

250715 [프로젝트] 추천 시스템 설계하기

1. 추천 시스템 이름

→ "리뷰를 많이 남긴 상품들이에요" / "최대 할인율 기반 추천 시스템"

2. 추천 시스템의 테마: 추천 시스템의 고유 컨셉에 대한 설명

→ 고객이 최대한의 할인율을 적용받을 수 있도록 할인율과 평점을 결합한 추천을 제공합니다.

3. 구현 로직: SQL 쿼리 설명 및 주요 로직 설명

→ SQL 코드 첨부

4. 결과

: 추천 결과의 샘플을 출력하고, 이에 대해서 간단하게 설명해주세요.

→ 빅쿼리 화면을 캡처하여 이미지를 제출하고, 이에 대해서 간단히 설명해주세요.

1.1 이름: 가성비 갑! 합리적 추천 시스템

1.2 테마:

최소 비용으로 최대 검증 신뢰를 확보할 수 있도록, "평점 x 리뷰 수"를 할인가로 나눈 가성비 지수를 기반으로 상품을 추천합니다. 이 시스템은 ****"지불하는 1원당 얻는 신뢰"*****를 수치화해 상위 상품을 고릅니다. 고객은 단순히 싸다는 이유만으로는 아니라, **높은 평점과 풍부한 리뷰로 검증된 제품을 최저 비용**으로 확보할 수 있어, 가성비에 대한 불안 없이 안심 쇼핑이 가능합니다.

- "나는 똑똑한 소비를 하고 싶다!"
- "가성비를 최대한으로 챙기되, 이미 수많은 사용자가 인정한 제품만 사고 싶다"
- → 가격 대비 신뢰의 극대화

1.3 구현 로직

```
WITH clean AS (  
  SELECT  
    product_name,  
    category,  
    about_product,  
    -- rating이 문자열일 때만 숫자로 변환  
    SAFE_CAST(rating AS FLOAT64) AS rating_num,  
    -- rating_count가 이미 INT64라면 그대로 사용  
    rating_count,  
    -- discounted_price가 FLOAT64라면 그대로 사용  
    discounted_price  
  FROM  
    `calcium-adapter-465502-n6.day4.amazon2`  
)
```

```

WHERE
  SAFE_CAST(rating AS FLOAT64) IS NOT NULL
),

value_score AS (
  SELECT
    product_name,
    category,
    about_product,
    rating_num,
    rating_count,
    discounted_price,
    -- 1원당 얻는 "검증 신뢰도"
    (rating_num * rating_count) / discounted_price AS value_index
  FROM
    clean
)

SELECT
  product_name,
  category,
  about_product,
  ROUND(value_index, 2) AS value_index,
  rating_num,
  rating_count,
  discounted_price
FROM
  value_score
ORDER BY
  value_index DESC
LIMIT 5;

```

1.3.1.

clean CTE: 원본 데이터 전처리

- **SAFE_CAST(rating AS FLOAT64)**
 - 문자열 타입의 **rating** 컬럼을 **FLOAT64** 로 변환.
 - 변환 실패 행(예: "N/A")은 **NULL** 이 되어 **WHERE** 절에서 제외.
- **rating_count** & **discounted_price**
 - 두 컬럼이 이미 숫자형(INT64/FLOAT64)이므로 별도 전처리 없이 그대로 활용.
- 필터
 - 평점이 숫자로 변환된 행만 남겨, 이후 계산 안정성 확보.

1.3.2. **value_score** CTE: 가성비 지수 계산

- 공식: $\text{value_index} = \text{rating_num} \times \text{rating_count} / \text{discounted_price}$

$$\text{value_index} = \frac{\text{rating_num} \times \text{rating_count}}{\text{price}}$$

- 분자: 평점 × 리뷰 수 → '검증된 신뢰도 총합'
- 분모: 할인가 → '소비 비용'
- → "1원당 얻는 리뷰 기반 신뢰도"를 수치화

1.3.3. 최종 SELECT : 상위 상품 추출 및 포매팅

- ROUND(...,2) 으로 지수 가독성 향상
- ORDER BY value_index DESC
 - 가성비가 가장 뛰어난 상품부터 나열
- LIMIT 5
 - 고객에게 제안할 톱5 리스트 완성

요약

- 목표: "최소 비용으로 최대 검증 신뢰를 확보"할 상품만 뽑는다.
- 전처리: 문자열 평점만 숫자로 바꾸고, 나머지 숫자 컬럼은 그대로 사용.
- 핵심 로직: (평점×리뷰수)÷가격(평점×리뷰수)÷가격 공식으로 가성비 지수 산출.
- 출력: 지수 상위 5개를 소수점 둘째 자리로 깔끔하게 보여줌.

이렇게 설계된 로직은, 고객이 "돈은 적게 쓰되 이미 많은 사람에게 검증된 제품"을 빠르고 명확하게 찾도록 돕습니다.

1.4 결과

행	product_name	category	about_product	value_index	rating_num	rating_count	discounted_price
1	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and moreCompatible with Blu-Ray players, computers, Apple TV, Roku, cable, PS4, Xbox One, and other HDMI-compatible devicesSolid copper conductors and full metal jacket shielding for durability and high-performance connectivitySupports Ethernet, 3D, 4K video and Audio Return	8578.45	4.4	426973	219.0
2	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and moreCompatible with Blu-Ray players, computers, Apple TV, Roku, cable, PS4, Xbox One, and other HDMI-compatible devicesSolid copper conductors and full metal jacket shielding for durability and high-performance connectivitySupports Ethernet, 3D, 4K video and Audio Return	8578.43	4.4	426972	219.0
3	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video,Black	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel (ARC)Connects Blu-ray players, Fire TV, Apple TV, PS4, PS3, Xbox one, Xbox 360, computers and other HDMI-enabled devices to TVs, displays, A/V receivers and moreCable allows you to share an Internet connection among multiple devices without the need for a separate Ethernet cableMeets HDMI 2.0	6079.87	4.4	426973	309.0
4	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video,Black	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Please select appropriate display resolution & refresh rate for proper display output. HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel (ARC)Connects Blu-ray players, Fire TV, Apple TV, PS4, PS3, Xbox one, Xbox 360, computers and other HDMI-enabled devices to TVs, displays, A/V	6079.87	4.4	426973	309.0
5	Pigeon Polypropylene Mini Handy and Compact Chopper with 3 Blades for Effortlessly Chopping Vegetables and Fruits for Your Kitchen (12420,	Home&Kitchen\Kitchen&Dining\KitchenTools\ManualChoppers&Choppers\Choppers	Unique string function to chop vegetables and fruits with easePackage Contents: 1-Piece Handy Chopper.Blade Length: 3 centimetersEco-friendly design, no electricity requiredWarranty: 30 Days Warranty, Only On Manufacturing DefectsIMade from unbreakable ABS plastic for long-lasting use, Material: Polypropylene (PP)Sturdy 3-blade design made from Stainless Steel, As this	5574.41	4.1	270563	199.0

순위	product_name	value_index	rating_num	rating_count	discounted_price
1	AmazonBasics Flexible Premium	8578.45	4.4	426,973	219.0

순위	product_name	value_index	rating_num	rating_count	discounted_price
	HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot				
2	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	8578.43	4.4	426,972	219.0
3	Amazon Basics High-Speed HDMI Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel	6079.87	4.4	426,973	309.0
4	Amazon Basics High-Speed HDMI Cable, 6 Feet – Supports Ethernet, 3D, 4K video, Black	6079.87	4.4	426,973	309.0
5	Pigeon Polypropylene Mini Handy and Compact Chopper with 3 Blades	5574.41	4.1	270,563	199.0

• 극단적 리뷰 파워 + 저렴한 가격

상위 4개(HDMI 케이블) 모두 **42만 건 이상의 리뷰**를 바탕으로 평점 4.4를 기록하면서, 200~300원 대의 매우 낮은 가격에 판매되고 있습니다.

$(4.4 \times 426,973) \div 219 \approx 8,578$ 라는 수치는 "1원당 약 19건의 리뷰"를 확보했다는 의미로, **최소 비용** 으로 **최대 검증 신뢰**를 제공합니다.

• 의외의 히트 가성비 아이템

5위 Pigeon 소형 채소·과일 절단기는 **27만 건**의 리뷰와 평점 4.1을 바탕으로, 199원이라는 낮은 가격에 $270,563 \div 199 \approx 5,574$ 의 가성비 지수를 기록했습니다.

"전자·음향"이 아닌 "주방용품" 카테고리에서도 **압도적 가성비** 상품이 발견된 점이 흥미롭습니다.

• 고객에게 주는 가치

이 추천 시스템은 단순히 "할인율"이나 "평점" 중 하나만 보는 것이 아니라, **평점 × 리뷰 수 ÷ 가격** 공식을 통해, **"얼마나 많은 사람에게 검증받았는지를 가격 대비"** 정확히 측정합니다.

결과적으로, 고객은 ****싸지만 품질 불안****이 아니라

'값도 싸고, 많은 사용자에게 이미 검증된' 상품을 한눈에 확인할 수 있어, **쇼핑 결정에 대한 불확실성을 완벽히 제거**할 수 있습니다.

2.1 이름: 최대 할인율 기반 "리뷰를 많이 남긴 상품들"

2.2 테마:

카테고리 평균 할인율 대비 이례적인 할인 폭(Z-score) 과 평점 안정성(5점 만점 대비 비율) 을 결합해, 고객이 진짜 놓치면 손해인 딜만 골라드립니다. 고객은 "단순히 싸 보이는 딜"이 아니라, 할인 폭이 통계적으로도 이례적인 상품을 골라낼 수 있어, 최대 할인 + 검증된 품질 두 마리 토끼를 동시에 잡을 수 있습니다.

- "지금이 아니면 못사는, 놓치면 손해인 찐특가"
- "단순히 싸다고 좋은 게 아니라, 원래 안 싸던 상품이 평점까지 좋은 상황!"
- → 세일 타이밍과 검증 품질의 결합

2.3 구현 로직

```
WITH clean AS (
  SELECT
    product_name,
    category,
    about_product,
    SAFE_CAST(rating AS FLOAT64) AS rating_num,
    rating_count AS rating_count,
    discount_percentage AS discount_pct
  FROM
    `calcium-adapter-465502-n6.day4.amazon2`
  WHERE
    SAFE_CAST(rating AS FLOAT64) IS NOT NULL
),

-- 1) 카테고리별 평균·표준편차 계산, std_dp = 0인 카테고리는 제외
price_stats AS (
  SELECT
    category,
    AVG(discount_pct) AS avg_dp,
    STDDEV(discount_pct) AS std_dp
  FROM clean
  GROUP BY category
  HAVING STDDEV(discount_pct) <> 0
),

deal_score AS (
  SELECT
    c.*,

    -- 2) SAFE_DIVIDE 또는 NULLIF 로 0 분모 방지
    (
      SAFE_DIVIDE(c.discount_pct - ps.avg_dp, ps.std_dp) * 0.6
      + (c.rating_num / 5.0) * 0.4
    )
  FROM clean c
  JOIN price_stats ps ON c.category = ps.category
)
```

```

) AS deal_index

FROM clean c
JOIN price_stats ps USING(category)
)

SELECT
product_name,
category,
about_product,
ROUND(deal_index, 2) AS deal_index,
rating_num,
rating_count,
discount_pct
FROM deal_score
ORDER BY deal_index DESC
LIMIT 5;

```

단계	설명
①	SAFE_CAST(rating AS FLOAT64) 문자열인 rating 을 부드럽게 숫자형으로 변환합니다.
②	rating_count 이미 정수(INT64)이므로 별도 변환 없이 리뷰 수를 사용합니다.
③	discount_percentage 이미 실수(FLOAT64)이니 바로 '할인율'로 사용합니다.
④	WHERE ... IS NOT NULL rating 이 숫자로 변환되지 않는(잘못된) 행을 제거해 안정성 보장.
⑤	AVG(discount_pct) AS avg_dp 카테고리별 평균 할인율을 계산합니다.
⑥	STDDEV(discount_pct) AS std_dp 카테고리별 할인율 표준편차(분산의 제곱근)를 계산합니다.
⑦	HAVING STDDEV(...) <> 0 할인율이 모두 동일해 분산이 0인 그룹은 추천 대상에서 제외.
⑧	deal_index 계산 $(discount_pct - avg_dp) / std_dp$ 로 할인 이례성(Z-score)을 구한 뒤 0.6 비중을 곱하고, $(rating_num/5.0)$ 평점 비율에 0.4를 곱해 합산합니다.
⑨	ROUND(deal_index,2) 소수점 둘째 자리까지 반올림해 지수를 보기 좋게 포맷합니다.
⑩	ORDER BY deal_index DESC LIMIT 5 지수가 가장 높은 상위 5개 상품을 내림차순으로 추출합니다.

2.3.1. 데이터 정제(**clean**)

- 평점만 숫자로 변환 → 비정상 데이터 제거
- 리뷰 수·할인율은 이미 숫자형이므로 전처리 불필요

2.3.2. 카테고리별 통계(**price_stats**)

- 평균 할인율(**avg_dp**)·표준편차(**std_dp**) 계산
- 분산이 0인(할인율이 모두 동일한) 카테고리는 제외

2.3.3. 지수 계산(**deal_score**)

- 할인폭 이례성(Z-score) × 0.6
- 평점 비율(5점 만점 대비) × 0.4
- 두 값을 합산해 **deal_index** 생성

2.3.4. 최종 추출

- deal_index 내림차순 정렬 → 상위 5개 상품 추천

2.4 결과

행	product_name	category	about_product	deal_index	rating_num	rating_count	discount_pct
1	Samsung Galaxy S20 FE 5G (Cloud Navy, 8GB RAM, 128GB Storage) with No Cost EMI & Additional Exchange Offers	Electronics Mobles&Accessories Smartphones&Accessories Smartphones	5G Ready powered by Qualcomm Snapdragon 865 Octa-Core processor, 8GB RAM, 128GB internal memory expandable up to 1TB, Android 11.0 operating system and dual SIM Triple Rear Camera Setup- 12MP (Dual Pixel) OIS F1.8 Wide Rear Camera + BMP OIS Tele Camera + 10MP Ultra Wide 30X Space Zoom, Single Take & Night Mode 23MP F2.2 Front Punch Hole Camera 5.5 inch(14.43 centimeters) Infinity-O Super AMOLED Display with 120Hz Refresh rate, 1080 x 2400 (FHD+) Resolution 4500 mAh battery (Non-removable) with Super Fast Charging, Fast Wireless Charging & Finger Print sensor P68 Rated, MicroSD	2.11	4.2	27790	0.49
2	Goodscity Garment Steamer for Clothes, Steam Iron Press - Vertical & Horizontal Steaming up to 22g/min, 1200 Watt, 230 ml Water tank & 30 sec Fast Heating (GC 111)	Home&Kitchen Kitchen&HomeAppliances VacuumCleaning Ironing Irons, Steamers Accessories IroningStations	Lightweight & Portable: Dimensions: 223 x 158 x 126 mm. You can steam your clothes and fabrics vertically or horizontally without the weight of the water tank holding you back. This steamer is travel friendly and can easily fit it into your suitcase. Detachable Water Tank: The steamer comes with an easy-to-detach water tank with the capacity of 230 ml which allows it to run for at least 10 minutes One Button Function: You just need to hold on to one button while steaming and as soon as you let go, the vapour immediately stops which allows you to save on water. The blue light at the back indicates that the steamer is	1.79	4.2	611	0.62
3	Kitchen Kit Electric Kettle, 1.8L Stainless Steel Tea Kettle, Fast Boil Water Warmer with Auto Shut Off and Boil Dry Protection Tech	Home&Kitchen Kitchen&HomeAppliances SmallKitchenAppliances Kettles&HotWaterDispensers Kettle & Toaster Sets	.The tough pure metal stainless-steel build makes it very sturdy and the inside part is only made from premium food grade stainless steel that keep your water out of the contamination from plastic materials and chemicals	1.72	3.4	1066	0.76
4	TCL 80 cm (32 inches) HD Ready Certified Android Smart LED TV 32S5205 (Black)	Electronics HomeTheater, TV&Video Televisions SmartTelevisions	Resolution: HD Ready (1366 x 768) Refresh Rate: 60 hertz Connectivity: 2 HDMI Ports to connect set top box 1 USB ports to connect hard drives or other USB devices IR Port to control connected devices like DVD / STB Players Sound: 16 Watts Output Powerful Speakers with Dolby Audio Smart TV Features : Android R 11 Packed with amazing smart features such as Google Voice Search, Google App Store and Built-in Chromecast, Google Play Store, YouTube, Google cast, Netflix. Display : LED Smart TV Ultra bright screen for flawless picture quality even in bright rooms Warranty information: 2 year comprehensive warranty provided by	1.64	4.1	398	0.65
5	iFFALCON 80 cm (32 inches) HD Ready Smart LED TV 32F53 (Black)	Electronics HomeTheater, TV&Video Televisions SmartTelevisions	Resolution: HD Ready (1366 x 768) Refresh Rate: 60 hertz Connectivity: 2 HDMI Ports to connect set top box 1 USB ports to connect hard drives or other USB devices IR Port to control connected devices like DVD / STB Players Sound: 16 Watts Output Powerful Speakers with Dolby Audio Smart TV Features : Android R 11 1 GB RAM 8 GB ROM Packed with amazing smart features such as Google Voice Search, Google App Store and Built-in Chromecast, Google Play Store, YouTube, Google cast, Netflix. Display: A+ Grade Panel 178 Degree Wide Viewing Angle HDR Warranty information: 1 Year warranty provided by the	1.6	4.2	1269	0.64

순위	product_name	deal_index	rating_num	rating_count	discount_pct
1	Samsung Galaxy S20 FE 5G (Cloud Navy, 8GB RAM, 128GB Storage) with No Cost EMI & Exchange Offers	2.11	4.2	27,790	0.49
2	Goodscity Garment Steamer for Clothes, Steam Iron Press – Vertical & Horizontal (GC 111)	1.79	4.2	611	0.62
3	Kitchen Kit Electric Kettle, 1.8L Stainless Steel Tea Kettle, Fast Boil Water Warmer	1.72	3.4	1,066	0.76
4	TCL 80 cm (32 inches) HD Ready Certified Android Smart LED TV 32S5205 (Black)	1.64	4.1	398	0.65
5	iFFALCON 80 cm (32 inches) HD Ready Smart LED TV 32F53 (Black)	1.60	4.2	1,269	0.64

- **1위 Galaxy S20 FE 5G(2.11):** 할인율이 평균 대비 0.49%로 이례적이고, 2.7만 개가 넘는 리뷰로 검증된 안정적 딜입니다.
- **2~3위:** 소형 가전(스티머·전기 주전자) 카테고리에서도 “평균보다 큰 할인 + 검증된 평점”을 달성해 상위권에 올랐습니다.

- **4~5위 TV 제품:** 신형 스마트 TV 모델이 비교적 짧은 기간에 눈에 띄는 할인 혜택을 받으며, 리뷰·평점 안정성까지 갖춰진 딜입니다.

3.1 이름: 당신의 예산, 딱 맞춤 추천 시스템

3.2 테마

고객이 설정한 **최대 예산** 내에서, 평점과 리뷰 검증도를 결합해 **예산 대비 만족도**가 가장 높은 상품을 추천합니다. 예산을 초과할 걱정 없이, **최소 비용으로 최대 만족**을 보장하기 때문에 “돈이 부족해서 원하는 상품을 못 산다”는 고민을 완전히 해소해 줍니다. 예산 범위 내, 실패 없는 베스트 선택!

3.3 구현 로직

```
-- 1) 사용자 예산 설정 (예: 20,000원)
DECLARE max_budget FLOAT64 DEFAULT 20000;

-- 2) 예산 필터링 및 숫자형 컬럼 정제
WITH clean AS (
  SELECT
    product_name,
    category,
    about_product,
    SAFE_CAST(rating AS FLOAT64) AS rating_num, -- 문자열 평점을 숫자로
    rating_count, -- INT64 그대로 사용
    discounted_price AS discounted_price -- 이미 FLOAT64
  FROM
    `calcium-adapter-465502-n6.day4.amazon2`
  WHERE
    SAFE_CAST(rating AS FLOAT64) IS NOT NULL
    AND discounted_price <= max_budget -- 예산 내 필터링
),

-- 3) 만족도 점수 계산
score AS (
  SELECT
    product_name,
    category,
    about_product,
    rating_num,
    rating_count,
    discounted_price,
    rating_num * LN(rating_count + 1) AS satisfaction_score -- 로그 스케일 적용
  FROM clean
```


)

-- 4) 상위 5개 추천 추출

```
SELECT
  product_name,
  category,
  about_product,
  ROUND(satisfaction_score, 2) AS satisfaction_score,
  rating_num,
  rating_count,
  discounted_price
FROM score
ORDER BY satisfaction_score DESC
LIMIT 5;
```

단계	코드	설명
①	<code>SAFE_CAST(rating AS FLOAT64)</code>	문자열인 평점을 안전하게 숫자(<code>FLOAT64</code>)로 변환합니다. 변환 실패 시 <code>NULL</code> 처리.
②	<code>rating_count</code>	리뷰 수는 이미 정수(<code>INT64</code>)이므로 별도 전처리 없이 그대로 사용합니다.
③	<code>discounted_price</code>	할인가도 이미 숫자(<code>FLOAT64</code>) 타입이므로, 추가 전처리 없이 바로 사용합니다.
④	<code>WHERE SAFE_CAST(rating...) IS NOT NULL</code>	평점이 숫자로 변환되지 않는 행(예: 빈 문자열, "N/A")을 제거해 안정성을 확보합니다.
⑤	<code>AND discounted_price <= max_budget</code>	고객이 입력한 예산(<code>max_budget</code>)을 초과하는 상품을 미리 걸러냅니다.
⑥	<code>rating_num * LN(rating_count + 1)</code>	평점과 리뷰수를 결합한 만족도 점수 . 로그 스케일(<code>LN</code>)을 사용해, 리뷰 수가 많은 제품에 과도한 가중치가 쏠리지 않도록 조정합니다.
⑦	<code>ROUND(..., 2)</code>	결과의 가독성을 위해 소수점 둘째 자리까지 반올림합니다.
⑧	<code>ORDER BY satisfaction_score DESC</code>	높은 만족도 점수 순으로 상품을 정렬합니다.
⑨	<code>LIMIT 5</code>	상위 5개 상품만 최종 추천 리스트로 제공합니다.

3.3.1. 예산 필터(`discounted_price ≤ max_budget`)로 초과 지출을 방지

3.3.2.

평점 변환 + 리뷰 로그 가중치로 "예산 내에서 얼마나 검증된 만족을 얻을 수 있는지"를 정량화

3.3.3.

소수점 반올림, 내림차순 정렬, 5개 제한으로 결과를 깔끔하게 제시

3.4 결과

행	product_name	category	about_product	satisfaction_score	rating_num	rating_count	discounted_price
1	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video, Black	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Please select appropriate display resolution & refresh rate for proper display output. HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel (ARC)\Connects Blu-ray players, Fire TV, Apple TV, PS4, PS3, Xbox one, Xbox 360, computers and other HDMI-enabled devices to TVs, displays, A/V receivers and more\Cable allows you to share an Internet	57.04	4.4	426973	309.0
2	Amazon Basics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and more\Compatible with Blu-Ray players, computers, Apple TV, Roku, cable, PS4, Xbox One, and other HDMI-compatible devices\Solid copper conductors and full metal jacket shielding for durability and high-performance connectivity\Supports Ethernet, 3D, 4K video and Audio	57.04	4.4	426973	219.0
3	Amazon Basics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and more\Compatible with Blu-Ray players, computers, Apple TV, Roku, cable, PS4, Xbox One, and other HDMI-compatible devices\Solid copper conductors and full metal jacket shielding for durability and high-performance connectivity\Supports Ethernet, 3D, 4K video and Audio	57.04	4.4	426972	219.0
4	Amazon Basics High-Speed HD...	Electronics\HomeTheater,TV&Video\Accessories\Cables\HDMI Cables	HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel (ARC)\Connects Blu-ray players, Fire TV, Apple TV, PS4, PS3, Xbox one, Xbox 360, computers and other HDMI-enabled devices to TVs, displays, A/V receivers and more\Cable allows you to share an Internet connection among multiple devices without the need for a separate Ethernet	57.04	4.4	426973	309.0
5	SanDisk Extreme SD UHS I 64GB Card for 4K Video for DSLR and Mirrorless Cameras 170MB/s Read & 80MB/s Write	Electronics\Accessories\Memor...	Save time with card offload speeds upto 170mb/s powered by SanDisk quickflow technology\Perfect for shooting 4K UHD video and sequential burst mode photography\Capture uninterrupted video with UHS Class 3 (U3) and video speed class 30 (V30)\Built for and tested in harsh conditions	55.04	4.5	205052	939.0

1~4위 HDMI 케이블 제품군은 **평점 4.4×ln(리뷰수+1)** 공식으로 계산된 만족도 점수(≈57.04)가 압도적으로 높습니다. 특히 **42만 건이 넘는 리뷰**에 비해 가격(219~309원)이 매우 저렴해, "20,000원 이하 예산"으로도 **최고의 검증된 만족감**을 얻을 수 있습니다.

5위 SanDisk 메모리 카드는 **리뷰 20만 건**과 평점 4.5를 반영한 두 번째로 높은 점수(≈55.04)를 기록하여 **실속파 콘텐츠 크리에이터**에게 최적의 선택입니다.

4.1 이름: 지구를 사랑하는 친환경 그린 초이스 시스템

4.2 테마

'친환경', 'eco', 'green' 키워드를 담은 상품 중, **환경 가치 지수**를 계산해 추천합니다. **지속 가능성**을 중시하는 고객은, **환경보호와 검증된 품질, 합리적 가격**이라는 세 박자를 모두 만족시키는 선택을 할 수 있습니다. **about_product**에 '친환경'(eco/green) 키워드가 포함된 상품 중, **환경 가치 지수**를 계산해 **지속 가능·품질·가격 경쟁력**을 동시에 잡는 상품을 추천합니다.

- "환경도 챙기고, 품질과 가격도 다 잡는 착한 소비!"
- → 가치 소비를 추구하는 소비자를 위한 합리적 '친환경+품질+가격' 추천

4.3 구현 로직

```
-- 1) 텍스트에 '친환경' 키워드 포함된 상품만 추출
WITH clean AS (
SELECT
    product_name,
    category,
    about_product,
    SAFE_CAST(rating AS FLOAT64) AS rating_num,    -- 문자열 평점 → 숫자
    rating_count AS rating_count,    -- 이미 INT64, 그대로 사용
```

```

discounted_price          AS discounted_price  -- 이미 FLOAT64, 그대로 사용
FROM
`calcium-adapter-465502-n6.day4.amazon2`
WHERE
SAFE_CAST(rating AS FLOAT64) IS NOT NULL
AND (
  LOWER(about_product) LIKE '%친환경%'
  OR LOWER(about_product) LIKE '%eco%'
  OR LOWER(about_product) LIKE '%green%'
)
),

-- 2) 카테고리별 평균가 계산
stats AS (
  SELECT
    category,
    AVG(discounted_price) AS avg_price
  FROM clean
  GROUP BY category
),

-- 3) 환경 가치 지수(eco_index) 계산
green_score AS (
  SELECT
    c.*,
    -- (1) 평점 × ln(리뷰수+1)
    -- (2) 1 - (가격/카테고리평균가) 을 0~10 스케일로 변환
    rating_num * LN(rating_count + 1)
    + (1 - discounted_price / s.avg_price) * 10 AS eco_index
  FROM clean c
  JOIN stats s USING(category)
)

-- 4) 최상위 5개 추출
SELECT
  product_name,
  category,
  about_product,
  ROUND(eco_index, 2) AS eco_index,  -- 소수점 둘째 자리 반올림
  rating_num,
  rating_count,
  discounted_price
FROM green_score
ORDER BY eco_index DESC
LIMIT 5;

```

4.3.1. 친환경 필터

- `about_product` 에 '친환경', 'eco', 'green' 키워드가 하나라도 포함된 상품만 선별합니다.

4.3.2. 평점 안전 변환

- `SAFE_CAST(rating AS FLOAT64)` 로 문자열 평점을 숫자형으로 바꿔, 유효하지 않은 평점은 제거(`WHERE IS NOT NULL`).

4.3.3. 리뷰 수·가격 그대로 사용

- `rating_count` , `discounted_price` 가 이미 숫자형(INT64/FLOAT64)이므로, 문자열 전처리 함수 없이 그대로 사용합니다.

4.3.4. 카테고리별 평균가 계산

- 같은 카테고리 내 평균 `discounted_price` 를 구해, 가격 경쟁력 비교의 기준으로 삼습니다.

4.3.5. eco_index 산출

- 검증된 평판: `rating_num * LN(rating_count + 1)` 로 평점과 리뷰 수를 결합(리뷰 수가 지나치게 많아도 로그 스케일로 안정화).
- 가격 이점: `(1 - discounted_price / avg_price) * 10` 으로, 평균가 대비 얼마나 더 저렴한지를 0~10 척도로 환산.
- 두 값을 합산해 "친환경 + 품질 + 가격 경쟁력"을 모두 반영한 **eco_index** 생성.

4.3.6. 결과 정렬 & 제한

- `ORDER BY eco_index DESC LIMIT 5` 로, **eco_index**가 가장 높은 상위 5개 상품을 추천합니다.

4.4 결과

행	product_name	category	about_product	eco_index	rating_num	rating_count	discounted_price
1	SanDisk Ultra microSD UHS-I Ca...	ElectronicsAccessories/Memor...	Ideal for Android smartphones and tablets, and MIL camerasUp to 1TB to store even more hours of Full HD videoUp to 120MB/s transfer speeds let you move up to 1000 photos in a minute [32GB-1TB]Load apps faster with A1-rated performance, Operating Temperature-13°F to 185°FClass 10 for Full HD video recording and playback	54.57	4.4	67259	369.0
2	Swiffer Instant Electric Water Heater Faucet Tap Home-Kitchen Instantaneous Water Heater Tank less for Tap, LED Electric Head Water	Home&Kitchen/Heating,Cooling&AirQuality/WaterHeaters&Geyers/InstantWaterHeaters	✓Quick Electric Hot Water Tap Heating tube: high-purity copper liner heating element Rated voltage: 220V/50HZ Rated Power: 3000W/✓When you receive Swiffer Water tankless instant electric tap water heater, please do not test it in your hands, you need install it well and then test, or it will be dangerous. Please operate it according to the instructions.✓Support and well-being of dual-use	53.25	4.8	53803	1439.0
3	SanDisk Ultra® microSDXC™ UHS-I Card, 64GB, 140MB/s R, 10 Y Warranty, for Smartphones	ElectronicsAccessories/Memor...	Ideal storage for Android smartphones and tabletsUp to 140MB/s transfer speeds to move up to 1000 photos per minuteLoad apps faster with A1-rated performanceClass 10 for Full HD video recording and playbackCompatible with SanDisk MobileMate USB 3.0 microSD card reader	52.22	4.4	67259	569.0
4	SanDisk Ultra® microSDXC™ UHS-I Card, 64GB, 140MB/s R, 10 Y Warranty, for Smartphones	ElectronicsAccessories/Memor...	Ideal storage for Android smartphones and tabletsUp to 140MB/s transfer speeds to move up to 1000 photos per minuteLoad apps faster with A1-rated performanceClass 10 for Full HD video recording and playbackCompatible with SanDisk MobileMate USB 3.0 microSD card reader	52.22	4.4	67262	569.0
5	SanDisk Ultra microSD UHS-I Ca...	ElectronicsAccessories/Memor...	Ideal for Android smartphones and tablets, and MIL camerasUp to 1TB to store even more hours of Full HD videoUp to 120MB/s transfer speeds let you move up to 1000 photos in a minute [32GB-1TB]Load apps faster with A1-rated performanceClass 10 for Full HD video recording and playback	51.28	4.4	67260	649.0

• eco_index 계산

1. 검증된 평판: `rating_num × ln(rating_count + 1)`
2. 가격 경쟁력: `(1 - discounted_price/카테고리평균가) × 10`

이 둘을 더해 "친환경+품질+가격"의 복합 지표를 만들었습니다.

• 상위 결과 특징

- **SanDisk 메모리 카드** 제품군이 압도적으로 많은 리뷰(≈6.7만 건)와 높은 평점(4.4)을 바탕으로, 평균가 대비 저렴한 가격에서 최고의 **eco_index**를 기록했습니다.
- **Swiffer 물 히터**(2위)는 리뷰 수도 많고 평점(4.8)이 높아, 평균가 대비 합리적 가격으로 '친환경+성능' 가치를 잘 보여줍니다.

5.1 이름: 베스트 퍼포먼스 시스템

5.2 테마

- 이 추천 결과는 “나만 모르는 숨겨진 국민 필수템”, “실제 사람들이 가장 많이 써보고, 만족한 상품만 보고 싶다”라는 소비자의 근본적인 니즈에 완벽하게 부합합니다.
- 제품을 고를 때 가장 큰 걱정인 “이거 사도 괜찮을까?” → 이미 수십만 명이 ‘직접 경험’하고 평점을 준, 집단 지성 기반의 검증 결과만 보여주기 때문에 실패 없는 선택, 안심 구매를 보장합니다.
- 오직 “많은 리뷰 + 높은 평점”이라는 두 가지 기준으로 실패 없는 최강 스테디셀러만을 추천합니다.
- 실제 추천된 상품도 HDMI 케이블과 SanDisk 메모리 카드처럼 시장에서 이미 입증된 제품들로, “정말 검증된 제품만 쓰고 싶다”는 고객에게 최고의 신뢰를 제공합니다.

5.3 구현 로직

```
-- 1) 데이터 정제: 평점과 리뷰수, 가격이 유효한 데이터만 남김
WITH clean AS (
  SELECT
    product_name,
    category,
    about_product,
    SAFE_CAST(rating AS FLOAT64) AS rating_num,      -- ① 평점 문자열 → 숫자 변환
    rating_count,                                     -- ② 리뷰수 INT64 그대로 사용
    discounted_price                                -- ③ 가격도 FLOAT64 그대로 사용
  FROM
    `calcium-adapter-465502-n6.day4.amazon2`
  WHERE
    SAFE_CAST(rating AS FLOAT64) IS NOT NULL          -- ④ 평점 없는 행 제외
    AND rating_count IS NOT NULL                      -- ⑤ 리뷰수 없는 행 제외
    AND discounted_price IS NOT NULL                  -- ⑥ 가격 없는 행 제외
),

-- 2) 베스트 퍼포먼스 인덱스 산출: 평점 × 로그(리뷰수+1)
performance_score AS (
  SELECT
    product_name,
    category,
    about_product,
    rating_num,
    rating_count,
    discounted_price,
    rating_num * LN(rating_count + 1) AS performance_index -- ⑦ 평점 × ln(리뷰수+1)
  FROM clean
```

)

-- 3) 상위 5개 추출

```
SELECT
  product_name,
  category,
  about_product,
  ROUND(performance_index, 2) AS performance_index,    -- ⑧ 보기 좋게 반올림
  rating_num,
  rating_count,
  discounted_price
FROM performance_score
ORDER BY performance_index DESC                        -- ⑨ 내림차순 정렬
LIMIT 5;                                              -- ⑩ 5개만 추출
```

단계	코드	설명
①	SAFE_CAST(rating AS FLOAT64)	평점(rating)이 만약 문자열이라도, 안전하게 숫자형으로 변환 (예: "4.5" → 4.5)
②	rating_count	리뷰수는 이미 INT64(정수)이므로, 추가 전처리 없이 바로 사용
③	discounted_price	가격도 FLOAT64로 이미 저장되어 있으면 그대로 사용
④~⑥	IS NOT NULL	평점, 리뷰수, 가격 중 하나라도 결측(Null)이면 데이터에서 제외 (계산 오류 방지)
⑦	rating_num * LN(rating_count + 1)	베스트 퍼포먼스 인덱스 공식: 평점이 높을수록, 리뷰수가 많을수록 값이 커집니다. 리뷰수가 많을수록 지수는 늘지만, 로그 함수 를 써서 극단적 쏠림 현상 방지
⑧	ROUND(..., 2)	결과의 가독성을 위해 소수점 둘째 자리까지 반올림
⑨	ORDER BY performance_index DESC	베스트 퍼포먼스 인덱스가 높은 순으로 정렬(신뢰성 최상위 우선)
⑩	LIMIT 5	상위 5개 상품만 추천 결과로 노출

- 많은 리뷰수와 높은 평점이 결합된 상품이 “베스트 퍼포먼스”로 추출
- 로그 스케일을 적용해, 리뷰수 1만/10만/100만 같은 차이도 합리적으로 반영
- 가격, 할인, 키워드, 예산 등 기타 요소 없이 “집단 검증력” 하나만을 절대 기준으로 삼음
- 심플하고 명확한 지수로 고객이 “실패 없는 선택”을 할 수 있게 도움

“누가 봐도 신뢰할 만한 상품”만을 골라주며, **가격, 할인, 예산 등 여타 조건을 배제**하고, 집단의 경험(리뷰), 만족(평점)이라는 두 가지 핵심만을 정량적으로 결합한다는 점이 다른 추천 시스템들과 뚜렷하게 차별화됩니다.

5.4 결과

행	product_name	category	about_product	performance_index	rating_num	rating_count	discounted_price
1	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video, Black	Electronics HomeTheater,TV&Video Accessories Cables HDMI Cables	Please select appropriate display resolution & refresh rate for proper display output. HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K	57.04	4.4	426973	309.0
2	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics HomeTheater,TV&Video Accessories Cables HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and moreCompatible with	57.04	4.4	426973	219.0
3	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	Electronics HomeTheater,TV&Video Accessories Cables HDMI Cables	Flexible, lightweight HDMI cable for connecting media devices to playback display such as HDTVs, projectors, and moreCompatible with	57.04	4.4	426972	219.0

4	Amazon Basics High-Speed HD...	Electronics HomeTheater,TV&Video Accessories Cables HDMI Cables	HDMI A Male to A Male Cable: Supports Ethernet, 3D, 4K video and Audio Return Channel (ARC) Connects Blu-ray players, Fire TV, Apple	57.04	4.4	426973	309.0	
5	SanDisk Extreme SD UHS I 64GB Card for 4K Video for DSLR and Mirrorless Cameras 170MB/s Read & 80MB/s Write	Electronics Accessories Memor...	Save time with card offload speeds upto 170mb/s powered by SanDisk quickflow technology Perfect for shooting 4K UHD video and	55.04	4.5	205052	939.0	

순위	product_name	performance_index	rating_num	rating_count	discounted_price
1	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video, Black	57.04	4.4	426,973	309.0
2	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	57.04	4.4	426,973	219.0
3	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	57.04	4.4	426,972	219.0
4	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video, Black	57.04	4.4	426,973	309.0
5	SanDisk Extreme SD UHS I 64GB Card for 4K Video for DSLR...	55.04	4.5	205,052	939.0

• HDMI 케이블(1~4위)

- 평점 4.4, 리뷰 42만 6천 건이라는 압도적 신뢰도,
- 이 제품들이 장기간 많은 사용자들에게 선택되고, 꾸준히 높은 평점을 유지했다는 점에서
"누구나 안심하고 구매할 수 있는 국민템"임을 증명합니다.
- 가격(219~309원)도 합리적이지만, 이 시스템은 가격과 무관하게 **오로지 검증력**만을 기준으로 추천합니다.

• SanDisk 메모리 카드(5위)

- 평점 4.5, 리뷰 20만 5천 건.

- IT 주변기기 중에서도 가장 폭넓게 사랑받으며,
"대중적 검증을 통과한 고성능 스테디셀러"라는 점에서 높은 순위를 기록합니다.

전체 결과의 의미

위 다섯 가지의 시스템은 모두

구매 결정의 불확실성을 제거하고, 고객이 **진짜 중요하게 여기는 기준**(가성비·할인·예산·환경·집단검증)을 데이터로 증명해 주기 때문에, 단순히 추천을 넘어 **고객 신뢰와 만족도를 극대화**합니다.

- 모든 시스템이 서로 다른 '추천 기준'을 적용
 - 가격, 할인, 예산, 환경, 집단 검증 등 각기 다른 관점에서 최적의 상품만 선별
- 실제 결과는 **대체로 평점·리뷰수 우수 상품**이 공통적으로 강세
→ 하지만, **각 시스템의 기준에 따라 "상위 추천 상품의 순위와 성격"이 확실히 달라짐**
- **고객 입장에 따라**
 - "똑똑한 가성비 소비",
 - "지금만 가능한 진짜 딜",
 - "내 예산을 지켜주는 검증템",
 - "환경까지 생각한 현명한 구매",
 - "수많은 사람에게 검증된 국민 필수템" **자신의 상황/가치에 맞는 선택**이 가능