

Dateneingabe und Transformation

Übersicht

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

- 2.0 Allgemeine Syntax
- 2.1 Eingabe über die Tastatur
- 2.2 Transformationen
- 2.3 Eingabe eines externen ASCII-Files
- 2.4 Varianten der INPUT-Anweisung,
Formatierungselemente
- 2.5 Ein- und Ausgabe von SAS-Systemfiles
- 2.6 Zusammenfügen von Files
- 2.7 OUTPUT-Anweisung
- 2.8 DO-Schleifen im DATA-Step

Dateneingabe und Transformation

Allgemeine Syntax

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

```
DATA <dateiname <(dateioptionen)>>;
```

...

```
RUN;
```

<... > kennzeichnet optionale Parameter

Externes File

```
INFILE ' ... ' ;  
INPUT ... ;
```

Tastatur

```
INPUT ... ;  
CARDS;
```

Daten

```
;
```

SAS-System-File

```
SET SAS-dateiname;
```

+ zusätzliche Anweisungen

Programmbeispiele: Eingabe... .sas

Dateneingabe und Transformation

Eingabe über die Tastatur

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

```
DATA Eingabe1;
```

```
INPUT a $ x y z;
```

```
s = x + y + z;
```

```
CARDS;
```

```
b 1 2 3
```

```
c 4 5 6
```

```
d 7 8 9 ;
```

```
RUN;
```

```
/* Erläuterung dazu: siehe Datei Eingabe.sas. */
```

```
PROC PRINT; RUN;
```

Mit PROC PRINT wird die gesamte erzeugte Datei
ausgedruckt ins Output-Fenster.

Dateneingabe und Transformation

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Aktivierung des Programms

- klicken auf MännchenLogoGrafik oder
- klicken auf 'run' → 'submit' oder
- F3-Taste

Die Datei Eingabe1 hat

3 Beobachtungen (Individuen, Wertesätze)
5 Merkmale (Variablen) a, x, y, z und s.

Dateneingabe und Transformation

Alternative Besichtigung der Daten

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Solutions

Analysis

Interactive Data Analysis

je nach DATA-Kommando:

	Bibliothek	Dateiname
DATA Eingabe1;	WORK	Eingabe1
DATA sasuser.Eing1;	SASUSER	Eing1
DATA;	WORK	DATA1 DATA2 ...

Bemerkung:

Dateien, die sich im Arbeitsverzeichnis WORK befinden, werden am Ende der Sitzung gelöscht.

Die Variante "DATA sasuser.Eing1;" nicht verwenden.

Dateneingabe und Transformation

Automatisch generierte Variablen

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

N oder *obs*

gibt die aktuelle Beobachtungsnummer an.

ERROR_

- Nichtzulässige mathematische Operationen führen zu *_ERROR_* = 1 und das Ergebnis wird auf “.” (missing value) gesetzt. (vgl. Beispiel Eingabe2)
- Schlimmere Fehler führen zu höherem *_ERROR_-*Wert.

Dateneingabe und Transformation

Transformationen

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

- immer nach der INPUT-Anweisung angeben!

IF THEN ELSE und logische Operationen

vgl. Programm Eingabe2

Funktionen

vgl. Programm Eingabe3

Arithmetische Operationen

+, -, *, /, **

IF(log. Ausdruck)

nur bestimmte Wertesätze einlesen

Es werden nur die Wertesätze eingelesen, die die logische Bedingung erfüllen.

Dateneingabe und Transformation

IF THEN ELSE

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

jeweils nur eine Anweisung ausführen

```
IF (log. Ausdruck) THEN Anweisung;  
ELSE Anweisung;
```

jeweils mehrere Anweisungen ausführen

- IF (log. Ausdruck) THEN Anweisung;
ELSE DO
Anweisung1; Anweisung2; ... END;
- IF (log. Ausdruck) THEN DO
Anweisung1; ... END;
ELSE DO
Anweisung1; Anweisung2; ... END;

Eingabe durch externes File (ASCII)

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

DATA Eingabe4;
INFILE 'Pfadname';
INPUT Variablen;
evtl. Transformationen;
RUN;

- Diese Eingabe ist formatfrei, d.h. die Variablen sind im Rohdatenfile durch Leerzeichen getrennt.
- Sind die Eingabedaten durch ein anderes Zeichen, z.B. ';', getrennt, dann ist in der INFILE-Anweisung die Option DELIMITER=';' (oder DLM=';') anzugeben.

Tabulatorzeichen: DLM= ' 09 ' X ;

- Bedingungen:
fehlende Werte: . (Punkt)
alphanumerische Zeichenketten dürfen keine
Leerzeichen enthalten.
- Die INPUT-Anweisung kann auch abgekürzt
werden, z.B.
INPUT V1-V7;

Eingabe durch externes File (EXCEL)

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

```
PROC IMPORT datafile="... .xls";  
out Dateiname; /*SAS-Datei*/  
getnames=no; /*Variablennamen werden nicht  
übernommen*/  
  
sheet=spreadsheetname;  
RUN;
```

Wichtige Varianten der INPUT-Anweisung

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

- bisher: formatfrei
INPUT a \$ b \$ c d;
- formatiert-spaltenorientiert
INPUT a \$ 1-10 b \$ 11 c 13-14 .1;
- formatiert-über die Zeichenlänge
INPUT a \$10. b \$ 1. c 2. d 5.1;

Eingabeformate

w. 2. standard numerisch

w.d 2.1 standard numerisch mit Dezimalstelle

\$w. \$10 Zeichenlänge

Nachgestelltes \$-Zeichen steht für Zeichenketten.

Eingabe5

Eingabe6 (komplexere Formate)

Weitere Formatierungselemente

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Spaltenzeiger

@n: Zeige auf Spalte n (z.B. @12)

+n: Setze den Zeiger n Positionen weiter

Zeilenzeiger

n: Zeige auf Spalte 1 der n-ten Zeile

Zeilenhalter

@ (nachgestellt) Datenzeile wird von mehreren
INPUT-Anweisungen gelesen

@@ (nachgestellt) Aus einer Eingabezeile werden
mehrere Beobachtungen
gelesen

Ein- u. Ausgabe von SAS-Files

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Abspeichern einer permanenten SAS-Datei

```
DATA sasuser.banknote; /* Eine Datei mit  
dem Namen 'banknote' wird im SAS-internen  
Verzeichnis 'sasuser' gespeichert */  
<INFILE ' Pfadname der einzulesenden Datei;>  
INPUT Formatangaben;  
<CARDS;  
Daten (zeilenweise); >  
RUN;
```

Ein- u. Ausgabe von SAS-Files

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Abspeichern einer permanenten SAS-Datei

```
DATA sasuser.banknote; /* Eine Datei mit  
dem Namen 'banknote' wird im SAS-internen  
Verzeichnis 'sasuser' gespeichert */  
<INFILE ' Pfadname der einzulesenden Datei;>  
INPUT Formatangaben;  
<CARDS;  
    Daten (zeilenweise); >  
RUN;
```

Einlesen einer SAS-Datei

```
DATA banknote1;  
    SET sasuser.banknote < (Optionen)>;  
RUN;
```

Ein- u. Ausgabe von SAS- Files

Einige Optionen

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

DROP = Varname(n);	Weglassen von Variablen
KEEP = Varname(n);	nur diese Variablen werden verwendet
FIRSTOBS=integer;	1. zu verarbeitender Wertesatz
OBS = integer;	letzter zu verarbeitender Wertesatz
RENAME = (alter Varname = neuer Varname);	

Ausgabe

Formatierte Ausgabe

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

```
DATA;  
Pi=3.141592;  
FORMAT Pi 5.3;  
OUTPUT;  
STOP;  
RUN;  
Standard: 8 Zeichen.
```

Ausgabe

Formatierte Ausgabe

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

```
DATA;  
Pi=3.141592;  
FORMAT Pi 5.3;  
OUTPUT;  
STOP;  
RUN;  
Standard: 8 Zeichen.
```

Längere Variablennamen

vor die INPUT-Anweisung:

```
LENGTH Var.name $länge;  
z.B. LENGTH Var.name $12;
```

Zusammenfügen von Files

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Files 'untereinander'

SASfile_1

...

SASfile_n

```
DATA; /* Eingabe_Banknote13.sas */
```

```
SET SASfile_1 <(options)>
```

```
... SASfile_n<(options)>;
```

```
RUN;
```

Zusammenfügen von Files

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Files 'untereinander'

SASfile_1

...

SASfile_n

```
DATA; /* Eingabe_Banknote13.sas */
```

```
SET SASfile_1 <(options)>
```

```
... SASfile_n<(options)>;
```

```
RUN;
```

Files 'nebeneinander'

SASfile_1 ... SASfile_n

```
DATA; /* Eingabe_Banknote34.sas */
```

```
SET SASfile_1; SET SASfile_2;
```

```
... SET SASfile_n; RUN;
```

Sortieren und Zusammenfügen von Dateien

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Sortieren von Dateien

PROC SORT DATA=SASfile; BY nr; RUN;
nr gibt das Merkmal an, nach dem sortiert
werden soll.

Sortieren und Zusammenfügen von Dateien

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Sortieren von Dateien

PROC SORT DATA=SASfile; BY nr; RUN;
nr gibt das Merkmal an, nach dem sortiert werden soll.

Zusammenfügen von Dateien

MERGE SASfile_1 SASfile_2; BY nr; RUN;
Die Dateien müssen nach dem Merkmal nr sortiert sein!

Wie bei SET sind auch hier Optionen möglich.

Output-Anweisung

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

- dient der Ausgabe von Dateien
- es können mehrere Dateien gleichzeitig ausgegeben werden
- die Namen der auszugebenden Dateien erscheinen im DATA-Step.

```
Eingabe12.sas
```

DO-Schleifen im DATA-Step

Werkzeuge der
empirischen
Forschung

W. Kössler

Einleitung

Datenbehandlung

Syntax

Tastatur

Transformationen

Externes File

Input-Anweisung

SAS-Files

Zusammenfügen

Output-Anweisung

DO-Schleifen

Allgemeine Syntax

- DO Indexvariable = Anfangswert
 <TO Endwert> <BY Schrittweite>;
 END;
- DO WHILE (Ausdruck) | UNTIL (Ausdruck);

Eingabe13.sas

Eingabe14.sas

Eingabe15.sas