

발표대본

이후 발표를 맡은 김서기입니다. 다음으로는 쿼리에 대해서 발표하겠습니다.

I'm Kim Seogi, the presenter afterwards. Next, I will present the query.

이 쿼리는 당신이 셔츠, 블라우스 등으로 구성되어 있는 상의 종류 별로 색상이 어떤 것이 있는지 확인하길 원할때 사용하는것 입니다.사진에 서처럼 민트 셔츠는 157개, 퍼플색 셔츠는 91개 등 결과를 확인할 수 있었습니다.

This query is used when you want to see what colors are available for each type of shirt, blouse, etc. As shown in the picture, we could check the results of 157 mint shirts and 91 purple shirts.

이 쿼리는 당신이 상의 종류 별로 각각의 옷에서 소재는 어떤 것이 있는지 확인하길 원할때 사용하는것 입니다. unwind를 사용해 상의 배열을 분리 후,match를 사용해 빈 배열 제거를 하고 , unwind를 사용해 소재 배열을 분리 해주었습니다. 이후 상의의 소재로 group을 나누어 준 뒤, 갯수가 많은 순서대로 정렬을 해주었습니다. 마지막으로 한눈에 보기 쉽게 project로 변수를 정리하였습니다. 사진에서처럼 우븐 소재는 26000개, 저지 소재는 20000개 ...로 이루어져 있습니다.

This query is for each type of top, the material in each piece of clothing is used when you want to check what material is there.

After removing the top arrangement using unwind, we removed the empty arrangement using match, and separated the material arrangement using unwind.

After distributing the group with the material of the top, they arranged it in the order of the number of items. Lastly, to make it easy to see

I organized the variables with the project. As shown in the

picture, the woven material consists of 26000 pieces and the jersey material consists of 20000 pieces...

이 쿼리는 당신이 상의와 하의의 최고의 조합을 찾고싶을때 사용하는 것입니다.

상의 하의 카테고리를 기준으로 group을 나누어 주고, 각각의 개수를 sum변수에 저장했습니다.

이후 갯수가 많은 순서대로 를 기준으로 오름차순으로 정렬 후 한눈에 보기 쉽게 project로 변수를 정리하였습니다.

This query is used when you want to find the combination of top and bottom. The groups were divided based on the top and bottom categories, and each number was stored in the sum variable. After that, the variables were sorted in ascending order based on the order of high number and then organized into projects for easy viewing.

이 쿼리는 아우터별로 기장이 어떻게 되어 있는지 확인하기 위한 query입니다.

unwind를 사용해 아우터 배열을 분리 후, match를 사용해 아우터 기장의 빈 배열 제거를 했습니다.그 후 아우터의 기장을 기준으로 그룹을 나누어 준 뒤, 갯수가 많은 순서대로 정렬을 해주었습니다. 하프는 6957, 노말은 6455 ...로 이루어져 있습니다.

This query is to check the length of each outer. After removing the outer arrangement using unwind, we removed the empty arrangement of the outer length using match. After that, the groups were divided based on the length of the outerwear, and the order was arranged in the order of the number. The harp consists of 7000 and the normal 6500...

이 쿼리는 하의중에서 핏이 스키니인 하의를 찾을때 사용하는 쿼리입니다.

This query is used to find bottoms that are fit skinny in the bottoms.

이 쿼리는 상의중에서 핏이 루즈인 상의를 찾을 때 사용하는 쿼리입니다.

This query is used to find a top with a loose fit in the top.

다음은 최적화에 대해서 설명하겠습니다.

첫번째는 replication 입니다.

primary를 27017 포트로, secondary를 2개를 하여 27020, 27021 포트로 사용하였습니다. 데이터가 올바르게 분산되어 들어갔음을 알 수 있습니다.

The first is replication.

I used the primary as 27017 port and the secondary as 27020 and 27021 port. You can see that the data has been distributed correctly.

다음은 인덱싱 입니다

싱글 인덱스와 컴파운드 인덱싱을 사용했습니다. 그 결과 다음과 같이 실행시간이 감소했습니다.

Single index and compound indexing were used. As a result, the execution time has decreased as follows

시각화 하는 쿼리도 감소하도록 인덱싱을 하였습니다.

We also indexed the visualization queries to reduce time

