

Datasets

1. **Product Table**: This contains a list of Product IDs and corresponding ratings for a number of product features. These ratings represent how high or low the products rank with respect to the various product features.
2. **Person Table**: This contains a list of Person IDs and corresponding ratings for a number of product features. These ratings represent the preference of each person with respect to the various product features.
3. **Feature Category Lookup**: This contains a lookup table showing the groupings of each feature within categories.

Objective

1. For each user in the **Person Table**, we want to rank each product in the **Product Table** in order of the user's preference for the product's features.

Steps

1. Per category in the Feature Category Lookup table:
 - a. Select/subset the Features that belong to the Category. E.g. Feature_34, Feature_35 belong to Category_B.
 - b. Find the cosine similarity between the above selected features in the **Product Table** and the same selected features in the **Person Table**, per user. Cosine similarity/distance provides a measure of the similarity between two vectors. More details and examples are available on the internet on web pages such as this:
<https://www.statology.org/cosine-similarity-r/>
2. Output will be a table that contains columns for Person_ID, Product_ID, Category names and values representing the cosine similarity values. See format of final output below.
3. Perform a weighted average of the values in each row using the following weightings:
 - a. Category_A: 2
 - b. Category_B: 1
 - c. Category_C: 5
 - d. Category_D: 10
4. Sort the products, per Person, from highest to lowest according to Weighted_Average value.

Output Format

Person_ID	Product_ID	Category_A	Category_B	Category_C	Category_D	Weighted_Average
Person_1	Product_1
Person_1	Product_2
Person_1	Product_3
...