# source(file = "Script/proc/rank\_Script.R", encoding = "UTF-8")

#

# # 로지스틱 회귀 : rank\_Script.R

# plot(prf, main = "ROC Curve")

# 전반적인 상황 확보

# 기간별 주요 참여 활동(중복응답) 꺾은선 그래프

df <- read.csv(file = "DB/data\_combined.csv", header = T, stringsAsFactors = F)

MainAct <- read.csv("DB/고려요인.csv", stringsAsFactors = F, header = T)

colnames(MainAct)

str(MainAct)

# ggplot을 위해 년도만 가져옴

library(dplyr)

end\_Idx <- length(MainAct$기간)

# 년도 그룹

MainAct

library(ggplot2)

library(reshape2)

# melt\_MainAct <- melt(data = MainAct, id = c("기간"))

# ggplot로 그리기

y <- 2014

yTemp <- 0

# ggplot 4년치

for (idx in y:(y+4)) {

yTemp <- MainAct %>% filter(substr(기간, 1, 4) == idx)

yTemp$기간 <- substr(yTemp$기간, 5, 6)

melt\_MainAct <- melt(data = yTemp, id = c("기간"))

print(

ggplot(data = melt\_MainAct, aes(x = 기간 , y = value, group = variable, color = variable)) +

geom\_line(size = 1.5) +

labs(title = paste(idx, "재방문 시 고려요인(중복)")))

}

# 주요 관광수요 산업 컬럼을 구분하기 위해 카이제곱을 실시 해야함

# 이를 위해 범주형 데이터가 필요함. 이는 만족도 범주형을 가지고옴

df\_Temp <- read.csv("DB/만족도.csv", header = T, stringsAsFactors = T)

length(df\_Temp)

# 범주형 데이터 합치기

df\_Temp

str(MainAct)

MuxDf <- cbind(MainAct[-1], df\_Temp)

MuxDf

# 전반 적인 만족도와 나머지 데이터간의 관련성 여부

# 귀무 가설(H0) : 재방문시 고려요인 과 전반적인 만족도 걸럼과 관계가 있다.

# 연구 가설(H1) : 재방문시 고려요인 과 전반적인 만족도 걸럼과 관계가 없다.

# install.packages('gmodels')

library(gmodels)

end <- 0

end <- length(MainAct[-1]) # 컬럼 갯수 만큼 카이검정

chiList <- 0

for (idx in 1:end) {

ta <- CrossTable(unlist(MuxDf[idx]), MuxDf$만족도, chisq = T)

chiList[idx] <- ta$chisq$p.value

}

# 데이터 정렬 및 검증

length(chiList)

chi\_df <- data.frame(p\_Value=chiList)

chi\_df <- chi\_df %>% mutate(result = ifelse(p\_Value>=0.05, "채택", "기각"))

chi\_df <- chi\_df %>% arrange(desc(p\_Value))

row.names(chi\_df) <- colnames(MainAct[-1])

chi\_df

# 카이제곱 검정을 실시 하였고 재방문 고려요인과 만족도 간의 관계가 증명 되었다.

# p\_Value result

# 휴양휴식 0.7294646 채택

# 유흥\_오락 0.6443329 채택

# 뷰티관광 0.6092449 채택

# 쇼핑.1 0.5931206 채택

# 음식\_미식탐방 0.5622067 채택

# 자연풍경감상 0.5573205 채택

# 역사\_문화유적 0.5084509 채택

# 패션\_유행\_.등\_.세련된\_문화 0.4700696 채택

# K.POP\_한류스타\_팬미팅 0.4313990 채택

# 경제적인\_여행비용 0.4200565 채택

# 유흥\_놀이시설 0.4049757 채택

# 숙박시설\_편리한.교통 0.3984646 채택

# 이미용\_서비스 0.3710314 채택

# 기후\_뚜렷한\_사계절 0.3234574 채택

# 이후 내림차순으로 P-Value를 정렬하여 이를 통해 잠정적인 상승이 이루어진 요인을 추출함.

# P-Value 값이 가장 큰 요인을 가지고 안정적인 상승세를 이룬 주요 산업을 가지고 정책을 고민 해봄

# Val 설명

# MainAct : 고려요인

# MuxDf : 고려요인 x 만족도

reDf <- read.csv("DB/재방문횟수.csv", header = T, stringsAsFactors = F)

reDf

endIdx <- length(reDf)

model <- list()

# 재방문 횟수 별 컬럼 요인

# 포뮬러 미동작

# for (idx in 1:endIdx) {

# tempDf <- data.frame()

# tempDf <- cbind(MainAct[-1] ,reDf[idx])

# name <- colnames(reDf[idx])

# fo <- name ~ .

# model\_temp <- lm(formula = fo, data = tempDf)

# model <- c(model, model\_temp)

# }

df01 <- cbind(MainAct[-1] ,reDf[1])

df02 <- cbind(MainAct[-1] ,reDf[2])

df03 <- cbind(MainAct[-1] ,reDf[3])

df04 <- cbind(MainAct[-1] ,reDf[4])

modelX1 <- lm(formula = X1회 ~ ., data = df01)

modelX2 <- lm(formula = X2회 ~ ., data = df02)

modelX3 <- lm(formula = X3회 ~ ., data = df03)

modelX4 <- lm(formula = X4회이상 ~ ., data = df04)

modelX1$coefficients

modelX2$coefficients

modelX3$coefficients

modelX4$coefficients

# > modelX1$coefficients

# (Intercept) 휴양휴식

# -15.12584903 -0.30769897

# 유흥\_오락 뷰티관광

# 0.57903869 1.72454818

# 쇼핑.1 음식\_미식탐방

# 0.61130324 -0.30517035

# 자연풍경감상 역사\_문화유적

# 0.32429982 -0.29787166

# 패션\_유행\_.등\_.세련된\_문화 K.POP\_한류스타\_팬미팅

# 0.22488922 -0.09528488

# 경제적인\_여행비용 유흥\_놀이시설

# 0.02799473 0.27884229

# 숙박시설\_편리한.교통 이미용\_서비스

# -1.92615428 -1.09113882

# 기후\_뚜렷한\_사계절

# -0.21042785

# > modelX2$coefficients

# (Intercept) 휴양휴식

# -38.464478221 0.179250870

# 유흥\_오락 뷰티관광

# -0.097617381 0.063342873

# 쇼핑.1 음식\_미식탐방

# 0.108931551 0.115660654

# 자연풍경감상 역사\_문화유적

# -0.064091918 0.016847522

# 패션\_유행\_.등\_.세련된\_문화 K.POP\_한류스타\_팬미팅

# -0.009930368 0.040898039

# 경제적인\_여행비용 유흥\_놀이시설

# -0.072794050 0.260049767

# 숙박시설\_편리한.교통 이미용\_서비스

# 0.334763370 0.063524408

# 기후\_뚜렷한\_사계절

# 0.036686316

# > modelX3$coefficients

# (Intercept) 휴양휴식

# -14.672365791 0.097016962

# 유흥\_오락 뷰티관광

# 0.036435422 0.151887691

# 쇼핑.1 음식\_미식탐방

# 0.040072251 0.044663141

# 자연풍경감상 역사\_문화유적

# -0.024301653 0.038909315

# 패션\_유행\_.등\_.세련된\_문화 K.POP\_한류스타\_팬미팅

# 0.002208102 -0.048826780

# 경제적인\_여행비용 유흥\_놀이시설

# -0.127192381 0.088460702

# 숙박시설\_편리한.교통 이미용\_서비스

# 0.482939735 -0.075684852

# 기후\_뚜렷한\_사계절

# -0.043230518

# > modelX4$coefficients

# (Intercept) 휴양휴식

# -37.63044373 0.33519511

# 유흥\_오락 뷰티관광

# 0.10940390 0.67865665

# 쇼핑.1 음식\_미식탐방

# -0.02611421 0.19347051

# 자연풍경감상 역사\_문화유적

# -0.02939098 0.19699802

# 패션\_유행\_.등\_.세련된\_문화 K.POP\_한류스타\_팬미팅

# -0.07031369 0.04496613

# 경제적인\_여행비용 유흥\_놀이시설

# -0.04198566 -0.12206319

# 숙박시설\_편리한.교통 이미용\_서비스

# 0.85681595 -0.23551648

# 기후\_뚜렷한\_사계절

# 0.16685777

# 4년간 재방문을 하게 만든 고려요인

# 1회 방문시 고려요인 뷰티 관광

# 2회 방문시 고려요인 숙박시설\_편리한.교통

# 3회 방문시 고려요인 숙박시설\_편리한.교통

# 4회 방문시 고려요인 숙박시설\_편리한.교통

# 방한 방문객은 만족과 상관없이 방문 횟수가 늘면 늘수록 대한민국의 교통및 숙박시설

# 에 강한 만족감을 나타내고 있다.

# p\_Value result

# 휴양휴식 0.7294646 채택

# 유흥\_오락 0.6443329 채택

# 뷰티관광 0.6092449 채택

# 쇼핑.1 0.5931206 채택

# 음식\_미식탐방 0.5622067 채택

# 자연풍경감상 0.5573205 채택

# 역사\_문화유적 0.5084509 채택

# 패션\_유행\_.등\_.세련된\_문화 0.4700696 채택

# K.POP\_한류스타\_팬미팅 0.4313990 채택

# 경제적인\_여행비용 0.4200565 채택

# 유흥\_놀이시설 0.4049757 채택

# 숙박시설\_편리한.교통 0.3984646 채택

# 이미용\_서비스 0.3710314 채택

# 기후\_뚜렷한\_사계절 0.3234574 채택

# 휴양휴식 과 숙박시설 간의 루트를 정부차원에서 지원하여 개발해야한다.

# 연령층 대로 카이검정을 실시하게 된다면 현재 방한 외국인 의 주요요인이 바뀌게 될 수도 있다.

# 정책 개편 방안 주제 Top 5 요인

# 숙박시설 및 편리한 교통

# - 휴양휴식

# - 유흥\_오락

# - 뷰티 관광

# - 쇼핑

# - 음식\_미식탐방

# 이며 다른 결과와 취합해 상세 정책및 개편안 에 대한 고민을 할 수 있 을 것 같다.