Conjoint Analysis

- Research question: To what extent does each component (factor) contribute to the total utility of a product?

- Assumption:

   Total utility = Sum of all partial utilities

- Data base of the Conjoint Analysis are preferences of the interviewed subjects

- Important application: Design of a new product according to the requirements of the market

Conjoint Analysis

- Factors and their values are defined by the researcher in advance

- The various combinations of the factor values yield fictive products that are being ranked by the interviewed persons

- With Conjoint Analysis it is possible to derive metric partial utilities from the ranking results

- The summation of these partial utilities therefore results in metric total utilities

Conjoint Analysis

- Independent variables: Object attributes

- Dependent variable: Preferences of the interviewed person for the fictive products

- The utility structure of a number of persons can be computed through aggregation of the single results

Conjoint Analysis

❖ **Factors and Factor Values**

Important for the choice of factors and their values are:

- ➢ Relevance
- ➢ Interference
- ➢ Independence
- ➢ Realisability
- ➢ Compensatory relationships of the various factor values
- ➢ They do not constitute exclusion criteria
- ➢ Terminableness

Conjoint Analysis

- *Example: Attitudes towards dishwashing products*
  1. Clean: glass/dishes clean
  2. Shiny: glass/dishes shiny
  3. Smell: Non-perfumed/lemon fresh/intensive lemon fresh
  4. Quantity: small/medium/x-large
  5. Packaging: loose in box/tab in plastic/tab in dissolving plastic
  6. Design: single/multi-colored/multi-colored + ball

Conjoint Analysis

❖ **Survey Design**

1. Definition of the incentives

   Incentive = Combination of factor values that is presented for rating

   ➢ Profile method: incentive results from the combination of the different factor values of each factor

   ➢ Two-factor method (Trade-off Analysis): Only two factors are used at a time. For each of these pairs a Trade-off-Matrix is set up

   Which method is used depends on how demanding the interview is for the respondent, the expenditure of time and the closeness to reality

Conjoint Analysis

❖ **Survey Design**

2. Number of incentives

   ➢ Can increase very fast depending on the number of factors and their values. It is advisable to choose a subset (reduced design) out of all theoretically possible incentives (full design). There shouldn't be more than 20 fictive products.

   ➢ Choice of the number can be based on random selection (not very common) or systematic selection

   ➢ Possible are symmetric (= all factors have the same number of values) and asymmetric (divergent number of values) designs

   ➢ Addelman developed several "Basic Plans" for the construction of reduced asymmetric designs

Conjoint Analysis

- *Example (cont.): Attitudes towards dishwashing products*
  - The full design consists of 2x2x3x3x3x3=324 different incentives
  - The reduced design consists of 16 incentives/cards, therefore we get a ranking from 1…16 for each person

Clean	Shiny	Smell	Quantity	Packaging	Design	STATUS_	CARD_
1,00	2,00	1,00	2,00	2,00	1,00	0	1
2,00	1,00	2,00	2,00	1,00	2,00	0	2
1,00	2,00	1,00	1,00	1,00	2,00	0	3
2,00	2,00	2,00	1,00	2,00	3,00	0	4
2,00	2,00	1,00	1,00	1,00	1,00	0	5
1,00	1,00	3,00	1,00	2,00	2,00	0	6
2,00	1,00	1,00	3,00	2,00	1,00	0	7
2,00	2,00	3,00	2,00	1,00	1,00	0	8
1,00	1,00	1,00	1,00	1,00	1,00	0	9
2,00	1,00	1,00	1,00	1,00	3,00	0	10
1,00	2,00	3,00	3,00	1,00	3,00	0	11
2,00	1,00	3,00	1,00	3,00	1,00	0	12
1,00	1,00	2,00	3,00	1,00	1,00	0	13
1,00	2,00	2,00	1,00	3,00	1,00	0	14
1,00	1,00	1,00	2,00	3,00	3,00	0	15
2,00	2,00	1,00	3,00	3,00	2,00	0	16

Conjoint Analysis

➢ The column "Card_" shows the numbering of the cards

➢ The column "Status_" can show the values 0, 1 or 2.

incentives that are part of the reduced design get the number 0

A value of 1 tells us that the corresponding card is a so-called "holdout card" that isn't used for the estimation of the utility values, but for validation.

A value of 2 indicates a simulation card that isn't presented to the interviewed persons

Conjoint Analysis

❖ **Rating of the incentives**

Possibilities:

➢ Ranking

➢ Rough classification into groups of different utility with succeeding ranking within these groups. Aggregation of these results leads to a total ranking. Used when there is a large number of incentives.

➢ Rating scales

➢ Paired comparison

Conjoint Analysis

❖ **Estimation of the utility values**

➢ Conjoint Analysis is used to determine partial utilities ("partworths") β for all factor values based upon the ranked data

➢ Furthermore, with this partworths it is possible to compute the metric total utilities y of all incentives and the relative importance of the single object attributes

➢ Individual Conjoint Analysis: For each person utility values are computed

➢ Combined Conjoint Analysis: Only one value for each factor category

Conjoint Analysis

❖ **Estimation of the utility values**

➢ The additive model of the Conjoint Analysis is defined as:

$$y_k = \sum_{j=1}^{J} \sum_{m=1}^{M_j} \beta_{jm} \cdot x_{jm}$$

with

y_k : estimated total utility for incentive k

β_{jm} : partial utility for value (category) m of factor j

$$x_{jm} = \begin{cases} 1 \text{ if incentive k has value m of factor j} \\ 0 \text{ else} \end{cases}$$

Conjoint Analysis

❖ **Estimation of the utility values**

➢ Target criterion for the determination of the partial utilities: The resulting total utilities should yield a good representation of the empirically ranked data

➢ Related procedure for the determination of the partial utilities: monotonous analysis of variance

➢ Metric solution → metric ANOVA

Non-metric solution → monotonous ANOVA (not covered here)

Conjoint Analysis

❖ **Estimation of the utility values**

➢ Model for the metric solution

$$y_k = \mu + \sum_{j=1}^{J} \sum_{m=1}^{M_j} \beta_{jm} \cdot x_{jm}$$

Steps needed to get the solution:

1. Determination of μ (= average of all ranks)
2. Determination of the mean rank for each factor value (= which ranks did the interviewed person assign on the average in conjunction with this attribute?)

Conjoint Analysis

❖ **Estimation of the utility values**

3. The average utility of a factor value then equals the difference of its average empirical rank value and μ

   The partial utilities β are OLS

4. Now it is possible to compute the total utility of the incentives

➢ Missing values are being replaced by zeroes

➢ There shouldn't be too much missing values in order to prevent the analysis from breaking down

Conjoint Analysis

❖ **Aggregation of the utility values**

➢ For the comparability of the individual analyses of the single persons it is necessary to carry out a standardisation:

1. The factor value that yields the lowest utility contribution is set to zero

2. The differences between the single partial utility values and the lowest partial utility value are computed:

$$\beta_{jm}^{*} = \beta_{jm} - \beta_{j}^{Min}$$

Conjoint Analysis

❖ **Aggregation of the utility values**

    3.   The total utility value of the incentive that shows the strongest preference among all persons is set to one. The partial utility values are computed accordingly:

$$\hat{\beta}_{jm} = \frac{\beta^{*}_{jm}}{\displaystyle\sum_{J=1}^{J} \max_{m}\left\{\beta^{*}_{jm}\right\}}$$

➢   The importance of an attribute for the preference change results from the utility range of the several values of the attribute

Conjoint Analysis

❖ **Aggregation of the utility values**

> ➢ A high range value means that by varying the concerning attribute a meaningful change of the total utility can occur

> ➢ The relative importance of a factor can be obtained by weighting on the sum of the ranges

Conjoint Analysis
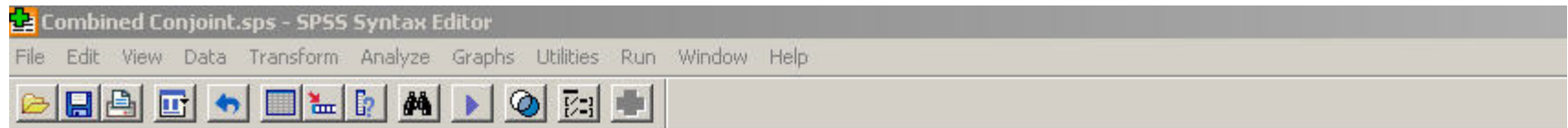
❖ **Aggregation of the utility values**

➢ The largest standardised partial utility values allow to make a statement about the relative importance of the attributes at the same time

➢ Aggregation always implies the loss of information

➢ If the aggregated utility structures are highly heterogeneous, this loss of information is substantial

➢ In this case, more homogeneous subgroups should be formed by performing a Cluster Analysis

Conjoint Analysis

- *Example (cont.): Attitudes towards dishwashing products*

  - There is no graphical user interface available in SPSS that would allow the performance of a Conjoint Analysis. The SPSS-Syntax has to be used in order to retrieve the required procedure CONJOINT.

  - After the Syntax-window has been opened, the following commands should be entered:

Conjoint Analysis

- *Example (cont.): Attitudes towards dishwashing products*


```
Combined Conjoint.sps - SPSS Syntax Editor
File   Edit   View   Data   Transform   Analyze   Graphs   Utilities   Run   Window   Help

TITLE "Conjoint Analysis".

CONJOINT
        plan    = "C:\download\CJ\Design.sav"
        /data    = "C:\download\CJ\Q2new.sav"
        /factors = Clean (DISCRETE) Shiny (DISCRETE) Smell (DISCRETE) Quantity (DISCRETE) Packaging (DISCRETE) Design (DISCRETE)
        /rank    = Q2a Q2b Q2c Q2d Q2e Q2f Q2g Q2h Q2i Q2j Q2k Q2l Q2m Q2n Q2o Q2p
        /print   = all
        /utility = "C:\download\CJ\Total Utility.sav"
        /plot    = SUMMARY.

SUBTITLE "Utility Values".

get file "C:\download\CJ\Total Utility.sav".

LIST.
```

Conjoint Analysis

The commands in the syntax have the following meaning:

➢ With the TITLE – statement it is possible to define a title for the results in the output window

➢ The actual Conjoint Analysis is performed with help of the procedure CONJOINT. There are various subcommands within this procedure:

- The PLAN subcommand tells CONJOINT which file contains the data for the reduced design

Each of the following commands is preceded by a slash (/):

Conjoint Analysis

- The subcommand DATA tells CONJOINT which file contains the rankings of the respondents

- FACTORS is used to define the relationship between the factors and the preference values (ranks) of the interviewed persons. The relationships can be DISCRETE, LINEAR, IDEAL or ANTIIDEAL. In this example, all of the relationships are DISCRETE because only categorical data is used. It means that no assumptions are made concerning the relationship between variables and ranks.

- The subcommand RANK tells CONJOINT that the data is coded in such a way that the sequence of the variables corresponds to the sequence of the cards

Conjoint Analysis

Alternatively, the subcommands SCORE or SEQUENCE can be used.

- By the UTILITY subcommand, a file with the name "Utility.sav" is being created, where various information is saved. This information contains the following variables:
  - CONSTANT: The constant term of the conjoint estimation
  - Several partial utility values
  - SCORE1 to SCORE16: Total utility values of the reduced design

➢ With the SUBTITLE command, a new subsection of the output is introduced

➢ With the "get file" statement the contents of the "Utility.sav" file are printed in the output window

For our example, the results would be the following:

Utilities

		Utility Estimate	Std. Error
Clean	Special additive to make glass clean	-,230	,091
	Special additive to make dishes clean	,230	,091
Shiny	special additive to make glass shiny	,227	,091
	special additive to make dishes shiny	-,227	,091
Smell	non-perfumed	-,083	,122
	lemon fresh	,639	,143
	intensive lemon fresh	-,555	,143
Quantity	standard (small 22 tabs)	-,830	,122
	standard (medium 44 tabs)	,834	,143
	standard (X-lage 88 tabs)	-,004	,143
Packaging	packaged loose in a box	-,299	,122
	single packaged tab in plastic	,430	,143
	single packaged tab in dissolving plasti	-,130	,143
Design	single cubicle	-,616	,122
	multi-colored cubicle	,236	,143
	multi-colored cubicle + integrated ball	,380	,143
(Constant)		8,957	,110

Importance Values

Clean	8,356
Shiny	8,254
Smell	21,727
Quantity	30,281
Packaging	13,265
Design	18,118

Averaged Importance Score

Correlations[a]

	Value	Sig.
Pearson's R	,981	,000
Kendall's tau	,933	,000

a. Correlations between observed and estimated preferences

Conjoint Analysis

❖ **Interpretation of the results**

➤ The term "Averaged importance score" under the box with the importance values tells us that a Combined Conjoint Analysis has been performed. These values give us a measure (in percent) of the relative importance of the single factors for the determination of the utilities. We can see that "quantity" is the most important factor whereas "shiny" is the least important.

➤ The Correlation coefficients are a measure for the quality of reproduction of the empirical data by the results of the Conjoint analysis.

➤ The estimated partial values for the factor values are given in the column "Utility Estimate"

Conjoint Analysis

❖ **Interpretation of the results**

➢ Summaries of Importance and Utility in form of bar charts are also given

➢ The partial utility values allow the computation of the total utility values for all the incentives. For example, the total utility value for card 1 would be:

8,96-0,23-0,23-0,08+0,83+0,43-0,62=9,06

➢ The constant term μ can be interpreted as base utility, and the other factor values contrast with μ in positive or negative direction

Conjoint Analysis

❖ **Interpretation of the results**

➢ Following this, the total utility values for all the cards can be computed similarly. Like it was said above, the results of this computations can be found in the "Utility.sav" file (Score1 to Score16):

```
The variables are listed in the following order:

LINE    1: CONSTANT Clean1 Clean2 Shiny1 Shiny2 Smell1 Smell2

LINE    2: Smell3 Quantity1 Quantity2 Quantity3 Packaging1 Packaging2 Packaging3

LINE    3: Design1 Design2 Design3 SCORE1 SCORE2 SCORE3 SCORE4

LINE    4: SCORE5 SCORE6 SCORE7 SCORE8 SCORE9 SCORE10 SCORE11

LINE    5: SCORE12 SCORE13 SCORE14 SCORE15 SCORE16
□
```

CONSTANT:	8,96	-,23	,23	,23	-,23	-,08	,64
Smell3:	-,56	-,83	,83	,00	-,30	,43	-,13
Design1:	-,62	,24	,38	9,07	10,82	7,52	9,58
SCORE5:	7,13	8,23	9,14	8,32	7,13	8,58	8,02
SCORE12:	7,28	8,67	7,56	9,96	8,98		

```
Number of cases read:  1    Number of cases listed:  1
```

Conjoint Analysis

❖ **Interpretation of the results**

➢ We can see that card 2 has the highest total utility (10,82). A product that makes the dishes clean, the glass shiny, has a lemon-fresh smell, a multi-colored design and is packaged loose in a medium sized box therefore is the most preferred product out of the 16 alternatives.

➢ Here, just the results for the reduced design are given. But it is also possible to compute total utilities for all the products of the complete design, even though they were not part of the interview.

Conjoint Analysis

❖ **Interpretation of the results**

➢ It is therefore possible to find out that, for example, a product that is similar to card 2, except that the box contains single packaged tabs, would have an even higher total utility than card 2.

Conjoint Analysis

❖ **Clustering the partial utility values**

➢ As aforementioned, a Cluster Analysis should be performed before aggregating the results if the aggregated utility structures are not homogeneous

➢ To get the data that is needed to do this, a new line has to be inserted into the syntax:

```
TITLE "Conjoint Analysis".

CONJOINT
        plan      = "C:\download\CJ\Design.sav"
        /data     = "C:\download\CJ\Q2new.sav"
        /factors = Clean (DISCRETE) Shiny (DISCRETE) Smell (DISCRETE) Quantity (DISCRETE) Packaging (DISCRETE) Design (DISCRETE)
        /subject = Person
        /rank     = Q2a Q2b Q2c Q2d Q2e Q2f Q2g Q2h Q2i Q2j Q2k Q2l Q2m Q2n Q2o Q2p
        /print    = all
        /utility = "C:\download\CJ\Person Utility.sav"
        /plot    = SUMMARY.

SUBTITLE "Utility Values".

get file "C:\download\CJ\Person Utility.sav".

LIST.
```
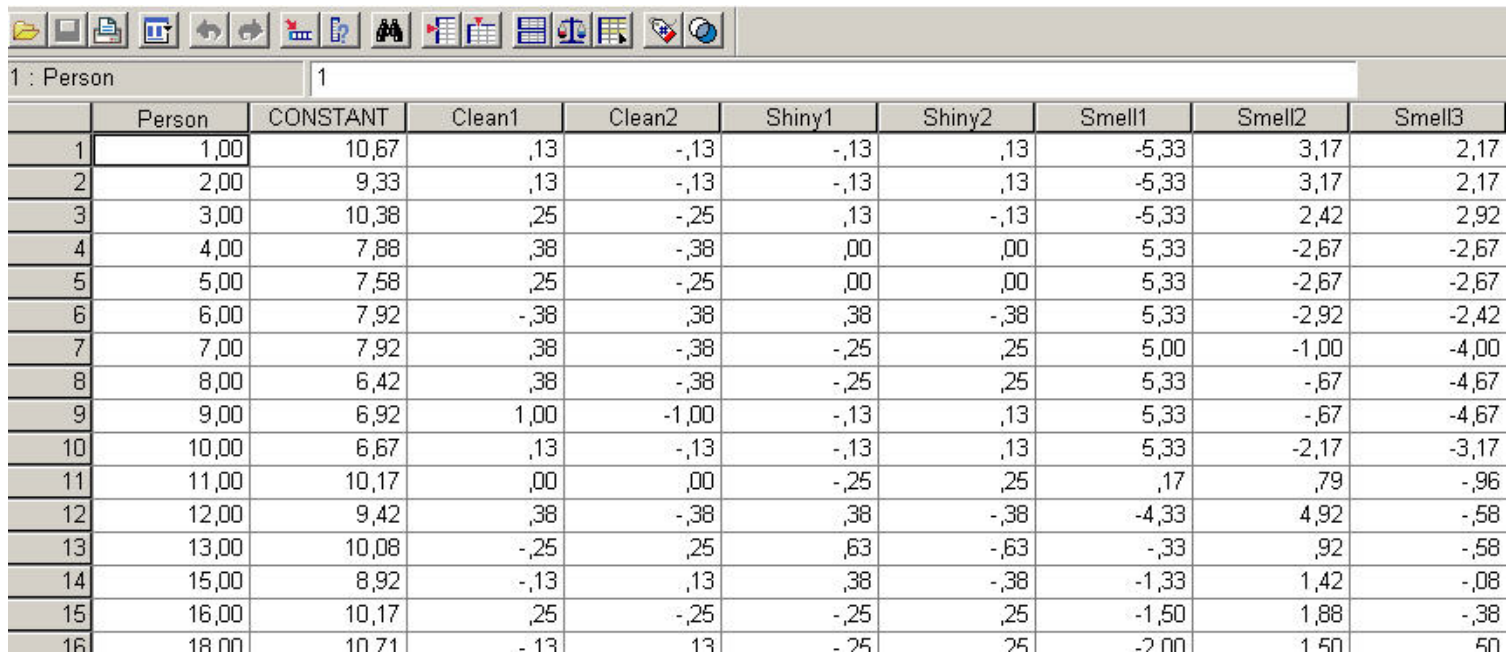
Conjoint Analysis

❖ **Clustering the partial utility values**

➢ The results for each person are computed and saved into the file "Person Utility.sav":

	Person	CONSTANT	Clean1	Clean2	Shiny1	Shiny2	Smell1	Smell2	Smell3
1	1,00	10,67	,13	-,13	-,13	,13	-5,33	3,17	2,17
2	2,00	9,33	,13	-,13	-,13	,13	-5,33	3,17	2,17
3	3,00	10,38	,25	-,25	,13	-,13	-5,33	2,42	2,92
4	4,00	7,88	,38	-,38	,00	,00	5,33	-2,67	-2,67
5	5,00	7,58	,25	-,25	,00	,00	5,33	-2,67	-2,67
6	6,00	7,92	-,38	,38	,38	-,38	5,33	-2,92	-2,42
7	7,00	7,92	,38	-,38	-,25	,25	5,00	-1,00	-4,00
8	8,00	6,42	,38	-,38	-,25	,25	5,33	-,67	-4,67
9	9,00	6,92	1,00	-1,00	-,13	,13	5,33	-,67	-4,67
10	10,00	6,67	,13	-,13	-,13	,13	5,33	-2,17	-3,17
11	11,00	10,17	,00	,00	-,25	,25	,17	,79	-,96
12	12,00	9,42	,38	-,38	,38	-,38	-4,33	4,92	-,58
13	13,00	10,08	-,25	,25	,63	-,63	-,33	,92	-,58
14	15,00	8,92	-,13	,13	,38	-,38	-1,33	1,42	-,08
15	16,00	10,17	,25	-,25	-,25	,25	-1,50	1,88	-,38
16	18,00	10,71	-,13	,13	-,25	,25	-2,00	1,50	,50

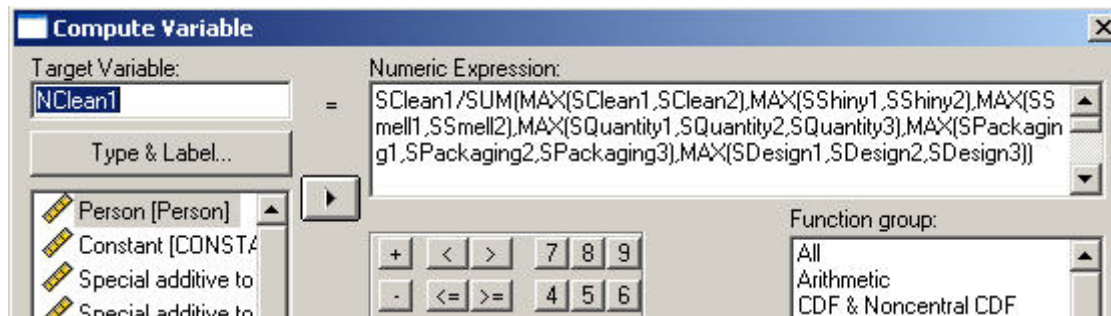# Conjoint Analysis

❖ **Clustering the partial utility values**

➢ Now the formulae given in the section "Aggregation of the utility values" are used to calculate the standardised partial utility values (**Transform > Compute)**. It is reasonable to do this in two steps:

1. Compute the differences between the single partial utility values and the lowest partial utility value:
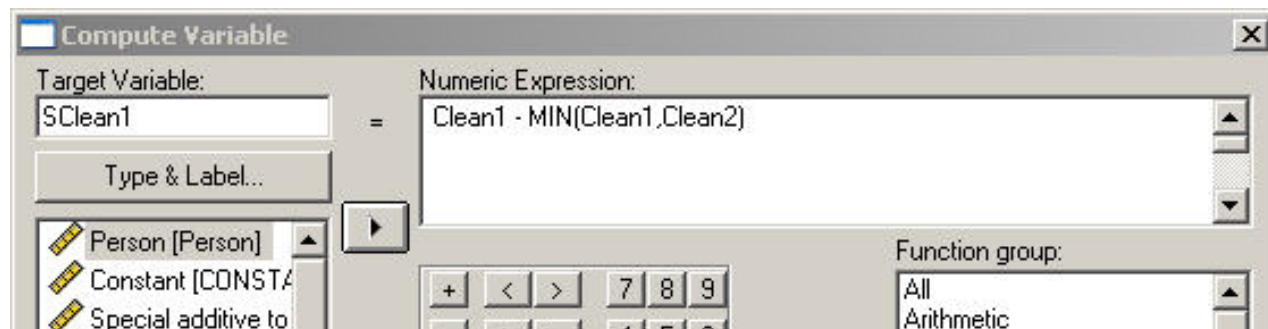
Conjoint Analysis

❖ **Clustering the partial utility values**

   2. Set the most preferred incentive to 1 and compute the standardised partial utility values accordingly:

➢ These two steps have to be repeated for each factor value. After that, the partial utility values and the standardised partial utility values can be attached to the original data file

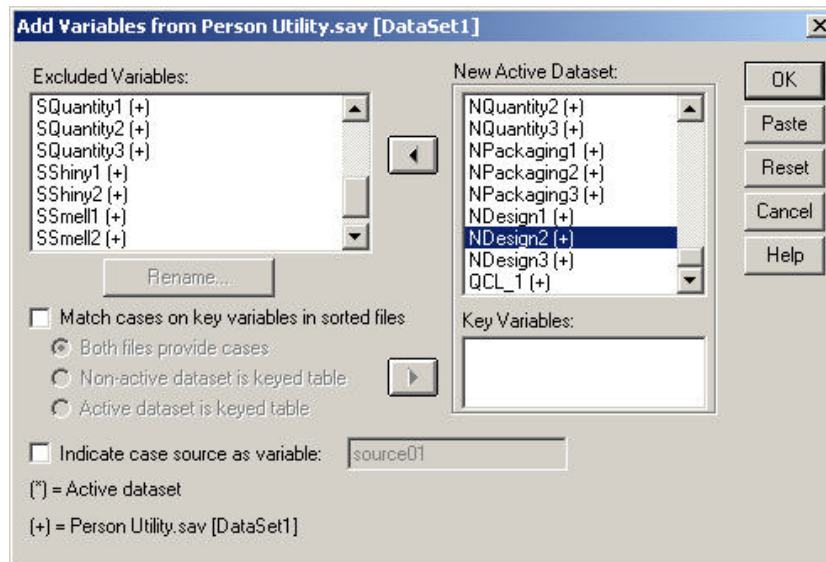# Conjoint Analysis

❖ **Clustering the partial utility values**

➢ To do this, use **Data > Merge files > Add variables**

➢ If both data files are opened and the original data file is the active one, the option "An open dataset" can be chosen. Otherwise there is the possibility to browse for the corresponding file:
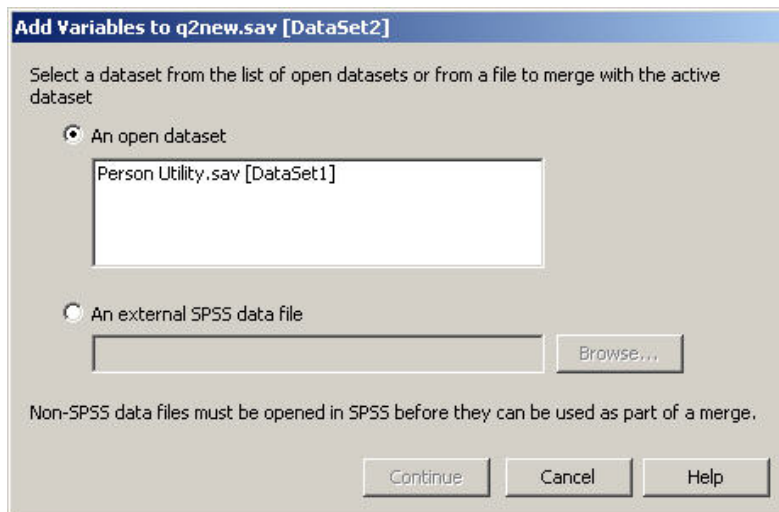
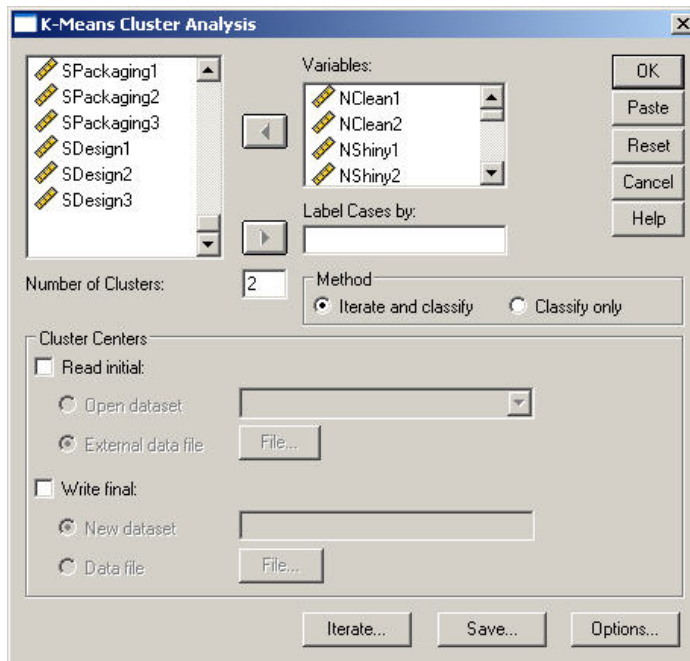# Conjoint Analysis

❖ **Clustering the partial utility values**

➢ We can exclude variables from the dataset that aren't needed anymore (i.e., the SClean1 to SDesign3 variables). The variables that aren't excluded form the new data set:

Conjoint Analysis

❖ **Clustering the partial utility values**

➢ The next step is the actual performance of the Cluster Analysis using the menu **Analyze > Classify > K-means Cluster**

Conjoint Analysis

❖ **Clustering the partial utility values**

➢ All the standardised variables from NClean1 to NDesign3 are used as input

➢ To save the results of the Cluster Analysis in the data file, the box "Cluster Membership" in the "Save"-menu has to be checked

➢ It is necessary to try different numbers of clusters to find out what the (as regards content) best solution is

➢ Cross-classifying the variable for the Cluster membership with certain demographic variables is needed to assess the best Cluster solution

Conjoint Analysis

❖ **Clustering the partial utility values**

➢ If the Cluster solution has finally been found, it is possible to compute an Combined Conjoint Analysis for each Cluster separately (using **Data > Select Cases**)