# What is the most popular item?

For retailers, product knowledge can mean more sales. The more you know your customer, the easier it is to identify opportunities to sell them products and target them with appropriate offers. For online stores, effectively understanding your customers relied on analyzing customer reviews and transaction records or selling patterns. To run a successful online business, it is essential for E-commerce retailers to know their favorite items.

The MapReduce process analyzed the Women's Clothing E-Commerce dataset. This dataset contains 23,486 online reviews from various retail store with 10 feature variables. Each row corresponds to a customer review, and to determine the favorable products, 3 relative variables are analyzed:
- **Transaction ID:** Unique transaction ID
- **Clothing ID:** Integer Categorical variable that refers to the specific piece purchased
- **Rating:** Positive Ordinal Integer variable for the product score granted by the customer from 1 Worst, to 5 Best.
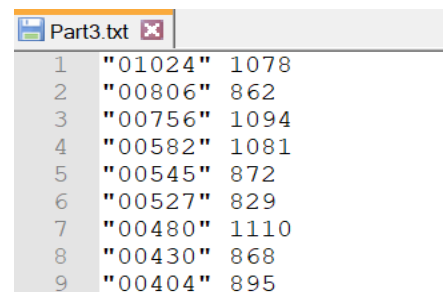
In the MapReduce procedure, we are assigned to calculate the selling amount of each product. The Clothing ID, which represents each product, was named "ItemName," treated as an integer variable. To know the favorite product, we need to count the occurrence for a specific product from the 230,000 transactions records.

The MapReduce Process was deconstructed as follow:
1. Find out ItemName in all transaction records
2. Count ItemName from all transaction records
3. Sort the result descending according to each products' selling records.
4. Generate a txt file that has a "Counts, ItemName" structure.



```
Part3.txt
1    "01024"  1078
2    "00806"  862
3    "00756"  1094
4    "00582"  1081
5    "00545"  872
6    "00527"  829
7    "00480"  1110
8    "00430"  868
9    "00404"  895
```

After the MapReduce process, we were able to find the Top 3 products which were Clothing ID: 1078, 862, 1094; Combined with other variables in the dataset, we were able to know the best selling products were dresses, knits, and dresses.

For women clothing retailers, catching the fashion design of the dresses of 1078 and 1094 are critical because it may stand for the fashion trend and knowing the materials and fabrics of 862 would help them identify their customer's requirements on softener and warmness of similarly designed knits. As a result, through analyzing which one is the favorite product, retailers can better identify the market demand and provide the right product to their customers.