

# 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.

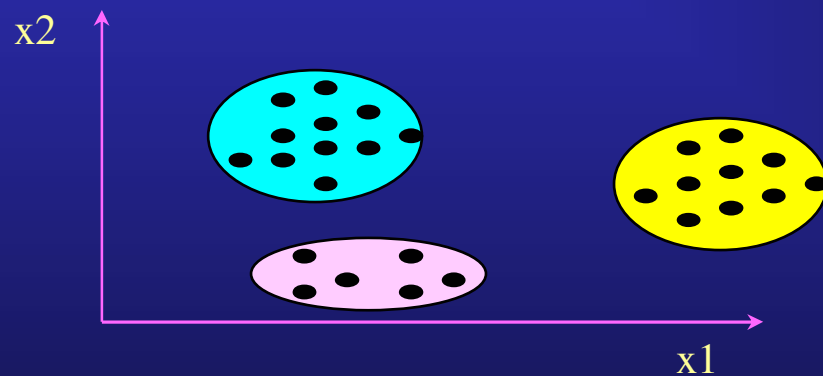
Q: Why do we need C.A when we have the Cross-Tabulation techniques?

## 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

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



# Segmentation variables

## Possible bases for segmentation:

- Dimensions that are outputs of Factor Analysis.
- Exploratory research.
- VALSE variables price sensitivities
- Heavy-light users
- Demographic variables
- Psychographic variables.

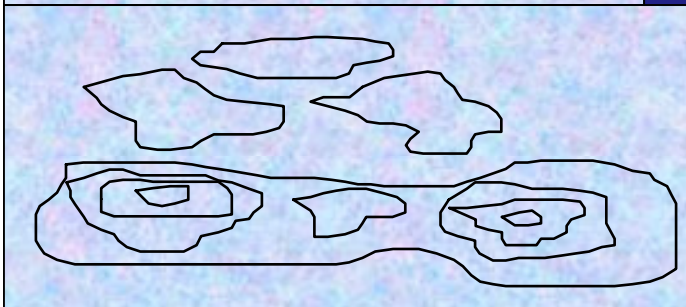
# Overview C.A

		Variables			
		X1	X2	X3	X4
Objects	O1				
	O2				
	O3				
	O4				
	O5				

1) n objects measured on p variables

		Objects			
		O1	O2	O3	O4
Objects	O1				
	O2				
	O3				
	O4				
	O5				

2) Transforming to nxn similarity (distance) matrix.



3) Cluster formation:

a) N.H.C.A (mutually exclusive clusters)

or

b) H.C.A (hierarchical clusters).

		Objects			
		O1	O2	O3	O4
Clusters	C1				
	C2				
	C3				

4) Cluster profiles

# Measures of similarity

General types of similarity measure are available:

- Distance measures
- Correlation measures
- Agreement or matching-type measures

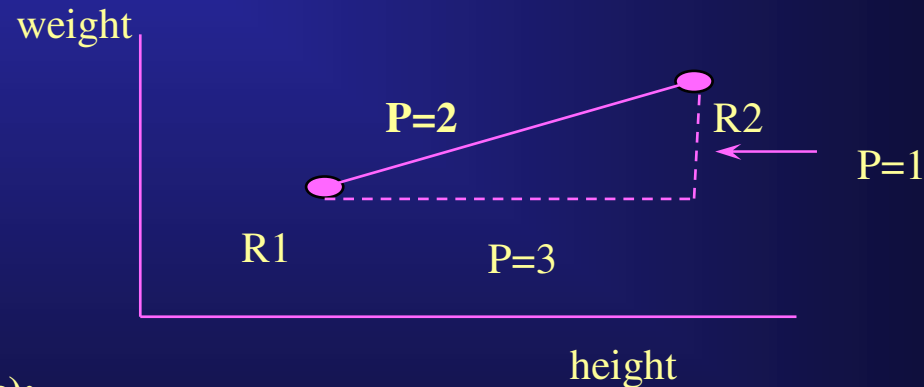
## Distance measure:

The most common is the Euclidian measure:  $d_{ij} = \sqrt{\sum_n (X_{ik_n} - X_{jk_n})^2}$

Where  $d_{ij}$  is the distance between objects  $i$  and  $j$ .  $X_{ik_n}$  represent the scores of objects  $i$  and  $j$  on variable  $k_n$ .

## Problems:

- 1) The variables may be measured by different units (suggest a solution)
- 2) The variables may be correlated (suggest a solution)



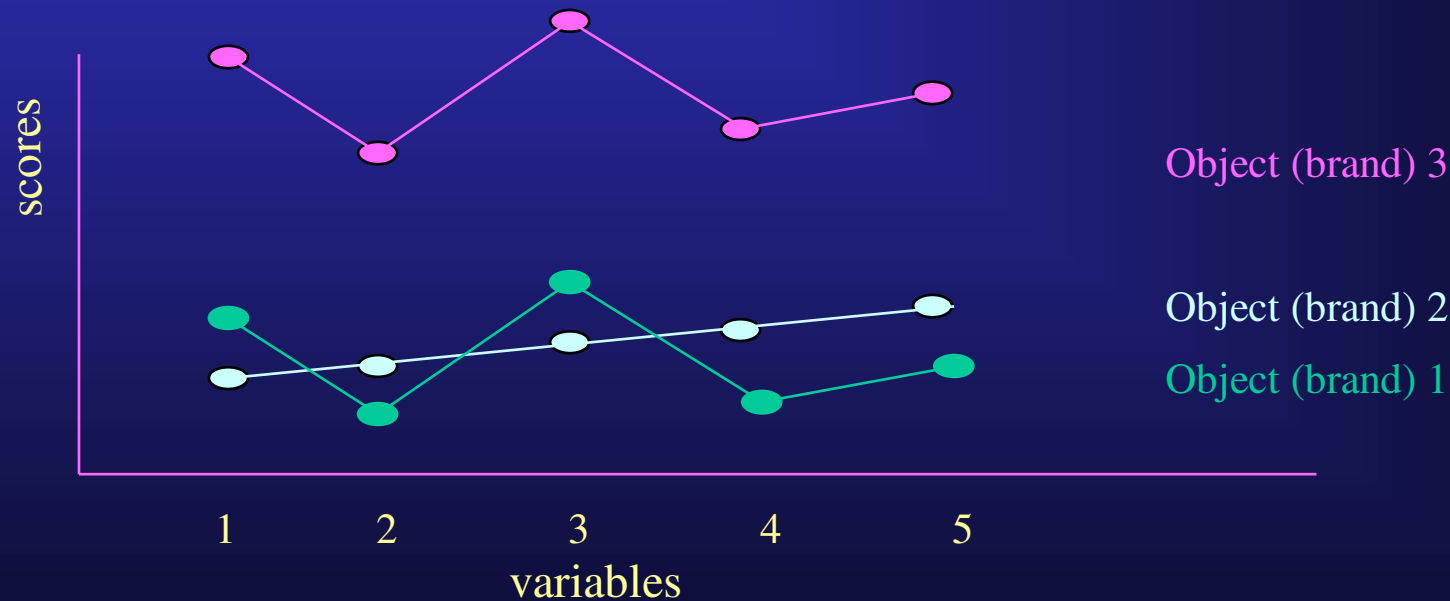
## The general form (the Minkowski metric):

$$d_{ij} = \left[ \sum_n (X_{ik_n} - X_{jk_n})^p \right]^{1/p}$$

When  $p=2$  this is the same as the Euclidian. When  $p=1$  - it is sometimes called “the taxicab” (why?)

# Correlation measures

The interpretation of a distance measure differs from that of a correlation measure. Consider an example below; profiles of 3 objects (brands) on 5 variables (attributes) are shown in the diagram. Using a distance measure brand 1 and 2 will be judged as most similar (closer numerical values on all 5 attributes). However using a correlation measure brand 1 and 3 would be judged as most similar because their responses are perfectly correlated although further apart on the scales. **Hence the two types of measures might yield different clusters when applied on the same data.**



Q1: when each of the measures is more appropriate?

## H.C.A Vs. N.H.C.A...

**Hierarchical clusters** are nested tree-like structures, and usually reflect a development sequence. Each person, product or occasion is treated as a separate and distinct cluster to begin with. They are merged using an appropriate similarity measure until every object belongs to a large cluster. It may help for “seeing the market structure” in terms of brands.

For a set of 100 persons the H.C.A will start with 100 clusters, each containing 1 object and finish with 1 cluster.

**Non-hierarchical methods** cluster a data set into a single classification of a number of clusters fewer than the number of objects. The number of the cluster may be specified a-priori or determined as part of the clustering method.

# Methods of clustering

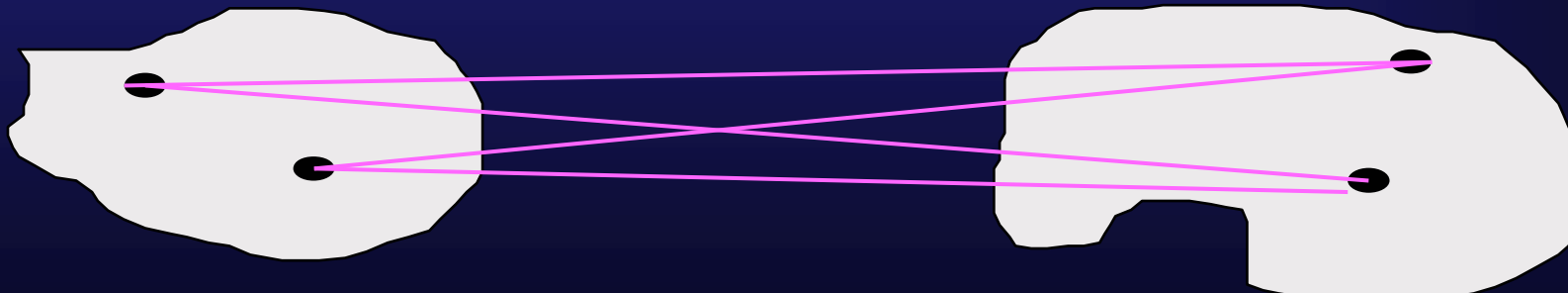
Minimum Distance (single linkage)



Maximum Distance (Complete linkage)

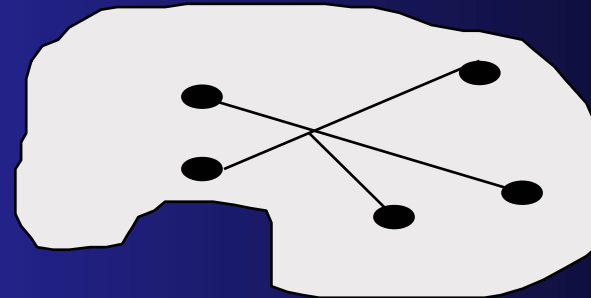
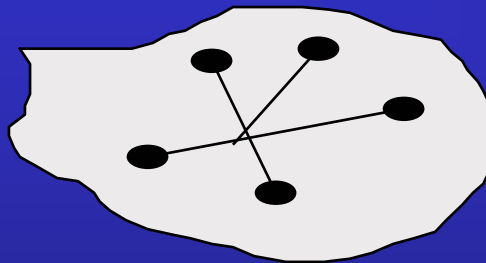


Average Distance (Average linkage) - the most common

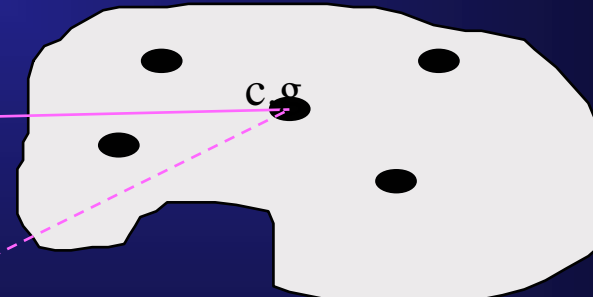
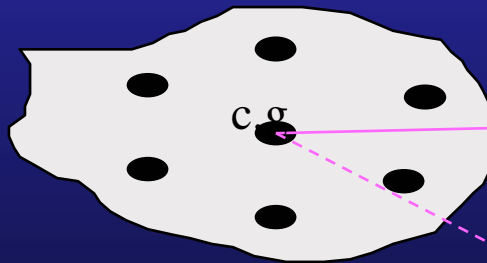


# Other agglomerative clustering methods

Ward's method

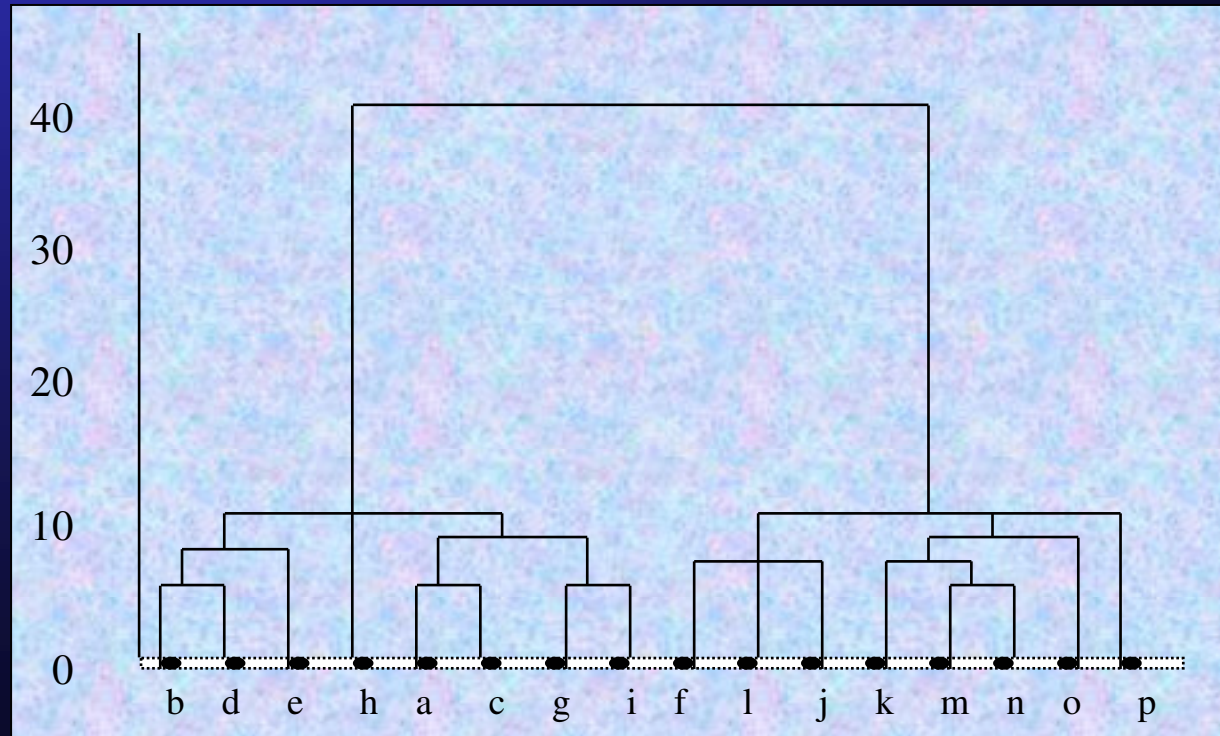
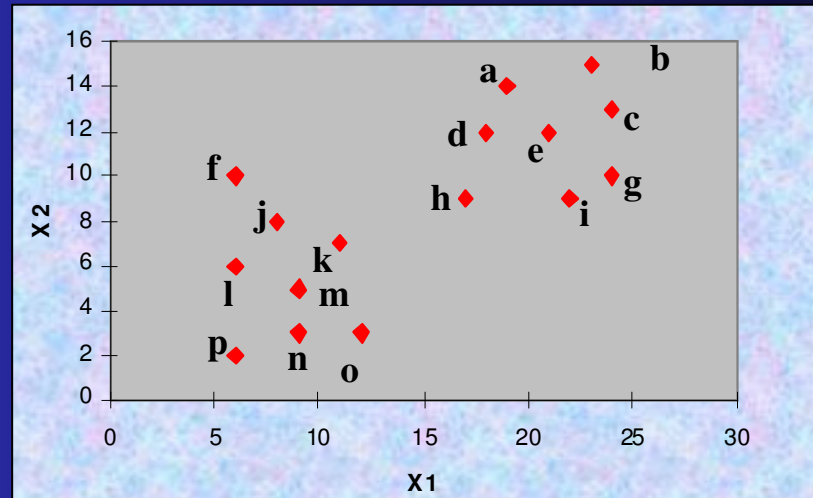


Centroid method



# Dendograms of H.C.A

point	X1	X2
a	23	15
b	19	14
c	24	13
d	18	12
e	21	12
f	6	10
g	24	10
h	17	9
l	22	9
j	8	8
k	11	7
l	6	6
m	9	5
n	9	3
o	12	3
p	6	2



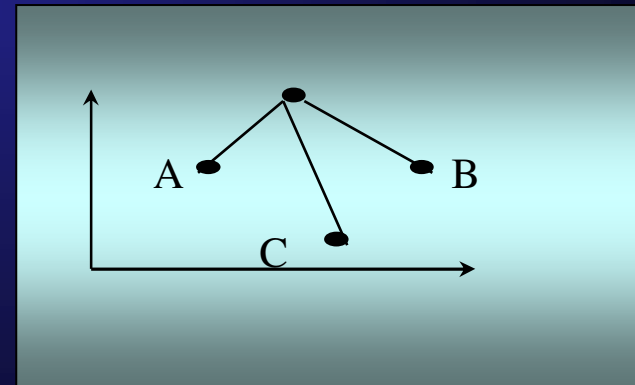
# N.H.C.A

- **K-Means Clustering (the most common).**

- **Methods based on trace**

- Object may be reallocated
- Iterative process of optimizing a certain criterion.
- Most common - the number of cluster has to be previously determined based on a-priori knowledge.

- Random pick: A, B, C
- Distance of point to all the cluster kernels
- assign the point to a cluster (A, B or C)
- Recompute the cluster kernel.
- Compute for all points and determine 3 cluster averaged centers.
- For the new (3) centers start all the computation again until convergence achieved.



# Number of clusters

- Input from H.C.A
- Run a large number of clusters to remove outliers
- As a rule of thumb, each cluster should have at least 50 consumers
- Can you interpret the clusters?

# A summarizing example - Clustering consumers based on attributes toward shopping.

Based on past research, six attitudinal variables were identified. Consumers were asked to express their degree of agreement with the following statements on a 7 point scale (1=disagree, 7 =agree):

- V1: Shopping is fun
- V2: Shopping is bad for your budget
- V3: I combine shopping with eating out.
- V4: I try to get the best buys while shopping.
- V5: I don't care about shopping
- V6: You can save a lot of money by comparing prices.

Case #	V1	V2	V3	V4	V5	V6
1	6	4	7	3	2	3
2	2	3	1	4	5	4
3	7	2	6	4	1	3
4	4	6	4	5	3	6
5	1	3	2	2	6	4
6	6	4	6	3	3	4
7	5	3	6	3	3	4
8	7	3	7	4	1	4
9	2	4	3	3	6	3
10	3	5	3	6	4	6
11	1	3	2	3	5	3
12	5	4	5	4	2	4
13	2	2	1	5	4	4
14	4	6	4	6	4	7
15	6	5	4	2	1	4
16	3	5	4	6	4	7
17	4	4	7	2	2	5
18	3	7	2	6	4	3
19	4	6	3	7	2	7
20	2	3	2	2	7	2

# C.A - A recommended approach

**Hierarchical Cluster Analysis**



**Decide how many clusters**



**Non Hierarchical Cluster Analysis**



**Validate cluster solution**



**Interpret findings**

# Positioning

- A product position is its unique imprint in the mind of the respondent. It applies to concepts, products or companies. A positioning may be changed through appropriate repositioning strategies.

- Objectives:

- to “see” our brand against the determinant attributes.
- to “see” competing brands against the determinant attributes
- to “see” all brands against buyer ideal points.

- Important decisions

- what brands should be positioned?
- what categories are involved (substitutable)?
- what are the appropriate attributes?

Q1: Give example of good repositions in Israel.

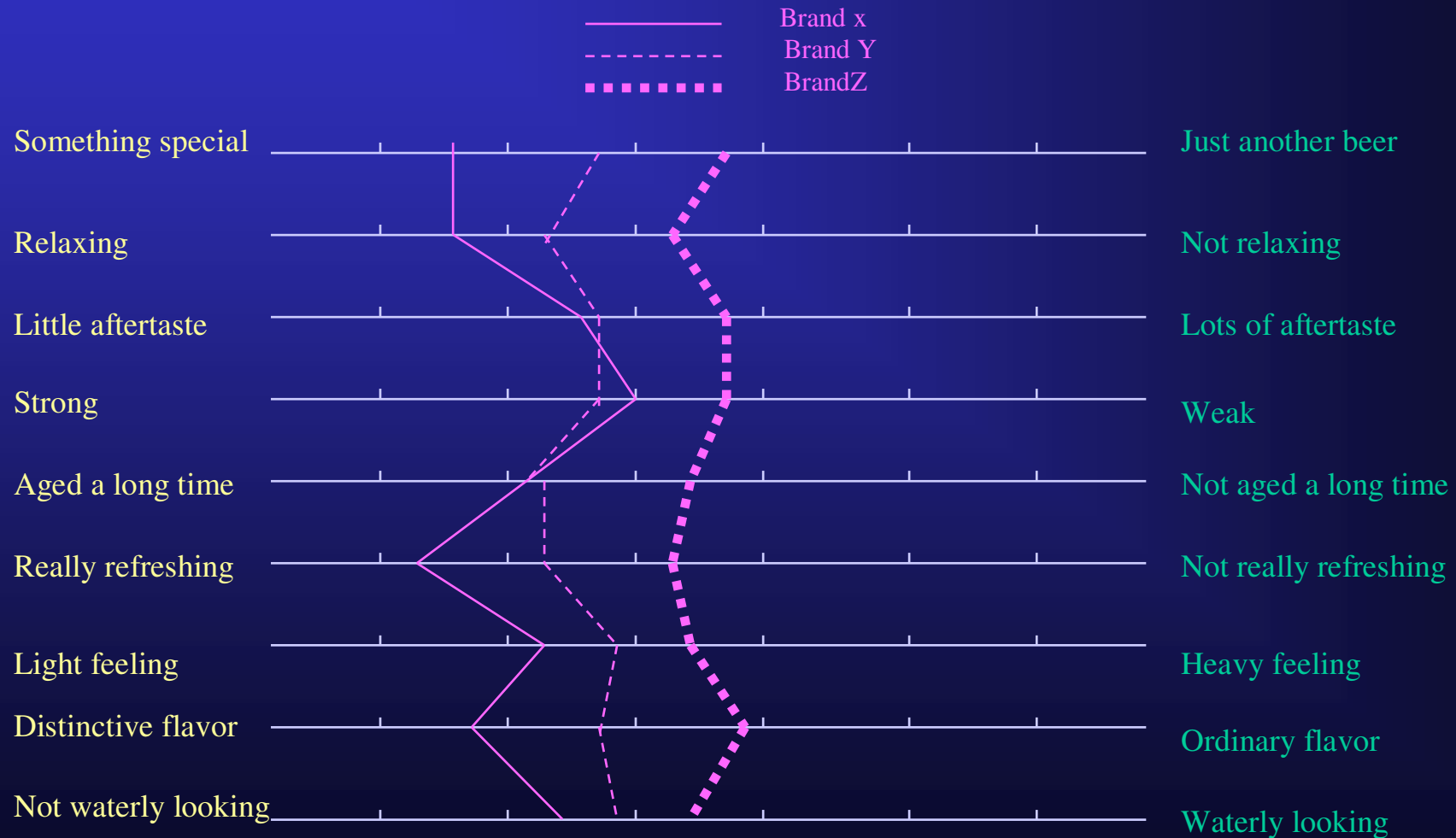
## Steps in positioning research

- Identify the relevant set of **competitive** products and brands which satisfy the **same** customer need
- Obtain demographic and other descriptive information to ascertain perceptual differences by segments.
- Analyze the data and present the results using simple representations such as: semantic differential plots, quadrant maps, importance/performance profiles, or use perceptual mapping techniques such as: specialized multidimensional scaling procedures, discriminant analysis, factor analysis, correspondence analysis.

# Profile analysis

## Profile analysis of a beer brand images

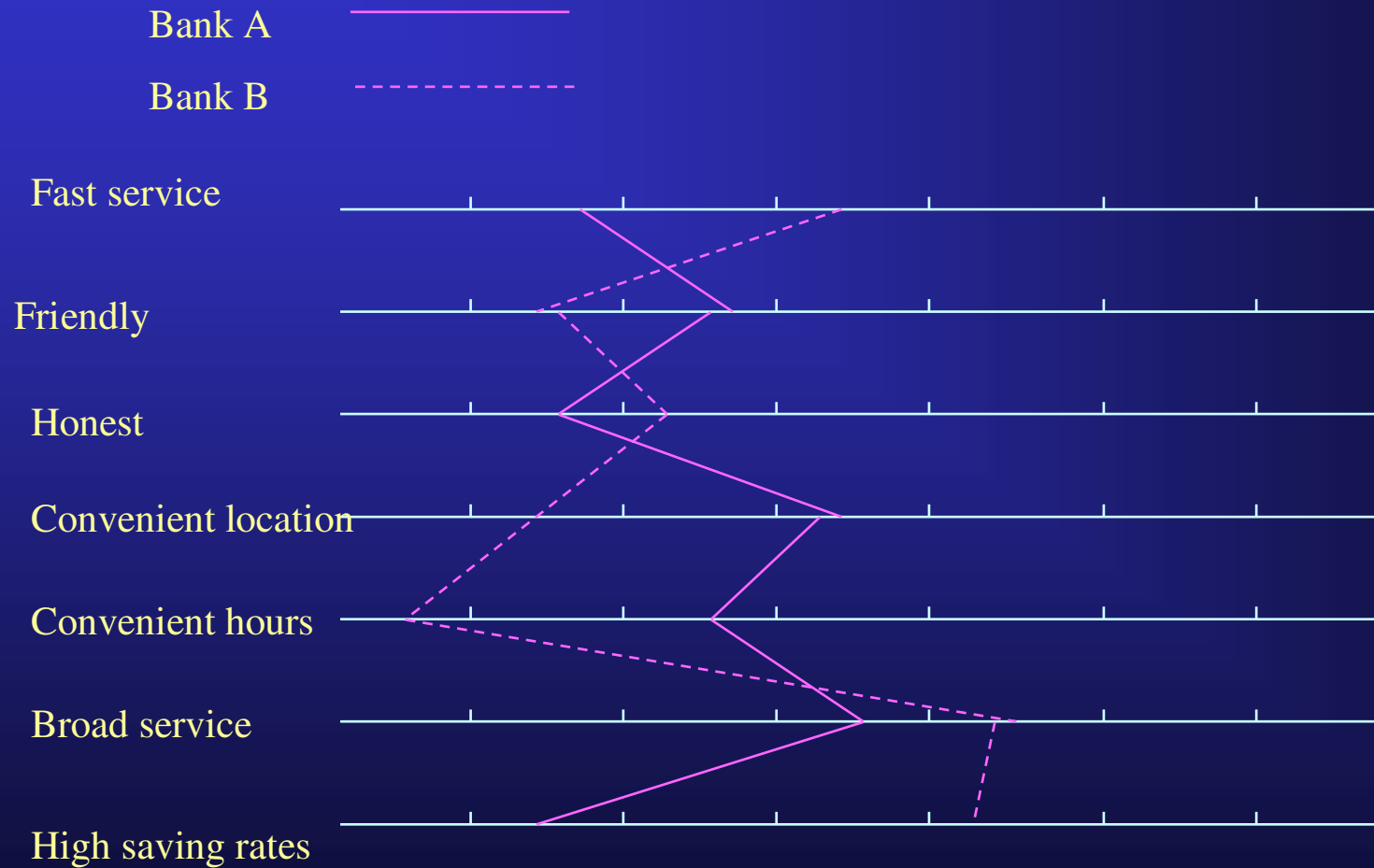
(source: William A. Mindak, "Fitting the Semantic Differential of the Marketing Problem", JM April 1962 p. 28-33)



## Profile analysis - Questions

- Describe the differences between the competing brands
- What can you learn from the analysis?
- Briefly describe possible marketing offers for each brand
- How would you acquire the information needed for the “snake plots”?

# Profile analysis - example2

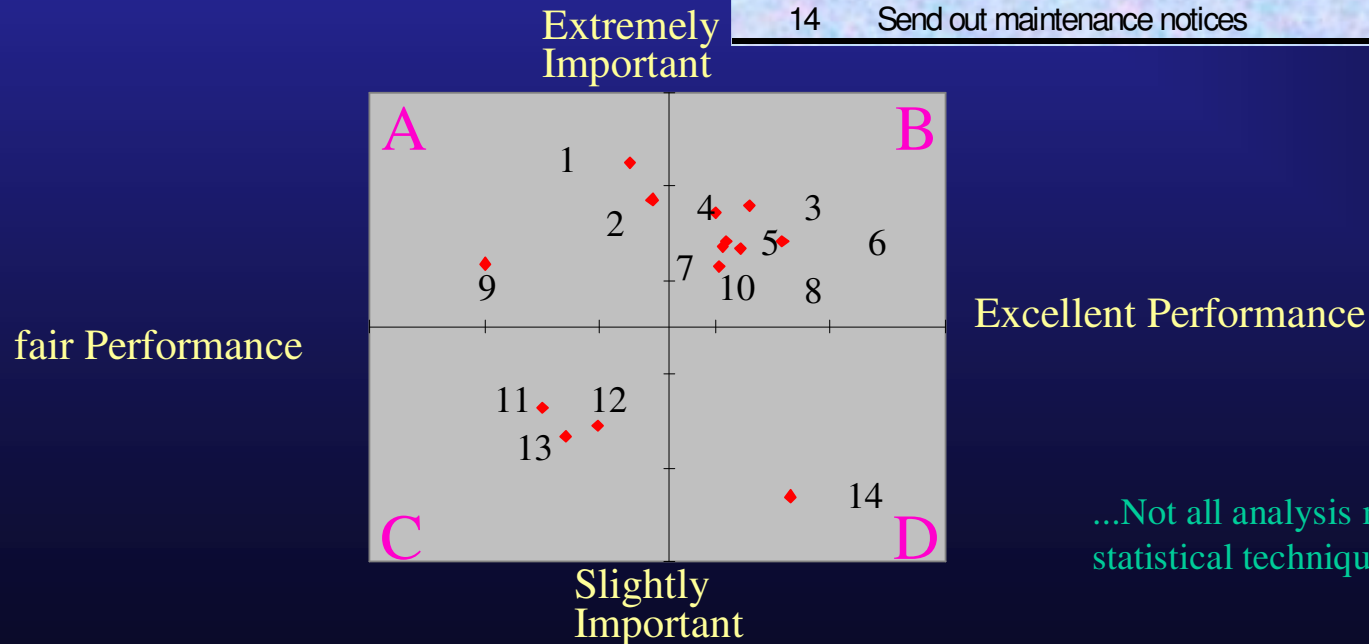


# Importance-Performance analysis

(adapted from JM, 41 J.A. Martilla and J.C James "Importance-Performance analysis[January 1977 p. 77-9])

An automobile dealer that less of 40% of its new car buyers remained loyal service customer after 6000 miles service.

Attribute	Attribute Description	Mean Importance	Mean Performance
1	Job done right the first time	3.83	2.63
2	Fast action on complaints	3.63	2.73
3	Prompt Warranty work	3.6	3.15
4	Able to do any job needed	3.56	3
5	Service available when needed	3.41	3.05
6	Courteous and friendly service	3.41	3.29
7	Car ready when promised	3.38	3.03
8	Perform only necessary work	3.37	3.11
9	Low prices on service	3.29	2
10	Clean up after service work	3.27	3.02
11	Convenient to home	2.52	2.25
12	Convenient to work	2.43	2.49
13	Courtesy buses and rental cars	2.37	2.35
14	Send out maintenance notices	2.05	3.33



...Not all analysis must involve sophisticated statistical techniques.

# Importance-Performance analysis - Interpretation

- **A - “Concentrate here”**

Customers feel that low service prices (attribute 1) are very important but indicate low satisfaction with the dealer performance.

- **B - Keep with the good work**

Customer value courteous and friendly service (attribute 6) and are pleased with the dealer's performance.

- **C - Low priority**

- The dealer is rated low in terms of providing courtesy buses and rental cars (attribute 13), but customers do not perceive this feature to be very important.

- **D - possible overkill**

The dealer is judged to be doing a good job of sending out maintenance notices (attribute 14) but customers attach only slight importance to them. (However there may be other good reasons for continuing this practice.)

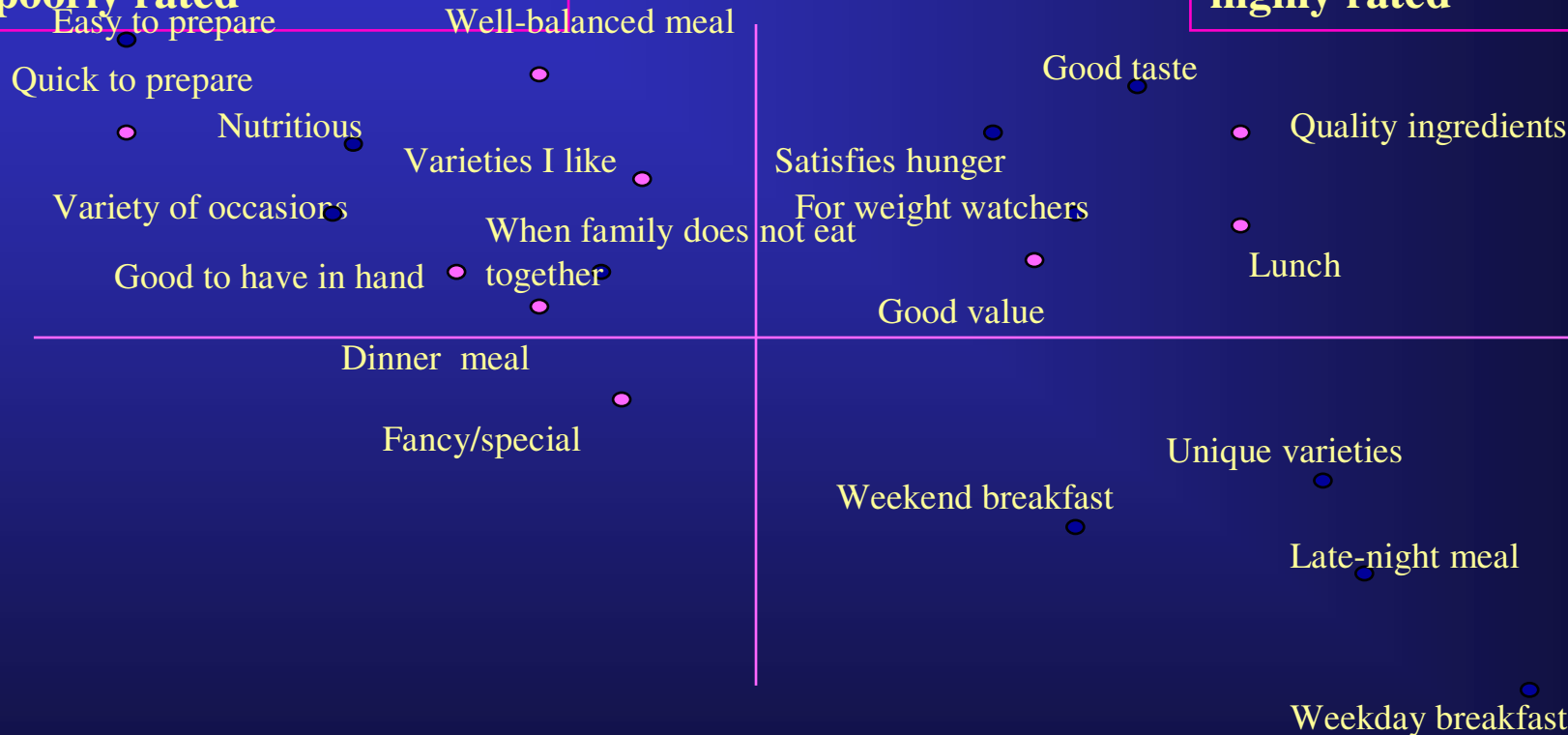
## **Advantages**

This is a relatively low cost technique and easily understood by information users, it can provide management with a useful focus for developing marketing strategies.

# Importance-Performance analysis, Example 2

**Highly important and poorly rated**

**Highly important and highly rated**



**Not important and poorly rated**

**Not important and highly rated**

Q: What is the product class? describe the brand's perceptions.

# Multidimensional Scaling (MDS)

A set of techniques to transform (dis)similarities and preferences among objects into distances by placing them in a multi-dimensional space.

It creates a spatial representation of (dis)similarity data.

It allows embedding ideal-points and property-vectors in the spatial representation, and estimating weights for individual differences.

## **What is it used for?**

- To uncover “hidden structure” in the data: Perceptual dimensions, competitors, clusters, and attributes.
- To identify and measure extent of competition/market structure.
- To facilitate modeling of choice.
- To evaluate and position concepts, stores, sale-force etc.
- To facilitate product planning and testing.
- To summarize test, and track advertising and image research.
- To track structural shifts in customer perceptions and preferences over time.

# Key decisions in MDS

- Marketing variables: Product/brands, individual/segments of consumers, attribute/occasions.
- What are the relations that should be analyzed?
- How to assess the proximity's to scale?
- Which analysis procedure (algorithm) to use?
- How many dimensions to retain?
- What method to use for visually representing the data?
- How to interpret the configuration?

# Similarity and distance

1) a is identical to b or it has some degree of similarity to it.

$$d(a, b) \geq 0$$

2) a is the most similar to a.

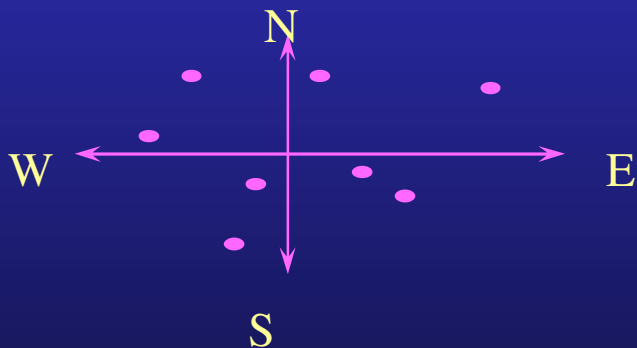
$$d(a, a) = 0$$

3) a is similar to b as b is similar to a.

$$d(a, b) = d(b, a)$$

## Representation of cities relations.

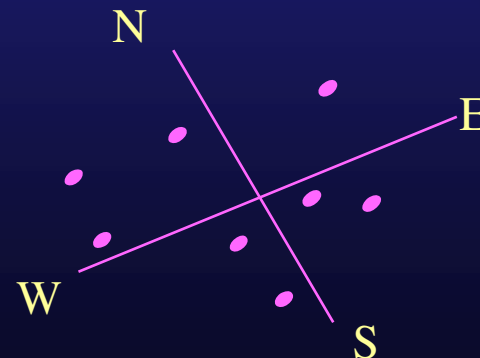
Geographic locations of cities



Airline distances

	a	b	c	d	e
a	X				
b		X			
c			X		
d				X	
e					X

MD-scale space for distances between cities.



# From similarity rankings to a map

	K-MART	PENNEYS	SEARS	WALMART	WARDS	WOOLWORTH
K-MART		12	11	1	7	3
PENNEYS			5	15	4	10
SEARS				13	6	14
WALMART					9	2
WARDS						8
WOOLWORTH						

*Dimension 1 - ??*



*Dimension 2 - ??*

# How it is done?

**The problem:** Given  $n(n-1)/2$  pairs on  $n$  objects with a measure of similarity between them we want to find a representation of the  $n$  points in a space of the smallest possible dimensionality such that the given proximity measure are monotonically related to the distances between the points in the spatial representation.

**The method:** An iterative process designed to adjust the positions of  $n$  points in an initial and perhaps arbitrary configuration until an explicitly defined measure of departure from the desired condition of monotonicity is minimized.

**Determination of the proper number of dimensions:** Most of the methods are designed to find the optimum configuration in a space of a prespecified number of dimensions. If the researcher does not know in advance the proper number a trial and error procedure is needed in which several configuration (with different dimensionality) are generated and the optimum one is chosen. note that large dimensionality offers a better fit while the low dimensionality solutions offer better parsimony, visualizability and stability.

# The thinking stage

## Interpretation of the resulting representation

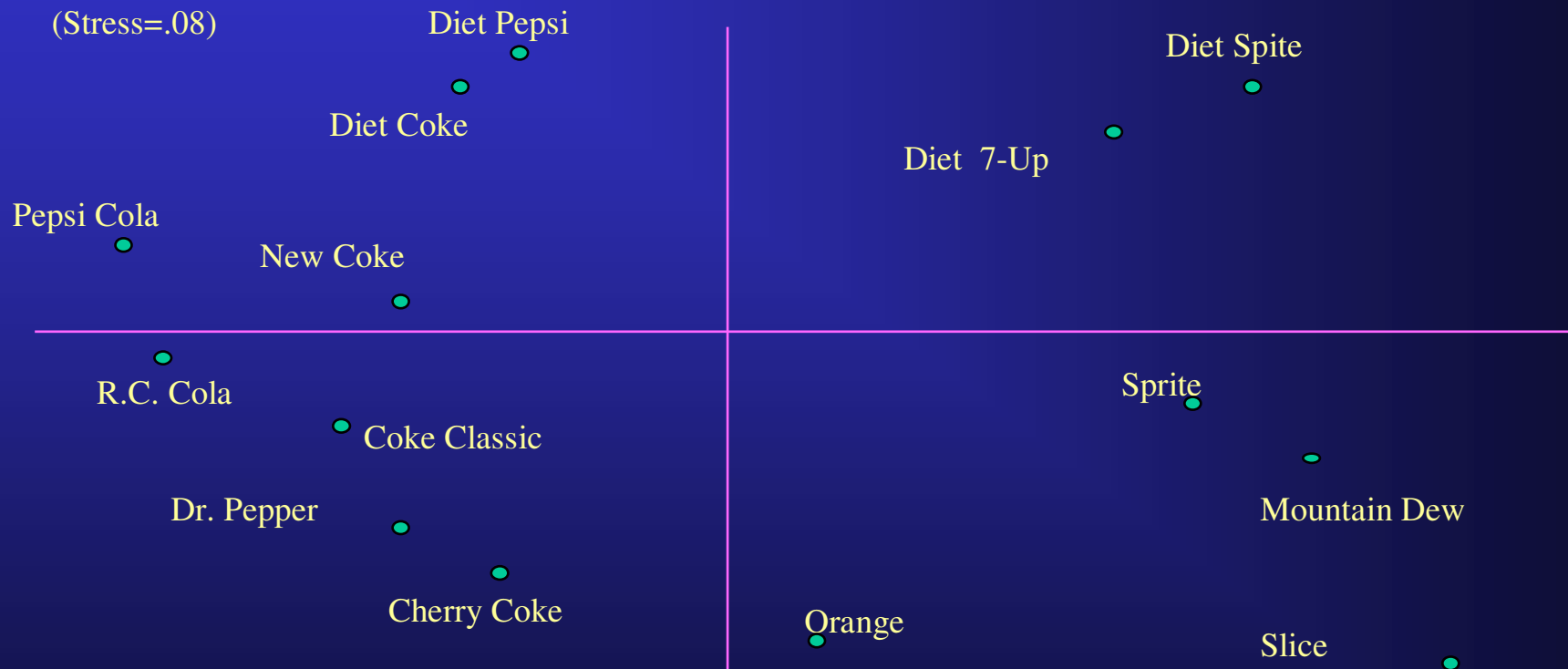
The central purpose of MDS is to find a spatial configuration that represents the structure originally hidden in the given matrix of proximity data, in a more accessible form to the human eye. One should therefore search for substantively significant interpretations for salient features of the resulting spatial representation as follows:

**Axes or directions:** since the orientation of the axes is entirely arbitrary, one should look for rotated or even oblique axes that may be readily interpretable.

**Cluster:** Whether or not there is a compelling interpretation for any axes there may be a set or hierarchical system of clusters that is readily interpretable.

**Other features:** kinds of orderly patterns (such as arrangement of points around the perimeter of a circle), Imagination and open mind are required...

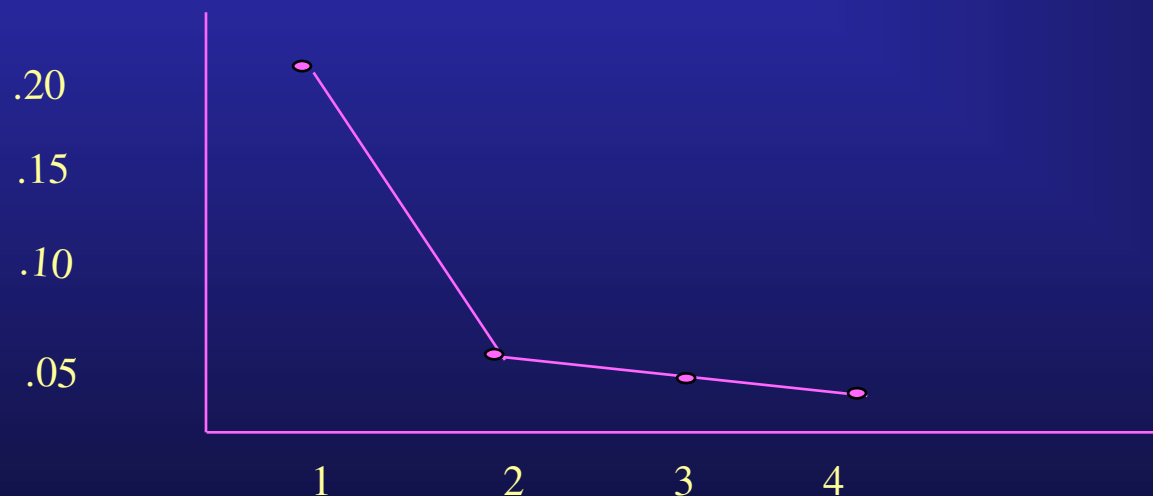
# Perceptual Map of 15 soda beverages



Q: Find interpretations for the dimensions

# How many dimensions to retain?

- Generally the lowest dimensionality is desired. However, oversimplification can be very misleading. The best approach is to select the fewest number of dimensions that faithfully reproduces the structure in the data.
- The quantitative measure is the “stress”. Low stress and elbows in a plot of stress Vs. # of dimensions (see below) indicate for a good fit and a structure in the data.

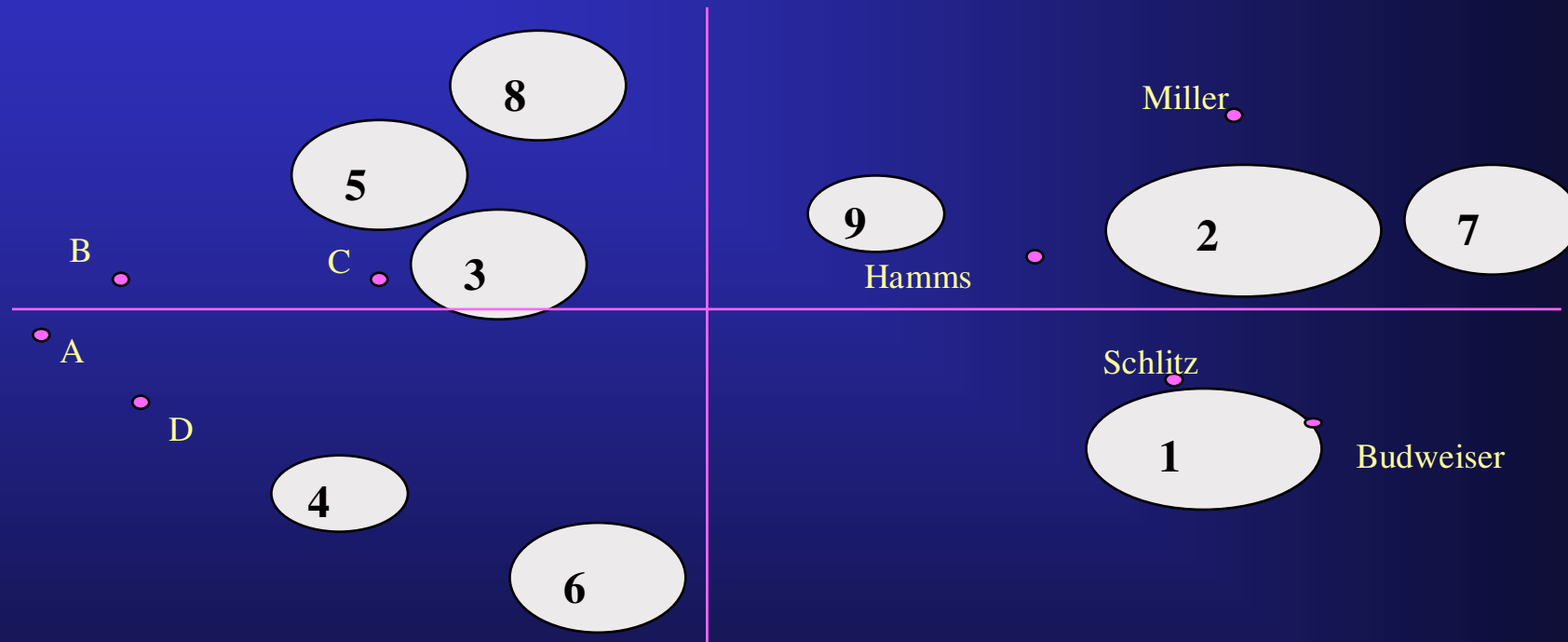


*In this example two dimensions can be selected.*

# Ideal Point(s)

## Distribution of ideal points in product space.

Source (Richard M. Johnson, "Market Segmentation - A Strategic Management Tool". JMR, 9 (February 1971), 16.



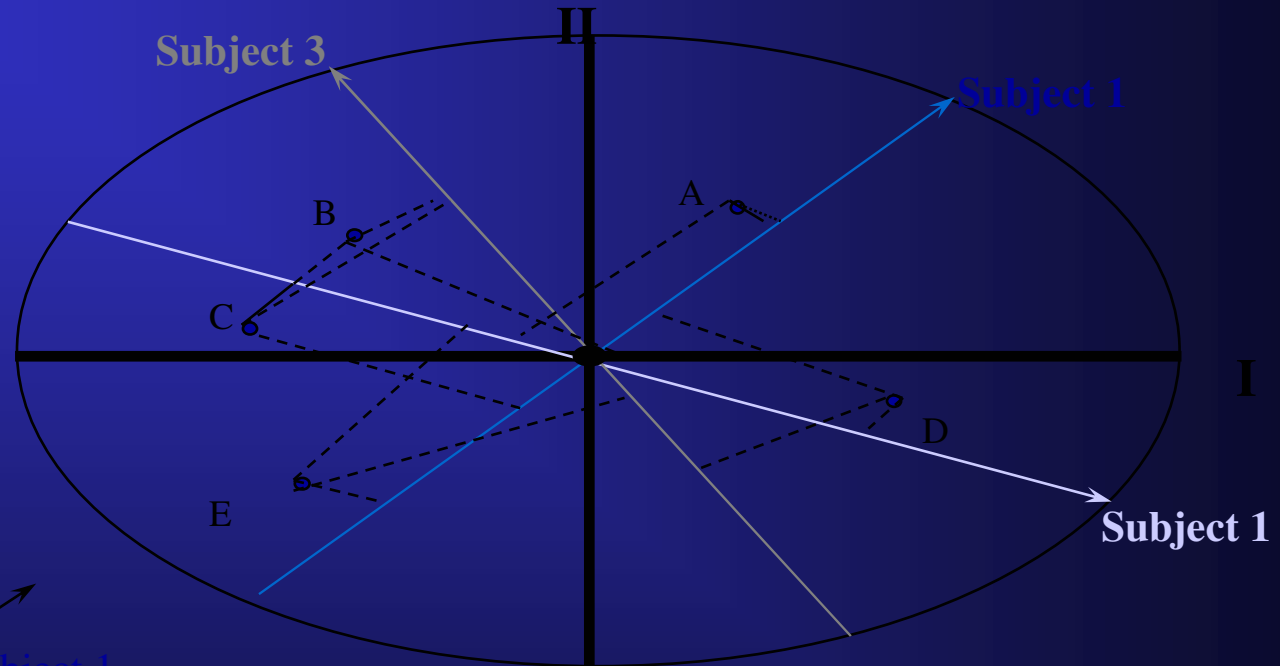
Q: What can be learned from the above map? what may be its shortcomings when dealing with a "new to the world" product

# Illustrative Vector Model and Isopreference Curves

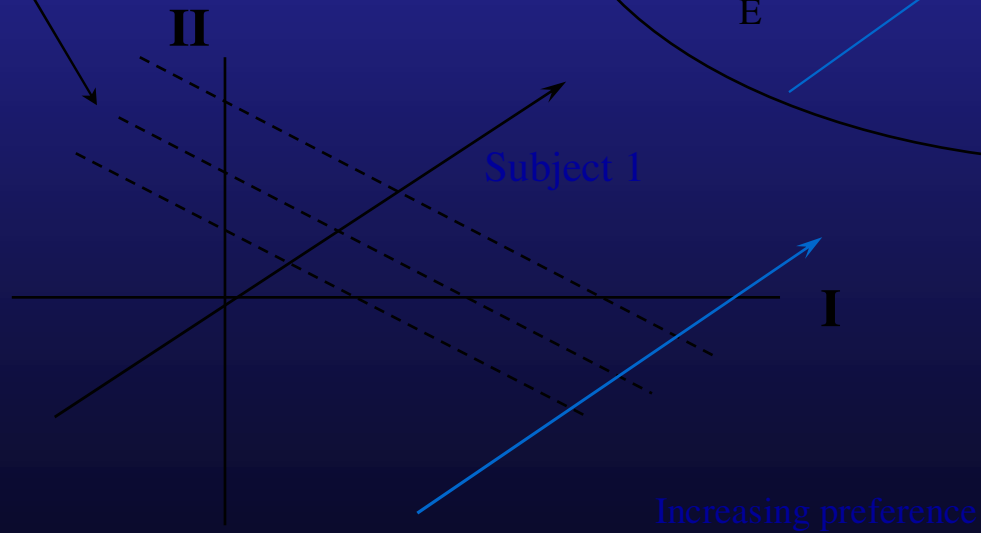
Subject 1:  $A > D > B > C > E$

Subject 2:  $D > A > E > C > B$

Subject 3:  $B > A > C > E > D$



Isopreference lines



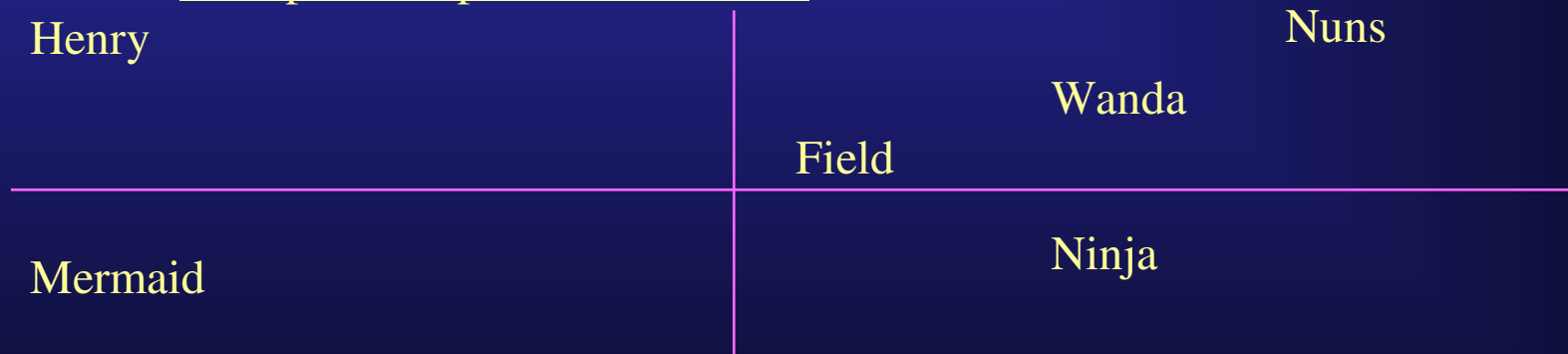
# Mapping the Movie Market: An OS Example

Respondents' ranking of similarity of six movies

*(Henry the V, Fish called Wanda, Nuns on the run, The little mermaid, Field of dreams and Ninja turtles)*

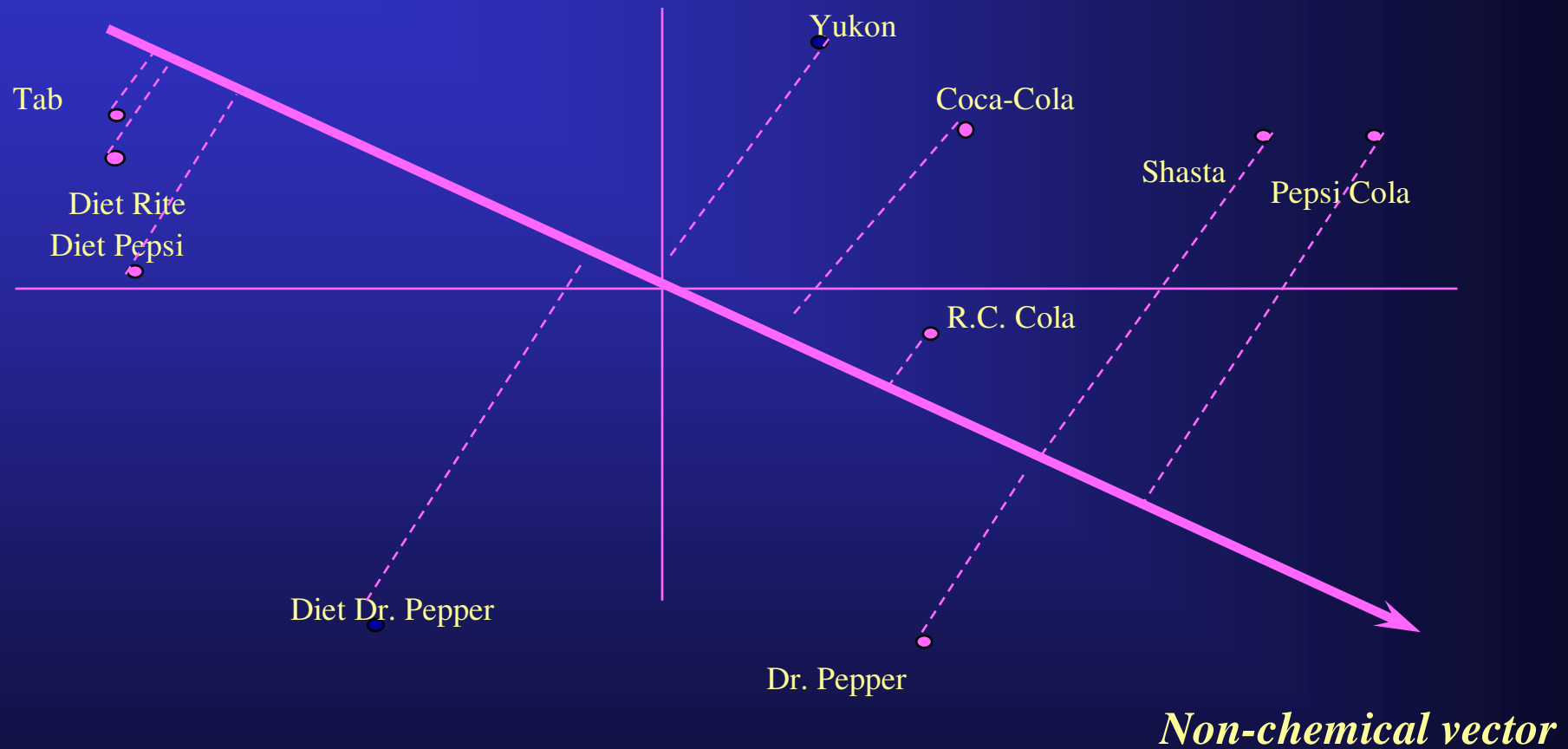
	Wanda	Nuns	Mermaid	Field	Ninja
Henry V	11	12	10	6	13
Wanda	-	1	14	2	5
Nuns	-	-	15	3	6
Mermaid	-	-	-	8	9
Field	-	-	-	-	4
Ninja	-	-	-	-	-

Perceptual Map of movie market



Q: Can you “name” the axes?

# Example - the “non chemical vector” in

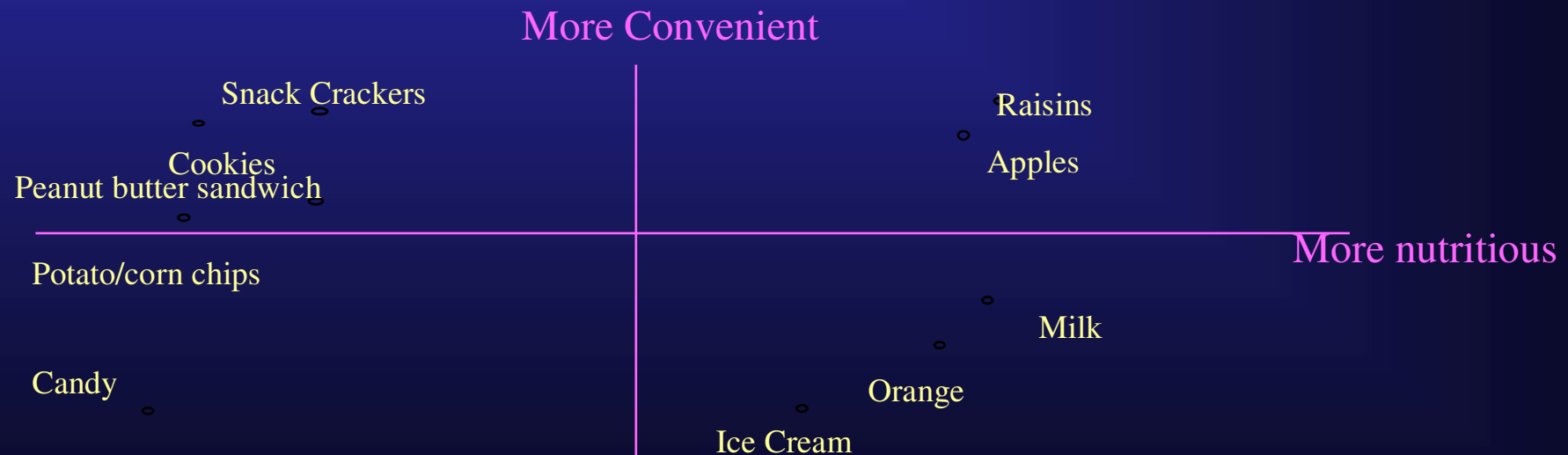


# Positioning map by using Factor Analysis

Factor loadings and importance rate for snack foods

	Characteristics	Factor I	Factor II	Mean Importance Rating
1	Filling/not filling	0.317	0.073	2.9
2	Fattening/not fattening	0.424	-0.009	2.64
3	Juicy/dry	0.301	0.125	3.28
4	Bad/good for complexion	0.645	0.104	2.19
5	Messy/not messy to eat	0.204	0.664	2.67
6	Expensive/inexpensive	0.244	0.347	2.43
7	Good/bad for teeth	0.762	0.056	1.53
8	Oily/not oily	0.516	0.24	2.65
9	Gives/doesn't give energy	0.541	0.165	2.221
10	Easy/hard to eat out of hand	-0.069	0.796	2.83
11	Nourishing/not nourishing	0.565	0.116	1.7
12	Stains/does not stain clothing, furniture	0.25	0.664	2.55
13	Easy/hard to serve	0.046	0.747	2.87
14	My children like it/dislike it	0.071	0.243	2.86

scale: 1=extremely important, 2=very important, 3=fairly important, 4= of little importance

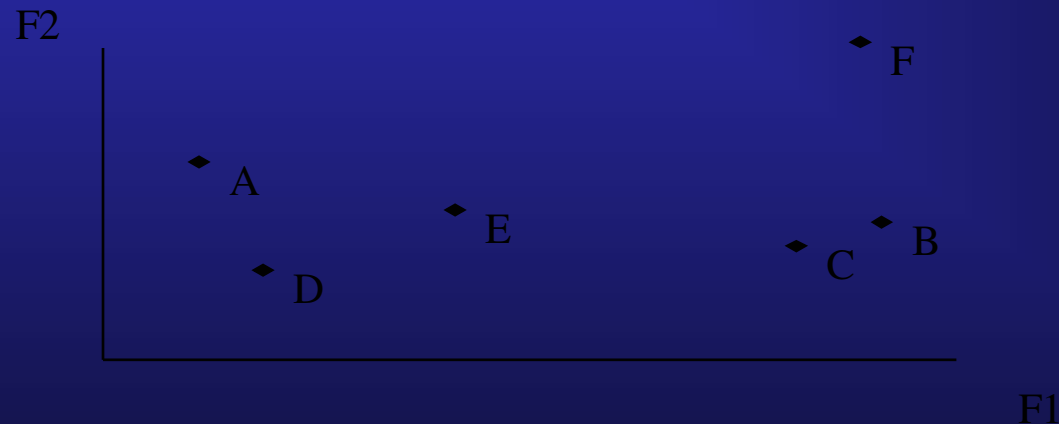


## Discriminant mapping

Consider six banks evaluated on 13 attributes on a 10 points scale (assume metric):

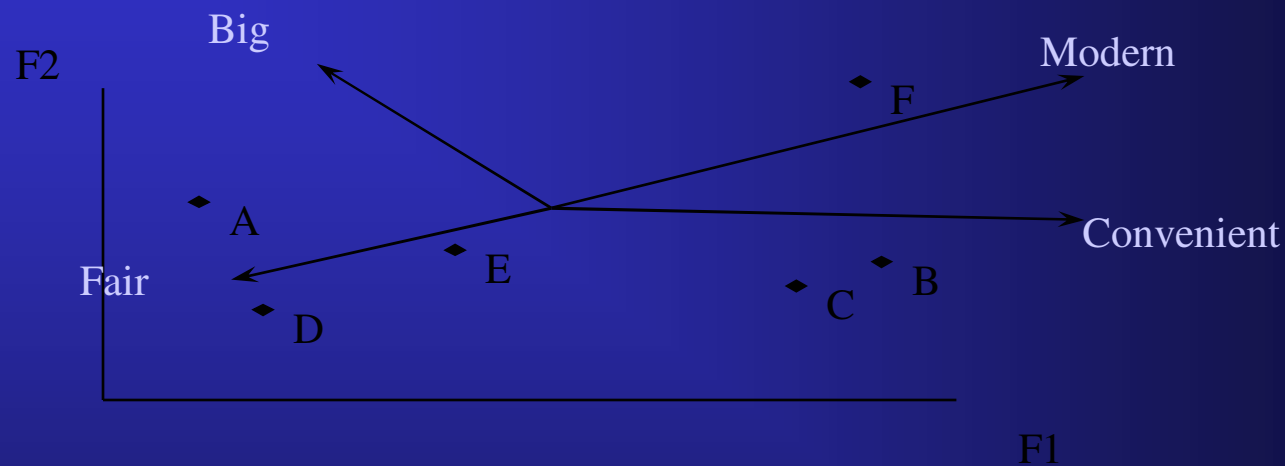
Convenient hours, progressive, handles accounts accurately, convenient locations, personal interest in customers, big, fair, active in local affairs, fast service, friendly, well managed, modern, courteous employees.

Multiple discriminant analysis was performed and the group centroids on the first two functions are illustrated below:



This does not reveal how the banks differ in terms of the original attributes (although this could be partially inferred by examining the standardized coefficients of the canonical functions). It is possible to insert attribute vectors on this map such that the projections of the group means reflects the relative ratings of the attribute for that group. The length of the vector can represent the ability to discriminate among the groups.

## Discriminant mapping

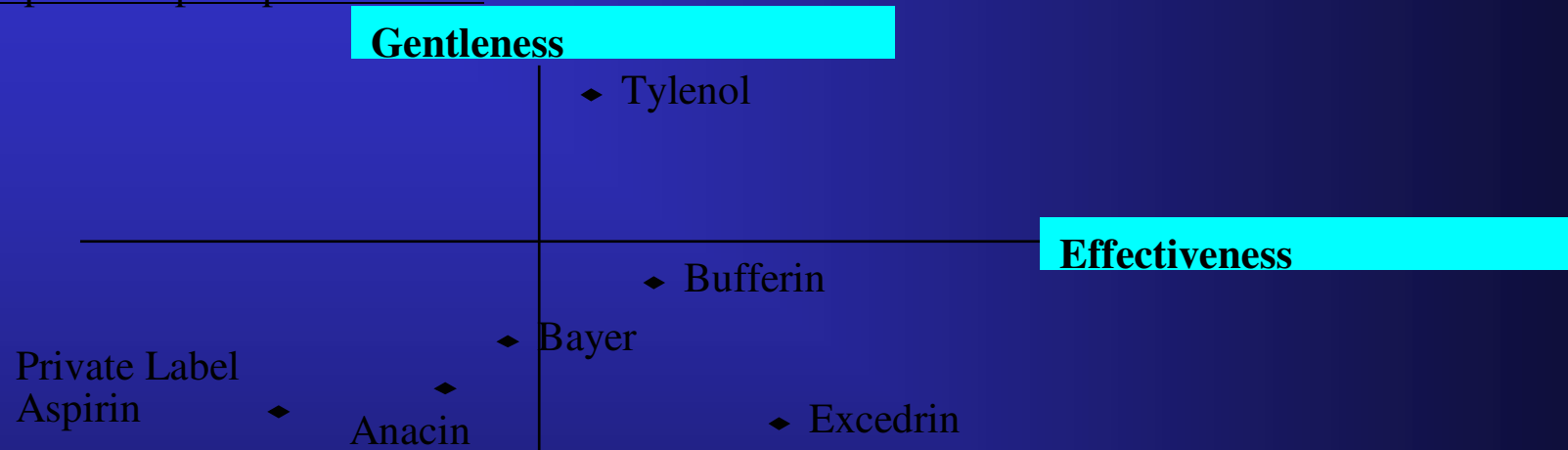


### How to do this:

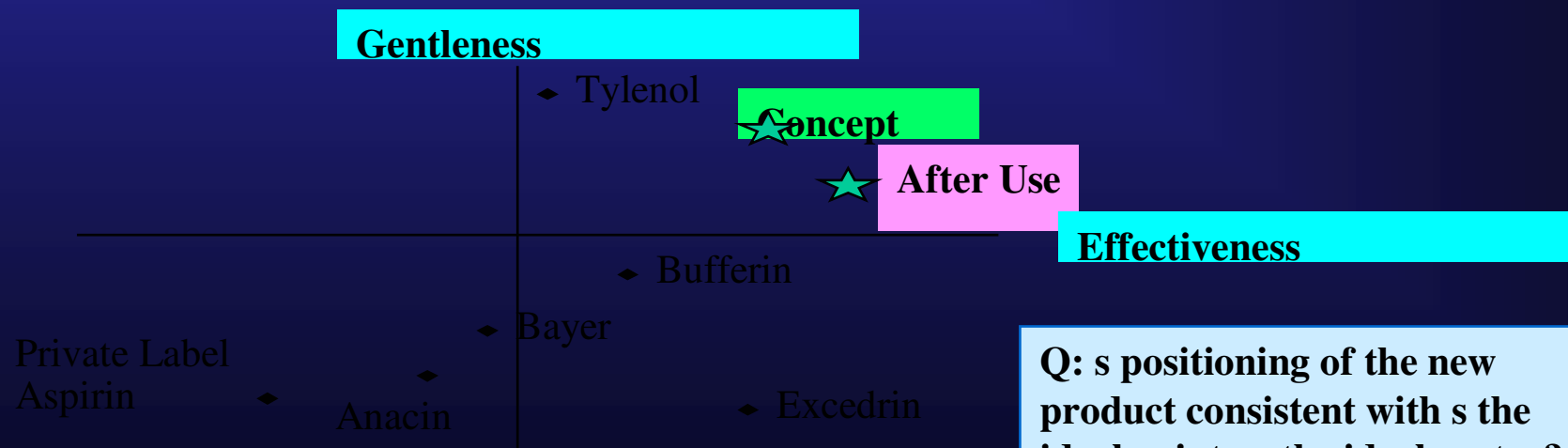
- Obtain the correlation between the original attribute scores and the discriminant scores on each discriminant function.
- Use as the origin the mean for all groups on both discriminant functions.
- Multiply the correlation by the F ratio for the particular attribute. The larger the F ratio the more discriminating that attribute so it will appear as a longer vector on the map. The vector's relative position is determined by the correlation with each axis (discriminant function).

# Benefit Segmentation

Perceptual Map of pain reliever.

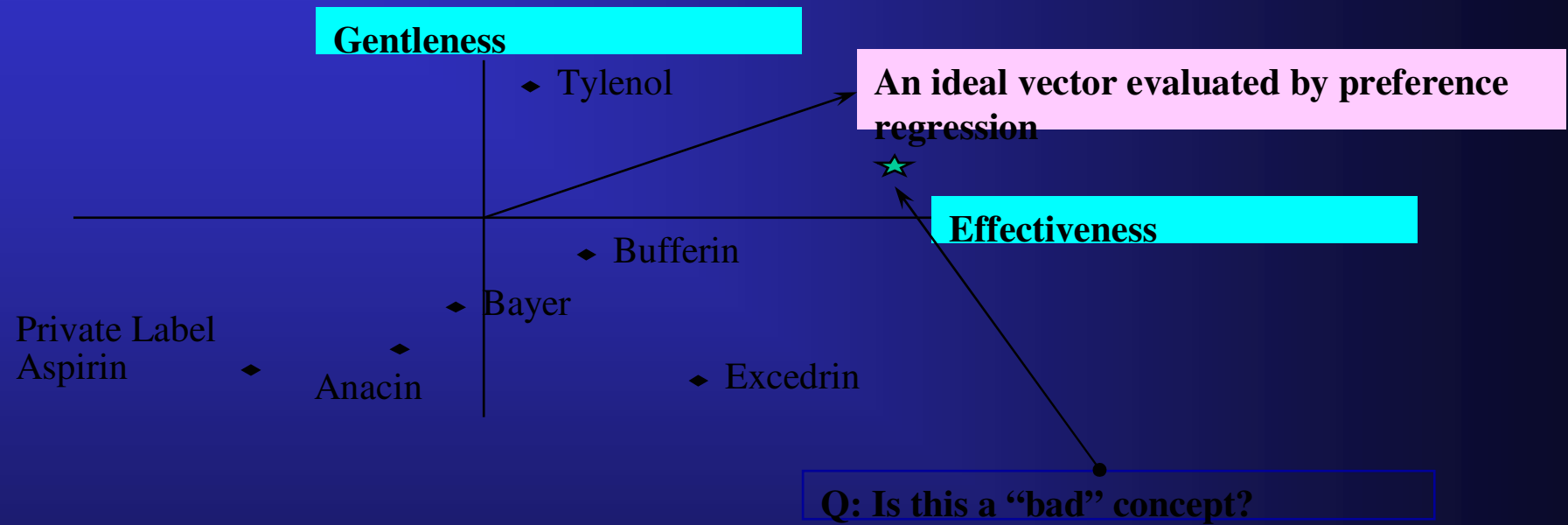


**Q: What would be the best “place” for a new product?**



**Q: s positioning of the new product consistent with s the ideal point or the ideal vector?**

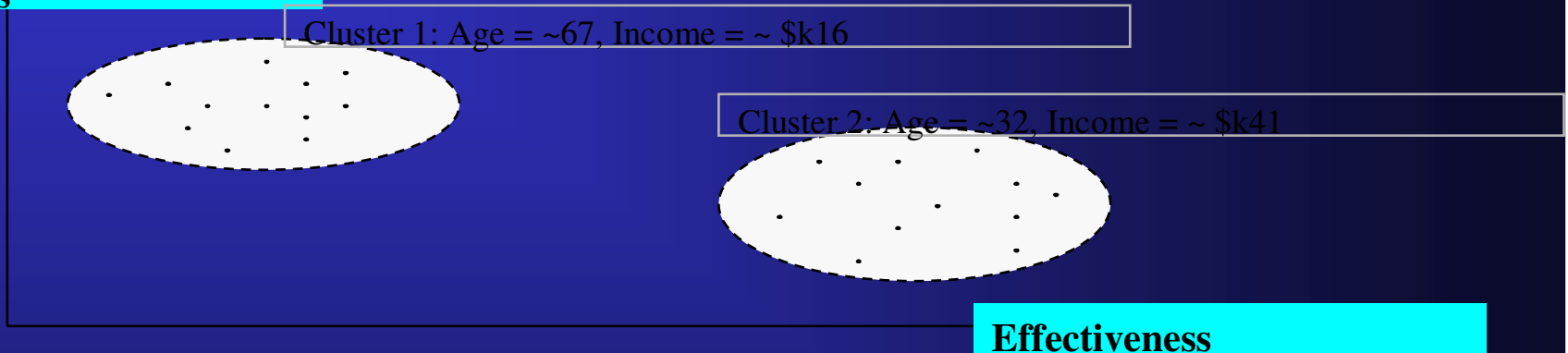
# Benefit Segmentation



# Benefit Segmentation - Positioning by Segments

Hypothetical Cluster analysis to identify Benefit Segments for pain relievers

**Gentleness**



Benefit Segmentation of pain relievers

