

The background features a complex, abstract design. It includes a network of thin, light-colored lines forming a web-like structure. Overlaid on this are various data visualizations: a grid of small, light-colored squares on the left, a series of small, light-colored circles in the center, and a series of small, light-colored triangles on the right. The overall color palette is muted, with shades of gray, light blue, and light green.

Basic Concepts: Frequent Patterns and Association Rules

Basic Concepts: Frequent Itemsets (Patterns)

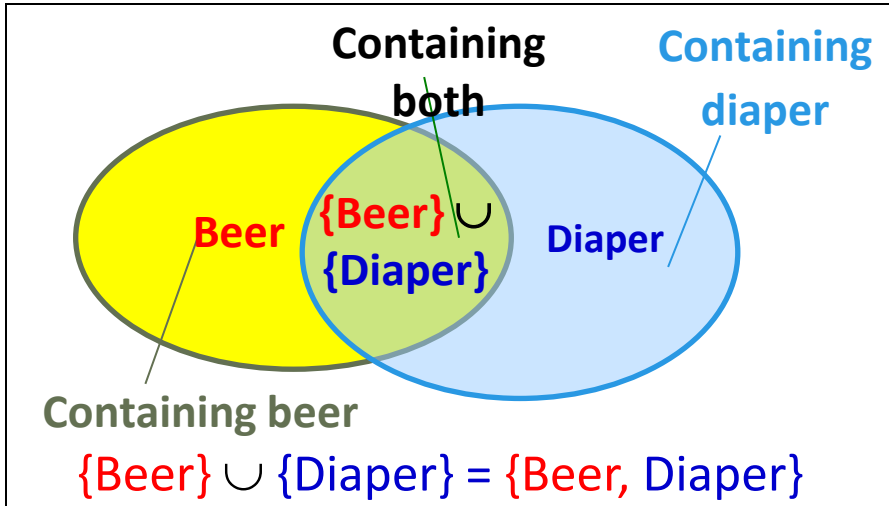
- ❑ **Itemset**: A set of one or more items
- ❑ **k-itemset**: $X = \{x_1, \dots, x_k\}$
- ❑ **(absolute) support (count)** of X :
Frequency or the number of occurrences of an itemset X
- ❑ **(relative) support**, s : The fraction of transactions that contains X (i.e., the **probability** that a transaction contains X)
- ❑ An itemset X is **frequent** if the support of X is no less than a *minsup* threshold (denoted as σ)

Tid	Items bought
10	Beer, Nuts, Diaper
20	Beer, Coffee, Diaper
30	Beer, Diaper, Eggs
40	Nuts, Eggs, Milk
50	Nuts, Coffee, Diaper, Eggs, Milk

- ❑ Let *minsup* = 50%
- ❑ Freq. 1-itemsets:
 - ❑ Beer: 3 (60%); Nuts: 3 (60%)
 - ❑ Diaper: 4 (80%); Eggs: 3 (60%)
- ❑ Freq. 2-itemsets:
 - ❑ {Beer, Diaper}: 3 (60%)

From Frequent Itemsets to Association Rules

Tid	Items bought
10	Beer, Nuts, Diaper
20	Beer, Coffee, Diaper
30	Beer, Diaper, Eggs
40	Nuts, Eggs, Milk
50	Nuts, Coffee, Diaper, Eggs, Milk



Note: Itemset: $X \cup Y$, a subtle notation!

- Association rules: $X \rightarrow Y (s, c)$
 - **Support**, s : The probability that a transaction contains $X \cup Y$
 - **Confidence**, c : The conditional probability that a transaction containing X also contains Y
 - $c = \text{sup}(X \cup Y) / \text{sup}(X)$
- **Association rule mining**: Find **all** of the rules, $X \rightarrow Y$, with minimum support and confidence
- Frequent itemsets: Let $\text{minsup} = 50\%$
 - Freq. 1-itemsets: Beer: 3, Nuts: 3, Diaper: 4, Eggs: 3
 - Freq. 2-itemsets: $\{\text{Beer, Diaper}\}$: 3
- Association rules: Let $\text{minconf} = 50\%$
 - $\text{Beer} \rightarrow \text{Diaper}$ (60%, 100%)
 - $\text{Diaper} \rightarrow \text{Beer}$ (60%, 75%) (Q: Are these all rules?)