wdv-notes 365
Stand: 14.JUL.1995 (4.)
Wiss.Datenverarbeitung
FREIE UNIVERSITÄT BERLIN

StatView: Bearbeiten von Daten.

© 1994–1995 *Karl-Heinz Dittberner*

MAC

In den wdv-notes Nr. 306 wurde das herausragende Statistikprogramm ***StatView*** vorgestellt [1–2]. In diesem Programm ist auch eine Tabellenkalkulation eingebaut, so daß sich die Eingabedaten leicht in neue Variablen umrechnen lassen.

In diesem Merkblatt werden die wichtigsten Werkzeuge für die Bearbeitung von Daten zusammengestellt. Alle Angaben basieren auf der Version 4.0.2. Eine ausführliche Beschreibung findet man im Kapitel 5 und 6 von [2].

1. Management der Daten:

Mit den Befehlen des Menüs »Manage« [1] lassen sich die Daten im Dataset, in der Datentabelle bearbeiten, d. h. ergänzen und umrechnen. Mit dem Befehl »Exclude Row« können bestimmte (markierte) Zeilen (Beobachtungen) von den weiteren Analysen ausgeschlossen werden. Mit »Include Row« kann das wieder aufgehoben werden.

Der Befehl »Formula…« ermöglicht das Berechnen neuer Variablen (Punkt 2), mit »Recode…« können vorhandene Variablen nach bestimmten Bedingungen umgewandelt werden (Abb. 2). Außerdem lassen sich die Daten sortieren (Befehl »Sort…«), Zufallszahlen (Befehl »Random Numbers…«) und Datenreihen (Befehl »Series…«) erzeugen.

2. Erzeugen neuer Variablen:

Mit Formeln lassen sich neue Variablen (Spalten) aufgrund der bereits vorhandenen Variablen berechnen. Im Formel-Fenster links oben (Abb. 1) findet man eine Liste aller Variablen (C = continous, kontinuierliche Variable, N = nominale Variable), die in der aktuellen, d. h. der obersten aktiven Datentabelle enthalten sind. Darunter befindet sich die Liste der in StatView eingebauten Funktionen. Die Bedeutung der Funktionsargumente wird im Fenster »Hints« [1] angegeben.

Eine Formel wird im Definitionsfenster mit Hilfe des Tastenfelds sowie durch Auswahl und Doppelklick der Variablen (Namen mit Sonderzeichen werden in "…" geschrieben) oder Funktionen eingegeben. Als Platzhalter für noch fehlende Argumente dient das „?“. Kommentare können in (* … *) zugefügt werden.

Durch Mausklick auf die Taste »Compute« wird die neue Variable berechnet, d. h. die neue Spalte erzeugt. Soll eine solche Spalte wieder gelöscht werden, dann muß sie nur in der Tabelle markiert werden und die Delete-Taste des Tastenfelds betätigt werden.

Mit der Taste »Attributes« können verschiedene Eigenschaften der Variablen eingestellt werden, z. B. der Name der Variablen, der Datentyp [1], die Datenklasse [1]. In den folgenden Punkten 3–12 wird eine Übersicht aller zur Verfügung stehenden Operatoren und Funktionen, sortiert nach Sachgebieten gegeben.

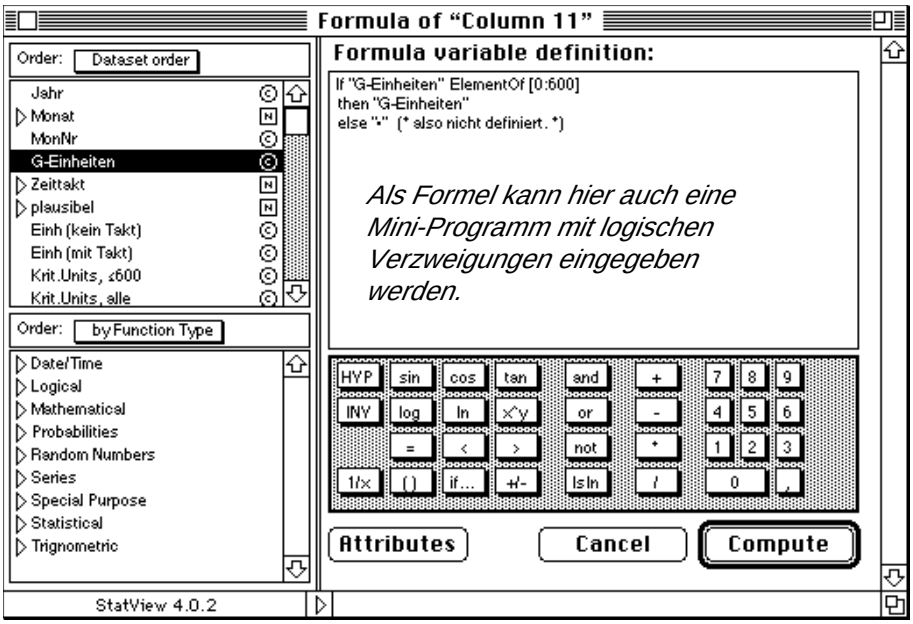


Abb. 1: Das Fenster zur Eingabe von Formeln, mit denen die Daten in einer Spalte (neue Variable) aufgrund anderer Variablen und Parameter berechnet werden. Dieses Fenster wird mit dem Befehl »Formula…« aus dem Menü »Manage« [1] *gestartet.*

3. Operatoren:

<	*kleiner*
=	*gleich*
>	*größer*
ElementOf	*Element von*
Is	*ist*
IsIn	*Entspricht ElementOf*
IsNot	*ist nicht*
≠	*ungleich*
≤	*kleiner oder gleich*
≥	*größer oder gleich*

4. Datum/Zeit-Funktionen:

Es wird hier eine Zeitzahl verwendet. Das ist ein LongInteger, der in Sekunden die Zeit angibt, die seit dem 1. Januar 1904 um 12:01:00 Uhr vergangen ist.

- **Date** (?; ?; ?) *Zeitzahl aus Jahr, Monat, Tag ermitteln.*
- **DateDifference** (?; ?; *u*) *Datendifferenz in Jahren (u=1), Monaten (u=2), Wochen (u=3), Tagen (u=4), Stunden (u=5), Minuten (u=6) oder Sekunden (u=7).*
- **Day** (?) *Tag aus Zeitzahl.*
- **DayOfYear** (?) *Tag im Jahr aus Zeitzahl.*
- **Hour** (?) *Stunde aus Zeitzahl.*
- **Minute** (?) *Minute aus Zeitzahl.*
- **Month** (?) *Monat aus Zeitzahl.*
- **Second** (?) *Sekunde aus Zeitzahl.*
- **Time** (?; ?; ?) *Zeitzahl aus Stunde, Minute, Sekunde.*
- **Weekday** (?) *Wochentag aus Zeitzahl.*
- **WeekOfYear** (?) *Wochennummer im Jahr.*
- **Year** (?) *Jahr aus Zeitzahl.*

5. Logische Funktionen:

- **false** *Falsch (Konstante), NEIN.*
- **if** ? **then** ? **else** ? *Logische Abfrage.*
- **IsMissing** (?) *fehlt?*
- **IsRowExcluded** *Zeile eingeschloss.?*
- **IsRowIncluded** *Zeile ausgeschloss.?*
- **NOT** (?) *NICHT*
- **true** *Richtig (Konstante), JA.*
- ? **<** ? *kleiner als*
- ? **=** ? *gleich*
- ? **>** ? *größer als*
- ? **AND** ? *UND*
- ? **&** ? *UND*
- ? **ElementOf** ? *Element von*
- ? **Is** ? *ist*
- ? **IsNot** ? *ist nicht*
- ? **OR** ? *ODER*
- ? **|** ? *ODER*
- ? **XOR** ? *Exklusives ODER*
- ? **≠** ? *ungleich*
- ? **<>** ? *ungleich*
- ? **≤** ? *kleiner oder gleich*
- ? **<=** ? *kleiner oder gleich*
- ? **≥** ? *größer oder gleich*
- ? **<=** ? *größer oder gleich*

Wichtiger Hinweis

Für das erfolgreiche Arbeiten mit StatView ist es besonders wichtig, die Datentabelle mit den Variablen und Parametern sehr sorgfältig zu planen. Für viele der Analysen werden bestimmte Datentypen und eine ganz bestimmte Organisation der Daten vorausgesetzt.

6. Mathematische Funktionen:

Bezüglich der Bedeutung des Arguments *AllRows* siehe den Vorspann zum Punkt 11.

- **Average** (?; …) *Mitteln.*
- **AverageIgnoreMissing** (?; …) *Mitteln, wobei fehlende Werte nicht berücksichtigt werden.*
- **Ceil** (?) *Runden nach oben.*
- **Combinations** (?; ?) *Kombination.*
- **CumProduct** (?) *Kumulat. Produkt.*
- **CumSum** (?) *Kumulat. Summe.*
- **CumSumSquares** (?) *Kum.Quadr.Sum.*
- **Difference** (?; 1; 1) *Differenz.*
- **Div** (?; ?) *Integer-Divison.*
- **DotProduct** (?; ?) *Skalarprodukt.*
- **e** *Eulersche Zahl.*
- **Erf** (?) *Fehlerfunktion.*
- **Factorial** (?) *xxx.*
- **Floor** (?) *Runden nach unten.*
- **Lag** (?; *a*) *Um a Zeilen verschieben.*
- **Ln** (?) *Logarithmus Basis e.*
- **Log** (?) *Logarithmus Basis 10.*
- **LogB** (?; *b*) *Logarithmus Basis b.*
- **Mod** (?; ?) *Modulo.*
- **MovingAverage** (?; ?) *Gleit. Mitteln.*
- **Norm** (?; *AllRows*) *Euklid. Norm.*
- **Percentages** (?; *AllRows*) *Prozente.*
- **Permutations** (?; ?) *Permutation.*
- **Remainder** (?; ?) *Rest.*
- **Round** (?) *Runden.*
- **Sqrt** (?) *Quadratwurzel.*
- **Sum** (?; …) *Summe.*
- **SumIgnoreMissing** (?; …) *Ohne Fehl.*
- **Trunc** (?) *Abschneiden.*
- **|** ? **|** *Betrag von*
- π *or* **Pi** *Pi (Konstante).*
- **(** ? **)** *Klammern von Ausdrücken.*
- **+** ? *Positives Vorzeichen.*
- **–** ? *Negatives Vorzeichen.*
- ? ***** ? *Multiplikation.*
- ? **+** ? *Addition.*
- ? **–** ? *Subtraktion.*
- ? **/** ? *Division.*
- ? **^** ? *Potenzierung.*
- ? ****** ? *Potenzierung.*

7. Wahrscheinlichkeiten:

- **ProbBinomial** (?; ?; ?) *Binomialvertlg.*
- **ProbChiSquare** (?; 1) χ^2*-Verteilung.*
- **ProbF** (?; 1; 1) *F-Verteilung.*
- **ProbNormal** (?; 0; 1) *Normal-Verteilg.*
- **Probt** (?; 1) *t-Verteilung.*
- **ReturnChiSquare** (?) *xxx.*
- **ReturnF** (?; 1; 1) *xxx.*
- **ReturnNormal** (?; 0; 1) *xxx.*
- **ReturnT** (?; 1) *xxx.*

8. Zufallszahlen:

- **RandomBeta** (1, 1) *β-verteilt.*
- **RandomBinomial** (?; ?) *Binomialverteilt.*
- **RandomChiSquare** (1) χ^2*-verteilt.*
- **RandomExponential** (1) *exp-verteilt.*
- **RandomF** (1; 1) *F-verteilt.*
- **RandomGamma** (1) *γ-verteilt.*
- **RandomGaussian** (0; 1) *normalverteilt.*
- **RandomInclusion** (?) *xxx.*
- **RandomNormal** (0; 1) *normalverteilt.*
- **RandomPoisson** (1) *Poisson-verteilt.*
- **RandomT** (1) *t-verteilt.*
- **RandomUniform** (0; 1) *gleichverteilt.*
- **RandomUniformInteger** (?; ?)

9. Reihen (Series):

- **BinomialCoeffs**
- **CubicSeries** (1; 0; 0; 1)
- **ExponentialSeries** (1)
- **FibonacciSeries**
- **GeometricSeries** (1; 2)
- **LinearSeries** (1; 1)
- **QuadraticSeries** (1; 0; 1)
- **QuarticSeries** (1; 0; 0; 0; 1)
- **RowNumber** *Lfd. Zeilennummer.*

10. Spezielle Funktionen:

- **[** ? **:** ? **)** *Bereich ohne oberen Wert.*
- **[** ? **:** ? **]** *Bereich.*
- **{** … **}** *Menge mit Elementen.*
- ≤ ? *kleiner oder gleich*
- **<=** ? *kleiner oder gleich*
- ≥ ? *größer oder gleich*
- **>=** ? *größer oder gleich*
- **(** ? **:** ? **)** *Bereich ohne die Eckwerte.*
- **(** ? **:** ? **]** *Bereich ohne unteren Wert.*
- **<** ? *kleiner*
- **>** ? *größer*

11. Statistische Funktionen:

Das erste Argument in diesen Funktionen ist immer eine Variable. Viele Funktionen erlauben die Angabe, welche Zeilen (Beobachtungen) in die Berechnung einbezogen werden sollen. Hier können angegeben werden *AllRows* (Default), *OnlyIncludedRows*, *OnlyExcludedRows*.

- **BoxCox** (?; ?) *Box-Cox Transformation.*
- **CoeffOfVariation** (?; *AllRows*)
- **Correlation** (?; ?; *AllRows*)
- **Count** (?; *AllRows*)
- **Covariance** (?; ?; *AllRows*)
- **Geometric Mean** (?; *AllRows*)
- **Groups** (?; …)
- **HarmonicMean** (?; *AllRows*)
- **LogOdds** (?) *Log-Odds Transformation.*
- **MAD** (?; *AllRows*) *Abs. Abweichung vom Medianwert.*
- **Maximum** (?; *AllRows*) *Maximaler Wert.*
- **Mean** (?; *AllRows*) *Mittelwert.*
- **Median** (?; *AllRows*) *Medianwert.*
- **Minimum** (?; *AllRows*) *Minimaler Wert.*
- **Mode** (?; *AllRows*)
- **NumberMissing** (?; *AllRows*)
- **NumberOfRows** *Anzahl der Zeilen.*
- **OneGroupChiSquare** (?; ?; ?)
- **Percentile** (?; ?; *AllRows*)
- **Range** (?; *AllRows*) *Bereich.*
- **Rank** (?; *AllRows*) *Rang.*
- **StandardDeviation** (?; *AllRows*) *Standardabweichung.*
- **StandardError** (?; *AllRows*) *Fehler.*
- **StandardScores** (?; *AllRows*)
- **SumOfColumn** (?; *AllRows*)
- **SumOfSquares** (?; *AllRows*)
- **TrimmedMean** (?; ?; *AllRows*)
- **Variance** (?; *AllRows*) *Streuung.*

12. Trigonometrische Funktionen:

Alle Argumente werden in Bogenmaß (Radian) erwartet. Die Ergebnisse der Arcus-Funktionen sind ebenfalls in Bogenmaß.

- **ArcCos** (?) *Arcus Cosinus.*
- **ArcCosh** (?) *Arcus Hyperbelcosinus.*
- **ArcCot** (?) *Arcus Cotangens.*
- **ArcCsc** (?) *Arcus Cosecans.*
- **ArcSec** (?) *Arcus Secans.*
- **ArcSin** (?) *Arcus Sinus.*
- **ArcSinh** (?) *Arcus Hyperbelsinus.*
- **ArcTan** (?) *Arcus Tangens.*
- **ArcTanh** (?) *Arcus Hyperbeltangens.*
- **Cos** (?) *Cosinus.*
- **Cosh** (?) *Hyperbelcosinus.*
- **Cot** (?) *Cotangens.*
- **Csc** (?) *Cosecans.*
- **DegToRad** (?) *Grad in Bogenmaß.*
- **RadToDeg** (?) *Bogenmaß in Grad.*
- **Sec** (?) *Secans.*
- **Sin** (?) *Sinus.*
- **Sinh** (?) *Hyperbelsinus.*
- **Tan** (?) *Tangens.*
- **Tanh** (?) *Hyperbeltangens.*

13. Sonstige Hinweise:

■ StatView unterscheidet bei den Formeln zwischen statischen, dynamischen und vom Benutzer eingegebenen (User entered) Formeln. Nur die dynamischen Formeln führen zu einer sofortigen Neuberechnung bei Änderungen von Daten oder Formel.

14. Literatur:

[1] Dittberner, K.-H.: Wissenschaftliche Statistik mit StatView 4.0. FU Berlin (IfP): wdv-notes Nr. 306, 1994–1995.

[2] Saga, S. and Rocco, T.: StatView for the Macintosh. The ultimate integrated data analysis and presentation system. Berkeley (USA): Abacus Concepts Inc. 1992. ISBN: 0-944800-03-3. – [CM 239].

[3] Dittberner, K.-H.: StatView: Toolbox. FU Berlin (IfP): wdv-notes Nr. 367, 1995.

Notizen:

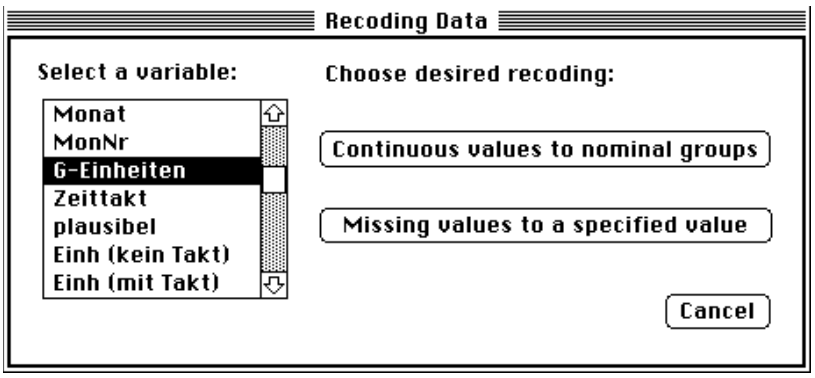


Abb. 2. ►

© khd-research • 2004 • Eine Publikation von http://www.khd-research.net/ • File: Notes/wdv-notes_365.pdf