

NoSQL project

1. Scrape all product data from TIKI website and store in a database in MongoDB

2. Make a backup file of the database

3. Import these data field into MySQL for further use

Product name

CategoryID

Category

Sub Category

Short description

Desciption

Filter out the extra html tags in the description

URL

Rating

Sales number

Price

4. Visualize the data

How number of products in each category

Compare the origin rate of products

Top 10 most sold, highest rating, and lowest price products

5. Get all the ingredient data and save it to a CSV file as format: product_id, ingredient