

제 17회 SAS 분석 챔피언십 1차 제안서

팀 번호	SA1089		
팀 장	이소연	팀원 명	이혜승
연락처	010-4632-3342	이메일	leesy960617@gmail.com

1. 홈쇼핑 매출에 영향을 미치는 원인과 판매실적 증대 및 그 활용에 대한 시각화 방안 제시 (1장 이내)

Column name	Label	설명
BDCT_STRT_DTIME	방송시작시간	
ORI_PD_CD, PD_1	PD 코드, PD 코드 1	PD 코드와 독립적이지 않을 경우 PD 코드와 SH 코드 변수 둘 중 하나 제거
ORI_SH_CD, SH_1, SH_2	SH 코드, SH 코드 1, SH 코드 2	PD 코드와 독립적이지 않을 경우 PD 코드와 SH 코드 변수 둘 중 하나 제거
GOODS_CLSS_MID_CD	상품중분류명	

1. 홈쇼핑 매출에 영향을 미칠 것으로 추정되는 요인을 조사한 결과, 경제상황, 월(月), 요일, 우천여부, 휴일여부, 방송시작시간, PD, 쇼호스트, 할인율, 브랜드 선호도, 상품분류 등이 있었음¹
2. 홈쇼핑 매출에 영향을 미칠 것으로 추정되는 요인 중 방송시작시간, PD 코드(PD1 코드), SH 코드(SH1 코드, SH2 코드), 상품중분류명이 기존에 주어진 데이터에 존재하고 경제상황, 월(月), 요일, 우천여부, 휴일여부, 브랜드 선호도는 파생변수 및 외부변수임
3. 방송시작시간, PD 코드(PD1 코드), SH 코드(SH1 코드, SH2 코드), 상품중분류명이 서로 독립인지 다음과 같은 방법을 통해 확인하고 시각화 함
 - HEATMAP + 상관계수 분석 → 다중공산성 제거
 - PD, SH 코드 분석: 피디와 쇼호스트의 조합에 따른 매출액 분석(분산분석)
 - PD, SH 가 독립적이지 않을 것으로 예상됨 (독립적이지 않다면 둘 중 하나 제거)
 - 기존 변수들이 유의하다는 것을 시각화(Conditional plot)를 통해 증명

¹ 생활경제국 표시광고팀, 『TV 홈쇼핑 이용실태와 개선방안』, 2002.5, p21 표 4-13

2. 판매실적 예측을 위한 파생변수 생성, 외부변수 발굴, 예측 모형 개발 (1장 이내)

Column name	Label		변수 종류	설명
BDCT_DT	방송 일자	경제상황 (분기별)	외부변수	소비자심리지수 사용
MONTH		월(月)	파생변수	1~12월 총 12개의 달
DAY		요일	파생변수	일요일~토요일 총 7개의 요일
RAIN_YN		우천 여부	외부변수	비우천 시: 0 / 우천 시: 1 (기상청 자료 참고)
HOLIDAY_YN		휴일 여부	파생변수	비휴일: 0 / 휴일: 1
BDCT_STRT_DTIME	방송시작시간		기존변수	황금시간대 찾기
ORI_PD_CD, PD_1	PD코드, PD코드1		기존변수	
ORI_SH_CD, SH_1, SH_2	SH코드, SH코드1, SH코드2		기존변수	
TVGGRP_SALE_PRC	TV상품그룹 판매가	할인율	파생 및 외부변수	(시중가격 - 판매가격) / 시중가격
BRAND_PREFER	브랜드인지도		외부변수	포털 웹크롤링을 통해 해당 상품의 '네이버 쇼핑' 검색 내역 수로 파악

- 기존 변수 외에 경제상황, 월(月), 요일, 우천 여부, 휴일 여부, 할인율, 브랜드선호도 변수 추가
- 월(月), 요일, 휴일 여부 변수는 기존 방송일자 변수에서 파생된 파생변수
 - 달 별로 매출액을 산출하면, 매출이 큰 달들이 발견될 것임(4,5,6 월로 추정)
 - 요일 별로 매출액을 산출하면, 매출이 큰 요일들이 발견될 것임(월>목>토요일)²
 - 휴일 여부가 매출액에 영향을 크게 미칠것으로 판단됨
- 할인율은 기존 TV 상품그룹판매가와 외부 자료에 의한 시중가격에서 파생하여 생성된 파생변수이자 외부변수
- 브랜드인지도는 웹사이트 크롤링을 통해 '네이버 쇼핑'의 검색 내역 수를 외부변수로 사용
- 범주형 자료와 수치형 자료로 구분하여 분석함
 - 범주형 자료: 월(月), 요일, 우천 여부, 휴일 여부, 방송시작시간, PD 코드, SH 코드, 상품중분류명
 - 수치형 자료: 경제상황, 할인율, 브랜드인지도
- 범주형 자료를 Class 로 분류하고 수치형 자료를 회귀분석을 통해 분석하여 각 범주에 해당하는 적합한 모델을 생성함

² 『중앙일보』, <https://news.joins.com/article/18230416> 기사에서 인사이트를 얻음

3. 예상되는 분석결과와 그에 관한 활용방안 아이디어 제시 (1장 이내)

수치형 자료인 경제상황, 할인율, 브랜드인지도 변수를 활용하여 회귀분석을 한 결과, 계수가 가장 큰 변수를 영향력이 가장 큰 변수라고 판단 할 수 있다. 또한 유의미한 변수를 선정해서 다중회귀분석 및 추가변수 그림을 활용하여 자세히 살펴본다. 다음 수치형 자료 세 가지 변수 중에 경제상황은 임의로 변경이 불가능하기 때문에 할인율과 브랜드인지도에 따라 모델 간 어떤 차이가 있는지 살펴본다. 할인율, 브랜드인지도가 각각 가장 유의미한 변수로 선택되는 경우 다음과 같은 마케팅을 할 수 있다.

- 할인율이 가장 유의미한 변수인 경우
: 소비자들이 저렴한 가격 때문에 홈쇼핑을 이용한다는 근거이기 때문에 소비자들이 더 저렴한 가격이라고 느낄 수 있도록 어플리케이션 전용 쿠폰이나 L.point 를 활용한 마케팅을 한다. 이 때, 인앱결제 비율과 ARS 결제 비율을 비교하는 추가 분석을 한다면 더욱 적합한 마케팅을 할 수 있을 것이라고 보여진다.
- 브랜드인지도가 가장 유의미한 변수인 경우
: 신생 브랜드보다 브랜드인지도가 높은 상품을 위주로 편성하도록 한다.

범주형 자료인 방송시작시간, 월(月), 요일, 우천 여부, 휴일 여부, 방송시작시간, PD 코드, SH 코드, 상품중분류명으로 분류하여 매출액으로 HEATMAP 을 작성하여 확인해본다. 이에 따라 PD 와 SH 의 조합, 해당 날짜 및 날씨, 방송시작시간과 상품중분류가 매출에 끼치는 영향을 살펴본다. 이 중 매출에 큰 영향을 미칠 것으로 예상되는 변수와 그 조합은 다음과 같다.

- PD 와 SH 의 조합
: PD 와 SH 는 말을 수 있는 프로그램의 수가 한정적이므로 함께 했을 때 시너지효과를 발휘할 수 있는 PD 와 SH 를 찾아서 매칭해준다. PD, SH 의 매칭 결과를 토대로 매출이 높았던 상품들과 2 차 매칭을 시켜준다.
- 요일, 방송시작시간과 상품중분류의 조합
: 황금 요일과 황금 방송 시간대는 한정되어 있기 때문에 하루 혹은 일주일 단위의 매출을 따져 다른 시간대에도 잘 팔리는 상품은 다른 시간대로 편성을 하고 황금 시간대에 유난히 매출이 높았던 상품들을 그 시간에 편성하도록 한다.
- 휴일과 상품 중분류의 조합
: 휴일에 유난히 매출이 높은 상품들이 있다면 연휴 기간에 편성을 하도록 한다. 이외에 장기 연휴의 경우 연휴 1~2 주 전, 1~7 일 전, 연휴 기간 내, 연휴 기간 이후 별로 나눠 상품들의 매출 변화를 살펴본다.
- 월(月)과 상품중분류의 조합
: 전체적으로 매출이 낮은 달에는 충동구매를 일으키기 쉬운 저렴한 상품들을 위주로 편성한다.

[별첨] 추가가 필요할 경우에만 제출 (최대 3장 이내)

- 1번 각주에 작성한 논문의 표를 첨부함

【표IV-13.】 상품구입을 결정하는데 영향을 미친 요소

단위 : 명, (%)

구분	영향을 미치지 않음		보통	영향을 미침		합계
	전혀 영향을 미치지 않았다	영향을 미치지 않았다		영향을 미쳤다	아주 많이 영향을 미쳤다	
브랜드 유명도	95(19.4)		145(29.6)	250(51.0)		490(100.0)
	30(6.1)	65(13.3)	145(29.6)	203(41.4)	47(9.6)	
가격	33(6.7)		110(22.4)	349(70.8)		492(100.0)
	8(1.6)	25(5.1)	110(22.4)	276(56.0)	73(14.8)	
경품	195(40.5)		155(32.2)	132(27.4)		482(100.0)
	64(13.3)	131(27.2)	155(32.2)	114(23.7)	18(3.7)	
사은품	174(36.3)		139(29.0)	167(34.8)		480(100.0)
	57(11.9)	117(24.4)	139(29.0)	143(29.8)	24(5.0)	
결제조건	100(20.4)		118(24.1)	272(55.5)		490(100.0)
	36(7.3)	64(13.1)	118(24.1)	205(41.8)	67(13.7)	
쇼호스트의 권유	154(31.4)		163(33.3)	173(35.3)		490(100.0)
	47(9.6)	107(21.8)	163(33.3)	149(30.4)	24(4.9)	
원산지 및 제조처	98(20.0)		159(32.4)	234(47.6)		491(100.0)
	20(4.1)	78(15.9)	159(32.4)	196(39.9)	38(7.7)	
제품성능 및 제품시연	28(5.8)		123(25.3)	336(69.0)		487(100.0)
	10(2.1)	18(3.7)	123(25.3)	259(53.2)	77(15.8)	

- 3번 각주에 작성한 인터넷 기사의 사진을 첨부함
 - 기사: '쇼핑은 월요일 집중'... 티몬 상반기 요일별 매출 중 가장 높아

티몬 요일별 매출 추이 기준 2015년 상반기



