

PROC FREQ für einfache Häufigkeiten

zum Berechnen absoluter und relativer Häufigkeiten, absoluter und relativer Summenhäufigkeiten

Allgemeine Form:

```
PROC FREQ DATA=name ;  
  TABLES variablenliste / Optionen ;  
RUN ;
```

Beispiel und Beschreibung der Programm-Statements:

```
TITLE1 'Einfache Häufigkeiten' ;  
TITLE2 '-----' ;  
PROC FREQ DATA=beispiel ;  
  TABLES geburtsmonat geschlecht ;  
RUN ;
```

Die Prozedur beginnt mit PROC endet mit RUN ;

Das ; am Ende eines Befehls beendet eine Anweisung innerhalb der Prozedur.

FREQ ist der Name der Prozedur.

DATA = beispiel legt das zu verwendende Datenfile fest (im Beispiel beispiel).

TABLES ist das gehörende Schlüsselwort und steht für Tabellen.

Ohne TABLES - Statement berechnet SAS Häufigkeiten für alle Merkmale in der Datendatei.

Es sind beliebig viele TABLES-Statements in einer PROC FREQ möglich.

Die Aufzählung der Merkmalsnamen darf in einem TABLES-Statement auch über mehrere Zeilen gehen. Das Semikolon kommt dann nach der letzten Variablen.

variablenliste wird mit den Namen der zu verarbeitenden Merkmale überschrieben. Trennen Sie die Merkmalsnamen mit Leerzeichen!!!

Mit dem TITLE-Befehl stellen Sie den Ergebnissen Überschriften voran. Der Befehl ist nicht zwingend notwendig, bringt aber Struktur in die Auswertung. **Achtung** - SAS übernimmt Titel in nachfolgende Prozeduren, wenn dort das TITLE-Statement fehlt.

Eine Beschreibung der Ergebnisse der Prozedur, die SAS im Output-Fenster zeigt, finden Sie auf der Rückseite.

Starten Sie die Programmzeilen aus dem Beispiel, zeigt SAS im Output-Fenster folgende Informationen:

The FREQ Procedure

Geburtsmonat

GEBURTSMONAT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Januar	7	7.69	7	7.69
Februar	2	2.20	9	9.89
Maerz	8	8.79	17	18.68
April	9	9.89	26	28.57
Mai	14	15.38	40	43.96
Juni	10	10.99	50	54.95
Juli	5	5.49	55	60.44
August	10	10.99	65	71.43
September	5	5.49	70	76.92
Oktober	7	7.69	77	84.62
November	6	6.59	83	91.21
Dezember	8	8.79	91	100.00

Frequency Missing = 2

Geschlecht

GESCHLECHT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
weiblich	62	66.67	62	66.67
maennlich	31	33.33	93	100.00

Beschreibung des Outputs:

Die Häufigkeiten beziehen sich auf die Gesamtzahl der Beobachtungen abzüglich der Datensätze mit fehlenden Werten. Im Beispiel fehlen bei zwei Beobachtungen die Angaben für den Geburtsmonat. SAS zeigt dies mit der Meldung **Frequency Missing = 2**.

Frequency = absolute Häufigkeiten.

Beispiel Geburtsmonat: 7 im Januar, 2 im Februar, 8 im März, ...

Beispiel Geschlecht: 62 Frauen und 31 Männer

Percent = relative Häufigkeiten.

Beispiel Geburtsmonat: 7,69% im Januar, 2,20% im Februar, 8,79% im März, ...

Beispiel Geschlecht: 62 Frauen und 31 Männer

Cumulative Frequency = absolute Summenhäufigkeiten.

Aufaddieren der absoluten Häufigkeiten in aufsteigender Reihenfolge beginnend bei der kleinsten Ausprägung. Im Beispiel kann man damit beispielsweise die Aussage treffen, dass 50 von 91 Personen im ersten Halbjahr geboren sind.

Beispiel Geburtsmonat: 7 im Januar, 9 im Januar und Februar, 17 im 1. Vierteljahr, ...

Beispiel Geschlecht: hier macht diese Angabe keinen Sinn.

Cumulative Percent = relative Summenhäufigkeiten.

Aufaddieren der relativen Häufigkeiten in aufsteigender Reihenfolge beginnend bei der kleinsten Ausprägung. Im Beispiel kann man damit beispielsweise die Aussage treffen, dass 54,95% der Personen im ersten Halbjahr geboren sind.

Beispiel Geburtsmonat: 7,69% im Januar, 9,89% im Januar und Februar, ...

Beispiel Geschlecht: hier macht diese Angabe keinen Sinn.

Beschreibung nützlicher Optionen:

- Option **NOCUM**: SAS unterdrückt die Berechnung der Summenhäufigkeiten:

```
PROC FREQ DATA=beispiel ;
  TABLES geburtsmonat / NOCUM ;
RUN ;
```

⇒ Die Berechnungen für **Cumulative Frequency** und **Cumulative Percent** werden nicht ausgegeben:

The FREQ Procedure

Geburtsmonat

GEBURTSMONAT	Frequency	Percent
Januar	7	7.69
Februar	2	2.20
Maerz	8	8.79
April	9	9.89
Mai	14	15.38
Juni	10	10.99
Juli	5	5.49
August	10	10.99
September	5	5.49
Oktober	7	7.69
November	6	6.59
Dezember	8	8.79

Frequency Missing = 2

- Option **NOPERCENT**: SAS unterlässt die Ausgabe der relativen Häufigkeiten und der relativen Summenhäufigkeiten:

```
PROC FREQ DATA=beispiel ;
  TABLES geburtsmonat / NOPERCENT ;
RUN ;
```

⇒ Die Angaben für **Percent** und **Cumulative Percent** fehlen:

The FREQ Procedure

Geburtsmonat

GEBURTSMONAT	Frequency	Cumulative Frequency
Januar	7	7
Februar	2	9
Maerz	8	17
April	9	26
Mai	14	40
Juni	10	50
Juli	5	55
August	10	65
September	5	70
Oktober	7	77
November	6	83
Dezember	8	91

Frequency Missing = 2