

FINDING THE RELATIONSHIPS

Ventje Jeremias Lewi Engel, M.T.
Prodi Informatika
Institut Teknologi Harapan Bangsa

2020

PERKULIAHAN

Akan ada kuliah dengan video yang diupload

Akan ada tugas tertulis maupun tugas praktikum

Menggunakan Jupyter Notebook

Biasakan diri dengan analisis eksplorasi data yang sudah dipelajari di kuliah Probstat I

Plagiasi kode (lebih dari 60%) dan tugas (lebih dari 40%) mengakibatkan mahasiswa otomatis mendapat nilai E.

TUJUAN PEMBELAJARAN

Setelah mengikuti perkuliahan ini, mahasiswa diharapkan mampu:

1. Mahasiswa menjelaskan hubungan antardata
2. Mahasiswa membedakan berbagai teknik analisis statistika multivariat



IMPORTANT!

OUTLINE

Tipe Data

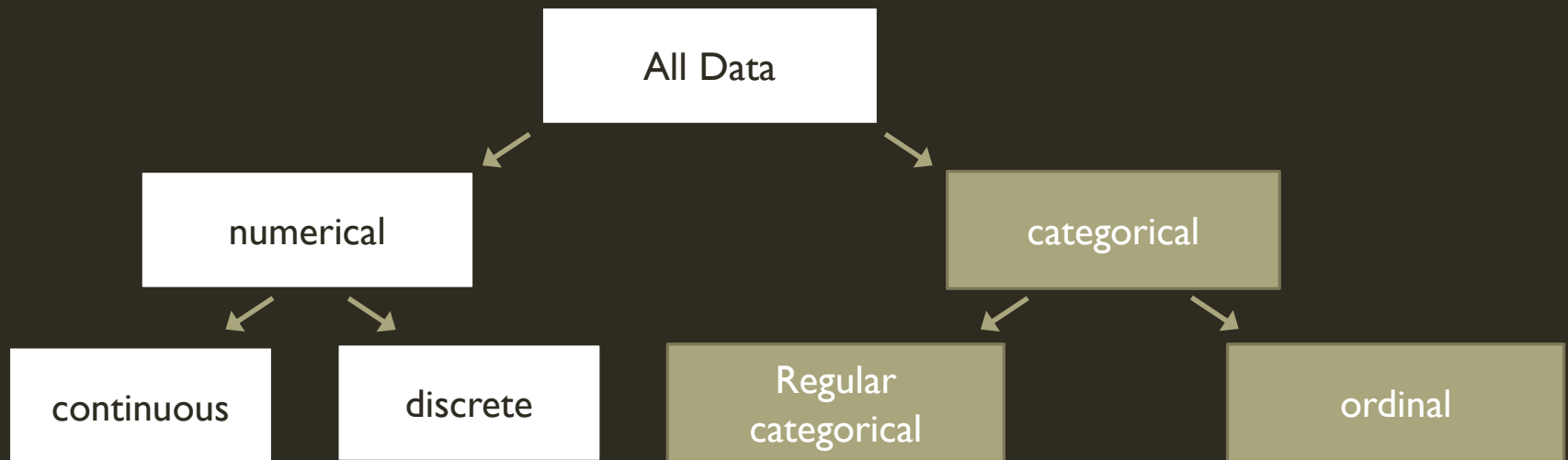
Hubungan Antardata

Dependence vs Interdependence Method

TABEL DATA

country	cr_req	cr_compl y	ud_re q	ud_comply	...	hemispher e	hdi
Argentina	21	100	134	32	...	southern	Very high
Australia	10	40	361	73	...	southern	Very high
Belgium	<10	100	90	67	...	northern	Very high
Brazil	224	67	703	82	...	southern	High
...
United States	92	63	5950	93	...	northern	Very high

TIPE DATA



TABEL DATA

country	cr_req	cr_compl y	ud_re q	ud_comply	...	hemispher e	hdi
Argentina	21	100	134	32	...	southern	Very high
Australia	10	40	361	73	...	southern	Very high
Belgium	<10	100	90	67	...	northern	Very high
Brazil	224	67	703	82	...	southern	High
...
United States	92	63	5950	93	...	northern	Very high

NONMETRIC DATA: NOMINAL & ORDINAL

Data Non Metrik (kualitatif) adalah data yang hanya menunjukkan karakteristik suatu objek. terdiri atas dua kategori yaitu:

1. Nominal

Data yang mengelompokkan objek berdasarkan karakteristik yang bersifat mutually exclusive (tidak memiliki irisan). Contoh: data jenis kelamin membagi objek dalam pria dan wanita.

2. Ordinal

Data yang mengelompokkan objek berdasarkan karakteristik yang mutually exclusive, dan mengurutkannya. Contoh: data hobi diurutkan dari yang paling disukai sampai yang tidak begitu disukai

METRIC DATA: INTERVAL & RATIO

Data Metrik (kuantitatif) adalah data yang menunjukkan karakteristik objek dan jumlah atau derajatnya. Data jenis ini terbagi dua kategori yaitu:

1. Interval

Data yang diurutkan dengan jarak (besaran) yang sama, namun tidak memiliki nilai nol absolut. Contoh: data temperatur yang diukur dalam derajat Celcius.

2. Ratio

Data yang memiliki skala interval dengan nilai 0 mutlak (tidak ada/nihil). Contoh: data berat badan, usia, pendapatan.

OUTLINE

Tipe Data

Hubungan Antardata

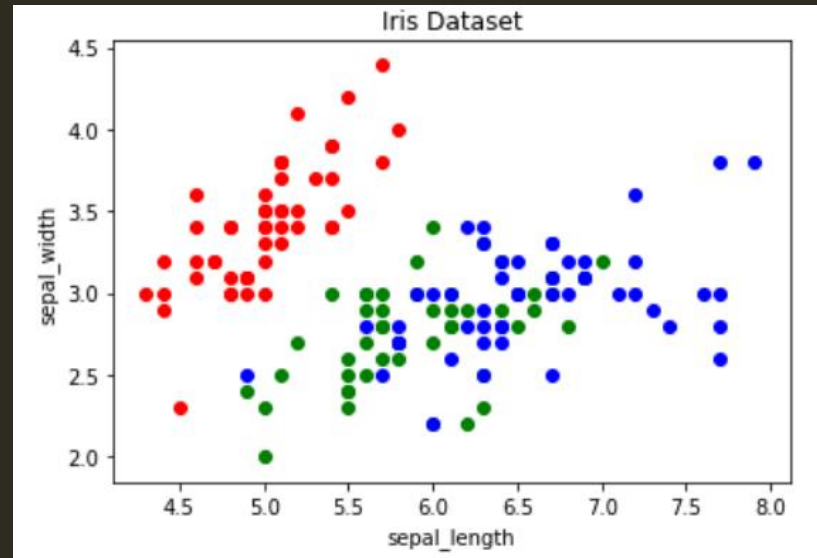
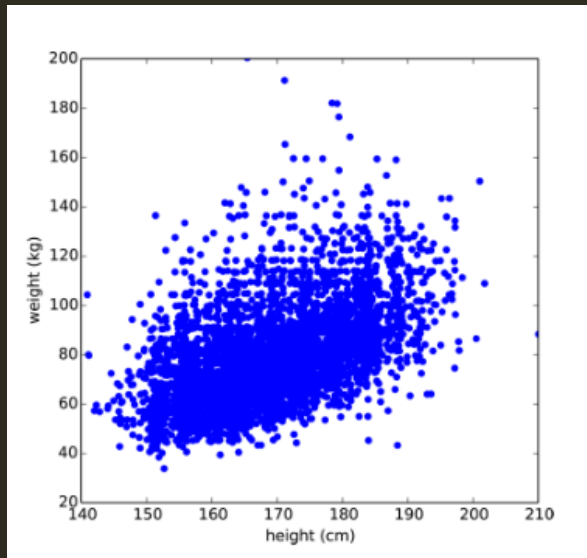
Dependence vs Interdependence Method

MELIHAT HUBUNGAN ANTARDATA

Terdapat 2 metode untuk melihat kira-kira hubungan yang ada antara dua data/varianbel.

1. Grafis

Sebagian besar peneliti senang menggunakan Scatter Plot untuk menggambarkan hubungan antardata.



MELIHAT HUBUNGAN ANTAR DATA

Terdapat 2 metode untuk melihat kira-kira hubungan yang ada antara dua data/variabel.

2. Hitung Korelasi

Kita bisa melihat korelasi 2 data atau variabel dengan menghitung korelasinya. Pertama, cek dulu secara grafis dan tentukan apakah hubungan yang terjadi **linier atau non-linier**.

Bila kemungkinan hubungan linier → gunakan korelasi Pearson

Bila kemungkinan hubungan non-linier → gunakan korelasi Spearman

Lebih lanjut: Baca Ref 2 Chapter 7 bagian Pearson's Correlation dan Spearman Correlation.



HUBUNGAN ANTAR VARIABLE DILIHAT SECARA DEPENDENSI

Terdapat dua jenis penelitian ditinjau dari hubungan antar variable:

1. DEPENDENCE

Satu atau lebih variable memiliki ketergantungan (dependence) pada satu atau lebih variable lainnya. Variable dependen (variable akibat) tersebut dapat di PREDIKSI atau di JELASKAN oleh variable independen (variable penyebab)

2. INTERDEPENDENCE

Tidak ada satu variable atau satu kelompok variable yang didefinisikan tergantung atau tidak tergantung dari variable lainnya. Prosedur ini akan mempelajari seluruh variable secara bersamaan (simultan)

DALAM HUBUNGAN **DEPENDENCE** TERDAPAT INDEPENDENT & DEPENDENT VARIABLES



ISTILAH LAIN

- Faktor penyebab
- Predictor
- Determinan
- Antecedent

ISTILAH LAIN

- Akibat
- Criterion

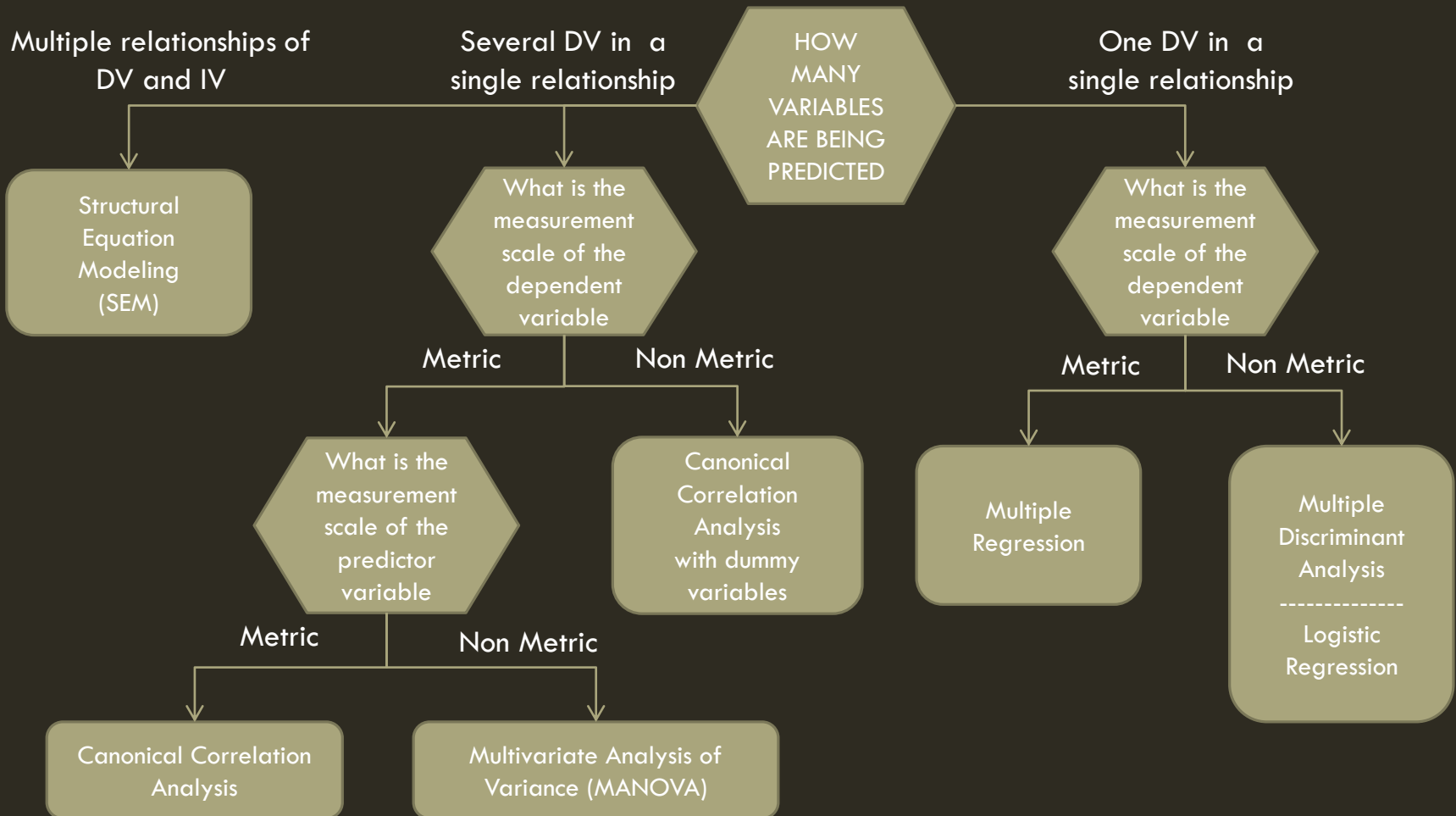
OUTLINE

Tipe Data

Hubungan Antardata

Dependence vs Interdependence Methods

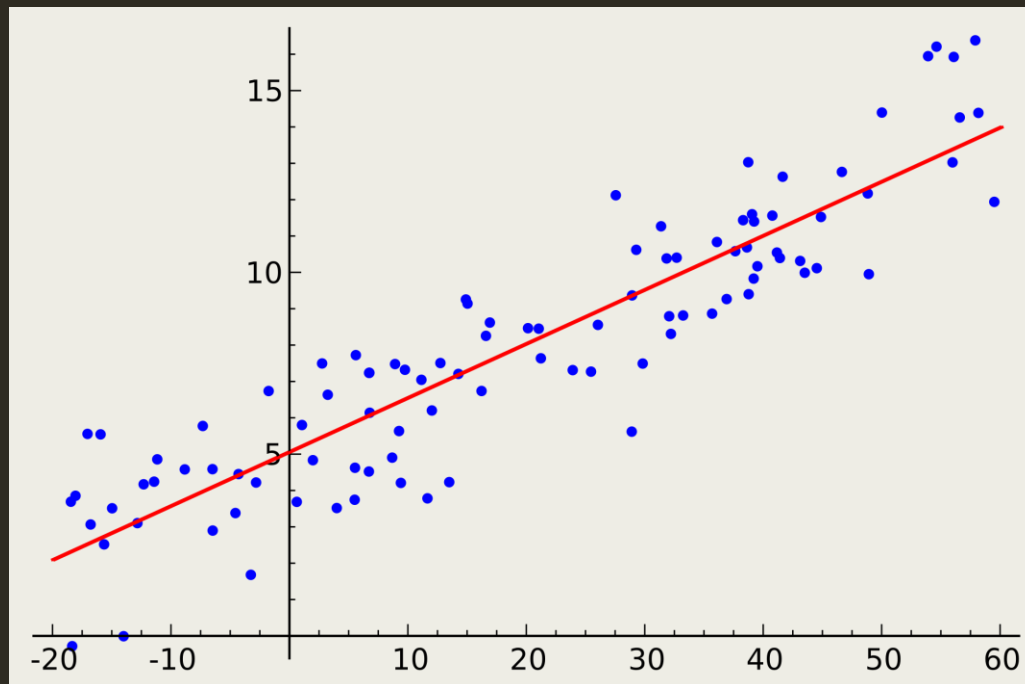
DEPENDENCE METHODS





DEPENDENCE METHOD

REGRESSION ANALYSIS

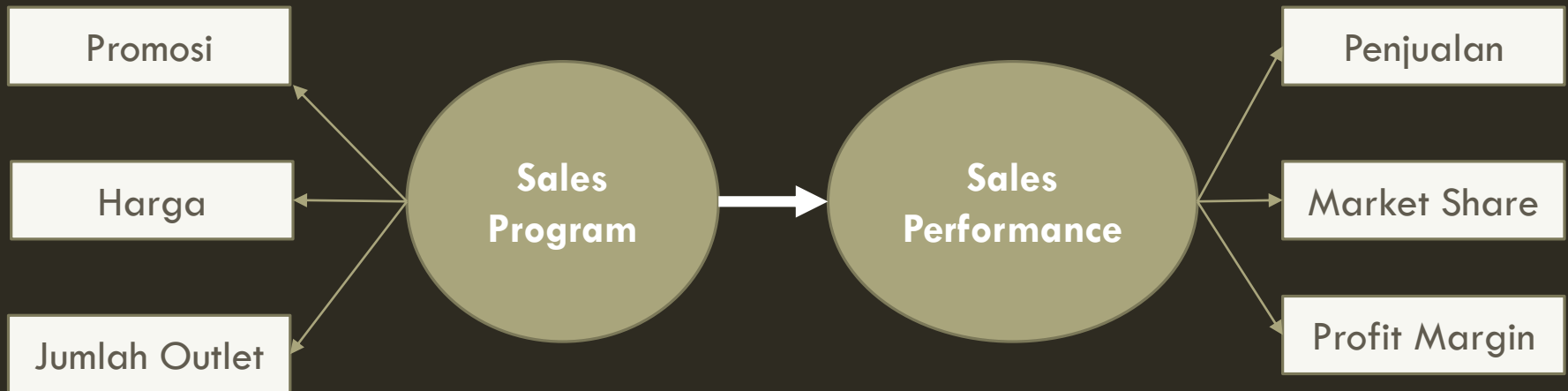


One Dependent Variable in a Single Relationship

Contoh: Apakah tinggi badan mempengaruhi berat badan ?

DEPENDENCE METHOD

CANONICAL CORRELATION ANALYSIS

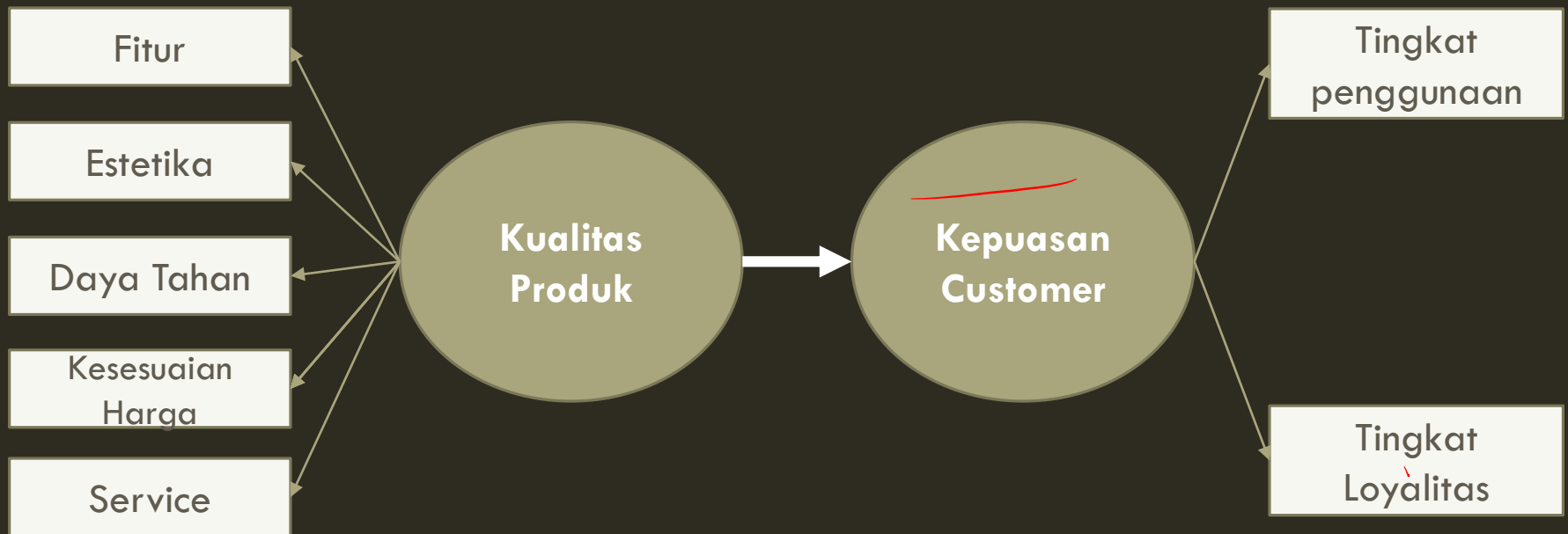


Several Dependent Variables in Single Relationship (Metric)

Contoh: Apakah SALES PROGRAM (Promosi, Harga serta Jumlah Outlet) mempengaruhi SALES PERFORMANCE (Penjualan, Market Share, Profit Margin)

DEPENDENCE METHOD

CANONICAL CORRELATION ANALYSIS W/ DUMMY VARIABLES

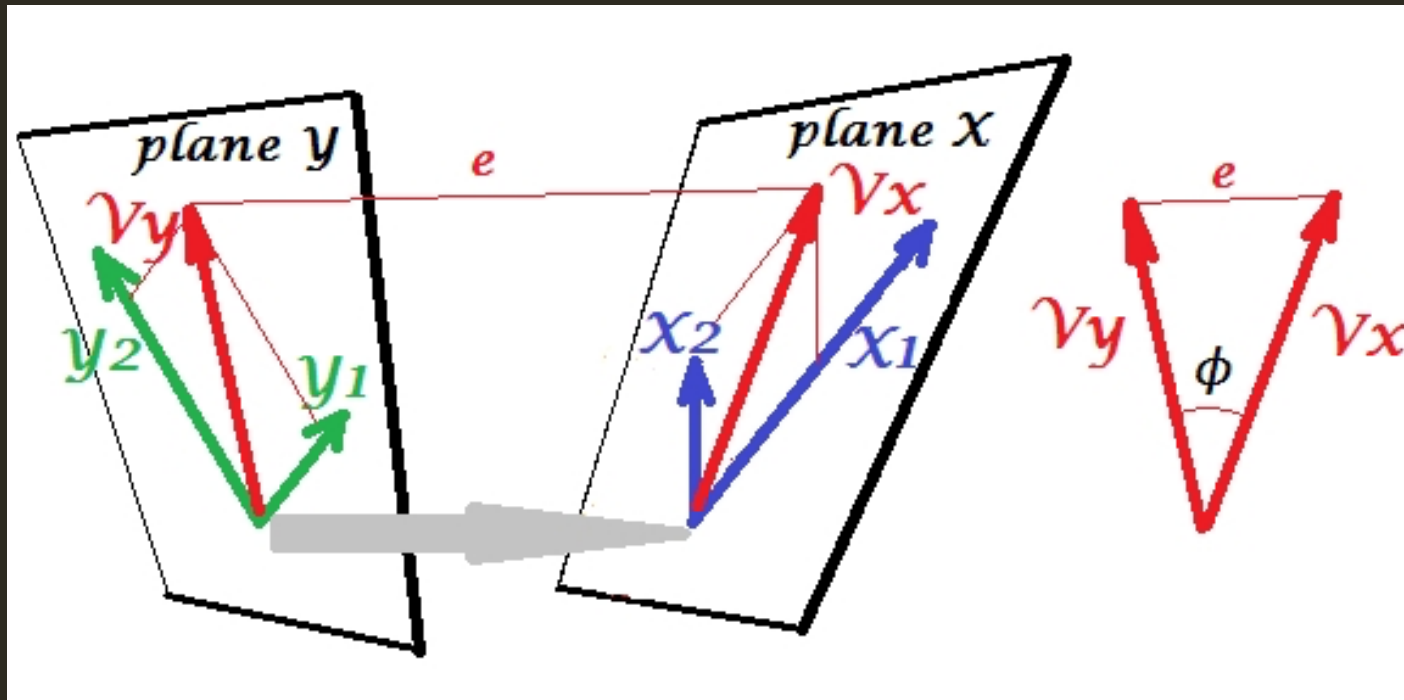


Several Dependent Variables in a Single Relationship (Non Metric)

Contoh: Apakah fitur, estetika, daya tahan, kesesuaian harga produk, after sales service (IV) Mempengaruhi : tingkat penggunaan (DV1), loyalitas customer (DV2)?

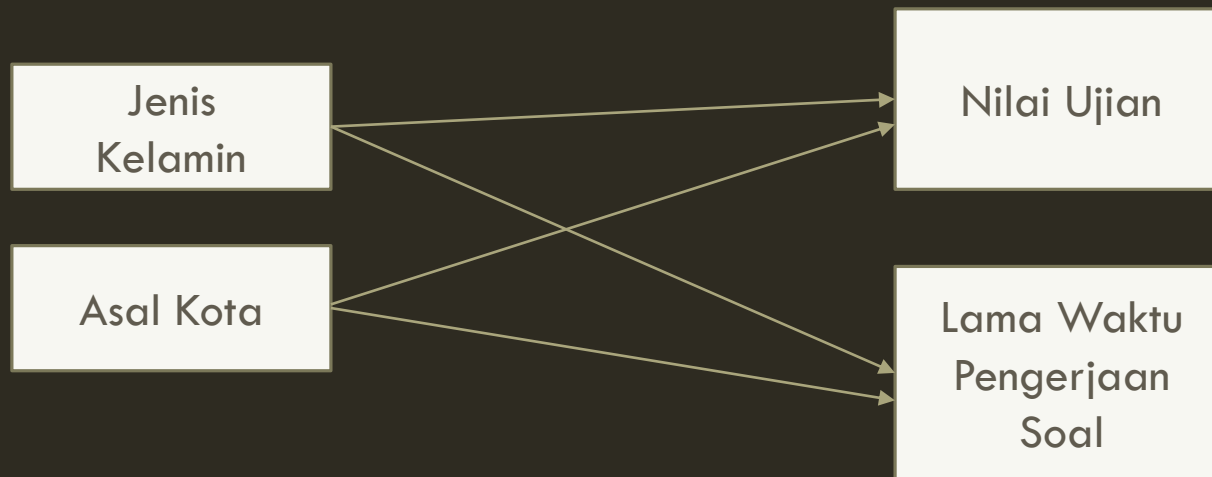
DEPENDENCE METHOD

CANONICAL CORRELATION ANALYSIS



DEPENDENCE METHOD

MULTIVARIATE ANALYSIS OF VARIANCE (MANOVA)

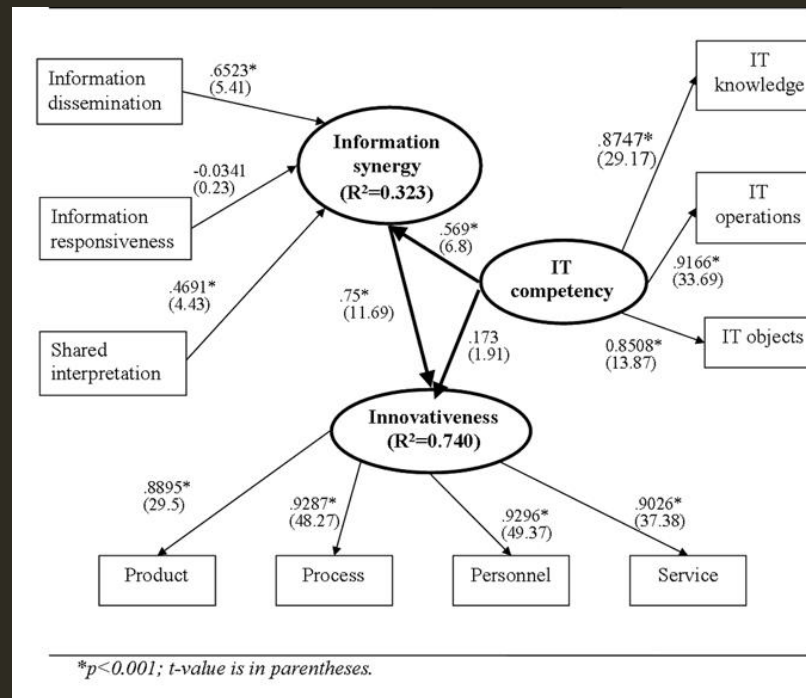


Several Dependent Variables in Single Relationship for non-metric Predictor (Independent Variables)

Contoh: Apakah metode pengajaran dan jenis kelamin mempengaruhi nilai rata-rata dan lama waktu pengerjaan soal ujian?

DEPENDENCE METHOD

STRUCTURAL EQUATION MODELING



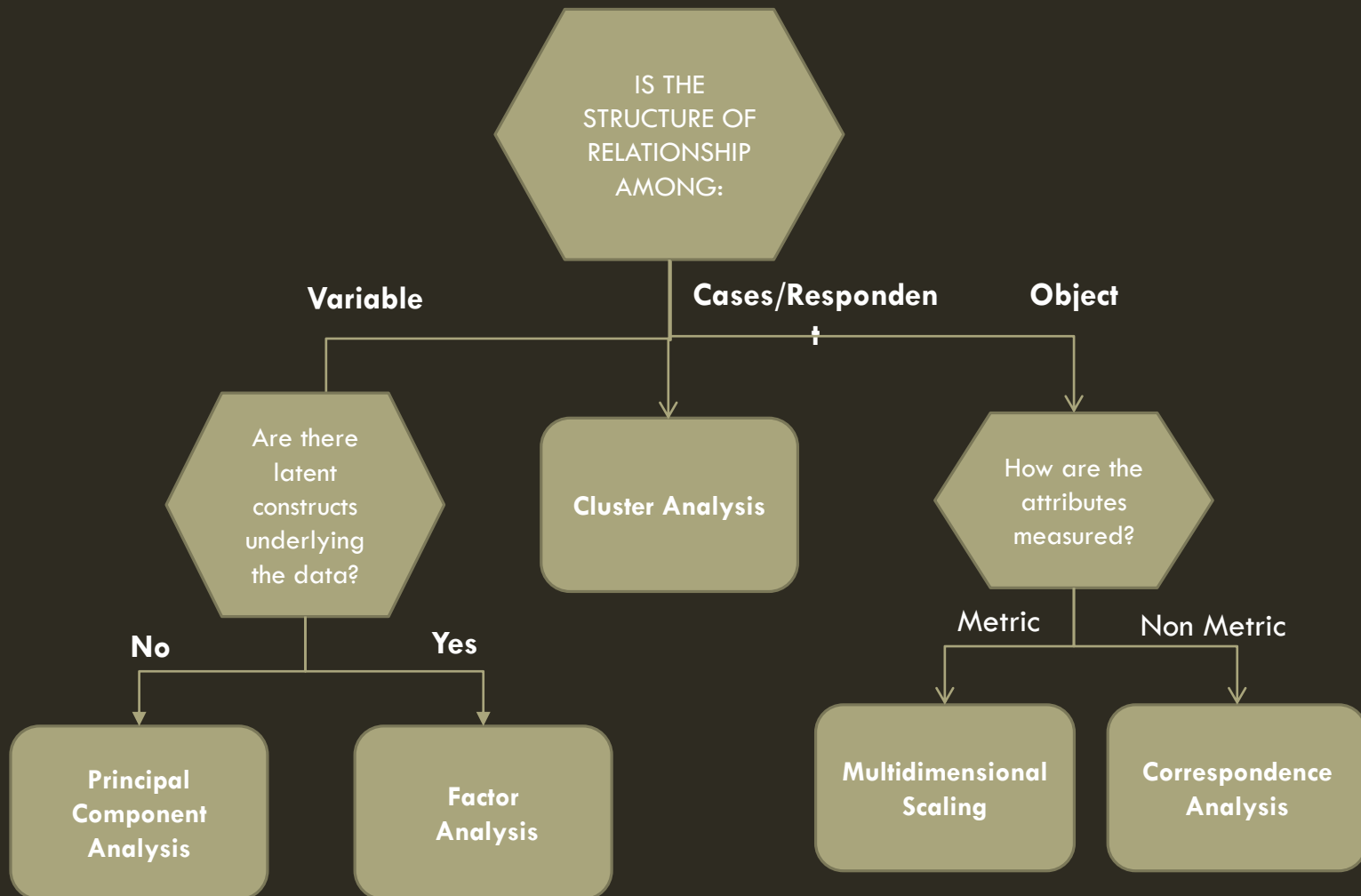
Multiple Relationships of Dependent and Independent Variables

Contoh: (1) apakah Kompetensi IT mempengaruhi Sinergi Informasi?

(2) apakah Kompetensi IT mempengaruhi Inovasi?

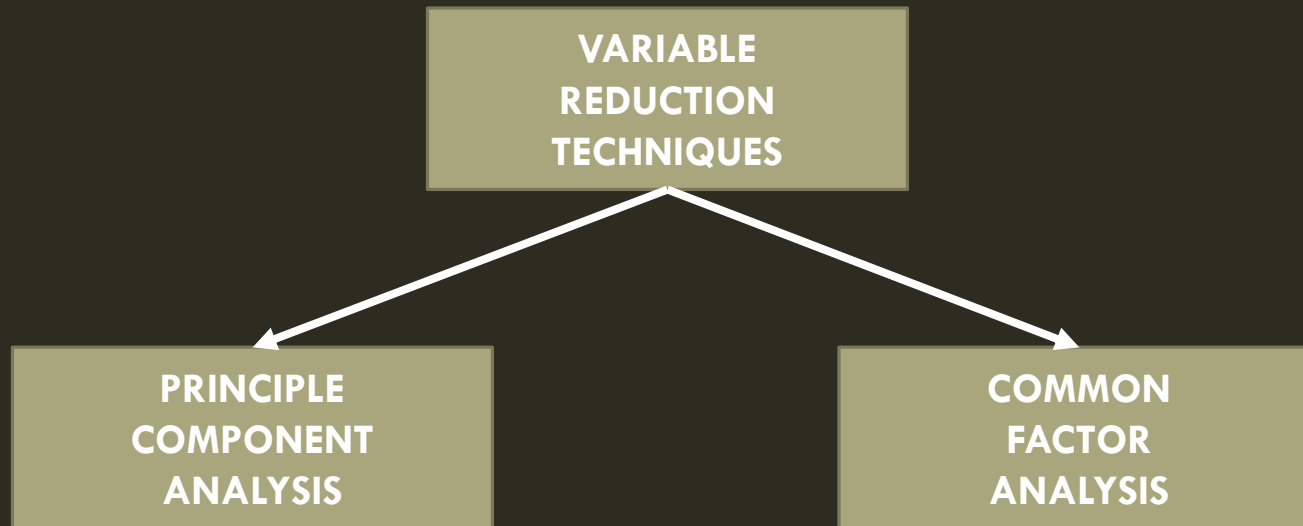
(3) apakah Sinergi Informasi mempengaruhi Inovasi?

INTERDEPENDENCE METHOD



INTERDEPENDENCE METHOD

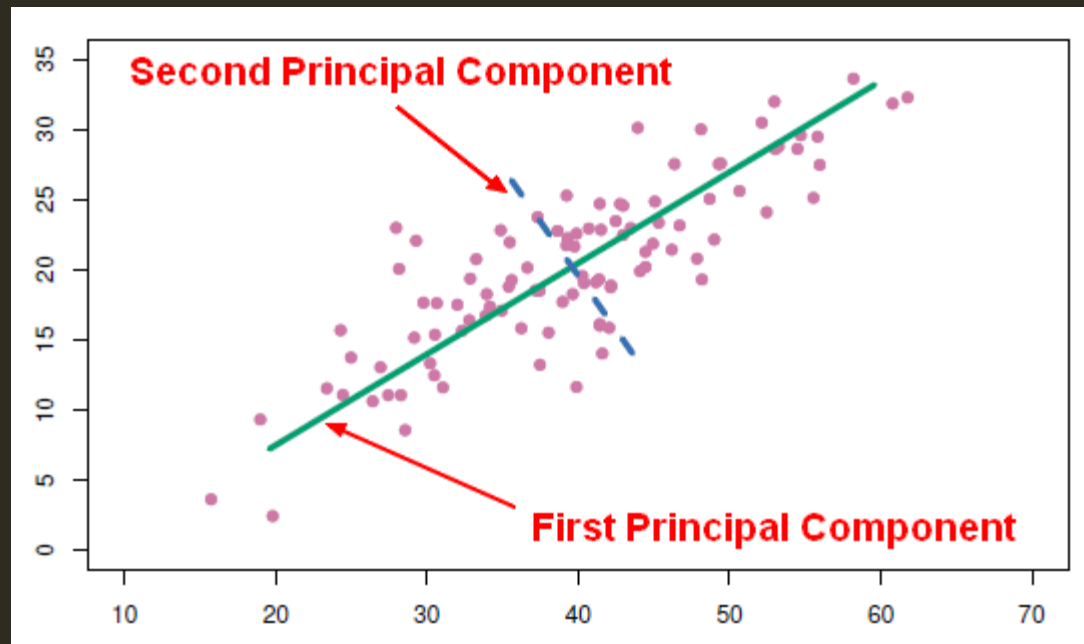
PCA VS. FACTOR ANALYSIS



- Analysis done on **Total Variance**
 - Goal: to summarize to minimum number of components for prediction
- Analysis done only on **Common Variance**
 - Goal: to identify underlying factors (dimensions) to reflect what the variables have in common

INTERDEPENDENCE METHOD

PRINCIPAL COMPONENT ANALYSIS

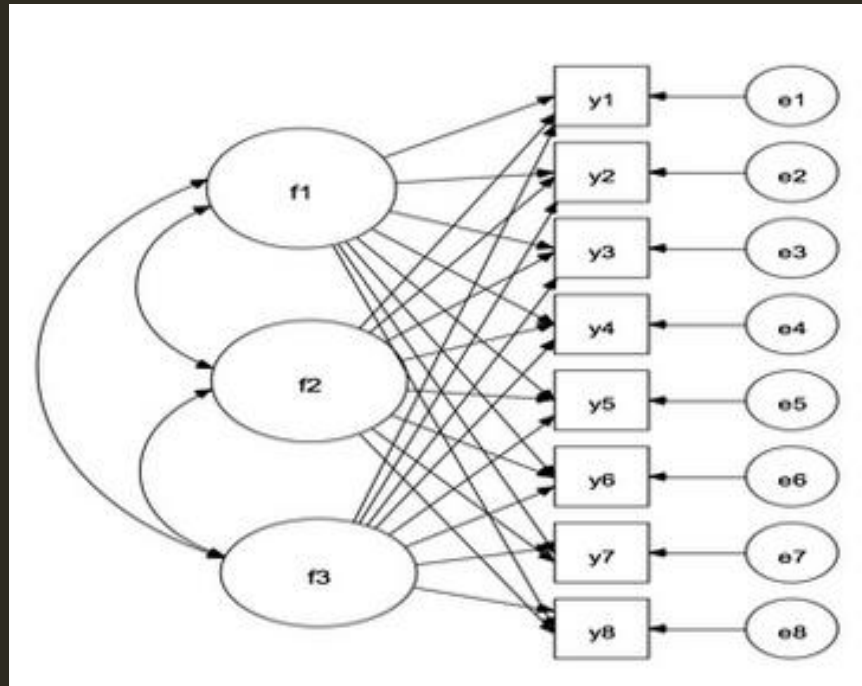


Identifikasi komponen utama berdasarkan **kombinasi linier variabel-variabel** yang memiliki variansi maksimum.

Contoh: Berapa komponen utama yang didapat dari berbagai variabel yang diteliti?

INTERDEPENDENCE METHOD

EXPLORATORY FACTOR ANALYSIS (EFA)

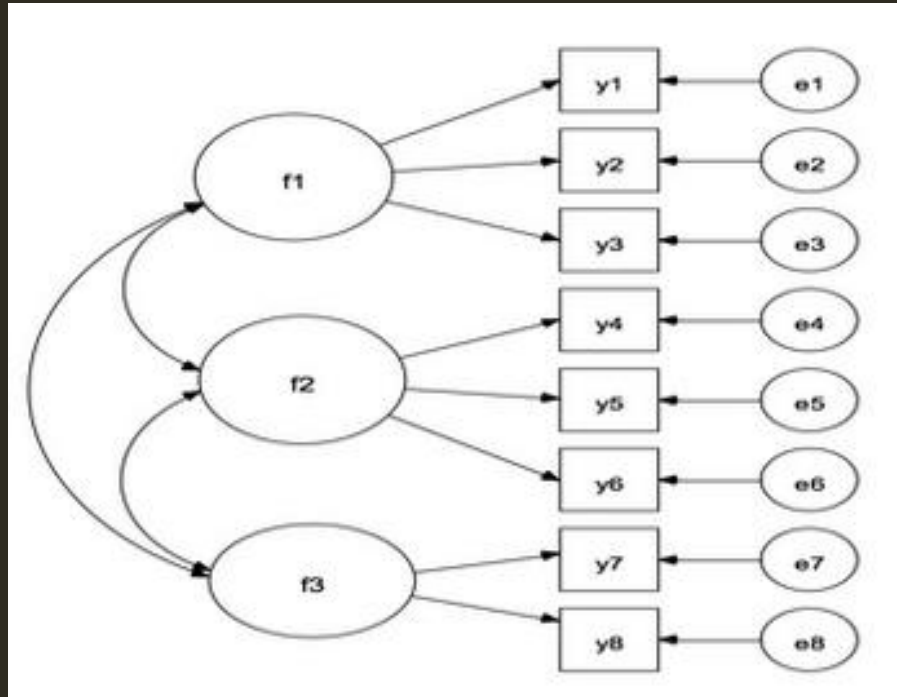


Mendapatkan pengelompokan **beberapa variable** ke dalam beberapa **faktor utama laten**, berdasarkan korelasi antar variabel-variabel tersebut.

Contoh: Pengelompokan berbagai berbagai tanda-tanda sakit flu yang sama menjadi beberapa indikasi utama.

INTERDEPENDENCE METHOD

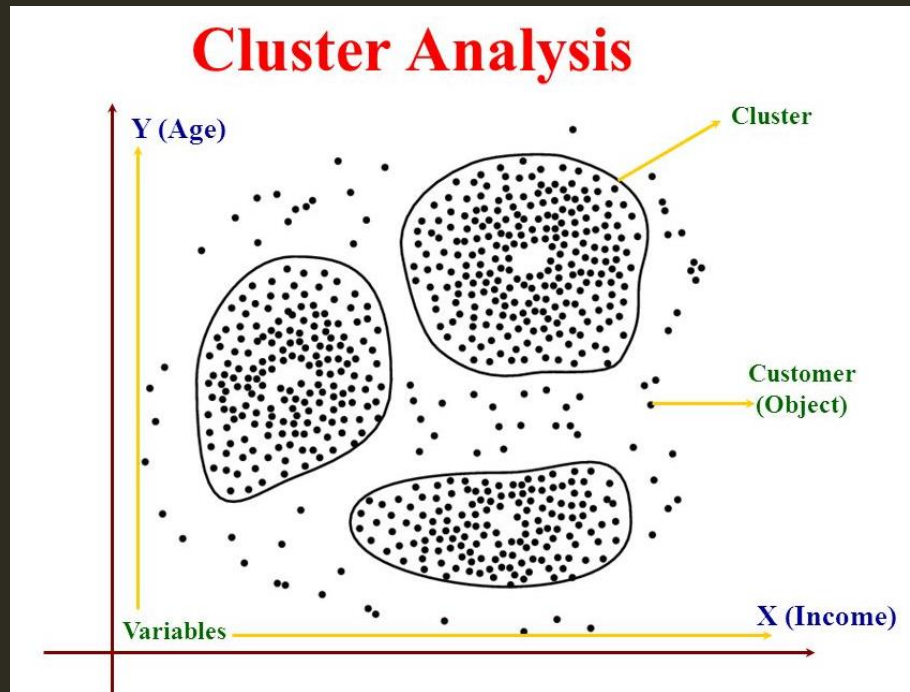
CONFIRMATORY FACTOR ANALYSIS (CFA)



Menguji kecocokan data dengan pengelompokan **beberapa variable** ke dalam **beberapa faktor utama**, dengan jumlah faktor yang sudah ditentukan sebelumnya.
Contoh: Apakah kecepatan layanan, keramahan layanan, akurasi layanan merupakan tiga faktor utama (sudah diduga sebelumnya ada 3 faktor) dari kualitas layanan Bank?

INTERDEPENDENCE METHOD

CLUSTER ANALYSIS



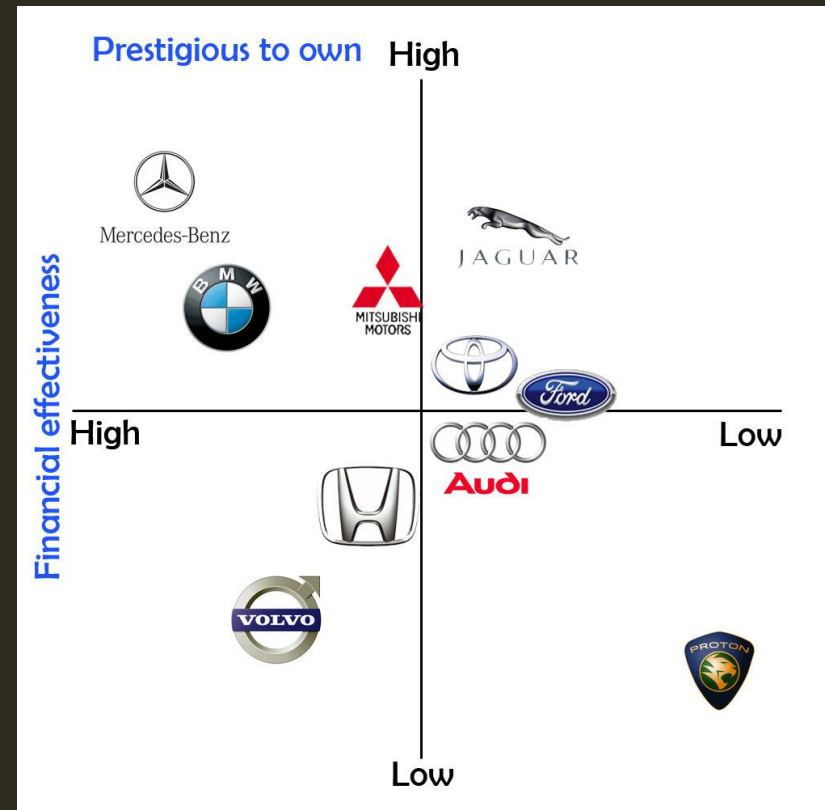
Identifikasi kelompok **responden (cases)** ke dalam beberapa kelompok utama sejenis berdasarkan data yang ada.

Contoh: Ada berapa segmen pembeli sepeda motor di Indonesia?

INTERDEPENDENCE METHOD MULTIDIMENSIONAL SCALING (MDS)

Menempatkan **OBJECTS** ke dalam **PERCEPTUAL MAP**, berdasarkan kemiripan atau perbedaan antar object (distance). Semua variabelnya **METRICS**.

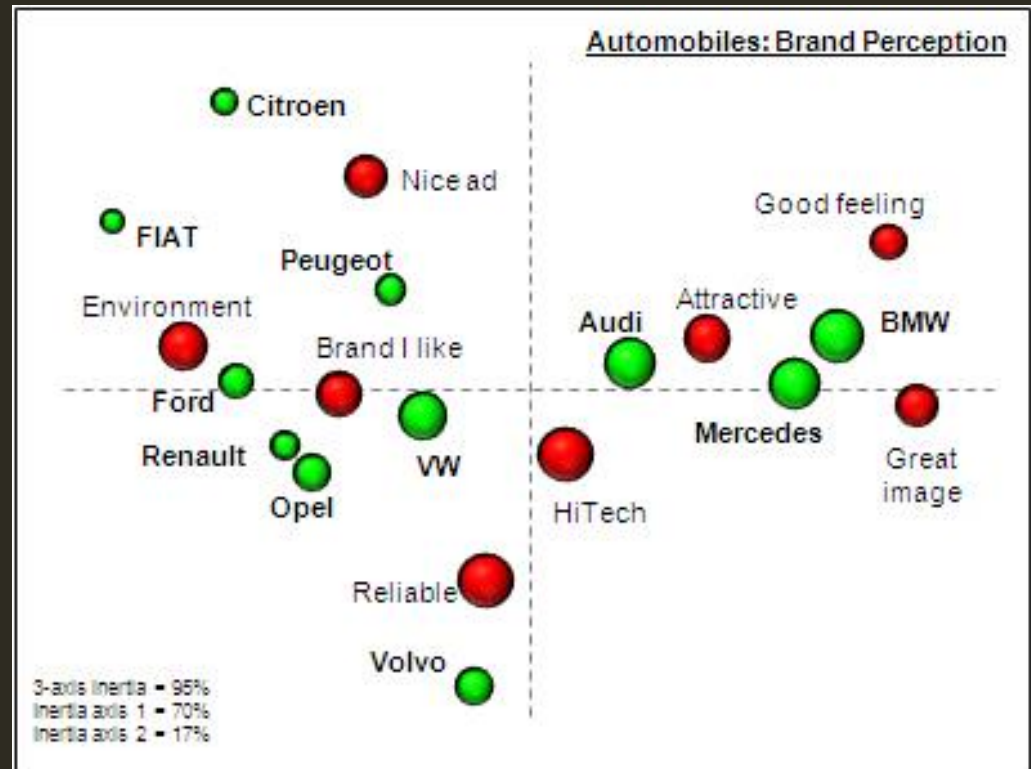
Contoh: Bagaimana posisi masing-masing merek mobil ditinjau dari aspek harga & gengsi menurut persepsi pelanggan?



INTERDEPENDENCE METHOD CORRESPONDENCE ANALYSIS (CA)

Menempatkan **OBJECTS** ke dalam **PERCEPTUAL MAP**, apabila variable-nya **NON METRICS**.

Contoh: Bagaimana posisi suatu produk dibandingkan dengan pesaing-pesaingnya?



THANK YOU

Ventje Jeremias Lewi Engel, M.T.
Prodi Informatika
Institut Teknologi Harapan Bangsa

