



**EXAMINATION: MULTIVARIATE ANALYSIS**  
**EXAMINER: PROF. DR. B. ERICHSON**

WS 2003/04

You are allowed to use a pocket calculator (in accordance with the instructions given by the examination office) and a translating dictionary from your native language to English (without any notes written in it). The answers to all questions should be made in one language, please use English or German.  
**All of the 4 exam questions must be answered** (the estimated time for each question is given).  
**This examination has 3 pages.**

**Question 1 (20 Min.)**

An empirical market research study was performed to analyse consumer perceptions of competing margarine brands. The respondents had to judge selected brands on relevant attributes: unsaturated fatty acid, calory content, vitamin content, shelf life and price.

- a) Explain the task of Factor Analysis.
- b) A key element of factor analysis are Eigenvalues. What is their meaning and how are they related to factor loadings?
- c) Look at the following tables. How many factors would you extract and why?

**Communalities**

	Initial	Extraction
unsaturated fatty acid	1,000	,931
calory content	1,000	,736
vitamin content	1,000	,927
shelf life	1,000	,993
price	1,000	,992

Extraction Method: Principal Component Analysis.

**Total Variance Explained**

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2,645	52,903	52,903
2	1,934	38,678	91,581
3	,369	7,374	98,955
4	3,930E-02	,786	99,741
5	1,296E-02	,259	100,000

Extraction Method: Principal Component Analysis.

- d) Interpret the following two tables from a SPSS-Output. What distinguishes the right table from the left one?

**Component Matrix<sup>a</sup>**

	Component	
	1	2
unsaturated fatty acid	,937	-,229
vitamin content	,929	-,254
calory content	,843	-,160
price	,277	,957
shelf life	,342	,936

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

**Rotated Component Matrix<sup>a</sup>**

	Component	
	1	2
unsaturated fatty acid	,964	3,287E-02
vitamin content	,963	6,726E-03
calory content	,855	7,396E-02
price	7,896E-03	,996
shelf life	7,646E-02	,994

Extraction Method: Principal Component Analysis.  
 Rotation Method: Varimax with Kaiser Normalization.

- e) Try to interpret the two factors or components!

**Please turn the page**



**Question 2 (15 Min.)**

To learn about the effects of different in-store locations (Normal-shelf, Extra-shelf or Cool-shelf) on sales of margarine an experiment was performed in three different supermarkets. The following table shows the sales quantities per 1000 customers.

	Day 1	Day 2	Day 3	Day 4	Day 5	Mean
Supermarket 1 "Normal-shelf"	47	39	40	46	45	43.4
Supermarket 2 "Extra-shelf"	68	65	63	59	67	64.4
Supermarket 3 "Cool-shelf"	59	50	51	48	53	52.2
Total Mean						53.3

- Which analysis method should be used to analyse the data? Give reasons.
- Explain the terms "explained variation" and "unexplained variation" for the given experiment.
- Compute the explained variation for the data above.
- Please test the significance of the observed effects by using the SPSS Output and F-table given below.

ANOVA<sup>a,b</sup>

			Unique Method	
			Sum of Squares	df
Menge Margarine	Main Effects	Plazierung	1112,133	2
	Model		1112,133	2
	Residual		175,200	12
	Total		1287,333	14

- Menge Margarine by Plazierung
- All effects entered simultaneously

F-Werte-Tabelle (Signifikanzniveau 10%):

v <sub>1</sub>	1	2	3	4	5	6	7	8	9	10	12
v <sub>2</sub>											
5	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34	3,32	3,30	3,27
6	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98	2,96	2,94	2,90
7	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75	2,72	2,70	2,67
8	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59	2,56	2,54	2,50
9	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47	2,44	2,42	2,38
10	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38	2,35	2,32	2,28
20	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00	1,96	1,94	1,89
30	2,88	2,49	2,28	2,14	2,05	1,98	1,93	1,88	1,85	1,82	1,77
40	2,84	2,44	2,23	2,09	2,00	1,93	1,87	1,83	1,79	1,76	1,71

v<sub>1</sub> = Freiheitsgrade im Zähler; v<sub>2</sub> = Freiheitsgrade im Nenner

**Question 3 (10 Min.)**

- Explain shortly the basic idea of cluster analysis.
- Perform one agglomeration step using the single linkage method on the following distance matrix:

	1	2	3	4	5	6
1	0					
2	0.9	0				
3	0.3	0.6	0			
4	0.8	0.3	0.5	0		
5	0.7	0.3	0.4	0.1	0	
6	0.5	0.4	0.2	0.3	0.2	0

Please turn the page



**Question 3 (15 Min.)**

The following SPSS-Output is given:

Model Summary <sup>b</sup>					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,962 <sup>a</sup>	,926	,888	150,12600	3,490

a. Predictors: (Constant), VISITS, PROMOT, PRICE  
b. Dependent Variable: SALES

ANOVA <sup>b</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1681029	3	560342,900	24,862	,001 <sup>a</sup>
	Residual	135226,9	6	22537,817		
	Total	1816256	9			

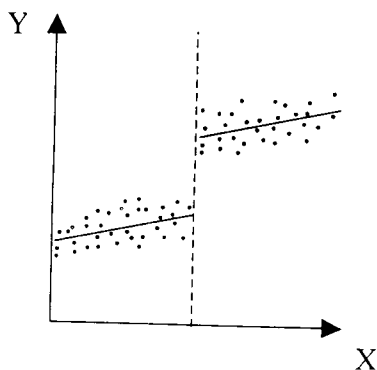
a. Predictors: (Constant), VISITS, PROMOT, PRICE  
b. Dependent Variable: SALES

Coefficients <sup>a</sup>						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-6,866	673,205		-,010	,992
	PRICE	9,927	38,164	,034	,260	,803
	PROMOT	,655	,103	,794	6,382	,001
	VISITS	11,085	4,428	,345	2,504	,046

a. Dependent Variable: SALES

- a) Identify the Data Analysis Method.
- b) Write down the estimated function.
- c) Judge the global goodness of fit of this function.
- d) How can you check the significance of a regressor? Does the variable VISITS show a significant effect here?
- e) How can a structural breakage in the data (see example beneath) be modelled? **Help:** Which kind of variables is necessary?



**Good Luck!**