

PROJECT 4 PART 1 (2 POINTS, DUE 3/1)

- Given the data set on the right, compute the following values **WITHOUT** using computer. Show your work.

- Find the support of
 $\{\text{Style}=\text{Cute}, \text{Size}=\text{M}, \text{Season}=\text{Summer}\}$
- Find the confidence for the following implication rules.
 - $\{\text{Style}=\text{Cute}, \text{Size}=\text{M}\} \rightarrow \{\text{Season}=\text{Summer}\}$
 - $\{\text{Style}=\text{Cute}, \text{Season}=\text{Summer}\} \rightarrow \{\text{Size}=\text{M}\}$

Dress ID	Style	Size	Season	Recommendation
1	Casual	L	Autumn	Y
2	Vintage	M	Spring	Y
3	Casual	M	Summer	Y
4	Brief	M	Summer	Y
5	Casual	S	Winter	Y
6	Cute	S	Summer	Y
7	Casual	M	Winter	Y
8	Cute	S	Summer	Y
9	Flare	S	Spring	Y
10	Casual	M	Winter	Y
11	Casual	L	Autumn	N
12	Casual	M	Summer	N
13	Casual	L	Summer	N
14	Cute	M	Summer	N
15	Casual	M	Autumn	N
16	Cute	M	Summer	N
17	Cute	M	Autumn	N
18	Casual	L	Spring	N
19	Cute	L	Summer	N
20	Brief	M	Spring	N

PROJECT 4 PART 2 (8 POINTS, DUE 3/9)

- Download the data from
https://archive.ics.uci.edu/ml/datasets/dresses_attribute_sales
- Read the data (using R and excel) and make it ready to use.
- Analyze the associations rules for attributes “Price”, “Size”, “Season”, “Waistline” and “Recommendation”. **Explain what the program does in each step.**

1. Frequent Itemset Generation

- a) Construct a data frame to record the support for all possible combinations of 3 values from attributes “Price”, “Size”, “Season”, “Waistline” and “Recommendation”.
- b) Find the most frequent 3-value itemset (the one has the maximum support value).

2. Rules Generation

- a) Make a list of **all 12** implication rules (**including the binary partitions $\{A, B\} \rightarrow \{C\}$ and $\{C\} \rightarrow \{A, B\}$, as well as the 2-element rules $\{A\} \rightarrow \{B\}$**) generated from the most frequent itemset found in 1.
- b) Compute the confidence for every implication rule in 2. (a).
- c) Find the implication rules with the highest confidence.
- d) Explain in words what your **mined** result mean?