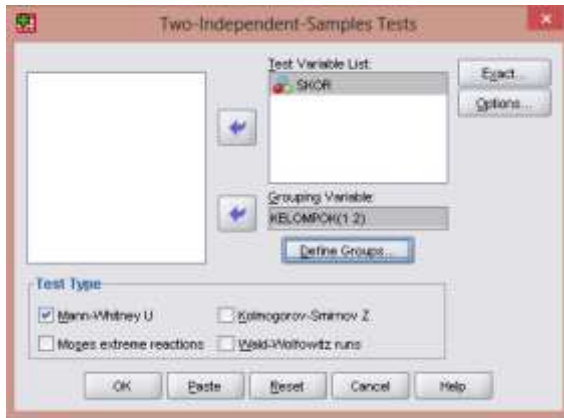
Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan dalam distribusi nilai antara kedua kelompok.
- b. Hipotesis Alternatif ( $H_1$ ): Terdapat perbedaan dalam distribusi nilai antara kedua kelompok.

Uji Mann Whitney dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test
- b. klik 2-Independent Samples
- c. Masukkan variabelnya: skor diletakkan di test variable. Kelompok diletakkan di grouping variabel
- d. Klik OK

Uji Mann-Whitney digunakan untuk menguji perbedaan distribusi antara dua kelompok independen yang tidak memenuhi asumsi normalitas. Uji ini merupakan alternatif non-parametrik untuk uji t dua sampel, yang berguna saat data berbentuk ordinal atau interval yang tidak terdistribusi normal. Hasil uji ini memberikan informasi tentang perbedaan median antar kelompok.



Gambar 1. Uji Mann Whitney

Jika nilai  $p < \alpha$ , hasilnya signifikan, yang berarti kita menolak  $H_0$  dan menganggap ada perbedaan distribusi antara kedua kelompok. Sebaliknya, jika  $p > \alpha$ , maka kita gagal menolak  $H_0$ , dan tidak ada bukti perbedaan signifikan antara kedua kelompok.

Kelebihan uji Mann-Whitney :

- a. Tidak memerlukan asumsi distribusi normal, sehingga cocok untuk data ordinal atau data yang berdistribusi non-normal.
- b. Lebih tahan terhadap outlier dibandingkan uji parametrik seperti uji t.

Kekurangan uji Mann-Whitney :

- a. Tidak dapat digunakan untuk membandingkan lebih dari dua kelompok (untuk itu, biasanya digunakan uji Kruskal-Wallis).
- b. Kurang efisien pada sampel yang sangat kecil dibandingkan dengan uji parametrik jika data sebenarnya berdistribusi normal.

## 8.2 Uji Wilcoxon Signed-Rank Test

Uji Wilcoxon Signed-Rank Test adalah uji statistik non-

parametrik yang digunakan untuk membandingkan dua sampel yang berpasangan (misalnya, sebelum dan sesudah suatu perlakuan) dan menentukan apakah terdapat perbedaan signifikan antara dua kondisi tersebut. Uji ini adalah alternatif non-parametrik untuk uji t berpasangan, digunakan ketika data tidak berdistribusi normal atau hanya berskala ordinal (Wahyudin et al., 2022).

Uji Wilcoxon Signed-Rank Test digunakan ketika:

- a. Data berasal dari dua pengukuran yang berpasangan atau berulang pada subjek yang sama (misalnya, nilai tes sebelum dan sesudah pelatihan).
- b. Data tidak memenuhi asumsi distribusi normal, yang merupakan persyaratan untuk uji t berpasangan.
- c. Pengukuran berskala ordinal atau interval, namun tidak normal

Langkah-langkah dalam uji Wilcoxon Signed-Rank Test:

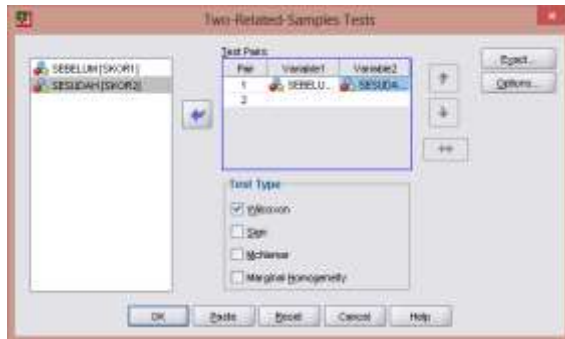
- a. Menghitung selisih D
- b. Membuang selisih 0
- c. Merangking selisih
- d. Menambahkan tanda pada peringkat
- e. Menghitung statistic Wilcoxon
- f. Menentukan signifikansi

Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan dalam distribusi antara dua pengukuran yang berpasangan.
- b. Hipotesis Alternatif ( $H_1$ ): Terdapat perbedaan dalam distribusi antara dua pengukuran yang berpasangan.

Uji Wilcoxon Signed-Rank Test dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test
- b. klik 2- Related Samples
- c. Masukkan variabelnya
- d. Klik OK



Gambar 2. Uji Wilcoxon Signed-Rank Test

Jika nilai  $p < \alpha$ , kita menolak  $H_0$  dan menyimpulkan bahwa ada perbedaan signifikan antara kedua kondisi. Jika  $p > \alpha$ , kita gagal menolak  $H_0$  dan menyimpulkan bahwa tidak ada bukti perbedaan signifikan.

Kelebihan uji Wilcoxon Signed-Rank Test:

- a. Tidak memerlukan asumsi normalitas, sehingga lebih fleksibel dibandingkan uji t berpasangan.
- b. Cocok untuk data ordinal atau data yang berdistribusi non-normal

Kekurangan uji Wilcoxon Signed-Rank Test:

- a. Hanya dapat digunakan untuk data berpasangan, tidak untuk data independen.
- b. Kurang efisien untuk sampel kecil jika data berdistribusi normal, di mana uji t berpasangan akan lebih tepat.

### 8.3 Uji Kruskal-Wallis

Uji Kruskal-Wallis adalah uji statistik non-parametrik yang digunakan untuk membandingkan tiga atau lebih kelompok independen dan menentukan apakah terdapat perbedaan signifikan di antara mereka. Uji ini merupakan alternatif non-parametrik dari

analisis varians satu arah (ANOVA) dan digunakan ketika data tidak memenuhi asumsi normalitas atau ketika data berskala ordinal (Santoso, 2010).

Uji Kruskal-Wallis digunakan ketika:

- a. Terdapat tiga atau lebih kelompok yang saling independen.
- b. Data dalam kelompok tidak memenuhi asumsi normalitas yang diperlukan untuk ANOVA.
- c. Skala data bersifat ordinal atau interval tetapi tidak berdistribusi normal.
- d. Anda ingin mengetahui apakah terdapat perbedaan yang signifikan di antara kelompok-kelompok tersebut.

Langkah-langkah dalam uji Kruskal-Wallis:

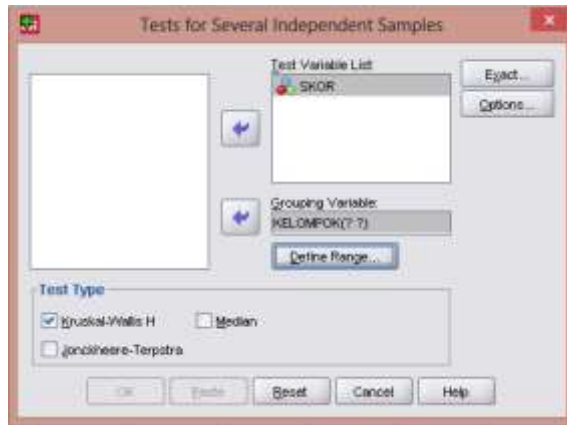
- a. Menggabungkan dan Merangking Data
- b. Menghitung Jumlah Peringkat untuk Setiap Kelompok ( $R_j$ )\*\*:
- c. Menghitung Statistik Kruskal-Wallis ( $H$ )\*\*:
- d. Menentukan signifikansi

Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan dalam distribusi nilai antara tiga atau lebih kelompok.
- b. Hipotesis Alternatif ( $H_1$ ): Setidaknya ada satu kelompok yang memiliki distribusi nilai yang berbeda dari yang lainnya.

Uji Kruskal-Wallis dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test
- b. klik Independent Samples
- c. Masukkan variabelnya
- d. Klik OK



Gambar 3. Uji Kruskal-Wallis

Jika  $p < \alpha$ , maka hasilnya signifikan, yang berarti kita menolak  $H_0$  dan menganggap ada perbedaan antara beberapa kelompok. Jika  $p > \alpha$ , maka kita gagal menolak  $H_0$  dan tidak ada bukti perbedaan signifikan di antara kelompok-kelompok.

Kelebihan uji Kruskal-Wallis:

- a. Tidak memerlukan asumsi normalitas, sehingga cocok untuk data ordinal atau data yang berdistribusi non-normal.
- b. Mampu menguji lebih dari dua kelompok independen, berbeda dengan uji Mann-Whitney yang hanya untuk dua kelompok.

Kekurangan uji Kruskal-Wallis:

- a. Jika hasilnya signifikan, uji ini tidak menunjukkan kelompok mana yang berbeda. Untuk menentukan kelompok spesifik yang berbeda, perlu dilakukan uji lanjut (post-hoc), seperti uji Dunn.
- b. Kurang efisien untuk sampel kecil jika data berdistribusi normal; dalam kasus seperti itu, ANOVA lebih sesuai.

## 8.4 Uji Friedman

Uji statistik Friedman adalah uji non-parametrik yang digunakan untuk menganalisis data berulang atau berpasangan dari tiga atau lebih kelompok, mirip dengan ANOVA berulang tetapi tanpa asumsi normalitas. Uji ini digunakan untuk menentukan apakah terdapat perbedaan signifikan antara beberapa pengukuran atau perlakuan yang diterapkan pada subjek yang sama (Prasetyo, 2021).

Uji Friedman digunakan ketika:

- a. Data terdiri dari beberapa pengukuran berulang pada subjek yang sama, atau dari data yang berpasangan pada tiga atau lebih kondisi.
- b. Asumsi normalitas tidak terpenuhi, sehingga ANOVA berulang tidak dapat digunakan.
- c. Data berada pada skala ordinal atau lebih tinggi.
- d. Contoh kasus di mana uji ini relevan adalah ketika menilai efek beberapa jenis perawatan pada pasien yang sama atau mengukur kinerja suatu kelompok pada beberapa waktu atau kondisi yang berbeda.

Langkah-langkah dalam uji Friedman:

- a. Merangking Data
- b. Menghitung Jumlah Peringkat untuk Setiap Perlakuan
- c. Menghitung Statistik Friedman
- d. Menentukan signifikansi

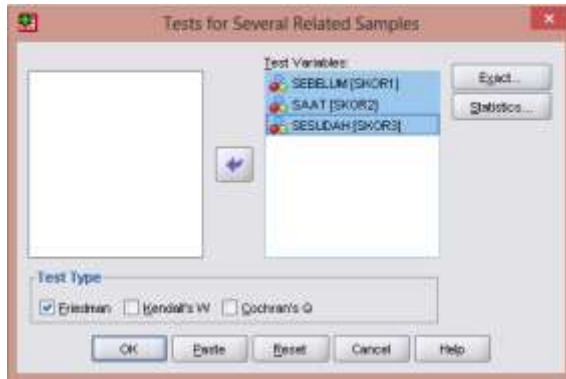
Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan dalam distribusi nilai antara beberapa perlakuan atau kondisi.
- b. Hipotesis Alternatif ( $H_1$ ): Setidaknya ada satu perlakuan atau kondisi yang memiliki distribusi nilai yang berbeda dari yang lainnya.

Uji Friedman dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test

- b. klik k-Related Samples
- c. Masukkan variabelnya
- d. Klik OK



Gambar 4. Uji Friedman

Jika nilai  $p < \alpha$ , hasilnya signifikan, yang berarti kita menolak  $H_0$  dan menganggap ada perbedaan di antara perlakuan atau kondisi. Jika  $p > \alpha$ , kita gagal menolak  $H_0$ , menunjukkan tidak ada bukti perbedaan signifikan antara perlakuan.

Kelebihan uji Friedman:

- a. Tidak memerlukan asumsi normalitas, sehingga cocok untuk data ordinal atau data yang berdistribusi non-normal.
- b. Dapat menangani lebih dari dua kondisi atau perlakuan berulang pada subjek yang sama.

Kekurangan uji Friedman:

- a. Jika hasilnya signifikan, uji ini tidak menunjukkan perlakuan atau kondisi mana yang berbeda. Untuk mengetahui perbandingan spesifik antara kondisi, perlu dilakukan uji post-hoc (misalnya, uji Wilcoxon Signed-Rank berpasangan).
- b. Tidak cocok untuk data independen; uji ini hanya untuk data berulang atau berpasangan.

## 8.5 Uji Chi Square

Uji statistik Chi-Square ( $\chi^2$ ) adalah uji non-parametrik yang digunakan untuk menentukan apakah terdapat hubungan atau asosiasi yang signifikan antara dua variabel kategori (nominal atau ordinal). Uji ini menguji perbedaan antara frekuensi yang diobservasi dan frekuensi yang diharapkan, sehingga dapat digunakan untuk mengukur hubungan antar variabel atau menguji kesesuaian distribusi data (Wintolo, 2024).

Ada dua jenis utama uji Chi-Square:

1. Chi-Square Tabel 2x2: jenis uji chi-square yang digunakan untuk menguji asosiasi antara dua variabel kategoris yang masing-masing hanya memiliki dua kategori. Tabel 2x2 memiliki dua baris dan dua kolom yang menunjukkan kombinasi kategori dari kedua variabel tersebut.
2. Chi-Square Tabel > 2x2: tabel kontingensi yang memiliki lebih dari dua kategori dalam satu atau kedua variabel yang diuji. Misalnya, tabel 3x3 menunjukkan tiga kategori untuk masing-masing variabel, sementara tabel 2x3 menunjukkan dua kategori untuk satu variabel dan tiga kategori untuk variabel lainnya.

Uji Chi-Square digunakan ketika:

- a. Data berada dalam bentuk kategori atau nominal.
- b. Ingin mengetahui apakah terdapat asosiasi antara dua variabel kategori (Test of Independence) atau apakah distribusi frekuensi sesuai dengan distribusi yang diharapkan (Goodness-of-Fit Test).
- c. Setiap kategori atau sel dalam tabel kontingensi memiliki nilai frekuensi yang mencukupi (disarankan lebih dari 5 untuk hasil yang akurat).

Langkah-langkah dalam uji Chi-Square:

- a. Membuat Tabel Kontingensi

- b. Menghitung Frekuensi yang Diharapkan
- c. Menghitung Statistik Chi-Square
- d. Menentukan signifikansi

Intepretasi hasil:

Chi Square Tabel 2x2

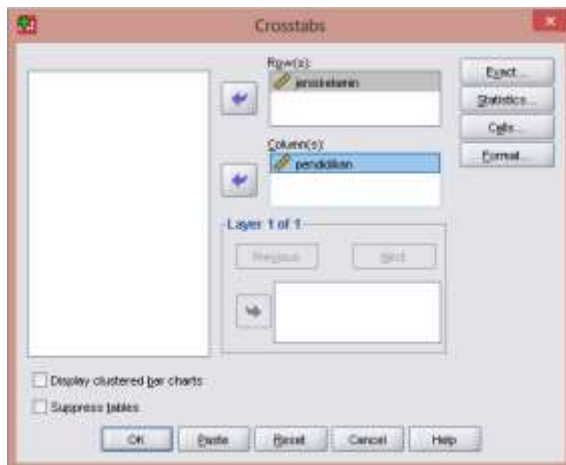
- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan antara dua kelompok.
- b. Hipotesis Alternatif ( $H_1$ ): Terdapat perbedaan antara dua kelompok.

Chi Square Tabel >2x2

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan antara lebih dari dua kelompok.
- b. Hipotesis Alternatif ( $H_1$ ): Terdapat perbedaan antara lebih dari dua kelompok.

Uji Chi-Square dalam aplikasi uji statistik:

- a. Klik Analyze, pilih descriptive statistic
- b. klik crosstab
- c. Masukkan variabelnya
- d. Klik statistic, pilih Chi Square
- e. Klik Ok



Gambar 5. Uji Chi Square

Jika nilai  $p < \alpha$ , hasilnya signifikan, yang berarti kita menolak  $H_0$  dan menyimpulkan adanya perbedaan signifikan. Sebaliknya, jika  $p > \alpha$ , kita gagal menolak  $H_0$  dan menyimpulkan tidak ada bukti perbedaan yang signifikan.

Kelebihan uji Chi Square:

- a. Tidak memerlukan asumsi distribusi normal, sehingga cocok untuk data kategori.
- b. Mudah digunakan dan diterapkan pada berbagai data kategori dalam penelitian sosial dan sains.

Kekurangan uji Chi Square:

- a. Sensitif terhadap ukuran sampel kecil; hasil bisa kurang akurat jika frekuensi sel sangat rendah.
- b. Tidak memberikan informasi tentang kekuatan atau arah hubungan, hanya menunjukkan ada atau tidaknya hubungan.

## 8.6 Uji McNemar

Uji statistik McNemar adalah uji statistik non-parametrik yang digunakan untuk menganalisis data kategori berpasangan, khususnya ketika ingin melihat perubahan atau perbedaan proporsi antara dua kondisi pada sampel yang sama. Uji ini sering digunakan untuk data biner atau dikotomis, seperti "ya/tidak" atau "sukses/gagal", dan cocok untuk studi pra-pasca atau studi yang melibatkan pengamatan berpasangan (matched pairs) (Syah, 2007).

Uji McNemar digunakan ketika:

- a. Ingin membandingkan dua pengukuran berulang atau berpasangan dalam kategori biner pada sampel yang sama.
- b. Data terdiri dari dua kategori yang berlawanan (misalnya, "sembuh/tidak sembuh", "lulus/tidak lulus").
- c. Setiap pasangan observasi berasal dari subjek yang sama yang diukur dua kali, atau dari dua kondisi yang saling berhubungan.

- d. Contoh umum adalah dalam studi medis yang ingin melihat apakah suatu pengobatan meningkatkan kondisi pasien dari “belum sembuh” menjadi “sembuh”.

Langkah-langkah dalam uji McNemar:

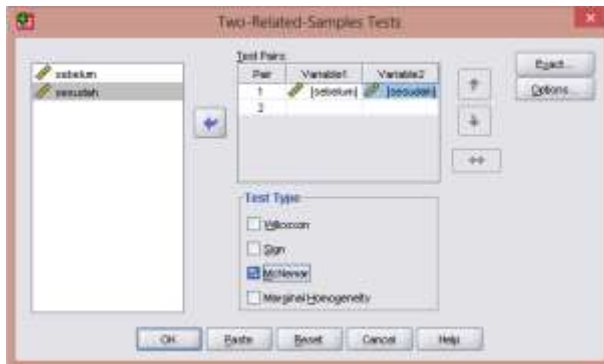
- a. Menghitung Statistik McNemar
- b. Menentukan signifikansi
- c. Menggunakan Koreksi Yates untuk Ukuran Sampel Kecil (Opsional)

Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan proporsi antara kondisi sebelum (A) atau sesudah (B).
- b. Hipotesis Alternatif ( $H_1$ ): ada perbedaan proporsi antara kondisi sebelum (A) atau sesudah (B).

Uji Mc Nemar dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test
- b. klik 2-Related Samples
- c. Masukkan variabelnya
- d. Klik OK



Gambar 6. Uji Mc Nemar

Jika nilai  $p < \alpha$ , kita menolak  $H_0$  dan menyimpulkan bahwa terdapat perbedaan signifikan antara kedua kondisi. Jika  $p > \alpha$ , maka kita gagal

menolak  $H_0$ , yang berarti tidak ada bukti perbedaan yang signifikan.

Kelebihan uji McNemar:

- a. Ideal untuk data kategori berpasangan dan mudah digunakan untuk studi pra-pasca pada data biner.
- b. Tidak memerlukan asumsi distribusi normal, sehingga cocok untuk data nominal.

Kekurangan uji McNemar:

- a. Hanya dapat digunakan untuk data kategori biner dan berpasangan.
- b. Sensitif terhadap ukuran sampel kecil, meskipun koreksi Yates dapat membantu.

## 8.7 Uji Cochran's Q

Uji statistik Cochran's Q adalah uji statistik non-parametrik yang digunakan untuk menganalisis data kategori biner (misalnya, "ya/tidak" atau "sukses/gagal") pada tiga atau lebih kondisi berpasangan. Uji ini merupakan perpanjangan dari uji McNemar yang dapat digunakan untuk membandingkan lebih dari dua kelompok yang berhubungan, seperti dalam studi yang melibatkan pengukuran berulang pada subjek yang sama di beberapa kondisi atau waktu (Usman, 2006).

Uji Cochran's Q digunakan ketika:

- a. Data terdiri dari variabel kategori biner pada tiga atau lebih kondisi yang diukur pada subjek yang sama.
- b. Ingin mengetahui apakah terdapat perbedaan signifikan dalam proporsi respons biner antar beberapa kondisi.
- c. Setiap pasangan observasi berasal dari subjek yang sama yang diukur dalam beberapa kondisi yang saling berhubungan.
- d. Contoh penerapan uji ini adalah untuk melihat apakah ada perbedaan efektivitas dari tiga atau lebih jenis pengobatan yang diberikan kepada pasien yang sama dalam suatu

penelitian klinis

Langkah-langkah dalam uji Cochran's Q:

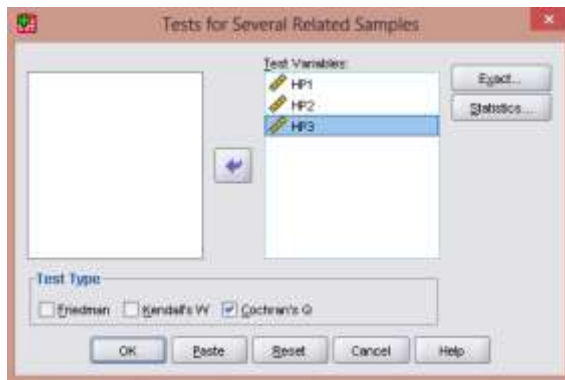
- a. Menghitung Statistik Cochran's Q
- b. Menentukan signifikansi
- c. Uji Lanjut (Post-Hoc) Jika Diperlukan

Intepretasi hasil:

- a. Hipotesis Nol ( $H_0$ ): Tidak ada perbedaan proporsi antara beberapa kondisi
- b. Hipotesis Alternatif ( $H_1$ ): ada perbedaan proporsi antara beberapa kondisi

Uji Cochran's Q dalam aplikasi uji statistik:

- a. Klik Analyze, pilih nonparametric test
- b. klik k-Related Samples
- c. Masukkan variabelnya
- d. Klik OK



Gambar 7. Uji Cochran's Q

Jika nilai  $p < \alpha$ , kita menolak  $H_0$ , yang berarti ada perbedaan signifikan dalam proporsi respons antara kondisi-kondisi tersebut. Jika  $p > \alpha$ , kita gagal menolak  $H_0$ , yang menunjukkan tidak ada bukti perbedaan signifikan antar kondisi.

Kelebihan uji Cochran's Q:

- a. Dapat digunakan untuk data biner dengan lebih dari dua kondisi yang berhubungan.
- b. Tidak memerlukan asumsi normalitas, sehingga cocok untuk data kategori.

Kekurangan uji Cochran's Q:

- a. Hanya dapat digunakan untuk data kategori biner.
- b. Jika hasilnya signifikan, tidak menunjukkan kondisi mana yang berbeda tanpa uji lanjut.

## 8.8 Uji Korelasi Spearman

Uji korelasi Spearman adalah uji statistik non-parametrik yang digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel ordinal atau dua variabel dengan distribusi yang tidak normal. Uji ini menghitung korelasi berdasarkan peringkat data, sehingga tidak terpengaruh oleh distribusi data atau adanya outlier, yang membuatnya cocok untuk data non-linear atau data yang tidak memenuhi asumsi normalitas (Akbar et al., 2023).

Uji korelasi Spearman digunakan ketika:

- a. Ingin mengetahui kekuatan dan arah hubungan antara dua variabel ordinal, interval, atau rasio.
- b. Data tidak berdistribusi normal atau mengandung outlier yang dapat mengganggu analisis korelasi Pearson.
- c. Hubungan antara variabel diduga tidak linear tetapi monoton (cenderung bergerak ke arah yang sama secara konsisten).

Langkah-langkah dalam uji korelasi Spearman:

- a. Mengubah Data ke dalam Peringkat
- b. Menghitung Selisih Peringkat
- c. Menghitung Koefisien Korelasi Spearman
- d. Menentukan signifikansi

Intepretasi hasil:

- a. Nilai  $r$  berkisar antara  $-1$  dan  $1$ .  
Jika  $r = +1$  berarti terdapat hubungan positif sempurna (monoton) antara dua variabel (jika satu meningkat, yang lain juga meningkat).
- b. Jika  $r = -1$  berarti terdapat hubungan negatif sempurna (monoton) antara dua variabel (jika satu meningkat, yang lain menurun).
- c. Jika  $r = 0$  berarti tidak ada hubungan monoton yang signifikan antara kedua variabel.
- d. Hipotesis Nol ( $H_0$ ): Tidak ada hubungan antara dua variabel.
- c. Hipotesis Alternatif ( $H_1$ ): ada hubungan antara dua variabel.

Uji Korelasi Spearman dalam aplikasi uji statistik:

- a. Klik Analyze, pilih correlate
- b. klik bivariate
- c. Masukkan variabelnya
- d. Centang Spearman
- e. Klik OK



Gambar 8. Uji Korelasi Spearman

Jika nilai  $p < \alpha$  (biasanya 0,05), kita menolak  $H_0$  dan menyimpulkan bahwa ada korelasi signifikan antara kedua variabel.

Kelebihan uji korelasi Spearman:

- a. Tidak memerlukan asumsi normalitas, sehingga cocok untuk data ordinal atau data dengan distribusi yang tidak normal.
- b. Tidak sensitif terhadap outlier, karena bekerja dengan peringkat data.
- c. Cocok untuk hubungan monoton, baik linear maupun non-linear.

Kekurangan uji korelasi Spearman:

- a. Tidak memberikan informasi tentang ketepatan hubungan linier (hanya menunjukkan apakah terdapat hubungan monoton).
- b. Kurang cocok untuk data nominal atau data kategorikal yang tidak dapat diurutkan.

### 8.9 Uji Koefisien Phi

Uji koefisien phi (atau koefisien  $\phi$ ) adalah ukuran korelasi yang digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel biner (dikotomis). Uji ini dapat digunakan dalam tabel kontingensi 2x2 untuk menentukan hubungan antara variabel biner, seperti “ya/tidak,” “lulus/gagal,” atau “positif/negatif”. Koefisien phi merupakan kasus khusus dari korelasi Pearson untuk data yang terdiri dari dua variabel dikotomis (Yolanda et al., 2024).

Uji koefisien phi digunakan ketika:

- a. Kedua variabel bersifat dikotomis (memiliki dua kategori).
- b. Ingin mengetahui kekuatan hubungan antara dua variabel biner.
- c. Data disajikan dalam tabel kontingensi 2x2.

Langkah-langkah dalam uji koefisien phi:

- a. Membuat tabel kontingensi

- b. Menghitung nilai Chi Square
- c. Menghitung Koefisien Phi
- d. Menentukan signifikansi

Intepretasi hasil:

- e. Nilai  $r$  berkisar antara  $-1$  dan  $1$ .  
Jika  $r = +1$  berarti terdapat hubungan positif sempurna (monoton) antara dua variabel (jika satu meningkat, yang lain juga meningkat).
- f. Jika  $r = -1$  berarti terdapat hubungan negatif sempurna (monoton) antara dua variabel (jika satu meningkat, yang lain menurun).
- g. Jika  $r = 0$  berarti tidak ada hubungan monoton yang signifikan antara kedua variabel.
- h. Hipotesis Nol ( $H_0$ ): Tidak ada hubungan antara dua variabel.
- d. Hipotesis Alternatif ( $H_1$ ): ada hubungan antara dua variabel.

Uji Korelasi Spearman dalam aplikasi uji statistik:

- a. Klik Analyze, pilih descriptive statistic
- b. klik crosstab
- c. Masukkan variabelnya
- d. Centang Phi and Cramer's V
- e. Klik OK



Gambar 9. Uji Koefisien Phi

Untuk menguji apakah koefisien phi signifikan, dapat menggunakan uji Chi-Square pada tabel kontingensi 2x2. Jika nilai Chi-Square signifikan ( $p < \alpha$ , biasanya 0,05), maka hubungan antara kedua variabel dianggap signifikan.

Kelebihan uji koefisien phi:

- a. Sederhana dan mudah digunakan untuk data biner dalam tabel kontingensi 2x2.
- b. Tidak memerlukan asumsi normalitas, karena koefisien phi adalah ukuran korelasi non-parametrik.

Kekurangan uji koefisien phi:

- a. Terbatas pada tabel 2x2 dan data yang bersifat dikotomis.
- b. Nilai phi cenderung menurun ketika jumlah kategori lebih dari dua atau ketika tabel lebih besar dari 2x2.

## DAFTAR PUSTAKA

- Akbar, R., Sukmawati, U. S., & Katsirin, K. (2023). Analisis Data Penelitian Kuantitatif: Pengujian Hipotesis Asosiatif Korelasi. *Jurnal Pelita Nusantara*, 1(3), 430–448.
- Hulu, V. T., & Kurniawan, R. (2021). *Memahami dengan Mudah Statistik Nonparametrik Bidang Kesehatan: Penerapan Software SPSS dan STATCAL*. Prenada Media.
- Pandriadi, P., Van Harling, V. N., Wahab, A., Vaulina, S., Sutjiningtyas, S., Ningsih, E. K., Setyono, B. D. H., Rizqi, V., Harisuddin, M. I., & Gaffar, S. (2023). *STATISTIKA DASAR*. Penerbit Widina.
- Prasetyo, B. (2021). Instrumen penelitian. *Metode Penelitian Pendekatan Kuantitatif*, 131.
- Santoso, S. (2010). *Statistik nonparametrik*. Elex Media Komputindo.
- Syah, D. (2007). *Pengantar statistik pendidikan*.
- Usman, H. (2006). *Pengantar statistik*.
- Wahyudin, M. P., Rismaningsih, F., Si, S. P., Ul'fah Hernaeny, M. P., Anggraeni, E. F., Si, S. P., Astuti, F., Saka, B. G. M., Eka Hendrayani, S. E., & Oktavia, Y. (2022). *Pengantar Statistika 2*. Media Sains Indonesia.
- Wintolo, D. (2024). *Pengantar Statistik dan Geostatistik*. UGM PRESS.
- Yolanda, F., Egianto, F., Armita, F., Wahyuni, L. A., Cahyani, R., Rahayu, S., & Saputri, T. (2024). Studi Literatur: Korelasi Bivariat Menggunakan Uji Korelasi Koefisien Kontingensi. *Jurnal Pendidikan Tambusai*, 8(2).  
<https://jptam.org/index.php/jptam/article/view/15041>

# REGRESI LINEAR

*Oleh Dr. Hwihanus, SE., MM., CMA*

## 9.1. Pengertian dan Persamaan Regresi Linear

Metode regresi linier pertama kali diperkenalkan oleh Sir Francis Galton pada abad ke-19 dalam konteks mempelajari korelasi tinggi badan antar generasi yang selanjutnya Karl Pearson (abad 20) mengembangkan konsep kuadrat terkecil (*least squares*) yang memperkirakan korelasi dan memperdalam pemahaman regresi sebagai teknik statistik yang disebut analisis data. Sejak saat itu, regresi linier telah berkembang pesat dan kini menjadi alat analisis utama dalam berbagai bidang, baik untuk tujuan akademis maupun praktis (Padilah & Adam, 2019). Seiring dengan perkembangan komputasi dan teknologi, penggunaan regresi semakin mudah melalui berbagai perangkat lunak seperti SPSS (Hajarisman & Herlina, 2022), R, Python (Sholeh et al., 2022), dan Excel.

Regresi linear adalah metode statistik yang digunakan untuk memodelkan hubungan antara variabel terikat (*dependent variable*) dan satu atau lebih variabel bebas (*independent variables*) dengan tujuan (1) memprediksi nilai variabel terikat berdasarkan nilai variabel bebas, (2) menentukan seberapa besar perubahan pada variabel terikat ketika variabel bebas berubah, (3) digunakan dalam riset ekonomi, pemasaran, kesehatan dan kebijakan publik dalam mengavaluasi efektivitas program atau kebijakan, (4) membantu peneliti memahami interaksi dan pengaruh antara beberapa variabel. Pada regresi linear, hubungan ini diasumsikan berbentuk linier, artinya

perubahan pada variabel bebas akan menghasilkan perubahan proporsional pada variabel terikat (Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, 2011).

Regresi linier dapat didefinisikan sebagai teknik analisis data yang memprediksi nilai data yang tidak diketahui dengan menggunakan nilai data lain yang terkait dan diketahui. Model ini menggambarkan hubungan antara variabel dengan persamaan linear, contohnya:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

## 9.2. Penerapan Bidang Regresi Linear

Penerapan regresi linear telah meluas ke berbagai bidang dengan kemampuan yang bersifat fleksibel dalam memodelkan dan memprediksi hubungan antara variabel. Regresi linear dapat membuat prediksi, menganalisis pengaruh variabel, dan memahami pola-pola yang terdapat dalam data. Beberapa bidang yang menggunakan regresi linear adalah

1. Ekonomi dan keuangan, regresi linear digunakan untuk memodelkan hubungan antar variabel ekonomi dan keuangan yang kompleks.
2. Manajemen bisnis dan pemasaran untuk mengoptimalkan strategi pemasaran, meramalkan penjualan dan mendukung pengambilan keputusan.
3. Ilmu Sosial dan Psikologi yang memahami hubungan sosial dan perilaku individu atau kesehatan mental.
4. Kesehatan dan Kedokteran secara klinis maupun epidemiologi.
5. Teknik dan industri dalam pengendalian kualitas, perencanaan produksi dan pemeliharaan prediktif
6. Lingkungan dan geografi terutama hubungan variabel iklim, polusi udara, pengelolaan sumber daya dan perubahan lahan.
7. Pendidikan untuk mengevaluasi prestasi akademik siswa, program pendidikan dan kepuasan siswa terhadap fasilitas atau lingkungan pendidikan

### 9.3. Regresi Linear Sederhana

Regresi linear sederhana adalah metode statistik yang digunakan untuk memahami dan memodelkan hubungan antara dua variabel yaitu satu variabel independen (prediktor) dan satu variabel dependen (respons) sebagai garis lurus yang paling mendekati data yang disebut "garis terbaik" dengan tujuan utama untuk memahami bagaimana perubahan variabel independen mempengaruhi variabel dependen, serta untuk melakukan prediksi terhadap nilai variabel dependen berdasarkan nilai variabel independen.

Konsep dasar model regresi linear sederhana dinyatakan dalam persamaan

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Di mana:

Y adalah variabel dependen (yang diprediksi).

X adalah variabel independen (yang mempengaruhi).

$\beta_0$  adalah konstanta (nilai Y ketika  $X = 0$ ).

$\beta_1$  adalah koefisien regresi yang menunjukkan seberapa besar perubahan Y ketika X berubah satu unit

$\epsilon$  adalah error atau residu yang menunjukkan penyimpangan antara nilai prediksi dan nilai aktua

### 9.4. Interpretasi Koefisien

Interpretasi koefisien dalam regresi linear membantu memahami hubungan kuantitatif antara variabel independen dan variabel dependen yang terdiri dari intersep (a) dan kemiringan (b)

1. Intersep (a) atau konstanta adalah nilai Y ketika  $X=0$ , persamaan regresi linear yaitu  $Y=a+bX$  sehingga intersep dilambangkan dengan a yang menunjukkan titik regresi memotong sumbu Y.

Intersep menunjukkan titik awal atau nilai dasar variabel dependen saat variabel independen bernilai nol.

Contoh: Jika model regresi menunjukkan bahwa  $Y=10+5X$

Intersep adalah 10, yang berarti ketika  $X=0$ , nilai Y adalah 10. Jika Y adalah penjualan bulanan dan X adalah jumlah iklan yang ditempatkan, maka intersep menunjukkan bahwa jika tidak ada iklan (saat  $X=0$ ), perusahaan masih akan memiliki penjualan sebesar 10 unit karena faktor lain di luar pengeluaran iklan.

2. Kemiringan ( $b$ ) atau koefisien regresi, koefisien yang menunjukkan perubahan pada variabel dependen  $Y$  untuk setiap perubahan satu unit pada variabel independen  $X$ . Dalam persamaan  $Y=a+bX$ , kemiringan dilambangkan dengan  $b$ . Jika  $b>0$ , maka variabel dependen  $Y$  meningkat ketika variabel independen  $X$  meningkat dan menunjukkan hubungan positif antara  $X$  dan  $Y$ . Jika  $b<0$ , maka variabel dependen  $Y$  menurun ketika variabel independen  $X$  meningkat, menunjukkan hubungan negatif antara  $X$  dan  $Y$ . Besaran nilai  $b$  menggambarkan seberapa besar perubahan pada  $Y$  untuk setiap perubahan satu unit pada  $X$ .

### 9.5. Langkah-langkah Melakukan Regresi Linear Sederhana

Melakukan regresi linear sederhana melibatkan serangkaian langkah untuk menganalisis hubungan antara dua variabel yaitu variabel independen dan variabel dependen yang mencakup pemahaman data, pembuatan model, hingga interpretasi hasil. Berikut adalah langkah-langkah untuk melakukan analisis regresi linear sederhana:

1. Pahami data dan tujuan analisis untuk memastikan bahwa regresi linear adalah metode yang sesuai dengan adanya hubungan antara variabel independen ( $X$ ) dan variabel dependen ( $Y$ ) yang bersifat numerik atau kontinu dan membuat prediksi.
2. Kumpulkan dan persiapkan data dalam regresi harus lengkap dan memenuhi beberapa kondisi tidak ada nilai yang hilang (*missing values*) dan tidak mengandung outlier yang ekstrem, karena regresi linear sangat sensitif terhadap outliers.
3. Plot data untuk melihat hubungan linier untuk melakukan visualisasi data dengan scatter plot untuk melihat apakah terdapat pola linier antara variabel independen ( $X$ ) dan variabel dependen ( $Y$ ). Jika terdapat pola yang cenderung linier, maka regresi linear adalah pilihan yang tepat atau jika pola terlihat melengkung atau tidak linier, metode regresi lain mungkin lebih cocok.
4. Hitung Koefisien *Regresi Ordinary Least Squares* (OLS), untuk model regresi  $Y=\beta_0+\beta_1X+\epsilon$  dimana koefisien ini bisa dihitung secara manual dengan rumus OLS atau menggunakan perangkat lunak statistik seperti Python, R, atau Excel.

5. Buat Model Regresi sederhana  $\hat{Y} = \beta_0 + \beta_1 X$  untuk memprediksi nilai Y berdasarkan nilai X.
6. Evaluasi model regresi diperlukan untuk melihat seberapa baik model tersebut dalam memprediksi data melalui (a) koefisien determinasi ( $R^2$ ) yang menunjukkan proporsi variansi dalam Y yang dapat dijelaskan oleh X. Nilai  $R^2$  berkisar antara 0 hingga 1, maka semakin dekat ke 1 maka semakin baik model tersebut menjelaskan variasi dalam data. (b) standard Error yang menunjukkan deviasi standar dari residuals, atau seberapa jauh nilai prediksi dari nilai aktual. Dan (c) Signifikansi koefisien merupakan uji apakah koefisien  $\beta_1$  signifikan secara statistik dalam memprediksi Y dengan menggunakan uji t untuk menentukan apakah p-value di bawah ambang batas (umumnya 0,05).
7. Periksa asumsi-asumsi model regresi dengan memastikan model regresi memenuhi linearitas, independensi residuals, homoskedastisitas dan normalitas residuals
8. Interpretasikan hasil regresi dengan interpretasi koefisien  $\beta_1$  (slope) untuk memahami seberapa besar pengaruh variabel independen terhadap variabel dependen,  $R^2$  untuk mengetahui seberapa besar variasi yang dijelaskan oleh model dan jika p-value koefisien rendah (biasanya  $< 0,05$ ), maka variabel independen memiliki pengaruh yang signifikan terhadap variabel dependen.
9. Gunakan model untuk prediksi nilai Y berdasarkan nilai X yang baru untuk mendapatkan prediksi  $\hat{Y}$ .
10. Buat kesimpulan dan rekomendasi untuk menjawab tujuan analisis awal. Jika analisis bertujuan untuk menunjukkan pengaruh, kesimpulan harus menjelaskan hubungan yang ditemukan. Jika tujuannya untuk prediksi, kesimpulan dapat mencakup keandalan model untuk memprediksi variabel dependen.

#### 9.6. Metode Ordinary Least Squares (OLS)

Metode Ordinary Least Squares (OLS) adalah teknik yang umum digunakan dalam regresi linear untuk menemukan estimasi parameter atau koefisien model regresi linear. Tujuan utama metode OLS adalah

meminimalkan jumlah kuadrat selisih antara nilai yang diobservasi dan nilai yang diprediksi model. OLS adalah metode mendasar dan banyak digunakan dan bersifat sederhana dan matematis dalam menghasilkan estimasi yang terbaik dalam banyak kasus, terutama ketika asumsi-asumsi model terpenuhi.

Metode Ordinary Least Squares (OLS) dimulai dari konsep dasar hingga asumsi dan penerapannya dalam analisis regresi dengan tahapan sebagai berikut

1. Dasar-Dasar Metode OLS dinyatakan dalam persamaan

$$Y = \beta_0 + \beta_1 X + \epsilon$$

Tujuan OLS adalah menentukan nilai  $\beta_0$  dan  $\beta_1$  sehingga garis regresi yang dihasilkan memberikan estimasi terbaik terhadap hubungan antara X dan Y.

2. Konsep Kuadrat Terkecil (*Least Squares*), OLS bekerja dengan cara meminimalkan jumlah kuadrat dari selisih antara nilai Y aktual dan nilai Y yang diprediksi oleh model yang menghasilkan selisih antara nilai aktual dan nilai prediksi yang disebut "error" atau "residual." OLS meminimalkan *sum of squared residuals* (jumlah kuadrat error) yang dinyatakan sebagai:

$$SS_{\text{residuals}} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

di mana:

$Y_i$  adalah nilai aktual dari variabel dependen untuk observasi ke- $i$ ,  
 $\hat{Y}_i$  adalah nilai prediksi dari variabel dependen untuk observasi ke- $i$ ,

$n$  adalah jumlah total observasi.

Meminimalkan  $SS_{\text{residual}}$ , OLS memastikan bahwa garis regresi yang dihasilkan memiliki jarak total yang paling kecil antara titik data aktual dan garis regresi.

3. Menghitung Koefisien OLS Regresi Linear Sederhana

Dalam kasus regresi linear sederhana (satu variabel prediktor), koefisien  $\beta_0$  dan  $\beta_1$  dapat dihitung dengan rumus:

$$\beta_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X}$$

di mana:

$\bar{X}$  adalah rata-rata dari variabel independen X,

$\bar{Y}$  adalah rata-rata dari variabel dependen Y.

4. Untuk hasil yang valid dan interpretasi yang tepat, OLS bergantung pada beberapa asumsi dasar agar estimasi OLS menjadi *Best Linear Unbiased Estimator* (BLUE), yang berarti estimasi terbaik, linear, dan tak bias berupa
  - Linearitas yang menyatakan hubungan antara variabel dependen dan variabel independen harus linier.
  - Independensi error harus independen satu sama lain yang berarti tidak ada pola dalam urutan data, atau error harus acak.
  - Homoskedastisitas dimana variansi error harus konstan untuk semua nilai variabel independen. Jika tidak akan terjadi heteroskedastisitas, yang mengindikasikan bahwa variansi error berubah-ubah pada level variabel prediktor yang berbeda.
  - Error atau residual harus berdistribusi normal. Asumsi ini sangat penting untuk analisis inferensial, seperti pengujian hipotesis dan pembuatan interval kepercayaan.
5. Keunggulan Metode OLS adalah
  - Sederhana dan mudah diimplementasikan pada software statistik dan pemrograman, seperti Python, R, dan Excel, mendukung OLS.
  - Estimasi yang tidak bias
  - Penggunaan yang luas dalam berbagai model statistik dan diterapkan di berbagai bidang penelitian dan industri.
6. Limitasi Metode OLS terutama ketika asumsi-asumsi dasar tidak terpenuhi:
  - Sensitif terhadap outliers dimana OLS cenderung terpengaruh oleh data yang ekstrem atau outliers dan bersifat bias.

- Multikolinieritas (Regresi Berganda) adanya korelasi yang tinggi antara variabel independen dan menyebabkan estimasi koefisien menjadi tidak stabil.
  - Heteroskedastisitas, adanya varians dari parameter estimasi tidak minimum, sehingga mengurangi keandalan estimasi.
  - Ketidaksiharian data nonlinear pada hubungan antara variabel tidak linier, regresi linear dan metode OLS tidak akan menghasilkan estimasi yang baik.
7. Penerapan OLS dalam analisis data bertujuan untuk
- Dalam aplikasi bisnis dan keuangan, untuk memprediksi hasil ke depan berdasarkan data historis.
  - Dalam ilmu sosial, untuk menganalisis hubungan kausal antara variabel independen dan dependen.
  - OLS memungkinkan analisis lebih lanjut terhadap kecocokan model (fit) dan validasi asumsi, melalui statistik seperti  $R^2$  (koefisien determinasi) yang mampu menjelaskan varians variabel dependen.

### 9.7. Regresi Linear Berganda

Regresi linear berganda adalah teknik atau alat statistik yang digunakan untuk menganalisis hubungan antara satu variabel dependen (Y) dan dua atau lebih variabel independen ( $X_1, X_2, \dots, X_n$ ). Regresi linear berganda adalah metode yang efektif untuk memahami hubungan antara satu variabel dependen dan beberapa variabel independen, serta membuat prediksi. Model regresi ini merupakan pengembangan dari regresi linear sederhana yang hanya melibatkan satu variabel independen.

Regresi linear berganda sering digunakan dalam berbagai bidang, termasuk ekonomi, pemasaran, ilmu sosial, dan sains untuk memahami faktor-faktor yang memengaruhi variabel tertentu atau untuk membuat prediksi dengan persamaan model di tulis sebagai berikut

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon$$

Di mana:

- Y adalah variabel dependen (target) yang ingin diprediksi.
- $X_1, X_2, \dots, X_n$  adalah variabel-variabel independen yang digunakan untuk memprediksi Y.
- $\beta_0$  adalah intercept atau konstanta, yang menunjukkan nilai rata-rata dari Y ketika semua X bernilai nol.
- $\beta_1, \beta_2, \dots, \beta_n$  adalah koefisien regresi untuk masing-masing variabel independen, yang menunjukkan seberapa besar perubahan Y untuk setiap perubahan satu unit pada X, dengan asumsi variabel lain konstan.
- $\epsilon$  adalah error atau residual, yang merupakan selisih antara nilai aktual dan nilai yang diprediksi oleh model.

### 9.8. Tujuan dan Fungsi Regresi Linear Berganda

Regresi linear berganda digunakan untuk memahami bagaimana setiap variabel independen yang memengaruhi variabel dependen dan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen.

Adapun tujuan regresi linear berganda adalah untuk

1. Memahami dan memprediksi hubungan antara variabel independen (*predictors*) dan variabel dependen (*outcome*). Model ini bisa digunakan untuk memprediksi nilai variabel dependen berdasarkan kombinasi beberapa variabel independen untuk memperkirakan hasil atau performa di masa depan berdasarkan data historis.
2. Regresi linear berganda membantu kita memahami sejauh mana masing-masing variabel independen memengaruhi variabel dependen.
3. Regresi linear berganda, bisa mengevaluasi signifikansi masing-masing variabel independen dan untuk mengetahui apakah variabel tersebut benar-benar berpengaruh pada variabel dependen atau tidak.
4. Dalam beberapa kasus, regresi linear berganda dapat mengidentifikasi interaksi antara variabel independen untuk melihat apakah kombinasi dua atau lebih variabel independen memiliki efek yang lebih besar atau lebih kecil pada variabel

dependen.

5. Hasil analisis regresi linear berganda dapat membantu dalam pengambilan keputusan yang lebih baik.

### **9.9. Asumsi Regresi Linear Berganda**

Beberapa asumsi penting yang harus dipenuhi dalam regresi linear berganda adalah untuk memastikan hasil analisis valid dan interpretasi yang akurat yang digunakan sebagai fondasi dasar agar model regresi dapat digunakan secara efektif untuk menjelaskan hubungan antarvariabel atau untuk memprediksi variabel dependen dengan baik (Prasetyo, 2022) (Sugito & Marliyana, 2021).

1. Linearitas, hubungan antara variabel independen dan variabel dependen harus linear. Jika hubungan ini tidak linear, hasil prediksi dan interpretasi bisa jadi kurang akurat
2. Independensi, observasi dalam dataset harus independen. Jika data terkait satu sama lain (misalnya data berulang atau berseri waktu), bisa terjadi bias.
3. Homoskedastisitas, varians error (residual) harus konstan untuk semua nilai variabel independen. Jika varians berubah-ubah (heteroskedastisitas), model bisa menghasilkan hasil yang bias.
4. Normalitas, error atau residual sebaiknya mengikuti distribusi normal. Asumsi ini terutama penting untuk validitas pengujian statistik, seperti uji t dan uji F.
5. Tidak ada multikolinearitas, variabel independen sebaiknya tidak memiliki korelasi yang sangat tinggi antara satu sama lain. Jika ada korelasi tinggi (multikolinearitas), model bisa mengalami kesulitan dalam menentukan kontribusi setiap variabel secara independen.
6. Tidak ada outlier atau pengaruh observasi yang kuat
7. Tidak ada variabel yang penting yang ditinggalkan (omitted variable bias) yang menyatakan bahwa semua variabel yang memengaruhi variabel dependen telah dimasukkan dalam model regresi. Jika ada variabel penting yang ditinggalkan, model bisa menjadi bias dan koefisien yang dihasilkan tidak akan akurat.

Secara ringkas Asumsi dalam Regresi Linear Berganda:

Asumsi	Deskripsi	Cara Memeriksa
Linearitas	Hubungan antara variabel independen dan dependen harus linear.	Scatter plot, plot residual
Independensi	Setiap observasi dan error harus independen satu sama lain.	Uji Durbin-Watson (untuk data deret waktu)
Homoskedastisitas	Varians error harus konstan di seluruh nilai variabel independen.	Plot residual terhadap prediksi
Normalitas Error	Error atau residual harus terdistribusi normal.	Q-Q plot, uji Shapiro-Wilk, uji Kolmogorov-Smirnov
Tidak Ada Multikolinearitas	Variabel independen tidak boleh memiliki korelasi tinggi satu sama lain.	Variance Inflation Factor (VIF), korelasi antarvariabel
Tidak Ada Outlier atau Observasi Berpengaruh Kuat	Tidak ada observasi yang mengganggu hasil regresi.	Plot residual, leverage, Cook's Distance
Tidak Ada Omitted Variable Bias	Semua variabel yang relevan sudah termasuk dalam model.	Analisis teoritis, eksplorasi data mendalam

### 9.10. Koefisien Regresi dan Interpretasi

Koefisien regresi linear berganda memainkan peranan penting dalam memahami pengaruh variabel independen terhadap variabel dependen dan memberikan informasi bagaimana perubahan pada satu variabel independen tertentu akan memengaruhi variabel dependen, dengan asumsi bahwa variabel independen lainnya tetap konstan.

- a. Makna dari Koefisien Regresi ( $\beta_i$ ), menunjukkan seberapa besar pengaruh dari variabel independen terhadap variabel dependen. Setiap koefisien  $\beta_i$  dalam persamaan regresi mewakili rata-rata perubahan dalam variabel dependen Y untuk setiap kenaikan satu unit pada variabel independen  $X_i$ , dengan asumsi bahwa semua variabel independen lainnya tetap konstan.
- b. Signifikansi statistik koefisien regresi ditentukan untuk apakah

koefisien tersebut signifikan secara statistik, yang berarti bahwa perubahan dalam variabel independen benar-benar memiliki pengaruh terhadap variabel dependen. Signifikansi ini biasanya diuji dengan uji  $t$  untuk setiap koefisien regresi, dengan hipotesis Hipotesis nol ( $H_0$ ), koefisien regresi sama dengan nol ( $\beta_i=0$ ) yang berarti bahwa variabel independen  $X_i$  tidak memiliki pengaruh signifikan terhadap variabel dependen  $Y$ .

Hipotesis alternatif ( $H_1$ ), koefisien regresi tidak sama dengan nol ( $\beta_i \neq 0$ ), yang berarti bahwa variabel independen  $X_i$  memiliki pengaruh signifikan terhadap variabel dependen  $Y$ .

Nilai  $p$ -value dari uji  $t$  ini digunakan untuk menguji signifikansi. Jika  $p$ -value lebih kecil dari tingkat signifikansi tertentu (misalnya, 0,05), kita dapat menolak hipotesis nol dan menyimpulkan bahwa variabel independen tersebut berpengaruh signifikan terhadap variabel dependen.

- c. Tanda positif dan negatif pada koefisien, bila tanda positif koefisien menunjukkan hubungan positif antara variabel independen dan variabel dependen. Artinya, jika nilai variabel independen meningkat, nilai variabel dependen juga akan meningkat. Sedangkan koefisien negatif, menunjukkan hubungan negatif antara variabel independen dan variabel dependen. Artinya, jika nilai variabel independen meningkat, nilai variabel dependen akan menurun.
- d. Interpretasi nilai koefisien yang kecil atau mendekati nol hal ini menunjukkan variabel berarti bahwa variabel independen tersebut memiliki pengaruh yang sangat kecil atau bahkan tidak berpengaruh terhadap variabel dependen. Dalam hal ini, variabel tersebut mungkin bisa diabaikan atau dikeluarkan dari model untuk menyederhanakan analisis.
- e. Dalam regresi linear berganda, variabel independen bisa memiliki unit yang berbeda-beda, untuk memudahkan perbandingan pengaruh setiap variabel independen terhadap variabel dependen yang disebut koefisien standar adalah koefisien regresi yang dihasilkan setelah variabel diubah ke skala yang sama, biasanya dalam satuan standar deviasi. Kegunaan koefisien standar adalah

membandingkan seberapa besar pengaruh relatif dari masing-masing variabel independen dalam model. Koefisien standar yang lebih besar menunjukkan pengaruh yang lebih kuat terhadap variabel dependen dibandingkan dengan koefisien standar yang lebih kecil.

- f. Dalam beberapa kasus, efek dari satu variabel independen terhadap variabel dependen bisa berbeda tergantung pada nilai variabel independen lainnya, hal ini disebut efek interaksi dan bisa dimasukkan dalam model regresi linear berganda sebagai variabel interaksi.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 (X_1 \times X_2) + \epsilon$$

### 9.11. Analisis Signifikansi

Analisis signifikan adalah proses untuk menentukan apakah hubungan atau perbedaan yang diamati dalam data cukup kuat untuk dianggap **nyata** dan **bukan hanya kebetulan**. Signifikansi memberikan dasar untuk membuat keputusan atau kesimpulan berdasarkan data, terutama dalam pengujian hipotesis (Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, 2011) (Mona et al., 2015).

Signifikansi statistik menentukan apakah hasil atau pola yang diamati dalam sampel data cukup kuat untuk digeneralisasikan ke populasi, menggunakan pengujian hipotesis yang membandingkan hipotesis nol (tidak ada efek atau hubungan yang signifikan) dengan hipotesis alternatif (ada efek atau perbedaan yang signifikan dan berhubungan antar variabel). Uji signifikansi menghasilkan p-value, yang menunjukkan probabilitas bahwa hasil yang diperoleh setidaknya se-ekstrem hasil aktual akan terjadi jika hipotesis nol benar.

Tingkat signifikansi ( $\alpha$ ) adalah ambang batas yang digunakan untuk menilai apakah hasil uji statistik signifikan. Nilai  $\alpha$  umumnya dipilih 0,05 atau 0,01, yang berarti menerima peluang 5% atau 1% untuk salah menolak hipotesis nol (menggambil risiko kesalahan tipe I). Jika  $p\text{-value} \leq \alpha$  menunjukkan hipotesis nol ditolak, dan hasil dianggap signifikan secara statistik dan jika  $p\text{-value} > \alpha$  berarti Hipotesis nol tidak ditolak, dan hasil dianggap tidak signifikan secara statistik.

### Langkah-Langkah dalam Uji Signifikansi

1. Merumuskan Hipotesis, menentukan hipotesis nol ( $H_0$ ) dan hipotesis alternatif ( $H_1$ ).
2. Menentukan Tingkat Signifikansi ( $\alpha$ ), pilih tingkat signifikansi, biasanya 0,05 atau 0,01, tergantung pada seberapa besar risiko kesalahan yang dapat diterima.
3. Menghitung Statistik Uji yang sesuai dan berdasarkan jenis data dan tujuan analisis meliputi:  
Uji t digunakan untuk uji hipotesis rata-rata atau koefisien regresi.  
Uji z digunakan untuk pengujian proporsi atau rata-rata dalam sampel besar.  
ANOVA (Analysis of Variance) digunakan untuk membandingkan rata-rata lebih dari dua kelompok.  
Uji Chi-square digunakan untuk menguji asosiasi antara variabel kategorik.  
Statistik uji yang dihitung ini akan dibandingkan dengan distribusi teoretis untuk menentukan p-value.
4. Menginterpretasikan P-value yang menunjukkan probabilitas hasil yang diamati yang akan terjadi. Hasilnya dibandingkan dengan tingkat signifikansi yang telah ditetapkan. Jika  $p\text{-value} \leq \alpha$  maka hipotesis nol ditolak, yang menunjukkan bahwa hasilnya signifikan dan jika  $p\text{-value} > \alpha$ , hipotesis nol tidak ditolak, yang menunjukkan bahwa hasilnya tidak signifikan.

Interpretasi hasil signifikan sangat penting dalam analisis statistik. Hasil signifikan berarti bahwa ada cukup bukti untuk menyatakan bahwa efek yang diamati bukan karena kebetulan, namun hal ini tidak menunjukkan besarnya efek atau hubungan tersebut secara praktis. Oleh karena itu, interpretasi hasil analisis harus mempertimbangkan baik signifikansi statistik maupun signifikansi praktis atau ukuran efek (*effect size*).

Ukuran efek (*effect size*) menunjukkan besarnya hubungan atau perbedaan dan dapat memberikan konteks tambahan pada signifikansi statistik meliputi:

- Cohen's d digunakan untuk mengukur ukuran efek antara dua kelompok.

- R-squared dalam regresi linear yang menunjukkan proporsi varians variabel dependen yang dijelaskan oleh variabel independen.
- Eta-squared dalam ANOVA, menunjukkan proporsi total varians yang dijelaskan oleh varians antar kelompok.

Signifikansi dalam regresi linear berganda, signifikansi setiap koefisien regresi diperiksa untuk menentukan apakah variabel independen tersebut memiliki pengaruh yang signifikan terhadap variabel dependen yang meliputi:

- Menghitung Uji t untuk Koefisien dimana setiap koefisien diuji dengan uji t untuk menentukan apakah ada perbedaan signifikan dari nol.
- Membandingkan P-value dengan  $\alpha$ , jika p-value untuk koefisien tersebut di bawah tingkat signifikansi, maka koefisien tersebut dianggap signifikan.
- Melihat R-squared dan Adjusted R-squared untuk menilai seberapa baik model secara keseluruhan menjelaskan variabel dependen. R-squared yang lebih tinggi menunjukkan model yang lebih baik, tetapi harus diperhatikan adjusted R-squared untuk menghindari overfitting pada variabel tambahan yang tidak signifikan.

### 9.12. Koefisien Determinasi ( $R^2$ )

Koefisien Determinasi ( $R^2$ ) adalah ukuran statistik dalam analisis regresi yang menunjukkan seberapa baik model regresi menjelaskan variabilitas variabel dependen (Sugito & Marliyana, 2021). Nilai  $R^2$  berkisar dari 0 hingga 1, dan dapat diinterpretasikan sebagai proporsi varians dalam variabel dependen yang dijelaskan oleh variabel independen dalam model dengan rumusan

$$R^2 = 1 - \frac{SSE}{SST}$$

di mana:

SSE (Sum of Squares Error) adalah jumlah kuadrat sisa (error) antara nilai aktual dan nilai prediksi model.

SST (Total Sum of Squares) adalah total variasi dari rata-rata variabel dependen.

Rumus di atas menunjukkan bahwa  $R^2$  dihitung dengan melihat

seberapa besar pengurangan varians dalam variabel dependen yang dapat dijelaskan oleh model regresi dibandingkan dengan varians total.

#### Interpretasi Nilai $R^2$

$R^2 = 0$  menunjukkan model regresi tidak dapat menjelaskan variasi variabel dependen sama sekali. Dengan kata lain, prediksi model sama sekali tidak lebih baik daripada menggunakan rata-rata variabel dependen.

$R^2 = 1$  menunjukkan model regresi menjelaskan 100% variasi variabel dependen, artinya nilai prediksi model sangat sesuai dengan data aktual tanpa error.

$0 < R^2 < 1$  menyatakan bahwa nilai antara 0 dan 1 menunjukkan sebagian variabilitas dalam variabel dependen dijelaskan oleh variabel independen dalam model. Misalnya,  $R^2$  sebesar 0,75 berarti bahwa 75% variansi dalam variabel dependen dapat dijelaskan oleh variabel independen yang digunakan dalam model.

#### $R^2$ dan Kebaikan dari Model (*Goodness of Fit*)

Koefisien determinasi menunjukkan kebaikan dari model atau *goodness of fit* bila nilai  $R^2$  yang lebih tinggi berarti menunjukkan bahwa model memiliki kemampuan lebih baik untuk menjelaskan variabilitas variabel dependen, tetapi nilai  $R^2$  yang tinggi tidak selalu berarti model tersebut lebih baik atau lebih valid dalam semua situasi (Hajarisman & Herlina, 2022).

Overfitting terjadi apabila pada model yang terlalu rumit (misalnya, menambahkan terlalu banyak variabel independen), nilai  $R^2$  mungkin tinggi, tetapi model tersebut mungkin tidak dapat memprediksi data baru dengan baik, hal ini dikenal sebagai overfitting, di mana model terlalu cocok dengan data sampel tetapi tidak generik untuk data lain.

Adjusted  $R^2$  digunakan untuk mengatasi overfitting, kegunaan Adjusted  $R^2$  akan menyesuaikan nilai  $R^2$  dengan mempertimbangkan jumlah variabel independen dan ukuran sampel. Adjusted  $R^2$  akan lebih rendah dari  $R^2$  jika penambahan variabel baru tidak

meningkatkan kemampuan prediksi model.

Keterbatasan Koefisien Determinasi ( $R^2$ )

$R^2$  adalah ukuran yang berguna untuk melihat seberapa baik model menjelaskan variabel dependen, tetapi juga memiliki keterbatasan:

- a. Tidak menunjukkan signifikansi statistik,  $R^2$  hanya menunjukkan proporsi varians yang dijelaskan, bukan apakah variabel independen signifikan secara statistik.
- b. Tidak mengukur kausalitas dimana  $R^2$  hanya menggambarkan kekuatan hubungan, bukan hubungan sebab-akibat antara variabel independen dan dependen.
- c. Tidak selalu cocok untuk model nonlinear, dalam regresi linear,  $R^2$  adalah metrik yang baik, tetapi dalam model nonlinear,  $R^2$  mungkin tidak menunjukkan fit model dengan baik.

Perbedaan antara  $R^2$  dan Adjusted  $R^2$

Adjusted  $R^2$  adalah versi modifikasi dari  $R^2$  yang disesuaikan dengan jumlah variabel independen dalam model. Rumus Adjusted  $R^2$  adalah:

$$\text{Adjusted } R^2 = 1 - \left( \frac{(1 - R^2)(n - 1)}{n - k - 1} \right)$$

di mana:

n adalah jumlah sampel,

k adalah jumlah variabel independen.

Adjusted  $R^2$  memberikan penalti terhadap penambahan variabel independen yang tidak memberikan kontribusi signifikan, sehingga menjadi ukuran yang lebih baik untuk membandingkan model dengan jumlah variabel independen yang berbeda.

### 9.13. Aplikasi Regresi Linear dalam SPSS

Berikut ini perusahaan PT Sindikat akan melakukan analisa dengan beberapa item yang ada kaitan terhadap penjualan. Berapa besar pengaruhnya? Apakah data ini sesuai dengan kaidah statistik analisa regresi berganda. Berikan penjelasan lengkap apa yang Anda

peroleh sesuai pembelajaran diatas.

	Penjualan	Biaya Iklan	Biaya Jual	Biaya Produksi	Harga Barang	Volume Barang
1	5.860	558	334	5.556	64	100
2	8.772	1.010	220	3.775	121	38
3	8.092	1.181	528	3.034	127	117
4	5.466	975	366	5.253	136	30
5	9.426	1.199	473	4.955	111	137
6	5.200	830	641	5.327	109	62
7	9.875	1.139	251	5.931	51	43
8	8.104	1.005	467	3.197	50	50
9	9.598	847	494	4.930	97	77
10	7.454	972	585	5.799	61	49
11	9.642	1.485	541	3.268	135	120
12	8.120	1.284	250	3.369	106	113
13	5.016	603	352	3.635	78	53
14	6.243	1.428	385	4.129	127	31
15	7.205	892	262	4.603	141	79
16	6.446	614	367	6.112	70	110
17	9.649	1.111	242	4.153	133	42
18	8.901	1.299	581	3.337	140	129
19	6.013	1.073	403	5.778	150	33
20	6.240	996	769	6.392	131	74
21	11.850	1.367	301	7.117	61	52
22	9.725	1.206	560	3.836	60	60
23	11.518	1.016	593	5.916	116	92
24	12.053	1.856	676	4.085	169	150
25	10.150	1.605	313	4.211	133	141
26	6.270	754	440	4.544	98	66
27	7.804	1.785	481	5.161	159	39
28	9.006	1.115	328	5.754	176	99

29	8.058	767	459	7.640	88	138
30	12.062	1.389	303	5.191	166	52

## DAFTAR PUSTAKA

- Hajarisman, N., & Herlina, M. (2022). *Buku Ajar Analisis Regresi dan Aplikasinya menggunakan SPSS Program Studi Statistika* (Issue March). <https://doi.org/10.13140/RG.2.2.14988.80008>
- Mona, M., Kekenusa, J., & Prang, J. (2015). Penggunaan Regresi Linear Berganda untuk Menganalisis Pendapatan Petani Kelapa. Studi Kasus: Petani Kelapa Di Desa Beo, Kecamatan Beo Kabupaten Talaud. *D'CARTESIAN*, 4(2), 196. <https://doi.org/10.35799/dc.4.2.2015.9211>
- Padilah, T. N., & Adam, R. I. (2019). Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang. *FIBONACCI: Jurnal Pendidikan Matematika Dan Matematika*, 5(2), 117. <https://doi.org/10.24853/fbc.5.2.117-128>
- Prasetyo, R. A. (2022). Analisis Regresi Linear Berganda Untuk Melihat Faktor Yang Berpengaruh Terhadap Kemiskinan di Provinsi Sumatera Barat. *Journal of Mathematics UNP*, 7(2), 62. <https://doi.org/10.24036/unpjomath.v7i2.12777>
- Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, K. Y. (2011). *Probability & Statistics for Engineers & Scientists. In Sustainability (Switzerland)* (NINTH EDIT, Vol. 11, Issue 1). Prentice Hall. [https://faculty.ksu.edu.sa/sites/default/files/probability\\_and\\_statistics\\_for\\_engineers\\_and\\_scientisst.pdf](https://faculty.ksu.edu.sa/sites/default/files/probability_and_statistics_for_engineers_and_scientisst.pdf)
- Sholeh, M., Suraya, S., & Andayati, D. (2022). Machine Linear untuk Analisis Regresi Linier Biaya Asuransi Kesehatan dengan Menggunakan Python Jupyter Notebook. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 8(1), 20. <https://doi.org/10.26418/jp.v8i1.48822>
- Sugito, S., & Marliyana, S. D. (2021). Uji Performa Spektrofotometer Serapan Atom Thermo Ice 3000 Terhadap Logam Pb Menggunakan CRM 500 dan CRM 697 di UPT Laboratorium Terpadu UNS. *Indonesian Journal of Laboratory*, 4(2), 67. <https://doi.org/10.22146/ijl.v4i2.67438>



# ANALISIS FAKTOR

*Oleh Dr. Ir. La Ode Muhamad Nurrahmad Arsyad, S.T., M.T.*

## **10.1. Latar Belakang Penggunaan SPSS dalam Analisis Faktor**

SPSS (*Statistical Package for the Social Sciences*) adalah salah satu perangkat lunak statistik yang sangat populer, khususnya dalam penelitian sosial dan ilmu pengetahuan lainnya. Sejak dikembangkan, SPSS telah menjadi alat yang penting bagi para peneliti untuk melakukan berbagai analisis data yang kompleks secara efisien. Kemampuannya dalam mengolah data, menghasilkan statistik deskriptif, dan melakukan berbagai uji statistik menjadikan SPSS pilihan utama bagi banyak akademisi, praktisi, dan profesional yang ingin menganalisis data secara mendalam. Penggunaan SPSS dalam analisis faktor sangat membantu dalam menyederhanakan struktur data, mengidentifikasi variabel-variabel utama, dan mengekstraksi faktor yang signifikan yang sering kali sulit dicapai tanpa bantuan perangkat lunak ini.

Analisis faktor sendiri adalah teknik statistik yang digunakan untuk mengidentifikasi hubungan antar variabel dan mengelompokkan variabel-variabel yang memiliki kesamaan menjadi satu atau beberapa faktor yang lebih mudah dianalisis. Teknik ini biasanya diterapkan ketika peneliti memiliki banyak variabel dan ingin menyederhanakannya menjadi beberapa faktor yang lebih ringkas namun tetap representatif. Dalam penelitian akademik maupun praktis, analisis faktor digunakan untuk memahami struktur internal dari data yang kompleks dan membantu mengungkapkan pola

tersembunyi. SPSS mempermudah proses ini dengan menyediakan fitur dan alat analisis yang memungkinkan pengguna menjalankan analisis faktor dengan lebih cepat dan akurat.

SPSS memberikan fleksibilitas dalam memilih metode ekstraksi faktor dan metode rotasi, yang merupakan aspek penting dalam analisis faktor. Metode ekstraksi, seperti *principal component analysis* (PCA) atau *common factor analysis*, memungkinkan peneliti memilih pendekatan yang paling sesuai dengan tujuan penelitian mereka. Sementara itu, metode rotasi, baik orthogonal maupun oblique, membantu menginterpretasikan faktor yang dihasilkan agar lebih bermakna secara substantif. Kemampuan SPSS untuk mengakomodasi berbagai metode ini memungkinkan peneliti menyesuaikan analisis sesuai dengan kebutuhan spesifik mereka, yang pada akhirnya meningkatkan kualitas interpretasi dan hasil penelitian.

Di samping itu, SPSS memungkinkan pengguna untuk mengevaluasi kualitas data sebelum melakukan analisis faktor, misalnya dengan menyediakan uji statistik seperti KMO (Kaiser-Meyer-Olkin) dan uji Bartlett. Keduanya merupakan langkah penting untuk memastikan bahwa data yang digunakan sesuai dengan asumsi-asumsi analisis faktor. Uji KMO mengukur kecukupan sampel, sedangkan uji Bartlett menilai apakah variabel-variabel memiliki korelasi yang cukup tinggi untuk dapat diolah dalam analisis faktor. Jika hasil dari kedua uji ini memadai, maka peneliti dapat yakin bahwa data mereka memenuhi syarat untuk dilanjutkan dalam analisis faktor yang lebih mendalam.

Perkembangan teknologi dan meningkatnya ketersediaan data besar (*big data*) dalam berbagai bidang menjadikan SPSS semakin relevan dalam era digital ini. Banyak organisasi dan institusi yang kini mengandalkan data untuk mendukung keputusan dan strategi mereka. Dalam konteks ini, analisis faktor dengan SPSS menjadi semakin penting karena dapat membantu menyaring informasi yang benar-benar relevan dari jumlah variabel yang banyak. Sebagai contoh, dalam riset pasar, analisis faktor sering kali digunakan untuk memahami preferensi konsumen atau pola perilaku yang mendasar, sehingga perusahaan dapat merancang produk dan layanan yang lebih

sesuai dengan kebutuhan pelanggan mereka.

Selain untuk akademisi, SPSS juga menjadi alat penting bagi praktisi di berbagai bidang, seperti pemasaran, keuangan, dan sumber daya manusia. Para profesional ini menggunakan analisis faktor untuk mendapatkan wawasan strategis yang berguna dalam membuat keputusan berbasis data. Sebagai contoh, dalam departemen sumber daya manusia, analisis faktor dapat digunakan untuk mengidentifikasi faktor-faktor utama yang mempengaruhi kepuasan karyawan atau produktivitas di tempat kerja. Dengan hasil analisis faktor yang disederhanakan, manajer HR dapat lebih mudah menentukan intervensi yang tepat untuk meningkatkan kinerja organisasi.

Implementasi SPSS dalam analisis faktor juga mendukung penelitian yang lebih efisien, terutama dalam konteks akademik. Banyak mahasiswa dan dosen yang mengandalkan SPSS untuk menyelesaikan penelitian mereka dengan lebih cepat dan efektif. Dengan antarmuka yang relatif mudah dipahami dan prosedur yang jelas, SPSS memungkinkan pengguna pemula hingga tingkat lanjut untuk menggunakan alat-alat analisis statistik tanpa memerlukan kemampuan pemrograman yang mendalam. Hal ini menjadikan SPSS sebagai solusi ideal untuk berbagai kalangan peneliti, mulai dari mahasiswa yang baru belajar statistik hingga peneliti profesional yang memerlukan hasil cepat dan akurat.

## **10.2. Dasar-Dasar Teori Analisis Faktor**

Analisis faktor adalah teknik statistik yang sangat penting dalam riset multivariat, yang berfungsi untuk mengidentifikasi hubungan antara sejumlah besar variabel dan mengelompokkan variabel yang memiliki kesamaan ke dalam beberapa faktor yang lebih sederhana. Konsep dasarnya adalah bahwa sejumlah variabel yang tampaknya berbeda dapat diwakili oleh faktor-faktor yang mendasarinya, yang membantu menyederhanakan interpretasi data yang kompleks. Sebagai contoh, dalam penelitian psikologi, beberapa variabel yang mengukur perilaku sosial dapat dikelompokkan menjadi satu faktor yang disebut “sosialisasi.” Proses ini memudahkan peneliti untuk mengidentifikasi struktur data yang sebenarnya, menjadikannya alat

penting dalam berbagai bidang penelitian.

Teknik analisis faktor memiliki dua jenis utama: Analisis Faktor Eksploratori (*Exploratory Factor Analysis* atau *EFA*) dan Analisis Faktor Konfirmatori (*Confirmatory Factor Analysis* atau *CFA*). *EFA* digunakan ketika peneliti belum mengetahui struktur hubungan antar variabel dan ingin menemukan pola atau faktor yang muncul dari data secara alami. Sebaliknya, *CFA* digunakan ketika peneliti sudah memiliki hipotesis tertentu tentang faktor-faktor tersebut dan ingin mengonfirmasi apakah data mendukung hipotesis tersebut. Kedua teknik ini memiliki peran penting dalam penelitian, di mana *EFA* lebih digunakan dalam tahap eksplorasi awal, sedangkan *CFA* sering digunakan untuk verifikasi model pada tahap lanjutan.

Asumsi dalam analisis faktor adalah aspek penting yang harus diperhatikan untuk memastikan validitas dan reliabilitas hasil. Beberapa asumsi utama dalam analisis faktor meliputi linearitas hubungan antar variabel, normalitas data, dan kecukupan korelasi antar variabel. Kecukupan korelasi biasanya diuji melalui *KMO (Kaiser-Meyer-Olkin)* dan uji Bartlett. *KMO* mengukur tingkat kecukupan sampel, sedangkan uji Bartlett memeriksa apakah terdapat korelasi yang cukup kuat antar variabel. Tanpa memenuhi asumsi-asumsi ini, hasil analisis faktor bisa menjadi tidak valid atau tidak akurat, sehingga sangat penting untuk memastikan bahwa data sesuai dengan asumsi yang dibutuhkan sebelum melanjutkan analisis.

Salah satu metode yang sering digunakan dalam analisis faktor adalah Metode Ekstraksi. Metode ekstraksi utama yang banyak digunakan dalam *SPSS* termasuk *Principal Component Analysis (PCA)* dan *Common Factor Analysis (CFA)*. *PCA* sering digunakan sebagai teknik penyederhanaan data karena mampu mengidentifikasi komponen utama yang mewakili variabel secara keseluruhan. Di sisi lain, *Common Factor Analysis* lebih berfokus pada faktor laten yang mendasari variabel yang diamati. Pemilihan metode ekstraksi tergantung pada tujuan penelitian, dengan *PCA* lebih cocok untuk analisis eksplorasi umum, sementara *CFA* lebih cocok untuk mengungkapkan struktur laten yang lebih dalam.

Selain metode ekstraksi, Metode Rotasi juga sangat penting

dalam analisis faktor. Rotasi faktor bertujuan untuk mempermudah interpretasi faktor dengan mengatur ulang matriks faktor sehingga setiap variabel memiliki bobot yang signifikan pada satu faktor tertentu. Metode rotasi dibedakan menjadi dua jenis, yaitu rotasi *orthogonal* (seperti *varimax*) dan rotasi *oblique* (seperti *promax*). Rotasi *orthogonal* digunakan ketika faktor-faktor dianggap tidak berkorelasi, sementara rotasi *oblique* memungkinkan adanya korelasi antar faktor. Pemilihan jenis rotasi ini tergantung pada asumsi peneliti tentang hubungan antar faktor.

Interpretasi faktor adalah langkah krusial dalam analisis faktor, yang bertujuan untuk memberikan makna atau label pada faktor yang dihasilkan berdasarkan variabel yang memiliki bobot tinggi dalam faktor tersebut. Proses interpretasi ini memerlukan pemahaman yang mendalam tentang konteks penelitian, karena label yang diberikan pada faktor harus mencerminkan makna yang relevan bagi penelitian tersebut. Sebagai contoh, dalam analisis faktor yang menghasilkan variabel-variabel seperti "keterbukaan" dan "ekstrovertsi" dalam satu faktor, peneliti mungkin menamai faktor tersebut sebagai "kepribadian sosial." Proses ini penting untuk menghasilkan pemahaman yang komprehensif dan valid tentang hasil analisis.

Analisis faktor juga memiliki beberapa ukuran kecocokan data yang membantu dalam mengevaluasi sejauh mana model yang dihasilkan cocok dengan data. Selain KMO dan uji Bartlett, beberapa indeks lain yang dapat digunakan adalah nilai eigen dan total varians yang dijelaskan. Nilai eigen menunjukkan kontribusi faktor terhadap varians total dalam data, sedangkan total varians yang dijelaskan mengukur persentase varians keseluruhan yang diwakili oleh faktor-faktor yang dipilih. Semakin tinggi persentase varians yang dijelaskan, semakin baik faktor-faktor tersebut dalam mewakili data.

Implementasi analisis faktor di SPSS dimulai dengan memasukkan data yang telah disiapkan ke dalam perangkat lunak, memilih metode ekstraksi dan rotasi, serta menjalankan analisis. Setelah hasil diperoleh, pengguna dapat menginterpretasi tabel output yang menyajikan informasi seperti *communalities*, nilai eigen, dan bobot faktor. *Communalities* menunjukkan seberapa besar

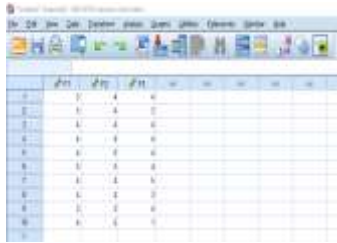
variabel tersebut dapat dijelaskan oleh faktor yang dihasilkan, sedangkan nilai eigen membantu menentukan jumlah faktor yang tepat. SPSS menyediakan berbagai opsi dan fitur untuk menyesuaikan analisis faktor sesuai dengan kebutuhan penelitian.

### 10.3. Penggunaan SPSS dalam Melakukan Analisis Faktor

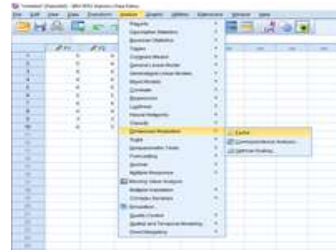
SPSS adalah perangkat lunak statistik yang umum digunakan dalam analisis data, terutama untuk riset dibidang ilmu sosial. Dalam konteks analisis faktor, SPSS menyediakan berbagai alat dan fitur yang memungkinkan pengguna untuk melakukan analisis faktor secara sistematis, mulai dari pengolahan data hingga interpretasi hasil. Bab ini bertujuan memberikan panduan langkah demi langkah untuk menggunakan SPSS dalam melakukan analisis faktor. Proses ini dimulai dari persiapan data, pemeriksaan asumsi, memilih metode analisis faktor, hingga memahami hasil yang ditampilkan dalam tabel output SPSS.

#### A. Langkah-langkah Analisis Faktor dengan SPSS

- 1) Buka lembar kerja baru SPSS, lalu klik *Variable View* untuk mengisi *Name*, *Decimals*, *Labels* dan *Measure*.
- 2) Selanjutnya klik *Data View*, kemudian masukan data *Service Quality* (F1), *Persepsi Harga* (F2), dan *Slot Time* (F3) (Gambar 7.1)
- 3) Dari menu SPSS klik *Analyze*⇒*Dimension Reduction*⇒*Factor* (Gambar 7.2)
- 4) Muncul kotak dialog “*Factor Analysis*” selanjutnya, masukkan semua variabel ke kotak *Variables*, lalu klik *Descriptives*. (Gambar 7.3)
- 5) Muncul kotak dialog “*Factor Analysis: Descriptives*” kemudian berikan tanda centang (✓) pada *initial solution*, *KMO and Bartlett’s test of sphericity*, dan *Anti-image*, lalu klik *Continue*. (Gambar 7.4)



Gambar 7.1



Gambar 7.2

- 6) Selanjutnya klik *extraction*, maka muncul kotak dialog “*Factor Analysis: Extraction*” berikan tanda centang (v) pada *Unrotated factor solution* dan *scree plot*, lalu klik *Continue*. (Gambar 7.5)



Gambar 7.3



Gambar 7.4

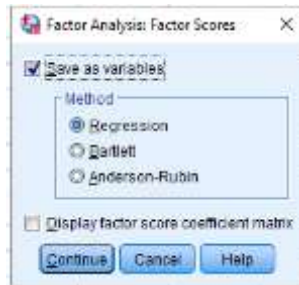
- 7) Berikutnya klik *Rotation*, maka muncul kotak dialog “*Factor Analysis: Rotation*” lalu aktifkan *Varimax* dan beri tanda centang (v) pada *Rotated solution*, lalu klik *Continue*. (Gambar 7.6)
- 8) Klik *Scores*, maka muncul kotak dialog “*Factor Analysis: Factor Score*” lalu aktifkan *Save as Variable*, lalu klik *Continue*. (Gambar 7.7)



Gambar 7.5



Gambar 7.6



Gambar 7.7

9) Klik “OK”, maka akan muncul *output Factor Analysis SPSS*

**KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.593
Bartlett's Test of Sphericity	Approx. Chi-Square	6.992
	df	3
	Sig.	.007

Gambar 7.8

Catatan:

Tabel output KMO and Bartlett's Test berguna untuk mengetahui kelayakan suatu variabel, apakah dapat diproses lebih lanjut menggunakan teknik analisis faktor ini atau tidak, yaitu dengan melihat nilai KMO MSA (Kaiser-Meyer-Olkin *Measure of Sampling Adequacy*). Jika nilai KMO MSA lebih besar dari 0,50, maka teknik analisis faktor dapat dilanjutkan, begitu pula sebaliknya, berdasarkan output diatas diketahui nilai KMO MSA sebesar 0,593 > 0,50 dan nilai Bartlett's Test of Sphericity (Sig) 0,007 < 0,05, maka analisis faktor dalam kasus ini dapat dilanjutkan karena telah memenuhi persyaratan pertama.

**Anti-image Matrices**

		Service Quality	Harga	Slot Time
Anti-Image Covariance	Service Quality	.704	-.334	-.197
	Harga	-.334	.745	-.069
	Slot Time	-.197	-.069	.887
Anti-Image Correlation	Service Quality	.565 <sup>a</sup>	-.481	-.250
	Harga	-.481	.582 <sup>a</sup>	-.095
	Slot Time	-.250	-.095	.888 <sup>a</sup>

a. Measures of Sampling Adequacy(MSA)

Gambar 7.9

*Anti-image Matrices* berguna untuk mengetahui dan

menentukan variabel mana saja yang layak dipakai dalam analisis faktor. Perhatikan bagian *Anti-image Correlation*, pada tabel diatas terdapat kode huruf (a) yang artinya tanda untuk Measure of Sampling Adequacy (MSA). Diketahui nilai MSA dari masing-masing variabel yang diteliti adalah sebagai berikut:

- a) *Service quality* sebesar 0,565
- b) Harga 0,580
- c) Slot Time 0,698

Persyaratan yang harus dipenuhi dalam analisis faktor adalah nilai MSA > 0,5. Dari hasil diatas diketahui bahwa nilai MSA untuk setiap variabel yang diteliti adalah > 0,5, maka persyaratan kedua dalam analisis faktor ini pun terpenuhi.

**Communalities**

	Initial	Extraction
Service Quality	1,000	,698
Harga	1,000	,618
Slot Time	1,000	,403

Extraction Method: Principal Component Analysis.

Gambar 7.10

Tabel Communalities ini menunjukkan nilai variabel yang diteliti apakah mampu untuk menjelaskan faktor atau tidak. Variabel dianggap mampu menjelaskan faktor jika nilai Extraction lebih besar dari 0,5, dengan demikian dapat disimpulkan bahwa semua variabel dapat dipakai untuk menjelaskan faktor.

**Total Variance Explained**

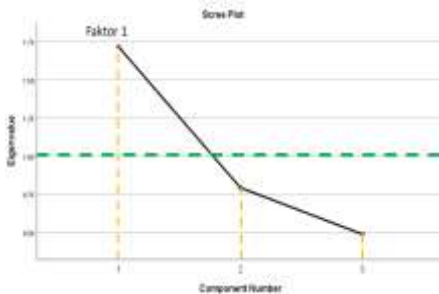
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1,719	57,286	57,286	1,719	57,286	57,286
2	,793	26,446	83,732			
3	,488	16,268	100,000			

Extraction Method: Principal Component Analysis.

Gambar 7.11

Tabel 7.11 diatas menunjukkan nilai masing-masing variabel yang dianalisis. Dalam kasus ini ada 3 variabel berarti ada 3 *Component* yang dianalisis. Ada dua macam analisis untuk

menjelaskan suatu varian, yaitu *Initial Eigenvalues* dan *Extraction Sums of Squared Loadings*. Pada varian Initial Eigenvalues menunjukkan faktor yang terbentuk. Apabila semua faktor yang dijumlahkan menunjukkan nilai variabel ( $1,719 + 0,793 + 0,488 = 3$  variabel). Sedangkan pada bagian Extraction Sums of Squared Loadings menunjukkan jumlah variasi atau banyaknya faktor yang dapat terbentuk, pada hasil diatas terdapat 1 (satu) variasi faktor, yaitu 1,719. Tabel 7.11 diatas pada bagian Initial Eigenvalues, maka ada 1 (satu) faktor yang dapat terbentuk dari 3 variabel yang dianalisis. Dimana syarat untuk menjadi sebuah faktor, maka nilai Eigenvalues harus lebih besar 1. Nilai Eigenvalues Component sebesar  $1,719 > 1$ , maka menjadi faktor dan mampu menjelaskan 57,286% variasi. Nilai total Component 2,3 tidak dihitung sebab nilai Eigenvalues Component 2 dan  $3 < 1$ , maka tidak menjadi faktor.



Gambar 7.12

**Component Matrix<sup>a</sup>**

	Component
	1
Service Quality	.835
Harga	.786
Slot Time	.634

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

Gambar 7.13

Gambar 7.12 Scree Plot ini dapat juga menunjukkan jumlah faktor yang terbentuk. Caranya dengan melihat titik Component yang memiliki nilai Eigenvalue  $> 1$ . Dari gambar Scree Plot diatas terdapat 1 titik Component yang memiliki nilai Eigenvalue  $> 1$ , maka dapat diartikan bahwa ada 1 (satu) faktor yang dapat terbentuk.

Gambar 7.13 Component Matrix ini menunjukkan nilai korelasi atau hubungan antara masing-masing variabel dengan faktor

yang akan terbentuk.

Dengan melihat ilustrasi diatas, maka dapat disimpulkan yang dapat diambil dalam analisis faktor ini adalah “Service Quality”

## **B. Tahapan Interpretasi Hasil Analisis Faktor**

Tahapan interpretasi hasil dalam analisis faktor bertujuan memberikan makna atau pemahaman yang lebih mendalam terhadap faktor-faktor yang telah terbentuk dari hasil analisis data. Proses ini melibatkan analisis berbagai *output*, termasuk nilai eigen, bobot faktor, communalities, serta rotasi faktor yang membantu menentukan variabel-variabel mana saja yang membentuk faktor tertentu. Interpretasi yang baik memerlukan pemahaman yang menyeluruh mengenai konteks penelitian, karena tujuan utama dari analisis faktor adalah menyederhanakan data tanpa kehilangan makna yang mendasarinya.

- a) Dalam interpretasi adalah memeriksa nilai eigen untuk menentukan jumlah faktor yang akan diinterpretasikan. Nilai eigen menunjukkan seberapa besar kontribusi faktor terhadap varians total dalam data. Umumnya, hanya faktor dengan nilai eigen di atas 1 yang dipertahankan untuk interpretasi, karena faktor dengan nilai lebih rendah dianggap tidak signifikan dalam menjelaskan varians data. Faktor-faktor dengan nilai eigen tinggi lebih mudah dipahami karena mereka berperan besar dalam mewakili data secara keseluruhan. Selain itu, peneliti juga bisa menggunakan scree plot untuk membantu memilih jumlah faktor yang tepat dengan melihat perubahan nilai eigen yang signifikan.
- b) Setelah menentukan jumlah faktor yang relevan, langkah selanjutnya adalah melihat bobot faktor untuk setiap variabel dalam matriks faktor. Bobot faktor ini menunjukkan seberapa besar kontribusi variabel terhadap faktor tertentu. Variabel dengan bobot tinggi pada satu faktor menunjukkan keterkaitan yang kuat dengan faktor tersebut, sehingga variabel tersebut dapat dianggap sebagai representasi utama dari faktor tersebut. Dalam beberapa kasus, variabel bisa

- memiliki bobot yang signifikan pada lebih dari satu faktor, yang memerlukan pertimbangan khusus untuk menentukan peran variabel tersebut dalam interpretasi.
- c) Selanjutnya, peneliti harus memeriksa hasil rotasi faktor, yang membantu dalam memperjelas hubungan antara variabel dan faktor. Rotasi, baik orthogonal (seperti varimax) atau oblique (seperti promax), bertujuan untuk mengoptimalkan bobot faktor sehingga setiap variabel memiliki bobot tertinggi pada satu faktor tertentu. Rotasi ini sangat penting karena membantu memisahkan variabel dengan lebih jelas ke dalam faktor yang sesuai, mempermudah peneliti dalam memahami dan menginterpretasi struktur faktor yang muncul. Dengan rotasi yang tepat, setiap faktor akan memiliki kelompok variabel yang lebih jelas dan terfokus, sehingga interpretasinya lebih akurat.
  - d) Setelah memeriksa hasil rotasi, tahapan berikutnya adalah memberikan nama atau label pada faktor berdasarkan variabel-variabel yang dominan dalam faktor tersebut. Pemberian nama ini penting agar faktor-faktor yang ditemukan memiliki makna yang relevan dan kontekstual sesuai dengan tujuan penelitian. Proses penamaan ini melibatkan analisis mendalam terhadap variabel yang membentuk faktor, serta pemahaman tentang bagaimana variabel tersebut saling terkait. Misalnya, jika faktor terdiri dari variabel "motivasi," "keterlibatan," dan "produktivitas," peneliti mungkin menamai faktor tersebut sebagai "kinerja kerja." Nama atau label faktor harus menggambarkan makna yang tepat untuk memudahkan interpretasi bagi pembaca.
  - e) Langkah terakhir dalam interpretasi adalah menilai reliabilitas faktor untuk memastikan bahwa faktor tersebut dapat diandalkan sebagai representasi dari data yang lebih luas. SPSS menyediakan uji reliabilitas seperti Cronbach's Alpha untuk mengukur konsistensi internal faktor. Faktor dengan nilai *Cronbach's Alpha* yang tinggi (umumnya lebih dari 0,7) dianggap reliabel, yang berarti bahwa variabel-variabel dalam

faktor tersebut berkorelasi cukup kuat dan dapat diandalkan. Reliabilitas yang tinggi meningkatkan validitas hasil analisis, sehingga hasil interpretasi dapat digunakan dalam penelitian atau aplikasi praktis dengan keyakinan yang lebih besar.

Interpretasi hasil analisis faktor merupakan proses yang komprehensif dan melibatkan beberapa langkah kritis untuk memastikan bahwa faktor-faktor yang dihasilkan benar-benar merepresentasikan struktur data yang akurat. Proses ini tidak hanya mengandalkan output statistik, tetapi juga memerlukan pemahaman yang mendalam mengenai konteks penelitian agar faktor-faktor yang dihasilkan relevan dan bermakna. Dengan mengikuti langkah-langkah interpretasi ini, peneliti dapat memperoleh wawasan yang lebih dalam dan menyajikan hasil yang lebih informatif bagi pembaca atau pengguna data lainnya.

#### **10.4. Implikasi Analisis Faktor Bagi Akademisi dan Praktisi**

Analisis faktor adalah metode statistik yang sangat berguna bagi akademisi untuk memahami struktur data secara mendalam dan mengidentifikasi pola yang tersembunyi dalam variabel. Bagi akademisi, teknik ini memudahkan pemetaan dimensi-dimensi konsep yang kompleks, seperti sikap, motivasi, atau perilaku. Implikasi bagi penelitian akademik mencakup kemampuan untuk mereduksi data, sehingga variabel-variabel yang saling berkorelasi dapat dikelompokkan menjadi faktor-faktor utama yang lebih mudah dianalisis. Dengan demikian, analisis faktor membantu akademisi dalam menyederhanakan data dan memperoleh insight baru dari data yang tersedia.

Selain itu, bagi akademisi, analisis faktor juga dapat digunakan untuk mengembangkan atau mengonfirmasi instrumen pengukuran dalam penelitian kuantitatif. Misalnya, ketika membuat kuesioner, akademisi dapat menggunakan analisis faktor untuk menentukan validitas konstruk, sehingga dapat dipastikan bahwa setiap dimensi atau faktor benar-benar merepresentasikan konsep yang ingin diukur. Dengan validitas yang tinggi, instrumen tersebut dapat diandalkan dalam penelitian-penelitian selanjutnya. Ini memberikan landasan

yang kuat bagi peneliti dalam mengukur konsep secara akurat, serta menjamin hasil penelitian yang lebih terpercaya.

Analisis faktor juga memiliki implikasi signifikan bagi praktisi di berbagai bidang seperti pemasaran, manajemen sumber daya manusia, dan psikologi. Dalam pemasaran, misalnya, analisis faktor memungkinkan praktisi untuk memahami preferensi pelanggan dengan lebih baik. Dengan mengidentifikasi faktor-faktor yang mendasari keputusan konsumen, perusahaan dapat mengembangkan strategi pemasaran yang lebih tepat sasaran. Faktor-faktor tersebut dapat berupa dimensi seperti “kualitas produk,” “harga,” atau “layanan pelanggan,” yang diperoleh dari variabel-variabel dalam survei pelanggan. Ini membantu perusahaan merancang pendekatan pemasaran yang lebih sesuai dengan keinginan pelanggan.

Di bidang sumber daya manusia, analisis faktor dapat membantu praktisi dalam mengevaluasi dimensi-dimensi kompetensi karyawan atau mengidentifikasi aspek-aspek yang berpengaruh terhadap kepuasan dan kinerja kerja. Misalnya, dengan melakukan analisis faktor terhadap data penilaian kinerja, manajer dapat mengelompokkan aspek-aspek seperti keterampilan teknis, soft skills, dan keterlibatan karyawan sebagai faktor utama yang memengaruhi kinerja. Dengan pemahaman ini, manajer dapat merancang program pelatihan dan pengembangan yang lebih efektif, serta melakukan perencanaan sumber daya manusia yang lebih strategis.

Bagi psikolog atau praktisi di bidang sosial lainnya, analisis faktor memungkinkan mereka untuk memahami dimensi-dimensi psikologis atau sosial yang tersembunyi dalam data survei atau kuesioner. Sebagai contoh, dalam pengembangan skala kepribadian, analisis faktor membantu mengidentifikasi dimensi-dimensi inti dari kepribadian, seperti ekstroversi, neurotisme, atau keterbukaan. Ini memungkinkan psikolog untuk mengembangkan alat ukur yang lebih akurat dan berguna dalam diagnosis atau asesmen psikologis. Hal ini tentunya berdampak pada peningkatan kualitas layanan psikologis yang diberikan kepada klien.

Analisis faktor juga mendukung pengambilan keputusan yang lebih baik di level manajerial dan kebijakan. Praktisi di bidang bisnis,

pemerintahan, atau pendidikan dapat memanfaatkan analisis faktor untuk mereduksi kompleksitas data dan fokus pada faktor-faktor yang paling berpengaruh terhadap masalah yang dihadapi. Sebagai contoh, dalam sektor pendidikan, analisis faktor dapat digunakan untuk mengevaluasi kualitas layanan pendidikan dengan mengidentifikasi dimensi-dimensi utama yang mempengaruhi kepuasan siswa, seperti kualitas pengajaran, fasilitas, dan layanan administrasi. Dengan mengetahui faktor-faktor ini, pengambil kebijakan dapat merancang program peningkatan yang lebih terfokus dan tepat sasaran.

Dalam konteks riset dan pengembangan produk, analisis faktor memberikan keuntungan bagi praktisi dengan menyajikan dimensi yang menjadi prioritas dalam pengembangan produk. Dengan mengidentifikasi faktor utama yang berpengaruh terhadap kepuasan pengguna atau efisiensi produk, tim R&D dapat membuat keputusan yang lebih informasional dalam mengembangkan fitur atau meningkatkan produk. Implikasi ini memungkinkan perusahaan untuk membuat inovasi yang lebih relevan dengan kebutuhan pasar, serta memberikan pengalaman produk yang lebih baik kepada konsumen.

Analisis faktor memberikan kontribusi yang besar dalam meningkatkan kualitas riset akademis maupun efektivitas aplikasi praktis di berbagai bidang. Teknik ini memungkinkan akademisi dan praktisi untuk mereduksi data dan memperoleh pemahaman yang lebih komprehensif dari kumpulan data yang kompleks. Dengan demikian, analisis faktor tidak hanya membantu dalam memahami struktur data, tetapi juga memberikan landasan yang kuat untuk merancang intervensi, strategi, atau kebijakan yang lebih baik.

#### **10.5. Arah Pengembangan Metode Analisis Faktor di Masa Depan**

Di masa depan, pengembangan metode analisis faktor diperkirakan akan semakin kompleks seiring dengan meningkatnya kebutuhan akan analisis data yang lebih mendalam dan akurat. Salah satu arah pengembangan yang penting adalah integrasi antara analisis faktor dengan machine learning dan teknologi kecerdasan buatan (AI). Penggabungan ini memungkinkan analisis faktor yang lebih cepat dan otomatis, terutama untuk data dalam jumlah besar (big data) yang

tidak mungkin dianalisis secara manual. Dengan menggunakan algoritma machine learning, analisis faktor dapat diproses dalam skala yang lebih besar dan diterapkan dalam berbagai konteks yang lebih luas, termasuk data yang lebih variatif dan tidak terstruktur.

Pengembangan analisis faktor ke depan juga akan lebih fokus pada penerapan data besar dan data tak terstruktur. Saat ini, analisis faktor umumnya digunakan pada data terstruktur yang telah dibersihkan dan diformat dengan baik. Namun, dengan perkembangan teknologi data seperti text mining dan sentiment analysis, analisis faktor akan berkembang untuk dapat diterapkan pada data teks, gambar, atau bahkan data suara. Hal ini membuka peluang untuk memanfaatkan analisis faktor pada data sosial media, analisis survei yang lebih komprehensif, dan bahkan dalam analisis data multimedia, yang akan sangat berguna di bidang pemasaran digital dan ilmu sosial.

Selain itu, pengembangan analisis faktor kemungkinan akan lebih diarahkan pada metode yang lebih adaptif dan fleksibel untuk berbagai jenis data. Metode tradisional seperti Principal Component Analysis (PCA) dan Exploratory Factor Analysis (EFA) mungkin akan digabungkan dengan metode baru seperti Bayesian Factor Analysis dan Confirmatory Factor Analysis (CFA) yang lebih fleksibel. Metode-metode ini mampu mengakomodasi data dengan distribusi yang tidak normal dan menangani variabel yang saling berkorelasi tinggi dengan lebih baik. Adaptabilitas ini sangat penting untuk analisis data yang kompleks dan dinamis, terutama dalam lingkungan penelitian yang memerlukan penyesuaian terhadap data yang heterogen.

Kemampuan untuk menangani masalah multikolinearitas dan heterogenitas data juga menjadi fokus pengembangan analisis faktor. Masalah-masalah ini sering muncul dalam data sosial yang memiliki variabel dengan hubungan yang rumit. Metode analisis faktor di masa depan diprediksi akan memiliki kemampuan untuk meminimalkan dampak multikolinearitas melalui algoritma yang lebih canggih, seperti penalized regression atau regularization methods. Dengan cara ini, analisis faktor dapat memberikan hasil yang lebih akurat dan reliabel, bahkan pada data dengan variabel yang sangat berkorelasi.

Pengembangan analisis faktor juga akan mengarah pada

peningkatan visualisasi hasil. Visualisasi data yang intuitif menjadi semakin penting karena membantu peneliti dan praktisi untuk memahami hasil analisis faktor dengan lebih mudah. Di masa depan, kemungkinan besar akan ada peningkatan penggunaan teknik visualisasi interaktif yang memungkinkan pengguna untuk mengeksplorasi hasil analisis secara dinamis. Misalnya, grafik heatmap dan network graph dapat memberikan pandangan yang lebih mendalam mengenai hubungan antar variabel dalam faktor-faktor yang terbentuk, sehingga hasil analisis lebih mudah dimengerti oleh audiens yang lebih luas.

Di sisi lain, dalam analisis faktor di masa depan, perhatian juga akan diarahkan pada interpretasi hasil yang lebih otomatis dan komprehensif. Dengan bantuan teknologi pemrosesan bahasa alami (NLP), metode analisis faktor akan mampu memberikan interpretasi otomatis dari faktor-faktor yang terbentuk, sehingga peneliti tidak hanya diberikan angka, tetapi juga penjelasan terkait hasil analisis. Pendekatan ini akan menghemat waktu dan meningkatkan pemahaman pengguna terhadap hasil analisis, terutama bagi mereka yang bukan ahli statistik atau akademisi.

Dalam hal penerapan praktis, metode analisis faktor di masa depan diprediksi akan lebih responsif terhadap kebutuhan industri dan bisnis. Saat ini, analisis faktor telah digunakan dalam pemasaran, psikologi, dan manajemen, tetapi dengan perkembangan metode, penerapan analisis faktor bisa diperluas ke bidang-bidang lain seperti manajemen risiko, kesehatan masyarakat, dan pendidikan. Contohnya, di bidang kesehatan, analisis faktor dapat membantu mengidentifikasi faktor risiko penyakit kronis, sedangkan di pendidikan, metode ini dapat digunakan untuk menganalisis faktor-faktor yang berpengaruh pada prestasi akademik siswa.

Pengembangan metode analisis faktor di masa depan akan sangat dipengaruhi oleh kemajuan teknologi dan kebutuhan akan analisis data yang lebih mendalam dan real-time. Dengan inovasi dalam teknologi data dan machine learning, metode ini diprediksi akan menjadi semakin adaptif, mudah diakses, dan relevan bagi berbagai bidang penelitian dan industri. Pengembangan ini akan

memungkinkan analisis faktor untuk tetap menjadi metode yang handal dalam membantu akademisi, praktisi, dan pengambil keputusan memahami data yang kompleks dan mengambil keputusan yang lebih baik berdasarkan analisis yang akurat.

## DAFTAR PUSTAKA

- Agung, I. G. N., 2001. *Statistika: Analisis Kausal Berdasarkan Data Kategorik*. Jakarta: PT Raja Grafindo Persada.
- Agung, I. G. N., 2006. *Statistika: Penerapan Model Rerata Sel Multivariat dan Model Ekonometri dengan SPSS*. Jakarta: Yayasan Sad Satria Bhakti.
- Anderson, E., 1954. Efficient and Inefficient Methods of Measuring Specific Differences. In *Statistics and Mathematics in Biology*, O. Kempthorne (Ed.). Ames: Iowa State College Press, pp. 98-107.
- Anderson, E., 1957. A Semi-graphical Method for the Analysis of Complex Problems. *Proceeding of the National Academy of Sciences, U. S.A.*, 43, 923-927. [Reprinted in *Technometrics*, 2, 387-393(1960)]
- Anderson, E., 1960. Some Stochastic Process Models for Intelligence Test Scores. In *Mathematical Methods in the Social Sciences.*, (K. J. Arrow et al. Eds.). Stanford, CA: Stanford University Press, pp. 205-220.
- Anderson, T. W., 1984. *An Introduction to Multivariate Statistical Analysis* (2nd ed.). New York: John Wiley.
- Anderson, T. W., 1963. Asymptotic Theory for Principal Components Analysis. *Annals of Mathematical Statistics*, 34, 122-148.
- Anderson, R. A., Dennis J. Sweeney, & Thomas A. Williams. 1996. *Manajemen Sains: Pendekatan Kuantitatif untuk Pengambilan Keputusan Manajemen*. Edisi Ketujuh. Alih Bahasa: Ancella A. Hermawan, MBA. Jakarta: Penerbit Erlangga.
- Andrews, D. F., 1972. Plots of High Dimensional Data, *Biometrics*, 28, 125-136
- Bartlett, M. S., 1954. A Note on Multiplying Factors for Various Chi-Squared Approximations. *Journal of the Royal Statistical Society (B)*, 16, 296-298.
- Bartlett, M. S., 1937. The Statistical Conception of Mental Factors. *British Journal of Psychology*, 28,97-104.
- Basri, M. 2006. *Pengaruh Budaya Lokal terhadap Kinerja Pelayanan Perijinan di Sulawesi Selatan*. Data hasil penelitian untuk disertasi doktor. Makassar: Program Pascasarjana Universitas

Hasanuddin.

- Becker, R. A., Cleveland, W. S., & Wilks, A. R., 1987. Dynamic Graphic for Data Analysis, *Statistical Science*, 2, no. 4, 355 – 395.
- Boruch, R. F. & Wolins, L., 1970. A Procedure for Estimation of Traits, Method, and Error Variance Attributable to a Measure. *Educational and Psychological Measurement*, 30(3), 547-5..
- Bruckener, L. A., 1978. On Chernoff faces. In *Graphical Representation of Multivariate Data*, P. C. C. Wang (Ed.). New York: Academic Press, pp. 93-121.
- Chernoff, H., 1973. Using Faces to represent Poits in KDimensional Space Graphycally, *Journal of the American Statistical Association*, 68, no. 342, 361 – 368.
- Dillon, W. R. & Goldstein, M. 1984. *Multivariate Analysis: Methods and Applications*. New York: John Wiley & Sons, Inc.
- Dun, L. C., 1928. The Effect of Inbreeding on the Bones of the Fowl. *Storrs Agricultural Experimental Station Bulletin*, 52, 1-112.
- Everitt, B., 1978. *Graphical Techniques for Multivariate Data*. New York: North-Holland.
- Fienberg, S. E., 1979. Graphical Methods in Statistics, *The American Statistician*, 33 (November), 165-178
- Gebhardt, F. 1971. Maximum Likelihood Solution to Factor Analysis When Some Factors are Completely Specified. *Psychometrica*, 36, 155-163.
- Girschick, M. A., 1939. On the Sampling Theory of Roots of Determinantal Equations. *Annals of Mathematical Statistics*, 10, 203-224.
- Gnanadesikan, R., 1977. *Methods for Statistical Data Analysis of Multivariate Observations*. New York: Wiley.
- Gorsuch, R. L., 1983. *Factor Analysis*. Second Edition. London: Lawrence Erlbaum Associates, Publishers.
- Gower, J. C. & Hand, D. J., 1996. *Biplot*. London: Chapman and Hall.
- Harmon, H. H., 1967. *Modern Factor Analysis*. Chicago: University of Chicago Press.
- Hirschey, M. & Wichern, D. W., 1984. Accounting and Market-value Measure of Profitability: Consistency, Determinats and Uses.

- Journal of Business and Economic Statistics, 2, no. 4, 375-383.
- Holzinger, K. J. & H. H. Harman, 1941. Factor Analysis: A Synthesis of Factorial Methods. Chicago: University of Chicago Press.
- Huff, D. L. & Black, W., 1978. A Multivariate Graphical Display for Regional Analysis. In Graphical Representation of Multivariate Data, P. C. C. Wang (Ed.). New York: Academic Press, pp. 199-218.
- Johnson, R., 1980. Elementary Statistics. 3rd Edition. Belmont, California: Duxbury Press, a Division of Wadsworth, Inc.
- Johnson, W., 1987. The Detection of Influential Observations for Allocation, separation, and the Determination of probabilities in Bayesian Framework. Journal of Business and Economic Statistics, 5, no. 3, 369–381
- Johnson, R. A. & Wichern, D. W., 1988, First ed. 1998 Fourth ed. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall, Englewood Cliffs.
- Jolicoeur, P., 1963. The Multivariate Generalization of the Allometry Equation. Biometrics, 19, 497-499.
- Jolicoeur, P. & Mosimann, J. E., 1960. Size and Shape Variation in the Painted Turtle: A Principal Component Analysis. Growth, 24, 339-354.
- Joreskog, K. G., 1969. A General Approach to Confirmatory Maximum Likelihood Factor Analysis, Psychometrika, 34, 183.
- Joreskog, K. G., 1971. Statistical Analysis of Sets of Congeneric Tests. Psychometrika, 36(2), 109-133.
- Joreskog, K. G., Gruvaeus, G. & Van Thillo, M. 1970. ACOUS: A General Computer Program for Analysis of Covariance Structures, In Research Bulletin, Princeton: Educational testing Service.
- Joreskog, K. G. & Lawley, D. D. 1968. New Methods in Maximum Likelihood Factor Analysis, British Journal of Mathematical and Statistical Psychology, 21, 85-96.
- Kaiser, H. F., 1958. The Varimax Criterion for Analytic Rotation in Factor Analysis. Psychometrika, 23, 187-200.
- Kelloway, E. K., 1998. Using LISREL for Structural Equation Modeling. New Delhi: SAGE Publications.

- Kerlinger, Fred N., 1973. *Foundations of Behavioral Research*, Second Edition. New York: Holt, Rinehart and Winston, Inc.
- Kim, J. & Mueller, C. W., 1978. *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills, London: Sage Publications.
- King, B., 1966. Market and Industry Factors in Stock Price Behavior. *Journal of Business*, 39, 139-190.
- Kourti, T. & McGregor, J., 1996. Multivariate SPC Methods for Process and Product Monitoring. *Journal of Quality Technology*, 28, 409-428.
- Kruskal, J. B., 1964. Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis, *Psychometrika*, 29, no. 1, 1-27.
- Lancaster, H. O., 1969. *The Chi-squared Distribution*. New York: John Wiley and Sons.
- Lawley, D. N., 1963. Testing a Set of Correlation Coefficients for Equality. *Annals of Mathematical Statistics*, 34, 149-151.
- Lawley, D. N. & Maxwell, A. E., 1971. *Factor Analysis as a Statistical Method* (2d ed.). New York: American Elsevier Publishing Co.
- Lebart, L., Morineau, A., & Warwick, K. M., 1984. *Multivariate Descriptive Statistical Analysis*. Translated by Elisabeth Moraillon Berry New York: John Wiley & Sons.
- Linden, M. 1977. A Factor Analysis Study of Olympic Decathlon Data. *Research Quarterly*, 48, no. 3, 562-568.
- MacQueen, J. B., 1967. Some Methods for Classification and Analysis Multivariate Observations, *Proceedings of 5<sup>th</sup> Berkeley Symposium on Mathematical Statistics and Probability*, 1, Berkeley, California: University of California Press, 281-297.
- Rosenberg, M. & Simmons, R. G., 1971. *Black and White Selfesteem: The Urban School Child*. Rose Monograph Series. Washington: American Sociological Association.
- Sharma, S. 1996. *Applied Multivariate Techniques*. New York: John Wiley & Sons, Inc.
- Smith, M. & Taffler, R., 1984. Improving the Communication Function of Published Accounting Statements. *Accounting and Business Research*, 14, no. 45, 139 – 146.
- Spearman, C. 1927. *The Abilities of Man*. New York: McMillan

Company.

- Stoetzel, J., 1960. A Factor Analysis of Liquor Preference. *Journal of Advertising Research*, 1, 7-11
- Supranto, J., 2004. *Analisis Multivariat: Arti & Interpretasi* Jakarta: Penerbit Rineka Cipta.
- Suryabrata, S. 1982. "Dasar-dasar Analisis Faktor", *Metodolog Penelitian: Analisis Kuantitatif*. Yogyakarta: Lembaga Pendidikan Doktor Universitas Gajah Mada.
- Thurstone, L. L. 1947. *Multiple Factor Analysis*. Chicago University of Chicago Press.
- Tiro, M. A., 2000a. *Dasar-dasar Statistika*. Makassar: State University of Makassar Press.
- Tiro, M. A., 2000b. *Analisis Korelasi dan regresi*. Makassar: State University of Makassar Press.
- Tiro, M. A., 2000c. *Analisis Regresi dengan Data Kategori*. Makassar: State University of Makassar Press.
- Tiro, M. A., 2005. *Analisis Korelasi dengan Data Kategori*. Makassar: Andira Publisher.
- Tukey, J. W., 1977. *Explanatory Data Analysis*. Reading, Mass.: Addison-Wesley.
- Wainer, H. & Thissen, D., 1981. *Graphical Data Analysis*. *Annual Review of Psychology*, 32. 191 – 241.
- Weihs, C. & Schmidli, H., 1990. *Online Exploratory Graphical Analysis: Routine Searching for Structure*, *Statistical Science*, 5, no. 2, 175 – 226.
- Wright, S., 1954. *The Interpretation of Multivariate System*. In *Statistics and Mathematics in Biology*, edited by O. Kempthorne and others. Ames, IA: Iowa State University Press, 11 – 33.

# ANALISIS KLUSTER

*Oleh Sri Hidayati L, S.KM., M.Kes.*

## 11.1 Pengantar

*Cluster analysis* atau analisis kluster adalah metode statistik yang bertujuan mengelompokkan objek-objek ke dalam grup yang memiliki karakteristik serupa. Metode ini sering digunakan dalam berbagai bidang seperti pemasaran, biologi, pendidikan, dan kesehatan masyarakat untuk memahami pola, hubungan, dan segmentasi dalam data. Analisis kluster membantu peneliti menyederhanakan kompleksitas data dengan mengelompokkan objek ke dalam gugus atau cluster berdasarkan kesamaan karakteristik tertentu (UMA, 2022).

Analisis cluster adalah cara untuk mengelompokkan objek data yang memiliki ciri-ciri yang sama. Ini menciptakan kelompok di mana item di dalamnya samadan berbeda dari yang lain. Metode ini menggunakan ukuran jarak untuk mengetahui kemiripan suatu benda. Ini adalah jenis pembelajaran tanpa pengawasan yang membantu menemukan pola dalam data tanpa label. Teknik analisis kluster adalah kunci dalam analisis data. Mereka membantu menemukan pola dalam data Anda. Metode pengelompokan seperti partisi, hierarki, dan berbasis kepadatan digunakan. Setiap metode memiliki algoritma tersendiri untuk membuat cluster berdasarkan sifat-sifat tertentu (Expert, 2024).

Teknik ini secara umum digunakan pada beberapa bidang yaitu Pengenalan Pola, Analisis Data Spasial (Membuat peta tematik dalam

GIS dengan mengelompokkan ruang fitur dan mendeteksi kluster spasial dan menjelaskannya dalam penggalian data spasial), Pengolahan gambar, Bidang ekonomi (Penelitian pasar), dan *World Wide Web (WWW)* menggunakan teknik ini dalam klasifikasi dokumen atau mengelompokkan data *weblog* untuk menemukan kelompok dengan pola akses yang sama. Contoh aplikasi analisis kluster:

1. Pemasaran: Membantu produsen dan penjual dalam menemukan kelompok-kelompok yang berbeda dalam basis pelanggan mereka, dan kemudian menggunakan teknik ini untuk mengembangkan program pemasaran yang telah ditargetkan.
2. Penggunaan lahan: Identifikasi area dengan penggunaan lahan yang sama dalam basis data observasi bumi.
3. Asuransi: Mengidentifikasi kelompok pemegang polis asuransi kendaraan bermotor dengan biaya klaim yang rata-rata tinggi.
4. Perencanaan kota: Mengidentifikasi kelompok rumah berdasarkan tipe rumah, nilai, dan lokasi geografisnya.
5. Studi gempa bumi: Pusat gempa bumi yang teramati harus dikelompokkan di sepanjang patahan benua.

Apa yang dimaksud dengan *Clustering* yang baik? Metode pengelompokan yang baik akan menghasilkan kluster yang berkualitas tinggi dengan kesamaan intra-kelas yang tinggi dan kemiripan antar-kelas yang rendah. Kualitas hasil pengelompokan bergantung pada ukuran kemiripan yang digunakan oleh metode dan implementasinya. Kualitas metode clustering juga diukur dari kemampuannya untuk menemukan beberapa atau semua pola tersembunyi (*Cluster Analysis*, 2016).

## **11.2 Metode Cluster Analisis**

Dewasa ini sudah banyak metode-metode clustering dikembangkan dalam mengatasi permasalahan pengelompokan data. Namun, pada saat ini hanya fokus pada dua metode clustering saja, yaitu Metode Hirarki dan Metode Non-Hirarki (Bustami, 2015).  
Macam-macam metode clustering:

1. Berbasis Metode Statistik

- a. Metode Hirarki
- b. Metode Non Hirarki
2. Berbasis Logika Fuzzy
  - a. Metode Fuzzy C-Means
  - b. Metode Fuzzy Subtractive Clustering
3. Berbasis Neural Network
4. Berbasis Optimasi Centroid atau Lebar Cluster
5. Metode Genetik Algoritma (GA)

a. Metode Hirarki

Dalam dunia analisis data, terdapat dua pendekatan utama yang dapat digunakan untuk melakukan analisis cluster, yaitu metode hierarki dan metode non-hierarki. Metode hierarki dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Proses ini kemudian dilanjutkan dengan menggabungkan objek lain yang memiliki kedekatan kedua, dan seterusnya. Hasil akhirnya adalah pembentukan struktur yang menyerupai pohon, di mana terdapat tingkatan yang jelas antara objek-objek, mulai dari yang paling mirip hingga yang paling tidak mirip. Dengan demikian, semua objek pada akhirnya akan terorganisir dalam sebuah kluster. Untuk mempermudah pemahaman mengenai proses hierarki ini, biasanya digunakan dendogram, yang berfungsi untuk memperjelas hubungan antar objek dalam analisis tersebut.

Teknik pengelompokan pada metode hierarki menggunakan teknik penggabungan (*agglomerative*) dan teknik pembagian (*divisive*). Dalam pendekatan *agglomerative*, proses dimulai dengan setiap objek sebagai kluster terpisah, kemudian objek-objek yang paling mirip digabungkan secara bertahap hingga terbentuk satu kluster besar. Sebaliknya, metode *divisive* memulai dengan satu kluster besar yang kemudian dibagi menjadi kluster-kluster yang lebih kecil berdasarkan kriteria tertentu (Ls et al., 2021; Matdoan & Delsen, 2020).

Salah satu keunggulan dari metode hierarki adalah kemampuannya untuk memberikan gambaran visual tentang hubungan antar kluster melalui dendogram. Dendogram adalah

diagram pohon yang menunjukkan bagaimana kluster dibentuk dan hubungan antar kluster tersebut. Dengan menggunakan dendrogram, peneliti dapat dengan mudah mengidentifikasi tingkat kemiripan antar objek dan menentukan jumlah kluster yang optimal (Faisal & Rifai, 2023; Ls et al., 2021). Metode ini juga memungkinkan peneliti untuk mengeksplorasi data secara mendalam, karena setiap langkah dalam proses pengelompokan dapat dianalisis dan dipahami dengan lebih baik (Matdoan & Delsen, 2020).

Dalam praktiknya, terdapat beberapa teknik pengukuran jarak yang dapat digunakan dalam metode hierarki, seperti single linkage, complete linkage, average linkage, dan Ward's method. Setiap teknik memiliki cara yang berbeda dalam menghitung jarak antar kluster, yang dapat mempengaruhi hasil akhir dari analisis kluster. Misalnya, metode Ward's dikenal efektif dalam menghasilkan kluster yang lebih homogen, sehingga sering digunakan dalam analisis yang memerlukan presisi tinggi (Koesnadi et al., 2023; Ls et al., 2021).

Secara keseluruhan, metode hierarki dalam analisis kluster menawarkan pendekatan yang sistematis dan terstruktur untuk mengelompokkan data, yang sangat berguna dalam berbagai bidang penelitian, termasuk kesehatan, pendidikan, dan ekonomi. Dengan kemampuannya untuk memberikan wawasan yang mendalam tentang struktur data, metode ini menjadi alat yang penting dalam analisis multivariat (Ls et al., 2021; Matdoan & Delsen, 2020).

#### b. Metode Non Hirarki

Metode non-hierarki dalam analisis kluster adalah teknik yang digunakan untuk mengelompokkan data tanpa membentuk struktur hierarkis. Salah satu metode yang paling terkenal dalam kategori ini adalah K-means, yang bekerja dengan cara membagi data ke dalam sejumlah kluster yang telah ditentukan sebelumnya oleh peneliti. Dalam K-means, proses dimulai dengan pemilihan sejumlah titik pusat (centroid) secara acak, dan kemudian setiap objek dalam dataset akan dikelompokkan ke dalam kluster berdasarkan kedekatannya dengan centroid tersebut. Proses ini

diulang hingga posisi centroid stabil dan tidak ada perubahan signifikan dalam pengelompokan (10060119034 & Nur Azizah Komara Rifai, 2024; Rosmini et al., 2018).

Salah satu kelebihan dari metode non-hierarki adalah kemampuannya untuk menangani dataset yang besar dengan efisien. K-means, misalnya, dikenal karena kecepatan dan kesederhanaannya dalam mengelompokkan data numerik. Namun, metode ini juga memiliki kelemahan, seperti ketergantungan pada pemilihan awal centroid yang dapat mempengaruhi hasil akhir kluster. Oleh karena itu, sering kali disarankan untuk menjalankan algoritma ini beberapa kali dengan pemilihan centroid yang berbeda untuk mendapatkan hasil yang lebih konsisten (Cupak, 2022; Park et al., 2011).

Selain K-means, terdapat juga metode lain dalam kategori non-hierarki, seperti K-medoids, yang lebih tahan terhadap outlier karena menggunakan objek yang ada dalam dataset sebagai pusat kluster, bukan titik centroid yang dihitung. Metode ini dapat digunakan untuk data yang mencakup variabel numerik dan kategorik, sehingga memberikan fleksibilitas lebih dalam pengelompokan (10060119034 & Nur Azizah Komara Rifai, 2024; Thaha et al., 2023). Metode non-hierarki juga sering digunakan dalam kombinasi dengan metode hierarki untuk meningkatkan akurasi dan efektivitas analisis kluster, seperti yang terlihat dalam penelitian yang menggabungkan K-means dengan hasil dari metode Ward (Park et al., 2011; Fauziah & Achmad, 2023).

Secara keseluruhan, metode non-hierarki menawarkan pendekatan yang praktis dan efisien untuk analisis kluster, terutama ketika peneliti memiliki pemahaman yang jelas tentang jumlah kluster yang diinginkan. Dengan berbagai teknik yang tersedia, peneliti dapat memilih metode yang paling sesuai dengan karakteristik data yang mereka analisis (Cupak, 2022; Fitriyani & Iswani Achmad, 2023).

### **11.3 Langkah-Langkah SPSS Metode Cluster Analysis**

Analisis cluster adalah teknik statistik yang digunakan untuk mengelompokkan data menjadi beberapa kelompok (cluster)

berdasarkan kesamaan karakteristik. SPSS menyediakan beberapa metode untuk melakukan analisis cluster, yaitu metode hierarkis dan non-hierarkis (k-means).

#### 1. Metode Hierarkis

Metode ini membangun hierarki cluster secara bertahap, mulai dari setiap data sebagai cluster tunggal hingga semua data bergabung menjadi satu cluster. Langkah-langkah:

##### a. Standarisasi Data:

- 1) Pastikan semua variabel yang digunakan memiliki skala yang sama. Jika tidak, lakukan standarisasi (misalnya, z-score).
- 2) Menu: Analyze -> Descriptive Statistics -> Descriptives

##### b. Analisis Cluster:

- 1) Menu: Analyze -> Classify -> Hierarchical Cluster
- 2) Variable: Masukkan variabel yang sudah distandarisasi ke dalam kotak "Variables".
- 3) Method: Pilih metode linkage (single linkage, complete linkage, average linkage, dsb.) dan metode jarak (Euclidean distance, Manhattan distance, dsb.).

##### c. Interpretasi Hasil:

- 1) Dendrogram: Visualisasi hasil clustering yang menunjukkan hubungan antara cluster.
- 2) Agglomeration Schedule: Tabel yang menunjukkan tahapan penggabungan cluster.
- 3) Cluster Membership: Tabel yang menunjukkan anggota setiap cluster.

**Latihan 1:** Peneliti akan melakukan identifikasi terhadap berbagai merk susu bubuk yang dijual di pasaran. Variabel penelitian sebagai berikut:

- a. Merk susu
- b. Kandungan lemak (g/100 gr susu)
- c. Karbohidrat (g/100 gr susu)
- d. Mineral (g/100 gr susu)
- e. Energi (kilokalori/100 gr susu)

	MerkSusu	Lemak	Karbohidrat	Mineral	Energi
1	Oat Quaker	11.00	57.00	.96	370.00
2	Nestle Carnation	6.00	9.70	1.60	119.00
3	Logi Havermout	9.16	67.33	.00	386.67
4	Ovaltine	8.40	75.00	1.28	416.00
5	Milo	10.00	66.00	4.50	400.00
6	Dancow Balita	23.67	44.67	4.67	476.67
7	Frisian Flag Instant	26.00	40.00	5.80	496.00
8	Frisian Flag Full Cream	28.00	38.40	5.60	506.00
9	Frisian Flag Coklat	13.00	69.70	3.00	444.00
10	Frisian Flag Madu	18.00	52.80	5.00	459.00
11	Dancow Coklat	110.00	65.25	4.50	425.00
12	Dancow Full Cream	25.90	40.33	5.92	495.00
13	Indomilk Full Cream	26.00	36.90	5.60	505.60
14	Indomilk Coklat	14.00	62.50	5.00	446.50
15	Presteo Rendah Lemak	1.00	50.00	8.00	357.00
16	Anlene Kalsium Tinggi	.80	48.80	5.70	340.00
17	Tropicana Slim	.00	50.50	1.92	347.00
18	Protifar	1.00	27.80	7.20	366.00

Gambar 1. Data view

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	MerkSusu	String	23	0	Merk Susu	None	None	23	Left	Nominal	Input
2	Lemak	Numeric	12	2	Kandungan lemak (g/100 g susu)	None	None	12	Right	Scale	Input
3	Karbohidrat	Numeric	12	2	Karbohidrat (g/100 g susu)	None	None	12	Right	Scale	Input
4	Mineral	Numeric	12	2	Mineral (g/100 g susu)	None	None	12	Right	Scale	Input
5	Energi	Numeric	12	2	Energi (kalori/100 g susu)	None	None	12	Right	Scale	Input

Gambar 2. Variable view

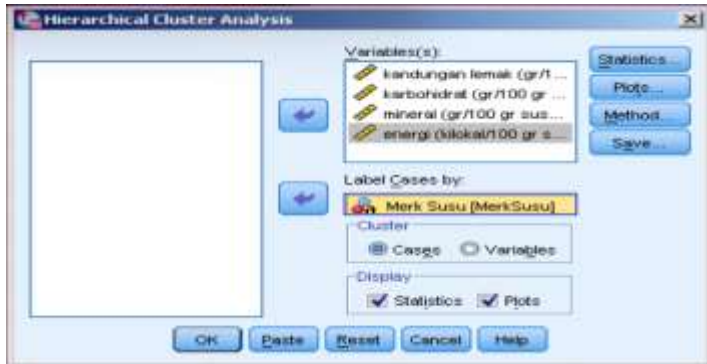
Langkah-langkah:

1. Transformasi data (standarisasi dengan Z-Score)

Bila satuan data sangat bervariasi, akan menyebabkan bias, sehingga data asli harus ditransformasi dengan cara distandarisasi sebelum dilakukan

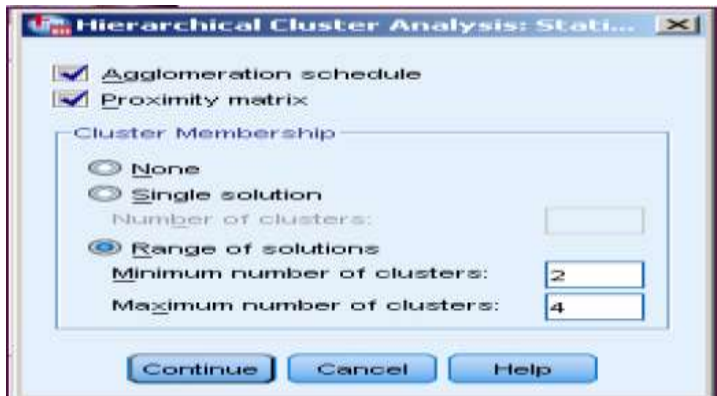
Transformasi dilakukan terhadap variabel ke bentuk z-score. Pada Hierarchical Cluster transformasi bisa dilakukan langsung pada Method di menu utama Cluster.

- *Analyze* lalu *Classify* lalu *Hierarchical Cluster*
- Masukkan semua variabel interval-rasio ke kotak Variable(s): lemak, karbohidrat, mineral dan energi
- Masukkan variabel susu ke dalam *Label Cases by:*
- Pada *Cluster* aktifkan *Cases*
- Pada *Display* aktifkan *Statistics* dan *Plots*



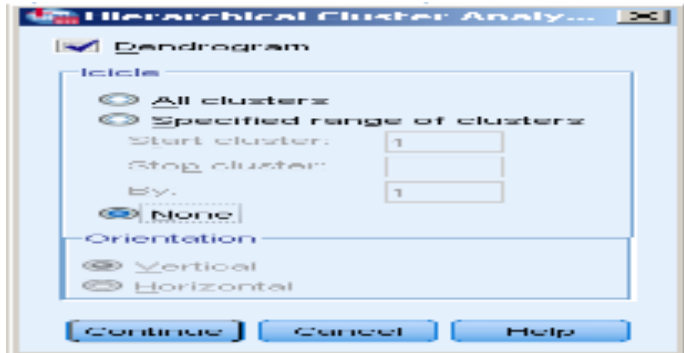
**Gambar 3.** Analyze variable

- Pada kotak *Statistics*, Aktifkan *Agglomeration schedule* dan *proximity matrix* (menampilkan jarak antar variabel)
- Pada *Cluster Membership*, klik pada *Range of solution*, isi bilangan 2 pada *Minimum number of clusters* dan 4 pada *Maximum number of cluster*
- Klik *Continue*



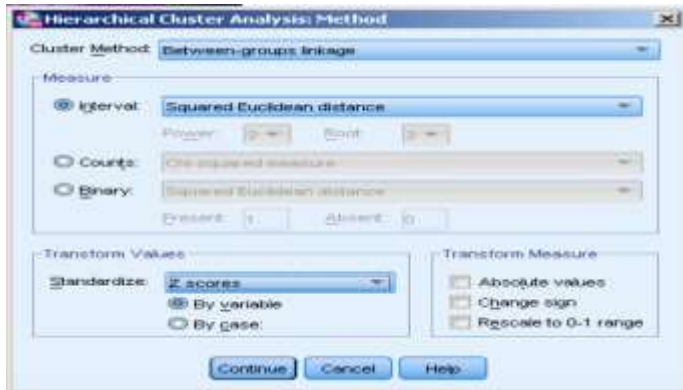
**Gambar 4.** Statistics

- Klik kotak *Plots*, aktifkan *Dendogram* pada *Icicle* aktifkan *None*
- Klik *Continue*



Gambar 5. Dendrogram

- Klik kotak *Method*, pada *Cluster Method*: pilih *Between-groups linkage*
- Klik tampilan menu utama Cluster, klik OK



Gambar 6. Cluster Method

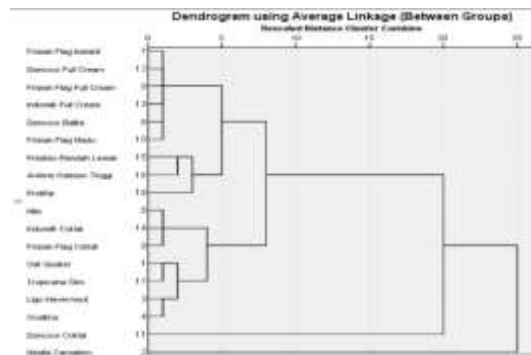
## 2. Output



- Anggota *cluster* 4: Dancow Coklat

Cluster Membership			
Case	4 Clusters	3 Clusters	2 Clusters
1:Oat Guaker	1	1	1
2:Nestle Carnation	2	2	2
3:Ligo Havermout	1	1	1
4:Ovaltine	1	1	1
5:Milo	1	1	1
6:Dancow Balita	3	1	1
7:Frisian Flag Instant	3	1	1
8:Frisian Flag Full Cream	3	1	1
9:Frisian Flag Coklat	1	1	1
10:Frisian Flag Madu	3	1	1
11:Dancow Coklat	4	3	1
12:Dancow Full Cream	3	1	1
13:Indomilk Full Cream	3	1	1
14:Indomilk Coklat	1	1	1
15:Presteo Rendah Lemak	3	1	1
16:Anlene Kalsium Tinggi	3	1	1
17:Tropicana Slim	1	1	1
18:Prolifor	3	1	1

Gambar 9. Cluster Membership



Gambar 10. Dendrogram using Average Linkage

## 2. Metode Non Hierarkis (K-Means)

Metode ini membagi data menjadi jumlah cluster yang telah ditentukan sebelumnya.

### a. Standarisasi Data:

- 1) Pastikan semua variabel yang digunakan memiliki skala yang sama. Jika tidak, lakukan standarisasi (misalnya, z-score) Pastikan semua variabel yang digunakan memiliki skala yang sama. Jika tidak, lakukan standarisasi (misalnya, z-score)
- 2) Menu: *Analyze -> Descriptive Statistics -> Descriptives*

- b. Analisis Cluster:
  - 1) Menu: *Analyze -> Classify -> K-Means Cluster*
  - 2) *Number of Clusters*: Tentukan jumlah cluster yang diinginkan.
  - 3) Pilih metode iterasi (*iterate and classify*).
- c. Interpretasi Hasil:
  - 1) *Cluster Membership*: Tabel yang menunjukkan anggota setiap cluster.
  - 2) *Centroids*: Nilai rata-rata setiap variabel untuk setiap cluster.

**Latihan 2:** Peneliti akan mengelompokkan pelanggan rumah sakit berdasarkan profil pelanggan, yaitu:

- a. Usia (tahun)
- b. Jumlah anak (orang)
- c. Penghasilan (Rp/bulan)
- d. Kegiatan membaca koran (jam/minggu)
- e. kegiatan menonton TV (jam/minggu)
- f. Jumlah motor yang dimiliki (buah)
- g. Jumlah mobil yang dimiliki (buah)
- h. Jumlah kartu kredit yang dimiliki (buah)
- i. Tingkat pembelian barang (frekuensi/minggu)
- j. Konsumsi membeli barang (Rp./bulan)
- k. Jumlah jam kerja(jam/minggu)
- l. Jumlah kegiatan belanja (jam/minggu)

	Kategori	Daerah	Status	Usia	Anak	Income	motor	Tv	Motor	Mobil	Kredit	Beli	Konsumsi	kerja	Shopping
1	Jakarta	1	1	26	0	2550000	10	20	1	1	0	4	800000	40	50
2	Bandung	1	2	26	0	750000	11	18	0	1	0	4	200000	40	12
3	Tegal	5	3	25	2	300000	0	16	0	0	1	2	80000	15	7
4	Semarang	1	3	40	2	750000	9	15	2	1	4	6	225000	55	20
5	Yogyakarta	2	2	26	0	600000	7	11	1	1	2	3	160000	30	10
6	Solo	2	1	30	0	400000	6	14	1	0	1	1	180000	30	8
7	Makassar	3	2	26	0	250000	6	17	0	0	1	2	75000	16	6
8	Mekong	2	2	34	0	400000	5	20	1	1	2	2	120000	40	11
9	Palembang	2	3	26	2	400000	5	19	1	0	2	2	190000	45	14
10	Jakarta	5	3	24	3	300000	3	15	1	0	1	1	90000	14	6
11	Bora	5	1	20	0	350000	4	11	1	0	0	1	100000	10	6
12	Karangany	3	2	26	0	350000	3	14	1	0	1	1	100000	15	5
13	Jakarta	1	2	30	0	1500000	8	18	2	0	0	3	400000	70	21
14	Bandung	1	2	30	0	1000000	4	12	1	1	3	0	300000	45	11
15	Surabaya	1	0	42	2	1400000	9	10	1	1	4	6	480000	60	16
16	Makassar	0	4	26	1	400000	6	11	1	0	1	1	120000	14	9
17	Palembang	3	1	26	0	300000	3	17	1	1	1	2	90000	15	6
18	Tuban	5	2	26	0	250000	4	12	1	0	1	2	70000	11	3
19	Jakarta	1	2	42	4	1000000	5	16	2	1	3	0	400000	42	14
20	Semarang	1	2	31	0	1000000	7	16	2	1	4	0	300000	50	16
21	Surabaya	1	1	32	0	400000	8	20	2	0	4	0	180000	55	20
22	Surabaya	1	1	34	0	800000	6	21	2	1	3	0	370000	45	14
23	Gianyar	0	1	40	0	1000000	5	22	1	1	1	2	300000	12	6
24	Pati	3	1	29	0	300000	2	16	2	0	0	2	90000	12	6
25	Cepu	3	3	27	2	250000	2	16	1	0	1	3	75000	15	6
26	Wongasri	3	2	26	0	300000	3	19	1	0	1	3	80000	16	6

Gambar 11. Data View

	Name	Type	Width	Decimals	Label	Values	Missing	Column	Align	Measure
1	Kategori	String	12	0	Kategori tempat tinggal	None	None	7	Left	Nominal
2	Daerah	Numeric	12	0	Daerah tempat tinggal	{1, kota b.	None	5	Right	Scale
3	Status	Numeric	12	0	Status pernikahan	{1, belum	None	4	Right	Scale
4	Usia	Numeric	12	0	Usia	None	None	3	Right	Scale
5	Anak	Numeric	12	0	Jumlah anak	None	None	4	Right	Scale
6	Income	Numeric	12	0	Penghasilan rata-rata per bulan	None	None	6	Right	Scale
7	Motor	Numeric	12	0	Jumlah jam membaca koran setiap minggu	None	None	4	Right	Scale
8	Tv	Numeric	12	0	Jumlah permesoran tv setiap minggu	None	None	2	Right	Scale
9	Mobil	Numeric	12	0	Jumlah motor yang dimiliki	None	None	3	Right	Scale
10	Mobil	Numeric	12	0	Jumlah mobil yang dimiliki	None	None	3	Right	Scale
11	Kartukredit	Numeric	12	0	Jumlah kartu kredit atau ATM yang dimiliki	None	None	3	Right	Scale
12	Beli	Numeric	12	0	Tingkat pembelian barang setiap minggu	None	None	3	Right	Scale
13	Konsumsi	Numeric	12	0	Tingkat pengeluaran bulanan	None	None	6	Right	Scale
14	Kerja	Numeric	12	0	Jumlah jam kerja setiap minggu	None	None	3	Right	Scale
15	Shopping	Numeric	12	0	Jumlah jam berbelanja setiap minggu	None	None	6	Right	Scale

Gambar 12. Variable View

Langkah-langkah:

1. Transformasi data (standarisasi dengan Z-Score)

Bila satuan data sangat bervariasi, akan menyebabkan bias, sehingga data asli harus ditransformasi dengan cara distandarisasi sebelum dilakukan analisis. Transformasi dilakukan terhadap variabel ke bentuk z-score.

- Analyze => Descriptive Statistics=> Descriptive
- Masukkan semua variabel interval rasio ke kotak Variable(s): usia, anak, income, koran, tv, motor, mobil,
- k\_kredit, beli, konsumsi, kerja dan shopping.

- Aktifkan (klik pada kotak) pilihan *Save Standardized values as variables*
- Klik OK



○ Gambar 13. Descriptive

- Muncul 12 variabel baru dengan imbuhan 'Z' di depan setiap variabel (*Zusia, Zanak, Zincome, Zkoran, Ztv, Zmotor, Zmobil, Zkredit, Zbeli, Zkonsumsi, Zkerja dan Zshopping*).

	Shopping	Zusia	Zanak	Zincome	Zkoran	ZTV	Zmotor	Zmobil	Zkredit	Zbeli	Zkonsumsi	Zkerja	Zshopping
1	10	-80764	-58772	286288	165225	117175	-38311	105815	63411	13275	271299	46532	-13879
2	12	-54735	-58772	12154	208863	59145	130036	105815	63411	13275	14736	46532	24185
3	7	-80764	112903	-79587	-28965	01116	-208657	-90784	-79583	-63444	-77627	-86740	-70954
4	20	-59679	112903	12154	129686	-27889	130036	105815	135099	-90001	14736	126477	176540
5	10	79531	-58772	-38813	50310	-143668	-38311	105815	-88271	-25382	-36577	-06765	-13879
6	9	-30617	-58772	-49007	11672	-68914	-38311	-90784	-88271	-25382	-48839	-06765	-13879
7	5	-80764	-58772	-89780	11672	30131	-208657	-90784	-79583	-63444	-67889	-86740	-108930
8	11	-96794	-58772	-90200	-28965	117175	-38311	105815	-88271	-63444	-57102	-06765	-05113
9	14	-54735	112903	-49007	-28965	88180	-38311	-90784	63411	-63444	-48839	73181	62089
10	5	-144803	192481	-79587	-108243	27889	-38311	-90784	-79583	-101805	-77627	-83209	-108930
11	6	-188912	-58772	-80288	-68925	-143668	-38311	-90784	-151636	-101805	-70785	-13358	-89840
12	5	79531	-58772	-60389	-108243	-68914	-38311	-90784	-79583	-101805	-70785	-86740	-108930
13	21	-00617	-58772	169065	88848	01116	130036	-90784	208775	-25382	168674	206423	196032
14	11	79531	-58772	53121	129986	-14843	-38311	105815	63411	51640	66149	73181	05113
15	18	181738	112903	165865	129986	-128173	-38311	105815	135099	30001	168674	153126	138056
16	9	-63846	26565	-60260	-28965	-143668	-38311	-90784	-79583	-101805	-53681	-92029	-13879
17	6	-22675	-58772	-79587	-108243	30131	-38311	105815	-79583	-63444	-77627	-86740	-89840
18	3	-80764	-58772	-89780	-68925	-14843	-38311	-90784	-79583	-63444	-67889	-108930	-146522
19	14	181738	26577	165865	-28965	-27889	130036	105815	63411	168724	168674	57182	62089
20	18	15413	-58772	53121	50310	88180	130036	105815	135099	216885	68048	153126	138056

Gambar 14. Data View

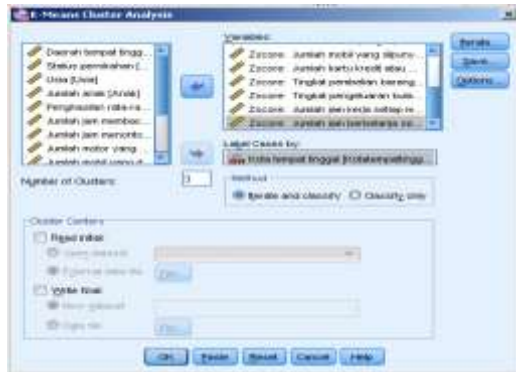
## 2. Analisis gugus

- *Analyze => Classify => K-Means Cluster*
- Masukkan semua variabel yang sudah distandarisasi (*Zusia, Zanak, Zincome, Zkoran, Ztv, Zmotor, Zmobil, Zkredit, Zbeli, Zkonsumsi, Zkerja dan Zshopping*).

Zmotor, Zmobil, Zk\_kredit, Zbeli, Zkonsumsi, Zkerja dan Zshopping).

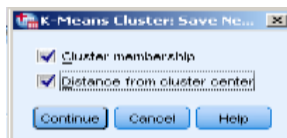
- Masukkan variabel kota tempat tinggal ke *Label Cases by*:

Isi jumlah cluster sebesar 3 yang akan dibentuk ke *Number of Cluster*. Jumlah cluster biasanya 2 – 5, tergantung tujuan penelitian dan faktor subyektif lain dari peneliti.



Gambar 15. Mean Cluster Analysis

- Klik kotak *Save*.
- Kotak dialog *Save* memungkinkan hasil analisis *cluster* disimpan dalam bentuk variabel baru di SPSS Data Editor.
- Aktifkan kotak *Cluster membership* dan *Distance from cluster center*. Klik tombol *Continue*.

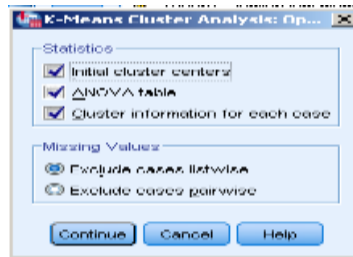


Gambar 15. Save box

- Klik kotak *Options*.
- Pada *Statistics* aktifkan *Initial cluster centers*

dan *Anova table*. Klik *Continue*.

- Pada menu utama *Cluster*, klik *OK*.



**Gambar 15.** *Option box*

### 3. *Output*

- *Output* disimpan dengan nama *K\_Means Cluster.spv*
- *Output* Final Cluster Centers menunjukkan bahwa ke-3 cluster karakteristik memiliki yang berbeda.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Usia	26	20	42	30.04	6.238
Jumlah anak	26	0	4	89	1.158
Penghasilan rata-rata per bulan	26	250000	2800000	690384.82	490513.859
Jumlah jam membaca koran setiap minggu	26	2	11	5.73	2.523
Jumlah jam menonton TV setiap minggu	26	10	22	15.96	3.447
Jumlah motor yang dimiliki	26	0	2	1.23	.587
Jumlah mobil yang dimiliki	26	0	1	.46	.508
Jumlah kartu kredit atau ATM yang dimiliki	26	0	5	2.12	1.395
Tingkat pembelian barang setiap minggu	26	1	9	3.85	2.607
Tingkat pengeluaran bulanan	26	75000	800000	283461.54	148162.712
Jumlah jam kerja setiap minggu	26	10	70	31.27	18.769
Jumlah jam berbelanja setiap minggu	26	3	21	10.73	5.268
Valid N (listwise)	26				

**Gambar 15.** *Output*

Initial Cluster Centers			
	Cluster		
	1	2	3
Zscore: Usia	-80764	1.91738	-1.60912
Zscore: Jumlah anak	-59772	2.85577	-59772
Zscore: Penghasilan rata-rata per bulan	2.68088	1.65055	-89393
Zscore: Jumlah jam membaca koran setiap minggu	1.89225	-28866	-68605
Zscore: Jumlah jam menonton TV setiap minggu	1.17175	-27899	-1.43958
Zscore: Jumlah motor yang dimiliki	-39311	1.31036	-39311
Zscore: Jumlah mobil yang dimiliki	1.05915	1.05915	-90784
Zscore: Jumlah kartu kredit atau ATM yang dimiliki	63411	63411	-1.51635
Zscore: Tingkat pembelian barang setiap minggu	1.3279	1.66724	-1.01805
Zscore: Tingkat pengeluaran bulanan	2.71299	1.68874	-70765
Zscore: Jumlah jam kerja setiap minggu	46532	57192	-1.13358
Zscore: Jumlah jam berbelanja setiap minggu	-1.3879	62089	-89946

Iteration History <sup>a</sup>			
Iteration	Change in Cluster Centers		
	1	2	3
1	0.000	1.115	0.000
2	0.00	.000	.000

a. Convergence achieved due to no or small change in cluster centers. The maximum absolute coordinate change for any center is .000. The convergence criterion is 2. The minimum distance between initial centers is 5.773.

**Gambar 16.** Initial Cluster Centers

Final Cluster Centers			
	Cluster		
	1	2	3
Zscore: Usia	.18619	1.08919	-.49774
Zscore: Jumlah anak	-.59772	.55344	-.02214
Zscore: Penghasilan rata-rata per bulan	1.14088	.83507	-.71432
Zscore: Jumlah jam membaca koran setiap minggu	1.13731	.56916	-.60677
Zscore: Jumlah jam menonton TV setiap minggu	.47540	.20459	-.24030
Zscore: Jumlah motor yang dimiliki	.28828	1.02645	-.50867
Zscore: Jumlah mobil yang dimiliki	.66575	.73132	-.51444
Zscore: Jumlah kartu kredit atau ATM yang dimiliki	.63411	1.11199	-.65617
Zscore: Tingkat pembelian barang setiap minggu	-.02066	1.60330	-.63444
Zscore: Tingkat pengeluaran bulanan	1.17361	.74601	-.68961
Zscore: Jumlah jam kerja setiap minggu	.56126	1.14930	-.64680
Zscore: Jumlah jam berbelanja setiap minggu	.24105	1.28660	-.59459

**Gambar 17.** Final Cluster Centers

Contoh variabel Usia:

- Cluster 1 mempunyai usia  $z = 0,18619$ ,
- Usia pada cluster 2 ( $z = 1,08919$ ),
- Usia pada cluster 3 ( $z = 0,49774$ )

Jika  $x = m + z \cdot \theta$  dan

	N	Minimum	Maximum	Mean	Std. Deviation
Usia	26	20	42	30,04	6,238



Gambar 18. Descriptive Statistic

Maka:

- Rerata usia C1=  $30,04 + (0,18619 * 6,238) = 31,2$  th
- Rerata usia C2=  $30,04 + (1,08919 * 6,238) = 36,8$  th
- Rerata usia C3=  $30,04 + (0,49774 * 6,238) = 33,1$  th

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore: Usia	0,504	2	608	23	9,047	,001
Zscore: Jumlah anak	1,816	2	829	23	1,954	,164
Zscore: Penghasilan rata-rata per bulan	9,173	2	289	23	31,706	,000
Zscore: Jumlah jam membaca koran setiap minggu	0,967	2	481	23	14,480	,000
Zscore: Jumlah jam menonton TV setiap minggu	1,128	2	989	23	1,136	,339
Zscore: Jumlah motor yang dipunyai	5,294	2	827	23	6,448	,002
Zscore: Jumlah mobil yang dipunyai	4,697	2	678	23	6,923	,004
Zscore: Jumlah kartu kredit atau ATM yang dimiliki	7,944	2	398	23	20,052	,000
Zscore: Tingkat pembelian barang setiap minggu	10,732	2	154	23	69,790	,000
Zscore: Tingkat pengeluaran bulanan	8,880	2	332	23	26,128	,000
Zscore: Jumlah jam kerja setiap minggu	7,868	2	401	23	19,668	,000
Zscore: Jumlah jam berbelanja setiap minggu	7,755	2	413	23	16,786	,000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Gambar 19. Hasil Uji Anova

Analisis perbedaan dengan Anova, didapatkan bahwa:

- Tingkat pembelian merupakan yang paling berbeda antar cluster (mempunyai F terbesar sebesar 69,790 dengan  $p = \text{Sig.} = 0,000 < 0,05$ ).
- Jumlah jam menonton TV tidak berbeda antar cluster ( $F = 1,136$  dengan  $p = \text{Sig.} = 0,339 > 0,05$ ).

Cluster Membership			
Case Number	Kota tempat tinggal	Cluster	Distance
1	Jakarta utara	1	2.699
2	Bandung	1	2.212
3	Tegal	3	2.089
4	Semarang	3	1.788
5	Yogya	3	2.897
6	Solo	3	1.529
7	Banjaregara	3	2.072
8	Medun	3	2.489
9	Pekalongan	3	2.831
10	Jepara	3	2.421
11	Blora	3	2.097
12	Kerawang	3	1.722
13	Jakarta selatan	1	3.412
14	Bandung	1	2.129
15	Surabaya	2	3.199
16	Nagelang	3	1.495
17	Parakan	3	1.883
18	Tuban	3	1.571
19	Jakarta selatan	2	3.115
20	Semarang	2	1.820
21	Surabaya	2	2.874
22	Surabaya	2	2.282
23	Ciamis	1	3.587
24	Pas	3	2.359
25	Cepu	3	1.815
26	Wonogri	3	1.815

Distances between final Cluster Centers			
Cluster	1	2	3
1			
2	2.769		4.182
3	4.182	5.262	

Number of Cases in each Cluster	
Cluster	1
1	8.000
2	8.000
3	15.000
Valid	28.000
Missing	.000

Gambar 20. Hasil Cluster Membership

- Data bisa juga dilihat di *Data View*, kolom terakhir

	ZBel	ZKonsumsi	ZKeja	ZShipping	OCL_1	OCL_2
1	13279	2,71299	46532	-13879	1	2,69893
2	13279	14736	46532	24105	1	2,22208
3	-63444	-77627	-86710	-70854	3	2,08897
4	90001	14736	1,25477	1,76040	2	1,75801
5	-25082	-36577	-06765	-13879	3	2,89705
6	-25082	-46839	-06765	-32871	3	1,52943
7	-63444	-87889	-86710	-1,08838	3	2,07236
8	-63444	-57102	-06765	05113	3	2,49824
9	-63444	-46839	73181	62089	3	2,83133
10	-1,01805	-77627	-92039	-1,08838	3	2,42051
11	-1,01805	-70785	-1,13368	-89846	3	2,09662
12	-1,01805	-70785	-86710	-1,08838	3	1,72225
13	-25082	1,66674	2,06423	1,95032	1	3,41201
14	51640	66049	73181	05113	1	2,12884
15	90001	1,66674	1,53126	1,38056	2	3,10903
16	-1,01805	-53681	-92039	-32871	3	1,49536
17	-63444	-77627	-86710	-89846	3	1,88276
18	-63444	-87889	-1,08029	-1,46822	3	1,57130
19	1,66724	1,66674	57192	62089	2	3,11509
20	2,05085	66049	1,53126	1,38056	2	1,81975
21	2,05085	-1,6052	1,26477	1,76040	2	2,67373
22	2,05085	45524	73181	81081	2	2,26211
23	-63444	66049	-92039	-89846	1	3,95705
24	-63444	-77627	-1,02699	-51862	3	2,35876

Gambar 21. Final Data View

## DAFTAR PUSTAKA

- 10060119034, D. A. P. S., & Nur Azizah Komara Rifai. (2024). Penerapan Metode Simple K-Medoid (SKM) dalam Mengklasterkan SMA di Kabupaten Karanganyar Tahun 2023. *Bandung Conference Series: Statistics*, 4(1). <https://doi.org/10.29313/bcss.v4i1.11098>
- Bustami, B. (2015). Analisis dan Perbandingan Kualitas Pengelompokan Dokumen (Document Clustering) Dengan Menggunakan Metode K-Means Dan K-Medians. *Elkawnie: Journal of Islamic Science and Technology*, 1(2), Article 2. <https://doi.org/10.22373/ekw.v1i2.536>
- Cluster analysis*. (2016, March 18). SlideShare. <https://www.slideshare.net/slideshow/cluster-analysis-59734953/59734953>
- Cupak, A. (2022). Regionalization of low flow for chosen catchments of the upper Vistula river basin using non-hierarchical cluster analysis. *Időjárás*, 126(1). <https://doi.org/10.28974/idojaras.2022.1.2>
- Expert. (2024, August 11). *Teknik Analisis Cluster: Mengelompokkan Data dengan Cara yang Bermakna untuk 2024-2025*. Editverse. <http://www.editverse.com/id/teknik-analisis-klaster-mengelompokkan-data-secara-bermakna-untuk-tahun-2024-2025/>
- Faisal, S. A., & Rifai, N. A. K. (2023). Penerapan Metode Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokan Indikator Indeks Pembangunan Manusia Tahun 2021 di Jawa Barat. *Bandung Conference Series: Statistics*, 3(1). <https://doi.org/10.29313/bcss.v3i1.6327>
- Fauziyah, W. M., & Achmad, A. I. (2023). Penerapan Analisis Cluster Hybrid untuk Pengelompokan Kabupaten/Kota di Provinsi

- Jawa Barat Berdasarkan Indikator Kemiskinan Tahun 2022. *Bandung Conference Series: Statistics*, 3(2).  
<https://doi.org/10.29313/bcss.v3i2.8610>
- Fitriyani, E. N., & Iswani Achmad, A. (2023). Penerapan Analisis K-Medoids Cluster untuk Mengelompokkan Wilayah di Provinsi Jawa Barat Berdasarkan Fasilitas Kesehatan Tahun 2021. *Bandung Conference Series: Statistics*, 3(2).  
<https://doi.org/10.29313/bcss.v3i2.8080>
- Koesnadi, G. L., Handayani, K. T., Marwanda, N. D., Qomaryah, P. M., Amelia, D., Mardianto, M. F. F., & Ana, E. (2023). Pengelompokan Provinsi Berdasarkan Prioritas Potensi Sektor Maritim Indikator Blue Economy Menggunakan Analisis Cluster Average Linkage. *Jurnal Sains Matematika Dan Statistika*, 9(1).  
<https://doi.org/10.24014/jsms.v9i1.20291>
- Ls, D., Lesnussa, Y. A., Talakua, M. W., & Matdoan, M. Y. (2021). Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Pendidikan dengan Menggunakan Metode Ward. *Jurnal Statistika Dan Aplikasinya*, 5(1). <https://doi.org/10.21009/jsa.05105>
- Matdoan, M. Y., & Delsen, M. S. N. V. (2020). Penerapan Analisis Cluster Dengan Metode Hierarki Untuk Klasifikasi Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia. *Statmat : Jurnal Statistika Dan Matematika*, 2(2).  
<https://doi.org/10.32493/sm.v2i2.4740>
- Park, H., Lim, C.-M., Bhardwaj, V., & Kim, Y.-K. (2011). Benefit segmentation of TV home shoppers. *International Journal of Retail & Distribution Management*, 39(1).  
<https://doi.org/10.1108/095905511111104459>
- Rosmini, R., Fadlil, A., & Sunardi, S. (2018). Implementasi Metode K-

Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah. *It Journal Research and Development*, 3(1).

[https://doi.org/10.25299/itjrd.2018.vol3\(1\).1773](https://doi.org/10.25299/itjrd.2018.vol3(1).1773)

Thaha, I., Wahyuni, M. S., Sutamrin, S., & Mu'adz, A. M. (2023).

Perbandingan Penggerombolan Tingkat Pencemaran Udara dengan K- Medoid dan CLARA berdasarkan Indeks Kualitas Udara (IKU) di Provinsi Sulawesi Selatan. *Proximal: Jurnal Penelitian Matematika Dan Pendidikan Matematika*, 7(1).

<https://doi.org/10.30605/proximal.v7i1.3074>

UMA. (2022, July 26). Mengenal Analisis Cluster: Definisi, Metode dan

Cara Penggunaan. *Lembaga Penelitian Dan Pengabdian Masyarakat*.

[https://lp2m.uma.ac.id/2022/07/26/mengenal-analisis-](https://lp2m.uma.ac.id/2022/07/26/mengenal-analisis-cluster-definisi-metode-dan-cara-penggunaan/)

[cluster-definisi-metode-dan-cara-penggunaan/](https://lp2m.uma.ac.id/2022/07/26/mengenal-analisis-cluster-definisi-metode-dan-cara-penggunaan/)

## BIODATA PENULIS



**Ely Mulyati, ST., MT**

Dosen Program Studi Teknik Sipil  
Fakultas Sains dan Teknologi Universitas Bina Darma

Penulis lahir di Palembang tanggal 24 Agustus 1977 adalah dosen tetap Program Studi Teknik Sipil Fakultas Saint Dan Teknologi Universitas Bina Darma. Menyelesaikan S-1 di Universitas Muhammadiyah Yogyakarta dan S-2 di Universitas Sriwijaya.

Mengajar merupakan suatu pekerjaan dan bagian perjalanan hidup yang menyenangkan dengan segala tantangan yang dihadapi, namun waktu dan pengalaman mengajarkan untuk menjadi lebih baik dan menjalankan perintah NYA. Dengan memegang prinsip “Sebaik-baiknya manusia adalah manusia yang bermanfaat bagi orang banyak”. Semoga Buku ini dapat bermanfaat dan digunakan oleh berbagai pihak yang membutuhkannya, dan membawa barokah bagi para penulis semua.

Penulis dapat dihubungi melalui e-mail: [ely.mazpar@gmail.com](mailto:ely.mazpar@gmail.com)

## **PROFIL PENULIS**



**M. Rosihan Arsyad**  
**Mahasiswa Teknik Sipil**

## **PROFIL PENULIS**



**Sri Maryati, SST., M.Kes**  
**Dosen Prodi Pendidikan Bidan Tahap Sarjana**  
**STIKes Budi Luhur Cimahi**

Penulis Bernama lengkap Sri Maryati, lahir 36 tahun yang lalu di Ciamis dan sekarang menetap di Ciparay. Riwayat Pendidikan terakhir yaitu S2 Manajemen Kesehatan Ibu dan Anak di Universitas Diponegoro, lulus pada tahun 2014. Penulis merupakan salah satu dosen yang bekerja di salah satu Institusi Swasta di Kota Cimahi yaitu STIKes Budi Luhur dan mengajar mata kuliah promosi kesehatan dan metodologi penelitian. Ini merupakan karya buku ketiganya. Penulis sebelumnya telah mengembangkan diri dalam melakukan Tri dharma Perguruan Tinggi yaitu dalam pengajaran, penelitian dan pengabdian kepada masyarakat.

## **PROFIL PENULIS**



**Leni Gustina, SE, MM**  
**Dosen Prodi Manajemen**  
**Universitas Putra Indonesia YPTK Padang**

Lahir di Padang, 30 Agustus 1980. Dosen di Universitas Putra Indonesia YPTK Padang dari tahun 2017 sampai sekarang. Menyelesaikan S-1 di Sekolah Tinggi Ilmu Ekonomi Keuangan Perbankan dan Pembangunan (STIE-KBP Padang) tahun 2004, S-2 Konsentrasi Manajemen Sumber Daya Manusia di Universitas Putra Indonesia YPTK Padang tahun 2009. Pernah Bekerja di Instansi Pemerintahan Kota Padang sebagai tenaga Honorer di Dinas Pasar Kota Padang sejak 2006-2019.

## **PROFIL PENULIS**



**Linda Prasetyaning Widayanti, S.KM., M.Kes**  
**Dosen Fakultas Psikologi dan Kesehatan**  
**Universitas Islam Negeri Sunan Ampel**

Penulis lahir di Mojokerto pada bulan April 1987. Pada tahun 2009 menyelesaikan Pendidikan S1 di Fakultas Kesehatan Masyarakat Universitas Airlangga Surabaya. Pada tahun 2010 bekerja sebagai staf HRD di RS Muhammadiyah Hasanah Mojokerto. Setahun kemudian diterima di Stikes Mojokerto sebagai dosen muda. Pada tahun 2012 menamatkan Pendidikan S2 di Fakultas Kesehatan Masyarakat Universitas Airlangga Surabaya dan diterima menjadi dosen Pegawai Negeri Sipil di wilayah Kementrian Agama UIN Sunan Ampel Surabaya pada tahun 2014. Penulis aktif menjadi anggota PERSAKMI (Persatuan Sarjana Kesehatan Masyarakat Indonesia) sebagai organisasi yang bergerak di bidang kesehatan masyarakat. Saat ini penulis tinggal di Surabaya dan aktif menjadi Dosen di Fakultas Psikologi dan Kesehatan Program Studi Gizi UIN Sunan Ampel Surabaya.

Judul Penelitian:

---

1. Unveiling The Vital Role of Flavonoids in Combating Bad Cholesterol: A Scientific Approach
2. Determining the Effect of Orange Peel Extract in Water on Total Cholesterol Fluctuations in HFD-Induced Mice
3. Risiko Tinggi Dislipidemia pada Remaja Obesitas: Menggali Faktor dan Implikasinya terhadap Kesehatan Kardiovaskular
4. A Review Article: Control Caloric Intake in Autoimmune Survivors
5. Pengetahuan Gizi Seimbang dan Olahraga Terhadap Kenaikan Berat Badan Selama Pandemi COVID-19 pada Remaja
6. Pengaruh Indeks Massa Tubuh, Asupan Lemak, Dan Protein Terhadap Kadar Kolesterol Remaja
7. The effectiveness of zinc micronutrients from pumpkin (*Cucurbita moschata* D) extract on the testosterone levels of mice (*Mus musculus* L)

## PROFIL PENULIS



**Dr. Hwihanus, SE., MM., CMA**  
**Dosen Fakultas Ekonomi dan Bisnis**  
**Universitas 17 Agustus 1945 Surabaya**

Lahir di Surabaya, Jenjang Pendidikan S1 ditempuh di Fakultas Ekonomi Universitas 17 Agustus 1945 Surabaya lulus tahun 1992. Pendidikan S2 Magister Manajemen Fakultas Ekonomi Universitas 17 Agustus 1945 Surabaya, lulus tahun 1997 dan Pendidikan S3 Doktor Ilmu Ekonomi Fakultas Ekonomi dan Bisnis Universitas 17 Agustus 1945 Surabaya lulus tahun 2020. Saat ini menjabat sebagai Dosen Fakultas Ekonomi Universitas 17 Agustus 1945 Surabaya sejak 1993 sampai sekarang. Beberapa buku yang sudah di terbitkan:

1. Antecedents Value of the Firm, Penerbit LPPM Untag Surabaya, tahun 2019
2. Akuntansi Internasional Perspektif Aktivitas, Penerbit Revka Prima Media, tahun 2019
3. Pembelajaran Bahasa Mandarin (xue Hanyu), Penerbit Zahira Media Publisher, tahun 2021
4. Panduan/Petunjuk Manajemen Badan Usaha Milik Desa Simoketawang pada desa Ramah Anak dalam Bingkai Patriot Merah Putih, Penerbit CV. Global Edukasi Teknologi, Tahun 2022
5. Ilmu Akuntansi, Penerbit PT. Literasi Nusantara Abadi Group, Tahun 2023

6. Pengantar Akuntansi Perusahaan, Penerbit CV. Gita Lentera, 2023
7. Ilmu Administrasi Publik (Suatu Pengantar), Penerbit PT. Adikarya Pratama Globalindo, 2023
8. Pengantar Bisnis: Perspektif Ekonomi Global, Penerbit CV. Gita Lentera, 2023
9. Pengantar Ilmu Manajemen: Organisasi dan Perkembangannya, Penerbit CV. Gita Lentera, 2023
10. Akuntansi Keuangan Menengah, Penerbit CV. Gita Lentera, 2024
11. Akuntansi Keuangan Lanjutan, Penerbit CV. Gita Lentera, 2024
12. Sistem Informasi Manajemen, Penerbit CV. Gita Lentera, 2024
13. Akuntansi Sektor Publik, Penerbit CV. Gita Lentera, 2024
14. Akuntansi Manajemen Strategis, Penerbit CV. Gita Lentera, 2024
15. Akuntansi Manajemen, Penerbit CV. Gita Lentera, 2024

## PROFIL PENULIS



**Dr.Ir. La Ode Muhamad Nurrahmad Arsyad, S.T.,M.T.,IPM.,ASEAN Eng.,APEC Eng**

Dosen Program Studi Teknik Sipil  
Fakultas Teknik Universitas Halu Oleo

Penulis merupakan Dosen di Jurusan Teknik Sipil Fakultas Teknik Universitas Halu Oleo Sejak Tahun 2008. Sebagai seorang yang sepenuhnya mengabdikan dirinya sebagai dosen, selain pendidikan formal yang telah ditempuhnya penulis juga mengikuti berbagai pelatihan dan kompetensi untuk meningkatkan kinerja dosen, khusus di bidang pengajaran, penelitian dan pengabdian kepada masyarakat. Penulis juga aktif pada berbagai bidang organisasi masyarakat, antara lain Ikatan Ahli Bandar Udara Indonesia (IABI) dan Forum Studi Transportasi Antar Perguruan Tinggi (FSTPT). penulis bekerja sebagai tenaga ahli konsultan Perencanaan Kebandarudaraan, perkerasan jalan raya, penulis juga aktif melakukan penelitian yang diterbitkan di berbagai jurnal nasional maupun internasional, khususnya di bidang transportasi.

Penulis dapat dihubungi melalui e-mail: [arsyadjr@uho.ac.id](mailto:arsyadjr@uho.ac.id)

# PENGANTAR SPSS

## Teori, Implementasi dan Interpretasi

*"Statistical analysis is one of the key elements of any research, and SPSS has been the go-to software for both students and professionals alike due to its versatility and user-friendly interface" (Pallant, 2020).*

**Buku ini merupakan panduan komprehensif untuk memahami konsep-konsep dasar dalam analisis data menggunakan SPSS, serta memberikan langkah-langkah praktis dalam mengimplementasikan teknik-teknik statistik yang penting dalam penelitian. Dimulai dengan pengenalan mengenai SPSS sebagai perangkat lunak statistik yang populer, buku ini menjelaskan berbagai jenis variabel dan data yang dapat dikelola di dalam SPSS. Pembaca akan diajarkan bagaimana cara mentransformasikan data, baik itu untuk keperluan pengolahan data awal maupun untuk meningkatkan kualitas analisis. Selanjutnya, buku ini mengupas berbagai teknik analisis deskriptif, yang memungkinkan pengguna untuk menggambarkan dan menyimpulkan data dengan mudah.**

**Buku ini juga membahas berbagai teknik statistik lanjutan yang dapat digunakan untuk menguji validitas dan reliabilitas data, uji asumsi klasik, serta uji hipotesis parametrik dan non-parametrik. Untuk penelitian yang lebih mendalam, pembaca akan diberikan pengetahuan tentang regresi linier, regresi logistik, serta analisis multivariat seperti analisis faktor dan kluster. Setiap bab dilengkapi dengan ilustrasi praktis dan contoh kasus nyata, memungkinkan pembaca untuk mengimplementasikan setiap teknik dalam penelitian mereka. Dengan pendekatan yang sistematis dan mudah diikuti, buku ini sangat cocok bagi mahasiswa, peneliti, dan praktisi yang ingin meningkatkan pemahaman dan kemampuan mereka dalam menggunakan SPSS untuk analisis data yang lebih tepat dan akurat.**



**Penerbit**  
**Gita Lentera**

Office1: Perm. Permata hijau regency blok F/1 kelurahan Pisang  
kecamatan Pauh kota Padang, Sumatera Barat  
Office2: Jl Weling no 120 Gejayan, Yogyakarta  
Cp. Admin: +62323-8699-7194  
git4lentera4@gmail.com www.git4lentera.com

ISBN 978-634-7072-09-2



Anggota IKAPI  
No. 042/SBA/2023