6. Collaborative filtering recommendation method.

**Brief description of the method:**

Collaborative filtering is a recommendation method used to predict user’s interests by analyze other similar users, making recommendation based on similar users’ interests. This method can be classified into two main categories: user-based and item-based collaborative filtering (<https://medium.com/@pragya19126/grocery-recommendation-using-collaborative-filtering-fb30fed3fdca>).

User-based Collaborative Filtering:

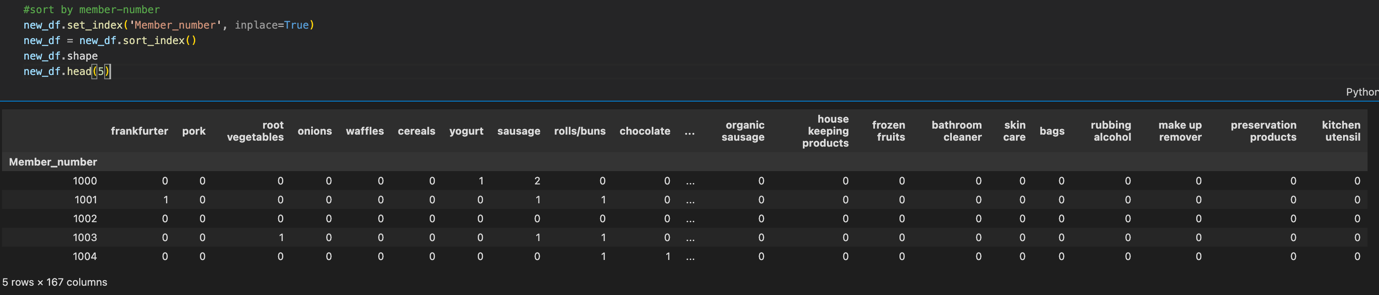
This method recommends items to a user based on the preferences of similar users. For example, if User A and User B have similar interests, and User A likes a particular item that User B has not yet rated, the system might recommend that item to User B.

Item-based Collaborative Filtering:

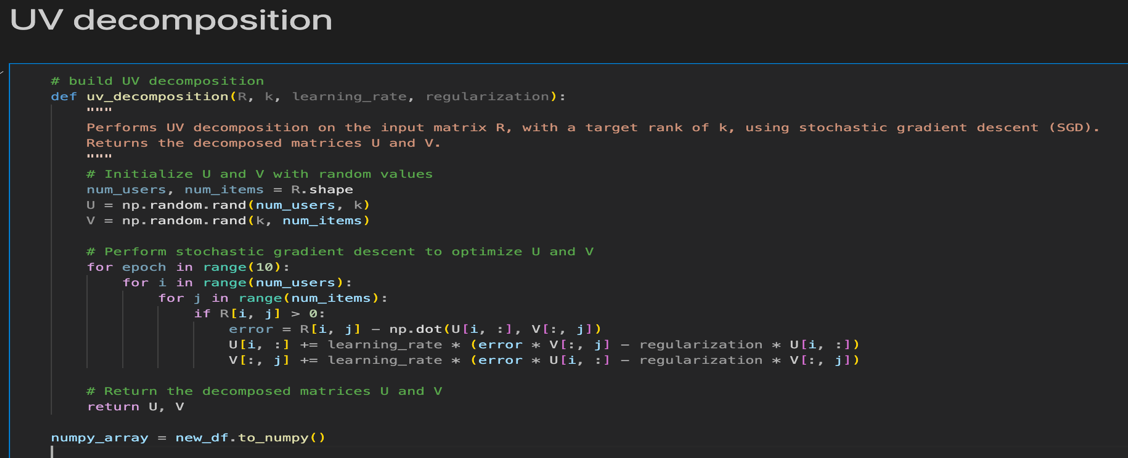
This method recommends items based on the similarity between items. For instance, if a user likes Item X and Item Y is similar to Item X, then Item Y might be recommended to the user.

In our recommendation system, we are using User-based Collaborative Filtering to generate recommendation. Steps are below:

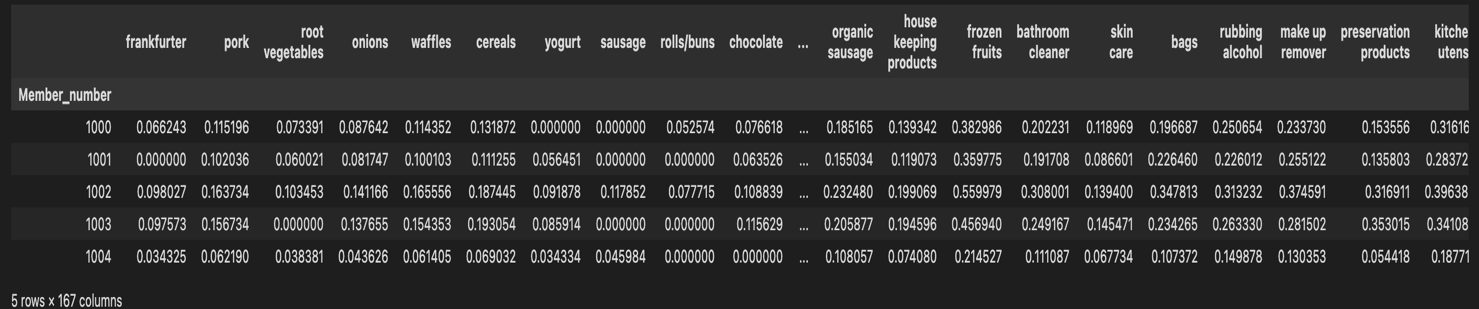
First, we processed the data, building a data frame one row a member, following with the items member had bought. We did this processing because we were finding the similar users and make recommendation based on what they bought before.



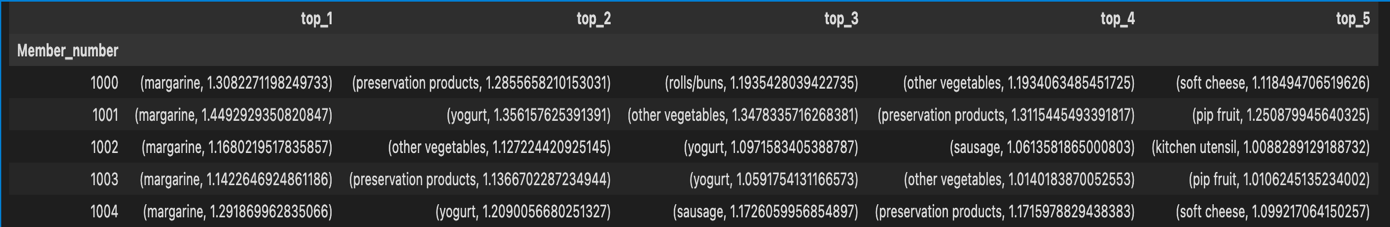
Second, we used UV decomposition template from Wrokshop12 to process data. The basic idea of UV decomposition is to decompose a large user-item interaction matrix into two lower-dimensional matrices, here referred to as U (user feature matrix) and V (item feature matrix).



Third, after UV decomposition we can obtain output data, each entry is the “possibility” (it is not really the possibility, but same idea) for a particular member will like a particular item.



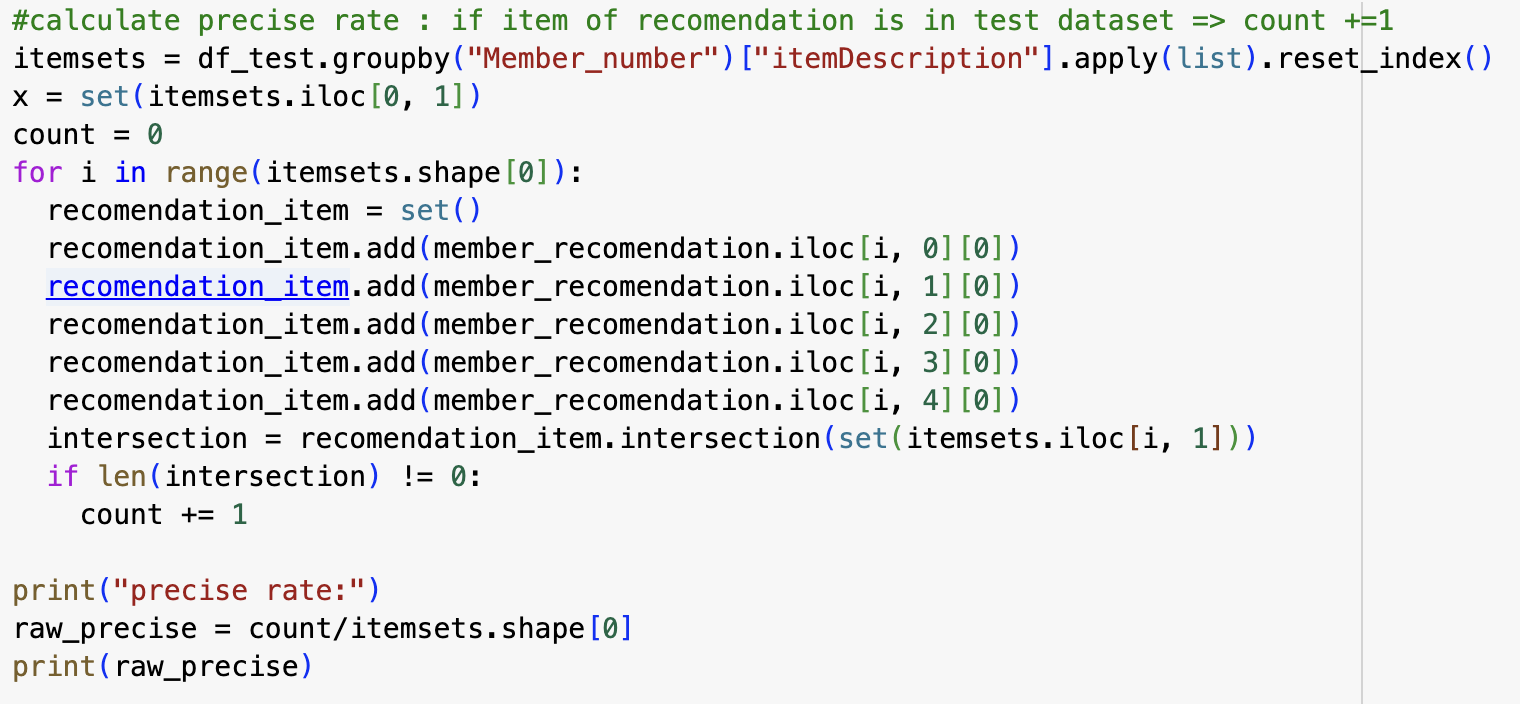
Finally, we choose top 5 biggest number in a row to make recommendation for each member.



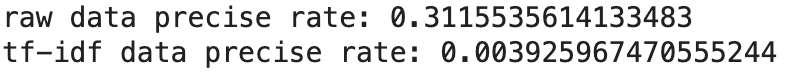
We also tried TF-IDF to process our data before we put them into UV decomposition. TF-IDF is used to reduce the importance of an item that everyone needs it; therefore, we would not obtain same item in most. However, we found the result is worse than non-TF-IDF data. Thus, we used non-TF-IDF as final result and keep TF-IDF as comparison.

**Metrics used for evaluation on test dataset:**

We make each recommendation of member a set. If the any item in the set is bought by that member in testing data. We said it “successful recommendation”. We count how many members we were successfully predicted, then divided by how many members in total; we can obtain precise rate.



Below is the metrics result of our recommendation:



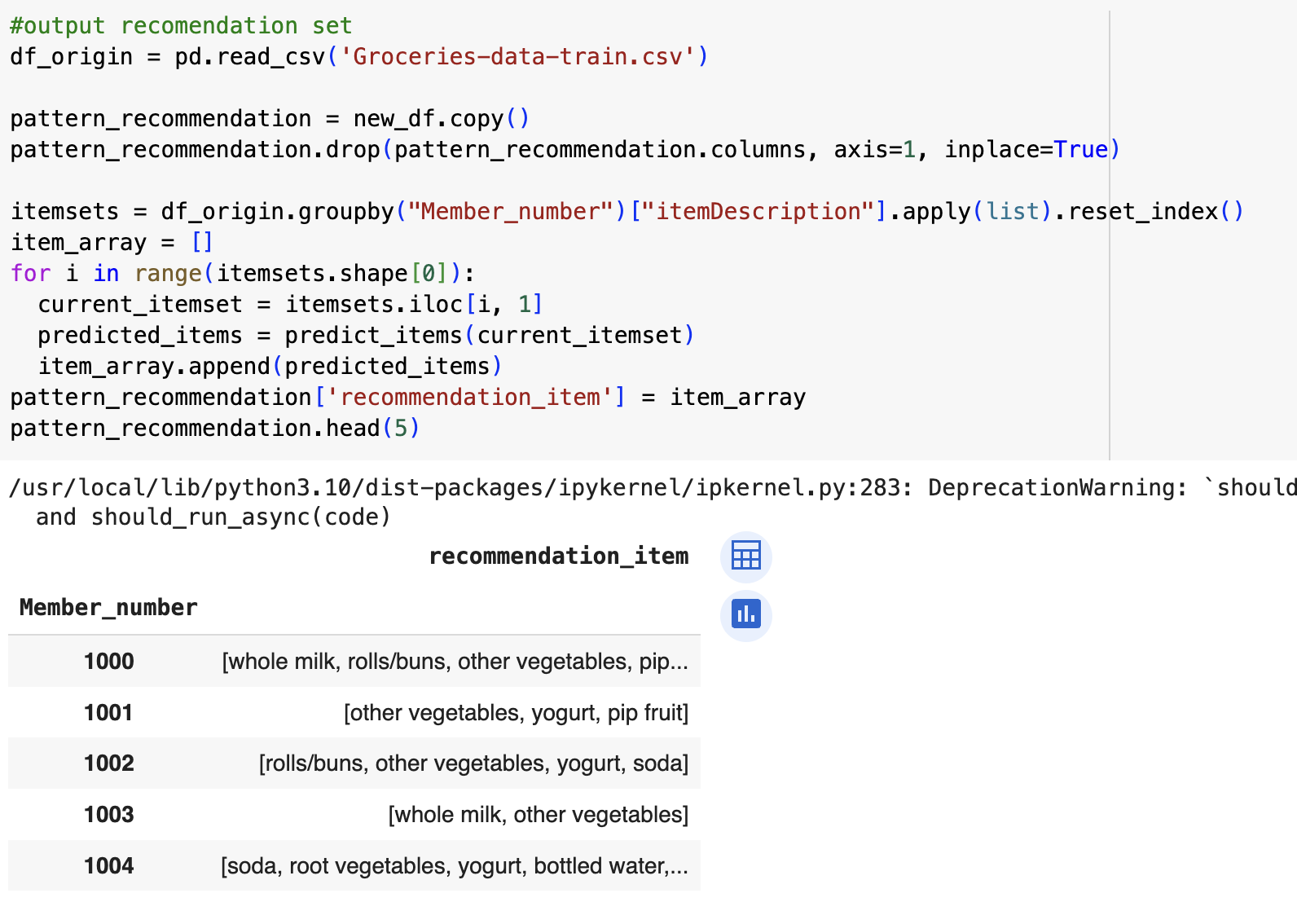
7. Recommendation method from frequent patterns.

**Brief description of the method:**

Frequent pattern is a pattern show frequent sets in particular data. We can get useful information from this pattern. For example, in grocery, what sort of items customer tend to buy them together.

**How this method is planned to work with patters/association rules as an input, and what will the output of this method:**

We used pattern as a template. If we imagine the task is to make recommendation from the items in a cart. The item in a cart is the items each member had bought in training data. We need to use these items set to predict/recommend items for each member. Therefore, we input sets from each member into Frequent pattern. The pattern will output another set, this set is the highest possibility a member would buy. We used this set as recommendation for every member. The result is below:



Then, we measure the performance of this method using precise rate above. The precise rate of recommendation from frequent pattern is below.

