

1. For the transactional database below, iterate through the Apriori algorithm using a minimum support threshold, $s = 2$.
2. Given a transactional database, give the support of two sequences: $\langle D \rangle$ and $\langle C(ED) \rangle$.
3. Given a transactional database and support threshold $s = 3$, mine the f-projected database. No need to further iterate.
4. Give the appropriate algorithm to use on this case (2/20)

Recommender System / AdaBoost / Gradient Boosted Trees / Sequence Mining / Association Rules Mining / Clustering

- a) To be able to contact people with the right commercials
 - b) on a movie database
 - c) when you want to compare players with a similar playstyle, find players with similar playstyle
 - a) the PageRank algorithm - how to construct a network of webpages
 - b) To discover certain style of play of a team
5. Exercise on DTW (Dynamic Time Warping): Given following numbers/matrix, do DTW with warppwindow $w=2$, give the connections and the flow.
 6. Data mining can be used as a pre-processing step for machine learning. Could association rule mining and clustering be used to construct features for machine learning? If you think so, explain how these features would be constructed? If not, why?
 7. Sales information has been received from a large retailer (E.g. Colruyt, Carrefour ...). Researcher analyze divorce rate. They state that more butter consumption leads to higher divorce rate. (Correlation vs causation) Is this statement valid? Why (not)?