

Datenmanagement

Berechnen von Variablen



Transformieren → Berechnen

- Es muss die Zielvariable angegeben werden. Die Berechnung kann für ausgewählte Fälle erfolgen. Unter „numerischem Ausdruck“ wird die Berechnung eingegeben.

Arithmetische Ausdrücke

Priorität	Operator	Bedeutung
1	**	Exponieren
2	*	Multiplizieren
3	/	Dividieren
4	+	Addieren
5	-	Subtrahieren

Anwendung

Fre 212 to v217.

Exe.

COMPUTE Internetnutzung=(v212+v213+v214+v215+v216+v217)/6.

EXECUTE.

Funktionen

- Es stehen in SPSS zahlreiche Funktionen zur Verfügung.

Beispiele für arithmetischer Funktionen

Funktion	Bedeutung
abs(x)	absoluter Betrag
rnd(x)	gerundeter Wert
sqrt(x)	Quadratwurzel

Beispiele für statistische Funktionen

Funktion	Bedeutung	Voreinstellung für n
sum.n(x, x, . . .)	Summe der Werte der Argumente	1
mean.n(x, x, . . .)	Mittelwert	1
min.n(x, x, . . .)	Minimum	1
max.n(x, x, . . .)	Maximum	1

- Der optionale Zusatz **.n kann** angegeben werden, um festzulegen, wie viele Argumente mindestens einen gültigen Wert vorweisen müssen, damit der Funktionswert berechnet werden kann.

Anwendung

```
COMPUTE Internetnutzung_2=MEAN(v212,v213, v214, v215, v216, v217).  
EXECUTE.
```

```
COMPUTE Internetnutzung_3=MEAN.5(v212,v213, v214, v215, v216, v217).  
EXECUTE.
```

```
FREQUENCIES VARIABLES=Internetnutzung Internetnutzung_2 Internetnutzung_3  
/ORDER=ANALYSIS.
```

Umkodieren

Automatisch umkodieren

Es gibt die Option selbst die Einteilung anhand verschiedener Kriterien (z.B. die Anzahl der Quartile) zu definieren.



Transformieren → Visuelles Klassieren (Bereichseinteiler)

Umkodieren



Transformieren → Umkodieren in andere (!) Variable

Anwendung

Visuelles Klassieren.

*v58.

```
RECODE v58 (MISSING=COPY) (LO THRU 39.0=1) (LO THRU 56.0=2) (LO THRU HI=3)
(ELSE=SYSMIS) INTO alter_gr.
```

```
VARIABLE LABELS alter_gr 'ALTER: BEFRAGTE<R> (Klassiert)'.
FORMAT alter_gr (F5.0).
```

```
VALUE LABELS alter_gr 1 ' 2 ' 3 '.
```

```
MISSING VALUES alter_gr (997.0 THRU HI).
```

```
VARIABLE LEVEL alter_gr (ORDINAL).
```

```
EXECUTE
```

```
FREQUENCIES VARIABLES=alter_gr
```

```
/ORDER=ANALYSIS.
```

```
RECODE v58 (0 thru 18=1) (18 thru 65=2) (66 thru Highest=3) INTO alter_gr_2.
```

```
EXECUTE.
```

```
FREQUENCIES VARIABLES=alter_gr_2
```

```
/ORDER=ANALYSIS.
```

Dummyvariablen

Vorteil:

- Es können auch kategoriale Variablen im Modell berücksichtigt werden
- Es können nicht lineare Zusammenhänge aufgedeckt werden

Eine mögliche Strategie zur Anpassung des Modells an nichtlineare Beziehungen in Daten ist die Konstruktion von Dummy-Variablen. Dabei müssen immer $n_i - 1$ Variablen im Modell integriert werden (n = Anzahl der Kategorien).

Interpretation:

- Bei dichotomen Variablen (z.B. Geschlecht 0 = männlich und 1 = weiblich) steht der Regressionskoeffizient für den Unterschied zwischen den beiden Ausprägungen.
- Bei mehr als 2 Ausprägungen wird eine als Referenzkategorie deklariert. Diese wird nicht mit ins Modell einbezogen

Anwendung: Alter zu Dummyvariablen

- alter_unter18 = 1, wenn Alter zwischen 0 und 18, ansonsten = 0.
- alter_19bis65 = 1, wenn Alter zwischen 19 und 65, ansonsten = 0.
- alter_65aelter = 1, wenn Alter größer 65, ansonsten = 0.

```
RECODE alter_gr_2 (0 thru 18=1) (ELSE=0) INTO alter_gr_2_unter18.  
EXECUTE.
```

...

Interpretation

- Ist nur die Dummy Variable enthalten, entspricht die Konstante dem Wert der Referenzkategorie (bei dichotomen Variablen Ausprägung = 0, bei mehreren Dummyvariablen, die Variable, die nicht im Modell enthalten ist).
- Der Regressionskoeffizient repräsentiert den Unterschied der jeweiligen Kategorie zur Referenzkategorie.

Aufgabe zum 20.01.2009

1. Überarbeiten Sie Ihr Modell.
2. Ziehen Sie eine kleine Stichprobe aus dem Allbusdatensatz ($n = 100$) und rechnen Sie das Modell erneut.
3. Wie würden Sie das Modell nun interpretieren?
4. Verwenden Sie das im Seminar erworbene Wissen und diskutieren Sie die Befunde z.B. anhand alternative Signifikanzniveaus, Effektstärken, Poweranalyse. Sie können dafür auch wieder auf die bivariate Ebene zurückkehren und nur den Zusammenhang von 2 Variablen untersuchen.
5. Senden Sie mir Ihre Ergebnisse bis zum 19.01.2009