# Main Quest 2 템플릿


평가 항목
추천시스템 1
- 1. 추천시스템의 이름
- 2. 추천시스템의 테마 및 목적
- 3. 구현 로직 (코드, 쿼리 및 구조도)
- 4. 결과
- 5. 참고자료

추천시스템 2
- 1. 추천시스템의 이름
- 2. 추천시스템의 테마 및 목적
- 3. 구현 로직 (코드, 쿼리 및 구조도)
- 4. 결과
- 5. 참고자료

추천시스템 3
- 1. 추천시스템의 이름
- 2. 추천시스템의 테마 및 목적
- 3. 구현 로직 (코드, 쿼리 및 구조도)
- 4. 결과
- 5. 참고자료

추천시스템 4
- 1. 추천시스템의 이름
- 2. 추천시스템의 테마 및 목적
- 3. 구현 로직 (코드, 쿼리 및 구조도)
- 4. 결과
- 5. 참고자료

추천시스템 5
- 1. 추천시스템의 이름
- 2. 추천시스템의 테마 및 목적
- 3. 구현 로직 (코드, 쿼리 및 구조도)
- 4. 결과
- 5. 참고자료

회고


## 평가 항목

| 채점 기준 | 설명 |
|---|---|
| 1. 추천 시스템의 창의성과 복잡성 | **5가지의 추천 시스템 아이디어가 모두 독창적인가?** |
| 2. SQL 구현의 정확성 및 효율성 | **SQL 쿼리가 데이터의 목적에 맞게 설계되었는가?** |
| 3. 보고서 구성 및 설명력 | **논리적으로 추천 시스템 구현 로직을 설명했는가?** |

## 추천시스템 1

### 1. 추천시스템의 이름

- Keyword-Based Budget Finder

### 2. 추천시스템의 테마 및 목적

- 사용자가 입력한 특정 키워드(예 : Monitor)가 포함된 제품(제품명) 중에서 가장 저렴한 상품를 추천합니다.
- 첫 페이지에는 10개의 검색 결과를 보여 주고, 나머지 결과는 페이지로 나누어서 제공
- 검색 결과 노출 정보
  - 상품명(product_name)
  - 상품 소개(about_product)
  - 상품 링크(product_link)
  - 할인 가격(discounted_price)
- 상품 링크 클릭 시 해당 상품의 상세 페이지로 이동하여 상품 확인 및 즉각적인 구매 유도 가능.

## 3. 구현 로직 (코드, 쿼리 및 구조도)

- Bigquery SQL

```
SELECT product_name, about_product, product_link, discounted_price
FROM modulabs_amazon.data
WHERE product_name LIKE '%Monitor%'
ORDER BY discounted_price ASC
LIMIT 10;
```

  - 사용자가 입력한 Keyword를 포함하고 있는 열을 조회하기 위해 WHERE 절에 `LIKE "%(keyword)%"` 적용
  - sql문에서는 검색 결과 확인을 위해 상위 10개 조회 적용.

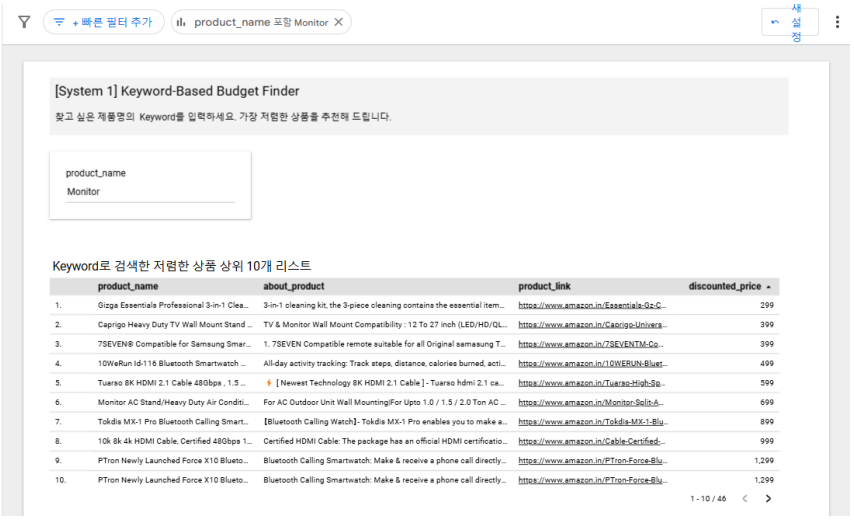
## 4. 결과

- SQL 조회 결과

| 행 | product_name | about_product | product_link | discounted_price |
|---|---|---|---|---|
| 1 | Gizga Essentials Professional 3-in-1 Cleaning Kit for Camera, Lens, Binocular, Laptop, TV, Monitor, Smartphone, Tablet (Includes: Cleaning Liquid 100ml, Plush Microfiber Cloth, Dust Removal Brush) | 3-in-1 cleaning kit, the 3-piece cleaning contains the essential items necessary for the proper maintenance and care of your optical equipment. It easily cleans dirt and dust. Effectively removes smudges and fingerprints from lenses, LCD and | https://www.amazon.in/Essentials-Gz-Ck-101-Professional-Micro-Fiber-Antibacterial/dp/B01IBRHE3E/ref=sr_1_144?qid=1672903001&s=computers&sr=1-144 | 299.0 |
| 2 | Caprigo Heavy Duty TV Wall Mount Stand for 12 to 27 inches LED/LCD/Monitor Screen's, Full Motion Rotatable Universal TV & Monitor Wall Mount Bracket with Swivel & Tilt Adjustments (Single Arm - M416) | TV & Monitor Wall Mount Compatibility : 12 To 27 inch (LED/HD/QLED/LCD/UHD/3D/4K/Monitor/Curved/Smart Screens) From All Universal Brands Like Samsung, BenQ, LG, Acer, Dell, Sony, Lenovo, HP, | https://www.amazon.in/Caprigo-Universal-Monitor-Rotatable-Black-M416/dp/B083GQGT3Z/ref=sr_1_243?qid=1672909136&s=electronics&sr=1-243 | 399.0 |
| 3 | 7SEVEN® Compatible for Samsung Smart 4K | 1. 7SEVEN Compatible remote suitable for all | https://www.amazon.in/7SEVENTM-Compatible- | 399.0 |

페이지당 결과 수: 50 1 – 10 (전체 10행)

- LookerStudio




## 5. 참고자료

-

---

# 추천시스템 2

## 1. 추천시스템의 이름

- Top-Rated Product Spotlight

## 2. 추천시스템의 테마 및 목적

- 제품의 평점(rating)과 평점 개수(rating_count)를 각각 10단계로 세분화하고 이 중 최상위 등급에 해당하는 제품만을 추천합니다.
- 품질과 신뢰도, 인기도 등과 관련 있는 rating, rating_count 지표를 활용하여 고객들에게 명확하고 직관적인 상품 선택 옵션을 제공할 수 있음.
- 제품 할인 정보와 상세 페이지 링크를 같이 제공하여 즉각적인 구매 행동으로 이어질 수 있음.

## 3. 구현 로직 (코드, 쿼리 및 구조도)

- Bigquery 기본 sql문

```
WITH RankedData AS (
  SELECT
    NTILE(10) OVER (ORDER BY rating DESC) AS rating_Rank,
    NTILE(10) OVER (ORDER BY rating_count DESC) AS rating_count_Rank,
    rating, rating_count,
    product_name, discounted_price, product_link
  FROM modulabs_amazon.data
)
SELECT *
FROM RankedData
WHERE rating_Rank = 1 AND rating_count_RANK = 1;
```

- LookerStudio에서 활용할 임시 테이블 생성 sql문

```
CREATE OR REPLACE TABLE modulabs_amazon.res2 AS
  SELECT
    NTILE(10) OVER (ORDER BY rating DESC) AS rating_Rank,
    NTILE(10) OVER (ORDER BY rating_count DESC) AS rating_count_Rank,
    rating, rating_count,
    product_name, discounted_price, product_link
  FROM modulabs_amazon.data;
```

- rating과 rating_count를 내림차순으로 정렬한 다음 NTILE(10) 적용
- 1 : 최상위등급, 10 : 최하위등급

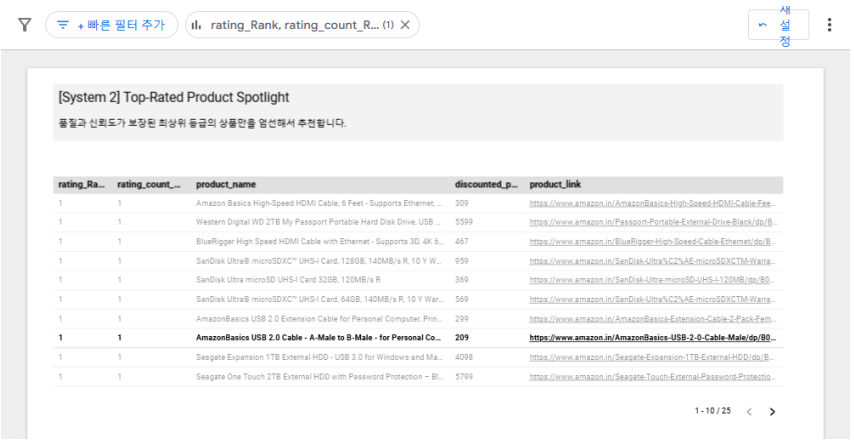
## 4. 결과

- SQL 조회 결과 : 총 25개

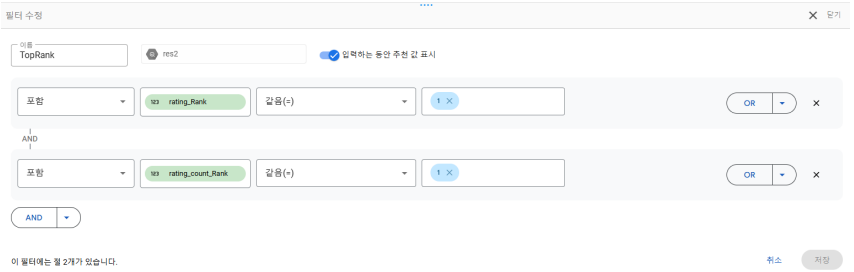
| 행 | rating_Rank | rating_count_Rank | rating | rating_count | product_name | discounted_price | product_link |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 4.4 | 426973 | Amazon Basics High-Speed HD... | 309.0 | https://www.amazon.in/AmazonBasics-High-Speed-Cable-2-Pack-Black/dp/B014I8SX4Y/ref=sr_1_73? |
| 2 | 1 | 1 | 4.5 | 54405 | Logitech M235 Wireless Mouse, 1000 DPI Optical Tracking, 12 Month Life Battery, Compatible with Windows, Mac, | 699.0 | https://www.amazon.in/Logitech-M235-Wireless-Mouse-Grey/dp/B004IO5BMQ/ref=sr_1_163? |
| 3 | 1 | 1 | 4.4 | 67259 | SanDisk Ultra® microSDXC™ UHS-I Card, 128GB, 140MB/s | 959.0 | https://www.amazon.in/SanDisk-Ultra%C2%AE- |

페이지당 결과 수: 50 1 – 25 (전체 25행)

- LookerStudio 화면




** TopRank 필터 적용




## 5. 참고자료

- 

---

# 추천시스템 3

## 1. 추천시스템의 이름

- Top User-Based Product Recommender

## 2. 추천시스템의 테마 및 목적

- 가장 많은 리뷰를 남긴 고객이 구매한 상품 정보(상품명, 상품 가격, 상품 링크)를 추천합니다.

## 3. 구현 로직 (코드, 쿼리 및 구조도)

- 구현 process
  - raw data에서 “user_id”(리뷰 작성자 아이디들) 컬럼의 데이터를 UNNEST 함수를 사용하여 개별 데이터 행으로 분리 → 임시 테이블(TMP)로 저장
  - 임시 테이블(TMP)에서 그룹 함수를 사용하여 사용자 아이디별 리뷰 개수, 구매액(상품 가격의 합으로 가정) 계산
  - 리뷰 개수와 구매액을 기준으로 내림차순 정렬 후 Top User 아이디 조회
  - 해당 user_id로 리뷰 작성된 상품 정보를 조회 후 출력
- Bigquery sql 문

```
WITH TMP AS (
  SELECT *
  FROM modulabs_amazon.data,
  UNNEST(SPLIT(user_id, ',')) AS user_id_value
)
SELECT DISTINCT product_name, discounted_price, product_link
FROM TMP
WHERE user_id_value = (
  SELECT user_id_value
  FROM TMP
  GROUP BY user_id_value
  ORDER BY COUNT(product_id) DESC, SUM(discounted_price) DESC
  LIMIT 1
)
ORDER BY discounted_price DESC;
```

### 4. 결과

- Query 실행 결과 : 모두 11개 상품 조회

| 행 | product_name | discounted_price | product_link |
|---|---|---|---|
| 1 | Zebronics Zeb-Companion 107 USB Wireless Keyboard and Mouse Set with Nano Receiver (Black) | 699.0 | https://www.amazon.in/Zebronics-Zeb-Companion-107-Wireless-Keyboard/dp/B087FXHB6J/ref=sr_1_48?qid=1672902996&s=computers&sr=1-48 |
| 2 | FLiX (Beetel) 3in1 (Type C\|Micro\|Iphone Lightening) Textured Pattern 3A Fast Charging Cable with QC & PD Support for Type C,Micro USB & Lightning Iphone Cable,Made in India,1.5 Meter Long | 182.0 | https://www.amazon.in/FLiX-Textured-charging-Lightning-Smartphones/dp/B08N1WL9XW/ref=sr_1_293?qid=1672909139&s=electronics&sr=1-293 |
| 3 | FLiX (Beetel) USB to Type C PVC Data Sync & 2A Smartphone Fast Charging Cable, Made in India, 480Mbps Data Sync, Tough Cable, 1 Meter Long USB Cable for USB Type C Devices Black XCD-C12 | 139.0 | https://www.amazon.in/Beetel-Smartphone-Charging-480Mbps-Xcd-C12/dp/B09NL4DJ2Z/ref=sr_1_170?qid=1672895799&s=electronics&sr=1-170 |

더보기

페이지당 결과 수: 50 1 – 11 (전체 11행)

### 5. 참고자료

-

---

## 추천시스템 4

### 1. 추천시스템의 이름

- Category-Wise Top-Rated Recommender

### 2. 추천시스템의 테마 및 목적

- 제품 카테고리별로 고객의 평점이 가장 높은 상위 5개 제품을 추천합니다.

### 3. 구현 로직 (코드, 쿼리 및 구조도)

- 구현 process
  - UNNEST() 함수를 'category' 컬럼에 적용하여 세부 카테고리 값을 가진 개별 데이터행 분리
  - GROUP BY와 ROW_NUMBER 함수를 이용해서 세부 카테고리별로 평균 rating  상위 5개 상품 정보(product_id)를 추출 → 임시 테이블(res5) 생성
  - (LookerStudio) 임시 테이블(res5)와 raw_data를 **데이터 혼합(LEFT JOIN)**하여 "조인테이블_5" 생성

- **res5** 데이트 테이블을 이용하여 컨트롤(드롭다운 목록) 생성
- **조인테이블5** 데이터 테이블을 이용해서 드롭다운 목록에서 선택된 세부 카테고리의 상품 정보 출력

- Bigquery SQL 문

```
WITH TMP AS(
  SELECT *,
    CAST(rating AS FLOAT64) AS float_rating
  FROM modulabs_amazon.data,
    UNNEST(SPLIT(category, '|')) AS category_value
),
TMP_ranked AS (
  SELECT category_value, product_id, AVG(float_rating) AS rating_avg,
    ROW_NUMBER() OVER (PARTITION BY category_value ORDER BY AVG(float_rating) DESC) AS R_number
  FROM TMP
  GROUP BY category_value, product_id
)

SELECT DISTINCT category_value, product_id, rating_avg, R_number
FROM TMP_ranked
WHERE R_number <= 5
ORDER BY category_value, rating_avg DESC;
```

-

## 4. 결과

- res5 테이블 조회 결과

쿼리 결과

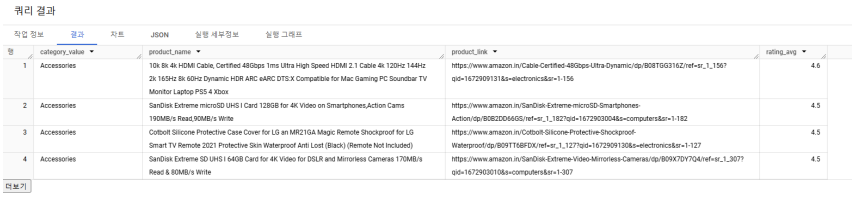
결과 저장 다음에서 열기

작업 정보 결과 차트 JSON 실행 세부정보 실행 그래프

| 행 | category_value | product_id | rating_avg | R_number |
|---|---|---|---|---|
| 1 | 3DGlasses | B097JVLW3L | 3.5 | 1 |
| 2 | AVReceivers&Amplifiers | B0978V2CP6 | 4.0 | 1 |
| 3 | Accessories | B08TGG316Z | 4.6 | 1 |
| 4 | Accessories | B0B2DD66GS | 4.5 | 3 |
| 5 | Accessories | B09TT6BFDX | 4.5 | 5 |
| 6 | Accessories | B09X7DY7Q4 | 4.5 | 2 |
| 7 | Accessories | B00R1P3B4O | 4.5 | 4 |
| 8 | Accessories&Peripherals | B0BP7XLX48 | 5.0 | 3 |

페이지당 결과 수: 50 1 – 50 (전체 991행)

- 조인 테이블(res5 & raw_data)에서 category_value가 " Accessories"인 데이터 출력 결과



쿼리 결과

작업 정보 결과 차트 JSON 실행 세부정보 실행 그래프

| 행 | category_value | product_name | product_link | rating_avg |
|---|---|---|---|---|
| 1 | Accessories | 10k 8k 4k HDMI Cable, Certified 48Gbps 1ms Ultra High Speed HDMI 2.1 Cable 4k 120Hz 144Hz 2k 165Hz 8k 60Hz Dynamic HDR ARC eARC DTS:X Compatible for Mac Gaming PC Soundbar TV Monitor Laptop PS5 4 Xbox | https://www.amazon.in/Cable-Certified-48Gbps-Ultra-Dynamic/dp/B08TGG316Z/ref=sr_1_156?qid=1672909131&s=electronics&sr=1-156 | 4.6 |
| 2 | Accessories | SanDisk Extreme microSD UHS I Card 128GB for 4K Video on Smartphones,Action Cams 190MB/s Read,90MB/s Write | https://www.amazon.in/SanDisk-Extreme-microSD-Smartphones-Action/dp/B0B2DD66GS/ref=sr_1_182?qid=1672903004&s=computers&sr=1-182 | 4.5 |
| 3 | Accessories | Cotbolt Silicone Protective Case Cover for LG an MR21GA Magic Remote Shockproof for LG Smart TV Remote 2021 Protective Skin Waterproof Anti Lost (Black) (Remote Not Included) | https://www.amazon.in/Cotbolt-Silicone-Protective-Shockproof-Waterproof/dp/B09TT6BFDX/ref=sr_1_127?qid=1672909130&s=electronics&sr=1-127 | 4.5 |
| 4 | Accessories | SanDisk Extreme SD UHS I 64GB Card for 4K Video for DSLR and Mirrorless Cameras 170MB/s Read & 80MB/s Write | https://www.amazon.in/SanDisk-Extreme-Video-Mirrorless-Cameras/dp/B09X7DY7Q4/ref=sr_1_307?qid=1672903010&s=computers&sr=1-307 | 4.5 |

더보기

- Bigquery SQL 문

```
SELECT product_name, discounted_price, product_link, positive_score, negative_score, total_score, float
FROM `modulabs_amazon.combined_data`
ORDER BY total_score DESC, float_rating DESC
LIMIT 20;
```

### 5. 참고자료

- 

---

## 추천시스템 5

### 1. 추천시스템의 이름

- Sentiment-Driven Recommender

### 2. 추천시스템의 테마 및 목적

- 고객들이 작성한 리뷰 테스트 데이터를 활용한 감정 분석을 통해 단순한 평점이나 판매량 이외에 보다 정교한 추천 시스템을 구현.

### 3. 구현 로직 (코드, 쿼리 및 구조도)

- raw_data의 “review_content” 컬럼에 있는 데이터를 Chatgpt로 이용해서 긍정/부정/종합점수 분석을 진행함. (결과는 csv 파일로 다운로드

첨부된 csv 파일에서 "review_content" 컬럼에는 위에서 분석한 것과 같은 리뷰 데이터가 들어 있습니다.
"review_content" 컬럼에 있는 리뷰 텍스트 데이터를 하나의 리뷰 데이터라고 간주하고,
그 리뷰 데이터에 대해서 긍정 점수와 부정 점수를 측정하고 종합 점수를 계산해서 별도의 csv 파일에 저장해 줘. 저장하는 csv 파일에는
* product_id
* 리뷰 긍정 점수
* 리뷰 부정 점수
* 리뷰 종합 점수
가 들어 있어야 함.

Analyzed

Analyzed

리뷰 데이터 분석이 완료되었습니다. 결과는 CSV 파일로 저장되었으며, 아래 링크를 통해 다운로드할 수 있습니다:

다운로드: analyzed_reviews.csv

*리뷰 텍스트 데이터에 대해서 정확한 감정 분석을 진행하기는 힘들지만, 감정 분석 결과에 대한 샘플 데이터를 얻는 다는 것에 만족을 할 수 밖에 없을 듯 합니다.*

- 위에서 얻은 리뷰 점수 데이터를 Bigquery에 테이블로 추가하고, raw_data와 조인을 해서 하나의 테이블로 생성함. (combined_data)

### 4. 결과

- Bigquery 조회 결과 : total_score DESC, float_rating DESC / LIMIT 20

쿼리 결과

결과 저장 다음에서 열기

작업 정보 결과 차트 JSON 실행 세부정보 실행 그래프

| 행 | product_name | discounted_price | product_link | positive_score | negative_score | total_score | float_rating |
|---|---|---|---|---|---|---|---|
| 1 | Amazon Basics Wireless Mouse \| 2.4 GHz Connection, 1600 DPI \| Type - C Adapter \| Upto 12 Months of Battery Life \| | 499.0 | https://www.amazon.in/Wireless-Connection-Battery-Ambidextrous-Suitable/dp/B09ZHCJDP1/ref=sr_1_209? | 5 | 0 | 5 | 5.0 |
| 2 | Melbon VM-905 2000-Watt Room Heater (ISI Certified, White Color) Ideal Electric Fan Heater for Small to Medium Room/Area (Plastic Body) | 998.0 | https://www.amazon.in/Melbon-Blower-Heater-2000-Watt-White/dp/B09P1MFKG1/ref=sr_1_313? | 5 | 0 | 5 | 4.6 |
| 3 | VRPRIME Lint Roller Lint | 499.0 | https://www.amazon.in/VR | 5 | 0 | 5 | 4.6 |

페이지당 결과 수: 50 1 – 20 (전체 20행)

### 5. 참고자료

- 

---

## 회고

- 인사이트
  - 기본적으로 단순한 raw data 이지만, 이런 저런 고민을 하면 좋은 추천 시스템에 대한 아이디어를 만들어 낼 수 있을 듯 합니다.
  - Looker Studio의 “컨트롤 - 드롭다운 목록” 기능을 활용하여 필터를 적용할 수 있음.

- 회고
  - 고객의 입장에서 어떤 추천 로직이 필요한지에 대해서 고민을 할 수 있는 좋은 프로젝트 였다고 생각합니다.
  - LookerStudio의 고급 기능 등을 좀 더 배워야 할 듯 합니다..