

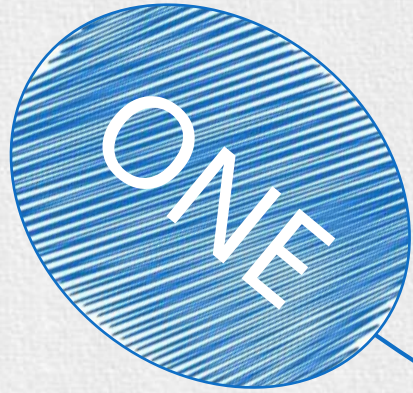
국내여행 활성화

2조 김민지 나혜성 이지영
이혜원 이소정 조수연





- 1 주제 선정이유 및 설명
- 2 데이터 수집/ 데이터 처리
- 3 MCA
- 4 Decision Tree
- 5 맞춤 활동 추천



주제 선정이유 및 설명



"국내 여행 2년 연속 하락, 해외여행 지출 증가 추세"

"한국인 해외여행 지출 크게 늘었다"

정유미 기자 youme@kyunghyang.com



특히 국제선 여행객은 7.9% 증가한 4556만 명으로 역대 1위를 기록했고 국내선 여행객은 1600만 명으로 1% 증가했다. 계절적 요인(겨울방학·연휴), 저비용항공사 공급력 확대(19.6%), 노선 다변화, 내·외국인 여행객 증가 등이 주요원인으로 분석된다.

한국경제연구원(이하 한경연)은 2011년~2017년 가계의 해외소비 추이를 분석한 결과 국내 소비는 연평균 1.6% 증가한 반면 해외소비는 연평균 10.1% 늘었으며 2011년 한해를 제외하고 해외소비 증가율이 국내소비 증가율보다 높았다고 26일 밝혔다. 특히 관광분야는 수입보다 지출이 많아 서비스수지(일반여행 부문) 적자가 2011년 대비 4.3배나 늘었다.

프리미엄 2018년 복숭아 농장 직거래 햇살바다

올해 상반기 해외여행객 4556만명 역대 최고

올해 상반기 해외여행객 4556만명 역대 최고

국토부, 중국노선 15.6% 늘어...일본노선 4% 증가불구 향후 감소 전망

김진강기자(kjk5608@skyedaily.com)

기사입력 2019-08-04 13:49:27

특히 국제선 여행객은 7.9% 증가한 4556만 명으로 역대 1위를 기록했고 국내선 여행객은 1600만 명으로 1% 증가했다. 계절적 요인(겨울방학·연휴), 저비용항공사 공급력 확대(19.6%), 노선 다변화, 내·외국인 여행객 증가 등이 주요원인으로 분석된다.



기록하여 단일 반기별 실적

특히 국제선 여행객은 7.9% 증가한 4556만 명으로 역대 1위를 기록했고 국내선 여행객은 1600만 명으로 1% 증가했다. 계절적 요인(겨울방학·연휴), 저비용항공사 공급력 확대(19.6%), 노선 다변화, 내·외국인 여행객 증가 등이 주요원인으로 분석된다.

분석주제

맞춤 활동 추천을 통한 국내 여행 활성화

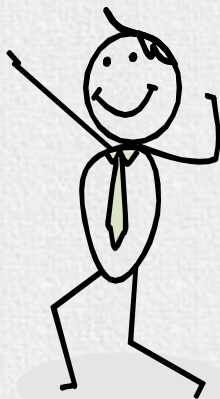
이용 데이터

신한카드(내국인) 데이터

- 거주지 주소(SI)
- 카드 가맹점(SIDO)
- 가맹점 업종(GB2)
- 성별(SEX_CCD)
- 연령대(CLN_AGE_R)
- 이용시간대(APV_TS_DL_TM_R)
- 이용건수(USEC)

분석전략

대분류 재설정
분석 목적에 맞는 데이터 재구성
MCA를 통한 연관성 높은 변수 추출
의사결정 나무를 통한 활동추천



TWO

데이터 수집 / 데이터 처리



기존 데이터 처리 및 변환

1 | 데이터 모으기

내국인 데이터 16,17,18

2 | 데이터 추출

sidon ≠ gu 데이터 지우기

3 | tappl를 통한 데이터 통합

같은 분류로 VLM, USEC 통합

자택시	가맹시	가맹구	업종 대분류	업종 소분류	성별
연령대	이용년월	요일	이용시간대	취급액	이용건수

시/도	가맹점	요일	시간대
취급액	이용건수	성별	연령대

si	sido	gu	GB3	GB2	SEX_CCD	CLN_AGE_R	TA_YM	DAW_CCD_R	APV_TS_DL_TM_R	VLM	USEC	
1	강원	경기	구리시	여행	숙박	M	20	201601	WHITE	취침	147180	4
2	강원	서울	용인구	여행	숙박	M	20	201601	RED	취침	468300	17
3	강원	강원	태백시	여행	숙박	M	20	201601	RED	활동	2519900	49
4	강원	인천	종구	취미오락	외식	M	20	201601	RED	취침	520036	31
5	강원	부산	사상구	취미오락	외식	M	20	201601	RED	활동	2293778	115
6	강원	강원	영월군	취미오락	외식	M	20	201601	RED	활동	16109966	1199
7	강원	경기	수원시	취미오락	유통	M	20	201601	RED	휴식	10447550	254
8	강원	강원	춘천시	취미오락	유통	M	20	201601	RED	활동	4125946	383
9	강원	경기	성남시	여행	체험	F	20	201601	WHITE	활동	1739400	4
10	강원	서울	강남구	여행	체험	F	20	201601	WHITE	활동	2908366	71
11	강원	경기	성남시	취미오락	교육훈련	M	20	201601	WHITE	활동	1588652	8
12	강원	광주	광안구	취미오락	종합쇼핑	F	20	201601	WHITE	휴식	245077	22
13	강원	강원	평창군	취미오락	종합쇼핑	F	20	201601	WHITE	휴식	4068635	280
14	강원	강원	강릉시	취미오락	종합쇼핑	F	20	201601	WHITE	활동	192610675	9018
15	강원	전북	익산시	취미오락	패션쇼핑	M	20	201601	RED	활동	1689537	35
16	강원	경기	고양시	취미오락	패션쇼핑	F	20	201601	RED	활동	1594896	49
17	강원	경기	수원시	취미오락	패션쇼핑	F	20	201601	WHITE	휴식	2016366	53
18	강원	경기	포천시	취미오락	패션쇼핑	M	20	201601	RED	활동	3321808	57
19	강원	인천	부평구	취미오락	패션쇼핑	F	20	201601	WHITE	휴식	1479248	66
20	강원	부산	남구	취미오락	패션쇼핑	F	20	201601	RED	활동	2105120	124
21	강원	인천	남동구	스포츠활동	레저스포츠	F	20	201601	RED	휴식	317552	13
22	강원	강원	철원군	스포츠활동	레저스포츠	M	20	201601	WHITE	휴식	5590610	294
23	강원	강원	동해시	취미오락	인터넷게임	M	20	201601	WHITE	활동	1630576	200
24	강원	충북	청주시	스포츠활동	스포츠용품구매	F	20	201601	WHITE	활동	1317484	26
25	강원	강원	동해시	스포츠활동	스포츠용품구매	F	20	201601	WHITE	휴식	14128388	98



	sido	GB	DAY	TIME	VLM	USEC	SEX	AGE
1	강원	골프	RED	취침	12263691	438	F	20
2	강원	골프	RED	활동	1194861431	44161	F	20
3	강원	골프	RED	휴식	195336121	6903	F	20
4	강원	골프	WHITE	취침	23803150	512	F	20
5	강원	골프	WHITE	활동	2094602728	52751	F	20
6	강원	골프	WHITE	휴식	409640690	9223	F	20
7	강원	골프	RED	취침	44085190	898	F	30
8	강원	골프	RED	활동	3830109237	73424	F	30
9	강원	골프	RED	휴식	672999906	11533	F	30
10	강원	골프	WHITE	취침	89379500	1188	F	30
11	강원	골프	WHITE	활동	7260775464	90745	F	30
12	강원	골프	WHITE	휴식	1110037756	16475	F	30
13	강원	골프	RED	취침	137018048	1830	F	40
14	강원	골프	RED	활동	10239753289	105217	F	40
15	강원	골프	RED	휴식	1473119730	16524	F	40
16	강원	골프	WHITE	취침	239636993	3193	F	40
17	강원	골프	WHITE	활동	22744319169	216315	F	40
18	강원	골프	WHITE	휴식	3273384588	34972	F	40
19	강원	골프	RED	취침	149841469	1545	F	50
20	강원	골프	RED	활동	9383480206	66823	F	50
21	강원	골프	RED	휴식	1336891391	10452	F	50
22	강원	골프	WHITE	취침	307373153	3613	F	50
23	강원	골프	WHITE	활동	33061692543	272873	F	50

Raw Data

Refind Data

1 데이터 모으기

	si	sido	gu	GB3	GB2	SEX_CCD	CLN_AGE_R	TA_YM	DAW_CCD_R	APV_TS_DL_TM_R	VLM	USEC
1	강원	경기	구리시	여행	숙박	M	20	201601	WHITE	취침	147180	4
2	강원	서울	용산구	여행	숙박	M	20	201601	RED	취침	468300	17
3	강원	강원	태백시	여행	숙박	M	20	201601	RED	활동	2519900	49
4	강원	인천	중구	취미오락	외식	M	20	201601	RED	취침	520036	31
5	강원	부산	사상구	취미오락	외식	M	20	201601	RED	활동	2293778	115
6	강원	강원	영월군	취미오락	외식	M	20	201601	RED	활동	16109966	1199
7	강원	경기	수원시	취미오락	유흥	M	20	201601	RED	휴식	10447550	254
8	강원	강원	춘천시	취미오락	유흥	M	20	201601	RED	활동	4125946	383
9	강원	경기	성남시	여행	체험	F	20	201601	WHITE	활동	1739400	4
10	강원	서울	강남구	여행	체험	F	20	201601	WHITE	활동	2908366	71
11	강원	경기	성남시	취미오락	교육훈련	M	20	201601	WHITE	활동	1588652	8
12	강원	광주	광산구	취미오락	종합쇼핑	F	20	201601	WHITE	휴식	245077	22
13	강원	강원	평창군	취미오락	종합쇼핑	F	20	201601	WHITE	휴식	4068635	280
14	강원	강원	강릉시	취미오락	종합쇼핑	F	20	201601	WHITE	활동	192610675	9018

2 국내 여행 데이터 추출

si(자택시) 와 sido(카드 가맹점 주소) 다른 데이터
만 추출

	sido	GB	SEX_AGE	DAY	TIME	VLM	USEC
1	강원	골프	F20	RED	취침	12263691	438
2	강원	골프	F20	RED	활동	1194861431	44161
3	강원	골프	F20	RED	휴식	195336121	6903
4	강원	골프	F20	WHITE	취침	23803150	512
5	강원	골프	F20	WHITE	활동	2094602728	52751
6	강원	골프	F20	WHITE	휴식	409640690	9223
7	강원	골프	F30	RED	취침	44085190	898
8	강원	골프	F30	RED	활동	3830109237	73424
9	강원	골프	F30	RED	휴식	672999906	11533
10	강원	골프	F30	WHITE	취침	89379500	1188
11	강원	골프	F30	WHITE	활동	7260775464	90745
12	강원	골프	F30	WHITE	휴식	1110037756	16475
13	강원	골프	F40	RED	취침	137018048	1830
14	강원	골프	F40	RED	활동	10239753289	105217

TWO

데이터 수집/ 데이터 처리

3 **tapply**를 통한 데이터 통합

tapply(local\$USEC,

계산하고자 하는 변수

list(local_5\$sido,local_5\$GB1,local_5\$SEX_AGE,local_5\$DAY,local_5\$TIME), sum)

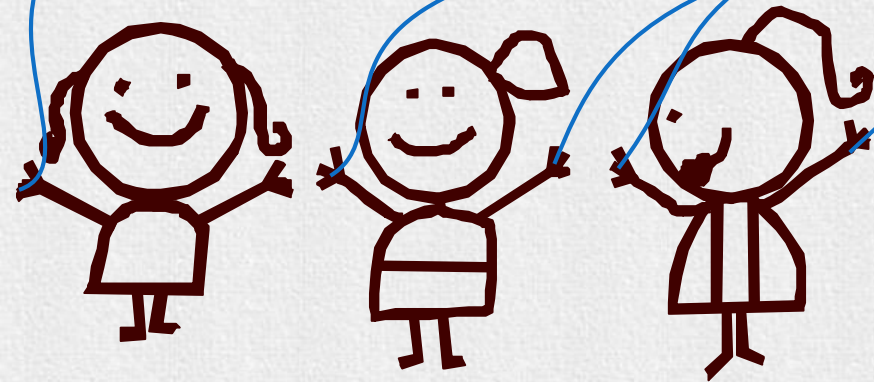
조건 변수

계산 함수

T H R E E

다중대응분석(MCA)

분할표에 포함된 정보의 연관성을
이해하기 쉽게 그래프로 표현

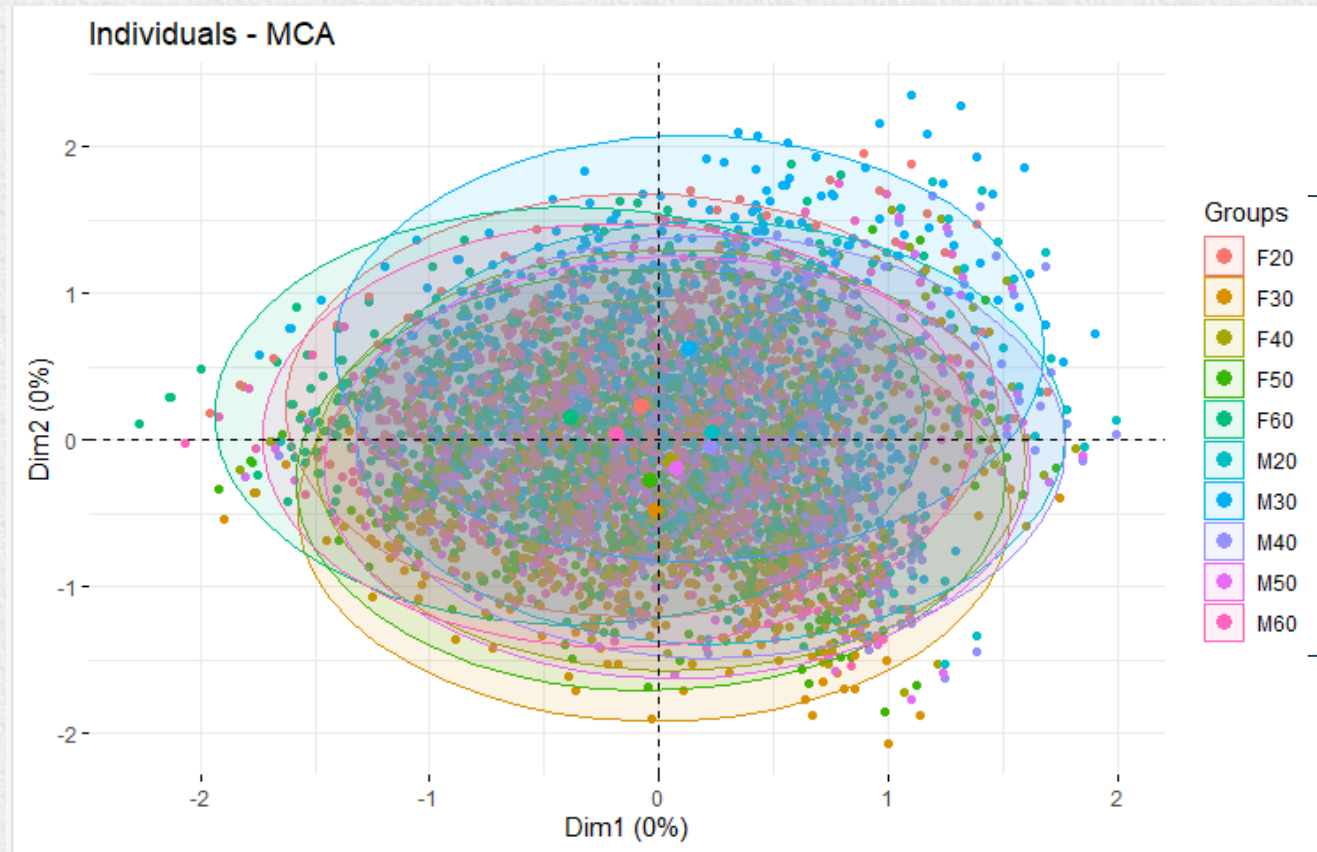


시행착오



MCA (모든 변수)

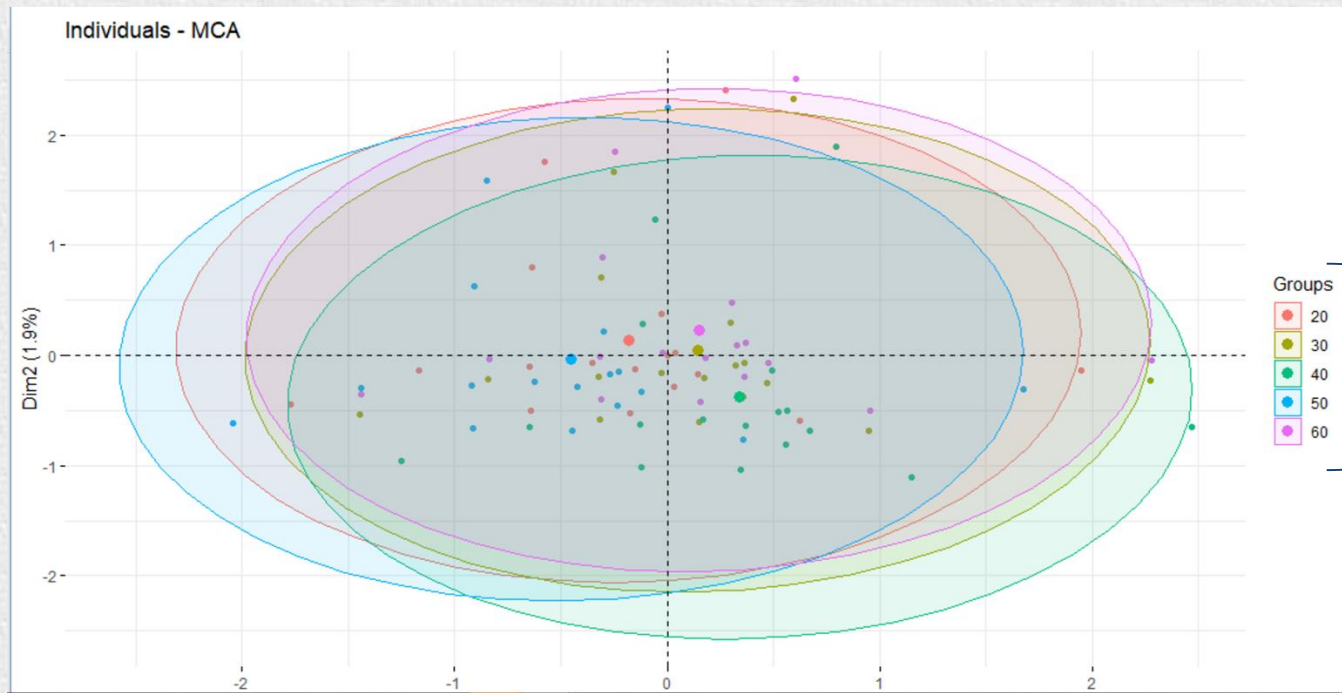
THREE



성별/연령대로
차이 없음

MCA (side&age)

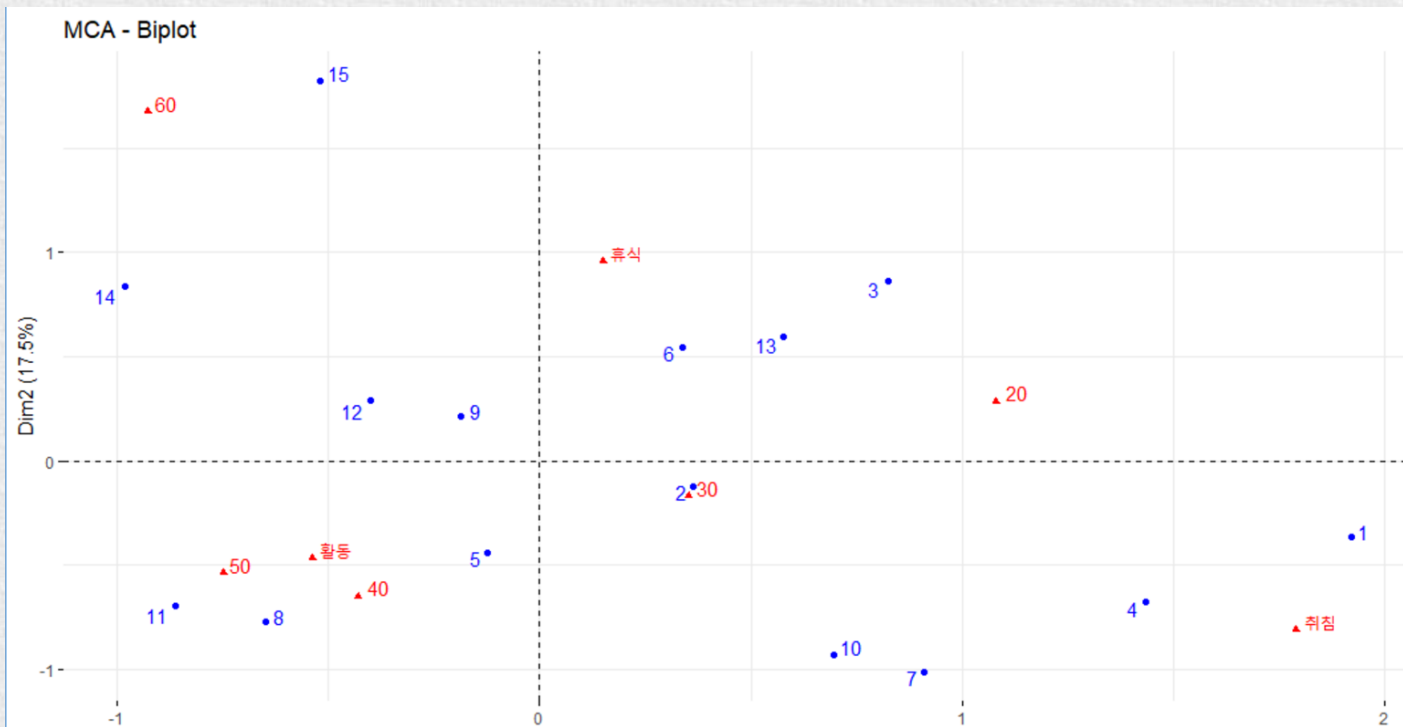
THREE



연령대별로
여행 지역 **차이 없음**

MCA (time&age)

THREE



군집화 가능하나 데이터 수가 **15개**밖에 안됨



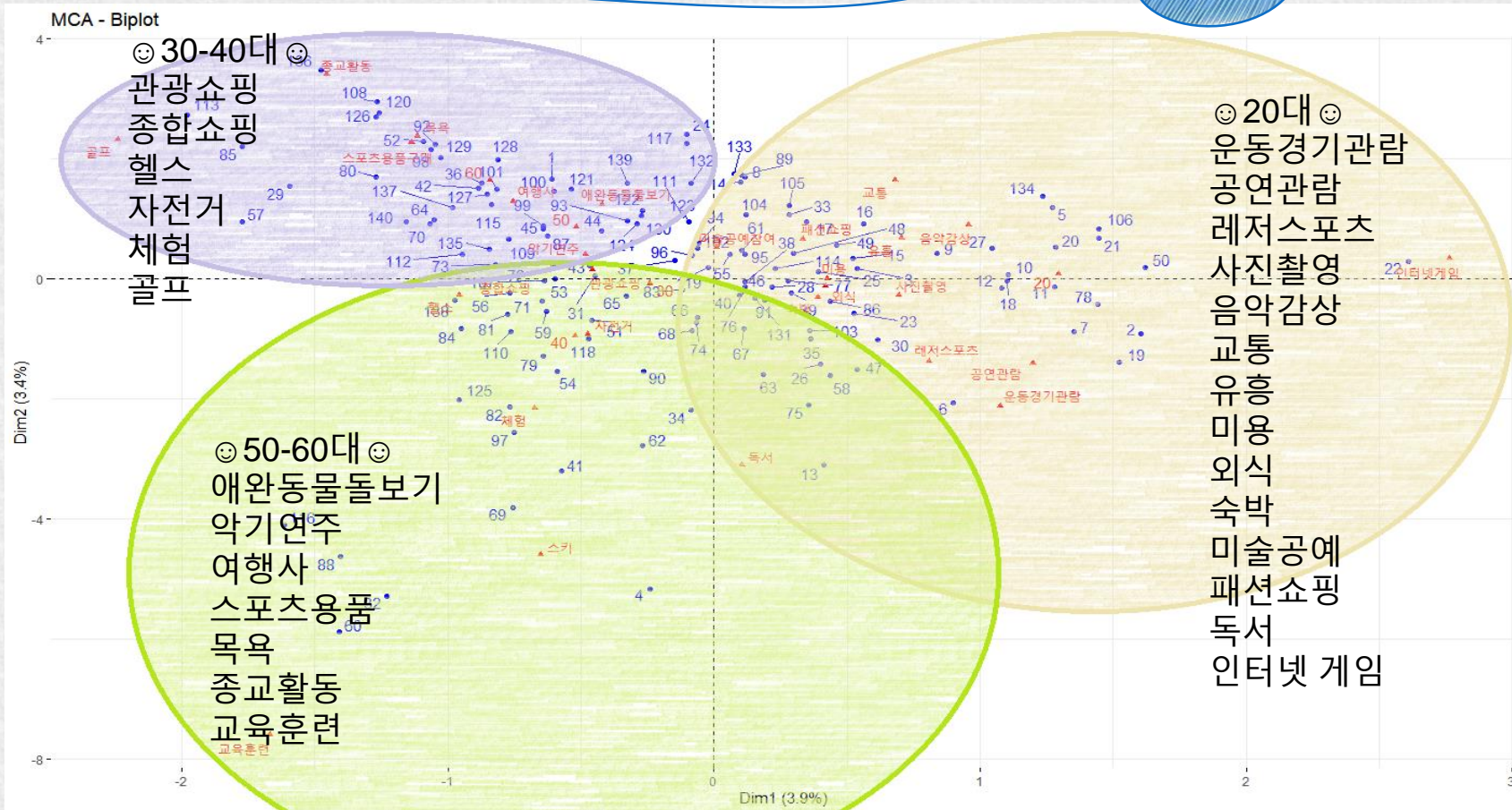
일반화 하기엔 데이터 수가 적어 **적합하지 않**다고 판단

최종선택



MCA(연령대&활동)

THREE



```
install.packages(c("FactoMineR", "factoextra"))
```

MCA MCA 시각화

```
library("FactoMineR")  
library("factoextra")
```

#MCA 실행#

```
res_local <- MCA(local, graph=FALSE, row.w=USEC)
```

USEC 가중값 설정

#시각화#

```
fviz_screplot(res_shc16_1, addlabels = TRUE, ylim = c(0, 45))
```

#전체 그림#

```
fviz_mca_biplot(res_shc16_1,  
                repel = TRUE,  
                ggtheme = theme_minimal())
```


#각 개체들 그래프#

```
fviz_mca_ind(res_shc16_1, col.ind="steelblue")  
fviz_mca_ind(res_shc16_1, col.ind = "cos2", repel = TRUE)
```

#연령으로 나누기, 그룹핑#

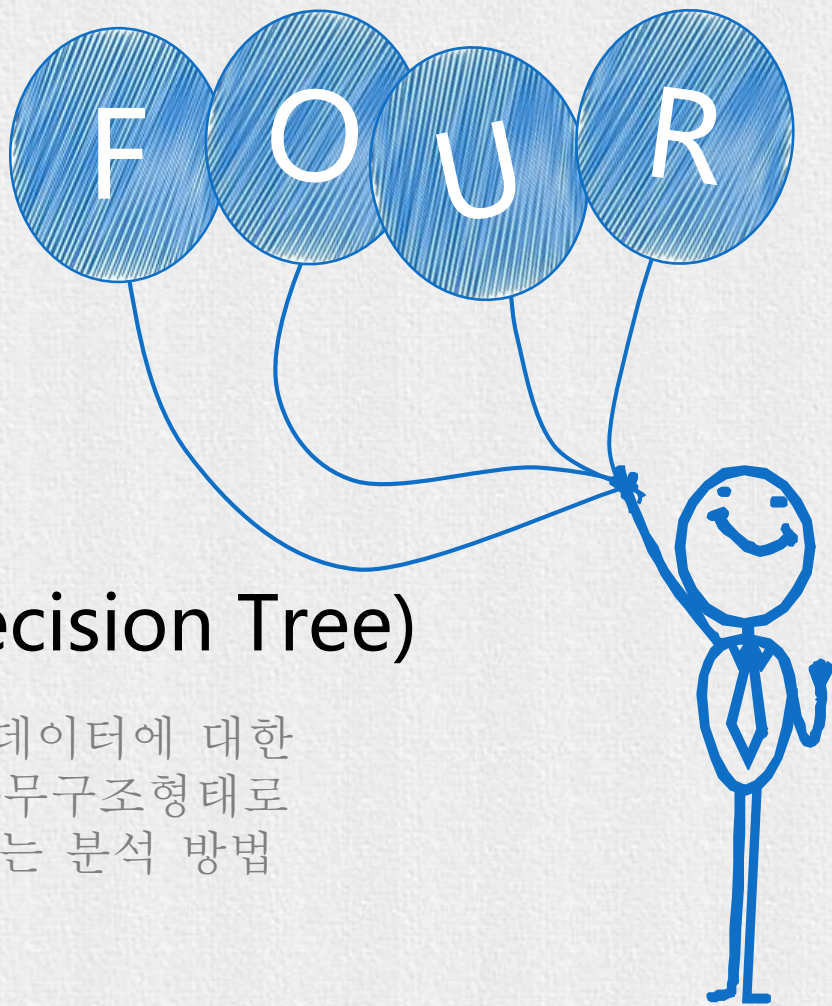
```
age <- as.factor(local[, "age"])  
fviz_mca_ind(res_local, label="none", habillage=age)
```

#연령으로 그룹핑#

```
p <- fviz_mca_ind(res_local, label="none", habillage=age, addEllipses=TRUE, ellipse.level=0.95)  
print(p)
```

#시도별로 표시, 그룹핑#

```
sido <- as.factor(local[, "sido"])  
q <- fviz_mca_ind(res_local, label="none", habillage=sido, addEllipses=TRUE, ellipse.level=0.95)  
print(q)
```

의사결정나무(Decision Tree)

지도학습의 일종으로, 데이터에 대한
분류를 기초로 하여 나무구조형태로
데이터에서 규칙을 찾는 분석 방법

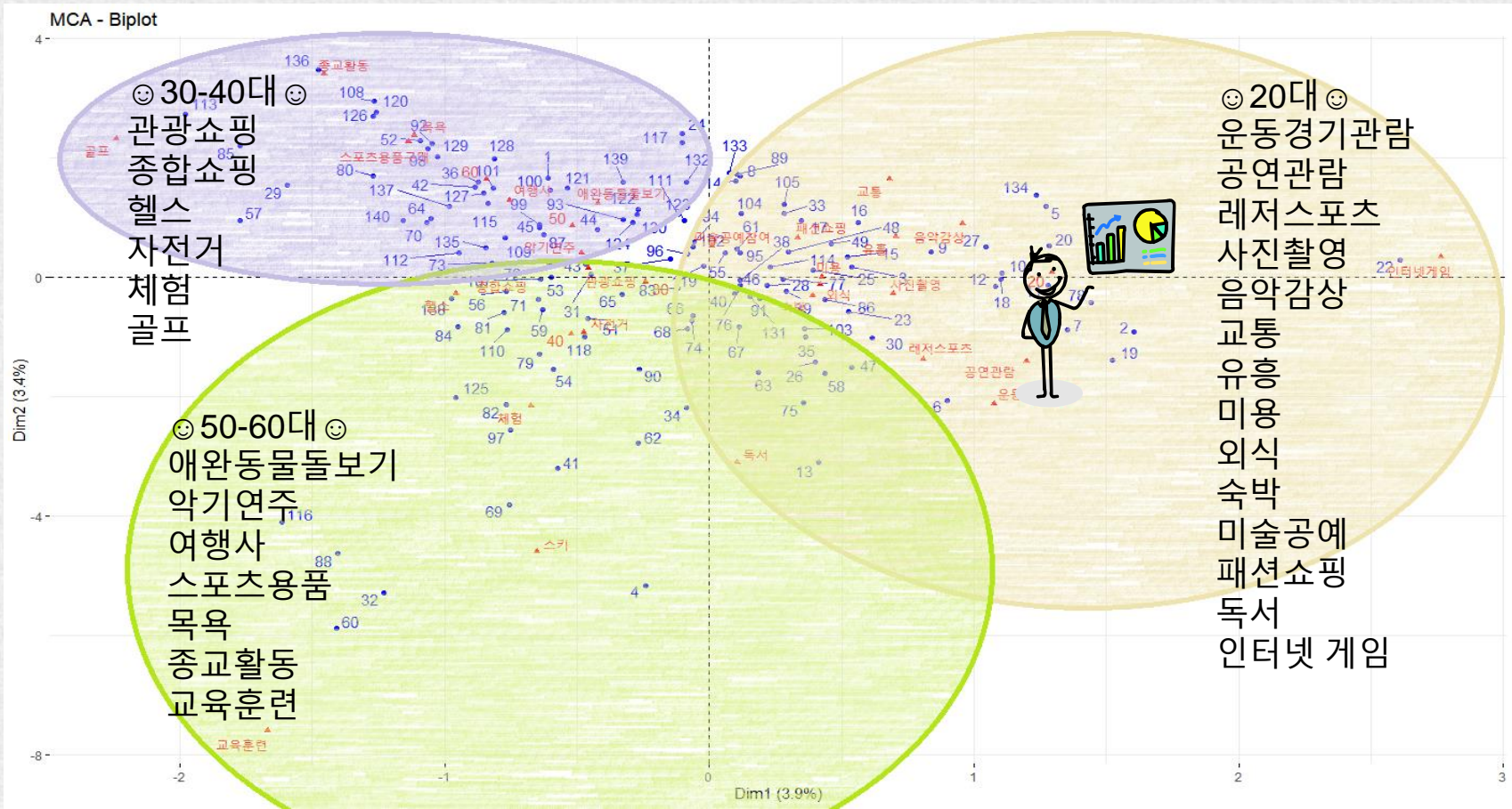


* $BUD = VLM / USEC$

변수명	변수설명	변수값
sido	카드 가맹점 주소	
GB	가맹점 소분류	
DAY	요일	a : 주말 b : 주중
TIME	이용시간대	a : 취침 (02-07) b : 활동 (07-19) c : 휴식 (19-02)
SEX	성별	M : 남성 F : 여성
AGE	나이	20 30 - 40 50 - 60
BUD	예산	

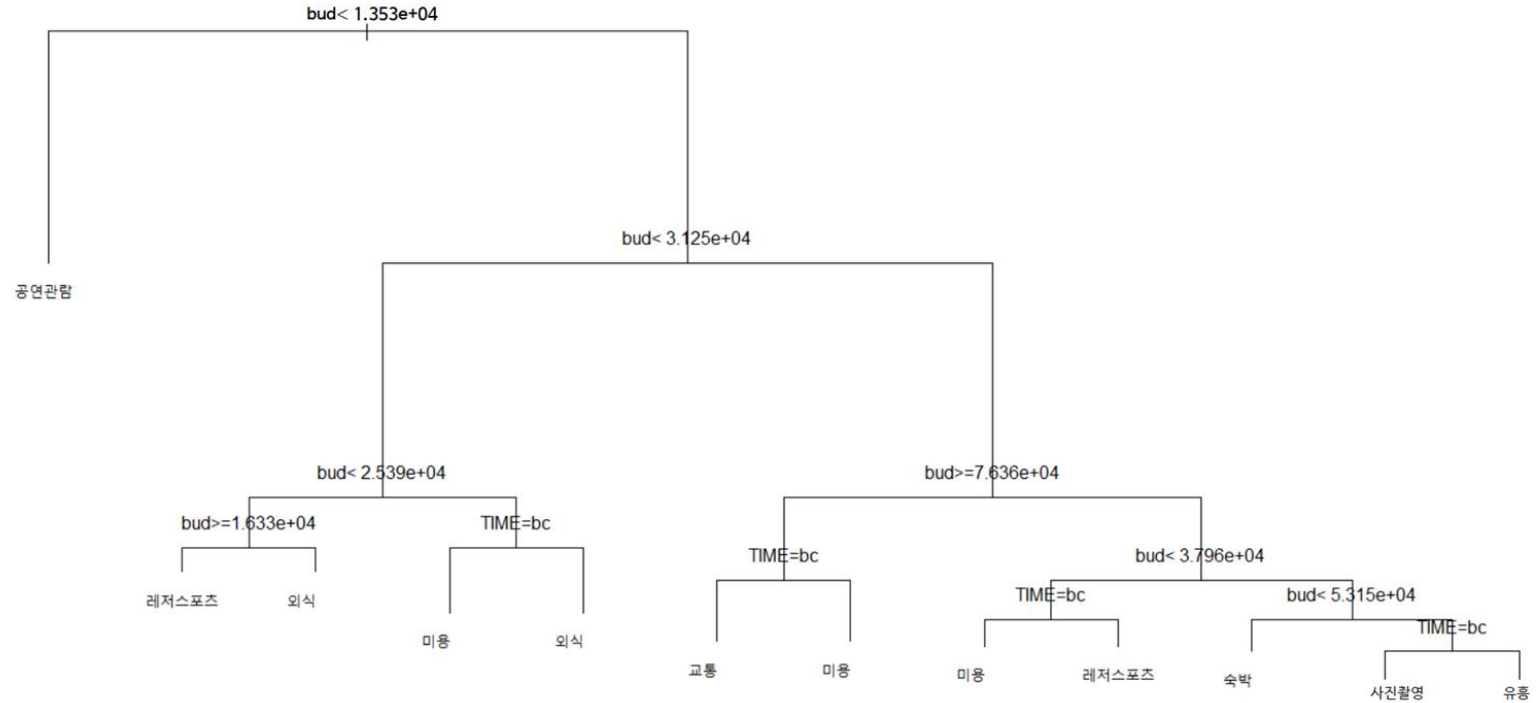
Decision Tree

FOUR



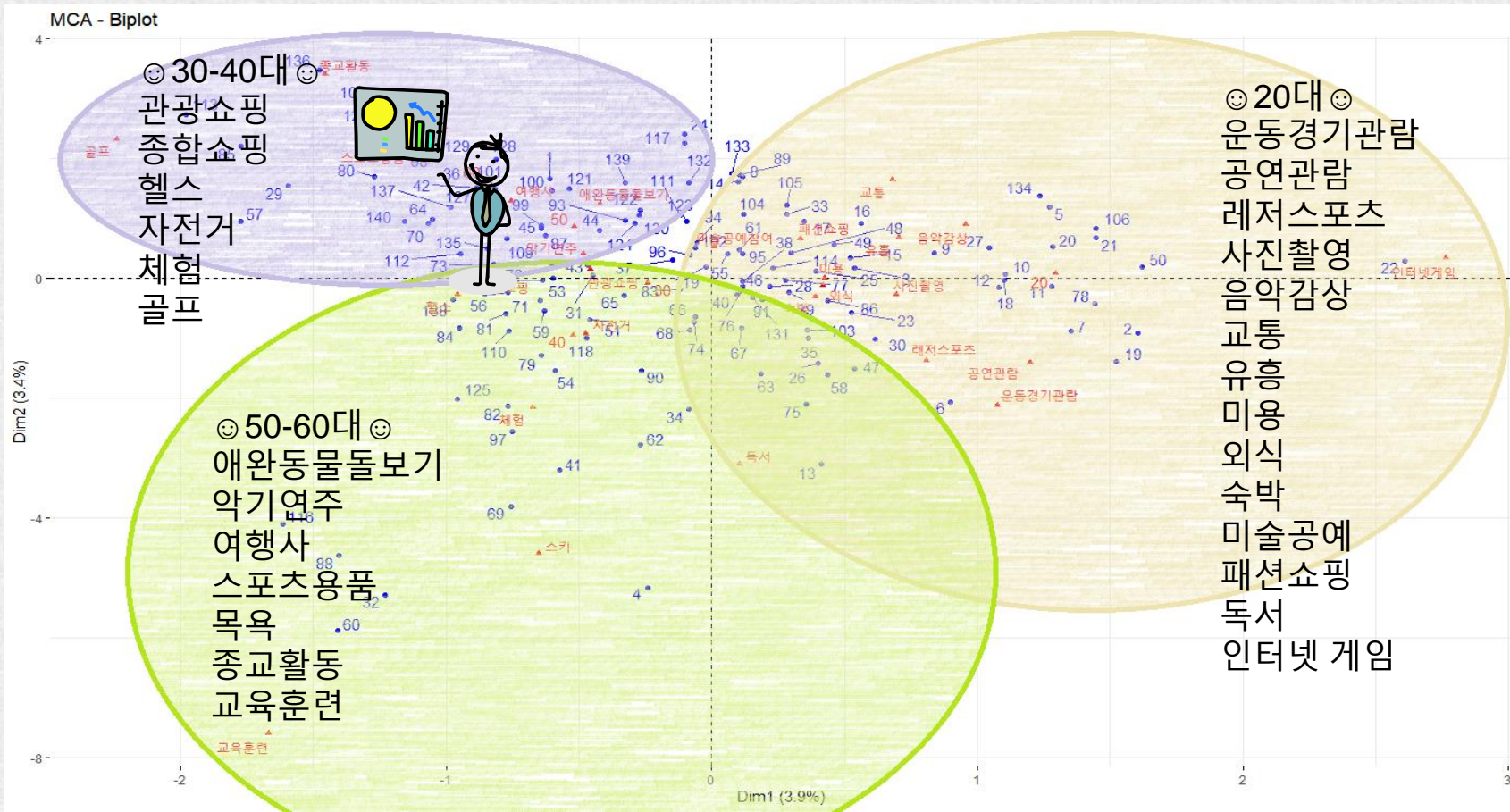
Decision Tree (20대)

FOUR



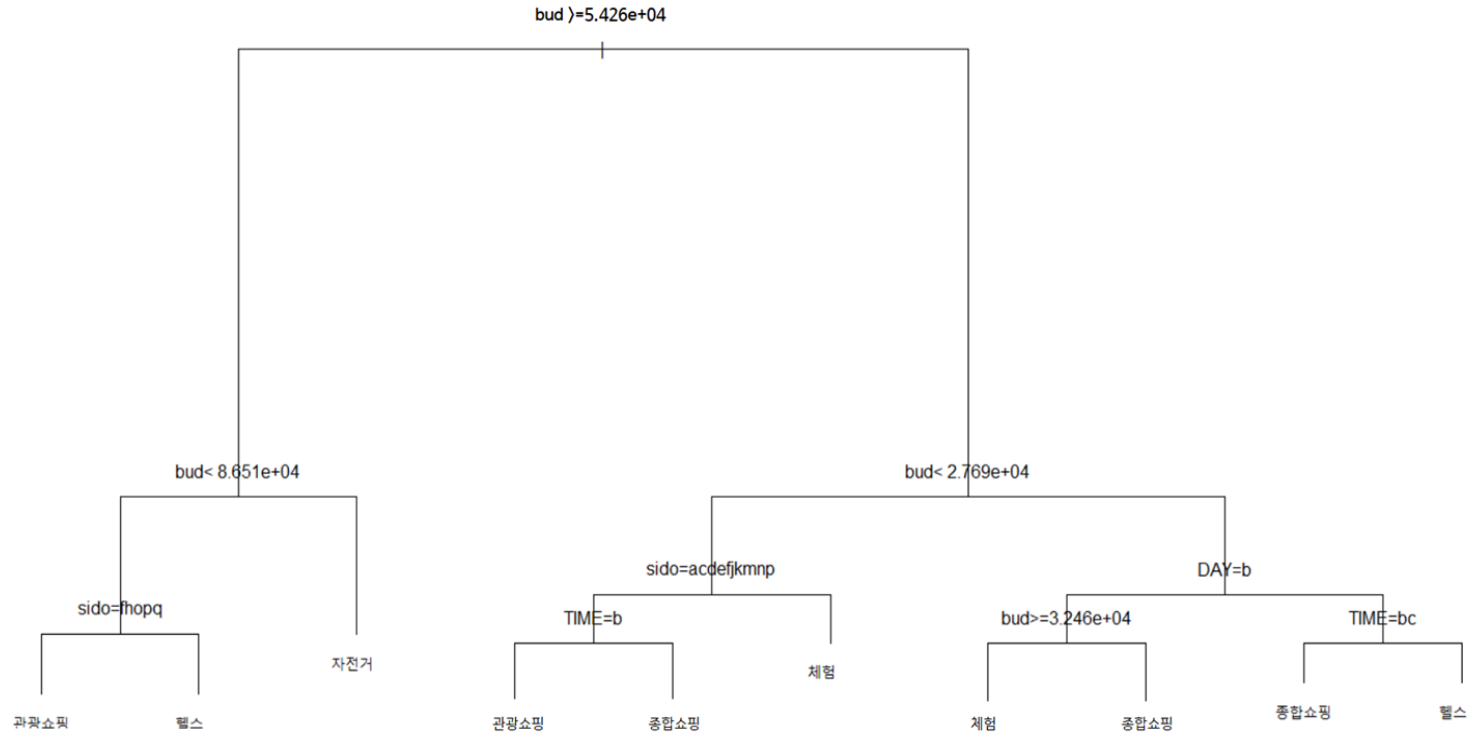
Decision Tree

FOUR



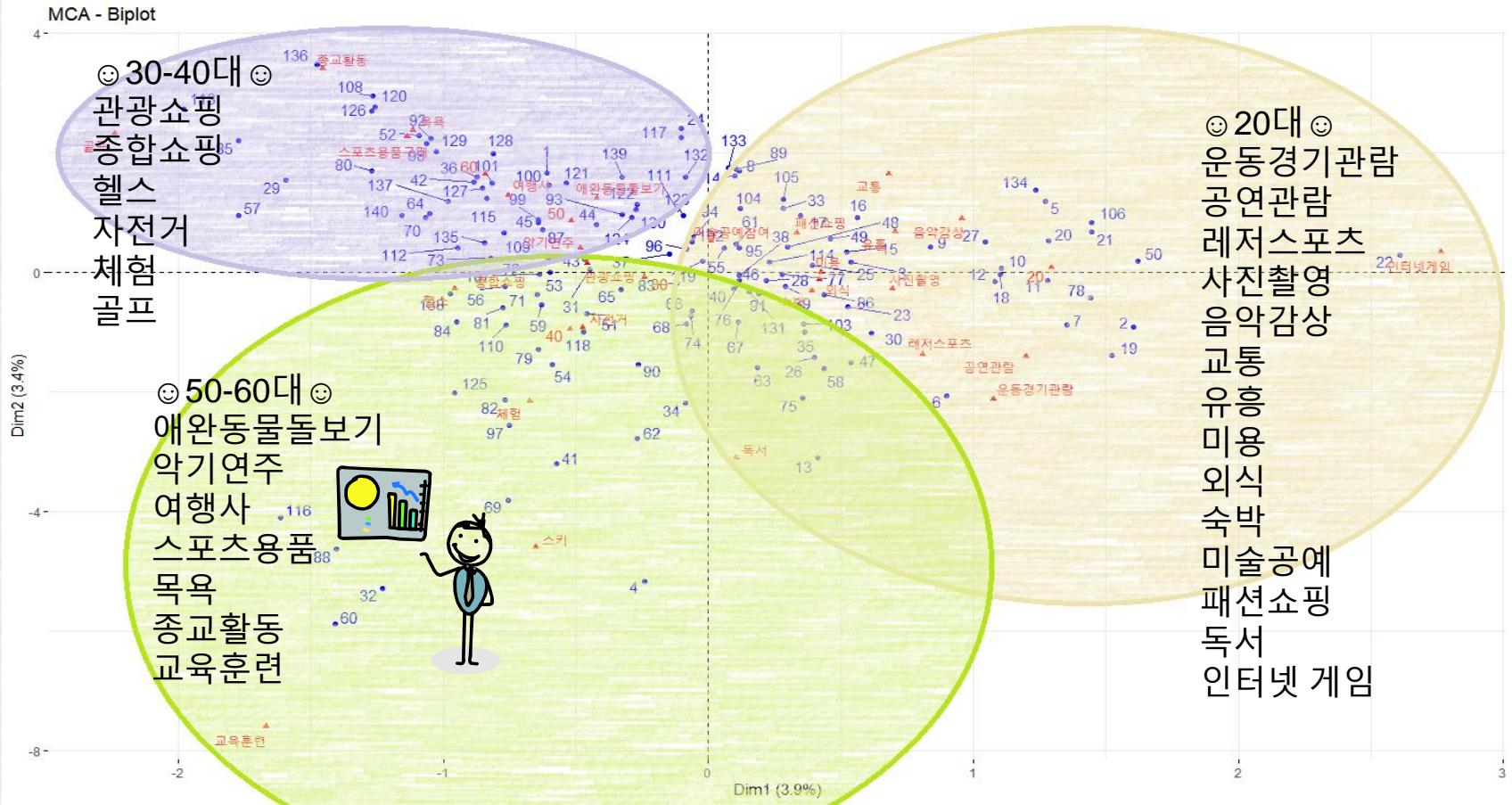
Decision Tree (30-40대)

FOUR



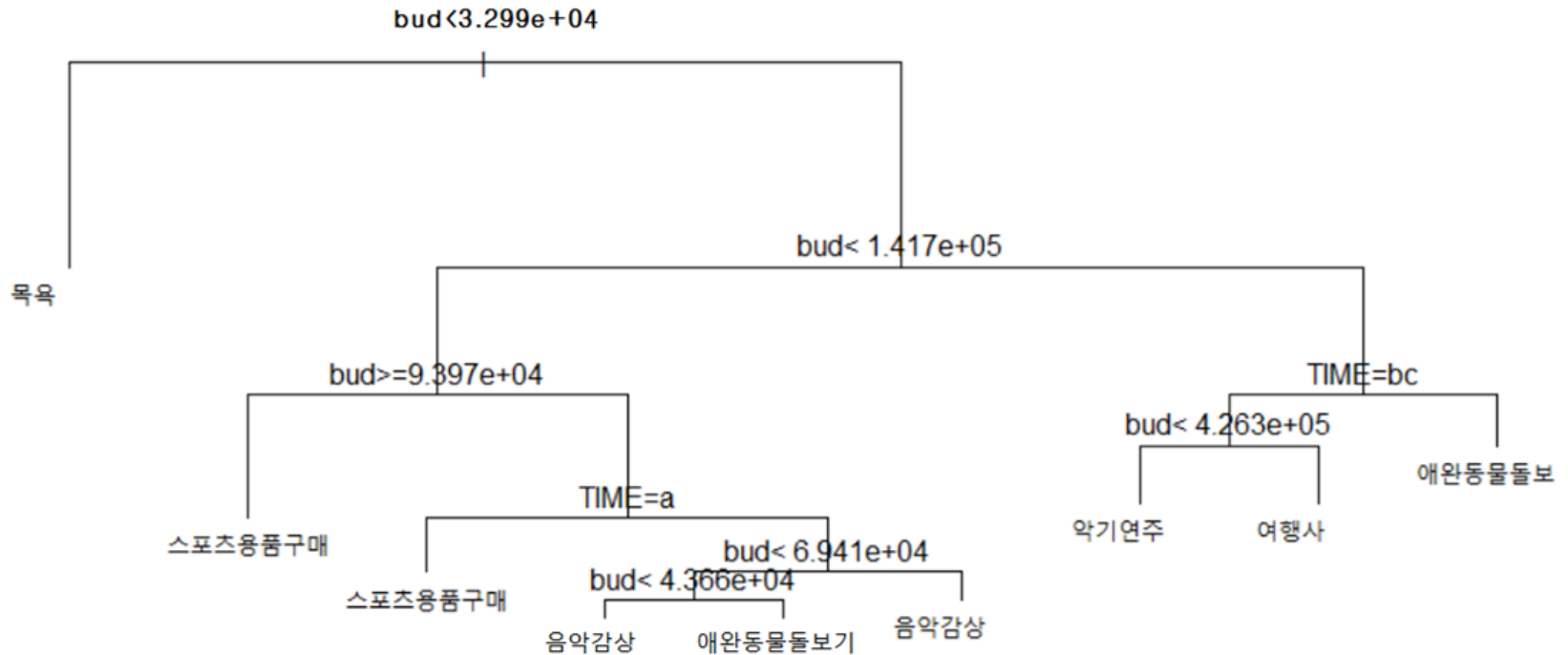
Decision Tree

FOUR



Decision Tree (50-60대)

FOUR



```
library('dplyr')
```

```
# 활동 필터링#
```

```
for20<-local_final %>% filter(GB %in% c("운동경기관람", "공연관람", "레저스포츠", "사진촬영", "음악감상", "교통", "유흥", "미용", "외식", "숙박"))
```

```
#연령별로 필터링#
```

```
for20<-for20%>%filter(AGE==20)
```

```
# 새로운 변수 설정#
```

```
for20_new <- for20 %>% mutate(bud=VLM/USEC)
```

```
#변수 삭제#
```

```
for20_new<-for20_new%>%select(-USEC,-VLM)
```


Decision Tree (R code)

FOUR

#결정트리생성#

```
install.packages("rpart")
```

```
library(rpart)
```

```
tree20_new <- rpart(GB~ ., data = for20_new, method="class" )
```

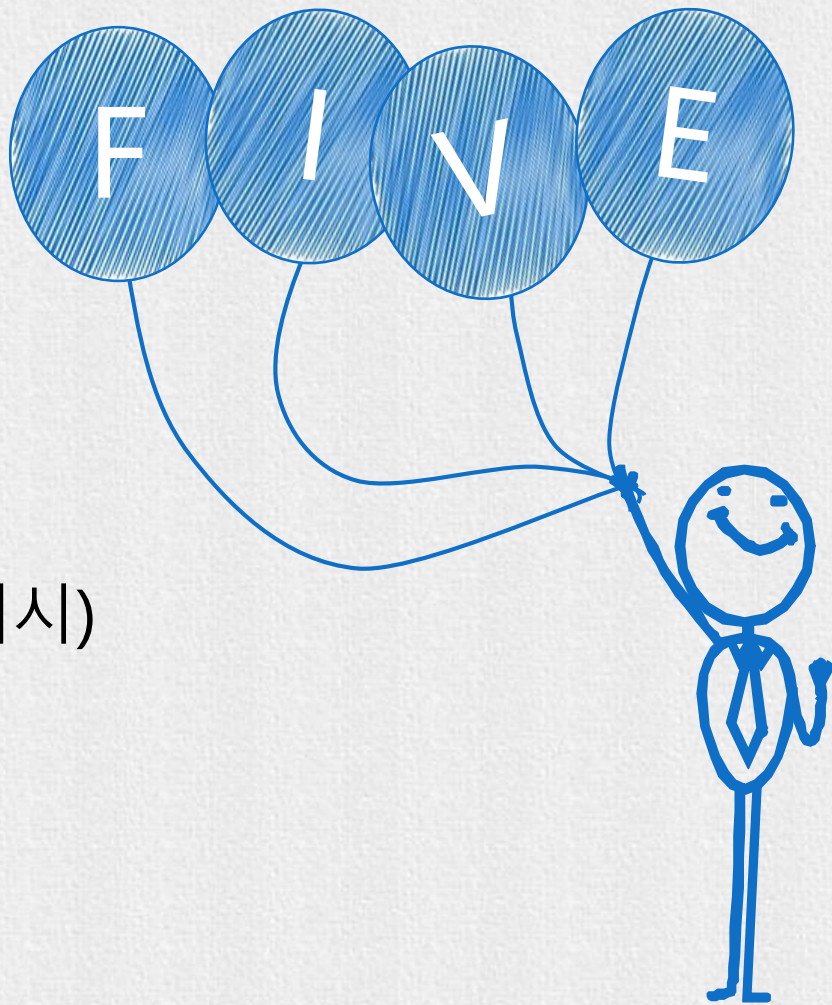
#Decision Tree그리기#

```
plot(tree20_new)
```

#텍스트 추가#

```
text(tree20_new)
```

맞춤 활동 추천 (예시)





안녕!

안식년을 맞아 힐링 여행을 가고 싶은데
뭘 하고 놀아야할지 모르겠네?

돈은 충분히 있으니 25만원 이상 오케이~

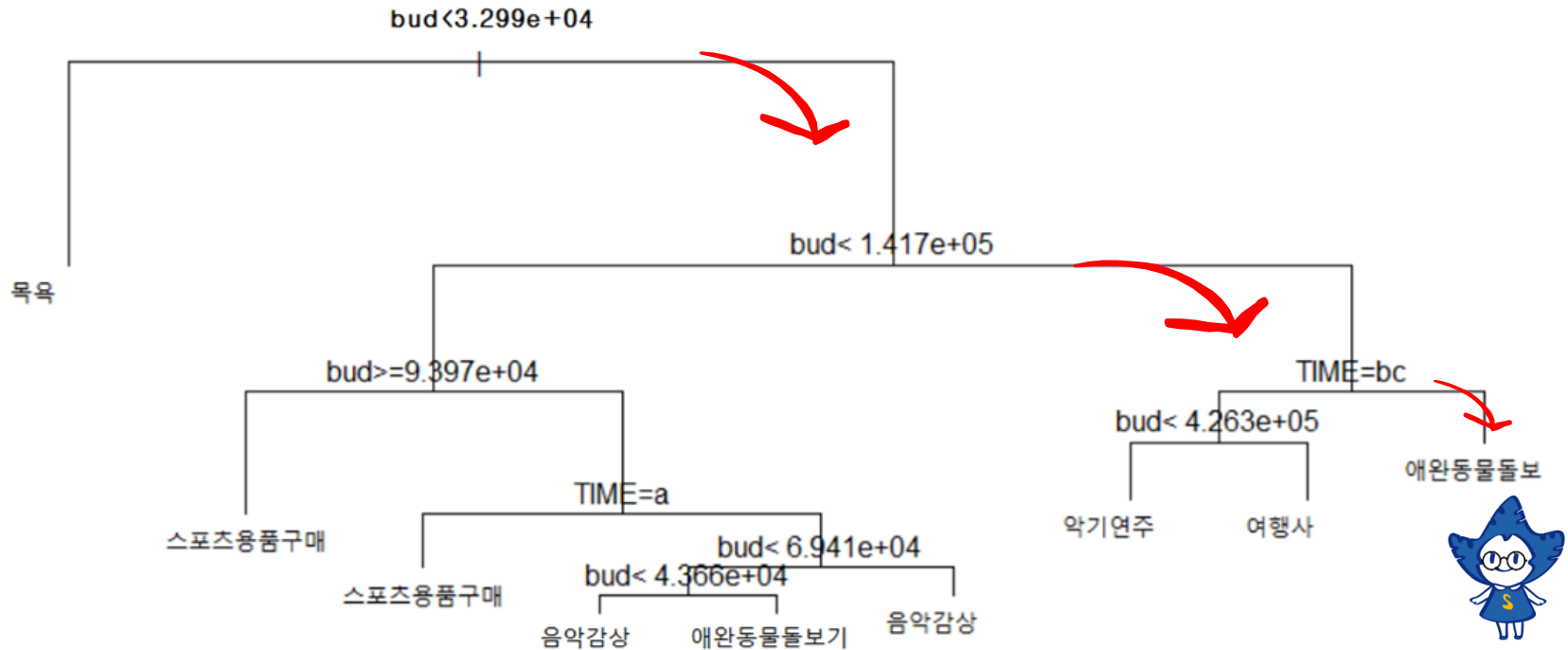
쏘리~ 새벽부터 부지런히 가야겠어

우리 학생들이 만든 추천시스템 한번 써볼까? 허허~

통계학과 학과장 여눈송(5X)

Decision Tree (50-60대)

FIVE



THANKS FOR WHATCHING

2조 김민지 나혜성 이지영
이혜원 이소정 조수연

