

Projects Portfolio

Retailer 데이터 분석

정량 / POS 데이터 분석

역할 ●●●●●

독일 내 Retailer 분석을 통해 이상 가격과 판매량 간 관련성 입증.

분석 배경

Q

A사 B모델이 600유로에서 1200유로 사이에서 판매가 되고 있는데, 우리 현지는 600유로만 이야기를 한다. 또는 현재 POS 데이터로 보면 해당 제품은 평균가인 800유로로만 판매되는 것처럼 보인다.

실제 스토어 단위 별 어떤 가격대에서 얼마나 많은 제품들이 판매되는지, 최적 가격대는 무엇일까?

문제 정의 & 솔루션 설계

Problem Statement

Solution

Problems

실 판매 가격대 분석

주요 모델의 가격 구간 탐색

- ✓ 자사 및 경쟁사 주요 모델 Store 별 실 판매 가격대 분석
- ✓ Weekly / Store별 세부 데이터 분석
- ✓ 주별로 가장 빈번하게 형성되고 있는 가격대 도출

- ✓ 주요 모델의 최적 가격대 도출

1

GfK가 보유하고 있는 Weekly / Outlet 정보 활용

2

Seasonality 보정을 통해, 냉장고 시장 내 주 별 Seasonality 제거

3

주 별 box plot을 통해 Outlier성 가격으로 판매한 Store 탐색
평균 가격대의 변동 탐색

4

주 별 판매 패턴 탐색 후, Sales Value가 최대치가 되는 Price 도출 완료



가장 빈번하게 형성된 가격대 + 최다 판매 가격대 파악
최적의 가격대를 선정하는 데 도움

분석 과정

WEEKLY DISTRIBUTION 탐색(Weekly) → 경쟁사 CPI 분석 (Weekly) → 최적 가격 구간 탐색(Total)

Weekly Distribution 탐색 (Store 단위)

- 주별로 outlier성 가격대로 판매한 store 수 분석
 - 주별 가격대 분포 / outlier성 점포 수 분석
- 정상 가격대 분포에 따라, 일부 store내에서 outlier성 가격이 발생한 것인지, 해당 주차 내 시장 환경 변화로 가격 편차가 발생했는지 여부 확인 가능



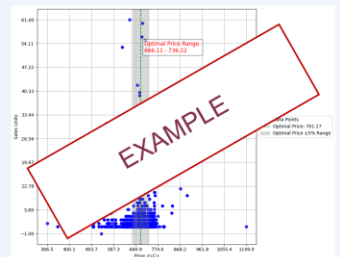
경쟁사와 CPI 분석

- 해당 브랜드의 핵심 모델과 동일한 용량대/동일한 type의 경쟁사 모델 간 가격 경쟁력 분석
- 가격 경쟁력과 더불어, seasonality를 제외한 sales unit과의 상관관계 도출 가능
- 전체 CPI 뿐 아니라, 주요 스토어 단위의 CPI 분석 가능



최적 가격 구간 탐색

- 독립변수를 가격, 종속변수를 Sales Value로 상정한 후, Sales Value가 최대치가 되는 가격 및 가격 구간 산출
- 주 별 Price와 Unit, Outlier성 가격대 변화 뿐 아니라, 전반적인 시장 내 가격 경쟁력을 도모할 수 있는 가격 구간에 대한 인사이트 도출 가능



Retailer 데이터 분석

정량 / POS 데이터 분석

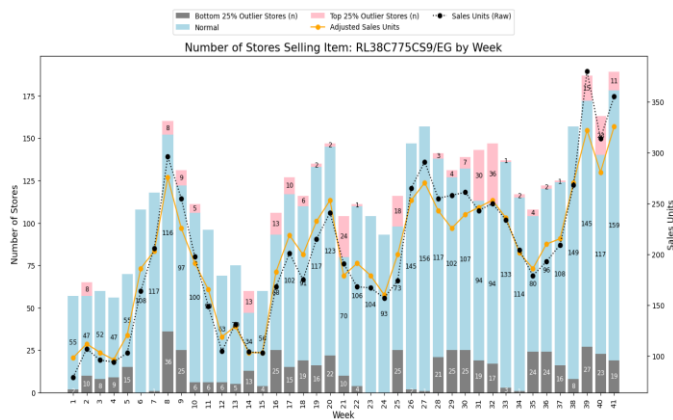
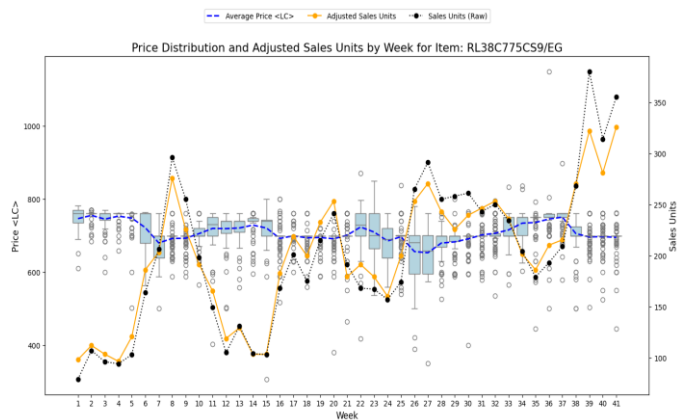
역할 ●●●●●●

독일 내 Retailer 분석을 통해 이상 가격과 판매량 간 관련성 입증.

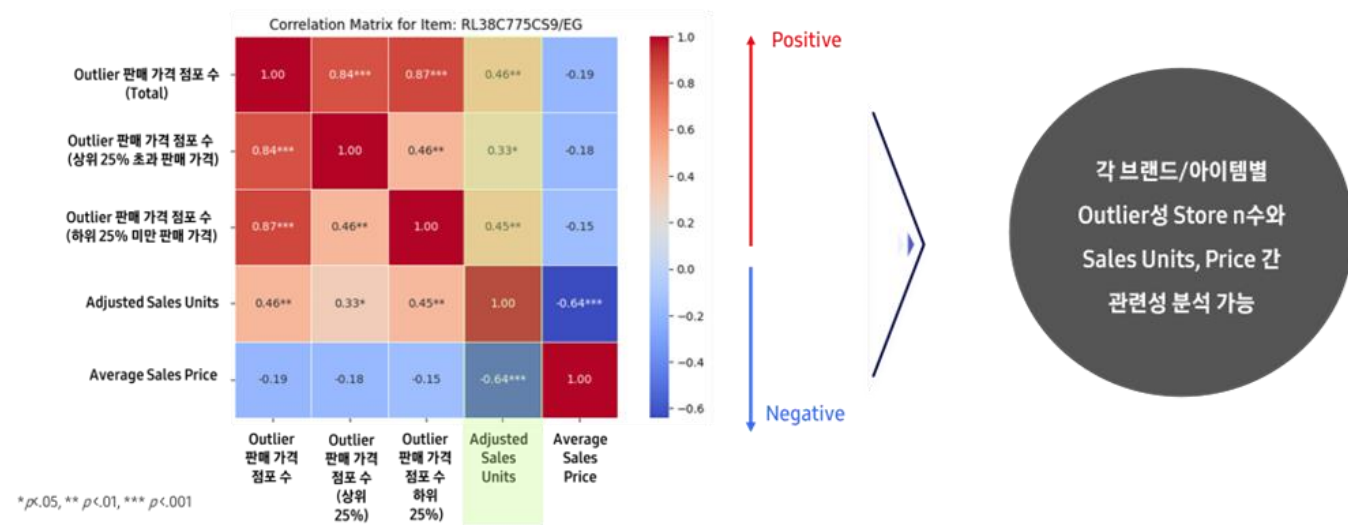
분석 결과

분석 결과 1 - Weekly Distribution 분석

사분위 기준 상위 25%, 하위 25% 범위 초과는 이상치로 간주



분석 결과 2 - Outlier Store 수와 Sales Units 간 상관관계 분석



분석 결과 3 - 최적 가격대 분석

변수 설정



성과

- Weekly Data 구독
- 현재 제공되는 Weekly 데이터로 확인할 수 없는 인사이트 제공을 통해 고객사의 만족도 제고

영합 ○ ○ ○ ○ ○



- [illegible]

오히려 ● ● ● ● ●

- ## 정량분석

식품 브랜드 동적 군집화

정형데이터 분석

역할 ●●●●●

시간의 흐름에 따라 식품 시장의 군집화 변화를 살펴본 후, 유사 제품 간 경쟁관계인지/유사관계인지 여부 판별

배경 조사 | 식품시장에 영향 미치는 요인 탐색 후 데이터셋 구성

식품 시장 내에서 유사한 느낌의 제품들이 있는데, 이 제품이 서로 경쟁 관계인지/혹은 시장을 같이 성장시키는 관계인지 궁금합니다.

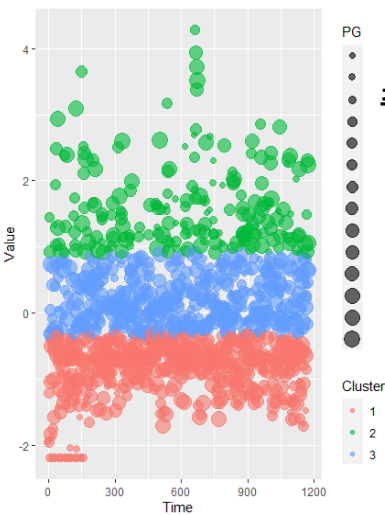


- 군집화를 통해 **통계적으로 유사 군집**인지 파악을 목적으로 함
- 군집분석 후 **상관분석**을 통해 **각 제품들 간 positive/negative 관계 여부** 파악
- 시간의 흐름에 따라 군집이 안정적인지 이해하기 위해 동적 SOM과 DBA/DTW 분석 활용

분석 과정

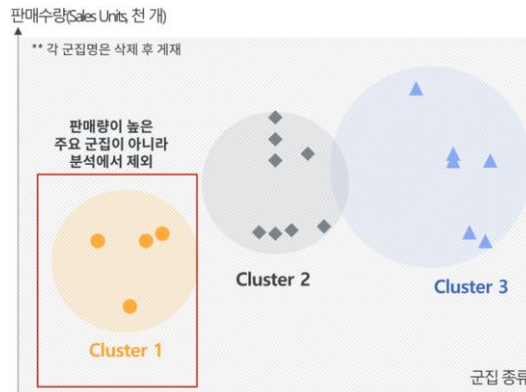
동적 SOM(Self Organizing Map) & DBA K-means Clustering

시간의 흐름을 반영한 Clustering 결과, 총 3개의 군집이 시간의 흐름과 관계 없이 안정적으로 유지되었음



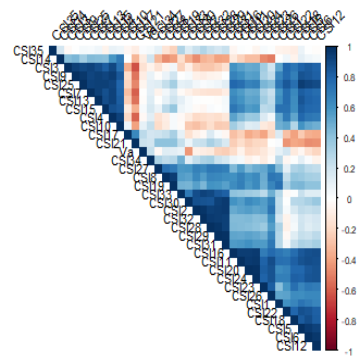
SOM > 1차 군집 요약 >
K-means Clustering >
실루엣 계수 확인
실루엣 계수: 0.64

주성분분석(PCA)를 통해 군집 내 품목 확인



일부 판매 수량이 높은
제품군만 Dot 표기

군집 내 제품들의 상관분석을 통해 관계성 파악



Insights

- 동적 군집화를 통해 군집의 안정성 검증 / 실루엣 계수 검증을 통해 모델의 타당성 확인
- 종단 시계열 분석을 통해 시간의 흐름에 따라 안정적으로 군집이 구성되어 있다는 점을 확인
- 각 유사 식품군별 경쟁 관계 / 공생 관계를 확인하여 제조사의 효율적인 제품군 출시 계획에 도움

가구 브랜드 감성분석

비정형데이터 분석

역할 ●●●●●

COVID-19 이후로 사무의자/인테리어 제품 관심도 높아진 가운데, 해당 가구 브랜드가 새로운 제품 라인 출시 및 인플루언서를 통한 광고 시작. 관련 VoC 수집을 통해 가구 브랜드에 대한 온라인 인식/광고 효과를 정성적으로 파악

Crawling을 통한 VoC 수집

온라인 커뮤니티(D모 사이트 가구갤러리, Youtube 등) 댓글 크롤링
크롤링 시 Rselenium을 활용한 동적 크롤링을 활용함

유튜브 영상 url: <https://www.youtube.com/watch?v=4DLCaQZ6VC4>

아주 달다>이거 색을 거 같아요> 매우 신선하다>>방금 죽은 거 같다> 두 종류를 잘 절충했다>>교잡종 같다> 이렇게 칭찬받지 않은 칭찬이 하나 더 생겨났다

침착한 클라쓰.. 꾸준한 감학지식 설명화로 광고까지 따왔는데 광고담당자도 설명회 ppt 알차게 준비해서 거부감없이 광고로 잘 넘어갔음ㅋㅋㅋ 덕분에 의자 살 생각 1도 없었는데 시디즈 찾아볼ㅋㅋ

이 날 방송내내 너무 웃겼음 청자들 이악물고 비싼 물건 찾아내면서 탈락 외치더니 다리미좌는 책상도 사주라 하고ㅋㅋ

시디즈가 사무용으로 지조있게 같을았는데 게이밍으로든 뛰어드는 도전정신 맘에들어

다리미좌 책상 보자마자 아무도 반대 안하는게 스니웃김ㅋㅋㅋㅋ

집이 추운 사람을 위해서 열선 게이밍 의자도 만들어줘임ㅋㅋ

아 근데 디자인이 진짜 이쁘긴 하다 객적으로 게이밍 의자 디자인이 항상 아쉬웠는데 이 의자라는게 허리때문에 예민한 문제인데 꼭 알아보고 나한테 맞는게 아니냐고 생각했는데 의자라는게 유명한 게이밍의자도 마음에 드는게 없고 사무용도 마음에 드는게 하나도 없어서 이 의자라는게 정말 좋은 거 같음

다른 의자와 같은 기능이 있고, 쾌적하게 쥘 수 있는 무지개색이랑 디자인이 마음에 드는게 좋겠다

예전부터 생각했던건데 의자 헤드레스트 부분에 사이드 스피커나 리모컨이 있으면 좋을거 같음 출시여도 무신 리어 스피커 나오는거 보면 의자에도 적용 가능할듯 싶은데 말이죠

진짜 광고영상엔 채팅창이 없어도 이미 수년간 학습이된 내 시야엔 채팅창이 보이네요 ㅋㅋㅋㅋ

명품의자 회사에서 게이밍을 내놓네 ㅋㅋ 사고싶다...

침착맨식 홍보법 언제보나 경이로운건지 신기한건지 재밌다 ㅋㅋㅋㅋ

허리디스크가 있어서 허리건강에 안좋다니 이번락은 패스지만 경험이 쌓이면서 게이밍 의자 개발 노하우가 좀더 쌓이면 고려해보겠습니다!

게이밍 의자를 바라보는 해결사녀굴면의 시각은 역시 다르다...!

클링시스템있는건 ㄹㅇ 깔끔들긴했네.

Example

텍스트마이닝 & 감성분석

KoNLP + SimplePOS09 사용하여 Token화

단어를 형태소 단위로 분류한 후 특수문자/분석에 활용하지 않을 단어 삭제 후 Token화

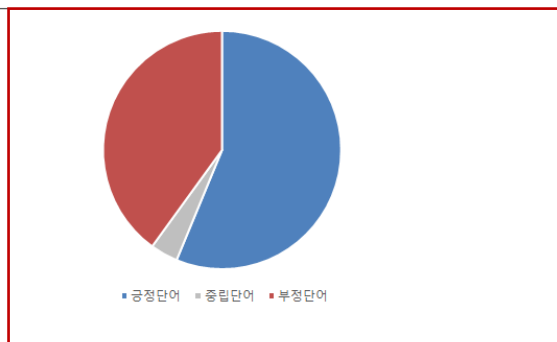
sentence	word
1	신선하다
1	죽은
1	잘
2	꾸준한
2	잘
3	비싼

감성분석

군산대학교 감성어 사전을 활용하여 단어별 감성점수 부여 후 감성분석 실시

sentence	word	value	sentiment
1	신선하다	1	positive
1	죽은	-1	negative
1	잘	1	positive
2	꾸준한	2	positive
2	잘	1	positive
3	비싼	-2	negative
6	추운	-1	negative
8	저는	-1	negative
8	유명한	1	positive
8	없어서	-1	negative
8	장점을	2	positive
8	좋은	2	positive
21	비싼	-2	negative
23	아무지개	2	positive
24	익숙하고	1	positive
25	비싼	-2	negative
25	아름	-2	negative
29	줄다	2	positive
31	편하고	2	positive
36	친구	1	positive
49	좋다고	2	positive
50	잘	1	positive
52	중요한	1	positive
52	바르게	2	positive
52	ㅎㅎ	1	positive
60	싸움	-1	negative

Count	
긍정단어	45
중립단어	3
부정단어	32



→ 긍정 단어 비중이 상대적으로 높음

Insights

- 해당 브랜드는 대중성이 높은 브랜드이지만, 최근 하이엔드 가구 브랜드에 대한 니즈 증가와 더불어, 가성비 브랜드에 대한 니즈 또한 증가하고 있기에, 단지 대중성을 가지고서는 시장 내에서 포지셔닝을 하기 어렵다는 사실을 도출하여, 추후 마케팅 방법의 보완점을 제안
- 하지만, 감성분석의 경우 형태소의 맥락적 의미(예: 반어법)를 고려하지 못하기에, 추가 VoC 검토 필요성이 있음

골프 브랜드 제안 프로젝트

비정형데이터 분석

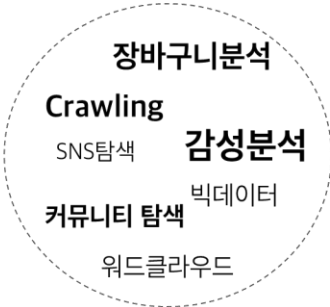
역할 ●●●●○

골프 브랜드 인지도/광고효과 조사 제안서 작성 시, 설문 개발/Segments 구성을 위해 온라인 기사/VoC분석을 통해 효율적인 설문 조사 도모

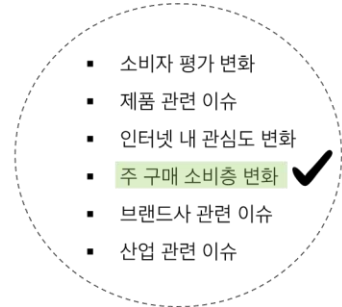
배경 조사 | 골프 시장 상황을 고려한 분석 방법 결정



어떤 아이템을 보고 싶은가?



어떤 방법을 사용할 것인가?



고려해야 할 점은 무엇인가?

- ✓ 2022년 이후, 골프 용품의 주 소비자 반열에 점차 젊은 연령층의 비중이 증가하기 시작함(예: 골린이, MZ)
- ✓ 유튜브, SNS, 골프 커뮤니티의 발달 + 중고시장 / 가성비 등의 키워드 함께 등장

In conclusion, 소비층의 변화에 따른 소비자 평가의 변화 / 브랜드의 Status변화 등이 예상되기에, 기사 탐색을 통한 골프 용품 관련 키워드 중심성과 더불어 커뮤니티 탐색을 통한 소비자들의 VOC를 동시에 확인하는 전략을 세움

Crawling & 사회적 연결망 분석

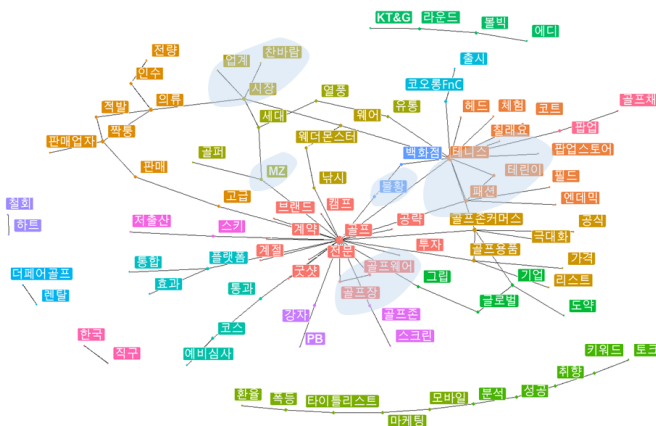
온라인 기사 탐색 및 사회적 연결망 분석을 통해 골프 관련 최신 이슈 탐색

Apriori + 고유벡터 중심성 고려해서 사회적 연결망 분석 수행 → 브랜드명이라는 중심 단어와 밀접한 단어를 찾기 위함

[결과 해석]

연령 / 관련 아이템 / 특징 도출

- ✓ 가장 중심적인 키워드는 '골프'이며, 해당 단어를 중심으로 테니스/MZ/골프웨어/골프장/패션/계절/고급 등의 단어가 연결
- ✓ 골프 용품과 관련하여, 짝퉁/적발/불황 등의 부정적인 단어들도 나타남
- ✓ 골프용품 예상 소비자층이 테니스/낚시 등의 다양한 스포츠에도 관심이 있을 것으로 예측됨



Questions

- ? 연도별 골프 용품에 관한 정보량은 어떠한가
- ? 가장 인기를 얻고 있는 아이템은 무엇인가
- ? 연령대별로 호불호가 있는 브랜드가 있는가
- ? 필드용/실내용으로 선호되는 브랜드가 있는가
- ? 골프 아이템과 관련된 VOC는 무엇이 있는가
- ? 최근 인기 스포츠 대비 골프의 위상은 어떠한가

Strategies

정보량 대조

커뮤니티/
SNS 탐색
문헌고찰

⇒ 브랜드 관련해서 짝퉁/적발/불황/찬바람 등의 부정어도 동시에 주요 단어로 등장했기에, VoC를 통해 더 세부적인 원인 분석 / 감성분석을 통해 부정적인 단어가 타 브랜드에서도 나타나는지 확인 필요

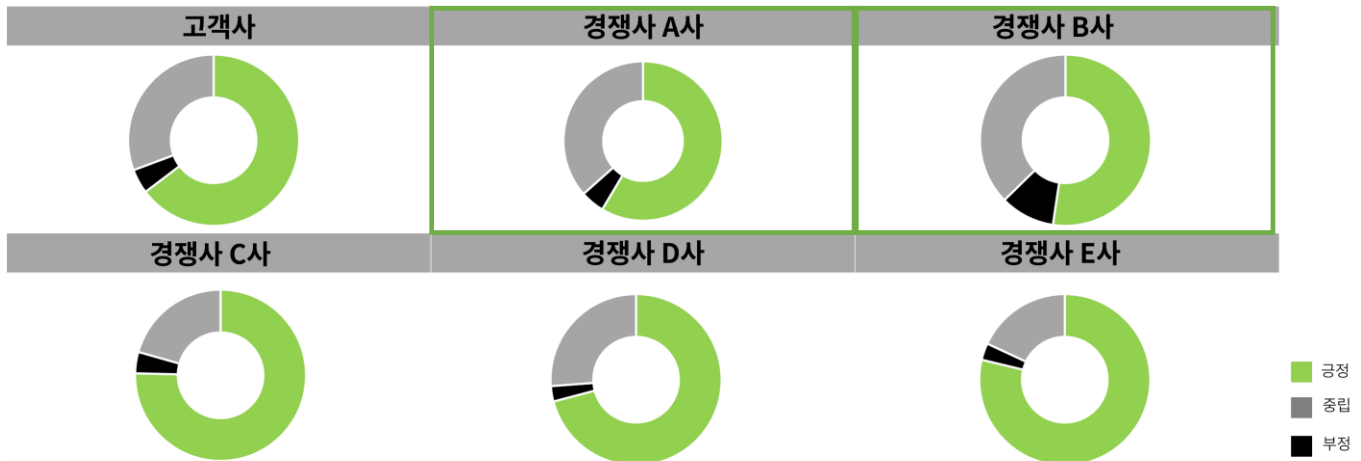
골프 브랜드 인지도/광고효과 조사 제안서 작성 시, 설문 개발/Segments 구성을 위해 온라인 기사/VoC분석을 통해 효율적인 설문 조사 도모

감성분석 & 워드클라우드 분석 수행

온라인 커뮤니티(D모 사이트 골프갤러리, R모 사이트, Youtube 등) 댓글 크롤링
크롤링 시 Rselenium을 활용한 동적 크롤링을 활용함

감성분석 결과

경쟁사 선정 기준은 매출액 & 정보량 대조 후 유사한 정보량을 보이는 브랜드 채택 후 추가 분석 수행



워드 클라우드 & VoC 분석 결과

[고객사]



[경쟁사 A사]



[경쟁사 B사]



(-) “백화점 타이틀 피팅 스튜디오에서 헤드만 주문한 게
왔다. 저 선명한 **마데인 차이나.. 비싸게 호갱**”
(골프 커뮤니티)

(-) 감성에 뽕 잘 팔아서 프로들
스폰하니까 투어사용률 1위지. 막 입문하는 사람들한테는
근본인 느낌 오케이. 근데 **작년 T200은 알리 째민 줄..**”
(골프 커뮤니티)

(+) “진짜 제대로 된 샤프트 피팅 받고 싶으면
한테 가라”(골프 커뮤니티)

(+) 하면 떠오르는 건 **거리와 내구성**"(유티브)
(+) 진짜 안 깨져요. **헤드가 튼튼한 것**
같아요"(유티브)

- (-) “왜 이렇게 **비싸까?**” (유튜브)
- (-) **쓰는 프로가 있긴 한가..**”(골프 커뮤니티)
- (-) “**거품**이 어마무시~ 웬만하면 간지말이라
알겠는데 **넌 비싸**”(유튜브)
- (-) “원래 럭셔리 이미지였는데 요즘 **싸마이** 느낌”(유튜브)

(+) !타구음 좋은데요 탕탕탕!“ (골프 커뮤니티)

- (-) “함참 찬란하다 아집으로 **혁신에 실패**”(유튜브)
- (-) “예전엔 좋았지만 지금은 **파산 뒤 기술이 사라진** 지 오래죠. 시니어에게 특화되었을 뿐”(유튜브)
- (-) 약간 **시니어 틀딱** 느낌..”(골프 커뮤니티)
- (-) **쉬운 채는 아니죠**”(골프 커뮤니티)

(+) “아이언 타감 1위는 . . .” (골프 커뮤니티)
 (+) “어르신들이 자주 쓰는 덴 이유가 있다. 카본
 샤프트가 **알아서 땀땀 쳐주는 느낌**. 진짜
잘 뜨고 오지게 편하긴 하더라!”(골프 커뮤니티)

Insights

- 텍스트마이닝을 체계적으로 활용하여 별도 파일럿 조사 없이 효율적으로 경쟁 브랜드 채택
- 해당 브랜드의 주요 고객층을 가늠할 수 있는 정보 수집을 해서, Segments 구분에 도움을 줌
- 정량조사로는 알 수 없는 보다 심도 있고 생생한 자료 수집
- 브랜드별 인식 차이 뿐 아니라 골프 산업과 관련된 최신 이슈를 본 조사에 반영 가능
- **제안 수주 및 골프 브랜드 관련 조사 수행(2023)**

대시보드 제작 프로젝트

타사에서는 클라이언트들이 직접 데이터를 쉽게 확인할 수 있는 시각화 서비스를 제공 중임. 자사에서 데이터 시각화에 대한 필요성이 제기되어 몇 가지 대시보드 샘플 안을 제작함

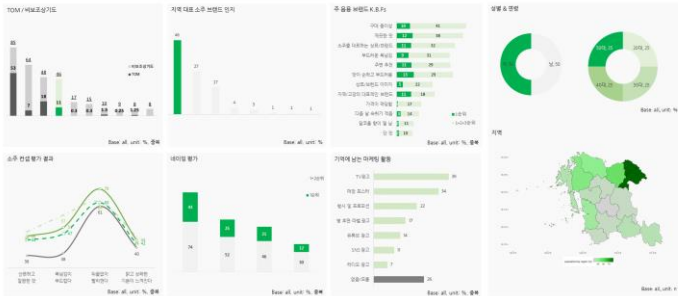
비정형데이터 분석

역할 ●●●●●●●●

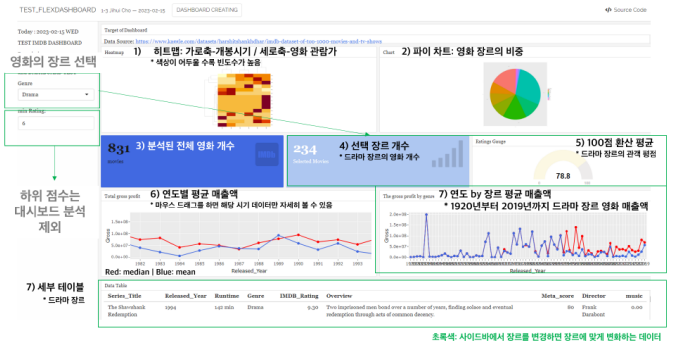
각 Tool별로 대시보드를 제작한 후 장/단점 분석

Excel Dashboard [Pivot / Slice / GIS Data 사용]

지역 소주 소비량 조사 결과

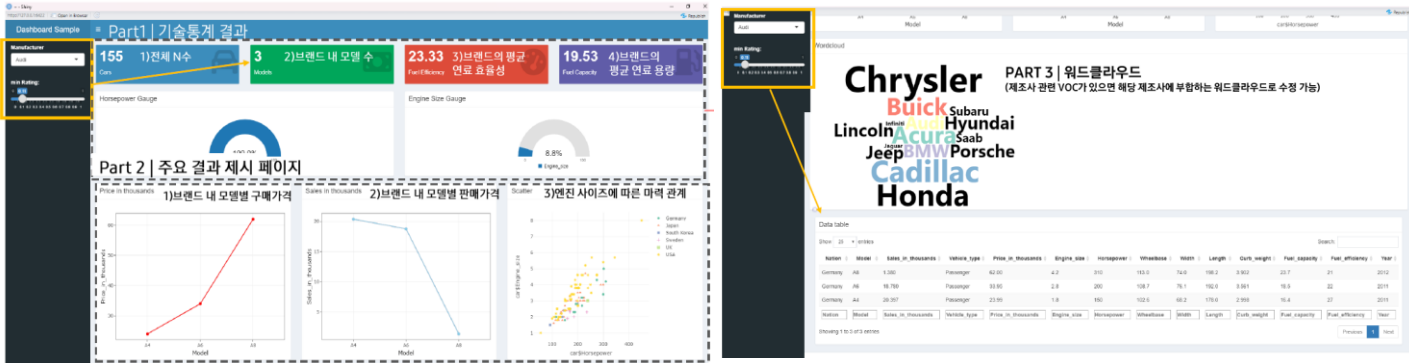


R FlexDashboard



주목: 사이드바에서 장르를 변경하면 장르에 맞게 변화하는 데이터

R Shiny Dashboard



각 대시보드의 장/단점 분석

장점

단점

Excel Dashboard

- 프로그래밍 언어에 익숙하지 않아도 쉽게 제작 가능
- 익숙한 차트/쉬운 차트 제공 가능
- 웹 리포팅 필요 없이 파일로 제공 가능

R Flex & Shiny Dashboard

- 웹 페이지 형태로 제공 가능
- R언어에 익숙하면 내부 구조를 쉽게 변경할 수 있음
- 즉각적인 유지/보수 가능
- 엑셀/PPT로 제공이 어려운 차트 제작 가능

- 수작업이 필요하며 작업 시간이 오래 걸려서 비효율적임
- 원본 데이터 유출 가능성
- 유지/보수를 즉각적으로 반영하기 어려움
- R언어에 대한 지식 필요
- 엑셀/PPT 보고서 대비 디자인이 예쁘지 않음 (Tableau나 Power BI 툴로 보완)
- 웹 리포팅이 제한적이며, Flexdashboard의 경우 markdown 사용법을 알아야 함

Insights

- 타 부서/제작팀 등의 인력 지원을 적게 들어서 자체적으로 대시보드 리포팅 가능
- 별도 툴 (Tableau, PowerBI) 교육 없이 엑셀 등 익숙한 프로그램을 통해서도 대시보드 제작이 가능
- 웹 리포팅 한계 / 실질적으로 클라이언트 서비스가 되지 않음 (IT팀 인력 한계)
- R언어나 Python언어를 활용할 수 있는 연구원이 부족하여 프로젝트 중단