Multivariate Data Analysis: Overview and Applications

Multivariate Analysis

- **What is it?**
- **Why use it?**
  - ✓ **Measurement**
  - ✓ **Explanation & Prediction**
  - ✓ **Hypothesis Testing**

Basic Concepts of Multivariate Analysis:

- ❑ ***The Variate***

- ❑ ***Measurement Scales***
  - ▪ ***Nonmetric***
  - ▪ ***Metric***
- ❑ ***Multivariate Measurement***
- ❑ ***Measurement Error***
- ❑ ***Types of Techniques***

Variate $(Y') = X_1W_1 + X_2W_2 + \ldots + X_nW_n$

Each respondent has a variate value (Y').

The Y' _value_ is a _linear combination_ of the entire set of variables that best achieves the statistical objective.

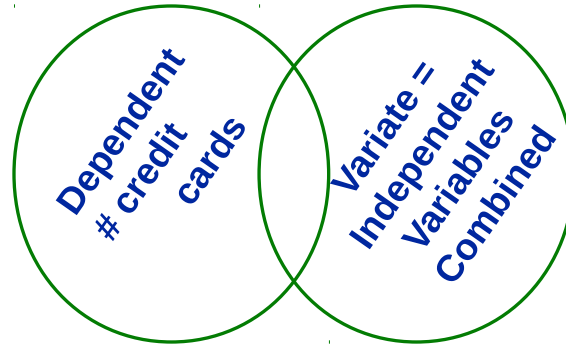**_Potential Independent Variables_:**

X_1 = income
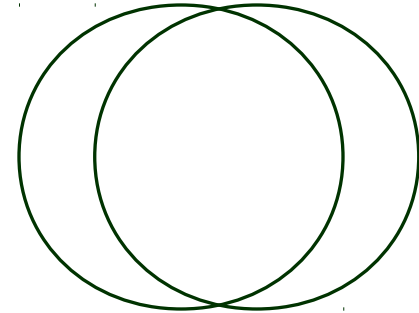X_2 = education
X_3 = family size
X_4 = occupation
X_5 = ? ?

Regression

"Bad"

Dependent # credit cards

Variate = Independent Variables Combined

"Good"

Discriminant

"Good"

Gp #1

Variate Values

Gp #2

Variate Values

"Bad"

Factor

"Bad"

Var #1

Var # 2

Var # 3

Var # 4

Var # 5

Var # 6

Small Overlap

"Good"

Variables 1, 3 & 6

Variables 2, 4 & 5
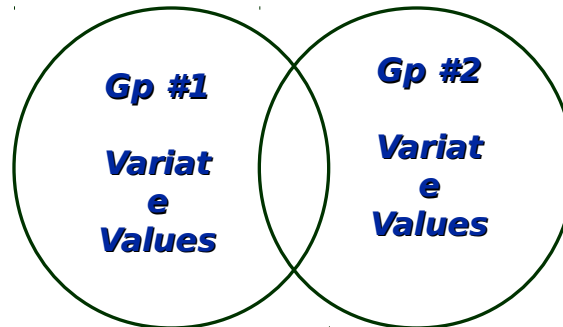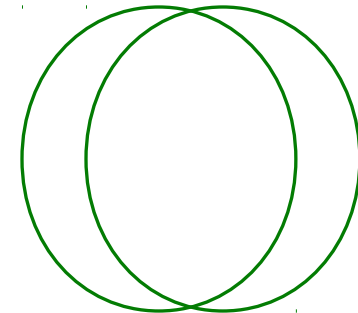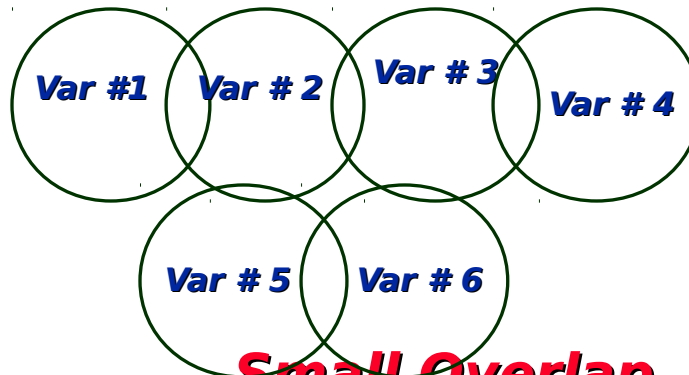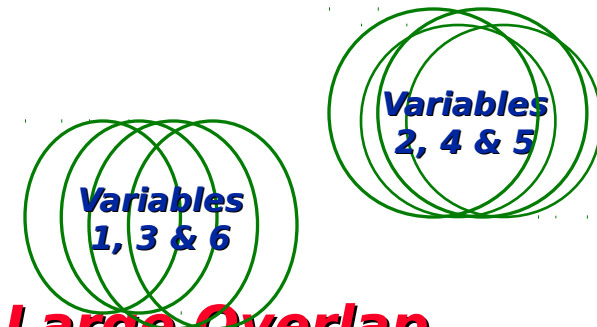
Large Overlap

Types of Data and Measurement Scales

```
                    ┌──────────┐
                    │   Data   │
                    └──────────┘
                         │
         ┌───────────────┴───────────────┐
┌──────────────┐                  ┌──────────────┐
│  Nonmetric   │                  │    Metric    │
│     or       │                  │     or       │
│ Qualitative  │                  │ Quantitative │
└──────────────┘                  └──────────────┘
       │                                 │
   ┌───┴────┐                        ┌───┴────┐
┌────────┐ ┌────────┐          ┌──────────┐ ┌────────┐
│Nominal │ │Ordinal │          │ Interval │ │ Ratio  │
│ Scale  │ │ Scale  │          │  Scale   │ │ Scale  │
└────────┘ └────────┘          └──────────┘ └────────┘
```
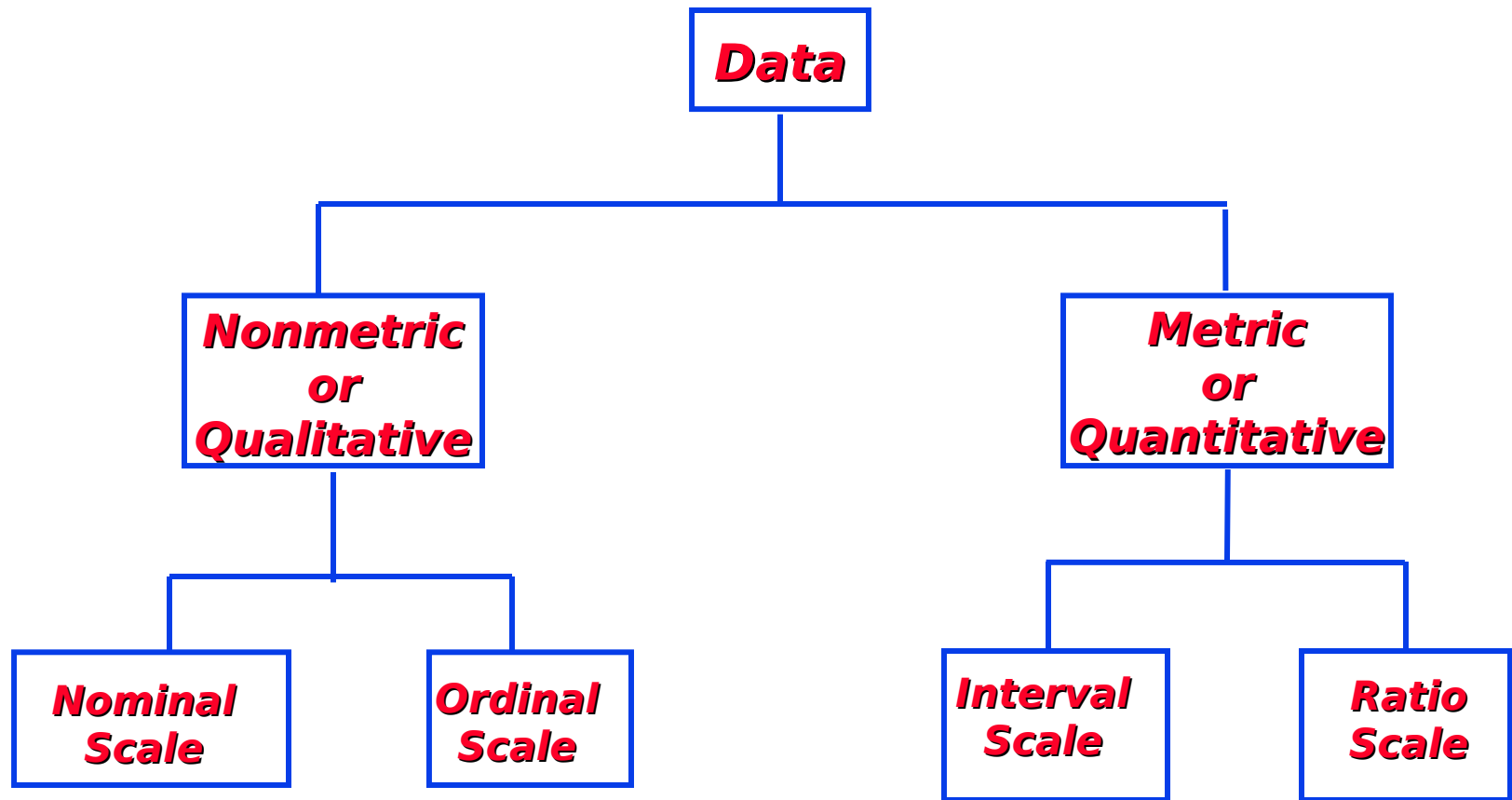
Multivariate Measurement

Multi-Item Constructs:

Independent

- ❑ Attitudes
- ❑ Psychographics
- ❑ Expectations
- ❑ Opinions

Dependent

- ❑ Likely Future Usage/Patronage
- ❑ Loyalty/Satisfaction
- ❑ Likely to Recommend
- ❑ Commitment – Job & Organization

Multivariate Measurement

"General Satisfaction Measures"

1. How satisfied are you with ____ ?
2. How likely are you to purchase/return ____ in the future?
3. How likely are you to recommend ____ to a friend?

"Actionable Satisfaction Measures"

How satisfied are you with . . . .
1. service
2. product/service selection
3. facility
4. personnel
5. hours
6. location

The Secure Customer Index™

% Very Satisfied   **% Definitely Buy Again**

"Secure Customers"

% Definitely Recommend to Others
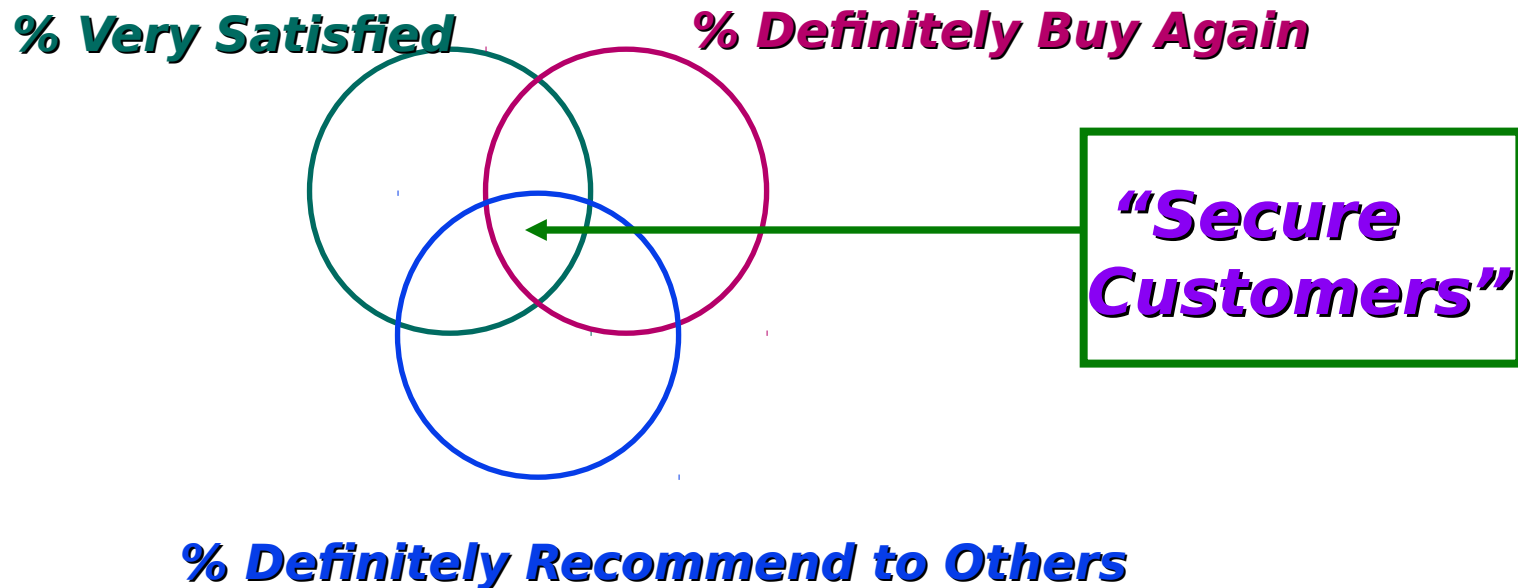
The Secure Customer Index ™
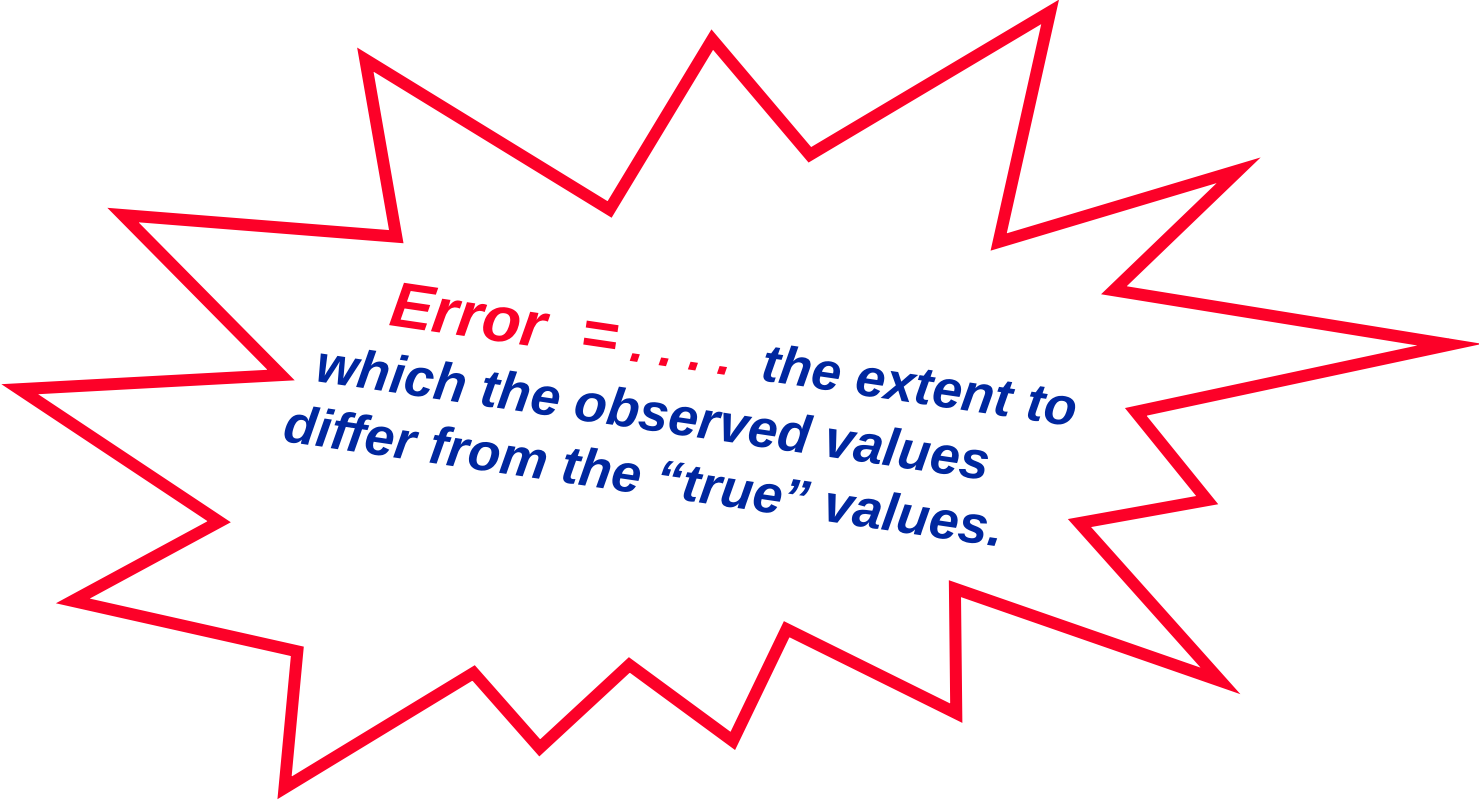
Three Components:

1. Percentage of customers who are VERY SATISFIED.

2. Percentage of customers who will DEFINITELY REPURCHASE
   your company's products/services in the future.

3. Percentage of customers who will DEFINITELY RECOMMEND
   your company to others if the occasion arises.

Very Dissatisfied Definitely Not Repurchase Definitely Not Recommend				Very Satisfied Definitely Repurchase Definitely Recommend

Burke, Inc. ™

Multivariate Measurement
Multi-Item Constructs

Sources of Items/Statements:

- Published Scales?

- Prior Research ?

- Theory?

- Practice – Current Business Situation ?

- Intuition ?

Types of Errors:

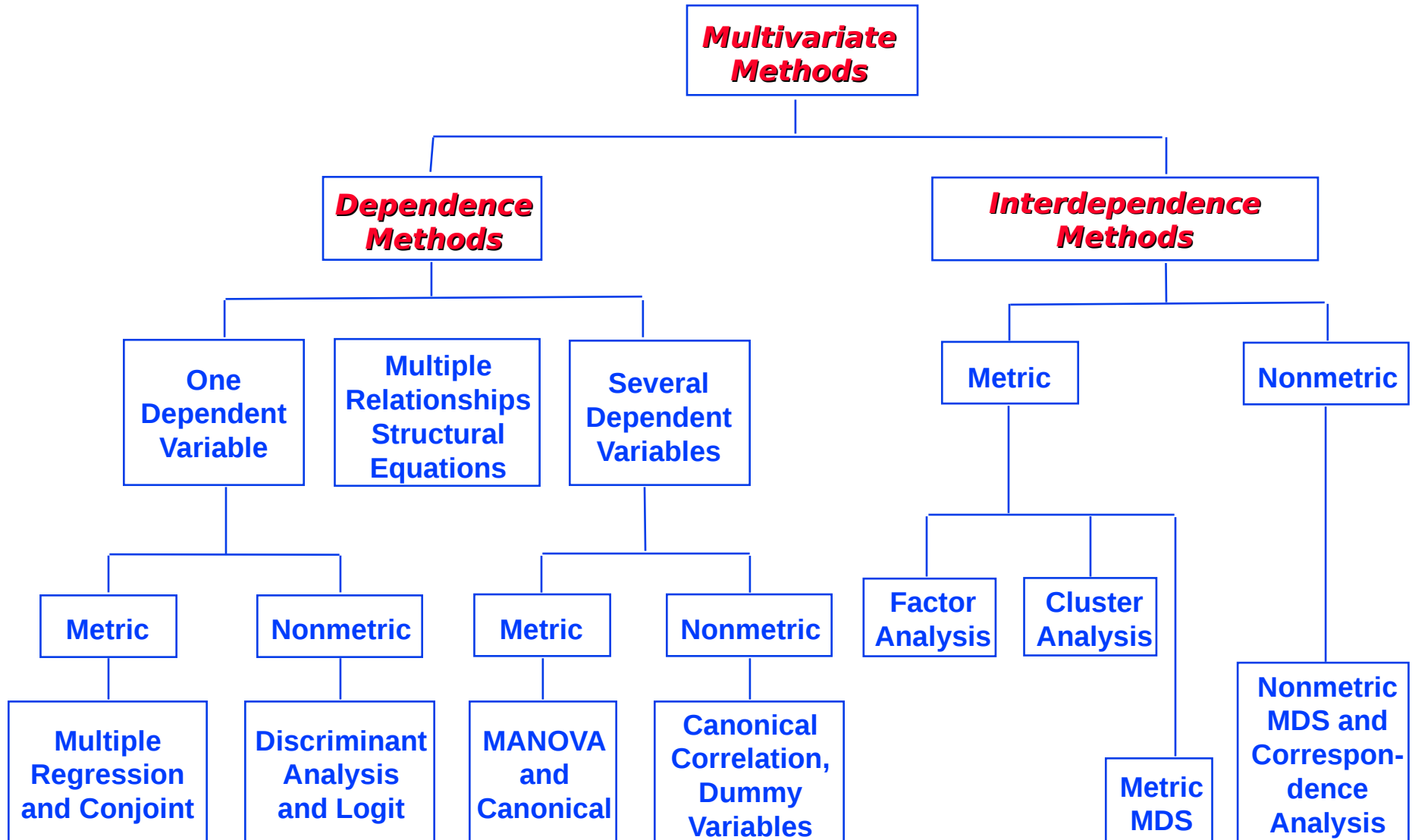- **Non-response** = *refusal, sampling, problem definition, etc.*
- **Response** = *respondent or interviewer*
- **Data Collection**
  - *Construct Development*
  - *Scaling Measurement*
  - *Questionnaire Design/Sequence, etc.*
- **Data Analysis**
- **Interpretation**

Two Types of Multivariate Techniques:

1. **Dependence**

2. **Interdependence**

Classification of Multivariate Methods:

Multiple Regression

A metric dependent variable

is predicted by several

metric independent variables.

Multiple Regression

Dependent = # of credit cards

Independent Variables

X_1 = income

X_2 = education

X_3 = family size

X_4 = occupation

X_5 = ? ?

Example of Multiple Regression Application in Financial Services Industry

Outcome Measures:

1. Customer Satisfaction

2. Likely to Recommend

3. Future Purchases

Bank Selection Factors:

1. Trust
2. Competent Employees
3. Excellent Customer Service
4. Good Financial Services
5. Friendly Employees
6. Interest Rates Paid
7. Convenient Locations
8. Interest Rates Charged
9. Care about Community
10. Open When You Want
11. Innovative Services

Discriminant Analysis

A non-metric (categorical)

dependent variable is predicted

by

several metric independent

variables.

Examples:

• Gender – Male vs. Female

• Heavy Users vs. Light Users

• Purchasers vs. Non-purchasers

• Good Credit Risk vs. Poor Credit Risk

• Member vs. Non-Member

Discriminant Analysis

Dependent Variable = **Credit Risk**

(Good Risk vs. Bad Risk)

Independent Variables:

X_1 = income

X_2 = education

X_3 = family size

X_4 = occupation

X_5 = ? ?

Example of Discriminant Analysis Application in Consumer Products Industry

Bath Soap

Product Features:

- **Pleasant smell.**
- **Skin creme feel.**
- **Lathers well.**
- **Cleans well.**
- **Deodorant.**
- **Rinses off easily.**
- **Moisturizing.**
- **No soap residue in soap dish.**
- **No ring around sink or tub.**

Outcome Measure:
- ✓ **Will Purchase**
- ✓ **Will Not Purchase**

MANOVA

Several metric dependent variables are predicted by a set of nonmetric (categorical) independent variables.

ANOVA/MANOVA

Metric Dependent Variables

Company

Progressive   _ _ _ _ _   Not Progressive

Modern   _ _ _ _ _   Traditional

Product

High Quality   _ _ _ _ _   Low Quality

Expensive   _ _ _ _ _   Not Expensive

Salesforce

Friendly   _ _ _ _ _   Not Friendly

Knowledgeable   _ _ _ _ _   Not Knowledgeable

ANOVA/MANOVA

Non-Metric Independent Variables

Type of Ad   =   Treatment

X_1 Humorous   Non-Humorous

X_2 Color   Black & White

X_3 People   No People

X_4 Black   White
        Models   Mixed Models
            Only   Only
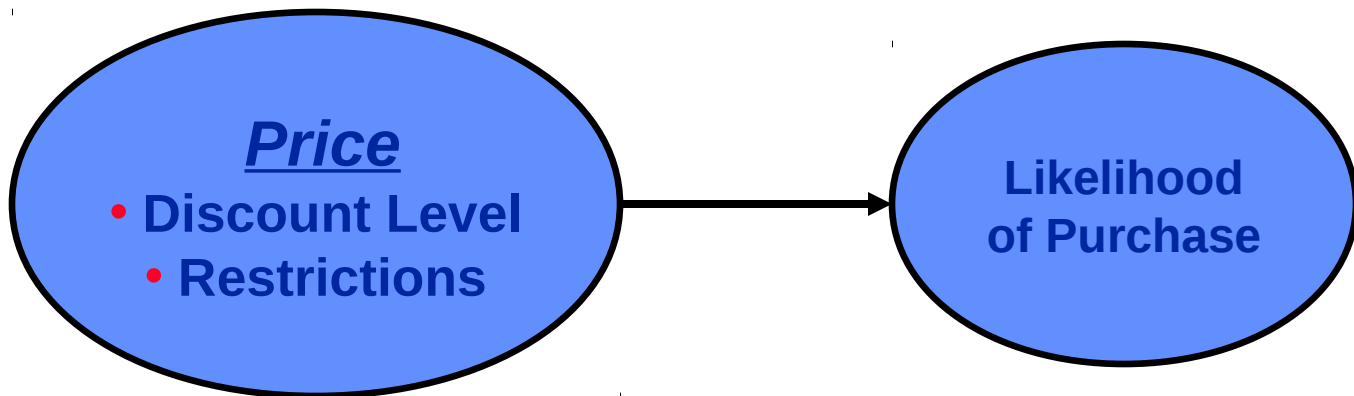
CANONICAL ANALYSIS

Several metric dependent variables

are predicted by several metric independent

variables.

CANONICAL ANALYSIS

Credit Usage   =   **_Demos, etc._**

✔ **# cards**

✔ **average balance**

✔ **amount of purchases**

✔ **types of purchases**
  - **food**
  - **lodging**
  - **cars**
  - **travel**
  - **appliances**

✔ **age**

✔ **income**

✔ **education**

✔ **gender**

✔ **interest rate**

✔ **debt amount**

✔ **# late payments**

CONJOINT ANALYSIS

. . . is used to understand respondents' preferences for products and services.

In doing this, it determines the importance of _both_:

attributes and

levels of attributes

. . . based on a smaller subset of combinations of attributes and levels.

Conjoint Analysis

Three Automobile Attributes:

Price	Style	Color
low	Sports Car	red
medium	Sedan	blue
high	SUV	black

3 x 3 x 3 = 27

Vanilla Coke ?

Cherry Coke ?

C 2 ?

Conjoint Analysis

Typical Applications:

❖ **Soft Drinks**
❖ **Candy Bars**
❖ **Cereals**
❖ **Beer**
❖ **Apartment Buildings; Condos**
❖ **Solvents; Cleaning Fluids**

Factor Analysis

. . . . analyzes the structure of the interrelationships among a large number of variables to determine a set of common underlying dimensions (factors).

Cluster Analysis

. . . . groups objects (respondents, products, firms, variables, etc.) so that each object is similar to the other objects in the cluster and different from objects in all the other clusters.

Multivariate Analysis Learning Checkpoint:

1. **What is multivariate analysis?**

2. **Why use multivariate analysis?**

3. **Why is knowledge of measurement scales important in using multivariate analysis?**

4. **What basic issues need to be examined when using multivariate analysis?**

5. **Describe the process for applying multivariate analysis.**

Multidimensional Scaling
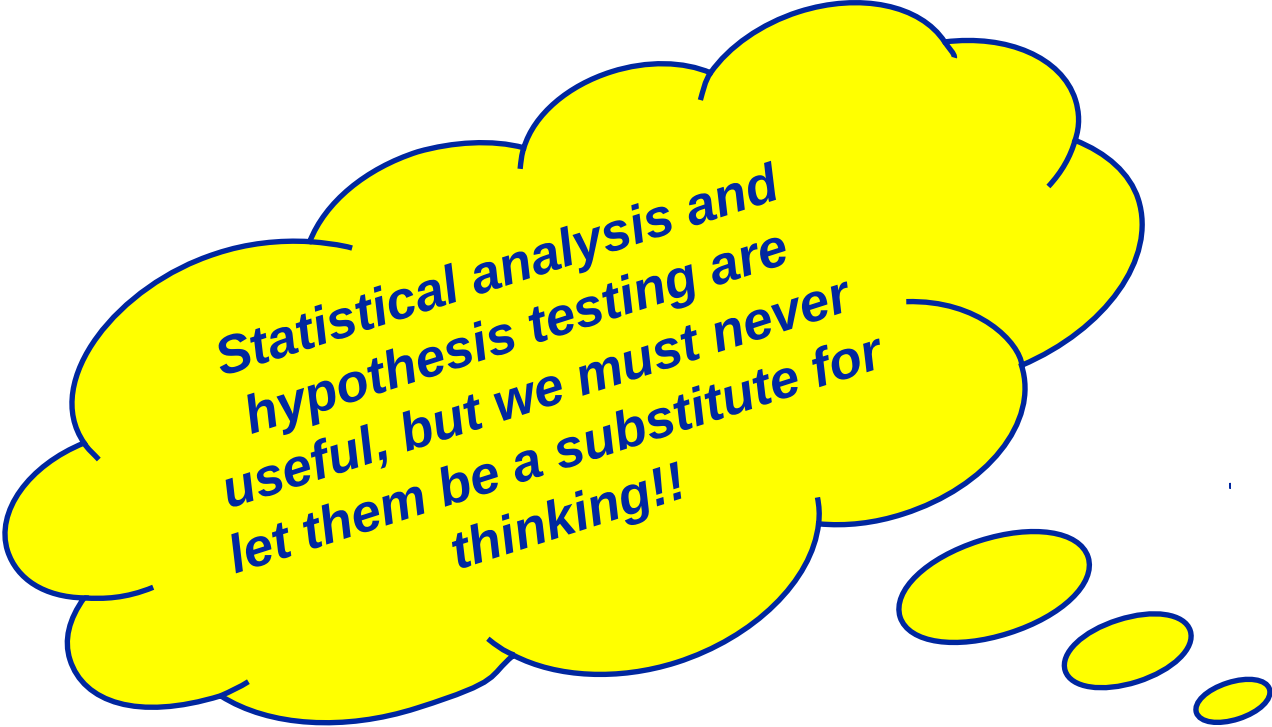
. . . identifies "unrecognized" dimensions that affect purchase behavior based on customer judgments of:

- *similarities or*
- *preferences*

and transforms these into distances represented as perceptual maps.

Correspondence Analysis

. . . uses non-metric data and evaluates either linear or non-linear relationships in an effort to develop a perceptual map representing the association between objects (firms, products, etc.)   and a set of descriptive characteristics of the objects.

Guidelines for Multivariate Analysis and Interpretation:

- ❑ **Establish Practical Significance as well as Statistical Significance.**

- ❑ **Sample Size Affects all Results.**

- ❑ **Know Your Data.**

- ❑ **Strive for Model Parsimony.**

- ❑ **Look at Your Errors.**

- ❑ **Validate Your Results.**

A Structured Approach to Multivariate Model Building:

Stage 1: **Define the Research Problem, Objectives, and**

Multivariate Technique(s) to be Used

Stage 2: **Develop the Analysis Plan**

Stage 3: **Evaluate the Assumptions Underlying the**

Multivariate Technique(s)

Stage 4: **Estimate the Multivariate Model and Assess**

Overall Model Fit

Stage 5: **Interpret the Variate(s)**

Stage 6: **Validate the Multivariate Model**

Description of Employee Survey Variables

XYZ Restaurant

Variable Description	Variable Type

Work Environment Measures

X_1	I am paid fairly for the work I do.	Metric
X_2	I am doing the kind of work I want.	Metric
X_3	My supervisor gives credit an praise for work well done.	Metric
X_4	There is a lot of cooperation among the members of my work group.	Metric
X_5	My job allows me to learn new skills.	Metric
X_6	My supervisor recognizes my potential.	Metric
X_7	My work gives me a sense of accomplishment.	Metric
X_8	My immediate work group functions as a team.	Metric
X_9	My pay reflects the effort I put into doing my work.	Metric
X_{10}	My supervisor is friendly and helpful.	Metric
X_{11}	The members of my work group have the skills and/or training to do their job well.	Metric
X_{12}	The benefits I receive are reasonable.	Metric

Relationship Measures

X_{13}	Loyalty – I have a sense of loyalty to Samouel's restaurant.	Metric
X_{14}	Effort – I am willing to put in a great deal of effort beyond that expected to help Samouel's restaurant to be successful.	Metric
X_{15}	Proud – I am proud to tell others that I work for Samouel's restaurant.	Metric

Classification Variables

X_{16}	Intention to Search	Metric
X_{17}	Length of Time an Employee	Nonmetric
X_{18}	Work Type = Part-Time vs. Full-Time	Nonmetric
X_{19}	Gender	Nonmetric
X_{20}	Age	Nonmetric
X_{21}	Performance	Metric

XYZ Restaurant

Variable Description	Variable Type
Restaurant Perceptions	
X_1 Excellent Food Quality	Metric
X_2 Attractive Interior	Metric
X_3 Generous Portions	Metric
X_4 Excellent Food Taste	Metric
X_5 Good Value for the Money	Metric
X_6 Friendly Employees	Metric
X_7 Appears Clean & Neat	Metric
X_8 Fun Place to Go	Metric
X_9 Wide Variety of menu Items	Metric
X_{10} Reasonable Prices	Metric
X_{11} Courteous Employees	Metric
X_{12} Competent Employees	Metric
Selection Factor Rankings	
X_{13} Food Quality	Nonmetric
X_{14} Atmosphere	Nonmetric
X_{15} Prices	Nonmetric
X_{16} Employees	Nonmetric
Relationship Variables	
X_{17} Satisfaction	Metric
X_{18} Likely to Return in Future	Metric
X_{19} Recommend to Friend	Metric
X_{20} Frequency of Patronage	Nonmetric
X_{21} Length of Time a Customer	Nonmetric
Classification Variables	
X_{22} Gender	Nonmetric
X_{23} Age	Nonmetric
X_{24} Income	Nonmetric
X_{25} Competitor	Nonmetric
X_{26} Which AD Viewed (#1, 2 or 3)	Nonmetric
X_{27} AD Rating	Metric
X_{28} Respondents that Viewed Ads	Nonmetric

DESCRIPTION OF DATABASE VARIABLES

Variable Description	Variable Type

PERCEPTIONS OF HATCO

X_1	Delivery speed	Metric
X_2	Price level	Metric
X_3	Price flexibility	Metric
X_4	Manufacturer's image	Metric
X_5	Overall service	Metric
X_6	Salesforce image	Metric
X_7	Product quality	Metric

PURCHASE OUTCOMES

X_9	Usage level	Metric
X_{10}	Satisfaction level	Metric

PURCHASER CHARACTERISTICS

X_8	Size of firm	Nonmetric
X_{11}	Specification buying	Nonmetric
X_{12}	Structure of procurement	Nonmetric
X_{13}	Type of industry	Nonmetric
X_{14}	Type of buying situation	Nonmetric

Multivariate Analysis
Learning Checkpoint:

1. **What is multivariate analysis?**
2. **Why use multivariate analysis?**
3. **Why is knowledge of measurement scales important in using multivariate analysis?**
4. **What basic issues need to be examined when using multivariate analysis?**
5. **Describe the process for applying multivariate analysis.**