

Häufigkeitstabellen mit PROC SQL – Eine Alternative?

Murat Ipek

PAREXEL Int., Deutschland; murat.ipek@parexel.com

Häufigkeitstabellen dienen der übersichtlichen Darstellung von Daten. Der Leser möchte neben den absoluten Werten eines Merkmals – wie beispielsweise die Anzahl der männlichen Teilnehmer an einer Befragung – auch deren Verteilung in der erhobenen Stichprobe erfahren – wie beispielsweise 23 der 59 befragten ($\approx 39\%$) sind männlich.

Bei einem eindeutigen Merkmal wie das Geschlecht ist die Berechnung unkompliziert und kann i.d.R. mit einem einzigen Prozedurenschritt in SAS berechnet werden. Betrachtet man Merkmale die wiederholt erhoben/beobachtet werden, wie beispielsweise das Auftreten von Kopfschmerzen in einer Interventionsstudie, so möchte man neben der absoluten und relativen Anzahl der Ereignisse auch wissen, bei wie vielen Personen mindestens einmal Kopfschmerzen beobachtet wurden.

Eine Darstellung dieser Art ist im Vorfeld verbunden mit einigen Datenmanipulationen um die Daten in die gewünschte Struktur aufzubereiten. Neben Summary-Prozeduren bietet SAS mit PROC SQL eine alternative Möglichkeit Häufigkeitstabellen effizient und ohne Datenmanipulation zu erstellen.