

Second Homework Assignment

Dillard's is a major retail chain with several stores. Their point-of-sales (POS) data over a period of time is available in the database in the schema called pos. There are 5 tables in the schema.

The data dictionary is provided in a separate file.

The retailer is interested in rearranging the floors of the stores. In retail parlance, to change the planograms. For budgetary reasons (manpower) they can only make at most 20 moves across the entire chain. (A move consists of moving one SKU to a different position. This is the first time the company is performing this analysis and thus it is highly unlikely that SKUs are already appropriately close to each other.) The assignment is to find the 100 SKUs that are best candidates to modify the planograms.

Suggested process to undertake:

1. Understand the data (features of the data)
2. Perform data exploration (number of SKUs, number of items per basket, number of stores, etc)
3. Solve the problem – there is a lot of data so you might want to carefully select only a subset

Note that the data set is relatively large (not “big data”). If you will have problems querying the entire data set, be creative (in a clever way select a subset of data – based on the knowledge of the problem).