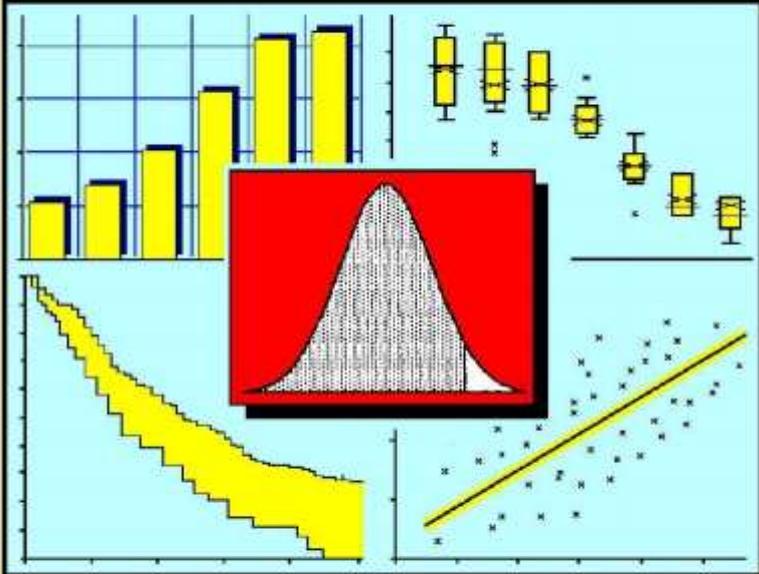


O. Hoffmann

PC-STATISTIK



Das schnelle Auswertungssystem für
deskriptive und analytische Statistik
mit uni- und multivariaten Verfahren

KURZ-EINFÜHRUNG

INHALT

| | | |
|----------|-------------------------------------|-----------|
| 1 | EINLEITUNG | 2 |
| 2 | KURZE EINFÜHRUNG IN PCS | 2 |
| 2.1 | Sofort mit der Arbeit beginnen | 2 |
| 2.2 | Definition der Dateistruktur | 4 |
| 2.3 | Dateneingabe | 6 |
| 2.4 | Gruppenbildung | 7 |
| 2.5 | Auswertung: Statistische Kenngrößen | 11 |
| 2.6 | Klasseneinteilungen definieren | 18 |
| 2.7 | Auswertung: Häufigkeitsauszählung | 20 |
| 3 | STATISTISCHER TASCHENRECHNER | 23 |
| 3.1 | Unabhängige Stichproben, t-Test | 24 |
| 3.2 | Odds Ratio | 25 |
| 4 | OPTIONEN | 26 |
| 5 | DRUCKMANAGER | 27 |
| 6 | PROTOKOLLDATTEI | 28 |
| 7 | VON PCS4 NACH PCS5 | 29 |

1 EINLEITUNG

PC-Statistik ist nunmehr seit über zwölf Jahren auf dem Markt und hat seinen überzeugten Anwenderkreis gefunden. Aus diesem Kreis sind auch viele Anregungen und Wünsche an den Autor herangetragen worden, was letztlich zur Entwicklung der vorliegenden Version führte.

Auch bei dieser Version wurde wieder versucht, die bisherige PCS-Philosophie so weit als möglich beizubehalten. Das Arbeiten mit dieser neuen Version gestaltet sich genauso einfach und schnell wie die Verwendung der bisherigen PCS-Versionen. Für den vertrauten Anwender bereitet damit die Umstellung von PCS4 auf PCS5 keine Probleme, zumal keinerlei Umsetzungsläufe erforderlich sind. Und der Anfänger findet wie bisher ein Auswertungssystem vor, welches ihm einen leichten Einstieg ermöglicht.

Nach wie vor ist PCS gestaltet worden für eine Zielgruppe, die sich eben gerade nicht aus Spezialisten auf dem Gebiet der Statistik zusammensetzt, sondern eher aus Anfängern, die auch mit den theoretischen Grundlagen nicht so vertraut sind. Besondere Unterstützung erhält dieser Anwenderkreis durch kurze und verständliche Einführungen in die theoretischen Grundlagen der zur Verfügung gestellten Auswertungsverfahren.

Eine wesentliche Änderung bringt die neue Programmversion in diesem Zusammenhang aber doch noch mit sich: Dem Trend der Zeit folgend wird das Einsteiger-Handbuch, welches die wichtigsten Aspekte der Bedienung aufzeigt, nur noch als Windows-Help-File zur Verfügung gestellt. Alle weiteren wichtigen Informationen zur Bedienung und zur Theorie finden Sie in der Online-Hilfe, die auch dementsprechend in diese beiden Teile aufgegliedert ist. Wir hoffen, dadurch auch wesentlich flexibler reagieren zu können, wenn Änderungen und Erweiterungen anstehen.

Allen, die PC-Statistik als Tester oder mit kritischen Beiträgen oder Verbesserungsvorschlägen auf dem Weg von Version 1 bis zur vorliegenden Version wohlwollend begleitet haben, sei an dieser Stelle herzlich gedankt.

2 KURZE EINFÜHRUNG IN PCS

Bevor Sie sich über die Online-Hilfe näher mit allen Möglichkeiten und den Details von PCS beschäftigen, soll Ihnen ein kleines Beispiel demonstrieren, wie einfach der Einstieg in das Programm und das Arbeiten damit ist. Die Arbeitsweise von PC-Statistik entspricht in weiten Bereichen nämlich genau dem Vorgehen, das man bei einer Auswertung per Hand anwenden würde.

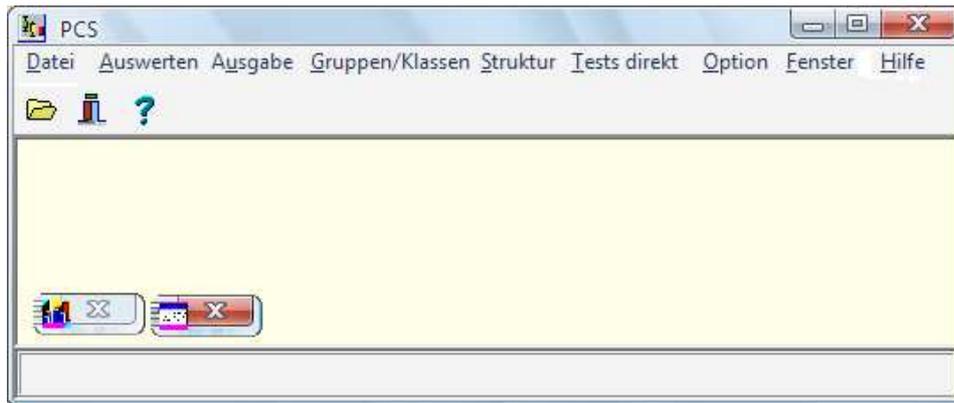
2.1 Sofort mit der Arbeit beginnen

Am Anfang steht zunächst die Datensammlung, völlig unabhängig von Rechnertyp und eingesetztem Programm. Nehmen wir an, dass wir in einer Schulklasse eine Umfrage durchgeführt haben, wobei für jedes Kind Geschlecht, Mathematik-Note, die Anzahl der Kinder in der Familie und die Körpergröße registriert wurden. Unser Ausgangsmaterial können wir z.B. folgendermaßen in einer Tabelle zusammenstellen, wobei wir uns hier auf insgesamt 10 Kinder beschränken wollen:

| lfd.Nr. | Geschlecht | Mathe-Note | Anz. Kinder | Größe (cm) |
|---------|------------|------------|-------------|------------|
| 1 | m | 4 | 1 | 155 |
| 2 | w | 3 | 2 | 140 |
| 3 | m | 2 | 4 | 165 |
| 4 | m | 4 | 1 | 155 |
| 5 | w | 1 | 1 | 160 |
| 6 | w | 5 | 1 | 145 |
| 7 | w | 1 | 2 | 150 |
| 8 | w | 3 | 3 | 130 |
| 9 | m | 2 | 2 | 147 |
| 10 | w | 2 | 1 | 153 |

Diese Daten wollen wir nun mit Hilfe von PCS statistisch auswerten. Dazu müssen sie in den Rechner eingegeben werden. Starten wir also das Programm. Es erscheint zunächst der Grundbildschirm mit dem

Menue und drei Beschleunigungsschaltern auf dem Schalterpanel. Lassen Sie den Mauscursor kurz über einem der Schalter verharren, erscheint ein Kurzhinweis zur Funktion des jeweiligen Schalters.



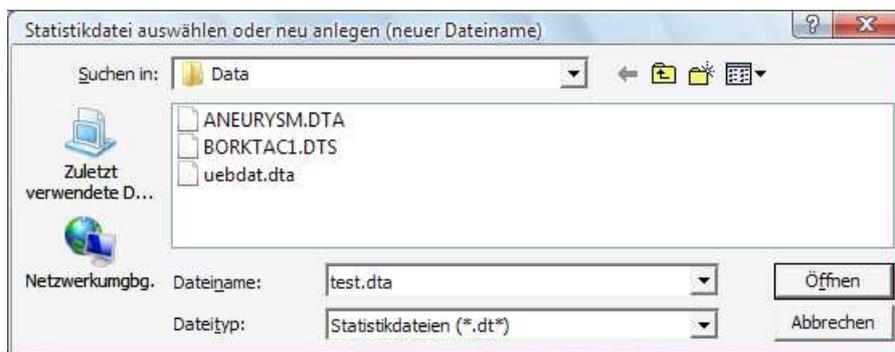
Im linken unteren Bereich werden zwei auf Symbolgröße verkleinerte Fenster angezeigt. Doppelklicken auf eines der Symbole vergrößert das ausgewählte Fenster. Zu sehen gibt es allerdings nicht viel. Im Tabellenfenster steht „Keine Ausgabe“ und das Graphikfenster erfreut den Betrachter durch eine einheitliche graue Fläche. (das ist auch nicht weiter verwunderlich, denn was soll beim momentanen Stand der Dinge schon angezeigt, bzw. ausgegeben werden?)

Öffnen wir jetzt einfach das Datei-Menue und wählen wir daraus den Punkt „Dateneingabe“, wird PCS sehr schnell die Mitarbeit verweigern. Um eine ordentliche Dateneingabe und Speicherung zu gewährleisten, benötigt PCS doch ein gewisses Minimum an Informationen über die zu erwartenden Daten. Dazu gehört, dass Sie zunächst festlegen, unter welchem Namen die Daten gespeichert werden sollen.

Wir wählen daher entweder im Datei-Menue den Punkt „Auswählen“ oder - was viel schneller geht - klicken den Schalter mit Symbol für das Öffnen einer Datei an:



Der Dialog zum Auswählen einer zu bearbeitenden Statistikdatei erscheint:



Hier kann man eine der schon vorhandenen Dateien für die Bearbeitung auswählen. (PCS-Statistikdateien haben alle die Endung .dt*, weswegen auch nur solche Dateien zur Auswahl angeboten werden.) Wir wollen momentan aber keine existierende Datei bearbeiten, sondern eine ganz neue Datei anlegen. Dazu geben wir im Editierfeld unter der Dateinamenliste einfach den gewünschten Namen ein, z.B. test.dta. (Wir müssen uns natürlich an die oben erwähnte Regel halten, dass Statistikdateien eine Namensendung der Form .dt* besitzen.)

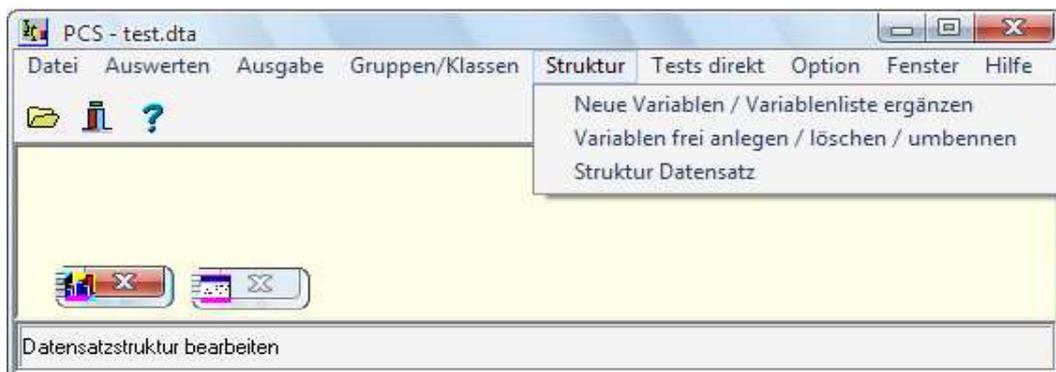
Die Datei test.dta existiert im Moment nicht. Darauf weist uns PCS unmittelbar hin und fragt zur Sicherheit nach, ob test.dta auch tatsächlich neu angelegt werden soll.



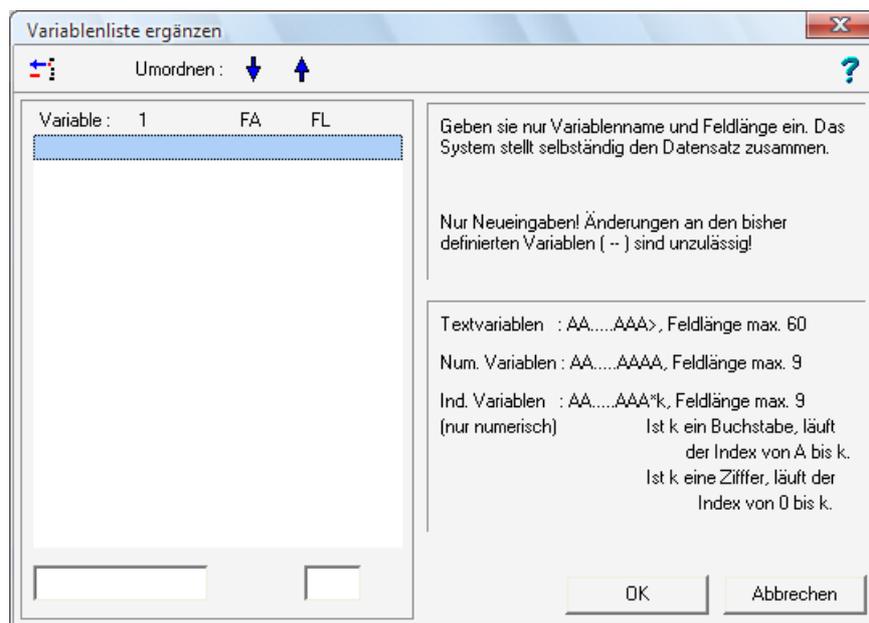
2.2 Definition der Dateistruktur

Damit ist zunächst einmal die neue Datei angelegt. Aber auch jetzt können wir noch nicht mit der Dateneingabe beginnen. Schließlich fehlen PCS immer noch jegliche Angaben über die Struktur der Daten. Diese Angaben muss man dem System einmalig „beibringen“. Sie werden dann in einer zusätzlichen Parameterdatei gespeichert, deren Namen sich aus dem Hauptteil des Dateinamens und dem Zusatz (Extension) .prm zusammensetzt. Um bei unserem Beispiel zu bleiben: Die Strukturangaben werden in einer Datei test.prm gespeichert. Wenn Sie Sicherheitskopien Ihrer Datenbestände anlegen, sollten Sie daher sicherstellen, dass auch die Parameterdatei test.prm gesichert wird.

Da es um die Festlegung der Datenstruktur geht, liegt es nahe, das Struktur-Menue zu öffnen:



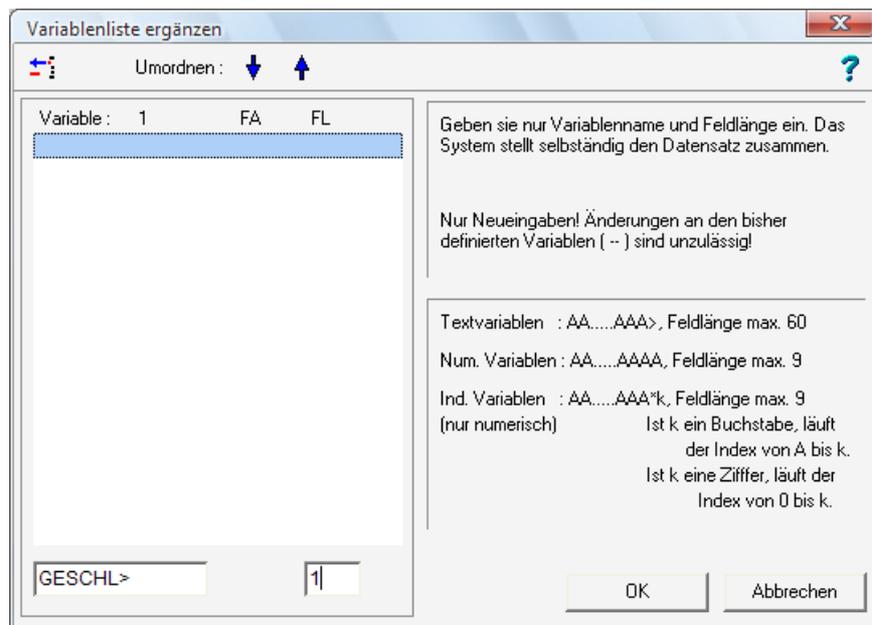
Hier wählen wir den ersten Punkt aus : „Neue Variablen/Variablenliste ergänzen“, denn wir befinden uns ja noch ganz am Anfang unserer Arbeit und müssen die Liste der Variablen erst neu anlegen. Es erscheint anschließend der Dialog zum Eintragen der Variablen mit ihren zugehörigen Strukturangaben:



Unter PCS können Sie die Variablen, also die Spalten Ihrer Ausgangstabelle, mit Abkürzungen belegen, die maximal 10 Zeichen lang sein dürfen, mit einem Buchstaben beginnen müssen (keine Umlaute) und weder Komma noch Doppelpunkt noch eckige Klammern beinhalten dürfen. Sie sollten für Ihre Variablen sinnvolle und leicht verständliche Abkürzungen festlegen, die dieser Regel entsprechen.

Zusätzlich ist noch folgender Punkt zu beachten: Primär wird in der Statistik mit numerischen Größen gearbeitet. In der Praxis lassen sich viele Sachverhalte aber häufig durch eine alphanumerische Verschlüsselung (Text) verständlicher darstellen. So haben wir in unseren Daten das Geschlecht nicht mit 1 und 2 verschlüsselt (wie in der Medizin häufig verwendet), sondern mit m und w, was unmittelbar zu verstehen ist. Damit PCS solche Variablen nicht wie Zahlenangaben behandelt, muss der Name einer Textvariablen mit einem >-Zeichen enden.

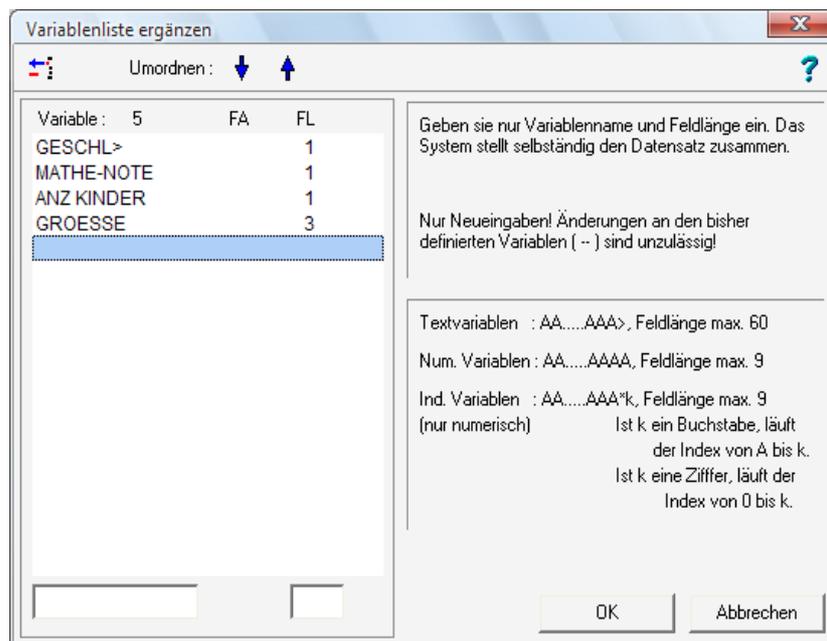
Geben wir für unser Beispiel also zunächst die Variable GESCHL> vor, indem wir diesen Namen einfach in das Editierfeld am unteren Rand eintragen.



PCS benötigt nicht nur diese Abkürzung, sondern auch eine Angabe darüber, wie viel Stellen für jede Variable zu reservieren sind. Das Editierfeld für die Feldlänge (FL) befindet sich direkt neben dem Namensfeld, und mit der Eingabe- oder Tab-Taste gelangen wir dorthin.

Wir geben nun für die Feldlänge eine 1 ein. Die anschließende Betätigung der Eingabetaste bewirkt die Übernahme der so definierten Variablen in die Variablenliste im oberen Teil (sofern kein Fehler festgestellt wird) und bereitet die Eingabefelder durch Löschung auf die nächste Variable vor.

So geben wir nun der Reihe nach die Definitionen für alle Variablen vor:



Die Eingabe der Variablen wird durch Klicken auf den OK-Schalter beendet. **Wichtig hierbei:** Nachdem man Bezeichnung und Feldlänge für die letzte Variable eingegeben hat, muss erst noch die Eingabetaste betätigt werden, damit diese Vorgaben auch in die Variablenliste übernommen werden. Erst dann sollte man den Dialog beenden.

Das Ergebnis unserer Bemühungen können wir uns auch sofort anzeigen lassen. Dazu wählen wir im Strukturmenue den Punkt „Struktur Datensatz“. Es wird nun die komplette Datensatzstruktur präsentiert:



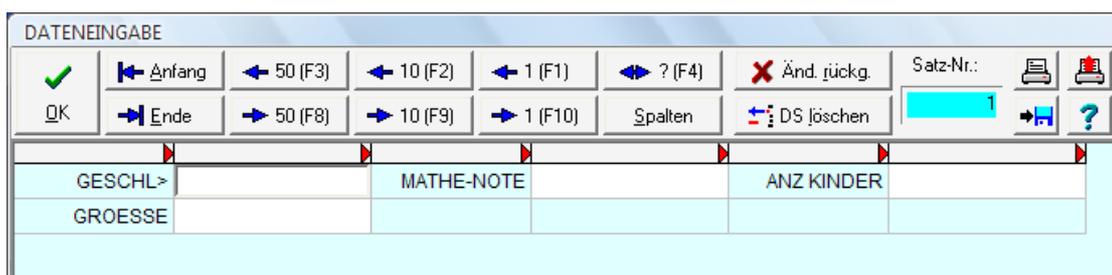
| Bezeichnung | FA.FL | Bezeichnung | FA.FL | Bezeichnung | FA.FL |
|-------------|-------|-------------|-------|-------------|-------|
| GESCHL> | 1.1 | ANZ KINDER | 3.1 | | |
| MATHE-NOTE | 2.1 | GROESSE | 4.3 | | |

Bei den Angaben der Form x.x, die zu jeder Variablen aufgeführt sind, handelt sich um Feldanfang und Feldlänge. 1.1 bei der Variablen GESCHL> bedeutet demnach, dass das Geschlecht im Datensatz ab Position 1 mit einer Länge von einem Zeichen gespeichert wird. Die Mathematiknote wird, da unmittelbar auf GESCHLECHT folgend, ab Stelle 2 abgelegt. Auch für diese Variable reicht ein Feld der Länge 1 aus.

Man erkennt also, wie PCS alle Variablen hintereinander anordnet. In ca. 98% aller Anwendungsfälle ist aber dieses Detailwissen für den Benutzer unerheblich, da er fast immer nur mit den Variablenbezeichnungen arbeitet.

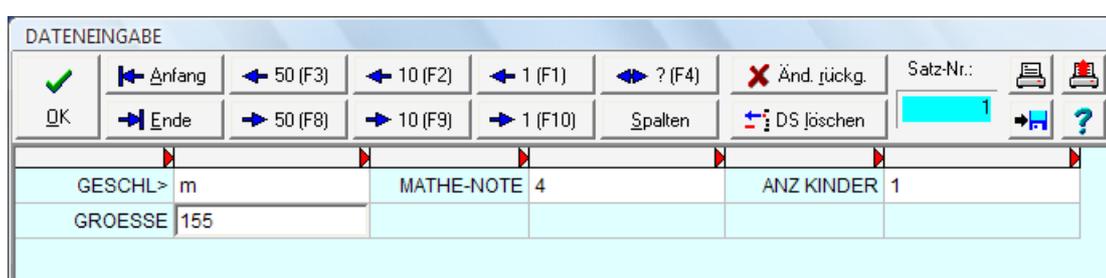
2.3 Dateneingabe

Nach den in 2.1 und 2.2 beschriebenen Vorarbeiten können wir endlich zur Dateneingabe schreiten. Dazu rufen wir im Dateimenue den Punkt „Dateneingabe“ auf und gelangen in den entsprechenden Editor:



| GESCHL> | MATHE-NOTE | ANZ KINDER |
|---------|------------|------------|
| m | | 1 |
| GROESSE | | |

Die Bedienung des Editors ist denkbar einfach. Man tippt einfach der Reihe nach die Werte ein, wobei man mit der Tab-Taste oder mit Return ein Feld weitergelangt. Mit Shift-Tab geht man ein Feld zurück, mit den Cursor- und Bild-Steuertasten kann man sich im Datensatz bewegen. Für unser Beispiel sieht die Eingabe des ersten Datensatzes also folgendermaßen aus:



| GESCHL> | MATHE-NOTE | ANZ KINDER |
|---------|------------|------------|
| m | 4 | 1 |
| GROESSE | 155 | |

Haben Sie das letzte Feld eines Datensatzes ausgefüllt und Return betätigt, wird automatisch zum nächsten Datensatz weitergeschaltet. Eine neue leere Maske steht wieder zur Verfügung, und die Datensatznummer (rechts oben) hat sich um 1 erhöht.

Innerhalb der Datei bewegt man sich mit Hilfe der Funktionstasten oder durch Betätigen der entsprechenden Schalter, wobei man um 1, 10 oder 50 Datensätze nach vorn oder nach hinten weiterblättern kann. Zusätzlich kann man direkt auf das Ende oder den Anfang der Datei positionieren.

Will man z.B. einer bestehenden Datei neue Daten hinzufügen, bewegt man sich erst zum letzten Datensatz und anschließend noch einen Satz weiter, um eine leere Eingabemaske zu erhalten.

Man hat auch die Möglichkeit, die Datei auf einen speziell vorgegebenen Datensatz zu positionieren. Dazu klickt man den Schalter



an oder betätigt die F4-Taste. Sie werden anschließend aufgefordert, die gewünschte Satznummer anzugeben.

Sie können ganze Datensätze löschen oder Änderungen in einem Eingabefeld oder im ganzen Datensatz rückgängig machen. Beendet wird die Eingabe über den OK-Schalter.

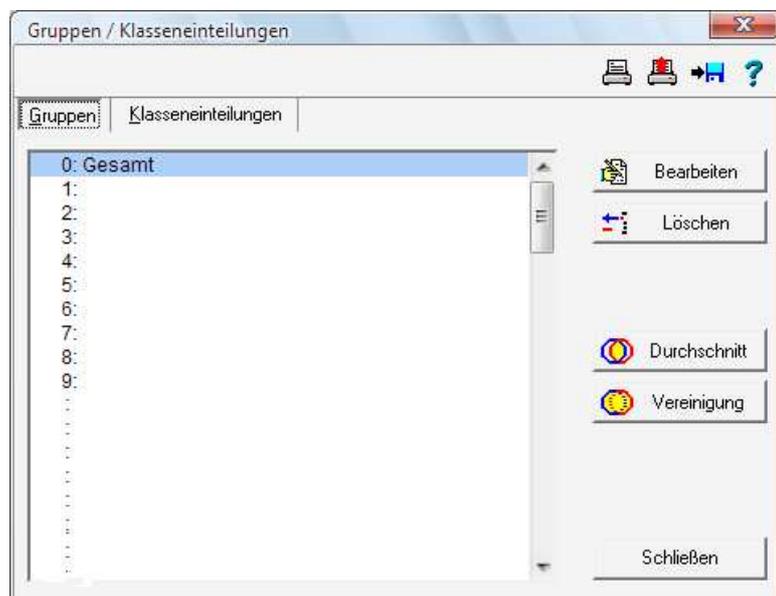
Üblicherweise präsentiert PCS die Eingabefelder in einer dreispaltigen Anordnung. Über den Schalter „Spalten“ können Sie aber auch eine ein- oder zweispaltige Dateneingabemaske realisieren. Die Breite der Bezeichnungs- und Eingabefelder können Sie auch nach Belieben ändern, indem Sie die Trennlinien der Spalten oberhalb des Editorbereichs mit der Maus anklicken und anschließend bei gedrückter Maustaste verschieben.

Gewisse Standardkontrollen bei der Dateneingabe erfolgen automatisch, so z.B. die Überprüfung, ob ein Eintrag in ein als numerisch gekennzeichnetes Feld auch tatsächlich numerisch ist. Individuelle Eingabekontrollen ermöglicht der Punkt „Eingabekontrolle“ im Dateimenu.

2.4 Gruppenbildung

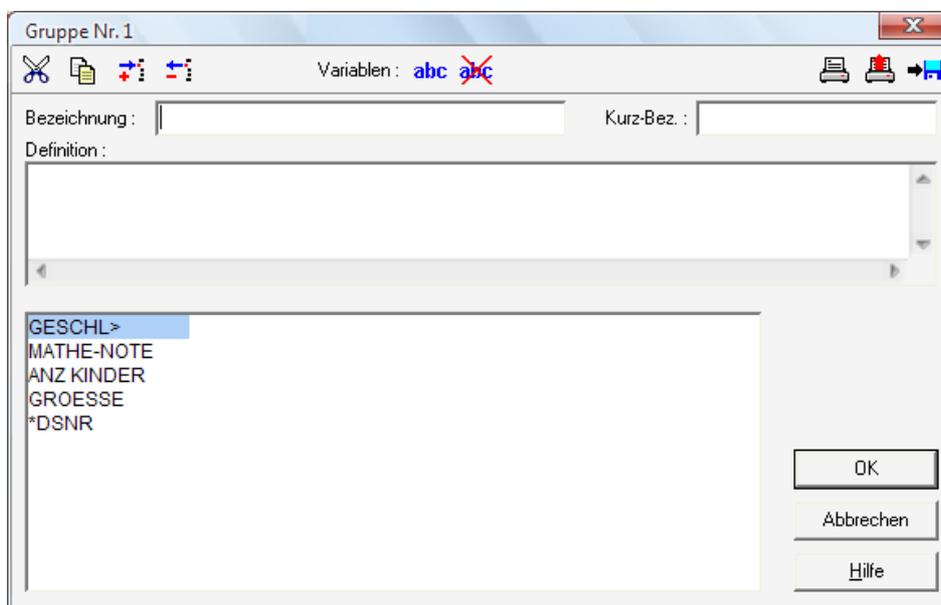
Bevor man mit der Auswertung beginnt, sollte man erst einmal den Computer abschalten und sich in einer „stillen Stunde“ ein Auswertungskonzept zurechtlegen. PCS kann Ihnen zwar die statistischen Berechnungen abnehmen, es kann Sie aber nicht der Mühe entheben, sinnvolle Fragestellungen in Bezug auf Ihr Datenmaterial zu entwickeln. Gehen wir einmal davon aus, dass uns interessiert, wie die mittlere Körpergröße bei Jungen und Mädchen aussieht und ob hier statistisch signifikante Unterschiede zu verzeichnen sind.

PCS hat zum momentanen Zeitpunkt allerdings noch keinerlei Information darüber, wie Jungen und Mädchen zu unterscheiden sind. Das Programm muss also erst in die Lage versetzt werden, die Gruppe der Jungen und die Gruppe der Mädchen erkennen zu können. Dazu muss PCS vermittelt werden, nach welchen Kriterien ein Datensatz der einen oder der anderen Gruppe zugeordnet wird. Die Möglichkeiten hierzu bietet PCS, wenn wir in der Hauptmenseleiste den Punkt „Gruppen/Klassen“ anwählen. Es erscheint der folgende Dialog:



Er enthält eine Liste der bisher bekannten Gruppen. Belegt ist zum jetzigen Zeitpunkt nur die Gruppe 0, die mit „Gesamt“ bezeichnet ist und alle Datensätze umfasst. PCS erwartet die Auswahl der zu bearbeitenden Gruppennummer. Insgesamt können die Gruppennummern 1 bis 255 frei belegt werden.

Wir entschließen uns, dem System beizubringen, wie es männliche Schüler erkennt. Diese Gruppe „Jungen“ soll unter der Nummer 1 gespeichert werden. Dazu klicken wir die Zeile mit der Nummer 1 doppelt an, oder wir wählen die Gruppe 1 durch einfaches Anklicken aus und betätigen anschließend den Schalter „Bearbeiten“. Damit gelangen wir in den Dialog-Editor für die Gruppeneinstellung:



Der Dialog stellt Editierfelder für die Bezeichnung der Gruppe (max. 35 Zeichen) und eine Kurzbezeichnung (max. 10 Zeichen) zur Verfügung. Letztere ist erforderlich, da nicht immer die vollständige Gruppenbezeichnung in den Ergebnistabellen Platz hat.

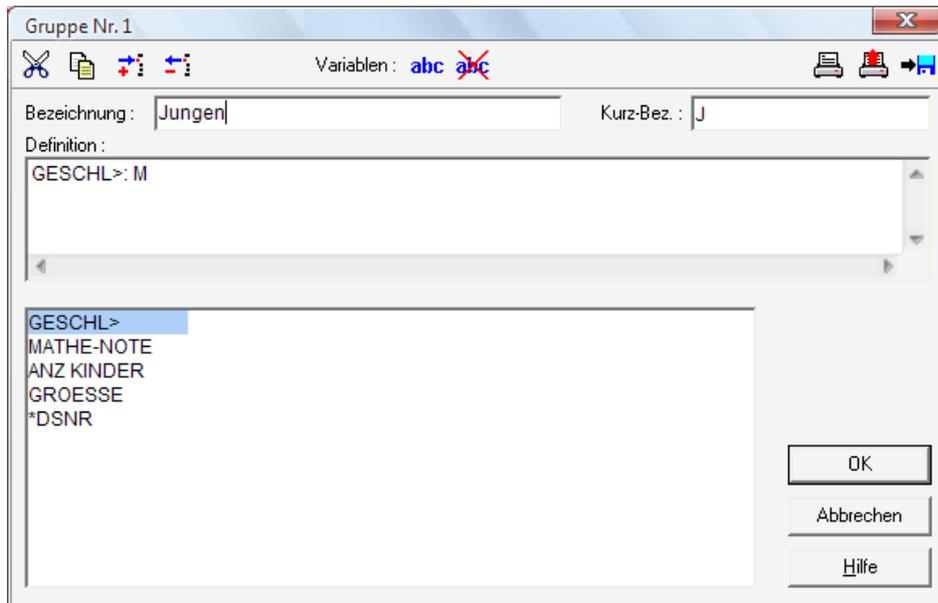
Im mittleren Bereich befindet sich der Editor zur Bearbeitung der genauen Definition der Gruppe. Dabei muss man die bisher vereinbarten Variablen verwenden, die daher im unteren Teil vollständig angezeigt werden. Die Variablen sind zunächst in der gleichen Reihenfolge aufgelistet, wie sie auch in der Eingabemaske erscheinen.

Mit dem Schalter  bewirkt man eine alphabetische Sortierung, mit  wird die ursprüngliche Reihenfolge wieder hergestellt.

Die Variablenliste enthält zusätzlich immer die Variable *DSNR, die automatisch erzeugt wird. *DSNR enthält die laufende Datensatznummer in der Datei. Diese wird aber nicht im Datensatz gespeichert, da sich bei Umordnung der Datei (z.B. durch Sortieren) der Wert von *DSNR ja ändert.

Für unser Beispiel wählen wir als Gruppenbezeichnung „Jungen“ und tragen dies in das Bezeichnungsfeld ein. Als Kurzbezeichnung tragen wir ein schlichtes „J“ ein. Mit Tab gelangen wir in den eigentlichen Definitionsbereich des Editors.

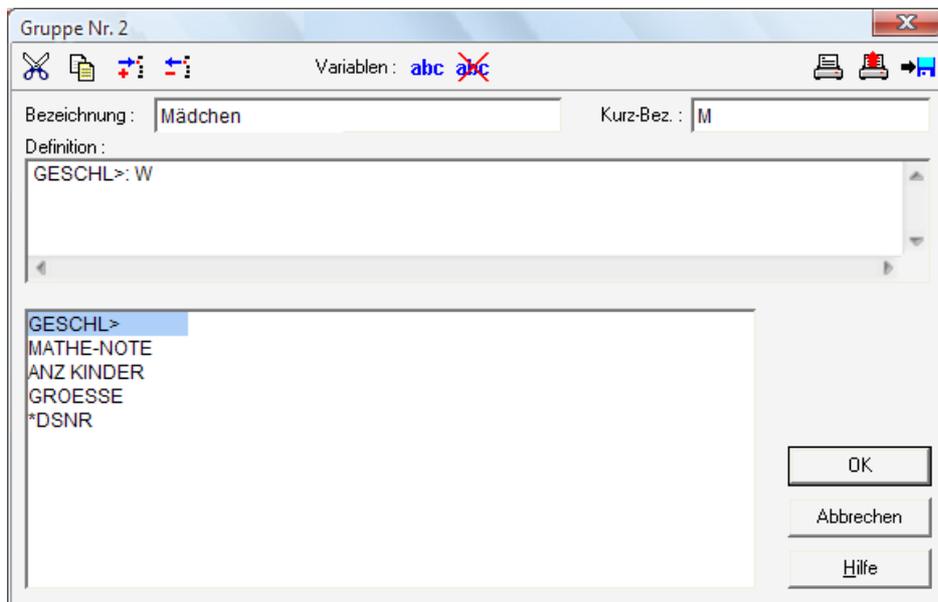
Bei der Gruppenbezeichnung und der Kurzbezeichnung war es gleichgültig, was wir eingegeben haben, handelt es sich doch nur um Texte, die in den Ergebnistabellen und Graphiken ausgegeben werden. Bei der Definition einer Gruppe wird aber die Einhaltung bestimmter Regeln verlangt, so z.B. die Verwendung der vereinbarten Variablennamen, da PCS ansonsten mit „Unverständnis“ reagiert. Nun müssen wir dem System nur noch die Regel beibringen, nach der Jungen zu erkennen sind. Für den Leser der Tabelle ist es ganz klar: Ein Junge ist durch ein m im Feld Geschlecht gekennzeichnet. Genau das müssen wir PCS vermitteln. Dies geschieht auf eine denkbar einfache Weise, nämlich indem wir den Variablennamen und den zugehörigen Inhalt angeben, beides durch einen Doppelpunkt getrennt. Wir tippen also ein:



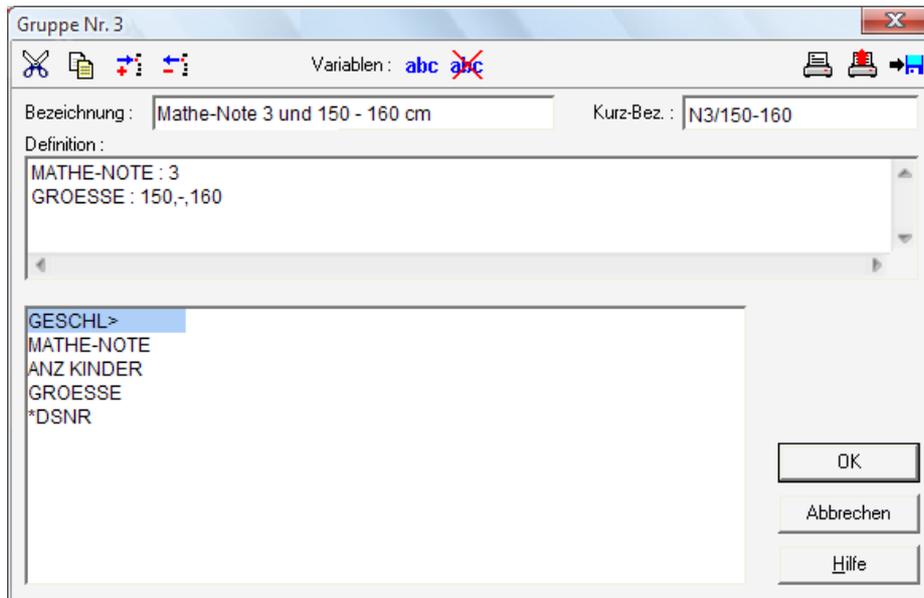
Bei der Gruppengröße werden alle Eingaben direkt in Großbuchstaben umgesetzt. Die Erkennungslogik von PCS unterscheidet nicht zwischen Groß- und Kleinschrift. Es ist also gleichgültig, ob Sie bei der Dateneingabe m oder M im Geschlechtsfeld eingegeben haben.

Wer Tipparbeit sparen will, kann die Variablennamen auch direkt aus der Variablenliste übernehmen: Im Definitionsektor wird die Position markiert, ab welcher der Name erscheinen soll. Anschließend wird der gewünschte Name in der Variablenliste doppelt angeklickt und so in den Definitionsektor übernommen.

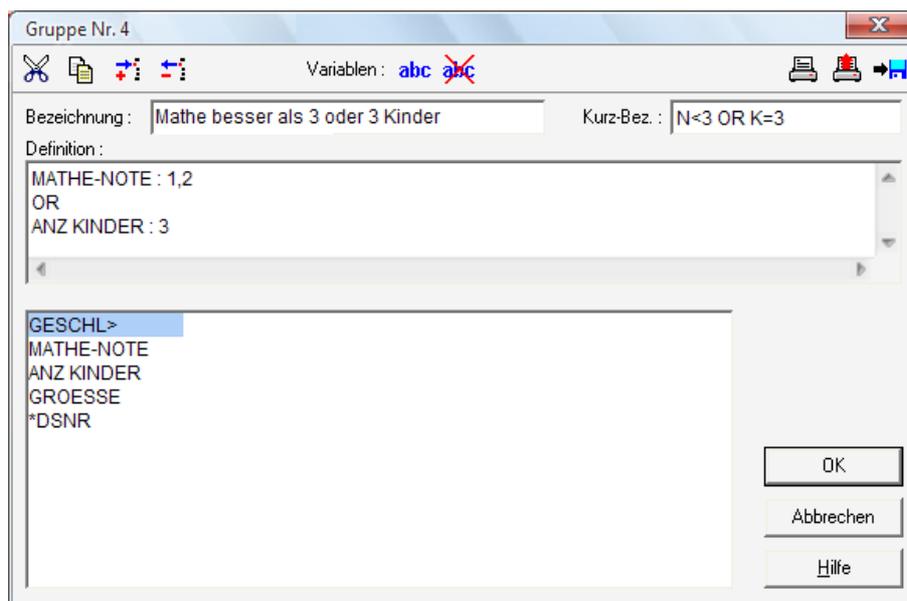
Mit dieser Eingabe ist die Definition für die Gruppe „Jungen“ komplett. Wir schließen mit OK die Eingabe ab. Es erscheint wieder die Gruppenübersicht, wobei jetzt unter Nummer 1 die Gruppe „Jungen“ aufgeführt ist. Erneut wird die Auswahl einer Gruppennummer erwartet. Wählen wir die 2 aus, um die Gruppe „Mädchen“ zu definieren. Es ist unschwer nachzuvollziehen, dass die komplette Eingabe der Gruppengröße folgendermaßen aussehen muss:



Weil wir gerade dabei sind, wollen wir uns gleich im Anlegen komplexerer Gruppen üben. Definieren wir unter der Nummer 3 die Gruppe der Schüler, welche in Mathematik die Note 3 haben und zwischen 150 und 160 cm (inclusive) groß sind. Hier müssen zwei Bedingungen erfüllt sein, die wir im Definitionsbereich des Editors einfach untereinander schreiben:



Die erste Zeile der Definition ist einfach zu verstehen: Im Feld Mathe-Note muss eine 3 stehen. Die Bedingung „Körpergröße zwischen 150 und 160 cm“ legt ein Intervall für die Variable GROESSE fest. Ein solches Intervall wird bei der Eingabe durch untere und obere Grenze, getrennt durch „-,“, definiert, womit auch die zweite Eingabe verständlich sein dürfte. Definieren wir nun unter Nummer 4 die Gruppe der Schüler, die in Mathematik besser als 3 abgeschnitten haben, oder bei denen 3 Kinder in der Familie sind. Die komplette Definition sieht wie folgt aus:

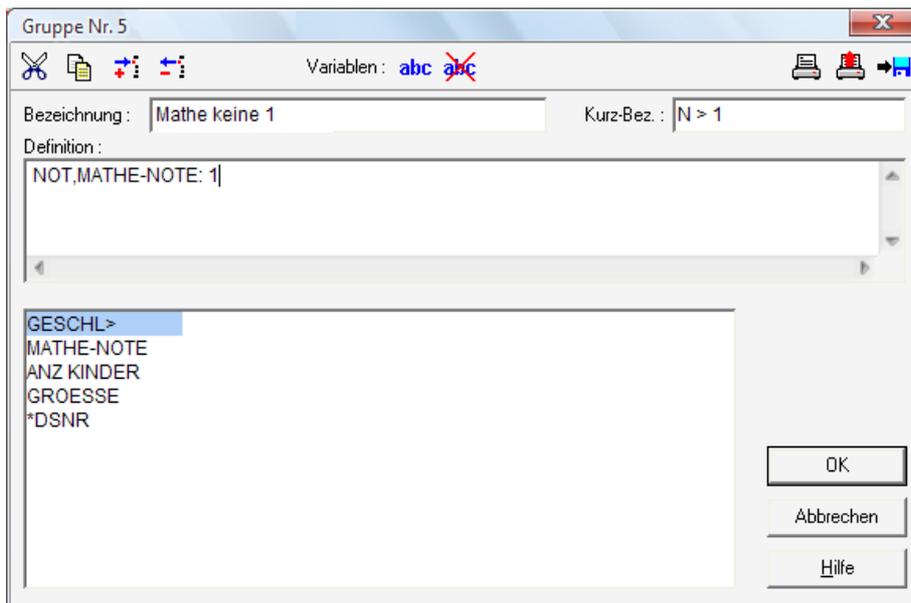


In Mathematik besser als 3 bedeutet, im Feld Mathe-Note eine 1 oder eine 2 stehen zu haben. Genau dies besagt die erste Zeile der Gruppendifinition. Hinter dem Doppelpunkt steht jetzt nicht nur eine Größe und auch kein Intervall, sondern es werden einfach, durch Komma getrennt, die zulässigen Werte in diesem Feld aufgezählt. Die Eingabe der zweiten Bedingung ist sofort zu verstehen. Im Unterschied zum vorherigen Beispiel sollen aber jetzt nicht beide Bedingungen gleichzeitig erfüllt sein, sondern ein Schüler soll zur Gruppe 4 gehören, wenn er die erste *oder* die zweite Bedingung (*oder beide*) erfüllt. Deswegen stehen die beiden Bedingungen hier nicht einfach untereinander, sondern sie sind mit einem in einer Extrazeile aufgeführten **OR** verbunden.

Noch ein letztes Beispiel: Unter Nummer 5 wollen wir die Gruppe der Schüler definieren, die in Mathematik keine 1 haben. Es dürfte klar sein, dass dies gleichbedeutend ist mit der Bedingung eine 2, 3, 4, 5 oder 6 im Feld Mathe-Note stehen zu haben. Man könnte also einfach als Definition

MATHE-NOTE:2,3,4,5,6

eingeben. Es geht aber auch anders, indem ein vorangestelltes NOT zur Negation benutzt wird:



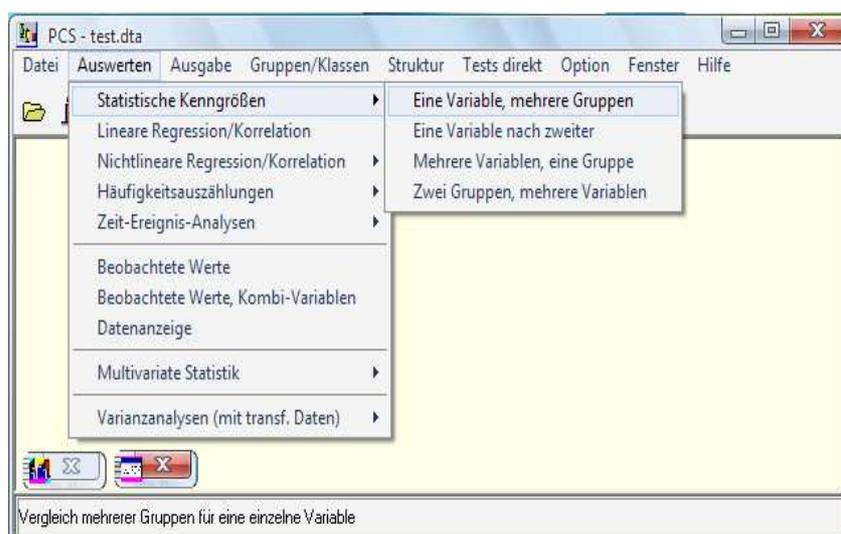
Dies möge zunächst als Einstieg genügen. Komplexere Methoden der Gruppenbildung, z.B. über Vereinigung oder Durchschnitt zweier Gruppen werden ausführlich in der Online-Hilfe erläutert.

Wir beenden die Gruppendefinitionen, indem wir den Schalter „Schließen“ betätigen und landen wieder auf der PCS-Oberfläche.

2.5 Auswertung: Statistische Kenngrößen

Nachdem die notwendigen Vorarbeiten hinsichtlich der Gruppendefinitionen erledigt sind, wollen wir zur ursprünglichen Aufgabenstellung zurückkehren, nämlich dem Vergleich von Jungen und Mädchen bezüglich der Körpergröße.

Wählen wir dazu in der Hauptmeneuleiste den Punkt „Auswerten“ an, werden im zugehörigen Menue die verfügbaren Analysemöglichkeiten aufgeführt.



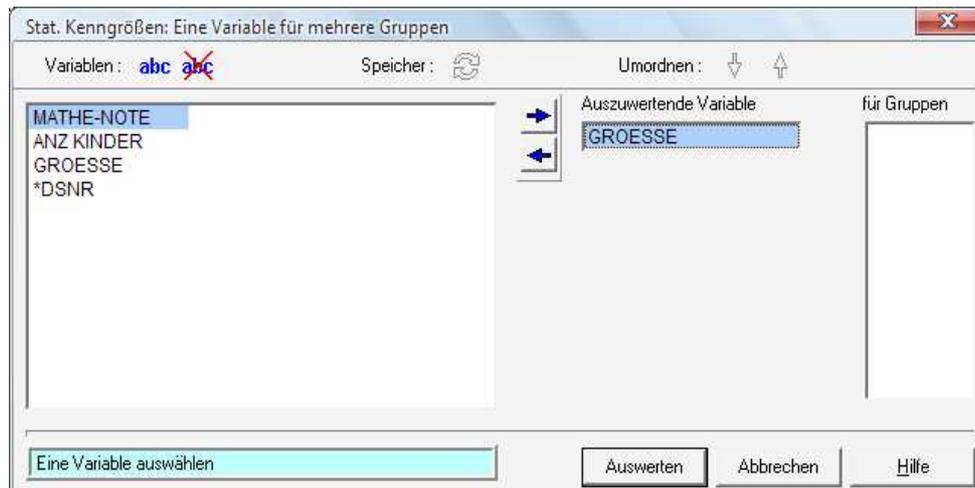
Aus der Vielzahl von Auswertungsmöglichkeiten, die sich auch noch in Untermenues auffächern, wählen wir nun zunächst den Punkt „Statistische Kenngrößen“ und im darauf erscheinenden Untermenue den Punkt „Eine Variable, mehrere Gruppen“. Ziel ist es ja, für eine Variable, nämlich GROESSE, die statistischen Kenngrößen

wie Mittelwert, Standardabweichung, etc., für die zwei Gruppen „Jungen“ und „Mädchen“ zu ermitteln, um auf dieser Basis einen entsprechenden statistischen Test durchzuführen.

Generell sind bei jeder Auswertung einige Fragen zu klären:

- Welche Variablen sind einzubeziehen?
- Welche Klasseneinteilung ist evtl. bei Auswertung der Variablen zu berücksichtigen (s. 2.6)?
- Wie müssen die Variablen eventuell transformiert werden (wichtig z.B. bei Varianzanalysen)?
- Welche Gruppen müssen einbezogen werden?

Nach Auswahl des Auswertungsverfahrens öffnet sich ein Dialogfenster, in welchem die für die jeweilige Methode spezifischen Parameter auszuwählen sind:

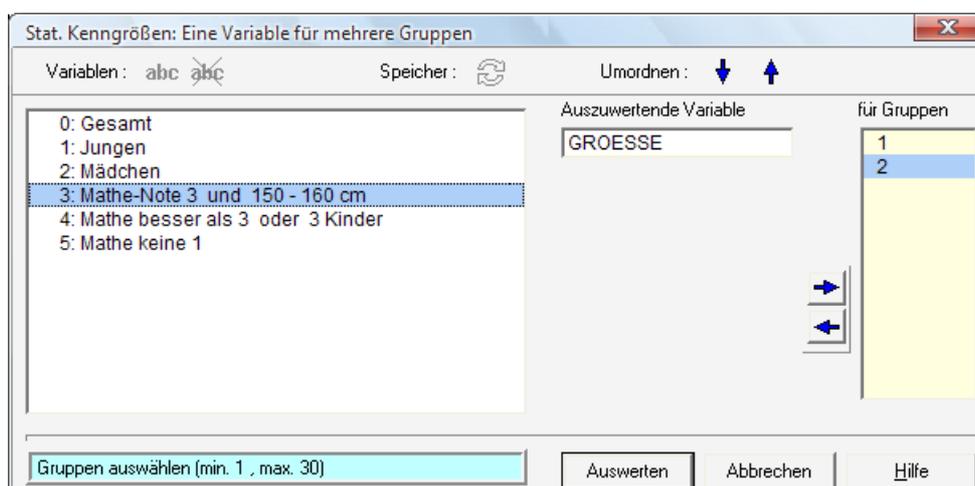


Das Fenster für den Auswertungsdialog enthält im linken Teil eine Liste, dessen Inhalt von der Methode abhängt und davon, welcher Bereich im rechten Teil angewählt ist. So enthält die Auflistung der Variablen hier keine Textvariablen, d.h. GESCHL> wird überhaupt nicht zur Auswahl angeboten, denn schließlich macht es wenig Sinn, Mittelwert und andere Kenngrößen aus Texten zu berechnen.

Üblicherweise wird mit der Auswahl der auszuwertenden Variablen begonnen. Sie können aber auch den Bereich Gruppen im rechten Teil anklicken. Sofort werden in der Liste im linken Teil die Gruppen zur Auswahl aufgeführt und das Panel mit den Schaltern zum Übertragen der Items wird neben dem Gruppenbereich positioniert.

Bleiben wir bei der Auswahl der Variablen. Unter der gewählten Methode kann nur eine einzelne Variable ausgewertet werden. Wir markieren zunächst die Variable GROESSE. Durch Betätigen des Schalters „Pfeil nach rechts“ wird die Variable nun in den rechten Teil in das Feld für die Variable übertragen. Man kann eine einzelne Variable aber auch durch Doppelklick auswählen und übertragen.

Da nur eine einzige Variable auszuwerten ist, fokussiert PCS jetzt automatisch im rechten Teil auf das Gruppenfeld. (Bei einer Auswahl von mehreren Variablen muss der Benutzer dies eigenhändig erledigen, indem er entweder den Gruppenbereich anklickt oder auf dem Schalterpanel den Schalter mit dem Häkchen betätigt.)



Nun werden die zu berücksichtigenden Gruppen ausgewählt, wobei man auf verschiedene Arten vorgehen kann. Wir können durch Doppelklicken auf Gruppe 1 erst diese übernehmen, anschließend durch Doppelklicken auf Gruppe 2 die nächste.

Oder Sie markieren bei gedrückter Maustaste die beiden Gruppen 1 und 2 und übernehmen sie anschließend durch Betätigen des Schalters „Pfeil nach rechts“.

Haben Sie sich bei der Auswahl vertan, können Sie in analoger Weise mit Hilfe des Schalters „Pfeil nach links“ einzelne Items wieder von der Auswertung ausschließen. Zusätzlich haben Sie die Möglichkeit die Auswahl in der Reihenfolge umzuordnen, und zwar bewegen Sie das markierte Element über die Schalter „Pfeil nach unten“, bzw. „Pfeil nach oben“.

Wollen Sie mehrere Auswertungen mit gleichen Variablen oder Gruppen ausführen, können Sie die letzte Auswahl aus einem Speicher abrufen (Schalter „Pfeile im Kreis“) und eventuell übernehmen. Damit ersparen Sie sich z.B. die erneute Auswahl von 30 Variablen oder Gruppen.

Der Auswahldialog wird durch Betätigen von „Auswerten“ abgeschlossen. (Sollten Angaben fehlen oder falsch sein, meldet sich PCS mit einem Warnton und positioniert das nun rot gefärbte Schalterpanel vor dem entsprechenden Auswahlbereich.)

Jetzt startet PCS mit der eigentlichen Auswertung, indem alle gespeicherten Datensätze durchgearbeitet werden, auf Zugehörigkeit zu einer der Gruppen überprüft werden, und eventuell die erforderlichen Rechenoperationen durchgeführt werden. Dabei wird in einem gesonderten Fenster die laufende Nummer des gerade verarbeiteten Datensatzes angezeigt (in 20-er-Schritten), so dass Sie den Fortgang der Arbeit beobachten können. Die wenigen Datensätze unserer Testdatei sind schnell abgearbeitet. Als Ergebnis erscheint auf dem Bildschirm:

| GRUPPE | M | SD | SEM | MEDIAN | N |
|--------|--------|-------|------|--------|---|
| A: J | 155.50 | 7.37 | 3.69 | 155.0 | 4 |
| B: M | 146.33 | 10.52 | 4.29 | 147.5 | 6 |

Der mehrseitigen Tabelle können wir für die beiden Gruppen, die hier nur mit Kurzbezeichnungen angeführt sind, zunächst Mittelwert(M), Standardabweichung(SD), Standardfehler des Mittelwertes(SEM), Median und die Anzahl berücksichtigter Werte(N) entnehmen. Weitere Seiten kann man durch Anklicken der Tabulatoren oberhalb der Tabellen anwählen, wie z.B.:

Stat. Kenngrößen : GROESSE

| GRUPPE | MEDIAN | MIN | 10% | 90% | MAX |
|--------|--------|-----|-------|-------|-----|
| A: J | 155.0 | 147 | 147.0 | 165.0 | 165 |
| B: M | 147.5 | 130 | 130.0 | 160.0 | 160 |

oder

Stat. Kenngrößen : GROESSE

| GRUPPE | MEDIAN | WH UNTEN | 25%(BXU) | 75%(BXO) | WH OBEN |
|--------|--------|----------|----------|----------|---------|
| A: J | 155.0 | 147 | 151.0 | 160.0 | 165 |
| B: M | 147.5 | 130 | 140.0 | 153.0 | 160 |

Bei den Box&Whisker-Plots hängt die Berechnung der effektiven Whiskergrenzen (WHU, WHO) von den Einstellungen im Dialog „Irrtumswahrscheinlichkeiten, Faktoren, Quantile“ ab (s. Kap. 4). Entsprechendes gilt für die Seiten $M \pm f \cdot SD$ (bzw. $M \pm f \cdot SE$):

Stat. Kenngrößen : GROESSE

| GRUPPE | M | M-2*SD | M+2*SD | M-3*SD | M+3*SD |
|--------|--------|--------|--------|--------|--------|
| A: J | 155.50 | 140.76 | 170.24 | 133.39 | 177.61 |
| B: M | 146.33 | 125.29 | 167.37 | 114.77 | 177.89 |

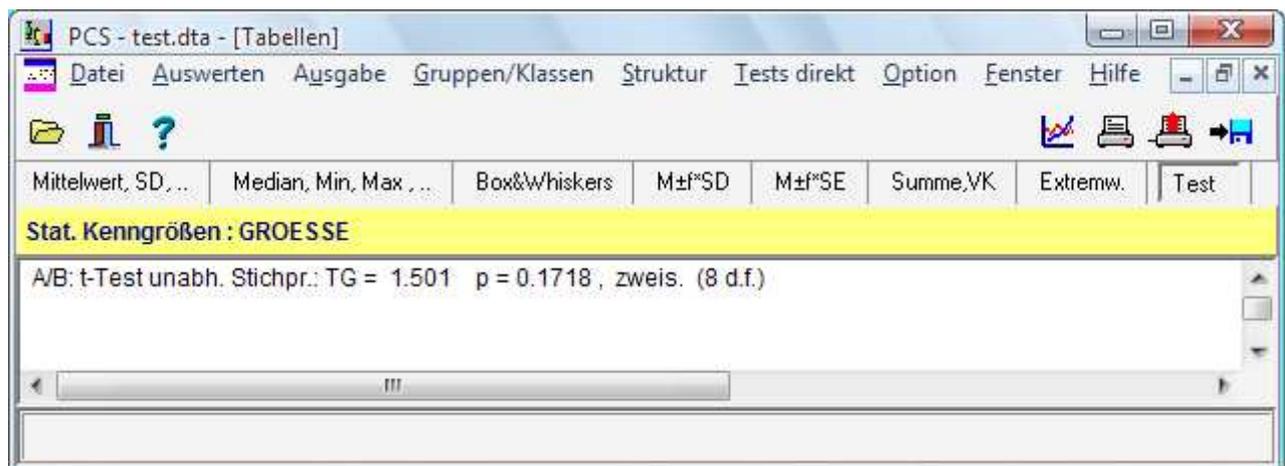
Von besonderem Interesse für unsere Fragestellung ist die Seite „Test“, wo das Ergebnis eines geeigneten Tests zum Vergleich der ausgewählten Gruppen bezüglich der Variablen GROESSE ausgewiesen wird:



Auch der Inhalt dieser Seite ist abhängig von den Einstellungen unter „Irrtumswahrscheinlichkeiten, Faktoren, Quantile“. Als Standardeinstellung bei Neuanlage einer Statistikdatei werden nichtparametrische Tests, d.h. Rangtests als bevorzugte Testmethode gewählt. In diesem Fall ist das also der Wilcoxon-Test für unabhängige Stichproben.

Zu jedem Testergebnis wird das zugehörige p ausgewiesen, sowie ein Hinweis auf die Art der Durchführung (einseitig oder zweiseitig) und falls sinnvoll, die Anzahl der Freiheitsgrade (d.f.).

Hätte man unter „Irrtumswahrscheinlichkeiten, Faktoren, Quantile“ als bevorzugte Testart „parametrische Tests“ und „zweiseitig“ gewählt, wäre der t-Test für unabhängige Stichproben zur Anwendung gekommen, und die Seite „Test“ hätte folgendermaßen ausgesehen:



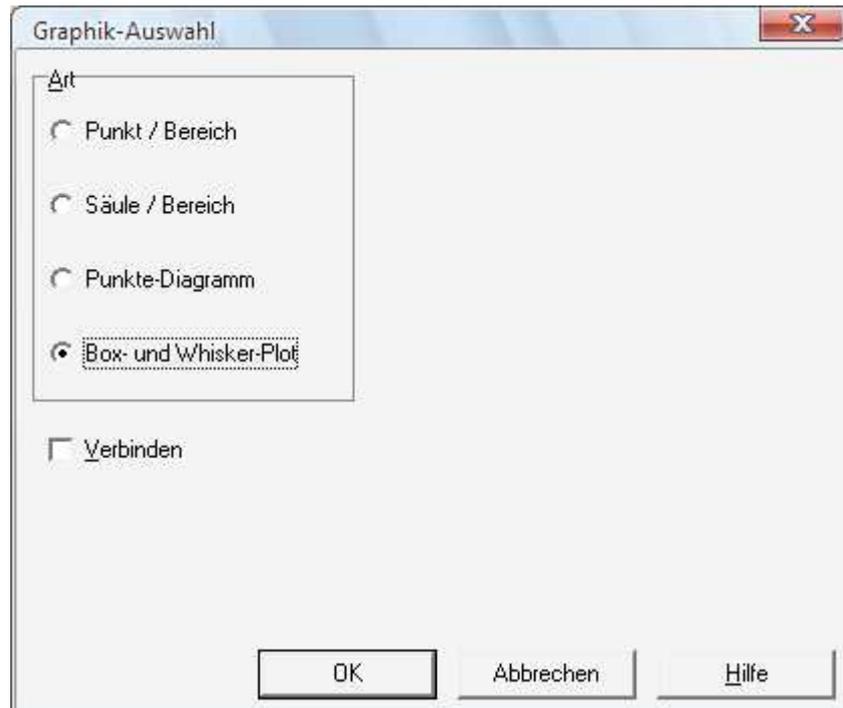
Inzwischen stehen auf der Schalterleiste oberhalb des Tabellenfensters auch weitere Schalter zu Verfügung. Entsprechend ist auch das Ausgabe-Menue um einige Punkte erweitert worden. Sie haben hier die Möglichkeit, die Tabelle jeweils Seite für Seite auszudrucken, wobei der PCS-Druckmanager (s. Kap. 5) für sparsamen Umgang mit dem Papier sorgt. Sie können einen Blattvorschub veranlassen, oder die jeweils aktive Seite in einer Protokolldatei (s. Kap. 6) ablegen.

Ferner haben Sie jetzt die Möglichkeit, eine graphische Präsentation der Ergebnisse abzurufen. Betätigen Sie den entsprechenden Schalter , bzw. wählen Sie den Punkt „Graphik“ im Ausgabe-Menue, öffnet sich ein Dialog zur Auswahl der gewünschten Darstellungsart.

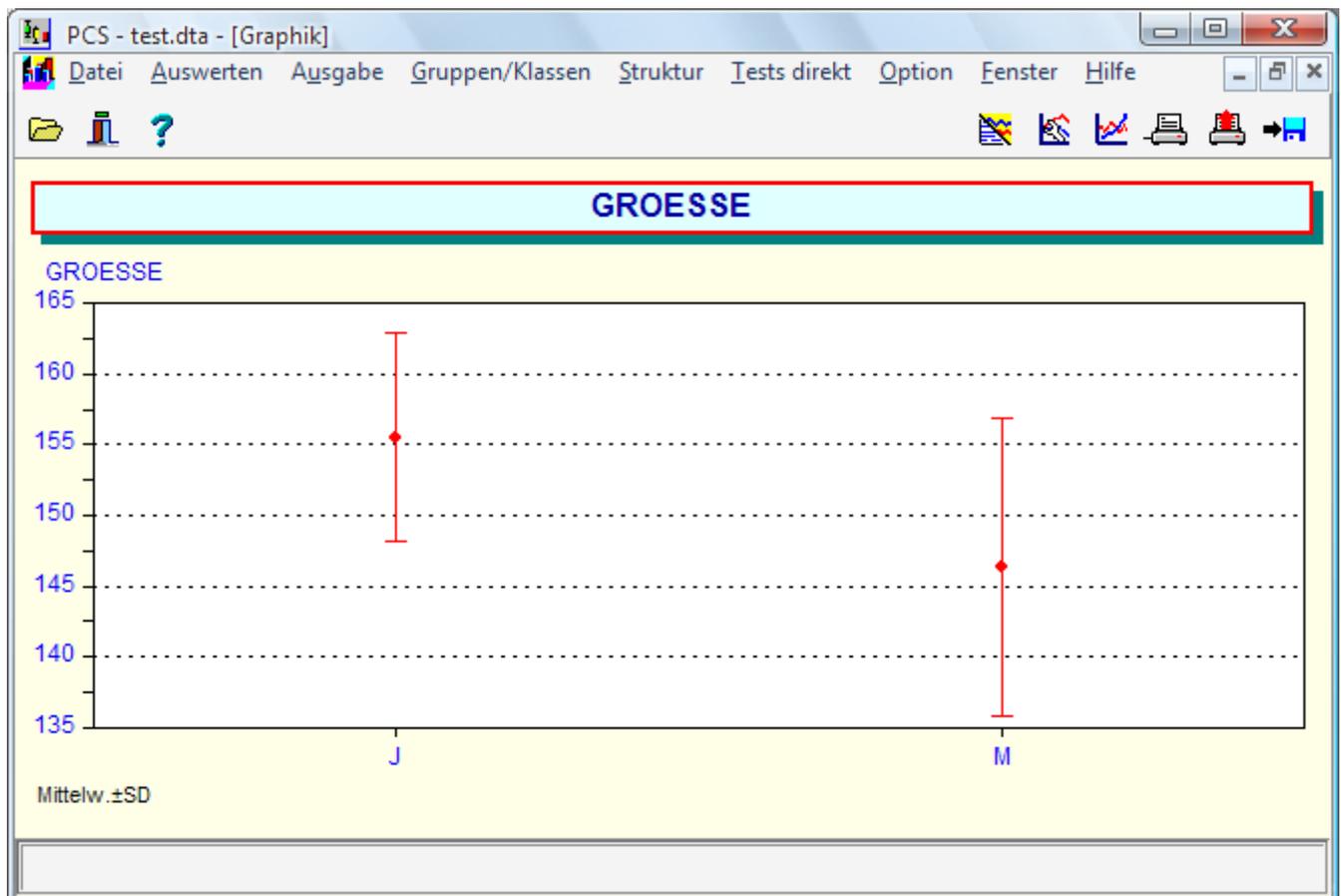
Hier legen Sie im linken Bereich zunächst fest, welchen Bildtyp Sie grundsätzlich wünschen, wobei gewählt werden kann zwischen der Darstellung des Mittelwertes mit einem Streubereich, einem Punktediagramm und Box&Whisker-Plots. Bei Mittelwerten mit Streubereich haben Sie zudem die Auswahl unter mehreren Streubereichen. Zusätzlich können Sie die Mittelwerte miteinander verbinden.

Bei Box&Whisker-Plots müssen Sie noch entscheiden, ob die Mediane oder die Mittelwerte miteinander verbunden werden sollen, falls überhaupt.

Die darzustellenden Streubereiche werden wiederum teilweise beeinflusst durch die Einstellungen im Dialog „Irrtumswahrscheinlichkeiten, Faktoren, Quantile“ (s. Kap. 4).



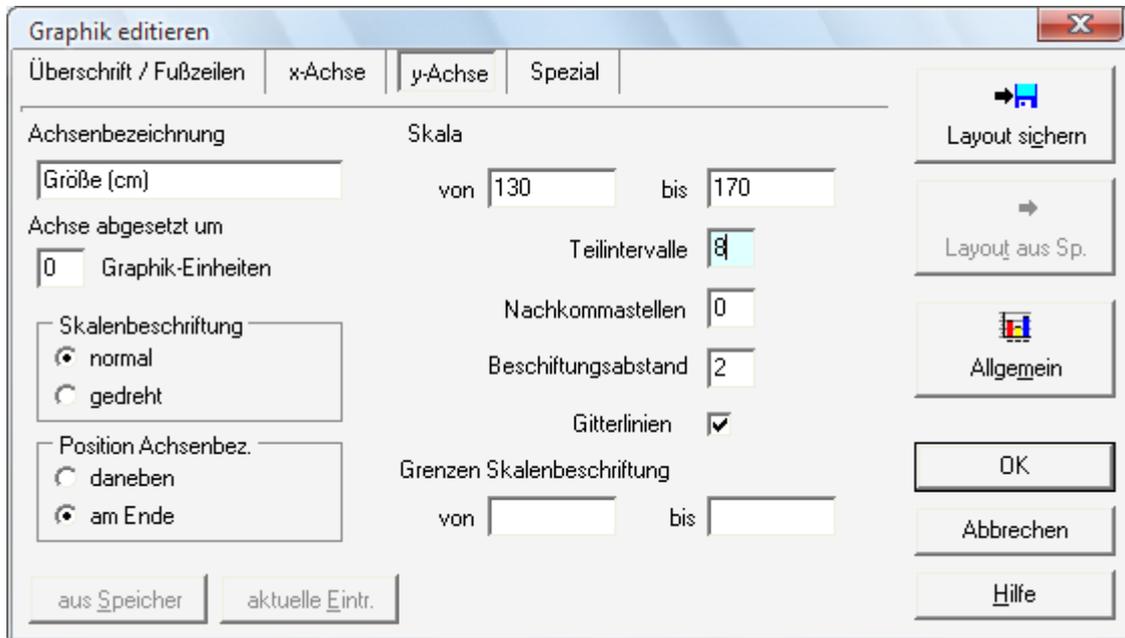
Mit der hier angezeigten Auswahl wird nach Betätigen von OK die folgende Graphik erzeugt:



Die Schalterkonfiguration im rechten Bereich der Schalterleiste hat sich wieder verändert (entsprechend auch das Ausgabemenu). Neben den üblichen Schaltern für Drucken, Blattvorschub und Kopie auf Platte findet man

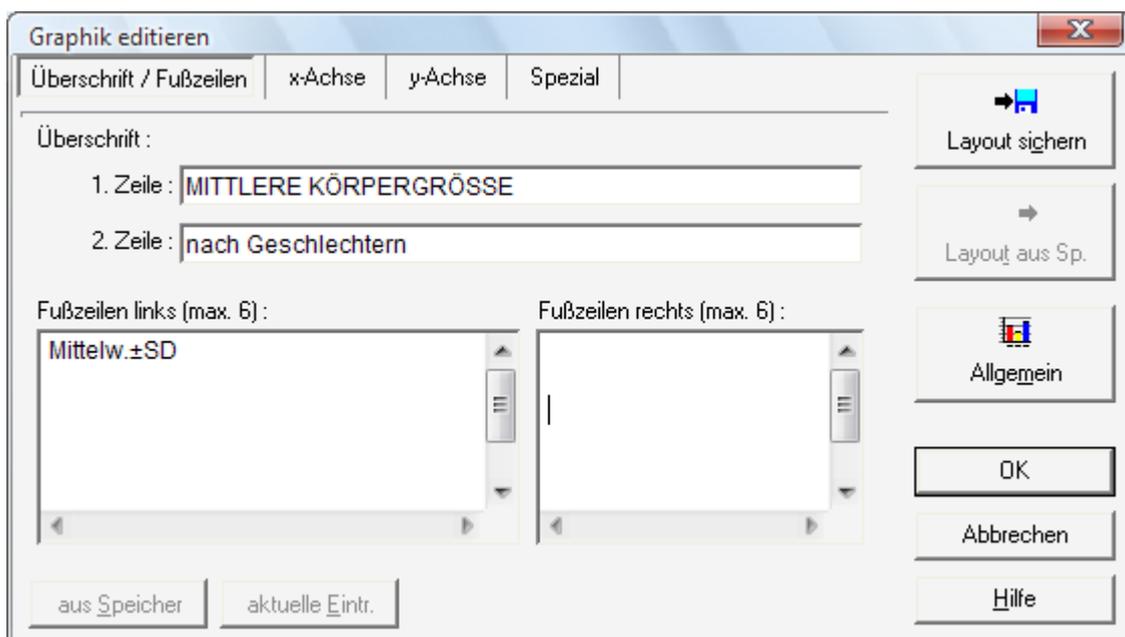
einen weiteren Schalter  zum schnellen Umschalten zwischen Tabellenanzeige und Graphik, sowie einen Schalter, über welchen eine Bearbeitung der Graphik ermöglicht wird. Bei Betätigung dieses Schalters  gelangt man in einen Dialog, aus welchem man die zu bearbeitenden Elemente der Graphik über Tabulatoren auswählen kann:

Wählen wir z.B. die y-Achse zur Bearbeitung aus, weil uns eventuell die von PCS automatisch gewählte Skaleneinteilung nicht gefällt. Im nachfolgenden Dialog können wir dann die y-Achse nach unseren Vorstellungen gestalten:



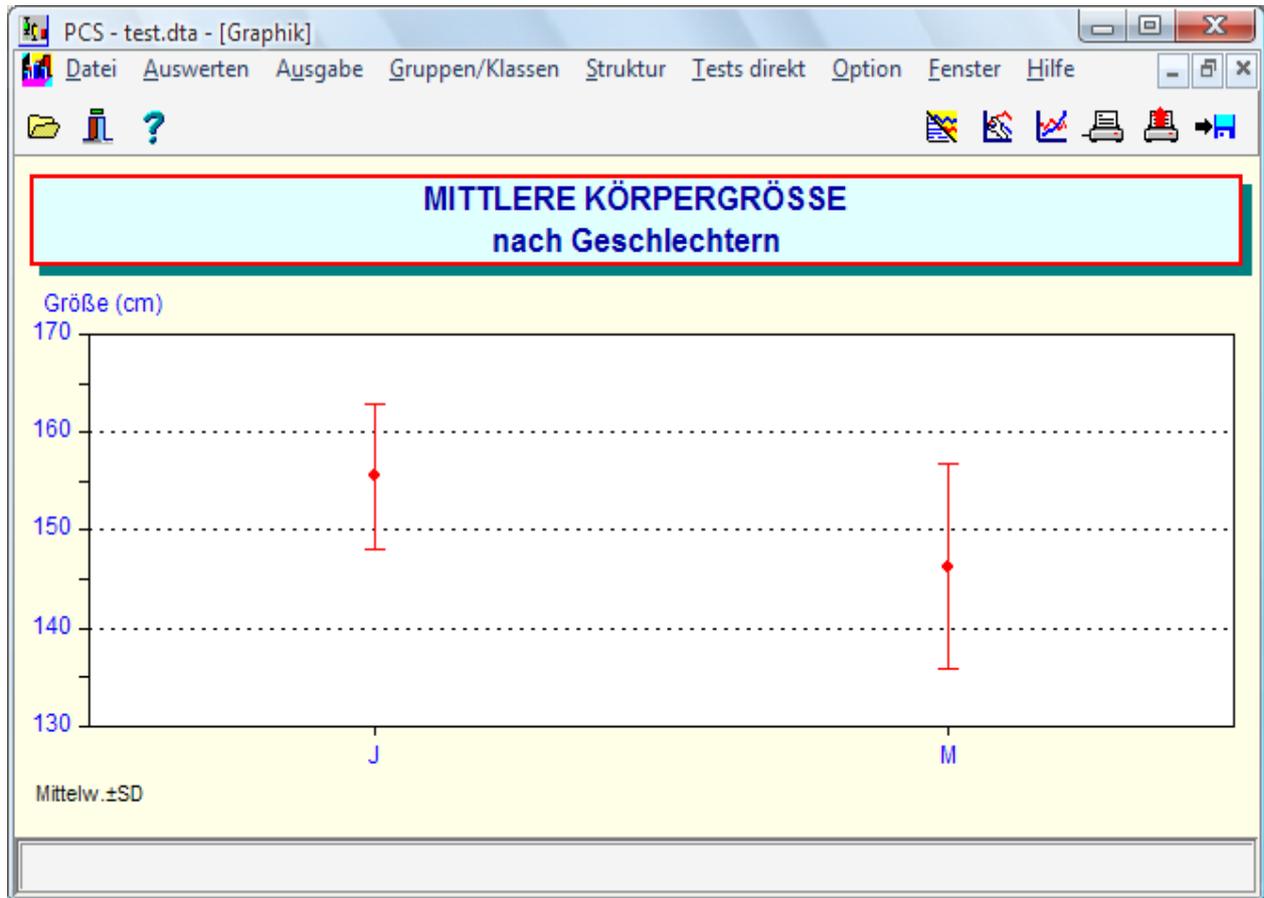
Hier haben wir z.B. einen Darstellungsbereich von 130 – 170 cm gewählt, wobei dieser in 8 Intervalle eingeteilt werden soll. Jeder zweite Teilstrich der Skala soll beschriftet werden. D.h. jedes Teilintervall umfasst 5 cm, die Beschriftung erfolgt in 10 cm-Abständen, also benötigen wir für die Beschriftung keine (0) Nachkommastellen. Die Achsenbezeichnung ist auch noch verändert worden.

Auch die Bildüberschrift soll verändert werden. Dazu wird der Tabulator „Überschrift/Fußzeilen“ im Graphik-Editor-Dialog angeklickt, und im nachfolgenden Dialog die maximal zweizeilige Überschrift eingegeben:



Sollen bestimmte Einstellungen auch bei nachfolgenden Graphiken zur Anwendung kommen, kann man das momentane Layout durch Anklicken des Schalters „Layout sichern“ speichern. Bei weiteren Graphiken (gleicher Art) steht dann im Graphik-Editor-Dialog ein Schalter zur Verfügung („Layout aus Sp.“), über welchen man das gespeicherte Layout wieder laden kann.

Beenden wir den Editor-Dialog mit OK, erscheint nun die neue Graphik:

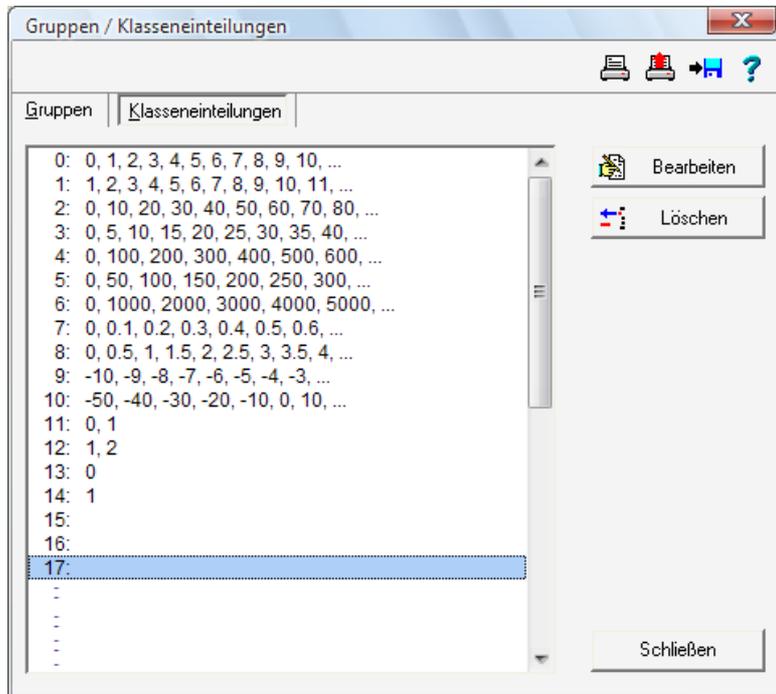


2.6 Klasseneinteilungen definieren

Widmen wir uns einer neuen Aufgabenstellung: Und zwar benötigen wir einen Überblick über die Verteilung der Körpergröße, unabhängig vom Geschlecht der Schüler. Dazu muss als erstes festgelegt werden, welche Intervalle man in Betracht ziehen will. Nehmen wir an, dass wir die folgende Einteilung wünschen:

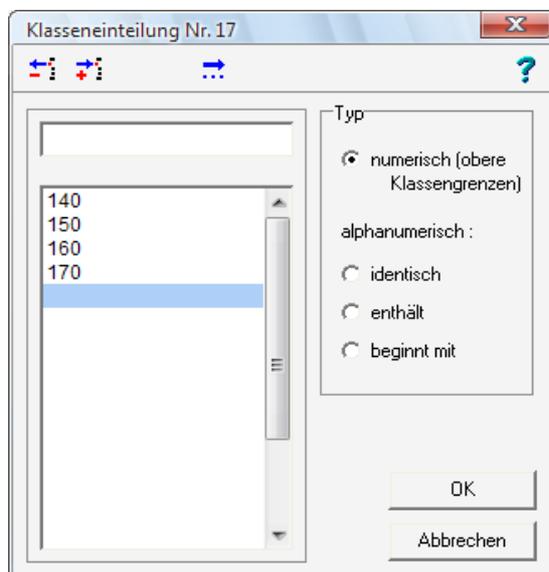
bis 140 cm , 141 - 150 cm , 151 - 160 cm , 161-170 cm , über 170 cm .

Der Statistiker spricht hier von einer Klasseneinteilung, und zur Bestimmung der Häufigkeitsverteilung müssen die Daten entsprechend klassiert werden. Die Aufgabe besteht jetzt nur noch darin, unsere Vorstellung an PCS zu übermitteln, denn von sich aus kann PCS beim besten Willen nicht wissen, welche Wünsche wir hinsichtlich der Klasseneinteilung haben. Dazu wird in der Hauptmenueleiste „Gruppen/Klassen“ und im darauf sich öffnenden Dialog der Tabulator „Klasseneinteilungen“ angeklickt. Es erscheint eine Liste der Klasseneinteilungen (einige generiert PCS selbständig beim Erzeugen einer neuen Datei).



In der Liste sind die maximal 80 Klasseneinteilungen einzeln aufgeführt, und zwar jeweils die oberen Klassengrenzen. Wird die Zeile mit ... fortgesetzt, heißt dies, dass noch weitere Werte in der Reihe folgen. Die hier aufgeführten und von PCS schon generierten Klasseneinteilungen sind jederzeit durch Einteilungen eigener Wahl ersetzbar. (Ausnahme: Klasseneinteilung 0, die nicht verändert werden kann.)

Ähnlich wie bei der Definition von Gruppen müssen Sie nun die zu bearbeitende Klasseneinteilung auswählen. Wir entscheiden uns für die Belegung einer freien Klasseneinteilung, wählen die Nummer 17 aus (Doppelklick oder einfacher Klick und Schalter „Bearbeiten“) und gelangen in den Dialog zur Bearbeitung von Klasseneinteilungen:



An der Typ-Einstellung ändern wir nichts, denn die Klasseneinteilung ist für die Auswertung einer numerischen Variablen gedacht, also sind die Klassengrenzen auch numerisch.

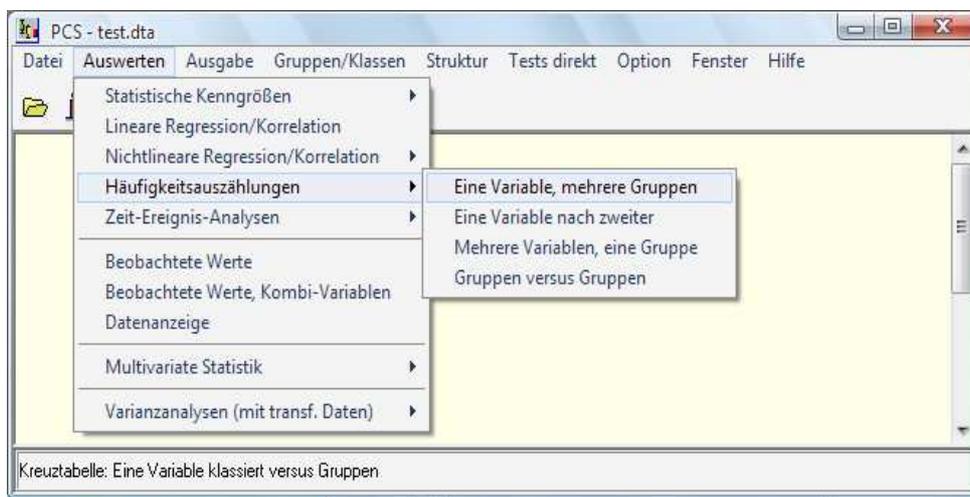
Entsprechend dem Hinweis für numerische Klasseneinteilungen tippen wir jetzt jeweils die oberen Klassengrenzen ein (mit abschließendem Return), und zwar in aufsteigender Folge. Die letzte Klassengrenze, die wir eingeben können, ist 170. PCS berücksichtigt von sich aus immer eine zusätzliche Klasse mit allen Werten größer als die letzte eingegebene Klassengrenze. Der Bereich „über 170 cm“ wird also automatisch mit einbezogen.

Statt alle Klassengrenzen einzeln einzutippen hätten wir auch mit dem Fortsetzungsschalter  arbeiten können, der das schnelle Erzeugen gleich abständiger Klassengrenzen ermöglicht. Hat man die ersten beiden Klassengrenzen eingegeben, werden durch Betätigen des Fortsetzungsschalters automatisch neue Klassen mit gleicher Breite wie die vorhergehende generiert. (Bei unseren 4 Werten hält sich die Zeitersparnis allerdings in Grenzen.)

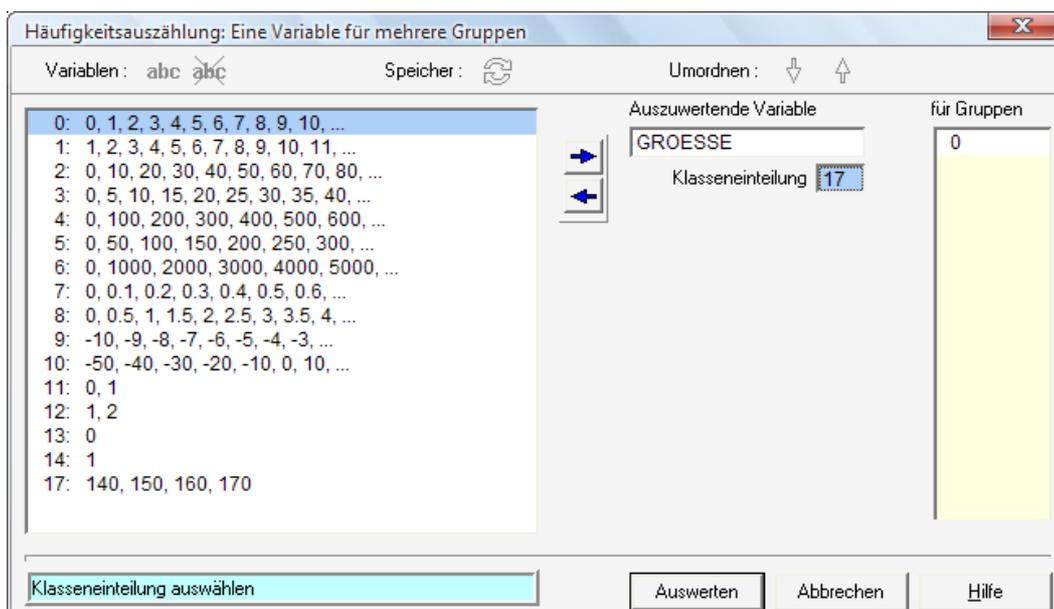
Beenden wir die Eingabe der oberen Klassengrenzen mit OK. In der Liste der Klasseneinteilungen finden wir jetzt unter Nummer 17 die neu definierte Einteilung. Mit „Schließen“ verlassen wir diesen Auswahldialog und besinnen uns auf unsere Aufgabenstellung.

2.7 Auswertung: Häufigkeitsauszählung

Die zu bewältigende Aufgabenstellung besteht im Prinzip darin, das Datenmaterial zu durchforsten und auszuzählen, wieviel Werte in den jeweiligen Klassen liegen. Dazu wählen wir im Auswertungsmenue den Punkt „Häufigkeitsauszählungen“, was uns zum nächsten Menue führt:

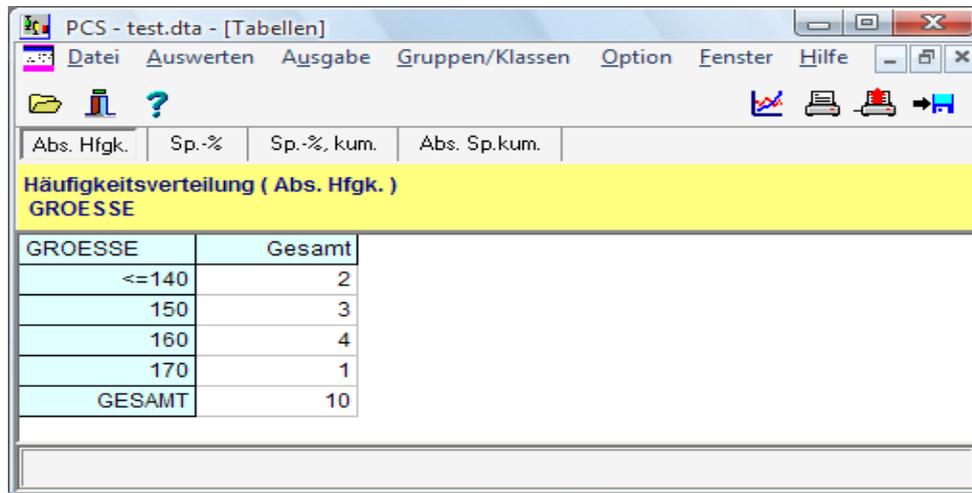


Hier kämen der erste oder dritte Menüpunkt in Frage: Wir interessieren uns nur für eine Variable und nur für eine Gruppe, nämlich die aller Schüler. Dies wäre sowohl von „Eine Variable, mehrere Gruppen“ als auch von „Mehrere Variablen, eine Gruppe“ abgedeckt. Entschließen wir uns für den ersten Menüpunkt, gelangen wir wieder in den Auswertungsdialog.



Ähnlich wie beim Auswertungsbeispiel in Kap. 2.5 wählen wir die auszuwertende Variable aus, nämlich GROESSE. Nach Auswahl der Variablen wird sofort auf das Feld Klasseneinteilung positioniert, und in der Liste im linken Bereich erscheinen die Klasseneinteilungen. Ausgewählt wird die zuvor angelegte Nummer 17. Bevor wir „Auswerten“ betätigen, wird noch die Gruppe 0 als auszuwertende Gruppe bestimmt.

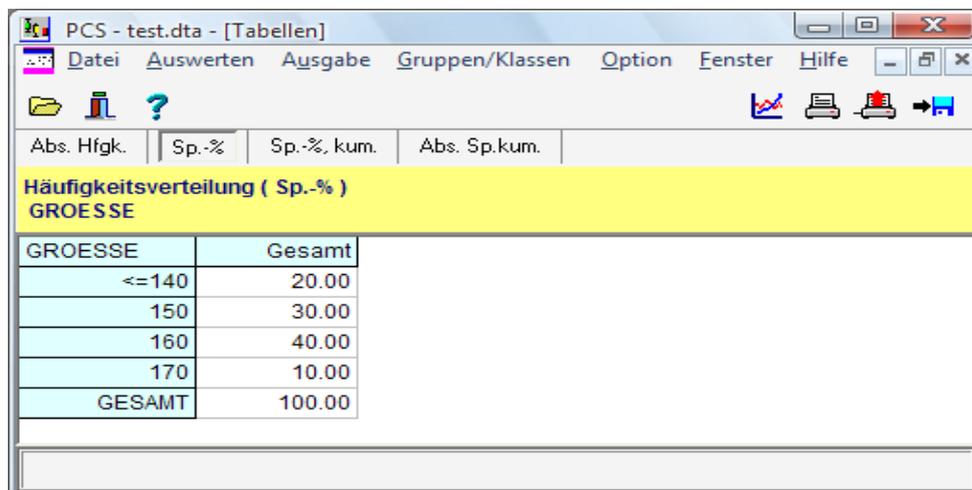
Die Verarbeitung wird nun unter Anzeige der laufenden Satznummern ausgeführt. Die Ergebnisse werden wieder auf mehreren Seiten angezeigt, wobei zunächst die absoluten Häufigkeiten präsentiert werden:



| GROESSE | Gesamt |
|---------|--------|
| <=140 | 2 |
| 150 | 3 |
| 160 | 4 |
| 170 | 1 |
| GESAMT | 10 |

Es sind also 2 Schüler bis 140 cm groß, 3 Schüler weisen eine Körpergröße zwischen 140 cm (exclusive) und 150 cm (inclusive) auf, 4 Schüler liegen im Bereich von über 150 bis 160 cm, während noch einer mehr als 160 cm und maximal 170 cm groß ist. Mit einer Körpergröße über 170 cm hat PCS keinen Schüler gefunden, weswegen diese Klasse in der Tabelle überhaupt nicht aufgeführt wird.

Durch Anklicken des Tabulators „Sp.-%“ erhält man eine Tabelle mit Prozentwerten, und zwar bezüglich der Spaltensumme. (Bei Tabellen mit mehreren Spalten gibt es auch die Möglichkeit, Prozentwerte bezüglich der Zeilensummen zu berechnen, sowie die Berechnung von Erwartungswerten und die Durchführung eines geeigneten statistischen Tests.)



| GROESSE | Gesamt |
|---------|--------|
| <=140 | 20.00 |
| 150 | 30.00 |
| 160 | 40.00 |
| 170 | 10.00 |
| GESAMT | 100.00 |

Prozentangaben sind häufig sinnvoller als absolute Werte, da der Vergleich mit anderen Tabellen gleichen Aufbaus erleichtert wird. Manchmal interessiert nicht die Verteilung selbst, sondern die kumulierten Häufigkeiten, die mit Fragestellungen der Art „Wie viel Prozent (oder wie viele) der Schüler sind bis 160 cm groß“ verknüpft sind. Natürlich lässt sich dies durch Addition der Prozentwerte beantworten, einfacher geht es jedoch über PCS, da auch diese Tabellen über die entsprechenden Tabulatoren erreichbar sind.

Auch hier können Sie über den entsprechenden Schalter  eine graphische Darstellung abrufen. Diese wird immer aus der aktuell angewählten Tabelle entwickelt. Ist also z.B. die Tabelle mit den %-Angaben bezüglich der Spaltensumme angewählt, erhält man eine Graphik mit der prozentualen (relativen) Häufigkeitsverteilung.

Üblicherweise gelangt man bei Aufruf einer Graphik zunächst wieder in einen Auswahldialog, wo man entscheiden kann, ob Zeilen oder Spalten dargestellt werden sollen, und welche. Für unser Beispiel gibt es aber keine sinnvollen Variationsmöglichkeiten. Da die Tabelle nur eine Spalte enthält, wird man in einer Graphik auch logischerweise nur deren Präsentation erwarten.

Rufen wir also von der Tabelle „Sp.-%“ aus die Graphik auf, wird der Graphik-Auswahl-Dialog übergangen und sofort die einzig praktikable graphische Darstellung präsentiert:



Diese Graphik weist noch einen kleinen Schönheitsfehler auf: Die Klassengrenzen stehen in der Mitte unter den Balken, was sinnvoll wäre, wenn diese Zahlen für irgendwelche Kategorien stünden. Vorteilhafter ist hier aber, die Markierungen zwischen den einzelnen Balken anzubringen. Mit PCS ist dies sehr einfach zu

bewerkstelligen: Rufen Sie mit dem Schalter  den Graphik-Editor auf und klicken Sie dort „x-Achse“ an.

Graphik editieren

Überschrift / Fußzeilen | x-Achse | y-Achse | Spezial

Achsenbezeichnung: GROESSE

Achse abgesetzt um: 0 Graphik-Einheiten

Skalenbeschriftung: normal gedreht

Position Achsenbez.: daneben am Ende

Skala

| L | M | G | Skalentext |
|---|---|---|------------|
| 1 | x | | 140 |
| 1 | x | | 150 |
| 1 | x | | 160 |
| 1 | x | | 170 |

Skalenverschiebung nach rechts

aus Speicher | aktuelle EINTR.

Layout sichern

Layout aus Sp.

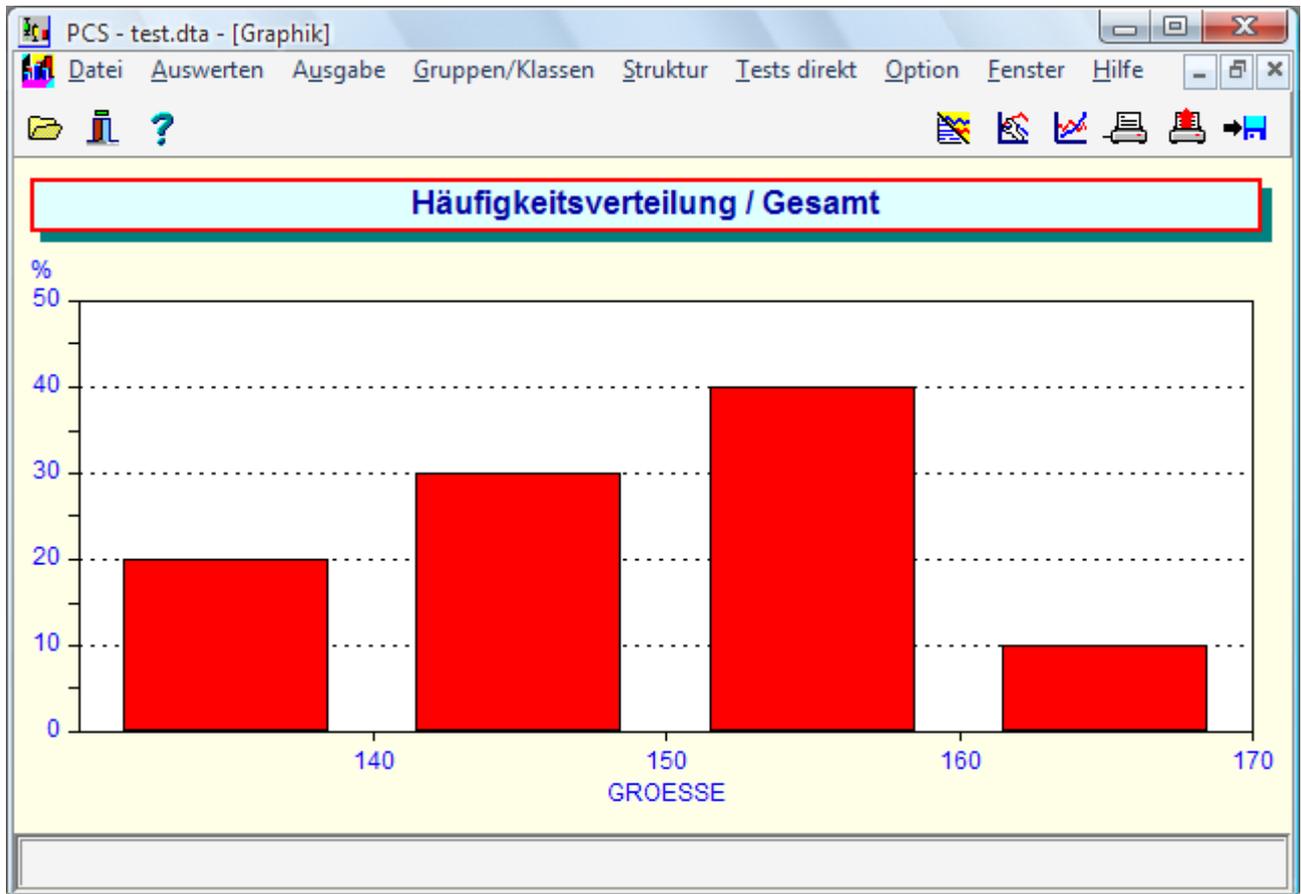
Allgemein

OK

Abbrechen

Hilfe

Auf der Seite „x-Achse“ markieren Sie nun einfach „Skalenverschiebung nach rechts“. Beenden Sie jetzt das Editieren der Graphik, erhalten Sie die verbesserte Abbildung:



Natürlich können Sie auch noch weitere Änderungen an der Graphik vornehmen. Gefallen Ihnen z.B. die Farben nicht, oder sollen die Balken direkt aneinander stoßen, betätigen Sie den Schalter „Allgemein“. Dann gelangen Sie in einen umfangreichen Dialog, der sich über mehrere Seiten erstreckt und Ihnen damit eine Vielzahl von Einstellungsmöglichkeiten eröffnet.

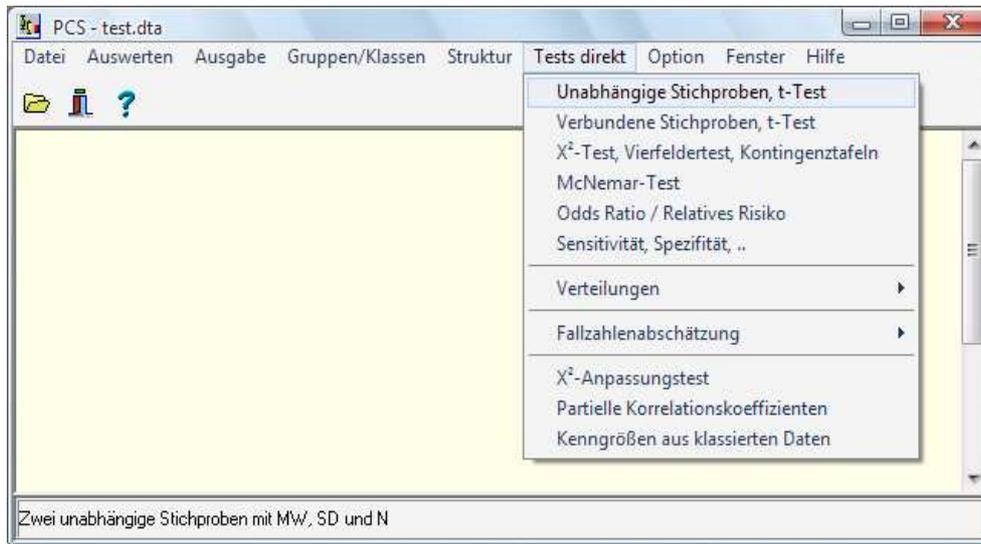
PCS ordnet von sich aus die Elemente einer Graphik so an, dass der zur Verfügung stehende Platz möglichst optimal genutzt wird, d.h. die Größe des Ausschnitts aus der xy-Ebene wird so gewählt, dass die darum herum angeordneten Elemente wie Überschrift, Legende, Fußzeilen, Achsenbeschriftungen, etc. im sichtbaren Bereich bleiben.

Möchten Sie das ändern, klicken Sie vom Graphik-Editor aus die Seite „Spezial“ an. Entfernen Sie nun in dem darauffolgenden Dialog die Markierung bei „Größe xy-Ebene selbständig ermitteln“. Jetzt werden die Werte in den Editorfeldern für die linke untere, bzw. rechte obere Ecke der xy-Ebene wirksam. Diese Werte können Sie nun nach Ihren eigenen Vorstellungen ändern, wobei die Angaben immer in % der gesamten Bildgröße (Breite oder Höhe) erfolgen müssen.

3 STATISTISCHER TASCHENRECHNER

Neben der Analyse Ihres eventuell sehr umfangreichen Datenmaterials bietet Ihnen PCS auch eine Art „Taschenrechnerfunktion“ für eine Reihe ausgewählter Tests an. Unter dem Menüpunkt „Tests direkt“ werden Verfahren zur Verfügung gestellt, die sich dadurch auszeichnen, dass sie nur eine geringe Zahl von Eingabedaten erfordern, für die auch keine Speicherung vorgesehen ist.

Das Untermenü zu „Tests direkt“ bietet Ihnen eine Reihe von Verfahren an, für welche jeweils zunächst ein Editor für die Dateneingabe zur Verfügung gestellt wird.



Auf alle hier angebotenen Verfahren können wir nicht eingehen; Sie finden ja auch genügend Hinweise in der Online-Hilfe, insbesondere auch zu den theoretischen Hintergründen der statistischen Testmethoden. Es sollen daher nur beispielhaft zwei Verfahren beschrieben werden.

3.1 Unabhängige Stichproben, t-Test

Dieser Programmteil erlaubt die Durchführung eines t-Tests für unabhängige Stichproben. Voraussetzung ist, dass für beide Stichproben Mittelwert, Standardabweichung und Stichprobenumfang bekannt sind. (Dieser Punkt ist z.B. dann sinnvoll anwendbar, wenn Sie eigene Ergebnisse, die Sie mit PCS ermittelt haben, mit Angaben aus der Literatur vergleichen wollen. In der Literatur finden Sie nämlich i.d.R. nur noch die zusammenfassenden Kennzahlen und können nicht mehr auf die Einzeldaten zurückgreifen.) Die Eingabe könnte dann z.B. folgendermaßen aussehen:

The dialog box 't-Test, unabhängige Stichproben' contains the following input data:

| | MW | SD | N |
|--------------|----|----|----|
| Stichprobe 1 | 21 | 2 | 8 |
| Stichprobe 2 | 18 | 4 | 10 |

Nach Betätigung von OK wird die folgende Ausgabe präsentiert:

The output window displays the following table and text:

| | MW | SD | SEM | N |
|--------------|----|----|--------|----|
| Stichprobe 1 | 21 | 2 | 0.7071 | 8 |
| Stichprobe 2 | 18 | 4 | 1.2649 | 10 |

t-Test, unabhängige Stichproben (gleiche Varianz):
 Testgröße: 1.929 p = 0.0723, zweis. (16 d.f.)

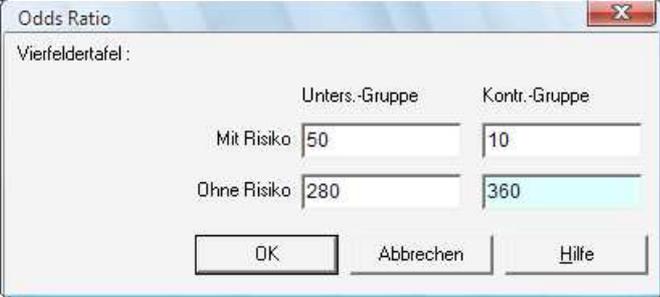
Es werden die eingegebenen Werte angezeigt, ergänzt um den Standardfehler des Mittelwertes (SEM), sowie das Testresultat (Testgröße) mit zugehörigem p-Wert, die Testvariante (einseitig oder zweiseitig) und die Freiheitsgrade (d.f. = degrees of freedom).

Bei der Durchführung des t-Tests wird intern geprüft, ob von der Annahme gleicher Varianzen für beide Stichproben ausgegangen werden kann. Trifft dies zu, erscheint das Testergebnis in obiger Form, nämlich mit dem Hinweis „(gleiche Varianz)“. Andernfalls wird eine modifizierte Form des t-Tests durchgerechnet (s. Online-Hilfe) und zusätzlich das Ergebnis des Tests auf Varianzgleichheit ausgewiesen.

Bei der Ergebnispräsentation erscheint in der Schalterleiste ein zusätzlicher Schalter, der sich durch ein fett geschriebenes T auszeichnet . Über diesen Schalter können Sie erneut das gleiche Testverfahren aufrufen, wobei im Eingabe-Editor die zuletzt vorgegebenen Werte zur Verfügung gestellt werden. Dies ist dann vorteilhaft, wenn Sie z.B. nur einen der Werte abändern wollen, weil Ihre ursprüngliche Eingabe vielleicht falsch war, oder weil Sie vielleicht einfach nur einmal sehen wollen, „was wäre wenn?“.

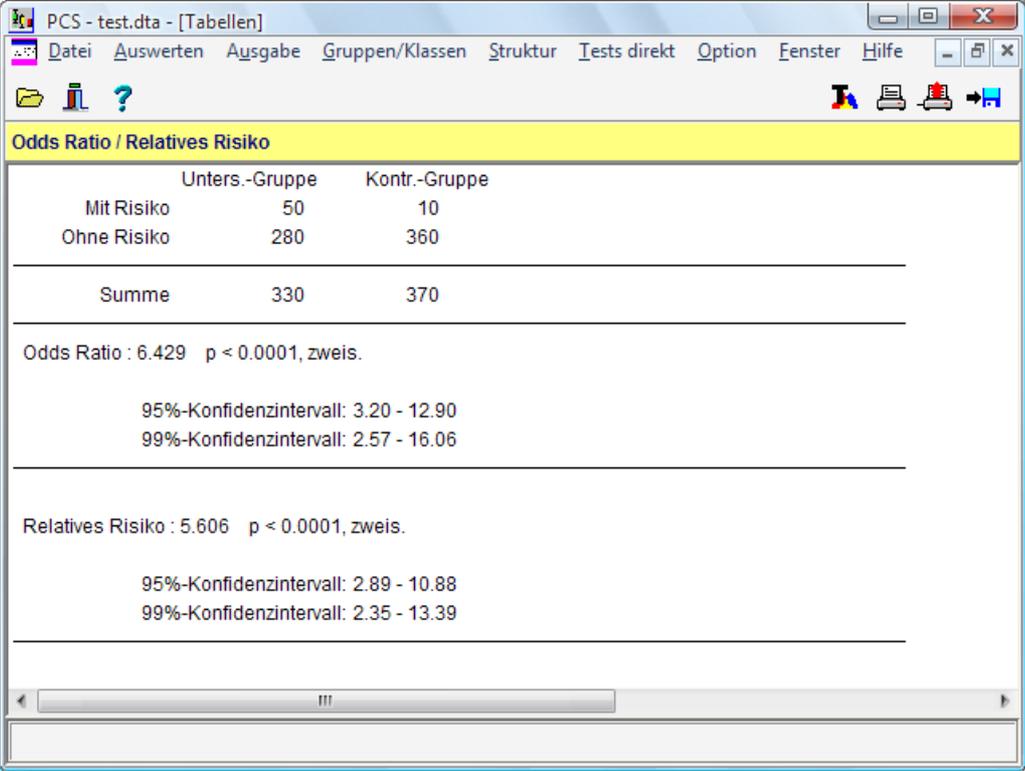
3.2 Odds Ratio

Ausgangspunkt ist eine Vierfeldertafel, in welche für Analyse- und Kontroll-Gruppe die Anzahl der Fälle mit und ohne Risiko eingetragen wird:



| | Unters.-Gruppe | Kontr.-Gruppe |
|-------------|----------------|---------------|
| Mit Risiko | 50 | 10 |
| Ohne Risiko | 280 | 360 |

Für die Ergebnispräsentation ergänzt das Programm die Tafel um die Gruppensummen und ermittelt näherungsweise das relative Risiko (odds ratio) der Untersuchungsgruppe im Vergleich zur Kontrollgruppe:



| | Unters.-Gruppe | Kontr.-Gruppe |
|-------------|----------------|---------------|
| Mit Risiko | 50 | 10 |
| Ohne Risiko | 280 | 360 |
| Summe | 330 | 370 |

Odds Ratio : 6.429 p < 0.0001, zweis.

95%-Konfidenzintervall: 3.20 - 12.90
99%-Konfidenzintervall: 2.57 - 16.06

Relatives Risiko : 5.606 p < 0.0001, zweis.

95%-Konfidenzintervall: 2.89 - 10.88
99%-Konfidenzintervall: 2.35 - 13.39

Der zugehörige p-Wert unter der Nullhypothese gleichen Risikos wird ausgewiesen, sowie die Grenzen des pA%- und pB%-Konfidenzintervalls, wobei pA und pB von den individuell eingestellten Irrtumswahrscheinlichkeiten abhängen (s. Kap. 4).

4 OPTIONEN

Das Option-Menue bietet Ihnen verschiedene Einstellmöglichkeiten. Unter „Text/Tabellen“ können Sie Farben für die Ausgabe von Tabellen, für den Dateneingabe-Editor und weitere Elemente der PCS-Oberfläche auswählen. Zusätzlich wählen Sie hier Standardschriften für die Bildschirm- und Druckerausgabe aus.

Der Menüpunkt „Graphik“ führt zu einem umfangreichen mehrseitigen Dialog, mit dessen Hilfe Sie generelle Vorgaben für die graphischen Präsentationen festlegen. Dazu gehören Farben, Schraffuren und Linienarten für die einzelnen Datenreihen, wobei für Schwarz-Weiß-Druck Optionen unabhängig von der Bildschirmausgabe gewählt werden können. Während Sie am Bildschirm also z.B. die Datenreihen durch unterschiedliche Farben unterscheiden, können Sie für den Schwarz-Weiß-Druck unterschiedliche Linienarten und/oder Schraffuren vorgeben.

Hier legen Sie auch fest, welche Schriftart für Graphiken gewählt wird, und welche Schriftgrößen für die einzelnen Elemente einer Graphik zu verwenden sind. Sie bestimmen die Breite der Boxes bei Box&Whisker-Plots, die Breite von Säulen in Säulendiagrammen und wie stark sich die Säulen überlappen, wenn mehrere Häufigkeitsverteilungen dargestellt werden.

Unter „Ausgabe“ legen Sie fest, mit welcher Auflösung Graphiken als Bitmaps gespeichert werden, und ob dies in Farbe oder in Schwarz-Weiß (entsprechend Druckereinstellungen) erfolgen soll. Ferner legen Sie die grundsätzliche Ausrichtung beim Druck fest (Hoch-/Querformat) und ob vor jedem Ausdruck der Windows-spezifische Drucker-Dialog aufgerufen werden soll.

Die hier vorgenommenen Einstellungen gelten generell für den Betrieb von PCS, unabhängig davon, welche Statistikdatei ausgewertet wird. Bei jeder Graphik haben Sie über den Schalter „Allgemein“ im Graphik-Editor-Dialog noch die Möglichkeit, mit Hilfe eines fast ebenso umfangreichen Dialogs individuelle, nur die spezifische Graphik betreffende Einstellungen vorzunehmen.

Anders verhält es sich dagegen mit dem Punkt „Irrtumswahrscheinlichkeiten“ im Option-Menue. Die Einstellungen, die Sie hier vornehmen gelten nur für Ihr Projekt, bzw. Ihre Statistikdatei, und sie werden darum auch in der zu Ihrer Statistikdatei gehörenden Parameterdatei gespeichert.

Hier stellen Sie nicht nur die Irrtumswahrscheinlichkeiten ein, sondern noch eine Reihe weiterer projektspezifischer Parameter. Der Menüpunkt „Irrtumswahrscheinlichkeiten“ führt zu folgendem Dialog:

Irrtumswahrscheinlichkeiten, Faktoren, Quantile

Testart

einseitig

zweiseitig

Bevorzugte Testmethode

parametrisch (t-Test, etc.)

nichtparametrisch (Rangtests)

Whisker-Faktor B&W-Plots

1 1,5 2 2,5 3

Irrtumswahrscheinlichkeit A : 0.05

B : 0.01

Faktor A (M±SD) : 2

Faktor B (M±SD) : 3

Quantil A (in %) : 10

Quantil B (in %) : 90

OK

Abbrechen

Hilfe

Unter „Testart“ entscheiden Sie sich für einseitige oder zweiseitige Tests (s. Online-Hilfe). Empfehlenswert sind zweiseitige Tests, obwohl einseitige Tests schneller zu „signifikanten“ Ergebnissen führen.

Als „Bevorzugte Testmethode“ sollte der nichtparametrische Test (Rangtest) gewählt werden. Diese Vorgabe führt dazu, dass PCS wann immer möglich verteilungsfreie Testverfahren benutzt. Diese Testverfahren sind eher konservativ, d.h. sie neigen dazu, die Nullhypothese länger beizubehalten. Dafür ersparen sie dem Anwender jede Diskussion darüber, ob die getesteten Daten tatsächlich der geforderten Verteilung (meist Normalverteilung) genügen.

Daneben legen Sie in diesem Dialog individuelle Irrtumswahrscheinlichkeiten p_A und p_B fest, wobei p_B immer kleiner als p_A sein muß. Werden bei den Auswertungen die p -Werte nicht explizit ausgegeben, werden Ergebnisse mit einem $p < p_A$ durch + und solche mit $p < p_B$ mit x gekennzeichnet.

Sie legen ferner fest, welche Faktoren für die Bestimmung von Streubereichen zu verwenden sind. Durch diese Einstellung wird z.B. die tabellarische und graphische Ausgabe von statistischen Kenngrößen beeinflusst.

Zusätzlich können Sie zwei Quantile festlegen, auf die in Tabellen und Graphiken statistischer Kenngrößen Bezug genommen wird.

Im unteren Bereich legen Sie noch fest, welcher Faktor bei der Bestimmung vorläufiger Whisker-Grenzen bei Box&Whisker-Plots zur Anwendung kommt. Standardmäßig ist hier der Faktor 1 eingestellt.

5 DRUCKMANAGER

Das Drucken unter Windows ist häufig reine Papierverschwendung. Besteht eine Ergebnistabelle nur aus wenigen Zeilen, wird immer ein ganzes Blatt Papier ausgegeben. PCS sorgt mit dem integrierten Druckmanager dagegen für einen sparsameren Umgang mit dem Papier. (Durch Anwählen von „Option / Druck-Speicherautomatik“ können Sie den Druckmanager wahlweise aktivieren, bzw. deaktivieren.)

Wenn Sie z.B. nach der Anzeige einer Ergebnistabelle den Schalter für „Drucken“  betätigen, passiert häufig nichts, d.h. der Drucker bleibt „stumm“. Das hängt damit zusammen, dass PCS alle Druckaufträge sammelt und in einem Druckerpuffer ablegt. Erst wenn eine Druckerseite voll ist, weil z.B. der neu angeforderte Ausdruck nicht mehr auf die Seite passt, wird tatsächlich gedruckt. PCS bemüht sich dabei jeden Druckauftrag auf einer Seite unterzubringen.

Wenn Sie PCS beenden, kann es vorkommen, dass im Druckerpuffer noch Aufträge vorhanden sind. Dann meldet sich PCS vor dem Schließen des Programms, weist darauf hin, dass der Druckerpuffer nicht leer ist, und fragt, ob der Puffer noch ausgedruckt werden soll.

Eine zusätzliche Eingriffsmöglichkeit haben Sie über den Schalter „Blattvorschub“ . Betätigen Sie diesen Schalter wird der Druckerpuffer entleert, d.h. ausgedruckt, und zwar unabhängig vom „Füllungszustand“ des Druckerpuffers.

Ein Wechsel der auszuwertenden Statistikdatei führt automatisch zum Ausdrucken des Druckpuffers. Dies ist notwendig, da die auf dem Drucker ausgegebenen Seiten im Kopf mit Dateiname, Uhrzeit und Seitennummer versehen werden.

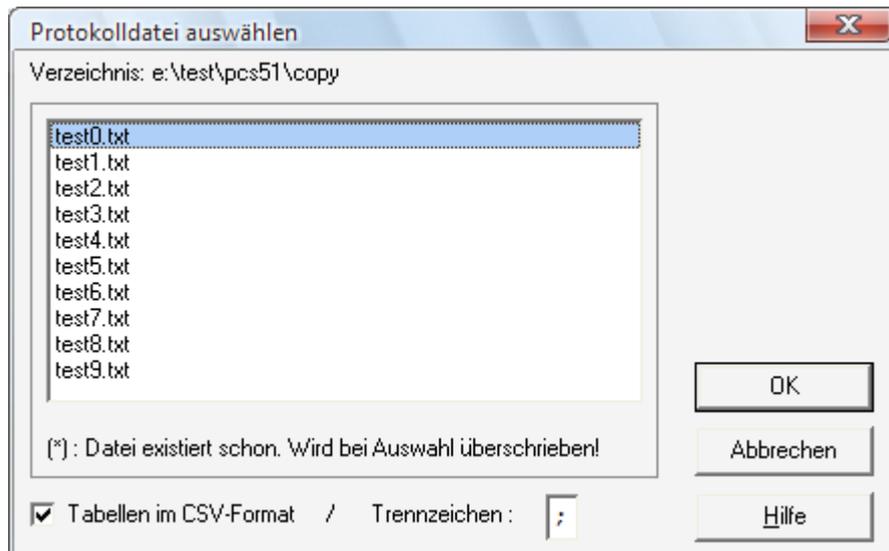
Graphiken werden vom Druckmanager nicht verwaltet. Sie werden generell auf einer vollständigen Seite ausgegeben, auch wenn Sie z.B. Hochformat ausgewählt haben und die Graphik eigentlich nur die obere Hälfte der Seite belegt.

6 PROTOKOLLDATEI

Häufig möchte man die Auswertungsergebnisse nicht sofort drucken, sondern sie lieber in eine Datei kopieren, um sie später per Textverarbeitung in einen Report oder einen wissenschaftlichen Artikel zu integrieren. Die Möglichkeit hierzu bietet der äußerste rechte Schalter  auf dem Schalterpanel.

Dieser Schalter erlaubt eine Protokollierung der Auswertungen durch fortlaufendes Abspeichern der tabellarischen Ergebnisse in einer ausgewählten Datei.

Betätigen Sie während einer Arbeitssitzung (oder nach Wechsel der auszuwertenden Datei) zum ersten Mal diesen Schalter, werden Sie aufgefordert, eine Protokolldatei auszuwählen:



Der Name einer Protokolldatei ergibt sich aus dem Namens-Hauptteil Ihrer Statistikdatei erweitert durch eine der Ziffern 0 bis 9, d.h. Sie können immer unter 10 Protokolldateien auswählen. Die Protokolldateien besitzen die Namenserweiterung .txt, damit sie z.B. einfach per Doppelklick mit dem Editor (NOTEPAD) gelesen werden können.

Protokolldateien werden im Unterverzeichnis XXXX\COPY abgelegt, wobei XXXX das Unterverzeichnis ist, in welchem PCS installiert wurde.) Schon existierende Protokolldateien sind im Auswahldialog mit * gekennzeichnet, um ein unbeabsichtigtes Überschreiben dieser Protokolldateien zu erschweren. Da nach Auswahl einer Protokolldatei das Protokoll nämlich neu begonnen wird, würde eine schon vorhandene Protokolldatei überschrieben.

Weitere Betätigungen des Schalters zum Kopieren auf Platte hängen dann die gewünschte Ausgabe einfach an die beim ersten Mal ausgewählte Protokolldatei an, so dass durch einfaches Anklicken des Kopierschalters die ganze Auswertungssitzung protokolliert werden kann. Bei mehrseitigen Tabellen muss allerdings jede Seite getrennt auf Platte kopiert werden!

Bei der ersten Auswahl einer Protokolldatei müssen Sie noch festlegen, ob Tabellen im CSV-Format ausgegeben werden sollen. Normalerweise werden Tabellen in der Protokolldatei spaltengerecht ausgegeben, indem ausreichend Zwischenräume berücksichtigt werden. Derart formatierte Ausgaben lassen sich aber in einer Textverarbeitung nur schwer wieder in Tabellenform bringen. Günstiger ist da das CSV-Format, d.h. die Werte werden unmittelbar hintereinander ausgegeben, nur getrennt durch das (im Dialog ebenfalls festzulegende) Trennzeichen.

Graphiken werden dagegen nicht in einer Datei gesammelt, sondern sie werden jeweils einzeln im BMP- oder JPEG-Format gespeichert, wobei Sie jedes mal den gewünschten Dateinamen und den Dateityp in einem entsprechenden Dialog angeben müssen. (Natürlich muss der Anwender dann auch den Überblick darüber behalten, welches Bild unter welchem Namen abgelegt wurde.)

7 VON PCS4 NACH PCS5

Haben Sie bisher PCS4 benutzt, sind keine besonderen Umsetzungsläufe erforderlich. Beim ersten Arbeiten mit einer PCS4-Statistikdatei führt PCS5 automatisch geringfügige Änderungen in der Parameterdatei durch. Wollen sie eventuell später doch noch einmal unter PCS4 auswerten, sollten Sie vor Einsatz von PCS5 Sicherungskopien Ihrer Statistikdateien anfertigen.