| EXAMINATION: | MULTIVARIATE ANALYSIS | WS 2005/06 |
| EXAMINER: | PROF. DR. B. ERICHSON | |

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 4 pages.

Question 1 (10 Min.)

In an empirical market research experiment the likeability of 3 different advertising spots was measured by ratings of female and male respondents. The results in aggregated form are shown in the following table.

Descriptive Statistics

Dependent Variable: likeability score

advertising spot	gender of respondent	Mean	Std. Deviation	N
spot 1	female	4,90	,994	10
	male	6,70	1,703	10
	Total	5,80	1,642	20
spot 2	female	7,60	1,265	10
	male	3,70	1,494	10
	Total	5,65	2,412	20
spot 3	female	8,50	1,354	10
	male	9,20	,789	10
	Total	8,85	1,137	20
Total	female	7,00	1,948	30
	male	6,53	2,649	30
	Total	6,77	2,317	60

Tests of Between-Subjects Effects

Dependent Variable: likeability score

Source	Type III Sum of Squares	df
Corrected Model	225,133(a)	5
Intercept	2747,267	1
spot	130,433	2
gender	3,267	1
spot * gender	91,433	2
Error	91,600	54
Total	3064,000	60
Corrected Total	316,733	59

a R Squared = ,711 (Adjusted R Squared = ,684)

F-value-table (statistical significance level 10%):

v_1 / v_2	1	2	3	4	5	6	7	8	9	10	12
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_1 = degrees of freedom in numerator;
v_2 = degrees of freedom in denominator

a) Which analysis method should be used to analyze the data? Give some reasons!

b) Specify the full factorial model!

c) Formulate the specific null hypotheses and test them by using the given tables!

Please turn the page

Question 2   (15 Min.)

An empirical market research study was performed to analyse consumer perceptions of competing cola brands. The respondents had to judge selected brands on 7 relevant attributes.

a) Explain the purpose and the steps of Factor Analysis!

b) Communalities and Eigenvalues are key elements of factor analysis. What are their meanings and how are they related to factor loadings? Look at the following table. Compute the different Communalities and Eigenvalues! Interpret the two factors! Would you extract a third factor? Justify your answer!

Rotated Component Matrix(a)

	Component	
	1	2
satisfy thirst	,362	,822
low price	-,839	-,440
sweet	-,033	-,951
high quality	,851	,494
fresh	,557	,786
the brand does a lot of advertising	,961	,228
the brand is very present	,945	,059

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 3 iterations.

c) Explain briefly the meaning of factor rotation! Are eigenvalues, communalities and explained variance affected by factor rotation?

d) Interpret the following plot!
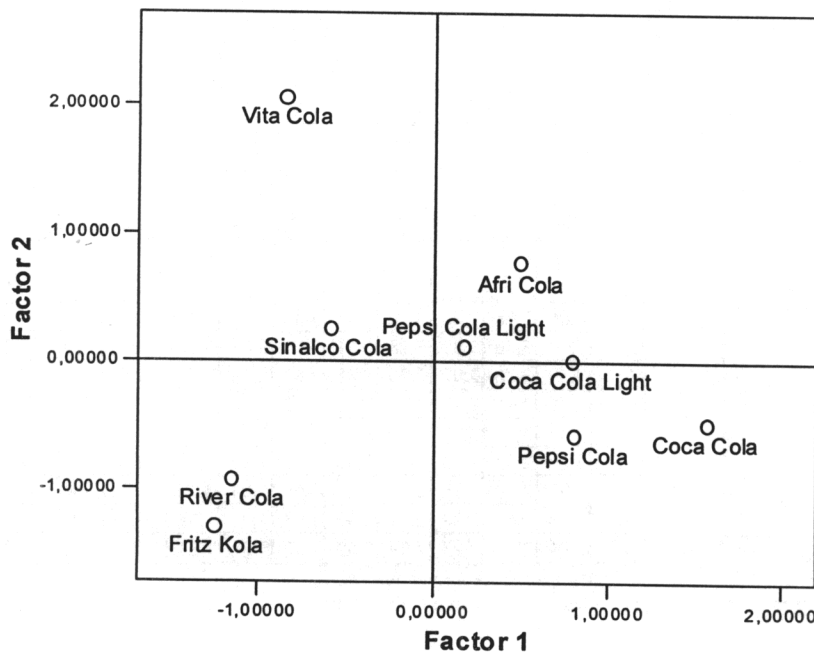
Question 3 (20 Min.)

A SPSS analysis leads to the following Output:

ANOVA

Model		Sum of Squares	df
1	Regression	1470744,255	3
	Residual	84255,745	6
	Total	1555000,000	9

a Predictors: (Constant), promotion spendings in thousand Euro, sales area in square metres, price in Euro
b Dependent Variable: sales in units

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	2491,665	940,117	
	sales area in square metres	,432	,150	,452
	price in Euro	-127,876	59,778	-,542
	promotion spendings in thousand Euro	9,155	31,650	,055

a Dependent Variable: sales

a) Identify the Data Analysis Method! Explain the purpose of this method!

b) Write down the estimated functio! What will change if the price is measured in Cent?

c) Judge the global goodness of fit of this function by
- computing R^2
- performing a F-Test (assume the theoretical F-Value: $F_{3,6}(0,95)=4,76$).

d) Judge the importance of the variables! Which variable has the strongest influence on sales?

e) Test the significance of the coefficients (using the t-table)! Which variables are statistically significant at the 5% level in a two-sided test?

		Significance Level				
1-Tailed:		,10	,05	,025	,01	.005
2-Tailed:		,20	,10	,05	,02	.01
D F	1	3,078	6,314	12,706	31,821	63,657
e r	2	1,886	2,920	4,303	6,965	9,925
g e	3	1,638	2,353	3,182	4,541	5,841
r e	4	1,533	2,132	2,776	3,747	4,604
e d	5	1,476	2,015	2,571	3,365	4,032
e o	6	1,440	1,943	2,447	3,143	3,707
s m	7	1,415	1,895	2,365	2,998	3,499
	8	1,397	1,860	2,306	2,896	3,355
o	9	1,383	1,833	2,262	2,821	3,250
f	10	1,372	1,812	2,228	2,764	3,169

f) Construct a 95% confidence interval of β_2 (price in Euro)!

g) What does Heteroskedasticity mean?

Please turn the page

Question 4 (15 Min.)

a) Explain shortly the main purposes of cluster analysis!

b) Cluster analysis can be based on different proximity measures. Classify the following proximity measures: City-Block, Tanimoto, Q-Correlation, Euclidean, Russel/Rao, Minkowski, and Simple Matching into the following table:

	Nominal Data	Metric Data
Similarity Measures		
Distance Measures		

c) Perform <u>one</u> agglomeration step using the Complete-Linkage-Method on the following matrix:

Proximity Matrix

Case	Binary Squared Euclidean Distance				
	1:Nokia N90	2:Siemens M131	3:Sony Ericsson W	4:Motorola V180	5:Samsung SGH-Z50
1:Nokia N90	0	6	5	8	3
2:Siemens M131	6	0	7	4	7
3:Sony Ericsson W	5	7	0	7	6
4:Motorola V180	8	4	7	0	7
5:Samsung SGH-Z50	3	7	6	7	0

This is a dissimilarity matrix

d) Use the following table to draw the dendrogram!

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	5	3,000	0	0	3
2	2	4	4,000	0	0	4
3	1	3	6,000	1	0	4
4	1	2	8,000	3	2	0

Good Luck!