wdv-notes 306
Stand: 15.JUL.1995 (3.)
Wiss.Datenverarbeitung
FREIE UNIVERSITÄT BERLIN

Wissenschaftliche Statistik mit *StatView* 4.0.

© 1994–1996 *Karl-Heinz Dittberner*

MAC

Das Programm ***StatView*** von Abacus Concepts, Berkeley (USA) enthält alle Werkzeuge, die Wissenschaftler, Ingenieure und Forscher brauchen, um ihre Daten unkompliziert, statistisch einwandfrei analysieren und darstellen zu können [1].

Viele der Möglichkeiten des Macintosh Betriebssystems werden bereits von ***StatView*** genutzt. Unter den für den Mac erhältlichen Statistik-Programmen ist ***StatView*** mit Abstand die benutzerfreundlichste Anwendung, die außerdem eine integrierte Tabellenkalkulation [5] enthält.

In diesem Merkblatt wird ein kurzer Überblick über die allgemeine Eigenschaften des ***StatView*** in der Version 4.0.2 sowie eine grafische Übersicht der Hauptmenüs gegeben.

1. Welche statistischen Analysen sind möglich?

Das ***StatView*** beherrscht 11 Arten statistischer Analysen und 6 Diagrammarten (siehe Punkt 2). In der folgenden Übersicht wird auf die ausführliche Beschreibung in dem jeweiligen Kapitel der Originaldokumentation verwiesen [1].

- ***Statistische Kennwerte*** (Descriptive Statistics – Kapitel 11): Berechnet werden u. a. Mittelwert, Median, Streuung, Standardabweichung, Fehler, Varianzkoeffizient, Bereich, Schiefe, Kurtosis. Ausgabe erfolgt als Tabelle.
- ***Häufigkeitsverteilung*** (Frequency Distribution – Kapitel 12): Histogramme, Kreisdiagramme und Z-Score Histogramme.
- ***Stichprobentests: Vergleich einer Stichprobe mit der hypothetischen Grundgesamtheit*** (One Sample Analysis – Kapitel 13): Student t-Test, Chiquadrat-Test, ein- oder zweiseitig, auch mit Berechnung des Vertrauensbereichs.
- ***Stichprobentests: Vergleich zweier gepaarten Stichproben*** (Paired Comparisons – Kapitel 14): Gepaarter Student t-Test, Z-Test, ein- oder zweiseitig, auch mit Berechnung des Vertrauensbereichs. Ausgabe erfolgt als Tabelle.
- ***Stichprobentests: Vergleich zweier ungepaarter Stichproben*** (Unpaired Comparisons – Kapitel 15): Ungepaarter Student t-Test, F-Test, ein- oder zweiseitig, auch mit Berechnung des Vertrauensbereichs. Ausgabe erfolgt als Tabelle.
- ***Korrelation und Kovarianz*** (Correlation + Covariance – Kapitel 16): Korrelationskoeffizient, Korrelationsmatrix, Fishers Transformation (r to z), Bartletts Test, Kovarianz.
- ***Regression*** (Regression – Kapitel 17): Einfache und mehrfache Regression, Polynom Regression, stufenweise Regression, Residuen. Ausgabe erfolgt als Tabelle, auch als Plot.
- ***Varianzanalyse*** (ANOVA – Kapitel 18).
- ***Kontingenztabellen*** (Contingency Tables – Kapitel 19).
- ***Nichtparametrische Tests*** (Nonparametrics – Kapitel 20): Vorzeichentest, Mann-Whitney Test, Kolmogorov-Smirnov Test, Wald-Wolfowitz Test, Wilcoxon Test, Spearman Test, Kendalls Test, Kruskal-Wallis Test, Friedman Test. Ausgaben erfolgen als Tabellen.
- ***Faktorenanalyse*** (Factor Analysis – Kapitel 21): Principal Component Analysis u. Transformationen.

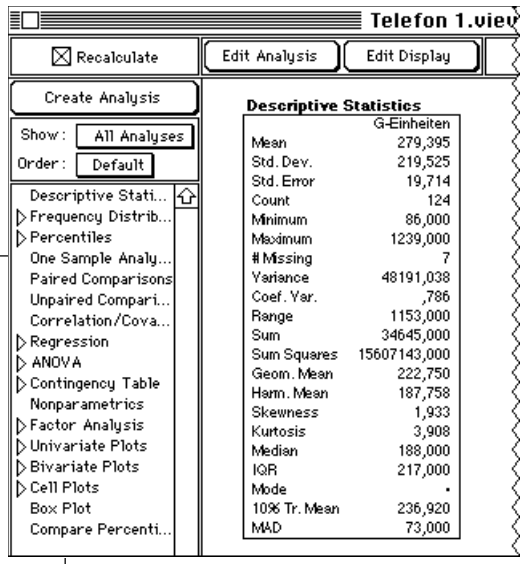


Abb. 1: Das Arbeitsfenster von StatView. Hier mit der Ausgabe von statistischen Kennwerten einer Stichprobe aus Abb. 3.

2. Welche Diagramme können produziert werden?

StatView erlaubt neben den bereits unter Punkt 1 genannten Diagrammen 6 weitere Diagrammarten zur Darstellung von Daten.

- ***y-Diagramme*** (Univariate Plots – Kapitel 22): Punkt-Diagramme (Scatter Plots), Linien-Diagramme, Säulen-Diagramme.
- ***xy-Diagramme*** (Bivariate Plots – Kapitel 22): Punkt-Diagramm, auch mit Verbindungslinien und mit Regressionslinien, auch mit hyperbolischen Vertrauensbereichsgrenzen. Auch logarithmische Achsen.
- ***Balken-Diagramme*** (Cell Plots – Kapitel 22): Vertikale Balken, auch mit Fehlerbereich, auch Punkte und Linien sind möglich.
- ***Box-Diagramme*** (Box Plots – Kapitel 22): Darstellung der 10., 25., 50., 75. und 90. Prozentile einer Variablen.
- ***Prozentile-Diagramme*** (Percentile Plots – Kapitel 22): Beobachtete Werte einer Variablen über ihrer Prozentile aufgetragen.
- ***Vergleichende Prozentile-Diagramme*** (Compare Percentile Plots – Kapitel 22).

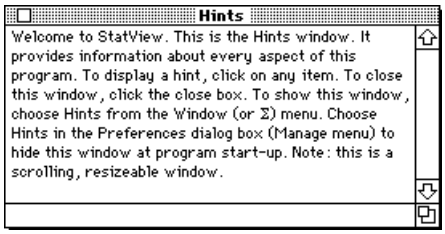


Abb. 2: Das Online-Hilfefenster „Hints" von StatView.

3. Das Starten:

Der Start von ***StatView*** erfolgt durch Doppelklick auf das Icon. Liegt das Programm-Icon im Fenster des »KlickStarters«, dann reicht ein einfacher Mausklick aus.

4. Die Hauptmenüs:

Die Hauptmenüs sind auf der Seite 2 oben mit allen Befehlen abgebildet. Mit dem Befehl »New« aus dem Menü »File« wird eine leere Datentabelle (dataset) zur Eingabe erzeugt, eine vorhandene wird mit »Open…« geöffnet.

5. Die Dateneingabe:

Die Dateneingabe erfolgt in einer Tabelle (Abb. 3). Zuerst muß der Kopf der Tabelle mit allen Angaben zu den Variablen ausgefüllt werden.

Jede Spalte dieser Tabelle stellt eine Variable dar, und jede Zeile ist eine Beobachtung (observation). Es lassen sich auch kompakte Variablen – das sind Variablen mit mehreren Unterspalten – verwenden, was beispielsweise für die ANOVAs benötigt wird. Mit der Taste »Compact« (Abb. 3) können mehrere nebeneinanderliegende Spalten zu ei-

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 Tests werden bestimmte Datentypen und eine ganz bestimmte Organisation der Daten vorausgesetzt.

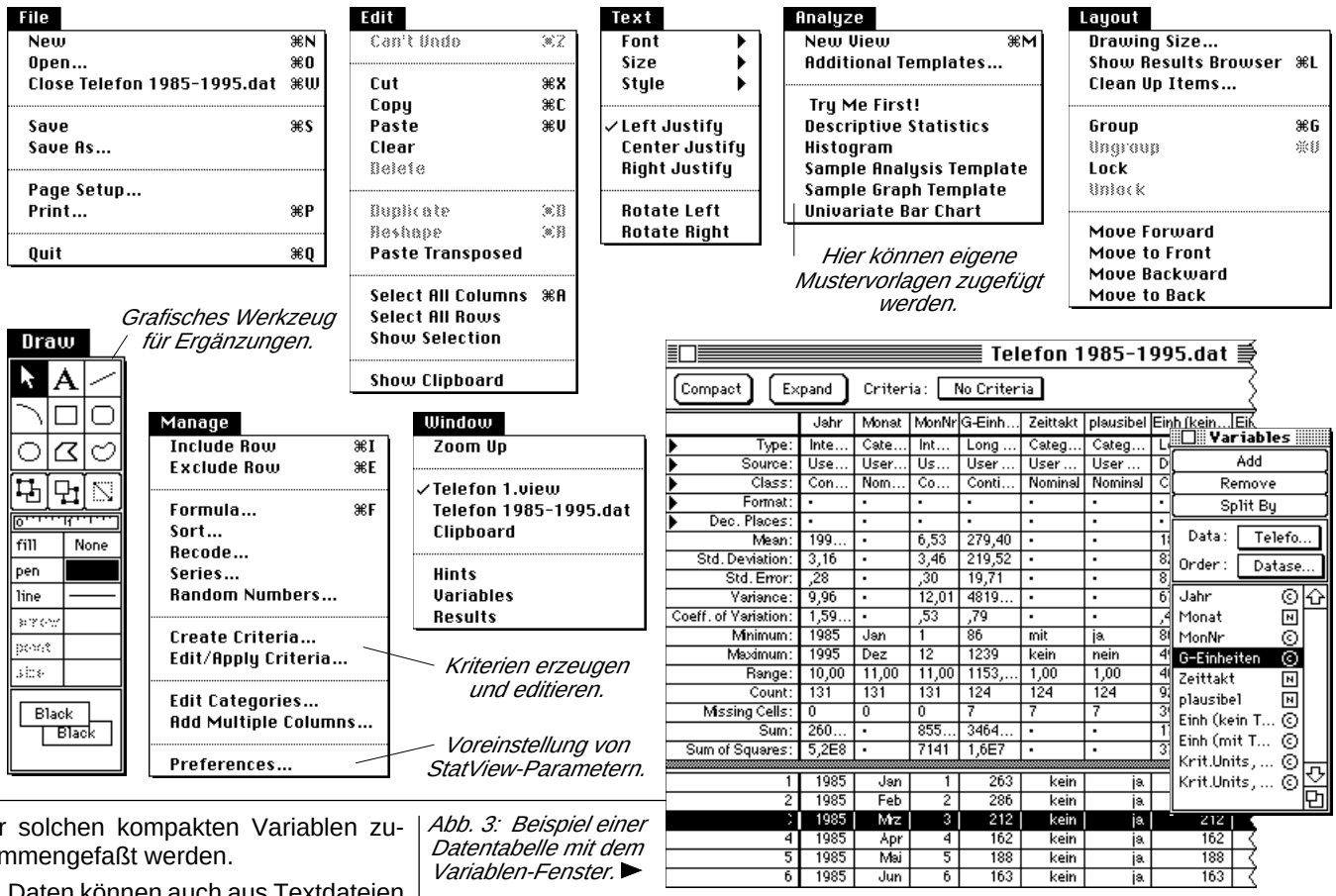


Abb. 3: Beispiel einer Datentabelle mit dem Variablen-Fenster. ►

ner solchen kompakten Variablen zusammengefaßt werden.

Daten können auch aus Textdateien importiert werden. Die Daten lassen sich auch aus einem anderen Fenster einer anderen Applikation über die Zwischenablage per „Cut & Paste"-Technik in die Datentabelle einfügen. Sobald alle Daten eingegeben sind, können diese mit den Befehlen aus dem View-Fenster (Abb. 1, linke Seite) analysiert werden.

Besonders vorteilhaft ist, daß sich beliebige Kriterien formelmäßig definieren lassen. Mit solchen Kriterien kann die Analyse auf (den Kriterien genügende) Teilmengen der Datentabelle beschränkt werden. Es können beliebig viele Kriterien definiert werden.

6. Datentypen und -klassen:

Es werden die folgenden Datentypen unterschieden: ***Integer*** – ganze Zahlen ($-32\,767 \le z \le 32\,767$), ***Long Integer*** – ganze Zahlen ($2\,147\,483\,647 \le z \le 2\,147\,483\,647$), ***Real*** – Gleitkommazahlen, ***Date/Time*** – Datums- und Zeitdaten, ***Currency*** – Daten, die als Währung angezeigt werden, ***String*** – alphanumerische Daten max. 255 Zeichen lang (z. B. Namen), ***Category*** – alphanumerische Bezeichnungen für die Kategorien von Stichproben.

Es gibt drei Datenklassen: ***Kontinuierliche*** (Beispiel: Körpergewicht), ***nominale*** (Legen auch eine Gruppenzugehörigkeit fest. Beispiel: Geschlecht) und ***informative*** (Beispiel: Namen) Daten.

7. Das Arbeitsfenster (View):

Alle Analysen erfolgen in einem View-Fenster (Abb. 1), das mit dem Befehl »New View« aus dem Menü »Analyze« erzeugt wird. Durch Doppelklick auf die Analyseart im View-Fenster wird zunächst ein Template der Analyse in das View-Fenster gelegt, in dem weitere Handlungsanweisungen mitgeteilt werden.

Die gewünschten Variablen werden im Variablen-Fenster (Abb. 3) ausgewählt. Sind alle erforderlichen Variablen und Parameter festgelegt, wird die Analyse berechnet und im View-Fenster angezeigt.

8. Export der Ergebnisse:

Zwar lassen sich alle Ergebnisse auf Papier ausdrucken, dennoch benötigt man oft die Ergebnisse an anderer Stelle zur Weiterverarbeitung. Deshalb können die Ergebnisse auch als Dateien exportiert werden.

Alle Tabellen, auch die in den View-Fenstern erzeugten, lassen sich mit dem Befehl »Save As…« aus dem Menü »File« als Textdateien exportieren. Will man jedoch die grafische Anordnung der Tabelle oder die Plots exportieren, so müssen diese als Dateien im PICT-Format gesichert werden, was mit dem Popup-Menü in der File-Dialogbox eingestellt wird. Solche Dateien können beispielsweise vom PageMaker [4] importiert und in einem Dokument positioniert werden.

9. Sonstige Hinweise:

■ Das Handbuch [1] enthält im Kapitel 2 ein ausführliches Tutorial. Es empfiehlt sich sehr, diese Beispiele durchzuarbeiten. Ein Lektüre-Muß ist auch das Kapitel 3, in dem über die richtige Verwendung von Datasets, den Datentabellen berichtet wird.

■ Jedes View-Fenster kann auch als eine Mustervorlage für spätere Analysen gesichert werden. Legt man eine solche Datei im Ordner »StatView Templates« ab, dann kann diese Vorlage in das Menü »Analyze« aufgenommen werden.

■ Zum *StatView* sind Ergänzungsmodule lieferbar: QC Tool, Survival Analysis, SuperANOVA. Diese stehen aber zur Zeit nicht zur Verfügung.

■ Ein Vergleich von [2] mit ***StatView*** ist interessant und zeigt, daß die Computerwelt in den letzten 30 Jahren doch sehr erhebliche Fortschritte gemacht hat, und man heute nicht einmal mehr die Geigy-Tabellen [3] zum Nachschlagen braucht. –

10. Literatur:

[1] 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. – Dieses Manual steht beim WDVMAC im Raum K 60 im Regal. Es darf dort nicht entfernt werden. – [CHB 854/1].

[2] Dittberner, K.-H.: Berechnung statistischer Kennwerte einer Stichprobe. FU Berlin (IfP): wdv-notes Nr. 218, 1967–1994.

[3] Documenta Geigy: Wissenschaftliche Tabellen. Basel: Geigy 1969 (7. Auflage).

[4] Dittberner, K.-H.: MAC-Software: Aldus PageMaker V 5.0. FU Berlin (IfP): wdv-notes Nr. 238, 1993–1994.

[5] Dittberner, K.-H.: StatView: Bearbeiten von Daten. FU Berlin (IfP): wdv-notes Nr. 365, 1994–1995.