# Segmentation

- *Why do we segment?*
- *When it is mostly important?*

## A Definition

Market Segmentation is concerned with individual or intergroup differences in response to marketing mix variables. The managerial presumption is that if these response differences exist, can be identified, are reasonably stable over time and the segments can be efficiently reached the firm may increase its sales and profits beyond those obtained by assuming market homogeneity.

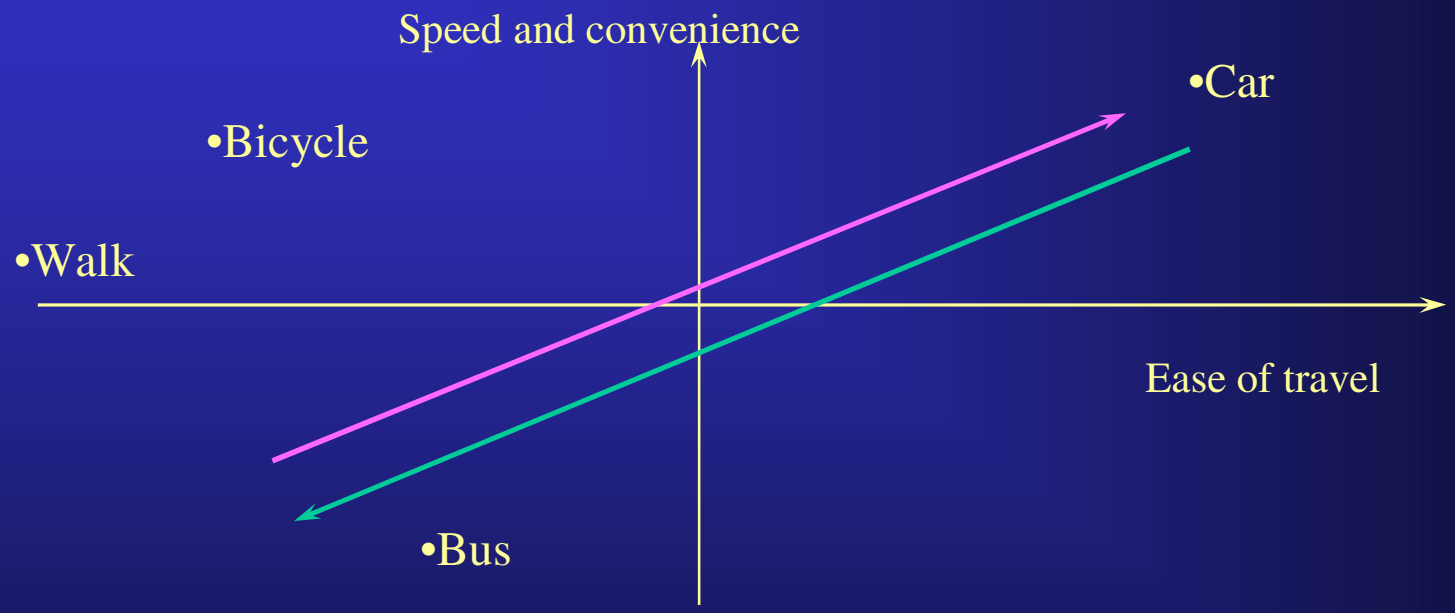
## Du-Pont's Definition

"A group of customers anywhere along the distribution chain who have common needs and values - who will respond similarly to our offerings and who are large enough to be strategically important to our business."

# דוגמא לפילוח שוק

| פלח שוק | מאפיינים דמוגרפיים | מאפיינים התנהגותיים | מאפיינים פסיכוגרפיים | מותגים מועדפים |
|---|---|---|---|---|
| לקוחות מוכווני חיסכון | גברים | משתמשים כבדים | עצמאיים ומוכווני ערך | מותגים במבצע (נאמנות נמוכה) |
| מוכווני מניעת עששת | משפחות גדולות | משתמשים כבדים | שמרניים ובעלי מודעות לבריאות | קרסט |
| מוכווני הופעה חיצונית | צעירים | מעשנים | פעילים בחברה | אקווה פרש |
| טעם ורעננות | צעירים או ילדים | | נהנתנים, אוהבי נוחות | קולגייט |

# Identification of strategic Benefits from perceptual Mapping (Hauzer and Wishniewski, 1979)




Interpret the conceptual map:

What would be your strategy if you were a car dealer, how would you position a new transportation service.

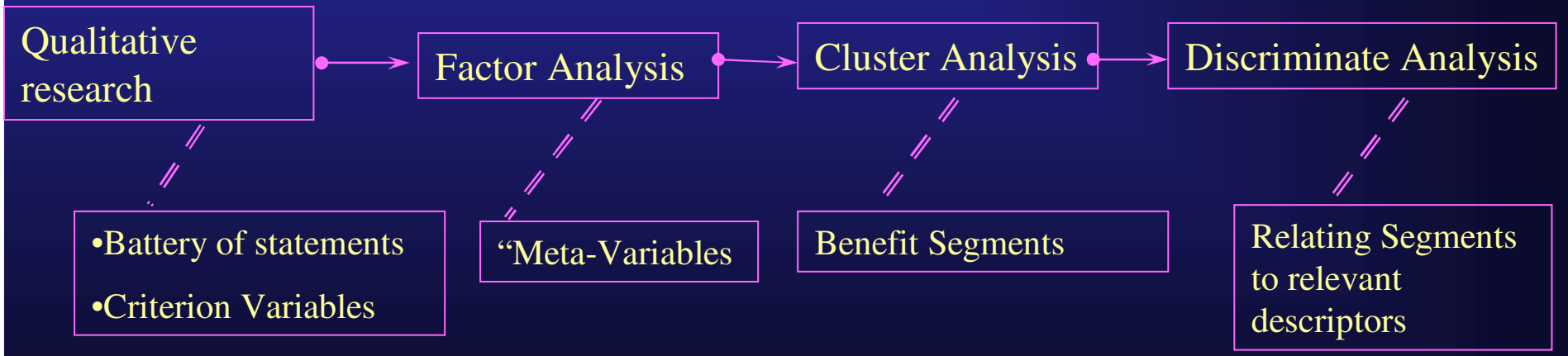
# Segmentation - Overview

Approaches

**A-Priori Segmentation** - grouping customers according to predefined criteria (e.g., demographics, psychographics, behavior, lifestyle).

**Post-hoc Segmentation** - grouping customers according to the similarity of their multivariate profiles which include variables such as attributes, benefits sought, preferences.

## How can we create a segment?



# From Attributes to Factors - Brief Explanation of Factor Analysis

F.A finds combinations that offer reduction: It creates new, fewer in number, variables from original set. It helps to identify conceptual or benefit dimensions underlying expressed measure of product perceptions and preferences.

**Example**: Reduce the data about pupils in an ordinary school in Tel-Aviv.

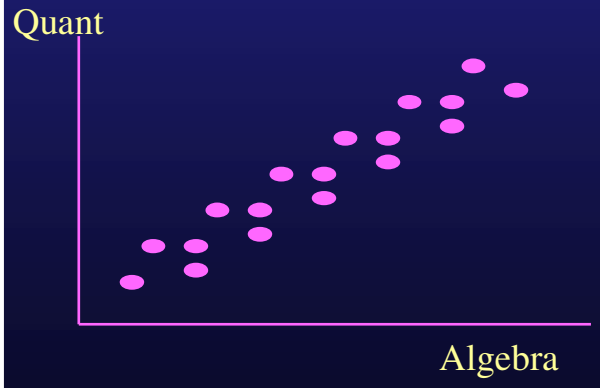
| **Attributes** | **Factors** |
|---|---|
| Weight | F1: |
| Height | |
| Education | F 2: |
| Income | |

# Intuitive understanding of Factor Analysis

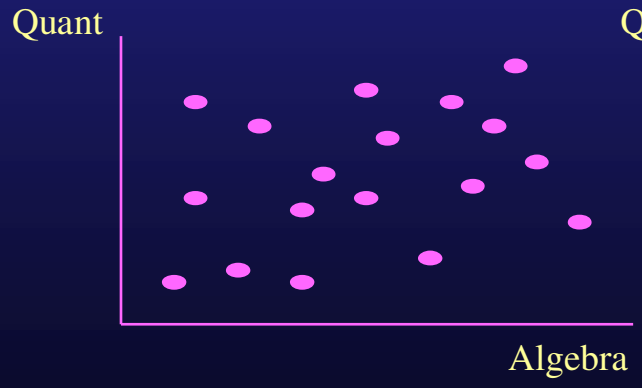
| | Algebra | Geometry | Qaunt | Writing | Spelling | Reading |
|---|---|---|---|---|---|---|
| John | 10 | | | 90 | | |
| Jane | 90 | | | 10 | | |

Q: Can you estimate Jane's achievements in geometry and spelling? What is your confidence with this prediction?

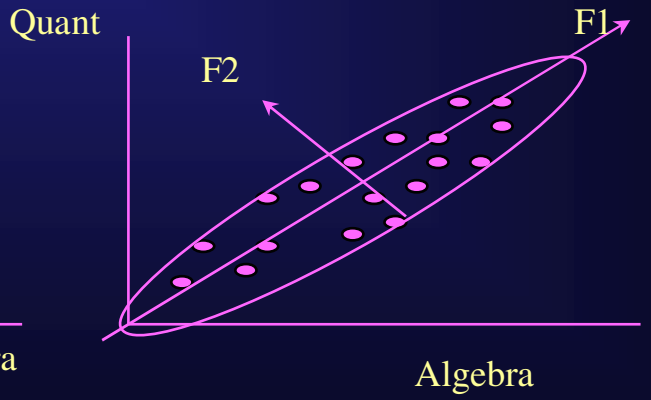
High correlation



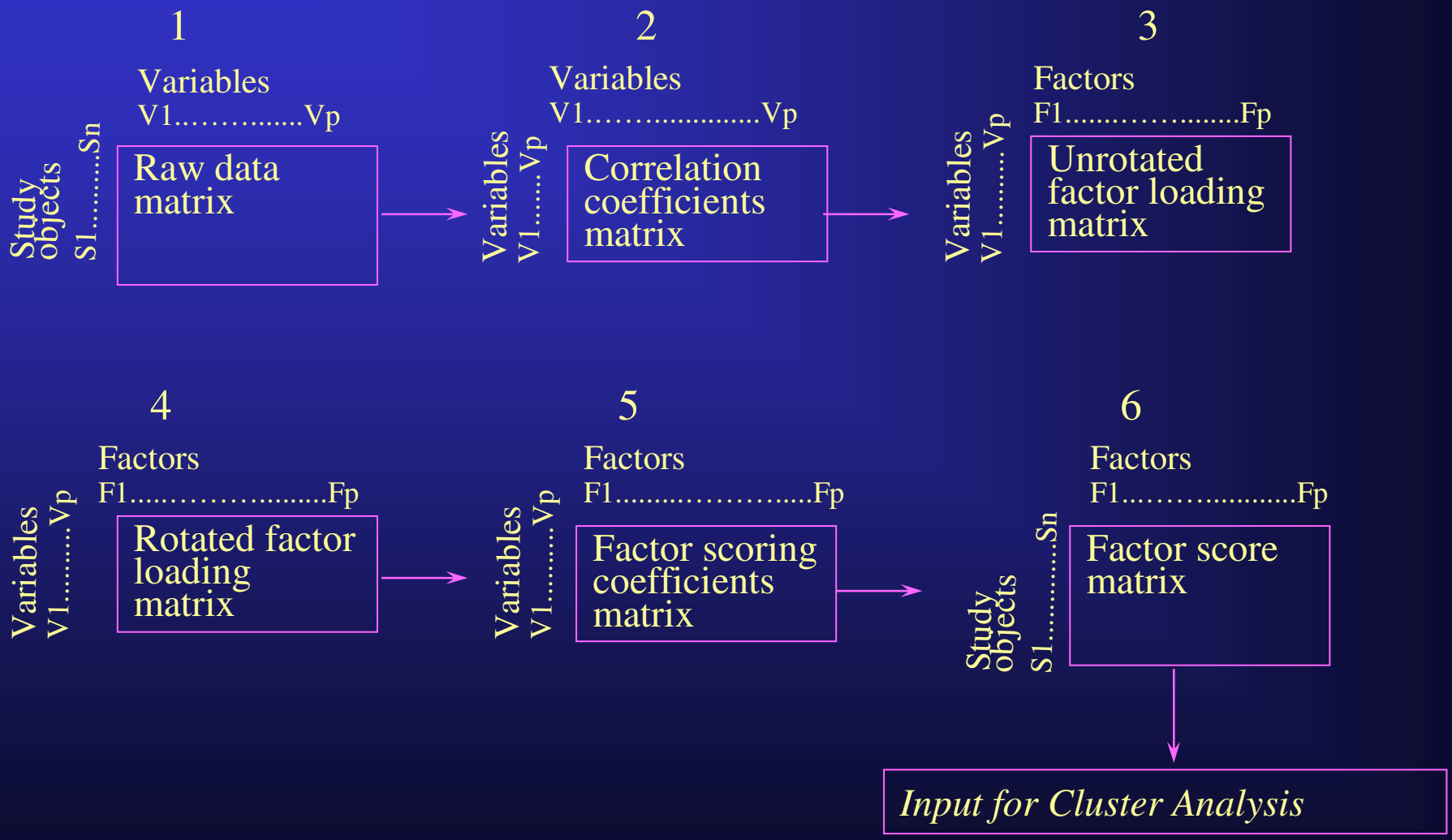

No correlation




An attempt to capture most of the information with two factors



# Factor Analysis - Overview



# Factor Analysis - The key Concepts

## The Principle Component Model

$Z_1 = a_{11}F_1 + a_{12}F_2 + \ldots\ldots\ldots\ldots a_{1n}F_n$
$Z_1 = a_{21}F_1 + a_{22}F_2 + \ldots\ldots\ldots\ldots a_{2n}F_n$
.
.
.
$Z_n = a_{n1}F_1 + a_{n2}F_2 + \ldots\ldots\ldots\ldots a_{nn}F_n$

$\Rightarrow$ $Z_j = a_{j1}F_i + a_{j2}F_2 + \ldots\ldots a_{jn}F_n$

*Z is standardized (mean= 0, S.D=1)*

Zi - Variable, Fi - factor, aij - loadings or correlation between Z and F's

While both models look similar to regression equation, the predictor variables in factor analysis (F1..Fn) are actually hypothetical concepts, values which must be estimated from the observed data. Technically the basic problem is to estimate the pattern loadings a1..an.

**Structure Loadings** are the correlations between variables and factors. For the case of uncorrelated factors the structure loadings and the pattern loadings (a..) are identical.

Orthogonality (uncorrelated factors):

$$S_j^2 = 1 = a_{j1}^2 + \ldots a_{jn}^2$$

Where a squared "a" represent the contribution of Factor Fa to the variance of Zj.

and $S^2$ is the total variance.

# Continued

Communality: The variance of variable Zj which is "common" (shared) with other variables.

**Communality (Zj)= $a^2_{j1}$+ $a^2_{j2}$ + $a^2_{j3}$ +...**

What is the contribution of each factor to the communality of $Z_3$?

What is the contribution of the factor to the total communality of?

# Example Factor Analysis

In a psychografic test 500 respondents were asked to indicate the degree to which each of the several psychografic statements described them. A 10 points scale was used and the data appears as shown, for convenience only 5 statements will be used.

| | Statement | Respondents | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4....500 | |
| 1 | I often try the latest hairdo styles when they change. | 8 | 4 | 1 | .... | 8 |
| 2 | I shop a lot for specials. | 3 | 6 | 8 | .... | 6 |
| 3 | When I must choose between the two, I usually dress for fashion, not for comfort. | 9 | 2 | 1 | .... | 8 |
| 4 | A Person can save a lot of money by shopping around for bargaines | 4 | 6 | 9 | .... | 6 |
| 5 | I find myself checking the prices in the grocery store even for small items. | 2 | 5 | 8 | .... | 6 |

Q1: How do you code the table into SPSS?

# The correlation matrix

*Variables*

| | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|
| **1** | 1 | .01 | .97 | .44 | .02 |
| **2** | .01 | 1 | .15 | .69 | .86 |
| **3** | .97 | .15 | 1 | .51 | .12 |
| **4** | .44 | .69 | .51 | 1 | .78 |
| **5** | .02 | .86 | .12 | .78 | 1 |

*Variables*

Q2: What are the characteristics of this matrix?

Q3: What is the interpretation of the high value of the following correlation factors: $r_{13}$=.97, $r_{24}$=.69, $r_{45}$=.78, $r_{25}$=.86?

# Principal component analysis for the psychografic data

In essence we can define each of the five (psychografic) variables in terms of five hypothetical factors (how?). However this is simply replacing the variables and result in loosing the parsimonious representation. We would prefer a situation in which few factors account for a significant portion of the total variance.

| | P1 | P2 | P3 | P4 | P5 | Variance |
|---|---|---|---|---|---|---|
| **1** | .581 | .8064 | .0276 | -.0645 | -.0852 | 1 |
| **2** | .7671 | -.5448 | .3193 | .1118 | -.0216 | 1 |
| **3** | .6724 | .7260 | .1149 | -.0072 | .0862 | 1 |
| **4** | .9324 | -.1043 | -.3078 | .1582 | .000 | 1 |
| **5** | .7911 | -.5582 | -.0647 | -.2413 | .0102 | 1 |
| **Variance** | 2.8733 | 1.7966 | .2148 | .100 | .0153 | 5.000 |
| **%** | 57.5 | 35.9 | 4.3 | 2.0 | 0.3 | 100 |

# Principle component analysis for the psychografic data (continue)

Using Only first two components:

Communality for variable 1 = $.58^2+.81^2=.99$

Variance explained by factor 1 = $.58^2+.77^2+.67^2+.93^2+.79^2=2.87$

Total variance in the data = 5 (why?)

Percent of the variance explained = 2.87/5=57.5%

**The total variance explained by factors 1 and 2 = 57.5+25.9=93.4%**

# The terminology that often used in the context of principle component analysis

**A Factor** is a dimension underlying several variables. Analytical, it is a linear combination of the variables:

$F_1=W_1X_1+W_2X_2+...$

Where: $F_1$ - factor1, $X_j$ - the variables of the study (5 in our example), $W_j$ - weights used to combine the individual scores.

The various methods of factor analysis are distinguished by the manner in which the weights Wj are determined.

**A Factor score**: The score of a respondent on a factor. If we decide to settle with two factors we will have two factor scores for each of the 500 respondents.

**A Factor loading**: The correlation between a factor and a variable. In our example the variables #2, #4, #5 are highly correlated (loaded) with factor #1.

**Labeling Factors**: The art of segmentation; consists of selecting a term which best describes all the variables that load highly a factor. Factor #1 may be labeled as "price conscious": and factor #2 as " fashion conscious".

**The proportion of total variance of a certain variable** accounted for by a factor may be obtained by squaring the loading. In our example factor #1 explains $.9234^2=86.94\%$ of the variance in variable 4.

# The terminology that often used in the context of principle component analysis (Continued)

**Communality ($h^2$)**: the proportion of the total variance of a certain variable explained by the factors, it equals the sum of the squared loadings for a variable. In our example the communality of variable #4 (by using two factors is $.9324^2+.1043^2=.8802$.

**The Eigen Value:** the sum of squares in one column. It gives a measure of the amount of variation in all variables accounted for by one factor. In our example the eigen value of factor #1 = $.58^2+.77^2+.67^2+.79^2=2.87$ Note that the two factors together explain 93.4% of the total variance of the data (why?).

# How many factors to retain in the final solution?

There are no completely generalizable rules governing this situation. The idea is to stop the factoring when reaching factors that account for trivial variance.

**There are two dominant heuristics:**

- **The Eigen value greater then 1**: use factors that their eigen value is greater then one
- **The Scree test**:


*The curve levels off with very minor differences between successive factors*

# Rotation

Objective: Rotation of factors is undertaken to facilitate their interpretation and assign to verbal labels. It simplifies the rows and/or columns of the factor matrix in such a way that the factor loadings become closer to 0 or to 1.

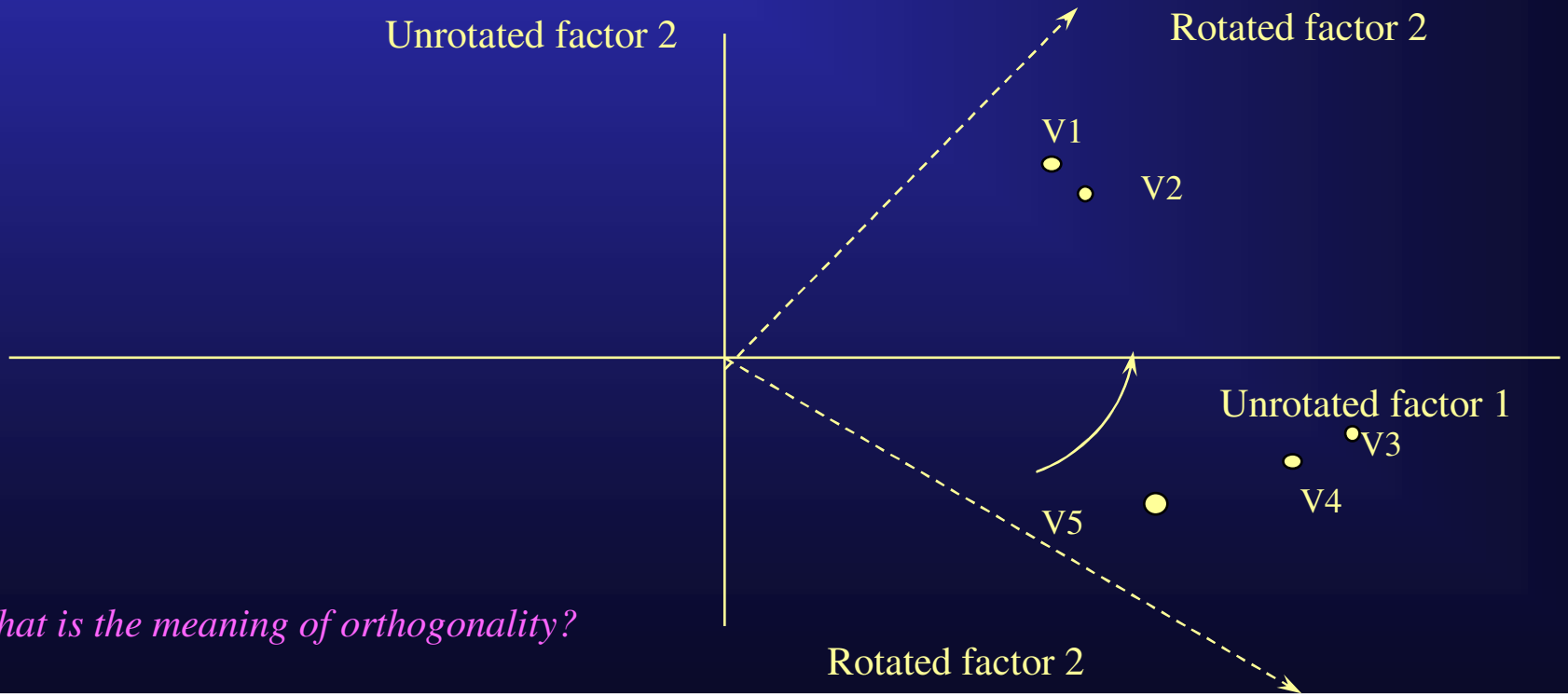
**An example for an "Ideal" rotation would be:**

| Variable | $F_1$ | $F_2$ |
|---|---|---|
| $X_1$ | 1 | 0 |
| $X_2$ | 1 | 0 |
| $X_3$ | 0 | 1 |
| $X_4$ | 0 | 1 |
| $X_5$ | 1 | 0 |

Normally, this is not the situation. The process of rotation does not alter either the number of the factors or the total variance accounted for all the factors. Even the spatial configuration of the variables defined by the preliminary solution is not altered. **Only the perspective is changed - we can easily assign variables to factors.**

# Comparison between rotated and unrotated factor loadings in an Orthogonal Rotation

| Variable | Unrotated Factor loadings | | Rotated Factor loadings | |
|---|---|---|---|---|
| | F1 | F2 | F3 | F4 |
| V1 | .50 | .80 | .03 | .94 |
| V2 | .60 | .70 | .16 | .90 |
| V3 | .90 | -.25 | .90 | .24 |
| V4 | .80 | -.30 | .84 | .15 |
| V5 | .60 | -.50 | .76 | -.13 |




*Q: what is the meaning of orthogonality?*

# The art of labeling

- A cut off loading of .5 is often used by researches as a rule of thumb to designate high loading to be used in a factor description.
- After grouping variables to factors the marketing “art” begins.
- The name should include what is as well as what is not involved in a factor.
- The variables should be reordered in terms of their loadings under each factor to facilitate labeling and interpretation.

*Examples and discussions are provided (26-33)*

# One of the F.A application: Creating perceptual maps . An example of a car

1) From an exploratory research derive the variables to analyze and the competing brands

Attributes

A1 - Appeals to other people

A2 - Attractive looking

A3 - Expensive looking

A4 - Exciting

A5 - Very reliable

A6 - Well engineered

A7 - Trend setting

A8 - Has all the latest features

A9 - Luxurious

A10 - Distinctive looking

A11 - Nameplate you can trust

A12 - Conservative looking

A13 - Family vehicle

A14 - Basic transportation

A15 - High Quality

Cars

1) Buick Century

2) Ford Taurus

3) Oldsmobile Cutlass Supreme

4) Ford Thunderbird

5) Chevrolet Celebrity

6) Honda Accord

7) Pontiac Grand AM

8) Chevrolet Corsica

9) Ford Tempo

10) Toyota Camry

# Perceptual maps for a car

2) Obtain Data on the variables (consumers ratings):

| | | | Attributes | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Respondent | Brand | Preference | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | A10 ….. |
| 1 | 1 | 4 | 5 | 7 | 7 | 8 | 5 | 6 | 8 | 5 | 6 | 8 |
| 1 | 2 | 8 | 4 | 6 | 5 | 6 | 5 | 6 | 4 | 5 | 5 | 5 |
| 1 | 3 | 7 | 3 | 6 | 7 | 8 | 5 | 5 | 8 | 7 | 8 | 6 |
| …… | | | | | | | | | | | | |
| 1 | 10 | 8 | 6 | 6 | 6 | 6 | 4 | 5 | 5 | 6 | 5 | 5 |
| 2 | 1 | 4 | 6 | 6 | 6 | 6 | 6 | 4 | 5 | 7 | 8 | 7 |
| 2 | 2 | …… | | | | | | | | | | |

……..

The data matrix:

Attribute car 1 car 2 car 3……

1

2

…..

# Use the computer in the following stages

3) Develop a correlation Matrix.

4) Find Factors and Factor loadings

5) Select the appropriate number of Factors, use the following heuristics: a) 60%-70% b) Communality more then 1 c) the scree plot.

6) Modify the initial factors analysis solution (use rotation).

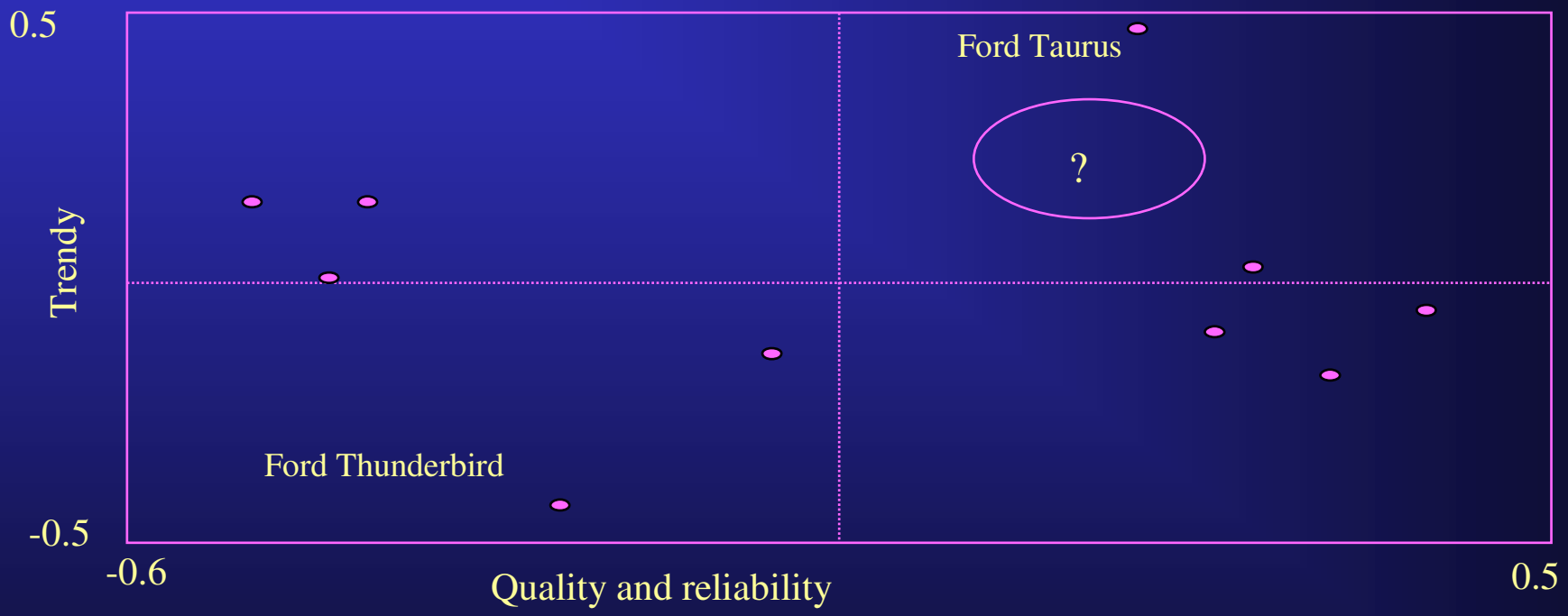
7) Label the Factors,

8) Obtain the Factors scores (for each of the brands):

| | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| 1) Buick Century | .265 | .017 | .157 |
| 2) Ford Taurus | .233 | .474 | .299 |
| 3) Oldsmobile Cutlass Supreme | .536 | -.037 | -.736 |
| 4) Ford Thunderbird | -.302 | -.456 | .513 |
| 5) Chevrolet Celebrity | -.393 | .194 | -.078 |
| 6) Honda Accord | .434 | -.139 | -.835 |
| 7) Pontiac Grand AM | -.584 | .197 | .196 |
| 8) Chevrolet Corsica | …….... | | |
| 9) Ford Tempo | | | |
| 10) Toyota Camry | | | |

# The perceptual maps

9) Use the factor scores to draw a perceptual map



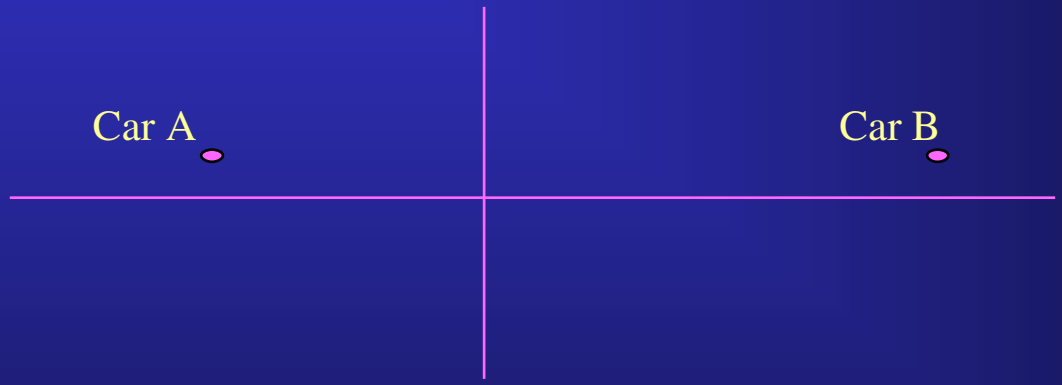

Q: what is missing here?

# The Ideal point(s)

10) Find the ideal points and plot them on the perceptual map.

**The basic intuition:**

A perceptual map with two cars

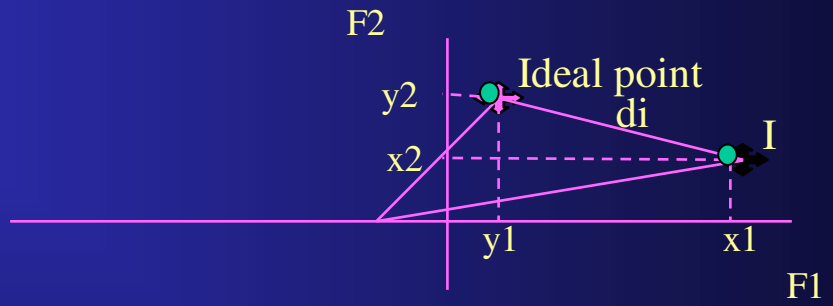



Q: Assuming that respondent 1 gave Car A the preference rank 10 and for car B 2. Where would you "place" Pi on the perceptual map?

# A Preference Regression method

$$P_i = a - bd_i^2 + e$$

Pi - preference for I, di is the distance of brand I from the Ideal point




For clarity assume that there are 2 factors:

$$d_i = \sqrt{(y_1 - x_{1i})^2 + (y_2 - x_{2i})^2}$$

And now for some algebra:

$$\begin{aligned}
&P_i = a - b\left[(y_1 - x_{1i})^2 + (y_2 - x_{2i})^2\right] + e = \\
&= a - b\left[y_1^2 + y_2^2 + x_{1i}^2 + x_{2i}^2 - 2y_1 x_{1i} - 2y_2 x_{2i}\right] + e = \\
&a - b\left(y_1^2 + y_2^2\right) - b\left(x_{1i}^2 + x_{2i}^2\right) + 2by_1 x_{1i} + 2by_2 x_{2i} + e = \\
&A_0 + B_0\left(x_{1i}^2 + x_{2i}^2\right) + B_1 x_{1i} + B_2 x_{2i} + e
\end{aligned}$$

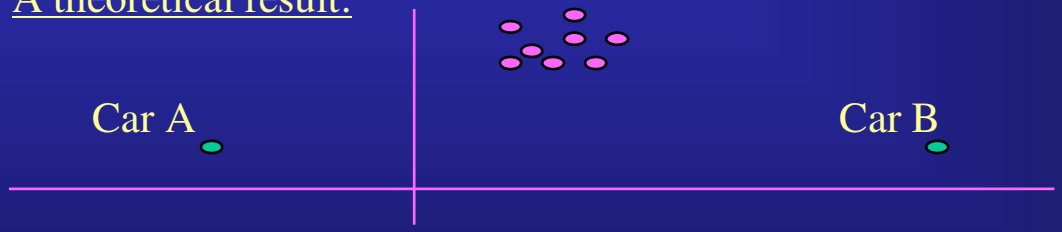
*This is a regression equation for the ideal point. $A_0$, $B_0$, $B_1$, $B_2$ can be obtained*

# Finding the Ideal Point for each customer

From the regression equation for each customer one can obtain the coordinates of the ideal point from the following set of equations:

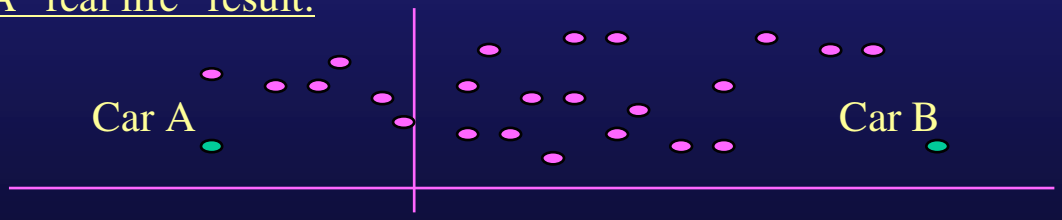
$$\begin{cases} B_0 = -b \\ B_1 = 2by_1 \\ B_2 = 2by_2 \end{cases} \Rightarrow \left( y_1 = -\frac{B_1}{2B_0}, y_2 = -\frac{B_2}{2B_0} \right)$$

A theoretical result:




**"Ideal point" potential**

A "real life" result:




*Q: How can we obtain a market ideal point in the "real life" situation?*

# Cluster Analysis

- C.A is a set of techniques which Classify, based on observed characteristics, an heterogeneous aggregate of people, objects or variables, into more homogeneous groups.
- C.A is useful to identify market segments, competitors in market structure analysis, matched cities in test market etc.

Q: Why do we need C.A when we have the Cross-Tabulation techniques?

# Steps involved in C.A

- Select a representative and adequately large sample of persons, products, or occasions.
- Select a representative set of attributes from a carefully specified field.
- Describe or measure each person, product, or occasion in terms of the attribute variables.
- Choose a suitable metric and convert the variables into compatible units.
- Select an appropriate index and assess the similarity between pairs of person, product or occasion profiles.
- Select and apply an appropriate clustering algorithm to the similarity matrix after choosing a cluster model.
- Compute the characteristic mean profiles of each cluster and interpret the findings.

# The basic intuition behind C.A

$$Minimize\left(\frac{Within\ \ cluster\ \ \text{var}\ iance}{Between\ \ cluster\ \ \text{var}\ iance}\right)$$