

Anleitung zur Bedienung des  
Statistik-Programms  
IBM© SPSS© Statistics Version 23

-Oktober 2019-

© IMBEI Mainz 2019

## Inhalt

|   |    |
|---|----|
| 1. Die Benutzer-Oberfläche von IBM© SPSS© 23                                | 1  |
| 1.1. Der Daten -Editor  | 1  |
| 1.2. Der Ausgabe-Viewer (das Ausgabe-Fenster)                               | 2  |
| 1.3. Daten-Erweiterung und Daten-Auswahl                                    | 5  |
| 2. Spezielle Auswertungen mit SPSS  | 9  |
| 2.1. Bestimmung absoluter und relativer Häufigkeiten                        | 9  |
| 2.2. Vierfeldertafeln, Exakter Fisher-Test, Chi-Quadrat-Unabhängigkeitstest | 9  |
| 2.3. Statistische Maßzahlen für stetige Variablen                           | 12 |
| 2.4. Boxplot  | 13 |
| 2.5. Der t-Test für unverbundene Stichproben                                | 14 |
| 2.6. Unverbundener Wilcoxon-Test (Mann-Whitney-U-Test)                      | 16 |
| 2.7. t-Test für verbundene Stichproben                                      | 16 |
| 2.8. Streudiagramm (Scatterplot)  | 19 |
| 2.9. Lineare Regression   | 20 |
| 2.10. Korrelations-Koeffizient nach Pearson                                 | 21 |
| 2.11. Partieller Korrelationskoeffizient                                    | 22 |
| 2.12. Logistische Regression  | 22 |
| 2.13. Analyse von Überlebenszeiten, KAPLAN-MEIER, Log-Rank-Test             | 26 |
| 2.14. Regression für zensierte Beobachtungen (Cox-Regression)               | 29 |
| 2.15. Bland-Altman-Plot   | 33 |
| 2.16. Kappa-Maß   | 34 |
| 2.17. Einfaktorielle ANOVA  | 35 |
| 2.18. Kruskal-Wallis-Test   | 36 |
| 2.19. Friedman-Test   | 37 |

|  |    |
|--|----|
| 4. Beschreibung der in dieser Anleitung benutzten Datenmengen                        |    |
| 3.1. Datenmenge HDF.SAV  | 38 |
| 3.2. Datenmenge KARZINOM.SAV   | 38 |
| 3.3. Datenmenge SCHWIMMBADNUTZUNG.SAV  | 39 |
| 3.4. Datenmenge SPORT_LMK.SAV  | 39 |
| 3.5. Datenmenge VERDINUM.SAV   | 40 |
| 4. Index der wichtigsten Befehlsfolgen   | 41 |
| 5. Wegweiser zum Aufspüren von Kennzahlen der statistischen<br>Auswertung in SPSS 23 | 43 |

# 1. Die Benutzer-Oberfläche von SPSS

SPSS ist ein statistisches Analyse-Paket. In diesem Programm arbeitet man normalerweise mit zwei Fenstern, dem Daten-Editor-Fenster und dem Ausgabe-Fenster. Für den Fenster-Wechsel ist jeweils der Menü-Befehl **Fenster** oder die Task-Leiste zu benutzen.

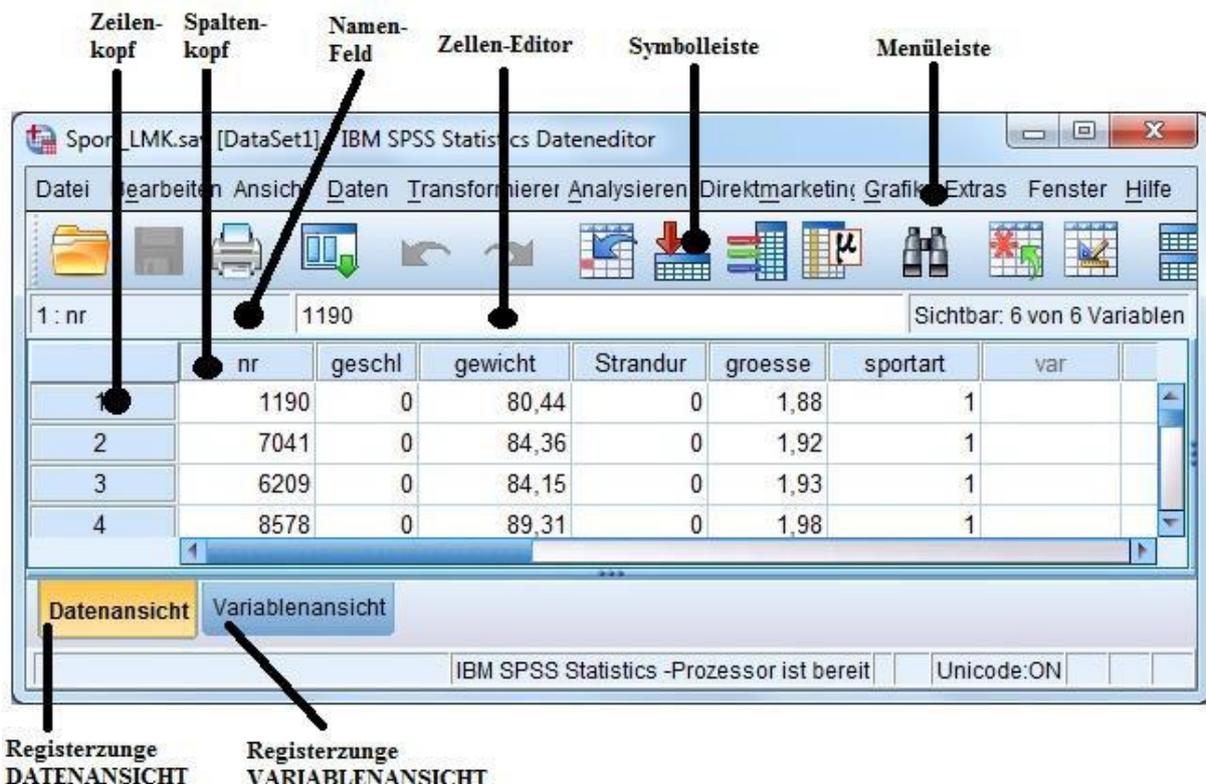
## 1.1. Der Daten -Editor

### 1.1.1. Aufbau des Daten-Editors

Der SPSS-Daten-Editor (auch "Daten-Fenster") ist in der Datenansicht für das Eingeben, Betrachten und Bearbeiten von Daten gedacht. Eingegebene Daten werden hier in tabellarischer Form dargestellt. Die Auswertung der Daten erfolgt normalerweise über Menübefehle und diese zugeordneten Dialogfelder.

In der ebenfalls tabellarischen Variablenansicht werden die Eigenschaften der Variablen festgelegt: der Name (kurzgehalten), Variablen-Typ (numerisch oder Datum oder Text), Anzahl der Zeichen und der Dezimalstellen, ein beschreibender Variablenname (Variablenname im Klartext) sowie Wertbezeichnungen (etwa "nein" für 0, "ja" für 1), sinnvoll nur bei kategorialen Variablen. Die restlichen Spalten der Variablenansicht werden nur selten benutzt und daher hier nichtberücksichtigt.

Der Daten-Editor (das Datenfenster) besteht aus folgenden Regionen:



Im Zeilenkopf steht die Nummer der Zeile. Im Spaltenkopf steht der Name der Variable. Die Zeilennummer der aktiven (markierten) Zelle sowie der dazugehörige Variablenname erscheinen im Namenfeld. Im Zellen-Editor kann man bei Bedarf Änderungen an dem Inhalt einer markierten Zelle vornehmen. Die Symbolleiste erlaubt den schnellen Zugriff auf häufig gebrauchte Kommandos. Über die Befehle der Menüleiste gelangt man in die verschiedensten Auswertungs-Optionen. Über die Registerzungen kann die Ansicht gewechselt werden.

### 1.1.2. Eine Datenmenge in den Daten-Editor laden

Nach dem Start von SPSS ist der Daten-Editor leer. Man kann nun entweder neue Daten eingeben oder eine bereits bestehende Datenmenge von der Festplatte in den Editor laden. Da man meist mehrere Sitzungen zur Auswertung derselben Datenmenge braucht, sei hier die Vorgehensweise des Ladens (des Öffnens) einer bereits bestehenden Datenmenge geschildert.

Menü-Befehlsfolge **Datei, Öffnen, Daten** ★ Dialogfeld *Daten öffnen*; darin wird die Liste der SPSS-Datenmengen dargestellt. Diese tragen immer die Dateinamen-Erweiterung **.sav**. Nun muss man die gewünschte Datenmenge durch Mausklick links markieren und die Befehls-Schaltfläche **Öffnen** betätigen ★ Datenmenge erscheint im SPSS-Editor und ist bereit zur Auswertung.

## 1.2. Der Ausgabe-Viewer (das Ausgabe-Fenster)

### 1.2.1. Der Aufbau des Ausgabe-Viewers

Der Ausgabe-Viewer ist das Darstellungsmedium für Auswertungsergebnisse, die in Form von Tabellen, Statistiken und Diagrammen angezeigt werden. Dieses Fenster wird bei der ersten Auswertung der Daten automatisch geöffnet. Es besitzt ebenso wie der Daten-Editor sowohl eine Menü- als auch eine Symbolleiste. Die zur Verfügung stehenden Befehle unterscheiden sich teilweise von denen des Daten-Editors, da sie speziell dem Zweck der Bearbeitung der Auswertungsergebnisse angepasst sind.

Die Darstellungszone ist durch eine senkrechte Rahmenleiste in zwei Flächen aufgeteilt. In der schmaleren, linken Hälfte ist eine Hierarchie der angezeigten Objekte (Tabellen, Statistiken, Diagramme) eingeblendet; in der breiteren, rechten Hälfte befinden sich diese Objekte selbst. Der Trennrahmen kann - mit gedrückter linker Maustaste - verschoben werden, damit man in der rechten Fensterhälfte mehr sieht. Hier ein verkleinertes Beispiel eines Ausgabe-Viewers:

The screenshot shows the SPSS Output Viewer window with three labels pointing to specific parts: 'Hierarchie der Objekte (Gliederungsrahmen)' points to the left-hand tree view, 'Rahmentrenner' points to the vertical separator between the tree and the main content, and 'Ergebnis-Darstellung' points to the main content area.

The main content area displays the following statistics for 'Geschlecht':

| Statistiken |     |
|-------------|-----|
| Geschlecht  |     |
| N           | 442 |
| Gültig      | 442 |
| Fehlend     | 0   |

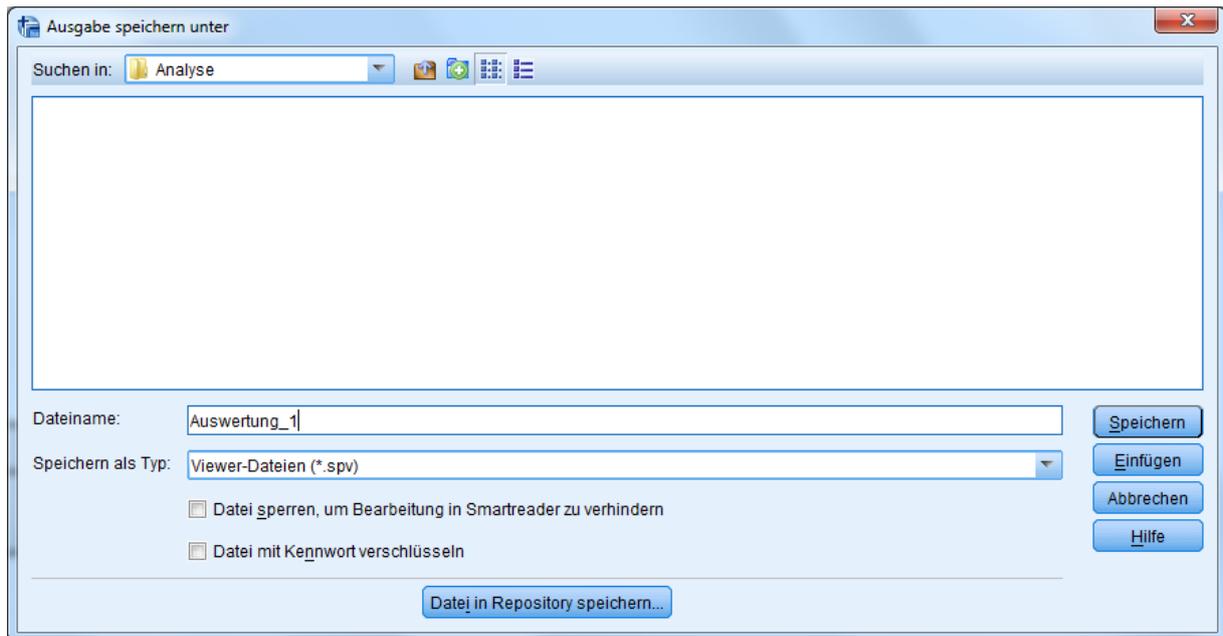
  

|             |          | Geschlecht |         |                 |                     |
|-------------|----------|------------|---------|-----------------|---------------------|
|             |          | Häufigkeit | Prozent | Gültige Prozent | Kumulative Prozente |
| Gültig      | männlich | 195        | 44,1    | 44,1            | 44,1                |
|             | weiblich | 247        | 55,9    | 55,9            | 100,0               |
| Gesamtsumme |          | 442        | 100,0   | 100,0           |                     |

## 1.2.2. Bedienung des Ausgabe-Viewers

### 1.2.2.1. Abspeichern des Inhaltes des Ausgabe-Viewers

Um den gesamten Inhalt des Ausgabe-Fensters speichern zu können, muss man sich im Ausgabe-Fenster befinden. Betätigen Sie die Befehlsfolge **Datei, Speichern unter** ★ Dialogfeld *Ausgabe Speichern unter*. In diesem Dialogfeld müssen Sie nur den verlangten Dateinamen (hier: user10) im Feld **Dateiname** eintragen und dann die Befehls-Schaltfläche **Speichern** betätigen. Hier der Zustand des Dialogfeldes unmittelbar vor dem Kommando **Speichern** :



#### HINWEISE:

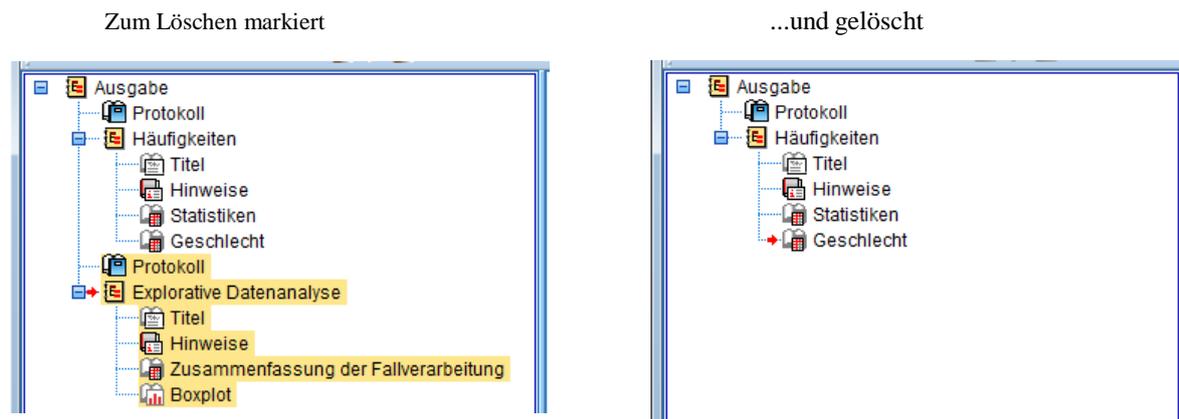
- Der Datei-Typ ist *ViewerDatei* mit der Erweiterung **.spv**. In diesem Feld **Dateityp** sollte nichts verändert werden.
- Der unter Suchen in: abgebildete Name des Ordners bezieht sich auf ein Verzeichnis auf der Festplatte des Autorensrechners und ist nur als Beispiel zu verstehen. Alle SPSS-Nutzer sollten für Ihre Auswertung einen speziellen Ordner erstellen, der evtl. auch mit weiteren Unterordnern versehen werden kann.

### 1.2.2.2. Löschung von überflüssigen Objekten im Ausgabe-Viewer

Häufig kommt es vor, dass man mehr Objekte, als für die Lösung der jeweiligen Aufgabe notwendig ist, im Ausgabe-Fenster stehen hat. Der Grund dafür kann etwa sein, dass man einen Auswertungsbefehl doppelt gegeben oder die falsche Variable benutzt hat oder etwas ausprobierte, das nicht zum gewünschten Ergebnis führte. In diesen Fällen sollten alle überflüssigen Objekte aus dem Ausgabe-Fenster gelöscht werden, bevor der gesamte Fenster-Inhalt als Datei gespeichert wird.

Unter der Annahme, dass das Objekt "Explorative Datenanalyse" in dem Ausgabe-Fenster überflüssig sei, hier ein Beispiel für die Verfahrensweise beim Löschen einzelner Objekte aus dem Ergebnis-Fenster: Man markiere die zu löschenden Objekte in der Hierarchiedarstellung der linken Fensterseite. Sodann ist die Befehlsfolge **Bearbeiten, Löschen** zu betätigen ★ die markierten Objekte sind aus dem Ausgabe-Fenster gelöscht.

Hier eine kleine exemplarische Darstellung:



Wenn die Ausgabe den erforderlichen Umfang aufweist, kann sie wie unter 1.2.2.1 mit dem vorgesehenen Namen abgespeichert werden.

### 1.2.2.3. Bearbeiten von Grafiken

Diagramme, also Grafiken zur Veranschaulichung von Verteilungen und Verläufen, werden ebenfalls als Objekte im Ausgabe-Viewer dargestellt. SPSS hat für jede Grafikform einen eigenen Darstellungs-Standard. Wenn man damit nicht zufrieden ist - seine Grafik etwa noch mit einem erklärenden Titel versehen oder Farben ändern will - so muss diese Grafik nochmal "von Hand" nachbearbeitet werden, dies natürlich ebenfalls im SPSS-Viewer. Dafür ist mitten in die Grafik ein Doppelklick mit linker Maustaste durchzuführen. Es öffnet sich der SPSS-Diagramm-Editor, welcher seinerseits in einem eigenen Fenster ausgeführt wird. Auch das Diagramm-Editor-Fenster besitzt in seiner Menüleiste eigene, auf die Zwecke der Grafik-Bearbeitung ausgerichtete Befehle. Eine SPSS-Grafik besteht aus einzelnen Objekten - Boxen, Linien, Achsen, Achsenbeschriftungsfeldern etc. Vor seiner Bearbeitung ist das jeweilige Objekt mit Einfachklick links zu markieren. Hier die wichtigsten Befehlsfolgen zur Bearbeitung markierter Objekte:

#### Bearbeiten, Eigenschaften

Öffnet das Dialogfeld *Eigenschaften* des jeweils markierten Objekts

#### Optionen

Ermöglicht das Einfügen von

- Bezugslinien für X - / Y - Achse
- Textfeldern
- Rastern (Gitternetzlinien)

Zur zügigen Bearbeitung von markierten Objekten im Diagramm-Fenster kann auch die Benutzung des Kontext-Menüs mit der rechten Maustaste empfohlen werden.

Einige Modifikationsbefehle für Diagramme sind auch über die Symbolleiste des Diagramm-Editors - markierungsabhängig - erreichbar:



### 1.3. Daten-Erweiterung und Daten-Auswahl

#### 1.3.1. Eine neue Variable erzeugen, deren Ausprägungen eingetragen werden müssen

Bei der Erstellung einer Daten-Tabelle erfolgt zunächst die Variablen-Definition in der Variablenansicht des Daten-Editors. Danach erfolgt die Dateneingabe der Datenansicht Daten-Editors. Es ist aber darüber hinaus nützlich zu wissen, wie bei Bedarf nachträglich eine neue Variable hinzugefügt werden kann.

Die Erzeugung einer neuen Variablen wird in der Tabelle Variablenansicht des Dateneditors vollzogen. Die Angaben zur neuen Variablen müssen in dieser Tabelle in eine eigene Zeile - denn jede Variable belegt in der Variablenansicht eine eigene Zeile - eingetragen werden. In dieser Zeile wären auch bei Bedarf in den entsprechenden Zellen die Anzahl der Dezimalstellen wie auch Variablenbezeichnung (verständlicher Klurname der Variablen) und Wertebezeichnungen (Benennung der Ausprägungen einer kategorialen Variablen) zu vereinbaren.

Für die Positionierung der neuen Variablen gibt es zwei Möglichkeiten: 1) Sie wird in die nächste freie Zeile der Variablenliste eingetragen, also unten angeschlossen. In der Datenansicht erscheint sie dann in der letzten Spalte. 2) Sie wird, weil von der Reihenfolge in der Datenansicht her vielleicht praktischer, an geeigneter Stelle in die Liste eingefügt. Dafür ist der Zeilenkopf derjenigen Variablen mit rechts anzuklicken, vor welcher die neue Variable eingefügt werden soll, gefolgt vom Kommando **Variable einfügen**. In der Datenansicht sind dann die Ausprägungen in die Spalte der neuen Variablen Zelle für Zelle einzutragen.

#### 1.3.2. Neue Variable aus bestehenden Werten errechnen lassen: Kategorisieren einer stetigen Variablen

Hier verwendete Datenmenge: SPORT\_LMK.SAV

Die Befehlsfolge **Transformieren, Umkodieren in andere Variablen** führt in das Dialogfeld *Umkodieren in andere Variablen*. Aus der Variablenliste ist die Eingabevariable auszuwählen und mit der Schaltfläche **Übernehmen** in das Feld **Numerische Var.-> Ausgabevar.-> Ausgabevar.** einzusteuern:



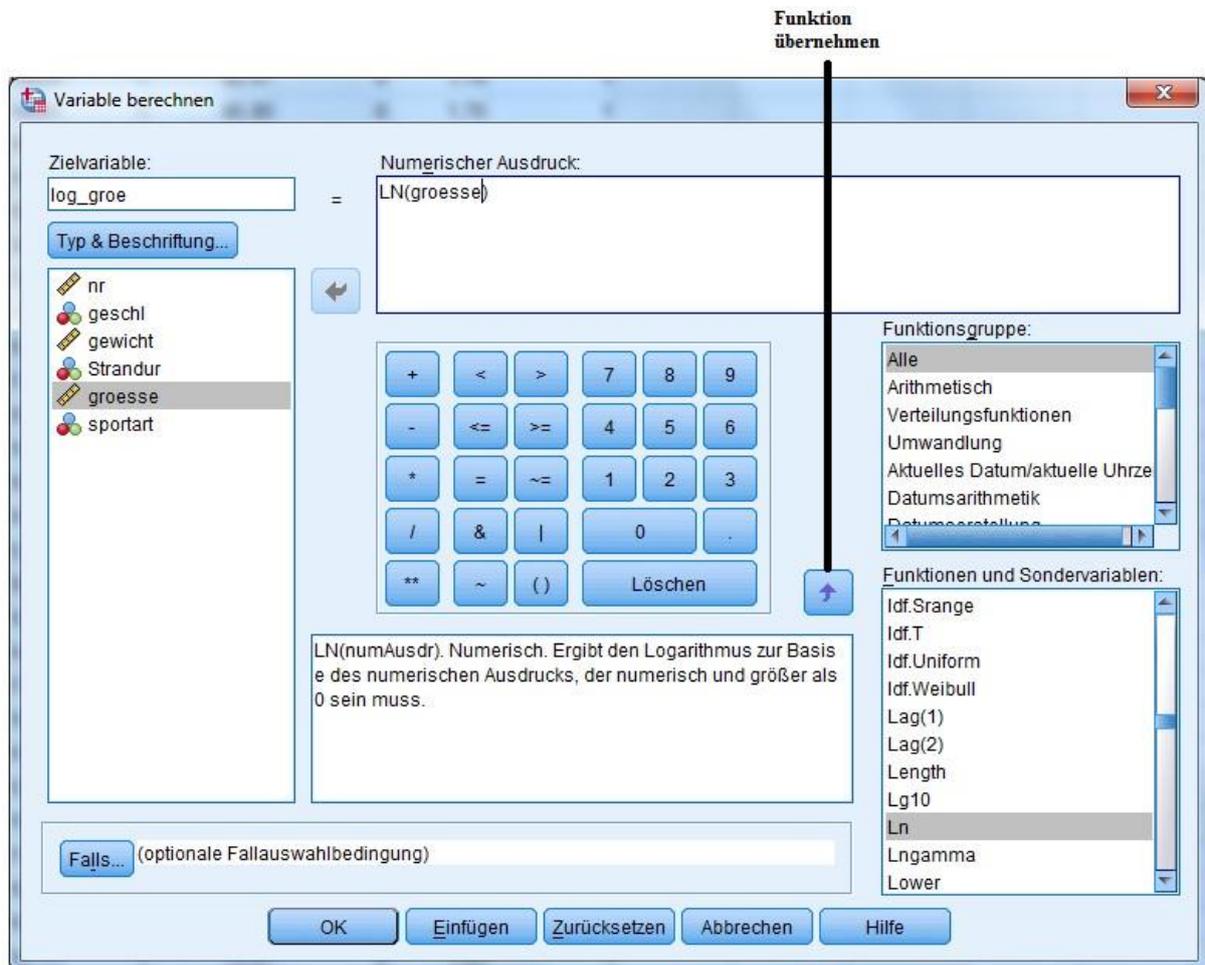
Nach der Eingabe des neuen Variablennamens muss unter **Alte und neue Werte** die stetige Quellvariable kategorial umkodiert werden. **Weiter** führt zurück in den Dialog *Umkodieren in andere Variablen*. Hier muss die Umkodierung durch **Ändern** abgeschlossen werden. Mit **OK** wird die neue Variable Bestandteil der Datentabelle.

### 1.3.3 Eine neue Variable aus einer alten Variablen berechnen

Hier verwendete Datenmenge: SPORT\_LMK.SAV  
Hier verwendete Variable: GROESSE

Beispiel: logarithmische Transformation

Die Befehlsfolge **Transformieren, Variable berechnen** ★ Dialogfeld *Variable berechnen*. In das Feld **Zielvariable** ist der Name der neuen Variablen einzutragen: log\_groe. Im Feld Funktionsgruppe **Alle** aktivieren; im Feld **Funktionen** muss die Option **LN** markiert und mittels des Schalters **Übernehmen** in das Feld **Numerischer Ausdruck** eingesteuert werden. Anstelle des Fragezeichens zwischen den Klammern muss der Originalname der Quell-Variablen eingetragen werden, wie unten ersichtlich:



Wenn das Dialogfeld die obige Gestalt besitzt, kann die Transformation mit **OK** ausgeführt werden. SPSS erzeugt nun eine neue Variable des Namens log\_groe für den logarithmierten Inhalt der Quell-Variablen groesse.

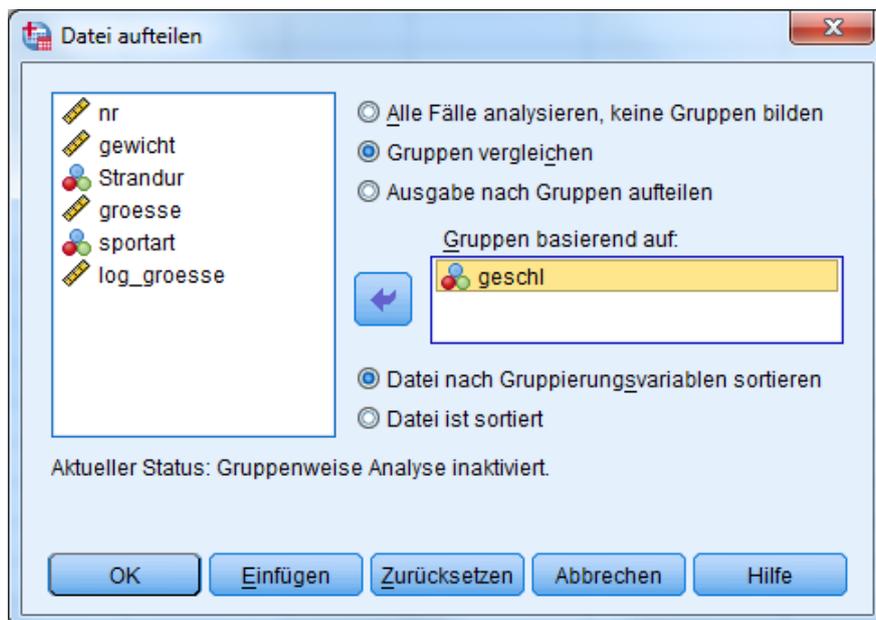
|   | nr   | geschl | gewicht | Strandur | groesse | sportart | log_groe |
|---|------|--------|---------|----------|---------|----------|----------|
| 1 | 1190 | 0      | 80,44   | 0        | 1,88    | 1        | ,63      |
| 2 | 7041 | 0      | 84,36   | 0        | 1,92    | 1        | ,65      |
| 3 | 6209 | 0      | 84,15   | 0        | 1,93    | 1        | ,66      |

### 1.3.4. Datei aufteilen

Hier verwendete Datenmenge: SPORT\_LMK.SAV

In dieser Datenmenge wurden sowohl Sportlerinnen als auch Sportler erfasst. Das jeweilige Geschlecht der Probanden wurde im Merkmal *geschl* aufgenommen, das jeweilige Gewicht im Merkmal *gewicht*. Um z. B. die Quartile für das Merkmal *gewicht* getrennt nach dem Geschlecht der ProbandInnen ausgeben zu lassen, wird die Datenmenge in SPSS in zwei Gruppen aufgeteilt, und zwar gemäß den Ausprägungen im Merkmal *geschl* (0 oder 1).

Befehlsfolge: **Daten, Aufgeteilte Datei** ★ Dialogfeld *Datei aufteilen*. Dort ist die Option **Ausgabe nach Gruppen aufteilen** zu aktivieren und die Gruppierungs-Variable, also *geschl*, in der Variablenliste zu markieren. Durch Klick auf die Schaltfläche **Übernehmen** wird die Variable in das Feld **Gruppen basierend auf** aufgenommen:



Nach **OK** wird jede nun durchgeführte Auswertung nach Gruppen getrennt durchgeführt, so auch z. B. die Berechnung der Quartile für die Variable *gewicht*:

#### Statistiken

| Gewicht  |            |         |         |
|----------|------------|---------|---------|
| männlich | N          | Gültig  | 195     |
|          |            | Fehlend | 0       |
|          | Perzentile | 25      | 84,3617 |
|          |            | 50      | 90,4219 |
| 75       |            | 95,7461 |         |
| weiblich | N          | Gültig  | 247     |
|          |            | Fehlend | 0       |
|          | Perzentile | 25      | 66,7276 |
|          |            | 50      | 70,8234 |
| 75       |            | 78,5279 |         |

HINWEIS: die Datei-Aufteilung nach Gruppen kann rückgängig gemacht werden mit der Befehlsfolge **Daten, Aufgeteilte Datei, Alle Fälle analysieren, keine Gruppen bilden, OK**. Ansonsten wird die Datei-Aufteilung beibehalten!

### 1.3.5. Fälle auswählen: ein Subkollektiv definieren

Hier verwendete Datenmenge: SPORT\_LMK.SAV  
Hier verwendetes Subkollektiv: SPORTART=2 & geschl =1

Will man nur Statistiken für eine bestimmte Untergruppe der gesamten Datenmenge berechnet haben - in diesem Beispiel sollen nur die Mannschaftssportlerinnen bei der Auswertung berücksichtigt werden - so kann das entsprechende Dialogfeld *Fälle auswählen* mit der Befehlsfolge **Daten, Fälle auswählen** aufgerufen werden. Dasselbst ist die Option **Falls Bedingung zutrifft** und daraufhin die Schaltfläche **Falls** zu betätigen. Es öffnet sich das untergeordnete Dialogfeld *Fälle auswählen: Falls*, das mit etwas Übung ausschließlich mit der Maus bedient werden kann. Die erforderlichen Einträge können aber auch getippt werden. Wenn, wie im vorliegenden Falle, mehr als eine Variable mit einschränkender Bedingung ausgewählt werden muss, so sind diese Bedingungen mit dem Zeichen **&** bzw. "and" zu verketten. Wie auch immer für die Auswahl des Subkollektivs der Mannschaftssportlerinnen vorgegangen wird, ob mit Mausclick auf die Schaltflächen oder durch Eintragung der Bedingungen über die Tastatur, zuletzt muss das Dialogfeld die folgende Gestalt besitzen:



Nach Betätigung von **Weiter** gelangt man wieder in das übergeordnete Dialogfeld *Fälle auswählen*, wo die Auswahl mit **OK** zu bestätigen ist.

|                | nr              | geschl       | gewicht          | Strandur     | groesse         | sportart     | log_groe       |  |
|----------------|-----------------|--------------|------------------|--------------|-----------------|--------------|----------------|--|
| 335            | 10755           | 1            | 83,78            | 1            | 1,90            | 2            | ,64            |  |
| 336            | 10420           | 1            | 87,21            | 1            | 1,90            | 2            | ,64            |  |
| <del>337</del> | <del>6538</del> | <del>0</del> | <del>80,10</del> | <del>0</del> | <del>1,73</del> | <del>3</del> | <del>,55</del> |  |

SPSS hat nun eine dichotome "Filter"-Variable namens *filter\_\$* erzeugt, die nur die Ausprägungen 1 (ausgewählt) und 0 (nicht ausgewählt) beinhaltet. -

**HINWEIS:** Die Befehlsfolge **Daten, Fälle auswählen, Alle Fälle, [OK]** hebt die getroffene Auswahl wieder auf.

## 2. Spezielle Auswertungen mit SPSS

### 2.1. Bestimmung absoluter und relativer Häufigkeiten

Hier verwendete Datenmenge: VERDINUM.SAV  
 Hier verwendete Variablen: THER("Behandlung"),  
 RESP8 ("Therapie-Erfolg nach 8 Wochen")

Die Befehlsfolge **Analysieren, Deskriptive Statistiken, Häufigkeiten** führt in das Dialogfeld *Häufigkeiten*. Dort sind in das Feld **Variablen** die beiden interessierenden Merkmale einzutragen. Mit **OK** wird folgender Output erzeugt:

#### Häufigkeiten

Statistiken

|   |         | Behandlung | Therapieerfolg nach 8 Wochen |
|---|---------|------------|------------------------------|
| N | Gültig  | 369        | 353                          |
|   | Fehlend | 0          | 16                           |

#### Häufigkeitstabelle

Behandlung

|        |                    | Häufigkeit | Prozent | Gültige Prozent | Kumulative Prozente |
|--------|--------------------|------------|---------|-----------------|---------------------|
| Gültig | Hydrochlorothiazid | 187        | 50,7    | 50,7            | 50,7                |
|        | Verapamil          | 182        | 49,3    | 49,3            | 100,0               |
|        | Gesamtsumme        | 369        | 100,0   | 100,0           |                     |

Therapieerfolg nach 8 Wochen

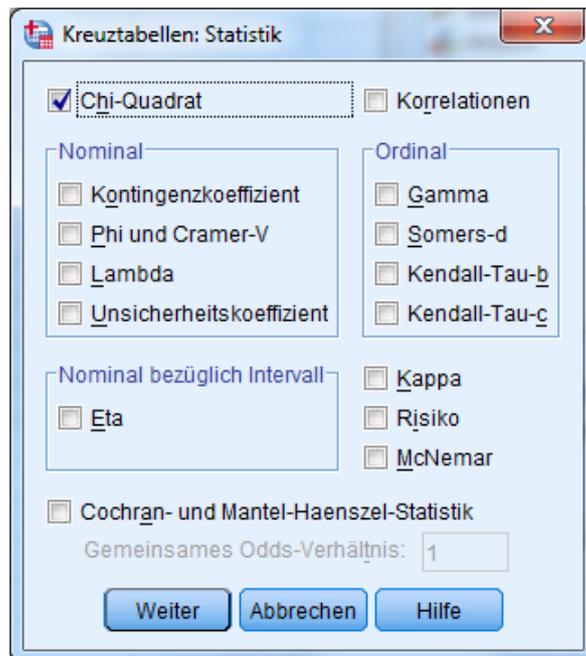
|             |             | Häufigkeit | Prozent | Gültige Prozent | Kumulative Prozente |
|-------------|-------------|------------|---------|-----------------|---------------------|
| Gültig      | nein        | 176        | 47,7    | 49,9            | 49,9                |
|             | ja          | 177        | 48,0    | 50,1            | 100,0               |
|             | Gesamtsumme | 353        | 95,7    | 100,0           |                     |
| Fehlend     | System      | 16         | 4,3     |                 |                     |
| Gesamtsumme |             | 369        | 100,0   |                 |                     |

Häufigkeit: absolute Häufigkeiten  
 Prozent: relative Häufigkeiten, wenn Fehlwerte nicht ausgeschlossen werden  
 Gültige Prozente: relative Häufigkeiten, wenn Fehlwerte ausgeschlossen werden

### 2.2. Vierfeldertafeln, Exakter Fisher-Test, Chi-Quadrat-Unabhängigkeitstest

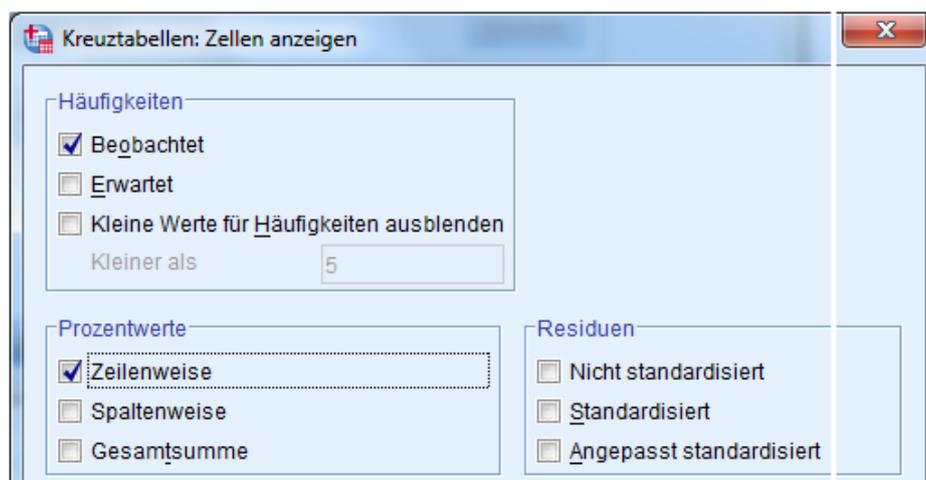
Hier verwendete Datenmenge: VERDINUM.SAV  
 Hier verwendetes Subkollektiv: alter < 45 Jahre  
 Hier verwendete Variable: Einflussgröße THER ("Therapie"),  
 Zielgröße RESP8 ("Therapie-Erfolg nach 8 Wochen").

Zur Durchführung des exakten Fisher-Tests ist zunächst folgende Befehlsfolge abzusetzen: **Analysieren, Deskriptive Statistiken, Kreuztabellen** ★ Dialogfeld *Kreuztabellen*. Dort ist die Einflussgröße *ther* als Zeilenvariable in das Feld **Zeilen**: und die Zielgröße *resp8* als Spaltenvariable in das Feld **Spalten**: einzutragen. - Zur Ausgabe des exakten Fisher-Tests ist an dieser Stelle die Schaltfläche **Statistiken** anzuklicken ★ Dialogfeld *Kreuztabellen: Statistik*. Dort ist die Optionen **Chi-Quadrat** zu aktivieren. Falls sinnvoll, können an dieser Stelle auch noch das Relative Risiko oder das Kappa-Maß oder der McNemar-Test aufgerufen werden:



Danach ist die Schaltfläche **Weiter** zu betätigen:

Im übergeordneten Dialogfeld *Kreuztabellen* muss noch die Schaltfläche **Zellen** angeklickt werden ★ Dialogfeld *Kreuztabellen: Zellen anzeigen*. Hier wird die Ausgabe von Zeilenprozenten angefordert durch Aktivierung der Option **Prozentwerte: Zeilenweise**. Wenn das Dialogfeld die folgende Gestalt besitzt, kann **Weiter** gegeben werden:



Zurück im übergeordneten Dialogfeld *Kreuztabellen* kann nun, da alle erforderlichen Einstellungen vollzogen sind, **OK** gegeben werden. Sodann erscheint dieser Output:

### Verarbeitete Fälle

|   | Fälle  |         |         |         |        |         |
|---|--------|---------|---------|---------|--------|---------|
|   | Gültig |         | Fehlend |         | Gesamt |         |
|   | N      | Prozent | N       | Prozent | N      | Prozent |
| Behandlung * Therapieerfolg nach 8 Wochen | 100    | 96,2%   | 4       | 3,8%    | 104    | 100,0%  |

Erläuterung zur Tabelle *Verarbeitete Fälle*: Bei 4 Von 104 Patienten konnte keine Untersuchung der Response nach 8 Wochen durchgeführt werden. Somit ergibt sich eine effektive Fallzahl von 100 Patienten, für die der exakte Fisher-Test und der Risikoschätzer berechnet werden können.

### Behandlung \* Therapieerfolg nach 8 Wochen Kreuztabelle

|            |                    |                  | Therapieerfolg nach 8 Wochen |       | Gesamt |
|------------|--------------------|------------------|------------------------------|-------|--------|
|            |                    |                  | nein                         | ja    |        |
| Behandlung | Hydrochlorothiazid | Anzahl           | 36                           | 19    | 55     |
|            |                    | % von Behandlung | 65,5%                        | 34,5% | 100,0% |
|            | Verapamil          | Anzahl           | 23                           | 22    | 45     |
|            |                    | % von Behandlung | 51,1%                        | 48,9% | 100,0% |
| Gesamt     |                    | Anzahl           | 59                           | 41    | 100    |
|            |                    | % von Behandlung | 59,0%                        | 41,0% | 100,0% |

Erläuterung zur Kreuztabelle: hier sind zwei Häufigkeitsarten zu finden: die absolute Häufigkeit in den Zeilen "Anzahl" und die relative Häufigkeit in Prozent in den Zeilen "% von Behandlung". Von 55 Patienten, die mit Hydrochlorothiazid behandelt wurden, hatten also 19 (35%) eine Response nach 8 Wochen, verglichen mit 22 von 45 (49%) der mit Verapamil behandelten Patienten.

### Chi-Quadrat-Tests

|                                    | Wert               | df | Asymptotische Signifikanz (2-seitig) | Exakte Signifikanz (2-seitig) | Exakte Signifikanz (1-seitig) |
|------------------------------------|--------------------|----|--------------------------------------|-------------------------------|-------------------------------|
| Chi-Quadrat nach Pearson           | 2,105 <sup>b</sup> | 1  | ,147                                 |                               |                               |
| Kontinuitätskorrektur <sup>a</sup> | 1,554              | 1  | ,213                                 |                               |                               |
| Likelihood-Quotient                | 2,106              | 1  | ,147                                 |                               |                               |
| Exakter Test nach Fisher           |                    |    |                                      | ,159                          | ,106                          |
| Zusammenhang linear-mit-linear     | 2,084              | 1  | ,149                                 |                               |                               |
| Anzahl der gültigen Fälle          | 100                |    |                                      |                               |                               |

a. Wird nur für eine 2x2-Tabelle berechnet

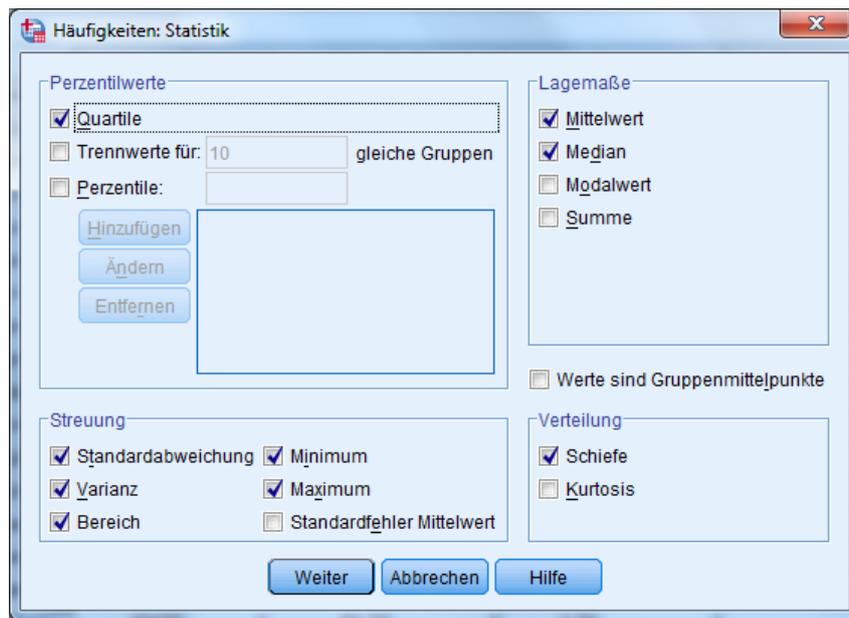
b. 0 Zellen (,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 18,45.

Erläuterung zur Tabelle *Chi-Quadrat-Tests*: In diesem Zusammenhang ist die Zeile *exakter Test nach Fisher* wichtig. In der Spalte *Exakte Signifikanz (2-seitig)* ist der zweiseitige p-Wert mit 0,159 größer als das vorgegebene Signifikanzniveau von 0,05. Damit kann man auf keinen Unterschied der Responderraten schließen.

### 2.3. Statistische Maßzahlen für stetige Variablen

Hier verwendete Datenmenge: SPORT\_LMK  
 Hier verwendetes Subkollektiv: SPORTART=2 & geschl =1  
 Hier verwendete Variable: GROESSE (Körpergröße in Meter)

Zunächst ist die Befehlsfolge **Analysieren, Deskriptive Statistiken, Häufigkeiten** zu aktivieren. Im daraufhin erscheinenden Dialogfeld *Häufigkeiten* ist die interessierende Variable in der Liste auszuwählen und in das Feld **Variable(n)** zu übernehmen. Aus Platzgründen soll keine Häufigkeitstabelle ausgegeben werden, deshalb ist die Option **Häufigkeitstabellen anzeigen** mit Mausclick zu deaktivieren. Die Schaltfläche **Statistik** führt in folgendes Auswahl-Dialogfeld, wo die gewünschten Maßzahlen wie folgend dargestellt zu markieren sind:



Nach Betätigung von **Weiter** gelangt man wieder in das übergeordnete Dialogfeld *Häufigkeiten*, wo entweder über die Schaltfläche **Diagramme** eine graphische Darstellung oder aber, falls dies nicht gewünscht ist, durch Betätigung von **OK**, die Ausgabe der Maßzahlen an- gefordert werden kann:

| Statistiken                |         |        |
|----------------------------|---------|--------|
| Körpergröße                |         |        |
| N                          | Gültig  | 120    |
|                            | Fehlend | 0      |
| Mittelwert                 |         | 1,7726 |
| Median                     |         | 1,7845 |
| Standardabweichung         |         | ,08177 |
| Varianz                    |         | ,007   |
| Schiefe                    |         | -,138  |
| Standardfehler der Schiefe |         | ,221   |
| Bereich                    |         | ,27    |
| Minimum                    |         | 1,63   |
| Maximum                    |         | 1,90   |
| Perzentile                 | 25      | 1,6994 |
|                            | 50      | 1,7845 |
|                            | 75      | 1,8455 |

## 2.4. Boxplot

Hier verwendete Datenmenge:

SPORT\_LMK.SAV

Hier verwendete Variablen:

GROESSE (Körpergröße in Meter),  
GESCHL (Geschlecht)

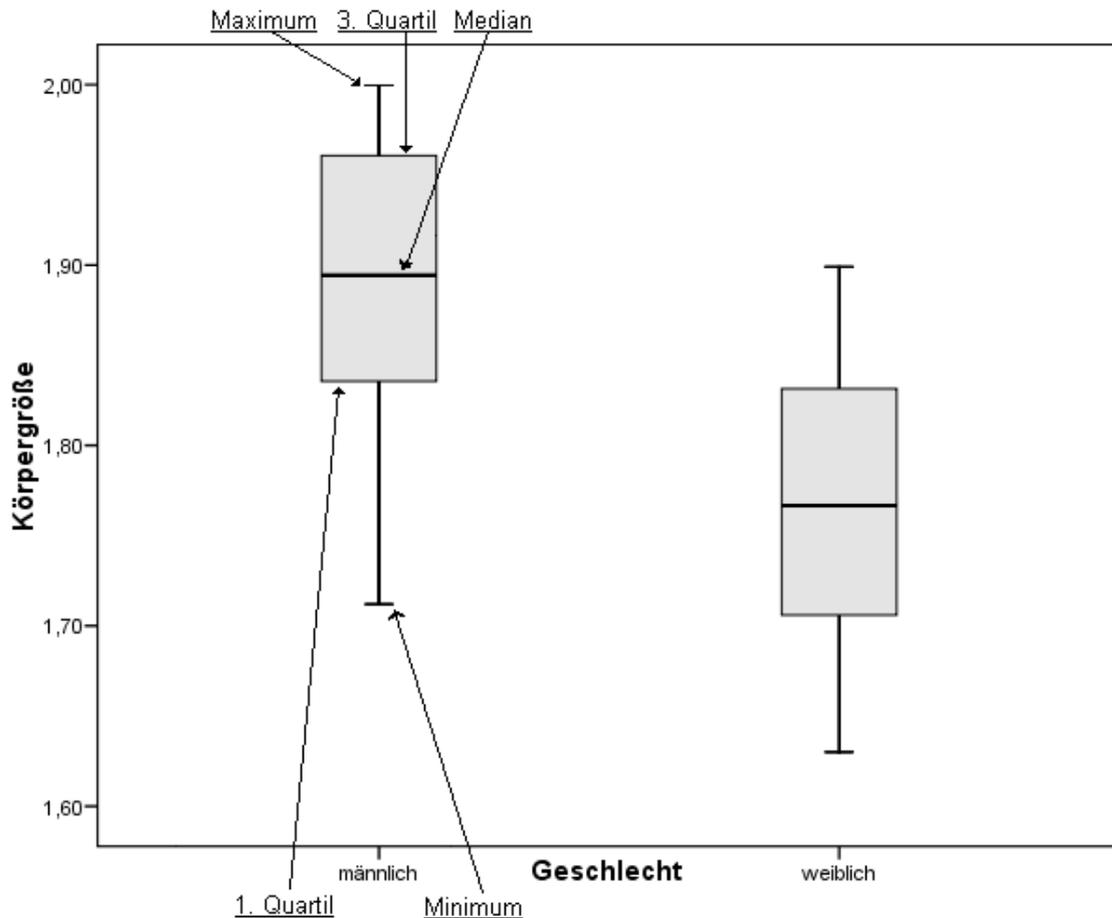
Zur optischen Visualisierung der Verteilung einer Variablen kann auch in diesem Zusammenhang der Grafiktyp Boxplot beitragen. Eine solche Grafik wird wie folgt erzeugt: Befehls- folge **Grafik, Veraltete Dialogfelder, Boxplot** ★ es öffnet sich das Dialogfeld *Boxplot*:



Man möchte nun die Verteilung einer stetigen Variablen aufgeteilt nach dem Inhalt einer qualitativ-nominalen Variablen kennenlernen: **Einfach, Definieren** ★ Dialogfeld *Einfachen Boxplot definieren: Auswertung über Kategorien einer Variablen* (Ausschnitt):



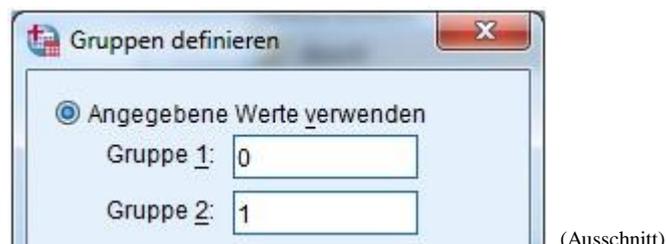
Als aufzutragende Variable wurde in diesem Beispiel *groesse* und als Einteilungsvariable *geschl* gewählt. Auf **OK** hin erzeugt SPSS die Grafik, welche noch bearbeitet werden kann (Farbe, Beschriftung, evtl. Skalierung):



## 2.5. Der t-Test für unverbundene Stichproben

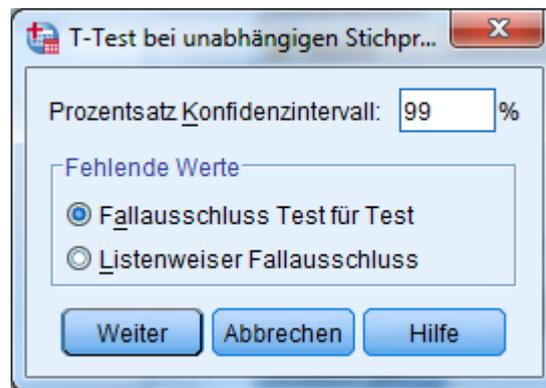
|                                |  |
|--------------------------------|--|
| Hier verwendete Datenmenge:    | VERDINUM.SAV   |
| Hier verwendete Subkollektive: | Behandlung mit HCT oder mit VER                          |
| Hier verwendete Variable:      | DIAST0 ("Diastolischer Blutdruck bei Behandlungsbeginn") |

Zum Mittelwertvergleich mit Hilfe des t-Tests für unverbundene Stichproben ist folgendermaßen vorzugehen: **Analysieren, Mittelwerte vergleichen, T-Test bei unabhängigen Stichproben** ★ Dialogfeld *T-Test bei unabhängigen Stichproben*. In das Feld **Testvariable(n)** ist die interessierende Variable, hier *diast0*, einzutragen. In das Feld **Gruppenvariable** muss diejenige Variable eingetragen werden, in welcher die Medikation für jeden Patienten vermerkt ist: *ther*. SPSS verlangt hier nochmals die Angabe der Ausprägungen in der Gruppenvariable. Deshalb ist an dieser Stelle die Schaltfläche **Gruppen definieren** zu betätigen. Die beiden Gruppen wären folgendermaßen festzulegen:



Dann ist die Schaltfläche **Weiter** zu betätigen. Im übergeordneten Dialogfeld *T-Test bei unabhängigen Stichproben* kann, falls dies gewünscht wird, noch mit der Schaltfläche **Optionen**

ein Dialogfeld aufgerufen werden, das die Veränderung des Konfidenzintervalls z. B. von 95% auf 99% erlaubt (nur bei Bedarf umstellen, ansonsten bei 95% belassen):



Nach **Weiter** und **OK** im übergeordneten Dialogfeld *T-Test bei unabhängigen Stichproben* erzeugt SPSS folgenden Output:

**Gruppenstatistik**

| Behandlung                                | H                  | Mittelwert | Standardabweichung | Standardfehler Mittelwert |
|---|--------------------|------------|--------------------|---------------------------|
| Diast. Blutdruck bei<br>Behandlungsbeginn | Hydrochlorothiazid | 187        | 104,43             | 6,364                     |
|   | Verapamil          | 182        | 104,25             | 6,079                     |

Erläuterungen zur Tabelle *Gruppenstatistiken*: Sie enthält für die Variable diast0 getrennt für die Gruppen "Hydrochlorothiazid" und "Verapamil" die Werte für den Stichprobenumfang (H), für den arithmetischen Mittelwert, die empirische Standardabweichung und den Standardfehler des Mittelwertes.

**Test bei unabhängigen Stichproben**

| Diast. Blutdruck bei<br>Behandlungsbeginn | Levene-Test der<br>Varianzgleichheit |      | T-Test für die Mittelwertgleichheit |         |                 |                         |                             |   |        |
|---|--------------------------------------|------|-------------------------------------|---------|-----------------|-------------------------|-----------------------------|---|--------|
|   | F                                    | Sig. | t                                   | df      | Sig. (2-seitig) | Mittelwert<br>differenz | Standardfehler<br>differenz | 95% Konfidenzintervall der<br>Differenz |        |
|   |                                      |      |                                     |         |                 |                         |                             | Unterer                                 | Oberer |
| Varianzgleichheit<br>angenommen           | ,730                                 | ,393 | ,287                                | 367     | ,774            | ,186                    | ,648                        | -1,089                                  | 1,461  |
| Varianzgleichheit nicht<br>angenommen     |                                      |      | ,287                                | 366,871 | ,774            | ,186                    | ,648                        | -1,088                                  | 1,460  |

Erläuterungen zur Tabelle *Test bei unabhängigen Stichproben*: In der ersten Zeile sind die Ergebnisse des klassischen t-Tests sowie ein Konfidenzintervall für die Mittelwertdifferenz dargestellt. In der zweiten Zeile findet man die Ergebnisse des modifizierten Tests nach Welch sowie ein modifiziertes Konfidenzintervall. Kann von einer Gleichheit der Varianzen ausgegangen werden, so verwendet man die Ergebnisse der ersten Zeile. Andernfalls sind die Angaben der zweiten Zeile zu entnehmen. Zur Entscheidung, welcher der beiden Tests geeigneter ist, kann man den Levène-Test verwenden.

|                  |                    |   |
|------------------|--------------------|---|
| <u>Begriffe:</u> | F                  | Prüfgröße des Levène-Tests auf Gleichheit der Varianzen in beiden Gruppen |
|                  | Signifikanz        | p-Wert des Levène-Tests   |
|                  | T                  | Prüfgröße des t-Tests   |
|                  | df                 | Freiheitsgrade  |
|                  | Sig. (2-seitig)    | zweiseitiger p-Wert des entsprechenden t-Tests                            |
|                  | Mittlere Differenz | Mittelwertdifferenz   |

## 2.6. Unverbundener Wilcoxon-Test (Mann-Whitney-U-Test)

Hier verwendete Datenmenge: VERDINUM.SAV  
 Hier verwendete Subkollektive: Behandlung mit Hydrochlorothiazid oder mit Verapamil  
 Hier verwendete Variable: DIAST0 (Diastolischer Blutdruck bei Behandlungsbeginn), THER (Behandlung)

Die Befehlsfolge **Analysieren, Nichtparametrische Tests, Alte Dialogfelder, Zwei unabhängige Stichproben** führt in das Dialogfeld *Tests bei zwei unabhängigen Stichproben*. Die interessierende Variable, hier *diast0*, ist in das Feld **Testvariablen** und die Gruppierungsvariable, hier *ther*, in das Feld **Gruppenvariable** einzusteuern. Nunmehr verlangt SPSS die Angabe der Ausprägungen für die beiden Subgruppen im Dialogfeld *Gruppen definieren* (wie beim t-Test). Nach Betätigung von **OK** im übergeordneten Dialogfeld *Tests bei zwei unabhängigen Stichproben* erscheint folgender Output:

**Ränge**

|  | Behandlung         | H   | Mittlerer Rang | Summe der Ränge |
|--|--------------------|-----|----------------|-----------------|
| Diast. Blutdruck bei Behandlungsbeginn | Hydrochlorothiazid | 187 | 186,04         | 34789,00        |
|  | Verapamil          | 182 | 183,93         | 33476,00        |
|  | Gesamtsumme        | 369 |                |                 |

**Teststatistiken<sup>a</sup>**

|                        | Diast. Blutdruck bei Behandlungsbeginn |
|------------------------|--|
| Mann-Whitney-U-Test    | 16823,000                              |
| Wilcoxon-W             | 33476,000                              |
| U                      | -,190                                  |
| Asymp. Sig. (2-seitig) | ,850                                   |

a. Gruppierungsvariable: Behandlung

### Erläuterungen zum SPSS-Ausdruck:

|                                      |  |
|--------------------------------------|--|
| H                                    | Fallzahlen                             |
| Mittlerer Rang                       | Mittelwerte der Ränge                  |
| Mann-Whitney-U                       | Prüfgröße des Mann-Whitney-Tests       |
| Wilcoxon-W                           | Prüfgröße des Wilcoxon-Tests           |
| Z                                    | Standardisierte Prüfgröße              |
| Asymptotische Signifikanz (2-seitig) | Zweiseitiger p-Wert des Wilcoxon-Tests |

## 2.7. t-Test für verbundene Stichproben

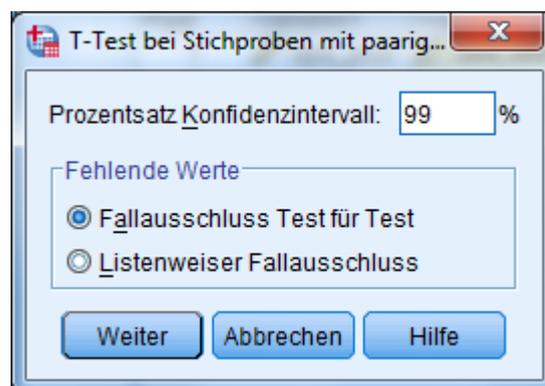
Hier verwendete Datenmenge: HFD.SAV  
 Hier verwendete Variablen: SPA\_PROX (Knochendichte am proximalen Messort), SPA\_DIST

(Knochendichte am distalen Messort)

Um einen t-Test für verbundene ("gepaarte") Stichproben durchzuführen, ist folgende Befehlssequenz abzusetzen: **Analysieren, Mittelwerte vergleichen, T-Test bei verbundenen Stichproben** ★ Dialogfeld *T-Test bei Stichproben mit paarigen Werten*. An dieser Stelle sind die beiden interessierenden Variablen in der Liste zu markieren und mittels der Schaltfläche **Übernehmen** in das Feld **Paarige Variablen** einzusteuern:



Falls man beim t-Test für verbundene Stichproben das in SPSS hierfür voreingestellte Konfidenzintervall von 95% z. B. in 99% umändern will, so ist die Schaltfläche **Optionen** zu betätigen ★ Dialogfeld *T-Test bei Stichproben mit paarigen Werten: Optionen*, wo im Feld **Prozentsatz Konfidenzintervall** die Zahl 99 eingetragen werden muss (nur bei Bedarf, ansonsten bei 95% belassen) :



Mit dem Befehl **Weiter** gelangt man wieder in das übergeordnete Dialogfeld. Hier ist, um die Auswertung durchzuführen, **OK** zu betätigen, woraufhin SPSS folgenden Output erzeugt:

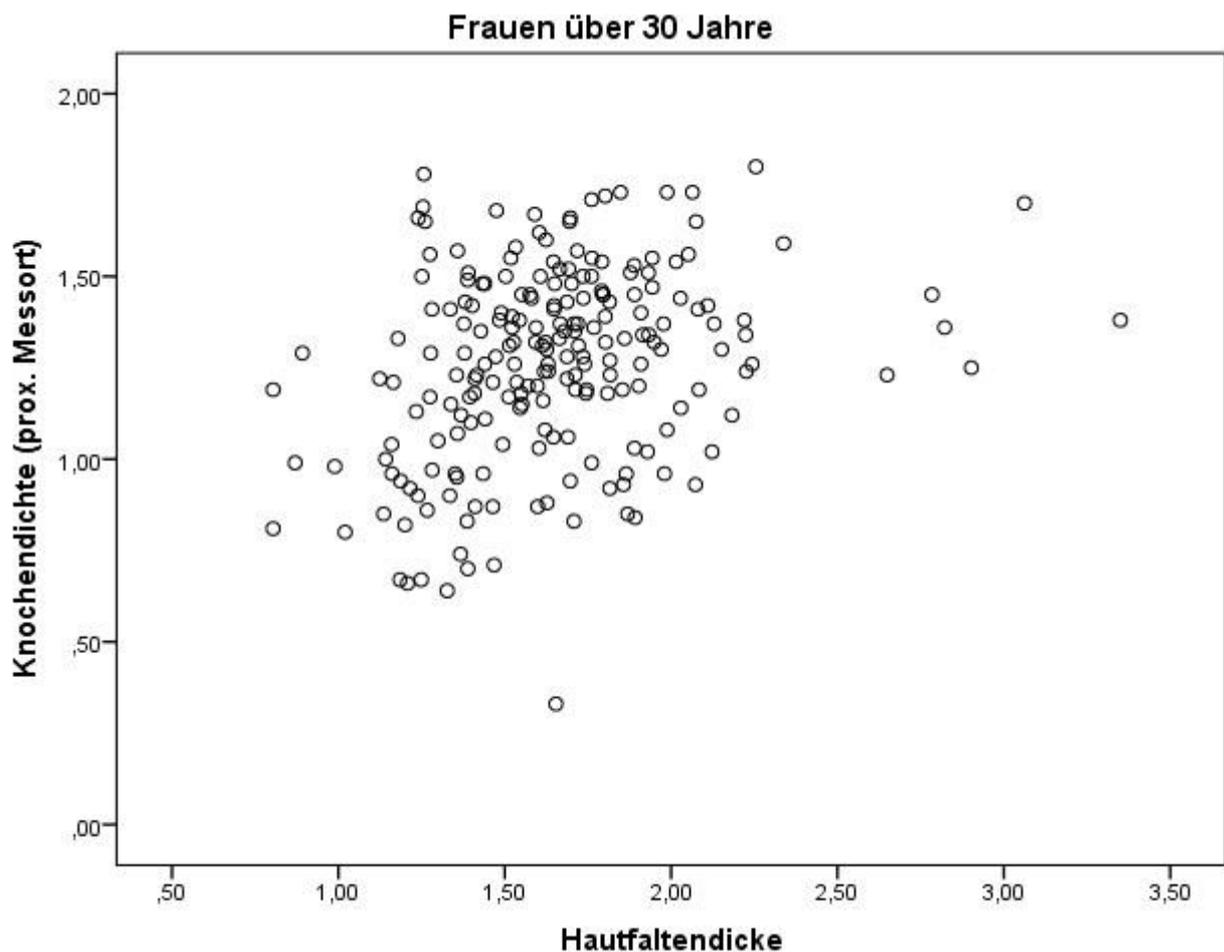


## 2.8. Streudiagramm (Scatterplot)

Hier verwendete Datenmenge: HFD.SAV  
Hier verwendete Variablen: SPA\_PROX, HFD  
Hier verwendetes Subkollektiv: GESCHL=2 & ALTER > 30

Bei einem Streu- oder auch Punktediagramm (Scatterplot) wird die gemeinsame Verteilung zweier stetiger Merkmale in einem Koordinatensystem dargestellt. Mit der Befehlsfolge **Grafik, Alte Dialogfelder, Streu-/Punktdiagramm, Einfaches S.** wird als grafische Darstellung ein Streudiagramm gewählt. Mittels **definieren** wird in diesem Beispiel die Variable *hfd* auf die x-Achse und die Variable *spa\_prox* auf die y-Achse gelegt. Mit dem Befehl **Titel** kann eine zweizeilige Diagramm-Überschrift eingegeben werden. Nach **OK** erscheint das Streudiagramm.

Nach Aufhellung des Hintergrundes durch Nachbearbeitung stellt sich die Grafik dar wie folgt:



Meist wird noch eine Regressionsgerade zur grafischen Darstellung hinzugefügt, um den Trend der Verteilung der Punkte zu visualisieren: **Elemente, Anpassungslinie bei Gesamtsumme**. Die oft nicht gewünschte, bei SPSS 22 aber automatisch in die Linie eingezeichnete Regressionsgleichung kann nach Markierung entfernt werden mit **Bearbeiten, Eigenschaften, Bezugslinie, Option Beschriftung zu Linie hinzufügen** deaktivieren, **Anwenden, Schließen**.

Vom Grafik-Editiermodus aus geht es zurück ins Ausgabe-Fenster mit **Datei, Schließen**.

## 2.9. Lineare Regression

Hier verwendete Variablen: SPA\_PROX (abhängig), HFD (unabhängig)  
 Hier verwendetes Subkollektiv: GESCHL = 2 & DIABETES = 0

Befehlsfolge: **Analysieren, Regression, Linear** ★ Dialogfeld *Lineare Regression*. Hier sind die abhängige und die unabhängige(n) Variable(n) in die entsprechenden Felder einzutragen. Für dieses Beispiel bleibt die vorgelegte Option **Einschluss** im Listenfeld **Methode** beibehalten. Nach **OK** wird folgende Ausgabe erzeugt:

**Eingegebene/Entfernte Variablen<sup>a</sup>**

| Modell | Eingegebene Variablen        | Entfernte Variablen | Methode   |
|--------|------------------------------|---------------------|-----------|
| 1      | Hautfaldendicke <sup>b</sup> |                     | Aufnehmen |

a. Abhängige Variable: Knochendichte (prox. Messort)

b. Alle angeforderten Variablen wurden eingegeben.

**Modellübersicht**

| Modell | R                 | R-Quadrat | Angepasstes R-Quadrat | Standardfehler der Schätzung |
|--------|-------------------|-----------|-----------------------|------------------------------|
| 1      | ,343 <sup>a</sup> | ,118      | ,113                  | ,24932                       |

a. Prädiktoren: (Konstante), Hautfaldendicke

[...]

**Koeffizienten<sup>a</sup>**

| Modell |                 | Nicht standardisierte Koeffizienten |                | Standardisierte Koeffizienten | t      | Sig. |
|--------|-----------------|-------------------------------------|----------------|-------------------------------|--------|------|
|        |                 | B                                   | Standardfehler | Beta                          |        |      |
| 1      | (Konstante)     | ,873                                | ,078           |                               | 11,122 | ,000 |
|        | Hautfaldendicke | ,239                                | ,045           | ,343                          | 5,264  | ,000 |

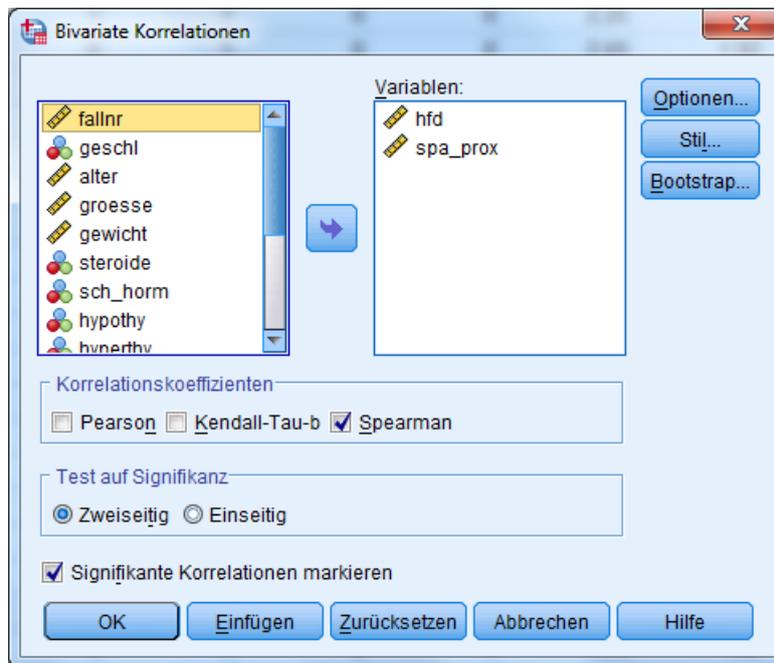
a. Abhängige Variable: Knochendichte (prox. Messort)

Unter "Modellübersicht" findet man den Korrelationskoeffizienten (R) und das lineare Bestimmtheitsmaß (R-Quadrat). Die Koeffizienten der Regressionsgeraden erscheinen im Abschnitt "Koeffizienten" in der Spalte B, zunächst der Achsenabschnitt  $a=0,873$ , dann der Regressionskoeffizient  $b=0,239$ .

## 2.10. Korrelations-Koeffizient nach Pearson

Hier verwendete Datenmenge: HFD.SAV  
 Hier verwendetes Subkollektiv: geschl=2 & diabetes =0  
 Hier verwendete Variablen: HFD, SPA\_PROX

Befehlsfolge: **Analysieren, Korrelation, Bivariat** ★ Dialogfeld *Bivariate Korrelationen*



Dort sind die entsprechenden Merkmale aus der Variablenliste in das Feld **Variablen** einzusteuern. Die Option **Spearman** im Feld **Korrelationskoeffizienten** ist dann zu aktivieren, wenn bei einem oder gar beiden zu korrelierenden Merkmalen keine annähernde Normalverteilung vorliegt. Im Falle annähernder Normalverteilung bei beiden Variablen darf der **Pearson**-Korrelationskoeffizient berechnet werden. **Zweiseitig** im Feld **Test auf Signifikanz** sowie **Signifikante Korrelationen markieren** sind bereits voreingestellt. Die Auswertung kann mit **OK** veranlasst werden.

★ Es erscheint als Output die Matrix der Korrelationskoeffizienten (jeweils 1. Zeile):

|              |                               | Hautfaldendicke         | Knochendichte (prox. Messort) |
|--------------|-------------------------------|-------------------------|-------------------------------|
| Spearman-Rho | Hautfaldendicke               | Korrelationskoeffizient | 1,000                         |
|              |                               | Sig. (2-seitig)         | ,339**                        |
|              |                               | N                       | ,000                          |
|              | Knochendichte (prox. Messort) | Korrelationskoeffizient | 210                           |
|              |                               | Sig. (2-seitig)         | ,339**                        |
|              |                               | N                       | ,000                          |
|              |                               | 210                     | 210                           |

\*\* Korrelation ist bei Niveau 0,01 signifikant (zweiseitig).

Der Korrelationskoeffizient ist in der ersten Zeile zu finden, er beträgt hier 0.339. Der p-Wert des Korrelationskoeffizienten findet sich in der Zeile "Sig. (2-seitig)" und ist so klein, dass er bei drei Nachkommastellen nur als 0,000 dargestellt wird, d h.:  $p < 0,001$ .

## 2.11. Partiieller Korrelationskoeffizient

Hier verwendete Variablen: HFD, SPA\_PROX  
 Kontrollvariable: ALTER  
 Hier verwendetes Subkollektiv: GESCHL = 2 & DIABETES = 0

Befehlsfolge: **Analysieren, Korrelation, partiell** ★ Dialogfeld *Partielle Korrelationen*. Die weiteren Optionen **Zweiseitig** und **Tatsächliches Signifikanzniveau anzeigen** sind hier bereits von der Voreinstellung her aktiviert. Die Variablen *hfd* und *spa\_prox* sind ins Feld **Variablen**, die Variable ALTER ins Feld **Kontrollvariablen** einzusteuern, danach kann **OK** gegeben werden. - SPSS erzeugt folgenden Output:

**Korrelationen**

| Kontrollvariablen              |                                |                        | Hautfaltendicke | Knochen­dichte (prox. Messort) |
|--------------------------------|--------------------------------|------------------------|-----------------|--------------------------------|
| alter                          | Hautfaltendicke                | Korrelation            | 1,000           | ,076                           |
|                                |                                | Signifikanz (2-seitig) | .               | ,274                           |
|                                |                                | df                     | 0               | 207                            |
| Knochen­dichte (prox. Messort) | Knochen­dichte (prox. Messort) | Korrelation            | ,076            | 1,000                          |
|                                |                                | Signifikanz (2-seitig) | ,274            | .                              |
|                                |                                | df                     | 207             | 0                              |

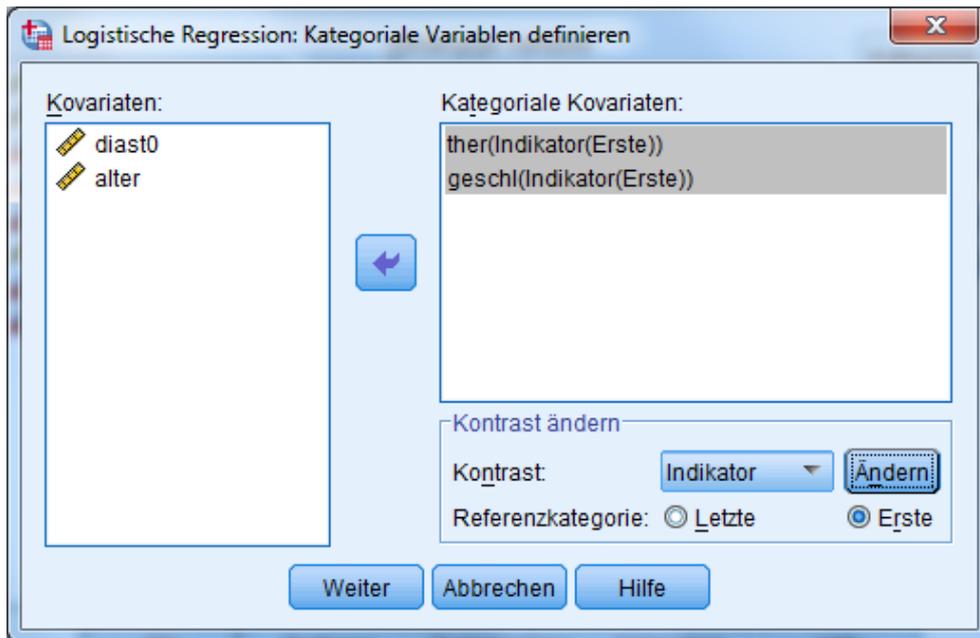
Der partielle Korrelationskoeffizient ist wieder in der ersten Zeile zu finden, er beträgt hier 0.076.

Der p-Wert des Korrelationskoeffizienten beträgt bei dieser Analyse (mit Kontrollvariable) 0,274.

## 2.12. Logistische Regression

Hier verwendete Datenmenge: VERDINUM.SAV  
 Hier verwendete Variablen: RESP8, DIAST0, ALTER, THER, VBEH, GESCHL

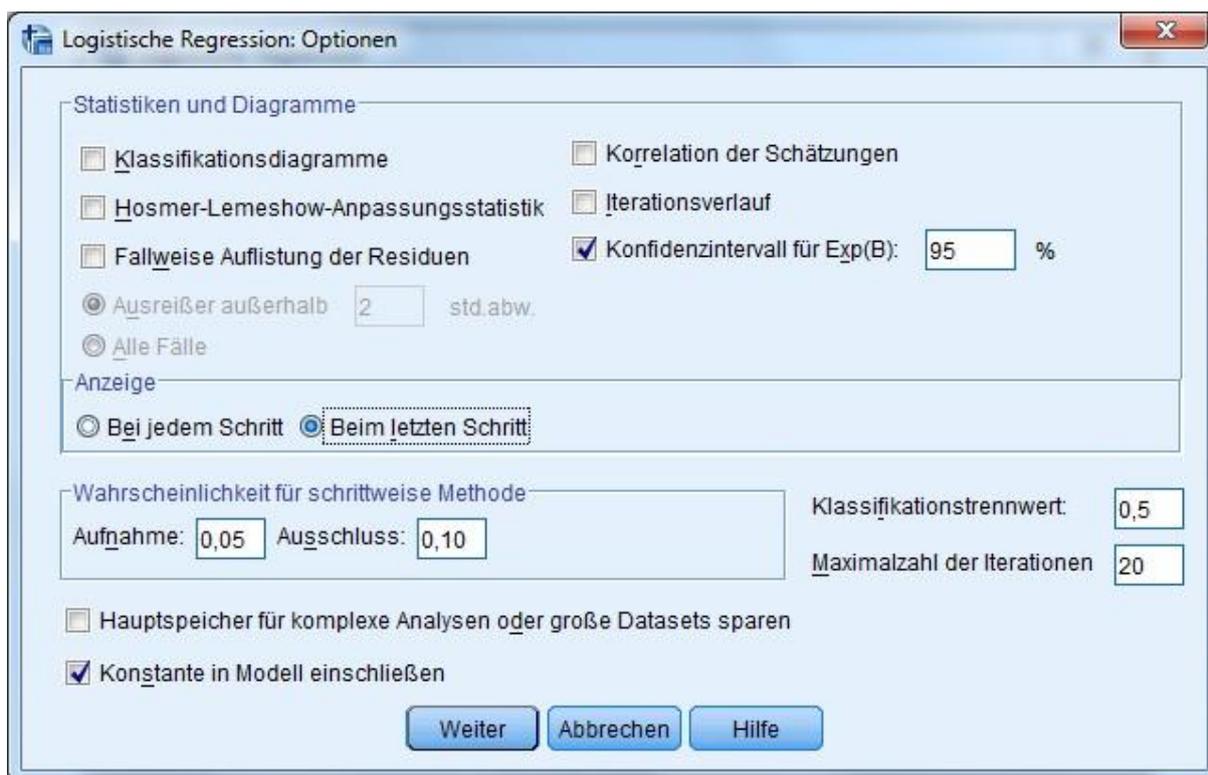
Die Logistische Regression ist durchzuführen mit **Analysieren, Regression, Binär logistisch** ★ Dialogfeld *Logistische Regression*. Dort ist zunächst die abhängige Variable *resp8* in das entsprechende Feld einzutragen. In das Feld **Kovariaten** sind die unabhängigen Variablen *diast0*, *alter*, *ther* und *geschl* einzusteuern. Danach muss über die Schaltfläche **Kategorial** das Dialogfeld *Kategoriale Variablen definieren* aufgerufen werden. Dort müssen die Variablen *ther* und *geschl* zusammen markiert und in das Feld **Kategoriale Kovariaten** eingesteuert werden. Sodann ist im Feld **Kontrast ändern** der Radio-Button auf **Referenzkategorie: Erste** zu setzen und auf **Ändern** zu klicken. Wenn das Dialogfeld die folgende Gestalt aufweist...



kann man mit **Weiter** in das übergeordnete Dialogfeld *Logistische Regression* zurückkehren. Im Listenfeld **Methode** soll die Selektionsmethode **Einschluss** aktiviert sein. Wenn diese Einstellungen vollzogen sind, sollte das Dialogfeld aussehen wie folgt:



Nunmehr ist noch die Schaltfläche **Optionen** zu betätigen ★ Dialogfeld *Logistische Regression: Optionen*. Das Konfidenzintervall von 95% wird eingestellt durch Aktivierung des Kontrollfeldes **Konfidenzint. für Exp(B)**. Der **Hosmer-Lemeshow-Test** ist ebenfalls aufzurufen. Im Feld **Anzeigen** soll die Ausgabe des Ergebnisses nur **Beim letzten Schritt** erfolgen. Dafür ist die entsprechende Option durch Anklicken des Radio-Buttons zu aktivieren, so dass das Dialogfeld *Logistische Regression: Optionen* schließlich das folgende Aussehen aufweist:



Der Befehl **Weiter** führt zurück ins übergeordnete Dialogfeld *Logistische Regression*, wo noch **OK** zu geben ist. SPSS erzeugt daraufhin folgenden Output:

#### Zusammenfassung der Falverarbeitung

| Ungewichtete Fälle <sup>a</sup> |                       | H   | Prozent |
|---------------------------------|-----------------------|-----|---------|
| Ausgewählte Fälle               | Einbezogen in Analyse | 352 | 95,4    |
|                                 | Fehlende Fälle        | 17  | 4,6     |
|                                 | Gesamtsumme           | 369 | 100,0   |
| Nicht ausgewählte Fälle         |                       | 0   | ,0      |
| Gesamtsumme                     |                       | 369 | 100,0   |

a. Wenn die Gewichtung in Kraft ist, finden Sie in der Klassifikationstabelle die Gesamtzahl von Fällen.

#### Codierung abhängiger Variablen

| Ursprünglicher Wert | Interner Wert |
|---------------------|---------------|
| nein                | 0             |
| ja                  | 1             |

#### Codierungen kategorialer Variablen

|            |                    | Häufigkeit | Parameter<br>codierung |
|------------|--------------------|------------|------------------------|
|            |                    |            | (1)                    |
| Geschlecht | 1                  | 187        | ,000                   |
|            | 2                  | 165        | 1,000                  |
| Behandlung | Hydrochlorothiazid | 178        | ,000                   |
|            | Verapamil          | 174        | 1,000                  |

**Klassifikationstabelle<sup>a,b</sup>**

| Beobachtet        |                              |      | Vorhersagewert               |     |                     |
|-------------------|------------------------------|------|------------------------------|-----|---------------------|
|                   |                              |      | Therapieerfolg nach 8 Wochen |     | Prozentsatz richtig |
|                   |                              |      | nein                         | ja  |                     |
| Schritt 0         | Therapieerfolg nach 8 Wochen | nein | 0                            | 175 | ,0                  |
|                   |                              | ja   | 0                            | 177 | 100,0               |
| Gesamtprozentsatz |                              |      |                              |     | 50,3                |

a. Die Konstante ist im Modell enthalten.

b. Der Trennwert ist ,500

## Block 0: Anfangsblock

**Klassifikationstabelle<sup>a,b</sup>**

| Beobachtet        |                              |      | Vorhersagewert               |     |                     |
|-------------------|------------------------------|------|------------------------------|-----|---------------------|
|                   |                              |      | Therapieerfolg nach 8 Wochen |     | Prozentsatz richtig |
|                   |                              |      | nein                         | ja  |                     |
| Schritt 0         | Therapieerfolg nach 8 Wochen | nein | 0                            | 175 | ,0                  |
|                   |                              | ja   | 0                            | 177 | 100,0               |
| Gesamtprozentsatz |                              |      |                              |     | 50,3                |

a. Die Konstante ist im Modell enthalten.

b. Der Trennwert ist ,500

**Variablen in der Gleichung**

|           |           | B    | Standardfehler | Wald | df | Sig. | Exp(B) |
|-----------|-----------|------|----------------|------|----|------|--------|
| Schritt 0 | Konstante | ,011 | ,107           | ,011 | 1  | ,915 | 1,011  |

**Nicht in der Gleichung vorhandene Variablen**

|                 |           |           | Score  | df | Sig. |
|-----------------|-----------|-----------|--------|----|------|
| Schritt 0       | Variablen | diast0    | 33,922 | 1  | ,000 |
|                 |           | alter     | 9,814  | 1  | ,002 |
|                 |           | ther(1)   | 8,292  | 1  | ,004 |
|                 |           | geschl(1) | ,402   | 1  | ,526 |
| Gesamtstatistik |           |           | 48,856 | 4  | ,000 |

## Block 1: Methode = Eingabe

**Omnibustests der Modellkoeffizienten**

|           |         | Chi-Quadrat | df | Sig. |
|-----------|---------|-------------|----|------|
| Schritt 1 | Schritt | 52,172      | 4  | ,000 |
|           | Block   | 52,172      | 4  | ,000 |
|           | Modell  | 52,172      | 4  | ,000 |

**Modellübersicht**

| Schritt | -2 Log-Likelihood    | R-Quadrat nach Cox & Snell | R-Quadrat nach Nagelkerke |
|---------|----------------------|----------------------------|---------------------------|
| 1       | 435,792 <sup>a</sup> | ,138                       | ,184                      |

a. Die Schätzung wurde bei Iteration Nummer 4 beendet, da Parameterschätzungen sich um weniger als ,001 geändert haben.

Klassifikationstabelle<sup>a</sup>

| Beobachtet        |                              |      | Vorhersagewert               |     |                     |
|-------------------|------------------------------|------|------------------------------|-----|---------------------|
|                   |                              |      | Therapieerfolg nach 8 Wochen |     | Prozentsatz richtig |
|                   |                              |      | nein                         | ja  |                     |
| Schritt 1         | Therapieerfolg nach 8 Wochen | nein | 109                          | 66  | 62,3                |
|                   |                              | ja   | 55                           | 122 | 68,9                |
| Gesamtprozentsatz |                              |      |                              |     | 65,6                |

a. Der Trennwert ist ,500

Variablen in der Gleichung

|                        | B         | Standardfehler | Wald  | df     | Sig. | Exp(B) | 95% Konfidenzintervall für EXP (B) |        |       |
|------------------------|-----------|----------------|-------|--------|------|--------|------------------------------------|--------|-------|
|                        |           |                |       |        |      |        | Unterer                            | Oberer |       |
| Schritt 1 <sup>a</sup> | diast0    | ,112           | ,020  | 31,601 | 1    | ,000   | ,894                               | ,860   | ,930  |
|                        | alter     | ,031           | ,011  | 8,074  | 1    | ,004   | 1,031                              | 1,010  | 1,054 |
|                        | ther(1)   | ,574           | ,232  | 6,099  | 1    | ,014   | 1,775                              | 1,126  | 2,800 |
|                        | geschl(1) | -,070          | ,232  | ,092   | 1    | ,762   | ,932                               | ,592   | 1,468 |
|                        | Konstante | 9,899          | 2,103 | 22,148 | 1    | ,000   | 19907,933                          |        |       |

a. In Schritt 1 eingegebene Variable(n): diast0, alter, ther, geschl.

### Erläuterungen zum SPSS-Ausdruck:

Unter *Codierung abhängiger/ kategorialer Variablen* ist abzulesen, wie die Merkmale intern kodiert werden. Daraus wird ersichtlich, dass die Wahrscheinlichkeit des Auftretens der Response modelliert wird ("ja" wird der 1 zugeordnet). Die *Klassifizierungstabelle* dient zur Beurteilung der Modellgüte (Grundlage zur Berechnung von Sensitivität und Spezifität des Vorhersagemodells), ebenso wie der p-Wert des Hosmer-Lemeshow-Tests weiter unten.

Die letzte Tabelle enthält die zentralen Aussagen zur Bewertung der Einflussgrößen. Unter *Sig.* wie *Signifikanz* ist der p-Wert des zweiseitigen Tests für die jeweiligen Einflussgrößen angegeben. Der kleinste p-Wert (,000) ist mit dem diastolischen Blutdruck bei Behandlungsbeginn verbunden. Von den anderen ausgewählten Merkmalen sind auf dem 5%-Niveau ebenfalls signifikant mit der Response assoziiert: *alter* und *ther*.

Das Vorzeichen und der Wert der zu den Merkmalen gehörenden Parameter sind in der Spalte **B** abzulesen. Unter **Exp(B)** ist das Odds Ratio der Merkmale zu finden. Zu beachten ist, dass SPSS bei nichtnumerischen Variablen die interne Kodierung benutzt (s. o.). Demzufolge steigt beispielsweise bei der Therapie die Responsewahrscheinlichkeit beim Übergang von *hct* (0) auf *ver* (1) und daher führt das Antihypertensivum Verapamil zu besseren Ergebnissen. Bei älteren Patienten ist wegen des positiven Vorzeichens von **B** (bzw. weil das Odds Ratio größer eins ist) von einer besseren Wirksamkeit der Antihypertensiva auszugehen. Das negative Vorzeichen für den Effekt des diastolischen Blutdrucks bei Behandlungsbeginn bedeutet, dass die Wahrscheinlichkeit einer Response bei höherem initialem Blutdruck abnimmt.

### 2.13. Analyse von Überlebenszeiten, KAPLAN-MEIER, Log-Rank-Test

Hier verwendete Datenmenge:

KARZINOM\_NEU.SAV

Hier verwendete Subkollektive:

Alter kleiner oder gleich 45 Jahre oder  
Alter größer als 45 Jahre (verschlüsselt  
in Variable AGEGROUP)

Hier verwendete Variablen:

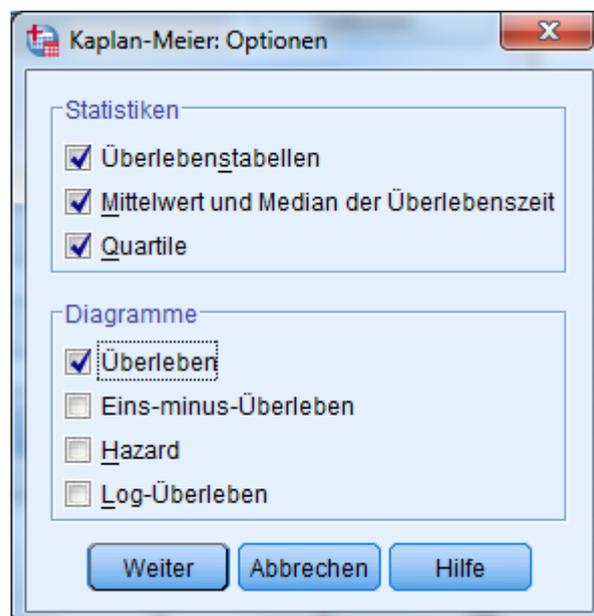
Zeitvariable BEOBREZ (Beobachtungszeit bis zum evtl. eingetretenen Rezidiv);  
Statusvariable REZIDIV (Rezidiv oder Tod aufgetreten; 0=nein, 1=ja)

Mit der Befehlsfolge **Analysieren, Überleben, Kaplan-Meier** gelangt man in das Dialogfeld *Kaplan-Meier*. Dort ist in das Feld **Zeit** die Zeitvariable *beobrez* und in das Feld **Status** die Statusvariable *rezidiv* einzutragen. Dann ist die Schaltfläche **Ereignis definieren** zu betätigen, da SPSS die Ausprägung des interessierenden Ereignisses mitgeteilt werden muss - in diesem Falle ist es die Ausprägung 1 für "Rezidiv oder Tod aufgetreten=ja".

Nun muss in die Zeile **Faktor** die Faktor-Variablen *agegroup* eingetragen werden. Durch Betätigung der Schaltfläche **Faktor vergleichen** gelangt man in das Dialogfeld *Kaplan-Meier: Faktorstufen vergleichen*, wo, wie im Folgenden dargestellt, der Log-Rank-Test zu aktivieren ist:



Ein Klick auf **Weiter** führt zurück in das übergeordnete Dialogfeld *Kaplan-Meier*. Dort ist unter der Schaltfläche **Optionen** das untergeordnete Dialogfeld *Kaplan-Meier: Optionen* aufzurufen. In diesem Dialogfeld sind die Optionen **Überlebenstabellen** und **Mittelwert und Median der Überlebenszeit** schon vorgebelegt. Außerdem ist im Feld **Statistiken** die Option **Quartile** sowie im Feld **Diagramme** die Option **Überleben** zur Darstellung einer Grafik der Überlebensfunktionen zu aktivieren. Wenn das Dialogfeld folgendes Aussehen aufweist...



...kann man mit **Weiter** zurück in das übergeordnete Dialogfeld *Kaplan-Meier* gelangen. Nach Betätigung von **Ok** erzeugt SPSS folgenden Output (die Überlebenstabellen sind aus Platzgründen in der Abbildung stark gekürzt):

### Zusammenfassung der Fallverarbeitung

| Altersklasse | Gesamtzahl | Anzahl der Ereignisse | Zensiert |         |
|--------------|------------|-----------------------|----------|---------|
|              |            |                       | H        | Prozent |
| bis 45 J.    | 36         | 6                     | 30       | 83,3%   |
| über 45 J.   | 157        | 48                    | 109      | 69,4%   |
| Insgesamt    | 193        | 54                    | 139      | 72,0%   |

[...]

### Überlebenstabelle

| Altersklasse | Zeit | Status | Kumulativer Anteil der Überlebenden zu 'Zeit' |                | Anzahl der kumulativen Ereignisse | Anzahl der verbleibenden Fälle |     |
|--------------|------|--------|---|----------------|-----------------------------------|--------------------------------|-----|
|              |      |        | Schätzung                                     | Standardfehler |                                   |                                |     |
| bis 45 J.    | 1    | 3,000  | 1   | ,972           | ,027                              | 1                              | 35  |
|              | 2    | 3,000  | 0   | .              | .                                 | 1                              | 34  |
|              | 3    | 5,000  | 0   | .              | .                                 | 1                              | 33  |
| über 45 J.   | 28   | 3,000  | 1   | ,992           | ,008                              | 1                              | 129 |
|              | 29   | 3,000  | 0   | .              | .                                 | 1                              | 128 |
|              | 30   | 3,000  | 0   | .              | .                                 | 1                              | 127 |

### Mittelwerte und Mediane für Überlebenszeit

| Altersklasse | Mittelwert <sup>a</sup> |                |                        |            | Median    |                |                        |            |
|--------------|-------------------------|----------------|------------------------|------------|-----------|----------------|------------------------|------------|
|              | Schätzung               | Standardfehler | 95%-Konfidenzintervall |            | Schätzung | Standardfehler | 95%-Konfidenzintervall |            |
|              |                         |                | Untergrenze            | Obergrenze |           |                | Untergrenze            | Obergrenze |
| bis 45 J.    | 55,559                  | 8,337          | 39,219                 | 71,899     | .         | .              | .                      | .          |
| über 45 J.   | 35,744                  | 3,472          | 28,938                 | 42,550     | 26,000    | 2,884          | 20,348                 | 31,652     |
| Insgesamt    | 40,989                  | 3,660          | 33,816                 | 48,163     | 34,000    | 5,587          | 23,049                 | 44,951     |

a. Wenn die Schätzung zensiert ist, wird sie auf die größte Überlebenszeit begrenzt.

### Perzentile

| Altersklasse | 50,0%     |                | 75,0%     |                |
|--------------|-----------|----------------|-----------|----------------|
|              | Schätzung | Standardfehler | Schätzung | Standardfehler |
| bis 45 J.    | .         | .              | 47,000    | 18,165         |
| über 45 J.   | 26,000    | 2,884          | 14,000    | 1,775          |
| Insgesamt    | 34,000    | 5,587          | 16,000    | 1,482          |

### Gesamtvergleiche

|                       | Chi-Quadrat | df | Sig. |
|-----------------------|-------------|----|------|
| Log Rank (Mantel-Cox) | 5,695       | 1  | ,017 |

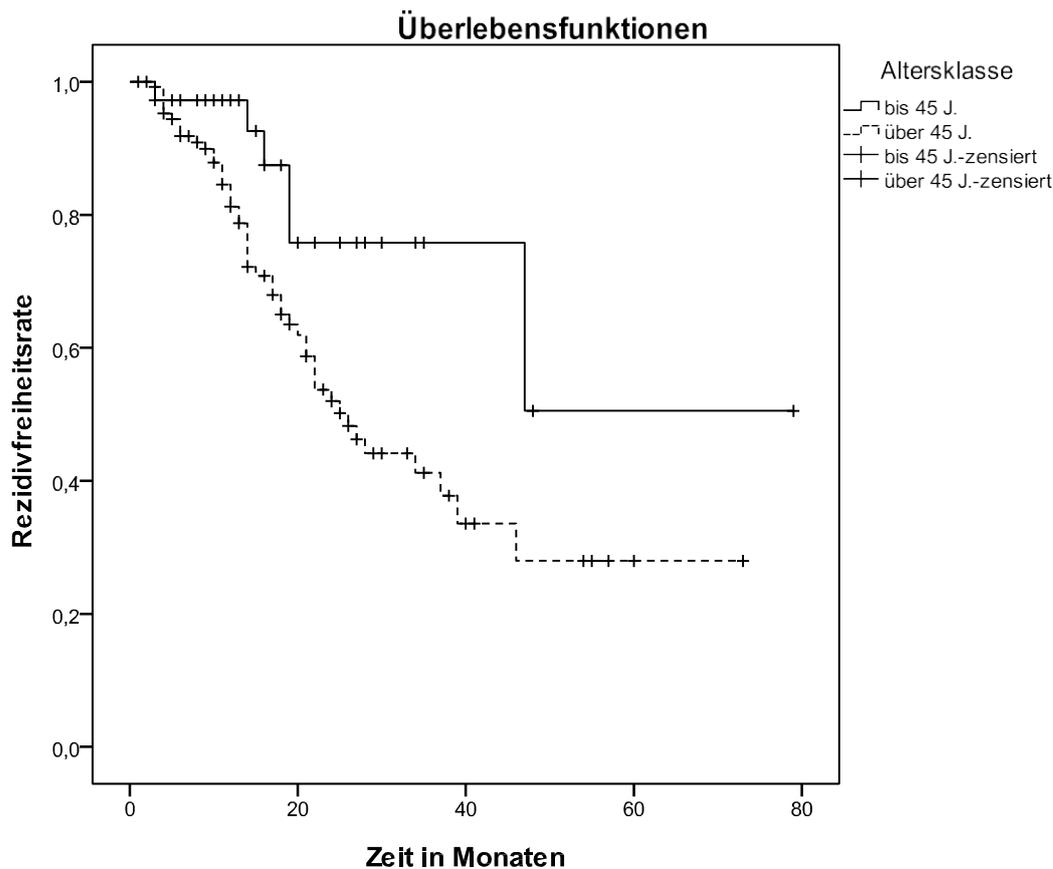
Test auf Gleichheit der Überlebensverteilungen für verschiedene Stufen von Altersklasse.

### Erläuterungen zum SPSS-Output:

- Mittelwert/Schätzung geschätzte mittlere Zeit, rezidivfrei zu überleben (oft nicht sehr aussagekräftig)
- Median/Schätzung geschätzte mediane Zeit, rezidivfrei zu überleben
- Standardfehler zugehörige geschätzte Standardabweichungen
- Perzentile (25% / 50% / 75%) Zeitpunkt, bei dem die geschätzte Wahrscheinlichkeit, rezidivfrei zu überleben, 25%, 50% oder 75% erreicht.
- Die Ergebnisse des Log-Rank-Tests findet man im Abschnitt **Gesamtvergleiche** in der Zeile "Log-Rank (Mantel-Cox)":
- Chi-Quadrat Prüfgröße
- df Freiheitsgrade
- Sig. p-Wert ("Signifikanz") des zweiseitigen Tests

### Erläuterungen zur Grafik:

Die Grafik der Rezidivfreiheitsfunktionen muss nachbearbeitet werden (Doppelklick auf Grafik). Zur Änderung der Linienfarbe in schwarz sind die Überlebensfunktionen einzeln zu markieren; das Dialogfeld *Eigenschaften* ist über das Kontextmenü (rechte Maustaste) erreichbar. Die Achsenbeschriftungen können durch Doppelklick editiert, bei Bedarf (2. Überschrift) kann ein Textfeld eingefügt werden. Die Legende lässt sich verschieben. Nachdem auch die Überschrift eingesetzt, die Linienfarbe in schwarz geändert und die Achsenbeschriftung entsprechend angepasst wurde, sollte die Grafik folgende Gestalt aufweisen:



### 2.14. Regression für zensierte Beobachtungen (Cox-Regression)

Hier verwendete Datenmenge:

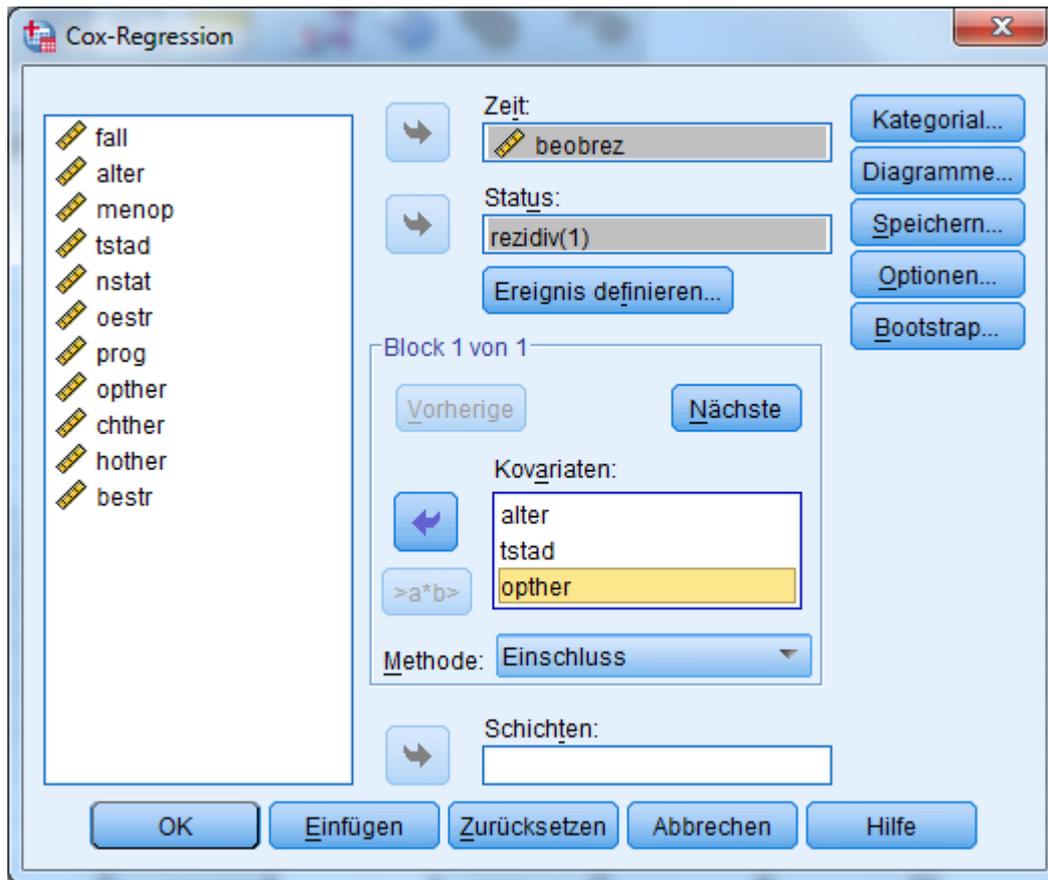
KARZINOM.SAV

Hier verwendete Variablen:

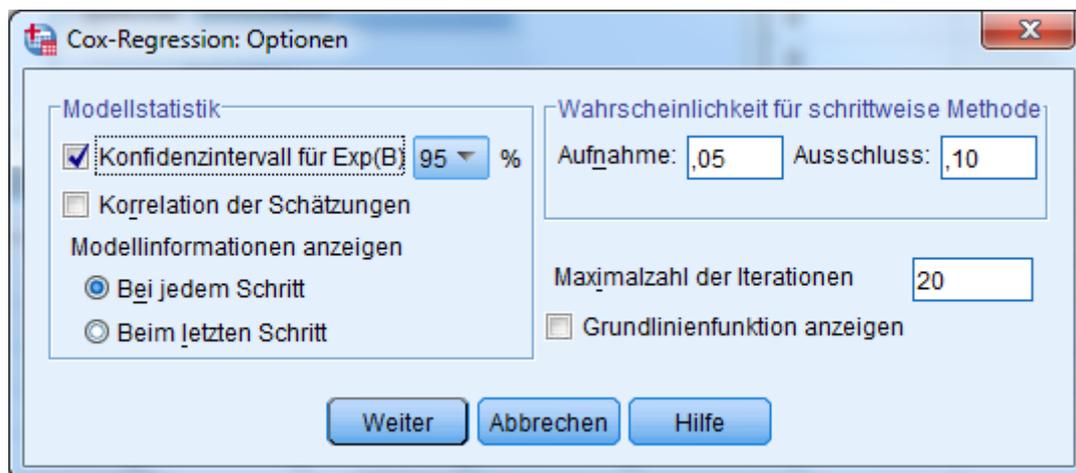
BEOBREZ, REZIDIV, ALTER, TSTAD,  
OPTHER.

Fehlwerte ("keine Angabe") wurden ausgeschlossen  
(s. 1.3.5. und Datenbeschreibung S. 34/35).

Mit der Befehlsfolge **Analysieren, Überlebensanalyse, Cox-Regression** gelangt man in das Dialogfeld *Cox-Regression*. Die Zeit- und die Status-Variable sind genauso wie in 2.13. zu behandeln. Man trägt die interessierenden Kovariablen in das entsprechende Feld ein (hier: *alter, tstad, opther*). Die Methode **Einschluß** soll beibehalten werden:



Unter den **Optionen** wählt man durch anklicken des Kontrollfeldes (erscheint ein Häkchen!) das 95%-Konfidenzintervall für Exp(B):



Mit **Weiter** gelangt man wieder in das übergeordnete Dialogfeld *Cox-Regression*. Ein Klick auf die Schaltfläche **Kategorial** führt in das Dialogfeld *Cox-Regression: Kategoriale Kovariaten definieren*. Dort sind die Variablen *tstad* und *opther* in das Feld **Kategoriale Kovariaten** einzusteuern. Beide Variablen müssen markiert sein; dann im Feld **Kontrast ändern** den Radio-Button auf die Option **Referenzkategorie: Erste** setzen und die Schaltfläche **Ändern** anklicken.

Nachdem die beschriebenen Einstellungen getroffen sind, sollte das Dialogfeld folgende Gestalt besitzen:



Ein Klick auf **Weiter** führt ins übergeordnete Dialogfeld *Cox-Regression*, wo zur Ausführung der Berechnung noch **OK** zu geben ist. Es erscheint nun folgender Output:

## Cox-Regression

### Zusammenfassung der Fallverarbeitung

|                             |  | H   | Prozent |
|-----------------------------|--|-----|---------|
| In Analyse verfügbare Fälle | Ereignis <sup>a</sup>                                  | 54  | 28,0%   |
|                             | Zensiert   | 112 | 58,0%   |
|                             | Gesamtsumme  | 166 | 86,0%   |
| Nicht verwendete Fälle      | Fälle mit fehlenden Werten                             | 0   | 0,0%    |
|                             | Fälle mit negativer Zeit                               | 0   | 0,0%    |
|                             | Zensierte Fälle vor dem ersten Ereignis in der Schicht | 27  | 14,0%   |
|                             | Gesamtsumme  | 27  | 14,0%   |
| Gesamtsumme                 |  | 193 | 100,0%  |

a. Abhängige Variable: beobrez

### Codierungen kategorialer Variablen<sup>a,c</sup>

|                     |   | Häufigkeit | (1) | (2) | (3) | (4) |
|---------------------|---|------------|-----|-----|-----|-----|
| tstad <sup>b</sup>  | 1 | 19         | 0   | 0   | 0   | 0   |
|                     | 2 | 66         | 1   | 0   | 0   | 0   |
|                     | 3 | 41         | 0   | 1   | 0   | 0   |
|                     | 4 | 65         | 0   | 0   | 1   | 0   |
|                     | 5 | 2          | 0   | 0   | 0   | 1   |
| opther <sup>b</sup> | 1 | 172        | 0   | 0   | 0   |     |
|                     | 2 | 14         | 1   | 0   | 0   |     |
|                     | 3 | 5          | 0   | 1   | 0   |     |
|                     | 4 | 2          | 0   | 0   | 1   |     |

a. Kategorievariable: tstad

b. Codierung von Indikatorparametern

c. Kategorievariable: opther

## Block 0: Anfangsblock

### Omnibustests der Modellkoeffizienten

|                   |
|-------------------|
| -2 Log-Likelihood |
| 457,697           |

## Block 1: Methode = Eingabe

### Omnibustests der Modellkoeffizienten<sup>a</sup>

| -2 Log-Likelihood | Gesamt (Score) |                |      | Änderung vom vorherigen Schritt |                |      | Änderung vom vorherigen Block |                |
|-------------------|----------------|----------------|------|---------------------------------|----------------|------|-------------------------------|----------------|
|                   | Chi-Quadrat    | Freiheitsgrade | Sig. | Chi-Quadrat                     | Freiheitsgrade | Sig. | Chi-Quadrat                   | Freiheitsgrade |
| 435,865           | 26,799         | 7              | ,000 | 21,832                          | 7              | ,003 | 21,832                        | 7              |

a. Anfangsblock 1: Methode = Eingabe

### Variablen in der Gleichung<sup>b</sup>

|           | B     | SE    | Wald   | Freiheitsgrade | Sig. | Exp(B) | 95,0% Konfidenzintervall für Exp(B) |         |
|-----------|-------|-------|--------|----------------|------|--------|-------------------------------------|---------|
|           |       |       |        |                |      |        | Unterer                             | Oberer  |
| alter     | -,001 | ,012  | ,004   | 1              | ,950 | ,999   | ,977                                | 1,023   |
| tstad     |       |       | 21,245 | 4              | ,000 |        |                                     |         |
| tstad(1)  | ,574  | ,769  | ,558   | 1              | ,455 | 1,776  | ,394                                | 8,011   |
| tstad(2)  | ,721  | ,790  | ,833   | 1              | ,361 | 2,056  | ,437                                | 9,665   |
| tstad(3)  | 1,722 | ,756  | 5,187  | 1              | ,023 | 5,593  | 1,271                               | 24,610  |
| tstad(4)  | 3,027 | 1,037 | 8,517  | 1              | ,004 | 20,641 | 2,703                               | 157,641 |
| opther    |       |       | 4,948  | 2 <sup>a</sup> | ,084 |        |                                     |         |
| opther(1) | ,362  | ,532  | ,463   | 1              | ,496 | 1,436  | ,506                                | 4,074   |
| opther(2) | 1,356 | ,628  | 4,666  | 1              | ,031 | 3,879  | 1,134                               | 13,272  |

a. Freiheitsgrad aufgrund konstanter oder linear abhängiger Kovariaten reduziert

b. Konstant oder linear abhängige Kovariaten opther(3) = tstad(4) ;

### Kovariate Mittelwerte

|           | Mittelwert |
|-----------|------------|
| alter     | 56,602     |
| tstad(1)  | ,331       |
| tstad(2)  | ,211       |
| tstad(3)  | ,361       |
| tstad(4)  | ,012       |
| opther(1) | ,078       |
| opther(2) | ,030       |
| opther(3) | ,012       |

## Erläuterung zum SPSS-Output:

Im Abschnitt "Variablen in der Gleichung" findet man in den entsprechenden Spalten

|                |   |
|----------------|---|
| B              | Regressionskoeffizient der Cox-Regression                                 |
| SE             | zugehöriger Standardfehler  |
| Wald           | Prüfgröße   |
| df             | Freiheitsgrade  |
| Signifikanz    | p-Wert des zugehörigen Tests  |
| Exp(B)         | Hazard Ratio, wenn die entsprechende Variable um eine Einheit erhöht wird |
| Untere / Obere | untere / obere Grenze des Konfidenzintervalls für das relative Risiko     |

## 2.15. Bland-Altman-Plot

Hier verwendete Datenmenge: ANTIMOD.SAV (wurde nur für diese Anleitung erstellt)  
Hier verwendete Variablen: ANTIK ("Erfasste Messwerte, altes Gerät"),  
MODERN ("Erfasste Messwerte, modernes Gerät")

Der Bland-Altman-Plot ist wie das schon in 2.8. behandelte Streudiagramm (Scatterplot) eine Punktwolke. Überdies aber enthält ein fertiger Bland-Altman-Plot drei zusätzliche von Werten der Y-Achse ausgehende waagerechte Bezugslinien. Die mittlere Bezugslinie befindet sich in Höhe des Mittelwertes der Differenzen der Messwerte. Über dieser "Mittelwert-Linie" wird eine weitere Gerade eingetragen, deren Wert sich errechnet aus dem Mittelwert plus der doppelten Standardabweichung des auf Y aufgetragenen Merkmals (Differenzen der Messwerte). Unter der „Mittelwert-Linie“ wird die dritte Gerade eingetragen, welche sich aus dem Mittelwert minus der doppelten Standardabweichung errechnet. Die Region innerhalb der oberen und der unteren Bezugslinie enthält unter Annahme der Normalverteilung 95% der Werte.

### Vorgehensweise bei Erzeugung des Bland-Altman-Plots:

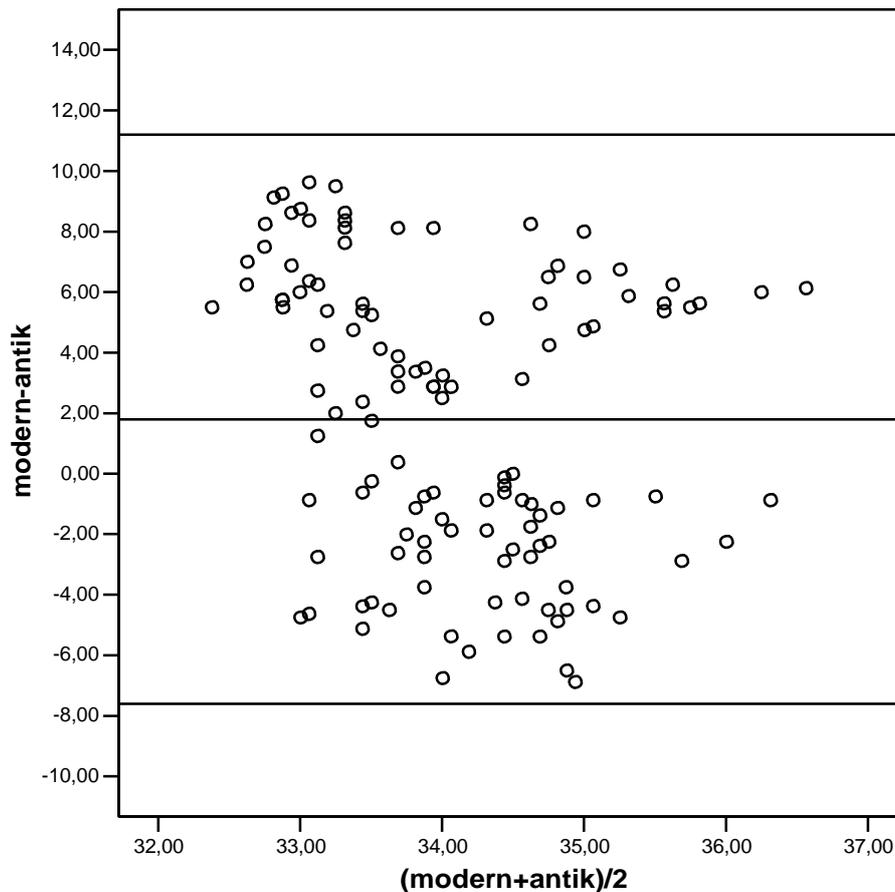
a) Punktwolke erzeugen und editieren: Befehlsfolge **Grafik, Alte Dialogfelder, Streu-/Punkt-Diagramm, Einfaches Streudiagramm, Definieren** => Dialogfeld *Einfaches Streudiagramm*. In die Eingabezeile für die X-Achse wird die Variable des Mittelwertes der einzelnen Messungen eingesteuert, während die Variable der Differenzen der Messwerte auf die Y-Achse zu legen ist. Das entstandene Streudiagramm muss nun mit Doppelklick (mitten hinein) editiert werden.

b) Skala der Y-Achse anpassen: da die Bezugslinie für den höchsten Wert (11,1865) nicht sichtbar ist, wird jetzt die Skalierung geändert mit **Bearbeiten, Y-Achse auswählen, Register Skala**. Deaktivieren Sie das Kontrollfeld **Auto** vor der Eingabezeile **Maximum** und geben Sie dort unter **Benutzerdefiniert** den Wert 14 ein. Verkleinern Sie zudem die Unterteilung der Y-Achsenkala von 5 auf 2 wie folgt: deaktivieren Sie das Kontrollfeld **Auto** vor der Eingabezeile **Erstes Inkrement** und geben Sie dort unter **Benutzerdefiniert** den Wert 2 ein. Beenden Sie die Anpassung der Y-Achse mit **Anwenden, Schließen**.

c) Auftragung der drei Geraden auf der Y-Achse, nachdem der Mittelwert und die Standardabweichung der Differenzen berechnet wurden:

- ✓ Für den Mittelwert der Differenzen (hier im Beispiel: 1,7883);
- ✓ Für Mittelwert plus 2-fache Standardabweichung (hier im Beispiel: 11,1865);
- ✓ Für Mittelwert minus 2-fache Standardabweichung (hier im Beispiel: -7,6099).

Befehlsfolge: **Optionen, Bezugslinie für Y-Achse** ==> Dialogfeld *Eigenschaften*, Register **Bezugslinie**. Es wird automatisch bei beliebigem Wert eine Linie erzeugt, deren neue Position nun in der Eingabezeile **Achsenposition** einzutragen ist: 1,7883. Nun muss **Anwenden** und **Schließen** befohlen werden. Dieser Vorgang ist für die beiden anderen Positionierungswerte analog zu wiederholen. Zum Schluss sollte die Grafik folgende Gestalt aufweisen:



Nach Schließen des Grafik-Editors ist der Bland-Altman-Plot Bestandteil der Ausgabe.

## 2.16. Kappa-Maß, McNemar-Test

Hier verwendete Datenmenge: BEFUNDE.SAV (wurde nur für diese Anleitung erstellt) Hier verwendete Variablen: BEFUND\_1 ("Werte von Befunder 1", Ausprägungen: "0=unauffällig, 1=auffällig, 3=pathologisch")  
BEFUND\_2 ("Werte von Befunder 2")

Soll zwecks Feststellung von Unterschieden zwischen den kategorialen Ergebnissen zweier Befunder ein McNemar-Test sowie das dazugehörige Kappa-Maß berechnet werden, so ist folgendermaßen vorzugehen:

Befehlsfolge **Analysieren, Deskriptive Statistiken, Kreuztabellen**: ==> Dialogfeld *Kreuztabellen*. In dessen Variablen-Eingabezeilen (**Zeilen:** / **Spalten:**) sind die beiden Merkmale einzusteuern, wobei die Reihenfolge beliebig ist. Danach ist die Schaltfläche **Statistiken** anzuklicken, um in das Dialogfeld *Kreuztabellen: Statistik* zu gelangen, wo durch Anklicken des jeweiligen Kontrollfeldes das Kappa-Maß und der McNemar-Test aufzurufen wären. Der Befehl **Weiter** führt zurück ins übergeordnete Dialogfeld *Kreuztabellen*. Dort können entweder durch Betätigung der Schaltfläche **Zellen...** noch zusätzliche Informationen

für die Kreuztabelle (**Zeilen-/Spaltenprozent**e) hervorgerufen oder durch Klick auf **OK** die Berechnung von McNemar-Test und Kappa-Maß direkt angestoßen werden. Es wird daraufhin folgender Output ausgegeben (Auszug):

- für den McNemar-Test:

**Chi-Quadrat-Tests**

|                           | Wert   | df | Asymp. Sig. (zweiseitig) |
|---------------------------|--------|----|--------------------------|
| McNemar-Bowker-Test       | 11,727 | 3  | ,008                     |
| Anzahl der gültigen Fälle | 159    |    |                          |

p-Wert des McNemar-Tests

- für das Kappa-Maß:

**Kappa-Maß**

**Symmetrische Maße**

|                                      | Wert | Asymp. Standardfehler <sup>a</sup> | Näherungsweise A <sup>b</sup> | Näherungsweise Sig. |
|--------------------------------------|------|------------------------------------|-------------------------------|---------------------|
| Maß für die Übereinstimmung<br>Kappa | ,197 | ,060                               | 3,511                         | ,000                |
| Anzahl der gültigen Fälle            | 159  |                                    |                               |                     |

a. Die Nullhypothese wird nicht vorausgesetzt.

b. Unter Annahme der Nullhypothese wird der asymptotische Standardfehler verwendet.

## 2.17. Einfaktorielle ANOVA

Hier verwendete Datenmenge: KARZINOM.SAV

Hier verwendete Variablen: ALTER, MENOP (Menopausenstatus)

Die Befehlsfolge **Analysieren, Mittelwerte vergleichen, Einfaktorielle Varianzanalyse** führt in das Dialogfeld *Einfaktorielle Varianzanalyse*. Die interessierende Variable, hier *alter*, ist in das Feld **Abhängige Variablen** und die Gruppierungsvariable, hier *menop*, in die Zeile **Faktor** einzusteuern. Nach Betätigung von **OK** erscheint folgender Output:

**ANOVA**

alter

|                       | Quadratsumme | df  | Mittel der Quadrate | F       | Sig. |
|-----------------------|--------------|-----|---------------------|---------|------|
| Zwischen Gruppen      | 17348,065    | 2   | 8674,033            | 117,761 | ,000 |
| Innerhalb der Gruppen | 13994,950    | 190 | 73,658              |         |      |
| Gesamtsumme           | 31343,016    | 192 |                     |         |      |

Erläuterungen zum SPSS-Ausdruck:

F

Prüfgröße

Signifikanz

p-Wert der Varianzanalyse

## 2.18. Kruskal-Wallis-Test

Hier verwendete Datenmenge: KLWA\_DEMO.SAV (wurde nur für diese Anleitung erstellt)

Hier verwendete Variablen: MESSWERT, GRUPPE

Die Befehlsfolge **Analysieren, Nichtparametrische Tests, Alte Dialogfelder, K unabhängige Stichproben** führt in das Dialogfeld *Tests bei mehreren unabhängigen Stichproben*. Die interessierende Variable, hier MESSWERT, ist in das Feld **Testvariablen** und die Gruppierungsvariable, hier GRUPPE, in das Feld **Gruppierungsvariable** einzusteuern. Danach ist der Dialog **Bereich definieren** zu aktivieren. Unter *Bereich für Gruppierungsvariable* verlangt SPSS die Angabe der niedrigsten und der höchsten Ausprägung der Gruppierungsvariablen. Nach Betätigung von **Weiter** sowie von **OK** im übergeordneten Dialogfeld *Tests bei mehreren unabhängigen Stichproben* erscheint folgender Output (Auszug):

**Teststatistiken<sup>a,b</sup>**

|             | messwert |
|-------------|----------|
| Chi-Quadrat | 183,006  |
| df          | 3        |
| Asymp. Sig. | ,000     |

a. Kruskal-Wallis-Test

b. Gruppierungsvariable:  
gruppe

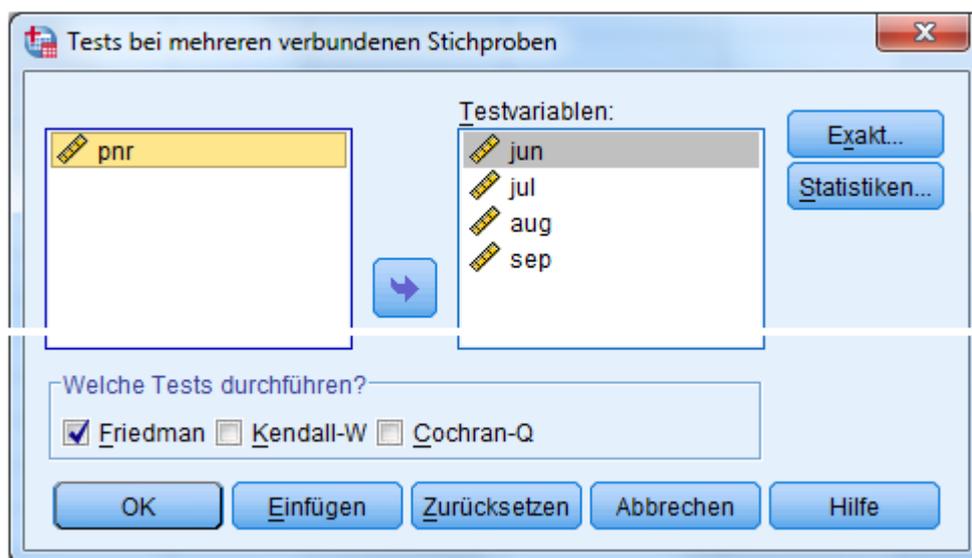
Unter **Asymptotische Signifikanz** ist der p-Wert des Kruskal-Wallis-Tests verzeichnet.

## 2.19. Friedman-Test

Hier verwendete Datenmenge: SCHWIMMBADNUTZUNG (wurde nur für diese Anleitung erstellt)

Hier verwendete Variablen: PNR (ProbandInnennummer)  
JUN, JUL AUG; SEP (Anzahl der Schwimmbadbesuche der einzelnen ProbandInnen in den entsprechenden Monaten)

Mit der Befehlsfolge **Analysieren, Nichtparametrische Tests, Alte Dialogfelder, K verbundene Stichproben** gelangt man in das Dialogfeld *Tests bei mehreren verbundenen Stichproben*, wo der Friedman-Test im Kontrollfeld schon aktiviert ist. In das Feld **Testvariablen** sind die interessierenden Variablen einzusteuern, in diesem Falle JUN, JUL, AUG und SEP. Wenn das Dialogfeld sich wie folgt darstellt...



...kann **OK** betätigt werden. Im Ausgabe-Fenster erscheint daraufhin die Liste der mittleren Ränge aller Variablen sowie die Teststatistik für den Friedman-Test:

| Ränge                          |                |
|--------------------------------|----------------|
|                                | Mittlerer Rang |
| Schwimmbadbesuche im Juni      | 1,47           |
| Schwimmbadbesuche im Juli      | 2,77           |
| Schwimmbadbesuche im August    | 3,11           |
| Schwimmbadbesuche im September | 2,64           |

| Teststatistiken <sup>a</sup> |         |
|------------------------------|---------|
| H                            | 297     |
| Chi-Quadrat                  | 279,135 |
| df                           | 3       |
| Asymp. Sig.                  | ,000    |

a. Friedman-Test

Erläuterungen zum SPSS-Ausdruck:

**Chi-Quadrat:** Prüfgröße des Friedman-Tests

**df:** Freiheitsgrade

**Asymptotische Signifikanz:** p-Wert des Friedman-Tests.

### 3. Beschreibung der in dieser Anleitung benutzten Datenmengen

#### 3.1. Datenmenge HDF.SAV

Ein aussagekräftiges diagnostisches Verfahren zur Erkennung der Osteoporose (besonders bei Frauen) ist die Messung der Knochendichte, die jedoch mit nicht unerheblichem Aufwand verbunden ist. Daher wurde die Frage diskutiert, ob man durch die wesentlich einfachere Bestimmung der Hautfaltendicke (HFD) mittels einer neuen Methode relativ zuverlässig auf die Knochendichte und damit auf eine mögliche Osteoporose Gefährdung schließen kann.

In einer Dissertation (R. Zawalski, 1997) wurden für insgesamt 222 weibliche und 47 männliche Patienten im Alter von 14 - 82 Jahren die HFD, die Knochendichte nach verschiedenen Methoden, sowie verschiedene Faktoren, die einen Einfluss auf die genannten Größen haben könnten, erhoben. Im Einzelnen besteht die Datenmenge HFD aus folgenden Variablen:

| Variablenname | Inhalt   |
|---------------|--|
| FALLNR        | fortlaufende Nummer der Patienten  |
| GESCHL        | Geschlecht (1=männlich, 2=weiblich)  |
| ALTER         | Alter in vollendeten Jahren  |
| GROESSE       | Körpergröße in cm  |
| GEWICHT       | Körpergewicht in kg  |
| BMI           | Body Mass Index = Körpergewicht [kg], dividiert durch Körperlänge <sup>2</sup> [m <sup>2</sup> ] |
| STEROIDE      | Einnahme von Steroiden (0=nein, 1=ja)  |
| SCH_HORM      | Einnahme von Schilddrüsenhormonen (0=nein, 1=ja)   |
| HYPOTHY       | Vorliegen einer Hypothyreose (0=nein, 1=ja)  |
| HYPERTHY      | Vorliegen einer Hyperthyreose (0=nein, 1=ja)   |
| DIABETES      | Vorliegen von Diabetes (0=nein, 1=ja)  |
| HFD           | Hautfaltendicke am Handrücken [mm]   |
| SPA_PROX      | Knochendichte am proximalen Messort [mg/cm <sup>2</sup> ]  |
| SPA_DIST      | Knochendichte am distalen Messort [mg/cm <sup>2</sup> ]  |
| DPA_LWS       | Knochendichte an der Lendenwirbelsäule [mg/cm <sup>2</sup> ]                                     |
| DPA_LIFE      | Knochendichte am linken Femurhals [mg/cm <sup>2</sup> ]  |
| DPA_REFE      | Knochendichte am rechten Femurhals [mg/cm <sup>2</sup> ]   |

Die SPA-Werte wurden mit Single-Photonen-Absorptiometrie, die DPA-Werte mit Dual-Photonen-Absorptiometrie bestimmt; letztere wurden nur für 110 bzw. 111 Frauen erhoben.

#### 3.2. Datenmenge KARZINOM.SAV

In einer retrospektiven Studie (Dissertation H. Rößler, Mainz 1986) sollte untersucht werden, ob der Rezeptorstatus für Östrogen und Progesteron eine prognostische Bedeutung für das Auftreten von Rezidiven bei Frauen mit Mamma-Karzinom hat. Außerdem wurden weitere mögliche prognostische Faktoren sowie die verschiedenen Therapien, die die Frauen erhalten hatten, erhoben. Insgesamt standen die Daten von 193 Frauen mit Mamma-Karzinom, bei denen die Therapie zunächst zur Remission geführt hatte, zur Verfügung.

| Variablenname | Inhalt  |
|---------------|---|
| FALL          | Fall-Nummer   |
| ALTER         | Alter (in vollendeten Jahren) (99=keine Ang.)   |
| MENOP         | Menopausenstatus (1=Prae, 2=Peri, 3=Post, 9=keine Ang.)   |
| TSTAD         | Tumorstadium (0=T0, 1=T1, 2=T2, 3=T3, 4=T4, 5=keine Ang.)                                       |
| NSTAT         | Lymphknotenstatus (0=N0, 1=N1, 2=N2, 3=N3, 4=keine Ang.)  |
| OESTR         | Rezeptorstatus Oestrogen (1=positiv, 2=negativ, 9=keine Ang.)                                   |
| PROG          | Rezeptorstatus Progesteron (1=positiv, 2=negativ, 9=keine Ang.)                                 |
| OPTHER        | Operative Therapie (1=radikal, 2=eingeschränkt, 3=Probeexzision, 4=Punktion, 9=keine Ang.)      |
| CHTHER        | Chemotherapie (0=nein, 1=therapeutisch, 2=adjuvant, 9=keine Ang.)                               |
| HOTHER        | Hormontherapie (0=nein, 1=ja, 9=keine Ang.)   |
| BESTR         | Bestrahlung (0=nein, 1=ja, 9=keine Ang.)  |
| REZIDIV       | Rezidiv oder Tod aufgetreten (0=nein, 1=ja)   |
| BEOBREZ       | Beobachtungszeit bis zum (evtl.) eingetretenen Rezidiv (Monate)                                 |
| AGEGROUP      | Altersgruppe (1 für <=45 Jahre, 2 für > 45 Jahre), nur in Datenmenge KARZINOM_NEU.SAV vorhanden |

### 3.3. Datenmenge SCHWIMMBADNUTZUNG

Diese fiktive Datenmenge enthält 297 Fälle. Sie dient zu Lehr- und Demonstrationszwecken.

| Variablenname | Inhalt                                    |
|---------------|---|
| PNR           | PNR (ProbandInnennummer)                  |
| JUN           | Anzahl der Schwimmbadbesuche im Juni      |
| JUL           | Anzahl der Schwimmbadbesuche im Juli      |
| AUG           | Anzahl der Schwimmbadbesuche im August    |
| SEP           | Anzahl der Schwimmbadbesuche im September |

### 3.4. Datenmenge SPORT\_LMK

Diese fiktive Datenmenge enthält 442 Fälle. Sie dient zu Lehr- und Demonstrationszwecken.

| Variablenname | Inhalt  |
|---------------|---|
| NR            | Nummer des Falles   |
| GESCHL        | Geschlecht der SportlerIn / des Sportlers (0=männlich, 1=weiblich)      |
| GEWICHT       | Körpergewicht in kg   |
| STRANDUR      | normalerweise Strandurlaub durchgeführt (0=nein, 1=ja)                  |
| GROESSE       | Körpergröße in Meter  |
| SPORTART      | ausgeübte Sportart (1=Leichtathletik, 2=Mannschaftssport, 3=Kraftsport) |

### 3.5. Datenmenge VERDINUM.SAV

In der VERDI-Studie (Holzgreve et al.[1989], British Medical Journal 299, 881-886) wurde der Kalziumantagonist Verapamil mit dem Diuretikum Hydrochlorothiazid bezüglich der anti-hypertensiven Wirksamkeit verglichen. Es nahmen insgesamt 369 Patienten an der Studie teil, für die nach dem Zufallsprinzip festgelegt wurde, welche von beiden Behandlungen sie erhielten. Bei jedem Patienten wurde der Behandlungseffekt nach 8, 24 und 48 Wochen qualitativ beurteilt (Zielkriterium: diastolischer Blutdruck < 90 mmHg). Für einige Patienten war der Behandlungserfolg ("Response") nicht beurteilbar, da sie ihre Teilnahme an der Studie vorzeitig beendeten.

| <b>Variablenname</b> | <b>Inhalt</b>   |
|----------------------|---|
| PTNNR                | Fortlaufende Patientenummer                           |
| ALTER                | Lebensalter bei Behandlungsbeginn [Jahre]             |
| THER                 | Behandlung [HCT: Hydrochlorothiazid / VER: Verapamil] |
| GESCHL               | Geschlecht [m: männlich / w: weiblich]                |
| VBEH                 | Vorbehandlung mit Antihypertensiva [ja/nein]          |
| DIAST0               | Diastolischer Blutdruck bei Behandlungsbeginn [mmHg]  |
| RESP8                | Therapieerfolg nach 8 Wochen [ja/nein]                |
| RESP24               | Therapieerfolg nach 24 Wochen [ja/nein]               |
| RESP48               | Therapieerfolg nach 48 Wochen [ja/nein]               |

## 4. Index der wichtigsten Befehlsfolgen

|  |         |
|--|---------|
| Absolute und relative Häufigkeiten (kategoriale Variable(n))   | [S. 10] |
| <b>Analysieren, Deskriptive Statistiken, Häufigkeiten...</b>   |         |
| Balkendiagramm   | [ - ]   |
| <b>Grafiken, Veraltete Dialogfelder, Balken...</b>   |         |
| Boxplot  | [S. 14] |
| <b>Grafiken, Veraltete Dialogfelder, Boxplot...</b>  |         |
| Chi-Quadrat-Unabhängigkeitstest  | [S. 10] |
| <b>Analysieren, Deskriptive Statistiken, Kreuztabellen, Variablen eintragen, Statistik, Chi-Quadrat, Weiter, OK</b>  |         |
| Cox-Regression   | [S. 29] |
| <b>Analysieren, Überlebensanalyse, Cox-Regression</b>  |         |
| Datei aufteilen  | [S. 8]  |
| <b>Daten, Datei aufteilen, Ausgabe nach Gruppen aufteilen...</b>   |         |
| Datei aufteilen rückgängig machen (aufheben)   | [S. 8]  |
| <b>Daten, Datei aufteilen, Alle Fälle analysieren, OK</b>  |         |
| Datei öffnen (im Daten-Editor)   | [S. 2]  |
| <b>Datei, Öffnen, Datei markieren, OK</b>  |         |
| Datei speichern (Ausgabe-Viewer-Fenster)   | [S. 4]  |
| <b>Datei, Speichern unter, Dateiname eintragen, OK</b>   |         |
| Einfaktorielle ANOVA   | [S. 35] |
| <b>Analysieren, Mittelwerte vergleichen, Einfaktorielle ANOVA, Variablen eintragen, OK</b>   |         |
| Exakter Fisher-Test  | [S. 10] |
| <b>Analysieren, Deskriptive Statistiken, Kreuztabellen, Variablen eintragen, Statistik, Chi-Quadrat, Weiter, OK</b>  |         |
| Fälle auswählen  | [S. 9]  |
| <b>Daten, Fälle auswählen, Falls Bedingung zutrifft, Falls</b>   |         |
| Fallauswahl rückgängig machen (aufheben)   | [S. 9]  |
| <b>Daten, Fälle auswählen, Alle Fälle</b>  |         |
| Histogramm   |         |
| <b>Grafiken, Veraltete Dialogfelder, Histogramm...</b>   |         |
| Kaplan-Meier Überlebensraten-Schätzung   | [S. 26] |
| <b>Analysieren, Überlebensanalyse, Kaplan-Meyer</b>  |         |
| Kappa-Maß  | [S. 33] |
| <b>Analysieren, Deskriptive Statistiken, Kreuztabellen, Statistik, Kontrollfeld Kappa aktivieren, Weiter, OK</b>   |         |
| Kategorisieren einer stetigen Variablen  | [S. 6]  |
| <b>Transformieren, Umkodieren, in andere Variablen</b>   |         |
| Korrelationskoeffizient nach Pearson   | [S. 21] |
| <b>Analysieren, Korrelation, Bivariat (Pearson ist vorbelegt)</b>  |         |
| Korrelationskoeffizient nach Spearman  | [ - ]   |
| <b>Analysieren, Korrelation, Bivariat, Pearson deaktivieren, Spearman aktivieren</b>   |         |
| Korrelationskoeffizient, partiell  | [S. 22] |
| <b>Analysieren, Korrelation, Partiiell</b>   |         |
| Kruskal-Wallis-Test  | [S. 35] |
| <b>Analysieren, Nichtparametrische Tests, Alte Dialogfelder, K unabhängige Stichproben, Variablen eintragen, Bereich definieren, Weiter, OK</b>                        |         |
| Lineare Regression   | [S. 21] |
| <b>Analysieren, Regression, Linear</b>   |         |
| Logistische Regression   | [S. 23] |
| <b>Analysieren, Regression, Binär logistisch</b>   |         |
| McNemar-Test   | [S. 33] |
| <b>Analysieren, Deskriptive Statistiken, Kreuztabellen, Statistik, Kontrollfeld McNemar aktivieren, Weiter, OK</b>   |         |
| Neue Variable aus bestehender Variablen berechnen  | [S. 7]  |
| <b>Transformieren, Berechnen</b>   |         |
| Statistische Maßzahlen (stetige Variable(n))   | [S. 13] |
| <b>Analysieren, Deskriptive Statistiken, Häufigkeiten, Variablen einsteuern, Häufigkeitstabellen anzeigen deaktivieren, Statistik, Maßzahlen auswählen, Weiter, OK</b> |         |
| Streudiagramm (Scatterplot)  | [S. 20] |
| <b>Grafiken, Veraltete Dialogfelder, Streudiagramm, Einfach</b>  |         |
| T-Test für unverbundene Stichproben  | [S. 15] |
| <b>Analysieren, Mittelwerte vergleichen, T-Test bei unabhängigen</b>   |         |

|   |   |         |
|---|---|---------|
|   | <b>Stichproben</b>  |         |
| T-Test für verbundene Stichproben                                   |   | [S. 18] |
|   | <b>Analysieren, Mittelwerte vergleichen, T-Test bei gepaarten Stichproben</b> |         |
| Vierfeldertafeln (Kreuztabellen)                                    |   | [S. 10] |
|   | <b>Analysieren, Deskriptive Statistiken, Kreuztabellen</b>                    |         |
| Wilcoxon-Test (=“Mann-Whitney-U-Test“) für unverbundene Stichproben |   | [S. 17] |
|   | <b>Analysieren, Nichtparametrische Tests, Zwei unabhängige Stichproben</b>    |         |
| Wilcoxon-Test für verbundene Stichproben                            |   | [ - ]   |
|   | <b>Analysieren, Nichtparametrische Tests, Zwei verbundene Stichproben</b>     |         |

## 5. Wegweiser zum Aufspüren von Kennzahlen der statistischen Auswertung in SPSS 23

Chi-Quadrat Unabhängigkeits-Test, p-Wert: Zeile Chi-Quadrat nach Pearson / Spalte Asymptotische Signifikanz (2-seitig)

Chi-Quadrat-Tests

|                                | Wert                | df | Asymptotische Signifikanz (2-seitig) |
|--------------------------------|---------------------|----|--------------------------------------|
| Chi-Quadrat nach Pearson       | 47,094 <sup>a</sup> | 6  | ,000                                 |
| Likelihood-Quotient            | 55,191              | 6  | ,000                                 |
| Zusammenhang linear-mit-linear | 35,708              | 1  | ,000                                 |
| Anzahl der gültigen Fälle      | 413                 |    |                                      |

a. 3 Zellen (25,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist ,92.

Cox-Regression, p-Werte: Spalte Signifikanz

### Einfaktorielle ANOVA

ONEWAY ANOVA

alter

|                       | Quadratsumme | df  | Mittel der Quadrate | F       | Signifikanz |
|-----------------------|--------------|-----|---------------------|---------|-------------|
| Zwischen den Gruppen  | 17348,065    | 2   | 8674,033            | 117,761 | ,000        |
| Innerhalb der Gruppen | 13994,950    | 190 | 73,658              |         |             |
| Gesamt                | 31343,016    | 192 |                     |         |             |

**Exakter Fisher-Test, p-Wert:** Zeile **Exakter Test nach Fisher** / Spalte **Exakte Signifikanz (2-seitig)**

Chi-Quadrat-Tests

|                                    | Wert               | df | Asymptotische Signifikanz (2-seitig) | Exakte Signifikanz (2-seitig) | Exakte Signifikanz (1-seitig) |
|------------------------------------|--------------------|----|--------------------------------------|-------------------------------|-------------------------------|
| Chi-Quadrat nach Pearson           | 2,105 <sup>b</sup> | 1  | ,147                                 |                               |                               |
| Kontinuitätskorrektur <sup>a</sup> | 1,554              | 1  | ,213                                 |                               |                               |
| Likelihood-Quotient                | 2,106              | 1  | ,147                                 |                               |                               |
| Exakter Test nach Fisher           |                    |    |                                      | ,159                          | ,106                          |
| Zusammenhang linear-mit-linear     | 2,084              | 1  | ,149                                 |                               |                               |
| Anzahl der gültigen Fälle          | 100                |    |                                      |                               |                               |

a. Wird nur für eine 2x2-Tabelle berechnet

b. 0 Zellen (,0%) haben eine erwartete Häufigkeit kleiner 5. Die minimale erwartete Häufigkeit ist 18,45.

Statistik für Test<sup>a</sup>

**Friedman-Test, p-Wert:**  
Zeile **Asymptotische Signifikanz**

|                           |         |
|---------------------------|---------|
| N                         | 297     |
| Chi-Quadrat               | 279,135 |
| df                        | 3       |
| Asymptotische Signifikanz | ,000    |

a. Friedman-Test

**Kappa-Maß:** Zeile **Maß der Übereinstimmung Kappa** / Spalte **Wert Korrelationskoeffizient** (bivariat) nach Pearson oder Spearman:  
Zeile **Korrelationskoeffizient** / Spalte **"Variablenname"**  
 (hier: **spa\_prox**)

Korrelationen

|              |          | hfd                     | spa_prox |
|--------------|----------|-------------------------|----------|
| Spearman-Rho | hfd      | Korrelationskoeffizient | 1,000    |
|              |          | Sig. (2-seitig)         | ,380**   |
|              |          | N                       | 269      |
| spa_prox     | spa_prox | Korrelationskoeffizient | ,380**   |
|              |          | Sig. (2-seitig)         | 1,000    |
|              |          | N                       | 269      |

\*\* . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

### Kruskal-Wallis-Test

Statistik für Test<sup>a, b</sup>

|                           | Messwert |
|---------------------------|----------|
| Chi-Quadrat               | 183,006  |
| df                        | 3        |
| Asymptotische Signifikanz | ,000     |

a. Kruskal-Wallis-Test

b. Gruppenvariable: Gruppe

p-Wert: Zeile **Asymptotische Signifikanz**, Spalte **"Messwert"**

**Lineare Regression**, p-Werte: Spalte **Signifikanz**

**Logistische Regression**, p-Werte: Spalte **Sig.**

**Log-Rank-Test:** letzte Tabelle der Ausgabe in Zeile **Log Rank (Mantel-Cox)** / Spalte **Sig.**

**Gesamtvergleiche**

|                       | Chi-Quadrat | Freiheitsgrade | Sig. |
|-----------------------|-------------|----------------|------|
| Log Rank (Mantel-Cox) | 15,497      | 3              | ,001 |

Test auf Gleichheit der Überlebensverteilungen für die verschiedenen Stufen von tstad.

**McNemar-Test**, p-Wert: Zeile **McNemar-Bowker-Test** / Spalte **Asymptotische Signifikanz (2-seitig)**

**Chi-Quadrat-Tests**

|                           | Wert   | df | Asymptotische Signifikanz (2-seitig) |
|---------------------------|--------|----|--------------------------------------|
| McNemar-Bowker-Test       | 11,727 | 3  | ,008                                 |
| Anzahl der gültigen Fälle | 159    |    |                                      |

**t-Test** (unverbunden), p-Wert: Zeile **Varianzen sind nicht gleich** / Spalte **Sig (2-seitig)**

**Test bei unabhängigen Stichproben**

|  |                             | Levene-Test der Varianzgleichheit |             | T-Test für die Mittelwertgleichheit |         |                 |                    |                              |                                      |       |
|--|-----------------------------|-----------------------------------|-------------|-------------------------------------|---------|-----------------|--------------------|------------------------------|--------------------------------------|-------|
|  |                             | F                                 | Signifikanz | T                                   | df      | Sig. (2-seitig) | Mittlere Differenz | Standardfehler der Differenz | 99% Konfidenzintervall der Differenz |       |
|  |                             |                                   |             |                                     |         |                 |                    |                              | Untere                               | Obere |
| Diast. Blutdruck bei Behandlungsbeginn | Varianzen sind gleich       | ,730                              | ,393        | ,287                                | 367     | ,774            | ,186               | ,648                         | -1,492                               | 1,864 |
|  | Varianzen sind nicht gleich |                                   |             | ,287                                | 366,871 | ,774            | ,186               | ,648                         | -1,491                               | 1,863 |

**t-Test** (verbunden), p-Wert: Spalte **Sig (2-seitig)**

Test bei gepaarten Stichproben

|                                    | Gepaarte Differenzen |                    |                                 |                                      |       | T      | df  | Sig. (2-seitig) |
|------------------------------------|----------------------|--------------------|---------------------------------|--------------------------------------|-------|--------|-----|-----------------|
|                                    | Mittelwert           | Standardabweichung | Standardfehler des Mittelwertes | 99% Konfidenzintervall der Differenz |       |        |     |                 |
|                                    |                      |                    |                                 | Untere                               | Obere |        |     |                 |
| Paaren 1<br>spa_prox -<br>spa_dist | ,35186               | ,12705             | ,00775                          | ,3318                                | ,3720 | 45,421 | 268 | ,000            |

**Vorzeichentest**, p-Wert: Zeile **Asymptotische Signifikanz (2-seitig)** / Spalte **"Variablenamen"**  
 (hier: Gewicht in g Frühjahr 07 - Gewicht in g Herbst 06)

Statistik für Test<sup>a</sup>

|   |  |
|---|--|
|   | Gewicht in g<br>Frühjahr 07 -<br>Gewicht in g<br>Herbst 06 |
| Z                                       | -9,540   |
| Asymptotische<br>Signifikanz (2-seitig) | ,003   |

**Wilcoxon-Test** (unverbunden), p-Wert: Zeile **Asymptotische Signifikanz (2-seitig)** / Spalte **"Variablenname"**  
 (hier: Diast.

a. Vorzeichentest

Blutdruck bei Behandlungsbeginn)

Statistik für Test<sup>a</sup>

|   | Diast.<br>Blutdruck bei<br>Behandlung<br>sbeginn |
|---|--|
| Mann-Whitney-U                          | 18823,000  |
| Wilcoxon-W                              | 33476,000  |
| Z                                       | -,190  |
| Asymptotische<br>Signifikanz (2-seitig) | ,850   |

a. Gruppenvariable: Behandlung

**Wilcoxon-Test** (verbunden), p-Wert: Zeile Asymptotische Signifikanz (2-seitig) / Spalte "Variablennamen"

(hier: Gewicht in g Frühjahr 07 –Gewicht in g Herbst 06)

Statistik für Test<sup>b</sup>

|   | Gewicht in g<br>Frühjahr 07 -<br>Gewicht in g<br>Herbst 06 |
|---|--|
| Z                                       | -8,406 <sup>a</sup>  |
| Asymptotische<br>Signifikanz (2-seitig) | ,000   |

a. Basiert auf positiven Rängen.

b. Wilcoxon-Test