<엔지니어링DB 및 실습 – Project 최종 보고서>

**최적 광고 매체, 타겟, 비히클 추천**

수업 시간: 수 D 금 D

조장: 201620239 최유림

조원: 201620237 강윤환

201620231 전세환

201520240 변현배

**목차**

1. **서론**
   1. 프로젝트 배경 및 목표
   2. 프로젝트 수행 과정
2. **본론**
   1. 데이터 조사
      1. 소비자행태조사(MCR) 공공 데이터(KOBACO)
      2. 시간대 별 TV, 라디오 매체 채널 별 광고 단가
      3. 시간대 별 TV, 라디오 매체 채널 별 시청률
      4. 인터넷, 신문, 잡지 회사 별 광고 단가
      5. 외부 광고 단가
   2. 데이터 베이스 설계 및 구축 과정
      1. 요구 사항 분석 및 개념적 설계
      2. 엔티티 구현
      3. 정규화 & 엔티티 통합 및 분리
      4. ER Modeling

2.2.4.1 관계설정

* + - 1. Extended ER model
  1. 계층화 의사결정법(AHP)을 이용한 가중치 선정
     1. AHP 기법의 기본 절차
     2. 실습에서의 AHP 기법 적용

2.4. 광고 매체 추천 SQL 쿼리 문 작성

2.4.1 뷰를 사용한 적합도 테이블 생성

2.4.2 고객이 타겟을 특정화 하였을 때

2.4.3 고객이 타겟을 특정화 하지 않았을 때

2.4.3.1 MySQL 데이터 추출

2.4.3.2 파이썬 연동

2.4.3.3 단가 뷰 생성

* + 1. 비히클 추천 SQL
  1. 테스트 결과 및 해석

2.5.1 테스트 결과

2.5.2 결과 해석

1. **결론**

3.1. 데이터 베이스

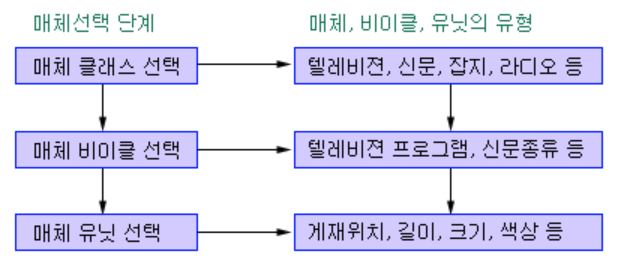
3.2. 데이터 분석

3.3. 기대효과

1. **참고문헌 및 출처**
2. **서론**
   1. **프로젝트 배경**

TV, 라디오, 신문과 같은 전통적인 매체부터 SNS나 블로그, 모바일 앱과 같은 현대적인 매체에 이르기까지 소비자에게 상품을 홍보하기 위해 다양한 광고가 이뤄지고 있다. 2016년부터 총 광고비의 변화를 살펴보면 광고시장의 규모가 점점 커져왔으며, 앞으로도 이러한 흐름이 계속될 것이라고 예측할 수 있다.

하지만 광고가 이뤄지기 위한 매체 선택 단계는 매체 클래스 선택, 매체 비히클 선택, 매체 유닛 선택과 같은 세 단계로 이뤄지게 되는데 이때 기존의 사례, 관례, 경험 같은 요인이 크게 작용한다. 우리는 점점 광고시장의 규모가 확대되고 있음에도 불구하고 이러한 전통적인 방식의 매체 선택 방식으로 광고 매체를 결정하는 것이 효과적이지 않다고 판단했다.



<광고 선정 절차>

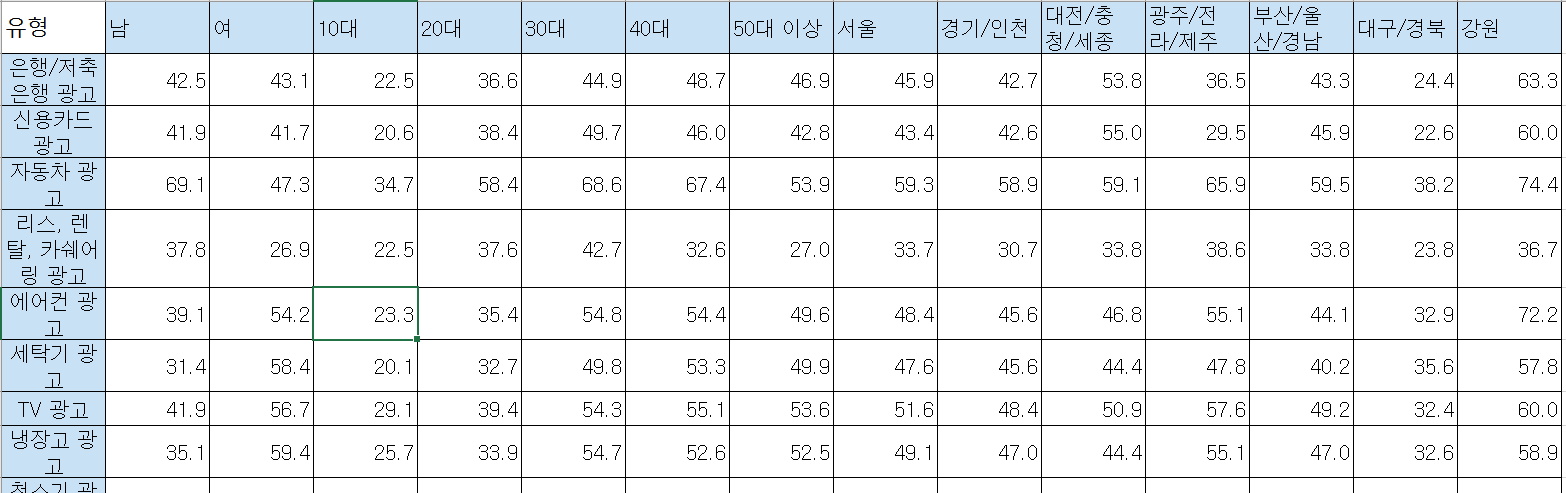
* 1. **프로젝트 목표**

우리는 기존의 관례, 경험에 의존하던 광고 매체 선택 단계를 발전시키려 한다. 소비자행태조사(MCR) 공공데이터를 기반으로 데이터를 수집한후 데이터 분석을 통해 품목별, 타겟별 최적의 광고 효과를 낼 수 있게 광고 매체 적합도를 계산하여 보다 합리적이고 효율적인 방법으로 광고 매체 선택을 할 수 있도록 할 것이다.

* 1. **프로젝트 수행 과정**

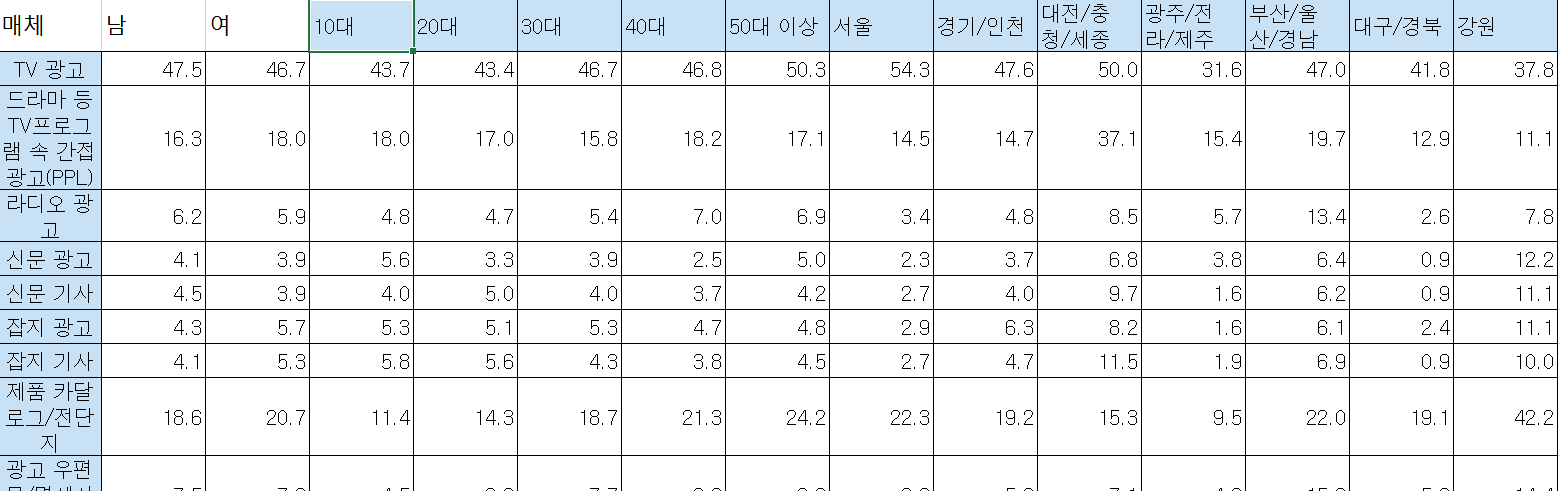
품목별 적합한 광고 매체와 광고 타겟을 선정하기 위해 소비자행태조사(MCR) 공공 데이터를 수집하였고 광고 매체의 가성비를 고려하기 위해 매체별 광고단가와 매체별 시청률 데이터를 수집하였다. 수집한 데이터를 기반으로 ER Modeling과 정규화 과정 등을 거쳐 데이터베이스를 구축하고, AHP기법과 SQL 질의문을 활용해 최적의 광고 매체를 선택할 수 있도록 했다.

1. **본론**
   1. **데이터 조사**
      1. **소비자행태조사(MCR) 공공 데이터(KOBACO)**
2. 품목 별 광고 관심도



광고 품목들을 광고하였을 때, 어떠한 성별, 연령대, 거주 지역의 사람이 관심을 어느 정도로 보였는지에 대한 데이터를 나타낸 것이다. 데이터 값들은 해당 품목의 광고에 대하여 각 유형의 사람들 중 몇 %가 긍정 응답을 하였는지를 의미한다.

1. 매체 접점(구매)



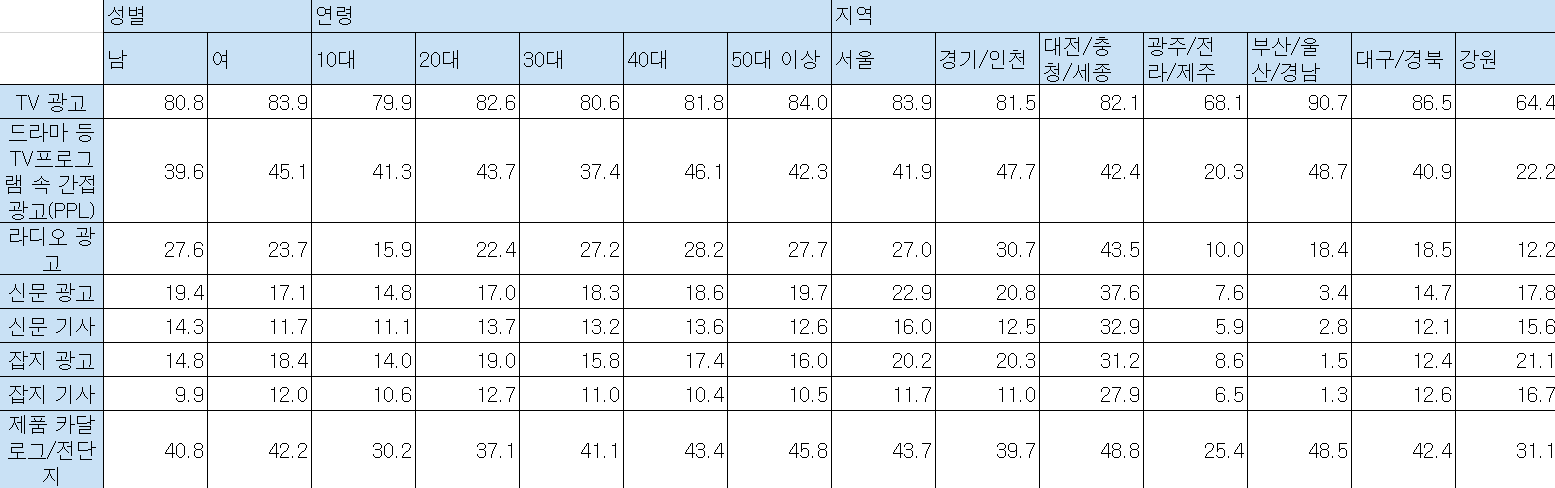
어떠한 품목을 여러 매체들에 광고를 하였을 때, 해당 매체의 광고를 시청한 각 유형의 사람들 중 몇 %가 실제로 품목을 구매했는지에 대한 데이터들로, 각 매체들에 광고를 했을 때 실제 구매에 얼마나 영향을 미쳤는지를 나타낸 것이다.

1. 매체 접점(호감)



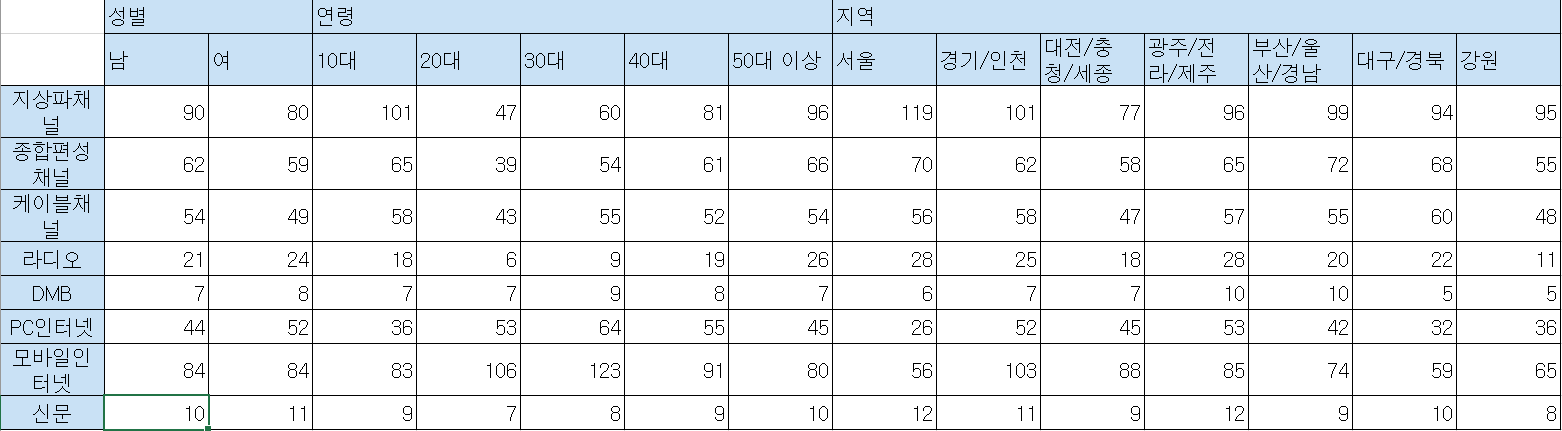
여러 매체들의 광고들을 사람들이 접했을 때, 각 유형의 사람들 중 몇 %가 해당 매체의 광고에 대하여 긍정 응답을 하였는지에 대한 데이터들을 나타낸 것이다.

1. 매체 접점(인지)



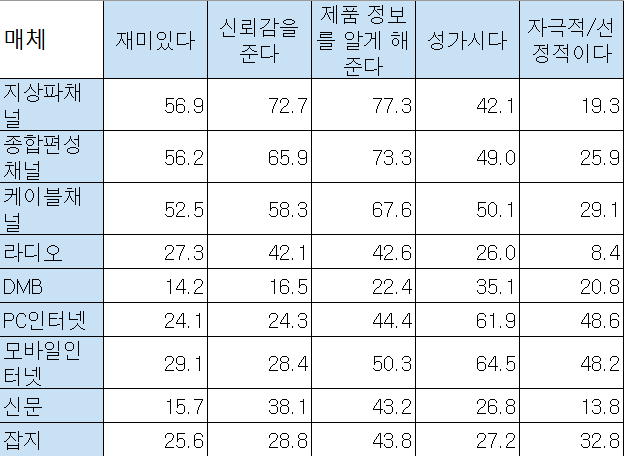
여러 매체들의 광고들을 사람들이 접했을 때, 각 유형의 사람들 중 몇 %가 해당 매체의 광고에 주목하였는지에 대한 데이터를 나타낸 것이다.

1. 매체 이용시간 상세(평일, 2019)



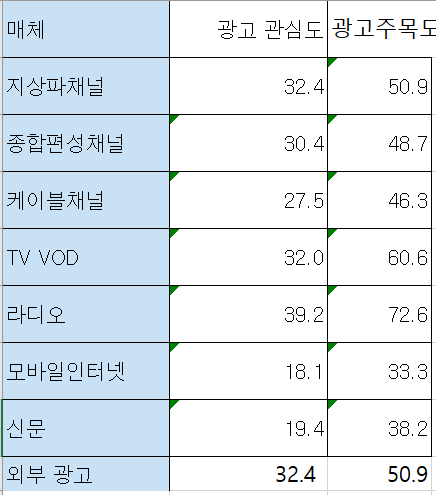
여러 매체들을 각 유형의 사람들이 평일에 몇 분 이용하는지를 몇 분 이용하는지를 나타낸 것이다.

1. 매체 별 광고 이미지 평가



매체그룹들에 대한 광고 이미지 평가를 하여서 각 항목 당 선택한 사람들의 비율을 나타낸 것이다.

1. 매체그룹별 광고 관심, 주목도 평가



매체 그룹에 다한 광고 관심도와 광고 주목도를 조사하여 해당 매체 그룹에 관심, 주목하는 사람들의 비율을 나타낸 것이다.

* + 1. **시간대 별 TV, 라디오 매체 채널 별 광고 단가**





<KBS, JTBC 시간대별 광고 단가표>

TV, 라디오 매체의 경우 광고 단가에 따라 시간대를 등급으로 나눠서 가격을 책정하는 방식이다. 각 방송사 별로 공개되어 있는 광고 단가 정보를 통해 시간대별 광고 단가를 각 방송사별로 조사하였다.

* + 1. **시간대 별 TV, 라디오 매체 채널 별 시청률**

TV 시청률 조사 기관인 닐슨 코리아에서 공개된 채널 별 시청률 데이터를 수집하였다.



* + 1. **인터넷, 신문, 잡지 회사 별 광고 단가**

인터넷, 신문, 잡지 매체의 회사별 광고 단가를 각 회사의 사이트에 공개되어 있는 광고 단가 정보를 통해 조사하였다.

<Top Gear 잡지 광고 단가표>

* + 1. **외부 광고 단가**

광고 대행사에서 공개된 외부 광고 단가표를 조사하여 수집하였다.



<등급별 버스 부착 광고 단가표>

* 1. **데이터 베이스 설계 및 구축 과정**
     1. **요구 사항 분석 및 개념적 설계**

1단계) 요구사항 분석: 사용자가 수행하는 업무를 분석하여 필요한 데이터가 무엇인지, 그 데이터에 어떤 처리가 필요한지를 알아내고, 사용자의 요구 사항을 수집하여 요구 사항 명세서를 작성한다.

2단계) 개념적 설계: 요구 사항 분석 결과를 기반으로 데이터 모델을 이용, 개체와 속성을 추출하여 개체 간의 관계를 결정할 수 있도록 한다.

1. **광고 기획사 아산기획에 광고 의뢰를 위해 회원으로 가입하려면 회원 아이디, 비밀번호, 이름, 전화번호, 소속을 입력해야 한다.**

개체: 회원

속성: 회원 아이디, 비밀번호, 이름, 전화번호, 소속

1. **회원은 회원 아이디로 식별한다.**

회원 개체의 기본 키는 회원 아이디이다.

1. **회원이 광고 최적 적합도를 의뢰하면 의뢰에 대한 의뢰번호, 제품명, 품목, 의뢰일자 정보를 유지해야 한다. 이때 회원의 요구에 따라 광고 타겟을 특정하거나 광고 예산을 설정할 수 있다.**

개체: 회원, 광고 적합도

관계: 의뢰

“의뢰” 관계의 속성: 의뢰번호, 제품명, 품목, 의뢰일자, 광고 타겟, 광고 예산.

1. **의뢰는 의뢰번호로 식별한다.**

관계 “의뢰”의 기본 키는 의뢰번호이다.

1. **회원은 여러 개의 광고 최적 적합도를 의뢰할 수 있고, 광고 최적 적합도 하나는 한 명의 회원만 의뢰할 수 있다.**

개체 “회원”과 개체 “광고적합도”가 맺는 관계는 일대다 (1:n) 관계임을 알 수가 있습니다

.

1. **광고 최적 적합도는 의뢰번호, 광고 최적 적합치, 매체, 광고 타겟 정보를 유지해야 한다.**

개체: 광고 최적 적합도

속성: 의뢰번호, 광고 최적 적합치, 매체, 광고 타겟

1. **광고 최적 적합도는 의뢰번호로 식별한다.**

개체 ‘광고 최적 적합도’의 기본 키는 의뢰번호이다.

1. **성별과 연령대에 따른 각 광고 품목에 대한 품목별 광고관심도를 유지해야 한다.**

개체: 품목별 광고 관심도

속성: 성별, 연령대, 광고 품목

1. **품목별 광고 관심도는 광고 품목으로 식별한다.**

개체 “품목별 광고 관심도”의 기본 키는 광고 품목임을 알 수가 있다.

1. **성별과 연령대에 따른 각 매체에 대한 매체별 인지도를 유지해야한다.**

개체: 매체 별 인지도

속성: 성별, 연령대, 매체

1. **매체별 인지도는 매체로 식별한다.**

개체 “매체별 인지도”의 기본 키는 매체임을 알 수가 있다.

1. **성별과 연령대에 따른 각 매체에 대한 매체별 호감도를 유지해야한다.**

개체: 매체별 호감도

속성: 성별, 연령대, 매체

1. **매체별 호감도는 매체로 식별한다.**

개체 “매체별 호감도”의 기본 키는 매체임을 알 수가 있다.

1. **성별과 연령대에 따른 각 매체에 대한 매체별 구매 결정 영향력을 유지해야한다.**

개체: 매체별 구매 결정 영향력

속성: 성별, 연령대, 매체

1. **매체별 구매 결정 영향력은 매체로 식별한다.**

개체 “매체별 구매 결정 영향력”의 기본 키는 매체임을 알 수가 있다.

1. **성별과 연령대에 따른 각 매체 그룹들에 대한 매체별 이용시간을 유지해야한다.**

개체: 매체별 이용시간

속성: 성별, 연령대, 매체 그룹들(TV광고, 라디오 광고 등등 여러 개)

1. **매체별 이용시간은 매체 그룹으로 식별한다.**

개체 “매체별 이용시간(평일)”의 기본 키는 매체 그룹임을 알 수가 있다.

1. **각 매체그룹들에 대한 매체 그룹별 광고 이미지 평가는 5가지 평가기준(재미있다, 신뢰감을 준다, 제품 정보를 알게 해 준다, 성가시다, 자극적/선정적이다)을 유지해야한다.**

개체: 매체 그룹별 광고 이미지 평가

속성: 매체그룹(TV광고, 라디오 광고 등등 여러 개), 5가지 평가 기준(재미있다, 신뢰감을 준다, 제품 정보를 알게 해 준다, 성가시다, 자극적/선정적이다)

1. **매체 그룹별 광고 이미지 평가는 매체그룹으로 식별한다.**

개체 “매체 그룹별 광고 이미지 평가”의 기본 키는 매체 그룹(TV광고, 라디오 광고 등등 여러 개)임을 알 수가 있다.

1. **(TV, 라디오 제외)매체 그룹의 단가표는 매체, 비히클별 단가 정보를 유지해야한다.**

개체: 매체 그룹의 단가표(TV, 라디오 제외)

속성: TV, 라디오를 제외한 매체들(신문, 잡지, SNS 등등), 비히클별 단가

1. **(TV, 라디오 제외)매체 그룹의 단가표는 매체로 식별한다.**

개체 “매체 그룹의 단가표(TV, 라디오 제외)”의 기본 키는 매체임을 알 수 있다.

1. **(TV-지상파, 종합편성, 케이블/라디오)매체 그룹의 단가표의 경우 지상파 채널 (KBS, MBC, SBS), 종합편성 채널(TV CHOSUN, JTBC, CHANNEL\_A), 케이블 채널(OCN, SPOTV, YTN), 라디오 채널(KBS RADIO, MBC RADIO, SBS RADIO)의 시간대별 단가 정보를 유지해야 한다.**

개체: 매체의 시간대별 단가(TV, 라디오)

속성: TV(지상파, 종합편성, 케이블)채널별 단가, 라디오 채널별 단가, 시간대

1. **(TV-지상파, 종합편성, 케이블/라디오)매체 그룹의 단가표의 경우 광고 시간대로 식별한다.**

개체 “(TV-지상파, 종합편성, 케이블/라디오)매체 그룹의 단가표”의 기본 키는 광고 시간대임을 알 수 있다.

1. **광고 집중도는 매체 그룹, 매체 별 광고관심도와 광고 주목도를 유지해야한다.**

개체: 광고 집중도

속성: 매체 그룹, 광고 관심도, 광고 주목도

1. **광고 집중도는 매체 그룹으로 식별한다.**

개체 “광고 집중도”의 기본 키는 매체 그룹임을 알 수 있다.

1. **광고 매체 그룹은 매체와 매체 그룹 데이터를 유지해야 한다.**

개체: 광고 매체 그룹

속성: 매체, 매체 그룹

1. **광고 매체 그룹은 매체로 식별한다.**

개체 “광고 매체 그룹”의 기본 키는 매체임을 알 수 있다.

1. **TV매체, 라디오 매체의 시청률표는 12개의 채널에 대한 시간대별 시청률을 유지해야 한다.**

개체: TV, 라디오 매체 시청률표

속성: 시간대, TV 채널별(공중파, 종편, 케이블) 시청률, 라디오 채널별 청취율

1. **TV매체, 라디오 매체의 시청률표는 시간대로 식별한다.**

개체 “TV매체, 라디오 매체의 시청률표”의 기본키는 시간대임을 알 수 있다.

1. **TV매체, 라디오 매체의 단가표는 12개의 채널에 대한 시간대별 단가를 유지해야 한다.**

개체: TV, 라디오 매체 단가표

속성: 시간대, TV 채널별(공중파, 종편, 케이블) 단가, 라디오 채널별 단가

1. **TV매체, 라디오 매체의 단가표는 시간대로 식별한다.**

개체 “TV매체, 라디오 매체의 단가표”의 기본키는 시간대임을 알 수 있다.

1. **성별과 연령대에 따른 제품의 광고 매체 적합도를 측정하기 위해 품목별 광고 관심도, 매체별 인지도, 호감도, 구매결정 영향력, 광고 이미지 평가, 광고 집중도, 이용시간 데이터를 사용한다.**

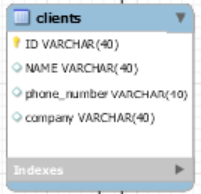
개체: 광고 매체 적합도

속성: 제품, 매체, 타겟별 광고 매체 적합도

1. **의뢰의 타겟 성별과 연령대가 특정되지 않을 경우 광고 매체 적합도를 통해 최적의 광고 매체와 광고 타겟을 선정한다.**
2. **의뢰의 타겟 성별과 연령대가 특정 지어질 경우 광고 매체 적합도를 통해 최적의 광고 매체를 선정한다.** 
   * 1. **엔티티 구현**

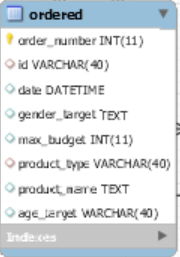
Entity는 데이터베이스에서 표현하려고 하는 개체로서 서로 구별할 수 있으며, 업무를 수행하는 데 필요한 데이터를 특성이 유사한 것끼리 모아 놓은 독립적인 집합이다. 하나의 Entity는 하나 이상의 attribute로 구성되며, 각 속성이 Entity의 상태를 기술해 준다. 이 집합을 어떻게 정의하는가에 따라 Entity를 결정할 수 있다.

1. **Clients(회원) 엔티티**



광고를 의뢰하는 고객의 기본정보를 유지하는 테이블이다. 속성은 ID와 고객의 이름, 전화번호, 소속회사를 나타내는 열로 나누었다. 기본키는 ID로 설정했다.

1. **Ordered(의뢰서) 엔티티**



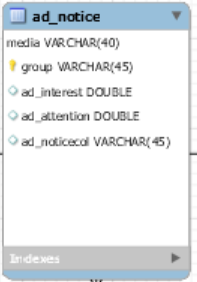
고객으로부터 의뢰받은 정보를 유지하는 테이블이다. 속성은 의뢰 번호와 의뢰 고객의 id, 의뢰 날짜, 성별과 연령대 타겟, 품목분류와 품목명, 최대 예산을 나타내는 열로 나누었다. 기본키는 order\_number로 설정했다.

1. **Fit(광고 적합도) 엔티티**



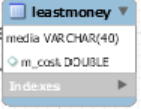
Fit 테이블은 품목, 매체와 연령, 성별에 따른 광고 적합도를 나타낸 테이블이다. 속성은 연령, 성별 광고 적합도이고, 기본키는 품목과 매체 복합키로 설정했다.

1. **Ad\_notice(광고 관심, 주목도) 엔티티**



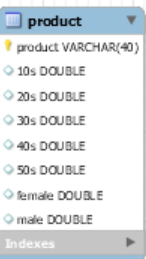
Ad\_notice 테이블은 광고 매체에 따른 소지바의 광고 관심도와 광고 집중도를 나타낸 그래프이다. 속성은 매체와 매체 그룹, 광고 관심도와 광고 집둥도로 설정했다. 기본키는 매체로 설정했다.

1. **Least money(최소 광고 예산) 엔티티**



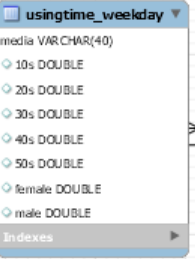
Least money 테이블은 각 매체의 최소 광고 비용을 나타낸 그래프이다. 속성은 최소 광고 비용으로 설정하였고, 기본키는 매체로 설정했다.

1. **Product(품목별 광고 관심도) 엔티티**



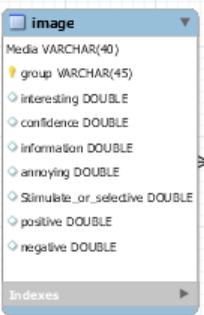
Product 테이블은 품목에 따른 소비자의 광고 관심도를 나타낸 테이블로, 품목과 연령대, 성별에 따른 광고 관심도를 나타낸다. 속성으로는 품목과 연령대별, 성별 광고 관심도이고, 기본키는 product로 설정했다.

1. **Usingtime\_weekday(매체별 이용률) 엔티티**



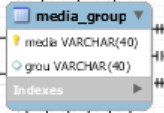
Usingtime\_weekday 테이블은 매체에 따른 연령별, 성별 이용률을 나타낸 테이블이다. 매체, 연령대, 성별을 속성으로 가지고 기본키는 매체이다.

1. **Image(매체별 이미지 평가) 엔티티**



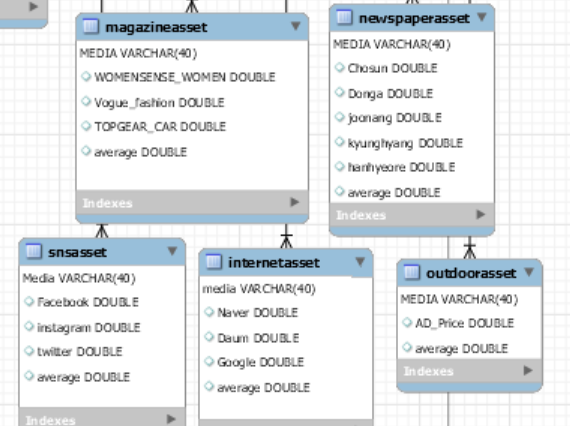
위의 테이블은 각 매체들의 광고가 사람들로부터 어떠한 평가를 받았는지를 나타내는 테이블이다. Group은 각 매체를 나타내고, media는 각 매체들에 관하여 좀더 세부적으로 들어갔을 때, 매체의 ppl 광고, sns광고 등 세부적으로 어떠한 것들이 있는지를 나타낸다. 기본 키 값은 속성 이름을 파란색으로 칠해서 나타내고, 위에서는 media이다.

1. **Media\_group (매체 그룹) 엔티티**



Media\_group 테이블은 매체와 매체 그룹을 속성으로 가지며 두 속성이 pk로 composite key 속성이다. 매체가 속한 매체그룹을 나타낸다.

1. **Magazineasset, newspaperasset, snsasset, internetasset, outdoorasset(잡지, 신문, SNS, 인터넷, 외부광고 단가) 엔티티**



Internet asset 테이블은 매체에 따른 인터넷 광고 단가를 나타내는 테이블이다. 기본 키 값은 속성 이름을 파란색으로 칠해서 나타내고, 기본키는 media이다.

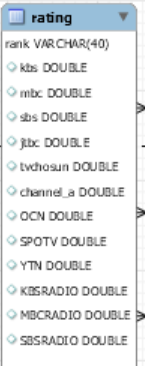
Magazine asset 테이블은 매체에 따른 잡지 광고 단가를 나타내는 테이블이다. 기본 키 값은 속성 이름을 파란색으로 칠해서 나타내고, 기본키는 media이다.

Newspaper\_asset 테이블은 매체에 따른 신문 광고 단가를 나타내는 테이블이다. 기본 키 값은 속성 이름을 파란색으로 칠해서 나타내고, 기본키는 media이다.

Outdoor\_asset 테이블은 Outdoor 그룹에 속한 매체의 광고 단가를 나타낸 테이블로 속성은 광고단가, 기본키는 media로 설정하였다.

SNS\_asset 테이블은 SNS 비히클별 SNS의 광고 단가를 나타낸 그래프이다. 각 비히클별 단가와 단가 평균을 속성으로 설정했고, 기본키는 Media로 설정했다.

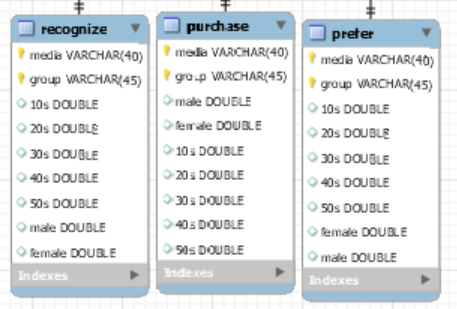
1. **Rating(시청률) 엔티티**



Rating 테이블은 tv, radio 채널의 시청률을 특정한 기준을 가진 rank별로 나누어

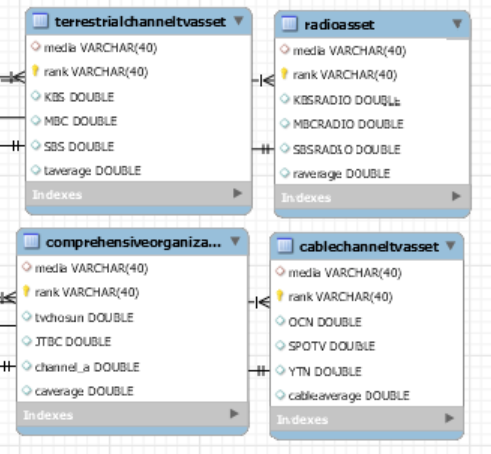
나타냈다. rank, 채널별 시청률을 속성으로 가지고 기본키는 rank이다.

1. **recognize, purchase, prefer(매체 인지,구매,호감) 엔티티**



세 테이블은 각각 매체와 연령, 성별에 따른 매체 호감도, 인지도, 구매 결정 영향력을 나타낸 테이블이다. 세 테이블 모두 같은 형식으로, 속성을 매체 그룹, 연령대별, 성별 특성치로 설정했다. 기본키는 media와 group 복합키로 설정했다.

1. **Terrestrial channel tv asset, radio asset, comprehensive organization channel tv asset, Cable channel asset(지상파, 라디오, 종편, 케이블 광고 단가) 엔티티**



Terrestrial\_channel\_tv\_asset 테이블은 랭크에 따른 지상파 TV 채널 비히클 별 광고 단가를 나타낸 테이블이다. 속성은 매체 그룹과 지상파 채널 별 광고 단가, 평균 광고 단가로 설정했으며 기본키는 rank로 설정했다.

Cable channel tv asset 테이블은 시간대로 구분되는 rank에 따른 케이블 tv 채널들의 단가를 나타내는 테이블이다. 속성은 rank와 매체, 케이블 채널 비히클별 광고 단가 정보로 설정했다. 기본키는 rank로 설정했다.

Comprehensive organization channel tv asset 테이블은 시간대로 구분되는 rank에 따른 종합편성채널 tv의 광고 단가를 나타내는 테이블이다. 속성은 rank와 매체, 종합편성채널 비히클별 광고 단가 정보로 설정했다. 기본키는 rank로 설정했다.

Radio\_asset 테이블은 Radio 비히클별 Radio의 광고 단가를 나타낸 그래프이다. 매체와 각 비히클별 단가, 단가 평균을 속성으로 설정했고, 기본키는 rank로 설정했다.

* + 1. **정규화 & 엔티티 통합 및 분리**

1. Client 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

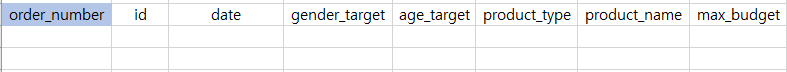
2NF: 릴레이션의 기본키인 ID를 제외한 모든 속성이 ID에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치 종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Ordered 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

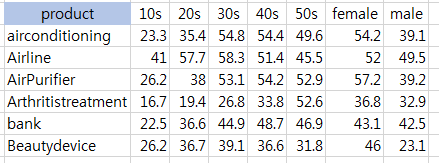
2NF: 릴레이션의 기본키인 order\_number를 제외한 모든 속성이 order\_number에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치 종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Product 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

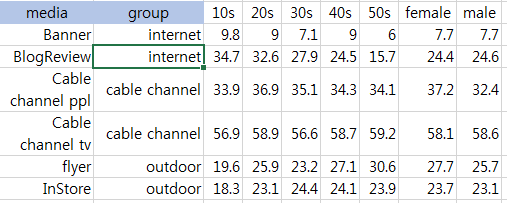
2NF: 릴레이션의 기본키인 Product를 제외한 모든 속성이 product에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

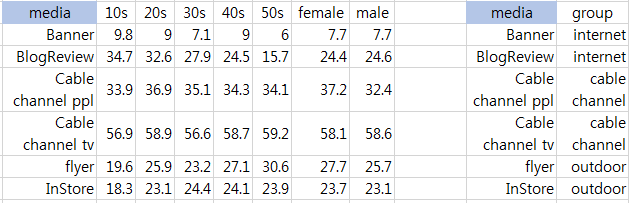
4NF: 릴레이션에 다치 종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. prefer, recognize, purchase 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션의 연령대와 성별 속성의 값이 기본키인 {media, group}에 완전 함수 종속되지 못하고 기본키의 일부분인 media에 부분 함수 종속되므로 2NF를 만족하지 않는다. 2NF를 만족하기 위해 부분 함수 종속을 제거한 릴레이션은 아래와 같다.

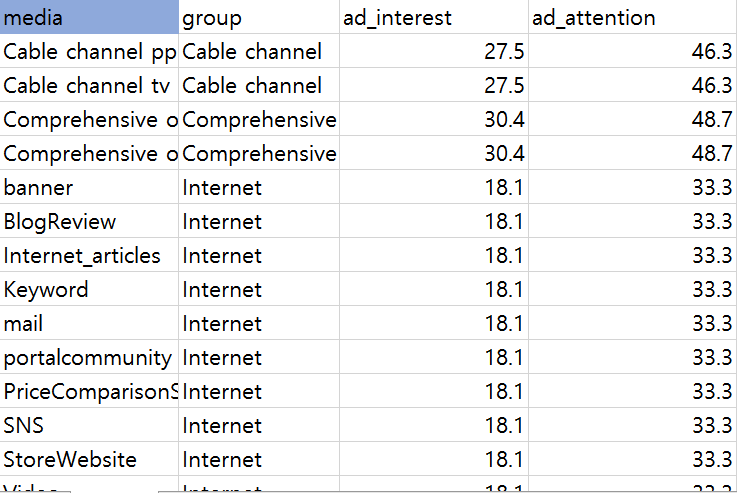


3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치 종속 관계가 존재하지 않으므로 4NF를 만족한다.

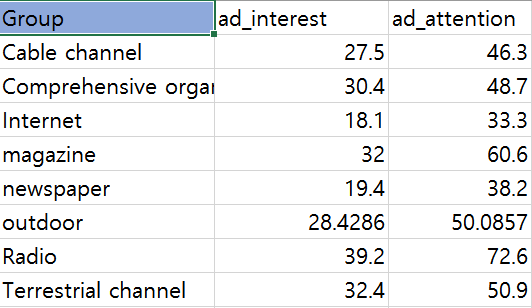
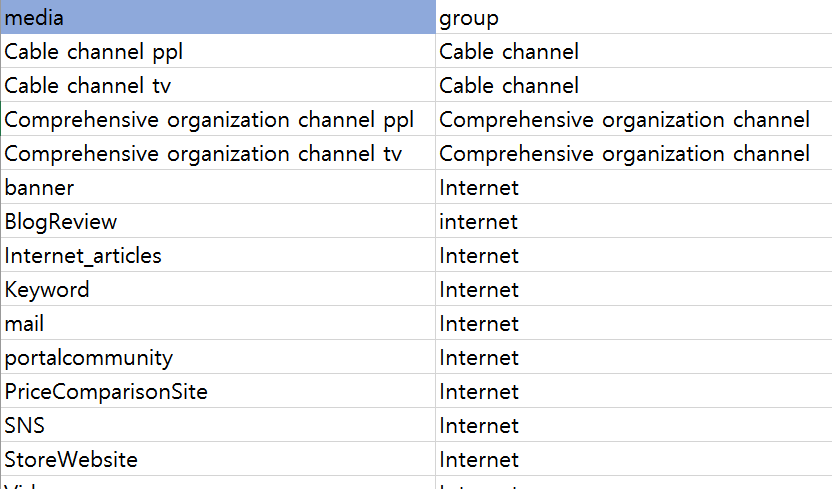
1. Ad\_notice 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

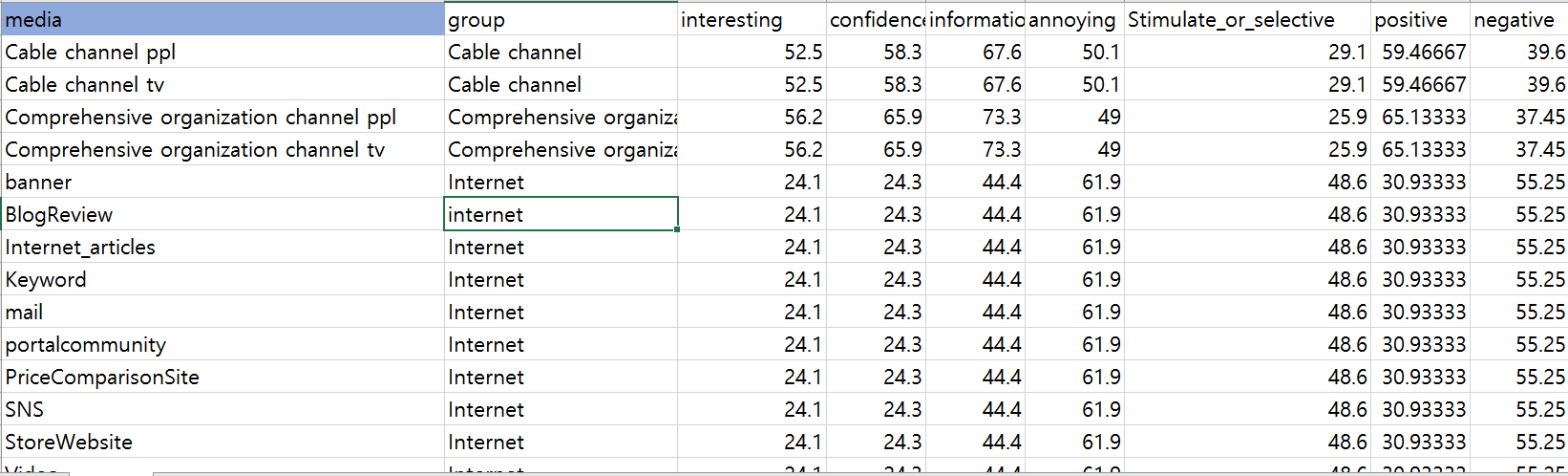
3NF: 위의 테이블은 각 매체에서 광고를 하였을 때 사람들이 얼마나 이를 주목하고 관심을 가졌는지를 나타내므로 ad\_interest, ad\_attention 속성들은 모두 기본 키가 아닌 속성인 group에 의하여 결정이 된다. 따라서 ad\_interest, ad\_attention 속성들은 media라는 기본 키에 이행적으로 함수 종속되어 있어서 3NF를 만족시키지 못해서 테이블을 분해하는 작업이 필요하다. 3NF를 만족하도록 테이블을 분해하면 다음과 같다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

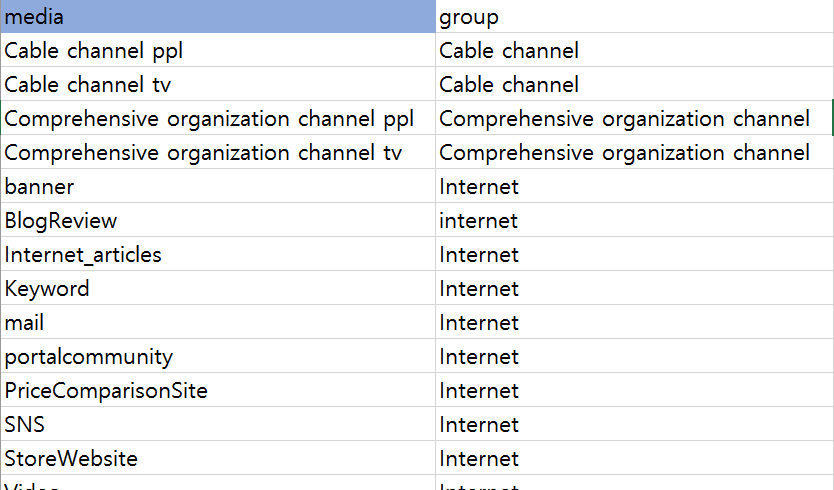
1. image 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 위의 테이블은 각 매체들의 광고가 사람들로부터 어떤 평가를 받았는지를 나타내므로 interesting~negative 평들은 모두 기본 키가 아닌 속성인 group에 의하여 결정이 된다. 따라서 이미지 평가와 관련된 속성들은 media라는 기본 키에 이행적으로 함수 종속 되어 있어서 3NF를 만족 시키지 못해서 테이블을 분해하는 작업이 필요하다. 3NF를 만족하도록 테이블을 분해하면 다음과 같다.

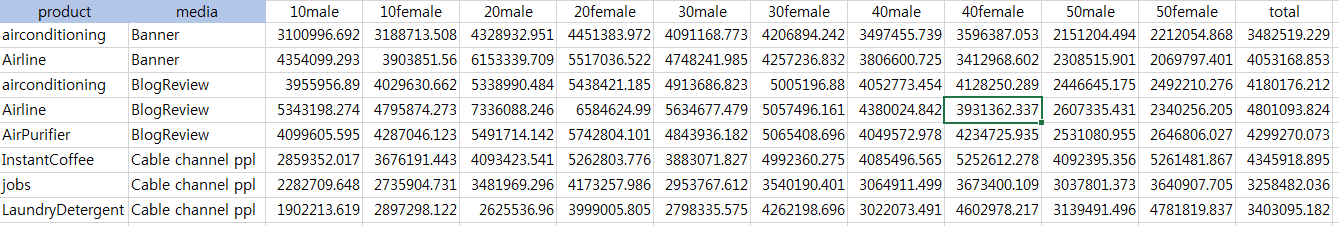




BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Fit 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션의 기본키인 {product, media}를 제외한 모든 속성이 기본키의 일부분에 부분 함수 종속되지 않고 기본키 전체에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Leastmoney 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

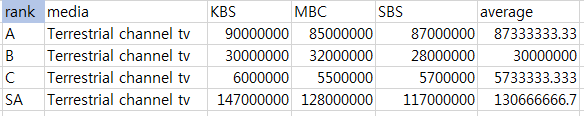
2NF: 릴레이션의 기본키인 media를 제외한 m\_cost 속성이 기본키에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Terrestrial\_channel\_tv\_asset 테이블.



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

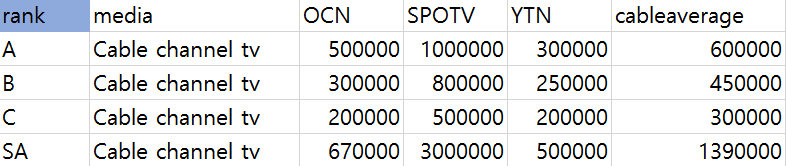
2NF: 릴레이션의 기본키인 media를 제외한 광고 단가 속성이 기본키에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Cable channel tv asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

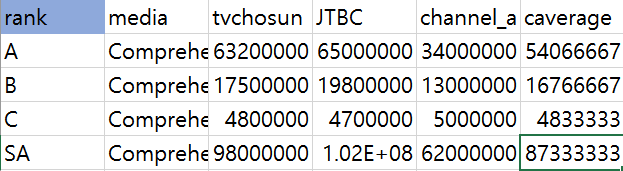
2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Comprehensive organization channel tv asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

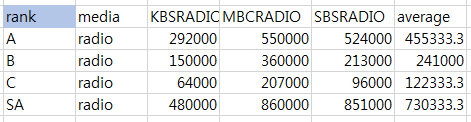
2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Radio\_asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션의 기본키인 media를 제외한 비히클 별 단가 속성들이 기본키에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Internet\_asset



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

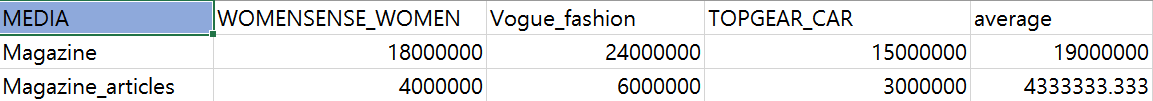
2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Magazine\_asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

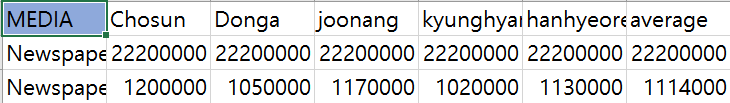
2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Newspaper asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해되지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

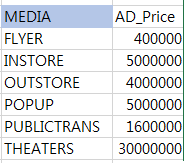
2NF: 릴레이션이 기본 키를 오직 1개만 가지므로 기본 키가 아닌 모든 속성은 기본 키에 완전 함수 종속이 된다. 따라서 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. Outdoor\_asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

2NF: 릴레이션의 기본키인 media를 제외한 광고 단가 속성이 기본키에 완전 함수 종속되므로 2NF를 만족한다. r

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

1. SNS\_asset 테이블



1NF: 릴레이션의 모든 속성이 더는 분해하지 않는 원자 값을 가진다. 따라서 1NF를 만족한다.

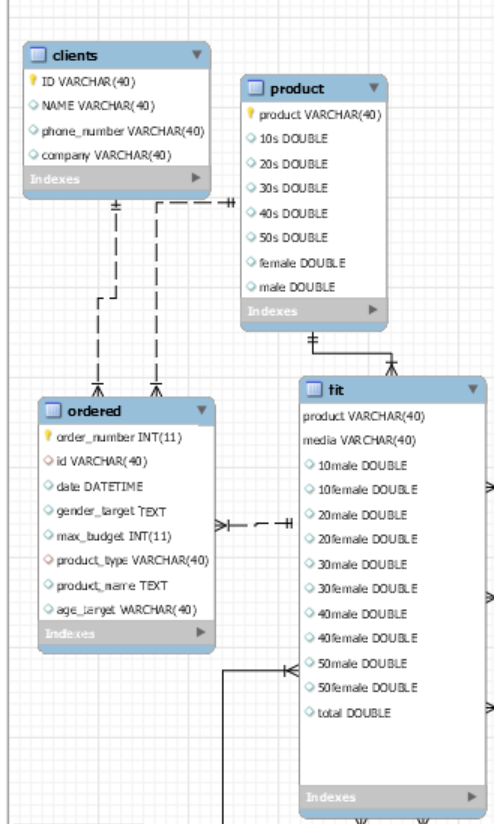
2NF: 릴레이션의 기본키인 media를 제외한 비히클 별 단가 속성들이 기본키에 완전 함수 종속되므로 2NF를 만족한다.

3NF: 릴레이션에 이행적 함수 종속 관계가 존재하지 않으므로 3NF를 만족한다.

BCNF: 릴레이션에 후보키가 아닌 결정자가 존재하지 않으므로 BCNF를 만족한다.

4NF: 릴레이션에 다치종속 관계가 존재하지 않으므로 4NF를 만족한다.

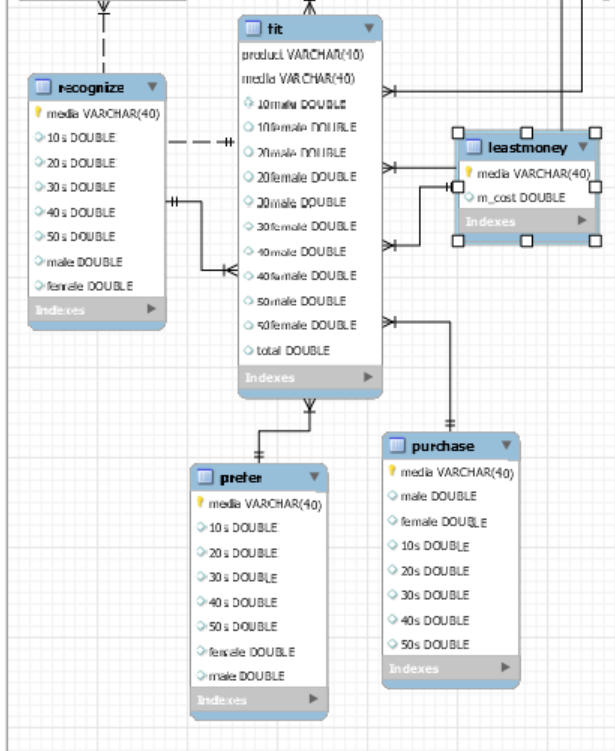
* + 1. **ER Modeling**
       1. **관계 설정**



Clients는 고객에 관한 정보를 담고 있는 테이블이고 ordered는 고객의 의뢰서에 관한 정보를 담은 테이블이다. Client의 pk값인 id가 ordered 테이블의 pk값으로 참조 되지 않았기에 weak한 관계이고 non-ID Dependent 하다. 하나의 Client가 여러 개의 의뢰서를 요청할 수 있기에 1:N 관계이다.

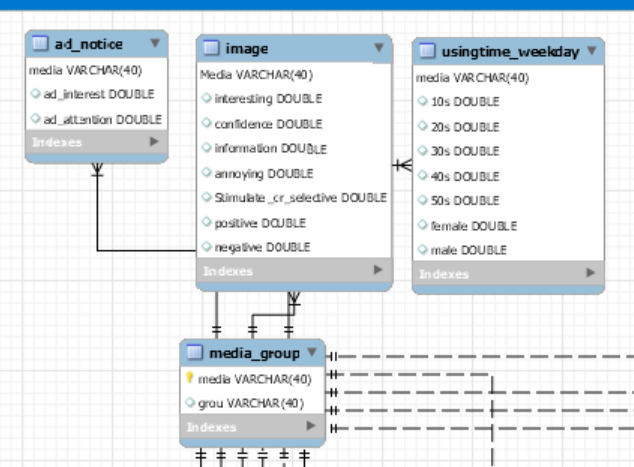
Fit (적합도) 테이블은 10대~50대 남녀별 광고 매체 적합도 정보를 담고 있는 테이블이다. Product과 media가 composite 키로 pk키의 역할을 한다. Fit 테이블과 ordered 테이블은 weak한 관계이고 non-ID Dependent하다. 하나의 fit (적합도)는 여러 개의 의뢰서에 해당될 수 있기에 1:N 관계이다.

Product 테이블은 10대~50대 남녀의 광고 품목별 광고 관심도 정보를 포함하고 있다. Fit 테이블과는 pk값이 pk값으로 참조되기에 strong하고 ID-Dependent하다. Ordered 테이블과는 weak한 관계이고 non-ID Dependent 하다. 품목별 광고 관심도로 여러 개의 광고 적합도를 얻을 수 있기에 FIT 테이블과 1:N 관계이다. 품목별 광고 관심도가 여러 개의 의뢰서에 해당 될 수 있기에 Ordered 테이블과 1:N 관계이다.



Recognize 테이블은 10대~50대 남녀의 광고 매체에 대한 인지도 prefer 테이블은 10~50대 남녀의 광고 매체에 대한 선호도 purchase 테이블은 10~50대 남녀의 광고 매체에 대한 구매결정도에 관한 데이터를 포함한다. 이 세 수치 테이블의 pk값이 fit테이블의 pk값으로 참조되었기에 strong하고 ID-Dependent하다. 세 수치 테이블로부터 여러 개의 적합도 테이블을 얻을 수 있기에 모두 1:N 관계이다.

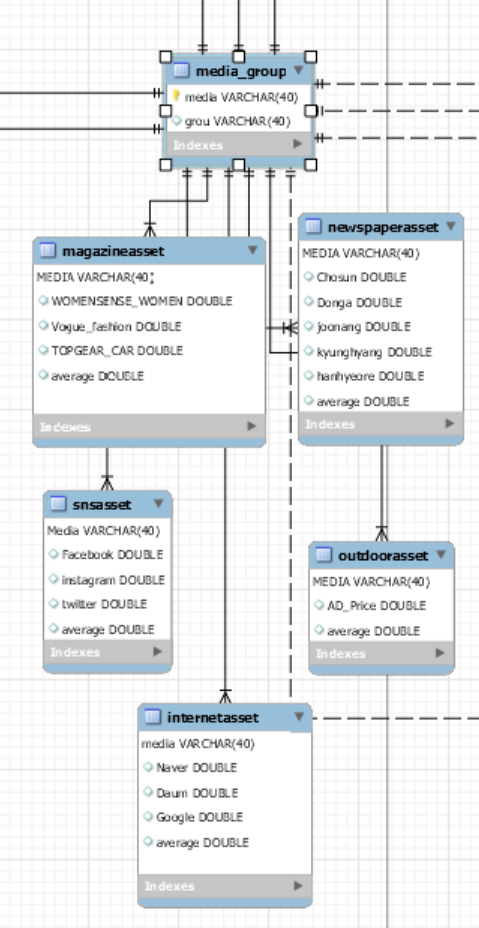
Least money 테이블은 광고 매체들의 최저 광고 제작예산에 대한 정보를 포함한다. pk값이 fit테이블의 pk값으로 참조 되었기에 strong하고 ID-Dependent하다. 최저 광고 제작예산은 열러 광고 적합도를 결정 하는데 쓰이기에 1:N 관계이다.



Media group 테이블은 매체와 매체 그룹에 대한 정보를 포함한다.

Ad\_notice 테이블은 광고 그룹에 대한 광고 관심, 주목도 정보를 포함한다. Image 테이블은 광고 그룹에 대한 5가지 평가기준(재미있다, 신뢰감을 준다, 제품 정보를 알게 해 준다, 성가시다, 자극적/선정적이다) 정보를 포함한다. Usingtime\_weekday 테이블은 10대~50대 남녀의 광고 그룹별 사용시간 정보를 포함한다.

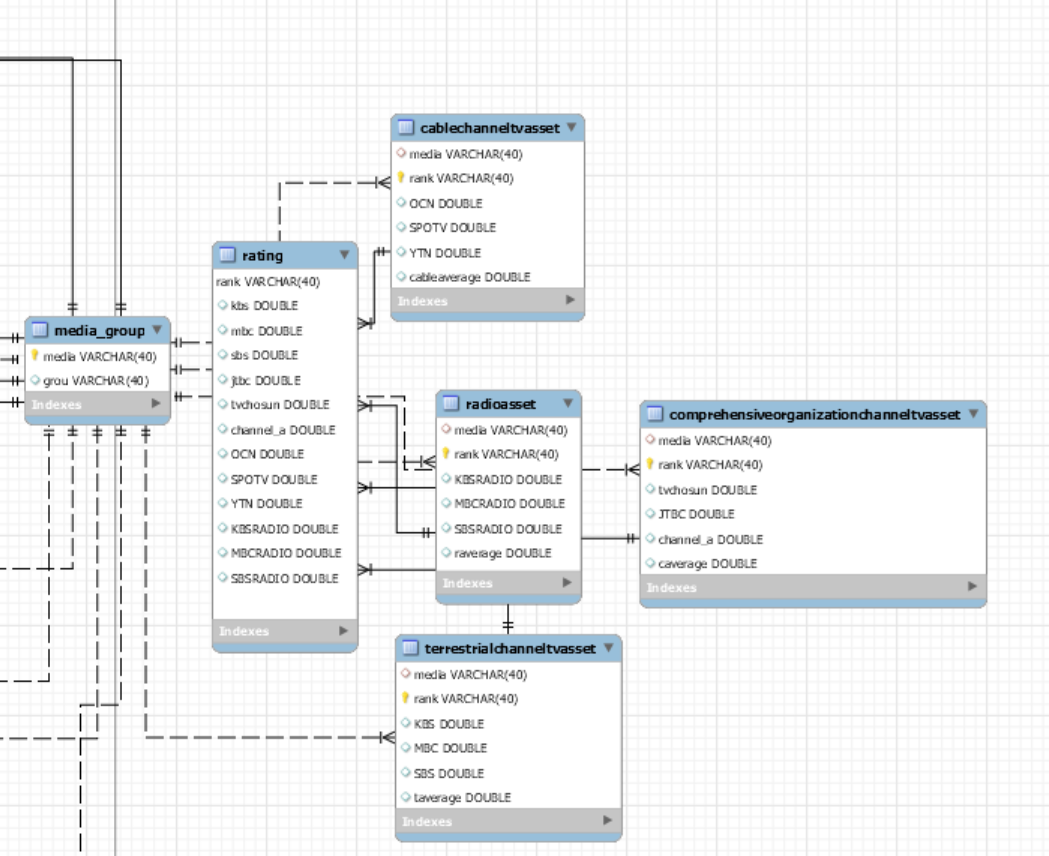
Media group 테이블의 pk값은 ad\_notice, image, usingtime\_weekday 테이블의 pk값으로 참조 되기에 테이블 들의 관계는 strong하고 ID-Dependent하다. 하나의 media\_group은 여러 개의 ad\_notice, image, usingtime\_weekday 정보를 가질 수 있기에 모두 1:N 관계이다.



Outdoor asset, SNS asset, Internet asset, Newspaper asset,, magazine asset 테이블은 각 각 매체 그룹의 비히클의 단가에 관한 정보를 포함한다.

Media\_group 테이블에서 pk값을 참조 받기에 테이블 들의 관계는 strong하고 ID-Dependent하다.

하나의 media\_group에 여러 단가 테이블이 해당될 수 있기에 media\_group과 Outdoor asset, SNS asset, Internet asset, Newspaper asset,, magazine asset 테이블은 1: N 관계이다.



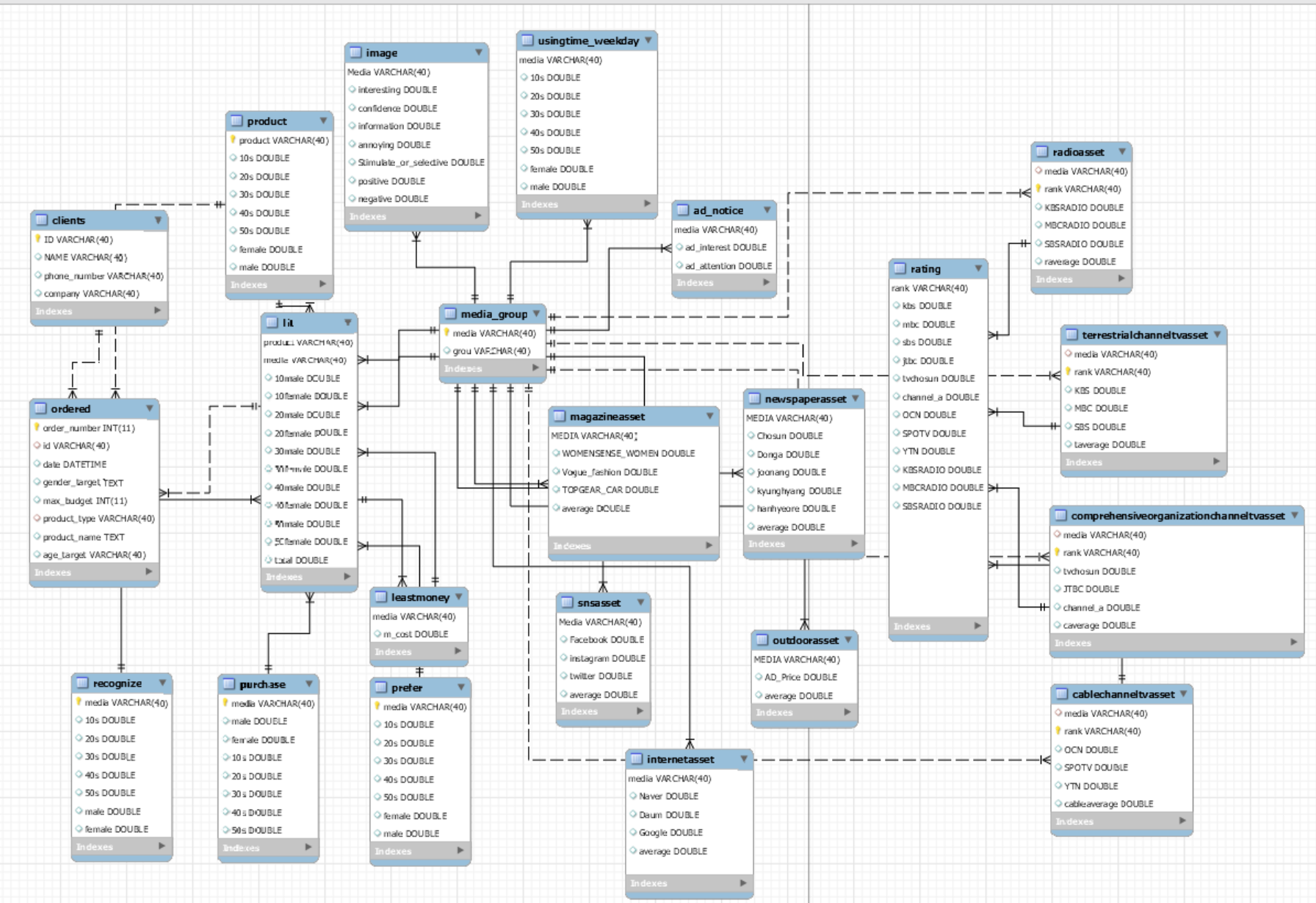
Rating 테이블은 tv 채널과 라디오 채널의 시간대 Rank 별 시청률 정보를 포함한다. Cablechanneltv\_asset, terestrialchanneltv\_asset, comprehensive organizationchanneltv\_asset , radio\_asset 테이블은 시간대 Rank 별 비히클 채널들의 단가 정보를 포함한다. Rating 테이블은 Cablechanneltv\_asset, terestrialchanneltv\_asset, comprehensive organizationchanneltv\_asset, radio\_asset 테이블의 pk값을 참조 받기에 테이블 들의 관계는 strong하고 ID-Dependent하다.

Cablechanneltv\_asset, terestrialchanneltv\_asset, comprehensive organizationchanneltv\_asset, radio\_asset 단가 테이블들은 여러 rating 시청률 테이블을 가질 수 있기에 모두 1:N 관계이다.

Cablechanneltv\_asset, terestrialchanneltv\_asset, comprehensive organizationchanneltv\_asset, radio\_asset 단가 테이블들은 media\_group 테이블의 pk 값을 pk값으로 참조 받지 않기에 테이블들은 weak한 관계이고 non-ID Dependent하다. 하나의 media\_group은 여러 단가 테이블을 가질수 있기에 media\_group테이블과 Cablechanneltv\_asset, terestrialchanneltv\_asset, comprehensive organizationchanneltv\_asset, radio\_asset 단가 테이블들은 1:N 관계이다.

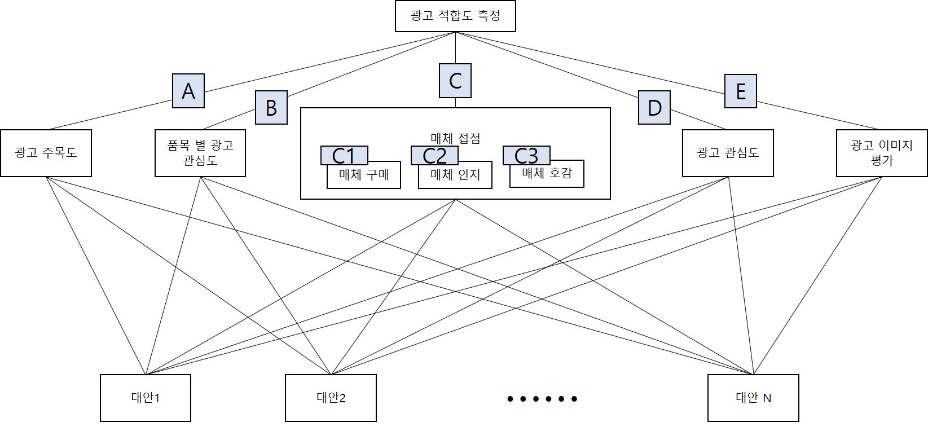
* + - 1. **Extended ER model**

다음 사진은 ER Model의 전체적인 모습이다.



* 1. **계층화 의사결정법(AHP)을 이용한 가중치 선정**

AHP(Analytic Hierarchy Process)은 1971년 Saaty에 의해 만들어진, 설정한 목표에 대하여 세부 항목들로 계층화를 해서 의사결정을 하는 기법으로, 행렬을 사용하여 각 평가 기준들에 대한 가중치를 부여하는 기법이다. 각 평가 항목들에 가중치를 부여하기 이전에, 어떠한 하나의 의사결정 문제에 대하여 계층화를 하여 평가 기준들을 몇 개 선정하는데, 여러 대안들에 있어서 이 평가 기준들이 어느 정도로 중요한지에 대한 상대적 중요도를 비교하고, 또한 평가 기준들이 의사 결정 문제에 대하여 상대적으로 어느 정도 중요한지 알아보기 위해 쌍대비교를 통해 평가 기준들 간의 상대적 중요도를 측정한다. 다음 그림은 광고 적합도 측정 문제를 계층화하여 나타낸 모습이다.



광고할 매체를 추천해주는 데이터 베이스를 구축하기 위하여 여러 데이터들을 검색하는 과정에서, 우리 조원들은 어떠한 매체가 광고를 하기에 적합하다고 판단하는 데에는 광고 주목도, 품목 별 광고 관심도, 매체 접점, 광고 관심도, 광고 이미지 평가 항목들이 영향을 미칠 거라고 추측할 수가 있었다. 또한 ‘사람들이 얼마나 매체를 접하였는가’를 나타내는 매체 접점 항목은, 얼마나 많은 사람들이 매체 광고를 통하여 제품을 구매하였는지를 나타내는 매체 구매 항목, 매체 호감 항목, 매체 광고를 얼마나 많은 사람들이 인지하였는지를 나타내는 매체 인지 항목, 이 3가지가 영향을 미칠 것이라 판단하였고, 그리하여 이 3가지 항목을 매체 접점 항목의 하위 계층에 두게 되었습니다(그래서 사실 상 다른 항목들처럼 선으로 표시해야 되나, 선으로 표시하지 않고 위와 같이 표현하는 것이 덜 복잡해 보일 것이라 생각되어 선 없이 표현하였다). A~E는 ‘광고 적합도’를 판단하는 데에 5가지 평가 기준들이 어느 정도 비중이 있는지를 나타내는 가중치다(A+B+C+D+E=1). 마찬가지로 C1~C3은 ‘매체 접점’이 어느 정도인지 판단하는 데에 3가지 평가 기준들이 어느 정도 비중이 있는지를 나타내는 가중치다(C1+C2+C3=1, 따라서 ‘광고 적합도’를 판단하는 데에 매체 구매는 C\*C1정도의 가중치를, 매체 인지는 C\*C2의 가중치를, 매체 호감은 C\*C3의 가중치를 갖는다). 우리는 이러한 AHP 기법을 python을 통해 구현을 하여서 각 항목의 가중치를 구하고자 했다.

* + 1. **AHP 기법의 기본 절차**

1. 각 계층 내 쌍대비교 설문 실시

앞서 본 그림처럼 주어진 목표(광고 적합도 측정)를 계층 구조로 분해하였다면, 같은 계층에 있는 요소들끼리 쌍대 비교를 하는 과정을 거쳐야 한다(). 즉, 한 계층에서 서로 다른 두 요소끼리 1대 1로 비교를 해서 어느 요소가 다른 요소에 비해 얼마나 중요하다고 생각되는지를 1~9 사이의 점수를 매겨서 상대적 중요도를 결정하게 되는 것이다. 이 때, 이러한 중요도를 결정하기 위해 적당한 인원의 사람들을 표본으로 삼아서 설문 조사를 집행한다. 아래 쌍대 비교표와 같이 같은 계층의 두 항목 A, B 중에서 두 항목의 상위 항목(상위 계층에 있는)에 영향을 더 많이 미치는, 다시 말해 중요도가 더 높다고 생각 되는 항목에 중요도 점수를 매기는 것이다.



두 항목의 중요도가 동등하다고 생각되는 경우에는 1점, 한 항목의 중요도가 약간 더 높다고 생각되는 경우 3점, 한 항목의 중요도가 상당한 차이가 있다고 생각되는 경우 5점, 한 항목이 확실하게 더 중요하다 생각되는 경우 7점, 마지막으로 한 요소가 절대적으로 더 중요하다고 여겨질 때는 9점을 매기게 되는 것입니다(2,4,6,8은 1,3,5,7,9 사이의 중요도라고 판단되는 경우). 더 중요하다고 판단되는 항목에 2~9 사이의 점수를 주고, 덜 중요한 항목에 대해서는 더 중요한 항목에 준 점수의 역수를 점수로 매긴다. 이렇게 한 계층의 모든 항목(5개)들에 대하여

총 5C2 = 10 번 진행하게 된다. 모든 항목들에 대하여 쌍대 비교를 완료하였다면, 매긴 점수들을 토대로 쌍대 비교 행렬을 작성하게 된다. 이 때, 만일 A항목과 B항목을 비교했을 때, A항목이 약간 더 중요도가 높다고 판단했을 때, (A,B)=3, (B,A)=1/3을 행렬에 넣게 되는 것이다.

(쌍대 비교 행렬) = [AIJ], (i=1,…n, j=1,…,n, Aji = 1/Aij)

한 명의 사람에 대해서만 설문 조사를 한다면 그 사람이 매긴 점수에 관한 쌍대 비교 행렬을 AHP분석에 바로 사용하면 되지만, 다수의 사람에 대하여 설문 조사를 할 경우에는 각각의 쌍대 비교 행렬의 모든 요소들의 산술평균이나 기하평균을 계산하여 나온 새로운 행렬로 AHP분석을 하게 된다. 우리 조는 10명을 대상으로 설문조사를 하였고, 각 쌍대 비교 행렬의 모든 요소들을 기하평균으로 계산하였다.

* 기하평균이란?

n개의 양수 A1, A2,…,An의 기하평균을 G라고 하였을 때,

G = 이다.

1. 고유치 방법(Eigenvalue Method)을 이용하여 쌍대 비교된 요소들의 가중치 추정

쌍대 비교를 완료하였다면, 쌍대 비교 행렬을 이용하여 각 요소들의 가중치를 추정한다. 쌍대 비교 행렬(A)을 자체 행렬 곱을 한 뒤, 의 각 행의 합계와 모든 요소의 합계를 구합니다. 이를 통해 eigen vector를 구하는데, eigen vector는 각 항목에 대한 가중치를 의미합니다. 즉, 해당 항목의 가중치(eigen vector)를 구하기 위해서는 (행렬의 해당 항목의 행의 합계)/(행렬의 모든 요소의 합계)를 구하면 되고, 모든 항목들의 eigen vector들의 합은 1이 된다.

1. 일관성 검토

각 항목에 대한 가중치를 구했다면, 마무리 단계로 해당 설문 조사가 일관성 있게 진행되었는지를 검토해야 한다. CI/RI=CR 값이 나왔을 때, CR 값이 0.1을 넘는 경우에 일관성이 없다고 판단을 하고, 이러한 결과가 나왔을 때 설문 조사를 다시 실시해야 하는 것입니다. CI(일관성지수)는 쌍대 비교 행렬의 각 열에 구한 가중치 값들을 각각 곱한 행렬(B)을 구해서 B 행렬의 각 행의 합계를 구하고, 구한 행의 합계들을 각 항목의 가중치들로 나눈다. 그렇게 나눈 값들의 평균을 구하면 lambda max값이 나오는데, 이 때 CI = (n=항목의 개수) 가 되는 것이다. 그리고 RI는 무작위 일관성 지수(Random consistency Index)를 말하며 1에서 9까지 정수들을 무작위 추출하여 역수 행렬을 작성한 후 이로부터 일치지수를 구한 것을 말하며, Satty에 의해 행렬의 크기에 따른 RI 값이 정리되어 있어 이를 참고하였다.

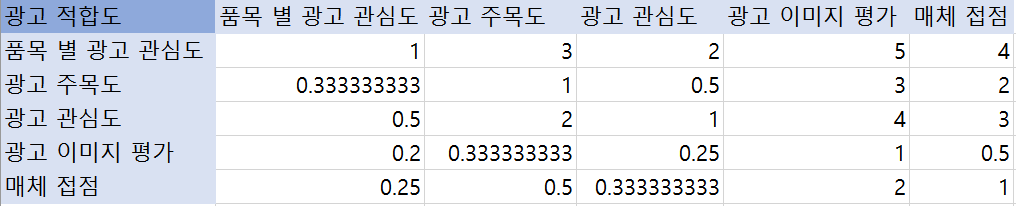
* + 1. **실습에서의 AHP 기법 적용**

1. 각 계층 내 쌍대비교 설문 실시

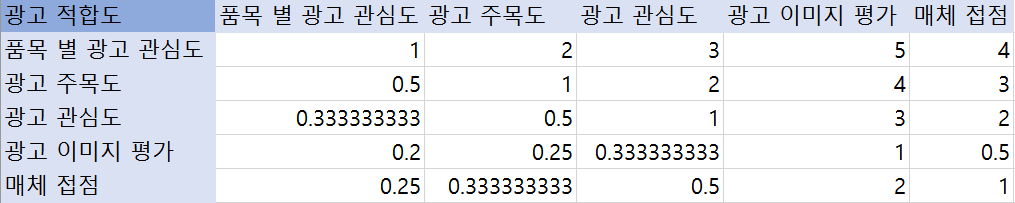
우리 조원들은 이러한 설문 조사를 10명의 사람들을 표본으로 뽑아서 진행하게 되었다. 10명에 대하여 설문 조사를 진행해 본 결과는 다음과 같았다.

① 계층 1(광고 주목도, 품목 별 광고 관심도, 매체 접점, 광고 관심도, 광고 이미지 평가)의 요소 간의 쌍대 비교

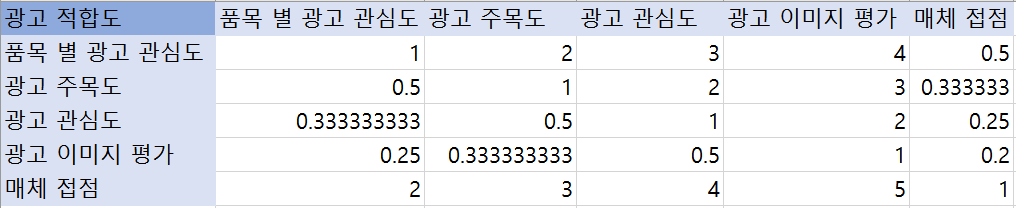
-Survey1



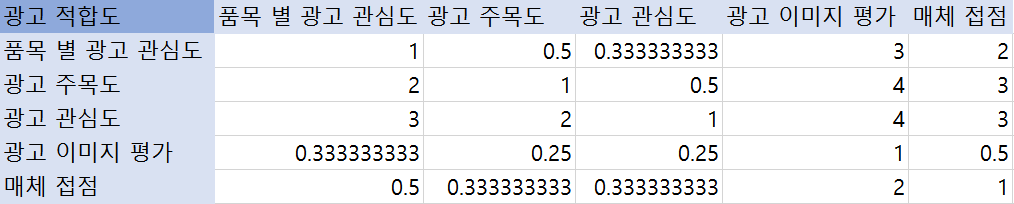
-Survey2



-Survey3



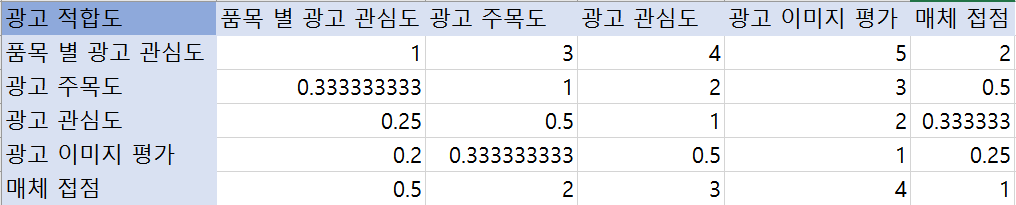
-Survey4



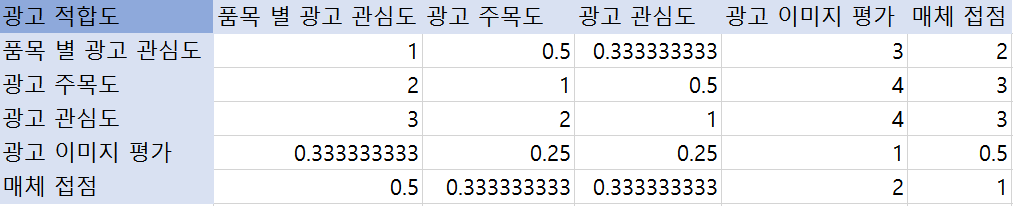
-Survey5



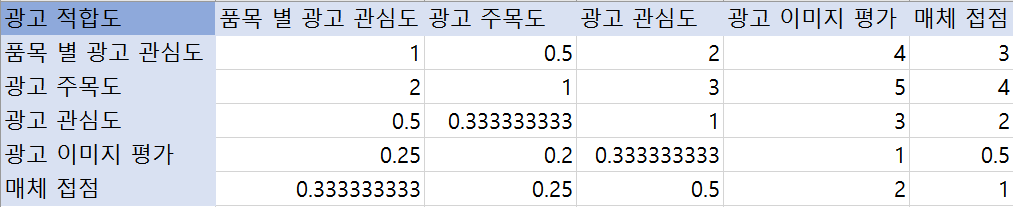
-Survey6



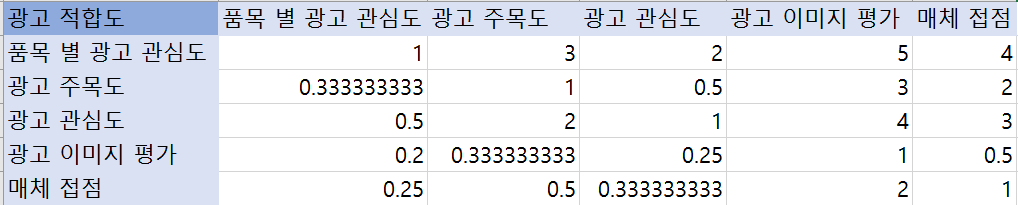
-Survey7



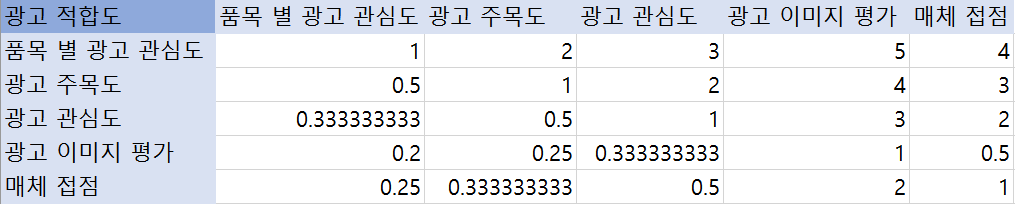
-Survey8



-Survey9



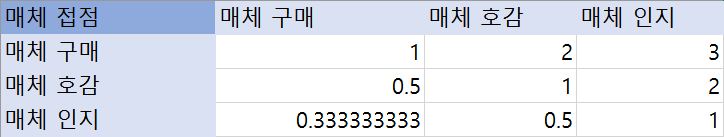
-Survey10



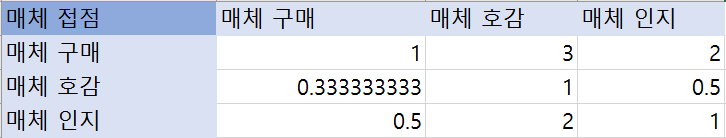
한 명의 사람에 대해서만 설문 조사를 한다면 그 사람이 매긴 점수에 관한 쌍대 비교 행렬을 AHP분석에 바로 사용하면 되지만, 다수의 사람에 대하여 설문 조사를 할 경우에는 각각의 쌍대 비교 행렬의 모든 요소들의 산술평균이나 기하평균을 계산하여 나온 새로운 행렬로 AHP분석을 하게 된다. 우리 조는 위의 10개의 행렬에 대한 각 요소의 기하평균을 나타내는 새로운 행렬을 만들어서 AHP분석을 시행하였다.

② 계층 2(매체 구매, 매체 인지, 매체 호감)의 요소 간의 쌍대 비교

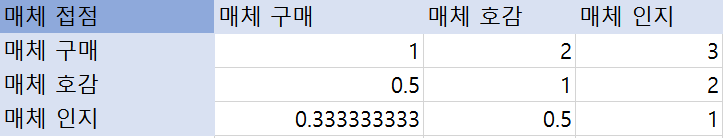
-Survey1



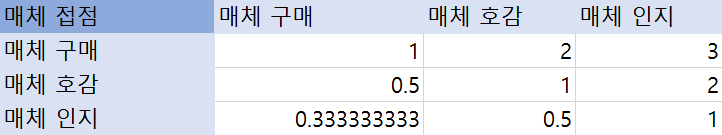
-Survey2



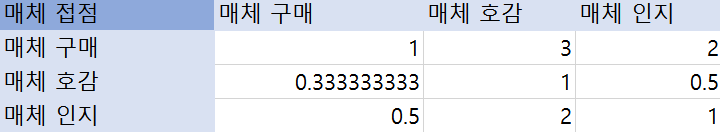
-Survey3



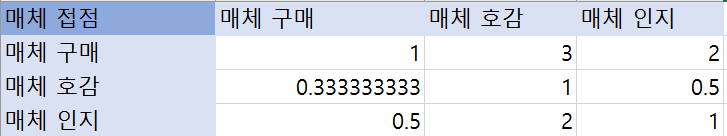
-Survey4



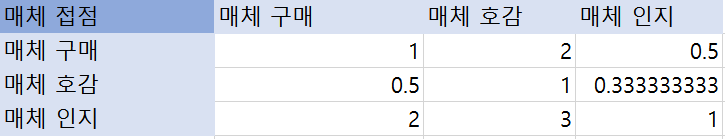
-Survey5



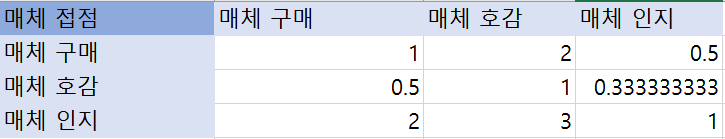
-Survey6



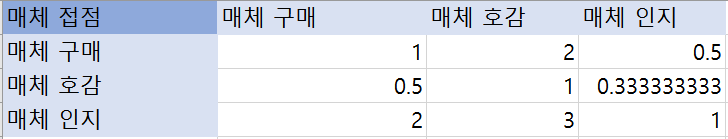
-Survey7



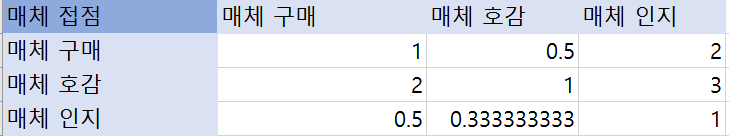
-Survey8



-Survey9

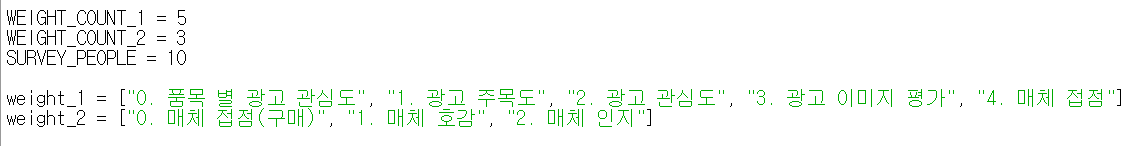


-Survey10



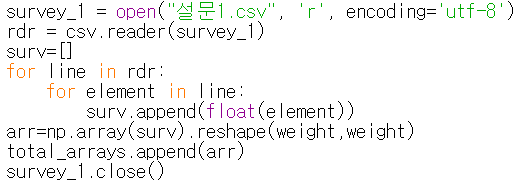
이렇게 10명에게 설문 조사를 한 결과를 excel에 정리한 뒤, 이를 csv 형태로 변환하여 파이썬에 넘겨준다. 파이썬에서 Compare 함수를 만들었는데, Compare함수는 가중치를 매길 항목의 리스트와 항목의 개수를 매개변수로 받아서 여러 쌍대 비교 행렬에 대한 기하 평균을 계산한 행렬을 반환하는 역할을 한다.





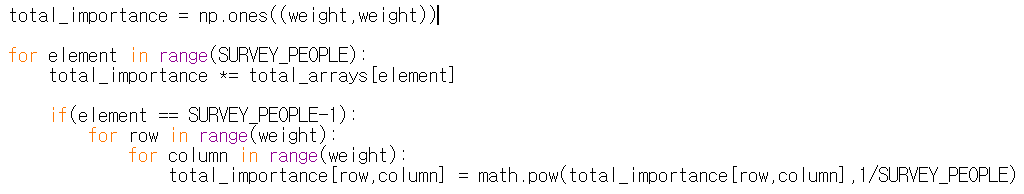


계층1과 계층2의 항목들과 항목의 개수를 전달한다.



(위의 코드를 설문 조사 20개에 대하여 만든다. 위에서 본 설문 조사에 관한 csv 파일들을 읽어와서 이를 파이썬에서 행렬(array)의 형식으로 변환하는 과정이다.)

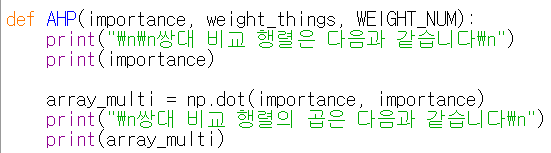
우리는 한명이 아닌 다수의 사람들에게 설문 조사를 진행하였으므로, 위에서 만들어진 각 행렬들(Survey\_1~10간의 평균, Survey\_11~20간의 평균)의 각 요소들의 기하 평균을 구해야 합니다.



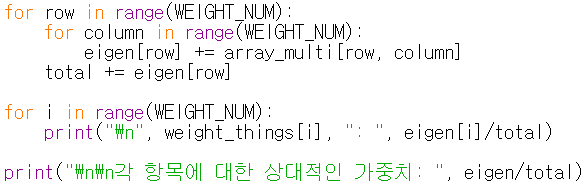
그리하여 각 요소들을 각각 모두 곱한 뒤, 1/(설문 조사를 진행한 사람 수) 를 제곱하여 기하 평균을 구한다.

1. 고유치 방법(Eigenvalue Method)을 이용하여 쌍대 비교된 요소들의 가중치 추정

AHP 함수에서는 Compare 함수에서 구한 쌍대 비교 행렬과, 쌍대 비교 행렬의 자체 행렬 곱을 한 행렬 값을 반환한다.

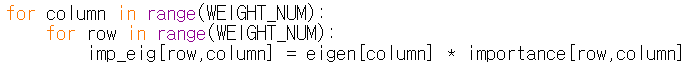


그런 다음 각 행의 합계와 모든 요소의 합계를 구해서 (각 행의 합계)/(모든 요소의 합계)를 각 항목에 대해 구해서 각 항목의 가중치를 다음과 같이 구해줍니다.



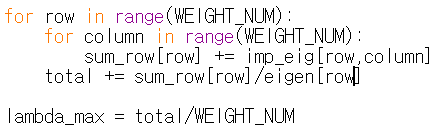
1. 일관성 검토

CI 값을 구하기 위해 쌍대 비교 행렬의 각 열과 가중치(eigen vector) 행렬 값을 각각 곱해준다.

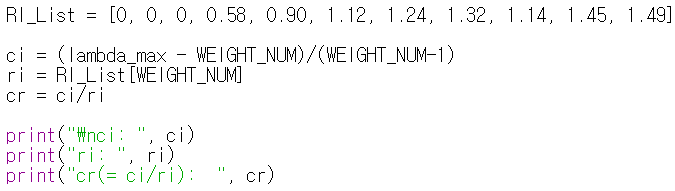


그렇게 하여 구한 새로운 행렬(B)의 각 행의 합계를 구하고, 각 행의 합계를 각 항목의 가중치로 나눈다.

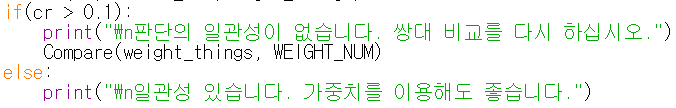
그렇게 나눈 값들을 모두 합하여(total) 평균을 낸다. 그 평균 값이 lambda\_max가 되는 것이다.



행렬의 크기에 따른 RI 값들을 나타낸 리스트 이다. 여기에서는 행렬의 크기가(=항목의 수) 각각 1.12, 0.58 이므로 해당 값을 사용한다.

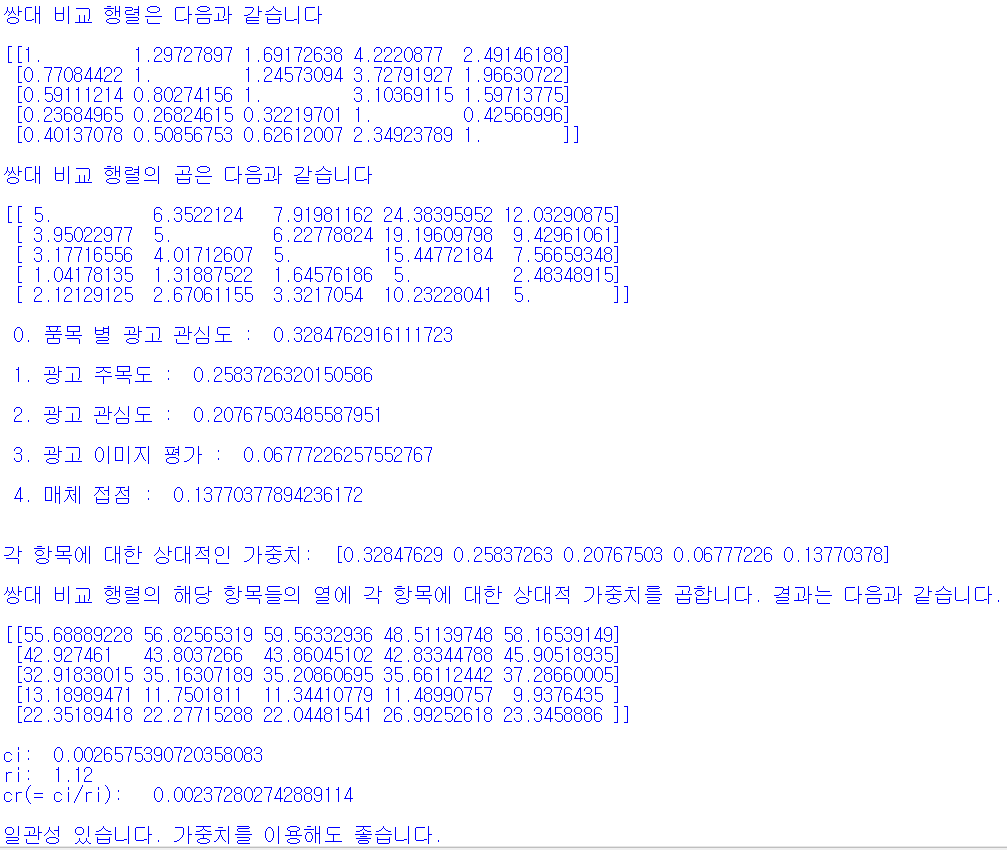


이렇게 나온 CR이 0.1을 넘지 않는다면 일관성을 보이는 것이므로 가중치를 사용하도록 한다.

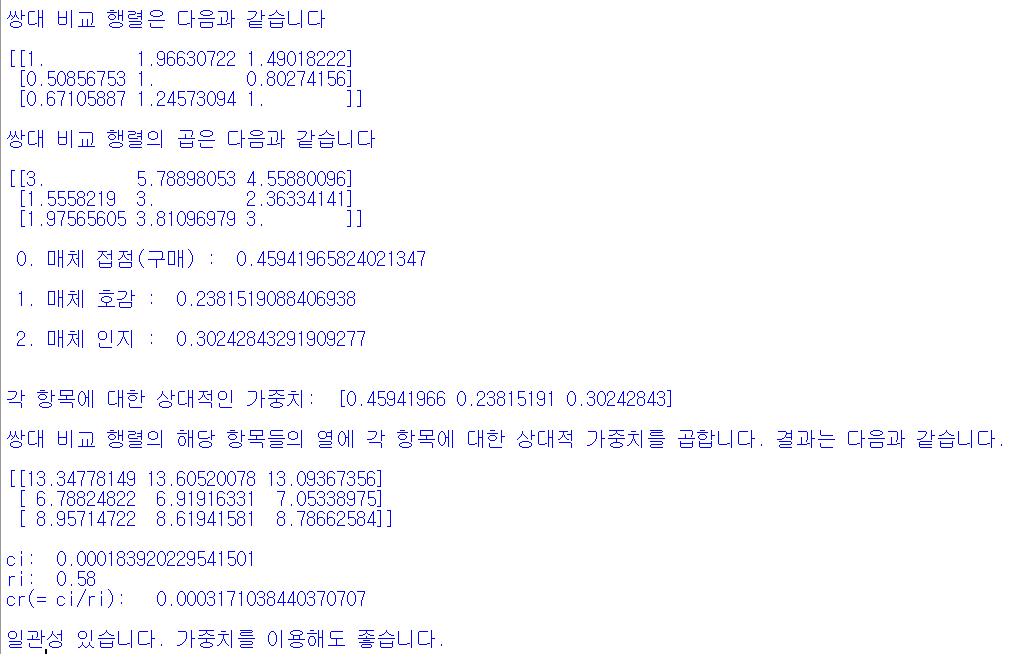


그리하여 나온 결과는 다음과 같습니다.

1. 계층 1(광고 주목도, 품목 별 광고 관심도, 매체 접점, 광고 관심도, 광고 이미지 평가)



1. 계층2(매체 구매, 매체 인지, 매체 호감)

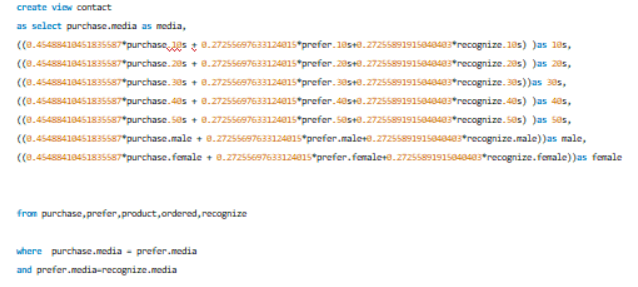


* 1. **광고 매체 추천 SQL 쿼리 문 작성**
     1. **뷰를 사용한 적합도 테이블 생성**

1. 뷰를 이용한 광고 적합도 테이블 생성

릴레이션 구현, 정규성 검정 등을 통해 얻은 테이블을 기반으로 여러 종류의 뷰를 생성했다. 이를 통해 테이블 간의 복잡한 계산을 단계별로 나누어 보다 쉽게 연산을 수행할 수 있도록 했다.

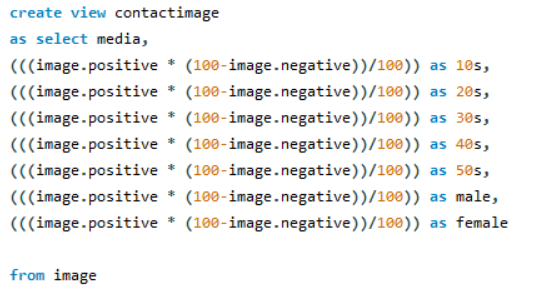
1. 연령대별 매체 접점 지수 계산



매체 호감도, 매체 인지도, 매체 구매 결정 영향력 테이블 값에 AHP기법을 사용하여 구한 각 테이블의 가중치를 곱한 뒤 평균을 구해 연령대별 매체 접점 뷰를 생성한다

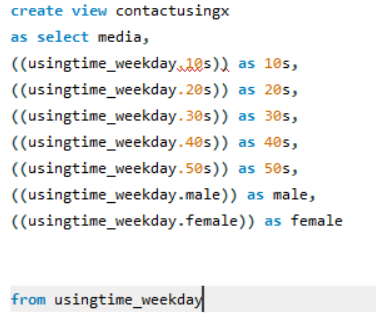
1. 연령대별 매체 이미지 평가 뷰

매체에 대한 이미지를 연령대별로 계산한 뷰를 생성



매체별 이미지 테이블의 긍정 응답비율과 부정 응답비율을 구한 뒤, 부정 응답비율을 긍정 응답비율로 변환하여 곱해 연령대별 매체 이미지 평가 뷰를 생성한다.

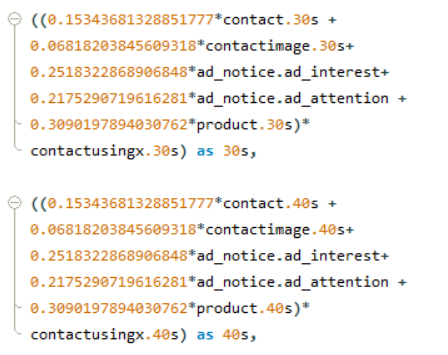
1. 연령대별 매체 이용시간 뷰 생성

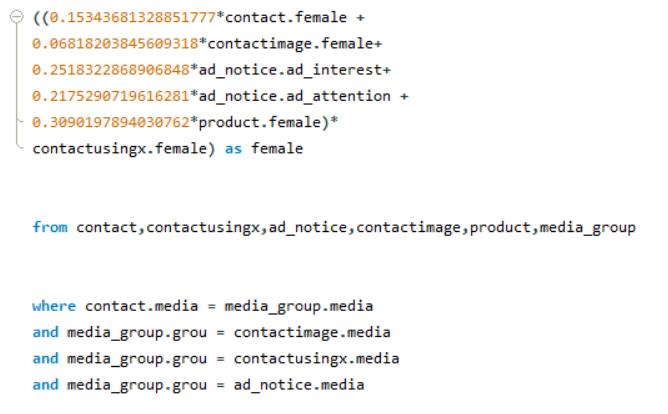


연령대별 매체 이용 시간에 대한 뷰를 생성한다.

매체를 그룹으로 묶어 각 매체 그룹당 각 연령대의 사용시간을 뷰에 입력하였다.

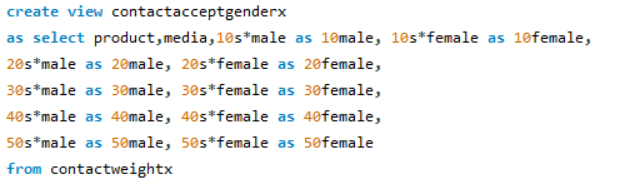
1. 적합도 계산 뷰

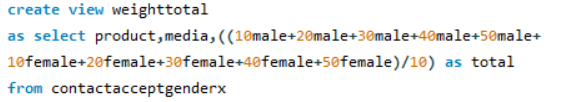


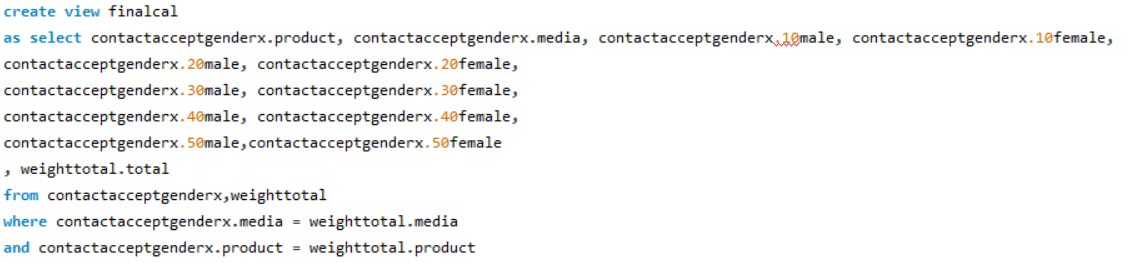


앞서 구한 연령대별 매체 접점 뷰와 연령대별 매체 이미지 평가 뷰, 연령대별 매체 이용시간 뷰에 추가로 광고 관심 주목도 테이블을 이용 광고 적합도 뷰를 생성한다. 이때 연령대별 매체 이용시간 뷰를 제외한 각각의 뷰와 테이블에 AHP기법으로 구한 가중치를 모두 곱한 뒤 더하고, 연령대별 매체 이용시간을 곱해 광고 적합도를 계산, 뷰에 삽입한다.

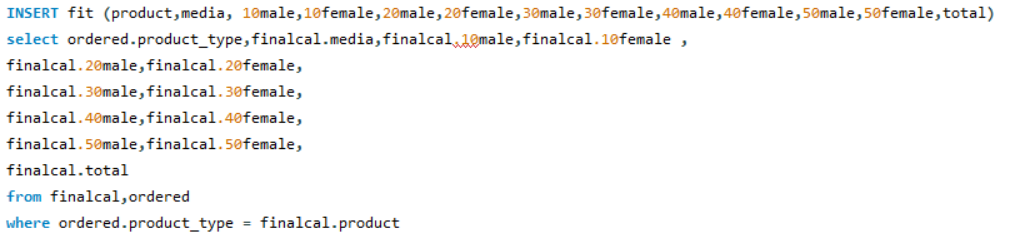
1. 광고 적합도 뷰의 연령대, 성별 속성 변환 후 테이블 삽입







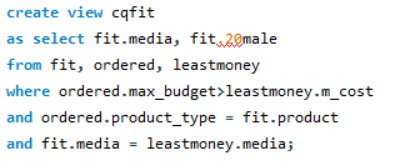
기존 광고 적합도의 분리되어 있는 연령대와 성별 속성을 계산을 통해 통합하여 총 10개의 속성으로 변환한다.

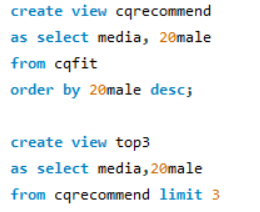


이후 속성이 변환된 광고 적합도의 뷰 데이터를 insert, select 코드를 사용하여 테이블로 삽입한다. 이 과정을 마지막으로 뷰를 사용한 광고 적합도 계산 및 광고 적합도 테이블 생성이 완료된다.

* + 1. **고객이 타겟을 특정화 하였을 때** ex) 20대 남자

뷰를 사용하면 계속 쿼리로 만드는 것이 아닌 값이 바뀌면 같이 바뀌어 손쉽게 바뀐 값의 결과를 알 수 있다.





고객이 사용할 수 있는 최대 예산 금액이 매체 별 최소광고제작 비용보다 적은 광고 매체들을 제외한 광고 적합도를 나타내는 뷰를 만들고 그 뷰 테이블에서 특정화한 광고 타겟의 광고 적합도가 가장 높은 3개의 광고 매체를 출력하는 뷰를 생성한다.

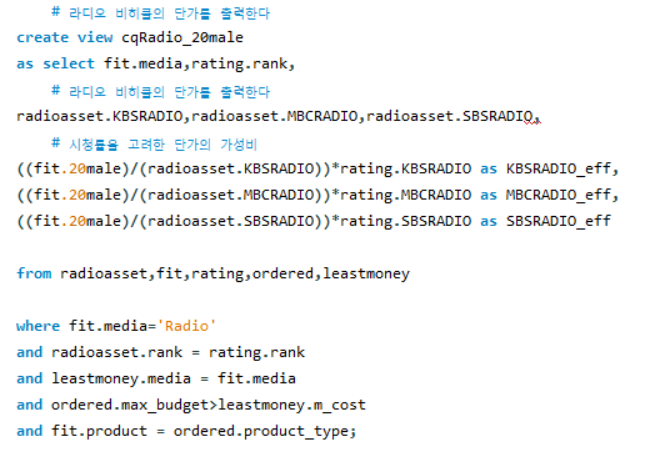
1. 지상파, 케이블, 종합편성 TV 채널의 광고 단가, 가성비 제공





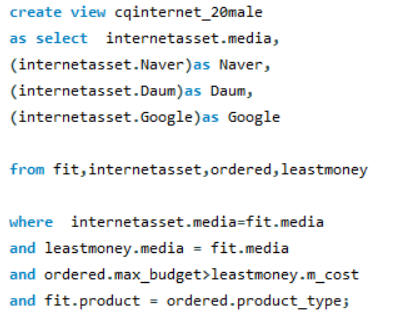
가장 높은 적합도를 지닌 매체 목록에 지상파 채널 tv, 케이블 채널 tv, 종합 편성 채널 tv가 있다면 광고 매체의 비히클별 가격과 시청률을 고려하여 광고 가성비를 나타낸 뷰를 생성한다. 이 세개의 뷰는 타겟에 따라 가성비가 달라지기에 만약 20대 남성이 아니라 다른 타겟이라면 20male부분을 다른 것으로 바꾸어 주면 된다. (40대 여성이라면 20male->40female)

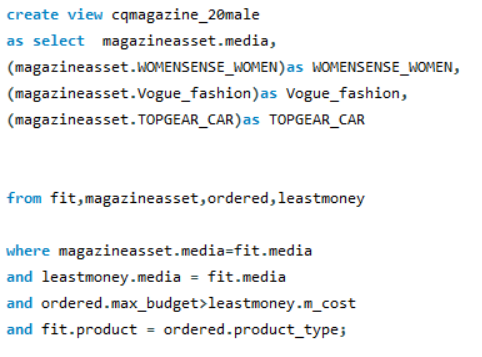
1. 라디오 매체의 단가, 가성비

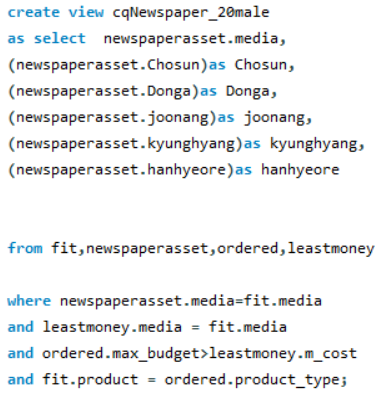


적합도의 3순위 내에 라디오가 있다면 각 라디오의 비히클의 단가와 가성비를 나타낸 뷰를 생성한다. 광고 매체의 비히클별 가격과 시청률을 고려하여 광고 가성비를 나타낸 뷰를 생성한다. 이 뷰는 타겟에 따라 가성비가 달라지기에 만약 20대 남성이 아니라 다른 타겟이라면 20male부분을 다른 것으로 바꾸어 주면 된다. (40대 여성이라면 20male->40female)

1. 인터넷, 잡지, 뉴스 매체의 단가 제공



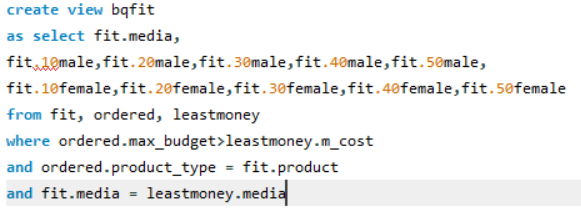




광고 적합도가 높은 매체 목록의 3순위 내에 전단지, 매장내부, 매장외부, 팝업, 대중교통, 극장이 있다면 각 매체 별 비히클의 단가를 나타낸 뷰를 생성한다.

* + 1. **고객이 타겟을 특정화 하지 않았을 때**

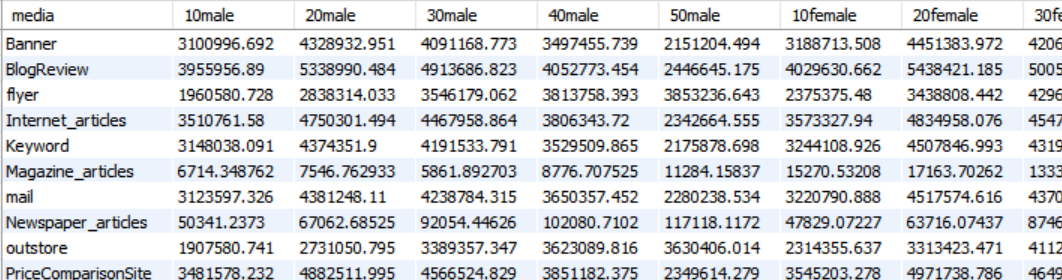
타겟이 정해져 있는 쿼리와 적합도에서 모든 대상으로 하는 것 외에는 같은 쿼리를 사용하여 뷰를 생성하였다.



타겟이 정해지지 않았기에 모든 연령과 성별의 적합도를 불러오는 뷰를 생성하였다.

* + - 1. **전체 타겟의 적합도에서 5순위 선정하기**

앞서 AHP 분석을 통하여 각 항목들의 가중치를 구하고, 이를 기반으로 각 성별 및 연령대 별로 광고 적합도를 구해서 fit 테이블 안에 넣었습니다. 그러고 난 뒤, fit 테이블 안의 값들 중, 의뢰인이 설정한 최대 지출 가능한 금액이 해당 매체 광고를 하는 데에 드는 최소 비용보다 큰 값들만을 골라서(즉, 고객이 의뢰한 금액 선 안에서 가능한 매체 광고들만 걸러서) 생선된 bqfit을 사용하여 위의 뷰를 출력하면 다음과 같습니다.



위의 bqfit 테이블에 있는 여러 연령대 및 성별의 광고 적합도 값들 중 가장 적합도가 높은 값들 5개를 뽑아서 5개의 값에 해당하는 연령대 및 성별 5개, 그리고 5개의 광고 적합도 값에 해당하는 매체 5개를 선정하고자 한다. 이 때, 파이썬과 mySQL을 연동을 시킴으로써 상위 5개의 적합도 값을 뽑아냈다.

**2.4.3.2. 파이썬 연동**

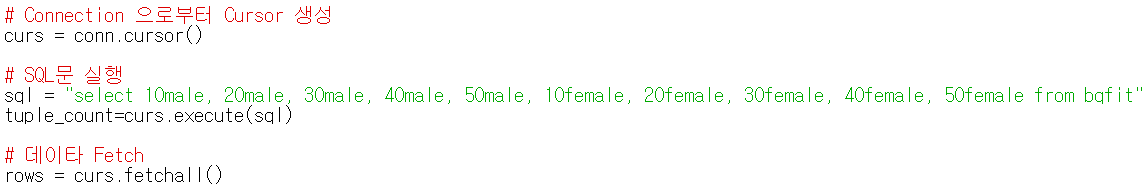
파이썬과 mySQL을 연동시키기 위해 우선 pymysql 라이브러리를 컴퓨터에 설치한 뒤에 불러온다. 또한 불러온 뷰를 행렬의 형태로 변환하기 위해서 numpy도 함께 import 한다.



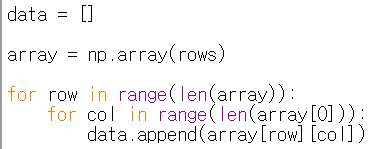
Import 해온 뒤에 pymysql의 connect를 통해 mySQL과 파이썬을 함께 연동 시킨다.



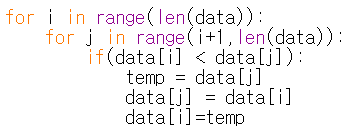
mySQL과 파이썬의 connection을 만들었다면, connection으로부터 Cursor를 생성한다. Cursor까지 모두 생성했다면 원하는 sql을 입력하여 execute함수에 넘겨줍니다. 우리는 bqfit 뷰의 모든 요소들 중에서 상위 5개의 적합도를 선정하고자 하므로 bqfit 뷰로부터 10~50대 여성, 남성의 적합도를 검색하는 sql를 작성한다. Sql을 실행하는 curs.execute(sql)함수는 실행하는 투플의 개수만큼 실행이 되므로 해당 값을 tuple의 개수를 나타내는 tuple\_count 변수에 넘겨서 이후에 행렬을 사용하는 데에 쓰이게 됩니다.



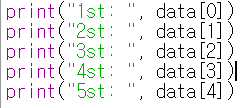
뷰의 데이터들은 모두 가져와서 rows라는 변수에 할당하게 됩니다.



가져온 데이터들을 할당한 변수 rows를 이용해서 뷰의 데이터들을 행렬의 형태로 변환하여 준다. 데이터들을 행렬의 형태로 변환한 뒤, 행렬의 각 요소들(=데이터)들을 빈 리스트 data에 추가시켜준다.

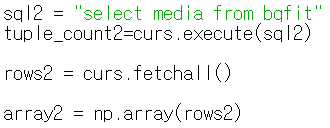


모든 데이터들을 data 리스트에 넣은 뒤, 데이터들, 즉 모든 광고 적합도를 큰 수에서 작은 수 순서대로 내림차순으로 정렬한다.



따라서 위와 같이 출력하게 되면 가장 높은 광고 적합도 5위를 출력하게 된다.

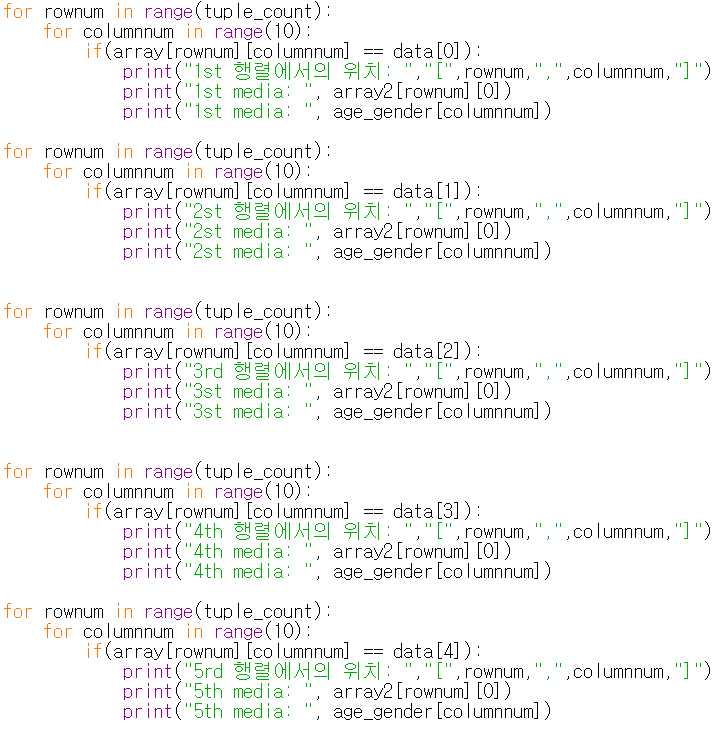
가장 높은 적합도 5개를 뽑았다면, 5개의 적합도에 해당하는 매체 또한 뽑아줘야 한다. 따라서 bqfit 뷰로부터 media 속성들을 검색하여 행렬의 형태로 바꿔준다.



또한 해당 적합도의 연령대 및 나이도 함께 검색하고자 하므로 아래와 같은 리스트를 만들어서 해당 적합도가 위치한 열의 정보를 이용하여 어떠한 품목의 광고가 어느 연령대 및 성별에 적합하였는지를 파악한다.



앞서 구한 5개의 적합도에 대하여, 각각 어느 행과 열에 위치하는지 알아내서, 해당 적합도와 같은 행에 위치한 매체가 무엇인지, 그리고 같은 열에 위치한 연령대 및 성별이 무엇인지를 파악합니다.



파이썬과 mySQL 연동 작업을 마친 뒤에는 connection을 닫아주면서 마무리한다.

**2.4.3.3. 단가 뷰 생성**

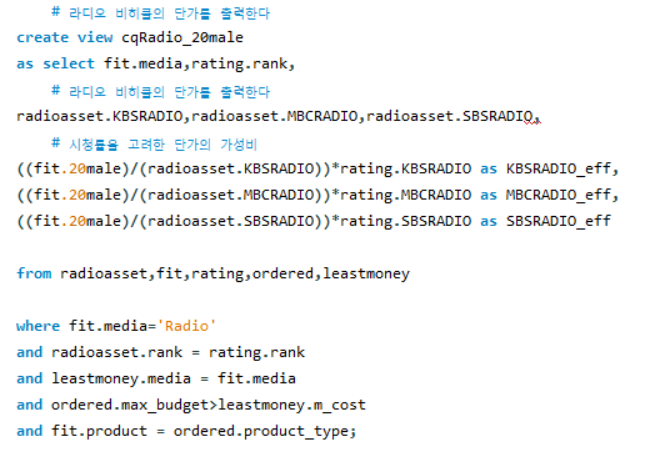
1) 지상파, 케이블, 종합편성 TV 채널의 광고 단가, 가성비 제공





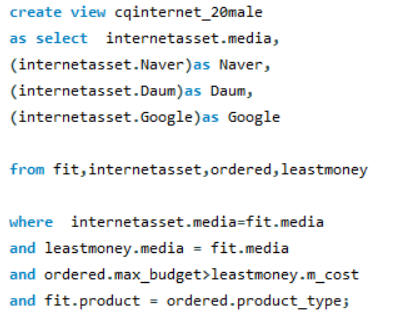
가장 높은 적합도를 지닌 매체 목록을 파이선을 통해 5개를 추출한 후 그안에 지상파 채널 tv, 케이블 채널 tv, 종합 편성 채널 tv가 있다면 광고 매체의 비히클별 가격과 시청률을 고려하여 광고 가성비를 나타낸 뷰를 생성한다. 이 세개의 뷰는 타겟에 따라 가성비가 달라지기에 적합도 순위안에 해당하는 타겟으로 바꾸어 주어야 한다. 만약 10대 남성 지상파 tv라면 위의 쿼리에서 20male -> 10male로 바꾸고 40male 또한 적합도 선정 순위 안에 있다면 40male로 바꾸어 값을 출력해야한다. (이 예시는 20대 남성이 지상파tv, 캐이블tv, 종합 편성tv의 적합도 값이 선정 순위안에 있다고 가정하였을 때의 쿼리문이다)

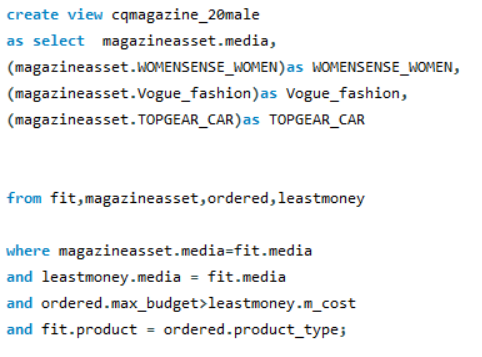
2) 라디오 매체의 단가, 가성비

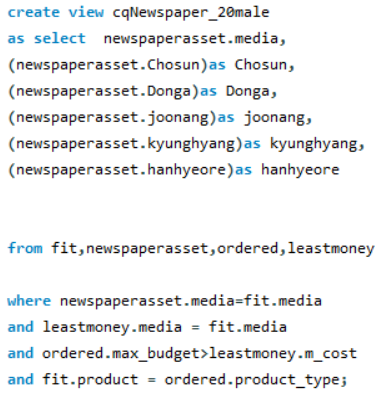


가장 높은 적합도를 지닌 매체 목록을 파이선을 통해 5개를 추출한 후 그 안에 라디오가 있다면 각 라디오의 비히클의 단가와 가성비를 나타낸 뷰를 생성한다. 광고 매체의 비히클별 가격과 시청률을 고려하여 광고 가성비를 나타낸 뷰를 생성한다. 이 뷰는 타겟에 따라 가성비가 달라지기에 적합도 순위안에 해당하는 타겟으로 바꾸어 주어야 한다. 만약 10대 남성 지상파 tv라면 위의 쿼리에서 20male -> 10male로 바꾸고 40male 또한 적합도 선정 순위 안에 있다면 40male로 바꾸어 값을 출력해야 한다. (이 예시는 20대 남성이 지상파tv, 케이블 tv, 종합 편성tv의 적합도 값이 선정 순위안에 있다고 가정하였을 때의 쿼리문이다)

3) 인터넷, 잡지, 뉴스 매체의 단가 제공

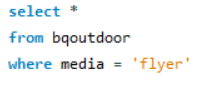




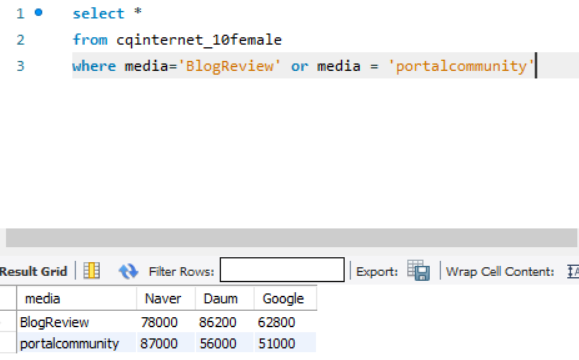


가장 높은 적합도를 지닌 매체 목록을 파이선을 통해 5개를 추출한 후 그 안에 전단지, 매장내부, 매장외부, 팝업, 대중교통, 극장이 있다면 각 매체 별 비히클의 단가를 나타낸 뷰를 생성한다.

* + 1. **비히클 추천 SQL**



타겟을 선정하는 것과 선정하지 않은 상황에서 적합도 순위에 해당하는 미디어의 비히클을 출력하기 위하여 해당 미디어가 속해져 있는 뷰에서 그 미디어의 값 만을 출력한다. (적합도 선정 순위 안에 전단지가 선정되어 전단지의 비히클의 단가를 출력한다)



적합도 선정 순위 안에 인터넷 기사와 메일이 선정되어 인터넷 기사와 메일의 비히클의 단가를 cqinternet\_10female 뷰에서 출력한다.

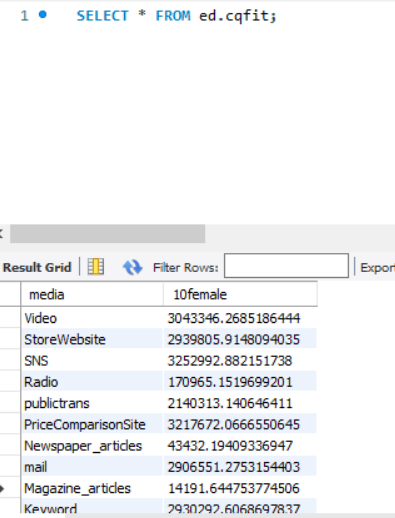
* + 1. **테스트 결과 해석**
       1. **테스트 결과**

위의 쿼리를 통하여 주문서를 작성해보고 그에 따른 결과를 도출해보았다. 쿼리를 통하여 뷰를 생성해 두었기에 특별한 경우가 아니라면 뷰를 호출하여 값을 알 수 있다.

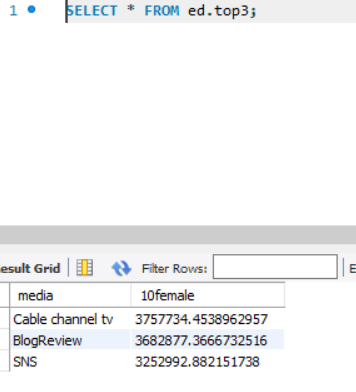
**케이스1) 섬유 유연제 신제품 광고를 10대 여성을 타겟으로 400만원의 투자 예산으로 하려 한다. 주어진 조건에서의 최적의 광고 매체를 추천하겠다.**



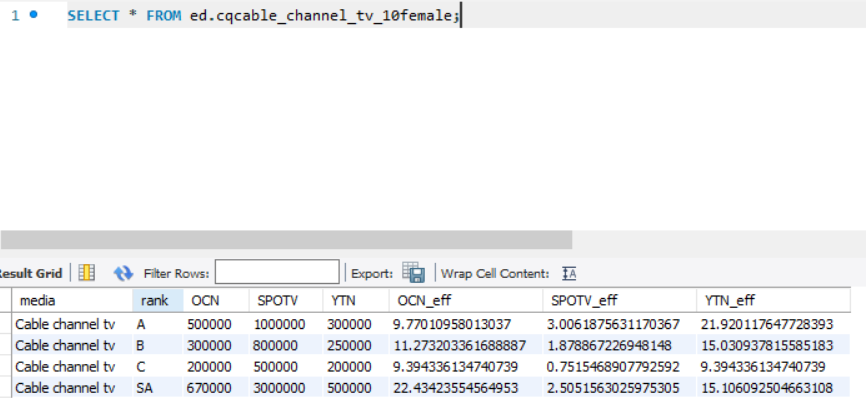
<Case1의 주문서>



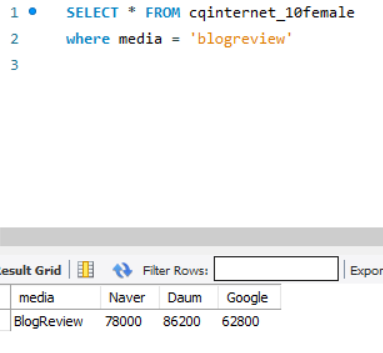
10대 여자를 대상으로 섬유유연제 광고를 실시할 때의 광고 매체 적합도 뷰를 출력한다.



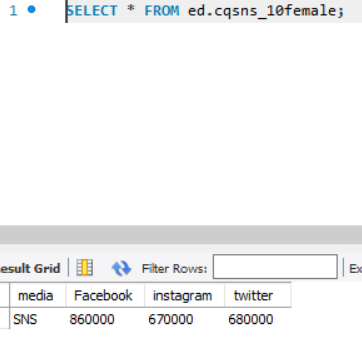
적합도를 내림차순으로 정렬하여 높은 순위 광고 매체 3가지를 출력한다. 높은 순위 3가지는 케이블 TV, 블로그 리뷰, SNS가 출력되었다.



<케이블 TV 비히클별 단가표>



<블로그 리뷰 비히클별 단가표>



<SNS 비히클별 단가표>

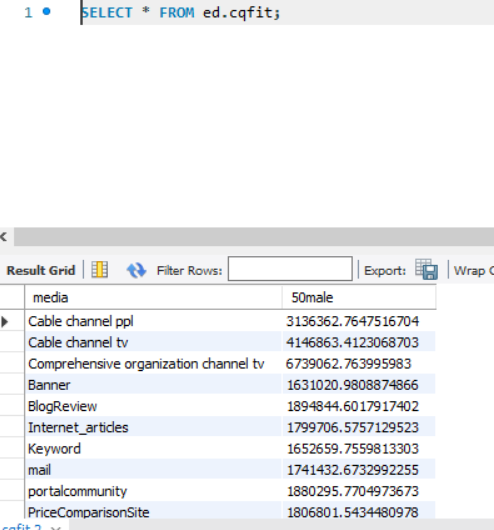
적합도가 높은 3가지 매체의 비히클별 가격을 출력하였다. 결과는 위와 같다. 예산이 충분하다면 OCN을 통한 케이블 광고가 효과가 좋을 것이고 만약 예산을 절약하고 싶다면 google을 통한 블로그 리뷰를 광고 방법으로 선택하는 것이 좋을 것이다.

10대 여성에게 4백만원의 예산 금액으로 광고하는 것은 광고효과가 좋은 TV 매체중에서 광고단가가 낮은 편인 케이블tv 매체와 10대의 특징 중 하나인 많은 인터넷 접촉횟수로 효과를 볼 수 있는 블로그 리뷰나 SNS에 광고를 하는 것이 최적의 광고 방법이라고 예상한다

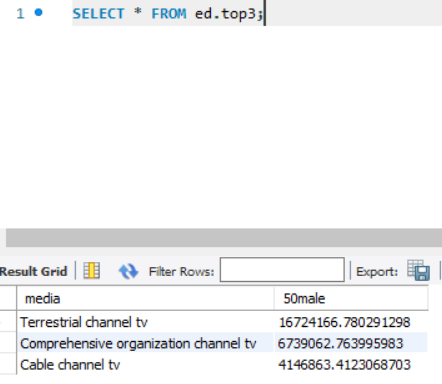
**케이스2) 자동차 신제품 광고를 50대 남성을 타겟으로 5억원의 투자 예산으로 하려 한다. 주어진 조건에서의 최적의 광고 매체를 추천하겠다.**



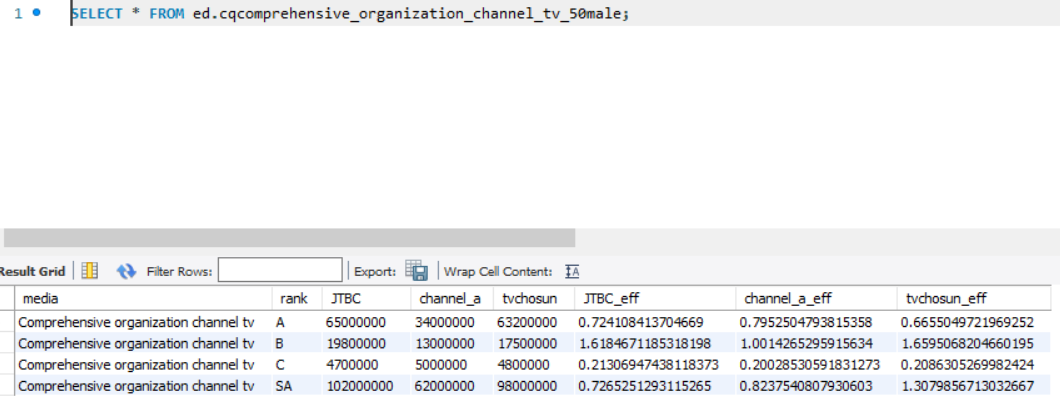
<Case2의 주문서>



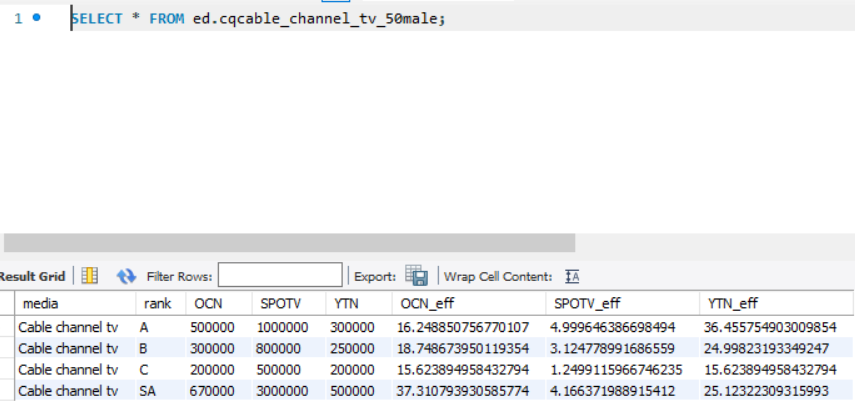
50대 남성을 대상으로 자동차 광고를 실시할 때의 광고 매체 적합도 뷰를 출력한다.



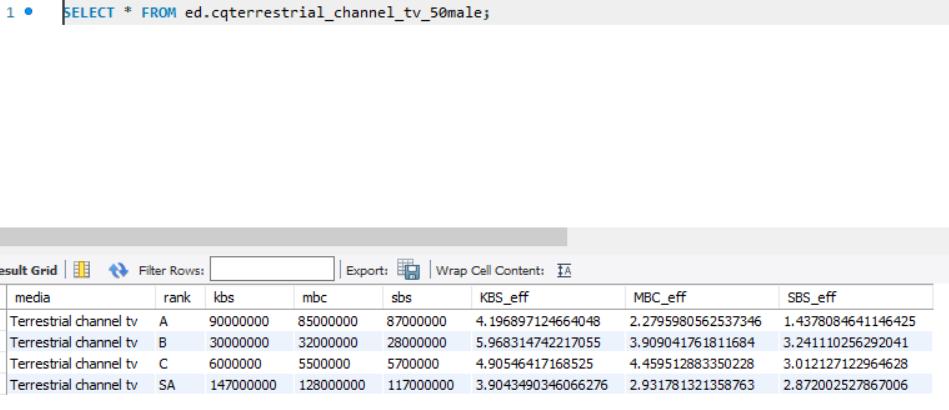
광고 매체 적합도중 상위 3개를 뽑아보았는데 지상파tv, 종합편성tv, 케이블tv가 나오게 되었다



<종합편성채널 TV 비히클별 단가표>



<케이블 채널 TV 비히클별 단가표>

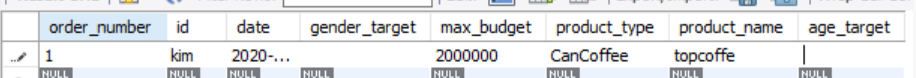


<지상파 채널 TV 비히클별 단가표>

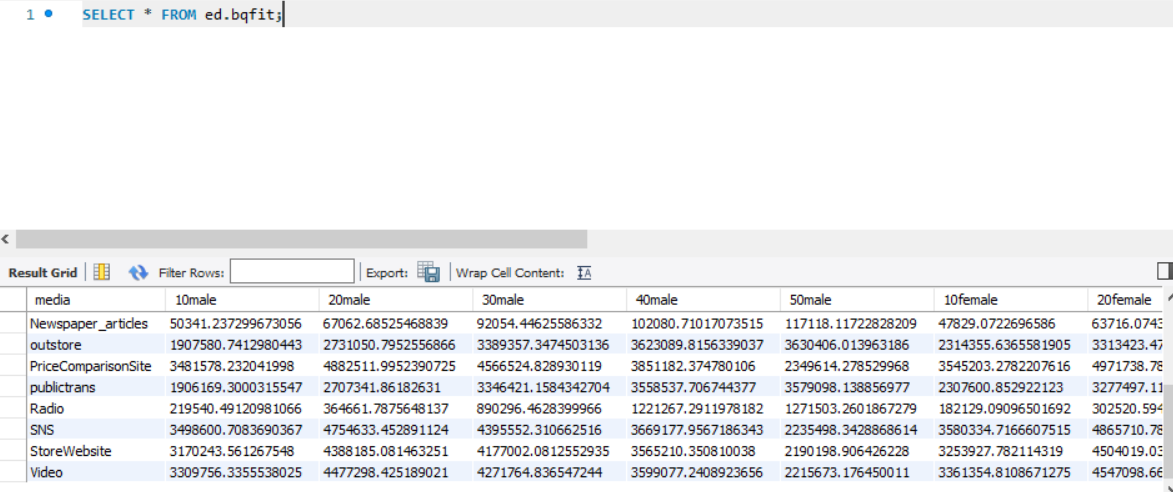
50대 남성을 대상으로 차를 광고하고 광고 예산이 충분히 많을 경우 tv를 통한 광고가 가장 적합도가 높게 나왔다. 적합도가 제일 높게 나온 것은 지상파tv였으나 광고비가 굉장히 비싸다는 단점이 있다. 하지만 충분한 투자 가능 금액이 있어 지상파 TV매체에 광고할 수 있다면 엄청난 효과를 볼 수 있을 것이다. 그리고 50대 남성을 대상으로 하였기에 구매력도 높기에 높은 효과가 예상된다.

비용 효율적인 측면관점에서 생각한다면 케이블, 지상파, 종합편성tv의 순서대로 광고하는 것이 적합하다.

**케이스3) 캔커피 신제품 광고를 광고 타겟을 지정하지 않고 이백만원의 투자 예산으로 하려 한다. 주어진 조건에서의 최적의 광고 매체와 광고 타겟을 추천하겠다.**



<Case3의 주문서>

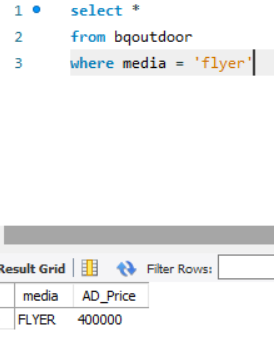
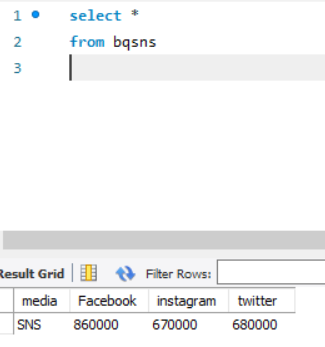
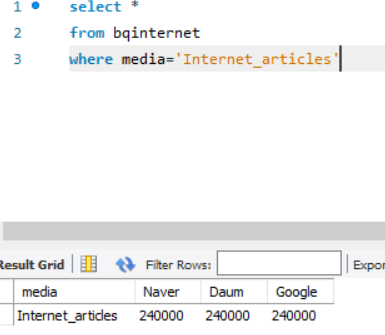


광고 타겟을 정하지 않았기에 모든 연령과 성별의 적합도를 불러왔다.

위에서 설명한 파이썬을 연동하여 사용한 방법을 통하여 광고 매체, 광고 타겟 적합도 5순위를 선정하였다. 결과는 아래와 같이 나오게 되었다.



캔커피의 광고 적합도가 가장 높은 매체와 타겟은 20대 여자 SNS, 20대여 인터넷기사, 20대 남자 SNS, 20대 남자 인터넷 기사, 50대 여자 전단지가 적합도가 높게 나왔다. 구매력에서는 30,40대가 높지만 200만원이라는 적은 돈으로 30,40대가 주로 시청하고 효과가 좋은 tv매체들이 삭제되어 적은 돈으로 높은 효율을 낼 수 있는 20대 남녀의 SNS와 인터넷 기사가 적합도가 좋게 나왔다. 50대 여자에게 광고 매체로 전단지 또한 높은 적합도를 나타냈다.



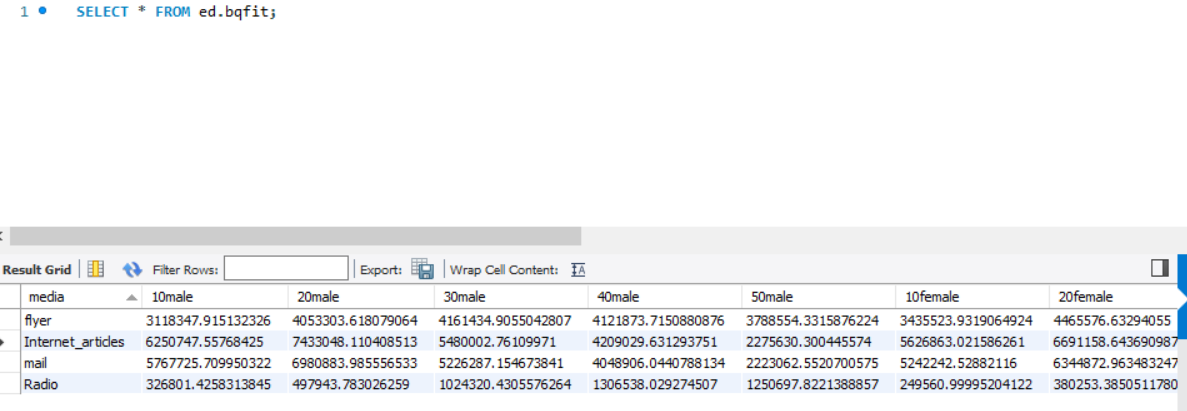
<인터넷기사 비히클별 단가> <SNS 비히클별 단가> <전단지 비히클별 단가>

SNS가 다른 미디어에 비해 2배가량 비싸지만 높은 효과를 보고 싶다면 SNS의 페이스북, 인스타그램, 트위터를 선택하는 것이고 최고의 가성비의 광고를 위해서라면 네이버, 다음, 구글에 인터넷 기사를 광고를 내는 것이 합리적이라고 생각된다.

**케이스4) 배달음식 신제품 광고를 광고 타겟을 지정하지 않고 오십 만원의 투자 예산으로 하려 한다. 주어진 조건에서의 최적의 광고 매체와 광고 타겟을 추천하겠다.**



<Case4의 주문서>

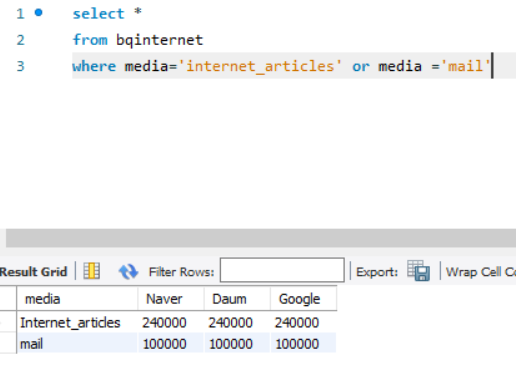


광고 타겟을 정하지 않았기에 모든 연령과 성별의 적합도를 불러왔다.

위의 파이선을 사용한 방법을 통하여 광고 매체, 광고 타겟 5순위를 선정하였다. 결과는 아래와 같이 나오게 되었다.



50만원의 적은 돈으로 광고를 하는 것이기에 선택지가 굉장히 적다. 전단지, 인터넷 기사, 메일, 라디오가 가능한데 그 중 적합도가 가장 높았던 미디어와 타겟은 20대 남녀에게 인터넷 기사와 메일을 통한 광고와 10대 남자에게 인터넷 기사로 광고하는 것이 적합도가 높게 나왔다. 이를 보면 젊은 세대의 사람들이 장년층에 비하여 인터넷 기사 광고에 영향을 많이 받는 것을 알 수 있다. 이는 많은 인터넷 사용시간과 높은 배달음식 이용횟수로 인하여 이러한 결과가 나온 것이라 생각한다



인터넷 기사와 메일의 경우 3곳의 가격이 동일하기에 광고의 작전에 따라 결과가 달라질 것이라고 예측된다. 한곳에 집중적으로 하여 자주 접촉되거나 3곳에 골고루 나누어 많은 사용자에게 접촉될 수 있도록 하는 것으로 나뉠 것이다.

* + - 1. **결과 해석**

특정 품목의 광고의 적합도 즉 광고 효과는 광고 매체에 따라 크게 달라지는 것을 테스트 결과 데이터를 통해 볼 수 있었다.

광고 타겟이 특정화 된 상태에서 최적의 광고 매체를 선정할 때 거의 대부분의 품목의 경우에서 TV 종류, 특히 지상파 TV로 광고하는 것이 광고 효과가 가장 좋다고 나왔고 잡지에 광고하는 것이 제일 광고 효과가 떨어지는 것을 볼 수 있었다. 모든 광고 품목의 광고 매체를 선정할 때 지상파 TV를 매체로 광고하면 좋지만 광고 비용이 높다는 점 때문에 모든 고객이 광고 매체로 선정할 수 없다. 고객 마다 사용할 광고 예산 금액이 다르기에 한정된 예산내에서 가장 적합도가 높은 매체를 선정해야 한다.

위의 예시로 섬유유연제 신제품을 광고 타겟을 10대 여자로 한정하여 광고하고자 할 때 광고예산안에서 가장 적합한 매체는 케이블 TV, 블로그 리뷰, SNS로 계산되었다. 케이블 TV중 뷰티 관련 프로그램을 많이 진행하는 채널의 경우 비교적 적은 돈으로 높은 광고 효과를 누릴 수 있을 것이고 블로그 리뷰나 SNS의 경우 10대 여자가 사용시간이 높은 매체에 해당하기에 광고 효과가 높을 것이다. 광고 매체 예측치 모델을 통해 나온 결과가 꽤나 적절하고 합당하게 나왔기에 설계한 모델의 예측 성능이 상당히 높다고 판단할 수 있다.

광고 타겟을 특정하지 않고 적합한 광고 매체와 타겟을 선정할 때 타겟 별 광고 적합도의 총합을 보면 30, 40대를 타겟으로 광고를 하는 것이 광고효과가 제일 높고 20, 50대가 그 다음으로 높고 10대가 가장 낮게 나오게 되었다. 하지만 이것 만으로 광고 타겟을 30, 40대로만 한정하여 선정하는 것은 적절하지 않다. 품목 별로 가장 높은 광고 효과를 보이는 타겟이 다르고 연령대 별로 가장 적합한 광고 매체가 다르기 때문이다.

30, 40대에게 광고 효과가 제일 높은 것은 매체는 TV인데 TV는 광고제작비가 비싸기 때문에 예산 금액을 고려하여 판단해야 한다. 반면에 광고제작 단가가 비교적 저렴한 인터넷 매체를 이용한 광고에서는 10, 20대의 광고 효과가 30, 40대보다 높게 나왔기 때문에 광고할 제품의 주 구매층이 10, 20대인 경우 인터넷 매체에 광고를 하는 것이 더 적합한 선택일 수 있다. 이를 통해서 광고 매체 선정은 구매력과 같은 단순한 수치로 판단하는 것이 아니라 자신의 투자 가능 금액에서의 광고할 품목의 광고 효과가 제일 좋은 매체를 선정한 뒤 그 매체에 많이 접촉하는 타겟과 구매력이 좋은 타겟과 같이 다양한 요소들을 고려하여 광고 타겟을 선정해야 최대의 광고효과를 볼 수 있다는 것을 알 수 있었다.

1. **결론**

3.1. 데이터 베이스

프로젝트 목표를 달성하기 위해 한국방송광고진흥공사 (kobaco)에서 제공하는 소비자행태조사(MCR) 데이터 서비스의 공공데이터를 수집, 데이터베이스를 구현했다. 데이터베이스를 구현하기 위해 가장 먼저 요구 사항 분석 및 개념적 설계 과정을 거쳤는데, 이 과정이 데이터베이스 구현을 위해 가장 중요한 부분이라고 생각하게 되었다. 사용자의 요구사항이 무엇인지, 어떤 업무를 수행하게 되는지, 수집한 데이터에 어떤 처리가 필요한지를 확실하게 알 수 있어야 프로젝트 목표 달성을 위한 데이터베이스를 설계할 수 있을 것이다. 이후 엔티티를 구현하고 정규화 과정을 통해 엔티티를 통합하거나 분리할 수 있었다. 강의에서는 BCNF를 비롯한 4NF, 5NF까지 배웠지만 실질적으로는 3NF나 BCNF까지 구현하는 것이 가장 일반적이고 기본적인 정규형이라고 느꼈고, 4NF나 5NF와 같은 정규형의 정규화는 특수한 경우, 필요에 따라 수행된다는 것을 알 수 있었다. 정규화 과정을 마친 뒤, 관계 데이터 모델을 통해 각 엔티티 간의 관계를 설정하여 각 테이블의 기본 키와 외래 키 관계 등을 설정할 수 있었다. 이후 수집한 데이터를 앞선 과정에 맞는 형식으로 엑셀에 정리한 뒤, AutoSet과 MySQL Workbench를 통해 데이터베이스로 가져오고 각 테이블의 기본 키와 외래 키를 설정, 관계를 설정하여 데이터베이스를 구현할 수 있었다.

데이터베이스를 구현하면서 정해진 절차에 맞게 차례대로 과정을 수행하는 것이 매우 중요했다. 중간의 한 단계라도 제대로 이뤄지지 않을 경우 다음 단계를 정상적으로 수행할 수 없었고, 이렇게 구현된 데이터베이스의 경우 프로젝트 목표를 수행하기에 적합하지 않았다. SQL 질의 등을 통해 데이터베이스를 어떻게 조작하는 지도 굉장히 중요하지만, 목표에 맞는 데이터베이스를 설계하고 구현하는 것이 가장 중요하다고 생각하게 되었다.

3.2. 데이터 분석

기본적으로 어떤식의 광고가 제일 효과가 좋을지 생각만 하는 것이 아니라 실제 소비자의 데이터를 사용하여 가중치를 적용시켜 감각적으로 광고를 추천하는 것이 아니라 눈에 보이는 확실한 증거인 데이터로 고객에게 맞는 광고 방법을 추천하려 하였다. 다양한 데이터들을 단순히 나열하는 것이 아닌 설문조사를 통해 얻어낸 데이터를 AHP기법을 사용하여 어떠한 인자가 더욱 중요한지 알아보아 가중치를 다르게 적용시켰다. 그 가중치에 맞게 적합도를 계산해내었고 적합도를 통해 추천을 하는 구조를 만들었다. 시작할 때 TV가 가장 효과적인 매체일 것이라고 예측하였는데 역시 TV의 적합도는 확실히 높았다. 하지만 모두가 생각하는 가장 높은 효과의 매체이기에 광고 제작 비용이 너무나 비싸 일반적인 광고주는 TV를 선택할 수 없다. 이에 광고주의 최대 예산에 맞는 광고 방법을 추천할 수 있게 광고 제작 비용이 최대예산보다 비싼 매체는 제외하는 쿼리를 작성하여 더욱 고객에게 최적화된 데이터베이스를 만들 수 있었다. 그리고 고객이 타겟을 정하고 광고를 진행할 수 있게 그 타겟의 광고 적합도를 내림차순으로 정렬하여 보기 쉽게 나타내었고 상위 3개의 매체를 추천해주었다. 이 각 3개의 매체는 각각 3개의 비히클을 가지고 있어 다양한 광고 전략 수립도 가능하게 하여 최대의 효과를 낼 수 있게 하였다. 만약 타겟을 정하지 않고 타겟과 매체를 동시에 추천받고 싶은 경우를 대비하여 파이썬과 연동하여 가장 높은 적합도 5개를 추천해주었다. 이와같이 다양한 수리 알고리즘과 계산 쿼리로 고객에 최적화된 베이스를 구축하기 위해 많은 자료를 찾아보고 공부하고 실제로 응용하여 보며 적용이 무사히 되었을 때 뿌듯함을 느꼈다.

3.3. 기대효과

실제 소비자의 데이터를 다양한 기법으로 만들어낸 광고 매체 추천 데이터베이스이기에 많은 광고주들이 실제 광고에 직접 투자해 보고 결과를 받는 것이 아니라 이것을 통하여 미리 어느 정도의 결과가 나올지 예상하고 자신의 예산에 맞는 최적의 광고 작전을 짠 후 광고 투자에 들어간다면 자본의 낭비가 최소한으로 줄어듦과 동시에 최대의 광고 효율을 얻을 수 있을 것이다. 또한 이것은 광고주에게만 이점이 되는 것이 아니라 그 광고를 불필요하게 느껴 귀찮게 느끼는 사람에게 최소한으로 전달되고 그 제품에 대해 관심 있는 사람에게 전달이 잘되어 광고주와 소비자가 서로 이득을 갖는 윈윈 구조를 생성할 수 있을 것이다.

그동안 E-R 모델링 및 다이어그램, 릴레이션, SQL과 같은 이론들을 배웠다면, 이번 실습을 통하여 우리가 배웠던 이론들을 실제로 적용해 볼 수가 있는 기회를 가질 수가 있었다. 8장에서 배웠던 내용들을 기반으로 요구 사항 명세서를 작성해서 여러 개체들을 생성하고, 9장에서 배운 정규화 개념을 기반으로 하여 테이블들을 분해 시키는 작업을 거쳐서 만들어진 최종 테이블들로 4장에서 배웠던 E-R 모델링 이론을 통하여 광고 적합도 테이블, 매체 테이블, 시청률 테이블들 등 간의 릴레이션을 생성하여 E-R 다이어그램을 만들어 보았다. 또한 엔지니어링 DB에서 배웠던 내용 뿐만 아니라 최적화 응용 시간에 배웠던 파이썬 내용을 토대로 AHP 기법을 구현 해보고, MySQL과 연동을 시켜서 이번 수업 시간에 배웠던 내용에 적용을 해보는 유익한 경험이 되었던 것 같다. 스케일이 큰 프로젝트였던 만큼 많은 의논이 필요하였고, 그리하여서 상황 상 많이 만나기 어려웠던 상황임에도 팀원들 모두 시간을 많이 내주어서 프로젝트에 관하여 많은 의논을 나누고, 많이 고민을 하였다. 특히 팀플 기간 중 8할 이상은 E-R 모델링과 릴레이션 관련하여 의논을 나누는 데에 보냈던 것 같은데, E-R 모델링 파트가 수업 내용 중 앞 부분에 위치해 있었던 만큼 가장 기본적인 부분임에도 사람마다 관계를 해석하는 방향이 다르다 보니 응용을 할 때 생각보다 어렵게 느껴졌던 것 같다. 또한 수업 시간에 SQL 부분을 배울 때 큰 어려움을 느끼지 못했지만, 실제로 저희가 SQL문을 만들어서 작성을 해보니 오류가 나온 적도 많았다. 그러한 부분에 대하여 왜 오류가 난 것인지 강의노트를 다시 보면서 공부하고, 인터넷에 검색도 해보면서 더욱 정확하게 배우게 되었던 것 같다. SQL 뿐만 아니라 다른 부분에 있어서도 실제로 적용을 하면서 여러 문제에 부딪혀 보면서 깨닫게 된 점도 많았던 것 같다. 이번 과제를 하면서 광고 적합성 관련 데이터들을(매체, 품목 별 관심도 등) 조사하면서 광고를 하게 될 때 생각보다 고려해야 될 점이 많다는 것을 깨달을 수가 있었다. 저희가 프로젝트를 진행할 때에도 테이블이 10개가 넘게 매우 많이 나왔을 만큼 광고를 해야 될 때 고려해야 될 사항들을 많이 탐색해서 데이터 분석에 사용하였지만, 실제로 그 보다도 고려해야 될 사항들이 더 많을 거라고 생각이 된다. 이번 실습에는 시간 상 더 많은 부분을 탐색해보지 못해서 많이 아쉽지만, 다음에 기회가 될 때 더 많은 데이터들을 탐색해서 분석하게 되면 이후에 실제로 광고 매체를 찾는 데에 적용을 할 때 조금 더 구체적인 정보를 얻을 수 있지 않을까 기대가 된다.

1. **참고문헌 및 출처**

KOBACO(<https://adstat.kobaco.co.kr/mcr/portal/mainPage.do>)

<AHP를 이용한 선박 평형 수 처리장치의 최적선정에 관한 연구>

(RISS, 학위 논문(박사)-부경대학교 대학원: 기술경영협동과정 2016.8, 저자: 이상원)

<델파이 기법(Delphi)과 계층적 의사결정방법(AHP)의 적용을 통한 전통정원의 보존상태 평가지표 개발>

(RISS, 학위 논문(박사)-성균관대학교 일반대학원: 조경학과 2011.2, 저자: 안진성)

한국방송광고진흥공사(kobaco), <https://www.kobaco.co.kr/>

황보제. (2000). “제품유형에 따른 광고매체별 소비자 구매의도: 전통매체광고와 뉴미디어 광고의 효과”, 석사학위논문, 한성대학교, 서울, 대한민국, 35-37p

배성덕. (2016). “빅데이터 분석을 통한 모바일 광고플랫폼의 광고효과 연구: 광고특성, 매체특성, 상품특성을 중심으로”, 석사학위논문, 국민대학교, 서울, 대한민국, 21-24p

이민섭. (2014). “빅데이터 분석 시스템을 이용한 TV의 시간대별 시청률과 광고효율성 분석”, 석사학위논문, 서울대학교 대학원, 서울, 대한민국, 32-34p

최창훈. (2014). “구매시점광고가 소비자의 구매의도에 미치는 영향에 관한 연구”, 석사학위논문, 울산대학교, 울산, 대한민국, 18p

김태환, (2004). “매체별 광고효과 및 브랜드 파워 차이에- 따른 광고 – 시장점유율 관계분석”, 석사학위논문, 중앙대학교 대학원, 서울, 대한민국, 8-15p

윤덕균. (2014). “기업진단론 23강 AHP” KOCW(고등교육 교수학습자료 공동활용 체제, http://contents.kocw.or.kr/KOCW/document/2014/hanyang/yundeokkyun/24.pdf