

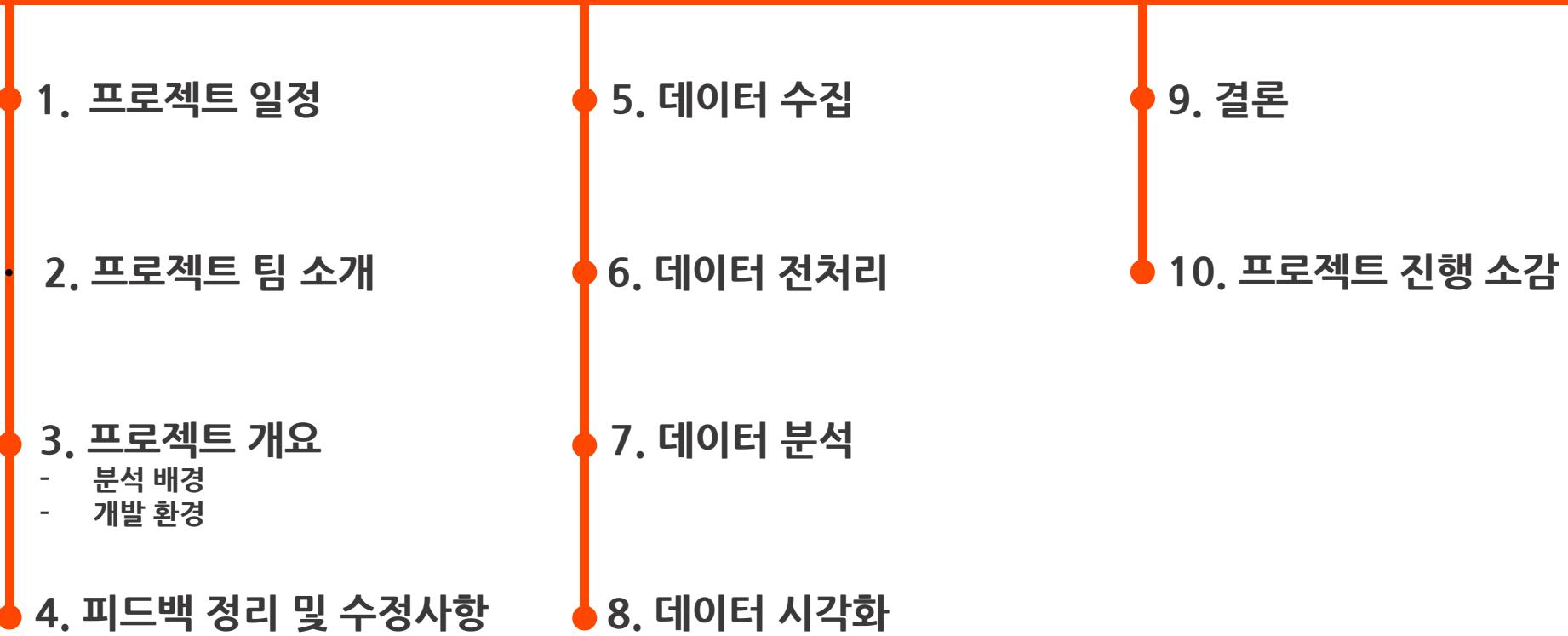


김건희, 김민석, 김준형, 오세랑, 이중호

29cm 식단개선

(29cm 매출 성장을 위한 소스 다각화)

목차

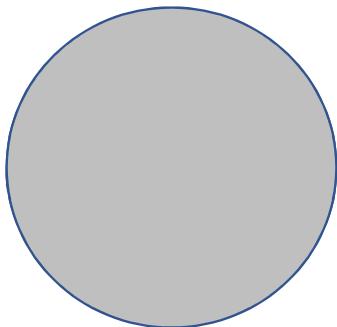
- 
1. 프로젝트 일정
 2. 프로젝트 팀 소개
 3. 프로젝트 개요
 - 분석 배경
 - 개발 환경
 4. 피드백 정리 및 수정사항
 5. 데이터 수집
 6. 데이터 전처리
 7. 데이터 분석
 8. 데이터 시각화
 9. 결론
 10. 프로젝트 진행 소감

1. 프로젝트 일정

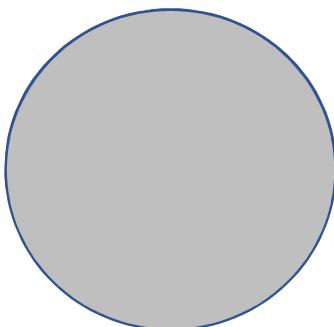
	1주차					2주차					3주차			
	5/9	5/10	5/11	5/12	5/13	5/16	5/17	5/18	5/19	5/20	5/23	5/24	5/25	5/26
사전 조사 및 주제 설정														
데이터 수집														
데이터 전처리														
데이터 분석														
데이터 시각화														
발표 준비 및 결론 도출														

2. 프로젝트 팀 소개

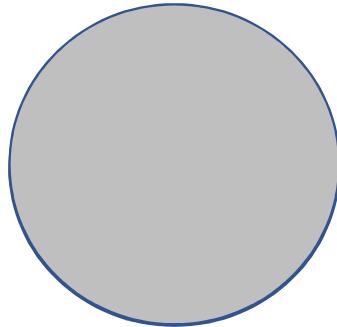
숨마쿰라우데



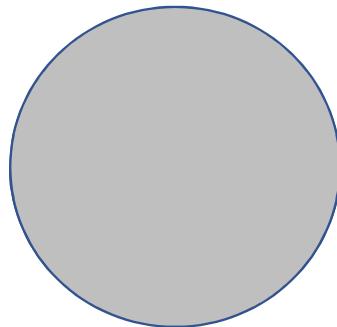
김건회
데이터 전처리
(Python)



김민석
크롤링
(Python)
발표



김준형
데이터 시각
(Tableau)



오세랑
데이터 시각화
(Tableau)



이중호
데이터 시각화
(Tableau)

2. 프로젝트 팀 소개

[본인의 역할]

1. 기획안에 맞추어 적용할 수 있는 핵심지표 제안 및 설정(RFM)
2. Python과 Tableau를 동시에 사용함으로 인해 같은 지표에 대해 세부식이 두 갈래로 나누어질 때 중요한 판단 요소를 확인 후 논리적으로 설득하여 하나의 수식으로 통합
3. R, F, M 각 지표에 대해 데이터를 분류한 후 분포를 확인하여 어떻게 나누면 좋을지 정성적 판단에 대한 기여
 - > RFM으로 Segment 분류를 함에 있어 매우 중요하고 시간을 많이 쏟았던 부분 (RFM 점수 몇 분위로 Segment를 몇 분류 나눌 것인지)
4. Tableau 내 Filter 중첩과 관련하여 구현이 안 되던 부분을 결국 해결함
 - > Context filter가 필요했던 상황에서 대시보드 작업상에는 해당 filter 없이 적용을 했어야 했는데 이를 VLOD 문제를 해결하고 결국 재구매율과 판매전환율을 한 그래프에 올릴 수 있게 되었음
5. 이외 데이터 탐색, 분석, 전처리, 커뮤니케이션 등



이중수

데이터 시각화
(Tableau)

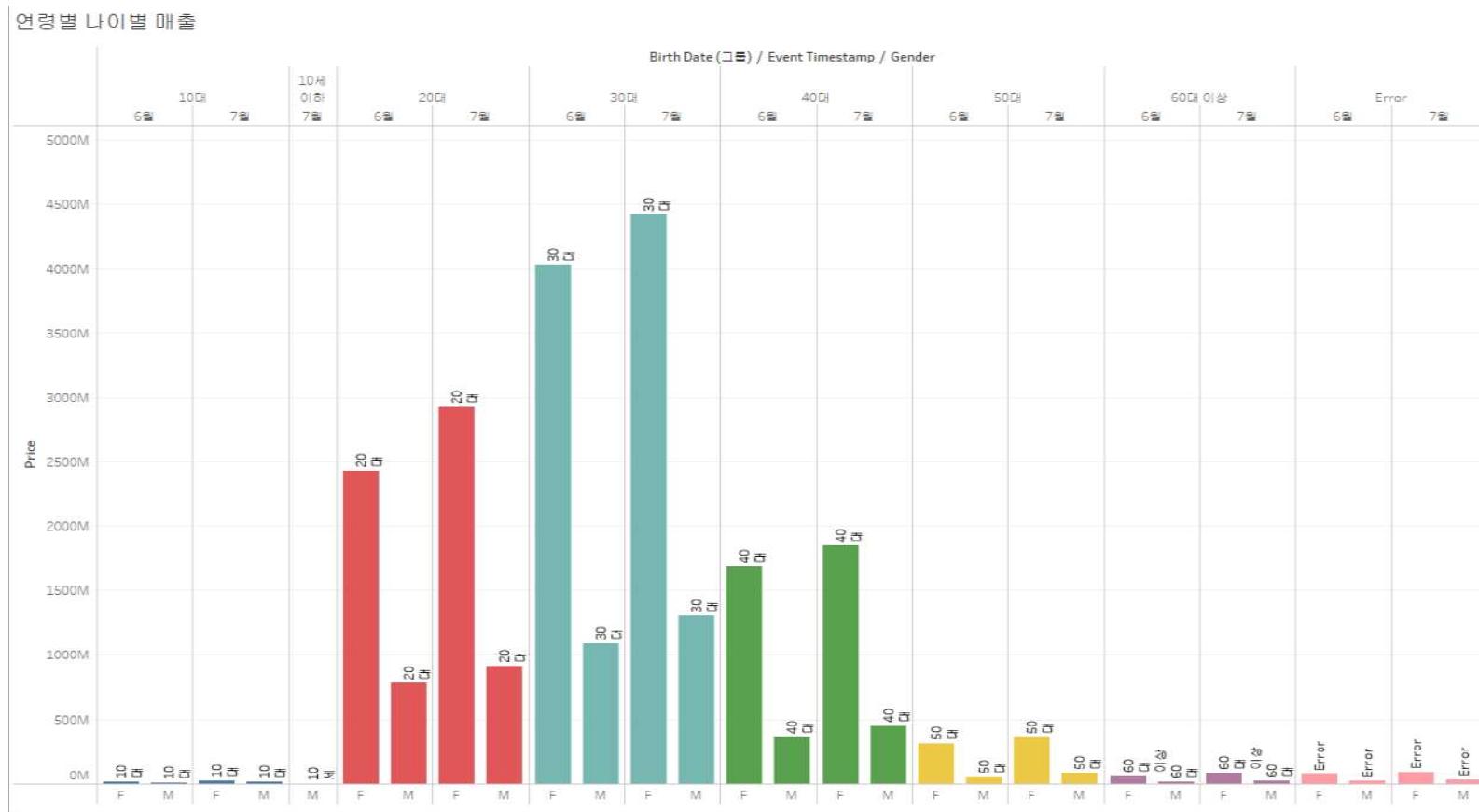
3. 프로젝트 개요

프로젝트 주제 선정 및 분석 배경

- 1) 프로젝트 주제 : 29cm 매출 다각화를 통한 성장
- 2) 주제 선정 및 분석 배경
 - User 분석 : 2030 여성에 매출 집중
 - Category 분석 : 의류, 잡화가 매출의 약 85%를 견인
 - Brand 분석 : 소수의 브랜드가 매출 대부분을 견인

3. 프로젝트 개요

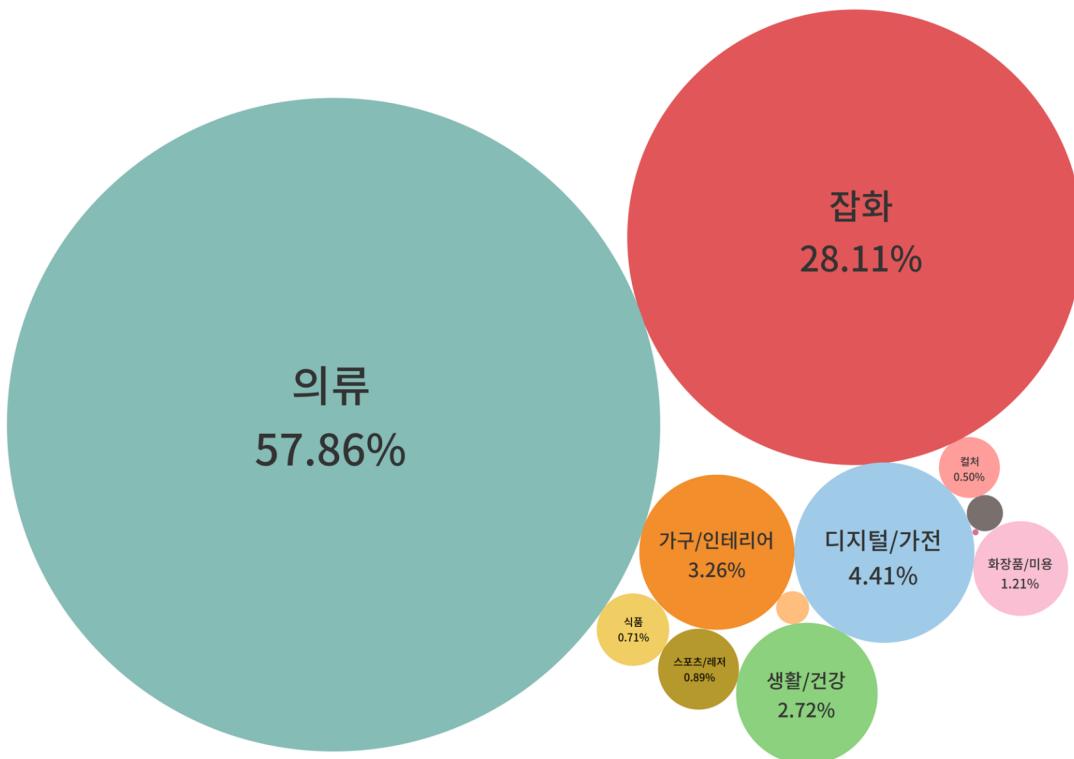
- User 분석 : 2030 여성에 매출 집중



Megabyte School 데이터 분석가 취업 완성 과정

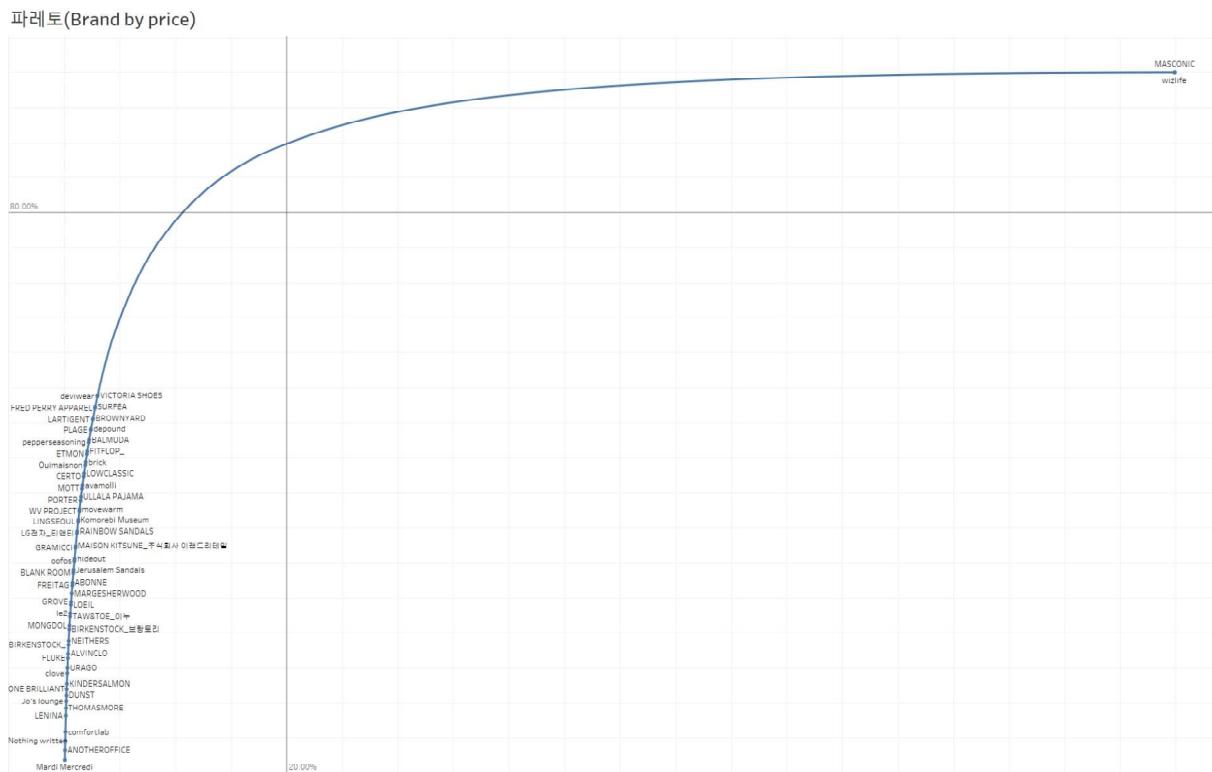
3. 프로젝트 개요

- Category 분석 : 의류, 잡화가 매출의 약 85%를 견인



3. 프로젝트 개요

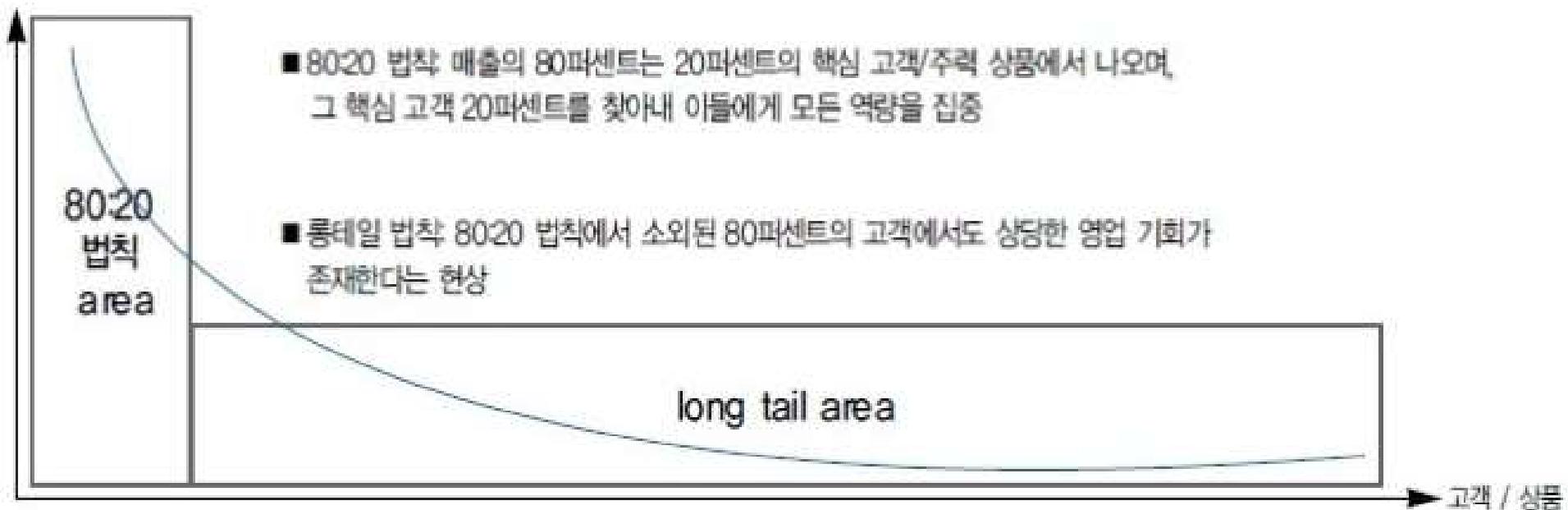
- Brand 분석 : 소수의 브랜드가 매출 대부분을 견인



3. 프로젝트 개요

현황 분석

- 상품 카테고리, 브랜드, 고객층 각각에 대해 **파레토 현상**이 두드러지게 나타남을 확인
- 온라인플랫폼에서 **롱 테일 현상**이 나타나는 점을 근거로 29cm의 매출 성장 가능성을 발견

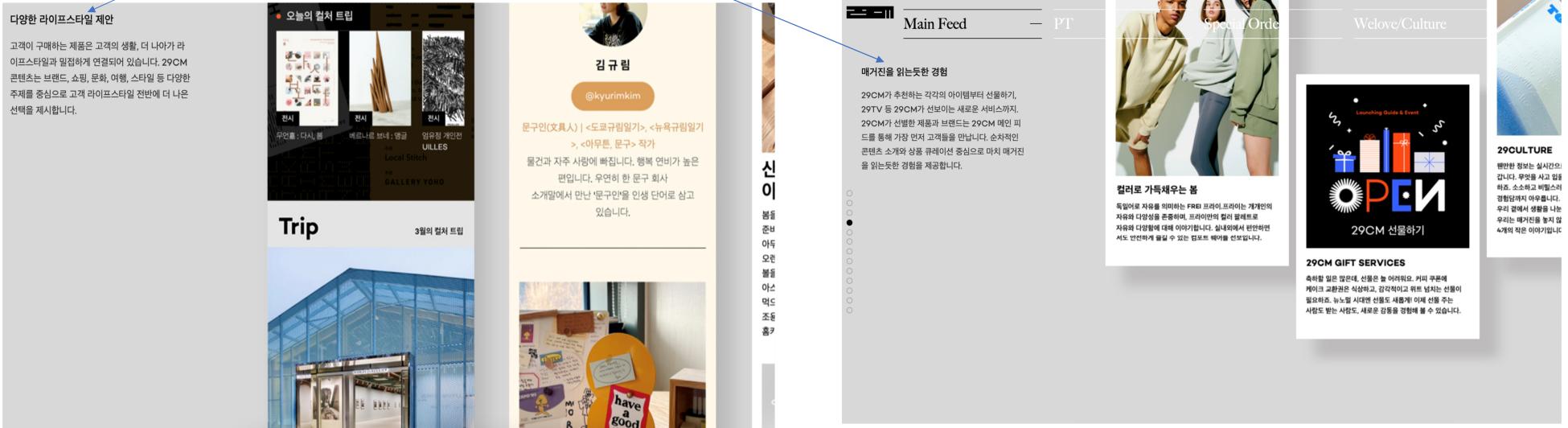


롱테일 관련 논문 <https://www.dbpia.co.kr/Journal/articleDetail?nodeId=NODE09451949>

3. 프로젝트 개요

프로젝트 기획 의도 및 인사이트

- 29cm의 특성을 파악, 회사 방향성에 적합한 **매출 다각화** 방안 제시
- 주 매출처인 카테고리 및 고객층이 아닌 부분에서 봉 테일 현상을 유도해낼 수 있는지를 탐구



3. 프로젝트 개요

개발 환경

Language	Library	Tool
Python	Numpy Pandas Time Seaborn Matplotlib Requests	Jupyter Notebook Tableau GitHub Notion Slack

4. 피드백 정리 및 수정사항

기획안 피드백 및 멘토링

1. 기획안 발표에 대한 피드백에 대한 수정 및 개발 사항

- 1) 전처리 관련 : Brand_Name 오류, 카테고리 오분류 → 파이썬 코드 및 정성적 분류로 해결
- 2) 비주력 카테고리 탐색 방법 : 외부데이터 활용 및 내부데이터 분석으로 지표 생성 및 판단
- 3) 마케팅 관련 비용 대비 수익 고민 : 단기적 비용 증가 < 장기적인 브랜드 성장에서 오는 이익

2. 멘토링을 통해 받은 피드백에 대한 수정 및 개발 사항

- 1) 비효율적인 코드 관련 개선 : for 문 → def, apply 문으로 해결
- 2) BrandName 오류 해결 : 규칙성 파악 후 비규칙적인 오류는 정성적인 판단 및 정규 표현식으로 수정
예시) 규칙성 - 언더바, 괄호 포함 등, 비규칙성 - 브랜드 이름이 'Null', 언더바 포함 등
- 3) 태블로 관련 : 이중축 / 계산된 필드 중첩 활용

5. 데이터 수집

데이터 수집 및 설명

- 내부 데이터 : 29cm GA기반 데이터 (events, products, users)

1. sampled_events.csv

- Event_name['purchase_success']에 대한 고유 ID값이 없어 어려움이 있었음

Session id	세션
Event Timestamp	이벤트 발생 시간
Event Name	이벤트 종류(클릭, 장바구니, 구매, 찜)
User No	사용자 고유번호
Item No	제품 고유번호
Device Type	사용자 핸드폰 타입
Mobile Brand Name	사용자 핸드폰 브랜드
Mobile Model Name	사용자 핸드폰 모델
Mobile Marketing Name	사용자 핸드폰 기종
Operating System Version	사용자 핸드폰 OS
Country	국가
Region	지역
Platform	사용자 핸드폰 OS

2. sampled_products.csv

- Category1, 2 종복 체크

Brand Name	브랜드 이름
Category1 Code	제품 대분류 코드
Category1 Name	제품 대분류 이름
Category2 Code	제품 중분류 코드
Category2 Name	제품 중분류 이름
Category3 Code	제품 소분류 코드
Category3 Name	제품 소분류 이름
Image Name	제품 사진
Item Name	제품 이름
Item No	제품 고유코드
Brand No	브랜드 고유코드
Price	가격

3. sampled_users

- Year(Birth_Date)가 8888 또는 9999인 경우는 Nan 처리

- Birth_Date의 Year 기준으로 작업을 진행하는 경우 외에는 제거하지 않음
(제거하기에는 매출 비중이 큼)

Birth Date	사용자 생년월일
Gender	사용자 성별
User No	사용자 고유코드

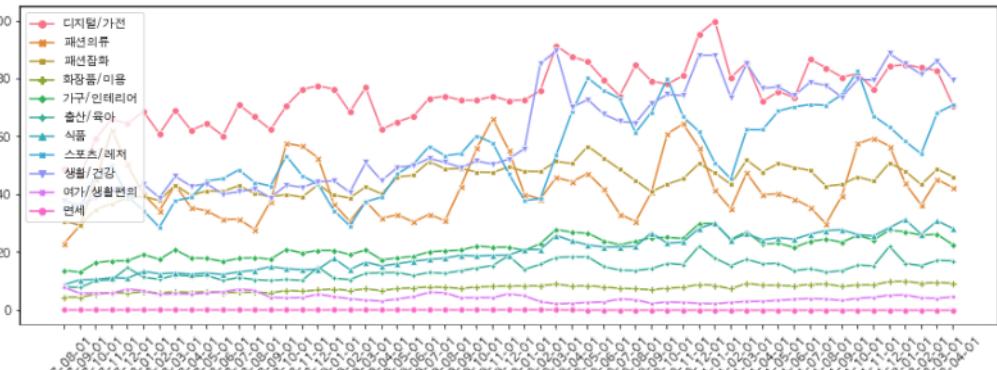
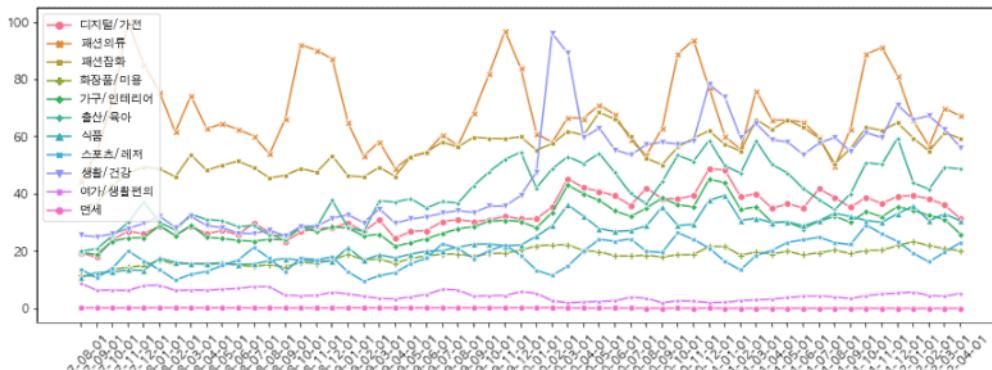
5. 데이터 수집

데이터 수집 및 설명

- 외부 데이터 : 네이버 쇼핑 API 활용을 통한 쇼핑 검색어 트렌드 (비주력 카테고리 유추 위한 데이터)

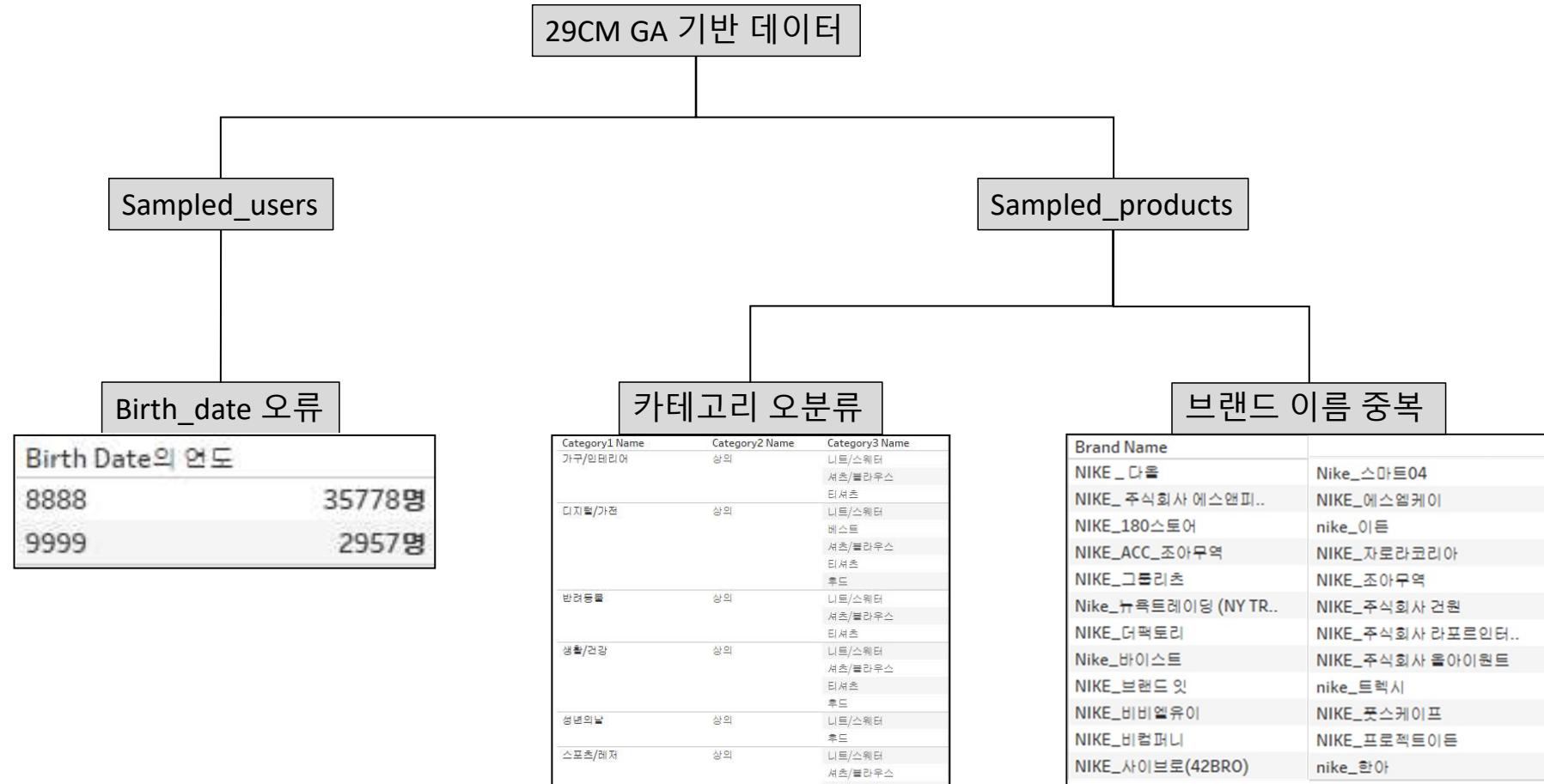
활용

- 2017/08 - 2022/04 기간 네이버 쇼핑 검색어 비율(최고점이 100)을 파악
- [남성] → [디지털/가전] 카테고리가 가장 많이 검색됨을 확인
- [의류], [스포츠/레저] → 계절성 파악
- [생활/건강], [스포츠/레저] → 코로나 시즌 변동 확인
(코로나로 [스포츠] 관련 급감 후 휴트 관련 제품 상승, [생활/건강]에서 마스크 및 위생 관련 제품 급증)



6. 데이터 전처리

원본데이터 문제 인식

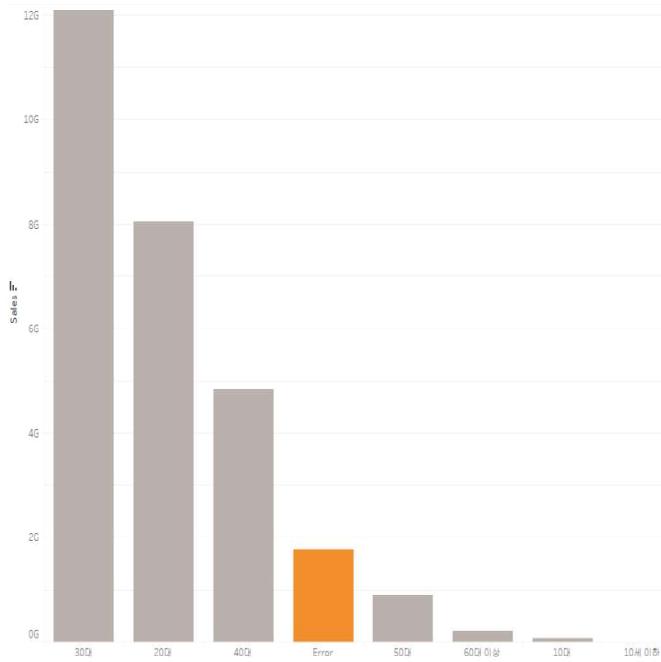


6. 데이터 전처리

1) Birth_date 전처리

Birth_Date : 9999,8888, Null 등의 생일은 Nan 처리 후 연령대 분석에 있어서는 제외하여 분석

세대별 매출



Error 전처리

Birth_date 시간 처리

```
In [45]: 1 #8888,9999, null 등의 데이터 nat로 사용  
2 df['birth_date'] = pd.to_datetime(df['birth_date'], errors = 'coerce')  
  
In [47]: 1 #생일 없는 경우  
2 df['birth_date'].isnull().sum()  
  
Out[47]: 335636  
  
In [48]: 1 def convert_datetime(row):  
2     if pd.isnull(row):  
3         return np.nan  
4     else:  
5         return pd.to_datetime(row)  
  
In [49]: 1 df['birth_date'] = df['birth_date'].apply(convert_datetime)  
  
In [ ]: 1 #df csv로  
2 df.to_csv('df.csv')  
  
In [50]: 1 df['birth_date'].describe()
```

Out[50]: count 5544771
unique 15790
top 1978-01-23 00:00:00
freq 8041
first 1923-12-12 00:00:00
last 2017-07-17 00:00:00
Name: birth_date, dtype: object

전처리 결과

10대	1,606
10세 이하	2
20대	88,347
30대	95,274
40대	34,770
50대	6,438
60대 이상	1,604
Error	26,917
총합계	254,958

6. 데이터 전처리

2) 카테고리 전처리

카테고리 오분류 : category2_name이 잘못된 category1_name 하위 카테고리로 들어가 있는 경우 전처리

전처리 코드

```
In [38]: 1 #클린 category 딕셔너리
2 dic = {'가공식품': '식품', '건강증정용품': '생활/건강', '계절가전': '디지털/가전', '공학기기/용품': ''
In [39]: 1 for key, value in dic.items():
2     df.loc[df['category2_name']==key, 'category1_name'] = value
In [40]: 1 #확인
2 df[df['category2_name']=='지갑']['category1_name'].unique()
Out[40]: array(['정화'], dtype=object)
In [ ]: 1 df[df['category1_name']=='의류']['category2_name'].unique()
In [41]: 1 #kids에 속하는 게 맞는 경우는 키즈로
2 kids = {'기방': '잡화', '하의': '의류', '주얼리': '잡화', '스포츠웨어': '의류', '신발': '정화', '아우터': '의류'}
In [42]: 1 df[df['category1_name']=='키즈']
...
In [43]: 1 for key, value in kids.items():
2     df.loc[(df['category2_name']==key) & (df['category1_name']!='키즈'), 'category1_n
In [44]: 1 #키즈 이미 확인
2 df[df['category1_name']=='키즈']
...
In [ ]: 1 #새 데이터 파일
2 df.to_csv('category_edited')
```

전처리 전

Category1 Name	Category2 Name	Category3 Name	의류	살의	나트/스웨터
가구/인테리어	상의	나트/스웨터			베스트
		셔츠/블라우스			셔츠/블라우스
		티셔츠			스웻셔츠
디지털/가전	상의	나트/스웨터			카디igan
		베스트			티셔츠
		셔츠/블라우스			hoodie
		티셔츠			
		후드			
반려동물	상의	나트/스웨터			나트/스웨터
		셔츠/블라우스			베스트
		티셔츠			셔츠/블라우스
		hoodie			티셔츠
생활/건강	상의	나트/스웨터			카디igan
		셔츠/블라우스			티셔츠
		티셔츠			hoodie
		후드			
성녀의날	상의	나트/스웨터			나트/스웨터
		후드			셔츠/블라우스
					스웻셔츠
					카디igan
					티셔츠
					hoodie
스포츠/레저	상의	나트/스웨터			
		셔츠/블라우스			
		티셔츠			
		hoodie			
식품	상의	베스트			나트/스웨터
		셔츠/블라우스			셔츠/블라우스
		카디igan			스웻셔츠
		티셔츠			조끼
					카디igan
					티셔츠
					hoodie

전처리 후

Category1..	Category2..	Category3 Name	의류	상의	나트/스웨터
					베스트
					셔츠/블라우스
					스웻셔츠
					카디igan
					티셔츠
					hoodie
					나트/스웨터
					셔츠/블라우스
					스웻셔츠
					카디igan
					티셔츠
					hoodie
					나트/스웨터
					셔츠/블라우스
					스웻셔츠
					조끼
					카디igan
					티셔츠
					hoodie

6. 데이터 전처리

3) 브랜드 전처리

브랜드 이름 : 중복된 브랜드이지만 유통사, 휴먼 에러 등의 문제로 전처리

전처리 코드

```
In [134]: 1 def underbar_all(x):
2     for i in ['_$', '_B', '_SDF', '_LS.*', '_[().*','_INC']:
3         p=re.compile(i)
4         underbar_only_ = p.findall(x)
5         if len(underbar_only_) > 0 :
6             return x.replace('_', '')
```



```
In [123]: 1 # brand_name_df.loc[(brand_name_df.brand_name.str.contains('_')) & (brand_name_df.ne
```



```
In [135]: 1 # brand_name_df.loc[(brand_name_df.brand_name.str.contains('_')) & (brand_name_df.ne
2 # brand_name_df.loc[(brand_name_df.brand_name.str.contains('_')) & (brand_name_df.ne
3 # brand_name_df.loc[(brand_name_df.brand_name.str.contains('_')) & (brand_name_df.ne
4 brand_name_df.loc[(brand_name_df.brand_name.str.contains('_')) & (brand_name_df.new_
```



```
In [136]: 1 brand_name_df.loc[brand_name_df.new_brand_name.isnull(),'new_brand_name']=brand_name
```

전처리 전

Brand Name	
NIKE_다음	Nike_스마트04
NIKE_주식회사 에스엔피..	NIKE_에스엘케이
NIKE_180스토어	nike_이든
NIKE_ACC_조아무역	NIKE_자로라코리아
NIKE_그辱리츠	NIKE_조아무역
Nike_뉴욕트레이딩 (NY TR..	NIKE_주식회사 건원
NIKE_더팩토리	NIKE_주식회사 라포르인터..
Nike_바이스트	NIKE_주식회사 블아이원트
NIKE_브랜드잇	nike_트렉시
NIKE_비비엘유이	NIKE_풋스케이프
NIKE_비컴퍼니	NIKE_프로젝트이든
NIKE_사이브로(42BRO)	nike_한아

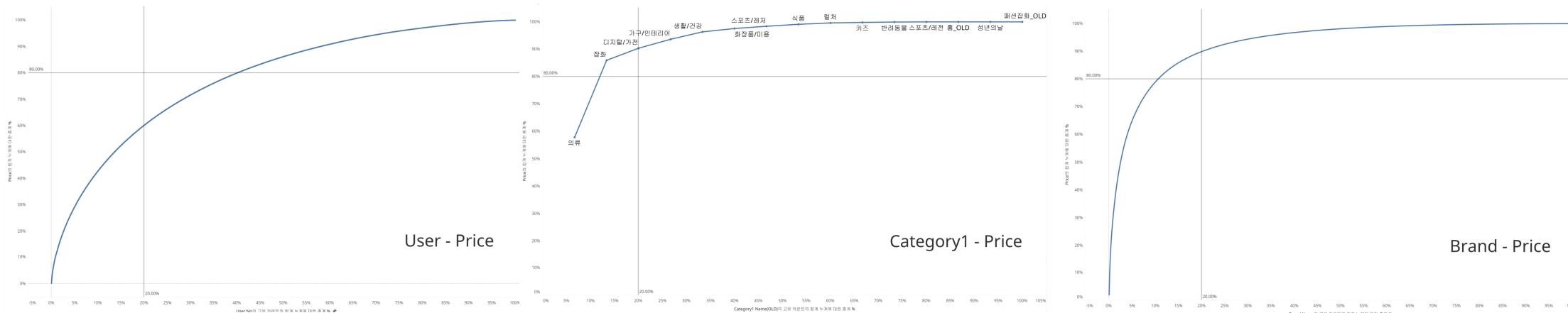
전처리 후

NIKE	22,607
총합계	22,607

7. 데이터 분석

1) 파레토 분석

- 고 객 : 상위 고객들이 존재하지만 파레토 현상이 나타나는 정도는 아님
- 카테고리 : 대부분류 및 중분류 모두 [의류], [잡화] 카테고리가 상위를 차지하며 **파레토 현상** 존재 확인
- 브랜드 : 상위 브랜드 소수가 매출 대부분을 견인하는 **파레토 현상** 존재 확인



⇒ ‘라이프 스타일’을 추구한다는 29cm가 소외된 다수(고객/카테고리/브랜드)에 대한 마케팅을 해야 함

7. 데이터 분석

2) 코호트 분석

- 가설

- 29CM는 매거진의 성향으로 활용되기 때문에 재방문율이 높을 것임

- 분석 결과

- 주어진 데이터 기간 내(2달)

고객 재방문율 $\rightarrow \{(254,958 - 98,317) / 254,958\} \approx 61\%$

- 7일 내 재방문율 = 37% *소수점 첫째자리 버림

⇒ 재방문율이 낮지 않음을 확인

코호트 1 Visited	Null	1	2	3	4	60	61	62	총합계
2021년 7월 1일	5,044	518	254	151	46				5,279
2021년 7월 2일	1,953	387	202	151	46				3,759
2021년 7월 3일	1,921	516	171	141	46				3,879
2021년 7월 4일	2,486	427	214	141	46				4,602
2021년 7월 5일	1,854	412	163	101	46				3,515
2021년 7월 6일	1,688	335	164	9	46				3,273
2021년 7월 7일	1,519	310	138	8	46				2,843
2021년 7월 8일	1,269	244	138	11	46				2,464
2021년 7월 9일	1,212	227	135	10	46				2,300
2021년 7월 10일	1,514	383	133	11	46				2,789
2021년 7월 11일	2,050	369	177	12	46				3,678
2021년 7월 12일	1,660	336	141	9	46				2,927
2021년 7월 13일	1,641	377	147	9	46				2,967
2021년 7월 14일	1,661	320	141	10	46				2,861
2021년 7월 15일	1,537	251	141	8	46				2,541
2021년 7월 16일	1,326	277	117	10	46				2,311
2021년 7월 17일	1,493	313	121	10	46				2,495
2021년 7월 18일	1,780	329	158	11	46				2,932
2021년 7월 19일	1,659	295	134	8	46				2,559
2021년 7월 20일	1,660	286	118	7	46				2,555
2021년 7월 21일	1,723	237	142	8	46				2,572
2021년 7월 22일	1,549	220	101	7	46				2,268
2021년 7월 23일	1,191	203	88	7	46				1,812
2021년 7월 24일	1,211	273	102	6	46				1,873
2021년 7월 25일	1,574	251	124	5	46				2,253
2021년 7월 26일	1,590	223	100	6	46				2,177
2021년 7월 27일	1,443	198	68	5	46				1,928
2021년 7월 28일	1,357	172	76	4	46				1,779
2021년 7월 29일	1,192	178	54	5	46				1,579
2021년 7월 30일	1,373	148	70	5	46				1,699
2021년 7월 31일	1,203	178	77	5	46				1,520
2021년 8월 1일	1,600	210	82	3	46				1,926
2021년 8월 2일	1,550	191	35	4	46				1,776
2021년 8월 3일	1,706	116							1,822
2021년 8월 4일	670								670
총합계	98,317	35,417	17,385	12,604	46	30	4	4	254,958

7. 데이터 분석

3) RFM 분석

- **Recency** : 고객이 얼마나 최근에 방문했는가
 - 08/04 (마지막 데이터 일자)와 가장 최근 구매일자와의 Datediff
- **Frequency** : 고객이 얼마나 자주 방문했는가
 - 구매 일자 건수
- **Monetary** : 고객이 얼마나 구매하였는가
 - 총 구매 금액

❖구매 event 기준

User No	RFM Recency	RFM Frequency	RFM Monetary
0/6pBZeSfIo3vbd+wr..	44	1	59,000
0/7cVYQ CtKEs/ChF4oV..	1	1	175,000
0/9VTm6s0PhPUuvSL2..	14	2	160,000
0//5YrgWlnKdUcHtZaA..	35	1	109,000
0//mBlrBxjYBUGCs6ztO..	62	1	109,000
0/+vEOhv6Dqj8eaIOM..	4	16	1,289,000
0/AQ1rw5KSBPY85Xko..	59	1	175,000
0/aTwtN2l13dHyXTMr..	60	1	42,000
0/biEln9CiKR5dMaDIfa..	5	2	152,300
0/BXUFmUppArEfhimB..	28	3	208,000
0/eqTzvW4UB+NYg2lu..	9	4	533,000
0/ev8fCuloviNEs0pAI4..	43	1	13,000
0/ggn+8rIrUG/rDeqS..	63	1	124,000
0/h7KW/1ej37lzbjqOH..	1	7	2,618,100
0/iGikvRaJS8jXzEBqL4O..	6	1	88,000
0/iXnasUjjXp2KiP+BuLe..	23	1	43,000
0/jR2VGsGuqrKpmhC0l..	3	3	463,000
0/mBsCnehiLD1LSEu6..	26	3	47,900
0/N0/Do4yfij/tmDzju0..	36	2	394,000

7. 데이터 분석

3) RFM 분석

Segment 선정 과정 : RFM 각 점수를 통해 고객 세그먼트

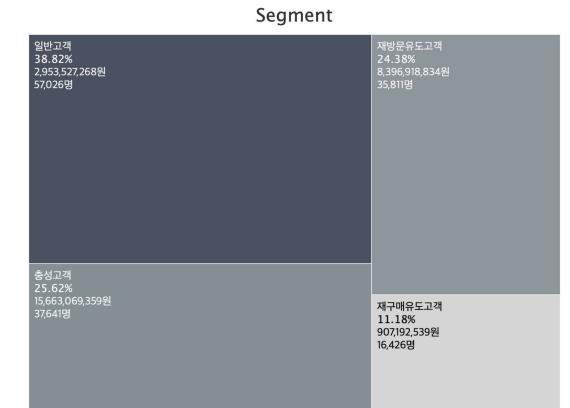
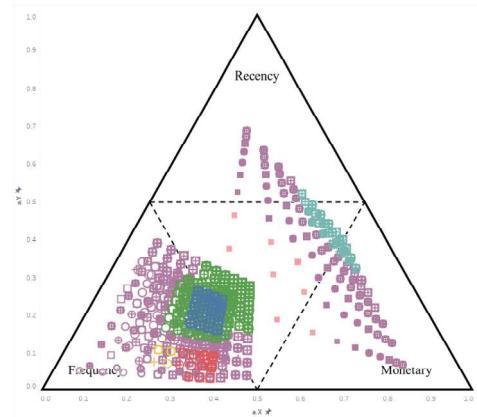
- 4분위 점수(8분류) ◀ 고매출 유저 다수, 세분화 점수 부여 필요
- 10분위 (8분류) ◀ Frequency 점수 분포 문제 (삼각도표 → 추후 User 개선 지표로 사용)
- 4분위(8분류) ◀ 과도한 세분화 → 심각한 소규모 인원 세그먼트 발생 → 프로젝트 목적에 따라 분류 단순화 필요
- ↓ - 4분위(4분류) ◀ 분석에서 주로 사용

4분위 점수
(8분류)

10분위 점수
(8분류)

4분위 점수
(8분류)

4분위 점수
(4분류)



7. 데이터 분석

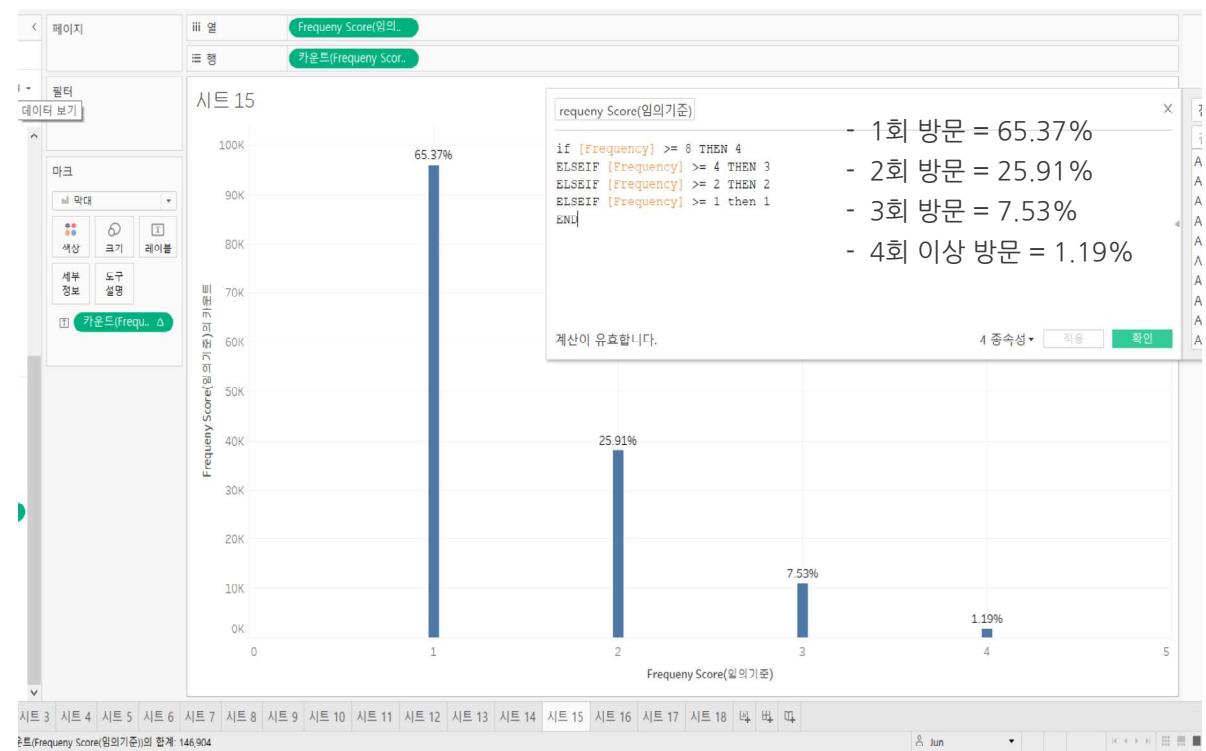
* RFM 분석에서의 가장 큰 걸림돌 → ‘Frequency’

- 주어진 데이터의 기간이 2달치임에 따라 Frequency의 분포가 우측 그림과 같았고,

어떤 방식으로도 4분위 위로는 구분 불가하여 RFM 모든 수치를 4분위로 나누고

customer segment는 기존의 분류가 아닌
고객 특성에 대한 대응을 위한 군집으로 분류함

- 충성고객
- 재구매필요고객
- 재방문필요고객
- 일반고객



7. 데이터 분석

4) 구매전환율, 재구매율

- 구매전환율 = (기준 별 구매 event / 총 event)
- 재구매율 = (기준 별 2회 이상 구매 유저 / 구매이력이 있는 유저)
 - ❖ Category1,2 ,Brand 기준

◦ 유저 예외 사항 및 사유

- 구매전환율 : 소수의 event를 발생시키면서
구매 event만 존재하는 고객 다수 존재

- 재구매율 : Frequency 와 의미 중첩

+) 2030 남성 분석 : 재구매율의 경우 구매한 user가 적은 category2를
제외하기 위해 구매자가 20명 이상인 category2만 표현

시트 34 (2)

User No	Category2 Name	Event Timesta..	Session Id	
0/7cVYQCtKEs/ChF4oVz4A..	상의	2021년 8월 4일	70777388caf7c1ad7f5c3e9..	0
0//mBlrBxjY8UGCs6ztOTA..	모자	2021년 6월 4일	63c031e987b587ec821a47..	0
0/aTwtN2I13dHyXTMraNj..	상의	2021년 6월 6일	5345cdf98f50b6b010bd81..	0
0/BXUfMUpArEfhimBug..	원피스/诘프수트	2021년 7월 8일	cbb66d97b2fdb6ecee992e..	0
티켓		2021년 6월 18일	4549427ab501fca3449d64..	1
		2021년 6월 22일	715d91778d743b59134e3c..	1
0/iXnasUjjXp2KiP+BuLeQ==	가방	2021년 7월 13일	af5604111beecd28211b2bc..	0
0/jR2VGsGuqrKpmhC0ISc..	가방	2021년 7월 20일	bd60b7054e3090a6a7b3fc..	0
	건강식품	2021년 7월 20일	bd60b7054e3090a6a7b3fc..	0
	상의	2021년 7월 12일	8dcf0bb99b636e860943bc..	0
	신발	2021년 8월 2일	56b65de416c2c508f3b532..	0
	하의	2021년 7월 12일	8dcf0bb99b636e860943bc..	0
0/mBsCnehiLD1LSEu6dfiA..	바디케어	2021년 7월 2일	3096f517f6ef05f0ad1b565..	0
	목실품	2021년 6월 18일	977f12bce26cd2997d1610..	0
	휴대폰 악세서리	2021년 7월 10일	e590b746282f6de6fffb68..	0

8. 데이터 시각화

Tableau 대시보드 전체 보기



8. 데이터 시각화

1. 29CM 현황분석

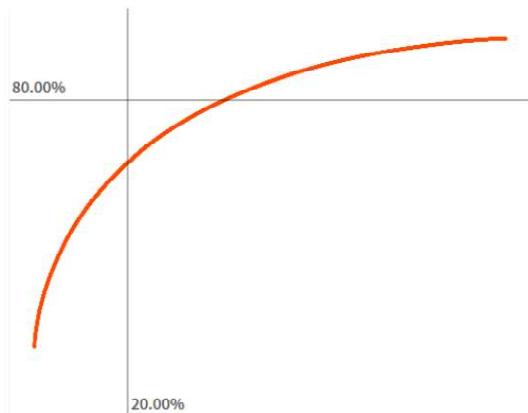


8. 데이터 시각화

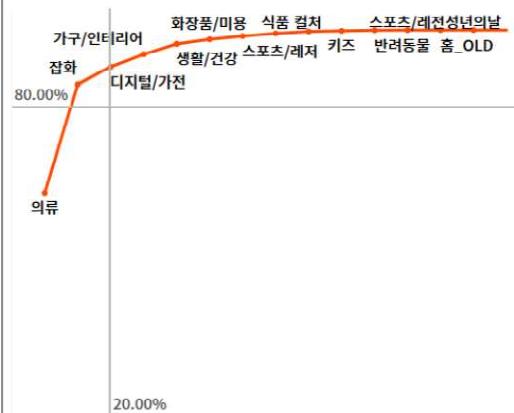
2. 29CM 문제정의

현황분석 문제정의 User Category Brand

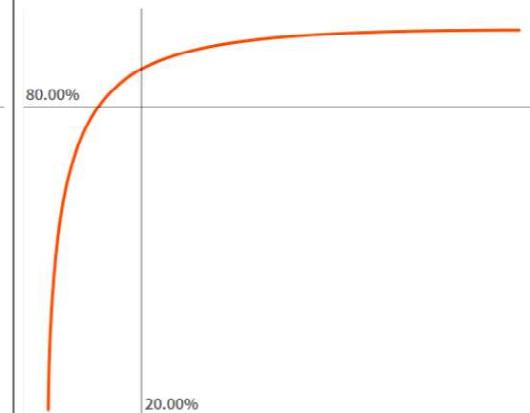
파레토(User by price)



파레토(Category by price)



파레토(Brand by price)



매출 비중(나이-성별)



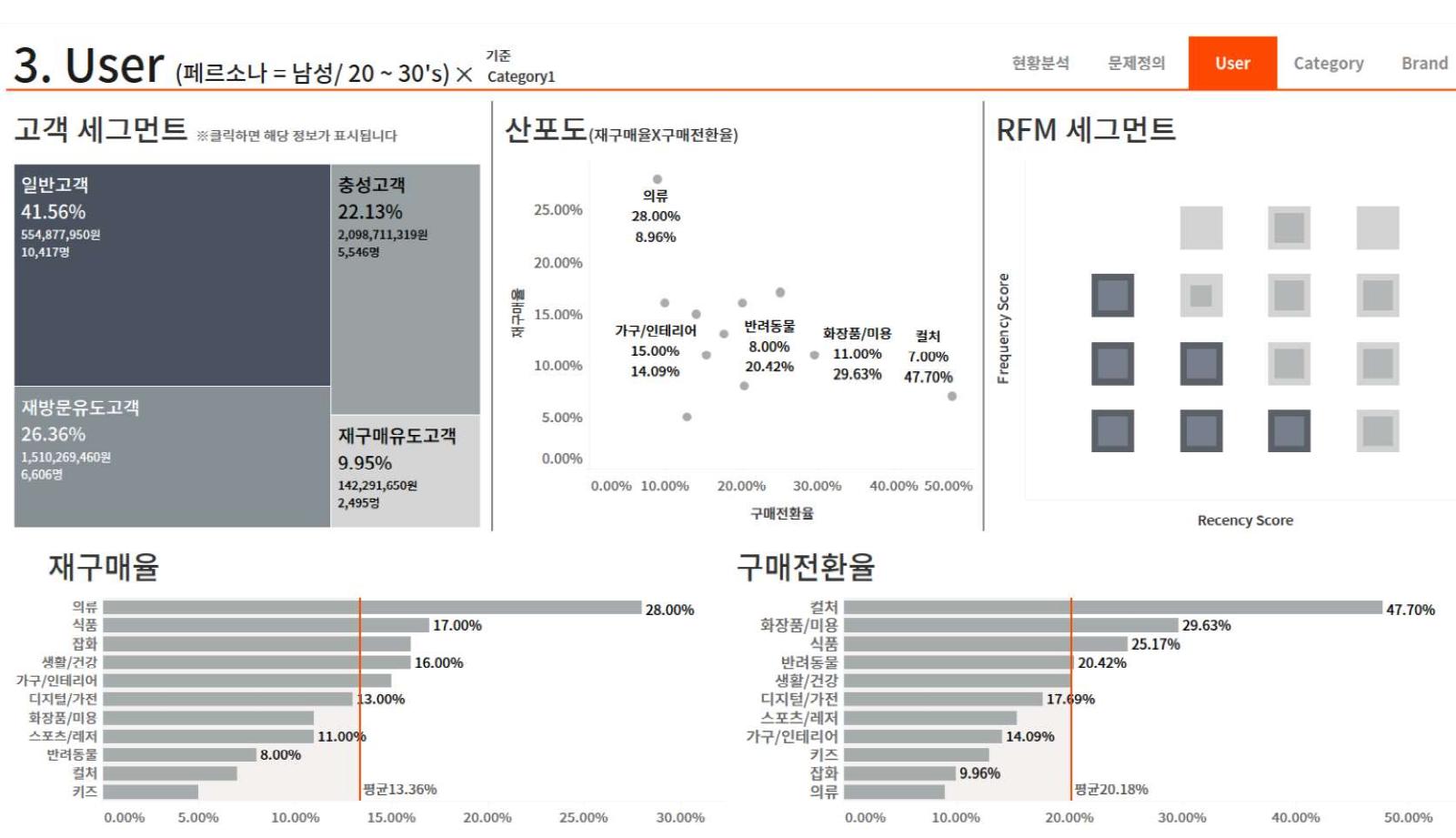
매출 순위(by Category)

1위	의류	57.86%
2위	잡화	28.11%
3위	디지털/가전	4.41%
4위	가구/인테리어	
5위	생활/건강	
6위	화장품/미용	
7위	스포츠/레저	
8위	식품	
9위	컬처	
10위	키즈	

매출 순위(by Brand)

1 위	MARDI MERCREDI	14.612%
2 위	ANOTHEROFFICE	10.711%
3 위	NOTHING WRITTEN	9.791%
4 위	COMFORTLAB	
5 위	FACADE PATTERN	
6 위	LENINA	
7 위	BIRKENSTOCK	
8 위	THOMASMORE	
9 위	JO'S LOUNGE	
10 위	DUNST	

8. 데이터 시각화



8. 데이터 시각화

4. CATEGORY (의류,잡화 제외)

현황분석 문제정의 User Category Brand

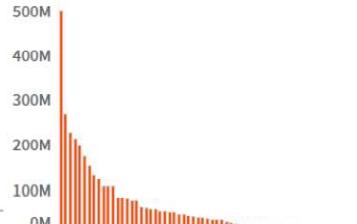
카테고리 산포도

※클릭하면 해당 정보가 표시됩니다

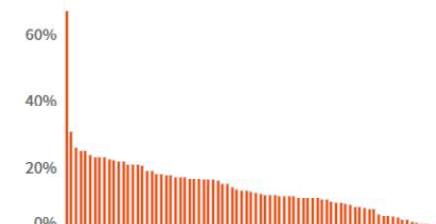


캠핑

매출
858,515,190원
4위



구매전환율
11.06%
44위



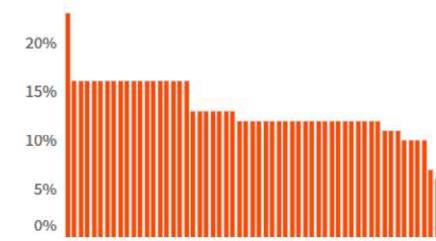
이벤트수

168,738건
6위

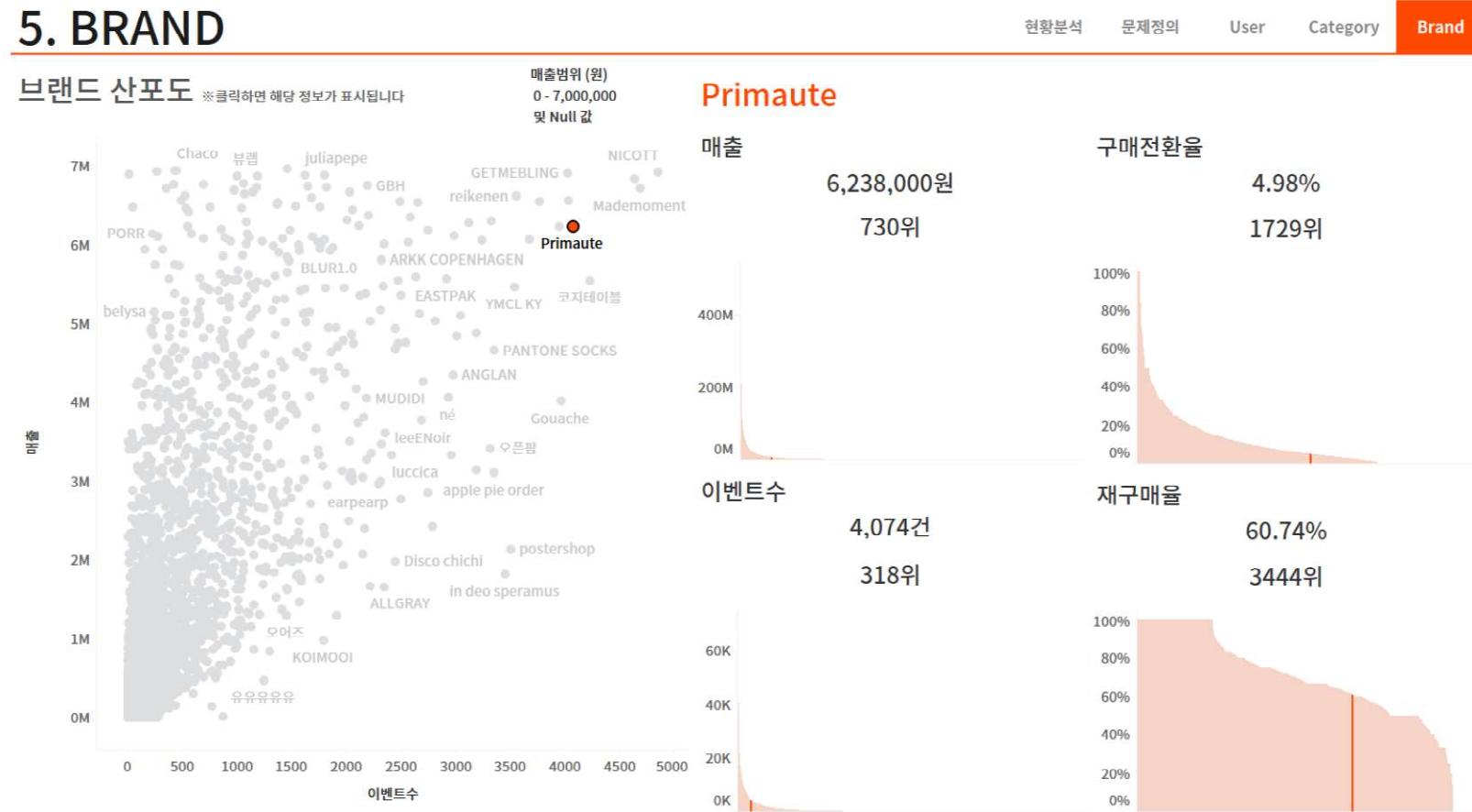


재구매율

38.0%
54위



8. 데이터 시각화



9. 결론

User

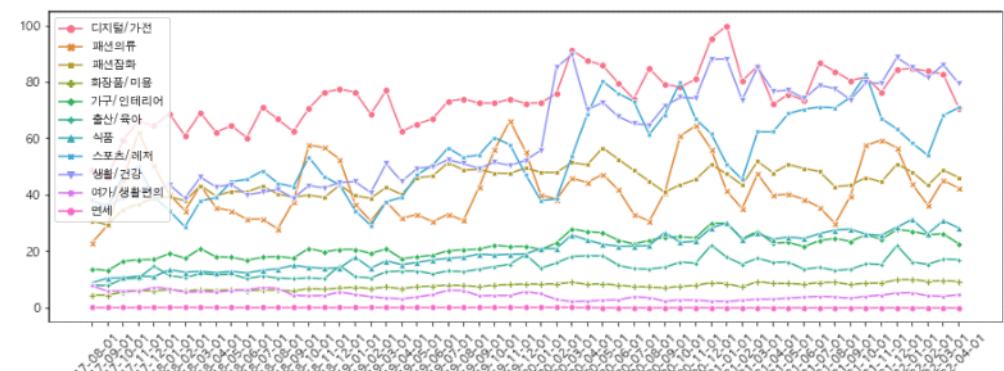
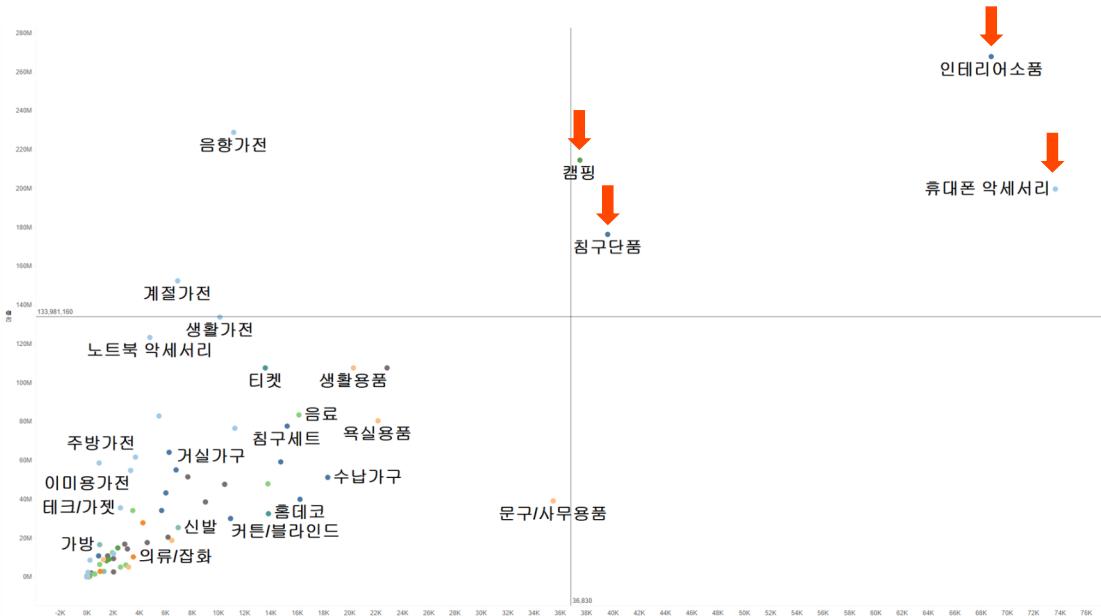
⇒ 세그먼트 별 지표를 통해 특성을 확인

- 충성고객 ◀ 전체적으로 우수한 지표, 다만 한 지표로 치우쳐진 카테고리의 경우
미진한 부분을 높여줄 수 있는 마케팅
- 재구매 유도고객 ◀ 방문은 찾으나 구매금액이 낮은 고객에 대해서는 ‘디지털/가전’처럼
가격이 높은 카테고리에 대한 광고 노출로 매출 상승 유도
- 재방문 유도고객 ◀ 구매 금액은 큰 고객이기 때문에 재방문 유도로 매출 상승의 기회.
구매 금액에 따른 할인 쿠폰 추가 제공 등으로 재방문 및 충성고객으로 이어질 수 있도록 유도
- 일반고객 ◀ 전체적으로 낮은 지표, 재구매율이 특히나 낮은 모습.
재방문을 유도하기 위한 첫 방문 쿠폰, 메일 알림 등의 마케팅 더욱 활성화

9. 결론

Category

⇒ 매출, 이벤트 수 상위지표인 1사분면 위 4개의 카테고리 공략

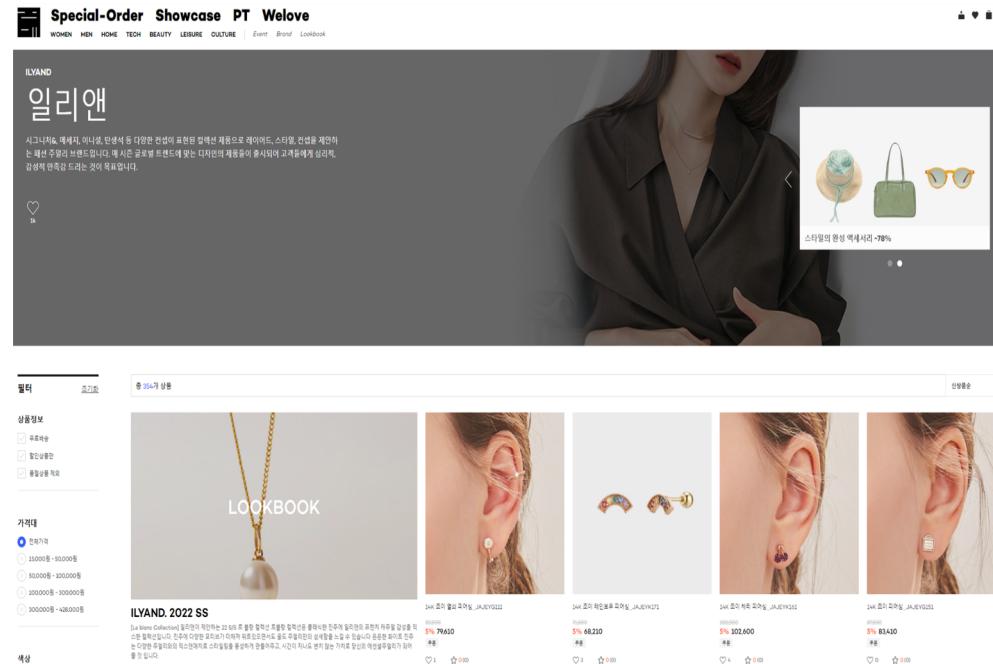


❖ 외부 데이터를 통해 카테고리 트렌드 검증

9. 결론

Brand

일리엔

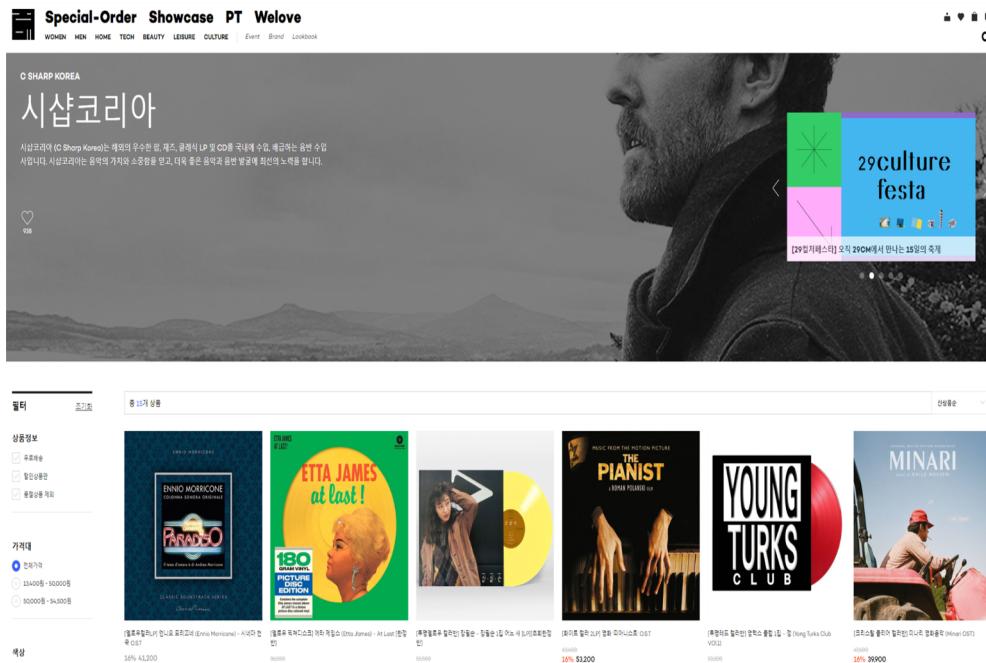


- 가장 높은 매출의 상품 → 소비자 탄생석 관련 (선물로 추정)
- 그 다음은 연예인 착용 상품
- 일반 상품들은 증정 상품이 있는 제품들의 매출이 큼
- 위의 세 가지를 바탕으로 **단순 쥬얼리 상품 사진만 올리기 보다는 연예인 착용샷 및 증정 이벤트 쪽으로 추천**

9. 결론

Brand

시샵코리아



- 높은 매출에 비해 Event 수 부족
- LP를 테마별로 테마관을 만들어 분류 추천
(영화ost, 싱어송라이터, 팝 종류가 무분별하게 섞여있음)
- 데이터에서 분석 결과로는 국내 싱어송라이터 보다는
매출이 높은 팝과 영화OST 쪽으로 컨텐츠 확장 추천

9. 결론

Brand

로우 클래식

The screenshot shows the homepage of the Low Classic website. At the top, there's a navigation bar with links for Special-Order, Showcase, PT, Welove, WOMEN, MEN, HOME, TECH, BEAUTY, LEISURE, CULTURE, Event, Brand, and Lookbook. Below the navigation is a section titled 'LOW CLASSIC' featuring a large image of a woman in a tan blazer. To the right of the main image is a smaller inset showing a woman holding a white tote bag with the text 'Darkness and Light'. On the left side of the main image, there's descriptive text about the brand's focus on classic styles and quality. Below the main image is a 'LOOKBOOK' section with three female models. At the bottom of the page is a product grid displaying five items: '20S STITCH SHIRT_LIGHT BLUE', '20S STITCH SHIRT_DEEP BIEGE', '20S DOUBLE VEST_NORV', and '20S DOUBLE VEST_BLUE GREEN'. Each item has its price (e.g., 215,000), color, and a small image.

- 클래식한 의류, 가방, 신발, 액세서리 제품 판매
- 제품의 구성, 수, 질이 모두 좋지만 상품 연계가 안 됨
- 옷을 살 때 모델이 같이 신은 신발 및 가방 정보가 연계되는 시스템 추천

9. 결론

Brand

소미노

The screenshot shows a product showcase for PT Welove. At the top, there's a navigation bar with categories: WOMEN, MEN, HOME, TECH, BEAUTY, LEISURE, CULTURE, Event, Brand, Lookbook. Below the navigation is a banner for 'SOMINO' featuring a person in a white shirt and a small inset image of various health products. The main content area displays six product cards with images and prices:

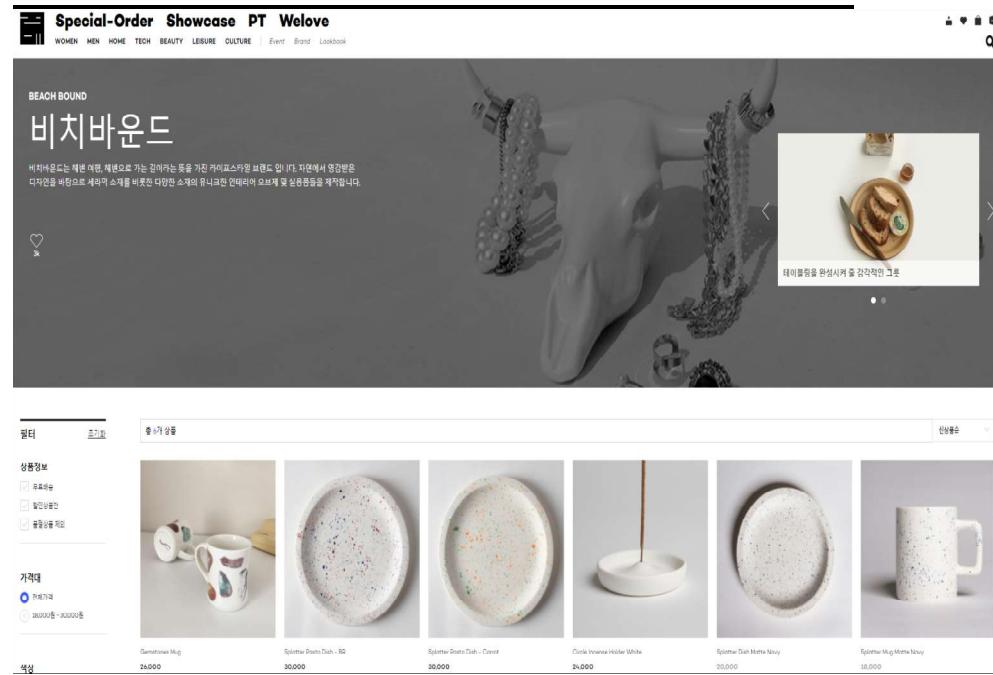
제품명	가격
육질 우유종_100g	37,000
수제 동물원 양우유_200g	35,000
밀고 카운트 양유_200g	109,000
육질 240g	75,000
소융한 나를 찾은 살로우_Lounge	163,000
우주우는 소나노를 알찬스对不起_Oriday	30,000

- 자연주의 건강 식품 판매
- 주 매출 상품은 휴식과 건강 관리를 위한 건강식품군
- 소미노 두유는 2개월간 매출 3위로 앞의 두 상품에게 밀리지만 **클릭 수는 1위**
- 식품류는 직접 먹어볼 수 없기 때문에 **20개 묶음 상품 외에 하루 1끼로 일주일 단위 저격 7개 묶음 판매 제안**

9. 결론

Brand

비치바운드

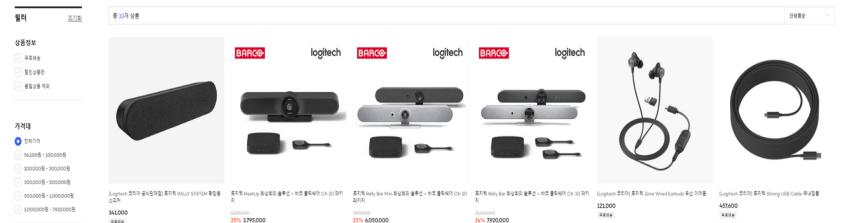
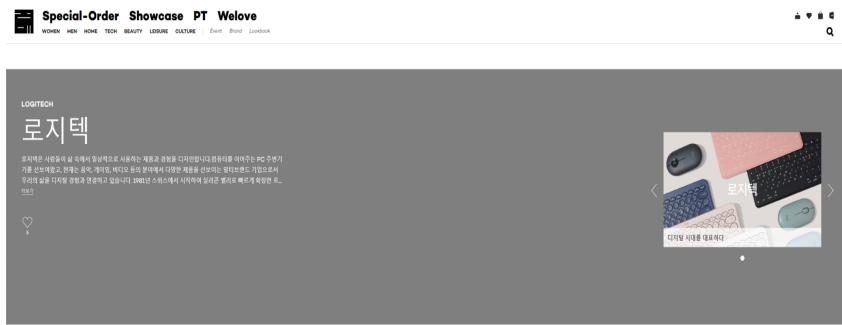


- 컵, 컵받침으로 이루어진 7개의 상품만으로 2개월간 약 700만원 이상의 매출
- 브랜드 상품들의 클릭 수를 보면 지금 **선택된 약소 브랜드 중 모든 상품 빠짐없이 높은 클릭 수**
- **제품 양을 좀 늘려보기**
- **특정 컬렉션적인 요소를 반영해보기**

9. 결론

Brand

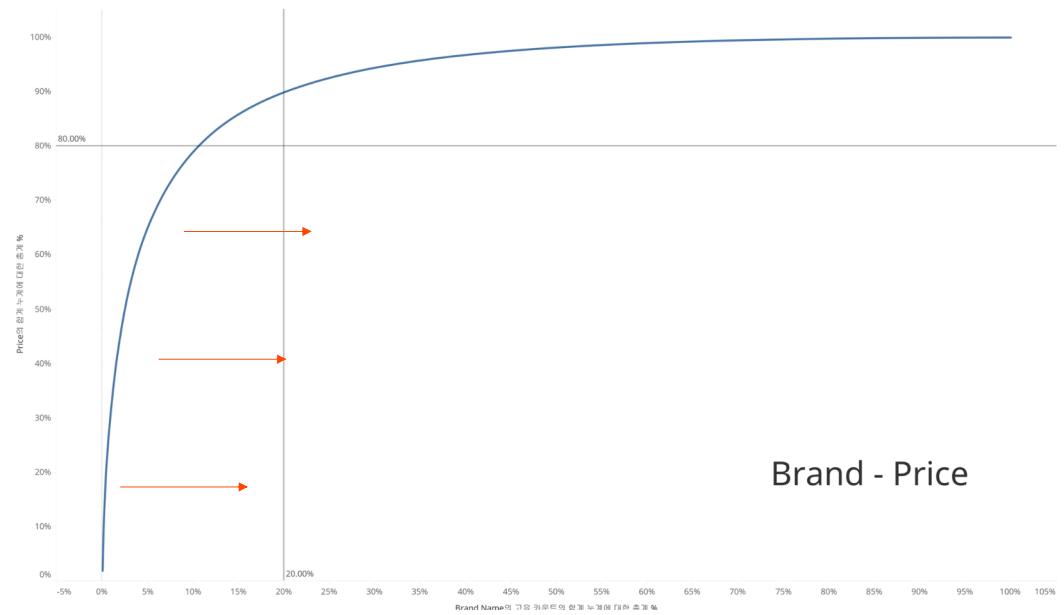
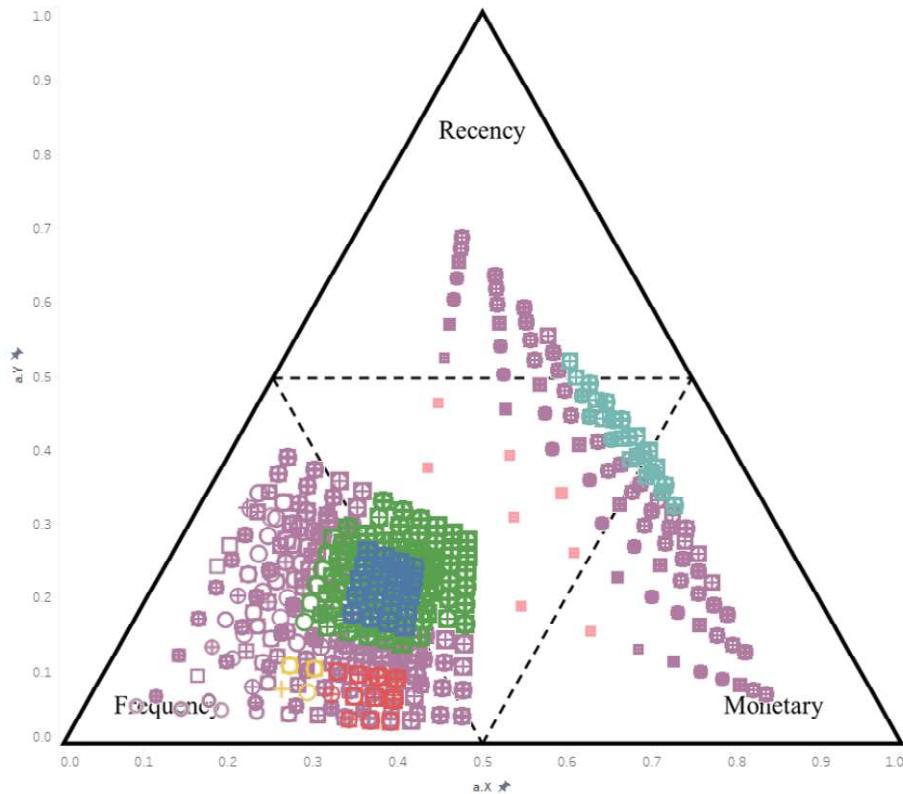
로지텍



- 게임 시장이 날마다 커져가는 가운데 PC주변기기는 물론 게이밍 기기 면에서 시장 점유율 1~2위를 하는 브랜드
- 회사 비전인 라이프스타일을 연계한 콜라보 런칭 기획 제안

9. 결론

성과 측정 지표



10. 프로젝트 진행 소감

[프로젝트 관련]

- 아쉬운 점
 - 데이터 검증을 더 꼼꼼하게 할 시간이 조금 부족했음
 - 고객의 페이지 방문, 마케팅 정보 등 세부적인 GA 데이터가 있었다면 더 면밀하게 분석 할 수 있었을 것임
 - 데이터의 기간이 짧아 경향성이나 일반화, 시간 관련 분석이 어려웠음
 - 성과 측정 할 수 있는 지표를 생각하기가 어려웠음
- 만족스러운 점
 - 2달 데이터로는 한계가 있으나 추후 기간이 늘었을 때 지속적으로 활용 가능한 대시보드를 작성함
 - RFM에서 충분한 논의를 거쳐 정성적 판단으로 분류한 세그먼트를 활용할 수 있었음

10. 프로젝트 진행 소감

[데이터 분석 관련]

- 배운 점
 - 데이터 전처리의 필요성과 데이터 엔지니어의 역량이 중요함
 - 시각화를 생각하는 것과 수행하는 것은 다름
 - 데이터 분석가에게 스킬만이 아니라 도메인 지식, 리터러시 등이 매우 중요함
 - 의사 소통에 할애하는 시간이 생각보다 많음
 - 작업 속도와 효율 면에서 SQL의 중요성을 느낌

[팀 관련]

- 팀 프로젝트였기에 혼자서는 하지 못했을 굉장히 많은 내용에 대해 심도 있게 토의하고 다룰 수 있었음
- 집단지성의 힘을 경험함

감사합니다.

내용과 관련하여 질문이 있으시다면
아래 메일로 연락해주시기 바랍니다.

EMAIL ► jungho_lee5@naver.com