

Knowledge Discovery in Databases
WS 2017/18

Übungsblatt 3: Repräsentation und Frequent Itemsets

Aufgabe 3-1 Schemes

Skizzieren Sie eine Data Warehouse Applikation für ein Universitäts-Informationssystem. Der Fokus soll auf **Vorlesungen** liegen, die von einem **Professor** zu einer bestimmten **Zeit**, in einem bestimmten **Raum** gehalten werden. **Studiengang**, **Adressen** und **Lehrstühle** sind ebenfalls von Interesse. Professoren gehören immer einem Lehrstuhl an. Geben Sie Fakten- und Klassifikationstabelle im Star- und im Snowflake-Schema an.

Aufgabe 3-2 Binning

Geben Sie equi-width- und equi-height Histogramme mit Binsgröße 3 für folgende Datenpunkte an:

1; 1; 4; 4; 5; 5; 5; 5; 6; 7; 7; 9

Aufgabe 3-3 Frequent Itemset Mining

Sei D eine Transaktionsdatenbank mit folgenden Transaktionen.

TID	items_bought
T1	{K, A, D, B}
T2	{D, A, C, E, B}
T3	{C, A, B, E}
T4	{B, A, D}

Außerdem sei $min_sup = 60\%$.

Bestimmen Sie alle Frequent Itemsets mit dem Apriori-Algorithmus.