Aufgabenstellung „Warenkorbanalyse“

Ein Supermarkt verkauft Waren zehn verschiedener Kategorien (siehe unten). In jeder Kategorie gibt es mehrere Ladenhüter (Waren die sich sehr schlecht verkaufen). Um den Gewinn zu optimieren, beschließt die Geschäftsleitung, eine Warenkorbanalyse durchführen zu lassen. Dazu wird ein Bonusprogramm namens „BACKPAY“ eingeführt, mit dem die Kunden mittels Kundenkarte Bonuspunkte sammeln und später gegen winzige Prämien eintauschen können. Viel wichtiger für die Geschäftsführung: damit können die Daten für die Warenkorbanalyse gesammelt werden. Anhand der Analyse sollen nun Kombinationsangebote (Sonderangebote, die aus zwei unterschiedlichen Produkten bestehen, z.B. Bier (Kategorie Alkohol) und Windeln (Kategorie Toilettenartikel) zum „Kombinationssonderpreis“ zusammengestellt werden.
Die Analyse der Daten und Zusammenstellung geeigneter Kombi-Angebote ist Ihre Aufgabe!

Vorgehensweise:

Führen Sie eine Warenkorbanalyse durch: Ermitteln Sie statistisch relevante Assoziationsregeln. Sammeln Sie weitere nützliche Informationen, indem Sie die gefundenen Assoziationsregeln weiter analysieren. Clustern Sie hierzu die Transaktionen, die zum Support einer gefundenen Regel beitragen, nach den zusätzlichen Attributen. Machen Sie Vorschläge zur Zusammenstellung geeigneter Kombi-Angebote.

Milestones:

<u>Milestone 1 (Abnahme: 8.1./12.1.03)</u>

- Implementierung geeigneter Datenstrukturen (Itemsets)
- Implementierung des Apriori-Algorithmus

<u>Milestone 2 (Abnahme: 29.1./3.2.03)</u>

- Bestimmung von Assoziationsregeln aus den gefundenen *frequent itemsets*

Datensatz:

Im Datensatz „shopping“ finden Sie anonymisierte Daten zum Einkaufsverhalten der Kunden des Supermarkts. Jede Transaktion ist in einer separaten Zeile gespeichert, alle Attribute durch Komma getrennt. Die ersten zehn Attribute repräsentieren die gekauften Waren, sortiert nach:

- Fertiggerichte (ready made)
- Tiefkühlkost (frozen foods)
- Alkohol (alcohol)
- Gemüsefrischwaren (fresh vegetables)
- Milch (milk)
- Backwaren (bakery goods)
- Frischfleisch (fresh meat)
- Toilettenartikel (toiletries)
- Imbisse (snacks)
- Konserven (tinned goods)

Zusätzlich zu den gekauften Items wurden für jeden Käufer noch zusätzliche Attribute wie Geschlecht (gender), Altersgruppe (age), Familienstand (marital), Kinder (children), Arbeitsverhältnis (working) aggregiert.