STATISTIK MULTIVARIAT UNTUK EKONOMI DAN BISNIS: di Lengkapi dengan Program IBM SPSS

Penulis:

Dr. Joni Wilson Sitopu, M.Pd.
Dr. Ir. Yongker Baali, M.Si
Rida Ristiyana, S.E., M.Ak., CIQnR., C.FR., C.Ftax., C.Ed., CTTr.
Anna Sofia Atichasari, SE.,M.Si.,CMA
Dr. Hj. Utin Nina Hermina., SE., MS.i.
Hartina Husain, S.Si., M.Stat.
Amalia Nur Chasanah, S.E, M.M.
Mario Donald Bani, S.P., M. Biotech. (Adv.)
Gregorio Antonny Bani

GET PRESS INDONESIA

STATISTIK MULTIVARIAT UNTUK EKONOMI DAN BISNIS: di Lengkapi dengan Program IBM SPSS

Penulis :
Dr. Joni Wilson Sitopu, M.Pd.
Dr. Ir. Yongker Baali, M.Si
Rida Ristiyana, S.E., M.Ak., CIQnR., C.FR., C.Ftax., C.Ed., CTTr.
Anna Sofia Atichasari, SE.,M.Si.,CMA
Dr. Hj. Utin Nina Hermina., SE., MS.i.
Hartina Husain, S.Si., M.Stat.
Amalia Nur Chasanah, S.E, M.M.
Mario Donald Bani, S.P., M. Biotech. (Adv.)
Gregorio Antonny Bani

ISBN :
Editor : Nanny Mayasari, S.Pd., M.Pd., CQMS.
Penyunting: Yuliatri M.Hum.
Desain Sampul dan Tata Letak : Atyka, S.Pd.

Penerbit : Get Press Indonesia
Anggota IKAPI No. 033/SBA/2022

Redaksi :
Jl. Palarik Air Pacah RT 001 RW 006
Kelurahan Air Pacah Kecamatan Koto Tangah
Padang Sumatera Barat
Website : www.getpress.co.id
Email : globaleksekutifteknologi@gmail.com

Cetakan pertama, Oktober 2023

KATA PENGANTAR

Alhamdulillah dengan mengucap rasa syukur kepada Allah SWT yang telah memberikan rahmat, karunia, dan hidayah-Nya berupa ilmu, kesehatan dan keselamatan. Atas kehendak-Nya, penulis dapat menyelesaikan buku dengan judul **STATISTIK MULTIVARIAT UNTUK EKONOMI DAN BISNIS: di Lengkapi dengan Program IBM SPSS.**

Buku ini berbeda dengan buku-buku penelitian lainnya, karena buku ini di lengkapi dengan program IBM SPSS untuk mempermudah dalam mengolah data penelitian. Buku ini membahas berbagai topik penting, seperti skala pengukuran, regresi sederhana, analisis faktor, analisis klaser, analisis conjoin, analisis diskriminan, analisis MANOVA. Setiap bab di lengkapi dengan contoh kasus praktis.

Buku ini mengurai keseluruhan proses penelitian dan pengolahan data hasil penelitian. Semoga buku ini memberikan banyak manfaat untuk mahasiswa, akademisi, praktisi, dan menjadi ladang pahala para penulisnya.

Padang, Oktober 2023
Penulis

Nanny Mayasari, S.Pd., M.Pd., CQMS.

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1

PENGANTAR ANALISIS MULTIVARIAT

Oleh Dr. Joni Wilson Sitopu, M.Pd.

1.1. Pendahuluan

Analisis statistik multivariat adalah penggunaan metode statistik matematika untuk mempelajari dan memecahkan masalah teori dan metode multi-indeks. Sejak 20 tahun terakhir, teknologi aplikasi komputer dan kebutuhan mendesak untuk penelitian, teknik analisis statistik multivariat banyak digunakan dalam bidang geologi, meteorologi, hidrologi, kedokteran, industri, pertanian dan ekonomi dan banyak bidang lainnya, telah menjadi solusi praktis masalah dengan cara yang efektif. (dalam KMT Lasmiatun., Solehudin., dkk., (2023)). Konsep yang mendasari penelitian ilmiah berupa istilah yang dikenal dengan variabel. Variabel ini bentuknya dapat berupa nilai angka dalam penelitian kuantitatif atau nilai mutu dalam penelitian kualitatif. Variabel dari suatu penelitian merupakan kegiatan pengujian cocok atau tidaknya teori dan fakta empiris yang ada di dalam dunia nyata. (dalam Muhammad Hasan., Suhelayanti., dkk., (2022)).

Analisis Multivariat adalah metode pengolahan variabel dalam jumlah yang banyak, dimana tujuannya adalah untuk mencari pengaruh variabel-variabel tersebut terhadap suatu obyek secara simultan atau serentak. Metode analisis multivariat adalah suatu metode statistika yang tujuan digunakannya adalah untuk menganalisis data yang terdiri dari banyak variabel serta diduga antar variabel tersebut saling berhubungan satu sama

lain. Analisis multivariat adalah salah satu dari teknik statistik yang diterapkan untuk memahami struktur data dalam dimensi tinggi. Dimana variabel-variabel yang dimaksud tersebut saling terkait satu sama lain. Berdasarkan beberapa definisi Analisis Multivariat di atas, maka statistikian menyimpulkan bahwa yang dimaksud dengan analisis multivariat adalah suatu analisis yang melibatkan varibel dalam jumlah lebih dari atau sama dengan 3 variabel. Dimana minimal ada satu variabel terikat dan lebih dari satu variabel bebas serta terdapat korelasi atau keterikatan antara satu variabel dengan variabel lainnya. Maka dapat diartikan bahwa Analisis Multivariat juga merupakan analisis yang melibatkan cara perhitungan yang kompleks. tujuannya adalah agar dapat memahami struktu data berdimensi tinggi dan saling berhubungan satu sama lain.

1.2. Jenis Data Dalam Analisis Multivariat

Ada berbagai jenis skala pengukuran, dan jenis data yang dikumpulkan menentukan jenis skala pengukuran yang akan digunakan untuk pengukuran statistik. Skala pengukuran ini berjumlah empat yaitu; skala nominal, skala ordinal, skala interval, dan skala rasio. Skala pengukuran digunakan untuk mengukur data kualitatif dan kuantitatif. Dengan skala nominal dan ordinal yang digunakan untuk mengukur data kualitatif sedangkan skala interval dan rasio digunakan untuk mengukur data kuantitatif (dalam Aysyah Rengganis, Nana Harlina Haruna., dkk., (2022)). Dalam Analisis Multivariat terdapat jenis data atau skala data. Skala data yang digunakan ada dua macam, yaitu data metrik dan data non metrik. Data metrik adalah data yang bersifat numerik atau berisi angka-angka dan dapat dilakukan perhitungan matematis di dalamnya, misal nilai ujian, tingkat IQ, berat badan, dll.

1.2.1. Data Metrik

Data metrik disebut juga dengan data numerik atau data kuantitatif. Dalam hal ini data metrik ada 2 macam, yaitu data interval dan data rasio.

a. **Interval**
   Skala interval adalah skala numerik di mana kita mengetahui urutan dan perbedaan yang tepat antara nilai-nilai tersebut. Contoh klasik dari skala interval adalah suhu Celcius karena perbedaan antara setiap nilai adalah sama. Misalnya, perbedaan antara 60 dan 50 derajat adalah 10 derajat yang terukur, seperti perbedaan antara 80 dan 70 derajat. Contoh umum adalah mengukur suhu pada skala Fahrenheit. Dapat digunakan untuk menghitung mean, median, modus, range, dan standar deviasi. Skala interval bagus karena bidang analisis statistik pada kumpulan data ini terbuka. Misalnya, tendensi sentral dapat diukur dengan modus, median, atau rata-rata; standar deviasi juga dapat dihitung, skala interval tidak hanya memberi tahu kita tentang urutan, tetapi juga tentang nilai di antara setiap item.
b. **Ratio**
   Skala rasio adalah tingkat tertinggi dalam pengukuran data karena skala ini memberi tahu kita tentang urutan, nilai pasti antar unit, dan juga memiliki interval yang sama dan nol mutlak yang memungkinkan untuk berbagai statistik deskriptif dan inferensial untuk diterapkan. Contoh variabel rasio yang baik meliputi tinggi, berat, dan durasi. Skala rasio memberikan banyak kemungkinan dalam hal analisis statistik. Variabel-variabel ini dapat ditambahkan, dikurangi, dikalikan, dibagi (rasio). Tendensi sentral dapat diukur dengan modus, median, atau rata-rata; ukuran dispersi, seperti standar deviasi dan koefisien variasi juga dapat dihitung dari skala rasio.

1.2.2. Data Nonmetrik

Data non metrik adalah data non numerik atau disebut juga data kualitatif atau data kategorik. Yang termasuk jenis data non metrik ini, yaitu data nominal dan data ordinal.

a. **Nominal,** skala nominal digunakan untuk pelabelan variabel, tanpa nilai kuantitatif. Skala nominal bisa disebut label. terdapat beberapa contoh di bawah ini. Perhatikan bahwa semua skala ini saling eksklusif (tidak tumpang tindih) dan tidak ada yang memiliki signifikansi numerik. Skala nominal mirip dengan nama atau label. Pada contoh di bawah ini, diukur dalam skala nominal. Pada contoh di bawah ini diukur dalam skala nominal

Tabel 1.1 contoh Pengukuran dalam Skala Nominal

Jenis Kelamin	Warna	Kota
L . Laki-laki P. Perempuan	1. Hitam 2. Putih 3. Merah 4. Kuning 5. dll	A. Medan B. Tebing Tinggi C. Pematang Siantar

(Sumber: Contoh Skala Nominal, 2023)

Memberi label L sebagai Laki-laki, P Perempuan atau 1 Hitam, 2 Putih, dstnya sama sekali tidak berarti salah satu atribut lebih baik dari yang lain. label atau kategori tidak menunjukan jumlah, hanya digunakan sebagai identitas untuk memudahkan analisis data.

b. **Ordinal**

Dengan skala ordinal yaitu urutan nilai adalah yang penting dan signifikan, tetapi perbedaan antara masing-masing nilai tidak terlalu diketahui. Skala ordinal biasanya mengukur konsep non-numerik seperti kepuasan, kebahagiaan, ketidaknyamanan, dll. Misalnya cara terbaik untuk menentukan tendensi sentral pada sekumpulan data ordinal adalah dengan menggunakan modus atau median. Misalnya: Direktur/Pejabat Rumah Sakit mungkin perlu bertanya kepada pasiennya: Bagaimana Anda menilai pelayanan di Rumah Sakit kami? 1) Sangat baik sekali; 2) baik sekali; 3) baik; 4) Kurang baik; 5)Tidak baik

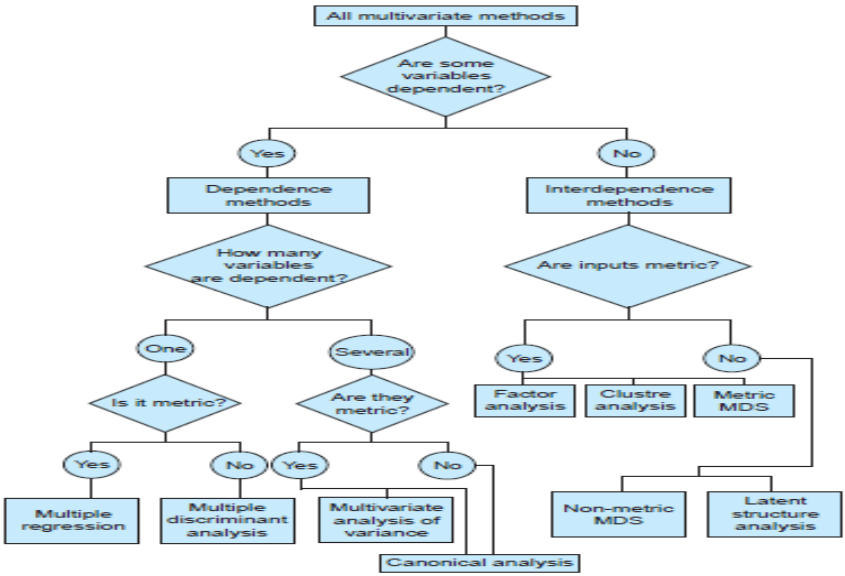
Atribut dalam contoh ini terdaftar dalam urutan menurun.

1.3. Klasifikasi Metode Data Analysis

Metode analisis multivariat adalah suatu metode statistika yang tujuan digunakannya adalah untuk menganalisis data yang terdiri dari banyak variabel serta diduga antar variabel tersebut saling berhubungan satu sama lain. Analisis multivariat adalah salah satu dari teknik statistik yang diterapkan untuk memahami struktur data dalam dimensi tinggi. Dimana variabel-variabel yang dimaksud tersebut saling terkait satu sama lain. (dalam Elfira Rahmadani., Mohan Taufiq Mashuri., Joni Wilson Sitopu., dkk., (2023)).

Metode analisis multivariat yang dapat dengan mudah diklasifikasikan menjadi dua kategori besar yaitu, metode ketergantungan dan metode saling ketergantungan. Jenis klasifikasi tergantung pada pertanyaan: Apakah beberapa variabel yang terlibat tergantung pada orang lain? Jika jawabannya 'ya', kita memiliki metode ketergantungan; tetapi jika jawabannya adalah 'tidak', kita memiliki metode saling ketergantungan. Dua pertanyaan lagi relevan untuk memahami sifat teknik multivariat.

Pertama, jika beberapa variabel bergantung, pertanyaannya adalah berapa banyak variabel yang bergantung? Pertanyaan lainnya adalah, apakah datanya metrik atau non-metrik? Ini berarti apakah datanya kuantitatif, dikumpulkan dalam skala interval atau rasio, atau apakah datanya kualitatif, dikumpulkan dalam skala nominal atau ordinal. Metode yang akan digunakan untuk situasi tertentu bergantung pada jawaban atas semua pertanyaan. Diagram alir sifat dari beberapa metode multivariat penting seperti yang ditunjukkan pada gambar. di bawah ini.



Gambar 1.1 Diagram Alir 1
(Sumber: Jadish N. Sheth)

Jadi, kita memiliki dua jenis metode multivariat: satu jenis untuk data yang berisi variabel dependen dan independen, dan jenis lainnya untuk data yang berisi beberapa variabel tanpa hubungan ketergantungan. Dalam kategori sebelumnya termasuk metode seperti analisis regresi berganda, analisis diskriminan berganda, analisis varians multivariat dan analisis kanonik, sedangkan dalam kategori terakhir kita menempatkan metode seperti analisis faktor, analisis kluster, penskalaan multidimensi atau MDS (baik metrik maupun non-metrik) dan analisis struktur laten.

1.3.1. Dependence

Metode *dependence* digunakan jika satu atau beberapa variabel bergantung pada variabel yang lain. Ketergantungan melihat sebab dan akibat; dengan kata lain, dapatkah nilai dari dua atau lebih variabel independen digunakan untuk menjelaskan, menggambarkan, atau memprediksi nilai variabel dependen lainnya? Sebagai contoh sederhana, variabel

dependen, berat badan dapat diprediksi oleh variabel independen seperti tinggi dan usia.

Metode *dependence* termasuk di antaranya adalah:

a. **Multiple Regression Analysis**
   Analisis regresi berganda digunakan ketika ada satu variabel dependen (Y) dan dua atau lebih variabel independen (X). Dalam metode ini, variabel independen ditandai dengan nomor. Misalnya, X_1, X_2, X_3, dst.
b. **Linear Probability Model**
   Linear probability model (LPM) merupakan metode regresi linear yang *outcome* atau hasilnya berupa bilangan biner (0 atau 1), dan satu atau lebih variabel independen digunakan untuk memprediksi hasilnya.
c. **Multivariate Analysis of Variance and Covariance**
   Multivariate Analysis of Variance (MANOVA) adalah perluasan dari metode *Analysis of Variance* (ANOVA). Metode ini mengukur efek dari dua atau lebih variabel independen terhadap dua atau lebih variabel dependen. Sederhananya, data yang memiliki dua variabel X dan dua variabel Y dihitung dengan metode ini.

1.3.2. Interdependence

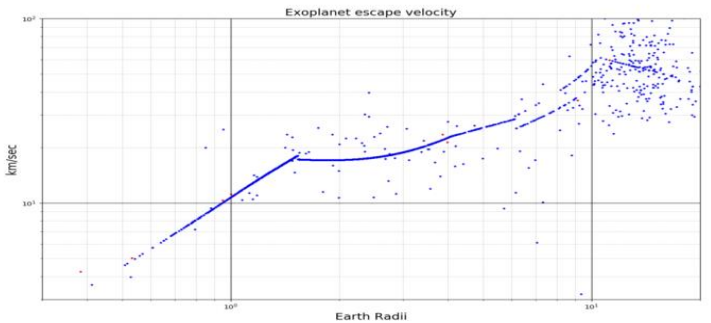
Metode interdependence (saling ketergantungan) digunakan untuk memahami susunan struktural dan pola yang mendasari dalam kumpulan data. Dalam hal ini, tidak ada variabel yang bergantung pada yang lain, jadi kita tidak mencari hubungan sebab akibat. Sebaliknya, metode interdependence berusaha memberi makna pada sekumpulan variabel atau mengelompokkannya bersama-sama dengan cara yang bermakna. Jadi: Salah satunya adalah tentang pengaruh variabel tertentu pada yang lain, sedangkan yang lainnya adalah tentang struktur kumpulan data. Dengan mempertimbangkan beberapa metode analisis multivariat yang berguna, yaitu :

a. Regresi linier berganda
b. Regresi logistik berganda

c. Analisis varians multivariat (MANOVA)
d. Analisis factor
e. Analisis klister

1. Regresi linier berganda

Regresi linier berganda adalah metode ketergantungan yang melihat hubungan antara satu variabel dependen dan dua atau lebih variabel independen. Model regresi berganda akan memberi tahu kita sejauh mana setiap variabel independen memiliki hubungan linier dengan variabel dependen. Hal ini berguna karena membantu kita memahami faktor mana yang mungkin mepengaruhi hasil tertentu, memungkinkan kita memperkirakan hasil di masa mendatang. Misalnya sebagai analis data, kita dapat menggunakan regresi berganda untuk memprediksi pertumbuhan tanaman. Dalam contoh ini, pertumbuhan tanaman adalah variabel dependen kita dan kita ingin melihat bagaimana berbagai faktor mempengaruhinya. Variabel independen kita bisa berupa curah hujan, suhu, jumlah sinar matahari, dan jumlah pupuk yang ditambahkan ke tanah. Model regresi berganda akan menunjukkan kepada kita proporsi varian dalam pertumbuhan tanaman yang diperhitungkan oleh masing-masing variabel independen.



Gambar 1.2 exoplanet escape velocity
(Sumber *: Wikimedia Commons)*

2. Regresi logistik berganda

Analisis regresi logistik digunakan untuk menghitung (dan memprediksi) probabilitas kejadian biner yang terjadi. Hasil biner adalah hasil yang hanya memiliki dua kemungkinan hasil; apakah peristiwa itu terjadi (1) atau tidak (0). Jadi, berdasarkan serangkaian variabel independen, regresi logistik dapat memprediksi seberapa besar kemungkinan skenario tertentu akan muncul. Ini juga digunakan untuk klasifikasi. Andaikan jika kita bekerja sebagai analis di sektor asuransi dan kita perlu memprediksi seberapa besar kemungkinan setiap pelanggan potensial akan mengajukan klaim. kita dapat memasukkan berbagai variabel independen ke dalam model kita, seperti usia, apakah mereka memiliki kondisi kesehatan yang serius atau tidak, pekerjaan mereka, dan seterusnya. Dengan menggunakan variabel-variabel ini, analisis regresi logistik akan menghitung kemungkinan kejadian (membuat klaim) terjadi.

3. Analisis Varians Multivariat (MANOVA)

Multivariate analysis of variance (MANOVA) digunakan untuk mengukur pengaruh beberapa variabel independen terhadap dua atau lebih variabel dependen. Dengan MANOVA, variabel independen bersifat kategoris, sedangkan variabel dependen bersifat metrik. Variabel kategori adalah variabel yang termasuk dalam kategori berbeda, misalnya, variabel status pekerjaan dapat dikategorikan ke dalam unit-unit tertentu, seperti bekerja penuh waktu, bekerja paruh waktu, menganggur, dan lainya. Variabel metrik diukur secara kuantitatif dan nilainya numerik. Dalam analisis MANOVA, kita melihat berbagai kombinasi variabel independen untuk membandingkan perbedaan pengaruhnya terhadap variabel dependen.

4. Analisis faktor

Analisis faktor adalah metode saling ketergantungan yang berupaya mengurangi jumlah variabel dalam kumpulan data. Jika kita memiliki terlalu banyak variabel, akan sulit untuk

menemukan pola dalam data kita. Pada saat yang sama, model yang dibuat menggunakan kumpulan data dengan terlalu banyak variabel rentan terhadap overfitting. Overfitting adalah kesalahan pemodelan yang terjadi ketika model cocok terlalu dekat dan spesifik dengan kumpulan data tertentu, membuatnya kurang dapat digeneralisasikan untuk kumpulan data di masa mendatang, dan dengan demikian berpotensi kurang akurat dalam prediksi yang dibuatnya. Analisis faktor bekerja dengan mendeteksi kumpulan variabel yang berkorelasi tinggi satu sama lain. Variabel-variabel ini kemudian dapat diringkas menjadi satu variabel. Analis data akan sering melakukan analisis faktor untuk menyiapkan data untuk analisis selanjutnya.

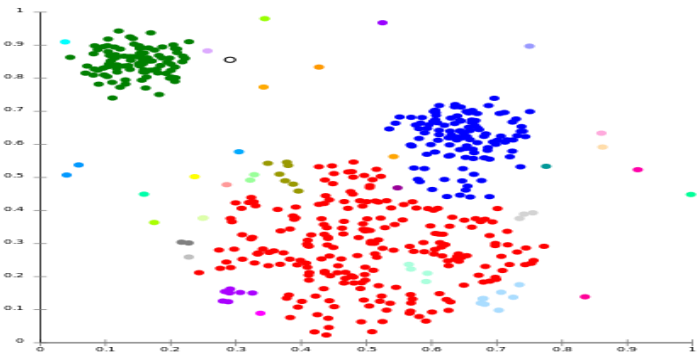
Andaikan kita memiliki kumpulan data yang berisi data yang berkaitan dengan pendapatan, tingkat pendidikan, dan pekerjaan seseorang. kita mungkin menemukan tingkat korelasi yang tinggi di antara masing-masing variabel, dan dengan demikian mereduksinya menjadi faktor tunggal yaitu status sosial ekonomi. Kita mungkin juga memiliki data tentang seberapa senang kita dengan layanan pelanggan, seberapa besar kita menyukai produk tertentu, dan seberapa besar kemungkinan kita merekomendasikan produk tersebut ke teman. Masing-masing variabel ini dapat dikelompokkan ke dalam faktor tunggal yaitu kepuasan pelanggan (selama ditemukan berkorelasi kuat satu sama lain). Meskipun kita telah mengurangi beberapa titik data menjadi satu faktor saja, kita tidak benar-benar kehilangan informasi apa pun, faktor-faktor ini cukup menangkap dan mewakili masing-masing variabel yang bersangkutan. Dengan kumpulan data yang disederhanakan, kita dapat untuk melakukan analisis lebih lanjut.

5. Analisis klaster

Metode saling ketergantungan lainnya, analisis klaster digunakan untuk mengelompokkan item serupa dalam kumpulan data ke dalam klaster. Saat mengelompokkan data ke

dalam cluster, tujuannya adalah agar variabel dalam satu cluster lebih mirip satu sama lain daripada variabel di cluster lain. Hal ini diukur dalam jarak intracluster dan antarcluster. Jarak Intracluster melihat jarak antar titik data dalam satu cluster. Ihal ini harus kecil. Jarak Intercluster melihat jarak antar titik data dalam cluster yang berbeda. Hal ini idealnya berukuran besar.

Analisis klaster membantu kita memahami bagaimana data dalam sampel kita didistribusikan dan dapat menemukan pola. Contoh analisis klaster adalah segmentasi audiens. Jika kita bekerja di bidang pemasaran, kita dapat menggunakan analisis kluster untuk menentukan kelompok pelanggan berbeda yang dapat memperoleh manfaat dari kampanye yang lebih bertarget. Sebagai analis layanan kesehatan, kita dapat menggunakan analisis kluster untuk mengeksplorasi apakah faktor gaya hidup tertentu atau lokasi geografis terkait dengan kasus penyakit tertentu yang lebih tinggi atau lebih rendah. Karena merupakan metode saling ketergantungan, analisis klaster sering dilakukan pada tahap awal analisis data.



Gambar 1.3 Analisis Klaser

(Sumber : Wikimedia Commons)

1.4. Penggunaan Metode multivariate

Tabel 1.2 Metode Dependence

Variabel	Dependence			
	1 variabel		2 atau lebih variabel	
	Metrik	Nonmetrik	Metrik	Nonmetrik
Independence 1 variabel				
Metrik	Regresi	Analisis diskriminant	Korelasi Kanonik	MDA (Multiple-group Discriminat Analysis)
Nonmetrik	t-test	Analisis Diskrit disriminant	MANOVA (Multivariate Analiysis of Varians)	Dikrit MDA
2 atau lebih			Korelasi Kanonik	MDA (Multiple-group Discriminat Analysis)
Metrik	Multiple Regresion	Analisis diskriminant Regresi Logistik	MANOVA (Multivariate Analiysis of Varians)	Dikrit MDA
Nonmetrik	ANOVA	Analisis Diskrit disriminant Analisis Konjoint (MOMANOVA)	Korelasi Kanonik	MDA (Multiple-group Discriminat Analysis)

(Sumber: Joseph F. Hair JR., dkk, 2010).

Tabel 1.3 Metode independence 1

Metode	Variabel Bebas	Variabel Terikat
Analisis Logistik	Nominal atau ordinal	Nominal atau ordinal
Analisis Varians	Nominal atau ordinal	Interval atau ratio
Analisis diskriminat	Interval atau ratio	Nominal atau ordinal
Analisis Regresi dan analisis korelasi kanonik	Interval atau ratio	Interval atau ratio

(Sumber: Joseph F. Hair JR., dkk, 2010).

Tabel 1.4 Metode independence 2

Variabel	Jenis data	
	Metrik	Nonmetrik
2 Variabel	Simple correlation	-Two-way contingency table -Loglinear Models
Lebih dari 2	Principal components	multiway contingency table
	Factor analysis	Loglinear Models
		Correspondence analysis

(Sumber: Joseph F. Hair JR., dkk, 2010).

Tabel 1.5 Metode independence 3

Metode	Variabel Pertama	Variabel Kedua
Analisis Loglinear	Nominal atau ordinal	Nominal atau ordinal
Analisis kelompok (cluster)	Nominal atau ordinal	Interval atau ratio
Analisis Komponen Utama dan Analisis Faktor	Interval atau ratio	Interval atau ratio

(Sumber: Joseph F. Hair JR., dkk, 2010).

Tergantung pada skala pengkuran data yang digunakan

DAFTAR PUSTAKA

Aysyah Rengganis., Nana Harlina Haruna,., dkk., (2022). *Penelitian dan Pengembangan*. Medan. Yayasan Kita Menulis.

Elfira Rahmadani, Mohan Taufiq Mashuri, Joni Wilson Sitopu., (2023). *Statistika Pendidikan*. Padang. PT GLOBAL EKSEKUTIF TEKNOLOGI.

Joseph F. Hair JR., William C. Black., dkk. (2010). Multivariate Data Analisis. Pearson Prentice Hall.

KMT Lasmiatun., Solehudin., dkk., (2023). *Manajemen Dan Analisis Data*. Padang. PT GLOBAL EKSEKUTIF TEKNOLOGI.

Muhammad Hasan., Suhelayanti., dkk., (2022). *Pengantar Riset Pendidikan*. Medan. Yayasan Kita Menulis.

BAB 2

SKALA PENGUKURAN DAN METODE ANALISIS DATA

Oleh Dr. Ir. Yongker Baali, M.Si.

2.1. Pendahuluan

Pengenalan Tentang Skala Pengukuran dalam Penelitian, Skala pengukuran adalah alat atau metode yang digunakan untuk mengukur atau menggambarkan karakteristik atau variabel dalam penelitian. Skala pengukuran ini membantu peneliti untuk mengumpulkan data yang relevan dan menyajikan hasil penelitian dengan cara yang dapat diinterpretasikan.

Berikut adalah beberapa jenis skala pengukuran yang umum digunakan dalam penelitian (Sekaran dan Bougie, 2016):

1. Skala Nominal: digunakan untuk mengkategorikan atau mengklasifikasikan data ke dalam kelompok-kelompok yang berbeda berdasarkan atribut atau karakteristik tertentu. Contohnya adalah jenis kelamin (laki-laki, perempuan), status perkawinan (menikah, belum menikah), atau jenis pekerjaan (dokter, guru, pengusaha).
2. Skala Ordinal: memungkinkan pengurutan atau peringkat data berdasarkan urutan atau tingkatan tertentu, tetapi tidak memberikan informasi tentang jarak antara peringkat tersebut. Contohnya adalah tingkat kepuasan (sangat puas, puas, tidak puas) atau tingkat pendidikan (SD, SMP, SMA, perguruan tinggi).

3. Skala Interval: memberikan informasi tentang peringkat data serta jarak antara peringkat tersebut dengan titik nol yang tidak bermakna. Pada skala ini, perbedaan antara nilai-nilai memiliki arti yang bermakna. Contohnya adalah suhu dalam derajat Celsius atau Fahrenheit.
4. Skala Rasio: memiliki semua karakteristik skala interval, tetapi juga memiliki titik nol yang bermakna. Oleh karena itu, skala rasio dapat digunakan untuk mengukur perbandingan dan perbedaan proporsional antara nilai-nilai. Contohnya adalah usia, tinggi badan, atau pendapatan.

2.2. Jenis-Jenis Skala Pengukuran

Berikut adalah penjelasan mengenai jenis-jenis skala pengukuran, yaitu:

1. Skala Nominal: adalah jenis skala pengukuran yang digunakan untuk mengkategorikan data ke dalam kelompok-kelompok yang berbeda berdasarkan atribut atau karakteristik tertentu. Skala ini hanya menyediakan label atau nama untuk setiap kategori, tetapi tidak ada urutan atau tingkatan yang diberikan pada data tersebut. Contoh penggunaan skala nominal adalah jenis kelamin, agama, atau kategori warna. Skala nominal tidak mengizinkan perbandingan atau operasi matematis di antara kategori-kategori tersebut (Hair *dkk.*, 2019).
2. Skala Ordinal: memungkinkan pengurutan atau peringkat data berdasarkan urutan atau tingkatan tertentu. Pada skala ini, data diberi label dengan tingkatan atau peringkat, tetapi jarak antara peringkat tidak memiliki makna yang sama. Contoh penggunaan skala ordinal adalah tingkat kepuasan pelanggan (sangat puas, puas, tidak puas) atau tingkat pendidikan (SD, SMP, SMA, perguruan tinggi). Skala ordinal dapat menyediakan informasi tentang urutan, tetapi tidak memberikan informasi tentang perbedaan antara tingkatan tersebut (Polit dan Beck, 2017)
3. Skala Interval: memberikan informasi tentang peringkat data serta jarak antara peringkat tersebut dengan titik nol

yang tidak bermakna. Pada skala ini, perbedaan antara nilai-nilai memiliki arti yang bermakna. Skala interval memberikan kesempatan untuk melakukan operasi matematika seperti penambahan atau pengurangan. Contoh penggunaan skala interval adalah suhu dalam derajat Celsius atau Fahrenheit. Namun, perbandingan rasio antara nilai-nilai tidak dapat dihitung pada skala interval (Cresswell, 2014)

4. Skala Rasio: memiliki semua karakteristik skala interval, tetapi juga memiliki titik nol yang bermakna. Dengan adanya titik nol, skala rasio memungkinkan perbandingan dan perbedaan proporsional antara nilai-nilai. Contoh penggunaan skala rasio adalah usia, berat badan, atau pendapatan. Pada skala ini, semua operasi matematika dapat dilakukan, termasuk perbandingan rasio antara nilai-nilai (Sekaran dan Bougie, 2016).

2.2.1. Kelebihan dan Kelemahan Jenis Skala Pengukuran

Berikut adalah uraian mengenai kelebihan dan kelemahan masing-masing jenis skala pengukuran, yaitu skala nominal, ordinal, interval, dan rasio:

1. Skala Nominal:
   Kelebihan: Skala nominal memberikan kemampuan untuk mengkategorikan data secara jelas dan sederhana. Data yang dikategorikan dapat digunakan untuk melakukan analisis statistik dasar, seperti frekuensi dan persentase.
   Kelemahan: Skala nominal tidak mengizinkan perbandingan atau operasi matematis antara kategori-kategori. Penggunaan skala ini membatasi analisis statistik yang lebih lanjut, seperti penghitungan rata-rata atau uji perbedaan signifikan.
2. Skala Ordinal:
   Kelebihan: Skala ordinal memberikan informasi tentang urutan atau tingkatan data. Data dapat diurutkan dan

digunakan untuk analisis perbandingan sederhana. Skala ini mempertimbangkan perbedaan tingkat antara kategori-kategori.

Kelemahan: Skala ordinal tidak memberikan informasi yang cukup detail tentang perbedaan antara tingkatan. Jarak antara peringkat tidak memiliki makna yang sama, sehingga analisis statistik yang melibatkan operasi matematika yang kompleks menjadi terbatas (Sekaran dan Bougie, 2016).

3. Skala Interval:

   Kelebihan: Skala interval memungkinkan pengukuran yang lebih akurat dengan mempertimbangkan perbedaan antara nilai-nilai. Skala ini memungkinkan penggunaan operasi matematika seperti penambahan dan pengurangan. Analisis statistik yang lebih maju, seperti penghitungan rata-rata dan uji statistik, dapat dilakukan pada skala interval.

   Kelemahan: Skala interval tidak memiliki titik nol yang bermakna, sehingga perbandingan rasio antara nilai-nilai tidak dapat dilakukan. Hasil pengukuran pada skala ini dapat memberikan kesan bahwa perbandingan absolut terhadap nol mungkin ada, padahal titik nolnya sebenarnya tidak ada atau hanya merupakan titik acak.

4. Skala Rasio:

   Kelebihan: Skala rasio memiliki semua kelebihan skala interval dan memiliki titik nol yang bermakna. Skala ini memungkinkan perbandingan rasio antara nilai-nilai dan operasi matematika yang lebih lengkap. Data pada skala rasio dapat digunakan untuk analisis statistik yang lebih lanjut, seperti perbandingan proporsional dan perhitungan rasio.

   Kelemahan: Skala rasio sering kali memerlukan instrumen pengukuran yang lebih presisi dan akurat. Selain itu, pada beberapa kasus, titik nol yang bermakna mungkin tidak tercapai dalam pengukuran tertentu, misalnya pada suhu dalam derajat Celcius.

2.2.2. Proses Pemilihan Skala Pengukuran Penelitian

Proses pemilihan skala pengukuran yang sesuai untuk variabel penelitian melibatkan pertimbangan berbagai faktor yang berkaitan dengan sifat variabel, tujuan penelitian, dan analisis statistik yang akan dilakukan yaitu:

1. Pertimbangkan Sifat Variabel:
   a. Identifikasi apakah variabel bersifat kategorikal (nominal atau ordinal) atau bersifat numerik (interval atau rasio).
   b. Jika variabel bersifat kategorikal, pertimbangkan apakah informasi urutan (ordinal) dibutuhkan atau hanya informasi kategori saja (nominal).
   c. Jika variabel bersifat numerik, pertimbangkan apakah diperlukan perhitungan rasio (rasio) atau cukup perhitungan relatif (interval).
2. Pertimbangkan Tujuan Penelitian:
   a. Pahami tujuan penelitian dan pertanyaan penelitian yang ingin dijawab.
   b. Tentukan apakah skala pengukuran yang dipilih dapat memberikan informasi yang relevan untuk mencapai tujuan penelitian.
   c. Pastikan skala pengukuran dapat memberikan tingkat keakuratan dan detail yang dibutuhkan untuk analisis data (Sekaran dan Bougie, 2016).
3. Pertimbangkan Analisis Statistik yang Dilakukan:
   a. Pertimbangkan metode analisis statistik yang akan digunakan dalam penelitian.
   b. Pastikan skala pengukuran yang dipilih sesuai dengan persyaratan analisis statistik tersebut.
   c. Perhatikan apakah analisis statistik yang akan dilakukan memerlukan operasi matematika yang kompleks, seperti penghitungan rata-rata, uji perbedaan, atau analisis regresi (Suharsimi, 2017).
4. Lakukan Uji Praktis dan Uji Kesesuaian:
   a. Lakukan uji praktis untuk memastikan skala pengukuran yang dipilih dapat digunakan dengan mudah oleh responden dan peneliti.

b. Lakukan uji kesesuaian untuk memverifikasi apakah skala pengukuran yang dipilih dapat menghasilkan data yang valid dan reliabel (Sugiyono, 2019).

2.3. Teknik Pengumpulan Data

Tinjauan Tentang Teknik Pengumpulan Data yang Umum Digunakan. Teknik pengumpulan data merupakan langkah penting dalam proses penelitian. Berikut adalah uraian tentang teknik pengumpulan data umum yang digunakan, seperti wawancara, kuesioner, observasi, dan studi dokumentasi yaitu (Sugiyono, 2019):

1. Wawancara: adalah teknik pengumpulan data yang melibatkan interaksi langsung antara peneliti dan responden. Wawancara dapat dilakukan secara tatap muka atau melalui telepon. Teknik ini memungkinkan peneliti untuk mendapatkan informasi yang mendalam dan konteks yang lebih luas. Wawancara juga memungkinkan adanya dialog antara peneliti dan responden, sehingga memungkinkan klarifikasi dan pertanyaan tambahan.
2. Kuesioner: adalah teknik pengumpulan data yang melibatkan pemberian serangkaian pertanyaan tertulis kepada responden. Kuesioner dapat diberikan secara langsung, melalui pos, atau secara online. Teknik ini memberikan kesempatan untuk mengumpulkan data dari jumlah responden yang besar dalam waktu yang relatif singkat. Kuesioner yang dirancang dengan baik akan menghasilkan data yang konsisten dan terstruktur.
3. Observasi: adalah teknik pengumpulan data yang melibatkan pengamatan langsung terhadap objek atau subjek penelitian. Observasi dapat dilakukan dengan mengamati perilaku, interaksi, atau kejadian yang terjadi secara alami. Observasi dapat dilakukan dengan menggunakan instrumen pengamatan yang telah ditentukan sebelumnya. Teknik ini memungkinkan peneliti untuk mengumpulkan data tentang perilaku yang sebenarnya terjadi tanpa ada pengaruh dari persepsi atau laporan subjek.

4. Studi Dokumentasi: melibatkan pengumpulan data dari dokumen-dokumen yang relevan dengan penelitian. Dokumen tersebut dapat berupa jurnal ilmiah, laporan, catatan, arsip, atau dokumen resmi lainnya. Teknik ini memungkinkan peneliti untuk mendapatkan informasi yang sudah ada sebelumnya, seperti data historis, kebijakan, atau dokumentasi kegiatan tertentu. Studi dokumentasi dapat menjadi sumber data yang berharga untuk mendukung penelitian.

a. Kelebihan dan Kelemahan Masing-masing Teknik Pengumpulan Data

Berikut adalah uraian tentang kelebihan dan kelemahan masing-masing teknik pengumpulan data yang umum digunakan yaitu:

1. Wawancara
   a. Kelebihan:
      - Memungkinkan peneliti untuk memperoleh informasi mendalam dan konteks yang lebih luas.
      - Memungkinkan dialog dan interaksi langsung antara peneliti dan responden.
      - Memungkinkan klarifikasi dan pertanyaan tambahan.
   b. Kelemahan:
      - Waktu yang dibutuhkan untuk melakukan wawancara mungkin lebih lama dibandingkan teknik pengumpulan data lainnya.
      - Dapat terpengaruh oleh persepsi atau sikap subjek wawancara.
2. Kuesioner
   a. Kelebihan:
      - Dapat mengumpulkan data dari jumlah responden yang besar dalam waktu yang relatif singkat.
      - Data yang dikumpulkan dapat terstruktur dan konsisten.
      - Mudah untuk dianalisis secara statistik.
   b. Kelemahan:

- Respon dari responden dapat terbatas atau tidak akurat.
- Memerlukan keterampilan merancang kuesioner yang baik untuk menghindari bias atau kesalahan interpretasi.

3. Observasi
   a. Kelebihan:
      - Memungkinkan pengamatan langsung terhadap perilaku atau kejadian yang terjadi secara alami.
      - Data yang dikumpulkan dapat objektif karena tidak ada pengaruh dari persepsi subjek.
      - Dapat memberikan data tentang konteks yang lebih kaya.
   b. Kelemahan:
      - Observasi mungkin memerlukan waktu yang lama dan menghabiskan sumber daya.
      - Beberapa perilaku atau kejadian mungkin sulit untuk diobservasi.
      - Subjek yang diamati dapat merasa tidak nyaman atau mempengaruhi perilaku mereka.
4. Studi Dokumentasi
   a. Kelebihan:
      - Memanfaatkan data yang telah ada, seperti laporan, dokumen resmi, atau arsip.
      - Dapat memberikan informasi historis atau data yang konsisten.
      - Membantu dalam memahami konteks dan kebijakan terkait penelitian.
   b. Kelemahan:
      - Keterbatasan dalam akses ke dokumen yang diperlukan.
      - Validitas dan akurasi data yang dikumpulkan dapat bervariasi tergantung pada sumber dokumen.
      - Tergantung pada kualitas dan ketersediaan dokumen yang relevan.

b. Faktor-faktor yang Harus Dipertimbangkan dalam Memilih Teknik Pengumpulan Data yang Sesuai untuk Penelitian

Berikut adalah faktor-faktor yang harus dipertimbangkan dalam memilih teknik pengumpulan data yang sesuai untuk penelitian yaitu:

1. Tujuan Penelitian:
   - Menentukan apakah tujuan penelitian adalah mengumpulkan data kuantitatif atau kualitatif.
   - Menentukan apakah penelitian membutuhkan data yang bersifat deskriptif, eksploratif, atau verifikatif.
   - Menyesuaikan teknik pengumpulan data dengan tujuan penelitian yang ingin dicapai (Misalnya, apakah tujuan penelitian adalah mengukur fenomena, menjelaskan pola, atau memahami persepsi).
2. Sifat Variabel Penelitian:
   - Mengidentifikasi jenis variabel yang akan diteliti, apakah bersifat kuantitatif atau kualitatif.
   - Mengidentifikasi apakah variabel penelitian dapat diukur secara langsung atau memerlukan pendekatan yang lebih subjektif.
   - Menyesuaikan teknik pengumpulan data dengan karakteristik variabel tersebut (Misalnya, wawancara untuk data kualitatif dan kuesioner untuk data kuantitatif).
3. Karakteristik Responden atau Subjek Penelitian:
   - Mempertimbangkan karakteristik populasi atau sampel penelitian, seperti usia, pendidikan, budaya, atau latar belakang sosial.
   - Menyesuaikan teknik pengumpulan data agar cocok dengan kelompok responden yang diteliti.
4. Ketersediaan Sumber Daya:
   - Mempertimbangkan ketersediaan sumber daya manusia, waktu, dana, dan infrastruktur yang diperlukan untuk melaksanakan teknik pengumpulan data tertentu.

- Menyesuaikan teknik pengumpulan data dengan ketersediaan sumber daya yang dimiliki (Misalnya, waktu, dana, tenaga, dan infrastruktur).
- Memilih teknik yang paling efisien dan efektif dalam konteks ketersediaan sumber daya.

5. Validitas dan Reliabilitas Data:
   - Menilai validitas dan reliabilitas data yang dapat diperoleh dari teknik pengumpulan data yang dipilih.
   - Memastikan teknik pengumpulan data yang dipilih dapat menghasilkan data yang valid dan reliabel.
   - Melakukan uji coba atau validasi terhadap instrumen pengumpulan data untuk memastikan kualitas data yang diperoleh.
6. Etika Penelitian:
   - Memperhatikan kepatuhan terhadap aspek etika dalam pengumpulan data, seperti hak privasi dan kerahasiaan data responden.
   - Memastikan bahwa teknik yang digunakan tidak melanggar etika penelitian.

2.4. Analisis Data Deskriptif

2.4.1. Konsep Analisis Data Deskriptif

Analisis data deskriptif adalah suatu proses dalam penelitian yang bertujuan untuk menggambarkan, mengorganisir, dan meringkas data yang telah dikumpulkan. Melalui analisis deskriptif, peneliti dapat memahami karakteristik data, melihat pola-pola yang ada, dan memberikan gambaran umum tentang variabel-variabel yang diteliti.

Berikut adalah uraian tentang konsep analisis data deskriptif yaitu (Bungin, 2016):

1. Pengertian Analisis Data Deskriptif: Analisis data deskriptif merupakan teknik statistik yang digunakan untuk menyajikan dan menggambarkan data secara numerik

maupun grafis. Tujuan utamanya adalah untuk memberikan gambaran yang jelas tentang data yang ada, seperti distribusi frekuensi, ukuran pemusatan, ukuran variabilitas, dan hubungan antar variabel.

2. Metode Analisis Data Deskriptif:
   - Statistik Deskriptif: Melibatkan penggunaan ukuran-ukuran statistik seperti mean (rata-rata), median (tengah), dan modus (nilai yang paling sering muncul) untuk menggambarkan karakteristik pusat data.
   - Grafik dan Diagram: Menggunakan visualisasi seperti diagram batang, diagram lingkaran, histogram, dan diagram pencar untuk membantu memahami pola-pola data dan memvisualisasikan distribusi variabel.
3. Interpretasi Hasil Analisis Data Deskriptif: Hasil analisis data deskriptif digunakan untuk memberikan gambaran dan pemahaman yang lebih baik tentang variabel-variabel yang diteliti. Misalnya, mengidentifikasi pemusatan data (apakah data cenderung berpusat di sekitar nilai tertentu), variasi data (sejauh mana data tersebar), atau kecenderungan kategori yang dominan.
4. Kelebihan Analisis Data Deskriptif:
   - Memberikan gambaran yang jelas tentang karakteristik data yang ada.
   - Membantu mengenali pola, tren, atau kecenderungan yang terdapat dalam data.
   - Menyajikan data dengan cara yang mudah dipahami oleh pembaca atau pemangku kepentingan.
   - Memberikan dasar yang kuat untuk analisis lebih lanjut.
5. Kelemahan Analisis Data Deskriptif:
   - Tidak memberikan penjelasan tentang hubungan sebab-akibat antara variabel.
   - Tidak menghasilkan kesimpulan atau generalisasi yang kuat seperti pada analisis inferensial.

2.4.2. Teknik dan Metode dalam Analisis Data Deskriptif

Teknik dan metode yang digunakan dalam analisis data deskriptif meliputi frekuensi, persentase, mean, median, modus, dan visualisasi data. Berikut adalah uraian tentang teknik dan metode tersebut (Bungin, 2016):

1. Frekuensi: adalah teknik yang digunakan untuk menghitung jumlah kemunculan atau distribusi frekuensi dari setiap nilai atau kategori dalam data. Frekuensi memberikan gambaran tentang seberapa sering suatu nilai atau kategori muncul dalam dataset.
2. Persentase: digunakan untuk menggambarkan proporsi atau perbandingan antara jumlah kemunculan suatu nilai atau kategori dengan total keseluruhan data. Persentase memberikan informasi tentang kontribusi relatif setiap nilai atau kategori dalam dataset.
3. Mean (Rata-rata): adalah nilai rata-rata dari sejumlah data yang dihitung dengan menjumlahkan semua nilai dan membaginya dengan jumlah data yang ada. Mean memberikan gambaran tentang nilai tengah atau pusat dari data.
4. Median (Tengah): adalah nilai tengah yang membagi data menjadi dua bagian yang sama jumlahnya. Median digunakan untuk menggambarkan nilai tengah dalam data, terlepas dari adanya outlier atau nilai ekstrim.
5. Modus (Nilai yang paling sering muncul): adalah nilai yang paling sering muncul dalam dataset. Modus digunakan untuk mengidentifikasi nilai atau kategori dominan dalam data.
6. Visualisasi Data: melibatkan penggunaan grafik atau diagram untuk mempresentasikan data secara visual. Beberapa metode visualisasi yang umum digunakan meliputi diagram batang, diagram lingkaran, histogram, dan diagram pencar. Visualisasi data membantu memahami pola, tren, dan distribusi data dengan lebih jelas dan mudah dipahami.

2.4.3. Interpretasi Hasil Analisis Data Deskriptif

Interpretasi hasil analisis data deskriptif merupakan proses untuk menggambarkan dan memberikan pemahaman tentang karakteristik data yang telah dianalisis yaitu:

1. Menginterpretasikan Frekuensi dan Persentase: Frekuensi dan persentase digunakan untuk menggambarkan distribusi nilai atau kategori dalam dataset. Interpretasi dapat dilakukan dengan melihat nilai frekuensi atau persentase tertinggi, nilai yang paling sering muncul, atau pola distribusi yang menonjol. Hal ini membantu dalam mengidentifikasi nilai dominan atau kategori yang signifikan dalam dataset.
2. Menginterpretasikan Mean (Rata-rata): Mean digunakan untuk memahami nilai tengah atau pusat dari data. Interpretasi dapat dilakukan dengan membandingkan mean dengan nilai lainnya dalam dataset atau dengan menggunakan nilai referensi tertentu. Hal ini membantu dalam mengevaluasi sejauh mana data cenderung mendekati atau menjauhi nilai rata-rata.
3. Menginterpretasikan Median (Tengah): Median digunakan untuk memahami nilai tengah dalam data. Interpretasi dapat dilakukan dengan membandingkan median dengan nilai ekstrem atau menggunakan nilai median sebagai titik acuan. Hal ini membantu dalam melihat distribusi data secara keseluruhan dan mengidentifikasi nilai yang lebih representatif daripada mean dalam situasi tertentu.
4. Menginterpretasikan Modus (Nilai yang paling sering muncul): Modus digunakan untuk mengidentifikasi nilai atau kategori dominan dalam data. Interpretasi dapat dilakukan dengan menyoroti nilai modus dan melihat seberapa sering nilai tersebut muncul. Hal ini membantu dalam menggambarkan nilai yang paling umum atau yang memiliki frekuensi tertinggi dalam dataset.
5. Menginterpretasikan Visualisasi Data: Visualisasi data membantu dalam memahami pola dan karakteristik data secara lebih intuitif. Interpretasi dapat dilakukan dengan mengamati pola, tren, atau distribusi yang terlihat dalam

grafik atau diagram. Hal ini membantu dalam menyampaikan informasi dengan cara yang lebih visual dan mudah dipahami.

2.5. Metode Analisis Data Inferensial

Pengenalan tentang analisis data inferensial, analisis data inferensial adalah metode statistik yang digunakan untuk membuat inferensi atau kesimpulan tentang populasi berdasarkan data sampel yang terbatas. Analisis data inferensial melibatkan dua tahap penting, yaitu estimasi dan pengujian hipotesis. Pada tahap estimasi, para peneliti menggunakan data sampel untuk mengestimasi parameter populasi yang tidak dapat diukur langsung. Sementara itu, pada tahap pengujian hipotesis, para peneliti menguji klaim atau hipotesis yang diajukan tentang populasi berdasarkan data sampel yang telah dikumpulkan.

Pada analisis data inferensial, pemilihan teknik statistik yang tepat menjadi kunci. Beberapa teknik yang umum digunakan dalam analisis data inferensial meliputi uji t-parametrik, uji chi-square, uji ANOVA (Analysis of Variance), dan regresi. Teknik-teknik ini digunakan untuk menguji perbedaan atau hubungan antara variabel, menguji signifikansi hipotesis, dan membuat generalisasi tentang populasi berdasarkan data sampel. Pendekatan analisis data inferensial juga melibatkan penggunaan tingkat signifikansi dan interval kepercayaan. Tingkat signifikansi menentukan seberapa kecil probabilitas kesalahan yang diterima saat menolak atau gagal menolak hipotesis. Interval kepercayaan adalah rentang nilai yang digunakan untuk mengestimasi parameter populasi dengan tingkat kepercayaan tertentu.

Pada umumnya, analisis data inferensial digunakan dalam penelitian ilmiah, survei, atau percobaan di berbagai bidang, termasuk kedokteran, ekonomi, psikologi, dan sains sosial. Tujuannya adalah untuk membuat kesimpulan yang lebih luas dan menggeneralisasikan temuan dari sampel ke populasi yang lebih besar.

Teknik dan metode yang digunakan dalam analisis data inferensial mencakup uji hipotesis, uji beda, uji regresi, dan analisis varians (ANOVA). Berikut ini adalah uraian mengenai teknik dan metode tersebut:

1. Uji Hipotesis: adalah teknik yang digunakan untuk menguji klaim atau hipotesis yang diajukan tentang parameter populasi. Pada uji hipotesis, hipotesis nol (H0) diajukan untuk diuji kebenarannya berdasarkan data sampel yang telah dikumpulkan. Berbagai jenis uji hipotesis yang umum digunakan meliputi uji t, uji chi-square, uji F, dan uji z. Uji hipotesis ini memberikan dasar untuk membuat kesimpulan tentang signifikansi statistik dan mengambil keputusan mengenai hipotesis yang diajukan.
2. Uji Beda: digunakan untuk membandingkan rata-rata atau proporsi antara dua kelompok atau lebih. Uji beda yang umum digunakan antara lain uji t-paired, uji t-independent, uji chi-square untuk beda proporsi, dan uji ANOVA untuk beda rata-rata antara tiga kelompok atau lebih. Teknik ini membantu mengidentifikasi perbedaan yang signifikan antara kelompok-kelompok tersebut dan memberikan pemahaman yang lebih mendalam tentang hubungan antara variabel.
3. Uji Regresi: digunakan untuk menganalisis hubungan antara satu atau lebih variabel independen dengan variabel dependen. Metode regresi linier, regresi logistik, dan regresi multinomial merupakan beberapa teknik regresi yang sering digunakan. Uji regresi membantu dalam memahami pengaruh variabel independen terhadap variabel dependen, mengidentifikasi hubungan fungsional antara variabel, dan memprediksi nilai variabel dependen berdasarkan nilai variabel independen.
4. Analisis Of Varians (ANOVA): digunakan untuk membandingkan rata-rata antara tiga kelompok atau lebih. ANOVA memungkinkan kita untuk menentukan apakah terdapat perbedaan signifikan antara kelompok-kelompok tersebut. Beberapa jenis ANOVA yang sering digunakan meliputi ANOVA satu arah, ANOVA dua arah, dan ANOVA

faktorial. Teknik ini membantu dalam memahami perbedaan yang signifikan antara kelompok-kelompok dan menyediakan informasi tentang variabel yang berkontribusi terhadap perbedaan tersebut.

Interpretasi hasil analisis data inferensial dan pengambilan kesimpulan adalah langkah penting dalam penelitian. Berikut ini adalah uraian mengenai interpretasi hasil analisis data inferensial dan pengambilan kesimpulan yaitu (Sudjana, 2015):

1. Interpretasi Hasil Analisis Data Inferensial: melibatkan penafsiran signifikansi statistik dan estimasi parameter. Hasil analisis data inferensial mencakup nilai-nilai p-value, interval kepercayaan, dan ukuran efek. P-value digunakan untuk menentukan apakah terdapat perbedaan yang signifikan antara kelompok-kelompok atau hubungan yang signifikan antara variabel. Interval kepercayaan digunakan untuk mengestimasi rentang nilai yang mungkin berisi parameter populasi. Ukuran efek (*effect size*) digunakan untuk mengukur kekuatan hubungan atau perbedaan antara variabel.
2. Pengambilan Kesimpulan: didasarkan pada hasil analisis data inferensial dan melibatkan evaluasi signifikansi statistik, interpretasi ukuran efek, dan relevansi temuan penelitian. Kesimpulan yang diambil harus mempertimbangkan konteks penelitian, tujuan penelitian, dan implikasi praktis dari temuan. Kesimpulan harus didukung oleh bukti statistik yang kuat dan pemahaman yang komprehensif terhadap analisis data inferensial yang dilakukan.

DAFTAR PUSTAKA

Bungin, B. (2016) *Metodologi penelitian kuantitatif: Komunikasi, ekonomi, dan kebijakan publik serta ilmu sosial lainnya*. Surabaya: Prenada Media Group.

Cresswell, J. . (2014) *Research design: Qualitative, quantitative, and mixed methods approaches*. California: Sage Publications.

Hair, J. . *et al.* (2019) *Multivariate Data Analysis (8th ed.)*. Boston: Cengage Learning.

Polit, D. . and Beck, C. . (2017) *Nursing Research: Generating and Assessing Evidence for Nursing Practice. 10th Edition*. Philadelphia: Wolters Kluwer Health.

Sekaran, U. and Bougie, R. (2016) *Research methods for business: A skill-building approach*. New York: John Wiley & Sons.

Sudjana (2015) *Metode statistika*. Bandung: PT. Tarsito.

Sugiyono (2019) *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.

Suharsimi, A. (2017) *Prosedur penelitian: Suatu pendekatan praktik*. Jakarta: Rineka Cipta.

BAB 3

REGRESI SEDERHANA & KORELASI

Oleh
Rida Ristiyana, S.E., M.Ak., CIQnR., C.FR., C.Ftax., C.Ed., CTTr.

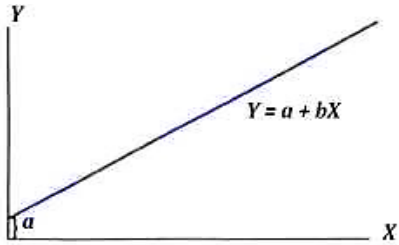
3.1. Pendahuluan

Dari berbagai aspek dan lingkup kehidupan, kita sering menjumpai adanya korelasi satu hal dengan hal lainnya atau satu variabel dengan variabel lain. Bab 3 ini membahas mengenai regresi linier sederhana dan korelasi sederhana. Disebut sederhana sebab melibatkan hanya pada satu variabel bebas dan satu variabel terikat. Peneliti yang ingin mengetahui hubungan antar dua variabel yang dilihat baik dari kekuatan dan bentuk hubungan perlu mempelajari bab ini.

3.2. Regresi Linier Sederhana

Definisi Regresi Linier Sederhana, analisis regresi adalah analisis yang digunakan untuk mengukur ada atau tidaknya hubungan antar variabel. Regresi linier adalah regresi yang memiliki variabel bebas dengan pangkat paling tinggi adalah satu (Ghozali, 2018). Regresi liner sederhana adalah analisis yang digunakan untuk melihat hubungan fungsional variabel bebas dan variabel terikat yang dinyatakan dalam persamaan matematik dan garis serta memiliki hubungan yang searah (Kemendikbud, 2023). Persamaan regresi ini dapat berbentuk garis lurus (linier) dan tidak lurus atau non linier (tidak lurus).

Persamaan regresi linier sederhana adalah suatu model persamaan yang mendeskripsikan hubungan variabel bebas *(predictor)* dengan variabel terikat *(response)* (Ghozali, 2021) yang digambarkan dengan garis lurus seperti pada Gambar 3.1.



Gambar 3.1 Persamaan Regresi Linier Sederhana

Secara matematik persamaan regresi liner sederhana yaitu :

$$\mathbf{Y = a + b\,X}$$

Keterangan :

Y = Garis regresi/variabel *response*

a = Konstanta *(intersep),* potongan dengan sumber vertikal

b = konstanta regresi *(slope)*

X = Variabel bebas *(predictor)*

Besarnya konstanta menggunakan persamaan berikut (Yuliara, 2016) :

$$a = \frac{(\sum Y_i)(\sum X_i^2) - (\sum X_i)(\sum X_i Y_i)}{n\sum X_i^2 - (\sum X_i)^2}$$

$$b = \frac{n\ (\sum X_i\, Y_i) - (\sum X_i)\,(\sum Y_i)}{n \sum X_i^2\ -\ (\sum X_i)^2}$$

n = jumlah data

3.2.1. Tujuan dan Syarat Kelayakan

RLS Memiliki tujuan diantaranya (Hidayat, 2022) :

a) Menghitung nilai estimasi serta nilai dari variabel terikat yang disesuaikan dengna dengan nilai pada variabel bebas
b) Untuk menguji dugaaan sementara (hipotesis)

c) Untuk meramalkan nilai rata-rata dari variabel bebas dengan menyesuaikan nilai variabel bebas diluar jangkauan sampel.

Saat melakukan regresi linier sederhana harus memenuhi syarat kelayakan (Ghozali, 2018), yaitu :

a. Jumlah sampel yang dipakai harus sama
b. Jumlah dari variabel bebas berjumlah 1
c. Nilai residual wajib berdistribusi normal
d. Adanya hubungan linier atas variabel bebas dan terikat
e. Tidak terkena heterokedastisitas
f. Tidak terkena autokorelasi

3.2.2. Langkah-Langkah Analisis RLS

Terdapat delapan tahapan/langkah yang harus dilakukan dalam menganalisis regresi linier sederhana secara manual, yaitu :

a. Menetapkan tujuan dari analisis regresi linier sederhana
b. Mengidentifikasi variabel *predictor* dan *response*
c. Mengumpulkan/kolektif data (susun bentuk tabel)
d. Menghitung X2, XY dan total dari masing-masing
e. Menghitung a dan b (menggunakan rumus)
f. Menyusun model persamaan garis regresi
g. Memprediksi variabel *predictor* dan *response*
h. Menetapkan taraf signifikansi dan melakukan pengujian Uji t.

Sebagai ilustrasi untuk mempermudah pemahaman diberikan ilustrasi (contoh) seperti dibawah ini.

Contoh ke-1 :
Pertanyaan :
Penelitian mengenai apakah konsumsi jumlah kalori per hari dapat mempengaruhi berat badan dari 10 mahasiswa. Deskripsi terlihat pada Tabel 3.1.

Tabel 3.1 Deskripsi Jumlah kalori dan berat badan

No.	Nama Mahasiswa	Kalori/ hari (X)	Berat Badan (Y)
1	Sebastian	530	89
2	Marie	300	48
3	Sinta	358	56
4	Yudha	510	72
5	Teti	302	54
6	Ani	300	42
7	Ridwan	387	60
8	Aqita	527	85
9	Melani	415	63
10	Yudi	512	74

(Sumber: Penulis, 2023. Diolah)

Analisislah menggunakan RLS ?

Penyelesaian :

a. Menetapkan tujuan dari analisis regresi linier sederhana
   Untuk mengetahui apakah jumlah kalori per hari dapat mempengaruhi berat badan mahasiswa.
b. Mengidentifikasi variabel *predictor* dan *response*
   Variabel X = Jumlah kalori per hari
   Variabel Y = Berat badan
c. Mengumpulkan/kolektif data (susun bentuk tabel) dan Menghitung X2, XY dan total dari masing-masing
   Untuk memudahkan disusun tabel bantuan seperti berikut :

Tabel 3.2 Kolektif Data

No.	X	$_X$2	Y	$_Y$2	XY
1	530	280900	89	7921	47170
2	300	90000	48	2304	14400
3	358	128164	56	3136	20048
4	510	260100	72	5184	36720
5	302	91204	54	2916	16308
6	300	90000	42	1764	12600
7	387	149769	60	3600	23220
8	527	277729	85	7225	44795
9	415	172225	63	3969	26145
10	512	262144	74	5476	37888
Σ	4141	1802235	643	43495	279294

(Sumber : Penulis, 2023. Diolah)

d. Menghitung a dan b (menggunakan rumus)

$$\mathbf{b} = \frac{n\ (\sum X_i\, Y_i) - (\sum X_i)(\sum Y_i)}{n \sum X_i^{\,2} - (\sum X_i)^2}$$

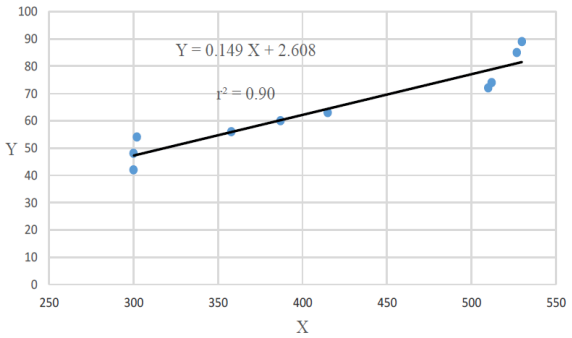
$$= \frac{10(279294) - (4141)(643)}{10(1802235) - (4141)^2} = \frac{130227}{874469} \cong 0{,}14892 \approx 0{,}149$$

$$\mathbf{a} = \frac{(\sum Y_i)(\sum X_i^{\,2}) - (\sum X_i)(\sum X_i Y_i)}{n \sum X_i^{\,2} - (\sum X_i)^2}$$

$$= \frac{(643)(1802235) - (4141)(279294)}{10(1802235) - (4141)^2} = \frac{2280651}{874469} \cong 2{,}608$$

Atau dengan menggunakan cara lain untuk mencari nilai a, yaitu : a = $\bar{Y}$ − b $\bar{X}$ = 64,3 − 0,149(414,1) ≅ 2,608, sehingga persamaan menjadi Y = 2,608 + 0,149 X.

e. Menyusun model persamaan garis regresi



Gambar 3.2 Garis Regresi Hubungan X dengan Y

(Sumber: Penulis, 2023. Diolah)

f. Memprediksi variabel *predictor* dan *response*

Dilakukan analisis korelasi yang hasilnya merupakan koefisien korelasi. Formula korelasi (Yuliara, 2016) yaitu :

$$r = \frac{n\sum_{i=1}^{n} Xi\,Yi - \left(\sum_{i=1}^{n} Xi\right)\left(\sum_{i=1}^{n} Yi\right)}{\sqrt{\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} Xi\right)^2\right]\left[n\sum_{i=1}^{n} Y_i^2 - \left(\sum_{i=1}^{n} Yi\right)^2\right]}}$$

Dengan rumus diatas maka koefisien korelasinya adalah :

$$r = \frac{n\sum_{i=1}^{n} X_i\,Y_i - \left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{n} Y_i\right)}{\sqrt{\left[n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right]\left[n\sum_{i=1}^{n} Y_i^2 - \left(\sum_{i=1}^{n} Y_i\right)^2\right]}}$$

$$= \frac{10(279294) - (4141)(643)}{\sqrt{\left[10(1802235) - (4141)^2\right]\left[10(43495) - (643)^2\right]}}$$

$$= \frac{130277}{137120{,}2318} = 0{,}95$$

Dari hasil diatas menunjukkan bahwa hubungan sangat kuat (95%) antara variabel bebas dengan variabel terikat. Ini artinya konsumsi jumlah kalori per hari sangat mempengaruhi berat badan mahasiswa. Sedangkan untuk koefisien determinasi menunjukkan bahwa variabel bebeas

mampu menjelaskan variabel terikat sebesar 90% (r^2 = 0,90).

g. Menetapkan taraf signifikansi & melakukan pengujian Uji t
Diketahui r^2 = 0,90, jumlah data n = 10
Hipotesis yang diajukan :
H0 : β = 0; variabel X tidak berdampak signifikan terhadap Y
H1 : β ≠ 0; variable X berdampak signifikan terhadap Y
Tingkat signifikansi = 5%
Nilai t_{hit} =

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0{,}95\sqrt{10-2}}{\sqrt{1-0{,}90}} = 8{,}497$$

Sehingga t_{hit} = 8,497, derajat kebebasan df = n – k = 10-2 = 8
Gunakan tabel uji t sig 5%, df = 8 maka diperoleh t_{tab} = 2,306. Sehingga jika dibandingkan $t_{hit} > t_{tab.}$ Ini artinya ada pengaruh/dampak yang signifikan antara variabel bebas dengan variabel terikat.

3.2.3. Studi Kasus RLS Dengan SPSS

Contoh ke-2 :

Pertanyaan :

Peneliti ingin mengetahui dampak stres kerja pada kinerja pegawai dengan melibatkan 12 responden dan sudah diuji reliabilitas dan validitas. Data yang sudah disajikan pada tabel 3.3 sudah lolos asumsi klasik. Hipotesis yang diajukan adalah "terdapat dampak stres kerja pada kinerja pegawai. Data penelitian sebagai berikut :

Tabel 3.3 Data Studi Kasus

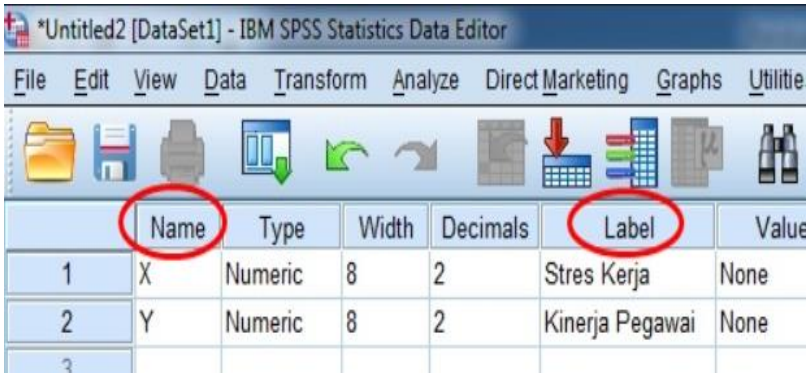
No	Stres Kerja (X)	Kinerja Pegawai (Y)
1	28	21
2	20	24
3	21	27
4	23	22
5	17	26
6	25	24
7	22	23
8	19	25
9	27	21
10	25	22
11	24	25
12	17	28

(Sumber : Penulis, 2023. Diolah)

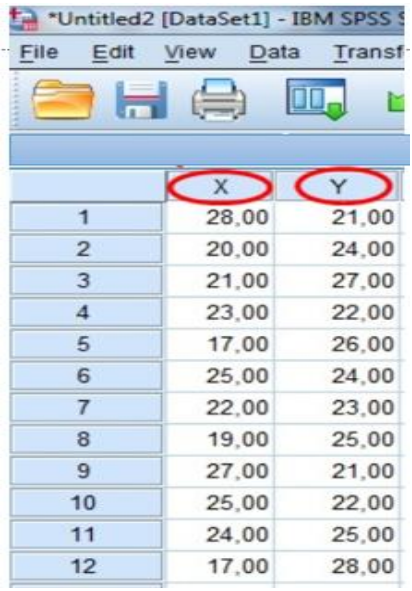
Bagaiman proses olah data dan interpretasi menggunakan SPSS

Penyelesaian :

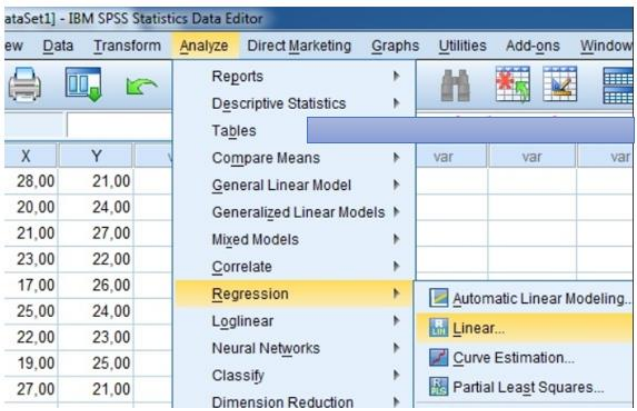
a) Install aplikasi SPSS dan buka lembar kerja dan pilh variabel *view*. Pada kolom name ketik simbol dari variabel bebas dan terikat, pada labael diketik deskripsi nama variabel.



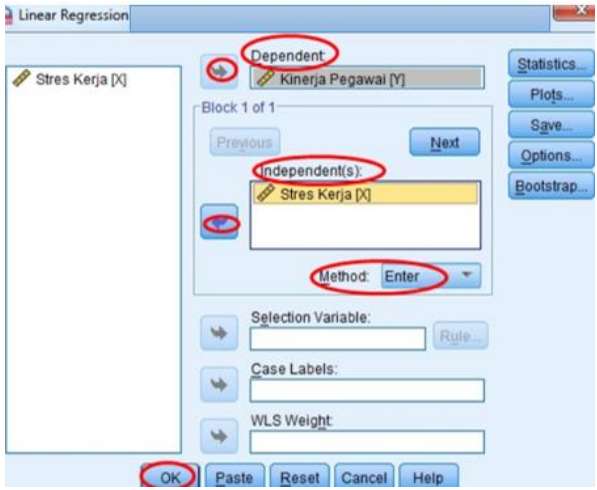
b) Pada *Data view* masukan data mentahnya (excel)



c) Klik menu *Analyze – Regression – Linier*



d) Klik kotak dialog *Linier Regression*, masukan variabel bebas ke kotak independent dan variabel terikat ke kotak dependent. Metode Enter lalu OK.



e) Muncul *output* seperti :

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Stres Kerja[b]	.	Enter

a. Dependent Variable: Kinerja Pegawai

b. All requested variables entered.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,813[a]	,661	,627	1,40170

a. Predictors: (Constant), Stres Kerja

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	38,352	1	38,352	19,520	,001[b]
	Residual	19,648	10	1,965		
	Total	58,000	11			

a. Dependent Variable: Kinerja Pegawai

b. Predictors: (Constant), Stres Kerja

Coefficients[a]

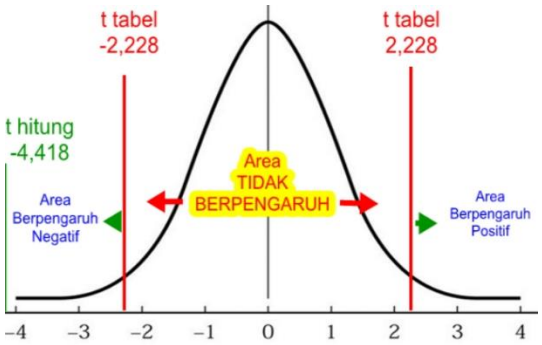
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	35,420	2,616		13,538	,000
	Stres Kerja	-,511	,116	-,813	-4,418	,001

a. Dependent Variable: Kinerja Pegawai

Berdasarkan pada hasil tabel *coefficient* maka dapat dibuat persamaan regresi liner sederhana yaitu : Y = 35,420 – 0,511 X. Angka konstan (a) yang menunjukkan apabila stres kerja tidak ada maka kinerja pegawai akan konsisten senilai 35,420. Angka koefisien regresi (b) artinya tiap penambahan 1% tingkat stres kerja akan justru meningkatkan kinerja pegawai

senilai -0,511. Karena nilai pada koef.regresi minus (-) maka stres kerja berdampak negatif pada kinerja pegawai.

Uji hipotesis merujuk juga pada tabel koefisien. Diketahui signifikansi sebesar 0,001 < prob (0,05) maka dapat disimpulkan H0 ditolak dan Ha diterima yang artinya ”stres kerja berdampak pada kinerja pegawai”. Nilai t hit -4,418 > t tabel (2,228).



Gambar 3.3 Kurva Stres Kerja dengan Kinerja Pegawai

(Sumber : Penulis, 2023. Diolah)

Dari hasil kurva tersebut dapat disimpulkan -4.418 ada pada area negatif. Ini artinya terdapat dampak negatif antara stres kerja dengan kinerja pegawai.

3.3. Korelasi Sederhana

3.3.1. Definisi Korelasi Sederhana

Definisi korelasi sederhana adalah teknik statistik yang digunakan untuk mengetahui bentuk hubungan antar variabel bebas dan terikat sekaligus mengukur kekuatan dari hubungan tersebut yang memiliki hubungan timbal balik (hubungan dua arah) (Kemendikbud, 2023). Bentuk hubungan bisa positif atau negatif sedangkan kekuatan hubungan bisa erat, lemah maupun tidak erat (Ghozali, 2018).

3.3.2. Jenis-Jenis Korelasi antar Dua Variabel

Korelasi memiliki variasi bentuk hubungan, (Esa Unggul, 2018) yaitu :

a. Korelasi Positif : Jika variabel bebas meningkat maka variabel terikat juga cenderung meningkat dan sebaliknya.
b. Korelasi Negatif : Jika variabel bebas meningkat maka variabel terikat cenderung menurun dan sebaliknya.
c. Korelasi Sempurna : Jika kenaikan/penurunan variabel bebas selalu sebanding dengan kenaikan/penurunan variabel terikat.
d. Tidak ada korelasi : Jika tidak adanya hubungan antara kedua variabel.

Contoh ke-3 :

Pertanyaan :

Perusahaan Aqilla.Corp. ingin melihat hubungan hasil penjualan dengan biaya iklan. Berikut datanya (Tabel 3.4).

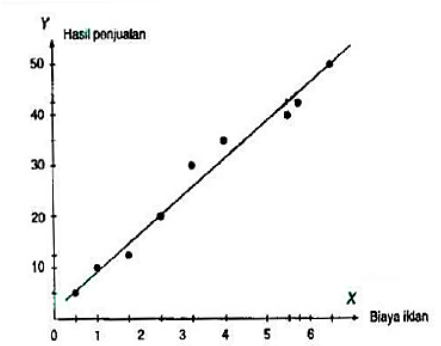
Tabel 3.4 Data Hasil penjualan dan Biaya Iklan

Biaya Iklan (Juta Rp)	**Biaya Penjualan (Juta Rp)**
0,50	5,00
1,00	10,00
1,75	12,50
2,50	20,00
3,25	30,00
4,00	35,00
5,50	40,00
5,75	42,50
6,50	50,00

(Sumber : Penulis, 2023. Diolah)

Buatlah diagram dari data tersebut dan masuk kedalam jenis korelasi yang mana ?

Penyelesaian :



Gambar 3.4 Diagram Penjualan dan Iklan

(Sumber : Penulis, 2023. Diolah)

Dari Gambar 3.4 menunjukkan bahwa hubungan penjualan dengan iklan adalah korelasi positif, semakin meningkat biaya penjualan maka semakin meningkat pula biaya iklan.

3.3.3. Koefisien Korelasi dan Koefisien Determinasi

Koefisien korelasi (r) adalah Indeks/bilangan yang digunakan untuk mengukur keeratan suatu hubungan antar variabel. Berikut formulanya (Esa Unggul, 2018):

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

Keterangan :

X = Variabel bebas/independen

Y = Variabel terikat/dependen

n = Banyaknya sampel

nilai dari r antara -1 dan 1 ($-1 \leq r \leq 1$)

Koefisien determinasi adalah seberapa jauh/besar semua variabel bebas dapat menjelaskan varians dari variabel terikatnya. Umumnya dengan mengkuadratkan nilai koefisien korelasi. Nilai koefisien determinasi range 0 – 1. Contoh : nilai r sebesar 0,6 maka koefisien determinasi 0,6 x 0,6 = 0,36. Artinya variabel bebas mampu menjelaskan varians dari variabel terikatnya sebesar 36%. Artinya 64% (100%-36%) varians variabel dependen yang dapat diterangkan oleh faktor lain.

3.3.4. Interval Keeratan Korelasi Antar Variabel

Interval keeratan pada hubungan antar variabel dibedakan menjadi (Esa Unggul, 2018; Ghozali, 2021):

a. r = 0 Korelasi tidak ada
b. $0 < r \leq 0{,}20$ Korelasi sangat lemah
c. $0{,}20 < r \leq 0{,}40$ Korelasi lemah sekali
d. $0{,}40 < r \leq 0{,}70$ Korelasi cukup kuat
e. $0{,}70 < r \leq 0{,}90$ Korelasi Kuat
f. $0{,}90 < r < 1{,}00$ Korelasi Sangat Kuat
g. $r = 1$ Korelasi sempurna

3.3.5. Jenis-Jenis Koefisien Korelasi

Adapun koefisien korelasi dibedakan menjadi :

1. Koefisien Korelasi *Pearson Correlation :*
   Digunakan untuk mengukur suatu hubungan dari dua variabel baik yang data rasio maupun interval (Universitas Negeri Yogyakarta, 2022) Biasanya simbolnya adalah 'R' dan bersifat parametrik kuantitatif. Metode ini bisa ditempuh dengan 2 cara yaitu :
   a. Metode *Least Square*

$$r = \frac{n\Sigma XY - \Sigma X . \Sigma Y}{\sqrt{\left(n\Sigma X^2 - (\Sigma X)^2\right)\left(n\Sigma Y^2 - (\Sigma Y)^2\right)}}$$

b. Metode *Product Moment Correlation*

Digunakan untuk skala ordinal dan bersifat nonparametrik statitik (Universitas Negeri Yogyakarta, 2022).

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 . \Sigma y^2}}$$

Keterangan :
r = koefisien korelasi
x = deviasi rata-rata variabel X
y = deviasai rata-rata variabel Y

2. Koefisien Korelasi *Rank Spearman*
3. Indeks yang digunakan untuk mengukur keeratan hubungan diantara dua variabel yang berupa data ordinal. Biasanya disimbolkan dengan 'Rs'. Berikut formulanya :

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}$$

Keterangan :
r_s = koefisien korelasi rank spearman
d = selisih dalam ranking
n = banyaknya pasangan rank

4. Koefisien Penentu (KP) / Koefisien Determinasi (R^2)
Digunakan untuk menjelasakan seberapa besar dampak suatu nilai variabel bebas pada naik/turunnya nilai variabel terikat. Formulanya adalah :

$$\mathbf{KP = R^2 = (KK)^2 \ X\ 100\%}$$

Keterangan :

KK = Koefisien korelasi

Jika koefisien korelasinya merupakan koefisien korelasi pearson (r) maka koefisien penentunya yaitu :

$$\mathbf{KP = R^2 = r^2 \ X \ 100\%}$$

3.3.6. Studi Kasus Korelasi Pada SPSS

Contoh ke-4 :

Peneliti ingin menguji hubungan antara motivasi dan minat dengan prestasi belajar dari 12 mahasiswa. Berikut datanya :

Tabel 3.5 Data Motivasi, Minat dan Prestasi

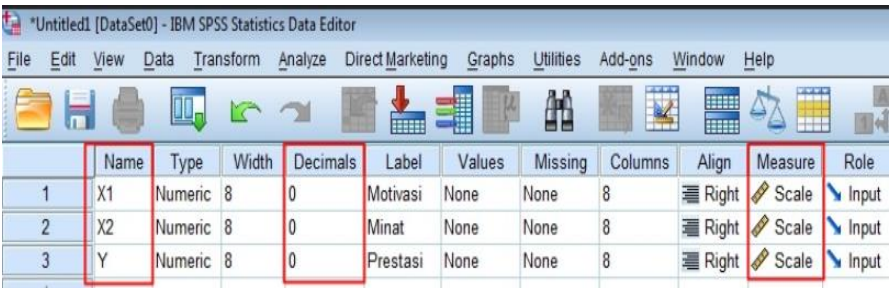
No. Responden	Motivasi	Minat	Prestasi
1	75	75	80
2	60	70	75
3	65	70	75
4	75	80	90
5	65	75	85
6	80	80	85
7	75	85	95
8	80	88	95
9	65	75	80
10	80	75	90
11	60	65	75
12	65	70	75

(Sumber : Raharjo, 2021)

Bagaiman proses olah datanya dan interpretasi uji *Pearson Corellation* menggunakan SPSS ?

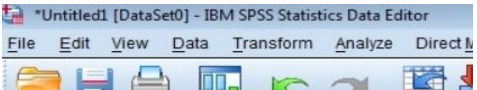
Penyelesaian :

a) Install SPSS dan buka SPSS, pilih menu variabel *view* dan ketik simbol variabel yang ada. *Decimals* dirubah menjadi 0, label diberi nama deskripsi variabel dan *measure* ganti *scale.*



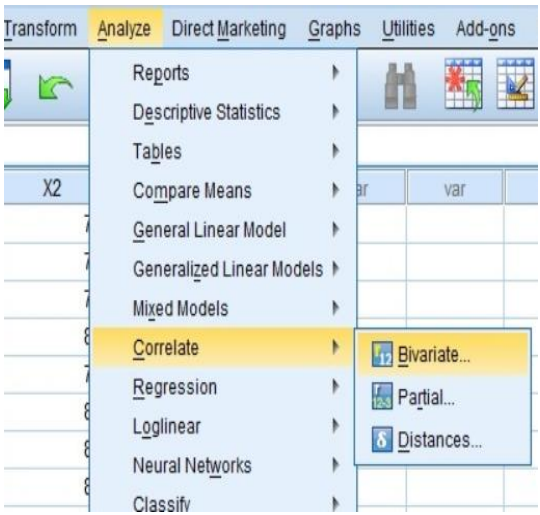
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	X1	Numeric	8	0	Motivasi	None	None	8	Right	Scale	Input
2	X2	Numeric	8	0	Minat	None	None	8	Right	Scale	Input
3	Y	Numeric	8	0	Prestasi	None	None	8	Right	Scale	Input

b) Klik Data *view* masukan data mentah motivasi, Minat dan Prestasi.

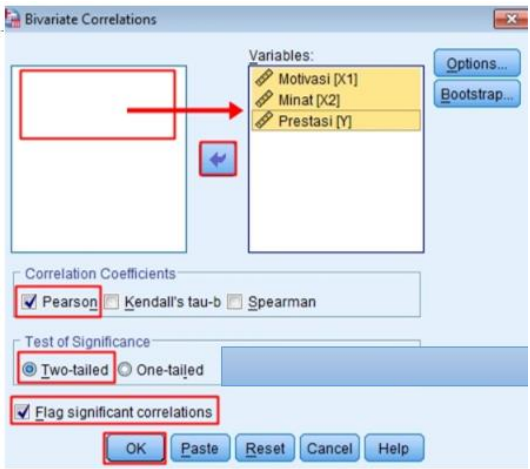


	X1	X2	Y
1	75	75	80
2	60	70	75
3	65	70	75
4	75	80	90
5	65	75	85
6	80	80	85
7	75	85	95
8	80	88	95
9	65	75	80
10	80	75	90
11	60	65	75
12	65	70	75

c) Pilih menu *Analyze – Correlate – Bivariate*



d) Muncul kotak dialog bivariate correlations, masukan semua variabel bebasa dan terikat pada kotak variables. Pada ceklisan pilih Pearson Two-tailed dan Flag significant correlations lalu OK.



Correlations

		Motivasi	Minat	Prestasi
Motivasi	Pearson Correlation	1	,788**	,796**
	Sig. (2-tailed)		,002	,002
	N	12	12	12
Minat	Pearson Correlation	,788**	1	,908**
	Sig. (2-tailed)	,002		,000
	N	12	12	12
Prestasi	Pearson Correlation	,796**	,908**	1
	Sig. (2-tailed)	,002	,000	
	N	12	12	12

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

Berdasarkan pada hasil diatas menunjukkan bahwa :

1. Nilai sig motivasi dengan prestasi 0,002 < 0,05 yang artinya ada korelasi yang signifikan antara motivasi dengan prestasi. Hubungan minat dengan prestasi sig. 0,000 < 0,005 yang mengindikasikan ada korelasi yang signifikan antara minat dengan prestasi.
2. Nilai r hit motivasi dengan prestasi 0,796 > r tabel 0,576. Disimpulkan ada korelasi antara motivasi dengan prestasi. Kemudian r hit minat dengan prestasi 0,908 > r tabel 0,576. Disimpulkan ada korelasi antara minat dengan prestasi. Sebab r hit bernilai positif sehingga hubungan dua variabel atau motivasi dan minat yang terus meningkat akan mampu meningkatkan prestasi belajar.
3. Output pearson correlation ada dua tanda bintang. Ini menandakan ada korelasi pada variabel yang ada dengan signifikansi 1%.

DAFTAR PUSTAKA

Esa Unggul, U. (2018). *Korelasi dan Regresi Linier Sederhana*. Https://Lms-Paralel.Esaunggul.Ac.Id/. https://lms-paralel.esaunggul.ac.id/pluginfile.php?file=/193372/mod_resource/content/5/6_7450_esa155_042019_pdf.pdf

Ghozali, I. (2018). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 25* (9th ed.). Universitas Diponegoro. https://onesearch.id/Record/IOS2851.slims-19545

Ghozali, I. (2021). *Aplikasi Analisis Multivariat dengan Program IBM SPSS 26* (10th ed.). Badan Penerbit Universitas Diponegoro. http://imamghozali.com/produk-78-.html

Hidayat, A. (2022). *Regresi Linear Sederhana dengan SPSS*. Https://Www.Statistikian.Com/. https://www.statistikian.com/2012/08/regresi-linear-sederhana-dengan-spss.html

Kemendikbud. (2023). *Korelasi Linier Sederhana dan Regresi Linier*. Https://Lmsspada.Kemdikbud.Go.Id/. https://lmsspada.kemdikbud.go.id/pluginfile.php/651977/mod_resource/content/2/BAB 5. Korelasi Linear Sederhana dan Regresi Linear.pdf

Raharjo, S. (2021). *Cara Melakukan Analisis Korelasi Bivariate Pearson dengan SPSS*. Https://Www.Spssindonesia.Com/. https://www.spssindonesia.com/2014/02/analisis-korelasi-dengan-spss.html

Universitas Negeri Yogyakarta. (2022). *Analisis Korelasi dan Regresi*. Https://Pendidikan-Akuntansi.Fe.Uny.Ac.Id/. https://pendidikan akuntansi.fe.uny.ac.id/sites/pendidikan-akuntansi.fe.uny.ac.id/files/Korelasi dan Regresi.pdf

Yuliara, I. M. (2016). *Modul Regresi Linier Sederhana*. Https://Simdos.Unud.Ac.Id/. https://simdos.unud.ac.id/uploads/file_pendidikan_1_dir/3218126438990fa0771ddb555f70be42.pdf

BAB 4

ANALISIS FAKTOR

Oleh Anna Sofia Atichasari, SE.,M.Si.,CMA.

4.1. Konsep Dasar Analisis Faktor

Suatu jenis studi yang disebut analisis faktor digunakan untuk menunjukkan fitur atau pola penting dalam suatu fenomena. Memadatkan kekayaan informasi dari sebagian besar variabel ke kumpulan faktor yang lebih mudah dikelola adalah tujuan dasar analisis faktor. Untuk mempelajari hubungan antara berbagai variabel dan untuk memperluas deskripsi variabel-variabel ini, analisis faktor, sebuah teknik statistik, dapat digunakan. Analisis faktor berusaha memadatkan data atau meringkas informasi yang ada di dalam kumpulan variabel asli ke dalam kumpulan variabel yang lebih kecil untuk mengungkap dimensi komposit atau faktor dengan jumlah informasi yang hilang paling sedikit.

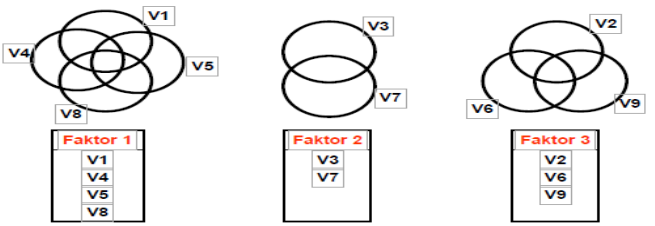
Setelah data dibersihkan, data tersebut kemudian dilakukan analisis faktor dengan menggunakan teknik rotasi varimax. Pendekatan varimax ortogonal bertujuan untuk mengurangi (membuat sesedikit mungkin) jumlah variabel yang memiliki loading tinggi pada satu komponen. Faktor-faktor yang dihasilkan oleh rotasi ortogonal tidak memiliki korelasi satu sama lain. Secara teori, analisis faktor juga digunakan untuk mereduksi data, yang merupakan proses meringkas banyak variabel menjadi jumlah yang lebih kecil dan menunjuk variabel-variabel tersebut sebagai faktor. Tingkat varians yang sepenuhnya menjelaskan semua makna yang ada pada data yang ada akan terungkap dari parameter-parameter ini. Ketika melakukan penelitian eksploratori (penelitian eksplanatori),

analisis faktor digunakan karena faktor-faktor yang mempengaruhi variabel belum ditentukan secara menyeluruh. Terdapat ketidakpastian mengenai faktor-faktor yang memengaruhi suatu variabel (penelitian eksplanatori). Analisis faktor juga dapat digunakan untuk menilai keandalan suatu kumpulan survei. Contohnya sebuah indikator akan rusak jika tidak mengelompok pada variabelnya, melainkan mengelompok pada variabel lain.

Pada intinya, analisis faktor berusaha untuk mengidentifikasi sejumlah komponen yang terbatas dengan karakteristik sebagai berikut:

1. Mampu menginterpretasikan variasi data sebaik mungkin.
2. Variabel-variabelnya independen satu sama lain.
3. Setiap faktor dapat diinterpretasikan dengan mudah.

Semua variabel akan diperiksa hubungannya (interdependensi) dalam analisis faktor ini, yang akan mengarah pada pengelompokan atau, lebih tepatnya, abstraksi dari banyak variabel menjadi hanya beberapa variabel atau faktor baru. Dengan beberapa komponen ini, akan lebih mudah untuk ditangani. Inilah yang disebut dengan Exploratory Research. Untuk mengelompokkan atau lebih tepatnya mengabstraksikan banyak variabel menjadi hanya beberapa variabel atau faktor baru, semua variabel yang sudah ada akan dilihat keterkaitannya (saling ketergantungan antar variabel) dalam analisis faktor ini. Akan lebih mudah untuk menangani beberapa variabel ini.

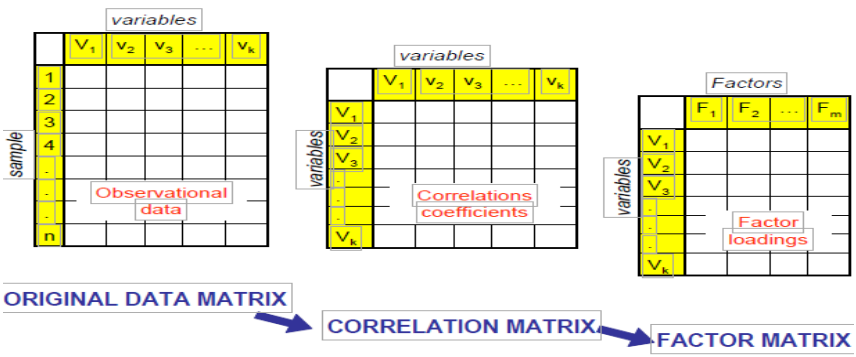


Gambar 4.1 Exploratory Research

(Sumber: Penulis, 2023)

Konsep pada Analisis Faktor yaitu sebagai berikut :

1. Melakukan Reduksi, Pengurangan dari banyak variabel menjadi beberapa variabel daripada mengkorelasikan variabel-variabel dependen dan independen.
2. Metode yang digunakan disebut teknik interdependensi, dan ini melibatkan melihat setiap satu set hubungan yang sama. Korelasi r = 1 dan r = 0 digunakan dalam teori ini. Hal ini digunakan untuk membedakan antara variabel yang memiliki korelasi dan yang tidak.
3. Jumlah variasi yang dikontribusikan oleh satu variabel kepada variabel lainnya dikenal sebagai "komunitas", menurut analisis faktor.
4. Setiap variabel akan memiliki faktor umum (jumlahnya sedikit) dan Faktor unik karena adanya kovariasi di antara variabel-variabel yang ditentukan (faktor tidak dapat dilihat dengan mudah).
5. Terdapat koefisien skor faktor, sehingga faktor 1 menjelaskan sebagian besar variabel dan faktor 2 menjelaskan sebagian besar variasi yang tersisa setelah faktor 1 diperhitungkan. Faktor 2 dan Faktor 1 tidak berhubungan..



Gambar 4.2 Prosedur Umum Dalam Analisis Faktor

(Sumber: Penulis, 2023)

Studi tentang saling ketergantungan antar variabel dikenal sebagai analisis faktor. Tujuannya adalah untuk mengidentifikasi

faktor-faktor umum di antara variabel-variabel asli dengan mengidentifikasi sekumpulan variabel baru yang jumlahnya lebih sedikit daripada variabel-variabel asli. Selama analisis faktor. Pemrosesan menjadi lebih sederhana dengan membagi sejumlah besar variabel menjadi beberapa komponen dengan sifat dan kualitas yang sebanding. Proses pengelompokan melibatkan penghitungan korelasi dari kumpulan variabel, menempatkan variabel yang memiliki korelasi kuat pada satu komponen, dan variabel yang memiliki korelasi relatif rendah pada komponen lainnya.

Analisis saling ketergantungan multivariat mencakup analisis komponen, yang metode operasi utamanya adalah menggabungkan variabel-variabel yang berkorelasi ke dalam satu atau lebih faktor umum, di mana satu faktor tidak berkorelasi dengan faktor lainnya. Principal Axis Factoring (PAF) dan Principal Component Factoring (PCF) adalah dua teknik estimasi parameter yang umum digunakan dalam analisis faktor, sebuah komponen dari analisis saling ketergantungan multivariat, analisis faktor mengumpulkan data yang saling berhubungan ke dalam satu atau beberapa faktor umum, di mana faktor yang satu tidak berhubungan dengan faktor yang lain. Principal Axis Factoring (PAF) dan Principal Component Factoring (PCF) adalah dua teknik estimasi parameter yang umum digunakan dalam analisis faktor. Berikut ini adalah daftar langkah-langkah utama dalam analisis faktor:

4.1.1. Uji *Determinant of Correlation Matrix*

Jika determinan mendekati 0, matriks korelasi dianggap berada di antara variabel yang terhubung satu sama lain. Komputasi menghasilkan nilai Uji Determinan Matriks Korelasi sebesar 0,06 sebagai hasilnya. Angka ini sangat mendekati nol dan mulai mendekati matriks korelasi antar variabel terkait.

4.1.2.Kaiser Meyer Olkin Measure of Sampling (KMO)

KMO atau yang disebut dengan Kaiser Meyer Olkin Measure of Sampling merupakan ukuran seberapa jauh koefisien korelasi dan koefisien korelasi parsialnya satu sama lain. Jika jumlah kuadrat koefisien korelasi parsial antara semua pasangan variabel kecil dibandingkan dengan jumlah kuadrat koefisien korelasi, nilai KMO akan mendekati 1. Jika angka KMO lebih dari 0.5.146, maka dianggap memadai. Dan nilai KMO adalah 0.580, sesuai dengan temuan. Hasilnya, kriteria KMO terpenuhi karena memiliki nilai yang lebih tinggi dari 0,5.

4.1.3. Bartlett Test of Sphericity

Uji statistik yang disebut Bartlett Test of Sphericity digunakan untuk menentukan apakah variabel-variabel dalam populasi terkait atau tidak terkait satu sama lain. Dengan kata lain, matriks identitas, di mana setiap variabel memiliki korelasi yang tepat dengan dirinya sendiri (r = 1) namun tidak memiliki hubungan sama sekali dengan variabel lain (r = 0), ini yang disebut dengan matriks korelasi populasi. Jika matriks identitas telah dihasilkan dari matriks korelasi, maka matriks tersebut akan lolos uji Bartlett atau tidak. Hubungan antar variabel sangat penting dalam analisis factor, karena tujuan metode ini adalah menggabungkan sejumlah variabel menjadi satu komponen. Ketika matriks identitas adalah matriks korelasi, maka tidak ada hubungan antara variabel, sehingga analisis faktor tidak dapat dilakukan. Sekumpulan variabel dikelompokkan dengan menentukan korelasinya, menggabungkan variabel-variabel yang berkorelasi tinggi dengan satu faktor dengan variabel-variabel yang berkorelasi rendah dengan faktor lainnya.

Situasi yang membutuhkan penerapan analisis faktor antara lain sebagai berikut :

1. Mengidentifikasi/mengenali dimensi atau karakteristik yang mendasari hubungan sekelompok variabel.

2. Kumpulan variabel yang terhubung akan digantikan oleh satu set variabel baru yang tidak terlalu terkait (independen) dalam penelitian multivariat berikutnya, seperti analisis diskriminan dan analisis regresi berganda.
3. Untuk digunakan dalam studi multivariat tambahan, pilih dan identifikasi kelompok variabel yang penting dari kelompok variabel yang lebih besar.

Berikut ini adalah persamaan maupun rumus analisis faktor:

X1 = Ai1F1 + Ai2F2 +Ai3F3 +..... + ViUi

Keterangan :

Fi : variabel terstandart ke-I

Ai1 : koefisien regresi dari variabel I pada common faktor ke I

Vi : koefisien regresi terstandar dari variabel I pada faktor unik ke I

F : common faktor

Ui : variabel unik variabel ke-I

M : jumlah common faktor

Karena korelasi adalah prinsip dasar dari analisis faktor, pernyataan berikut ini dibuat tentang pendekatan statistik korelasi:

1. Harus ada tingkat korelasi atau hubungan yang cukup antara variabel-variabel independen.
2. Kekuatan korelasi persial, yaitu korelasi antara dua variabel sementara semua faktor lain dianggap konstan.
3. Ukuran Measure Sampling Adequacy atau Barlett Test of Spericity digunakan untuk menilai keefektifan pengujian matriks korelasi.

Tahap selanjutnya yaitu melakukan prosedur analisis faktor setelah sampel diperoleh dan uji asumsi dilalui. Prosedur ini meliputi pengujian variabel-variabel yang akan diteliti serta

variabel-variabel terpilih dengan menggunakan Measure Sampling Adequacy dan Barlett Test of Spericity.

1. Melaksanakan langkah mendasar dari analisis faktor, yaitu membuat faktor atau beberapa faktor dari variabel yang telah lulus uji untuk variabel sebelumnya..
2. Menyediakan teknik rotasi faktor untuk variabel yang dibuat. Variabel yang masuk ke dalam elemen tertentu diperjelas dengan rotasi.
3. Interpretasi terhadap faktor atau faktor-faktor yang telah dikembangkan dan dianggap mewakili variabel-variabel komponen..
4. Validasi temuan faktor digunakan untuk memeriksa validitas faktor yang dibuat.

Semua variabel yang saat ini digunakan dimasukkan, dan berbagai macam tes kemudian dijalankan. Menurut penalaran pengujian, sebuah variabel akan memiliki korelasi yang cukup kuat dengan variabel lain jika variabel tersebut memiliki kecenderungan untuk mengelompok atau membentuk sebuah faktor. Sebaliknya, variabel yang memiliki korelasi yang lemah dengan faktor lain tidak akan dipertimbangkan untuk dimasukkan dalam penyelidikan selanjutnya. Angka KMO harus lebih dari 0,5 dan signifikan harus di bawah 0,05 untuk uji KMO dan uji Barlett.

Dengan kriteria yaitu MSA = 1, variabel dapat diprediksi secara akurat oleh variabel lain sedangkan angka uji MSA harus berada di antara 0 dan 1.

a. Jika MSA lebih dari 0.5, variabel masih dapat diprediksi dan diteliti lebih lanjut.
b. Namun apabila MSA kurang dari 0.5, tidak ada analisis atau prediksi lebih lanjut terhadap variabel.

Bartlett's sphericity merupakan uji statistik yang digunakan untuk memastikan apakah variabel-variabel dalam populasi tidak berkorelasi. Dalam contoh multivariat, uji Bartlett berusaha untuk memastikan apakah ada hubungan antara

variabel. Matriks korelasi antara variabel identik dengan matriks identitas jika variabel X1, X 2,... Xn adalah independen (saling bebas). Uji Barlett dan KMO. Hanya ukuran kecukupan Kaisaer Meyer Olkin (KMO) dan uji kebulatan Berett yang dapat digunakan untuk menarik kesimpulan yang valid secara statistik mengenai apakah analisis faktor dilakukan atau tidak. Kegunaan analisis faktor dipertanyakan oleh hasil uji KMO, dengan kisaran nilai antara 0,5 hingga 1,0. Analisis faktor (kesesuaian). Jika KMO lebih dari 0,5, analisis faktor dapat dilakukan; jika tidak, analisis faktor tidak dapat dilakukan.

Tahap berikuntnya yaitu membangun faktor untuk menemukan struktur yang mendasari keterkaitan antara variabel-variabel asli setelah variabel-variabel tersebut diidentifikasi dan dipilih serta kalkulasi korelasi sudah memenuhi persyaratan untuk analisis. ***Principal Component*** merupakan teknik yang banyak digunakan dalam analisis faktor eksplorasi.

4.2. Metode Principal Component

Menentukan struktur yang menjadi dasar dari variabel asli dalam penelitian dan mengurangi struktur set variabel pertama melalui reduksi data adalah tujuan khusus dari analisis faktor komponen utama. Tingkat variasi data diperhitungkan saat menghitung komponen utama. Seluruh varians matriks faktor serta angka 1 membentuk diagonal matriks korelasi. Ketika menentukan jumlah komponen yang paling tidak harus memperhitungkan variasi maksimum dalam data yang akan digunakan dalam analisis multivariat berikutnya, analisis principal component disarankan.

4.3. Kriteria Penentuan Jumlah Faktor

Analisis faktor berusaha untuk membuat lebih sedikit dalam analisis faktor daripada jumlah total variabel yang dipertimbangkan. Metode berdasarkan nilai eigen, presentasi

varians, dan scree plot digunakan untuk menentukan berapa banyak faktor yang diperoleh dalam investigasi ini. Berdasarkan nilai eigen, kriteria pertama yang digunakan. Nilai eigen menampilkan tingkat varians yang terkait dengan komponen tertentu. Variabel asli tidak lebih baik dari faktor dengan nilai eigen kurang dari 1, dan faktor dengan nilai eigen lebih dari satu akan disimpan. Ekstraksi berakhir ketika nilai eigen terakhir dengan nilai yang lebih tinggi atau sama dengan satu dipilih.

Berdasarkan persentase varians, kriteria kedua digunakan. Jumlah total variasi yang diperoleh digunakan untuk menghitung jumlah faktor yang akan digunakan. Ekstraksi faktor dapat dihentikan jika persentase total varians (lebih besar dari setengah total varians variabel awal) sudah memadai. Scree plot digunakan untuk menetapkan kriteria ketiga. Hubungan antara komponen dan nilai eigennya digambarkan pada grafik yang disebut scree plot. Kriteria ini dibuat dengan memplotkan nilai eigen versus jumlah item yang akan diekstrak. Sumbu horizontal menunjukkan jumlah komponen, sedangkan sumbu vertikal menunjukkan nilai eigen. Penurunan plot nilai eigen digunakan untuk menghitung jumlah elemen dalam kriteria ini. Ketika scree mulai mendatar atau rata dan nilai eigen berada pada nilai yang lebih tinggi dari satu dan kurang dari satu, maka terdapat titik penghentian untuk mengekstraksi jumlah elemen; pada saat ini, jumlah komponen yang dapat diambil diindikasikan.

4.4. Rotasi Faktor

Tujuan utama dari proses rotasi yaitu peningkatan interpretabilitas dan penyederhanaan faktor. Teknik rotasi ortogonal dan metode rotasi miring adalah dua pendekatan rotasi untuk analisis faktor yang masih dieksplorasi oleh beberapa akademisi. Rotasi ortogonal merupakan rotasi yang membuat sumbu yang bersangkutan tegak lurus atau sejajar dengan sumbu lainnya. Karena sumbu-sumbu tegak lurus satu sama lain setelah rotasi ini, maka setiap komponen tidak bergantung pada komponen lainnya. Ketika mengurangi jumlah variabel tanpa memperhitungkan signifikansi elemen yang

diambil, rotasi ortogonal digunakan. Sumbu tegak lurus tidak lagi dipertahankan oleh teknik rotasi miring. Dengan rotasi ini, sumbu faktor tidak tegak lurus satu sama lain, oleh karena itu korelasi antara komponen masih dipertimbangkan. Untuk mendapatkan jumlah komponen yang signifikan secara teoritis, rotasi miring digunakan. Ada beberapa langkah analitis untuk teknik rotasi ortogonal, termasuk pendekatan quartimax, varimax, dan equimax. Menyederhanakan baris-baris matriks faktor adalah tujuan utama dari pendekatan rotasi kuartimax. Sebuah variabel akan memiliki muatan faktor yang tinggi pada salah satu faktor setelah dilakukan rotasi nilai muatan faktor, sedangkan faktor lainnya dikurangi seminimal mungkin. Tujuan dari teknik rotasi ini adalah untuk membuat struktur baris matriks menjadi lebih sederhana. Metode ini tidak sering digunakan oleh para peneliti karena tidak dapat digunakan untuk menghasilkan struktur yang jelas. Titik rotasi pada akhirnya akan dikalahkan oleh aspek yang terlalu mencakup semuanya dari strategi ini.

Pendekatan varimax memusatkan analisisnya pada penyederhanaan kolom-kolom matriks faktor. Sebuah kolom dapat disederhanakan sampai batas maksimum jika hanya berisi nilai 0 dan 1. Pada setiap kolom matriks, pendekatan ini memiliki kecenderungan untuk menghasilkan beberapa nilai muatan faktor yang tinggi (mendekati -1 atau +1) serta beberapa nilai muatan faktor yang mendekati 0. Jika korelasi antara elemen atau variabel adalah +1 atau -1, yang menunjukkan hubungan sempurna yang positif atau negatif, logika penafsiran akan lebih sederhana.

Hubungan yang sangat buruk diwakili oleh nilai 0. Untuk membuat nilai faktor loading yang besar atau faktor lain sekecil mungkin, pendekatan varimax digunakan. Jika dibandingkan dengan struktur sebelumnya, struktur ini lebih sederhana. Pendekatan quartimax dan varimax digabungkan dalam metode equimax. Tujuan dari pendekatan ini adalah untuk merampingkan baris atau kolom matriks faktor. Pada tahap awal pengembangannya, strategi ini tidak populer atau sering

digunakan. Pendekatan varimax dari daftar di atas akan diterapkan dalam investigasi ini.

4.5. Interpretasi Hasil Analisis Faktor

Proses interpretasi melibatkan penjelasan pola-pola deskriptif, menempatkan analisis yang telah dilakukan sebelumnya ke dalam perspektif, dan mencari hubungan dan keterkaitan di antara deskripsi data yang telah dikumpulkan. Memberi nama faktor yang direduksi dan menghitung skor faktor jika tujuannya adalah untuk mereduksi data. Membandingkan nilai faktor loading pada variabel-variabel dalam faktor yang dibuat untuk mendapatkan nilai faktor loading untuk setiap variabel.

4.6. Kriteria penentuan signifikansi faktor loading

Tabel 4.1 Pedoman Mengidentifikasi Nilai Loading Faktor

Nilai Loading Faktor yang dianggap signifikan	Ukuran sampel yang dibutuhkan
0,3	35
0,3	25
0,4	20
0,4	15
0,5	12
0,5	10
0,6	85
0,6	70
0,7	60
0,7	50

(Sumber: Penulis, 2023)

4.7. Pemberian Nama Faktor

Variabel-variabel yang diteliti membentuk masing-masing faktor setelah faktor tersebut terbentuk secara lengkap, dan setiap faktor kemudian diberi nama berdasarkan sifat-sifat yang dimiliki oleh masing-masing komponennya. Ketika memberi nama sebuah faktor, kita melihat apa yang mendukung dan

secara akurat menangkap karakteristik variabel asli yang dikumpulkan menjadi satu komponen. variabel-variabel yang dikumpulkan menjadi satu komponen. Menerapkan generalisasi pada variabel-variabel awal ini adalah tindakan yang dapat dilakukan.

4.8. Validasi Hasil Analisis Faktor

Menguji stabilitas analisis ini adalah langkah terakhir dalam analisis faktor. Konfirmasi temuan faktor adalah nama konvensional untuk pengujian ini. Validitas hasil analisis faktor dalam penelitian ini diperiksa pada tahap ini dengan membagi sampel penuh menjadi dua bagian yang sama. Metode analisis faktor yang sama pada metode Principal Component yang kemudian digunakan untuk melakukan validasi pada setiap komponen sampel. Temuan validasi ditafsirkan sebagai menunjukkan bahwa jumlah faktor yang dibuat di kedua bagian sampel sesuai dengan jumlah faktor yang ditemukan dari analisis faktor yang dilakukan pada seluruh sampel.

DAFTAR PUSTAKA

Agus Eko Sujianto, *Aplikasi Statistik dengan SPSS 16.0*. Jakarta: Prestasi Pustaka Publisher, 2009

Imam Ghozali, Aplikasi Analisis Multivariatif dengan Program IBM SPSS 21 Update PLS Regresi. Semarang: Badan Penerbit Universitas Diponegoro, 2013

Jonathan Sarwono, Metode Penelitian Kuantitatif Dan Kualitatif. Yogyakarta: Graha Ilmu, 2006.

Mudrajat kuncoro,*Metode Riset untuk Bisnis dan Ekonomi*, Jakarta : Erlangga ; 2009.

Naresh K. Malhotra, *Riset Pemasaran: Pendekatan Terapan, Jilid 2*, Jakarta: Indeks, 2006.

Simamora Bilson, Analisis Multivariat Pemasaran, Jakarta: Gramedia Pustaka Utama, 2005.

Wiratmanto, Skripsi, Analisis Faktor dan Penerapannya dalam Mengidentifikasi Faktor- faktor yang Mempengaruhi Kepuasan Konsumen Terhadap Penjualan Media Pembelajaran.

Wiratna Sujarweni, *SPSS Untuk Penelitian*, Yogyakarta: Pustaka Baru Press, 2014.

BAB 5

ANALISIS KLASER

Oleh Dr. Hj. Utin Nina Hermina., SE., MS.i.

5.1. Pendahuluan

Pada penelitian kuantitatif kita mengenal ada analisis klaser atau nama lain dari analisis cluster, analisis klaser ini masuk dalam analisis multivariant dengan metode interdependensi, yang kegunaan dari analisis klaser tersebut di arahkan untuk meringkas data penelitian, bukan untuk menghitung keeratan hubungan antara variabel prediktor dengan kriterium. . (Santosa, 2016: 165). Analisis Klaser meneliti seluruh hubungan interdependensi yang tidak membedakan variabel bebas dan variabel terikat (independent variable and dependent variable). Perbedaan terhadap variabel tersebut umumnya terjadi dalam analisis regresi berganda, analisis varian, analisis diskriminan, yang mana kita dapat mengetahui pengaruh dari setiap variable bebas, secara individu maupun Bersama sama pada variabel tak bebas.

Pada analisis klaser digunakan untuk mengelompokkan objek dan kadang-kadang variabel kedalam kelompok yang homogen/relatif homogen yang disebut klaser . Objek di dalam klaster harus homogen atau relatif homogen yang memilikin kemiripan satu sama lain, sedangkan antar-klaser harus heterogen atau bervariasi, berbeda satu sama lain, klaser harus dibentuk berdasarkan data. Variabel di mana pengklaseran (clustering) dilakukan harus dipilih berdasarkan pada hasil penelitian sebelumnya, teori, hepotesis yang telah diuji, atau pertimbangan dari peneliti yang sudah berpengalaman. Ukuran jarak atau kemiripan (distance of similarity) yang tepat harus dipilih. Ukuran yang paling sering dipergunakan ialah jarak

yuklidean (*euclidean distance*) atau kuadratnya (jarak dipangkatkan dua). (Supranto, 2010:169).

Pada bab ini kita akan membahas langkah yang harus ditempuh dalam menggunakan analisis klaser, yang akan diilustrasikan dalam konteks pengklaser-an hierarki, yang dilanjutkan dengan pembahasan metode analisis klaser.

5.2. Tujuan Analisis Klaser

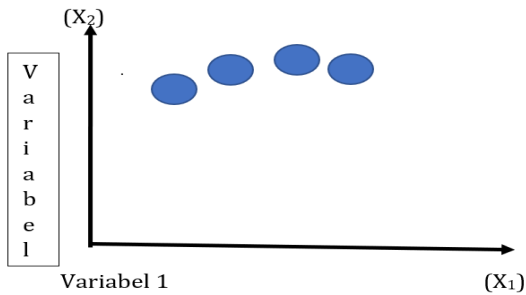
Analisis klaser memiliki tujuan :

a. Untuk mengklasifikasi objek dari penelitian seperti orang, produk, usaha, kedalam kelompok-kelompok yang relatif homogen berdasarkan variabel yang dipertimbangkan untuk diteliti. Obyek pada kelompok yang diteliti haruslah memiliki kesamaan dan berbeda jauh dengan obyek pada kelompok lainnya.
b. tujuan kedua adalah mengelompokan obyek-obyek berdasarkan karakteristiknya. Pengelompokkan tersebut berdasarkan obyek penelitian atau secara individu, yang setiap obyek yang memiliki karakteristik yang sama akan berada dalam klaser yang sama. yang memiliki kesamaan setiap klaser-klaser yang terbentuk akan mempunyai karakteristik yang cenderung sama (homogen) dan pengelompokkan tersebut berdasarkan variable-variabel yang akan di teliti (Usman dan sobari (2013) dalam S. Kartikawati : 2018:1) .

5.3. Analisis Klaser

Pada Metode analisis multivariat dikenal salah satunya adalah analisis klaser. Analisis klaser diartikan sebagai analisis yang digunakan untuk mengklasifikasi obyek atau observasi yang berbeda dalam suatu kelompok (klaster) sehingga memiliki kesamaan antara obyek dalam kelompok yang maksimal dan kesamaan dalam kelompok yang minimal. Analisis klaser dikenal juga sebagai analisis klasifikasi atau taksonomi numeric (numerical taxonomy), dimana setiap obyek hanya masuk ke

dalam satu klaser saja, tidak tumpang tindih dengan obyek yang lain. Ditunjukkan dengan gambar 5.1 dibawah ini.



Gambar 5.1 Pengklasteran Ideal

(Sumber : Suprapto : 2010: 143)

Keterangan pada gambar 4.1, memperlihatkan hasil klaser yang ideal. Yang diperlihatkan dari setiap obyek hanya masuk menjadi anggota salah satu klaser, tidak menjadi anggota dari dua atau lebih klaser. Gambar di atas juga menunjukan kondisi dimana klaser dipisahkan secara berbeda pada dua vaiabel. Pada variabel 1 (X_1) dan variabel 2 (X_2).

Analisis klaser yang tepat adalah jika ia memiliki kriteria;

a. yang homegenitas dalam kelompok (klaser) yang besar atau tinggi
b. Yang Heterogenitas antar kelompok (klaser) yang besar atau tinggi

5.3.1. Melakukan Analisis Klaster

Proses melakukan analisis klaster dapat dilakukan untuk memudahkan penelitian, sehingga peneliti akan sistematis didalam mengerjakan penelitiannya. Tujuan dilakukan anasisis klaser untuk membuat kelompok (klaser) yang diharapkan dapat memaksimalkan kemiripan didalam kelompok dan memaksimalkan perbedaan antar kelompok. Hal Pertama yang dapat dilakukan oleh peneliti dengan melalui perumusan

masalah yang kemudian dilanjutkan dengan melihat ukuran dan jarak dari masing-masing klaser tersebut. Ukuran jarak akan menentukan ketepatan dalam memilihdan menentukan kemiripan atau ketidak miripan dari obyek penelitian. Untuk menentukan banyaknya klaser yang diperlukan, memerlukan pertimbangan subyektif dari si peneliti, karena dari hasil perhitungan secara obyektif. Pemetaan dengan menggunakan analisis klaser akan diperoleh gambaran tentang kelompok-kelompok obyek dengan karakteritik semirip mungkin diantara grup sangat berbeda. (Solimun:2017:27)

Gambar 5.2 Prosedur Analisis Klaser

(Sumber : Supranto : 2010: 147)

Berikut ini kita akan uraikan langkah-langkah dalam pengklaseran; Perumusan masalah. Dimulai dari perumusan masalah pengklaseran dengan mendefiniskan variabel-variabel di telah dipergunakan untuk dasar pengklaseran. Bagian ini merupakan yang terpenting dalam perumusan klaser. Melakukan pemilihan variabel yang akan dipergunakan untuk

melakukan pengklaseran. Sebagai ilustrasi, berikut ini dilakukan pengelompokan penjual kuliner berdasarkan respon mereka terhadap integritas komunikasi pemasaran untuk meningkatkan penjualan secara online pada masa pandemi.

Berikut ini merupakan contoh kasus implementasi dari K-means klaser, pada bidang pemasaran. Yaitu analisis Integrated Marketing Communication dalam meningkatkan penjualan online. Penelitian ini menggunakan komunikasi pemasaran pada masa pandemi Covid 19 di tahun 2020. Peneliti ingin meneliti komunikasi pemasaran yang terintegrasi melalui lima variabelnya, yaitu Advertising, Sales Promotion, Personal Selling, Public Relations, dan Direct and Online Marketing. Yang diteliti adalah pemilik usaha kuliner yang pada masa pandemi masih dapat melakukan penjualannya dan mengalihkan usaha kuliner konvensional menjadi penjualan secara online.

Pada penelitian ini peneliti ingin mengetahui saluran komunikasi yang disajikan merupakan komunikasi pemasaran yang berbeda dengan komunikasi secara langsung, komunikasi pemasaran ini merupakan promosi yang dapat menggambarkan produk yang dijual, harga yang keliatan murah (dengan diskon khusus). Untuk itu peneliti ingin mengetahui factor apa saja yang dapat meningkatkan penjualan secara online.

Responden yaitu pemilik usaha kuliner, diminta mengisi kuesioner untuk mendapatkan jawaban mereka ke dalam suatu pertanyaan yang berbentuk skala yang diukur dengan 5 (lima) item pertanyaan menggunakan 5 point yang dimulai dari sangat tidak setuju (1) hingga sangat setuju (5). Akan ada point netral (3) sehingga akan ada 5 point, mulai dari sangat tidak setuju (1) hingga sangat setuju (5).

Adapun pertanyaannya adalah (Hermina:2021),

1. Variabel 1: Saya menggunakan media social untuk mengiklankan produk makanan
2. Variabel 2: Menggunakan iklan dimedia social sangat membantu dalam melakukan penjualan secara online
3. Variabel 3: Saya membuat promosi dengan penjualan sistem paket dengan harga yang lebih murah

4. Variabel 4: Saya melakukan promosi dengan menawarkan harga yang relative murah dibandingkan produk sejenis
5. Variabel 5: Saya memberikan insetif ongkir gratis untuk pelanggan pertama atau pelanggan setia
6. Variabel 6: Penjualan dengan direct & online marketing harus memperhatikan situs yang digunakan
7. Variabel 7: Saya memilih menggunakan media social dikarenakan real-time (mendapatkan respon langsung)
8. Variabel 8: Sistem Direct & Online Marketing merupakan sistem yang tepat dimasa pandemic ini
9. Variabel 9: Saya mengetahui Penjualan online karena pesan komunikasi pemasarannya
10. Variabel 10: Transaksi penjualan online tepat dimasa pandemic ini.

Tabel 5.1 Pengaruh Integrasi Komunikasi Pemasaran terhadap penjualan online

No	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10
1	5	5	5	4	4	5	5	5	4	5
2	4	5	5	4	5	5	5	5	4	5
3	5	5	5	5	3	5	4	5	3	5
4	5	5	5	3	4	5	4	5	4	5
5	5	5	3	4	3	5	4	5	4	5
6	5	5	4	3	4	5	4	5	4	5
7	4	5	2	3	5	5	3	5	4	4
8	5	5	5	4	5	4	5	5	4	4
9	5	5	5	4	5	4	5	4	4	4
10	5	5	3	4	3	3	5	5	3	5
11	5	5	4	4	5	5	4	5	4	5
12	5	4	5	4	5	5	4	5	4	5
13	5	5	5	4	5	5	5	4	5	5
14	5	5	4	4	5	4	4	5	5	5
15	4	5	4	4	5	5	4	5	5	4
16	4	4	4	4	5	5	4	5	5	4
17	5	4	5	3	4	3	4	5	4	4
18	5	4	5	4	4	5	5	4	4	5
19	5	5	2	4	4	3	4	5	3	5
20	5	4	5	4	4	5	4	4	5	5
21	5	5	5	5	4	5	4	5	5	4
22	5	5	5	5	4	4	4	4	5	5

No	Var 1	Var 2	Var 3	Var 4	Var 5	Var 6	Var 7	Var 8	Var 9	Var 10
23	4	5	5	5	4	4	5	5	4	5
24	5	5	5	5	4	4	5	4	4	4
25	5	5	4	5	3	3	4	5	5	4
26	5	5	4	5	4	5	5	5	4	5
27	5	5	5	3	4	5	4	4	5	4
28	5	5	5	4	4	5	5	4	5	5
29	5	5	5	5	4	5	4	4	5	5
30	4	5	5	5	4	4	5	4	5	4

(Sumber: Hermina, 2021)

Data diperoleh dari 30 orang sampel responden yang disajikan dalam table 4.1. perhatikan dalam praktik tersebut, pengklaseran dilakukan dalam sampel yang lebih besar 50 orang responden, penggunaan sampel n=30, dilakukan untuk ilustrasi. Analisis klaser digunakan untuk mengelompokkan responden yang memilih integrasi komunikasi pemasaran untuk meningkatkan penjualan secara online pada masa pandemik. Pengklaseran didasarkan pada variabel dari integrasi komunikasi pemasaran.

1. Memilih ukuran jarak (similaritas)
   Pada langkah yang kedua, yaitu memilih ukuran jarak, beberapa ukuran diperlukan untuk menentukan obyek yang mirip pada klaser yang sama. Dan pendekatan ini di nyatakan dalam jarak (distance) di antara pasangan obyek. Menggunakan ukuran jarak yang berbeda akan menghasilkan pengklaseran yang berbeda pula. Oleh sebab itu di anjurkan untuk memakai ukuran lain dan kemudian dibandingkan hasilnya.

2. Memilih suatu prosedur pengklaseran
   Pada tahapan ini, prosedur pengklaseran dapat mengunakan metode hierarki dan dapat juga non hierarki. Pengklaseran hierarki ditandai dengan pengembangan suatu hierarki atau struktur mirip pohon. Jenis prosedur pengklaseran kedua, non hierarki disebut juga K-means clustering. Metode ini meliputi *sequential threshold, parallel threshold* dan

optimising parttioning. *sequential threshold method*, pusat klaser dipilih dan semua obyek dalam suatu *prespesified threshold value* dari pusat, digabung bersama, dilanjutkan untuk suatu pusat klaser baru atau *seed* dipilih, dan proses diulangi, pada titik-titik yang belum diklaserkan

3. Penentuan banyaknya klaser
   Isu utama dalam analisis klaser yaitu dengan menentukan berapa banyaknya klaser. Berikut ini beberapa petunjuk yang dapat digunakan dalam menentukan banyaknya klaser
   a. Pertimbangan secara teoritis, konseptual, praktis
   b. Di dalam pengklaseran hierarki, jarak di mana klaster digabung bisa dipergunakan sebagai kriteria.
   c. Pada pengklaseran non hierarki, rasio jumlah varian dalam klaser dengan jumlah varian antar klaser dapat diplotkan melawan banyaknya klaser.
   d. Besarnya relatif klaser seharusnya bermanfaat.

4. Menginterprestasikan dan memprofil klaser
   Pada langkah kelima ini, dilakukan interprestasi dan memprofil klaser yang terdiri dari pengkajian mengenai centroids yaitu rata-rata nilai obyek yang terdapat pada klaser di setiap variabel. Nilai centroids dapat membuat kita menguraikan setiap klaser dengan cara memberikan nama atau label

5. Mengakses Validitas klaser
   Untuk mengakses kehandalan dan kesahihan, dan pemecahan pengklaseran sangat komplek karena harus melalui pengujian keandalan dan kesahihannya. Perlu dilakukan pertimbangan dalam analisis klaser, jangan sampai jika ada pemecahan pengklaseran tanpa seberapa penilaian tentang kehandalan dan kesahihannya.
   Prosedur berikut memberiokan pengecekan pada mutu hasil pengklaseran.
   a. Lakukan analisis klaser pada data yang sama dengan menggunakan ukuran jarak yang berbeda, lakukan

perbandingan dilintas ukuran untuk menentukan stablitas pemecahan.

b. Gunakan metode klaser yang berbeda dan perbandingkan hasilnya
c. Bagi data secara acak menjadi dua bagian
d. Hilangkan beberapa variabel secara acak dan lakukan pengklaseran yang didasarkan pada sisa variabel, bandingkan hasilnya dengan hasil pengklaseran yang asli
e. Pada klaser non-hierarki, pemecahan mungkin tergantung pada urutan obyek dalam seluruh data. Lanjutkan dengan multiple rund dan gunakan urutan obyek yang berbeda sampai penyelesaian menjadi stabil.

(Suprapto :2010:161)

5.4. Metode Analisis Klaser

Prosedur pengklaseran antara hierarki atau non hierarki mempunyai ciri sebagai pengembangan struktur yang mirip pohon bercabang (tree like structure). Berdasarkan metode dalam analisis klaser dapat dibagi menjadi dua jenis, yaitu :

5.4.1. Metode Analisis Klaser non Hierarki (tidak bertingkat)

Pada Analisis klaser non hieraki dilakukan proses pengklaseran secara langsung di semua observasi yang ada di dataset, sehingga dapat ditemukan klaser yang terjadi dalam satu level dan tidak bertingkat. Pada metode ini mengelompokkan obyek dalam kelompok sehingga jarak antara individu ke pusat kelompok minimum.

Pada metode ini dimulai dengan melakukan penentuan terlebih dahulu jumlah klaser yang diinginkan (dua, tiga , atau lebih), setelah itu dari jumlah klaser yang telah ditentukan, maka proses klaser dapat dilakukan tanpa mengikuti proses hierarki. Metode non hierarki sering disebut sebagai pengklaseran sebanyak k rata-rata (K-means Clustering). Metode ini bisa

dirinci sebagai *sequential threshold, paralel threshold dan optimizing partitioning.* (suprapto :2010 :170)

5.4.2. Metode Analisis Klaser Hierarki (bertingkat)

Sedangkan pada analisis klaser hierarki analisis dilakukan dengan memakai proses bertingkat. Pada klaser hierakri memiliki ciri-ciri ;

a. *Single Linkage*
   Merupakan metode dimana proses klaser yang didasarkan pada jarak terdekat diantara obyeknya, jika kedua obyek tersebut terpisah oleh jarak yang singkat, makan kedua obyek tersebut akan bergabung menjadi satu klaser, dan seterusnya.
b. *Complete Linkage*
   Pada metode complete linkage dilakukan berdasarkan jarak maksimum atau terjauh.
c. *Average Linkage*
   Metode Average menggunakan kriteria dalam mengukur jarak yaitu melalui merata-ratakan jarak seluruh individu dalam satu kelompok dengan jarak seluruh individu dalam kelompok lain.

5.4.3. Langkah Proses Analisis Klaser Hierarki

Pada proses analisis klaser hierarki dengan menggunakan pendekatan aglomeratif, yaitu ; (Husain :

1. Tentukan perhitungan jarak yang akan digunakan dengan dasar penggabungan klaser
2. Membuat definisi untuk setiap titik data (observasi) untuk menjadi satu klaser sendiri
3. Mengabungkan dua klaser yang terdekat yang memiliki kesamaan paling tinggi menjadi satu klaser
4. Melakukan pengulangan langkah no 2 dilakukan secara bertahap, sehingga mensisakan satu klaser terbesar yang berisikan keseluruhan dari titik data yang ada.

5.4.4. Tipe Analisis Klaser

1. Well separated cluster
   Didefinisikan sebagai sehimpunan titik yang memiliki kemiripan dengan titik lain dalam cluster daripada cluster lain.
2. Center-based clsuter
   Adalah sebuah cluster yang didalamnya memiliki anggota-anggota dengan kemiripan paling dekat dengan pusat cluster daripada pusat cluster lainnya.
3. Pada pusat cluster terdiri dari ;
   Centroid : rata-rata dari semua tituk dalam cluster
   Medoid : memilih titik sebagai titik tengah
4. Density-based cluster, merupakan area padat titik, yang dipisahkan dengan area kepadatan rendah, dari area kepadatan tinggi blainnhya. Teknik ini digunakan ketika cluster tidak teratur atau saling berkaitan, dan ketika noise dan outliers hadir. (Supranto : 2010 : 167)

DAFTAR PUSTAKA

Supranto.J, 2010. Analisis Multivariat arti dan inteprestasi. Jakarta. PT Rineka Cipta.

Santoso, Joko Agus, 2016. Statistik Multivariat. Yogyakarta. Penerbit Kepel Press Putri Arsita

Solimun,MS. Fernandes, Rinaldi, AA.,Nurjannah,. 2017. Metode Statistika Mutivariat. Malang. UB Press

Hussein, Saddam. 2021. Belajar Statistika, Machine Learning, geospasialis.com

Hermina, Utin Nina, 2021. Penelitian Analisis IMC (Integrated Marketing Communication) untuk meningkatkan Penjualan Online pada Masa Pandemi di Kota Pontianak

BAB 6

ANALISIS KONJOIN

Oleh Hartina Husain, S.Si., M.Stat.

6.1. Pendahuluan

Pada pertengahan tahun 1970-an, metode analisis konjoin mulai menarik perhatian dan sering digunakan dalam riset pemasaran. Pada umumnya, analisis konjoin merupakan analisis untuk mengetahui kecenderungan konsumen menentukan keputusan memilih suatu produk atau jasa berdasarkan karakteristik/fitur (atribut) yang dimiliki. Analisis ini pun membantu dalam mendapatkan komposisi atribut-atribut dari suatu produk melalui perilaku konsumen. Dengan demikian, analisis ini mampu mengidentifikasi mana atribut yang sifatnya dominan atau memiliki kontribusi lebih terhadap keputuan pembelian produk atau jasa (Widayat, 2018). Seiring perkembangannya, sejak tahun 1990-an aplikasi penggunaan analisis konjoin meningkat cukup signifikan pada berbagai bidang.

Memahami analisis konjoin bukanlah hal yang mudah. Melalui pemahama sebuah eksperimen, analisis konjoin jauh lebih mudah dan sederhana. Misalnya seorang ahli kimia pada perusahaan X yang bergerak pada pembuatan *body lotion* ingin mengetahui efek dari suhu dan tekenan terhadap viskositas (kekentalan) *body lotion* Analisis ANOVA (*Analysis of Variance*) seringkali digunakan dalam mengukur variabel tersebut. Sementara dalam kaitannya dengan tingkah laku manusia, seringkali dibutuhkan eksperimen menggunakan variabel atau faktor yang terkendali. Contohnya, apakah *body lotion* yang didesain oleh suatu perusahaan memiliki bau harum menyengat atau tingkat keharumannya biasa saja ? Bagaimana promosi

yang dilakukan pada produk berupa *body lotion* ? Apakah sama seperti promosi parfum, sabun, *deodorant* ? Berapa harga yang dipasarkan oleh produk *body lotion* tersebut ? Adanya variabel atau faktor terkontrol membuat analisis konjoin tepat digunakan pada contoh kasus tersebut. Analisis konjoin pada dasarnya terdiri atas presentasi konsumen terhadap serangkaian produk dengan atribut-atribut yang bervariasi . Dari beberapa pilihan yang diberikan, konsumen kemudian diminta untuk menentukan produk mana yang dipilih. Kemudian dari data yang dikumpulkan, analisis konjoin menggunakan teknik dalam statistika untuk mengetahui bagaimana preferensi konsumen pada kombinasi atribut yang berbeda. Dari hasil tersebut akan didapatkan atribut yang memberikan sumbangsi paling tinggi terhadap keinginan konsumen dalam membeli produk atau jasa . Oleh karenanya, pada proses analisisnya memberikan ukuran kuantitatif terhadap kepentingan relatif dan tingkat kepentingan setiap atribut pada suatu produk (Nugroho, 2008).

6.2. Model Analisis Konjoin

Analisis konjoin merupakan pengembangan teknik multivariat yang digunakan untuk memahami bagaimana perkembangan preferensi konsumen/responden terhadap suatu objek (produk atau layanan) . Hal ini didasarkan pada premis sederhana bahwa konsumen mengevaluasi nilai dari suatu objek berdasarkan jumlah kombinasi terpisah yang terbentuk dari kombinasi atribut-atribut (Hair *et al.*, 2010).

Model analisis konjoin sebagai bagian dari *multivariate dependence menthod* dapat dituliskan menjadi:

$$Y = X_1 + X_2 + X_3 + \cdots + X_k$$

Dimana

Y : Data metrik atau non-metrik

$X_1 + X_2 + X_3 + \cdots + X_k$: Data nonmetrik

Variabel dependen (Y) merupakan pendapat keseluruhan dari responden terhadap beberapa sekian level dan faktor pada sebuah produk atau layanan. Sementara itu variabel independen (X) dikatakan sebagai faktor berupa data nonmetrik yang tidak lain yaitu bagian dari faktor/level (Dwi Santoso, 2016).

Model dasar analisis konjoin untuk pilihan responden (c_i) untuk setiap atribut dan level atribut dapat dituliskan dengan persamaan berikut (Novi, Sigit and Jose, 2007):

$$c_i = \beta_0 + \sum_{r=1}^{s} p_{rk_{ri}}$$

Dimana :

c_i : *Utility total*

β_0 : Intersep model respon

$p_{rk_{ri}}$: Sumbangan *utility* dari atribut ke – r level ke-k_{ri}

s : Jumlah atribut

Adapun *utility functions* dapat diestimasi menggunakan beberapa faktor berikut ini :

1. Faktor Diskrit

$$\hat{p}_{rk} = \begin{cases} \hat{\alpha}_{rk} & ; k = 1,2,\cdots,m_j - 1 \\ -\sum_{r=1}^{m_j-1} \hat{\alpha}_{rk} & ; k = m_j \end{cases}$$

2. Faktor Linier

$$\hat{p}_{rk} = \hat{\beta}_r x_k$$

3. Faktor ideal-point

$$\hat{p}_{rk} = \hat{\gamma}_{r1} z_{rk} + \hat{\gamma}_{r2} z^2{}_{rk}$$

Prosedur yang paling sederhana dan sangat populer digunakan dalam mengestimasi model dasar analisis konjoin

yaitu *dummy variable regression.* Metode ini menjadikan variabel bebas sebagai variabel *dummy* dari atribut. Sementara itu, atribut yang terdiri atas level sebanyak k_r diberi kode dan dinyatakan dalam $k_r - 1$ variabel *dummy*.

Tata cara penilaian atribut menggunakan variabel *dummy* pada regresi linear berganda yaitu :

1. Jika data yang digunakan berasal dari penilaian rancangan stimuli sebelumnya yang menggunakan data metrik, maka regresi dengan variabel *dummy* diestimasi menggunakan metode *Ordinary Least Square* (OLS).
2. Jika penilaian stimuli menggunakan rangking stimuli, maka dilakukan terlebih dahulu transformasi data menjadi skala interval menggunakan *Multidemensial Scaling* (MDS) atau *monotonic regression* . Setelah itu barulah dilanjutkan regresi dengan variabel *dummy.*
3. Jika penilaian dilakukan secara terpisah pada tiap-tiap atribut (level atribut), maka variabel dependennya dinyatakan menjadi intesital pilihan atau aktual pembelian. Cara ini dikenal juga dengan sebutan *discret choice* yang menggunakan logit model dalam estimasinya.

Sebagai contoh yaitu pada atribut ke-r dengan k_r, dapat dituliskan variabel *dummy* seperti pada Tabel 6.1 berikut:

Tabel 6.1 Variabel Dummy Atribut ke – r dan level ke k_r

Level	X_1	X_2	$\cdots$	X_{k_r-1}
1	1	0	$\cdots$	0
2	0	1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$k_r - 1$	0	0	$\cdots$	1
k_r	0	0	$\cdots$	0

(Sumber: Penulis, 2023. Diolah)

Tabel 6.1 menunjukkan contoh variabel dummy dengan r atribut dan sebanyak k_r level. Adapun variabel *dummy* bernilai 1 jika level yang bersangkutan ada dan bernilai 0 jika tidak ada.

Dengan demikian, nilai $\hat{p}_{rk}$ dapat menggunakan persamaan regresi linear berganda dengan variabel *dummy* dalam proses estimasinya. Hal ini mengakibatkan persamaan regresinya menjadi:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_i X_{ki} + \varepsilon_i$$

6.3. Istilah dalam Analisis Konjoin

Istilah-istilah yang ada pada analisis konjoin sebagai berikut:

1. *Parth-worth functions*, juga dikenal sebagai *utility functions*, yaitu fungsi yang menggambarkan kontribusi setiap level atribut terhadap preferensi konsumen.
2. *Relative importance weight* yaitu nilai yang menggambarkan atribut mana yang lebih penting atau memiliki pengaruh yang lebih besar dalam hal keputusan pembelian oleh konsumen.
3. Level Atribut yaitu nilai yang menunjukkan tingkatan dari setiap atribut.
4. *Full profiles* atau *complete profiles* yaitu kombinasi atribut yang komprehensif sebagai bentuk representasi dari preferensi konsumen.
5. *Cyclical design* yaitu desain yang digunakan untuk mengurangi banyaknya pasangan kombinasi atribut yang harus dibandingkan.
6. *Factional factorial design* yaitu desain yang digunakan untuk mengurangi kompleksitas profil stimulus yang harus dievaluasi.
7. *Orthogonal arrays* yaitu kelas desain *factorial* yang digunakan untuk mengetahui efek utama dari masing-masing atribut secara efisien. Dengan demikian, jumlah

kombinasi yang perlu dievaluasi dapat dikurangi sehingga mampu menghemat usaha dan waktu.

8. *Internal validity* yaitu ukuran validitas yang merujuk sejauh mana suatu eksperimen atau studi dapat menyimpulkan hubungan sebab akibat antara atribut yang dievaluasi dengan preferensi konsumen secara akurat.

6.4. Metodologi Analisis Konjoin

Terdapat beberapa proses yang harus dilakukan dalam analisis konjoin sebagai berikut:

1. Menentukan atribut dan level atribut

Tahap awal ini dilakukan dengan mengidentifikasi atribut-atribut yang relevan mengenai produk atau layanan yang ingin diteliti. Adapun atribut-atribut tersebut merupakan faktor-faktor yang mempengaruhi preferensi atau keputusan konsumen terhadap suatu produk atau layanan.

Pemilihan level-level atribut ini sangat penting karena mencakup variasi yang memadai serta representatif dalam konteks produk ataupun layanan yang diteliti. Melalui penentuan atribut dan level atribut yang sesuai maka kita dapat memastikan bahwa aspek-aspek penting dari suatu produk atau layanan tercakup dalam analisis konjoin dan mampu memberikan pemahaman yang lebih baik mengenai preferensi konsumen.

Misalnya dilakukan analisis konjoin terhadap preferensi konsumen terhadap laptop diperoleh seperti Tabel 6.1. Dengan kata lain, menentukan atribut dan level dengan cermat adalah langkah kunci dalam memastikan bahwa analisis konjoin dapat memberikan pemahaman yang lebih baik tentang preferensi konsumen terhadap produk atau layanan yang di teliti.

Tabel 6.2 Contoh Atribut dan Level Atribut

Level Atribut	Atribut			
	Merek	Harga	Ukuran Layar	Kapasitas Penyimpanan
	Lenovo	Murah	Kecil (13 inci)	256 GB
	HP	Sedang	Sedang (15 inci)	512 GB
	Dell	Mahal	Besar (17 inci)	1 TB

(Sumber: Penulis, 2023. Diolah)

Tabel 6.2 menunjukkan atribut yang terdiri atas merek, harga, ukuran layar, dan kapasitas penyimpanan dimana masing-masing atribut memiliki level atribut. Sepeti pada atribut merek yang terdiri atas tiga level atribut yaitu Lenovo, HP, dan Dell. Kemudian pada atribut harga terdiri atas level murah,sedang, mahal, begitupun pada atribut lainnya terdiri atas 3 level berbeda.

2. Merancang matriks konjoin

Tahap ini merupakan langkah yang krusial untuk menyusun matriks yang berisi kombinasi antara atribut dan level atribut yang telah ditentukan pada tahap sebelumnya atau disebut juga dengan *full profiles.* Misalnya pada Tabel 6.2 terdiri atas 4 atribut dengan masing-masing 3 level atributnya. Dengan demikian, *full profiles* pada pada preferensi konsumen terhadap laptop sebanyak $3^4 = 81$ kombinasi atribut yang mungkin. Dalam hal ini, contoh *full profiles* yaitu :

Tabel 6.3 Contoh full profiles

Nomor	Merek	Harga	Ukuran Layar	Kapasitas Penyimpanan
1	Lenovo	Murah	Kecil (13 inci)	256 GB
2	Lenovo	Murah	Kecil (13 inci)	512 GB
3	Lenovo	Murah	Kecil (13 inci)	1 TB
4	Lenovo	Murah	Sedang(15 inci)	256 GB
5	Lenovo	Murah	Sedang(15 inci)	512 GB
6	Lenovo	Murah	Sedang(15 inci)	1 TB
⋮	⋮	⋮	⋮	⋮
81	Dell	Mahal	Besar(17 inci)	1 TB

(Sumber: Penulis, 2023. Diolah)

Tabel 6.3 menunjukkan pada nomor 1, kombinasi produk yang mungkin disukai oleh konsumen yaitu laptop lenovo harga murah dengan ukuran layar kecil (13 inci) dan kapasitas penyimpanan 256 GB. Kemudian pada nomor 2 yaitu kombinasi lainnya dimana konsumen mungkin menyukai laptop lenovo harga murah dengan ukuran layar kecil (13 inci) tetapi kapasitas penyimpanannya yaitu 512 GB.

3. Melakukan survei dan pengumpulan data

Tahap ini dilakukan survei kepada responden yang menjadi sampel representatif dari populasi yang ingin diteliti. Survei dilakukan dengan memberikan pertanyaan kepada responden dengan meminta mereka memilih produk mana yang paling disukai atau yang paling sesuai dengan preferensi mereka. Survei dapat menggunakan metode pengumpulan data berupa kuisioner, wawancara langsung ataupun daring, bergantung pada kebutuhan penelitian. Peneliti perlu memastikan bahwa instruksi dan pertanyaan yang diberikan kepada responden telah dipahami dengan jelas. Dengan demikian, data yang

diperoleh valid dan dapat diandalkan. Selain itu, pemilihan sampel representatif juga perlu dipastikan agar data yang dikumpulkan dapat mewakili preferensi konsumen secara keseluruhan. Sebagai contoh, pada Tabel 8.2 yang terdiri atas 81 kombinasi ini survei dapat diisi dimana responden memberikan rangking yaitu 1 sampai 81. Angka 1 untuk kombinasi yang paling disukai oleh responden dan angka 81 untuk kombinasi yang paling tidak disukai responden.

4. Melakukan analisis data

Pada tahap ini, pilihan responden dalam survei sebelumnya dianalisis dengan mengidentifikasi bobot dan preferensi relatif atribut-atribut yang diteliti. Dengan demikian kita dapat mengetahui kontribusi dari atribut-atribut tersebut dalam mempengaruhi preferensi konsumen. Hal ini dapat dilakukan dengan menggunakan *Parth-worth functions* atau *utility functions.* Utilitas memberikan ukuran terhadap gambaran tingkat keinginan atau preferensi relatif terhadap atribut atau level atribut tertentu.

5. Melakukan interpretasi hasil

Interpretasi hasil merupakan langkah terakhir dalam analisis konjoin yang berisikan hasil identifikasi atribut yang paling krusial berdasarkan analisis data. Hasil ini yang kemudian dijadikan sebagai pedoman dalam pengambilan keputusan, perancangan produk yang lebih sesuai, dan pengembangan strategi pemasaran.

6.5. Kegunaan Analisis Konjoin

Beberapa kegunaan analisis konjoin dikaitkan dalam konteks ekonomi dan bisnis yaitu (Rao, 2014) :

1. Mengidentifikasi preferensi dari konsumen
   Informasi mengenai preferensi konsumen mampu memberikan pemahaman yang lebih baik mengenai kebutuhan, keinginan, dan pemahaman yang lebih baik mengenai pasar. Hal ini memungkinkan suatu perusahaan dalam membuat perencanaan dan strategi pemasaran untuk

memenuhi permintaan konsumen secara maksimal. Begitupun pengembangan produk yang sesuai dengan dengan preferensi pasar mampu menigkatkan kepuasan konsumen akan produk atau layanan karena sesuai dengan ekspektasi mereka.

2. Mengetahui pentingnya atribut
   Analisis konjoin mampu menyajikan atribut yang paling berpengaruh dalam pengambilan keputusan konsumen. Perusahaan dapat berfokus pada peningkatan daya tarik produk atau layanan melalui atribut-atribut yang dominan mempengaruhi preferensi konsumen.
3. Melakukan pengembangan produk dan inovasi
   Hasil analisis konjoin memberikan wawasan baru untuk merancang produk atau layanan yang sesuai dengan preferensi konsumen. Perusahaan dapat membuat inovasi berupa kombinasi atribut yang diinginkan konsumen atau menambahkan fitur-fitur pada produk/layanan yang sudah ada sesuai preferensi pasar.
4. Segmentasi pasar
   Pasar dapat dibagi menjadi beberapa segmen seuai dengan karakteristik dan preferensi konsumen yang serupa. Dengan demikian, pasar luas yang heterogen dapat menjadi pasar yang jauh lebih kecil yang memiliki kebutuhan, keinginan, dan perilaku konsumen yang sama.
5. Pengambilan keputusan pemasaran
   Analisis ini dapat membantu perusahaan dalam mengoptimalkan strategi harga, merancang promosi, dan mengambil keputusan yang berpotensi dalam meningkatkan keberhasilan pemasaran produk.

Sementara itu, menurut dewanti (2007) analisis konjoin dapat digunakan sebagai :

1. Membuat rancangan harga
2. Melakukan prediksi mengenai tingkat penjualan ataupun penggunaan suatu produk
3. Melakukan uji coba konsep produk baru
4. Membuat segmentasi preferensi konsumen

5. Merancang strategi promosi
6. Mengembangkan produk baik itu produk yang baru maupun pembaharuan produk yang sudah ada sebelumnya.

DAFTAR PUSTAKA

Dewanti, paramita (2007) 'Penerapan regresi peubah dummy pada analisis konjoin(conjoint analysis) dalam pengembangan produk baru'.

Dwi Santoso, A. (2016) 'Analisis konjoin terhadap preferensi pengguna layanan perpustakaan universitas negeri semarang', *Skripsi*, 1(1), pp. 44–47.

Hair, J. *et al.* (2010) 'Multivariate Data Analysis.pdf', *Australia : Cengage*, p. 758.

Novi, S., Sigit, N. and Jose, R. (2007) 'Analisis Konjoin Preferensi Mahasiswa Dalam Pemilihan Flashdisk', *Jurnal Statistik* [Preprint], (1998). Available at: http://jurnal.untad.ac.id/jurnal/index.php/ejurnalfmipa/article/download/4000/2954.

Nugroho, S. (2008) *Statistika Multivariat Terapan*. UNIB Press.

Rao, V.R. (2014) *Applied conjoint analysis*, *Applied Conjoint Analysis*. Available at: https://doi.org/10.1007/978-3-540-87753-0.

Widayat (2018) *Statistika Multivariat (pada Bidang Manajemen dan Bisnis)*. UMMPress.

BAB 7

ANALISIS REGRESI LOGISTIK

Oleh Amalia Nur Chasanah, S.E., M.M.

7.1. Pendahuluan

Analisis regresi logistik merupakan satu diantara sekian pendekatan yang dipakai untuk membangun sebuah model prediksi. Namun, berbeda dengan regresi linear, pada regresi logistik ini menggunakan variabel dependen yang berstandar dikotomis. Skala dikotomis ini mengukur melalui dua tahapan, seperti : ya dan tidak, sukses dan gagal, bangkrut dan tidak serta lain sebagainya. Contoh kasus yang dapat diselesaikan dengan regresi logistik antara lain :

Prediksi perusahaan akan bangkrut atau tidak dengan menggunakan ukuran perusahaan dan rasio keuangannya, di nataranya:

a. Probabilitas seseorang terkena penyakit jantung dengan menggunakan ukuran kebiasaan merokok, kolesterol dan jenis kelaminnya.
b. Probabilitas seorang pelamar diterima pada perusahaan berdasarkan tingkat pendidikan, pengalaman kerja dan jenis kelamin.

Kasus yang ada sewajarnya bisa dibereskan dengan menggunakan penguraian diskriminan, tetapi opini distribusi normal mutivariat tidak bisa terpenuhi lantaran variabel independennya adalah kombinasi dari variabel metrik serta nonmetrik. Sehingga, analisis logistik menjadi pemecahan jika asumsi distribusi normal multivariat tidak dapat dipenuhi. Asumsi yang ada pada regresi logistik :

1. Variabel dependen dan independent tidak berhubungan linier.
2. Tidak ada asumsi normalitas multivariat.
3. Tidak ada asumsi homokedastisitas.
4. Variabel independent bisa berskala nonmetrik.
5. Variabel dependen harus berskala nominal.
6. Variabel independen tak seharusnya mempunyai jenis yang serupa antar variabel yang satu dengan yang lainnya
7. Kumpulan pada variabel independen wajib berkaraktersitik khusus
8. Sampel yang dibutuhkan dengan total yang relative besar

7.2. Model Persamaan Regresi Logistik

Model persamaan dalam regresi logistik sama seperti pada persamaan OLS yaitu : Y = B0 + B1X + e. e merupakan kesalahan varian atau residual. Dalam model persamaan logistik tidak memakai penafsiran yang sejenis dengan persamaan OLS. Persamaan yang sudah terjadi adalah :

$$\text{Ln}\left(\frac{p}{1-p}\right) = b0 + b1\,X1 + e$$

$$p = \frac{1}{1+e^{-(b0+b1X1)}}$$

exp atau ditulis “e” merupakan bentuk eksponen.

Pada model persamaan di atas, maka sukar menginterpretasikan koefisien regresinya seperti pada persamaan OLS karena probailitas p dan dan variabel X1 memiliki hubungan nonlinear.

Maka menggunakan nilai Odds Ratio atau yang sering kita dengar dengan singkatan Exp(B) atau OR.

Contoh :

nilai slope regresi sebanyak 0,8 sehingga Exp (B) adalah :

$2{,}72^{0{,}8} = 2{,}23$

Besarnya nilai Exp(B) dapat dijelaskan seperti di bawah ini:

Jika nilai Exp (B) merupakan pengaruh jenis kelamin pada keberhasilan melamar pada perusahaan yaitu sebanyak 2,23, sehingga dapat dijelaskan bahwa laki-laki lebih berpeluang diterima perusahaan dibandingkan perempuan. Interpretasi ini diartikan ketika terdapat tanda pada tiap variabel seperti di bawah ini:

1. Variabel bebas yang berjenis kelamin: 0 wanita dan 1 laki-laki
2. Variabel terikat keberhasilan diterima perusahaan: 0 untuk tidak diterima, 1 untuk diterima.

7.3. Metode Uji Regresi Logistik

Secara umum, terrdapat beberapa metode yang bisa digunakan dalam SPSS yaitu :

1. Simultan: Biasanya disebut dengan metode enter, yakni mencantumkan seluruh variabel bebas ke dalam pola dengan cara berbarengan.
2. Hirarki: Mencantumkan variabel dengan cara satu demi satu, diawali dengan mencantumkan variabel kontrol terlebih dahulu sebelum variabel prediktor.
3. Stepwise: Disebut forward conditional, merupakan teknik dimana variabel bebas yang disortir agar mendapatkan hasil model persamaan terbaik.

7.4. Uji Regresi Logistik Menggunakan SPSS

Diketahui penelitian yang berjudul Pengaruh tingkat pendidikan dan jenis kelamin terhadap Keberhasilan Pelamar pada perusahaan. Di mana variabel bebas ada dua yakni tingkat pendidikan dan jenis kelamin. variabel terikatnya yaitu kejadian keberhasilan pelamar pekerjaan.

Jenis kelamin kode 1 laki-laki, perempuan kode 0. Variabel keberhasilan diterima perusahaan, kode 1 diterima dan kode 0 tidak diterima. Data sebagai berikut:

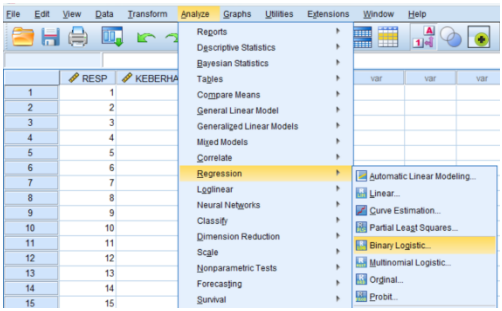
Tabel 7.1 Data Tingkat Keberhasilan Pelamar

Resp	Keberhasilan	Jenis Kelamin	Pendidikan
1	1	0	4
2	0	0	5
3	1	0	5
4	1	1	3
5	1	1	4
6	1	1	4
7	1	1	6
8	0	0	2
9	0	0	3
10	0	1	5
11	1	0	6
12	0	1	4
13	0	1	3
14	0	1	5
15	0	0	4
16	1	0	6
17	1	1	6
18	0	0	2
19	1	1	4
20	0	0	1
21	0	0	2
22	1	1	5
23	1	1	6
24	1	1	4
25	0	1	3

(Sumber : Penulis, 2023. Diolah)

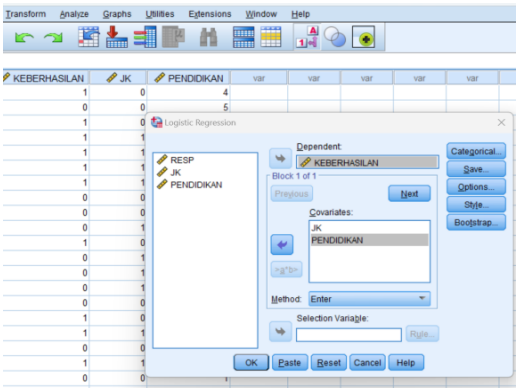
Langkah analisis:

1. Analyze, Regression, Binary Logistic



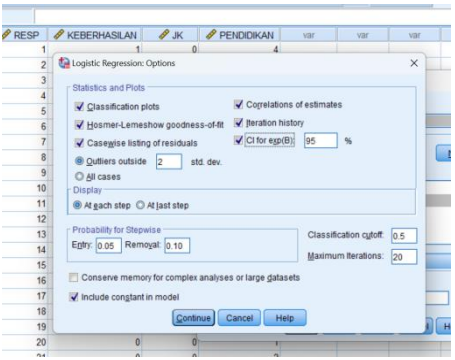
Gambar 7.1 Langkah Regresi

2. Masukkan keberhasilan ke dependen, dan JK dan pendidikan ke covariat



Gambar 7.2 Memasukkan Variabel

3. Pilih option, centang semua pilihan yang ada



Gambar 7.3 Option

4. Continue, OK

Hasil output dapat dilihat sebagai berikut :

Tabel 7.2 Case Processing Summary

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	25	100.0
	Missing Cases	0	.0
	Total	25	100.0
Unselected Cases		0	.0
Total		25	100.0

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

(Sumber: Data Primer diolah, 2023)

Output *Case Processing Summary* menerangkan bahwa seluruh peristiwa atau case menjadi sampel sejumlah 25, maka tidak ada sampel yang missing.

Tabel 7.3 Encoding Variabel Dependen

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

(Sumber : Data Primer diolah, 2023)

Output di atas menjelaskan bahwa data yang dipakai pada variabel dependen yakni keberhasilan, dimana kode : 0 adalah tidak berhasil dan kode : 1 adalah berhasil.

Block 0: Beginning Block

Tabel 7.4 Iteration History

Iteration History[a,b,c]

Iteration		-2 Log likelihood	Coefficients Constant
Step 0	1	34.617	.080
	2	34.617	.080

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 34.617

c. Estimation terminated at iteration number 2 because parameter estimates changed by less than .001.

(Sumber: Data Primer diolah, 2023)

Tabel 7.5 Tabel Klasifikasi

Classification Table[a,b]

Observed			Predicted		
			KEBERHASILAN		
			0	1	Percentage Correct
Step 0	KEBERHASILAN	0	0	12	.0
		1	0	13	100.0
	Overall Percentage				52.0

a. Constant is included in the model.

b. The cut value is .500

(Sumber: Data Primer diolah, 2023)

Tabel 7.6 Variable in the Equation

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.080	.400	.040	1	.842	1.083

(Sumber: Data Primer diolah, 2023)

Tabel 7.7 Variables Not in the Equation

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	JK	1.924	1	.165
		PENDIDIKAN	7.974	1	.005
	Overall Statistics		8.365	2	.015

(Sumber: Data Primer diolah, 2023)

Output di atas merupakan Blok 0 atau blok permulaan adalah proses inisialisasi artinya variabel JK dan Pendidikan belum dimasukkan ke dalam model penelitian. Dengan kata lain, model ini adalah model persamaan logistik yang hanya menggunakan konstanta saja untuk memprediksi responden masuk ke dalam kategori Berhasil atau tidak berhasil dalam melamar pekerjaan. Dari nilai signifikansi, diketahui konstanta

yang dihasilkan adalah 0.842 (> 0.05), hal ini berarti bahwa dengan menggunakan model persamaan sederhana (hanya konstanta saja) belum mampu memberikan penjelasan tentang keberhasilan dalam melamar pekerjaan. Selanjutnya dapat dilihat pada output Blok 1.

Block 1: Method = Enter

Tabel 7.8 Iteration History (Block 1)

Iteration History[a,b,c,d]

Iteration		-2 Log likelihood	Coefficients		
			Constant	JK	PENDIDIKAN
Step 1	1	25.384	-3.266	.524	.748
	2	24.709	-4.539	.789	1.012
	3	24.676	-4.902	.875	1.084
	4	24.676	-4.926	.881	1.088
	5	24.676	-4.926	.881	1.088

a. Method: Enter

b. Constant is included in the model.

c. Initial -2 Log Likelihood: 34.617

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

(Sumber: Data Primer diolah, 2023)

Tabel 7.9 Omnibus Test

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	9.941	2	.007
	Block	9.941	2	.007
	Model	9.941	2	.007

(Sumber: Data Primer diolah, 2023)

Tabel 7.10 Model Summary (Block 1)

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	24.676[a]	.328	.438

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

(Sumber: Data Primer diolah, 2023)

Dari tabel di atas, dapat dilihat bahwa model dengan memasukkan dua variabel independen ternyata telah terjadi perubahan dalam penaksiran parameter (-2 Log likelihood) sebesar 24.676. Jika dilihat nilai R-square sebesar 0.328 atau 32.8% (Cox & Snell) dan 0.438 atau 43.8% (Nagelkerke). Dengan demikian dapat ditafsirkan bahwa dengan dua variabel, yaitu JK dan Pendidikan maka keberhasilan dalam melamar pekerjaan dapat dijelaskan sebesar 43.8%.

Tabel 7.11 Hosmer and Lemeshow Test

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.344	6	.386

(Sumber: Data Primer diolah, 2023)

Tabel 7.11 adalah pengujian chi-square dari Hosmer and Lemeshow test. Tetapi pada implementasinya sudah dilaksanakan perubahan (modifikasi). Hipotesisnya yaitu :

H0 = Model sudah fit dalam menerangkan data (Goodness of fit)

H1 = Model tidak fit dalam menerangkan data

Kualifikasi pengujian:

Apabila skor pada p-value signifikansi (> 0.05) maka terima H0

Hasil pengujian chi-square yang didapatkan mempunyai nilai p-signifikansi sebanyak 0.3865 (> 0.05) sehingga terima H0. Kesimpulan dari pengujian ini adalah model telah fit dalam menerangkan data (goodness of fit).

GAMBAR JANGAN TERPOTONG KA

Tabel 7.12 Contingency Table

Contingency Table for Hosmer and Lemeshow Test

		KEBERHASILAN = 0		KEBERHASILAN = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	4	3.799	0	.201	4
	2	3	2.897	1	1.103	4
	3	1	1.279	1	.721	2
	4	1	2.117	4	2.883	5
	5	1	.748	1	1.252	2
	6	2	.595	1	2.405	3
	7	0	.335	2	1.665	2
	8	0	.231	3	2.769	3

(Sumber: Data Primer diolah, 2023)

Tabel 7.13 Tabel Klasifikasi (Block 1)

Classification Table[a]

Observed			Predicted		
			KEBERHASILAN		
			0	1	Percentage Correct
Step 1	KEBERHASILAN	0	8	4	66.7
		1	2	11	84.6
	Overall Percentage				76.0

a. The cut value is .500

(Sumber: Data Primer diolah, 2023)

Gambar jangan terpotong ka

Tabel 7.14 Variables in the Equation (Block 1)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	JK	.881	1.027	.736	1	.391	2.413	.322	18.063
	PENDIDIKAN	1.088	.478	5.181	1	.023	2.970	1.163	7.581
	Constant	-4.926	2.218	4.933	1	.026	.007		

a. Variable(s) entered on step 1: JK, PENDIDIKAN.

(Sumber: Data Primer diolah, 2023)

Kriteria uji: Tolak hipotesis nol (H0) jika nilai p-value signifikansi lebih kecil dari 0.05 (< 0.05) Dari tabel di atas merupakan tabel utama dari analisis data dengan menggunakan regresi logistik. Nilai p-value signifikansi variabel JK sebesar 0.391 > 0.05 maka menerima H0.

Dapat disimpulkan bahwa tidak terdapat pengaruh yang signifikan Jenis Kelamin terhadap Keberhasilan melamar pekerjaan dengan nilai koefisien pengaruh sebesar 0.881. Nilai p-value signifikansi variabel pendidikan sebesar 0.023 < 0.05 maka tolak H0 yang membuktikan bahwa terdapat pengaruh yang signifikan pendidikan terhadap keberhasilan melamar pekerjaan dengan nilai koefisien pengaruh sebesar 1.088. Model persamaan regresi logistik :

Keberhasilan = -4.926 + 0.881 JK + 1.088 Pendidikan

Interpretasi Analisis Regresi Logistik

Hasil persamaan regresi logistik tersebut tidak dapatdiinterpretasikan secara langsung dari nilai koefisiennya seperti regresi linier dengan OLS. Interpretasi dapat dilakukan dengan melihat nilai dari exp(B) atau nilai eksponen dari koefisien persamaan regresi yang terbentuk. Dari exp(B1) dapat dilihat bahwa Laki-laki (1) mempunyai kesempatan berhasil diterima 2.413 kali lebih dibandingkan dengan responden perempuan (0). Nilai exp(B2) sebesar 2.970 artinya bahwa semakin tinggi pendidikan akan menyebabkan perubahan sebesar 2.970 pada keberhasilan diterima perusahaan. Dengan demikian bahwa jika ada peningkatan pendidikan dari rendah ke tinggi akan meningkatkan probabilitas berhasil diterima perusahaan sebesar 2.970 kali.

DAFTAR PUSTAKA

Firdaus, M. (2018) *Ekonometrika Suatu Pendekatan Aplikatif* Jakarta: Bumi Aksara.

Ghozali, I. (2013) *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 21*. Semarang: BP UNDIP.

Gujarati, D. (2004) *Basic Econometrics*. New York: McGraw Hill, Inc.

Priyastma, R. (2020) *The Book of SPSS*. Yogyakarta: START UP.

Singgih, S. (1999) *SPSS Mengolah data statistik secara profesional*. Jakarta: PT. Elex Media Komputindo.

BAB 8

ANALSIS DISKRIMINAN

Oleh Mario Donald Bani, S.P., M. Biotech. (Adv.)

8.1. Pendahuluan

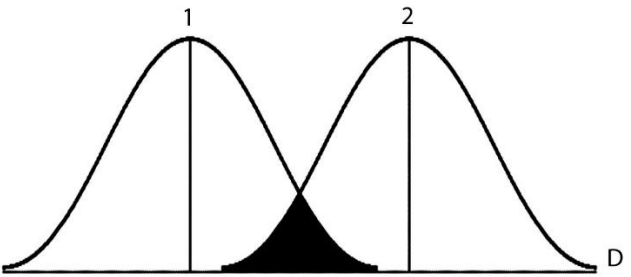
Analisis Diskriminan adalah salah satu metode analisis yang disediakan dalam ilmu statistika untuk mengukur pengaruh dari satu variabel terikat yang memiliki data kategoris (nominal dan ordinal) berdasarkan data yang ditampilkan oleh variabel bebas yang menggunakan skala numeris (rasio dan interval). Pada prinsipnya analisis diskriminan merupakan bagian dari analisis multivariat yang menggunakan *dependence method* dan mirip dengan regresi linier berganda, namun sebagaimana yang dijelaskan sebelumnya di atas, bahwa analisis diskriminan digunakan untuk menetapkan data kualitatif dengan menggunakan data kuantitatif (Hair *et al.*, 2010).

Model matematis dari analisis diskriminan ditunjukkan dalam persamaan sebagai berikut:

$D = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + ... + b_kX_k$ (Johnson and Wichern, 2007)
Dimana:
D = skor diskriminan
B = koefisien diskriminasi atau bobot
X = prediktor atau variabel bebas

Dalam analisis diskriminan nilai koefisien b diestimasi, agar nilai D dapat diperoleh. Pengelompokkan variabel terikat telah ditentukan terlebih dahulu (*apriori*), sehingga ketika rasio jumlah kuadrat antar kelompok telah mencapai skor

diskriminan maksimum, maka jumlah kuadarat setiap kelompok akan digunakan peneliti untuk menentukan kelompok-kelompok dari setiap variabel yang diukur (Gozali, 2018). Dengan kata lain pengelompokkan tersebut dilakukan berdasarkan nilai rata-rata kuadarat dari masing-masing variabel akan membentuk kurva yang saling berpotongan sebagaimana ditunjukkan pada Gambar 11.1 berikut ini.



Gambar 8.1 Kurva Diskriminan

(Sumber: Hair *et al.*, 2010)

Analisis diskriminan merupakan salah satu metode yang digunakan agar seorang peneliti dapat membedakan kelompok-kelompok dalam suatu populasi. Tujuan dari analisis diskriminan adalah mengelompokkan objek dalam kelas tertentu dengan menggunakan variabel terikatnya (D). Sebagai contoh terdapat seseorang yang ingin membeli sepeda motor dengan merek tertentu dan menggunakan harga, ukuran, berat, akselerasi, kecepatan hingga torsi mesin sebagai variabel bebas dan merk motor yang dipilihnya sebagai variabel terikat. Dalam menyusun penilaian dari seseorang tersebut, maka digunakanlah sebuah fungsi matematis yang kemudian membuat jarak antar kelompok semakin tinggi, sehingga dapat dijadikan acuan untuk mengelompokkan jenis-jenis motor yang akan dipilih.

Fungsi lain dari diskiminan dapat menggunakan fungsi diskriminan fisher, karena dapat menentukan nilai tengah dari jumlah kuadrat setiap kelompok (Johnson and Wichern, 2007).

Apabila terdapat dua kelompok 1 dan 2, maka secara matematis dapat ditulis sebagai berikut:

$D = \frac{n_1\overline{X}_2 + n_2\overline{X}_1}{n_1 + n_2}$ (Hair *et al.*, 2010)

Dimana:

D = Nilai kritis diskriminan

n_1 = jumlah objek dalam kelompok 1

n_2 = jumlah objek dalam kelompok 2

$\overline{X}_1$ = rata-rata nilai untuk kelompok 1

$\overline{X}_2$ = rata-rata nilai untuk kelompok 1

Analsisi diskriminan Fisher menitikberatkan fokusnya pada pemisahan variabel ke dalam bentuk kelompok tertentu di dalam sebuah populasi (Rahmadeni Rahmadeni, 2019). Model matematis dari Analisis Diskriminan Model Fisher, ditunjukkan pada persamaan berikut ini:

$$\sum_{m=1}^{r}(\mathrm{y_m} - \bar{\mathrm{y}}_{\mathrm{km}})^2 = \sum_{m=1}^{r}[\, a_{\mathrm{m}}^{\mathrm{T}}(x - \bar{x}_k)]^2 \;\leq\; \sum_{m=1}^{r}[\, a_{\mathrm{m}}^{\mathrm{T}}(x - \bar{x}_h)]^2$$

Untuk semua $h \neq k$; $r \leq s$; $s = min\ (h\text{-}1, p)$, maka:

y_m = skor diskriminan ke-m dari objek

$\bar{y}_m$ = rata-rata diskriminan ke –m dari kelompok ke –k (k = 1, 2, ..., h).

a_m^T = vektor koefisien fungsi diskriminan ke –m

x = vektor data pengamatan dari objek yang digolongkan.

$\bar{x}_k$ = vektor nilai rata-rata variabel pembeda dari kelompok ke-k.

r = banyaknya fungsi diskriminan yang digunakan dalam penggolongan.

s = banyaknya fungsi diskriminan yang mugkin dibentuk dalam analisis diskriminan.

h = banyaknya kelompok (populasi) yang dipelajari.

p = banyaknya variabel pembeda dalam analisis diskriminan.

(Suryanto, 1988).

Agar dapat diterapkan, maka analisis diskriminan harus memenuhi asumsi utama, antara lain:

1. Berdistribusi normal
2. Matriks kovarian antar kelompok harus sama
3. Tidak terjadi mulikolinearitas
4. Tidak ada data ekstrim (Gozali, 2018)

Dalam penerapannya, diketahui bahwa penggunaan teknik analsis diskriminan dapat dijadikan acuan peneliti dari berbagai keilmuan dalam melakukan kajian di bidang keilmuan lain, contohnya: seorang analis perusahaan jasa transportasi dapat melakukan analisis jenis kendaraan yang sesuai dan tidak sesuai berdasarkan data sekunder mengenai spesifikasi kendaraan tanpa harus menjadi ahli di bidang otomotif atau sejenisnya. Contoh lain yang sudah pernah dilakukan adalah dengan menerapkan analisis diskriminan untuk mengetahui tipe dari suatu UMKM dengan menggunakan omset, modal dan jumlah tenaga kerja sebagai variabel bebas (Sam, 2019).

Analisis ini juga dapat diterapkan untuk memprediksikan mana saja kelompok jenis kredit macet atau tidak berdasarkan variabel lama pinjaman dan jumlah gaji seseorang (Rahmadeni Rahmadeni, 2019). Pada intinya adalah analisis diskriminan dapat digunakan untuk berbagai macam situasi dan kondisi untuk mengelompokkan variabel tertentu dalam bentuk data kualitatif berdasarkan data kuantitatif yang diperoleh oleh peneliti. Agar mudah dipahami, maka Bab ini akan mengulas mengenai analisis diskriminan sehingga dapat dijadikan referensi bagi pembaca.

8.2. Korelasi Kanonik Diskriminan

Korelasi kanonis (*canonical correlation*). Analisis ini dilakukan untuk mengukur seberapa kuat hubungan setiap kelompok yang dibentuk berdasarkan nilai diskriminan. Analisis ini bertujuan untuk menentukan variabel mana saja yang memiliki hubungan yang paling kuat dengan variabel

terikat untuk membentuk kelompok dari skor diskriminan yang diperoleh. Model matematis dari analisis korelasi kanonik adalah:

$R_{cm} = \sqrt{\lambda_m/(1 + \lambda_m)}$ (Andri and Goejantoro, 2014)

Dimana:

R_{cm} = korelasi kanonik antara fungsi diskriminan ke-m dan (h-1) variabel *dummy* yang membatasi anggota-anggota dari h kelompok.

λ_m = *eigenvalue* ke-m, m = 1,2...,s,

s = min (h-1, p), h = banyaknya kelompok

p = banyaknya variabel pembeda dalam fungsi diskriminan.

8.3. Nilai Tengah dan Pembatas Kelompok

Nilai tengah merupakan nilai tengah rata-rata skor diskriminan untuk setiap kelompok yang dibentuk. Jadi jumlah nilai tengah sama banyak dengan jumlah kelompok yang dibentuk. Nilai tengah dilambangkan dengan C yang berarti *centroid*. Rumus yang digunakan untuk menghitung *centroid* adalah sebagai berikut:

$$C_n = \frac{\sum D_k}{n_k}$$

Nilai tengah atau *centroid* yang diperoleh digunakan sebagai acuan dalam mengelompokkan objek, yaitu dengan cara menentukan nilai *cutting score* atau *cut-off*. Nilai *cutting score* digunakan untuk menentukan batas pemisah antar kelompok. Contohnya, apabila terdapat dua kelompok dengan nilai *cutting score* 0,20, maka jumlah anggota suatu kelompok 1 dam 2 ditentukan dengan melihat apakah skor diskriminan dari masing-masing anggota kelompok tersebut berada di atas atau bawah nilai *cutting score*. Dalam menentukan nilai *cutting score*

atau *cut-off*, maka nilai *centroid* akan digunakan dengan rumus sebagai berikut:

$cut - off = \frac{n_1 C_1 + n_2 C_2}{n_1 + n_2}$ (Andri and Goejantoro, 2014)

8.4. Model Analisis Diskriminan

Agar lebih mudah melakukan analisis diskriminan, maka peneliti pada umumnya menggunakan pendekatan matriks untuk menetapkan klasifikasi yang membedakan variabel dalam sebuah kelompok dan antar kelompok. Hal tersebut dilakukan untuk mendapatkan fungsi diskriminan berdasarkan nilai diskriminan yang diperoleh, sehingga perbedaan relatif antar kelompok terhadap perbedaan di dalam kelompok mencapai nilai tertinggi (Tatham, et.al, 1998). Berdasarkan model matematis analisis diskriminan, maka apabila terdapat data berukuran *n* yang diperoleh secara acak dari populasi *p* yang berukuran *n* = 1, 2, ..., *k,* maka berdasarkan nilai *p* pada setiap variabel adalah X_1, X_2, ..., X_p, dapat membentuk matriks data dalam bentuk *p* x n_k dari populasi p_k yang ditunjukkan melalui persamaan sebagai berikut:

$$\underset{(pxn_k)}{X} = \begin{bmatrix} x_{1n1} & x_{1n_2} & \cdots & x_{1n_k} \\ x_{2n_2} & x_{2n_2} & \cdots & x_{2n_k} \\ \vdots & \vdots & \cdots & \vdots \\ x_{pn_1} & x_{pn_2} & \cdots & x_{pn_k} \end{bmatrix}$$ (Johnson and Wichern, 2007)

Berdasarkan model matematis analisis diskriminan, maka model analisis diskriminan didefinisikan dalam bentuk persamaan $Y=a^TX$, sehingga perbedaan antar kelompok dari *Y* adalah a^TBa dan perbedaan dalam kelompok *Y* adalah a^TWa. Oleh karena itu, nilai λ didefinisikan sebagai tingkat perbedaan relatif antara kelompok terhadap perbedaan yang berada di dalam kelompok yang kemudian ditunjukkan melalui persamaan berikut ini:

$\lambda = \frac{a^T Ba}{a^T Wa} \leftrightarrow \lambda\, aT\, Wa \;=\; aT\, Ba$ (Gaspersz, 1995)

Berdasarkan persamaan di atas, maka masalah yang selanjutnya dalam analisis diskriminan adalah penentuan vektor pembobot *a* yang mencapai nilai λ tertitinggi. Berdasarkan aturan matematis, agar nilai vektor pembobot *a* dapat diperoleh untuk menjadikan nilai λ mencapai taraf tertinggi, maka berlaku rumus:

$$\frac{\partial \lambda}{\partial a} = \frac{\{(2Ba)(a^T Wa) - (2Wa)(a^T Ba)\}}{(a^T Wa)^2}$$

$$\frac{\partial \lambda}{\partial a} = \frac{2\{(Ba)(a^T Wa) - (Wa)(a^T Ba)\}}{(a^T Wa)^2}$$

$$\frac{\partial \lambda}{\partial a} = \frac{2\{(Ba)(a^T Wa) - (Wa)(\lambda a^T Wa)\}}{(a^T Wa)^2}$$

$$\frac{\partial \lambda}{\partial a} = \frac{2(Ba - \lambda Wa)}{(a^T Wa)} = 0$$ (Andri and Goejantoro, 2014)

Hasil penurunan rumus di atas dapat digunakan untuk menentukan vektor koefisiensi diskriminan, yang secara singkat disajikan melalui persamaan sebagai berikut:

$(B - \lambda W)\, a=0$ (Johnson and Wichern, 2007)

Dengan mengasumsikan matrik *W* adalah *non-singular* agar dapat ditentukan nilai inversnya, yaitu W^{-1}, maka:

$(W^{-1} B - \lambda I)a = 0$ (Gaspersz, 1995)

Pada dasarnya koefesien pembobot fungsi diskriminan a^T diperoleh melalui persamaan sebagai berikut:

$(B - \lambda W)$a = $(W^{-1} B - \lambda I)a = 0$ (Andri and Goejantoro, 2014)

Dimana:
B = matrik kovarian antara kelompok
W = matrik kovarian dalam kelompok

λ = akar ciri (*eigenvalue*)
a = vektor ciri

Apabila dipenuhi syarat determinan adalah $|W^{-1} B - \lambda I| = 0$, apabila persamaan sebelumnya di atas diberi batasan $a^T Wa = 1$, maka akan ditemukan solusi dari persamaan sebelumnya, sehingga menjadikan nilai $\lambda = a^T Ba$. Dengan demikian, maka nilai maksimum λ dapat memberi gambaran yang jelas mengenai tingkat perbedaan yang masimum antar kelompok.

8.5. Uji Normalitas Diskriminan

Uji normalitas data merupakan syarat utama dalam melakukan analisis diskriminan dan harus dilakukan pada semua variabel secara bersamaan. Namun demikian, uji normalitas dapat juga dilakukan per variabel, dengan asumsi bahwa apabila setiap variabel berdistribusi normal, maka seharusnya data semua variabel juga demikian.

Dalam melakukan uji normalitas, maka peneliti menggunakan asumsi statistika bahwa:

1. Jika nilai P-Value lebih besar dari nilai alpha yang ditentukan, maka data yang diperoleh berdistribusi normal.
2. Sebaliknya, jika nilai P-Value lebih kecil dari alpha yang ditentukan, maka data yang diperoleh tidak berdistribusi normal.

Diketahui pula, apabila jumlah data yang diperoleh kecil atau jumlah sampel yang ditentukan terlalu kecil, maka analisis akan sulit mencapai kenormalan ganda. Oleh karena itu, analisis yang harus dilakukan adalah dengan melakukan uji vektor rata-rata, apabila data yang diperoleh memenuhi asumsi bahwa nilai uji kovarian data berdasarkan matriks kovariansi yang ditampilkan dari nilai *Log Determinants* tidak berbeda jauh, sehingga dianggap bahwa matriks kovarian relatif sama untuk kedua kelompok (Mattjik and Sumertajaya, 2011).

8.6. Matriks Kovarian

Variansi merupakan ukuran perbedaan yang diperoleh dari nilai tengah deviasi kuadratik dari nilai tengah sesungguhnya data. Nilai variansi atau perbedaan akan menunjukkan ukuran penyebaran data. Nilai kovarians akan menunjukkan tingkat perbedaan dari dua atau lebih variabel data yang diambil secara acak, sehingga menjadikan data tersebut lebih mudah untuk dikelompokkan.

Berdasarkan data yang diperoleh, maka matriks kovarian disusun berdasarkan pada asumsi dasar bahwa $X = (X_1, .., X_p)^T$ adalah sebuah vektor p dari variabel acak, yang disebut juga vektor acak. Setiap anggota dari variabel X merupakan suatu variabel acak yang memiliki keragaman X_1 (j = 1, …, p) dengan nilai $\mu_j = E[X_j]$ dan varians $\sigma_j^2 = E[(X_j - \mu_j)]$. Nilai harapan X adalah sebagai vektor dari nilai-nilai yang diharapkan dari komponennya, yaitu:

$$E[X] = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu \quad \text{(Andri and Goejantoro, 2014)}$$

Matrik kovarian populasi X ditunjukkan melalui persamaan sebagai berikut:

$$Cov[X] = E\left[(X_j - \mu_j)(X_k - \mu_k)\right]$$

$$Cov[X] = E\left\{ \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \pi_2 \\ \vdots \\ X_P - \mu_P \end{bmatrix} [X_1 - \mu_1 \quad X_2 - \mu_2 \quad \cdots \quad X_P - \mu_P] \right\}$$

$$Cov[X] \begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \dots & (X_1 - \mu_1)(X_P - \mu_P) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \dots & (X_2 - \mu_2)(X_P - \mu_P) \\ \vdots & \vdots & \dots & \vdots \\ (X_P - \mu_P)(X_1 - \mu_1) & (X_P - \mu_P)(X_2 - \mu_2) & \dots & (X_P - \mu_P)^2 \end{bmatrix}$$

Untuk j = 1, 2 . . . ,p dan k = 1, 2, . . . ,p (Johnson and Wichern, 2007)

Unsur-unsur diagonal sama dengan $E\left[(X_j - \mu_j)^2\right] = \sigma_j^2$. Unsur-unsur *off-diagonal* sama dengan $E\left[(X_j - \mu_j)(X_k - \mu_k)\right] = Cov$ (X_j, X_k) = σ_{jk} adalah kovarian atau peragam antara variabel X_j dan X_k. Berdaarkan pernyataan tersebut, maka matrik kovarian ditunjukkan pada persamaan sebagai berikut:

$$Cov\ (X) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{p1} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}$$ (Johnson and Wichern, 2007)

Uji matriks kovaruan pada umumnya menggunakan uji *Box's M* yang menggunakan pendekatan *Chi-square*, dimana derajat bebasnya adalah $\frac{1}{2}(k-1)k(p+1)$ dengan p merupakan jumlah variabel bebas dan k merupakan jumlah kelompok yang terbentuk. Uji hipotesis yang digunakan dalam analisis ini adalah:

$H_0 = \Sigma_1 = \Sigma_2 = \dots = \Sigma_k$, yang berarti jumlah matriks kovarian yang terbentuk adalah sama

$H_1 = \Sigma_1 \neq \Sigma_2 = \dots = \Sigma_k$, yang berarti terdapat satu atau lebih matriks kovarian yang berbeda antar kelompok.

Model persamaan yang digunakan untu menguji hipotesis di atas adalah:

C = (1-u) M (Johnson and Wichern, 2007)

$C = (1-u) \sum_{i=1}^{k}(n_1 - 1)ln|S| - \sum_{i=1}^{k}(n_1 - 1)ln|S_i|$

Diaman, apabila nilai $C \geq \mathcal{X}^2_{\alpha,\frac{1}{2}(k-1)p(p+1)}$ berarti tolak H_0 dan terima H_1 dengan taraf kepercayaan α, karena terdapat matriks kovarian yang berbeda dari kelompok k yang terbentuk (Alifta Ainurrochmah; Memi Nor Hayati; Satriya, 2019).

8.7. Uji Vektor Rata-rata

Dalam melakukan uji vektor rata-rata, maka pendekatan *chi-square* (χ^2) dapat digunakan (Anonymous, 2005)(Andri and Goejantoro, 2014). Berdasarkan hal tersebut, maka asumsi yang digunakan dalam pengambilan keputusan adalah:

1. Jika nilai uji *V* lebih kecil atau sama dengan nilai tabel *chi-square* ($\chi^2_{a,v=p(h-1)}$, maka *H0* diterima karena tidak ada perbedaan nilai rata-rata antar kelompok, sehingga data yang diperoleh tidak dapat membenuk fungsi diskriminan.
2. Jika nilai uji *V* lebih besar dari nilai table *chi-square* ($\chi^2_{a,v=p(h-1)}$), maka *H0* ditolak karena terdapat perbedaan dari nilai rata-rata antar kelompok, sehingga data yang diperoleh dapat membentuk fungsi diskriminan.

8.8. Analisis Diskriminan Berganda

Analisis diskriminan berganda perlu dilakukan apabila data variabel terikat (Y) lebih dari satu, sehingga diperlukan fungsi diskriminan sebanyak (*n*-1) untuk *n* kategori, maka persamaan dibuat dalam bentuk matriks sebagai berikut:

$$\underset{(pxn_1)}{X^{(1)}} = (X_{11}, X_{12}, \dots., X_{1n_1})$$

$$\underset{(pxn_2)}{X^{(2)}} = (X_{21}, X_{22}, \dots., X_{2n_2}$$

Dari data matriks di atas dapat ditentukan vektor nilai rata-rata contoh dan matriks ragam peragam (variance-covariance) berikut:

$$\underset{(px1)}{\bar{X}_1} = \frac{1}{n_1}\sum_{j=1}^{n_1} X_{1j}$$

$$\underset{(px\,p)}{S_1} = \frac{1}{n_1 - 1}\sum_{j=1}^{n_1}(X_{1j} - \bar{X}_1)(X_{1j} - \bar{X}_1)'$$

$$\underset{(px1)}{\bar{X}_2} = \frac{1}{n_2}\sum_{j=1}^{n_1} X_{2j}$$

$$\underset{(px\,p)}{S_2} = \frac{1}{n_2 - 1}\sum_{j=1}^{n_2}(X_{2j} - \bar{X}_2)(X_{2j} - \bar{X}_2)'$$

Karena diasumsikan bahwa populasi induk memiliki peragam yang sama, yaitu Σ, maka matriks kelompok sampel digabung untuk memperoleh penduga bagi Σ melalui rata-rata terbobot berikut:

$$S_G \frac{(n_1 - 1)\ .\ S_1 + (n_2 - 1)S_2}{(n_1 + n_2 - 2)}$$

Pengujian perbedaan vektor nilai rata-rata di antara dua populasi dilakukan dengan jalan merumuskan hipotesis berikut:

H_0 : $\underline{U_1} = \underline{U_2}$; artinya vektor nilai rata-rata dari populasi 1 sama dengan dari populasi 2.
H_1 : $\underline{U_1} \neq \underline{U_2}$; artinya kedua vektor nilai rata-rata berbeda.

Pengujian terhadap hipotesis dilakukan menggunakan uji statistic T^2-Hotelling yang dirumuskan sebagai berikut:

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\underline{\bar{X}_1} - \underline{\bar{X}_2})' S_G^{-1} (\underline{\bar{X}_1} - \underline{\bar{X}_2})$$

8.9. Pembentuk Fungsi Diskriminan

Perbedaan satuan pengukuran dapat memberikan pengaruh terhadap besaran koefisien pembobot pada fungsi diskriminan, sehingga pengaruh tersebut perlu dihilamgkan dengan cara penentuan koefisien baku diskriminan dengan menggunakan persamaan sebagai berikut:

$a^*_{mi} = \sqrt{w_{ii}}\ a_m : i = 1,2, \dots, P$ (Gaspersz, 1995)
m = 1, 2, … ,s
s = min (h -1, p)

Dimana:

a^*_{mi} = koefisien diskriminan baku dari variabel ke-*i* dalam fungsi diskriminan ke-*m*.

w_{ii} = elemen diagonal utama dalam matrik *W*

a_{mi} = koefisien diskriminan tak baku dari ke-*i* dalam fungsi diskriminan ke- *m.*

8.10. Peranan Relatif Fungsi Diskriminan

Setiap fungsi diskriminan yang dibentuk berperan untuk mengklasifikasikan setiap data dalam kelompok-kelompok yang sudah ditentukan sebelumnya. Diaman pembagian data-data tersebut diukur dengan menggunakan persentasi relatif *eigenvalue* yang disajikan melalui persamaan sebagai berikut:

Peranan relatif $Y_i = \frac{\lambda_i}{\sum_{m=1}^{S} \lambda_m}$ x 100% (Gaspersz, 1995)

s = min (h-1, p)

8.11. Penilaian Validitas Diskriminan

Skor diskriminan yang diperoleh perlu diuji validitasnya melalui persamaan sebagai berikut:

Persentase ketepatan klasifikasi = $\frac{n_{1C} + n_{2C}}{n} \times 100\%$ dimana $n = n_{1C} + n_{2C} + n_{1M} + n_{2M}$

Dengan :

n_{1C} = banyak pengamatan π_1 yang dikelompokkan secara benar sebagai π_1

n_{1M} = banyak pengamatan π_1 yang salah dikelompokkan sebagai π_2

n_{2C} = banyak pengamatan π_2 yang dikelompokkan secara benar sebagai π_2

n_{2M} = banyak pengamatan π_2 yang salah

dikelompokkan sebagai π

Asumsi yang digunakan adalah apabila niliai ketepatan klasifikasi berada di atas 50%, maka tingkat validitas diskriminan dianggap tinggi (Santoso, 2002).

8.12. Langkah-langkah Analisis Diskriminan

Beberapa tahapan utama dalam melakukan analisis diskriminan, antara lain:

1. Merumuskan masalah
   Peneliti harus menentukan terlebih dahulu alasan penggunaan analisis diskriminan dan menentukan variabel yang akan diuji melalui analisis diskriminan.
2. Melakukan estimasi koefisien fungsi diskriminan
   Secara umum terdapat dua cara yang dilakukan untuk menentukan fungsi diskriminan, yakni:
   a. Metode langsung
      Metode ini dilakukan dengan cara menggunakan variabel prediktor secara bersamaan, dimana variabel prediktor dimasukkan tanpa memperhatikan kekuatan diskriminan dari setiap variabel. Metode hanya dapat dilakukan apabila setiap variabel yang digunakan merupakan variabel yang memiliki dasar teori yang kuat.
   b. Metode bertahap (*Stepwise method*)
      Dilakukan secara bertahap dengan mempertimbangkan kekuatan setiap variabel prediktor dalam pengelompokkan data.

3. Melakukan interpretasi hasil
   Langkah-langkah dalam analisis diskriminan dengan menggunakan software SSPS adalah sebagai berikit:
   a. Memasukkan variabel di kotak *variable view*

b. Melakukan input data setiap variabel di kota *data view*
c. Setelah itu, klik *analyze,* lalu *classify* dan klik *discriminant*, kemudian masukkan variabel bebas dan variabel groupingnya.
d. Di kota *variabel grouping* terdapat kotak *define range*, di situ dapat ditentukan nilai minimum dan maximum dalam kelompok
e. Pilih *statistics*. Peneliti tinggal menentukan hal apa saja yang akan diuji dengan mencentang *means*, *box's M* dan *univariate ANOVA*. Metode yang digunakan unutk menentukan koefisien diskriminan dapat dipilih ingin menggunakan *fisher's* dan juga *unstardardized*. Matriks data yang akan dibentuk juga dapat dipilih melalui opsi kotak *matrices.* Setelah itu klik *continue.*
f. Pilik *method*, pilih *wilks' lambda*.
g. Lalu yang terakhir adalah klik *OK*
h. Setelah itu akan keluar hasil *output* analisis dalam kotak *discriminant, group statistics, pooled within-groups matrices(a), eigenvalue, test of equality of groups means, log determinants, wilks' lambda, standadized canonical discriminant function coefficient, structure matrix, functions at group centroid.*
i. Interpretasi analisis diskriminan dimulai dari perbedaan rata-rata variabel setiap grup dan rata-rata total, dimana apabila nilai rata-rata antar grup berbeda, maka setiap variabel memiliki fungsi yang sama dalam mengelompokkan data. Hal tersebut juga akan berlaku terbalik apabila nilai rata-rata variabel antar grup sama.
j. Nilai standar deviasi kelompok juga harus lebih rendah dari nilai satndar deviasi total, sehingga hal tersebut merupakan data yang diperoleh memenuhi asumsi homogenitas. Sangat baik kalau standar deviasi dalam grup lebih rendah daripada standar deviasi total.

k. Nilai *standardized coefficient* dan dan struktur matrix, menunjukkan kekuatan setiap variabel dalam membentuk fungsi diskriminasi dari yang paling tinggi hingga yang paling rendah.
l. *Pooled within-group correlation matrix* mengindikasikan korelasi antar variabel prediktor.
m. *Test of equality of group means* menunjukkan nilai signifikansi F untuk dapat dibandingkan dengan nilai alpha yang digunakan oleh peneliti.
n. Nilai lain dapat diperhatikan untuk peneliti, sehingga dapat memperkaya pembahasan hasil uji yang dilakukannya dalam penelitian.

4. Melakukan uji signifikansi
   Uji signifikansi diskriminan dilakukan dengan menggunakan nilai wilks‘ lambda, yang kemudian diubah ke data chi-square.
5. Melakukan validasi fungsi diskriminan
   Validasi fungsi diskriminan dilakukan dengan melihat nilai korelasi kanonik, dimana apabila nilai korelasi makin mendekati 1, maka setiap fungsi diskriminan yang dibenuk pun akan semakin kuat.

DAFTAR PUSTAKA

Alifta Ainurrochmah; Memi Nor Hayati; Satriya, A.M.A. (2019) 'Perbandingan Klasifikasi Analisis Diskriminan Fisher Dan Metode Naive Bayes', *Jurnal Aplikasi Statistika & Komputasi Statistik*, 11(2), pp. 37–48.

Andri, J. and Goejantoro, R. (2014) 'Analisis Diskriminan Pada Konsumen Air Mineral Dalam Kemasan (AMDK) Discriminant Analysis In Consumer Bottled Of Mineral Water (Mineral Water)', 5, pp. 91–98.

Gaspersz, V. (1995) *Teknik Analisis Dalam Penelitian Percobaan*. 1st edn. Bandung: Tarsito.

Gozali, I. (2018) *Aplikasi Analisis Multivariate (Edisi 9)*, *Semarang: BP Universitas Diponegoro*.

Hair, J. *et al.* (2010) *Multivariate Data Analysis.pdf*. 7th edn, *Australia : Cengage*. 7th edn. Cengage.

Johnson, R.A. and Wichern, D.W. (2007) *Applied Multivariate Statistical Analysis.: Pearson Prentice Hall*. 6th edn, *Pearson Prentice Hall*. 6th edn. New Jersey: Perason Prentice Hall.

Mattjik, A.A. and Sumertajaya, I.M. (2011) *Sidik Peubah Ganda dengan Menggunakan SAS*, *Sidik Peubah Ganda Dengan menggunakan SAS*.

Rahmadeni Rahmadeni, M.S. susandi susandi susandi (2019) 'Analisis Diskriminan Fisher Untuk Klasifikasi Risiko Kredit', in *Seminar Nasional Teknologi Informasi Komunikasi dan Industri*, pp. 478–481. Available at: http://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/downloadSuppFile/7962/944.

Sam, M. (2019) 'yang diperoleh dari Dinas Koperasi , UMKM , Perindustrian dan Perdagangan . Variabel yang Data yang digunakan dalam penelitian ini adalah data

sekunder berupa 250 profil UMKM digunakan dalam penelitian ini mencakup variabel bebas berupa Modal (X 1), Oms', in *Prosiding Seminar Nasional Penelitian & Pengabdian Kepada Masyarakat 2019*. Palopo, pp. 119–124.

Santoso, S. (2002) *Buku Latihan SPSS Staistik Multivariat*. Jakarta: PT Elex Media Komputindo.

Suryanto (1988) *Metode Statistik Multivariat*. Jakarta: P2LPTK.

BAB 9

ANALISIS MANOVA

Oleh Gregorio Antonny Bani

9.1. Pendahuluan

Manova (*Multi Analysis of Varian*) merupakan bentuk analisis yang lebih luas dari Anova (*Analysis of Varians*), dimana analisis ini digunakan untuk mengukur pengaruh dari variabel bebas yang dapat berjumlah lebih dari satu terhadap variabel bebas yang berjumlah lebih dari satu (Johnson and Wichern, 2007). Manova merupakan analisis berdimensi tinggi karena digunakan banyak variabel yang berhubungan satu dengan yang lainnya, serta bentuk hubungan dari variabel-variabel dapat diuji secara simultan atau sekaligus, dan juga secara parsial atau satu per satu (Riswan; Khairudin, 2019). Perhitungan Manova difokuskan pengujian nilai signifikansi yang menghitung perbedaan rata-rata secara simultan antara kelompok yang terbagi dalam dua atau lebih variabel terikat (Tabachnick, 2007).

Hal yang perlu dipahami oleh peneliti adalah bahwa dalam setiap upaya mengungkapkan suatu permasalahan secara statistika, terkadang pengamatan perlu dilakukan dari berbagai sisi. Berdasarkan hal tersebut, maka secara bersamaan terbentuk peubah acak secara serentak yang bersifat multivariate. Misalkan ε merupakan suatu bentuk eksperimental dalam ruang sampel S, dimana $X_1 = X_{i(S)}$ untuk i = 1, 2, …, n dan untuk setiap $s \in S$. Maka $X_1, X_2, \ldots, X_n$ merupakan bentuk peubah acak multivariat atau vector acak yang berdimensi tinggi (Tirta, 2004). Model matematis dari MANOVA mengikuti persamaan regresi linier berganda, sebagaimana ditunjukkan pada persamaan berikut ini:

$$X_{ij} = \mu + Y_i + \beta_i + \varepsilon_{ij}$$

Dimana:

X_{ij} = pengamatan pada baris ke-*I* dan kolom ke-*j*

μ = nilai tengah rata-rata

Y_i = pengaruh baris ke-*i*

β_i = pengaruh kolom ke-*j*

ε_{ij} = galat (Santoso, 2019)

Berdasarkan pengamatan di atas, maka diketahui bahwa uji MANOVA didasarkan pada pertimbangan keragaman antar contoh dalam baris dan keragaman antar pengamatan dalam kolom. Oleh karena itu, persamaan uji MANOVA disajikan dalam bentuk Tabel 11.1 berikut ini:

Tabel 9.1 Uji MANOVA (Multi analysis of varians)

Sumber	**db**	**JK**	**KT**	**F**
Baris	(n-1)	$\sum_{i=1}^{n}(\sum_{j=1}^{k} X_{ij})^2/k - FK$	$JK_{Baris}/(n-1)$	KT_{Baris}/KT_G
Kolom	(k-1)	$\sum_{j=1}^{k}(\sum_{i=1}^{n} X_{ij})^2/n - FK$	$JK_{Kolom}/(k-1)$	KT_{Kolom}/KT_G
Galat	(n-1) (k-1)	$JK_G = JK_{Total} - JK_{Baris} - JK_{Kolom}$	$JK_G/(n-1)(k-1)$	
Total	nk-1	$\sum_{i=1}^{n}(\sum_{j=1}^{k} X_{ij})^2 - FK$		

(Sumber: Sulistiyowati, 2017)

FK = Faktor koreksi dari persamaan

$$FK = \sum_{i=1}^{n}(\sum_{j=1}^{k} X_{ij})^2/nk$$

Asumsi dasar yang digunakan untuk menguji hipotesis MANOVA adalah H_0 akan ditolak apabila rasio generalisasi hasil analisis ragam lebih kecil dari taraf kepercayaannya (Johnson and Wichern, 2007). Pada awal mulanya sebagian besar aplikasi teknik analisis multivariat telah dilakukan dalam bidang ilmu psikologi dan eksata. Namun, seiring dengan perkembangan minat metode analisis multivariat yang kini telah menyebar ke berbagai bidang lain seperti bidang social dan humaniora, karena model analisis dari statistik multivariat dianggap mampu memecahkan masalah yang rumit dengan pemodelan yang lebih mudah diinterpretasikan. Dalam penggunaanya, MANOVA sering dianggap memiliki beberapa hal yang lebih baik dari ANOVA, dimana MANOVA memiliki kelebihan, seperti dapat melakukan analisis pada semua variabel terikat sekaligus, sehingga kesalahan tipe I (α) dalam pengambilan keputusan uji statistik dapat dihindari dan mengungkapkan perbedaan yang tidak ditemukan dengan analisis yang dilakukan secara terpisah dari setiap variabel (Stevens and Pituch, 2016).

Berdasarkan hal tersebut, maka setiap peneliti memiliki kesempatan yang lebih dalam mengungangkapkan tingkat signifikansi dari variabel yang diuji karena terdapat perbedaan dari perlakuan dan interaksi dari setiap perlakuan, sehingga dapat memberi informasi yang lebih baik dari hasil penilitian yang dilakukan. MANOVA memiliki dasar asumsi yang sama ANOVA, tetapi jangkauannya lebih banyak untuk kasus yang bersifat mutivariat. Oleh karena itu, beberapa dasar asumsi dari MANOVA, antara lain:

1. Independensi
   Pengamatan yang dilakukan terhadapa variabel harus independen agar dapat diketahui secara siginifikan pengaruh dari setiap perlakuan dan interaksi dari setiap perlakuan.
2. Teknik sampling acak.
   Hal ini harus dilakukan agar memenuhi kaidah probabilitas dan interval dari setiap pengamatan dapat ditemukan dengan baik.
3. Normalitas multivariate.

Data yang diperoleh harus berdistribusi normal dalam kaidah multivariate secara kelompok.

4. Homogenitas matriks kovariansi.
   Variansi dari setiap variabel terikat harus sama pada setiap kelompok dengan tingkat korelasi setiap varibael harus signifikan, agar dapat diketahui tingkat perbedaan yang jelas dari setiap matriks kovarian (Sutrisno and Wulandari, 2018).

Berdasarkan hal tersebut, maka perlu dilakukan uji pendahuluan terlebih dahulu untuk memenuji kaidah asumsi sebagaiamana yang telah disebutkan di atas. Hal tersebut bertujuan agar uji MAOVA yang dapat diinterpretasikan secara baik dan benar, serta memenuhi kaidah akademis yang dibutuhkan dari setiap penelitian. Bab ini akan menguraikan tentang uji MANOVA agar dapat mudah dipahami dan dapat digunakan sebagai informasi bagi peneliti yang hendak melakukan penelitian.

9.2. Nomalitas Multivariat

Dalam uji MANOVA, ditemukan lebih dari satu variabel terikat. Oleh karena itu, uji asumsi Normalitas harus dilakukan secara simultan. Terdapat berbagai macam cara yang dapat dilakukan dalam melakukan analisis normalitas, namun cara yang paling sederhana adalah membandingan model penyebaran data dalam kurva normal. Pendekatan lain yang dapat dilakukan adalah dengan menggunakan pendekatan *Chi-square* ($\mathcal{X}^2$) melalui persamaan sebagai berikut:

$$\mathcal{X}^2 = \frac{(f_i - f_h)^2}{f_h}$$

Dimana:

$\mathcal{X}^2$ = chi square hitung
f_1 = frekuensi yang diharapkan
f_h = frekuensi observasi

Asumsi yang digunakan apabila nilai $\mathcal{X}^2$ hitung lebih kecil dari nilai $\mathcal{X}^2$ tabel, maka data berdistribusi normal. Hal yang berkebalikan akan berlaku apabila nilai $\mathcal{X}^2$ hitung lebih besar dari nilai $\mathcal{X}^2$ tabel (Prasojo, 2019).

9.3. Homoskedastisitas Multivariat

Pengujian Homoskedastisitas dalam uji MANOVA merupakan uji homogenitas yang terdiri dari 2 bentuk, yakni yakni uji homogenitas varian dan uji homogenitas matriks kovarian, dimana secara matematis akan dihitung dengan menggunakan uji Levene sebagaimana ditunjukkan dalam persamaan sebagai berikut:

$$W = \frac{(n-k)\sum_{i=1}^{k} N_i\,(\bar{Z}\imath\,.\,\bar{Z}n)^2}{(k-1)\sum_{i=1}^{k}(Zij\,.\,Zi)^2}$$

Dimana:
Z_i = nilai tengah kelompok ke-i
Z = nilai tengah keseluruhan data
(Iqbal *et al.*, 2020)

Asumsi yang digunakan dalam uji ini adalah apabila nilai W lebih besar dari nilai F (α, k-1, N-k) maka data yang diperoleh dianggap homogen dan begitu juga berlaku sebaliknya.

9.4. Statistik MANOVA

Dalam melakukan uji statistic MANOVA maka terdapat beberapa pilihan yang dapat dilakukan, antara lain: *Wilks' Lambda, Pillai, Lawley-Hotelling*, dan *Roy's Largest Root*. Seiring dengan perkembangan dunia teknologi kommputer, maka kini telah banyak software statistic yang menyajikan model analisis statistik MANOVA tersebut di atas dengan cara yang lebih

mudah dan pada umumnya semua analisis yang tersedia selalu menunjukkan kesimpulan yang sama, dimana keempat metode yang tersedia melakukan pengujian terhadap nilai nilai eigen dan matriks kovariansi dari data yang diperoleh (Rencher, 2002).

1. Uji *Wilks' Lambda*

   $A = \frac{[E]}{[H+E]}$ (Hidayati, 2019)

   Asumsi yang digunakan dalam uji *Wilks' Lambda* adalah apabila nilai A lebih kecil sama dengan $A_{\alpha,\ p,\ vH,\ VE}$, maka terdapat perbedaan yang signifikan dari hasil analisis ragam setap variabel

   Dimana:
   p = jumlah variabel
   V_H = derajat bebas hipotesis
   V_E = derajat bebas error

   Karakteristik dari uji *Wilks' Lambda*, antara lain:
   a. Nilai $V_E \geq p$.
   b. Derajat bebas V_E dan V_H bernilai sama dengan uji ANOVA, yaitu $V_H=k-1$ dan $V_B=k.n-1$
   c. Posisi parameter p dan V_H dapat ditukar kedudukannya satu sama lain
   d. Nilai uji Lambda pada nilai eigen λ_s dan $E^{-1}H$, adalah $A = \prod_{i=1}^{s} \frac{1}{1+\lambda_i}$
   e. Nilai A berkisar antara $0 \leq A \leq 1$
   f. Nilai table kritis akan naik ketika nilai p juga naik
   g. Hipotesis ditolak apabila nilai A melebihi nilai α dan distribusi F (Sauddin, Islam and Alauddin, 2018)

2. Uji *Pillai*

   $V^s = tr[H(H+E)^{-1}]$ (Rencher, 2002)

   Asumsi yang digunakan adalah apabila nilai V^s lebih kecil sama dengan nilai α, maka tidak ada signifikansi dalam ragam dan begitu juga berlaku sebaliknya (Sauddin, Islam and Alauddin, 2018).

3. Uji *Lawley-Hotelling*
   $U^s = tr(E^{-1}H)$ (Rencher, 2002)

   Asumsi yang digunakan dalam pengujian *Lawley-Hotelling* adalah apabila nilai U^s dihitung lebih kecil dari Tabel, maka dianggap tidak ada signifikansi dalam ragam dan begitu juga akan berlaku sebaliknya (Sauddin, Islam and Alauddin, 2018).

4. Uji *Roy's Largest Root*
   $N = \frac{1}{2}(V_E - p - 1)$ (Hidayati, 2019)

   Bentuk lain dari uji *Roy's Largest Root* adalah:

   $$\theta = \frac{\lambda}{1+\lambda}$$

   Nilai kritis dapat ditemukan dalam Tabel θ, dimana apabila nilai θ_{hitung} lebih kecil dari nilai θ_{table}, maka tidak ada signifikansi dalam ragam dan begitu juga sebaliknya (Sauddin, Islam and Alauddin, 2018)

Perhitungan yang perlu dilakukan adalah melakukan perhitungan nilai tengah rata-rata dan analisis hasil ragam, yang disajikan pada Tabel 11.2 berikut ini:

Tabel 9.2 Rangkuman Matriks Analisis Ragam MANOVA

Sumber	**Matriks Ragam**	**db**
Faktor A	$H_A = (A\widehat{M})'[A(W'W)^{-1}A']^1 A\widehat{M}$	V_{HA}=*a-1*
Faktor B	$H_B = (B\widehat{M})'[B(W'W)^{-1}B']^1 B\widehat{M}$	V_{HB}=*b-1*
Interaksi	$H_{AB} = (C\widehat{M})'[C(W'W)^{-1}C']^1 C\widehat{M}$	V_{HAB}=*(a-1)(b-1)*
Galat	$E = (X - W\widehat{M})'(X - W\widehat{M})$	
Total	$T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n_{ij}}(x_{ijk} - \bar{x})(x_{ijk} - \bar{x})'$	*N-1*

(Sumber: Sutrisno and Wulandari, 2018)

Dimana:

a = jumlah baris
b = jumlah kolom
N = jumlah data
X = matriks *N.p*
$\widehat{M}$ = matriks *(ab).p*

Dalam uji MANOVA juga dapat dilakukan 2 arah, dengan model matematis pada persamaan sebagai berikut:

$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} = \mu_{ij} + \varepsilon_{ijk}$ (Santoso, 2019)

Dimana:

α_i = efek level ke-*i* dari setiap variabel dalam Y_{ijk}
β_j = efek level ke-*i* dari setiap variabel dalam Y_{ijk}
γ_{ij} = interaksi dari setiap variabel

Berdasarkan persamaan di atas, maka diperoleh jumlah kuadrat dan matriks hasil merupakan kali kelompok dari:

$T = H_A + H_B + H_{AB} + E$

Sehingga diperoleh matriks:

$E = T - H_A - H_B - H_{AB}$ (Sauddin, Islam and Alauddin, 2018)

DAFTAR PUSTAKA

Hidayati, N. (2019) '(Manova) Pada Rancangan Acak Kelompok Lengkap Dasar (Rakld)', *Jurnal Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu* [Preprint]. Available at: http://sigitnugroho.id/e-Skripsi/0801 Kajian Prosedur Multivariat Analysis Of Variance (MANOVA) pada Rancangan Acak Kelompok Lengkap Dasar (RAKLD) .pdf.

Iqbal, M. *et al.* (2020) 'Analisis MANOVA Satu Arah untuk Melihat Perbedaan Status Gizi Balita Berdasarkan Wilayah Pembangunan Utama di Indonesia Tahun 2017', *Journal of Data Analysis*, 3(1), pp. 50–61.

Johnson, R.A. and Wichern, D.W. (2007) *Applied Multivariate Statistical Analysis.: Pearson Prentice Hall*. 6th edn, *Pearson Prentice Hall*. 6th edn. New Jersey: Perason Prentice Hall.

Prasojo, B.H.N.A. (2019) *Statistik Bisnis*. 1st edn. Edited by Sumartik. Sidoarjo: UMSIDA Press.

Rencher, A.C. (2002) *Methods of Multivariate Analysis*. 2nd edn. New York: John Wiley & Sons, Inc.

Riswan; Khairudin (2019) *Statistik Multivariate*. Lampung: CV. Anugrah Utama Raharja.

Santoso, I.H. (2019) *Statistik II*. 1st edn. Edited by R.S. Bahtiar. Surabaya: UWKS Press.

Sauddin, A., Islam, U. and Alauddin, N. (2018) *Step by Step MANOVA Using R*. Available at: https://doi.org/10.13140/RG.2.2.30862.20807.

Stevens, K.A. and Pituch, J.P. (2016) *Applied multivariate statistics for the social sciences - Analyses with SAS and*

IBM's SPSS Sixth. 6th edn. New York: Routledge. Available at: https://www.cambridge.org/core/product/identifier/CBO9781107415324A009/type/book_part.

Sulistiyowati, W.C.C.A. (2017) *Statistika Dasar (Konsep dan Aplikasinya)*. 1st edn. Edited by S.B. Sartika. Sidoarjo: UMSIDA Press.

Sutrisno, S. and Wulandari, D. (2018) 'Multivariate Analysis of Variance (MANOVA) untuk Memperkaya Hasil Penelitian Pendidikan', *AKSIOMA : Jurnal Matematika dan Pendidikan Matematika*, 9(1), p. 37. Available at: https://doi.org/10.26877/aks.v9i1.2472.

Tabachnick, B.G.L.S.F. (2007) *Using Multivariate Statistics.* 6th edn. Edited by C. Campanella. New Jersey: Pearson Education, Inc. Available at: https://doi.org/10.1037/022267.

Tirta, I.M. (2004) *Pengantar Statistika Matematika*. 2nd edn. Jember: Penerbit FMIPA Universitas Jember.

BIODATA PENULIS

Dr. Joni Wilson Sitopu, M.Pd.
Dosen Universitas Simalungun

Joni Wilson Sitopu, lahir di Medan, Pendidikan S1 Prodi Matematika FMIPA USU Medan, S2 Prodi Pendidikan Matematika UNIMED Medan dan S3 Prodi Ilmu Matematika FMIPA USU Medan. Dosen di FKIP, SPS Prodi S2 PW, IPS dan Manajemen, serta Kepala Lembaga Akreditasi Internal (LAI) Universitas Simalungun. Penulis aktif mengikuti sebagai pemakalah pada Seminar Nasional dan internasional, aktif mengikuti Seminar Nasional dan internasional dan aktif menulis Buku, Jurnal OJS, Sinta, dan Scopus.

BIODATA PENULIS

Dr. Ir. Yongker Baali, M.Si
Dosen Program Studi Manajemen
Fakultas IPTEK dan Keguruan Universitas Trinita

Penulis dilahirkan di Madolay BANGKEP-SULTENG pada tanggal 12 Agustus 1965. Penulis adalah dosen tetap pada Program Studi Manajemen Fakultas IPTEK dan Keguruan Universitas Trinita. Menyelesaikan Studi S1 Jurusan Teknik Industri Fakultas Teknologi Industri pada Institut Teknologi Minaesa Tomohon, tahun 1992 memperoleh gelar Insinyur (Ir.).

Menyelesaikan Studi S2 Jurusan Pengelolaan Sumberdaya Pembangunan minat Manajemen Perusahaan pada Program Pascasarjana UNSRAT Manado, tahun 2008 memperoleh gelar Magister Sains (M.Si).Menyelesaikan Studi S3 Program Doktor Kajian Lingkungan dan Pembangunan pada Program Pascasarjana Universitas Brawijaya Malang, tahun 2014 memperoleh gelar Doktor (Dr.).

BIODATA PENULIS

Rida Ristiyana, S.E., M.Ak., CIQnR., C.FR., C.Ftax., C.Ed., CTTr.
Dosen Tetap Program Studi Akuntansi
Fakultas Ekonomi dan Bisnis
Universitas Islam Syekh-Yusuf (UNIS) Tangerang

Penulis adalah dosen yang telah tersertifikasi sebagai dosen profesional. Ia adalah dosen tetap pada Program Studi Akuntansi, Fakultas Ekonomi dan Bisnis, Universitas Islam Syekh-Yusuf (UNIS) Tangerang. Ia menyelesaikan Pendidikan S-1 Akuntansi di Universitas Stikubank (UNISBANK) Semarang pada tahun 2013 dan menyelesaikan Pendidikan S-2 Akuntansi di Universitas Mercu Buana (UMB) Jakarta pada tahun 2016. Pada 2 pendidikan tersebut memperoleh predikat *Cumlaude*. Pada 2021 telah menyelesaikan sertifikasi profesi peneliti. Penulis memiliki kepakaran di bidang Akuntansi, Pajak, Keuangan dan untuk mewujudkan karir sebagai dosen profesional, penulis pun aktif sebagai peneliti di bidang kepakarannya dan hasil penelitian telah didanai oleh internal perguruan tinggi serta dipublikasikan pada jurnal-jurnal terakreditasi baik nasional maupun internasional. Selain itu, penulis juga menjadi reviewer pada dewan redaksi di beberapa OJS. Penulis juga aktif menjadi pemakalah di berbagai kegiatan ilmiah dan menjadi narasumber pada workshop/seminar/lokakarya tertentu. Di sisi lain, penulis juga aktif dalam menulis buku sekaligus menjadi editor buku (editing naskah) dengan harapan dapat memberikan kontribusi positif

bagi bangsa dan negara yang nantinya dapat menjadi ilmu jariyah dan ladang pahala demi mencerdaskan anak bangsa. Email Penulis: rristiyana@unis.ac.id. Adapun karya buku yang telah terbit meliputi topik berikut :1). Analisis Laporan Keuangan - Penilaian Kinerja Perusahaan Dengan Pendekatan Rasio; 2). Isu-isu Kontemporer Akuntansi Manajemen - Sebagai Alat Perencanaan, Pengendalian dan Pengambilan Keputusan; 3). Bunga Rampai-Kebijakan Perpajakan di Indonesia di Masa Pandemi COVID-19; 4). Pengaturan Pengelolaan Keuangan Perusahaan-Implementasi Strategi Dalam Keputusan Pendanaan dan Pengendalian Keuangan; 5). Implementasi Pengelolaan Keuangan Daerah - Tata kelola menuju Pemerintahan Yang Baik; 6). Konsep, Teori dan Implikasi Rencana Kerja dan Penganggaran; 7). Tinjauan Fungsi Manajemen Keuangan Perusahaan - Pengelolaan Unsur-unsur Keuangan Perusahaan; 8). Konsep, Prinsip dan Implementasi Keuangan Syariah; 9). Dasar-dasar Akuntansi Syariah; 10). Dasar-dasar Wawasan Kewirausahaan; 11). Akuntansi Keuangan Tingkat Menengah; 12). Konsep dan Sistem Akuntansi Biaya; 13). Akuntansi Suatu Pengantar; 14). Pengantar Ilmu Ekonomi; 15). Mengukur Kinerja Perusahaan Melalui Analisis Rasio Keuangan; 16).Mengenal Bank dan Lembaga Keuangan Lainnya; 17). Pengetahuan Dasar Pasar Modal dan Investasi; 18). Sistem dan Strategi Dalam Konteks Pengendalian; 19). Pengantar Perpajakan; 20). Perpajakan; 21). Akuntansi Sektor Publik; 22). Akuntansi Manajemen; 23). Pengantar Audit; 24). Pengantar MSDM; 25). Kewirausahaan; 26). Manajemen SDM Era Society 5.0; 27). Usaha Mikro, Kecil dan Menengah (UMKM); 28). Sistem Akuntansi; 29). Praktikum Akuntansi Syariah; 30). UMKM Membangun Ekonomi Kreatif; 31). Teori Akuntansi; 32). Manajemen Perpajakan; 33). Ekonomi Moneter; 34). Metodologi Penelitian Kuantitatif dan Kualitatif; 35). Metodologi Penelitian Kuantitatif: Dengan Aplikasi IBM SPSS; 36). Manajemen Koperasi dan UMKM; 37). Strategi dan Meta Analisis Dalam Penelitian Bisnis; 38). Metodologi Penelitian Ekonomi dan Bisnis: Dilengkapi dengan Analisis Regresi SPSS dan SEM-PLS; 39). Metodologi Penelitian Kuantitatif: Dengan Aplikasi IBM SPSS. Sedangkan buku yang sudah dieditori mencakup : a). Akuntansi

Internasional; b). Auditing; c). Analisis Laporan Keuangan; d). Metodologi Penelitian Bisnis; e).Manajemen SDM; f). Manajemen Risiko; g).Perpajakan Internasional; h). Portofolio dan Investasi; i). Kewirausahaan; j). Pengantar Manajemen; k). Dasar-dasar Manajemen Keuangan; l).Dasar-dasar Akuntansi Manajemen; m).Manajemen Perpajakan; n). Dasar-Dasar Perpajakan; o). Akuntansi Keuangan Daerah Berbasis Akrual.

BIODATA PENULIS

Anna Sofia Atichasari, SE.,M.Si.,CMA
asatichasari@unis.ac.id
Dosen Program Studi Akuntansi Fakultas Ekonomi&Bisnis
Universitas Islam Syekh-Yusuf

Penulis lahir di Solo, 15 Oktober 1980. Penulis menekuni bidang akuntansi sejak tahun 1999. Menyelesaikan pendidikan S1 dan S2 ilmu akuntansi di Trisakti Dan memulai karir mengajar sebagai Asisten Dosen serta dipercaya sebagai koordinator Praktikum. Pernah bergabung juga menjadi Auditor pada Kantor Akuntan Publik. Mulai berkarier sebagai dosen tahun 2006 sekaligus kaprodi akuntansi pada perguruan tinggi swasta yaitu Universitas Bung Karno sejak tahun 2010. Pada tahun 2017 diamanahkan kaprodi akuntansi dan berlanjut sebagai wakil dekan bidang akademik pada Universitas Islam Syekh-Yusuf Tangerang di Tahun 2021. Beberapa buku sudah di luncurkan terutama bidang Akuntansi keuangan, akuntansi syariah, akuntansi manajemen dan metodologi penelitian. Publikasi penelitian dan artikel bidang akuntansi juga dilakukan. Aktif tergabung pada Forum Dosen Akuntansi Wilayah Banten dan terlibat menjadi pengurus IAI KAPd tingkat nasional.

BIODATA PENULIS

Dr. Hj. Utin Nina Hermina., SE., MS.i
Dosen Jurusan Administrasi Bisnis
Politeknik Negeri Pontianak

Penulis lahir di Pontianak tanggal 30 September 1971. Penulis adalah dosen tetap pada Program Studi Administrasi Negara Jurusan Administrasi Bisnis Politeknik Negeri Pontianak. Menyelesaikan pendidikan S1 pada Jurusan Ekonomi Manajemen dan menyelesaikan Pendiidkan S2 pada Jurusan Ilmu Ekonomi dan Manajemen Universitas Padjadjaran Bandung. Dan Menyelesaikan Pendidikan S3 pada Jurusan Ekonomi dan Bisnis Universitas Persada Indonesia (UPI YAI) di Jakarta.. Mengajar pada mata kuliah Pemasaran, Azas-azas Manajemen, dan Metodelogi Penelitian, dan juga aktif membuat Penelitian dan Pengabdian Pada Masyarakat.

BIODATA PENULIS

Hartina Husain, S.Si., M.Stat
Dosen Program Studi Tadris Matematika
Fakultas Tarbiyah Institut Agama Islam Negeri Parepare

Penulis lahir pada 23 Mei 1996 di Samata, Sulawesi Selatan. Penulis meraih gelar sarjana dari Universitas Hasanuddin tahun 2017 pada jurusan Statistika. Kemudian penulis melanjutkan studi magister di Surabaya tepatnya di Institut Teknologi Sepuluh Nopember dengan jurusan yang sama. Setelah kelulusannya di tahun 2021, penulis aktif bekerja sebagai Dosen Tadris Matematika di Institut Agama Islam Negeri Parepare. Aktif menulis artikel atau jurnal terkait pemodelan statistika, regresi non parametrik , dan analisis meta. Selain itu, penulis tertarik mempelajari mengenai *data analyst*, *data science* dan *big data*.

BIODATA PENULIS

Amalia Nur Chasanah, S.E, M.M.
Dosen Program Studi Manajemen
Fakultas Ekonomi dan Bisnis Universitas Dian Nuswantoro

Penulis lahir di Semarang 14 September 1983. Penulis adalah dosen tetap pada Program Studi Manajemen Fakultas Ekonomi dan Bisnis, Universitas Dian Nuswantoro Semarang. Menyelesaikan pendidikan S1 pada Program Studi Ilmu Ekonomi dan Studi Pembangunan dan melanjutkan S2 pada Program Studi Magister Manajemen di Universitas Diponegoro.

BIODATA PENULIS

Mario Donald Bani, S.P., M. Biotech. (Adv.)
Dosen Program Studi Bioteknologi
Institut Bio Scientia Internasional Indonesia

Penulis lahir di Kupang, Nusa Tenggara Timur pada tanggal 29 Mei 1986. Penulis adalah dosen tetap pada Program Studi Bioteknologi, Institut Bio Scientia Internasional Indonesia. Penulis menyelesaikan pendidikan S1 pada Jurusan Budidaya Pertanian, Fakultas Pertanian, Universitas Nusa Cendana, Kupang dan melanjutkan S2 pada Jurusan Bioteknologi, School of Chemistry and Molecular Biosciences, the University of Queensland, Australia. Penulis menekuni bidang bioteknologi lingkungan dan pertanian.