

The background of the slide is a complex, abstract composition. It features a network of thin, reddish-brown lines forming a web-like structure. Scattered throughout are numerous small, colored dots in shades of green, blue, and orange. On the left side, there is a vertical strip with a grid of small, light-colored squares. In the upper left corner, there is a small inset image showing a cluster of orange and red dots. The overall aesthetic is technical and data-driven.

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
X = Beer, Nuts, Diaper
Beer: 3, Nuts: 3, Diaper: 4

□ **(relative) support**, s: The fraction of transactions that contains X (i.e., the **probability** that a transaction contains X)

X = Beer, $\frac{3}{5} \rightarrow$ transaction count

□ 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%)

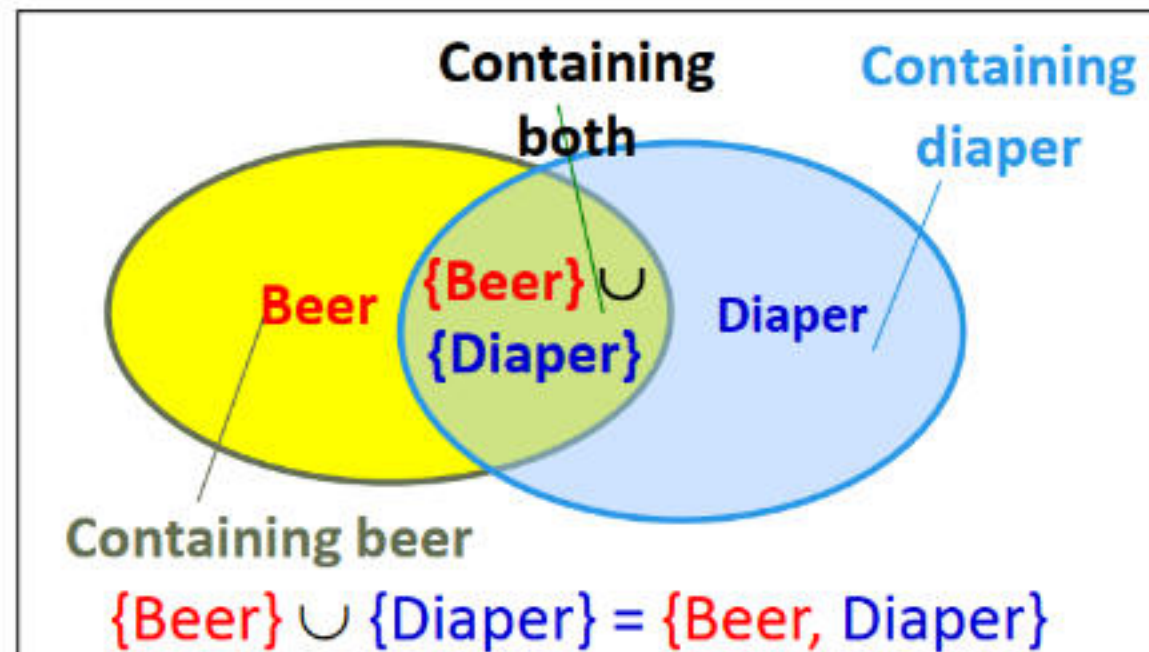
□ Diaper: 4 (80%); Eggs: 3 (60%)

□ 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: {Beer, Diaper}: 3
- Association rules: Let $\text{minconf} = 50\%$
 - $Beer \rightarrow Diaper$ (60%, 100%)
 - $Diaper \rightarrow Beer$ (60%, 75%) (Q: Are these all rules?)