

project_이수아

* 아마존 데이터셋 이해하기

product_id	STRING
product_name	STRING
category	STRING
discounted_price	FLOAT
actual_price	FLOAT
discount_percentage	FLOAT
rating	STRING
rating_count	INTEGER
about_product	STRING
user_id	STRING
user_name	STRING
review_id	STRING
review_title	STRING
review_content	STRING
img_link	STRING
product_link	STRING

추천시스템 1.

1) 이름 : "진짜 만족도가 높은 상품을 찾아라!"

2)테마 : 제품 리뷰 평점이 높아도 실제 사용 고객수가 적은 경우는 평점을 구매의 척도로 삼기에는

신뢰성이 많이 떨어진다. 높은 평점을 유지하는 제품 중 실제 사용고객수도 많은 제품만 선별해서 추천해주는 시스템이 있을때 소비자들의 구매 시간을 단축시켜 줄 수 있으며, 제품 선택에서 실패할 확률을 낮춰줄 수 있다.

3)핵심 질문 : "많은 사람이 평가했는데도 여전히 평점이 높은 상품은 뭘까?"

4)구현 로직: rating , rating_count

SQL 로직: WHERE rating >= 4.5 AND rating_count >= 500 ,

ORDER BY rating DESC, rating_count DESC

```
SELECT product_id, product_name,
       cast(rating as FLOAT64) AS rating,
       cast(replace(cast(rating_count as string), ',', '')) as int64 AS rating_count,
       discounted_price,
       actual_price,
       category,
       product_link
FROM `modulabs-project-470807.Day4.Amazon`
WHERE
safe_cast(rating as FLOAT64) >= 4.5
AND cast(replace(cast(rating_count as string), ',', '')) as int64 >= 500
ORDER BY
safe_cast(rating as FLOAT64) DESC,
cast(replace(cast(rating_count as string), ',', '')) as int64 DESC;
```

5)결과 :

특정 기준(별점 4.5 이상, 평가 수 500 이상)을 충족하는 제품 중 평점과 평가수를 내림차순으로 정렬하여 , 평점도 높으면서 . 리뷰수도 많은 신뢰성이 높은 상품 순으로 추천하는 시스템이다.

(평점이 높은 순 우선, 동률이면 리뷰수가 많은 순)

구매만족도가 높은 상품 순으로 상단에서 쉽게 확인 할 수 있다.

이는 잠재 고객들의 제품 신뢰도와 서비스 편의성을 높여 실제 구매력을 높일 수 있다.

작업 정보	결과	시각화	JSON	실행 세부정보	실행 그래프			
행	product_id	product_name	rating	rating_count	discounted_price	actual_price	category	product_link
1	B0BR4F878Q	Swiffer Instant Electric Water Heater Faucet Tap Home-Kitchen Instantaneous Water Heater Tankless for Tap, LED Electric Head Water Heaters Tail Gallon Comfort(3000W) ((Pack of 1))	4.8	53803	1439.0	1999.0	Home&Kitchen Heating,Cooling&AirQuality WaterHeaters&Geysers InstantWaterHeaters	https://www.amazon.in/Swiffer-Instant-Electric-Home-Kitchen-Instantaneous/dp/B0BR4F878Q/ref=sr_1_150?qid=1672923597&s=kitchen&sr=1-150
2	B0B53DS4TF	Instant Pot Air Fryer, Vortex 2QT, Touch Control Panel, 360° EvenCrisp™ Technology, Uses 95 % less Oil, 4-in-1 Appliance: Air Fry, Roast, Bake, Reheat (Vortex 1.97Litre, Black)	4.8	3964	4995.0	20049.0	Home&Kitchen Kitchen&HomeAppliances SmallKitchenAppliances DeepFatFryers AirFryers	https://www.amazon.in/Instant-Vortex-2QT-EvenCrisp-TM-Technology/dp/B0B53DS4TF/ref=sr_1_316?qid=1672923609&s=kitchen&sr=1-316
3	B0B23LW7NV	Spigen EZ Fit Tempered Glass Screen Protector for iPhone 14 Pro Max - 2 Pack (Sensor Protection)	4.7	7779	999.0	2899.0	Electronics Mobles&Accessories MobileAccessories Maintenance,Upkeep&Repairs ScreenProtectors	https://www.amazon.in/Spigen-Tempered-Screen-Protector-iPhone/dp/B0B23LW7NV/ref=sr_1_130?qid=1672895784&s=electronics&sr=1-130
4	B09WN3SRC7	Sony Bravia 164 cm (65 inches) 4K Ultra HD Smart LED Google TV KD-65X74K (Black)	4.7	5935	77990.0	139900.0	Electronics HomeTheater,TV&Video Televisions SmartTelevisions	https://www.amazon.in/Sony-Bravia-inches-Google-KD-65X74K/dp/B09WN3SRC7/ref=

제품 ID	제품 이름	평점	리뷰 수	할인 가격	실제 가격	카테고리
B0BR4F878Q	Swiffer 인스턴트 전기 수도꼭지용 온수기, LED 전기 헤드 수온기, 3000W ((1개입))	4.8	53,803	₹1,439	₹1,999	주방 & 가전
B0B53DS4TF	인스턴트팟 에어프라이어 Vortex 2QT, 터치패널, 4-in-1 (에어프라이, 로스트, 베이킹, 리히트)	4.8	3,964	₹4,995	₹20,049	주방 가전
B0B23LW7NV	Spigen EZ Fit 아이폰14 Pro Max 강화유리 보호 필름 (2개입)	4.7	7,779	₹999	₹2,899	전자제품/휴대폰 액세서리
B09WN3SRC7	Sony Bravia 164cm(65인치) 4K 울트라 HD 스마트 LED 구글 TV KD-65X74K (블랙)	4.7	5,935	₹77,990	₹139,900	TV & 전자제품

추천시스템 2.

1) 이름 : "숨겨진 보석을 찾아라! 가성비와 만족도 두 마리 토끼를 잡아라!"

2)테마 : " 할인율이 높다고 무조건 좋은 상품은 아닙니다.

할인율이 높은 상품 순서대로 보여주는 사이트에서 가성비와 품질 모두 만족하는 상품을 찾기위해서는 많은 시간을 소요해야합니다.

바쁜 현대사회에서 소비자들은 할인율이 아무리 높아도 quality가 떨어지는 상품들이 계속 나오는 화면을 보면서 스크롤을 내리다가 지쳐서 구매를 포기하는 경우가 종종 생긴다.

'높은 할인율' 과 '높은 평점'을 모두 만족하는 상품 순서로 보여주는 사이트에서 소비자는 불필요한 시간과 비용을 줄이리 수 있고, 판매자는 구매자가 진정 만족하면서 '득템'할 수 있는 인기 사이트를 갖게 될 것이다.

3)핵심 질문 : " 할인도 많이 해주면서, 사람들의 만족도도 높은 상품 은 뭘까?"

4)구현 로직: : discount_percentage , rating

SQL 로직 : WHERE discount_percentage >= 0.5 AND rating >= 4.0

ORDER BY discount_percentage DESC

```
SELECT
  product_id,
  product_name,
  CAST(REPLACE(CAST(discount_percentage AS STRING), '%', '' ) AS FLOAT64) AS discount_percentage,
  CAST(rating AS FLOAT64) AS rating,
  discounted_price,
  actual_price,
  category,
  product_link
FROM `modulabs-project-470807.Day4.Amazon`
WHERE
  CAST(REPLACE(CAST(discount_percentage AS STRING), '%', '' ) AS FLOAT64) >= 0.5
  AND SAFE_CAST(rating AS FLOAT64) >= 4.0
ORDER BY
  discount_percentage DESC,
  rating DESC;
```

5)결과 : '추천시스템 2' 는 데이터셋에서 할인율이 50% 이상이고, 고객 평점이 4.0 이상인 모든 제품 정보를 할인율이 높은 순서대로, 그리고 할인율이 같을 경우 평점이 높은 순서대로 보여줍니다.

평소 구매할 생각이 없었던 제품이라도 할인률이 높고 제품 만족도까지 높은 가성비 좋은 잇템을 발견하게 되어, 소비자는 "할인 혜택을 받아 가성비와 높은 구매 평점으로 인한 신뢰도가 있는 제품을 구매할 수 있어 만족도가 높아질 것입니다.

더불어 소비자는 좋은 제품을 싸게 살 수 있어 이익이고, 판매자는 별다른 마케팅이 없어도 이 추천시스템을 통해 제품이 노출되면서 판매실적이 높아 질 수 있어 이익을 볼 수 있습니다.

	product_id	product_name	discount_percent...	rating	discounted_price	actual_price	category
1	B097C564GC	rtis [2 Pack] Mini USB C Type C Adapter Plug, Type C Female to USB A Male Charger Charging Cable Adapter Converter compatible for	0.94	4.3	294.0	4999.0	Computers&Accessories Accessories&Peripherals Adapters USBtoUSBAdapters
2	B0BF57RN3K	Fire-Boltt Ninja Call Pro Plus 1.83" Smart Watch with Bluetooth Calling, AI Voice Assistance, 100 Sports Modes IP67 Rating, 240*280	0.91	4.2	1799.0	19999.0	Electronics WearableTechnology SmartWatches
3	B0BF54972T	Fire-Boltt Ninja Call Pro Plus 1.83" Smart Watch with Bluetooth Calling, AI Voice Assistance, 100 Sports Modes IP67 Rating, 240*280	0.91	4.2	1799.0	19999.0	Electronics WearableTechnology SmartWatches

순위	제품 ID	제품명	할인율	평점	할인가(₹)	정상가(₹)	카테고리 ¹⁾
1	B097C564GC	[2개 세트] 미니 USB C 타입 C 어댑터 플러그, Type C Female → USB A Male 충전 케이블 어댑터, 호환용 컨버터	94%	4.3	₹294	₹49,999	컴퓨터 & 액세서리 > 주변기기 > 어댑터 > USB 변환 어댑터
2	B0BF57RN3K	Fire-Boltt Ninja Call Pro Plus 1.83" 스마트워치, 블루투스 통화, AI 음성 지원, 100개 스포츠 모드, IP67 방수 등급	91%	4.2	₹1,799	₹19,999	전자제품 > 웨어러블 기술 > 스마트워치
3	B0BF54972T	Fire-Boltt Ninja Call Pro Plus 1.83" 스마트워치, 블루투스 통화, AI 음성 지원, 100개 스포츠 모드, IP67 방수 등급	91%	4.2	₹1,799	₹19,999	전자제품 > 웨어러블 기술 > 스마트워치

추천시스템 3.

1) 이름 : " 관심 카테고리 구매 베스트!!!"

2)테마 : 소비자별로 자주 찾는 특정 카테고리 내에서 맞춤 추천을 해준다면, 소비자 입장에서 여러 사이트를 돌아다니면서 물품을 찾는데 시간을 낭비하지 않을 수 있다.
 자주 찾는 사이트에서 이미 축적된 데이터로 소비 성향 별 맞춤 물건을 추천해 준다면 빠르고 효율적으로 자주 이용하는 물건을 구매할 수 있게 된다.
 판매자 입장에서도 단골 고객을 유치하여 판매율을 높일 수 있다.

3)핵심 질문 : "특정 카테고리에서 가장 '핫'한 아이템은 뭘까?"

4)구현 로직: `category` , `rating` , `rating_count`

SQL 로직: `UNNEST(SPLIT(category, '|'))`
`WHERE category_value = 'Camping&Hiking'`

```

WITH CategoryProducts AS (
  SELECT
    product_id,
    product_name,
    rating,
    rating_count,
    cat
  FROM
    `modulabs-project-470807.Day4.Amazon`,
    UNNEST(SPLIT(category, '|')) AS cat
)
SELECT
  product_id,
  product_name,
  rating,
  rating_count,
  cat
FROM
  CategoryProducts
WHERE
  cat = 'Electronics'
  AND SAFE_CAST(rating AS FLOAT64) > 4.0
  AND SAFE_CAST(rating_count AS INT64) >= 500
ORDER BY
  SAFE_CAST(rating_count AS INT64) DESC
LIMIT 10;

```

5)결과 :

고객이 자주 이용하는 'Electronics' 카테고리에서 평점도 높고 리뷰수도 많은 구매 베스트 상품들을 순서대로 보여줌으로서 빠르고 편리하게 그리고 만족도가 높은 구매를 할 수 있다.

데이터를 분석해보면 , 고객이 자주 이용하는 'Electronics' 카테고리에서는

- AmazonBasics HDMI 케이블 제품이 여러 길아·팩으로 상위권을 차지.
- boAt 브랜드 유선 이어폰 시리즈가 인기가 높음.
- Redmi 보급형 스마트폰 시리즈도 많은 리뷰를 기록.
- **평점(rating)**
 - 상위 4개 HDMI 케이블: 평점 **4.4** (매우 높음)
 - 나머지 이어폰·휴대폰: 평점 **4.1**
- **리뷰 개수(rating_count)**
 - HDMI 케이블은 **42만 개 이상 리뷰**로 인기가 많습니다.
 - 이어폰도 **36만 개 이상 리뷰**, 스마트폰 시리즈는 **31만 개 이상 리뷰**로 모두 대중적인 인기 상품입니다.

행	product_id	product_name	rating	rating_c...	cat
1	B07KSMBL2H	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	4.4	426973	Electronics
2	B014I8SSD0	Amazon Basics High-Speed HDMI Cable, 6 Feet - Supports Ethernet, 3D, 4K video,Black	4.4	426973	Electronics
3	B014I8SX4Y	Amazon Basics High-Speed HDMI Cable, 6 Feet (2-Pack),B...	4.4	426973	Electronics
4	B07KSMBL2H	AmazonBasics Flexible Premium HDMI Cable (Black, 4K@60Hz, 18Gbps), 3-Foot	4.4	426972	Electronics
5	B07GQD4K6L	boAt Bassheads 100 in Ear Wired Earphones with Mic(Furious Red)	4.1	363713	Electronics
6	B07GPXXNNG	boAt Bassheads 100 in Ear Wired Earphones with Mic(Taffy Pink)	4.1	363713	Electronics
7	B071Z8M4KX	boAt BassHeads 100 in-Ear Wired Headphones with Mic (...)	4.1	363711	Electronics
8	B09GFPVD9Y	Redmi 9 Activ (Carbon Black, 4GB RAM, 64GB Storage) Octa-core Helio G35 5000 mAh Battery	4.1	313836	Electronics
9	B09GLXVH9	Redmi 9A Sport (Coral Green, 2GB RAM, 32GB Storage) 2GHz Octa-core Helio G25 Processor 5000 mAh Battery	4.1	313836	Electronics
10	B09GFM8CGS	Redmi 9A Sport (Carbon Black, 2GB RAM, 32GB Storage) 2GHz Octa-core Helio G25 Processor 5000 mAh Battery	4.1	313832	Electronics

상품 ID (product_id)	상품명 (product_name)	평점 (rating)	리뷰 수 (rating_count)	카테고리 (cat)
B07KSMBL2H	아마존베이식스 플렉시블 프리미엄 HDMI 케이블 (검정, 4K@60Hz, 18Gbps), 0.9m	4.4	426,973	전자제품
B014I8SSD0	아마존베이식스 하이스피드 HDMI 케이블, 1.8m - 이더넷/3D/4K 지원, 검정	4.4	426,973	전자제품
B014I8SXY4	아마존베이식스 하이스피드 HDMI 케이블 (2개 세트), 검정	4.4	426,973	전자제품
B07KSMBL2H	아마존베이식스 플렉시블 프리미엄 HDMI 케이블 (검정, 4K@60Hz, 18Gbps), 0.9m	4.4	426,972	전자제품
B07GQD4K6L	boAt Bassheads 100 인이어 유선 이어폰, 마이크 포함 (퓨리어스 레드)	4.1	363,713	전자제품
B07GPXXNNG	boAt Bassheads 100 인이어 유선 이어폰, 마이크 포함 (테피핑크)	4.1	363,713	전자제품
B071Z8M4KX	boAt BassHeads 100 인이어 유선 이어폰, 마이크 포함 (...)	4.1	363,711	전자제품
B09GFPVD9Y	Redmi 9 Activ (카본 블랙, 4GB RAM, 64GB 저장공간)	4.1	313,836	전자제품

추천시스템 4:

1) 이름 : " 오늘의 인기상품 ! 요즘 인기있는 핫한 아이템"

2)테마 : 짧은 구매 기간 동안 가장 많은 리뷰가 달리는 상품은 요즘 트렌드가 잘 반영된 상품으로

주목받고 있을 확률이 높습니다. 무심코 들어가본 사이트에서 최근 가장 인기있는 핫한 아이템이 소개되면 잠재 구매자의 흥미를 유발하고 이는 구매로 이어 질 수 있습니다.

3)핵심 질문 : "요즘 사람들이 가장 많이 리뷰를 남기는 상품은 뭘까?"

4)구현 로직 : `user_id` (또는 `review_id`)

SQL 로직 : `ARRAY_LENGTH(SPLIT(user_id, ','))` , 리뷰의 총개수를 계산 이 값을 기준으로 정렬하면 '가장 리뷰가 많이 달린' 상품 순위를 얻을 수 있습니다.

5)결과 :

```
-- 예시 쿼리
SELECT
  product_name,
  rating,
  ARRAY_LENGTH(SPLIT(user_id, ',')) AS total_reviews -- user_id 개수를 세어 리뷰 수로 활용
FROM
  `your_project.your_dataset.your_table`
ORDER BY
  total_reviews DESC
LIMIT 20;
```

예상 출력 결과:

product_name	rating	total_reviews
Apple iPhone 15 Pro Max	4.2	28547
Samsung Galaxy S24 Ultra	4.4	23891
Sony WH-1000XM5 Wireless Headphones	4.6	19832
Apple MacBook Pro 16-inch	4.3	17246
Nintendo Switch OLED	4.5	16834
Amazon Echo Dot (5th Gen)	4.1	15923
Instant Pot Duo 7-in-1 Electric Pressure Cooker	4.7	14567
Dyson V15 Detect Cordless Vacuum	4.4	13892

추천시스템5.

1) 이름 : "이런 상품은 어때요? 이걸 산 사람들이 함께 본 상품 추천 "

2)테마 : ****'A** 상품에 리뷰를 남긴 사람들은 B 상품에도 리뷰를 남겼다******는 사실을 발견하면,
A 상품을 보는 사람에게 B 상품을 추천할 수 있습니다.

3)핵심 질문 : "특정 상품을 리뷰한 사람들은 또 어떤 상품들을 리뷰했을까?"

4)구현 로직 : `product_id` , `user_id`

1. UNNEST를 사용해 (product_id, unnested_user_id) 페어를 만듭니다.
2. 특정 상품(예: 'B07JW9H4J1')을 리뷰한 user_id 목록을 찾습니다.
3. 이 user_id 목록을 가진 사람들이 리뷰한 '다른' 상품들을 모두 찾아 집계하고, 가장 많이 등장한 상품 순으로 추천합니다. WITH 절을 사용하여 구현할 수 있습니다.

5)결과 :

```

-- 예시 쿼리 (협업 필터링)
WITH UserProductPairs AS (
  SELECT
    product_id,
    product_name,
    user_id_split
  FROM
    `your_project.your_dataset.your_table`,
    UNNEST(SPLIT(user_id, ',')) AS user_id_split
),
TargetUsers AS (
  SELECT DISTINCT user_id_split
  FROM UserProductPairs
  WHERE product_id = 'B07JW9H4J1' -- 타겟 상품 ID
)
SELECT
  upp.product_name,
  COUNT(*) AS co_purchase_count,
  upp.product_id
FROM
  UserProductPairs upp
JOIN
  TargetUsers tu ON upp.user_id_split = tu.user_id_split
WHERE
  upp.product_id != 'B07JW9H4J1' -- 타겟 상품 제외
GROUP BY
  upp.product_name, upp.product_id
ORDER BY
  co_purchase_count DESC
LIMIT 10;

```

5)결과 :

예상 출력 결과:

product_name	co_purchase_count	product_id
USB-C to Lightning Cable	1247	B08H5QMZPX
iPhone 15 Pro Max Case	1089	B09WMTKLVS
Wireless Charging Pad	892	B07K6DQTS1
MagSafe Car Mount	756	B08Y8QMZPX
Screen Protector for iPhone 15	643	B07H8QMZPX
Apple AirPods Pro (2nd Gen)	587	B0BDHB9Y8P
Portable Power Bank 20000mAh	521	B08H93GNMJ