

[https://bsu.instructure.com/courses/152094/files/16917845?module_item_id=4761888&fd_c](https://bsu.instructure.com/courses/152094/files/16917845?module_item_id=4761888&fd_cookie_set=1)
ookie_set=1Homework 5

Overview

This store offers personal care items for the shopper's convenience.

6 Distinct Products

1. Toothbrush
2. Toothpaste
3. Shampoo
4. Body wash
5. Deodorant
6. Ibuprofen

Shopping Baskets

Basket	Product
1 (2 items)	Toothbrush, toothpaste
2 (3 items)	Shampoo, deodorant, body wash
3 (4 items)	Toothbrush, toothpaste, shampoo, ibuprofen
4 (5 items)	Toothpaste, shampoo, body wash, deodorant, ibuprofen

How many rules can be discovered given $msv = 0\%$ and $mcv = 0\%$? Justify/Explain why.

Apriori algorithms are used to make observations in data and find 'rules' or patterns in the data. msv and mcv set parameters on these rules. Since both are set to 0% , this indicates that we would be interested in looking at every possible combination, disregarding if it's a rule that happens frequently or infrequently, so many potential rules can be discovered given $msv = 0\%$ and $mcv = 0\%$.

List two strong rules, and justify/explain why.

{Toothbrush} → {Toothpaste}

2 out of the 4 baskets, so half of the baskets in the dataset contain toothbrush and toothpaste. Both of these baskets which contain the toothbrush item also contain toothpaste, which means that whenever a toothbrush is added to a cart, toothpaste will always also be added. This 100% confidence makes this a strong rule.

{Bodywash} → {Shampoo}

Similarly 2 out of 4 baskets contain both body wash and shampoo. Every time bodywash is purchased, shampoo is purchased along with it. This confidence of 100% indicates that it is a strong rule. Although there are 3 total baskets that contain shampoo, when body wash is present in the basket, shampoo is too making this a strong rule.

List one impossible/unlikely rule, and justify/explain why.

{Ibuprofen} → {Toothbrush}

Ibuprofen was purchased in 2 out of the 4 total baskets and in neither instance was there also the purchase of a toothbrush. When Ibuprofen was purchased, none of the customers purchased toothbrushes. This makes the rule impossible, or extremely unlikely.

Summary/What we've learned

This homework assignment made me pay closer attention to analyzing patterns in data. Applying the apriori algorithm to a relatively simple case such as in this homework made me better understand how the algorithm works. It's interesting to now be able to understand how big companies like Walmart or Amazon make suggestions based on what is currently in your cart, and how the suggestions can vary given differing msv and mcv values. Being able to recognise the different algorithms in data mining gives me a bigger appreciation for its applications in big data.