

SQL final project

추천 시스템 1

1. 추천 시스템 이름 : 높은 할인율 & 높은 평점 에 기반한 추천

2. 추천 시스템의 테마 :

고객은 합리적인 가격과 품질을 동시에 추구하기에,

단순히 가격이 저렴한 상품이 아니라, **할인율이 크면서 평점도 일정 수준(4.0 이상)을 만족하는 상품을** 추

3. 구현 로직

```
SELECT
  product_id,
  product_name,
  actual_price,
  discounted_price,
  discount_percentage,
  SAFE_CAST(rating AS FLOAT64) AS rating
FROM `verdant-bond-470202-k2.final_project.amazon`
WHERE SAFE_CAST(rating AS FLOAT64) >= 4.0
ORDER BY discount_percentage DESC, rating DESC
LIMIT 50;
```

4. 결과 해석

| 작업 정보 | | 결과 | 시각화 | JSON | 실행 세부정보 | 실행 그래프 |
|-------|------------|---|--------------|------------------|---------------------|--------|
| 행 | product_id | product_name | actual_price | discounted_price | discount_percent... | rating |
| 1 | B097C564GC | rtss [2 Pack] Mini USB C Type C Adapter Plug, Type C Female to USB A Male Charger Charging | 4999.0 | 294.0 | 0.94 | 4.3 |
| 2 | B0BF54972T | Fire-Boltt Ninja Call Pro Plus 1.83" Smart Watch with Bluetooth Calling, AI Voice Assistance, 100 | 19999.0 | 1799.0 | 0.91 | 4.2 |
| 3 | B0BF4YBLPX | Fire-Boltt Ninja Call Pro Plus 1.83" Smart | 19999.0 | 1799.0 | 0.91 | 4.2 |

- 평점 4.0 이상인 상품 + 할인율 높은 상품을 할인율에 대한 내림차순으로 정렬
- 해당 조건을 만족하는 상위 50개 물건을 추천한다.

추천 시스템 2

1. 추천 시스템 이름 : 카테고리별 베스트셀러 추천 시스템

2. 추천 시스템의 테마 :

고객이 특정 카테고리 안에서 신뢰할 수 있는 상품을 찾으려 하기 위해, 평점과 리뷰 수를 결합한 점수를 기준으로 추천합니다. 단순히 평점만 높거나 리뷰 수만 많은 상품 대신, **평점 × log(리뷰 수)** 지표를 활용하여 두 요소를 동시에 반영합니다. 이를 통해 리뷰가 적지만 평점이 높은 상품의 과대평가 문제, 리뷰 수만 많은 상품의 저평가 문제를 보완할 수 있습니다.

3. 구현 로직

- **category** 컬럼을 기준으로 분리하여 세부 카테고리 단위 분석
- 평점을 숫자로 변환 (**SAFE_CAST**) 후, 평점이 NULL이 아닌 상품만 고려
- 가중 점수 = 평점 × LOG(1 + 리뷰 수) 계산
- 각 카테고리별로 점수가 높은 상위 5개 상품을 추천

```
SELECT
  category_value AS category,
```

```

product_id,
product_name,
SAFE_CAST(rating AS FLOAT64) AS rating,
rating_count,
ROUND(SAFE_CAST(rating AS FLOAT64) * LOG(1 + rating_count), 2) AS weighted_score
FROM `verdant-bond-470202-k2.final_project.amazon`,
    UNNEST(SPLIT(category, '|')) AS category_value
WHERE SAFE_CAST(rating AS FLOAT64) IS NOT NULL
QUALIFY ROW_NUMBER() OVER (
    PARTITION BY category_value
    ORDER BY weighted_score DESC
) <= 5;

```

4. 결과

| 행 | category | product_id | product_name | rating | rating_count | weighted_score |
|---|-----------|-------------|--|--------|--------------|----------------|
| 1 | Laundr... | B01LY9W8AF | Cello Eliza Plastic Laundry Bag/Basket, 50 litres, Light Grey | 4.2 | 7274 | 37.35 |
| 2 | Laundr... | B095XCRD... | Esquire Laundry ... | 4.3 | 5911 | 37.34 |
| 3 | Laundr... | B083J64CBB | Kuber Industries Waterproof Canvas Laundry Bag/Hamper M | 4.0 | 10234 | 36.93 |
| 4 | Laundr... | B0814P4L98 | PrettyKrafts | 4.0 | 5380 | 34.36 |

- 위 쿼리를 실행하면 **카테고리별 상위 5개 베스트셀러**가 추출됩니다.
- 추천 상품은 단순히 평점이 높은 것만이 아니라, 리뷰 수까지 함께 고려했습니다.
- 고객이 자신이 원하는 카테고리 내에서 가장 신뢰할 수 있는 베스트셀러를 추천 받게 됩니다.

추천 시스템 3

1. 추천 시스템 이름 : 카테고리별 사용자들의 찐추천 꿀템_리뷰기반

2. 추천 시스템의 테마 : 정량적 지표(평점, 리뷰수) + 정성적 지표(리뷰 텍스트 분석) 을 종합적으로 고려하여 진짜 사용한 사람들이 진심으로 추천한 물건들을 선별하고자 했다. 구체적으로는 평점이 일정 기준 이상이고 리뷰 수가 충분히 많은 상품을 먼저 필터링한 뒤, 리뷰 내용에서 “excellent”와 같은 긍정 단어가 많이 언급된 상품을 카테고리별로 추천한다.

3. 구현 로직

- `rating` 을 숫자로 변환하고, `rating >= 4.0` 및 `rating_count >= 50` 이상인 상품만 1차 필터링
- `review_content` 를 단어 단위로 분리
- 긍정 단어 사전과 매칭하여 긍정 단어 수를 상품별로 집계
- 가중 점수 = 평점 × LOG(1 + 리뷰 수) + 긍정 단어 카운트
- 카테고리별로 점수가 높은 상위 5개 상품을 추천

```

WITH base AS (
SELECT
    product_id,
    product_name,
    rating_count,
    SAFE_CAST(rating AS FLOAT64) AS rating,
    category,
    review_content
FROM `verdant-bond-470202-k2.final_project.amazon`
WHERE SAFE_CAST(rating AS FLOAT64) >= 4.0

```

```

    AND rating_count >= 50
),

-- 리뷰 내용을 단어 단위로 분리
review_words AS (
    SELECT
        product_id,
        product_name,
        rating,
        rating_count,
        category_value AS category,
        word
    FROM base,
        UNNEST(SPLIT(category, '|')) AS category_value,
        UNNEST(SPLIT(review_content, ' ')) AS word
),

-- 긍정 단어 사전 정의
positive_words AS (
    SELECT word FROM UNNEST([
        '좋다','추천','만족','최고','굿','excellent','perfect'
    ]) AS word
),

-- 상품별 긍정 단어 카운트
positive_counts AS (
    SELECT
        r.product_id,
        r.product_name,
        r.category,
        r.rating,
        r.rating_count,
        COUNT(*) AS positive_word_count
    FROM review_words r
    JOIN positive_words p
    ON r.word LIKE p.word
    GROUP BY r.product_id, r.product_name, r.category, r.rating, r.rating_count
)

-- 카테고리별 추천 상품 추출
SELECT
    category,
    product_id,
    product_name,
    rating,
    rating_count,
    positive_word_count,
    ROUND(rating * LOG(1 + rating_count) + positive_word_count, 2) AS weighted_score
FROM positive_counts
QUALIFY ROW_NUMBER() OVER (PARTITION BY category ORDER BY weighted_score DESC) <= 5;

```

4. 결과

| 행 | category | product_id | product_name | rating | rating_count | positive_word_count | weighted_score |
|---|------------|------------|--|--------|--------------|---------------------|----------------|
| 1 | GamingMice | B08LT9B... | Logitech G102 USB Light Sync Gaming Mouse with Customizable | 4.5 | 10541 | 1 | 42.68 |
| 2 | GamingMice | B08CHZ... | Redgear A-15 Wired Gaming Mouse with Upto 6400 DPI, RGB & Driver | 4.3 | 15790 | 1 | 42.57 |
| 3 | GamingMice | B08D64... | Ant Esports Gaming Mouse | 4.2 | 8537 | 1 | 39.02 |

- 카테고리별로 평점이 높고 리뷰 수가 충분하며, 리뷰 내용이 실제로 긍정적인 상품 상위 5개가 추출됩니다.

예시:

- *Laptops* → 평점 4.6, 리뷰 수 15,000, 긍정 단어 20,000회 이상 → 최상위
- *Headphones* → 평점 4.8, 리뷰 수 200, 긍정 단어 1,000회 이상 → 상위권

추천 시스템 4

1. 추천 시스템 이름 : 비슷한 상품 추천해주기

2. 추천 시스템의 테마 :

고객이 특정 상품을 조회했을 때, 동일 카테고리에서 평점이 높은 유사 상품을 함께 추천합니다. 이 방식은 고객이 원래 보고 있던 제품과 **비슷하지만 조금 다른 선택지**를 제시하여, 구매 전환율과 교차판매(Cross-selling)를 높이는 데 효과적입니다.

3. 구현 로직

- 기존 상품과 같은 카테고리에 속한 다른 상품을 탐색
- 기존 상품은 제외하고, 동일 카테고리 내 평점이 높은 상품을 추천
- 카테고리별 상위 5개를 추천

```
SELECT
  a.product_id,
  a.product_name,
  b.product_id AS similar_product_id,
  b.product_name AS similar_product_name,
  SAFE_CAST(b.rating AS FLOAT64) AS rating,
  b.rating_count
FROM `verdant-bond-470202-k2.final_project.amazon` a
JOIN `verdant-bond-470202-k2.final_project.amazon` b
ON a.category = b.category AND a.product_id != b.product_id
WHERE SAFE_CAST(b.rating AS FLOAT64) >= 4.0
QUALIFY ROW_NUMBER() OVER (PARTITION BY a.product_id ORDER BY b.rating DESC, b.rating_count DESC) <= 5;
```

4. 결과

| 행 | product_id | product_name | similar_product_id | similar_product_name | rating | rating_count |
|---|------------|---|--------------------|---|--------|--------------|
| 1 | B01LY9W8AF | Cello Eliza Plastic Laundry Bag/Basket, 50 litres, Light Grey | B095XCRDQW | Esquire Laundry Bas... | 4.3 | 5911 |
| 2 | B01LY9W8AF | Cello Eliza Plastic Laundry Bag/Basket, 50 litres, Light Grey | B08VGFX2B6 | Kuber Industries Waterproof Round Non Wovan Laundry Bag/Hamper Metali | 4.1 | 3688 |
| 3 | B01LY9W8AF | Cello Eliza Plastic Laundry Bag/Basket, | B08VGM3YMF | Heart Home Waterproof Round | 4.1 | 1996 |

- 특정 상품 페이지를 본 경우, 같은 카테고리의 **평점이 높고 리뷰 수가 많은 다른 상품** 5개가 추천됩니다.
- 예: *Laptop A*를 조회하면, 같은 카테고리 내 평점 4.5 이상, 리뷰 5,000건 이상인 **Laptop B, C, D...**가 추천됨
- 여러 대안을 쉽게 비교할 수 있어 선택의 폭이 넓어집니다.

추천 시스템 5

1. 추천 시스템 이름 : **충성고객 맞춤추천_리뷰 자주 남긴 고객을 위한 추**

2. 추천 시스템의 테마 :

리뷰를 많이 남기는 고객은 충성도가 높고, 플랫폼에 대한 참여도가 큼니다. 이 고객군을 식별한 뒤, 그들에게 **특별한 맞춤 상품**을 추천하면, 고객 경험을 향상시키고 장기적 충성도를 강화할 수 있습니다.

3. 구현 로직

- `user_id`와 `review_id`를 분리하여 고객별 리뷰 개수를 집계
- 리뷰 개수가 일정 기준(예: 10개 이상) 이상인 고객을 충성 고객으로 정의
- 이 고객에게 최근 평점이 높은 인기 상품을 우선 추천

```
WITH user_reviews AS (
  SELECT
    user_id,
    COUNT(review_id) AS review_count
  FROM `verdant-bond-470202-k2.final_project.amazon`,
    UNNEST(SPLIT(user_id, ',')) AS user_id,
    UNNEST(SPLIT(review_id, ',')) AS review_id
  GROUP BY user_id
)
SELECT
  ur.user_id,
  p.product_id,
  p.product_name,
  SAFE_CAST(p.rating AS FLOAT64) AS rating,
  p.rating_count
FROM user_reviews ur
JOIN `verdant-bond-470202-k2.final_project.amazon` p
ON ur.review_count >= 10
WHERE SAFE_CAST(p.rating AS FLOAT64) >= 4.5
ORDER BY ur.review_count DESC, p.rating DESC
LIMIT 100;
```

4. 결과

| 행 | user_id | product_id | product_name | rating | rating_count |
|---|------------|-------------|--|--------|--------------|
| 1 | AE55KTF... | B0BQRJ3... | REDTECH USB-C to Lightning Cable 3.3FT, [Apple MFi Certified] Lightning to Type C Fast Charging Cord Compatible with iPhone | 5.0 | null |
| 2 | AE55KTF... | B0BP7XLX... | Syncwire LTG to USB Cable for Fast Charging Compatible with Phone 5/ 5C/ 5S/ 6/ 6S/ 7/8/ X/XR/XS Max/ 11/12/ 13 Series and Pad Air/Mini, Pod & | 5.0 | 5 |
| 3 | AG5DWP... | B0BQRJ3... | REDTECH USB-C to Lightning Cable 3.3FT, [Apple MFi | 5.0 | null |

- 리뷰를 10개 이상 남긴 고객을 대상으로, **평점 4.5 이상인 인기 상품**이 추천됩니다.

- 예:
 - 리뷰를 50개 남긴 *user_123*에게 최신 평점 4.8 상품이 노출
 - VIP 고객은 “나를 위한 맞춤 추천”이라는 경험을 얻게 되고, 플랫폼은 충성 고객의 만족도를 높일 수 있음