

DASAR-DASAR STATISTIKA

(Konsep dan Metode Analisis)



Penulis :

Dr. I Wayan Terimajaya, S.E., M.M., M.H

Ni Luh Sintya Dewi, S.Pd., M.Pd

Dr. Ture Simamora, S.Pt., M.Si

Loso Judijanto, S.Si., M.M., M.Stat

Radha Krisnamurti Sigamura, M.Pd

+5.2% Dr. Nurhayati, S.E., M.E

Sri Yani Kusumastuti

Dr. Raymond Bahana, S.T., M.Sc

Dr. Laurensius Laka, M.Psi., Psikolog

Anindita Henindya Permatasari, S.Si, M.Mat

Admi Salma, M.Si

Ni Putu Sudarsani, SE., M.Si

Ardiansyah, S.Pt., M.Si

Basri, S.Sos.I., M.M

SONPEDIA.COM

PT. Sonpedia Publishing Indonesia

DASAR-DASAR STATISTIKA

(Konsep dan Metode Analisis)

Penulis :

Dr. I Wayan Terimajaya, S.E., M.M., M.H
Ni Luh Sintya Dewi, S.Pd., M.Pd
Dr. Ture Simamora, S.Pt., M.Si
Loso Judijanto, S.Si., M.M., M.Stat
Radha Krisnamurti Sigamura, M.Pd
Dr. Nurhayati, S.E., M.E
Sri Yani Kusumastuti
Dr. Raymond Bahana, S.T., M.Sc
Dr. Laurensius Laka, M.Psi., Psikolog
Anindita Henindya Permatasari, S.Si, M.Mat
Admi Salma, M.Si
Ni Putu Sudarsani, SE., M.Si
Ardiansyah, SPT., MSi
Basri, S.Sos.l., M.M

Penerbit:

SONPEDIA
Publishing Indonesia

DASAR-DASAR STATISTIKA
(Konsep dan Metode Analisis)

Penulis :

Dr. I Wayan Terimajaya, S.E., M.M., M.H
Ni Luh Sintya Dewi, S.Pd., M.Pd
Dr. Ture Simamora, S.Pt., M.Si
Loso Judijanto, S.Si., M.M., M.Stat
Radha Krisnamurti Sigamura, M.Pd
Dr. Nurhayati, S.E., M.E
Sri Yani Kusumastuti
Dr. Raymond Bahana, S.T., M.Sc
Dr. Laurensius Laka, M.Psi., Psikolog
Anindita Henindya Permatasari, S.Si, M.Mat
Admi Salma, M.Si
Ni Putu Sudarsani, SE., M.Si
Ardiansyah, SPT., MSi
Basri, S.Sos.l., M.M

ISBN : 978-623-8634-41-5

Editor :

Sepriano & Efitra

Penyunting :

Inayah Uzma

Desain sampul dan Tata Letak :

Yayan Agusdi

Penerbit :

PT. Sonpedia Publishing Indonesia

Redaksi :

Jl. Kenali Jaya No 166 Kota Jambi 36129 Tel +6282177858344

Email: sonpediapublishing@gmail.com

Website: www.buku.sonpedia.com

Anggota IKAPI : 006/JBI/2023

Cetakan Pertama, Mei 2024

Hak cipta dilindungi undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara
Apapun tanpa ijin dari penerbit

KATA PENGANTAR

Puji syukur kepada Tuhan yang Maha Esa, telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan buku ini dengan baik. Buku ini berjudul “***DASAR-DASAR STATISTIKA : Konsep dan Metode Analisis***”. Tidak lupa kami ucapkan terima kasih bagi semua pihak yang telah membantu dalam penulisan dan penerbitan buku ini.

Statistika adalah ilmu yang sangat penting dalam berbagai bidang ilmu pengetahuan, mulai dari ekonomi, sosial, teknik, hingga ilmu kesehatan. Kemampuan untuk memahami dan menerapkan konsep-konsep statistika dasar merupakan modal penting dalam analisis data dan pengambilan keputusan yang didasarkan pada data. Oleh karena itu, buku ini dirancang untuk memberikan pemahaman yang kuat tentang konsep-konsep dasar statistika dan metode analisis.

Buku ini adalah panduan komprehensif yang membahas konsep-konsep dasar statistika dan metode analisis yang penting dalam penelitian ilmiah. Topik-topik utama yang dibahas meliputi pengantar statistika, tujuan dan ruang lingkupnya, serta sumber dan kualitas data. Buku ini juga mengajarkan cara mendeskripsikan data, memahami konsep probabilitas, dan menggunakan distribusi probabilitas dalam analisis statistik. Pembahasan mendalam mengenai pengujian hipotesis, konsep dasar regresi dan korelasi, serta metode analisis regresi linear sederhana dan berganda memberikan dasar yang kuat bagi pembaca untuk memahami hubungan antar variabel.

Selain itu, buku ini juga membahas penggunaan software statistik SPSS, yang memudahkan pembaca dalam menerapkan analisis statistik secara praktis. Dengan berbagai contoh nyata, buku ini dirancang untuk membantu pembaca, terutama pemula, memahami

dan menerapkan konsep serta teknik statistik. Tujuannya adalah agar pembaca dapat melakukan analisis data secara efektif, meningkatkan kualitas penelitian, dan membuat keputusan yang didasarkan pada data yang valid dan reliabel.

Buku ini mungkin masih terdapat kekurangan dan kelemahan. Oleh karena itu, saran dan kritik para pemerhati sungguh penulis harapkan. Semoga buku ini memberikan manfaat dan menambah khasanah ilmu pengetahuan.

Bandung, Mei 2024
Penulis

DAFTAR ISI

KATA PENGANTAR	ii
DAFTAR ISI.....	iv
BAGIAN 1 PENGANTAR STATISTIKA.....	1
A. PENGERTIAN STATISTIKA.....	1
B. BALASAN MEMPELAJARI STATISTIKA	3
C. PEMANFAATAN STATISTIKA DALAM PENELITIAN SOSIAL	5
D. JENIS-JENIS STATISTIKA	15
BAGIAN 2 TUJUAN DAN RUANG LINGKUP STATISTIKA.....	17
A. DEFINISI STATISTIK & STATISTIKA.....	17
B. TUJUAN STATISTIKA.....	21
C. KLASIFIKASI STATISTIKA	22
D. FUNGSI DAN KEGUNAAN STATISTIKA.....	24
E. PERANAN STATISTIKA DALAM PENELITIAN	26
BAGIAN 3 METODE PENGUMPULAN DATA.....	28
A. PENGERTIAN, TUJUAN DAN BENTUK METODE PENGUMPULAN DATA.....	28
B. METODE SURVEI	31
C. METODE OBSERVASI.....	32
D. METODE WAWANCARA.....	35
E. METODE DOKUMENTASI	39
BAGIAN 4 SUMBER DATA DALAM STATISTIKA.....	42
A. DATA PRIMER (SENSUS, SURVEI, <i>BIG DATA</i>)	42
B. DATA SEKUNDER.....	46
C. JENIS DATA (DATA KUANTITATIF DAN KUALITATIF)	50

D. PENYAJIAN DATA (DATA <i>CROSS SECTION</i> , <i>TIME SERIES</i> ,	
DAN PANEL)	54
E. SKALA DATA (NOMINAL, ORDINAL, INTERVAL, RASIO)	59
BAGIAN 5 KUALITAS DATA	65
A. PENGERTIAN KUALITAS DATA	65
B. DATA	67
C. KARAKTERISTIK KUALITAS DATA	69
D. DATA STATISTIKA	73
BAGIAN 6 DESKRIPSI DATA	77
A. KONSEP DATA	77
B. ANALISIS DESKRIPSI	80
C. DISTRIBUSI FREKUENSI	82
D. GEJALA PUSAT	88
E. UKURAN PENYEBARAN	91
F. UKURAN KEMENCENGAN	94
G. UKURAN KERUNCINGAN.....	96
BAGIAN 7 KONSEP PROBABILITAS	98
A. DEFINISI PROBABILITAS, EKSPERIMEN, HASIL (<i>OUTCOME</i>),.....	
DAN KEJADIAN (<i>EVENTS</i>)	98
B. PENDEKATAN PENGHITUNGAN PROBABILITAS.....	100
C. PROBABILITAS MENGGUNAKAN ATURAN PENJUMLAHAN.....	105
D. PROBABILITAS MENGGUNAKAN ATURAN PERKALIAN.....	108
E. PROBABILITAS MENGGUNAKAN TABEL KONTIGENSI.....	110
F. PROBALITAS MENGGUNAKAN TEOREMA BAYES	111
G. PRINSIP PENGHITUNGAN	114

BAGIAN 8	DISTRIBUSI PROBABILITAS DISKRIT	118
A.	PENGERTIAN DISTRIBUSI PROBABILITAS DISKRIT.....	118
B.	DISTRIBUSI BINOMIAL	118
C.	DISTRIBUSI BERNOULLI.....	120
D.	DISTRIBUSI MULTINOMIAL.....	120
E.	DISTRIBUSI POISSON.....	122
F.	DISTRIBUSI GEOMETRIK.....	123
G.	DISTRIBUSI BINOMIAL NEGATIF (PASCAL).....	125
H.	DISTRIBUSI SERAGAM DISKRIT (DISCRETE UNIFORM).....	127
I.	DISTRIBUSI HIPERGEOMETRIK.....	128
BAGIAN 9	PENGUJIAN HIPOTESIS.....	131
A.	PENGEMBANGAN HIPOTESIS.....	131
B.	KONSEP DASAR PENGUJIAN HIPOTESIS.....	133
BAGIAN 10	PENGANTAR BASIS DATA.....	146
A.	ANALISIS KORELASI	146
B.	ANALISIS REGRESI	157
C.	PERBEDAAN ANTARA KORELASI DAN REGRESI	163
BAGIAN 11	REGRESI LINEAR SEDERHANA.....	165
A.	PENGANTAR REGRESI LINEAR SEDERHANA.....	165
B.	TUJUAN DAN MANFAAT REGRESI LINEAR SEDERHANA	167
C.	MODEL REGRESI LINEAR SEDERHANA	168
D.	PENDUGA PARAMETER REGRESI LINEAR SEDERHANA.....	169
E.	ANALISIS KERAGAMAN REGRESI.....	172
F.	SELANG KEPERCAYAAN.....	173
G.	KOEFISIEN DETERMINASI.....	174

H.	ASUMSI REGRESI LINEAR SEDERHANA	175
BAGIAN 12 REGRESI LINEAR BERGANDA		177
A.	PENGANTAR.....	177
B.	PENDUGAAN TITIK KOEFISIEN REGRESI LINEAR BERGANDA....	178
C.	KOEFISIEN DETERMINASI REGRESI LINIER BERGANDA (R^2)	180
D.	KOEFISIEN KORELASI GANDA.....	182
E.	ASUMSI-ASUMSI MODEL REGRESI LINIER BERGANDA.....	182
F.	ESTIMASI PARAMETER MODEL REGRESI LINIER BERGANDA.....	183
G.	PENGUJIAN PARAMETER MODEL REGRESI LINIER	
	BERGANDA.....	183
H.	PELANGGARAN-PELANGGARAN TERHADAP ASUMSI.....	
	REGRESI LINIER BERGANDA.....	186
BAGIAN 13 PENGANTAR ANALISI VARIAN (ANOVA)		191
A.	KONSEP DASAR ANOVA.....	191
B.	ASUMSI ANOVA	192
C.	JENIS-JENIS ANOVA.....	194
D.	SUMBER KERAGAMAN (VARIASI) ANOVA	194
E.	INTERPRETASI HASIL DAN PENGUJIAN HIPOTESIS DALAM	
	ANOVA	195
F.	PROSEDUR ANOVA	195
G.	KELEBIHAN DAN KEKURANGAN ANOVA.....	199
H.	UJI LANJUT (POST HOC TESTS).....	199
I.	PENGGUNAAN ANOVA DALAM PENELITIAN	200
J.	APLIKASI UNTUK ANALISIS ANOVA.....	201

BAGIAN 14 PENGANTAR SPSS.....	202
A. PENGERTIAN DAN SEJARAH SPSS	202
B. PENELITIAN STATISTIK DAN SPSS.....	204
C. PENYUSUNAN DATA	206
DAFTAR PUSTAKA	215
TENTANG PENULIS	225

BAGIAN 1

PENGANTAR STATISTIKA

A. PENGERTIAN STATISTIKA

Statistika memiliki sejarah yang panjang dalam sejarah peradaban manusia. Pada jaman sebelum masehi, bangsa-bangsa di Mesopotamia, Mesir, dan Cina telah mengumpulkan data statistik untuk memperoleh informasi tentang berapa pajak yang harus dibayar oleh setiap penduduk, berapa hasil pertanian yang mampu diproduksi, berapa cepat atlet lari dan sebagainya. Pada abad pertengahan, lembaga Gereja menggunakan statistika untuk mencatat jumlah kelahiran, kematian, dan perkawinan (Purwanto, 2018).

Statistika yang dimulai dengan pengumpulan dan penyajian data, kemudian semakin berkembang dengan ditemukannya teori probabilitas dan teori pengambilan keputusan yang sangat dibutuhkan dalam kehidupan sehari-hari agar efisien pada semua bidang, baik sosial, ekonomi, politik, manajemen, maupun teknik. Pada tahun 1950-an, statistika memasuki wilayah pengambilan keputusan melalui proses generalisasi dan peramalan dengan memperhatikan faktor risiko dan ketidakpastian. Kenyataan itu sebenarnya sudah diramalkan oleh seorang ahli statistik H. G. Wells yang hidup pada tahun 1800-an yang mengatakan “berpikir secara statistika suatu saat akan menjadi suatu kemampuan atau keahlian

yang sangat diperlukandalam masyarakat yang efisien, seperti halnya kebutuhan manusia untuk membaca dan menulis” (Lind, 2002).

Contoh pentingnya mempelajari statistik: pemasaran berbagai macam produk dipengaruhi oleh faktor tempat, karena itu harus memindahkan produk dari tempat produksi ke tempat konsumen. Perpindahan tersebut dapat mengakibatkan kerusakan, dan biaya. Produsen sering kali harus memberikan garansi terhadap barang yang diproduksi. Berapa lama barang harus digaransi? Statistika mengajarkan tentang probabilitas dan distribusi probabilitas yang memungkinkan untuk mengetahui umur dan kualitas barang, sehingga memperkecil risiko untuk garansi. Apabila Anda membeli barang elektronik misalnya, akan diberikan garansi mulai dari 1 sampai 2 tahun. Namun demikian, produsen juga telah memperhitungkan kemungkinan harus menggaransi barang tersebut, dan biasanya tidak lebih dari 5% dari total produksi (Purwanto, 2018).

Sementara itu, statistika adalah ilmu mengumpulkan, menata, menyajikan, menganalisis, dan menginterpretasikan data menjadi informasi untuk membantu pengambilan keputusan yang efektif. Istilah statistika dapat pula diartikan sebagai metode untuk mengumpulkan, mengolah, menyajikan, menganalisis, dan menginterpretasikan data dalam bentuk angka-angka (Dajan, 1995).

B. BALASAN MEMPELAJARI STATISTIKA

Mengapa statistika perlu dipelajari? Statistika memiliki kegunaan yang luas bagi pengambilan keputusan yang tepat di berbagai bidang kehidupan. Karena, sekurang-kurangnya ada dua alasan penting untuk mempelajari statistika.

Pertama, statistika memberikan pengetahuan dan kemampuan kepada seseorang untuk melakukan evaluasi terhadap data. Dengan pengetahuan statistika yang dimiliki, seseorang dapat menerima, meragukan bahkan menolak (kebenaran, keberlakuan) suatu data.

Dalam kehidupan sehari-hari, sebenarnya kita berhadapan dengan statistika. Contoh yang dapat kita temukan dengan mudah akhir-akhir ini adalah hasil jajak pendapat (*polling*) yang disajikan oleh sejumlah media cetak, baik surat kabar maupun masalah di ibu kota. Beberapa hasil jajak pendapat tersebut melakukan inferensi berdasarkan sampel yang ditarik. Inferensi yang diperoleh dari hasil jajak pendapat tersebut beberapa ada yang valid, namun ada pula yang tidak valid. Selain masalah validitas ini kita juga perlu memperhatikan masalah sampel karena terdapat jajak pendapat yang dilakukan dengan jumlah (besar) sampel yang tidak memadai. Untuk dapat menilai kebenaran atau keberlakuan hasil (data) penelitian tersebut, kita memerlukan statistika. Meskipun demikian, statistika dapat dengan mudah digunakan untuk menyampaikan hasil yang berbeda dengan keadaan sebenarnya jika mereka yang memanfaatkan hasil atau temuan suatu penelitian tidak memahami statistika.

Kedua, bagi mahasiswa ilmu-ilmu sosial, statistika amat bermanfaat bagi dunia kerja kelak. Bidang pekerjaan yang berkaitan dengan ilmu-ilmu sosial akan menuntut Anda untuk mampu melakukan hal-hal, seperti membuat interpretasi dari temuan suatu survei atau suatu data statistik, menerapkan metode-metode analisis statistik dalam menyusun inferensi, memberikan gambaran karakteristik dari unit analisis suatu penelitian (baik individu/perorangan kelompok maupun organisasi) atau menyusun laporan berdasarkan analisis-analisis statistik. Misalkan, dalam suatu penelitian tentang "Pengaruh televisi pada penggunaan waktu luang anak di desa transisi", peneliti ingin mengetahui, berdasarkan sampel yang ditarik, proporsi dari anak yang keluarganya memiliki televisi. Selanjutnya, peneliti dapat juga melakukan perbandingan karakteristik penggunaan waktu luang anak di antara kelompok anak yang keluarganya memiliki dan yang tidak memiliki televisi di desa tersebut. Dalam analisisnya, peneliti dapat mengetahui apakah perbedaan di antara 2 kelompok anak tersebut merupakan perbedaan sebenarnya atau hanya merupakan suatu kebetulan sebagai akibat adanya variasi acak (*random*) dalam data sampel.

Manfaat atau kegunaan statistika tentu saja tidak terbatas pada lapangan kerja di bidang ilmu-ilmu sosial. Statistika juga dimanfaatkan secara luas, baik dalam bidang ilmu alam, biologi, dunia usaha, dan dunia industri. Di bidang ilmu sosial dan ilmu alam, bagi para ilmuwan menggunakan statistika setidaknya untuk 3 kepentingan, yaitu:

1. Pengumpulan data (melalui survei atau kegiatan eksperimen),
2. Pengujian hipotesis dan
3. Pengembangan teori.

Pada dunia usaha maupun industri, data sampel digunakan untuk kepentingan meramalkan penjualan dan keuntungan maupun untuk kepentingan pengawasan kualitas produksi.

C. PEMANFAATAN STATISTIKA DALAM PENELITIAN SOSIAL

Berdasarkan penjelasan mengenai alasan mempelajari statistika kita sudah dapat melihat bahwa statistika merupakan bagian yang tak terpisahkan dari berbagai bidang ilmu maupun dunia kerja. Sebelum kita membahas kegunaan statistika. Secara khusus, bagi peneliti sosial terlebih dahulu kita akan membahas secara ringkas kegunaan penelitian serta prinsip-prinsip dasar penelitian dalam ilmu-ilmu social (Arifin, 2010)

Ilmu-ilmu sosial sebagai bidang studi berbeda dalam banyak hal. Salah satu di antaranya perbedaan pada aspek dari dunia (kenyataan) sosial yang mereka pelajari. Ahli sosiologi, misalnya membahas kenyataan tersebut mulai dari kelompok-kelompok (misalnya kelas sosial), organisasi (misalnya birokrasi) sampai masyarakat (misalnya masyarakat industri) yang terdapat di dalamnya. Walaupun aspek yang menjadi fokus kajian masing-masing bidang ilmu tersebut berbeda, namun pada umumnya para

ahli tersebut memiliki kepentingan yang sama, yaitu memahami kehidupan sosial dari manusia dan berusaha mengembangkan pengetahuan (ilmu) tentang hal tersebut. Berdasarkan pengetahuan tersebut para ahli berharap dapat melakukan prediksi.

Sebagai bidang ilmu yang berusaha memahami kehidupan sosial, para ahli atau peneliti di bidang tersebut harus berusaha bersikap objektif. Namun hal ini, tidaklah mudah dilakukan dalam kajian bidang ilmu-ilmu sosial. karena kehidupan sosial manusia itu bersifat kompleks dan para ahli atau peneliti sosial mungkin saja merupakan bagian dari kehidupan sosial yang dipelajarinya. Akibatnya, dapat saja pengetahuan tentang kenyataan social yang dimiliki para ahli atau peneliti tersebut mempengaruhi usaha pemahamannya. Agar dapat bersikap objektif, sebagai salah satu prinsip dasar (ilmiah) penelitian, seorang peneliti atau para ahli diharapkan mempelajari kehidupan sosial manusia dan segala bentuk perwujudannya seperti apa adanya, bukan seperti yang diinginkan atau diharapkannya. Prinsip objektif ini memang sulit diterapkan, namun tidaklah mustahil untuk dicapai oleh para ahli atau peneliti.

Proses dalam Penelitian Ilmu-ilmu Sosial. Usaha memahami kehidupan sosial menuntut agar para ahli atau peneliti sampai tingkat tertentu melakukan konseptualisasi dan pengukuran terhadap kenyataan sosial yang ingin dipahaminya. Sebagai contoh, seorang peneliti ingin mengetahui "Faktor-faktor yang mempengaruhi minat belajar siswa SMP dan SMA". Proses untuk memahami kehidupan (gejala) sosial tersebut biasanya diawali

dengan sebuah atau beberapa pertanyaan yang berkaitan dengan beberapa aspek dari gejala sosial yang diteliti tersebut. Misalnya, apakah terdapat hubungan antara acara yang ditayangkan stasiun televisi swasta terhadap minat belajar siswa di kedua jenjang sekolah tersebut? Apakah orang tua memiliki peran dalam mempengaruhi minat belajar siswa tersebut? Hal-hal apa lagi yang diperkirakan dapat mempengaruhi minat belajar siswa SMP dan SMA? Setelah mengajukan sebuah atau serangkaian pertanyaan, peneliti menetapkan, secara spesifik, apa yang ingin ditelitinya, kapan dan di mana penelitian tersebut akan dilakukan serta bagaimana informasi yang relevan atau data dari penelitian tersebut akan dikumpulkan. Jadi, secara singkat penelitian (ilmiah) merupakan suatu kegiatan yang mutlak dilakukan jika seseorang (ahli ataupun peneliti) ingin mendapatkan pemahaman yang valid dari suatu gejala sosial, di mana penelitian tersebut sedapat mungkin diusahakan bersifat objektif.

Dalam kaitannya dengan penelitian tentang minat belajar siswa, untuk mendapatkan pemahaman yang valid, peneliti akan mengumpulkan datanya (fakta yang ditemukan di lapangan), selanjutnya melakukan analisis terhadap data tersebut, dan menyusun klasifikasi terhadap data secara sistematis. Dalam penelitian tersebut, peneliti dapat menggambarkan atau menjelaskan apakah telah terjadi pergeseran peran orang tua dalam mempengaruhi minat belajar siswa. Faktor-faktor lain apa saja yang lebih berperan dalam mempengaruhi minat belajar mereka.

Sepanjang proses penelitian, seorang peneliti dihadapkan pada dua pertanyaan penting. Pada awal proses, pertanyaan yang diajukan adalah variabel-variabel penting apa saja yang harus dipelajari? Di akhir proses, pertanyaan yang diajukan adalah penjelasan seperti apa (terhadap gejala yang diteliti) yang dianggap memadai atau memuaskan? Untuk mendapatkan jawaban dari kedua pertanyaan penting tersebut peneliti perlu melakukan pengamatan (penelitian) langsung, yaitu melalui kajian literatur atau studi kepustakaan maupun mencari informasi yang relevan tentang gejala yang diteliti dari berbagai sumber (misalnya melalui wawancara terhadap para ahli yang mempelajari gejala yang diteliti).

Masih berkaitan dengan usaha memahami kehidupan sosial, pemahaman sebagai suatu gejala adalah suatu hal yang tidak mudah dijelaskan. Pada dasarnya, seseorang (peneliti) dikatakan memiliki pemahaman terhadap suatu gejala jika orang (peneliti) tersebut memiliki fakta (data) yang “masuk akal” dan dapat dihubungkan dengan suatu kerangka teoretis. Dari fakta yang dikaitkan dengan teori yang digunakan sebagai rujukan, peneliti selanjutnya dapat mengembangkan suatu pernyataan yang merupakan penjelasan umum dari pada (*general explanatory*). Pernyataan-pernyataan yang dihasilkan jika dikaitkan dengan contoh penelitian di atas, dapat menjelaskan dengan cukup rinci tentang gejala minat belajar siswa dan faktor-faktor yang mempengaruhinya.

Dari usaha memahami kehidupan sosial, selanjutnya seorang peneliti biasanya berusaha melakukan prediksi, misalnya benar-benar memahami gejala minat belajar siswa (seperti sebagaimana kecenderungan minat belajar yang dimiliki siswa, sejauh mana peran orang tua mengalami pergeseran dalam mempengaruhi minat belajar anaknya serta faktor-faktor apa saja yang turut mempengaruhinya) seorang peneliti diharapkan dapat melakukan prediksi dengan tingkat akurasi yang cukup baik. Tingkat akurasi dalam prediksi dipengaruhi oleh beberapa hal, yaitu tingkat pengetahuan yang dimiliki peneliti mengenai gejala yang dipelajarinya, jumlah faktor yang diteliti, tingkat spesifikasi (kekhususan) dari kondisi gejala yang diteliti serta jumlah (banyaknya) lokasi (masyarakat) yang diteliti.

Usaha melakukan prediksi ini, pada dasarnya merupakan suatu proses disebarluaskannya atau diterbitkannya hasil penelitian yang membahas kecenderungan minat belajar siswa SMP dan SMA akan mendorong peneliti lain untuk meneliti kecenderungan tersebut di lokasi (lingkungan) yang karakteristiknya relatif sama, selanjutnya meneliti kembali hal tersebut pada lokasi (lingkungan) yang berbeda. Jika kecenderungan yang sama terbukti pada penelitian-penelitian selanjutnya, para peneliti yang mempelajarinya akan semakin yakin dengan temuan (hasil) penelitian tentang kecenderungan tersebut. Dengan berjalannya waktu, pengaruh dari faktor-faktor lain terus dipelajari dan metode-metode pengujian (statistika) baru juga dilakukan. Proses seperti ini terus berlangsung

karena mungkin saja kecenderungan yang ditemukan dalam penelitian terdahulu salah diartikan. Dengan demikian, kita perlu selalu menyadari bahwa hasil (temuan) dalam penelitian ilmu-ilmu sosial bersifat tentatif. Temuan tersebut selalu dapat mengalami perubahan atau perbaikan tergantung pada informasi (data) tambahan yang ditemukan lebih lanjut. Adanya kondisi temuan seperti ini mendorong timbulnya pertanyaan: “Kapan seorang peneliti sosial dapat menganggap bahwa pengetahuannya sudah cukup atau memadai?” Jawabannya prediksi karena hal tersebut dapat digunakan sebagai ukuran untuk menguji cukup tidaknya pengetahuan yang dimiliki peneliti untuk suatu gejala. Untuk itu, pertanyaan yang perlu diajukan: “Apakah peneliti telah cukup mempelajari gejala yang ditelitinya (misalnya minat belajar siswa) sehingga ia dapat menyusun prediksi temuan dari penelitian berikutnya walaupun hasil (temuan) dari penelitian berikutnya walaupun hasil (temuan) dari penelitian tersebut belum diperoleh?” Pembahasan tentang proses dalam penelitian sosial telah memperlihatkan bahwa kemampuan seseorang memahami dan melakukan prediksi suatu gejala secara ilmiah tergantung pada (kegiatan) penelitian. Suatu kegiatan penelitian biasanya diawali dengan merencanakan suatu rancangan penelitian (*research design*), yang mencakup identifikasi dari variabel-variabel yang akan diteliti dan menetapkan prosedur yang akan dilalui dalam pengumpulan dan analisis data. Jika rancangan penelitian disusun dengan baik dan diikuti dengan cermat oleh peneliti maka jawaban atas pertanyaan-pertanyaan penting yang diajukan dapat ditemukan

peneliti. Selanjutnya, rancangan penelitian tersebut berfungsi sebagai pembatas dari kegiatan penelitian yang dilakukan.

Berikut ini adalah penjelasan ringkas dari suatu kegiatan yang umumnya dilaksanakan dalam suatu penelitian sosial (penelitian kuantitatif). Bagian pertama dari kegiatan penelitian disebut dengan tahap persiapan. Mengawali kegiatannya, seorang peneliti biasanya memiliki suatu hipotesis atau dugaan yang bersifat ilmiah mengenai suatu kecenderungan, pola maupun hubungan-hubungan yang ada dari suatu gejala sosial. Hipotesis atau dugaan tersebut mencakup variabel-variabel yang ingin diteliti. Kegiatan selanjutnya, melakukan identifikasi (merumuskan batasan lagi) populasi penelitian serta menetapkan jenis instrumen pengumpulan data yang akan digunakan. Jika peneliti akan melakukan pengumpulan datanya melalui wawancara tatap muka perorangan (*face-to-face interview*) maka instrumen yang dikembangkannya adalah pedoman wawancara atau yang dikenal dengan kuesioner (daftar pertanyaan). Instrumen penelitian ini tentu akan berbeda jika pengumpulan datanya dilakukan dengan mengedarkan (mengirimkan melalui pos) daftar pertanyaan kepada para responden (*self-administered study*).

Bagian kedua dari suatu kegiatan penelitian dikenal dengan pengolahan data. Data yang dikumpulkan peneliti dapat merupakan data primer maupun data sekunder. Jawaban responden yang tercatat dalam kuesioner (daftar pertanyaan) adalah salah satu contoh data primer. Data yang diperoleh dari

institusi atau lembaga lain, misalnya kepolisian ataupun data sensus dari BPS merupakan contoh dari data sekunder. Dalam tahap pengolahan data, peneliti melakukan klasifikasi (pengelompokan). Proses ini mensyaratkan adanya suatu bentuk kuantifikasi. Hal ini diperlukan karena data yang telah dikumpulkan harus disusun dalam bentuk (cara) tertentu (reduksi data) sehingga dari data yang telah disederhanakan tersebut peneliti dapat melihat adanya kecenderungan, pola atau hubungan-hubungan.

Selanjutnya, setelah melakukan klasifikasi (penyederhanaan) data, peneliti melakukan analisis terhadap datanya. Pada kegiatan analisis inilah statistika akan memainkan peranan penting. Berdasarkan (karakteristik) data, peneliti menetapkan cara-cara perhitungan yang sesuai sehingga berdasarkan data (informasi) yang ada peneliti dapat menyusun kesimpulan atau membuat generalisasi.

Dari penjelasan mengenai proses dalam penelitian sosial, terlihat jelas bahwa statistika (sebagai alat bantu) adalah bagian yang tak terpisahkan dari ilmu-ilmu sosial. Peranan statistika tidak terbatas pada tahapan analisis data, tetapi juga digunakan dalam tahapan persiapan (penyusunan rancangan dalam penelitian) dan tahapan pengumpulan data. Tanpa alat bantu statistika menjadi amat, sulit bahkan mustahil dilakukan. Untuk itu, pada bagian berikut ini akan dibahas bagaimana para peneliti sosial memanfaatkan statistika sebagai alat bantu.

Pemanfaatan statistika oleh peneliti sosial dalam kegiatan penelitian dapat digunakan untuk 4 keperluan.

Pertama, statistika digunakan untuk menyusun, meringkas atau menyederhanakan data. Data yang diperoleh dari suatu penelitian survei dengan topik tertentu (misalnya gaya hidup kelompok profesional di perkotaan) biasanya tidak hanya besar dalam jumlah respondennya, tetapi juga mencakup banyak informasi. Terhadap data yang diperoleh tersebut, peneliti memerlukan cara untuk menyusun dan menyederhanakannya agar kegiatan penelitian tersebut dapat dilanjutkan.

Kedua, statistika digunakan untuk membantu peneliti dalam merancang (merencanakan) kegiatan survei atau eksperimen yang dapat memperkecil biaya untuk mendapatkan informasi dalam jumlah tertentu. Kegunaan statistika yang kedua ini berkaitan dengan metodologi dan inferensi (penarikan kesimpulan) secara statistika. Dalam kondisi ini, banyak waktu yang dicurahkan untuk membahas pengambilan keputusan dalam teknik penarikan sampel atau sampling dan penarikan kesimpulan (inferensi). Pada pengambilan keputusan yang pertama, peneliti perlu menetapkan jumlah data yang diperlukan dan bagaimana memperolehnya, sedangkan pada pengambilan keputusan kedua, peneliti menetapkan bagaimana inferensi (penarikan kesimpulan) tersebut dibuat. Penarikan sampel terhadap data maupun dalam kegiatan survei atau eksperimen akan menghasilkan informasi, namun pada saat yang sama memiliki konsekuensi biaya. Untuk menyesuaikan

biaya, kualitas, dan kuantitas informasi yang diperoleh dari suatu penelitian. Peneliti dapat melakukan perubahan -sampai pada tingkat tertentu- pada prosedur kegiatan penelitiannya. Perubahan yang mungkin dilakukan adalah pada bagaimana cara data (responden) dipilih dan berapa banyak informasi yang akan dikumpulkan dari setiap sumber. Modifikasi sederhana dalam pemilihan data (responden) sering kali dapat menghemat biaya dalam prosedur penarikan sampel. Pengetahuan yang dimiliki tentang statistika dan teknik-teknik dalam statistika dapat membantunya dalam pengambilan keputusan mengenai teknik penarikan sampel.

Ketiga, dengan pengetahuan mengenai statistika peneliti dapat menetapkan metode yang terbaik dalam penarikan kesimpulan (inferensi) sesuai dengan teknik penarikan sampel tertentu. Metode inferensi dalam suatu kegiatan penelitian survei atau eksperimen terdiri dari dua kemungkinan, yaitu peneliti melakukan prediksi atau membuat keputusan tentang beberapa karakteristik dari populasi atau *universe* yang menjadi pusat perhatian peneliti. Metode inferensi dalam ilmu-ilmu sosial bervariasi, untuk itu pilihan metode inferensi harus disesuaikan dengan kondisi yang ada.

Kegunaan statistika yang terakhir (keempat) adalah dalam mengukur baik tidaknya (*goodness*) sebuah inferensi (penarikan kesimpulan). Selain membantu peneliti untuk menggambarkan data, sumbangan utama statistika lainnya adalah dalam melakukan evaluasi terhadap baik tidaknya suatu inferensi. Hampir setiap

orang dapat merancang suatu metode inferensi, namun sering kali kita sulit menentukan sejauh mana inferensi tersebut (*reliabel*) baik. Selanjutnya, peneliti juga perlu menyadari bahwa ia memiliki kemungkinan untuk menarik kesimpulan yang keliru dalam penelitiannya yang bertujuan membuat keputusan berdasarkan data sampel tentang suatu karakteristik dari sebuah populasi atau *universe*.

D. JENIS-JENIS STATISTIKA

Berdasarkan definisi tersebut, statistika dibagi dalam dua jenis. Berdasarkan aktivitas yang dilakukan, statistika dapat dibedakan menjadi statistika deskriptif (*descriptive statistics*) dan statistika inferensia (*inferential statistics*).

Menurut Sugiyono (2018) Statistika deskriptif membahas cara-cara pengumpulan data, penyederhanaan angka-angka pengamatan yang diperoleh (meringkas dan menyajikan), serta melakukan pengukuran pemusatan dan penyebaran untuk memperoleh informasi yang lebih menarik, berguna, dan lebih mudah dipahami. Statistika deskriptif dalam penelitian ilmu-ilmu sosial memiliki tiga kegunaan, yaitu *pertama*, dengan statistika deskriptif, kumpulan data yang diperoleh akan tersaji dengan ringkas dan rapi serta dapat memberikan informasi inti dari kumpulan data yang ada. *Kedua*, statistika deskriptif memungkinkan peneliti menyajikan ataupun menggambarkan datanya dengan teknik grafik maupun

teknik numerik. *Ketiga*, statistika deskriptif memungkinkan peneliti mengukur dua karakteristik dari setiap respondennya dan selanjutnya meneliti hubungan di antara kedua karakteristik (variabel) tersebut.

Menurut Ghozali (2018) Statistika inferensia membahas cara menganalisis data serta mengambil kesimpulan (yang pada dasarnya berkaitan dengan estimasi parameter dan pengujian hipotesis). Metode statistika inferensia adalah metode yang berkaitan dengan analisis sebagian data sampai ke peramalan atau penarikan kesimpulan mengenai keseluruhan data. Sebagian data suatu variabel dikenal sebagai sampel, sedangkan keseluruhan datanya adalah populasi.

Dalam statistika inferensia diadakan pendugaan parameter, membuat hipotesis, serta menguji hipotesis tersebut sampai pada pembuatan kesimpulan yang berlaku umum. Metode ini sering disebut juga statistic induktif, karena kesimpulan yang ditarik didasarkan pada informasi dari sebagian data saja. Statistika inferensia memiliki 4 karakteristik, yaitu pengamatan secara acak, teknik (cara) penarikan sampel (*sampling*), data dalam bentuk angka (*numerical data*), dan tujuan umum inferensia (*common inferential objective*).

BAGIAN 2

TUJUAN DAN RUANG LINGKUP STATISTIKA

A. DEFINISI STATISTIK & STATISTIKA

Pada umumnya orang tidak membedakan antara statistika dan statistik. Secara etimologi, kata statistik berasal dari bahasa latin: “status”, yang artinya negara, atau kata “staat” dalam bahasa Belanda. Pada mulanya, kata statistik diartikan sebagai bahan keterangan atau data, baik data kuantitatif ataupun data kualitatif yang dibutuhkan oleh suatu negara. Dalam kamus Bahasa Indonesia, statistika diartikan dalam dua pemaknaan. Pertama statistika sebagai “ilmu statistik” dan kedua statistika diartikan sebagai “ukuran yang diperoleh atau berasal dari sampel”.

Sudjana (2000) menyatakan kata statistik dipakai untuk menyatakan kumpulan data, bilangan maupun non bilangan yang disusun dalam tabel dan atau diagram yang melukiskan atau menggambarkan suatu persoalan. Kata statistik juga mengandung pengertian lain yakni dipakai untuk menyatakan “ukuran” sebagai wakil dari kumpulan data mengenai sesuatu hal. Ukuran ini didapat berdasarkan perhitungan menggunakan kumpulan sebagian data yang diambil dari keseluruhan tentang persoalan tersebut. Selanjutnya dijelaskan Sudjana (2000) bahwa statistika adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisaannya dan penarikan kesimpulan

berdasarkan kumpulan data dan penganalisaan yang dilakukan. Dalam perkembangan selanjutnya, statistik diartikan sebagai kumpulan bahan keterangan yang berbentuk angka-angka (Sudijono, 2004). Pengertian statistik sebagai data yang berbentuk angka-angka masih digunakan oleh berbagai bidang kehidupan, misalnya pada bidang lalu lintas, digunakan istilah statistik kecelakaan, di sekolah-sekolah digunakan istilah data statistik.

Supardi (2013) memaknai statistik merupakan seperangkat metode yang membahas tentang:

- 1) Bagaimana cara mengumpulkan data yang dapat memberikan informasi yang optimal
- 2) Bagaimana cara meringkas, mengolah dan menyajikan data
- 3) Bagaimana cara melakukan analisis terhadap sekumpulan data sehingga dari analisis itu timbul strategi-strategi tertentu
- 4) Bagaimana cara mengambil kesimpulan dan menyarankan keputusan yang sebaiknya diambil atas dasar strategi yang ada
- 5) Bagaimana menentukan besarnya resiko kekeliruan yang mungkin terjadi jika mengambil keputusan atas dasar strategi tersebut.

Sebagai metode penelitian ilmiah, statistik memiliki tiga ciri khas yang utama yaitu:

- 1) Statistik bekerja dengan angka-angka
Statistik hanya dapat digunakan sebagai metode penelitian ilmiah, apabila data yang dikumpulkan merupakan data yang

berwujud angka-angka. Angka-angka dalam statistik mempunyai dua arti, yaitu angka yang menunjukkan jumlah dan angka yang menunjukkan nilai atau harga. Dalam arti kedua ini, angka merupakan kualitas sesuatu. Misalnya, angka kecerdasan, angka prestasi belajar, berat badan, tinggi badan, dan lain-lain.

2) Statistik bersifat objektif

Dalam hal ini, Hadi (2002) dengan tegas mengemukakan bahwa kerja statistik menutup pintu masuknya unsur-unsur subjektif yang dapat menyulap keinginan menjadi kenyataan, tidak dapat berbicara lain kecuali apa adanya. Adapun apa arti dan bagaimana menggunakan kenyataan-kenyataan statistik itu, adalah persoalan-persoalan lain yang berada di luar kompetensi statistik. Statistik sebagai alat, jelas tidak dapat berbuat lain kecuali apa adanya. Karena itu dapat dikatakan, bahwa statistik bersifat objektif. Hal ini tidak berarti bahwa setiap penelitian menggunakan metode statistik, akan menghasilkan sesuatu hasil yang benar-benar objektif. Objektif atau tidak suatu hasil kerja statistik, masih harus ditentukan oleh relevansi dari teknik yang digunakan dengan keadaan atau jenis data yang sedang dihadapi. Untuk itu diperlukan pemahaman yang mendalam dari si peneliti tentang statistik yang akan digunakan. Misalnya tidak relevan dengan nilai rata-rata hitungannya, sehingga hasilnya menjadi tidak objektif. Untuk itu, mungkin lebih tepat menggunakan median atau modus.

3) Statistik bersifat universal

Statistik bersifat universal, maksudnya bahwa statistik dapat digunakan hampir pada semua bidang penelitian. Penelitian-penelitian dalam ilmu eksakta, sosial dan budaya, semuanya dapat menggunakan statistik dengan keyakinan penuh.

Statistika beda halnya dengan statistik, statistika yang dalam bahasa Inggris “statistics” (ilmu statistik), ilmu tentang cara-cara mengumpulkan, mentabulasi dan menggolongkan, menganalisis dan mencari keterangan yang berarti dari data yang berupa angka. Statistika merupakan ilmu pengetahuan yang berhubungan dengan cara-cara mengumpulkan, menabulasi, menggolong-golongkan, menganalisis, dan mencari keterangan yang berarti dari data yang berupa bilangan-bilangan atau angka, sehingga dapat ditarik suatu kesimpulan atau keputusan tertentu. Selain itu, Statistika juga merupakan cabang ilmu matematika terapan yang terdiri dari teori dan metoda mengenai bagaimana cara mengumpulkan, mengukur, mengklasifikasi, menghitung, menjelaskan, mensintesis, menganalisis, dan menafsirkan data yang diperoleh secara sistematis. Dengan demikian, didalamnya terdiri dari sekumpulan prosedur mengenai bagaimana cara:

- Mengumpulkan data
- Meringkas data
- Mengolah data
- Menyajikan data

- Menarik kesimpulan dan interpretasi data berdasarkan kumpulan data dan hasil analisisnya

B. TUJUAN STATISTIKA

Adapun tujuan statistika yaitu :

a) Menjelaskan Hubungan Antar Variabel

Dalam data kuantitatif, setelah data dikumpulkan, kemudian data tersebut di olah dan disajikan, untuk pengambilan keputusan dalam suatu organisasi. Oleh karena itu diperlukan kemampuan untuk menyaring jumlah yang besar tersebut agar dapat menjabarkan hubungan antar variabel dalam pengambilan keputusan.

Sebagai contoh mengenai kebutuhan analisis statistik dalam menjelaskan hubungan antar variabel, jika ada seorang entrepreneur, dengan mengumpulkan data pendapatan dan biaya dapat membandingkan hasil pengembalian investasinya dalam satu periode dengan data dari periode-periode sebelumnya.

b) Membuat Keputusan Lebih Baik

Seseorang bisa menggunakan dan memberdayakan ilmu statistik dalam alat bantu sehingga menghasilkan keputusan yang lebih baik dari kondisi ketidakpastian sebelumnya.

c) Mengatasi Perubahan-Perubahan

Globalisasi memungkinkan terjadinya perubahan dalam berbagai hal. Tetapi data statistik dapat membantu dalam mengatasi perubahan yang terjadi.

d) Membuat Rencana dan Ramalan-ramalan

Perencanaan adalah serangkaian tindakan yang akan dilakukan di masa yang akan datang, dengan didasari perkiraan tentang kejadian-kejadian atau hubungan-hubungan di masa yang akan datang. Sehingga diperlukan proses atau teknik peramalan untuk memperoleh perkiraan tentang masa depan. Prosedur statistik ini dapat membantu meningkatkan proses peramalan (forecasting).

C. KLASIFIKASI STATISTIKA

Menurut (Ananda dan Fadhli, 2018) statistika dapat diklasifikasikan berdasarkan fungsi. Berdasarkan fungsinya, statistika dapat dibedakan menjadi dua bagian, yaitu:

- a. Statistika deskriptif (descriptive statistics), yaitu statistik yang mempelajari tata cara pengumpulan, penyusunan, penyajian, dan analisa data penelitian yang berwujud angka-angka, agar dapat memberikan gambaran yang teratur, ringkas dan jelas mengenai suatu gejala, keadaan dan peristiwa sehingga dapat ditarik makna tertentu. Statistik deskriptif menggunakan data numerik dan grafis untuk mengenali pola sejumlah data, merangkum informasi yang terdapat dalam data tersebut, dan menyajikan informasi sesuai dengan yang diinginkan.
- b. Statistika inferensial (inferensial statistics), yaitu statistik yang mempelajari tata cara penarikan kesimpulan mengenai karakteristik populasi, berdasarkan data kuantitatif yang

diperoleh dari sampel penelitian. Statistika inferensial juga menganalisis data sampel dan hasilnya akan digeneralisasi di tingkat populasi tempat sampel diambil. Definisi lain dari statistika inferensial adalah metode statistik mengenai tata cara penarikan kesimpulan pada sejumlah populasi berdasarkan analisa pada sampel.

Berdasarkan ruang lingkup penggunaannya, statistika dapat diklasifikasikan menjadi:

- a. Statistik pendidikan, yaitu statistik yang digunakan atau diterapkan pada bidang dan disiplin ilmu pendidikan.
- b. Statistik sosial, yaitu statistik yang digunakan atau diterapkan pada bidang dan disiplin ilmu sosial.
- c. Statistik ekonomi, yaitu statistik yang digunakan atau diterapkan pada bidang dan disiplin ilmu ekonomi.
- d. Statistik kesehatan, yaitu statistik yang digunakan atau diterapkan pada bidang dan disiplin ilmu kesehatan.
- e. Statistik pertanian, yaitu statistik yang digunakan atau diterapkan pada bidang dan disiplin ilmu pertanian.
- f. Statistik bidang ilmu/kajian lainnya.

Berdasarkan indikator yang dianalisis, statistika dapat diklasifikasikan menjadi dua jenis yaitu:

- a. Statistika parametrik, adalah statistik yang parameter populasinya harus memenuhi syarat-syarat tertentu, seperti data berskala interval atau rasio, pengambilan sampel harus random,

berdistribusi normal, memiliki varians yang homogen, dan model regresinya linear. Dalam statistik parametrik, indikator yang dianalisis yaitu parameter-parameter dari ukuran objek yang bersangkutan. Uji statistik yang dapat digunakan pada statistika parametrik, antara lain: uji-Z (1 atau 2 sampel), uji-T (1 atau 2 sampel), korelasi pearson, one or two way ANOVA test, dan analisis regresi.

- b. Statistika Nonparametrik adalah statistik yang parameter populasinya bebas dari syarat-syarat tertentu yang harus dipenuhi, sebagaimana halnya dengan statistik parametrik. Dalam statistik nonparametrik, indikator-indikator yang dianalisis adalah sisi lain dari parameter ukuran objek yang diteliti. Analisis statistika nonparametrik antara lain: uji tanda peringkat Wilcoxon dan uji U mann-withney (untuk 1 sampai 2 kelompok), uji Kruskal-Wallis (untuk kelompok lebih dari 2), uji korelasi Rank Spearman dan Kendali Tau, dan uji Chi-kuadrat

D. FUNGSI DAN KEGUNAAN STATISTIKA

1) Fungsi Statistika

Poin-poin berikut ini menjelaskan fungsi statistika, menurut (Hartanto & Yuliani, 2019). yaitu:

- a) Mengumpulkan dan menyajikan data secara sistematis.
- b) Menjadikan data yang kompleks agar menjadi sederhana sehingga mudah dipahami.

- c) Mengklasifikasikan data.
- d) Memberikan dasar dan tehnik dalam membuat perbandingan.
- e) Mempelajari hubungan antar berbagai fenomena.
- f) Untuk menunjukkan tren perilaku.
- g) Merumuskan hipotesis dan mengujinya.
- h) Untuk menarik kesimpulan rasional.

Berdasarkan penggolongan statistik, adapun fungsi statistik dapat dirinci sebagai berikut (Rosyadi, 2018) :

- a) Fungsi statistik dapat dirinci sebagai berikut.

Fungsi Statistik Deskriptif adalah untuk dapat memahami, mendeskripsikan, menerangkan data/atau peristiwa yang dikumpulkan dalam suatu penelitian dan tidak sampai pada generalisasi atau pengambilan kesimpulan mengenai keseluruhan populasi yang diselidiki.

- b) Fungsi Statistik Inferensial adalah untuk meramalkan dan mengontrol. Statistik Inferensial ini mempelajari tata cara penarikan kesimpulan mengenai keseluruhan atau populasi berdasarkan data atau gejala dan peristiwa yang ada dalam suatu penelitian.

2) Kegunaan Statistika

Adapun kegunaan statistika menurut (Rosyadi, 2018) yaitu:

- a) Memperoleh gambaran baik gambaran secara khusus maupun gambaran secara umum tentang suatu gejala, keadaan atau peristiwa.

- b) Mengikuti perkembangan atau pasang surut mengenai gejala, keadaan atau peristiwa tersebut dari waktu ke waktu.
- c) Melakukan pengujian, apakah gejala yang satu berbeda dengan gejala yang lain ataukah tidak, jika terdapat perbedaan itu merupakan perbedaan yang berarti atau perbedaan itu terjadi hanya secara kebetulan saja.
- d) Mengetahui apakah yang satu ada hubungannya dengan gejala lain.
- e) Menyusun laporan yang berupa data kuantitatif dengan teratur, ringkas dan jelas.
- f) Menarik kesimpulan secara logis, mengambil keputusan secara tepat dan mantap

E. PERANAN STATISTIKA DALAM PENELITIAN

Dalam penelitian statistik berperan sebagai alat dalam merancang penelitian, menganalisis datanya, dan menarik kesimpulan darinya. Sebagian besar studi penelitian menghasilkan sejumlah besar data mentah yang harus direduksi agar data tersebut dapat dibaca dan dipahami dengan mudah serta dapat digunakan untuk analisis lebih lanjut.

Menurut Sugiyono (2017) menjelaskan peranan statistik dalam penelitian sebagai berikut:

- 1) Alat untuk menghitung besarnya anggota sampel yang diambil dari suatu populasi. Dengan demikian jumlah sampel yang diperlukan dapat dipertanggungjawabkan.
- 2) Alat untuk menguji validitas dan reliabilitas instrumen. Sebelum digunakan untuk penelitian, perlu diuji validitas dan reliabilitasnya terlebih dahulu.
- 3) Teknik-teknik untuk menyajikan data, sehingga data lebih komunikatif. Teknik-teknik penyajian data ini seperti tabel, grafik, diagram lingkaran, dan piktogram.
- 4) Alat untuk analisis data, seperti menguji hipotesis penelitian yang diajukan. Dalam hal ini statistik yang digunakan antara lain korelasi, regresi, t-test, anava, dan lain sebagainya.

BAGIAN 3

METODE PENGUMPULAN DATA

A. PENGERTIAN, TUJUAN DAN BENTUK METODE PENGUMPULAN DATA

Pengumpulan data adalah salah satu langkah penting dalam mendukung keberhasilan penelitian. Pengumpulan data merupakan proses pengambilan data guna mendapatkan informasi yang relevan sesuai topik yang diteliti. Untuk mengumpulkan data secara akurat, peneliti harus menggunakan metode atau teknik pengumpulan data. Aan Komariah dan Djaman Satori (2011) menyatakan metode pengumpulan data adalah proses sistemasi yang dilakukan peneliti untuk mendapatkan data yang diinginkan. Menurut Riduwan dan Akdon (2010), metode pengumpulan data adalah teknik atau langkah yang dilakukan untuk memperoleh data. Dengan demikian, metode pengumpulan data adalah langkah-langkah yang tersusun secara sistematis yang dilakukan untuk mengumpulkan informasi tersebut.

Dalam penelitian, seringkali peneliti salah kaprah menyamakan instrumen penelitian dengan metode pengumpulan data. Pada hal dua konteks tersebut berbeda. Pada penelitian kuantitatif dan kuantitatif instrumen penelitian merupakan alat utama untuk memperoleh data atau informasi penelitian. Berbeda dengan metode pengumpulan data yang merupakan teknik yang digunakan

oleh seorang peneliti dalam memperoleh data penelitian. Oleh karena itu, metode pengumpulan data yang tepat memungkinkan peneliti untuk mendapatkan informasi yang akurat.

Menurut Creswell (2015), peneliti dalam menggunakan metode pengumpulan data yang baik perlu memperhatikan tahapan-tahapan pengumpulan data. Adapun tahapan-tahapan tersebut antara lain:

1. Identifikasi Informan

Melakukan identifikasi terhadap informan yang akan digunakan sebagai sumber informasi penelitian.

2. Identifikasi Lokasi Penelitian

Melakukan identifikasi lokasi penelitian guna memahami kondisi dan situasi di lokasi penelitian.

3. Mengurus Perizinan

Melakukan pengurusan perizinan sebelum dilakukan pengumpulan data. Peneliti perlu mengurus izin melaksanakan penelitian di lokasi sehingga mempermudah akses ke informan penelitian.

4. Pengadministrasian Data Penelitian

Melakukan pengadministrasian data penelitian yang telah diperoleh guna dilakukan analisis selanjutnya.

Tujuan umum metode pengumpulan data adalah untuk mengumpulkan informasi yang dapat mendukung analisis dan interpretasi hasil penelitian. Selain itu, terdapat tujuan khusus dari metode pengumpulan data yaitu:

1. Memperoleh Data yang Akurat

Metode pengumpulan data ditentukan berdasarkan pertimbangan keakuratan pengambilan data.

2. Mengumpulkan Data yang Relevan

Metode pengumpulan data ditentukan oleh peneliti untuk mengumpulkan data yang relevan dengan pertanyaan penelitian.

3. Memahami Konteks

Metode pengumpulan data ditujukan untuk memungkinkan peneliti untuk memahami konteks sosial budaya di lingkungan atau lokasi penelitian

4. Memperoleh Pemahaman yang Mendalam

Metode pengumpulan data juga ditujukan untuk memperoleh pemahaman yang mendalam tentang fenomena yang diteliti.

Peneliti mesti mampu menentukan metode pengumpulan data yang tepat sesuai topik yang diteliti. Pemilihan metode pengumpulan data bisa didasarkan pada perumusan tujuan penelitian, peninjauan pustaka yang relevan untuk mengetahui metode-metode pengumpulan data yang sudah digunakan pada penelitian yang relevan dengan topik yang diteliti, pertimbangan karakteristik populasi dan sampel penelitian, pertimbangan ketersediaan sumberdaya yang dimiliki, serta pertimbangan kelebihan dan kelemahan metode.

Menurut Sugiyono (2020), terdapat beberapa metode pengumpulan data penelitian yaitu observasi, wawancara, dokumentasi. Beberapa metode pengumpulan data yang umum

digunakan adalah survei, observasi, wawancara dan studi dokumentasi.

B. METODE SURVEI

Metode survei merupakan metode pengumpulan data penelitian dengan populasi besar dan sampel yang relatif kecil. Karakteristik metode survei antara lain informasi penelitian dikumpulkan dari kelompok besar orang guna mendeskripsikan berbagai aspek, informasi diperoleh dari pengajuan pertanyaan, dan informasi diperoleh dari sampel bukan dari populasi. Metode pengumpulan data berdasarkan survei menggunakan kuesioner atau daftar pertanyaan dari sampel penelitian. Kelebihan dan kelemahan metode survei sebagai berikut:

1. Kelebihan Survei

Kelebihan metode survei adalah metode ini dapat mengumpulkan data dari banyak responden sekaligus. Metode survei juga bersifat fleksibel artinya pengumpulan data dapat dilakukan dengan menyesuaikan situasi dan kondisi responden. Dengan kata lain pengumpulan data dapat dilaksanakan dengan melalui *online* atau *offline*. Selain itu, teknik pengumpulan data dengan survei mudah dilakukan oleh peneliti.

2. Kekurangan Survei

Kekurangan metode survei adalah metode pengumpulan data tersebut memiliki peluang bias akibat responden tidak bersedia

memberikan informasi yang akurat. Metode pengumpulan data ini juga memiliki peluang penyimpangan makna oleh penyusunan pertanyaan atau interpretasi responden penelitian.

C. METODE OBSERVASI

Observasi adalah metode pengumpulan data yang dilakukan melalui sebuah pengamatan dan pencatatan terhadap suatu keadaan atau perilaku subyek yang diteliti. Observasi dapat juga diartikan pengumpulan data melalui pengamatan langsung terhadap subyek penelitian. Creswell (2015) menyatakan bahwa metode observasi merupakan teknik pengumpulan data yang didasarkan pada pengamatan atau meneliti dengan melihat langsung situasi dan kondisi subyek penelitian. Menurut sugiyono (2020), observasi merupakan sebuah proses yang tersusun secara sistematis dan kompleks berdasarkan proses biologi dan psikologis.

Kelebihan dan kelemahan metode observasi sebagai berikut:

1. Kelebihan Observasi

Kelebihan metode observasi adalah metode ini bisa mengumpulkan data dari responden dengan jumlah yang besar, metode observasi bisa mengumpulak data dari responden yang memiliki keragaman latar belakang, serta metode ini dapat mengumpulkan data yang representatif dari populasi.

2. Kekurangan Observasi

Kekurangan metode observasi adalah metode ini berpotensi memiliki tingkat bias atas informasi yang diperoleh, metode observasi membutuhkan biaya yang cukup besar serta waktu penelitian yang relatif lebih lama. Metode observasi juga memiliki peluang terjadinya penyimpangan dalam pemilihan responden penelitian.

Menurut Creswell (2015), metode pengumpulan data secara observasi agar berjalan dengan lancar membutuhkan tahapan-tahapan pelaksanaan antara lain menentukan tempat atau lokasi observasi, mengunjungi lokasi observasi, menentukan siapa dan apa yang akan diobservasi, menentukan peran observasi, melakukan berbagai observasi untuk memperoleh pemahaman yang utuh, merancang media yang digunakan untuk observasi, membuat catatan lapangan deskriptif dan reflektif, menunjukkan eksistensi peneliti di lingkungan observasi.

Metode observasi dapat dikelompokkan menjadi 3 macam yaitu observasi partisipatif, observasi terstruktur dan tersamar, dan observasi tidak terstruktur Sanafiah Faisal (2020). Penjelasan macam-macam metode observasi tersebut sebagai berikut:

1. Observasi Partisipatif

Metode observasi partisipatif adalah metode pengumpulan data dimana peneliti terlibat aktif dalam kegiatan sehari-hari orang atau subyek yang diteliti. Peneliti dalam pengumpulan data terlibat langsung melakukan apa yang dikerjakan oleh subyek penelitian. Metode ini memberikan peluang pengambilan data

yang presisi, lengkap dan memberikan informasi yang relevan dengan data yang dibutuhkan peneliti.

2. Observasi Terus Terang dan Tersamar

Metode observasi terus terang dan tersamar adalah metode pengambilan data dengan peneliti bertindak terus terang dan memberikan informasi kepada responden atau subyek penelitian bahwa peneliti sedang melaksanakan penelitian. Responden mengetahui seluruh aktivitas peneliti mulai dari awal hingga akhir pengambilan data. Namun pada saat tertentu, peneliti tidak terus terang atau tersamar dalam melakukan observasi. Tindakan ini dilakukan bilamana peneliti menginginkan ada data yang perlu dijaga kerahasiannya.

3. Observasi Tidak Tersruktur

Metode observasi tidak tersruktur adalah metode pengumpulan data yang dilakukan tidak secara sistematis tentang apa yang harus di observasi. Metode ini mementingkan hal-hal seperti isi dan situasi pengamatan yang dinamis, melakukan pencatatan pengamatan, mencatat data yang diperoleh ketika pengamatan sedang berjalan. Peneliti biasanya menggunakan alat bantu rekaman dan kamera untuk mengumpulkan data dari responden penelitian. Metode tidak tersruktur direncanakan agar tidak terjadi kekakuan dalam pengambilan data dan menciptakan suasana yang rileks dan hubungan baik antara responden dengan peneliti.

Metode observasi tidak terstruktur juga digunakan peneliti bilamana fokus penelitian belum jelas atau peneliti menginginkan fokus pengambilan data berkembang sesuai dengan kondisi pada saat pengamatan. Observasi tidak terstruktur tidak menerapkan aturan yang baku dalam pelaksanaan pengamatan, namun peneliti menerapkan rambu-rambu pengamatan.

D. METODE WAWANCARA

Metode wawancara merupakan metode pengumpulan data yang dilakukan dengan pengajuan sejumlah pertanyaan secara lisan kepada responden penelitian. Metode wawancara juga diartikan sebagai bentuk metode pengumpulan data dengan bertanya langsung kepada responden penelitian. Dalam perkembangan teknologi saat ini, metode wawancara langsung tidak lagi satu-satunya metode wawancara yang dilakukan. Metode wawancara dengan peneliti berkomunikasi dengan responden menggunakan telepon atau jaringan internet juga dapat dilakukan dengan dasar pertimbangan tertentu.

Peneliti harus memperhatikan beberapa hal sebelum memilih menggunakan metode wawancara yaitu subyek atau responden penelitian yang berkaitan dengan penelitian, informasi dari responden harus dapat dipercaya dan menjelaskan kondisi yang sebenarnya, pemahaman responden atas pertanyaan yang disampaikan peneliti. Dalam melakukan wawancara peneliti harus

mampu menghindari penggunaan kata yang bermakna ganda sehingga tidak terjadi kesalahpahaman oleh responden atau subyek penelitian. Peneliti juga perlu mempersiapkan pertanyaan yang singkat dan jelas. Pertanyaan yang terlalu panjang dapat mempengaruhi terjadinya bias dalam pengumpulan data. Adapun kelebihan dan kelemahan metode wawancara dijelaskan sebagai berikut:

1. Kelebihan Wawancara

Menurut Sekaran dan Bougie (2016), kelebihan metode wawancara yaitu, pertama peneliti yang menggunakan metode wawancara dapat menciptakan suasana hubungan yang lebih akrab dan nyaman dengan responden penelitian. Kondisi ini akan memberikan respon positif dan menghasilkan tanggapan yang lebih baik terutama topik yang dirasa sensitif. Kedua, peneliti yang melakukan wawancara mempunyai waktu yang lebih banyak untuk berinteraksi dan mengajukan pertanyaan, menggali informasi secara mendalam dan mengatur pertanyaan yang akan menjadi informasi kunci. Ketiga, peneliti bisa merubah pemilihan kata dalam pertanyaan dari yang direncanakan semula bilamana ketika saat mewawancara terdapat kata yang bisa memberikan pemahaman ganda atau sulit dipahami oleh responden. Peneliti perlu menyesuaikan sesesuai konteks dan situasi dilapangan. Peneliti juga dalam mewawancara tatap muka akan bisa memperhatikan bahasa tubuh dari responden penelitian. Keempat, mutu sampel yang diambil dalam wawancara biasanya lebih tinggi dibandingkan

dengan metode pengumpulan data lainnya. Kelima, peneliti dalam melakukan wawancara tidak perlu membutuhkan banyak responden penelitian untuk memperoleh informasi yang relevan dengan penelitian. Keenam, metode wawancara minim gangguan yang berpotensi untuk mengganggu pengambilan data yang terkadang muncul dalam FGD. Ketujuh, metode pengumpulan data berdasarkan wawancara memiliki peluang untuk memperoleh informasi atau data yang sangat mendalam. Kedelapan, peneliti bisa menyusul jadwal dan tempat wawancara lebih awal sebelum dilakukan wawancara dengan responden sesuai prosedur atau rancangan penelitian.

2. Kelemahan wawancara

Menurut Sekaran dan Bougie (2016), kelemahan metode wawancara yaitu pertama, metode pengumpulan data secara wawancara relatif membutuhkan waktu yang lama. Peneliti dalam melakukan wawancara kepada responden perlu melakukan wawancara mendalam guna memperoleh data yang sebenarnya. Data yang sudah dikumpul perlu diurutkan, diatur kemudian dianalisis dan dilaporkan. Kedua, peneliti dalam melakukan wawancara membutuhkan tenaga dan biaya yang relatif lebih besar dibandingkan dengan metode pengumpulan data lainnya. Ketiga, peluang terjadinya bias dalam wawancara sangat mungkin terjadi akibat kesalahan menafsirkan pertanyaan atau memilih kata dalam mengajukan pertanyaan. Keempat, peneliti yang tidak mempunyai keterampilan dan pengalaman sebagai pewawancara akan berpeluang proses wawancara yang

dilakukan tidak jauh berbeda dengan metode survei atau kuesioner. Kelima, pemilihan responden penelitian perlu kajian dan kehati-hatian untuk mengantisipasi terjadinya bias. Proses ini tentu mengharuskan peneliti melakukan indentifikasi dan penilaian yang tentunya membutuhkan waktu yang relatif lama.

Metode pengumpulan data berdasarkan wawancara biasanya dilakukan dengan maksud untuk mengidentifikasi, menyelidiki dan mengeksplorasi berbagai faktor dalam situasi yang menjadi sentral dari masalah yang kompleks. Wawancara juga ditujukan untuk melakukan pengujian hipotesis atau memberikan hipotesis baru terhadap variabel atau peubah penelitian dan hubungannya. Dengan demikian, peneliti dapat menerapkan jenis wawancara sesuai konteks dan tujuan penelitian yang diharapkan. Adapun jenis wawancara yang umum dalam penelitian dijelaskan sebagai berikut:

1. Wawancara Terstruktur

Peneliti memilih wawancara terstruktur dikarenakan informasi yang akan dibutuhkan sudah diketahui sejak awal. Bagi peneliti yang baru menggunakan metode wawancara alangkah lebih baiknya menggunakan wawancara terstruktur yang dapat disiapkan terlebih dahulu. Pertanyaan-pertanyaan wawancara terstruktur umumnya bersifat tertutup artinya peneliti melakukan wawancara dengan memberikan pilihan jawaban.

2. Wawancara Semi Terstruktur

Wawancara semi terstruktur merupakan metode pengumpulan data dengan menggunakan pedoman wawancara sebagai acuan.

Peneliti tidak terikat dengan daftar pertanyaan yang kaku. Peneliti bisa menyesuaikan dengan situasi dan kondisi responden dalam mengajukan, menambah dan melakukan formulasi pertanyaan. Pada umumnya, wawancara semi terstruktur bersifat pertanyaan terbuka. Peneliti memberikan kebebasan kepada untuk menjawab sesuai pemahaman dan pengalaman yang dimiliki responden penelitian.

3. Wawancara Tidak Tersruktur

Wawancara tidak terstruktur adalah metode pengumpulan data wawancara bersifat bebas artinya pengumpulan data tidak menggunakan pedoman daftar pertanyaan dan pilihan jawaban. Peneliti biasanya menentukan topik atau tujuan umum dari wawancara. Pengumpulan data bersifat wawancara tidak terstruktur memberikan kebebasan kepada responden untuk menyampaikan informasi sesuai pertanyaan yang diajukan. Peneliti lebih berperan aktif sebagai pendengar yang memberikan tanggapan non verbal. Pertanyaan dalam wawancara tidak terstruktur umumnya bersifat eksploratif yaitu untuk mengetahui pandangan, sikap dan perasaan responden penelitian terhadap suatu hal.

E. METODE DOKUMENTASI

Metode dokumentasi adalah metode pengumpulan data yang dilakukan secara tidak langsung ke subyek atau responden

penelitian. Pengumpulan data dilakukan dengan studi dokumen. Dokumen merupakan catatan tertulis yang memuat pernyataan tertulis yang disusun oleh seseorang atau lembaga atas sebuah peristiwa yang bertujuan sebagai sumber data atau informasi. Metode dokumentasi memiliki kelebihan dan kelemahan sebagai berikut:

1. Kelebihan Dokumentasi

Kelebihan metode dokumentasi yaitu pertama, peneliti yang sulit menjangkau responden atau subyek penelitian melalui pengumpulan data secara dokumentasi dapat memberikan alternatif jalan pengumpulan data. Kedua, penelitian yang dilakukan dengan menggunakan data masa lampau menjadikan metode dokumentasi sebagai pilihan terbaik. Ketiga pengumpulan data secara dokumentasi bersifat tidak reaktif karena studi dokumentasi dilakukan kepada benda mati maka data yang dibutuhkan tidak dipengaruhi kehadiran peneliti. Keempat, penelitian dengan metode dokumentasi memberikan peluang untuk pengambilan sampel yang lebih besar dengan biaya yang cenderung kecil.

2. Kelemahan Dokumentasi

Kelemahan metode pengumpulan data secara dokumentasi yaitu pertama, pengumpulan data secara dokumentasi memiliki peluang bias karena dokumen yang tersedia bisa ditulis secara berlebihan atau ada data yang disembunyikan. Hal ini akan mempengaruhi keakuratan hasil penelitian. Kedua, dokumen bisa saja tidak bisa diakses semua orang. Peneliti bisa saja

kesulitan dalam mengumpulkan dokumen-dokumen yang relevan dengan penelitian. Ketiga, data yang termuat dalam dokumen seringkali kurang lengkap. Keempat, dokumen memiliki format yang tidak baku sehingga peneliti akan menghadapi kendala disebabkan tujuan penulisan dokumen berbeda dengan tujuan penelitian.

Peneliti dalam menggunakan metode dokumentasi perlu mengatasi kelemahan-kelemahan tersebut. Peneliti perlu melakukan observasi terhadap data yang telah diperoleh. Penting sekali, peneliti juga memperhatikan waktu pembuatan dokumen dan koherensinya dengan tujuan penelitian.

BAGIAN 4

SUMBER DATA DALAM STATISTIKA

A. DATA PRIMER (SENSUS, SURVEI, *BIG DATA*)

1. Sensus

Data primer khususnya yang diperoleh melalui sensus merupakan pondasi utama dalam statistika dalam menyajikan informasi yang akurat serta bisa diandalkan. Sensus adalah upaya pemerintah atau badan statistika untuk mengumpulkan data dari seluruh populasi atau elemen yang ingin diteliti. Dalam konteks statistika, data primer dari sensus memiliki peran yang sangat penting karena memberikan gambaran menyeluruh tentang karakteristik populasi yang diteliti.

Sensus adalah salah satu sumber data utama dalam statistika yang memberikan informasi yang tidak bisa didapatkan dari sumber data lain (Wang, Wu and Liu, 2019). Dalam sebuah sensus penduduk, misalnya, data primer yang dikumpulkan mencakup informasi tentang demografi, pendidikan, pekerjaan, dan berbagai variabel lainnya yang relevan untuk analisis statistika. Hal ini membuat sensus menjadi sumber data yang sangat berharga bagi para peneliti dan pengambil kebijakan.

Salah satu keunggulan utama data primer dari sensus adalah tingkat keakuratannya. Karena sensus mencakup seluruh populasi, data

yang diperoleh lebih akurat dibandingkan dengan metode sampling atau survei (Li, Xu and Zhang, 2020). Dengan demikian keputusan yang diambil berdasarkan data primer ini memiliki dasar yang kuat dan dapat diandalkan.

Meskipun data primer dari sensus memberikan gambaran yang komprehensif, proses pengumpulan data ini mempunyai tantangan khusus pula. Salah satu masalah pokok yaitu memastikan partisipasi yang baik dari responden. Dalam sensus penduduk, misalnya, pemerintah harus melakukan kampanye yang efektif untuk meningkatkan tingkat partisipasi masyarakat.

Dalam rangka memastikan kualitas dan keandalan data primer dari sensus, perlu upaya yang berkelanjutan dalam pengembangan metodologi pengumpulan data serta pengelolaan serta analisis data yang sesuai.

2. Survei

Survei adalah metode pengumpulan data yang melibatkan pengambilan sampel dari populasi tertentu untuk mendapatkan pemahaman tentang karakteristik, pendapat, atau perilaku populasi. Survei merupakan instrumen yang efektif untuk mengumpulkan data primer dari populasi yang luas dan beragam (Smith, Johnson and Williams, 2021). Dalam survei, responden diminta untuk menjawab serangkaian pertanyaan yang dirancang untuk mencakup berbagai aspek yang ingin diteliti. Dengan demikian data primer yang diperoleh dari survei dengan *sample*

diharapkan juga dapat menyajikan gambaran yang akurat mengenai populasi yang diamati.

Salah satu keunggulan utama data primer dari survei adalah kemampuannya untuk menghasilkan data yang spesifik dan terperinci tentang subjek yang diteliti. Survei memungkinkan peneliti untuk mengumpulkan informasi seperti pendapat, preferensi, dan sikap responden (Jones, Brown and Davis, 2019) secara lebih efisien daripada sensus.

Proses pengumpulan data primer melalui survei mempunyai tantangan khusus. Salah satu masalah terpenting yaitu memastikan representativitas *sample* yang diambil (Brown, Wilson and Garcia, 2020). *Sample* yang tidak representatif dapat menghasilkan data yang bias dan tidak akurat sehingga mempengaruhi validitas analisis statistika yang dilakukan.

Dalam mengatasi tantangan tersebut perlu adanya perencanaan yang matang dalam desain survei, termasuk dalam pemilihan *sample* dan formulasi pertanyaan. Hal ini kritikal dalam memastikan bahwa data primer yang didapatkan berdasarkan *sample* memiliki kualitas yang tinggi dan dapat diandalkan untuk analisis statistika yang akurat.

Melalui survei peneliti dan pengambil kebijakan dapat mengumpulkan informasi yang spesifik dan terperinci tentang populasi yang diteliti yang kemudian bisa dimanfaatkan dalam melakukan pengambilan keputusan dengan lebih efektif.

3. *Big Data*

Big data merupakan sekumpulan data bervolume luar biasa besar serta kompleks yang diperoleh berdasarkan kemajuan teknologi informasi dan komunikasi. Dalam statistika data primer dari *big data* memiliki peran yang semakin penting dalam memberikan wawasan yang dalam mengenai kecenderungan, pola, serta relasi di pada data. *Big data* memiliki potensi besar dalam menyediakan data primer yang beragam dan luas dari berbagai sumber termasuk sensor, media sosial, serta transaksi bisnis (Chen, Wang and Liu, 2023). Dengan volume data yang besar, *big data* memungkinkan analisis statistika yang lebih mendalam dan akurat, yang bisa dimanfaatkan dalam memprediksi secara lebih akurat tentang perilaku atau kejadian di masa depan.

Salah satu keunggulan utama data primer dari *big data* yaitu kemampuan dalam menyajikan pemahaman secara *real-time*. Data primer yang diperoleh dari *big data* dapat memberikan informasi tentang kejadian atau tren yang sedang terjadi saat ini sehingga memungkinkan pengambil keputusan untuk bertindak secara cepat dan responsif (Li, Zhang and Wu, 2022).

Penggunaan data primer dari *big data* juga menimbulkan tantangan spesifik. Salah satu tantangan utama yaitu dalam pengelolaan serta analisis data yang kompleks (Wang, Xu and Brown, 2021). *Big data* sering kali terdiri dari berbagai jenis data yang heterogen dan tidak

terstruktur sehingga memerlukan teknik analisis yang canggih untuk mengekstrak informasi yang berarti.

Dalam mengatasi tantangan tersebut pengembangan teknik analisis data dan algoritma baru menjadi sangat penting. Penggunaan teknologi seperti *machine learning* dan *data mining* telah menjadi kunci dalam mengolah data primer dari *big data* secara efisien dan efektif. Dengan teknik-teknik ini peneliti dan pengambil keputusan dapat mengidentifikasi pola-pola laten dalam *big data* serta mengambil tindakan yang sesuai (Zhang *et al.*, 2020).

B. DATA SEKUNDER

1. Pengertian Data Sekunder

Data sekunder merujuk pada data yang telah dikumpulkan dan dipublikasikan sebelumnya oleh pihak lain untuk tujuan tertentu. Dalam konteks statistika data sekunder sering kali digunakan untuk melengkapi atau memperluas analisis yang dilakukan, atau untuk menyediakan dasar perbandingan dengan data primer.

Data sekunder dapat mencakup berbagai jenis informasi termasuk data survei, data administratif, data akademik, dan banyak lagi (Smith, Johnson and Williams, 2021). Data ini telah dikumpulkan oleh lembaga atau individu lain untuk tujuan tertentu seperti penelitian, analisis pasar, atau kepentingan administratif.

Salah satu keunggulan utama data sekunder adalah ketersediaannya yang relatif mudah dan murah dibandingkan dengan pengumpulan data primer. Data sekunder dapat diakses melalui berbagai sumber termasuk jurnal ilmiah, database online, dan laporan riset yang dipublikasikan (Jones, Brown and Davis, 2019). Hal tersebut memfasilitasi peneliti dalam mengakses informasi tanpa perlu melakukan pengumpulan data sendiri.

Meskipun demikian penting untuk diingat bahwa data sekunder juga memiliki beberapa keterbatasan. Salah satu keterbatasannya adalah terbatasnya kontrol peneliti terhadap kualitas dan validitas data. Kualitas data sekunder dapat bervariasi tergantung pada sumbernya, dan peneliti perlu melakukan evaluasi yang cermat terhadap keandalan data sebelum menggunakan data tersebut dalam analisis.

Dalam penggunaan data sekunder peneliti juga harus memperhatikan etika dalam penggunaan data yang sudah ada. Adalah penting untuk menghormati hak cipta dan privasi serta memastikan bahwa penggunaan data sekunder tidak melanggar kebijakan atau regulasi.

2. Metode Pengumpulan Data Sekunder

Metode pengumpulan data sekunder adalah proses mendapatkan informasi dari sumber-sumber yang sudah ada sebelumnya untuk digunakan dalam analisis statistika. Dalam statistika penggunaan data sekunder sering kali menjadi pilihan yang efisien dan efektif

karena menghindari biaya dan waktu yang diperlukan untuk mengumpulkan data primer.

Ada metodologi yang lazim dalam memperoleh data sekunder (Wang, Xu and Liu, 2021). Salah satunya adalah dengan melakukan penelusuran literatur, di mana peneliti mencari informasi dari jurnal ilmiah, buku, laporan riset, serta dokumen-dokumen lain yang terkait sesuai tema. Metode ini memungkinkan peneliti untuk mengakses informasi yang terkini dalam bidang tertentu.

Penggunaan *database* juga menjadi metode yang populer dalam mengumpulkan data sekunder. Banyak lembaga dan organisasi yang menyediakan akses ke basis data yang mencakup berbagai jenis informasi mulai dari data ekonomi dan sosial hingga data kesehatan dan lingkungan (Jones, Brown and Davis, 2019). Dengan mengakses *database* ini peneliti bisa mendapatkan data yang luas serta terstruktur bagi kepentingan analisis statistika.

Teknologi digital juga telah memungkinkan penggunaan data sekunder dari platform online. Internet dan media sosial menjadi sumber informasi yang penting dalam mengumpulkan data sekunder, terutama dalam konteks penelitian tentang perilaku konsumen, tren sosial, dan opini publik. Dengan memanfaatkan data dari platform online, peneliti dapat mengakses informasi yang *real-time* dan relevan.

Meskipun metode pengumpulan data sekunder dapat memberikan akses ke informasi yang beragam dan luas, peneliti perlu

memperhatikan keandalan dan validitas data yang diperoleh. Evaluasi yang cermat terhadap sumber data sekunder menjadi penting untuk menjamin bahwa data yang diperoleh memiliki kualitas yang baik.

3. Keuntungan dan Tantangan Data Sekunder

Penggunaan data sekunder dalam statistika memiliki sejumlah keuntungan yang signifikan, namun juga melibatkan beberapa tantangan yang perlu diatasi. Memahami kedua sisi ini penting bagi peneliti dan analis statistika.

Salah satu keuntungan utama penggunaan data sekunder adalah efisiensi waktu dan biaya. Menggunakan data yang sudah ada sebelumnya memungkinkan peneliti untuk menghemat waktu yang seharusnya digunakan untuk mengumpulkan data baru. Hal ini juga mengurangi biaya yang terkait dengan proses pengumpulan data, seperti biaya survei atau pengumpulan dari lapangan.

Penggunaan data sekunder memungkinkan peneliti untuk mengakses informasi yang luas dan beragam. Data sekunder dapat mencakup berbagai aspek yang relevan untuk penelitian mulai dari data ekonomi dan sosial hingga data kesehatan dan lingkungan (ones, Brown and Davis, 2019). Dengan demikian peneliti dapat memiliki akses ke data yang lebih lengkap dan komprehensif.

Penggunaan data sekunder juga melibatkan sejumlah tantangan. Salah satu tantangan utama adalah kualitas dan validitas data.

Kualitas data sekunder dapat bervariasi tergantung pada sumber dan peneliti perlu melakukan evaluasi yang cermat terhadap keandalan data sebelum menggunakan dalam analisis statistika (Brown, Wilson and Garcia, 2020). Ada juga risiko bahwa data sekunder mungkin tidak sepenuhnya sesuai dengan kebutuhan penelitian.

Tantangan lain adalah dalam memastikan representativitas data. Dalam beberapa kasus data sekunder mungkin tidak mencakup seluruh populasi yang ingin diteliti sehingga menghasilkan kesimpulan yang tidak sepenuhnya representatif. Oleh karena itu peneliti perlu mempertimbangkan keberagaman dan distribusi *sample* data sekunder dalam analisis. Dengan memahami keuntungan dan tantangan penggunaan data sekunder peneliti dapat mengambil langkah-langkah yang tepat untuk menjamin bahwa data sekunder yang dimanfaatkan pada analisis statistika memiliki kualitas yang baik dan dapat diandalkan.

C. JENIS DATA (DATA KUANTITATIF DAN KUALITATIF)

1. Data Kuantitatif

Data kuantitatif merupakan salah satu jenis data yang diperoleh dalam bentuk angka atau nilai numerik yang memungkinkan dilakukan analisis statistika untuk mengeksplorasi pola, hubungan, dan tren dalam data. Dalam statistika data kuantitatif mempunyai peranan yang sangat kritical menjadi sumber informasi yang

digunakan untuk membuat generalisasi dan prediksi tentang fenomena yang diamati.

Data kuantitatif dapat diperoleh dari berbagai sumber termasuk survei, eksperimen, rekaman administratif, dan banyak lagi (Smith, Johnson and Williams, 2021). Data ini memberikan dasar empiris bagi analisis statistika yang memungkinkan peneliti untuk menguji hipotesis, membuat inferensia, dan mengidentifikasi pola serta melakukan prediksi atas fenomena yang diamati.

Salah satu keunggulan utama data kuantitatif adalah kemampuannya untuk diolah dan dianalisis secara sistematis. Data kuantitatif dapat diatur dalam bentuk tabel, grafik, atau diagram yang memungkinkan analisis statistika yang lebih terstruktur dan komprehensif (Jones, Brown and Davis, 2019). Dengan mendayagunakan teknik statistika yang sesuai peneliti mampu melakukan identifikasi relasi kausalitas ataupun asosiatif antar variabel-variabel yang diamati.

Data kuantitatif juga memungkinkan analisis statistika yang lebih lanjut melalui teknik-teknik seperti regresi, uji hipotesis, dan analisis varians. Teknik-teknik ini memungkinkan peneliti untuk mengukur kekuatan hubungan antara variabel, menguji signifikansi perbedaan antara kelompok, dan membuat prediksi tentang perilaku atau kejadian di masa depan (Brown, Wilson and Garcia, 2020).

Penggunaan data kuantitatif juga melibatkan sejumlah tantangan yang harus diatasi. Salah satu masalah utama yaitu dalam menjamin

akurasi serta keandalan data. Kesalahan pengukuran atau pelaporan dapat menghasilkan data yang bias atau tidak dapat diandalkan sehingga mempengaruhi validitas analisis statistika yang dilakukan.

Interpretasi data kuantitatif juga memerlukan pemahaman yang cermat terhadap konteks dan metodologi yang digunakan dalam pengumpulan data. Tanpa pemahaman yang memadai tentang karakteristik dan keterbatasan data, analisis statistika dapat menghasilkan kesimpulan yang salah atau tidak tepat.

Dengan memahami peran, keunggulan, dan tantangan data kuantitatif, peneliti bisa mendayagunakan data ini dengan efektif dalam mendapatkan wawasan yang lebih dalam tentang fenomena yang diteliti. Dengan kombinasi penggunaan teknik analisis yang tepat dan pemahaman yang mendalam tentang data yang digunakan, data kuantitatif dapat menjadi alat yang kuat dalam statistika untuk membuat inferensia yang berarti dan mendukung pengambilan keputusan yang informasional.

2. Data Kualitatif

Data kualitatif merupakan jenis data yang tidak berupa numerik dan biasanya didapatkan dalam suatu observasi, wawancara, atau analisis teks yang menghasilkan deskripsi atau interpretasi tentang fenomena yang diamati. Dalam statistika data kualitatif memiliki peran yang penting dalam memahami konteks, makna, dan kompleksitas dari suatu fenomena, serta dapat digunakan untuk mendukung analisis statistika yang lebih mendalam.

Data kualitatif memungkinkan peneliti untuk menjelajahi dan menggali kedalaman tentang persepsi, sikap, dan pengalaman individu atau kelompok terhadap suatu topik. Dengan menggunakan teknik seperti wawancara mendalam, observasi partisipan, atau analisis isi, data kualitatif memberikan wawasan yang kaya dan mendalam tentang fenomena yang diamati.

Salah satu keunggulan utama data kualitatif yaitu kemampuan dalam merepresentasikan nuansa serta konteks dari sebuah realita. Data kualitatif memungkinkan peneliti untuk memahami bagaimana individu atau kelompok berinteraksi dengan lingkungan, bagaimana responden memberikan makna terhadap pengalaman, dan bagaimana responden merespons perubahan dalam konteks tertentu (Jones, Brown and Davis, 2019).

Data kualitatif memungkinkan peneliti untuk mengeksplorasi dan mengidentifikasi pola atau tema yang muncul dalam data. Melalui analisis yang mendalam, peneliti dapat mengidentifikasi pola-pola yang mendasari perilaku atau kepercayaan, serta memahami faktor-faktor yang mempengaruhi variasi dalam pengalaman individu atau kelompok (Brown, Wilson and Garcia, 2020).

Penggunaan data kualitatif juga melibatkan sejumlah tantangan yang harus diatasi. Salah satu masalah yaitu dalam menjamin validitas serta reliabilitas data. Karena data kualitatif bisa jadi bersifat subjektif dan interpretatif, peneliti perlu memastikan bahwa

data yang diperoleh relevan dan dapat diandalkan untuk membuat generalisasi yang berarti (Wang, Xu and Brown, 2021).

Tantangan lainnya adalah dalam mengelola dan menganalisis data kualitatif yang kompleks. Data kualitatif sering kali terdiri dari berbagai jenis informasi yang heterogen, dan peneliti perlu menggunakan teknik analisis yang tepat untuk mengorganisir dan menginterpretasikan dengan benar. Dalam hal ini lazimnya data kualitatif dikonversikan menjadi data kuantitatif agar dapat dilakukan analisis statistika.

D. PENYAJIAN DATA (DATA *CROSS SECTION*, *TIME SERIES*, DAN PANEL)

1. Data *Cross Section*

Penyajian data *cross section* merupakan salah satu pendekatan dalam statistika yang melibatkan pengumpulan data pada suatu titik waktu tertentu dari berbagai individu, unit, atau variabel yang beragam. Dalam konteks statistika data *cross section* memberikan gambaran tentang karakteristik, perilaku, atau hubungan antara variabel pada suatu titik waktu tertentu.

Data *cross section* sering digunakan untuk menggambarkan variasi dalam populasi pada suatu saat tertentu (Smith, Johnson and Williams, 2021). Misalnya dalam survei penduduk yang dilakukan pada satu titik waktu tertentu data *cross section* dapat memberikan

informasi tentang distribusi demografi, kondisi ekonomi, atau preferensi sosial pada saat survei dilakukan.

Salah satu keunggulan utama data *cross section* adalah kemampuan untuk memberikan gambaran yang komprehensif tentang kondisi pada satu titik waktu tertentu. Data *cross section* dapat memberikan pemahaman yang lebih dalam tentang variasi dan keberagaman dalam populasi pada saat tertentu yang kemudian dapat digunakan untuk membuat inferensi tentang fenomena yang diamati.

Data *cross section* juga memungkinkan analisis statistika yang lebih luas dan mendalam. Data *cross section* dapat digunakan untuk menganalisis hubungan antara variabel, atau menguji hipotesis tentang faktor-faktor yang mempengaruhi fenomena yang diamati.

Penggunaan data *cross section* juga melibatkan sejumlah tantangan yang harus ditangani. Salah satu masalah pokok yaitu dalam memastikan representativitas *sample* yang diambil. *Sample* yang tidak representatif dapat menghasilkan simpulan yang tidak akurat ataupun tidak terandalkan sehingga mempengaruhi validitas analisis statistika yang dilakukan (Wang, Xu and Brown, 2021).

Tantangan lainnya adalah dalam mengatasi masalah keberagaman dan heterogenitas dalam data *cross section*. Karena data *cross section* sering kali mencakup berbagai jenis individu atau unit yang berbeda, peneliti perlu menggunakan teknik analisis yang tepat untuk mengelompokkan atau membandingkan data dengan benar.

2. *Time Series*

Penyajian data *time series* adalah pendekatan dalam statistika yang melibatkan pengumpulan data secara berkala selama periode waktu tertentu yang kemudian digunakan untuk menganalisis perubahan atau tren dalam variabel yang diamati sepanjang waktu. Dalam statistika data *time series* memberikan wawasan tentang bagaimana suatu variabel berubah seiring waktu dan memungkinkan analisis tentang pola atau perilaku jangka panjang.

Data *time series* sering digunakan untuk memonitor dan menganalisis perubahan dalam ekonomi, keuangan, cuaca, atau kesehatan selama periode waktu tertentu. Misalnya data *time series* tentang harga saham dapat digunakan untuk mengidentifikasi tren pasar atau pola fluktuasi harga sepanjang waktu.

Salah satu keunggulan utama data *time series* adalah kemampuan untuk menunjukkan pola atau tren jangka panjang dalam data. Dengan menganalisis data *time series*, dapat diidentifikasi pola musiman, tren jangka panjang, atau fluktuasi acak dalam variabel yang diamati yang kemudian dapat digunakan untuk membuat prediksi tentang perilaku masa depan (Jones, Brown and Davis, 2019).

Data *time series* juga memungkinkan analisis statistika yang lebih mendalam dan kompleks. Teknik seperti analisis spektral, dekomposisi, atau model regresi *time series* dapat digunakan untuk mengidentifikasi dan memodelkan hubungan antara variabel dalam

data *time series* sehingga menghasilkan wawasan yang semakin dalam mengenai aspek-aspek yang berpengaruh dan menyebabkan perubahan seiring waktu.

Pendayagunaan data *time series* juga melibatkan sejumlah tantangan yang perlu diatasi. Salah satu masalah utama yaitu dalam memastikan konsistensi dan kualitas data selama periode waktu yang diamati. Kesalahan pengukuran atau pemrosesan data dapat menghasilkan ketidakakuratan dalam analisis *time series* sehingga mempengaruhi validitas kesimpulan yang diambil.

Tantangan lainnya adalah dalam menghadapi fluktuasi acak atau gangguan yang mungkin terjadi dalam data *time series*. Dalam menganalisis data *time series* peneliti perlu menggunakan teknik-teknik statistika yang tepat untuk mengidentifikasi dan mengurangi efek gangguan atau *noise* dalam data sehingga memperoleh estimasi yang lebih akurat (Li, Zhang and Wu, 2022).

Dengan memahami keunggulan dan tantangan data *time series* peneliti bisa mendayagunakan data *time series* dengan efektif untuk memahami perubahan seiring waktu dalam fenomena yang diteliti. Dengan kombinasi penggunaan teknik analisis yang tepat dan pemahaman yang mendalam tentang data yang digunakan, data *time series* dapat menjadi alat yang kuat dalam statistika untuk mengidentifikasi tren, memprediksi perilaku masa depan, dan mendukung pengambilan keputusan yang informasional.

3. Panel

Penyajian data panel adalah metode dalam statistika yang melibatkan pengumpulan data dari unit-unit yang sama pada beberapa titik waktu yang berbeda. Dalam pengamatan data panel, unit-unit tersebut dapat berupa individu, rumah tangga, perusahaan, atau negara, yang diamati pada beberapa waktu yang berbeda untuk menganalisis perubahan atau tren dalam variabel yang diamati.

Data panel sering digunakan untuk menganalisis perubahan seiring waktu dalam unit-unit yang diamati. Misalnya dalam studi ekonomi, data panel dapat digunakan untuk menganalisis dampak kebijakan atau faktor-faktor ekonomi terhadap perilaku perusahaan atau pasar pada beberapa periode waktu.

Salah satu keunggulan utama data panel adalah kemampuan untuk menangkap variasi antar unit dan perubahan seiring waktu dalam unit yang sama. Dengan menggunakan data panel, peneliti dapat mengidentifikasi dan memodelkan perubahan dalam hubungan antara variabel serta menguji hipotesis tentang faktor-faktor yang mempengaruhi perubahan tersebut (Jones, Brown and Davis, 2019).

Data panel juga memungkinkan analisis statistika yang lebih mendalam dan kompleks. Dengan menggunakan teknik analisis panel seperti model regresi panel atau analisis *variance* (ANOVA) panel, peneliti dapat mengidentifikasi dan memahami faktor-faktor

yang mempengaruhi perubahan dalam variabel yang diamati sepanjang waktu.

Penggunaan data panel juga melibatkan sejumlah masalah yang harus ditangani. Salah satu masalah pokok yaitu dalam mengatasi masalah keberagaman antar unit dan waktu dalam data panel. Peneliti perlu menggunakan teknik analisis yang tepat untuk mengendalikan faktor-faktor yang tidak diamati atau efek yang tersembunyi dalam data panel sehingga memperoleh estimasi yang lebih akurat (Wang, Xu and Brown, 2021).

Tantangan lainnya adalah dalam mengelola dan menganalisis data panel yang kompleks. Data panel sering kali terdiri dari berbagai variabel dan observasi yang heterogen dan peneliti perlu menggunakan teknik-teknik statistika yang tepat untuk mengorganiskan dan menganalisis data dengan benar.

E. SKALA DATA (NOMINAL, ORDINAL, INTERVAL, RASIO)

1. Nominal

Skala data nominal yaitu salah satu jenis skala dalam statistika yang didayagunakan untuk mengukur ataupun melakukan klasifikasi data suatu variabel ke dalam kategori-kategori tertentu yang tidak mempunyai makna urutan ataupun tingkatan. Dalam skala ini data hanya bisa dikategorikan menjadi kelompok atau klasifikasi

tertentu tanpa memperhitungkan tingkat atau derajat relatif di antara kategori-kategori tersebut.

Data nominal sering digunakan untuk mengklasifikasikan data pada suatu variabel ke dalam kelompok-kelompok yang tidak memiliki urutan atau hierarki tertentu (Smith, Johnson and Williams, 2021). Contohnya dalam survei demografi, jenis kelamin (pria, wanita) atau status perkawinan (belum menikah, menikah, bercerai) merupakan contoh variabel yang dapat diukur menggunakan skala nominal.

Salah satu ciri terpenting dalam skala data nominal yaitu bahwa kategori-kategori yang diaplikasikan adalah saling eksklusif serta tidak saling beririsan. Setiap unit data hanya bisa dimasukkan ke dalam suatu kategori saja, serta tidak ada peringkat atau tingkat yang dihubungkan dengan kategori tersebut (Jones, Brown and Davis, 2019).

Penggunaan skala data nominal juga memiliki sejumlah keterbatasan. Skala nominal tidak memperhitungkan perbedaan intensitas atau derajat antara kategori-kategori yang digunakan sehingga informasi tentang perbedaan kuantitatif antara unit data hilang dalam analisis (Brown, Wilson and Garcia, 2020).

Meskipun demikian skala data nominal tetap menjadi alat yang penting dalam statistika untuk mengklasifikasikan dan mengorganisasikan data ke dalam kelompok-kelompok yang relevan. Dengan pemahaman yang baik tentang sifat serta ciri skala data

nominal, peneliti bisa mendayagunakan data nominal dengan efektif.

2. Ordinal

Skala data ordinal yaitu jenis skala dalam statistika yang digunakan untuk mengukur ataupun mengklasifikasikan data suatu variabel ke dalam kategori-kategori yang mempunyai urutan ataupun tingkatan relatif tertentu. Data tidak hanya dikategorikan menjadi beberapa kelompok, tetapi juga memiliki tingkatan atau derajat yang mengindikasikan perbedaan relatif antara kategori-kategori tersebut.

Data ordinal sering digunakan untuk mengklasifikasikan variabel ke dalam kelompok-kelompok yang memiliki urutan atau hierarki tertentu, tetapi dalam hal ini belum ada interval yang konsisten di antara kategori-kategori tersebut (Smith, Johnson and Williams, 2021). Contoh variabel yang diukur menggunakan skala ordinal yaitu tingkatan pendidikan (SD, SMP, SMA, perguruan tinggi) ataupun tingkatan kepuasan (sangat puas, puas, netral, tidak puas, sangat tidak puas).

Salah satu ciri pokok dalam skala data ordinal yaitu bahwa kategori-kategori yang dipakai sudah memiliki urutan atau tingkatan yang dikenali, tetapi interval antara kategori-kategori tersebut tidak selalu konsisten. Meskipun ada urutan yang jelas, perbedaan antara tingkatan tidak selalu sama sehingga sulit untuk melakukan operasi matematika yang valid pada data ordinal.

Skala ordinal memungkinkan peneliti untuk mengidentifikasi pola atau tren dalam data yang memiliki urutan atau hierarki tertentu serta untuk melakukan perbandingan relatif antara kategori-kategori yang diamati (Brown, Wilson and Garcia, 2020).

Meskipun memiliki keterbatasan dalam penggunaan operasi matematika, skala data ordinal tetap menjadi alat yang penting dalam statistika untuk mengorganisasikan dan mengklasifikasikan data ke dalam kelompok-kelompok yang relevan.

3. Interval

Skala data interval merupakan jenis skala dalam statistika yang memungkinkan pengukuran variabel dengan memperhitungkan perbedaan sebenarnya antara nilai-nilai yang diamati. Dalam skala ini perbedaan antara setiap titik pada skala dianggap sama, tetapi tidak ada titik nol absolut yang menjadi referensi pada variabel tersebut. Data interval sering digunakan dalam pengukuran suhu, waktu, atau skor tes (Smith, Johnson and Williams, 2021). Dalam pengukuran suhu Celsius, misalnya, perbedaan antara 20°C dan 30°C dianggap sama dengan perbedaan antara 30°C dan 40°C tetapi tidak ada titik nol absolut yang menjadi referensi ketiadaan suhu.

Salah satu ciri pokok dalam skala data interval yaitu bahwa perbedaan antara setiap titik pada skala dianggap sama dan operasi matematika seperti penambahan atau pengurangan dapat dilakukan secara valid. Dengan menggunakan data interval, peneliti dapat

melakukan perhitungan rata-rata, standar deviasi, atau mengukur perubahan relatif antara nilai-nilai yang diamati.

Adalah penting untuk dicatat bahwa skala data interval belum mempunyai titik nol absolut sebagai referensi yang menyatakan ketiadaan variabel tersebut. Keberadaan nol pada skala interval tidak memiliki makna absolut, tetapi hanya merupakan salah satu titik pada skala yang dipilih untuk tujuan referensi.

Meskipun memiliki keterbatasan ini, skala data interval tetap memberikan manfaat dalam analisis statistika. Data interval memungkinkan peneliti untuk melakukan perbandingan relatif yang lebih akurat antara nilai-nilai yang diamati, serta untuk mengidentifikasi pola atau tren dalam data dengan lebih tepat (Wang, Xu and Brown, 2021).

4. Rasio

Skala data rasio merupakan jenis skala dalam statistika yang memiliki semua karakteristik skala interval namun sudah memiliki titik nol yang menyatakan ketiadaan variabel tersebut secara absolut. Dalam skala ini perbedaan antara setiap titik pada skala dianggap sama, dan titik nol memiliki makna yang signifikan karena menunjukkan ketiadaan variabel yang diukur.

Data rasio sering digunakan dalam pengukuran yang melibatkan ukuran, bobot, panjang, atau jumlah. Contoh penggunaannya adalah dalam pengukuran berat badan, di mana nilai nol

menunjukkan ketiadaan berat badan, dan perbedaan antara setiap titik pada skala dianggap sama.

Salah satu ciri pokok dalam skala data rasio yaitu bahwa titik nol pada skala memiliki makna yang absolut. Keberadaan titik nol pada skala rasio memungkinkan peneliti untuk melakukan operasi matematika yang lebih lanjut, seperti perkalian atau pembagian, dan untuk membuat perbandingan absolut antara nilai-nilai yang diamati (Jones, Brown and Davis, 2019).

Seperti skala interval skala data rasio juga memerlukan kehati-hatian dalam penggunaannya. Meskipun titik nol pada skala rasio memiliki makna absolut, peneliti perlu memperhatikan bahwa beberapa variabel mungkin memiliki nol yang ditentukan secara arbitrarium, seperti dalam pengukuran suhu dalam skala Kelvin.

Meskipun memiliki keterbatasan ini skala data rasio tetap memberikan manfaat yang signifikan dalam analisis statistika. Data rasio memungkinkan peneliti untuk membuat perbandingan absolut yang akurat antara nilai-nilai yang diamati serta untuk melakukan operasi matematika yang lebih lanjut dalam analisis data secara statistika (Wang, Xu and Brown, 2021).

BAGIAN 5

KUALITAS DATA

A. PENGERTIAN KUALITAS DATA

Secara etimolog data adalah bentuk jamak dari datum, yang berarti "sesuatu yang diberikan" dan partisip masa lampau netral dari dare, yang berarti "memberi", kata data seringkali digunakan dalam bentuk tunggal, dengan "data-data" sebagai bentuk jamaknya. Data yang merupakan kumpulan angka, fenomena atau keadaan lainnya. Dalam bahasa sehari-hari kata "data" paling sering digunakan dalam bentuk tunggal sebagai kata benda massal. Menurut Kuswandi dan E. Mutiara mendeskripsikan data sebagai sekumpulan informasi yang diperoleh dari pengamatan, dapat berupa simbol, angka, dan properti. Dan menurut Margareta Rahayuningsih menyatakan bahwa data adalah angka, simbol, teks, atau gambar yang dapat dikumpulkan dan diproses untuk menghasilkan informasi.

Dan menurut Arikunto Suharsimi mendefinisikan data sebagai serangkaian fakta dan angka yang dapat digunakan sebagai komponen untuk menyusun informasi. Menurut Slamet Riyadi, data adalah kumpulan informasi yang diperoleh melalui pengamatan, bisa berbentuk angka atau simbol.

Dapat disimpulkan data merupakan sekumpulan Fakta yang dikumpulkan yang kemudian di deskripsikan dalam bentuk angka, diagram, serta disampaikan dalam bentuk tulisan dan lainnya. Dan Perlu ditekankan bahwa data bersifat mentah, sehingga untuk memastikan kebenaran, akurasi, ketepatan waktu, dan cakupannya, sangat penting untuk mengandalkan data yang telah terpercaya. Oleh karena sifat mentahnya, informasi yang diperoleh dari data belum sepenuhnya utuh. Dengan demikian, sering kali diperlukan pengolahan data agar dapat menghasilkan informasi yang mudah dipahami. Sedangkan secara umum kualitas dapat pula di artikan sebagai mutu yang menunjukkan tingkat baik atau buruknya sesuatu.

Kualitas data tidak dapat didefinisikan dengan satu kata saja, namun dapat didefinisikan dengan beberapa kata. Kualitas data tidak dapat ditentukan hanya dengan beberapa kata penentu yang berkontribusi terhadap kualitasnya. Kualitas data adalah faktor yang memberikan nilai tambah pada berbagai cara penggunaan data. Kualitas dan data sangat berhubungan karena sekumpulan fakta bisa dikatakan baik dan buruk tergantung dari kualitasnya. Dan dampak langsung kualitas data akan langsung terlihat dan mempengaruhi aspek yang dinilai. menggunakan data yang kurang berkualitas akan memberikan hasil yang buruk dan menyebabkan lebih banyak kerugian. Keputusan yang diambil berdasarkan data yang tidak akurat mempunyai dampak buruk jangka panjang. Hal ini memungkinkan Anda memahami bahwa data tersebut memiliki signifikansi yang kuat.

B. DATA

Metode pengumpulan data adalah teknik atau cara yang digunakan oleh peneliti untuk menghimpun informasi yang diperlukan guna mencapai tujuan penelitian. Tujuan dari pengumpulan data ini adalah untuk mendapatkan data yang relevan dengan penelitian yang sedang dilakukan. Selain itu, instrumen pengumpulan data adalah alat yang digunakan dalam proses pengumpulan data. Beberapa metode pengumpulan data yang perlu dipahami meliputi:

1. Studi dokumen adalah metode pengumpulan data secara tidak langsung yang memfokuskan pada pembahasan subjek penelitian. Tinjauan dokumen melibatkan penelitian berbagai jenis dokumen yang relevan untuk analisis dokumen.
2. Wawancara adalah salah satu teknik pengumpulan data yang melibatkan interaksi tatap muka antara peneliti dan responden. Pertanyaan diajukan dan dijawab secara langsung, meskipun dengan perkembangan teknologi, wawancara juga dapat dilakukan melalui media seperti telepon, email, atau Skype.
3. Kuesioner merupakan metode pengumpulan data dengan memberikan pertanyaan atau pernyataan tertulis kepada responden untuk dijawab. Metode ini efisien jika peneliti mengetahui variabel yang akan diukur dan memahami harapan dari responden.
4. Pengamatan atau observasi merupakan metode pengumpulan data kompleks karena melibatkan faktor-faktor yang beragam

dalam implementasinya. Observasi tidak hanya digunakan untuk mengukur sikap responden, tetapi juga untuk merekam berbagai fenomena yang terjadi selama pengumpulan data.

5. Metode pengumpulan data eksperimental dilakukan melalui penyelidikan dan penelitian

Berdasarkan Cara Mendapatkannya data terbagi menjadi 2 yaitu.

1. Data Primer

Data primer merujuk pada informasi yang diperoleh secara langsung dari objek penelitian, baik oleh suatu organisasi maupun perorangan. Contohnya, data dari survei, wawancara, dan kuesioner.

2. Data Sekunder

Data sekunder merujuk pada informasi yang diperoleh dari sumber yang telah ada sebelumnya. Individu tidak perlu mengumpulkan data langsung dari objek penelitian, melainkan dapat memanfaatkan hasil penelitian sebelumnya dalam bentuk grafik, tabel, atau diagram. Misalnya, data penyakit spesifik, sensus penduduk, dan lainnya.

Berdasarkan Sumber data terbagi menjadi;

1. Data Internal

Data internal diperoleh langsung dari organisasi atau lokasi penelitian

2. Data Eksternal

Data eksternal diperoleh dari sumber di luar organisasi atau lokasi penelitian. Jenis data ini sering digunakan untuk membandingkan satu tempat dengan tempat lainnya, seperti

Jika ditinjau berdasarkan sifatnya, data terbagi menjadi dua, yaitu data kuantitatif dan data kualitatif yang semuanya di sebut dengan Data Statistik. Dari dinilainya, data di bagi menjadi dua yaitu Data Distrik dan Data Kontinum. Perhitungan terhadap suatu variabel menghasilkan data distrik. Sedangkan pengukuran terhadap suatu variabel akan menghasilkan data kontinum.

Data distrik (Kategori/Nominal) adalah gejala yang hanya dapat dibagi (Dikelompokan) menurut jenisnya. Angka yang dilekatkan pada variabel distrik adalah angka kuantitatif yang dihasilkan dari perhitungan (penjumlahan, dan data kontinum adalah data yang hanyadapat dibagi (dikelompokkan) menurut tingkatan besar kecilnya. Data kontinum adalah data yang diperoleh melalui pengukuran. Data kontinum dapat dikelompokkan menjadi tiga ordinal, interval dan rasio.

C. KARAKTERISTIK KUALITAS DATA

Kualitas data adalah ukuran seberapa cocok suatu kumpulan data untuk memenuhi tujuan spesifiknya. Ukuran kualitas data didasarkan pada karakteristik kualitas data seperti keakuratan, kelengkapan, konsistensi, validitas, keunikan, dan ketepatan waktu.

Kumpulan data yang baik harus mempertimbangkan kualitas seperti kesesuaian, kredibilitas, relevansi, keandalan, atau kegunaan. 4 Prinsip Mutlak tentang Kualitas Data menurut Philip B. Crosby antara lain :

1. Defenisi Kualitas Data adalah Kesesuaian.
2. Sistem mutu adalah pencegahan, bukan sekedar pemeriksaan atau penilaian.
3. Standar performa adalah Zero Defects (relative terhadap persyaratan)
4. Pengukuran kualitas adalah harga ketidaksesuaian bukan indeks.

Dalam proses menghasilkan Data yang berkualitas BPS (Badan Pusat Statistik) melakukan 3 proses, yaitu *Quality Governance* yaitu pendekatan yang strategis dalam mengelola, mengamankan, dan memastikan kualitas data secara efektif. Selanjutnya melakukan *Quality Management* yang bertujuan untuk menjaga kualitas data yang tinggi. Dan terakhir *Corporate Infrastructure* yaitu melakukan proses yang merupakan salah satu solusi untuk membuat sebuah sistem yang mengintegrasikan. Berikut syarat umum dari kualitas data;

1. Relevansi, sejauh mana data memenuhi kebutuhan pengguna.
2. Akurasi, kemampuan data dalam menjelaskan fenomena secara tepat.

3. Aktualitas dan ketetapan waktu, Aktualitas mengacu pada perbedaan antara waktu suatu data/informasi statistik dihasilkan dengan waktu data/informasi tersebut didiseminasikan
4. Ketepatan waktu mengacu pada diseminasi data sesuai dengan jadwal yang dijanjikan.
5. Koherensi dan keterbandingan, Koherensi merujuk kepada kemampuan data untuk digabungkan dan digunakan secara bersama sedangkan Keterbandingan merujuk kepada kemampuan data untuk dapat dibandingkan dari waktu ke waktu dan antar wilayah.
6. Aksesibilitas, Seberapa mudah pengguna dapat mengakses data/informasi (berikut metadata)
7. Interpretabilitas, Pemahaman dan kejelasan dari suatu data.

Dimensi kualitas data berperilaku dan diukur secara berbeda di setiap tingkat data. Secara teknis, semua metrik kualitas data terbagi dalam dua kategori besar: yang pertama berkaitan dengan karakteristik intrinsik data, sedangkan yang kedua berkaitan dengan karakteristik kontekstualnya, dan data berkualitas tinggi harus memiliki Karakteristik berikut :

- Ketepatan. Data tersebut dengan benar mewakili entitas atau peristiwa yang seharusnya diwakilinya, dan data tersebut berasal dari sumber yang dapat diverifikasi dan dipercaya.
- Konsistensi. Datanya seragam di seluruh sistem dan kumpulan data, dan tidak ada konflik antara nilai data yang sama di sistem atau kumpulan data yang berbeda.

- Keabsahan. Data tersebut sesuai dengan aturan dan parameter bisnis yang ditentukan, yang memastikan bahwa data terstruktur dengan benar dan berisi nilai yang seharusnya.
- Kelengkapan. Data mencakup semua nilai dan tipe data yang diharapkan ada di dalamnya, termasuk metadata apa pun yang harus menyertai kumpulan data.
- Ketepatan waktu. Data tersebut terkini (relatif terhadap persyaratan spesifiknya) dan tersedia untuk digunakan saat diperlukan.
- Keunikan. Data tidak berisi catatan duplikat dalam satu kumpulan data, dan setiap catatan dapat diidentifikasi secara unik.

Data berkualitas tinggi tentu saja bebas dari masalah kualitas data, seperti data yang terduplikasi, data yang tidak lengkap, data yang tidak konsisten, data yang salah, data yang tidak terdefinisi dengan baik, data yang tidak terorganisir dengan baik, hingga keamanan data yang buruk. Manfaat Kualitas Data yang baik sebagai berikut;

1. Pengambilan keputusan Data berkualitas melampaui fakta akurasi dan relevansi. Sebaliknya, hal ini menambah keyakinan pada pengambil keputusan bahwa keputusan tidak dibuat berdasarkan firasat dan dorongan hati. Sebaliknya, keputusan dibuat berdasarkan data fisik, perhitungan, dan analisis yang tepat.
2. Informasinya tidak perlu bersifat tidak langsung dan menyesatkan dengan data yang berkualitas. Dapat digunakan

untuk membuat sasaran langsung di area tertentu. Hal ini berkontribusi dengan menambahkan fokus dan membuat prediksi dan penilaian yang lebih baik.

3. Dengan data yang berkualitas, rencana tindakan dapat dibuat untuk pertumbuhan dan pembangunan.
4. Lebih sedikit waktu yang dihabiskan untuk merekonsiliasi data karena kualitas data dapat diandalkan dan akurat, hal ini menghilangkan seluruh proses rekonsiliasi data. Tentu saja data tersebut harus diperiksa, namun dapat mempersingkat lamanya proses validasi data dengan sumbernya.
5. kualitas data berasal dari pemeliharaan keakuratan dan relevansi, kecil kemungkinan terjadinya kesalahan dalam data. Ketika data berkualitas dihasilkan, hal itu menimbulkan kepercayaan dalam mengambil keputusan karena kepercayaan terhadap sistem. Oleh karena itu, data yang berkualitas mengarah pada keputusan yang berkualitas.

Data berkualitas tinggi menghasilkan kolaborasi yang lebih baik, komunikasi yang lebih jelas, hubungan pemangku kepentingan yang lebih baik, dan organisasi yang lebih sehat secara keseluruhan.

D. DATA STATISTIKA

Data statistik adalah konsep penting yang perlu dipahami untuk menerapkan pengukuran statistik dengan benar, dan akan

menghasilkan perhitungan yang mendekati real bila kualitas data memenuhi standar.

Skala pengukuran merupakan konsep penting dalam penelitian dan statistik. Ia digunakan untuk mengukur variabel dalam sebuah penelitian. Menurut Prof. Asep Saefudin dan Lukmanul Hakim (2022) dalam statistik, skala pengukuran merupakan penentu dalam memilih metode atau rumus yang akan digunakan. Berbeda skala pengukuran maka akan berbeda rumus serta teknik-teknik yang akan digunakan:

1. Skala Nominal

Skala ini termasuk dalam pengukuran kualitatif dan merupakan yang paling sederhana diantara ketiga jenis lainnya.

2. Skala Ordinal

Skala ini juga termasuk ke dalam skala pengukuran kualitatif. Ia menggunakan urutan tertentu guna memberi peringkat, dan efektif untuk data yang memerlukan urutan ketika melakukan evaluasi.

3. Skala Interval

Skala interval adalah pengukuran dengan skala numerik di mana nilai-nilai yang berdekatan memiliki jarak yang sama dan diukur sepanjang skala.

4. Skala Rasio

Rasio adalah skala pengukuran kuantitatif yang memudahkan kita untuk mencari perbedaan antar variabel dan memberi urutan data. Skala pengukuran ini memungkinkan untuk

melakukan semua jenis perhitungan dan menarik kesimpulan pasti. Perbedaannya dengan interval adalah rasio tidak memiliki nilai negatif.

Jika ditinjau berdasarkan sifatnya, data terbagi menjadi dua, yaitu data kuantitatif dan data kualitatif yang semuanya di sebut dengan Data Statistik. Dari dinilainya, data di bagi menjadi dua yaitu Data Distrik dan Data Kontinum. Perhitungan terhadap suatu variabel menghasilkan data distrik. Sedangkan pengukuran terhadap suatu variabel akan menghasilkan data kontinum.

Data distrik (Kategori/Nominal) adalah gejala yang hanya dapat dibagi (Dikelompokan) menurut jenisnya. Angka yang dilekatkan pada variabel distrik adalah angka kuantitatif yang dihasilkan dari perhitungan (penjumlahan, dan data kontinum adalah data yang hanyadapat dibagi (dikelompokkan) menurut tingkatan besar kecilnya. Data kontinum adalah data yang diperoleh melalui pengukuran. Data kontinum dapat dikelompokkan menjadi tiga ordinal, interval dan rasio.

Pengukuran kualitas data *sebaris* adalah pengukuran data berkelanjutan yang diintegrasikan ke dalam pemrosesan data. Hal ini difokuskan pada penemuan atau konfirmasi karakteristik fisik dan isi data. Bergantung pada ukuran dan kompleksitas data yang Anda kelola serta frekuensi pembaruannya, pengukuran in-line kemungkinan besar akan ditargetkan pada subkumpulan data dan hubungan penting. Tujuan dari pengawasan data secara berkala

adalah untuk mengidentifikasi perubahan signifikan dalam pola data yang mungkin disebabkan oleh masalah kualitas data. Dalam kasus seperti ini, pemantauan harus memiliki kendali bawaan yang dapat mendeteksi ketika suatu proses tidak terkendali dan mempunyai kemampuan untuk mempengaruhi proses tersebut.

BAGIAN 6

DESKRIPSI DATA

A. KONSEP DATA

Didalam statistik data merupakan salah satu komponen yang penting di dalam memecahkan masalah-masalah statistik. Salah satu kegunaan dari data adalah dapat digunakan untuk menentukan alat analisis statistik apa yang sesuai untuk digunakan. Terdapat beberapa jenis data di dalam statistik tergantung penggolongannya dapat dilihat pada Tabel 1.

Tabel 6.1 Jenis-Jenis Data

Penggolongan Data		Pengertian	Contoh
Sifatnya	Data Kuantitatif	Data yang bersifat numerik (angka). data yang berbentuk angka pasti, yang dapat diukur (<i>measurable</i>) atau dapat dihitung menggunakan variabel angka atau bilangan	Tingkat Suku Bunga, Inflasi, Jumlah Uang Beredar, Upah Minimum Provinsi

	Data Kualitatif	Data non-numerik yang dikuantitatifkan (dikoding). Data dari penjelasan kata verbal sehingga tidak dapat dianalisis dalam bentuk bilangan atau angka.	Jenis Kelamin, Persepsi Konsumen, Pendidikan, Agama
Periode Waktu	Data <i>Time Series</i>	Data yang pengamatannya dilakukan dari waktu ke waktu (satu objek, banyak waktu)	Laba PT. X selama 5 tahun, Jumlah Uang beredar selama 10 tahun
	Data <i>Cross Section</i>	Data yang pengamatannya dilakukan pada satu waktu dengan banyak objek	Harga Saham Bank Persero, Data Jenis Kelamin PT. X
	Data <i>Pooling</i>	Data yang merupakan	Inflasi 34 Provinsi di

		gabungan antara data <i>time series</i> dan data <i>cross-section</i> (banyak objek, banyak waktu)	Indonesia Selama 5 Tahun, Harga Saham PT. X Periode 2010-2020
Cara Pengambilan	Data Primer	Data yang berasal dari sumber aslinya, dikumpulkan khusus untuk keperluan riset	Data Kuesioner, Data <i>Indep Interview</i> , Data hasil FGD
	Data Sekunder	Data yang dikumpulkan melalui pihak lain yang berasal dari sumber <i>internal</i> atau <i>eksternal</i> dari suatu organisasi	Data-data yang diperoleh dari Lembaga resmi seperti BPS, BI, BEI

Dalam sebuah penelitian, seseorang peneliti akan selalu membutuhkan data, karena melalui data peneliti akan mendapatkan informasi yang terpercaya. Oleh karena itu data dibutuhkan untuk memberikan penjelasan yang berkaitan dengan masalah tertentu. Terdapat beberapa fungsi dari data dalam suatu penelitian, antara lain:

1. Data digunakan sebagai bahan evaluasi dalam menanggapi suatu masalah tertentu
2. Data berfungsi untuk memecahkan sebuah masalah dan menentukan suatu kebijakan serta keputusan
3. Data yang diperoleh juga dapat digunakan sebagai acuan dalam setiap implementasi suatu kegiatan atau aktivitas.
4. Data berfungsi sebagai dasar suatu penelitian atau perencanaan dalam sebuah kegiatan.

B. ANALISIS DESKRIPSI

Deskripsi data merupakan upaya untuk menampilkan data agar data dapat ditunjukkan secara baik dan mudah. Dalam deskripsi data penyusunan terhadap data meliputi bentuk tampilan yang mudah untuk dibaca serta lengkap. Deskripsi data akan memberikan gambaran yang nyata dari responden, sehingga lebih mudah di mengerti oleh peneliti atau pembaca yang tertarik dengan hasil penelitian yang dilakukan.

Tujuan dilakukan analisis deskriptif dengan menggunakan teknik statistika adalah untuk meringkas data agar menjadi lebih mudah dilihat dan dimengerti. Analisis data yang paling sederhana dan sering digunakan oleh peneliti adalah menganalisis data yang ada dengan menggunakan prinsip-prinsip deskriptif. Dengan menganalisis secara deskriptif ini mereka dapat mempresentasikan secara ringkas, sederhana, dan lebih mudah dimengerti.

Tujuan dari melakukan analisis deskriptif dengan menggunakan teknik statistika adalah untuk meringkas data agar menjadi lebih mudah dilihat dan dimengerti. Berdasarkan teori statistik, statistik akan terbagi menjadi dua, yaitu statistik deskriptif dan statistik inferensial. Statistik Deskriptif merupakan suatu metode yang berkaitan dengan mengorganisir, menyimpulkan dan mempresentasikan data dalam bentuk lebih yang lebih informatif misalkan dalam bentuk tabel dan grafis. Statistik Inferensial merupakan metode yang menggunakan rumus statistik dan hasil perhitungannya dijadikan sebagai pedoman dalam membuat kesimpulan secara umum atau mengeneralisasikannya.

Analisis data yang paling sederhana dan sering digunakan oleh peneliti adalah menganalisis data dengan menggunakan prinsip-prinsip deskriptif. Dengan menganalisis secara deskriptif ini mereka dapat mempresentasikan secara ringkas, sederhana, dan lebih mudah dimengerti. Dalam mendeskripsikan informasi dari data yang digunakan dalam suatu penelitian akan dibagi dalam dua macam, tergantung sifat dari sebuah data, apakah data Kuantitatif atau data Kualitatif. Penggunaan data kualitatif, maka deskripsi data dilakukan dengan cara menyusun dan mengelompokkan data yang ada kedalam kelas-kelas tertentu atau disebut dengan Distribusi Frekuensi Kategorikal. Penggunaan data kuantitatif, maka cara mendeskripsi data dapat dilakukan dengan menggunakan statistika deskriptif yang mencakup Distribusi Frekuensi Numerikal, Gejala Pusat (*Central Tendency*), Ukuran Penyebaran, Kemencengan dan

Keruncingan, sehingga dapat menampilkan ringkasan data yang lebih informatif.

C. DISTRIBUSI FREKUENSI

Distribusi Frekuensi juga dapat diartikan sebagai rangkaian data angka yang bersifat kuantitatif atau kualitatif. Rangkaian data angka yang bersifat kuantitatif akan disebut sebagai Distribusi Frekuensi Numerikal dan yang bersifat kualitatif akan disebut sebagai Distribusi Frekuensi Kategorikal. Tabel didalam distribusi frekuensi merupakan alat penyajian data statistika yang berbentuk baris dan kolom. Alat penyajian data statistik yang berbentuk kolom dan baris yang di dalamnya dimuat angka-angka yang menggambarkan pembagian frekuensi dari variabel yang sedang dijadikan objek penelitian.

(1) Distribusi Frekuensi Kategorikal

Merupakan distribusi frekuensi yang pembagian kelas-kelasnya didasarkan pada golongan data yang bersifat kualitatif. Data Kualitatif merupakan data non-numerik yang dikuantitatifkan dengan tujuan Penelitian tertentu.

Contoh:

- Jenis kelamin (1: Laki-Laki dan 2: Perempuan);
- Tingkat Pendidikan (1: SD, 2: SMP, 3: SMA, 4: Diploma, 5: Sarjana dan 6: Pasca Sarjana);

- Persepsi Kepuasan (1: Sangat Tidak Puas, 2: Tidak Puas, 3: Cukup Puas, 4: Puas dan 5: Sangat Puas).

Membentuk Distribusi Frekuensi Kategorikal

Contoh: Berikut adalah data jenis kelamin dari 30 orang karyawan di PT. Sentosa

L L L P P L P P L P
 P L L P L P P L P P
 P P L L L P P P L L

Keterangan: L = Laki-Laki dan P = Perempuan

Pertanyaan: Bentuklah distribusi frekuensi PT. Sentosa Berdasarkan jenis kelamin?

Langkah-langkah pengerjaan, maka tentukan dulu jumlah kelas yang akan dibentuk, untuk jenis kelamin sudah jelas bahwa kategori nya ada dua yaitu kategori laki-laki dan Perempuan maka kelas yang akan dibentuk sebanyak 2 kelas, sehingga distribusi frekuensi sebagai berikut:

Tabel 6.2. Distribusi Frekuensi Berdasarkan Jenis Kelamin

Kelas	Frekuensi Absolut (Orang)	Frekuensi Relatif (%)
Laki-Laki	14	47%
Perempuan	16	53%
Jumlah	30	100%

Berdasarkan Tabel 2 yang memberikan informasi mengenai Jenis Kelamin PT. Sentosa, dapat disimpulkan dari 30 karyawan PT. Sentosa sebanyak 14 orang berjenis kelamin laki-laki atau sebanyak 47% dari total karyawan berjenis kelamin laki-laki. Karyawan yang berjenis kelamin Perempuan ada sebanyak 16 orang dari 30 karyawan PT. Sentosa atau sebanyak 53% karyawan PT. Sentosa berjenis kelamin perempuan. Disimpulkan pada PT. Sentosa karyawan Perempuan lebih banyak dibandingkan karyawan laki-laki.

(2) Distribusi Frekuensi Numerikal

Merupakan distribusi frekuensi yang pembagian kelas-kelasnya didasarkan pada golongan data yang bersifat kuantitatif. Data Kuantitatif merupakan data numerik (angka).

Contoh: Data Inflasi, Nilai Ujian, Umur, Harga Saham, Berat Badan, Tinggi Badan, Tingkat Suku Bunga, Nilai Tukar, Laba, Penjualan.

Membentuk Distribusi Frekuensi Numerikal

Contoh: Berikut adalah data penjualan dari 80 orang karyawan di PT. Agung Maju yang bergerak dalam bidang obat-obatan (Juta Rp) selama bulan Januari 2024

47	61	70	72	75	80	82	87	91	94
43	59	68	71	74	79	81	86	90	93
45	60	68	72	74	79	82	86	90	93
51	65	70	73	76	80	83	88	91	96
38	58	67	71	74	78	81	85	88	93

50	61	70	72	75	80	83	88	91	95
54	66	70	73	76	80	83	88	92	97
35	57	67	71	74	77	81	84	88	92

Pertanyaan: Berdasarkan data diatas bentuklah distribusi frekuensi frekuensi atau susunlah data menjadi data berkelompok?

Berdasarkan soal diatas diketahui bahwa terdapat sampel (n) sebanyak 80 orang karyawan, maka untuk membentuk distribusi frekuensi ada beberapa tahap yang dapat dilakukan yaitu:

(a) Tentukan Jumlah Kelas

Aturan Struges sering digunakan untuk membentuk data dalam distribusi frekuensi. Dalam sebuah tabel distribusi frekuensi akan memuat data yang telah diatur dalam kelas-kelas dengan interval (panjang kelas) yang sama. Penentuan banyak kelas pada penyusunan tabel distribusi frekuensi kelompok dalam aturan Sturges dapat menggunakan rumus:

$k = 1 + 3,322 \log n.$ Dimana:
 k = jumlah kelas pada tabel distribusi frekuensi
 n = banyaknya sampel

berdasarkan soal diatas maka $k = 1 + 3,322 \text{ Log}(80)$
 $k = 1 + 3,322 (1,903)$
 $k = 7,322$
 dibulatkan menjadi 7

hasil perhitungan jumlah kelas yang harus dibentuk sebanyak 7 kelas.

- (b) Tentukan Besarnya Range

Range merupakan jarak antara data terkecil dengan data terbesar dari distribusi data yang ada. Rumus yang dapat digunakan adalah:

Range = Nilai Terbesar – Nilai Terkecil

Pada soal diketahui penjualan terkecil PT. Sentosa sebesar Rp. 35 Juta dan nilai terbesar Rp. 97 Juta. Maka perhitungan mencari Range = Rp. 97 Juta – Rp. 35 Juta = Rp. 62 Juta.

Hasil perhitungan menunjukkan range penjualan dari 80 karyawan PT. Sentosa sebesar Rp. 62 Juta.

- (c) Tentukan Panjang Kelas (Interval Kelas)

Interval kelas merupakan perbedaan antara suatu class boundary dengan *class boundary* sebelumnya. Panjang interval kelas (ci) dapat menggunakan rumus:

$ci = \frac{Range}{\sum Kelas}$ Pada tahap sebelumnya kita sudah menghitung bahwa jumlah kelas yang harus terbentuk sebanyak 7 kelas dan *Range* sebesar 62, maka $ci = 62/7 = 8,857$ dibulatkan menjadi 9.

Hasil perhitungan menunjukkan panjang interval setiap kelas yang akan berisikan informasi penjualan dari 80 karyawan PT. Sentosa sebesar Rp. 9 Juta.

(d) Membentuk Kelas

Dalam membentuk kelas hal-hal yang harus diperhatikan adalah pastikan semua data harus bisa masuk didalam kelas yang dibentuk. Nilai terkecil harus masuk pada kelas pertama dan nilai terbesar harus masuk pada kelas terakhir. Pastikan tidak terjadi *double accounting* dalam memasukkan data kesetiap kelas, maksudnya pastikan interval setiap kelas tidak memiliki nilai rentang yang sama, untuk menghindari hal tersebut maka kita dapat mengurangi atau menambahkan diujung nilai.

(e) Mencari Frekuensi Tiap-Tiap Kelas

Dilakukan dengan cara memasukkan nilai dari data mentah kedalam kelas-kelas nya.

Berdasarkan Tabel 3 kita akan mendapatkan informasi yang lebih mudah dalam mengidentifikasi karakteristik penjualan dari PT. Agung Maju selama bulan Januari 2024. Informasi yang bisa didapatkan dari total 80 karyawan PT. Agung Maju diketahui bahwa paling banyak berada pada interval nilai penjualan antara Rp. 80 Juta sampai Rp. 88 Juta yaitu sebanyak 22 karyawan atau 27,5%. Paling sedikit sebanyak 3 orang karyawan atau 3,75% dari total karyawan memiliki interval nilai penjualan antara Rp. 35 Juta sampai dengan Rp. 43 Juta.

Tabel 3. Distribusi Frekuensi Penjualan PT. Agung Maju

Kelas	Penjualan (Juta Rp)	Tally	Frekuensi Absolut (Orang)	Frekuensi Relatif (%)
I	35 – 43	III	3	3,75%
II	44 – 52	IIII	4	5,00%
III	53 – 61	IIII II	7	8,75%
IV	62 – 70	IIII IIIII	10	12,5%
V	71 – 79	IIII IIIII IIIII IIIII	20	25,0%
VI	80 – 88	IIII IIIII IIIII IIIII II	22	27,5%
VII	89 – 97	IIII IIIII IIII	14	17,5%
	Jumlah		80	100%

D. GEJALA PUSAT

Permasalahan yang sering terjadi disaat menggunakan distribusi frekuensi adalah pada saat menggunakan data yang cenderung homogen, sehingga memberikan kesulitan dalam pengambilan kesimpulan, sehingga diperlukan pengukuran lain yang sifatnya eksplisit untuk menjelaskan karkateristik data sebuah Penelitian. Terdapat tiga pengukuran eksplisit untuk menjelaskan karakteristik dari sebuah data, yaitu: Gejala Pusat (*Central Tendency*), Ukuran Penyebaran (*Dispersion*), Ukuran Kemencengan (*Skewness*) dan Keruncingan (*Kurtosis*).

Gejala Pusat merupakan suatu ukuran lokasi atau ukuran nilai rata-rata yang mempunyai kecenderungan memusat, sehingga sering disebut ukuran kecenderungan memusat (*measures of central*

tendency). Secara umum ukuran lokasi yang digunakan adalah rata-rata, median, modus, kuartil, desil, persentil:

1. Rata-rata

Merupakan jumlah semua data dibagi dengan banyaknya data. Secara formulasi, rata-rata hitung dinyatakan dengan rumus sebagai berikut:

Rata-Rata Hitung Populasi:

$$\mu = \frac{\sum X}{N}$$

Dimana:
 μ = rata-rata populasi
 $\sum X$ = jumlah nilai seluruh observasi populasi
 N = Jumlah populasi

Rata-Rata Hitung Sample:

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

Dimana:
 \bar{x} = rata-rata sampel
 $\sum x$ = jumlah nilai seluruh observasi sampel
 n = Jumlah sampel

Contoh:

Terdapat pedagang UMKM di Jakarta Barat, diambil sampel sebanyak 10 pedagang UMKM untuk nilai produksinya (Juta Rp) per bulan pada bulan Januari 2024.

UMKM	1	2	3	4	5	6	7	8	9	10
Penjualan	20	15	25	17	15	17	10	12	17	20

hitunglah rata-rata produksi dari 10 UMKM tersebut!

$$\bar{x} = \frac{\sum x}{n} = \frac{20 + 15 + 25 + 17 + 15 + 17 + 10 + 12 + 17 + 20}{10} = 16,80$$

Berdasarkan hasil perhitungan, rata-rata produksi UMKM di Jakarta Barat adalah Rp 16,80 Juta

2. Median

Merupakan data tengah atau data yang membagi barisan data menjadi dua sama banyak. Jika banyak datanya ganjil, maka median adalah data yang paling tengah setelah pengurutan (membesar atau mengecil), jika banyak datanya genap maka median adalah rata-rata dari dua nilai tengahnya setelah pengurutan. Penggunaan median sangat cocok untuk data yang memiliki sifat tidak terpengaruhi oleh nilai ekstrim dan data yang heterogen. Secara formulasi, median dinyatakan dengan rumus sebagai berikut:

$$\text{Letak Median} = \left(\frac{n + 1}{2} \right) \quad n = \text{jumlah observasi}$$

Menggunakan data produksi UMKM di Jakarta Barat, hitunglah nilai tengah (median)! Langkah-langkah pengerjaan sebagai berikut:

- Urutkan data dari terkecil ke terbesar

10 12 15 15 17 17 17 20 20 25

- Kemudian tentukan letak median menggunakan rumus:
 $(10+1)/2 = 5,5$ maka nilai median terletak di data ke 5 dan ke 6 yaitu $(17+17)/2 = 17$

Berdasarkan hasil perhitungan nilai median untuk produksi UMKM di wilayah Jakarta Barat sebesar Rp 17 Juta.

3. Modus

Merupakan nilai dari observasi yang paling sering atau banyak muncul.

Menggunakan data produksi UMKM di Jakarta Barat, hitunglah nilai modus nya! Langkah-langkah pengerjaan sebagai berikut:

Urutkan data dari terkecil ke terbesar

10 12 15 15 17 17 17 20 20 25

Diketahui bahwa nilai yang paling sering ada nilai 17

E. UKURAN PENYEBARAN

Ukuran Penyebaran atau ukuran dispersi atau ukuran penyimpangan merupakan ukuran yang menunjukkan besar kecilnya perbedaan data dari rata-rata. Ukuran penyebaran mampu melihat seberapa jauh penyimpangan nilai data dari nilai pusatnya atau ukuran yang mampu menunjukkan seberapa banyak nilai data yang berbeda dengan nilai pusatnya. Kelemahan dari ukuran gejala pusat adalah kurang mampu membandingkan beberapa rangkaian data jika tidak cermat dalam menganalisa mampu memberikan hasil yang menyesatkan. Oleh karena itu ukuran penyebaran digunakan sebagai pelengkap dari ukuran gejala pusat dalam menggambarkan sekumpulan data, dimana mampu menggambarkan sekumpulan data yang lebih jelas dan akurat. Ukuran penyebaran membantu mengetahui sejauh mana suatu nilai menyebar dari nilai tengahnya, semakin kecil semakin besar.

Terdapat beberapa pendekatan perhitungan untuk mengukur penyebaran sekumpulan data antara lain yang seringkali digunakan adalah dengan pendekatan Varian dan Standar Deviasi.

1. Standar Deviasi

Digunakan untuk mengetahui homogenitas dari data serta dapat menjadi proksi dari sebuah ukuran risiko atau stabilitas data. Ukuran variabilitas yang sering digunakan adalah simpangan baku atau standar deviasi. Standar Deviasi adalah akar dari varian. Secara formulasi, standar deviasi dinyatakan dengan rumus sebagai berikut:

Standar Deviasi Populasi:

$$\sigma = \sqrt{\frac{\sum [X_i - \mu]^2}{N}}$$

Dimana:

- σ = standar deviasi populasi
- X_i = nilai observasi ke-i
- μ = nilai rata-rata populasi
- N = jumlah populasi

Standar Deviasi Sample:

$$s = \sqrt{\frac{\sum [X_i - \bar{x}]^2}{n - 1}}$$

Dimana:

- s = standar deviasi sampel
- X_i = nilai observasi ke-i
- \bar{x} = nilai rata-rata sampel
- n = jumlah sample

Menggunakan data produksi UMKM Jakarta Barat, hitunglah penyebaran data menggunakan standar deviasi!

x_i	$x_i - \bar{x}$	$[X_i - \bar{x}]^2$
20	$(20-16,8) = 3,2$	10,2
15	$(15-16,8) = -1,8$	3,2
25	$(25-16,8) = 8,2$	67,2

17	$(17-16,8) = 0,2$	0,0
15	$(15-16,8) = -1,8$	3,2
17	$(17-16,8) = 0,2$	0,0
10	$(10-16,8) = -6,8$	46,2
12	$(12-16,8) = -4,8$	23,0
17	$(17-16,8) = 0,2$	0,0
20	$(20-16,8) = 3,2$	10,2
$\bar{x} = 16,8$		$[X_i - \bar{x}]^2 = 163,6$

$$\text{Maka, } s = \sqrt{\frac{163,6}{10 - 1}} = 4,3$$

Jadi besarnya standar deviasi dari produksi UMKM Jakarta Barat adalah 4,3. Hal ini berarti rata-rata penyebaran setiap observasi terhadap rata-ratanya sebesar Rp 4,3 Juta.

2. Varian

Merupakan rata-rata hitung dari deviasi kuadrat setiap data terhadap rata-rata hitungnya. Varians adalah alat ukur variabilitas serangkaian data yang dihitung dengan mencari rata-rata selisih atau beda kuadrat antara data observasi dengan pusat datanya. Secara formulasi, varian dinyatakan dengan rumus sebagai berikut:

Varian Populasi:

$$\sigma^2 = \frac{\sum [X_i - \mu]^2}{N}$$

Dimana:

- σ^2 = Varian populasi
- X_i = nilai observasi ke-i
- μ = nilai rata-rata populasi
- N = jumlah populasi

Varian Sample:

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

Dimana:

s^2	=	Varian sample
x_i	=	nilai observasi ke-i
\bar{x}	=	nilai rata-rata sample
n	=	jumlah sample

Menggunakan data produksi UMKM Jakarta Barat, lakukan perhitungan penyebaran data menggunakan varian. Hasil perhitungan standar deviasi sebesar 4,3 maka besarnya nilai penyebaran menggunakan rumus varian $s^2 = (4,3)^2 = 18,2$ angka ini menunjukkan rata-rata kuadrat penyebaran setiap observasi terhadap rata-ratanya sebesar Rp 18,2 Juta.

F. UKURAN KEMENCENGAN

Ukuran Kemencengan merupakan tingkat ketidaksimetrisan atau kejauhan simetri dari sebuah distribusi. Sebuah distribusi yang tidak simetris akan memiliki rata-rata, median, dan modus yang tidak sama besarnya ($\text{Mean} \neq \text{Median} \neq \text{Modus}$), sehingga distribusi akan terkonsentrasi pada salah satu sisi dan kurvanya akan condong. Berdasarkan formulasi Pearson, rumus untuk koefisien skewness dinyatakan sebagai berikut:

$$Sk = \frac{3 [\text{mean} - \text{median}]}{\text{standar deviasi}}$$

Nilai koefisien skewness berkisar dari -3 sampai +3, apabila nilai koefisien skewness negatif menunjukkan kurva condong ke kanan (condong negatif), apabila koefisien skewness positif menunjukkan kurva condong ke kiri (condong positif). Nilai yang mendekati nol maka distribusi frekuensi tersebut semakin simetris, bila koefisien kecondongan positif, ekor kanan distribusi frekuensinya lebih panjang dari ekor kirinya, bila koefisien kecondongan negatif, ekor kiri distribusi frekuensinya lebih panjang dari ekor kanannya.

Menggunakan data produksi UMKM Jakarta Barat, lakukanlah pengujian tingkat kemecengan dengan koefisien skewness dari distribusi data produksi UMKM. Berdasarkan hasil perhitungan didapatkan nilai rata-rata sebesar 16,8 dan nilai median sebesar 17 serta nilai standar deviasi sebesar 4,3 maka nilai koefisien skewness sebesar:

$$Sk = \frac{3 [16,8 - 17]}{4,3} = -0,14$$

Berdasarkan hasil perhitungan diketahui nilai koefisien skewness sebesar -0,14 (negatif) hal ini menunjukkan bahwa distribusi data menceng kanan, artinya data produksi UMKM banyak berada pada batas atas (nilai tinggi).

G. UKURAN KERUNCINGAN

Ukuran Keruncingan merupakan tingkat kepuncakan dari sebuah distribusi yang biasanya diambil secara relatif terhadap suatu distribusi normal. Ukuran keruncingan (kurtosis) sekumpulan data atau suatu distribusi adalah suatu ukuran yang dapat digunakan untuk menentukan runcing tidaknya kurva suatu distribusi. Formulasi perhitungan koefisien alpha 4 adalah sebagai berikut:

$$\alpha_4 = \frac{\frac{1}{n} \sum [x_i - \bar{x}]^4}{s^4}$$

Dimana:

\bar{x}	= nilai rata-rata sample
x_i	= nilai-nilai setiap observasi
n	= jumlah sample
s	= standar deviasi sample

Perhitungan nilai koefisien alpha 4 biasa digunakan untuk mengukur keruncingan distribusi suatu data. Koefisien alpha 4 merupakan rata-rata dari selisih antara data-data dengan rata-rata (mean) pangkat empat dibagi standar deviasi pangkat 4.

Jenis-jenis keruncingan kurva ada tiga, yaitu:

- Leptokurtis = Sangat runcing ($\alpha_4 > 3$)
- Mesokurtis = Keruncingan sedang ($\alpha_4 = 3$)
- Platikurtis = Kurva datar ($\alpha_4 < 3$)

Menggunakan data produksi UMKM Jakarta Barat, tentukan tingkat keruncingan dengan menggunakan rumus koefisien alpha 4! Berdasarkan hasil perhitungan sebelumnya diketahui besarnya standar deviasi (s) sebesar 4,3.

x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^4$
20	10,2	104,9
15	3,2	10,5
25	67,2	4521,2
17	0,0	0,0
15	3,2	10,5
17	0,0	0,0
10	46,2	2138,1
12	23,0	530,8
17	0,0	0,0
20	10,2	104,9
$\bar{x} = 16,8$		$(x_i - \bar{x})^4 = 7.420,9$

$$\text{maka, } \alpha_4 = \frac{1}{10} [7.420,9] = 2,17$$

Didapatkan hasil nilai koefisien alpha 4 sebesar $2,17 < 3$ maka distribusi landai atau tumpul (platikurtis).

BAGIAN 7

KONSEP PROBABILITAS

Ada ketidakpastian dalam pengambilan keputusan yang disebabkan adanya risiko. Untuk itu diperlukan pemahaman tentang teori probabilitas, atau sering disebut sebagai ilmu ketidakpastian. Teori probabilitas memungkinkan pengambil keputusan untuk menganalisis risiko dan meminimalkan kerugian yang ada. Misalnya, dalam memasarkan produk baru atau menerima kiriman barang yang mungkin berisi bagian yang rusak. Karena konsep probabilitas sangat penting dalam bidang statistik inferensi, bab ini memperkenalkan bahasa dasar probabilitas, termasuk istilah-istilah seperti eksperimen, peristiwa, probabilitas subjektif, dan aturan perjumlahan dan perkalian.

A. DEFINISI PROBABILITAS, EKSPERIMEN, HASIL (*OUTCOME*), DAN KEJADIAN (*EVENTS*)

Probabilitas, eksperimen, hasil, dan kejadian merupakan konsep dasar dalam teori probabilitas yang digunakan untuk memahami dan memprediksi kemungkinan terjadinya suatu peristiwa dalam suatu situasi tertentu.

1. Probabilitas

Probabilitas adalah suatu nilai numerik yang menunjukkan kemungkinan terjadinya suatu kejadian dalam suatu eksperimen.

Nilai ini dapat berkisar antara 0 (mustahil) hingga 1 (pasti). Probabilitas dihitung dengan membagi jumlah hasil yang menguntungkan dengan total kemungkinan hasil dalam suatu eksperimen.

Contoh:

- (a) Melempar koin: Probabilitas mendapatkan sisi atas adalah $1/2$, karena hanya ada satu sisi atas dari dua sisi koin yang mungkin muncul.
- (b) Mengambil kartu dari satu set berisi 52 kartu: Probabilitas mendapatkan kartu As adalah $4/52$, karena ada 4 kartu As dari 52 kartu yang mungkin muncul.

2. Eksperimen

Eksperimen adalah suatu proses yang dilakukan untuk mengamati suatu fenomena dan menghasilkan hasil. Eksperimen dapat dilakukan secara acak atau terkontrol.

Contoh:

- (a) Melempar koin
- (b) Mengambil kartu dari satu set
- (c) Memutar roda roulette
- (d) Mengadakan percobaan kimia

3. Hasil (Outcome)

Hasil adalah salah satu kemungkinan yang muncul dalam suatu eksperimen. Hasil dapat berupa angka, simbol, atau deskripsi dari suatu peristiwa.

Contoh:

- (a) Dalam eksperimen melempar koin, hasil yang mungkin muncul adalah sisi atas atau sisi bawah.
- (b) Dalam eksperimen mengambil kartu dari satu set, hasil yang mungkin muncul adalah salah satu dari 52 kartu di dalam set tersebut.

4. Kejadian (Events)

Kejadian adalah suatu subset dari hasil yang mungkin muncul dalam suatu eksperimen. Kejadian dapat berupa satu hasil atau lebih. Kejadian dapat diklasifikasikan sebagai saling melingkupi atau saling lepas.

Contoh:

- (a) Dalam eksperimen melempar koin, kejadian "mendapatkan sisi atas" adalah satu hasil.
- (b) Dalam eksperimen mengambil kartu dari satu set, kejadian "mendapatkan kartu As" atau "mendapatkan kartu berwarna merah" adalah dua kejadian yang saling lepas.

B. PENDEKATAN PENGHITUNGAN PROBABILITAS

Bagian ini menjelaskan tiga cara untuk menghitung probabilitas suatu peristiwa: klasik, empiris, dan subjektif. Metode klasik dan empiris bersifat objektif dan didasarkan pada informasi dan data. Metode subjektif didasarkan pada keyakinan atau perkiraan seseorang terhadap kemungkinan suatu peristiwa:

1. Probabilitas Klasik

Probabilitas klasik adalah metode untuk menghitung probabilitas suatu kejadian dalam suatu eksperimen dengan cara menghitung rasio antara jumlah hasil yang menguntungkan dengan total kemungkinan hasil dalam eksperimen tersebut. Metode ini hanya dapat digunakan pada eksperimen yang memenuhi tiga syarat berikut:

- a. Eksperimen harus memiliki hasil yang pasti dan sama. Artinya, setiap kali eksperimen dilakukan dengan cara yang sama, akan selalu menghasilkan hasil yang sama.
- b. Semua hasil yang mungkin muncul harus sama peluangnya. Artinya, tidak ada hasil yang lebih mungkin muncul daripada hasil lainnya.
- c. Kejadian yang ingin dihitung probabilitasnya harus terdefinisi dengan jelas. Artinya, harus jelas mana yang termasuk dalam kejadian tersebut dan mana yang tidak.

Rumus Probabilitas Klasik:

Probabilitas (P) suatu kejadian (A) dalam suatu eksperimen dengan n hasil yang mungkin dapat dihitung dengan rumus berikut:

$$P(A) = n(A) / n$$

di mana:

P(A) adalah probabilitas kejadian A

n(A) adalah jumlah hasil yang menguntungkan (hasil yang termasuk dalam kejadian A)

n adalah total kemungkinan hasil dalam eksperimen

Contoh: Mengambil sebuah kartu dari satu set berisi 52 kartu. Berapakah probabilitas mendapatkan kartu As?

Solusi: Eksperimen ini memenuhi ketiga syarat probabilitas klasik. Ada 52 hasil yang mungkin muncul dalam eksperimen ini, yaitu setiap kartu di dalam set. Semua kartu memiliki peluang yang sama untuk muncul. Kejadian yang ingin dihitung probabilitasnya adalah mendapatkan kartu As. Ada 4 kartu As dalam set ini. Berdasarkan rumus probabilitas klasik, probabilitas mendapatkan kartu As adalah:

$$P(\text{kartu As}) = n(\text{kartu As}) / n = 4 / 52 = 1/13$$

2. Probabilitas Empiris

Probabilitas empiris, juga dikenal sebagai probabilitas frekuensi atau probabilitas eksperimental, adalah metode untuk menghitung probabilitas suatu kejadian dalam suatu eksperimen dengan cara menghitung rasio antara jumlah kemunculan kejadian tersebut dengan jumlah total percobaan yang dilakukan. Metode ini dapat digunakan pada eksperimen apa pun, terlepas dari apakah eksperimen tersebut memenuhi syarat-syarat probabilitas klasik atau tidak.

Rumus Probabilitas Empiris:

Probabilitas (P) suatu kejadian (A) dalam suatu eksperimen dengan n percobaan dapat dihitung dengan rumus berikut:

$$P(A) = n(A) / n$$

di mana:

P(A) adalah probabilitas kejadian A

$n(A)$ adalah jumlah kemunculan kejadian A dalam n percobaan
 n adalah total percobaan yang dilakukan:

Contoh: Mengambil sebuah kartu dari satu set berisi 52 kartu sebanyak 100 kali. Berapakah probabilitas empiris mendapatkan kartu As?

Solusi: Eksperimen ini dapat diulang sebanyak 100 kali. Pada setiap pengambilan, ada 52 hasil yang mungkin muncul, yaitu setiap kartu di dalam set. Kejadian yang ingin dihitung probabilitasnya adalah mendapatkan kartu As. Misalkan dalam 100 pengambilan, kartu As muncul sebanyak 12 kali. Berdasarkan rumus probabilitas empiris, probabilitas empiris mendapatkan kartu As adalah:

$$P(\text{kartu As}) = n(\text{kartu As}) / n = 12 / 100 = 0,12$$

Probabilitas empiris adalah metode yang lebih umum dan fleksibel dibandingkan dengan probabilitas klasik. Metode ini dapat digunakan pada berbagai jenis eksperimen, termasuk eksperimen yang tidak memenuhi syarat-syarat probabilitas klasik. Namun, perlu diingat bahwa probabilitas empiris merupakan perkiraan dari probabilitas yang sebenarnya. Semakin banyak percobaan yang dilakukan, semakin akurat perkiraan probabilitas empiris tersebut.

3. Probabilitas Subjektif

Probabilitas subjektif, juga dikenal sebagai probabilitas Bayesian atau probabilitas personal, adalah metode untuk menghitung probabilitas suatu kejadian berdasarkan penilaian, kepercayaan, atau pengetahuan individu tentang kejadian tersebut. Metode ini tidak didasarkan pada data empiris atau eksperimen, melainkan

pada penilaian subjektif individu terhadap kemungkinan terjadinya suatu peristiwa.

Faktor-faktor yang Mempengaruhi Probabilitas Subjektif

- a. Individu akan menggunakan pengetahuan dan pengalaman mereka tentang kejadian serupa di masa lalu untuk menilai probabilitasnya.
- b. Individu akan mempertimbangkan informasi yang tersedia bagi mereka saat membuat penilaian mereka.
- c. Kepercayaan dan prasangka individu dapat memengaruhi penilaian mereka tentang probabilitas suatu kejadian.
- d. Emosi dan intuisi individu juga dapat berperan dalam penilaian mereka tentang probabilitas suatu kejadian.

Contoh:

- a. Seorang dokter mungkin memiliki probabilitas subjektif yang berbeda untuk pasien yang sama tergantung pada pengetahuan dan pengalamannya dalam menangani kasus serupa.
- b. Seorang investor mungkin memiliki probabilitas subjektif yang berbeda untuk nilai saham yang sama tergantung pada informasi dan analisisnya.

Keuntungan Probabilitas Subjektif

- a. Probabilitas subjektif dapat digunakan untuk menghitung probabilitas kejadian yang sulit atau tidak mungkin atau fleksibel diukur secara objektif.
- b. Probabilitas subjektif dapat diperbarui seiring dengan munculnya informasi baru.

- c. Probabilitas subjektif dapat mempertimbangkan faktor-faktor subjektif seperti kepercayaan, prasangka, dan emosi.

Kekurangan Probabilitas Subjektif

- a. Probabilitas subjektif dapat tidak objektif dan bias karena dipengaruhi oleh penilaian individu.
- b. Probabilitas subjektif sulit untuk diukur dan dikuantifikasi secara akurat.
- c. Probabilitas subjektif dapat berbeda-beda antar individu, sehingga sulit untuk mencapai konsensus.

Probabilitas subjektif tidak objektif dan dapat bias karena dipengaruhi oleh penilaian individu. Probabilitas subjektif harus digunakan dengan hati-hati dan dikombinasikan dengan metode lain, seperti probabilitas klasik atau probabilitas empiris, untuk mendapatkan penilaian yang lebih objektif dan akurat.

C. PROBABILITAS MENGGUNAKAN ATURAN PENJUMLAHAN

Probabilitas aturan penjumlahan adalah rumus yang digunakan untuk menghitung peluang terjadinya **satu atau lebih** kejadian dalam suatu ruang sampel. Ada dua jenis aturan penjumlahan peluang, yaitu untuk kejadian saling lepas dan kejadian tidak saling lepas:

1. Aturan Penjumlahan Peluang untuk Kejadian Saling Lepas

Dua kejadian dikatakan saling lepas (*mutually exclusive*) jika terjadinya satu kejadian tidak mempengaruhi kemungkinan terjadinya kejadian lain. Dengan kata lain, tidak mungkin kedua kejadian terjadi secara bersamaan.

Rumus aturan penjumlahan peluang untuk kejadian saling lepas adalah:

$$P(A \cup B) = P(A) + P(B)$$

di mana:

$P(A \cup B)$ adalah peluang terjadinya satu atau lebih kejadian A dan B.

$P(A)$ adalah peluang terjadinya kejadian A.

$P(B)$ adalah peluang terjadinya kejadian B.

Contoh: Mengambil kartu dari satu set berisi 52 kartu

A = mendapatkan kartu As

B = mendapatkan kartu berwarna merah

$$P(A) = 4/52 = 1/13$$

$$P(B) = 26/52 = 1/2$$

$$P(A \cup B) = P(A) + P(B) = 1/13 + 1/2 = 2/26 + 13/26 = 15/26$$

Penjumlahan peluang terambilnya kartu berwarna merah ditulis $P(A)$, ditambah probabilitas bahwa kartu tersebut bukan berwarna merah, ditulis $P(\sim A)$ dan dibaca “bukan A”, secara logis harus sama dengan 1. Ditulis: $P(A) + P(\sim A) = 1$. Aturan ini yang disebut aturan komplemen atau pelengkap.

Aturan pelengkap ini digunakan untuk menentukan peluang terjadinya suatu peristiwa dengan mengurangkan peluang tidak terjadinya suatu peristiwa dengan 1. Aturan ini berguna karena terkadang lebih mudah menghitung peluang terjadinya suatu peristiwa dengan menentukan peluang tidak terjadinya suatu peristiwa dan mengurangkannya hasil dari 1. Perhatikan bahwa kejadian A dan $\sim A$ saling lepas dan lengkap secara kolektif. Oleh karena itu, probabilitas A dan $\sim A$ berjumlah 1.

2. Aturan Penjumlahan Peluang untuk Kejadian Tidak Saling Lepas

Dua kejadian dikatakan tidak saling lepas (*non-mutually exclusive*) jika terjadinya satu kejadian dapat mempengaruhi kemungkinan terjadinya kejadian lain. Dengan kata lain, kedua kejadian dapat terjadi secara bersamaan.

Rumus aturan penjumlahan peluang untuk kejadian tidak saling lepas adalah:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

di mana:

$P(A \text{ atau } B)$ adalah peluang terjadinya satu atau lebih kejadian A dan B .

$P(A)$ adalah peluang terjadinya kejadian A .

$P(B)$ adalah peluang terjadinya kejadian B .

$P(A \cap B)$ adalah peluang terjadinya kedua kejadian A dan B secara bersamaan.

Contoh: Mengambil kartu dari satu set berisi 52 kartu:

A = mendapatkan kartu As

B = mendapatkan kartu berwarna hati

$$P(A) = 4/52 = 1/13$$

$$P(B) = 13/52$$

$$P(A \cap B) = 1/52$$

$$P(A \cup B) = P(A)+P(B) - P(A \cap B) = 1/13 + 13/52 - 1/52 = 4/13$$

D. PROBABILITAS MENGGUNAKAN ATURAN PERKALIAN

Dalam teori probabilitas, aturan perkalian peluang digunakan untuk menghitung probabilitas terjadinya dua atau lebih kejadian secara bersyarat. Dua macam aturan perkalian peluang yaitu untuk kejadian saling lepas dan tidak saling lepas.

1. Aturan Perkalian untuk Kejadian Saling Lepas

Kejadian saling lepas (mutually exclusive events) adalah kejadian yang tidak dapat terjadi secara bersamaan.

Rumus:
$$P(A \cap B) = P(A) * P(B)$$

di mana:

$P(A \cap B)$ adalah probabilitas terjadinya A dan B secara bersamaan.

$P(A)$ adalah probabilitas terjadinya A.

$P(B)$ adalah probabilitas terjadinya B.

Contoh: Lempar koin dua kali. Peluang mendapatkan sisi kepala pada lemparan pertama dan ekor pada lemparan kedua adalah:

$$P(\text{kepala di lemparan pertama} \cap \text{ekor di lemparan kedua}) = \\ P(\text{kepala di lemparan pertama}) * P(\text{ekor di lemparan kedua}) = 1/2 \\ * 1/2 = 1/4$$

2. Aturan Perkalian untuk Kejadian Tidak Saling Lepas

Kejadian tidak saling lepas (non-mutually exclusive events) adalah kejadian yang dapat terjadi secara bersamaan.

Rumus:
$$P(A \cap B) = P(A) * P(B|A)$$

di mana:

$P(A \cap B)$ adalah probabilitas terjadinya A dan B secara bersamaan.

$P(A)$ adalah probabilitas terjadinya A.

$P(B|A)$ adalah probabilitas terjadinya B dengan syarat A telah terjadi.

Contoh: Ambil dua kartu dari satu set kartu remi yang ter kocok dengan baik. Peluang mendapatkan kartu As hati dan kartu Raja sekop adalah:

$$P(\text{As hati} \cap \text{Raja sekop}) = P(\text{As hati}) * P(\text{Raja sekop} | \text{As hati})$$

$$P(\text{As hati}) = 1/52 \text{ (probabilitas awal mendapatkan kartu As hati)}$$

$$P(\text{Raja sekop} | \text{As hati}) = 1/51 \text{ (probabilitas mendapatkan kartu Raja sekop setelah mengambil kartu As hati)}$$

$$P(\text{As hati} \cap \text{Raja sekop}) = 1/52 * 1/51 \approx 0,0004$$

E. PROBABILITAS MENGGUNAKAN TABEL KONTIGENSI

Tabel kontingensi adalah tabel yang digunakan untuk menampilkan hubungan antara dua atau lebih variabel kategorikal. Probabilitas yang dihitung dari tabel kontingensi dapat digunakan untuk memahami hubungan antara variabel-variabel tersebut dan untuk membuat prediksi tentang kemungkinan hasil di masa depan.

Tabel ini dapat digunakan untuk menghitung berbagai macam probabilitas, termasuk:

1. Probabilitas marginal: Probabilitas kemunculan setiap kategori dalam suatu variabel.
2. Probabilitas bersyarat: Probabilitas kemunculan suatu kategori dalam suatu variabel dengan asumsi kategori lain telah terjadi.
3. Probabilitas gabungan: Probabilitas kemunculan dua atau lebih kategori dari dua atau lebih variabel secara bersamaan.

Contoh: Misalkan kita memiliki tabel kontingensi yang menunjukkan jenis kelamin dan hasil ujian matematika dari 100 siswa:

Jenis Kelamin	Lulus	Tidak Lulus	Total
Laki-laki	30	20	50
Perempuan	25	25	50
Total	55	45	100

Probabilitas marginal:

- a. Probabilitas seorang siswa laki-laki adalah $50/100 = 0,5$.
- b. Probabilitas seorang siswa perempuan adalah $50/100 = 0,5$.
- c. Probabilitas seorang siswa lulus adalah $55/100 = 0,55$.

d. Probabilitas seorang siswa tidak lulus adalah $45/100 = 0,45$.

Probabilitas bersyarat:

a. Probabilitas seorang siswa lulus dengan asumsi dia laki-laki adalah $30/50 = 0,6$.

b. Probabilitas seorang siswa tidak lulus dengan asumsi dia laki-laki adalah $20/50 = 0,4$.

c. Probabilitas seorang siswa lulus dengan asumsi dia perempuan adalah $25/50 = 0,5$.

d. Probabilitas seorang siswa tidak lulus dengan asumsi dia perempuan adalah $25/50 = 0,5$.

Probabilitas gabungan:

a. Probabilitas seorang siswa laki-laki dan lulus adalah $30/100 = 0,3$.

b. Probabilitas seorang siswa laki-laki dan tidak lulus adalah $20/100 = 0,2$.

c. Probabilitas seorang siswa perempuan dan lulus adalah $25/100 = 0,25$.

d. Probabilitas seorang siswa perempuan dan tidak lulus adalah $25/100 = 0,25$.

F. PROBABILITAS MENGGUNAKAN TEOREMA BAYES

Teorema Bayes adalah konsep penting dalam teori probabilitas yang memungkinkan kita untuk memperbarui probabilitas awal tentang suatu kejadian berdasarkan informasi atau bukti baru.

Konsep Dasar:

- a. Hipotesis: Kemungkinan penjelasan atau prediksi tentang suatu fenomena.
- b. Probabilitas Prior: Probabilitas awal tentang suatu hipotesis sebelum mempertimbangkan bukti baru.
- c. Probabilitas Posterior: Probabilitas yang diperbarui untuk suatu hipotesis setelah mempertimbangkan bukti baru.
- d. Bukti: Informasi yang diperoleh yang dapat mendukung atau menolak suatu hipotesis.

Rumus Teorema Bayes

$$P(B | A) = (P(A | B) * P(B)) / P(A)$$

di mana:

$P(B|A)$ adalah probabilitas posterior, probabilitas hipotesis B setelah kita mengetahui bukti A.

$P(A|B)$ adalah probabilitas sebaliknya, probabilitas bukti A terjadi dengan syarat hipotesis B benar.

$P(B)$ adalah probabilitas prior, probabilitas awal hipotesis B sebelum mempertimbangkan bukti A.

$P(A)$ adalah probabilitas total dari bukti A, terlepas dari hipotesis mana yang benar.

Teorema Bayes memungkinkan kita untuk menghitung probabilitas posterior, yang mencerminkan tingkat kepercayaan yang diperbarui terhadap suatu hipotesis setelah mempertimbangkan bukti baru.

Contoh Penggunaan:

- a. Diagnosis penyakit: Dokter dapat menggunakan teorema Bayes untuk memperbarui diagnosis awal mereka berdasarkan hasil tes medis.
- b. Klasifikasi spam email: Filter spam email dapat menggunakan teorema Bayes untuk menilai apakah email baru adalah spam atau bukan, berdasarkan kata-kata yang ada di dalamnya.
- c. Analisis keandalan produk: Insinyur dapat menggunakan teorema Bayes untuk memperbarui perkiraan mereka tentang tingkat kegagalan produk berdasarkan laporan kerusakan.

Contoh: Seorang pasien datang ke dokter dengan gejala demam dan batuk. Dokter memiliki dua hipotesis: pasien menderita flu atau pasien menderita pneumonia.

Peluang Prior: Berdasarkan data prevalensi pneumonia di wilayah tersebut diketahui bahwa probabilitas Flu: 0,7 dan probabilitas Pneumonia: 0,3.

Bukti: Gejala demam dan batuk

Peluang Bersyarat:

- a. $P(\text{demam dan batuk} \mid \text{Flu}) = 0,8$ (Kemungkinan pasien dengan flu mengalami demam dan batuk)
- b. $P(\text{demam dan batuk} \mid \text{Pneumonia}) = 0,9$ (Kemungkinan pasien dengan pneumonia mengalami demam dan batuk)

Solusi:

$$\begin{aligned}
 P(\text{Pneumonia} \mid \text{demam dan batuk}) &= (P(\text{demam dan batuk} \mid \text{Pneumonia}) * P(\text{Pneumonia})) / P(\text{demam dan batuk}) \\
 &= (0,9 * 0,3) / (0,8 * 0,7 + 0,9 * 0,3)
 \end{aligned}$$

$$= 0.27 / 0,83$$

$$= 0,3253$$

Atau dengan tabel

Hipotesis	Probabilitas Prior	Probabilitas bersyarat	Probabilitas gabungan	Probabilitas Posterior
	$P(A_i)$	$P(B A_i)$	$P(A_i \text{ dan } B)$	$P(A_i B)$
Flu	0,7	0,8	0,56	0,6745
Pneumonia	0,3	0,9	0,27	0,3253
			$P(B) = 0,83$	

Berdasarkan teorema Bayes, probabilitas posterior bahwa pasien menderita pneumonia adalah 0,3253. Artinya, setelah mempertimbangkan gejala demam dan batuk, dokter lebih yakin bahwa pasien menderita pneumonia daripada flu.

G. PRINSIP PENGHITUNGAN

Jika jumlah kemungkinan hasil dalam suatu percobaan kecil, maka relatif mudah untuk menghitungnya. Misalnya, ada enam kemungkinan hasil pelemparan sebuah dadu, yaitu: 1, 2, 3, 4, 5, dan 6. Akan tetapi, jika terdapat banyak kemungkinan hasil, misalnya banyaknya kepala dan ekor pada suatu percobaan yang pelemparannya sepuluh kali, maka akan butuh waktu untuk menghitung semua kemungkinan. Hasilnya bisa muncul semua kepala, satu kepala dan sembilan ekor, dua kepala dan delapan ekor, dan seterusnya. Untuk memudahkan penghitungan, kami

menjelaskan tiga rumus: rumus perkalian, rumus permutasi, dan rumus kombinasi.

Beberapa prinsip penghitungan yang paling umum digunakan:

1. Aturan Perkalian

Aturan perkalian menyatakan bahwa untuk menghitung jumlah kemungkinan hasil yang diperoleh dari dua kejadian yang saling bebas, kita kalikan probabilitas masing-masing kejadian.

Rumus:

$$P(A \cap B) = P(A) * P(B)$$

di mana:

$P(A \cap B)$ adalah probabilitas terjadinya kedua kejadian A dan B secara bersamaan.

$P(A)$ adalah probabilitas terjadinya kejadian A.

$P(B)$ adalah probabilitas terjadinya kejadian B.

Contoh:

Melempar koin dua kali:

A = mendapatkan sisi atas pada lemparan pertama

B = mendapatkan sisi atas pada lemparan kedua

$$P(A) = 1/2$$

$$P(B) = 1/2$$

$$P(A \cap B) = P(A) * P(B) = 1/2 * 1/2 = 1/4$$

Mengambil dua kartu dari satu set berisi 52 kartu

A = mendapatkan kartu As pada pengambilan pertama

B = mendapatkan kartu Raja pada pengambilan kedua

$$P(A) = 4/52 = 1/13$$

$P(B|A) = 3/51$ (karena setelah mengambil kartu As, tersisa 51 kartu dan 3 kartu Raja)

$$P(A \cap B) = P(A) * P(B|A) = 1/13 * 3/51 = 1/234$$

2. Aturan Permutasi

Aturan permutasi menyatakan bahwa untuk menghitung jumlah kemungkinan urutan dari n elemen yang berbeda dengan memperhatikan urutan, kita kalikan faktorial dari n .

Rumus: ${}_n P_r = \frac{n!}{(n-r)!}$

di mana:

$n!$ adalah faktorial dari n , $n! = n * (n - 1) * (n - 2) * \dots * 2 * 1$

n adalah jumlah elemen.

Contoh:

Mengatur 3 buku yang berbeda di rak:

$$n = 3 \text{ (jumlah buku)}$$

$$3! = 3 * 2 * 1 = 6 \text{ (jumlah kemungkinan urutan)}$$

Memilih 3 orang dari 10 orang untuk menjadi anggota tim:

$$n = 10 \text{ (jumlah orang)}$$

$$k = 3 \text{ (jumlah anggota tim)}$$

$${}_{10}P_3 = 10!/(10 - 3)! = 10*9*8 = 720 \text{ (jumlah kemungkinan tim)}$$

3. Aturan Kombinasi

Aturan kombinasi menyatakan bahwa untuk menghitung jumlah kemungkinan memilih k elemen dari n elemen yang berbeda, tanpa memperhatikan urutan pemilihan, kita gunakan rumus kombinasi.

Rumus: ${}_n C_k = \frac{n!}{k!(n-k)!}$

di mana:

${}_n C_k$ adalah kombinasi dari n elemen yang diambil k elemen.

n adalah jumlah elemen.

k adalah jumlah elemen yang dipilih.

Contoh:

Memilih 2 buah apel dari 5 buah apel yang berbeda:

$n = 5$ (jumlah apel)

$k = 2$ (jumlah apel yang dipilih)

${}_5 C_2 = 5! / (2! * (5 - 2)!) = 10$ (jumlah kemungkinan kombinasi)

BAGIAN 8

DISTRIBUSI PROBABILITAS DISKRIT

A. PENGERTIAN DISTRIBUSI PROBABILITAS DISKRIT

Distribusi probabilitas diskrit adalah sebuah konsep dalam statistik yang menggambarkan probabilitas untuk semua kemungkinan hasil dari suatu variabel acak dimana variabel tersebut dapat mempunyai nilai berbeda yang jumlahnya dapat terbatas atau tak terbatas. Jumlah total dari semua nilai probabilitas adalah 1 atau 100%.

Distribusi probabilitas diskrit yang paling umum adalah binomial, Bernoulli, multinomial, dan Poisson. Setiap distribusi memiliki aturan dan parameter sendiri yang mempengaruhi probabilitas dari setiap nilai dalam semua kemungkinan hasil.

B. DISTRIBUSI BINOMIAL

Distribusi binomial adalah distribusi probabilitas diskrit yang hanya memberikan dua kemungkinan hasil dalam suatu eksperimen, yaitu sukses atau gagal. Contohnya, jika sebuah koin (mata uang rupiah) dilempar, hanya ada dua kemungkinan hasil: gambar atau angka, dan jika ada eksperimen yang dilakukan, maka hanya ada dua kemungkinan hasil: sukses atau gagal. Distribusi ini disebut juga distribusi probabilitas binomial.

Rumus:

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x}$$

Dimana:

$P(x)$ adalah hasil probabilitas dengan jumlah sukses adalah x .

x adalah jumlah total eksperimen yang sukses.

n adalah jumlah total eksperimen atau percobaan.

p adalah probabilitas sukses dari setiap eksperimen.

$(1 - p)$ adalah probabilitas gagal dari setiap eksperimen.

Contoh soal:

Berapakah probabilitas atau peluang mendapatkan sisi gambar tiga kali dalam lima kali pelemparan sebuah koin?

Jawaban:

Koin memiliki dua kemungkinan hasil dalam setiap lemparan, sisi gambar atau sisi angka.

$p = 0,5$ (peluang mendapatkan sisi gambar di setiap eksperimen)

$(1 - p) = 1 - 0,5 = 0,5$ (peluang mendapatkan sisi angka di setiap eksperimen)

$x = 3$ (jumlah total sukses)

$n = 5$ (jumlah total pelemparan atau eksperimen)

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1 - p)^{n-x}$$

$$P(3) = \frac{5!}{3!(5-3)!} 0,5^3 \times 0,5^{5-3} = \frac{3!}{3!2!} \times 0,125 \times 0,25$$

$$P(3) = \frac{120}{6 \times 2} \times 0,125 \times 0,25 = 0,3125 = 31,25\%$$

Probabilitas mendapatkan sisi gambar tiga kali dalam lima kali pelemparan sebuah koin adalah 31,25%

C. DISTRIBUSI BERNOULLI

Distribusi Bernoulli adalah distribusi probabilitas diskrit dalam statistik yang memodelkan satu eksperimen atau uji coba dengan dua kemungkinan hasil yaitu sukses atau gagal. Distribusi Bernoulli diberi nama dari ahli matematika asal Swiss yaitu Jacob Bernoulli.

Rumus:

$$P(x) = p^x (1 - p)^{1-x}$$

Dimana:

$P(x)$ adalah hasil probabilitas sukses yaitu x .

p adalah probabilitas sukses dari eksperimen

x = eksperimen sukses.

Contoh:

Seorang pelatih basket percaya bahwa peluang pemainnya melakukan tembakan bebas dengan sukses dalam satu percobaan adalah 0,75.

Analisa menggunakan distribusi Bernoulli untuk situasi ini:

Probabilitas bahwa pemain berhasil melakukan tembakan bebas (sukses) adalah $P(x = 1) = 0,75$.

Probabilitas bahwa pemain tidak berhasil melakukan tembakan bebas (gagal) adalah $P(x = 0) = 1 - 0,75 = 0,25$.

D. DISTRIBUSI MULTINOMIAL

Distribusi multinomial adalah distribusi probabilitas diskrit yang merupakan pengembangan dari distribusi binomial dimana terdapat

lebih dari dua kemungkinan hasil untuk setiap eksperimen, tidak hanya sukses dan gagal saja. Setiap eksperimen dapat menghasilkan satu dari k kemungkinan hasil.

Rumus:

$$P(X) = \frac{n!}{x_1!x_2!\dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$

Dimana:

$P(X)$ adalah hasil probabilitas dengan jumlah sukses adalah X .

x_1, x_2, \dots, x_k adalah kemungkinan hasil.

n adalah jumlah total eksperimen atau percobaan.

$x_1 + x_2 + \dots + x_k = n$.

k adalah probabilitas sukses dari setiap eksperimen.

p_1, p_2, \dots, p_k adalah probabilitas sukses dari setiap kemungkinan.

Contoh soal:

Pada pemilu dengan tiga kandidat di sebuah negara, kandidat A memperoleh 20% suara, kandidat B memperoleh 30% suara, dan kandidat C memperoleh 50% suara. Jika enam pemilih dipilih secara acak untuk sampel, berapakah probabilitas atau pulang bila dalam sampel terdapat tepat satu pendukung calon A, dua pendukung calon B, dan tiga pendukung calon C?

Jawaban:

$A = x_1 = 1; B = x_2 = 2$ dan $C = x_3 = 3$

$k = 3$ (total kandidat)

$n = x_1 + x_2 + x_3 = 1 + 2 + 3 = 6$

$p_A = 20\% = 0,2$; $p_B = 30\% = 0,3$ dan $p_C = 50\% = 0,5$

$$P(X) = \frac{n!}{x_1!x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$

$$P(x_1 = 1, x_2 = 2, x_3 = 3) = \frac{6!}{1!2!3!} 0,2^1 \times 0,3^2 \times 0,5^3$$

$$P(x_1 = 1, x_2 = 2, x_3 = 3) = 0,135 = 13,5\%$$

Probabilitas bila dalam sampel terdapat tepat satu pendukung calon A, dua pendukung calon B, dan tiga pendukung calon C adalah 13,5%.

E. DISTRIBUSI POISSON

Distribusi Poisson adalah distribusi probabilitas diskrit yang digunakan untuk memodelkan jumlah kejadian yang terjadi dalam interval waktu atau ruang tertentu, dengan tingkat rata-rata kejadian yang diketahui. Distribusi poisson adalah kasus khusus dari distribusi binomial, dimana distribusi binomial akan menjadi distribusi poisson ketika n mendekati tak hingga dan p mendekati nol.

Rumus:

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

Dimana:

$P(x)$ adalah hasil probabilitas

λ adalah jumlah rata-rata kejadian yang terjadi dalam interval tersebut.

e adalah logaritma natural yang nilainya 2,718281828459
 x adalah bilangan bulat non-negatif yang mewakili jumlah kejadian yang diamati.

Contoh soal:

Jika biasanya rata-rata 10 pelanggan datang ke sebuah toko per jam ($\lambda = 10$). Dengan menggunakan distribusi Poisson, berapa probabilitas bila jumlah kedatangan adalah 8 pelanggan dalam satu jam di saat tertentu?

Jawaban:

Probabilitas dengan 8 pelanggan ($x = 8$)

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-10} 10^8}{8!} \approx 0,1126 = 11,26\%$$

Probabilitas bila jumlah kedatangan adalah 8 pelanggan adalah 11,26%

F. DISTRIBUSI GEOMETRIK

Distribusi geometri adalah distribusi probabilitas diskrit yang digunakan dalam statistik yang mengamati jumlah eksperimen atau percobaan yang diperlukan untuk mencapai keberhasilan pertama dalam serangkaian percobaan independen, dimana setiap percobaan hanya memiliki dua kemungkinan hasil (sukses atau gagal). Tidak ada batasan maksimum banyaknya eksperimen yang dilakukan sampai menghasilkan sukses.

Rumus:

$$P(x = k) = (1 - p)^{k-1} p$$

Dimana:

x adalah hasil probabilitas setelah mendapat hasil sukses yang pertama.

p adalah probabilitas sukses dari tiap eksperimen

k adalah jumlah total eksperimen sampai mendapat hasil sukses yang pertama.

Contoh soal:

Sebuah koin dilempar (dimana probabilitas mendapatkan sisi gambar, p, adalah 0,5) berulang kali hingga didapatkan sisi gambar untuk pertama kalinya. Hitung probabilitas untuk nilai k yang berbeda sampai mendapatkan gambar pertama (k lemparan).

Jawaban:

$p = 0,5$ (probabilitas mendapat sisi gambar tiap lemparan)

$$(1 - p) = 1 - 0,5 = 0,5$$

$$P(x = k) = (1 - p)^{k-1} p = (0,5)^{k-1} \times 0,5$$

Sekarang, mari kita hitung probabilitas untuk nilai k yang berbeda (jumlah lemparan hingga sisi gambar pertama).

- Probabilitas mendapat sisi gambar pada lemparan koin pertama ($k = 1$):

$$P(x = 1) = (0,5)^{1-1} \times 0,5 = 1 \times 0,5 = 0,5 = 50\%$$

- Probabilitas mendapat sisi gambar pada lemparan koin kedua ($k = 2$):

$$P(x = 1) = (0,5)^{2-1} \times 0,5 = 0,5 \times 0,5 = 0,25 = 25\%$$

- Probabilitas mendapat sisi gambar pada lemparan koin ketiga ($k = 3$):

$$P(x = 1) = (0,5)^{3-1} \times 0,5 = 0,25 \times 0,5 = 0,125 = 12,5\%$$

dan seterusnya

G. DISTRIBUSI BINOMIAL NEGATIF (PASCAL)

Distribusi binomial negatif, juga dikenal sebagai distribusi Pascal atau distribusi Pólya, adalah distribusi probabilitas diskrit yang digunakan untuk memodelkan jumlah percobaan yang diperlukan untuk mencapai jumlah keberhasilan yang telah ditentukan dalam sebuah rangkaian uji coba Bernoulli yang independen dan identik. Distribusi ini digunakan untuk menggambarkan situasi di mana kita tertarik untuk mengetahui jumlah percobaan yang diperlukan untuk memperoleh jumlah keberhasilan yang tetap, di mana setiap percobaan menghasilkan keberhasilan atau kegagalan. Distribusi binomial negatif mirip dengan distribusi binomial, tetapi dengan pengulangan yang terus-menerus.

Rumus:

$$P(X = r) = \binom{r-1}{k-1} p^k (1-p)^{r-k}$$

$$P(X = r) = \frac{(r-1)!}{(k-1)!((r-1)-(k-1))!} p^k (1-p)^{r-k}$$

Dimana:

X adalah variabel acak yang mewakili jumlah eksperimen hingga mencapai k keberhasilan tercapai.

r adalah jumlah total eksperimen yang diperlukan untuk memperoleh k keberhasilan.

p adalah probabilitas sukses dari tiap eksperimen

k adalah jumlah keberhasilan yang ingin dicapai.

Contoh soal:

Seorang pemain bola basket memiliki rata-rata akurasi lemparan tiga poin sebesar 40%. Berapa banyak percobaan yang dibutuhkan oleh pemain ini untuk mencetak 5 lemparan tiga poin?

Jawaban:

Dalam contoh ini, dapat digunakan distribusi binomial negatif dengan parameter sebagai berikut:

$p=0,4$ (probabilitas keberhasilan dalam setiap lemparan tiga poin)

$k=5$ (jumlah keberhasilan yang ingin dicapai)

Misal ingin dicari probabilitas seorang pemain membutuhkan tepat 10 percobaan untuk mencetak 5 lemparan tiga poin. Maka substitusi nilai $r=10$ ke dalam rumus:

$$P(X = r) = \binom{r-1}{k-1} p^k (1-p)^{r-k}$$

$$P(X = 10) = \binom{10-1}{5-1} 0,4^5 (1-0,4)^{10-5}$$

$$P(X = 10) = \binom{9}{4} 0,4^5 \times 0,6^5 = \frac{9!}{4!(9-4)!} 0,4^5 \times 0,6^5$$

$$P(X = 10) = \frac{9 \times 8 \times 7 \times 6}{4 \times 3 \times 2 \times 1} 0,4^5 \times 0,6^5 = \frac{3024}{24} \times 0,01024 \times 0,07776$$

$$P(X = 10) = 0,1 = 10\%$$

Probabilitas seorang pemain membutuhkan tepat 10 percobaan untuk mencetak 5 lemparan tiga poin adalah 10%

H. DISTRIBUSI SERAGAM DISKRIT (DISCRETE UNIFORM)

Distribusi seragam diskrit adalah distribusi probabilitas diskrit dimana setiap kemungkinan nilai yang terjadi dalam rentang tertentu memiliki peluang yang sama untuk muncul. Setumpuk kartu memiliki distribusi yang seragam karena kemungkinan terambilnya kartu dengan gambar hati, gambar keriting, gambar wajik (berlian), atau gambar sekop (waru) memiliki kemungkinan yang sama. Sebuah koin juga mempunyai distribusi yang seragam karena peluang munculnya sisi gambar atau sisi angka pada pelemparan koin adalah sama.

Rumus:

$$P(x) = \frac{1}{n} \text{ , untuk } x = 1, 2, 3, \dots, n$$

Dimana:

x adalah kemungkinan yang dipilih dari keseluruhan kemungkinan hasil.

n adalah jumlah total kemungkinan hasil.

Contoh:

Misalkan dilempar sebuah dadu dengan sisi enam. Setiap kemungkinan hasil (1, 2, 3, 4, 5 dan 6) mempunyai peluang yang sama yaitu $\frac{1}{6}$ untuk terpilih. Ini mewakili distribusi seragam yang terpisah karena keenam hasil mempunyai kemungkinan yang sama, masing-masing $\frac{1}{6}$.

I. DISTRIBUSI HIPERGEOMETRIK

Distribusi hipergeometrik adalah distribusi probabilitas diskrit yang digunakan ketika ingin menghitung probabilitas terambilnya sejumlah "keberhasilan" (objek dengan tipe tertentu) dari suatu populasi tanpa pengembalian. Artinya setiap pengambilan mempengaruhi probabilitas pengambilan berikutnya karena objek tidak dimasukkan kembali ke dalam populasi awal.

Rumus:

$$P(X = k) = \frac{C_k^K C_{n-k}^{N-K}}{C_n^N} = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$
$$P(X = k) = \left(\frac{\left(\frac{K!}{k!(K-k)!} \right) \left(\frac{N-K!}{(n-k)!((N-K)-(n-k))!} \right)}{\left(\frac{N!}{n!(N-n)!} \right)} \right)$$

Dimana:

C adalah kombinasi (Combination), kombinasi adalah cara penyusunan objek sampel tanpa memperhatikan urutan.

P(X = k) adalah hasil probabilitas.

N adalah jumlah total objek dalam populasi.

K adalah jumlah total objek keberhasilan dalam populasi.

n adalah jumlah total objek yang dijadikan sampel (tanpa pengembalian) dari populasi.

k adalah jumlah objek sampel yang berhasil (dari tipe tertentu).

Contoh soal:

Misalkan kartu remi dengan jumlah awal 52 kartu. Suatu eksperimen mengambil 5 buah kartu berurutan tanpa pengembalian, berapa probabilitas atau peluang terambilnya tepat 2 kartu as (Ace Card)?

Jawaban:

Jumlah awal kartu (N): 52 kartu

Jumlah total kartu as dalam kartu remi (K): 4 (karena ada 4 kartu as di dalam 52 kartu)

Jumlah total kartu yang ditarik (n): 5 (mengambil 5 kartu tanpa pengembalian)

Jumlah kartu as yang ingin ditarik (k): 2 (ingin mendapat tepat 2 kartu as)

Probabilitas terambilnya tepat 2 as dari 5 kartu yang ditarik dapat dihitung dengan menggunakan rumus distribusi hipergeometri:

$$P(X = k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$
$$P(X = 2) = \frac{\binom{4}{2} \binom{52-4}{5-2}}{\binom{52}{5}} = \frac{\binom{4}{2} \binom{48}{3}}{\binom{52}{5}}$$

$$P(X = 2) = \left(\frac{\left(\frac{4!}{2!(4-2)!} \right) \left(\frac{52-4!}{5-2!((52-4)-(5-2))!} \right)}{\left(\frac{52!}{5!(52-5)!} \right)} \right)$$

$$P(X = 2) = \left(\frac{\left(\frac{24}{2 \times 2} \right) \left(\frac{48 \times 47 \times 46}{3 \times 2 \times 1} \right)}{\left(\frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2 \times 1} \right)} \right) = \frac{6 \times 17296}{2598960}$$

$$P(X = 2) = 0,0404 = 4,04\%$$

Probabilitas atau peluang terambilnya tepat 2 kartu as adalah 4,04%

BAGIAN 9

PENGUJIAN HIPOTESIS

A. PENGEMBANGAN HIPOTESIS

Memahami pengujian hipotesis sangat penting untuk membuat keputusan yang tepat, memecahkan masalah, dan memajukan pengetahuan di berbagai disiplin ilmu. Sebelum mengulas lebih jauh mengenai “Pengujian Hipotesis” ini, Ditinjau dari asal-usulnya, kata hipotesis (*hypothesis*, bentuk pluralnya adalah *hypotheses*), tersusun dari kata “*hypo*” yang berarti tentatif, perihal verifikasi atau pokok yang akan dibuktikan, dan “*thesis*” yang berarti pernyataan mengenai solusi terhadap masalah. Jadi, hipotesis menawarkan solusi terhadap masalah, dan karena beberapa alasan sehingga diperlukan verifikasi secara empiris (Singh, 2006). Dalam arti yang operasional, menurut Leavy (2017) hipotesis adalah pernyataan yang meramalkan bagaimana variabel berhubungan satu sama lain dan dapat diuji melalui penelitian. Pandangan senada dikemukakan Sekaran dan Bougie (2016), bahwa hipotesis adalah pernyataan tentatif yang dapat diuji, yang memprediksi apa yang diharapkan peneliti untuk ditemukan dalam data empiris.

Kadangkala hipotesis didefinisikan secara sederhana sebagai “hubungan yang diduga ada antara dua atau lebih variabel yang dinyatakan dalam bentuk pernyataan” (Laka, 2023). Tentu saja, tidaklah salah dengan menyebutkan praduga “ada”, tetapi

pernyataan probabilitas yang mengandung “kebenaran dengan kemungkinan terbesar terjadi”, tentu bukan asumsi dengan fondasi proposisi yang lemah. Pernyataan itu merupakan “*smart guess*” yang berisi “pernyataan formal” yang ingin diketahui buktinya dengan cara yang ilmiah. Cara ilmiah dimaksud mengalir dari hubungan yang melekat pada komponen kajian pustaka bergerak menuju ke kerangka konseptual atau kerangka teoretis, untuk kemudian membuahkan rumusan hipotesis, sebagaimana gambar berikut ini.



Gambar 9.1. Hubungan Kajian Pustaka, Kerangka Konseptual – atau Kerangka Teoretis – dengan Hipotesis.

Setelah peneliti mengidentifikasi variabel-variabel penting hasil kajian pustaka dan menetapkan hubungannya melalui penalaran logis yang dispesifikkan dalam kerangka konseptual – atau kerangka teoretis – pada saat yang demikian, peneliti berada dalam posisi siap untuk menguji apakah hubungan tersebut memang benar. Jadi, hipotesis diturunkan dari kajian pustaka, dan melalui kerangka konseptual – atau kerangka teoretis – dihasilkanlah pernyataan berupa asumsi-asumsi yang bersifat tentatif. Upaya merumuskan

pernyataan yang bersifat tentatif, yang dapat diuji disebut pengembangan hipotesis (*hypothesis development*), yang pada gilirannya akan diuji secara empiris di lapangan.

B. KONSEP DASAR PENGUJIAN HIPOTESIS

Terdapat beberapa “konsep dasar” yang memerlukan peneliti pemahaman dalam rangka pengujian hipotesis. Beberapa konsep dasar dimaksud, antara lain:

1. Hipotesis Nol (H_0) dan Hipotesis Alternatif (H_a)

Hipotesis penelitian adalah jawaban “sementara” terhadap pertanyaan penelitian (*research question*), atau merupakan ‘anggapan dasar’ peneliti terhadap masalah yang sedang dikaji. Untuk “membuktikan” kebenaran dari hipotesis penelitian tersebut, peneliti mengumpulkan data guna menunjukkan bukti benarnya. Untuk menunjukkan bukti benarnya, peneliti membutuhkan alat bantu untuk mengolah dan menganalisis data itu, alat bantu dimaksud disebut statistik (*statistic*). Bedakan antara *statistics* – yang berarti ilmu statistik (statistika), yaitu ilmu tentang pengolahan dan penafsiran data kuantitatif – dengan statistik (*statistic*) yang berarti “ukuran yang diperoleh atau berasal dari sampel, sebagai lawan dari istilah parameter yang berarti ukuran yang diperoleh atau berasal dari populasi”. Jadi, dapat kiranya ditegaskan bahwa “statistik” adalah ukuran apa pun juga yang berkaitan dengan data

sampel, sementara “parameter” merupakan ukuran apa pun juga yang berkenaan dengan data populasi.

Dalam konteks statistik, “hipotesis adalah pernyataan mengenai keadaan data yang ada pada populasi (parameter) yang akan diuji sejauh mana kebenarannya berdasarkan data yang diperoleh dari sampel (statistik)”. Atas dasar inilah sehingga statistik yang diuji adalah hipotesis nihil atau hipotesis nol (*null hypothesis*), hipotesis yang meniadakan (*nullify*) hubungan antar kelompok, atau meniadakan perbedaan antara “parameter” dan “statistik” yang dibandingkan. Jadi yang dimaksud dengan data statistik adalah data sampel yang digunakan untuk mengestimasi nilai sebenarnya pada populasi. Hipotesis statistik dibutuhkan, untuk menguji apakah hipotesis penelitian yang diuji dari data sampel itu dapat digeneralisasikan untuk populasinya atau tidak. Sebagai konsekuensinya – sejauh dilakukan dengan benar – maka kesimpulan yang diperoleh dari analisis data sampel itu sah diberlakukan untuk populasinya.

Umpamanya peneliti membandingkan keunggulan metode A dengan metode B, lalu melandaskan dengan asumsi bahwa kedua metode tersebut “sama-sama baik”, maka asumsi ini disebut hipotesis nol (H_0), karena “meniadakan” perbedaan kedua metode tersebut. Berlawanan dengan ini, peneliti mungkin memprediksi bahwa metode A lebih superior atau metode B lebih inferior, dalam hal ini peneliti menyatakan apa yang disebut sebagai hipotesis alternatif (*alternative hypothesis*) (H_a atau kadangkala ditulis

dengan notasi H_1), karena secara implisit menyatakan “ada” perbedaan kedua metode tersebut. H_a adalah pernyataan “prediksi” yang dirumuskan peneliti secara spesifik, yang menyatakan apa yang diharapkan akan terjadi dalam penelitian. Misalnya, H_0 menyatakan “tidak ada pengaruh *self-esteem* terhadap perilaku prososial”, maka H_a menyatakan yang sebaliknya, “ada pengaruh *self-esteem* terhadap perilaku prososial.” Jadi, H_a merupakan hipotesis yang didukung peneliti, hipotesis yang mengandung pernyataan prediksi sang peneliti. Dapat kiranya ditegaskan bahwa H_0 menggunakan praduga “tidak ada” (tidak ada perbedaan, tidak ada hubungan), sebaliknya H_a menggunakan praduga “ada” (ada perbedaan, ada hubungan). H_0 dinyatakan dalam kalimat negatif, sebaliknya H_a dinyatakan dalam kalimat positif.

Dalam analisis data penelitian, statistik digunakan tidak untuk langsung menguji H_a yang berisi prediksi peneliti, “akan tetapi digunakan untuk menolak atau menerima hipotesis nol (H_0). Penerimaan atau penolakan H_a merupakan konsekuensi dari penolakan atau penerimaan H_0 , sebab H_0 yang diuji”. Dengan demikian jelaslah bahwa H_0 adalah hipotesis yang menjadi “patokan” diterima atau ditolaknya hipotesis (jika H_0 diterima maka H_a ditolak, jika H_0 ditolak maka H_a diterima). Dengan kata lain, teknik komputasi statistika selalu menguji H_0 , sehingga H_0 disebut juga hipotesis statistika. Teknik statistik yang digunakan adalah statistik inferensial atau statistik induktif, kadangkala disebut juga statistik probabilitas, yaitu teknik statistik yang digunakan

untuk menganalisis data sampel yang hasilnya digeneralisasikan untuk populasi. “Untuk dapat mengambil kesimpulan apakah akan menerima atau menolak suatu hipotesis berdasarkan penyelidikan sampel, harus ada “pedoman” yaitu dengan membandingkan antara harga statistik (yang dihitung dari data sampel) dengan harga parameter yang dihipotesiskan”. Kalau perbedaan kedua harga itu cukup kecil, berarti H_0 diterima. Akan tetapi, jika perbedaan kedua harga itu cukup besar, maka H_0 ditolak. Kata “cukup” di sini bermakna kualitatif, ukuran kuantitatifnya merujuk pada tingkat signifikansi sebagaimana bahasan berikut ini.

2. Tingkat Signifikansi

Perlu dipahami bahwa istilah “signifikan” dalam pengertian statistik, berbeda maknanya dengan pengertian sehari-hari. Dalam pengertian sehari-hari, signifikan acapkali disepadankan dengan kata “bermakna (*meaningful*) atau penting (*important*)”, sedangkan dalam pengertian statistik, signifikan berarti kemampuan digeneralisasikan dengan tingkat kesalahan tertentu. Apa yang digeneralisasikan di sini adalah data sampel yang digeneralisasikan untuk populasinya. Umpama “ada pengaruh signifikan, artinya pengaruh tersebut dapat digeneralisasikan dengan tingkat kesalahan tertentu”, “ada hubungan signifikan, artinya hubungan tersebut dapat digeneralisasikan dengan tingkat kesalahan tertentu”, “ada perbedaan signifikan, artinya perbedaan tersebut dapat digeneralisasikan dengan tingkat kesalahan tertentu.” Dengan kata lain, istilah signifikan artinya hipotesis penelitian yang telah terbukti

pada suatu sampel (baik deskriptif, komparatif, maupun asosiatif) dapat diberlakukan atau digeneralisasikan untuk populasinya. Istilah signifikan juga menunjukkan makna perbedaan atau hubungan yang diuji, yang terjadi bukan karena *random error* atau karena kebetulan saja.

Ingatlah bahwa semua kesimpulan ilmiah (*scientific conclusions*) harus merupakan pernyataan yang mengandung kebenaran dengan probabilitas tinggi. Oleh sebab itu ‘tingkat signifikansi’ menjadi konsep yang sangat penting dalam pengujian hipotesis. Dalam praktiknya, tingkat signifikansi atau “peluang terjadinya kesalahan” ditetapkan peneliti sebelum hipotesis diuji. Dengan kata lain, peneliti menetapkan terlebih dahulu berapa besar risiko kesalahan penolakan H_0 yang hendak ditanggung. Suatu kesimpulan yang ditarik dari data sampel yang akan digeneralisasikan untuk populasinya, selalu mengandung peluang (*probability* = p) terjadinya “salah” dan “benar”. Ukuran “kesalahan” dan “kebenaran yang dipercaya” dinyatakan dalam ukuran persentase (%), bukan prosentase. Hipotesis statistik merupakan anggapan atau pernyataan yang mengandung salah dan benar. Kesimpulan penelitian yang disandarkan pada keputusan statistik, tidak dapat ditopang “benar” seratus persen. Peluang “kesalahan” dan “kepercayaan” inilah yang disebut tingkat signifikansi (*level of significance*). Apabila peluang kesalahan sebesar 1% maka taraf kepercayaannya sebesar 99%. Apabila peluang kesalahan sebesar 5% maka tentu taraf kepercayaannya sebesar 95%. Pada

umumnya, para peneliti ilmu sosial dan humaniora menetapkan “*rules of thumb*” peluang terjadinya kesalahan sebesar 5%. Jadi, sewaktu peneliti menyatakan penolakan terhadap H_0 , harus dipahami bahwa penolakan itu mengandung risiko kesalahan sebesar taraf signifikansi yang ditetapkan peneliti. Apabila peneliti menetapkan tingkat signifikansi sebesar 5%, maka hal ini menyiratkan bahwa peneliti bersedia mengambil risiko kesalahan sebesar 5% untuk menolak H_0 , ketika H_0 seharusnya diterima, dengan probabilitas sama dengan atau kurang dari 5 kali dalam 100. Jika peneliti menetapkan tingkat signifikansi pada 0,01, itu berarti H_0 akan ditolak dengan probabilitas sama dengan atau kurang dari 1 kali dalam 100. Peneliti mengatakan bahwa dia bersedia membatasi kemungkinan menolak H_0 ketika H_0 itu benar (kesalahan Tipe I) menjadi 1 dalam 100. Umpama tingkat signifikansi ditetapkan pada 0,0001, itu artinya probabilitas membuat kesalahan Tipe I adalah 1 dari 10.000 atau kurang.

Dengan demikian, tingkat signifikansi adalah nilai maksimum probabilitas menolak H_0 yang seharusnya diterima. Untuk mendeteksi signifikan atau tidak, peneliti dapat menempuh dua cara. Pertama, pendekatan klasik, dengan “mengacu pada tabel” sesuai teknik analisis yang digunakan. Pada setiap tabel sudah disediakan untuk taraf signifikansi berapa persen suatu hasil analisis dapat digeneralisasikan. Ambil contoh uji t, dengan ketentuan, apabila nilai t-hitung > t-tabel dengan taraf signifikansi (α) tertentu, misalnya 5%, maka artinya H_0 ditolak, H_a diterima. Akan tetapi,

jika yang terjadi sebaliknya (nilai t -hitung $\leq t$ -tabel), maka artinya H_0 diterima, H_a ditolak. Jadi, pendekatan klasik ini dengan membandingkan nilai statistik hitung dengan nilai pada tabelnya. Kedua, pendekatan probabilistik, yang memanfaatkan program komputer, menampilkan nilai probabilitas (*probability value*). Patokan peneliti adalah nilai signifikansi atau disebut juga *probability values (p-values)* ini. Nilai p adalah nilai yang menunjukkan besarnya “peluang salah” menolak H_0 berdasarkan data penelitian. Dengan pendekatan probabilistik ini, peneliti membandingkan nilai p dengan nilai α (α). Apabila nilai $p \leq \alpha$ maka keputusan H_0 ditolak, H_a diterima. Sebaliknya, apabila nilai $p > \alpha$ maka keputusannya adalah H_0 diterima, H_a ditolak, karena berarti tidak signifikan secara statistik. Nilai- p adalah tingkat di mana nilai yang diberikan dari statistik uji (seperti uji t misalnya) berada pada batas antara daerah penerimaan dan penolakan. Dengan demikian, peneliti tentu berharap *p-value* yang kecil ($p < \alpha$), sebab apabila nilai p -nya kecil, artinya adanya perbedaan pada hasil penelitian itu menunjukkan adanya perbedaan pada populasinya, dan perbedaan itu terjadi bukan karena faktor kebetulan (*by chance*).

3. Dua Kesalahan dalam Menguji Hipotesis

Ketika peneliti mengestimasi data populasi berdasarkan data sampel, sangat mungkin terjadi kesalahan. Hal ini terjadi, sebab sampel yang menjadi basis pengujian hipotesis hanyalah “elemen” dari populasi, yang proses pencuplikannya tidak dapat

menggambarkan secara “tepat” karakteristik populasi. Oleh karena teknik sampling manapun tidak dapat menggambarkan secara tepat, sehingga muncul kesalahan yang disebut kesalahan sampel (*sampling error*):

a. Kesalahan Tipe 1 (Type 1 error)

Kesalahan Tipe I adalah kesalahan yang terjadi ketika peneliti memutuskan menolak H_0 yang benar, yang seharusnya diterima. Jika H_0 yang benar tersebut ditolak, maka penolakan H_0 yang benar inilah yang disebut kesalahan Tipe I. Kesalahan Tipe I ini seperti “alarm palsu”, peneliti mengira dirinya memiliki “sesuatu” padahal tidak ada apa-apa di sana. Peneliti menyimpulkan “ada perbedaan” padahal sesungguhnya tidak ada perbedaan, atau menyimpulkan “ada hubungan” padahal sesungguhnya tidak ada hubungan. Besarnya peluang melakukan kesalahan Tipe I disebut “taraf signifikansi” (*level of significance*), yang diberi simbol p atau simbol α yang dinyatakan dalam proporsi atau persentase. Probabilitas pengujian melakukan kesalahan tipe I disebut juga kesalahan alpha (α error), disebut juga *size of test*. Peluang untuk tidak membuat kesalahan Tipe I adalah sebesar $1 - \alpha$. Misalnya, $\alpha = 0,05$ maka $1 - 0,05 = 0,95$ atau 95%. Angka sebesar 95% inilah yang disebut taraf kepercayaan (*level of confidence*).

b. Kesalahan Tipe II (Type II error)

Kesalahan Tipe II adalah kesalahan yang terjadi ketika peneliti memutuskan menerima H_0 yang salah, yang seharusnya ditolak. Jika H_0 yang salah tersebut diterima, maka penerimaan H_0 yang

salah itu disebut kesalahan Tipe II, disebut juga kesalahan beta (β error). Jadi kesalahan Tipe II ini semacam ada “miss” (salah menangkap), dimana peneliti menyimpulkan tidak ada apa-apa padahal ada sesuatu, menyimpulkan “tidak ada perbedaan” padahal sesungguhnya ada perbedaan, atau menyimpulkan “tidak ada hubungan” padahal sesungguhnya ada hubungan. Besarnya peluang membuat kesalahan Tipe II yang diberi simbol huruf Yunani β ini juga dinyatakan dalam bentuk proporsi dan persentase, sedangkan harga $(1 - \beta)100\%$ dikenal sebagai tingkat kekuatan uji (*power of the test*).

4. Bentuk Pengujian Hipotesis

Ditinjau dari arah distribusi datanya, dikenal (1) hipotesis tidak terarah (*nondirectional hypothesis*), dan (2) hipotesis terarah (*directional hypothesis*) yang mencakup uji pihak kiri, dan uji pihak kanan. Jenis uji mana yang akan digunakan tergantung pada bunyi kalimat hipotesis yang “diturunkan” peneliti dari kerangka konseptual atau kerangka teoretis penelitiannya.

a. Hipotesis tidak terarah (*nondirectional hypothesis*) disebut juga

Uji Dua Pihak, Uji Dua Ekor, atau Uji Dua Sisi (Two Tail Test)

Hipotesis nondireksional ini berisi pernyataan tentang “adanya perbedaan atau adanya hubungan. Apabila mengenai perbedaan, maka hipotesis dua arah ini hanya menyatakan bahwa kelompok A berbeda dari kelompok B, tanpa menyatakan secara eksplisit kelompok mana yang lebih dari yang lain”. Apabila mengenai hubungan, maka hipotesis dua

arah ini hanya menyatakan bahwa variabel X berhubungan atau berkorelasi dengan variabel Y, tanpa menyatakan secara eksplisit apakah hubungan (korelasi) itu negatif atau positif. Terkait dengan prosedur pengujiannya, disebut dengan bermacam nama, di sini disebut uji dua pihak. Uji dua pihak “digambarkan sebagai uji hipotesis, di mana wilayah penolakan (area kritis) berada di kedua ujung distribusi normal. Ini menentukan apakah sampel yang diuji berada di dalam ataukah di luar kisaran nilai tertentu. H_a akan diterima sebagai pengganti H_0 , jika nilai yang dihitung jatuh pada salah satu dari dua pihak dari distribusi probabilitas”. Uji dua pihak ini digunakan apabila H_0 berbunyi “sama dengan” ($H_0 =$), dan H_a berbunyi “tidak sama dengan” ($H_a \neq$).

- b. Hipotesis terarah (directional hypothesis) disebut juga Uji Satu Pihak, Uji Satu Ekor, atau Uji Satu Sisi (One Tail Test)

Hipotesis terarah adalah rumusan hipotesis yang sudah “jelas arahnya”, sudah jelas hasil yang diharapkan mengarah pada ke salah satu sudut (ke kiri atau ke kanan) pada distribusi datanya. Hipotesis ini menyatakan dengan tegas apakah hubungan itu positif atau negatif. Hubungan “positif” antara dua variabel maksudnya “kenaikan skor/nilai pada salah satu variabel diikuti kenaikan skor/nilai pada variabel lainnya, atau penurunan skor/nilai pada salah satu variabel diikuti penurunan skor/nilai pada variabel lainnya.” Sebaliknya, pada hubungan antar variabel yang “negatif”, kenaikan skor/nilai pada satu variabel, justru diikuti penurunan skor/nilai pada variabel lainnya. Oleh

karena langsung “mengarah” sehingga disebut juga “hipotesis langsung”. Umpama menyatakan perbedaan dua kelompok sampel tentang suatu variabel, maka perbedaan itu juga menegaskan kelompok mana yang memiliki skor/nilai tentang suatu variabel yang lebih tinggi atau rendah dari skor/nilai kelompok lainnya. Uji satu pihak digunakan ketika peneliti akan menguji, misalnya apakah rata-rata populasi lebih rendah atau lebih tinggi dari beberapa nilai yang dihipotesiskan.

1) Uji Pihak Kiri (*Left Tail Test*)

Oleh karena “daerah penolakannya” berada di ujung kiri kurva distribusi normal, maka uji hipotesisnya disebut uji pihak kiri. Uji pihak kiri digunakan apabila rumusan H_0 dinyatakan “lebih besar atau sama dengan” (\geq), dan H_a berbunyi “lebih kecil” ($<$).

2) Uji Pihak Kanan (*Right Tail Test*)

Oleh karena “daerah penolakannya” berada di ujung kanan kurva distribusi normal, maka uji hipotesisnya disebut uji pihak kanan. Uji pihak kanan digunakan apabila rumusan H_0 dinyatakan “lebih kecil atau sama dengan” (\leq), dan H_a berbunyi “lebih besar” ($>$).

Berdasarkan uraian di atas, dapat disimpulkan bahwa langkah umum dalam pengujian hipotesis dilakukan dengan pola kerja berurutan, yaitu:

- a. Menentukan formulasi hipotesis (baik H_0 maupun H_a).
Dirumuskan dalam bentuk kalimat pernyataan (hipotesis

- penelitian), dan dalam bentuk simbol statistik (hipotesis statistik). Pada saat yang sama, dari rumusan hipotesisnya, peneliti juga menentukan alternatif pengujian, apakah tidak terarah (*two tail test*) atau terarah (*one tail test*).
- b. Menentukan jenis uji statistik yang sesuai. Bagaimanapun juga, statistik hanyalah alat bantu, penelitilah yang memutuskan penggunaan alat bantu tersebut. Ada banyak ragam jenis uji statistik yang dapat digunakan (misalnya, uji korelasi, regresi, uji-F, uji-t, uji-z, anova, *chi-square*, dan sebagainya), masing-masing mempunyai persyaratan tertentu yang harus dipenuhi.
 - c. Menentukan batas atau tingkat signifikansi (*level of significance*). Dalam hal ini, peneliti menetapkan batas toleransi dalam menerima kesalahan hasil pengujian, atau secara singkat disebut juga “tingkat kesalahan”. Ilmu sosial dan humaniora, *rules of thumb* biasanya dengan nilai α sebesar 5% sedangkan bidang Kesehatan, umumnya menggunakan nilai α yang lebih kecil, biasanya sebesar 1%. Makin besar tingkat kesalahan yang ditolerir, tentu makin kecil tingkat kepercayaan (*confidence interval*) hasil pengujian, atau makin besar α , makin kecil β . Perlu diingat bahwa tingkat signifikansi ditetapkan sebelum pengumpulan data.
 - d. Menentukan kriteria pengujian. Artinya menentukan daerah terima/tolak. Di sini, peneliti dapat menggunakan pendekatan klasik (membandingkan “hasil” hitung uji statistik dengan tabelnya), dapat pula dengan pendekatan probabilistik, yaitu membandingkan nilai signifikansi atau probabilitas hitung

dengan nilai α yang telah ditetapkan peneliti. Jika nilai signifikansi (Sig.) $< \alpha$ (0,05) maka H_0 ditolak. Sebaliknya, jika nilai signifikansi (Sig.) $> \alpha$ (0,05) maka artinya H_0 diterima. Pada langkah ini, penting untuk diingatkan kepada peneliti pemula untuk lebih berhati-hati “membunyikan” hasil analisis statistiknya. Oleh karena sudah ditetapkan besar α -nya di awal (sebelum ambil data) maka konsekuensinya, semua hasil komputasi yang menghasilkan statistik lebih kecil dari nilai α yang ditetapkan diberi label “signifikan”, tidak peduli berapa pun kecilnya, bukan “sangat” signifikan. Sebaliknya, semua hasil komputasi yang menghasilkan statistik lebih besar dari nilai α yang ditetapkan diberi label “tidak signifikan”, tidak peduli berapa pun besarnya, bukan “sangat” tidak signifikan.

- e. Menuliskan kesimpulan berdasarkan keputusan hasil uji. Dalam arti, apakah menerima atau menolak hipotesis yang diajukan sesuai dengan kriteria pengujiannya.

BAGIAN 10

PENGANTAR BASIS DATA

A. ANALISIS KORELASI

1. Pengertian Analisis Korelasi

Analisis korelasi adalah teknik statistik untuk mengetahui arah hubungan, kuat hubungan, dan signifikansi kuatnya hubungan antara dua variabel atau lebih. Dua variabel dikatakan berkorelasi apabila perubahan pada salah satu variabel disertai dengan perubahan pada variabel lainnya secara linier, baik dalam arah yang sama ataupun arah yang sebaliknya. Kuatnya hubungan dan arah hubungan linier antara dua variabel dinyatakan dengan ukuran statistik yang dinamakan koefisien korelasi. Ketika fluktuasi suatu variabel secara andal menyebabkan fluktuasi serupa pada variabel lain, sering kali terdapat kecenderungan untuk berpikir bahwa perubahan pada satu variabel menyebabkan perubahan pada variabel lainnya.

Kelebihan analisis korelasi adalah sebagai berikut.

- a. Analisis korelasi membantu dalam memperoleh secara tepat derajat dan arah hubungan linier antara dua variabel.
- b. Efek korelasi adalah mengurangi rentang ketidakpastian prediksi. Prediksi berdasarkan analisis korelasi akan lebih dapat diandalkan dan mendekati kenyataan.

- c. Analisis korelasi membantu dalam menemukan variabel-variabel penting yang menjadi sandaran variabel-variabel lain, dapat mengungkapkan hubungan-hubungan yang menyebabkan gangguan-gangguan menyebar sehingga bisa dicari upaya-upaya stabilisasi.

Kelemahan analisis korelasi adalah sebagai berikut.

- a. Selalu mengasumsikan hubungan linier.
- b. Nilai koefisien korelasi dipengaruhi oleh nilai ekstrim data.
- c. Tidak dapat digunakan untuk mengetahui korelasi dalam distribusi frekuensi yang berkelompok.
- d. Tidak dapat menunjukkan sebab dan akibat. Hanya karena dua variabel berkorelasi bukan berarti variabel yang satu menyebabkan variabel yang lain.

2. Jenis-Jenis Korelasi

Korelasi dijelaskan atau diklasifikasikan dalam beberapa cara berbeda. Tiga yang paling penting adalah:

- a. Korelasi positif, negatif dan nol.

Korelasinya positif (langsung) atau negatif (sebaliknya) tergantung pada arah perubahan variabelnya. Korelasi dikatakan positif jika kedua variabel berfluktuasi dalam arah yang sama. Artinya, jika suatu variabel meningkat maka rata-rata variabel lainnya juga meningkat, atau jika salah satu variabel menurun rata-rata variabel lainnya juga menurun. Misalnya hubungan antara tinggi badan dan berat badan

sekelompok orang serta hubungan antara pendapatan dan pengeluaran.

Korelasi dikatakan negatif jika kedua variabel bervariasi dalam arah yang berlawanan. Artinya jika satu variabel meningkat, maka variabel lainnya menurun, atau jika salah satu variabel menurun maka variabel lainnya meningkat. Misalnya hubungan antara harga suatu produk dan permintaannya serta hubungan antara jam kerja dan waktu luang.

Korelasi nol sebenarnya bukan jenis korelasi tetapi tetap disebut korelasi nol atau tidak ada korelasi. Jika tidak ditemukan hubungan apa pun antar variabel, maka dikatakan korelasi nol. Artinya perubahan nilai suatu variabel tidak mempengaruhi atau mengubah nilai variabel lainnya. Misalnya, hubungan antara berat badan seseorang dan kecerdasan adalah korelasi nol atau tidak ada sama sekali.

b. Korelasi sederhana, parsial dan berganda.

Perbedaan antara korelasi sederhana, parsial dan berganda didasarkan pada jumlah variabel yang diteliti. Korelasi dikatakan sederhana jika hanya dua variabel yang diteliti. Misalnya, ketika seseorang mempelajari hubungan antara nilai yang diperoleh siswa dan kehadiran siswa di kelas.

Korelasi parsial adalah ketika seseorang mempelajari tiga variabel atau lebih tetapi menganggap hanya dua variabel yang saling mempengaruhi dan pengaruh variabel lain yang

mempengaruhi dianggap konstan. Misalnya, dalam contoh hubungan antara nilai siswa dan kehadiran di kelas, variabel lain yang mempengaruhi seperti efektivitas pengajaran guru, penggunaan alat bantu mengajar dianggap konstan. Korelasi berganda adalah ketika tiga atau lebih variabel diteliti. Misalnya, hubungan antara nilai siswa, kehadiran siswa, efektivitas guru, dan penggunaan alat bantu.

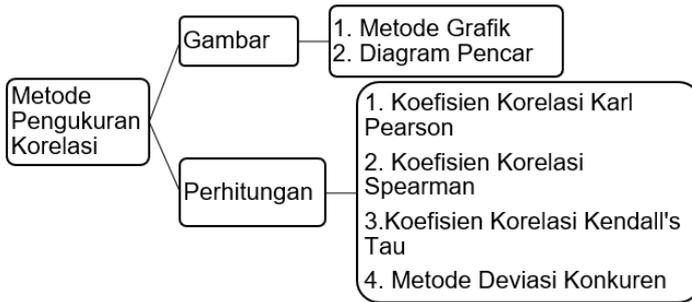
c. Korelasi linier dan nonlinier.

Korelasi linier atau nonlinier bergantung pada seberapa konstan rasio perubahan antar variabel. Korelasi dikatakan linier jika besarnya perubahan suatu variabel memiliki perbandingan yang konstan terhadap besarnya perubahan variabel lain. Jika variabel-variabel tersebut diplot pada grafik, semua titik yang diplot akan berada pada satu garis lurus. Misalnya, jika diasumsikan untuk memproduksi satu unit barang jadi diperlukan 10 unit bahan baku, maka untuk memproduksi 2 unit barang jadi diperlukan dua kali lipat dari satu unit tersebut.

Korelasi dikatakan nonlinier jika besarnya perubahan suatu variabel tidak mempunyai perbandingan yang konstan terhadap besarnya perubahan variabel lain. Jika variabel-variabel tersebut diplot pada grafik, titik-titiknya akan berada pada kurva dan bukan pada garis lurus. Misalnya, jika kita melipatgandakan jumlah penayangan iklan, maka volume penjualan belum tentu menjadi dua kali lipat.

3. Metode Pengukuran Korelasi

Terdapat berbagai metode pengukuran korelasi, yang dapat direpresentasikan seperti Gambar 10.1.



Gambar 10.1. Metode pengukuran korelasi

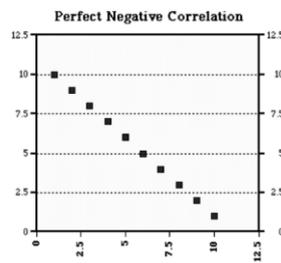
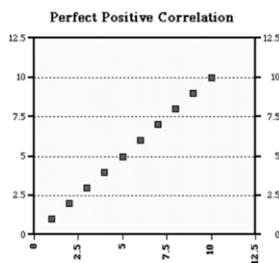
Di antara cara-cara tersebut yang akan kita bahas hanyalah diagram pencar dan koefisien korelasi Karl Pearson.

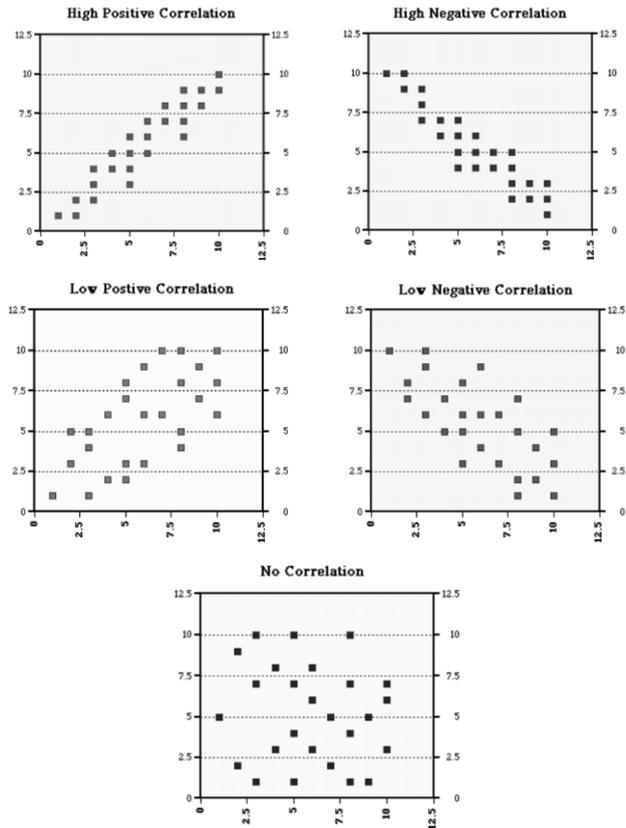
a. Diagram pencar

Diagram pencar adalah metode visual pengukuran korelasi. Ini adalah representasi diagram data bivariat untuk memastikan hubungan antara dua variabel. Dengan metode ini, data yang diberikan diplot pada grafik, yaitu untuk setiap pasangan nilai X dan Y diberi titik sehingga diperoleh poin sebanyak observasi. Biasanya variabel bebas ditunjukkan pada sumbu X sedangkan variabel terikat ditunjukkan pada sumbu Y. Setelah nilai diplot pada grafik, maka jenis korelasi antara variabel X dan Y akan terungkap.

- **Korelasi Positif Sempurna:** Dalam hal ini, titik-titik akan terbentuk pada garis lurus yang turun dari sudut kiri bawah ke sudut kanan atas.

- Korelasi Negatif Sempurna: Dalam hal ini, titik-titik akan terbentuk pada garis lurus yang naik dari pojok kiri atas ke pojok kanan bawah.
- Korelasi Positif Tinggi: Dalam hal ini, titik-titik yang diplot berada dalam lajur sempit, di mana titik-titik menunjukkan kecenderungan naik dari sudut kiri bawah ke sudut kanan atas.
- Korelasi Negatif Tinggi: Dalam hal ini, titik-titik yang diplot berada dalam lajur sempit, di mana titik-titik tersebut menunjukkan kecenderungan turun dari sudut kiri atas ke sudut kanan bawah.
- Korelasi Positif Rendah: Jika titik-titik tersebar luas pada diagram, titik-titiknya naik dari sudut kiri ke sudut kanan atas.
- Korelasi Negatif Rendah: Jika titik-titik tersebar luas pada diagram, titik-titiknya turun dari sudut kiri atas ke sudut kanan bawah.
- Korelasi Nol (Tidak Ada): Jika titik-titik yang diplot tersebar secara sembarangan pada grafik.





Gambar 10.2. Klasifikasi korelasi dalam diagram pencar

Contoh 1:

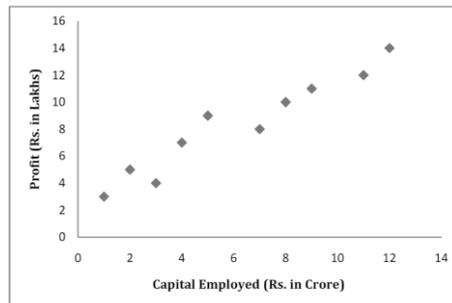
Diberikan pasangan nilai berikut:

Modal (puluhan juta rupiah)	1	2	3	4	5	7	8	9	11	12
Laba (puluhan juta rupiah)	3	5	4	7	9	8	10	11	12	14

- Gambarlah diagram pencar.
- Apakah ada korelasi antara laba dan modal yang digunakan?
Apakah itu positif atau negatif? Apakah tinggi atau rendah?

Solusi:

- a) Diagram pencar dari permasalahan pada Contoh 1 ditunjukkan pada Gambar 10.3.



Gambar 10.3. Diagram pencar dari Contoh 1

- b) Dari pengamatan diagram pencar dapat dikatakan bahwa variabel-variabel tersebut berkorelasi positif. Pada diagram titik-titiknya cenderung naik dari pojok kiri bawah ke pojok kanan atas. Titik-titik yang diplot berada dalam lajur sempit yang menunjukkan bahwa kasus tersebut mempunyai korelasi positif tinggi.

b. Koefisien korelasi Karl Pearson

Koefisien korelasi digunakan untuk mengukur kuatnya hubungan antara dua variabel. Pada populasi, koefisien korelasi atau kuatnya hubungan antara variabel x dan y , ditulis ρ (ρ dibaca rho). Nilai ρ berkisar antara -1 hingga +1. Pada sampel, koefisien korelasi atau kuatnya hubungan antara variabel x dan y ditulis r . Nilai r berkisar antara -1 hingga +1. Apabila r bernilai positif maka hubungan antar variabel x dan y bersifat searah. Apabila r bernilai negatif maka hubungan antar variabel x dan y bersifat berlawanan arah. Misalnya korelasi $r = 0,8$ menunjukkan hubungan positif dan kuat

antara dua variabel, sedangkan korelasi $r = -0,3$ menunjukkan hubungan negatif dan lemah. Korelasi yang mendekati nol menunjukkan tidak adanya hubungan linier antara dua variabel kontinu.

Metode perhitungan koefisien korelasi Karl Pearson didasarkan pada kovarians dua variabel. Jika kedua variabel yang diteliti adalah x dan y , maka rumus yang dikemukakan oleh Karl Pearson berikut ini dapat digunakan untuk mengukur kuat hubungan korelasinya.

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2} \cdot \sqrt{\sum(y-\bar{y})^2}} \text{ atau}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

Rumus yang berbeda di atas dapat digunakan dalam situasi yang berbeda tergantung pada informasi yang diberikan dalam soal.

Contoh 2:

Carilah koefisien korelasi antara umur dengan kebiasaan bermain siswa berikut dengan menggunakan metode koefisien korelasi Karl Pearson.

Umur	15	16	17	18	19	20
Jumlah siswa	250	200	150	120	100	80
Jumlah siswa yang biasa bermain	200	150	90	48	30	12

Solusi:

Untuk mengetahui hubungan antara umur dengan kebiasaan bermain siswa, perlu dihitung persentase siswa yang mempunyai kebiasaan bermain.

$$\text{Persentase kebiasaan bermain} = \frac{\text{Jumlah siswa yang biasa bermain}}{\text{Jumlah Siswa}} * 100$$

100

Diasumsikan umur siswa adalah variabel x dan persentase kebiasaan bermain adalah variabel y .

Tabel 10.1. Perhitungan koefisien korelasi Karl Pearson untuk Contoh 2

x	y	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
15	80	-2,5	6,25	30	900	-75
16	75	-1,5	2,25	25	625	-37,5
17	60	-0,5	0,25	10	100	-5
18	40	0,5	0,25	-10	100	-5
19	30	1,5	2,25	-20	400	-30
20	15	2,5	6,25	-35	1225	-87,5
$\Sigma = 105$	$\Sigma = 300$		$\Sigma = 17,5$		$\Sigma = 3350$	$\Sigma = -240$
$\bar{x} = \frac{\Sigma x}{n} = \frac{105}{6} = 17,5$				$\bar{y} = \frac{\Sigma y}{n} = \frac{300}{6} = 50$		
$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2} \cdot \sqrt{\Sigma(y-\bar{y})^2}} = \frac{-240}{\sqrt{17,5 \cdot 3350}} = -0,9912$						

Interpretasi: Dari perhitungan di atas terlihat jelas adanya korelasi negatif yang tinggi yaitu $r = -0,9912$. Artinya, kebiasaan bermain di kalangan siswa semakin berkurang seiring bertambahnya usia.

Contoh 3:

Tentukan koefisien korelasi Karl Pearson antara modal yang digunakan dan laba yang diperoleh dari data berikut.

Modal (dalam juta rupiah)	10	20	30	40	50	60	70	80	90	100
Laba (dalam juta rupiah)	2	4	8	5	10	15	14	20	22	50

Solusi:

Misalkan modal adalah variabel x dan laba adalah variabel y .

Tabel 10.2. Perhitungan koefisien korelasi Karl Pearson Contoh 3

x	y	x^2	y^2	xy
10	2	100	4	20
20	4	400	16	80
30	8	900	64	240
40	5	1600	25	200
50	10	2500	100	500
60	15	3600	225	900
70	14	4900	196	980
80	20	6400	400	1600
90	22	8100	484	1980
100	50	10000	2500	5000
$\Sigma = 550$	$\Sigma = 150$	$\Sigma = 38500$	$\Sigma = 4014$	$\Sigma = 11500$
$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n(\Sigma x^2) - (\Sigma x)^2][n(\Sigma y^2) - (\Sigma y)^2]}} =$ $\frac{(10 \cdot 11500) - (550 \cdot 150)}{\sqrt{[(10 \cdot 38500) - (550)^2][(10 \cdot 4014) - (150)^2]}} = 0,8519$				

Interpretasi: Dari perhitungan di atas terlihat jelas bahwa terdapat korelasi positif yang tinggi yaitu $r = 0,8519$. Artinya, peningkatan modal menyebabkan peningkatan laba yang diperoleh.

B. ANALISIS REGRESI

1. Pengertian Analisis Regresi

Analisis regresi adalah alat statistik untuk mempelajari sifat dan tingkat hubungan fungsional antara dua variabel atau lebih dan untuk memperkirakan (atau memprediksi) nilai variabel terikat yang tidak diketahui dari nilai variabel bebas yang diketahui. Variabel yang menjadi dasar prediksi variabel lain disebut variabel terikat dan variabel yang diprediksi disebut variabel bebas. Regresi akan selalu menghasilkan hubungan, terlepas dari apakah variabel tersebut benar-benar terkait secara kausal atau tidak, penting juga untuk memilih variabel bebas dengan cermat.

Misalnya, jika kita mengetahui bahwa dua variabel harga (X) dan permintaan (Y) berkaitan erat, kita dapat mengetahui nilai X yang paling mungkin untuk nilai Y tertentu atau nilai Y yang paling mungkin untuk nilai X tertentu. Demikian pula, jika kita mengetahui bahwa besaran pajak dan kenaikan harga suatu barang mempunyai hubungan yang erat, maka kita dapat mengetahui harga yang diharapkan untuk sejumlah pungutan pajak tertentu.

Kegunaan analisis regresi antara lain sebagai berikut.

- a. Memberikan perkiraan nilai variabel terikat dari nilai variabel bebas.
- b. Digunakan untuk mengukur bagaimana perubahan suatu variabel akan mempengaruhi variabel lainnya
- c. Digunakan untuk memperoleh ukuran kesalahan dalam menggunakan garis regresi sebagai dasar estimasi.
- d. Dengan bantuan analisis regresi, kita dapat memperoleh ukuran korelasi yang terjalin antara kedua variabel.

2. Jenis-Jenis Regresi

Analisis regresi mencakup beberapa jenis, seperti linier sederhana, linier berganda, dan nonlinier.

a) Regresi linier sederhana

Regresi linier sederhana menggunakan satu variabel bebas untuk menjelaskan atau memprediksikan hasil variabel terikat. Misalnya, jumlah karyawan berpengaruh terhadap lamanya suatu pekerjaan dapat diselesaikan. Model regresi linier sederhana dinyatakan dengan menggunakan persamaan berikut:

$$Y = a + bX + \varepsilon$$

dengan:

Y = variabel yang ingin diprediksi (variabel terikat).

X = variabel yang digunakan untuk memprediksi Y (variabel bebas).

a = intersep.

b = kemiringan (slope).

ε = galat (kesalahan)

b) Regresi linier berganda

Regresi linier berganda menggunakan beberapa variabel bebas untuk menjelaskan atau memprediksikan hasil variabel terikat. Misalnya, luas lahan dan curah hujan berpengaruh terhadap hasil panen padi. Model regresi linier berganda dinyatakan dengan menggunakan persamaan berikut:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n + \varepsilon$$

dengan:

Y = variabel yang ingin diprediksi (variabel terikat).

$X_k, k = 1, 2, \dots, n$ = variabel yang digunakan untuk memprediksi Y (variabel bebas).

a = intersep.

$b_k, k = 1, 2, \dots, n$ = koefisien regresi untuk $X_k, k = 1, 2, \dots, n$

ε = galat (kesalahan)

c) Regresi Nonlinier

Regresi nonlinier adalah suatu bentuk analisis regresi di mana data disesuaikan dengan suatu model dan kemudian dinyatakan sebagai fungsi matematika nonlinier (logistik, polinomial, dan sebagainya). Salah satu contoh penggunaan regresi nonlinier adalah memprediksi pertumbuhan populasi dari waktu ke waktu. Model pertumbuhan populasi logistik dapat memberikan perkiraan populasi untuk periode yang tidak diukur, dan prediksi pertumbuhan populasi di masa depan. Variabel independen dan dependen yang digunakan dalam regresi nonlinier harus bersifat kuantitatif.

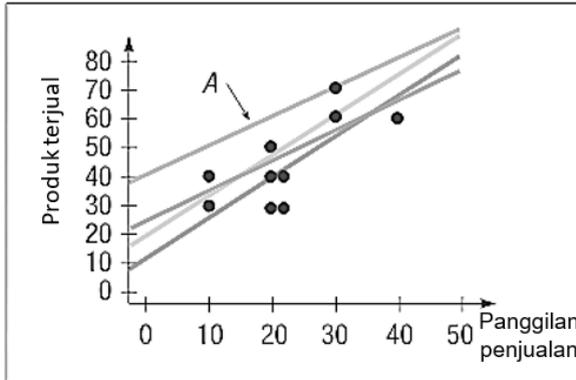
3. Garis Regresi Linear

Setelah mengidentifikasi dua variabel yang berkorelasi, selanjutnya ingin memodelkan hubungan ini. Dalam hal ini, digunakan satu variabel sebagai variabel prediktor atau penjelas atau bebas untuk menjelaskan variabel lainnya, yaitu variabel respon atau variabel terikat. Untuk melakukan hal ini, diperlukan hubungan yang baik antara kedua variabel. Model tersebut kemudian dapat digunakan untuk memprediksi perubahan variabel terikat. Hubungan yang kuat antara variabel bebas dan variabel terikat menghasilkan model yang baik.

Dalam analisis regresi, data digunakan untuk menentukan posisi garis yang paling mewakili hubungan antara kedua variabel. Pendekatannya adalah menggunakan diagram pencar untuk memposisikan garis secara visual. Pada diagram pencar, garis yang digambar melalui titik-titik data mengilustrasikan bahwa garis tersebut mungkin sesuai dengan data. Variabel penjelas (atau variabel bebas) selalu berada pada sumbu x. Variabel respon (atau variabel terikat) selalu berada pada sumbu y.

Garis yang ditarik lurus melalui titik-titik data mempunyai satu kelemahan yaitu posisinya sebagian didasarkan pada penilaian orang yang menggambar garis tersebut. Sebagai contoh, pada Gambar 3 diberikan diagram pencar dengan garis lurus yang mewakili penilaian empat orang. Semua garis kecuali garis A tampaknya masuk akal. Artinya, setiap garis dipusatkan di antara data grafik. Namun, masing-masing garis akan menghasilkan

perkiraan jumlah produk terjual yang berbeda untuk sejumlah panggilan penjualan tertentu.



Gambar 10.3. Diagram pencar dengan empat garis lurus

Dalam analisis regresi, diperlukan metode yang menghasilkan satu garis regresi terbaik. Metode ini disebut “prinsip kuadrat terkecil”. Metode ini memberikan garis "paling pas". Prinsip kuadrat terkecil merupakan prosedur matematis yang menggunakan data untuk memposisikan garis dengan tujuan meminimalkan jumlah kuadrat jarak vertikal antara nilai variabel terikat aktual (y aktual) dan nilai prediksi y .

Persamaan garis regresi linier mempunyai bentuk

$$\hat{y} = a + bx$$

dengan

\hat{y} (baca y topi) = nilai prediksi variabel y untuk nilai x yang dipilih.

a = titik potong (intersep) dari y . Ini adalah nilai prediksi variabel y ketika $x = 0$.

b = kemiringan (slope) garis regresi, atau rata-rata perubahan y untuk setiap perubahan sebesar satu satuan (baik kenaikan maupun penurunan) pada variabel bebas x .

x = nilai variabel bebas yang dipilih.

Tujuan analisis regresi adalah menentukan nilai a dan b untuk mendapatkan persamaan linier yang paling sesuai dengan data.

Contoh 4:

Seorang ahli hidrologi membuat model untuk memprediksi aliran volume Sungai (galon per menit) di persimpangan jembatan dengan variabel bebasnya adalah curah hujan harian (dalam milimeter). Model tersebut berupa persamaan garis regresi linier

$$\hat{y} = 1,6 + 29x.$$

Dapat diketahui dari persamaan tersebut bahwa titik potong (intersep) dari y sebesar 1,6. Artinya, pada hari tanpa curah hujan, akan terdapat 1,6 galon air/menit. mengalir di sungai itu. Kemudian, kemiringan (slope) sebesar 29 menunjukkan bahwa jika hujan turun sebanyak satu milimeter pada hari itu, maka aliran sungai akan bertambah sebesar 29 galon/menit. Misalkan, ingin diprediksi rata-rata aliran sungai jika pada suatu hari hujan turun 0,45 milimeter. Hal tersebut dapat ditentukan sebagai berikut:

$$\hat{y} = 1,6 + 29x = 1,6 + (29 * 0,45) = 14,65 \text{ galon/menit.}$$

Jadi, rata-rata aliran sungai jika pada suatu hari hujan turun 0,45 milimeter adalah sebesar 14,65 galon/menit.

C. PERBEDAAN ANTARA KORELASI DAN REGRESI

Korelasi dan regresi sama-sama digunakan sebagai pengukuran statistik untuk mendapatkan pemahaman yang baik tentang hubungan antar variabel. Jika koefisien korelasinya negatif (atau positif) maka kemiringan garis regresi juga akan negatif (atau positif). Tabel 3 berikut menyoroti perbedaan utama antara korelasi dan regresi.

Tabel 10.3. Perbedaan antara korelasi dan regresi

No	Korelasi	Regresi
1	Mengukur arah dan kuat hubungan antar variabel.	Mengukur sifat dan tingkat hubungan rata-rata antara dua variabel atau lebih berdasarkan unit data asli.
2	Merupakan ukuran relatif dari hubungan antar variabel.	Merupakan ukuran mutlak dari hubungan antar variabel.
3	Tidak bergantung pada satuan pengukuran.	Tidak terlepas dari satuan pengukuran.
4	Bukan alat peramalan.	Merupakan alat peramalan.
5	Mungkin terdapat korelasi nol atau tidak ada korelasi.	Tidak ada regresi nol.

6	Korelasi digunakan untuk mengetahui apakah variabel berhubungan atau tidak.	Regresi digunakan untuk menggambarkan secara numerik apakah variabel bebas berpengaruh terhadap perubahan variabel terikat.
7	Variabel-variabel tersebut dapat digunakan secara bergantian atau dipertukarkan.	Variabel bebas dan terikat tidak dapat dipertukarkan.
8	Analisis korelasi menggunakan nilai numerik bertanda untuk memperkirakan kekuatan hubungan antar variabel.	Analisis regresi menggunakan persamaan garis regresi untuk memprediksi nilai variabel terikat dari variabel bebas.

BAGIAN 11

REGRESI LINEAR SEDERHANA

A. PENGANTAR REGRESI LINEAR SEDERHANA

Dalam kehidupan sehari-hari banyak hal yang terjadi karena pengaruh hal lainnya, seperti dalam kasus properti, harga suatu rumah dipengaruhi oleh luas lahannya. Dalam kasus ini, luas lahan adalah komponen yang mempengaruhi terhadap harga rumah, sementara harga rumah adalah yang menerima pengaruh atau yang dipengaruhi oleh luas lahan. Dalam kasus pendidikan, nilai ujian siswa dipengaruhi oleh lama belajarnya. Nilai ujian adalah menerima pengaruh dan lama belajar adalah yang memberi pengaruh terhadap nilai ujian. Untuk mengkaji pengaruh tersebut dan menggambarannya dalam bentuk permodelan, digunakan suatu analisis yang disebut Regresi.

Regresi pertama kali dikenalkan oleh seorang ahli Antropologi dan Meteorologi asal Inggris, Sir Francis Galton (1822-1911). Analisis Regresi merupakan salah satu metode statistika yang digunakan untuk melihat pengaruh suatu variabel terhadap variabel lainnya. Selain itu analisis regresi juga digunakan untuk memodelkan hubungan antar variabel bebas dan variabel terikat. Saat ini, analisis regresi sudah diaplikasikan pada banyak bidang, seperti bidang ekonomi, manajemen, pendidikan bahkan sosial.

Variabel yang memberikan pengaruh disebut variabel bebas, dan disimbolkan dengan X , dan variabel yang menerima pengaruh disebut variabel terikat dan disimbolkan dengan Y . Jadi analisis regresi digunakan untuk melihat pengaruh variabel bebas terhadap variabel terikat. Pengaruh yang dikaji dalam regresi hanya berfokus pada pengaruh variabel bebas terhadap variabel terikat. Pada saat variabel bebas berpengaruh terhadap variabel terikat, maka tidak berarti bahwa variabel terikat juga berpegaruh pada variabel bebas.

Ada beberapa jenis analisis regresi. Jika dipandang dari jumlah variabel bebas yang dimiliki, maka regresi dibagi menjadi dua, yaitu regresi sederhana dan regresi berganda. Dikatakan sederhana jika dalam regresi tersebut hanya melibatkan satu variabel bebas. Namun Jika dalam regresi tersebut melibatkan lebih dari satu variabel bebas, maka disebut regresi berganda. Jika melihat dari kelinearitannya, maka regresi dibagi menjadi regresi linear dan regresi non-linear. Regresi linear adalah regresi yang ordo atau pangkat dari variabel bebasnya satu, sehingga apabila digambarkan menggunakan plot akan membentuk garis lurus. Regresi non-linear adalah regresi yang pangkat variabel bebasnya selain satu. Contoh regresi lainnya adalah regresi logistik, dan regresi data panel.

Secara general, kegunaan masing-masing analisis regresi ini sama, yaitu untuk melihat pengaruh variabel bebas terhadap variabel terikat. Pada bagian ini, akan berfokus membahas regresi linear sederhana.

Analisis regresi linear sederhana adalah analisis yang digunakan untuk melihat pengaruh satu variabel bebas terhadap variabel terikat. Jika data pada variabel bebas dan variabel terikat digambarkan pada scatterplot dan membentuk garis lurus, maka hal itu menunjukkan bahwa kedua variabel tersebut memiliki hubungan linear. Jenis data yang digunakan dalam analisis regresi linear sederhana adalah data numerik.

B. TUJUAN DAN MANFAAT REGRESI LINEAR SEDERHANA

Ada beberapa tujuan dan manfaat dari analisis regresi linear sederhana diantaranya, yaitu :

1. Mengetahu pengaruh variabel bebas terhadap variabel terikat
Dalam analisis regresi, salah satu tujuannya adalah untuk mengetahui apakah variabel bebas berpengaruh terhadap variabel terikat. Untuk mengetahui hal ini, bisa dilakukan uji F.
2. Membentuk Model Regresi Linear sederhana
Model ini menggambarkan pengaruh variabel bebas terhadap variabel terikatnya.
3. Prediksi
Dari model regresi yang diperoleh, dapat dilakukan prediksi nilai tengah variabel terikat saat nilai variabel bebas diketahui

C. MODEL REGRESI LINEAR SEDERHANA

Regresi linear sederhana melibatkan satu variabel bebas X yang bersifat tetap, dan satu variabel terikat Y , yang merupakan variabel acak. Data disajikan sebagai sepasang pengamatan $\{(x_i, y_i); i = 1, 2, \dots, n\}$. Model regresi linear sederhana akan menghasilkan garis lurus yang terbentuk dari pasangan data tersebut. Adapun model regresi linear sederhana untuk populasi adalah

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

dimana β_0 adalah intersep dan β_1 disebut slope, nilai β_0 dan β_1 disebut koefisien regresi, sementara ε adalah kompon error/galat acak. Error ε diasumsikan $E(\varepsilon)=0$ dan $\text{var}(\varepsilon)=\sigma^2$ dan $\varepsilon \sim N(0, \sigma^2)$, selain itu ε juga diasumsikan tidak berkorelasi antar error, atau dapat dikatakan bahwa nilai suatu error tidak bergantung pada nilai error lainnya.

Koefisien regresi β_0 dan β_1 adalah dua parameter yang nilainya tidak diketahui namun dapat diduga menggunakan data sampel. Adapun dugaan untuk β_0 , dan β_1 dilambangkan dengan b_0 , dan b_1 , Sementara nilai $\mu_{Y|X}$ akan diduga dengan \hat{y} dari bentuk persamaan garis regresi yang dihasilkan oleh data sampel.

$$\hat{y} = b_0 + b_1 X$$

Dalam kondisi real, data populasi sangat jarang tersedia, sehingga data sampel yang umumnya digunakan untuk menduga kondisi populasi. Begitu pula dalam analisis regresi linear sederhana,

dimana nilai parameter akan diduga dengan data sampel. Adapun model regresi liner sederhana untuk sampel adalah

$$y_i = b_0 + b_1X + e_i$$

dimana $e_i = y_i - \hat{y}$ adalah sisaan. Sisaan dan galat adalah dua hal yang berbeda, dimana galat/error merupakan selisih dari nilai aktual variabel Y dan nilai dugaan dari data populasi, sementara sisaan adalah selisih nilai aktual variabel Y dengan nilai dugaan dari data sampel.

Perlu diingat bahwa persamaan regresi dan model regresi memiliki perbedaan. Dalam persamaan regresi, tidak ada unsur error atau sisaan, contoh persamaan garis regresi $\hat{y} = b_0 + b_1X$. Sementara dalam model regresi, terdapat unsur error atau sisaan di dalamnya, seperti $y_i = b_0 + b_1X + e_i$.

D. PENDUGA PARAMETER REGRESI LINEAR SEDERHANA

Metode kuadrat terkecil adalah metode yang umum digunakan untuk menduga parameter β_0 dan β_1 pada regresi linear sederhana. Metode ini dilakukan dengan cara meminimumkan jumlah kuadrat galat dengan melakukan turunan parsial terhadap $\sum e_i^2$ sedemikian sehingga diperoleh dugaan parameter b_0 , dan b_1 , sebagai berikut:

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

dimana $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ dan $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Penduga parameter b_0 dapat diinterpretasikan sebagai rata-rata nilai Y saat tidak ada pengaruh dari variabel X. Sementara b_1 dapat diinterpretasikan sebagai rata-rata kenaikan y sebesar b_1 setiap nilai x naik satu satuan.

Contoh: Misalkan seorang pengusaha kosmetik ingin melihat pengaruh dari biaya promosi (juta rupiah) terhadap volume barang yang terjual (ribu item). Maka diambil 10 sampel dengan dua buah variabel yaitu biaya promosi (X) dan Jumlah barang terjual (Y).

X	12	14	13	12	15	13	14	15	13	14
Y	56	62	60	61	65	66	60	63	65	62

Dari data, diketahui banyak data $n = 10$, $\sum_{i=1}^{10} x_i = 135$; $\sum_{i=1}^{10} y_i = 620$; $\sum_{i=1}^{10} x_i y_i = 8383$; dan $\sum_{i=1}^{10} x_i^2 = 1833$. Nilai yang sudah diperoleh, digunakan untuk mendapatkan nilai b_1

$$b_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$= \frac{10(8383) - (135)(620)}{10(1833) - (135)^2}$$

$$= 1,238$$

Setelah diperoleh nilai $b_1 = 1,238$, selanjutnya hitung nilai b_0

$$b_0 = \bar{y} - b_1\bar{x} = 62 - (1,238)(13,5) = 45,28$$

Sehingga persamaan regresi yang memperlihatkan pengaruh biaya promosi (X) terhadap volume penjualan (Y) adalah

$$\hat{Y} = 45,28 + 1,238X$$

Dari persamaan tersebut, nilai $b_0 = 45,28$ dapat diinterpretasikan sebagai rata-rata volume penjualan saat tidak ada biaya promosi adalah sebanyak 45,28 ribu item. Jadi sekitar 45280 item terjual tanpa adanya biaya promosi yang dikeluarkan oleh pemilik perusahaan kosmetik tersebut. Nilai $b_1 = 1,238$ dapat diinterpretasikan bahwa setiap kenaikan satu juta rupiah biaya produksi (tergantung satuan variabel Y), maka volume penjualan akan meningkat sebanyak 1,238 ribu item atau sebanyak 1238 item.

Persamaan regresi tersebut juga dapat digunakan untuk melakukan prediksi saat nilai x diketahui, misalnya jika biaya produksi yang dikeluarkan $x=10$, maka $\hat{Y} = 45,28 + 1,238(10) = 57,7$. Jadi saat biaya promosi yang dikeluarkan adalah 10 juta, maka volume barang yang terjual adalah 57,7 ribu item.

E. ANALISIS KERAGAMAN REGRESI

Analisis keragaman regresi atau lebih dikenal dengan Analisis Variansi atau anova digunakan untuk melihat apakah variabel bebas berpengaruh terhadap variabel terikatnya dengan menguraikan keragaman yang dimiliki variabel Y. Jumlah kuadrat regresi (JKR), Jumlah kuadrat Galat (JKG) dan jumlah kuadrat Total (JKT) adalah komponen yang akan digunakan, dimana:

$$JKT = JKR + JKG$$

Dengan $JKT = \sum (y_i - \bar{y})^2$ menggambarkan total kuadrat dari selisih nilai aktual ke nilai rata-rata. $JKR = \sum (\hat{y}_i - \bar{y})^2$, menggambarkan total kuadrat dari selisih nilai dugaan ke nilai rata-rata. $JKG = \sum (y_i - \hat{y}_i)^2$, menggambarkan total kuadrat dari selisih nilai aktual ke nilai dugaan.

Nilai JKT, JKR dan JKG akan disajikan dalam tabel anova sebagai berikut. Selain memuat Nilai jumlah Kuadrat, pada tabel anova juga memuat sumber keragaman, derajat bebas, dan nilai kuadrat Tengah

Tabel 11.1 Tabel Anova untuk regresi linear sederhana

Sumber variasi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F hitung
Regresi	1	JKR	$KTR = \frac{JKR}{1}$	$\frac{KTR}{KTG}$
Galat	n-2	JKG	$KTG = \frac{JKG}{n-2}$	

Total	n-1	JKT		
-------	-----	-----	--	--

Untuk menentukan apakah variabel bebas berpengaruh terhadap variabel terikat, maka akan dilakukan uji hipotesis. Adapun hipotesis yang digunakan adalah:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Statistik Uji yang digunakan adalah Uji F

$$F = \frac{KTR}{KTG}$$

Kriteria penolakan H_0 pada taraf nyata α adalah ketika F hitung $> F_{(a,1,n-2)}$, dan terima H_0 untuk kondisi lainnya. Ketika terjadi penolakan H_0 maka dapat dikatakan bahwa variabel bebas mempengaruhi variabel terikat. Jika gagal tolak H_0 , maka disimpulkan bahwa belum cukup bukti mengatakan bahwa variabel bebas mempengaruhi variabel terikat.

F. SELANG KEPERCAYAAN

Pada bagian ini akan dibahas tentang selang kepercayaan dari paramer regresi linear sederhana. Selang kepercayaan atau confident interval (CI) merupakan interval antar dua nilai, dimana parameter

Selang kepercayaan sebesar $(1 - \alpha)100\%$ untuk slope β_1 adalah

$$\hat{\beta}_1 - t_{\frac{\alpha}{2}, n-2} se(\hat{\beta}_1) \leq \hat{\beta}_1 \leq \hat{\beta}_1 + t_{\frac{\alpha}{2}, n-2} se(\hat{\beta}_1)$$

Dan Selang kepercayaan sebesar $(1-\alpha)100\%$ untuk intersep β_0 adalah

$$\hat{\beta}_0 - t_{\frac{\alpha}{2}, n-2} se(\hat{\beta}_0) \leq \hat{\beta}_0 \leq \hat{\beta}_0 + t_{\frac{\alpha}{2}, n-2} se(\hat{\beta}_0)$$

Dimana nilai $\hat{\beta}_0 = b_0$, $\hat{\beta}_1 = b_1$, yang merupakan nilai penduga parameter regresi. Nilai ini telah diperoleh sebelumnya saat membangun model regresi linear sederhana. Selanjutnya untuk nilai $se(\hat{\beta}_1) = \sqrt{\frac{KTG}{\sum(x_i - \bar{x})^2}}$ dan nilai $se(\hat{\beta}_0) = \sqrt{\frac{KTG \sum x_i^2}{n \sum(x_i - \bar{x})^2}}$, nilai KTG sendiri sudah diperoleh pada tabel anova, sementara nilai yang $\sum x_i^2$ dan $\sum(x_i - \bar{x})^2$ diperoleh dari data variabel bebas.

G. KOEFISIEN DETERMINASI

Setelah mendapatkan pendugaan model regresi, maka akan muncul pertanyaan seberapa besar keragaman dari variabel terikat Y yang mampu dijelaskan oleh model. Pertanyaan ini akan terjawab dari nilai Koefisien detrminasi yang dilambangkan dengan R^2 . Nilai R^2 berkisar antara 0 sampai 100%. Semakin mendekati 100%, maka akan semakin besar kergaman yang mampu dijelaskan oleh model.

Nilai R^2 dapat dihitung dengan memanfaatkan beberpaa nilai yang telah didapat pada tabel anova, dimana:

$$R^2 = JKR / JKT$$

H. ASUMSI REGRESI LINEAR SEDERHANA

Ada beberapa asumsi yang harus dipenuhi dalam analisis regresi linear sederhana.

1. Uji Normalitas

Uji normalitas dilakukan pada data sisaan/ residual e , bukan pada data di variabel bebas maupun variabel terikat. Uji normalitas adalah uji yang digunakan untuk melihat apakah residual berdistribusi normal atau tidak. Model regresi yang baik adalah model regresi yang memiliki residual yang berdistribusi normal. Asumsi ini diperlukan untuk menguji signifikansi dan pembentukan selang kepercayaan.

Pengujian normalitas residual dapat dilakukan dengan histogram, dengan plot, uji Chi-Square atau uji kolmogorov-Smirnov. Jika asumsi normalitas terlanggar, maka dapat dilakukan pemangkasan outlier, penambahan data, dan transformasi data, seperti transformasi akar, logaritma dan lainnya.

2. Uji heteroskedastisitas

Kesamaan atau homogenitas varian pada penduga kuadrat terkecil merupakan suatu asumsi yang harus terpenuhi. Homogenitas varian menggambarkan kondisi bahwa variansi antar residual adalah sama, yang disebut juga homoskedastisitas. Jika varian dari residual tidak sama, maka disebut heteroskedastisitas. ini akan menyebabkan hilangnya presisi dalam pendugaan. Hal ini tentu tidak diharapkan terjadi. Untuk mendeteksi heteroskedastisitas, dapat dilakukan dengan

beberapa cara, seperti menggunakan scatterplot, uji Glejser, uji Park. Cara penanganan heteroskedastisitas adalah dengan melakukan transformasi data dan WLS.

3. Autokorelasi

Autokorelasi adalah suatu kondisi dimana antar residual saling berkorelasi. Kasus autokorelasi sering muncul pada data yang berkaitan dengan waktu atau data time series. Kemunculan autokorelasi menyebabkan varian sampel tidak mampu menjadi penduga yang baik bagi varian populasi, hal lain yang terjadi adalah tidak akuratnya dugaan parameter regresi.

Dalam analisis regresi, autokorelasi bukanlah kondisi yang diharapkan. Untuk mendeteksi terjadinya autokorelasi dapat digunakan uji Durbin Watson. Jika setelah pengujian diketahui bahwa terjadi autokorelasi, maka sapat dilakukan transformasi Cochrane Orcutt.

4. Uji Linieritas

Uji linearitas digunakan untuk mengetahui apakah model regresi yang terbentuk memiliki hubungan yang linear atau tidak.

Dalam analisis regresi, ada asumsi lain yang harus dipenuhi, yaitu terkait multikolinieritas. Dalam analisis regresi, diharapkan tidak terjadi multikolinieritas. Multikolinieritas sendiri adalah kondisi dimana suatu variabel bebas berkorelasi dengan variabel lainnya. Namun, dalam analisis regresi linear sederhana, tidak perlu dilakukan pengecekan asumsi ini karena hanya terdapat satu variabel bebas dalam regresi linear sederhana.

BAGIAN 12

REGRESI LINEAR BERGANDA

A. PENGANTAR

Dalam bab sebelumnya telah diuraikan model regresi linear sederhana yaitu model regresi dua variabel yang merupakan model regresi yang menjelaskan hubungan dan pengaruh antara satu variabel tidaklah demikian hubungan dan pengaruh tersebut sering tidak mencukupi atau kurang realitis. Dalam contoh penjualan sebagai variabel terikat dan biaya promosi sebagai variabel bebas, misalnya diasumsikan penjualan bukan saja dipengaruhi oleh biaya promosi, padahal dalam teori ekonomi dapat dijelaskan bahwa penjualan juga dapat dipengaruhi oleh variabel lain antara lain harga dan banyak outlet.

Berdasarkan persoalan tersebut maka pada bab ini, pembahasan akan diperluas mengenai model regresi linear berganda (majemuk) yaitu model regresi yang melibatkan lebih dari pada satu variabel bebas dan satu variabel terikat Y , dengan dua atau lebih variabel bebas (X_1, X_2, \dots, X_n). Model regresi berganda yang paling sederhana adalah regresi tiga variabel yaitu model regresi dengan satu variabel terikat (Y) dan dua variabel bebas (X_1, X_2). Dan model regresi ini yang kita bahas pada bab ini.

Tujuan daripada bab ini adalah mahasiswa diharapkan dapat memahami dan menganalisis regresi linear berganda dan dapat melakukan inferensi dan memberikan intepretasi terhadap koefisien regresi.

B. PENDUGAAN TITIK KOEFISIEN REGRESI LINIEAR BERGANDA

Bentuk umum model regresi populasi tiga variabel dapat dibuat sebagai berikut:

$$E(Y_i | X_{1i}, X_{2i}) = \beta_{0.12} + \beta_{0.1.2}X_{1i} + \beta_{02.1}X_{2i} \dots \dots \dots (11.1)$$

Bentuk stokastiknya dapat dibuat adalah sebagai berikut:

$$Y_i = E(Y_i | X_{1i}, X_{2i}) + U_i = \beta_{0.12} + \beta_{0.1.2}X_{1i} + \beta_{02.1}X_{2i} + U_i \dots \dots \dots (*)$$

Y adalah sebagai variabel terikat, X₁ dan X₂ merupakan variabel bebas, sedangkan U adalah faktor gangguan yang stokastik, dan i menyatakan observasi (pengamatan) yang ke-i. X_{1i} berarti observasi ke-i atas variabel X₁. X_{2i} berarti observasi ke-i atas variabel X₂. Angka indeks bawah (titik bawah) yang dicantumkan pada koefisien, dapat diintepretasikan sebagai berikut:

1. Tik-bawah 1 menyatakan variabel terikat Y
2. Tik-bawah 2 menyatakan variabel bebas X₁
3. Tik-bawah 3 menyatakan variabel bebas X₂

β_{0.12} adalah intersep yang menunjukkan pengaruh atau efek rata-rata semua variabel yang tidak dimasukan ke dalam model

terhadap Y. Meskipun interpretasi mekanismenya adalah rata-rata nilai Y ketika X_1 dan X_2 disamakan dengan nol. Koefisien $\beta_{0.1,2}$ dan $\beta_{02,1}$ disebut koefisien regresi parsial.

Persamaan (*) di atas dapat ditulis dalam bentuk yang lebih sederhana sebagai berikut:

$$Y_i = E(Y_i | X_{1i}, X_{2i}) + U_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + U_i \dots \dots \dots (11.2)$$

Sedangkan bidang regresi sampelnya (dalam bentuk yang lebih sederhana sebagai penduga dari bidang regresi populasi, dinyatakan sebagai berikut:

$$\hat{Y} = b_0 + b_1 X_{1i} + b_2 X_{2i} \dots \dots \dots (11.3)$$

Bentuk stokastiknya adalah

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + e_i = \hat{Y} + e_i \dots \dots \dots (11.4)$$

Keterangan:

- \hat{Y}_i = merupakan penduga $E(Y_i | X_{1i}, X_{2i})$
- $b_0 = b_{0.12}$ merupakan penduga $\beta_0 = \beta_{0.12}$
- $b_1 = b_{0.12}$ merupakan penduga $\beta_1 = \beta_{0.12}$
- $b_2 = b_{0.21}$ merupakan penduga $\beta_2 = \beta_{0.21}$
- e_i = merupakan penduga U_i

Interpretasi terhadap nilai b_0 , b_1 dan b_2 misalnya persamaan regresi sampel sebagai penduga persamaan regresi populasi sebagai berikut ini:

$$\hat{Y} = b_0 + b_1 X_{1i} + b_2 X_{2i}$$

Interpretasi terhadap nilai b_0 , b_1 dan b_2

b_0 = menyatakan nilai rata-rata variabel Y atau $E(Y_i | X_{1i}, X_{2i})$, bila $X_1 = 0$ dan $X_2 = 0$

b_1 = menyatakan perubahan nilai rata-rata variabel Y karena perubahan 1 unit X_1 dengan asumsi variabel lain tetap atau konstan.

b_2 = menyatakan perubahan nilai rata-rata variabel Y karena perubahan 1 unit X_2 dengan asumsi variabel lain tetap atau konstan.

C. KOEFISIEN DETERMINASI REGRESI LINIER BERGANDA (R^2)

Pengujian koefisien determinasi ini dilakukan dengan maksud mengukur kemampuan model dalam menerangkan seberapa pengaruh variabel independen secara bersama-sama (stimultan) mempengaruhi variabel dependen yang dapat diindikasikan oleh nilai *adjusted R – Squared* (Ghozali, 2016). Koefisien determinasi menunjukkan sejauh mana kontribusi variabel bebas dalam model regresi mampu menjelaskan variasi dari variabel terikatnya. Koefisien determinasi dapat dilihat melalui nilai R-square (R^2) pada tabel Model Summary.

Menurut Ghozali (2016) nilai koefisien determinasi yang kecil memiliki arti bahwa kemampuan variabel- variabel independen dalam menjelaskan variabel dependen sangat terbatas, Sebaliknya jika nilai mendekati 1 (satu) dan menjauhi 0 (nol) memiliki arti

bahwa variabel – variabel independen memiliki kemampuan memberikan semua informasi yang dibutuhkan untuk memprediksi variabel dependen (Ghozali, 2016).

Uji koefisien determinasi dilakukan untuk mengetahui seberapa besar variabel endogen secara simultan mampu menjelaskan variabel eksogen. Semakin tinggi nilai R^2 berarti semakin baik model prediksi dari model penelitian yang diajukan. Uji koefisien determinasi (R^2) dilakukan untuk menentukan dan memprediksi seberapa besar atau penting kontribusi pengaruh yang diberikan oleh variabel independen secara bersama-sama terhadap variabel dependen. Nilai koefisien determinasi yaitu antara 0 dan 1. Jika nilai mendekati 1, artinya variabel independen memberikan hampir semua informasi yang dibutuhkan untuk memprediksi variabel dependen. Namun, jika nilai R^2 semakin kecil, artinya kemampuan variabel – variabel independen dalam menjelaskan variabel dependen cukup terbatas (Ghozali, 2016).

Menurut Chin (1998), nilai *R-Square* dikategorikan kuat jika lebih dari 0,67, moderat jika lebih dari 0,33 tetapi lebih rendah dari 0,67, dan lemah jika lebih dari 0,19 tetapi lebih rendah dari 0,33.

Untuk menghitung nilai koefisien determinasi regresi berganda dapat dihitung dengan rumus (Gujarati, 2016):

$$R^2 = \frac{b_1 \sum y_i x_{2i} + b_2 \sum y_i x_{3i}}{\sum y_i^2}$$

D. KOEFISIEN KORELASI GANDA

Dalam analisis regresi dua variabel, tingkat keeratan hubungan antara variabel terikat dengan variabel bebasnya diukur oleh koefisien korelasi r . Sedangkan dalam analisis regresi dengan tiga variabel tingkat keeratn hubungan antara variabel bebas dengan variabel terikatnya secara bersama-sama diukur oleh koefisien korelasi berganda (R). Meskipun r dapat bernilai positif atau negatif, namun R selalu bernilai positif. Dalam praktek R kurang penting yang lebih penting adalah R^2 dan nilai R dapat dihitung dengan rumus adalah sebagai berikut:

$$R = \sqrt{R^2}$$

Keterangan:

R = koefisien korelasi ganda

E. ASUMSI-ASUMSI MODEL REGRESI LINIER BERGANDA

Menurut Gujarati (2016) menyatakan bahwa asumsi-asumsi mendasar yang harus dipenuhi pada model regresi linier berganda adalah sebagai berikut:

1. Model regresinya adalah linier dalam parameter.
2. Nilai rata-rata dari *error* adalah nol.
3. Variansi dari *error* adalah konstan (homoskedastik).
4. Tidak terjadi autokorelasi pada *error*.
5. Tidak terjadi multikolinieritas pada variabel bebas.
6. *Error* berdistribusi normal.

F. ESTIMASI PARAMETER MODEL REGRESI LINIER BERGANDA

Estimasi parameter ini bertujuan untuk mendapatkan model regresi linier berganda yang akan digunakan dalam analisis. Pada materi pelatihan ini, metode yang digunakan untuk mengestimasi parameter model regresi linier berganda adalah metode kuadrat terkecil atau sering juga disebut dengan metode *ordinary least square* (OLS). Metode OLS ini bertujuan meminimumkan jumlah kuadrat *error*. Berdasarkan persamaan (2.2) dapat diperoleh penaksir (*estimator*) OLS untuk β adalah sebagai berikut (Kutner, *et.al.*, 2004):

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Penaksir OLS pada persamaan (2.3) merupakan penaksir yang tidak bias, linier dan terbaik (*best linear unbiased estimator/BLUE*) (Sembiring, 2003; Gujarati, 2003; Greene, 2003 dan Widarjono, 2007).

G. PENGUJIAN PARAMETER MODEL REGRESI LINIER BERGANDA

Pengujian parameter ini bertujuan untuk mengetahui ada atau tidaknya pengaruh variabel bebas terhadap variabel tidak bebas, baik secara serentak maupun secara parsial:

1. Pengujian Parameter Secara Serentak (Simultan)

Prosedur pengujian parameter secara simultan adalah sebagai berikut:

- a. Membuat hipotesis.

$$H_0 = \beta_1 = \beta_2 = \dots = \beta_{n-1} = 0$$

H_1 : Tidak semua β_k sama dengan nol, untuk $k = 1, 2, \dots, p-1$.

(Kutner, *et.al.*, 2004)

atau:

H_0 : Variabel X_1, X_2, \dots, X_n secara simultan tidak berpengaruh terhadap variabel terikat

H_1 : Variabel X_1, X_2, \dots, X_n secara simultan berpengaruh terhadap variabel terikat

- b. Menentukan tingkat signifikansi (α).

Tingkat signifikansi (α) yang seringkali digunakan dalam penelitian adalah 5%.

- c. Menentukan statistik uji.

Statistik uji yang digunakan adalah:

$$F = \text{RKR/RKE}$$

dengan:

RKR adalah rata-rata kuadrat regresi (dapat diperoleh dari Tabel Analisis Variansi).

RKE adalah rata-rata kuadrat *error* (dapat diperoleh dari Tabel Analisis Variansi).

- d. Menentukan daerah kritik (penolakan H_0).

Daerah kritik yang digunakan adalah H_0 ditolak bila F hitung $>$ F tabel)

Ho diterima jika $F \text{ hitung} \leq F \text{ table}$

Selain dari daerah kritik di atas, dapat juga digunakan daerah kritik yang lain yaitu jika nilai peluang (Sig.) < tingkat signifikansi (α), maka H_0 ditolak.

e. Menarik kesimpulan

Kesimpulan dapat dilakukan dengan membandingkan nilai F hitung dengan F tabel atau membandingkan tingkat signifikansinya dengan α

2. Pengujian Parameter Secara Individu

Prosedur pengujian untuk parameter secara parsial adalah sebagai berikut:

b. Membuat hipotesis

$H_0 : \beta_i = 0$

$H_1 : \beta_i \neq 0$ untuk $i = 1, 2, \dots, p-1$. (Kutner, *et.al.*, 2004)

atau:

H_0 : Variabel bebas ke- i tidak berpengaruh terhadap variabel terikat

H_1 : Variabel bebas ke- i berpengaruh terhadap variabel terikat

untuk $i = 1, 2, \dots, p-1$.

c. Menentukan tingkat signifikansi (α)

Tingkat signifikansi (α) yang seringkali digunakan dalam penelitian adalah 5%.

d. Menentukan statistik uji.

Statistik uji yang digunakan adalah dengan formulasi (Gujarati, 2016):

$$t_i = \frac{\beta_i}{Se(\beta_i)}$$

Keterangan:

t_i = Besarnya nilai t-hitung

β = Koefisien variabel bebas

$Se(\beta_i)$ = Standar error β_i

e. Menentukan daerah kritik (penolakan H_0).

H_0 diterima jika = $t_{hitung} \leq t_{\alpha} ; (n-k)$

H_0 ditolak jika = $t_{hitung} > t_{\alpha} ; (n-k)$

f. Kesimpulan

Membandingkan t hitung dengan t tabel

H. PELANGGARAN-PELANGGARAN TERHADAP ASUMSI REGRESI LINIER BERGANDA

Dalam analisis regresi linier berganda terdapat beberapa pelanggaran-pelanggaran yang seringkali dilakukan terhadap asumsi-asumsinya, diantaranya diuraikan berikut ini:

1. Multikolinieritas

Multikolinieritas adalah terjadinya hubungan linier antara variabel bebas dalam suatu model regresi linier berganda (Gujarati, 2016). Hubungan linier antara variabel bebas dapat terjadi dalam bentuk

hubungan linier yang sempurna (*perfect*) dan hubungan linier yang kurang sempurna (*imperfect*).

Adapun dampak adanya multikolinieritas dalam model regresi linier berganda adalah (Gujarati, 2016 dan Widarjono, 2007):

- a. Penaksir OLS masih bersifat BLUE, tetapi mempunyai variansi dan kovariansi yang yang besar sehingga sulit mendapatkan taksiran (estimasi) yang tepat.
- b. Akibat penaksir OLS mempunyai variansi dan kovariansi yang yang besar, menyebabkan interval estimasi akan cenderung lebih lebar dan nilai hitung statistik uji t akan kecil, sehingga membuat variabel bebas secara statistik tidak signifikan mempengaruhi variabel tidak bebas.
- c. Walaupun secara individu variabel bebas tidak berpengaruh terhadap variabel tidak bebas melalui uji t, tetapi nilai koefisien determinasi (R^2) masih bisa relatif tinggi.

Selanjutnya untuk mendeteksi adanya multikolinieritas dalam model regresi linier berganda dapat digunakan nilai *variance inflation factor* (VIF) dan *tolerance* (TOL) dengan ketentuan jika nilai VIF melebihi angka 10, maka terjadi multikolinieritas dalam model regresi. Kemudian jika nilai TOL sama dengan 1, maka tidak terjadi multikolinieritas dalam model regresi.

2. Heteroskedastisitas

Heteroskedastisitas adalah variansi dari *error* model regresi tidak konstan atau variansi antar *error* yang satu dengan *error* yang lain berbeda (Widarjono, 2007).

Dampak adanya heteroskedastisitas dalam model regresi adalah walaupun estimator OLS masih linier dan tidak bias, tetapi tidak lagi mempunyai variansi yang minimum dan menyebabkan perhitungan *standard error* metode OLS tidak bisa dipercaya kebenarannya. Selain itu interval estimasi maupun pengujian hipotesis yang didasarkan pada distribusi t maupun F tidak bisa lagi dipercaya untuk evaluasi hasil regresi.

Akibat dari dampak heteroskedastisitas tersebut menyebabkan estimator OLS tidak menghasilkan estimator yang BLUE dan hanya menghasilkan estimator OLS yang *linear unbiased estimator* (LUE). Selanjutnya dilakukan deteksi masalah heteroskedastisitas dalam model regresi. Salah satu cara yang dapat digunakan untuk mendeteksi adanya heteroskedastisitas dalam model regresi adalah dengan Metode Glejser. Glejser merupakan seorang ahli ekonometrika dan mengatakan bahwa nilai variansi variabel *error* model regresi tergantung dari variabel bebas. Selanjutnya untuk mengetahui apakah pola variabel *error* mengandung heteroskedastisitas Glejser menyarankan untuk melakukan regresi nilai mutlak residual dengan variabel bebas. Jika hasil uji F dari model regresi yang diperoleh tidak signifikan, maka tidak ada heteroskedastisitas dalam model regresi (Widarjono, 2007).

3. Uji Autokorelasi

Uji autokorelasi yang dilakukan dalam penelitian ini bertujuan untuk mengetahui apakah dalam model regresi linear yang diperoleh terdapat korelasi antara kesalahan pengganggu pada data. Jika terjadi autokorelasi, maka persamaan tersebut menjadi tidak baik atau tidak layak dipakai prediksi. Ukuran dalam menentukan ada tidaknya gejala autokorelasi dilakukan dengan uji *Durbin-Watson* (DW) dengan formulasi (Gujarati, 2016):

$$D - W = \frac{\sum (e_t - e_{t-1})}{\sum e_t^2}$$

Menentukan gejala autokorelasi melalui *Durbin-Watson* (DW) dengan ketentuan sebagai berikut (Aglifari, 2016):

- a. Terjadi gejala autokorelasi positif apabila nilai DW di bawah -2 (DW < -2).
- b. Terjadi gejala autokorelasi negatif apabila nilai DW diatas 2 (DW > 2).
- c. Tidak terjadi gejala autokorelasi apabila nilai DW berada diantara -2 dan 2 (-2 < DW < 2).

4. Uji Normalitas

Uji normalitas adalah uji asumsi klasik yang digunakan untuk mengetahui apakah model regresi yang didapat memiliki residual berdistribusi normal atau tidak. Model regresi yang baik jika residual model regresi yang didapat berdistribusi normal. Uji normalitas dapat dilakukan dengan uji histogram, uji normal P Plot,

dan uji Chi Square. Uji normalitas dengan Chi Square dilakukan dengan formulasi (Gujarati, 2016):

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

Dalam penelitian ini cara untuk mengetahui apakah data penelitian berdistribusi normal atau tidak dapat dilakukan dengan melihat *normal probability plots*. Pengujian ini dilakukan dengan cara melihat titik-titik persebaran data terhadap garis diagonal pada grafik histogram. Kriteria pengambilan keputusan dengan melihat grafik histogram adalah sebagai berikut (Aglifari, 2016):

- a. Apabila titik-titik persebaran data pada histogram mengikuti garis diagonal dan tidak menyebar menjauh, maka dapat disimpulkan bahwa data yang digunakan dalam model berdistribusi normal begitu juga sebaliknya.
- b. Apabila terdapat garis melengkung keatas seperti membentuk gunung dan terlihat sempurna dengan kaki yang simetris, maka dapat disimpulkan bahwa data yang digunakan dalam model berdistribusi normal begitu juga sebaliknya.

BAGIAN 13

PENGANTAR ANALISI VARIAN (ANOVA)

A. KONSEP DASAR ANOVA

Analisis varian (ANOVA) atau sidik ragam adalah metode statistik yang powerful untuk menguji hipotesis atau signifikansi perbedaan rata-rata dari dua atau lebih kelompok dalam sebuah populasi secara simultan.

ANOVA didasarkan pada pemisahan total varian yang ada dalam data menjadi dua komponen, yaitu variasi antara kelompok (*between groups*) bisa disebut perlakuan dan variasi dalam kelompok (*within groups*) bisa disebut nilai residual atau error (galat) untuk menentukan apakah terdapat perbedaan yang signifikan secara statistik antar kelompok dalam populasi. Jika rata-rata antar kelompok berbeda secara signifikan, maka dapat disimpulkan bahwa terdapat perbedaan yang berarti dalam populasi yang diteliti.

Teknik ini banyak digunakan dalam berbagai bidang penelitian, seperti pertanian, sains, psikologi, pendidikan, sosial dan ekonomi. ANOVA mudah dipelajari dan dipahami, sehingga banyak digunakan oleh para peneliti. Pemanfaatan ANOVA dalam berbagai bidang, seperti:

1. Penelitian

ANOVA dapat digunakan untuk menguji efektivitas suatu perlakuan atau intervensi, membandingkan kelompok populasi

berdasarkan karakteristik tertentu, dan menganalisis interaksi antara faktor-faktor yang berbeda.

2. Bisnis

ANOVA dapat digunakan untuk menganalisis data penjualan, membandingkan kinerja produk atau layanan yang berbeda, dan mengevaluasi efektivitas strategi pemasaran.

3. Manufaktur

ANOVA dapat digunakan untuk mengontrol kualitas produk, mengidentifikasi penyebab cacat, dan mengoptimalkan proses produksi.

Dengan memenuhi asumsi yang diperlukan dan memahami prosedur yang tepat, ANOVA dapat memberikan informasi yang berharga dalam menganalisis data penelitian dan mengambil keputusan yang tepat. Pemahaman yang baik tentang ANOVA dan penggunaan aplikasi yang sesuai akan membantu peneliti dalam menganalisis data secara efektif dan efisien.

B. ASUMSI ANOVA

Sebelum menggunakan ANOVA, sangat penting untuk memastikan bahwa data memenuhi asumsi-asumsi dasar berikut:

1. Data Berdistribusi Normal

Distribusi data dalam setiap kelompok harus mendekati distribusi normal. Sampel harus diambil secara acak dari populasi

yang berdistribusi normal. Ini diperlukan sehingga nilai error terdistribusi secara normal.

2. Varian Homogen

Varians data dalam setiap kelompok harus sama. Ini diperlukan sehingga nilai error akan memiliki ragam homogen untuk setiap taraf perlakuan

3. Independensi

Observasi dalam setiap kelompok harus independen satu sama lain. Unit satuan percobaan ditentukan dan ditempatkan secara acak pada setiap kelompok perlakuan. Ini diperlukan sehingga nilai error independen (saling bebas) satu sama lain.

4. Aditif

Pengaruh dari faktor perlakuan dan lingkungan dan error bersifat aditif, tinggi rendahnya respons diakibatkan dari pengaruh penambahan perlakuan atau kelompok. Nilai Respons merupakan nilai rata-rata umum ditambah dengan penambahan dari perlakuan dan error.

Jadi, asumsi-asumsi yang wajib dipenuhi dalam melakukan ANOVA adalah, Normalitas, Homogen, Independensi, dan Aditif. Jika asumsi ini tidak terpenuhi, maka hasil analisis ANOVA mungkin tidak valid.

C. JENIS-JENIS ANOVA

Terdapat beberapa jenis ANOVA, yang paling umum dan utama adalah:

1. ANOVA Satu Arah (One-way ANOVA)

Digunakan untuk menguji dan membandingkan perbedaan rata-rata dua atau lebih kelompok yang hanya memiliki satu variabel independent atau faktor.

2. ANOVA Dua Arah (Two-way ANOVA)

Digunakan untuk menguji dan membandingkan perbedaan rata-rata dua atau lebih kelompok yang memiliki dua variabel independent atau faktor.

D. SUMBER KERAGAMAN (VARIASI) ANOVA

Dalam analisis ANOVA, total variasi dalam data terbagi menjadi beberapa sumber keragaman (SK), yaitu:

3. Variasi antar kelompok (between groups)/ Perlakuan:

Variasi ini menunjukkan perbedaan rata-rata antar kelompok.

4. Variasi dalam kelompok (within groups)/ Error:

Variasi ini menunjukkan perbedaan individu dalam kelompok yang sama.

5. Variasi total

Merupakan jumlah variasi antar kelompok dan variasi dalam kelompok.

E. INTERPRETASI HASIL DAN PENGUJIAN HIPOTESIS DALAM ANOVA

Hasil ANOVA disajikan dalam tabel yang disebut tabel ANOVA. Tabel ini berisi nilai F, nilai P, dan efek ukuran. Nilai P menunjukkan probabilitas mendapatkan nilai F yang sama besar atau lebih besar dari nilai yang diamati jika hipotesis nol benar. Nilai efek ukuran menunjukkan besarnya efek variabel independen terhadap variabel dependen.

ANOVA menggunakan uji statistik F untuk menguji hipotesis tentang perbedaan rata-rata antar kelompok. Pengujian hipotesis dilakukan dengan membandingkan nilai F hitung (rasio variasi antar kelompok dengan variasi dalam kelompok) dengan nilai F kritis dari tabel distribusi F. Jika nilai F hitung lebih besar dari nilai F kritis, maka hipotesis nol (H_0) ditolak (tidak ada perbedaan rata-rata antara kelompok), yang berarti terdapat perbedaan signifikan secara statistik antara rata-rata kelompok dalam populasi.

F. PROSEDUR ANOVA

Secara umum, prosedur untuk melakukan analisis ANOVA adalah sebagai berikut:

1. Menentukan Hipotesis nol (H_0) dan Hipotesis alternatif (H_1).
Hipotesis nol / H_0 secara umum tidak terdapat perbedaan antara rata-rata kelompok dalam populasi (perlakuan tidak berpengaruh terhadap respon), sedangkan Hipotesis alternatif /

- H1 terdapat perbedaan antara rata-rata kelompok dalam populasi (perlakuan berpengaruh terhadap respon yang diukur).
2. Menghitung jumlah kuadrat-JK (*sum of squares*) untuk variasi antar kelompok (perlakuan)- JKP, variasi dalam kelompok (error) - JKE, dan variasi total - JKT.
 3. Menghitung derajat kebebasan-db (*degree of freedom*) tiap-tiap sumber keragaman / variasi (perlakuan, error, dan total)
 4. Menghitung rata-rata kuadrat/ kuadrat tengah - KT (*mean square*) dengan membagi jumlah kuadrat dengan derajat kebebasannya.
 5. Menghitung nilai F hitung dengan membagi rata-rata kuadrat antar kelompok (perlakuan) dengan rata-rata kuadrat dalam kelompok (error).
 6. Membandingkan nilai F hitung dengan nilai F kritis dan mengambil keputusan untuk menolak atau menerima hipotesis nol (H_0). Secara manual, jika probabilitas $< 5\%$ (0.05) atau selang kepercayaan 95%:
 - a. Signifikan: $F \text{ hitung} > F \text{ kritis} (0.05)$ maka tolak H_0 dan terima H_1 , sehingga terdapat perbedaan rata-rata antara kelompok atau perlakuan berpengaruh terhadap respon yang diukur.
 - b. Non Signifikan: $F \text{ hitung} < F \text{ kritis} (0.05)$ maka terima H_0 dan tolak H_1 , sehingga tidak terdapat perbedaan rata-rata antara kelompok atau perlakuan tidak berpengaruh terhadap respon yang diukur.

Contoh:

Anova pada rancangan acak lengkap (RAL) dengan persamaan linear:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

y_{ij} = respon parameter dari perlakuan ke-i dan ulangan ke-j

μ = rata-rata populasi

α_i = pengaruh perlakuan ke-i

ϵ_{ij} = pengaruh error / galat dari perlakuan ke-i dan ulangan ke-k

Selanjutnya, prosedur analisis ANOVA mencari jumlah kuadrat (JK) dengan rumus berikut (t=perlakuan; r=ulangan):

$$FK = (\text{Total data})^2 / \text{banyaknya data} = (\sum X_{..})^2 / n$$

$$JKT = (\text{Tiap data dikuadratkan lalu dijumlah semua}) - FK \\ = \sum (X_{..}^2) - FK$$

$$JKP = (\text{Total tiap perlakuan dikuadratkan lalu dijumlah semua}) / \text{ulangan} - FK = \sum (X_{t.}^2) / r - FK$$

$$JKE = JKT - JKP$$

Kemudian hasil perhitungan dimasukkan ke dalam table ANOVA

SK	db	JK	KT	Fhit	F _{0,05}
Perlakuan (P)	t-1	JKP	JKP/dbP	KTP/KTE	(0,05;dbP;dbE)
Error (E)	tx(r-1)	JKE	Jke/dbE		
Total	(txr)-1	JKT			

Tabel Distribusi F (α 0.05)

α (0.05)	db Perlakuan					
db Error	1	2	3	4	5	10
1	161.45	199.50	215.71	224.58	230.16	241.88
2	18.51	19.00	19.16	19.25	19.30	19.40
3	10.13	9.55	9.28	9.12	9.01	8.79
4	7.71	6.94	6.59	6.39	6.26	5.96
5	6.61	5.79	5.41	5.19	5.05	4.74
6	5.99	5.14	4.76	4.53	4.39	4.06
7	5.59	4.74	4.35	4.12	3.97	3.64
8	5.32	4.46	4.07	3.84	3.69	3.35
9	5.12	4.26	3.86	3.63	3.48	3.14
10	4.96	4.10	3.71	3.48	3.33	2.98
11	4.84	3.98	3.59	3.36	3.20	2.85
12	4.75	3.89	3.49	3.26	3.11	2.75
13	4.67	3.81	3.41	3.18	3.03	2.67
14	4.60	3.74	3.34	3.11	2.96	2.60
15	4.54	3.68	3.29	3.06	2.90	2.54
16	4.49	3.63	3.24	3.01	2.85	2.49
17	4.45	3.59	3.20	2.96	2.81	2.45
18	4.41	3.55	3.16	2.93	2.77	2.41
19	4.38	3.52	3.13	2.90	2.74	2.38
20	4.35	3.49	3.10	2.87	2.71	2.35
21	4.32	3.47	3.07	2.84	2.68	2.32
22	4.30	3.44	3.05	2.82	2.66	2.30
23	4.28	3.42	3.03	2.80	2.64	2.27
24	4.26	3.40	3.01	2.78	2.62	2.25
25	4.24	3.39	2.99	2.76	2.60	2.24
			...			
50	4.03	3.18	2.79	2.56	2.40	2.03

G. KELEBIHAN DAN KEKURANGAN ANOVA

ANOVA memiliki beberapa kelebihan, yaitu: mudah dipelajari dan dipahami, dapat digunakan untuk membandingkan perbedaan rata-rata dari lebih dari dua kelompok secara simultan, memiliki kekuatan uji yang lebih besar dibandingkan dengan melakukan beberapa uji t secara terpisah, dan dapat mendeteksi perbedaan yang signifikan meskipun ukuran sampel kecil.

Namun, ANOVA juga memiliki beberapa kelemahan, yaitu: memerlukan data yang memenuhi asumsi dasar, asumsi normalitas dan homogenitas varians harus terpenuhi, tidak dapat memberikan informasi tentang pasangan kelompok mana yang berbeda secara signifikan, dan sensitif terhadap outlier atau nilai-nilai ekstrem dalam data.

H. UJI LANJUT (POST HOC TESTS)

Jika hasil analisis ANOVA menunjukkan adanya perbedaan signifikan antara kelompok, maka perlu dilakukan uji lanjutan (post hoc tests) untuk mengetahui pasangan kelompok mana yang berbeda secara signifikan. Hasil uji lanjut dapat menunjukkan pada taraf perlakuan yang memiliki efek lebih baik pada respon. Beberapa uji lanjutan yang umum digunakan antara lain uji Fisher (LSD/BNT), uji Tukey (HSD/BNJ), uji Duncan (DMRT), uji Contrasts, dan uji Polinomial Ortogonal sesuai dengan sifat perlakuan.

1. PENGGUNAAN ANOVA DALAM PENELITIAN

ANOVA banyak digunakan dalam berbagai bidang penelitian, seperti:

1. Pertanian

Membandingkan rata-rata hasil produksi telur, susu, bobot badan ternak, konsumsi pakan, panen dari beberapa varietas tanaman atau perlakuan pupuk yang berbeda.

2. Biologi

Membandingkan rata-rata pertumbuhan bakteri atau sel dengan perlakuan yang berbeda, populasi mikroba, dan variasi gen pada makhluk hidup

3. Psikologi

Membandingkan rata-rata skor tes atau kuesioner antara kelompok yang berbeda, pencegahan, dan pengobatan gangguan mental.

4. Ilmu social dan ekonomi

Membandingkan rata-rata pendapatan, tingkat kepuasan, atau variabel lain antara kelompok yang berbeda.

Berikut adalah contoh penggunaan ANOVA:

1. Seorang peneliti ingin mengetahui apakah terdapat perbedaan tinggi dan bobot badan rata-rata antara mahasiswa yang mengikuti program latihan fisik selama 12 minggu dengan mahasiswa yang tidak mengikuti program tersebut. Peneliti dapat menggunakan ANOVA Satu Arah untuk menguji hipotesis ini.

2. Seorang peneliti ingin menguji efektivitas tiga jenis pupuk berbeda terhadap pertumbuhan tanaman sorghum. Peneliti melakukan percobaan dengan menanam sorghum di tiga petak tanah yang berbeda dan memberikan pupuk yang berbeda pada setiap petak. Setelah panen, peneliti mengukur tinggi tanaman sorghum di setiap petak. Peneliti dapat menggunakan ANOVA untuk menguji apakah terdapat perbedaan yang signifikan secara statistik dalam tinggi tanaman sorghum antara ketiga jenis pupuk. rata-rata dari dua atau lebih kelompok dalam sebuah populasi secara simultan.

J. APLIKASI UNTUK ANALISIS ANOVA

Terdapat beberapa aplikasi statistik yang dapat digunakan untuk analisis ANOVA, antara lain:

1. SPSS (Statistical Package for the Social Sciences)
2. R (bahasa pemrograman untuk analisis statistik)
3. SAS OnDemand for Academic (online)
4. Minitab
5. Microsoft Excel (dengan add-in Data Analysis Tool Pack)

Pemilihan aplikasi statistik tergantung pada kebutuhan penelitian, ketersediaan sumber daya, dan preferensi pengguna.

BAGIAN 14

PENGANTAR SPSS

A. PENGERTIAN DAN SEJARAH SPSS

1. Pengertian SPSS

SPSS adalah singkatan dari *Statistical Package for the Social Sciences*. Ini adalah perangkat lunak yang digunakan untuk analisis statistik. Awalnya dikembangkan oleh IBM Corporation, SPSS sekarang merupakan salah satu perangkat lunak analisis data yang paling populer di dunia, terutama di bidang penelitian sosial, bisnis, dan ilmu sosial lainnya.

SPSS memungkinkan pengguna untuk melakukan berbagai macam analisis statistik, mulai dari analisis deskriptif sederhana hingga analisis yang lebih kompleks seperti regresi, analisis varians, dan analisis faktor. Ini menyediakan antarmuka pengguna grafis yang mudah digunakan yang memungkinkan pengguna untuk mengimpor data dari berbagai sumber, melakukan analisis statistik, dan membuat visualisasi data dengan mudah.

Dalam SPSS merupakan program aplikasi yang mampu melakukan analisis statistik tingkat lanjut dan sistem pengelolaan data dalam lingkungan grafis dengan menu deskriptif dan kotak dialog sederhana untuk memudahkan penggunaannya. Banyak operasi yang dapat dilakukan hanya dengan menunjuk dan mengklik. SPSS

banyak digunakan dalam berbagai riset pasar, pengendalian kualitas dan peningkatan serta penelitian ilmiah. SPSS pertama kali dirilis dalam versi PC (dapat digunakan pada komputer desktop) dengan nama SPSS/PC+ (versi DOS). Namun popularitas sistem operasi Windows. SPSS akan merilis versi Windows (dari versi 6.0 hingga versi terbaru). Perangkat lunak SPSS dibuat dan dikembangkan oleh SPSS Inc., kemudian diakuisisi oleh IBM Corporation. Keunggulan software komputer ini adalah mudah digunakan untuk pemrosesan dan analisis data statistik. Fungsinya antara lain IBM SPSS Data Collection untuk pengumpulan data. Field, Andy (2013).

2. Sejarah SPSS

SPSS (Statistical Package for the Social Sciences) adalah perangkat lunak yang digunakan untuk analisis statistik. Awal mula SPSS pertama kali dikembangkan oleh Norman H. Nie, C. Hadlai "Tex" Hull, dan Dale H. Bent pada tahun 1968 di Universitas Stanford sebagai proyek riset. Awalnya, SPSS dirancang untuk membantu para peneliti dalam analisis data dalam ilmu sosial. Kemudian pada 1975 SPSS di komersialisasi dirilis sebagai perangkat lunak berbayar pertama yang tersedia untuk analisis statistik. Versi komersial pertama dikenal sebagai SPSS 4.0. Seiring berjalannya waktu, SPSS terus mengalami perkembangan dan peningkatan fitur. Hal ini memungkinkan pengguna untuk melakukan berbagai jenis analisis statistik, mulai dari analisis deskriptif sederhana hingga teknik analisis yang lebih kompleks seperti regresi, ANOVA, dan analisis faktor. Nie, Norman H., C. Hadlai Hull, dan Dale H. Bent (1970)

Akuisisi oleh IBM: Pada tahun 2009, IBM mengakuisisi SPSS Inc. seharga sekitar \$1,2 miliar. Akuisisi ini memperkuat portofolio perangkat lunak analisis data IBM dan meningkatkan integrasi antara SPSS dan produk-produk IBM lainnya. Integrasi dengan IBM: Setelah akuisisi, SPSS terus diperbarui dan diintegrasikan lebih dalam dengan produk-produk perangkat lunak IBM lainnya, seperti IBM Watson Analytics dan IBM Cognos Analytics. Integrasi ini membantu organisasi memanfaatkan analisis data yang lebih kuat dan komprehensif.

Perkembangan terbaru SPSS terus menjadi salah satu perangkat lunak analisis statistik yang paling populer di dunia, digunakan oleh organisasi di berbagai industri dan bidang penelitian. Perkembangan terbaru mencakup peningkatan fitur analisis, peningkatan keamanan data, dan peningkatan integrasi dengan teknologi baru seperti komputasi awan dan kecerdasan buatan. Sejak awal mula pengembangannya hingga menjadi produk yang diakuisisi oleh IBM, SPSS telah memainkan peran penting dalam analisis data di berbagai bidang, mulai dari ilmu sosial dan kesehatan hingga bisnis dan pemasaran.

B. PENELITIAN STATISTIK DAN SPSS

Program SPSS merupakan software yang digunakan pada kegiatan statistik mulai dari pengolahan data hingga analisis

data. Penggunaan SPSS memerlukan pengetahuan tentang beberapa elemen statistik yang berkaitan dengan statistik, antara lain:

1. Populasi

Populasi secara umum dapat diartikan sebagai kumpulan data yang mengidentifikasi suatu fenomena. Statistik populasi tidak terbatas pada permasalahan orang atau perusahaan, namun mempunyai cakupan yang lebih luas, misalnya populasi ayam pada suatu daerah, populasi bakteri “y” pada roti, dan lain-lain. Populasinya bisa sangat besar hingga dikatakan tidak terbatas, seperti populasi oksigen dunia, populasi plankton di lautan, dan lain-lain. Barbie and Earl R. (2016)

2. Sampel

Sampel didefinisikan sebagai bagian dari suatu populasi atau sekumpulan data yang diambil atau dipilih dari suatu populasi. Pada dasarnya suatu populasi adalah bagian dari suatu populasi, atau suatu populasi dapat dibagi menjadi berbagai jenis sampel. Fowler, Floyd J. (2013)

3. Variabel

Saat mendeskripsikan suatu populasi, tidak perlu mengetahui semua properti populasi. Untuk mendeskripsikan suatu populasi, hanya satu atau lebih properti yang disebut variabel yang perlu diketahui. Neumen, W. Lawrence (2013)

C. PENYUSUNAN DATA

1. Persiapan Data

Langkah awal yang dilakukan dalam pengolahan data maka terlebih dahulu dengan mempersiapkan variabel penelitian dan data-data yang akan diinput kedalam program SPSS. Setelah data-data yang akan diinput telah siap untuk diolah maka ada aturan yang harus diikuti seperti : tipe variabel (*numeric, date, string* dan sebagainya) serta ukuran data yang digunakan.

2. Ukuran Data

Ukuran data sangat penting diketahui sebab akan menentukan analisis apa yang akan digunakan. Pada SPSS ada tiga ukuran data, yaitu: nominal, ordinal dan scale.

- ❖ Nominal adalah data yang diperoleh dari hasil kategorisasi dan menunjukkan kesejajaran dengan menggunakan jenis kelamin. Misalnya: 1 = pria dan 2 = wanita.
- ❖ Ordinal adalah data yang diperoleh dari kategorisasi dan mempunyai peringkat dengan skala likert Misalnya: 1 = sangat tidak setuju, 2 = tidak setuju, 3 = cukup setuju, 4 = setuju 5 = sangat setuju.
- ❖ Scale adalah data yang didapatkan bukan dari kategorisasi, data ini berupa data interval dan rasio. Pada data ini dapat dilakukan operasi matematis (penjumlahan, pengurangan, perkalian dan pembagian). Misalnya: data pendapatan, data tinggi badan, nilai siswa, biaya produksi, penjualan, dan sebagainya.

3. Variabel Penelitian

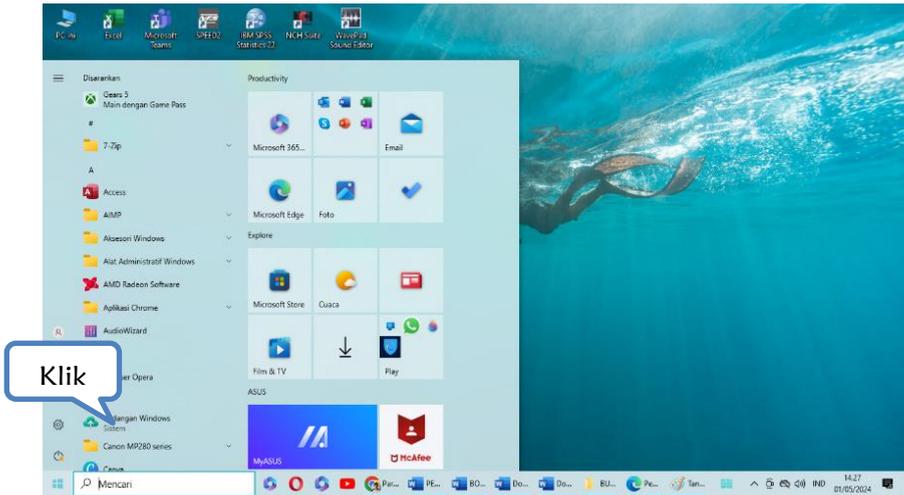
Dalam sebuah penelitian variabel merupakan konsep yang nilainya dapat bervariasi dan berubah-ubah. Ada beberapa jenis variabel yang biasa digunakan sebagai berikut :

- ❖ Variabel bebas (*independen*) yaitu variabel yang nilainya mampu mempengaruhi variabel terikat. Seperti kompetensi, motivasi, lingkungan kerja, budaya kerja, kualitas pelayanan, pelatihan, gaya kepemimpinan dan lain-lain
- ❖ Variabel terikat (*dependen*) yaitu variabel yang nilainya dipengaruhi oleh variabel bebas. Seperti kepuasan kerja, kinerja, keputusan pembelian, harga saham, hasil belajar, tingkat penjualan dan lain-lain.
- ❖ Variabel moderasi yaitu variabel yang mempengaruhi hubungan antara variabel independen dengan variabel dependen seperti : motivasi, Customer experience, loyalitas dan lain-lain.
- ❖ Variabel kontrol yaitu variabel yang dikendalikan atau variabel yang nilainya dibuat tetap, hal ini agar tidak dipengaruhi oleh variabel lain

4. Cara Membuka Program SPSS

Cara pertama :

Klik start > all program > IBM SPSS Statistik > IBM SPSS Statistik



Cara Kedua: klik ganda icon IBM SPSS Statistics di dekstop



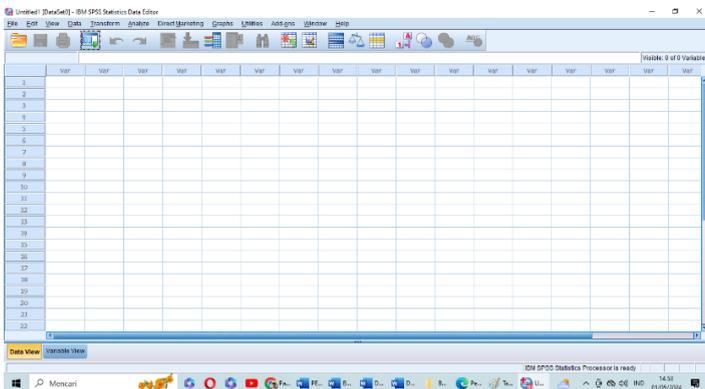
Maka akan muncul tampilan IBM SPSS sebagai Berikut :



Selanjutnya muncul tampilan IBM SPSS Statistics Viewer berikut:

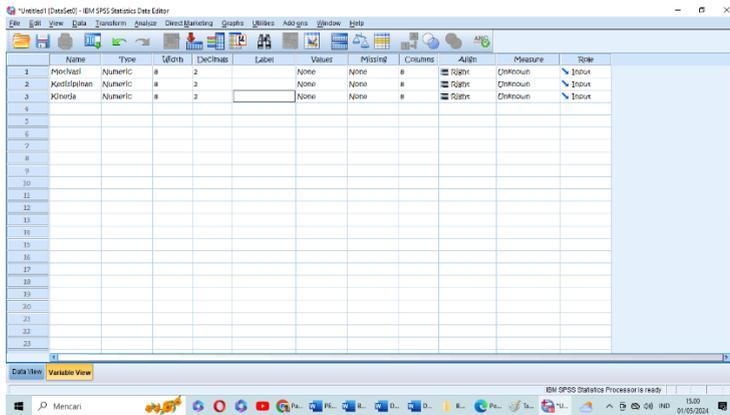


Kemudian setelah itu tampilan area kerja SPSS seperti dibawah ini :



5. Memasukkan Data

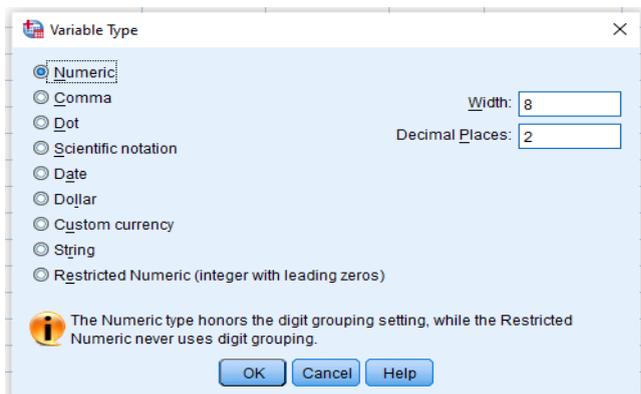
Langkah 1 : menuliskan variabel pada halaman variabel views



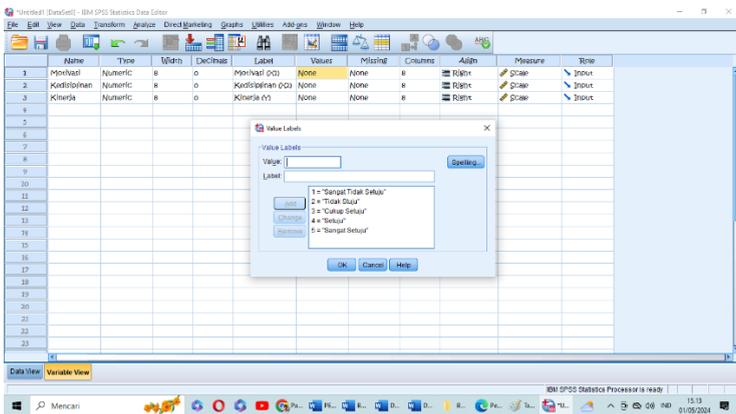
Keterangan:

- ❖ Name : nama variabel
- ❖ Type : menentukan tipe variabel.

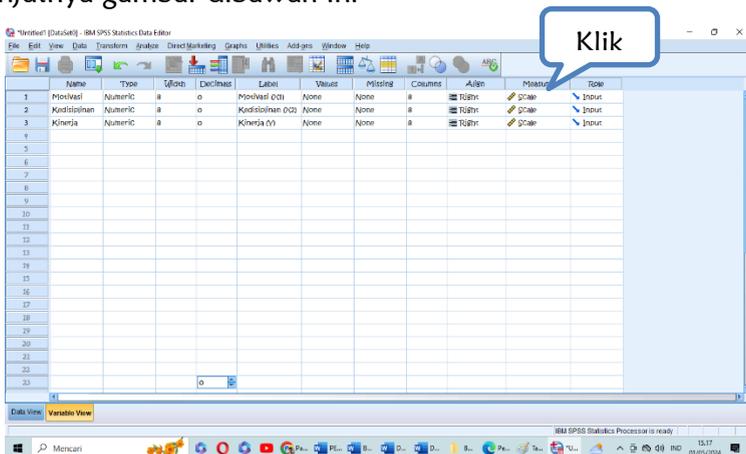
Pada IBM SPSS Statistics terdapat 9 variabel, yang paling banyak digunakan yaitu numeric (data berupa angka) dan string (bukan angka). Maka pilihan Type yang digunakan tetap numeric. Seperti gambar dibawah ini :



- ❖ Width : ukuran menentukan lebar kolom secara default adalah 8.
- ❖ Decimals : menentukan banyaknya angka di belakang koma.
- ❖ Label : memberikan penjelasan nama variabel.
- ❖ Values : memberi value (nilai) untuk data tipe nominal atau ordinal. Misal: 1 = Sangat Tidak Setuju, 2 = Tidak Setuju 3 = Cukup Setuju 4 = Setuju 5 = Sangat Setuju.



Selanjutnya gambar dibawah ini



- ❖ Decimals : menentukan angka dibelakang koma, pilih 0.
- ❖ Label : menuliskan variabel dengan X1, X2,....dan Y
- ❖ Columns : menentukan lebar kolom pada kolom data pada halaman Data View. (Tidak Perlu diubah)
- ❖ Align : menentukan letak data left, Center dan Right (pilih Right)
- ❖ Measures : menentukan ukuran data : nominal, ordinal dan scale (Pilih Scale).

Langkah 2 : Penginputan data pada halaman view

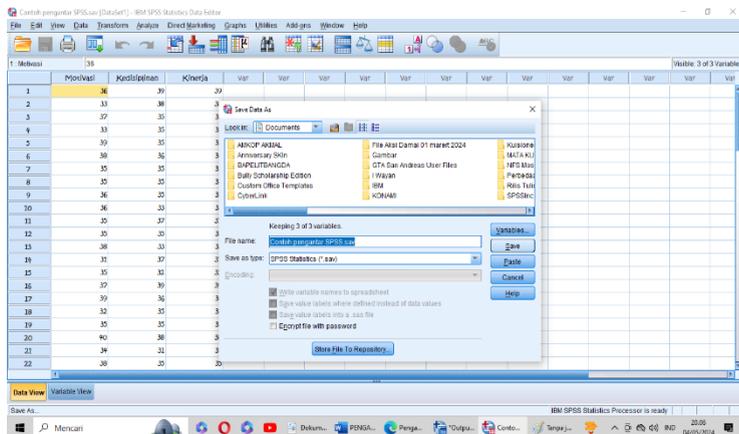
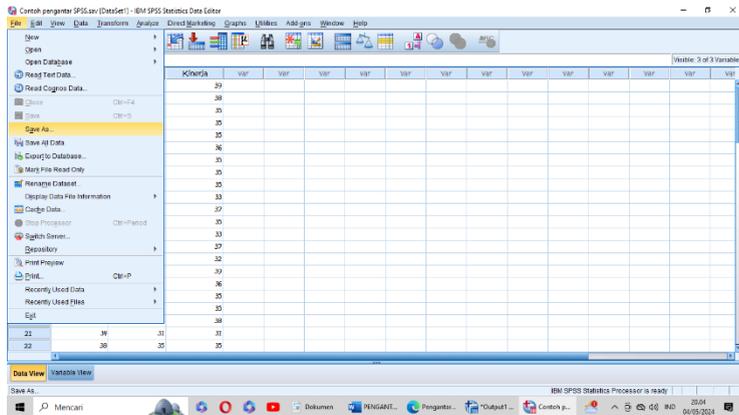
	MonKwasi	KondisiJalan	Kinerja	Var											
1	36	39	39												
2	33	38	38												
3	37	35	35												
4	33	35	35												
5	39	35	35												
6	38	36	36												
7	35	35	35												
8	35	35	35												
9	36	35	35												
10	36	33	33												
11	35	37	37												
12	35	35	35												
13	38	33	33												
14	31	37	37												
15	35	32	32												
16	37	39	39												
17	39	36	36												
18	32	35	35												
19	35	35	35												
20	38	38	38												
21	34	31	31												
22	38	35	35												

6. Menyimpan Data

Langkah selanjutnya setelah penginputan pada data view maka selanjutnya data tersebut kita simpan dengan langkah-langkah sebagai berikut:

- ❖ Klik File > Save As, kemudian terbuka kotak dialog Save Data As.
- ❖ pada Save In, pilih direktori/folder untuk menyimpan data anda. Ketikan : Membangun Data pada SPSS di File Name,

sialhkan ketik nama file yang akan disimpan selanjutnya klik Save.

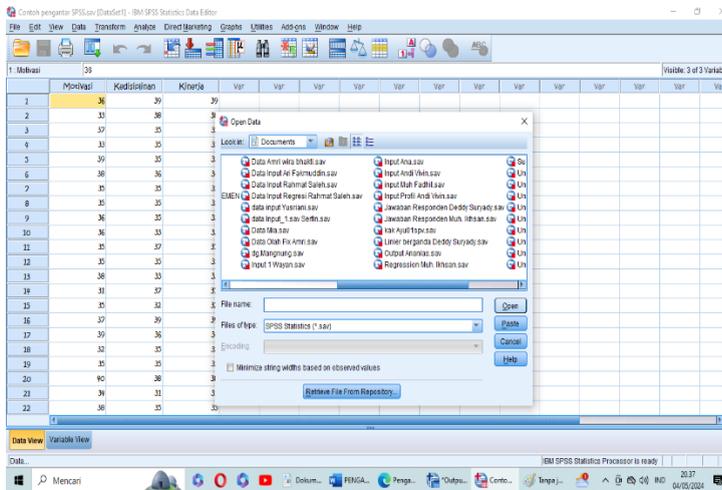
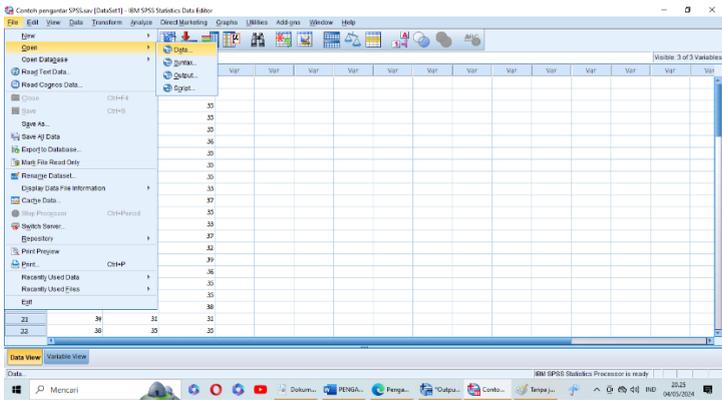


7. Membuka File Data Save

Untuk membuka data yang telah disimpan sebelumnya maka langkah-langkah yang dilakukan adalah sebagai berikut :

- ❖ Terlebih dahulu silahkan buka Software Aplikasi SPSS kemudian Klik File > Open > Data > Kemudian akan terbuka kotak dialog Open Data. Pada kolom Log In > silahkan memilih judul

yang telah di simpan sebelumnya > kemudian silahkan klik Open



DAFTAR PUSTAKA

- Aan Komariah, Djam'an Satori. 2011. *Metode Penelitian Kualitatif*. Bandung: Alfabeta.
- Algifari. 2016. *Ekonometrika*. Cetakan ke 1. Yogyakarta: BPFE.
- Ananda, R., & Fadhli, M. 2018. *Statistik Pendidikan : teori dan praktik dalam pendidikan*. Medan: Widya Puspita.
- Anas Sudijono. 2004. *Pengantar Statistik Pendidikan*. Jakarta: Raja Grafindo Persada
- Babbie, Earl R. (2016) "The Practice of Social Research." Cengage Learning.
- Brown, E., Wilson, S. and Garcia, M. (2020a) 'Challenges in Analyzing Qualitative Data', *Journal of Methodological Studies*, 25(1), pp. 120–135.
- Brown, E., Wilson, S. and Garcia, M. (2020b) 'Challenges in Analyzing Quantitative Data', *Journal of Methodological Studies*, 25(1), pp. 120–135.
- Brown, E., Wilson, S. and Garcia, M. (2020c) 'Challenges in Sample Representativeness in Surveys', *Journal of Methodological Studies*, 25(1), pp. 120–135.
- Brown, E., Wilson, S. and Garcia, M. (2020d) 'Challenges in the Quality of Secondary Data', *Journal of Methodological Studies*, 25(1), pp. 120–135.
- Brown, E., Wilson, S. and Garcia, M. (2020e) 'Limitations and Applications of Ordinal Scale Data', *Journal of Methodological Studies*, 25(1), pp. 120–135.

- Brown, E., Wilson, S. and Garcia, M. (2020f) 'Limitations of Nominal Scale Data', *Journal of Methodological Studies*, 25(1), pp. 120–135.
- Chen, Y., Wang, H. and Liu, M. (2023) 'The Potential of Big Data in Providing Diverse and Extensive Primary Data', *International Journal of Big Data Research*, 5(1), pp. 78–92.
- Chin, W. W. (1998). *The Partial Least Squares Approach to Structural Equation Modeling*. *Modern Methods for Business Research*, 295, 336
- Creswell, John. 2015. *Penelitian Kualitatif & Desain Riset*. Yogyakarta : Pustaka Pelajar.
- Dajan, Anto. (1995). *Pengantar Metode Statistik Jilid I*. Jakarta: LP3S.
- David R. Anderson, Dennis J. Sweeney, Thomas A. Williams (2020). *Statistics for Business and Economics*, 14th Edition. Boston: Cengage Learning
- David R. Anderson, Dennis J. Sweeney, Thomas A. Williams (2020). *Statistics for Business and Economics*, 14th Edition. Boston: Cengage Learning
- Devore, J. L. 2011. *Probability and Statistics for Engineering and the Sciences* (8th ed.). Brooks/Cole.
- Diez, D. M., Barr, C. D., & Cetinkaya-Rundel, M. 2015. *OpenIntro Statistics* (3rd ed.). OpenIntro, Inc.
- Douglas Lind , William Marchal, Samuel Wathen (2020). *Statistical Techniques in Business and Economics* (18th Edition). McGraw-Hill Education

- Douglas Lind , William Marchal, Samuel Wathen (2020). *Statistical Techniques in Business and Economics (18th Edition)*. McGraw-Hill Education
- Draper, Norman. Smith, Harry.1981. *Applied Regression Analysis*, ed 2. New York: John Wiley & Sons, Inc
- Fowler, Floyd J. (2013) "Survey Research Methods." SAGE Publications.
- Ghozali, I. (2016) *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23*. Edisi 8. Semarang: Badan Penerbit Universitas Diponegoro.
- Ghozali, Imam. 2018. *Aplikasi Analisis Multivariate dengan Program IBM SPSS*. 25. Badan Penerbit Universitas Diponegoro: Semarang.
- Gomez, Kwanchai A. dan Gomez, Arturo A. 1984. *Statistical Procedures for Agricultural Research*. 2nd Ed. Canada: John Wiley & Sons, Inc.
- Grinstead, C. M., & Snell, J. L. 1997. *Introduction to Probability (2nd ed.)*. American Mathematical Society.
- Gujarati, N.D. 2016. *Basic Econometrics*. 4th ed. New York: McGraw-Hill Companies, Inc.
- Gunawan,Imam. 2016. *Pengantar Statistik Inferensial*.Raja Grafindo Persada. Jakarta
- Hadi Sutrisno. 2002. *Metodologi Riset*. Yogyakarta: Andi Offset
- Hajarisman, Nusar. 2014. *Analisis Regresi Lanjut*. Bandung: P2U LPPM UNISBA

- Hartanto, D., & Sri Yuliani. 2019. *Statistik Riset Pendidikan; Dilengkapi Analisis SPSS*. In Cahaya Firdaus Publishing and Printing. Cahaya Firdaus
- Hull, C. Hadlai, Norman H. Nie, dan Dale H. Bent. (1975) "Development of the Generalized Model." *SPSS: A Statistical Package for the Social Sciences*, edited by Norman H. Nie, C. Hadlai Hull, dan Dale H. Bent, McGraw-Hill, pp. 35–74
- Irianto, Agus. 2015. *Statistika Konsep dasar, Aplikasi, dan Pengembangannya*, Edisi Keempat. Jakarta: PRENADAMEDIA GRUP
- Jainuri M, (2019). *Pengantar Aplikasi Komputer SPSS*, Edisi Revisi. Web Hira Institute.
- Jones, A., Brown, K. and Davis, C. (2019a) 'Accessibility of Secondary Data', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019b) 'Accessing Databases for Secondary Data', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019c) 'Advantages of Nominal Scale Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019d) 'Advantages of Panel Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019e) 'Advantages of Qualitative Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.

- Jones, A., Brown, K. and Davis, C. (2019f) 'Advantages of Quantitative Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019g) 'Advantages of Ratio Scale Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019h) 'Advantages of Time Series Data Analysis', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Jones, A., Brown, K. and Davis, C. (2019i) 'Specificity and Detail in Survey Data', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Kuncoro, Haryo. 2018. *Statistika Deskriptif untuk Analisis Ekonomi*. Jakarta: Bumi Aksara
- Kutner, M.H., C.J. Nachtsheim., dan J. Neter. 2004. *Applied Linear Regression Models*. 4th New York: McGraw-Hill Companies, Inc.
- Laka, L. (2023). *Metodologi penelitian dengan pendekatan kuantitatif*. Jilid 2. Yogyakarta: Deepublish.
- Leavy, P. (2017). *Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches*. New York, NY: The Guilford Press.
- Li, X., Xu, J. and Zhang, L. (2020) 'Accuracy Evaluation of Census Data', *Journal of Data Science*, 18(3), pp. 503–518.
- Li, X., Zhang, Q. and Wu, Y. (2022a) 'Managing Random Fluctuations in Time Series Data', *Journal of Data Analytics*, 8(2), pp. 215–230.

- Li, X., Zhang, Q. and Wu, Y. (2022b) 'Real-Time Insights from Big Data', *Journal of Data Analytics*, 8(2), pp. 215–230.
- Lind, A. Dauglass, William G. Marchal and Robert D. Mason, (2002).
- Lind, Douglas A., Marchal, William G., and Wathen, Samuel A.. 2012. *Statistical Techniques in Business & Economics*. Fifteenth Edition. New York: McGraw-Hill
- M. Husni Arifin, 2012. *Konsep-konsep Dasar Statistika*, Penerbit BPFE Yogyakarta.
- Mendenhall, W., Beaver, R. J., & Beaver, B. M. 2012. *Introduction to Probability and Statistics (14th ed.)*. Cengage Learning.
- Modern, Jakarta: Salemba Empat.
- Montgomery, Douglas C. 2020. *Design and Analysis of Experiments*. 10th Ed. New Jersey; Wiley.
- Montgomery, Douglas C. Peck, Elizabeth A. Vining, G.G. 2012. *Introduction to Linear Regression Analysis*. New York: John Wiley & Sons, Inc
- Moore, D. S., Notz, W. I, & Flinger, M. A. 2013. *The Basic Practice of Statistics*. Sixth edition. New York: W. H. Freeman and Company
- Myers, Raymond. 1986. *Classical and Modern Regression with Applications*, Second Edition. Boston: PWS Publishers
- Neuman, W. Lawrence. (2013) "Social Research Methods: Qualitative and Quantitative Approaches." Pearson.
- Nie, Norman H., C. Hadlai Hull, dan Dale H. Bent. (1970) "SPSS: Statistical Package for the Social Sciences." McGraw-Hill.

- ones, A., Brown, K. and Davis, C. (2019) 'Accessibility and Diversity of Secondary Data', *International Journal of Research Methods*, 15(3), pp. 401–415.
- Petrie, A. dan Watson, P. 2013. *Statistics for Veterinary and Animal Science*. 3rd Ed. London: Wiley-Blackwell.
- Purwanto, Suharyadi, 2018, *Statistika untuk Ekonomi dan Keuangan*
- Riduwan, Akdon. 2010. *Rumus dan Data dalam Analisis Data Statistika*. Bandung: Alfabeta.
- Roflin, E., Rohana, dan Riana, F., 2022. *Analisis Korelasi dan Regresi*. Pekalongan: Nasya Expanding Management
- Ross, S. M. 2014. *A First Course in Probability* (9th ed.). Pearson.
- Rosyadi, S. 2018. *Revolusi industri 4.0: Peluang dan Tantangan Bagi Alumni Universitas Terbuka*. Fakultas Ilmu Sosial Dan Ilmu Politik.
- Santoso, S. 2000. *Buku Latihan SPSS Statistik Parametrik*. Jakarta: Elex Media Komputindo.
- Sekaran, Bougie. 2016. *Metode Penelitian untuk Bisnis: Pendekatan Membangun Keterampilan*. Edisi ke-7. Wiley & Sons: Sussex Barat.
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. (7th Ed.). Chichester, West Sussex, United Kingdom: John Wiley & Sons.
- Sembiring, R.K. 2003. *Analisis Regresi*. Edisi Kedua. Bandung: Institut Teknologi Bandung.
- Singh, Y.K. (2006). *Fundamental of Research Methodology and Statistics*. New Delhi: New Age International Publishers.

- Smartstat. 2022. Asumsi-asumsi Anova Satu Faktor. <https://www.smartstat.info/materi/rancangan-percobaan/asumsi-analisis-variabel/asumsi-asumsi-anova-satu-faktor.html>
- Smith, J., Johnson, M. and Williams, R. (2021a) 'Secondary Data as a Source in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021b) 'The Effectiveness of Surveys in Data Collection', *Journal of Survey Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021c) 'The Role of Cross Sectional Data in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021d) 'The Role of Interval Scale Data in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021e) 'The Role of Nominal Scale Data in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021f) 'The Role of Ordinal Scale Data in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Smith, J., Johnson, M. and Williams, R. (2021g) 'The Role of Quantitative Data in Statistics', *Journal of Statistical Research*, 10(2), pp. 215–230.
- Statistical Techniques in Business & Economics. McGraw-Hill Irwin. Hal, 1-20

- Sudjana. 2000. Metode Statistika. Bandung: Gramedia Pustaka Utama.
- Sugiyono. 2017. Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung : Alfabeta, CV.
- Sugiyono. 2018. Metode Penelitian Kuantitatif, Kualitatif, dan R&D, penerbit. Alfabeta, Bandung.
- Sugiyono. 2020. Metode Penelitian Kuantitatif, Kualitatif dan R&D. Bandung: Alfabeta.
- Supardi. 2013. Aplikasi Statistika dalam Penelitian Konsep Statistika yang Lebih Komprehensif. Jakarta: Change Publication.
- Supardi. 2014. Aplikasi Statistik dalam Penelitian. Smart. Jakarta
- Wackerly, D., Mendenhall, W., & Scheaffer, R. L. 2014. Mathematical Statistics with Applications (7th ed.). Cengage Learning.
- Walpole, Ronald E. Myers, Raymond H. 1995. Ilmu peluang dan statistika untuk insinyur dan ilmuan Edisi-4. Bandung: Penerbit ITB Bandung
- Walpole, Ronald E., Myers, R. H., Myers, L., and Ye, Keying. 2002. Probability & Statistics for Engineers & Scientists. Ninth Edition. New Jersey: Prentice-Hall
- Wang, H., Wu, Y. and Liu, Y. (2019) 'The Role of Census in Data Collection', International Journal of Population Research [Preprint].
- Wang, H., Xu, J. and Liu, M. (2021) 'Methods of Collecting Secondary Data', Journal of Data Collection Methods, 8(2), pp. 215–230.

- Wang, L., Xu, J. and Brown, E. (2021a) 'Challenges in Managing and Analyzing Complex Big Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Wang, L., Xu, J. and Brown, E. (2021b) 'Ensuring Consistency and Quality of Panel Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Wang, L., Xu, J. and Brown, E. (2021c) 'Ensuring Representativeness of Cross Sectional Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Wang, L., Xu, J. and Brown, E. (2021d) 'Ensuring Validity of Qualitative Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Wang, L., Xu, J. and Brown, E. (2021e) 'Understanding the Interval Scale Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Wang, L., Xu, J. and Brown, E. (2021f) 'Understanding the Ratio Scale Data', *Journal of Big Data Management*, 15(3), pp. 401–415.
- Widarjono, A. 2007. *Ekonometrika: Teori dan Aplikasi untuk Ekonomi dan Bisnis*. Edisi Kedua. Yogyakarta: Ekonisia Fakultas Ekonomi Universitas Islam Indonesia.
- Zhang, Y. et al. (2020) 'The Impact of Artificial Intelligence and Blockchain on the Accounting Profession', *IEEE Access*, 8. Available at: <https://doi.org/10.1109/ACCESS.2020.3000505>.

TENTANG PENULIS

Penulis Bagian 1



Dr. I Wayan Terimajaya, SE., MM., MH

Seorang Penulis dan Dosen Prodi Kewirausahaan Fakultas Ekonomi dan Bisnis Universitas Tabanan, Bali. Lahir di Badung, 10 Pebruari 1967 Badung. Penulis merupakan anak pertama dari empat bersaudara dari pasangan bapak I Ketut Sugi (Alm) dan Ibu Ni Made Sulatri (Alm). Menamatkan pendidikan program Sarjana (S1) di Universitas Tabanan prodi Ekonomi Pembangunan dan menyelesaikan program Pasca Sarjana (S2) di Universitas Udayana, Denpasar prodi Magister Manajemen di bidang Manajemen Keuangan. Menyelesaikan program Pasca Sarjana (S2) di Universitas Mahendradatta, Denpasar prodi Magister Hukum di bidang Ilmu Hukum. Selanjutnya menyelesaikan program Pasca Sarjana (S3) di Universitas Hindu Indonesia, Denpasar program (Doktor) prodi Manajemen Pendidikan. Pengalaman mengajar selain di Universitas Tabanan juga mengajar di Universitas Dhyana Pura, Universitas Tri Atmamulya, dan Universitas Hindu Indonesia. Penulis juga sebagai Kelompok Ahli Kabupaten Tabanan.

Penulis Bagian 2



Ni Luh Sintya Dewi, S.Pd., M.Pd

Lahir di Sempidi pada tanggal 29 September 1993. Penulis merupakan putri satu-satunya dari pasangan Bapak I Nyoman Sumber, SH dan Ibu Ni Made Murda, S.Pd, AUD. Penulis menyelesaikan pendidikan S1 Prodi Pendidikan Matematika pada tahun 2015 dan menyelesaikan pendidikan S2 Prodi Pendidikan

Matematika pada tahun 2019. Sekarang ini penulis merupakan Dosen Tetap di Universitas Tabanan di Prodi Agribisnis Fakultas Sains dan Teknologi. Selain mengajar, juga aktif dalam kegiatan ilmiah seperti penelitian dan pengabdian kepada masyarakat.

Penulis Bagian 3



Dr. Ture Simamora, S.Pt., M.Si

seorang Penulis dan Dosen Prodi Peternakan Universitas Timor, Kefamenanu. Lahir di desa Lumban Toruan, 12 Oktober 1989 Sumut. Penulis merupakan anak ketiga dari lima bersaudara dari pasangan bapak Santun Simamora dan Ibu Bunga Purba. Ia menamatkan pendidikan program Sarjana (S1) di Universitas Andalas Padang prodi Peternakan dan menyelesaikan program Pasca Sarjana (S2) di Institut Pertanian Bogor prodi Ilmu Produksi dan Teknologi Peternakan. Ia menyelesaikan program Pasca Sarjana (S3) di Institut Pertanian Bogor Prodi Ilmu Penyuluhan Pembangunan Konsentrasi Penyuluhan Peternakan. Saat ini Penulis menjabat sebagai Koordinator Program Studi Peternakan Fakultas Pertanian, Sains dan Kesehatan Universitas Timor.

Penulis Bagian 4



Loso Judijanto, SSI, MM, MStats

adalah peneliti yang bekerja pada lembaga penelitian **IPOSS Jakarta**. Penulis dilahirkan di Magetan pada tanggal 19 Januari 1971. Penulis menamatkan pendidikan *Master of Statistics* di *the University of New South Wales, Sydney, Australia* pada tahun 1998 dengan dukungan beasiswa ADCOS (*Australian Development Cooperation Scholarship*) dari Australia.

Sebelumnya penulis menyelesaikan Magister Manajemen di Universitas Indonesia pada tahun 1995 dengan dukungan beasiswa dari Bank Internasional Indonesia. Pendidikan sarjana diselesaikan di Institut Pertanian Bogor pada Jurusan Statistika – FMIPA pada tahun 1993 dengan dukungan beasiswa dari KPS-Pertamina. Penulis menamatkan Pendidikan dasar hingga SMA di Maospati, Sepanjang karirnya, Penulis pernah ditugaskan untuk menjadi anggota Dewan Komisaris dan/atau Komite Audit pada beberapa perusahaan/lembaga yang bergerak di berbagai sektor antara lain pengelolaan pelabuhan laut, telekomunikasi seluler, perbankan, pengembangan infrastruktur, sekuritas, pembiayaan infrastruktur, perkebunan, pertambangan batu bara, properti dan rekreasi, dan pengelolaan dana perkebunan. Penulis memiliki minat dalam riset di bidang kebijakan publik, ekonomi, keuangan, *human capital*, dan *corporate governance*. Penulis dapat dihubungi melalui e-mail di: losojudijantobumn@gmail.com.

Penulis Bagian 5



Radha Krisnamurti Sigamura, M.Pd

seorang Penulis dan Dosen Fakultas Teknik dan Fakultas Pertanian di Universitas Musi Rawas. Lahir di desa Tahah Pilih, 18 April 1994 Sumatra Selatan. Penulis merupakan anak pertama dari 4 bersaudara dari pasangan bapak Edi Elfision dan Ibu Suarti. ia menamatkan pendidikan program Sarjana (S1) di Universitas PGRI Palembang prodi Pendidikan Matematika dan menyelesaikan program Pasca Sarjana (S2) di Universitas Indraprasta PGRI Jakarta prodi

Pendidikan Matematika dan Ilmu Pengetahuan Alam.

Penulis Bagian 6



Dr. Nurhayati, SE, ME

Merupakan dosen tetap di Program Studi Ekonomi Pembangunan Fakultas Ekonomi dan Bisnis Universitas Trisakti Jakarta. Lulus dari Program Sarjana (S1) Program Studi Ilmu Ekonomi dan Sudi Pembangunan Universitas Trisakti dan menyelesaikan program Pasca Sarjana (S2) di Program Magister Perencanaan dan Kebijakan Publik Universitas Indonesia dan melanjutkan Pasca Sarjana (S3) di Program Kebijakan Publik Universitas Trisakti.

Pengalaman mengajar Statistika, Ekonometrika dan Praktikum Alat Analisis Kuantitatif. Banyak menulis artikel di bidang Ekonomi, Regional, dan Pembangunan Berkelanjutan. Penulis aktif sebagai pengurus Jurnal sebagai Managing Editor pada Jurnal Media Ekonomi. Penulis juga aktif sebagai Ketua Lembaga Pengolahan Data dan Statistik di Fakultas Ekonomi dan Bisnis Universitas Trisakti.

Penulis Bagian 7



Sri Yani Kusumastuti

Dosen tetap di Program Studi Ekonomi Pembangunan Fakultas Ekonomi dan Bisnis Universitas Trisakti Jakarta. Lulus dari Program Sarjana (S1) Program Studi Ilmu Ekonomi dan Sudi Pembangunan Universitas Gadjah Mada dan menyelesaikan program Pasca Sarjana (S2) di Program Magister Sains Ilmu Ekonomi Universitas Gadjah Mada. Pengalaman mengajar Mikroekonomika, Statistika, dan Ekonometrika. Banyak menulis artikel di

bidang Ekonomi, Keuangan, dan Perbankan, Menjadi Editor di

beberapa buku, antara lain: *Kinerja perdagangan luar negeri Indonesia pada masa krisis: suatu kajian empiris (2002)*, Desain eksperimental: aplikasi dalam riset ilmu ekonomi, manajemen dan akuntansi (2011). Penulis buku: Lembaga Jasa Keuangan di Indonesia (2018).

Penulis Bagian 8



Dr. Raymond Bahana, S.T., M.Sc.

Dosen di Computer Science Department, School of Computing & Creative Arts, Binus University sejak tahun 2012. Memperoleh gelar Sarjana Teknik di bidang Teknik Elektro dengan peminatan Sistem Komputer dari UNIKA Atma Jaya, Jakarta pada tahun 1995. Pada tahun 2002 meraih gelar Master of Science di bidang Computer Engineering dari University of Central Florida di Orlando, USA. Gelar doktoral diperoleh pada tahun 2023 di bidang Computer Science dari Binus University, Jakarta. Di bidang penelitian, banyak melakukan riset di topik serious game dengan dosen-dosen dari berbagai jurusan maupun dosen-dosen dari universitas dalam dan luar negeri serta dengan pengembang boardgame dan digital game di Indonesia.

Penulis Bagian 9



Dr. Laurensius Laka, M.Psi., Psikolog, lahir di Sintang, Kalbar. Lulus SMA tahun 1987, selanjutnya menjadi abdi negara di jajaran TNI-AD. Di tengah pengabdian, muncul semacam kontemplasi mengenai keeratan hubungan antara dunia pendidikan dan kemajuan bangsa. Bertolak dari kontemplasi tersebut, menghantarkan penulis menempuh S-1 Psikologi

di Universitas Wisnuwardhana Malang, lulus 2006; S-2 Magister Profesi Psikologi di Universitas Airlangga, lulus 2009; dan S-3 Ilmu Psikologi juga di Universitas Airlangga, lulus 2015. Beberapa buku yang dipublikasikan, “Bunga Rampai Tantangan Merdeka Belajar Kampus Merdeka di Era Industri 4.0 (2022), Metodologi Penelitian dengan Pendekatan Kuantitatif, Jilid 1 (2022), Kesehatan Mental Masyarakat (*book chapter*, 2023), Metodologi Penelitian dengan Pendekatan Kuantitatif, Jilid 2 (2023), editor dan penulis buku ‘*Goal Orientation dan Self-Efficacy dalam Self-Regulated Learning*’ (*book chapter*, 2023), editor dan penulis buku ‘*Psikologi Pendidikan: Teori dan Aplikasi*’ (*book chapter*, 2023), Pengembangan Bahan Ajar (*book chapter*, 2023), Metodologi Penelitian Kualitatif (*book chapter*, 2023), Metodologi dan Teknik Penulisan Ilmiah (*book chapter*, 2024), Panduan Praktis Penulisan Karya Tulis Ilmiah (*book chapter*, 2024), Metode & Model Pembelajaran Inovatif (*book chapter*, 2024), dan Pendidikan Karakter Gen Z di Era Digital (*book chapter*, 2024). Saat ini penulis selaku Kaprodi program magister Prodi Pastoral STP-IPI Malang. Selain itu juga berbagi pengetahuan dalam bidang psikologi di Fakultas Psikologi Universitas Yudharta Pasuruan, sebuah Perguruan Tinggi yang terkenal dengan jargon *The Multicultural University*, dalam naungan Pondok Pesantren Ngalah – Pasuruan. Penulis dapat dihubungi via e-mail: laurensiuslaka78@gmail.com

Penulis Bagian 10



Anindita Henindya Permatasari, S.Si, M.Mat.

Lahir di Semarang, 23 Mei 1993. Lulus S1 Matematika Universitas Diponegoro tahun 2015. Lulus S2 Matematika Universitas Diponegoro tahun 2018. Sejak tahun 2019 diterima menjadi dosen tetap Departemen Matematika Fakultas Sains dan Matematika Universitas Diponegoro Semarang. Penulis

merupakan anggota Kelompok Bidang Keahlian (KBK) Matematika Terapan yang berfokus pada kajian tentang terapan matematika di bidang IPTEK, sosial dan industri. Selain mengajar, penulis juga aktif melaksanakan kegiatan penelitian dan pengabdian masyarakat. Tema penelitian yang menjadi fokus penulis adalah pemodelan penyakit menular, sistem dinamik dan kontrol optimal, serta pemodelan stokastik.

Penulis Bagian 11



Admi Salma, M.Si

seorang Dosen Prodi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang. Lahir di Kota Solok, Sumatera Barat, 25 Desember 1990. Penulis merupakan anak kedua dari tiga bersaudara dari pasangan bapak Admiral, S.Pd (Alm) dan Ibu Kartelina. Penulis menamatkan pendidikan program Sarjana (S1) di Universitas Negeri Padang prodi Pendidikan Matematika dan menyelesaikan program Pasca Sarjana (S2) di

Intitut Pertanian Bogor prodi Statistikka Terapan .

Penulis Bagian 12



Ni Putu Sudarsani, SE., M.Si

seorang Penulis dan Dosen Prodi Ekonomi Pembangunan Fakultas Ekonomi dan Bisnis Universitas Tabanan, Provinsi Bali. Lahir di Mengwi, 29 Oktober 1985 Badung. Penulis merupakan anak kelima dari lima bersaudara dari pasangan bapak I Made Dana dan Ibu Ni Made Sumarni. ia menamatkan pendidikan program Sarjana (S1) di Universitas Tabanan, Fakultas Ekonomi dan Bisnis Prodi Ekonomi,

menyelesaikan program Pasca Sarjana (S2) di Universitas Udayana, Denpasar, Bali Pada prodi Ekonomi. Dan sedang mengikuti program Doktor di Universitas Hindu Negeri I Gusti Bagus Sugriwa, Denpasar Bali.

Penulis Bagian 13



Ardiansyah, SPt. MSi.

Merupakan Dosen Jurusan Peternakan, Fakultas Pertanian, Universitas Mulawarman, Samarinda, Kalimantan Timur (2022-sekarang). Ia menamatkan Pendidikan S1 (2013) Jurusan Ilmu Nutrisi dan Teknologi Pakan dan S2 (2016) Jurusan Ilmu Nutrisi dan Pakan di Fakultas Peternakan, IPB University, Bogor, Jawa Barat. Semasa kuliah aktif sebagai asisten praktikum mata kuliah Rancangan Percobaan dan Formulasi Ransum. Saat ini mengampu beberapa matakuliah seperti Matematika terapan, Statistika, Metode Penelitian, dan Rancangan Percobaan.

Penulis Bagian 14



Basri, S.Sos.I., M.M

Penulis adalah Dosen Prodi Bisnis dan Manajemen Ritel pada kampus Politeknik Indonesia Kota Makassar. Lahir di Makassar, 16 April 1977 Sulawesi Selatan. Penulis merupakan anak kedelapan dari delapan bersaudara dari pasangan bapak Hamjah Daeng Gassing dan Andi Boddi Daeng Rannu. ia menamatkan pendidikan program Sarjana (S1) di Universitas Islam Negeri (UIN) Alauddin Makassar prodi Komunikasi Penyiaran Islam dan menyelesaikan program Pasca Sarjana (S2) di STIE YPUP Makasar prodi

Manajemen konsentrasi di bidang Manajemen Sumber Daya Manusia.

Penerbit :

PT. Sonpedia Publishing Indonesia

Buku Gudang Ilmu, Membaca Solusi
Kebodohan, Menulis Cara Terbaik
Mengikat Ilmu. Everyday New Books

SONPEDIA.COM
PT. Sonpedia Publishing Indonesia

Redaksi :

Jl. Kenali Jaya No 166

Kota Jambi 36129

Tel +6282177858344

Email: sonpediapublishing@gmail.com

Website: www.sonpedia.com