

## AUFGABE 11

In der Datei MARGARINE.SAV (ist in AUFGABEN.ZIP enthalten)<sup>1</sup> sind für 11 Butter- und Margarinemarken (Sanella, Homa etc.) Werte für 10 Variablen (Streichfähigkeit, Haltbarkeit etc.) erfasst. Die einzelnen Marken wurden von 32 Probanden hinsichtlich jeder der Eigenschaftsvariablen auf einer siebenstufigen Ratingskala von hoch bis niedrig beurteilt. Als Messniveau wird Intervallskalenniveau angenommen. Aus den 32 Urteilen über eine Eigenschaft einer Marke wurde ein Durchschnittswert gebildet. Die Durchschnittswerte bilden die Variablenwerte der SPSS-Datei.

- a. Berechnen Sie die quadrierten Euklidischen Distanzen zwischen den Butter- und Margarinemarken. Dabei sollen alle 10 Variablen einbezogen werden. Welche beiden Marken haben die kleinste und welche beiden Marken haben die größte Distanz?
- b. Führen Sie mit allen 10 Variablen eine hierarchische Clusteranalyse durch. Wählen Sie als Distanzmaß die quadrierte Euklidische Distanz und als Clustermethode das Single-Linkage-Verfahren (Nächstgelegener Nachbar). Erläutern Sie den hierarchischen Prozess der Clusterbildung in den einzelnen Stufen. Begründen Sie, warum die Autoren Backhaus u. a. die Marke Delicado als Ausreißer bewerten und deshalb für ein Clustern der Butter- und Margarinemarken empfehlen, diese Marke auszuschließen.
- c. Führen Sie unter Ausschluss der Marke Delicado eine hierarchische Clusteranalyse durch. Alle 10 Variablen sollen einbezogen werden. Als Distanzmaß soll die quadrierte Euklidische Distanz und als Methode sowohl das Wardverfahren als auch das Zentroid-Clustering verwendet werden. Beurteilen Sie die Ergebnisse.
- d. Überprüfen Sie die gefundene Clusterlösung der Butter- und Margarinemarken mit der Clusterzentrenanalyse (K-Means-Cluster). Dabei soll die Marke Delicado nicht einbezogen werden.

---

<sup>1</sup> Die Daten entstammen K. Backhaus, B. Erichson, W. Plinke, R. Weiber, Multivariate Analysemethoden.